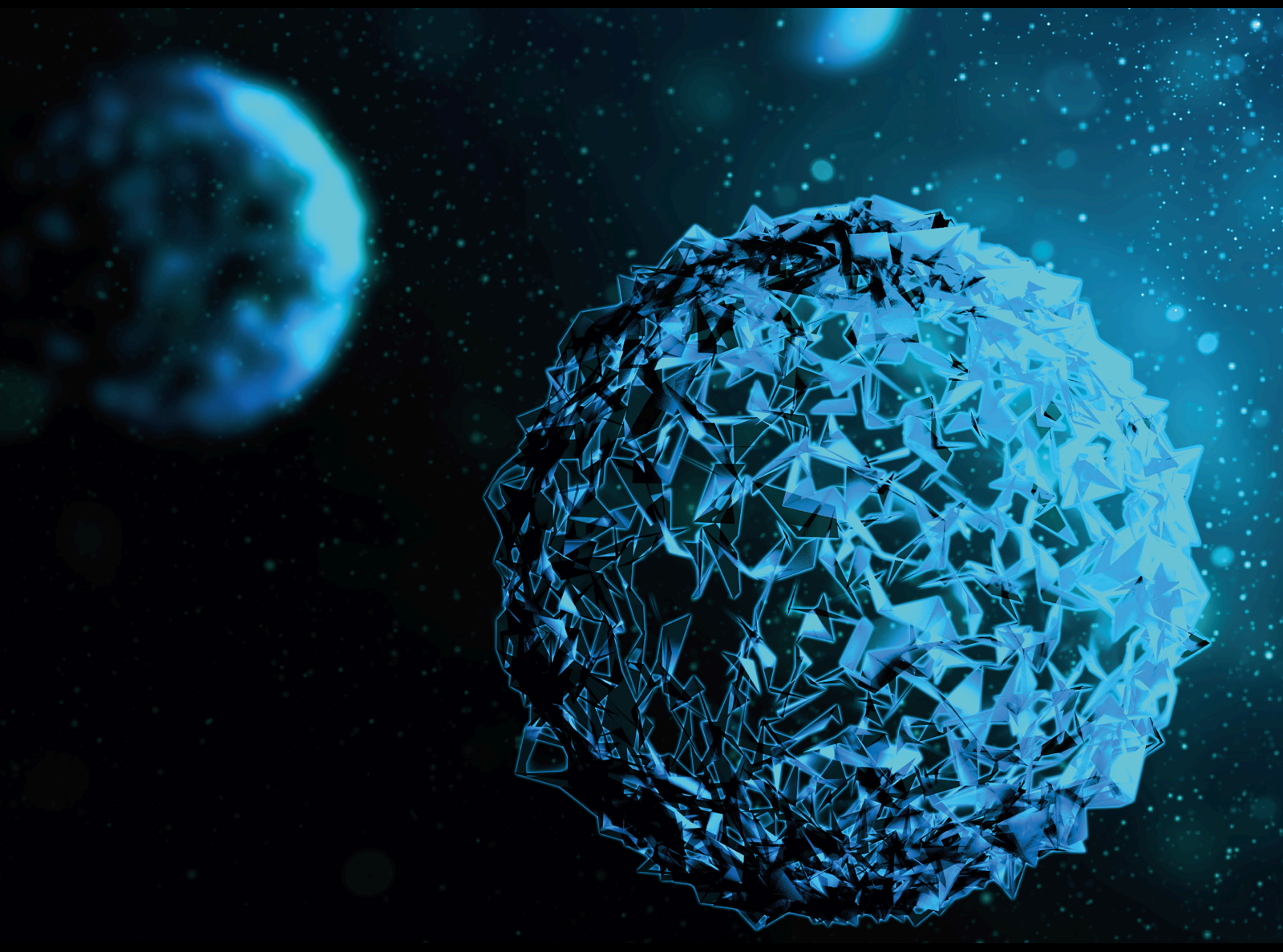
BioMed Research International

Applications of Computational Methods for Prediction, Prevention, Diagnosis, and Treatment of Neurodegenerative Diseases

Lead Guest Editor: Min Tang
Guest Editors: Wenqiang Chen, Jialiang Yang, and Ling Kui



Applications of Computational Methods for Prediction, Prevention, Diagnosis, and Treatment of Neurodegenerative Diseases

BioMed Research International

Applications of Computational Methods for Prediction, Prevention, Diagnosis, and Treatment of Neurodegenerative Diseases

Lead Guest Editor: Min Tang
Guest Editors: Wenqiang Chen, Jialiang Yang, and Ling Kui

Section Editors

Penny A. Asbell, USA
David Bernardo, Spain
Gerald Brandacher, USA
Kim Bridle, Australia
Laura Chronopoulou, Italy
Gerald A. Colvin, USA
Aaron S. Dumont, USA
Pierfrancesco Franco, Italy
Raj P. Kandpal, USA
Fabrizio Montecucco, Italy
Mangesh S. Pednekar, India
Letterio S. Politi, USA
Jinsong Ren, China
William B. Rodgers, USA
Harry W. Schroeder, USA
Andrea Scribante, Italy
Germán Vicente-Rodriguez, Spain
Momiao Xiong, USA
Hui Zhang, China

Academic Editors

Bioinformatics

Contents

Retracted: Hydroxyphenyl Butanone Induces Cell Cycle Arrest through Inhibition of GSK3β in Colorectal Cancer
BioMed Research International
Retraction (1 page), Article ID 9898173, Volume 2024 (2024)

Retracted: Bioinformatics Analyses Reveals a Comprehensive Landscape of CXC Chemokine Family Functions in Non-Small Cell Lung Cancer
BioMed Research International
Retraction (1 page), Article ID 9841854, Volume 2024 (2024)

Retracted: Identifying COVID-19-Specific Transcriptomic Biomarkers with Machine Learning Methods
BioMed Research International
Retraction (1 page), Article ID 9878305, Volume 2023 (2023)

Retracted: Identification of Tumor Tissue of Origin with RNA-Seq Data and Using Gradient Boosting Strategy
BioMed Research International
Retraction (1 page), Article ID 9865973, Volume 2023 (2023)

Retracted: Clustering of Brain Tumor Based on Analysis of MRI Images Using Robust Principal Component Analysis (ROBPCA) Algorithm
BioMed Research International
Retraction (1 page), Article ID 9859342, Volume 2023 (2023)

Retracted: Genome-Wide Screening Identifies Prognostic Long Noncoding RNAs in Hepatocellular Carcinoma
BioMed Research International
Retraction (1 page), Article ID 9841321, Volume 2023 (2023)

Retracted: A Nomogram Predicting the Prognosis of Renal Cell Carcinoma Patients with Lung Metastases
BioMed Research International
Retraction (1 page), Article ID 9832720, Volume 2023 (2023)

Retracted: A Computational Framework to Identify Transcriptional and Network Differences between Hepatocellular Carcinoma and Normal Liver Tissue and Their Applications in Repositioning Drugs
BioMed Research International
Retraction (1 page), Article ID 9832414, Volume 2023 (2023)

Retracted: Serum IL-37 Level Is Associated with Rheumatoid Arthritis and Disease Activity: A Meta-Analysis
BioMed Research International
Retraction (1 page), Article ID 9827814, Volume 2023 (2023)

Retracted: Potential Drug Prediction of Glioblastoma Based on Drug Perturbation-Induced Gene Expression Signatures
BioMed Research International
Retraction (1 page), Article ID 9827010, Volume 2023 (2023)

Retracted: Protocatechualdehyde Inhibits the Osteoclast Differentiation of RAW264.7 and BMM Cells by Regulating NF-*κ*B and MAPK Activity
BioMed Research International
Retraction (1 page), Article ID 9820815, Volume 2023 (2023)

Retracted: NREM Sleep EEG Characteristics Correlate to the Mild Cognitive Impairment in Patients with Parkinsonism
BioMed Research International
Retraction (1 page), Article ID 9806215, Volume 2023 (2023)

Retracted: Identification of Key Genes and Immune Infiltrate in Nonalcoholic Steatohepatitis: A Bioinformatic Analysis
BioMed Research International
Retraction (1 page), Article ID 9803894, Volume 2023 (2023)

Retracted: RUNX3 Expression Level Is Correlated with the Clinical and Pathological Characteristics in Endometrial Cancer: A Systematic Review and Meta-analysis
BioMed Research International
Retraction (1 page), Article ID 9789534, Volume 2023 (2023)

Retracted: A Novel Missense Variant in the Gene *PPP2R5D* Causes a Rare Neurodevelopmental Disorder with Increased Phenotype
BioMed Research International
Retraction (1 page), Article ID 9781948, Volume 2023 (2023)

Retracted: Retinal OCT Texture Analysis for Differentiating Healthy Controls from Multiple Sclerosis (MS) with/without Optic Neuritis
BioMed Research International
Retraction (1 page), Article ID 9763086, Volume 2023 (2023)

Retracted: Microbiome-Metabolomics Reveals Endogenous Alterations of Energy Metabolism by the Dushen Tang to Attenuate D-Galactose-Induced Memory Impairment in Rats
BioMed Research International
Retraction (1 page), Article ID 9759620, Volume 2023 (2023)

Retracted: A Correlation Analysis between SNPs and ROIs of Alzheimer's Disease Based on Deep Learning
BioMed Research International
Retraction (1 page), Article ID 9819525, Volume 2023 (2023)

Contents

Retracted: Comparison of 10-Year Survival Outcomes for Early Single Hepatocellular Carcinoma following Different Treatments
BioMed Research International
Retraction (1 page), Article ID 9827803, Volume 2023 (2023)

Retracted: A Systematic Review of Lymphangioleiomyomatosis on Diagnosis and Molecular Mechanism
BioMed Research International
Retraction (1 page), Article ID 9823193, Volume 2023 (2023)

Retracted: The Potential Function of Super Enhancers in Human Bone Marrow Mesenchymal Stem Cells during Osteogenic Differentiation
BioMed Research International
Retraction (1 page), Article ID 9816124, Volume 2023 (2023)

Retracted: Cannabinoid WIN 55,212-2 Inhibits Human Glioma Cell Growth by Triggering ROS-Mediated Signal Pathways
BioMed Research International
Retraction (1 page), Article ID 9803081, Volume 2023 (2023)

Retracted: Effects of Tai Chi on Patients with Mild Cognitive Impairment: A Systematic Review and Meta-analysis of Randomized Controlled Trials
BioMed Research International
Retraction (1 page), Article ID 9789542, Volume 2023 (2023)

Retracted: A novel TUBG1 mutation with neurodevelopmental disorder caused by malformations of cortical development
BioMed Research International
Retraction (1 page), Article ID 9761369, Volume 2023 (2023)

Retracted: Shared Genetic and Epigenetic Mechanisms between the Osteogenic Differentiation of Dental Pulp Stem Cells and Bone Marrow Stem Cells
BioMed Research International
Retraction (1 page), Article ID 9872581, Volume 2023 (2023)

[Retracted] A Computational Framework to Identify Transcriptional and Network Differences between Hepatocellular Carcinoma and Normal Liver Tissue and Their Applications in Repositioning Drugs
Aimin Hu, Zheng Wei, Zuxiang Zheng, Bichao Luo, Jieming Yi, Xinhong Zhou, and Changjiang Zeng
Research Article (11 pages), Article ID 9921195, Volume 2021 (2021)

[Retracted] Identification of Key Genes and Immune Infiltrate in Nonalcoholic Steatohepatitis: A Bioinformatic Analysis
Zhen-yu Jiang, Yi Zhou, Lu Zhou, Shao-wei Li, and Bang-mao Wang
Research Article (15 pages), Article ID 7561645, Volume 2021 (2021)

[Retracted] Clustering of Brain Tumor Based on Analysis of MRI Images Using Robust Principal Component Analysis (ROBPCA) Algorithm
Ali Hamzenejad, Saeid Jafarzadeh Ghoushchi, and Vahid Baradaran
Research Article (11 pages), Article ID 5516819, Volume 2021 (2021)

[Retracted] NREM Sleep EEG Characteristics Correlate to the Mild Cognitive Impairment in Patients with Parkinsonism
Cheng Zhang, Luhua Wei, Fengqingyang Zeng, Tingwei Zhang, Yunchuang Sun, Yane Shen, Guangfa Wang, Jing Ma, and Jue Zhang
Research Article (10 pages), Article ID 5561974, Volume 2021 (2021)

[Retracted] Protocatechualdehyde Inhibits the Osteoclast Differentiation of RAW264.7 and BMM Cells by Regulating NF-κB and MAPK Activity
Yunyun Qu, Xin Liu, Shuai Zong, Huanxin Sun, Shuang Liu, and Yueran Zhao
Research Article (11 pages), Article ID 6108999, Volume 2021 (2021)

[Retracted] RUNX3 Expression Level Is Correlated with the Clinical and Pathological Characteristics in Endometrial Cancer: A Systematic Review and Meta-analysis
Zhen Liu, Zhi-pan Hong, and Shu-xue Xi
Review Article (7 pages), Article ID 9995384, Volume 2021 (2021)

[Retracted] Retinal OCT Texture Analysis for Differentiating Healthy Controls from Multiple Sclerosis (MS) with/without Optic Neuritis
Hamidreza Dehghan Tazarjani, Zahra Amini, Rahele Kafieh, Fereshteh Ashtari, and Erfan Sadeghi
Research Article (13 pages), Article ID 5579018, Volume 2021 (2021)

[Retracted] Identifying COVID-19-Specific Transcriptomic Biomarkers with Machine Learning Methods
Lei Chen, Zhandong Li, Tao Zeng, Yu-Hang Zhang, KaiYan Feng, Tao Huang, and Yu-Dong Cai
Research Article (11 pages), Article ID 9939134, Volume 2021 (2021)

[Retracted] Hydroxyphenyl Butanone Induces Cell Cycle Arrest through Inhibition of GSK3β in Colorectal Cancer
Songyan Zhang, Yunfeng Wang, Haopeng Zhang, Chengming Sun, Shuwei Dang, and Ming Liu
Research Article (13 pages), Article ID 9981815, Volume 2021 (2021)

[Retracted] Microbiome-Metabolomics Reveals Endogenous Alterations of Energy Metabolism by the Dushen Tang to Attenuate D-Galactose-Induced Memory Impairment in Rats
Jifeng Wang, Min He, Wenjun Guo, Yanhong Zhang, Xin Sui, Jianan Lin, Xiaoran Liu, Hui Li, Jing Li, Qing Yang, Mo Kan, Zhuang Zhang, Sitong Ming, Xiaobo Qu, and Na Li
Research Article (17 pages), Article ID 6649085, Volume 2021 (2021)

Contents

[Retracted] Genome-Wide Screening Identifies Prognostic Long Noncoding RNAs in Hepatocellular Carcinoma
Yujie Feng, Xiao Hu, Kai Ma, Bingyuan Zhang, and Chuandong Sun
Research Article (16 pages), Article ID 6640652, Volume 2021 (2021)

[Retracted] Cannabinoid WIN 55,212-2 Inhibits Human Glioma Cell Growth by Triggering ROS-Mediated Signal Pathways
Kun Wang, Qian Wang, Qinghao Li, Zhaoqiang Zhang, Jing Gao, Cundong Fan, Baoliang Sun, and Qingbin Ni
Research Article (11 pages), Article ID 6612592, Volume 2021 (2021)

[Retracted] Effects of Tai Chi on Patients with Mild Cognitive Impairment: A Systematic Review and Meta-analysis of Randomized Controlled Trials
Run Lin, Shaoyang Cui, Juan Yang, Huijun Yang, Zitong Feng, Dietlind L. Wahner-Roedler, Xuan Zhou, Manisha Salinas, Molly J. Mallory, Alexander Do, Sara E. Bublitz, Tony Y. Chon, Chunzhi Tang, Brent A. Bauer, and Mingzhu Xu
Review Article (10 pages), Article ID 5530149, Volume 2021 (2021)

[Retracted] Comparison of 10-Year Survival Outcomes for Early Single Hepatocellular Carcinoma following Different Treatments
Fanyu Meng, Haoyun Zhang, Haiwen Peng, and Shichun Lu
Research Article (8 pages), Article ID 6638117, Volume 2021 (2021)

[Retracted] A Nomogram Predicting the Prognosis of Renal Cell Carcinoma Patients with Lung Metastases
Xinyu Sheng, Xuan Lu, Jian Wu, Lu Chen, and Hongcui Cao
Research Article (11 pages), Article ID 6627562, Volume 2021 (2021)

[Retracted] A novel *TUBG1* mutation with neurodevelopmental disorder caused by malformations of cortical development
Ru Shen, Zhen Zhang, Yu Zhuang, Xiaohong Yang, and Lifen Duan
Research Article (8 pages), Article ID 6644274, Volume 2021 (2021)

[Retracted] Identification of Tumor Tissue of Origin with RNA-Seq Data and Using Gradient Boosting Strategy
Ruixi Li, Bo Liao, Bo Wang, Chan Dai, Xin Liang, Geng Tian, and Fangxiang Wu
Research Article (14 pages), Article ID 6653793, Volume 2021 (2021)

[Retracted] A Novel Missense Variant in the Gene *PPP2R5D* Causes a Rare Neurodevelopmental Disorder with Increased Phenotype
Lulu Yan, Ru Shen, Zongfu Cao, Chunxiao Han, Yuxin Zhang, Yingwen Liu, Xiangchun Yang, Min Xie, and Haibo Li
Research Article (7 pages), Article ID 6661860, Volume 2021 (2021)

[Retracted] A Systematic Review of Lymphangioleiomyomatosis on Diagnosis and Molecular Mechanism
Xiaotong Dong, Lvcheng Jin, Ailan Wang, Liping Wu, Xintong Fan, Qian Hou, Tianbao Li, Ruilian Zhao, and Yunxiang Zhang
Review Article (8 pages), Article ID 6612776, Volume 2021 (2021)

[Retracted] A Correlation Analysis between SNPs and ROIs of Alzheimer's Disease Based on Deep Learning
Juan Zhou, Linfeng Hu, Yu Jiang, and Liyue Liu
Research Article (13 pages), Article ID 8890513, Volume 2021 (2021)

[Retracted] Shared Genetic and Epigenetic Mechanisms between the Osteogenic Differentiation of Dental Pulp Stem Cells and Bone Marrow Stem Cells
Sebastian Gaus, Hanluo Li, Simin Li, Qian Wang, Tina Kottek, Sebastian Hahnel, Xiangqiong Liu, Yupei Deng, Dirk Ziebolz, Rainer Haak, Gerhard Schmalz, Lei Liu, Vuk Savkovic, and Bernd Lethaus
Research Article (25 pages), Article ID 6697810, Volume 2021 (2021)

[Retracted] Serum IL-37 Level Is Associated with Rheumatoid Arthritis and Disease Activity: A Meta-Analysis
Shengnan Cao, Haojun Shi, Guodong Sun, Yuanzhen Chen, Guangjian Hou, Dandan Wang, and Bin Shi
Review Article (7 pages), Article ID 6653439, Volume 2021 (2021)

[Retracted] Potential Drug Prediction of Glioblastoma Based on Drug Perturbation-Induced Gene Expression Signatures
Bochi Zhu, Xijing Mao, and Yuhong Man
Research Article (11 pages), Article ID 6659701, Volume 2021 (2021)

[Retracted] Bioinformatics Analyses Reveals a Comprehensive Landscape of CXC Chemokine Family Functions in Non-Small Cell Lung Cancer
He Tian, Liyu Wang, Yu Liu, Yalong Wang, Yujia Zheng, Tao Fan, Bo Zheng, Fengwei Tan, Qi Xue, Shugeng Gao, Chunxiang Li, and Jie He
Research Article (34 pages), Article ID 6686158, Volume 2021 (2021)

[Retracted] The Potential Function of Super Enhancers in Human Bone Marrow Mesenchymal Stem Cells during Osteogenic Differentiation
Zhijie Huang, Bo Jia, Qian Wang, Naishun Wang, and Jianjiang Zhao
Research Article (11 pages), Article ID 6614762, Volume 2021 (2021)

Hindawi
BioMed Research International
Volume 2024, Article ID 9898173, 1 page
https://doi.org/10.1155/2024/9898173

Retraction

Retracted: Hydroxyphenyl Butanone Induces Cell Cycle Arrest through Inhibition of GSK3β in Colorectal Cancer

BioMed Research International

Received 8 January 2024; Accepted 8 January 2024; Published 9 January 2024

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Zhang, Y. Wang, H. Zhang, C. Sun, S. Dang, and M. Liu, "Hydroxyphenyl Butanone Induces Cell Cycle Arrest through Inhibition of GSK3β in Colorectal Cancer," *BioMed Research International*, vol. 2021, Article ID 9981815, 13 pages, 2021.

Hindawi
BioMed Research International
Volume 2024, Article ID 9841854, 1 page
https://doi.org/10.1155/2024/9841854

Retraction

Retracted: Bioinformatics Analyses Reveals a Comprehensive Landscape of CXC Chemokine Family Functions in Non-Small Cell Lung Cancer

BioMed Research International

Received 8 January 2024; Accepted 8 January 2024; Published 9 January 2024

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Tian, L. Wang, Y. Liu et al., "Bioinformatics Analyses Reveals a Comprehensive Landscape of CXC Chemokine Family Functions in Non-Small Cell Lung Cancer," *BioMed Research International*, vol. 2021, Article ID 6686158, 34 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9878305, 1 page
https://doi.org/10.1155/2023/9878305

Retraction

Retracted: Identifying COVID-19-Specific Transcriptomic Biomarkers with Machine Learning Methods

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Chen, Z. Li, T. Zeng et al., "Identifying COVID-19-Specific Transcriptomic Biomarkers with Machine Learning Methods," *BioMed Research International*, vol. 2021, Article ID 9939134, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9865973, 1 page
https://doi.org/10.1155/2023/9865973

Retraction

Retracted: Identification of Tumor Tissue of Origin with RNA-Seq Data and Using Gradient Boosting Strategy

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] R. Li, B. Liao, B. Wang et al., "Identification of Tumor Tissue of Origin with RNA-Seq Data and Using Gradient Boosting Strategy," *BioMed Research International*, vol. 2021, Article ID 6653793, 14 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9859342, 1 page
https://doi.org/10.1155/2023/9859342

Retraction

Retracted: Clustering of Brain Tumor Based on Analysis of MRI Images Using Robust Principal Component Analysis (ROBPCA) Algorithm

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] A. Hamzenejad, S. J. Ghoushchi, and V. Baradaran, "Clustering of Brain Tumor Based on Analysis of MRI Images Using Robust Principal Component Analysis (ROBPCA) Algorithm," *BioMed Research International*, vol. 2021, Article ID 5516819, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9841321, 1 page
https://doi.org/10.1155/2023/9841321

Retraction

Retracted: Genome-Wide Screening Identifies Prognostic Long Noncoding RNAs in Hepatocellular Carcinoma

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Feng, X. Hu, K. Ma, B. Zhang, and C. Sun, "Genome-Wide Screening Identifies Prognostic Long Noncoding RNAs in Hepatocellular Carcinoma," *BioMed Research International*, vol. 2021, Article ID 6640652, 16 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9832720, 1 page
https://doi.org/10.1155/2023/9832720

Retraction

Retracted: A Nomogram Predicting the Prognosis of Renal Cell Carcinoma Patients with Lung Metastases

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] X. Sheng, X. Lu, J. Wu, L. Chen, and H. Cao, "A Nomogram Predicting the Prognosis of Renal Cell Carcinoma Patients with Lung Metastases," *BioMed Research International*, vol. 2021, Article ID 6627562, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9832414, 1 page
https://doi.org/10.1155/2023/9832414

Retraction

Retracted: A Computational Framework to Identify Transcriptional and Network Differences between Hepatocellular Carcinoma and Normal Liver Tissue and Their Applications in Repositioning Drugs

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] A. Hu, Z. Wei, Z. Zheng et al., "A Computational Framework to Identify Transcriptional and Network Differences between Hepatocellular Carcinoma and Normal Liver Tissue and Their Applications in Repositioning Drugs," *BioMed Research International*, vol. 2021, Article ID 9921195, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9827814, 1 page
https://doi.org/10.1155/2023/9827814

Retraction

Retracted: Serum IL-37 Level Is Associated with Rheumatoid Arthritis and Disease Activity: A Meta-Analysis

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Cao, H. Shi, G. Sun et al., "Serum IL-37 Level Is Associated with Rheumatoid Arthritis and Disease Activity: A Meta-Analysis," *BioMed Research International*, vol. 2021, Article ID 6653439, 7 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9827010, 1 page
https://doi.org/10.1155/2023/9827010

Retraction

Retracted: Potential Drug Prediction of Glioblastoma Based on Drug Perturbation-Induced Gene Expression Signatures

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] B. Zhu, X. Mao, and Y. Man, "Potential Drug Prediction of Glioblastoma Based on Drug Perturbation-Induced Gene Expression Signatures," *BioMed Research International*, vol. 2021, Article ID 6659701, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9820815, 1 page
https://doi.org/10.1155/2023/9820815

Retraction

Retracted: Protocatechualdehyde Inhibits the Osteoclast Differentiation of RAW264.7 and BMM Cells by Regulating NF-κB and MAPK Activity

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Qu, X. Liu, S. Zong, H. Sun, S. Liu, and Y. Zhao, "Protocatechualdehyde Inhibits the Osteoclast Differentiation of RAW264.7 and BMM Cells by Regulating NF-κB and MAPK Activity," *BioMed Research International*, vol. 2021, Article ID 6108999, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9806215, 1 page
https://doi.org/10.1155/2023/9806215

Retraction

Retracted: NREM Sleep EEG Characteristics Correlate to the Mild Cognitive Impairment in Patients with Parkinsonism

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] C. Zhang, L. Wei, F. Zeng et al., "NREM Sleep EEG Characteristics Correlate to the Mild Cognitive Impairment in Patients with Parkinsonism," *BioMed Research International*, vol. 2021, Article ID 5561974, 10 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9803894, 1 page
https://doi.org/10.1155/2023/9803894

Retraction

Retracted: Identification of Key Genes and Immune Infiltrate in Nonalcoholic Steatohepatitis: A Bioinformatic Analysis

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z.-y. Jiang, Y. Zhou, Z. Lu, S.-w. Li, and B.-m. Wang, "Identification of Key Genes and Immune Infiltrate in Nonalcoholic Steatohepatitis: A Bioinformatic Analysis," *BioMed Research International*, vol. 2021, Article ID 7561645, 15 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9789534, 1 page
https://doi.org/10.1155/2023/9789534

Retraction

Retracted: RUNX3 Expression Level Is Correlated with the Clinical and Pathological Characteristics in Endometrial Cancer: A Systematic Review and Meta-analysis

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Liu, Z. Hong, and S. Xi, "RUNX3 Expression Level Is Correlated with the Clinical and Pathological Characteristics in Endometrial Cancer: A Systematic Review and Meta-analysis," *BioMed Research International*, vol. 2021, Article ID 9995384, 7 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9781948, 1 page
https://doi.org/10.1155/2023/9781948

Retraction

Retracted: A Novel Missense Variant in the Gene *PPP2R5D* Causes a Rare Neurodevelopmental Disorder with Increased Phenotype

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Yan, R. Shen, Z. Cao et al., "A Novel Missense Variant in the Gene *PPP2R5D* Causes a Rare Neurodevelopmental Disorder with Increased Phenotype," *BioMed Research International*, vol. 2021, Article ID 6661860, 7 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9763086, 1 page
https://doi.org/10.1155/2023/9763086

Retraction

Retracted: Retinal OCT Texture Analysis for Differentiating Healthy Controls from Multiple Sclerosis (MS) with/without Optic Neuritis

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. D. Tazarjani, Z. Amini, R. Kafieh, F. Ashtari, and E. Sadeghi, "Retinal OCT Texture Analysis for Differentiating Healthy Controls from Multiple Sclerosis (MS) with/without Optic Neuritis," *BioMed Research International*, vol. 2021, Article ID 5579018, 13 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9759620, 1 page
https://doi.org/10.1155/2023/9759620

Retraction

Retracted: Microbiome-Metabolomics Reveals Endogenous Alterations of Energy Metabolism by the Dushen Tang to Attenuate D-Galactose-Induced Memory Impairment in Rats

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Wang, M. He, W. Guo et al., "Microbiome-Metabolomics Reveals Endogenous Alterations of Energy Metabolism by the Dushen Tang to Attenuate D-Galactose-Induced Memory Impairment in Rats," *BioMed Research International*, vol. 2021, Article ID 6649085, 17 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9819525, 1 page
https://doi.org/10.1155/2023/9819525

Retraction

Retracted: A Correlation Analysis between SNPs and ROIs of Alzheimer's Disease Based on Deep Learning

BioMed Research International

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Zhou, L. Hu, Y. Jiang, and L. Liu, "A Correlation Analysis between SNPs and ROIs of Alzheimer's Disease Based on Deep Learning," *BioMed Research International*, vol. 2021, Article ID 8890513, 13 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9827803, 1 page
https://doi.org/10.1155/2023/9827803

Retraction

Retracted: Comparison of 10-Year Survival Outcomes for Early Single Hepatocellular Carcinoma following Different Treatments

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] F. Meng, H. Zhang, H. Peng, and S. Lu, "Comparison of 10-Year Survival Outcomes for Early Single Hepatocellular Carcinoma following Different Treatments," *BioMed Research International*, vol. 2021, Article ID 6638117, 8 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9823193, 1 page
https://doi.org/10.1155/2023/9823193

Retraction

Retracted: A Systematic Review of Lymphangioleiomyomatosis on Diagnosis and Molecular Mechanism

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] X. Dong, L. Jin, A. Wang et al., "A Systematic Review of Lymphangioleiomyomatosis on Diagnosis and Molecular Mechanism," *BioMed Research International*, vol. 2021, Article ID 6612776, 8 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9816124, 1 page
https://doi.org/10.1155/2023/9816124

Retraction

Retracted: The Potential Function of Super Enhancers in Human Bone Marrow Mesenchymal Stem Cells during Osteogenic Differentiation

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Huang, B. Jia, Q. Wang, N. Wang, and J. Zhao, "The Potential Function of Super Enhancers in Human Bone Marrow Mesenchymal Stem Cells during Osteogenic Differentiation," *BioMed Research International*, vol. 2021, Article ID 6614762, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9803081, 1 page
https://doi.org/10.1155/2023/9803081

Retraction

Retracted: Cannabinoid WIN 55,212-2 Inhibits Human Glioma Cell Growth by Triggering ROS-Mediated Signal Pathways

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] K. Wang, Q. Wang, Q. Li et al., "Cannabinoid WIN 55,212-2 Inhibits Human Glioma Cell Growth by Triggering ROS-Mediated Signal Pathways," *BioMed Research International*, vol. 2021, Article ID 6612592, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9789542, 1 page
https://doi.org/10.1155/2023/9789542

Retraction

Retracted: Effects of Tai Chi on Patients with Mild Cognitive Impairment: A Systematic Review and Meta-analysis of Randomized Controlled Trials

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] R. Lin, S. Cui, J. Yang et al., "Effects of Tai Chi on Patients with Mild Cognitive Impairment: A Systematic Review and Meta-analysis of Randomized Controlled Trials," *BioMed Research International*, vol. 2021, Article ID 5530149, 10 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9761369, 1 page
https://doi.org/10.1155/2023/9761369

Retraction

Retracted: A novel TUBG1 mutation with neurodevelopmental disorder caused by malformations of cortical development

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] R. Shen, Z. Zhang, Y. Zhuang, X. Yang, and L. Duan, "A novel *TUBG1* mutation with neurodevelopmental disorder caused by malformations of cortical development," *BioMed Research International*, vol. 2021, Article ID 6644274, 8 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9872581, 1 page
https://doi.org/10.1155/2023/9872581

Retraction

Retracted: Shared Genetic and Epigenetic Mechanisms between the Osteogenic Differentiation of Dental Pulp Stem Cells and Bone Marrow Stem Cells

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Gaus, H. Li, S. Li et al., "Shared Genetic and Epigenetic Mechanisms between the Osteogenic Differentiation of Dental Pulp Stem Cells and Bone Marrow Stem Cells," *BioMed Research International*, vol. 2021, Article ID 6697810, 25 pages, 2021.

Hindawi
BioMed Research International
Volume 2023, Article ID 9832414, 1 page
https://doi.org/10.1155/2023/9832414

Retraction

Retracted: A Computational Framework to Identify Transcriptional and Network Differences between Hepatocellular Carcinoma and Normal Liver Tissue and Their Applications in Repositioning Drugs

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] A. Hu, Z. Wei, Z. Zheng et al., "A Computational Framework to Identify Transcriptional and Network Differences between Hepatocellular Carcinoma and Normal Liver Tissue and Their Applications in Repositioning Drugs," *BioMed Research International*, vol. 2021, Article ID 9921195, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 9921195, 11 pages
https://doi.org/10.1155/2021/9921195

Research Article

A Computational Framework to Identify Transcriptional and Network Differences between Hepatocellular Carcinoma and Normal Liver Tissue and Their Applications in Repositioning Drugs

Aimin Hu, Zheng Wei, Zuxiang Zheng, Bichao Luo, Jieming Yi, Xinhong Zhou, and Changjiang Zeng

Xiantao First People's Hospital Affiliated to Yangtze University, Xiantao 433000, China

Correspondence should be addressed to Zheng Wei; weizheng95852@163.com

Received 1 April 2021; Revised 29 June 2021; Accepted 30 June 2021; Published 23 September 2021

Academic Editor: Min Tang

Hepatocellular carcinoma (HCC) is one of the most common and lethal malignancies worldwide. Although there have been extensive studies on the molecular mechanisms of its carcinogenesis, FDA-approved drugs for HCC are rare. Side effects, development time, and cost of these drugs are the major bottlenecks, which can be partially overcome by drug repositioning. In this study, we developed a computational framework to study the mechanisms of HCC carcinogenesis, in which drug perturbation-induced gene expression signatures were utilized for repositioning of potential drugs. Specifically, we first performed differential expression analysis and coexpression network module analysis on the HCC dataset from The Cancer Genome Atlas database. Differential gene expression analysis identified 1,337 differentially expressed genes between HCC and adjacent normal tissues, which were significantly enriched in functions related to various pathways, including α-adrenergic receptor activity pathway and epinephrine binding pathway. Weighted gene correlation network analysis (WGCNA) suggested that the number of coexpression modules was higher in HCC tissues than in normal tissues. Finally, by correlating differentially expressed genes with drug perturbation-related signatures, we prioritized a few potential drugs, including nutlin and eribulin, for the treatment of hepatocellular carcinoma. The drugs have been reported by a few experimental studies to be effective in killing cancer cells.

1. Introduction

Liver cancer was the fifth most common cancer in 2012, accounting for 9.1% of global cancer deaths [1]. Most liver cancers (83%) are diagnosed in less developed countries, mainly in Asia, Africa, and Southern Europe. The vast majority (75%-90%) of primary liver cancers are hepatocellular carcinomas (HCCs), the most common and deadly malignant tumor worldwide [2, 3]. HCC usually occurs in cases with liver cirrhosis caused by viral infection and chronic inflammatory liver disease caused by exposure to chemical carcinogens. Known risk factors for HCC include chronic hepatitis B virus (HBV) and hepatitis C virus (HCV) infection, dietary aflatoxin exposure, fatty liver disease, alcoholic cirrhosis, obesity, smoking, diabetes, and iron overload [4, 5]. Patients are often diagnosed at advanced stages of liver cancer, and chemotherapy and immunotherapy are the most feasible treatment options.

Unfortunately, despite extensive research on the molecular mechanism of liver carcinogenesis, there are still only a few effective treatment options. Very little is known about the pathogenesis of most types of cancers, including liver cancer [6–8]. In the past few decades, about 130-180 anticancer drugs have been approved by the US FDA for use in clinical treatment [9]. Despite the increasing research on cancer drugs using model species (supported by nonhuman data) and nearly 1,000 drugs having been formulated using combinations of FDA-approved anticancer drugs, there is

still uncertainty about the efficacy of these drugs in the treatment of cancer due to insufficient understanding of the molecular mechanisms of the disease [10–13]. This uncertainty poses the challenge of expensive clinical trials for the pharmaceutical industry. Thus, research and development of new drugs for the treatment of cancer and rational use of such drugs are still a big scientific issue. An alternative way is to perform drug repositioning, that is, identifying novel usage of existing drugs. Drug repositioning has been widely used in cancers [14–16] and other diseases including the recent outbreaking COVID-19 [17, 18].

In this study, we carried out in-depth mining of second-generation sequencing transcriptomics data on liver cancer in The Cancer Genome Atlas (TCGA) database, modularized genes through differential expression analysis, compared module changes before and after liver cancer, and studied the corresponding genes in the module-enriched functional pathways to understand changes in gene regulation in liver cancer. In addition, we also compared the changes in gene regulation before and after liver cancer with the small-molecule compounds or drugs reported to interfere with the gene regulation in model species and selected the small-molecule compounds or drugs which caused reversed changes as the potential drugs for the treatment of cancer. Through these studies, we hope to develop systematic research and treatment programs for liver cancer in the future.

2. Materials and Methods

2.1. Gene Expression Data. The sample data (liver cancer) were obtained from TCGA database [19]. We collected RNA sequencing data (read count data) of all samples under the project ID of TCGA-LIHC. After excluding samples with low data quality (if the vial in the sample ID is B, it means formalin-fixed paraffin-embedded tissue, which has been proved to be ineffective for sequencing analysis, and this will be removed), we obtained a total of 421 liver cancer samples (50 normal samples and 371 cancer samples), in order to ensure the reliability of subsequent analysis. The significance of the difference between the cancer sample and the normal sample was compared, and the result was corrected by the Bonferoni method. The result showed that the difference between the groups was not significant (Supplementary Figure S1).

2.2. Dataset of Differentially Expressed Genes Affected by Drugs or Small Molecules. The dataset of differentially expressed genes with which drugs or small molecules could interfere was collected from Crowd Extracted Expression of Differential Signatures (CREEDS) (http://amp.pharm.mssm.edu/creeds) [20], which contained 8590 gene expression characteristics as affected by various small-molecule compounds or drugs.

2.3. Differential Expression Gene Extraction and Functional Enrichment Analysis. We applied a pipeline similar to Gao et al. [21]. Specifically, we used principal component analysis (PCA) to screen all samples of liver cancer and excluded the outliers to reduce sample disturbance. A total of 13 outliers were excluded. Subsequently, read counts were used to call differential expression genes between liver cancer and normal samples by using an R package DESeq2 [22] (adjusted p value $<1e\text{-}5$ was set as the threshold). We used ClueGO in Cytoscape (https://cytoscape.org/) to perform pathway enrichment analysis on the above differential gene expression genes [23]. The datasets used were from the GO, KEGG, and Reactome databases [24–26], and 0.05 was selected as the significance threshold.

2.4. Construction of Cancer and Normal Gene Coexpression Networks. We classified liver cancer samples and normal samples by hierarchical clustering provided by weighted gene correlation network analysis (WGCNA) and removed abnormal samples to construct a coexpression network [27]. The soft threshold was set as follows: normal group, 5 and liver cancer group, 5.

2.5. Analysis of Gene Regulatory Networks. Clusters of gene modules were obtained by WGCNA. The correlation of gene expression in each module was relatively high for genes belonging to the same regulator subnetwork and participating in the same functional regulation. We can understand the influence of the occurrence of liver cancer on the synergy between genes by comparing the number of clustering modules before and after liver cancer. We performed functional enrichment analysis on all modules gathered in the normal sample group (except the gray module; genes in the gray module were not related, or the correlation was not significant) and found functional pathways related to cancer regulation; these functional pathways were used as benchmarks. We compared the number of genes and the related changes in the modules enriched in these functional pathways in the corresponding liver cancer sample group to explore the internal regulatory relationship of related pathways before and after cancer and to understand the pathogenesis of cancer. In order to reveal the main functional pathways of each module, we use the ClueGO cyREST tool for functional enrichment analysis [28]. For gene modules with specific functions, we use DGCA (for differential gene correlation analysis, a comprehensive R package for differential gene correlation analysis) [29] to calculate the correlations between genes in the modules in the liver cancer group and the normal group. The correlation between genes in the final module is visualized by Cytoscape.

2.6. Analysis of Potential Applicability of Drugs. 8590 drug perturbation-induced gene expression signatures collected in CREEDS were used in our analysis. Signatures from CREEDS were tanked using Fisher's exact test. We calculated the significance of overlap between the upregulation and downregulation of genes caused by drugs and the upregulated and downregulated genes in normal and cancer samples, respectively. The drugs were ranked on the basis of overlap observed between the genes induced by the drug and the differentially expressed genes in liver cancer.

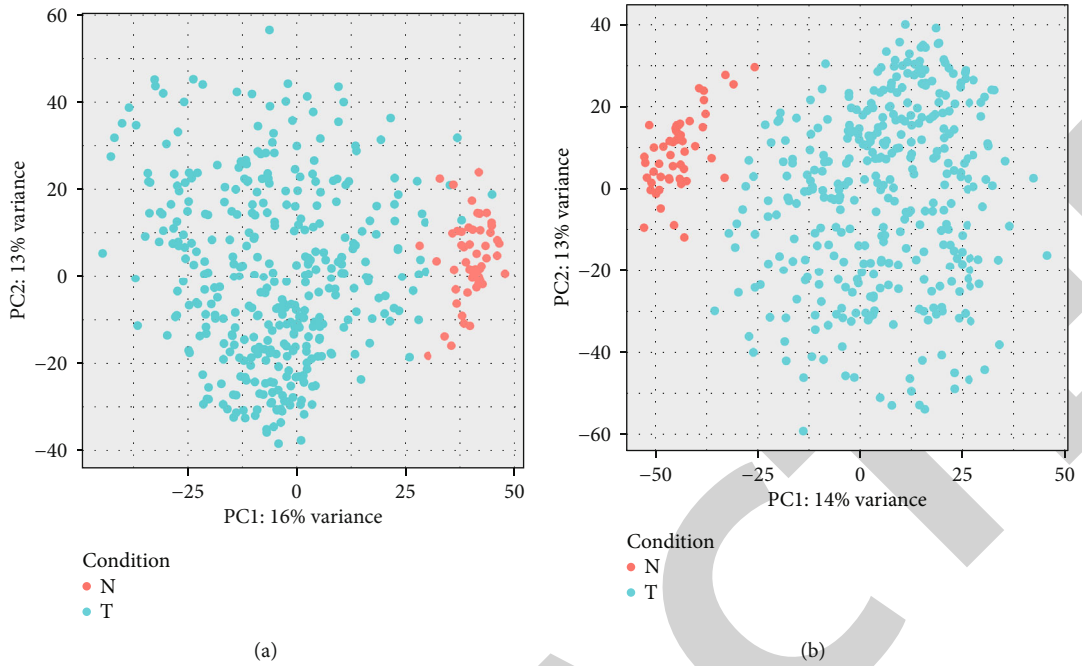


Figure 1: Principal component analysis (PCA) of liver cancer samples. We can obtain the distribution of the samples on the principal component axis through the matrix decomposition method, which can be subsequently used to preprocess the sample data and remove outliers. (a) Original sample distribution; (b) filtered sample distribution; the red dot N indicates a normal sample point, and the blue dot T indicates a liver cancer sample point.

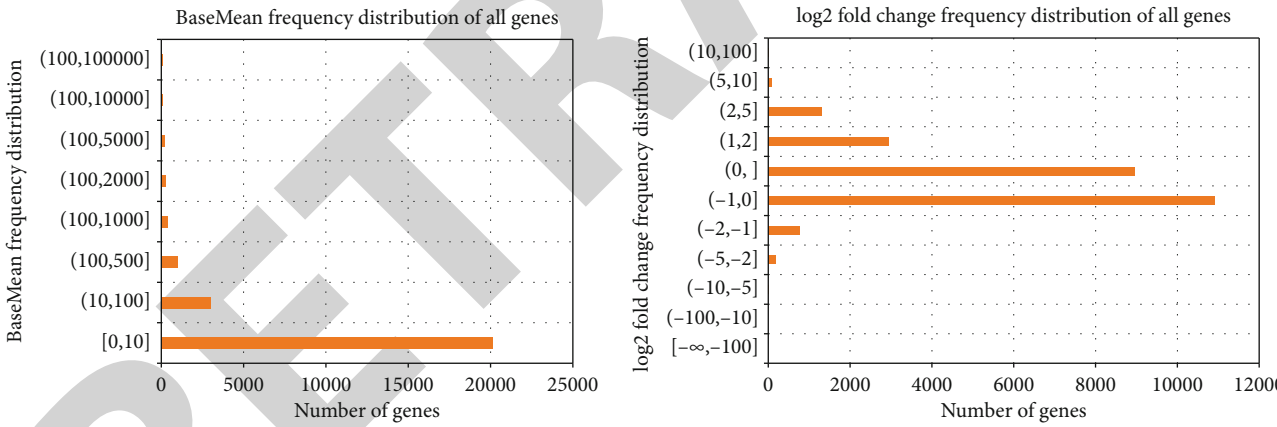


Figure 2: Data distribution plot for DESeq2 results. (a) Base mean frequency distribution of all genes and distribution of the mean for the gene's read counts in all samples; (b) log2 fold change frequency distribution of all genes and distribution of the mean value for the gene's fold change in all samples.

3. Result

3.1. Differential Gene Analysis Identified Important HCC Genes Enriched in Many GO Terms. Results of PCA for all samples have been presented in Figure 1. DESeq2 was used after data processing. In liver cancer samples, we obtained 12,867 significantly differentially expressed genes compared to normal samples, with 4,533 upregulated and 8,334 downregulated genes.

Using the selected threshold screening (adjusted p value $<1e\text{-}5$), many significantly differentially expressed genes were detected. We decided to increase the screening criteria (adjusted p value $\leq 1e\text{-}5$, base mean ≥ 10, and $|\log 2$ fold change $|\geq 1$) according to the data distribution (Figure 2). We obtained 1,337 significantly differentially expressed genes, with 1,041 upregulated and 296 downregulated genes, compared with normal samples. Enrichment analysis results have been presented in Supplementary Table S1 and S2, including "α-adrenergic receptor activity" (GO:0004936), "adrenergic receptor activity" (GO:0004935), "epinephrine binding" (GO:0051379), "copper ions" (GO:0010273), "detoxification of inorganic compounds" (GO:0061687), "stress response to metal ions" (GO:0097501), "stress response to copper ion" (GO:0004935) in GO dataset,

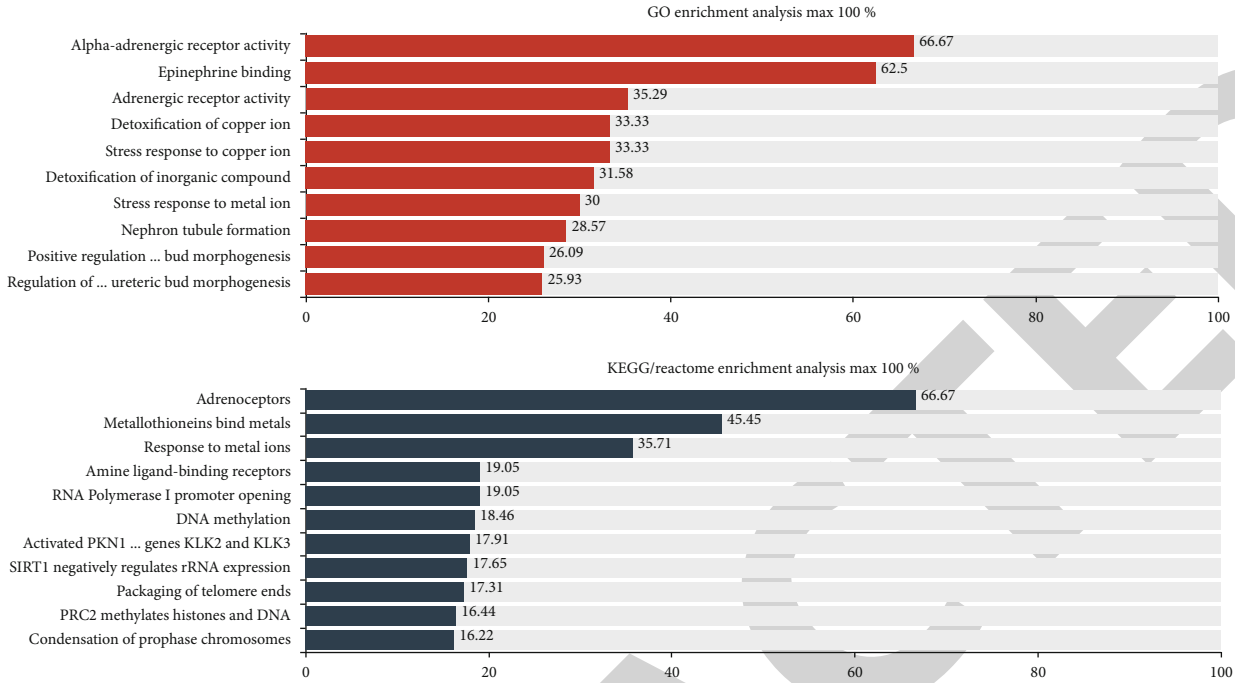


Figure 3: The top ten enrichment analysis results with the highest proportion of genes in GO, KEGG, and Reactome. The x-axis depicts the percentage of enriched differentially expressed genes in the corresponding pathway, and the y-axis presents the name of the pathway. GO enrichment analysis is on the top, whereas KEGG and Reactome enrichment analysis are merged below.

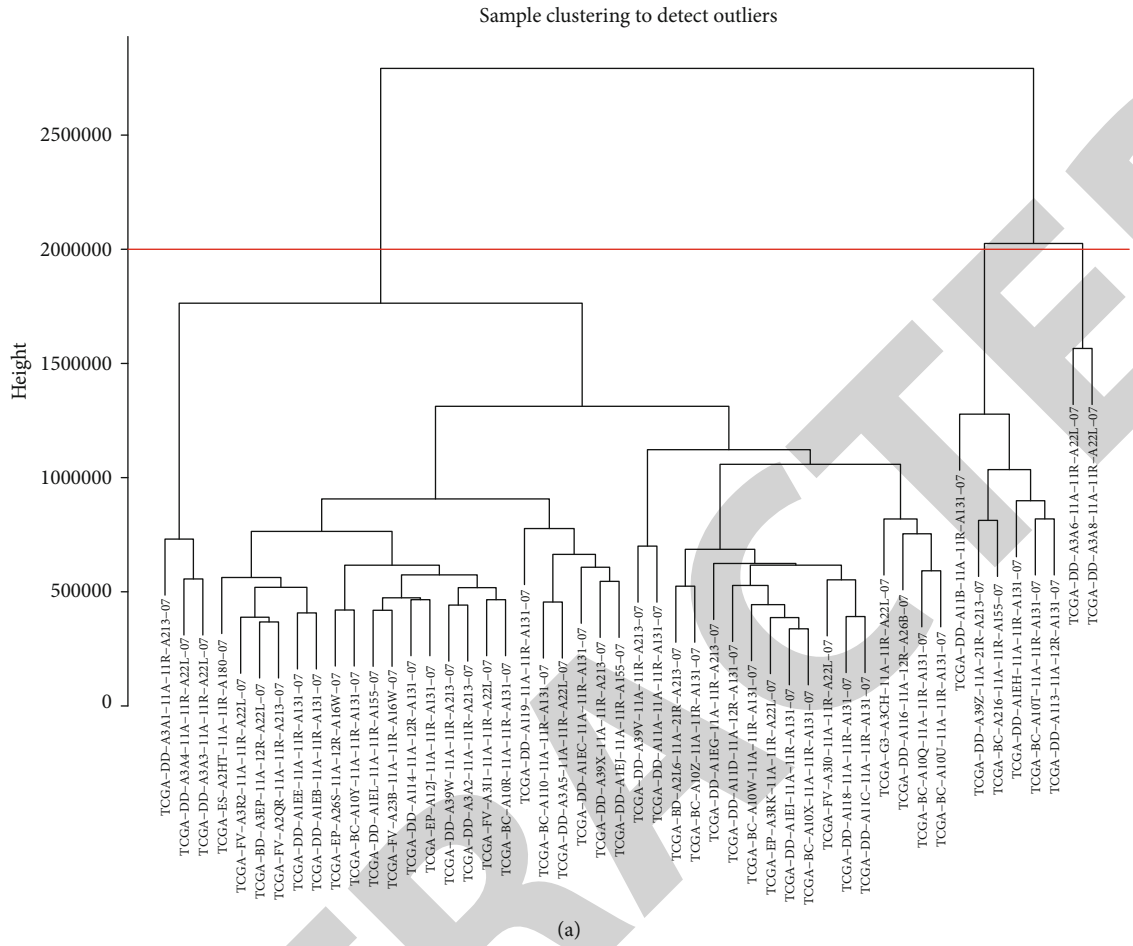
"metallothioneins bind metals" (R-HSA:5661231) [30], "RNA Polymerase I Promoter Opening" (R-HSA:73728) [31], and "SIRT1 negatively regulates rRNA expression" (R-HSA:427359) [32] in Reactome dataset, and "Neuroactive ligand-receptor interaction" (KEGG:04080), "Systemic lupus erythematosus" (KEGG:05322), and "Alcoholism" (KEGG:05034) in KEGG dataset. We have shown the top 10 pathways with the highest proportion of genes in the enrichment results of GO, KEGG, and Reactome (Figure 3).

3.2. HCC-Perturbed Coexpression Modules Identified by Gene Coexpression Submodule Analysis. We used the original data of differential gene analysis as the input data of WGCNA to conduct gene coexpression network analysis. The input data were divided into a normal group and a liver cancer group and analyzed separately to observe the synergistic effect between gene expression under different conditions. We removed the genes whose expression level was 0 in all samples, deleted the sample points of the partial segregation group according to the hierarchical clustering results of the samples (Figures 4(a) and 4(b)), and then performed coexpression network analysis. We finally obtained 84 gene modules in the liver cancer group and 45 gene modules in the normal group (Figures 4(c) and 4(d)). We found that in the liver cancer group, the number of submodules was significantly higher, the synergy of gene expression was lower, and the cell regulatory system tended to be disordered as compared with the normal group. The main functional pathways of each module revealed by functional enrichment analysis using ClueGO cyREST tool have been presented in Supplementary Table S3 and S4. There were 1,857 instances of annotation data for the submodules of the normal group and only 357 instances of annotation data for the submodules of the liver cancer group. Although the number of cancer submodules was higher, the corresponding functions of the modules were lower, which resulted in decline in system robustness.

In the analysis of gene coexpression network modules, compared with the normal group, the number of submodules in the liver cancer group is significantly increased, the synergy of gene expression is lower, and the cell regulation system tends to be chaotic. From this, we infer that the robustness of the system is reduced, cancerous cells cannot complete all the functions of normal cells, and the system function is mainly inclined to the direction of cell proliferation, for example.

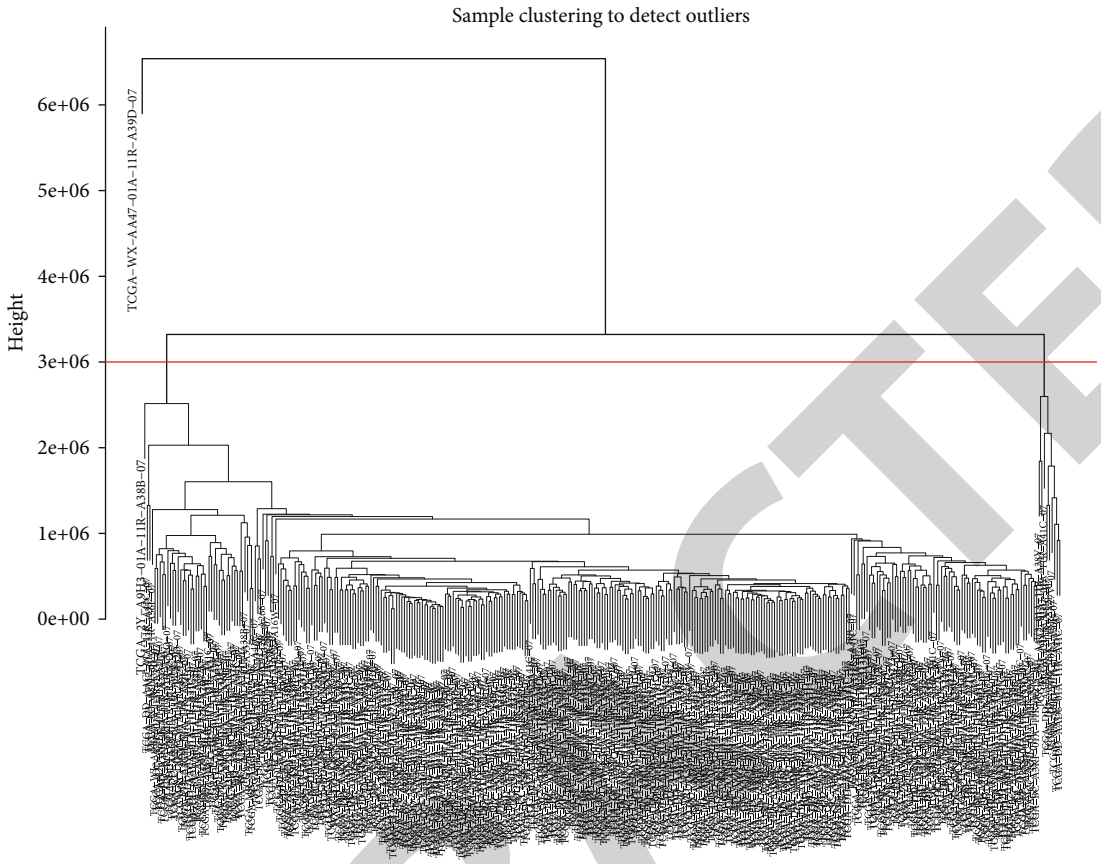
In order to further understand how the coexpression module in liver cancer samples differed from the normal group, we compared the module sets of the two groups of samples. For each coexpression module in the liver cancer group, we selected a module in the normal group with the largest gene overlap corresponding to it. Total 84 coexpression network modules of liver cancer genes were mapped to 16 gene coexpression network modules from the normal group (Figure 5 and Supplementary Table S5 and S6).

In normal samples, the "pink" module gene was enriched in the adrenergic receptor activity, which was reported to be related to cancer [33]. In order to better represent the gene correlation within the modules, the "pink" module (637 genes) in the normal group was selected. We screened the gene pair correlation results calculated by DGCA (absolute value of correlation greater than 0.3, and p value <0.05) and plotted the network diagram (Figure 6).

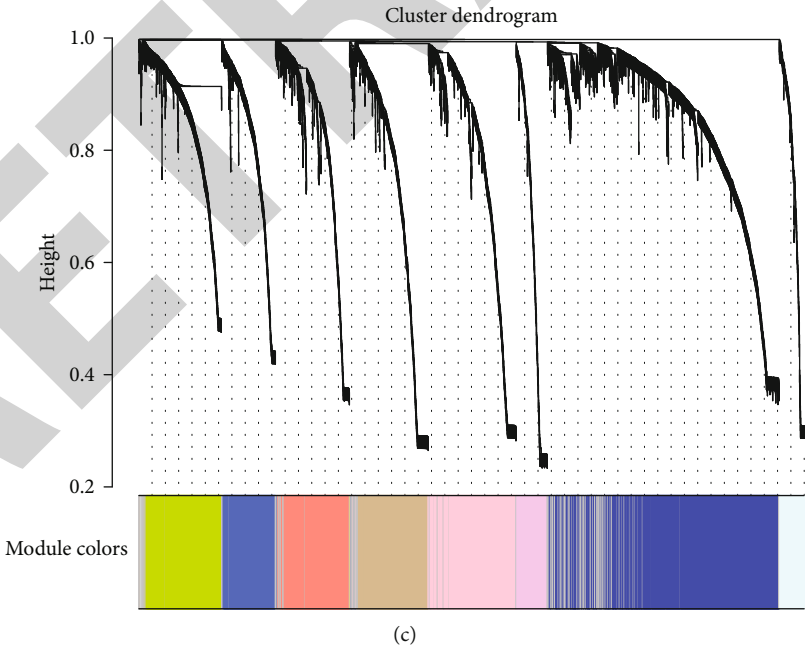


(a)

Figure 4: Continued.



(b)



(c)

Figure 4: Continued.

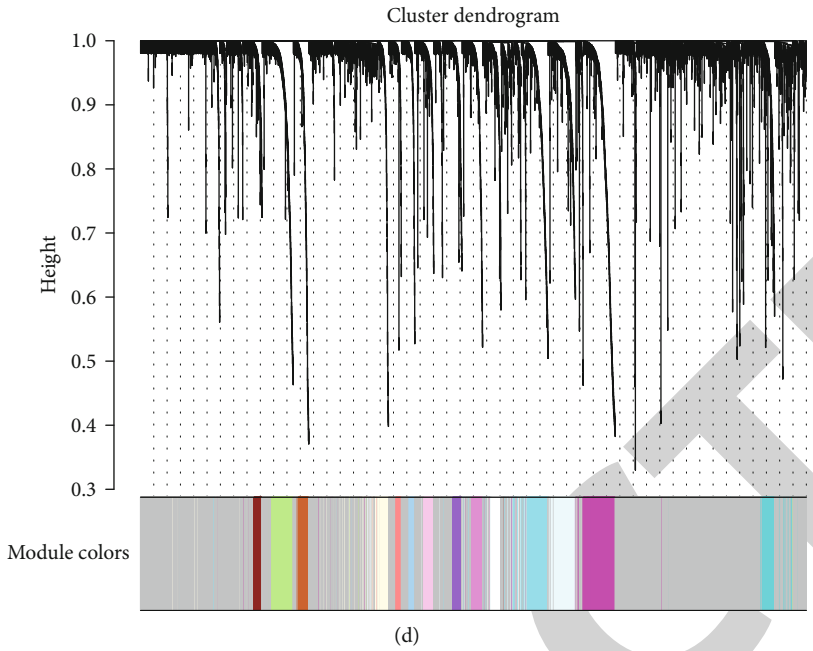


Figure 4: WGCNA analysis results. (a) Hierarchical clustering results of the normal group (cutoff = $2e-6$). (b) Hierarchical clustering results of the liver cancer group (cutoff = $3e-6$); both (a) and (b) select the branch with the most samples for subsequent analysis. (c) Normal group coexpression module (45 gene modules, except gray module). (d) Liver cancer sample coexpression module (84 gene modules, except gray module).

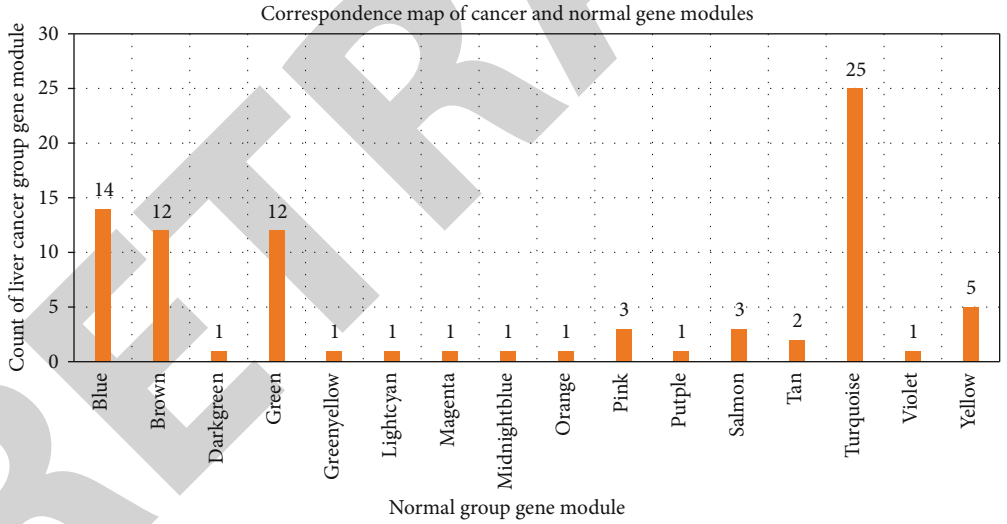


Figure 5: Correspondence map of cancer and normal gene modules. The x-axis is the gene coexpression module of the normal group, and the y-axis is the corresponding number of gene coexpression modules of each module in the liver cancer group.

By analyzing the pathway enrichment of differentially expressed genes and comparing the changes in the coexpression module in liver cancer and normal tissues, we could understand the pathogenesis of cancer at the system level, which provides a reference for our drug therapy. By interfering with the expression of differentially expressed genes, we can reverse the differentially expressed genes in liver cancer cells and restore their normal expression levels.

3.3. Repositioning of Potential Drugs for Treating HCC. We used drug perturbation-induced gene expression signatures obtained from CREEDS to compare the genes whose expression was significantly different ($p < 0.00001$) in liver cancer and normal samples and calculated the intersection of the genes in each drug perturbation-induced gene expression signature with significantly different genes. The results of the comparison are shown in Supplementary Table S7, S8, and S9. The number of overlaps between the genes upregulated by drugs and the genes downregulated in liver cancer samples was sorted from large to small (p value <0.05, 3,576 results). The number of overlaps between the genes downregulated by drugs and the genes upregulated

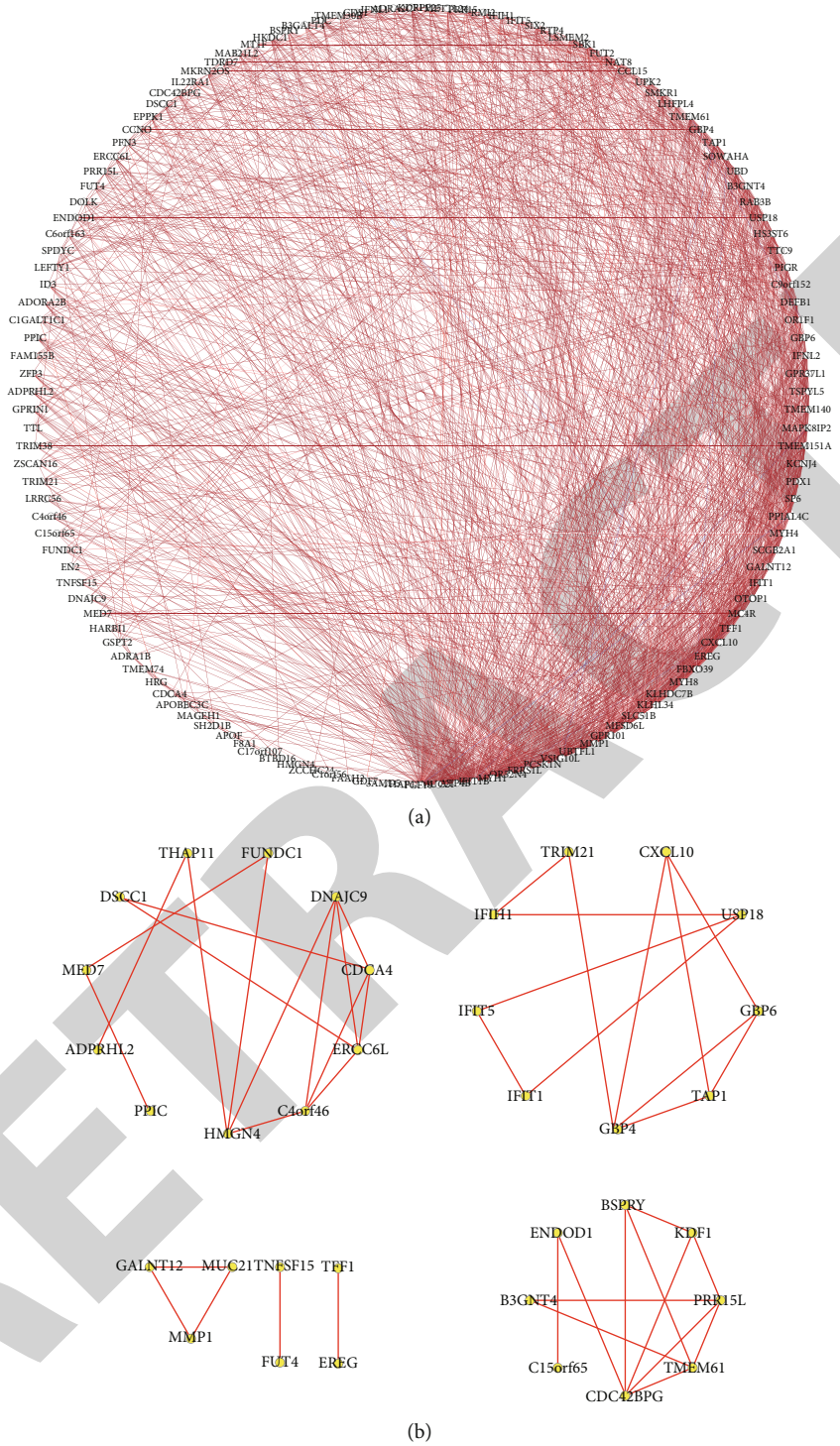


FIGURE 6: Correlation of genes in the pink module in cancer and normal samples. (a) Normal group gene association analysis; (b) cancer group gene association analysis. Data with an absolute value of the correlation coefficient greater than 0.6 and a confidence level of less than 0.05 are shown. The red edges indicate a positive correlation, and the blue indicates a negative correlation. Most genes are positively related.

in liver cancer samples was also sorted from large to small (p value <0.05, 4 results). Under each screening condition, the top five drug or small-molecule perturbations were observed (duplicate tags were skipped) to see if they could treat liver cancer (Table 1). Most of the genes regulated by these drugs were significantly differentially expressed in gastric cancer (Table S4). This may provide new ideas and directions for the use of drugs.

Table 1: List of potential therapeutic drugs for liver cancer.

Type	Drug/small molecule	Possible effect	Evidence (DOI)
a	Nutlin	Nutlin kills cancer cells via mitochondrial p53	10.4161/cc.8.11.8746 10.1124/mol.114.091603 10.1016/j.yexcr.2017.04.009
a	Paclitaxel|eribulin	Antimitotic agent	10.1159/000489067 10.1093/annonc/mdu435.56
a	17β-Estradiol	17β-Estradiol suppresses HCC cell proliferation and xenograft tumor development by inducing apoptosis	10.1371/journal.pone.0153863 10.1124/mol.63.6.1373
a	Enzalutamide	Inhibiting androgen; inhibition of tumor growth in response to androgen stimulation	10.1158/0008-5472.SABCS13-P2-09-05 10.1002/cncr.30336
b	6α-Methylprednisolone	Strong anti-inflammatory effect	10.1002/ddr.430020113 10.1111/nmo.12391
b	Histone deacetylase inhibitor	Effect on cell cycle regulation in hepatoma cells	10.7150/jca.34091 10.5392/IJoC.2012.8.4.074 10.1002/jcp.1087
b	4,4′-Diaminodiphenylmethane|inhibitor|(2Z,3Z)-bis{amino[(2-aminophenyl)sulfanyl]methylidene}butanedinitrile	Inhibitors of hepatocarcinogenesis	10.1093/carcin/2.10.1033
b	L-Proline residue|DNA damage-inducible transcript 3 protein	L-Proline-rich proteins and critical L-proline residues have a positive effect on the treatment of liver cancer	10.3892/mmr.1.4.459
b	Hemin	Inhibition of proliferation and induction of apoptosis of hepatoma cells	10.1196/annals.1299.055

*Type "a" refers to the drugs that are used to improve the downregulated genes in liver cancer; type "b" refers to the drugs that are used to improve the upregulated genes in liver cancer. The drug label and possible mechanism of action are shown in the table, and literature support is provided.

4. Discussion

Analyzing the pathways enriched by the differential genes of human HCC (Supplementary TableS1 and S2), we found that there are many pathways related to α-adrenergic receptors, such as "α-adrenergic receptor activity" (GO:0004936), "adrenergic receptor activity" (GO:0004935), and "epinephrine binding" (GO:0051379). We suspect that this is obviously related to liver cancer. Through literature research, it has been reported that human HCC can cause profound changes in the hepatic α-adrenergic receptor signal transduction pathway and may lead to carbohydrate-related metabolic dysfunction and wasting syndrome in cancer patients [33]. There are also pathways related to metal ion metabolism, such as "detoxification pathway of copper ions" (GO:0010273), "detoxification of inorganic compounds" (GO:0061687), "stress response to metal ions" (GO:0097501), and "stress response to copper ion" (GO:0004935). The correlation between metal ions, such as copper, and liver cancer has also been reported [34]. Pathways that have been reported to be related to liver cancer also include "metallothioneins bind metals" (R-HSA:5661231) [30], "RNA Polymerase I Promoter Opening" (R-HSA:73728) [31], and "SIRT1 negatively regulates rRNA expression" (R-HSA:427359) [32]. These functional channels are all in our channel list, which also provides evidence to support the reliability of the data we analyze.

In the analysis of gene coexpression network modules, compared with the normal group, the number of submodules in the liver cancer group is significantly increased, the synergy of gene expression is lower, and the cell regulation system tends to be chaotic. From this, we infer that the robustness of the system is reduced. Cancerous cells cannot complete all the functions of normal cells, while the system functions are mainly tilted in a specific direction, such as cell proliferation.

During the screening of potential drugs, we discovered some potential small molecule drugs. The genes affected by these drugs overlap with the differentially expressed genes in liver cancer samples. And the genes whose expression is upregulated by drugs are significantly downregulated in liver cancer samples. Therefore, we believe that these drugs have the potential to affect gene expression in liver cancer samples and restore them to normal levels.

This paper proposes new ideas for the pathogenesis study and drug treatment for liver cancer. We used RNA sequencing read count data from liver cancer and normal samples as input files and gene expression characteristics induced by drug interference as reference files. By means of DESeq2 differential expression analysis, WGCNA gene coexpression network analysis, and comparative analysis of gene expression characteristics with drug effects, we explored the related pathway changes in liver cancer and obtained potential therapeutic drugs that highly matched the characteristics of alterations in gene expression due to liver cancer. This method is of great significance for systematically understanding the treatment mechanism, pathway change characteristics, and drug guidance for HCC. We will further analyze subnetwork modules, collect key gene sets related to cancer, and screen potential drugs for these key genes.

Because the liver cancer specimens collected by TCGA has come from multiple individuals and platforms, and the pathogenesis is different for each type of cancer, liver cancer contains multiple subtypes, which introduces certain errors and uncertainties in the analysis. In further studies, we will continue to improve this method and try to distinguish gene expression changes and pathway change characteristics of different subtypes. With the introduction and promotion of the concept of personalized medicine and the decreasing cost of next-generation sequencing, the proposed treatment method may become possible in the near future.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors' Contributions

Aimin Hu and Zheng Wei designed the project. Zuxiang Zheng, Bichao Luo, and Jieming Yi analyzed the data, performed the experiments, and wrote the manuscript. Xinhong Zhou and Changjiang Zeng modified and reviewed the manuscript.

Supplementary Materials

Table S1: GO enrichment analysis. Table S2: KEGG and Reactome pathway enrichment analysis. Table S3: Module Path Enrichment Analysis. Table S4: Module Path Enrichment Analysis. Table S5: statistics list. Table S6: gene coexpression module. Table S7: comparison results of disturbed gene sets of drugs. Table S8: drugs increase the expression of downregulated genes. Table S9: drugs increase the expression of upregulated genes. Supplementary Figure S1: box plot of RNA read count data from cancer samples and normal samples. *(Supplementary Materials)*

References

[1] A. B. Ryerson, C. R. Eheman, S. F. Altekruse et al., "Annual Report to the Nation on the Status of Cancer, 1975-2012, featuring the increasing incidence of liver cancer," *Cancer*, vol. 122, no. 9, pp. 1312–1337, 2016.

[2] M. C. Wong, J. Y. Jiang, W. B. Goggins et al., "International incidence and mortality trends of liver cancer: a global profile," *Scientific Reports*, vol. 7, no. 1, p. 45846, 2017.

[3] X.-R. Xu, J. Huang, Z.-G. Xu et al., "Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver," *Proceedings of the National Academy of Sciences*, vol. 98, no. 26, pp. 15089–15094, 2001.

[4] D. Anwanwan, S. K. Singh, S. Singh, V. Saikam, and R. Singh, "Challenges in liver cancer and possible treatment approaches," *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, vol. 1873, no. 1, p. 188314, 2020.

[5] Y. Cai, J. Yang, T. Huang, and M. W. Wang, "Editorial: computational methods in predicting complex disease associated genes and environmental factors," *Frontiers in Genetics*, vol. 12, p. 679651, 2021.

[6] H. Liu, C. Qiu, B. Wang et al., "Evaluating DNA methylation, gene expression, somatic mutation, and their combinations in inferring tumor tissue-of-origin," *Frontiers in Cell and Development Biology*, vol. 9, p. 619330, 2021.

[7] B. He, C. Dai, J. Lang et al., "A machine learning framework to trace tumor tissue-of-origin of 13 types of cancer based on DNA somatic mutation," *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, vol. 1866, no. 11, p. 165916, 2020.

[8] B. He, J. Lang, B. Wang et al., "TOOme: a novel computational framework to infer cancer tissue-of-origin by integrating both gene mutation and expression," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 394, 2020.

[9] C. Liu, D. Wei, J. Xiang et al., "An improved anticancer drug-response prediction based on an ensemble method integrating matrix completion and ridge regression," *Mol Ther Nucleic Acids*, vol. 21, pp. 676–686, 2020.

[10] J. Fu and H. Wang, "Precision diagnosis and treatment of liver cancer in China," *Cancer Letters*, vol. 412, pp. 283–288, 2018.

[11] J. Sun, Q. Wei, Y. Zhou, J. Wang, Q. Liu, and H. Xu, "A systematic analysis of FDA-approved anticancer drugs," *BMC Systems Biology*, vol. 11, no. S5, p. 87, 2017.

[12] E. Y. Chen, V. Raghunathan, and V. Prasad, "An overview of cancer drugs approved by the US Food and Drug Administration based on the surrogate end point of response rate," *JAMA Internal Medicine*, vol. 179, no. 7, pp. 915–921, 2019.

[13] X. Liu, J. Yang, Y. Zhang et al., "A systematic study on drug-response associated genes using baseline gene expressions of the Cancer Cell Line Encyclopedia," *Scientific Reports*, vol. 6, no. 1, p. 22811, 2016.

[14] J. Yang, S. Peng, B. Zhang et al., "Human geroprotector discovery by targeting the converging subnetworks of aging and age-related diseases," *Geroscience*, vol. 42, no. 1, pp. 353–372, 2020.

[15] X. Xu, H. Long, B. Xi et al., "Molecular network-based drug prediction in thyroid cancer," *International Journal of Molecular Sciences*, vol. 20, no. 2, p. 263, 2019.

Hindawi
BioMed Research International
Volume 2023, Article ID 9803894, 1 page
https://doi.org/10.1155/2023/9803894

Retraction

Retracted: Identification of Key Genes and Immune Infiltrate in Nonalcoholic Steatohepatitis: A Bioinformatic Analysis

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z.-y. Jiang, Y. Zhou, Z. Lu, S.-w. Li, and B.-m. Wang, "Identification of Key Genes and Immune Infiltrate in Nonalcoholic Steatohepatitis: A Bioinformatic Analysis," *BioMed Research International*, vol. 2021, Article ID 7561645, 15 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 7561645, 15 pages
https://doi.org/10.1155/2021/7561645

Research Article

Identification of Key Genes and Immune Infiltrate in Nonalcoholic Steatohepatitis: A Bioinformatic Analysis

Zhen-yu Jiang,[1,2,3] Yi Zhou,[1,2,3] Lu Zhou,[2] Shao-wei Li,[4,5,6] and Bang-mao Wang[2]

[1]*Department of Graduate School of Tianjin Medical University, Tianjin, China*
[2]*Department of Gastroenterology, General Hospital of Tianjin Medical University, Tianjin, China*
[3]*Department of Gastroenterology, The Second Affiliated Hospital of Baotou Medical College, Inner Mongolia University of Science and Technology, Baotou, China*
[4]*Key Laboratory of Minimally Invasive Techniques & Rapid Rehabilitation of Digestive System Tumor of Zhejiang Province, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, Zhejiang Province, China*
[5]*Department of Gastroenterology, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, Zhejiang Province, China*
[6]*Institute of Digestive Disease, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, Zhejiang Province, China*

Correspondence should be addressed to Shao-wei Li; li_shaowei81@hotmail.com and Bang-mao Wang; gi.tmuh@yeah.net

Received 15 April 2021; Revised 25 August 2021; Accepted 26 August 2021; Published 13 September 2021

Academic Editor: Min Tang

Background. Nonalcoholic steatohepatitis (NASH) can progress to cirrhosis and hepatic carcinoma and is closely associated with changes in the neurological environment. The discovery of new biomarkers would aid in the treatment of NASH. *Methods*. Data GSE89632 were downloaded from the Gene Expression Omnibus (GEO) database, and *R* package "limma" was used to identify differentially expressed genes (DEGs) for NASH vs. normal tissues. The STRING database was used to construct a protein-protein interaction (PPI) network, and the Cytoscape software program (Version 3.80) was used to visualize the PPI network and identify key genes. The immune infiltration of NASH was determined using the *R* package "CIBERSORT". *Results*. We screened 41 DEGs. GO and KEGG enrichment analyses of the DEGs revealed the enrichment of pathways related to NAFLD steatosis and inflammation. A PPI network analysis was also performed on the DEGs, and seven genes (MYC, CXCL8, FOS, SOCS1, SOCS3, IL6, and PTGS2) were identified as hub genes. An immune infiltration assessment revealed that macrophages M2, memory resting CD4+ T cells, and $\gamma\Delta$ T cells play important roles in the immune microenvironment of NASH, which may be mediated by the seven identified hub genes.

1. Introduction

Nonalcoholic fatty liver disease (NAFLD) is characterized by hepatic steatosis or the accumulation of triglycerides in hepatic cells [1] and is directly related to obesity, hypertension, insulin resistance, and increased levels of blood lipids [2]. The prevalence of NAFLD has been increasing in recent years as the world economy grows and lives improve, with a global prevalence of about 25%.

NAFLD can progress to a more severe form called nonalcoholic steatohepatitis (NASH). NASH is a serious consequence of the development of NAFLD and is characterized by the abnormal accumulation of liver fat and immune cell infiltration caused by chronic hepatitis and inflammation [3]. Most NASH appears to have the consequence of further developing into liver fibrosis and can lead to serious diseases, such as cirrhosis and hepatocellular carcinoma, some of which are also associated with an increased risk of cardiovascular disease. These factors significantly add to the clinical complexity of such patients and increase the difficulty of clinical management. Even more problematic is the fact that there is currently no effective treatment for NASH-related hepatocellular carcinoma (HCC) [4]. However, the factors influencing the progression from NAFLD to NASH and

even to HCC are unclear; although, multiple hypotheses have been proposed, including lipotoxicity, oxidative stress, and proinflammatory mediators [5].

In studies of animal models, it has been found that animals with high fat and cholesterol levels exhibit typical pathological features of NAFLD, including hepato- and splenomegaly, early NASH histopathology, hypercholesterolaemia, elevated serum liver enzyme levels, and increased levels of proinflammatory cytokines [6]. A recent study has shown that fat intake is strongly associated with cognitive decline and increased risk of dementia [7]. Given that the high-fat, high-cholesterol diet associated with NASH has the ability to alter the neural environment and thus cause decline of cognition, studying the mechanisms of NASH is urgent for finding potential therapeutic targets and developing new treatments. However, it is important to elucidate the mechanisms underlying NAFLD in detail, as the prevalence and progression mechanisms of NAFLD are still unclear [8].

Bioinformatics has been widely used as a new approach to explore different disease mechanisms. For instance, it can be used for the early diagnosis and prognostic assessment of cancer by screening tumor-related biomarkers from large data repositories [9]. The method is equally useful for identifying key genes associated with NAFLD and NASH, thus furthering the exploration of the mechanisms underlying the development of NASH [10]. The development of NASH can lead to intestinal malnutrition and brain communication disorders through the disruption of metabolic processes and neurotransmission. A bioinformatic analysis of NASH development will allow us to further investigate the impact of the hepatic-intestinal-brain axis on the development of NASH to develop prevention and treatment strategies.

Therefore, we screened for genes with significantly different expressions in NASH patients and healthy individuals using data obtained from public databases. We then identified key genes, analyzed their potential biological functions, and studied the immune infiltration of NASH to further explore the factors associated with NASH and HCC in the future.

2. Materials and Methods

2.1. Expression Profile Data Acquisition of NASH. Initially, we chose to search for NASH-associated gene expression datasets in the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/).To meet research needs, we set the following standards: (1) data must include normal and NAHS livers, and (2) the samples were from humans. Ultimately, we selected GSE89632 for our study. GSE89632 was based on the GPL14951 platform Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip. The dataset of GSE89632 includes 63 valid samples, including 20 with simple steatosis (SS), 19 with NASH and 24 health controls (HCs).

2.2. Data Process and DEGs of NASH Identified. We downloaded the gene expression matrix of GSE89632 and the annotation file of GPL14951. First, we replace the ID of the probe with the gene symbol, and if they have the same probe, the max one will be retained. According to the aim of our study, we retained expression profiling data for only 24 HCs and 19 NASH patients. As the data had already been processed in a "quantile normalized" fashion, we proceeded directly to the next analysis. The "limma" package was used to identify differentially expressed genes (DEGs) between NASH patients and HCs; genes were considered DEGs if they had $|\log FC| > 2$ and P value < 0.05.

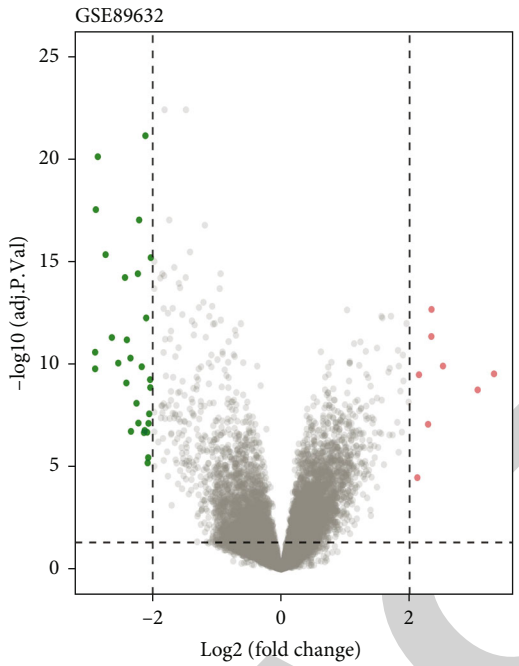
2.3. GO and KEGG Enrichment Analyses. GO and KEGG enrichment analyses were used to explore the potential biological functions of DEGs in NASH patients and HCs. The package "clusterprofiler" was used to implement the analysis procedure described above, and the analysis results were presented as bubble plots via the package "ggplot2."

2.4. Protein-Protein Interaction (PPI) Network and Hub Gene Identified. The PPI network of DEGs was constructed by the String dataset (https://www.string-db.org/). The minimum required interaction score was set to 0.4, and the disconnected nodes in the network were hidden. We exported the result obtained by the String database in "TSV" format and subsequently imported it into the Cytoscape software program (Version 3.80). The function "MCODE" of Cytoscape was used to explore the key module, and "cytohubba" was used to identify the hub genes from the network. Default settings were adopted for both "cytohubba" and "MCODE."

2.5. The Assessment of Immune Infiltration in NASH. CIBERSORT, a tool for evaluating the proportions of 22 subtypes of human immune cells, was used to evaluate the immune infiltration of NASH. We chose to calculate the immune infiltration using the "*R*" software program, among the perm was set at 1000. Results were visualized by multiple graphics. Boxplots were used to demonstrate the overall immune cell infiltration in NASH, and a stacked histogram was used to visualize infiltration in individual samples. A t-test was used to determine the different infiltration between NASH patients and HCs, and the results were depicted by boxplots. We constructed a correlation matrix of the seven hub genes and TME of NASH, and package "corrplot" was used for visualization.

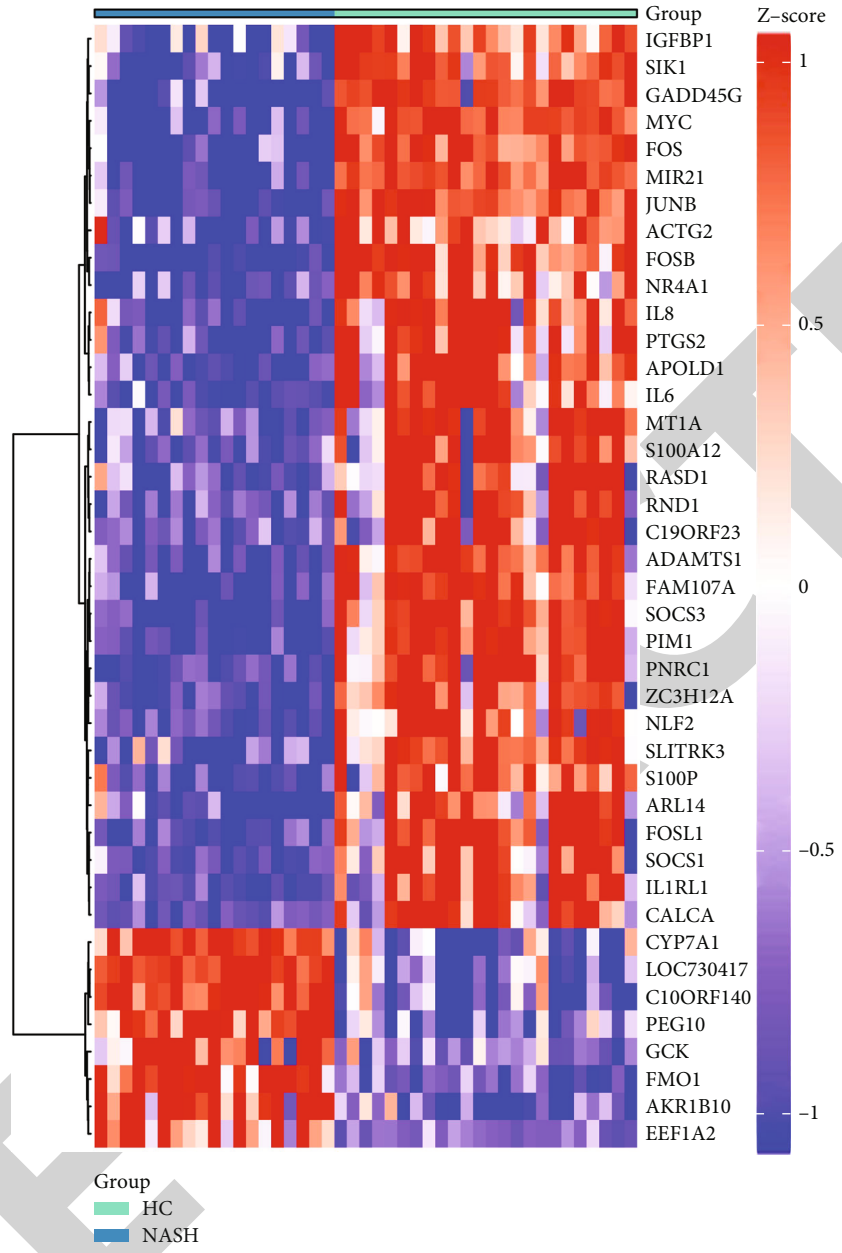
3. Results

3.1. Identification of DEGs in NASH. To guarantee the accuracy of the DEGs, we set a P value < 0.05 and [log 2 foldchange (log 2FC)] > 2 as the cut-off criteria. Ultimately, 41 DEGs were identified, most of which (33/41; 80.49%) were downregulated in NASH patients, with only a small proportion (8/41; 19.51%) upregulated. Volcano plots were used to depict DEGs, with green indicating downregulated genes, red upregulated genes, and gray genes with unsatisfactory conditions (Figure 1(a)). The heat map illustrates the expression profiles of the 41 DEGs in the 2 groups, with blue representing the NASH group, green the HCs, red the DEGs with upregulated expression, and dark blue the DEGs with the downregulated expression (Figure 1(b)).



(a)

Figure 1: Continued.

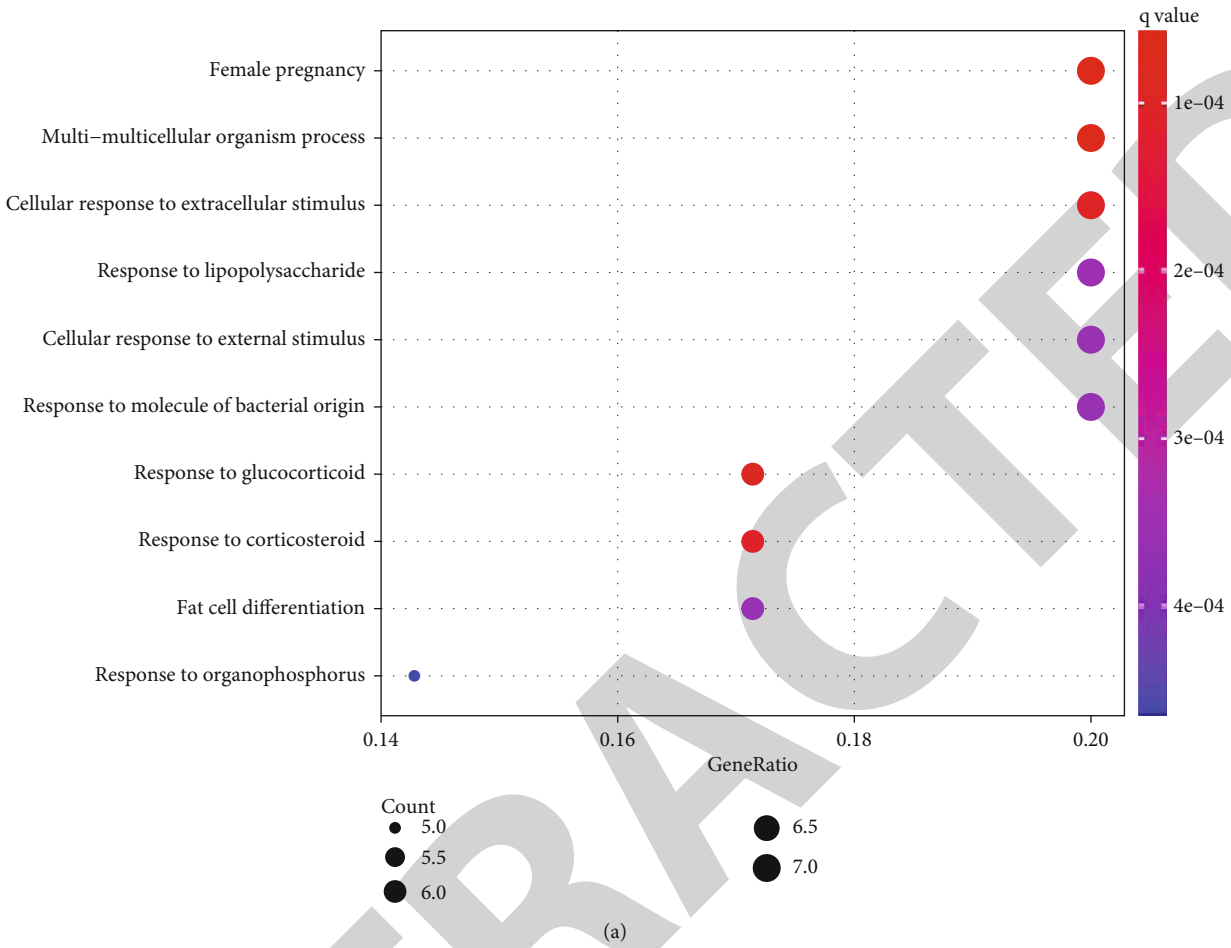


(b)

Figure 1: Differentially expressed genes (DEGs) in NASH vs. normal tissues. (a) A volcano plot of all DEGs. (b) Heat map of DGEs.

3.2. Results of the Function Enrichment Analysis of 41 DEGs. The occurrence of diseases often involves a variety of biological functions. We therefore used the 41 DEGs to describe significant pathways that might be involved in NASH. A GO enrichment analysis revealed the pathways of "female pregnancy," "multi–multicellular organism process," "cellular response to extracellular stimulus," "response to lipopolysaccharide," "cellular response to external stimulus," "response to molecule of bacterial origin," "response to glucocorticoid," "response to corticosteroid," "fat cell differentiation," and "response to organophosphorus" to be involved in the process of NASH (Figure 2(a)). Furthermore, the KEGG enrichment analysis suggested that the pathways of "osteoclast differentiation," "IL−17 signaling," "TNF signaling," "Janus kinase/signal transducer and activator of transcription (JAK/STAT) signaling," "microRNAs in cancer," "prolactin signaling," "growth hormone synthesis," "secretion and action," "type II diabetes mellitus," "leishmaniasis," and "colorectal cancer" might be biological pathways altered in NASH compared with normal tissues (Figure 2(b)).

3.3. Identification of the DEG PPI Network and Hub Genes. We set the minimum required interaction score at 0.4 when performing the PPI network analysis of DEGs using a string database, hiding the connected nodes in the network. This network totaled 39 nodes with 69 edges, and the vast



(a)

Figure 2: Continued.

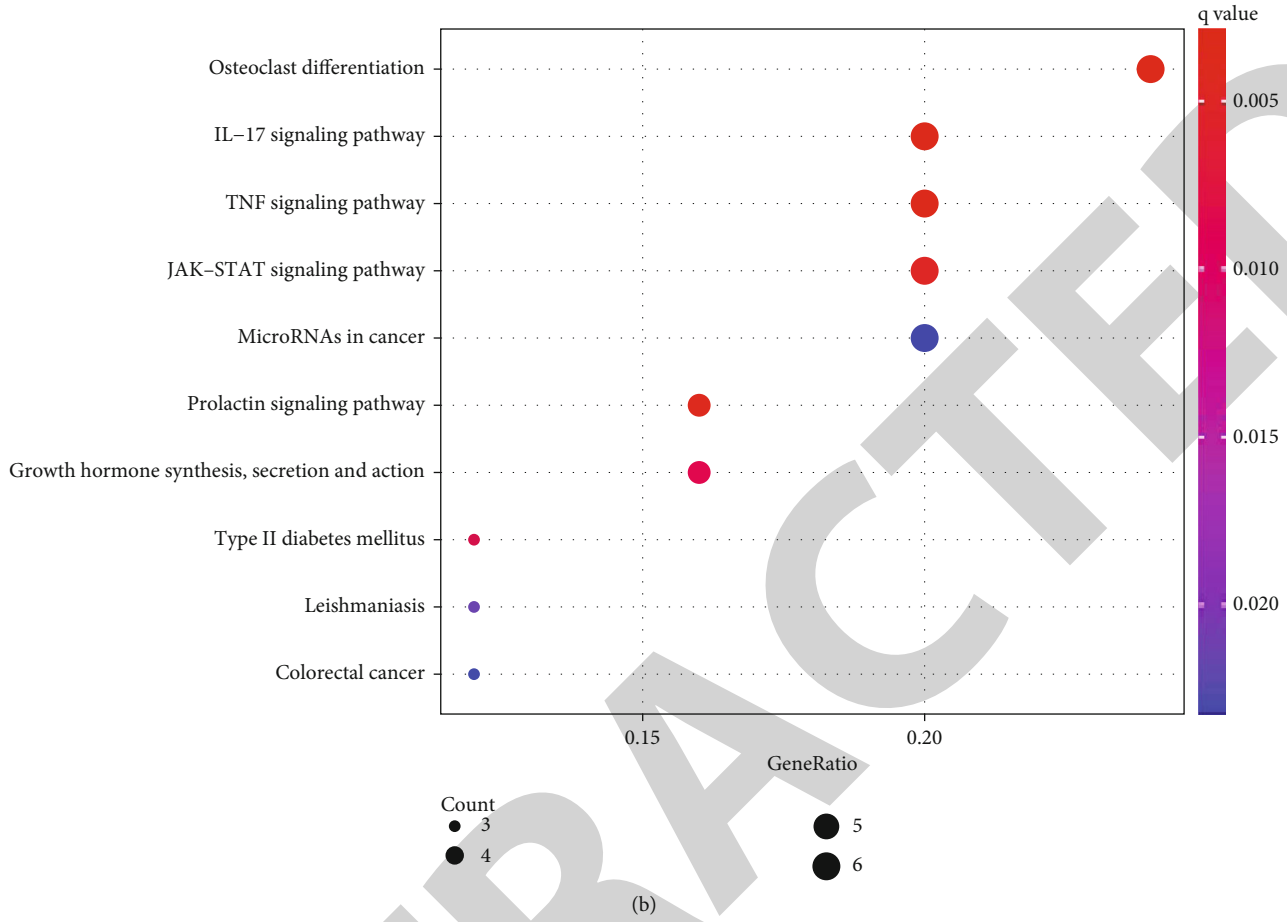


Figure 2: A functional enrichment analysis of DEGs shown by bubble plots. (a) The results of a GO enrichment analysis. (b) The results of a KEGG enrichment analysis.

majority of the constituent networks were downregulated genes. We visualized the network using the Cytoscape software program. As shown in Figure 3(a), blue represents downregulated genes, while red represents upregulated genes. We subsequently analyzed the module via the MCODE function and finally obtained two key modules (Figures 3(b) and 3(c)). The larger module scored 6.571, including "NR4A1," "MYC," "CXCL8," "FOS," "SOCS3," "IL6," "PTGS2," and "SOCS1." The other module included "FOSB," "JUNB," and "FOSL1," scoring 3. Goth core modules included downregulated genes, demonstrating the importance of downregulation and the interaction of these genes in NASH. To identify the key candidate genes from the network, we analyzed the main network using cytohubba and obtained the top 10 genes (see Table 1), of which 7 were included in the larger submodule; we identified these 7 genes as hub genes.

3.4. Immune Cell Infiltration in NASH and Its Relationship with Hub Genes. The 22 human immune cell proportions of NASH are shown in Figure 4(a). The top three categories were M2 macrophages, T cell CD4 memory resting, and $\gamma\Delta$T cells. These three categories accounted for a large proportion of NASH immune cell infiltration (>50%). Figure 4(b) shows the proportion of immune cell subtypes in individual samples, and it can be seen that there are differences in the infiltration of each sample, but macrophages and T cells comprised a larger proportion of each sample than other subtypes. Figure 4(c) shows the different immune cell occupancies in NASH and normal liver tissue. M2 macrophages, T cell CD4 memory resting, and $\gamma\Delta$T cells were significantly increased in NASH patients compared with HCs, which is again consistent with the high percentage of these cells in NASH patients, underscoring the importance of these three cell types in the immune microenvironment of NASH.

Subsequently, we explored the relationship between the seven hub genes and the tumor microenvironment (TME) in NASH. "Naïve B cells," "plasma cells," "$\gamma\Delta$T cells," "monocytes," "M2 macrophages," "resting mast cells," and "activated mast cells" showed a consistent correlation among the seven hub genes, indicating that these genes are involved in the formation of the specific immune cell infiltration pattern of NASH.

4. Discussion

NASH is characterized by the abnormal accumulation of liver fat and immune cell infiltration due to chronic inflammation, which can lead to liver diseases, such as cirrhosis and hepatocellular carcinoma, and is strongly associated

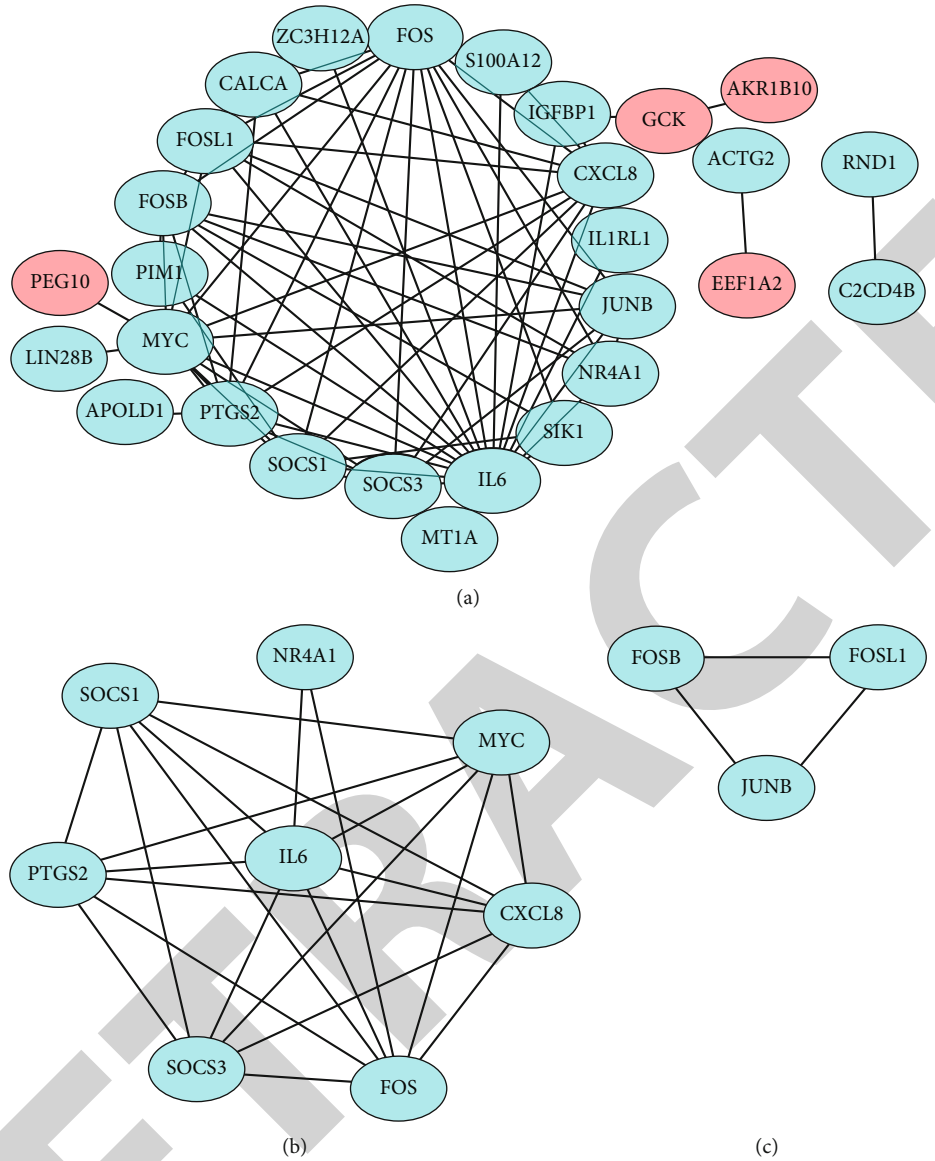


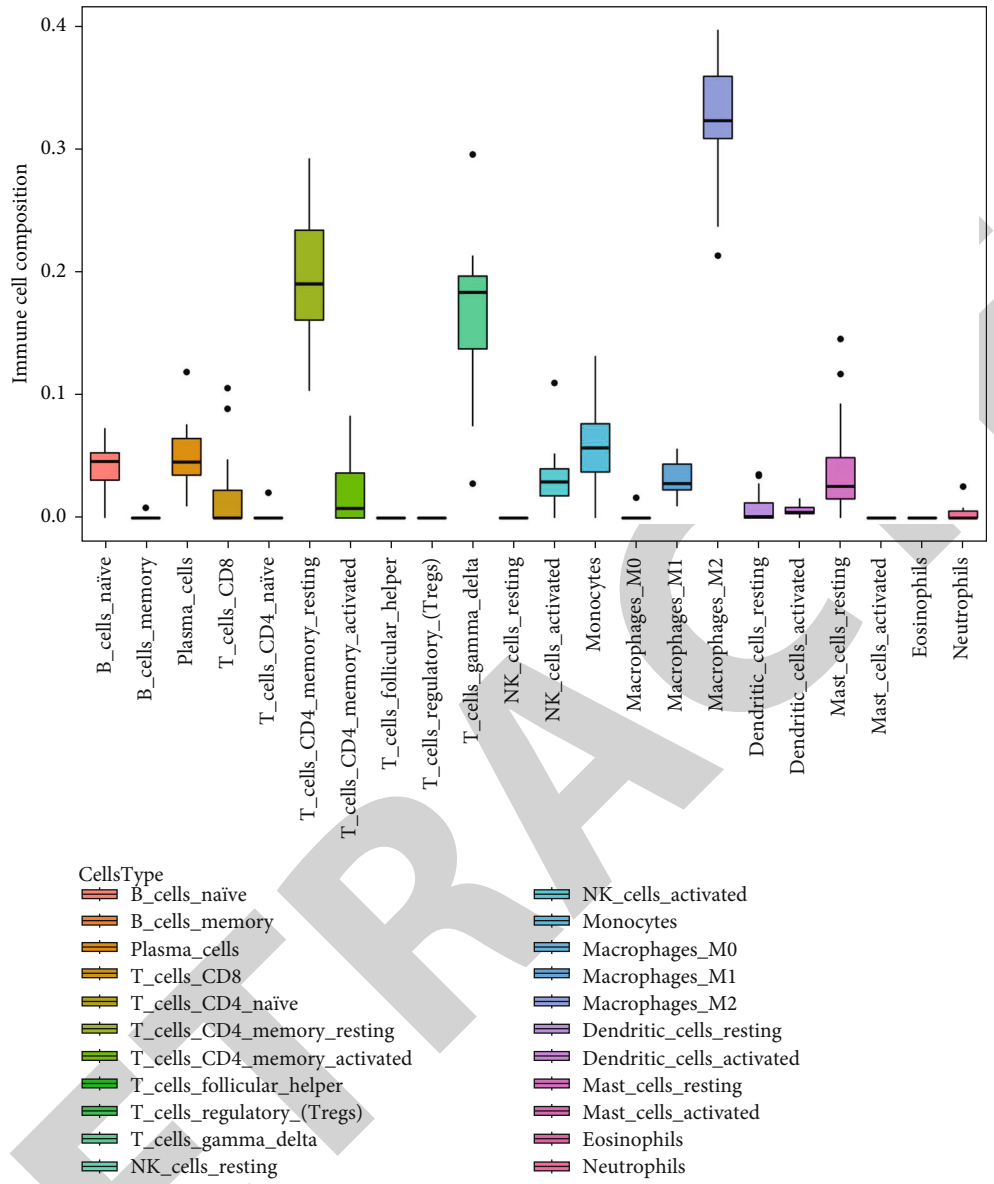
Figure 3: A functional protein-protein interaction (PPI) network. (a) PPI network of DEGs of NASH. Blue represents downregulated genes, while red represents upregulated genes. (b, c) Key modules of the PPI network.

Table 1: Top 10 DEGs in the network ranked by the MCC method.

Rank	Name	Score
1	IL6	1101
2	FOS	1082
3	MYC	926
4	CXCL8	774
5	PTGS2	769
6	SOCS3	744
7	SOCS1	726
8	FOSB	294
9	FOSL1	264
9	JUNB	264

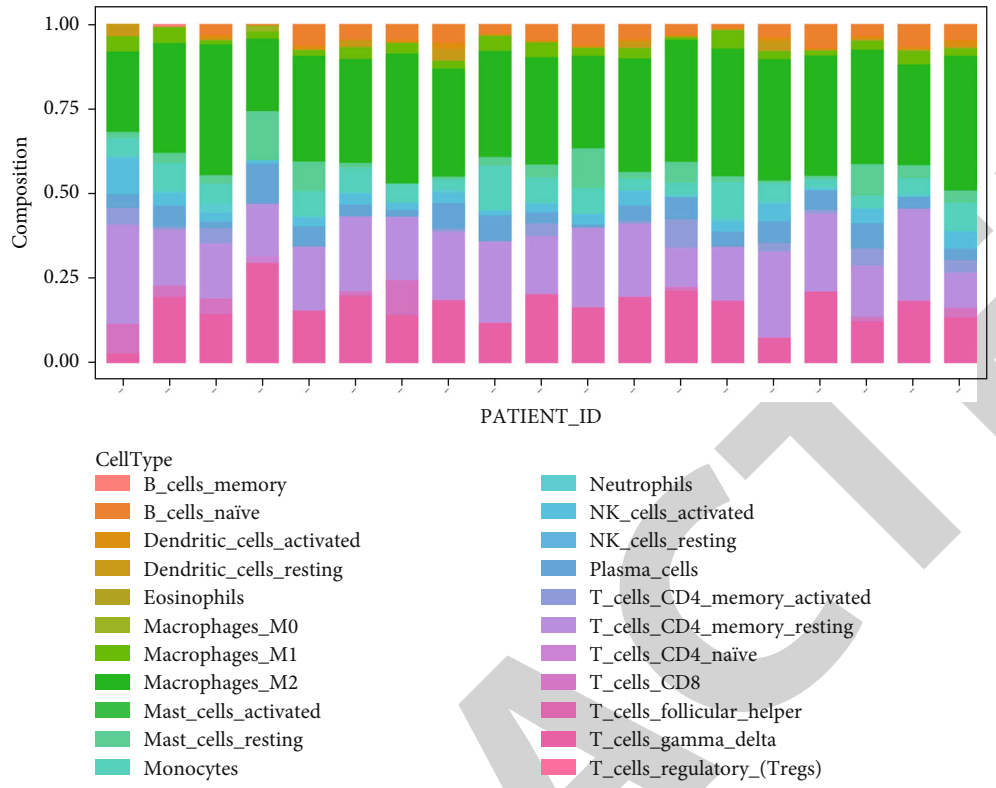
with cognitive decline and an increased risk of dementia in response to an altered neuroenvironment. The discovery of new biomarkers is therefore important for identifying potential therapeutic targets and developing new treatments.

To investigate the underlying biological mechanisms contributing to the development and progression of NASH, we performed GO and KEGG analyses of DEGs and identified many significant biological pathway enrichments. Th17 cells and IL-17 have been proven to promote the transformation from simple steatosis to steatohepatitis, which is closely related to NAFLD steatosis and inflammation [11]. The JAK-STAT signaling pathway has also been proven to be involved in the regulation of NASH. Abnormalities in the GH-JAK2-STAT5 signaling pathway altered the whole-body fat metabolism to promote peripheral lipolysis while enhancing hepatic fat synthesis, resulting in hepatic fat



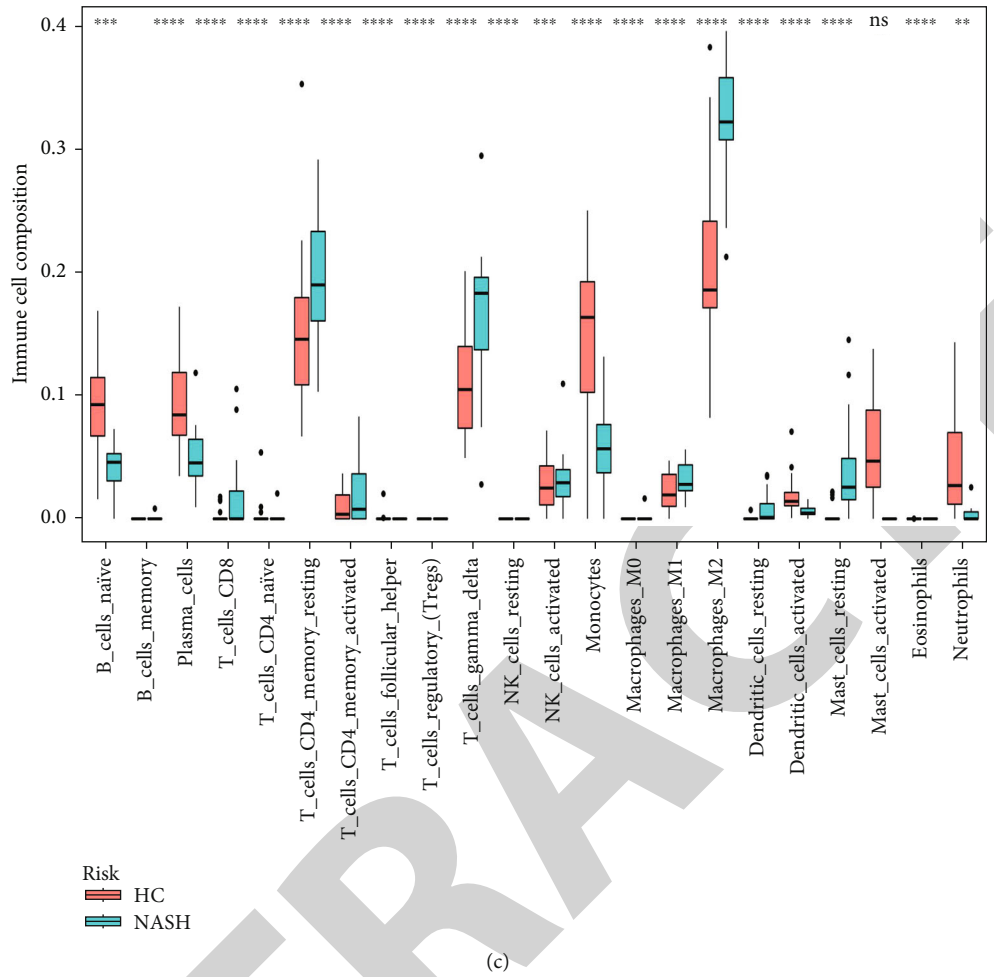
(a)

Figure 4: Continued.



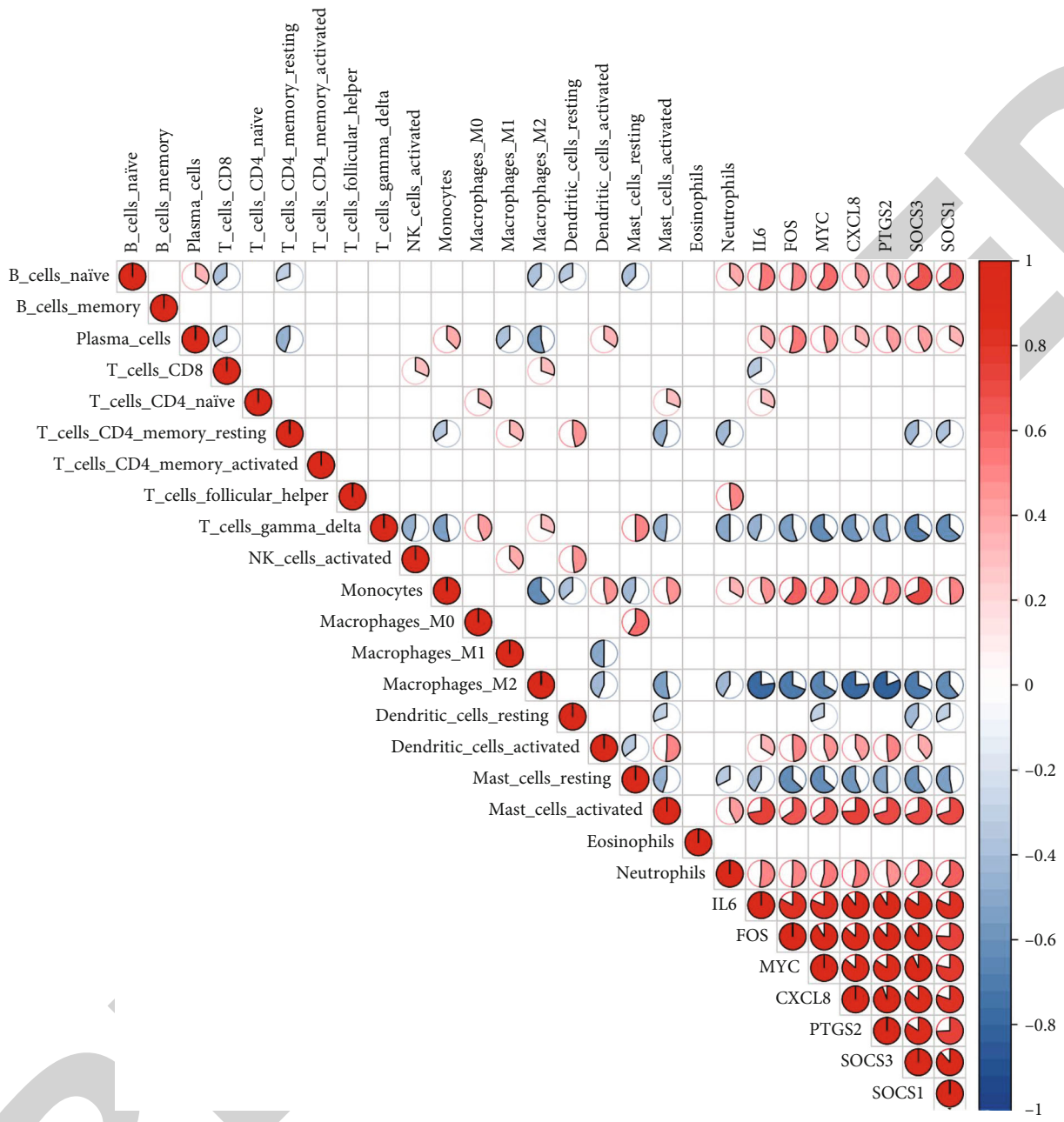
(b)

Figure 4: Continued.



(c)

Figure 4: Continued.



(d)

Figure 4: Immune infiltration in NASH. (a, b) The proportion of 22 immune cell subtypes in NASH. (c) The 22 subtypes of immune cells have different proportions in NASH and normal tissues. (d) Correlation matrix: blank represents no correlation, red represents a positive correlation, and blue represents a negative correlation.

accumulation and leading to the development of the NAFLD phenotype [12].

We also performed a PPI network analysis of DEGs and identified seven hub genes: "MYC," "CXCL8," "FOS," "SOCS1," "SOCS3," "IL6," and "PTGS2." MYC is a protooncogene encoding a nuclear-phosphorylated protein that plays a role in cell cycle progression, apoptosis, and cell transformation. Amplification of the MYC gene has been observed in a variety of cancers. In liver disease, highly aggressive and poorly differentiated HCC has been shown to be significantly associated with its overexpression [13]. Activation of the MYC gene has also been observed in HCC cells in mouse experiments [14]. After knockout of the MYC gene, the livers of mice developed NAFLD-specific and progressively similar features to NASH [15]. In our study, MYC was also found to be downregulated in NASH patients, which is consistent with the above studies and further suggests that MYC may be an important factor affecting hepatic lipid metabolism.

Chemokines are chemoattractants for leukocyte trafficking, growth, and activation in injured and inflammatory tissues. As inflammation is key in the transition from SS to NASH, chemokines may play an important role in the pathogenesis of NASH. There is growing evidence that the expression of chemokines and their receptors is elevated in the livers of obese patients with advanced steatosis and NASH [16]. CXCL8 is a member of the CXC chemokine subfamily secreted by these cells, which primarily regulates the recruitment of neutrophils to sites of inflammation. It has been shown that serum CXCL8 levels are significantly higher in patients with NASH than in those with SS or HCs. Meanwhile, CXCL8 is an essential multifunctional cytokine that promotes tumor metastasis and invasion [17]. It has been confirmed that HCC patients with an increased expression of CXCL8 always had a worse outcome than those with a low expression [18]. CXCL8 was mapped to the pathways related to LPS, TRL3, B cells, monocytes, and the KEGG toll-like receptor signaling pathway. However, the specific role of CXCL8 in the pathogenesis of NASH remains unclear, meriting a further indepth study [19].

The fos gene, or c-fos, is a major member of the fos gene family, which regulates transcription from promoters containing AP-1 activation elements. C-fos acts as a natural partner of c-jun to promote cell growth. However, the overexpression of c-jun protein in the absence of c-fos may lead to the abnormal formation of homodimer transcriptional complexes, thus disrupting the normal regulation mechanism of gene expression [20]. In studies of liver disease, c-jun has been shown to be strongly expressed in the liver of patients with acute hepatitis. In a NASH mouse model, c-Jun/AP-1 activation was found to be a key regulator of liver changes. An increase in c-jun may promote the development and progression of NASH [21]. Since the DNA binding domain of c-fos is homologous to the c-Jun DNA binding domain, we may speculate that c-fos also plays an important role in the occurrence and development of NASH [20]. In the occurrence and development of HCC, the liver-specific c-fos expression leads to reversible premalignant hepatocyte transformation and enhanced dEn-carcinogenesis. c-fos-expressing livers show necrotic foci, immune cell infiltration, and an altered hepatocyte morphology. Furthermore, increased proliferation, dedifferentiation, activation of the DNA damage response, and gene signatures of aggressive HCCs are also observed. Mechanistically, c-fos decreases the expression and activity of the nuclear receptor LXrα, leading to increased hepatic cholesterol and the accumulation of toxic oxysterols and bile acids [22]. Therefore, the role of c-fos in the disease progression of NASH or HCC is contradictory and mysterious, meriting further research [23].

Cytokine signaling repressor proteins (SOCS1-SOCS7 and cytokine-induced SH-2 containing proteins) are induced by proinflammatory cytokines and regulate cytokine signaling through the JAK/STAT pathway [24]. SOCS1 and SOCS3, as suppressors of cytokine signaling, have both been found to be associated with insulin resistance. However, the role of SOCS3 is more paradoxical and deserves to be explored [25, 26]. In NASH, the increased inflammatory response leads to an increase in inflammatory factors, such as interleukin-6 (IL-6)3 and tumor necrosis factor-α (TNF-α), which in turn leads to an increase in the SOCS3 expression [27]. SOCS3 protein is thought to promote insulin resistance by inhibiting insulin and leptin signaling in the inflammatory response. However, in mice, hepatic SOCS3 deficiency was found to enhance hepatic insulin sensitivity but paradoxically promotes NAFLD and inflammation in a hypertrophic environment with high glucose and fatty acids [28]. These findings reveal that SOCS3 plays an important negative regulatory role in insulin sensitivity, adipogenesis, and energy homeostasis as well as having a complex role in promoting the development of NASH. At the same time, SOCS3 is known to inhibit cytokine signaling via JAK/signal transducers, activators of transcription, NF-κB, and the focal adhesion kinase (FAK) signaling pathway. Substantial data have demonstrated the link between the SOCS3 regulation of inflammation and its suppressor activity on tumor initiation and development. In SOCS3 conditional knockout mice, deletion of the SOCS3 gene promoted carcinogen-induced hepatic tumor development through the activation of STAT3 and resistance to apoptosis. In HCC patient samples, the expression of SOCS3 was reduced compared with surrounding non-HCC regions [29]. Recently, hypermethylation of SOCS3 was shown to be associated with a poor clinical outcome in HCC patients with hepatitis B virus (HBV) infection backgrounds but not in those with hepatitis C virus (HCV) or no virus infection [30].

IL-6 is an inflammatory factor produced primarily at sites of acute and chronic inflammation. The IL-6 expression is significantly higher in steatotic livers than in normal livers, suggesting that IL-6 is involved in the progression of NASH [31]. It has also been found that IL-6 levels correlate significantly with NASH severity [32]. However, our results do not seem to support the above, as we found the IL-6 expression to be downregulated in NASH. The role of IL-6 in liver injury is complex and contradictory. IL-6 has been found to promote hepatocyte regeneration [33], impede apoptosis, and repair damaged cells [34] after liver injury. This suggests that the role of IL-6 in liver cells is multidirectional and not merely proinflammatory, which is well worth further study.

Prostaglandin-endoperoxide synthase (PTGS), also known as cyclooxygenase, is a key enzyme in prostaglandin biosynthesis. There are two isozymes of PTGS: constitutive PTGS1 and inducible PTGS2, which together are involved in prostaglandin biosynthesis during inflammation and mitosis. PTGS2, or cyclooxygenase-2, was found to be significantly associated with iron death, a new type of cell death that has been recently reported and extensively studied. Iron death is characterized by the contribution of iron to the development of oxidative cellular damage. Current studies suggest that it may be associated with different types of chronic liver diseases, including HCV, NASH, and HCC, as well as DILI. An imbalance in iron metabolism and reactive oxygen species- (ROS-) induced lipid peroxidation is thought to be mechanisms involved in the liver injury sustained in these diseases [29, 35, 36]. Furthermore, PTGS2 has been identified as a marker of iron sagging, along with

other markers, including changes in NADPH levels and lipid peroxidation [37, 38]. However, the effect of PTGS2 on NASH has yet to be investigated.

NASH has a chronic inflammatory manifestation and has been shown to have multiple immune cells directly associated with it. We analyzed the immune infiltration in the liver of NASH patients compared to normal patients using CIBERSORT and found significant differences in multiple subtypes of immune cells between HCs and NASH patients. The correlation between the seven hub genes and immune cells also revealed consistency in the expression of these genes with multiple immune cells. For example, all seven hub genes were underexpressed in NASH patients compared with HCs (Figure 3), and the expression of these genes was positively correlated with naïve B cells (Figure 4(d)), while the proportion of naïve B cells was reduced in NASH patients compared with HCs (Figure 4(c)). Taken together, we conclude that the low expression of these hub genes may mediate the decline in multiple immune cells associated with NASH. Those genes may be involved in shaping the immune microenvironment characteristic of NASH, but the exact mechanisms remain to be investigated.

The immune infiltration analysis showed that the proportion of M2 macrophages, memory resting CD4+ T cells, and $\gamma\Delta$ T cells was significantly elevated in NASH compared with normal hepatic cells, indicating the importance of these three types of immune cells in the immune microenvironment of NASH. The pathology of NASH is characterized by lipid-induced hepatocyte apoptosis and inflammatory cell infiltration, some of which are activated macrophages [39]. Macrophages have a unique phenotype and function. They are generally considered to be involved in the inflammatory response to NASH [40, 41]. Excessive lipid accumulation in the liver promotes the activation of Kupffer cells (KCs) and the production of inflammatory mediators (TNF-α and IL-1β, among others), leading to increased insulin resistance, liver inflammation, and the formation of the blood-brain barrier, which eventually results in the development of NASH [42]. M1 macrophages have the capacity to deliver antigens and secrete proinflammatory cytokines, whereas M2 macrophages downregulate the immune response by secreting the inhibitory cytokines IL-10, transforming growth factor-β (TGF-β,) and mannose receptor (Mrc) [43, 44]. Dysregulation of M1/M2 macrophages can lead to chronic inflammation, cancer, and NAFLD [45]. In recent years, it has been suggested that M2 macrophages may play a protective role against NAFLD by promoting apoptosis of M1 macrophages. However, the observation in our study that M2-like cells are much more abundant than normal cells in NASH seems to contradict the above view. This may be explained by the plasticity and flexibility of mononuclear phagocytes and their activation state, i.e., that polarized M1/M2 macrophages can reverse the phenotype to some extent both in vivo and in vitro [45, 46]. Another view is that the increase in M2-like cells is a consequence rather than a cause of NASH. The progression of NAFL to NASH is accompanied by apoptosis and repair of hepatocytes, and the increase in M2 macrophages may thus be related to hepatocyte repair and chronic inflammation of the liver.

As NASH is a chronic inflammatory disease, hepatic CD4+ T cell infiltration becomes a feature of NASH. An increase in human CD4+ central and effector memory T cells in the liver and peripheral blood has been observed, accompanied by a significant upregulation of proinflammatory cytokines (IL-17A and interferon gamma). Furthermore, the clearance of CD4+ T cells was shown to reduce liver inflammation and fibrosis in mice, and a study by Rau et al. highlighted an increase in CD4+ memory T cells at sites of intrahepatic fibrosis in patients with NAFLD, demonstrating that CD4+ memory T cell subsets are important drivers of progression from steatosis to fibrosis in NAFLD. However, there is a lack of directly relevant studies and explanations for the role of memory resting CD4+ T cells in NASH. Our findings may serve as a hint for future directions.

$\gamma\Delta$ T cells belong to a unique subset of innate lymphocytes. They can produce at least two distinct subsets: interferon- (IFN-) γ- and interleukin- (IL-) 17-producing $\gamma\Delta$ T cells. $\gamma\Delta$ T cells play a protective role in several pathophysiological processes, such as pathogen clearance, tumor surveillance, and tissue repair, while playing the opposite role in autoimmunity, allergy and carcinogenesis [47]. Thus, the role of $\gamma\Delta$ T cells in maintaining immune homeostasis and tissue homeostasis is complex and bidirectional, with these cells acting as both "protectors" and "destroyers." There is growing evidence that $\gamma\Delta$ T cells in the liver respond to liver-targeted injury and regulate the development of liver disease [48, 49], thus playing an important role in maintaining liver physiological homeostasis as well as regulating liver pathological processes. NAFLD is a chronic progression of the liver from SS to cirrhosis, while NASH is an important pathological change in the middle of this progression [50]. It has been shown that IL-17 accelerates the progression of NAFLD by recruiting neutrophils and inducing ROS [51]. $\gamma\Delta$ T cells are the main supplier of IL-17A [52]. An increase in IL-17A-producing $\gamma\Delta$ T cells can be seen in mice with NAFLD. $\gamma\Delta$ T cell deficiency protects mice from NAFLD damage, characterized by steatohepatitis IL-17A in CD161+ $\gamma\Delta$ T cells [48]. Thus, IL-17A-producing $\gamma\Delta$ T cells are a major regulator of the progression of NAFLD. However, there is still a lack of corresponding studies on the mechanisms underlying the direct and indirect effects of $\gamma\Delta$ T-cell subsets on NASH. The mechanism underlying the activation of $\gamma\Delta$ T-cell subsets also remains a mystery. Our findings of an increased and active expression of $\gamma\Delta$ T cells in NASH patients corroborate previous findings and may highlight a potential therapeutic approach to blocking the progression of SS to NASH and eventually to NAFLD, via blocking or eliminating IL-17A-producing $\gamma\Delta$ T cells.

5. Conclusion

We identified seven hub genes that are closely related to the development of NASH through a bioinformatics analysis and also clarified the immune infiltration of NASH through an immune cell subtype analysis, providing a useful

reference for future research on NASH biomarkers and the immune microenvironment.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there is no conflict of interest.

Authors' Contributions

Zhen-yu Jiang and Yi Zhou contributed equally to this work.

Acknowledgments

This work was supported in part by the Program of Inner Mongolia Autonomous Region Tumor Biotherapy Collaborative Innovation Center, Medical Science and Technology Project of Zhejiang Province (2021PY083), Program of Taizhou Science and Technology Grant (20ywb29), Major Research Program of Taizhou Enze Medical Center Grant (19EZZDA2), Open Project Program of Key Laboratory of Minimally Invasive Techniques & Rapid Rehabilitation of Digestive System Tumor of Zhejiang Province (21SZDSYS01, 21SZDSYS09), and Key Technology Research and Development Program of Zhejiang Province (2019C03040).

References

[1] S. Saadati, B. Hatami, Z. Yari et al., "The effects of curcumin supplementation on liver enzymes, lipid profile, glucose homeostasis, and hepatic steatosis and fibrosis in patients with non- alcoholic fatty liver disease," *European Journal of Clinical Nutrition*, vol. 73, no. 3, pp. 441–449, 2019.

[2] M. Dharmalingam and P. G. Yamasandhi, "Nonalcoholic fatty liver disease and type 2 diabetes mellitus," *Indian Journal of Endocrinology and Metabolism*, vol. 22, no. 3, pp. 421–428, 2018.

[3] S. Y. Koo, E. J. Park, and C. W. Lee, "Immunological distinctions between nonalcoholic steatohepatitis and hepatocellular carcinoma," *Experimental & Molecular Medicine*, vol. 52, no. 8, pp. 1209–1219, 2020.

[4] L. Liu, "Clinical features of hepatocellular carcinoma with hepatitis B virus among patients on Nucleos(t) ide analog therapy," *Infectious Agents and Cancer*, vol. 15, no. 1, p. 8, 2020.

[5] B. Cicuéndez, I. Ruiz-Garrido, A. Mora, and G. Sabio, "Stress kinases in the development of liver steatosis and hepatocellular carcinoma," *Molecular Metabolism*, vol. 50, article 101190, 2021.

[6] L. N. Tu, M. R. Showalter, T. Cajka et al., "Metabolomic characteristics of cholesterol-induced non-obese nonalcoholic fatty liver disease in mice," *Scientific Reports*, vol. 7, no. 1, p. 6120, 2017.

[7] S. G. Higarza, S. Arboleya, M. Gueimonde, E. Gómez-Lázaro, J. L. Arias, and N. Arias, "Neurobehavioral dysfunction in non-alcoholic steatohepatitis is associated with hyperammonemia, gut dysbiosis, and metabolic and functional brain regional deficits," *PLoS One*, vol. 14, no. 9, article e0223019, 2019.

[8] S. L. Friedman, B. A. Neuschwander-Tetri, M. Rinella, and A. J. Sanyal, "Mechanisms of NAFLD development and therapeutic strategies," *Nature Medicine*, vol. 24, no. 7, pp. 908–922, 2018.

[9] G. Feng, X. P. Li, C. Y. Niu et al., "Bioinformatics analysis reveals novel core genes associated with nonalcoholic fatty liver disease and nonalcoholic steatohepatitis," *Gene*, vol. 742, p. 144549, 2020.

[10] C. Cai, X. Song, and C. Yu, "Identification of genes in hepatocellular carcinoma induced by non-alcoholic fatty liver disease," *Cancer Biomarkers*, vol. 29, no. 1, pp. 69–78, 2020.

[11] Y. Tang, Z. Bian, L. Zhao et al., "Interleukin-17 exacerbates hepatic steatosis and inflammation in non-alcoholic fatty liver disease," *Clinical and Experimental Immunology*, vol. 166, no. 2, pp. 281–290, 2011.

[12] D. Kaltenecker, M. Themanns, K. M. Mueller et al., "Hepatic growth hormone - JAK2 - STAT5 signalling: Metabolic function, non- alcoholic fatty liver disease and hepatocellular carcinoma progression," *Cytokine*, vol. 124, article 154569, 2019.

[13] P. Kaposi-Novak, L. Libbrecht, H. G. Woo et al., "Central role of c-Myc during malignant conversion in human hepatocarcinogenesis," *Cancer Research*, vol. 69, no. 7, pp. 2775–2782, 2009.

[14] Z. J. Brown, Q. Fu, C. Ma et al., "Carnitine palmitoyltransferase gene upregulation by linoleic acid induces $CD4^+$ T cell apoptosis promoting HCC development," *Cell Death & Disease*, vol. 9, no. 6, p. 620, 2018.

[15] L. R. Edmunds, P. A. Otero, L. Sharma et al., "Abnormal lipid processing but normal long-term repopulation potential of myc-/- hepatocytes," *Oncotarget*, vol. 7, no. 21, pp. 30379–30395, 2016.

[16] Y. S. Roh and E. Seki, "Chemokines and chemokine receptors in the development of NAFLD," *Advances in Experimental Medicine and Biology*, vol. 1061, pp. 45–53, 2018.

[17] Q. Liu, A. Li, Y. Tian et al., "The CXCL8-CXCR1/2 pathways in cancer," *Cytokine & Growth Factor Reviews*, vol. 31, pp. 61–71, 2016.

[18] J. Zhu, Y. Zhou, L. Wang et al., "CXCL5/CXCL8 is a promising potential prognostic and tumor microenvironment-related cluster in hepatocellular carcinoma," *Journal of Gastrointestinal Oncology*, vol. 11, no. 6, pp. 1364–1380, 2020.

[19] K. L. Donnelly, C. I. Smith, S. J. Schwarzenberg, J. Jessurun, M. D. Boldt, and E. J. Parks, "Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease," *The Journal of Clinical Investigation*, vol. 115, no. 5, pp. 1343–1351, 2005.

[20] T. D. Halazonetis, K. Georgopoulos, M. E. Greenberg, and P. Leder, "c-Jun dimerizes with itself and with c-Fos, forming complexes of different DNA binding affinities," *Cell*, vol. 55, no. 5, pp. 917–924, 1988.

[21] C. Dorn, J. C. Engelmann, M. Saugspier et al., "Increased expression of c-Jun in nonalcoholic fatty liver disease," *Laboratory Investigation*, vol. 94, no. 4, pp. 394–408, 2014.

[22] L. Bakiri, R. Hamacher, O. Graña et al., "Liver carcinogenesis by FOS-dependent inflammation and cholesterol dysregulation," *The Journal of Experimental Medicine*, vol. 214, no. 5, pp. 1387–1409, 2017.

[23] S. C. Hasenfuss, L. Bakiri, M. K. Thomsen, E. G. Williams, J. Auwerx, and E. F. Wagner, "Regulation of steatohepatitis

Hindawi
BioMed Research International
Volume 2023, Article ID 9859342, 1 page
https://doi.org/10.1155/2023/9859342

Retraction

Retracted: Clustering of Brain Tumor Based on Analysis of MRI Images Using Robust Principal Component Analysis (ROBPCA) Algorithm

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] A. Hamzenejad, S. J. Ghoushchi, and V. Baradaran, "Clustering of Brain Tumor Based on Analysis of MRI Images Using Robust Principal Component Analysis (ROBPCA) Algorithm," *BioMed Research International*, vol. 2021, Article ID 5516819, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 5516819, 11 pages
https://doi.org/10.1155/2021/5516819

Research Article

Clustering of Brain Tumor Based on Analysis of MRI Images Using Robust Principal Component Analysis (ROBPCA) Algorithm

Ali Hamzenejad,[1] **Saeid Jafarzadeh Ghoushchi**,[2] **and Vahid Baradaran**[1]

[1]*Department of Industrial Engineering, Islamic Azad University, Tehran North Branch, Tehran, Iran*
[2]*Faculty of Industrial Engineering, Urmia University of Technology, Urmia, Iran*

Correspondence should be addressed to Saeid Jafarzadeh Ghoushchi; s.jafarzadeh@uut.ac.ir

Received 8 February 2021; Revised 26 February 2021; Accepted 16 August 2021; Published 31 August 2021

Academic Editor: Min Tang

Automated detection of brain tumor location is essential for both medical and analytical uses. In this paper, we clustered brain MRI images to detect tumor location. To obtain perfect results, we presented an unsupervised robust PCA algorithm to clustered images. The proposed method clusters brain MR image pixels to four leverages. The algorithm is implemented for five brain diseases such as glioma, Huntington, meningioma, Pick, and Alzheimer's. We used ten images of each disease to validate the optimal identification rate. According to the results obtained, 2% of the data in the bad leverage part of the image were determined, which acceptably discerned the tumor. Results show that this method has the potential to detect tumor location for brain disease with high sensitivity. Moreover, results show that the method for the Glioma images has approximately better results than others. However, according to the ROC curve for all selected diseases, the present method can find lesion location.

1. Introduction

Imaging techniques used in medicine allow for a relatively accurate diagnosis of diseases without surgery by direct observation of body tissues. Today, many nanotechnology imaging techniques have been improved. Medical imaging is a technique for obtaining images of body components for medical purposes such as diagnosing or studying diseases. The disease can be treated more successfully; that is, the patient's treatment is faster and better with fewer problems and pain and lower costs. Another goal of imaging is to examine the disease's progression and the effectiveness of the treatment [1, 2].

In most literature algorithms for the clustering and classification of MRI, brain MRI is classified by several methods. In [3], the two-dimensional wavelet transforms (DWT) and the Principal Component Analysis (PCA) were employed for feature extraction. The classification they utilized included normal and abnormal brain MR [4]. Chaplot et al. [5] used an approximate two-dimensional subband DWT in MRI images of the brain as new features in the research, in which Daubechies filters were used as a filtering fraction. They considered Alzheimer's disease abnormal. Their studies found that SVM has better results with radial base function and the polynomial core than linear neural networks and SVM [5]. Hackmack et al. [6] (2012) used multidimensional complex wavelet transformations to use MRI images of the brain and then SVM linearly determined whether the brain is multiscale or healthy. The results show that low-frequency scales contain more information than high-frequency values [6]. El-Dahshan et al. [7] used an approximate two-dimensional DWT subband of brain-magnetic MRI in 2015 to compute the feature vector. In their research, Maitra et al. (2008) employed Slantlet deformation, which utilizes its improved DWT variant to obtain the related properties of the MRI of the brain. The fuzzy C-meaning tool was utilized to evaluate brain MRI to determine a natural individual or Alzheimer's disease depending

on the features of the image histogram [8]. To distinguish the brain's MRI from a T1-T2 weight, Ramathilagam et al. [9] used a c–means updated algorithm. The authors recommended repeating the dist-max algorithm using the dist-max default algorithm before implementing the algorithm since the c-means standard factor is extremely sensitive to the noise-induced region during extraction [9]. Another study by Lehmann et al., 2017, deals with healthy Alzheimer's patients in the mild and moderate stages of the disease. In this study, 39 different EEG signal features were recorded for individuals with rest and closed eyes. Different classifiers were used to compare differentiating methods of Alzheimer's patients from healthy people [10–12].

In a Kim et al. study, Alzheimer's diagnosis is based on the EEG signal of people employing genetic algorithms and neural networks. The ability to distinguish patients with Alzheimer's from the moderate stage of stable people with a reliable 82 percent of the EEG signal was one of the noteworthy points in this research [13]. An entropy estimate measured the relation of EEG signal disorder to stable individuals with brain tumor patients. In this study, it has been shown that in patients with moderate to severe moderate-to-severe brain tumors, the degree of EEG entropy disorder decreases significantly, most of which are observed in P3 and P4 canals. The findings suggest that low-frequency patients have a more robust power range than normal individuals with low EEG signals, such as the delta and theta bands [14, 15].

Metaheuristic optimization algorithms have become increasingly appealing in recent studies. Metaheuristics are used to identify high-quality solutions to a rising variety of complicated real-world problems, such as combinatorial ones because they can address multiple-objective multiple-solution and nonlinear formulations [16–20]. Optimization methods are at the basis of a variety of vital activities, and they can be applied to a variety of image segmentation problems in medicine [21–25]. There are some optimization methods based on metaheuristic algorithms that can be used for feature extraction in image processing [26–31]. According to methods, Nowinski et al. used the GWR method for the brain's detection. The results show that GWR-based analysis is useful for describing the natural brain, determining areas of interest, and determining healthy age [32]. Haegelen et al. image patterns of T1 weight and T2 were investigated for magnetic resonance imaging with a mean MR image of 57 patients with Parkinson's disease. Animals were better registered than SyN on the left thalamus and better than the right-thalamus patch-based approach [33]. In 2016, Zacharaki et al. studied machine learning algorithms that automatically recognize the related characteristics and are desirable for distinguishing a brain tumor [34]. A research was completed by Fritzsche et al. [35] of 15 brain tumor patients, and of 18 MCI patients, ten were stable for three years, and eight were stable for 15 healthy patients for three years. The diagnosis is also increased with the examination limited to the left brain hemisphere (83.3% accuracy, sensitivity 70%, and 100% attribute). Manual procedures and manual volumes were 66.7% (100.4%) and 72.2% (60.60% 87.5%), respectively [35]. Zarei et al. reviewed the structural tensor scans of MRI scans from 16 patients with AD and 22 healthy volunteers [36]. Using morphometric MRI imaging, Devanand et al. [37] analyzed the local adjustments of the hippocampus, Parahipokamp grains, and entorhinal cortex in determining the conversion from the moderate cognitive disability of MCI to brain tumor of AD [38]. In CT images, such as poor borders, touching organs, and heterogeneity of the liver, Ranjbarzadeh and Saadi proposed a method to resolve liver and tumor segmentation problems. For improved edge extraction, they used the Eight Directional Kirsch filter [39]. In other work, they proposed an algorithm based on the study of the influence of speckle noise on the extended histogram in the speckle noise corrupted image [40]. A new robust algorithm for brain disease classification has been introduced by Hamzenejad et al. The empirical results showed that, while requiring a smaller number of classification attributes, the proposed algorithm obtained a high classification rate and better practices than algorithms recently introduced [41]. Moreover, other research related to brain tumor are RPCA [42] and Optimized Quantum Matched-Filter Technique [43]. Also, biological uses of machine learning are prevalent, for instance, diagnosis of gastric cancer tumor [44], foot fatigue [45], lung tumor [46], tuberculous [47], thyroid nodules [48], Parkinson [49], and paraquat-poisoned patients [50].

In this paper, we used the robust PCA algorithm to cluster MRI images pixel automatically. We used five brain diseases to determine tumor location. The results are compared with fuzzy C-means (FCM) method based on performance analysis criteria.

2. Methods and the Material

2.1. Robust Principal Component Analysis (Robust PCA or ROBPCA). One of the essential straight linear algebra results is the analysis of principal components as the nonparametric and straightforward approach is to derive meaningful details from confounding sets. The transformation of the PCA is obtained by decreasing the least-squares error, assuming that the vector of properties is $X \in R^d$ (d-dimensional space), then the space of the reduced features. $Y \in R^h$. If the vector of properties is $X \in R^d$, then, the least-squares error can be in the form of

$$\mathrm{MSE} = E\left[\left\|x - \widehat{x(r)}\right\|^2\right], \tag{1}$$

where x is the input vector of information, and $\widehat{x(r)}$ is the decoding function for transforming r to x. In PCA, we are looking for a conversion that obtains the least-squares error, which means the following conversion is desired.

$$T_D = \arg\min\{MSE\} = \arg\min_{\mathrm{r}}\left\{E\left[\left\|x - \widehat{x(r)}\right\|^2\right]\right\}. \tag{2}$$

It means that we want to find the r vector values so that

the MSE is as minimal as possible. As mentioned, the PCA is aimed at identifying a linear T transformation that generates the least-square error and improves the maximization of this linear $T^T \operatorname{cov}_{x-\hat{x}} T$ transformation in which $\operatorname{cov}_{x-\hat{x}}$ is the data covariance matrix with an average of x-zero. The PCA then resolves the particular problem, and the T-matrix is formed by the vector of unique features of the linear transformation columns. The information is illustrated in the lower dimension $Y = (x - \hat{x})$, namely, Y. According to the above, MSE has at least $d - h$ value; that is, a particular large vector corresponding to multiple eigenvalues should be taken to reduce the dimension or the category objective for the minimum MSE. Here, it is necessary to address some of the covariance matrix properties because the main problem is to find the particular values of the covariance matrix. For example, in three dimensions, the covariance matrix is as follows:

$$C = \begin{pmatrix} \operatorname{cov}(x,x) & \operatorname{cov}(x,y) & \operatorname{cov}(x,z) \\ \operatorname{cov}(y,x) & \operatorname{cov}(y,y) & \operatorname{cov}(y,z) \\ \operatorname{cov}(z,x) & \operatorname{cov}(z,y) & \operatorname{cov}(z,z) \end{pmatrix}. \tag{3}$$

Robust PCA's primary goal is to obtain components that do not affect the outliers, and a powerful covariance estimate matrix has replaced the covariance matrix.

If the primary data of the problem is a matrix $n \times p$ or $X = X_{n,p}$, N is the number of observations, and p specifies the number of variables. The ROBPCA method includes the following three steps:

Step 1. The data in its preprocessing mode is placed in smaller spaces to a maximum of $n - 1$.

Step 2. The initial covariance matrix S_0 is created, and the number of components k maintained in the sequence is used, which is the product of the k-dimensional space in which the data apply. After placing the data in the designated spaces, the dispersion matrix of k is the eigenvalue of $l_1, \cdots, l_k$. Eigenvectors are the main component of a powerful problem. The k-dimensional component is the k-dimensional space decomposition problem that is decomposed from the original p-space, which by putting together all the eigenvectors of the matrix $p \times k$ or $P_{p,k}$. The variable vector p determines the estimate. Also, $\hat{\mu}$ is known as a powerful center. Ranks obtained as $n \times k$ matrices from the data or $T_{n,k}$ are specified as follows:

$$T_{n,k} = \left(X_{n,p} - 1_n \hat{\mu}^T\right) P_{p,k}, \tag{4}$$

where 1_n is unit diagonal matrix. Also, k produces a powerful factor component of the powerful S dispersion matrix as much as $p \times p$ with the k rank as follows:

$$S = P_{p,k} L_{k,k} P_{p,k}^T, \tag{5}$$

where $L_{k,k}$ is a diagonal matrix of diameters of eigenvalues $l_1, \cdots, l_k$. The ROBPCA method, like the CPCA, is perpendicular to the spatial and permutable method as if the transformation or rotational transformation (rotation and reverse) is applied to it. Power centers also move or rotate, while ranks remain constant under these transformations. $A_{p,p}$ is a fully ranked orthogonal matrix $A^T = \mathrm{A}^{-1}$. Then, μ_x and $P_{p,k}$ are the ROBPCA centers and loading matrix orthogonal data, respectively. Consequently, the converted centers and the scanning matrix are, respectively, equal to $XA^T + 1_n v^T = A\mu_x + v$ and AP. Also, ratings after the conversion are fixed as follows:

$$\begin{aligned} T\left(XA^T + 1_n v^T\right) &= \left(XA^T + 1_n v^T - 1_n (A\hat{\mu}_x + v)^T AP \right. \\ &= \left(X - 1_n \hat{\mu}_x^T\right) P = T(X). \end{aligned} \tag{6}$$

In all ROBPCA methods, two goals are considered: (a) find a linear combination of variables, even for outlier data; (b) find the type and number of an outlier. The different types of outliers are shown in Figure 1. In this figure, the mapping of data based on robust PCA is depicted. The surface is the 2D PCA space such that data is distributed on it and is fitted by two principal components. Regarding the position of the data, the data points are belonging to such leverages.

The number of levers is defined as follows;

(1) *Regular Observation or Regular Leverage*. Refers to data in a homogeneous group in the PCA space

(2) *Good Leverage*. Refers to missing or missing data from the original data in the PCA space, such as points 1 and 4

(3) *Orthogonal Layer*. Refers to data perpendicular to the PCA space, which is not detectable in the data itself, such as point 5

(4) *Bad Leverage*. Data is provided perpendicularly to the PCA space or away from the original or regular data, such as points 2 and 3

A diagram or layer map must be drawn up to detect regular observations with three outliers for high dimension data. The horizontal axis of the diagonal detection layer is the distance between the score distance (SD) and the vertical axis of the orthogonal distance (OD) defined as follows:

$$\mathrm{SD}_i = \sqrt{\sum_{j=1}^{k} \frac{t_{ij}^2}{l_j}}. \tag{7}$$

The rank t_{ij} is obtained from the matrix $T_{n,k}$. If $k = 1$, then, SD is equal to the standard score of $t_i / \sqrt{l_1}$. Also, OD is defined as follows:

$$\mathrm{OD}_i = \left|\left| x_i - \hat{\mu} - P_{p,k} t_i' \right|\right| \tag{8}$$

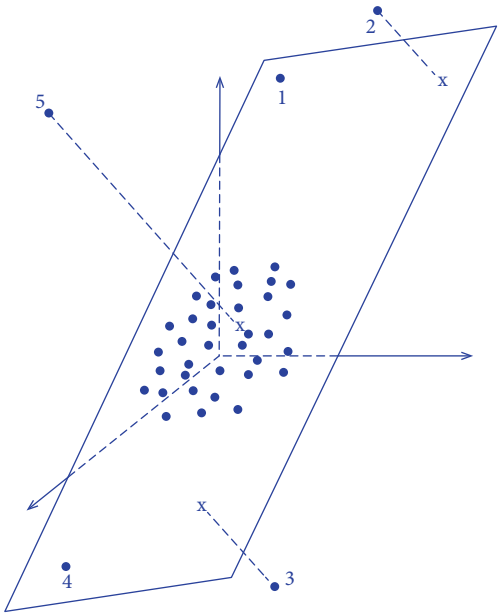


Figure 1: Identify the different types of outliers in ROBPCA.

Input: $y_{m\times m} = \{m \times m\} \in R^2$
Output: *SD*, *OD*
Step 1: Reshape matrix y to vector
$x_i = \{y_1, y_2, \cdots, y_{m^2}\} \in R$
Step 2: Compute Eq. (4) and Eq. (5):
Step 4: Compute Eq. (6):
Step 5: Compute SD:
$SD_i = \sqrt{\sum_{j=1}^{k} (t_{ij}^2 / l_j)}$
Step 6: Compute OD:
$OD_i = \|x_i - \widehat{\mu} - P_{p,k} t_i'\|$
Step 7: Plot diagnosis plot (SD, OD).

Algorithm 1: Algorithm robust PCA.

The i-th observation is given as the vector x_i (p-variable). Also, t_i' is equal to i-th row of the matrix T_n. For the separation of an outlier, the horizontal line $\sqrt{\chi^2_{k,0.75}}$ for $k > 1$, $\pm\sqrt{\chi^2_{1,0.975}}$ for $k = 1$ because the square of the Mahalanobis's decentralization for the data with a normal distribution, it is difficult to determine the vertical cut-off line because there is no information on the distribution of the perpendicular.

3. Results and Discussion

3.1. Dataset. In this article, we used five brain disorders to apply the proposed approach. The dataset includes Alzheimer's, glioma, Huntington, meningioma, and sarcoma. Such images of diseases include MRI photographs from the repository of Harvard Medical School [51]. All images come from $T2$-weighted MR brain images in the axial plane and have 256×256 pixels. Each image is processed separately and analyzed in an unsupervised technique [52].

3.2. Results of Proposed Clustering Method. This paper proposes a robust algorithm to determine the tumor location in a magnetic brain image (MRI). MRI image pixels are categorized based on four good, bad, orthogonal, and regular leverages based on the singular value decomposition (SVD) method. The scores obtained from the leverages in the robust PCA algorithm indicate the tumor location. We used five brain diseases, glioma, meningioma, Pick, Huntington, and Alzheimer, to validate this method. Also, ten images such as sensitivity, accuracy, precision, and fallout are recorded. As explained in the method section, the primary purpose of the ROBPCA is to obtain components that do not affect the outlier, and the covariance matrix is replaced by a powerful covariance estimate matrix, which constructs the input image based on four good, bad, orthogonal, and regular. To use ROBPCA, the input parameters of the problem are as follows:

(i) k. The number of principal components that will be used in this issue. The number of components is selected using the criteria in Hubert and Engelen [53]

(ii) $Kmax$. Maximum number of principal components to be computed, which is considered a default by 10

(iii) α. The robustness parameter is considered to be 0.75 by default

(iv) h. The number of anomalies the algorithm has to withstand. Which is obtained from the $n - h$ equation

$$h = [\alpha * n] + 1, \tag{9}$$

which n represents the sample number. In this case, the value $k = 0$, k max $= 10$, $\alpha = 0.75$, and $h = 65536$. The higher the α, the more accurate the calculations for uncontaminated data would be. On the other hand, setting a lower value for α would increase the algorithm's robustness in abnormal points. After applying the ROBPCA method, the results are as shown in Figure 2.

Figure 2 shows the graph of pixel clustering results using the ROBPCA method; therefore, each point is the single image's pixels. The results of the ROBPCA analysis include the clustering of pixels into four levers. Two vertical and horizontal cut-off lines separate four leverages. The lower left, lower right, upper left, and upper right levers are regular, good, orthogonal, and bad leverages. The output graph shows the orientation distance (OD) and the distance (SD). The cut-off line for the axis of intermediate distance is 3.338, and the cutting-axis line is 3.0172.

According to the graph of the results of Figure 2, most sample data are located in two orthogonal and regular layers. Regular observations or regular leverage were called data placed in a homogeneous group in the PCA space. The vertical or orthogonal layer was said to be orthogonal to the data that is perpendicular to the PCA space, which is not visible to the data itself.

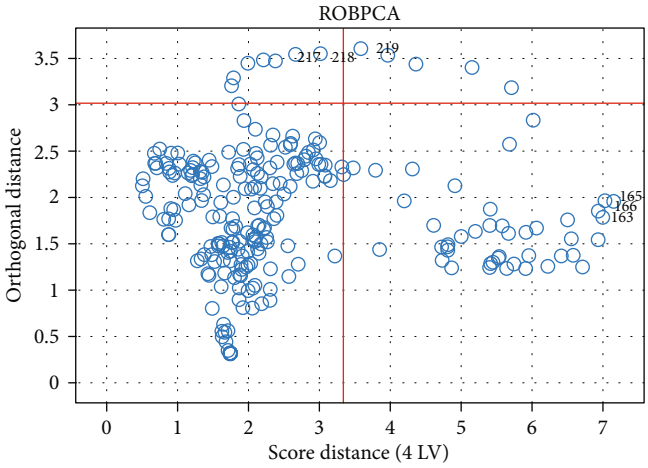


Figure 2: ROBPCA results chart for brain tumor image sample. Upper right side: bad leverage; upper left: orthogonal layer; lower right: good leverage; lower left: regular layer.

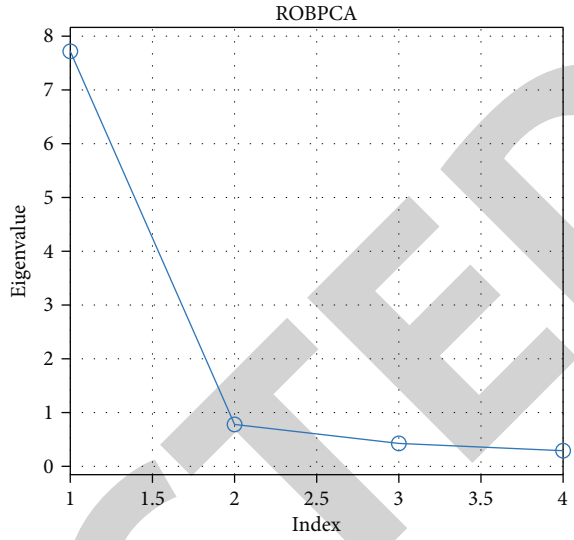


Figure 3: The scree plot of the robust PCA method.

According to the diagram, 72.7% of data is in the regular lever and 25.3% in the good laver or lexical cluster data that are not directly detected before and 2% in the regular lever. According to the scree plot in Figure 3, four are the best classes using brain images. The scree plot shows the eigenvalue of the covariance matrix of robust PCA. A covariance matrix's eigenvectors and eigenvalues form the "core" of a robust PCA. The new function space directions are determined by the eigenvectors (principal components), and the eigenvalues determine their magnitude. In other words, along the new function axis, the eigenvalues describe the variance of the data. Robust PCA's approach is to perform the covariance matrix's decomposition, a matrix where each element defines the covariance between two characteristics.

The classical PCA method is an approach for reducing the input data's dimension or reducing input variables of features. The robust PCA can reduce the number of features. However, robust PCA is different from PCA. One of the important characteristics of the robust PCA is to separate data into four leverages. This property does not exist in classical PCA. In robust PCA, input data is converted to a single vector then single variables are separated into four leverages. These leverage based on unsupervised clustering property separate data with the same behaviors. Therefore, this method does classical PCA work and applies an unsupervised clustering method to the input data. It can be added that classification and segmentation are machine learning and supervised methods. However, our method is an unsupervised clustering approach. The results of this method can be used in machine learning ground-truth images for the output layer.

After separating the leverage, the image matrix's transformation into the initial image is shown in Figures 4–8. Figures 4–8 show the results of leverage separation for meningioma, glioma, Alzheimer's Huntington, and Pick MRI images. Regarding Figures 4–8, the black parts are in a regular lever and the gray interior parts of the red box representing the vertical lever data. According to the results, tumor location is extracted from bad leverage, and the tiny pixels are removed from the selection. The orthogonal leverage is the complement of regular leverage to construct an original image. In the other use of robust PCA, the bad and good leverages are used as outlier or noise data to clear the input image. However, in this paper, we used outliers to detect tumor location.

3.3. Performance Analysis. According to Figure 4, the tumor location of meningioma is approximately well detected. In this case, $\alpha = 0.7$ and colored place from bad leverage have coincided over the original image. In this method, the pixels with the white color using $\alpha = 0.7$ coefficient help to detect optimal location. About glioma and Alzheimer's, this method has the potential for clustering.

To compare with other methods and analyze, we used ground truth images of an automated method of FCM comparison. Therefore, the performance analysis is calculated based on this comparison as follows:

(i) *True Positive (TP)*. The clustered pixels of the tumor are similar to FCM clustered location

(ii) *False Positive (FP)*. Healthy pixels are detected with discrepancy with FCM

(iii) *True Negative (TN)*. Healthy pixels are detected like FCM

(iv) *False Negative (FN)*. The clustered pixels of the tumor are detected with discrepancy with FCM

The sensitivity of devising the percentages of TP pixels into the sum of TP and FN pixels is defined as follows:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{10}$$

Likewise, the performance analysis's specificity results

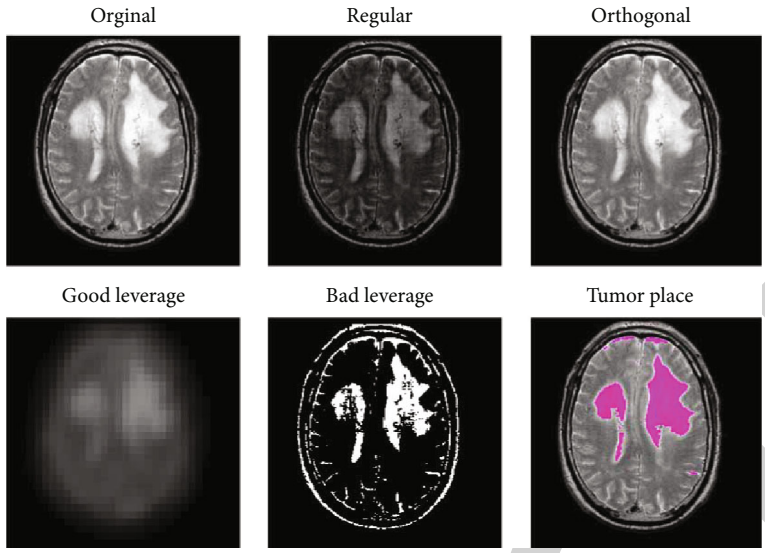


Figure 4: Separation of four levers using the ROBPCA method for meningioma.

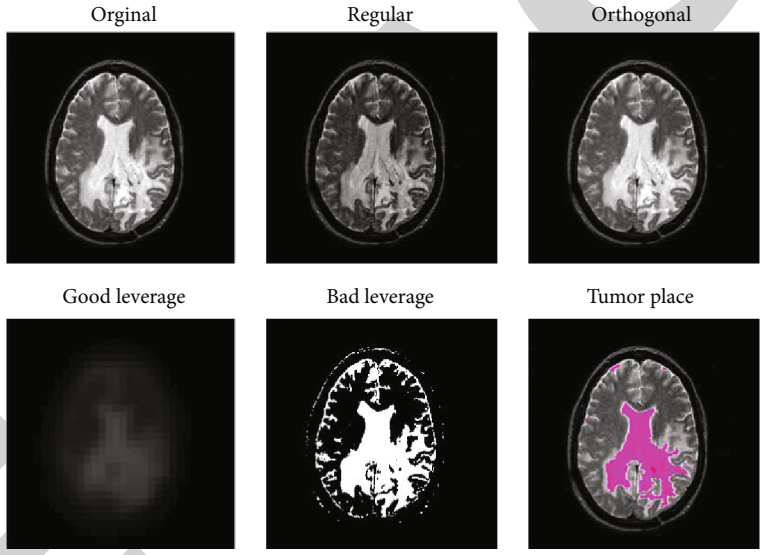


Figure 5: Separation of four levers using the ROBPCA method for glioma.

are the number of false-positive pixels and true negative cases in actual negative cases.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \tag{11}$$

$$\text{Precision, PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{12}$$

$$\text{Accuracy(ACC)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{13}$$

$$\text{Fall} - \text{out(FPR)} = \frac{\text{FP}}{\text{FP} + \text{FN}}. \tag{14}$$

Figure 9 shows the comparison of our presented clustering method with the automated FCM approach. Based on the results of robust PCA, the red pixels show the tumor location. These values are calculated for each disease with ten images. Based on the Pick results, some brain liquid is detected as a tumor that can control with a coefficient and Huntington disease because of nonwhite pixels; despite low $\alpha = 0.6$, some worse points are selected. However, with $\alpha = 0.8$, the selection is correct optimally. The sensitivity, fallout, accuracy, precision, and specificity are explained in Table 1. The results illustrated that high sensitivity is demonstrated for glioma brain image; also, this parameter is low for Huntington. It means that this method for Huntington has not high potential to detect tumor location. However, it is perfect for glioma. The specificity and accuracy for all cases are 0.9 and 0.5, respectively. Also, the precision of the approach is nearly is 0.89. The minimum and maximum fallout are shown for Huntington and glioma, respectively. According to the receptive operating characteristic (ROC) curve, we explain that the technique is plausible for all cases.

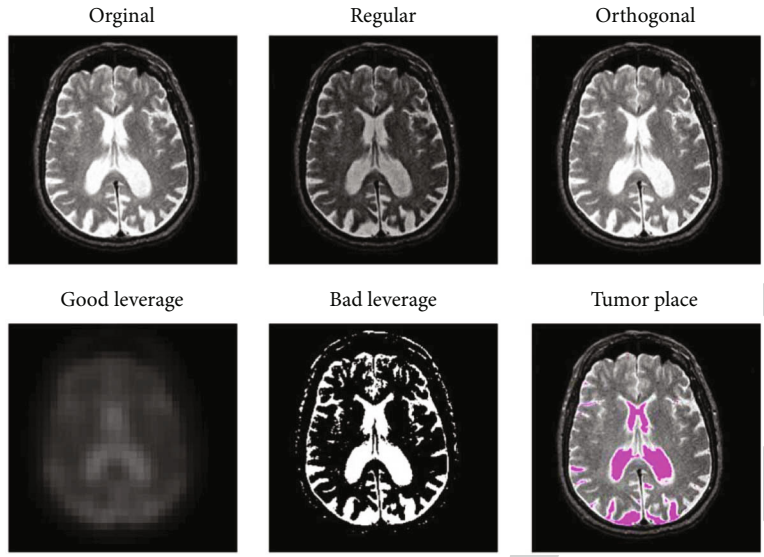


Figure 6: Separation of four levers using the ROBPCA method for Alzheimer.

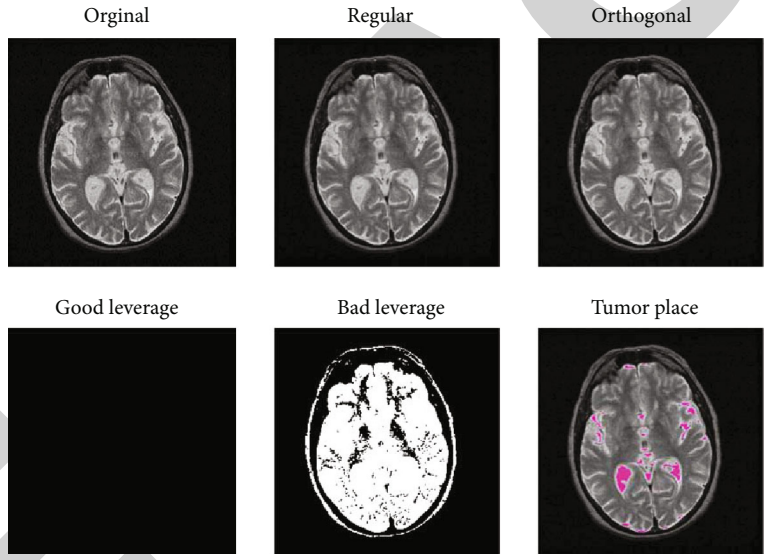


Figure 7: Separation of four levers using the ROBPCA method for Huntington.

Because all cases have sensitivity greater than fallout, all the cases are above the guess line. Regarding Figure 10, perfect result should have minimum fallout and maximum sensitivity. This method is the potential to detect tumor location for glioma better than others.

4. Discussion

Classical PCA is a data-simplification method that uses multivariant databases to lower dimensions for research. It works by calculating the eigenvalues and eigenvalues of a correlation or covariance matrix. In the classical PCA, input data is transformed to a new plane that each coordinates include the greatest variances. Maintaining principal components with the greatest variances and ignoring those with the minor variances contributes to reducing the data dimension while retaining those characteristics of the data set that add the most to its variance.

ROBPCA approach combines projection pursuit with robust scatter matrix estimation and yields more accurate estimates for noncontaminated data and more robust estimates at contaminated data. It is helpful for the analysis of regression data with outliers and multicollinearity. In the case of outliers, the robust PCA produces more concise variants than the nonrobust PCA. Because of the large number of variables in these models, stable PCA solutions to large-scale cointegration models in undersized samples with outliers are of particular interest. In the ROBPCA, the original image is assumed to be contained in an $M \times N$ data matrix $I = I_{M \times N}$, where M denotes the width of the image and N is the height. After that, the ROBPCA process is broken down into three main stages. First, the data are preprocessed such that the transformed data are contained inside a

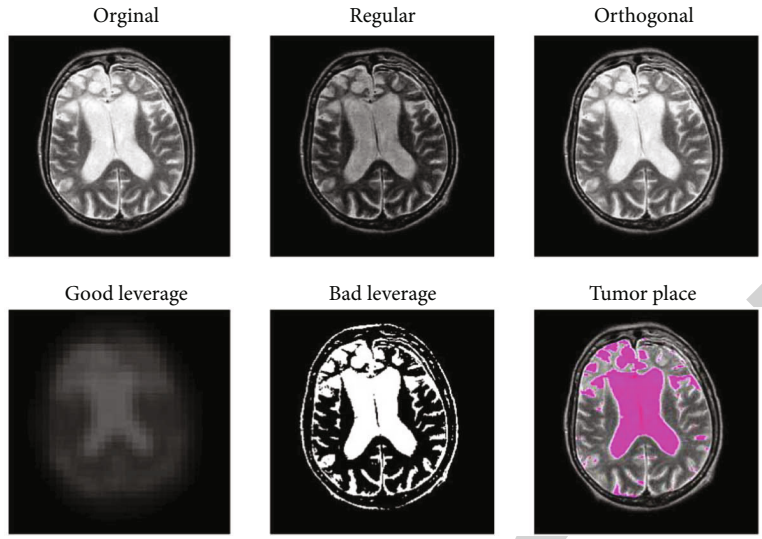


Figure 8: Separation of four levers using the ROBPCA method for Pick.

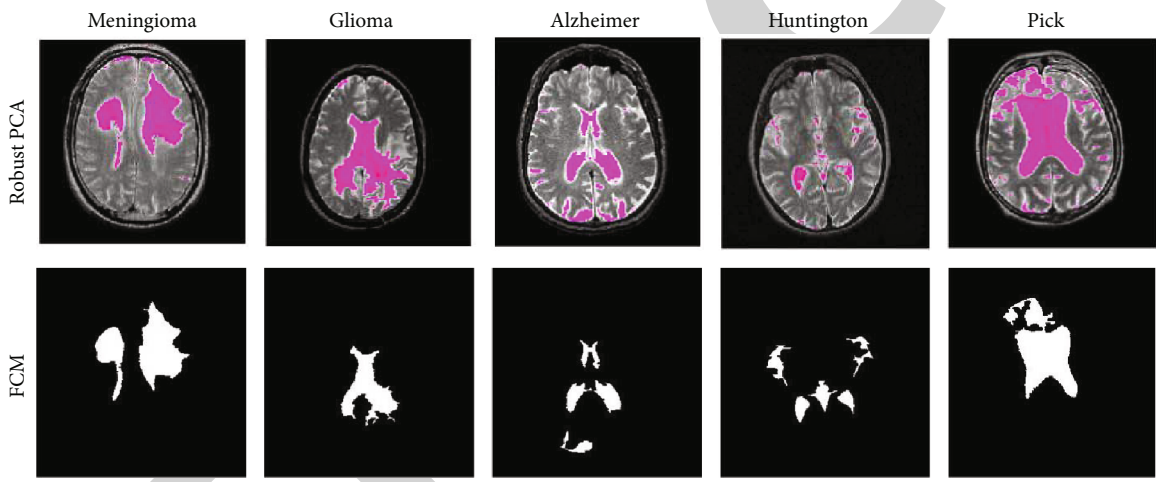


Figure 9: The comparison of the presented method with the FCM approach.

Table 1: The results of detection parameters for $\alpha = 0.8$.

	Sensitivity	Specificity	Precision	Accuracy	Fall-out
Meningioma	0.81	0.67	0.89	0.73	0.33
Glioma	0.90	0.5	0.9	0.70	0.5
Alzheimer	0.83	0.67	0.89	0.73	0.33
Huntington	0.42	0.86	0.8	0.63	0.14
Pick	0.50	0.83	0.83	0.66	0.17

subspace of at most n dimensions. Following that, a provisional covariance matrix is created and used to choose the number of components k that will be preserved in the sequel, resulting in a k-dimensional subspace that matches the data well. The data points are then projected onto this subspace, where their position and scatter matrix are robustly determined, and the k nonzero eigenvalues are calculated.

The ROBPCA has two aspects: (1) to find certain linear combinations of the original variables that contain the majority of the details, even though there are outliers, and (2) to identify and classify outliers. Therefore, in this paper, we used the second properties of ROBPCA to find tumor location. In the classical PCA, this property does not exist and is used for dimension reduction; however, the ROBPCA clusters data to the leverages. Based on the nature of MRI images, three standard colors are presented, black spots are dominated by empty space, white color illustrates the tumor location, and gray points contain other parts of MRI image

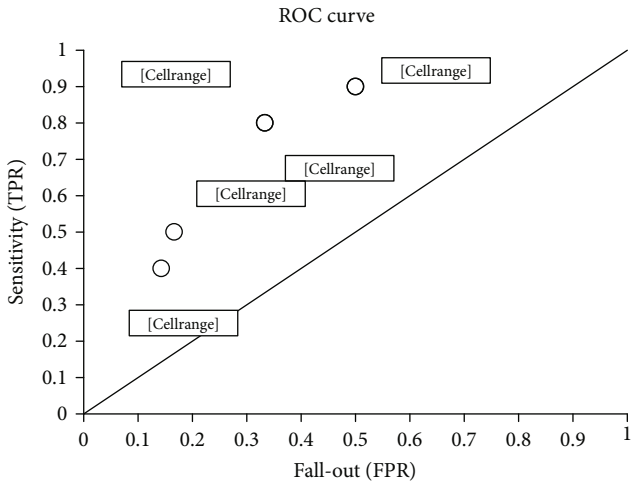


Figure 10: Receiver operating characteristic (ROC) curve.

such as gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). Regarding the primary duty of the ROBPCA, the outliers are clustered to three good, bad, and orthogonal leverage, and nonoutlier data is located in regular leverage. This leverage segments the tumor location. The main goal of this paper is to use the clustering properties of ROBPCA to segment MRI images and separate the tumor location as types of outliers. There is no proper performance analysis method for clustering methods because of the unsupervised nature of the method. However, to compare and validate data, we used the results of the FCM method with classification performance analysis. The black points are considered as negative indexes and white pixels (bad leverage) as the positive index. We used 50 images for evaluating the performance of the automated clustering methods. The final performance results illustrate the effectiveness of the presented methods.

5. Conclusion

In this paper, a robust algorithm for the determination of tumor location is presented. We use the robust PCA algorithm to cluster image pixels. Robust PCA's results are the clustering of MRI images to four leverages, consisting of regular, orthogonal, good, and bad leverage. In this paper, orthogonal leverage approximately can estimate tumor location. This paper used glioma, meningioma, Pick, Huntington, and Alzheimer's brain diseases to determine the tumor's optimal location. We used ten images of each disease to validate the optimal identification rate. According to the results obtained, 2% of the data in the bad leverage part of the image are determined, which acceptably discerned the tumor. Also, 25.3 percent of the data are located in the orthogonal lever, showing the brain's central and healthy parts. Furthermore, 72.7 percent of the image's data are in the black part that shows other parts of the images. Results show that this method has the potential to detect tumor location for brain disease with high sensitivity. Moreover, results show that the method for the glioma brain images has approximately better results than others. However, according to the ROC curve for all selected diseases, the present method is acceptable.

Data Availability

The dataset is available online: http://www.med.harvard.edu/AANLIB/.

Conflicts of Interest

The authors declare that they have no financial or other conflicts of interest in this research and publication.

References

[1] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982.

[2] N. Abdullah, U. K. Ngah, and S. A. Aziz, "Image classification of brain MRI using support vector machine," in *2011 IEEE International Conference on Imaging Systems and Techniques*, pp. 242–247, Batu Ferringhi, Malaysia, 2011.

[3] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[4] Y. Zhang, Z. Dong, L. Wu, and S. Wang, "A hybrid method for MRI brain image classification," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10049–10053, 2011.

[5] S. Chaplot, L. Patnaik, and N. Jagannathan, "Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network," *Biomedical Signal Processing and Control*, vol. 1, no. 1, pp. 86–92, 2006.

[6] K. Hackmack, F. Paul, M. Weygandt, C. Allefeld, J.-D. Haynes, and Alzheimer's Disease Neuroimaging Initiative, "Multi-scale classification of disease using structural MRI and wavelet transform," *NeuroImage*, vol. 62, no. 1, pp. 48–58, 2012.

[7] E. S. El-Dahshan, T. Hosny, and A. B. Salem, "Hybrid intelligent techniques for MRI brain images classification," *Digital Signal Processing*, vol. 20, pp. 433–441, 2015.

[8] M. Maitra and A. Chatterjee, "Hybrid multiresolution Slantlet transform and fuzzy _c_ -means clustering approach for normal-pathological brain MR image segregation," *Medical Engineering & Physics*, vol. 30, no. 5, pp. 615–623, 2008.

[9] S. Ramathilagam, R. Pandiyarajan, A. Sathya, R. Devi, and S. R. Kannan, "Modified fuzzy-means algorithm for segmentation of T1-T2-weighted brain MRI," *Journal of Computational and Applied Mathematics*, vol. 235, pp. 1578–1586, 2016.

[10] C. Lehmann, T. Koenig, V. Jelic et al., "Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG)," *Journal of Neuroscience Methods*, vol. 161, no. 2, pp. 342–350, 2017.

[11] Z. Hidasi, B. Czigler, P. Salacz, É. Csibri, and M. Molnár, "Changes of EEG spectra and coherence following performance in a cognitive task in Alzheimer's disease," *International Journal of Psychophysiology*, vol. 65, no. 3, pp. 252–260, 2017.

[12] M. J. Hogan, G. R. Swanwick, J. Kaiser, M. Rowan, and B. Lawlor, "Memory-related EEG power and coherence reductions in mild Alzheimer's disease," *International Journal of Psychophysiology*, vol. 49, no. 2, pp. 147–163, 2013.

[13] H. T. Kim, B. Y. Kim, E. H. Park et al., "Computerized recognition of Alzheimer disease-EEG using genetic algorithms and

neural network," *Future Generation Computer Systems*, vol. 21, no. 7, pp. 1124–1130, 2015.

[14] D. Abásolo, R. Hornero, P. Espino, J. Poza, C. I. Sánchez, and R. de la Rosa, "Analysis of regularity in the EEG background activity of Alzheimer's disease patients with approximate entropy," *Clinical Neurophysiology*, vol. 116, no. 8, pp. 1826–1834, 2015.

[15] A. Gholipour, J. A. Estroff, C. E. Barnewolt, S. A. Connolly, and S. K. Warfield, "Fetal brain volumetry through MRI volumetric reconstruction and segmentation," *International Journal of Computer Assisted Radiology and Surgery*, vol. 6, no. 3, pp. 329–339, 2016.

[16] H. Chen, A. A. Heidari, H. Chen, M. Wang, Z. Pan, and A. H. Gandomi, "Multi-population differential evolution-assisted Harris hawks optimization: framework and case studies," *Future Generation Computer Systems*, vol. 111, pp. 175–198, 2020.

[17] M. Wang and H. Chen, "Chaotic multi-swarm whale optimizer boosted support vector machine for medical diagnosis," *Applied Soft Computing*, vol. 88, article 105946, 2020.

[18] Y. Xu, H. Chen, J. Luo, Q. Zhang, S. Jiao, and X. Zhang, "Enhanced moth-flame optimizer with mutation strategy for global optimization," *Information Sciences*, vol. 492, pp. 181–203, 2019.

[19] M. Wang, H. Chen, B. Yang et al., "Toward an optimal kernel extreme learning machine using a chaotic moth-flame optimization strategy with applications in medical diagnoses," *Neurocomputing*, vol. 267, pp. 69–84, 2017.

[20] X. Zhao, X. Zhang, Z. Cai et al., "Chaos enhanced grey wolf optimization wrapped ELM for diagnosis of paraquat- poisoned patients," *Computational Biology and Chemistry*, vol. 78, pp. 481–490, 2019.

[21] H.-L. Chen, G. Wang, C. Ma, Z.-N. Cai, W.-B. Liu, and S.-J. Wang, "An efficient hybrid kernel extreme learning machine approach for early diagnosis of Parkinson's disease," *Neurocomputing*, vol. 184, pp. 131–144, 2016.

[22] Y. Zhang, R. Liu, X. Wang, H. Chen, and C. Li, "Boosted binary Harris hawks optimizer and feature selection," *Engineering with Computers*, 2020.

[23] D. Zhao, L. Liu, F. Yu et al., "Chaotic random spare ant colony optimization for multi-threshold image segmentation of 2D Kapur entropy," *Knowledge-Based Systems*, vol. 216, article 106510, 2021.

[24] Y. Zhang, R. Liu, A. A. Heidari et al., "Towards augmented kernel extreme learning models for bankruptcy prediction: algorithmic behavior and comprehensive analysis," *Neurocomputing*, vol. 430, pp. 185–212, 2021.

[25] X. Zhao, D. Li, B. Yang, C. Ma, Y. Zhu, and H. Chen, "Feature selection based on improved ant colony optimization for online detection of foreign fiber in cotton," *Applied Soft Computing*, vol. 24, pp. 585–596, 2014.

[26] J. Tu, H. Chen, J. Liu et al., "Evolutionary biogeography-based whale optimization methods with communication structure: towards measuring the balance," *Knowledge-Based Systems*, vol. 212, article 106642, 2021.

[27] C. Yu, M. Chen, K. Cheng et al., "SGOA: annealing-behaved grasshopper optimizer for global tasks," *Engineering with Computers*, 2021.

[28] C. Li, L. Hou, B. Sharma et al., "Developing a new intelligent system for the diagnosis of tuberculous pleural effusion," *Computer Methods & Programs in Biomedicine*, vol. 153, pp. 211–225, 2018.

[29] J. Xia, H. Chen, Q. Li et al., "Ultrasound-based differentiation of malignant and benign thyroid nodules: an extreme learning machine approach," *Computer Methods and Programs in Biomedicine*, vol. 147, pp. 37–49, 2017.

[30] L. Shen, H. Chen, Z. Yu et al., "Evolving support vector machines using fruit fly optimization for medical data classification," *Knowledge-Based Systems*, vol. 96, pp. 61–75, 2016.

[31] L. Hu, G. Hong, J. Ma, X. Wang, and H. Chen, "An efficient machine learning approach for diagnosis of paraquat-poisoned patients," *Computers in Biology and Medicine*, vol. 59, pp. 116–124, 2015.

[32] W. L. Nowinski, V. Gupta, W. Y. Chan, Y. Y. Sitoh, and K. Sim, "Use of normative distribution of gray to white matter ratio in orthogonal planes in human brain studies and computer-assisted neuroradiology," *International Journal of Computer Assisted Radiology and Surgery*, vol. 6, no. 4, pp. 489–505, 2016.

[33] C. Haegelen, P. Coupé, V. Fonov et al., "Automated segmentation of basal ganglia and deep brain structures in MRI of Parkinson's disease," *International Journal of Computer Assisted Radiology and Surgery*, vol. 8, no. 1, pp. 99–110, 2013.

[34] E. I. Zacharaki, V. G. Kanas, and C. Davatzikos, "Investigating machine learning techniques for MRI-based classification of brain neoplasms," *International Journal of Computer Assisted Radiology and Surgery*, vol. 6, no. 6, pp. 821–828, 2016.

[35] K. H. Fritzsche, B. Stieltjes, S. Schlindwein, T. van Bruggen, M. Essig, and H. P. Meinzer, "Automated MR morphometry to predict Alzheimer's disease in mild cognitive impairment," *International Journal of Computer Assisted Radiology and Surgery*, vol. 5, no. 6, pp. 623–632, 2018.

[36] M. Zarei, B. Patenaude, J. Damoiseaux et al., "Combining shape and connectivity analysis: an MRI study of thalamic degeneration in Alzheimer's disease," *NeuroImage*, vol. 49, no. 1, pp. 1–8, 2015.

[37] D. P. Devanand, R. Bansal, J. Liu, X. Hao, G. Pradhaban, and B. S. Peterson, "MRI hippocampal and entorhinal cortex mapping in predicting conversion to Alzheimer's disease," *NeuroImage*, vol. 60, no. 3, pp. 1622–1629, 2012.

[38] E. Westman, A. Simmons, Y. Zhang et al., "Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls," *NeuroImage*, vol. 54, no. 2, pp. 1178–1187, 2011.

[39] R. Ranjbarzadeh and S. B. Saadi, "Automated liver and tumor segmentation based on concave and convex points using fuzzy _c_ -means and mean shift clustering," *Measurement*, vol. 150, p. 107086, 2020.

[40] R. Ranjbarzadeh, S. B. Saadi, and A. Amirabadi, "LNPSS: SAR image despeckling based on local and non-local features using patch shape selection and edges linking," *Measurement*, vol. 164, p. 107989, 2020.

[41] A. Hamzenejad, S. Jafarzadeh Ghoushchi, V. Baradaran, and A. Mardani, "A robust algorithm for classification and diagnosis of brain disease using local linear approximation and generalized autoregressive conditional heteroscedasticity model," *Mathematics*, vol. 8, no. 8, p. 1268, 2020.

[42] M. Ahmadi, A. Sharifi, M. Jafarian Fard, and N. Soleimani, "Detection of brain lesion location in MRI images using convolutional neural network and robust PCA," *International Journal of Neuroscience*, pp. 1–12, 2021.

[43] M. Ahmadi, A. Sharifi, S. Hassantabar, and S. Enayati, "QAIS-DSNN: tumor area segmentation of MRI image with

Hindawi
BioMed Research International
Volume 2023, Article ID 9806215, 1 page
https://doi.org/10.1155/2023/9806215

Retraction

Retracted: NREM Sleep EEG Characteristics Correlate to the Mild Cognitive Impairment in Patients with Parkinsonism

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] C. Zhang, L. Wei, F. Zeng et al., "NREM Sleep EEG Characteristics Correlate to the Mild Cognitive Impairment in Patients with Parkinsonism," *BioMed Research International*, vol. 2021, Article ID 5561974, 10 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 5561974, 10 pages
https://doi.org/10.1155/2021/5561974

Research Article

NREM Sleep EEG Characteristics Correlate to the Mild Cognitive Impairment in Patients with Parkinsonism

Cheng Zhang[1], **Luhua Wei**[2], **Fengqingyang Zeng**[3], **Tingwei Zhang**[3], **Yunchuang Sun**[2], **Yane Shen**[1], **Guangfa Wang**[1], **Jing Ma**[1] and **Jue Zhang**[3]

[1]*Department of Respiratory and Critical Care Medicine, Peking University First Hospital, Beijing 100034, China*
[2]*Department of Neurology, Peking University First Hospital, Beijing 100034, China*
[3]*Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China*

Correspondence should be addressed to Jing Ma; majjmail@163.com

Cheng Zhang and Luhua Wei contributed equally to this work.

Received 7 January 2021; Revised 16 April 2021; Accepted 3 July 2021; Published 26 July 2021

Academic Editor: Min Tang

Early identification and diagnosis of mild cognitive impairment (MCI) in patients with parkinsonism (PDS) are critical. The aim of this study was to identify biomarkers of MCI in PDS using conventional electroencephalogram (EEG) power spectral analysis and detrended fluctuation analysis (DFA). In this retrospective study, patients with PDS who underwent an overnight polysomnography (PSG) study in our hospital from 2019 to 2020 were enrolled. Patients with PDS assessed by clinical examination and questionnaires were divided into two groups: the PDS with normal cognitive function (PDS-NC) group and the PDS with MCI (PDS-MCI) group. Sleep EEG signals were extracted and purified from the PSG and subjected to a conventional power spectral analysis, as well as detrended fluctuation analysis (DFA) during wakefulness, nonrapid eye movement (NREM) sleep, and rapid eye movement (REM) sleep. Forty patients with PDS were enrolled, including 25 with PDS-NC and 15 with PDS-MCI. Results revealed that compared with PDS-NC patients, patients with PDS-MCI had a reduced fast ratio ((alpha + beta)/(delta + theta)) and increased DFA during NREM sleep. DFA during NREM was diagnostic of PDS-MCI, with an area under the receiver operating characteristic curve of 0.753 (95% CI: 0.592–0.914) ($p < 0.05$). Mild cognitive dysfunction was positively correlated with NREM-DFA ($r = 0.426$, $p = 0.007$) and negatively correlated with an NREM-fast ratio ($r = -0.524$, $p = 0.001$). This suggested that altered EEG activity during NREM sleep is associated with MCI in patients with PDS. NREM sleep EEG characteristics of the power spectral analysis and DFA correlate to MCI. Slowing of EEG activity during NREM sleep may reflect contribution to the decline in NREM physiological function and is therefore a marker in patients with PDS-MCI.

1. Introduction

Parkinsonism (PDS) is a syndrome with various causes characterized by bradykinesia, rigidity, and rest tremor [1], usually seen in neurodegenerative disorders, such as Parkinson's disease (PD), multiple system atrophy, and Lewy body dementia. Apart from the motor symptoms, patients with PDS also present with several nonmotor symptoms, which at times can be more debilitating than motor dysfunction [2]. Cognitive impairment is a common and troublesome nonmotor complication frequently encountered by clinical practitioners. Approximately 30% of patients with Parkinson's disease experience mild cognitive impairment (MCI) [3].

The cumulative prevalence of PD dementia is approximately 75–90% in patients with a disease course exceeding 10 years. This increases the mortality rate and severely impacts the quality of life [4]. Similar conditions are seen in atypical parkinsonian syndromes [5–7]. Importantly, cognitive impairment without dementia can persist for years

before the onset of dementia in parkinsonian patients [8], which provides a potential monitoring and therapeutic window. Therefore, the early detection of cognitive deficits is important.

Currently, in clinical practice, cognitive function is usually assessed using questionnaires or scales. These are well-established methods and easy to use; however, the specificity and sensitivity of many scales remain unsatisfactory [9] due to their subjective nature as well as the complexity of the cognitive profile in PDS. This is worse when it comes to MCI.

The quantitative analysis of electroencephalogram (EEG) signals provides a method for assessing cognitive function in various diseases. For example, numerous studies focusing on EEG have demonstrated correlations between quantitative EEG (qEEG) biomarkers and cognitive function in Alzheimer's disease [10]. Some studies have also shown that qEEG is a very promising predictor of PD-related cognitive decline [11].

However, previous studies have mainly focused on the association between resting awake EEG alterations and cognitive dysfunction in patients with PDS and rarely addressed sleep EEG. However, sleep EEG also provides a window of insight into cognitive dysfunction [12].

For example, previous studies have shown that alterations in sleep EEG are associated with brain function in the aging population [12]. Decreased slow-wave, theta, and sigma activities during nonrapid eye movement (NREM) sleep and changes in spindle characteristics predicted the very early onset of cognitive impairment [12]. It has also been shown that the power of alpha activity during NREM sleep is associated with intelligence in the general population [13].

It remains unclear, however, whether sleep EEG is associated with mild cognitive impairment in PDS patients.

Nowadays, power spectral analysis is the most common method used in these EEG studies.

In addition, due to the nonlinear dynamics and nonstationarity of EEG signals limiting the conventional power spectral analysis, a simpler alternative method, the detrended fluctuation analysis (DFA) method, has been used for various nonstationary physiological signals in recent years [14]. This is a method for calculating the long-term correlation of nonstationary time series proposed by Peng et al. in 1994 [15] in their study of DNA sequences. The advantage of the DFA method over many other analysis methods is that it can detect long-range correlations implicitly in unstable time series and can effectively minimize interference due to signal instability. Recent studies have applied DFA in the field of EEG and cognition, and it can be used as an alternative to conventional EEG spectral analysis as a marker of cognitive impairment in patients with obstructive sleep apnea (OSA) or Alzheimer's disease [14, 16].

Therefore, this study sought to determine sleep EEG biomarkers of polysomnography (PSG) related to MCI in PDS patients. We hypothesized that the power spectral analysis features of overnight sleep EEG are related to cognitive function in patients with PDS.

2. Materials and Methods

Patients with PDS who underwent overnight PSG monitoring were enrolled in this study, and their overnight EEG was analyzed using conventional spectral analysis and DFA. Differences between patients with normal cognitive function and those with MCI were compared to identify EEG markers reflecting cognitive impairment.

2.1. Study Population. A total of 45 consecutive patients who were admitted to our neurology department with PDS from August 2019 to August 2020 and underwent overnight PSG in our sleep lab were enrolled in this study.

This retrospective study was approved by the ethics committee of Peking University First Hospital, and the requirement for obtaining informed consent was waived (ethics approval no. 2019[181]). The study adhered to the Declaration of Helsinki, and patient confidentiality was maintained.

Inclusion criteria are as follows:

(1) Compliance with the criteria for PDS includes bradykinesia, in combination with either resting tremor, rigidity, or both [1]

(2) PDS was not secondary to identified structural, toxic, or metabolic disorders or conditions

Exclusion criteria are as follows:

(1) Poor quality of PSG-EEG signals, high impedance, excessive electromyography (EMG), or presence of motor artifacts

(2) Dementia cases where substantial impairment is present in one or more cognitive domains and the impairment must be sufficient to interfere with independence in everyday activities [1]

(3) Other medical conditions that cause EEG changes, such as epilepsy

(4) Other medical conditions that may affect cognition, such as stroke

(5) Extremely severe motor symptoms affecting patients' cooperation with cognitive evaluations

Among the 45 patients with PDS, we excluded two patients with high signal impedance, excessive myoelectricity, or motion artifacts; moreover, three cases of PDS with dementia were excluded. A total of 40 patients were finally enrolled.

2.2. Cognitive Function Evaluation. All patients were interviewed, and a detailed medical and medication history was acquired. The cognitive function was evaluated based on clinical examination, the Montreal Cognitive Assessment (MoCA) and Mini-Mental State Examination (MMSE). All cognitive evaluations were performed when PD patients were in the "on" state and when the other parkinsonian patients were in relatively better conditions. Subjects were stratified into two groups based on cognitive evaluations: PDS with

MCI (PDS-MCI) group and PDS with normal cognition (PDS-NC) group. There are no diagnostic criteria for PDS-MCI. We defined the PDS-MCI group based on the criteria of mild neurocognitive disorders [17] and MCI in PD [17]: parkinsonism with modest impairment in one or more cognitive domains that are not sufficient to interfere significantly with functional independence, although subtle difficulties in complex functional tasks may be present, and cognitive deficits on either formal neuropsychological testing or a scale of global cognitive abilities. More specifically, the cutoff value of MoCA for PDS-MCI is 21–25 points. The MMSE was used to exclude dementia with a cutoff value of 23/24 points. The PDS-NC group was defined as having parkinsonism without dementia or MCI, as described above.

2.3. Overnight PSG. Sleep apneas were confirmed by an overnight PSG (Compumedics, E-Series, Australia). Six EEG signals (C3-M2, C4-M1, F3-M2, F4-M1, O1-M2, and O2-M1), two channels of electrooculography (EOG) signals (E1-M2, E2-M2), chin EMG (EMG1-EMG2, EMG1-EMG3), electrocardiography (ECG), respiration (nasal pressure, airflow), oxygen saturation (SpO_2), abdominal and chest movement, and leg movements were recorded according to the American Academy of Sleep Medicine (AASM) guidelines [13, 14].

Sleep stage and respiratory events were analyzed according to the AASM 2.3 guidelines [13, 14]. Sleep stages were divided into the NREM (N1, N2, and N3) sleep, rapid eye movement (REM) sleep, and wake stages.

Sleep-related parameters (total sleep time (TST), sleep efficiency (SE), sleep latency (SL), wake time after sleep onset (WASO), and the proportion of each sleep period to total sleep (N1/TST, N2/TST, N3/TST)) were calculated and reported in the PSG study report.

2.4. Spectral Analysis and DFA of PSG-EEG Signals

2.4.1. EEG Spectral Analysis. We used noninvasive electrodes to record scalp EEG signals from sleeping participants, as described in the overnight PSG method. All EEG signals were recorded at a sampling rate greater than 200 Hz and stored in the EDF format. A band-pass filter of 0.3–35.0 Hz was applied to the original signals. The signals of two channels (C4-M1, O1-M2) were analyzed. MATLAB software (MATLAB 2018a) was used for EEG preprocessing, artifact removal, and data analysis. We first estimated the power spectrum in each frequency band (delta, 0.5–4.0 Hz; theta, 4.0–8.0 Hz; alpha, 8.0–13.0 Hz; sigma, 12.0–16.0 Hz; and beta, 13.0–30.0 Hz) for each 30 s segment using the Welch method [18]. This is a modification of the periodogram method that produces more reliable power spectrum estimates by applying windows and overlaps [19]. We used the Hamming window function with a window width of two full cycles of the lowest interest frequency (0.5 Hz) to avoid underestimation, and the overlap was set to half the window width. Next, the absolute power of the segmented signal bands was obtained by numerical integration of the estimated power spectrum by the Simpson method. The relative power and fast ratio were also calculated and defined as follows: (alpha + beta)/(delta + theta).

Table 1: Demographic characteristics and sleep parameters monitored by PSG.

	PDS-NC ($n = 25$)	PDS-MCI ($n = 15$)
Age (years)	65.80 ± 7.89	68.53 ± 11.11
BMI (kg/m^2)	23.59 (21.79, 26.68)	25.51 (20.70, 27.68)
AHI (/h)	11.90 (2.30, 17.80)	4.40 (2.80, 15.30)
TST (min)	279.70 ± 94.97	234.50 ± 91.25
SE (%)	58.70 (48.15, 76.10)	59.00 (39.30, 68.80)
Sleep latency (min)	22.50 (9.38, 48.13)	36.00 (13.00, 52.63)
WASO (min)	137.38 ± 76.52	158.97 ± 83.03
R/TST (%)	16.93 ± 7.68	15.77 ± 8.08
N1/TST (%)	13.10 ± 6.73	17.25 ± 10.73
N2/TST (%)	58.47 ± 10.87	53.01 ± 16.24
N3/TST (%)	10.20 (5.85, 15.70)	8.00 (3.90, 21.90)

$p > 0.05$ PDS-NC vs. PDS-MCI. Data representing age, TST, WASO, R/TST, N1/TST, and N2/TST are expressed as mean ± standard deviation, and t -tests were used for between-group comparison. Data representing AHI, BMI, SE, sleep latency, and N3/TST are expressed as the median (25th-75th percentile), and nonparametric tests were used for between-group comparison. Abbreviations: BMI: body mass index; AHI: apnea and hypopnea index; TST: total sleep time; SE: sleep efficacy; WASO: wake time after sleep onset; R: rapid eye movement sleep; N1: nonrapid eye movement 1 sleep; N1: nonrapid eye movement 2 sleep; N3: nonrapid eye movement 3 sleep.

Table 2: The EEG power spectral analysis and DFA in PDS-NC and PDS-MCI patients.

	PDS-NC	PDS-MCI	P
alpha-NREM (C4)	0.08±0.032	0.05±0.015	0.032
alpha-NREM (O1)	0.08±0.034	0.05±0.023	0.022
alpha-wake (C4)	0.15±0.068	0.10±0.061	0.037
alpha-wake (O1)	0.17±0.094	0.11±0.041	0.016
sigma-wake (C4)	0.04±0.015	0.03±0.015	0.045
sigma-wake (O1)	0.04±0.017	0.03±0.012	0.038
beta-wake (C4)	0.13±0.050	0.08±0.031	0.005
beta-wake (O1)	0.12±0.041	0.08±0.032	0.003
fast ratio-NREM (C4)	1.12±0.60	0.70±0.32	0.022
fast ratio-NREM (O1)	1.24±0.55	0.65±0.24	0.001
fast ratio-wake (C4)	1.27±0.702	0.82±0.435	0.031
fast ratio-wake (O1)	1.40±0.64	0.73±0.28	0.001
DFA-NREM (C4)	0.76±0.12	0.87±0.14	0.011
DFA-NREM (O1)	0.80±0.12	0.90±0.12	0.005
DFA-wake (C4)	0.76±0.13	0.87±0.16	0.026
DFA-wake (O1)	0.80±0.12	0.90±0.11	0.012

Abbreviations: PDS-NC: parkinsonism with normal cognitive function; PDS-MCI: parkinsonism with mild cognitive impairment; NREM: nonrapid eye movement; DFA: detrended fluctuation analysis. Fast ratio: (alpha + beta)/(delta + theta).

2.4.2. DFA. DFA is a method for calculating the long-term correlation of nonstationary time series proposed by Peng et al. in 1994 [15]. The obtained scaling coefficient α describes the relationship between the future and historical trends of the signal; for example, white noise $\alpha = 0.5$ and

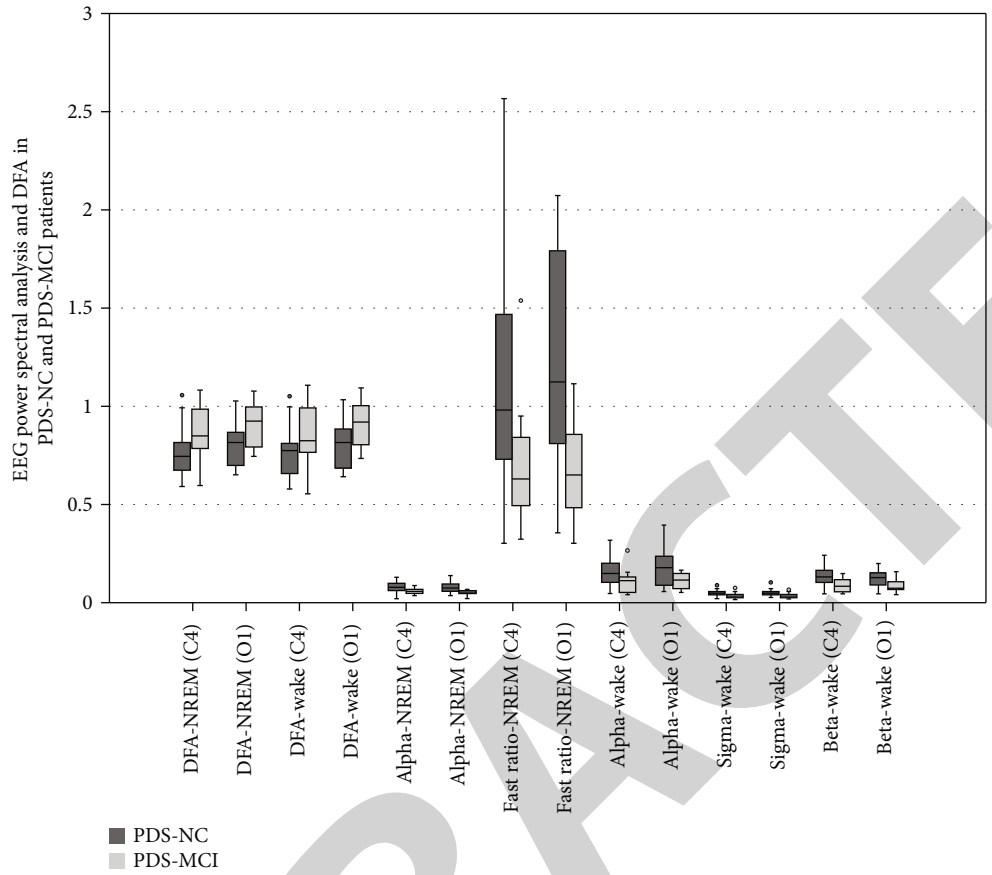


Figure 1: The EEG power spectral analysis and DFA in PDS-NC and PDS-MCI patients (a summary of Table 2, $p < 0.05$).

fractal Brownian noise $\alpha = 1.5$. DFA is suitable for EEG data analysis because of its nonstationarity. Recent studies have demonstrated its advantages in sleep EEG analysis, such as sleep staging [20] and sleep apnea detection [21]. In this study, we performed DFA of the segmented sleep EEG following preprocessing. First, we calculated the scaling coefficient α for all segments of the wake, NREM sleep, and REM sleep stages. We then averaged them as typical values of the corresponding stages for subsequent statistical analysis. The coefficient was obtained by a first-order fit to 50 logarithmically spaced scales of each EEG segment.

2.5. Statistical Methods. The normality of distributions was assessed by the Shapiro-Wilk test.

Data are expressed as mean ± standard deviation, and t -tests were used if the variables conformed to a normal distribution (alpha, beta, theta, sigma, delta, DFA, age, TST, WASO, R/TST, N1/TST, and N2/TST). Nonnormally distributed continuous variables (fast ratio, AHI, BMI, SE (%), sleep latency, and N3/TST (%)) were analyzed using a nonparametric test and presented as the median (25th-75th percentile).

A Spearman correlation was used to analyze the correlation between cognitive impairment and power spectral analysis and DFA. The values of DFA in NREM sleep were used to diagnose cognitive dysfunction using receiver operating characteristic (ROC) curve analysis. A p value of < 0.05 indicated a significant difference.

SPSS 17.0 software (SPSS Inc., Chicago, IL, USA) was used for statistical analysis.

3. Results

3.1. Demographic Characteristics and Comparison of Sleep Parameters Monitored by PSG. Forty patients with PDS (age: 46–84 years; mean ± standard deviation: 66.83 ± 9.18 years; and 25 men and 15 women) were included: 15 cases of PD, 13 cases of MSA, and 12 cases of PDS with unknown etiologies. Based on cognitive evaluations, participants were stratified into two groups: 25 patients with normal cognitive function (PDS-NC) and 15 with mild cognitive impairment (PDS-MCI).

No significant differences were revealed between the two groups in terms of age, sex, composition ratio, BMI, AHI, and sleep-related parameters (total sleep time (TST), sleep efficiency (SE), sleep latency (SL), wake time after sleep onset (WASO), and the proportion of each sleep period to total sleep) (Table 1).

Of the 40 patients evaluated, 18 had REM sleep without atonia (RWA) on PSG, and 22 had no RWA. In the PDS-NC group, 11 had RWA and 14 had no RWA; in the PDS-MCI group, 7 had RWA and 8 had no RWA. The

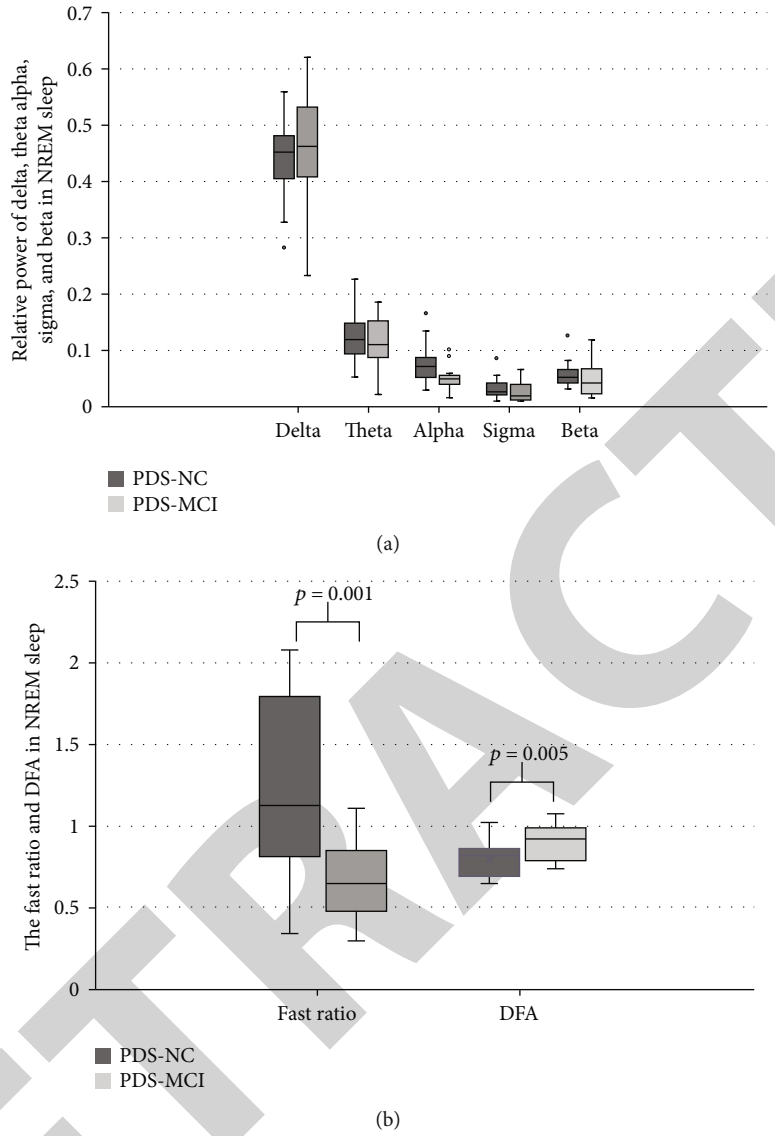


Figure 2: Comparisons of the power spectral analysis and DFA in NREM sleep between the PDS-MCI and PDS-NC groups.

composition ratio of RWA in the two groups was not significantly different ($p > 0.05$).

3.2. Comparisons of the Power Spectral Analysis and DFA between the PDS-MCI and PDS-NC Patients

3.2.1. General Summary of the Comparisons of the Power Spectral Analysis and DFA. The alpha activity and fast ratio ((alpha + beta)/(delta + theta)) were lower during wakefulness and NREM sleep in patients with PDS-MCI than in those with PDS-NC ($p < 0.05$).

The sigma and beta activities were lower during wakefulness in patients with PDS-MCI than in those with PDS-NC ($p < 0.05$).

DFA increased in patients with PDS-MCI during wakefulness and NREM sleep compared to those with PDS-NC ($p < 0.05$).

The other power spectral analysis characteristics were not significantly different in the two groups ($p > 0.05$).

The results of channels C4 and O1 were similar. The data of the power spectral analysis and DFA (those of p values < 0.05) are shown in Table 2 and Figure 1.

All the variables analyzed in the study are summarized in the Supplementary Figure (see available here).

3.2.2. Power Spectral Analysis and DFA in the NREM Sleep, REM Sleep, and Wake Stages. The O1 lead was selected as the analysis lead. The p value was adjusted to 0.05/7 = 0.007 as there were seven indicators analyzed (alpha, beta, theta, sigma, delta, fast ratio, and DFA) in each sleep stage. $p < 0.007$ was considered significant.

(1) Comparisons of the Power Spectral Analysis and DFA in NREM Sleep between Both Groups. The results showed that PDS-MCI patients had a decreased EEG fast ratio (0.65 ± 0.24 vs. 1.24 ± 0.55, $p = 0.001$) and increased DFA (0.90 ± 0.12 vs. 0.80 ± 0.12, $p = 0.005$) during NREM compared to PDS-NC patients ($p < 0.007$) (Figure 1).

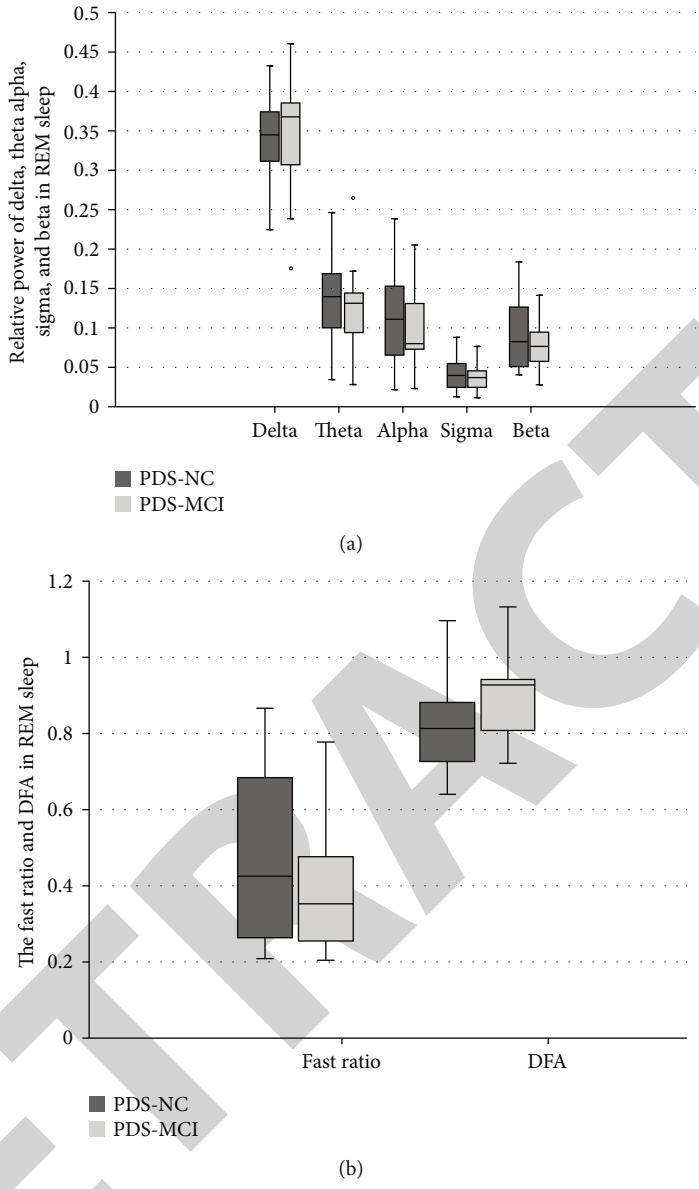


Figure 3: The comparisons of the power spectral analysis and DFA in REM sleep between the PDS-MCI and PDS-NC groups.

The alpha, beta, theta, sigma, and delta activities are not different between the PDS-MCI and PDS-NC groups ($p > 0.007$) (Figure 2).

(2) Comparisons of the Power Spectral Analysis and DFA in REM Sleep between the PDS-MCI and PDS-NC Groups. The alpha, beta, theta, sigma, and delta activities, EEG fast ratio, and DFA of REM sleep are not different between the PDS-MCI and PDS-NC groups ($p > 0.007$) (Figure 3).

(3) Comparisons of the Power Spectral Analysis and DFA during Wakefulness between the PDS-MCI and PDS-NC Groups. The beta activity and fast ratio were significantly lower in PDS-MCI patients (0.08 ± 0.032; 0.73 ± 0.28) than in PDS-MCI patients (0.12 ± 0.041; 1.40 ± 0.64) ($p = 0.003$; $p = 0.001$, respectively).

The alpha, theta, sigma, and delta activities and DFA during wakefulness did not differ between the PDS-MCI and PDS-NC groups ($p > 0.007$) (Figure 4).

3.3. Correlations between DFA and Fast Ratio in NREM Sleep and MCI. Mild cognitive dysfunction was positively correlated with DFA in NREM ($r = 0.426$, $p = 0.007$) and negatively correlated with the fast ratio in NREM ($r = -0.524$, $p = 0.001$) in channel O1 during NREM sleep.

3.4. ROC Curve of DFA and Fast Ratio in NREM Sleep for the Diagnosis of PDS-MCI. The area under the curve was 0.753 (95% CI: 0.592-0.914) for DFA applied to diagnose MCI in channel O1 during NREM sleep ($p = 0.009$). The cutoff value was 0.92 for channel O1, with sensitivity (60.0%) and specificity (91.7%), respectively (Figure 5).

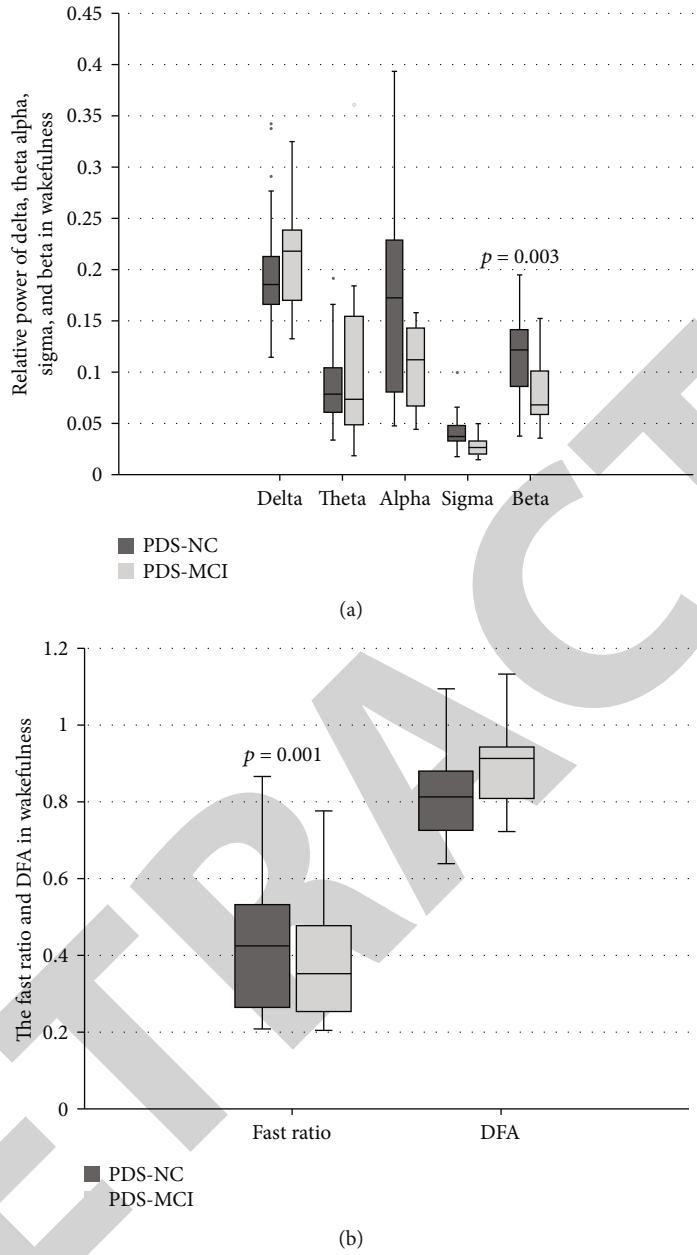


Figure 4: The comparisons of the power spectral analysis and DFA during wakefulness between the PDS-MCI and PDS-NC groups.

The area under the curve for the fast ratio was only 0.183 ($p = 0.002$).

4. Discussion

This study demonstrated that the power spectral analysis and DFA characteristics of NREM sleep EEG were related to MCI in patients with PDS, who showed a reduced fast ratio ((alpha + beta)/(delta + theta)) and increased DFA. DFA during NREM was diagnostic of PDS-MCI, with an area under the ROC curve (AUC) of 0.753 (95% CI: 0.592–0.914) ($p < 0.05$). Therefore, we believe that there are characteristic alterations of NREM sleep EEG in patients with MCI. DFA and the fast ratio can be used as biomarkers reflecting cognitive function.

Cognitive decline is common in the progression of parkinsonism [3]. Impaired cognition has a major impact on either quality of life or mortality in patients with PDS [3]. The early prediction of MCI in patients with PDS is important to intervene before the development of dementia.

Neuropsychological testing is undoubtedly an important method for determining the cognitive function of patients with PDS. However, it also has several limitations. For example, it is subjective; the repeated use of the scale can induce learning effects. Thus, various objective markers have been investigated for the early identification of PDS-MCI, such as epidermal and insulin-like growth factors in plasma/serum and β-amyloid in cerebrospinal fluid [22], and hippocampal atrophy on brain magnetic resonance imaging [3]. Despite the diversity of studies [23], no biomarker has yet

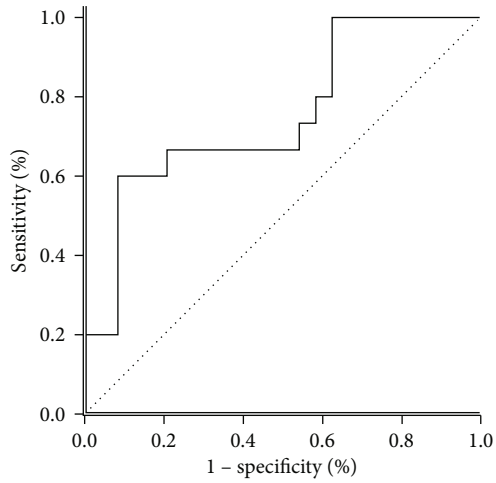


Figure 5: ROC curve of DFA in NREM sleep for the diagnosis of PDS-MCI with the AUC of 0.753 (95% CI: 0.592-0.914).

been validated. Among these markers, qEEG has shown good potential in identifying cognitive decline in PDS [11]. For example, theta power and alpha power help to distinguish PD with MCI from PD with normal cognition, and delta power and beta power help to distinguish PD with MCI from PD with dementia. In fact, besides the resting awake EEG, sleep EEG activity also relates to cognitive performance [24] and even general intelligence [25]. For example, some sleep spindle characteristics and sigma EEG power are related to improved learning and memory consolidation [26]. In addition, reductions in slow-wave, theta, and sigma activities and sleep spindle characteristics during NREM sleep may be related to MCI in older adults [12, 27].

In addition to conventional power spectral analysis, DFA is a simpler alternative method to quantify EEG. It is more suitable for the analysis of nonlinear and nonstationary physiological data [28] such as EEG and more easily removes artifacts [29]. Previous studies showed that OSA patients had a higher DFA during wakefulness than controls, which may be a marker of impaired performance of these patients [14]. DFA of EEG also performs as well as conventional power spectral analysis as a marker of impaired performance and sleepiness resulting from sleep loss [14]. DFA has been used to analyze EEG as a new alternative biomarker of cognitive impairment [14, 16]. It has been applied to the analysis of the EEG background activity in Alzheimer's disease [16, 30].

The relationship between the characteristics of sleep EEG power spectral analysis and DFA and cognitive dysfunction in patients with PDS remains unclear. In this study, we extracted and analyzed the PSG-EEG signals in the wake, NREM sleep, and REM sleep stages throughout the night. We sought to observe the changes in sleep EEG in MCI patients from the perspective of sleep. Our study found that PDS-MCI patients had a tendency of reduced alpha activity during NREM sleep. Studies have shown that alpha activity in NREM sleep is associated with intelligence in the general population [25]. The reduction in alpha activity in NREM sleep may reflect a change in the electrophysiological basis for memory consolidation and other cognitive activities. In fact, some studies have already shown that changes in EEG activities during NREM sleep are associated with a higher risk of early decline in cognitive function in older adults [12].

Another finding with regard to the EEG power spectral analysis is a decreased fast ratio ((alpha + beta)/(delta + theta)) during NREM sleep in the PDS-MCI group. Generally, slowing of resting awake EEG frequencies, which can be manifested as a decrease in alpha and beta activities and an increase in delta and theta activities, correlates with PD-related cognitive decline [14]. EEG slowing during wakefulness was also found in OSA patients with a higher EEG slowing ratio ((delta + theta)/(alpha + beta)), which may be linked to the subjective sleepiness of OSA patients [25]. The present study found that during NREM sleep, patients with PDS-MCI showed slowing of sleep EEG activity, which is similar to that in wakefulness, as previously reported. Many studies have shown that sleep, especially NREM sleep, promotes long-term consolidation of memories connecting the hippocampus and neocortex and is important for the maintenance of brain function [25]. Slowing of EEG activity during NREM sleep may reflect the decline in NREM physiological function and is therefore a potential marker in patients with PDS-MCI.

In addition, we found that a new parameter, DFA, was related to cognition in patients with PDS. Previous studies have shown that DFA can reflect the brain activity and complexity of an EEG. It increases during the transition from wake to sleep and continues to increase with deeper stages of NREM sleep [31]. A higher DFA reflects a lower EEG complexity. Our study showed that DFA in NREM sleep increased in PDS-MCI patients, reflecting a decrease in the complexity of EEG. The complexity of EEG is related to neurodegeneration and brain function. Recent studies have demonstrated that the decreased complexity of EEG is a predictor of neurodegeneration in patients with idiopathic rapid eye movement behavior disorders [32].

The complexity of EEG also correlates to the mild cognitive impairment of Alzheimer's disease (AD) [33]. Compared to healthy individuals, the EEG signals from AD patients have less complexity, and this complexity is considered a potential biomarker for the functional abnormalities in AD [33]. Our result suggests that reduced EEG complexity may also be associated with cognitive decline seen in patients with PDS; DFA is a promising marker for it.

The main limitations of the study are as follows. (1) In this retrospective study, NREM sleep EEG characteristics were correlated with MCI, but it remains to be investigated whether there is a causal relationship and whether it can predict MCI. (2) To assess whether MCI is independently correlated with these EEG parameters, a regression analysis would be more appropriate. However, the sample size was too small to allow for a regression analysis. This study is a preliminary study, and the sample size can be increased for subsequent analysis.

The early prediction of MCI in patients with PDS is important for early intervention prior to the development of dementia. Recent studies suggest that EEG changes precede the onset of the clinical signs of cognitive impairment

in PDS [11, 32]. This retrospective study points out that EEG characteristics and DFA parameters in NREM sleep are associated with MCI in patients with PDS. Future prospective studies are required to see whether DFA parameters during NREM sleep could predict early cognitive decline and thus provide a basis for early intervention.

5. Conclusion

In summary, this study demonstrated that the power spectral analysis and DFA characteristics of NREM sleep EEG were related to MCI in patients with PDS. DFA can be used as a biomarker reflective of cognitive function. Slowing of EEG activity during NREM sleep may reflect the decline in NREM physiological function and is therefore a marker in patients with PDS-MCI. The early prediction of MCI in patients with PDS is important for early interventions prior to the development of dementia.

Data Availability

Additional unpublished data from the study are available from the first or corresponding author on reasonable request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Authors' Contributions

Cheng Zhang and Luhua Wei contributed equally to the manuscript and are co-first author in this article.

Acknowledgments

We would like to express our appreciation to all the staff involved in the sleep study and cognitive assessment of PDS patients.

Supplementary Materials

Supplementary Figure. Fig. S: the variables of the power spectral analysis and DFA analyzed in the study are summarized in the Supplementary Figure. *(Supplementary Materials)*

References

[1] R. B. Postuma, D. Berg, M. Stern et al., "*MDS clinical diagnostic criteria for Parkinson's disease*," *Movement Disorders*, vol. 30, no. 12, pp. 1591–1601, 2015.

[2] M. T. Hayes, "*Parkinson's disease and parkinsonism*," *The American Journal of Medicine*, vol. 132, no. 7, pp. 802–807, 2019.

[3] M. Delgado-Alvarado, B. Gago, I. Navalpotro-Gomez, H. Jiménez-Urbieta, and M. C. Rodriguez-Oroz, "*Biomarkers for dementia and mild cognitive impairment in Parkinson's disease*," *Movement Disorders*, vol. 31, no. 6, pp. 861–881, 2016.

[4] J. Gratwicke, M. Jahanshahi, and T. Foltynie, "*Parkinson's disease dementia: a neural networks perspective*," *Brain*, vol. 138, no. 6, pp. 1454–1476, 2015.

[5] S. Koga, A. Parks, R. J. Uitti et al., "*Profile of cognitive impairment and underlying pathology in multiple system atrophy*," *Movement Disorders*, vol. 32, no. 3, pp. 405–413, 2017.

[6] A. Pilotto, S. Gazzina, A. Benussi et al., "*Mild cognitive impairment and progression to dementia in progressive supranuclear palsy*," *Neurodegenerative Diseases*, vol. 17, no. 6, pp. 286–291, 2017.

[7] N. Sakae, O. A. Santos, O. Pedraza et al., "*Clinical and pathologic features of cognitive-predominant corticobasal degeneration*," *Neurology*, vol. 95, no. 1, pp. e35–e45, 2020.

[8] G. S. Watson and J. B. Leverenz, "*Profile of cognitive impairment in Parkinson's disease*," *Brain Pathology*, vol. 20, no. 3, pp. 640–645, 2010.

[9] M. Skorvanek, J. G. Goldman, M. Jahanshahi et al., "*Global scales for cognitive screening in Parkinson's disease: critique and recommendations*," *Movement Disorders*, vol. 33, no. 2, pp. 208–218, 2018.

[10] U. Smailovic and V. Jelic, "*Neurophysiological markers of Alzheimer's disease: quantitative EEG approach*," *Neurol Ther*, vol. 8, Suppl 2, pp. 37–55, 2019.

[11] V. V. Cozac et al., "*Quantitative EEG and cognitive decline in Parkinson's disease*," *Parkinsons Dis*, vol. 2016, p. 9060649, 2016.

[12] J. Taillard, P. Sagaspe, C. Berthomier et al., "*Non-REM sleep characteristics predict early cognitive impairment in an aging population*," *Frontiers in Neurology*, vol. 10, p. 197, 2019.

[13] J. N. Caviness, R. L. Utianski, J. G. Hentz et al., "*Differential spectral quantitative electroencephalography patterns between control and Parkinson's disease cohorts*," *European Journal of Neurology*, vol. 23, no. 2, pp. 387–392, 2016.

[14] A. L. D'Rozario, J. W. Kim, K. K. H. Wong et al., "*A new EEG biomarker of neurobehavioural impairment and sleepiness in sleep apnea patients and controls during extended wakefulness*," *Clinical Neurophysiology*, vol. 124, no. 8, pp. 1605–1614, 2013.

[15] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, "*Mosaic organization of DNA nucleotides*," *Physical Review. E, Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 49, no. 2, pp. 1685–1689, 1994.

[16] D. Abasolo, R. Hornero, J. Escudero, and P. Espino, "*A study on the possible usefulness of detrended fluctuation analysis of the electroencephalogram background activity in Alzheimer's disease*," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 9, pp. 2171–2179, 2008.

[17] I. Litvan, J. G. Goldman, A. I. Tröster et al., "*Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines*," *Movement Disorders*, vol. 27, no. 3, pp. 349–356, 2012.

[18] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[19] K. Polat and S. Güneş, "Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform," *Applied Mathematics & Computation*, vol. 187, no. 2, pp. 1017–1026, 2007.

[20] A. F. Farag, S. M. El-Metwally, and A. A. A. Morsy, "Automated sleep staging using detrended fluctuation analysis of sleep EEG.," in Soft Computing Applications, 2013.

[21] J. Zhou, X. M. Wu, and W. J. Zeng, "*Automatic detection of sleep apnea based on EEG detrended fluctuation analysis and*

Hindawi
BioMed Research International
Volume 2023, Article ID 9820815, 1 page
https://doi.org/10.1155/2023/9820815

Retraction

Retracted: Protocatechualdehyde Inhibits the Osteoclast Differentiation of RAW264.7 and BMM Cells by Regulating NF-κB and MAPK Activity

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Qu, X. Liu, S. Zong, H. Sun, S. Liu, and Y. Zhao, "Protocatechualdehyde Inhibits the Osteoclast Differentiation of RAW264.7 and BMM Cells by Regulating NF-κB and MAPK Activity," *BioMed Research International*, vol. 2021, Article ID 6108999, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6108999, 11 pages
https://doi.org/10.1155/2021/6108999

Research Article

Protocatechualdehyde Inhibits the Osteoclast Differentiation of RAW264.7 and BMM Cells by Regulating NF-κB and MAPK Activity

Yunyun Qu,[1,2] Xin Liu,[1] Shuai Zong,[1] Huanxin Sun,[1] Shuang Liu,[1] and Yueran Zhao[1]

[1]*Department of Central Lab, Shandong Provincial Hospital, Cheeloo College of Medicine, Shandong University, Jinan, Shandong 250021, China*
[2]*The Second Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan, China*

Correspondence should be addressed to Yueran Zhao; yrzhao@sdu.edu.cn

Received 11 April 2021; Accepted 7 June 2021; Published 16 July 2021

Academic Editor: Min Tang

Protocatechualdehyde (PCA), an important component of *Salvia miltiorrhiza*, has many activities, such as anti-inflammatory and antisepsis activities. However, the role of PCA in osteoclasts is not clear. We used RAW264.7 cells (a mouse leukemic monocyte/macrophage cell line) and bone marrow macrophages (BMMs) to probe the role of PCA in osteoclasts and the underlying mechanism. The effects of PCA on cell activity were evaluated with CCK-8 assays. TRAP staining detected mature osteoclasts. Corning Osteo Assay Surface plates were used to examine absorption. The levels of RNA and protein were analyzed, respectively, using RT-PCR and Western blotting. PCA (5 μg/ml) was not toxic to the two cell types but reduced the formation of osteoclasts and bone absorption. Furthermore, PCA restrained the expression of mRNAs encoding proteins associated with osteoclasts and reduced the phosphorylation of proteins in important signaling pathways. The results indicate that PCA inhibits osteoclast differentiation by suppressing NF-κB and MAPK activity.

1. Introduction

Osteoporosis is a degenerative disease with an increasing risk due to aging and is one of the most common diseases in older men [1]. The disease can occur at any age but mainly occurs in elderly men. Approximately 200 million people suffer from osteoporosis worldwide, and this condition has become the seventh most common disease in the world.

Under normal circumstances, bone tissue maintains normal function by constantly updating metabolism. Due to abnormal metabolism, older patients experience bone loss. The increased activity of osteoclasts increases the loss of bone mass, and the attenuated function of osteoblasts decreases the formation of bone mass, leading to fractures in the bone microstructure and a negative balance in bone reconstruction [2]. Therefore, more bone loss than supplementation is the cytological basis for the occurrence of osteoporosis.

Currently, bisphosphonates are common agents used to treat osteoporosis and significantly inhibit the activity of osteoclasts. They have been recommended by many authoritative institutions worldwide due to their clear efficacy [3]. However, the safety of long-term bisphosphonate use has gradually become a concern. Long-term application of bisphosphonates was reported to cause mandibular osteonecrosis, atypical subtrochanteric or femoral shaft fracture, and other adverse reactions [4].

Salvia miltiorrhiza (SM) is a Chinese herbal medicine that has long been used to treat osteoporosis. Clinical trials have shown that SM alone is as effective as traditional drugs in the treatment of osteoporosis and has the advantage of being nontoxic [5]. *In vitro* experiments have confirmed that SM exerts a good protective effect on ovariectomy-induced osteoporosis and diabetic osteoporosis [6, 7]. Salvianolic acid B is a water-soluble extract of SM. It exerts a protective effect

on prednisone-induced osteoporosis [8], increasing bone cancellation, improving the bone trabecular structure, and increasing bone density.

Protocatechualdehyde (PCA) is the main degradation product of salvianolic acid B and has many activities, such as anti-inflammatory and antibacterial properties [9]. *In vivo*, the prodrug PCA restrains the destruction of the femoral condylar cartilage caused by adjuvant arthritis [10]. PCA significantly improves bone physical indexes, increases bone density and the levels of bone minerals and organic matter, increases the ratio of important organs, and effectively prevents the occurrence of osteoporosis caused by prednisone acetate in the growth period of rats [11]. However, the role of PCA in osteoclasts is not clear.

In this study, RAW264.7 cells (a mouse leukemic monocyte/macrophage cell line) and bone marrow macrophages (BMMs) were used to analyze the role of PCA in osteoclasts and the underlying mechanism. PCA has important theoretical significance and clinical application value in the search for economical and effective drug monomers to treat osteoporosis.

2. Materials and Methods

2.1. Materials. RAW264.7 cells were selected from ATCC (Manassas, VA, USA). BALB/c mice were purchased from Shandong University. PCA (purity > 98% by HPLC) was obtained from Sigma. Cell culture media were obtained from BI. Cytostimulators were acquired from R&D Systems. CCK-8 and TRAP kits were purchased from Sigma. 24-well bone plates were purchased from Corning Life Science (Lowell, MA, USA). Phalloidin was obtained from Thermo Fisher Scientific. The apoptosis kit was purchased from BD. All antibodies were acquired from Cell Signaling Technology (Beverly, MA, USA). Primers were designed by PrimerBank and synthesized by Shanghai Bioengineering Company.

2.2. Cell Culture. BMMs were acquired from both ends of cut femurs and tibias from approximately 6-week-old male BALB/c mice. The animal experiment was authorized by the Experimental Animal Ethics Review Committee of Shandong University. The marrow cavity was rinsed 3-5 times with α-MEM, and red blood cells were removed to obtain BMMs. The lysate was filtered through a 40 μm mesh filter and centrifuged to separate the cells. Samples from the mice were cultured overnight in α-MEM containing M-CSF (5 ng/ml) in a nonadherent culture plate, and then, cells suspended in the collected supernatant were cultured with M-CSF (30 ng/ml) for 3 days. The supernatant and suspended cells were discarded; the adherent cells were the BMMs.

RAW264.7 cells were cultured in complete DMEM. However, DMEM was substituted for α-MEM to culture osteoclasts induced by receptor activator of nuclear factor-κB ligand (RANKL).

2.3. Cytotoxicity Assay. CCK-8 assays were used to determine the toxicity of PCA toward cells. In 96-well plates, BMMs (2×10^4 cells/well) and RAW264.7 cells (1×10^4 cells/well) were seeded in the presence of PCA (0, 1, 10, 20, 30, or 50 μg/ml) for 5–6 days. CCK-8 (10 μl/well) was added to the supernatant and incubated for 0.5–1 h, and the absorbance was assayed at 452 nm by a Multiskan Go1510 microplate reader. The ratio of the experimental group to the control group represents cell activity [12].

Table 1: The conditions and primer sequences used for RT-PCR.

Target genes	Primer sequences	Tm	Size
GAPDH	F: AGGTCGGTGTGAACGGATTTG R: GGGGTCGTTGATGGCAACA	60	95
CTSK	F: CTCGGCGTTTAATTTGGGAGA R: TCGAGAGGGAGGTATTCTGAGT	60	164
NFATc1	F: GGAGAGTCCGAGAATCGAGAT R: TTGCAGCTAGGAAGTACGTCT	60	102
C-FOS	F: CGGGTTTCAACGCCGACTA R: TGGCACTAGAGACGGACAGAT	60	165
MMP-9	F: GCAGAGGCATACTTGTACCG R: TGATGTTATGATGGTCCCACTTG	60	229
DC-STAMP	F: GGGGACTTATGTGTTTCCACG R: ACAAAGCAACAGACTCCCAAAT	60	72

2.4. TRAP Staining. RAW264.7 cells (5×10^3 cells/well) were cultured in 48-well plates, and BMMs (2×10^4 cells/well) were seeded in 96-well plates. The next day, 50 ng/ml RANKL and 30 ng/ml M-CSF were added to the BMM culture media, and 50 ng/ml RANKL was added to the RAW264.7 culture media. In addition, PCA (0, 1, 2.5, or 5 μg/ml) were added to the medium. After 4–6 days, the cells were stained and identified according to the literature [13].

2.5. Absorption Pit Assay. We used Corning Osteo Assay Surface 24-well plates to assay absorption pits. BMMs (4×10^5 cells/well) and RAW264.7 cells (1×10^5 cells/well) were seeded and incubated for 24 h. Next, 50 ng/ml RANKL and 30 ng/ml M-CSF were added to the BMM culture media, and 50 ng/ml RANKL was added to the RAW264.7 culture media. In addition, PCA (1, 2.5, or 5 μg/ml) was added at the same time, and the cells were incubated for 4–8 days. After 4–8 days, a 10% bleach solution was used to wash the cells. A light microscope (Olympus) captured images of the absorption pits, and ImageJ software quantified the image [14].

2.6. Immunofluorescence Staining of the F-Actin Ring. BMMs (3×10^5 cells/well) were seeded in 24-well plates, and a series of PCA, M-CSF (30 ng/ml), and RANKL (50 ng/ml) were added to the medium. The cells were incubated for 4 days, and phalloidin was used to stain the cells after fixation and permeabilization. A fluorescence microscope was used to capture fluorescence images.

2.7. Flow Cytometry Analysis. RAW264.7 cells (5×10^5 cells/well) were seeded in 6-well plates and cultured with a series of PCA and 50 ng/ml RANKL. After 3–4 days, the apoptosis kit was used to stain the cells. Cell apoptosis was assessed using flow cytometry.

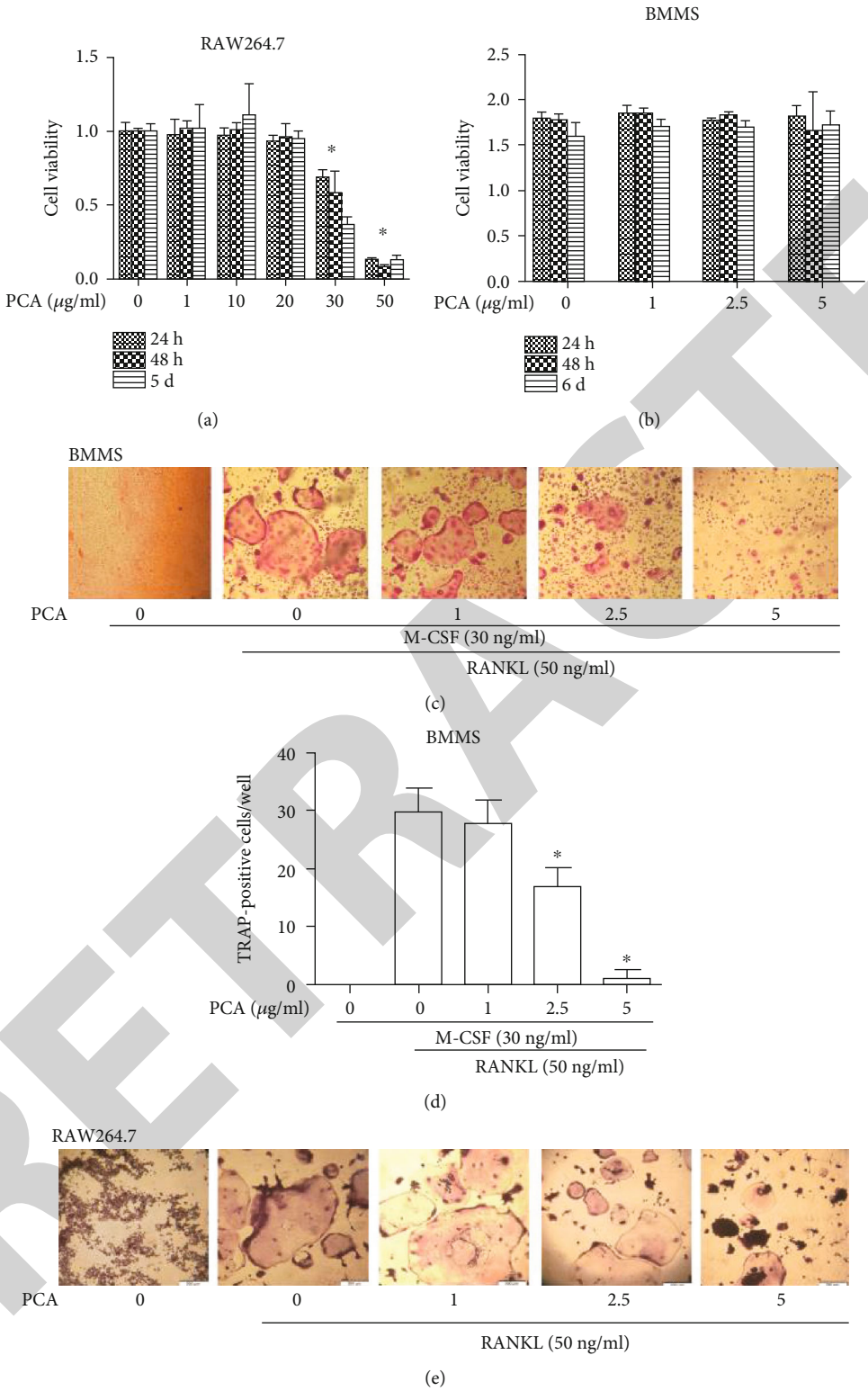


Figure 1: Continued.

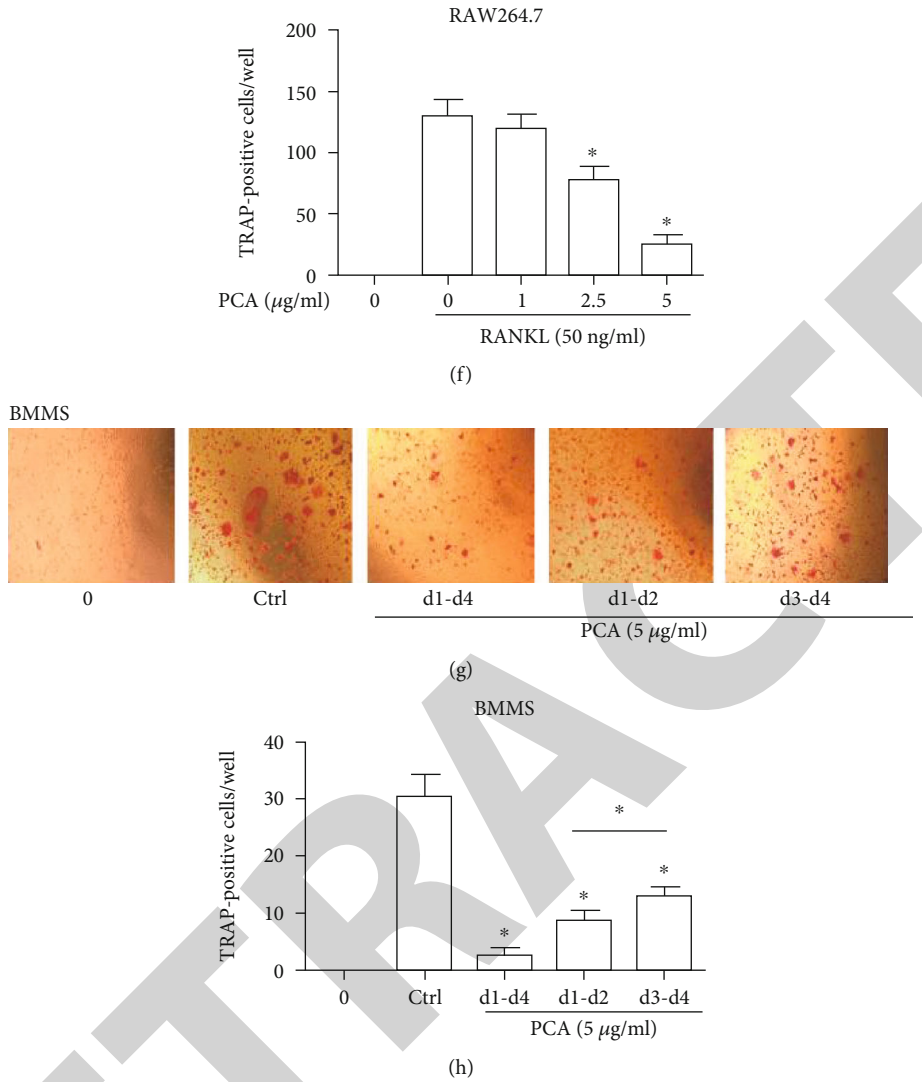


Figure 1: PCA attenuates osteoclast differentiation induced by RANKL. (a) Viability of RAW264.7 cells exposed to PCA at different time points. (b) Viability of BMM cells exposed to PCA. (c) BMMs were induced to osteoclasts in 4 days. (d) Number of osteoclasts per well. (e) RAW264.7 cells were induced to differentiate into osteoclasts after 5 days. (f) Number of osteoclasts per well. (g) PCA (5 μg/ml) was added to BMMs at three periods during the differentiation of osteoclasts. (h) Number of osteoclasts per well. A light microscope was used to acquire photomicrographs (magnification 10×10). Cells that were TRAP-positive and had more than three nuclei were considered osteoclasts. The quantities are presented as the mean ± SEM values ($n = 3$). $^{*}P < 0.01$ and $^{\#}P < 0.05$ compared with the control group.

2.8. RT-qPCR. RAW264.7 cells were seeded in 24-well plates (5×10^4 cells/well). The next day, the cells were pretreated with a series of PCA concentrations for 60 min before adding 50 ng/ml RANKL. After 5–6 days, the cells were collected. PCR was performed as previously described [15]. Primers were listed in Table 1.

2.9. Western Blot Analysis. RAW264.7 cells (2.5×10^6 cells/m) were plated in 6-well plates and then pretreated with PCA (1, 2.5, or 5 μg/ml) for 2 h before the addition of RANKL (50 ng/ml) and incubation for 10 min. Western blots were performed as previously described [11]. The levels of p-p38/p38, p-ERK/ERK, p-JNK/JNK, and p-p65/p65 are shown in histograms.

2.10. Statistical Analysis. The results are presented as means ± standard errors. Each experiment was repeated more than three times. SPSS19.0 software was used to analyze differences. $P < 0.05$ was defined as a statistically significant difference.

3. Results

3.1. PCA Attenuates Osteoclast Differentiation Induced by RANKL. The potential cytotoxicity of PCA toward RAW264.7 cells (Figure 1(a)) and BMMs (Figure 1(b)) was evaluated using CCK-8 assays. Based on the results, the highest concentration of PCA (5 μg/ml) has no effect on cell activity.

BMMs and RAW264.7 cells were successfully induced to differentiate into osteoclasts by RANKL. The osteoclasts were large and full, with vacuoles forming in the middle and

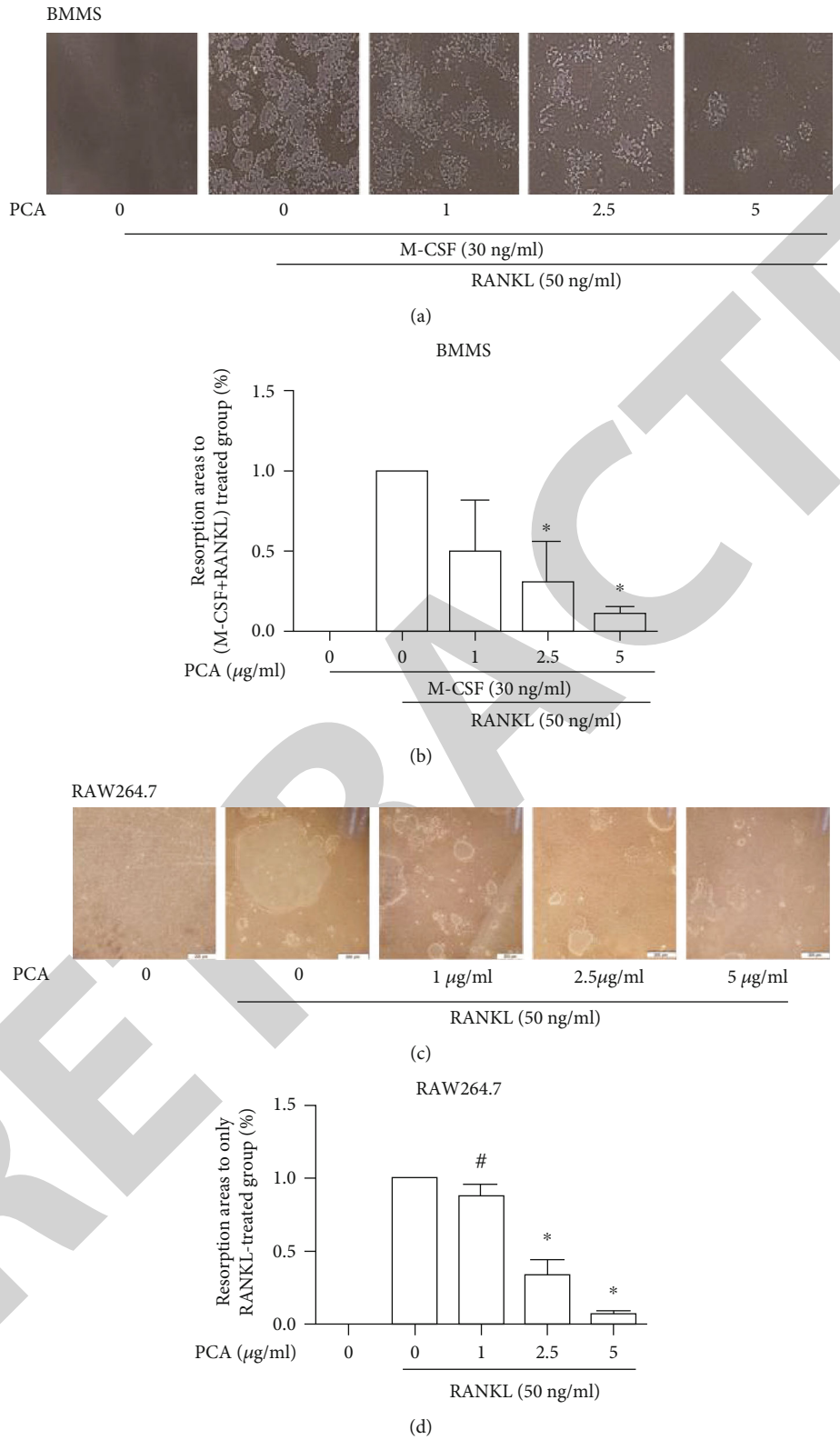


Figure 2: Continued.

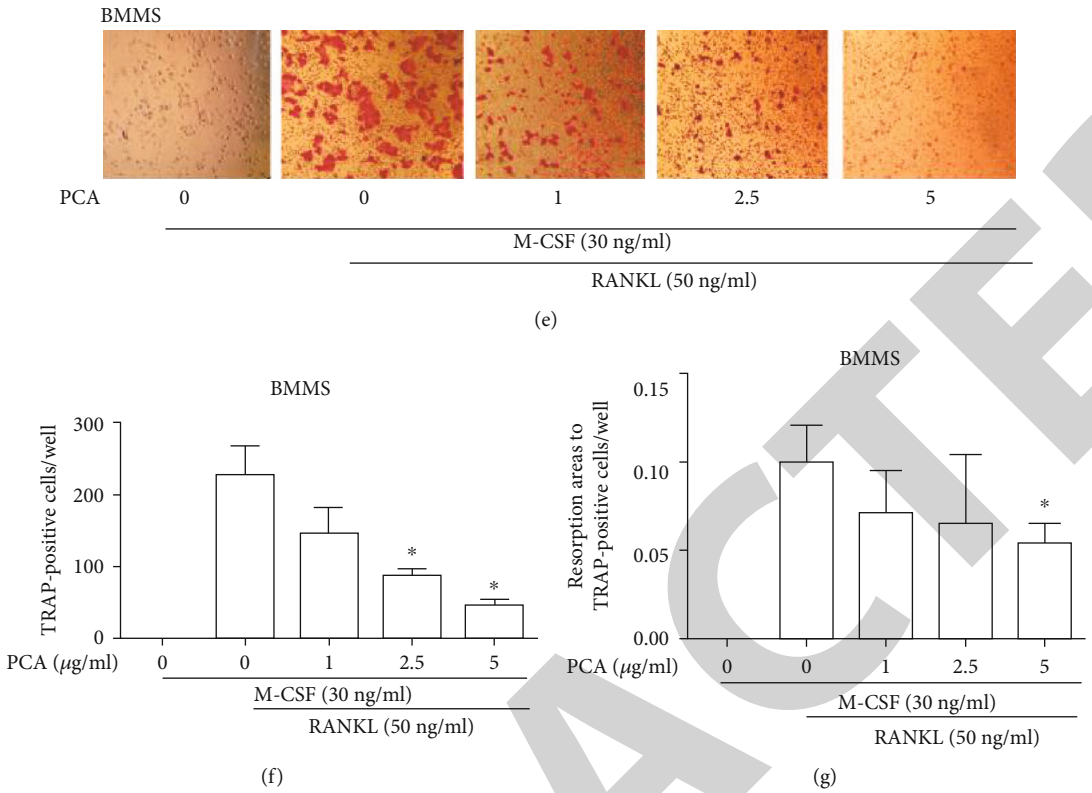


Figure 2: PCA reduced bone absorption. (a) BMMs were grown in Osteo Assay Surface plates for 4 days (magnification 4×10). (b) Absorption area per well. (c) RAW264.7 cells were grown in Osteo Assay Surface plates for 7–8 days (magnification 10×10). (d) Resorption area per well. (e) TRAP staining of osteoclasts differentiated from BMMs on Corning CaP-coated plates (magnification 10×10). (f) Number of osteoclast per well. (g) Analysis of the absorption area per osteoclast-like cell. Absorption pits were quantified using ImageJ software. $^{*}P < 0.01$ and $^{\#}P < 0.05$ compared with the group treated with RANKL alone.

additional nuclear accumulation throughout the cells. However, when PCA was present, the size of the osteoclasts decreased, the vacuoles disappeared, and the number of surrounding nuclei decreased. Furthermore, we used TRAP to stain osteoclasts. The test results indicated that PCA prevented osteoclastic differentiation not only in BMMs but also in RAW264.7 cells (Figures 1(c)–1(f)). Thus, PCA attenuates the generation of osteoclasts.

We added 5 μg/ml PCA to the medium and induced BMMs to differentiate in three periods to determine the stage of osteoclastogenesis at which PCA exerted its effect. Osteoclast formation was markedly inhibited by PCA in the earlier period (Figures 1(g) and 1(h)), whereas the inhibitory effect of PCA decreased in the later period. These data suggested that PCA suppresses osteoclast formation mainly in the earlier period of RANKL-induced differentiation.

3.2. PCA Decreases the RANKL-Induced Formation of Absorption Pits. The absorption of osteoclasts induced by RANKL was detected. The group stimulated with RANKL formed obvious bone absorption pits, but the number and area of absorption pits decreased in the presence of PCA (Figure 2). Therefore, PCA inhibits the function of osteoclasts.

3.3. PCA Suppresses the Size of F-Actin Rings and Does Not Affect Apoptosis. The size of the F-actin ring was measured, and PCA reduced the length of the F-actin rings (Figures 3(a) and 3(b)). However, the flow cytometry analysis showed that PCA did not affect the ratio of apoptotic cells (Figures 3(c) and 3(d)). These data suggested that the inhibitory effect of PCA is not attributable to apoptosis.

3.4. PCA Attenuates the mRNA Expression of RANKL-Induced Osteoclastic Markers. The RT-qPCR results showed the mRNA levels of factors produced by RANKL-stimulated osteoclasts. The RANKL-induced group exhibited substantially increased expression of these mRNAs (Figure 4). However, PCA significantly reduced the mRNA levels of these related factors.

3.5. PCA Suppresses the Activation of the MAPK Pathway. The MAPK signaling pathway exerts an irreplaceable effect on the generation of osteoclasts. Primary members of this signaling pathway are ERK, JNK, and p38. Phosphorylation of these proteins is the main indicator of activation. We measured the levels of these proteins produced during the formation of osteoclasts in cells treated with or without PCA. The RANKL-induced group exhibited substantially increased phosphorylation of these proteins. However, the changes were inhibited by PCA (Figure 5).

3.6. PCA Prevents RANKL-Induced NF-κB Activation. The change in p65 expression was examined. We measured the levels of p65 phosphorylation during RANKL-stimulated

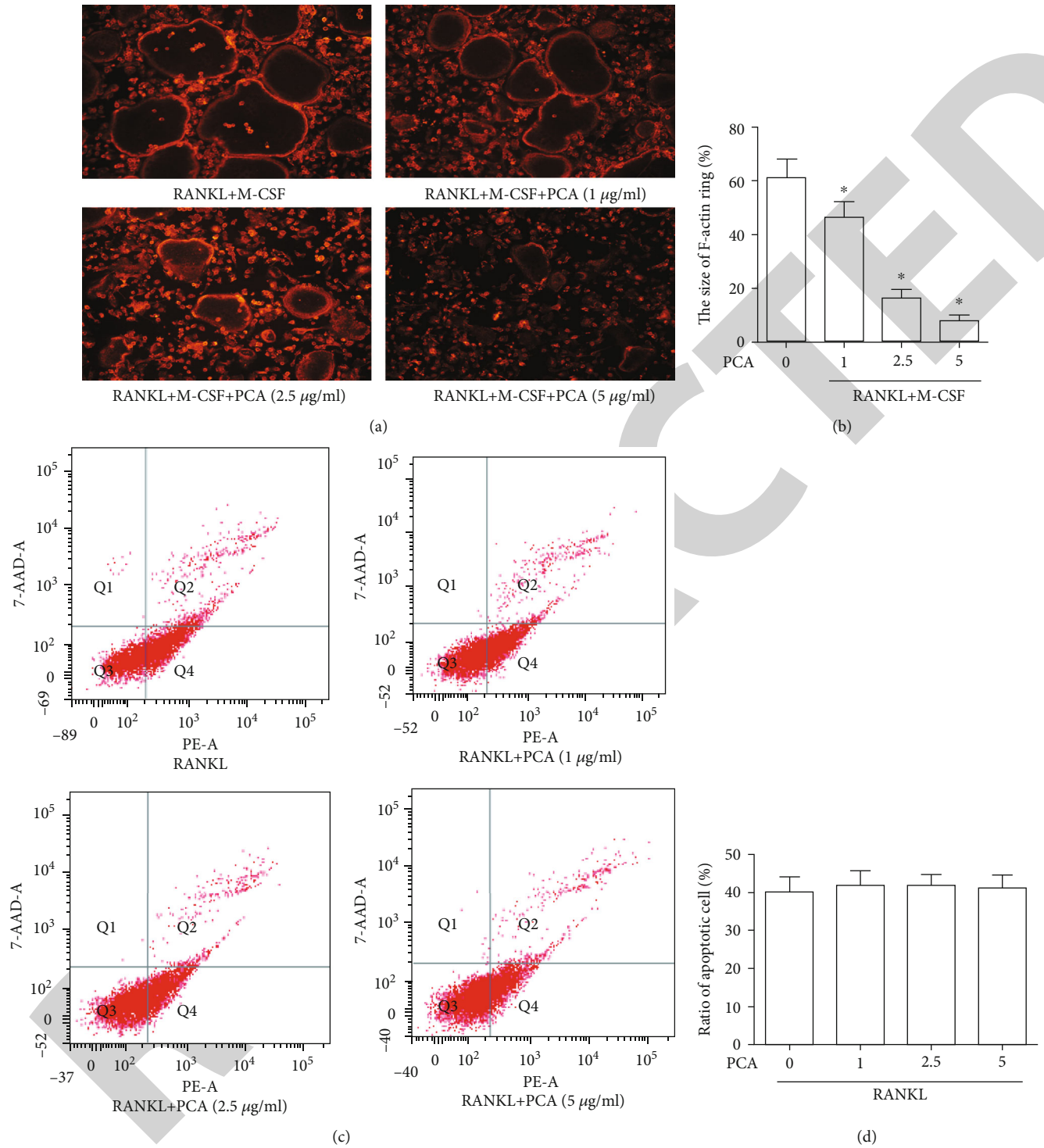


Figure 3: PCA reduces the size of F-actin rings. (a) A fluorescence microscope was used to capture the fluorescence images of F-actin rings (magnification 40x). (b) The F-actin ring size was quantified. $^{*}P < 0.01$ compared with the M-CSF+RANKL group. (c) Flow cytometry was used to detect cells. (d) The ratios of apoptotic cells to control cells. $^{*}P < 0.01$ compared with the control group.

differentiation of cells treated with or without PCA. As shown in Figure 6, the RANKL-stimulated group exhibited increased p65 phosphorylation. However, when PCA was present, the level of p65 phosphorylation was decreased. These results indicate that PCA prevents NF-κB activation induced by RANKL.

4. Discussion

Osteoclasts are derived from promonocytes in bone marrow hematopoietic stem cells and are mainly present in the periosteum. Osteoclasts are the only cells with bone absorption functions, initiating the process of bone remodeling and

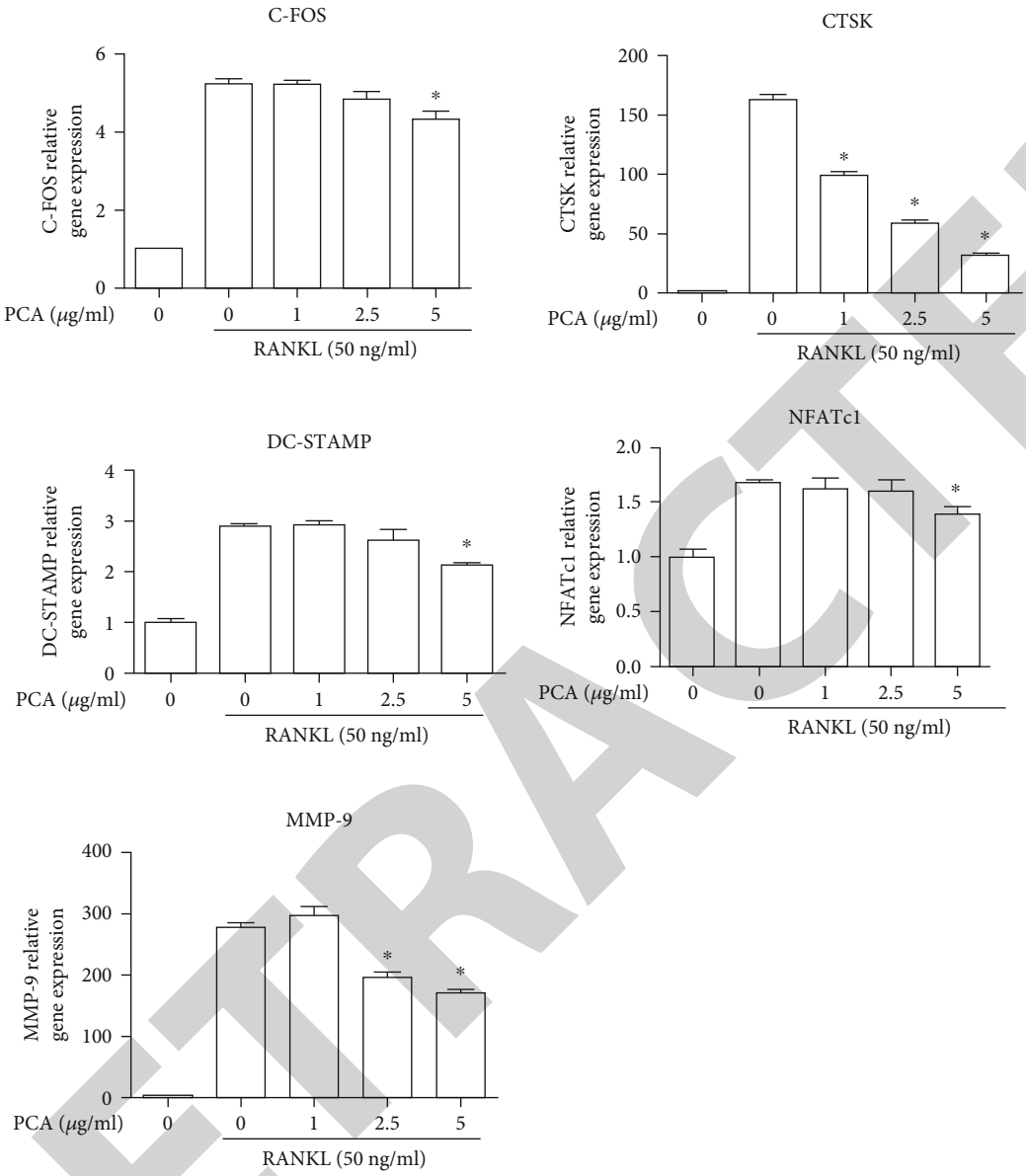


Figure 4: PCA inhibits mRNA expression in RANKL-stimulated osteoclasts. $^{*}P < 0.01$ compared with the RANKL group.

clearing old bone matrix [16]. Therefore, decreasing osteoclast differentiation or absorption may be an effective method to treat this type of pathological bone disease. This study was the first to show that PCA significantly inhibited the formation and function of RANKL-induced osteoclasts. In addition, PCA effectively inhibited the expression of osteoclast-associated marker genes stimulated by RANKL. Mechanistically, PCA suppressed RANKL-induced activation of the MAPK and NF-κB signaling pathways.

We used two types of osteoclast progenitors (BMMs and RAW264.7 cells) to study the roles of PCA in the induction and absorption function of osteoclasts [17]. Through the research, we discovered that RAW264.7 cells only proliferated in the absence of stimulation, and the proliferation of BMMs only gradually decreased. However, when stimulated with RANKL, these cells all differentiated into osteoclasts and showed absorption functions. This result is consistent with the literature [13, 14]. Furthermore, this effect was inhibited by PCA. These actions were neither related to the cytotoxic effects of PCA nor a cause of apoptosis. Therefore, we further examined the mechanism of PCA in osteoclasts.

A series of iconic factors are produced in the course of differentiation from precursor cells to mature osteoclasts and thus can be used as markers of osteoclasts and their differentiation stages. Cathepsin K (CTSK) is an amino acid protease expressed by osteoclasts. CTSK stimulates the expression of TRAP on osteoclasts *in vitro*; thus, CTSK is an important factor for determining osteoclast function [18]. The recombinant nuclear factor of activated T-cells, cytoplasmic 1(NFATc1) is the most important factor associated with NF-κB in the incipient stage of osteoclast formation. *In vitro* experiments confirmed that stem cells lacking

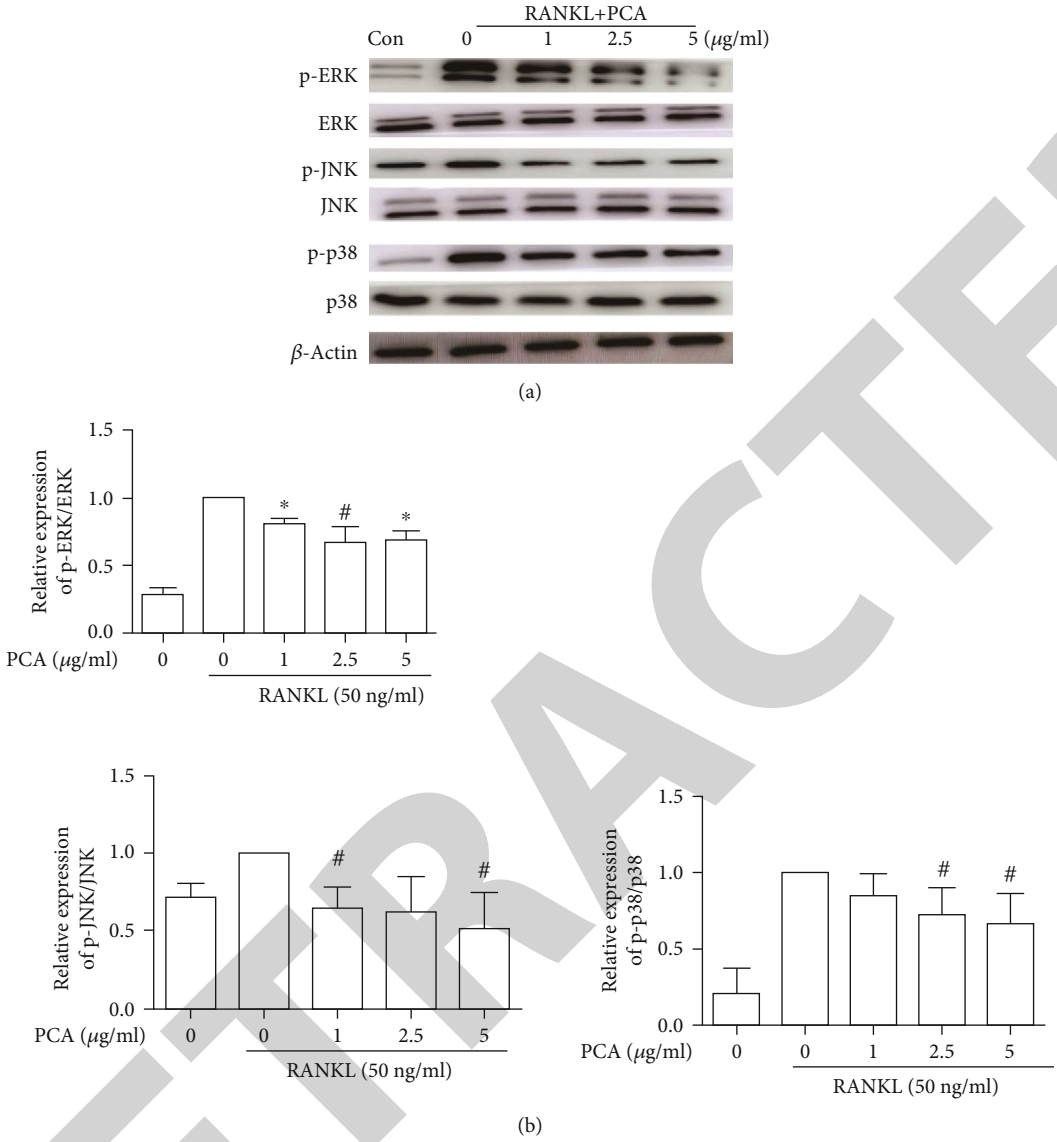


Figure 5: PCA suppresses the activation of the MAPK pathway. (a) Western blots showing protein levels. (b) The relative protein levels were compared with those in the group treated with RANKL alone. $^{*}P < 0.01$ and $^{\#}P < 0.05$ compared with the RANKL group.

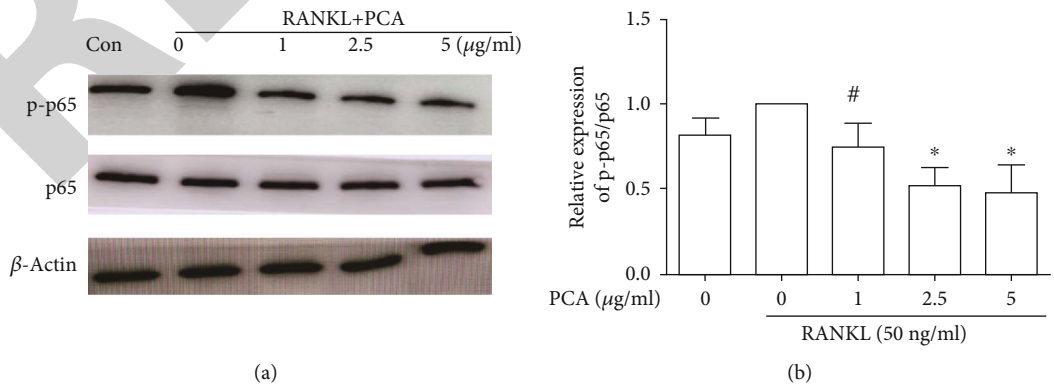


Figure 6: PCA prevents RANKL-induced NF-κB activation. (a) Western blots showing the levels of p-p65 and p65. (b) The protein levels were compared with those in the RANKL group. $^{*}P < 0.01$ and $^{\#}P < 0.05$ compared with the group treated with RANKL alone.

NFATc1 fail to differentiate into osteoclasts, but ectopic NFATc1 expression in the absence of RANKL also enables BMMs to differentiate into osteoclasts [19]. These results all suggest the importance of NFATc1. Matrix metalloproteinase (MMP-9), which is abundant in osteoclasts, is an important protease that stimulates bone absorption. Specifically, MMP-9 plays a crucial function in the active transcription of osteoclastogenic genes via the cleavage of the N-terminal tail of H3 under tight epigenetic regulation [20]. Dendritic cell-specific transmembrane protein (DC-STAMP), an important molecule involved in the fusion of mononuclear cells into osteoclasts, improves the absorption function of osteoclasts [21]. Cellular oncogene FOS (C-FOS) is a proto-oncogene associated with the development of osteosarcoma. During osteoclast differentiation, C-FOS is an indispensable downstream regulatory factor of RANKL, which accelerates the formation of osteoclasts by activating the downstream factor NFATc1 [22]. In the present study, PCA reduced the mRNA levels of these factors. Therefore, PCA may inhibit the formation of osteoclasts by downregulating the expression of these genes. However, these genes are downstream of the signaling pathway, and the mechanism by which PCA modulates these downstream genes remains unclear. Therefore, we detected the levels of proteins in the MAPK and NF-κB signaling pathways.

MAPK is an important signaling pathway involved in osteoclast differentiation [23]. This pathway mainly activates and regulates key downstream regulatory molecules associated with the formation of osteoclasts, including C-FOS and NFATc1 [24], and ultimately stimulates the differentiation of osteoclasts. The ERK, JNK, and p38 subfamilies are the main MAPK subfamilies. The binding of RANKL to RANK activates ERK through a series of reactions. ERK regulates the transcription of C-FOS after translocation to the nucleus, ultimately transforming mature macrophages into precursor osteoclasts [25]. Phosphorylated JNK activates AP-1 to promote the expression of MMP and ALP genes and stimulates the differentiation, survival, fusion, activation, and maturation of osteoclasts [26]. After the binding of RANKL to RANK, MEK6 phosphorylation is promoted and p38 is activated. Activated p38 is transferred from the cytoplasm to the nucleus, phosphorylating the transcription factor MITF and other proteins, ultimately promoting osteoclast differentiation [27]. In our study, PCA reduced the phosphorylation of ERK, JNK, and p38 in RAW264.7 cells treated with RANKL. The downregulation of MAPK expression induced by PCA was consistent with previously described results [28]. These results suggest that PCA may affect the fusion of precursor cells into osteoclasts by activating MAPK to suppress the expression of downstream factors.

NF-κB signaling is another indispensable pathway in osteoclast induction [29]. The trimer of RANK-RANKL-TRAF6 activates NF-κB. Then, NF-κB separates from I-κB and is quickly transported into the nucleus, binding to the promoters of the corresponding target genes and regulating osteoclast differentiation, maturation, and apoptosis [30, 31]. In the present study, PCA reduced the level of phosphorylated p65 in RANKL-induced osteoclasts. PCA suppressed RANKL-mediated osteoclast formation by inhibiting NF-κB activity. PCA possibly inhibits the transcription of NFATc1 by restraining the phosphorylation of NF-κB p65, thereby affecting the expression of downstream genes.

PCA is a water-soluble component of SM. SM and its other active ingredients are known to possess antiosteoporosis properties. Tanshinone is one of the most much-studied active components at present. Published studies have shown that tanshinone inhibits osteoclast activation and reduces bone absorption by destroying F-actin rings [32]. Cryptotanshinone also inhibits the formation of osteoclasts [33]. In our study, PCA reduced the number of TRAP-positive osteoclasts and the area of absorption pits, which proved that PCA inhibited the maturation and function of osteoclasts, suggesting that PCA has the potential to prevent bone loss and treat osteoporosis. Based on the findings of the present study, PCA exerts a synergistic effect with the abovementioned components of SM. This result further suggests that SM may also be an antiosteoporotic treatment that inhibits the maturation and function of osteoclasts.

Although some results were reported in our study, many other issues remain to be further explored. For example, how does PCA suppress the activation of MAPK and NF-κB? Does PCA bind to cognate receptors on the cell surface of precursor cells and subsequently inhibit cellular signaling pathways that lead to NF-κB and MAPK activation, or does PCA freely enter cells and directly inhibit NF-κB and MAPK activation? Other questions include the effectiveness of PCA *in vivo* and its pharmacokinetics, which all require further study.

In conclusion, our research shows that PCA suppresses osteoclast differentiation by inhibiting MAPK and NF-κB activation. Therefore, PCA potentially represents a cost-effective and alternative medicine for osteoporosis.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was funded by the National Key Research and Development Program of China (SQ2017YFSF080004), Natural Science Foundation of Shandong Province (ZR2017PH035), and Shandong Traditional Chinese Medicine Science and Technology Development Project (2019-0219).

References

[1] P. D. Miller, "Management of severe osteoporosis," *Expert Opinion on Pharmacotherapy*, vol. 17, no. 4, pp. 473–488, 2016.

[2] M. Tsukasaki and H. Takayanagi, "Osteoimmunology: evolving concepts in bone-immune interactions in health and

Hindawi
BioMed Research International
Volume 2023, Article ID 9789534, 1 page
https://doi.org/10.1155/2023/9789534

Retraction

Retracted: RUNX3 Expression Level Is Correlated with the Clinical and Pathological Characteristics in Endometrial Cancer: A Systematic Review and Meta-analysis

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Liu, Z. Hong, and S. Xi, "RUNX3 Expression Level Is Correlated with the Clinical and Pathological Characteristics in Endometrial Cancer: A Systematic Review and Meta-analysis," *BioMed Research International*, vol. 2021, Article ID 9995384, 7 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 9995384, 7 pages
https://doi.org/10.1155/2021/9995384

Review Article

RUNX3 Expression Level Is Correlated with the Clinical and Pathological Characteristics in Endometrial Cancer: A Systematic Review and Meta-analysis

Zhen Liu,[1] **Zhi-pan Hong**,[2] **and Shu-xue Xi**[3,4]

[1]*Department of Gynecology, Chifeng Municipal Hospital, Chifeng Clinical Medical School of Inner Mongolia Medical University, Chifeng, China*
[2]*Department of Tumor Surgery, Chifeng Municipal Hospital, Chifeng Clinical Medical School of Inner Mongolia Medical University, Chifeng, China*
[3]*Geneis (Beijing) Co. Ltd., Beijing 100102, China*
[4]*Qingdao Geneis Institute of Big Data Mining and Precision Medicine, Qingdao 266000, China*

Correspondence should be addressed to Zhi-pan Hong; hongzhipan1985@126.com

Received 2 April 2021; Revised 7 May 2021; Accepted 30 June 2021; Published 15 July 2021

Academic Editor: Min Tang

Human Runt-associated transcription factor 3 (RUNX3) plays an important role in the development and progression of endometrial cancer (EC). However, the clinical and pathological significance of RUNX3 in EC needs to be further studied. In order to clarify the clinical and pathological significance of RUNX3, a systematic review and meta-analysis was conducted in EC patients. Keywords RUNX3, endometrial cancer, and uterine cancer were searched in Cochrane Library, Web of Knowledge, PubMed, CBM, MEDLINE, and Chinese CNKI database for data up to Dec 31, 2018. References, abstracts, and meeting proceedings were manually searched in supplementary. Outcomes were various clinical and pathological features. The two reviewers performed the literature searching, data extracting, and method assessing independently. Meta-analysis was performed by RevMan5.3.0. A total of 563 EC patients were enrolled from eight studies. Meta-analysis results showed that the expression of RUNX3 has significant differences in these comparisons: lymph node (LN) metastasis vs. non-LN metastasis ($P = 0.26$), EC tissues vs. normal tissues ($P < 0.00001$), clinical stages I/II vs. II/IV ($P < 0.00001$), muscular infiltration < 1/2 vs. muscular infiltration ≥ 1/2 ($P < 0.00001$), and G1 vs. G2/G3 ($P < 0.00001$). The decreasing expression of RUNX3 is associated with poor TNM stage and muscular infiltration. It is indicated that RUNX3 was involved in the suppression effect of EC. However, further multicenter randomized controlled trials are needed considering the small sample size of the included trials.

1. Introduction

Endometrial cancer (EC) is a malignant tumour of uterine epithelial cells. EC is common in postmenopausal women over the age of 50, the incidence of which peaks at the age of 50–59. In developed countries, the incidence of EC occupies the first place in gynaecological tumours [1]. EC is a multifactor process in which the deletion of tumour suppressor genes and the activation of carcinogenesis are particularly important [2, 3]. Though there are many studies exploring cancer-associated biomarkers like genes and noncoding RNAs [4–7], the molecular mechanism of EC is largely unknown.

In recent years, the role of human Runt-associated transcription factor 3 (RUNX3) in the development and progression of EC has attracted increasing attention [8]. RUNX3 is located in human chromosome 1p36.11, and its total length is approximately 67 kb [9]. RUNX3 is an important tumour candidate suppressor gene that is closely related to gastric cancer, breast cancer, gallbladder cancer, and other malignant tumours [10–16]. However, despite numerous patients

with EC worldwide, RUNX3 in EC has not been definitively reported, especially on the correlation of its overexpression and its clinical significance in endometrial cancer. Here, we performed a systematic literature review and meta-analysis to determine the association between RUNX3 and the clinicopathological characteristics of EC.

2. Materials and Methods

2.1. Literature Information. A total of 412 articles were initially identified using the below search strategy. Based on titles and abstracts, 378 out of the 412 were excluded due to non-endometrial-related studies, nonoriginal articles (e.g., review and letter), and duplicate studies. After reading the full texts, we excluded 21 data that could not be extracted due to non-RUNX3-related studies, nonimmunohistochemical SP method, and disunited positive criteria. Eventually, 8 studies (2 in English and 6 in Chinese) were included in the meta-analysis (Figure 1) [17–22].

2.2. Study Characteristics. The eight studies included were based on the Asian population, 6 of which were conducted in China, 1 in Japan, and 1 in Korea, involving a total of 827 patients. The overall median age ranged at 28–72 years. The total positive rate of low RUNX3 expression by immunohistochemistry (IHC) is 68.07% from eight studies (from 26.4% to 82.7%). All detected specimens were derived from EC tissues by either biopsy or surgical resection and were confirmed by IHC on membrane protein level. The studies were divided into five groups according to the following criteria: (1) low RUNX3 expression in EC tissues or normal endometrial cancer, (2) low RUNX3 expression in positive and negative lymph node (LN) metastasis of EC tissues, (3) low RUNX3 expression in different clinical stages of EC tissues, (4) low RUNX3 expression in different infiltration degree of EC tissues, and (5) low RUNX3 expression in different pathological stages of EC tissues (Table 1).

2.3. Literature Search Strategy. The electronic databases Cochrane Library, PubMed, MEDLINE, Web of Knowledge, and Chinese CNKI and CBM were comprehensively searched for data dated up to Dec 31, 2018. The search string of PubMed was ("RUNX3" [Title/Abstract]) OR "RUNT" [Title/Abstract] AND (uterus carcinoma [Mesh Terms]) AND "carcinoma" [Mesh Terms] OR "endometrial cancer" [Title/Abstract]). The reference lists of relative articles were also screened to further identify potential studies.

2.4. Criteria for Selecting Data. We selected data with the following criteria to be included in this study: (1) the data include both case group and control samples, (2) the data is published together with some paper with full text, (3) standard methods like IHC were accompanied with the diagnosis of EC, (4) the amount and information of data were enough for an accurate evaluation on the hazard ratio for survival with 95% CI, and (5) RUNX3 expression study was conducted on primary EC tissue (not serum or other kinds of specimen). If there are some duplicated studies, we prefer the most informative or most recent one.

2.5. Data Extraction. The data were extracted similar to Li et al. [23].

2.6. Statistical Analysis. We used Cochrane RevMan 5.3.0 (the Cochrane Collaboration, Copenhagen) to perform statistical analyses. Specifically, we use risk ratio (RR) to present dichotomous data and mean difference with 95% CI to present continuous variables. The heterogeneity of the data was tested by the Chi-square test, and a data is deemed significant if $P < 0.1$. We used I-square to estimate total variation. We drew a funnel plot to assess the potential publication bias and used pooled estimates of RRs and their 95% CI to compare dichotomous measures. The significance threshold was set to be $P < 0.05$.

3. Results

3.1. EC Group vs. Control Group. Five studies [17, 18, 20–22] reported low RUNX3 expression in the EC group (EC tissues) and control group (normal gastric tissues). Meta-analysis via random effect model indicated that the expression rate of RUNX3 in the EC group was higher than that in the control group. The difference between the two groups was statistically significant (RR = 0.5, 95% CI (0.44, 0.57), $P < 0.00001$, Figure 2).

3.2. LN Metastasis of EC Tissues: Positive Group vs. Negative Group. Five studies [18–22] reported RUNX3 expression in positive and negative LN metastasis of EC tissues. Meta-analysis via random effect model showed that low RUNX3 expression rate in the positive group (LN+) was higher than that in the negative group (LN-). The difference between the two groups was not statistically significant (RR = 0.79, 95% CI (0.52, 1.19), $P = 0.26$, Figure 3).

3.3. TNM Stage of EC Tissues: III–IV Stage Groups vs. I–II Stage Groups. Five studies [18–22] reported the expression of RUNX3 in the II–IV stage groups and I–II stage groups of EC tissues. Meta-analysis via random effect model showed that low RUNX3 expression rate in the III–IV stages was higher than that in the I–II stage groups. The difference between the two groups was statistically significant (RR = 3.55, 95% CI (2.11, 5.98), $P < 0.00001$, Figure 4).

3.4. Muscular Infiltration of Endometrial Cancer: <1/2 Group vs. ≥1/2 Group. Five studies [18–22] reported low RUNX3 expression in the muscular infiltration < 1/2 group and ≥1/2 group in EC tissues. Meta-analysis via random effect model showed that low RUNX3 expression rate in the muscular infiltration of ≥1/2 group was higher than that in the <1/2 group. The difference between the two groups was statistically significant (RR = 2.38, 95% CI (1.71, 3.31), $P < 0.00001$, Figure 5).

3.5. Pathological Classification of Endometrial Cancer: G1 Group vs. G2–G3 Group. Five studies [18–22] reported low RUNX3 expression in the pathological classification of G1 group and G2–G3 groups in EC tissues. Meta-analysis via random effect model showed that low RUNX3 expression rate in the G1 group was higher than that in the G2–G3 groups. The

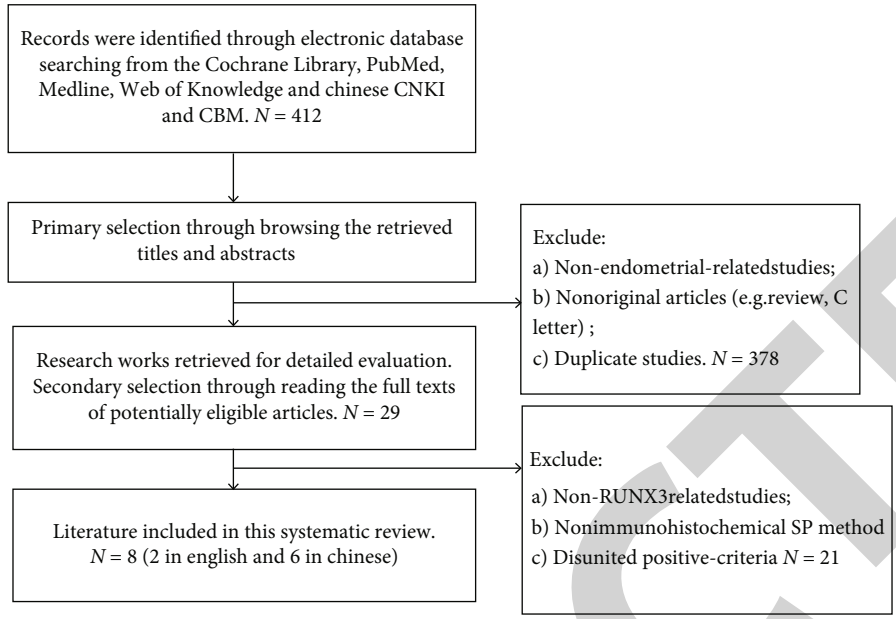


Figure 1: Flow chart for the selection of studies.

Table 1: General characteristics of included studies.

Studies	Year	Country	Cases (n)	Age	Method	RUNX3 expression rate (%)	Group
Dong	2018	Korea	61	—	IHC	49.1%	(2) (3) (4)
Zhang	2013	China	120	—	IHC	66.6%	(1) (2) (3) (4) (5)
Liu	2013	China	117	—	IHC	73.5%	(1) (2) (3) (4) (5)
Zhang	2012	China	100	25–81	IHC	80%	(1) (2) (3) (4) (5)
Guo	2011	China	162	38–72	IHC	26.4%	(1) (2) (3) (4) (5)
He	2010	China	98	28–72	IHC	79.5%	(1) (2)
Feng	2010	China	140		IHC	42.8%	(1) (2) (3) (4) (5)
Tatsuo	2008	Japan	29		IHC	82.75	(1)

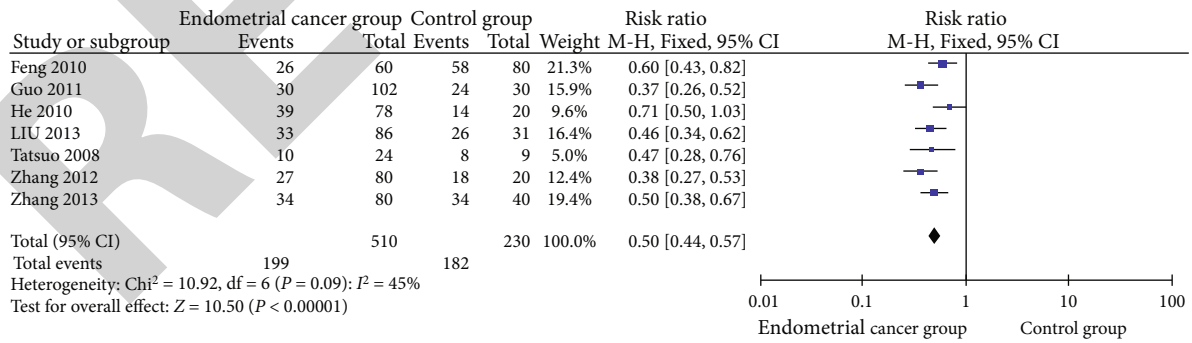


Figure 2: Meta-analysis of RUNX3 expression in the endometrial cancer and control group.

difference between the two groups was statistically significant (RR = 3.19, 95% CI (2.25, 4.51), $P < 0.00001$, Figure 6).

3.6. Publication Bias. A funnel plot of every two groups in comparison above was applied with RR as the x-axis and SE (RR) as the y-axis. The plot was symmetric, suggesting that publication bias was minimal (Figure 7).

4. Discussion

The RUNX3 gene is a member of the Runt family, which includes RUNX1, RUNX2, and RUNX3. RUNX3 is located on human chromosome 1p36.1. The total length of the gene is approximately 67 kb and includes two promoters and six exons [24]. Given the high GC content of RUNX3

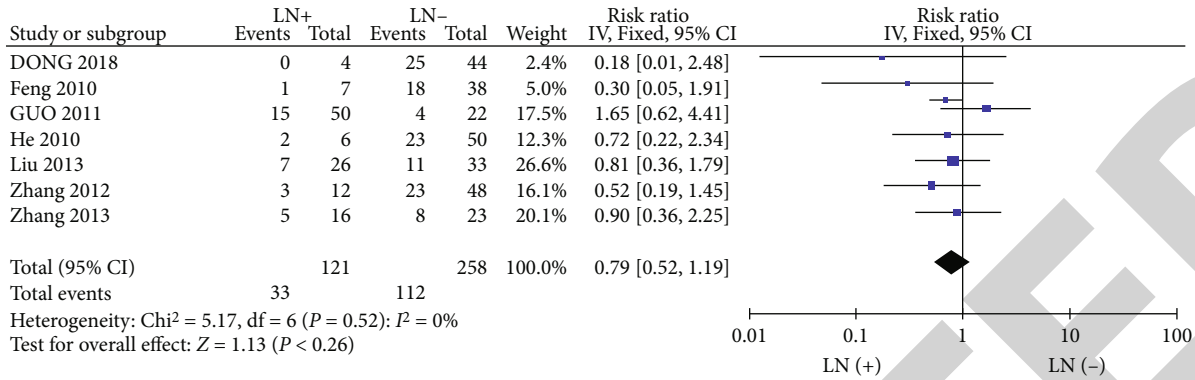


Figure 3: Meta-analysis of RUNX3 expression in LN (+) and LN (-) endometrial cancer.

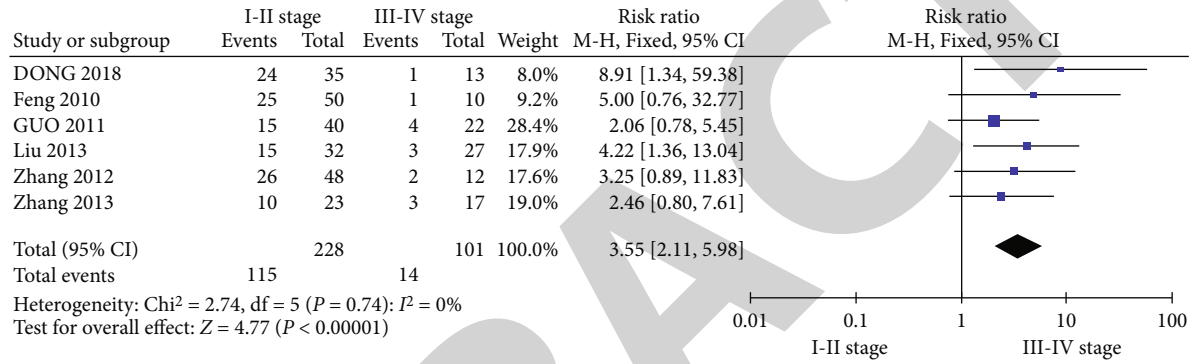


Figure 4: Meta-analysis of low RUNX3 expression in the II–IV stage groups and I–II stage groups.

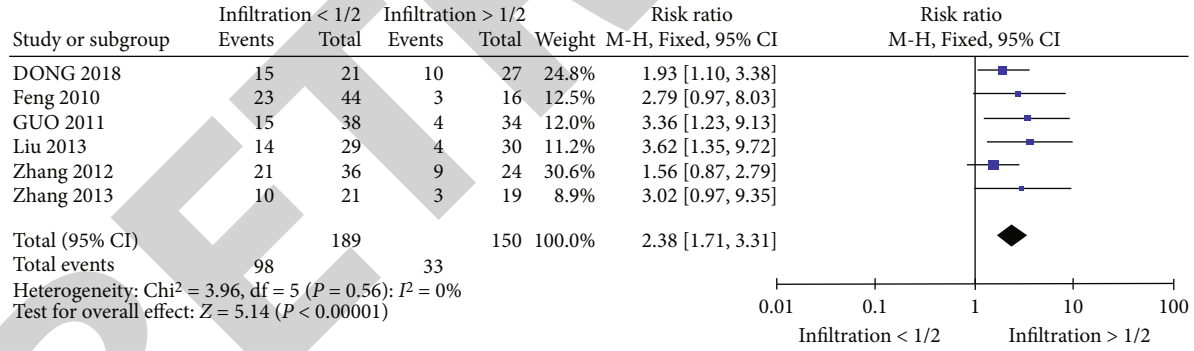


Figure 5: Meta-analysis of low RUNX3 expression in the muscular infiltration < 1/2 group and ≥1/2 group.

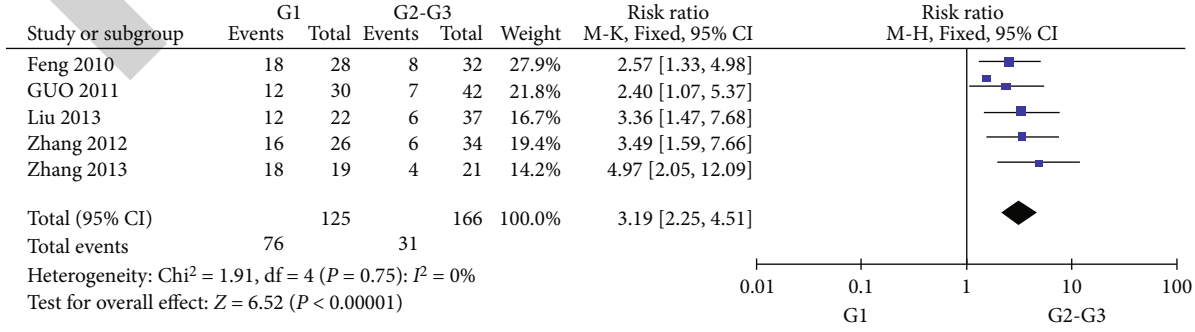


Figure 6: Meta-analysis of low RUNX3 expression in the G1 group and G2–G3 groups.

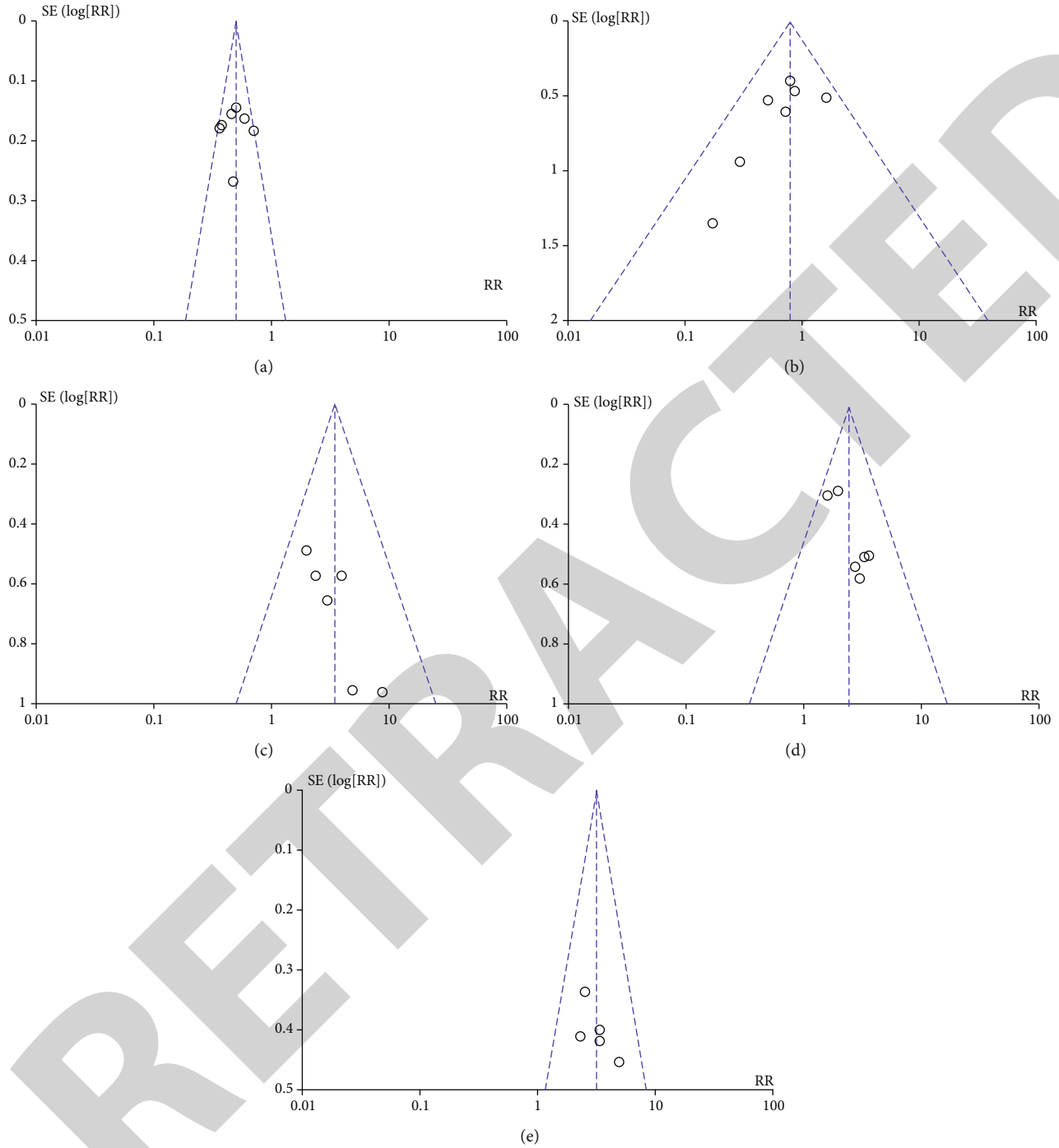


Figure 7: Funnel plot: (a) low RUNX3 expression in the EC group and control group; (b) low RUNX3 expression in the LN (+) and LN (-) EC group; (c) low RUNX3 expression in the III–IV stage groups and I–II stage groups; (d) low RUNX3 expression in the muscular infiltration of the EC < 1/2 group and ≥1/2 group; (e) low RUNX3 expression in the G1 group and G2–G3 groups in endometrial cancer.

promoter, which accounts for approximately 64%, methylation is likely to occur [25]. RUNX3 regulates cell growth and apoptosis and exhibits important and complex transcriptional effects on cell signalling and other biological effects. The lack of RUNX3 expression is closely related to the occurrence and development of various human malignant tumours [26, 27]. RUNX3's anticancer mechanism is closely related to the TGF-β pathway, which functions as a growth suppressor and apoptosis-inducing factor. TGF-β is a growth factor that inhibits many developmental and physiological processes. A disorder of its signal pathway is involved in the occurrence and development of many tumours [18, 28]. The RUNX3 expression is abnormal in various gastrointestinal tumours, such as gastric cancer [29]. Therefore, RUNX3 is currently considered as a tumour suppressor gene. The role of RUNX3 in endometrial carcinogenesis has recently attracted increasing attention from scholars worldwide [30].

Given the biological properties of RUNX3, some scholars believe that studying the action pathway and mechanism of RUNX3 can provide new targets and insights into the treatment of uterine malignant tumours in the clinic [21]. Despite various basic and clinical studies on RUNX3 and endometrial cancer, no consensus opinion has been reached in detail. In the current study, we systematically reviewed the correlation between low expression and the clinical significance of RUNX3 in endometrial cancer. We found that low RUNX3 expression rate in the EC group is higher than that in the control group by meta-analysis. Low RUNX3 expression was not related to LN metastasis but is associated to TNM stage, degree of muscular infiltration, and pathological degree. In conclusion, low RUNX3 expression and its clinical-pathological features are closely related in endometrial cancer. RUNX3 may play a critical role in the pathophysiology, integration, and complementation of endometrial cancer. Nevertheless, the translational potentials of these findings warrant further investigation.

Although this systematic review is aimed at providing the best possible estimate of the correlation between the low expression and clinical significance of RUNX3 in endometrial cancer, it suffers from several limitations. Firstly, the number of studies and patients included in the current meta-analysis is relatively small. Secondly, all the studies were based on the Asian population, including 6, 1, and 1 from China, Korea, and Japan, respectively. Significant differences, such as those in aetiology, biological features, clinical types, and prognosis in the risk of EC, exist among various ethnic groups within a given geographical area. Given the lack of statistics on the Western population, data on the low expression rate of RUNX3 in Western patients are unavailable. Second, EC is quite heterogeneous; the sequencing of single cell RNA data of EC patients might contribute to understanding in-depth mechanisms of EC [31–34]. Finally, it will be helpful for clinical usage to identify drugs targeting EC through drug repositioning or drugs targeting this gene [35–37]. In virtue of several limitations and inconsistent combined results, further large, well-designed prospective cohort studies with good exposure assessment are warranted to confirm the findings from our study and provide further evidence.

Data Availability

The analyzed data are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Zhi-pan Hong and Zhen Liu designed the overall study and prepared the manuscript. Shu-xue Xi reviewed the manuscript. All authors read and approved the final version. Zhen Liu and Zhi-pan Hong contributed equally to this work.

Acknowledgments

What is more, we followed the methods of Hu et al. [38].

References

[1] D. L. Kolbe, J. A. DeLoia, P. Porter-Gill et al., "Differential analysis of ovarian and endometrial cancers identifies a methylator phenotype," *PLoS One*, vol. 7, no. 3, article e32941, 2012.

[2] M. Wang and M. A. Zhihong, "Multivariate analysis of prognostic factors of endometrial carcinoma by COX proportional hazard regression model," *Chinese Journal of Clinical Oncology*, vol. 38, no. 6, pp. 344–348, 2011.

[3] E. J. Crosbie, M. Zwahlen, H. C. Kitchener, M. Egger, and A. G. Renehan, "Body mass index, hormone replacement therapy, and endometrial cancer risk: a meta-analysis," *Cancer Epidemiology Biomarkers & Prevention*, vol. 19, no. 12, pp. 3119–3130, 2010.

[4] W. Li, S. Wang, J. Xu, G. Mao, G. Tian, and J. Yang, "Inferring latent disease-lncRNA associations by faster matrix completion on a heterogeneous network," *Frontiers in Genetics*, vol. 10, 2019.

[5] X. Li, Y. Lin, C. Gu, and J. Yang, "FCMDAP: using miRNA family and cluster information to improve the prediction accuracy of disease related miRNAs," *BMC Systems Biology*, vol. 13, Supplement 2, 2019.

[6] X. Xiao, W. Zhu, B. Liao et al., "BPLLDA: predicting lncRNA-disease associations based on simple paths with limited lengths in a heterogeneous network," *Frontiers in Genetics*, vol. 9, 2018.

[7] M. Lu, X. Xu, B. Xi et al., "Molecular network-based identification of competing endogenous RNAs in thyroid carcinoma," *Genes*, vol. 9, no. 1, p. 44, 2018.

[8] S. Zhang, Q. Hao, and H. I. Sun, "Correlation studies between MTA1 protein expression and endometrial carcinoma," *Clinical Oncology and Cancer Research*, vol. 39, pp. 208–211, 2012.

[9] Q.-L. Li, K. Ito, C. Sakakura et al., "Causal Relationship between the Loss of _RUNX3_ Expression and Gastric Cancer," *Cell*, vol. 109, no. 1, pp. 113–124, 2002.

[10] Q. Lau, E. Raja, M. Salto-Tellez et al., "RUNX3 is frequently inactivated by dual mechanisms of protein mislocalization and promoter hypermethylation in breast cancer," *Cancer Research*, vol. 66, no. 13, pp. 6512–6520, 2006.

[11] T. Takahashi, N. Shivapurkar, E. Riquelme et al., "Aberrant promoter hypermethylation of multiple genes in gallbladder carcinoma and chronic cholecystitis," *Clinical Cancer Research*, vol. 10, no. 18, pp. 6126–6133, 2004.

[12] S. Bae and J.-K. Choi, "Tumor suppressor activity of RUNX3," *Oncogene*, vol. 23, no. 24, pp. 4336–4340, 2004.

[13] B. He, C. Dai, J. Lang et al., "A machine learning framework to trace tumor tissue-of-origin of 13 types of cancer based on DNA somatic mutation," *Biochimica et Biophysica Acta - Molecular Basis of Disease*, vol. 1866, no. 11, article 165916, 2020.

[14] B. He, J. Lang, B. Wang et al., "TOOme: a novel computational framework to infer cancer tissue-of-origin by integrating both gene mutation and expression," *Frontiers in Bioengineering and Biotechnology*, vol. 8, 2020.

[15] Z. Song, X. Chen, Y. Shi et al., "Evaluating the potential of T cell receptor repertoires in predicting the prognosis of

Hindawi
BioMed Research International
Volume 2023, Article ID 9763086, 1 page
https://doi.org/10.1155/2023/9763086

Retraction

Retracted: Retinal OCT Texture Analysis for Differentiating Healthy Controls from Multiple Sclerosis (MS) with/without Optic Neuritis

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. D. Tazarjani, Z. Amini, R. Kafieh, F. Ashtari, and E. Sadeghi, "Retinal OCT Texture Analysis for Differentiating Healthy Controls from Multiple Sclerosis (MS) with/without Optic Neuritis," *BioMed Research International*, vol. 2021, Article ID 5579018, 13 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 5579018, 13 pages
https://doi.org/10.1155/2021/5579018

Research Article

Retinal OCT Texture Analysis for Differentiating Healthy Controls from Multiple Sclerosis (MS) with/without Optic Neuritis

Hamidreza Dehghan Tazarjani,[1] Zahra Amini,[1] Rahele Kafieh,[1] Fereshteh Ashtari,[2] and Erfan Sadeghi[3]

[1]*Medical Image and Signal Processing Research Center, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Iran*
[2]*Isfahan Neurosciences Research Center, Isfahan University of Medical Sciences, Isfahan, Iran*
[3]*Department of Biostatistics and Epidemiology, Faculty of Health, Isfahan University of Medical Sciences, Isfahan, Iran*

Correspondence should be addressed to Zahra Amini; zahraamini64@yahoo.com.au

Received 12 February 2021; Revised 20 April 2021; Accepted 22 June 2021; Published 10 July 2021

Academic Editor: Min Tang

Multiple sclerosis (MS) is an inflammatory disease damaging the myelin sheath in the central and peripheral nervous system in the brain and spinal cord. Optic Neuritis (ON) is one of the most prevalent ocular demonstrations of MS. The current diagnosis protocol for MS is MRI, but newer modalities like Optical Coherence Tomography (OCT) are already of interest in early detection and progression analysis. OCT reveals the symptoms of MS in the Central Nervous System (CNS) through cross-sectional images from neural retinal layers. Previous works on OCT were mostly focused on the thickness of retinal layers; however, texture features seem also to have information in this regard. In this research, we introduce a new pipeline that constructs layer-stacked (LS) images containing data from each specific layer. A variety of texture features are then extracted from LS images to differentiate between healthy controls and ON/None-ON MS cases. Furthermore, the definition of texture extraction methods is tailored for this application. After performing a vast survey on available texture analysis methods, a treasury of powerful features is collected in this paper. As a primary work, this paper shows the ability of such features in the diagnosis of HC and MS (ON and None-ON) cases. Our findings show that the texture features are powerful to diagnose MS cases. Furthermore, adding information of conventional thickness values to texture features improves considerably the discrimination between most of the target groups including HC vs. MS, HC vs. MS-None-ON, and HC vs. MS-ON.

1. Introduction

Multiple sclerosis (MS) is an inflammatory disease damaging the myelin sheath in the central and peripheral nervous system in the brain and spinal cord. This disease causes the immune system to attack one or more proteins of the myelin structure and disrupts the ability of the nervous system to communicate and therefore brings about many physical signs and symptoms [1]. Those suffering from MS show neurological symptoms including disorders in the autonomic, visual, motor, and sensory nervous system [2]. Optic Neuritis (ON) is a common eye problem where inflammation or demyelination affects the optic nerve. It occurs when inflammation damages the optic nerve, a bundle of nerve fibers that transmits visual information from the eye to the brain. Signs and symptoms of ON can be the first indication of MS, or they can occur later in the course of MS. Not everyone who experiences ON goes on to develop further symptoms of MS, but a significant proportion does [3].

The current diagnosis protocol for MS is Magnetic Resonance Imaging (MRI); however, researchers are already looking for substitute methods to overcome MRI limitations like high cost, late-stage diagnosis, and inaccurate signs due to aging rather than MS [4]. The effects of MS on the Central Nervous System (CNS) make the retinal nerve fiber layer (RNFL) a proper candidate for being imaged instead of brain MRIs. The thickness of RNFL can be used to assess the existence of any damage in the CNS. Moreover, RNFL is

considered as one of the main retinal layers. The role of the remaining layers is not exactly known in the case of MS and needs more investigations.

Optical Coherence Tomography (OCT) is a noninvasive imaging modality to take cross-sectional images of biological tissues. Retinal OCT provides information on symptoms of many eye diseases such as macular degeneration, glaucoma, and diabetic retinopathy and helps ophthalmologists to diagnose and treat such diseases in a timely manner [5–7]. Parisi et al.'s study in 1998 on the diagnosis of MS using retinal OCT was the first work in this field [8]. He investigated whether there is a relationship between RNFL thickness and visual pathway function in patients with MS. Since then, a great deal of research has been done on thickness changes in different retinal layers and the possibility of their use for diagnosing MS. Petzold et al. in 2017 prepared a survey covering this topic and reviewed 110 articles from 1991 to 2016 and provided a good overview of the subject [9].

In recent years, in the field of retinal OCT image processing, much attention has been paid to extracting and using texture features of layers while these types of features have not yet been used widely for MS diagnosis and there are few works addressing this issue. As an example, Varga et al. in 2015 had a study investigating the differences in texture descriptors and optical properties of retinal tissue layers in patients with MS and evaluated their usefulness in the detection of neurodegenerative changes using OCT image segmentation [10]. The term texture in image processing and machine vision refers to the amount, type, and distribution of pixel brightness throughout the image along with the texture of the image [11]. Researchers have defined it as "A texture area in an image can be constructed with an irregular and varied spatial distribution of the intensity of the brightness or color [12]." In this regard, four general categories named statistical, structural, signal processing-based, and model-based features are usually used [11]. In this study, we want to examine the texture of OCT images, and we suspect that changes in the texture of the layers must occur before the thickness changes. It seems that the deterioration of axons in the retinal nerve fiber layer and changes in the texture layers can be determined by the noninvasive OCT method, making them possible to be used as a complementary diagnostic tool in addition to the existing methods for early detection of recurrent MS-ON and MS-None-ON [13].

Here is an overview of the literature investigating texture features in OCT images. In 2007, Baroni et al. investigated the possibility of discriminating retinal OCT image layers in texture processing using Grey-Level Cooccurrence Matrix (GLCM) feature extraction [14]. In 2014, Anantrasirichai et al. presented a new method for extracting the texture of OCT retinal images in glaucoma [15]. In 2018, Sawyer et al. examined the possibility of using texture analysis to classify ovarian OCT images [16]. In 2019, Nunes et al. used texture analysis of OCT data to define new biomarkers for MS, of course, only on one specific retinal layer [17].

The rest of this paper is as follows. The proposed method for texture extraction of retinal OCT layers is described in Section 2. The performance of the method is evaluated and discussed in Section 3. Finally, Section 4 presents the conclusions of the study.

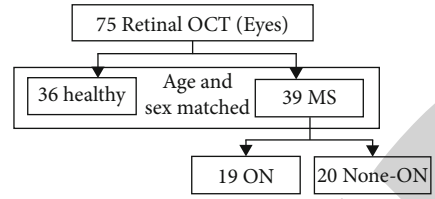


Figure 1: Detailed structure of the data.

2. Material and Method

2.1. Database. The data in this study is obtained from Spectralis Heidelberg HRA+OCT device in Faiz Hospital and Sadra Ophthalmology Center, Isfahan, Iran. The size of each B-scan is 496×480 pixels. For some subjects, data contains 19 B-scans, and for others, it includes 25 B-scans. OCT data includes 36 health control (HC) eyes and 39 patients suffering from MS (20 eyes suffering from MS with no history of ON (MS-None-ON) and 19 eyes suffering from MS with a history of ON (MS-ON)). HC and patient populations have matched gender and age approximately. A summary data flow diagram is presented in Figure 1.

2.2. Algorithm Flow. The workflow of the proposed method is shown in Figure 2. The first step is the preprocessing block in which the retinal delineation [18] is used to extract the layers. In the second block, layer-stacked (LS) images are created by stacking each specific layer from all B-scans of one subject. The third block is applied for masking the images as input to the next feature extraction block. Five different groups of texture features are utilized in this step. In the following, the most effective features are selected based on p value for distinguishing HC, MS-ON, and MS-None-ON population from retinal OCT layers around the fovea. Finally, in the last step, a classification between HC and abnormal population is performed. Each block of the proposed algorithm flow is elaborated below.

The sample output of preprocessing block is shown in Figure 3. To construct layer-stacked images, we consider that data for each subject consists of a number of B-scans, and each B-scan contains 10 layers, locations of which are obtained in preprocessing step. Accordingly, we construct 10 layer-stacked images by cutting and stacking each individual layer from all B-scans of one subject (Figures 4 and 5). A sample of layer-stacked images is demonstrated in Figure 6.

During texture calculation, boundary points in layer-stacked images have synthetic contrast which may fool the feature extraction method and lead to incorrect and outlier values. To solve this problem, an eliminating mask is developed to ignore pixels located on both sides of each individual layer.

Feature extraction is then performed on masked layer-stacked images. The features used in our work are GLCM, Local Binary Pattern (LBP), Local Directional Pattern (LDP), Local Optimal Oriented Pattern (LOOP), and fractal dimension. Finally, discriminant features are fed into

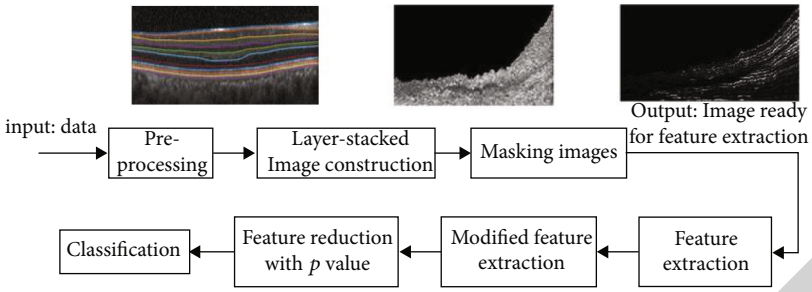


Figure 2: Algorithm flow of the proposed method.

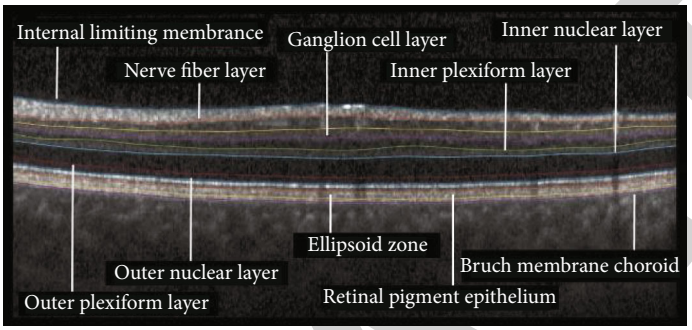


Figure 3: Interretinal layers (a sample output of preprocessing block).

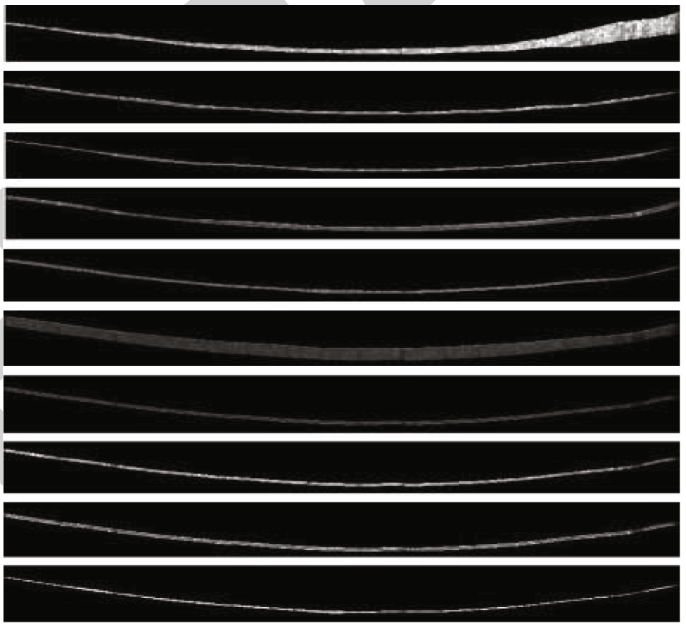


Figure 4: Sample of individual layers in one B-scan.

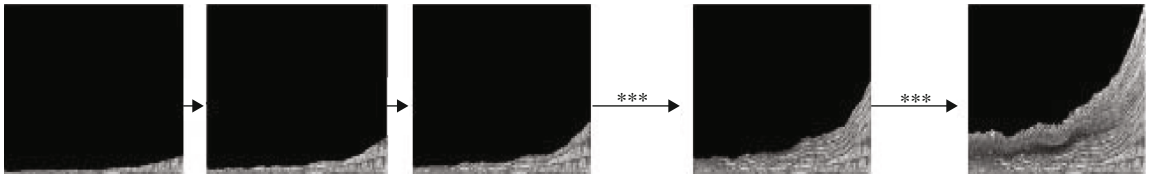


Figure 5: Construction process for layer-stacked (LS) images.

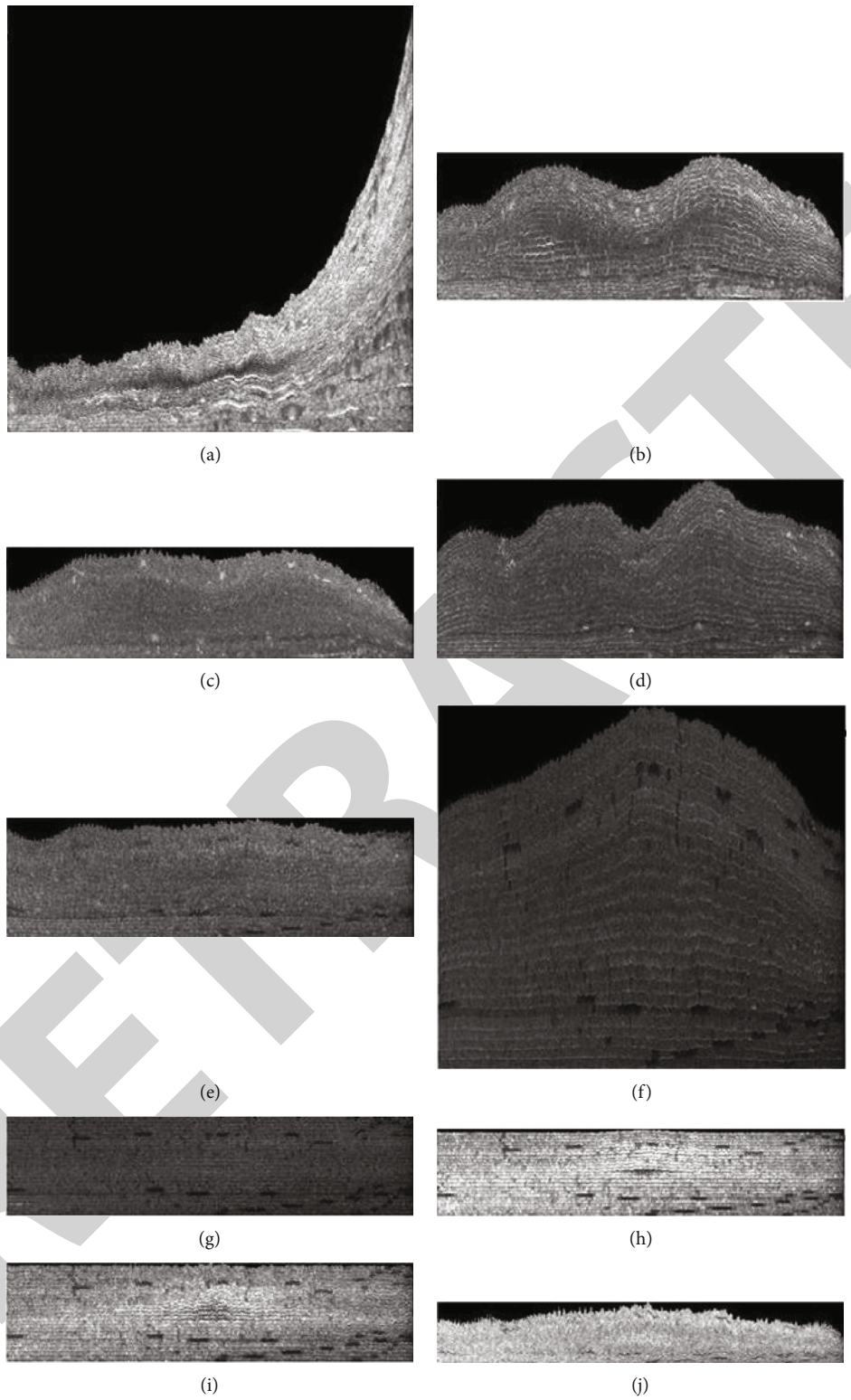


Figure 6: Layer-stacked (LS) images corresponding to each retinal layer. (a) First layers of all B-scans. (b) Second layers of all B-scans. (c) Third layers of all B-scans. (d) Fourth layers of all B-scans. (e) Fifth layers of all B-scans. (f) Sixth layers of all B-scans. (g) Seventh layers of all B-scans. (h) Eighth layers of all B-scans. (i) Ninth layers of all B-scans. (j) Tenth layers of all B-scans.

Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) classifiers for differentiating between HC and MS cases.

2.3. Texture Feature Extraction. Investigating texture features is an efficient way to characterize various properties, such as structure, orientation, roughness, smoothness, or regularity

of an image. Extracting features from masked layer-stacked images, we apply two categories of texture features including original and modified features.

2.3.1. Original Features. Different texture analysis methods are utilized in this research and elaborated in the next subsections. A set of features are then extracted according to Table 1.

(1) Grey-Level Cooccurrence Matrix. GLCM describes the spatial relationship between each intensity tone by considering changes between grey levels i and j at a particular displacement distance d and at a particular angle θ [15]. Here, we use a 256 quantization level and the distance is selected as one pixel with four distinct orientations (0, 45, 90, and 135 degrees). Furthermore, conditions of those pixels with 180 degrees in difference are considered to be the same.

(2) Local Binary Pattern. LBP is a method for describing the texture characteristics introduced in 1990 [19]. LBP compares the intensity of each pixel with neighboring pixels and determines the output value based on equation (1), where P is the number of neighboring points that are chosen, i.e., 8, i_p is the intensity of the neighborhood points, and i_c is the intensity of the central point. LBP_P calculates the output of LBP for P neighboring points.

$$\mathrm{LBP}_P(x_c, y_c) = \sum_{p=0}^{P-1} \mathrm{sign}\left(i_p - i_c\right)2^p, \qquad \mathrm{Sign}(x) = \begin{cases} 1, & \text{if } x \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

(3) Local Directional Pattern. A more robust to noise modified version of LBP is LDP which computes directional components for each pixel with Kirsch kernels and provides a measure of the strength of intensity variation in those directions [20]. For each central pixel located at (x_c, y_c) with intensity i_c, eight rotated versions of the Kirsch edge detector should be applied on neighboring pixels with intensities i_n $n = 0, 1, \cdots, 7$. Eight corresponding responses of the Kirsch masks are m_n $n = 0, 1, \cdots, 7$. m_k is the k^{th} highest Kirsch activation, and all the neighboring pixels having Kirsch response higher than m_k are assigned 1, and others 0. Then, the LDP value for the pixel (x_c, y_c) is given by

$$\mathrm{LDP}_k(x_c, y_c) = \sum_{n=0}^{7} \mathrm{sign}(m_n - m_k) \times 2^n. \tag{2}$$

(4) Local Oriented Optimization Pattern. LOOP offers a nonlinear combination of LBP and LDP that overcomes their individual problems while maintaining the strengths of each. Compared to LDP, LOOP assigns an exponential weight w_n to each of neighboring pixels. w_n is a digit between 0 and 7, according to the rank of the magnitude of m_n among the 8 Kirsch mask outputs [21]. The value of the LOOP in (x_c, y_c) is given by

$$\mathrm{LOOP}(x_c, y_c) = \sum_{n=0}^{7} \mathrm{sign}(i_n - i_c) \bullet 2^{w_n}. \tag{3}$$

(5) Fractal Analysis. Images with self-similarity characteristics are called fractal. The box-counting analysis is an appropriate method of fractal dimension estimation for images with or without self-similarity [22]. We have a basic equation for calculating fractal dimension given by equation (4), in which N is the number of boxes that cover the pattern, and r is the magnification or inverse value of the box size.

$$D = \frac{\log(N)}{\log(r)}. \tag{4}$$

A higher slope means that the object is more fractal, i.e., reduction in the size of the box reveals more complexity. The lower slope means that the object is closer to the straight line, i.e., less fractal, and the amount of details does not increase rapidly with increasing magnification.

2.3.2. Modified Features. Inserting zero values by masking layer-stacked images (third block in Figure 2) causes unwanted strip artifact. In order to solve this problem, we modify the output of the abovementioned texture analysis methods, to extract more accurate features. A list of used abbreviations in this paper and their explanations is shown in Table 2.

For GLCM, the first row and column of the output matrix (which represent unwanted zero pixels) are eliminated. The GLCM features listed in Table 1 can then be calculated from

$$\text{Energy (ene)} = \sum_{i=1}^{N}\sum_{j=1}^{N} p(i.j)^2,$$

$$\text{Entropy (ent)} = -\sum_{i=1}^{N}\sum_{j=1}^{N} p(i.j)\log p(i.j),$$

$$\text{Contrast (con)} = \sum_{i=1}^{N}\sum_{j=1}^{N} (i-j)^2 p(i.j),$$

$$\text{Homogeneity (hom)} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{p(i.j)}{1+(i-j)^2},$$

$$\text{Correlation (cor)} = \sum_{i=1}^{N}\sum_{j=1}^{N} \left(\frac{i-\mu_x}{\sigma_x}\right)\left(\frac{j-\mu_y}{\sigma_y}\right) p(i.j),$$

$$\text{Sum of squares} = \sum_{i=1}^{N}\sum_{j=1}^{N} (i-\mu)^2 p(i.j),$$

$$\text{Cluster shade} = \sum_{i=1}^{N}\sum_{j=1}^{N} (i+j-2\mu)^4 p(i.j),$$

Table 1: List of utilized features.

Texture analysis method	Features	Description	Texture analysis method	Features	Description
GLCM	Energy	Provides the sum of squared elements in the GLCM. It has values between 0 and 1	GLCM	Difference variance	Measures the dispersion (with regard to the mean) of the grey-level difference distribution of the image
	Entropy	Measure of randomness that can be used to characterize the texture of an image		Difference entropy	Measures the disorder related to the grey-level difference distribution of the image
	Contrast	A measure of intensity contrast between a pixel and its neighbor over the whole image		Maximum probability	Measures the maximum likelihood of producing the pixels of interest
	Homogeneity	Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal		IMC1	Measure of dependency between two random variables
	Correlation	A measure of how correlated a pixel is to its neighbor over the whole image		IMC2	
	Sum of squares	Measures the dispersion (with regard to the mean) of the grey-level distribution	LBP LDP LOOP	Mean	Measures mean and standard deviation of histograms
	Cluster shade	Characterizes the tendency of clustering of the pixels in the region of interest		Standard deviation	
	Cluster prominence			Dynamic range	Measures the ratio between the largest and smallest values
	Dissimilarity	A measure of distance between pairs of pixels in the region of interest		Kurtosis	Measure of the "tailedness" of the probability distribution
	Autocorrelation	Represents the degree of similarity between a given time series and a lagged version of it		Skewness	Measure of the asymmetry of the probability distribution
	Sum average	Measures the mean of the grey-level sum distribution of the image	Fractal dimension	Mean	Measures mean and standard deviation of the box-counting method
	Sum entropy	Measures the disorder related to the grey-level sum distribution of the image		Standard deviation	
	Sum variance	Measures the dispersion (with regard to the mean) of the grey-level sum distribution of the image			
	Inverse difference	A measure of local homogeneity of an image			

$$\text{Cluster Prominence} = \sum_{I=1}^{N}\sum_{J=1}^{N}(i+j-2\mu)^3 p(i.j),$$

$$\text{Dissimilarity} = \sum_{i=1}^{N}\sum_{j=1}^{N}|i-j|.p(i.j),$$

$$\text{Autocorrelation} = \sum_{i=1}^{N}\sum_{j=1}^{N}(i.j)P(i.j),$$

$$\text{Sum average (sa)} = \sum_{k=2}^{2N} kp_{x+y}(k),$$

$$\text{Sum entropy (se)} = -\sum_{k=2}^{2N} p_{x+y}(k)\log p_{x+y}(k),$$

$$\text{Sum variance (sv)} = \sum_{k=2}^{2N}\left(k-\mu_{x+y}\right)^2 P_{x+y(k)},$$

$$\text{Inverse difference} = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{p(i.j)}{1+|i-j|},$$

$$\text{Inverse difference moment} = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{p(i.j)}{1+(i-j)^2},$$

$$\text{Difference variance} = \sum_{k=0}^{N-1}\left(k-\mu_{x-y}\right)^2 P_{x-y}(k),$$

$$\text{Difference entropy} = \sum_{k=0}^{N-1} p_{x-y}(k)\log\left\{p_{x-y}(k)\right\},$$

$$\text{Maximum probability} = \max_{i.j} p(i.j),$$

$$\text{Information measure of correlation 1 (IMC1)} = \frac{HXY - HXY1}{\max(HX.HY)},$$

$$\text{Information measure of correlation 2 (IMC2)} = \sqrt{1-\exp\left[-2(HXY2-HXY)\right]}, \quad (5)$$

where element $[i, j]$ of the matrix is generated by counting the number of times a pixel with value i is adjacent to a pixel with value j and then dividing the entire matrix by the total number of such comparisons made. Each entry is therefore considered to be the probability that a pixel with value i will be found adjacent to a pixel of value j. μ_x, μ_y, σ_x, and σ_y are means and standard deviations. p_x and p_y are partial probability density functions. x and y are the coordinates (row and column) of an entry in the cooccurrence matrix, and $p_{x+y}(i)$ is the probability of cooccurrence matrix coordinates summing to $x + y$. HX and HY are the entropies of p_x and p_y. Finally, HXY, $HXY1$, and $HXY2$ are shown in

Table 2: A list of used abbreviations and their explanations.

Abbreviation	Explanation
MS	Multiple sclerosis
ON	Optic neuritis
OCT	Optical coherence tomography
CNS	Central nervous system
LSI	Layer-stacked images
MS-ON	Multiple sclerosis with optic neuritis
MS-None ON	Multiple sclerosis without optic neuritis
MRI	Magnetic resonance imaging
RNFL	Retinal nerve fiber layer
GLCM	Grey-level cooccurrence matrix
HC	Health control
LBP	Local binary pattern
LDP	Local directional pattern
LOOP	Local optimal oriented pattern
SVM	Support vector machine
LDA	Linear discriminant analysis
FD	Fractal dimension
LS image	Layer-stacked image

$$HXY = -\sum_i\sum_j p(i,j)\log\left(p(i,j)\right),$$
$$HXY1 = -\sum_i\sum_j p(i,j)\log\left\{P_X(i)p_y(j)\right\},$$
$$HXY2 = -\sum_i\sum_j p_x(i)p_y(j)\log\left\{P_X(i)p_y(j)\right\}. \quad (6)$$

In LBP, LDP, and LOOP methods, the features in Table 1 should be extracted from the histogram of the output. To solve the same problem of unwanted strip artifact, the first column of the histogram (which represent unwanted zero pixels) is eliminated. Finally, five statistical features including mean, standard deviation, dynamic range, kurtosis, and skewness are extracted.

The last category of texture analysis methods to be considered is fractal analysis. Here, we remove the black background above the layer-stacked images before performing the masking step. The mean and standard deviation of the fractal dimensions for each image is then reported.

2.4. Feature Selection and Classification. To handle the course of dimensionality problem caused by small number of available data compared to bunch of calculated features, more significant features are selected based on t-test and Bonferroni correction. The Bonferroni correction is an adjustment made to p values when several dependent or independent statistical tests are being performed simultaneously on a single data set. To perform a Bonferroni correction, the critical p value (α) is divided by the number of comparisons being made.

Here, considering that the majority of the features have been extracted from the GLCM matrix and this matrix has produced the features in four different angles, according to

Table 3: Evaluation of statistical significance of the extracted features, before feature selection. The t-test was used to identify which features show significant differences between healthy and MS (ON and None-ON) cases. The p values indicate the test rejection of the null hypothesis at 5% significance level, considering the Bonferroni correction (p value < 0.001).

Features	Layer	p (HC vs. MS-None-ON)	p (HC vs. MS-ON)	p (MS-ON vs. MS-None-ON)	p (HC vs. MS)	Features	Layer	p (HC vs. MS-None-ON)	p (HC vs. MS-ON)	p (MS-ON vs. MS-None-ON)	p (HC vs. MS)
Autocorrelation	2	<0.001	<0.001	0.966	<0.001	Fractal mean	1	<0.001	<0.001	0.832	0.051
Autocorrelation	3	<0.001	0.101	0.092	0.003	Fractal mean	4	<0.001	<0.001	0.799	0.955
Autocorrelation	4	<0.001	0.026	0.222	0.008	Fractal mean	5	<0.001	<0.001	0.862	0.896
Cluster prominence	3	<0.001	0.086	0.033	<0.001	Fractal mean	8	<0.001	<0.001	0.339	0.785
Cluster prominence	4	0.002	0.057	0.525	<0.001	Fractal mean	9	<0.001	<0.001	0.964	0.710
Cluster shade	2	0.038	<0.001	0.209	0.044	Fractal mean	10	<0.001	<0.001	0.514	0.408
Cluster shade	3	<0.001	0.025	0.054	<0.001	Fractal Std.	5	<0.001	<0.001	0.744	0.014
Cluster shade	4	<0.001	0.015	0.298	<0.001	Fractal Std.	8	0.113	0.001	0.276	<0.001
Contrast	1	0.341	<0.001	0.058	0.012	Fractal Std.	9	<0.001	<0.001	0.751	<0.001
Contrast	2	<0.001	<0.001	0.944	<0.001	Fractal Std.	10	<0.001	<0.001	1.000	<0.001
Contrast	3	<0.001	0.005	0.150	0.001	LBP mean	2	0.001	0.003	0.981	<0.001
Contrast	4	<0.001	0.001	0.398	<0.001	LBP mean	3	0.001	<0.001	0.950	0.017
Correlation	1	<0.001	<0.001	0.999	0.927	LBP mean	4	<0.001	<0.001	0.127	0.024
Correlation	2	<0.001	0.004	0.029	0.546	LBP mean	5	<0.001	<0.001	0.353	0.061
Difference entropy	2	<0.001	<0.001	0.929	0.951	LBP mean	6	<0.001	<0.001	0.941	0.335
Difference entropy	3	<0.001	<0.001	0.513	0.677	LBP mean	7	<0.001	<0.001	0.908	<0.001
Difference entropy	4	<0.001	<0.001	0.546	0.263	LBP mean	9	<0.001	<0.001	0.251	0.068
Difference entropy	5	<0.001	0.004	0.413	0.733	LBP Std.	4	<0.001	<0.001	0.409	0.319
Difference variance	2	<0.001	<0.001	0.999	0.001	LBP Std.	5	<0.001	<0.001	0.456	0.645
Difference variance	3	<0.001	0.001	0.169	<0.001	LBP Std.	6	<0.001	0.001	0.905	0.234
Difference variance	4	<0.001	<0.001	0.523	<0.001	LBP Std.	7	<0.001	<0.001	0.205	0.116
Dissimilarity	2	<0.001	<0.001	0.790	0.010	LBP Std.	9	<0.001	<0.001	0.417	0.150
Dissimilarity	3	<0.001	0.001	0.271	0.030	LBP dynamic range	3	0.525	0.002	0.088	<0.001
Dissimilarity	4	<0.001	0.001	0.380	0.013	LBP dynamic range	5	<0.001	0.004	0.567	0.164
Dissimilarity	5	<0.001	0.035	0.274	0.302	LBP dynamic range	6	<0.001	0.006	0.675	<0.001
Energy	2	<0.001	<0.001	0.999	<0.001	LBP kurtosis	2	<0.001	<0.001	0.804	0.968
Energy	3	<0.001	<0.001	0.921	<0.001	LBP kurtosis	3	<0.001	<0.001	0.941	0.048
Energy	4	<0.001	0.001	0.528	0.861	LBP kurtosis	4	0.001	<0.001	0.805	0.001
Energy	5	<0.001	0.019	0.329	0.008	LDP mean	2	0.459	0.659	0.961	<0.001
Entropy	2	<0.001	<0.001	0.896	0.903	LDP mean	4	<0.001	<0.001	0.564	0.067
Entropy	3	<0.001	<0.001	0.446	0.765	LDP mean	5	<0.001	<0.001	0.555	0.035

Table 3: Continued.

Features	Layer	p (HC vs. MS-None-ON)	p (HC vs. MS-ON)	p (MS-ON vs. MS-None-ON)	p (HC vs. MS)	Features	Layer	p (HC vs. MS-None-ON)	p (HC vs. MS-ON)	p (MS-ON vs. MS-None-ON)	p (HC vs. MS)
Entropy	4	<0.001	<0.001	0.453	0.274	LDP mean	6	<0.001	<0.001	0.121	0.799
Entropy	5	<0.001	0.029	0.222	0.871	LDP mean	7	<0.001	<0.001	0.883	0.047
Homogeneity	2	<0.001	<0.001	0.942	0.900	LDP mean	8	<0.001	<0.001	0.808	0.055
Homogeneity	3	<0.001	<0.001	0.681	0.690	LDP mean	9	<0.001	<0.001	0.219	0.037
Homogeneity	4	<0.001	<0.001	0.504	0.201	LDP mean	10	<0.001	<0.001	0.834	0.782
Homogeneity	5	<0.001	0.006	0.396	0.680	LDP skewness	2	<0.001	<0.001	0.813	0.530
IMC1	2	<0.001	<0.001	0.862	0.017	LDP skewness	3	<0.001	<0.001	0.930	0.085
IMC2	9	0.004	<0.001	0.619	0.043	LDP skewness	4	0.004	<0.001	0.633	0.002
Inverse difference moment normalized	1	0.319	<0.001	0.063	0.011	LDP Std.	2	<0.001	<0.001	0.276	0.018
Inverse difference moment normalized	2	<0.001	<0.001	0.920	<0.001	LDP Std.	4	<0.001	<0.001	0.652	0.083
Inverse difference moment normalized	3	<0.001	0.006	0.154	0.001	LDP Std.	5	<0.001	<0.001	0.882	0.025
Inverse difference moment normalized	4	<0.001	0.001	0.381	<0.001	LDP Std.	6	0.005	<0.001	0..093	0.597
Maximum probability	2	<0.001	<0.001	1.000	<0.001	LDP Std.	7	<0.001	<0.001	0.554	0.310
Maximum probability	3	<0.001	<0.001	0.876	<0.001	LDP Std.	8	<0.001	<0.001	0.989	0.870
Maximum probability	4	<0.001	0.068	0.293	0.430	LDP Std.	9	<0.001	<0.001	0.629	.882
Maximum probability	5	<0.001	0.062	0.222	0.647	LDP Std.	10	<0.001	<0.001	0.901	0.792
Sum average	2	<0.001	<0.001	0.814	0.005	LDP dynamic range	1	0.188	0.219	0.999	<0.001
Sum average	3	<0.001	0.013	0.194	0.028	LDP kurtosis	2	<0.001	<0.001	0.998	<0.001
Sum average	4	<0.001	0.006	0.383	0.048	LDP kurtosis	3	<0.001	<0.001	1.000	<0.001
Sum entropy	2	<0.001	<0.001	0.953	0.708	LDP kurtosis	4	<0.001	<0.001	0.801	0.466
Sum entropy	3	<0.001	0.002	0.402	0.938	LDP kurtosis	5	<0.001	0.003	0.781	0.001
Sum entropy	4	<0.001	0.001	0.514	0.384	LOOP mean	2	0.459	0.659	0.961	<0.001
Sum entropy	5	<0.001	0.064	0.182	0.806	LOOP mean	4	<0.001	<0.001	0.564	0.067
Sum of squares	2	<0.001	<0.001	0.528	0.001	LOOP mean	5	<0.001	<0.001	0.555	0.035
Sum of squares	3	<0.001	0.014	0.092	<0.001	LOOP mean	6	<0.001	<0.001	0.121	0.799
Sum of squares	4	<0.001	0.003	0.262	<0.001	LOOP mean	7	<0.001	<0.001	0.883	0.047
Sum variance	2	<0.001	<0.001	0.400	0.002	LOOP mean	8	<0.001	<0.001	0.808	0.055
Sum variance	3	<0.001	0.016	0.087	<0.001	LOOP mean	9	<0.001	<0.001	0.219	0.037
Sum variance	4	<0.001	0.004	0.248	<0.001	LOOP mean	10	<0.001	<0.001	0.834	0.782
LOOP Std.	2	<0.001	<0.001	0.964	0.013	LOOP skewness	2	<0.001	<0.001	0.998	<0.001
LOOP Std.	3	<0.001	<0.001	0.202	0.021	LOOP skewness	3	<0.001	<0.001	1.000	<0.001

Table 3: Continued.

Features	Layer	p (HC vs. MS-None-ON)	p (HC vs. MS-ON)	p (MS-ON vs. MS-None-ON)	p (HC vs. MS)	Features	Layer	p (HC vs. MS-None-ON)	p (HC vs. MS-ON)	p (MS-ON vs. MS-None-ON)	p (HC vs. MS)
LOOP Std.	4	<0.001	<0.001	0.800	0.218	LOOP skewness	4	<0.001	<0.001	0.818	0.641
LOOP Std.	5	<0.001	<0.001	0.830	0.019	LOOP skewness	5	<0.001	0.002	0.769	0.001
LOOP Std.	6	0.014	<0.001	0.091	0.548	LOOP Std.	9	<0.001	<0.001	0.538	0.390
LOOP Std.	7	<0.001	<0.001	0.576	0.425	LOOP Std.	10	<0.001	<0.001	0.786	0.799
LOOP Std.	8	<0.001	<0.001	0.463	0.316						

Bonferroni correction, the value of meaningful p value ($p < 0.005$) is divided by 4 and $p < 0.001$ is considered as a significant level for cut-off. After Bonferroni adjustment for multiple comparisons, features with $p < 0.001$ are selected as significant features. Then, two classification models, SVM and LDA, are utilized for differentiating between four possible groups including HC vs. MS, HC vs. MS-ON, HC vs. MS-None-ON, and MS-ON vs. MS-None-ON. A 10-fold cross-validation is used to evaluate accuracy, for each classification model.

3. Result

3.1. Feature Analysis. To evaluate the statistical significance of the extracted features, the t-test is used to identify which features show significant differences between healthy and MS (ON and None-ON) cases (Table 3). The p values indicate the test rejection of the null hypothesis at the 5% significance level, considering the Bonferroni correction (p value < 0.001). Frequencies of significant selected features for each retinal layer are also presented in Table 4.

3.2. Classification Result. According to Tables 3 and 4, 15 common significant features between three groups (HC, MS-ON, and MS-None-ON) are selected as input of each classifier. However, no significant feature is found for the last group (MS-ON vs. MS-None-ON). Then, the classification step is done to discriminate between four target groups including HC vs. MS, HC vs. MS-None-ON, HC vs. MS-ON, and MS-ON vs. MS-None-ON. The accuracy results obtained from our classifiers in different groups are shown in Table 5.

In addition, to prepare a fair comparison with the previous studies, we also test the performance of the two classification models using thickness features as input. As above-mentioned, utilizing texture features for our intended goal is totally novel and previous researches were only relying on thickness as discriminant features. Therefore, the thickness features are calculated as the average value of distance between two consecutive boundaries, which lead to 10 thickness values out of 11 retinal layer boundaries, and this thickness feature vector is fed also to each classifier. In summary, the following set of information are utilized as input of each two classifiers:

(I) 15 common selected texture features based on t-test and Bonferroni correction

(II) Thickness features

(III) Combination of I and II

As it can be found in Table 5, in cases I and II, SVM outperforms LAD. In analyzing the effect of texture and thickness features separately, it has to be mentioned that the best accuracy result for groups HC vs. MS and HC vs. MS-ON is found using texture features and SVM classifier. Meanwhile, for groups HC vs. MS-None-ON and MS-ON vs. MS-None-ON, thickness features and SVM classifier obtain the best accuracy.

Table 4: Frequency of significant selected features for each retinal layer.

Layers	Frequency			
	HC vs. MS	HC vs. MS-ON	HC vs. MS-None-ON	MS-ON vs. MS-None-ON
1	1	4	2	0
2	10	22	22	0
3	10	11	21	0
4	7	15	24	0
5	0	8	18	0
6	1	5	5	0
7	1	6	6	0
8	1	5	5	0
9	1	9	8	0
10	1	5	7	0
Total	33	90	117	0

Table 5: The accuracy of texture features, thickness, and combination of texture features and thicknesses.

Methods	Classifiers	HC vs. MS	HC vs. MS-ON	HC vs. MS-None-ON	MS-ON vs. MS-None-ON
Texture features	SVM	85.3	83.6	78.6	64.1
	LDA	72.0	74.6	64.3	48.8
Thicknesses	SVM	84.0	81.8	90.0	89.7
	LDA	64	69.1	73.3	82.1
Texture features & thicknesses	SVM	96.0	87.3	96.4	82
	LDA	100	98.2	96.5	56.4

Furthermore, the impressive point is that in case III and with the combination of texture and thickness features as input of the classifiers, the result improved considerably and also the performance of the LDA classifier is superior to the SVM performance in most of the conditions.

4. Conclusion

There are no specific tests for MS detection. Instead, a diagnosis of MS often relies on ruling out other conditions that might produce similar signs and symptoms, known as a differential diagnosis. Blood tests, spinal tap (lumbar puncture), evoked potential tests, and MRI are the conventional MS diagnosis methods. The first MR images of MS were produced in the early 1980s [23]. In most people with relapsing-remitting MS, the diagnosis is fairly straightforward and based on a pattern of symptoms consistent with the disease and confirmed by brain imaging scans such as MRI; however, MS diagnosis can be more difficult in patients with unusual symptoms or progressive disease.

MRI-based methods have been indeed the most successful techniques to estimate CNS damage up to the present, although it is becoming increasingly clear that due to the ability of direct visualization of retinal axons, OCT has become an extremely sensitive method for

imaging neurodegeneration in MS patients. Studies show thickness reduction in retinal layers of MS patients with and without history of ON by OCT image analysis [4, 9, 24–28]. Hence, OCT is suggested as an important tool for monitoring MS and also as a complementary method for MRI-based diagnosis techniques [29–32]. However, as mentioned above, the majority of previous works are on the thickness analysis of retinal layers. Here, by combining the information of thickness and texture of retinal layers, we prepared a more comprehensive analysis of OCT imaging performance in the diagnosis of MS with or without ON.

Indeed, texture analysis is a novel strategy for studying intrinsic changes in retinal layers during neurodegenerative diseases. MS, as one of the famous neurodegenerative disorders, is investigated in this research.

After performing a vast survey on available texture analysis methods, a treasury of powerful features is collected in this paper. As a primary work, this paper shows the ability of such features in discrimination of HC and MS (ON and None-ON) cases. Even with simple classification methods, the texture features are powerful to diagnose MS cases (from HC ones) with accuracy of 85.3% and 72% with SVM and LDA classifiers, respectively.

Another valuable point is that adding information of conventional thickness values to texture features improves the discrimination between most of the target groups including HC vs. MS, HC vs. MS-None-O, and HC vs. MS-ON. It should be noted that the results of the last group (MS-ON vs. MS-None-ON) are generally weaker than other groups due to the lack of significant discriminant texture features for this group.

Furthermore, the findings show that some layers like 2, 3, and 4 carry more texture information useful in separation of HC from MS cases. Such finding can be a start point for further investigation in this area.

Data Availability

The data will be available upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Vice-Chancellery for Research and Technology of Isfahan University of Medical Sciences under Grant 397766 and also by the National Institute for Medical Research Development under Grant 964582.

References

[1] R. S. Nowakowski, "Stable neuron numbers from cradle to grave," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 33, pp. 12219-12220, 2006.

[2] J. Nakahara, M. Maeda, S. Aiso, and N. Suzuki, "Current concepts in multiple sclerosis: autoimmunity versus oligodendrogliopathy," *Clinical Reviews in Allergy and Immunology*, vol. 42, no. 1, pp. 26–34, 2012.

[3] J. J. González-López, G. Rebolleda, M. Leal et al., "Comparative diagnostic accuracy of ganglion cell-inner plexiform and retinal nerve fiber layer thickness measures by Cirrus and spectralis optical coherence tomography in relapsing-remitting multiple sclerosis128517," *BioMed Research International*, vol. 2014, 2014.

[4] F. Barkhof, P. A. Calabresi, D. H. Miller, and S. C. Reingold, "Imaging outcomes for neuroprotection and repair in multiple sclerosis trials," *Nature Reviews Neurology*, vol. 5, no. 5, pp. 256–266, 2009.

[5] J. M. Schmitt, "Optical coherence tomography (OCT): a review," *IEEE Journal on Selected Topics in Quantum Electronics*, vol. 5, no. 4, pp. 1205–1215, 1999.

[6] H. Zimmermann, T. Oberwahrenbrock, A. U. Brandt, F. Paul, and J.-M. Dörr, "Optical coherence tomography for retinal imaging in multiple sclerosis," *Degenerative Neurological and Neuromuscular Disease*, vol. 4, pp. 153–162, 2014.

[7] J. Michalewski, Z. Michalewska, Z. Nawrocka, M. Bednarski, and J. Nawrocki, "Correlation of choroidal thickness and volume measurements with axial length and age using swept source optical coherence tomography and optical low-coherence reflectometry," *BioMed Research International*, vol. 2014, Article ID 639160, 2014.

[8] V. Parisi, G. Manni, M. Spadaro et al., "Correlation between morphological and functional retinal impairment in multiple sclerosis patients," *Investigative Ophthalmology and Visual Science*, vol. 40, no. 11, pp. 2520–2527, 1999.

[9] A. Petzold, L. J. Balcer, P. A. Calabresi et al., "Retinal layer segmentation in multiple sclerosis: a systematic review and meta-analysis," *The Lancet Neurology*, vol. 16, no. 10, pp. 797–812, 2017.

[10] B. E. Varga, W. Gao, K. L. Laurik et al., "Investigating tissue optical properties and texture descriptors of the retina in patients with multiple sclerosis," *PLoS ONE*, vol. 10, no. 11, p. e0143711, 2015.

[11] M. Mirmehdi, X. Xian, and J. Suri, *Handbook of Texture Analysis*, pp. 1–413, Imperial College Press, UK, 2008.

[12] M. Tuceryan and A. K. Jain, "Texture analysis," in *Pattern Recognition and Computer Vision*, pp. 207–248, World Scientific Publishing Co, 1998.

[13] S. Khalid, M. U. Akram, T. Hassan, A. Nasim, and A. Jameel, "Fully automated robust system to detect retinal edema, central serous chorioretinopathy, and age related macular degeneration from optical coherence tomography images," *BioMed Research International*, vol. 2017, Article ID 7148245, 2017.

[14] M. Baroni, S. Diciotti, A. Evangelisti, P. Fortunato, and A. La Torre, "Texture classification of retinal layers in optical coherence tomography," *IFMBE Proceedings*, vol. 16, no. 1, pp. 847–850, 2007.

[15] N. Anantrasirichai, A. Achim, J. E. Morgan, I. Erchova, and L. Nicholson, "SVM-based texture classification in optical coherence tomography," in *Proceedings-International Symposium on Biomedical Imaging*, pp. 1332–1335, San Francisco, CA, USA, 2013.

[16] T. W. Sawyer, S. Chandra, P. F. S. Rice, J. W. Koevary, and J. K. Barton, "Three-dimensional texture analysis of optical coherence tomography images of ovarian tissue," *Physics in Medicine and Biology*, vol. 63, no. 23, p. 235020, 2018.

Hindawi
BioMed Research International
Volume 2023, Article ID 9878305, 1 page
https://doi.org/10.1155/2023/9878305

Retraction

Retracted: Identifying COVID-19-Specific Transcriptomic Biomarkers with Machine Learning Methods

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Chen, Z. Li, T. Zeng et al., "Identifying COVID-19-Specific Transcriptomic Biomarkers with Machine Learning Methods," *BioMed Research International*, vol. 2021, Article ID 9939134, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 9939134, 11 pages
https://doi.org/10.1155/2021/9939134

Research Article

Identifying COVID-19-Specific Transcriptomic Biomarkers with Machine Learning Methods

Lei Chen,[1,2] **Zhandong Li**,[3] **Tao Zeng**,[4] **Yu-Hang Zhang**,[5] **KaiYan Feng**,[6] **Tao Huang**,[4,7] **and Yu-Dong Cai**[1]

[1]*School of Life Sciences, Shanghai University, shanghai 200444, China*
[2]*College of Information Engineering, Shanghai Maritime University, shanghai 201306, China*
[3]*College of Food Engineering, Jilin Engineering Normal University, Changchun 130052, China*
[4]*Bio-Med Big Data Center, CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, shanghai 200031, China*
[5]*Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA*
[6]*Department of Computer Science, Guangdong AIB Polytechnic College, Guangzhou 510507, China*
[7]*CAS Key Laboratory of Tissue Microenvironment and Tumor, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China*

Correspondence should be addressed to Tao Huang; tohuangtao@126.com and Yu-Dong Cai; cai_yud@126.com

Received 17 March 2021; Revised 3 June 2021; Accepted 24 June 2021; Published 7 July 2021

Academic Editor: Min Tang

COVID-19, a severe respiratory disease caused by a new type of coronavirus SARS-CoV-2, has been spreading all over the world. Patients infected with SARS-CoV-2 may have no pathogenic symptoms, i.e., presymptomatic patients and asymptomatic patients. Both patients could further spread the virus to other susceptible people, thereby making the control of COVID-19 difficult. The two major challenges for COVID-19 diagnosis at present are as follows: (1) patients could share similar symptoms with other respiratory infections, and (2) patients may not have any symptoms but could still spread the virus. Therefore, new biomarkers at different omics levels are required for the large-scale screening and diagnosis of COVID-19. Although some initial analyses could identify a group of candidate gene biomarkers for COVID-19, the previous work still could not identify biomarkers capable for clinical use in COVID-19, which requires disease-specific diagnosis compared with other multiple infectious diseases. As an extension of the previous study, optimized machine learning models were applied in the present study to identify some specific qualitative host biomarkers associated with COVID-19 infection on the basis of a publicly released transcriptomic dataset, which included healthy controls and patients with bacterial infection, influenza, COVID-19, and other kinds of coronavirus. This dataset was first analysed by Boruta, Max-Relevance and Min-Redundancy feature selection methods one by one, resulting in a feature list. This list was fed into the incremental feature selection method, incorporating one of the classification algorithms to extract essential biomarkers and build efficient classifiers and classification rules. The capacity of these findings to distinguish COVID-19 with other similar respiratory infectious diseases at the transcriptomic level was also validated, which may improve the efficacy and accuracy of COVID-19 diagnosis.

1. Introduction

COVID-19 is recognized as starting from the end of 2019. It is a severe respiratory disease caused by a new type of coronavirus SARS-CoV-2 and has been spreading all over the world [1–3]. By the end of January 2021, approximately 100 million cases and 2 million deaths have been reported worldwide [4], making COVID-19 one of the most widespread and deadly infectious diseases in human history. In the US alone, more than 26 million cases were reported [4]. Different from other severe diseases, COVID-19 hardly has typical symptoms that could be used for diagnosis. A wide range of disease-associated symptoms, such as respiratory or systematic, were reported to be associated with COVID-19, including fever,

cough, headache, diarrhea, and muscle or body aches [5, 6]. Moreover, patients infected with SARS-CoV-2 may have no pathogenic symptoms, i.e., presymptomatic patients and asymptomatic patients. In the early stage (first 2 days) of SARS-CoV-2 infection, patients may not have any COVID-19 associated symptoms, and they could be clustered as presymptomatic patients [7]. However, some patients may never have any symptoms but still have been infected by SARS-CoV-2, and they could be defined as asymptomatic patients. Both types of patients could further spread the virus to other susceptible people, thereby making the control of the COVID-19 pandemic difficult [8].

The two major challenges for COVID-19 diagnosis at present are as follows: (1) patients could share similar symptoms with other respiratory infections, and (2) patients may not have any symptoms but could still spread the virus. Therefore, identifying new biomarkers at different omics levels (genomic, transcriptomic, or proteomic levels) may be helpful for large-scale screening and diagnosis of COVID-19. Genomic analyses on COVID-19 mainly focused on the genomics of the virus and not the host by identifying the typical sequence of the ORF1ab, spike, ORF3a, envelope, membrane, and nucleocapsid of SARS-CoV-2 [9]. Meanwhile, many transcriptomic and proteomic analyses focused on the host, especially on the host–virus interaction-associated alterations in the host system. For example, in April 2020, a systematic study (GSE150728) on the expression pattern of immune-associated genes in lung tissue or related human lung cells during the infection of SARS-CoV-2 was presented, revealing that the selective death of type II pneumocytes caused by abnormal immune responses caused high morbidity and mortality in COVID-19 cases [10]. However, despite the encouraging results presented, this study has two obvious shortcomings: (1) the major findings were based on in vitro-cultured cell lines and only two patients each group were enrolled, and (2) only immune-associated genes were taken into consideration. As for other transcriptomic analyses, only single-cell subgroups, such as human lung cell lines [10], cardiomyocyte cells [11], and human bronchial organoids [12], have been analysed and discussed, and systematic transcriptomic analyses on lung tissue are lacking.

Although some initial analyses on such transcriptomic datasets could identify a group of candidate gene biomarkers, such as IFI6, TIMP1, and LGR6, for COVID-19 in the previous study [13], the dataset used did not contain normal controls and only divided patients into three rough groups: patients with COVID-19, those with other viral infections, and those without viral infections. Thus, the previous work could not fully identify biomarkers capable for clinical use in COVID-19, which requires disease-specific diagnosis compared with other multiple infectious diseases. As an extension of the previous study, a recent dataset released on the Gene Expression Omnibus (GEO) database (GSE161731) [14] was introduced for further analyses. These blood sample transcriptomic data of 195 subjects include 19 healthy controls and 23, 17, 77, and 59 patients with bacterial infection, influenza, COVID-19, and other kinds of coronavirus, respectively. The new dataset could be used to screen out potential transcriptomic biomarkers from the comprehensive lung tissue, and a comparison between COVID-19 and other infectious respiratory diseases could further help identify disease-specific biomarkers to distinguish COVID-19 from other similar diseases.

In this study, on the basis of the publicly released dataset, optimized machine learning models were applied to identify some specific qualitative host biomarkers associated with COVID-19 infection. Two powerful feature selection methods (Boruta [15] and Max-Relevance and Min-Redundancy (mRMR) [16]), were applied on this dataset one by one. A feature list was generated, which was further fed into the incremental feature selection (IFS) method [17]. Four classic classification algorithms were tried in the IFS method. As a result, we accessed some essential biomarkers, efficient classifiers, and classification rules. The capacity of these findings to distinguish COVID-19 with other similar respiratory infectious diseases at the transcriptomic level was validated, which could improve the efficacy and accuracy of COVID-19 diagnosis.

2. Materials and Methods

2.1. Data. The blood expression profiles of 15,379 genes in acute respiratory infection samples were downloaded from the GEO database under accession number GSE161731 [14]. A total of 195 samples with demographic information were included as follows: 19 healthy controls, 23 patients with bacterial pneumonia, 17 patients with influenza, 59 patients with seasonal coronavirus, and 77 patients with SARS-CoV-2 infection. The 15,379 genes are listed in Table S1. The processed transcript-per-million expression data were used for further analysis.

2.2. Boruta Feature Filtering. The investigated dataset involved lots of features/genes. Evidently, some are relevant to acute respiratory infection, whereas others are not. To extract the relevant features, the Boruta [15] method was employed.

Boruta is a random forest- (RF-) based feature select method. Given a dataset, a shuffled feature is added for each original feature. A RF classifier is built on a dataset with original and added features. According to the performance of RF, calculate the Z score of all features and find the maximum Z score among shuffled features (MZSA). Determine the original features as "important" if their Z scores are significantly higher than MZSA; whereas when Z scores of some features are signicantly lower than MZSA, they are labelled as "unimportant." The above procedures are executed several times until all original features are labelled as "important" or "unimportant," or the times of RF runs have reached a predefined number.

In this study, we adopted the program of Boruta retrieved from https://github.com/scikit-learn-contrib/boruta_py. It was run with its default parameters.

2.3. Max-Relevance and Min-Redundancy (mRMR) Feature Selection. mRMR [16] is a mutual information- (MI-) based feature selection approach to evaluate the importance of features. This method has wide applications in tackling several

biological and medical problems [13, 18–23]. For variables x and y, their MI can be calculated by

$$I(x, y) = \iint p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dxdy, \quad (1)$$

where $p(x)$ denotes the marginal probabilistic density of x, $p(x, y)$ represents the joint probabilistic density of x and y, respectively. A high MI means two variables have high associations. For a feature, its importance is reflected by its rank in a feature list. To generate such list, a loop procedure is included in the mRMR method. Initially, this list is empty. A feature is selected in each round and appended to this list. Such feature is selected by the following manner. For each nonselected feature, calculate its relevance to class labels, which is defined as the MI of it and class labels, and its redundancies to already-selected features, which is defined as the average MI of it and already-selected features. The feature with maximum difference of above two values is selected. The loop procedure stops until all features are selected. For convenience, this list was called mRMR feature list in this study.

In present study, the mRMR program downloaded from http://penglab.janelia.org/proj/mRMR/ was used. Such program was executed with its default parameters.

2.4. Incremental Feature Selection (IFS). IFS is a widely used approach integrated with supervised classifier (e.g., SVM) to determine the optimal feature number for classification model construction [17]. On the basis of the mRMR feature list available from mRMR, IFS could produce step-wise feature subsets in a given step interval s (i.e., 1). For instance, the first feature subset has the top-ranked s features, and then the second feature subset has the top-ranked $2 \times s$ features, and so on. For each candidate feature subset, a classifier could be built on the basis of the training sample data within such feature subset. In IFS, the optimal feature subset is obtained when a classifier could achieve the best performance measurement, evaluated by Matthew's correlation coefficient (MCC) [24], within 10-fold cross-validation [25] on such feature subset.

2.5. Candidate Classification Algorithms. The four classification algorithms were tried in the IFS method. Their brief descriptions are as follows.

2.5.1. RF. RF is an assembly prediction model that uses average prediction [26], which predicts the class label of a test sample dependent on the consensus prediction results from a series of decision trees (DTs). It is widely used in bioinformatics researches [27–31].

2.5.2. Support Vector Machine (SVM). SVM [32–38] consists of several computational steps. First, it transforms the original data from a low-dimensional data space to a high-dimensional data space. It could also transform the original nonlinear data pattern to new linear data pattern [39, 40]. Second, it divides the data points in the high-dimensional data space by maximizing the space interval among data points from different classes/labels. Finally, it predicts the test sample's class label by judging which space interval this new data point belongs to. Here, the SVM model construction adopted the SMO in Weka.

2.5.3. K-Nearest Neighbor (kNN). The computational steps of kNN [41] are as follows: first, it calculates the sample distance between a new sample and all training samples. Then, it ranks all training samples in accordance with these distance measurements. Next, it chooses the K-nearest training samples and estimates the class label distribution of these samples. Finally, it predicts the class label of new sample as the one with the largest distribution frequency. Here, the kNN model building adopted the Ibk in Weka.

2.5.4. DT. As a rule-based white-box classification and regression model, DT [42, 43] generally applies IF-TEHN format to indicate each feature's role and weight in a model and corresponding rule, which thereby provides interpretative rules. Here, the DT model learning adopted the CART algorithm with the Gini index in the Scikit-learn package.

2.6. MCC. MCC [24] can evaluate the classification performance of different models. For the multiclass problem faced in this work, MCC could be calculated using the following formula:

$$\mathrm{MCC} = \frac{\mathrm{cov}\,(X, Y)}{\sqrt{\mathrm{cov}\,(X, X)\,\mathrm{cov}\,(Y, Y)}}, \quad (2)$$

where data matrix X has binary values representing the predicted sample classes, data matrix Y has binary values indicating the true sample classes, and $\mathrm{cov}\,(\cdot, \cdot)$ calculates the two matrices' covariance. The value of MCC ranges from -1 to $+1$ [19], and it is equal to $+1$ when the classification model has the best performance.

3. Results

In this study, we applied several advanced computational methods to the blood expression profiles of acute respiratory infection samples. The whole procedures are illustrated in Figure 1. The detailed results are listed in this section.

3.1. Results of Boruta and mRMR Methods. Each acute respiratory infection sample was represented by the blood expression level of 15,379 genes, which are provided in Table S1. These features (genes) were first analysed by the Boruta method. 604 relevant features were extracted, which are listed in Table S2. Then, these features were evaluated by the mRMR method. A feature list, called mRMR feature list, was produced, which is also provided in Table S2.

3.2. Results of IFS Method. The mRMR feature list was fed into the IFS method, which incorporated one of four classification algorithms (RF, SVM, KNN, and DT). 604 feature subsets were constructed in the IFS method, each of which contained some top features in the mRMR feature list. On each feature subset, a classifier was built based on a given classification algorithm, which was further assessed by 10-fold cross-validation. The accuracy on each category, overall accuracy, and MCC were counted. The above measurements

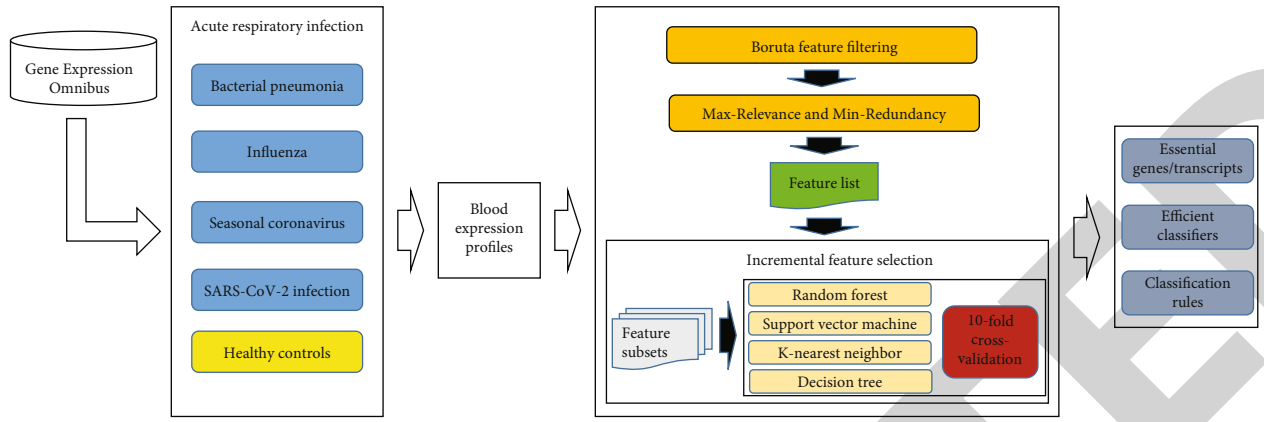


Figure 1: Entire procedures to investigate the blood expression profiles of acute respiratory infection samples. The profiles are retrieved from Gene Expression Omnibus. They are first analysed by Boruta and mRMR methods, resulting in a feature list. Such list is fed into the incremental feature selection method to extract essential biomarker genes/transcripts, build efficient classifiers, and construct classification rules.

obtained by all classification algorithms and constructed feature subsets are available in Table S3. For an easy observation, a curve was plotted for each classification algorithm, in which MCC was set as the Y-axis and number of features was set as the X-axis. These four curves are shown in Figure 2. For SVM, the highest MCC was 0.917, which was obtained by using top 168 features. Thus, the SVM classifier with these features was deemed as the optimum SVM classifier. The overall accuracy of such classifier was 0.938 (Table 1). The accuracies on five categories yielded by such classifier are illustrated in Figure 3. Samples in three categories were all correctly predicted. These results indicated the excellent performance of the optimum SVM classifier.

As for KNN and RF, the highest MCCs were 0.845 and 0.896 when the top 183 and 565, respectively, features were used. These MCCs were lower than that of the optimum SVM classifier. Likewise, the optimum KNN and RF classifiers were built with the corresponding top features. The overall accuracies of these two classifiers are listed in Table 1. They were also lower than that of the optimum SVM classifier. The accuracies on five categories yielded by these two classifiers were also generally lower than those of the optimum SVM classifier (see Figure 3).

In addition to the above-mentioned three black-box classification algorithms, we also employed a white-box classification algorithm, DT. The same procedure was done for this algorithm. The curve is shown in Figure 2. The highest MCC was 0.818 when top 511 features were adopted. Such MCC was lower than that of the optimum SVM/KNN/RF classifier. The overall accuracy was 0.867 (Table 1), also lower than that of the optimum SVM/KNN/RF classifier. Furthermore, the accuracies on five categories, as shown in Figure 3, were also generally lower than those of other three optimum classifiers. Although such DT classifier did not provide good performance, we can obtain more insights from such classifier, which would be listed in the following subsection.

3.3. Classification Rules. The best DT classifier adopted top 511 features. Thus, we used these 511 features to build a DT using all acute respiratory infection samples. 21 rules were extracted from this DT, which are listed in Table S4. Among these 21 rules, eight rules were for prediction of SARS-CoV-2 infection samples, which were most, followed by rules for seasonal coronavirus, influenza, healthy control, and bacterial pneumonia (see Figure 4). The discussion on these rules can be found in Discussion.

3.4. Functional Enrichment Analyses. The optimum SVM classifier adopted top 168 features (genes). Using these selected COVID19 associated genes as gene of interest and all genes in analyses as gene background, we performed GO and KEGG enrichment analyses using DAVID website (https://david.ncifcrf.gov/). The FDR threshold for significant enriched results is set as 0.05. All the significant results are presented in Table 2.

4. Discussion

The top-ranked features (genes/transcripts) and rules were identified by applying these optimal machine learning models. According to recent publications [44–51], several identified top-ranked features and rule-involved features have been confirmed to be associated with the infection of a specific kind of pathogen, thus validating the efficacy and accuracy of the prediction in the current work. The detailed discussion can be found below.

4.1. Transcripts Associated with Disease-Specific Diagnosis of Different Pathogens. The first identified gene in the prediction list is ***RPL6***. Together with some other ribosomal proteins, such as ***RPL3*** and ***RPS20***, *RPL6* has already been reported to have differential expression patterns under specific physical and pathological conditions [52, 53]. Early in 2006, ribosomal proteins have been shown to be associated with lung bacterial infections caused by pneumococcal

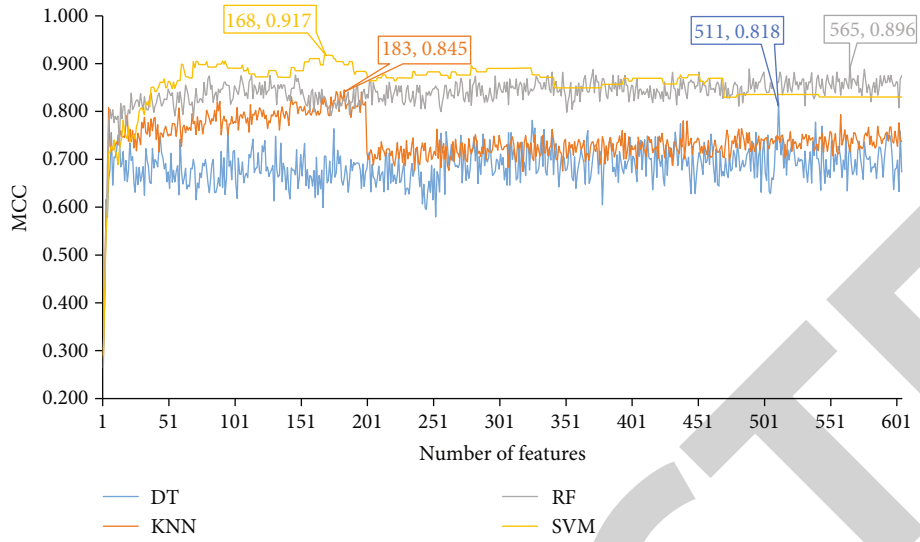


Figure 2: IFS curves with different classifiers on different numbers of gene features. The SVM provides the highest MCC of 0.917 when top 168 features are adopted.

Table 1: Performance of the optimum classifiers based on different classification algorithms.

Classification algorithm	Number of features	Overall accuracy	MCC
Decision tree	511	0.867	0.818
K-nearest neighbor	183	0.882	0.845
Support vector machine	168	0.938	0.917
Random forest	565	0.923	0.896

pneumonia in a mouse model [44]. As for influenza virus infections, in 2015, another independent study [45] at transcriptomic level has identified a group of ribosomal proteins, including *RPL6*, *RPL15*, *RPL17*, and *RPL22*, to have differential expression levels during influenza infections. As for coronavirus, including SARS-CoV-2, in 2020, a study [54] on the interactions between viral envelope protein and host cells confirmed that papain-like proteases, which are quite conserved in the coronavirus family, interact with the host ribosomal proteins. Therefore, ribosomal proteins, such as *RPL6*, *RPL3*, and *RPS20*, have differential expression levels during bacterial infection, influenza, and coronavirus infections, including COVID, thus making such transcripts potential biomarkers to distinguish patients with viral infections and normal controls.

The next identified gene is ***ZNF496***, an effective DNA-binding transcription factor in the lung under physical and pathological conditions [55]. With few validated reports on its associations with infections, it has only been shown to be associated with SARS-CoV-2 in a recent transcriptional regulatory network study [46] and identified as a potential therapeutic target, implying its potential significance for COVID-19 [46]. Therefore, such gene may also be a potential biomarker distinguishing patients with COVID-19 from others.

DYNLRB1, as another predicted biomarker candidate, has previously been reported to be associated with linking dynein to cargos and regulatory adapters for dynein functions [56]. Early in 2011, *DYNLRB1* has been confirmed to be associated with multiple viral infections in lung, including influenza virus but not coronavirus, in mouse models [47]. However, no direct evidence has shown that such gene is associated with bacterial or coronavirus infections (including COVID-19), indicating it may be a potential biomarker for influenza virus infection, which is also in agreement with the prediction.

TRBV20-1 is a transcript of the variable domain of T cell receptor, which participates in the antigen recognition and varies for different potential antigens, such as those from different pathogens, including influenza, bacteria, or coronavirus [48, 49]. Although gene *TRBV20-1* does not have tissue specificity, considering that T cell-mediated immune responses have shown to be associated with COVID infections and the predicted gene *TRBV20-1* has been confirmed to be expressed in lung, it is reasonable to speculate that *TRBV20-1* may participate in the COVID-mediated lung infections.

Apart from another ribosomal protein associated transcript ***RPL36AL***, ***PHOSPHO1***, as a potential regulator for phosphatase activity and phosphocholine phosphatase activity regulations in cells, has been predicted to have differential expression levels during infection with different pathogens. Phosphatase activity has been shown to be essential for the infection of bacteria [57], influenza [58], and SARS-CoV-2 [50]. In particular, in the study associated with SARS-CoV-2, *PHOSPHO1* has also been shown to be associated with immunomodulatory effects of the host against such virus [50]. Therefore, *PHOSPHO1* may also be one of the potential biomarker candidates with disease-specific diagnosis capacity.

TMEM165, as a widely reported transmembrane protein expressed in fibroblasts, has also been predicted to be associated with bacterial infections in lungs. Different from the

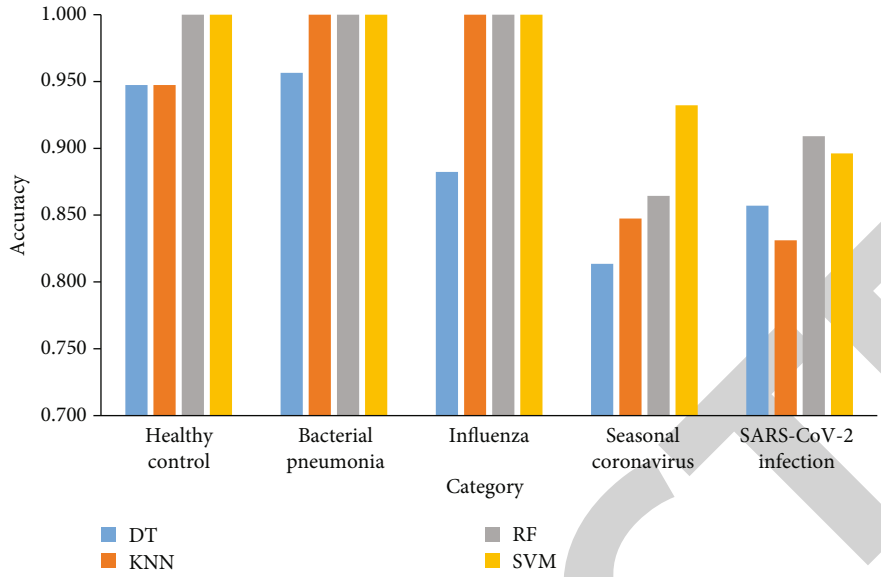


Figure 3: Performance of the optimum classifiers with four different classification algorithms on five categories.

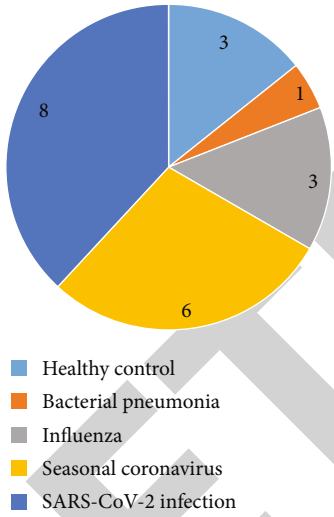


Figure 4: Pie chart to show the distribution of 21 classification rules on five categories.

genes discussed above, *TMEM165* has been shown to be not associated with viral infections, including influenza or coronavirus infection. In 2019, researchers have shown that *TMEM165* is associated with bacterial infections in yeast [59]. Further, another study confirmed that such gene is effective in the lung and associated with chronic bacterial infection and inflammation [51], thus corresponding with the prediction in the present work.

4.2. Quantitative Rules Associated with Disease-Specific Diagnosis of Different Pathogens. Apart from the above qualitative analyses, quantitative analyses were performed to establish accurate rules for disease classification. Here, the top rules of each group were selected for follow-up detailed discussion.

The first rule aims to identify patients with COVID-19 infection with decreased expression levels of ***SORT1***, ***RPL21P28***, ***SIDT2***, and ***TKT*** and a relatively high expression of **GZMB**. ***SORT1*** has been shown to be upregulated in almost all lung infections due to its specific relationships with neutrophil recruitment in lung tissues/surrounding vascular against pathogens, especially for bacterial infections [60–62]. By contrast, specifically, in COVID-19, a network based analyses has shown that such gene is associated with the infection of SARS-CoV-2 with a relatively low expression, corresponding with our predictions [63]. Similar decreased expression levels of *RPL21P28*, *SIDT2*, and *TKT* have also been validated in the transcriptomic analyses of COVID-19 host cells [64, 65]. Generally, *GZMB* has been widely reported to be expressed within cytotoxic CD8+ T cells. However, recent publications have also confirmed that in anti-virus CD4+ T cells, *GZMB* is also highly expressed which is detected using intracellular staining [66]. As specific for SARS-CoV-2 associated infections, in 2020, a specific single-cell transcriptomic analyses on SARS-CoV-2 host cells revealed that in reactive CD4+ T cells, ***GZMB*** turned out to be upregulated [67], corresponding with the prediction in the present study. Although based on our bulk analyses, we cannot confirm whether detected *GZMB* is derived from CD4+ or CD8+ T cells; however, as an SARS-CoV-2 viral infection-associated gene, the identification of such gene may also prove the validity of the prediction to a certain extent.

The next rule is aimed at identifying patients with other coronavirus infection with decreased expression levels of ***HK3***, ***CDKN1A***, ***HMGN3***, ***CACNA1l***, and ***ATP6V1D*** and an increased expression of ***SORT1***. As discussed above, *SORT1* has been shown to be associated with lung infections induced by multiple pathogens [60–62], including other coronavirus, thus explaining the high expression of such gene in this rule. *HK3*, which encodes the effective hexokinase 3 protein and participates in glucose metabolism pathways, has been predicted to be downregulated during coronavirus

Table 2: Gene Ontology and KEGG pathway enrichment results.

Index	Term	Category
1	SRP-dependent cotranslational protein targeting to membrane	GOTERM_BP_DIRECT
2	Viral transcription	GOTERM_BP_DIRECT
3	Nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	GOTERM_BP_DIRECT
4	Translational initiation	GOTERM_BP_DIRECT
5	Ribosome	KEGG_PATHWAY
6	Ribosome	GOTERM_CC_DIRECT
7	Translation	GOTERM_BP_DIRECT
8	Structural constituent of ribosome	GOTERM_MF_DIRECT
9	rRNA processing	GOTERM_BP_DIRECT
10	Cytosolic large ribosomal subunit	GOTERM_CC_DIRECT
11	Cytosolic small ribosomal subunit	GOTERM_CC_DIRECT
12	Poly(A) RNA binding	GOTERM_MF_DIRECT
13	Focal adhesion	GOTERM_CC_DIRECT
14	Membrane	GOTERM_CC_DIRECT
15	RNA binding	GOTERM_MF_DIRECT
16	Small ribosomal subunit	GOTERM_CC_DIRECT
17	Cytosol	GOTERM_CC_DIRECT
18	Intracellular ribonucleoprotein complex	GOTERM_CC_DIRECT
19	Extracellular exosome	GOTERM_CC_DIRECT
20	Extracellular matrix	GOTERM_CC_DIRECT
21	Ribosomal large subunit assembly	GOTERM_BP_DIRECT
22	Nucleolus	GOTERM_CC_DIRECT
23	Cytoplasmic translation	GOTERM_BP_DIRECT

infection [68], including SARS-CoV-2 infection [69]. As for the remaining four genes, *CKDN1A* has been directly reported to be positively associated with coronavirus infection and related complications [70]. Although no direct evidence confirmed the relationship between coronavirus infection and *HMGN3*, *CACNA1l*, and *ATP6V1D*, all these genes have been shown to be associated with infection-associated inflammation responses [71, 72], indicating their potential capacity for the prediction of coronavirus infections.

In rules associated with bacterial lung infections, the high expression levels of ***SORT1***, ***HK3***, and ***BAZ1A*** may be enough to identify patients with bacterial lung infections. As discussed above, a high expression of *SORT1* indicates the activation of neutrophil recruitment, which is quite common for bacterial infections [73] and different from COVID-19 infection. Meanwhile, *HK3* seems to be upregulated in lungs during bacterial infection, and such gene has been screened out as a host transcriptomic biomarker for the classification of bacteria and virus [74], thus corresponding with the prediction in the present work. Although no direct reports indicated the expression patterns of *BAZ1A* during bacterial infections, as mentioned above, neutrophil recruitment is quite common for bacterial infections.

With the involvement of effective biomarker candidates, such as ***SORT1***, ***HK3***, ***CDKN1A***, ***NLRC5***, and ***DACH1***, the next rule contributes to the identification of influenza virus infections. Similar with the previous rules, *SORT1*, *HK3*, and *CDKN1A* have been predicted to be associated with the identification of influenza virus infections. As discussed above, a high expression of *SORT1*, a downregulated *HK3*, and a high expression of *CDKN1A* are associated with viral infections [60–62, 68, 71, 72]. The upregulation of *NLRC5* and the activation of related pathways triggered by interactions between *NLRC5* and RIG-I initiate a robust antiviral response against influenza virus infection [75]. Therefore, a relatively increased expression of *NLRC5* during influenza virus infections is reasonable. As for *DACH1*, a recent comparable study [76] on COVID-19 infection, influenza virus infection, and normal controls revealed that after transcriptional regulation, the expression of *DACH1* was relatively increased in patients infected with influenza, thus validating the efficacy and accuracy of the newly presented computational methods.

Increased expression levels of ***RPL21P28*** and ***RTN1*** and a decreased expression of ***SORT1*** contribute to the rule for identifying healthy controls. The decreased expression of *SORT1* indicated no remarkable neutrophil recruitment, corresponding with the physical conditions of normal controls. *RPL21P28* has shown to be significantly differentially expressed in normal controls and tissues after infections, especially in human macrophages [77], which are generally activated during infections. Therefore, such gene could be summarized in this rule for the identification of normal controls. Similarly, *RTN1* has been shown to be associated with macrophage-mediated immune suppressants, different from the immune activators in the previously discussed rules [78], thereby validating the predictions on normal control.

4.3. Functional Enrichment Analyses Using DAVID (DAVID Bioinformatics Resources 6.8). Here, with the 168 selected COVID19 associated genes as gene of interest and all candidate genes as gene background, we performed functional enrichment analyses on GO terms and KEGG pathways using DAVID (Table 2) and selected the significant enriched results with FDR threshold as 0.05. According to the enriched results, multiple GO terms and KEGG pathways associated with RNA binding and replication via reverse transcription processes have been identified, meaning that selected genes are shown to be enriched in the RNA viral replication. Considering that COVID19 is a typical RNA virus, the enrichment results validated the reliability of the selected genes. Apart from that, we also identified multiple GO/KEGG terms associated with extracellular exosome/matrix. According to recent publications, extracellular microenvironment, especially for the vesicles outside the cells, is associated with the proliferation and spread of COVID-19 virus [79], validating our enrichment results.

All in all, the optimal blood-oriented features identified for the disease-specific diagnosis of COVID-19 and similar respiratory infectious pathogens have been validated. They are associated with their respective pathogens, and they even directly contribute to the pathogenesis according to recent publications. Therefore, the newly presented computational method in this study could be effective for the identification of COVID-19-associated biomarkers, and they could lay a solid foundation for further pathogenesis exploration on COVID-19-associated diseases.

5. Conclusion

In this study, a computational analysis was performed on an existing dataset of acute respiratory infection samples. The results included three parts. The first part was a set of genes/transcripts. They were highly related to one or more types of acute respiratory infection and can be latent biomarkers. The second part was the efficient classifiers, which can quickly identify the type of acute respiratory infection for a query sample. The third part was a set of classification rules, indicating different expression patterns on five types, giving more information to help us understand different types of acute respiratory infection.

Abbreviations

mRMR: Max-Relevance and Min-Redundancy
IFS: Incremental feature selection
RF: Random forest;
MI: Mutual information
MCC: Matthew's correlation coefficient
DT: Decision tree
SVM: Support vector machine
kNN: *K*-nearest neighbor

Data Availability

The data used to support the findings of this study have been deposited in the Gene Expression Omnibus repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE161731).

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Authors' Contributions

Lei Chen and Zhandong Li contributed equally to this work.

Acknowledgments

This research was funded by the Strategic Priority Research Program of Chinese Academy of Sciences (XDB38050200), the National Key R&D Program of China (2017YFC1201200 and 2018YFC0910403), Shanghai Municipal Science and Technology Major Project (2017SHZDZX01), the National Natural Science Foundation of China (31701151), Shanghai Sailing Program (16YF1413800), the Youth Innovation Promotion Association of Chinese Academy of Sciences (CAS) (2016245), and the Fund of the Key Laboratory of Tissue Microenvironment and Tumor of Chinese Academy of Sciences (202002).

Supplementary Materials

Supplementary 1. Table S1: 15937 features (genes) to represent each acute respiratory infection sample.

Supplementary 2. Table S2: mRMR feature list generated by mRMR method.

Supplementary 3. Table S3: performance of IFS with different classifiers.

Supplementary 4. Table S4: rules generated from DT analysis.

References

[1] K. Yuki, M. Fujiogi, and S. Koutsogiannaki, "COVID-19 pathophysiology: a review," *Clinical Immunology*, vol. 215, p. 108427, 2020.

[2] T. P. Velavan and C. G. Meyer, "The COVID-19 epidemic," *Tropical Medicine & International Health*, vol. 25, no. 3, pp. 278–280, 2020.

[3] K. Dhama, S. Khan, R. Tiwari et al., "Coronavirus disease 2019-COVID-19," *Clinical Microbiology Reviews*, vol. 33, no. 4, 2020.

[4] E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real time," *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533-534, 2020.

[5] H. A. Rothan and S. N. Byrareddy, "The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak," *Journal of Autoimmunity*, vol. 109, p. 102433, 2020.

[6] G. Pascarella, A. Strumia, C. Piliego et al., "COVID-19 diagnosis and management: a comprehensive review," *Journal of Internal Medicine*, vol. 288, no. 2, pp. 192–206, 2020.

[7] L. C. Tindale, M. Coombe, J. E. Stockdale et al., "Transmission interval estimates suggest pre-symptomatic spread of COVID-19," *MedRxiv*, 2020.

[8] H. Han, Z. Xu, X. Cheng et al., "Descriptive, retrospective study of the clinical characteristics of asymptomatic COVID-19 patients," *mSphere*, vol. 5, no. 5, 2020.

[9] M. Chiara, D. S. Horner, C. Gissi, and G. Pesole, "Comparative genomics suggests limited variability and similar evolutionary patterns between major clades of SARS-Cov-2," *BioRxiv*, 2020.

[10] D. Blanco-Melo, B. E. Nilsson-Payant, W. C. Liu et al., "Imbalanced host response to SARS-CoV-2 drives development of COVID-19," *Cell*, vol. 181, no. 5, pp. 1036–1045.e9, 2020.

[11] A. Sharma, G. Garcia Jr., Y. Wang et al., "Human iPSC-derived cardiomyocytes are susceptible to SARS-CoV-2 infection," *Cell Reports Medicine*, vol. 1, no. 4, article 100052, 2020.

[12] T. Suzuki, Y. Itoh, Y. Sakai et al., "Generation of human bronchial organoids for SARS-CoV-2 research," *BioRxiv*, 2020.

[13] Y.-H. Zhang, H. Li, T. Zeng et al., "Identifying transcriptomic signatures and rules for SARS-CoV-2 infection," *Frontiers in Cell and Developmental Biology*, vol. 8, p. 627302, 2021.

[14] M. T. McClain, F. J. Constantine, R. Henao et al., "Dysregulated transcriptional responses to SARS-CoV-2 in the periphery," *Nature Communications*, vol. 12, no. 1, p. 1079, 2021.

[15] M. Kursa and W. Rudnicki, "Feature selection with the Boruta package," *Journal of Statistical Software, Articles*, vol. 36, no. 11, pp. 1–13, 2010.

[16] Hanchuan Peng, Fuhui Long, and C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.

[17] H. A. Liu and R. Setiono, "Incremental feature selection," *Applied Intelligence*, vol. 9, no. 3, pp. 217–230, 1998.

[18] S. Zhang, T. Zeng, B. Hu et al., "Discriminating origin tissues of tumor cell lines by methylation signatures and Dysmethylated rules," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 507, 2020.

[19] S. Zhang, X. Y. Pan, T. Zeng et al., "Copy number variation pattern for discriminating MACROD2 states of colorectal cancer subtypes," *Frontiers in Bioengineering and Biotechnology*, vol. 7, p. 407, 2019.

[20] L. Chen, T. Zeng, X. Pan, Y. H. Zhang, T. Huang, and Y. D. Cai, "Identifying methylation pattern and genes associated with breast cancer subtypes," *International Journal of Molecular Sciences*, vol. 20, no. 17, p. 4269, 2019.

[21] S. He, F. Guo, Q. Zou, and HuiDing, "MRMD2.0: a Python tool for machine learning with feature ranking and reduction," *Current Bioinformatics*, vol. 15, no. 10, pp. 1213–1221, 2021.

[22] X. Zhao, L. Chen, and J. Lu, "A similarity-based method for prediction of drug side effects with heterogeneous information," *Mathematical Biosciences*, vol. 306, pp. 136–144, 2018.

[23] X. Pan, H. Li, T. Zeng et al., "Identification of protein subcellular localization with network and functional embeddings," *Frontiers in Genetics*, vol. 11, p. 626500, 2021.

[24] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," *Biochimica et Biophysica Acta (BBA) - Protein Structure*, vol. 405, no. 2, pp. 442–451, 1975.

[25] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *International joint Conference on artificial intelligence*, pp. 1137–1145, Montreal, QC, Canada, 1995.

[26] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[27] Z. B. Lv et al., "RF-PseU: a random forest predictor for RNA Pseudouridine sites," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 10, 2020.

[28] L. Xu, G. Liang, C. Liao, G.-D. Chen, and C.-C. Chang, "k-Skip-n-Gram-RF: a random forest based method for Alzheimer's disease protein identification," *Frontiers in Genetics*, vol. 10, p. 7, 2019.

[29] Y. Jia, R. Zhao, and L. Chen, "Similarity-based machine learning model for predicting the metabolic pathways of compounds," *IEEE Access*, vol. 8, pp. 130687–130696, 2020.

[30] H. Liang, L. Chen, X. Zhao, and X. Zhang, "Prediction of drug side effects with a refined negative sample selection strategy," *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 1573543, 16 pages, 2020.

[31] X. Y. Pan, T. Zeng, Y. H. Zhang et al., "Investigation and prediction of human interactome based on quantitative features," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 730, 2020.

[32] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[33] J. Li, L. Lu, Y. H. Zhang et al., "Identification of leukemia stem cell expression signatures through Monte Carlo feature selection strategy and support vector machine," *Cancer Gene Therapy*, vol. 27, no. 1-2, pp. 56–69, 2020.

[34] J.-P. Zhou, L. Chen, T. Wang, and M. Liu, "iATC-FRAKEL: a simple multi-label web server for recognizing anatomical therapeutic chemical classes of drugs with their fingerprints only," *Bioinformatics*, vol. 36, no. 11, pp. 3568-3569, 2020.

[35] H. Liu, B. Hu, L. Chen, and L. Lu, "Identifying protein subcellular location with embedding features learned from networks," *Current Proteomics*, vol. 17, 2020.

[36] L. Chen, S. Wang, Y. H. Zhang et al., "Identify key sequence features to improve CRISPR sgRNA efficacy," *IEEE Access*, vol. 5, pp. 26582–26590, 2017.

[37] L. Chen, X. Y. Pan, W. Guo et al., "Investigating the gene expression profiles of cells in seven embryonic stages with machine learning algorithms," *Genomics*, vol. 112, no. 3, pp. 2524–2534, 2020.

[38] Y. Zhu, B. Hu, L. Chen, and Q. Dai, "iMPTCE-Hnetwork: A Multilabel Classifier for Identifying Metabolic Pathway Types of Chemicals and Enzymes with a Heterogeneous Network," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 6683051, 12 pages, 2021.

[39] C. Meng, F. Guo, and Q. Zou, "CWLy-SVM: a support vector machine-based tool for identifying cell wall lytic enzymes," *Computational Biology and Chemistry*, vol. 87, p. 107304, 2020.

[40] M. Tahir and A. Idris, "MD-LBP: an efficient computational model for protein subcellular localization from HeLa cell lines using SVM," *Current Bioinformatics*, vol. 15, no. 3, pp. 204–211, 2020.

[41] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.

[42] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 3, pp. 660–674, 1991.

[43] Y.-H. Zhang, T. Zeng, L. Chen, T. Huang, and Y. D. Cai, "Determining protein-protein functional associations by functional rules based on gene ontology and KEGG pathway," *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1869, no. 6, article 140621, 2021.

[44] T. Sawa, S. Kimura, N. H. Honda et al., "Diagnostic usefulness of ribosomal protein L7/L12 for pneumococcal pneumonia in

a mouse model," *Journal of Clinical Microbiology*, vol. 51, no. 1, pp. 70–76, 2013.

[45] E. E. Davenport, R. D. Antrobus, P. J. Lillie, S. Gilbert, and J. C. Knight, "Transcriptomic profiling facilitates classification of response to influenza challenge," *Journal of Molecular Medicine*, vol. 93, no. 1, pp. 105–114, 2015.

[46] C. Su, S. Rousseau, and A. Emad, "Identification of COVID-19-relevant transcriptional regulatory networks and associated kinases as potential therapeutic targets," *bioRxiv*, 2020.

[47] J. Merino-Gracia, M. F. García-Mayoral, and I. Rodríguez-Crespo, "The association of viral proteins with host cell dynein components during virus infection," *The FEBS Journal*, vol. 278, no. 17, pp. 2997–3011, 2011.

[48] J. Rossjohn, S. Gras, J. J. Miles, S. J. Turner, D. I. Godfrey, and J. McCluskey, "T cell antigen receptor recognition of antigen-presenting molecules," *Annual Review of Immunology*, vol. 33, no. 1, pp. 169–200, 2015.

[49] L. Rowen, B. F. Koop, and L. Hood, "The complete 685-kilobase DNA sequence of the human beta T cell receptor locus," *Science*, vol. 272, no. 5269, pp. 1755–1762, 1996.

[50] M. J. Corley, C. Sugai, M. Schotsaert, R. E. Schwartz, and L. C. Ndhlovu, "Comparative in vitro transcriptomic analyses of COVID-19 candidate therapy hydroxychloroquine suggest limited immunomodulatory evidence of SARS-CoV-2 host response genes," *bioRxiv*, 2020.

[51] D. Polineni, H. Dang, P. J. Gallins et al., "Airway mucosal host defense is key to genomic regulation of cystic fibrosis lung disease severity," *American Journal of Respiratory and Critical Care Medicine*, vol. 197, no. 1, pp. 79–93, 2018.

[52] A. Anirudhan, P. I. Angulo-Bejarano, P. Paramasivam et al., "RPL6: a key molecule regulating zinc- and magnesium-bound metalloproteins of Parkinson's disease," *Frontiers in Neuroscience*, vol. 15, p. 631892, 2021.

[53] Q. Wu, Y. Gou, Q. Wang et al., "Downregulation of RPL6 by siRNA inhibits proliferation and cell cycle progression of human gastric cancer cell lines," *PLoS One*, vol. 6, no. 10, article e26401, 2011.

[54] G. Nallur, "Interaction of the SARS-COV2 envelope protein (E) with lysophosphatidic acid receptor 1 (LPAR1) and additional human proteins involved in inflammation, immunity, ADP ribosylation and vesicular transport. Immunity, ADP Ribosylation and Vesicular Transport," *SSRN Electronic Journal*, 2020.

[55] J. A. Browne, M. NandyMazumdar, A. Paranjapye, S. H. Leir, and A. Harris, "The Bromodomain Containing 8 (BRD8) Transcriptional Network in Human Lung Epithelial Cells," *Molecular and Cellular Endocrinology*, vol. 524, article 111169, 2021.

[56] B. Wanschers, R. van de Vorstenbosch, M. Wijers, B. Wieringa, S. M. King, and J. Fransen, "Rab6 family proteins interact with the dynein light chain protein DYNLRB1," *Cell Motility and the Cytoskeleton*, vol. 65, no. 3, pp. 183–196, 2008.

[57] M. Alhariri, M. A. Majrashi, A. H. Bahkali et al., "Efficacy of neutral and negatively charged liposome-loaded gentamicin on planktonic bacteria and biofilm communities," *International Journal of Nanomedicine*, vol. Volume 12, pp. 6949–6961, 2017.

[58] L. M. Al-Dalawi, *Effect of lipids on the infectivity of influenza A viruses*, University of Nottingham, 2019.

[59] E. Lebredonchel, M. Houdou, H. H. Hoffmann et al., "Investigating the functional link between TMEM165 and SPCA1," *Biochemical Journal*, vol. 476, no. 21, pp. 3281–3293, 2019.

[60] C. L. Vázquez, A. Rodgers, S. Herbst et al., "The proneurotrophin receptor sortilin is required for _Mycobacterium tuberculosis_ control by macrophages," *Scientific Reports*, vol. 6, no. 1, article 29332, 2016.

[61] Z. Zeng, H. B. Huang, L. L. Huang et al., "Regulation network and expression profiles of Epstein-Barr virus-encoded microRNAs and their potential target host genes in nasopharyngeal carcinomas," *Science China Life sciences*, vol. 57, no. 3, pp. 315–326, 2014.

[62] J. Ma, C. Chen, A. S. Barth, C. Cheadle, X. Guan, and L. Gao, "Lysosome and cytoskeleton pathways are robustly enriched in the blood of septic patients: a meta-analysis of transcriptomic data," *Mediators of Inflammation*, vol. 2015, Article ID 984825, 15 pages, 2015.

[63] A. M. Alshabi, I. A. Shaikh, B. M. Vastrad, and C. M. Vastrad, "Identification of differentially expressed genes and enriched pathways in SARS-CoV-2/COVID-19 using bioinformatics analysis," *Research Square*, 2020.

[64] S. Di Giorgio, F. Martignano, M. G. Torcia, G. Mattiuz, and S. G. Conticello, "Evidence for host-dependent RNA editing in the transcriptome of SARS-CoV-2," *Science Advances*, vol. 6, no. 25, article eabb5813, 2020.

[65] J. Sun, F. Ye, A. Wu et al., "Comparative transcriptome analysis reveals the intensive early stage responses of host cells to SARS-CoV-2 infection," *Frontiers in Microbiology*, vol. 11, p. 2881, 2020.

[66] L. Hua, S. Yao, D. Pham et al., "Cytokine-dependent induction of CD4+ T cells with cytotoxic potential during influenza virus infection," *Journal of Virology*, vol. 87, no. 21, pp. 11884–11893, 2013.

[67] B. J. Meckiff, C. Ramírez-Suástegui, V. Fajardo et al., "Single-cell transcriptomic analysis of SARS-CoV-2 reactive $CD4^+$ T cells," *SSRN Electronic Journal*, 2020.

[68] S. Miyamoto, A. N. Murphy, and J. H. Brown, "Akt mediates mitochondrial protection in cardiomyocytes through phosphorylation of mitochondrial hexokinase-II," *Cell Death & Differentiation*, vol. 15, no. 3, pp. 521–529, 2008.

[69] I. Ortea and J.-O. Bock, "Re-analysis of SARS-CoV-2 infected host cell proteomics time-course data by impact pathway analysis and network analysis. A potential link with inflammatory response," *BioRxiv*, 2020.

[70] J. Cinatl Jr., G. Hoever, B. Morgenstern et al., "Infection of cultured intestinal epithelial cells with severe acute respiratory syndrome coronavirus," *Cellular and Molecular Life Sciences CMLS*, vol. 61, no. 16, pp. 2100–2112, 2004.

[71] Y. Xia, N. Liu, X. Xie et al., "The macrophage-specific V-ATPase subunit ATP6V0D2 restricts inflammasome activation and bacterial infection by facilitating autophagosome-lysosome fusion," *Autophagy*, vol. 15, no. 6, pp. 960–975, 2019.

[72] D. Cornblath, "DS3. 1 neuromuscular manifestations of HIV infection," *Clinical Neurophysiology*, vol. 117, pp. 21–21, 2006.

[73] I. E. Galani and E. Andreakos, "Neutrophils in viral infections: current concepts and caveats," *Journal of Leukocyte Biology*, vol. 98, no. 4, pp. 557–564, 2015.

[74] T. E. Sweeney, H. R. Wong, and P. Khatri, "Robust classification of bacterial and viral infections via integrated host gene expression diagnostics," *Science Translational Medicine*, vol. 8, no. 346, p. 346ra91, 2016.

[75] P. Ranjan, N. Singh, A. Kumar et al., "NLRC5 interacts with RIG-I to induce a robust antiviral response against influenza virus infection," *European Journal of Immunology*, vol. 45, no. 3, pp. 758–772, 2015.

Hindawi
BioMed Research International
Volume 2024, Article ID 9898173, 1 page
https://doi.org/10.1155/2024/9898173

Retraction

Retracted: Hydroxyphenyl Butanone Induces Cell Cycle Arrest through Inhibition of GSK3β in Colorectal Cancer

BioMed Research International

Received 8 January 2024; Accepted 8 January 2024; Published 9 January 2024

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Zhang, Y. Wang, H. Zhang, C. Sun, S. Dang, and M. Liu, "Hydroxyphenyl Butanone Induces Cell Cycle Arrest through Inhibition of GSK3β in Colorectal Cancer," *BioMed Research International*, vol. 2021, Article ID 9981815, 13 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 9981815, 13 pages
https://doi.org/10.1155/2021/9981815

Research Article

Hydroxyphenyl Butanone Induces Cell Cycle Arrest through Inhibition of GSK3β in Colorectal Cancer

Songyan Zhang,[1] Yunfeng Wang,[2] Haopeng Zhang,[2] Chengming Sun,[1] Shuwei Dang,[2] and Ming Liu[1]

[1]*Department of General Surgery, The Second Affiliated Hospital of Harbin Medical University, Harbin 150086, China*
[2]*Department of General Surgery, The Fourth Affiliated Hospital of Harbin Medical University, Harbin 150001, China*

Correspondence should be addressed to Ming Liu; mingliu35@hrbmu.edu.cn

Received 1 April 2021; Revised 30 April 2021; Accepted 15 June 2021; Published 3 July 2021

Academic Editor: Min Tang

Background. Colorectal cancer (CRC) is among the top three gastrointestinal malignancy in morbidity and mortality. The abnormal activation of Wnt/β-catenin pathway is considered to be a key factor in the occurrence and development of CRC. Novel inhibitor discovery against key factor in WNT pathway is important for CRC treatment and prevention. *Methods*. Cell proliferation was detected after hydroxyphenyl butanone treatment in human colorectal cancer HCT116, LOVO, and normal colonic epithelial NCM460 cells. Colony formation, cell invasion ability, and cell cycle were detected with and without GSK-3β knockdown. *Results*. Hydroxyphenyl butanone induces cycle arresting on G1-S phase of colorectal cancer cell line through GSK3β in Wnt/β-catenin pathway and inhibits malignant biological manifestations of cell proliferation, colony formation, and invasion. The inhibition in the high concentration group is stronger than that in the low concentration group, and the antitumor effect is different for different tumor cells. Under the same concentration of natural hydroxyphenyl butanone, the inhibition on normal colonic epithelial cells is significantly lower than that on tumor cells. The natural hydroxyphenyl butanone with medium and low concentration could promote the proliferation of normal colonic epithelial cells. *Conclusion*. This study illustrated natural hydroxyphenyl butanone as new inhibitor of GSK3β and revealed the mechanisms underlying the inhibitory effects in colorectal cancer.

1. Introduction

Colorectal cancer (CRC) is one of the most common gastrointestinal malignancy. The morbidity and mortality of CRC are among the top three malignant tumors [1–3]. Although great improvement has been achieved in diagnosis and treatment of CRC, the 5-year survival rate of CRC is still too low, and about 47.2% CRC patients showed recurrence of tumor. Recent studies have suggested that the development of CRC has a close relationship with many pathways, such as Wnt/β-catenin, EGFR, p53, MAPK, Notch, and RhoA/ROCK, among which Wnt/β-catenin pathway might be most important. The abnormal activation of Wnt/β-catenin pathway is considered to be a key factor in the occurrence and development of CRC [4]. In addition, Tan et al. proved that the proliferation and apoptosis of CRC cells can be inhibited by the inactivation of Wnt/β-catenin pathway [5].

Wnts are a series of glycoproteins related to stimulation of receptor-mediated signal transduction pathways. The enrichment of nuclear β-catenin is one of the most critical factors to activate Wnt pathway. After targeted for phosphorylation of the N-terminal by glycogen synthase kinase 3β (GSK-3β) at Ser33/Ser37/Tr41 residues, the cytoplasmic level of β-catenin is regulated by ubiquitin-mediated proteolysis. The phosphorylation of β-catenin by GSK-3β occurs in an Axin-APC complex [6, 7], during which GSK-3β activity presents a negative correlation with the presence of Wnt pathway [8]. *APC* is the key regulator of Wnt pathway, and mutations of *APC* are observed in at least 70% of CRCs, especially in early colonic neoplasia [9, 10]. In addition, Ashok et al. found that luteolin inhibited HCT-15 cell proliferation by downregulation of the inactive form of GSK-3β [11]. However, though there are many studies on Wnt/β-catenin pathway, its exact role in CRCs still needs more investigation.

With the increasing of old population [12, 13], the demand for low-toxic and high-efficient antitumor treatment methods is increasing. According to statistics data [14], from 1940 to 2014, more than 75% of anticancer drugs directly came from natural products. Therefore, it is of great significance to search for new effective anticancer drugs from natural products to improve the prognosis of CRC patients. Recent research also revealed that many kinds of extracts from plants (in different parts, such as barks, leaves, and fruits) have the ability of anticolon cancer cells *in vivo* or *in vitro* [15–17]. Zhang et al. showed that azoxymethane-induced upregulation of WNT/β-catenin pathway was prevented by genistein, which reduced colon preneoplasia in rats [18]. Reabroi et al. showed that the anticancer activity of 19-O-triphenylmethyl andrographolide involves the inhibition of Wnt/β-catenin signaling via a GSK-3β-independent pathway [19]. Aisha et al. found that xanthone extracts caused significant growth inhibition of HCT 116 colorectal carcinoma cells by upregulating the MAPK/ERK, c-Myc/Max, and p53 pathways [20].

Red raspberry (*Rubus idaeus L.*) is a kind of economical berry crop that contains numerous potential healthy compounds, such as polyphenols, flavonoids, kaurane-type diterpenoids glycosides, oleanane, sodium-type triterpenoids, glycosides, and esters [21]. Related studies have shown that red raspberry phytochemical extracts [22] had many pharmacological effects. Raspberry ketone had the ability of increase norepinephrine-induced lipolysis in white adipocytes, which could prevent and improve obesity and fatty liver [23]. The extract of red raspberry fruit showed anti-inflammatory properties to reduce the degree of bone resorption, soft tissue swelling, and osteophyte formation, preventing articular destruction in treated animals [24]. Wang et al. tested the protective effect of raspberry ketone against nonalcoholic steatohepatitis and confirmed its function in protecting kidney function, weight loss, and antialcoholic steatohepatitis [25]. Meanwhile, research on the biological activity of red raspberry leaf polyphenols showed certain effects on inhibiting human laryngeal carcinoma (HEp2) and colon adenocarcinoma (SW 480) cell lines [26]. The phytochemical extracts of red raspberry had been proved to have certain preventive effect on many kinds of malignant tumors, such as esophageal cancer [27], breast cancer [28], cervix cancer [29], and prostate cancer [30]. Experiments showed that the increasing doses of red raspberry extract could lead to inhibition of hepatocellular carcinoma [31]. Black raspberry extract had the ability of reducing acute UVB-induced inflammation which could help to prevent human skin cancers [32]. David et al. reported the anti-inflammatory effect of raspberry on DSS-induced colonic injury, which is considered to be one of the precancerous lesions of CRC [33]. Results showed a significantly increase preventive effect of CRC and ulcerative colitis. Therefore, it is speculated that the phytochemical extract of red raspberry may have broad application prospects in CRC prevention and treatment.

Red raspberry phytochemical extract is the active component of red raspberry fruit extraction. In our previous studies, we identified the extraction method and found that hydroxyphenyl butanone is the main component of red raspberry phytochemical extract, which can play a variety of antitumor effects. In this study, we aimed at revealing natural inhibitor in colorectal cancer and discovering the biochemical mechanisms of natural hydroxyphenyl butanone.

2. Methods

2.1. Preparation of Natural Hydroxyphenyl Butanone. The 4-p-hydroxyphenyl-2-butanone (pure raspberry ketone) was purified by the School of Chemical Engineering, Harbin University of Technology. For natural hydroxyphenyl butanone, fresh red raspberry was purchased from Shangzhi City, Heilongjiang Province. In the preliminary study of our research group, 100 g red raspberry/300 ml absolute ethanol (1 : 3, w/V) was mixed, homogenized by polyteron homogenizer, vacuum filtered, and evaporated at 45°C to extract the chemical extract of red raspberry. The main component was determined as raspberry ketone (4-p-hydroxyphenyl-2-butanone) by HPLC, ellagic acid, and polysaccharide. This method takes into account the effectiveness of chemical extracts of red raspberry plants and the safety of long-term human application after extraction and has been applied to the scientific research on the antihuman hepatocarcinoma effect of chemical extracts of red raspberry plants and published research papers [34, 35].

2.2. Cell Culture. Human colorectal cancer HCT116, LOVO, and normal colonic epithelial NCM460 cell line (from the School of Life Science and Technology, Harbin Institute of Technology) was used firstly. Then, cell culture was performed by using 1640 medium (Gibco) supplemented with 10% serum (Gibco, Grand Island, NY, USA) at 37°C-5% CO_2 incubator.

2.3. CCK8 Method for Measuring Cell Proliferation. One thousand and fifty hundred cells/well in 96-well plates were plated into cells. After adhering to 200 μl 1640 medium, the medium was aspirated at 0 h, 24 h, 48 h, 72 h, and 96 h, and 100 μl of medium containing 10% CCK8 was added to each well. The OD value was read at a wavelength of 450 nm by a microplate reader for 4 hours. The 0 h OD value was used as the baseline, and the ratio of OD value to 0 h OD value at each time point was the cell proliferation multiplication. The cell proliferation multiple of the control group was base 1 at each time point, and the inhibition rate of each group was equal to one minus the proliferation ratio of each group/-the proliferation ratio of the control group.

2.4. Colony Formation and Cell Invasion Ability Test. The cells were seeded at 400 cells/well in a 6-well plate and routinely cultured until each single cell clone was more than 50 cells, and crystal violet was fixed. The number of single cell clones was counted. Transwell plate (3428, costar) was evenly spread in the upper chamber, and 3×10^5 cells/well were plated into the cells, and 1.5 ml of 1640 medium containing 0.5% serum was added, and 2.6 ml of 1640 medium containing 10% serum was added to the lower chamber. The cells were photographed under a crystal violet staining

microscope after incubation at 37°C-5% for 24 hours. Four field of view counts were randomly selected in each group.

2.5. Cell Cycle Detection and Immunofluorescence. The cells cultured in adherent cells were trypsinized without EDTA, and then, the cells were collected by centrifugation, fixed in 75% ethanol at 4°C for 24 hours. PI stained for 30 min, and the cell cycle was measured by flow cytometry. After adhering to the cells on the glass plate, 4% formaldehyde was fixed, 10% triton-X100 was permeated, 5% BSA was blocked for one hour, and 0.5% BSA diluted primary antibody (#8480S, cell signaling, 1 : 100 dilution) was incubated at 37°C for 2 hours; the secondary antibody diluted with 0.5% PBS was incubated at 37°C for 30 minutes, and the nuclei were counterstained with 0.1% diluted DAPI for 5 minutes. The antifluorescence quencher was mounted and photographed under a fluorescence microscope.

2.6. Key Factor Protein Level Detection. The cells were counted who has the same number, and total protein was extracted by adding RIPA lysate at a concentration of 1×10^6 cells/100 μl, electrophoresed on SDS-acrylamide gel, and blotted onto a 0.2 μm PVDF membrane. After 1 hour when 5% skim milk was sealed, 0.1% PBST was washed and then immersed in 0.1% PBST diluted primary antibody overnight; 0.1% PBST was washed, immersed in secondary antibody for 1 hour; 0.1% PBST was washed, and developer was added dropwise. Then, the ECL Western Blot Analysis System (Amersham Biosciences, Chalfont St. Giles, Buckinghamshire, UK) was involved to perform the parallel grayscale analysis.

2.7. Construction of Transfected Cell Lines with Downregulated Gene Expression. siRNA (Suzhou Gemma Gene) sequence is as follows:

Control group (NC)

Justice chain (5′-3′): UUCUCCGAACGUGUCACGU TT

Antisense strand (5′-3′): ACGUGACACGUUCGGA GAATT

GSK-3β downregulation gene group (GSK3β-SI)

Justice chain (5′-3′): GGACUAUGUUCCGGAAACA TT

Antisense strand (5′-3′): UGUUUCCGGAACAUAG UCCTT

Human colorectal cancer HCT116 and LOVO cells were evenly plated into 6-well plates and cultured for 24 hours. The cell density was about 50%. Premixed transfection solution (containing 75 pmol of siRNA, 7.5 μl of lipofectamin 2000, and 200 μl of Opti-MEM medium) was added to each well. Further experiments were carried out after incubation at 37°C-5% CO_2 for 48 hours.

2.8. RT-PCR. The adherent cultured cell samples were extracted with 1 μg of total RNA by using TRIzol, reversely transcribed into cDNA by reverse transcription kit and subjected to RT-PCR detection using a ViiATM7 real-time PCR instrument (Applied Biosystems).

GSK-3β gene primer sequence is as follows:

Table 1: Half inhibition concentration.

	Half inhibitory concentration	
	HCT116	LOVO
Red raspberry phytochemical extract	27 mg/ml	17 mg/ml
Natural hydroxyphenyl butanone	190 μg/ml	174 μg/ml
Fluorouracil	0.7 μg/ml	0.4 μg/ml

Upstream 5′-3′: AGACGCTCCCTGTGATTTATGT
Downstream 5′-3′: CCGATGGCAGATTCCAAAGG
GAPDH primer sequence is as follows:
Upstream 5′-AGGCTGTGGGCAAGGTCATC-3″
Downstream 5′-TCAGGTCCACCACTGACACG-3′

2.9. Positive Control. Fluorouracil (5-fluoro-2,4(1H,3H)-pyrimidine dione, 5-FU) is a classic antitumor drug for colorectal cancer. At present, most of the chemotherapy regimens for large intestine are centered on fluorouracil or drugs that can produce fluorouracil through in vivo transformation. And it is also the commonly used positive control drug in antitumor research of colorectal cancer [3]. In this study, fluorouracil was used as the positive control of antitumor effect, and pure raspberry ketone was used as the positive control of antitumor mechanism to study the action and mechanism of red raspberry phytochemical extract on colorectal cancer cells.

3. Results

3.1. Inhibition of Natural Hydroxyphenyl Butanone on Colorectal Cancer Cell Proliferation, Colony Formation, and Cell Invasion Ability. The proliferation inhibition effects of natural hydroxyphenyl butanone, raspberry ketone, and fluorouracil were detected by CCK8 method at different times and different concentrations in colorectal cancer HCT116 and LOVO cell lines. The optimal half-inhibitory concentration of the two colorectal cancer cell lines for 72 hours culturing was determined by gradient inhibition concentration experiments and calculated by the CompuSyn software (Table 1).

The calculated half-inhibitory concentrations and 1/2 and 1/4 half-inhibited concentrations of the different red raspberry plant chemical extracts, treated for 72 hours, were calculated as the high, medium, and low experimental group concentrations according to specific colorectal cancer cell line. The half inhibitory concentration of raspberry ketone and fluorouracil, treated for 72 hours, applied as the concentration of the positive control group (Table 2).

The cell proliferation of colorectal cancer HCT116, LOVO and normal human colonic epithelial cell line NCM460 was measured at the experimental concentration. The calculated inhibition rate was as follows: 48.60 ± 6.10% /9.62 ± 5.40% for the HCT116/NCM460 high-dose raspberry group, 21.78 ± 3.79%/−2.96 ± 5.50% for the middle dose raspberry group, and 12.15 ± 3.14%/−101.30 ± 12.28% for the low-dose raspberry group, while 53.65 ± 4.60%/21.45 ± 8.57% for the raspberry ketone group and 46.48 ± 8.59%/

Table 2: Concentration of the experimental group and positive control group for 72 hours.

Groups	HCT 116	LOVO
Red raspberry extract high concentration (HRB)	27 mg/ml	17 ml/ml
Red raspberry extract medium concentration (MRB)	13.5 ml/ml	8.5 mg/ml
Red raspberry extract low concentration (LRB)	6.25 mg/ml	4.25 mg/ml
Ketone (PRB)	190 μg/ml	174 μg/ml
Fluorouracil (5-Fu)	0.7 μg/ml	0.4 μg/ml

12.06 ± 6.19% in the fluorouracil group. For LOVO/NCM460 cell lines, the results were shown as the high-dose raspberry group 54.36 ± 2.65%/−1.48 ± 5.16%, medium-dose raspberry group 15.55 ± 3.75%/−8.31 ± 8.91%, low-dose raspberry group 6.11 ± 4.00%/−52.62 ± 10.36%, raspberry ketone group 50.39 ± 2.81%/16.37 ± 6.54%, and fluorouracil group 51.64 ± 6.92%/8.69 ± 5.19%. Our analysis results showed that both of the experimental group and the positive control group have an inhibitory effect on the proliferation of colorectal cancer cell lines, and the inhibitory effect is positively correlated with the concentration of the experimental group. The experimental group and the positive control group had higher inhibitory effect on the proliferation of colorectal cancer cell lines than on normal colon cell lines. The low and medium concentration experimental group could promote the proliferation of normal colon cell lines (Figure 1(a)).

After adding experimental concentration of red raspberry phytochemical extract and raspberry ketone positive control to the medium, the number of colony formation was as follows: HCT116 control group 27 ± 3.37, high-dose raspberry group 10.5 ± 2.38, medium-dose raspberry group 20.75 ± 2.22, low-dose raspberry group 27.75 ± 3.86, raspberry ketone group 11 ± 2.16, and fluorouracil group 12 ± 2.93. For LOVO cell lines are as follows: control group 34 ± 5.42, high-dose raspberry group 8.25 ± 0.96, medium-dose raspberry group 22.25 ± 1.50, low-dose raspberry group 29.75 ± 4.03, raspberry ketone group 11.25 ± 2.99, and fluorouracil group 12.25 ± 3.86. The colony forming ability of HCT116 and LOVO cells decreased in both of the experimental group and the positive control group, and the degree of decrease was positively correlated with the concentration of the experimental group. The colony forming ability of the low concentration experimental group was close to or equal to control group (Figure1(b)). Proliferation of colorectal cancer HCT116, LOVO, and normal human colonic epithelial cell line NCM460 was measured at the experimental concentration. The calculated inhibition rate was as follows: 48.60 ± 6.10%/9.62 ± 5.40% for the HCT116/NCM460 high-dose raspberry group, 21.78 ± 3.79%/−2.96 ± 5.50% for the middle dose raspberry group, and 12.15 ± 3.14%/−101.30 ± 12.28% for the low-dose raspberry group, while 53.65 ± 4.60 %/21.45 ± 8.57% for the raspberry ketone group and 46.48 ± 8.59%/12.06 ± 6.19% in the fluorouracil group. For LOVO/NCM460 cell lines, the results were shown as the high-dose raspberry group 54.36 ± 2.65%/−1.48 ± 5.16%, medium-dose raspberry group 15.55 ± 3.75%/−8.31 ± 8.91%, low-dose raspberry group 6.11 ± 4.00%/−52.62 ± 10.36%, raspberry ketone group 50.39 ± 2.81%/16.37 ± 6.54%, and fluorouracil group 51.64 ± 6.92%/8.69 ± 5.19%. Our analysis results showed that both of the experimental group and the positive control group have an inhibitory effect on the proliferation of colorectal cancer cell lines, and the inhibitory effect is positively correlated with the concentration of the experimental group. The experimental group and the positive control group had higher inhibitory effect on the proliferation of colorectal cancer cell lines than on normal colon cell lines. The low and medium concentration experimental group could promote the proliferation of normal colon cell lines (Figure 1(a)). After adding experimental concentration of red raspberry phytochemical extract and raspberry ketone positive control to the medium, the number of colony formation was as follows: HCT116 control group 27 ± 3.37, high-dose raspberry group 10.5 ± 2.38, medium-dose raspberry group 20.75 ± 2.22, low-dose raspberry group 27.75 ± 3.86, raspberry ketone group 11 ± 2.16, and fluorouracil group 12 ± 2.93. For LOVO cell lines are as follows: control group 34 ± 5.42, high-dose raspberry group 8.25 ± 0.96, medium-dose raspberry group 22.25 ± 1.50, low-dose raspberry group 29.75 ± 4.03, raspberry ketone group 11.25 ± 2.99, and fluorouracil group 12.25 ± 3.86. The colony forming ability of HCT116 and LOVO cells decreased in both of the experimental group and the positive control group, and the degree of decrease was positively correlated with the concentration of the experimental group. The colony forming ability of the low concentration experimental group was close to or equal to the control group (Figure 1(b)).

The transwell experimental cell invasion rate was detected in colon cancer HCT116 and LOVO cells with red raspberry plant chemical extracts or raspberry ketone treatment. The results of transwell were shown as HCT116 control group 27.00 ± 3.67%, high-dose raspberry group 10.50 ± 2.38%, medium-dose raspberry group 20.75 ± 2.22%, low-dose raspberry group 27.75 ± 3.86%, raspberry ketone group 11.00 ± 2.16%, and fluorouracil group 12.00 ± 2.94%; and LOVO control group 91.75 ± 14.41%, high-dose raspberry group 23.75 ± 5.12%, medium-dose raspberry group 52.00 ± 5.71%, low-dose raspberry group 76.00 ± 8.40%, raspberry ketone group 24.25 ± 3.11%, and fluorouracil group 27.25 ± 3.96%. The lower cell invasion ability was detected in the experimental group and the positive control group, and the degree of weakening was positively correlated with the concentration of the experimental group. The cell invasion ability of the low concentration experimental group was close to or equal to that of the control group (Figure 1(c)).

3.2. Natural Hydroxyphenyl Butanone Blocked G1-S Phase of Colorectal Cancer Cell Cycle through WNT Pathway. Applying qualitative analysis of immunofluorescence for detecting

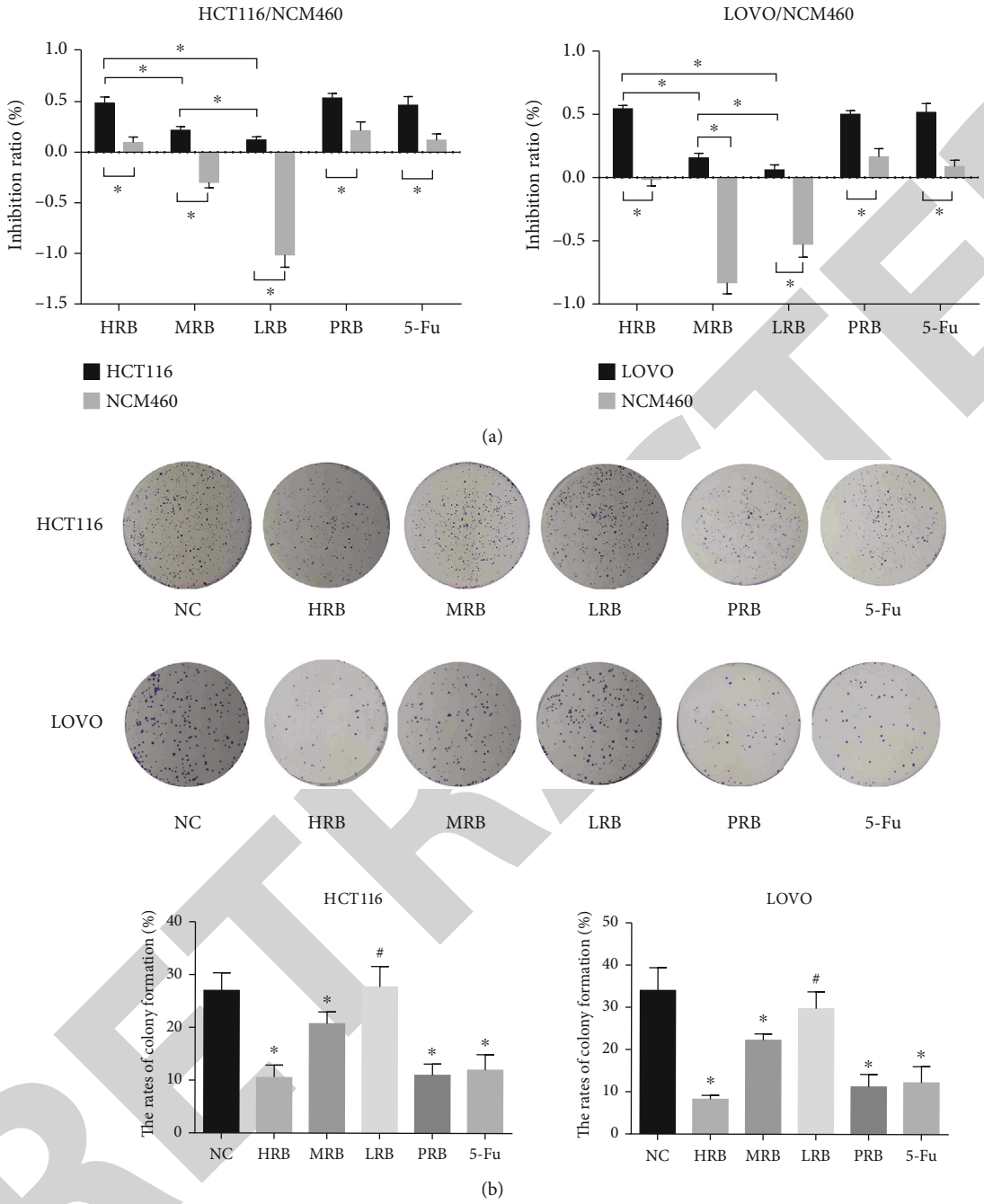


Figure 1: Continued.

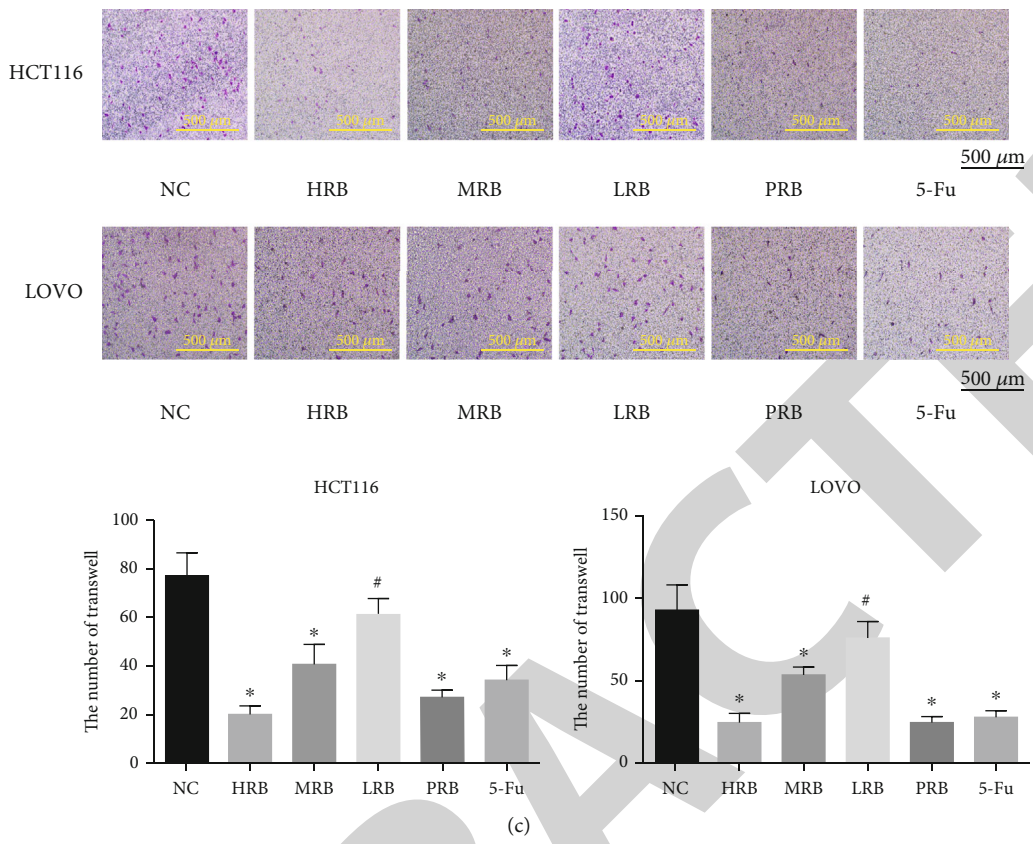
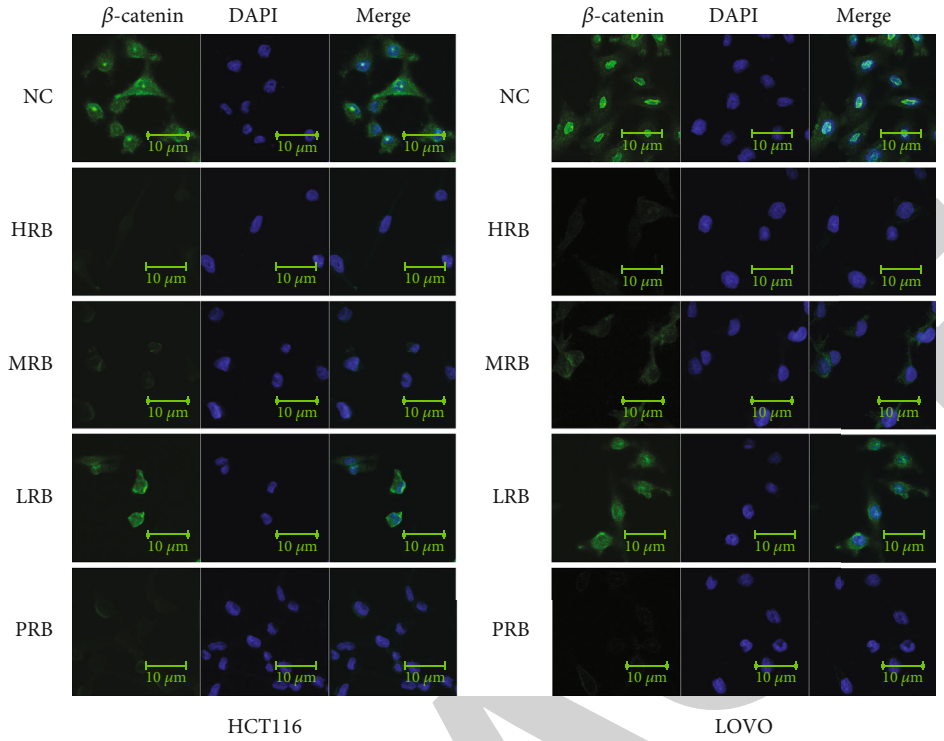


Figure 1: Effect of natural hydroxyphenyl butanone on proliferation, colony formation, and invasion of colorectal cancer HCT116 and LOVO cells. (a) Inhibition rate of chemical extracts of red raspberry with different concentrations, raspberry ketone, and fluorouracil on colorectal cancer HCT116, LOVO cells, and normal colonic epithelial NCM460 cells. (b) The effects of the chemical extracts of red raspberry with different concentrations, raspberry ketone, and fluorouracil on the colony formation of colorectal cancer HCT116 and LOVO cells. (c) Effects of different concentrations of red raspberry phytochemical extracts, raspberry ketone, and fluorouracil on colorectal cancer HCT116 and LOVO cell invasion ability. $^{*}p < 0.05$; $\#p > 0.05$.

the expression of WNT pathway key proteins, β-catenin showed decreasing expression in the experimental group and positive control group after red raspberry phytochemical extract and raspberry ketone treatment, while β-catenin protein plays a regulatory role in the nucleus by binding to downstream-related nuclear factors. Furthermore, the distribution of β-catenin protein transferred from cytoplasm and nucleus in the control group became mainly in the cytoplasm, and the distribution in the nucleus was significantly reduced (Figure 2(a)).

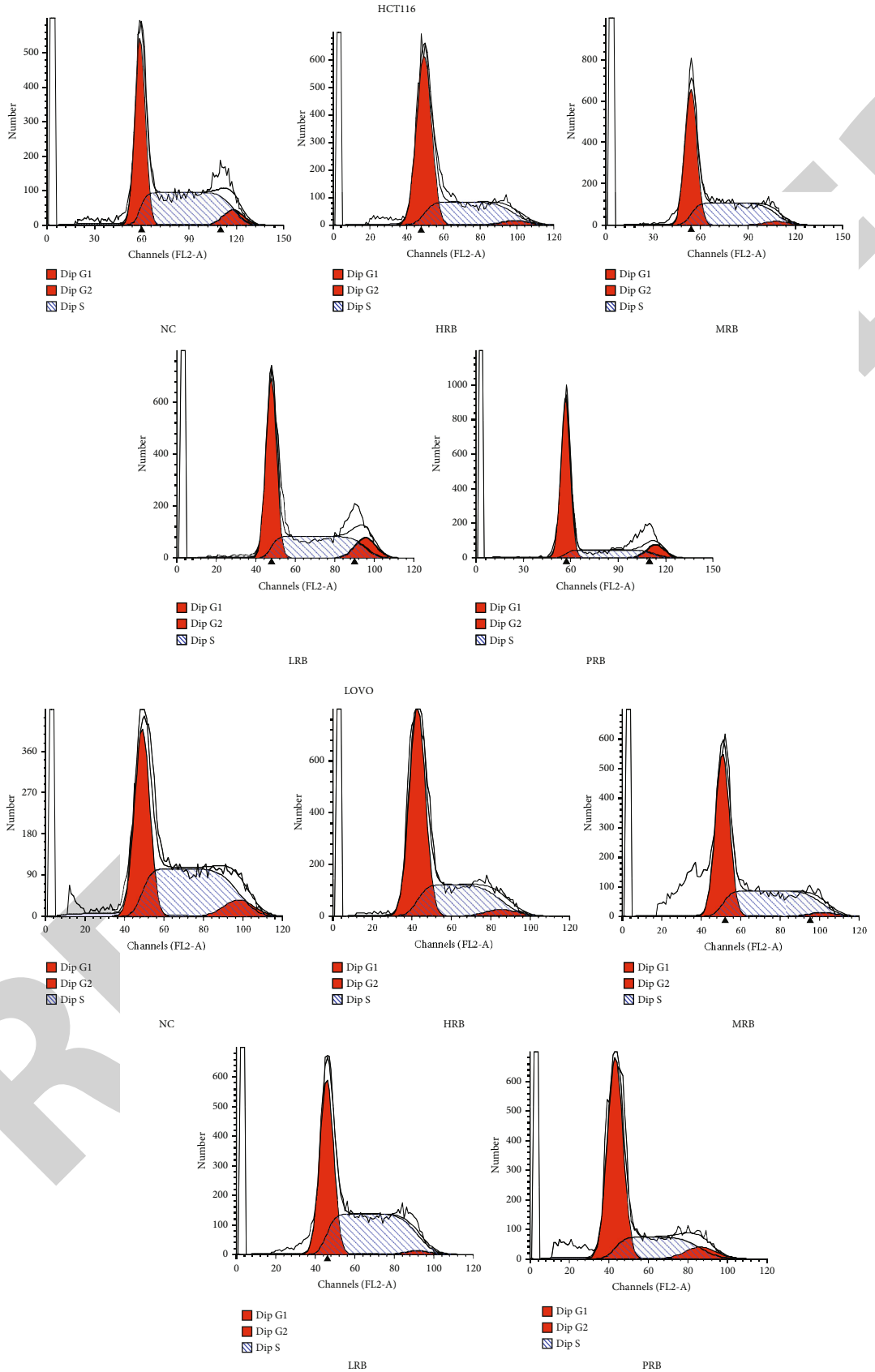
Flow cytometry was applied to determine the cell cycle distribution of each group (HCT16/LOVO): G1 phase control group $41.54 \pm 0.50\%/38.31 \pm 1.33\%$, high-dose raspberry group $57.78 \pm 2.04\%/59.22 \pm 1.28\%$, medium-dose raspberry group $50.00 \pm 0.59\%/52.00 \pm 2.00\%$, low-dose raspberry group $49.17 \pm 1.34\%/43.62 \pm 1.63\%$, and raspberry ketone group $65.15 \pm 0.91\%/61.77 \pm 1.11\%$; S-phase control group $51.07 \pm 2.12\%/50.70 \pm 1.99\%$, high-dose raspberry $39.20 \pm 1.58\%/36.32 \pm 1.35\%$, medium-dose raspberry group $46.43 \pm 2.27\%/43.00 \pm 3.00\%$, low-dose raspberry group $48.66 \pm 2.83\%/53.85 \pm 2.81\%$, and raspberry ketone group $23.56 \pm 0.79\%/30.64 \pm 1.28\%$; and G2 control group $7.38 \pm 1.70\%/11.09 \pm 1.58\%$, high-dose raspberry group $3.02 \pm 0.54\%/4.46 \pm 1.12\%$, medium-dose raspberry group $2.00 \pm 1.73\%/5.00 \pm 3.00\%$, low-dose raspberry group $2.17 \pm 1.73\%/2.53 \pm 1.58\%$, and raspberry ketone group $11.30 \pm 1.70\%/7.64 \pm 0.81\%$. The G1 phase cells in the experimental group and the positive control group increased, and the S phase cells decreased, suggesting that there is a G1-S phase cell cycle arrest; the degree of blockade is positively correlated with the concentration of the experimental group, and the concentration of the low concentration experimental group is close to or equivalent to the control group (Figures 2(b) and 2(c)).

The expressions of wnt1, wnt3a, wnt5a, β-catenin, CyClinD1, and CDK4 protein were downregulated in Western blot, and GSK3β and APC protein expressions were upregulated, resulting that WNT channel as a whole is in a state of reduced function after treatment. Red raspberry phytochemical extracts can cause the downregulation of the WNT pathway key protein β-catenin in the colorectal cancer cell and the decrease in the distribution of nuclei in its function region. The downregulation of the function of the WNT pathway and cell cycle-related proteins CyClinD1 and CDK4 downstream of the WNT pathway leads to cell cycle arrest (Figure 2(d)).



(a)

Figure 2: Continued.



(b)

Figure 2: Continued.

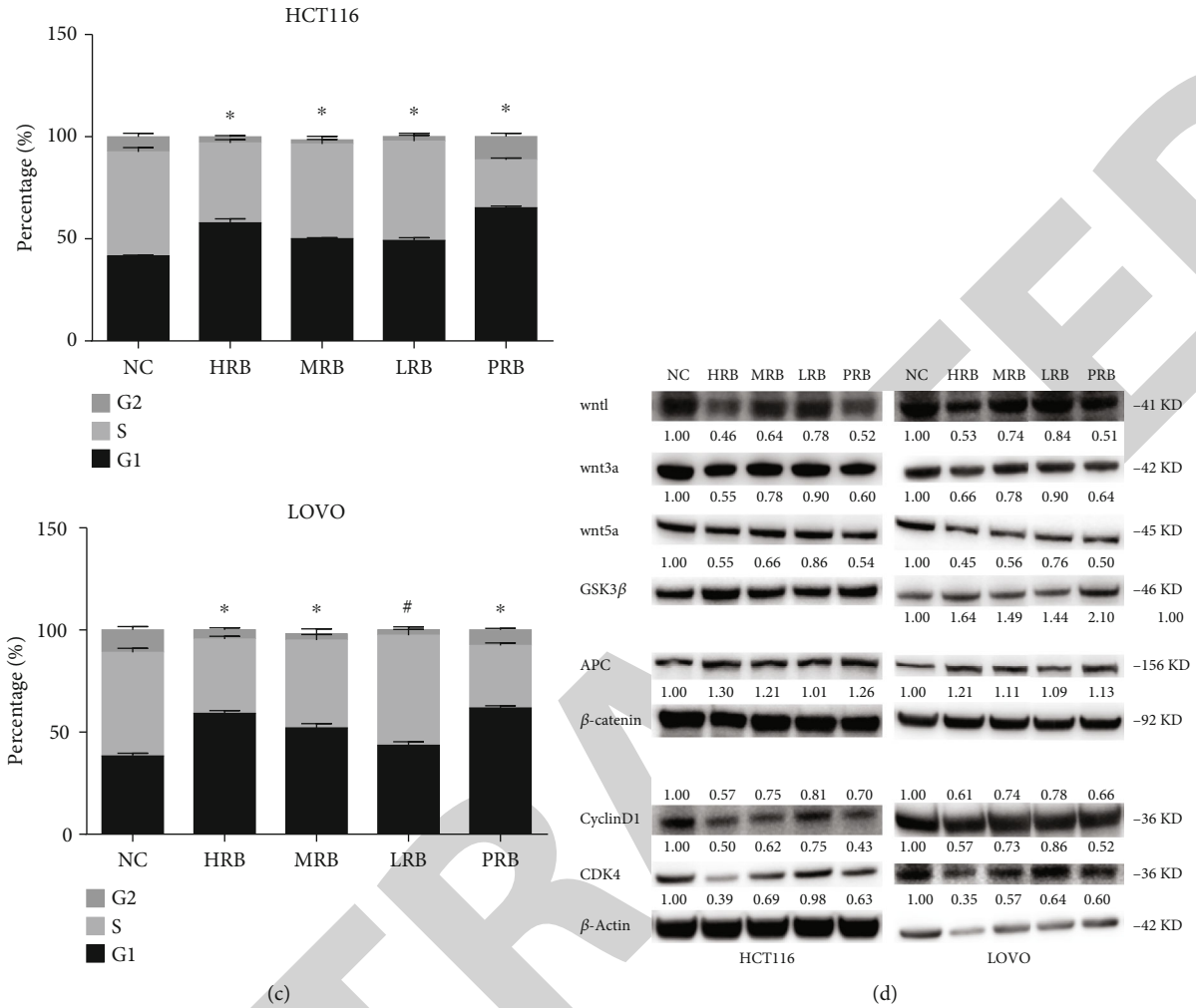


Figure 2: The effect of the cell cycle of HCT116 and LOVO in colorectal cancer and β-catenin raised by the chemical extract of red raspberry on expression and localization in cells and Wnt pathway. (a) Effects on expression and localization of β-catenin in cells with different concentrations of red raspberry phytochemical extracts and raspberry ketone. (b, c) Effects of different concentrations of red raspberry phytochemical extracts and raspberry ketone on cell cycle. (d) Effects of different concentrations of red raspberry phytochemical extracts and raspberry ketone on the expression of Wnt pathway proteins. $^{*}p < 0.05$; $\#p > 0.05$.

3.3. The Rescue of Inhibitory Effect of Natural Hydroxyphenyl Butanone by Blocking the WNT Pathway Key Protein GSK3β. The siRNA knockdown could reduce the GSK-3β RNA expression to 30% (HCT116 31.92 ± 7.92%, LOVO 28.52 ± 7.42%) of the control group in the colorectal cancer HCT116 and LOVO cell lines while the protein level expression is also significant downregulated (Figures 3(a) and 3(b)). After the expression of GSK-3β was downregulated by siRNA transfection, colorectal cancer HCT116 and LOVO cells were treated with red raspberry plant chemical extracts and raspberry ketones, WNT pathway protein constitutive proteins β-catenin, CyClinD1, and CDK4 on Western blot. The change from the control group is reduced (Figure 3(c)). The siRNA transfection blocking GSK-3β could rescue the effect of red raspberry plant chemical extracts on proteins related to cell cycle regulation downstream of GSK-3β in the WNT pathway. After the expression of GSK-3β was downregulated by siRNA transfection, the experimental concentrations of red raspberry plant chemical extract and raspberry ketone positive control were added to the medium, and β-catenin protein was compared with the control group by qualitative analysis of immunofluorescence. The difference in expression level could be rescued, and the difference in expression distribution in the nucleus and cytoplasm is also reduced (Figure 3(d)). Our result proved the chemical extract of red raspberry plants could affect the colorectal cancer cell cycle through the WNT pathway and exerts an inhibitory effect on the colorectal cancer cells.

As an important signaling mechanism for cell growth and development, the Wnt pathway is closely related to tumors, especially colorectal cancer. As an important transcriptional regulator in the Wnt pathway, β-catenin plays a significant role in the inhibition of malignant expression of tumor cells through LEF/TCF and other pathways with its increase in expression and import to nucleus. This study showed that the phytochemical extracts of raspberry and raspberry ketone could downregulate the expression of Wnt pathway protein and reduce the nuclear import of β-catenin.

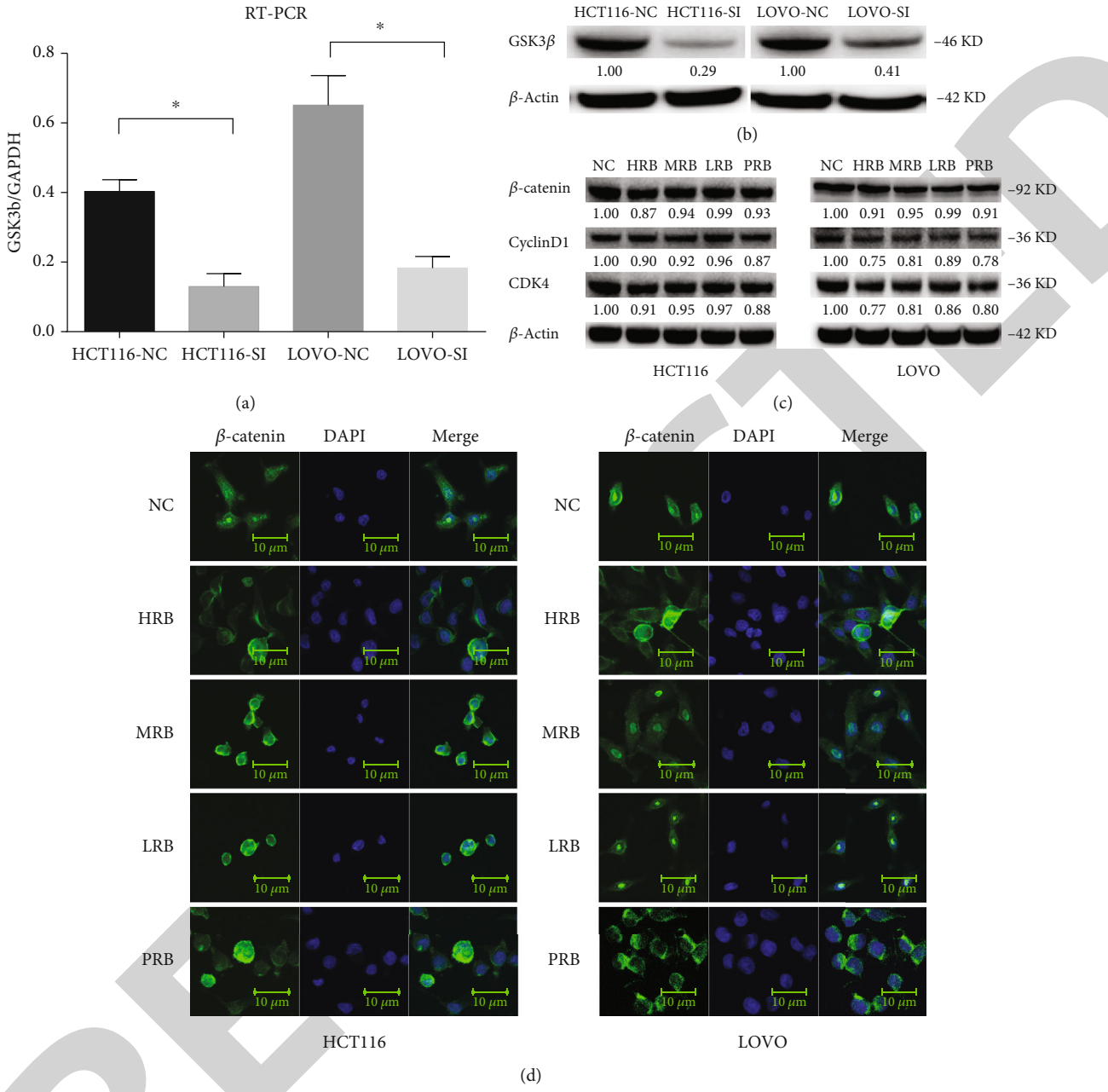


Figure 3: The effect of HCT116 and LOVO cells in colorectal cancer downregulated by GSK3β expression and β-catenin raised by the chemical extract of red raspberry on expression and localization in cells and Wnt pathway. (a) GSK3β expression downregulated RT-PCR detection. The expression level of GSK3β mRNA in the HCT116 transfected group was $31.92 \pm 7.92\%$ of that in the control group and that in the LOVO transfected group was $28.52 \pm 7.42\%$ of that in the control group by RT-PCR detection. (b) Western blot detection of downregulated levels of GSK3β expression. GSK3β mRNA expression level in the HCT116 transfected group was 29% of that in the control group by Western blot analysis and was 41% in the LOVO transfected group of that in the control group. (c) Under the same experimental conditions, effects of chemical extracts of raspberry with different concentrations and raspberry ketone on the Wnt pathway of colorectal cancer HCT116 and LOVO cells with Si-RNA transfection downregulated GSK3β expression. (d) Under the same experimental conditions, effects of different concentrations of red raspberry phytochemical extracts and raspberry ketone on the expression and localization in cells of β-catenin in Si-RNA transfection colorectal cancer HCT116 and LOVO cells downregulated by GSK3β expression. $^{*}p < 0.05$.

The use of siRNA to blocking the Wnt pathway was significantly weakened, indicating that the phytochemical extracts of red raspberry plants and raspberry ketone exert an antitumor effect through the Wnt pathway in colorectal cancer cell lines (Figure 4).

4. Discussion

Our group determined the mature method of phytochemical extracts of red raspberry plants through preliminary research and found that the phytochemical extracts of red raspberry

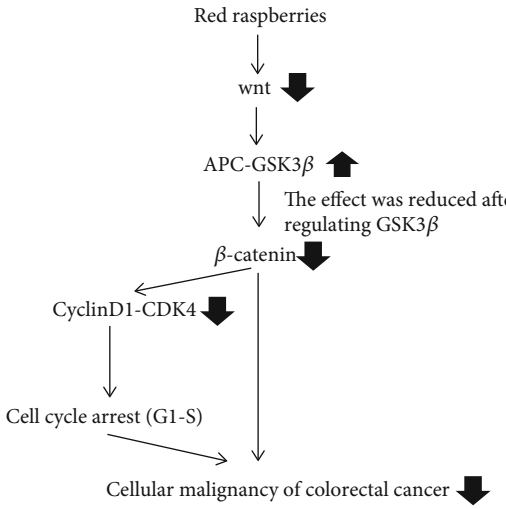


Figure 4: Effect of chemical extracts of raspberry plants on the Wnt pathway.

plants can inhibit the malignant manifestations of human hepatocellular carcinoma through PTEN/AKT pathway [19]. Studies have shown that raspberry ketone can inhibit the proliferation of a variety of tumor cells including colorectal cancer and lead to cell cycle G1-S phase arresting [36]; there are reports that raspberry powder can reduce the number of large intestinal abnormal crypt foci (ACF) induced by oxidation Azomethane (AOM) [37] and reduce the intestinal tumor number of APCmin/+ mice and inhibit cell proliferation activity of it [38], but the mechanisms of these effects produced by raspberry powder or raspberry ketone are not fully understood. The effects and mechanisms of phytochemical extracts of red raspberry on colorectal cancer cells are still lacking in related research, and there are no related reports whether the antitumor mechanism of raspberry phytochemical extracts and pure raspberry ketones in the above studies is exactly the same.

This study showed that the phytochemical extracts of raspberry plants and raspberry ketone can cause cell cycle G1-S arresting of colorectal cancer HCT116 and LOVO cells through Wnt pathway, inhibiting the proliferation, colony formation, and invasion. The progression of the cell cycle is tightly regulated by a complex network of cyclins, CDKs, and transcription factors at several irreversible regulatory points. The cell cycle is regulated by the CyclinD1/CDK4 complex in the early G1 [39]. CyclinD1/CDK4 is one of the important downstream regulatory pathways of the Wnt pathway, and its downregulation of expression can lead to cell cycle G1-S phase arresting. This study demonstrates that raspberry phytochemical extracts and raspberry ketones can cause cell cycle G1-S arresting by downregulating the expression of CyclinD1/CDK4 protein in colorectal cancer cell lines.

As a promoter of the Wnt signaling pathway, the Wnt family contains 14 proteins. This study found that red raspberry plant chemical extracts and raspberry ketones can cause downregulation of Wnt1, Wnt3a, and Wnt5a proteins, suggesting that it may have an inhibitory effect on multiple proteins of the Wnt family. Because the raspberry phytochemical extract and raspberry ketone may cause the downregulation of multiple Wnt family proteins, it is not suitable to verify the effect on the Wnt pathway by blocking a single Wnt family protein alone. In this study, the next level of the signal protein GSK-3β of the Wnt protein in the Wnt pathway was blocked for study.

Because we did not find an antitumor drug that is widely recognized and has the same antitumor mechanism as the chemical extract of the red raspberry plant, as a positive control, we chose a low concentration classic antitumor drug fluorouracil of colorectal cancer as a positive control for the cell phenotype to express the inhibition of chemical extracts of red raspberry plants on colorectal cancer cells. The mechanism of action of fluorouracil and red raspberry phytochemical extracts on colorectal cancer tumor cells was different in the cell cycle and Wnt pathway. It should not be used as a positive control. We only used pure raspberry ketone as a positive control.

Currently, clinically widely used colorectal cancer antitumor drugs such as platinum, paclitaxel, and fluorouracil are toxic to normal cells while exerting a killing effect on tumor cells. In this experiment, when the half-inhibitory concentration of tumor cells was reached, the inhibitory effect of chemical extracts of red raspberry on normal colon cells was weaker than that of fluorouracil. Under the same experimental concentration, proliferation inhibition of colorectal cancer HCT116 and LOVO cell lines exerted by chemical extracts of raspberry plants and raspberry ketone was significantly stronger than that of normal colonic epithelial cells. The chemical extracts of red raspberry in the low-intensity experimental group promoted the proliferation of normal colonic epithelial cells, suggesting that the inhibitory effect of chemical extracts of red raspberry on colorectal cancer cells may be a selective antitumor effect, which may be less toxic or even have a certain cytoprotective effect on normal colon tissue.

In the macrolevel, as a kind of widely used of raspberry fruit and functional food, its function mainly displays in the traditional medicine in the regulation of the body function, rather than the direct toxic effect of tumor; red raspberry has certain protective effect on normal colon tissue, which may be the macroeffect of microreaction in cell experiments. In terms of the microscopic molecular mechanism, this study found that red raspberry phytochemical extracts had inhibitory effects on Wnt pathway activity of colorectal cancer cells, and abnormal Wnt pathway as an important characteristic of colorectal cancer cells has been recognized as a result. There was no obvious abnormal activation of Wnt pathway in normal colon cells, which may be part of the reason why the chemical extract of red raspberry had weak toxicity and even certain protective effect on normal colon tissues. Of course, more detailed molecular mechanisms require a lot of further research is also one of the directions of further research.

Extraction pure raspberry ketone is difficult and expensive, and there is no research on the safety of drugs in humans. The chemical extract of red raspberry plant can cause the inhibition of colorectal tumor cells similar to pure raspberry ketone, and as a widely consumed food extract, its safety in human application is fully guaranteed. This study shows that the chemical extract of red raspberry has broad

application prospects as a functional food against colorectal cancer. Of course, the conversion to clinical applications requires a large number of animal and human trials.

Data Availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors' Contributions

ML conceived the idea for publication. SYZ and ML had intellectual input into the study design. SYZ, YFW, and HPZ performed the study. SYZ, CMS, and SWD provided intellectual input into preparation of the manuscript. All authors read and approved the final manuscript.

References

[1] M. Rani, B. P. Yang, and R. Nesbit, "Hepatitis B control by 2012 in the WHO Western Pacific Region: rationale and implications," *Bulletin of the World Health Organization*, vol. 87, no. 9, pp. 707–713, 2009.

[2] J. Ferlay, D. M. Parkin, and E. Steliarova-Foucher, "Estimates of cancer incidence and mortality in Europe in 2008," *European Journal of Cancer*, vol. 46, no. 4, pp. 765–781, 2010.

[3] W. Q. Chen, R. S. Zheng, P. D. Baade et al., "Cancer statistics in China, 2015," *CA: a cancer journal for clinicians*, vol. 66, no. 2, pp. 115–132, 2016.

[4] G. Perkins and P. Laurent-Puig, "Colorectal cancer biology," *La Revue du praticien*, vol. 65, no. 6, pp. 802–806, 2015.

[5] B. L. Tan, N. M. Esa, H. S. Rahman, H. Hamzah, and R. Karim, "Brewers' rice induces apoptosis in azoxymethane-induced colon carcinogenesis in rats via suppression of cell proliferation and the Wnt signaling pathway," *BMC Complementary and Alternative Medicine*, vol. 14, no. 1, p. 304, 2014.

[6] H. Aberle, A. Bauer, J. Stappert, A. Kispert, and R. Kemler, "Beta-catenin is a target for the ubiquitin-proteasome pathway," *The EMBO journal*, vol. 16, no. 13, pp. 3797–3804, 1997.

[7] M. J. Hart, R. de los Santos, I. N. Albert, B. Rubinfeld, and P. Polakis, "Downregulation of β-catenin by human Axin and its association with the APC tumor suppressor, β-catenin and GSK3β," *Current biology : CB*, vol. 8, no. 10, pp. 573–581, 1998.

[8] B. Rubinfeld, I. Albert, E. Porfiri, C. Fiol, S. Munemitsu, and P. Polakis, "Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly," *Science (New York, N.Y.)*, vol. 272, no. 5264, pp. 1023–1026, 1996.

[9] A. B. Sparks, P. J. Morin, B. Vogelstein, and K. W. Kinzler, "Mutational analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer," *Cancer Research*, vol. 58, no. 6, pp. 1130–1134, 1998.

[10] M. Bienz and H. Clevers, "Linking colorectal cancer to Wnt signaling," *Cell*, vol. 103, no. 2, pp. 311–320, 2000.

[11] Z. Li, W. Zhou, Y. Zhang et al., "ERK regulates HIF1α-mediated platinum resistance by directly targeting PHD2 in ovarian cancer," *Clinical Cancer Research*, vol. 25, no. 19, pp. 5947–5960, 2019.

[12] X. Q. Wang, J. D. Wang, H. H. Ma, J. Zhang, and X. J. Zhou, "Downregulation of miR-195 correlates with lymph node metastasis and poor prognosis in colorectal cancer," *Medical Oncology*, vol. 29, no. 2, article 9880, pp. 919–927, 2012.

[13] S. Tian, I. Simon, V. Moreno et al., "A combined oncogenic pathway signature of BRAF, KRAS and PI3KCA mutation improves colorectal cancer classification and cetuximab treatment prediction," *Gut*, vol. 62, no. 4, pp. 540–549, 2013.

[14] D. J. Newman and G. M. Cragg, "Natural products as sources of new drugs from 1981 to 2014," *Journal of Natural Products*, vol. 79, no. 3, pp. 629–661, 2016.

[15] S. Bera, B. Das, A. De et al., "Metabolite profiling and *in-vitro* colon cancer protective activity of *Cycas revolutacone* extract," *Natural Product Research*, vol. 34, no. 4, pp. 599–603, 2020.

[16] Y. Z. Hashim, J. Worthington, P. Allsopp et al., "Virgin olive oil phenolics extract inhibit invasion of HT115 human colon cancer cells in vitro and in vivo," *Food & Function*, vol. 5, no. 7, pp. 1513–1519, 2014.

[17] J. Sun, X. Fu, Y. Wang et al., "Erianin inhibits the proliferation of T47D cells by inhibiting cell cycles, inducing apoptosis and suppressing migration," *American Journal of Translational Research*, vol. 8, no. 7, pp. 3077–3086, 2016.

[18] Y. Zhang, Q. Li, D. Zhou, and H. Chen, "Genistein, a soya isoflavone, prevents azoxymethane-induced up-regulation of WNT/β-catenin signalling and reduces colon pre-neoplasia in rats," *The British Journal of Nutrition*, vol. 109, no. 1, pp. 33–42, 2013.

[19] S. Reabroi, R. Saeeng, N. Boonmuen et al., "The anti-cancer activity of an andrographolide analogue functions through a GSK-3β-independent Wnt/β-catenin signaling pathway in colorectal cancer cells," *Scientific Reports*, vol. 8, no. 1, pp. 7924–7924, 2018.

[20] J. Sun, X. Cai, M. M. Yung et al., "miR-137 mediates the functional link between c-Myc and EZH2 that regulates cisplatin resistance in ovarian cancer," *Oncogene*, vol. 38, no. 4, pp. 564–580, 2019.

[21] A. Pabst, D. Barron, E. Sémont, and P. Schreier, "4-Oxo-β-ionol and linalool glycosides from raspberry fruits," *Phytochemistry*, vol. 31, no. 12, pp. 4187–4190, 1992.

[22] Y. J. Surh, "Cancer chemoprevention with dietary phytochemicals," *Nature Reviews. Cancer*, vol. 3, no. 10, pp. 768–780, 2003.

[23] C. Morimoto, Y. Satoh, M. Hara, S. Inoue, T. Tsujita, and H. Okuda, "Anti-obese action of raspberry ketone," *Life Sciences*, vol. 77, no. 2, pp. 194–204, 2005.

[24] M. E. Figueira, M. B. Camara, R. Direito et al., "Chemical characterization of a red raspberry fruit extract and evaluation of its pharmacological effects in experimental models of acute inflammation and collagen-induced arthritis," *Food & Function*, vol. 5, no. 12, pp. 3241–3251, 2014.

[25] L. L. Wang, X. J. Meng, and F. Q. Zhang, "Raspberry ketone protects rats fed high-fat diets against nonalcoholic steatohepatitis," *Journal of medicinal food*, vol. 15, no. 5, pp. 495–503, 2012.

[26] K. Durgo, A. Belscak-Cvitanovic, A. Stancic, J. Franekic, and D. Komes, "The bioactive potential of red raspberry (Rubus

Hindawi
BioMed Research International
Volume 2023, Article ID 9759620, 1 page
https://doi.org/10.1155/2023/9759620

Retraction

Retracted: Microbiome-Metabolomics Reveals Endogenous Alterations of Energy Metabolism by the Dushen Tang to Attenuate D-Galactose-Induced Memory Impairment in Rats

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Wang, M. He, W. Guo et al., "Microbiome-Metabolomics Reveals Endogenous Alterations of Energy Metabolism by the Dushen Tang to Attenuate D-Galactose-Induced Memory Impairment in Rats," *BioMed Research International*, vol. 2021, Article ID 6649085, 17 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6649085, 17 pages
https://doi.org/10.1155/2021/6649085

Research Article

Microbiome-Metabolomics Reveals Endogenous Alterations of Energy Metabolism by the Dushen Tang to Attenuate D-Galactose-Induced Memory Impairment in Rats

Jifeng Wang,[1] **Min He**,[1,2] **Wenjun Guo**,[1] **Yanhong Zhang**,[1] **Xin Sui**,[1] **Jianan Lin**,[1] **Xiaoran Liu**,[1] **Hui Li**,[3] **Jing Li**,[1] **Qing Yang**,[1] **Mo Kan**,[1] **Zhuang Zhang**,[1] **Sitong Ming**,[1] **Xiaobo Qu**,[1] **and Na Li**[1]

[1]*Jilin Provincial Key Laboratory of Biomacromolecules of Chinese Medicine, Jilin Ginseng Academy, Changchun University of Chinese Medicine, Changchun, Jilin 130021, China*
[2]*Leiden University—European Center for Chinese Medicine, Leiden University, 2333CC Leiden, Netherlands*
[3]*Qian Wei Hospital of Jilin Province, Changchun, Jilin 130117, China*

Correspondence should be addressed to Na Li; lhaln@hotmail.com

Received 1 November 2020; Revised 23 April 2021; Accepted 12 May 2021; Published 27 May 2021

Academic Editor: Min Tang

Aging affects the brain function in elderly individuals, and Dushen Tang (DST) is widely used for the treatment of senile diseases. In this study, the protective effect of DST against memory impairment was evaluated through the Morris water maze (MWM) test and transmission electron microscopy (TEM). A joint analysis was also performed using LC-MS metabolomics and the microbiome. The MWM test showed that DST could significantly improve the spatial memory and learning abilities of rats with memory impairment, and the TEM analysis showed that DST could reduce neuronal damage in the hippocampus of rats with memory impairment. Ten potential biomarkers involving pyruvate metabolism, the synthesis and degradation of ketone bodies, and other metabolic pathways were identified by the metabolomic analysis, and it was found that 3-hydroxybutyric acid and lactic acid were involved in the activation of cAMP signaling pathways. The 16S rDNA sequencing results showed that DST could regulate the structure of the gut microbiota in rats with memory impairment, and these effects were manifested as changes in energy metabolism. These findings suggest that DST exerts a good therapeutic effect on rats with memory impairment and that this effect might be mainly achieved by improving energy metabolism. These findings might lead to the potential development of DST as a drug for the treatment of rats with memory impairment.

1. Introduction

Aging is one of the most important factors involved in the gradual functional decline and structural degeneration of the brain and has been implicated in memory impairment, which is a major symptom of Alzheimer's disease (AD) [1–3]. Due to the rapid aging of the population, memory impairment seriously affects the quality of life of a large number of elderly individuals [4, 5]. Recent studies of cerebral energy metabolism have shown that mitochondrial dysfunction, oxidative stress, excessive levels of advanced glycation end products, and neurotoxicity are involved in the pathogenesis of memory impairment [6, 7]. Studies of memory impairment and relevant pathogenesis mechanisms could provide novel perspectives and targets for exploring new drug candidates and new therapy strategies and for evaluating the efficacy of drugs against aging and AD.

Traditional Chinese herbal medicines have shown unique therapeutic effects on various diseases. In China, Dushen Tang (DST, also called "Ginseng decoction"), which is composed of a single herbal material (Panax ginseng C.A. Meyer, regarded as "the king of herbs" in China), was first published in "Shi Yao Quan Shu," which describes that its extraction method is decoction [8, 9]. DST has a long history of application and has been widely used, particularly for memory enhancement [10]. Modern pharmacological studies have

confirmed the multiple pharmacological effects of ginseng, and these effects include memory enhancement, immune modulation, antioxidant stress, and antiaging effects [11]. Specifically, ginseng restores D-gal-induced memory deficits by enhancing hippocampal LPA1 receptor expression, LTP, and neurogenesis [12]. Ginseng polysaccharides and ginsenosides, which are the main active ingredients of the DST prescription, exert favorable curative effects on AD and other aging-related diseases [13–15]. However, the human body is a complicated system that exhibits dynamic metabolic changes and is affected by interactions between the intestinal flora and the host. Although the therapeutic effect of ginseng against memory impairments has been confirmed in clinical settings in China, the potential mechanism through which DST affects endogenous metabolism and the microbiome remains unclear.

Long-term injections of D-galactose (D-gal) have been widely used to establish AD and aging models because these cause many aging- or AD-like pathological symptoms, such as memory impairments, reduced neurogenesis, increased levels of amyloid precursor protein, and oxidative damage [1, 16, 17]. Using such an animal model might aid the assessment of therapeutic effects and could provide more scientific evidence and biological interpretations of the effects of DST against memory impairment.

Metabolomics analyzes the small molecules in a biological sample, such as serum and urine, to elucidate associations between metabolites and research subjects and elucidate the aberrant metabolic pathways involved in complex diseases, and as a result, this approach can efficiently screen biomarkers and allow an in-depth analysis of the molecular mechanisms regulating health or disease [18]. However, host metabolism is regulated not only by its own genome but also by its symbiotic bacteria, and among these bacteria, the gut microbiota both participate in and regulate metabolic processes [19]. The combination of metabolomics and microbiological techniques could provide a powerful tool to further study the relationship between host metabolism and the gut microbiota and to unravel the metabolic phenotype variations associated with the gut microbiota perturbations in disease development [20].

A recent study illustrated the essential roles of the gut microbiota in central nervous system diseases, including Parkinson's disease, autism spectrum disorders, and AD [21]. A novel strategy of integrating 16S rRNA gene sequencing with UPLC-MS-based metabolomics has been applied to analyze the relationship among the gut microbiota, metabolites, and memory impairment [22–24], but this strategy has not been used to study the protective effects of DST against memory impairment induced by D-galactose in rats. Therefore, in this study, we used a rat model with memory impairment induced by D-galactose to study the potential protective effects of DST. These effects were assessed by the MWM test and transmission electron microscopy (TEM). Thereafter, the serum metabolites and the fecal microbial community were analyzed by UPLC-Q-TOF/MS and 16S rDNA, respectively. Multivariate statistical analyses, such as PLS-DA and principal coordinates analysis (PCoA), were used to identify specific biomarkers and to explore the changes in the gut microbiota, and the metabolic pathways related to memory impairment were explored using the KEGG database. An analysis combining the common changes in the flora and metabolites showed that DST can influence memory impairments by regulating several important intestinal bacteria and endogenous energy-related metabolites.

2. Materials and Methods

2.1. Materials and Preparation of Water Extracts of DST. Free-growing ginseng (natural dry rhizome of Panax ginseng C.A. Mey grown in a forest for 15 years) was purchased from Jilin zhi 'en tang ginseng special products Co., Ltd. China, and was validated by Dacheng Jiang (affiliated with the Changchun University of Chinese Medicine). The purchased ginseng (50 g) was immersed in 500 mL of deionized water for 2 h and then decocted three times (100°C, each extraction was conducted for 1 h). The decoction was filtered, combined together, and then condensed to 200 mL to obtain the final water extract.

2.2. Identification of the Main Compounds of DST. DST was transferred into a centrifuge tube and lyophilized with a freeze dryer. One hundred milligrams of the retained pellet was dissolved in 10 mL of 70% methanol solution and centrifuged at 14000 × g and 4°C for 10 min. A 4 μL of aliquot of the supernatant was used for LC-MS analysis. Five milligrams of ginsenosides R1, Rg1, Re, Rf, Rh1, Rb1, Rc, Ro, Rb2, and Rb3 was weighed and placed in a 100 mL of volumetric flask. Methanol was added, and the volume was increased to the mark on the tube to obtain the reference solution. A 4 μL of aliquot of the reference solution was used for LC-MS analysis.

Chromatographic separation was performed using a Vanquish Duo UHPLC system (Thermo Fisher Scientific, San Jose, CA, USA) with a Hypersil GOLDTM column (2.1 mm × 50 mm, 1.9 μm) at 30°C and an injection volume of 4 μL. The mobile phase consisted of 0.1% aqueous formic acid (phase A) and acetonitrile (phase B) and was administered using the following gradient program with a flow rate of 0.4 mL/min: 10% B (0-4 min), 10-25% B (4-8 min), 25-29% B (8-16 min), 29% B (16-25 min), 29-95% (B 25-27 min), and 95% B (27-30 min).

Mass acquisition was conducted with a Q-Orbitrap mass spectrometer equipped with an electrospray ionization (ESI) source (Thermo Fisher Scientific, San Jose, CA, USA) operating separately in the negative ion mode, and the profile data were recorded in the range of m/z 200-1500. Full-scan acquisition was performed at a resolution of 7000, whereas tandem MS information was acquired under the ddMS2 mode with a resolution of 35000. The key parameters of the ionization source were set as follows: capillary voltage, -3.5 kV in the negative ion mode; sheath gas flow, 60 arb; auxiliary gas flow, 15 arb; sweep gas flow, 2 arb; and capillary temperature, 320°C. Prior to sample analysis, the mass spectrometer was calibrated using the vendor-provided Pierce™ calibration solution (Thermo Fisher Scientific, San Jose, CA, USA). All samples were maintained at 4°C through the analysis.

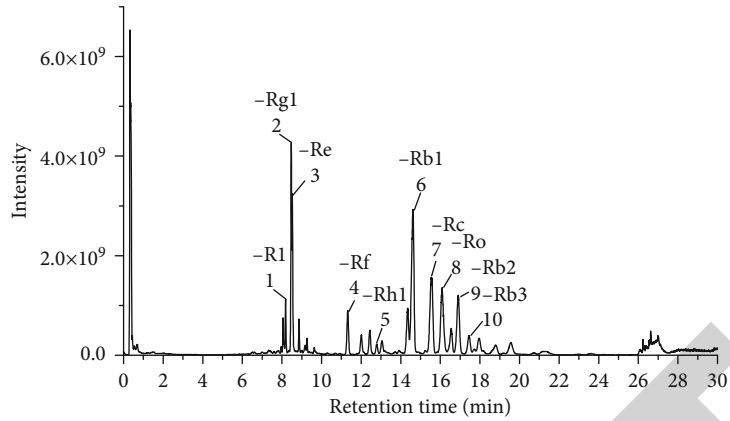


Figure 1: TIC chromatograms of Dushen Tang (DST).

The LC-MS raw chromatographic data were processed using the TraceFinderTM (TF) software from Thermo Fisher Scientific. The main compounds of DST were identified by matching the obtained data with accurate mass and MS/MS information of natural products acquired from Orbitrap Traditional Chinese Medicine Library (OTCML) databases, and some of these compounds were identified using reference substances.

2.3. Animal Modeling and Treatments. A total of 30 healthy adult male Sprague-Dawley rats (six weeks old, weighing 200 ± 10 g) were purchased from Changchun Yisi Experimental Animal Technology Co., Ltd. (China, License No.: SCXK(Ji)-2018-0007). After one week of adaptation feeding, the rats were randomly divided into the following three groups (10 rats in each group): healthy control group, model group, and DST group. All the rats had free access to standard rodent chow and water. The model and DST-treated rats were intraperitoneally injected with D-gal (daily dose of 500 mg/kg) for 7 weeks. Additionally, the rats in the DST group were orally administered DST (0.3 g/kg/d) during the experimental period, and the rats in the control and model groups were administered the same volume of saline.

2.4. Morris Water Maze Test. The MWM test was utilized to evaluate the spatial learning and memory capacities of the rats. This test forces rats to swim in water until they find a platform hidden underwater. Before the formal experiment, a training trial was conducted once per day for one week. On the test day, the rats were maintained on the platform for 10 s and then released in any of the four quadrants (once in each quadrant). The timer was set to 120 s, and the escape latency was time when the rat found the platform (up to 120 s). All behavioral data were obtained and examined using the SMART V3.0 video tracking system (Panlab Harvard Apparatus, Hugstetten, Germany).

2.5. Detection of Ultrastructural Changes in the Rat Hippocampus by Transmission Electron Microscopy. After the MWM test, the rats were perfused successively with 0.9% saline, 4% paraformaldehyde, and 1% glutaraldehyde via the heart. The skull was exposed, and the hippocampus was stripped, fixed in 2% paraformaldehyde-2.5% glutaraldehyde solution for 3 h at 4°C, washed thoroughly with 0.1 mol/L PBS, and fixed with 1% osmium tetroxide at 4°C for 1.5 h. After dehydration and embedding, semiultrathin sections with a thickness of 1-2 μm were prepared. The hippocampus was stained with methylene blue and positioned under an optical microscope, and ultrathin sections of 50–70 nm were prepared. After staining with uranyl acetate and lead citrate, the ultrastructure of hippocampal neurons was observed under a transmission electron microscope.

2.6. Metabolomic Analysis. Serum samples (100 μL) were extracted using methanol/acetonitrile/water (2 : 2 : 1, v/v), vortexed and shaken for 5 min, and centrifuged at 13,000 × g and 4°C for 15 min, and 100 μL of aliquots of the supernatant was obtained, freeze-dried, mixed well with 100 μL of water, and used for LC-MS analysis. Quality control (QC) samples were prepared by pooling equal volumes (10 μL) of serum from each sample for optimization of the chromatographic and MS conditions. Subsequently, a 10 μL of serum sample was analyzed using an Agilent 1290 Ultra Performance LC system equipped with a HILIC C18 column (Waters, ACQUITY UPLC BEH 100 mm × 2.1 mm, 1.7 μm) at 25°C. Phase A was aqueous and consisted of 25 mM of ammonium acetate, and phase B was acetonitrile. The mobile phase was administered using the following gradient program at a flow rate of 0.3 mL/min: 95% B (0–0.5 min), 95%–65% B (0.5–7 min), 65%–40% B (7–9 min), 40% B (9–10 min), 40%–95% B (10–11.1 min), and 95% B (11–16 min). All the samples were maintained at 4°C until further analysis.

Under the positive and negative ion modes, data were acquired over a full scan (range: 60–1200 Da) and MS2 scan (range: 25–1200 Da) using a Triple-TOF 5600 (AB SCIEX, USA) mass spectrometer. The mass spectrometry conditions consisted of the following: electrospray ionization (ESI); capillary voltage, 5.5 kV (ESI±); ion source gas 1, 60; ion source gas 2, 60; curtain gas, 30; and source temperature, 600°C.

The processed dataset was then fed to the SIMCA-P software package (v13.0, Umetrics, Umea, Sweden), and the normalized data were then used for principal component analysis (PCA) and partial least squares-discriminate analysis (PLS-DA). A PCA of the QC data was performed to

Table 1: Details of the compounds identified from DST.

No.	Target name	RT (min)	Formula	Expected (m/z)	Measured (m/z)	Peak number
1	Ginsenoside R1	8.18	C47H80O18	977.5327	977.531	1
2	Ginsenoside Rg1	8.47	C42H72O14	845.4904	845.4893	2
3	Ginsenoside Re	8.53	C48H82O18	991.5483	991.5465	3
4	Ginsenoside Rf	11.28	C42H72O14	845.4904	845.4893	4
5	Ginsenoside Rh1	12.78	C36H62O9	683.4376	683.4371	5
6	Ginsenoside Rb1	14.42	C54H92O23	1153.6011	1153.5995	6
7	Ginsenoside Rc	15.72	C53H90O22	1123.5906	1123.5895	7
8	Ginsenoside Ro	16.08	C48H76O19	955.4908	955.4901	8
9	Ginsenoside Rb2	16.94	C53H90O22	1123.5906	1123.5895	9
10	Ginsenoside Rb3	17.43	C53H90O22	1123.5906	1123.5895	10
11	Ginsenoside Rg2	26.62	C42H72O13	829.4955	829.4958	-
12	Mogroside II A2	11.33	C42H72O14	835.4616	835.4617	-
13	Notoginsenoside Fe	26.24	C47H80O17	961.5378	961.5386	-
14	Tenacissoside H	26.57	C42H66O14	793.438	793.4388	-
15	Cornuside	11.41	C24H30O14	577.133	577.135	-
16	Procyanidin B1	11.41	C30H26O12	577.1351	577.135	-
17	Tomatine	9.25	C50H83NO21	1032.5385	1032.5455	-
18	Picroside II	10.57	C23H28O13	547.1224	547.1248	-
19	Atractyloside A	8.62	C21H36O10	493.229	493.229	-
20	Dehydroandrographolide	26.86	C20H28O4	377.197	377.2005	-
21	Periplocin	12.14	C36H56O13	731.3415	731.3425	-
22	Melittoside	1.01	C21H32O15	523.1668	523.1701	-
23	Araloside A	17.65	C47H74O18	925.4802	925.481	-
24	Tubeimoside I	12.81	C63H98O29	1377.6332	1377.6445	-
25	Geniposidic acid	0.67	C16H22O10	373.114	373.1176	-
26	20(S)-Ginsenoside Rh2	28.76	C36H62O8	621.4372	621.4373	-
27	Orcinol gentiobioside	1.71	C19H28O12	447.1508	447.1512	-
28	Artemisinic acid	27.00	C15H22O2	233.1547	233.1543	-
29	Anemoside B4	14.66	C59H96O26	1219.6117	1219.6201	-
30	Demethylwedelolactone	26.77	C15H8O7	334.9964	334.9937	-
31	Dioscin	8.47	C45H72O16	913.4802	913.4755	-
32	Morroniside	1.44	C17H26O11	451.1457	451.1464	-
33	Licochalcone A	0.67	C21H22O4	373.1212	373.1176	-
34	Gracillin	12.02	C45H72O17	883.4697	883.4666	-
35	Tacrolimus	28.76	C44H69NO12	848.4802	848.4834	-
36	Eupalinolide A	1.01	C24H30O9	461.1817	461.1802	-
37	Jujuboside A	16.94	C58H94O26	1205.5961	1205.5946	-
38	Hosenkoside K	9.16	C54H92O25	1139.5855	1139.585	-
39	Methyl hexadecanoate	26.89	C17H34O2	315.2541	315.2543	-
40	Polyphyllin E	8.53	C51H82O20	1059.5382	1059.5337	-
41	Dihydrocucurbitacin F	27.03	C30H48O7	565.3382	565.3344	-
42	Asiaticoside	16.23	C48H78O19	957.5064	957.4971	-
43	Periplocymarin	18.35	C30H46O8	569.2887	569.2898	-
44	Macranthoside B	26.33	C53H86O22	1073.5538	1073.5568	-
45	Momordin Ic	26.74	C41H64O13	763.4274	763.4285	-
46	Kanzonol C	11.26	C25H28O4	391.1915	391.1888	-
47	Doxorubicin hydrochloride	11.41	C27H30NO11Cl	578.1435	578.1383	-
48	Morin	26.71	C15H10O7	337.012	337.0092	-

Table 1: Continued.

No.	Target name	RT (min)	Formula	Expected (m/z)	Measured (m/z)	Peak number
49	Obacunone	6.57	C26H30O7	499.1974	499.1954	-
50	Forsythoside E	3.59	C20H30O12	461.1664	461.1671	-
51	Ginsenoside Rk1	27.00	C42H70O12	765.4794	765.4786	-
52	Raddeanin A	9.23	C47H76O16	931.4827	931.4911	-
53	Protodioscin	8.82	C51H84O22	1047.5382	1047.5389	-
54	Sibiricose A6	2.06	C23H32O15	547.1668	547.1677	-
55	Ginsenoside Rh2	26.97	C36H62O8	667.4427	667.4435	-
56	Madecassoside	9.72	C48H78O20	973.5014	973.4927	-
57	Qingyangshengenin	7.90	C28H36O8	535.2104	535.2148	-
58	Arctiin	7.32	C27H34O11	579.2083	579.208	-
59	Cucurbitacin B	15.48	C32H46O8	593.2887	593.2876	-
60	L-Tryptophan	0.86	C11H12N2O2	203.0826	203.082	-
61	Ophiopogonin D	7.45	C44H70O16	853.4591	853.4564	-
62	Polyphyllin I	7.45	C44H70O16	853.4591	853.4564	-
63	Artesunate	1.89	C19H28O8	419.1478	419.1501	-
64	Morusin	1.89	C25H24O6	419.15	419.1501	-
65	Loganin	0.95	C17H26O10	389.1453	389.146	-
66	Handelin	9.02	C32H40O8	597.2705	597.2689	-
67	Gypenoside A	26.62	C46H74O17	897.4853	897.4833	-
68	Sophoricoside	7.21	C21H20O10	477.1038	477.1042	-
69	Eleutheroside B/syringin	1.01	C17H24O9	417.1402	417.1404	-
70	Gomisin D	6.27	C28H34O10	529.2079	529.2064	-
71	Helicid	0.50	C13H16O7	329.0878	329.0879	-
72	Iso-astragaloside I	11.28	C45H72O16	903.4514	903.4485	-
73	11-Oxomogroside V	13.76	C60H100O29	1283.6278	1283.6276	-
74	6-Shogaol	27.00	C17H24O3	275.1653	275.1654	-
75	Pseudoprotodioscin	8.05	C51H82O21	1075.5331	1075.5281	-
76	Vincristine sulfate	12.01	C46H56N4O10	869.3978	869.4009	-
77	Digitonin	16.11	C56H92O29	1263.5418	1263.5448	-
78	Irinotecan hydrochloride	8.05	C33H39N4O6Cl	585.2719	585.2675	-
79	Engeletin	7.19	C21H22O10	433.114	433.1142	-
80	Lanatoside C	8.21	C49H76O20	1029.4912	1029.4956	-

evaluate the stability of the instrument, and PLS-DA was employed to characterize the metabolic perturbations among different groups. The features were filtered based on the variable importance in projection (VIP) values (VIP > 2) and fold change values (FC > 2) obtained from the initial group comparisons. Student's *t*-test was further applied to assess the significances of the differences in these features between groups using the GraphPad Prism 6.0 software (La Jolla, CA, USA). A value of $p < 0.05$ was considered to indicate statistical significance, and this criterion was used for the selection of biomarkers. The MetaboAnalyst software, which is a web-based tool for the visualization of metabolomic data, was used for the metabolic pathway analysis, and Impala network tools were used to predict signal pathways.

2.7. Microbial Community Analysis. Rat feces were collected and immediately stored at -80°C. Prior to DNA extraction, the samples were thawed and mixed well. Microbial DNA was extracted using the E.Z.N.A. to analyze the taxonomic composition of the bacterial community. An aliquot (200 mg) of each fecal sample was suspended in a mixture of 40 μL of 10% N-lauroylsarcosine, 0.1 M Tris (pH 7.5), and 250 μL of guanidine thiocyanate. The universal primers 338F and 806R were used to amplify the 16S rDNA gene using a PCR machine. For library construction, the PCR mixture system (50 μL) contained 30 ng of DNA template, 4 μL of the PCR primer cocktail, 25 μL of PCR master mix (NEB Phusion High-Fidelity PCR Master Mix), and an appropriate volume of ddH2O. Amplification of each sample was performed in triplicate. The amplicons were analyzed using 2% agarose gels, purified with the AxyPrep DNA Gel Extraction Kit (Axygen Biosciences, Union City, CA, USA), and quantified using QuantiFluor-ST (Promega, Madison, WI, USA) according to the manufacturer's instructions, and sequencing was then performed with an Illumina HiSeq platform (Illumina, San Diego, CA, USA). The sequencing data were

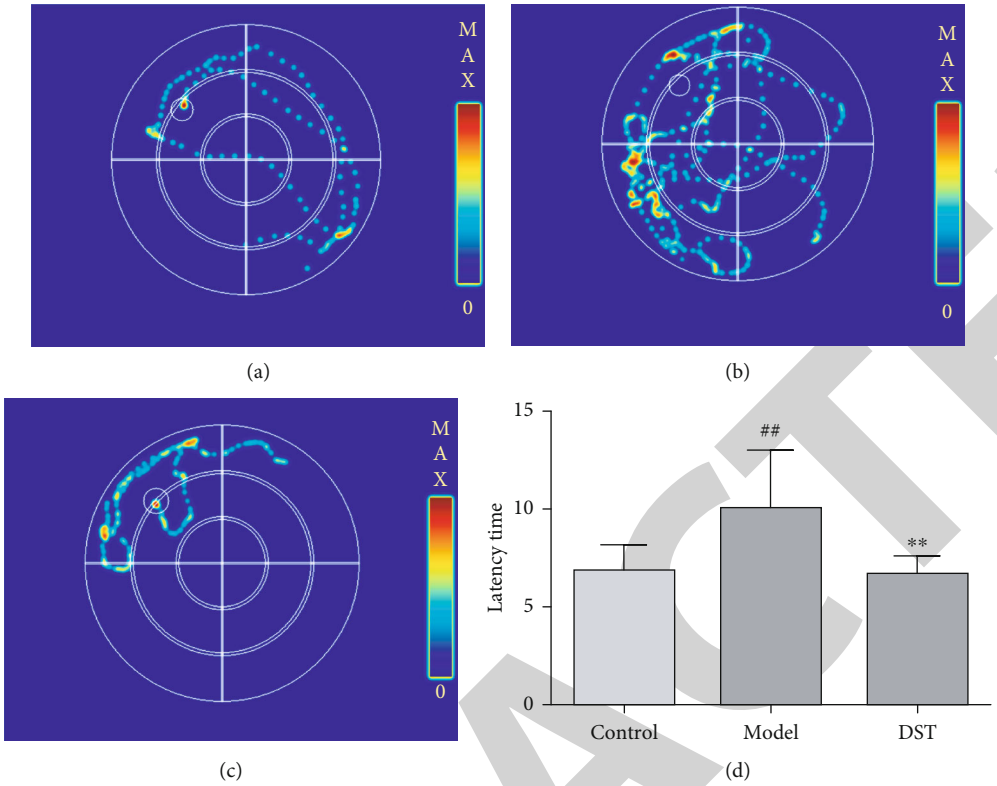


Figure 2: Effects of Dushen Tang (DST) on the escape latency and path traces in the Morris water maze test. (a–c) Path traces of each group in the Morris water maze test: (a) control group, (b) model group, and (c) Dushen Tang (DST) group. (d) Latency (time) in the Morris water maze test. All the values are expressed as the means (±SDs); $^{\#\#}p < 0.01$ vs. the control group; $^{**}p < 0.01$ vs. the model group.

filtered, trimmed, and used for the clustering of operational taxonomic units (OTUs) and taxonomic analysis. At the 97% similarity level, the final OTUs were generated based on the clustering results using Usearch (version 7.1). Weighted UniFrac distance-based PCoA was performed using the R package.

2.8. cAMP Signaling Pathway Verification. The proteins in the total cell lysate were separated using 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and were then transferred onto a polyvinylidene fluoride (PVDF) membrane. The membranes were blocked with 5% skim milk at 37°C for 2 h, washed with Tris-buffered saline and Tween 20 (TBST) buffer, and then probed with primary antibodies overnight at 4°C using β-actin as a loading control. After incubation with the appropriate secondary antibodies for 1 h at 37°C, protein bands were visualized using a chemiluminescence imaging system for quantitative analysis.

2.9. Statistical Analysis. Results are presented as the mean ± SEM. The statistical significance was analyzed using the GraphPad Prism 6.02 software (GraphPad Software Inc., San Diego, CA, USA). Statistical comparisons of western blot and MWM data were analyzed by two-tailed Student's t-test using the Statistical Package for Social Science program (SPSS 22.0, Chicago, IL, USA), and $p < 0.05$ was considered statistically significant.

3. Results

3.1. Main Compounds of DST. The main compounds of DST were identified by LC-MS. Under the present chromatographic and MS conditions, the compounds showed good separation within 30 min (Figure 1). The mass, retention time, and intensity of the peaks in each chromatogram were detected by TF. Only the compounds whose peak areas showed a higher signal than an established threshold (5.0×106) were considered the main compounds of DST. Comparisons with accurate mass and MS/MS data of natural products in OTCML databases identified a total of 80 main compounds in DST; however, some of the compounds were isomers, and thus, the retention time of the reference substance was used to distinguish the isomers. The 10 ginsenosides with isomers were identified by matching with the reference substance, marked on the spectrum, and were found at a high level in DST. The details of the main compounds identified in this study are summarized in Table 1.

3.2. Effect of DST on Rats in the Morris Water Maze Test. In the MWM, animals use their survival instinct to find an escape platform, and this test can thus be used to evaluate the spatial memory and learning abilities of rats. The trajectories of the rats in the model group were more complicated than those of the rats in the control group, and after DST treatment, the trajectories of the treated rats approached those of the rats in the control group (Figures 2(a)–2(c)).

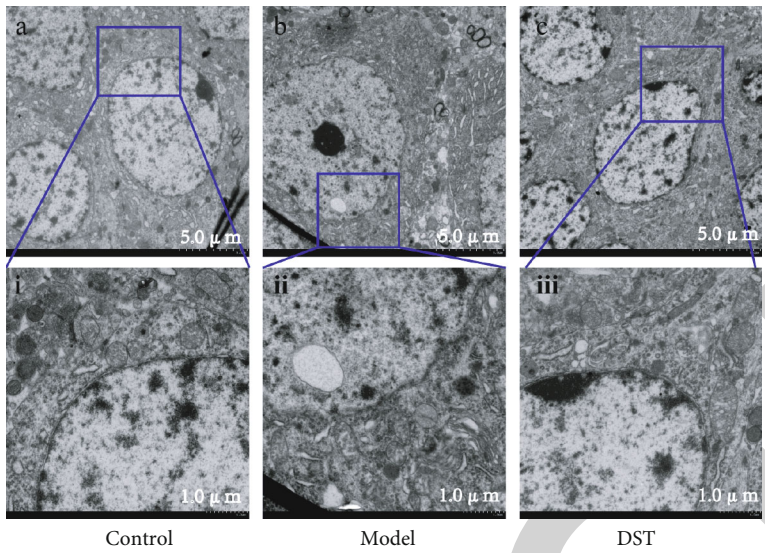


Figure 3: Ultrastructure of hippocampal neurons visualized under a transmission electron microscope. (a–c) Images at 1000x magnification, scale bar = 5.0 μm. (d–f) Images at 3000x magnification, scale bar = 1.0 μm.

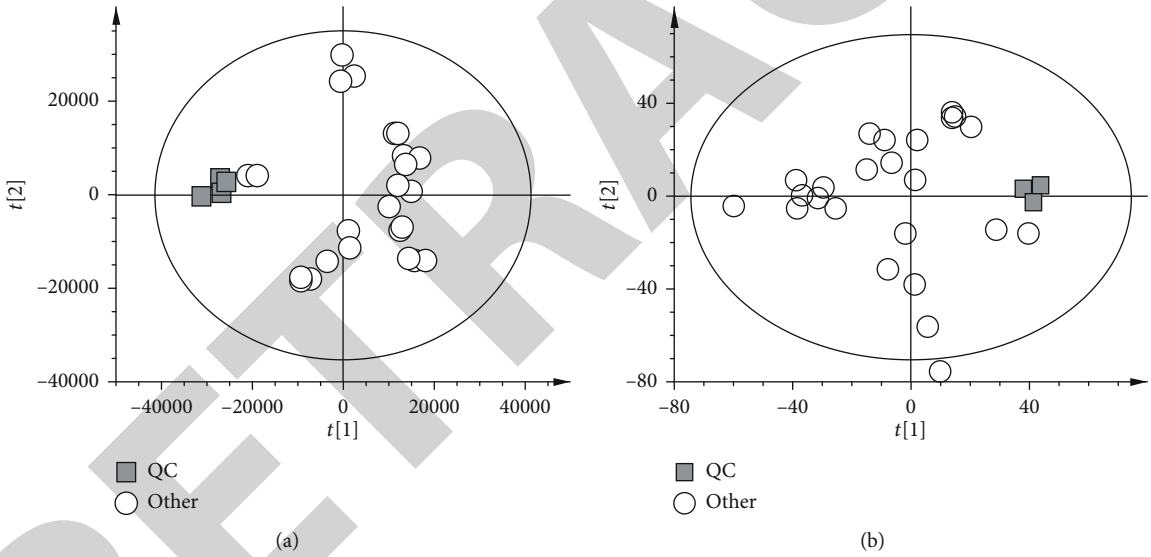


Figure 4: Principal component analysis (PCA) score plot of all the samples (○) highlighting the quality control (QC) (■) samples in the (a) positive and (b) negative ion modes.

The time that the rats in the model group needed to find the platform was significantly higher than that of the rats in the control group ($p < 0.01$), and that of the rats in the DST group was significantly shorter than that of the rats in the model group ($p < 0.01$) (Figure 2(d)). These findings suggest that DST could improve the spatial learning, memory, and cognitive abilities of rats with D-gal-induced memory impairment.

3.3. Effects of DST Treatment on the Ultrastructure of the Rat Hippocampus. To evaluate the protective effect of DST, the ultrastructure of neurons in the hippocampus of D-gal-induced rats was assessed by TEM. As shown in Figures 3(a) and 3(d), the hippocampal neurons of the rats in the control group were complete, clear, and well defined. These neurons exhibited a regular nuclear morphology, a smooth nuclear membrane, evenly distributed chromatin, clear nucleoli, abundant organelles, and normal-sized mitochondria. The model group showed obvious neuronal edema in the hippocampus; specifically, the perinuclear body was vacuolated, the nuclear envelope was ruptured, the electron density was high, the area of edema was enlarged, and mitochondria and other organelles were swollen (Figures 3(b) and 3(e)). Compared with those of the model rats, the hippocampal neurons of the rats in the DST group had a complete nuclear membrane, a uniform chromatin distribution, and normal-sized and normal-shaped mitochondria and organelles, which suggests that DST could reduce the neuronal

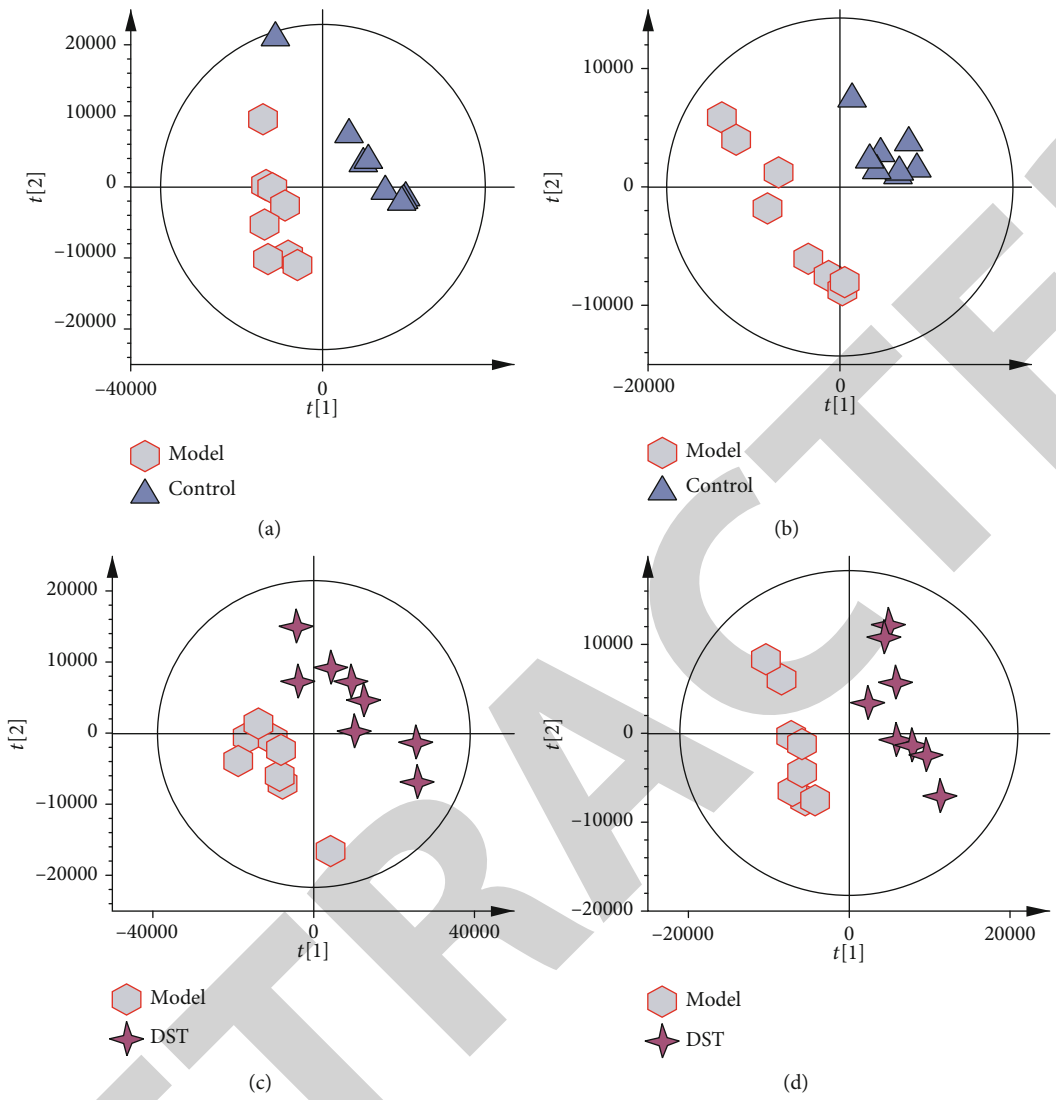


Figure 5: Partial least squares-discriminant analysis (PLS-DA) score plots of the control group (▲), model group (red hexagon), and Dushen Tang (DST) group (black diamond) based on data acquired in the (a, c) positive ion mode and in the (b, d) negative ion mode.

Table 2: Annotated potential biomarkers and the trends of their changes in serum.

No.	Compound name	m/z	Formula	Adduct	VIP	Model/control	DST/model
1	8-Hydroxycarapinic acid	509.15598	C26H30O8	[M+H]+	1.05956	$\downarrow^{*}$	$\uparrow^{\#\#\#}$
2	Pantothenic acid	220.11649	C9H17NO5	[M+H]+	1.93711	$\downarrow^{*}$	$\uparrow$
3	Creatinine	114.06604	C4H7N3O	[M+H]+	3.9793	$\uparrow^{*}$	$\downarrow$
4	Erucamide	338.33859	C22H43NO	[M+H]+	5.24912	$\uparrow^{***}$	$\downarrow^{\#\#}$
5	Threonic acid	135.02975	C4H8O5	[M-H]-	1.07714	$\downarrow^{*}$	$\uparrow$
6	Anserine	239.11552	C10H16N4O3	[M-H]-	3.35603	$\downarrow^{*}$	$\uparrow^{\#}$
7	3-Hydroxybutanoic acid	103.03831	C4H8O3	[M-H]-	4.24771	$\downarrow^{*}$	$\uparrow$
8	Phenacylamine hydrochloride	170.04254	C8H10ClNO	[M-H]-	2.82537	$\downarrow^{*}$	$\uparrow^{\#}$
9	Chlorothalonil-4-hydroxy	244.90831	C8HCl3N2O	[M-H]-	2.04014	$\downarrow^{**}$	$\uparrow^{\#}$
10	DL-Lactic acid	89.02377	C3H6O3	[M-H]-	3.35603	$\uparrow^{*}$	$\downarrow$

$^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$, the model group versus the control group; $^{\#}p < 0.05$, $^{\#\#}p < 0.01$, and $^{\#\#\#}p < 0.001$, the Dushen Tang (DST) group versus the model group.

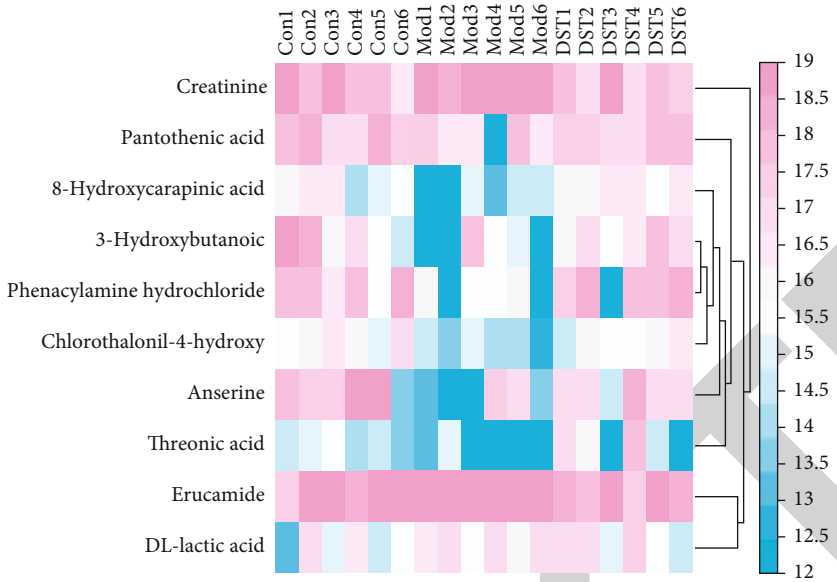


Figure 6: Heatmap showing the changes in the intensities of potential biomarkers.

damage in the hippocampus of rats with memory impairment (Figures 3(c) and 3(f)).

3.4. Metabolomic Profiling. It is important to verify the reproducibility and stability of the LC-MS method in metabolomic analysis [25]. Therefore, each QC sample was analyzed every six analytical runs to obtain a dataset for assessment of the reproducibility and stability of the method. Furthermore, after obtaining all the peaks from the serum samples and their alignment and normalization, the dataset containing information on the peaks of all the samples was used for PCA. As shown in Figure 4, the QC samples were clustered relatively tightly in comparison with all the samples, and this finding was obtained in both the positive and negative ion modes, which shows the reproducibility and stability of the LC-MS method.

The serum samples were analyzed by UPLC-Q-TOF/MS, and the raw data were subjected to multivariate statistical analysis through SIMCA-P to reveal the metabolite changes among the three groups and to evaluate the protective effect of DST against memory impairment. The differences in the metabolic profiles among the control, model, and DST groups were visualized by partial least squares-discriminant analysis (PLS-DA), and the results of the permutation test ($n = 100$) shown in Figure S1 revealed that the value of intercepts to the left was lower than the original value, which suggests that the established model was not overfit. The PLS-DA score plots (Figures 5(a) and 5(b)) showed that the model and control groups were completely separated by the principal components in both the positive and negative ion modes, which suggests that the blood metabolism of rats was disordered after model establishment. Furthermore, the cluster of the DST group was located far from that of the model group (Figures 5(c) and 5(d)), which indicates the protective effect of DST on memory repair, and this finding was consistent with the TEM and MWM findings.

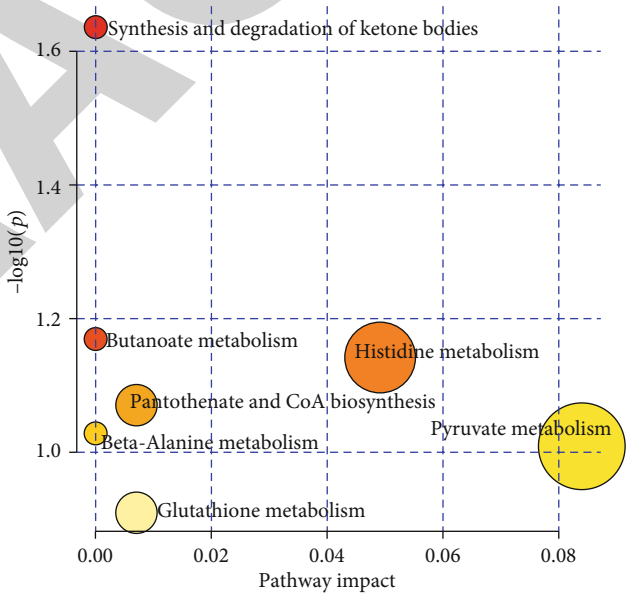


Figure 7: Bubble plots showing the altered metabolic pathways in rats with memory impairment.

The fold change (FC), p values, and VIP values were used to select metabolites with the greatest contribution to the separation between the control and model groups. A total of 33 metabolites, with $p < 0.05$ and FC > 2 or FC < 0.5, were identified as candidate biomarkers in the positive and negative ion modes. The structures of these candidate biomarkers were tentatively assigned by comparing the MS and MS/MS spectra with the detected metabolites in the Human Metabolome Database (HMDB), and the detailed information is shown in Table S1. Most of these candidate biomarkers were related to oxidative stress, neuroprotection, mitochondrial disorders, and Aβ deposition, which are involved in the pathogenesis of AD, and most of the

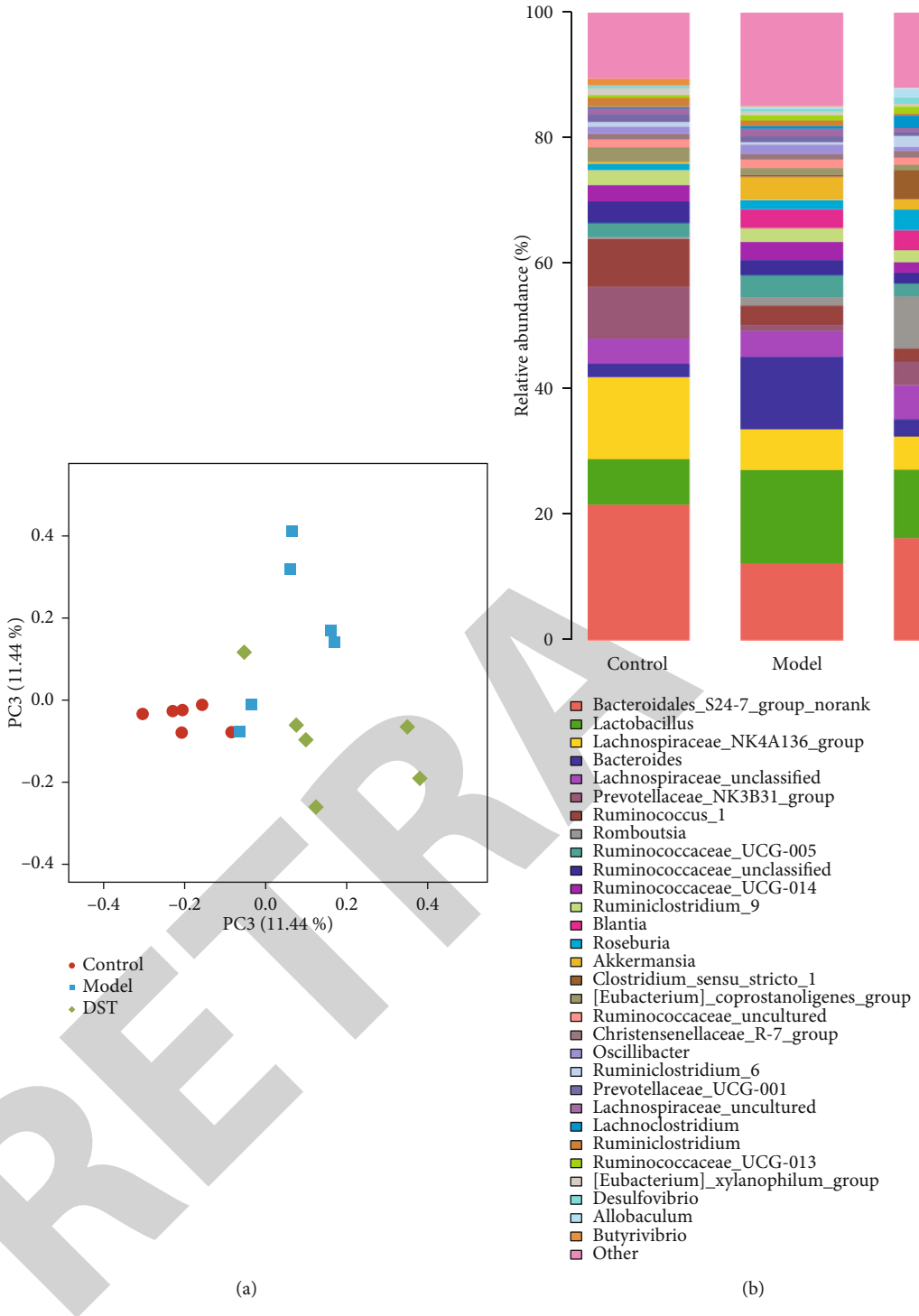


Figure 8: Continued.

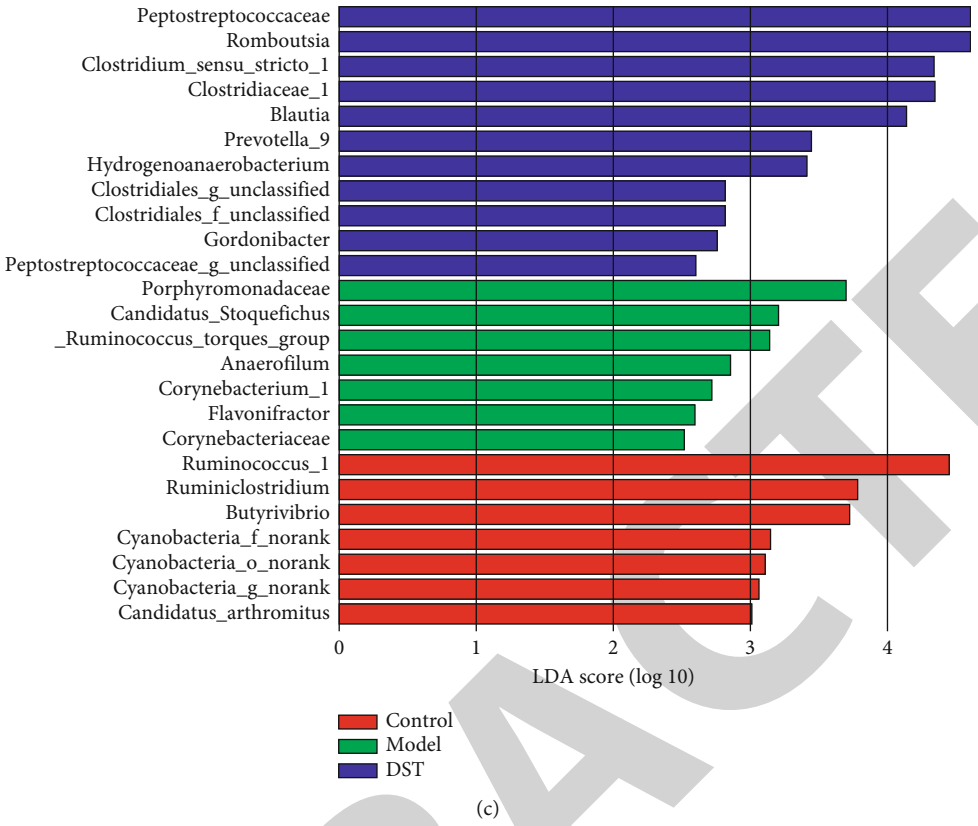


Figure 8: Fecal microbiota structure of the different groups. (a) Principal coordinates analysis (PCoA) plots of operational taxonomy units (OTUs) among the control group, model group, and Dushen Tang (DST) group. (b) Bacterial taxonomic profiling of the gut microbiota at the genus level in the different groups. (c) Differentially abundant microbial cladogram obtained by linear discriminant analysis effect size (LEfSe).

metabolic pathways involving these candidate biomarkers were related to the body's energy supply (Table S2). The candidate biomarkers with a VIP value larger than 1 were regarded as potential biomarkers (Table 2).

A heatmap was used to visualize the changes in the potential biomarkers among the groups and the clustering of the potential biomarkers (Figure 6). The color differences between the control and model groups indicate metabolic perturbations in the blood of rats with memory impairment, whereas in the rats treated with DST, the levels of these metabolites were partially recovered to those in the control group.

To reveal the metabolic pathways affected by memory impairment, a pathway analysis of the potential biomarkers was performed through a search of the KEGG database, and the results were visualized using R. The results showed that the relevant pathways influencing memory impairment, including pyruvate metabolism, synthesis and degradation of ketone bodies, pantothenate and CoA biosynthesis, glutathione metabolism, beta-alanine metabolism, butanoate metabolism, and histidine metabolism, were the most important metabolic pathways (Figure 7). By using Impala tools to analyze the signaling pathways involved in potential biomarkers, we found that 3-hydroxybutyric acid and lactic acid were involved in the activation of cAMP signaling pathways.

3.5. Microbial Community Analysis. After the blood metabolomics study, we also studied the effect of DST on the gut microbiota of D-gal-induced rats by performing 16S gene sequencing and found that the alterations induced by DST in the intestinal flora were mainly reflected at the genus level. The clustering of the microbiota structure in the different groups by weighted UniFrac distance-based PCoA showed good separation. The microbial composition of the rats with D-gal-induced memory impairment showed different clusters than that of the rats in the model group (Figure 8(a)). Furthermore, the abundance of microorganisms showed partial variation between the control and model groups, and DST treatment exerted a corrective effect on the abundances of some microbes, such as *Lactobacillus* and *Bacteroides* (Figure 8(b)). In addition, changes in the main microbial structure composition were found among the three groups (Figure 8(c)). In detail, *Ruminiclostridium*, *Butyrivibrio*, *Cyanobacteria_f_norank*, *Cyanobacteria_o_norank*, and *Cyanobacteria_g_norank* were the main functional microorganisms in the control group, and *Porphyromonadaceae*, *Anaerofilum*, *Corynebacterium_1*, *Flavonifractor*, *and Corynebacteriaceae* were the main functional microorganisms that showed alterations in the model group compared with the control group. Importantly, *Peptostreptococcaceae*, *Romboutsia*, *Clostridium_sensu_stricto_1*, *Clostridiaceae_1*, and *Hydrogenoanaerobacterium* were the main functional microorganisms identified in the DST group.

KEGG functional prediction analysis was performed using 18 different microbes at the genus level. The annotated

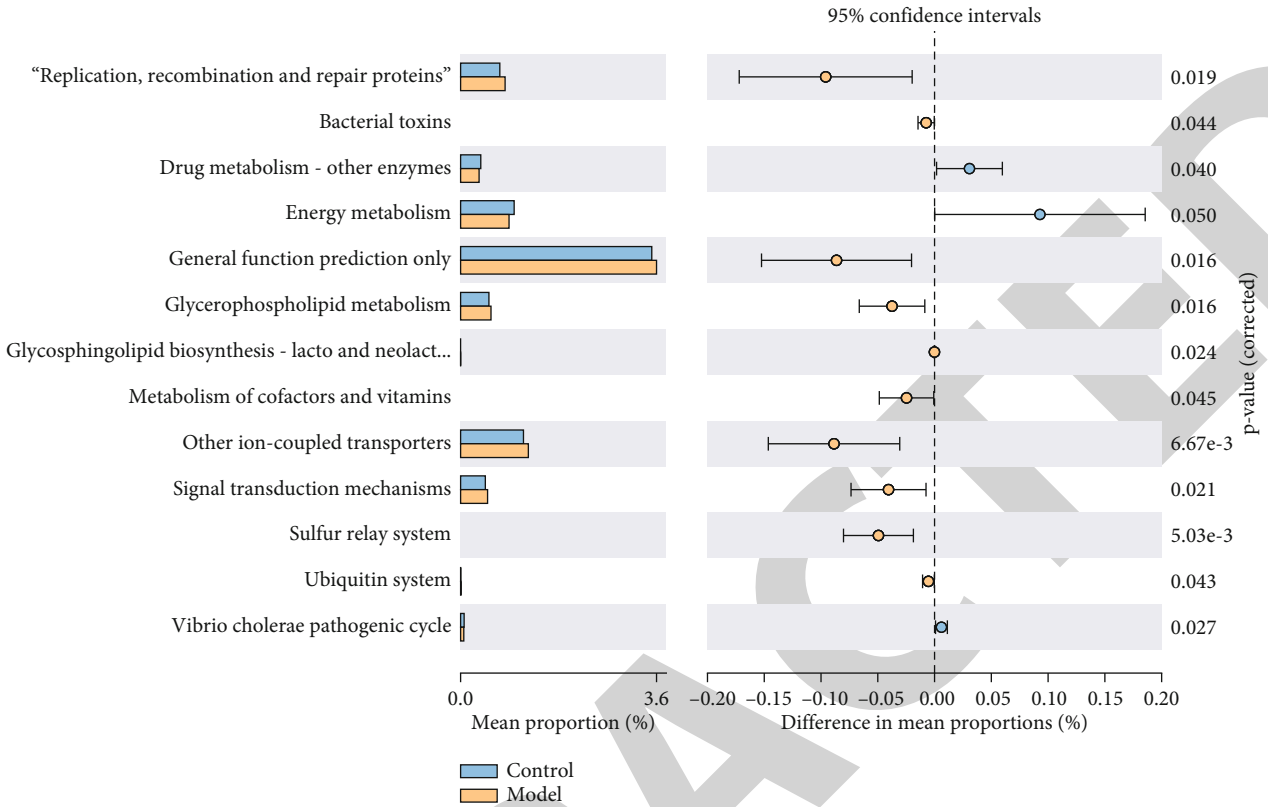


Figure 9: KEGG functional prediction analysis of the intestinal microbiota. The data are expressed as the means ± SEMs. The analysis shows the relative abundance of microbial genera based on Welch's test ($p < 0.05$). The colored circles represent 95% confidence intervals calculated using Welch's inverted method.

pathways represented potential pathways associated with the changes in the intestinal microbiota observed in the presence of memory impairment (Figure 9). These differential microbes were mainly related to energy metabolism, drug metabolism, other enzymes, signal transduction mechanisms, and sulfur relay systems.

3.6. The Influence of DST on Key Proteins in the cAMP Signaling Pathway. Our metabolomics study showed that the DST regulates energy-related especially cAMP signaling pathway. In order to validate the effect of DST on the cAMP signaling pathway, we evaluated additionally the expression of related proteins. As shown in Figure 10, the expression levels of cAMP, PKA, and ADCY, which are closely related to the nerve regeneration, were significantly reduced in the brain tissue of the rats in the model group compared with that in the control group ($p < 0.05$). Meanwhile, the expression level of GNAI was significantly increased in the brain tissue of the model group rats ($p < 0.05$). After DST treatment, the expression levels of cAMP, PKA, and ADCY were significantly increased ($p < 0.05$), while the expression level of GNAI was significantly reduced ($p < 0.05$). These results suggest that the DST improved memory impairment of the aging rats by activating the cAMP signaling pathway. Additionally, the expression of CREB and BDNF was detected, which two proteins are related to the cognitive function and are the downstream target of PKA. The results showed that compared with that in the control group, the expression of CREB and BDNF as significantly reduced in the brain tissue of the model group rats. After DST treatment, the expression of CREB and BDNF was significantly increased, indicating that DST can improve cognitive function and protect neurons by activating cAMP/PKA/CREB signaling pathways.

4. Discussion

Aging leads to brain damage and affects the quality of life of elderly populations, and DST is a widely used herbal decoction for senile diseases in China. In this study, the chemical profiles of DST were investigated by LC-MS-based chemical profiling. A total of 80 major ginsenosides were unambiguously identified in the DST group, and these included ginsenosides R1, Rg1, and Re. The MWM test and TEM analysis were used to assess the potential therapeutic effect of DST against memory impairment, and the results suggested that DST exerted good protective effects against D-galactose-induced memory impairment in rats. Subsequently, serum and fecal samples were analyzed by UPLC-MS and 16S rDNA sequencing to explore the mechanism through which DST affects metabolism and the intestinal flora to protect against cognitive impairment. Consistent with the MWM and TEM results, the PLS-DA score plot of the metabolites and PCoA score plot of the gut microbiota showed that the cluster of the DST group was separated from that of the model group. The metabolic results indicate that many endogenous energy-related metabolites (such as lactic acid

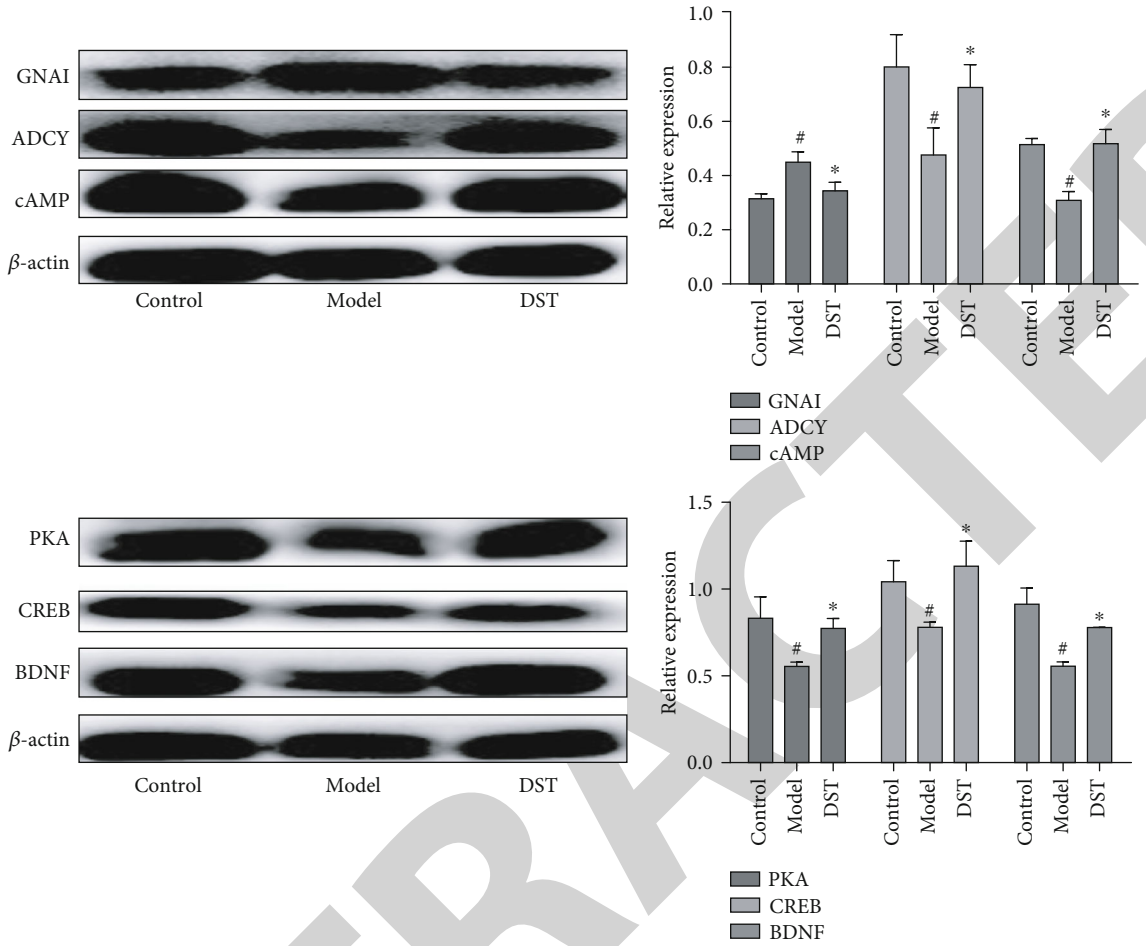


Figure 10: Expression levels of cAMP signal pathway-related proteins in rat brain tissues. The data are expressed as the means ± SEMs; $^{\#}p < 0.05$ vs. the control group; $^{*}p < 0.05$ vs. the model group.

and 3-hydroxybutyrate) that were dysregulated in the D-gal-induced rats could be corrected by the DST. The KEGG metabolic pathway analysis indicated that DST mainly regulates pyruvate metabolism and the synthesis and degradation of ketone bodies in rats with memory impairment. It was found that 3-hydroxybutyric acid and lactic acid were involved in the activation of cAMP signaling pathways.

Our study found that the levels of lactic acid, as an important biomarker of memory impairment, were increased in rats with memory impairment induced by D-gal. Lactic acid, which is the terminal product of lactic acid fermentation, is bioenergetically inefficient but favored under anoxic/hypoxic conditions and in pathological metabolic states [26]. Previous studies have confirmed that lactic acid is not only an energy substrate used preferentially by neurons but also plays an important role in memory formation [27], and this compound is found at higher levels in the frontal cortex and caudate nucleus of patients with AD [28]. These results are consistent with our findings.

It has been confirmed that patients with AD present with mitochondrial dysfunction and metabolic changes [29]. Mitochondrial dysfunction and a decline in respiratory chain function alter amyloid precursor protein processing, which leads to the production of pathogenic amyloid-β fragments [30–32]. Mitochondrial disorders are one of the main pathogenic factors of AD, and the size and shape of the mitochondria and organelles return to normal after DST treatment, which suggests the protective effect of DST [33, 34]. Lactic acid accumulation was observed in the model group with memory impairment, and this accumulation might have been observed due to decreased pyruvate dehydrogenase activity, which results in decreased catabolism of lactic acid and reduced mitochondrial respiration and thereby mitochondrial failure. Mitochondrial failure results in intracellular pH changes, synaptic dysfunction, amyloid deposition, and ultimately apoptosis [35, 36].

Mitochondrial dysfunction during aging is a trigger that induces many age-related changes in energy metabolism. Lactate and pyruvate are key intermediates in intracellular energy metabolic pathways and can indicate age-related imbalances in energy metabolism [37]. Lactate prevents cell aging through its lactate dehydrogenase-catalyzed oxidation to pyruvate, and pyruvate inhibits cell aging via its intramitochondrial metabolism [38]. This research has shown that DST can improve memory impairment in aging rats by regulating the metabolism of pyruvate. The mechanism of action of DST might involve acceleration of the oxidation of pyruvate by lactic acid to produce nicotinamide adenine

dinucleotide (NADH) and inhibit cell senescence. NADH is then converted into ATP through mitochondrial electron transfer and thereby provides energy to the cell [39]. Ketone bodies commonly refer to 3-hydroxybutyric acid together with acetoacetate and acetone [40] and exert neuroprotective effects on aging brain cells [30]; thus, CoA biosynthesis is highly correlated with energy metabolism. Notably, 3-hydroxybutanoic acid exerts a neuroprotective effect that supports the mitochondrial respiration system by reversing the inhibition of mitochondrial complex I or II [41]. High levels of 3-hydroxybutanoic acid have been shown to improve cognitive function in elderly and memory-impaired subjects [42, 43]. Additionally, 3-hydroxybutanoic acid can replace glucose as the major central nervous system fuel for the brain in humans [44], decrease oxidative stress by binding to specific receptors [45], promote the expression of brain-derived neurotrophic factor [46], and increase cerebral blood flow [47]. In our study, we found that DST increased the serum level of 3-hydroxybutyric acid, which confirmed that DST exerts beneficial effects on memory impairment by regulating the expression of 3-hydroxybutanoic acid and the related energy metabolism.

Additionally, according to our 16S rDNA sequencing results, we found that DST could correct the disturbance of the gut microbiota, including *Lactobacillus*, *Bacteroidales*, and *Bacteroides*, to improve the memory impairment induced by D-gal (Figure 8(b)). According to published references, *Bacteroidetes* are typically beneficial to human health due to their multiple capabilities, such as supporting glucose homeostasis [48–50], biosynthesizing polysaccharides, and transporting proteins, volatile fatty acids, and other nutrients [51, 52]. *Lactobacillus* can increase the activity of glutamate-cysteine ligase and the rate of glutathione synthesis, block superoxide anion and hydroxyl radical production, decrease the prooxidant content, and stimulate and reinforce the immune system [53–55]. In addition, by improving oxidative stress and neuroinflammation, *Lactobacillus* and *Bacteroidales* can clear plaques and decrease the plaque size and number [56]. *Lactobacillus* is mainly composed of probiotics, which can adjust the composition of the gut microflora and affect the gut barrier function [57]. *Lactobacillus* also leads to lower oxidative stress and inflammatory markers, both of which are the main factors responsible for aging, by affecting cytokine production [58–60]. *Lactobacillus* strains could exert the same mimetic antiaging effects of calorie restriction without an actual reduction in the calorie intake. AMPK can also be regarded as a novel AMPK activator [61] and regulates the dynamics of mitochondria to overcome oxidative stress by minimizing energy consumption and the exhaustion of resources, which results in the restoration of metabolic functions [62]. Furthermore, consistent with the metabolomics findings, the KEGG functional prediction analysis of the intestinal microbiota suggested that the important species of the gut microbiota that were regulated by DST (such as *Lactobacillus* and *Bacteroides*) are also mainly related to oxidative stress and energy metabolism, such as pyruvate, and the synthesis and degradation of ketone bodies.

Our study validated the effect of DST on cAMP signaling pathways from the protein expression. The cAMP is an important second messenger that mainly activates downstream cAMP-dependent protein kinase A (PKA) and promotes further activation of transcription factors, by which way fundamentally involves in brain mechanisms that require for the neuronal activity, nerve regeneration, energy production, metabolic processes, and synaptic physiology [63, 64]. In the cAMP signaling pathway, the transformation from ATP into cAMP is regulated by the ADCY, and the GNAI as the upstream target of ADCY prevents the generation of cAMP by inhibiting ADCY. The insufficient expression of ADCY by the D-gal induction can lead the reduction of the cAMP expression. The attenuated expression of the cAMP may further lead to the protein kinase (PKA) inactivation, following which to inhibit the expression of phosphorylated cAMP response element binding protein (CREB) [65]. The CREB is a molecular switch that is the core component of long-term memory. The reduction of CREB leads to a decline in memory ability. The phosphorylated CREB can stimulate the expression of brain-derived neurotrophic factor (BDNF), which belongs to neurotrophin family and is involved in learning and memory [66]. Our study found that the DST treatment promoted brain repair by participating protein synthesis for activating cAMP signaling pathway, especially by inhibiting the GNAI to activate the ADCY, following which to further activate its downstream target of cAMP and PKA, as well as the CREB and BDNF in rat brain tissue [67].

5. Conclusions

In conclusion, DST significantly improved the spatial memory and learning abilities of model rats and reduced the neuronal damage in the hippocampus of rats with memory impairment induced by D-gal. The 10 potential biomarkers identified by metabolomic analysis were related to pyruvate metabolism, the synthesis, and degradation of ketone bodies, pantothenate and CoA biosynthesis, and other metabolic pathways and were thus mainly related to energy metabolism. These biomarkers also play multiple functions in regulating oxidative stress, neuroprotection, mitochondrial disorders, and Aβ deposition related to aging and AD diseases. Furthermore, 16S rDNA sequencing followed by KEGG functional prediction analysis confirmed that the important species of the gut microbiota influenced by DST (such as Lactobacillus and Bacteroides) are also mainly related to oxidative stress and energy metabolism. In addition, this study has shown that DST inhibits nerve damage and exerts its antiaging effects by activating the cAMP signaling pathway.

Data Availability

All data are provided in this study, and raw data can be requested to the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

Na Li is financially supported by the National Natural Science Foundation of China (grant number 81704001) and Scientific and Technological Developing Project of Jilin Province (grant number 20200708056YY), and Min He is financially supported by the National Natural Science Foundation of China (grant number 82004030) and Scientific and Technological Developing Project of Jilin Province (grant number YDZJ202101ZYTS119).

Supplementary Materials

Figure S1: PLS-DA of the control and model groups based on the data obtained in the (a) positive ion mode and (b) the negative ion mode; PLS-DA of the model and DST groups based on the data obtained in the (c) positive ion mode and (d) the negative ion mode. Table S1: candidate biomarkers and their mode of action that might be related to memory impairment. Table S2: list of metabolic pathways related to candidate biomarkers ($p < 0.05$). *(Supplementary Materials)*

References

[1] T. Ali, H. Badshah, T. H. Kim, and M. O. Kim, "Melatonin attenuates D-galactose-induced memory impairment, neuroinflammation and neurodegeneration via RAGE/NF-K B/JNK signaling pathway in aging mouse model," *Journal of Pineal Research*, vol. 58, no. 1, pp. 71–85, 2015.

[2] Z. Sun, L. Sun, and L. Tu, "GABAB receptor-mediated PI3K/Akt signaling pathway alleviates oxidative stress and neuronal cell injury in a rat model of Alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 76, no. 4, pp. 1513–1526, 2020.

[3] S. Hira, U. Saleem, F. Anwar, Z. Raza, A. U. Rehman, and B. Ahmad, "In silico study and pharmacological evaluation of eplerinone as an anti-Alzheimer's drug in STZ-induced Alzheimer's disease model," *ACS Omega*, vol. 5, no. 23, pp. 13973–13983, 2020.

[4] L. Castillo-Mariqueo and L. Gimenez-Llort, "Indexes for flotation and circling, two non-search behaviors in the water maze, sensitive to d-galactose-induced accelerated aging and Alzheimer's disease," *Behavioural Brain Research*, vol. 377, article 112229, 2020.

[5] M. Peng, X. Fang, M. Miao, and T. Wang, "Effects of Wuweizi syrup on brain aging mice model induced by d-galactose," *Journal of King Saud University - Science*, vol. 32, no. 4, pp. 2426–2431, 2020.

[6] W. Xu, X. Liu, X. He et al., "Bajitianwan attenuates D-galactose-induced memory impairment and bone loss through suppression of oxidative stress in aging rat model," *Journal of Ethnopharmacology*, vol. 261, article 112992, 2020.

[7] H. Liu, X. Zhang, J. Xiao et al., "Astaxanthin attenuates d-galactose-induced brain aging in rats by ameliorating oxidative stress, mitochondrial dysfunction, and regulating metabolic markers," *Food & Function*, vol. 11, no. 5, pp. 4103–4113, 2020.

[8] D. Yang, X. Yang, H. Yan et al., "UPLC-MS/MS determination of twelve ginsenosides in Shenfu Tang and Dushen Tang," *International Journal of Analytical Chemistry*, vol. 2019, Article ID 6217125, 7 pages, 2019.

[9] S.-L. Li, S.-F. Lai, J.-Z. Song et al., "Decocting-induced chemical transformations and global quality of Du-Shen-Tang, the decoction of ginseng evaluated by UPLC-Q-TOF-MS/MS based chemical profiling approach," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 53, no. 4, pp. 946–957, 2010.

[10] S.-H. Choi, R. Lee, S. M. Nam et al., "Ginseng gintonin, aging societies, and geriatric brain diseases," *Integrative Medicine Research*, vol. 10, no. 1, article 100450, 2021.

[11] H.-J. Kim, S.-W. Jung, S.-Y. Kim et al., "*Panax* ginseng as an adjuvant treatment for Alzheimer's disease," *Journal of Ginseng Research*, vol. 42, no. 4, pp. 401–411, 2018.

[12] S. M. Nam, H. Hwang, M. Seo et al., "Gintonin attenuates D-galactose-induced hippocampal senescence by improving long-term hippocampal potentiation, neurogenesis, and cognitive functions," *Gerontology*, vol. 64, no. 6, pp. 562–575, 2018.

[13] K. F. Akhter, M. A. Mumin, E. M. K. Lui, and P. A. Charpentier, "Fabrication of fluorescent labeled ginseng polysaccharide nanoparticles for bioimaging and their immunomodulatory activity on macrophage cell lines," *International Journal of Biological Macromolecules*, vol. 109, pp. 254–262, 2018.

[14] Y. Zhang, X. Yang, S. Wang, and S. Song, "Ginsenoside Rg3 prevents cognitive impairment by improving mitochondrial dysfunction in the rat model of Alzheimer's disease," *Journal of Agricultural and Food Chemistry*, vol. 67, no. 36, pp. 10048–10058, 2019.

[15] R. Qi, R. Jiang, H. Xiao et al., "Ginsenoside Rg1 protects against d-galactose induced fatty liver disease in a mouse model via FOXO1 transcriptional factor," *Life Sciences*, vol. 254, article 117776, 2020.

[16] Q. Li, J. Zeng, M. Su, Y. He, and B. Zhu, "Acetylshikonin from Zicao attenuates cognitive impairment and hippocampus senescence in d-galactose-induced aging mouse model via upregulating the expression of SIRT1," *Brain Research Bulletin*, vol. 137, pp. 311–318, 2018.

[17] J. Budni, M. L. Garcez, F. Mina et al., "The oral administration of D-galactose induces abnormalities within the mitochondrial respiratory chain in the brain of rats," *Metabolic Brain Disease*, vol. 32, no. 3, pp. 811–817, 2017.

[18] M. M. Rinschen, J. Ivanisevic, M. Giera, and G. Siuzdak, "Identification of bioactive metabolites using activity metabolomics," *Nature Reviews. Molecular Cell Biology*, vol. 20, no. 6, pp. 353–367, 2019.

[19] Y. Wang, Z. Kuang, X. Yu, K. A. Ruhn, M. Kubo, and L. V. Hooper, "The intestinal microbiota regulates body composition through NFIL3 and the circadian clock," *Science*, vol. 357, no. 6354, pp. 912–916, 2017.

[20] Y.-L. Feng, G. Cao, D.-Q. Chen et al., "Microbiome-metabolomics reveals gut microbiota associated with glycine-conjugated metabolites and polyamine metabolism in chronic kidney disease," *Cellular and Molecular Life Sciences*, vol. 76, no. 24, pp. 4961–4978, 2019.

[21] Y. Fujii, T. T. T. Nguyen, Y. Fujimura et al., "Fecal metabolite of a gnotobiotic mouse transplanted with gut microbiota from a patient with Alzheimer's disease," *Bioscience, Biotechnology, and Biochemistry*, vol. 83, no. 11, pp. 2144–2152, 2019.

[22] D. Luo, K. Chen, J. Li et al., "Gut microbiota combined with metabolomics reveals the metabolic profile of the normal aging process and the anti-aging effect of FuFang Zhenshu TiaoZhi(FTZ) in mice," *Biomedicine & Pharmacotherapy*, vol. 121, article 109550, 2020.

[23] Y. Liu, T. Du, W. Zhang et al., "Modified Huang-Lian-Jie-Du decoction ameliorates Aβ synaptotoxicity in a murine model of Alzheimer's disease," *Oxidative Medicine and Cellular Longevity*, vol. 2019, Article ID 8340192, 27 pages, 2019.

[24] H. Zhou, J. Tai, H. Xu, X. Lu, and D. Meng, "Xanthoceraside could ameliorate Alzheimer's disease symptoms of rats by affecting the gut microbiota composition and modulating the endogenous metabolite levels," *Frontiers in Pharmacology*, vol. 10, article 1035, 2019.

[25] Y. Wang, W. Guo, Y. Liu et al., "Investigating the protective effect of gross saponins of Tribulus terrestris fruit against ischemic stroke in rat using metabolomics and network pharmacology," *Metabolites*, vol. 9, no. 10, p. 240, 2019.

[26] M. G. Vander Heiden, L. C. Cantley, and C. B. Thompson, "Understanding the Warburg effect: the metabolic requirements of cell proliferation," *Science*, vol. 324, no. 5930, pp. 1029–1033, 2009.

[27] A.-K. Bouzier-Sore, P. Voisin, P. Canioni, P. J. Magistretti, and L. Pellerin, "Lactate is a preferential oxidative energy substrate over glucose for neurons in culture," *Journal of Cerebral Blood Flow and Metabolism*, vol. 23, no. 11, pp. 1298–1306, 2003.

[28] P. Hong, X. Zhang, S. Gao, and P. Wang, "Role of monocarboxylate transporter 4 in Alzheimer disease," *Neurotoxicology*, vol. 76, pp. 191–199, 2020.

[29] R. H. Swerdlow, "Brain aging, Alzheimer's disease, and mitochondria," *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, vol. 1812, no. 12, pp. 1630–1639, 2011.

[30] M. Rusek, R. Pluta, M. Ułamek-Kozioł, and S. J. Czuczwar, "Ketogenic diet in Alzheimer's disease," *International Journal of Molecular Sciences*, vol. 20, no. 16, article 3892, 2019.

[31] H. M. Wilkins and R. H. Swerdlow, "Amyloid precursor protein processing and bioenergetics," *Brain Research Bulletin*, vol. 133, pp. 71–79, 2017.

[32] T. J. W. McDonald and M. C. Cervenka, "The expanding role of ketogenic diets in adult neurological disorders," *Brain Sciences*, vol. 8, no. 8, p. 148, 2018.

[33] J. O. Watzlawik, X. Hou, D. Fricova et al., "Sensitive ELISA-based detection method for the mitophagy marker p-S65-Ub in human cells, autopsy brain, and blood samples," *Autophagy*, pp. 1–16, 2020.

[34] L. Vaillant-Beuchot, A. Mary, R. Pardossi-Piquard et al., "Accumulation of amyloid precursor protein C-terminal fragments triggers mitochondrial structure, function, and mitophagy defects in Alzheimer's disease models and human brains," *Acta Neuropathologica*, vol. 141, no. 1, pp. 39–65, 2021.

[35] T. J. Gross, E. Doran, A. K. Cheema, E. Head, I. T. Lott, and M. Mapstone, "Plasma metabolites related to cellular energy metabolism are altered in adults with Down syndrome and Alzheimer's disease," *Developmental Neurobiology*, vol. 79, no. 7, pp. 622–638, 2019.

[36] J. Zheng, S.-J. Zheng, W.-J. Cai, L. Yu, B.-F. Yuan, and Y.-Q. Feng, "Stable isotope labeling combined with liquid chromatography-tandem mass spectrometry for comprehensive analysis of short-chain fatty acids," *Analytica Chimica Acta*, vol. 1070, pp. 51–59, 2019.

[37] S. Yanase, T. Ishii, K. Yasuda, and N. Ishii, "Metabolic biomarkers in nematode C. elegans during aging," *Advances in Experimental Medicine and Biology*, vol. 1134, pp. 163–175, 2019.

[38] Q. Li, D.-Q. Miao, P. Zhou et al., "Glucose metabolism in mouse cumulus cells prevents oocyte aging by maintaining both energy supply and the intracellular redox potential," *Biology of Reproduction*, vol. 84, no. 6, pp. 1111–1118, 2011.

[39] J. D. Biggers, D. G. Whittingham, and R. P. Donahue, "The pattern of energy metabolism in the mouse oöcyte and zygote," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 58, no. 2, pp. 560–567, 1967.

[40] C.-L. Hsieh, R. Koga, A. Furusho et al., "Enantioselective and simultaneous determination of lactate and 3-hydroxybutyrate in human plasma and urine using a narrow-bore online two-dimensional high-performance liquid chromatography system," *Journal of Separation Science*, vol. 41, no. 6, pp. 1298–1306, 2018.

[41] Y. Kashiwaya, T. Takeshima, N. Mori, K. Nakashima, K. Clarke, and R. L. Veech, "D-beta-hydroxybutyrate protects neurons in models of Alzheimer's and Parkinson's disease," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 10, pp. 5440–5444, 2000.

[42] M. A. Reger, S. T. Henderson, C. Hale et al., "Effects of β-hydroxybutyrate on cognition in memory-impaired adults," *Neurobiology of Aging*, vol. 25, no. 3, pp. 311–314, 2004.

[43] M. Ota, J. Matsuo, I. Ishida et al., "Effect of a ketogenic meal on cognitive function in elderly adults: potential for cognitive enhancement," *Psychopharmacology*, vol. 233, no. 21-22, pp. 3797–3802, 2016.

[44] N. Moller and K. Body, "Ketone body, 3-hydroxybutyrate: minor metabolite - major medical manifestations," *The Journal of Clinical Endocrinology and Metabolism*, vol. 105, no. 9, pp. 2884–2892, 2020.

[45] A. G. Ladurner, "Rheostat control of gene expression by metabolites," *Molecular Cell*, vol. 24, no. 1, pp. 1–11, 2006.

[46] S. F. Sleiman, J. Henry, R. Al-Haddad et al., "Exercise promotes the expression of brain derived neurotrophic factor (BDNF) through the action of the ketone body β-hydroxybutyrate," *eLife*, vol. 5, 2016.

[47] M. Maalouf, P. G. Sullivan, L. Davis, D. Y. Kim, and J. M. Rho, "Ketones inhibit mitochondrial production of reactive oxygen species production following glutamate excitotoxicity by increasing NADH oxidation," *Neuroscience*, vol. 145, no. 1, pp. 256–264, 2007.

[48] P. N. Alexandrov, J. M. Hill, Y. Zhao et al., "Aluminum-induced generation of lipopolysaccharide (LPS) from the human gastrointestinal (GI)-tract microbiome-resident *Bacteroides fragilis*," *Journal of Inorganic Biochemistry*, vol. 203, article 110886, 2020.

[49] H. L. Simpson and B. J. Campbell, "Review article: dietary fibre-microbiota interactions," *Alimentary Pharmacology & Therapeutics*, vol. 42, no. 2, pp. 158–179, 2015.

[50] A. Ticinesi, C. Tana, and A. Nouvenne, "The intestinal microbiome and its relevance for functionality in older persons," *Current Opinion in Clinical Nutrition and Metabolic Care*, vol. 22, no. 1, pp. 4–12, 2019.

[51] J. Durack and S. V. Lynch, "The gut microbiome: relationships with disease and opportunities for therapy," *The Journal of Experimental Medicine*, vol. 216, no. 1, pp. 20–40, 2019.

[52] P. Fathi and S. Wu, "Isolation, detection, and characterization of enterotoxigenic Bacteroides fragilis in clinical samples," *The Open Microbiology Journal*, vol. 10, no. 1, pp. 57–63, 2016.

[53] L. Desbonnet, L. Garrett, G. Clarke, J. Bienenstock, and T. G. Dinan, "The probiotic *Bifidobacteria infantis*: an assessment

Hindawi
BioMed Research International
Volume 2023, Article ID 9841321, 1 page
https://doi.org/10.1155/2023/9841321

Retraction

Retracted: Genome-Wide Screening Identifies Prognostic Long Noncoding RNAs in Hepatocellular Carcinoma

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Feng, X. Hu, K. Ma, B. Zhang, and C. Sun, "Genome-Wide Screening Identifies Prognostic Long Noncoding RNAs in Hepatocellular Carcinoma," *BioMed Research International*, vol. 2021, Article ID 6640652, 16 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6640652, 16 pages
https://doi.org/10.1155/2021/6640652

Research Article

Genome-Wide Screening Identifies Prognostic Long Noncoding RNAs in Hepatocellular Carcinoma

Yujie Feng, Xiao Hu, Kai Ma, Bingyuan Zhang, and Chuandong Sun

Department of Hepatobiliary and Pancreatic Surgery, The Affiliated Hospital of Qingdao University, Qingdao, Shandong Province 266003, China

Correspondence should be addressed to Bingyuan Zhang; bingyuanzhang@126.com
and Chuandong Sun; sunchuandong@hotmail.com

Received 19 October 2020; Revised 18 April 2021; Accepted 23 April 2021; Published 21 May 2021

Academic Editor: Min Tang

Hepatocellular carcinoma (HCC) is a common malignancy with a poor prognosis. Therefore, there is an urgent call for the investigation of novel biomarkers in HCC. In the present study, we identified 6 upregulated lncRNAs in HCC, including LINC01134, RHPN1-AS1, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1. Higher expression of these lncRNAs was correlated to a more advanced cancer stage and a poorer prognosis in HCC patients. Enrichment analysis revealed that these lncRNAs played a crucial role in HCC progression, possibly through a series of cancer-related biological processes, such as cell cycle, DNA replication, histone acetyltransferase complex, fatty acid oxidation, and lipid modification. Moreover, competing endogenous RNA (ceRNA) network analysis revealed that these lncRNAs could bind to certain miRNAs to promote HCC progression. Loss-of-function assays indicated that silencing of RHPN1-AS1 significantly suppressed HCC proliferation and migration. Though further validations are still needed, these identified lncRNAs could serve as valuable potential biomarkers for HCC prognosis.

1. Introduction

Hepatocellular carcinoma (HCC) has ranked as the second leading cause of cancer-related deaths worldwide [1], as more than 782,000 HCC-associated deaths are predicted to occur annually [2]. Previous studies have demonstrated that the upregulation of oncogenes and downregulation of tumor suppressors were related to HCC tumorigenesis. For example, ID1 enhanced tumor growth in HCC patients [3], while FABP4 inhibited tumor growth and invasion [4]. However, there is still an urgent call to explore new biomarkers in HCC, as novel therapeutic targets are needed for establishing future molecular therapies, in the hope of improving prognoses of HCC patients.

Noncoding RNAs (ncRNAs) have emerged as a very promising resource for the identification of prognostic biomarkers. ncRNAs longer than 200 bps are regarded as long noncoding RNAs (lncRNAs) [5]. Growing evidence has confirmed the association between HCC and lncRNAs [6]. For example, XIST [7–9] and LINC01138 [10] were reported to link to the regulation of HCC growth and metastasis. Specifically, lncRNA XIST regulates HCC tumor growth and migration by sponging miR-497-5p [9]. DANCR retained HCC stemness by suppressing CTNNB1 expression [11]. The LINC01138 drives carcinogenesis in HCC via activating arginine methyltransferase 5 [10]. Meanwhile, several lncRNAs have been demonstrated as potential prognostic biomarkers in HCC. For example, overexpression of lncRNA ENST00000429227.1, LINC00511, SNHG16, and AK001796 were associated with worse prognosis in HCC patients [12] . A further appreciation of the molecular functions and expression patterns of lncRNAs could help reveal novel biomarkers for HCC.

In this study, using gene expression data of HCC patients and healthy controls, we screened possible prognostic lncRNAs in HCC. Additionally, we utilized bioinformatic approaches to explore the potential functions of such lncRNAs and constructed competing endogenous RNA (ceRNA) networks to further demonstrate potential mechanisms concerning those identified lncRNAs in HCC

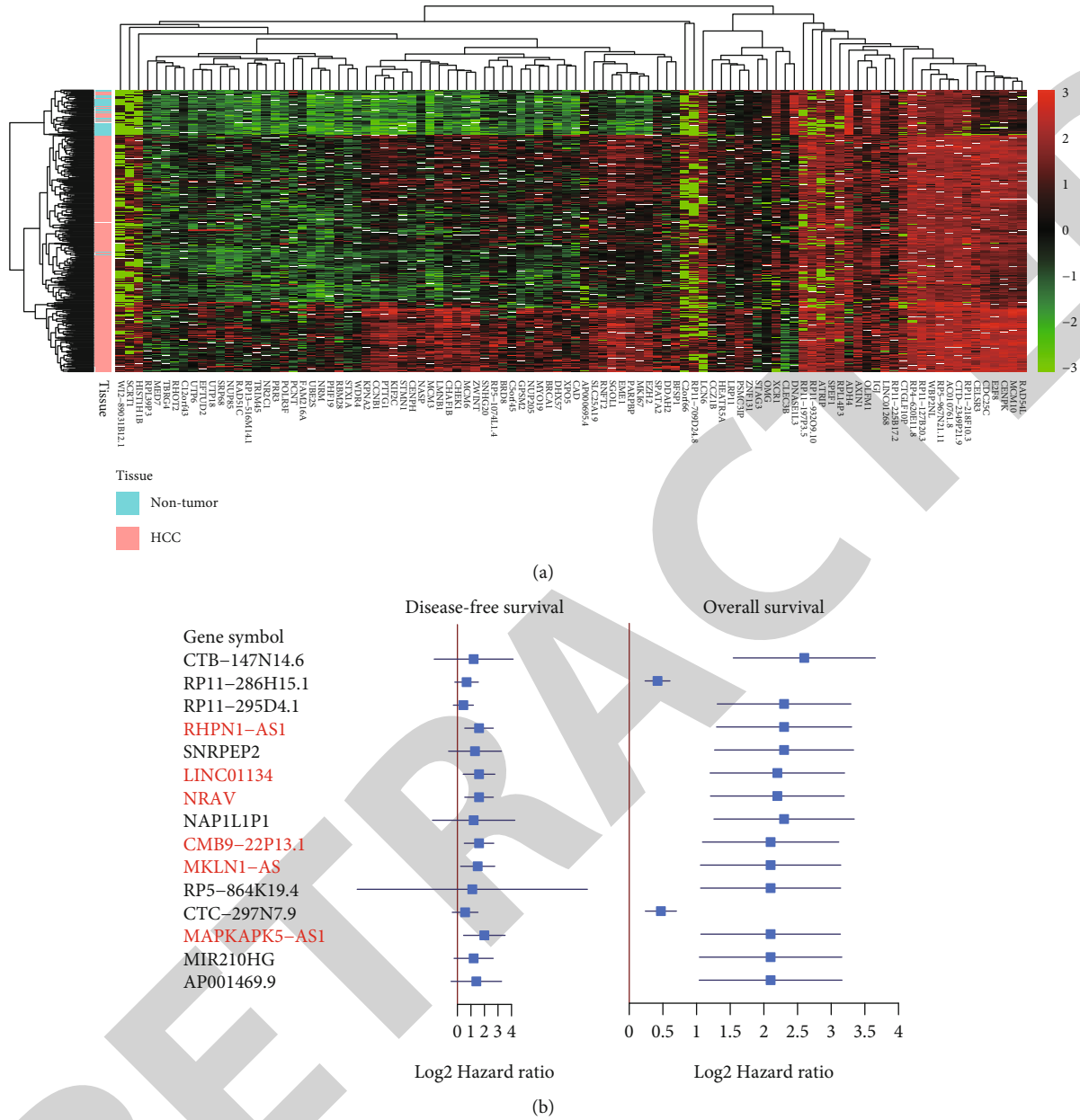


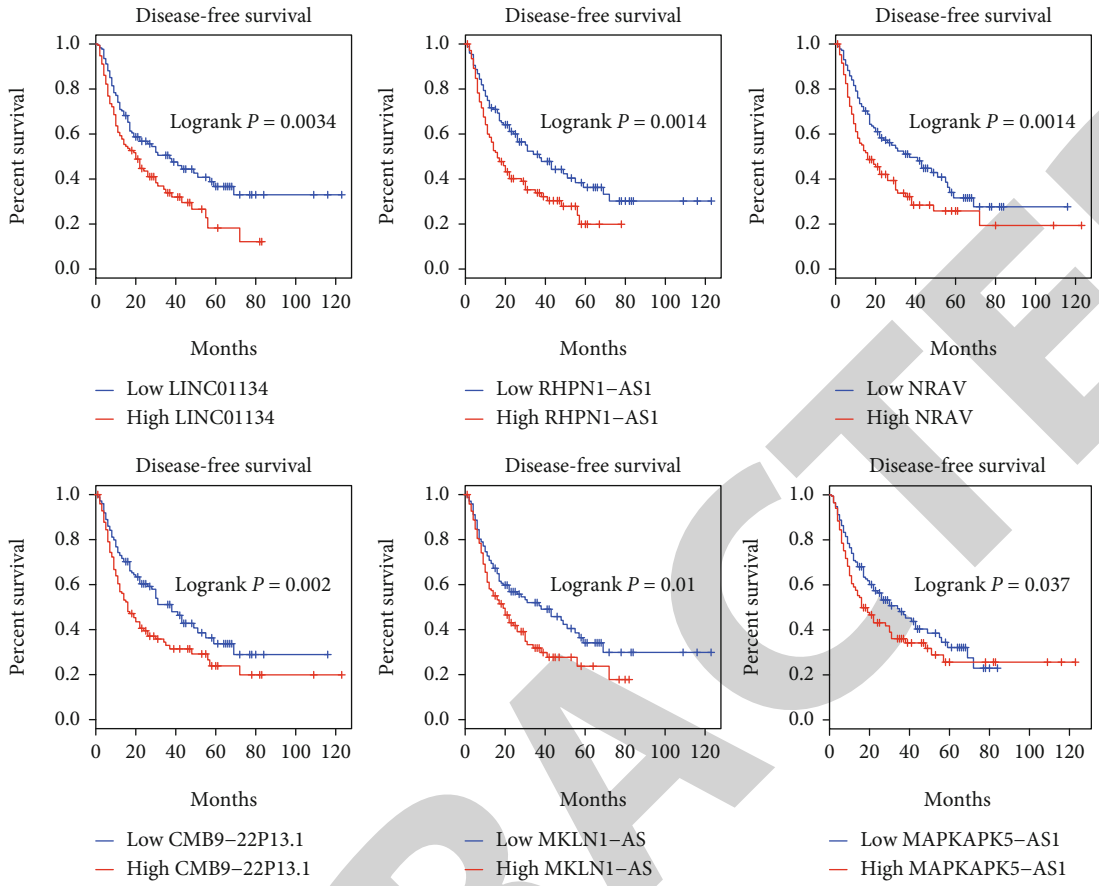
Figure 1: LINC01134, RHPN1-AS1, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1 were associated with both disease-free survival and overall survival time in HCC. (a) The expression profiles of the 100 differentially expressed lncRNAs in HCC and nontumor tissues. (b) The two forest plots of disease-free survival and overall survival for the 15 genes with significant association with overall survival. The gene symbols with red color are associated with both disease-free survival and overall survival.

development and progression. Loss-of-function assays were performed to investigate potential functions of key lncRNAs in HCC. We believe that this study could provide valuable markers in HCC and inspire further researches.

2. Materials and Methods

2.1. Clinical Samples. From April 2016 to March 2020, 6 paired HCC and adjacent normal tissues were collected at the Affiliated Hospital of Qingdao University. The ethics committee of this hospital approved this study before the enrollment of patients. All patients or their parents signed informed consent.

2.2. Data Preparation and Processing. The gene list of differently expressed lncRNAs in HCC was downloaded from the GEPIA dataset (http://gepia.cancer-pku.cn/). GEPIA is a newly developed interactive web server for analyzing the RNA sequencing expression data of 9,736 tumors and 8,587 normal samples from the TCGA and the GTEx projects using a standard processing pipeline, according to a previous report [13]. Specifically, GEPIA normalized gene expression by transcript per million (TPM), employed the R limma package to conduct differential expression analysis, and provided a list of differentially expressed genes ranked by the P values. lncRNAs with $\log 2 \mid FC \mid > 2.0$ and P value < 0.001 were considered as differentially expressed lncRNAs, and



(a)

Figure 2: Continued.

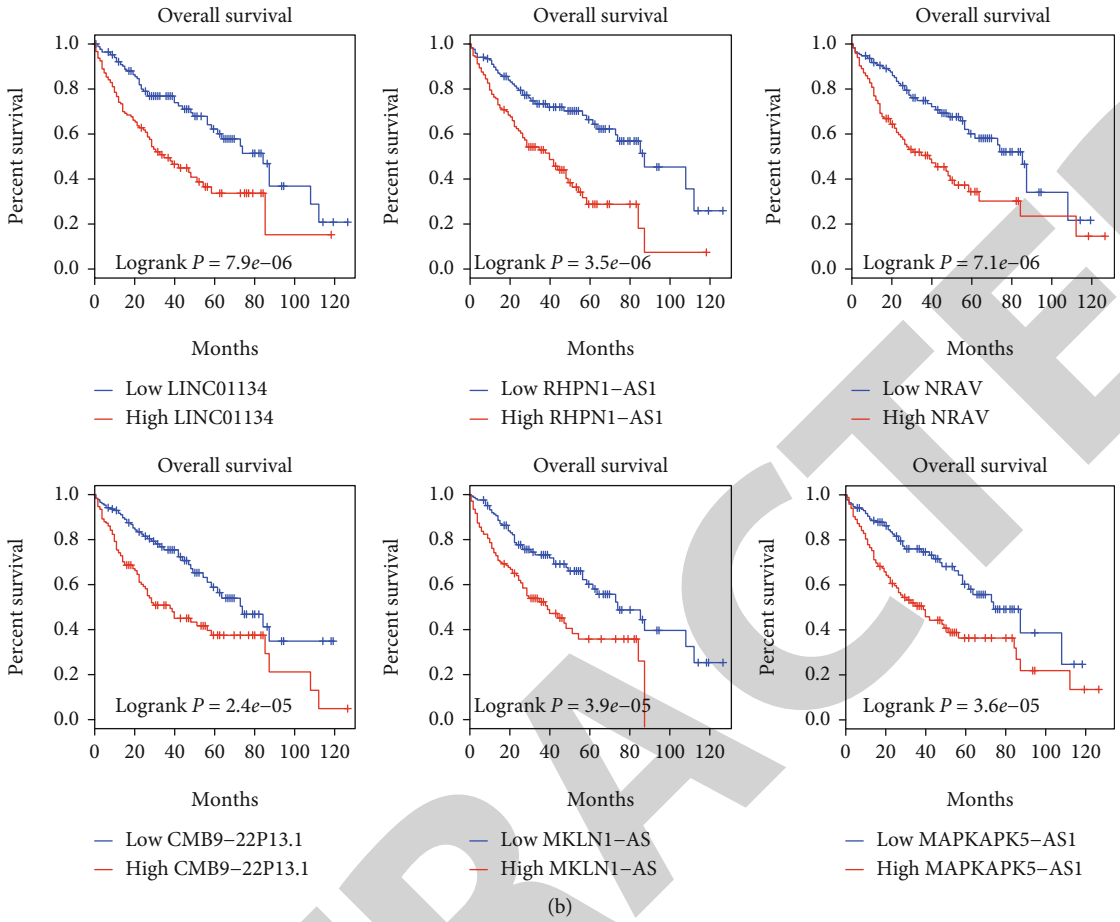


Figure 2: LINC01134, RHPN1-AS1, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1 were related to disease-free and overall survival time in HCC. The Kaplan-Meier (KM) curves of disease-free survival (a) and overall survival (b) for six prognostic lncRNAs in HCC.

the top 100 lncRNAs ranked by the *P* values were selected for downstream analysis. Moreover, those lncRNAs associated with survival time in HCC were also downloaded from the GEPIA website [13]. The median expression of each gene was considered as the cutoff for high and low expression. The Cox hazard regression analysis was conducted to test the association between the expression of any given gene and the survival time. The *P* values for survival analysis were calculated by log-rank test.

2.3. Construction of lncRNA-miRNA-mRNA Network and Further Analysis. The lncRNA-miRNA-mRNA network was built following several steps. First, the lncRNA-mRNA coexpression patterns were obtained from the GEPIA database [13]. The top 200 coexpressing genes by correlation analysis were selected as the potential targets of each lncRNA. Second, the miRNA-lncRNA and miRNA-mRNA interactions were predicted using the STARBASE database [14] . Finally, Cytoscape [15] software was used to visualize such lncRNA-miRNA-mRNA networks.

2.4. Pathway and Function Enrichment Analysis. Enrichment analyses were performed using the ClueGO plug-in [16], in order to investigate biological functions of six lncRNAs in HCC. The hypergeometric test was employed to determine the statistical significance.

2.5. Loss-of-Function Assays. HepG2 and Huh7 were obtained from Shanghai Institutes for Biological Sciences (Shanghai, China) and cultured according to the manufacturer's instruction. The following siRNAs were utilized in this study: si-RHPN1-AS1: 5′-ACAGCTATATCAGCCAACCAGAGT-3′; si-NC: 5′-GTTTACAACACGCTTCCTCTGA-3′. Transfection was performed using a Lipofectamine 2000 Reagent (Invitrogen). The real-time PCR was conducted according to published procedures in a previous report [17]. The following primers were utilized: RHPN1-AS1 5′-CTAGCCAGGAGGTTTCGC-3′ and 5′-TCCGCAACAAGCACACA-3′, LINC01134 5′-GGAGTTGGCTCCATCCTGAG-3′ and 5′-CTGGCATAGGGGTAACCTCA-3′, NRAV 5′-TCACTACTGCCCCAGGATCA-3′ and 5′-GGTGGTCACAGGACTCATGG-3′, CMB9-22P13.1 5′-AAGGCCCATGTAGCATCCC-3′ and 5′-TCTGTAAGGGAGAACCTGCCA-3′, MKLN-AS 5′-TCTGAAAGCAGCGCTTGGTA-3′ and 5′-GCGGAGTCCTCAAGGTATGG-3′, MAPKAPK5-AS1 5′-TCCCTAAGACACGCCGCATA-3′ and 5′-CGTGAATCTCCGCAGAGTGG-3′, ACTB 5′-CATGTACGTTGCTA

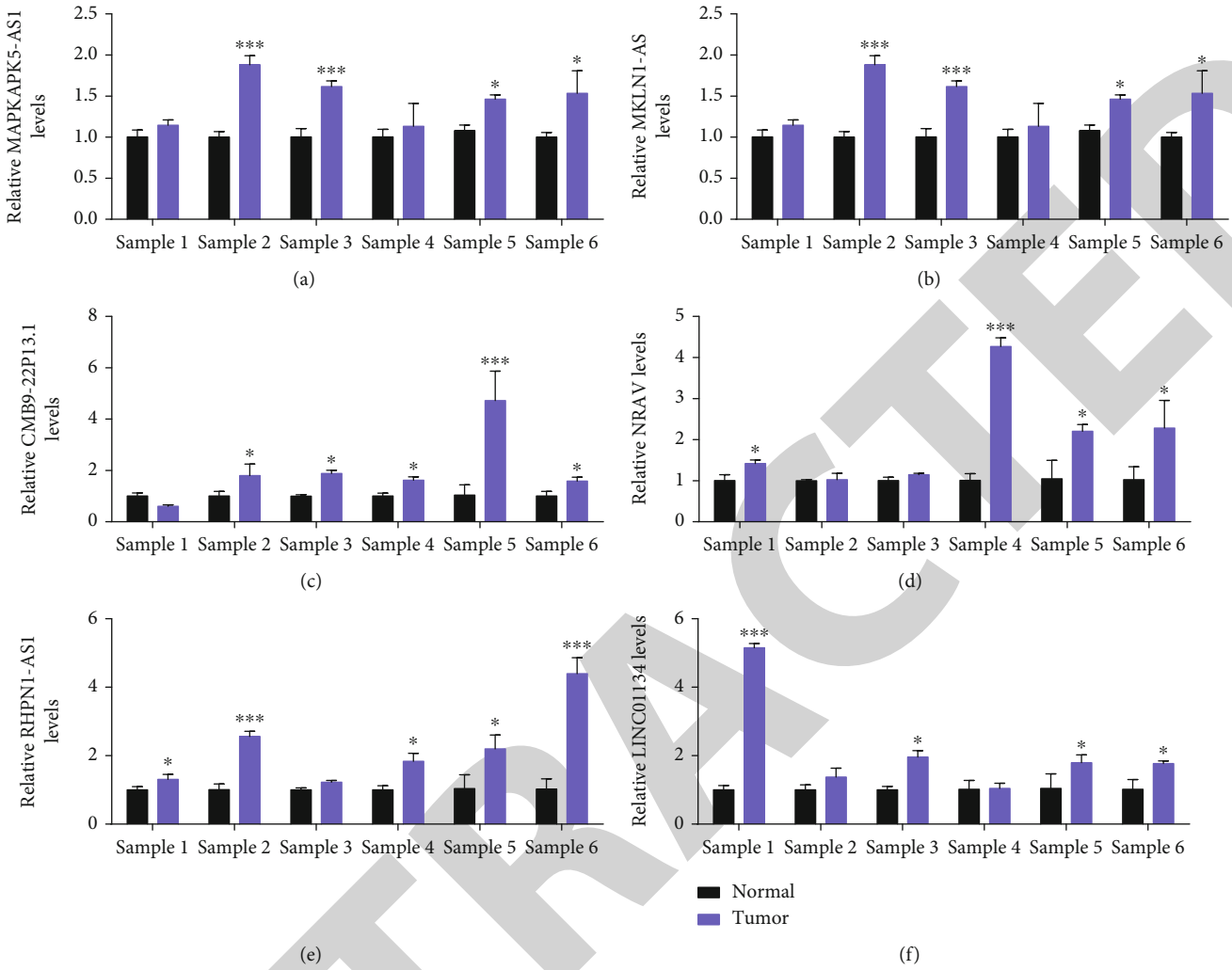


Figure 3: Upregulation of MAPKAPK5-AS1, MKLN1-AS, CMB9-22P13.1, NRAV, RHPN1-AS1, and LINC01134 in HCC and adjacent tissue samples. (a–f) MAPKAPK5-AS1 (a), MKLN1-AS (b), CMB9-22P13.1 (c), NRAV (d), RHPN1-AS1 (e), and LINC01134 (f) were upregulated in six pairs of HCC and adjacent normal tissues. The symbols of ∗, ∗∗, and ∗∗∗ indicate the P values < 0.05, 0.01, and 0.001, respectively.

TCCAGGC-3′ and 5′-CTCCTTAATGTCACGCACGAT-3′, and GAPDH 5′-CACCCACTCCTCCACCTTTG-3′ and 5′-CCACCACCCTGTTGCTGTAG-3′. In order to specify the effect of RHPN1-AS1 on cell viability, CCK-8 kit was utilized [18]. Meanwhile, a transwell assay was applied to detect cell migratory ability [17].

2.6. Statistical Analysis. The data analyses were conducted using GraphPad Prism 6.0 software (GraphPad Software Inc., USA). All values were denoted as the mean ± SD. Student's t test was used to evaluate differences between two groups, while the statistical significance from multiple comparisons was assessed with one-way analysis of variance (ANOVA), and P value < 0.05 was regarded as statistically significant.

3. Results

3.1. Screening of Prognostic lncRNAs in HCC. To identify the differentially expressed genes in HCC, we conducted differential expression analysis on the TCGA cohort using the GEPIA webserver. Specifically, we selected the top 100 upregulated lncRNAs ranked by the P values for downstream analysis (See Materials and Methods). As shown in Figure 1 (a), these lncRNAs had significantly higher expression levels in HCC tissues than in nontumor tissues. Survival analysis of these lncRNAs revealed that 15 lncRNAs, including CTB-147N14.6, RP11-286H15.1, RP11-295D4.1, RHPN1-AS1, SNRPEP2, LINC01134, NRAV, NAP1L1P1, CMB9-22P13.1, MKLN1-AS, RP5-864K19.4, CTC-297N7.9, MAPKAPK5-AS1, MIR210HG, and AP001469.9, were closely associated with overall survival (OS) time of patients with hepatocellular carcinoma (Figure 1(b), log-rank test, P value < 0.05). Similarly, survival analysis was also conducted to identify lncRNAs associated with disease-free survival (DFS). Consequently, only 6 lncRNAs were found to be significantly associated with shorter DFS in patients in HCC (Figures 1(b) and 2(a)), suggesting that these lncRNAs played more critical roles in the progression of HCC. It should be

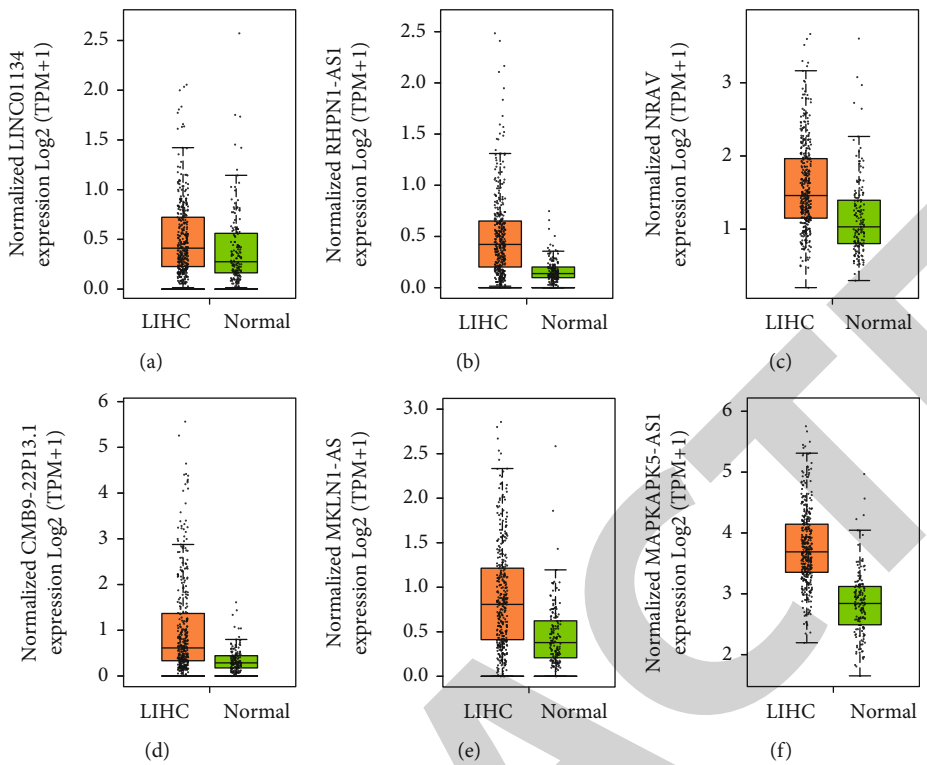


Figure 4: LINC01134, RHPN1-AS1, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1 were upregulated in HCC. (a–f) LINC01134 (a), RHPN1-AS1 (b), NRAV (c), CMB9-22P13.1 (d), MKLN1-AS (e), and MAPKAPK5-AS1 (f) were upregulated in HCC.

noted that higher expression of these six lncRNAs was associated with shorter OS and DFS, indicating that they might function as cancer-promoting regulators in HCC. In addition, we also investigated the prognostic values of top 100 downregulated lncRNAs in HCC and found that three lncRNAs, including CTD-2284J15.1, RP11-612B6.2, and RP4-601P9.2, were weakly associated with HCC overall survival (P value < 0.1, Supplementary Table (S1). These results indicated that the upregulated lncRNAs might play important roles in tumorigenesis or progression.

3.2. Validating the Upregulation of Prognostic lncRNAs in HCC Tissues. To validate the upregulation of prognostic lncRNAs in HCC tissues, we collected 6 pairs of HCC and adjacent normal tissues. Specifically, the six prognostic lncRNAs, including LINC01134, RHPN1-AS1, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1, were confirmed to be upregulated in most of HCC samples (Figure 3, Supplementary S2), suggesting that the upregulation of the six prognostic lncRNAs could also be observed in our tissue samples.

3.3. The Expression Patterns of Potential Tumor-Promoting lncRNAs in HCC and Nontumor Tissues. To examine the expression patterns of these six potential tumor-promoting lncRNAs in HCC, we further evaluated their differential expression levels between HCC and nontumor tissues. As illustrated in Figure 4, expressions of LINC01134, RHPN1-AS1, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1 were remarkably higher in HCC than in normal samples (Figures 4(a)–4(f)). Moreover, we also noticed that the expressions of these lncRNAs were higher in samples of more advanced HCC stages. Specifically, compared with stage I samples, LINC01134, RHPN1-AS1, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1 were expressed higher in II/III stage samples (P value < 0.05, Figures 5(a)–5(f)). However, differences in the expression of the above lncRNAs were not observed between the stage I and IV samples, mostly due to the small sample size of the stage IV group ($n = 5$). These results indicated that these prognostic lncRNAs were highly expressed in HCC with advanced stages.

3.4. Biological Functions of the Potential Tumor-Promoting lncRNAs in HCC. To elucidate the biological functions of these six prognostic lncRNAs, we conducted both correlation analysis and gene set enrichment analysis. Overall, the six prognostic lncRNAs were highly associated with genes involved in the cancer-associated pathways. Specifically, we observed that CMB9-22P13.1 was involved in regulating cell cycle DNA replication, cell cycle, unwinding of DNA, mitotic G1-G1/S phases, mismatch repair, and chromosome organization (Figure 6(a)). Moreover, LINC01134 was involved in regulating histone acetyltransferase complex through ATXN7L3, BRD8, and MRGBP and might regulate N-acetyltransferase activity through MED24, HCFC1, NAA40, and NAT9 (Figure 6(b)). MAPKAPK5-AS1 was associated with the regulation of metabolism of nucleotides, mRNA splicing, RNA processing, ribonucleoprotein complex assembly, and establishment of protein localization to the telomere (Figure 6(c)). NRAV was predicted to be involved in the

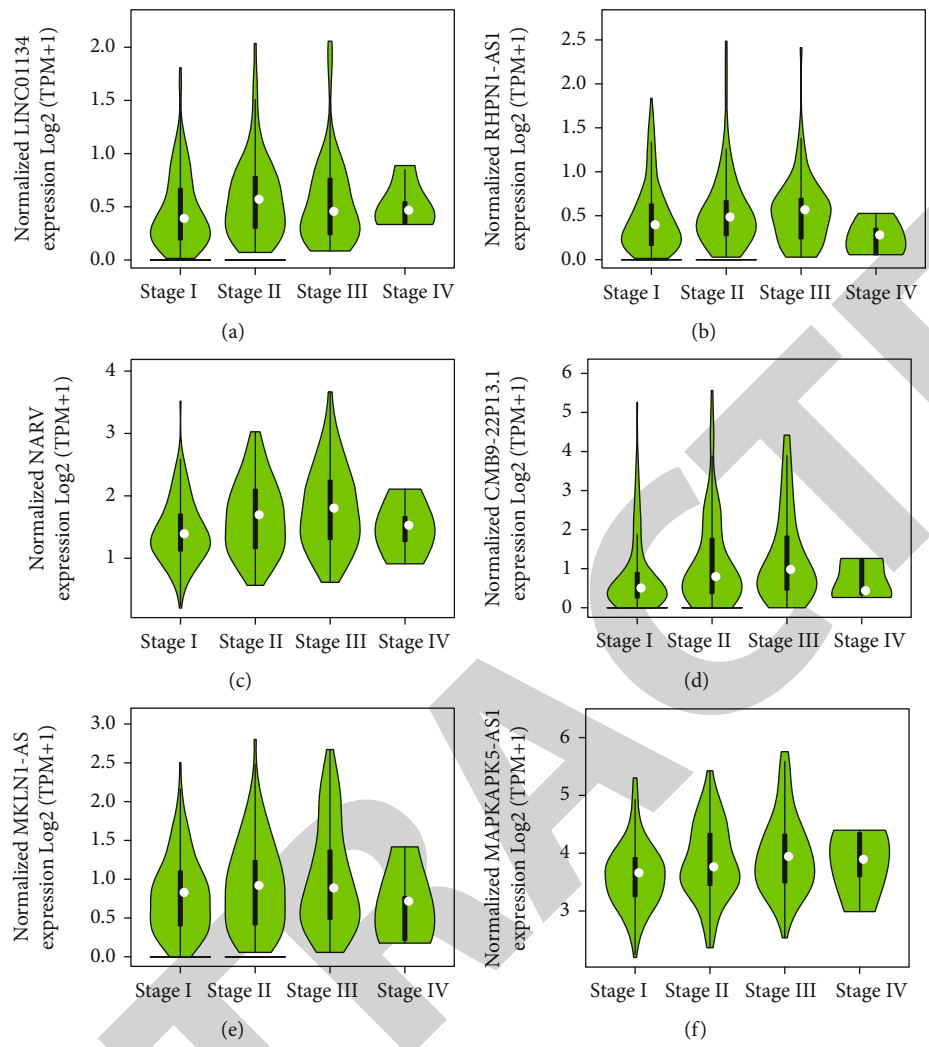
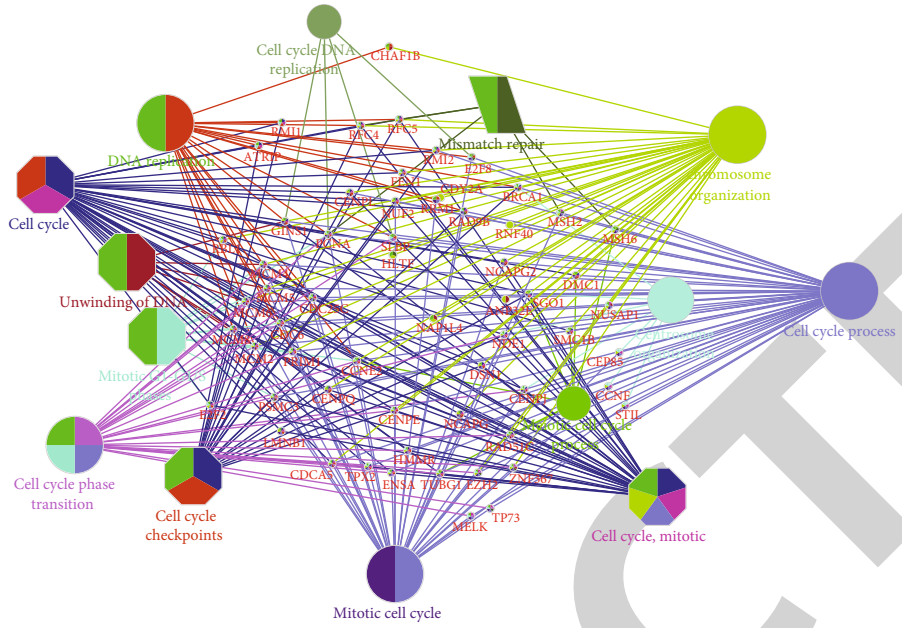


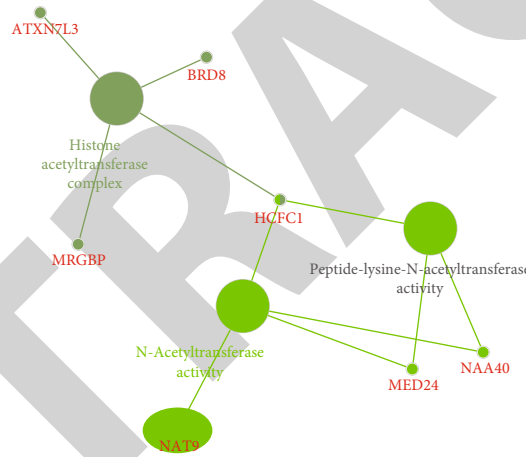
Figure 5: LINC01134, RHPN1-AS1, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1 were upregulated in advanced stage HCC. (a–f) LINC01134 (a), RHPN1-AS1 (b), NRAV (c), CMB9-22P13.1 (d), MKLN1-AS (e), and MAPKAPK5-AS1 (f) were upregulated in advanced stage HCC compared to low grade HCC.

regulation of acetyltransferase activity, mRNA binding, Golgi-to-ER retrograde transport, vasopressin-regulated water reabsorption, mitotic spindle organization, cilium organization, and translocation of GLUT4 to the plasma membrane (Figure 6(d)). RHPN1-AS1 was associated with DNA replication, snRNP assembly, mRNA metabolic process, mRNA splicing, and processing of capped intron-containing pre-mRNA (Figure 6(e)). MKLN1-AS was involved in the VEGFA-VEGFR2 pathway, endocytosis, Golgi-to-ER retrograde transport, RNA degradation, and regulation of RUNX2 expression and activity (Figure 6(e)). These results indicated that these lncRNAs might interact with genes involved in cancer-associated pathways listed above, thereby regulating the activities of these pathways.

3.5. The lncRNA-Mediated Competing Endogenous RNA Network in HCC. We constructed a series of competing endogenous RNA (ceRNA) networks in HCC using these six prognostic lncRNAs. The MKLN1-AS-mediated ceRNA network consisted of 14 miRNAs (including miR-761, miR-214-3p, miR-488-5p, miR-125b-5p, miR-370-3p, miR-22-3p, miR-125a-5p, miR-4782-3p, miR-3619-5p, miR-425-5p, miR-340-3p, miR-873-5p, miR-223-3p, and miR-510-5p) and 142 mRNAs (Figure 7(a)). MAPKAPK5-AS1-mediated ceRNA network consisted of 12 miRNAs (including miR-4306, miR-342-3p, miR-154-5p, miR-512-5p, miR-124-3p, miR-185-5p, miR-1271-5p, miR-4644, miR-182-5p, miR-96-5p, miR-362-5p, and miR-506-3p) and 55 mRNAs (Figure 7(b)). RHPN1-AS1-mediated ceRNA network consisted of 8 miRNAs (including miR-196a-5p, miR-342-3p, miR-345-5p, miR-377-3p, miR-339-5p, miR-196b-5p, miR-182-3p, and miR-486-5p) and 54 mRNAs (Figure 7(c)). Furthermore, LINC01134 was predicted to bind with miR-17-3p, miR-494-3p, miR-211-5p, miR-324-5p, miR-324-3p, miR-338-3p, miR-216b-5p, and miR-204-5p, thereby promoting the expression of 39 mRNAs (Figure 7(d)). The CMB9-22P13.1-mediated ceRNA network consisted of 2 miRNAs (including miR-299-3p and miR-522-3p) and 6 mRNAs (including MCM4, MELK, TP73, CKAP2L, TUB, and MBOAT1) (Figure 7(e)). These results indicated that

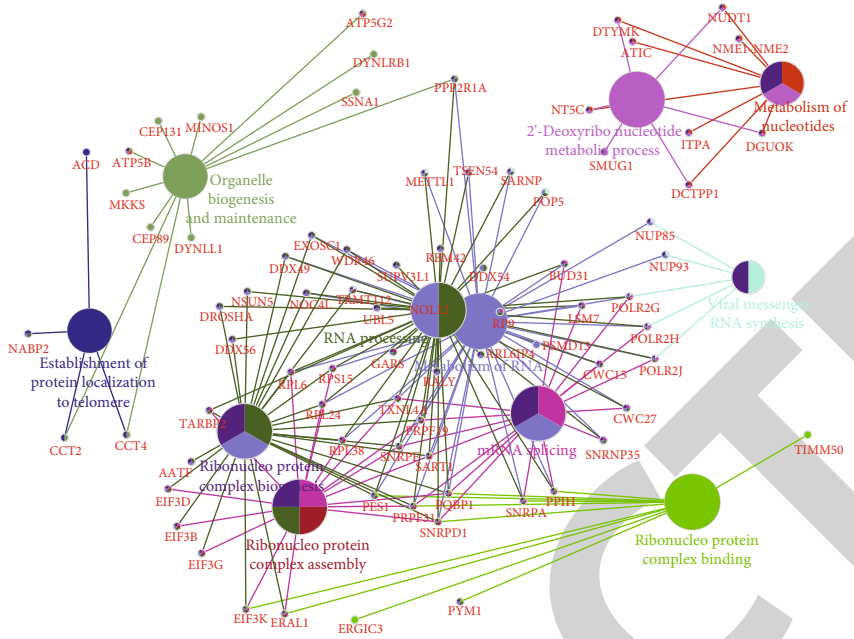


(a)

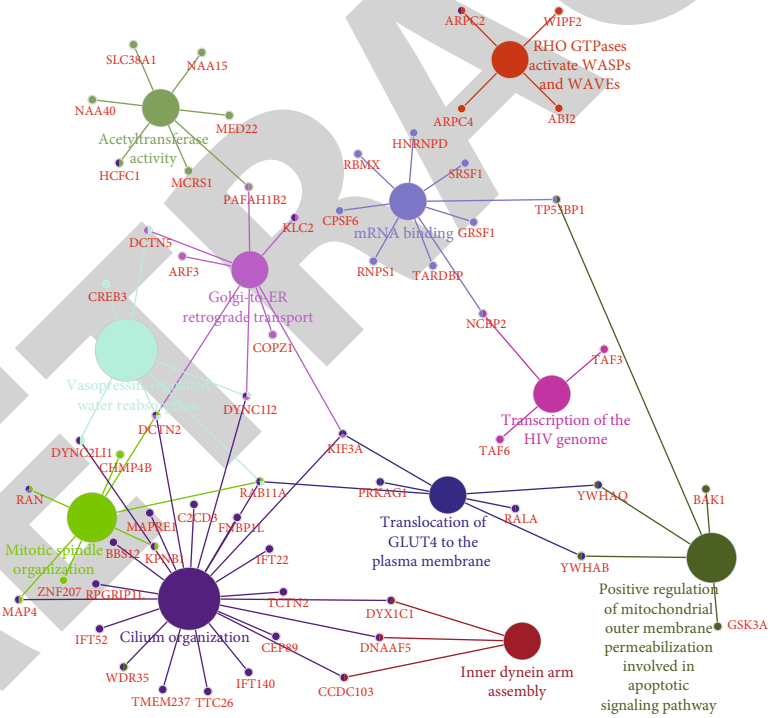


(b)

Figure 6: Continued.

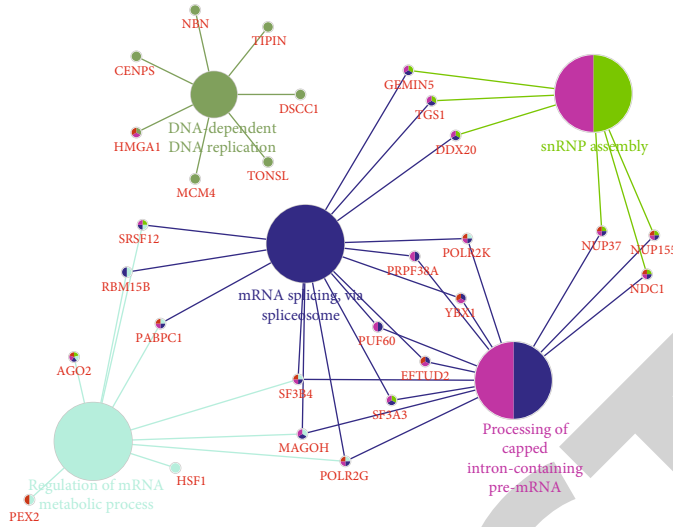


(c)

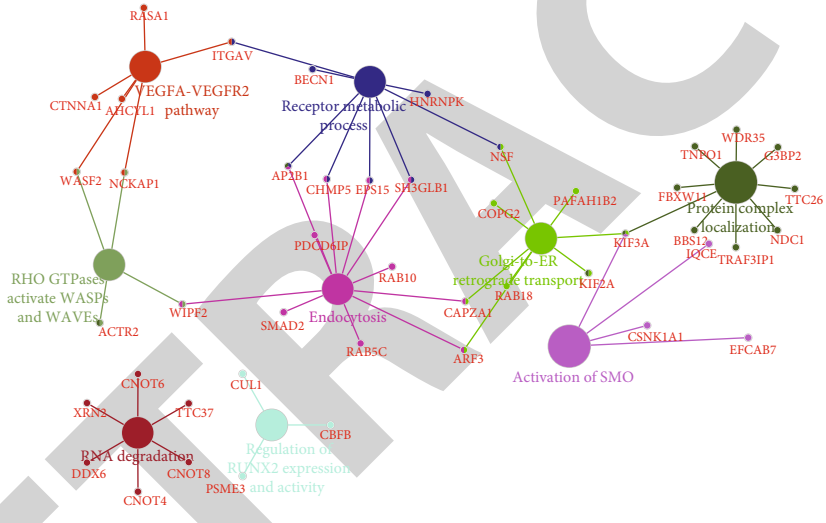


(d)

Figure 6: Continued.



(e)



(f)

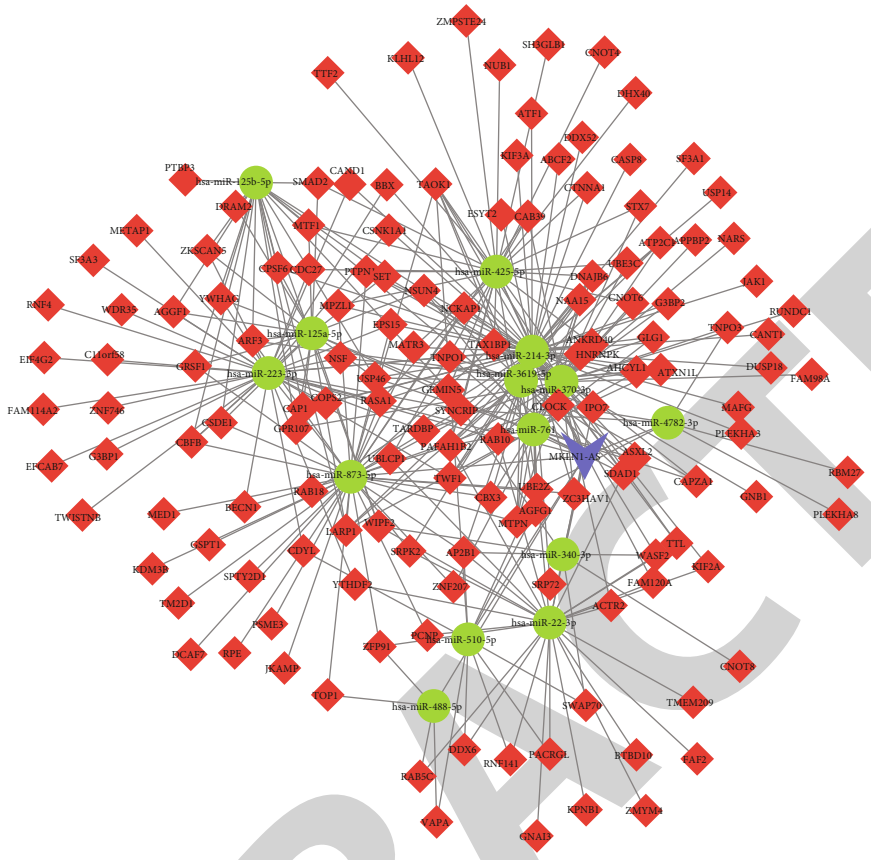
Figure 6: Enrichment analysis of differentially expressed lncRNAs in HCC. (a–f) We predicted the potential functions of CMB9-22P13.1 (a), LINC01134 (b), MAPKAPK5-AS1 (c), NARV (d), RHPN1-AS1 (e), and MKLN1-AS (f) in HCC.

these prognostic lncRNAs had the potential to sponge miRNAs, thereby regulating the expression levels of related protein-coding mRNAs.

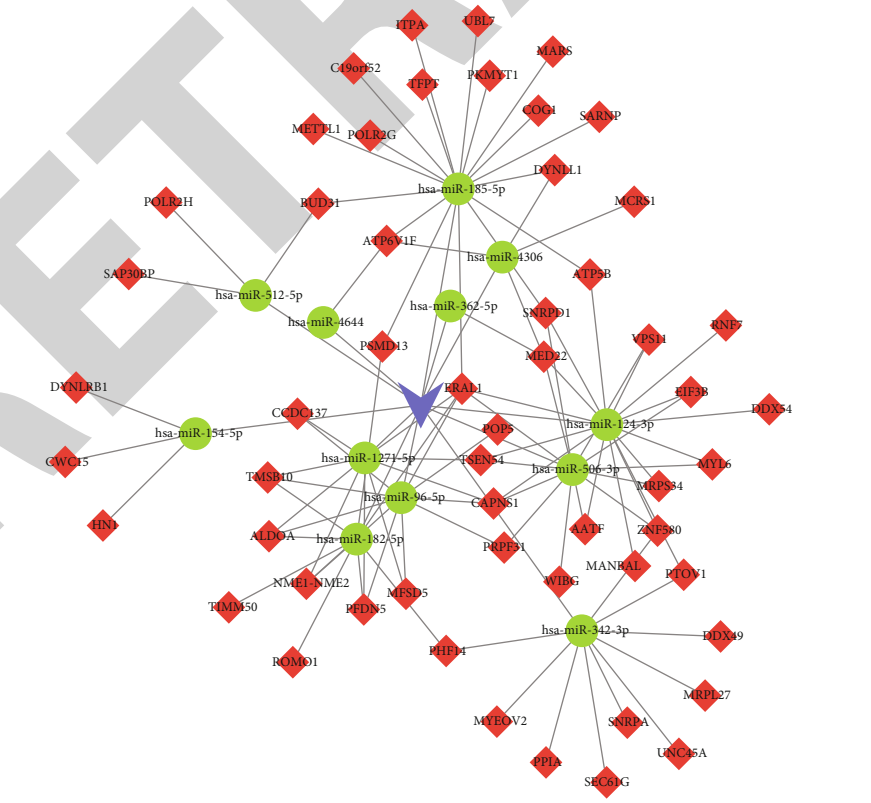
3.6. Inhibition of HCC Cell Proliferation and Migration In Vitro Caused by lncRNA RHPN1-AS1 Knockdown. As RHPN-AS1 achieved the most statistically significant association with both overall survival and disease-free survival, we further explored its role in HCC progression. Functional prediction revealed that this lncRNA was related to the regulation of DNA replication and mRNA splicing. Moreover, RHPN1-AS1 has been reported to act as an oncogene in multiple human cancers [19–21], such as nonsmall cell lung cancer, uveal melanoma, and breast cancer. However, the functional roles of this lncRNA in HCC remained largely unknown. The knockdown efficiency assay showed that the expression of RHPN1-AS1 was suppressed compared to the control group in HepG2 (Figure 8(a)) and Huh7 (Figure 8(c)) cells. Subsequently, the CCK-8 assay showed that the proliferation of both HepG2 (Figure 8(b)) and Huh7 (Figure 8(d)) cells was suppressed through RHPN1-AS1 knockdown. The transwell assay was also utilized to measure the migration ability of HCC cells. As illustrated in Figure 8, the number of migrating cells in the RHPN1-AS1 silencing group was remarkably reduced compared to that in the control groups of both HepG2 (Figures 8(e)-8(f)) and Huh7 (Figures 8(g)-8(h)) cells. These results indicated that RHPN1-AS1 silence could efficiently suppress HCC cell proliferation and migration.

4. Discussion

To date, a large number of long noncoding RNAs (lncRNAs) have been found to participate in carcinogenesis of several malignant tumors. To explore the functional roles of lncRNAs in HCC, we conducted a systematic analysis to



(a)



(b)

Figure 7: Continued.

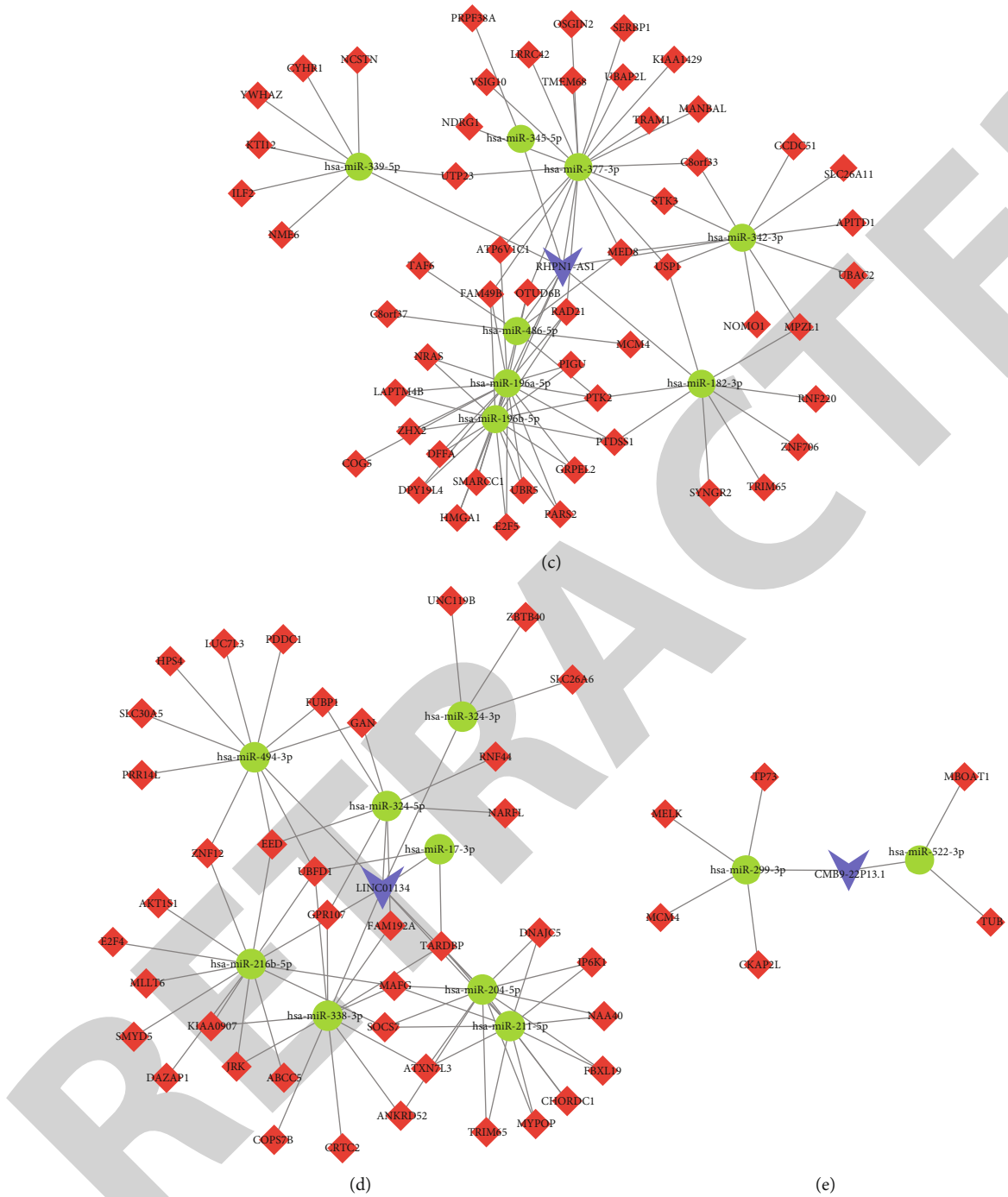


Figure 7: Construction of prognostic lncRNA-mediated competing endogenous RNA (ceRNA) network in HCC. The MKLN1-AS-, MAPKAPK5-AS1-, RHPN1-AS1-, LINC01134-, and CMB9-22P13.1-mediated ceRNA networks in HCC are displayed in (a–e). The red, green, and purple nodes represent the protein-coding mRNAs, miRNAs, and lncRNAs, respectively.

screen out prognostic lncRNAs. Specifically, we identified 6 upregulated lncRNAs in HCC samples, including LINC01134, RHPN1-AS1, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1. Higher expression levels of these lncRNAs were associated with advanced cancer stages and worse prognosis in HCC. Enrichment analysis revealed that these lncRNAs played a crucial role in HCC progression through affecting multiple cancer-related biological processes, including cell cycle, DNA replication, histone acetyltransferase complex, fatty acid oxidation, and lipid modification. Furthermore, ceRNA network analysis depicted that these lncRNAs could bind to a series of miRNAs to promote HCC progression.

lncRNAs have been reported to be key regulators in HCC. For example, Lnc-UCID enhanced cell cycle and growth [18]. HAND2-AS1 induced cancer stem cell self-

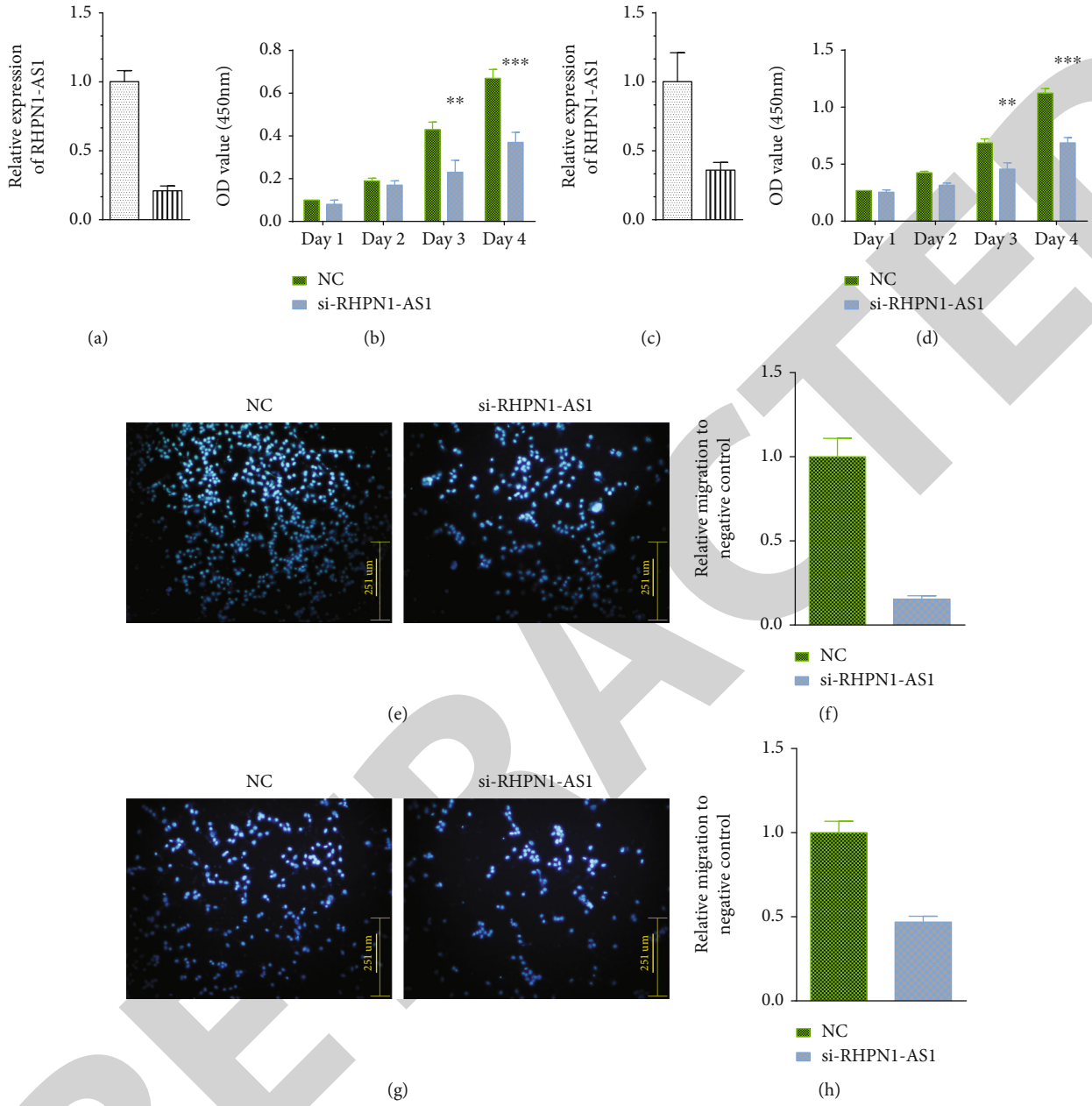


Figure 8: The proliferation and migration of HCC cells in vitro were inhibited by lncRNA RHPN1-AS1 knockdown. (a, c) The knockdown efficiency assay showed that the levels of RHPN1-AS1 were suppressed in HepG2 (a) and Huh7 cells (c) when compared to the siNC group. (b, d) RHPN1-AS1 knockdown could suppress both the HepG2 (b) and Huh7 (d) cell proliferation. (e, f) RHPN1-AS1 knockdown could suppress both the HepG2 cell migrations. (f–h) RHPN1-AS1 knockdown could suppress both the Huh7 cell migrations. The symbols of $*$, $**$, and $***$ indicate the P values < 0.05, 0.01, and 0.001, respectively.

renewal in HCC [22]. The abnormal expression of lncRNAs could help predict the progression and prognosis of HCC, and ncRNA CDKN2BAS is indicative of poor prognosis of HCC as it could lead to metastasis [23]. However, the functions of most lncRNAs in HCC remain unclear. This study for the first time showed that RHPN1-AS1, LINC01134, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1 were overexpressed in HCC, and a higher expression of these lncRNAs was observed in advanced-stage HCC samples as compared to early-stage HCC samples. Moreover, we demonstrated that higher expression levels of RHPN1-AS1, LINC01134, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1 were related to shorter OS and DFS time among HCC patients. Our study showed that these lncRNAs could serve as prognostic biomarkers for HCC.

Among these lncRNAs, RHPN1-AS1 was reported to act as an oncogene in multiple human cancers, including non-small cell lung cancer [19], uveal melanoma [20], and breast cancer [21]. LINC01134 and MKLN1-AS were found to promote HCC progression or metastasis by previous studies [24–28], suggesting that the two lncRNAs might be more specific in HCC, while NRAV and CMB9-22P13.1 were

rarely reported as an oncogene in cancers. Similar with RHPN1-AS1, MAPKAPK5-AS1 was frequently reported to act as tumor-promoting lncRNAs in multiple cancers like thyroid cancer, colorectal cancer, glioma, lung cancer, and HCC [29–33]. However, little was known when it comes to their roles in HCC. In this study, we conducted coexpression analysis and enrichment analysis of these lncRNAs in HCC. Very interestingly, we have found that these lncRNAs are important players in regulating HCC proliferation and metabolisms. LINC01134 was involved in regulating histone acetyltransferase complex, while the ceRNA network analysis showed that LINC01134 could bind to miR-17-3p and miR-494-3p, which had been key regulators in HCC progression. For example, miR-17 could promote hepatocarcinogenesis [34], and miR-494 could promote the development of HCC by targeting the SIRT3 and PI3K/AKT pathway [35]. RHPN1-AS1 was associated with the regulation of DNA replication and mRNA splicing, and miR-196a-5p was considered as a potential target of RHPN1-AS1. In HCC, microRNA-196a/-196b could affect the JAK/STAT pathway [36]. Meanwhile, CMB9-22P13.1 participated in regulating tumor-growth-related biological processes, including cell cycle and DNA replication. CMB9-22P13.1 also affected the activity of miR-299-3p and miR-522-3p. miR-522 has been reported to affect HCC proliferation by targeting DKK1 and SFRP2 [37]. MKLN1-AS was involved in regulating the VEGFA-VEGFR2 pathway, endocytosis, Golgi-to-ER retrograde transport, and RNA degradation. The ceRNA analysis showed that MKLN1-AS could bind to 14 miRNAs. Among them, miR-125b-5p could inhibit HCC cell growth and metastasis via TXNRD1, while miRNA-125a-5p could suppress HCC cell growth by targeting Bcl-2-like-2 protein. MAPKAPK5-AS1 was related to the regulation of metabolism of nucleotides. Furthermore, RHPN1-AS1 was reported to function as an oncogene in human cancers [38]. Previous studies have demonstrated that RHPN1-AS1 could sponge miRNAs to suppress their activities in cancer cells [38, 39]. Multiple miRNAs have been validated as the direct targets of RHPN1-AS1. For example, RHPN1-AS1 could act as a miR-596 sponge to induce LETM1 and promote ovarian cancer tumorigenesis and metastasis. In this study, RHPN1-AS1 was found overexpressed in HCC samples and correlated to worse prognosis of HCC. Knockdown of RHPN1-AS1 significantly suppressed HCC cell proliferation and migration, hinting that this lncRNA could act as an oncogenic lncRNA in HCC.

5. Conclusion

In conclusion, this study has demonstrated that LINC01134, RHPN1-AS1, NRAV, CMB9-22P13.1, MKLN1-AS, and MAPKAPK5-AS1 could serve as potential prognostic markers in HCC. We have conducted enrichment analyses and constructed ceRNA networks to provide functional implications of these prognostic lncRNAs in HCC. While some of these predictions still require further validations, loss-of-function assays have confirmed that silencing of RHPN1-AS1 significantly impaired HCC proliferative ability and migratory ability, and we believe that our study could cast light on the exploration of potential prognostic markers for HCC.

Data Availability

Previously reported gene expression and clinical data were used to support this study and are available at TCGA data portal (https://portal.gdc.cancer.gov/). These prior studies (and datasets) are cited at relevant places within the text as references.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

C. S. and B. Z. designed this study. Y.F., X.H., and K.M. conducted the data analysis, visualization, and experiments. Y.F. and C. S. contributed to the writing of the paper and setting of figures.

Supplementary Materials

Supplementary Table S1: the statistical significance of survival analysis for the three downregulated lncRNAs. Supplementary Table S2: the expression levels of six prognostic lncRNAs in the six pairs of HCC and nontumor tissues. *(Supplementary Materials)*

References

[1] Chinese Human Proteome Project (CNHPP) Consortium, Y. Jiang, A. Sun et al., "Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma," *Nature*, vol. 567, no. 7747, pp. 257–261, 2019.

[2] P. Rawla, T. Sunkara, P. Muralidharan, and J. P. Raj, "Update in global trends and aetiology of hepatocellular carcinoma," *Współczesna Onkologia*, vol. 22, no. 3, pp. 141–150, 2018.

[3] X. Yin, B. Tang, J. H. Li et al., "ID1 promotes hepatocellular carcinoma proliferation and confers chemoresistance to oxaliplatin by activating pentose phosphate pathway," *Journal of Experimental & Clinical Cancer Research*, vol. 36, no. 1, p. 166, 2017.

[4] C. Q. Zhong, X. P. Zhang, N. Ma et al., "FABP4 suppresses proliferation and invasion of hepatocellular carcinoma cells and predicts a poor prognosis for hepatocellular carcinoma," *Cancer Medicine*, vol. 7, no. 6, pp. 2629–2640, 2018.

[5] Y. Yang, L. Chen, J. Gu et al., "Recurrently deregulated lncRNAs in hepatocellular carcinoma," *Nature Communications*, vol. 8, no. 1, p. 14421, 2017.

[6] J. R. Prensner and A. M. Chinnaiyan, "The emergence of lncRNAs in cancer biology," *Cancer Discovery*, vol. 1, no. 5, pp. 391–407, 2011.

[7] W. G. Liu and Q. Xu, "Long non-coding RNA XIST promotes hepatocellular carcinoma progression by sponging miR-200b-3p," *European Review for Medical and Pharmacological Sciences*, vol. 23, no. 22, pp. 9857–9862, 2019.

[8] Q. Kong, S. Zhang, C. Liang et al., "LncRNA XIST functions as a molecular sponge of miR-194-5p to regulate MAPK1 expression in hepatocellular carcinoma cell," *Journal of Cellular Biochemistry*, vol. 119, no. 6, pp. 4458–4468, 2018.

[9] Y. Zhang, Z. Zhu, S. Huang et al., "lncRNA XIST regulates proliferation and migration of hepatocellular carcinoma cells

by acting as miR-497-5p molecular sponge and targeting PDCD4," *Cancer Cell International*, vol. 19, no. 1, p. 198, 2019.

[10] Z. Li, J. Zhang, X. Liu et al., "The LINC01138 drives malignancies via activating arginine methyltransferase 5 in hepatocellular carcinoma," *Nature Communications*, vol. 9, no. 1, p. 1572, 2018.

[11] S. X. Yuan, J. Wang, F. Yang et al., "Long noncoding RNA DANCR increases stemness features of hepatocellular carcinoma by derepression of CTNNB1," *Hepatology*, vol. 63, no. 2, pp. 499–511, 2016.

[12] Z. Tang, C. Li, B. Kang, G. Gao, C. Li, and Z. Zhang, "GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses," *Nucleic Acids Research*, vol. 45, no. W1, pp. W98–W102, 2017.

[13] J. H. Li, S. Liu, H. Zhou, L. H. Qu, and J. H. Yang, "starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data," *Nucleic Acids Research*, vol. 42, no. D1, pp. D92–D97, 2014.

[14] N. T. Doncheva, J. H. Morris, J. Gorodkin, and L. J. Jensen, "Cytoscape StringApp: network analysis and visualization of proteomics data," *Journal of Proteome Research*, vol. 18, no. 2, pp. 623–632, 2019.

[15] G. Bindea, B. Mlecnik, H. Hackl et al., "ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks," *Bioinformatics*, vol. 25, no. 8, pp. 1091–1093, 2009.

[16] F. Fu, X. Wan, D. Wang et al., "MicroRNA-19a acts as a prognostic marker and promotes prostate cancer progression via inhibiting VPS37A expression," *Oncotarget*, vol. 9, no. 2, pp. 1931–1943, 2018.

[17] P. Cui, J. Su, Q. Li, G. Xu, and N. Zhu, "LncRNA RHPN1-AS1 targeting miR-625/REG3A promotes cell proliferation and invasion of glioma cells," *Oncotargets and Therapy*, vol. - Volume 12, pp. 7911–7921, 2019.

[18] Y. L. Wang, J. Y. Liu, J. E. Yang et al., "Lnc-UCID promotes G1/S transition and hepatoma growth by preventing DHX9-mediated CDK6 down-regulation," *Hepatology*, vol. 70, no. 1, pp. 259–275, 2019.

[19] X. Li, X. Zhang, C. Yang, S. Cui, Q. Shen, and S. Xu, "The lncRNA RHPN1-AS1 downregulation promotes gefitinib resistance by targeting miR-299-3p/TNFSF12 pathway in NSCLC," *Cell Cycle*, vol. 17, no. 14, pp. 1772–1783, 2018.

[20] L. Lu, X. Yu, L. Zhang et al., "The long non-coding RNA RHPN1-AS1 promotes uveal melanoma progression," *International Journal of Molecular Sciences*, vol. 18, no. 1, p. 226, 2017.

[21] S. Zheng, P. Lv, J. Su, K. Miao, H. Xu, and M. Li, "Silencing of the long non-coding RNA RHPN1-AS1 suppresses the epithelial-to-mesenchymal transition and inhibits breast cancer progression," *American Journal of Translational Research*, vol. 11, no. 6, pp. 3505–3517, 2019.

[22] Y. Wang, P. Zhu, J. Luo et al., "LncRNA HAND2-AS1 promotes liver cancer stem cell self-renewal via BMP signaling," *The EMBO Journal*, vol. 38, no. 17, article e101110, 2019.

[23] J. Chen, X. Huang, W. Wang et al., "LncRNA CDKN2BAS predicts poor prognosis in patients with hepatocellular carcinoma and promotes metastasis via the miR-153-5p/ARHGAP18 signaling axis," *Aging (Albany NY)*, vol. 10, no. 11, pp. 3371–3381, 2018.

[24] X. Deng, Q. Bi, S. Chen et al., "Identification of a five-autophagy-related-lncRNA signature as a novel prognostic biomarker for hepatocellular carcinoma," *Frontiers in Molecular Biosciences*, vol. 7, p. 611626, 2021.

[25] W. Gao, X. Chen, W. Chi, and M. Xue, "Long non-coding RNA MKLN1-AS aggravates hepatocellular carcinoma progression by functioning as a molecular sponge for miR-654-3p, thereby promoting hepatoma-derived growth factor expression," *International Journal of Molecular Medicine*, vol. 46, no. 5, pp. 1743–1754, 2020.

[26] C. Wang, Y. Chen, K. Chen, and L. Zhang, "Long noncoding RNA LINC01134 promotes hepatocellular carcinoma metastasis via activating AKT1S1 and NF-κB signaling," *Frontiers in Cell and Development Biology*, vol. 8, p. 429, 2020.

[27] X. Zhao, Z. Bai, C. Li, C. Sheng, and H. Li, "Identification of a novel eight-lncRNA prognostic signature for HBV-HCC and analysis of their functions based on coexpression and ceRNA networks," *BioMed Research International*, vol. 2020, 14 pages, 2020.

[28] Z. Rong, Z. Wang, X. Wang, C. Qin, and W. Geng, "Molecular interplay between linc01134 and YY1 dictates hepatocellular carcinoma progression," *Journal of Experimental & Clinical Cancer Research*, vol. 39, no. 1, p. 61, 2020.

[29] L. Wang, L. Sun, R. Liu et al., "Long non-coding RNA MAPKAPK5-AS1/PLAGL2/HIF-1α signaling loop promotes hepatocellular carcinoma progression," *Journal of Experimental & Clinical Cancer Research*, vol. 40, no. 1, p. 72, 2021.

[30] Y. Zhou, S. Liu, Y. Luo, M. Zhang, X. Jiang, and Y. Xiong, "lncRNA MAPKAPK5-AS1 promotes proliferation and migration of thyroid cancer cell lines by targeting miR-519e-5p/YWHAH," *European Journal of Histochemistry*, vol. 64, no. 4, 2020.

[31] H. Zhang, Y. Wang, and J. Lu, "Identification of lung-adenocarcinoma-related long non-coding RNAs by random walking on a competing endogenous RNA network," *Ann Transl Med*, vol. 7, no. 14, p. 339, 2019.

[32] F. Luan, W. Chen, M. Chen et al., "An autophagy-related long non-coding RNA signature for glioma," *FEBS Open Bio*, vol. 9, no. 4, pp. 653–667, 2019.

[33] H. Ji, B. Hui, J. Wang et al., "Long noncoding RNA MAPKAPK5-AS1 promotes colorectal cancer proliferation by partly silencing p21 expression," *Cancer Science*, vol. 110, no. 1, pp. 72–85, 2019.

[34] H. Zhu, C. Han, and T. Wu, "MiR-17-92 cluster promotes hepatocarcinogenesis," *Carcinogenesis*, vol. 36, no. 10, pp. 1213–1222, 2015.

[35] J. Zhang, Y. Zhu, L. Hu, F. Yan, and J. Chen, "miR-494 induces EndMT and promotes the development of HCC (hepatocellular carcinoma) by targeting SIRT3/TGF-β/SMAD signaling pathway," *Scientific Reports*, vol. 9, no. 1, p. 7213, 2019.

[36] W. Ren, S. Wu, Y. Wu, T. Liu, X. Zhao, and Y. Li, "MicroRNA-196a/-196b regulate the progression of hepatocellular carcinoma through modulating the JAK/STAT pathway via targeting SOCS2," *Cell Death & Disease*, vol. 10, no. 5, p. 333, 2019.

[37] H. Zhang, C. Yu, M. Chen et al., "miR-522 contributes to cell proliferation of hepatocellular carcinoma by targeting DKK1 and SFRP2," *Tumour Biology*, vol. 37, no. 8, pp. 11321–11329, 2016.

Hindawi
BioMed Research International
Volume 2023, Article ID 9803081, 1 page
https://doi.org/10.1155/2023/9803081

Retraction

Retracted: Cannabinoid WIN 55,212-2 Inhibits Human Glioma Cell Growth by Triggering ROS-Mediated Signal Pathways

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] K. Wang, Q. Wang, Q. Li et al., "Cannabinoid WIN 55,212-2 Inhibits Human Glioma Cell Growth by Triggering ROS-Mediated Signal Pathways," *BioMed Research International*, vol. 2021, Article ID 6612592, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6612592, 11 pages
https://doi.org/10.1155/2021/6612592

Research Article

Cannabinoid WIN 55,212-2 Inhibits Human Glioma Cell Growth by Triggering ROS-Mediated Signal Pathways

Kun Wang,[1,2] Qian Wang,[1] Qinghao Li,[1] Zhaoqiang Zhang,[3] Jing Gao,[1] Cundong Fan,[4] Baoliang Sun,[4] and Qingbin Ni[1]

[1]*Postdoctoral Workstation, Taian City Central Hospital, Tai'an, Shandong 271000, China*
[2]*School of Medicine, Shandong University, Jinan, Shandong 270021, China*
[3]*Basic Medical College, Shandong First Medical University & Shandong Academy of Medical Sciences, Tai'an, Shandong 271000, China*
[4]*Key Lab of Cerebral Microcirculation in Universities of Shandong, Shandong First Medical University & Shandong Academy of Medical Sciences, Tai'an, Shandong 271000, China*

Correspondence should be addressed to Qingbin Ni; nqbin130@163.com

Received 23 December 2020; Revised 25 March 2021; Accepted 1 April 2021; Published 23 April 2021

Academic Editor: Min Tang

Glioblastoma is a highly invasive primary malignant tumor of the central nervous system. Cannabinoid analogue WIN 55,212-2 (WIN) exhibited a novel anticancer effect against human tumors. However, the anticancer potential and underlying mechanism of WIN against human glioma remain unclear. Herein, the anticancer efficiency and mechanism of WIN in U251 human glioma cells were investigated. The results showed that WIN dose-dependently inhibited U251 cell proliferation, migration, and invasion *in vitro*. WIN treatment also effectively suppressed U251 tumor spheroids growth *ex vivo*. Further studies found that WIN induced significant apoptosis as convinced by the caspase-3 activation and release of cytochrome C. Mechanism investigation revealed that WIN triggered ROS-mediated DNA damage and caused dysfunction of VEGF-AKT/FAK signal axis. However, ROS inhibition effectively attenuated WIN-induced DNA damage and dysfunction of VEGF-AKT/FAK signal axis and eventually improved U251 cell proliferation, migration, and invasion. Taken together, our findings validated that WIN had the potential to inhibit U251 cell proliferation, migration, and invasion and induce apoptosis by triggering ROS-dependent DNA damage and dysfunction of VEGF-AKT/FAK signal axis.

1. Introduction

Glioblastoma (GBM), originating from neuroepithelial tissue, represents the most common primary malignant tumor of the central nervous system, which shows high morbidity, high recurrence rate, and high mortality rate [1]. Although a series of advances have been made in the treatment of GBM in recent years, it is difficult to completely remove it due to its invasive growth mode and unclear boundary with the surrounding normal brain tissue. Therefore, GBM therapy is still a difficult problem, and it is necessary to find a better treatment of GBM.

The cannabis system is made up of cannabinoids and their receptors. It has been found that cannabinoids have analgesic, sedative, antispasmodic, antivomiting, reducing intraocular pressure, and other pharmacological effects [2]. There are three main types of cannabinoids, including endocannabinoids (such as 2-AG), natural cannabinoids (such as tetrahydrocannabinol (THC)), and synthetic cannabinoids (such as Win55212-2 (WIN)) [3]. Existing studies have shown that cannabinoids have an inhibitory effect on the occurrence and development of colorectal cancer [4], prostate cancer [5], non-small cell lung cancer [6], and breast cancer [7], all of which have good therapeutic effects. Among which, WIN can inhibit the proliferation and induce apoptosis of various tumor cells and inhibit the growth of tumors *in vivo*. However, the anticancer potential and mechanism of WIN against human GBM growth have not been reported yet.

ROS is an important factor in inducing apoptosis which can mediate DNA damage [8]. The VEGF-AKT/FAK signal pathway is an intracellular signaling axis which is important in regulating cell proliferation. Therefore, it is directly related to cancer invasion and longevity. Dysfunction of VEGF-AKT/FAK signaling pathway is often found in tumor cells, which are emerging as new potential therapeutic targets [9].

This study explored the anticancer effects of WIN on the proliferation, migration, and invasion of U251 cells and preliminarily studied its mechanism. The results indicated WIN had the potential to inhibit U251 cell proliferation, migration, and invasion and induce apoptosis by triggering ROS-dependent DNA damage and dysfunction of VEGF-AKT/-FAK signal axis.

2. Materials and Methods

2.1. Chemicals. DMEM medium, fetal bovine serum, and penicillin streptomycin were purchased from Invitrogen. WIN, MTT, DMSO, and DCFH-DA probes were purchased from Sigma. Annexin V-FITC Cell Apoptosis Assay Kit was purchased from Roche. The BCA Detection Kit was purchased from Biyuntian Biotechnology Research Institute, China. All antibodies were purchased from CST Inc. The Transwell Compartment System was purchased from Corning Costar (No.3422) and Matrigel from BD Inc.

2.2. Cell Culture and Drug Treatment. U251 human glioma cells were obtained from Chinese Academy of Sciences Type Culture Collection, and U251 cells were cultured in DMEM high-sugar medium containing 10% fetal bovine serum, and penicillin G sodium and streptomycin sulfate 100 μg/ml each. The cells were routinely cultured in an incubator (37°C, 5% CO_2, and 95% humidity).

2.3. Measurement of Cell Viability. Logarithmic proliferative U251 cells were collected and inoculated in a 96-well plate (5000 cells/well) for 24 h. After adhering to the wall, the cells were treated with 0-40 μM WIN for 48 h or with 20 μM WIN for 0-48 h. Then, 20 μl 5 mg/ml MTT solution was added, and the supernatant was discarded after 6 h incubation. Then, 150 μl DMSO was added to each well, and the light absorption value at 570 nm was measured. 3 duplicate wells were set for each test. Cell growth was reflected by the percentage of normal cell activity (%). Phase contrast microscope (Nikon Eclipse 80i) was used to observe the morphological changes of U251.

2.4. Cell Migration. U251 cells in logarithmic phase seeded in 6-well plate were cultured with DMEM medium supplemented with 10% FBS for 24 h. Cells after full monolayer were gently washed by medium for 3 times, and a 200 μl pipette head was used to scratch the monolayer of cells in a straight line. Then, cells were cultured with DMEM medium supplemented with 1% FBS and treated with 5, 10, and 20 μM WIN for 48 h. Cell migration was imaged at 0 h and 48 h. The migrated distance was measured, and the migrated rate was calculated. All experiment was repeated for 3 times.

2.5. Cell Invasion. Matrigel-procoated Transwell cell invasion assay was carried out to examine cell invasion. Firstly, the Transwell chamber was precoated with Matrigel diluent at 4°C and placed at 37°C for 3 h. Then, the chamber was added with 100 μl serum-free U251 cell suspension, and 600 μl DMEM medium containing 10% FBS was added below the chamber. After 24 h of culture, the chamber was taken out, the medium was discarded, and the cells in the upper chamber were erased with a cotton swab. The cells were gently washed and fixed with 4% paraformaldehyde for 20 min. After staining with crystal violet, cell invasion was detected by the microscope. All experiments were repeated three times.

2.6. Tumor Spheroids. U251 cells in logarithmic growth phase were inoculated into an ultralow adherent 6-well plate. Cells were gently shaken to disperse into a single cell. After 2 days of culture, one group was treated with 0-20 μM WIN for 48 h, followed by 0.01% final concentration (W/V) Cy3 for 6 h. The other group was treated with 20 μM WIN for 0-120 min, with 3 replicates. The tumor sphere volume and ROS production were measured by taking photos under a fluorescence microscope.

2.7. Detection of Cell Apoptosis. U251 cells were inoculated into a 6-well plate and treated with 20 μM WIN for 0-12 hours and Annexin V probe for 20 minutes in darkness. Then, cells were rinsed, and the degree of apoptosis was observed with a fluorescence microscope.

2.8. Examination of ROS Generation. U251 cells were inoculated into a 6-well plate and treated with 20 μM WIN for 0-120 min after adhering to the wall. Cells were labeled by DCFH-DA probe for 20 min in darkness; then, cells were rinsed, and ROS production was observed with a fluorescence microscope.

2.9. Western Blotting. U251 cells were inoculated in a 10 cm culture dish and treated with 20 μM WIN for 0-48 h. U251 cells in each group were collected and lysed, and the concentration of the protein samples was determined by the BCA protein kit. Next, the SDS-PAGE gel electrophoresis was used to separate the proteins, and then, the proteins were electrotransfered to the PVDF membrane. Subsequently, the proteins were sealed with TBST containing 5% skimmed milk powder at room temperature for 1 h, followed by a drop of primary antibody, incubated at 4°C overnight, and washed with 0.1% TBST for 10 min. Next, the membrane was incubated with secondary antibody at room temperature for 1 h and washed with TBST for 10 min. The ECL chemiluminescence method and gel imager were used for imaging.

2.10. Statistical Analysis. Each experiment was performed in triplicate and repeated for more than 3 times. Data were presented as the mean ± standard deviation ($x \pm s$). Statistical analysis was performed using the SPSS statistical package (SPSS 13.0 for Windows; SPSS, Inc., Chicago, IL, USA). The difference between two groups was analyzed by a two-tailed Student's t test. The difference among three or more groups was analyzed by one-way analysis of variance

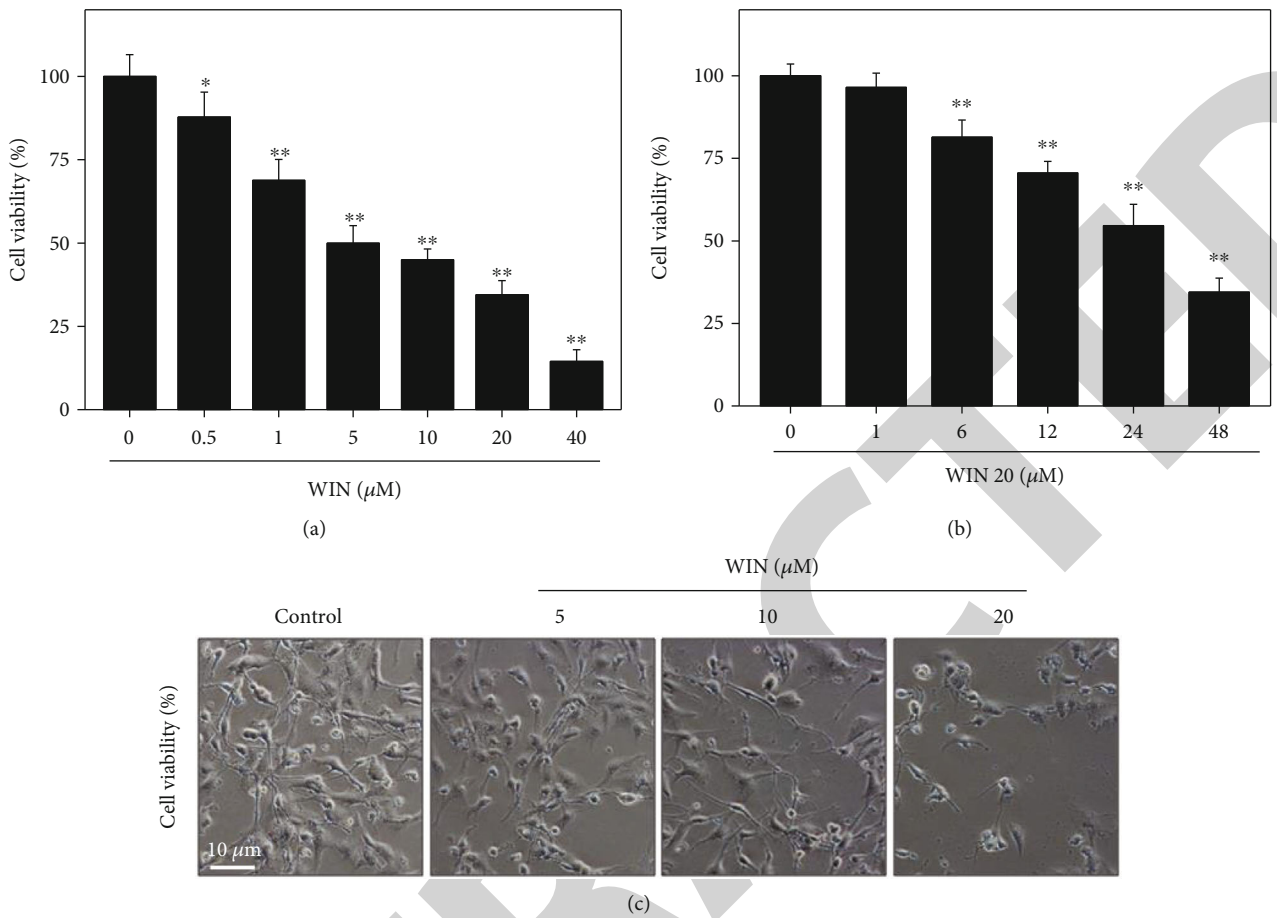


Figure 1: WIN inhibits U251 cell viability. WIN dose-dependently (a) and time-dependently (b) inhibited U251 cell viability. U251 cells seeded in 96-well plate were treated with 0-40 μM WIN for 48 h or cells were treated with 20 μM WIN for 0-48 h. Cell viability was detected by MTT assay. (c) Morphological changes. Cells after treatment were detected by microscope (magnification, $\times 200$). All experiments were repeated three times. Bar with "$*$" or "$**$" indicates the significant difference at $P < 0.05$ or $P < 0.01$ level, respectively.

(ANOVA) with multiple comparisons, and the bars with different letters mean statistically significant differences ($P < 0.05$). Bars with "$*$" or "$**$" represent statistically different at $P < 0.05$ or $P < 0.01$ level, respectively.

3. Results

3.1. WIN Inhibits GBM Cell Growth. Firstly, the MTT assay was used to detect the effect of WIN on U251 cell viability. As shown in Figure 1(a), U251 cells exposed to WIN (0-40 μM) for 48 h showed a dose-dependent decrease of cell viability. For example, treatment of U251 cells with 10 and 40 μM WIN significantly decreased the cell viability to 46.2% and 14.9%, respectively. The half inhibition rate concentration (IC_{50}) of WIN is 6.7 μM. Moreover, WIN treatment time-dependently inhibited U251 cell growth, which further confirmed WIN-induced growth inhibition against U251 cells (Figure 1(b)). Additionally, morphological changes of U251 cells by phase contrast were also detected. As shown in Figure 1(c), U251 cells treated with WIN showed dose-dependent cell shrinkage, reduction in attaching cell numbers, and loss of cell-to-cell contact, which further confirmed WIN-induced anticancer effect. Taken together, these results indicated that WIN could inhibit human glioma cell growth *in vitro*.

3.2. WIN Blocks the Migration and Invasion of U251 Cells. Cell scratch test and Transwell test were used to observe the changes of migration and invasion ability of U251 cells before and after WIN interference. As shown in Figures 2(a) and 2(c), U251 cells without WIN treatment displayed obvious cell migration and invasion. However, the migration and invasion ability of U251 cells treated with WIN were both significantly inhibited with a dose-dependent manner. The quantitative analysis results of cell migration and invasion further verified the effects (Figures 2(b) and 2(d)). These results indicated that WIN could block migration and invasion of U251 cells.

3.3. WIN Suppresses the Growth of U251 Tumor Spheroids Ex Vivo. To further confirm WIN-induced growth inhibition against human glioma cells, a tumor spheroid model was conducted to evaluate its anticancer efficiency *ex vivo*. As shown in Figure 3(a), treatment of tumor spheroids with WIN dose-dependently inhibited the growth of tumor spheroids, as convinced by the decreased volume of tumors

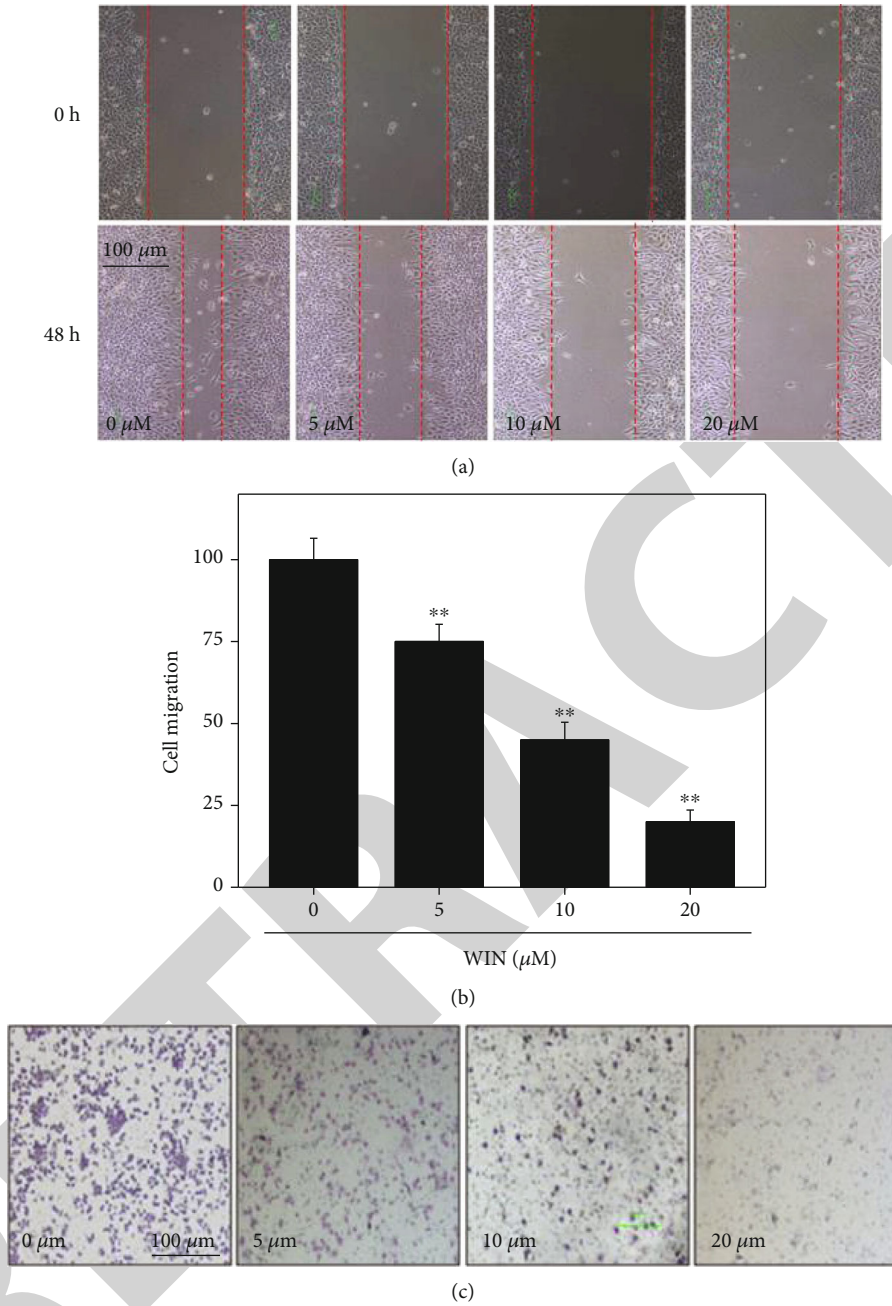


Figure 2: Continued.

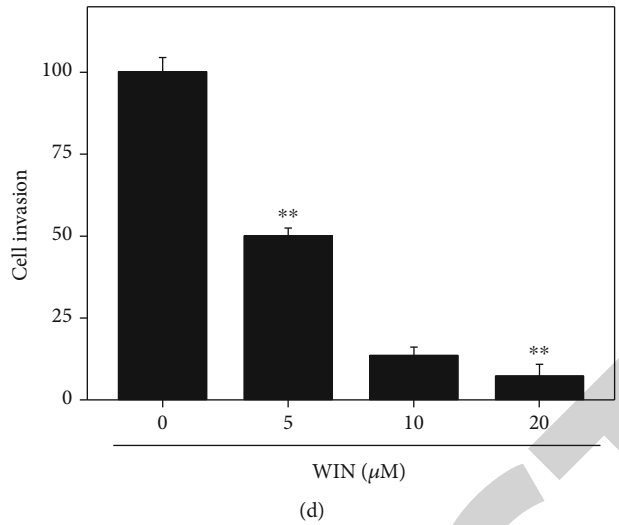


Figure 2: WIN inhibits migration and invasion of U251 cells. WIN dose-dependently inhibited U251 cell migration (a) and invasion (c). Quantitative analysis of cell migration (b) and invasion (d) of U251 cells. Cell migration was conducted in a 6-well plate by a wound-healing assay and observed at 0 h and 48 h. Cell invasion was conducted by Matrigel-procoated Transwell. Quantitative analysis of cell migration and invasion was conducted by manual calculating (% of control). All experiments were repeated three times. Bar with "**" indicates the significant difference at $P < 0.01$ level.

spheroids. The quantitative analysis of tumor spheroids volume further confirmed this effect (Figure 3(c)). Moreover, a time-dependent ROS production was also detected in tumor spheroids, as demonstrated by the increased green fluorescence (Figure 3(b)). The quantitative analysis of DCF fluorescence in tumor spheroids further convinced this effect (Figure 3(d)). Taken together, these results suggested that WIN could suppress U251 tumor spheroids growth *ex vivo* involving ROS production.

3.4. WIN Induces Apoptosis of U251 Cells. To investigate WIN-induced cell death mechanism of U251 cells, the Annexin V probe was used to detect early apoptosis of U251 cells induced by WIN. As shown in Figure 4(a), U251 cells after WIN treatment showed significant increase of green fluorescence with a time-dependent manner, indicating that WIN treatment induced early cell apoptosis in U251 cells. The quantitative analysis of Annexin V-positive cells further confirmed this conclusion (Figure 4(b)). Moreover, western blotting was also employed to explore the apoptotic mechanism induced by WIN. As shown in Figure 4(c), WIN treatment time-dependently triggered caspase-3 activation and cytochrome c release, which further confirmed WIN-induced apoptosis of U251 cells in protein level. Taken together, these results indicated that WIN could inhibit U251 cell growth by induction of apoptosis.

3.5. WIN Triggers ROS-Mediated DNA Damage. To further explore the anticancer mechanism induced by WIN, a ROS-detecting probe (DCFH-DA) was used to determine the intracellular ROS accumulation. As shown in Figure 5(a), intracellular ROS levels in U251 cells showed a time-dependent increase, as reflected by the increased green fluorescence. Quantitative analysis of ROS level further confirmed WIN-induced ROS generation (Figure 5(b)). Moreover, GSH level in WIN-treated U251 cells showed a dose-dependent decrease (Figure 5(c)). Excessive ROS can attack DNA and cause DNA damage. Herein, DNA-damaging axis was examined by the western blotting method. As shown in Figure 5(d), WIN treatment significantly increased the phosphorylation of ATR (Ser428), p53 (Ser15), and histone (Ser139), suggesting that WIN treatment caused DNA damage in U251 cells. However, ROS inhibition effectively attenuated WIN-induced DNA damage. As shown in Figure 5(e), pretreatment of cells with glutathione (GSH) significantly attenuated WIN-induced phosphorylation of ATR (Ser428), p53 (Ser15), and histone (Ser139). We speculated that ROS may act as an early event in WIN-induced growth inhibition against U251 cells. Taken together, these results suggested that WIN could trigger ROS-mediated DNA damage in U251 cells.

3.6. WIN Disturbs VEGF-AKT/FAK Signaling Axis. VEGF-AKT/FAK signal axis plays a key role in regulating cell proliferation, survival, migration, and apoptosis. Herein, the status of VEGF-AKT/FAK signal axis in WIN-treated U251 cells was examined by western blotting. As shown in Figure 6(a), WIN treatment significantly inhibited VEGF, Ser473-AKT, and Tyr397-FAK expression with a dose-dependent manner. No significant change of VEGFR2 expression was observed. However, ROS inhibition effectively improved expression of VEGF, Ser473-AKT, and Tyr397-FAK (Figure 6(b)). These results indicated that WIN inhibited U251 cell proliferation, migration, and invasion with a ROS-dependent manner, and ROS as an early apoptotic event was involved in WIN-induced anticancer mechanism in U251 cells.

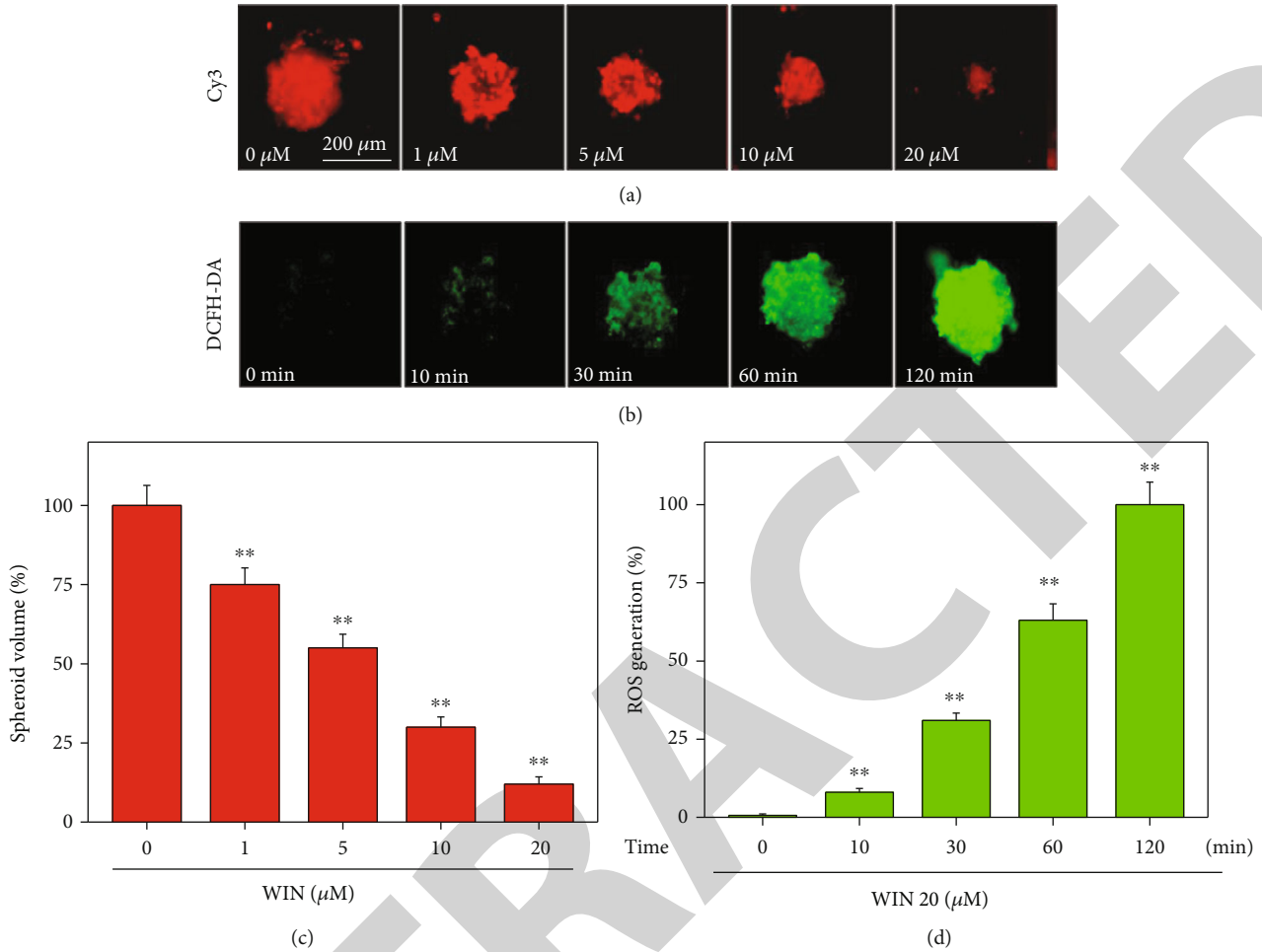


Figure 3: WIN inhibits tumor spheroids growth of U251 cells *ex vivo*. (a) WIN dose-dependently inhibited tumor spheroids growth of U251 cells *ex vivo*. U251 cells were cultured in an ultralow attachment 6-well plate for 2 days. Then, cells were treated with 0-20 μM WIN for 48 h. (b) Win time-dependently caused ROS production in tumor spheroids of U251. Tumor spheroids of U251 were prelabeled by 10 μM DCFH-DA for 15 min, and then tumor spheroids of U251 were treated with 20 μM WIN for 0-120 min. ROS production was monitored by fluorescence microscope (magnification, $\times 100$). Quantitative analysis of tumor spheroids volume (c) and ROS generation (d) were measured, respectively. All experiments were repeated three times. Bar with "**" indicates the significant difference at $P < 0.01$ level.

4. Discussion

Chemotherapy plays an irreplaceable role in therapy of human GBM. However, design of new anti-GBM drugs has become an urgent task due to the poor cure rate and high fatality rate. Previous studies have confirmed that, as a new approach, WIN provides a new idea for the treatment of GBM. Accumulated studies have confirmed that cannabinoids showed potential anticancer effects against human several cancers [4–7]. However, little information about WIN-mediated anticancer study in human glioma is available, and the underlying mechanism remains unclear. This study firstly explored WIN-induced anticancer activity and mechanism in human glioma, especially the role of ROS-mediated oxidative damage.

Targeted systemic treatment of tumors with cannabis has been a hot topic in recent years [10]. It has been found that the cannabis system can regulate the proliferation and apoptosis of tumor cells by affecting the signaling pathway [11–15]. With the in-depth study of the mechanism of action, the endocannabinoid system is likely to become a new target for tumor therapy, which has important clinical significance for improving the chemotherapy effect of glioma. Currently, cannabinoid receptor agonists such as Nabilone, Marinol, and the cannabinoid extract THC and CBD (Sativex) have been approved as adjuvant treatments for tumor chemotherapy [16]. However, natural cannabinoids are limited as antitumor drugs due to their certain tolerance and addictiveness.

Sildenafil mesylate (Win55212-2 (WIN)) is a highly specific synthetic cannabinoid receptor agonist [17], which has overcome the above shortcomings. It can exert strong hemp activity after binding to the receptor and is one of the representatives with promising application prospects [18]. It was found that WIN could relieve breast cancer [7], prostate cancer [15], and gastric cancer [19] in mice and prolong their survival by 50%. *In vitro* experiments showed that WIN

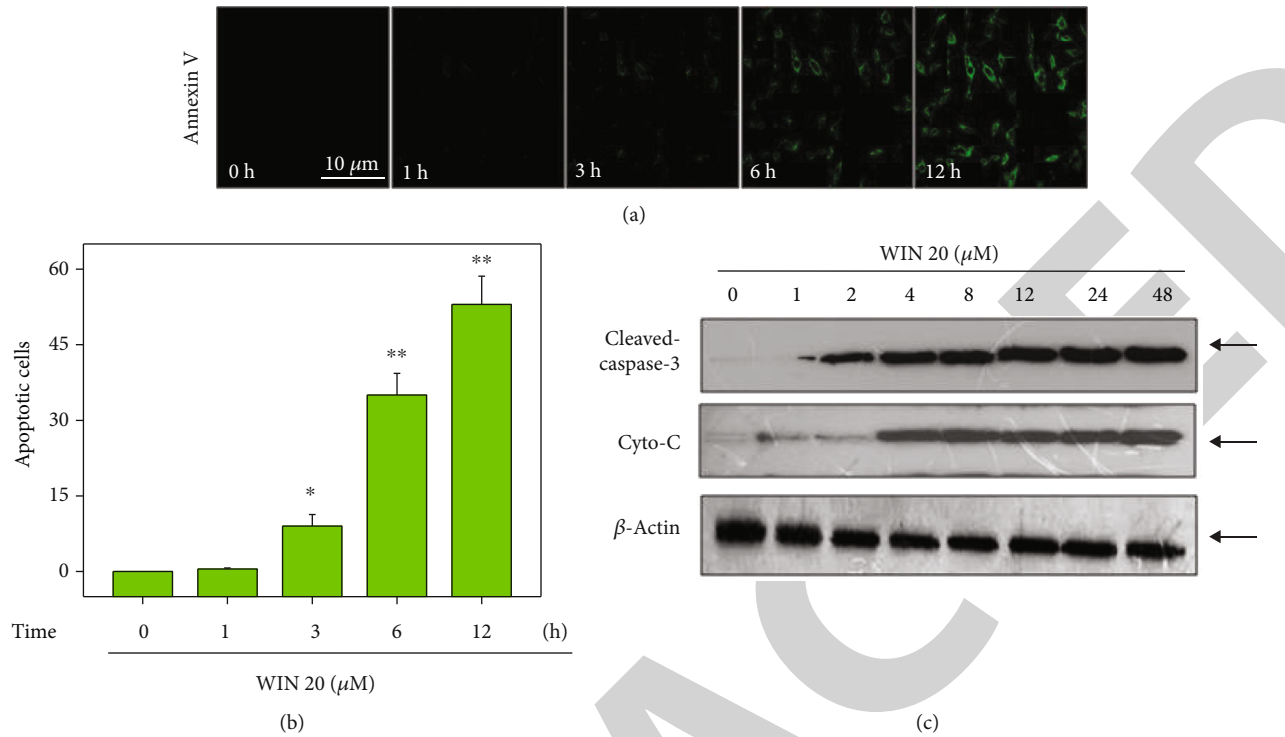


FIGURE 4: WIN induces apoptosis of U251 cells. (a) WIN time-dependently induces apoptosis of U251 cells. U251 cells seeded in 6-well plate were treated with 20 μM WIN for 0-12 h and loaded with Annexin V probe for 20 min in darkness. Then, cells were washed and the cell apoptosis were imaged by a fluorescence microscope (magnification, ×200). (b) Quantitative analysis of cell apoptosis. (c) WIN induced caspase-3 activation and Cyto-C release. Protein expression was detected by western blotting assay. All experiments were repeated three times. Bar with "*" or "**" indicates the significant difference at $P < 0.05$ or $P < 0.01$ level, respectively.

could induce death of prostate cancer cells, liver cancer cells, and other tumor cells [20, 21]. WIN inhibited DNA synthesis in T3 lymphocytes of human blastocysts and rodents, as well as in the Lewis lung cancer cells [22].

The ability of tumor cell migration and invasion is an important basis of tumor invasion and metastasis. GBM cells will continuously invade the surrounding brain tissue, and its invasion characteristics directly lead to low total surgical resection rate and high postoperative recurrence rate, which is directly related to poor prognosis of patients. Cannabinoids have been reported to inhibit both tumor cell migration and normal lymphocyte migration [23, 24]. Highly selective excitation of CBl/2 can inhibit tumor cell migration and metastasis without adverse effects on the immune system of tumor patients [25]. In this experiment, we first carried out cell scratch experiment and found that cell migration was significantly inhibited after WIN treatment of U251 cells. The Transwell cell invasion *in vitro* model is considered the most ideal determination of tumor cells *in vitro* invasion ability of experimental model at present [26]. In this study, we used Transwell assay to detect the effect of WIN on the invasion ability of U251 cells, and the results showed that WIN significantly inhibited the invasion ability of U251 cells. Compared with the single plane cultured tumor cells, 3D tumor ball can better simulate the signal transduction microenvironment between tumor cells and extracellular matrix; being one of the best *in vitro* models for cancer research, it is widely used in the study of tumor cell morphology, enrichment of cancer stem cells, cancer drug high-throughput screening, etc. [27]. This experiment showed that WIN inhibited the growth of U251 cells *in vitro* dose-dependently and promoted the production of ROS over time.

Apoptosis is considered the main cytotoxic mechanism by which antitumor drugs exert anticancer effects which is a complex process involving many factors and molecules. Mitochondria are the control center of cellular life activities, which is not only the center of cellular respiration chain and oxidative phosphorylation but also the regulation center of apoptosis [28]. Recently, it has been reported that CB receptor agonists are mitochondrial inhibitors, because they can reduce oxygen consumption and reduce mitochondrial membrane potential and mitochondrial hydrogen peroxide products, thus promoting cell apoptosis [29]. In this study, caspase-3 was continuously cleaved and Cyto-C was released, and cleaved cappase-3 and Cyto-C protein expression were significantly increased in U251 cells during the 0-48 h period of U251 cells treated with 20 μM WIN, suggesting that apoptosis induced by WIN might be related to the mitochondrial pathway of apoptosis. ROS is an important factor in inducing apoptosis [30]. Studies have confirmed that ROS generation was an early apoptotic event, which can

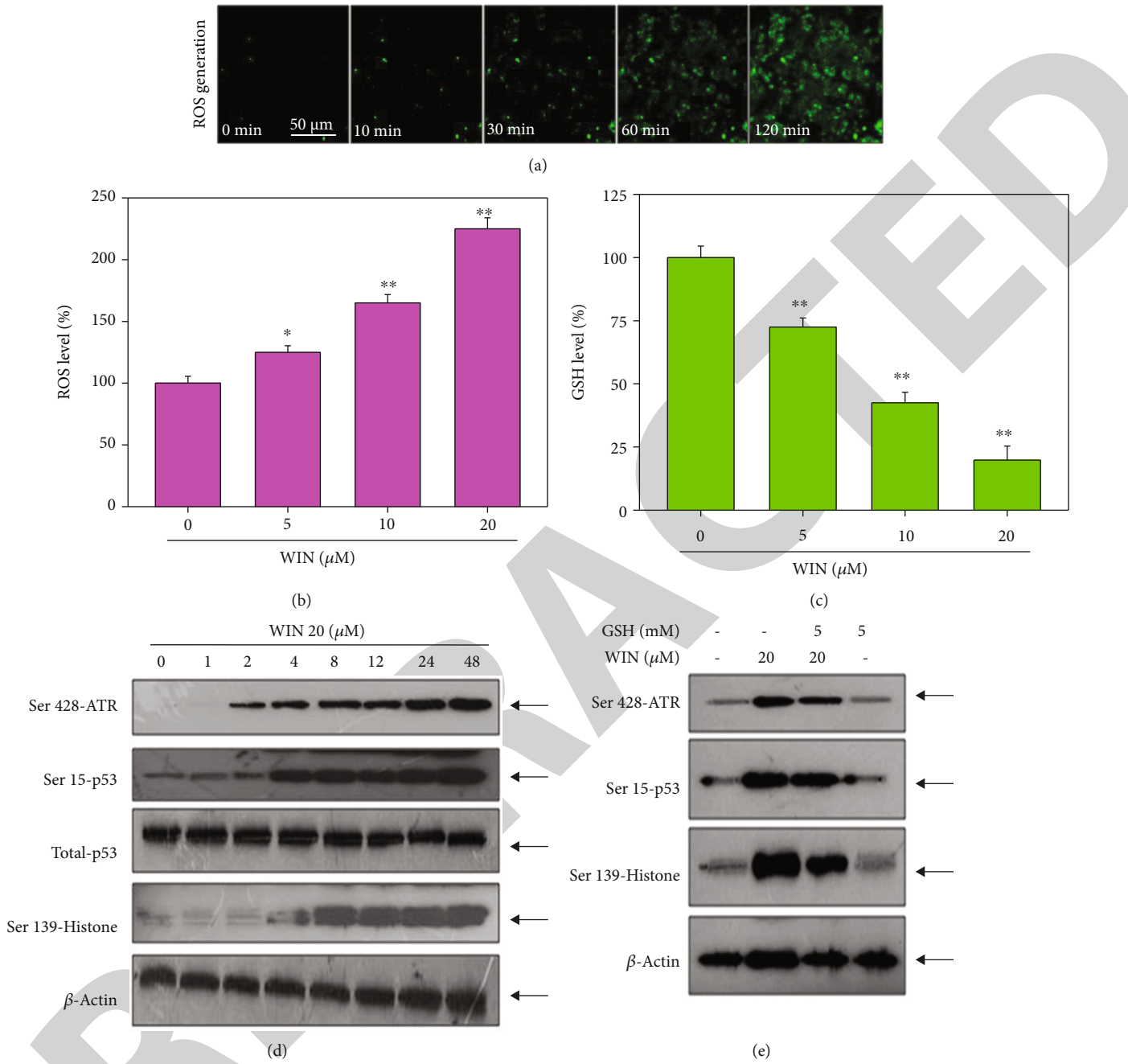


Figure 5: WIN triggers ROS-mediated DNA damage in U251 cells. (a) WIN caused time-dependent ROS generation in U251 cells. U251 cells seeded in 6-well plate were treated with 20 μM WIN for 0-120 min and loaded with DCFH-DA probe for 20 min in darkness. Then, cells were washed, and the ROS generation was imaged by a fluorescence microscope (magnification, ×200). (b) Quantitative analysis of ROS level. (c) GSH level. The ELIAS method was used to detect the GSH level in WIN-treated U251 cells. (d) WIN triggered time-dependent DNA damage in U251 cells. (e) ROS inhibition inhibited WIN-induced DNA damage in U251 cells. Cells were pretreated with GSH for 2 h and cotreated with 20 μM WIN for 48 h. Protein expression was detected by western blotting assay. All experiments were repeated three times.

be detected as early as 5 min [31–33]. Loss of mitochondrial membrane potential, opening of mitochondrial permeability transition pore (MPTP), and drug metabolism all contribute to the ROS generation [31–33]. In this study, it was found that WIN induced ROS production during U251 cell apoptosis, which triggered nonspecific signals in the downstream region and activated caspase-3 which leads to ultimately apoptosis, and this was further confirmed by the antioxidant GSH. To further explore its possible action mechanism, we examined the apoptosis-related protein p53. As an important tumor suppressor, P53 is involved in cell proliferation and apoptosis [34–36]. Western blotting exhibited an upregulation in the protein expression of Ser428-ATR, Ser15-p53, total-p53, and Ser139-Histone in U251 cells treated with WIN time-dependently (Figure 5(b)). Khan et al. [18] showed

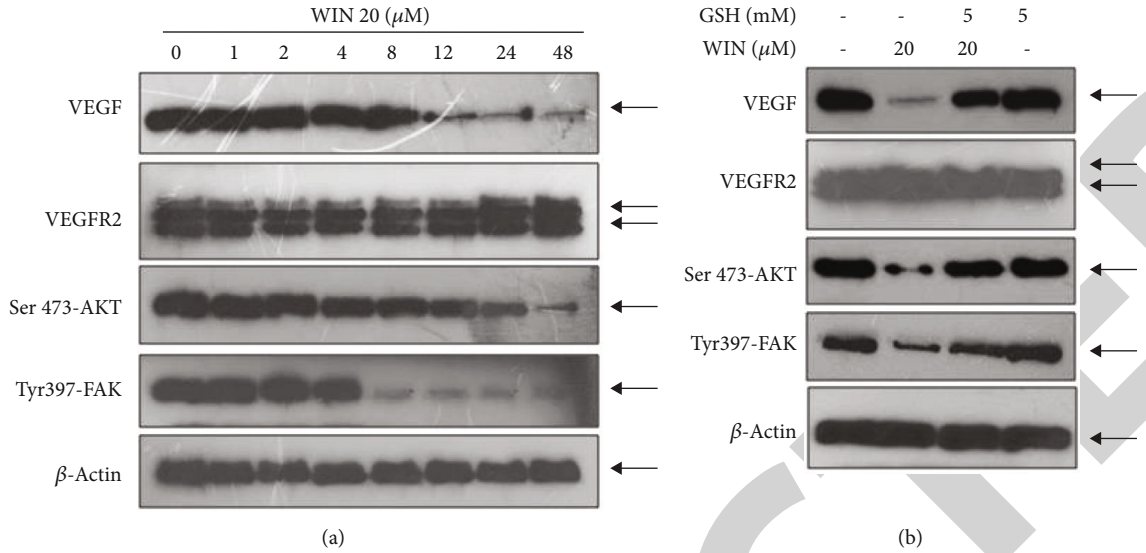


Figure 6: WIN inhibits VEGF-AKT/FAK signaling pathway in U251 cells. (a) WIN inhibited VEGF-AKT/FAK signaling pathway in a time-dependent manner in U251 cells. (b) ROS inhibition attenuated WIN-induced VEGF-AKT/FAK dysfunction in U251 cells. Cells were pretreated with GSH for 2 h and cotreated with 20 μM WIN for 48 h. Protein expression was detected by western blotting assay. All experiments were repeated three times.

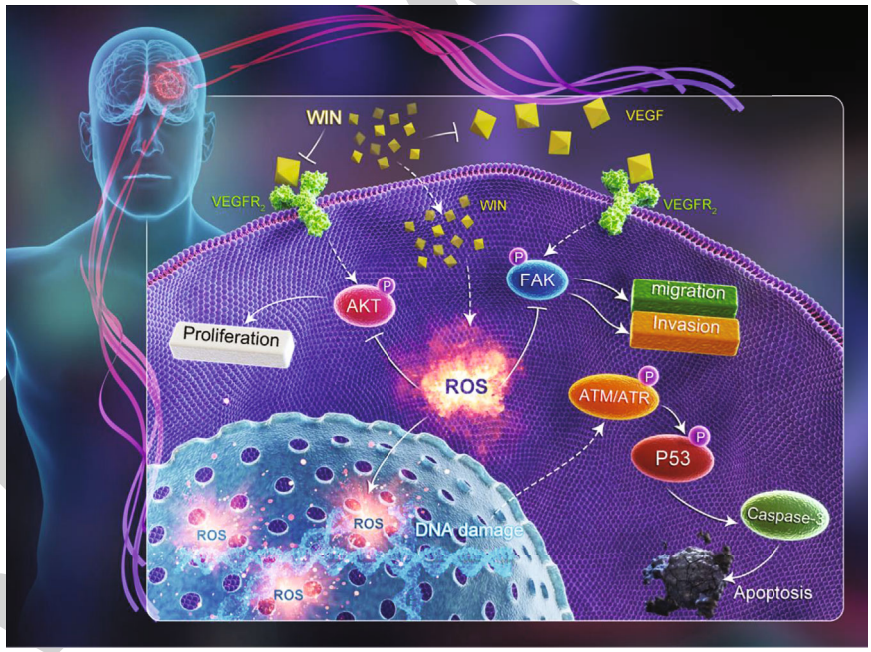


Figure 7: Proposed signal pathways. WIN induced ROS overproduction and subsequently caused DNA damage and dysfunction of VEGF-AKT/FAK signal axis, which eventually inhibited glioma cell migration, invasion, and proliferation.

that WIN could activate the classical apoptosis pathway of P53, increase the expression level of the proapoptotic protein Bax, and downregulate Bcl-2, which was consistent with the results of our experiment. It is known that VEGF-AKT/FAK signaling axis is the most important pathway regulating proliferation, migration, and invasion. Our experiment confirmed that WIN inhibited VEGF-AKT/FAK signaling pathway in U251 cells in a dose-dependent manner, thus resulting in the inhibition cell proliferation, migration, and invasion.

In summary, our findings provided evidence that WIN effectively inhibited human GBM cell proliferation, migration, and invasion and induces apoptosis in glioblastoma cells by triggering ROS-mediated DNA damage and regulating VEGF-AKT/FAK signals (Figure 7). This study provides a new idea of WIN for therapy of GBM with potential application in clinic. However, this study only explored the *in vitro* anticancer activity and mechanism in human glioma. The *in vivo* anticancer effect and mechanism need to be further evaluated.

Data Availability

The datasets used and/or analyzed during the current study are available within the manuscript.

Conflicts of Interest

The authors declare that there is no conflict of interest for all the authors.

Authors' Contributions

K Wang, CD Fan, and QB Ni designed the experiments. K Wang, Q Wang, MY Ni, ZQ Zhang, and J Gao performed the experiments. K Wang wrote the manuscript. All authors analyzed the data. All the authors have read and approved the final version of this manuscript. Kun Wang and Qian Wang contributed equally to this work.

Acknowledgments

This work was supported by the Natural Science Foundation of Shandong Province (No. ZR2020MH141, ZR2020MH134, ZR2019MH090), Shandong Post-Doctoral Innovation Project (No. 202003063), and Medical and Health Science and Technology Development Program of Shandong Province (No. 202003070813).

References

[1] Z. Kong, C. Jiang, R. Zhu et al., "^{18}F-FDG-PET-based radiomics features to distinguish primary central nervous system lymphoma from glioblastoma," *NeuroImage: Clinical*, vol. 23, article 101912, 2019.

[2] A. R. Dinu, A. F. Rogobete, T. Bratu et al., "*Cannabis sativa* revisited-crosstalk between microRNA expression, inflammation, oxidative stress, and endocannabinoid response system in critically ill patients with sepsis," *Cell*, vol. 9, no. 2, 2020.

[3] A. Kicman and M. Toczek, "The effects of cannabidiol, a non-intoxicating compound of cannabis, on the cardiovascular system in health and disease," *International Journal of Molecular Sciences*, vol. 21, no. 18, p. 6740, 2020.

[4] S. L. Calcaterra, A. N. Burnett-Hartman, J. D. Powers et al., "A population-based survey to assess the association between cannabis and quality of life among colorectal cancer survivors," *BMC Cancer*, vol. 20, no. 1, p. 373, 2020.

[5] O. Orellana-Serradell, C. E. Poblete, C. Sanchez et al., "Proapoptotic effect of endocannabinoids in prostate cancer cells," *Oncology Reports*, vol. 33, no. 4, pp. 1599–1608, 2015.

[6] E. Quoix and E. Lemarie, "Epidemiological novelties in lung cancer," *Revue des Maladies Respiratoires*, vol. 28, no. 8, pp. 1048–1058, 2011.

[7] V. Feinshtein, O. Erez, Z. Ben-Zvi et al., "Cannabidiol enhances xenobiotic permeability through the human placental barrier by direct inhibition of breast cancer resistance protein: an ex vivo study," *American Journal of Obstetrics and Gynecology*, vol. 209, no. 6, pp. 573.e1–573.e15, 2013.

[8] X. J. Wang, W. Chen, X. T. Fu et al., "Reversal of homocysteine-induced neurotoxicity in rat hippocampal neurons by astaxanthin: evidences for mitochondrial dysfunction and signaling crosstalk," *Cell Death & Discovery*, vol. 4, no. 50, pp. 1–12, 2019.

[9] H. L. Yang, V. Thiyagarajan, P. C. Shen et al., "Anti-EMT properties of CoQ0 attributed to PI3K/AKT/NFKB/MMP-9 signaling pathway through ROS-mediated apoptosis," *Journal of Experimental & Clinical Cancer Research*, vol. 38, no. 1, p. 186, 2019.

[10] K. Soderstrom, E. Soliman, and R. Van Dross, "Cannabinoids modulate neuronal activity and cancer by CB1 and CB2 receptor-independent mechanisms," *Frontiers in Pharmacology*, vol. 8, p. 720, 2017.

[11] S. Ghosh, S. Sheth, K. Sheehan et al., "The endocannabinoid/-cannabinoid receptor 2 system protects against cisplatin-induced hearing loss," *Frontiers in Cellular Neuroscience*, vol. 12, p. 271, 2018.

[12] B. M. Fonseca, G. Correia-da-Silva, and N. A. TeixeiraCannabinoid-induced cell death in endometrial cancer cells: involvement of TRPV1 receptors in apoptosis," *Journal of Physiology and Biochemistry*, vol. 74, no. 2, pp. 261–272, 2018.

[13] A. Capozzi, V. Mattei, S. Martellucci et al., "Anti-proliferative properties and proapoptotic function of new CB2 selective cannabinoid receptor agonist in Jurkat leukemia cells," *International Journal of Molecular Sciences*, vol. 19, no. 7, p. 1958, 2018.

[14] A. Bettiga, M. Aureli, G. Colciago et al., "Bladder cancer cell growth and motility implicate cannabinoid 2 receptor- mediated modifications of sphingolipids metabolism," *Scientific Reports*, vol. 7, no. 1, p. 42157, 2017.

[15] D. Roberto, L. H. Klotz, and V. Venkateswaran, "Cannabinoid WIN 55, 212-2 induces cell cycle arrest and apoptosis, and inhibits proliferation, migration, invasion, and tumor growth in prostate cancer in a cannabinoid-receptor 2 dependent manner," *Prostate*, vol. 79, no. 2, pp. 151–159, 2019.

[16] J. G. Latorre and E. B. Schmidt, "Cannabis, cannabinoids, and cerebral metabolism: potential applications in stroke and disorders of the central nervous system," *Current Cardiology Reports*, vol. 17, no. 9, p. 627, 2015.

[17] L. Muller, A. Radtke, J. Decker, M. Koch, and G. Belge, "The synthetic cannabinoid WIN 55, 212-2 elicits death in human cancer cell lines," *Anticancer Research*, vol. 37, no. 11, pp. 6341–6345, 2017.

[18] M. I. Khan, A. A. Sobocinska, K. K. Brodaczewska et al., "Involvement of the CB2 cannabinoid receptor in cell growth inhibition and G0/G1 cell cycle arrest via the cannabinoid agonist WIN 55,212-2 in renal cell carcinoma," *BMC Cancer*, vol. 18, no. 1, p. 583, 2018.

[19] X. Xian, L. Huang, B. Zhang, C. Wu, J. Cui, and Z. Wang, "WIN 55, 212-2 inhibits the epithelial mesenchymal transition of gastric cancer cells via COX-2 signals," *Cellular Physiology and Biochemistry*, vol. 39, no. 6, pp. 2149–2157, 2016.

[20] L. Pietrovito, M. Iozzo, M. Bacci, E. Giannoni, and P. Chiarugi, "Treatment with cannabinoids as a promising approach for impairing fibroblast activation and prostate cancer progression," *International Journal of Molecular Sciences*, vol. 21, no. 3, p. 787, 2020.

[21] B. J. Cridge and R. J. Rosengren, "Critical appraisal of the potential use of cannabinoids in cancer management," *Cancer Management and Research*, vol. 5, pp. 301–313, 2013.

Hindawi
BioMed Research International
Volume 2023, Article ID 9789542, 1 page
https://doi.org/10.1155/2023/9789542

Retraction

Retracted: Effects of Tai Chi on Patients with Mild Cognitive Impairment: A Systematic Review and Meta-analysis of Randomized Controlled Trials

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] R. Lin, S. Cui, J. Yang et al., "Effects of Tai Chi on Patients with Mild Cognitive Impairment: A Systematic Review and Meta-analysis of Randomized Controlled Trials," *BioMed Research International*, vol. 2021, Article ID 5530149, 10 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 5530149, 10 pages
https://doi.org/10.1155/2021/5530149

Review Article

Effects of Tai Chi on Patients with Mild Cognitive Impairment: A Systematic Review and Meta-analysis of Randomized Controlled Trials

Run Lin,[1,2] **Shaoyang Cui**,[1] **Juan Yang**,[3] **Huijun Yang**,[1,2] **Zitong Feng**,[4] **Dietlind L. Wahner-Roedler**,[3] **Xuan Zhou**,[5] **Manisha Salinas**,[6] **Molly J. Mallory**,[3] **Alexander Do**,[3] **Sara E. Bublitz**,[3] **Tony Y. Chon**,[3] **Chunzhi Tang**,[4] **Brent A. Bauer**,[3] **and Mingzhu Xu**[7]

[1]*Shenzhen Hospital of Guangzhou University of Chinese Medicine, Futian, Shenzhen, Guangdong, China*
[2]*Sixth Clinical Medical College, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China*
[3]*Division of General Internal Medicine, Mayo Clinic, Rochester, Minnesota, USA*
[4]*Clinical Medical College of Acupuncture Moxibustion and Rehabilitation, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China*
[5]*Formula-Pattern Research Center, School of Traditional Chinese Medicine, Jinan University, Guangzhou, Guangdong, China*
[6]*Division of Family Medicine, Mayo Clinic, Jacksonville, Florida, USA*
[7]*Shenzhen Hospital, Southern Medical University, Shenzhen, Guangdong, China*

Correspondence should be addressed to Brent A. Bauer; bauer.brent@mayo.edu and Mingzhu Xu; xulcui123@126.com

Received 24 February 2021; Revised 28 March 2021; Accepted 2 April 2021; Published 23 April 2021

Academic Editor: Min Tang

Background and Purpose. Mild cognitive impairment (MCI) is a common condition, which threatens the quality of life of older adults. Tai Chi (TC) is growing in popularity among patients with MCI. This study is aimed at evaluating the effectiveness and safety of TC in older adults with MCI. *Design*. Randomized controlled trials (RCTs) from multiple databases from inception to December 2020 published in English were searched. Two researchers independently performed eligible study screening and data extraction. The methodological quality was assessed with the Jadad score. Meta-analysis of RCTs on TC in the treatment of MCI was performed with RevMan Version 5.4.1. *Results*. Seven RCTs with 1265 participants were included. For most RCTs, the overall reporting of methodological quality was high. Results of the meta-analysis indicate that TC improved MCI patients' cognitive function significantly, including overall cognitive function (MD = −2.24, 95% CI -3.51 to -0.97, $P = 0.0005$), memory and learning (SMD = 0.83, 95% CI 0.22 to 1.45, $P = 0.008$), visuospatial ability (MD = 3.15, 95% CI 0.74 to 5.56, $P = 0.01$), executive functions (MD = 0.32, 95% CI 0.03 to 0.61, $P = 0.03$), and physical activity (MD = 18.78, 95% CI 10.80 to 26.76, $P < 0.00001$). However, no significant benefit was found for TC on psychological activity (MD = 0.17, 95% CI -0.62 to 0.96, $P = 0.36$) and biomarker improvement. *Conclusion*. The meta-analysis confirmed the clinical therapeutic effect of TC for MCI. More rigorous and long-term follow-up RCTs should be conducted in the future.

1. Introduction

Mild cognitive impairment (MCI) refers to a common condition in the elderly with a decline of memory, attention, and cognitive function which is more than expected considering age and education [1]. It is defined as an intermediate stage between normal aging cognitive decline and dementia [2]. A recent report shows that the incidence of MCI is 10%-25% in persons over the age of 65 [3]. With mild cognitive decline, most people suffer from progressive occurrence of

neuropsychiatric behaviors [4] and impaired engagement in activities of daily living [5], ultimately leading to a decreasing quality of life (QoL) and higher burdens on family care providers [6] and health professionals [7].

It has been estimated that 37.4% of MCI patients develop dementia within 27.09 ± 15.09 months of follow-up [8]. Therefore, it is critical to identify effective interventions that can delay cognitive decline in this vulnerable population.

To date, there are limited pharmacologic treatments with unclear risks to improve cognitive function or postpone the progression to dementia [9]. Physical activity has become a well-researched and verified behavior intervention for mild cognitive impairment and function and slowing the progression of dementia [10]. Previous reports have suggested that a physically active lifestyle could postpone or even prevent the progression of cognitive decline [11, 12]. Tai Chi (TC), also called Taiji or Taijiquan, is a type of psychophysiological exercise [13] based on traditional Chinese medicine. Relevant studies indicate that TC may be significantly conducive for older adults with MCI. TC is a series of rhythmic and continuous movements alongside cognitive stimulations such as movement recall [14, 15]. Compared with conventional rehabilitation therapy, TC is beneficial not only for neuromuscular activities but also for its associated characteristics, such as breathing control, mind-body relaxation, and mental focus, which may encourage MCI individuals to engage in TC exercises and maintain their physical and psychological well-being. Over the past decade, TC has become more globally widespread as a preferred therapeutic physical activity in patients with MCI [16], cognitive impairment [17], and dementia [18].

The number of clinical studies that explore cognitive effects of TC on patients with MCI has recently increased. A growing number of reviews have investigated both the generalized and specific correlation between TC and cognitive function [19–21], including a narrative review on cognitive function [19], evaluating global cognitive function [20], and a meta-analysis, which exclusively quantitatively analyzed the effect of TC on slowing cognitive progression in elders with MCI [21]. To date, there is no systematic review which has comprehensively analyzed the current literature.

Therefore, we synthesized evidence from RCTs that evaluated the impact of TC on participants with MCI within the areas of comprehensive cognition, physical function, neuropsychiatric aspects, and biomarkers.

2. Methods

According to our previous protocol, this review was performed with PROSPERO registration (http://www.crd.york.ac.uk/PROSPERO; registration number: CRD42019125104) and referenced from Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines [22].

2.1. Search Strategy. We comprehensively searched EBM Reviews, Embase, Scopus Medline, and Web of Science from database inception up to December 2020, identifying relevant studies published in the English language, using the following terms: "Tai Chi" ("tai-chi" or "t'ai-chi" or "tai-ji" or "taiji" or "taijiquan") and "cognition" (or "mild cognitive impairment" or "executive function" or "memory and learning" or "cognitive decline" or "early-stage dementia"). Relevant RCTs and reviews for additional studies were searched for further eligible studies. Two researchers independently performed the literature search.

2.2. Eligibility Criteria. Articles were selected for inclusion if they met the following criteria: (1) RCTs of TC intervention; (2) participants: adults age ≥ 55 years, diagnosed with MCI not caused by cerebral space-occupying lesion or craniocerebral trauma according to Petersen et al.'s criteria [3]; (3) interventions of controls which could be no treatment, sham, placebo, health education, or usual care; (4) outcomes which were comprehensive including overall cognitive function, memory and learning, visuospatial ability, executive functions, and physical activity and biomarkers; and (5) clinical trials reported in English. Non-RCTs including reviews, conference abstracts, protocols, cohorts, case series reports, and case-controls were excluded.

2.3. Data Extraction. Retrieved studies were imported into EndNote x9.1. The extracted data included (1) general information (the first authors, year of publication, and country), (2) characteristics of study (diagnostic criteria, sample size, age, intervention, and follow-up), and (3) outcome/measurement and elements for assessing the risk of bias. Data was reviewed by two investigators after reading the full-text studies independently. We extracted the mean (M) and standard deviation (SD) from four aspects of outcome to evaluate the effects of TC on MCI: the primary outcome: cognitive function (global cognitive function, memory and learning, visuospatial ability, and executive function), and the secondary outcome: physical activity, psychological evaluation, and related biomarkers.

2.4. Risk of Bias and Quality Assessment. The Cochrane Risk of Bias Tool for Randomized Controlled Trials [23] was used to assess potential sources of bias among the retrieved RCTs. The evaluation included six items on random sequence generation, allocation concealment, blinding of participants and researchers, blinding of outcome assessment, incomplete outcome data, selective reporting, and other bias. Each item was rated as a "high risk," "unclear risk," or "low risk" of bias [24]. Additionally, the methodological quality of each eligible study was assessed with the Jadad et al. scale [23]. The evaluation was determined in randomization, double blinding, and withdrawals and dropouts of participants. Studies with a total quality assessment score of 0–2 were regarded as low quality and 3–5 as high quality. Two authors performed the assessment independently, any discrepancy was resolved through consensus with the third author.

2.5. Data Synthesis and Statistical Analysis. Review Manager (RevMan) Version 5.4.1 (The Cochrane Collaboration, 2020) was used for statistical analysis. The dichotomous variables were expressed as the risk ratios (RRs) with 95% confidence intervals (CIs), and the continuous variables were estimated by mean difference (MD) or standardized mean difference (SMD) with 95% CIs. The statistical heterogeneity was performed by a standard chi-squared (χ^2) test and I^2 statistic

in this analysis. A fixed effects model was used if no significant heterogeneity existed ($P > 0.05$, $I^2 < 50\%$); a random effects model was selected if significant heterogeneity existed ($P \leq 0.05$, $I^2 \geq 50\%$). Publication bias was not assessed through funnel plots due to limited sample sizes. We detected the possible reasons for substantial heterogeneity by sensitivity analysis and subgroup analysis.

3. Results

3.1. Identification of Relevant Studies. The database search yielded 548 records. Of these, 4 records were removed as duplicates, and 501 records were excluded following screening of study titles and abstracts. The remaining 43 articles were further screened, and 1 article was removed due to no full-text availability; 35 articles were excluded as they did not meet the research criteria in this review. Outcome data of 2 RCTs [25, 26] were unavailable. Finally, 5 RCTs were included for meta-analysis. The literature selection process is presented in the flow chart (Figure 1).

3.2. Study Characteristics. There was a total of 7 studies which met inclusion criteria, published between 2012 and 2020. Three studies were conducted in Hongkong, China [25, 27, 28], two in Thailand [15, 16], one in Canada [26], and one in Turkey [29]. A total of 667 (TC group: 312; control group: 355) MCI participants were included in this analysis, of which most were diagnosed by Mini-Mental State Examination (MMSE), Clinical Dementia Rating (CDR), and Montreal Cognitive Assessment (MoCA). Five studies [15, 16, 26–28] compared TC and conventional rehabilitation modalities (including memory training, health education, stretching, and toning exercise). Two studies [25, 29] were comprised of comparisons between TC practitioners and patients without any physical exercises. Several kinds of TC were reported, including 10-form Yang-style Tai Chi [15, 16, 25], Taoist Tai Chi [26], simplified Yang-style Tai Chi [27, 28], and 24-form Tai Chi [29]. The duration of the intervention varied between 2 and 12 months, and duration of treatment session between 20 and 90 minutes. The information of all included RCTs' characteristics is summarized in Tables 1 and 2.

3.3. Methodological Quality and Risk of Bias. Detailed information on the evaluation of the methodological quality and risk of bias of each study was summarized with Revman. Six of the seven RCTs had a lower risk of bias of randomization allocation due to the specified reports of random-sequence generation [15, 16, 25, 27–29]. One [29] used the blocked randomization method, two trials [25, 28] used computer-generated random numbers, and three [15, 16, 27] used the randomization code. The other one had an unclear risk for randomization allocation due to a lack of detailed information [26]. Four RCTs were estimated at low-risk bias of allocation concealment [15, 16, 25, 28], one preferred grouping sequence list with password protection and stored on a computer [25], and three performed opaque sealed allocation [15, 16, 28]. Three trials had unclear allocation concealment without description in enough detail [26, 27, 29]. Five trials were blinded only in outcome assessment, without participants and personnel [15, 16, 25, 27, 28]. One trial had an unclear risk of performance bias as well as detection bias for its unclear blinding [26]. One trial was at unclear risk of performance bias and at low risk of detection bias [29]. Six trials had a low risk of bias with incomplete outcome data [15, 16, 25, 26, 28, 29], and the other one lacked detailing data of dropouts [27]. Three trials had a low risk of bias for their available prespecified outcomes with registration online [16, 25, 28], whereas the other four had a high-risk bias without reported prespecified outcomes [15, 26, 27, 29]. Three trials were estimated at low risk because of their good study design [16, 25, 28]. The other trials were assessed at high [26, 27] or unknown [15, 29] risk of bias in the other resource. For both poor unknown design of poor design allocation and selective reporting, two trials were estimated as high risk of bias in the other resource [26, 27]. The risk of bias of included studies was globally high (Figure 2). Four trials were evaluated as high quality because of their high jaded scores [15, 16, 25, 28], and the other three as poor quality [26, 27, 29]. The overall methodological quality of the RCTs was high (Table 1).

3.4. Effect of Interventions

3.4.1. Global Cognitive Function. The Chinese version of the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog) is the gold standard widely used in evaluating global cognitive function [30]. Two studies [27, 28] reported the effects of TC on global cognitive function assessed by the ADAS-Cog. The results of the data heterogeneity test $I^2 = 0\%$ showed improvement in the TC treatment group in MD scores of the global cognitive function, and the difference was statistically significant (MD = −2.24, 95% CI -3.51 to -0.97, $P = 0.0005$, $n = 272$, Figure 3).

3.5. Memory and Learning. The Logical Memory delayed recall was used to examine episodic memory [15]. Three studies [15, 16, 28] assessed the effects of TC on memory and learning ability with the Logical Memory delayed recall score. The results revealed a statistically significant mean effect size that favored TC over controls for memory and learning (SMD = 0.83, 95% CI 0.22 to 1.45, $P = 0.008$, $I^2 = 57\%$, $n = 126$, Figure 4).

3.6. Visuospatial Ability. The Block Design Test was used to evaluate visuospatial ability [16]. Two studies [15, 16] reported the effects of TC on visuospatial ability assessed by the Block Design Test. The results revealed a statistically significant mean effect size that favored TC over controls for visuospatial ability (MD = 3.15, 95% CI 0.74 to 5.56, $P = 0.01$, $I^2 = 0\%$, $n = 115$, Figure 5).

3.7. Executive Function. The digit span was used to assess executive function [16]. Three studies [15, 16, 27] involving 151 participants in the treatment group and 225 in the control group assessed the digit span. Study results revealed a statistically significant mean effect size that favored TC over controls for executive function (MD = 0.32, 95% CI 0.03 to 0.61, $P = 0.03$, $I^2 = 0\%$, $n = 376$, Figure 6).

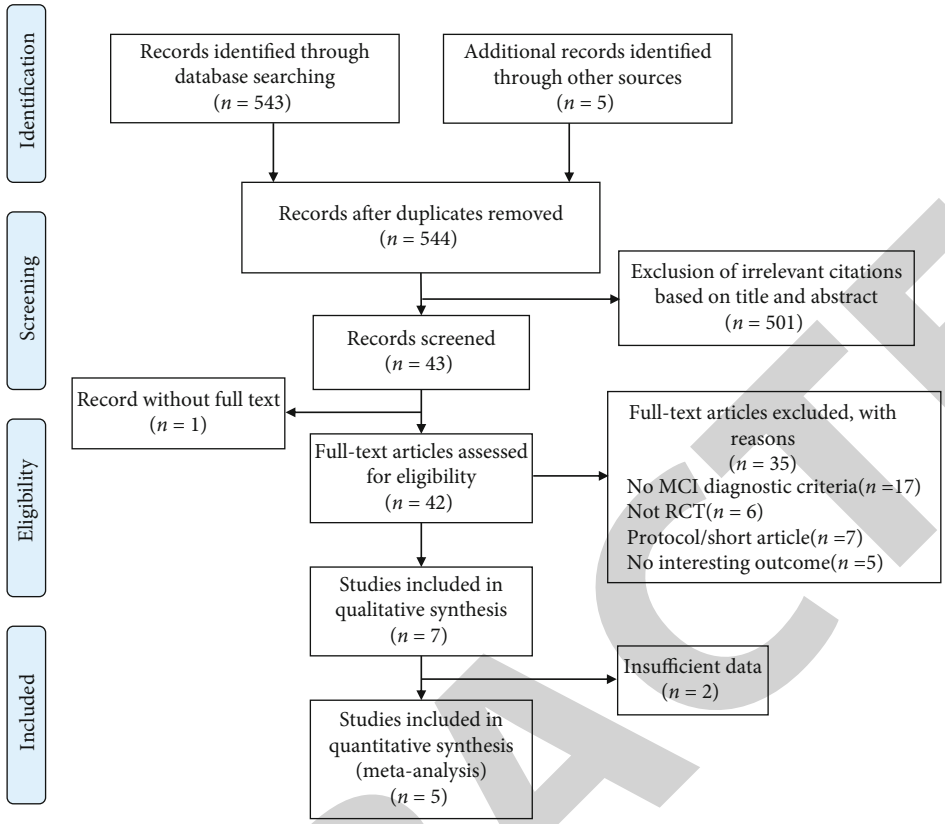


Figure 1: Flow diagram for study selection process.

3.8. Physical Activity. Physical Activity Scale for the Elderly (PASE), developed in 1993 to assess the status of physical activity in those with MCI, was used in the present study [31]. Two studies [28, 29] involving 25 participants in the treatment group and 28 in the control group assessed PASE. The results revealed a statistically significant mean effect size that favored TC over controls for physical activity (MD = 18.78, 95% CI 10.80 to 26.76, $P < 0.00001$, $I^2 = 0\%$, $n = 53$, Figure 7).

3.9. Psychological Evaluation. CSDD and GDS-15 were clinician-rated instruments assessing a range of biological and psychological symptoms associated with depression [27, 28]. Two studies with 272 participants showed the effects of TC on psychological activity, as measured by CSDD [27] and GDS-15 [28]. The results revealed no statistically significant mean effect size that favored TC over controls for psychological activity (MD = 0.17, 95% CI -0.62 to 0.96, $P = 0.68$, $I^2 = 0\%$, $n = 272$, Figure 8).

3.10. Biomarkers. Among all the included trials, only Sungkarat et al. [16] evaluated biomarkers and reported that the plasma brain-derived neurotrophic factor (BDNF) level was significantly increased in 33 older MCI adults, who completed 9 sessions of TC exercise guided by a certified instructor and then performed a consistent home practice with 50-minute video per session, thrice weekly, 72 sessions in 6 months totally ($P < 0.05$), whereas their interleukin-10 (IL-10) and tumor necrosis factor-α (TNF-α) were similar to those of their 33 counterparts ($P > 0.05$).

3.11. Adverse Events. Five studies reported AEs [15, 16, 25, 27, 28]. Death [27], leg pain [28], ill health [15, 16], and bone fractures [15, 16, 27] were reported. Except for study-unrelated bone fractures, it is unclear in the other studies whether AEs were related to TC intervention. Another trial reported no adverse effect [25]. Therefore, TC therapy appeared to be a safe treatment for patients with MCI.

4. Discussion

This study is the first review comprehensively evaluating the impact of TC on older adults with MCI. The included relevant studies [15, 16] revealed that during TC training, cognitive activities might contribute to the improvement of cognitive functions. Improving memory and executive functioning after TC exercise has been found to be associated with structural and functional changes in the cortex related to executive control processes, memory, and learning [32, 33]. Moreover, TC influences the cognitive functions of MCI patients in various degrees (global cognitive function (MD = −2.24), memory and learning (SMD = 0.83), visuospatial ability (MD = 3.15), and executive functions (MD = 0.32)). Specifically, the present study demonstrates that TC might help improve cognitive functions, as well as the physical abilities of people suffering from MCI. A positive

Table 1: Study characteristics of eligible studies.

Studies	Country	Diagnostic criteria	Participants IG/CG (M/F)	Mean age (y)	Intervention	Frequency & duration (Tai Chi)	Follow-up	Adverse event	Quality level
Sungkarat et al. [15]	Thailand	Amnestic MCI, MMSE ≥ 24, and MoCA < 26	33/33	IG: 68.3 (6.7) CG: 67.5 (7.3)	IG: 10-form TC CG: health education	3-week center-based + 12-week home-based TC (50 minutes per session, 3 times per week)	Not reported	One ill health, one study-unrelated ankle fracture	High (3/5)
Sungkarat et al. [16]	Thailand	Amnestic MCI, MMSE ≥ 24, and MoCA < 26	33/33	IG: 68.3 (6.7) CG: 67.5 (7.3)	IG: 10-form TC CG: health education	Practiced at home for 50 min/session, 3 times/week for 6 months	Not reported	One ill health, one study-unrelated ankle fracture	High (3/5)
Chan et al. [25]	China	MCI, MMSE 13–26, and CPSQI > 5	27/25	IG: 78.4 (7.1) CG: 82.2 (6.7)	IG: 10-form TC CG: usual activities	60-minute TCQ session twice a week for 2 months	6-month	None	High (3/5)
Fogarty et al. [26]	Canada	Amnestic MCI, MMSE	22/18	IG: 71.55 (9.33) CG: 72.61 (5.78)	IG: Taoist TC+MIP CG: memory intervention program (MIP)	2 × 90 min for 22 weeks	1-month follow-up session and a 3-month follow-up session	Not reported	Low (1/5)
Lam et al. [27]	China	CDR 0.5 or amnestic-mild MCI	92/169	IG: 77.2 (6.3) CG: 78.3 (6.6)	IG: 24-form simplified TC CG: stretching and toning exercise	3 × 30 min/week for 12 months	1 year	One died and one fall with bone fracture	Low (0/5)
Xu et al. [28]	China	MCI, HK-MoCA 19–21 (+1 point if education < 6 years)	6/6	IG: 76.43 (4.47) CG: 70.67 (4.23)	IG: 24-form simplified TC CG: health advice	12-week training + 12-week practice TC (30 minutes per session, 3 times per week)	6 months	One leg pain	High (3/5)
Birimoglu Okuyan and Deveci[29]	Turkey	MCI, MMSE, and MoCA < 25	20/22	IG: 74.21 (6.93) CG: 74.21 (6.93)	IG: Yang style of TC CG: not subjected to any physical practice	35-40 min/session, twice a week for 12 weeks	Not reported	Not reported	Low (1/5)

Note: MCI = mild cognitive impairment; MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment; HK-MoCA = Montreal Cognitive Assessment Hong Kong version; IG = intervention group; CG = control group; TC = Tai Chi.

Table 2: Outcome of eligible studies.

Studies	Cognitive function				Physical activity	Psychological evaluation	Biomarkers
	Global cognitive function	Memory and learning	Visuospatial ability	Executive function			
Sungkarat et al. [15]	*N	Logical Memory-delayed recall	Block Design Test	Digit span, TMT	*N	*N	*N
Sungkarat et al. [16]	*N	Logical Memory-delayed recall	e Block Design Test	Digit span, TMT	*N	*N	BDNF, TNF-α, IL-10
Chan et al. [25]#	MMSE	MIC	*N	NR	SF-12	*N	*N
Fogarty et al. [26]#	TEA, HVLT	RBMT, MAC-SR, digit symbol	*N	Digit span, TMT	RAPA scale, SF-36	*N	*N
Lam et al. [27]	CDR-SOB, ADAS-Cog, MMSE	*N	Visual span	Digit span	*N	CSDD, NPI	*N
Xu et al. [28]	ADAS-Cog, MoCA	Logical Memory delayed recall	*N	*N	EQ-VAS, PASE, EQ-5D	GDS-15, GAS-20	*N
Birimoglu Okuyan and Deveci [29]	(FaB scale) cognitive adaptations	*N	*N	*N	PASE	*N	*N

Note: CDR = Clinical Dementia Rating; ADAS-Cog = Alzheimer's Disease Assessment Scale-Cognitive Subscale; MMSE = Mini-Mental State Examination; TEA = Test of Everyday Attention; HVLT = Hopkins Verbal Learning Test; MoCA = Montreal Cognitive Assessment; MIC = Memory Inventory for Chinese; RMBT = Rivermead Behavioral Memory Test; MAC-SR = Memory Assessment Clinics Scale; TMT = Trail-Making Test; sF-12 = short-form 12; RAPA = Rapid Assessment of Physical Activity; SF-36 = RAND 36-Item Short-Form Health Survey-Medical Outcomes Study; PASE = Physical Activity Scale for the Elderly; EQ-VAS = EuroQol-visual analogue scale; EQ-5D = EuroQol-5 Dimension; NPI = Chinese Neuropsychiatric Inventory; CSDD = Cornell Scale for Depression in Dementia; BDNF = brain-derived neurotrophic factor; IL-10 = interleukin-10; TNF-α = tumor necrosis factor-α; * N = no report; #no data extracted.

trend in psychological ability was observed but failed to reach statistical significance.

To date, multifarious biomarkers have been available to support the diagnosis of neurodegenerative diseases in research and clinical settings. Diagnostic criteria for the preclinical asymptomatic phase of AD have been issued by The National Institute on Aging—Alzheimer's Association [34]. Biomarkers such as reduced cerebrospinal fluid (CSF) amyloid-β42 (Aβ42), increased cortical Aβ plaques or CSF tau levels, temporoparietal 18F-fluorodeoxyglucose positron emission tomography-computed tomography (18F-FDG PET-CT) hypometabolism, and medial temporal atrophy structures have been reported to be encompassed criteria contributing to MCI-AD diagnosis. AD biomarkers are essential to the identification of later dementia due to AD, and MCI cases show more than two pathologic biomarkers [35]. FDG-PET metabolism is lower in the MCI individuals which indicates that regional hypometabolism is closely linked with clinical progression of cognitive function, whereas entorhinal cortical thickness is thinner in MCI subjects. Diffusion tensor imaging might exceed CSF markers as a prophet of cognitive decline and cerebral atrophy. Another study indicates that AD biomarkers may aid in rehabilitation of MCI [36]; however, further research is needed to confirm these conclusions. Hitherto, there are very few studies on the mechanism of TC on biomarkers with MCI. In this review, only one included trial [16] mentioned the fluctuating plasma biomarkers of MCI, which suggests that TC is effective in favorably influencing biomarkers correlated with memory and learning. Other studies [32, 33] indicate that TC can enhance resting-state functional connectivity between the hippocampus and medial prefrontal cortex. Enhanced connectivity is significantly beneficial for memory ability and an increase in grey matter volume at the medial temporal lobe, putamen, and insula, which are essential regions for memory and cognition. TC could be an effective method for postponing or even possibly preventing the progression of cognitive decline during aging. Our review indicates that the physiological biomarker plasma BDNF might be responsible for the beneficial effects of TC on cognitive impairment. While this finding should be considered preliminary evidence due to the limited study number, further trials will help to substantiate its role in MCI biomarker improvement.

4.1. Strengths and Limitations. MCI is a transitional state between normal aging cognitive changes and dementia. TC, as an effective and special practice of mind-body exercise, has been described in many MCI case reports, but its exact mechanism is poorly understood. The efficacy of TC in MCI patients was confirmed from a set of high-quality studies in this analysis. Included RCTs were strictly screened, which means more rigorous study design and more reliable results. In addition, the current study analyzed the effects of TC not only on four specific cognitive domains but also on somatic and neuropsychiatric function qualitatively and quantitatively and the changes of MCI-related biomarkers. This should help to further investigate the comprehensive mechanism of TC in the treatment of older patients with MCI. A previously reported meta-analysis focused primarily on measuring the impacts of TC training on cognitive functioning of older adults with or without MCI [21]. In this study, we focused on older adults with MCI, which minimizes the potential bias from other diseases. We paid more attention to comprehensive physical and mental quality,

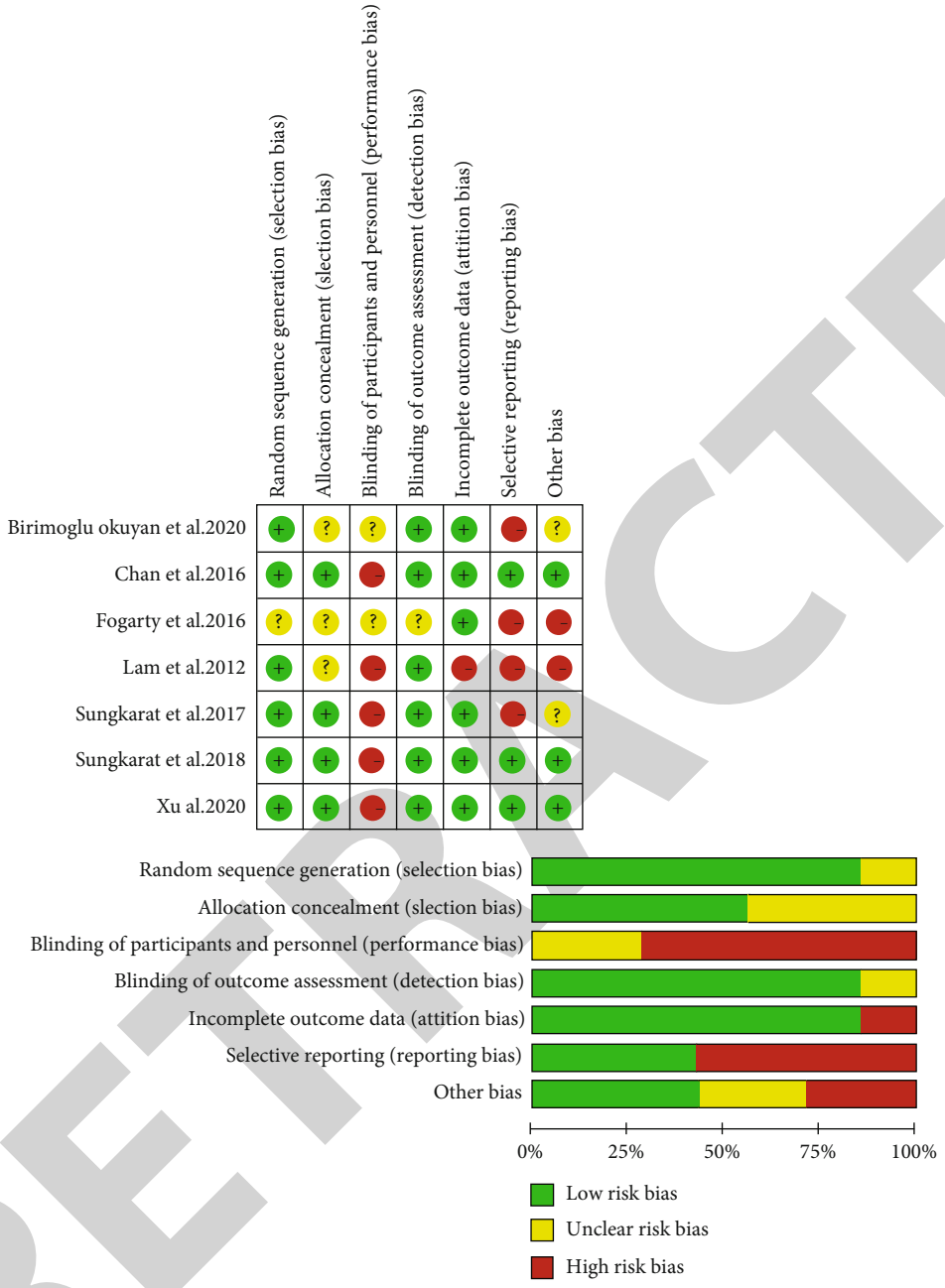


Figure 2: Risk of bias for each included study.

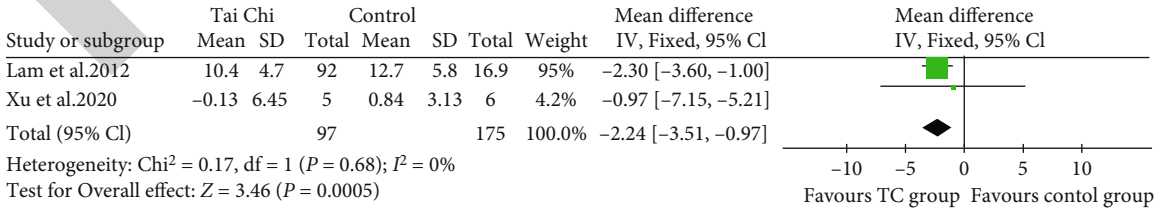


Figure 3: Forest plot for Tai Chi on global cognitive function.

which is conducive to the patients' return to society and family. These strategies support the conclusion of the efficacy of TC on cognition function of patients with MCI.

This review has some potential limitations. First, the number of included RCTs and sample size were limited with 7 RCTs including 1265 participants. Future research

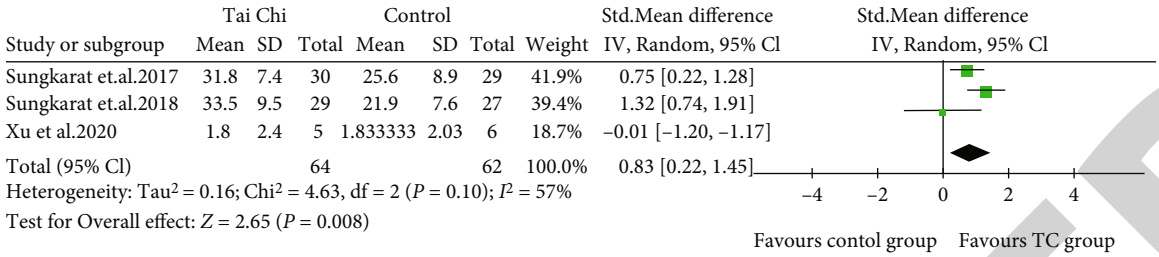


Figure 4: Forest plot for Tai Chi on memory and learning.

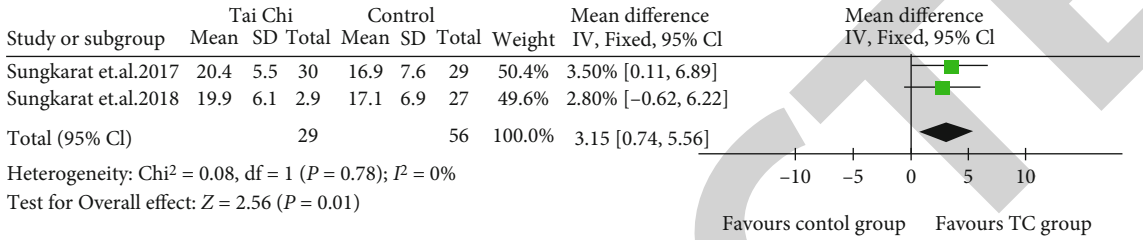


Figure 5: Forest plot for Tai Chi on visuospatial ability.

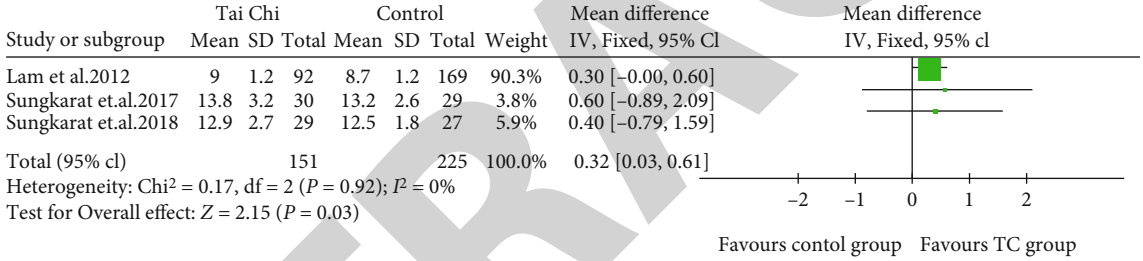


Figure 6: Forest plot for Tai Chi on executive function.

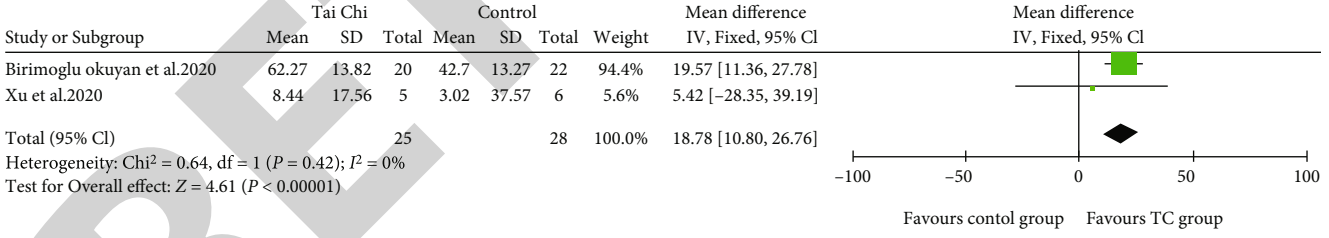


Figure 7: Forest plot for Tai Chi on physical activity.

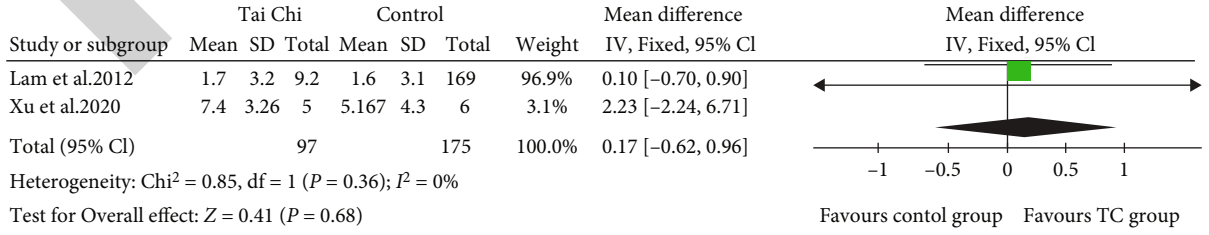


Figure 8: Forest plot for Tai Chi on psychological activity.

warrants more rigorous design of larger samples and multicenter RCTs. Second, the methodological quality assessment was far from being desirable. As an example, three included trials did not report detailed methods of allocation concealment [26, 27, 29]; four trials were at high risk of selective reporting, which could result in selection and reporting bias

with confounding of the results. No eligible trial was reported at lower risk of participants and personnel, so performance bias was inevitable. Additionally, studies identified from English databases reporting from non-English-speaking regions included in this review may to some degree cause a biased assessment of this topic. Our results suggest that TC is effective for MCI individuals, with heterogenous improvement across different cognitive domains, with little research regarding the effective mechanism of TC. Therefore, the results need to be interpreted with caution. Lastly, funnel plot analysis for the assessment of publication bias was not conducted due to the insufficient number of trials.

4.2. Practical Implications and Recommendations for Future Studies. This meta-analysis suggests that TC can safely improve cognitive function and physical activities in older adults with MCI when used with appropriate frequency and duration. It provides positive evidence for clinicians that this may be a conducive treatment for this population. TC may improve plasma BDNF, and therapeutic effects could be documented by serial measurements.

Originating from traditional Chinese medicine theories, TC training sessions in the evaluated studies varied regarding frequency, duration, and mode. Hence, in future studies, we recommend standardizing the TC treatment plan, including defined mode, duration, and frequency, to further investigate the comprehensive effectiveness of TC in MCI patients. Meanwhile, to promote population representativeness and avoid bias, recruiting criteria need to be more specific and systematic. Furthermore, the improvement between cognitive domains should be compared to further investigate the pertinency of TC therapy. Thus, more specific functional neuroimaging, vascular biochemical markers, and more sensitive and objective measurement methods are needed in the future.

5. Conclusions

This meta-analysis indicates that TC has positive clinical effects on cognitive function (global cognitive function, memory and learning, executive function, etc.), and physical abilities of older adults with MCI and provides a feasible approach to MCI management. Despite these positive results, it is hasty to arrive at a definite conclusion regarding the positive effect of TC for the treatment in older adults with MCI due to the general methodological quality and the heterogeneity of the included RCTs in this study. To provide stronger evidence, more multicenter, double-blinded, and placebo-controlled RCTs are required in the future.

Conflicts of Interest

The authors declare that there is no conflict of interest.

Authors' Contributions

Run Lin, Juan Yang, and Shaoyang Cui are joint first authors and contributed equally.

Acknowledgments

The authors would thank The HEAD Foundation, Singapore, for their support of this work. This study is supported by the Natural Science Foundation of China (No. 81904275), Natural Science Foundation of Guangdong Province (CN) (No. 2018A030313716), Shenzhen Special Fund for Introducing High-Level Medical Team Project (SZSM201502044), and Shenzhen science and technology innovation committee (JCY20190801182730737).

References

[1] S. A. Eshkoor, T. A. Hamid, C. Y. Mun, and C. K. Ng, "Mild cognitive impairment and its management in older people," *Clinical Interventions in Aging*, vol. 10, p. 687, 2015.

[2] H. Zhang, J. Wang, T. Sun et al., "A randomized controlled trial of combined executive function and memory training on the cognitive and noncognitive function of individuals with mild cognitive impairment: study rationale and protocol design," *Alzheimer's & Dementia*, vol. 4, no. 1, pp. 556–564, 2018.

[3] R. C. Petersen, O. Lopez, M. J. Armstrong et al., "Practice guideline update summary: mild cognitive impairment: report of the guideline development, dissemination, and implementation subcommittee of the American Academy of Neurology," *Neurology*, vol. 90, no. 3, pp. 126–135, 2018.

[4] Z. Ismail, E. E. Smith, Y. Geda et al., "Neuropsychiatric symptoms as early manifestations of emergent dementia: provisional diagnostic criteria for mild behavioral impairment," *Alzheimer's & Dementia*, vol. 12, no. 2, pp. 195–202, 2016.

[5] E. Cornelis, E. Gorus, I. Beyer, I. Bautmans, and P. De Vriendt, "Early diagnosis of mild cognitive impairment and mild dementia through basic and instrumental activities of daily living: development of a new evaluation tool," *PLoS Medicine*, vol. 14, no. 3, article e1002250, 2017.

[6] A. Morlett Paredes, P. B. Perrin, S. V. Peralta, M. E. Stolfi, E. Morelli, and J. C. Arango-Lasprilla, "Structural equation model linking dementia cognitive functioning, caregiver mental health, burden, and quality of informal care in Argentina," *Dementia*, vol. 16, no. 6, pp. 766–779, 2017.

[7] A. Martyr, S. M. Nelis, C. Quinn et al., "Living well with dementia: a systematic review and correlational meta-analysis of factors associated with quality of life, well-being and life satisfaction in people with dementia," *Psychological Medicine*, vol. 48, no. 13, pp. 2130–2139, 2018.

[8] P.-H. Chen, S.-J. Cheng, H.-C. Lin, C.-Y. Lee, and C.-H. Chou, "Risk factors for the progression of mild cognitive impairment in different types of neurodegenerative disorders," *Behavioural Neurology*, vol. 2018, Article ID 6929732, 8 pages, 2018.

[9] J. H. Jeong, H. R. Na, S. H. Choi et al., "Group- and home-based cognitive intervention for patients with mild cognitive impairment: a randomized controlled trial," *Psychotherapy and Psychosomatics*, vol. 85, no. 4, pp. 198–207, 2016.

[10] L. E. Bechard, "Physical Activity Participation in Mild Cognitive Impairment and Alzheimer's Disease," *Rehabilitation Sciences Institute, University of Toronto, Doctoral dissertation*, 2017.

[11] K. Bademli, N. Lok, M. Canbaz, and S. Lok, "Effects of physical activity program on cognitive function and sleep quality in elderly with mild cognitive impairment: a randomized

Hindawi
BioMed Research International
Volume 2023, Article ID 9827803, 1 page
https://doi.org/10.1155/2023/9827803

Retraction

Retracted: Comparison of 10-Year Survival Outcomes for Early Single Hepatocellular Carcinoma following Different Treatments

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] F. Meng, H. Zhang, H. Peng, and S. Lu, "Comparison of 10-Year Survival Outcomes for Early Single Hepatocellular Carcinoma following Different Treatments," *BioMed Research International*, vol. 2021, Article ID 6638117, 8 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6638117, 8 pages
https://doi.org/10.1155/2021/6638117

Research Article

Comparison of 10-Year Survival Outcomes for Early Single Hepatocellular Carcinoma following Different Treatments

Fanyu Meng,[1,2] **Haoyun Zhang**,[1] **Haiwen Peng**,[3] **and Shichun Lu**[1]

[1]*Department of Hepatobiliary Surgery, First Medical Center of Chinese People's Liberation Army (PLA) General Hospital, Beijing 100853, China*
[2]*Department of Hepatobiliary Surgery, Chifeng Clinical Medical School of Inner Mongolia Medical University, Chifeng, Inner Mongolia 024000, China*
[3]*Department of Orthopedics Surgery, First Medical Center of Chinese People's Liberation Army (PLA) General Hospital, Beijing 100853, China*

Correspondence should be addressed to Fanyu Meng; mfymfy2000@163.com

Received 28 November 2020; Revised 21 January 2021; Accepted 22 February 2021; Published 20 March 2021

Academic Editor: Min Tang

Introduction. To compare the actual 10-year survival outcomes of early single hepatocellular carcinoma (HCC) patients between 3 first-line treatments: radiofrequency ablation (RFA), surgical resection (SR), or transplantation (LT). *Methods*. A total of 1255 early single HCC patients retrieved from the Surveillance Epidemiology and End Results (SEER) database were included. Patients survived ≥10 years, and patients died <10 years were compared. Significant predictors associated with 10-year survival were identified by multivariate logistic regression analysis. The 10-year survival outcomes of 3 treatments were compared using multivariate model risk adjustment and inverse probability of treatment weighted (IPTW) adjustment. *Results*. Of the 1255 patients, 472 patients underwent SR, 259 patients underwent LT, and 524 patients underwent RFA. 149 patients achieved 10-year survival. Multivariate logistic regression analysis showed that age, race, treatment, and fibrosis score were significant predictors for 10-year survival, and LT had the best advantage of 10-year survival, followed by SR. Comparable 10-year survival outcomes were found between SR and RFA after IPTW. Then, a subgroup analysis was performed based on the tumor size, and the results showed that for ≤50 mm tumor, SR showed no significant advantages over RFA for 10-year survival. *Conclusions*. Estimates of the observational association of different treatments with 10-year survival are sensitive to the analytic method. LT showed the best outcomes for patients. No significant differences for 10-year survival were found between SR and RFA in the IPTW cohort. Subgroup analysis showed that for >50 mm tumor, SR showed significant advantages over RFA after IPTW.

1. Introduction

Hepatocellular carcinoma (HCC) is a common malignant tumor. It accounts for around 80% primary liver cancers worldwide [1]. HCC is estimated to be the fourth cause of cancer-related death overall worldwide [2]. Despite considerable progress in risk factor control, molecular profiles, early detection, diagnosis, treatment and so on, the incidence and cancer-related mortality in HCC patients are still increasing in many countries [3].

Several studies have focused on the 10-year survivors of HCC, and these studies only focus on one treatment [4–7]. Liver transplantation is the optimal treatment option for early HCC, but it was limited by lack of donor organs. One study reported that the actual 10-year survival rate for HCC patients who had liver transplantation was 41.1% [6]. For patients who underwent resection, the actual 10-year survival rate of patients was about 7.2%, whereas the actuarial survival quoted from the same studies was 26.8% [8]. Recurrence does not preclude long-term survival [9]. Different 10-year survival calculation methods lead to significant differences in results. The Kaplan-Meier method of actuarial survival analysis tends to overestimate survival outcomes [8]. Therefore, actual 10-year survivors were analyzed in this study.

Liver transplantation (LT), surgical resection (SR), and radiofrequency ablation (RFA) are recommended for early

HCC [10]. The three treatments were all analyzed in this study. A total of 1255 early single HCC patients were included, and we overcame the limitation of small sample size and long-term follow-up using data from the SEER database. In order to better compare the effect of treatments on 10-year survival, we used IPTW, a propensity scoring method to minimize the treatment selection deviation. We identify patient characteristics associated with 10-year survival using multivariate logistic regression analysis. The 10-year survival outcomes of the 3 treatments were also compared.

2. Methods

2.1. Database and Patient Selection. The current study was a retrospective cohort study of HCC patients identified from the Surveillance Epidemiology and End Results (SEER) database. The data (incidence-SEER 18 Regs Custom Data with additional treatment fields, Nov 2018 Sub, 1975 -2016 varying) was obtained by the SEER*Stat software (version 8.3.6; http://seer.cancer.gov/seerstat/). The following surgery codes were used: RFA : 16; SR : 20 to 25, 30, 36, 37, 50, 51, and 52; and LT : 61.

Patients with labeled primary sites of C22.0-liver and pathologically diagnosis (ICD-O-3 histologic type: 8170-8175) were carefully reviewed. The inclusion criteria were as follows: (1) first malignant tumor ($N = 93230$); (2) with detailed TNM stage information and no metastasis ($N = 45655$); (3) patients received RFA, SR, or LT were included ($N = 11243$); (4) single tumor ($N = 7638$); and (5) patients with at least 10-year follow-up were included, and patients who were alive but without 10-year follow-up were excluded ($N = 3754$). The exclusion criteria were as follows: (1) unknown race record ($N = 10$), (2) unknown exact tumor size ($N = 137$), and (3) unknown liver fibrosis score ($N = 2352$).

We collect sociodemographic data (age, sex, race), pathologic data (histologic type, grade, fibrosis score, tumor size, and T and N stages), and follow-up data for each patient. In the SEER database, the fibrosis score had been grouped into F0 (fibrosis score 0-4) and F1 (fibrosis score 5-6), respectively. Based on tumor size, tumor number, lymph node involvement, and vascular invasion, the AJCC 8th edition staging systems were derived as previously reported [11].

2.2. Statistical Analyses. Descriptive statistics are reported as number of events, mean ± SD, or median (interquartile range). Differences between the 10-year survivors and 10-year nonsurvivors were compared using Student's t-test or Mann–Whitney U test for continuous variables. For categorical variables, chi-square test or Fisher exact test were performed. Then, univariate and multivariate logistic regression analysis was performed to identify independent predictive factors for 10-year survival. Propensity scores were estimated with multinomial logistic regression, with treatments as outcomes and age, sex, T stage, race, grade, tumor size, fibrosis score, and regional lymph node involvement as pretreatment covariates. Inverse probability of treatment weighted was calculated with the estimated propensity scores [12]. The standardized mean differences (SMD) in covariate values between treatment groups were compared in the IPTWs samples [13]. All statistical analysis was performed using *R* 3.6.1 (https://www.r-project.org/) and SPSS 22.0.

3. Results

3.1. Patient Characteristics. This study focused on the actual 10-year survival, so patients who were alive but without 10-year follow-up were excluded in this study. Finally, a total of 1255 patients who met the inclusion criteria were included in this study. The median follow-up of the whole cohort was 135 months for overall survival. 149 (11.87%) of them achieved 10-year survival. All patients had an early T stage; of them, 269 patients were at T1a stage, 828 patients were at T1b, and 158 were T2. Among them, 472 patients underwent liver resection, 259 patients underwent liver transplantation, and 524 patients underwent RFA. The actual 10-year survival rate of patients who underwent LT was 34.0%. For patients who underwent SR and RFA, the actual 10-year survival rate was 9.1% and 3.4%, respectively. As shown in Table 1, patients who survive ≥10 years were younger than those who survive <10 years. Besides, significant differences were also found in T stage ($P = 0.021$), race ($P < 0.001$), tumor grade ($P = 0.011$), and treatment ($P < 0.001$) between the two groups (Table 1). The survival curves of patients stratified by treatment were shown in Figure 1(a). The median survival months for RFA, SR, and LT were 24, 30, and 67 months, respectively.

3.2. Treatment Characteristics and Logistic Regression Analysis of Unadjusted Cohort. Patients' baseline characteristics stratified by treatment were shown in Table 2. The baseline characteristics comparison of the unweighted cohort showed that age ($P < 0.001$), T stage ($P < 0.001$), race ($P < 0.001$), grade ($P < 0.001$), tumor size ($P < 0.001$), and fibrosis score ($P < 0.001$) were significantly different among the 3 treatments. Only sex ($P = 0.286$) and regional lymph node involvement ($P = 0.84$) showed no significant differences among the 3 treatments and with the SMD less than 0.1. Multivariate logistic regression analysis adjusting for all baseline patient characteristics showed that age, race, treatment, and fibrosis score were significant predictors for 10-year survival as shown in Table 3. While sex, T stage, grade, tumor size, and regional lymph node involvement were not significant predictors for 10-year survival.

3.3. Patient and Treatment Characteristics of IPTW-Adjusted Cohort. IPTW was used to balance the patient baseline characteristics of the 3 treatments. After IPTW, only age, tumor size, and fibrosis score differed significantly between the 3 treatments as shown in Table 2. The survival curves of weighted cohort stratified by treatment were shown in Figure 1(b).The change of SMD is shown in Figure 1(c). After IPTW, the actual 10-year survival rate of patients who underwent LT, SR, and RFA was 33.1%, 8.4% and 3.8%, respectively. The univariate logistic regression analysis for 10-year survival was performed in the IPTW cohort, and the results were shown in Table 4 (model 3). The results showed that

Table 1: Comparison of patient characteristics between 10-year survivors and 10-year nonsurvivors.

Variables	<10 year ($N = 1106$)	≥10 year ($N = 149$)	P
Age (mean (SD))	62.33 (9.75)	55.94 (10.22)	<0.001
Sex (%)			0.91
Male	855 (77.3)	114 (76.5)	
Female	251(22.7)	35(23.5)	
T stage (%)			0.021
T1a	224 (20.3)	45 (30.2)	
T1b	740 (66.9)	88 (59.1)	
T2	142 (12.8)	16 (10.7)	
Race (%)			<0.001
Black	136 (12.3)	4 (2.7)	
Other	264 (23.9)	52 (34.9)	
White	706 (63.8)	93 (62.4)	
Grade (%)			0.011
Well differentiated	200 (18.1)	37 (24.8)	
Moderately differentiated	361 (32.6)	59 (39.6)	
Poorly differentiated	138 (12.5)	17 (11.4)	
Unknown	407 (36.8)	36 (24.2)	
Tumor size (median (interquartile range))	32 (23-48)	30 (20-45)	0.044
Fibrosis score (%)			0.432
F0	332(30)	50 (33.6)	
F1	774 (70.0)	99 (66.4)	
Regional lymph node involvement (%)			0.335
No	1097 (99.2)	146 (98.0)	
Yes	9 (0.8)	3 (2.0)	
Therapy (%)			<0.001
RFA	506 (45.8)	18 (12.1)	
Resection	429 (38.8)	43 (28.9)	
Liver transplantation	171 (15.4)	88 (59.0)	

compared to RFA, patients who received LT have significant higher odds of 10-year survival (OR: 1.34, 95% CI: 1.25-1.45, $P < 0.001$), while patients who received SR had comparable odds of 10-year survival (OR: 1.05, 95% CI: 1.00-1.10, $P = 0.066$). IPTW cohort was further adjusted for age, tumor size, and fibrosis score in the multivariate logistic regression analysis, and the results showed that the 10-year survival odds for patients who received SR had no significant differences compared to RFA (OR: 1.04, 95% CI: 0.99-1.10, $P = 0.138$).

Unadjusted, multivariable adjusted, univariate IPTW-adjusted, and multivariable IPTW cohort logistic regression models for 10-year survival of the three treatments were shown in Table 4. Of the 4 methods used to compare 3 treatments, all showed that LT had the best advantage of 10-year survival compared to RFA, but the OR (95% CI) of the LT vs RFA decreased from 14.47 (8.66-25.44) in the model 1 to 1.34 (1.24-1.45) in the model 4. Similar results shown in SR vs RFA. These results suggested the presence of treatment selection bias. Moreover, after IPTW, patients who underwent RFA showed comparable 10-year survival with SR as shown in Table 4 (model 3 and model 4).

3.4. Subgroup Analysis. Although RFA showed comparable 10-year outcomes with SR in the entire IPWT cohort, the therapeutic effect of RFA is affected by the tumor size, of which the most commonly reported size cutoff values are 30 mm and 50 mm [14, 15]. Therefore, we stratified patients into different tumor size groups according to 30 mm and 50 mm. Then, after reweighting by IPTW, the 10-year survival outcomes were compared in each group. As shown in Table 5, for ≤30 mm tumor, LT showed significant advantages over SR ($P < 0.001$) and RFA ($P < 0.001$), but no significant differences were found between SR and RFA. For 30-50 mm tumor, only the comparison between LT and RFA showed significant differences, while no significant differences were found for SR vs RFA ($P = 0.090$) and LT vs SR ($P = 0.192$). For >50 mm tumor, LT showed the best outcomes, followed by SR.

4. Discussion

In this study, 1255 HCC patients with early and very early stage were included in this study. Univariate and multivariate

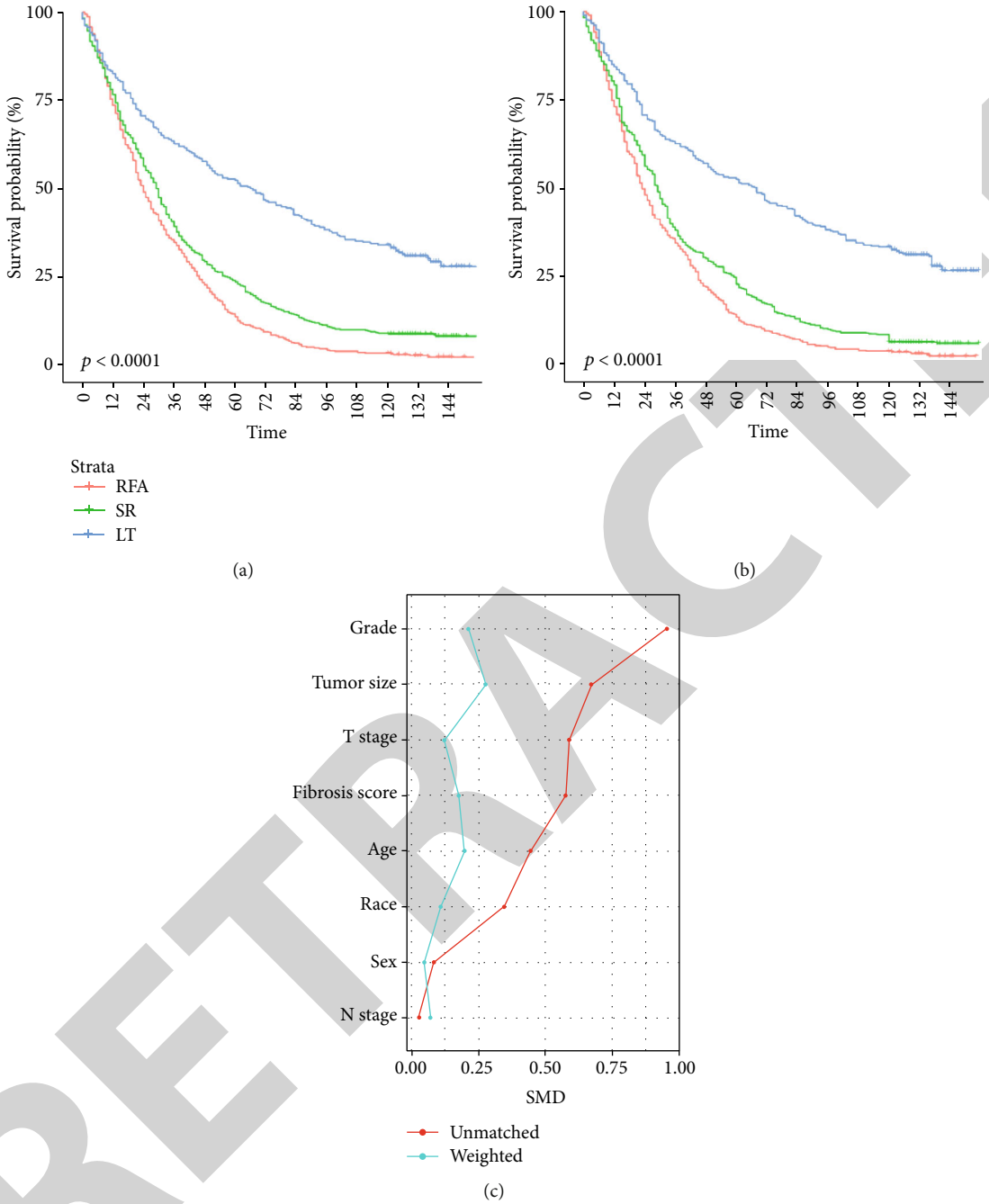


Figure 1: Overall survival of patients stratified by treatment (a). Overall survival of inverse probability treatment weighted-adjusted cohort stratified by treatment (b). Comparison of baseline characteristic standardized mean differences (SMD) of patients before and after inverse probability treatment weighting (c).

logistic regression analysis was used to compare the actual 10-year survival outcomes of 3 treatments for HCC patients. IPTW was used to balance the bias between 3 treatments.

Using multivariate logistic regression analysis adjusted for patient covariates, the significant predictors for 10-year survival of early-stage HCC were age, race, treatment, and fibrosis score. However, *T* stage and tumor size were not significant predictors for the early single HCC. Among these predictive factors, fibrosis score is an important one. Cirrhosis was reported to be an independent risk factor associated with actual 10-year survival after hepatectomy of HBV-related HCC [4]. In this study, we found that patients with high fibrosis score (OR 0.58, 95% CI: 0.36-0.92) were associated with less chance of l0-year survival compared with those who had low fibrosis score in the multivariate logistic regression analysis.

The comparisons of the outcomes between different treatments may be biased due to important baselines differences among patients, often as a result of treatment selection biases [16]. Therefore, to efficiently reduce the effect of

Table 2: Baseline characteristics of patients undergoing radiofrequency ablation, surgical resection, or liver transplantation in the unweighted and weighted study population.

	Unweighted cohort					Weighted cohort				
	Resection	Live transplantation	RFA	P	SMD	Resection	Live transplantation	RFA	P	SMD
	472	259	524			1060.73	934	1080.03		
Age (mean (SD))	62.79 (11.23)	56.85 (7.42)	62.80 (9.27)	<0.001	0.445	61.64 (10.38)	59.38 (7.74)	61.99 (9.19)	0.005	0.196
Sex = male (%)	355 (75.2)	208 (80.3)	406 (77.5)	0.286	0.082	806.7 (76.1)	737.1 (78.9)	830.5 (76.9)	0.778	0.046
T stage (%)				<0.001	0.588				0.531	0.122
T1a	47 (10.0)	94 (36.3)	128 (24.4)			206.1 (19.4)	220.3 (23.6)	238.6 (22.1)		
T1b	312 (66.1)	141 (54.4)	375 (71.6)			706.1 (66.6)	607.9 (65.1)	743.0 (68.8)		
T2	113 (23.9)	24 (9.3)	21 (4.0)			148.5 (14.0)	105.9 (11.3)	98.5 (9.1)		
Race (%)				<0.001	0.345				0.651	0.106
Black	58 (12.3)	17 (6.6)	65 (12.4)			117.5 (11.1)	92.2 (9.9)	116.6 (10.8)		
Other	158 (33.5)	42 (16.2)	116 (22.1)			292.0 (27.5)	200.3 (21.4)	260.4 (24.1)		
White	256 (54.2)	200 (77.2)	343 (65.5)			651.1 (61.4)	641.5 (68.7)	703.1 (65.1)		
Grade (%)				<0.001	0.956				0.108	0.212
Well differentiated	86 (18.2)	64 (24.7)	87 (16.6)			211.4 (19.9)	175.8 (18.8)	207.6 (19.2)		
Moderately differentiated	245 (51.9)	96 (37.1)	79 (15.1)			400.3 (37.7)	307.6 (32.9)	348.4 (32.3)		
Poorly differentiated	98 (20.8)	32 (12.4)	25 (4.8)			153.4 (14.5)	108.2 (11.6)	86.2 (8.0)		
Unknown	43 (9.1)	67 (25.9)	333 (63.5)			295.6 (27.9)	342.4 (36.7)	437.7 (40.5)		
Tumor size (mean (SD))	58.29 (43.14)	27.86 (14.33)	31.13 (14.85)	<0.001	0.671	42.99 (34.07)	32.49 (15.33)	34.86 (19.15)	<0.001	0.276
Fibrosis score = F1 (%)	225 (47.7)	221 (85.3)	427 (81.5)	<0.001	0.576	701.0 (66.1)	726.3 (77.8)	813.2 (75.3)	0.021	0.174
Lymph node involvement = yes (%)	4 (0.8)	2 (0.8)	6 (1.1)	0.84	0.026	7.3 (0.7)	4.1 (0.4)	15.3 (1.4)	0.28	0.069

treatment selection bias, IPTW was used in this study. IPTW effectively reduces the bias of baseline characteristics among treatment modalities as shown in Figure 1(c). After IPTW, the OR for 10-year survival of LT vs RFA decreased from 14.47 (95% CI: 8.66-25.44) in the model 1 to 1.34 (95% CI: 1.24-1.45) in the model 4 for early single nodule HCC patients. These results suggested that the estimation of the observational association between different treatments and 10-year HCC survival is highly sensitive to the analysis method. In each model, LT still showed the significant best survival advantage over other treatments. SR (OR:2.22, 95% CI: 1.15-4.39, $P = 0.019$) showed significant differences with RFA in the multivariate logistic regression model of the unadjusted cohort, while SR (OR: 1.04,95% CI: 0.99-1.10, $P = 0.138$) showed no significant differences with RFA in the IPTW cohort (model 4).

Only 20% of HCC patients are eligible for surgical resection [17]. Many HCC patients are not suitable for surgery due to long-term virus infection, impaired liver function, fibrosis, tumor location, and multifocal tumors. Therefore, more patients are candidates for curative therapy with RFA. Although RFA was often used in patients who were ineligible for SR, we found that for early single tumor, patients who underwent RFA showed comparable 10-year survival outcomes with SR. Similarly, one study reported that RFA could achieve long-term survival for as long as 10 years [7]. RFA can provide similar long-term survival results to SR of single nodular HCC when combined with multimodal treatment

Table 3: Univariate and multivariate logistic regression analysis of predictors for 10-year survival in all patients (unweighted cohort).

Variable	Univariate		Multivariate	
	OR (95% CI)	*P*	OR (95% CI)	*P*
Age, year	0.94 [0.92-0.95]	<0.001	0.95 [0.93-0.97]	<0.001
Sex (male vs female)	0.96 [0.64-1.45]	0.828	0.77 [0.49-1.22]	0.255
Race				
Black	1		1	
Other	6.70 [2.67-22.48]	<0.001	7.70 [2.90-26.82]	<0.001
White	4.48 [1.83-14.82]	0.004	3.71 [1.45-12.62]	0.015
T stage				
T1a	1			
T1b	0.59 [0.40-0.88]	0.008	0.99 [0.60-1.63]	0.957
T2	0.56 [0.30-1.01]	0.062	0.76 [0.35-1.59]	0.471
Grade				
Well differentiated	1		1	
Moderately differentiated	0.88 [0.57-1.39]	0.586	0.85 [0.51-1.42]	0.526
Poorly differentiated	0.67 [0.35-1.21]	0.194	0.61 [0.29-1.24]	0.180
Unknown	0.48 [0.29-0.78]	0.003	0.74 [0.42-1.31]	0.611
Tumor size, mm	0.99 [0.99-1.00]	0.065	1.00 [0.99-1.00]	0.496
Treatment				
RFA	1		1	
Resection	2.82 [1.63-5.07]	<0.001	2.22 [1.15-4.39]	0.019
Liver transplantation	14.47 [8.66-25.44]	<0.001	12.62 [7.22-23.12]	<0.001
Fibrosis score (F1 vs F0)	0.85 [0.59-1.23]	0.379	0.58 [0.36-0.92]	0.02
Regional lymph node involvement (yes vs no)	2.50 [0.55-8.50]	0.172	2.08 [0.39-8.89]	0.345

Table 4: Association of treatments and 10-year survival among HCC patients using different analytic models.

Models	OR (95% CI)	*P* value
Univariate logistic regression model of the unadjusted cohort (model 1)		
LT vs RFA	14.47 [8.66-25.44]	<0.001
SR vs RFA	2.82 [1.63-5.07]	<0.001
LT vs SR	5.13 [3.44-7.76]	<0.001
Multivariable logistic regression model of the unadjusted cohort (model 2)		
LT vs RFA	12.62 [7.22-23.12]	<0.001
SR vs RFA	2.22 [1.15-4.39]	0.019
LT vs SR	5.69 [3.38-9.80]	<0.001
IPTW-adjusted (model 3)		
LT vs RFA	1.34 [1.25-1.45]	<0.001
SR vs RFA	1.05 [1.00-1.10]	0.066
LT vs SR	1.28 [1.18-1.40]	<0.001
IPTW and adjusted for age, tumor size, and fibrosis score (model 4)		
LT vs RFA	1.34 [1.24-1.45]	<0.001
SR vs RFA	1.04 [0.99-1.10]	0.138
LT vs SR	1.28 [1.17-1.41]	<0.001

[18]. These results support RFA as first-line treatment for early HCC, especially for patients with ≤50 mm tumor as shown in the subgroup analysis (Table 5). Moreover, RFA is a safe treatment. Complications occurred in only 2.2% of the treatments [7]. Another study reported that the complication rate was 1.8% [19].

Although comparable outcomes between RFA and SR for small tumors (≤3 cm), RFA is associated with higher rates of

Table 5: Subgroup analysis according to the tumor size for 10-year survival outcomes between 3 treatments after reweighting by IPTW in each group.

Tumor size	Number of patients	Actual 10-year survival outcome		
		LT vs RFA	SR vs RFA	LT vs SR
≤30	613	1.35 [1.24-1.48] $P < 0.001$	1.00 [0.97-1.04] $P = 0.796$	1.35 [1.23-1.47] $P < 0.001$
30-50	378	1.25 [1.10-1.42] $P = 0.001$	1.12 [0.98-1.27] $P = 0.090$	1.12 [0.94-1.33] $P = 0.192$
>50	264	1.54 [1.14-2.07] $P = 0.005$	1.08 [1.03-1.13] $P = 0.001$	1.42 [1.05-1.91] $P = 0.022$

tumor recurrence and local disease progression, and the median 3- and 5-year survival rates were lower in this group [14]. In this study, the actual 10-year survival rate of SR (9.1%) was higher than RFA (3.4%). After IPTW, patients who underwent SR also showed better survival outcomes. The actual 10-year survival rate of SR and RFA was 8.4% and 3.8%, respectively. Therefore, surgical resection was still the preferred choice.

There are some limitations in this work. First, the variables supplied in the SEER databases were limited, and we could not get more detailed patient information like hepatitis and tumor location. Second, the database does not supply the actual fibrosis score, fibrosis score 0–4 means none to moderate fibrosis, and fibrosis score 5–6 means severe fibrosis or cirrhosis. Third, the database did not collect information on recurrence; thus, the impact of disease recurrence on patients' long-term survival could not be assessed. Fourth, the survival data was not restricted to cancer-related death; thus, patients in this study may die for other reasons.

Estimation of the observational link between different treatments and 10-year survival is sensitive to analytical methods. LT was the best treatment for early single HCC. Due to shortage of donors, SR and RFA both serve as effective treatments. No significant differences for 10-year survival were found between SR and RFA in the IPTW cohort especially for ≤50 mm tumor.

Data Availability

The analyzed data are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors have no conflicts of interest to declare.

Authors' Contributions

Fanyu Meng performed the conceptualization, supervision, data curation, and writing-original draft. Haoyun Zhang performed the conceptualization, methodology, formal analysis, and writing-original draft. Haiwen Peng contributed to the data curation, formal analysis, validation, and visualization. Shichun Lu contributed to the methodology, project administration, funding acquisition, and writing-review and editing. Fanyu Meng and Haoyun Zhang contributed equally to this work.

Acknowledgments

This paper was supported by the National Key R&D Program of China [Grant number 2017YFA0103003].

References

[1] H. B. el–Serag and K. L. Rudolph, "Hepatocellular carcinoma: epidemiology and molecular carcinogenesis," *Gastroenterology*, vol. 132, no. 7, pp. 2557–2576, 2007.

[2] Global Burden of Disease Cancer Collaboration, C. Fitzmaurice, C. Allen et al., "Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study," *JAMA Oncology*, vol. 3, no. 4, pp. 524–548, 2017.

[3] J. D. Yang, P. Hainaut, G. J. Gores, A. Amadou, A. Plymoth, and L. R. Roberts, "A global view of hepatocellular carcinoma: trends, risk, prevention and management," *Nature Reviews Gastroenterology & Hepatology*, vol. 16, no. 10, pp. 589–604, 2019.

[4] Z. L. Li, W. T. Yan, J. Zhang et al., "Identification of actual 10-year survival after hepatectomy of HBV-related hepatocellular carcinoma: a multicenter study," *Journal of Gastrointestinal Surgery*, vol. 23, no. 2, pp. 288–296, 2019.

[5] T. Ryu, Y. Takami, Y. Wada, T. Hara, S. Sasaki, and H. Saitsu, "Actual 10-year survival after surgical microwave ablation for hepatocellular carcinoma: a single-center experience in Japan," *Annals of Surgical Oncology*, vol. 26, no. 12, pp. 4126–4133, 2019.

[6] X. Li, L. Huang, and X. Leng, "Analysis of prognostic factors of more/equal to 10 years of survival for liver cancer patients after liver transplantation," *Journal of Cancer Research and Clinical Oncology*, vol. 144, no. 12, pp. 2465–2474, 2018.

[7] S. Shiina, R. Tateishi, T. Arano et al., "Radiofrequency ablation for hepatocellular carcinoma: 10-year outcome and prognostic factors," *The American Journal of Gastroenterology*, vol. 107, no. 4, pp. 569–577, 2012, quiz 78.

[8] A. M. Gluer, N. Cocco, J. M. Laurence et al., "Systematic review of actual 10-year survival following resection for hepatocellular carcinoma," *HPB*, vol. 14, no. 5, pp. 285–290, 2012.

[9] B. Franssen, G. Jibara, P. Tabrizian, M. E. Schwartz, and S. Roayaie, "Actual 10-year survival following hepatectomy for hepatocellular carcinoma," *Gastroenterology*, vol. 16, no. 9, pp. 830–835, 2014.

[10] J. Bruix, M. Sherman, and American Association for the Study of Liver Diseases, "Management of hepatocellular carcinoma: an update," *Hepatology*, vol. 53, no. 3, pp. 1020–1022, 2011.

[11] S. K. Kamarajah, T. L. Frankel, C. Sonnenday, C. S. Cho, and H. Nathan, "Critical evaluation of the American joint

Hindawi
BioMed Research International
Volume 2023, Article ID 9832720, 1 page
https://doi.org/10.1155/2023/9832720

Retraction

Retracted: A Nomogram Predicting the Prognosis of Renal Cell Carcinoma Patients with Lung Metastases

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] X. Sheng, X. Lu, J. Wu, L. Chen, and H. Cao, "A Nomogram Predicting the Prognosis of Renal Cell Carcinoma Patients with Lung Metastases," *BioMed Research International*, vol. 2021, Article ID 6627562, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6627562, 11 pages
https://doi.org/10.1155/2021/6627562

Research Article

A Nomogram Predicting the Prognosis of Renal Cell Carcinoma Patients with Lung Metastases

Xinyu Sheng,[1,2] **Xuan Lu**,[1,2] **Jian Wu**,[1,2] **Lu Chen**,[1,2] **and Hongcui Cao**[1,2,3]

[1]*State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, 79 Qingchun Rd, Hangzhou City 310003, China*
[2]*National Clinical Research Center for Infectious Diseases, 79 Qingchun Rd, Hangzhou City 310003, China*
[3]*Zhejiang Provincial Key Laboratory for Diagnosis and Treatment of Aging and Physic-chemical Injury Diseases, 79 Qingchun Rd, Hangzhou City 310003, China*

Correspondence should be addressed to Hongcui Cao; hccao@zju.edu.cn

Received 3 December 2020; Revised 23 February 2021; Accepted 6 March 2021; Published 19 March 2021

Academic Editor: Min Tang

Background. The optimal tool for predicting the survival of renal cell carcinoma (RCC) patients with lung metastases remains controversial. *Methods*. We selected patients diagnosed with RCC and lung metastases, from 2010 to 2015, from the Surveillance, Epidemiology, and End Results (SEER) database. After the selection of inclusion criteria and exclusion criterion, the rest of the patients were incorporated into model analysis. Least absolute shrinkage and selection operator (LASSO) regression was used to select the most important features for construction of a nomogram predicting cancer-specific survival. A calibration plot and the concordance index (*C*-index) were used to estimate nomogram efficacy in a validation cohort. The association between important factors selected by LASSO regression, and prognosis was assessed by the Kaplan-Meier (KM) survival curve. The receiver operating characteristic (ROC) curves were drawn to compare sensitivity and specificity between the nomogram we built and the TNM stage-based model. *Results*. A total of 1,369 patients met the inclusion criteria, but not the exclusion criteria. The LASSO regression model reduced 15 features to seven potential predictors of survival, including tumor grade, the extent of surgery, N and T status, histological profile, and brain and bone metastasis status. Such features had good discrimination in the KM survival curves. The nomogram showed excellent discriminatory power (*C*-index, 0.71; 95% confidence interval: 0.70 to 0.72) and good calibration in terms of both 1- and 2-year cancer-specific survival. The nomogram showed great discriminatory power (*C*-index 0.68) and adequate calibration when applied to the validation cohort. The areas under the curve (AUCs) of nomogram were 0.767 and 0.780, respectively, and the AUCs of TNM stage were 0.617 and 0.618 at 1 and 2 years, respectively. *Conclusions*. Our nomogram might play a major role in predicting the cancer-specific survival of RCC patients with lung metastases.

1. Introduction

Renal cell carcinoma (RCC), a common malignant tumor, accounts for 3.7% of all new tumor cases; RCC is more common in male than female patients, and there are 116,000 deaths annually according to the World Health Organization. The most common pathological type (85%) of RCC is clear cell carcinoma [1, 2]. It is important to accurately predict the survival of RCC patients; however, few efficient predictive tools are available [3–5], particularly for patients with metastatic RCC [6] (which remains incurable). Traditionally, immunochemotherapy has been used to treat nonmetastatic RCC, but the response rate was rarely more than 15%. The clinical prognosis of metastatic RCC is poor, and more studies are required [7]. Lung metastasis is the most common metastasis in RCC patients, but few studies have explored the survival of these patients. An accurate predictive tool would improve tumor control and patient quality of life.

The Surveillance, Epidemiology, and End Results (SEER) database (supported by the American National Cancer Institute) contains clinical and pathological data on cancer patients from 19 regions of the United States and is highly

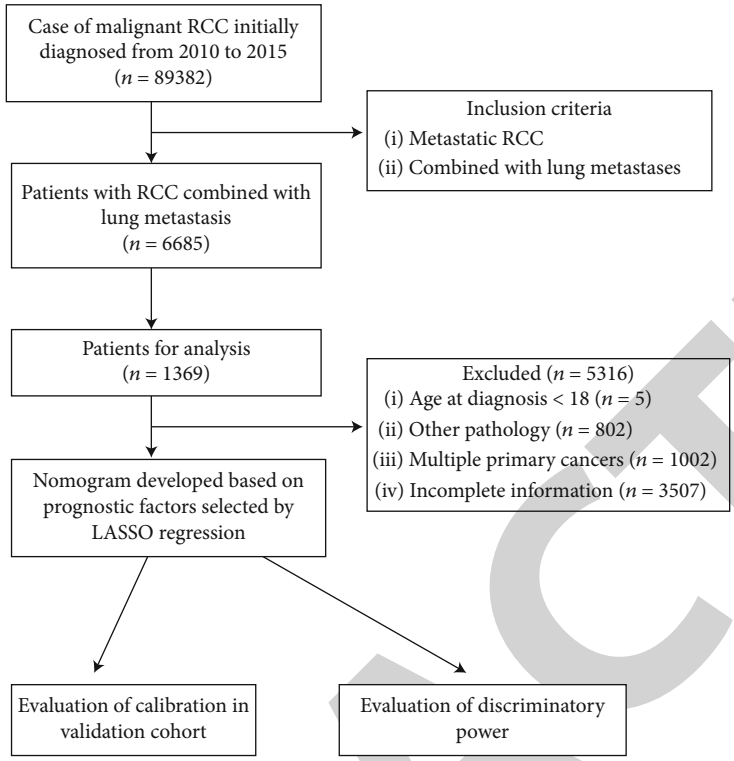


Figure 1: Flow chart of patient selection.

representative of the general population. Compared to most previous studies which had small sample sizes, our research relied on the SEER database enhances the credibility of this study.

A nomogram is a readily understood visual tool that transforms a complex regression equation to a simple graph; predictions are thus accessible and of high clinical utility. Nomograms find applications in both medical research and clinical practice [8]. The tumor-node-metastasis (TNM) staging system for RCC has been widely used to predict prognosis, but the clinical utility remains unclear [9–11]. Therefore, we developed a nomogram for predicting the cancer-specific survival (CSS) of RCC patients with lung metastases and compared its ROC curves to that of a TNM staging system-based model. Unlike most previous studies, we used a large patient sample that had been extensively evaluated and documented.

2. Materials and Methods

2.1. Patients. Data were obtained from the SEER database. The SEER database covers almost 28% of the US population and is the largest cancer database in the US. All cases are derived from the SEER Program (http://www.seer.cancer.gov) SEER*Stat database (version 8.3.5, accession number: 12099-Nov2018). Abundant information is available, including patient demographics and cancer characteristics. A total of 89,382 patients (including 12,187 stage M1 patients) were newly diagnosed with RCC from 2010 to 2015. The inclusion criteria were as follows: (1) RCC patients from 2010 to 2015 in the US and (2) combined with lung metastases. Thus, we included 1,369 RCC patients with lung metastases in the analysis (Figure 1). Exclusion criteria were as follows: history of other tumors (begin or malignant), age<18, and rare pathological features and data with incomplete information. The following five pathological types of RCC accounted for more than 90% of all cancers: clear cell, papillary, chromophobe, and sarcomatoid RCCs and collecting duct carcinoma (CDC). We excluded patients with other types of carcinoma.

2.2. Clinicopathological Features. The following data were collected: years at diagnosis and evaluation, age, race, sex, extent of surgery (none, partial nephrectomy (PN), or radical nephrectomy (RN)), tumor histology, histological grade, nodal (N) stage, insurance and marital status, laterality, tumor size, CSS, and follow-up duration. There were four histological grades (I–IV). Marital status was classified as "unmarried" (including "single (never married)" and "unmarried"), "married" (including "married under common law"), and "no longer married" ("widowed", "separated", and "divorced"). The median survival time was 10 months, and the maximum survival time was 70 months.

2.3. Survival Analysis and Statistical Methods. The primary outcome was the survival time (from hospital admission to the date of death or last follow-up). Clinicopathological features are summarized as median values with interquartile range (IQR), or as frequencies with percentages, for both the training and validation cohorts. The associations of the factors selected from the least absolute shrinkage and selection operator (LASSO) regression analysis with CSS were evaluated by drawing Kaplan-Meier curves. A nomogram

Table 1: Clinicopathological characteristics of RCC patients with lung metastases.

Characteristic	Primary cohort ($n = 976$)		Validation cohort ($n = 393$)		P value
	No. of patients	%	No. of patients	%	
Age (years)					0.72
Median	60		60		
Range	54-68		54-68		
Sex					0.252
Male	705	72.2	271	69	
Female	271	27.8	122	31	
Race					0.937
White	831	85.1	335	85.2	
Black	64	6.6	24	6.1	
Other	81	8.3	34	8.7	
Grade					0.928
I	29	3	10	2.5	
II	188	19.3	77	19.6	
III	405	41.5	158	40.2	
IV	354	36.3	148	37.7	
Laterality					1
Left	480	49.2	193	49.1	
Right	496	50.8	200	50.9	
T					0.859
T1a	19	1.9	6	1.5	
T1b	82	8.4	29	7.4	
T2a	89	9.1	43	10.9	
T2b	83	8.5	36	9.2	
T3	572	58.6	224	57	
T4	131	13.4	55	14	
N					0.875
N1	304	31.1	120	30.5	
N0	672	68.9	273	69.5	
Surgery					0.659
No surgery	231	23.7	84	21.4	
PN	19	1.9	8	2	
RN	726	74.4	301	76.6	
Tumor size (mm)					0.356
Median	95		100		
Range	75-120		75-120		
Pathological type					0.582
pRCC	28	2.9	12	3.1	
ccRCC	697	71.4	288	73.3	
RCC, NOS	168	17.2	53	13.5	
sRCC	8	0.8	3	0.8	
chRCC	63	6.5	30	7.6	
CDC	12	1.2	7	1.8	
Brain metastasis					0.421
Yes	116	11.9	40	10.2	
No	860	88.1	353	89.8	
Bone metastasis					<0.001
Yes	116	11.9	78	19.8	
No	860	88.1	315	80.2	

Table 1: Continued.

Characteristic	Primary cohort ($n = 976$) No. of patients	%	Validation cohort ($n = 393$) No. of patients	%	P value
Liver metastasis					0.523
Yes	131	13.4	47	12	
No	845	86.6	346	88	
Insurance status					0.93
Yes	936	95.9	378	96.2	
No	40	4.1	15	3.8	
Marital status					0.115
Unmarried	155	15.9	57	14.5	
Married	642	65.8	280	71.2	
Negative married	179	18.3	56	14.2	

Abbreviations: ccRCC: clear cell renal cell carcinoma; pRCC: papillary renal cell carcinoma; chRCC: chromophobe renal cell carcinoma; sRCC: sarcomatoid renal cell carcinoma; CDC: collecting duct carcinoma.

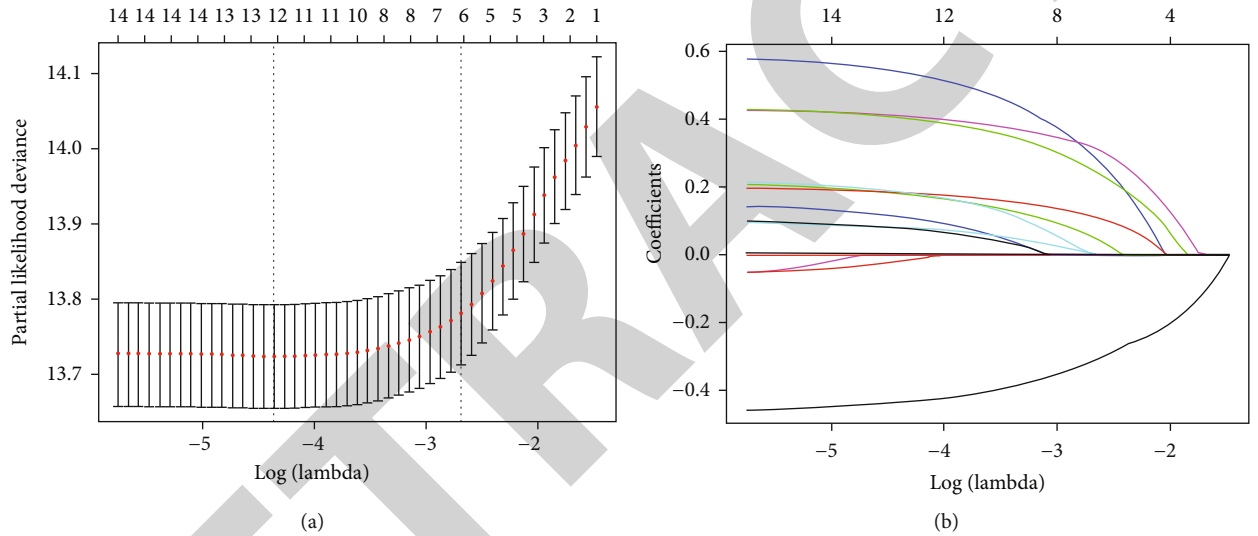


Figure 2: Feature selection using a least absolute shrinkage and selection operator (LASSO) regression model. (a) Tuning parameter (λ) selection for the LASSO model involved fivefold cross-validation using the minimal criteria. Dotted vertical lines were drawn at the optimal values; one standard error was added to each criterion to yield the 1-SE criteria. (b) The LASSO coefficients of the 15 features.

was created based on the results of the LASSO regression. Discriminatory power was quantified using the concordance index (C-index). The "rms" R package was used to evaluate the nomogram and draw calibration curves. All statistical analyses were performed using R software (ver. 3.5.3; R Foundation for Statistical Computing, Vienna, Austria). All tests were two-sided and a P value < 0.05 was considered statistical significance.

3. Results

3.1. Clinicopathological Characteristics. Of the 6,685 RCC patients who developed lung metastases from 2010 to 2015, 1,369 were included in the final analysis (976 in the training cohort and 393 in the validation cohort; Table 1). We considered the three major types of metastases (the brain, bone, and liver). The gross pathological types were clear cell, papillary, chromophobe, and sarcomatoid RCCs and CDC. The median age of both cohorts was 60 years, where RCC is more common in the elder. The median follow-up time was 10 months (range: 1–70 months) and the 1- and 2-year CSS rates were 53.9 and 34.8%, respectively.

3.2. Feature Selection and Prognostic Signature Building. We reduced the initial 15 features of the 976 patients in the training cohort to seven potential predictors of survival: T status (coefficient, 0.106) N status (coefficient, 0.440), tumor grade (coefficient, 0.211), extent of surgery (coefficient, -0.490), pathology (coefficient, 0.120), and brain and bone metastasis status (coefficient, 0.581 and 0.464) (Figure 2). Feature selection was performed using a LASSO binary logistic regression model. Dotted vertical lines were drawn at the optimal values (derived using the minimal criteria with one standard error (the "1-SE" criteria), see Figure 1). The final λ value was

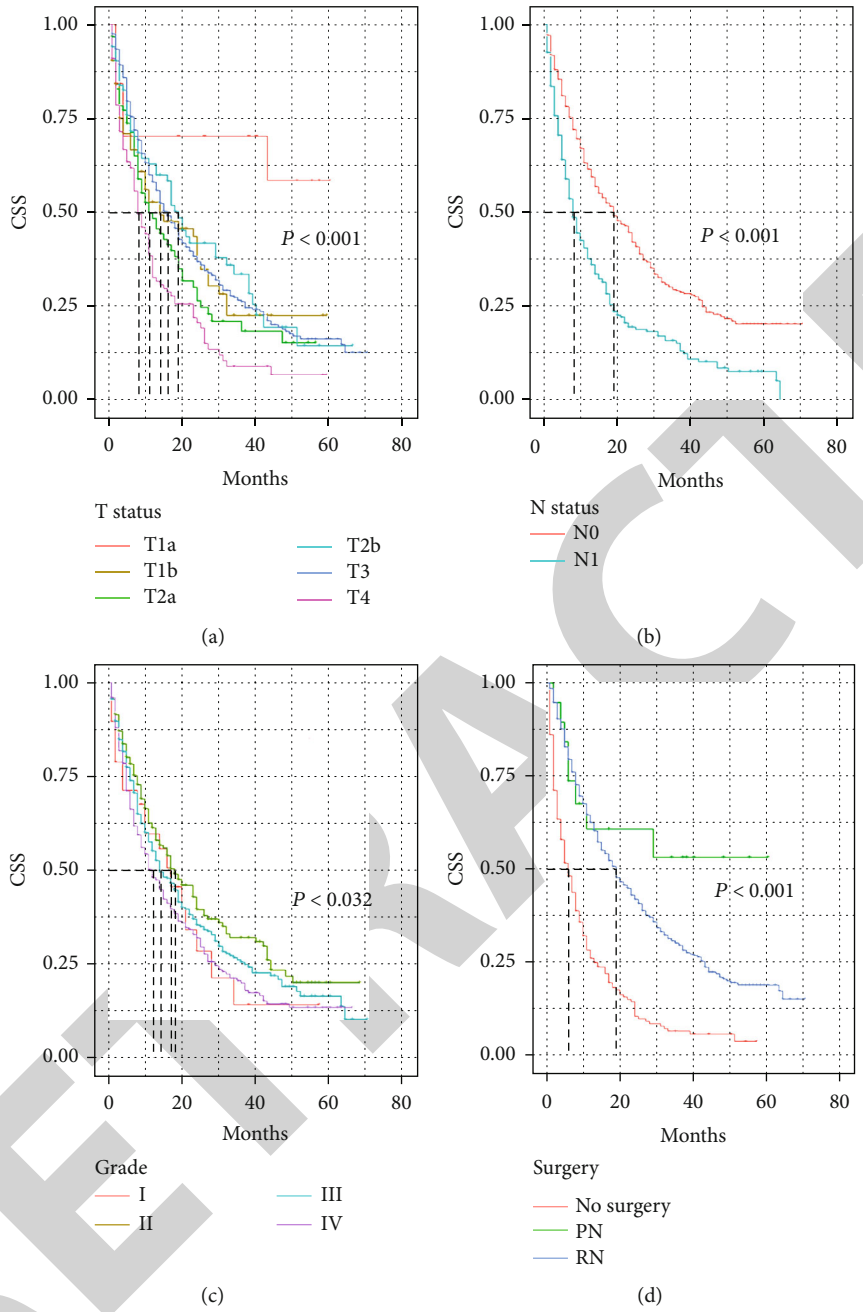


Figure 3: Continued.

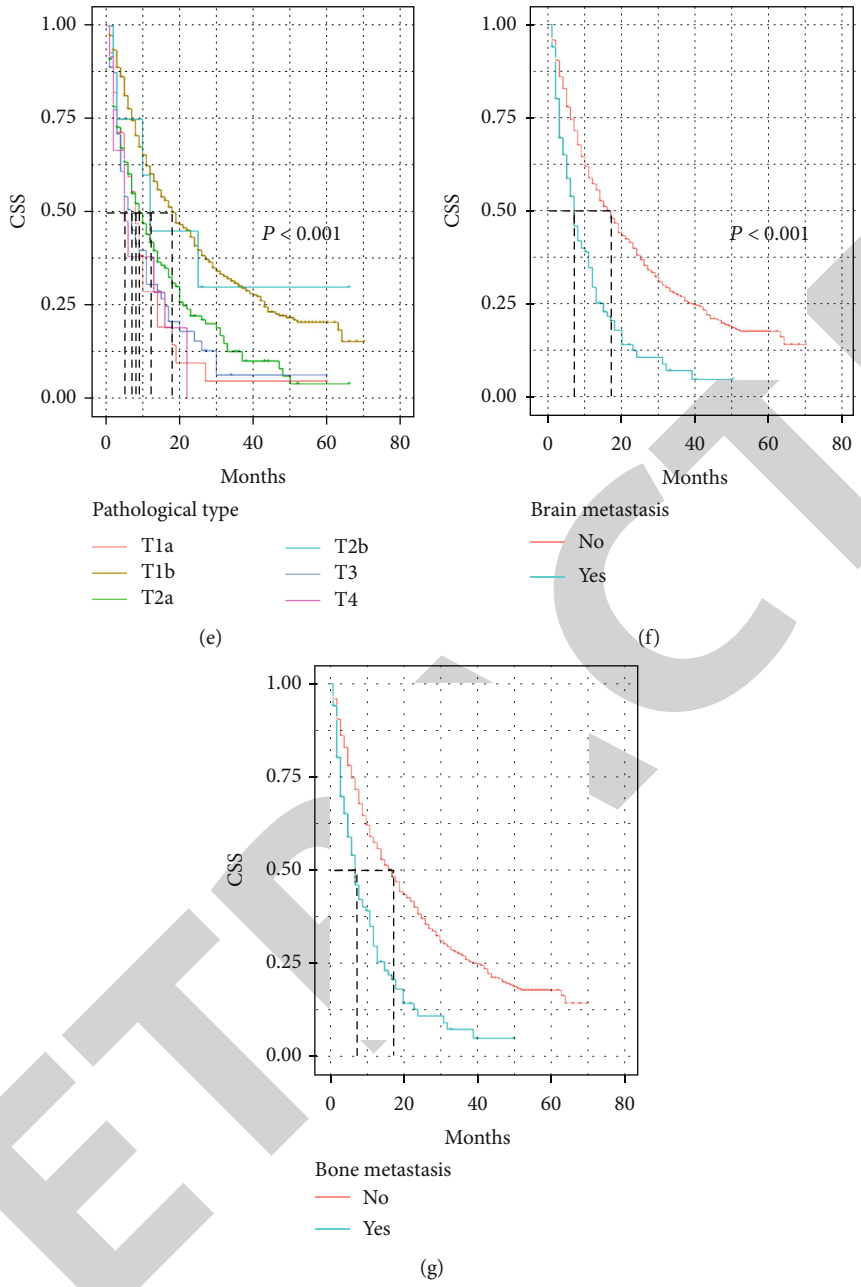


Figure 3: Kaplan-Meier curves of selected features. Seven features were included in the final model. The cancer-specific survival of renal cell carcinoma (RCC) patients with lung metastases was predicted by T status (a), N status (b), tumor grade (c), extent of surgery (d), pathology (e), brain metastasis status (f), and bone metastasis status (g).

0.068 and the log (λ) value was -2.687. We included the coefficient profiles of all 15 features in the diagram.

3.3. Kaplan-Meier Curves of Significant Features. All seven potential predictors of survival were found to be useful for predicting CSS in the training cohort (Figure 3). It was clear that patients lacking metastases had better CSS, which was also significantly enhanced by PN or RN. Patients receiving operation (either PN or RN) had a better CSS than those who did not. Not surprisingly, lymph node metastasis could not afford a significant survival benefit. The significant differences in tumor grade and histology among RCC types also affected the CSS outcomes.

3.4. A Prognostic Nomogram for CSS. The prognostic nomogram comprising all factors that significantly affected CSS is shown in Figure 4. The *C*-index for CSS prediction was 0.71 (95% confidence interval: 0.70 to 0.72). The calibration plots for the 1- and 2-year survival probabilities of RCC patients with lung metastases revealed excellent agreement between prediction and observation in training cohort (Figures 5(a) and 5(b)).

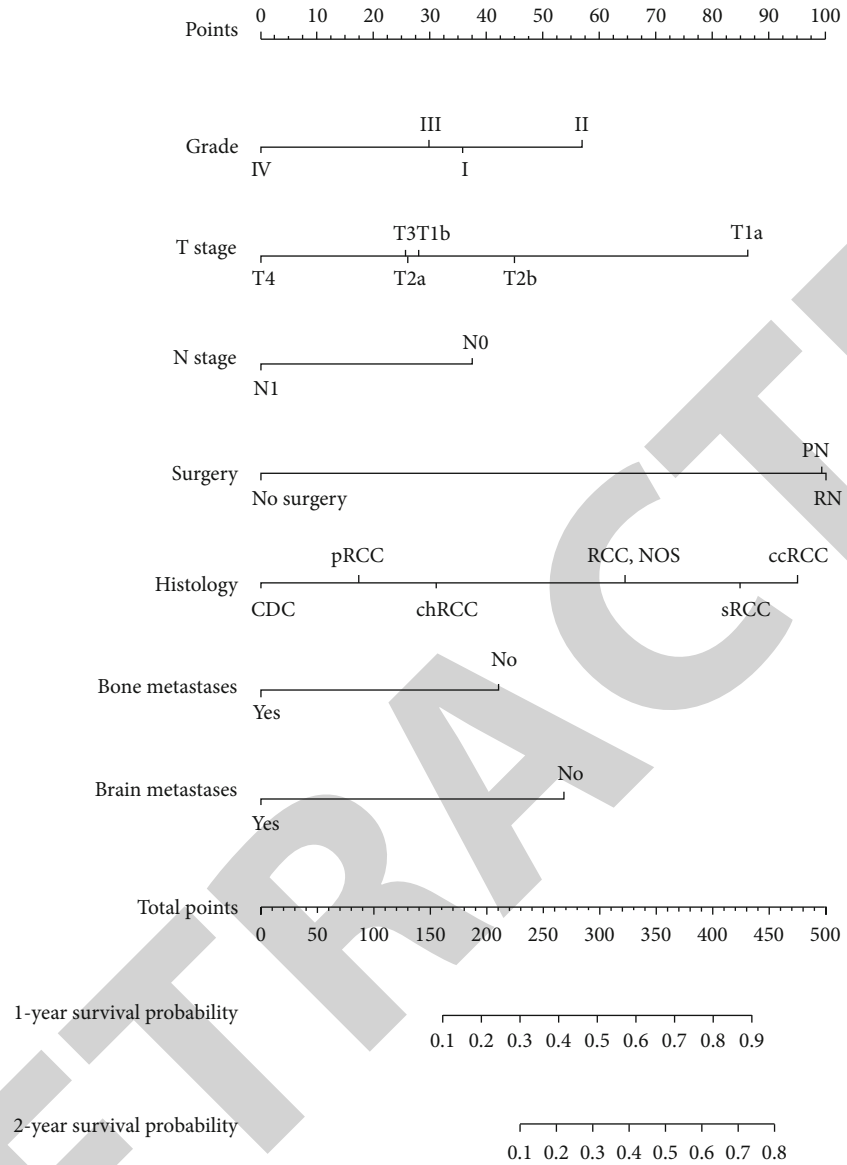


Figure 4: The prognostic nomogram. The nomogram was developed using the training cohort. The seven features selected by LASSO regression included T status, N status, tumor grade, extent of surgery, pathology, brain metastasis status, and bone metastasis status. The nomogram accurately predicted 1- and 2-year survival.

3.5. Validation of Predictive Accuracy. The follow-up time in the validation cohort was 10 months (range: 1–67 months) and the 1- and 2-year CSS rates were 56.3 and 43.1%, respectively. The *C*-index of the nomogram for CSS prediction was 0.68. The calibration plots of the survival probability of the validation cohort, for 1 and 2 years after RCC diagnosis, revealed good agreement between prediction and observation.

3.6. ROC Comparison with TNM Stage. We then applied the time-dependent ROC curves to compare sensitivity and specificity between the nomogram we built and the TNM stage at 1 and 2 years. The areas under the curve (AUCs) of nomogram were 0.767 and 0.780, respectively, and the AUCs of TNM stage were 0.617 and 0.618 at 1 and 2 years, respectively (Figure 6).

4. Discussion

In a LASSO regression, seven prognostic factors for RCC patients with lung metastases were identified, based on which a nomogram was constructed including tumor grade, extent of surgery, T and N status, histology, and brain and bone metastasis status. This was used to predict 1- and 2-year CSS in a large cohort of metastatic RCC patients drawn from the SEER database. The nomogram was validated, internally and externally, in terms of its discriminatory power and calibration.

There exist several outcome prediction models for RCC patients, though none of them is developed for patients with lung metastases. Due to the natural differences in tumor microenvironment, models set for metastatic and nonmetastatic RCC are different. The stage, size, grade, and necrosis

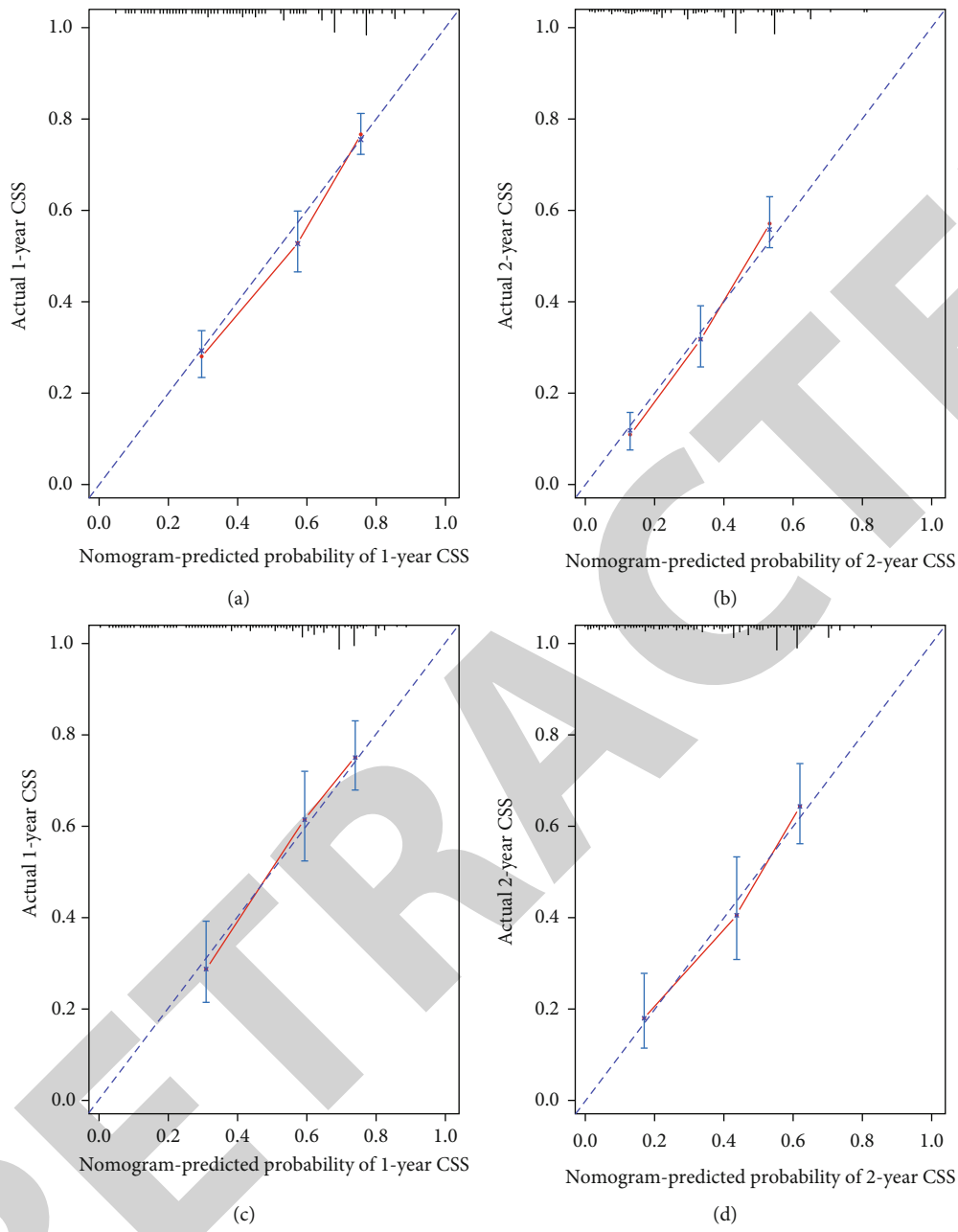


Figure 5: Calibration curves of the prognostic nomogram for each cohort. (a) Calibration curve of the nomogram for the training cohort at 1 year. (b) Calibration curve of the nomogram for the training cohort at 2 years. (c) Calibration curve of the nomogram for the validation cohort at 1 year. (d) Calibration curve of the nomogram for the validation cohort at 2 years.

(SSIGN) score and the University of California Los Angeles Integrated Staging System (UISS) are two widely used tools for prognosis prediction [12, 13]. SSIGN score is calculated to divide localized RCC patients into three groups and to predict 5-year metastasis-free survival rate. UISS can be used to predict 5-year disease-specific survival for both metastatic and nonmetastatic RCC patients, which includes T stage, Fuhrman's grade, and ECOG (Eastern Cooperative Oncology Group) status. Besides the two common tools for prediction, other three prognostic models are performed in the previous studies [14–16].

As for metastatic RCC, prediction tools are used for either selecting those patients who might benefit from adjuvant treatment or predicting survival rate. The patients care more about their survival time. Previous studies have launched some prediction models based on demographic and tumor pathology information [17, 18]. However, most of them just proposed prognostic factors, rather than integrating them into scores and presenting them in a concise form.

The effect of lung metastases on prognosis in RCC patients remained controversial. A research with 782 patients from Groupe Français d'Immunothérapie took diverse

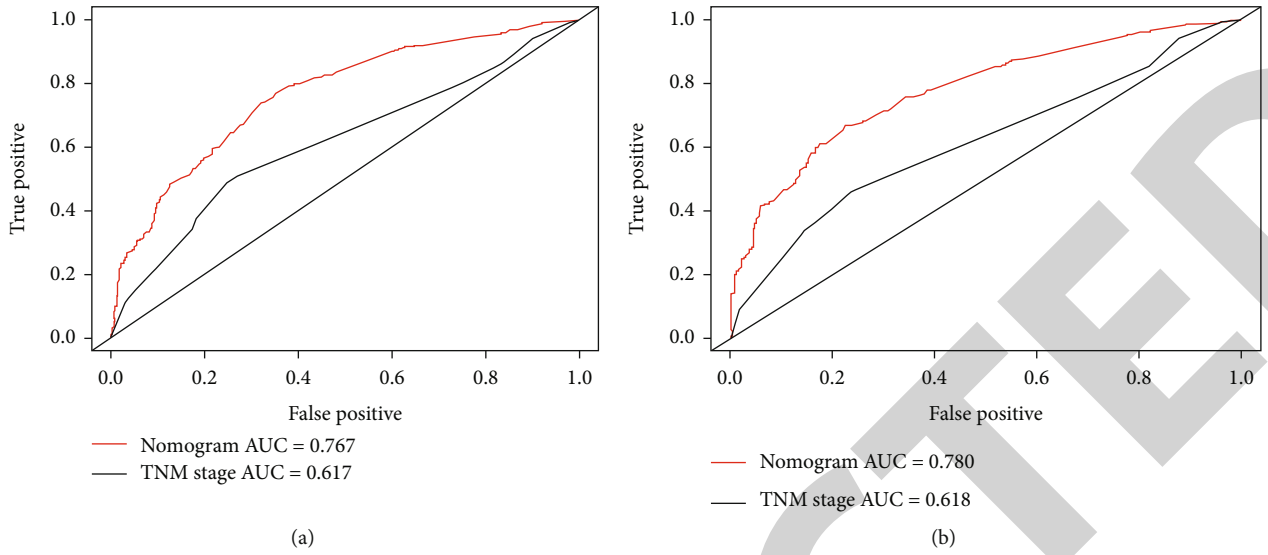


Figure 6: Time-dependent receiver operating characteristic (ROC) curves of the nomogram score and model based on TNM stage for CSS of RCC patients with lung metastases. (a) ROC at 1 year. The areas under the curve (AUCs) of the nomogram and TNM stage were 0.767 and 0.617, respectively. (b) ROC at 2 year. The areas under the curve (AUCs) of the nomogram and TNM stage were 0.780 and 0.618, respectively.

metastatic sites including the lung, mediastinum, bone, and liver into consideration, and the results showed that only the lung metastasis was not associated with overall survival [19]. Similarly, a retrospective study showed that the overall survival did not differ between patients with less than 8.5 months and more than 8.5 months [20]. However, as lung metastasis is the most common metastasis in RCC patients, it is hard to say that it has no effect on the overall survival of RCC. This can be corroborated by a study on adult metastatic RCC that lung metastasis together with weight loss and disease-free interval resulted in a worse survival [21]. Also, although there were several prognostic tools for RCC patients with lung metastasis, none of them was developed for RCC patients with lung metastases, which is exactly what we want to do in this study.

Currently, only few dedicated models predict the overall survival (OS) or CSS of RCC patients, particularly those with metastases. A model based on the TNM staging system is currently the default clinical model, and it evaluates tumor size and local extension of the primary tumor (T), regional lymph node involvement (N), and distant metastasis status (M) [22, 23]. Because we evaluated metastatic RCC patients, T and N status were retained in our nomogram. Although the TNM-based model is widely used to predict the survival of RCC patients, it is unclear whether the model is appropriate. A few studies found that models based on the TNM staging system were inaccurate in terms of predicting the RCC survival rate [24]. One study even reported no association between TNM staging system data and RCC patient outcomes. A study of more than 500 nonmetastatic RCC patients found that the prognostic utility of the TNM staging system was limited [25]. Few studies have focused on metastatic RCC; it is essential to establish a novel tool for predicting the outcomes (especially survival rates) of metastatic RCC patients. Most RCC metastases are in the lungs; these patients require special attention. The TNM staging system is clearly biased, emphasizing pathological indicators over other, such as clinical and treatment characteristics, which are more important from our perspectives.

We searched all the metastatic RCC prognostic models and compiled them into Table S1 in the Supplementary Material. Among these models, "performance status," "signs of inflammation," and some biochemical indicators including ALP, calcium, LDH, hemoglobin, and WBC count were not included in the SEER database, which is one limitation of analysis. However, although these models were developed for a while, none of them was compared with the traditional TNM staging system, which is worth to be taken into consideration. The nomogram we developed exhibited great discriminatory power and the CSS of the training and validation cohorts were well-calibrated. It proved that this prognostic model is stable and convincing. Also, we concerned the two types of surgical methods, PN and RN, since they might affect the patients' perioperative state and therefore affect the survival of RCC patients with lung metastases.

A nomogram is an intuitive visual tool for predicting disease survival and can be used to explain survival probability to patients in a straightforward manner. Many authors now use nomograms to predict disease (including RCC) survival rates [26, 27]. Nomograms containing more information than the TNM staging system; they better predict the outcomes of patients with pancreatic and hepatic carcinoma [28, 29] and RCC. We considered seven factors including pathological type and clinical and treatment characteristics, to minimize bias when predicting prognosis.

The strengths of our study were as follows. First, our nomogram is the only tools effective for predicting the CSS of RCC patients with lung metastases among existing models. Indeed, it is an effective bridge connecting doctors and patients. Second, we focused on a specialized cohort, i.e., RCC patients with lung metastases (the most common

metastasis in these patients). Third, our study is a good representative of the general population. The SEER database we used covers the entire United States and is thus much larger and more comprehensive than single-center databases; this feature enhances the credibility of our study.

This study also had some limitations. First, our research lacks external validation, though the SEER database compensates for the homogeneity of patients from single research center to some extent as it covers millions of patients all around US. Second, as the study was retrospective in nature, patient allocation to the training and validation cohorts may have been biased; randomized control trials are needed. Third, some biochemical indicators including ALP, calcium, LDH, hemoglobin, and WBC count were not included in the SEER database and the discriminatory power of our nomogram could not be compared with other prognostic models. Also, a larger cohort would have allowed us to detect small differences between patient subgroups. Finally, unknown factors may affect the CSS of RCC patients with lung metastases. More data are needed to confirm the effectiveness of the nomogram; to this end, a prospective study is planned.

5. Conclusion

The novel prognostic model based on the SEER database could predict the cancer-specific survival of RCC patients with lung metastases.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

No author has any conflict of interest in terms of the subject matter or the materials discussed in this paper.

Acknowledgments

We thank the SEER database owners for giving us the opportunity to analyze some of their data.

Supplementary Materials

Supplementary Table S1: characteristics in various prognostic models for metastatic RCC. *(Supplementary Materials)*

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2017," *CA: a Cancer Journal for Clinicians*, vol. 67, no. 1, pp. 7–30, 2017.

[2] U. Capitanio, K. Bensalah, A. Bex et al., "Epidemiology of renal cell carcinoma," *European Urology*, vol. 75, no. 1, pp. 74–84, 2019.

[3] G. Zhang, Y. Wu, J. Zhang et al., "Nomograms for predicting long-term overall survival and disease-specific survival of patients with clear cell renal cell carcinoma," *Oncotargets and Therapy*, vol. Volume 11, pp. 5535–5544, 2018.

[4] S. Buti, P. I. Karakiewicz, M. Bersanelli et al., "Validation of the GRade, Age, Nodes and Tumor (GRANT) score within the Surveillance Epidemiology and End Results (SEER) database: a new tool to predict survival in surgically treated renal cell carcinoma patients," *Scientific Reports*, vol. 9, no. 1, p. 13218, 2019.

[5] W. Zhou, C. Huang, and N. Yuan, "Prognostic nomograms based on log odds of positive lymph nodes for patients with renal cell carcinoma: a retrospective cohort study," *International journal of surgery (London, England)*, vol. 60, pp. 28–40, 2018.

[6] W. J. Xiao, Y. Zhu, B. Dai, H. L. Zhang, and D. W. Ye, "Assessment of survival of patients with metastatic clear cell renal cell carcinoma after radical cytoreductive nephrectomy versus no surgery: a seer analysis," *International braz j urol : official journal of the Brazilian Society of Urology*, vol. 41, no. 2, pp. 288–295, 2015.

[7] G. Rosiello, S. Knipper, C. Palumbo et al., "Unmarried status is a barrier for access to treatment in patients with metastatic renal cell carcinoma," *International Urology and Nephrology*, vol. 51, no. 12, pp. 2181–2188, 2019.

[8] A. Iasonos, D. Schrag, G. V. Raj, and K. S. Panageas, "How to build and interpret a nomogram for cancer prognosis," *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, vol. 26, no. 8, pp. 1364–1370, 2008.

[9] M. Sun, S. F. Shariat, C. Cheng et al., "Prognostic factors and predictive models in renal cell carcinoma: a contemporary review," *European Urology*, vol. 60, no. 4, pp. 644–661, 2011.

[10] B. Ljungberg, K. Bensalah, S. Canfield et al., "EAU guidelines on renal cell carcinoma: 2014 update," *European Urology*, vol. 67, no. 5, pp. 913–924, 2015.

[11] V. Ficarra, G. Novara, A. Galfano, and W. Artibani, "Neoplasm staging and organ-confined renal cell carcinoma: a systematic review," *European Urology*, vol. 46, no. 5, pp. 559–564, 2004.

[12] B. C. Leibovich, M. L. Blute, J. C. Cheville et al., "Prediction of progression after radical nephrectomy for patients with clear cell renal cell carcinoma: a stratification tool for prospective clinical trials," *Cancer*, vol. 97, no. 7, pp. 1663–1671, 2003.

[13] A. Zisman, A. J. Pantuck, F. Dorey et al., "Improved prognostication of renal cell carcinoma using an integrated staging system," *Journal of Clinical Oncology*, vol. 19, no. 6, pp. 1649–1657, 2001.

[14] P. I. Karakiewicz, A. Briganti, F. K.-H. Chun et al., "Multi-institutional validation of a new renal cancer–specific survival nomogram," *Journal of Clinical Oncology*, vol. 25, no. 11, pp. 1316–1322, 2007.

[15] M. I. C. H. A. E. L. W. KATTAN, V. I. C. T. O. R. REUTER, R. O. B. E. R. T. J. MOTZER, J. A. R. E. D. KATZ, and P. A. U. L. RUSSO, "A postoperative prognostic nomogram for renal cell carcinoma," *The Journal of Urology*, vol. 166, no. 1, pp. 63–67, 2001.

[16] M. A. X. I. M. I. L. I. A. N. O. SORBELLINI, M. I. C. H. A. E. L. W. KATTAN, M. A. R. K. E. SNYDER et al., "A postoperative prognostic nomogram predicting recurrence for patients with conventional clear cell renal cell carcinoma," *The Journal of Urology*, vol. 173, no. 1, pp. 48–51, 2005.

[17] R. J. Motzer, M. Mazumdar, J. Bacik, W. Berg, A. Amsterdam, and J. Ferrara, "Survival and prognostic stratification of 670 patients with advanced renal cell carcinoma," *Journal of Clinical Oncology*, vol. 17, no. 8, pp. 2530–2540, 1999.

Hindawi
BioMed Research International
Volume 2023, Article ID 9761369, 1 page
https://doi.org/10.1155/2023/9761369

Retraction

Retracted: A novel TUBG1 mutation with neurodevelopmental disorder caused by malformations of cortical development

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] R. Shen, Z. Zhang, Y. Zhuang, X. Yang, and L. Duan, "A novel *TUBG1* mutation with neurodevelopmental disorder caused by malformations of cortical development," *BioMed Research International*, vol. 2021, Article ID 6644274, 8 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6644274, 8 pages
https://doi.org/10.1155/2021/6644274

Research Article

A novel *TUBG1* mutation with neurodevelopmental disorder caused by malformations of cortical development

Ru Shen,[1] **Zhen Zhang**,[2] **Yu Zhuang**,[1] **Xiaohong Yang**,[1] **and Lifen Duan**[3]

[1]*Division of Laboratory, Kunming Children's Hospital affiliated with Kunming Medical University, Kunming 650028, China*
[2]*Institute of Pediatrics, The Kunming Children's Hospital, Key Laboratory of Child Critical Disease Research of Yunnan Province, Kunming 650028, China*
[3]*Division of Neurology, Kunming Children's Hospital affiliated with Kunming Medical University, Kunming 650028, China*

Correspondence should be addressed to Lifen Duan; duanlifen@etyy.cn

Received 16 December 2020; Revised 17 January 2021; Accepted 8 February 2021; Published 26 February 2021

Academic Editor: Min Tang

Neurodevelopmental disorder caused by malformations of cortical development is a rare neurological disease. Heterozygous missense variants in the *TUBG1* gene lead to malformations of human cortical development, which further result in intellectual disability, developmental retardation, and epilepsy. To the best of our knowledge, only thirteen patients and a total of nine pathogenic *TUBG1* variants have been described in the published literature. This study reports the case details and genetic data analysis of a girl (aged 8 years, 9 months) with developmental delay, psychomotor regression, epilepsy, and left external ear deformity. A novel *TUBG1* mutation was identified by whole-exome sequencing and Sanger sequencing, confirming that this mutation may be the cause of the neurodevelopmental disorders. This case report characterizes the phenotypic spectrum, molecular genetic findings, and functional consequences of novel pathogenic *TUBG1* variants in neurodevelopmental disorders caused by cortical development malformations.

1. Introduction

Neurodevelopmental disorders caused by malformations of cortical development (MCD) are rare diseases with high clinical heterogeneity. The main clinical features of MCD include motor impairment, intellectual disability, early-onset epilepsy, and cortical malformations (including lissencephaly, polymicrogyria, microlissencephaly, pachygyria, and simplified gyration) [1, 2]. There are many reasons for MCD such as KIF5C gene mutation [3]. However, the cortical and extracortical brain phenotypes observed are largely dependent on the specific tubulin gene affected.

Tubulin genes play a key role in several pathways of cortical development. Mutations in these genes may cause specific brain malformations strictly associated with epilepsy [4]. Mutations affecting seven genes encoding alpha- (*TUBA1A*), beta- (*TUBB2A*, *TUBB2B*, *TUBB3*, *TUBB4A*, and *TUBB*), and gamma-tubulin (*TUBG1*) have been identified in individuals with a range of malformations of cortical development (MCDs) [5].

In mammalian cells, γ-tubulin is highly conserved and often encoded by two genes. Humans have γ-tubulin 1 and γ-tubulin 2 *(TUBG1* and *TUBG2*) [6]. *TUBG1* is thought to be ubiquitously expressed, whereas *TUBG2* expression appears to be restricted to the brain [6, 7]. Mutations in the *TUBG1* gene located on chromosome 17q21 are one of the causes of neurodevelopmental disorders caused by MCD [8, 9]. *TUBG1* encodes γ-tubulin, which is highly expressed in the developing fetal brain as a component of centrosomes [8].

TUBG1 plays an integral role in microtubule nucleation, thereby affecting microtubule-dependent mitosis and brain development [10–12]. Poirier et al. introduced variants in the γ-tubulin gene that interfered with microtubule nucleation in *Saccharomyces cerevisiae*. The suppression of *TUBG1* in mice in utero also arrested neuronal migration [13]. Ivanova et al. studied seven MCD-related identified variants in *TUBG1*, and the study results suggest that disease-related *TUBG1* variants exert their pathogenicity by affecting microtubule dynamics rather than centrosomal positioning or nucleation ability [5].

TUBG1 mutations in neurodevelopmental disorders caused by MCD exhibit a relatively limited mutational spectrum [14]. Only nine pathogenic variants in thirteen cases of *TUBG1* mutations have been reported, and the main phenotypes are epilepsy combined with retardation of intelligence, language, and motor development. In this article, we describe a novel pathogenic variant of the *TUBG1* gene (c.751A>T p.N251Y) identified through whole-exome sequencing (WES) in a Chinese patient who presented epilepsy, intellectual disability, speech impairment, and left external ear deformity; these findings may serve to improve diagnosis, management, and research on neurodevelopmental disorders caused by MCD.

2. Materials and Methods

2.1. General Patient Information and Ethical Considerations. This study was reviewed by the Medical Ethics Committee of Kunming Children's Hospital affiliated with Kunming Medical University. Informed consent for genetic analysis was obtained from the parent. Clinical information and blood specimens were obtained from the patient and her family members.

The patient was a girl who was 8.75 years old and was the second child born to nonconsanguineous parents. She had a height of 1.17 meters, a weight of 18 kg, short stature, left auricle deformity, and pectus excavatum (Figure 1), and she had been diagnosed with epilepsy for 7 years. The patient experienced a relapse of epilepsy and came to the neurology outpatient department of our hospital. The parent complained that there was no obvious cause of the afebrile convulsions 5 times a day during the awake period when the child was 1 year old. The condition manifested as a sudden turning of the eyes, no response to calling, and limb twitching, (but no foaming at the mouth, or incontinence). Each episode lasted 10-20 minutes followed by relief, lethargy, and normal mental activity after wakening. The patient was diagnosed with epilepsy in other hospitals and was administered herbal medicine for more than 3 years to control seizures.

The patient had an attack every half month to a month while receiving medication. She was then treated with sodium valproate, topiramate, and levetiracetam. The drug dosage was unknown at that time. After taking the medicines for 3 years, she stopped taking it because of poor eating during the course of the treatment, and the seizures recurred after stopping the medicines.

The patient underwent thorough examination, including physical growth examination, laboratory examination, imaging examination, and behavior development assessment. Although the patient was 8.75 years old, her cognitive abilities were equivalent to those of a child aged 2-3 years. She was hyperactive and could not count or perform calculations. She could repeat adult language, but she could not communicate with others and could not eat alone or care for herself. She fell easily when walking fast. She was admitted to the neurology department of our hospital due to epilepsy and mental retardation.

2.2. Imaging Examination. The patient's conventional 12-lead electrocardiogram showed sinus arrhythmia and some T wave changes.

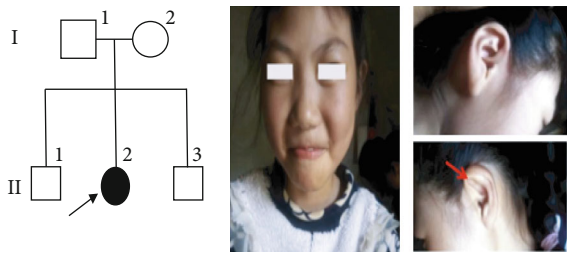
The patient's electroencephalogram (EEG) on August 16, 2019 showed that the EEG was abnormal when she was awake. Specifically, the background delta activity increased, and bilateral frontal and occipital sharp waves and sharp slow waves were released. Head CT showed that the white matter area of the bilateral lateral ventricles was reduced and the shape was irregular. CT of the patient's head showed that the white matter area of the bilateral lateral ventricles was reduced and that the shape was irregular.

Head MRI (August 16, 2019) revealed the following (Figure 2): (1) The gyrus of each lobe of the bilateral cerebral hemispheres is thick, the cortex is thickened, the surface is smooth, the sulci are few and shallow, and the bilateral ventricles are enlarged. No abnormal signal foci were found in the parenchyma of the brain. (2) The ventricular system was enlarged and the left temporal extracerebral space was widened. (3) The structure of the cerebellar hemisphere was poor.

2.3. Whole-Exome Sequencing and Genome Copy Number Analysis. First, we extracted genomic DNA from the peripheral blood of the proband and her family; NGS-based genetic testing was performed using Agilent SureSelect capture (Agilent, Santa Clara, California) and purified the DNA by amplification. We used the high-throughput sequencing platform (Illumina NovaSeq6000, USA) to detect the exon regions and flanking intron regions (20 bp) of 20099 genes in the human whole exome. Then, we compared the sequencing data with the human genome hg19 (GRCh37) reference sequence and evaluated the coverage and sequencing quality of the target area. We obtained 1523 genes related to the phenotype of the subject through multiple database searches.

This test analyzes the copy number variant sequencing (CNV-Seq) of large fragments (two or more consecutive exons) of the listed genes, and the CNVs have been verified by orthogonal experiments (qPCR or MLPA). When a detected pathogenic or possibly pathogenic variant is in an autosomal-recessive gene, the laboratory ensured that the coverage of the gene coding sequence reaches 100% through NGS and/or Sanger sequencing.

2.4. Analysis of Whole Exome Sequencing Data. First, we conducted mutation search and mutation annotation processes and found all possible mutations. Then, public databases (ESP6500, 1000 Genomes, ExAC, GnomAD) were used to exclude all mutations detected in the patient's sequencing data with a frequency greater than 0.001. By searching the OMIM, Swiss-Var, HGMD and ClinVar databases, we annotated reported disease genes and reported pathogenic sites. For missense mutations that are not recorded in the database, we used Polyphen-2, SIFT, CADD, and Mutationtaster protein software to predict the harmfulness of protein function. Finally, we combined the results of the WES and CNV-Seq analysis with the patient's phenotype, EEG and MRI analysis, and we performed genetic analysis according to the pathogenicity assessment results according to the ACMG guidelines.



(a) (b)

Figure 1: (a) The patient's family diagram. (b) The patient's facial features and images of both ears with left auricle deformity and a lack of an auricle structure.

According to the mutations found in WES, primers were designed, and Sanger sequencing was performed by polymerase chain reaction amplification to verify the mutation and inheritance pattern of WES.

3. Results

CNV-Seq detection did not find duplication or deletion of fragments. By WES, we analyzed the human GRCh37/hg19 genome database and found a heterozygous variant of *TUBG1* at chr17: 40765924 (NM_001070.4, *TUBG1*: c.751A>T; p.N251Y) adenine 751 was changed to thymine in the coding region and the aspartic acid encoding the protein became a tyrosine. The parent of the patients did not present the variant by WES, so the de novo variant was verified by Sanger sequencing (Figure 3(a)) and was not detected in the published database. The C.751A>T variant is not reported in HGMD or ClinVar databases or the published literature. Functional prediction indicated that the variant was more likely to cause changes in protein structure and function. Based on the patient's phenotype in conjunction with ACMG guidelines (PS2 + PM2 + PP3), this variant was evaluated as likely pathogenic.

Through analysis of protein conservation at *TUBG1:* c.751A>T; p.N251Y, the aspartic acid protein structure in this region of multiple species is highly conserved (Figure 3(b)).

We used Model V2.0 software to predict the complete structure of the *TUBG1* protein, and the visualization was produced using SAVES5.0. The three-dimensional structure diagram of wild-type and c.751A>T; p.N251Y *TUBG1* mutant protein reveals that the three-dimensional structure of the mutant protein was altered. This change may have a pathogenic effect on body function, leading to the occurrence of neurodevelopmental disorders caused by MCD (Figure 3(c)).

4. Discussion

Neurodevelopmental disorders caused by MCD are very rare neurological diseases. MCD represents a large group of brain cortical anomalies characterized by distinctive MRI findings [3]. The diagnosis of MCD is challenging due to frequent difficulties in defining etiologies. MRI is an essential diagnostic tool used in the assessment of specific epileptogenic cortical malformations or syndromes [14], but no specific cause could be detected. At present, thirteen cases of *TUBG1* mutations have been reported globally (Table 1). These cases cannot be diagnosed using chromosomal microarrays or biochemical screening for congenital metabolic abnormalities [15, 16]. In this study, we detected the etiology by whole-exome sequencing combined with CNV-Seq, but CNV-seq did not find duplications or deletions of fragment. We found that the *TUBG1* gene c.751A>T; p.N251Y is a novel pathogenic variant by WES, and the patient was finally diagnosed.

γ-Tubulin (*TUBG1*) was first described by Oakley and Oakley [17]. It belongs to the tubulin superfamily (including *TUBA1A*, *TUBB2A*, *TUBB2B*, *TUBB3*, *TUBB4A*, *TUBB*, and *TUBG1*). These tubulin gene mutations are related to a series of cortical malformations by disrupting normal microtubule interactions, which further affect the proliferation, migration, and differentiation of neuronal cells, as well as the growth and guidance of axons. Moreover, these genes are related to a series of brain malformations [18–20]. The *TUBG1* mutant affects the location of neurons and interferes with the movement of newborn neurons but does not affect the proliferation of progenitor cells [5]. The latest research suggests that novel heterozygous missense variants in the tubulin *TUBG1* gene lead to malformations of human cortical development, which further leads to intellectual disability and epilepsy [21].

Thirteen patients with different MRI features of MCD caused by mutation of *TUBG1* have been reported. Poirier et al. (2013) first reported the cases of three patients with unrelated cortical hypoplasia. Two patients showed severe malformation of the corpus callosum and cortical thickening and 1 patient presented a moderate malformation of the corpus callosum [2]. Brock et al. (2018) reported the cases of eight patients with cortical hypoplasia. One patient had diffuse anencephaly, and seven had corpus callosum malformations. Six of seven patients mainly had frontal lobe deformities, and 1 patient presented an enlarged frontal and parietooccipital lobe [1]. Yuen et al. (2019) reported one patient with corpus callosum malformation, and another with a decrease in the number of the cerebral sulci and gyri [22]. MRI analysis of this patient showed bilateral thickening of the cerebral hemispheres (parietal, occipital and temporal lobes). The cortex was thickened, and the surface was smooth. Only a few shallow sulci were noted, and the brain parenchyma was normal.

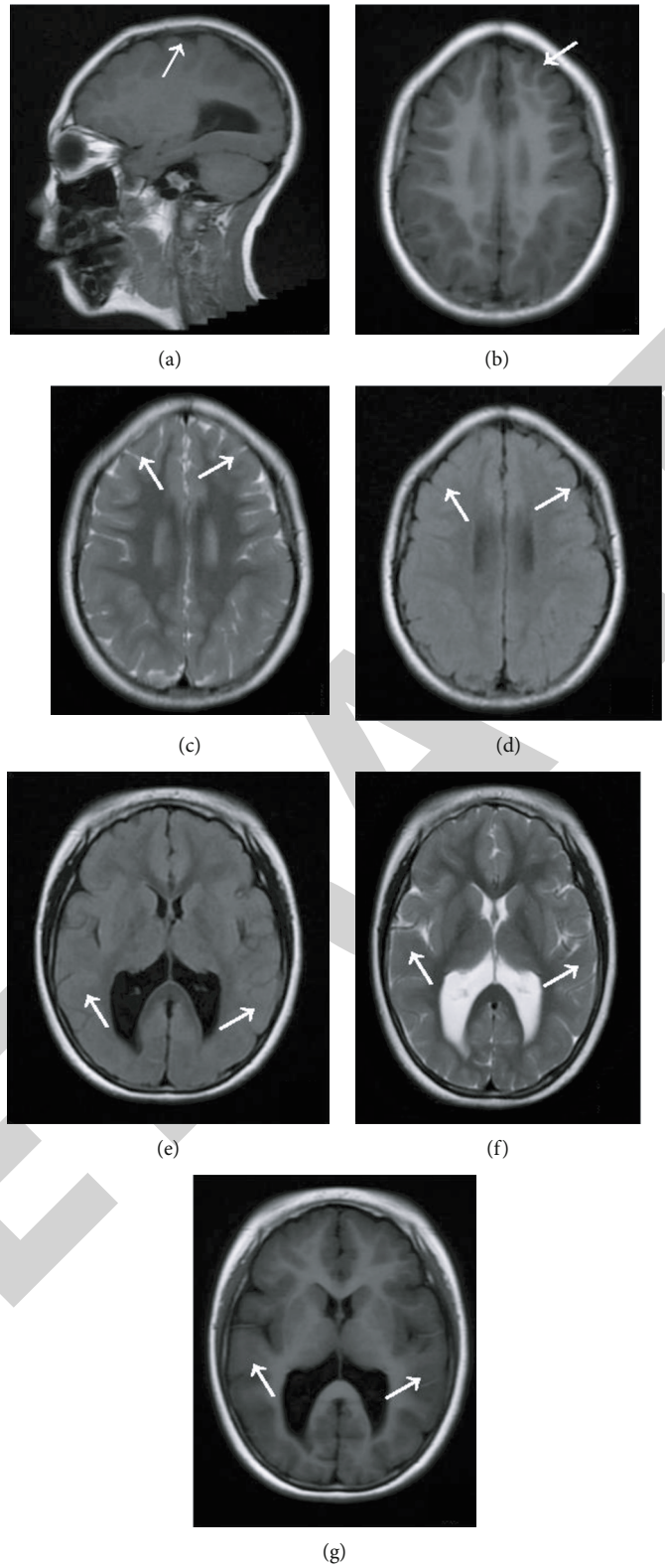
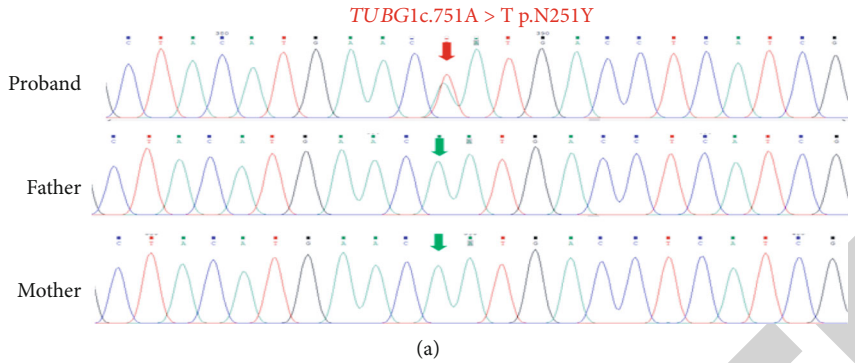


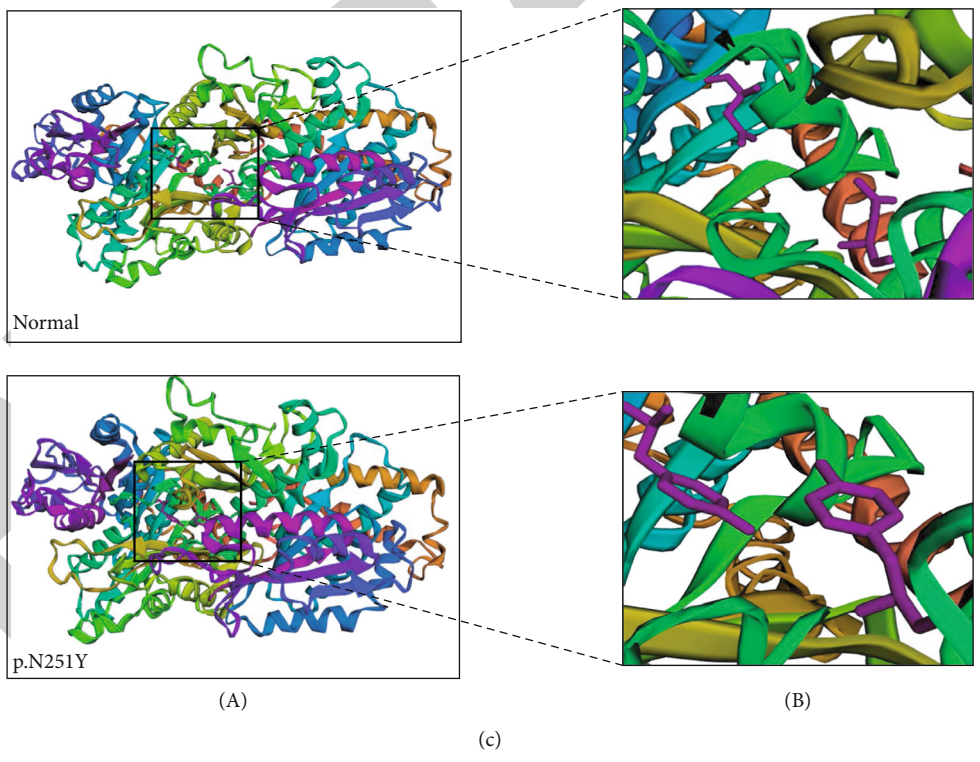
Figure 2: Magnetic resonance imaging of the patient's brain. (a) Female patient aged 8 years and 9 months. Magnetic resonance imaging of the brain showed the gyrus of each lobe of the bilateral cerebral hemispheres is thick, the cortex is thickened, the surface is smooth, the sulci are few and shallow, and the bilateral ventricles are enlarged. The white arrow indicates the deformity of the corpus callosum; (b, c, d) Show the T1W1, T2W1, and T2W1 FLAIR signals of the parietal lobe, respectively; (e, f, g) show the T1W1, T2W1, and T2W1 FLAIR signals of the temporal lobe and occipital lobe, respectively.



(a)

Species	Gene	aa	Alignment
Human	ENSP00000251413.2	251	T T T L R Y P G Y M N N D L I G L I A S L I P T
Mutated		251	T T L R Y P G Y M N Y D L I G L I A S L I P
Ptroglodytes	ENSPTRG00000009211	251	T T L R Y P G Y M N N D L I G L I A S L I P
Mmulatta	ENSMMUG00000016797	251	T T L R Y P G Y M N N D L I G L I A S L I P
Fcatus	ENSFCAG00000011674	251	T T L R Y P G Y M N N D L I G L I A S L I P
Ggallus	ENSGALG00000003150	220	T T T L R Y P G Y M N N D L I G L I A S L I P
Trubripes	ENSTRUG00000014769	251	T T L R Y P G Y M N N D L I G L I A S L I P
Drerio	ENSDARG00000015610	251	T T L R Y P G Y M N N D L I G L I A S L I P
Dmelanogaster	FBgn0260639	251	T T L R Y P S Y M N N N L I G L I A P L I P
Xtropicalis	ENSXETG00000014898	250	T T T L R Y P G Y M N N D L I G L I A S L I P

(b)



(c)

Figure 3: Analysis of the *TUBG1*:c.751A>T; p.N251Y mutation. (a) Sanger sequencing of the child suggested a de novo mutation in the proband. (b) *TUBG1* gene c.751A>T p.N251Y protein conservation analysis diagram. (c) Three-dimensional structures of *TUBG1* (WT) and *TUBG1* (p.N251Y) protein. (A) Complete protein structure of TUBG1 predicted by ModellerV2.0 and visualized with SAVES5.0. (B) The structures of p.251N and p.251Y are presented as purple sticks. The p.N251Y mutation turn forms the structure from an extended strand to a random coil as confirmed by the secondary structure prediction of *TUBG1* using the online HNN program. This system encompasses a neural network algorithm for secondary structure prediction. The 'h,' 'e,' and 'c' descriptors indicate the alpha helix, beta sheet, and random coil, respectively.

Table 1: Genotype and phenotype information of 13 previously reported *TUBG1* variants.

Case	Sex	Age at follow-up	Mutation	Mode of inheritance	Epileptic	Intellectual disability	Speech delay	Cortical dysgenesis (MRI)
This study	Female	8 years and 9 months	c.751A > T, p.Asn251Tyr	De novo	Yes	Severe	Can repeat words, cannot count or communicate	Bilateral thickening of the Guri of the lobes of the cerebral hemispheres, the cortex
Patient 1	Female	21 years	c.1160 T > C; p.Leu387Pro	De novo	Yes	Severe	Not available	Severe posterior predominant pachygyria/agyria (posterior agyria, frontal pachygyria), thickened cortex
Patient 2	Male	18 months	c.275A > G; p.Tyr92Cys	De novo	Yes	Severe	Not available	Severe posterior predominant pachygyria/agyria (posterior agyria, frontal pachygyria), thickened cortex
Patient 3	Female	31 years	c.991A > C; p.Thr331Pro	Father's DNA not available	Yes	Moderate	Not available	Posterior pachygyria, moderate posterior subcortical band heterotopia
Patient 4	Male	33 years	c.63C > A; p.Phe21Leu	De novo	No	Severe	Only sounds, no speech	Posterior predominant pachygyria (posterior frontal lobe and parieto-occipital cortex)
Patient 5	Male	21 years	c.985G > T; p.Asp329Tyr	Father's DNA not available	Yes	Severe	Nonverbal	Diffuse agyria
Patient 6	Female	19 months	c.776C > T; p.Ser259Leu	De novo	Yes	Not available	Delayed	Posterior predominant pachygyria (mild over frontal lobes, moderate over posterior lobes), cortex 10-13 mm thick
Patient 7	Female	14 years	c.776C > T; p.Ser259Leu	De novo	Yes	Not available	Nonverbal	Posterior predominant pachygyria, sparse cells over occipital lobes, cortex 13-15 mm thick
Patient 8	Female	11 years and 6 months	c.776C > T; p.Ser259Leu	Germline mosaicism in parent	Yes	Moderate	50 words	Posterior predominant pachygyria (mild over frontal lobe, moderate over temporal and occipital lobes), cortex 6-13 mm thick
Patient 9	Male	9 years and 6 months	c.776C > T; p.Ser259Leu	Germline mosaicism in parent	Yes	Moderate	Nonverbal	Posterior predominant pachygyria (mild over frontal lobe, moderate over temporal and occipital lobes), cortex >15 mm thick
Patient 10	Male	15 years	c.769A > T; p.Ile257Phe	De novo	Yes	Moderate	5 word sentences	Posterior predominant pachygyria (almost normal over frontal lobes, pachygyria over perisylvian and occipital lobes), cortex 6-10 mm thick
Patient 11	Female	18 months	c.776C > T; p.Ser259Leu	De novo	Yes	Severe	Nonverbal	Posterior predominant pachygyria (mild over frontal lobe, moderate over temporal and occipital lobe, deep parietal lobe infolding)
Patient 12	Female	10 years	c.202G > A; p.Asp68Asn	De novo	No	Moderate	Normal	Posterior predominant pachygyria, band heterotopia, nodular heterotopia
Patient 13	Male	6 months	c1021C > T; p.Arg341Trp	De novo	Yes	Moderate global delay	Moderate global delay	Reduced cortical sulci and gyri

Although the novel *TUBG1* variant has not been assessed in animal function experiments to confirm the relationship between its phenotype and genotype of MCD, we used Model V2.0 software to predict the structural integrity of *TUBG1* c.751A>T; p.N251Y variant protein and compared with the wild-type protein structure of the variant. We analyzed the putative effects of variants on protein structure and function. The three-dimensional structure diagram of wild-type and c.751A>T; p.N251Y *TUBG1* mutant protein reveals that the three-dimensional structure of the mutant protein is altered. The p.N251Y mutation transforms the structure from an extended strand to random coil as confirmed by the secondary structure prediction of *TUBG1* using the online HNN program. Based on the protein structure and function software prediction of the novel variant, we hypothesize that changes in the variant structure may affect the corresponding

functions to produce pathogenic effects and lead to neurodevelopmental disorders caused by MCD.

Although the complete phenotypic spectrum of tubular proteinosis is not completely clear, the MRI of the patient and the phenotype of left auricle deformity in this report increase the phenotypic characteristics of tubular protein lesions and also indicate the high clinical heterogeneity of neurodevelopmental disorders. Further functional studies will be focused on the relationship between this rare variant site and the phenotype to elucidate disease mechanisms that may facilitate the development of targeted treatments.

Retrospective analysis of the *TUBG1* variant reveals that is almost entirely a de novo mutation [1, 2, 22]. It is difficult to implement early screening, and this condition lacks effective treatment measures. Therefore, prenatal diagnosis during pregnancy is an effective means to block the continuation of the genetic factors implicated in the disease. At present, only symptomatic treatment is available. This patient mainly underwent antiepileptic drug treatment and rehabilitation training. No convulsions occured in the 1-year follow-up, and the condition did not develop.

5. Conclusions

This study clarifies the genetic etiology of the neurodevelopmental disorders in the patient and further reveals its characteristic phenotypic spectrum, molecular genetic findings, and functional consequences of novel pathogenic *TUBG1* variants in neurodevelopmental disorder caused by MCD. This study provides important guidance for research on the pathogenesis, treatment, prognostic evaluation, and reproductive implications of this disease.

Data Availability

The data used to support the findings of this study are all included within the article.

Conflicts of Interest

The authors declare that they have no competing interests, and all authors should confirm its accuracy.

Authors' Contributions

Ru Shen and Zhen Zhang performed data collection and wrote the manuscript. Yu Zhuang and Xiaohong Yang conducted the experiments and performed data analysis. Lifen Duan critically revised the manuscript. All authors have read and approved the manuscript. Ru Shen and Zhen Zhang are co-first authors.

Acknowledgments

We thank the patient and her family for their cooperation and contributions to this study. This work was generously supported by the Key R&D Program of Yunnan Provincial Department of Science and Technology-Special International Science and Technology Cooperation Project (No. 2018IA047); Yunnan Province Scientific and Technological Talents and Platform Program-Technical Innovation Talent Training Object Project Ru Shen (No. 202005AD160025); and Kunming Science and Technology Guarantee People's Livelihood Development Plan (No. 2019-1-S-25318000001074).

References

[1] S. Brock, K. Stouffs, E. Scalais et al., "Tubulinopathies continued: refining the phenotypic spectrum associated with variants in _TUBG1_," *European Journal of Human Genetics*, vol. 26, no. 8, pp. 1132–1142, 2018.

[2] K. Poirier, N. Lebrun, L. Broix et al., "Mutations in _TUBG1_ , _DYNC1H1_ , _KIF5C_ and _KIF2A_ cause malformations of cortical development and microcephaly," *Nature Genetics*, vol. 45, no. 6, pp. 639–647, 2013.

[3] S. Duquesne, M. C. Nassogne, P. Clapuyt, K. Stouffs, and Y. Sznajer, "Phenotype description in _KIF5C_ gene hot-spot mutations responsible for malformations of cortical development (MCD)," *European Journal of Medical Genetics*, vol. 63, no. 9, pp. 103991–103996, 2020.

[4] R. Romaniello, C. Zucca, F. Arrigoni et al., "Epilepsy in tubulinopathy: personal series and literature review," *Cell*, vol. 8, no. 7, pp. 669–714, 2019.

[5] E. L. Ivanova, J. G. Gilet, V. Sulimenko et al., "_TUBG1_ missense variants underlying cortical malformations disrupt neuronal locomotion and microtubule dynamics but not neurogenesis," *Nature Communications*, vol. 10, no. 1, pp. 2129–2134, 2019.

[6] A. Yuba-Kubo, A. Kubo, M. Hata, and S. Tsukita, "Gene knockout analysis of two γ-tubulin isoforms in mice," *Developmental Biology*, vol. 282, no. 2, pp. 361–373, 2005.

[7] E. Dráberová, V. Sulimenko, S. Vinopal et al., "Differential expression of human γ-tubulin isotypes during neuronal development and oxidative stress points to a γ-tubulin-2 prosurvival function," *The FASEB Journal*, vol. 31, no. 5, pp. 1828–1846, 2017.

[8] D. O. Wise, R. Krahe, and B. R. Oakley, "The γ-tubulin gene family in humans," *Genomics*, vol. 67, no. 2, pp. 164–170, 2000.

[9] M. Friedmann, L. T. Holth, H. Y. Zoghbi, and R. Reeves, "Organization, inducible-expression and chromosome localization of the human HMG-I(Y) nonhistone protein gene," *Nucleic Acids Research*, vol. 21, no. 18, pp. 4259–4267, 1993.

[10] N. Bahi-Buisson, K. Poirier, F. Fourniol et al., "The wide spectrum of tubulinopathies: what are the key features for the diagnosis?," *Brain*, vol. 137, no. 6, pp. 1676–1700, 2014.

[11] N. M. Mahoney, G. Goshima, A. D. Douglass, and R. D. Vale, "Making microtubules and mitotic spindles in cells without functional centrosomes," *Current Biology*, vol. 16, no. 6, pp. 564–569, 2006.

[12] C. Suri, T. W. Hendrickson, H. C. Joshi, and P. K. Naik, "Molecular insight into γ-γ tubulin lateral interactions within the γ-tubulin ring complex (γ-TuRC)," *Journal of Computer-Aided Molecular Design*, vol. 28, no. 9, pp. 961–972, 2014.

[13] S. Richards, N. Aziz, S. Bale et al., "Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology," *Genetics in Medicine*, vol. 17, no. 5, pp. 405–423, 2015.

[14] B. Thomas, N. Al Dossary, and E. Widjaja, "MRI of childhood epilepsy due to inborn errors of metabolism," *AJR. American*

Hindawi
BioMed Research International
Volume 2023, Article ID 9865973, 1 page
https://doi.org/10.1155/2023/9865973

Retraction

Retracted: Identification of Tumor Tissue of Origin with RNA-Seq Data and Using Gradient Boosting Strategy

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] R. Li, B. Liao, B. Wang et al., "Identification of Tumor Tissue of Origin with RNA-Seq Data and Using Gradient Boosting Strategy," *BioMed Research International*, vol. 2021, Article ID 6653793, 14 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6653793, 14 pages
https://doi.org/10.1155/2021/6653793

Research Article

Identification of Tumor Tissue of Origin with RNA-Seq Data and Using Gradient Boosting Strategy

Ruixi Li,[1,2,3] **Bo Liao**,[1,2,3] **Bo Wang**,[4,5] **Chan Dai**,[4,5] **Xin Liang**,[1,2,3] **Geng Tian**,[4,5] **and Fangxiang Wu**[1,2,3,6]

[1]*School of Mathematics and Statistics, Hainan Normal University, Haikou 570100, China*
[2]*Key Laboratory of Computational Science and Application of Hainan Province, Haikou 571158, China*
[3]*Key Laboratory of Data Science and Intelligence Education (Hainan Normal University), Ministry of Education, Haikou 571158, China*
[4]*Qingdao Geneis Institute of Big Data Mining and Precision Medicine, Qingdao 266000, China*
[5]*Geneis (Beijing) Co., Ltd., Beijing 100102, China*
[6]*Division of Biomedical Engineering, Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, SK, S7N5A9, Canada*

Correspondence should be addressed to Bo Liao; dragonbw@163.com

Received 30 October 2020; Revised 19 January 2021; Accepted 6 February 2021; Published 17 February 2021

Academic Editor: Min Tang

Background. Cancer of unknown primary (CUP) is a type of malignant tumor, which is histologically diagnosed as a metastatic carcinoma while the tissue-of-origin cannot be identified. CUP accounts for roughly 5% of all cancers. Traditional treatment for CUP is primarily broad-spectrum chemotherapy; however, the prognosis is relatively poor. Thus, it is of clinical importance to accurately infer the tissue-of-origin of CUP. *Methods*. We developed a gradient boosting framework to trace tissue-of-origin of 20 types of solid tumors. Specifically, we downloaded the expression profiles of 20,501 genes for 7713 samples from The Cancer Genome Atlas (TCGA), which were used as the training data set. The RNA-seq data of 79 tumor samples from 6 cancer types with known origins were also downloaded from the Gene Expression Omnibus (GEO) for an independent data set. *Results*. 400 genes were selected to train a gradient boosting model for identification of the primary site of the tumor. The overall 10-fold cross-validation accuracy of our method was 96.1% across 20 types of cancer, while the accuracy for the independent data set reached 83.5%. *Conclusion*. Our gradient boosting framework was proven to be accurate in identifying tumor tissue-of-origin on both training data and independent testing data, which might be of practical usage.

1. Introduction

Cancer of unknown primary (CUP) is a type of malignant tumor, histologically diagnosed as a metastatic carcinoma with no confidently anatomical primary site even after comprehensive evaluation. CUP accounts for approximately 3% to 5% of all tumors [1–4]. In general, primary cancer tissue can be identified at the same time as diagnosis. However, for some patients, it is relatively difficult to identify cancer tissue-of-origin since the markers for origin tracing is unidentifiable. Previous studies showed that less than 50% of CUPs could be accurately diagnosed [5–8]. Accurate classification of the tumor types according to anatomical and histological assays is urgent [9–11].

The patients diagnosed as CUP are treated by using traditional chemotherapy; however, prognoses of these patients are relatively poor. For a physician, accurate diagnosis can be a direct guide to individual surgical intervention as well as medication regimen. Furthermore, identification of the primary site of the tumor is relatively helpful for clinicians to design a targeted treatment plan, as well as improving survivals and quality of life [12, 13].

Currently, the diagnostic techniques primarily include comprehensive evaluation, imaging examination,

Initialization. Initialize with $f_0(x) = \arg\min_c \sum_{i=1}^{N} L(y_i, c)$.

For $t = 1$ to T:

Perform updates:

(1) Compute pseudo residual: $\tilde{y}_i = -(\partial L(y_i, f_{t-1}(x_i))/\partial f_{t-1}(x_i)), i = 1, 2, \cdots, N$

(2) Find the parameters of the beat weak learner:

$\omega_t = \arg\min_{\omega} \sum_{i=1}^{N} [\tilde{y}_i - h_i(x_i; \omega)]^2$.

(3) Choose the step-size ρ_t by line search:

$\rho_t = \arg\min_{\rho} \sum_{i=1}^{N} L(y_{i,f_{t-1}}(x_i) + \rho h_t(x_i; \omega_t))$.

(4) Update the model $f_t(x) = f_{t-1}(x) + \rho_t h_t(x; \omega_t)$

Output $f_T(x)$

Algorithm 1: Gradient boosting.

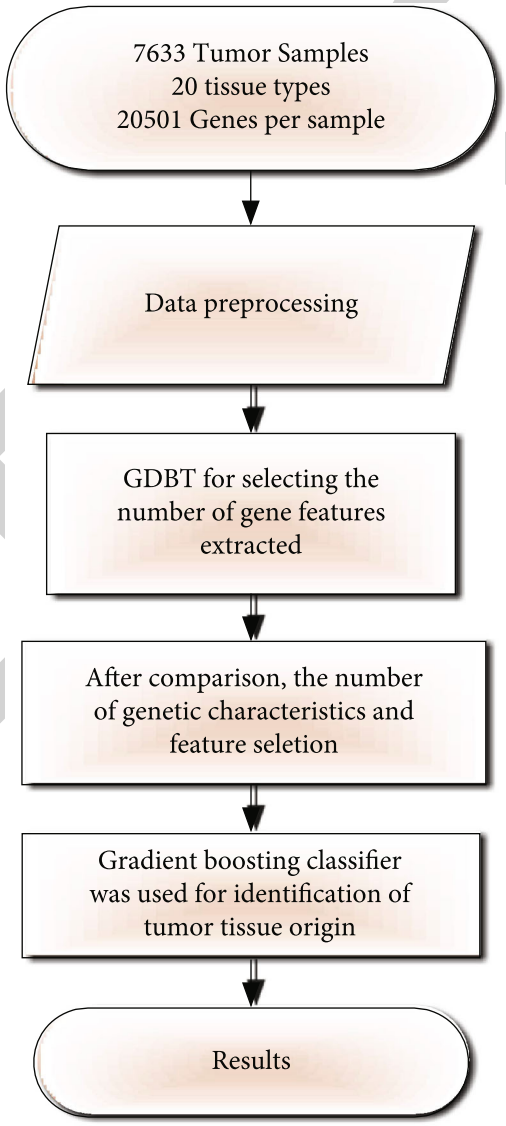


Figure 1: Flow chart of identification of tumor tissue origin.

Table 1: The disease name and sample number in TCGA data.

Disease	Code	Tumor samples	Percentage
Bladder urothelial carcinoma	BLCA	301	3.9025%
Breast invasive carcinoma	BRCA	1056	13.6912%
Cervical squamous cell carcinoma and endocervical adenocarcinoma	CESC	258	3.3450%
Colon adenocarcinoma	COAD	451	5.8473%
Glioblastoma multiforme	GBM	153	1.9837%
Head and neck squamous cell carcinoma	HNSC	480	6.2233%
Kidney renal clear cell carcinoma	KIRC	526	6.8197%
Kidney renal papillary cell carcinoma	KIRP	222	2.8783%
Acute myeloid leukemia	LAML	173	2.2430%
Brain lower grade glioma	LGG	439	5.6917%
Liver hepatocellular carcinoma	LIHC	294	3.8117%
Lung adenocarcinoma	LUAD	486	6.3011%
Lung squamous cell carcinoma	LUSC	428	5.5491%
Ovarian serous cystadenocarcinoma	OV	261	3.3839%
Pancreatic adenocarcinoma	PAAD	142	1.8410%
Prostate adenocarcinoma	PRAD	379	4.9138%
Rectum adenocarcinoma	READ	153	1.9837%
Skin cutaneous melanoma	SKCM	80	1.0372%
Stomach adenocarcinoma	STAD	415	5.3805%
Thyroid carcinoma	THCA	500	6.4826%
Uterine corpus endometrial carcinoma	UCEC	516	6.6900%
Total		7713	

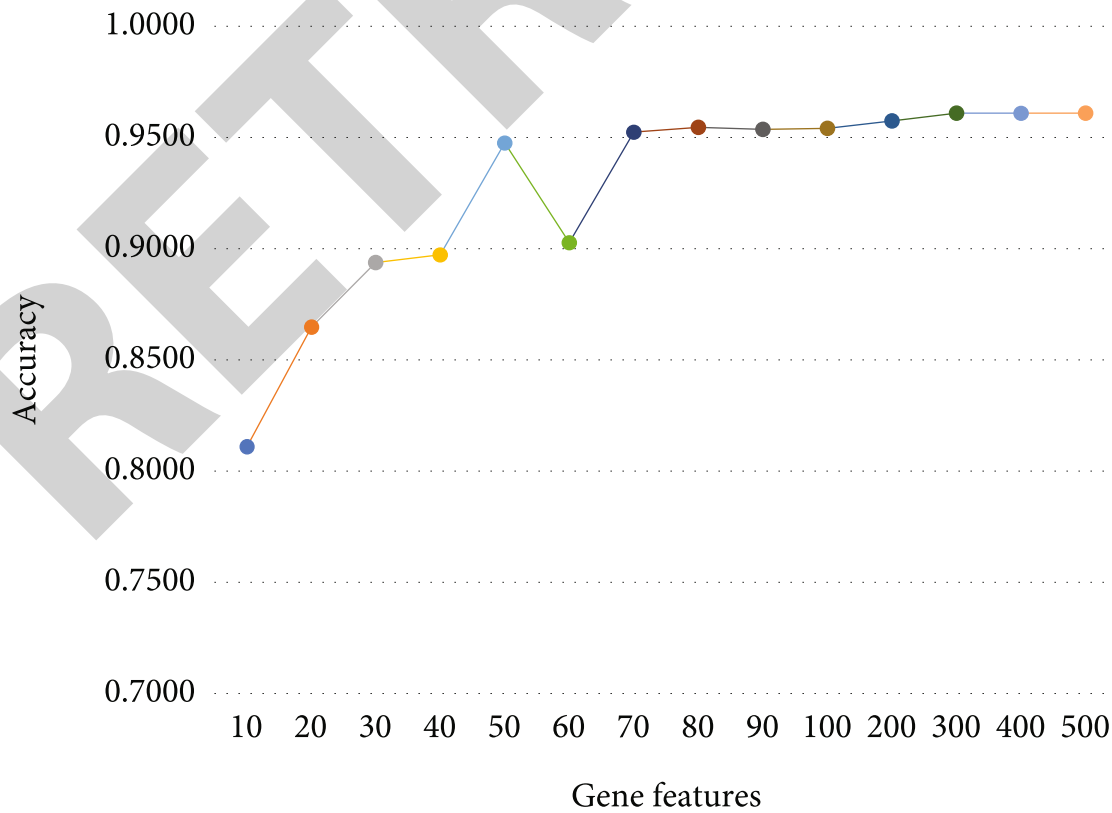


Figure 2: Accuracies of different numbers of genes with cross-validation.

	BLCA	BRCA	CESC	COAD	GBM	HNSC	KIRC	KIRP	LAML	LGG	LIHC	LUAD	LUSC	OV	PAAD	PRAD	READ	STAD	THCA	UCEC
BLCA	279		5	2		2				1		1	6			1	1		1	2
BRCA	4	1049	2												1					
CESC	7	1	237	1								1	1		2					8
COAD	1		1	421											1		26			1
GBM	2	1			147		1						1					1		
HNSC		1				478					1									
KIRC	4						506	14			2									
KIRP	4						10	208												
LAML									173											
LGG					1		1	1		436										
LIHC	1							2			290								1	
LUAD	2		1								1	466	15							1
LUSC	5	1	2			1					1	28	390							
OV	2													257						2
PAAD	2	1	1	1						1		1			133			2		
PRAD												1			1	377				
READ	2			81													70			
STAD				1														414		
THCA		1																	499	
UCEC	3		3	1				1				3								505

Figure 3: Confusion matrix of the classification using 400 genes.

pathological analysis, immunohistochemistry (IHC) panels, and genetic testing [2]. A gene expression-based test is considered as an adjunct test to an uncertain diagnosis of biopsy; moreover, it provides a new approach for the cancer diagnosis of predicting the prognosis of tumors [12]. Many cancerous cells retain features of their primary tissues of origin during metastasis; in other words, gene expression of metastatic cancer should be consistent with the gene expression of its primary tissue [14, 15]. It has been found that the gene expression profiles of metastatic tumors were different from the tissue of the metastatic site but more similar to those at the primary origin. A gene expression profile of the tissue origin is always retained during the process of tumor occurrence, development, and metastasis. Based on this theory, researchers developed a series of molecular markers of gene expression to trace the tissue origin of tumors.

CancerTYPE ID was a gene expression-based test, focusing on identifying the tissue of origin. This molecular test was based on real-time PCR technology by using the differential expression data of 92 genes in the tumor cells and classified tumors by matching the gene expression partem of tumor specimens to a database of 50 known tumor types and subtypes. The test compared genomic information from tumor samples with reference databases of more than 2000 tumors with definitive diagnoses. Gene expression profile analysis by using microarray data provided diagnoses of cancer types with high accuracy [7]. Another gene expression-based test named the Pathwork Tissue of Origin (TOO) test also contributes to improve the diagnosis of CUP. The Pathwork Tissue of Origin test applied a microarray-based expression profile of 2000 gene markers to assess the molecular similarity of the patient tumor with a panel of 15 known Genomic Test for Tumor Origin in formalin-fixed, paraffin-embedded (FFPE) tissues. This method primarily included two algorithms, one for standardization and the other for classification [2, 16].

RNA-seq is a high-throughput sequencing approach that sequences mRNA, small RNA, and noncoding RNA by using high-throughput sequencing technology. RNA-seq, characterized with more exact quantification, higher repeatability, wider examination area, and more credible analysis, can be used to study genome-wide differences in gene expression. In addition, it is considered as cost-effective. TOO was based on Array data, and CancerTYPE ID was conducted on the RT-PCR data; however, application of RT-PCR or Array has not only a higher cost but also a limited accuracy. Here, we conducted an experiment to identify the tissue of origin with a gradient boosting classifier [17] and RNA-seq technique.

2. Materials and Methods

2.1. Data Preparation. The Cancer Genome Atlas (TCGA) RNA-seq and array data include 20,501 genes from the ICGC Data Portal (https://dcc.icgc.org/releases/release_26/) download. In order to facilitate the follow-up work, we generated a $M * N$ matrix where M represents the sample size and N represents the number of genes. The matrix was generated by normalizing the expression value of each sample and each gene from TCGA. An independent data set, including 79 tumor samples from 6 cancer types with known origins,

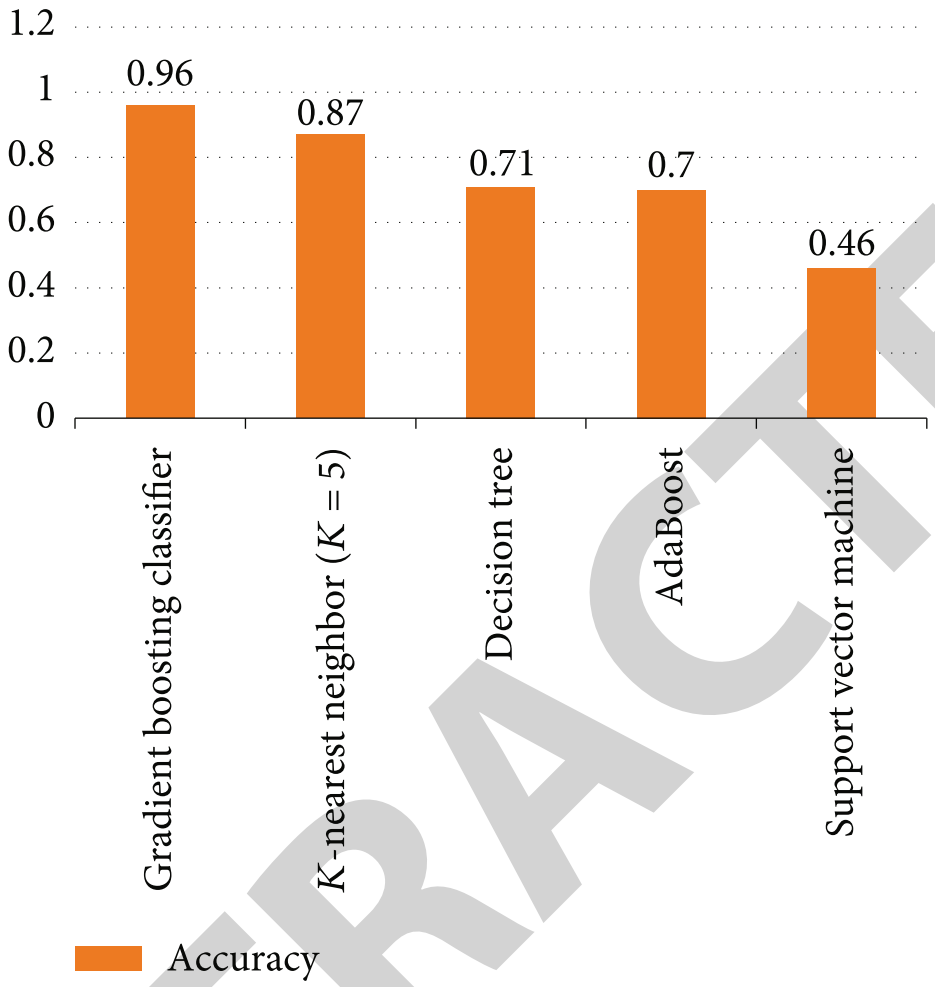


Figure 4: Accuracies of five different algorithms based on TCGA.

was also downloaded from the Gene Expression Omnibus (GEO). These samples belong to GSE8352, GSE8734, GSE11107, GSE11132, GSE4895, GSE6491, GSE7966, GSE7766, and GSE11843. The samples not included in the 20 cancers were excluded.

2.2. Gene Selection and Classification. We employed a gradient boosting algorithm for gene feature selection and final classification with cross-validation. Gradient boosting (GBDT) is a machine learning method for regression and classification in studies, which combines multiple weak learners into prediction models [18]. Furthermore, the weak learner is usually a decision tree. In the GBDT iteration, we assume that the strong learner obtained in the previous iteration is $f_{t-1}(x)$ and the loss function is $L(y, f_{t-1}(x))$. The goal of this round of iterations is to find a weak-learner $h_t(x)$ of the CART regression tree model and minimize the loss function $L(y, f_t(x) = f_{t-1}(x) + h_t(x))$ of this cycle. This iteration finds the decision tree, and therefore, the sample loss is as small as possible.

Major step in this machine learning method is to minimize the loss function L through optimization. In the t-th

Table 2: Correctly and incorrectly predicting the type of cancer.

Predicted_label	True_label	Matched_label
TCGA-BRCA	TCGA-BRCA	1
TCGA-BRCA	TCGA-BRCA	1
TCGA-LIHC	TCGA-BRCA	0
TCGA-BLCA	TCGA-BRCA	0
TCGA-BRCA	TCGA-BRCA	1
TCGA-BRCA	TCGA-BRCA	1
TCGA-BRCA	TCGA-BRCA	1
TCGA-UCEC	TCGA-CESC	0
TCGA-CESC	TCGA-CESC	1
TCGA-COAD	TCGA-COAD	1
TCGA-HNSC	TCGA-HNSC	1
TCGA-HNSC	TCGA-HNSC	1
TCGA-THCA	TCGA-THCA	1
TCGA-THCA	TCGA-THCA	1
TCGA-THCA	TCGA-THCA	1
TCGA-THCA	TCGA-THCA	1
TCGA-THCA	TCGA-THCA	1
TCGA-THCA	TCGA-THCA	1
TCGA-THCA	TCGA-THCA	1
TCGA-THCA	TCGA-THCA	1

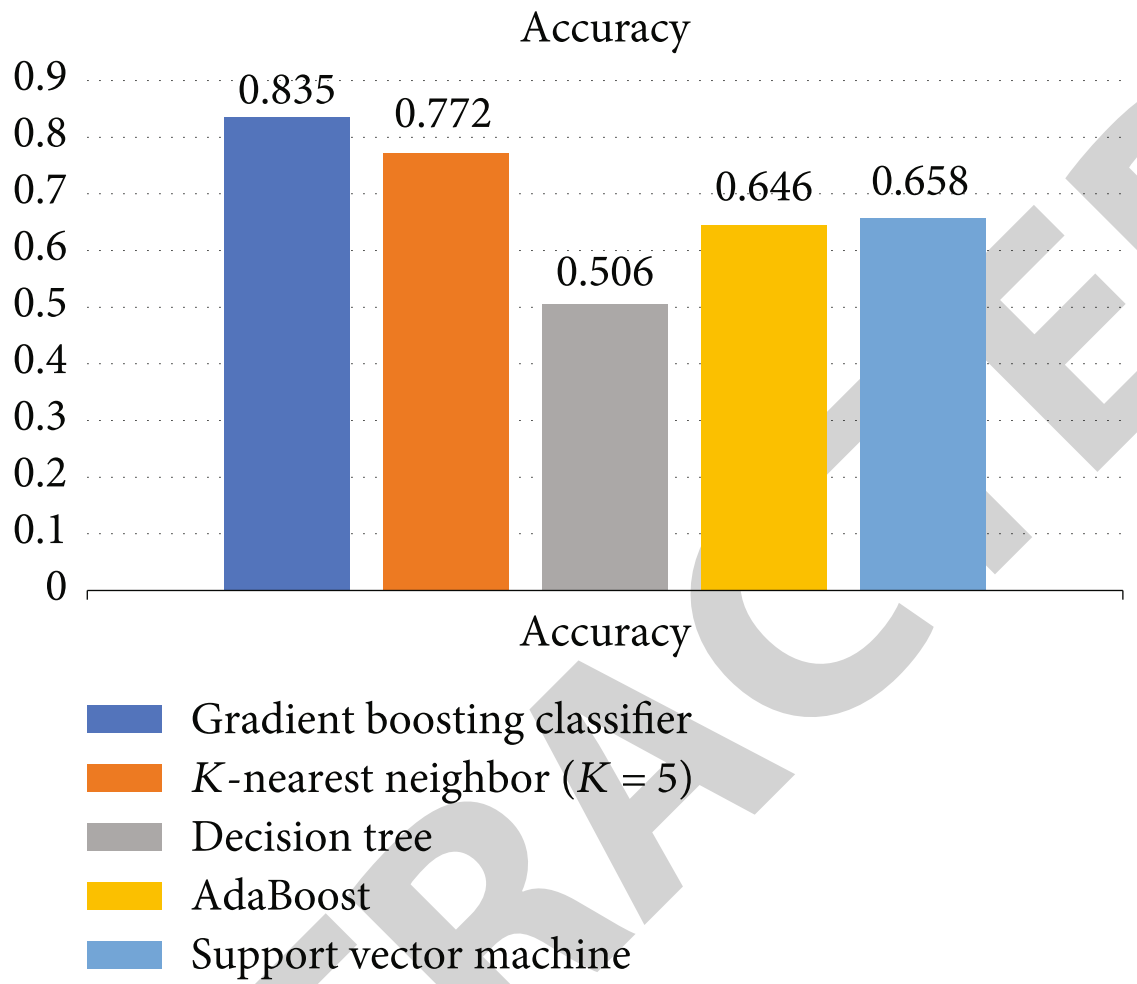


Figure 5: Accuracies of five different algorithms based on GEO.

iteration, the first $t-1$ base learners are all fixed,

$$f_t(x) = f_{t-1}(x) + \rho_t h_t(x). \quad (1)$$

Minimize loss function

$$L(f) = \sum_{i=1}^{N} L(y_i, f_t(x_i)). \quad (2)$$

The negative gradient of the loss function of the sample of the t wheel is expressed as

$$h_t \approx -\frac{\partial L(y_i, f_{t-1}(x_i))}{\partial f_{t-1}(x_i)}. \quad (3)$$

Input cancer sample training set:

$$T = \{(x_1, y_1), (x_2, y_2), \cdots x_N, y_N\}. \quad (4)$$

$N$is the number of 7633 cancer samples, the maximum number of iterations is T, the loss function is L, and output maximum learner is $f(x) = f_T(x)$.

For a single tree T, the following formula for the importance of each feature X_1 is used:

$$\Gamma_l^2(T) = \sum_{t=1}^{Q-1} i\wedge^2 I(v(t) = l), \quad (5)$$

where Q is the number of leaf nodes, Q-1 is the number of internal nodes, $X_{v(t)}$ is the splitting characteristic associated with the internal node$t$where$t$is for the cancer type, and$l$is the number of features. For each internal node t, the feature $X_{v(t)}$ is used to simulate and divide the feature space to obtain a square error reduction after splitting, that is, $i\wedge^2$. Finally, the importance of feature X_1 is summed up by the error reduction on all internal nodes. The more the total error is reduced, the more important this feature is. Because similar response values are in the same set, every node in the decision trees is a condition on a single gene. The more the total error decreases, the more important the feature becomes. For the integration of M trees, feature importance is the average of corresponding values of each tree.

Unlike GBDT, AdaBoost selects an exponential loss, while GBDT uses the classifying loss of function from the logistic loss $L(y, f(x)) = \log(1 + e^{-2yf(x)})$. We expected to

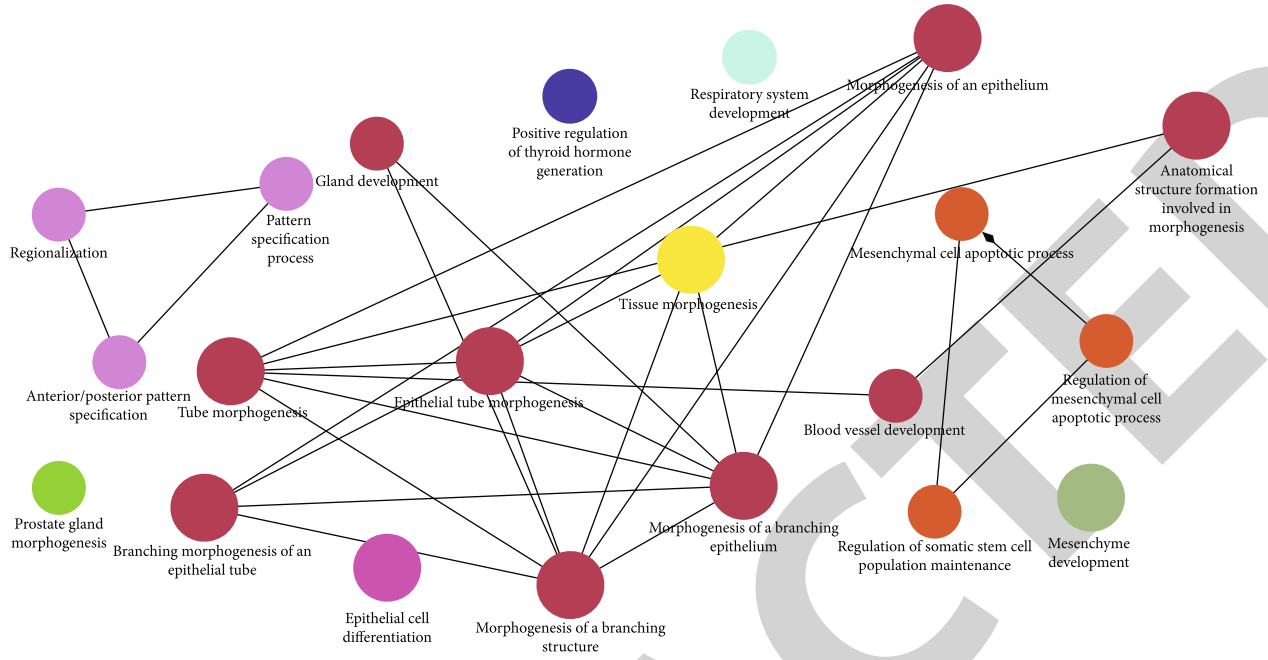


Figure 6: GO enrichment analysis about pathway.

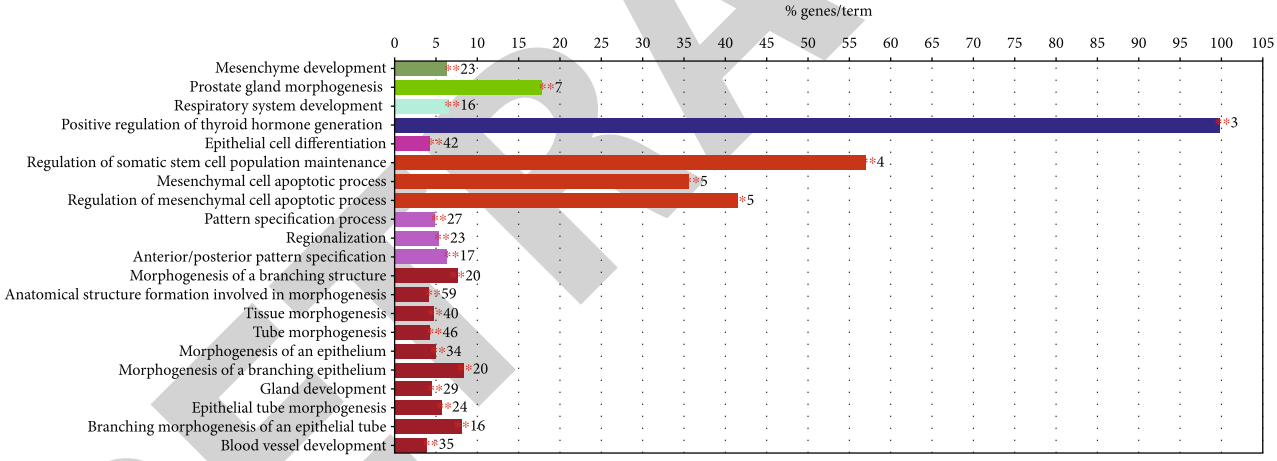


Figure 7: Specific cluster.

minimize the loss of the function; therefore, we used the derivative of the function to find the minimum value of the function. After getting $f_T(x)$, we have to do the probability estimation by $P = P(y = 1|x) = 1/(1 + e^{-2f(x)})$.

3. Results and Discussion

3.1. Workflow. The study process for identifying the tumor-of-origin was shown in Figure 1. Firstly, the expression profiles were downloaded from TCGA. A preprocess for the raw data was carried out before feature selection, which was performed by using the gradient boosting algorithm with 10-fold cross-validation. Then, final classification across 20 types of cancer was conducted by utilizing gradient boosting classifier, and the output of the model was displayed as an evaluation metric.

3.2. Data Preparation. From TCGA (Cancer Genome Atlas Research, 2008) data set, we downloaded expression profiles for 7713 RNA-seq samples covering 21 common cancers without metastasis [19]. Two samples were removed because of lack of clinical data. Then, we used RSEM to normalize these data. Table 1 summarized these data and showed the information for tumor samples.

372 metastasis samples containing 352 cases with SKCM were originally included in the test data set. However, the metastatic cases that originated from SKCM are relatively higher than those from other cancers. In order to reduce impact on the results, SKCM data were removed during data analysis.

3.3. 400 Genes Were Selected for Future Prediction. 20,501 genes across 7713 samples from the TCGA data set were

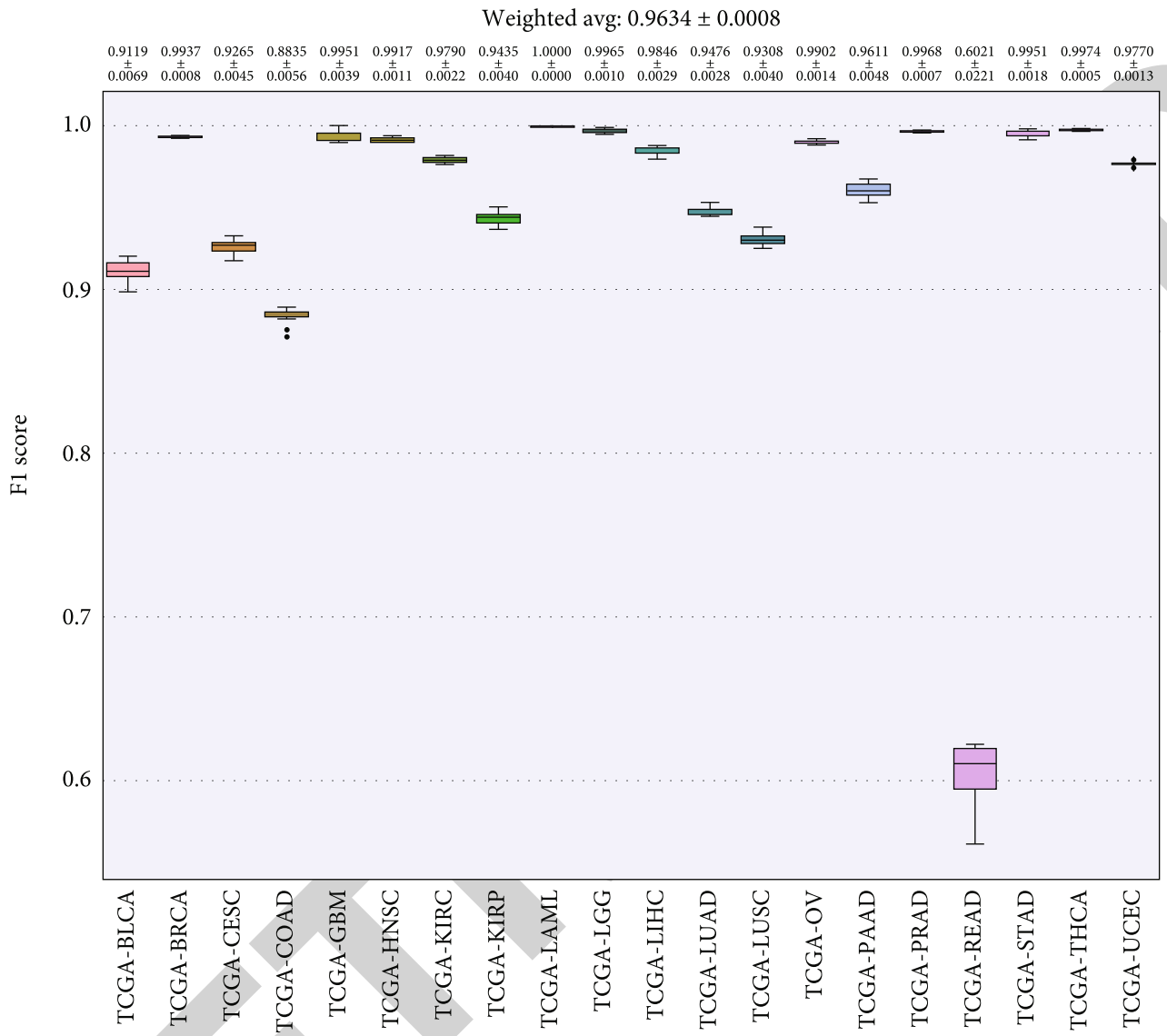


Figure 8: F1 score for each cancer.

included in the study. In order to reduce the model complexity, we performed feature selection. First, we ranked genes by importance scores calculated by gradient boosting algorithm, and the order was defined from high to low. We conducted a series of experiments, and the experimental results are shown in Figure 2. Based on the experimental results, we selected the feature number with highest accuracy. The top 400 gene features were extracted from each sample to construct a 7633 $\times$ 400 matrix [7]. This new matrix was the input for the classification of various cancers.

3.4. Classification. In the gene selection part, we got a 7633 $\times$ 400 matrix as the input matrix, and the corresponding gene expression profile of each sample was extracted. By using the GBDT method, we set n_estimators to 200. In fact, we also tested the estimator value from 100 to 300, and the results showed an upward trend followed downward trend and reached the maximum value at 200. Therefore, we finally chose 200 weak classifiers, which meant the number of decision trees was 200. The trained tree was used to select each cancer and returned the cancer which has been selected more times. We used the gene expression values as the training features to fit the cancer type as labels.

We adopted a 10-fold cross-validation in this study, which divided the data set into 10 subsets. Nine subsets were merged to a training set, and one subset was used as the test set. We repeated the algorithm ten times using the same gene features, and the average precision was 96.1%.

The confusion matrix is a standard format for precise evaluation, which is represented by an $M * M$ matrix. A confounding matrix can be used to judge the accuracy of the classifier classification and is presented in the form of a graph, so it is widely used to measure the success rate of classification. The confusion matrix is a summary of the predicted results of the classification problem. It can find errors in the classification model and understand the types

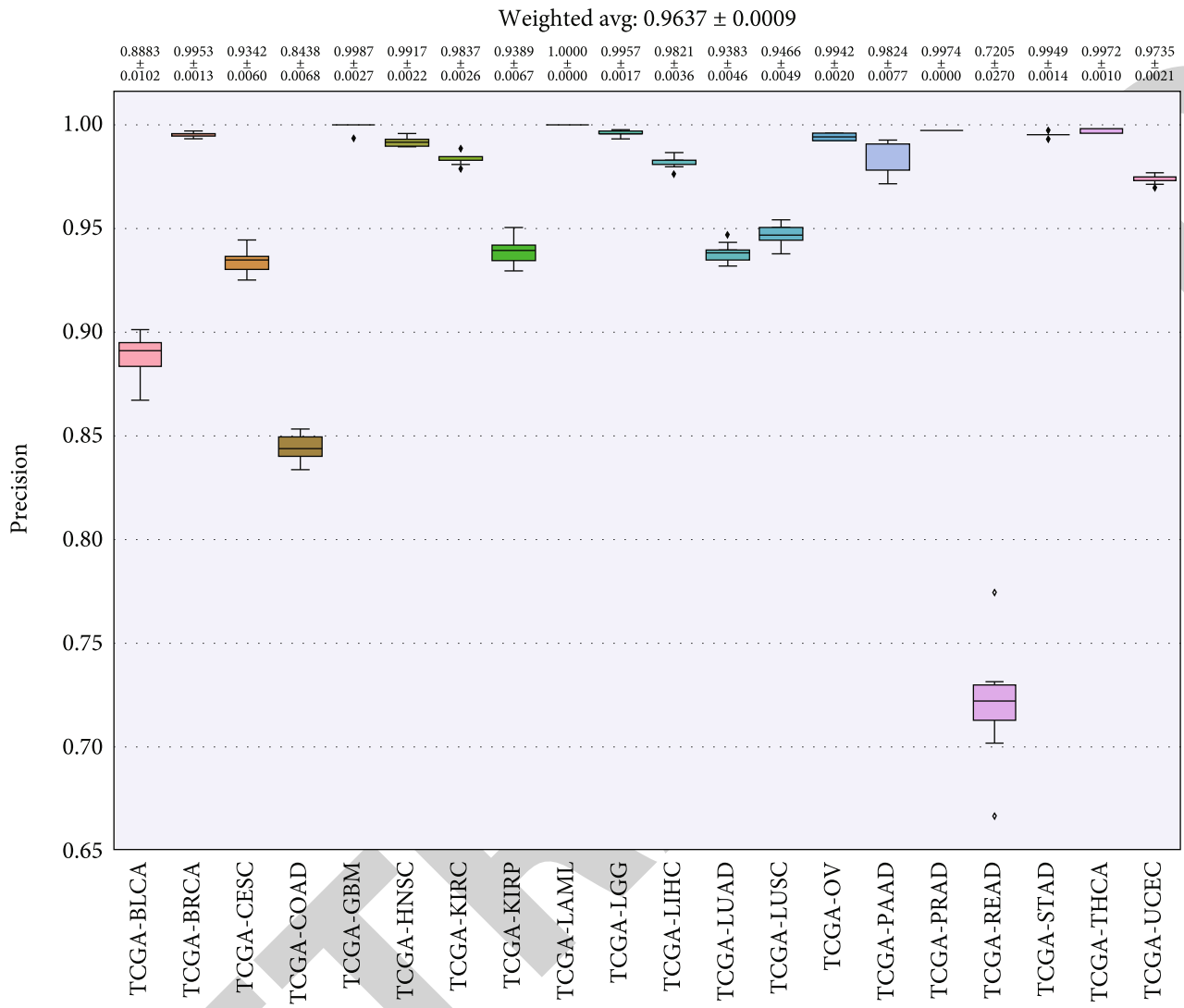


Figure 9: Precision for each cancer.

of errors that are occurring [20, 21]. The confusion matrix of the classification using 400 genes shown in Figure 3 exhibited the sample number of a certain type of cancer that was classified into another type.

We also made a comparison with K-nearest neighbor ($K = 5$) [22], decision tree [23], AdaBoost [24], and support vector machine [25]. The results are shown in Figure 4. The results of K-nearest neighbor ($K = 5$) are closer to GBDT; GBDT is significantly higher than the other methods.

Table 2 showed correct and incorrect predictions of each type of cancer. For example, it is TCGA-BRCA but was predicted to be TCGA-LIHC or TCGA-BLCA, and it is TCGA-CESC but was identified to be TCGA-UCEC. As shown in Table 1, BRCA, LIHC, BLCA, CESC, and UCEC, respectively, represented breast invasive carcinoma, liver hepatocellular carcinoma, bladder urothelial carcinoma, cervical squamous cell and carcinoma endocervical adenocarcinoma, and uterine corpus endometrial carcinoma. Except for the above cases, the overall prediction accuracy was reaching 85%.

In order to verify the generalization and robustness of the approach, we also downloaded data sets from GEO for independent validation. The data sets covered 6 cancer types, including BRCA, LUAD, PAAD, PRAD, STAD, and THCA. And the overall accuracy rate from the gradient boosting classifier reached 83.5%.We also made a comparison with K-nearest neighbor ($K = 5$), decision tree, AdaBoost, and support vector machine. The results are shown in Figure 5. The results of K-nearest neighbor ($K = 5$) are closer to GBDT; GBDT is significantly higher than the other methods.

Biological validation of the optimal biomarker signature was done by GO enrichment analysis. The Gene Ontology (GO) Consortium was formed to address the limited interoperability of genomic databases due to lack of progress [26]. Figures 6 and 7 are the result of the GO enrichment analysis. The enrichment results showed that the genes were significantly enriched in maintenance and regulation of cell differentiation during morphogenesis of human organs and suborgan tissues, such as cell differentiation in kidney and

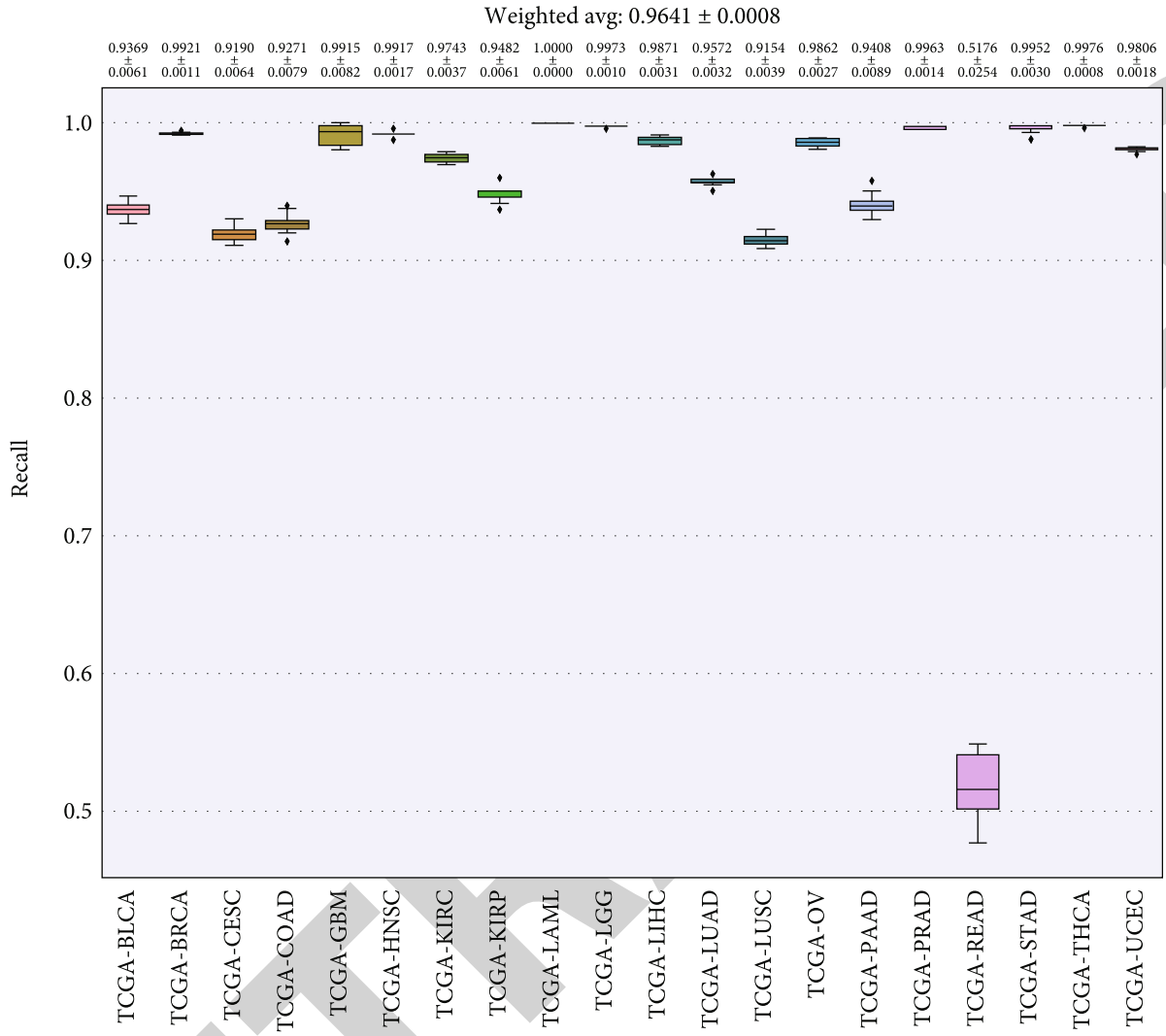


Figure 10: Recall for each cancer.

prostate gland morphogenesis, reproductive system development, urogenital system development, epithelial tube morphogenesis, and mesenchymal cells. There are other genes involved in the hormone-mediated signaling pathway, cell proliferation, angiogenesis and apoptosis, and thyroid hormone regulation. Overall, the genes were enriched in negatively regulating organ morphogenesis, positively regulating cell differentiation during morphogenesis, and inducing cell apoptosis. Remarkably, some genes involved in organ- or tissue-specific development are more likely to be differentially expressed in tumors and normal tissues. The HOXB13 gene which belongs to the HOX superfamily was highly enriched in prostate adenocarcinoma. Increased expression from the HoxB13 is indicative of an invasive or metastatic status as well as increases cellular migration and/or mobility. The HoxB13 expression level could be a potential marker to evaluate clinical diagnosis as well as patient prognosis [27–32].

In Figures 8–11, we presented the results of 10 times 10-fold cross-validation. Precision refers to the proportion of the correct model prediction among all results that the model prediction is positive. Recall refers to the ratio of the number of correctly predicted positive samples to the total number of true positive samples, that is, how many positive samples can be correctly identified from these samples. Specificity, which is relative to recall, refers to the ratio of correctly predicted negative samples to the total number of true negative samples. In other words, how many negative samples can be correctly identified from these samples. The F1 score is equivalent to the harmonic average of precision and precision. If any number of the recall and precision decreases, the F1 score will decrease.

A heat map is a visualization method to analyze the distribution of experimental data, which directly reflected the expression of 400 characteristic genes in cancer species. As shown in Figure 12, the expression levels of the top 50 characteristic genes in the cancer species were relatively average, among which C19orf33, CRYAB, ACTG2, ACTA2, IGFBP2, CSRP1, RAB34, SMS, MAGOH, C21orf33, IDI1, TRIM27, ACTL6A, and ILVBL gained higher expression, while OR14A16, CRP, and INS had lower expression.

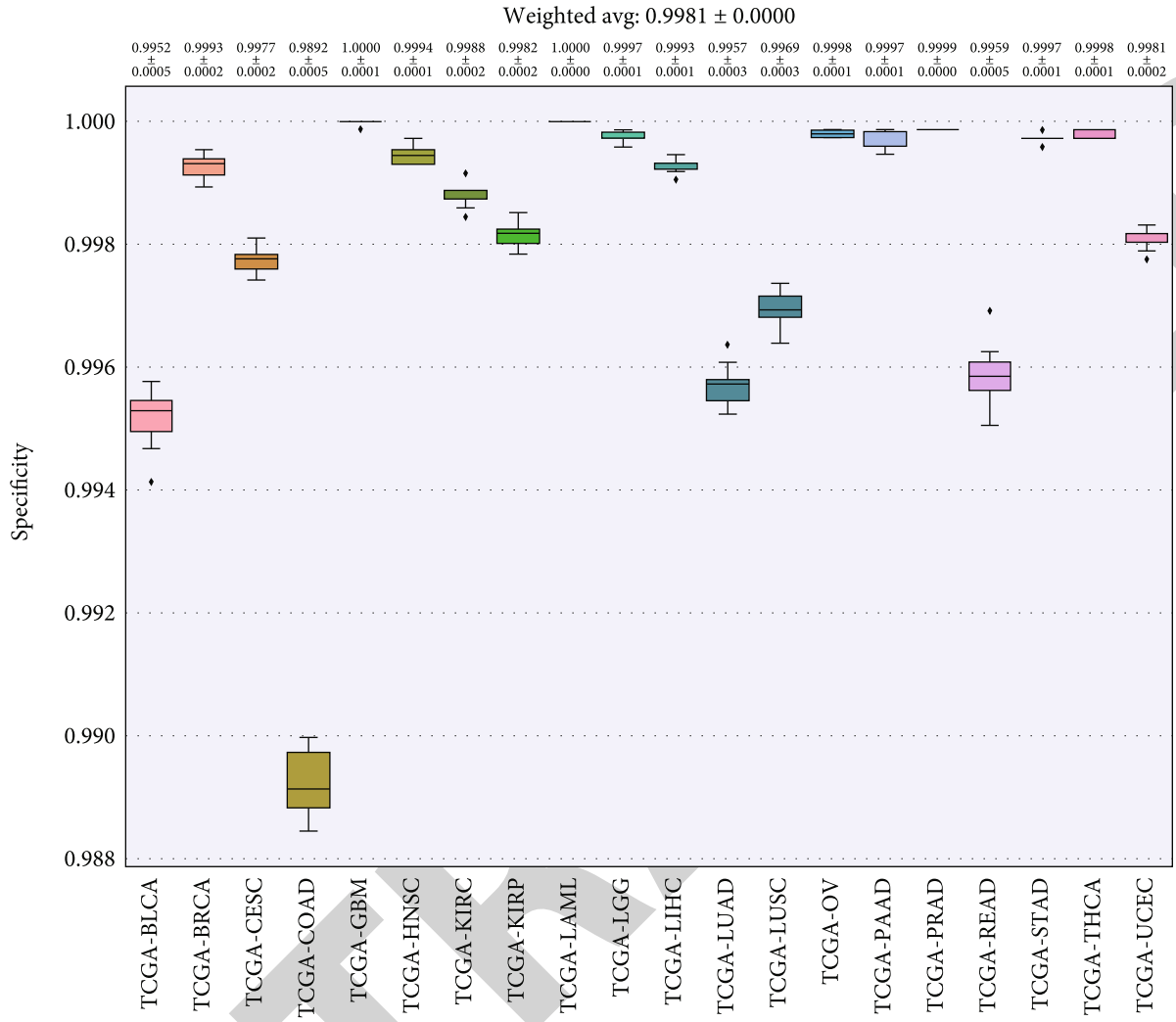


Figure 11: Specificity for each cancer.

3.5. Discussion. The treatment of cancers with unknown primary origin is mainly empirical chemotherapy, but the prognosis of patients is generally poor. A clear diagnosis directly determines the surgical method and scope, as well as the drug regimen of physicians. Because the method in this paper is based on sequencing, this approach guides medication in patients who are sequenced. For patients who are not sequenced, the next step of diagnosis and treatment should be determined according to the guidance of doctors. The diagnostic techniques primarily include comprehensive evaluation, imaging examination, pathological analysis, immunohistochemistry (IHC) panels, and genetic testing, but the treatment is less effective. The method proposed in this paper can be used to identify tumor tissue of origin, so as to provide doctors with help and appropriate drugs according to this.

We used GBDT to predict the tissue origin of the metastatic samples. GBDT can flexibly handle all kinds of data, including continuous value and discrete value. GBDT uses some robust loss functions and is relatively robust to outliers such as the Huber loss function and the quantile loss function. Because of the dependence among weak learners, it is difficult to carry out parallel training. Therefore, if the program runs too slowly with a large amount of data, it can achieve partial parallelism by adding self-sampling SGBT. The training data in this experiment was not parallel to the training data. Therefore, the results of this study might be influenced by the training method.

It was demonstrated that GBDT is a powerful method of ensemble learning. Breast cancer has a high mortality rate and is the most common cancer among women worldwide. Because of the high mortality rate of breast cancer patients, the most urgent need is to find appropriate biomarkers to determine the prognosis of breast cancer, especially BRCA (invasive breast cancer) [33]. Because basal cells like breast cancer, serous ovarian cancer, and lung squamous carcinoma have a high mRNA expression, tumors from different organs may have the same oncogenic driver events [34, 35]. Endometrial cancer is a common type of endometrial cancer, and the increase of age brings an increase in the incidence of UCES. Therefore, women aged between 45 and 65 are more likely to develop endometrial cancer than women of other ages [36, 37]. Cervical neoplasm is histologically classified like

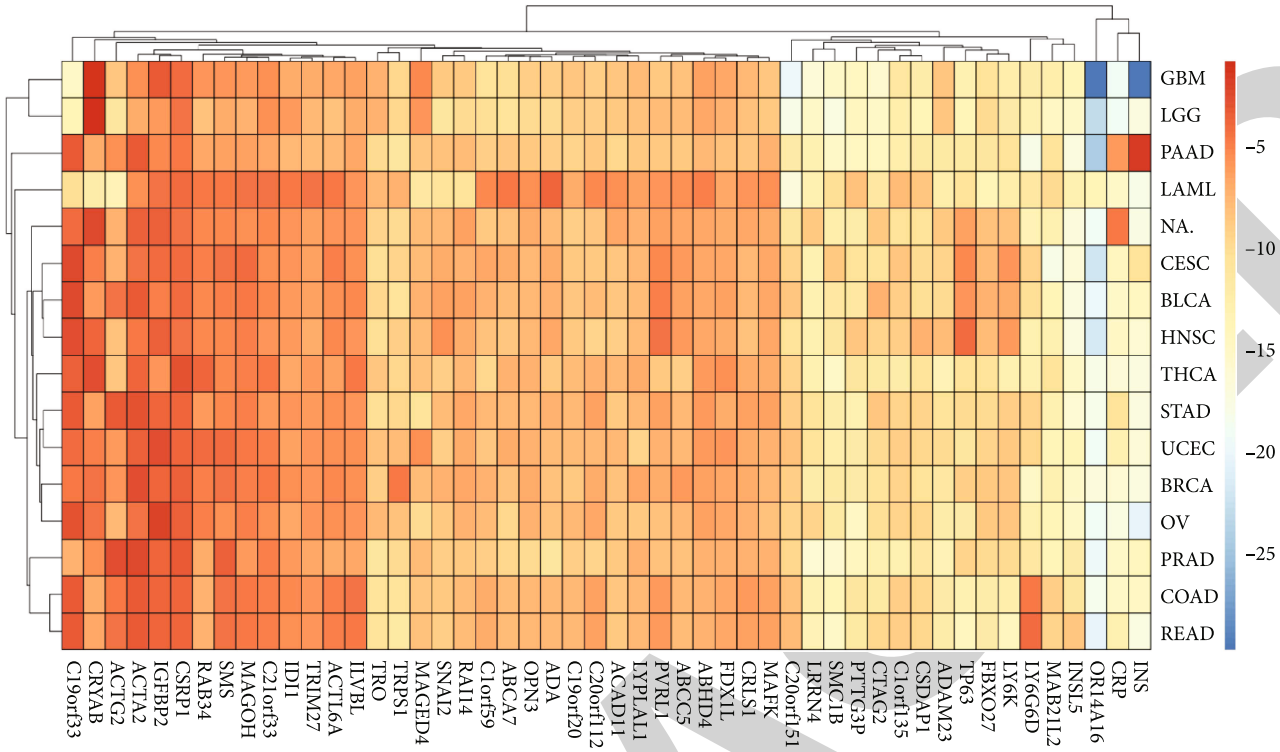


Figure 12: Average gene expression for each cancer.

squamous cell carcinoma, adenocarcinoma, and so on. Squamous cell carcinoma accounts for 85-90% of the total cervical cancer, and adenocarcinoma accounts for the rest [38].

Neither the TCGA test set nor GEO's independent test set was 100 percent accurate because a small percentage of cancers were misdiagnosed. The main reason for this error is that the two cancer species have similar characteristics and are easy to misjudge during classification, which is a key point that can be improved in the future.

Since this study was researched on gene expression profiles, it is easy to make an error prediction if the gene expressions of samples are similar. Therefore, the next step was to address this problem by increasing the sample number in types of cancer.

4. Conclusions

In conclusion, we applied a gradient boosting classifier to identify 20 tumor types based on expression profiles with a high accuracy, which might assist the pathologists in the diagnosis of cancers of unknown primary origins. Subsequent work has been to improve accuracy by increasing the number of samples of cancer types and improving methods.

Data Availability

The Cancer Genome Atlas (TCGA) RNA-seq and array data include 20,501 genes from the ICGC Data Portal (https://dcc.icgc.org/releases/release_26/) download. An independent data set was also downloaded from the Gene Expression Omnibus (GEO). These samples belong to GSE8352, GSE8734, GSE11107, GSE11132, GSE4895, GSE6491, GSE7966, GSE7766, and GSE11843.

Conflicts of Interest

There is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Nature Science Foundation of China (Grant Nos. 61863010, 11926412, 11926205, and 61873076); Natural Science Foundation of Hainan Province of China (Grant No. 119MS036); Innovative Research Projects for Graduate Students in Hainan Province (Grant No. hys2019-267).

Supplementary Materials

The file in the name of "The Selection of 400 Genes.docx" contains the names of the 400 genes selected. *(Supplementary Materials)*

References

[1] N. Pavlidis and K. Fizazi, "Cancer of unknown primary (CUP)," *Critical Reviews in Oncology/Hematology*, vol. 54, no. 3, pp. 243–250, 2005.

[2] R. D. Raji Pillai, C. Ted Rigl, J. Scott Nystrom, M. H. Miller, L. Buturovic, and W. D. Henner, "Validation and reproducibility of a microarray-based gene expression test for tumor identification in formalin-fixed, paraffin-embedded

specimens," *The Journal of Molecular Diagnostics*, vol. 13, pp. 48–56, 2011.

[3] R. W. Tothill, A. Kowalczyk, D. Rischin, A. Bousioutas, and A. J. Holloway, "An expression-based site of origin diagnostic method designed for clinical application to cancer of unknown origin," *Cancer Research*, vol. 65, no. 10, pp. 4031–4040, 2005.

[4] E. Briasoulis and N. Pavlidis, "Cancer of unknown primary origin," *The Oncologist*, vol. 2, no. 3, pp. 142–152, 1997.

[5] O. Guntinas-Lichius, J. Peter Klussmann, S. Dinh et al., "Diagnostic work-up and outcome of cervical metastases from an unknown primary," *Acta Oto-Laryngologica*, vol. 126, pp. 536–544, 2006.

[6] C. Tomuleasa, F. Zaharie, M. S. Muresan, L. Pop, and T. E. Ciuleanu, "How to diagnose and treat a cancer of unknown primary site," *Journal of Gastrointestinal Liver Diseases*, vol. 26, p. 69, 2017.

[7] X. J. Ma, R. Patel, X. Wang, R. Salunga, and M. Erlander, "Molecular classification of human cancers using a 92-gene real-time quantitative polymerase chain reaction assay," *Archives of Pathology Laboratory Medicine*, vol. 130, no. 4, pp. 465–473, 2006.

[8] K. Sheahan, J. C. O'Keane, A. Abramowitz et al., "Metastatic adenocarcinoma of an unknown primary site: a comparison of the relative contributions of morphology, minimal essential clinical data and CEA immunostaining status," *American Journal of Clinical Pathology*, vol. 99, no. 6, pp. 729–735, 1993.

[9] A. I. Su, J. B. Welsh, L. M. Sapinoso, S. G. Kern, and G. M. Hampton, "Molecular classification of human carcinomas by use of gene expression signatures," *Cancer Research*, vol. 61, no. 20, pp. 7388–7393, 2001.

[10] L. M. Weiss, P. Chu, B. E. Schroeder et al., "Blinded comparator study of immunohistochemical analysis versus a 92-gene cancer classifier in the diagnosis of the primary site in metastatic tumors," *Journal of Molecular Diagnostics*, vol. 15, no. 2, pp. 263–269, 2013.

[11] M. G. Erlander, X.-J. Ma, N. C. Kesty, L. Bao, R. Salunga, and C. A. Schnabel, "Performance and clinical evaluation of the 92-gene real-time PCR assay for tumor classification," *Journal of Molecular Diagnostics*, vol. 13, no. 5, pp. 493–503, 2011.

[12] G. Bloom, I. V. Yang, D. Boulware et al., "Multi-platform, multi-site, microarray-based human tumor classification," *The American Journal of Pathology*, vol. 164, pp. 9–16, 2004.

[13] S. Yang, "Gastric metastasis of ovarian serous cystadenocarcinoma," *International Medical Case Reports Journal*, vol. 11, pp. 201–204, 2018.

[14] S. Ramaswamy, "Multiclass Cancer Diagnosis Using Tumor Gene Expression Signatures," vol. 98, pp. 15149–15154, 2001.

[15] E. Meiri, W. C. Mueller, S. Rosenwald et al., "A second-generation microRNA-based assay for diagnosing tumor tissue origin," *The Oncologist*, vol. 17, no. 6, pp. 801–812, 2012.

[16] S. Hu, P. Chen, P. Gu, B. Wang, and H. Informatics, "A deep learning-based chemical system for QSAR prediction," *IEEE Journal of Biomedical*, vol. 24, pp. 3020–3028, 2020.

[17] J. H. Friedman, "Stochastic gradient boosting," *Computational Statistics Data Analysis*, vol. 38, no. 4, pp. 367–378, 2002.

[18] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, vol. 29, pp. 1189–1232, 2001.

[19] A. F. R. McLendon, D. Bigner, E. G. V. Meir et al., "Comprehensive genomic characterization defines human glioblastoma genes and core pathways," *Nature*, vol. 455, pp. 1061–1068, 2008.

[20] I. Düntsch and G. Gediga, "Confusion matrices and rough set data analysis," *Journal of Physics: Conference Series*, vol. 1229, 2019.

[21] M. J. Brusco and J. D. Cradit, "Graph coloring, minimum-diameter partitioning, and the analysis of confusion matrices," *Journal of Mathematical Psychology*, vol. 48, no. 5, pp. 301–309, 2004.

[22] L. E. Peterson, "*K*-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, article 1883, 2009.

[23] R. Tarter, "Valuation and treatment of adolescent substance abuse: a decision tree method," *American Journal of Drug and Alcohol Abuse*, vol. 16, pp. 1–46, 2009.

[24] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, pp. 119–139, 1997.

[25] C. Saunders, M. O. Stitson, J. Weston et al., "Support vector machine," *Computer Science*, vol. 1, pp. 1–28, 2002.

[26] M. Ashburner, C. Ball, J. A. Blake et al., "Gene Ontology: tool for the unification of biology," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.

[27] C. M. Ewing, A. M. Ray, E. M. Lange et al., "Germline mutations in HOXB13 and prostate-cancer risk," *The New England Journal of Medicine*, vol. 366, no. 2, pp. 141–149, 2012.

[28] B. Decker and E. A. Ostrander, "Dysregulation of the homeobox transcription factor gene HOXB13: role in prostate cancer," *Pharmacogenomics and personalized medicine*, vol. 7, pp. 193–201, 2014.

[29] C. Jung, "HOXB13 homeodomain protein suppresses the growth of prostate Cancer cells by the negative regulation of T-cell factor 4," *Cancer Research*, vol. 64, no. 9, pp. 3046–3051, 2004.

[30] R. Karlsson, M. Aly, M. Clements et al., "A population-based assessment of germline *HOXB13* G84E mutation and prostate cancer risk," *European Urology*, vol. 65, pp. 169–176, 2014.

[31] J. Sun, X. Cai, M. Yung et al., "miR-137 mediates the functional link between c-Myc and EZH2 that regulates cisplatin resistance in ovarian cancer," *Oncogene*, vol. 38, no. 4, pp. 564–580, 2019.

[32] Y. Zhang, Z. Li, Q. Hao et al., "The Cdk2-c-Myc-miR-571 axis regulates DNA replication and genomic stability by targeting geminin," *Cancer Research*, vol. 79, no. 19, pp. 4896–4910, 2019.

[33] Y. He, X. Li, Y. Meng, S. Fu, and H. Du, "A prognostic 11 long noncoding RNA expression signature for breast invasive carcinoma," *Journal of Cellular Biochemistry*, vol. 120, pp. 16692–16702, 2019.

[34] Cancer Genome Atlas N Network, "Comprehensive molecular portraits of human breast tumours," *Nature*, vol. 487, pp. 330–337, 2012.

[35] Q. Wang, M. Xu, Y. Sun et al., "Gene expression profiling for diagnosis of triple-negative breast cancer: a multicenter, retrospective cohort study," *Frontiers in Oncology*, vol. 9, p. 354, 2019.

[36] L. Shen, M. Liu, W. Liu, J. Cui, and C. Li, "Bioinformatics analysis of RNA sequencing data reveals multiple key genes in uterine corpus endometrial carcinoma," *Oncology Letters*, vol. 15, pp. 205–212, 2017.

Hindawi
BioMed Research International
Volume 2023, Article ID 9781948, 1 page
https://doi.org/10.1155/2023/9781948

Retraction

Retracted: A Novel Missense Variant in the Gene *PPP2R5D* Causes a Rare Neurodevelopmental Disorder with Increased Phenotype

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Yan, R. Shen, Z. Cao et al., "A Novel Missense Variant in the Gene *PPP2R5D* Causes a Rare Neurodevelopmental Disorder with Increased Phenotype," *BioMed Research International*, vol. 2021, Article ID 6661860, 7 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6661860, 7 pages
https://doi.org/10.1155/2021/6661860

Research Article

A Novel Missense Variant in the Gene *PPP2R5D* Causes a Rare Neurodevelopmental Disorder with Increased Phenotype

Lulu Yan,[1] Ru Shen,[2] Zongfu Cao,[3] Chunxiao Han,[1] Yuxin Zhang,[1] Yingwen Liu,[1] Xiangchun Yang,[1] Min Xie,[1] and Haibo Li[1]

[1]*The Central Laboratory of Birth Defects Prevention and Control, Ningbo Women and Children's Hospital, Ningbo, Zhejiang 315012, China*
[2]*Department of Laboratory Medicine of Kunming Children's Hospital, Key Laboratory of Child Critical Disease Research of Yunnan Province, Kunming, Yunnan 650034, China*
[3]*National Centre for Human Genetic Resource, National Research Institute for Family Planning, Beijing 100081, China*

Correspondence should be addressed to Haibo Li; lihaibo-775@163.com

Received 7 December 2020; Revised 1 January 2021; Accepted 1 February 2021; Published 12 February 2021

Academic Editor: Min Tang

PPP2R5D-related neurodevelopmental disorder, which is mainly caused by de novo missense variants in the *PPP2R5D* gene, is a rare autosomal dominant genetic disorder with about 100 patients and a total of thirteen pathogenic variants known to exist globally so far. Here, we present a 24-month-old Chinese boy with developmental delay and other common clinical characteristics of PPP2R5D-related neurodevelopmental disorder including hypotonia, macrocephaly, intellectual disability, speech impairment, and behavioral abnormality. Trio-whole exome sequencing (WES) and Sanger sequencing were performed to identify the causal gene variant. The pathogenicity of the variant was evaluated using bioinformatics tools. We identified a novel pathogenic variant in the *PPP2R5D* gene (c.620G>T, p.Trp207Leu). The variant is located in the variant hotspot region of this gene and is predicted to cause PPP2R5D protein dysfunction due to an increase in local hydrophobicity and unstable three-dimensional structure. We report a novel pathogenic variant of *PPP2R5D* associated with PPP2R5D-related neurodevelopmental disorder from a Chinese family. Our findings expanded the phenotypic and mutational spectrum of PPP2R5D-related neurodevelopmental disorder.

1. Introduction

Neurodevelopmental disorders (NDDs) are a group of neurological diseases including intellectual disability (ID), autism spectrum disorders (ASDs), and language disorders with common clinical presentations that can affect cognitive, social, and motor abilities [1, 2]. NDDs have a strong genetic component, but the specific genetic factors underlying these disorders are heterogeneous and remain undefined for a large proportion of affected individuals [3]. With the development of sequencing techniques, numerous pathogenic genes have been identified in patients with NDDs [4]. Recently, whole exome sequencing (WES) has evolved as an effective strategy to discover and characterize causal variants for a significant portion of NDD patients [5].

PPP2R5D is a member of the phosphatase-2A (PP2A) family of phosphatases with critical roles in development, including maintaining neurons and regulating neuronal signaling. The *PPP2R5D* gene is located at chromosome 6p21.1, encoding the B56δ regulatory subunit of PP2A, an intracellular serine/threonine phosphatase. Variants in *PPP2R5D* are associated with mild to severe global developmental delay, hypotonia, ID, ASD, and behavioral abnormalities. Among these, pronounced hypotonia with delay in gross motor skills is the most frequently occurring symptom of PPP2R5D-related neurodevelopmental disorder [6–9]. Ever since PPP2R5D-related NDD was identified by Del-Mazo in 1996, more than 100 individuals with PPP2R5D-related NDD have been described. The disorder has a relatively limited mutational spectrum with thirteen pathogenic variants

reported so far [10]. Here, we report a novel pathogenic variant in exon 5 of the *PPP2R5D* gene (c.620G>T, p.Trp207Leu) identified through trio-WES in a Chinese patient who presented hypotonia, macrocephaly, intellectual disability, speech impairment, and behavioral abnormalities. Meanwhile, some new significant clinical phenotypes possibly related to the *PPP2R5D* variants were also found.

2. Materials and Methods

2.1. The Patient General Information and Ethics Statement. Ethical approvals for the study were obtained from the Ningbo Women and Children's Hospital. Clinical information and blood specimens were obtained from the proband and related family members. The proband underwent thorough examination, including the physical growth and the behavior development assessment, MRI examination, laboratory examination, and EEG measurement. Written informed consent was provided by the proband's parents. The proband was born to nonconsanguineous parents after an uncomplicated pregnancy and delivery. Both parents of Chinese origin and the proband's brother were unaffected and healthy. There was no history of any neurodevelopmental abnormality in the family. The pedigree chart of his family is shown in Figure 1(a). The proband showed mild developmental delay (developmental quotient was 69.71), hypotonia, macrocephaly (head occipital frontal circumference was 47 cm > 2 SD), prominent forehead, and open mouth at 10 months of age (Figures 1(b) and 1(c)). He also had a rash because of an allergy to egg whites, milk, and crabs (Figures 1(d)–1(f)). All routine laboratory tests revealed normal findings, including those of blood routine examination, serum biochemical indices, and amino acid metabolic levels. The electroencephalograms were normal. Magnetic resonance imaging (MRI) revealed dysgenesis of the corpus callosum, a reduction in brain parenchyma, widening of the cerebral sulci and fissures, and enlargement of the ventricular system at 10 months of age. The cerebral structures improved visibly when reexamined at 24 months of age (Figures 1(g)–1(l)). Other symptoms included delayed walking, speech impairment, and behavioral issues such as tantrums and trouble adjusting to new situations. The proband began speaking "mom" and "dad" at the age of 22 months and walking at the age of 23 months after rehabilitation.

2.2. DNA Extraction and Whole Exome Sequencing. Genomic DNA was extracted from peripheral blood leukocytes with informed consent and hospital ethical approval, using a QIAamp DNA Blood Mini Kit (Qiagen, Valencia, CA, USA), according to the manufacturer's protocol. DNA samples from the proband and his parents were subjected to WES using a HiSeq 2500 sequencing platform (Illumina, San Diego, CA, USA). The low-quality reads were filtered out from all sequencing data using a quality score 20 (Q20); the filtered sequencing data were compared to the human reference genome (hg19) using BWA. Then, variants were called by GATK and annotated to public databases by Annovar. The minor allele frequencies (MAFs) of all known variants were also reported according to their presence in the dbSNP 1000 Genomes Project (1000 GP), Exome Aggregation Consortium, and Exome Variant Server databases. Databases such as OMIM, ClinVar, HGMD, and SwissVar were used to determine mutation harmfulness and pathogenicity where appropriate. All whole-exome variants were subjected to biological effect analysis, which included the use of programs such as SIFT, MutationTaster, PolyPhen-2, PROVEAN, Human Splicing Finder, and MaxEntScan to predict whether an amino acid substitution or indel has an important biological effect. Variants were selected after combining the clinical information and then classified as pathogenic, likely pathogenic, variant of uncertain significance, likely benign, or benign according to the American College of Medical Genetics and Genomics (ACMG) guidelines.

2.3. Validating by Sanger Sequencing. To validate the variant screened by WES, Sanger sequencing was performed for the proband and family members. Genomic DNA was amplified by PCR using the PCR primers (forward, 5′-TGTCCTCTT TGACTTCGTG-3′; reverse, 5′-TTCTTGGCTATGTTTG GC-3′), which were designed with Primer 5 software. PCR conditions were as follows: 94°C for 10 min of initial denaturation followed by 30 cycles of amplification of 30 s at 94°C, 30 s at 60°C, and 45 s at 72°C. After reacting with BigDye Terminator v3.1, the PCR products were run on an ABI 3500DX Genetic Analyzer (Applied Biosystems, Foster City, CA) and analyzed by Sequencing Analysis Software 6. After validating the variant, family segregation was performed for all individuals.

2.4. Bioinformatics Analysis and Pathogenicity Assessment. The PPP2R5D protein sequence was obtained from the NCBI Protein Database (https://www.ncbi.nlm.nih.gov/protein/) in FASTA format. Online tools PolyPhen-2, SIFT, MutationTaster, and PROVEAN were used to predict the pathogenicity of the missense variant. Multiple sequence alignment and conservative analyses were performed using MEGA software. The hydrophobic properties of wild-type and mutant PPP2R5D protein were analyzed by ProtScale (https://web.expasy.org/protscale/). The 3D structures of wild-type and mutant protein were generated using the I-TASSER online server and PyMOL tool (http://zhanglab.ccmb.med.umich.edu/I-TASSER/, http://www.pymol.org/). The variant pathogenicity was assessed in standards and guidelines recommended by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) [11].

3. Results

Trio-WES revealed 13 variants (Table 1) after filtering, among which *PPP2R5D* carried a de novo heterozygous variant (c.620G>T/p. Trp207Leu) (Figure 2(a)). The variant is located at a known variant hotspot (p. Trp207) within the B56 domain of this protein. The tryptophan at position 207 was highly conserved in different species (Figure 2(b)). By ProtScale analysis, it was found that the hydrophobicity of the protein increased because of the altered amino acid. Therefore, the change of this amino acid could generate

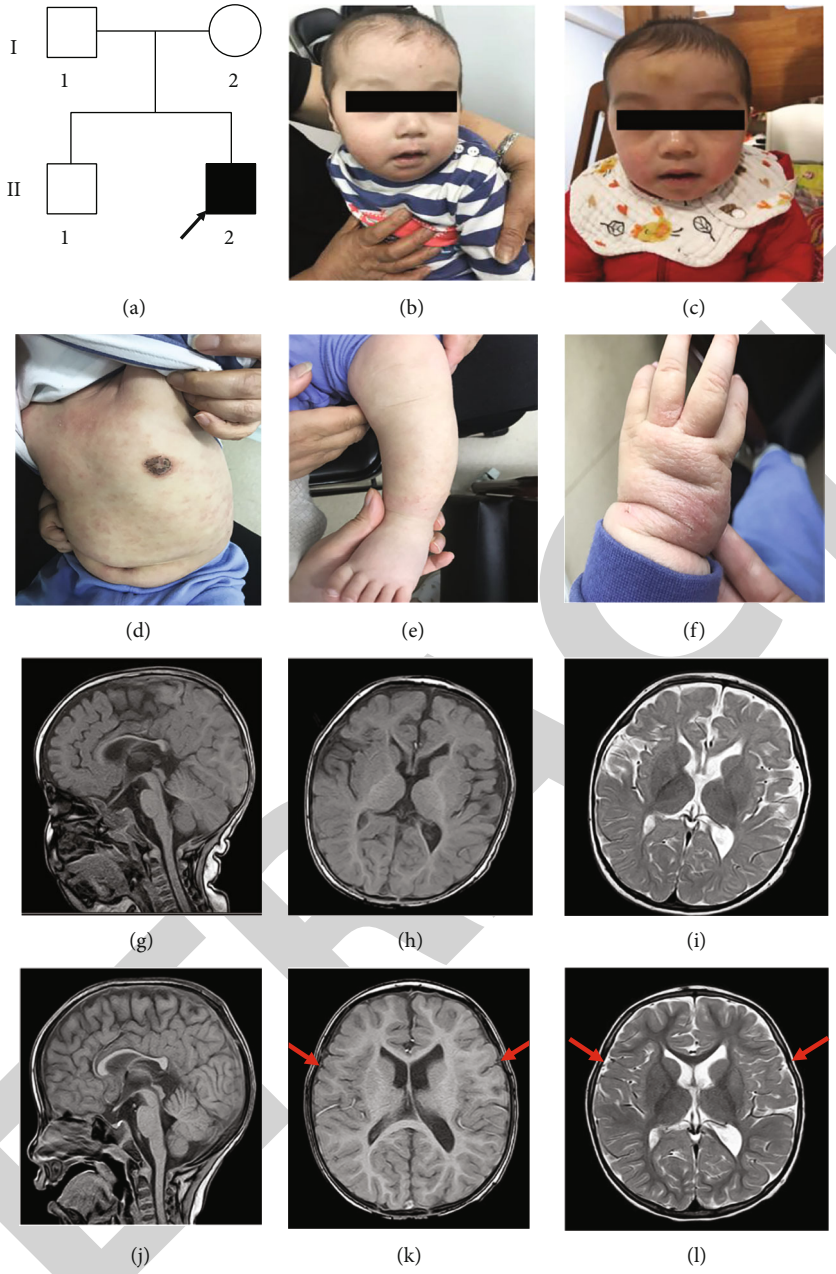


Figure 1: A Chinese boy presents NDD. (a) Family pedigree. Arrow indicates the proband (II2); squares indicate males; circles indicate females; the filled symbol indicates the affected individual. (b–f) Facial and skin features of the patient at 10/19 months of age. (b, c) The patient displays macrocephaly, prominent forehead, and open mouth. (d–f) Allergic rash. Mainly concentrated on the limbs and trunk. (g–l) Brain MRI of the proband. (g–i) MRI showed a hypoplastic corpus callosum, temporal lobe parenchymal atrophy, enlargement of the ventricular system, and retardation of myelination of the white matter. (j–l) The bilateral temporal cortex was clear and plump (red arrows).

alterations in the protein structure (Figure 2(c)). Furthermore, according to the I-TASSER tool prediction, there was a significant difference in the tertiary structural features between the wild-type and mutant PPP2R5D proteins. The Trp207Leu substitution impacted the formation of hydrogen bonds. Trp207 interacted with the Glu204 and Gln211 residues through hydrogen bonds, whereas the more hydrophobic residue leucine resulted in altered hydrogen bonding with Glu204 and Gln211 (Figures 3(a) and 3(b)). The amino acid replacement is not observed in unaffected family members and is absent from control databases, as well as 200 local controls. The majority of variant pathogenicity prediction tools supported the disease-causing consequence of this novel missense variant including a REVEL score of 0.7979. The variant is classified as pathogenic following the ACMG/AMP guidelines (PS2, PM1, PM2, PM5, PP2, and PP3) [8].

4. Discussion

PPP2R5D-related NDD is a recently characterized disorder. It is also known as Jordan's syndrome, named after Jordan Lang, the first child diagnosed in the United States. Since

Table 1: 13 variants in whole exome sequencing that were remaining after bioinformatic filtering.

Gene	Inheritance	Transcript	Exon/intron	Nucleotide alteration	Amino acid alteration	Chromosome position	Type	Disease's name (OMIM)	Proband	Father	Mother
PPP2R5D	[1]/[2]AD	NM_006245.3	Exon5	c.620G>T	p.W207L	chr6:42975031	Nonsynonymous SNV	Mental retardation, autosomal dominant 35	Het	/	/
CREBBP	[1]/[2]AD	NM_004380.2	Exon31	c.6746_6757 del	p.R2249_Q2252del	chr16:3778291-3778302:GCTGCTGCATGC:-	Nonframeshift deletion	Menke-Hennekam syndrome 1; Rubinstein-Taybi syndrome 1	Het	Het	/
PROKR2	[1]AD	NM_144773.3	Exon3	c.889G>A	p.V297I	chr20:5282952:C:T	Nonsynonymous SNV	Hypogonadotropic hypogonadism 3 with or without anosmia	Het	Het	/
UBA5	[1]AR	NM_024818.4	Exon11	c.1076G>A	p.G359D	chr3:132394712:G:A	Nonsynonymous SNV	Spinocerebellar ataxia, autosomal recessive 24; epileptic encephalopathy, early infantile, 44	Het	/	Het
YME1L1	1]AR [2]AR	NM_139312.2	Exon5	c.551A>T	p.H184L	chr10:27431366:T:A	Nonsynonymous SNV	Optic atrophy 11	Het	Het	/
NBAS	[1]AR [2]AR	NM_015909.3	Exon34	c.3938C>T	p.P1313L	chr2:15493828:G: A	Nonsynonymous SNV	Infantile liver failure syndrome 2; short stature, optic nerve atrophy, and Pelger-Huet anomaly	Het	Het	/
ARL13B	[1]AR	NM_182896.2	Exon2	c.106G>A	p.A36T	chr3:93714764:G:A	Nonsynonymous SNV	Joubert syndrome 8	Het	Het	/
ANK3	[1]AR	NM_020987.4	Exon37	c.6832A>C	p.K2278Q	chr10:61833807:T:G	Nonsynonymous SNV	Mental retardation, autosomal recessive, 37	Het	Het	/
MCEE	[1]AR	NM_032601.3	Exon2	c.275G>A	p.G92E	chr2:71351439:C:T	Nonsynonymous SNV	Methylmalonyl-CoA epimerase deficiency	Het	/	Het
UPB1	[1]AR	NM_016327.2	Exon2	c.142T>C	p.S48P	chr22:24896112:T:C	Nonsynonymous SNV	Beta-ureidopropionase deficiency	Het	Het	/
COL3A1	[1]AD [2]AR	NM_000090.3	Exon21	c.1472G>A	p.R491Q	chr2:189859788:G:A	Nonsynonymous SNV	Ehlers-Danlos syndrome, vascular type; polymicrogyria with or without vascular-type EDS	Het	/	Het
WFS1	[1]AD [2]AD [3]AR [4]AD [5]AD	NM_006005.3	Exon4	c.400G>A	p.A134T	chr4:6290798:G: A	Nonsynonymous SNV	Cataract 41; deafness, autosomal dominant 6/14/38; Wolfram syndrome 1; Wolfram-like syndrome, 54,46,autosomal dominant; diabetes mellitus, noninsulin-dependent, association with	Het	Het	/
ESCO2	[1]AR [2]AR	NM_001017420.2	Exon3	c.694C>T	p.R232C	chr8:27634519:C:T	Nonsynonymous SNV	Roberts syndrome; SC phocomelia syndrome	Het	/	Het

AD: autosomal dominant; AR: autosomal recessive; Het: heterozygous; Hom: homozygous; SNV: single nucleotide variants.

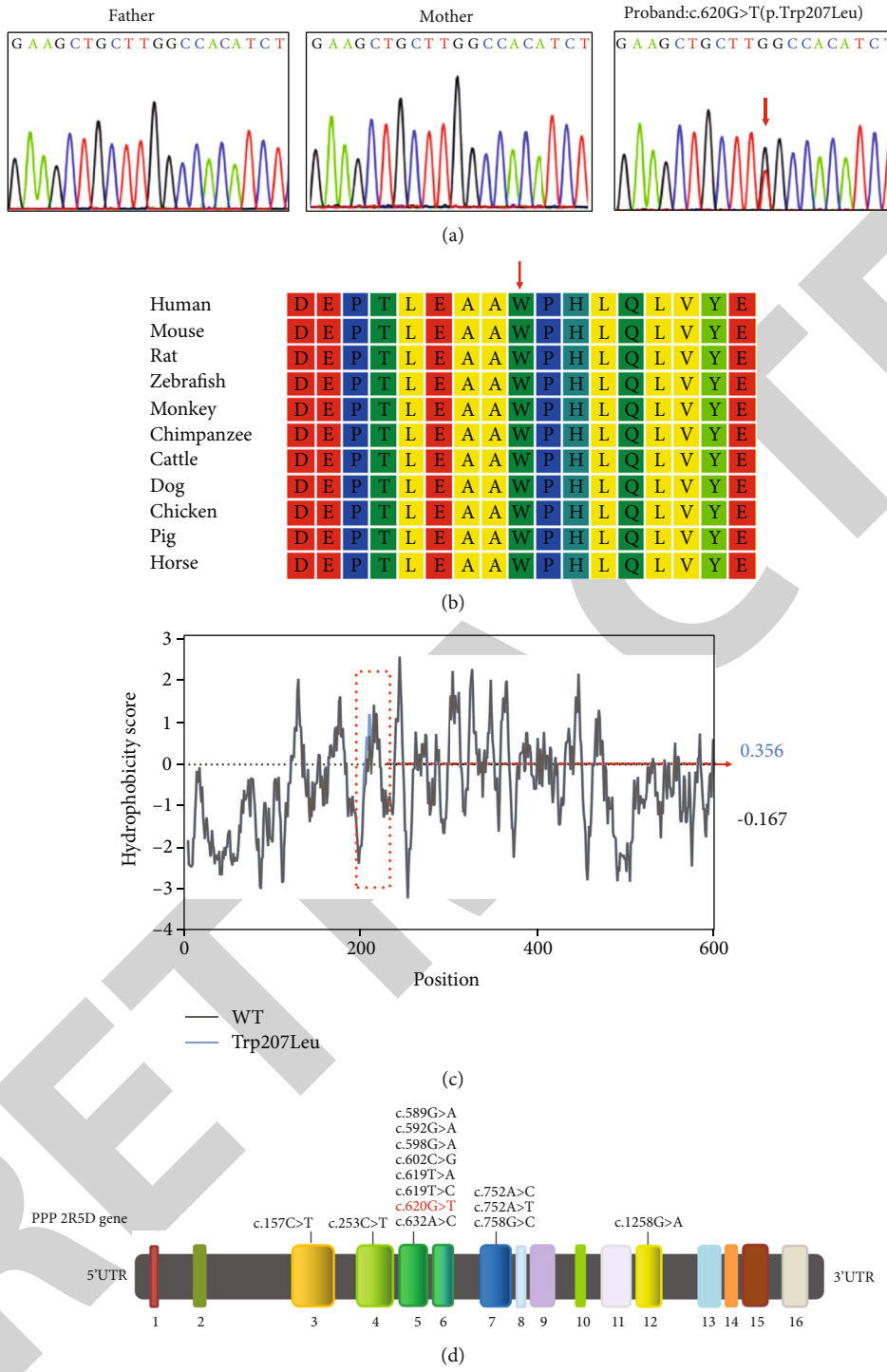
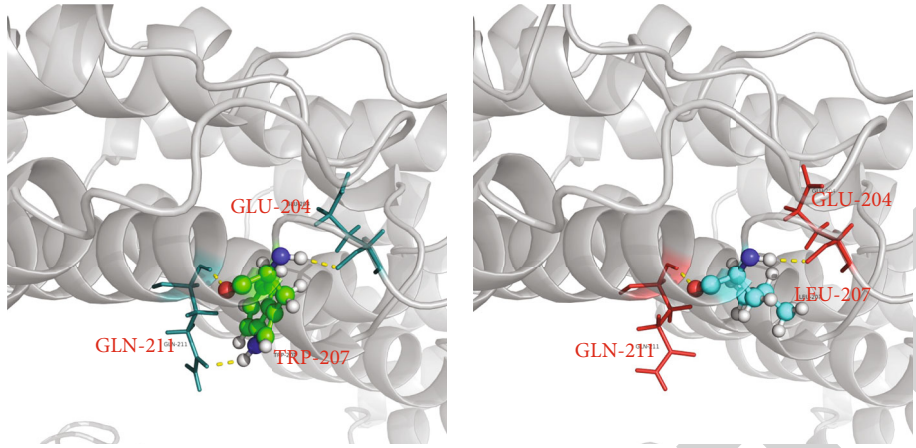


Figure 2: Identification of a missense variant in the *PPP2R5D* gene. (a) The novel heterozygous variant c.620G>T (p.Trp207Leu) in *PPP2R5D* was verified by Sanger sequencing. The variant is marked by a red arrow. (b) Protein alignment shows that the 207th tryptophan (marked by a red arrow) is highly conserved among different species. (c) Hydrophobicity of PPP2R5D protein. The prediction by ProtScale indicates an increase of local hydrophobicity around the site of p.Trp207Leu variant. WT: wild type. (d) The schematic diagram showing the gene structure of PPP2R5D. The variant in PPP2R5D occurs in exon 5, which causes G to T substitution at nucleotide 620 (NM_006245.3). Previously reported pathogenic variants are marked in black, and the variant identified in this study is highlighted in red.

then, more than 100 cases have been identified across the globe [10]. The clinical manifestations of individuals with *PPP2R5D* variants included developmental delay and/or intellectual disability (100%), macrocephaly (78.26%), hypotonia (73.91%), ophthalmologic abnormalities (38.89%), and seizures (26.09%) [12].

To the best of our knowledge, only two Chinese patients with macrocephaly and NDDs have been reported with the



(a) (b)

Figure 3: Structural modeling of PPP2R5D protein. The I-TASSER tool predicts the structural changes caused by the p.Trp207Leu variant. (a) Amino acid at position 207 of wild-type PPP2R5D protein interacted with the Glu204 and Gln211 residues through hydrogen bonds which are presented as yellow dashed lines. (b) The formation of hydrogen bonds is predictably altered by the substitution of tryptophan by leucine at position 207.

same heterozygous missense variant (c.G592>A) of the *PPP2R5D* gene [13]. We reported a third Chinese patient with a novel pathogenic variant c.620G>T (p.Trp207Leu). Both two Chinese patients diagnosed previously had similar phenotypes comparable to other patients with *PPP2R5D* variants, including moderate global developmental delay, mild ID, frontal bossing, hypertelorism, and a history of epilepsy [13]. Some phenotypes of our patient are highly consistent with other PPP2R5D-related NDD patients, such as the specific feature of macrocephaly, prominent forehead, hypotonia, delayed speech and language development, and mild ID. In terms of nonspecific brain MRI results, the overall brain development of our patient was slower than that of other children of the same age. In addition to the aforementioned clinical manifestations, our patient also exhibited a serious allergic rash of the skin, which has not been described in previous cases. It is not clear whether this clinical manifestation is part of the symptomatology of PPP2R5D-related NDD, and it could possibly represent a new clinical phenotype.

PP2A complex is a heterodimeric core enzyme, which is composed of a scaffolding subunit A, a regulatory subunit B, and a catalytic subunit C [14]. B56δ is one of the isoforms of subunit B and is highly expressed in the human brain [15]. There is evidence that PP2A-PPP2R5D holoenzyme is involved in controlling the activity of the PI3K/AKT signaling pathway that plays a role in the maintenance and functioning of neurons [8]. According to the previous report, all ID-associated *PPP2R5D* variants except Pro53Ser resulted in deficient holoenzyme formation, most likely through a dominant-negative mechanism [7].

To date, a limited variant spectrum of *PPP2R5D* involving 12 missense variants and 1 nonsense variant has been identified (https://www.ncbi.nlm.nih.gov/clinvar/), and most of the *PPP2R5D* variants are located in exon 5 (Figure 2(d)). Among these variants, the Trp207Arg pathogenic variant is positioned in a conserved acidic loop of the PPP2R5D subunit critical for PP2A-PPP2R5D holoenzyme formation. The patient with the Trp207Arg variant learned to walk at the age of 2. In addition to moderate ID, developmental delay, hypotonia, and macrocephaly, he(she) also had scoliosis, hip dysplasia, and fatigue [7]. In the present study, a de novo missense variant c.620G>T (p.Trp207Leu) in exon 5 of the *PPP2R5D* gene was successfully identified in a Chinese patient with PPP2R5D-related NDD. The Trp207 amino acid in PPP2R5D is highly conserved among 10 species based on sequence alignment. Analyses by all the variant pathogenicity prediction tools supported the disease-causing consequence of this novel missense variant. In addition, the Trp207Leu substitution impacts the formation of hydrogen bonds of the Trp207 with the Glu204 and Gln211 residues and is predicted to increase the local hydrophobicity and disrupt the PPP2R5D protein structure/function. We therefore concluded that the Trp207Leu variant in PPP2R5D was the genetic etiology underlying PPP2R5D-related NDD in the proband. The research by Houge et al. [7] demonstrated deficient holoenzyme formation of PP2A due to decreased association of the mutant PPP2R5D subunit to the subunit A or subunit C. The Trp207Leu variant in the *PPP2R5D* gene may similarly perturb interactions between subunits of the PP2A complex [16]. Further functional studies will be focused on this rare variant site for deeper understanding of its functional implications.

In conclusion, a de novo *PPP2R5D* variant was identified in a Chinese patient with NDD through WES, which expanded the mutational and phenotypic spectrum of PPP2R5D-related NDD. There are no known effective treatments for NDDs associated with *PPP2R5D* variants. However, early intervention and rehabilitation therapy are available, which helped our patient to regain some impaired abilities. Further studies are required to assess the impact of *PPP2R5D* variants on neuronal functions and to facilitate the development of novel therapeutic strategies.

Hindawi
BioMed Research International
Volume 2023, Article ID 9823193, 1 page
https://doi.org/10.1155/2023/9823193

Retraction

Retracted: A Systematic Review of Lymphangioleiomyomatosis on Diagnosis and Molecular Mechanism

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] X. Dong, L. Jin, A. Wang et al., "A Systematic Review of Lymphangioleiomyomatosis on Diagnosis and Molecular Mechanism," *BioMed Research International*, vol. 2021, Article ID 6612776, 8 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6612776, 8 pages
https://doi.org/10.1155/2021/6612776

Review Article

A Systematic Review of Lymphangioleiomyomatosis on Diagnosis and Molecular Mechanism

Xiaotong Dong,[1] Lvcheng Jin,[1] Ailan Wang,[2] Liping Wu,[3] Xintong Fan,[1] Qian Hou,[1] Tianbao Li,[2] Ruilian Zhao,[4] and Yunxiang Zhang[1]

[1]*Department of Pathology, First Affiliated Hospital of Weifang Medical University, Weifang People's Hospital, Weifang 261041, China*
[2]*Department of Science, Qingdao Geneis Institute of Big Data Mining and Precision Medicine, Qingdao, China*
[3]*Department of Respiratory, First Affiliated Hospital of Weifang Medical University, Weifang People's Hospital, Weifang 261041, China*
[4]*Department of General Practice, Affiliated Hospital of Weifang Medical University, Weifang 261031, China*

Correspondence should be addressed to Ruilian Zhao; zhaoruilian36@163.com and Yunxiang Zhang; zhangbing199592@163.com

Received 23 November 2020; Revised 19 January 2021; Accepted 28 January 2021; Published 11 February 2021

Academic Editor: Min Tang

Objective. Lymphangioleiomyomatosis (LAM) is a rare low-grade metastatic tumor; however, LAM patients were always found in young age with difficulty for diagnosis. Our study is aimed at observing the clinical characteristics of patients with lymphangiomatosis, including the clinical manifestations, imaging findings, histopathological features, and immunophenotype. *Methods*. We did a systematic review on LAM/PLAM cases, especially on male cases, and collected the clinical features and molecular mechanisms of PLAM based on previous findings. *Results*. Diagnosis criteria were summarized by combining CT scans, MRI, immunohistochemistry results, and gene sequencing results for effectively distinguishing between PLAM and similar diseases. Moreover, our study illustrated the molecular mechanism of PLAM as well as the signaling pathway involved in the disease initials. In addition, a male case was reported with differential diagnosis on the clinical manifestations, microscopic features, immunophenotypes, and genotypes. *Conclusion*. Our review will definitely improve the understanding of diagnosis and treatment in PLAM cases.

1. Introduction

Lymphangioleiomyomatosis (LAM), a rare low-grade metastatic tumor, is classified as perivascular epithelioid cell tumor in the WHO lung tumor classification [1, 2]. Kitaichi et al. studied more than 300 patients. Their average age of onset was 35 years. The pathological findings showed that the normal lung parenchyma was replaced by multiple small cysts with a diameter of 0.1 several centimeters and there were signs of smooth muscle hyperplasia around and inside the pulmonary lymphatic vessels, veins, and airways [3]. The clinical manifestations of LAM at the onset are relatively mild, most of which are progressive dyspnea, often accompanied by spontaneous pneumothorax [4], cough, chylothorax, peritoneal effusion, and hemoptysis [5]. Pulmonary lymphangioleiomyomatosis (PLAM) and axial lymphatic cystic mass called tumor LAM can lead to abdominal and pelvic lymph obstruction [6]. Tuberous sclerosis (TSC) is an inherited tumor syndrome associated with epilepsy, cognitive impairment, and multiorgan tumor formation. It has been reported that LAM occasionally occurs in patients with TSC [7]. Studies have found that LAM is often associated with angiomyolipoma in the kidneys and can increase the incidence of meningioma [8]. The rate of lung function decline in LAM patients may be 2 to 4 times higher than that in ordinary people [9, 10]. Most patients will have difficulty in breathing in daily activities and recurrent pneumothorax within 10 years after the onset of symptoms [11]. The disease

progression of LAM was not as fast as that of other cancers; however, clinical characteristics vary among patients. Respiratory failure was considered one of the main causes of the death in LAM patients; however, a prognostic approach for detecting respiratory failure has not been established [12].

At present, no specific cause of PLAM has been found and it is rare in male cases. Kazuhiro et al. reported a 17-year-old man with LAM after pneumothorax surgery. This patient was the youngest case reported, because histopathological diagnosis of pneumothorax in young men is not always possible; this case provides new and valuable clinical insights into pneumothorax in young men [13]. Another 22-year-old male patient with PLAM was reported to explore possible additional pathogenesis. The clinical, radiographic, and histopathological features of the patient were the same as those previously reported [14]. A 34-year-old man was diagnosed with a 2-year TSC and underwent transcatheter arterial embolization (TACE) for right renal hemorrhage caused by renal tumor. He was admitted to the hospital for frequent recurrent pneumothorax on the right side; chest computed tomography showed multiple small round, thin-walled cystic lesions in his both lungs. Recurrent pneumothorax is considered to be associated with PLAM disease with tuberous sclerosis (TSC-LAM). TSC-LAM is extremely rare; small bullae in early LAM may lead to missed diagnosis in some cases, so they are often diagnosed as spontaneous pneumothorax [15]. In high-resolution computed tomography (HRCT), a 37-year-old male patient had left pneumothorax and a large number of collapsed lungs with extensive parenchymal cysts in both lungs. This patient had normal phenotype and karyotype, suggesting that LAM may occur in normal male with unaffected TSC. Therefore, even if there is no sign of TSC, the possibility of LAM cannot be ruled out when a person has a diffuse cystic lung disease [16]. Kim et al. reported a man with PLAM and multiple hepatic angiomyolipoma, which led to the diagnosis of tuberous sclerosis. This unusual manifestation further broadens the clinical and pathological spectra of PLAM and hepatic angiomyolipoma. Therefore, the authors emphasized the distinction between multiple hepatic angiomyolipoma and hepatic tumor, while PLAM may be a cause of cystic lung disease even in a man [17]. Another report revealed a rare case of LAM in a male patient with tuberous sclerosis who developed pneumothorax after mechanical ventilation [18, 19]. Although these reports provide references for the diagnosis and treatment of PLAM, the specific causes of PLAM have not been discovered and few male patients have been reported. Therefore, there is an urgent need to understand the clinical manifestations and molecular mechanisms of PLAM in order to provide better diagnosis and treatment for patients.

In this study, we did a systemic review of LAM/PLAM disease in the aspects of diagnosis criteria and molecular mechanism. Furthermore, rare case of PLAM in male has been detailed presented with the clinical manifestations and immuno/genotypes serving as differential diagnosis reference for PLAM.

2. Diagnosis Criteria

In a clinical work, PLAM should be distinguished from the following diseases.

2.1. Idiopathic Pulmonary Fibrosis (IPF). The imaging manifestations of both PLAM and IPF are characterized by diffusely distributed balloons of varying sizes, but the distribution of the gas-bearing cavities of the two is obviously different. PLAM is evenly distributed throughout the lung field, while IPF is mostly distributed in the middle and lower lung fields. Moreover, PLAM often has ground glass-like density shadows, consolidation shadows, subpleural arc shadows, and honeycomb shadows on CT images. The main pathological manifestations of IPF are a distorted alveolar structure and pulmonary interstitial fibrosis, in which inflammatory cells and fibroblast foci can be detected; the normally distributed lung tissue is scattered between the lesions, showing the coexistence of new and old lesions, which is obviously different from PLAM.

2.2. Benign Metastatic Leiomyoma. Benign metastatic leiomyoma mostly manifested as nodules in the lung parenchyma, with cystic changes in the nodules, but no cystic spaces.

2.3. Lymphangioma. Lymphangioma occurs in neonates and children. Under the microscope, the lymphatic vessels are dilated and the number is increased. It contains lymph fluid and lymphocyte nests, and the boundary is clear.

2.4. Lymphangiomatosis. Lymphangiomatosis is characterized by complicated hyperplasia of lymphatic vessels. Hyperplastic lymphatic vessels often invade the lung parenchyma, mediastinum, and chest wall and often have multifocal growth. Lymphocyte stones and complex proliferation of the maze can be seen under the microscope. They are lined with flat endothelial cells, and macrophages containing blood-side proteins can be seen in part of the lung cavity. Fine-needle aspiration biopsy is a rapid, economical, and minimally invasive pathological diagnosis technique with high sensitivity and specificity. Although PLAM is relatively rare, fine-needle aspiration biopsy can still effectively identify PLAM and similar diseases by combining CT scans, MRI, immunohistochemistry results, and gene sequencing results (Figure 1).

The main imaging manifestations of PLAM are multiple (>10) thin-walled balloon-containing lesions, round or round shaped, with clear boundaries [20]. As the lesion progresses, the cystic cavity can gradually expand. The imaging findings of this case were diffuse multiple lung-containing airbags, round or quasiround, unequal in size, and with clear boundaries, which were consistent with the typical performance of PLAM. The final diagnosis of PLAM depends on histopathological examination. The typical histological characteristics are that the lesions are mostly distributed along the bronchi, blood vessels, and subpleura and that the lungs are diffused with multiple cystic cavities [21]. The nodular lesions are often located in the cyst wall. It is composed of immature smooth

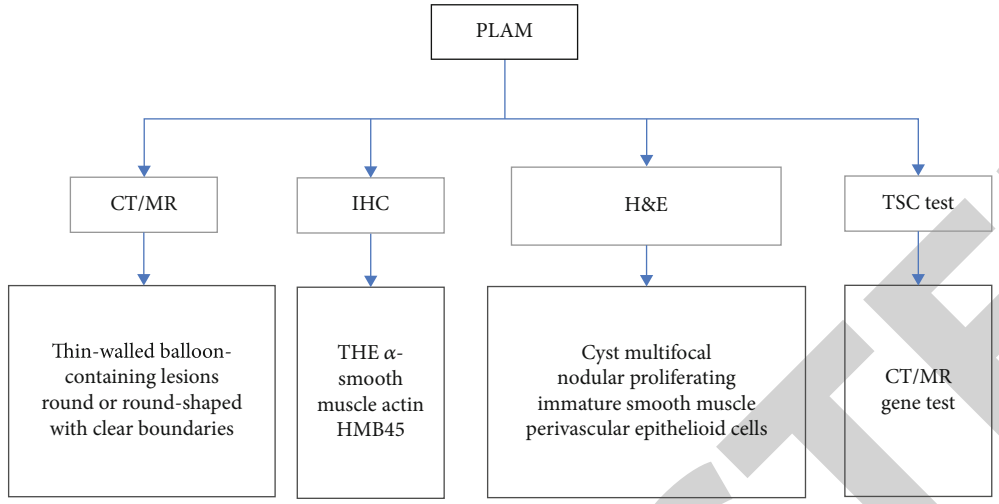


Figure 1: Tests for PLAM diagnosis process and criterion.

muscle cells and perivascular epithelioid cells, namely, LAM cells. LAM cells are rich in cytoplasm and eosinophilic and can infiltrate blood vessels and lymphatic vessels, causing pulmonary hemorrhage. Although the pathological examination of this case was a puncture tissue, the histological performance was typical. Under the microscope, some alveoli were expanded and merged to form pulmonary bullae. The blood vessels and lymphatic vessels expanded. The naive smooth muscle cells in the lung interstitial showed nodular hyperplasia. The cytoplasm is rich and eosinophilic. Immunohistochemical detection of PLAM showed that immature smooth muscle cells were stained positive for SMA, vimentin, desmin, and PCNA and epithelioid-like cells were stained positive for HMB45. In addition, LAM cells also showed positive staining for ER, PR, Melan-A, and actin [22], of which about 55%~75% were positive in perivascular epithelioid cell tumors. Serum VEGF-D levels may be associated with the clinical manifestations of PLAM, and one study reported that patients with lymphatic involvement had higher VEGF-D levels than those without lymphatic involvement [23, 24]. Most women with LAM present with cystic lung disease, and most need lung biopsy to confirm the diagnosis. Patients with a level of serum VEGF-D higher than 800 ng/L can be diagnosed as having LAM, and up to 70% of patients may avoid unnecessary invasive examination, while about 30% of patients with cystic lung diseases still need a lung biopsy for diagnosis.

Only 7 cases have been reported in the literature. Schiavina et al. reported a 17-year-old man with LAM after pneumothorax surgery (Table 1). PLAM is composed of sporadic PLAM disease and TSC-LAM. TSC and PLAM have similar molecular pathogenesis. TSC is an inherited disease caused by TSC1 or TSC2 gene mutation. Adult female TSC patients with PLAM are quite common, with an incidence rate of up to 30%–40%; the estimated incidence of sporadic PLAM disease in adult women is about 1/400,000. However, male patients with PLAM are extremely rare. Patients with TSC should receive genetic counseling before pregnancy.

3. Molecular Mechanism

PLAM disease falls into two distinct types: some patients occur sporadically named S-LAM, and others are TSC-LAM [25, 26]. In recent years, our knowledge about the disease pathogenesis of PLAM has improved substantially. In either case of LAM, the major hallmark is mutational inactivation of tumor suppressor genes TSC1 and TSC2 encoding hamartin and tuberin protein, respectively [27–29]. For TSC-LAM patients, mutants within TSC genes usually occur in germ cells as well as in somatic tissues. In contrast, the results of one research indicated that somatic mutants rather than germline mutations were identified within the TSC1 or TSC2 gene in six of nine S-LAMSC1/TSC2 which were newly identified in PLAM patients by next-generation sequencing [30, 31]. PLAM is the most common lung manifestation of TSC and the third most common cause of TSC-related death. Due to de novo germline heterozygous mutations, about two-thirds of TSC-LAM cases have no family history of LAM [32]. Compared with the TSC-LAM type, S-LAM-type PLAM usually has more severe symptoms, starting at the age of 20-30 years, and the average age at diagnosis is later (about 35 years; TSC-LAM type is 30 years). What is more is that according to the patient's statement, the onset of the disease can be presumed until the age of 10, which is the youngest reported by the male PLAM whose parents and brothers are healthy. In addition, the sporadic type is more prone to LAM and other lymphatic vessel complications [33] and the deterioration of lung function is more serious. As we all know, lung function is closely related to the prognosis of PLAM, so the prognosis of the sporadic type is worse than that of the TSC-LAM type [32]. However, our current knowledge regarding additional functional molecules involved in PLAM disease is not enough to explicit phenotypic heterogeneity between S-LAM and TSC-LAM.

The genes TSC1 and TSC2 were located on chromosomes 9q34 and 16p13.3, respectively [34, 35], and the proteins hamartin and tuberin, respectively, which form heterodimeric complex TSC1/2 to interact with dozens of proteins. One of the interacting proteins named TBC1D7 and

Table 1: A summary of the 7 cases of male pulmonary lymphangioleiomyomatosis.

Author	Age	Immunohistochemistry	TSC
Wakida K	17	SMA (+), D2-40 (+)	Yes
Kim NR	47	SMA (+), vimentin (+), desmin (+), HMB-45 (+), PR (+), ER (−), pancytokeratin (−), epithelial membrane antigen (−)	Yes
Schiavina M	37	HMB-45 (+), ER (+), PR (+)	No
Kang HW	22	HMB-45 (+), ER (+), PR (+)	Not detected
Kabi A	18	Not detected	Yes
Yamanaka S	34	SMA (+), HMB45 (+), ER (+), PGR (+)	Yes
Wang Chunyan	76	Desmin (+), CD34 (+), D2-40 (+), SMA (+), PR (+), HMB-45 (−), Melan-A (−), ER (−)	Not detected

TSC1/2 heterodimer bind together to form the TSC1-TSC2-TBC1D7 complex (TSC-TBC) [36], which regulates the mammalian target of the rapamycin (mTOR) signaling pathway in response to upstream signals such as growth factor stimulation and hypoxia. That interaction process is not directly between TSC-TBC and mTOR but instead through suppressing the activity of an intermediate protein Ras homolog enriched in brain (Rheb) by the GAP domain of tuberin [37, 38]. Rheb is a member of the GTPases superfamily and is vital in the regulation of cell growth and cycle. mTOR is conserved serine-threonine kinase which constitutes the catalytic component of two complexes named Mtorc1 and Mtorc2. The other six components of mTORC1 were regulatory-associated protein of mTOR (Raptor), proline-rich Akt substrate of 40 kDa (PRAS40), mammalian lethal with Sec13 protein 8 (mLST8), telomere maintenance 2 (TELO2), TELO2-interacting protein 1 homolog (TTI-1), and DEP domain-containing mTOR-interacting protein (Deptor) [39, 40]. mTORC1 was involved in phosphorylation of a couple of target proteins including CLIP-170, Grb10, Lipin-1, ATG1, 4E-BP, and S6K, positively regulating cell proliferation by promoting microtubule organization, lipid glycerolipid metabolism, cell autophagy, and protein synthesis [41]. Blocking the mTORC1 pathway with S6K1 inhibitor gave rise to the failed phosphorylation of eIF4B, resulting in decreases of the translation efficiency of Fra1 and subsequent induction of the expression of the E-cadherin transcriptional repressor zinc finger E-box-binding homeobox 1/2 (ZEB1/2) [42]. ZEB1/2 is essential to induce the process of epithelial-to mesenchymal transformation (EMT) [43] which contributes to tumor cell migration and invasion. mTORC2 was comprised by eight subunits: five of them are the same as in mTOC1, and two components Raptor and PRAS40 were absent and substituted by mammalian stress-activated protein kinase-interacting protein (mSIN1), rapamycin-insensitive companion of mTOR (Rictor), and Proter [41]. mTOC2 plays important roles in cellular metabolism, cell survival, and proliferation as well as cytoskeleton organization by phosphorylating target proteins Rho, AKT [44], PKC [44], and SGK1 [45]. Stress fibre formation was altered by RhoA activity increasing through upregulation of mTORC2 and followed by cellular migration and cell survival increase [46] (Figure 2).

PLAM disease affects almost exclusively women, and especially, it is exacerbated during pregnancy when estrogen (E2) increased. One study has confirmed that the E2-ERK2 signaling pathway contributes to LAM pathogenesis [42]. E2 binds to its acceptor ER, and proto-oncogene tyrosine-protein kinase Src (c-Src) is phosphorylated by ER, and subsequent c-Src activates Ras. Activated Ras regulates the ERK pathway, in turn triggering kinase activity of proteins Raf, MEK, and ERK [47–49]. E2-stimulated ERK promotes RNA transcription of the Fra1 gene and phosphorylation of Fra1 protein, which in turn inducted ZEB1/2. ERK-ZEB1/2 is critical for the EMT process and to control cell proliferation, migration, and invasion. In addition, ERK upregulates the mTORC1 pathway by directly interacting with or repressing the TSC1/2 complex [47]. Therefore, when drugs targeted on both the E2-ERK and mTORC signaling pathways are combined, treatment for PLAM disease might be more effective. The LAM cells in this case were mainly immature smooth muscle cells with fewer epithelioid cells, so they did not express HMB45 and Melan-A. Other immunohistochemical results were vimentin (+), desmin (+), SMA (+), D2-40 (+), ER (individual+), PR (partial+), CK (epithelial+). Immunohistochemical characteristics are consistent with literature reports.

4. Discussion

In a comprehensive literature, PLAM usually occurs in females; the difference in the patients reported in this study was that the onset occurred in men; this case was the second reported case of male sporadic PLAM. These patient clinical manifestations were not obvious, the onset was earlier, and the history was long for more than 40 years. A male patient has been mistaken for having pneumonia many times, and only simple symptomatic treatment has been carried out. In the past 20 years, imaging examinations have shown slow progress and pathological examination has not been performed to confirm the diagnosis. He missed the best time for treatment and ultimately leading to a poor prognosis.

The male patient, 50 years old, was reported to have chest tightness, asphyxia, and other symptoms more than 40 years ago without obvious reasons or inducements, and his condition worsened after activity. After taking cough and anti-inflammatory drugs, he did not improve. Except that he could not perform heavy physical labor, other activities were not restricted. Because the abovementioned symptoms have become worse recently and there is a small amount of hemoptysis, the patient was suspected of having a mass in

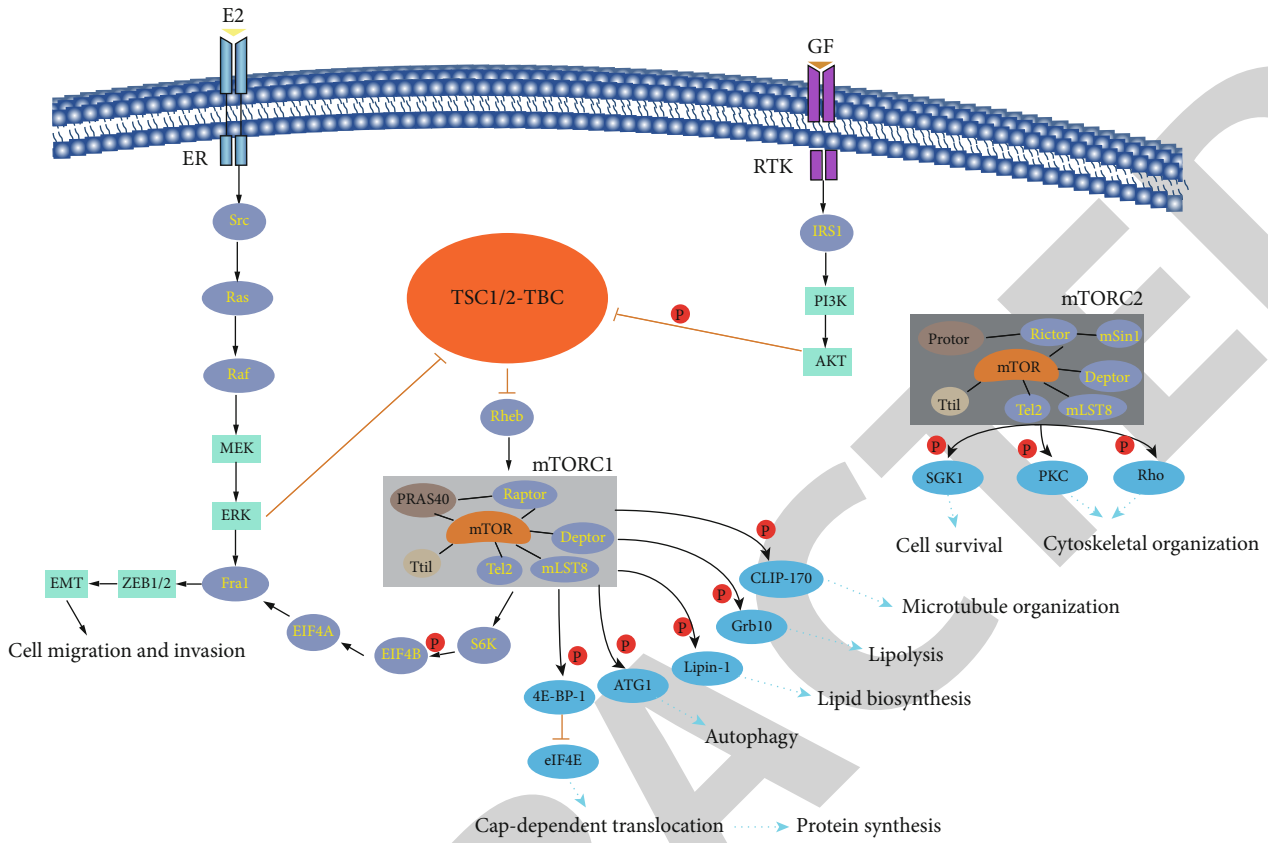


Figure 2: The signal pathway of the TSC1/2-TBC complex involved in PLAM.

his lungs and was admitted to the Department of Respiratory Medicine. A physical examination was performed when the patient was admitted to the hospital. The patient was thin, lacked energy, and had irregular breathing and movement rhythms in both lungs, with widened intercostal spaces, low breath sounds. Abnormal results of other examinations are as follows: D-dimer of 51550.00 ng/mL, prothrombin time of 15.6 s, prothrombin activity of 56.00%, and prothrombin INR value of 1.36 INR. Chest CT showed multiple air-bearing thin-walled cysts in both lungs, which are round quasicircular, with clear boundaries. Compared with the CT results in 1998, there was no obvious progress. Percutaneous lung biopsy was performed after admission.

In this case, the patient was onset in adolescence. CT scan of his chest showed multiple thin-walled round well-defined air-filled cysts. The histopathological morphology was alveolar expansion, fusion with each other to form bullae, expansion of blood vessels and lymphatic vessels, immature smooth muscle cells in the lung interstitial nodular hyperplasia, abundant cytoplasm, and eosinophilic. Immunohistochemistry shows vimentin (+), desmin (+), SMA (+), D2-40 (+), ER (+), PR (+), CK (+), Melan-A (−), and HMB-45 (−). Sanger sequencing showed no TSC1/TSC2 gene mutation (Figure 3).

The puncture tissues were routinely dehydrated in 4% neutral formaldehyde solution, embedded with paraffin, then sliced, HE stained, and Congo red stained, and the morphology of the lesion was observed under a microscope. Immunohistochemical staining was performed using the EnVision method on a Roche BenchMark XT Ventana automatic staining machine. Staining procedures were setup strictly in accordance with the staining instructions for automatic immunohistochemical detection. Positive staining photos were set in each chip as required. Desmin, Melan-A, vimentin, SMA, D2-40, ER, PR, CK, HMB-45, CD68, S-100, and NSE antibodies are rabbit anti-human monoclonal antibodies (1 : 50; Beijing Zhongshan Jinqiao, China). We performed first-generation sequencing to detect whether there are TSC1/TSC2 gene mutations, we used the QIAamp DNA FFPE Tissue Kit to extract patient DNA, we amplified the target fragments by PCR, we detected the amplification results by agarose gel electrophoresis, and we used the ABI 3130XL sequencer to detect the presence of gene mutation. The diagnosis approach is as follows:

(1) CT: the two lungs are filled with multiple gas-filled thin-walled cysts, which are round or round shaped, with varying sizes and clear boundaries (Figure 3(a1))

(2) microscopic features: alveolar expansion and fusion to form pulmonary bullae, dilation of blood vessels and lymphatic vessels, and immature smooth muscle cells in the lung interstitial nodular hyperplasia; these cells are very rich in cytoplasm and eosinophilic (Figure 3(a))

(3) immunohistochemistry and special staining: vimentin (+), desmin (+), SMA (+), D2-40 (+), ER (+), PR (+), CK (+), Melan-A (−), HMB-45 (−), CD68

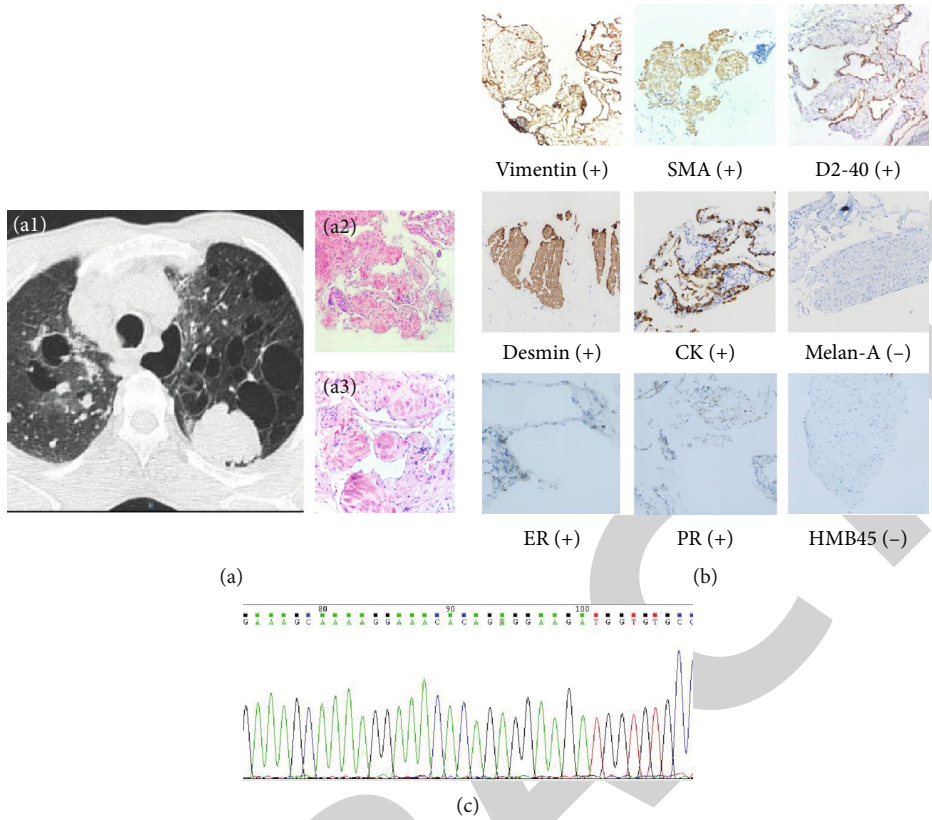


Figure 3: Pathological diagnosis in our study case. (a) CT axial position (pulmonary window), showing multiple lung types with varying sizes of multiple bullae and H&E slides for punctured lung tissue (top ×100, bottom ×400). (b) IHC staining of 9 markers in pathological diagnosis (vimentin, SMA, desmin, PR, HMB45: ×200, D2-40: ×400, ER: ×400, CK: ×100, and Melan-A: ×100). (c) Sanger sequencing for mutation detection around the TSC gene mutation region.

(−), S-100 (−), NSE (−), and Congo red stain (−) (Figure 3(b))

(4) generation sequencing: no TSC1/TSC2 gene mutation was detected in this case (Figure 3(c))

(5) diagnosis: based on clinical manifestations, microscopic features, immunohistochemical staining, special staining, genetic testing, etc., combined with imaging features, this case can be diagnosed as PLAM

(6) follow-up: the patient experienced sudden chest tightness and hemoptysis and was unable to lie down after a few hours of biopsy. There was no obvious pneumothorax on the chest radiograph. After active medical treatment, there was still repeated hemoptysis, low blood oxygen saturation, and respiratory failure. Due to severe lung lesions and a high risk of endotracheal intubation, the family of the patient refused to use endotracheal intubation ventilator-assisted ventilation treatment. The patient suddenly developed unconsciousness, respiratory arrest, and cardiac arrest symptoms on the second day after surgery. Rescue was invalid and clinical death was declared

In summary, the primary imaging findings of PLAM are multiple (>10) thin-walled balloon lesions, round or quasi-round, with well-defined boundaries. As the lesion progresses, the cystic space gradually expands. The typical histological features are bronchial, vascular, and subcolumnar lesions, with multiple cystic cavities diffused in the lungs. Immunohistochemistry of PLAM showed that immature smooth muscle cells were positively stained for SMA, vimentin, desmin, and PCNA, while epithelial cells were positive for HMB45. In addition, serum VEGF-D levels may be related to the clinical manifestations of PLAM and LAM can be diagnosed in patients with serum VEGF-D levels higher than 800 ng/L. The final diagnosis of PLAM should be based on clinical manifestations, microscopic features, immunohistochemical staining, special staining, genetic testing, etc. Our systematic review focuses on the diagnosis and molecular mechanisms of LAM and PLAM, as well as a special case study, providing detailed instruction in clinical practice and improving the understanding of potential drug discovery.

Data Availability

The data generated during the present study are available from the corresponding author on reasonable request.

Conflicts of Interest

All authors declare no conflict of interest.

Authors' Contributions

Authors Yunxiang Zhang and Ruilian Zhao conceived and designed the experiments. Xiaotong Dong, Lvcheng Jin, Liping Wu, Xintong Fan, and Qian Hou collected and analyzed the data. Xiaotong Dong, Ailan Wang, and Tianbao Li drafted the paper. All authors have read and approved this manuscript.

Acknowledgments

This work was supported by the Project of Shandong Province Medical and Health Science and Technology Development Plan (2017WS807) and Weifang Soft Science Research Project (2019RKX039).

References

[1] A. M. Taveira-DaSilva, W. K. Steagall, and J. Moss, *Lymphangioleiomyomatosis*, Springer, Berlin Heidelberg, 2011.

[2] W. D. Travis, E. Brambilla, A. G. Nicholson et al., "The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification," *Journal of Thoracic Oncology*, vol. 10, no. 9, pp. 1243–1260, 2015.

[3] M. Kitaichi, K. Nishimura, H. Itoh, and T. Izumi, "Pulmonary lymphangioleiomyomatosis: a report of 46 patients including a clinicopathologic study of prognostic factors," *American Journal of Respiratory and Critical Care Medicine*, vol. 151, no. 2, pp. 527–533, 1995.

[4] J. R. Mathieson, J. R. Mayo, C. A. Staples, and N. L. Müller, "Chronic diffuse infiltrative lung disease: comparison of diagnostic accuracy of CT and chest radiography," *Radiology*, vol. 171, no. 1, pp. 111–116, 1989.

[5] P. Crivelli, R. E. Ledda, S. Terraneo et al., "Role of thoracic imaging in the management of lymphangioleiomyomatosis," *Respiratory Medicine*, vol. 157, pp. 14–20, 2019.

[6] N. A. Avila, J. A. Kelly, S. C. Chu, A. J. Dwyer, and J. Moss, "Lymphangioleiomyomatosis: abdominopelvic CT and US findings," *Radiology*, vol. 216, no. 1, pp. 147–153, 2000.

[7] K. A. Orlova and P. B. Crino, "The tuberous sclerosis complex," *Annals of the New York Academy of Sciences*, vol. 1184, no. 1, pp. 87–105, 2010.

[8] J. Moss, R. DeCastro, N. J. Patronas, and A. Taveira-DaSilva, "Meningiomas in lymphangioleiomyomatosis," *JAMA*, vol. 286, no. 15, pp. 1879–1881, 2001.

[9] S. R. Johnson and A. E. Tattersfield, "Decline in lung function in lymphangioleiomyomatosis: relation to menopause and progesterone treatment," *American Journal of Respiratory and Critical Care Medicine*, vol. 160, no. 2, pp. 628–633, 1999.

[10] A. M. Taveira-DaSilva, M. P. Stylianou, C. J. Hedin, O. Hathaway, and J. Moss, "Decline in lung function in patients with lymphangioleiomyomatosis treated with or without progesterone," *Chest*, vol. 126, no. 6, pp. 1867–1874, 2004.

[11] S. R. Johnson, C. I. Whale, R. B. Hubbard, S. A. Lewis, and A. E. Tattersfield, "Survival and disease progression in UK patients with lymphangioleiomyomatosis," *Thorax*, vol. 59, no. 9, pp. 800–803, 2004.

[12] F. X. McCormack, N. Gupta, G. R. Finlay et al., "Official American Thoracic Society/Japanese Respiratory Society Clinical Practice Guidelines: lymphangioleiomyomatosis diagnosis and management," *American Journal of Respiratory and Critical Care Medicine*, vol. 194, no. 6, pp. 748–761, 2016.

[13] K. Wakida, Y. Watanabe, T. Kumasaka et al., "Lymphangioleiomyomatosis in a male," *Annals of Thoracic Surgery*, vol. 100, no. 3, pp. 1105–1107, 2015.

[14] H. W. Kang, C. J. Kim, S. K. Lee, K. S. Lee, C. S. Lee, and Y. H. Kim, "Pulmonary lymphangioleiomyomatosis in a male," *Journal of Korean Medical Science*, vol. 6, no. 1, pp. 83–85, 1991.

[15] Y. Sumitaka, M. Teruaki, and K. Masatoshi, "Two kinds of cystic lung lesions with pulmonary lymphangioleiomyomatosis in a male," *Annals of Thoracic & Cardiovascular Surgery*, vol. 23, no. 1, pp. 36–39, 2017.

[16] M. Schiavina, V. di Scioscio, P. Contini et al., "Pulmonary lymphangioleiomyomatosis in a karyotypically normal man without tuberous sclerosis complex," *American Journal of Respiratory and Critical Care Medicine*, vol. 176, no. 1, pp. 96–98, 2007.

[17] N. R. Kim, M. P. Chung, C. K. Park, K. S. Lee, and J. Han, "Pulmonary lymphangioleiomyomatosis and multiple hepatic angiomyolipomas in a man," *Pathology International*, vol. 53, no. 4, pp. 231–235, 2010.

[18] A. Kabi, S. Panda, S. Sama, S. Kumar, N. Kaeley, and S. Sogal P, "A rare pulmonary lymphangioleiomyomatosis disease in a male with tuberous sclerosis complex," *Respiratory Medicine Case Reports*, vol. 31, p. 101150, 2020.

[19] T. Li, X. Xu, J. Li et al., "Association of ACP1 gene polymorphisms and coronary artery disease in northeast Chinese population," *Journal of Genetics*, vol. 94, no. 1, pp. 125–128, 2015.

[20] S. R. Johnson, J. F. Cordier, R. Lazor et al., "European Respiratory Society guidelines for the diagnosis and management of lymphangioleiomyomatosis," *The European Respiratory Journal*, vol. 35, no. 1, pp. 14–26, 2010.

[21] J. Li, X. Liu, H. Chu et al., "Specific dephosphorylation of Janus Kinase 2 by protein tyrosine phosphatases," *Proteomics*, vol. 15, no. 1, pp. 68–76, 2015.

[22] L. M. Moir, "Lymphangioleiomyomatosis: current understanding and potential treatments," *Pharmacology & Therapeutics*, vol. 158, pp. 114–124, 2016.

[23] L. R. Young, Y. Inoue, and F. X. McCormack, "Diagnostic potential of serum VEGF-D for lymphangioleiomyomatosis," *New England Journal of Medicine*, vol. 358, no. 2, pp. 199-200, 2008.

[24] L. R. Young, R. VanDyke, P. M. Gulleman et al., "Serum vascular endothelial growth factor-D prospectively distinguishes lymphangioleiomyomatosis from other diseases," *Chest*, vol. 138, no. 3, pp. 674–681, 2010.

[25] G. D. Strizheva, T. Carsillo, W. D. Kruger, E. J. Sullivan, J. H. Ryu, and E. P. Henske, "The spectrum of mutations in TSC1 and TSC2 in women with tuberous sclerosis and lymphangiomyomatosis," *American Journal of Respiratory and Critical Care Medicine*, vol. 163, no. 1, pp. 253–258, 2001.

[26] T. A. Smolarek, L. L. Wessner, F. X. McCormack, J. C. Mylet, A. G. Menon, and E. P. Henske, "Evidence that lymphangiomyomatosis is caused by TSC2 mutations: chromosome 16p13 loss of heterozygosity in angiomyolipomas and lymph nodes from women with lymphangiomyomatosis," *The American Journal of Human Genetics*, vol. 62, no. 4, pp. 810–815, 1998.

[27] J. Yu, A. Astrinidis, and E. P. Henske, "Chromosome 16 loss of heterozygosity in tuberous sclerosis and sporadic

Hindawi
BioMed Research International
Volume 2023, Article ID 9819525, 1 page
https://doi.org/10.1155/2023/9819525

Retraction

Retracted: A Correlation Analysis between SNPs and ROIs of Alzheimer's Disease Based on Deep Learning

BioMed Research International

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Zhou, L. Hu, Y. Jiang, and L. Liu, "A Correlation Analysis between SNPs and ROIs of Alzheimer's Disease Based on Deep Learning," *BioMed Research International*, vol. 2021, Article ID 8890513, 13 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 8890513, 13 pages
https://doi.org/10.1155/2021/8890513

Research Article

A Correlation Analysis between SNPs and ROIs of Alzheimer's Disease Based on Deep Learning

Juan Zhou, Linfeng Hu, Yu Jiang, and Liyue Liu

School of Software, East China Jiaotong University, Nanchang 330013, China

Correspondence should be addressed to Liyue Liu; lly_nwpu@163.com

Received 20 September 2020; Revised 23 December 2020; Accepted 27 January 2021; Published 9 February 2021

Academic Editor: Min Tang

Motivation. At present, the research methods for image genetics of Alzheimer's disease based on machine learning are mainly divided into three steps: the first step is to preprocess the original image and gene information into digital signals that are easy to calculate; the second step is feature selection aiming at eliminating redundant signals and obtain representative features; and the third step is to build a learning model and predict the unknown data with regression or bivariate correlation analysis. This type of method requires manual extraction of feature single-nucleotide polymorphisms (SNPs), and the extraction process relies on empirical knowledge to a certain extent, such as linkage imbalance and gene function information in a group sparse model, which puts forward certain requirements for applicable scenarios and application personnel. To solve the problems of insufficient biological significance and large errors in the previous methods of association analysis and disease diagnosis, this paper presents a method of correlation analysis and disease diagnosis between SNP and region of interest (ROI) based on a deep learning model. It is a data-driven method, which has no obvious feature selection process. *Results*. The deep learning method adopted in this paper has no obvious feature extraction process relying on prior knowledge and model assumptions. From the results of correlation analysis between SNP and ROI, this method is complementary to other regression model methods in application scenarios. In order to improve the disease diagnosis performance of deep learning, we use the deep learning model to integrate SNP characteristics and ROI characteristics. The SNP feature, ROI feature, and SNP-ROI joint feature were input into the deep learning model and trained by cross-validation technique. The experimental results show that the SNP-ROI joint feature describes the information of the samples from different angles, which makes the diagnosis accuracy higher.

1. Introduction

Alzheimer's disease (AD) is a disease of brain tissue defect, which is manifested by cognitive impairment, memory decline, comprehension, and judgment impairment or loss [1]. Mild cognitive impairment (MCI) is considered an early stage of AD. Without scientific intervention and treatment, early patients with AD or MCI will continue to deteriorate, seriously affecting their quality of life and the development of society. With the implementation of the Human Genome Project (HGP), in recent years, the interdisciplinary application of mathematics, computer science, and biology has formed Bioinformatics. It converts genes, proteins, and other biological molecules into digital signals and then uses information science methods to process and analyze the information [2–7], so as to understand the pathogenesis of diseases. The pathogenesis of AD is complex and may be related to many concomitant diseases, age, and other factors. Imaging genetics is the study of the relationship between brain image variation and genetic variation, to characterize the pathogenesis of gene variation on brain structure and function. SNP is a polymorphism at the DNA level, which is the key source of the occurrence and development of AD. Magnetic resonance imaging (MRI) technology has been proved to be an effective method for the detection of a variety of mental diseases such as AD. The candidate brain regions that may be related to AD are called ROIs by researchers. The density, volume, and other morphological characteristics of ROIs are applied to determine whether there are abnormalities in individual brain structure or function [8]. The analysis and mining of genetic and medical data to study the pathogenesis of AD can help to improve the early diagnosis rate of AD and

provide support for the early detection and treatment of AD. At present, some methods of correlation analysis between SNP and brain ROI have been widely used to explore the pathogenesis and risk assessment of Alzheimer's disease [9, 10]. However, this strategy partially ignores the interrelationships between brain regions and may miss other important genetic variations that have not yet been reported.

In recent years, genome-wide association study (GWAS) has been applied to the study of different complex diseases globally [11, 12], and the relevant susceptible SNPs have been accurately identified and included in the GWAS Catalog [13]. With the generation of high-throughput whole-genome sequencing data, the role of data-driven genome-wide association research method on the pathogenesis of AD becomes more and more obvious [14–16]. However, with further research, it was found that the experimental results obtained by traditional GWAS are difficult to repeat, with low explanatory power and lack of heritability. Association analysis based on single variables can reveal some pathogenic loci or risk genes. For example, Westman et al. used the least square method to analyze MRI. In the experimental classification results, the accuracy of AD and normal controls was 87%, and that of MCI and normal controls was 71.8% [17]. Beheshti and Demirel calculated the Pearson correlation coefficient between the gray matter voxel features of MRI and the classification label, measured the correlation, and conducted feature ranking. By comparing different feature ranking methods, the final classification results of AD and normal controls were up to 88.8% [18, 19]. Most of the above studies are based on single-modal image data, but single-modal data usually only reflected part of the information related to brain abnormalities from a certain side, lacking statistical efficacy. The univariate association analysis ignored the weak markers, which produced significant changes by interacting with other molecules [20]. Multimodal neuroimaging data can provide complementary information and theoretically improve the accuracy of classification results. In order to systematically understand the formation mechanism of AD, multiscale, multimode, and heterogeneous data should be fused to mine the interaction between cross-omics variables [21, 22]. Some methods of data fusion based on ensemble classifier, dimension-based, and multicore learning have been proposed to establish a fusion predictor for other complex diseases. In addition, some studies have proposed improvement methods from the aspects of statistical learning [23] and regulatory relationship between SNP and gene [24].

With the development of computing hardware and the growth of data scale, the deep learning model [25] has been widely used in many application fields; for example, it has made remarkable achievements in biological and medical information processing, such as disease diagnosis [26–28]. So far, some risk genes that are significantly associated with AD have been excavated from the genomic level, but this may still be just the tip of the iceberg behind their complex genetic mechanisms. The complex interaction mechanism among genetic factors makes it difficult to understand the formation mechanism of AD, while the deep learning model has certain advantages for understanding nonlinear mapping. For AD classification, Nanni et al. first processed MRI features with different feature extraction methods to obtain multiple groups of features and then fused multiple groups of features with different combination methods to compare the differences of results generated by different methods [29]. Liu et al. used multimodal image data and deep learning network to extract depth features to achieve AD diagnosis, revealing the close relationship between changes in gray matter and AD disease [30]. Altaf et al. used Support Vector Machine (SVM), random forest, and K-Nearest Neighbor (KNN) to classify Alzheimer's disease and assigned a weight to each classifier. Finally, the classification results of each classifier were integrated and weighted [31]. Suk et al. proposed a method based on deep learning to distinguish NC (normal control) from AD, NC from MCI, and MCI from AD. However, the classification results are relatively insensitive to MCI [32].

In order to overcome the shortcomings of methods proposed by the previous researchers, we utilize a new strategy from the perspectives of feature fusion and automatic feature extraction. In this paper, we propose an analysis and diagnosis method of correlation between SNPs and ROIs based on deep learning. The method includes the following: on the hand of correlation of SNPs and ROIs, since deep learning model does not need to extract features manually, our method directly uses SNP data as input and uses the predicted value of ROIs as output to train the model; on the hand of diagnosis, first feature confusion is used on SNP data and ROI data, then a random forest algorithm is adopted for feature importance ranking, and finally a deep learning network is used to improve the performance of classification of disease state. Experimental results show that the error of our method is lower than that of other correlation analysis and diagnosis methods.

2. Methods

The study of the association between SNPs in the whole genome and ROIs in the brain region and predicting the patient's disease state is beneficial for early diagnosis and treatment of AD patients, but the current analysis and diagnosis method of the patient's disease state is almost always based on single-modal data, and the method may ignore the benefits of complementary information between SNPs and ROIs. In this paper, an analysis and diagnosis method of correlation between SNPs and ROIs based on deep learning is proposed, as shown in Figure 1, which is divided into three modules. Firstly, the SNPs and ROIs are confused with no feature information lost as much as possible. Then the random forest algorithm is used for feature selection. Finally, a deep learning network is constructed to predict the patient's disease state.

2.1. Random Forest Algorithm. Random forest (RF) is a popular ensemble machine learning method that has great application in both classification and regression tasks. Random is reflected in two aspects: the randomness of the sample and the randomness of the features. The implementation steps are as follows: firstly, the decision tree is constructed by randomly extracting part of the training set from the dataset through bootstrap

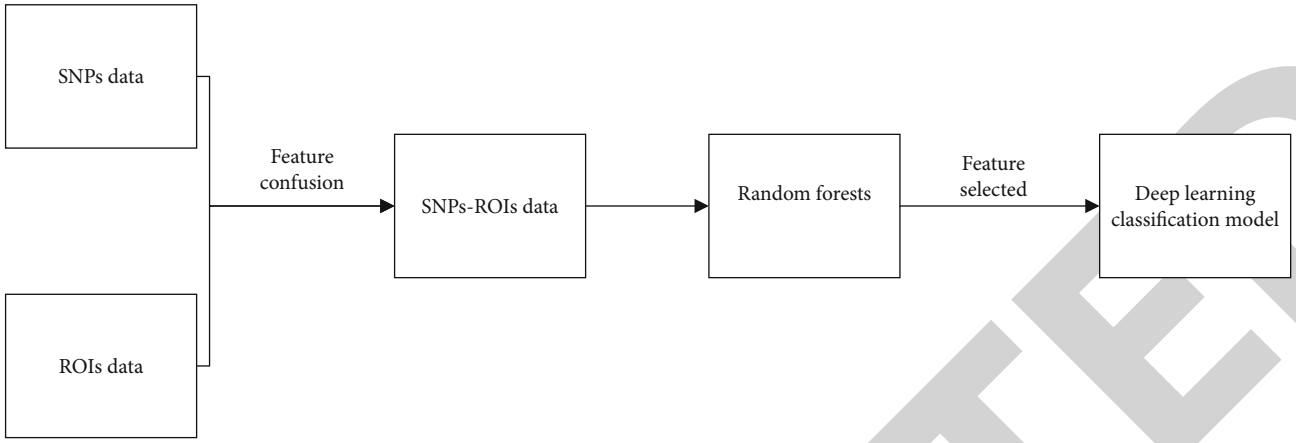


Figure 1: An analysis and diagnosis frame diagram.

technology; secondly, during the construction of the decision tree, features are randomly selected from the training set for splitting the nodes to ensure that they are the best partition. In the process of node splitting, there are usually Gini coefficient, information gain and information gain ratio to measure the goodness or badness of the partition.

RF uses only 66% of the original data to construct the decision tree. There is still about 1/3 of the data unutilized, which could be used to evaluate the performance of the decision tree and calculate the prediction error rate of the model, called out-of-bag data error. For each decision tree, select the corresponding out-of-bag data (out-of-bag (OOB)) to calculate the out-of-bag data error, noted as errOOB1. Randomly add noise interference to feature X of all samples of out-of-bag data OOB, and again calculate the out-of-bag data error, noted as errOOB2. Suppose there are N trees in the forests, then the importance of feature $X = (\sum \text{errOOB2} - \text{errOOB1})/N$. The reason why this value can illustrate the importance of the feature is that, if after adding random noise, the out-of-bag data accuracy drops significantly (that is to say, errOOB2 goes up), it means that this feature has a great impact on the prediction results of the sample. This in turn indicates a higher level of importance. On the basis of feature importance, calculate the importance of each feature, and rank them in descending order for feature selection.

2.2. Deep Learning Classification Model. With the rise of deep learning, it is now widely used in the medical field. The main manifestation is the diagnosis of diseases with the help of medical images, including the classification of diseases and the localization of lesion, early diagnosis of diseases, and screening. Deep learning originates from artificial neural networks, which are composed of multiple single-layer and nonlinear networks superimposed on each other; Deep Neural Network (DNN) relies on the relationship between layers, and each layer is a higher level of abstraction of the previous layer, which can train huge amounts of data and has the ability to learn the essential features of a dataset. Compared to traditional machine learning, deep learning has two major advantages: one is the data-driven automatic learning of features, when there are a large number of features, reducing the subjectivity and time of manual feature selection, and the second is that the model deeper than shallow models has a hierarchical structure of nonlinear features, thus contributing to better modeling of very complex data patterns. In recent years, it has also received increasing attention in the classification of medical images and disease prediction. Lu et al. [33] proposed a new framework based on deep learning, which used multimodal, multiscale deep neural network to diagnose individual AD. This method had an accuracy rate of 82.4% in identifying individuals with MCI, achieved a sensitivity of 94.23% in classifying individuals clinically diagnosed as AD, and had a specificity of 86.3% in nondementia control group. To address the situation where the multimodal data are not all complete, Thung et al. [34] proposed a multitasking deep learning model. Complete MRI data, incomplete PET data, and multimodal data such as demographic information (i.e., age, gender, and education level) and genetic information were used as inputs, and then the subnet weights were updated based on the availability of each modal data section. The results showed that the method was superior to LRMC [35] and iMSF [36] and could be extended to complex imaging data. The main types of deep learning are top-down supervised learning, such as Deep Convolutional Neural Network (DCNN) and bottom-up unsupervised learning, such as Stacked Auto Encoder (SAE). Both types of learning models can be used to classify patient disease states. In this paper, we use the former.

In this paper, consider the large number of applications of deep learning networks in related fields; we build a three-layer convolutional neural network, which is divided into an input layer, an implicit layer, and an output layer. It consists mainly of convolutional layers, pooling layers, and a fully connected layer. The role of the convolutional layer is local perception, which perceives each local feature firstly and then performs a higher level of local synthesis to obtain global information. The excitation layer is a nonlinear mapping of the output of the convolutional layer. The pooling layer is mainly used for feature downscaling, compressing the number of data and parameters, reducing overfitting, while improving the fault

tolerance of the model. The fully connected layer is used to get the final output through the Softmax function. It learns features from the sample effectively and avoids complex feature extraction processes. We use the Relu activation function for the first two layers because it iterates quickly and improves its generalization ability through the drop layer, and the last layer implements the classification of patient states through Softmax activation function. Finally, we evaluated the performance of the entire model, as well as a comparative analysis against models that did not perform biometric combinations.

Compared with traditional neural network activation functions, such as Sigmoid and Tanh functions, Relu function has following advantages: bionic principle makes it excellent for feature filtering, avoiding gradient explosion and gradient loss problems and simplifying the calculation process. Therefore, the Relu function is used as the activation function in this paper, and its definition is shown in

$$f(x) = \begin{cases} 0, & \text{for } x < 0, \\ x, & \text{for } x \geq 0. \end{cases} \tag{1}$$

Softmax is used in the process of multiclassification by taking the output of multiple neurons and mapping it to the (0,1) interval, which can be understood as probability, to perform multiclassification. The sum of probabilities for all classes is 1, and the class with the highest probability is selected as the classification result. The Softmax function is used as the activation function of the fully connected layer in this paper and for the probability that the sample vector X belongs to the j^{th} classification calculated as

$$P(y = j) = \frac{e^{x^T} Wj}{\sum_{k=1}^{K} e^{x^T Wk}}. \tag{2}$$

3. Experimental Data and Evaluation Measures

3.1. ADNI Datasets. The neuroimaging program of Alzheimer's disease is the most influential of the current AD studies. ADNI (Alzheimer's Disease Neuroimaging Initiative) database (http://adni.loni.usc.edu/) is internationally one of the most widely used sources of experimental data. This study has full permission for using the dataset. The ADNI collected multimodal data such as images (MRI and PET, Positron Emission Computed Tomography), biological sample data (genetic data, cognitive tests, and blood biomarkers), and clinical statistics. MRI image data mainly reflect the changes of brain structure, including original data and preprocessed image files. PET imaging data reflect metabolic activity. Biological sample data include blood, urine, and cerebrospinal fluid (CSF), while clinical statistics consist of clinical information on each subject, including demographic, physical, and cognitive assessment data. The genetic data were sequenced by high-throughput sequencing data, and the sequencing file format provided by ANDI was VCF (Variant Call Format), BAM (Binary Alignment Map), etc. Studies have shown that genetic factors play an important role in AD. ADNI integrates genetic, imaging, and clinical data into a data platform for analysis, so as to facilitate global researchers to further study the occurrence and development mechanism of AD.

Table. 1: Sample state coding diagram.

MCI	0	1	0
CN	1	0	0
AD	0	0	1

3.2. Experimental Data Preprocessing. The dataset used in this paper contains 632 samples, each of which has 486 SNPs and 56 ROIs. The evaluation index mainly adopts RSME (Root Mean Squared Error) and so on. Through analysis, it is found that there is a big difference between SNP and ROI data and there is a big difference in the value and range of result data among different ROIs in ROI data (for example, some ROIs are between -1000 and -100, while others are between 100 and 1000). Therefore, it is necessary to carry out normalization preprocessing to keep the data in the same range. The normalization preprocessing not only speeds up iterative convergence but also improves the accuracy. These advantages will be explained in the experimental results of correlation analysis.

Then, in this study, SNP data and ROI data were used as input to construct a classification model to predict the disease status of the samples (CN, MCI, and AD).

$$\left[\begin{array}{ccccccccc} s_{11} & \cdots & \cdots & s_{1p} & r_{11} & \cdots & r_{1q} & y_1 \\ s_{21} & & & \vdots & \vdots & & \vdots & \vdots \\ \vdots & & & \vdots & \vdots & & \vdots & \vdots \\ \underbrace{s_{n1} \quad \cdots \quad \cdots \quad s_{np}}_{X_{n*p}} & & & & \underbrace{r_{n1} \quad \cdots \quad r_{nq}}_{R_{n*q}} & & & \underbrace{y_n}_{Y_{n*1}} \end{array}\right], \tag{3}$$

$$\min \sum_{i=1}^{n} L(y_i, f(S)) \tag{4}$$

$$s.t. \min |S|, \tag{5}$$

where X_{n*p} matrix represents the SNP site value, where n represents the number of samples, p represents the number of SNPs, $S_{ij} \in \{0, 1, 2\}$, "0" represents the wild homozygous type, "1" represents the heterozygous type, and "2" represents the mutant homozygous type; R_{n*q} represents the ROI matrix, where n represents the number of samples, q represents the number of ROI, and its value is a continuous real number; and Y_{n*1} represents the label column of the sample. Equation (3) indicates that the task of multiclassification is to find a minimum SNP and ROI set S, and the accuracy of sample classification is the highest, in which L function is 0-1 loss function. The dataset adopted in this paper contains 632 samples, each of which has 486 SNPs and 56 ROIs, so there are 542 features. In order to improve the training efficiency

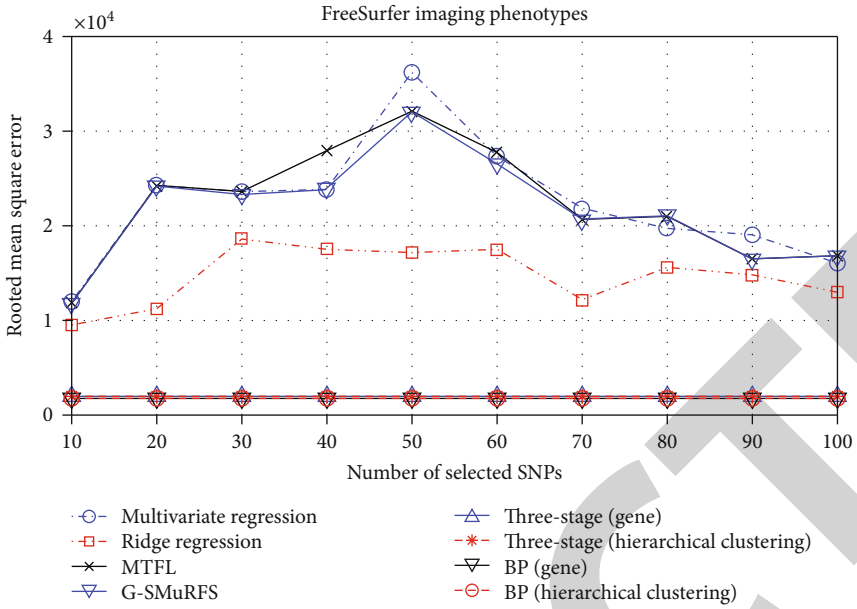


Figure 2: The comparison results of RMSE on nonnormalized data.

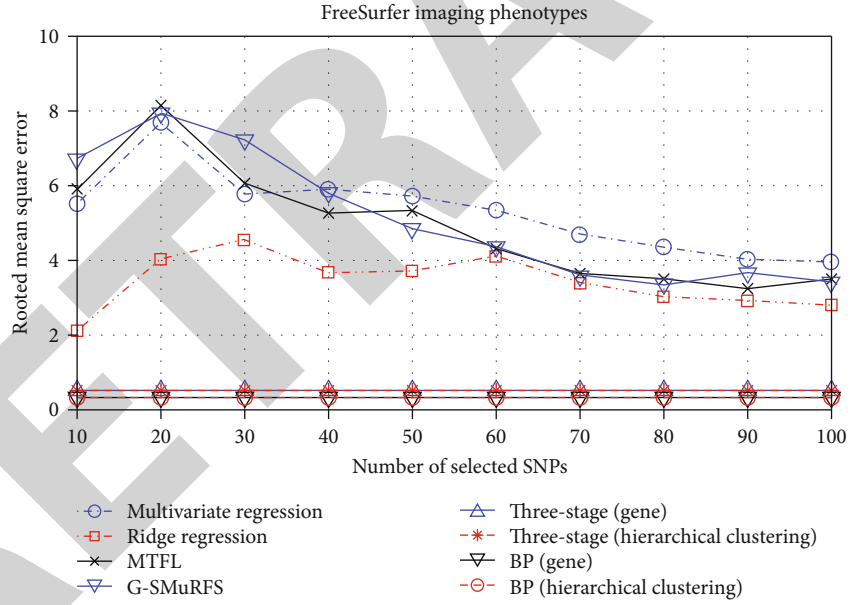


Figure 3: The comparison results of RMSE on normalized data.

of deep learning, the sample label coding method shown in Table 1 is adopted in this paper.

3.3. The Evaluation Index. The Receiver Operating Characteristic (ROC) curve has two main functions: (1) model selection—the best signal detection model, discard the second-best model; and (2) parameter setting—set the optimal threshold in the same model. In order to evaluate the performance of our method, the ROC curve is used to demonstrate the multiclassification performance of deep learning. The horizontal and vertical axes of ROC curve are FPR (false positive rate) and TPR (true positive rate), respectively.

$$\begin{aligned} \mathrm{TPR} &= \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \\ \mathrm{FPR} &= \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}. \end{aligned} \tag{6}$$

TP represents true positive, FN represents false negative, FP represents false positive, and TN represents true negative.

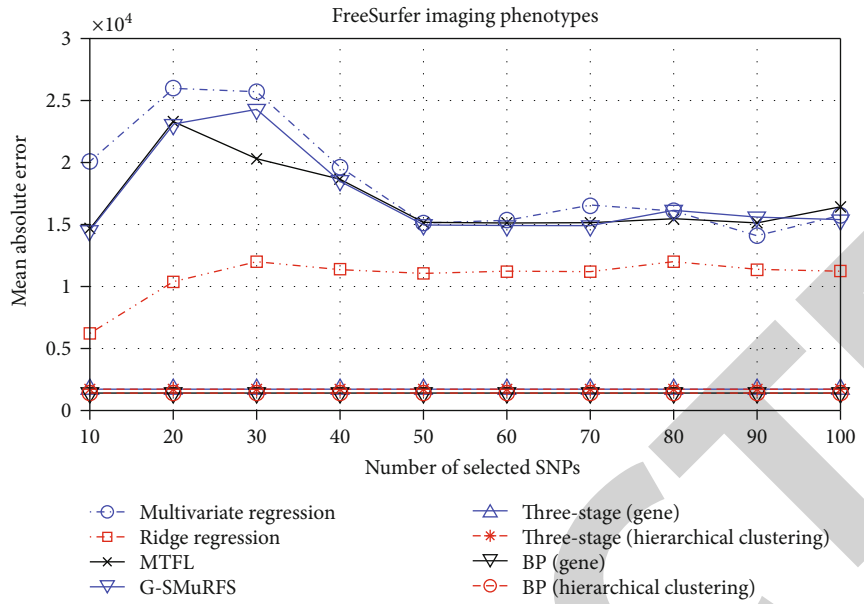


Figure 4: The comparison results of MAE (Mean Absolute Error) on nonnormalized data.

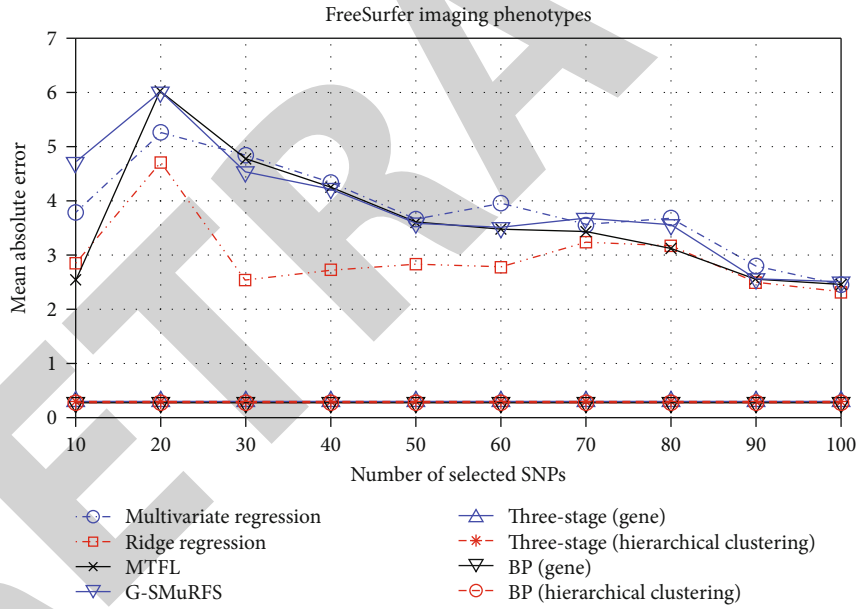


Figure 5: The comparison results of MAE on normalized data.

In the process of model training, 5-fold cross-validation method is adopted; that is, 80% subset samples in the data set are randomly selected as the training data, and the remaining 20% subset samples are used as the test data.

4. Experimental Results

4.1. Correlation Analysis Results of SNP and ROI Based on Deep Learning

4.1.1. Comparison of Normalized Pretreatment Results. To demonstrate the superiority of the proposed method, the deep learning method is compared with the three-stage method and group sparse model. Figures 2 and 3, respectively, show the ROI prediction results of various prediction models on RMSE indexes before and after data preprocessing. It can be found that the method is very similar in performance, but the method in this paper has no artificial feature extraction process.

Next, the normalization method is used to preprocess ROI data, and the results are shown in Figure 3. It can be found that the normalized pretreatment is beneficial to improve the efficiency of the regression method. The RMSE of the various methods decreases by several orders of magnitude, which also

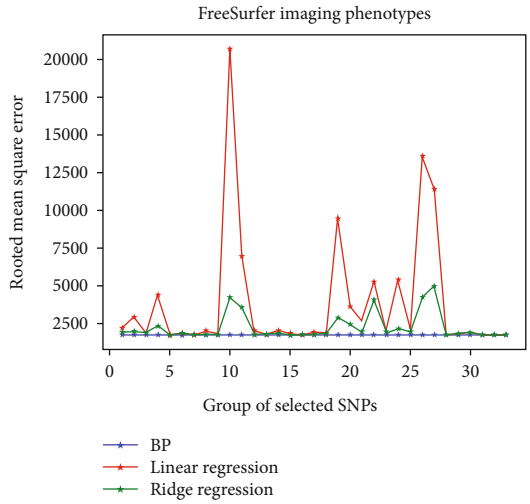


Figure 6: The regression results of feature group.

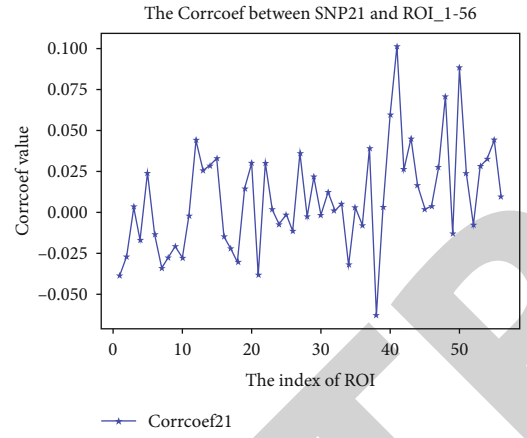


Figure 7: Correlation coefficient between SNP21 and all ROIs.

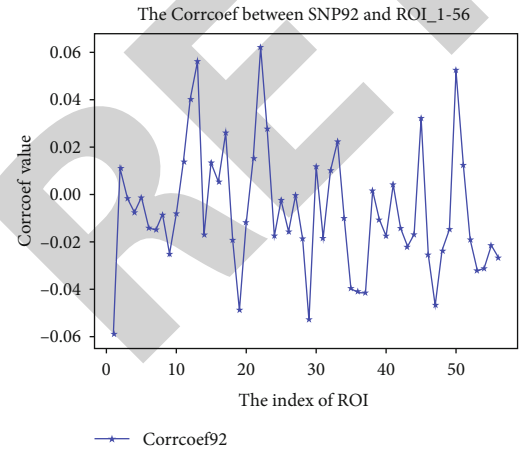


Figure 8: Correlation coefficient between SNP92 and all ROIs.

demonstrates the improvement of the model error by normalizing the data, further confirming the necessity of normalizing the ROI data, which has a very different and wide range of values stated in the analysis phase. The previous six methods

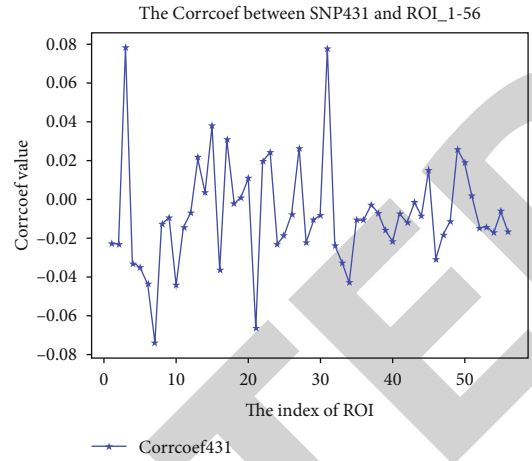


Figure 9: Correlation coefficient between SNP431 and all ROIs.

are proposed by previous researchers, while the last two are BP (Backward Propagation) neural network methods that we proposed on the basis of deep learning grouped by gene and hierarchical clustering. The final results also showed that our BP neural network method based on deep learning retains its superiority over other methods.

Figures 4 and 5, respectively, show the MAE results before and after data pretreatment. The performance of all regression analysis methods has been improved after pretreatment, among which the ridge regression method is more significant.

Comparing our method with the remaining competing methods, we find that the BP method demonstrates an advantage in predicting ROI phenotypes on both RMSE and MAE evaluation metrics, as evidenced by smaller regression errors.

4.1.2. Correlation Analysis between SNPs and ROIs. Firstly, ridge regression was used as the primary selection for SNPs, and the importance degree of SNPs was ranked according to their regression coefficient. After that all SNPs were divided into 33 groups by using the gene grouping data, and then the three regression methods were used for each group, respectively. Their regression error results are shown in Figure 6.

It can be found from Figure 6 that the deep learning method is superior to other regression analysis methods in almost every group of data. According to the gene grouping data, the SNPs (SNP21, SNP92, SNP431, SNP328, and SNP9) of the top 5 weight coefficients were in groups 2, 10, 26, 24, and 2, respectively. Next, the Pearson correlation coefficients between these key SNPs (the first 3, SNP21, SNP92, and SNP431) and ROI are shown, respectively, as shown in

$$r = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}}. \quad (7)$$

The Pearson correlation coefficients of the above SNPs and ROIs are shown in Figures 7, 8 and 9, indicating that different SNPs are complementary to ROIs, and the same SNPs have a strong negative correlation with different ROIs, while

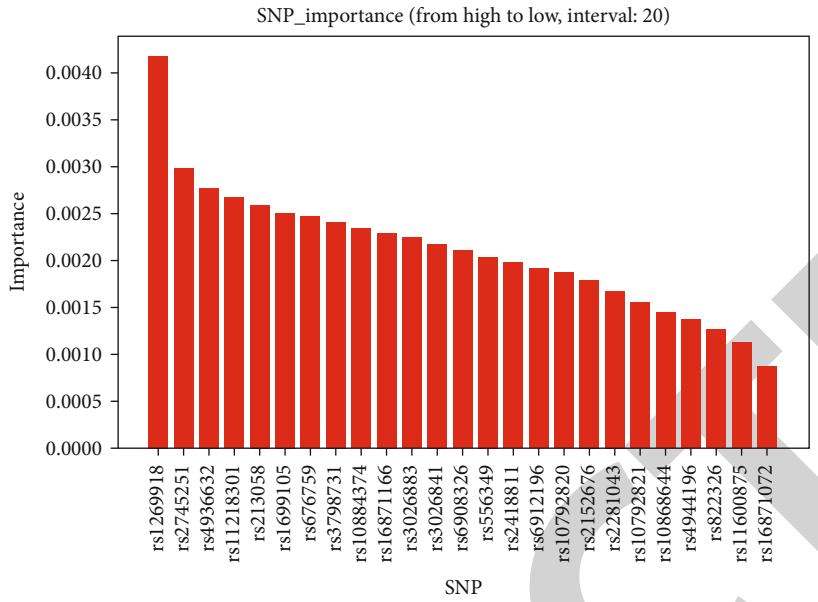


Figure 10: Correlation between SNP and sample label.

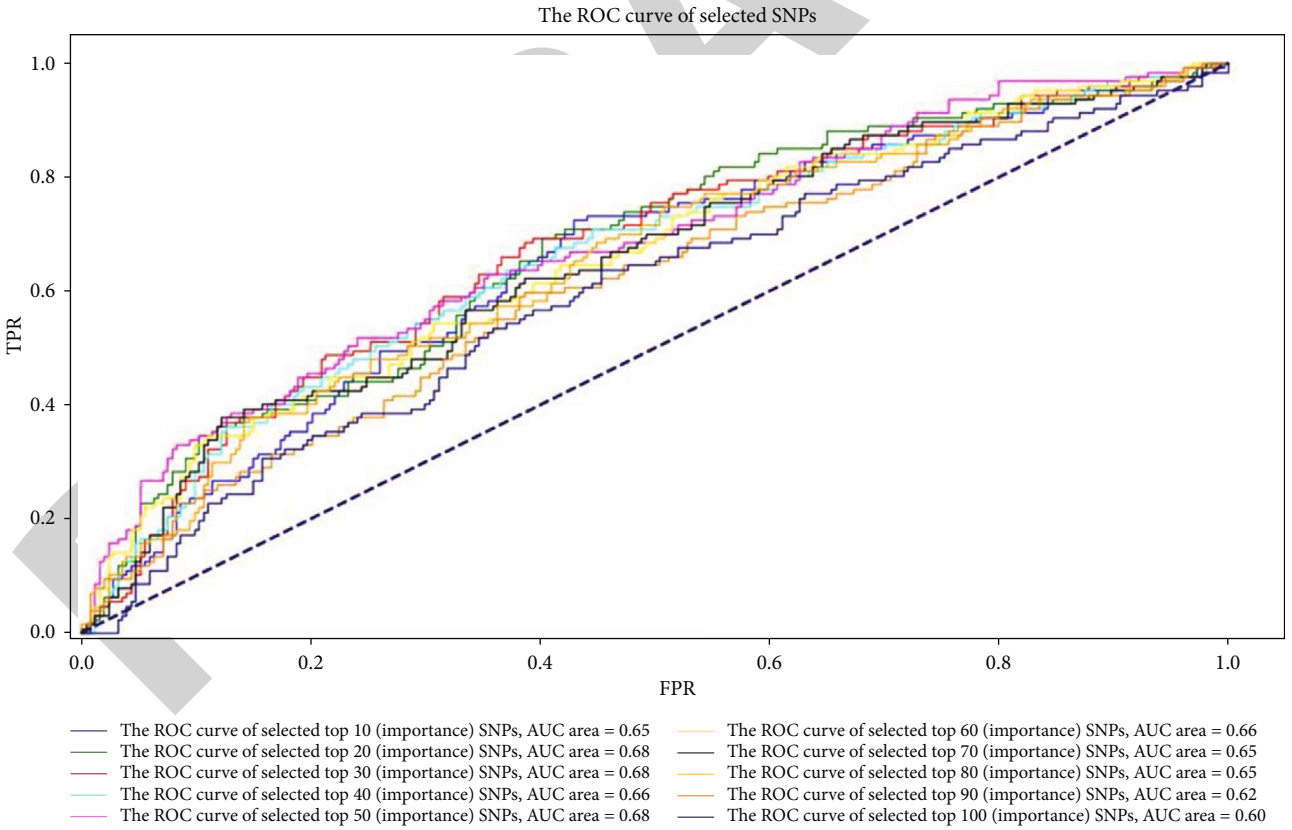


Figure 11: Multiclass classification results by SNPs.

others have a strong positive correlation. Of course, these data are statistics only, and the results depend in part on the sampling process.

4.2. Results of Disease Diagnosis Method Based on Deep Learning. To illustrate the advantages of multimodal data feature fusion, the multicategory performance is shown in

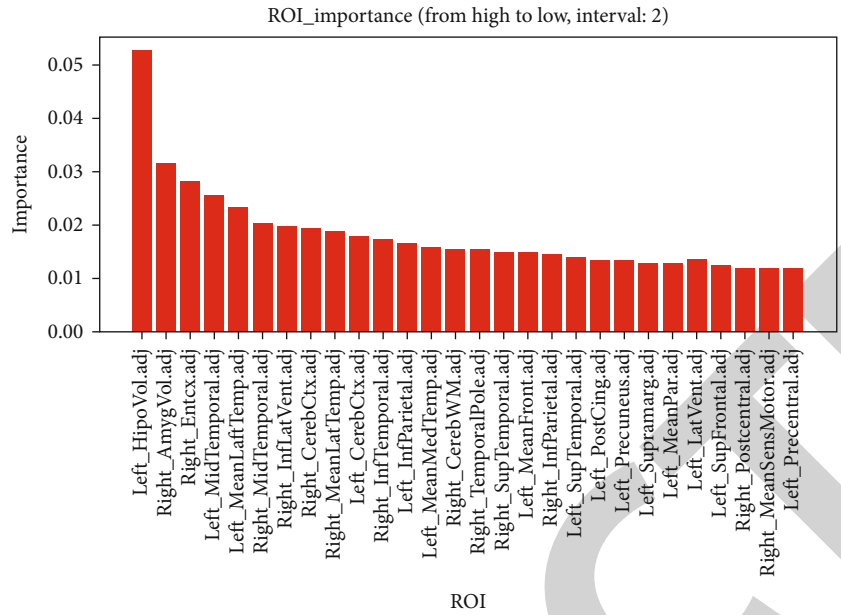


Figure 12: Correlation between ROIs and sample label.

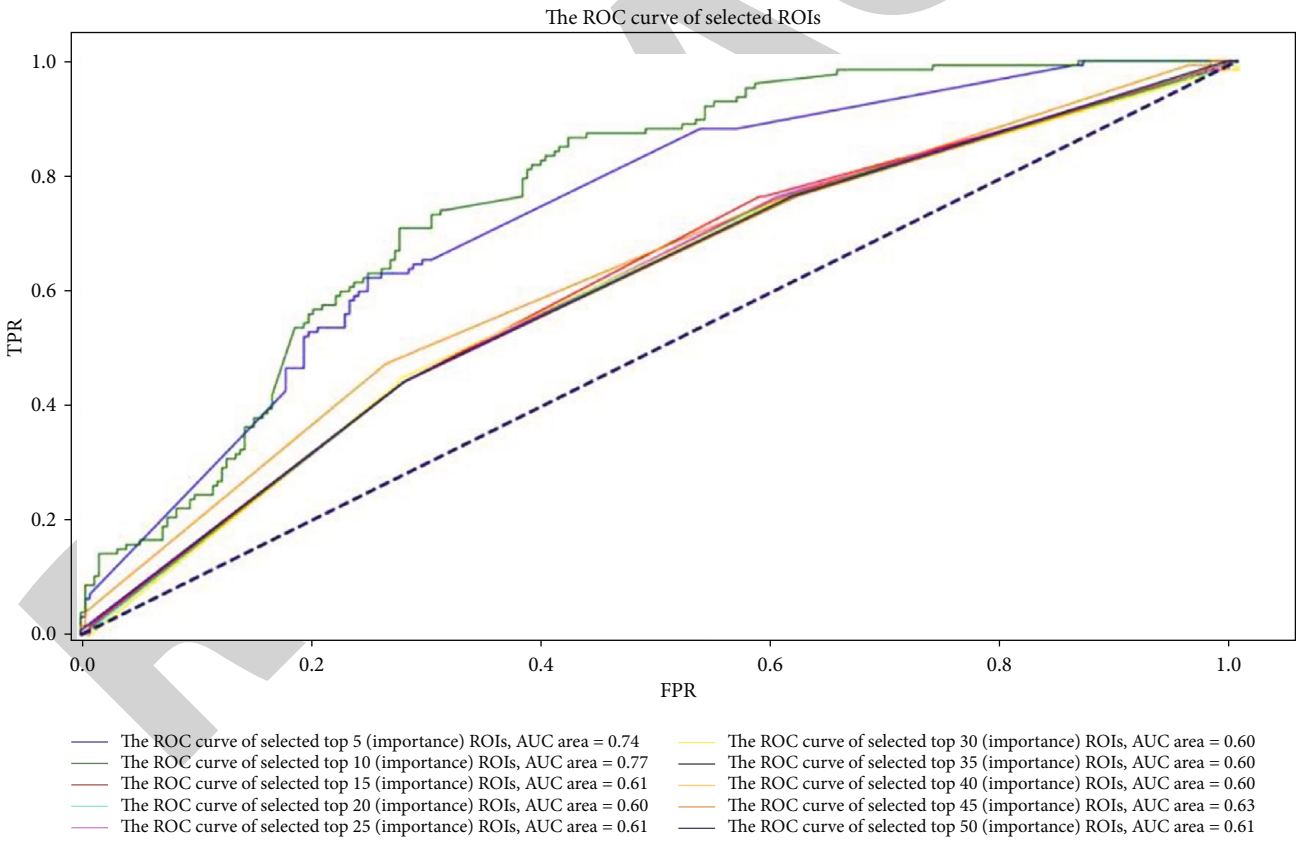


Figure 13: Multiclass classification results by ROIs.

the following three cases: prediction based on SNPs only, prediction based on ROIs, and SNP-ROI joint prediction.

4.2.1. Prediction Based on SNPs. In order to improve the training efficiency of deep learning model, the random forest method [37] was first used to evaluate the correlation between each SNP and the sample classification state, and then the correlation degree was ranked. The results are shown in Figure 10.

Due to the large number of SNPs, it is plotted at an interval of 20, and the other SNPs are omitted. The higher the correlation degree, the higher the contribution of the SNP to the

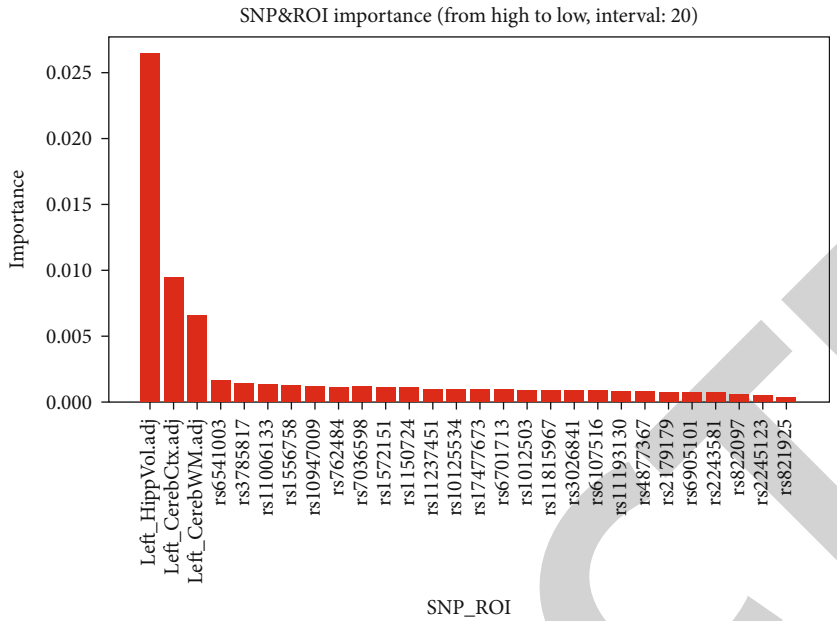


Figure 14: Correlation between ROI, SNP, and sample label.

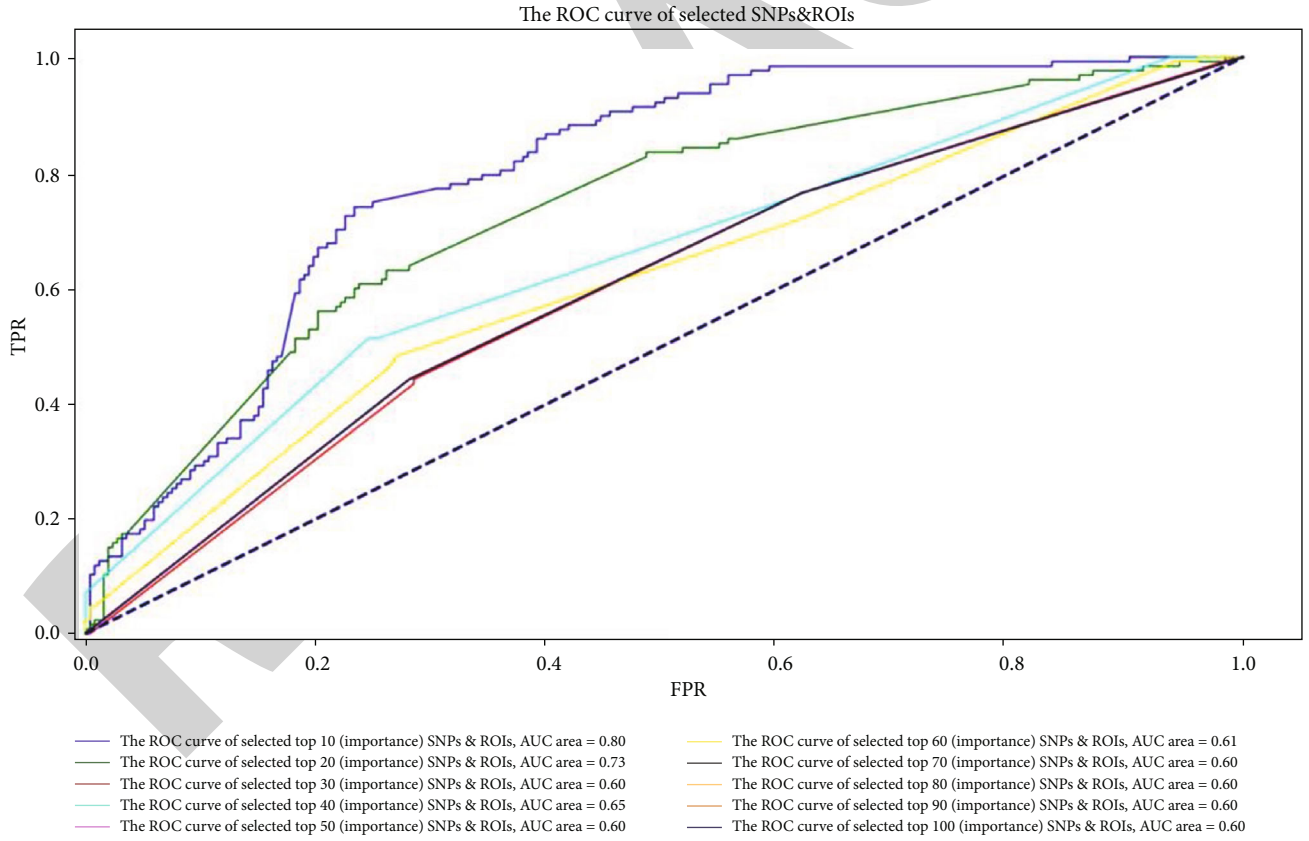


Figure 15: Multiclass classification results by ROI-SNP.

sample classification. In order to save training cost, Top K SNPs can be selected as the input of the deep learning classification model.

It can be found from Figure 11 that when feature weight is extracted from Top 10 SNPs, the effect is better than other feature combinations, and the AUC (Area Under Curve) area in the ROC (Receiver Operating Characteristic) curve is 0.6.

4.2.2. Prediction Based on ROIs. As in the previous section, the random forest method is still used to examine the degree

of correlation between a single ROI and the sample state, and all ROIs are sorted by the degree of correlation, as shown in Figure 12.

Because of the large number of ROIs, they are plotted at intervals of 2, with other ROIs omitted. Using the ROI weight ranking results generated by the random forest. Selected weight Top 5, 10, ..., 50 ROIs were used as input to the deep learning classifier model, and then 5-fold cross-validation was used for model training.

The experimental results in Figure 13 show that when the 10 ROIs in front of the weight are extracted as the feature input, the AUC area in the ROC curve reaches 0.77, and compared with Figure 10, the ROI feature is better than the SNP feature to describe the sample's disease state. This result is consistent with intuitive cognition, because ROI can directly describe the characteristics of the individual's disease, while SNP is genetic data, which is only a potential pathogenic factor for the sample's disease state.

4.2.3. ROI-SNP Jointly Predicting. ROIs can directly reflect the structural characteristics of the brain, while SNPs reflect the genetic characteristics of the sample. The former is more direct with the sample state, while the latter is a potential pathogenic factor, showing certain complementarity. Therefore, this paper intends to combine the two characteristics. Considering the combination of SNPs and ROIs, a random forest was used to calculate all the weights of SNPs and ROIs for ranking. The ranking results are shown in Figure 14.

The results in the figure reflect the above view that ROI is more directly related to the sample state. Using the weight ranking results generated by the random forest. Weight in front of 10, 20, ..., 100 SNPs and ROIs were used as joint feature input to train the deep learning classifier model, and the results are shown in Figure 15.

It can be found from Figure 15 that when the feature extraction weight ranks the top 10 SNPs or ROIs, the AUC area in the ROC curve reaches 0.8, which is better than the classification performance of SNP-only and ROI-only. The experimental results show that the combination of characteristics of different types of data is beneficial to provide complementary information, so as to obtain better sample classification accuracy.

According to the above ROC analysis results, with the increase of the number of features, the multiclassification results of various classification models show a certain degree of decline, which may be due to two reasons: (1) there is information redundancy, or even noise, between the features added later and the features added earlier, resulting in performance degradation; (2) due to the increase in the number of features, the deep learning classification model needs to consume more resources for training. If the training is insufficient, there may be underfitting of the model, resulting in performance degradation.

5. Conclusion

So far, some risk genes that are significantly associated with AD have been excavated from the genomic level, but this may still be just the tip of the iceberg behind their complex genetic mechanisms. Aiming at the problems of insufficient biological significance, large errors and inaccuracy of disease diagnosis in previous association analysis and disease diagnosis methods, we present a method of association analysis and disease diagnosis based on deep learning. Our method is a kind of data-driven method, which does not require prior knowledge to extract features manually, and the regression performance and multiclassification accuracy can also meet the application requirements. In addition, according to the experimental results of multiclassification tasks, the data fusion of complementary features is conducive to improving the accuracy of the model. In this paper, disease diagnosis can be regarded as a triad task. Each sample has three candidate states (normal, mild cognitive impairment, and AD). ROIs reflect the structural information of the brain, while SNPs reflect the genetic information of individuals, and the two information are complementary. In order to improve the disease diagnosis performance of deep learning, this paper uses the deep learning model to integrate SNP characteristics and ROI characteristics. On the experimental data set, SNP feature, ROI feature, and SNP-ROI joint feature are extracted, respectively, and these three features are input into the deep learning model, respectively, and trained by half fold cross-validation. The experimental results show that the SNP-ROI joint feature describes the information of the samples from different angles, which makes the diagnosis accuracy higher.

In this study, we proposed a correlation analysis of SNPs with ROIs and constructed a deep learning AD disease diagnostic model with SNP-ROI joint features. We uncovered a number of potentially pathogenic SNPs through correlation analysis and achieved an AUC of 80% with the SNP-ROI joint feature diagnostic model, as our model is data-driven and therefore does not rely on manually extracted features, which provides a clinical suggestion for existing AD diagnoses based on the physician's a priori judgement, and the improved diagnostic accuracy of the joint feature compared to a single feature, which gives us a research direction: firstly, the fusion of this genetic and imaging data for disease diagnosis is better than unimodal data; secondly, the physician's a priori information can be fused with other representative intermediate phenotypic features to further improve diagnostic quality.

Due to the limitation of computing resources, data, and data model, traditional image genetics research is mostly based on single mode image data. Since the single-modal brain imaging data only reflect some local information of brain structure or function, it is difficult to identify patients with early AD without obvious morphological changes. In addition, most of the studies used imaging genomics to investigate the genetic variation related to AD and only investigated the impact of genomic variation on AD. However, like other complex diseases, it is related to the interaction of multiple biomolecules. Only the analysis of omics data at a single level will make it difficult to explain the pathogenesis of AD. Therefore, we believe the following: (1) the multimodal brain image data packets contain more information than the single-modal image data packets, and the different modal image data have certain complementary information, so the establishment of multimodal brain image data fusion

analysis model is conducive to the accurate identification of early AD patients; (2) on the basis of in-depth mining of genome-wide SNP data of AD, the integration of other levels of omics data is conducive to a systematic and complete understanding of the occurrence and development process of AD; (3) the construction of biomolecular interaction network and the identification of its key feature modules are conducive to improving the performance of MCI/AD classification or risk assessment models and can also help to explain the molecular mechanism of diseases from the perspective of network modules and biological pathways; and (4) based on the powerful computing advantages of cloud platform and feature extraction advantages of deep learning model, it is helpful to carry out deep mining of AD multimode image data and multisource omics big data.

Data Availability

Data are available at http://adni.loni.usc.edu/.

Disclosure

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.ucla.edu/). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.ucla.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Conflicts of Interest

The authors confirm that this article content has no conflicts of interest.

Acknowledgments

This paper is partially supported by the JiangXi Provincial Natural Science Foundation of China (No. 20192ACB21004), the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (No. 20YJAZH142), and the Scientific and Technological Research Project of Education Department in Jiangxi Province (GJJ190356).

References

[1] H. W. Querfurth and F. M. Laferla, "Alzheimer's disease," *The New England Journal of Medicine*, vol. 362, no. 4, pp. 329–344, 2010.

[2] N. Vilor-Tejedor, M. A. Ikram, G. V. Roshchupkin et al., "Independent multiple factor association analysis for multiblock data in imaging genetics," *Neuroinformatics*, vol. 17, no. 4, pp. 583–592, 2019.

[3] F. S. Nathoo, L. Kong, H. Zhu, and for the Alzheimer's Disease Neuroimaging Initiative, "A review of statistical methods in imaging genetics," *Canadian Journal of Statistics*, vol. 47, no. 1, pp. 108–131, 2019.

[4] A. de Marvao, T. J. W. Dawes, and D. P. O'Regan, "Artificial intelligence for cardiac imaging-genetics research," *Frontiers in Cardiovascular Medicine*, vol. 6, 2020.

[5] C. Biffi, A. de Marvao, M. I. Attard et al., "Three-dimensional cardiovascular imaging-genetics: a mass univariate framework," *Bioinformatics*, vol. 34, no. 1, pp. 97–103, 2018.

[6] H. Janouschek, C. R. Eickhoff, T. W. Mühleisen, S. B. Eickhoff, and T. Nickl-Jockschat, "Using coordinate-based meta-analyses to explore structural imaging genetics," *Brain Structure and Function*, vol. 223, no. 7, pp. 3045–3061, 2018.

[7] J. Zhou, Y. Qiu, S. Chen et al., "A novel three-stage framework for association analysis between SNPs and brain regions," *Frontiers in Genetics*, vol. 11, article 572350, 2020.

[8] D. Zhang, X. Liu, J. Chen, B. Liu, and J. Wang, "Widespread increase of functional connectivity in Parkinson's disease with tremor: a resting-state fMRI study," *Frontiers in Aging Neuroscience*, vol. 7, p. 6, 2015.

[9] J. Yan, L. du, S. Kim et al., "Transcriptome-guided amyloid imaging genetic analysis via a novel structured sparse learning algorithm," *Bioinformatics*, vol. 30, no. 17, pp. i564–i571, 2014.

[10] X. Hao, X. Yao, S. Risacher et al., "Identifying candidate genetic associations with MRI-derived AD-related ROI via tree-guided sparse learning," *IEEE/ACM Transactions on Computational Biology & Bioinformatics*, vol. 16, no. 6, pp. 1986–1996, 2019.

[11] Alzheimer's Disease Neuroimaging Initiative, H. Hu, J. Li, J. Li, J. Yu, and L. Tan, "Genome-wide association study identified ATP6V1H locus influencing cerebrospinal fluid BACE activity," *BMC Medical Genetics*, vol. 19, no. 1, 2018.

[12] T. Zhou, K. H. Thung, M. Liu, and D. Shen, "Brain-wide genome-wide association study for Alzheimer's disease via joint projection learning and sparse regression model," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 1, pp. 165–175, 2019.

[13] D. Welter, J. MacArthur, J. Morales et al., "The NHGRI GWAS Catalog, a curated resource of SNP-trait associations," *Nucleic Acids Research*, vol. 42, no. D1, pp. D1001–D1006, 2014.

[14] H. Marei, A. Althani, J. Suhonen et al., "Common and rare genetic variants associated with Alzheimer's disease," *Journal of Cellular Physiology*, vol. 231, no. 7, pp. 1432–1437, 2016.

[15] A. J. Saykin, L. Shen, X. Yao et al., "Genetic studies of quantitative MCI and AD phenotypes in ADNI: progress, opportunities, and plans," *Alzheimer's & Dementia*, vol. 11, no. 7, pp. 792–814, 2015.

[16] C. M. Karch, C. Cruchaga, and A. M. Goate, "Alzheimer's disease genetics: from the bench to the clinic," *Neuron*, vol. 83, no. 1, pp. 11–26, 2014.

[17] E. Westman, A. Simmons, Y. Zhang et al., "Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls," *Neuro image*, vol. 54, no. 2, pp. 1178–1187, 2011.

[18] I. Beheshti and H. Demirel, "Probability distribution function-based classification of structural MRI for the detection of Alzheimer's disease," *Computers in Biology & Medicine*, vol. 64, pp. 208–216, 2015.

[19] I. Beheshti and H. Demirel, "Feature-ranking-based Alzheimer's disease classification from structural MRI," *Magnetic Resonance Imaging*, vol. 34, no. 3, pp. 252–263, 2016.

[20] W. Zhang, T. Zeng, and L. Chen, "EdgeMarker: Identifying differentially correlated molecule pairs as edge- biomarkers," *Journal of Theoretical Biology*, vol. 362, pp. 35–43, 2014.

Hindawi
BioMed Research International
Volume 2023, Article ID 9872581, 1 page
https://doi.org/10.1155/2023/9872581

Retraction

Retracted: Shared Genetic and Epigenetic Mechanisms between the Osteogenic Differentiation of Dental Pulp Stem Cells and Bone Marrow Stem Cells

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Gaus, H. Li, S. Li et al., "Shared Genetic and Epigenetic Mechanisms between the Osteogenic Differentiation of Dental Pulp Stem Cells and Bone Marrow Stem Cells," *BioMed Research International*, vol. 2021, Article ID 6697810, 25 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6697810, 25 pages
https://doi.org/10.1155/2021/6697810

Research Article

Shared Genetic and Epigenetic Mechanisms between the Osteogenic Differentiation of Dental Pulp Stem Cells and Bone Marrow Stem Cells

Sebastian Gaus,[1] **Hanluo Li**,[1] **Simin Li**,[2] **Qian Wang**,[3] **Tina Kottek**,[1] **Sebastian Hahnel**,[1] **Xiangqiong Liu**,[4] **Yupei Deng**,[4] **Dirk Ziebolz**,[2] **Rainer Haak**,[2] **Gerhard Schmalz**,[2] **Lei Liu**,[5] **Vuk Savkovic**,[1] **and Bernd Lethaus**[1]

[1]*Department of Cranio Maxillofacial Surgery, University Clinic Leipzig, Liebigstr. 12, Leipzig 04103, Germany*
[2]*Department of Cariology, Endodontology and Periodontology, University Leipzig, Liebigstr. 12, Leipzig 04103, Germany*
[3]*Department of Central Laboratory, Taian Central Hospital, Longtan Road No. 29, Taian, 271000 Shandong Province, China*
[4]*Department of Molecular Cell Biology, Beijing Tibetan Hospital, China Tibetology Research Center, 218 Anwaixiaoguanbeili Street, Chaoyang, Beijing 100029, China*
[5]*Department of Neurology, Shandong Provincial Third Hospital, Cheeloo Chollege of Medicine, Shandong University, Jinan, 100191 Shandong Province, China*

Correspondence should be addressed to Vuk Savkovic; vuk.savkovic@medizin.uni-leipzig.de

Received 29 November 2020; Revised 4 January 2021; Accepted 20 January 2021; Published 8 February 2021

Academic Editor: Min Tang

Objective. To identify the shared genetic and epigenetic mechanisms between the osteogenic differentiation of dental pulp stem cells (DPSC) and bone marrow stem cells (BMSC). *Materials and Methods*. The profiling datasets of miRNA expression in the osteogenic differentiation of mesenchymal stem cells from the dental pulp (DPSC) and bone marrow (BMSC) were searched in the Gene Expression Omnibus (GEO) database. The differential expression analysis was performed to identify differentially expressed miRNAs (DEmiRNAs) dysregulated in DPSC and BMSC osteodifferentiation. The target genes of the DEmiRNAs that were dysregulated in DPSC and BMSC osteodifferentiation were identified, followed by the identification of the signaling pathways and biological processes (BPs) of these target genes. Accordingly, the DEmiRNA-transcription factor (TFs) network and the DEmiRNAs-small molecular drug network involved in the DPSC and BMSC osteodifferentiation were constructed. *Results*. 16 dysregulated DEmiRNAs were found to be overlapped in the DPSC and BMSC osteodifferentiation, including 8 DEmiRNAs with a common expression pattern (8 upregulated DEmiRNAs (miR-101-3p, miR-143-3p, miR-145-3p/5p, miR-19a-3p, miR-34c-5p, miR-3607-3p, miR-378e, miR-671-3p, and miR-671-5p) and 1 downregulated DEmiRNA (miR-671-3p/5p)), as well as 8 DEmiRNAs with a different expression pattern (i.e., miR-1273g-3p, miR-146a-5p, miR-146b-5p, miR-337-3p, miR-382-3p, miR-4508, miR-4516, and miR-6087). Several signaling pathways (TNF, mTOR, Hippo, neutrophin, and pathways regulating pluripotency of stem cells), transcription factors (RUNX1, FOXA1, HIF1A, and MYC), and small molecule drugs (curcumin, docosahexaenoic acid (DHA), vitamin D3, arsenic trioxide, 5-fluorouracil (5-FU), and naringin) were identified as common regulators of both the DPSC and BMSC osteodifferentiation. *Conclusion*. Common genetic and epigenetic mechanisms are involved in the osteodifferentiation of DPSCs and BMSCs.

1. Introduction

Repairing bone defects remains a challenge for clinical practitioners to the present day. For the last decades, autologous bone transplantation has been a "gold standard" for treating bone defects in the dental field, such as insufficient bone volume for dental implants and maxillofacial defects [1, 2]. Nonetheless, the conventional treatment involves drawbacks such as donor site morbidity and limitation of bone volume [3], which calls for alternative approaches. Mesenchymal

stem cells appear to be a good match to the unmet needs of the conventional autologous bone transplantation. They are capable of acting in paracrine anti-inflammatory and trophic fashion, as well as of providing a cellular base for tissue replacement by virtue of their self-renewal and multilineage differentiation capacity [4]. The most commonly used MSCs are bone marrow stem cells (BMSCs), dental pulp stem cells (DPSCs), and adipose tissue stem cells (ADSCs). BMSCs have been the most commonly used type of stem cells for osteogenesis and new bone formation due to their tissue origin; however, BMSCs are obtained by means of invasive and painful bone marrow aspiration [5]. This downside has inspired the use of other less invasively obtained stem cell types. Among those, DPSCs obtained from the extracted third molar present a highly accessible alternative [6]. In addition, DPSCs have been demonstrated as highly capable of osteogenic differentiation under osteogenesis-inducing conditions [7], and they are therefore becoming a valuable alternative approach for transplantation-based bone regeneration [8]. Nevertheless, critical preclinical data necessary for understanding the genetic and epigenetic mechanisms involved in the osteogenic differentiation of DPSCs and BMSCs are still modestly represented.

Stem cells derived from different tissues tend to respond to the same stimulus differently and differentiate towards the tissue of their origin [9]. Accordingly, BMSCs display a high predisposition to progress towards osteogenic differentiation [9]. It has been shown that the osteogenic differentiation capacity of the BMSCs is higher than that of ADSCs [10, 11] and comparable between the BMSCs and DPSCs [12]. Having tremendous BMSCs and DPSCs osteogenic potential at disposal, it is necessary to gain a fuller understanding of genetic and epigenetic processes that underlie their differentiation capacity in order to utilize their osteogenic ability.

Many genetic and epigenetic factors are involved in the osteogenic differentiation of stem cells, such as messenger RNAs (mRNAs), microRNAs (miRNAs), and signaling pathways. As a significant component of the epigenetic modification, miRNAs are classified as short noncoding RNAs which can inhibit the expression of mRNAs by binding the 3′-untranslated region (UTR) of target mRNAs [13, 14]. The aberrant expression of many miRNAs (e.g., miR-16, miR-381, miR-20a, and miR-214) has been involved in the osteogenic differentiation of stem cells by inducing the dysregulation of osteogenesis-related signaling pathways (e.g., Wnt, BMP, MAPK, RUNX2, and Notch pathways) [15–22]. Although some research using RNA-sequencing technique have investigated the miRNA expression alterations during the osteodifferentiation process of BMSCs [23] and DPSCs [24], respectively, there is still no report which is aimed on identifying the genetic and epigenetic biomarkers shared between the osteogenic differentiation of both types of stem cells.

The present study is aimed at identifying the overlapping genetic and epigenetic mechanisms involved in the osteogenic differentiation of DPSC and BMSC. To this purpose, a series of bioinformatics analyses (e.g., differential expression analysis, DEmiRNA-target gene network construction, DEmiRNA-transcription factor construction, and functional enrichment analysis) were performed to investigate the major common genetic and epigenetic mechanisms of DPSC osteogenesis and BMSC osteogenesis.

2. Materials and Methods

2.1. Data Acquisition. The miRNA expression profiling datasets regarding the osteogenic differentiation of DPSC and BMSC were searched for in the Gene Expression Omnibus (GEO) database of NCBI [25]. The following inclusion criteria were defined for analyzing two study groups: an undifferentiated control group nominated as the day zero of differentiation (d0), and the experimental group examined after differentiation on day 14, with matched examination days for the DPSC and BMSC experimental group (d14). The datasets with a sample size of more than three for each group were included.

2.2. Analyzing Processes of the Present Research. Two datasets (GSE138180 and GSE107279), which investigated the miRNA expression profile of DPSC and BMSC, were selected for assessment. After selecting the corresponding datasets, a series of bioinformatics analyses were performed according to the flowchart of this study (Figure 1).

Firstly, the differential expression analysis was performed to identify the DEmiRNAs which were aberrantly expressed during the osteodifferentiation process of DPSCs and BMSCs, in order to identify the overlapping DEmiRNAs. Subsequently, the target genes of these common DEmiRNAs that were singled out in the first step were identified by constructing a target gene network of 16 overlapping DEmiRNAs. Furthermore, target genes of DEmiRNAs involved in the DPSC osteogenesis and BMSC osteogenesis were, respectively, extracted by searching miRNA-target interaction databases. The functions (i.e., biological processes and signaling pathways) of these target genes were identified by the means of functional enrichment analysis. In addition, transcription factors that potentially target the DEmiRNAs were identified by constructing the DEmiRNAs-transcription factor interaction network. Conclusively, the DEmiRNAs-small molecular drug targets network was constructed in order to identify the small molecular drugs that can influence the expression of DEmiRNAs involved in DPSC and BMSC osteodifferentiation.

2.3. Differential Expression Analysis for Identifying DEmiRNAs. Differential expression analysis (DEA) was used to identify the differentially expressed miRNAs (DEmiRNAs) of the two selected datasets. DEA was performed by using different packages in the R program depending on the different experimental types of datasets. The limma package was used to analyze continuous data such as microarray data, whereas the edgeR/DESeq/DESeq2 was used for count data such as high-throughput sequencing. The experimental type of the GSE138180 dataset presents the noncoding RNA profile analyzed by the means of microarray, whereas the experimental type of the GSE107279 dataset profiling of the noncoding RNA has been attained by high-throughput

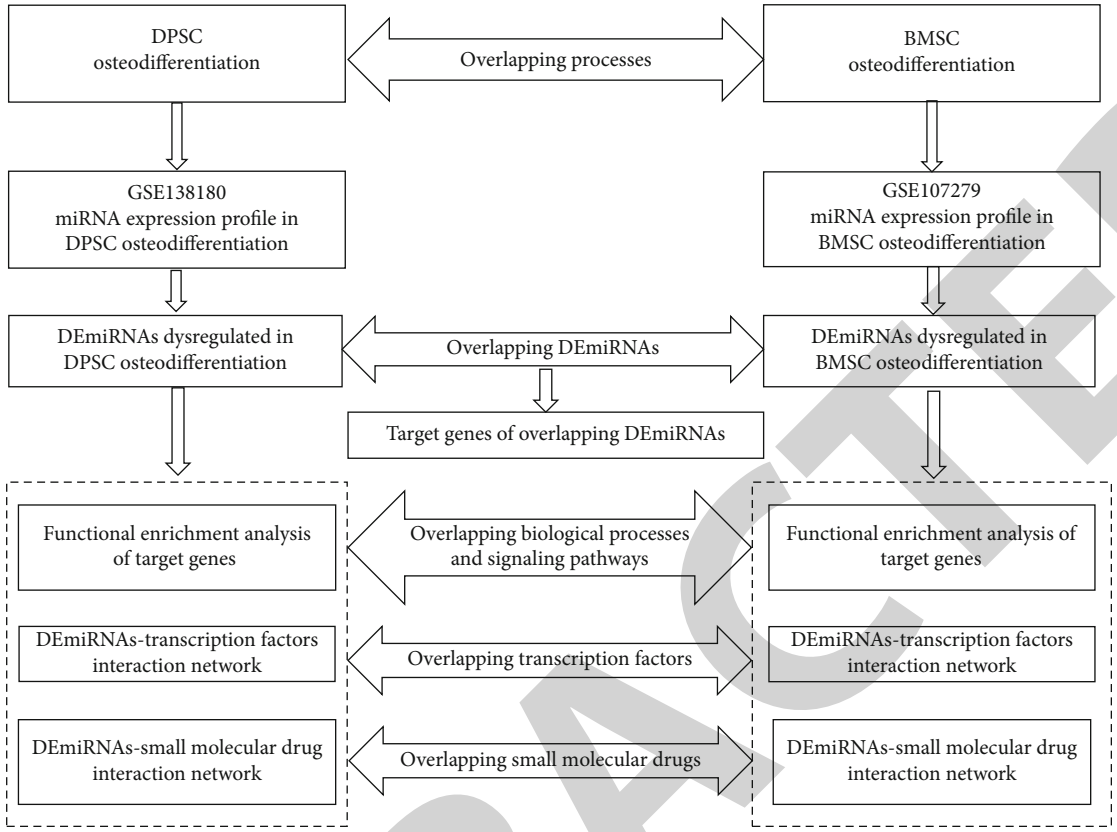


Figure 1: The study flowchart.

sequencing. Based on this, the DEA of the GSE138180 dataset was carried out using the limma package of the R program version 3.6.1 [26], whereas the DEA of the GSE107279 dataset was performed by using the DESeq2 package [27]. The miRNAs with a p value < 0.05 and $|\log FC| \geq 1$ were selected as differentially expressed. The DEmiRNAs with $\log FC \geq 1$ was defined as upregulated DEmiRNAs, while the DEmiRNAs with $\log FC \leq -1$ were defined as downregulated DEmiRNAs. The overlapping DEmiRNAs expressed in the osteogenic differentiation of DPSC and BMSC were found by uploading the list of DEmiRNAs in the Venn web tool webpage (http://bioinformatics.psb.ugent.be/webtools/Venn/).

2.4. Target Genes of the Overlapping DEmiRNAs in DPSC and BMSC Osteodifferentiation. After identifying the overlapped DEmiRNAs that were dysregulated in both DPSC and BMSC osteodifferentiation, their target genes were extracted by downloading the human miRNA-target interaction pairs from three databases including miRDB Version 6.0 [28], miRWalk [29], and TargetScan v7.1 [30]. The overlapping DEmiRNA-target gene interaction network was therefore constructed, and the target genes with the top degree were identified.

2.5. Functional Enrichment Analysis of Target Genes of DEmiRNAs. After obtaining the DEmiRNAs expressed during the osteogenic differentiation of DPSC and BMSC, target genes of DEmiRNAs involved in DPSC osteodifferentiation and BMSC osteodifferentiation were, respectively, extracted by downloading the human miRNA-target interaction pairs from three databases mentioned earlier [28–30]. The functional enrichment analysis (FEA) was performed on the target genes of DEmiRNAs expressed in the osteogenic differentiation of DPSC and BMSC. This analysis was conducted by using the clusterProfiler package of the R program [31]. The functions of the candidate target genes of the DEmiRNAs were explored by assessing their enrichment in Gene Ontology (GO) terms, in particular biological processes (BPs) and the pathway enrichment analysis according to the Kyoto Encyclopedia of Genes and Genomes (KEGG). The GO terms and KEGG pathways with a p value < 0.05 were considered to be significantly enriched. If the number of enriched BPs and pathways was greater than 30, only the top 30 with the highest p values were chosen to be visualized in the bar plot. If the number of enriched BPs and pathways was lower than 30, all of the BPs and pathways were visualized in the bar plot.

2.6. Construction of DEmiRNA-Transcription Factor (TF) Interaction Network. Based on the DEmiRNAs obtained by DEA, the DEmiRNA-transcription factor (TF) interaction pairs were derived from the TransmiR database [32]. The DEmiRNA-TF regulatory networks involved in the DPSC osteogenesis and BMSC osteogenesis were, respectively, plotted by using the Cytoscape software version 3.7.2. The topological characteristics of the nodes in these two networks were calculated. The top 30 TFs with the highest degree in

these two networks were selected, and their intersection was obtained. The intersection between the top 30 TFs of these two networks could be regarded as critical overlapping TFs that targeted DEmiRNAs dysregulated in both the DPSC and BMSC osteodifferentiation.

2.7. Construction of DEmiRNA-Small Molecular Drug Target Interaction Network. The SM2miR database provides information about experimentally validated linkage between miRNAs and bioactive small molecules that can influence expression levels of the miRNAs [33]. This database contains 2,925 relationship pairs between small molecule drugs and miRNAs in 17 species, and only the data related to human species were collected. A total of 2,756 human interaction pairs between miRNAs and small molecular drugs were collected from this database. The interaction pairs of (DEmiRNAs (DPSC osteodifferentiation)-small molecule targets) and (DEmiRNAs (BMSC osteodifferentiation)-small molecule targets) were, respectively, extracted from database SM2miR v1.0 [33]. Subsequently, the DEmiRNAs-small molecule drug targets networks were constructed for DPSC and BMSC osteodifferentiation by the means of Cytoscape software version 3.6. In these two networks, the expression patterns (up-/downregulation) of miRNAs were defined according to their expression in the DPSC and BMSC osteodifferentiation, rather than their expression patterns shown in the SM2miR database. Consequently, the intersections between these two networks were identified based on the overlapping DEmiRNA-small molecule interaction pairs and the overlapping small molecules.

3. Results

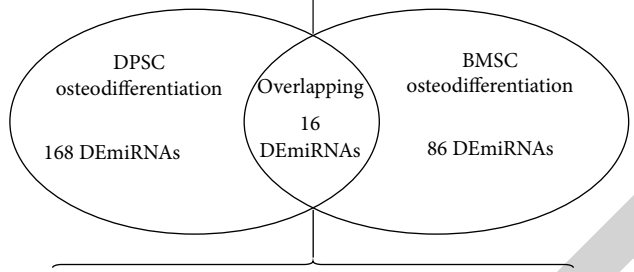
3.1. Dataset Procurement. Preselection of data relevant to the miRNA expression profiling of DPSC and BMSC in the course of the osteogenic differentiation process brought about two datasets (GSE138180 and GSE107279). The GSE138180 dataset (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE138180) reported the miRNA expression alteration of DPSC investigated by comparing the 14^{th} day of postdifferentiation with the 14^{th} day of culturing cells without differentiation. The GSE107279 dataset (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE107279) addressed the miRNA expression alteration of BMSC by comparing the 13^{th} day of postdifferentiation with day 0 (nondifferentiated cells while prior to differentiation).

3.2. The Overlapping DEmiRNAs during the Osteogenic Differentiation between DPSC and BMSC. After performing DEA based on two datasets (GSE138180 and GSE107279), 186 DEmiRNAs (122 upregulated and 64 downregulated) and 104 DEmiRNAs (52 upregulated and 52 downregulated) were identified as differentially expressed in the osteogenic differentiation processes of DPSC and BMSC, respectively (Figure 2. File S1 and S2 showed the dysregulation of all DEmiRNAs dysregulated in the osteodifferentiation of DPSCs and BMSCs. The top 30 DEmiRNAs of GSE138180 and GSE107279 were listed in Table 1, ranked in the ascending order of their p value. The explanation of the abbreviated parameters shown in Table 1 was listed as below: logFC: log2 fold change; AveExpr: average expression across all samples; t: logFC divided by its standard error; p.value: raw p value (based on t) from the test that logFC differs from 0; adj.p.Val: Benjamini-Hochberg false discovery rate adjusted p value; B: log-odds that miRNA is DE; expression patterns: DEmiRNA is upregulated or downregulated.

As shown in the Venn diagram (Figure 2), a total of 186 DEmiRNAs (122 upregulated and 64 downregulated) and 104 DEmiRNAs (52 upregulated and 52 downregulated) were identified to be involved in the osteodifferentiation of DPSCs and BMSCs, respectively. 16 DEmiRNAs were identified as relevant for the osteogenic differentiation of both DPSCs and BMSCs. Among these 16 overlapped DEmiNAs, eight DEmiRNAs (i.e., miR-101-3p, miR-143-3p, miR-145 (3p/5p), miR-19a-3p, miR-34c-5p, miR-3607-3p, miR-378e, and miR-671 (3p/5p)) were found to show the same expression patterns in the osteodifferentiation course (2 weeks) of both DPSCs and BMSCs, while another eight DEmiRNAs (i.e., miR-1273g-3p, miR-146a-5p, miR-4508, miR-4516, miR-6087, miR-146b-5p, miR-337-3p, and miR-382-3p) were found to show the different expression patterns. Among the eight DEmiRNAs with the same expression patterns, seven DEmiRNAs were found to be upregulated in the osteogenic differentiation course (~2 weeks) of both DPSCs and BMSCs: miR-101-3p, miR-143-3p, miR-145(3p/5p), miR-19a-3p, miR-34c-5p, miR-3607-3p, and miR-378e; and another one DEmiRNA (miR-671 (3p/5p)) was found to be downregulated in the osteogenic differentiation course (2 weeks) of both DPSCs and BMSCs. Another eight DEmiRNAs were divergently expressed in DPSCs and BMSCs throughout the osteogenic differentiation course (2 weeks), e.g., five miRNAs were found downregulated in DPSCs while upregulated in BMSCs (i.e., miR-1273g-3p, miR-146a-5p, miR-4508, miR-4516, and miR-6087), as well as three miRNAs upregulated in DPSCs while downregulated in BMSCs (i.e., miR-146b-5p, miR-337-3p, and miR-382-3p). Among these 16 DemiRNAs, 11 of them (i.e., miR-19a-3p, miR-3607-3p, miR-378e, miR-671 (3p/5p), miR-1273g-3p, miR-146b-5p, miR-337-3p, miR-4508, miR-4516, miR-6087, and miR-382-3p) were not supported by previous evidence to be involved in the osteodifferentiation of stem cells. The functions of other five DEmiRNAs (i.e., miR-101-3p, miR-143-3p, miR-145 (3p/5p), miR-34c-5p, and miR-146a-5p) in the osteodifferentiaton of DPSCs and BMSCs were summarized in Table 2.

3.3. The Target Genes of the 16 Overlapped DEmiRNAs. In order to identify the target genes of the 16 overlapping DEmiRNAs, their DEmiRNA-target interaction network was constructed (Figure 3). The topological characteristics of all nodes of this network were shown in File S3, and the topological features of the top 30 gene nodes of this network were shown in Table 3. As shown in Table 3, genes CCND2 (cyclin D2), THBS1 (thrombospondin-1), CCND1 (cyclin D1), IGF1R (insulin-like growth factor 1 receptor), REL (REL proto-oncogene, NF-KB subunit), and ELK4 (ETS transcription factor ELK4)) were identified as the target genes of 16 DEmiRNAs.

Datasets	Upregulated DEmiRNAs	Downregulated DEmiRNAs	Total DEmiRNAs
GSE138180 (DPSC osteodifferentiation)	122	64	186
GSE107279 (BMSC osteodifferentiation)	52	52	104



8 DEmiRNAs
same expression pattern in
DPSC and BMSC osteodifferentiation

7 DEmiRNAs upregulated	hsa-miR-101-3p hsa-miR-143-3p hsa-miR-145 (3p/5p) hsa-miR-19a-3p hsa-miR-34c-5p hsa-miR-3607-3p hsa-miR-378e
1 DEmiRNAs downregulated	hsa-miR-671 (3p/5p)

8 DEmiRNAs
different expression pattern in
DPSC and BMSC osteodifferentiation

DEmiRNAs Stem cells	DPSCs	BMSCs
hsa-miR-1273g-3p	down	up
hsa-miR-146a-5p	down	up
hsa-miR-146b-5p	up	down
hsa-miR-337-3p	up	down
hsa-miR-382-3p	up	down
hsa-miR-4508	down	up
hsa-miR-4516	down	up
hsa-miR-6087	down	up

Figure 2: 16 DEmiRNAs identified to be overlapped in the osteodifferentiation of DPSCs and BMSCs.

3.4. The Overlapping Biological Processes and Pathways Involved in the Osteogenic Differentiation of DPSCs and BMSCs. Subsequently, by the means of the functional enrichment analysis, the significantly enriched biological processes (BPs) and signaling pathways in target genes of the DEmiRNAs were identified for DPSCs and BMSCs (Figures 4 and 5). The overlapping BPs and pathways involved in the osteogenic differentiation of DPSCs and BMSCs were singled out. As shown in Figure 4, several BPs were identified as overlapping and therefore common for osteogenic differentiation of both DPSCs and BMSCs: p53 binding, protein heterodimerization activity, ATPase activity, GTPase binding, cadherin binding, and histone binding DPSCs and BMSCs. As shown in Figure 5, a total of 12 signaling pathways were identified as overlapped between DPSCs and BMSCs: TNF, mTOR, Hippo, neutrophin, pathways regulating pluripotency of stem cells, cell cycle, MAPK, focal adhesion, ubiquitin-mediated proteolysis, viral carcinogenesis, autophagy, and protein processing in endoplasmic reticulum and endocytosis.

3.5. The Overlapping Transcription Factors Targeted by DEmiRNAs between DPSCs and BMSCs. As shown in Figure 6, the DEmiRNA-transcription factor (TF) interaction network involved in both DPSCs and BMSCs were constructed. The top 30 TFs with the highest degree in these two networks were listed in Table 4 ranking in the descending order of degree. By comparing the top 30 TFs with the highest degree in these two networks shown in File S4 and File S5, a total of 26 transcription factors were found overlapped, e.g., TCF12, ERG, CEBPB, ELF1, TCF3, ARNTL, KDM2B, MYC, BRD4, CTCF, MAX, LARP7, HIF1A, E2F1, KDM5B, MAZ, PHF8, STAT1, EP300, AR, FOXA1, TFAP2C, RUNX1, ESR1, TRIM28, and SPI1.

3.6. The Overlapping Small Molecules Targeted by DEmiRNAs Dysregulated in DPSCs and BMSCs Osteodifferentiation. The DEmiRNA-small molecular drug target interaction network, respectively, involved in DPSCs and BMSCs osteodifferentiation were shown in Figures 7(a) and 6(b). In the SM2miR database, only 40 DEmiRNAs of DPSC osteodifferentiation and 27 DEmiRNAs of BMSC osteodifferentiation were found to be targeted and regulated by the small molecular drugs. The small molecular drugs in this database were not found to target and impact the expression of the other DEmiRNAs that were also dysregulated in DPSC and BMSC osteodifferentiation. By assessing the interaction pairs shown in Figures 7(a) and 7(b), we identified 13 interaction pairs consisting of four DEmiRNAs (hsa-miR-143-3p, hsa-miR-146b-5p, hsa-miR-34c-5p, and hsa-miR-671-5p). 13 small molecules were found to be shared between the two networks that are displayed in Table 5. Additionally, a total of 19 small molecule drugs were found to not only impact the expression of DEmiRNAs in DPSC osteodifferentiation but

Table 1: The top 30 DEmiRNAs expressed in the osteogenic differentiation of DPSCs and BMSCs, ranked in the ascending order of their p value.

The top 30 DEmiRNAs dysregulated in DPSC osteodifferentiation				
DEmiRNAs	logFC	p.value	Adj.p.Val	Expression pattern
hsa-let-7f-2-3p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-1468-3p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-153-3p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-212-5p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-2355-5p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-342-5p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-3658	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-450a-1-3p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-450b-3p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-4639-5p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-4730	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-508-3p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-548d-5p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-588	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-6077	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-99a-3p	4.542015233	$1.45E-06$	0.000106517	Up
hsa-miR-1203	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-1238-5p	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-128-2-5p	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-1288-5p	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-129-5p	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-138-1-3p	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-141-3p	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-1972	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-3074-5p	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-339-5p	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-33a-3p	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-34b-3p	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-369-3p	4.1591475	$2.09E-06$	0.000106517	Up
hsa-miR-382-3p	4.1591475	$2.09E-06$	0.000106517	Up
The top 30 DEmiRNAs dysregulated in BMSC osteodifferentiation				
DEmiRNAs	logFC	p.value	Adj.p.Val	Expression pattern
hsa-miR-3182	-5.37979	$5.96E-105$	$3.54E-102$	Up
hsa-miR-182-5p	2.561101	$1.58E-69$	$4.67E-67$	Down
hsa-miR-335-3p	5.160868	$3.61E-59$	$7.13E-57$	Down
hsa-miR-4284	-5.15177	$2.73E-56$	$4.05E-54$	Up
hsa-miR-92b-3p	1.590835	$1.49E-45$	$1.76E-43$	Down
hsa-miR-21-5p	-1.35202	$5.02E-44$	$4.96E-42$	Up
hsa-miR-101-3p	-1.49927	$2.63E-43$	$2.23E-41$	Up
hsa-miR-181a-5p	-1.3232	$2.30E-37$	$1.70E-35$	Up
hsa-miR-143-3p	-1.96642	$1.91E-34$	$1.26E-32$	Up
hsa-miR-210	2.509065	$8.42E-34$	$4.99E-32$	Down

Table 1: Continued.

hsa-miR-146a-5p	-3.4374	$1.93E-33$	$1.04E-31$	Up
hsa-miR-192-5p	-1.78251	$2.03E-31$	$1.00E-29$	Up
hsa-miR-146b-5p	2.019632	$4.61E-30$	$2.10E-28$	Down
hsa-miR-335-5p	4.772695	$3.12E-28$	$1.32E-26$	Down
hsa-miR-382-5p	2.241351	$1.20E-25$	$4.76E-24$	Down
hsa-miR-4485	-6.6002	$4.46E-24$	$1.65E-22$	Up
hsa-miR-27b-5p	1.588671	$2.77E-21$	$9.68E-20$	Down
hsa-miR-10b-5p	-1.21399	$4.59E-20$	$1.51E-18$	Up
hsa-miR-4532	-6.35341	$5.95E-19$	$1.86E-17$	Up
hsa-miR-34c-5p	-1.82262	$8.35E-18$	$2.36E-16$	Up
hsa-miR-181c-5p	-1.15476	$1.18E-17$	$3.19E-16$	Up
hsa-miR-409-3p	1.1162	$1.06E-15$	$2.73E-14$	Down
hsa-miR-218-5p	1.374343	$2.30E-15$	$5.61E-14$	Down
hsa-miR-22-5p	2.004899	$2.36E-15$	$5.61E-14$	Down
hsa-miR-133b	-6.46427	$4.93E-14$	$1.08E-12$	Up
hsa-miR-3195	-5.55444	$5.12E-14$	$1.09E-12$	Up
hsa-miR-133a	-6.42493	$6.09E-14$	$1.24E-12$	Up
hsa-miR-6723-5p	-3.89682	$7.78E-12$	$1.44E-10$	Up
hsa-miR-4497	-3.78672	$1.67E-11$	$3.00E-10$	Up
hsa-miR-4488	-7.47261	$2.43E-11$	$4.11E-10$	Up

Table 2: The dysregulation and function of five overlapping DEmiRNAs (miR-101-3p, miR-143-3p, miR-145 (3p/5p), miR-34c-5p, and miR-146a-5p) in the osteodifferentiation of BMSCs and DPSCs, respectively.

	BMSC osteodifferentiation	DPSC osteodifferentiation
miR-101-3p	The overexpression of miR-101 promoted the osteogenic differentiation of BMSCs by targeting EZH2/Wnt/β-catenin signaling [34].	No evidence.
miR-143-3p	The downregulation of miR-143 promoted the osteogenic differentiation of BMSCs by being competitively combined with lncRNA MALAT1 and upregulating Osterix (Osx) [35].	The downregulation of miR-143 promoted the osteogenic differentiation of DPSCs by activating the NF-κB [36] and OPG/RANKL signaling [37].
miR-145 (3p/5p)	The downregulation of miR-145 promoted the osteogenic differentiation of BMSCs by targeting semaphorin3A (SEMA3A) [38].	The downregulation of miR-145 promoted the odontoblast differentiation of DPSCs by targeting KLF4 and OSX [39].
miR-34c-5p	The upregulation of miR-34c inhibited osteoblast differentiation of BMSCs by targeting the Notch signaling [40].	No evidence.
miR-146a-5p	The overexpression of miR-146a inhibited the osteogenic ability of BMSCs by targeting Smad4 gene [41].	The overexpression of miR-146a-5p promoted osteodifferentiation of DPSCs by suppressing Notch signaling [42].

also that of DEmiRNAs in BMSC osteodifferentiation: 5-fluorouracil, anthocyanin, arsenic trioxide, ascorbate, atorvastatin, bromocriptine, curcumin, docosahexaenoic acid, ginsenoside Rh2, glucose, hesperidin, hydroxycamptothecin (HCPT), microcystin-LR (MC-LR), mistletoe lectin-I, naringin, proanthocyanin, progesterone, vitamin D3, and vorinostat (SAHA). Table 6 shows the targeting relationship between 19 overlapping small molecule drugs and their regulated DEmiRNAs, respectively, expressed in DPSC and BMSC osteodifferentiation process.

4. Discussion

This study identifies multiple genetic and epigenetic biomarkers common for osteogenic differentiation of DPSCs and BMSCs, including miRNAs, their target genes, transcription factors, signaling pathways, and small molecular drugs affecting those mentioned. The overlapping miRNAs with the same expression pattern in DPSC and BMSC osteodifferentiation were obtained by investigating the dysregulated DEmiRNAs in both DPSC and BMSC osteogenesis. The

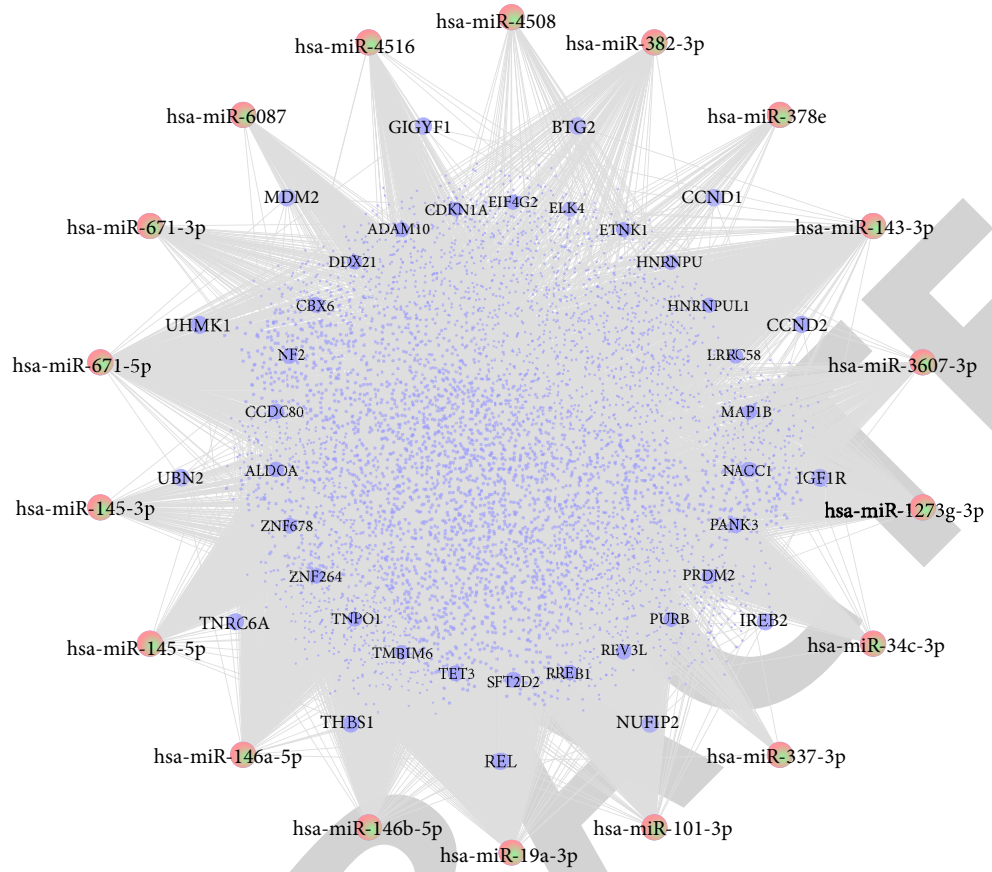


Figure 3: The overlapping 16 DEmiRNA-target gene network. The top 30 target genes with the highest degree were displayed enlarged in the network.

target genes of those shared DEmiRNAs were identified by constructing the shared DEmiRNA-target gene network. The TFs among those genes, signaling pathways involved, and small molecule drugs that predictively affect those pathways were identified as highly plausible common regulators of the osteodifferentiation process in both DPSCs and BMSCs. Their involvement in DPSC and BMSC osteodifferentiation will be discussed here.

This discussion addresses several of the eight miRNAs that were found to be differentially expressed, with the same expression pattern in the course of osteogenic differentiation of DPSCs and BMSCs (miR-101-3p, miR-143-3p, miR-145 (-3p/-5p), miR-19a-3p, miR-34c-5p, miR-3607-3p, miR-378e, and miR-671 (-3p/-5p)). Since there has not been any evidence showing the involvement of miR-19a, miR-3607, miR-378e, and miR-671 in osteogenesis of stem cells, we excluded them from further discussion in order to avoid speculation. The possible role of the other four miRNAs (miR-101, miR-143, miR-145, and miR-34 family (miR-34a/b/c)) will be interpreted in the following part of this section. Overexpression of miR-101 was shown to target EZH2/Wnt/β-Catenin signaling, hereby promoting the osteogenic differentiation in general [34]. Nevertheless, no evidence has been produced to corroborate the expression pattern and either promoting or an inhibiting role of miR-101 during the osteogenesis of DPSCs. The miR-143 has been designated with an evident inhibiting role in osteogenesis and odontogenesis. Downregulation of miR-143 has been shown to promote the osteogenic and odontogenic differentiation of DPSCs by activating the NF-κB [36] and OPG/RANKL signaling [37]. The decreased expression of miR-143 was also shown to promote the osteogenic differentiation of BMSCs by being competitively combined with lncRNA MALAT1 and further upregulating the expression of Osterix (Osx), suggesting that MALAT1-miR-143-Osx could be an integrative element in the ceRNA network of BMSC osteogenesis [35]. Opposite to the results shown in these previous studies, the present study shows an upregulation of miR-143 during the osteogenic differentiation in both BMSCs and DPSCs. miR-145 (-3p/-5p) appears to function the same way as miR-143. The downregulation of miR-145 can apparently promote the odontoblast differentiation of DPSC by targeting transcriptional factor KLF4 and OSX [39]. Also, the downregulation of miR-145 can promote osteogenic differentiation of BMSC by targeting semaphorin3A (SEMA3A), a known positive regulator of osteogenesis [38]. Similarly to the miR-143, the expression pattern of miR-145 shown in the previous work of others is not in agreement with the *in silico* data derived from the datasets analyzed in this study, which show that miR-145 is overexpressed in the course of osteodifferentiation of both DPSCs and BMSCs. The miR-34 family (miR-34a/b/c) has so far not been

Table 3: The topological characteristics of the top 30 target genes with the highest degree in the 16 overlapping DEmiRNA-target gene network. Typically, this table would be displaying only the top 30 target genes with the highest degree; since 26 target genes were ranked with the degree of 6, a total of 39 target genes with a degree ≥ 6 were finally listed in this table.

Name	Degree	Average shortest path length	Betweenness centrality	Closeness centrality	Topological coefficient
BTG2	8	2.431705	0.003363	0.411234	0.18598065
CCND2	8	2.517604	0.003716	0.397203	0.16845654
NUFIP2	8	2.449032	0.003768	0.408325	0.18394988
THBS1	8	2.231521	0.005344	0.448125	0.17855649
UBN2	8	2.587281	0.004042	0.386506	0.17526212
GIGYF1	8	2.589493	0.003609	0.386176	0.17635862
CCND1	7	2.485899	0.002815	0.402269	0.19450361
IGF1R	7	2.426175	0.003724	0.412171	0.19710456
IREB2	7	2.631152	0.00312	0.380062	0.20047142
REL	7	2.838341	0.002574	0.352318	0.19393939
TNRC6A	7	2.514286	0.002822	0.397727	0.20702541
UHMK1	7	2.499171	0.002866	0.400133	0.20764488
MDM2	7	2.519816	0.003737	0.396854	0.19915693
ADAM10	6	2.52682	0.001899	0.395754	0.24168969
CDKN1A	6	2.667281	0.001979	0.374914	0.2337037
EIF4G2	6	2.655853	0.001796	0.376527	0.23616084
ELK4	6	2.542304	0.001911	0.393344	0.24100872
ETNK1	6	2.726636	0.001859	0.366752	0.23112339
HNRNPU	6	2.652903	0.001945	0.376946	0.23472566
HNRNPUL1	6	2.610876	0.001817	0.383013	0.23927525
LRRC58	6	2.652903	0.001945	0.376946	0.23472566
MAP1B	6	2.741382	0.001822	0.364779	0.21069923
NACC1	6	2.780092	0.002673	0.3597	0.22723133
PANK3	6	2.901382	0.002215	0.344663	0.20067454
PRDM2	6	2.437235	0.002245	0.410301	0.22983744
PURB	6	2.785622	0.002027	0.358986	0.2084477
REV3L	6	2.652166	0.001623	0.37705	0.23674815
RREB1	6	2.823594	0.002115	0.354159	0.20654912
SFT2D2	6	2.652166	0.001623	0.37705	0.23674815
TET3	6	2.52682	0.001899	0.395754	0.24168969
TMBIM6	6	2.52682	0.001899	0.395754	0.24168969
TNPO1	6	2.567373	0.002674	0.389503	0.23314389
ZNF264	6	2.627834	0.001775	0.380542	0.23815304
ZNF678	6	2.57106	0.001949	0.388945	0.23931624
ALDOA	6	2.781198	0.002859	0.359557	0.21913299
CCDC80	6	2.662488	0.00286	0.375588	0.20698404
NF2	6	3.205161	0.002464	0.311997	0.2200685
CBX6	6	2.693456	0.001811	0.37127	0.2316048
DDX21	6	2.772719	0.001683	0.360657	0.22691611

designated as categorically pro- or counterosteogenic in DPSCs. Two surveys have shown the involvement of the miR-34 family in the osteogenic differentiation of BMSC [43, 44], however, with quite contradictory results about its expression pattern. Chen et al. showed that miR-34a could inhibit osteogenesis by suppressing regulators cell cycle and cell proliferation cyclin D1, CDK4, and CDK6 [43] and that the inhibition of miR-34a could facilitate osteogenesis of BMSCs. Contrary to that, Xin et al. showed that the upregulation of miR-34a could promote osteogenesis of BMSCs and reverse proinflammatory cytokine influence by targeting tumor necrosis factor-alpha (TNF-α) [45]. Altogether, the osteogenesis-relevant DEmiRNAs that were found overlapping in DPSCs and BMSCs are involved in signaling pathways that regulate differentiation and inflammatory processes by targeting their major regulators.

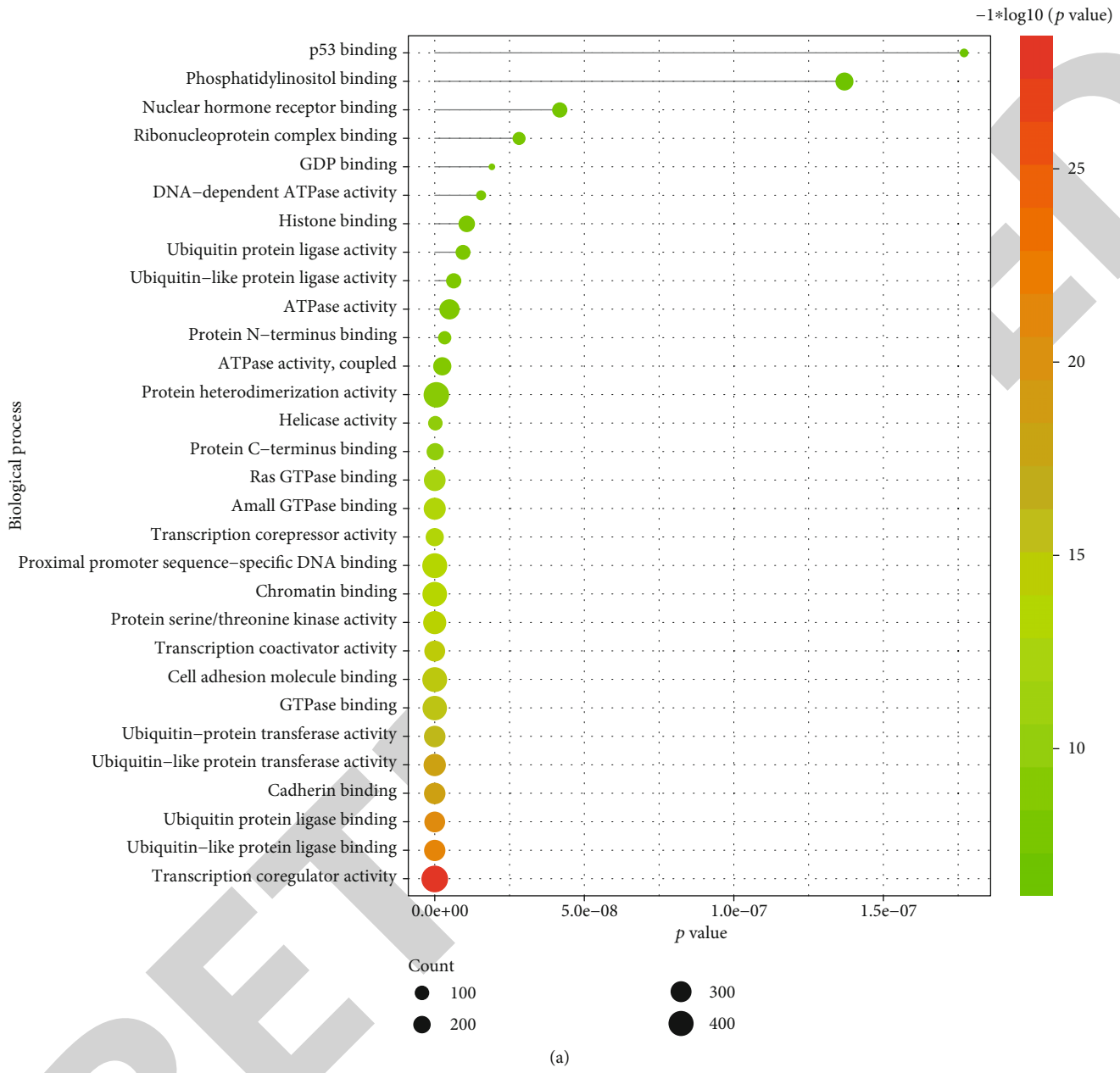


Figure 4: Continued.

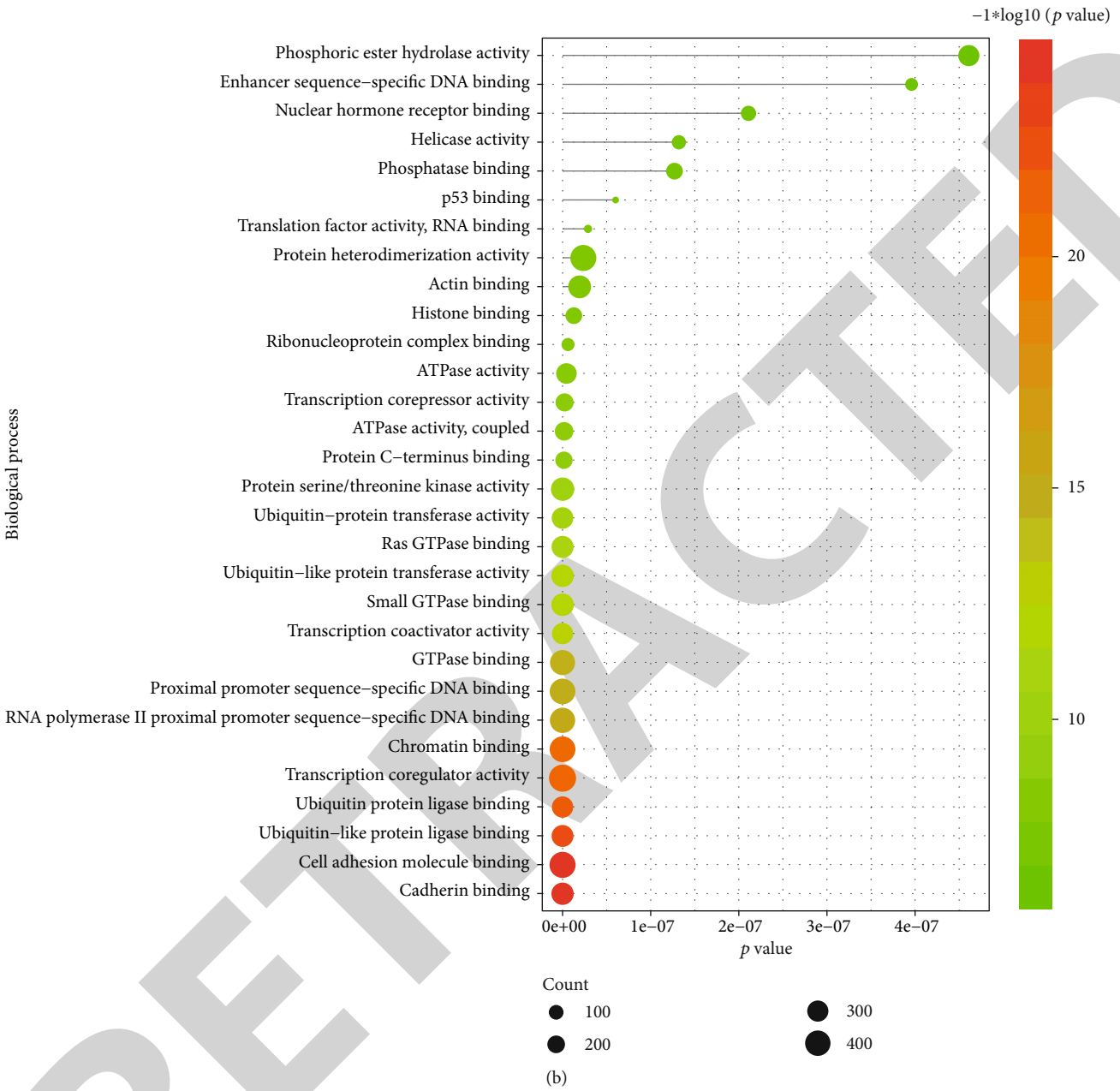
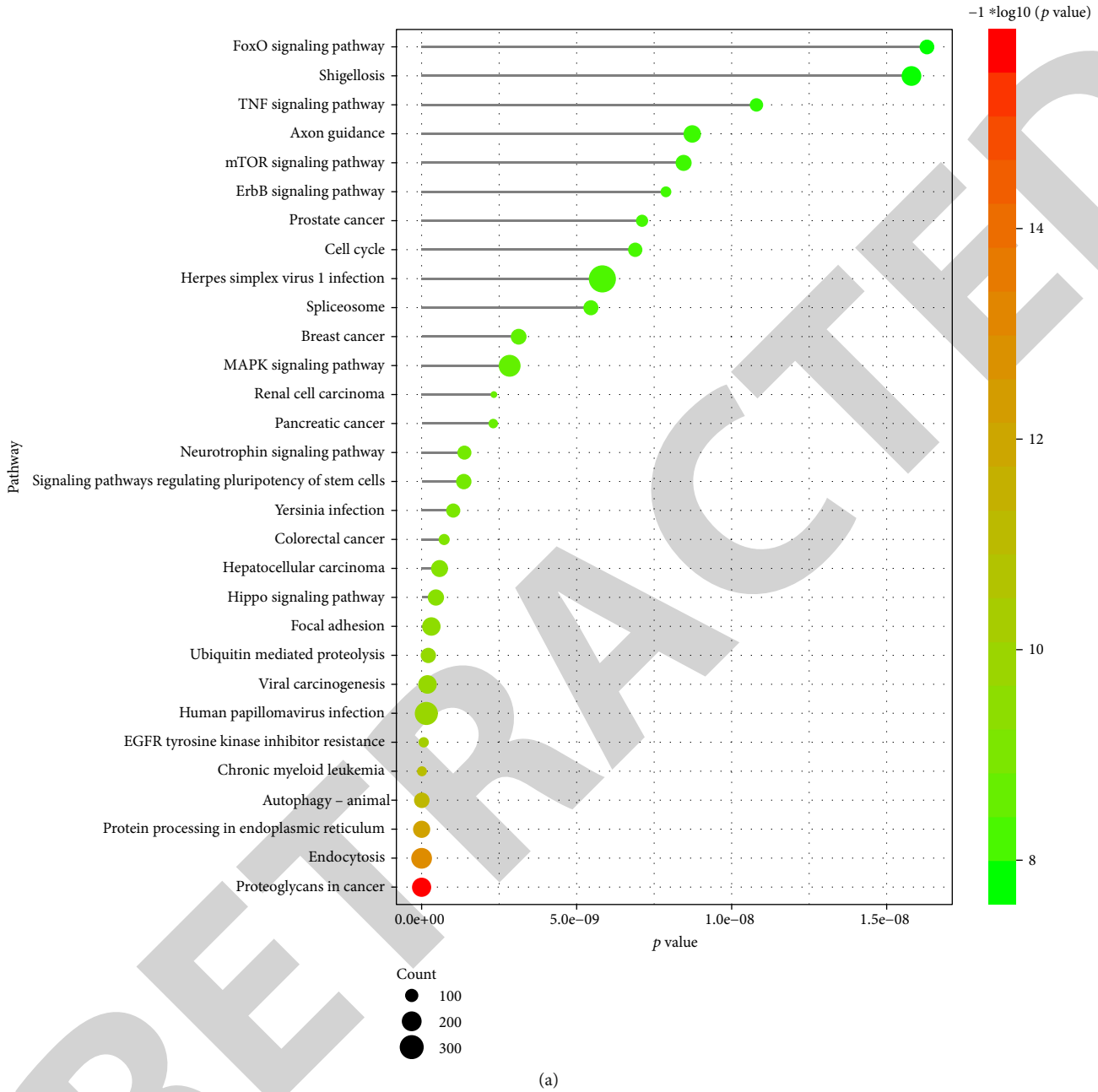


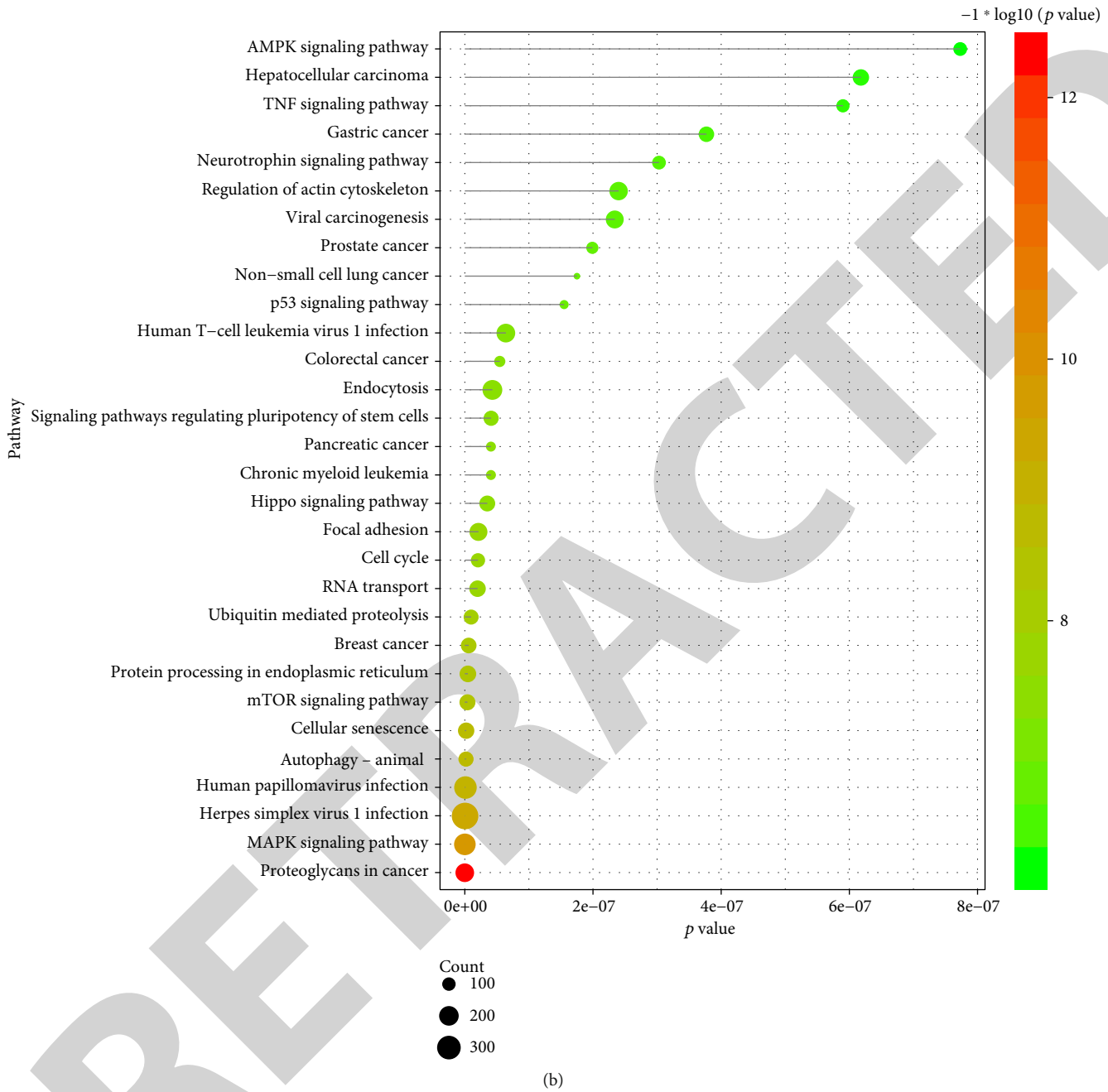
Figure 4: The enriched biological processes involved in the osteogenic differentiation of DPSC (a) and BMSC (b).

After identifying 16 overlapping DEmiRNAs, the target genes of these 16 DEmiRNAs were also identified by constructing the corresponding DEmiRNA-target gene network. Among the top 30 target genes of 16 overlapping DEmiRNAs, only some typical examples of genes will be discussed based on the previous literature evidence regarding their involvement in the osteodifferentiation of DPSCs and BMSCs (e.g., REL, CCND1/2, IGF1R, THBS1, and ELK4). REL gene family consists of RelA (p65), RelB, and c-Rel, subunits of NF-KB. Activation of NF-KB subunit RelA (p65) was shown to significantly promote inflammation and inhibit osteodifferentiation of BMSCs by promoting the ubiquitination and degradation of beta-catenin, indicating that inhibitors of RelA (p65) could be a novel therapeutic target for inhibiting inflammation and enhancing osteogenesis [46]. In accordance with the findings shown in BMSCs, studies investigating DPSCs also showed that the downregulation of NF-KB subunit RelA (p65) enhanced the odontogenic/osteogenic differentiation of DPSCs [47, 48]. Cyclin D genes are involved in cell cycle regulation. Silencing of cyclin D genes (CCND1/2) was shown to arrest the cell cycle, and its overexpression has been connected to the progression of the cell cycle [49]. That the data regarding the effects of cell cycle regulators on osteodifferentiation have so far been insufficiently coherent, CCND1/2 expression was significantly upregulated in the course of DPSC osteodifferentiation [50, 51], whereas during BMSC differentiation the genes involved in the cell



(a)

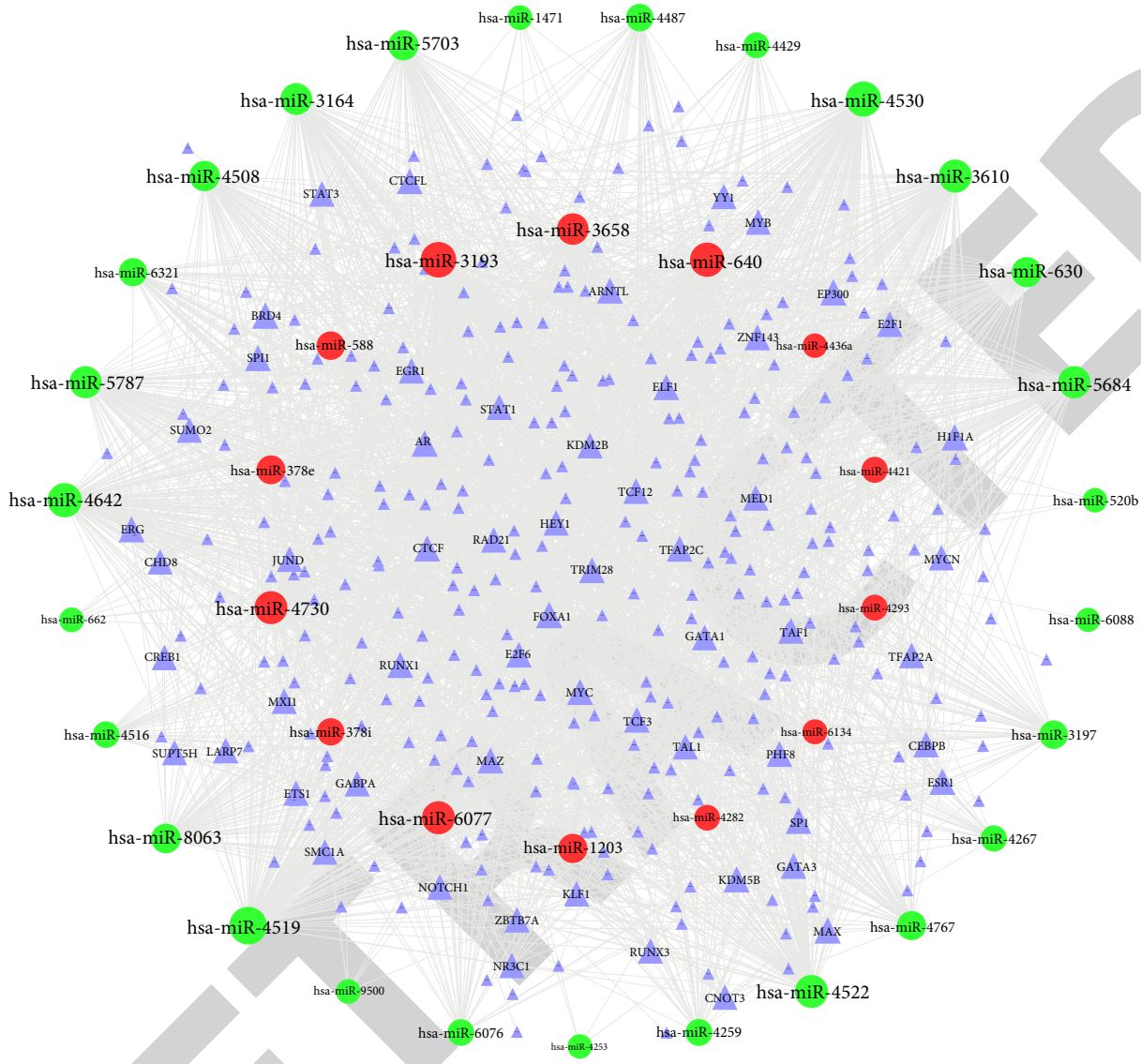
Figure 5: Continued.



(b)

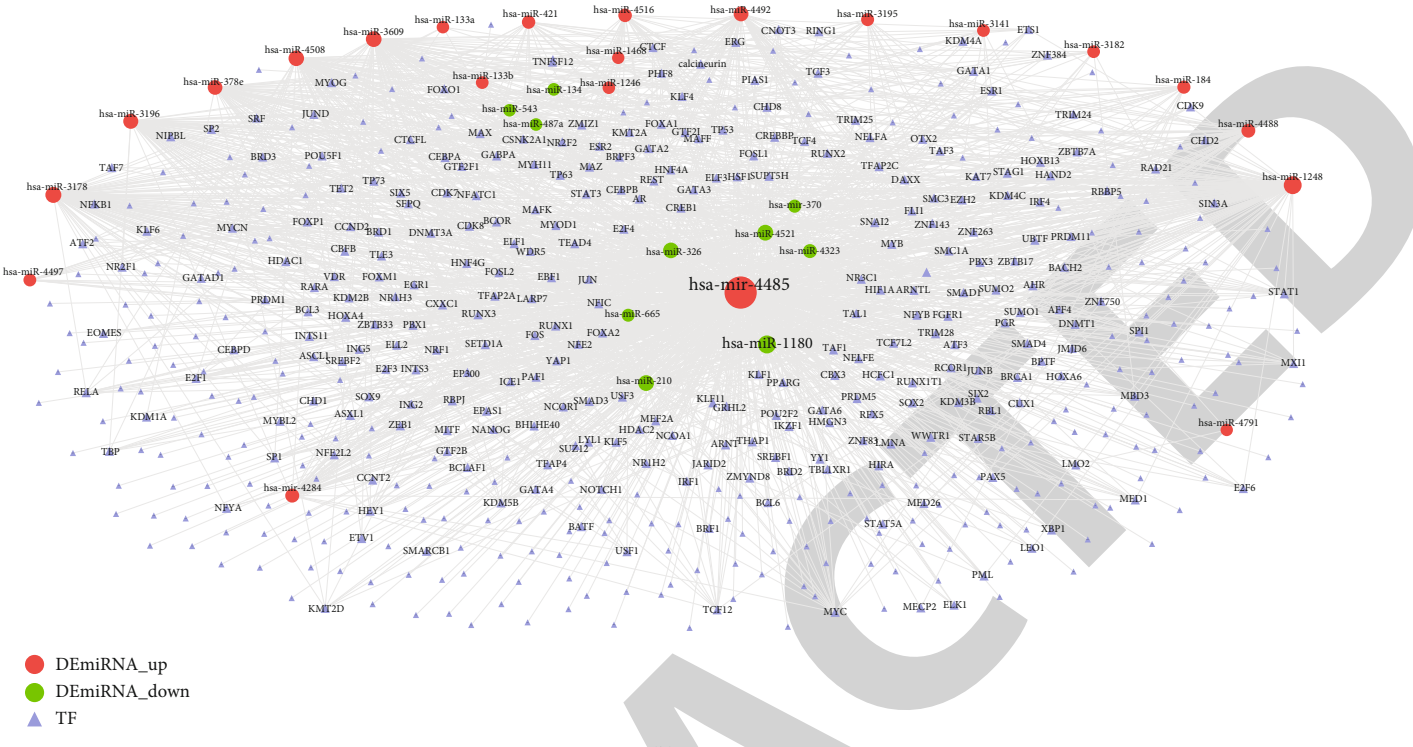
Figure 5: The enriched signaling pathway involved in the osteogenic differentiation of DPSC (a) and BMSC (b).

cycle were significantly downregulated [52]. IGF1R (insulin-like growth factor 1 receptor) can bind IGF1 with a high affinity [53]. The previous studies regarding BMSCs also showed that IGF1 promoted osteodifferentiation of BMSCs and could even be regarded as an alternative biomarker to bone morphogenetic protein-7 [54, 55]. In addition, the IGF axis (e.g., IGF-2 and IGFBP-2) was activated under osteogenic conditions in DPSCs, indicating its promoting role in DPSC osteodifferentiation [56]. THBS1 (thrombospondin-1) is a major regulator of latent TGF-β activation [57], which can inhibit osteoblast differentiation in BMSCs. Accordingly, THBS1 was found to inhibit the osteodifferentiation of BMSCs by activating the TGF-β pathway [58]. An analogue involvement of THBS1 in osteodifferentiation of DPSCs has not been reported yet. Another modulator of osteogenic differentiation ELK4 (ETS transcription factor ELK4) was also predicted to be among the target genes of 16 DEmiRNAs. Ets trasnscription factor consisting of Ets1 and Ets2 targets several osteogenic markers (e.g., osteocalcin, osteopontin, bone sialoprotein, and osteonectin) in BMSCs [59]. However, the evidence regarding the regulation of Ets transcription factor in the osteodifferentiation of DPSC is still lacking and thus could be a novel topic for future research.



(a)

Figure 6: Continued.



(b)

Figure 6: The DEmiRNA-TF interaction network involved in the osteogenic differentiation of DPSC (a) and BMSC (b).

Among the 13 transcription factors found to be overlapped between the top nodes of the DPSC and BMSC osteogenesis DEmiRNA-TF network, four TFs (RUNX1, FOXA1, HIF1A, and MYC) have been reported relevant for osteodifferentiation of stem cells. Upregulation of Runt-related transcription factor 1 (RUNX1) was shown to promote osteogenesis of BMSCs by activating the canonical Wnt/β-catenin pathway [60]. RUNX1 involvement in DPSC osteogenesis has not been shown; nevertheless, another member of the RUNX family (RUNX2) could enhance bone deposition and facilitate novel bone formation after transplantation of RUNX2-transfected DPSCs [61]. Forkhead box protein A1/2 (FOXA1/2) has been reported as hypermethylated in the course of BMSC osteogenesis [62] and its knockdown promoted BMSC osteodifferentiation [63], which has still not been reported in DPSCs. Hypoxic conditions are known to promote cell proliferation and enhance the osteogenic differentiation of both BMSC and DPSC by inducing the expression of HIF1A (hypoxia-inducible factor 1-alpha) and upregulating the osteogenic-related genes (e.g., Oct-4, Sox2, c-Myc, RUNX2, and PPARγ2) [64–66]. HIF1A overexpression in BMSC can induce the overexpression of proosteogenic genes and enhance the ALP activity even in normoxic conditions [67]. However, there is still no direct evidence on the regulating role of HIF1A in the osteodifferentiation process of DPSC. Overexpressed c-MYC was mainly shown to inhibit osteogenic differentiation of both BMSC [68] and DPSC [69]. In DPSCs, stably expressed c-MYC inhibited cell growth during osteogenic differentiation and also impaired the osteogenic phenotypes (i.e., ALP activity and expression of osteogenic marker genes) [69]. On the other hand, a microarray study of gene expression in BMSCs indicated that c-MYC mediated an upregulation of dexamethasone (DEX) expression and of the bone morphogenetic protein 2- (BMP2-) induced osteoblast differentiation [70]. Involvement of other overlapping TFs (e.g., CTCF, BRD4, MAX, ERG, KDM5B, MAZ, EP300, AR, and ESR1) in the osteogenic differentiation of BMSC and DPSC has not been reported yet; thus, these TFs could present novel investigating markers for future research.

A total of 12 overlapping signaling pathways were found to be enriched by the target genes of DEmiRNAs that are involved in the DPSC osteogenesis and BMSC osteogenesis. Among these 12 signaling pathways, five were selected as highly relevant (i.e., TNF, mTOR, Hippo, neutrophin, and signaling pathways regulating pluripotency of stem cells), as the most thoroughly researched and recognized as classic pathways in the osteogenesis of stem cells. Tumor necrosis factor (TNF) signaling cascade has been shown to affect osteogenic differentiation in DPSCs in a dose-dependent manner: proosteogenic at lower doses via activating NF-κB signaling pathway [71] and osteosuppressing at higher doses by modulating the Wnt/β-catenin signaling [72]. In BMSCs, the osteosuppressing effect of TNF-α is exerted through the downregulation of miR-34a and miR-21 expression [73], which are listed among the top 30 dysregulated DEmiRNAs in osteodifferentiation of BMSCs. The mTOR signaling pathway is activated by stimulation with dexamethasone and kaempferol in DPSCs and BMSCs, respectively, suggesting mTOR as a common pathway in the course of their osteogenesis [74, 75]. Hippo signaling pathway promotes osteogenic differentiation in DPSCs upon regulation by SOX2 and the

Table 4: DEmiRNA-transcription factor (TF) interaction network involved in DPSC and BMSC osteodifferentiation. The top 30 TFs with the highest degree of difference in the DEmiRNA-TF network of DPSCS and BMSCS osteodifferentiation were, respectively, ranked in the descending order of degree.

The top 30 TFs with the highest degree in DEmiRNAs-TF network of DPSC osteodifferentiation					
Name	Degree	Average shortest path length	Betweenness centrality	Closeness centrality	Topological coefficient
FOXA1	30	1.954545	0.015768	0.511628	0.223996
BRD4	29	2.008021	0.009367	0.498003	0.244796
AR	29	2.029412	0.012695	0.492754	0.226894
MAX	29	1.997326	0.009458	0.500669	0.245836
EP300	28	1.981283	0.010644	0.504723	0.248155
ESR1	27	2.024064	0.01016	0.494055	0.226708
HIF1A	27	1.986631	0.01079	0.503365	0.245075
ERG	26	2.008021	0.008387	0.498003	0.265101
KDM5B	26	1.986631	0.008951	0.503365	0.259557
CTCF	25	2.034759	0.009876	0.491459	0.25205
MYC	25	2.013369	0.008083	0.49668	0.27092
RUNX1	24	2.034759	0.006933	0.491459	0.277477
SPI1	23	2.072193	0.00707	0.482581	0.250994
ZNF143	22	2.066845	0.00628	0.483829	0.255913
MAZ	22	2.050802	0.006421	0.487614	0.279362
ELF1	22	2.05615	0.005904	0.486346	0.298782
ARNTL	22	2.109626	0.005538	0.474018	0.27185
SP1	22	2.061497	0.007389	0.485084	0.278977
TCF3	21	2.088235	0.005238	0.478873	0.281495
PHF8	21	2.136364	0.004093	0.468085	0.295021
CEBPB	21	2.125668	0.006218	0.47044	0.272461
TCF12	21	2.114973	0.005147	0.472819	0.289542
TFAP2C	21	2.136364	0.006262	0.468085	0.266791
E2F1	20	2.040107	0.006168	0.49017	0.295552
TAF1	20	2.141711	0.004164	0.466916	0.29772
CTCFL	20	2.131016	0.004453	0.46926	0.301133
MED1	19	2.195187	0.003279	0.455542	0.301484
JUND	19	2.18984	0.005712	0.456654	0.264742
NOTCH1	19	2.088235	0.004576	0.478873	0.309831
KDM2B	19	2.147059	0.003495	0.465753	0.310989
LARP7	19	2.104278	0.004425	0.475222	0.316458
STAT1	19	2.125668	0.004486	0.47044	0.303774
E2F6	19	2.088235	0.00484	0.478873	0.274578
GATA3	19	2.163102	0.004673	0.462299	0.286357
TRIM28	19	2.147059	0.004839	0.465753	0.297788
The top 30 TFs with the highest degree in DEmiRNA-TF network of BMSC osteodifferentiation					
Name	Degree	Average shortest path length	Betweenness centrality	Closeness centrality	Topological coefficient
MAX	20	2.021505	0.003652	0.494681	0.161047
AR	20	2.010753	0.010422	0.497326	0.149037
MYC	18	2.032258	0.002722	0.492063	0.175405
ESR1	18	2.02509	0.004109	0.493805	0.154739
MAZ	17	2.050179	0.003885	0.487762	0.163923
FOXA1	17	2.057348	0.004905	0.486063	0.139836
ERG	17	2.039427	0.002605	0.490334	0.182422
HIF1A	16	2.035842	0.00274	0.491197	0.186531
CREB1	16	2.046595	0.003777	0.488616	0.172636

Table 4: Continued.

The top 30 TFs with the highest degree in DEmiRNAs-TF network of DPSC osteodifferentiation					
Name	Degree	Average shortest path length	Betweenness centrality	Closeness centrality	Topological coefficient
EP300	15	2.050179	0.00182	0.487762	0.20078
CTCF	15	2.043011	0.002819	0.489474	0.177864
FLI1	14	2.057348	0.001999	0.486063	0.19308
BRD4	14	2.050179	0.001697	0.487762	0.205114
TP53	13	2.064516	0.002431	0.484375	0.200512
TFAP2C	13	2.071685	0.002276	0.482699	0.183618
TCF12	13	2.060932	0.001517	0.485217	0.214543
RUNX1	13	2.071685	0.001542	0.482699	0.210065
RAD21	13	2.0681	0.002309	0.483536	0.174057
KDM5B	13	2.057348	0.001707	0.486063	0.211276
CEBPB	13	2.071685	0.003446	0.482699	0.156718
TRIM28	12	2.0681	0.001403	0.483536	0.222766
TEAD4	12	2.064516	0.001834	0.484375	0.212077
STAT1	12	2.0681	0.001237	0.483536	0.229126
SPI1	12	2.075269	0.004133	0.481865	0.206451
PHF8	12	2.071685	0.001465	0.482699	0.222712
E2F1	12	2.0681	0.001526	0.483536	0.215917
TCF3	11	2.082437	0.001244	0.480207	0.225841
REST	11	2.057348	0.00601	0.486063	0.207061
NFE2	11	2.071685	0.00123	0.482699	0.231276
MYCN	11	2.071685	0.001787	0.482699	0.231098
LARP7	11	2.082437	0.001129	0.480207	0.235147
KDM2B	11	2.075269	0.00107	0.481865	0.23975
HEY1	11	2.078853	0.001289	0.481034	0.231113
ELF1	11	2.089606	0.001158	0.478559	0.227812
EBF1	11	2.075269	0.001284	0.481865	0.220856
ARNTL	11	2.082437	0.001177	0.480207	0.22781

downstream BMP signaling [76]. In murine BMSCs, the activation of this pathway was induced by calcitonin gene-related peptide (CGRP) and it promoted osteogenic differentiation. Neurotrophin-mediated signaling pathway with rain-derived neurotrophic factor (BDNF) as a key mediator is primarily active in neurogenesis [77] but also intensively involved in osteogenesis of both DPSCs and BMSCs. In DPSCs, the upregulation of BDNF was shown to promote the odontogenic differentiation of DPSCs towards odontoblast-like cells [78] and to indirectly promote osteogenesis in BMSCs [79]. In addition, a plethora of signaling pathways that regulate the pluripotency of stem cells provides the basic framework for any differentiation, including the endpoint bone. Among several key mediators, transforming growth factor-beta (TGF-β) signaling is of particular importance, being implicated in regulatory pathways relevant for osteodifferentiation as well. The activation of the TGF-β pathway promotes osteogenesis in DPSCs by upregulating osteoblast marker genes type I collagen, alkaline phosphatase (ALP), osteocalcin (OCN), and RUNX2 [80], whereas in BMSCs, its function regarding osteogenesis has been reported in controversial contexts [81, 82].

Apart from the genetic and epigenetic biomarkers, this research also identified 19 small molecular drugs with a simultaneous predictive impact on the DEmiRNAs that are dysregulated in both DPSC and BMSC osteodifferentiation. This section will focus on several of those small molecules that exhibit clear anti-inflammatory or proosteogenic effect in BMSCs: curcumin, docosahexaenoic acid (DHA), vitamin D3, arsenic trioxide, 5-fluorouracil (5-FU), and naringin. Curcumin, a natural phenolic biphenyl compound derived from the plant Curcuma longa, is an antioxidant with anti-inflammatory and osteogenesis-regulating effects [83, 84]. In BMSCs, curcumin facilitates differentiation towards osteoblasts by activating Akt/GSK3β and Wnt/β-catenin pathways [85]. Docosahexaenoic acid (DHA) belongs to the omega-3 polyunsaturated fatty acids, and it is well-known for being protective against inflammation. DHA promotes the osteodifferentiation of BMSCs by leading to the upregulation of proosteogenic proteins bone-sialoprotein 2 (BSP2), ALP, and RUNX2 [86]. The 1α,25-dihydroxyvitamin D_3, an active metabolite of vitamin D, is crucial for the maintenance of calcium homeostasis and balance of bone remodeling. It induces differentiation of human BMSCs and dental bud

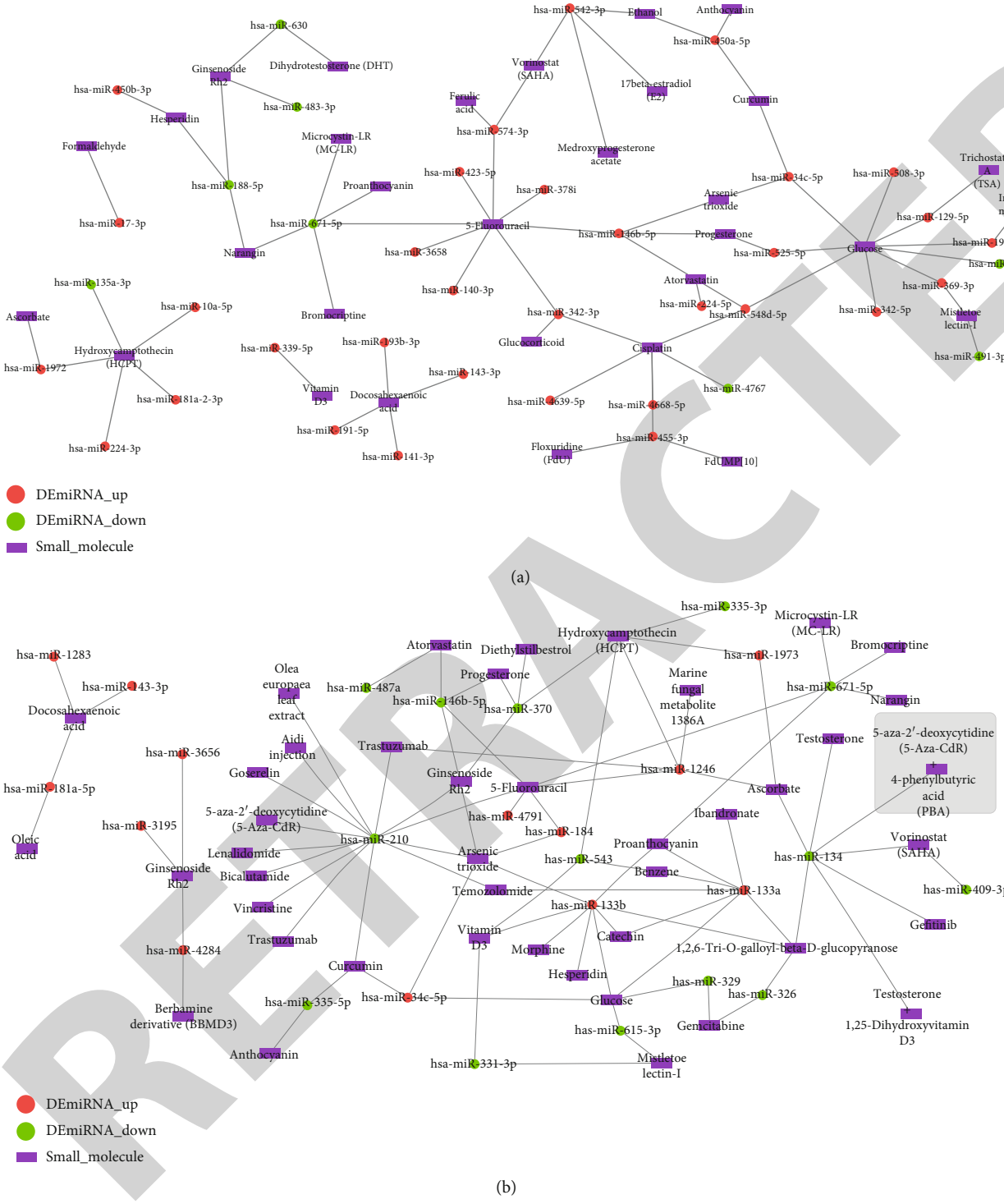


FIGURE 7: The DEmiRNA-small molecular drug targets the interaction network involved in the osteogenic differentiation of DPSC (a) and BMSC (b).

stem cells (DBSCs) towards osteoblasts by increasing the expression levels of osteogenic markers (RUNX2, collagen I (Coll I), ALP, osteopontin, and osteocalcin) [87, 88]. Arsenic trioxide (ATO) is a drug which has been commonly used in treating acute promyeloid leukemia (APL) [89]. Additionally, ATO was shown to promote the osteogenic differentiation of BMSC while inhibiting their adipogenic differentiation [90, 91]. The 5-fluorouracil (5-FU) is a commonly used drug in cancer treatment based on inducing the cell death of cycling cells [92]. BMSCs treated with 5-FU exhibited an enhanced osteodifferentiation [93]. Naringin is a dihydrotestosterone flavonoid compound known to promote osteodifferentiation of BMSCs by downregulating expression levels of peroxisome proliferator-activated

Table 5: The 13 overlapping DEmiRNA-small molecular drug target interaction pairs between the DEmiRNA-small molecular interaction networks of DPSC and BMSC osteodifferentiation.

DEmiRNAs dysregulated in both DPSC and BMSC osteodifferentiation	Small molecular targets
hsa-miR-143-3p	Docosahexaenoic acid
hsa-miR-146b-5p	5-fluorouracil
	Arsenic trioxide
	Atorvastatin
	Progesterone
hsa-miR-34c-5p	Arsenic trioxide
	Curcumin
	Glucose
	5-fluorouracil
hsa-miR-671-5p	Bromocriptine
	Microcystin-LR (MC-LR)
	Narangin
	Proanthocyanin

receptor γ (PPARγ) [94]. It has still not been demonstrated if the proosteogenic or anti-inflammatory effects of the mentioned small molecular drugs found in BMSCs exert the same impact in DPSCs. Nevertheless, since the selection of candidates is a result of multiply derived and highly overlapping calculations and predictions employed in this study, we are leaving room for a confident extrapolation of the drug effects on the DPSCs as well.

Apart from discussing the main outcomes, it is noteworthy to mention the limitations of this study. Although mRNA and lncRNA expression profiles related to the osteodifferentiation of DPSCs and BMSCs were available in public databases, they were not included in the analysis. Undoubtedly useful raw data from these datasets did not apply to the common differentially expressed miRNAs in DPSC and BMSC osteodifferentiation. Also, analyzing mRNA and lncRNA expression profiles would have required an entirely different, amiss the scope of the intended study.

The study design of the two analyzed datasets varied in terms of experimental conditions. The experimental group used to generate that the GSE138180 dataset was treated by odontogenic differentiation medium, whereas the control group was cultured in the nondifferentiating medium for two weeks. Even though the DPSCs in control were not stimulated to differentiate and they were therefore regarded as naive in the original study that generated GSE138180, the culturing conditions may have still left room for spontaneous differentiation [24]. The authors interpreted the absence of mineral nodes in the control group, which is a standard sign of the absence of osteogenic differentiation, as an absence of spontaneous differentiation altogether. By all means, this presents sufficient evidence that no osteodifferentiation took place in the control group [24]. Nevertheless, BMSCs can spontaneously differentiate to osteoblasts [95], which imposes a question if the DPSCs can do it too. The experimental design of the study generated the GSE107279 compared the two weeks of post-differentiation with day 0, hereby ensuring that the control group did not spontaneously differentiate. Hereby, the control groups in the two datasets 'nondifferentiated while not stimulated' in GSE138180 and 'nondifferentiated while prior to differentiation' in GSE107279 were compared as peer groups. In the lack of further osteogenesis-relevant datasets for BMSCs and DPSCs and considering the argument on the nondifferentiated control group provided by the authors of GSE138180, we carried out the comparative analysis of the two mentioned datasets. Optimally though, experimental set up should be established with a control group at d0 (without culturing and prior to any differentiation) and the experimental group on d14 (after culturing in the osteodifferentiation medium).

In addition, it is important to mention that the osteogenic induction media (50 mg/mL ascorbic acid, 100 nmol/L dexamethasone, and 10 mmol/L β-glycerolphosphate) used for GSE138180 could either induce DPSCs towards the osteogenic differentiation and form osteoblasts, or towards the odontogenic differentiation, hereby forming odontoblasts. The authors of GSE138180 regarded the odontogenic and osteogenic differentiation as the same process and reported an odontogenic differentiation of DPSCs in their publication [24] and osteogenic differentiation in the resulting GEO dataset summary (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE138180). In the future, potential confusion in investigating the osteogenic/odontogenic differentiation could easily be resolved by assessing odontogenic biomarkers such as dentin matrix acid phosphoprotein 1 (DMP1), dentin sialophosphoprotein (DSPP), and matrix extracellular phosphoglycoprotein (MEPE) [96]. To this day, much of the interpretation of previous work has been limited by not assessing those differential markers [97–102]. Except for this difference, the genetic and epigenetic mechanisms involved in the osteodifferentiation and odontodifferentiation of DPSCs are common [96]. For the purpose of this study, we adopted the reported statement that DPSCs underwent osteogenic differentiation that was comparable to the osteogenic differentiation of BMSCs.

Availibility of only one time point—2 weeks of osteodifferentiation—for analysis of the DPSC-endpoint-bone in GSE138180 presents another limitation in assessing the gradual process of differentiation. Assessing genetic and epigenetic status over several time points could be crucial for understanding the DPSC differentiation dynamics [103]. Analyzing further data points of the DPSC osteodifferentiation could be extremely helpful for their comparison with the already available multiple time points of the BMSC osteodifferentiation.

Last but not least, this study presents computational predictions based on the reported existing raw data. The predictive data have not been directly experimentally validated. Further validation of the biomarkers predicted in this research could be performed by polymerase chain reaction (PCR) and western blotting assays. Despite the stated limitations, the biomarkers identified in the study provide potential directions for future surveys of genetic and epigenetic

Table 6: The targeting relationship between 19 overlapping small molecule drugs and their regulated DEmiRNAs, respectively, expressed in DPSC and BMSC osteodifferentiation process. These 19 overlapping small molecular drugs can not only regulate the DEmiRNAs dysregulated in DPSC osteodifferentiation but also regulate the DEmiRNAs dysregulated in BMSC osteodifferentiation.

Small molecule	Regulated DEmiRNAs involved in DPSC osteodifferentiation	Regulated DEmiRNAs involved in BMSC osteodifferentiation
5-fluorouracil	hsa-miR-1246	hsa-miR-140-3p
	hsa-miR-146b-5p	hsa-miR-146b-5p
	hsa-miR-184	hsa-miR-342-3p
	hsa-miR-210	hsa-miR-3658
	hsa-miR-4791	hsa-miR-378i
	hsa-miR-671-5p	hsa-miR-423-5p
		hsa-miR-574-3p
		hsa-miR-671-5p
Anthocyanin	hsa-miR-335-5p	hsa-miR-450a-5p
	hsa-miR-133b	hsa-miR-146b-5p
	hsa-miR-146b-5p	hsa-miR-34c-5p
Arsenic trioxide	hsa-miR-184	
	hsa-miR-210	
	hsa-miR-34c-5p	
Ascorbate	hsa-miR-1246	hsa-miR-1972
	hsa-miR-134	
	hsa-miR-1973	
Atorvastatin	hsa-miR-146b-5p	hsa-miR-146b-5p
	hsa-miR-487a	hsa-miR-224-5p
		hsa-miR-548d-5p
Bromocriptine	hsa-miR-671-5p	hsa-miR-671-5p
	hsa-miR-210	hsa-miR-34c-5p
Curcumin	hsa-miR-335-5p	hsa-miR-450a-5p
	hsa-miR-34c-5p	
Docosahexaenoic acid	hsa-miR-1283	hsa-miR-141-3p
	hsa-miR-143-3p	hsa-miR-143-3p
	hsa-miR-181a-5p	hsa-miR-191-5p
		hsa-miR-193b-3p
Ginsenoside Rh2	hsa-miR-210	hsa-miR-188-5p
	hsa-miR-3195	hsa-miR-483-3p
	hsa-miR-3656	hsa-miR-630
	hsa-miR-370	
	hsa-miR-4284	
Glucose	hsa-miR-133a	hsa-miR-129-5p
	hsa-miR-133b	hsa-miR-199b-5p
	hsa-miR-329	hsa-miR-342-5p
	hsa-miR-34c-5p	hsa-miR-34c-5p
	hsa-miR-615-3p	hsa-miR-369-3p
		hsa-miR-508-3p
		hsa-miR-520b
		hsa-miR-525-5p
		hsa-miR-548d-5p
Hesperidin	hsa-miR-133b	hsa-miR-188-5p
		hsa-miR-450b-3p
Hydroxycamptothecin (HCPT)	hsa-miR-1246	hsa-miR-10a-5p
	hsa-miR-1973	hsa-miR-135a-3p

Table 6: Continued.

Small molecule	Regulated DEmiRNAs involved in DPSC osteodifferentiation	Regulated DEmiRNAs involved in BMSC osteodifferentiation
	hsa-miR-335-3p	hsa-miR-181a-2-3p
	hsa-miR-370	hsa-miR-1972
	hsa-miR-543	hsa-miR-224-3p
Microcystin-LR (MC-LR)	hsa-miR-671-5p	hsa-miR-671-5p
Mistletoe lectin-I	hsa-miR-331-3p	hsa-miR-369-3p
	hsa-miR-615-3p	hsa-miR-491-3p
Narangin	hsa-miR-671-5p	hsa-miR-188-5p
		hsa-miR-671-5p
	hsa-miR-133a	hsa-miR-671-5p
Proanthocyanin	hsa-miR-133b	
	hsa-miR-671-5p	
Progesterone	hsa-miR-146b-5p	hsa-miR-146b-5p
	hsa-miR-370	hsa-miR-525-5p
	hsa-miR-133b	hsa-miR-339-5p
Vitamin D3	hsa-miR-331-3p	
	hsa-miR-543	
Vorinostat (SAHA)	hsa-miR-134	hsa-miR-542-3p
	hsa-miR-409-3p	hsa-miR-574-3p

biomarkers as well as regulatory mediators involved in osteogenic differentiation of DPSCs.

Eventually, the DPSCs and BMSCs appear to share common DEmiRNAs and their targets, with cohesive outcomes on the genetic and epigenetic levels in terms of the osteodifferentiation process. This places them among common targets that could be directly or indirectly influenced by the mentioned selected small molecule drug candidates. The established correlations provide a base for a better understanding of the common molecular alterations during the osteogenic differentiation of DPSCs and BMSCs. Those could present an immense potential if used for tissue engineering purposes, especially drug delivery, herewith putatively contributing to the future therapeutical trends in bone regeneration.

5. Conclusions

Several miRNAs (miR-101-3p, miR-143-3p, miR-145-3p, miR-145-5p, miR-19a-3p, miR-34c-5p, miR-3607-3p, miR-378e, miR-671-3p, and miR-671-5p, miR-671-3p, and miR-671-5p), genes (REL, CCND1/2, IGF1R, THBS1, and ELK4), signaling pathways (TNF, mTOR, Hippo, neutrophin, and signaling pathways regulating pluripotency of stem cells), transcription factors (RUNX1, FOXA1, HIF1A, and MYC), and small molecular drugs (curcumin, docosahexaenoic acid (DHA), vitamin D3, arsenic trioxide, 5-fluorouracil (5-FU), and naringin) emerged as shared regulating factors during the osteogenic differentiation of DPSCs and BMSCs. The biomarkers and small molecular drugs identified in this study could be used for the genetic/epigenetic manipulation and drug delivery of DPSCs and therefore result in novel strategies for bone tissue engineering.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Authors' Contributions

Sebastian Gaus and Hanluo Li contributed equally as the first author. Vuk Savkovic and Bernd Lethaus contributed equally as the senior author. These three authors contributed equally as the corresponding author: Dr. Lei Liu, Dr. Vuk Savkovic, and Prof. Dr. Bernd Lethaus.

Acknowledgments

Simin Li (CSC No.: 201608080010) and Hanluo Li (CSC No.: 201308080064) were supported by the China Scholarship Council (CSC) for doctor study at University Leipzig.

Supplementary Materials

Supplementary 1. File S1 The DEmiRNAs dysregulated in the osteodifferentiation of DPSC.

Supplementary 2. File S2 The DEmiRNAs dysregulated in the osteodifferentiation of BMSC.

Supplementary 3. File S3 The topological characteristics of all nodes in the 16 shared DEmiRNAs-target gene network.

Supplementary 4. File S4 The topological characteristics of all nodes in the DEmiRNAs-transcription factor interaction network of DPSCs osteodifferentiation.

Supplementary 5. File S5 The topological characteristics of all nodes in the DEmiRNAs-transcription factor interaction network of BMSCs osteodifferentiation.

References

[1] S. Corbella, S. Taschieri, R. Weinstein, and M. Del Fabbro, "Histomorphometric outcomes after lateral sinus floor elevation procedure: a systematic review of the literature and meta-analysis," *Clinical Oral Implants Research*, vol. 27, no. 9, pp. 1106–1122, 2016.

[2] T. Fretwurst, L. M. Gad, K. Nelson, and R. Schmelzeisen, "Dentoalveolar reconstruction: modern approaches," *Current Opinion in Otolaryngology & Head and Neck Surgery*, vol. 23, no. 4, pp. 316–322, 2015.

[3] E. M. Younger and M. W. Chapman, "Morbidity at bone graft donor sites," *Journal of Orthopaedic Trauma*, vol. 3, no. 3, pp. 192–195, 1989.

[4] S. Shanbhag and V. Shanbhag, "Clinical applications of cell-based approaches in alveolar bone augmentation: a systematic review," *Clinical Implant Dentistry and Related Research*, vol. 17, Supplement 1, pp. e17–e34, 2015.

[5] M. F. Zahid, "Methods of reducing pain during bone marrow biopsy: a narrative review," *Annals of Palliative Medicine*, vol. 4, no. 4, pp. 184–193, 2015.

[6] P. M. Sunil and R. Manikandan, "Harvesting dental stem cells - overview," *Journal of Pharmacy & Bioallied Sciences*, vol. 7, Supplement 2, pp. S384–S386, 2015.

[7] M. Tatullo, M. Marrelli, K. M. Shakesheff, and L. J. White, "Dental pulp stem cells: function, isolation and applications in regenerative medicine," *Journal of Tissue Engineering and Regenerative Medicine*, vol. 9, no. 11, pp. 1205–1216, 2015.

[8] J. Jensen, C. Tvedesøe, J. H. D. Rölfing et al., "Dental pulp-derived stromal cells exhibit a higher osteogenic potency than bone marrow-derived stromal cells in vitro and in a porcine critical-size bone defect model," *SICOT-J*, vol. 2, p. 16, 2016.

[9] T. Pizzute, K. Lynch, and M. Pei, "Impact of tissue-specific stem cells on lineage-specific differentiation: a focus on the musculoskeletal system," *Stem Cell Reviews and Reports*, vol. 11, no. 1, pp. 119–132, 2015.

[10] G.-I. I. Im, Y.-W. Shin, and K.-B. Lee, "Do adipose tissue-derived mesenchymal stem cells have the same osteogenic and chondrogenic potential as bone marrow-derived cells?," *Osteoarthritis and Cartilage*, vol. 13, no. 10, pp. 845–853, 2005.

[11] L. Xu, Y. Liu, Y. Sun et al., "Tissue source determines the differentiation potentials of mesenchymal stem cells: a comparative study of human mesenchymal stem cells from bone marrow and adipose tissue," *Stem Cell Research & Therapy*, vol. 8, no. 1, p. 275, 2017.

[12] Y. C. Lee, Y. H. Chan, S. C. Hsieh, W. Z. Lew, and S. W. Feng, "Comparing the osteogenic potentials and bone regeneration capacities of bone marrow and dental pulp mesenchymal stem cells in a rabbit calvarial bone defect model," *International Journal of Molecular Sciences*, vol. 20, no. 20, p. 5015, 2019.

[13] Y. Xu, Z. He, M. Song, Y. Zhou, and Y. Shen, "A microRNA-switch controls dietary restriction-induced longevity through Wnt signaling," *EMBO Reports*, vol. 20, no. 5, article e46888, 2019.

[14] K. Felekkis, E. Touvana, Stefanou Ch, and C. Deltas, "microRNAs: a newly described class of encoded molecules that play a role in health and disease," *Hippokratia*, vol. 14, no. 4, pp. 236–240, 2010.

[15] J. Wang, S. Liu, J. Li, S. Zhao, and Z. Yi, "Roles for miRNAs in osteogenic differentiation of bone marrow mesenchymal stem cells," *Stem Cell Research & Therapy*, vol. 10, no. 1, p. 197, 2019.

[16] S. Peng, D. Gao, C. Gao, P. Wei, M. Niu, and C. Shuai, "MicroRNAs regulate signaling pathways in osteogenic differentiation of mesenchymal stem cells (Review)," *Molecular Medicine Reports*, vol. 14, no. 1, pp. 623–629, 2016.

[17] J. De Boer, R. Siddappa, C. Gaspar, A. Van Apeldoorn, R. Fodde, and C. Van Blitterswijk, "Wnt signaling inhibits osteogenic differentiation of human mesenchymal stem cells," *Bone*, vol. 34, no. 5, pp. 818–826, 2004.

[18] K. S. Houschyar, C. Tapking, M. R. Borrelli et al., "Wnt pathway in bone repair and regeneration–what do we know so far," *Frontiers in Cell and Development Biology*, vol. 6, p. 170, 2019.

[19] L. Duan, H. Zhao, Y. Xiong et al., "miR-16-2∗ interferes with WNT5A to regulate osteogenesis of mesenchymal stem cells," *Cellular Physiology and Biochemistry*, vol. 51, no. 3, pp. 1087–1102, 2018.

[20] J.-f. Zhang, W.-m. Fu, M.-l. He et al., "MiRNA-20a promotes osteogenic differentiation of human mesenchymal stem cells by co-regulating BMP signaling," *RNA Biology*, vol. 8, no. 5, pp. 829–838, 2014.

[21] H. Long, Y. Zhu, Z. Lin et al., "miR-381 modulates human bone mesenchymal stromal cells (BMSCs) osteogenesis via suppressing Wnt signaling pathway during atrophic nonunion development," *Cell Death & Disease*, vol. 10, no. 7, 2019.

[22] Y. Guo, L. Li, J. Gao, X. Chen, and Q. Sang, "miR-214 suppresses the osteogenic differentiation of bone marrow-derived mesenchymal stem cells and these effects are mediated through the inhibition of the JNK and p 38 pathways," *International Journal of Molecular Medicine*, vol. 39, no. 1, pp. 71–80, 2017.

[23] C.-C. Chang, M. T. Venø, L. Chen et al., "Global MicroRNA Profiling in Human Bone Marrow Skeletal–Stromal or Mesenchymal-Stem Cells Identified Candidates for Bone Regeneration," *Molecular Therapy*, vol. 26, no. 2, pp. 593–605, 2018.

[24] Z. Chen, K. Zhang, W. Qiu et al., "Genome-wide identification of long noncoding RNAs and their competing endogenous RNA networks involved in the odontogenic differentiation of human dental pulp stem cells," *Stem Cell Research & Therapy*, vol. 11, no. 1, p. 114, 2020.

[25] E. Clough and T. Barrett, "The Gene Expression Omnibus Database," *Methods in Molecular Biology*, vol. 1418, pp. 93–110, 2016.

[26] M. E. Ritchie, B. Phipson, Y. H. Di Wu et al., "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic acids research*, vol. 43, no. 7, p. e47, 2015.

[27] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2," *Genome biology*, vol. 15, no. 12, p. 550, 2014.

[28] N. Wong and X. Wang, "miRDB: an online resource for microRNA target prediction and functional annotations," *Nucleic Acids Research*, vol. 43, no. D1, pp. D146–D152, 2015.

[29] C. Sticht, C. De La Torre, A. Parveen, and N. Gretz, "miRWalk: an online resource for prediction of microRNA binding sites," *PloS one*, vol. 13, no. 10, article e0206239, 2018.

[30] V. Agarwal, G. W. Bell, J.-W. Nam, and D. P. Bartel, "Predicting effective microRNA target sites in mammalian mRNAs," *eLife*, vol. 4, 2015.

[31] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *OMICS*, vol. 16, no. 5, pp. 284–287, 2012.

[32] Z. Tong, Q. Cui, J. Wang, and Y. Zhou, "TransmiR v2.0: an updated transcription factor-microRNA regulation database," *Nucleic Acids Research*, vol. 47, no. D1, pp. D253–D258, 2019.

[33] X. Liu, S. Wang, F. Meng et al., "SM2miR: a database of the experimentally validated small molecules' effects on microRNA expression," *Bioinformatics*, vol. 29, no. 3, pp. 409–411, 2013.

[34] H. Wang, Y. Meng, Q. Cui et al., "MiR-101 Targets the EZH2/Wnt/β-Catenin the Pathway to Promote the Osteogenic Differentiation of Human Bone Marrow-Derived Mesenchymal Stem Cells," *Scientific Reports*, vol. 6, no. 1, 2016.

[35] Y. Gao, F. Xiao, C. Wang et al., "Long noncoding RNA MALAT1 promotes osterix expression to regulate osteogenic differentiation by targeting miRNA-143 in human bone marrow-derived mesenchymal stem cells," *Journal of Cellular Biochemistry*, vol. 119, no. 8, pp. 6986–6996, 2018.

[36] P. Zhang, W. Yang, G. Wang, and Y. Li, "miR-143 suppresses the osteogenic differentiation of dental pulp stem cells by inactivation of NF-κB signaling pathway via targeting TNF-α," *Archives of Oral Biology*, vol. 87, pp. 172–179, 2018.

[37] F. L. Zhan, X. Y. Liu, and X. B. Wang, "The role of MicroRNA-143-5p in the differentiation of dental pulp stem cells into odontoblasts by TargetingRunx2via theOPG/RANKLSignaling pathway," *Journal of Cellular Biochemistry*, vol. 119, no. 1, pp. 536–546, 2018.

[38] Y. Jin, F. Hong, Q. Bao et al., "MicroRNA-145 suppresses osteogenic differentiation of human jaw bone marrow mesenchymal stem cells partially via targeting semaphorin 3A," *Connective Tissue Research*, vol. 61, no. 6, pp. 577–585, 2020.

[39] H. Liu, H. Lin, L. Zhang et al., "miR-145 and miR-143 Regulate Odontoblast Differentiation through Targeting _Klf4_ and _Osx_ Genes in a Feedback Loop∗," *The Journal of Biological Chemistry*, vol. 288, no. 13, pp. 9261–9271, 2013.

[40] Y. Bae, T. Yang, H.-C. Zeng et al., "miRNA-34c regulates Notch signaling during bone development," *Human Molecular Genetics*, vol. 21, no. 13, pp. 2991–3000, 2012.

[41] W. Kuang, L. Zheng, X. Xu et al., "Dysregulation of the miR-146a-Smad4 axis impairs osteogenesis of bone mesenchymal stem cells under inflammation," *Bone research*, vol. 5, no. 1, 2017.

[42] Z. Qiu, S. Lin, X. Hu et al., "Involvement of miR-146a-5p/neurogenic locus notch homolog protein 1 in the proliferation and differentiation of STRO-1+human dental pulp stem cells," *European Journal of Oral Sciences*, vol. 127, no. 4, pp. 294–303, 2019.

[43] L. Chen, K. HolmstrØm, W. Qiu et al., "MicroRNA-34a inhibits osteoblast differentiation and in vivo bone formation of human stromal stem cells," *Stem Cells*, vol. 32, no. 4, pp. 902–912, 2014.

[44] B. C. Xin, Q. S. Wu, S. Jin, A. H. Luo, D. G. Sun, and F. Wang, "Berberine promotes osteogenic differentiation of human dental pulp stem cells through activating EGFR-MAPK-Runx 2 pathways," *Pathology Oncology Research*, vol. 26, no. 3, pp. 1677–1685, 2020.

[45] W. Xin, X. Wang, W. Zhang, H. Zhu, R. Dong, and J. Zhang, "Tumor necrosis Factor-α inhibits bone marrow stem cell differentiation into osteoblasts by downregulating microRNA-34a expression," *Annals of Clinical and Laboratory Science*, vol. 49, no. 3, pp. 324–329, 2019.

[46] J. Chang, F. Liu, M. Lee et al., "NF-κB inhibits osteogenic differentiation of mesenchymal stem cells by promoting β-catenin degradation," *Proceedings of the National Academy of Sciences*, vol. 110, no. 23, pp. 9469–9474, 2013.

[47] N. S. T. Hozhabri, M. D. Benson, M. D. Vu et al., "Decreasing NF-κB expression enhances odontoblastic differentiation and collagen expression in dental pulp stem cells exposed to inflammatory cytokines," *PLoS One*, vol. 10, no. 1, 2015.

[48] Y. Wang, M. Yan, Y. Yu, J. Wu, J. Yu, and Z. Fan, "Estrogen deficiency inhibits the odonto/osteogenic differentiation of dental pulp stem cells via activation of the NF-κB pathway," *Cell and Tissue Research*, vol. 352, no. 3, pp. 551–559, 2013.

[49] J. P. Alao, "The regulation of cyclin D1 degradation: roles in cancer development and the potential for therapeutic invention," *Molecular Cancer*, vol. 6, no. 1, p. 24, 2007.

[50] T. Guo, G. Cao, Y. Li et al., "Signals in stem cell differentiation on fluorapatite-modified scaffolds," *Journal of Dental Research*, vol. 97, no. 12, pp. 1331–1338, 2018.

[51] L. Liu, J. Ling, X. Wei, L. Wu, and Y. Xiao, "Stem cell regulatory gene expression in human adult dental pulp and periodontal ligament cells undergoing odontogenic/osteogenic differentiation," *Journal of Endodontics*, vol. 35, no. 10, pp. 1368–1376, 2009.

[52] H. Jiang, T. Hong, T. Wang et al., "Gene expression profiling of human bone marrow mesenchymal stem cells during osteogenic differentiation," *Journal of Cellular Physiology*, vol. 234, no. 5, pp. 7070–7077, 2019.

[53] J. Riedemann and V. Macaulay, "IGF1R signalling and its inhibition," *Endocrine-related cancer*, vol. 13, Supplement_1, pp. S33–S43, 2006.

[54] B. Reible, G. Schmidmaier, A. Moghaddam, and F. Westhauser, "Insulin-like growth factor-1 as a possible alternative to bone morphogenetic protein-7 to induce osteogenic differentiation of human mesenchymal stem cells in vitro," *International Journal of Molecular Sciences*, vol. 19, no. 6, p. 1674, 2018.

[55] P. Xue, X. Wu, L. Zhou et al., "IGF1 promotes osteogenic differentiation of mesenchymal stem cells derived from rat bone marrow by increasing TAZ expression," *Biochemical and Biophysical Research Communications*, vol. 433, no. 2, pp. 226–231, 2013.

[56] H. E. Alkharobi, H. Al-Khafaji, J. Beattie, D. A. Devine, and R. El-Gendy, "Insulin-like growth factor axis expression in dental pulp cells derived from carious teeth," *Frontiers in Bioengineering and Biotechnology*, vol. 6, 2018.

[57] J. E. Murphy-Ullrich and M. Poczatek, "Activation of latent TGF-β by thrombospondin-1: mechanisms and physiology,"

Cytokine & Growth Factor Reviews, vol. 11, no. 1-2, pp. 59–69, 2000.

[58] K. B. DuBose, M. Zayzafoon, and J. E. Murphy-Ullrich, "Thrombospondin-1 inhibits osteogenic differentiation of human mesenchymal stem cells through latent TGF-β activation," *Biochemical and Biophysical Research Communications*, vol. 422, no. 3, pp. 488–493, 2012.

[59] A. Raouf and A. Seth, "Ets transcription factors and targets in osteogenesis," *Oncogene*, vol. 19, no. 55, pp. 6455–6463, 2000.

[60] Y. Luo, Y. Zhang, G. Miao, Y. Zhang, Y. Liu, and Y. Huang, "Runx 1 regulates osteogenic differentiation of BMSCs by inhibiting adipogenesis through Wnt/β-catenin pathway," *Archives of Oral Biology*, vol. 97, pp. 176–184, 2019.

[61] G. Feng, J. Zhang, X. Feng et al., "Runx 2 modified dental pulp stem cells (DPSCs) enhance new bone formation during rapid distraction osteogenesis (DO)," *Differentiation*, vol. 92, no. 4, pp. 195–203, 2016.

[62] Y. Cao, H. Yang, L. Jin, J. Du, and Z. Fan, "Genome-wide DNA methylation analysis during osteogenic differentiation of human bone marrow mesenchymal stem cells," *Stem Cells International*, vol. 2018, 2018.

[63] C. Ye, M. Chen, E. Chen et al., "Knockdown of FOXA2 enhances the osteogenic differentiation of bone marrow-derived mesenchymal stem cells partly via activation of the ERK signalling pathway," *Cell Death & Disease*, vol. 9, no. 8, 2018.

[64] S. P. Hung, J. H. Ho, Y. R. V. Shih, T. Lo, and O. K. Lee, "Hypoxia promotes proliferation and osteogenic differentiation potentials of human mesenchymal stem cells," *Journal of Orthopaedic Research*, vol. 30, no. 2, pp. 260–266, 2012.

[65] S. Y. Kwon, S. Y. Chun, Y.-S. Ha et al., "Hypoxia enhances cell properties of human mesenchymal stem cells," *Tissue Engineering and Regenerative Medicine*, vol. 14, no. 5, pp. 595–604, 2017.

[66] Y. Zhou, W. Fan, and Y. Xiao, "The effect of hypoxia on the stemness and differentiation capacity of PDLC and DPC," *Bio Med research international*, vol. 2014, article 890675, 2014.

[67] D. Zou, W. Han, S. You et al., "In vitro study of enhanced osteogenesis induced by HIF-1α-transduced bone marrow stem cells," *Cell Proliferation*, vol. 44, no. 3, pp. 234–243, 2011.

[68] S. Melnik, N. Werth, S. Boeuf et al., "Impact of c-MYC expression on proliferation, differentiation, and risk of neoplastic transformation of human mesenchymal stromal cells," *Stem Cell Research & Therapy*, vol. 10, no. 1, p. 73, 2019.

[69] K. R. Park, H. M. Yun, I. J. Yeo, S. Cho, J. T. Hong, and Y. S. Jeong, "Peroxiredoxin 6 inhibits osteogenic differentiation and bone formation through human dental pulp stem cells and induces delayed bone development," *Antioxidants & Redox Signaling*, vol. 30, no. 17, pp. 1969–1982, 2019.

[70] E. Piek, L. S. Sleumer, E. P. van Someren et al., "Osteo-transcriptomics of human mesenchymal stem cells: accelerated gene expression and osteoblast differentiation induced by vitamin D reveals c-MYC as an enhancer of BMP2-induced osteogenesis," *Bone*, vol. 46, no. 3, pp. 613–627, 2010.

[71] X. Feng, G. Feng, J. Xing et al., "TNF-alpha triggers osteogenic differentiation of human dental pulp stem cells via the NF-kappa B signalling pathway," *Cell Biology International*, vol. 37, no. 12, pp. 1267–1275, 2013.

[72] Z. Qin, Z. Fang, L. Zhao, J. Chen, Y. Li, and G. Liu, "High dose of TNF-alpha suppressed osteogenic differentiation of human dental pulp stem cells by activating the Wnt/beta-catenin signaling," *Journal of Molecular Histology*, vol. 46, no. 4-5, pp. 409–420, 2015.

[73] N. Yang, G. Wang, C. Hu et al., "Tumor necrosis factor alpha suppresses the mesenchymal stem cell osteogenesis promoter miR-21 in estrogen deficiency-induced osteoporosis," *Journal of Bone and Mineral Research*, vol. 28, no. 3, pp. 559–573, 2013.

[74] E.-C. Kim, H.-C. Lim, O. H. Nam et al., "Delivery of dexamethasone from bioactive nanofiber matrices stimulates odontogenesis of human dental pulp cells through integrin/BMP/mTOR signaling pathways," *International Journal of Nanomedicine*, vol. 11, pp. 2557–2567, 2016.

[75] J. Zhao, J. Wu, B. Xu et al., "Kaempferol promotes bone formation in part via the mTOR signaling pathway," *Molecular Medicine Reports*, vol. 20, no. 6, pp. 5197–5207, 2019.

[76] J. Yuan, X. Liu, Y. Chen et al., "Effect of SOX2 on osteogenic differentiation of dental pulp stem cells," *Cellular and Molecular Biology (Noisy-le-Grand, France)*, vol. 63, no. 1, pp. 41–44, 2017.

[77] W. Fei, Z. Huiyu, D. Yuxin, L. Shiting, Z. Gang, and T. Yinghui, "Calcitonin gene-related peptide-induced osteogenic differentiation of mouse bone marrow stromal cells through Hippo pathway in vitro," *Hua Xi Kou Qiang Yi Xue Za Zhi*, vol. 34, no. 3, pp. 286–290, 2016.

[78] Y. Valverde, R. Narayanan, S. B. Alapati et al., "Poly (adenosine phosphate ribose) polymerase 1 inhibition enhances brain-derived neurotrophic factor secretion in dental pulp stem cell-derived odontoblastlike cells," *Journal of Endodontia*, vol. 44, no. 7, pp. 1121–1125, 2018.

[79] Q. Liu, L. Lei, T. Yu, T. Jiang, and Y. Kang, "Effect of brain-derived neurotrophic factor on the neurogenesis and osteogenesis in bone engineering," *Tissue Engineering. Part A*, vol. 24, no. 15-16, pp. 1283–1293, 2018.

[80] K. Yusa, O. Yamamoto, M. Iino et al., "Eluted zinc ions stimulate osteoblast differentiation and mineralization in human dental pulp stem cells for bone tissue engineering," *Archives of Oral Biology*, vol. 71, pp. 162–169, 2016.

[81] X.-. X. Zhao, X.-. L. An, X.-. C. Zhu et al., "Inhibiting transforming growth factor-beta signaling regulates in vitro maintenance and differentiation of bovine bone marrow mesenchymal stem cells," *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, vol. 330, no. 8, pp. 406–416, 2018.

[82] B. T. Hu and W. Z. Chen, "MOTS-c improves osteoporosis by promoting osteogenic differentiation of bone marrow mesenchymal stem cells via TGF-beta/Smad pathway," *European Review for Medical and Pharmacological Sciences*, vol. 22, no. 21, pp. 7156–7163, 2018.

[83] Q. Gu, Y. Cai, C. Huang, Q. Shi, and H. Yang, "Curcumin increases rat mesenchymal stem cell osteoblast differentiation but inhibits adipocyte differentiation," *Pharmacognosy Magazine*, vol. 8, no. 31, pp. 202–208, 2012.

[84] H. Li, L. Yue, H. Xu et al., "Curcumin suppresses osteogenesis by inducing miR-126a-3p and subsequently suppressing the WNT/LRP6 pathway," *Aging (Albany NY)*, vol. 11, no. 17, pp. 6983–6998, 2019.

[85] A. M. Gorabi, N. Kiaie, S. Hajighasemi, T. Jamialahmadi, M. Majeed, and A. Sahebkar, "The effect of curcumin on

Hindawi
BioMed Research International
Volume 2023, Article ID 9827814, 1 page
https://doi.org/10.1155/2023/9827814

Retraction

Retracted: Serum IL-37 Level Is Associated with Rheumatoid Arthritis and Disease Activity: A Meta-Analysis

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Cao, H. Shi, G. Sun et al., "Serum IL-37 Level Is Associated with Rheumatoid Arthritis and Disease Activity: A Meta-Analysis," *BioMed Research International*, vol. 2021, Article ID 6653439, 7 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6653439, 7 pages
https://doi.org/10.1155/2021/6653439

Review Article

Serum IL-37 Level Is Associated with Rheumatoid Arthritis and Disease Activity: A Meta-Analysis

Shengnan Cao,[1,2] Haojun Shi,[3] Guodong Sun,[4] Yuanzhen Chen,[2] Guangjian Hou,[2] Dandan Wang,[2] and Bin Shi[1,2]

[1]*School of Acupuncture-Tuina, Shandong University of Traditional Chinese Medicine, Jinan 250062, China*
[2]*Bone Biomechanics Engineering Laboratory of Shandong Province, Neck-Shoulder and Lumbocrural Pain Hospital of Shandong First Medical University, Jinan 250062, China*
[3]*Second Clinical Medical College, Henan University of Chinese Medicine, Henan 450016, China*
[4]*Department of Rehabilitation Medicine, The Third Affiliated Hospital of Shandong First Medical University, Jinan 250062, China*

Correspondence should be addressed to Dandan Wang; 158wdd@163.com and Bin Shi; sdyky-shibin@163.com

Received 16 December 2020; Revised 7 January 2021; Accepted 20 January 2021; Published 8 February 2021

Academic Editor: Min Tang

Interleukin-37 (IL-37) inhibits the pathogenesis of rheumatoid arthritis (RA) via downregulating proinflammatory cytokines. Accordingly, we performed an analysis to accurately assess the relationship between serum IL-37 cytokine levels and disease activity of RA. Subgroup analysis and sensitivity analysis were applied to explore the sources of heterogeneity. Correlation coefficient (r) was utilized to evaluate the relationship between IL-37 and disease activity of RA patients. Ten studies were included into the research. Functional analysis revealed elevated serum IL-37 concentrations in RA patients (SMD = 1.61, $P < 0.00001$). The relationship between serum IL-37 levels and disease activity was statistically significant (C-reactive protein: $r = 1.47$, $P = 0.0002$; erythrocyte sedimentation rate: $r = 1.55$, $P < 0.00001$; rheumatoid factor: $r = 1.40$, $P = 0.004$; tumor necrosis factor-α: $r = 1.64$, $P = 0.0003$; Disease Activity Score for 28 joints: $r = 1.63$, $P < 0.00001$; tender joint count: $r = 1.48$, $P < 0.00001$; and swollen joint count: $r = 1.52$, $P = 0.0003$), but anti-CCP was not significant (anti-CCP: $r = 0.98$, $P = 0.72$). In summary, these data are suggesting that the elevated serum level of IL-37 in RA is positively correlated with the disease activity of RA, suggesting a role for IL-37in the pathogenesis of RA.

1. Introduction

Rheumatoid arthritis (RA) has an estimated prevalence of 1% to 2% worldwide and seriously impacts patients' life quality [1, 2]. Pathological manifestation of RA is chronic inflammation of multiple synovial joints, and progressive inflammation can cause destruction of articular cartilage and bone, eventually leading to irreversible progressive joint deformity and loss of function [3]. In addition, studies showed that interleukin promoted disease development by activating the expression of genes involved in tissue degradation, synovial tissue proliferation, and joint erosion [4]. Targeting Interleukin-17 (IL-17) provides a potential therapy [5]. However, treatment of RA patients with monoclonal antibodies against IL-17 may cause adverse effects [6]. Therefore, the search for new cytokines related to RA and have therapeutic potential is still continuing.

IL-37 is a newly discovered natural immune response inhibitor and exerts anti-inflammatory effects by suppressing proinflammatory cytokines [7]. IL-37 includes five splice variants (IL-37a–e), which have intracellular and extracellular functions same as cytokines [8]. Research demonstrated that IL-37 suppressed certain cells such as macrophages, monocytes, and epithelial cells to secrete proinflammatory cytokines [8]. Abnormal IL-37 expressions in serum have been reported in several inflammation-related diseases, such as RA and ankylosing spondylitis [9, 10]. In in vitro experiments, IL-37 inhibited joint inflammation and significantly reduced the expression of IL-17, IL-1, and IL-6 in Th17 cells.

[11]. Studies demonstrate that IL-37 may play a critical role in the pathogenesis of RA. To define the role of IL-37 in the progression of RA, we performed the current meta-analysis.

2. Materials and Methods

2.1. Literature Search Strategy. This article follows the criteria of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses, with relevant research searched up to November 2020, including the Embase, Web of Science, PubMed, and Chinese National Knowledge Infrastructure (CNKI) databases. The keywords used for search were as follows: "Interleukin-37" or "Interleukin 37" or "IL-37" or "IL 37," "Rheumatoid Arthritis" or "RA" or "Arthritis, Rheumatoid." Publication languages were limited to English and Chinese. The search strategy and study selection were independently performed by two researchers on a database search, and the final research choices were agreed upon. Two researchers independently evaluated the same article to determine their eligibility for inclusion and resolved differences through consensus.

2.2. Inclusion and Exclusion Criteria. The included researches were based on the following criteria: (1) explored the relationship between serum IL-37 levels and RA; (2) observational studies, such as cross-sectional, case-control, or cohort designs in human; (3) measured the IL-37 level, and/or disease activity in patients and controls must be provided; and (4) the control group must be healthy controls (HCs). Excluded studies were based on the following criteria: (1) do not meet the inclusion criteria; (2) complete data not provided or cannot calculate the required data; (3) cannot obtain the full text or duplicated studies; (4) reviews or comments; and (5) animal research.

2.3. Data Extraction. The following information was collected from each study: publication year, country, ethnicity, specimen source, gender and age, case and control numbers, period of RA, and the serum IL-37 level. In addition, the correlation coefficient (r) between the IL-37 level and the disease activity was also collected, including erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), anticyclic citrullinated peptide (anti-CCP), rheumatoid factor (RF), Disease Activity Score for 28 joints (DAS28), tender joint count (TJC), swollen joint count (SJC), and TNF-α.

2.4. Quality Assessment. All studies were assessed using the Newcastle-Ottawa Scale (NOS) assessment scale [12]. The relevant data was independently extracted by two reviewers (Shengnan Cao and Haojun Shi), and disagreement was resolved through discussion and referred to a third reviewer (Bin Shi) if necessary. The scale includes nine items covering three dimensions. A point is awarded for each item that is satisfied by the study. The NOS scores range from 0 to 9, with higher scores indicating higher quality. In this study, score ≥ 6 was defined as high quality.

2.5. Statistical Analysis. We used the Review Manager (RevMan 5.3, Cochrane Collaboration, Nordic Cochrane Center, Copenhagen, Denmark) for statistical analyses. The difference between IL-37 in RA and HC was calculated by standardized mean differences (SMDs) and 95% confidence intervals (CIs). When some studies reported results using the median of the first and third quartiles (M (P_{25}, P_{75})), an approximation method was used to calculate the mean and standard deviation ($X \pm S$) [13]: $X \approx (P_{25} + M + P_{75})/3$ and $S = (P_{75} - P_{25})/1.35$. Besides, all standard errors of the correlation coefficient (SE) were calculated by the following formula [14]: $\mathrm{SE} = \sqrt{((1-r)/(n-1))^2}$ to combine the r of the random-effect model. Q-statistic ($P < 0.05$) and I^2 statistics ($I^2 > 50\%$) were used to assess the heterogeneity among studies. In individual studies, we used the Cochrane Collaboration's tool to assess the risk of bias. $P < 0.05$ means statistical significant.

3. Results

3.1. Included Studies and Quality Assessment. After a comprehensive literature search and verification, the author's research yielded a total of 390 records, of which 47 duplicate records were excluded. Also, 163 nonrelevant studies, 76 nonhuman studies, and 68 conference or comment abstracts were excluded. Finally, 10 studies were eligible for the systematic analysis [15–24] (Figure 1). In total, 731 RA patients were included in this study, of which 529 were females. Data characteristics of the included studies are in Table 1.

3.2. Outcomes and Sensitivity Analysis. Through the heterogeneity test, the results showed that the IL-37 level was significantly higher in RA (SMD = 1.61, $P < 0.00001$). Sensitivity analysis showed that the combined SMDs did not change significantly, indicating that our results were stable. The Cochrane Risk of Bias Tool was used to judge the risk of bias for each included study. Though all studies were peer-reviewed and did not address conflicts of interest, most of the included studies did not provide sufficient information to assess overall quality (Figure 2). Small subject numbers can increase heterogeneity and bias; the publication bias was examined using the funnel plot.

3.3. Subgroup Analysis. Since heterogeneity existed among included articles, subgroup analysis was performed according to the ethnicity, sample size, and disease period. For ethnicity, the serum IL-37 level of RA patients in Asians (SMD = 1.42, 95%CI = 1.27-1.58) and African (SMD = 1.75, 95%CI = 1.26-2.24). According to the sample size stratification, for RA patients, IL-37 levels were identified in both small sample size and large sample size (small: SMD = 4.53, 95% CI = 2.61-6.46; large: SMD = 1.67, 95%CI = 0.40-2.95). Furthermore, the significant difference was shown in the disease period (active: SMD = 3.67, 95%CI = 3.23-4.11; inactive: SMD = 1.86, 95%CI = 1.45-2.27) (Table 2).

3.4. Correlation of IL-37 Level and RA. The relationship between the IL-37 level and disease activity of RA was investigated. The serum IL-37 level was positively correlated with CRP, ESR, RF, and TNF-α levels (CRP: $r = 1.47$, $P = 0.0002$; ESR: $r = 1.55$, $P < 0.00001$; RF: $r = 1.40$, $P = 0.004$; TNF-α: $r = 1.64$, $P = 0.0003$) (Table 3). In addition to anti-CCP, the DAS28 scores, TJC counts, and SJC counts were also correlated with the IL-37 level (DAS28: $r = 1.63$, $P < 0.00001$;

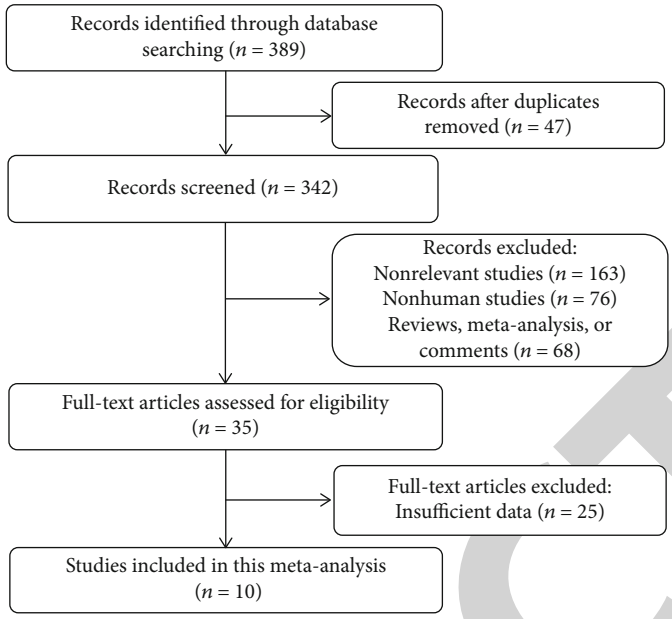


Figure 1: Flow diagram for study selection.

Table 1: Characteristics of eligible studies.

References	Country	Gender		Age		No.		NOS
		Experimental	Control	Experimental	Control	Experimental	Control	
Zhang and Li (2016) [23]	China	M: 17 F: 62	M: 34 F: 66	53.9 ± 9.4	51.7 ± 10.20	79	100	7
Chen and Wang (2014) [19]	China	M: 11 F: 57	M: 4 F: 16	54.41 ± 14.09	53.5 ± 10.55	68	20	7
Song et al. (2018) [16]	China	M: 7 F: 52	M: 17 F: 29	50.4 ± 14.7	44.7 ± 13.70	59	46	7
Yang et al. (2015) [18]	China	M: 83 F: 69	M: 60 F: 40	52.3 ± 15.6	52 ± 14.10	152	100	7
Wu et al. (2019) [21]	China	M: 13 F: 12	NA	48.4 ± 10.16	47.2 ± 9.98	25	25	6
Wan et al. (2016) [20]	China	M: 10 F: 35	M: 6 F: 15	41.3 ± 2.1	41.2 ± 2.30	45	21	6
Xu et al. (2017) [22]	China	M: 12 F: 47	NA	59.86 ± 11.50	NA	59	30	6
Xia et al. (2015) [17]	China	M: 22 F: 128	M: 6 F: 44	51.3 ± 18.6	45.2 ± 20.40	150	50	7
Ragab et al. (2019) [15]	Africa	M: 8 F: 40	M: 7 F: 35	48.67 ± 20.74	44.7 ± 13.70	48	42	6
Akram et al. (2018) [24]	Pakistan	M: 19 F: 27	M: 11 F: 9	54.33 ± 34.07	44.0 ± 12.59	46	20	6

M: male; F: female; No: number; NOS: Newcastle-Ottawa Scale; NA: not available.

TJC: $r = 1.48$, $P < 0.00001$; SJC: $r = 1.52$, $P = 0.0003$; anti-CCP: $r = 0.98$, $P = 0.72$) (Figure 3).

4. Discussion

In this study, we reviewed the literature about the association between serum IL-37 level and RA. The results consistently prove that the IL-37 level in RA was significantly elevated. At the same time, it demonstrated that the IL-37 level was associated with disease activity of RA.

RA is the most commonly diagnosed systemic inflammatory arthritis. Generally, IL-37 is an important anti-inflammatory cytokine involved in inflammation regulation and can inhibit the expression and function of proinflammatory cytokines [25]. Researchers reported that the plasma level of IL-37 was increased in RA [26]. Meanwhile, as demonstrated

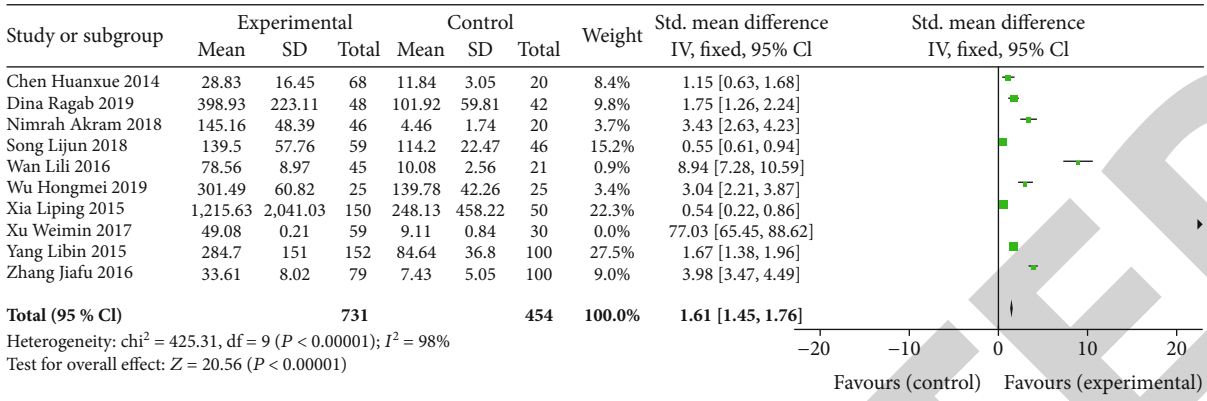


Figure 2: Forest plots for SMDs of IL-37 levels with RA and HCs.

Table 2: Results of subgroup analysis of included literatures.

Subgroup	*N*	SMD (95% CI)	*Z*	*P*	Heterogeneity		Model
					I^2 (%)	*P*	
Ethnicity							
Asians	9	1.42 (1.27, 1.58)	18.11	<0.00001	98%	<0.00001	R
African	1	1.75 (1.26, 2.24)	7.01	<0.00001	NA	NA	R
Sample size							
Large	4	1.67 (0.40, 2.95)	2.57	0.01	98%	<0.00001	R
Small	6	1.17 (0.96, 1.37)	11.28	<0.00001	96%	<0.00001	R
Period							
Active	3	3.67 (3.23, 4.11)	16.38	<0.00001	98%	<0.00001	R
Inactive	3	1.86 (1.45, 2.27)	8.93	<0.00001	97%	<0.00001	R
All	10	3.25 (2.21, 4.29)	6.13	<0.00001	98%	<0.00001	R

N: number; CI: confidence interval; R: random-effect model; SMD: standard mean difference; NA: not available.

Table 3: Correlation between serum IL-37 level and disease activity in RA patients.

Variable	No.	*r*	95% CI (LCI, UCI)	*Z*	*P*	Model
CRP	5	1.47	(1.20, 1.79)	3.73	0.0002	R
ESR	4	1.55	(1.41, 1.71)	9.13	<0.00001	R
Anti-CCP	3	0.98	(0.88, 1.09)	0.36	0.72	R
RF	4	1.40	(1.11, 1.75)	2.89	0.004	R
DAS28	6	1.63	(1.38, 1.93)	5.70	<0.00001	R
TJC	3	1.48	(1.25, 1.77)	4.46	<0.00001	R
SJC	3	1.52	(1.20, 1.93)	3.43	0.0006	R
TNF-α	2	1.64	(1.26, 2.13)	3.66	0.0003	R

No: number; CI: confidence interval; R: random-effect model; *r*: correlation coefficient.

by authors, IL-37 levels in serum and synovial fluid of RA patients were elevated markedly, whereas they were almost undetectable in healthy controls. Our meta-analysis results showed that when analyzed by ethnic subgroup, in Asian RA patients (mainly Chinese), IL-37 levels were significantly higher than in healthy people. IL-37 was expressed in the thymus, bone marrow, lymph nodes, NK cells, monocytes, and B cells [27]. Studies have shown that IL-37 could be upregulated under the action of inflammatory stimulation and cytokines, such as TLR agonists, IL-1β, IL-18, and TNF-α. IL-37 may mediate the negative feedback mechanism that inhibits excessive proinflammatory cytokines in RA patients [26]. Accordingly, the increase of anti-inflammatory cytokines IL-37 may be a potential mechanism to reduce the severity of joint inflammation and disease. Series studies have shown that IL-37 may be a potential biomarker for RA diagnosis, disease activity assessment, or curative effect observation [28].

RF form part of the differential diagnosis of arthropathies, and the main index of serological diagnosis of RA. Serological indicators of ESR and CRP are nonspecific inflammatory markers, which have been used to assess systemic inflammatory responses, and have proven diagnostic value for RA. DAS28 is a quantitative index for evaluating RA disease activity [29, 30]. Simultaneously, TJC and SJC are counting tools to detect joint tenderness and swelling. Therefore, we explored the potential relationship between IL-37 levels and related parameters. Results indicated that the CRP, ESR, RF, TNF-α, DAS28, TJC, and SJC levels of RA patients were positively correlated with the IL-37 level. Consistent with our findings, researchers reported a positive correlation between the plasma level of IL-37 and both CRP

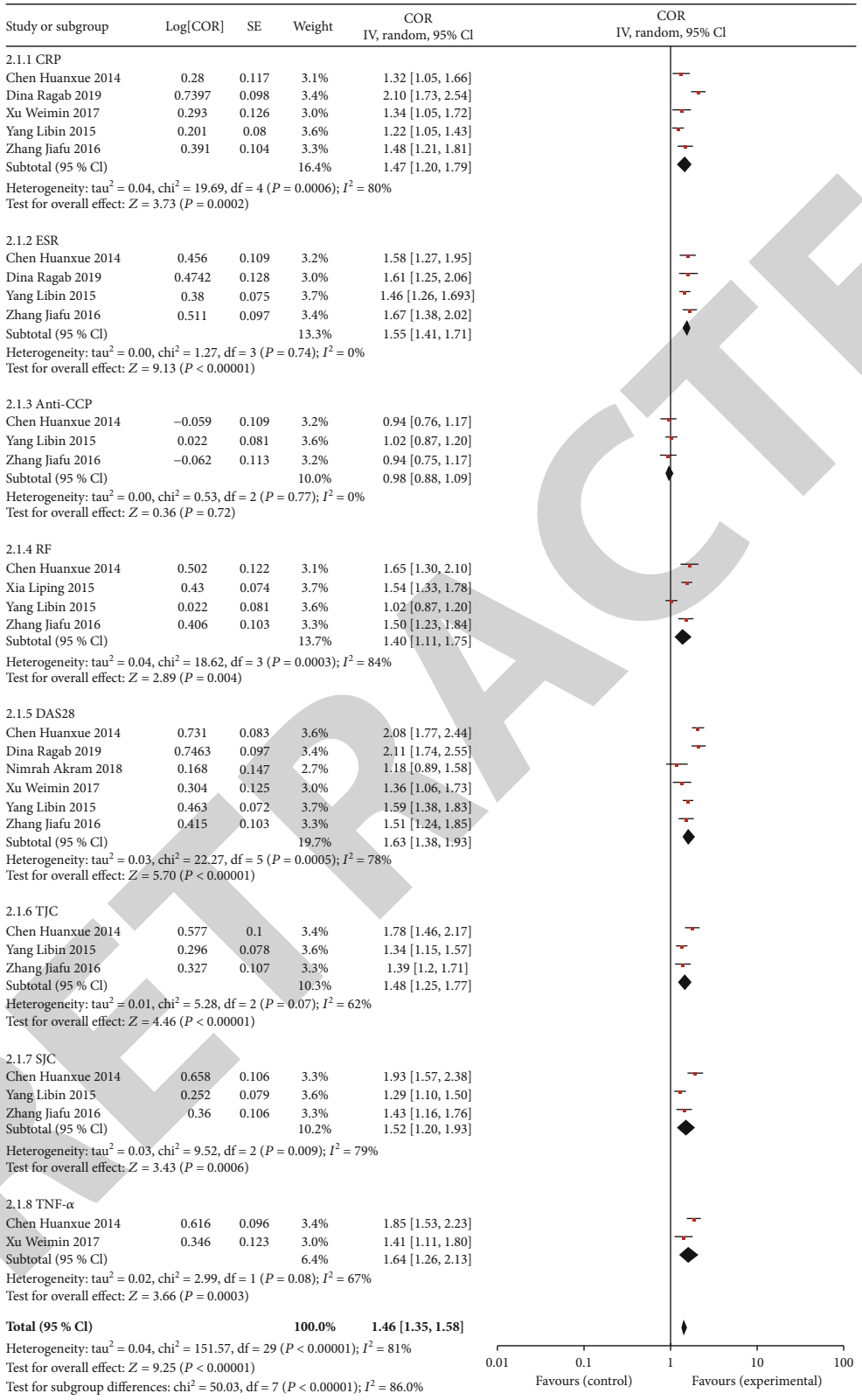


Figure 3: Forest plot of the C-reactive protein (CRP), erythrocyte sedimentation rate (ESR), anticyclic citrullinated peptide (anti-CCP), rheumatoid factor (RF), Disease Activity Score for 28 joints (DAS28), tender joint count (TJC), swollen joint count (SJC), and tumor necrosis factor-α (TNF-α).

and DAS28 [28]. The TNF-α is an important factor leading to the onset of RA, and with the increases of inflammatory response and the proinflammatory cytokine TNF-α, IL-37 also increases, which is closely related to disease activities and may play a protective role in the occurrence and development of RA [31, 32]. Anti-CCP antibody is a well-known serological marker [33]. In the diagnosis of early RA, it improves the diagnostic accuracy by measuring the anti-CCP antibody and RF [34]. However, this study used these clinical indicators to explore the relationship between IL-37 levels and RA disease activity and found that the anti-CCP level of RA patients was not correlated with the IL-37 level. This may be caused by the limited number of relevant studies and cases.

However, some limitations should be considered. The meta-analysis results of the included articles were relatively heterogeneous. After subgroup analysis, the source of heterogeneity was not found. This may be due to differences in experimental design and quality, which needs further investigation. At the same time, the insufficient number of related studies and cases is another limit that should also be considered. Finally, RA patients were mainly from Chinese, so our conclusions cannot be extrapolated to other populations.

In all, our meta-analysis provided that the serum IL-37 level is significantly elevated in RA. The serum IL-37 level is positively correlated with disease activity of RA, and further studies are needed to verify our findings.

Conflicts of Interest

All authors declare they have no conflicts of interest.

Authors' Contributions

All authors had access to the data and a role in writing the manuscript. All authors read and approved the final manuscript. Shengnan Cao and Haojun Shi contributed equally to this work.

Acknowledgments

This study was supported by the Distinguished Experts of Taishan Scholar Project (grant number ts201511074), Natural Science Foundation of Shandong Province (grant numbers ZR2019MH134 and ZR2018LH019), Academic Promotion Project of the Shandong First Medical University (grant number 2019QL003), China Postdoctoral Science Foundation (grant number 2019M662420), Shandong Province Postdoctoral Innovation Project (grant number 202002049), and TCM Technology Development Project of Shandong Province (grant numbers 2019-0543 and 2019-0545).

References

[1] E. Minichiello, L. Semerano, and M. C. Boissier, "Time trends in the incidence, prevalence, and severity of rheumatoid arthritis: a systematic literature review," *Joint, Bone, Spine*, vol. 83, no. 6, pp. 625–630, 2016.

[2] R. P. Zhou, X. S. Wu, Y. Y. Xie et al., "Functions of interleukin-34 and its emerging association with rheumatoid arthritis," *Immunology*, vol. 149, no. 4, pp. 362–373, 2016.

[3] E. H. El-Sayed, M. H. Saleh, M. H. Al-Shahaly, E. A. Toraih, and A. Fathy, "IL-37 gene variant (rs 3811047): a marker of disease activity in rheumatoid arthritis: a pilot study," *Autoimmunity*, vol. 51, no. 8, pp. 378–385, 2018.

[4] J. Sharma, S. Bhar, and C. S. Devi, "A review on interleukins: the key manipulators in rheumatoid arthritis," *Modern Rheumatology*, vol. 27, no. 5, pp. 723–746, 2017.

[5] R. Fleischmann, M. Schiff, D. van der Heijde et al., "Baricitinib, methotrexate, or combination in patients with rheumatoid arthritis and no or limited prior disease-modifying antirheumatic drug treatment," *Arthritis & Rheumatology*, vol. 69, no. 3, pp. 506–517, 2017.

[6] M. Wei and D. Duan, "Efficacy and safety of monoclonal antibodies targeting interleukin-17 pathway for inflammatory arthritis: a meta-analysis of randomized controlled clinical trials," *Drug Design, Development and Therapy*, vol. 10, pp. 2771–2777, 2016.

[7] W. D. Xu, Y. Zhao, and Y. Liu, "Insights into IL-37, the role in autoimmune diseases," *Autoimmunity Reviews*, vol. 14, no. 12, pp. 1170–1175, 2015.

[8] D. Boraschi, D. Lucchesi, S. Hainzl et al., "IL-37: a new anti-inflammatory cytokine of the IL-1 family," *European Cytokine Network*, vol. 22, no. 3, pp. 127–147, 2011.

[9] B. Chen, K. Huang, L. Ye et al., "Interleukin-37 is increased in ankylosing spondylitis patients and associated with disease activity," *Journal of Translational Medicine*, vol. 13, no. 1, p. 36, 2015.

[10] L. Ye, L. Ji, Z. Wen et al., "IL-37 inhibits the production of inflammatory cytokines in peripheral blood mononuclear cells of patients with systemic lupus erythematosus: its correlation with disease activity," *Journal of Translational Medicine*, vol. 12, no. 1, p. 69, 2014.

[11] X. Wang, K. Xu, S. Chen, Y. Li, and M. Li, "Role of interleukin-37 in inflammatory and autoimmune diseases," *Iranian journal of Immunology*, vol. 15, no. 3, pp. 165–174, 2018.

[12] A. Stang, "Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses," *European Journal of Epidemiology*, vol. 25, no. 9, pp. 603–605, 2010.

[13] X. Wan, W. Wang, J. Liu, and T. Tong, "Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range," *BMC Medical Research Methodology*, vol. 14, no. 1, p. 135, 2014.

[14] G. Cai, L. Wang, D. Fan et al., "Vitamin D in ankylosing spondylitis: review and meta-analysis," *Clinica Chimica Acta*, vol. 438, pp. 316–322, 2015.

[15] D. Ragab, S. Mobasher, and E. Shabaan, "Elevated levels of IL-37 correlate with T cell activation status in rheumatoid arthritis patients," *Cytokine*, vol. 113, pp. 305–310, 2019.

[16] L. Song, Y. Wang, Y. Sui et al., "High interleukin-37 (IL-37) expression and increased mucin-domain containing-3 (TIM-3) on peripheral T cells in patients with rheumatoid arthritis," *Medical Science Monitor*, vol. 24, pp. 5660–5667, 2018.

[17] L. Xia, H. Shen, and J. Lu, "Elevated serum and synovial fluid levels of interleukin-37 in patients with rheumatoid arthritis: attenuated the production of inflammatory cytokines," *Cytokine*, vol. 76, no. 2, pp. 553–557, 2015.

Hindawi
BioMed Research International
Volume 2023, Article ID 9827010, 1 page
https://doi.org/10.1155/2023/9827010

Retraction

Retracted: Potential Drug Prediction of Glioblastoma Based on Drug Perturbation-Induced Gene Expression Signatures

BioMed Research International

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] B. Zhu, X. Mao, and Y. Man, "Potential Drug Prediction of Glioblastoma Based on Drug Perturbation-Induced Gene Expression Signatures," *BioMed Research International*, vol. 2021, Article ID 6659701, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6659701, 11 pages
https://doi.org/10.1155/2021/6659701

Research Article

Potential Drug Prediction of Glioblastoma Based on Drug Perturbation-Induced Gene Expression Signatures

Bochi Zhu, Xijing Mao, and Yuhong Man

Department of Neurology, The Second Hospital of Jilin University, Changchun City, Jilin Province, 130041, China

Correspondence should be addressed to Yuhong Man; myh1996@163.com

Received 22 October 2020; Revised 11 January 2021; Accepted 12 January 2021; Published 27 January 2021

Academic Editor: Min Tang

Objectives. Glioblastoma (GBM) is a malignant brain tumor which is the most common and aggressive type of central nervous system cancer, with high morbidity and mortality. Despite lots of systematic studies on the molecular mechanism of glioblastoma, the pathogenesis is still unclear, and effective therapies are relatively rare with surgical resection as the frequently therapeutic intervention. Identification of fundamental molecules and gene networks associated with initiation is critical in glioblastoma drug discovery. In this study, an approach for the prediction of potential drug was developed based on perturbation-induced gene expression signatures. *Methods*. We first collected RNA-seq data of 12 pairs of glioblastoma samples and adjacent normal samples from the Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) were identified by DESeq2, and coexpression networks were analyzed with weighted gene correlation network analysis (WGCNA). Furthermore, key driver genes were detected based on the differentially expressed genes and potential chemotherapeutic drugs and targeted drugs were found by correlating the gene expression profiles with drug perturbation database. Finally, RNA-seq data of glioblastoma from The Cancer Genome Atlas (TCGA) dataset was collected as an independent validation dataset to verify our findings. *Results*. We identified 1771 significantly DEGs with 446 upregulated genes and 1325 downregulated genes. A total of 24 key drivers were found in the upregulated gene set, and 81 key drivers were found in the downregulated gene set. We screened the Crowd Extracted Expression of Differential Signatures (CREEDS) database to identify drug perturbations that could reverse the key factors of glioblastoma, and a total of 354 drugs were obtained with p value $< 10^{-10}$. Finally, 7 drugs that could turn down the expression of upregulated factors and 3 drugs that could reverse the expression of downregulated key factors were selected as potential glioblastoma drugs. In addition, similar results were obtained through the analysis of TCGA as independent dataset. *Conclusions*. In this study, we provided a framework of workflow for potential therapeutic drug discovery and predicted 10 potential drugs for glioblastoma therapy.

1. Introduction

Glioblastoma (GBM) is a malignant brain tumor which arises from glial cells and is the most common and aggressive type of central nervous system cancer worldwide [1, 2]. To date, surgical resection combined with radiation therapy and chemotherapy is still the frequently therapeutic intervention for GBM [3]. Although some advances in treatment of glioblastoma have been made in recent years, its prognosis is still poor because of its invasive and aggressive behavior. The annual incidence rate of glioblastoma in China is 5-8 per 10 million, and the five-year survival ratio is lower than 5% with a median survival time of 12.6 months. Despite extensive systematic studies on the mechanism of tumorigenesis, metastasis, and recurrence, the underlying mechanism of glioblastoma is still unclear [4] and effective drugs are relatively rare. Therefore, it is crucial to conduct further studies to identify fundamental molecules and gene networks associated with initiation and development, as well as prognosis of GBM, and explore more effective drugs.

High-throughput sequencing has been widely used in cancer research over the past decade, which greatly promotes the understanding of molecular genetics of glioblastoma. TCGA, a landmark cancer genomics program, which contains genomic, epigenomic, transcriptomic, and proteomic data, has also improved our ability to diagnose, treat, and

prevent cancer, including glioblastoma. Thus, gene expression profiling has become an objective and important method to classify tumors besides histological classification [5, 6]. A classification of GBM based on platelet-derived growth factor receptor (PDGFR) and the epidermal growth factor receptor (EGFR) has been built [7]. Moreover, molecular subclasses could be utilized to predict prognosis of patients. Methylation status of the O6-methylguanine DNA methyltransferase (MGMT) gene promoter and isocitrate dehydrogenase enzyme 1/2 (IDH1/2) mutation were the prognostic molecules which have been fully confirmed [8, 9]. In addition, IDH mutation status and telomerase reverse transcriptase (TERT) promoter mutation status could enhance prognostic stratification of patients with glioblastoma. Patients with MGMT promoter methylation were more sensitive to temozolomide (TMZ) chemotherapy and had a better prognosis [10]. BRAF (v-raf murine viral oncogene homolog B1) mutations were found in pilocytic astrocytomas and have been proved to be a new therapeutic target for inhibition of the mitogen-activated protein kinase (MAPK) cascade, but its prognostic significance is unknown [11]. Vemurafenib, as a BRAF inhibitor, can have a complete clinical regression of relapsed glioblastoma multiforme [12]. This success has brought great encouragement to the targeted therapies of glioblastoma.

On the other hand, new drug discovery faces many serious challenges, such as a long developing period, substantial cost, high attrition rates, and changing regulatory requirements, which can all contribute to lower yielding for the pharmaceutical industry and a less desirable choice for inventors [13]. Drug repurposing is a strategy for identifying new indication for approved drugs or preclinical compounds. Comparing to developing entirely new drugs from scratch, drug repurposing offers various advantages. The promise of drug repurposing is to accelerate translating the benefits including technology and enhanced knowledge of human disease into therapeutic advances bypassing more time and cost-efficiency. One of the most promising novel methodologies for drug repurposing is computational approaches, such as signature matching, genetic association, and pathway mapping [14, 15]. Omics data can be useful in figuring out not only the mechanism of disease but also pharmacology. By now, the omics technologies can be applied at any molecular levels from genes, RNA, and proteins to metabolites corresponding to genome, transcriptome, and metabolome, respectively. The most widely used omics in drug repurposing is transcriptome and based on which several computational approaches have been proposed [16, 17]. The majority of these strategies were based on transcription profile from cell lines given several large projects based on transcriptome such as LINCS (the library of integrated network-based cellular signatures) [18]. However, there are essential differences between cell lines and tissues, much less to organ and human beings. Therefore, results inferred from omics data based on tissues are more beneficial to drug repurposing.

In this study, the original RNA-seq data with clinical information of glioblastoma samples and adjacent normal samples was downloaded from GEO as testing data. We obtained DEGs by DESeq2, compared gene module changes, and retrieved the potential drugs from CREEDS. Through these studies, we design to understand the mechanism of tumorigenesis and develop systematic research programs for GBM in the future.

2. Materials and Methods

2.1. Data Collecting. We collected RNA sequencing data of glioblastoma from TCGA and GEO dataset, respectively. 12 pairs of tumor samples and adjacent normal samples were firstly collected from the project ID of GEO-GSE151352 as the testing dataset. For a better verification of our findings, we used an additional data from TCGA as an independent validation dataset which contained 156 tumor samples and 5 normal samples with an expression profile of 60483 genes of each sample.

2.2. Differential Gene Expression Analysis between Cancer and Normal Samples. As the input data, the raw read counts were extracted from TCGA dataset or GEO dataset and R package DESeq2 [19] was used for differential gene expression analysis with log2 | fold change | >THRESHOLD_FOLDCHANGE and adjusted p value < THRESHOLD_P_ADJ as threshold, where THRESHOLD_FOLDCHANGE and THRESHOLD_P_ADJ are shown in Table 1. The differentially expressed gene analysis was performed between two different conditions. To analyze the biological significance, the biological functions were annotated for the up- or down-regulated gene set. Gene Ontology [1, 20] was set as the reference database for the enrichment analysis. R package clusterProfiler [21] was used for the calculations. The R package ggplot2 was used to plot the enriched pathways or functions.

2.3. Weighted Gene Correlation Network Analysis and Key Driver Analysis. The genes with ENSEMBL type were converted to gene symbol using map Ids function from the R package clusterProfiler. For WGCNA [22], we used the R package WGCNA for the module discovery. Then, we performed key driver analysis (KDA) as described by Yang et al. [23]. The dynamic neighborhood search (DNS) was applied for KDA. Firstly, we generated a subnetwork NG that is within 2 steps of the nodes in the given gene set. Secondly, for each gene in the NG, DNS was used to find the genes that are within 2 steps of the gene. Thirdly, the hypergeometric test is performed to calculate the p value of the enrichment between the gene set from the second step and the input gene set, with the gene set from the first step as the background. The parameters used in the KDA were as follows: with adjusted p value < 0.05 for DEG and p value for subnet < 0.01 as threshold. We considered key driver genes as hub genes that connected to the up- or downregulated genes.

2.4. Drug Prediction. The DEGs and key drivers were applied into the CREEDS to find potential drugs. The dataset could be gained from http://amp.pharm.mssm.edu/creeds.

Table 1: Pipelines and parameters for this work.

Step	Goal	Method	Parameters
1	Differentially expressed genes discovery	DESeq2	For GEO dataset: log 2 | fold change | >1, adjusted p value < 0.05; for TCGA dataset: log 2 | fold change | >2, adjusted p value < 0.001.
2	Coexpression analysis	WGCNA	For the T group, the soft power threshold was set to 14. For the N group, the soft power threshold was set to 10.
3	Key driver analysis	Key connector analysis	The protein-protein interaction network was used as the reference network for KDA. Parameters were as follows: –nlayerExpansion for 1, –nlayerSearch for 6, –enrichedNodesPercentCut for -1.
4	Drug discovery	Hypergeometric test	The function phyper was used, and the lower tail parameter was set to true.

3. Results

3.1. Potential Drug Prediction Based on Drug Perturbation-Induced Gene Expression Signatures. We developed a new approach for drug prediction based on drug perturbation-induced gene expression signatures, including DEG analysis, Gene Ontology (GO) enrichment, and key driver analysis, as well as coexpression from WGCNA. Then, the key driver and coexpression genes were combined as key factors of glioblastoma to detect gene and drug perturbation signatures with network pharmacology (Figure 1).

3.2. Differentially Expressed Genes Were Gained and Enriched in Different Pathways. We compared the expression profile between the tumor group and the normal group by DEG analysis using DESeq2 [19]. We used the threshold log 2 | fold change | >1 and adjusted p value < 0.05 for adjusted p value. Finally, we got 1771 significantly different expression genes, with 446 upregulated genes (25.2%) and 1325 (74.8%) downregulated genes. The detailed information (log2FC, adjusted p value) of 1771 DEGs is listed in Supplementary Table 1. The top 10 upregulated genes and downregulated genes are shown, respectively, in Figure 2(a).

For a better understanding of functions of the upregulated and downregulated genes, we performed GO enrichment analysis on the two gene sets separately, and the threshold of the significance was adjusted p value < 0.05. We chose the top ten pathways with the highest proportion of genes for display (Figures 2(b) and 2(c)). The enrichment results showed that the upregulated genes mainly play roles in DNA conformation change, DNA packaging, chromatin assembly or disassembly pathways, extracellular matrix organization, nucleosome organization, and DNA replication-dependent nucleosome assembly or organization, which play an important role in cell differentiation, proliferation, and apoptosis, while the downregulated genes are mainly distributed in signal pathways related to the function of brain nerve cells, such as the modulation of chemical synaptic transmission and the regulation of transsynaptic signaling, which suggests that the occurrence of glioblastoma may affect the function of normal cells.

3.3. Weighted Gene Correlation Network Analysis Found Coexpression Modules Perturbed by GBM. To understand the relationship of differentially expressed genes, we attempted to modularize the genes into different modules using WGCNA. The DEG data of the normal group and the GBM group was analyzed separately to observe the synergistic effect of gene expressions under different conditions. The genes with zero expression were removed in all samples; then, the samples with a partially segregated group of the hierarchical clustering results were deleted, and then, weighted gene correlation network analysis was performed. We finally obtained 44 gene modules in the GBM group and 31 gene modules in the normal group. The number of modules and submodules in the GBM group increased, but the coexpression effect of genes decreased, and the signal pathway regulation system tended to be disordered. Then, we randomly selected 400 genes to construct a topological overlapping heat map (Figure 3(a)) and made functional enrichment analysis on the modules (Figures 3(b) and 3(c). As is shown, highly coexpressed genes mainly focused on the mRNA processing, translational initiation, and skeletal system development pathways.

3.4. Key Driver Analysis Found Key Drivers Regulating the Gene Sets. We performed key driver analysis of up- and downexpressed genes separately. 24 key drivers were found in the upregulated gene set, and 81 key drivers were found in the downregulated gene set. Top key drivers and corresponding up- or downregulated genes were selected as shown in Figure 4. CAMK2G is short for Calmodulin-Dependent Protein Kinase II Gamma and is one of the hub genes of downregulated genes (Figure 4(b)). Among its related pathways are neuroscience and translation regulation by alpha-1 adrenergic receptors. CDH1 (Cadherin 1) is another hub gene of downregulated genes and participated in regulation of Wnt-mediated beta catenin signaling and target gene transcription (Figure 4(a)). CACNA1G (Calcium Voltage-Gated Channel Subunit Alpha1 G) is a hub gene of upregulated genes that related to regulation of Wnt-mediated beta catenin signaling, target gene transcription, and sweet taste signaling pathways (Figure 4(b)). GLRA3 (Glycine Receptor Alpha 3) is a hub gene of downregulated genes and encodes a member of the ligand-gated ion channel protein family (Figure 4(c)). Its encoded protein is a member of the glycine receptor subfamily. We further discussed the mechanism of these genes in Discussion.

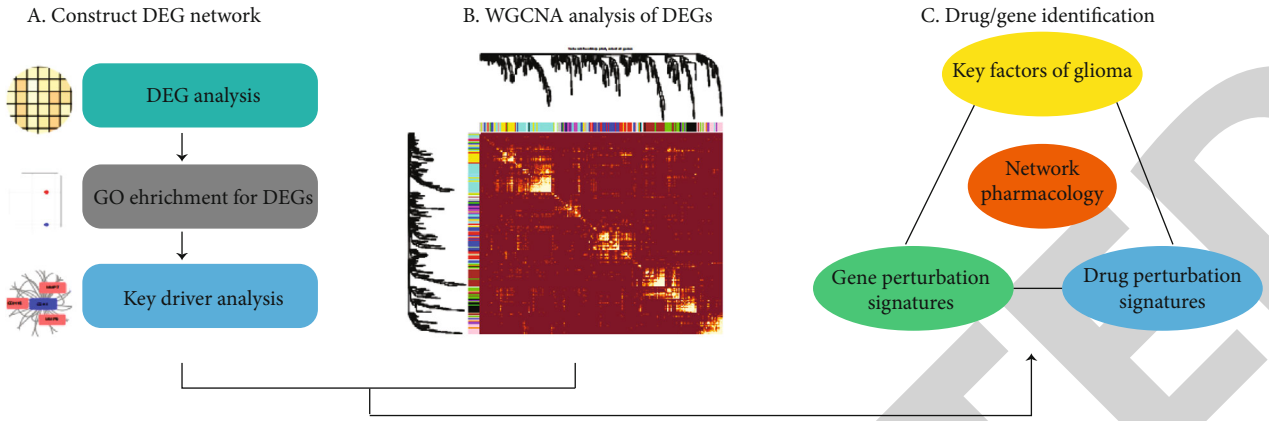


Figure 1: The overall pipeline of potential drug prediction based on drug perturbation-induced gene expression signatures: (a) DEG discovery and key driver analysis; (b) finding modules in gene expression profile by WGCNA; (c) finding drugs targeting glioblastoma by multiple strategies.

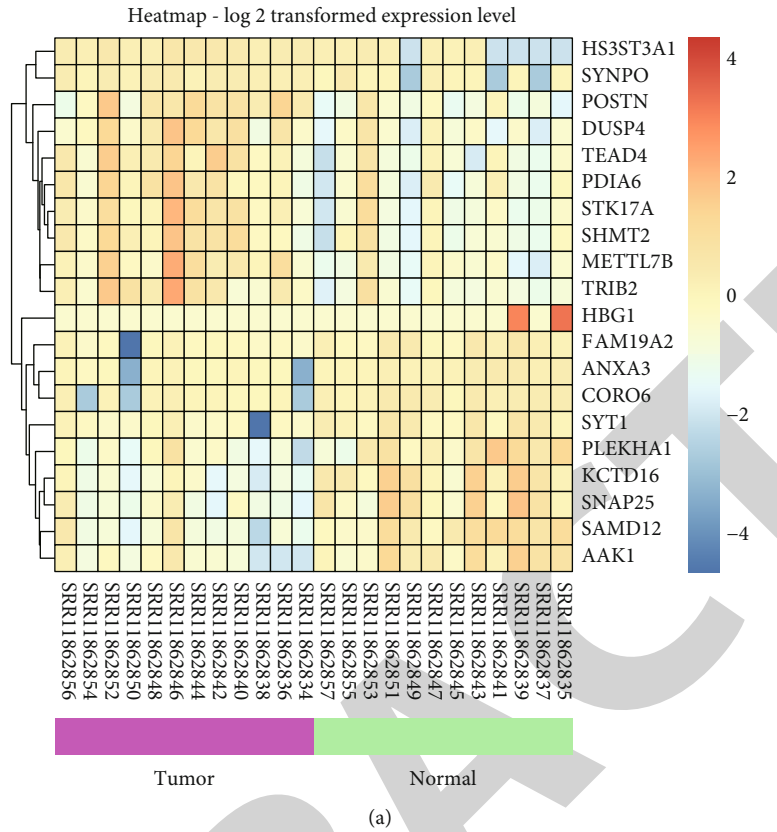
3.5. Several Potential Therapeutic Drugs for Glioblastoma Cancer Were Found. We screened the CREEDS to identify drug perturbations that could reverse the DEGs of glioblastoma because the CREEDS dataset contains thousands of single-drug perturbation-induced gene expression signatures collected from GEO. Those gene set-drug pairs with significant p value might be valuable to our research. 24 drugs were found when we put upregulated genes and key drivers of upregulated genes into CREEDS with p value $< 10^{-10}$ as the threshold. 330 drugs were obtained when downregulated genes and key drivers of downregulated genes were put into CREEDS with the same screening condition. We got 4 drugs that could reverse the expression of downregulated gene and 6 drugs that could reverse the expression of upregulated gene by reviewing the drug profile and previous studies (Table 2). These drugs may be used to treat glioblastoma clinically in the future, such as PD-0332991, which is a selective CDK4/6 inhibitor, and presented outstanding results in phase II clinical trials of estrogen receptor- (ER-) positive HER2-negative breast cancer. It also restricted cell proliferation in preclinical models of hepatocellular carcinoma via promoting a reversible cell cycle arrest [24]. More details of potential drugs were discussed in Discussion.

3.6. Independent Dataset Showed Similar Results. The data of GBM from TCGA was used as an independent validation. We obtained 5501 significantly DEGs with 2913 upregulated genes (52.95%) and 2588 (47.05%) downregulated genes. The top 10 pathways through GO enrichment analysis of different expression genes showed similar results on downregulated genes, while there are some major differences between the pathways from upregulated gene enrichment analysis of the two datasets (Figure 5). Based on the different expression genes of TCGA data, 242 and 32 key drivers associated with upregulation and downregulation genes were discovered, separately.

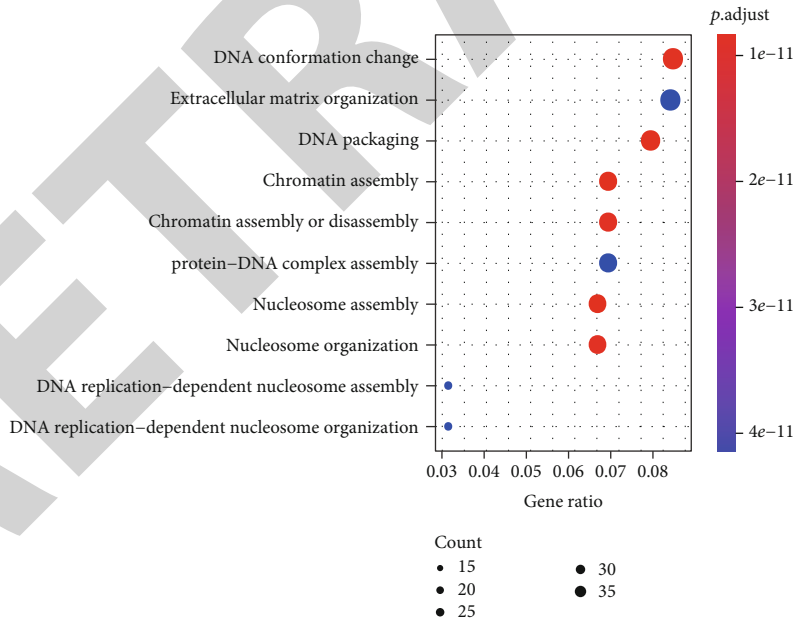
4. Discussion

GBM is a highly aggressive malignant brain tumor with a poor prognosis, and effective drugs targeted to GBM are relatively rare. New drug discovery is a long period-consuming and high-cost thing. Therefore, it is quite meaningful to find a new way to discover potential drugs to GBM. In this work, we have built a drug discovery process of glioblastoma based on differentially expressed gene analysis and predicted 10 potential drugs for glioblastoma therapy. RNA sequencing data of tumor samples and adjacent normal samples from patients with glioblastoma were used as input files and potential drug dataset as reference files. Furthermore, we used an independent verification set to prove that this method was feasible.

During the process of analysis, we explored the related signal pathway changes of glioblastoma by enrichment analysis. The upregulated genes mainly play roles in DNA conformation change, DNA packaging, chromatin assembly or disassembly pathways, etc. It is well known that the methylation of DNA can change the conformation of DNA molecules. DNA methyltransferases (DNMT1 and DNMT3b) could regulate and maintain the methylation of promoter and are overexpressed in human cancer. Additionally, it has been reported that DNA methyltransferase mediated transcriptional silencing in malignant glioblastoma [25]. It is suggested that defects of the chromatin architecture underlie GBM pathogenesis [26]. In addition to obtaining the relevant signaling pathways, we also obtained some key genes, such as CDH1, through key driver analyzing. CDH1 is a tumor suppressor gene and a tumor metastasis suppressor gene. It encodes a classical cadherin of the cadherin superfamily and mediates the adhesion between epithelial cells. It is closely related to the occurrence, development, and metastasis of malignant tumors from various epithelial sources. Loss of function of this gene is believed to contribute to the proliferation and metastasis of tumor cells. Mutations in this gene are correlated with gastric, breast, colorectal, and ovarian cancer. CAMK2G could support cancers through activating transcription factors such as AKT1, CREB, and CDK1/2 [27, 28]. Chai et al. reported CAMK2G was related to lung tumor by affecting stem-like traits and suggested these were mediated through NF-κB activation. GO annotations related to this gene include protein homodimerization activity and protein kinase activity [29].



(a)



(b)

Figure 2: Continued.

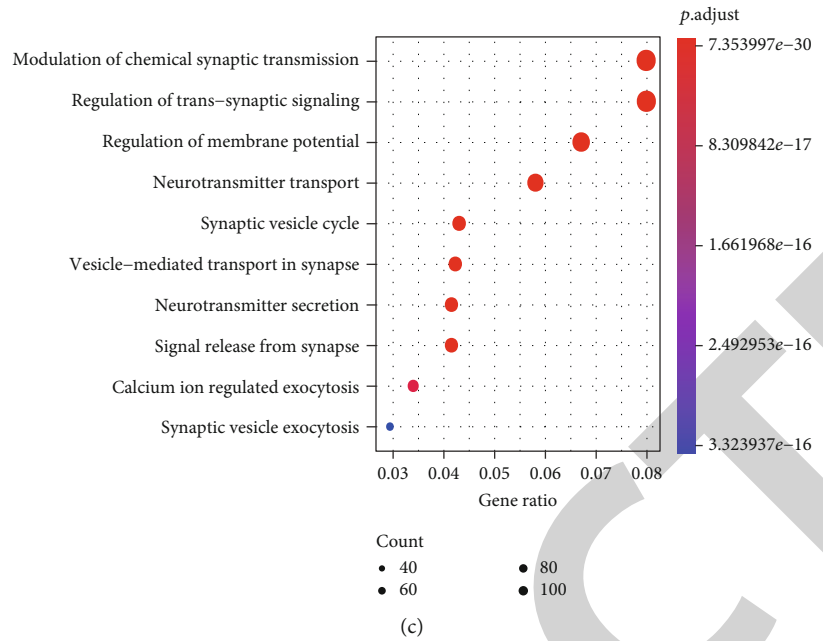


Figure 2: (a) TOP 10 differentially expressed genes by DESeq2. HS3ST3A1, SYNPO, POSTN, DUSP4, TEAD4, PDIA6, STK17A, SHMT2, METTL7B, and TRIB2 were the top 10 upregulated genes. HBG1, FAM19A2, ANXA3, CORO6, SYT1, PLEKHA1, KCTD16, SNAP25, SAMD12, and AAK1 were the top 10 downregulated genes. (b) The top ten pathways by GO enrichment analysis for upregulated genes and (c) downregulated genes between glioblastoma samples and normal samples. The x-axis is the ratio of enriched differential expression genes in the corresponding pathway, and the y-axis is the name of the pathway.

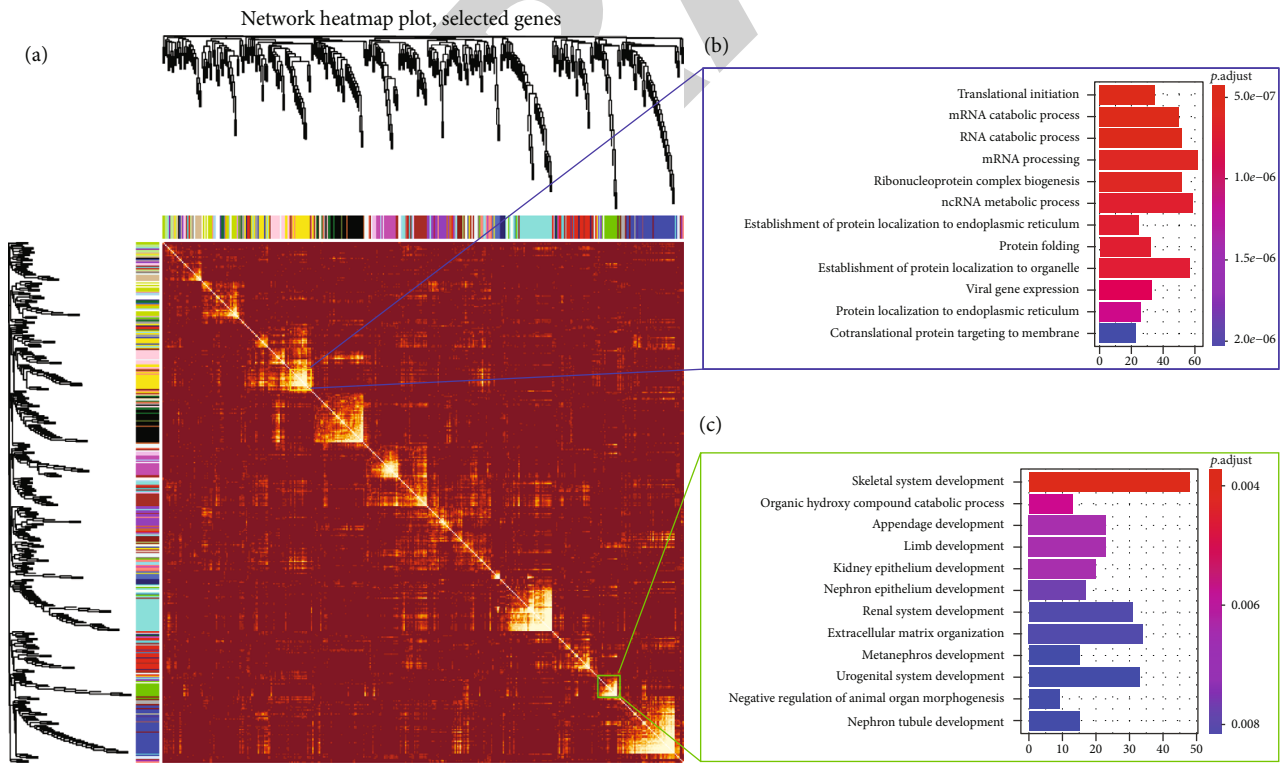
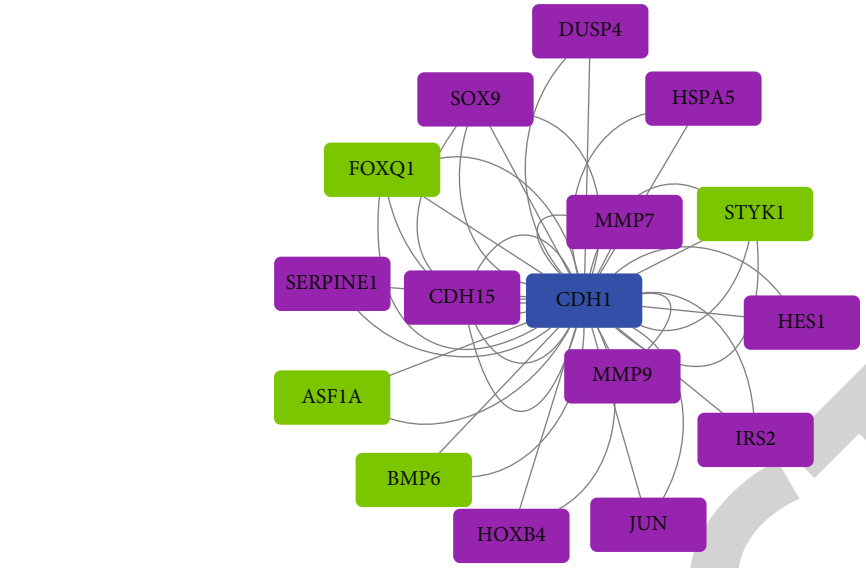
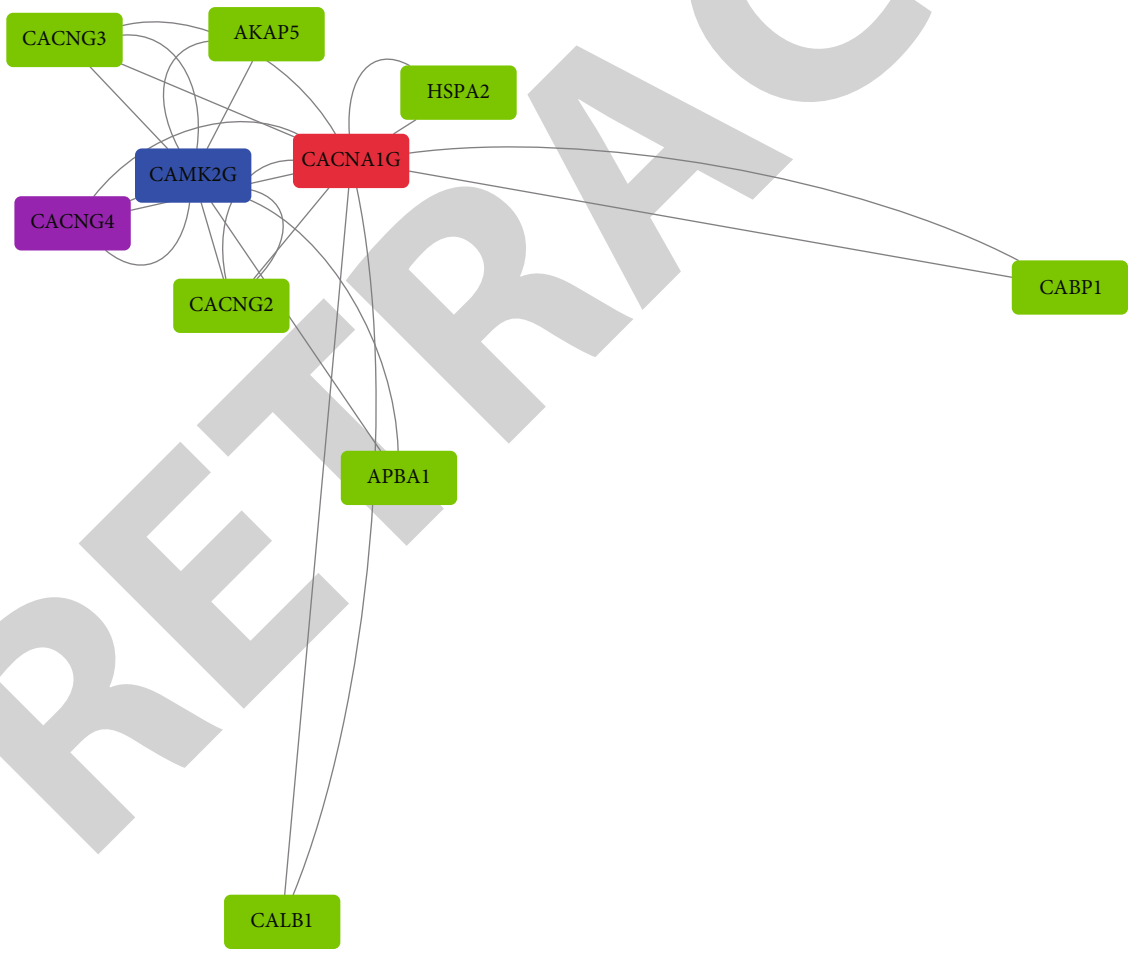


Figure 3: (a) Topological overlapping heat map by WGCNA. Since there are too many genes to visualize, only 400 randomly selected genes are shown in the topological overlapping heat map. Each module was enriched significantly in specific pathways. GO analysis of module 3 and module 11 are shown in (b) and (c).

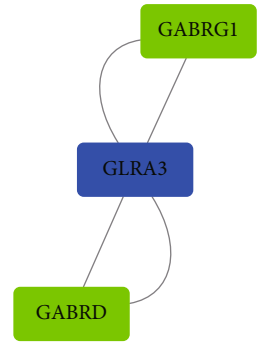


(a)



(b)

Figure 4: Continued.



(c)

Figure 4: A subnetwork concerning up-/downregulated genes and key drivers. Purple blocks present upregulated genes, red color accounts for key drivers of upregulated genes, green showed the downregulated genes, and blue illustrated key drivers of downregulated genes.

Table 2: List of potential therapeutic drugs for glioblastoma cancer.

Type	Drug/small molecule	Possible effect	Evidence (DOI)
Up	Curcumin	Curcumin-induced cell death is mediated both by the activation of cell death pathways and by the inhibition of growth/proliferation pathways	10.1208/s12248-009-9128-x
Up	APC/C inhibitors	Target cell cycle proteins for proteasomal-mediated degradation	10.1158/1541-7786.MCR-18-1361
Up	PDE9A inhibitor	Elevates central cGMP levels	10.1158/1535-7163 10.1111/j.1365-2184.2012.00819.x
Up	W-13	A calmodulin antagonist that inhibits cell growth and induces cell apoptosis	10.3892/or.2019.7022
Down	V-4084	Selectively inhibits MET kinase, affects cell cycle in both tumor and host	10.1186/s12967-015-0667-x 10.1093/noajnl/vdaa067
Down	Kaempferol	Well-known flavonoid, remarkable bioactivity against various malignant tumors	10.1371/journal.pone.0155264 10.3892/etm.2019.7886 10.3892/or.2014.3662
Down	PD-0332991	Selective CDK4/6 inhibitor, outstanding results in phase II clinical trials of estrogen receptor- (ER-) positive HER2-negative breast cancer	10.1136/gutjnl-2016-312268
Down	Dexamethasone	Induces tumor lysis syndrome	10.3747/co.21.1769 10.1097/00000542-200609000-00042
Down	All-trans-retinoic acid	Blocks the cell cycle, enhances apoptosis, and decreases gastric CSC properties	10.3390/ijms19113388
Down	Erlotinib	Blocks tumor cell growth by targeting EGFR	10.1177/1758834011427927 10.1056/NEJMoa050753

In the potential therapeutic drugs for glioblastoma cancer, 10 drugs were selected after gene-drug screening. Curcumin, a component of turmeric (Curcuma longa), had been proved as safe, affordable, and efficacious drug comparing with chemotherapeutic agents. Curcumin is highly fat dissolving and could induce cell death by activating cell death pathways as well as inducing inhibition of growth/proliferation process [30]. Anaphase-Promoting Complex (APC/C) inhibitors, such as proTAME and apcin, targeted cell cycle proteins for proteasomal-mediated degradation. The targets of APC/C are regulated throughout the cell cycle via two mutually exclusive activator proteins, CDH1 and CDC20 [31]. W-13 is a calmodulin antagonist that could inhibit cell growth and induce cell apoptosis, which may display antitumor effects by binding to CAMK2G protein [32]. DE9A inhibitor plays an important role in cell proliferation, differentiation, and apoptosis via the cGMP signaling pathway [33, 34]. BAY73-6691, a PDE9A inhibitor, can suppress breast cancer cell population growth and induce apoptosis [35]. Kaempferol is a well-known flavonoid, with remarkable bioactivity against various malignant tumors, such as non-small-cell lung cancer (NSCLC), breast cancer, hepatocellular carcinoma (HCC), ovarian cancer (OC), and gastric cancer (GC). However, the detailed mechanisms of kaempferol against numerous cancer types have remained elusive [36]. All-trans-retinoic acid could block the cell cycle, enhance

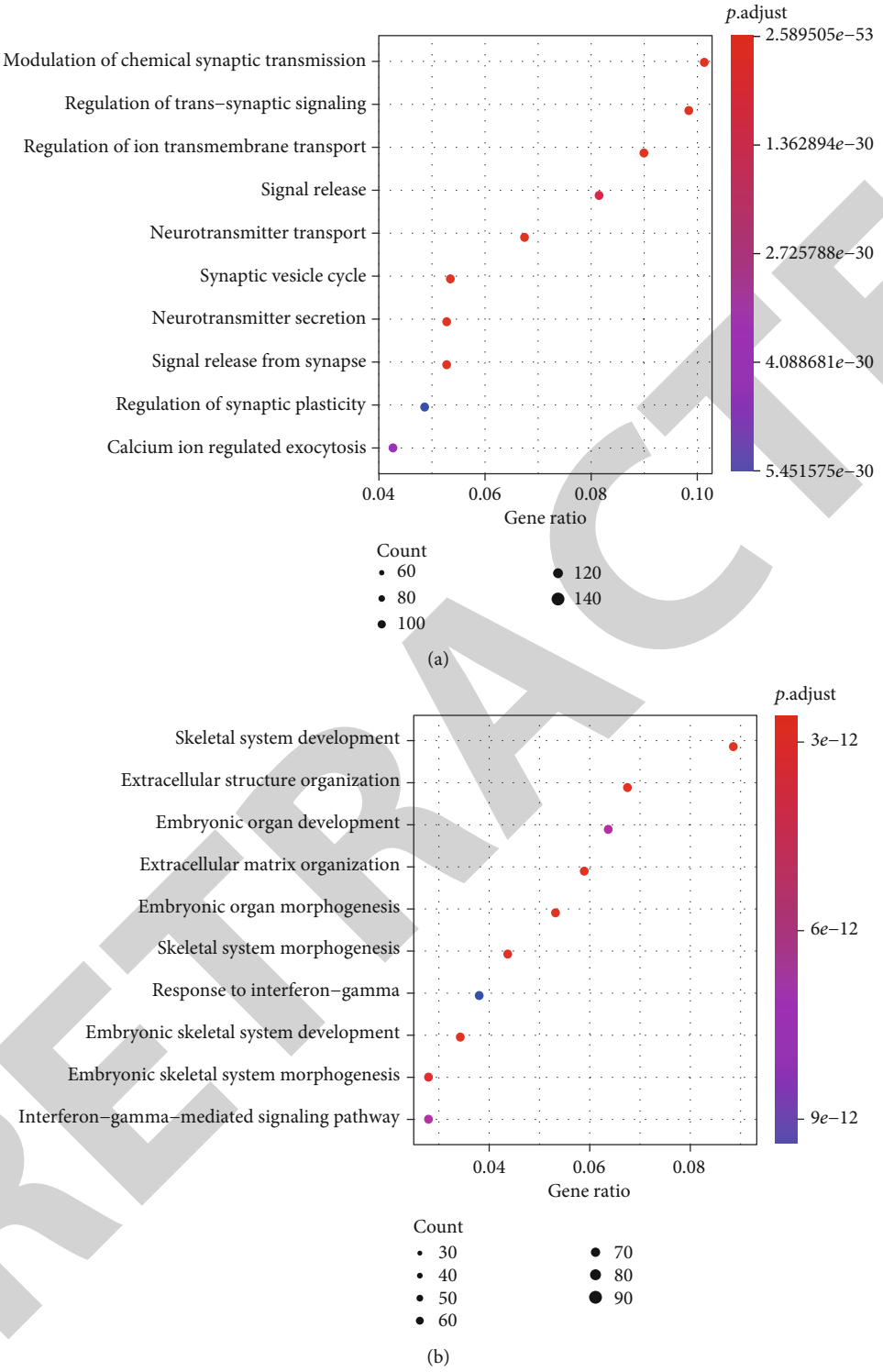


Figure 5: Differentially expressed genes by DESeq2 and enrichment analysis results of TCGA data. (a) The top ten pathways by GO enrichment analysis for downregulated genes and (c) upregulated genes between glioblastoma samples and normal samples. The x-axis is the ratio of enriched differential expression genes in the corresponding pathway, and the y-axis is the name of the pathway.

apoptosis, and decrease gastric cancer stem cell (CSC) properties [37]. Drugs with high fat solubility, small molecular weight, and simple chemical structure were believed to penetrate the blood-brain barrier easily. However, the potential therapeutic drugs against glioma with the ability to penetrate the blood-brain barrier are still unknown.

There were still several limitations in our study. The datasets were from public databases, the number of samples was

relatively small, and the accuracy of the results needed to be further verified. Furthermore, how many effects of a potential drug in the regulatory network still needs to be studied and verified using experimental data. We will continue to track the progress of glioblastoma research and verify the accuracy of the results.

5. Conclusions

In this study, we provided a framework of workflow for potential therapeutic drug discovery through a series of analysis processes and predicted 10 potential drugs for glioblastoma therapy. Whether these drugs are effective in patients with glioblastoma deserves further study.

Data Availability

The data of glioma in the Gene Expression Omnibus (GEO) dataset was collected from the project ID of GEO-GSE151352. The DEGs and key driver dataset were gained from http://amp.pharm.mssm.edu/creeds.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any potential conflict of interest.

Authors' Contributions

Yuhong Man designed the study. Bochi Zhu and Xijing Mao collected data, analyzed data, interpreted data, and wrote the manuscript. Yuhong Man reviewed the manuscript.

Supplementary Materials

Supplementary Table 1: the log2FC and adjusted p value of 1771 DEGs between GBM and normal samples. *(Supplementary Materials)*

References

[1] M. Ashburner, C. A. Ball, J. A. Blake et al., "Gene ontology: tool for the unification of biology," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.

[2] S. Deorah, C. F. Lynch, Z. A. Sibenaller, and T. C. Ryken, "Trends in brain cancer incidence and survival in the United States: Surveillance, Epidemiology, and End Results Program, 1973 to 2001," *Neurosurgical Focus*, vol. 20, no. 4, p. E1, 2006.

[3] M. Westphal and K. Lamszus, "The neurobiology of gliomas: from cell biology to the development of therapeutic approaches," *Nature Reviews. Neuroscience*, vol. 12, no. 9, pp. 495–508, 2011.

[4] K. C. Cotto, A. H. Wagner, Y.-Y. Feng et al., "DGIdb 3.0: a redesign and expansion of the drug-gene interaction database.

[5] T. Sorlie, C. M. Perou, R. Tibshirani et al., "Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 19, pp. 10869–10874, 2001.

[6] P. J. Valk, R. G. Verhaak, M. A. Beijen, and C. A. Erpelinck, "Prognostically useful gene-expression profiles in acute myeloid leukemia," *The New England Journal of Medicine*, vol. 350, no. 16, pp. 1617–1628, 2004.

[7] Y. Sun, W. Zhang, D. Chen et al., "A glioma classification scheme based on coexpression modules of EGFR and PDGFRA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 9, pp. 3538–3543, 2014.

[8] J. R. Chen, Y. Yao, H. Z. Xu, and Z. Y. Qin, "Isocitrate dehydrogenase (IDH)1/2 mutations as prognostic markers in patients with glioblastomas," *Medicine*, vol. 95, no. 9, article e2583, 2016.

[9] M. E. Hegi, L. Liu, J. G. Herman et al., "Correlation of O6-methylguanine methyltransferase (MGMT) promoter methylation with clinical outcomes in glioblastoma and clinical strategies to modulate MGMT activity," *Journal of Clinical Oncology*, vol. 26, no. 25, pp. 4189–4199, 2008.

[10] R. Stupp, W. P. Mason, M. J. van den Bent et al., "Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma," *The New England Journal of Medicine*, vol. 352, no. 10, pp. 987–996, 2005.

[11] Y. Shi, X. Mo, S. Hong, T. Li, B. Chen, and G. Chen, "Studying the role and molecular mechanisms of MAP4K3 in sorafenib resistance of hepatocellular carcinoma," *BioMed Research International*, vol. 2020, Article ID 4965670, 8 pages, 2020.

[12] G. W. Robinson, B. A. Orr, and A. Gajjar, "Complete clinical regression of a BRAF V600E-mutant pediatric glioblastoma multiforme after BRAF inhibitor therapy," *BMC Cancer*, vol. 14, no. 1, p. 258, 2014.

[13] S. Pushpakom, F. Iorio, P. A. Eyers et al., "Drug repurposing: progress, challenges and recommendations," *Nature Reviews. Drug Discovery*, vol. 18, no. 1, pp. 41–58, 2019.

[14] J. M. Pulley, J. P. Rhoads, R. N. Jerome et al., "Using what we already have: uncovering new drug repurposing strategies in existing omics data," *Annual Review of Pharmacology and Toxicology*, vol. 60, no. 1, pp. 333–352, 2020.

[15] T. Li, X. Xu, J. Li et al., "Association of ACP1 gene polymorphisms and coronary artery disease in northeast Chinese population," *Journal of Genetics*, vol. 94, no. 1, pp. 125–128, 2015.

[16] Q. Liu, R. Bonneville, T. Li, and V. X. Jin, "Transcription factor-associated combinatorial epigenetic pattern reveals higher transcriptional activity of TCF7L2-regulated intragenic enhancers," *BMC Genomics*, vol. 18, no. 1, p. 375, 2017.

[17] M. L. Goodman, G. M. Trinca, K. R. Walter et al., "Progesterone receptor attenuates STAT1-mediated IFN signaling in breast cancer," *Journal of Immunology*, vol. 202, no. 10, pp. 3076–3086, 2019.

[18] Q. Duan, C. Flynn, M. Niepel et al., "LINCS canvas browser: interactive web app to query, browse and interrogate LINCS L1000 gene expression signatures," *Nucleic Acids Research*, vol. 42, no. W1, pp. W449–W460, 2014.

[19] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2," *Genome Biology*, vol. 15, no. 12, 2014.

[20] The Gene Ontology Consortium, "The Gene Ontology resource: 20 years and still GOing strong," *Nucleic Acids Research*, vol. 47, pp. D330–D3D8, 2019.

[21] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *Omics*, vol. 16, no. 5, pp. 284–287, 2012.

[22] P. Langfelder and S. Horvath, "WGCNA: an R package for weighted correlation network analysis," *BMC Bioinformatics*, vol. 9, no. 1, p. 559, 2008.

Hindawi
BioMed Research International
Volume 2024, Article ID 9841854, 1 page
https://doi.org/10.1155/2024/9841854

Retraction

Retracted: Bioinformatics Analyses Reveals a Comprehensive Landscape of CXC Chemokine Family Functions in Non-Small Cell Lung Cancer

BioMed Research International

Received 8 January 2024; Accepted 8 January 2024; Published 9 January 2024

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Tian, L. Wang, Y. Liu et al., "Bioinformatics Analyses Reveals a Comprehensive Landscape of CXC Chemokine Family Functions in Non-Small Cell Lung Cancer," *BioMed Research International*, vol. 2021, Article ID 6686158, 34 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6686158, 34 pages
https://doi.org/10.1155/2021/6686158

Research Article

Bioinformatics Analyses Reveals a Comprehensive Landscape of CXC Chemokine Family Functions in Non-Small Cell Lung Cancer

He Tian,[1] Liyu Wang,[1] Yu Liu,[1] Yalong Wang,[2] Yujia Zheng,[1] Tao Fan,[1] Bo Zheng,[3] Fengwei Tan,[1] Qi Xue,[1] Shugeng Gao,[1] Chunxiang Li ,[1] and Jie He [1]

[1]*Department of Thoracic Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China*
[2]*Department of Anesthesiology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China*
[3]*Department of Pathology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China*

Correspondence should be addressed to Chunxiang Li; lichunxiang@cicams.ac.cn and Jie He; prof.jiehe@gmail.com

Received 20 October 2020; Revised 8 December 2020; Accepted 2 January 2021; Published 27 January 2021

Academic Editor: Jialiang Yang

Backgrounds. Lung cancer is a major source of tumor-related death each year with non-small cell lung cancer (NSCLC) being a prevalent subtype. The metastasis from NSCLC to the brain usually imposes many neuron disorders. Previous studies have suggested that communications among cancer cells and interstitial cells are essential in tumorigenesis and are influenced by chemokines. In the tumor microenvironment, CXC chemokines can participate in the shifting of immune cells and manage tumor cell condition, thus affecting the progression of cancer and patient destinies. However, the expression and values of CXC chemokine family in NSCLC have not been systematically illustrated using public databases. *Methods*. UALCAN, STRING, ONCOMINE, GeneMANIA, cBioPortal, GEPIA, TISIDB, TRRUST, TIMER, Kaplan-Meier Plotter, and R software were utilized in this study. *Results*. Based on the TIMER and UACLCAN databases, in LUAD patients, the expression levels of CXCL10, CXCL13, and CXCL14 were significantly elevated while the transcriptional levels of CXCL2/3/4/7/12/16 were significantly reduced; in LUSC patients, the expression levels of CXCL6/10/13/14 were significantly elevated while the expression levels of CXCL2/3/4/5/7/11/12/16/17 were significantly reduced. We found remarkable relevance between the pathological stages of LUAD patients and the expressions of CXCL8 (positive) and CXCL17 (negative). Similarly, there are significant correlations between the pathological stages of LUSC patients and the expressions of CXCL1/2/6/17. In LUAD, patients with low expression levels of CXCL1/4/7/8 and patients with high expression levels of CXCL12/14/16 were associated with a significantly better prognosis. But in LUSC, all correlations between chemokines and prognosis are statistically insignificant. Pairwise expression correlation analysis among CXC chemokines shows that there are 7 significant correlations (between CXCL1 and CXCL2, between CXCL1 and CXCL3, between CXCL1 and CXCL8, between CXCL2 and CXCL3, between CXCL4 and CXCL7, between CXCL9 and CXCL10, and between CXCL9 and CXCL11) in LUAD and 4 significant correlations (between CXCL1 and CXCL8, between CXCL2 and CXCL3, between CXCL4 and CXCL7, and between CXCL10 and CXCL11) in LUSC. Significant correlations between the expressions of CXC chemokines and the infiltration of six common types of immune cells were also discovered in both LUAD and LUSC. *Conclusions*. We provided a comprehensive landscape of the CXC chemokine family in LUAD and LUSC using the bioinformatics method and found differences between LUSC and LUAD in the field of CXC chemokines. Our study may help validate and identify known novel immunotherapeutic targets and prognostic biomarkers.

1. Introduction

Annually, about 1.8 million people develop lung cancer, and 1.6 million people die as a result of this disease [1]. Approximately 85% of patients are classified as NSCLC (non-small cell lung cancer), of which the main histological types are adenocarcinoma and squamous cell carcinoma. There are enormous genomic and clinical heterogeneities between LUAD and LUSC [2]. Squamous-cell lung cancer, constituting 25%-30% of NSCLC cases, is usually centrally located and more likely to invade large blood vessels and has rare mutations/alterations which can be targeted from afar, which lead to many challenges in the treatment [3, 4]. Adenocarcinoma of the lung comprises around 40% of all lung cancer and tends to occur in the periphery of the lung. It is easier to identify before it has spread outside of the lungs, so it has better prognosis than LUSC [5]. Many researchers have focused on the therapeutic targets of NSCLC, especially immune checkpoint suppressor and oncogene mutations, and many advancements have been made [6]; however, it is far from sufficient; more therapeutic targets and prognostic biomarkers need to be identified.

Endogenous chemokine ligands and G protein-coupled seven-transmembrane spanning signaling receptors are the main members of the chemokine superfamily; they are chemotactic cytokines that control the shifting and communication cells and tissues [7]. In cancer microenvironment, chemokines are mainly produced by tumor cells and immune cells and play important role in mediating immune cell trafficking and lymphoid tissue development; thus, they are involved in antitumor immunological responses [8]. It has been convinced that chemokines can affect many cancer-related biological processes including tumor angiogenesis, tumorigenesis, progression, and metastasis and even influence patients' clinical outcomes [9–11].

CXC chemokines participate in many biological progresses of malignant disease in several different organs, including the lung [12], breast [13], colorectal [14], and kidney [15]. In lung cancer, CXC chemokines are mainly involved in angiogenesis, immunoangiostasis, and metastases [16, 17]. CXCL1 [18], CXCL5 [19], CXCL8 [20], and CXCL12 [21] are CXC chemokines that have attracted much attention, since they are considered to facilitate lung cancer initiation, development, and metastasis by different mechanisms; however, for some CXC chemokines like CXCL4 [22, 23], CXCL14 [24, 25], and CXCL16 [26, 27], researchers have opposite opinions towards their roles in tumor development.

Generally, previous studies have sporadically characterized the function of some CXC chemokines in NSCLC, but our study is the first to give a comprehensive landscape of CXC chemokines in LUAD and LUSC using public database and bioinformatics skills. Technology development of second-generation gene sequencing and springing up of various databases will accelerate the macrolevel research of the CXC chemokine family in tumors.

Our study conducted a comprehensive and profound bioinformatics analysis of the expression of CXC chemokines in LUAD and LUSC and explored their roles as therapeutic targets and prognostic biomarkers on several credible public databases and R packages, thus offering more evidence to facilitate the clinical routine in selecting new drugs and investigating NSCLC patients' long-term outcome more accurately.

2. Materials and Methods

2.1. ONCOMINE. ONCOMINE (http://www.oncomine.org) is a translational bioinformatics service that provides powerful, genome-wide expression analysis. Datasets in ONCOMINE were collected from public repositories such as Gene Expression Omnibus (GEO) and Array Express by Compendia Bioscience (a prominent cancer bioinformatics company widely used by the pharmaceutical industry to identify novel gene targets for drug discovery and development), and the datasets are composed of microarray data of primary tumors, cell lines, or xenografts [28]. Data were extracted to evaluate the expression of CXC chemokines in lung cancer. In our study, a p value of 0.05, a fold change of 2, and a gene rank in the top 10% were set as the significance thresholds. Student's t test was used to analyze the difference in the expression of CXC chemokines in lung cancer.

2.2. GEPIA. GEPIA (http://gepia.cancer-pku.cn/index.html) is an analysis tool containing RNA sequence expression data of 9736 tumors and 8587 normal tissue samples, which was developed at Peking University. Data in GEPIA are extracted from TCGA and GETx projects using a standard pipeline, and the RNA-Seq datasets GEPIA used are based on the UCSC Xena project (http://xena.ucsc.edu) [29]. In this study, we performed pathological stage analysis with the "Single Gene Analysis" module and correlation analysis with the "Correlation Analysis" module. Multiple gene comparison analysis of CXC chemokines was performed with the "Multiple Gene Comparison" module, using the "LUAD" and "LUSC" datasets. 3D principle component analysis and survival map were performed with the "Dimensionality Reduction" module and "Survival Analysis" module of GEPIA2 (test), respectively. Student's t test was used to generate a p value for expression or pathological stage analysis, and we take 0.05 as the cutoff of the p value.

2.3. UALCAN. UALCAN (http://ualcan.path.uab.edu/) is a comprehensive web resource, providing analyses based on The Cancer Genome Atlas (TCGA) and MET500 cohort data [30]. In our study, expression data for CXC chemokines was obtained using the "Expression Analysis" module and "Methylation Analysis" module of UALCAN and the "LUAD" and "LUSC" datasets. Student's t test was used to generate a p value. We set 0.05 as the p value cutoff.

2.4. cBioPortal. cBioPortal (http://www.cbioportal.org) is a comprehensive web resource, which can visualize and analyze multidimensional cancer genomics data; it stores nonsynonymous mutations, DNA copy-number data, mRNA and microRNA expression data, protein-level and phosphoprotein-level data, DNA methylation data, and deidentified clinical data. The data in cBioPortal was based on TCGA database and a large number of published articles

[31]. Genetic change, coexpression, and the network module of CXC chemokines based on TCGA database were obtained from cBioPortal. A total of 556 LUAD samples (TCGA, Pan-Cancer Atlas) and 487 LUSC (TCGA, Pan-Cancer Atlas) were included. mRNA expression z scores (RNA Seq V2 RSEM) were obtained using a z score threshold of ±2.0.

2.5. GeneMANIA. GeneMANIA (http://www.genemania.org) is a website providing information for genetic interactions and proteins, pathways, coexpression, colocalization, and protein domain similarity of genes; it can find other genes that are related to input genes with a very large set of functional association data, and it relies on several data sources including GEO, BioGRID, EMBL-EBI, Pfam, Ensembl, NCBI, MGI, I2D, InParanoid, and Pathway Commons [32].

2.6. STRING. STRING (https://string-db.org/) is aimed at collecting, scoring, and integrating all publicly available sources of protein-protein interaction (PPI) data and at complementing these with computational predictions of potential functions; it relies on many credible resources including COG, Ensembl, Intact, RefSeq, PubMed, Reactome, DIP, BioGRID, MINT, KEGG, SGD, FlyBase, Swiss-Prot/UniProt, SWISS-MODEL, HUGO, OMIM, NCI/NaturePID, PDB, The Interactive Fly, BioCyc, Gene Ontology, and SIMAP [33]. We conducted a PPI network analysis of differentially expressed CXC chemokines to explore the interactions among them with STRING in July 2020.

2.7. TRRUST. TRRUST (https://www.grnpedia.org/trrust/) is a reliable, intuitive tool for transcriptional regulatory networks based on 11237 articles describing small-scale experimental studies of transcriptional regulations in PubMed. Containing 8444 transcription factor- (TF-) target regulatory relationships of 800 human TFs, the TRRUST database can provide information on how these interactions are regulated [34].

2.8. TIMER. TIMER (https://cistrome.shinyapps.io/timer/) is an online tool that evaluated the infiltration of different immune cells and their clinical impact, and its data is based on samples from TCGA [35]. In our study, the correlation between immune cell and CXC chemokine level was evaluated with "Gene module"; the "Diff Exp" module was applied to show the expression of CXC chemokines among different cancer types; somatic copy number alteration of CXC chemokines was performed with "SCNA" module; and "Survival module" was used to evaluate the correlation among clinical outcome and the infiltration of immune cells and CXC chemokine expression.

2.9. TISIDB. TISIDB (http://cis.hku.hk/TISIDB/) is a gene-based resource exploring interactions between tumor and immune contexture. It is based on several credible resources, including PubMed database, high throughput screening data, RNA, and exome sequencing datasets of patient cohorts with immunotherapy, TCGA, UniProt, GO, DrugBank, etc. TISIDB allows users to interrogate the function of a specific gene in tumor-immune interplay through literature mining and the data analysis of genome-wide screening and high-throughput profiling [36]. In our study, we used the "Chemokine" module of TISIDB to make Spearman expression correlations between 16 CXC chemokines in our research and several common chemokines (including the 16 CXC chemokines themselves in our study) among different types of cancer.

2.10. Kaplan-Meier Plotter. The Kaplan-Meier Plotter-Lung Cancer (https://kmplot.com/analysis/index.php?p=service&cancer=lung) is an online tool that can show how genes influence survival in lung cancer ($n = 3{,}452$). EGA, GEO, and TCGA are included as sources in this database. Initially, the main function of this website is to identify survival biomarkers based on meta-analysis [37]. In our study, 719 lung adenocarcinoma samples and 524 lung squamous-cell carcinoma samples were used to draw the OS survival curve of CXC chemokines in LUAD and LUSC, respectively.

2.11. Statistical Analysis. R software with package "clusterProfiler" was used for GO and KEGG analysis and plotting.

3. Results

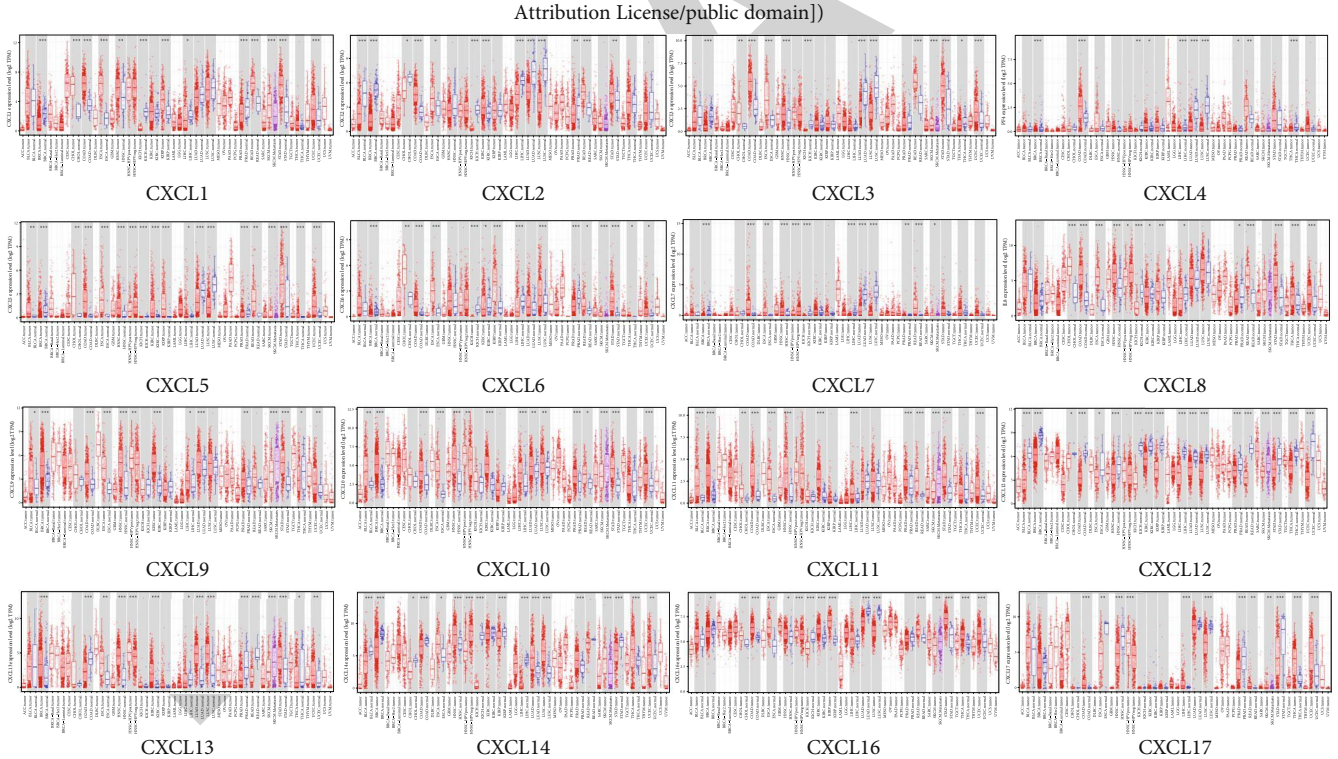
3.1. Aberrant Transcription and Methylation Level of CXC Chemokines in LUAD and LUSC Patients. Sixteen CXC chemokines (not including CXCL15) were retrieved from the ONCOMINE database. We first explored the mRNA expression levels of CXC chemokines in lung cancer and normal lung tissues with ONCOMINE (where LUAD and LUSC are not discriminated). Results are displayed in Figure 1(a). Based on the data from ONCOMINE, the expressions of CXCL14, CXCL13, and CXCL9 in lung cancer tissues were elevated with statistical significance while the transcriptional levels of CXCL1, CXCL2, CXCL3, CXCL4, CXCL5, CXCL7, CXCL12, CXCL16, and CXCL17 were significantly reduced in lung cancer vs. normal renal tissue. We also assessed the expression levels of CXC chemokines in cancer vs. normal tissue with TIMER where LUAD and LUSC are divided in Figure 1(b). In LUAD, CXCL9, CXCL10, CXCL13, and CXCL14 were significantly elevated and CXCL2, CXCL3, CXCL4, CXCL5, CXCL7, CXCL12, and CXCL16 were significantly reduced, while in LUSC, CXCL6, CXCL10, CXCL13, and CXCL14 were significantly elevated and CXCL2, CXCL3, CXCL4, CXCL5, CXCL7, CXCL12, CXCL16, and CXCL17 were significantly reduced. For triple verification, we visited the UALCAN database, using its "Expression analysis" module to explore the CXC chemokine transcriptional level in cancer vs. normal tissue in Figure 1(c), and the results are as follows: the transcriptional levels of CXCL2, CXCL3, CXCL4/PF4, CXCL7/PPBP, CXCL11, CXCL12, CXCL16, and CXCL17 were significantly reduced in both LUAD and LUSC tissues; the transcriptional levels of CXCL6, CXCL9, CXCL10, CXCL13, and CXCL14 were significantly elevated in both LUAD and LUSC tissues. For CXCL5, it was significantly reduced in LUSC tissues and reduced in LUAD tissues without statistical significance. From the TIMER and UALCAN databases, CXCL9/10/13/14 were elevated and

Analysis type by cancer	Cancer vs. normal CXCL1		Cancer vs. normal CXCL2		Cancer vs. normal CXCL3		Cancer vs. normal CXCL4		Cancer vs. normal CXCL5		Cancer vs. normal CXCL6		Cancer vs. normal CXCL7		Cancer vs. normal CXCL8		Cancer vs. normal CXCL9		Cancer vs. normal CXCL10		Cancer vs. normal CXCL11		Cancer vs. normal CXCL12		Cancer vs. normal CXCL13		Cancer vs. normal CXCL14		Cancer vs. normal CXCL16		Cancer vs. normal CXCL17	
Bladder cancer			1	2		1					1	2			1		2		2					5	1							
Brain and CNS cancer	5				1			1	1		1	1			4	1	1		2		1	1		4	1		1	2	8			
Breast cancer		15		27	1	9		4		7	1	5		4			15		16		12		2	16	2		1	8	2	2	1	
Cervical cancer	3				1				1		1				4		2		2		2			1	2			3	1			
Colorectal cancer	21		14	1	23		4		10		7		2		19		1	1	4		12			8		5		2	5		3	
Esophageal cancer	7		4		4		2				3		1		5		1		1		1		1					1	1			2
Gastric cancer	4								1	1					4		1				1			4				7	1			8
Head and neck cancer	10	1	3		3				1		2				5		6		7		6			3	4		2	1				2
Kidney cancer	1	3				2		2	2	2	1			7			4		4		3			7	1	2		1	3			
Leukemia		9	1	3	1	3	1	10	1				2	8	1	3	1					1	1	3	1			2		1		
Liver cancer	1	1		6					1		4			1	2		4	1	5		3		1	4				5				
Lung cancer		2		12		6		10		2				11			2							3	7		4		4		1	
Lymphoma			3	2				1			1					5	21		16		7		9		13	1	8		3			
Melanoma	1								1	1					1		1		1					1				2				
Myeloma	1			1		1		1		1						1								2								
Other cancer	3		2	2	1	1			3		1		2	1	6		6	1	11		1			2	3		2	1	3	2	1	
Ovarian cancer	1					1									1				1		1			1	2		2					
Pancreatic cancer			2		2				5		1		1		3		1		1					3	1		1		2		1	
Prostate cancer		1		1								1		1	1		1				1		2	1		2	4					
Sarcoma		1		7					1	3		1		2			2		4		2		2	2		2		8				1
Significant unique analyses	58	33	30	64	37	24	7	29	8	16	24	10	8	34	57	10	72	3	76		53	2	18	70	38	12	25	42	29	9	6	14
Total unique analyses	447		440		407		433		392		409		440		464		452		435		392		458		413		385		296		253	

1 5 10 10 5 1

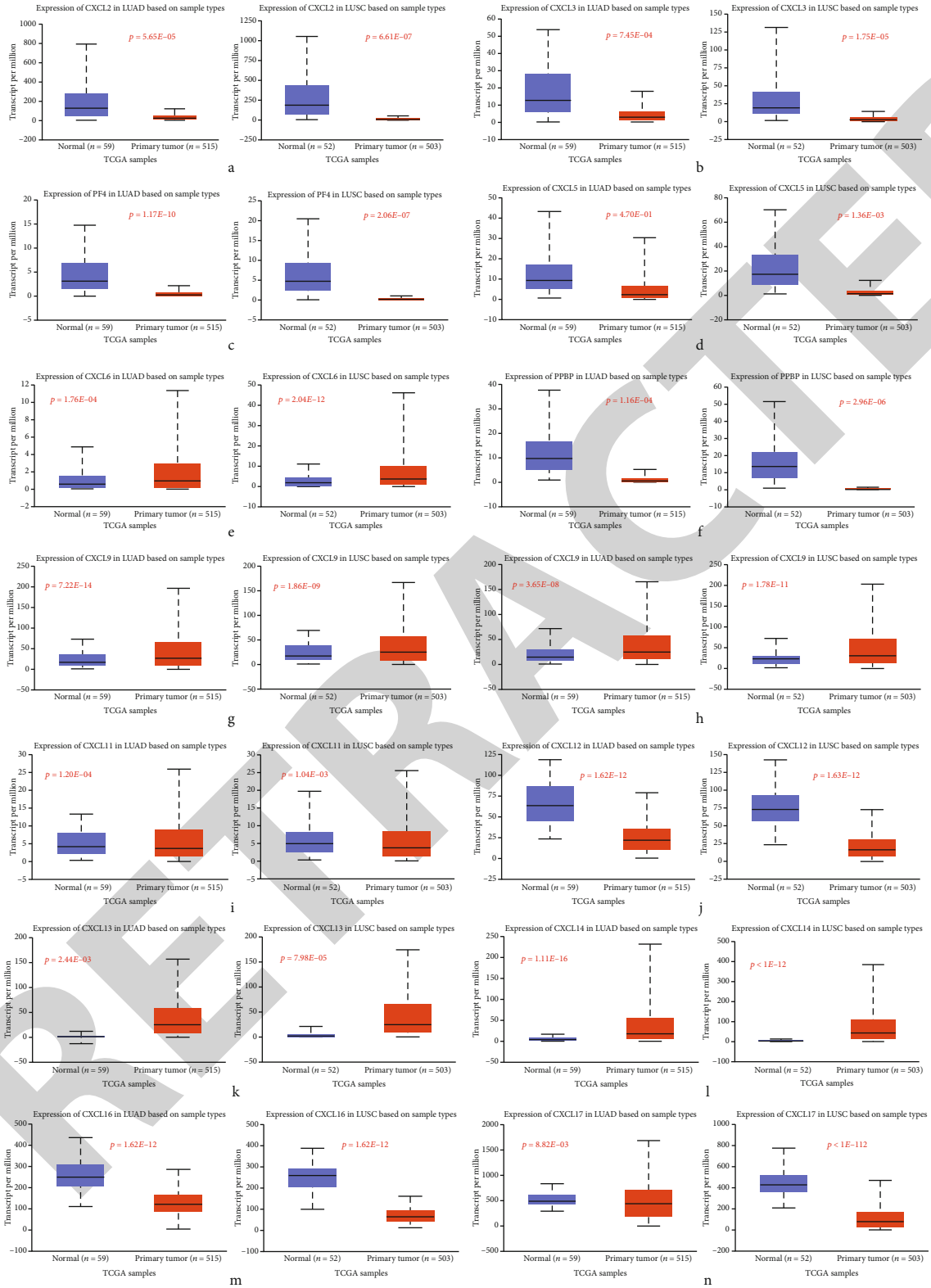
%

(a) mRNA levels of CXC chemokines in lung cancer (ONCOMINE). The figure shows the numbers of datasets with statistically significant mRNA overexpression (red) or downregulated expression (blue) of CXC chemokines (reproduced from Zeng et al. 2020 [under the Creative Commons Attribution License/public domain])



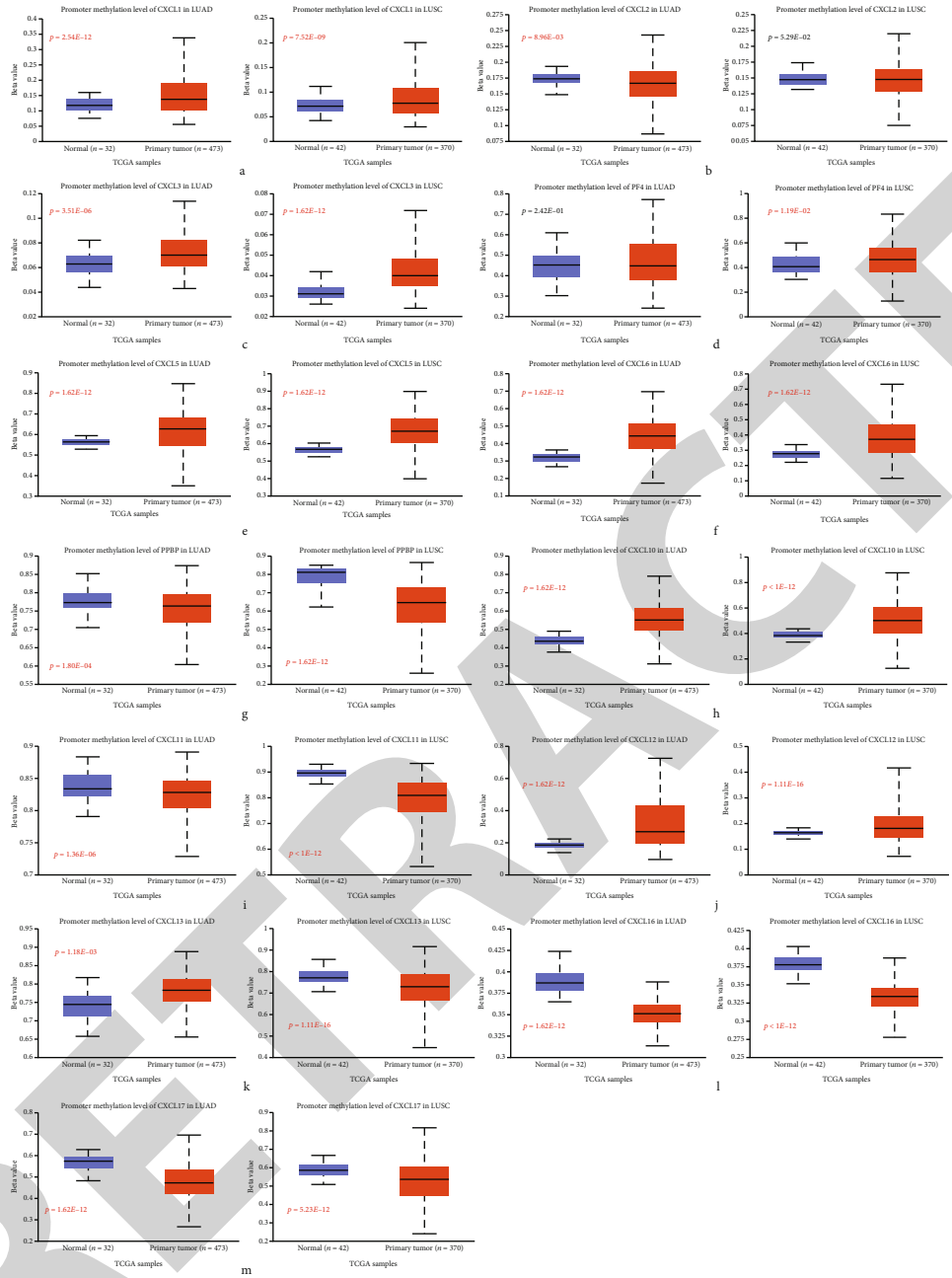
(b) Differential expression between tumor and adjacent normal tissues for CXC chemokines across all TCGA tumors types (TIMER); statistical significance was evaluated using the Wilcoxon test. Columns are gray when normal data are available

Figure 1: Continued.

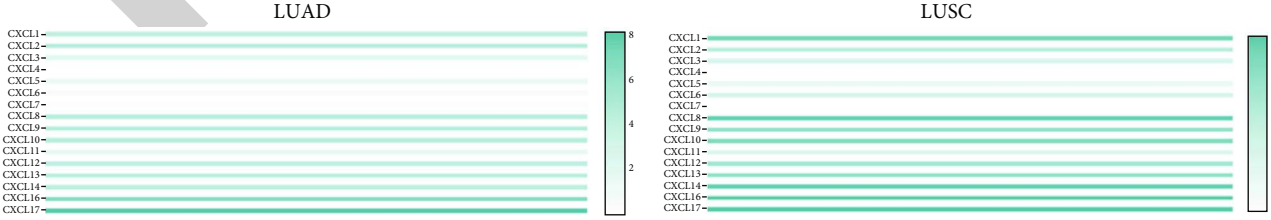


(c) The transcription of CXC chemokines in LUAD and LUSC (UALCAN): (A) CXCL2, (B) CXCL3, (C) CXCL4, (D) CXCL5, (E) CXCL6, (F) CXCL7, (G) CXCL9, (H) CXCL10, (I) CXCL11, (J) CXCL12, (K) CXCL13, (L) CXCL14, (M) CXCL16, and (N) CXCL17

Figure 1: Continued.



(d) The methylation level of CXC chemokines in LUAD and LUSC (UALCAN). (A) CXCL1 (B) CXCL2 (C) CXCL3 (D) CXCL4 (E) CXCL5 (F) CXCL6 (G) CXCL7 (H) CXCL10 (I) CXCL11 (J) CXCL12 (K) CXCL13 (L) CXCL16 (M) CXCL17



(e) The relative level of CXC chemokines in LUAD and LUSC

Figure 1: Expression of the CXC chemokine family in lung cancer.

CXCL2/3/4/7/12/16 were decreased in LUAD with significance and CXCL6/10/13/14 were elevated and CXCL2/3/4/5/7/12/16/17 were reduced in LUSC with significance.

In the UALCAN database, except for the expression level, we also explored the methylation level of CXC chemokines in LUAD and LUSC since epigenetic modifications are on the spotlight of tumor research; the results are presented in Figure 1(d). The methylation levels of CXCL7/PPBP, CXCL11, CXCL16, and CXCL17 were significantly reduced in both LUAD and LUSC tissues; the methylation levels of CXCL1, CXCL3, CXCL5, CXCL6, CXCL10, and CXCL12 were significantly elevated in both LUAD and LUSC tissues; the methylation level of CXCL2 was significantly decreased in LUAD, but its data is not significant in LUSC; the methylation level of CXCL4/PF4 was only significantly higher in LUSC. For CXCL13, its methylation level is significantly elevated in LUAD, but is significantly decreased in LUSC, which is the opposite trend in LUAD and LUSC.

We also compared the relative expression levels of CXC chemokines in LUAD and LUSC tissues using GEPIA in Figure 1(e) and found that among all CXC chemokines we evaluated, the relative level of CXCL17 was the highest in LUAD; those of CXCL1, CXCL8, CXCL10, CXCL14, CXCL16, and CXCL17 were higher in LUSC. We evaluated all the 16 CXC chemokines so that we could identify additional CXC chemokines associated with tumorigenesis, development, and clinical outcome in LUAD and LUSC.

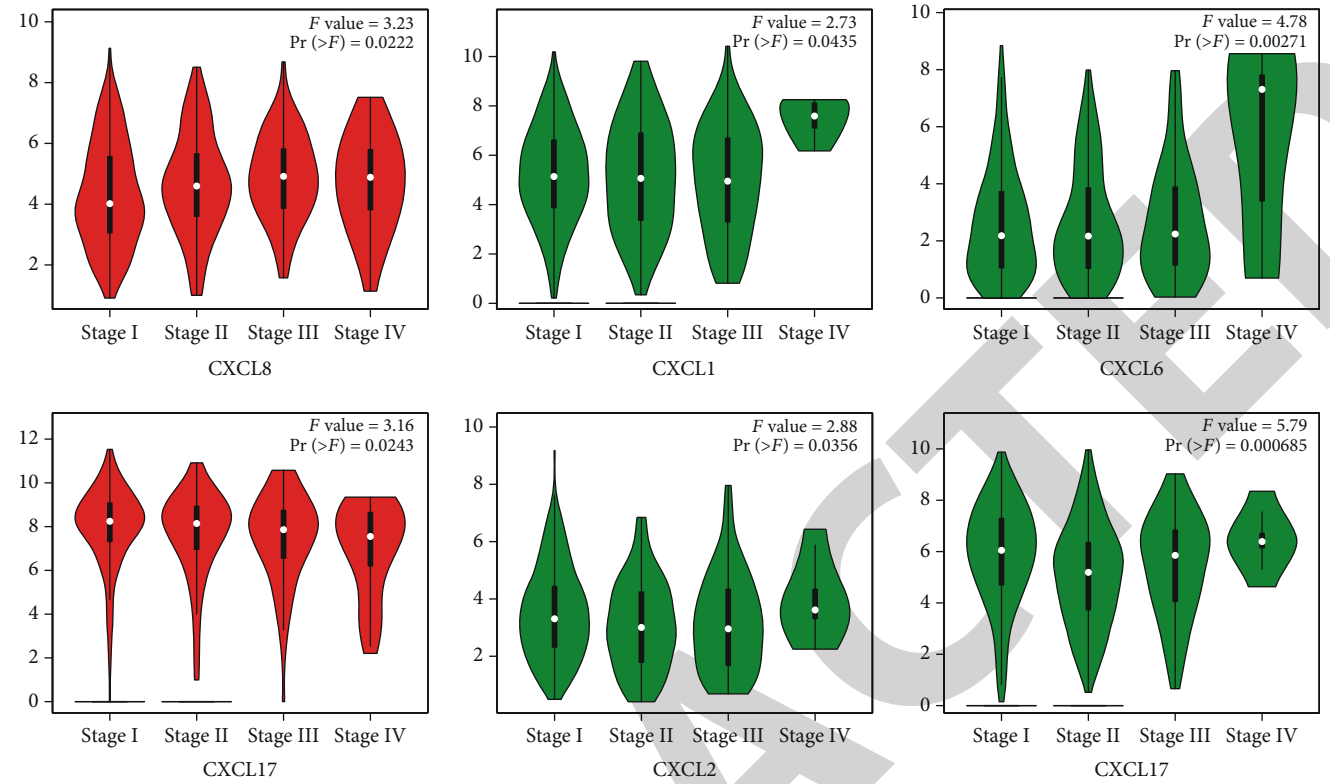
3.2. The Prognostic Value of CXC Chemokines in Patients with LUAD and LUSC. We firstly assessed the correlation between the CXC chemokines and the pathological stage of LUAD and LUSC patients in Figure 2(a). In LUAD, significant associations were found between the expressions of CXCL8 ($p = 0.022$) and CXCL17 ($p = 0.024$). As the tumor progressed, the expression of CXCL8 increased and the expression of CXCL17 decreased. In LUSC, we found a significant correlation between the expressions of CXCL1 ($p = 0.044$), CXCL2 ($p = 0.036$), CXCL6 ($p = 0.003$), and CXCL17 ($p \leq 0.001$), and as the tumor progressed to stage IV, the expression levels of CXCL1 and CXCL6 increased positively.

To evaluate the value of CXC chemokines in the progression of LUAD and LUSC, we assessed the correlation between CXC chemokines and clinical outcome using Kaplan-Meier Plotter-Lung Cancer. OS (overall survival) curves are presented. In LUAD (Figure 2(b)), patients with low expression levels of CXCL1 ($p = 0.024$), CXCL4 ($p = 0.043$), CXCL7 ($p = 0.034$), and CXCL8 ($p = 0.008$) were significantly associated with longer OS; patients with high expression levels of CXCL12 ($p \leq 0.001$), CXCL14 ($p = 0.007$), and CXCL16 ($p = 0.003$) were significantly associated with longer OS. In LUSC (Figure 2(c)), all of the 16 CXC chemokines in our study have no statistical significance correlation with clinical outcome, so we displayed the OS curve of the CXC chemokines that are significant in LUAD to make a comparison.

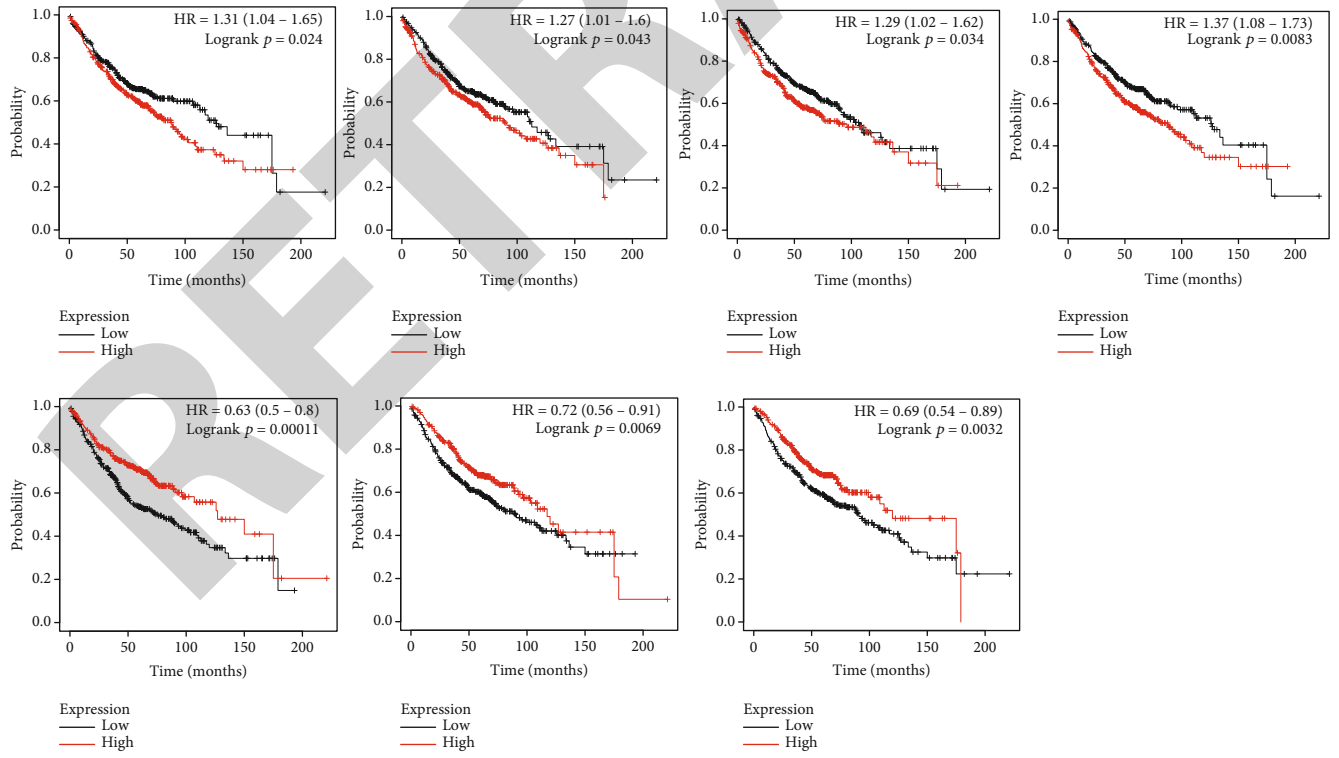
We also conducted a survival map using GEPIA2.0, which displayed the survival contribution of different CXC chemokines in multiple cancer types, including LUAD and LUSC (Figure 2(d).

We conducted a comprehensive analysis of the molecular characteristics of CXC chemokines. Using the cBioPortal database, CXCL1, CXCL2, CXCL3, CXCL4, CXCL5, CXCL6, CXCL7, CXCL8, CXCL9, CXCL10, CXCL11, CXCL12, CXCL13, CXCL14, CXCL16, and CXCL17 were mutated 4, 5, 5, 4, 4, 6, 5, 4, 6, 5, 5, 5, 6, 6, 5, and 6% of the queried LUAD samples, respectively (Figure 3(a)), and they were mutated in 7, 6, 7, 6, 7, 9, 7, 5, 7, 6, 6, 4, 6, 5, 6, and 7% of the queried LUSC samples, respectively (Figure 3(a)). Enhanced/declined mRNA expression was the most common change in these samples. Next, we conducted a correlation heat map of CXC chemokines in different types of cancer, including LUAD and LUSC using the "Chemokine" module of the TISIDB database (Figure 3(b)); blank areas in CXCL4 and CXCL17 were due to the data vacancy of these two CXC chemokines in TISIDB. There was a moderate to high correlation among CXCL1, CXCL2, CXCL3, CXCL4, CXCL5, CXCL6, and CXCL8 and another moderate to high correlation among CXCL9, CXCL10, CXCL11, and CXCL13. Using the "Correlation Analysis" module of GEPIA, we conducted gene-to-gene correlation analysis among CXC chemokines (Figure 3(c)). Under the standard "p value < 0.05 and R value ≥ 0.6," in LUAD, we found significant expression correlation between these pairs: CXCL1 and CXCL2 ($R = 0.6$), CXC1 and CXCL3 ($R = 0.69$), CXCL1 and CXCL8 ($R = 0.7$), CXCL2 and CXCL3 ($R = 0.79$), CXCL4 and CXCL7 ($R = 0.91$), CXCL9 and CXCL10 ($R = 0.6$), CXCL9 and CXCL11 ($R = 0.62$), CXCL10 and CXCL11 ($p < 0.05$; $R = 0.82$); in LUSC, we found significant expression correlation between these couples: CXCL1 and CXCL8 ($R = 0.68$), CXC2 and CXCL3 ($R = 0.79$), CXCL4 and CXCL7 ($R = 0.92$), and CXCL10 and CXCL11 ($R = 0.72$). Moreover, we made a PPI network analysis of CXC chemokines in humans with STRING in July 2020 to explore the potential interactions among them. As expected, several nodes (16) and several edges (111) were obtained in the PPI network (Figure 3(d)), and the PPI enrichment p value <1.0e-16. The function of these CXC chemokines was associated with the chemokine signaling pathway and the inflammatory reaction. Outcomes of GeneMANIA also showed that the functions of these CXC chemokines were mainly related to chemotaxis and several chemokine activities such as chemokine receptor binding process; it showed 20 related genes and 1512 total links (Figure 3(e)).

3.3. Function Enrichment and Pathway Analysis of CXC Chemokines in Patients with LUAD and LUSC. R software with package "clusterProfiler" was applied to analyze the functions of differentially expressed CXC chemokines and their neighboring genes in LUAD and LUSC, respectively (Figure 4). Figure 4(a) shows the top 10 most highly enriched GO items in LUAD. Among the 10 most highly enriched functions in the BP category, chemokine-mediated signaling pathway, response to chemokine, cellar response to chemokine, and cell chemotaxis were associated with the tumorigenesis and progression of LUAD. The external side of the plasma membrane, focal adhesion, and cell-substrate junction were

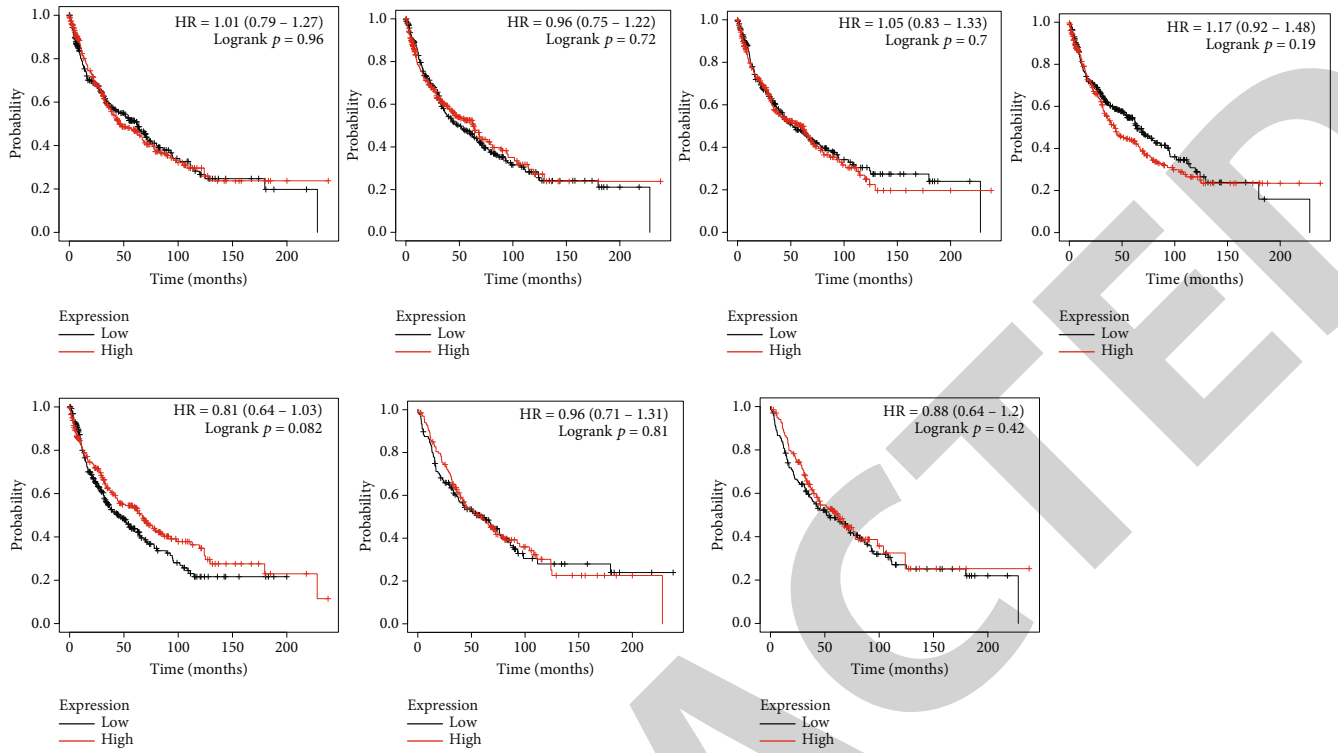


(a) Correlation between different expressed CXC chemokines and the pathological stage of LUAD (red color) and LUSC (green color) patients (GEPIA); $^{*}p < 0.05$

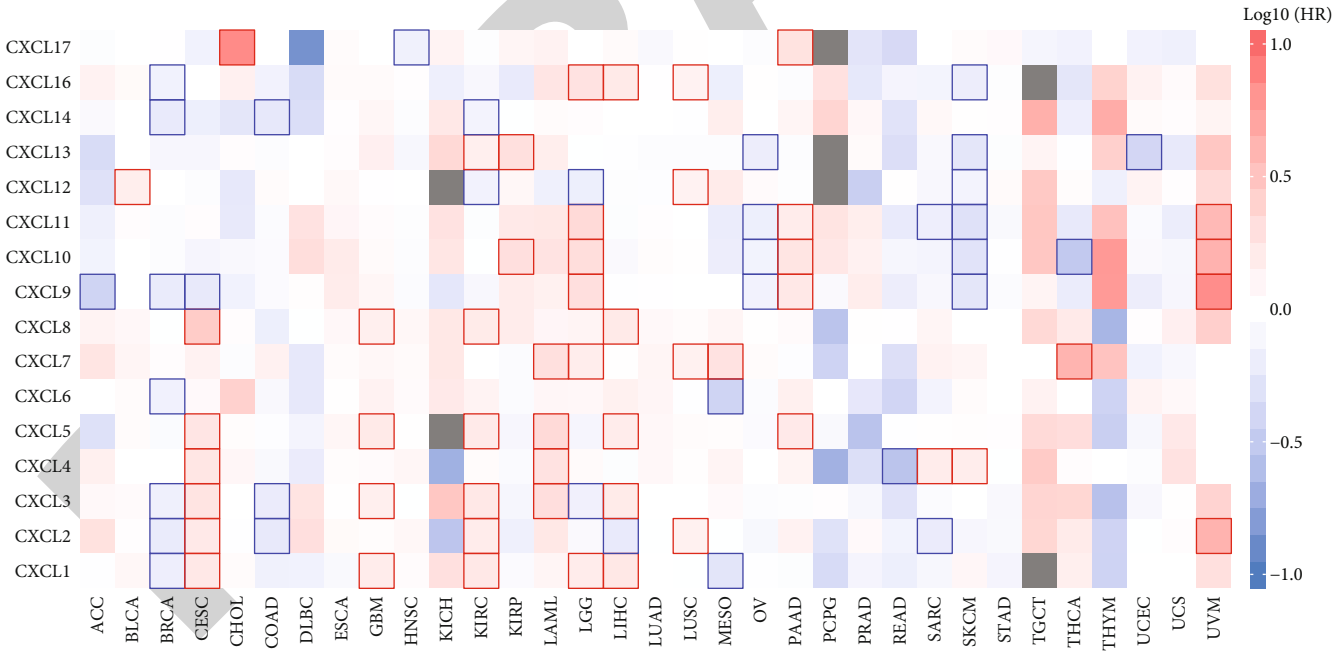


(b) The overall survival curve of different expressed CXC chemokines in LUAD patients. Upper from left to right are CXCL1, CXCL4, CXCL7, and CXCL8; lower from left to right are CXCL12, CXCL14, and CXCL16; $^{*}p < 0.05$

Figure 2: Continued.



(c) The overall survival curve of different expressed CXC chemokines in LUSC patients. Upper from left to right are CXCL1, CXCL4, CXCL7, and CXCL8; lower from left to right are CXCL12, CXCL14, and CXCL16

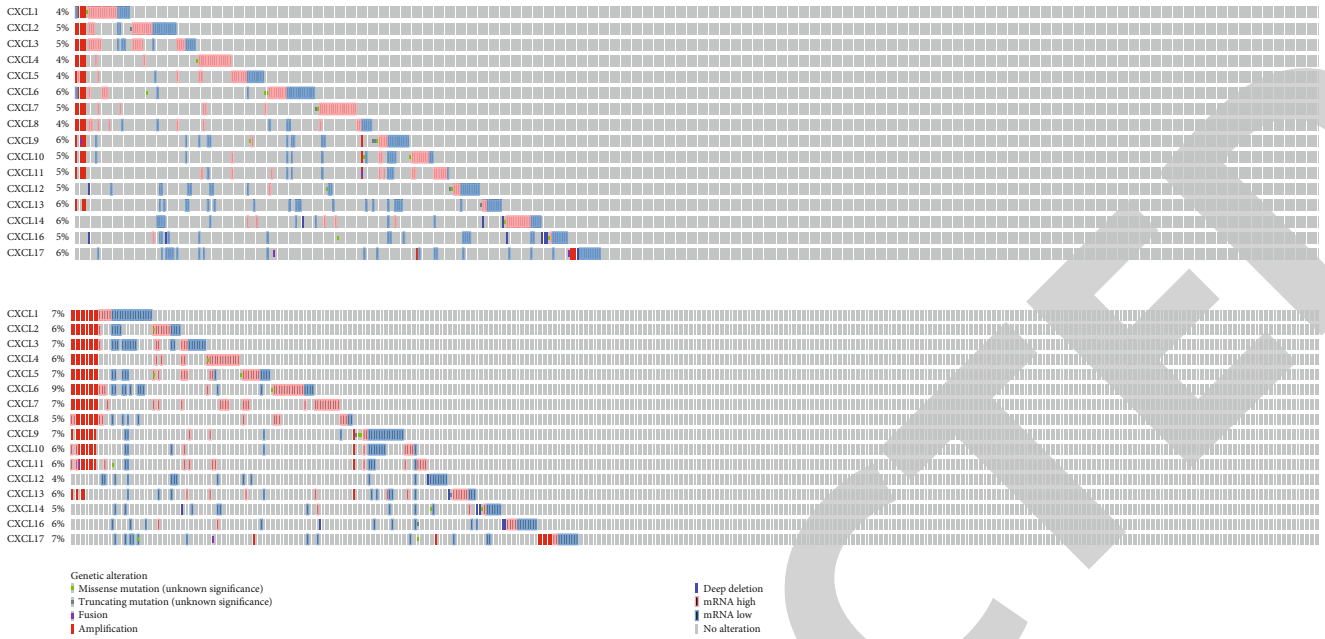


(d) Survival map, which showed the survival contribution of different CXC chemokines in multiple cancer types, estimated using the Mantel-Cox test (GEPIA)

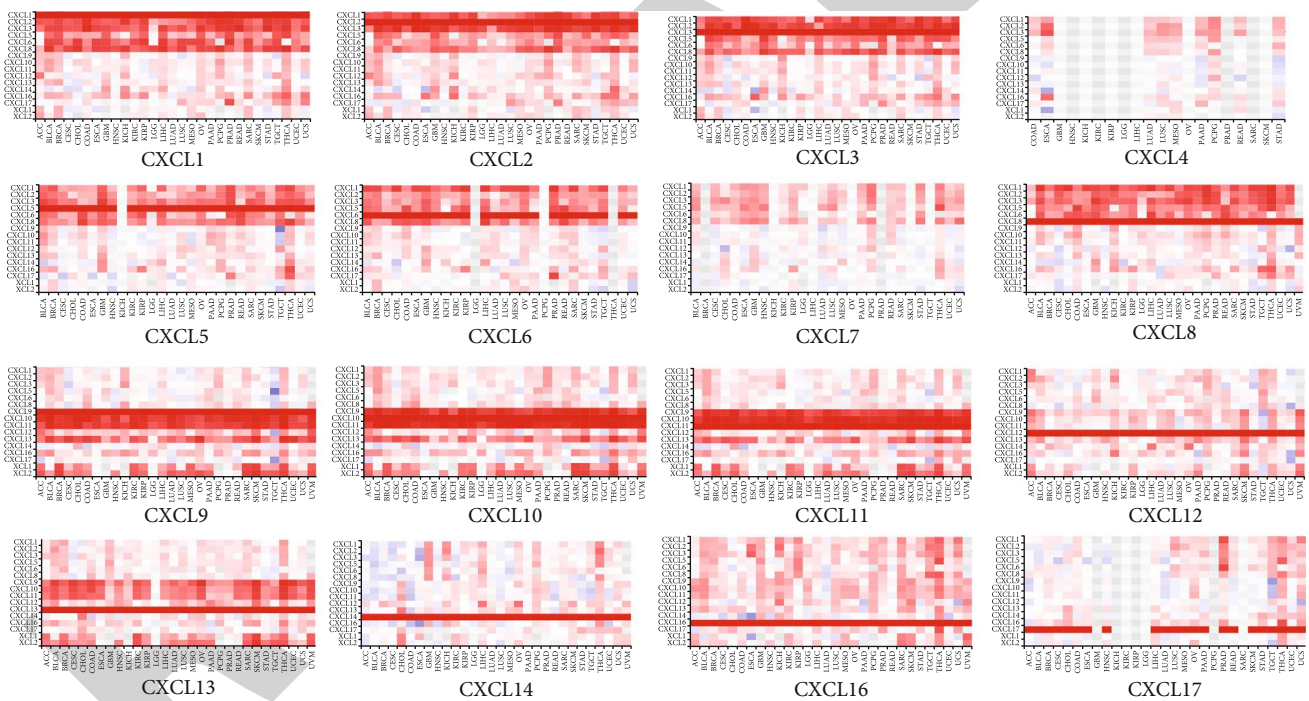
Figure 2: Prognostic value of different expressed CXC chemokines in lung cancer.

the 3 most highly enriched items in the CC category. In the molecular function part, the CXC chemokines and their neighboring genes were mainly enriched in chemokine receptor binding, chemokine activity, G protein-coupled receptor binding, cytokine receptor binding, and cytokine activity.

Figure 4(b) showed the KEGG pathway analyses of the CXC chemokine in LUAD. As expected, among the top 30 KEGG pathways, CXC chemokines were mainly enriched in the chemokine signaling pathway, viral protein interaction with cytokine and cytokine receptor, cytokine-cytokine receptor interaction, Kaposi sarcoma-associated herpesvirus

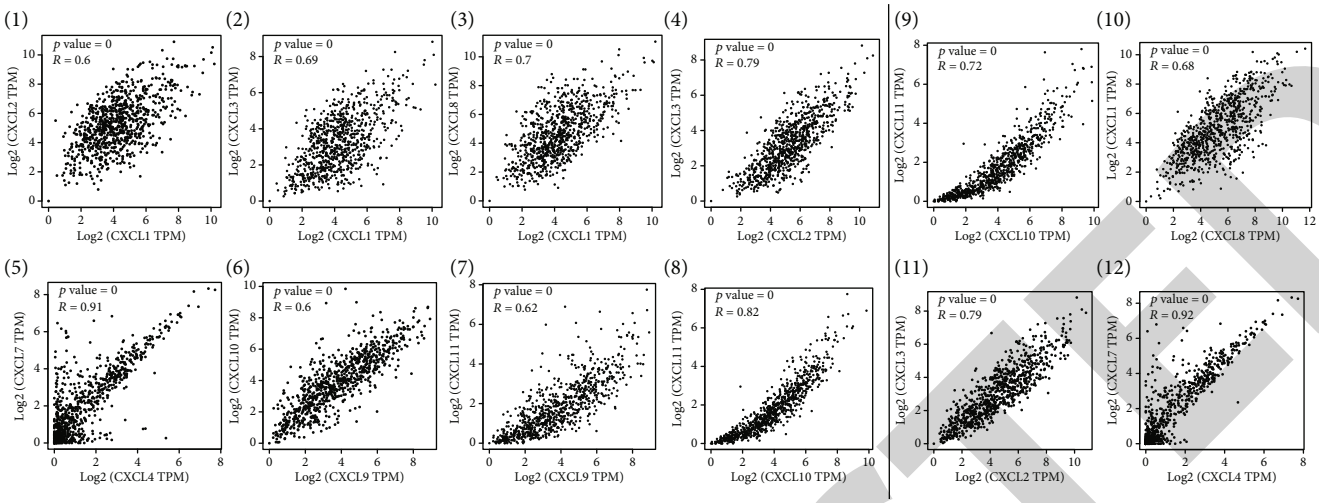


(a) Summary of alterations in CXC chemokines in LUAD (upper) and LUSC (lower) (cBioPortal)

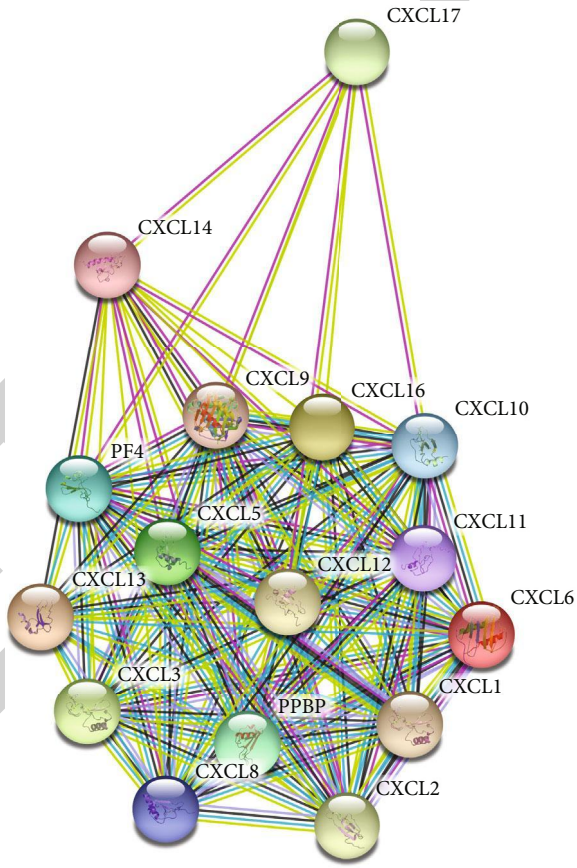


(b) Correlation heat map of different expressed CXC chemokines in different types of cancer, including LUAD and LUSC (TISIDB)

Figure 3: Continued.

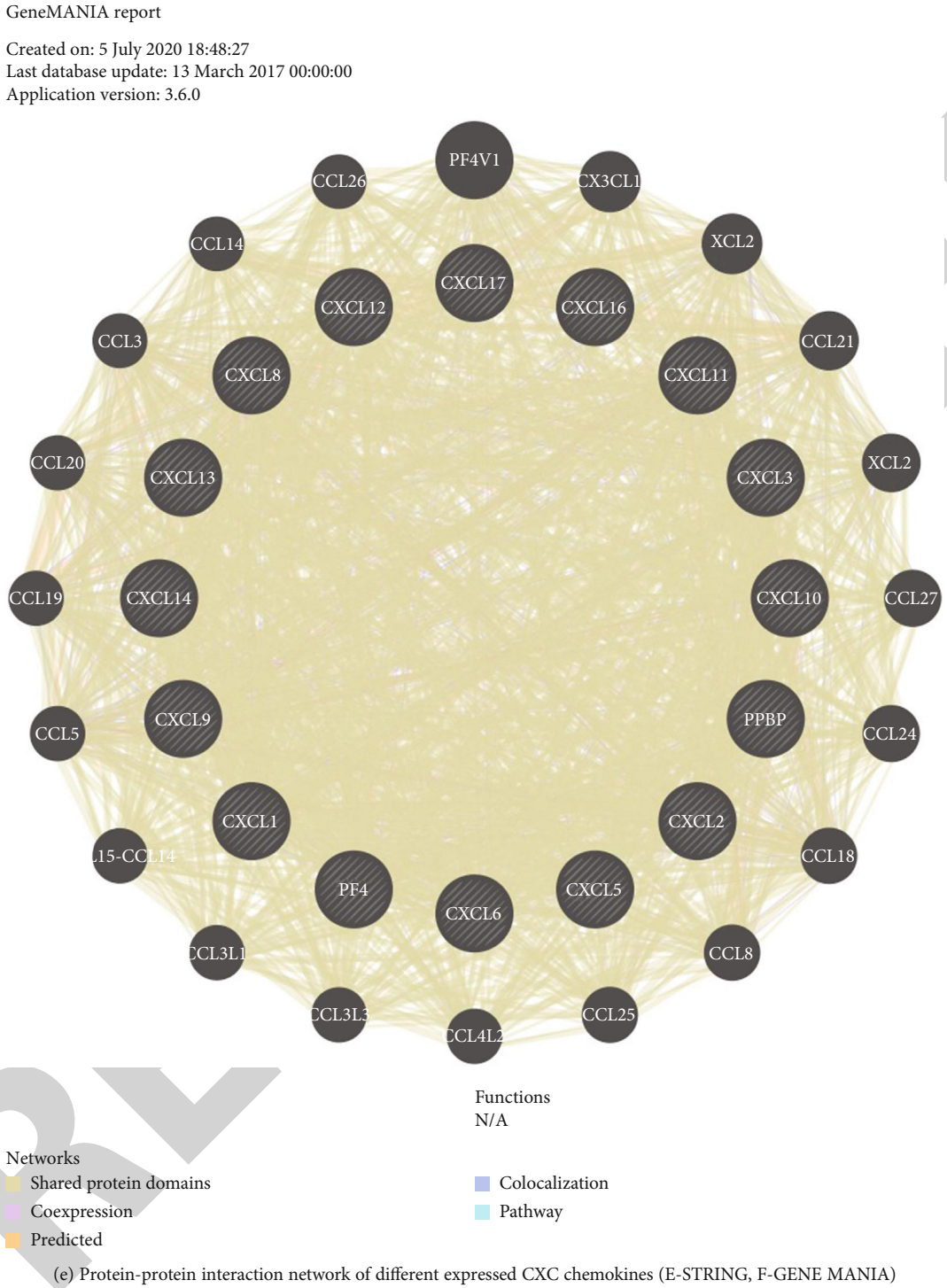


(c) Correlation analysis between different CXC chemokines with statistical significance ($p < 0.05$) in LUAD and LUSC (GEPIA). (1) to (8) were analyses in LUAD patients; (9) to (12) were analyses in LUSC patients



(d) Protein-protein interaction network of different expressed CXC chemokines (E-STRING, F-GENE MANIA)

Figure 3: Continued.



(e) Protein-protein interaction network of different expressed CXC chemokines (E-STRING, F-GENE MANIA)

Figure 3: Genetic alteration, interaction analyses, and neighbor gene network of different expressed CXC chemokines in LUAD and LUSC patients.

infection, and human cytomegalovirus infection. Figure 4(c) displayed the GO enrichment analysis in LUSC. The top 10 items are cell chemotaxis, chemokine-mediated signaling pathway, response to chemokine, cellular response to chemokine, leukocyte migration, myeloid leukocyte migration, positive regulation of chemotaxis, granulocyte chemotaxis, and positive regulation for leukocyte chemotaxis. The external side of the plasma membrane, focal adhesion, and cell-substrate junction were the 3 most highly enriched items in the CC category. In the molecular function category, the CXC chemokines and their neighboring genes were mainly enriched in G protein-coupled receptor binding, cytokine

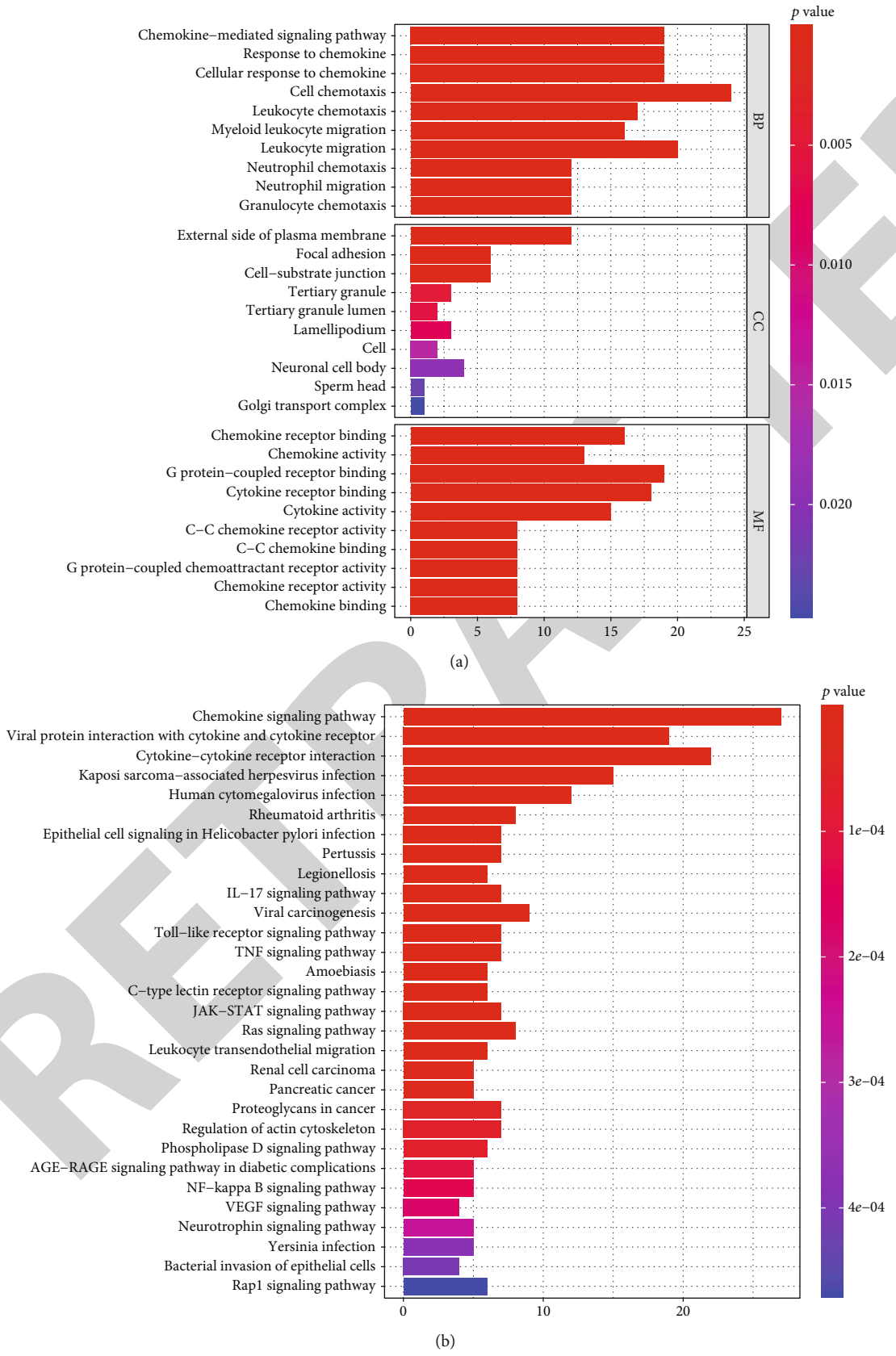


Figure 4: Continued.

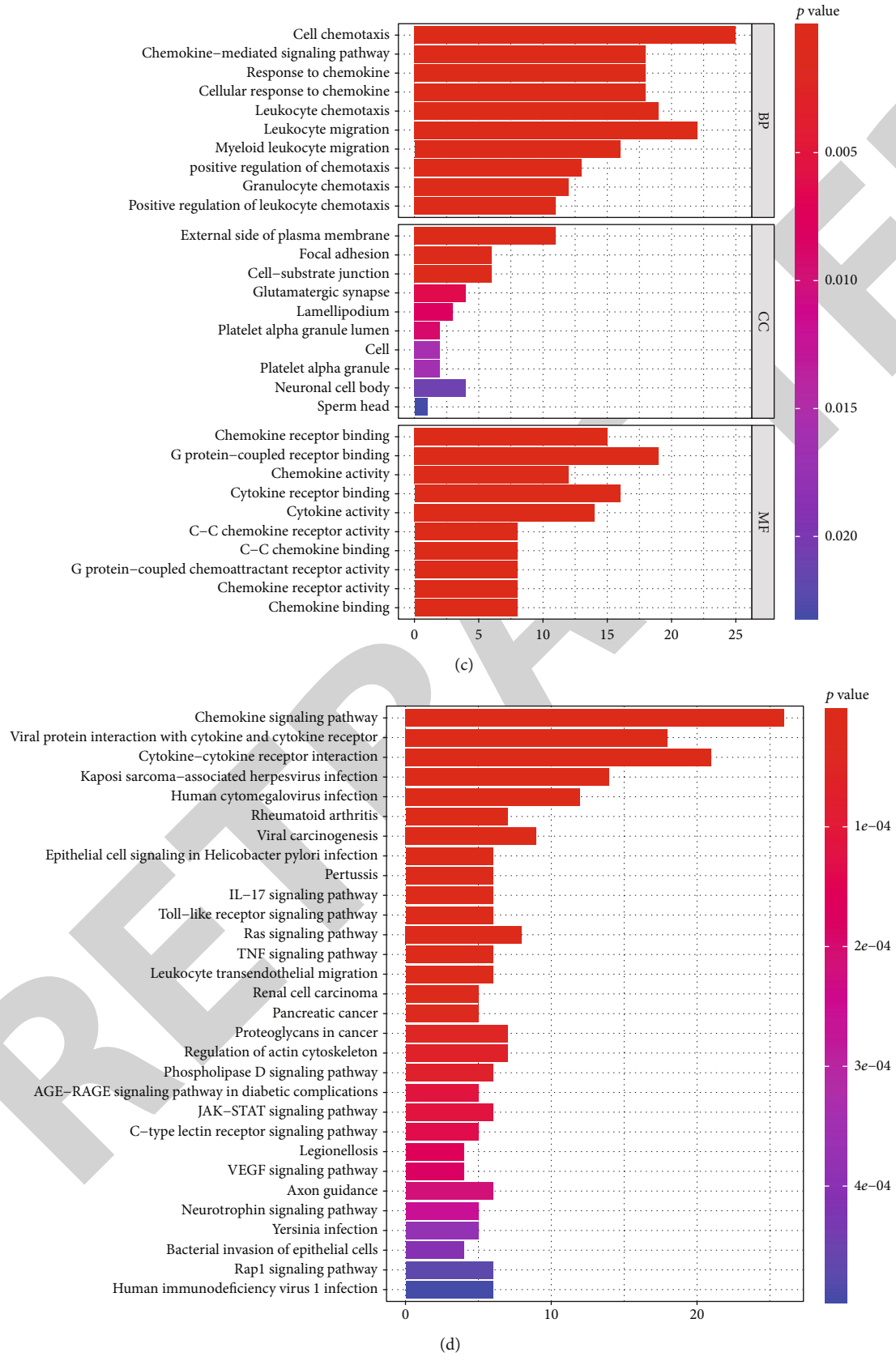


Figure 4: Bar plots of GO enrichment analysis in cellular component terms, biological process terms, and molecular function terms and KEGG enriched terms of different expressed CXC chemokines in LUAD and LUSC: (a, b) LUAD; (c, d) LUSC.

Table 1: Key regulated factor of CXC chemokines in human (TRRUST).

Key TF	Description	Regulated genes	p value	FDR
RELA	v-rel reticuloendotheliosis viral oncogene homolog A (avian)	CXCL1, CXCL2, CXCL5, CXCL8, CXCL10, CXCL12	4.22E-08	6.58E-08
NFKB1	Nuclear factor of kappa light polypeptide gene enhancer in B cell 1	CXCL1, CXCL2, CXCL5, CXCL8, CXCL10, CXCL12	4.39E-08	6.58E-08
SP1	Sp1 transcription factor	CXCL1, CXCL5, CXCL14	0.00461	0.00461

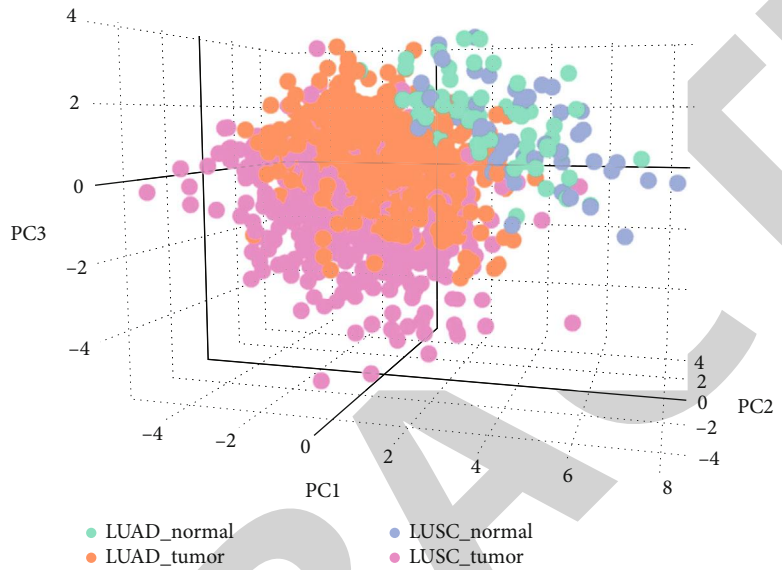
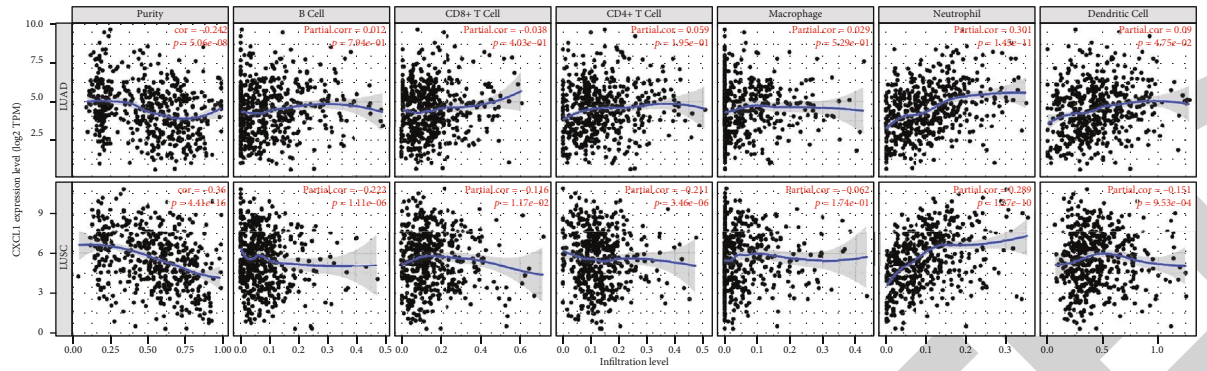


Figure 5: PCA (principal component analysis) on samples from TCGA LUAD and TCGA LUSC based on their expression of CXC chemokines (GEPIA2.0).

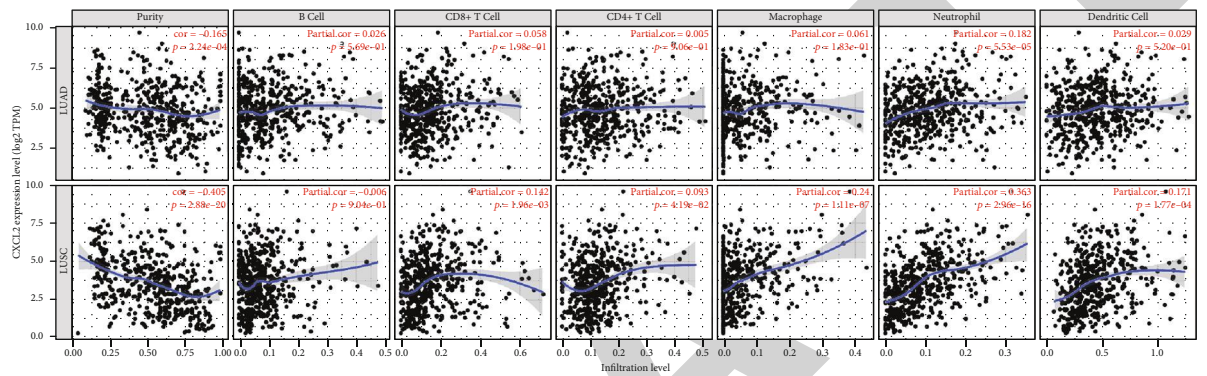
receptor binding, and chemokine receptor binding. Figure 4(d) is the KEGG pathway analyses of the CXC chemokine in LUSC where CXC chemokines were mainly enriched in the chemokine signaling pathway, cytokine-cytokine receptor interaction, and viral protein interaction with the cytokine and cytokine receptor.

3.4. Transcription Factor Targets and Principal Component Analysis of CXC Chemokines in Patients with LUAD and LUSC. Since there is significant difference in the expressions of CXC chemokines in LUAD/LUSC vs. normal tissue, we explored possible transcription factor targets of the CXC chemokines using the TRRUST database in Table 1. CXCL1, CXCL2, CXCL3, CXCL4, CXCL5, CXCL6, CXCL7, CXCL8, CXCL9, CXCL10, CXCL12, CXCL13, CXCL14, CXCL16, and CXCL17 were included in TRRUST. Three transcription factors (RELA, NFKB1, and SP1) were found to be related to the regulation of CXC chemokines. RELA and NFKB1 were the key transcription factors for CXCL1, CXCL2, CXCL5, CXCL8, CXCL10, and CXCL12. SP1 was the key transcription factor for CXCL1, CXCL5, and CXCL14. Figure 5 showed the 3D PCA (principal component analysis) on samples from the TCGA tumor and TCGA normal, including LUAD, LUAD normal, LUSC, and LUSC normal, based on their expression of CXC chemokines using the GEPIA2.0 database.

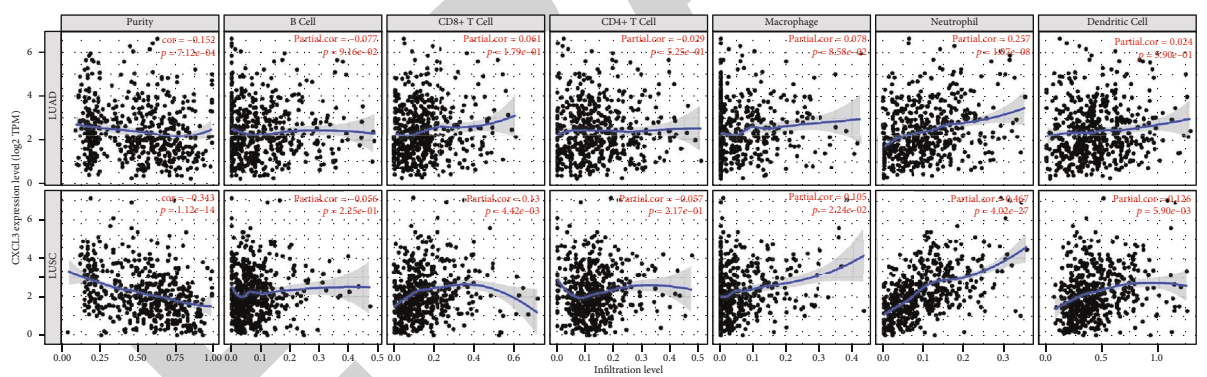
3.5. Immune Cell Infiltration Levels among Tumors with Different SCNA of CXC Chemokines in Patients with LUAD and LUSC. CXC chemokines are part of immune cell infiltration and inflammatory responses and so can affect the clinical outcome of LUAD and LUSC patients. Therefore, we explored the correlation between CXC chemokines and immune cell infiltration with the TIMER database. The correlation maps are displayed from CXCL1, CXCL2 to CXCL17 in Figure 6. To extract the key information more easily, we made summaries for Figure 6 in Tables 2 and 3, which are for LUAD and LUSC, respectively. CXCL9, CXCL10, CXCL11, CXCL12, and CXCL16 were chemokines that were positively correlated with all 6 types of immune cells in both LUAD and LUSC. In LUAD, expressions of CXCL4, CXCL5, CXCL6, CXCL7, and CXCL8 were negatively correlated with $CD4^+$ T cells. CXCL5, CXCL7, CXCL8, and CXCL17 were chemokines that had both positive and negative associations among different immune cells. In LUSC, all 16 CXC chemokines (except CXCL17) had positive or negative correlation with neutrophil cells, and CXCL13 was positively associated with all 6 types of immune cell. We also evaluated the correlation of CXC chemokines and immune cell infiltration (Tables 4(a) and 4(b)). The Cox proportional hazard model was adopted. B cells ($p < 0.01$) was significantly associated with the clinical outcome of LUAD patients (Table 4(a)). In LUSC patients, CXCL2 ($p < 0.05$), CXCL8 ($p < 0.05$), and



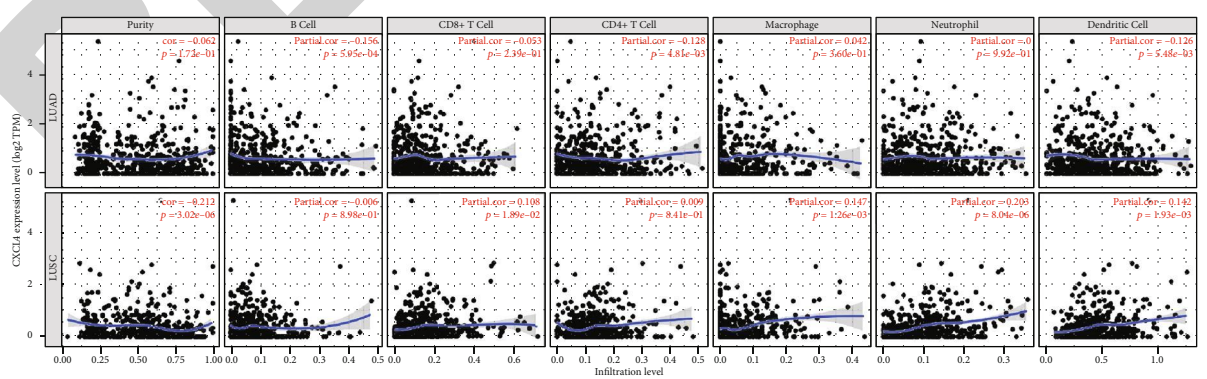
(a)



(b)

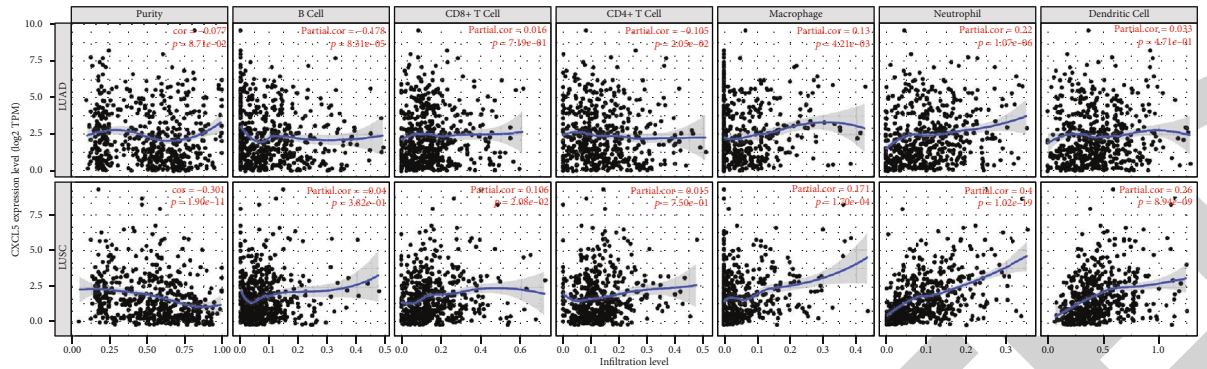


(c)

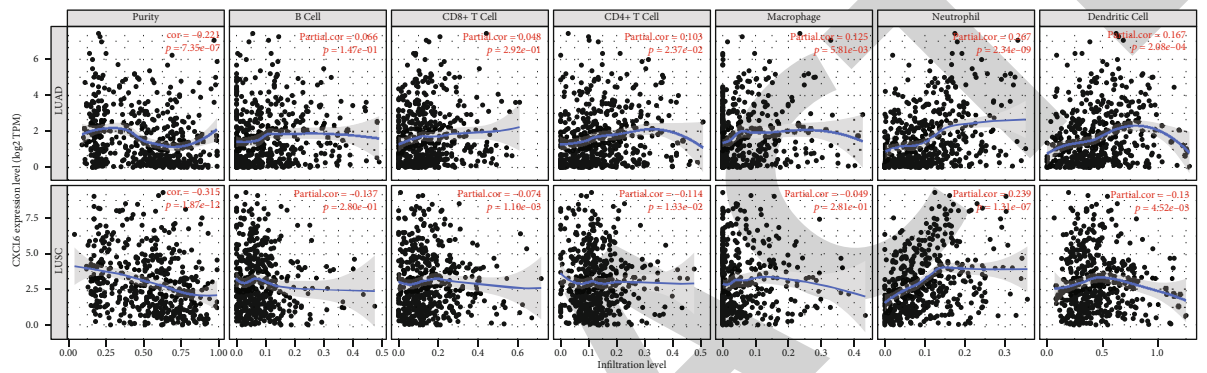


(d)

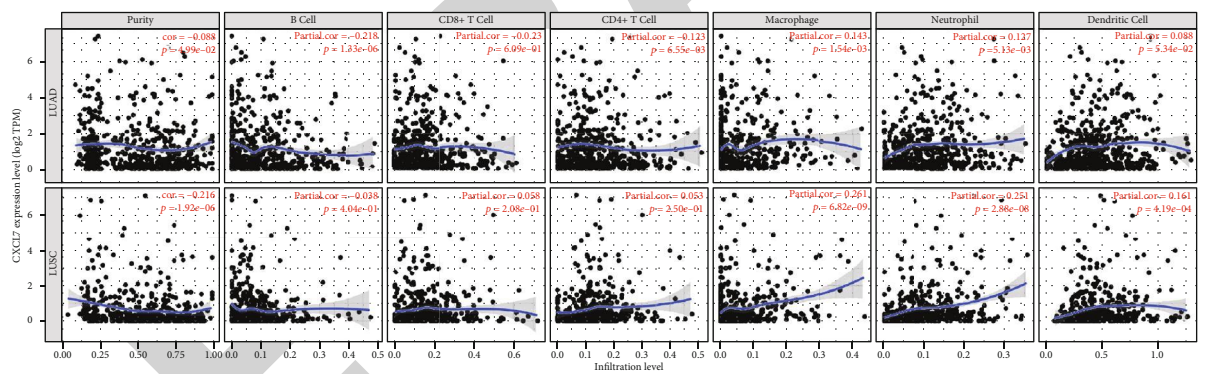
Figure 6: Continued.



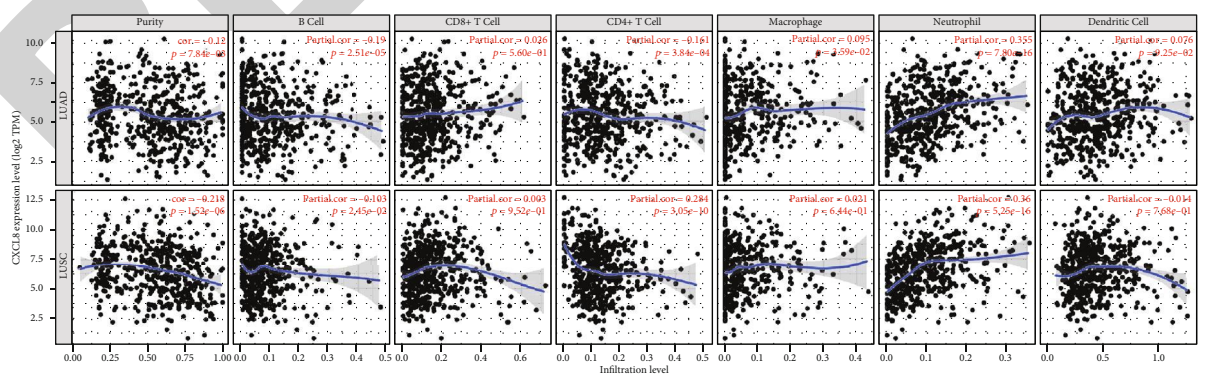
(e)



(f)

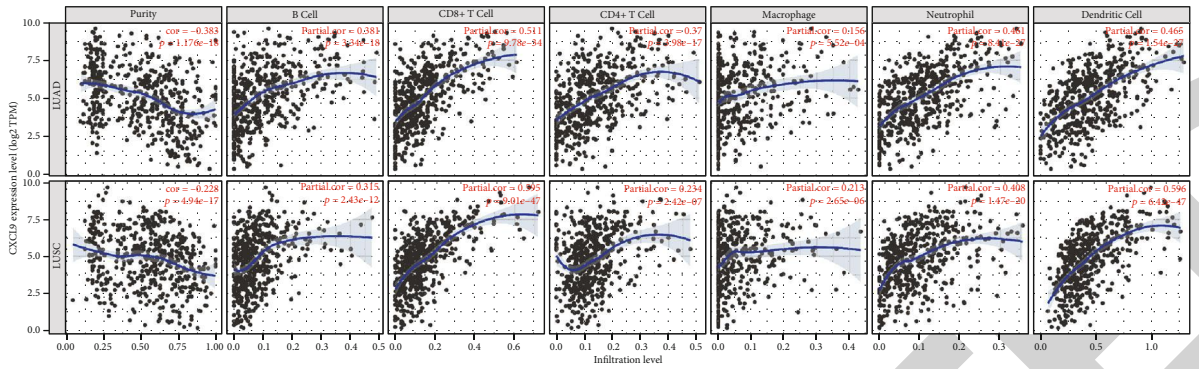


(g)

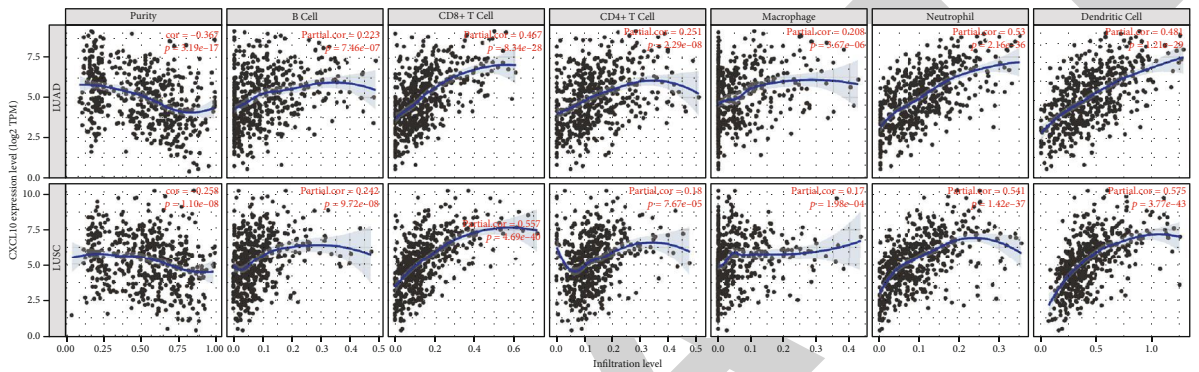


(h)

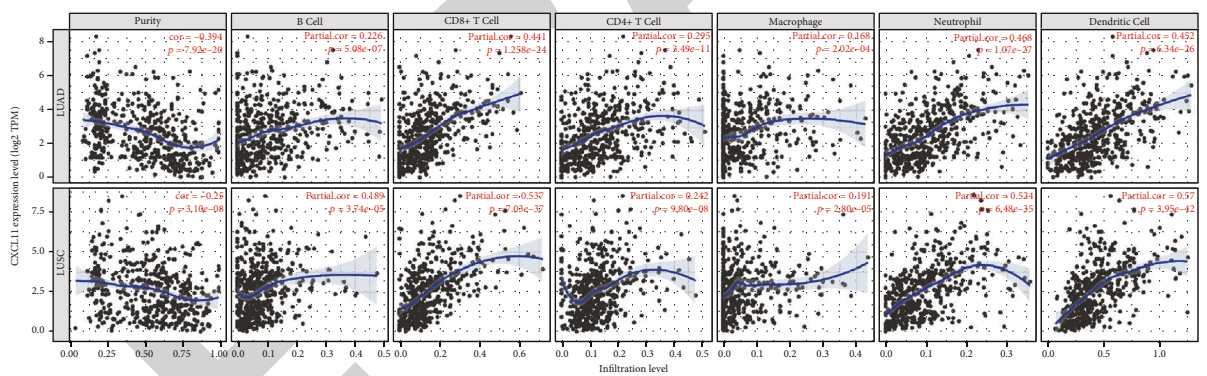
Figure 6: Continued.



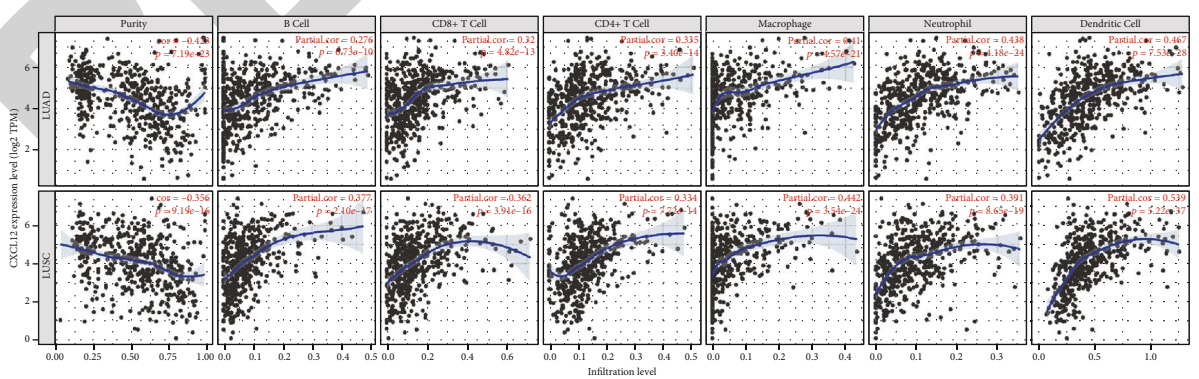
(i)



(j)

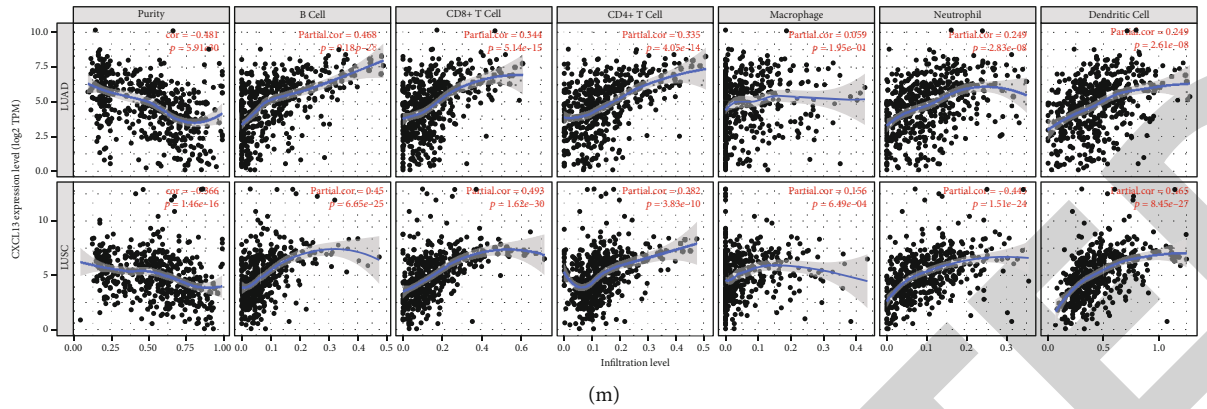


(k)

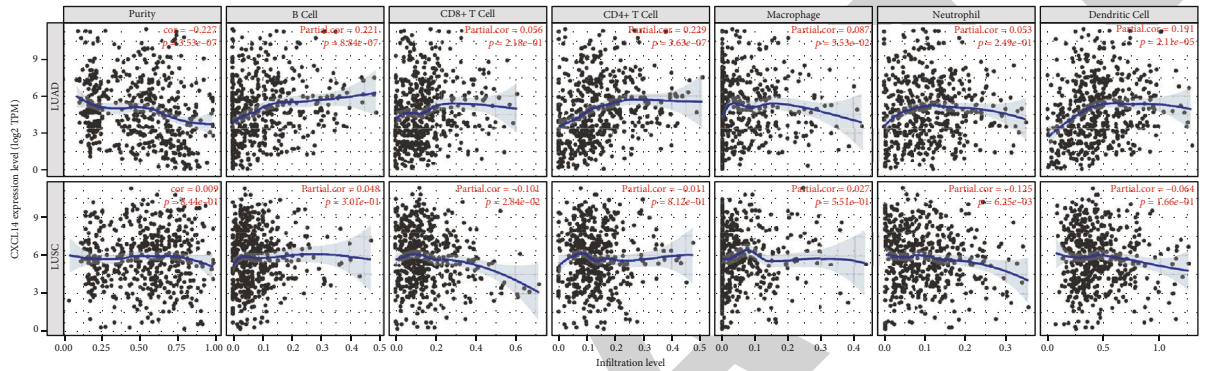


(l)

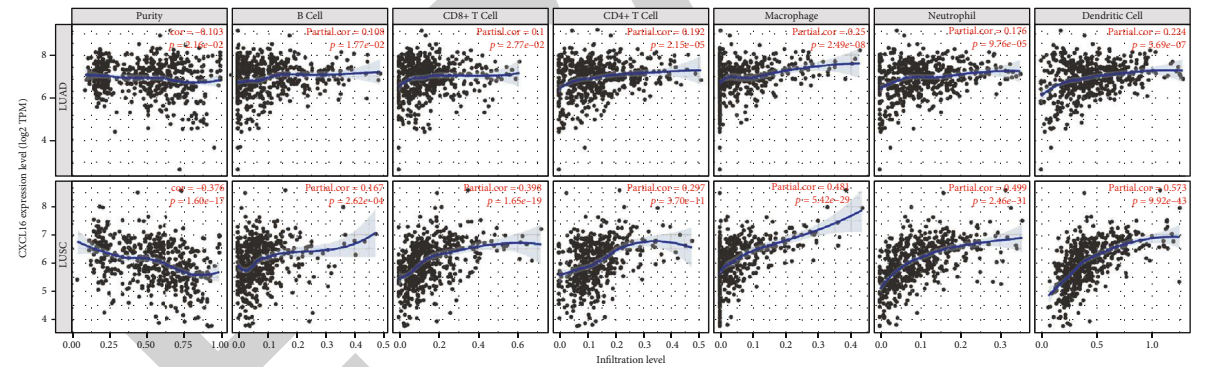
Figure 6: Continued.



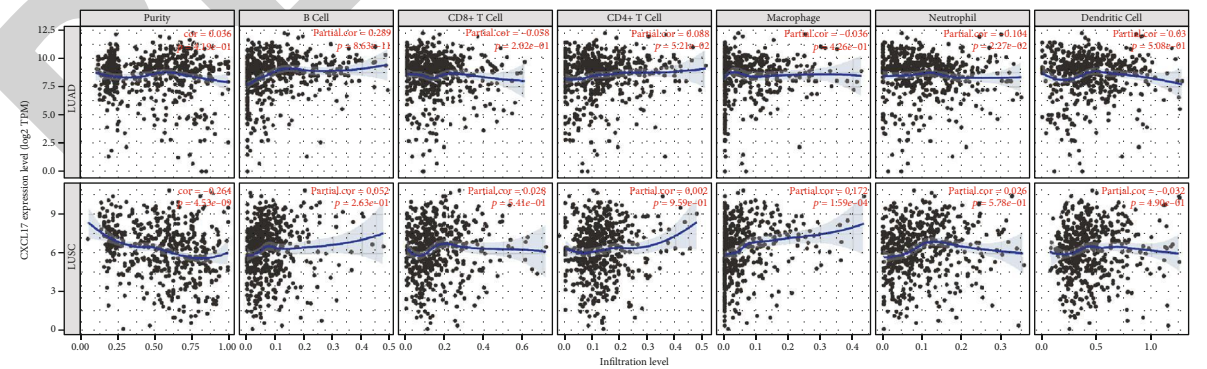
(m)



(n)



(o)



(p)

FIGURE 6: The correlation between different expressed CXC chemokines and immune cell infiltration in LUAD and LUSC patients (TIMER). The correlation between the abundance of immune cell and the expression of (a) CXCL1, (b) CXCL2, (c) CXCL3, (d) CXCL4, (e) CXCL5, (f) CXCL6, (g) CXCL7, (h) CXCL8, (i) CXCL9, (j) CXCL10, (k) CXCL11, (l) CXCL12, (m) CXCL13, (n) CXCL14, (o) CXCL16, and (p) CXCL7 in LUAD and LUSC patients.

Table 2: Summary of LUAD data in Figure 6.

	B cell	CD8+ cell	CD4+ cell	Macrophage	Neutrophil	Dendritic cell
CXCL1					*	*
CXCL2					*	
CXCL3						
CXCL4	**		**			**
CXCL5	**		**	*	*	
CXCL6			**	**	**	**
CXCL7	**		**	*	*	
CXCL8	**		**	*	*	
CXCL9	*	*	*	*	*	*
CXCL10	*	*	*	*	*	*
CXCL11	*	*	*	*	*	*
CXCL12	*	*	*	*	*	*
CXCL13	*	*	*		*	*
CXCL14	*		*			*
CXCL16	*	*	*	*	*	*
CXCL17	*				**	

*$p < 0.05$, positively correlated; **$p < 0.05$, negatively correlated.

Table 3: Summary of LUSC data in Figure 6.

	B cell	CD8+ cell	CD4+ cell	Macrophage	Neutrophil	Dendritic cell
CXCL1	*	*	*		*	*
CXCL2		*	*	*	*	*
CXCL3		*		*	*	*
CXCL4		**		**	**	**
CXCL5		**		*	*	*
CXCL6	**		**		**	**
CXCL7				*	*	*
CXCL8	**		**		*	
CXCL9	*	*	*	*	*	*
CXCL10	*	*	*	*	*	*
CXCL11	*	*	*	*	*	*
CXCL12	*	*	*	*	*	*
CXCL13	*	*	*	*	*	*
CXCL14		*			*	
CXCL16	*	*	*	*	*	*
CXCL17				*		

*$p < 0.05$, positively correlated; **$p < 0.05$, negatively correlated.

CXCL13 ($p < 0.05$) were significantly associated with the clinical outcome.

At last, we conducted the comparison of tumor infiltration levels among tumors with different SCNAs (somatic copy number alterations) for CXC chemokines in LUAD and LUSC samples (Figure 7) using the TIMER database. From the results, we found that in both LUAD and LUSC patients, all of the SCNA types that are significantly correlated with immune cells are (1) arm-level deletion, (2) arm-level gain, and (3) high amplification. In LUAD, all 16 CXC chemokines have a significant part in $CD4^+$ T cells (in which SCNA types are mainly arm-level depletion) and neutrophil cells (in which SCNA types are mainly high amplification). In LUSC, all 16 CXC chemokines have a significant part in $CD4^+$ T cells, macrophage cells, and neutrophil cells; predominant SCNA types of the three types are the

Table 4: The cox proportional hazard model of CXC chemokines and six tumor-infiltrating immune cells in LUAD and LUSC patients (TIMER): (a) LUAD; (b) LUSC.

(a)

	Coef	HR	95% CI_l	95% CI_u	*p* value	Sig
B_cell	-4.563	0.010	0.001	0.163	0.001	**
CD8_T cell	0.412	1.510	0.171	13.339	0.711	
CD4_T cell	2.862	17.503	0.715	428.624	0.079	
Macrophage	0.052	1.054	0.066	16.919	0.971	
Neutrophil	-2.638	0.072	0.000	11.525	0.309	
Dendritic	0.224	1.251	0.281	5.565	0.769	
CXCL1	-0.039	0.962	0.813	1.138	0.651	
CXCL2	-0.096	0.908	0.768	1.074	0.261	
CXCL3	0.231	1.259	0.978	1.622	0.074	
CXCL4	-0.107	0.898	0.742	1.088	0.274	
CXCL5	0.039	1.040	0.937	1.154	0.463	
CXCL6	-0.001	0.999	0.863	1.156	0.984	
CXCL7	-0.022	0.978	0.861	1.112	0.738	
CXCL8	0.068	1.071	0.919	1.247	0.382	
CXCL9	-0.008	0.992	0.789	1.248	0.948	
CXCL10	-0.172	0.842	0.592	1.197	0.339	
CXCL11	0.238	1.269	0.964	1.670	0.089	
CXCL12	-0.047	0.954	0.810	1.123	0.571	
CXCL13	-0.010	0.990	0.884	1.108	0.857	
CXCL14	0.043	1.044	0.976	1.117	0.210	
CXCL16	-0.047	0.954	0.770	1.182	0.666	
CXCL17	-0.074	0.929	0.855	1.009	0.079	

*$p < 0.05$, $p < 0.01$, and $p < 0.001$.

(b)

	Coef	HR	95% CI_l	95% CI_u	*p* value	Sig
B_cell	1.631	5.110	0.341	76.487	0.237	
CD8_T cell	0.627	1.873	0.266	13.192	0.529	
CD4_T cell	2.269	9.674	0.628	149.081	0.104	
Macrophage	-2.730	0.065	0.004	1.010	0.051	
Neutrophil	1.133	3.105	0.037	263.329	0.617	
Dendritic	0.626	1.870	0.404	8.653	0.423	
CXCL1	-0.109	0.896	0.779	1.031	0.127	
CXCL2	0.180	1.197	1.012	1.415	0.035	*
CXCL3	-0.174	0.840	0.654	1.078	0.172	
CXCL4	0.015	1.015	0.739	1.394	0.927	
CXCL5	-0.007	0.993	0.872	1.132	0.921	
CXCL6	-0.047	0.954	0.860	1.059	0.378	
CXCL7	0.095	1.099	0.936	1.291	0.249	
CXCL8	0.162	1.176	1.036	1.336	0.012	*
CXCL9	-0.165	0.848	0.718	1.002	0.053	
CXCL10	0.097	1.102	0.841	1.445	0.481	

Table 4: Continued.

	Coef	HR	95% CI_l	95% CI_u	p value	Sig
CXCL11	-0.050	0.951	0.755	1.197	0.667	
CXCL12	0.011	1.011	0.887	1.152	0.866	
CXCL13	-0.111	0.895	0.806	0.994	0.038	*
CXCL14	0.011	1.011	0.946	1.080	0.753	
CXCL16	-0.005	0.995	0.766	1.293	0.970	
CXCL17	-0.059	0.943	0.862	1.032	0.202	

*$p < 0.05$, $p < 0.01$, $p < 0.001$.

same—arm-level alterations. In LUAD samples, CXC14 are only chemokines of which SCNA is significantly correlated with $CD8^+$ T cells among 16 CXC chemokines.

4. Discussion

CXC chemokines and their receptors were initially identified as important regulators in inflammatory response [10]. In cancer, accumulative evidence has proven that CXC chemokines are critical in tumor initiation, angiogenesis, and progression [7], and with the rapid development of tumor immunology, the correlation among CXC chemokines, tumor microenvironment, and cancer immunotherapy have been reported [8, 11].

In NSCLC, much attention has been paid to the functions of some CXC chemokines. For example, the CXCL1 paracrine network was identified to be linked with cancer chemoresistance and metastasis in 2012 [18], and in 2020, researchers found that CXCL1 was an unfavorable prognosis factor negatively regulated by DACH1 in NSCLC using immunohistochemistry staining [38]. In 2004, researchers found that COX-2 (cyclooxygenase-2) contributes to the progression of NSCLC tumorigenesis by enhancing the expression of angiogenic chemokines CXCL8 and CXCL5 [39], both of which can contribute to lung cancer progression. As for the cancer treatment, combination of CXCL9 or CXCL10 with cisplatin improves therapeutic efficacy in solid tumors [40, 41]. Changes in the serum level of CXCL8 can reflect and predict the response to anti-PD-1 therapy in NSCLC [42]. Although the CXC chemokine family in lung cancer has been studied for years, there are still unknown corners and even controversies about several CXC chemokines. For instance, researchers have opposite views about the roles of CXCL4 [22, 43], CXCL14 [24, 44], and CXCL16 [27, 45], and we know little about the positions of CXCL2, CXCL6, and CXCL13 in NSCLC. Generally, although many researchers have studied the function of some CXC chemokines singly and have put up solid evidence, the comprehensive landscape of CXC chemokines and their behavior in LUAD and LUSC distinctly have not been well-portrayed yet.

Firstly, we explored the expression and methylation of CXC chemokines in 3 database and their correlation with pathological stages in LUAD and LUSC. In the ONCOMINE database where lung cancer types are not discriminated, we found that 12 genes were differentially expressed in lung cancer compared with normal tissue (upregulation of CXCL9/13/14; downregulation of CXCL1/2/3/4/5/7/12/16/17). And in the TIMER and UALCAN databases, where LUAD and LUSC are distinguished, we found that 9 genes were differentially expressed in LUAD compared with normal tissue (upregulation of CXCL10/13/14; downregulation of CXCL2/3/4/7/12/16), and in LUSC, 13 genes were differentially expressed compared with normal tissue (upregulation of CXCL6/10/13/14; downregulation of CXCL2/3/4/5/7/11/12/16/17). That is the whole expression condition of CXC chemokines in lung cancer. In the UALCAN database where we conducted the methylation analysis of CXC chemokines, we found that methylation levels of CXC1/3/5/6/10/12 were significantly elevated, and methylation levels of CXCL7/11/16/17 were significantly reduced in both LUAD and LUSC. Suzuki et al. reported that aberrant methylation of CXCL12 in NSCLC is associated with poor prognosis [46], which agrees with our research to some degree and inspires us that the methylation level of other CXC chemokines may be involved with tumor prognosis. Then, we explore the association between CXC chemokines and pathological stages and the clinical prognosis. We found that in LUAD, expression of CXCL8 increased and expression of CXCL17 decreased as the tumor progressed, and in LUSC, the expressions of CXCL1 and CXCL6 increased considerably when the tumor progressed to stage IV. These results imply that CXCL1/6/8/17 may predict the stage or tendency of tumors. In the UALCAN database, we found that LUAD patients with low expression of CXCL1/4/7/8 and LUAD patients with high expression of CXCL12/14/16 were significantly associated with better OS (overall survival), while no CXC chemokines in LUSC have statistical significance. These data demonstrate that differently expressed CXC chemokines may be important in LUAD. Yu et al. has identified CXCL1 as an unfavorable prognosis factor in NSCLC [38], which highly agrees with our results, and many researchers have believed that CXCL8 was associated with tumor progression, angiogenesis, and relapse for a long time [47]; therefore, our results may pave the way for CXCL8 to be an adverse prognosis factor in LUAD, even NSCLC. These data also suggest that in LUAD, CXCL4/7 may be adverse prognosis factors while CXCL12/14/16 may be favorable prognosis factors. As mentioned before, there are opposite views about the roles of CXCL4, CXCL14, and CXCL16 in NSCLC; our data gives more evidence on these controversies. So far, we do not know enough about CXCL7, and these data offer new information. For CXCL12, it has been believed that CXCL12 and its receptor, CXCR4, formed the CXC12/CXCR4 axis, and the axis was believed to contribute to tumor progression and metastasis in LUAD [21, 48], but in our study, LUAD patients with high

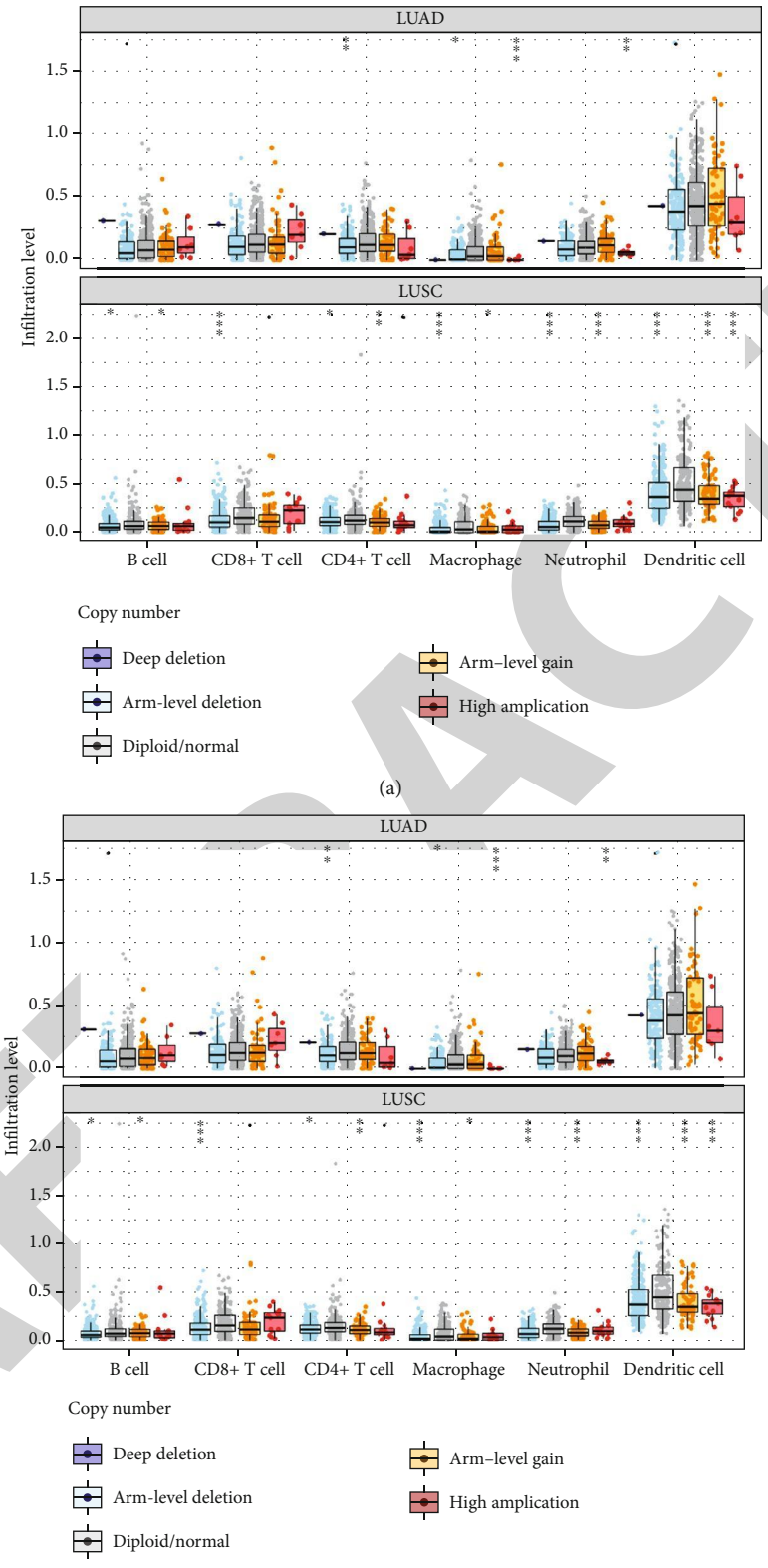
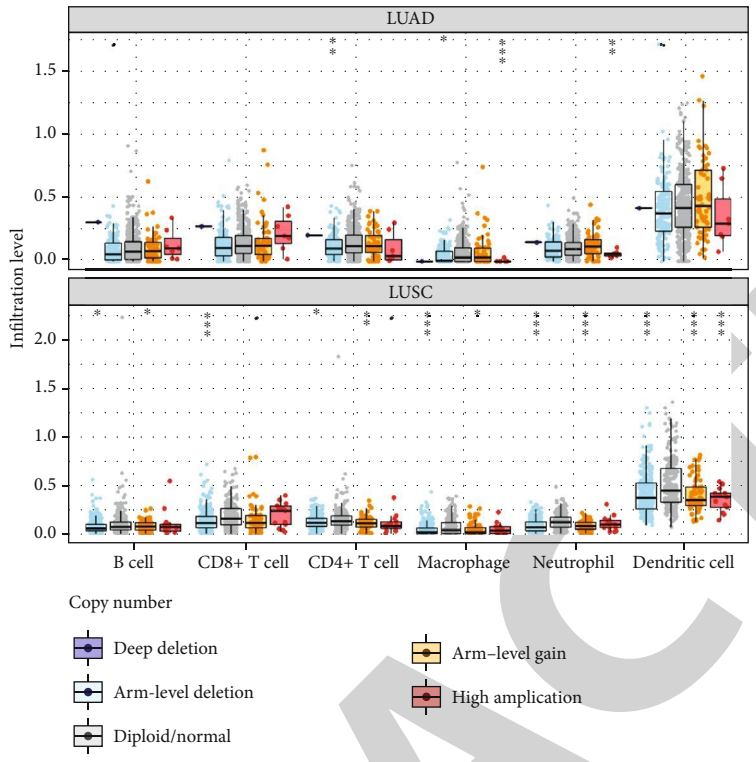
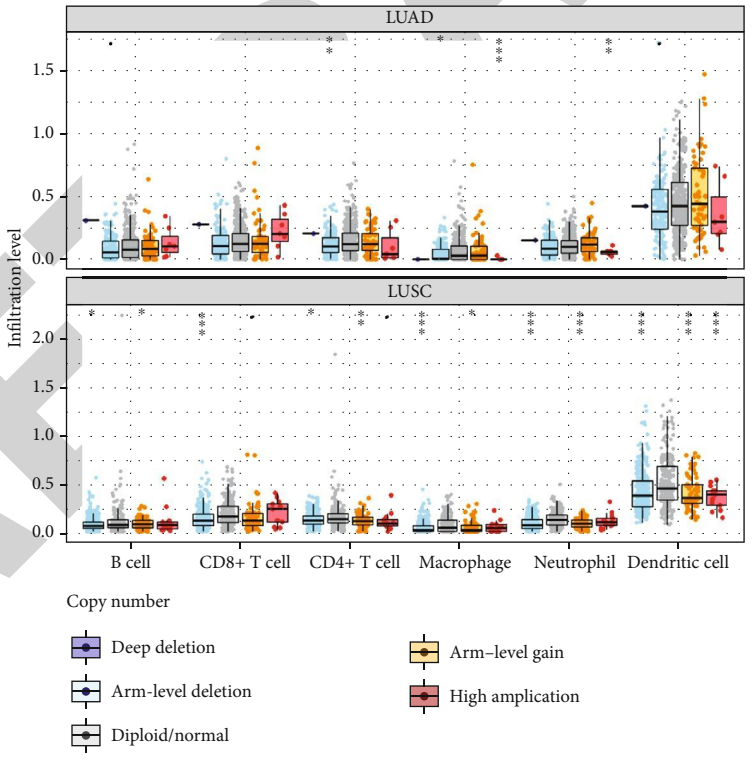


Figure 7: Continued.

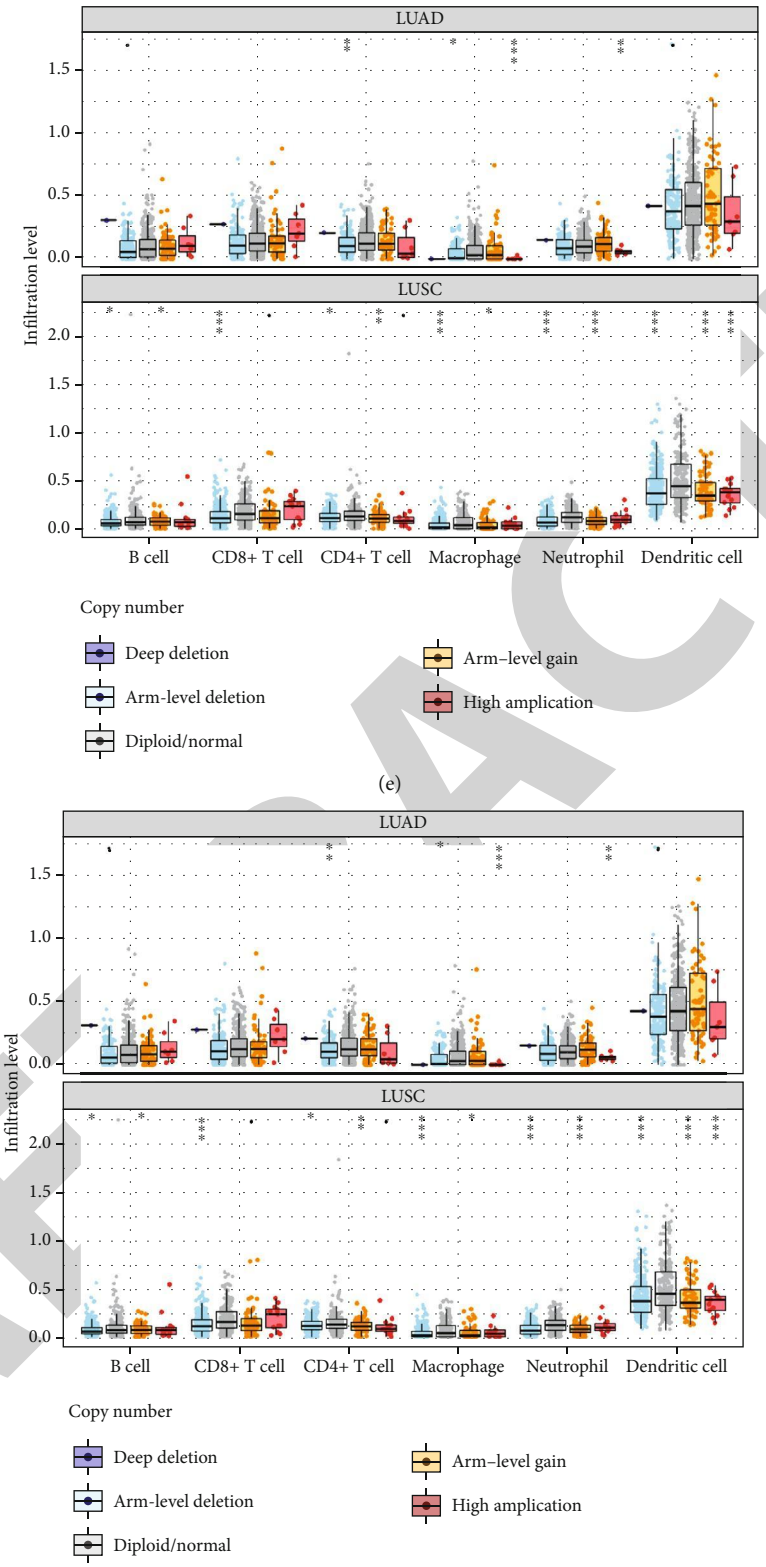


(c)



(d)

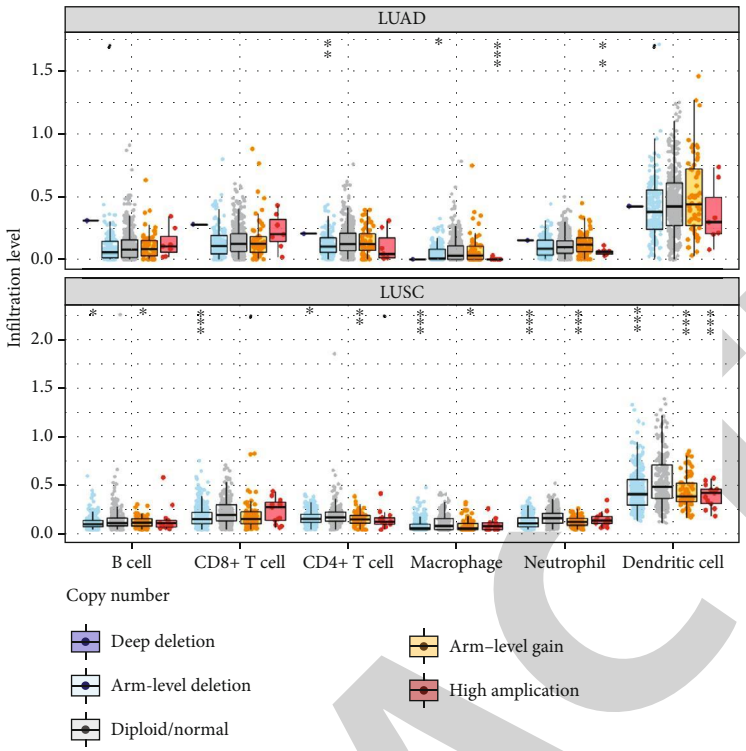
Figure 7: Continued.



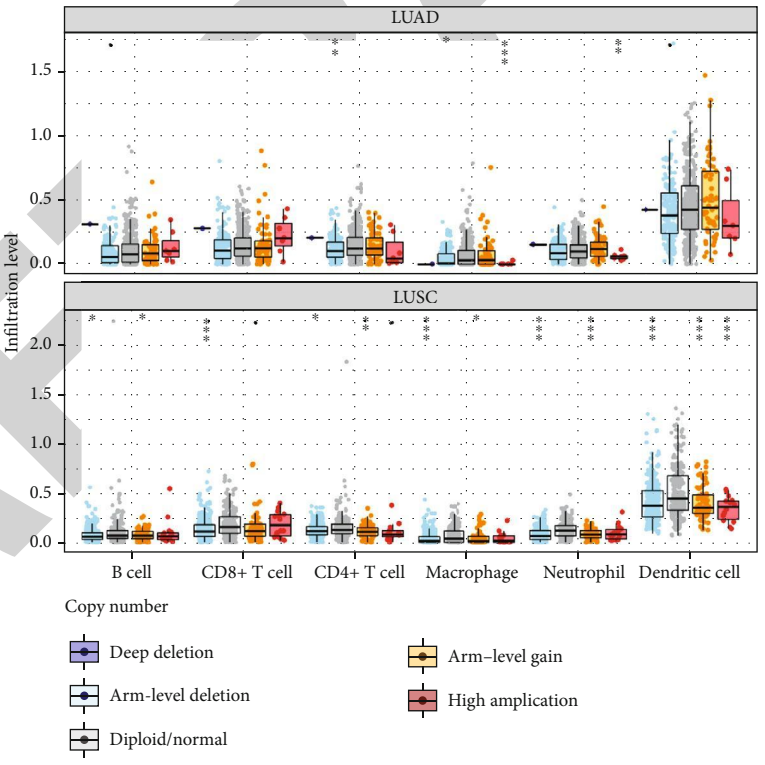
(e)

(f)

Figure 7: Continued.

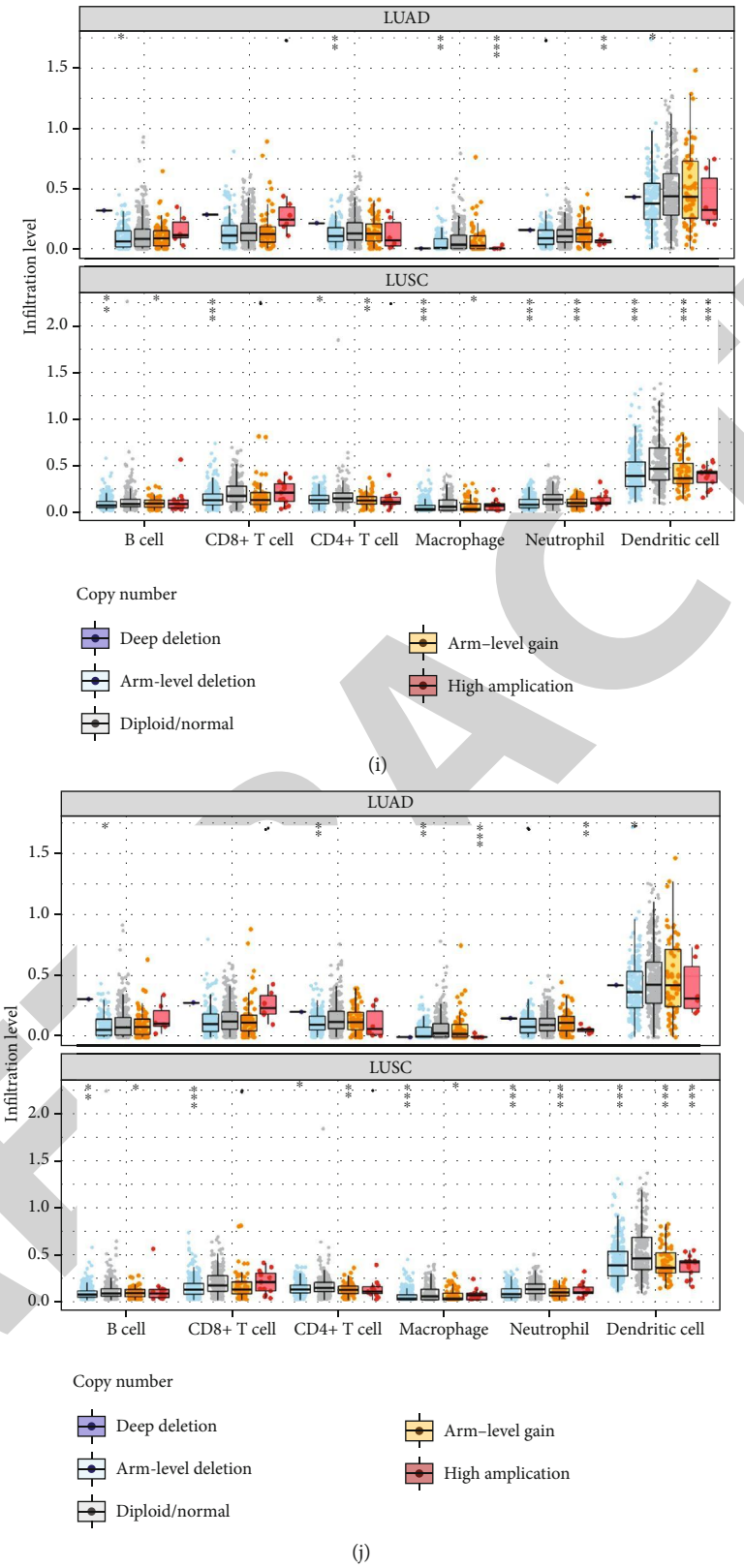


(g)



(h)

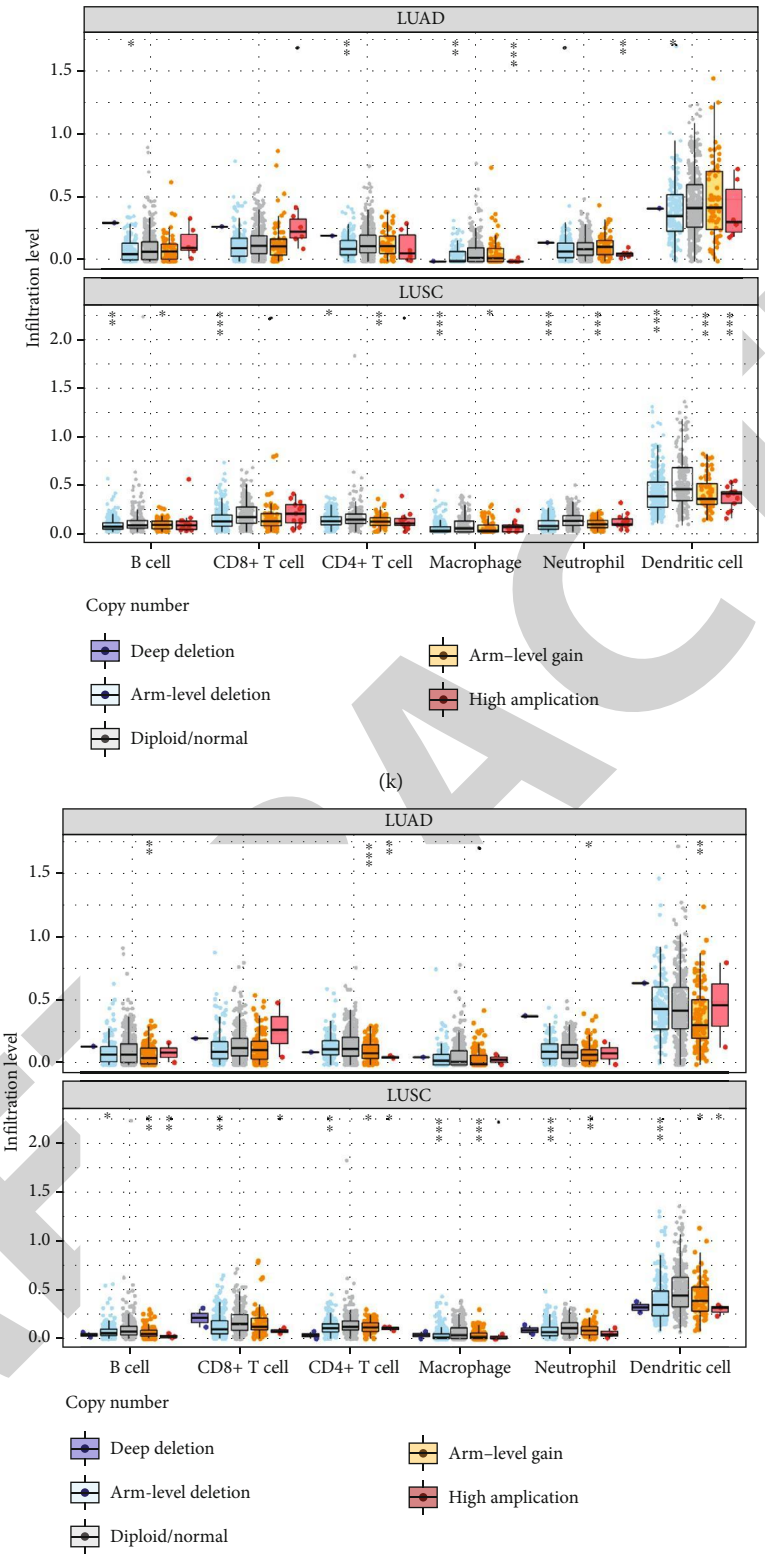
Figure 7: Continued.



(i)

(j)

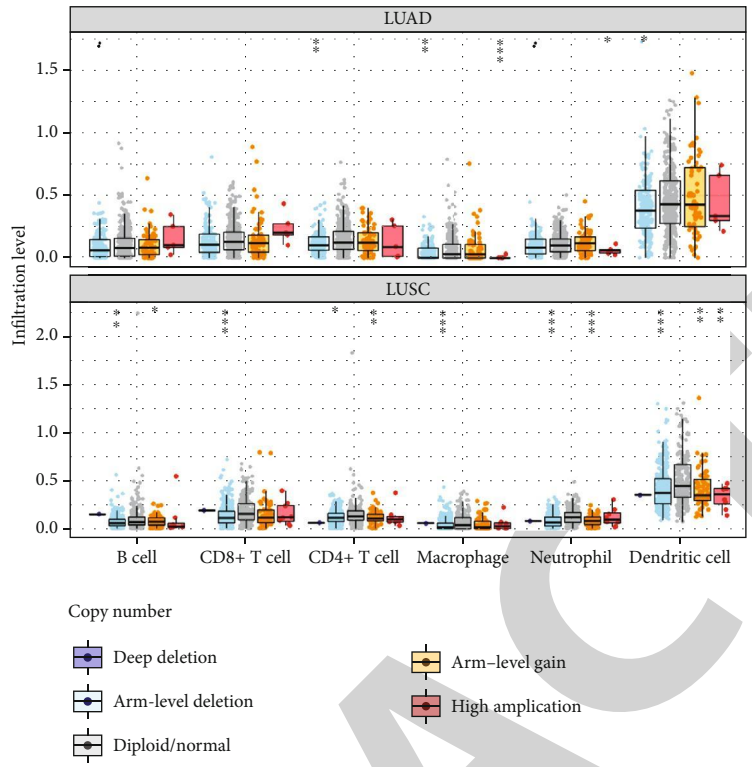
Figure 7: Continued.



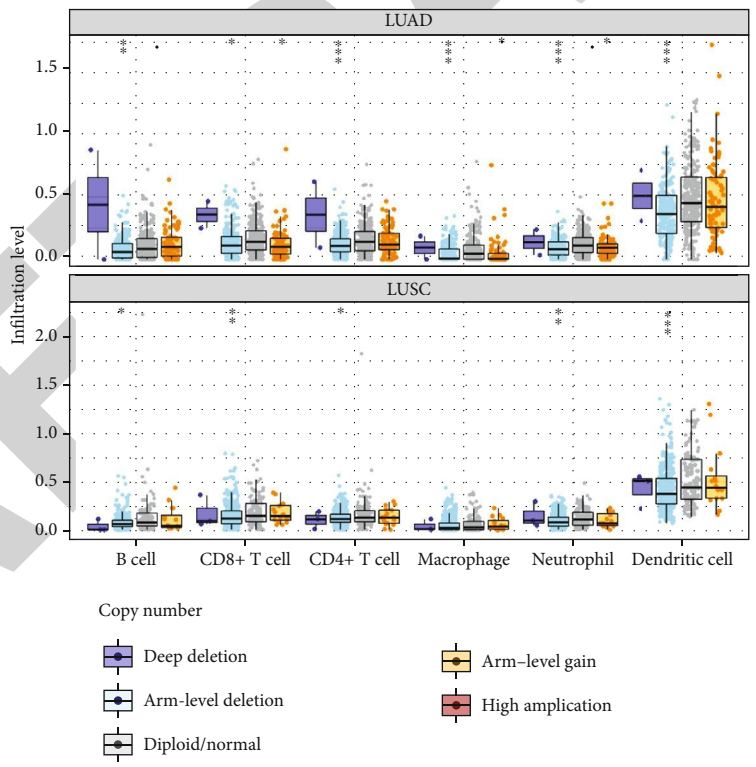
(k)

(l)

Figure 7: Continued.



(m)



(n)

Figure 7: Continued.

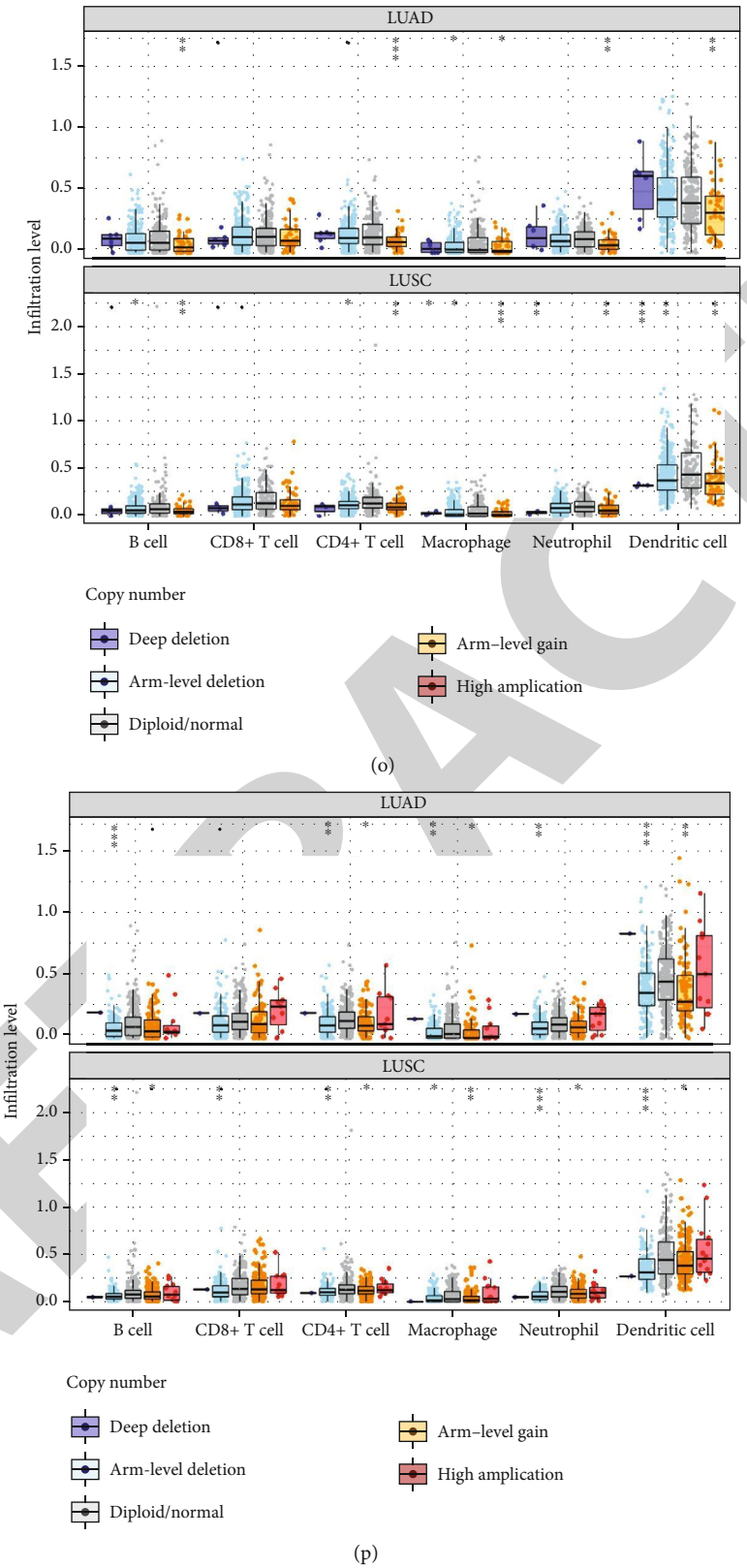


Figure 7: Comparison of tumor infiltration levels among tumors with different SCNAs (somatic copy number alterations) for CXC chemokines in LUAD and LUSC samples. SCNAs are defined by GISTIC2.0. Box plots are presented to show the distributions of each immune subset at each copy number status in LUAD and LUSC patients. The infiltration level for each SCNA category is compared with the normal using a two-sided Wilcoxon rank-sum test: (a) CXCL1, (b) CXCL2, (c) CXCL3, (d) CXCL4, (e) CXCL5, (f) CXCL6, (g) CXCL7, (h) CXCL8, (i) CXCL9, (j) CXCL10, (k) CXCL11, (l) CXCL12, (m) CXCL13, (n) CXCL14, (o) CXCL16, and (p) CXCL7.

CXCL12 expression were associated with better OS significantly; the reasons that lie behind these two contrary findings need further exploration. From the survival map, we found that CXCL2/7/12/16 made significant contribution to the survival in LUSC, and no CXC chemokines were significant in LUAD, which suggest that we expect more research on the roles of CXC chemokines in LUSC.

In lung cancer, since several chemokines were differentially expressed, we explored their molecular traits in LUAD and LUSC. Frequent genetic alterations are found in CXC chemokines, and the elevated/reduced mRNA expression was the type that had the most alteration in both LUAD and LUSC. Tumor initiation progression of lung cancer is complex and intricate, and genetic alteration plays a significant role in this story [49]. From correlation heat maps and gene-to-gene correlation analysis, we found tight corrections among these 16 CXC chemokines in both LUAD and LUSC, implying that these cytokines work synergistically in tumor initiation and progression. In the past 5 years, researchers have identified that CXCL1, CXCL5, and CXCL8 acted as tumor promoters in LUAD or NSCLC, and all of them could be antagonized by DACH1 (the human Dachshund homologue 1) [20, 38, 50], so we speculated if there were any correlations among them. This speculation was partly validated in our study, since we found significant correlations between CXCL1 and CXC8 in both LUAD and LUSC patients. Tight association between CXCL9 and CXCL10 in LUAD was found, and as we mentioned above, combination of CXCL9 or CXCL10 with cisplatin improves therapeutic efficacy in solid tumors [40, 41]; we have reason to believe that there may be a close relation between the CXCL9 and CXCL10. Similarly, Wang et al. found that combination of CXCL10 and CXCL11 led to synergistic antitumor effects [51], and correlations with statistical significance were found between CXCL10 and CXCL11 in both LUAD and LUSC; therefore, we can also speculate the unusual correlation between these two chemokines. As for the driver gene of NSCLC, Luppi et al. and Tsai et al. identified CXCL8 [52] and CXCL12 [53] which worked together with the EGFR gene in the progression of NSCLC, respectively; however, no significant association was12 and EGFR/12 and EGFR (Supplementary Figure 1).

Next, we focused on the function of CXC chemokines using GO enrichment analysis and KEGG pathway enrichment analysis. Not surprisingly, we found that the functions of these genes are mainly related to the chemokine signaling pathway, cytokine-cytokine receptor interactions, and viral protein interaction with cytokine and cytokine receptor in both LUAD and LUSC samples. It has been identified that chemokine signaling pathways are important in the progression, metastasis angiogenesis, senescence, epithelial-mesenchymal transition, and immune evasion of various cancers [54–56]. The roles of cytokine-cytokine receptor interactions are also pivotal in several tumor-associated biological processes [21, 57]. Viral protein can interact with the cytokine and cytokine receptor to affect and even subvert the function of the cytokine network and to regulate the immune response, which is important in cancer [58–60]. Generally, these GO and KEGG results suggest that the CXC chemokines are potential drug therapeutic targets in LUAD and LUSC.

The PCA analysis on samples from LUAD and LUSC samples in TCGA based on their expression of CXC chemokines reflects the ability of all the 16 CXC chemokines in differentiating the LUAD tumor/normal and LUSC tumor/-normal patients; from the 3D figure, we found that CXC chemokines did enable to make these discriminations. We also characterized the transcription factor targets for the 16 CXC chemokines and found that RELA, NFKB1, and SP1 may be the key transcription factors of CXC chemokines. Zeng et al. made the same conclusion without study about this [15], but we included more CXC chemokines in our study. RELA plays a pivotal role in regulating oncogene-induced senescence in preneoplastic lesions [61], is essential to link smoke-induced inflammation with lung cancer growth, and participates in the activation of Wnt/beta-signaling in tumors [62]. NFKB1 is identified as an inhibitor of tumors and inflammatory response; by reducing the aberrant activation of the NF-κB signaling pathway, it negatively regulated the tumorigenesis and progression of several types of cancers [63]. NKKB1 also interfered with diverse complex immunological progresses, associated with many autoinflammatory disorders [64]. SP1 is overexpressed in many cancers and implicated in inflammation, genomic instability, and epigenetic silencing [65]; it has also been reported to be a target in cancer chemotherapy [66]. In NSCL, SP1 could promote cancer progression by interacting with lncRNA LINCo1234 and OTUB1 [67].

The chemokine system can orchestrate the immune cell migration and position them properly in a spatiotemporal manner [68]. In NSCLC, accumulating evidence shows that immune cell infiltration could affect tumor initiation, progression, and prognosis and could be important determinants of response to immunotherapies [69–71]. In our study, we found a significant correlation between the expression of CXC chemokines and the infiltration of the six immune cell types, B cells, $CD8^+$ T cells, $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells, indicating that CXC chemokines not only are prognostic indicators but may also reflect the immune status. Gao et al. reported that CXCL11 could promote $CD8^+$ T cell infiltration in NSCLC [72], which is consistent with our findings in this study in that we found that CXCL11 was significantly correlated with $CD8^+$ T cell infiltration in both LUAD and LUSC.

Moreover, we explored the correlation between aneuploidy and the infiltration of the six immune cell types. In cancer, chromosomal instability underpins much of intratumor heterogeneity and can drive phenotypic adaptation during tumor evolution [73]; meanwhile, tumor aneuploidy is correlated with markers in immune evasion and indicates reduced response to immunotherapy [74]. Copy number alterations have been reported as risk factors, predictors, and even drivers of tumor progress and poor clinical outcome in various cancer types [75–78]. In our study, we found a significant association between the SCNA of CXC chemokines and the infiltration of the six immune cell types and therefore offer more information for the exploration of relationship between SCNA and CXC chemokines in lung cancer.

Generally, differences existed in many aspects between LUAD and LUSC; they have a distinct expression pattern in

several CXC chemokines, but it is hard to find a clear law to distinguish them, which might be due to the fact that our study was purely based on bioinformatics analysis; therefore, it was not effective enough to do the distinction, which also might be due to the fact that both LUAD and LUSC belong to non-small cell lung cancer; therefore, the disparity between them is just not so obvious.

Our study is the first to portray the overall view of CXC chemokines in non-small lung cancer using several public databases and bioinformatics analysis, which is our biggest highlight and main limitation. Therefore, independent cohort and in vitro or in vivo research should be performed to validate our results to a large extent. In conclusion, we hope that our results offer a comprehensive landscape and new insight of CXC chemokines in LUAD and LUSC, thus provide more information for the development of new immunotherapy medicine, facilitate the clinical routine of selecting drugs, and help the clinicians predict the prognosis of patients more accurately.

Data Availability

The datasets analyzed for this study can be found in the ONCOMINE, GEPIA, UALCAN, Kaplan-Meier Plotter, and cBioPortal web resources, and requests to further access to datasets can be directed to tianhepumc@163.com.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors' Contributions

JH and CL designed the study and revised the manuscript. HT, LW, YL, FT, and QX performed data analysis work and aided in writing the manuscript. YW, YZ, TF, SG, and BZ designed the study and assisted in writing the manuscript. All authors read and approved the final manuscript.

Acknowledgments

This work was supported by the Chinese National Natural Science Foundation (81972196), the National Key Basic Research Development Plan (2018YFC1312100, 2016YFC1303200), and the CAMS Innovation Fund for Medical Sciences (CIFMS) (2017-I2M-1-005).

Supplementary Materials

Figure S1: correlation analysis between CXC8/12 and EGFR expression chemokines: LUAD (upper); LUSC (lower). *(Supplementary Materials)*

References

[1] F. R. Hirsch, G. V. Scagliotti, J. L. Mulshine et al., "Lung cancer: current therapies and new targeted treatments," *Lancet*, vol. 389, no. 10066, pp. 299–311, 2017.

[2] R. S. Herbst, D. Morgensztern, and C. Boshoff, "The biology and management of non-small cell lung cancer," *Nature*, vol. 553, no. 7689, pp. 446–454, 2018.

[3] M. A. Socinski, C. Obasaju, D. Gandara et al., "Current and emergent therapy options for advanced squamous cell lung cancer," *Journal of Thoracic Oncology*, vol. 13, no. 2, pp. 165–183, 2018.

[4] The Cancer Genome Atlas Research Network, "Comprehensive genomic characterization of squamous cell lung cancers," *Nature*, vol. 489, no. 7417, pp. 519–525, 2012.

[5] C. Zappa and S. A. Mousa, "Non-small cell lung cancer: current treatment and future advances," *Translational Lung Cancer Research*, vol. 5, no. 3, pp. 288–300, 2016.

[6] K. C. Arbour and G. J. Riely, "Systemic therapy for locally advanced and metastatic non-small cell lung cancer: a review," *JAMA*, vol. 322, no. 8, pp. 764–774, 2019.

[7] M. T. Chow and A. D. Luster, "Chemokines in cancer," *Cancer Immunology Research*, vol. 2, no. 12, pp. 1125–1131, 2014.

[8] N. Nagarsheth, M. S. Wicha, and W. Zou, "Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy," *Nature Reviews. Immunology*, vol. 17, no. 9, pp. 559–572, 2017.

[9] F. Balkwill, "Cancer and the chemokine network," *Nature Reviews. Cancer*, vol. 4, no. 7, pp. 540–550, 2004.

[10] J. Vandercappellen, J. Van Damme, and S. Struyf, "The role of CXC chemokines and their receptors in cancer," *Cancer Letters*, vol. 267, no. 2, pp. 226–244, 2008.

[11] V. Mollica Poeta, M. Massara, A. Capucetti, and R. Bonecchi, "Chemokines and chemokine receptors: new targets for cancer immunotherapy," *Frontiers in Immunology*, vol. 10, p. 379, 2019.

[12] A. Spaks, "Role of CXC group chemokines in lung cancer development and progression," *Journal of Thoracic Disease*, vol. 9, Suppl 3, pp. S164–s171, 2017.

[13] E. Chen, X. Qin, K. Peng et al., "Identification of potential therapeutic targets among CXC chemokines in breast tumor microenvironment using integrative bioinformatics analysis," *Cellular Physiology and Biochemistry*, vol. 45, no. 5, pp. 1731–1746, 2018.

[14] Y. S. Lee, I. Choi, Y. Ning et al., "Interleukin-8 and its receptor CXCR2 in the tumour microenvironment promote colon cancer growth, progression and metastasis," *British Journal of Cancer*, vol. 106, no. 11, pp. 1833–1841, 2012.

[15] Q. Zeng, S. Sun, Y. Li, X. Li, Z. Li, and H. Liang, "Identification of therapeutic targets and prognostic biomarkers among CXC chemokines in the renal cell carcinoma microenvironment," *Frontiers in Oncology*, vol. 9, p. 1555, 2020.

[16] R. M. Strieter, P. J. Polverini, D. A. Arenberg et al., "Role of C-X-C chemokines as regulators of angiogenesis in lung cancer," *Journal of Leukocyte Biology*, vol. 57, no. 5, pp. 752–762, 1995.

[17] R. M. Strieter et al., "CXC chemokines: angiogenesis, immunoangiostasis, and metastases in lung cancer," *Annals of the New York Academy of Sciences*, vol. 1028, no. 1, pp. 351–360, 2004.

[18] S. Acharyya, T. Oskarsson, S. Vanharanta et al., "A CXCL1 paracrine network links cancer chemoresistance and metastasis," *Cell*, vol. 150, no. 1, pp. 165–178, 2012.

[19] K. Wu, S. Yu, Q. Liu, X. Bai, X. Zheng, and K. Wu, "The clinical significance of CXCL5 in non-small cell lung cancer," *Oncotargets and Therapy*, vol. 10, pp. 5561–5573, 2017.

[20] Q. Liu, A. Li, S. Yu et al., "DACH1 antagonizes CXCL8 to repress tumorigenesis of lung adenocarcinoma and improve prognosis," *Journal of Hematology & Oncology*, vol. 11, no. 1, p. 53, 2018.

[21] O. Wald, O. M. Shapira, and U. Izhar, "CXCR4/CXCL12 axis in non small cell lung cancer (NSCLC) pathologic roles and therapeutic potential," *Theranostics*, vol. 3, no. 1, pp. 26–33, 2013.

[22] F. Pucci, S. Rickelt, A. P. Newton et al., "PF4 promotes platelet production and lung cancer growth," *Cell Reports*, vol. 17, no. 7, pp. 1764–1772, 2016.

[23] J. Jian, Y. Pang, H. H. Yan et al., "Platelet factor 4 is produced by subsets of myeloid cells in premetastatic lung and inhibits tumor metastasis," *Oncotarget*, vol. 8, no. 17, pp. 27725–27739, 2017.

[24] C. R. Park, D. J. You, D. K. Kim et al., "CXCL14 enhances proliferation and migration of NCI-H460 human lung cancer cells overexpressing the glycoproteins containing heparan sulfate or sialic acid," *Journal of Cellular Biochemistry*, vol. 114, no. 5, pp. 1084–1096, 2013.

[25] R. Hata, K. Izukuri, Y. Kato et al., "Suppressed rate of carcinogenesis and decreases in tumour volume and lung metastasis in CXCL14/BRAK transgenic mice," *Scientific Reports*, vol. 5, no. 1, pp. 1–10, 2015.

[26] W. Hu, Y. Liu, W. Zhou, L. Si, and L. Ren, "CXCL16 and CXCR6 are coexpressed in human lung cancer in vivo and mediate the invasion of lung cancer cell lines in vitro," *PLoS One*, vol. 9, no. 6, p. e99056, 2014.

[27] S. M. Hald, Y. Kiselev, S. al-Saad et al., "Prognostic impact of CXCL16 and CXCR6 in non-small cell lung cancer: combined high CXCL16 expression in tumor stroma and cancer cells yields improved survival," *BMC Cancer*, vol. 15, no. 1, p. 441, 2015.

[28] D. R. Rhodes, J. Yu, K. Shanker et al., "ONCOMINE: a cancer microarray database and integrated data-mining platform," *Neoplasia*, vol. 6, no. 1, pp. 1–6, 2004.

[29] Z. Tang, C. Li, B. Kang, G. Gao, C. Li, and Z. Zhang, "GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses," *Nucleic Acids Research*, vol. 45, no. W1, pp. W98–w102, 2017.

[30] D. S. Chandrashekar, B. Bashel, S. A. H. Balasubramanya et al., "UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses," *Neoplasia*, vol. 19, no. 8, pp. 649–658, 2017.

[31] J. Gao, B. A. Aksoy, U. Dogrusoz et al., "Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal," *Science Signaling*, vol. 6, no. 269, p. pl1, 2013.

[32] D. Warde-Farley, S. L. Donaldson, O. Comes et al., "The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function," *Nucleic Acids Research*, vol. 38, suppl_2, pp. W214–W220, 2010.

[33] D. Szklarczyk, A. L. Gable, D. Lyon et al., "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Research*, vol. 47, no. D1, pp. D607–d613, 2019.

[34] H. Han, J. W. Cho, S. Lee et al., "TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions," *Nucleic Acids Research*, vol. 46, no. D1, pp. D380–d386, 2018.

[35] T. Li, J. Fan, B. Wang et al., "TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells," *Cancer Research*, vol. 77, no. 21, pp. e108–e110, 2017.

[36] B. Ru, C. N. Wong, Y. Tong et al., "TISIDB: an integrated repository portal for tumor-immune system interactions," *Bioinformatics*, vol. 35, no. 20, pp. 4200–4202, 2019.

[37] B. Győrffy, P. Surowiak, J. Budczies, and A. Lánczky, "Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer," *PLoS One*, vol. 8, no. 12, p. e82241, 2013.

[38] S. Yu, M. Yi, L. Xu, S. Qin, A. Li, and K. Wu, "CXCL1 as an unfavorable prognosis factor negatively regulated by DACH1 in non-small cell lung cancer," *Frontiers in Oncology*, vol. 9, p. 1515, 2019.

[39] M. Põld, L. X. Zhu, S. Sharma et al., "Cyclooxygenase-2-dependent expression of angiogenic CXC chemokines ENA-78/CXC ligand (CXCL) 5 and interleukin-8/CXCL8 in human non-small cell lung cancer," *Cancer Research*, vol. 64, no. 5, pp. 1853–1860, 2004.

[40] G. Li, L. Tian, J. M. Hou et al., "Improved therapeutic effectiveness by combining recombinant CXC chemokine ligand 10 with cisplatin in solid tumors," *Clinical Cancer Research*, vol. 11, no. 11, pp. 4217–4224, 2005.

[41] R. Zhang, L. Tian, L. J. Chen et al., "Combination of MIG (CXCL9) chemokine gene therapy with low-dose cisplatin improves therapeutic efficacy against murine carcinoma," *Gene Therapy*, vol. 13, no. 17, pp. 1263–1271, 2006.

[42] M. F. Sanmamed, J. L. Perez-Gracia, K. A. Schalper et al., "Changes in serum interleukin-8 (IL-8) levels reflect and predict response to anti-PD-1 treatment in melanoma and non-small-cell lung cancer patients," *Annals of Oncology*, vol. 28, no. 8, pp. 1988–1995, 2017.

[43] K. Yamaguchi, K. Ogawa, T. Katsube et al., "Platelet factor 4 gene transfection into tumor cells inhibits angiogenesis, tumor growth and metastasis," *Anticancer Research*, vol. 25, no. 2a, pp. 847–851, 2005.

[44] M. Tessema, D. M. Klinge, C. M. Yingling, K. Do, L. van Neste, and S. A. Belinsky, "Re-expression of CXCL14, a common target for epigenetic silencing in lung cancer, induces tumor necrosis," *Oncogene*, vol. 29, no. 37, pp. 5159–5170, 2010.

[45] Y. Shibata, N. Kobayashi, T. Sato, K. Nakashima, and T. Kaneko, "The clinical significance of CXCL16 in the treatment of advanced non-small cell lung cancer," *Thoracic Cancer*, vol. 11, no. 5, pp. 1258–1264, 2020.

[46] M. Suzuki, S. Mohamed, T. Nakajima et al., "Aberrant methylation of CXCL12 in non-small cell lung cancer is associated with an unfavorable prognosis," *International Journal of Oncology*, vol. 33, no. 1, pp. 113–119, 2008.

[47] A. YUAN, P. C. YANG, C. J. YU et al., "Interleukin-8 messenger ribonucleic acid expression correlates with tumor progression, tumor angiogenesis, patient survival, and timing of relapse in non-small-cell lung cancer," *American Journal of Respiratory and Critical Care Medicine*, vol. 162, no. 5, pp. 1957–1963, 2000.

[48] P. L. Wagner, E. Hyjek, M. F. Vazquez et al., "CXCL12 and CXCR4 in adenocarcinoma of the lung: association with metastasis and survival," *The Journal of Thoracic and Cardiovascular Surgery*, vol. 137, no. 3, pp. 615–621, 2009.

[49] M. T. Agulló-Ortuño, F. López-Ríos, and L. Paz-Ares, "Lung cancer genomic signatures," *Journal of Thoracic Oncology*, vol. 5, no. 10, pp. 1673–1691, 2010.

[50] N. Han, X. Yuan, H. Wu et al., "DACH1 inhibits lung adenocarcinoma invasion and tumor growth by repressing CXCL5 signaling," *Oncotarget*, vol. 6, no. 8, pp. 5877–5888, 2015.

Hindawi
BioMed Research International
Volume 2023, Article ID 9816124, 1 page
https://doi.org/10.1155/2023/9816124

Retraction

Retracted: The Potential Function of Super Enhancers in Human Bone Marrow Mesenchymal Stem Cells during Osteogenic Differentiation

BioMed Research International

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Huang, B. Jia, Q. Wang, N. Wang, and J. Zhao, "The Potential Function of Super Enhancers in Human Bone Marrow Mesenchymal Stem Cells during Osteogenic Differentiation," *BioMed Research International*, vol. 2021, Article ID 6614762, 11 pages, 2021.

Hindawi
BioMed Research International
Volume 2021, Article ID 6614762, 11 pages
https://doi.org/10.1155/2021/6614762

Research Article

The Potential Function of Super Enhancers in Human Bone Marrow Mesenchymal Stem Cells during Osteogenic Differentiation

Zhijie Huang[1], **Bo Jia**[1], **Qian Wang**[2], **Naishun Wang**[3] and **Jianjiang Zhao**[4]

[1]*Departments of Oral and Maxillofacial Surgery, Stomatological Hospital, Southern Medical University (Guangdong Provincial Stomatological Hospital), Guangzhou, Guangdong 510280, China*
[2]*Department of Central Laboratory, Taian Central Hospital, Longtan Road No. 29, Tai'an, 271000 Shandong Province, China*
[3]*Department of Orthopaedics, Taian Hospital of Traditional Chinese Medicine, Yingxuan Street No. 216, Taishan District, Tai'an, 271000 Shandong Province, China*
[4]*Departments of Oral and Maxillofacial Surgery, Shenzhen Stomatological Hospital, Southern Medical University (Shenzhen Dental Medical Center of Guangdong Province), Shenzhen, Guangdong 518001, China*

Correspondence should be addressed to Naishun Wang; 630109356@qq.com and Jianjiang Zhao; zjj2521@sina.com

Received 28 November 2020; Revised 7 January 2021; Accepted 12 January 2021; Published 27 January 2021

Academic Editor: Min Tang

Super enhancers (SEs) are large clusters of transcriptional activity enhancers, which drive and control the expression of cell identity genes, as well as differentiation of specific cell types. SEs have great application potential in pathogenic mechanism studies in developmental biology, cancer, and other diseases. However, the potential function and regulatory mechanism of SEs in the osteogenic differentiation of human bone marrow mesenchymal stem cells (hBMSCs) are unknown. Therefore, this study investigated the potential function of SEs in the osteogenic differentiation of hBMSCs and their target genes. Osteogenesis was induced in three hBMSCs groups for 14 days. Further, ChIP-seq was performed on cells before and after osteogenic differentiation. Two target genes were then selected from cells before and after osteogenic differentiation for RT-qPCR. Finally, the selected SE target genes were analyzed by bioinformatics. In total, 1,680 SEs were identified in hBMSCs. After 14 days of osteogenic induction, only 342 SEs were detected in cells, among which 1,380 unique SEs were detected in hBMSCs, 42 unique SEs were found in cells induced by osteoblast differentiation after 14 days, and 300 SEs were common in both groups. Further, 1,680 genes were found to be regulated by SEs in hBMSCs, including 1,094 genes with protein-coding function and 586 noncoding genes. Additionally, 342 genes were regulated by SEs in cells after 14 days of osteogenic differentiation induction, of which 223 and 119 had protein-coding and noncoding functions, respectively. KEGG analysis of SE target genes before and after osteogenic differentiation showed the TGF-β, PI3K-Akt, and ECM receptor signaling pathways as highly enriched and closely related to osteogenic differentiation. Further, RT-qPCR results of four selected target genes confirmed the sequencing results. Taken together, osteogenic differentiation of hBMSCs involves changes in multiple SEs, which may regulate the osteogenic differentiation of hBMSCs by regulating the expression of target genes.

1. Introduction

The maxillofacial bone defect is often caused by trauma, infection, congenital malformation, or cancer surgery, which not only affects the facial aesthetic, as well as mastication and swallowing functions, but also often results in psychological and social challenges for patients [1]. Due to the complex structure of the maxillofacial bone, repairing defects often pose significant challenges. Currently, bone transplantation is a common clinical practice with autologous bone

transplantation considered the "gold standard" for bone transplantation; however, this procedure has numerous adherent disadvantages, including damage to the donor site, limited available bone, and poor bone quality in patients with osteoporosis and other diseases [2]. Although synthetic bone can be used to replace autologous bone, specific limitations often arise, including immunogenicity, risk of infection, low bone induction, and poor bone integration [3]. Alternatively, bone tissue engineering techniques combine stem cells with biological scaffold materials to form reconstructed tissue with similar functional characteristics to natural bone tissue through specific bone inducing factor stimulation. This strategy is expected to become a novel method for bone repair soon [4] and consists of the following components: seed cells (most critical), growth factors, and biological scaffold materials. Human bone marrow mesenchymal stem cells (hBMSCs) are a diverse group of cells with multidirectional differentiation potential. When tissues or cells in the body are stimulated by the local microenvironment during periods of tissue damage or disease, they become stimulated to release cytokines, growth factors, etc., thus stimulating BMSCs to differentiate into various connective tissue cells, including articular cartilage, bone, muscle, and tendon tissues. In addition, these cells can be amplified and cultured *in vitro* or *in vivo* to ultimately differentiate into osteoblasts, chondrocytes, fat cells, tendon cells, muscle tubes, and nerve cells, among others [5]. Therefore, BMSCs are often employed as seed cells for bone tissue engineering. Hence, further investigation into the currently uncharacterized osteogenic differentiation mechanism of BMSCs may serve to promote the application of bone tissue engineering in clinical practice.

Super enhancers (SEs) are a large cluster of transcriptional enhancers [6] that span, by order of magnitude, more DNA than typical enhancers (TEs). In addition, many cell-specific master transcription factor-binding motifs are enriched in SEs, enabling SEs to bind to a larger number of transcription factors, cofactors, and transcriptional activity-related histone modification markers with higher density. Further, these characteristics provide SEs with stronger transcriptional activation activity and higher interference sensitivity, which can drive and control the expression of cell identity genes, as well as regulate specific expression patterns in different cell types, thereby offering great application potential in developmental biology research related to the pathogenic mechanisms in cancer and other diseases [6, 7]. However, the potential function and molecular mechanism of SEs in osteogenic differentiation of hBMSCs are not characterized yet.

In this study, osteogenic differentiation of three hBMSCs groups was carried out, and ChIP-seq was performed on cells before and after osteogenic differentiation in order to screen and identify SEs related to osteogenic differentiation of hBMSCs. Further, bioinformatics analysis was conducted for their target genes to predict their potential functions in osteogenic differentiation, laying the foundation for further investigations on the osteogenic differentiation of hBMSCs, and providing new regulatory targets for the construction of novel tissue-engineered bone grafts.

2. Materials and Methods

2.1. Culture and Osteogenic Differentiation of hBMSCs. Frozen primary hBMSCs (Saliai, Guangzhou, China) were rapidly placed in a water bath at 37°C until thawed. The primary hBMSCs were inoculated at a density of $5 \times 10^3/cm^2$ in a T-25 culture flask and cultured at 37°C in 5% CO_2. Once the cells reached 90–95% confluence, they were transferred to and cultured on 6-well plates. The cells were then divided into two groups: culture without treatment (D0 group) and culture with osteoblastic differentiation fluid treatment (Cyagen, Suzhou, China, HUXMA-90021), according to the manufacturer's instructions (D14 group). The culture medium was replaced every three days, and cells were collected after 14 days of culture.

2.2. Alizarin Red Staining. The culture medium for the D0 and D14 groups was absorbed and washed with PBS three times. Cells were fixed with 4% paraformaldehyde at 25°C for 30 min and rinsed with distilled water thrice. The cells were then stained with 0.1% alizarin red for 30 min at 37°C. Subsequently, the cells were washed with distilled water, dried, and observed under a microscope (Aote, Chongqing, China, BDS200-FL). Finally, the images were collected and analyzed.

2.3. RNA Extraction and RT-qPCR. Total RNA was extracted from D0 and D14 cells using the TRIzol kit (MRC, Ohio, USA, TR118-500) according to the manufacturer's instructions. The quality of the total RNA extracted was measured. The PCR reaction system was performed according to protocols provided with the qPCR detection kit (Vazyme, Nanjing, China, Q341-03). The initial activation of the qPCR reaction was conducted at 95°C for 120 s, and subsequently, 40 cycles were performed at the following settings: denaturation, 95°C, 30 s; annealing, 60°C, 30 s; extension, 72°C, 45 s. A final extension was performed at 75°C for 4 min and repeated thrice. *GAPDH* was used as an internal reference gene. After the qPCR reaction, Ct values of the target gene and internal reference gene were obtained through the real-time fluorescence curve. Finally, mRNA expression levels in the D0 and D14 groups were analyzed by $2^{-\Delta\Delta Ct}$. The primers for *RUNX2*, *OPN*, *COL1A1*, and *GAPDH* are provided in Table 1.

2.4. Western Blot. After osteogenic induction, cells were collected and lysed with RIPA buffer (Beyotime, Shanghai, China) to obtain total protein content. Protein levels were quantified using the Bradford kit (Pierce, Illinois, USA) and electrophoresis on a 10% SDS-polyacrylamide gradient gel. The proteins were then transferred to a polyvinylidene fluoride membrane (Millipore, Massachusetts, USA) and incubated in 5% bovine serum albumin for one hour. Further, protein expression was detected using anti-RUNX2 (CST, Massachusetts, USA, diluted 1 : 2,000), anti-OPN (Abcam, Cambridge, UK, diluted 1 : 2,000), and anti-COL1A1 (Abcam, diluted 1 : 2,000) antibodies. Antibody diluent was added to the membrane and incubated overnight at 4°C. Subsequently, the membranes were incubated at room temperature with secondary enzyme-labeled antibodies

Table 1: Primer sequences used in quantitative polymerase chain reaction validation. Osteogenesis-specific transcription factors.

Primer	Sequence (5'⟶3')
OPN q forward	AGACCCTTCCAAGTAAGTCC
OPN q reverse	TCATCTACATCATCAGAGTCGT
COL1A1 q forward	GCCAAGACGAAGACATCCCA
COL1A1 q reverse	GGCAGTTCTTGGTCTCGTCA
RUNX2 q forward	TGGTTACTGTCATGGCGGGTA
RUNX2 q reverse	TCTCAGATCGTTGAACCTTGCTA

(Abcam, ab6721, 1 : 2,000 dilution) for one hour. HRP signals were detected using a chemiluminescence reagent (Millipore). Protein expression was quantified using the BCA protein quantification method and was normalized against GAPDH (Abcam, diluted 1 : 500). The ImageJ software (National Institutes of Health, Bethesda, MD, USA) was used to process the image.

2.5. Chromatin Immunoprecipitation-Sequencing (ChIP-seq). In brief, 8×10^6 fresh cells were extracted and suspended in 10 mL medium. Chromatin immunoprecipitation was performed using the EpiTM chromatin immunoprecipitation kit (Epibiotek, Guangzhou, China, cat. no. R1802) according to the manufacturer's instructions. Cells were harvested and cross-linked with 1% formaldehyde for 10 min, followed by quenching with 0.125 M glycine for 5 min. Cells were lysed in 1 mL lysis buffer and rotated for 30 min at 4°C. Further, cell lysates were centrifuged at $2{,}400 \times g$ for 10 min at 4°C to isolate the nuclei. The nuclei were suspended in digestion buffer and subjected to enzymatic shearing to generate 200–500 bp chromatin fragments at 37°C for 10–15 min. Next, the shearing reaction was stopped, and the fragmented chromatin was centrifuged at $18{,}000 \times g$ for 10 min at 4°C. The supernatant was transferred to a new Eppendorf tube, and the chip reaction mix was added containing protein A/G magnetic beads, chip IP buffer, antibodies, and protease inhibitor cocktail and rotated at 4°C overnight. Protein A/G magnetic beads were then washed. The chromatin was eluted in the reverse cross-linking buffer, followed by incubation at 65°C for 3 h. Subsequently, the ChIP DNA was treated with RNase A and protease K at 37°C for 30 min and purified using Phenol chloroform. ChIP DNA was processed for library generation using the QIAseq Ultralow Input Library Kit (QIAGEN, Hilden, Germany) following the manufacturer's protocol.

2.6. Identification of Super Enhancers. ChIP-seq reads were aligned to the human reference genome by Bowtie Aligner [8]. ChIP-seq reads aligning to the region were extended by 200 bp, and the density of reads per base pair (bp) was calculated. The density of reads in each region was normalized to the total number of million mapped reads, producing read density in units of reads per million mapped reads per base pair (rpm/bp). ChIP-seq peaks were identified using MACS (Model-Based Analysis of ChIP-seq) by considering reads mapped only once at a given locus. Enhancers were defined as regions of ChIP-seq enrichment for H3K27ac. Closely spaced peaks within a range of 12.5 kb were merged together, followed by measurement of input and H3K27ac signals. These merged peaks were ranked by H3K27ac signal and then classified as SEs or TEs [6].

Table 2: Primer sequences used in quantitative polymerase chain reaction validation. SE target genes.

Primer	Sequence (5'⟶3')
LINC00460 q forward	TGGATGAACCACCATTGCCA
LINC00460 q reverse	TCGGGGTGACTTCAGAATGC
LINC00884 q forward	TGATTTTCGCTGCGTCCTCT
LINC00884 q reverse	GGCGGCTGAGTTCTGTAAGT
KCNMA1-AS1 q forward	ATGTTCTTCCCAACCTGCCAA
KCNMA1-AS1 q reverse	CTCAAACACGAGCGGACCAG
ASTN2-AS1 q forward	AAGTTGGCACACAGCACTCT
ASTN2-AS1 q reverse	ATTTCTACCTCCCGCTTGCC

2.7. Functional Enrichment Analysis. According to the location of the SE, target genes were predicted, and the functions and pathways of these genes were analyzed through multiple databases. GO (Gene Ontology) analysis was conducted using the DAVID database (Database for Annotation, Visualization, and Integrated Discovery; http://david.abcc.ncifcrf.gov/), in which the molecular function of target genes can be determined. The default algorithm of the database (Fisher's exact test) was used to obtain P values related to GO term enrichment, and $P < 0.05$ was considered statistically significant. KEGG (Kyoto Encyclopedia of Gene and Pathway Analysis) analysis (KEGG biological pathway database; http://www.genome.jp/) was used to identify biological pathways associated with target genes. Using the default algorithm in the database (Fisher's exact test), P values related to the enrichment of SE target genes in each biological pathway were obtained. $P < 0.05$ was considered statistically significant.

2.8. Validation of SE Target Gene Expression. Two target genes were selected from the D0 and D14 groups to verify the accuracy of sequencing results. The RT-qPCR method has been described previously. The target genes and their primers are provided in Table 2.

2.9. Statistical Analysis. The SPSS13.0 software was used for statistical analysis, and measurement data were expressed as mean ± standard deviation. Two independent sample Student's t-tests were conducted. In addition, one-way ANOVA and post hoc Student-Newman-Keuls test were used for comparison between groups. $P < 0.05$ was considered statistically significant.

3. Results

3.1. Osteogenic Differentiation of hBMSCs. After culturing primary hBMSCs to P2 generation, the cells exhibited long spindle morphology and unique growth characteristics. After

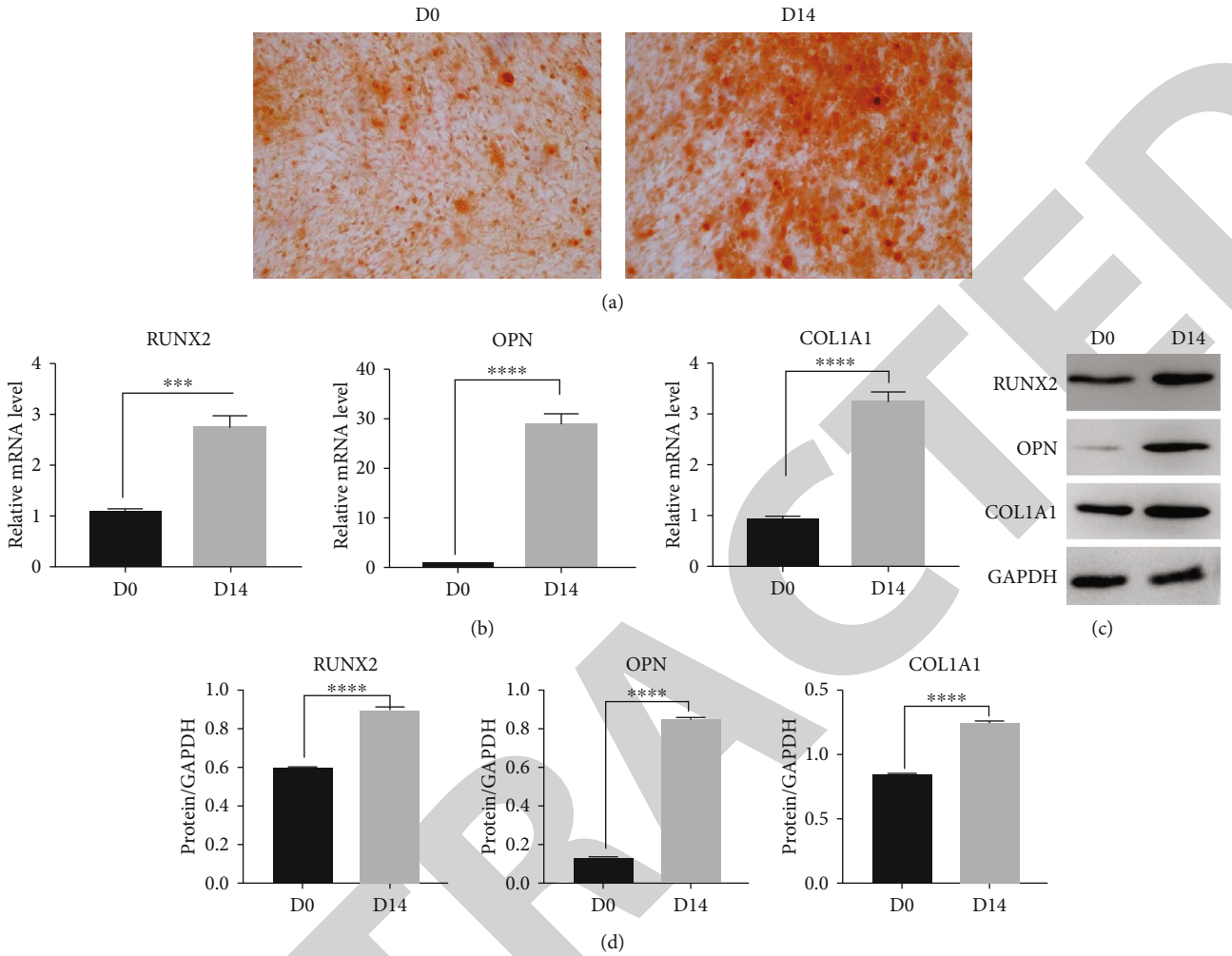


Figure 1: Identification of osteogenic differentiation of hBMSCs. (a) Alizarin red staining (magnification ×100). (b) Relative mRNA expression levels of osteogenesis-specific transcription factors RUNX2, OPN, and COL1A1, as detected by RT-qPCR. (c, d) Western blot (WB; c) and corresponding quantification (d) of osteogenesis-specific transcription factors RUNX2, OPN, and COL1A1. GAPDH was used as a loading control. ∗∗ indicates $P < 0.01$; ∗∗∗ indicates $P < 0.001$; ∗∗∗∗ indicates $P < 0.0001$.

14 days of osteogenic induction and culture, the cells began to aggregate and showed irregular morphology, including primarily polygonal or cubic shapes.

After 14 days of osteogenic induction and differentiation, alizarin red staining showed red calcium deposition nodules (Figure 1(a)). RNA and protein were extracted from hBMSCs before (D0) and 14 days after (D14) osteogenic-induced differentiation. Further, mRNA and protein expression levels of osteogenic-specific transcription factors (RUNX2, OPN, and COL1A1) were detected by qPCR and Western blot, respectively, for which RUNX2, OPN, and COL1A1 mRNA (Table 3, Figure 1(b)) and protein (Table 4, Figures 1(c) and 1(d)) expression levels were significantly higher after 14 days of osteogenic differentiation induction ($P < 0.05$).

3.2. Screening of SEs in hBMSCs before and after Osteogenic Differentiation by Chromatin Immunoprecipitation-Sequencing (ChIP-seq). In total, 31,470 enhancers were identified from the D0 group by the ChIP-seq method, including 1,680

Table 3: The RUNX2, OPN, and COL1A1 gene expression change during osteogenesis of hBMSCs.

Osteogenesis-specific transcription factors	D0	D14
RUNX2	1.07 ± 0.07	2.73 ± 0.20
OPN	0.92 ± 0.07	28.77 ± 1.89
COL1A1	0.92 ± 0.06	3.23 ± 0.17

Table 4: The RUNX2, OPN, and COL1A1 protein expression change during osteogenesis of hBMSCs.

Osteogenesis-specific transcription factors	D0	D14
RUNX2	0.59 ± 0.01	0.89 ± 0.01
OPN	0.12 ± 0.01	0.85 ± 0.01
COL1A1	0.84 ± 0.01	1.24 ± 0.01

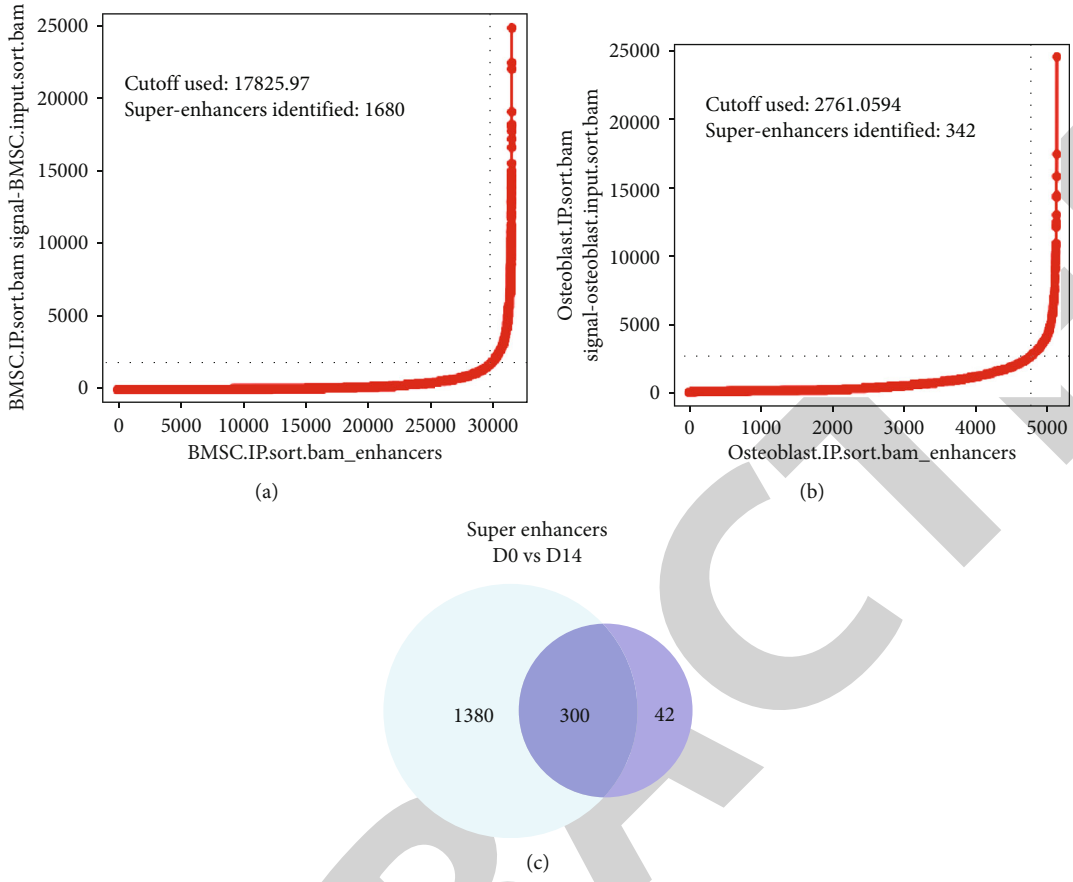


Figure 2: Super-enhancers in BMSCs (D0 group) and osteoblasts (D14 group). (a) Distribution of ChIP-seq density across enhancers with 1680 SEs and a cutoff value of 17825.97 in the D0 group (ranked by H3K27ac signal. (b) Distribution of ChIP-seq density across enhancers with 342 SEs and a cutoff value of 2761.06 in the D14 group (ranked by H3K27ac signal. (c) The comparison of D0 group and D14 group super-enhancers.

Table 5: Detailed information of super enhancers. D0 group.

ID	Chrom	Start	End	Size	Target gene	Gene type
4_BMSC_peak_18504_lociStitched	Chr11	76073282	76088184	14902	WNT11	Protein-coding
25_BMSC_peak_100453_lociStitched	Chr7	82131135	82211994	80859	LOC101927356	ncRNA
3_BMSC_peak_80926_lociStitched	Chr5	31360700	31378910	18210	DROSHA	Protein-coding
8_BMSC_peak_101958_lociStitched	Chr7	112973906	113001251	27345	RAT17	ncRNA

Table 6: Detailed information of super enhancers. D14 group.

ID	Chrom	Start	End	Size	Target gene	Gene type
2_osteoblast_peak_1318_lociStitched	Chr10	90674066	90676304	2238	LOC101926942	ncRNA
2_osteoblast_peak_3420_lociStitched	Chr15	89119260	89121050	1790	ABHD2	Protein-coding
2_osteoblast_peak_6765_lociStitched	Chr4	90317314	90319003	1689	CCSER1	Protein-coding
2_osteoblast_peak_8606_lociStitched	Chr7	18638278	18640106	1828	HDAC9	Protein-coding

SEs with a cutoff value of 17,825.97 (Figure 2(a), and Table 5). Detailed information is provided in Table 3. In the D14 group, 5,119 enhancers were identified, of which 342 were super-enhancers, with a cutoff value of 2,761.05 (Figure 2(b) and Table 6), among which 300 SEs were shared between the D0 and D14 groups, 1,380 SEs were unique to D0, and 42 SEs were unique to D14. Additionally, the target genes of SEs were analyzed, identifying

Table 7: The super enhancer target gene expression changes during osteogenesis of hBMSCs.

Super enhancer target gene	D0	D14
LINC00460	1.08 ± 0.06	0.54 ± 0.04
LINC00844	0.97 ± 0.02	0.11 ± 0.01
KCNMA1-AS1	1.04 ± 0.03	5.72 ± 0.52
ASTN2-AS1	1.12 ± 0.08	1.44 ± 0.12

1,680 SE target genes in the D0 group, including 1,094 protein-coding and 586 noncoding genes. Further, 342 SE target genes in the D14 group were identified, among which 223 were protein-coding, and 119 were noncoding genes. To validate the sequencing results, two unique SE target genes were selected randomly from the D0 (LINC00460 and LINC00844) and D14 (KCNMA1-AS1 and ASTN2-AS1) groups for RT-qPCR verification. Results showed that relative mRNA expression levels of LINC00460 and LINC00844 after osteogenesis differentiation were significantly decreased compared to expression before differentiation (Table 7, Figure 3(a)), whereas KCNMA1-AS1 and ASTN2-AS1 were significantly increased (Table 7, Figure 3(b)).

3.3. GO and KEGG Biological Pathway Enrichment Analyses. GO analysis was conducted for SE target genes in the D0 and D14 groups, at $P < 0.05$ and an enrichment score of -log (P value). For the D0 group, the top three GO terms in BP (biological process) were extracellular matrix organization, axon guidance, and cell adhesion, with 49, 61, and 67 genes, respectively (Figure 4(a)). In CC (cellular component), the top three GO terms were focal adhesion, proteinaceous extracellular matrix, and cytoplasm, with 56, 46, and 398 genes, respectively (Figure 4(a)). For MF (molecular function), the top three enrichment scores were for protein binding, collagen binding, and SH3 domain binding, with 497, 16, and 22 genes, respectively (Figure 4(a)). In the D14 group, the top GO enrichment terms in BP were regulation of interleukin-1β secretion, negative regulation of retinal ganglion cell axon guidance, and extracellular matrix disassembly, with 2, 2, and 7 genes, respectively (Figure 4(b)). In CC, the top three GO terms were basement membrane, ruffle, and focal adhesion, with 6, 6, and 12 genes, respectively (Figure 4(b)). In MF, the top three enrichment scores were for proteoglycan binding, peptide hormone binding, and methenyltetrahydrofolate cyclohydrolase activity, with 497, 16, and 22 genes, respectively (Figure 4(b)).

Through KEGG analysis of SE-related target genes in the D0 and D14 groups, at $P < 0.05$ and enrichment score of -log (P value), the top three pathways in the D0 group were determined to be proteoglycans in cancer, pathways in cancer, and TGF-β signaling pathway, which are closely related to osteogenic differentiation pathways (Figure 5(a)). In the D14 group, the top three pathways were ECM receptor interaction, proteoglycans in cancer, and PI3K-Akt signaling pathway, for which the ECM receptor interaction and PI3K-Akt signaling pathway were closely related to the osteogenic differentiation pathway (Figure 5(b)).

4. Discussion

Human bone marrow mesenchymal stem cells have the potential to differentiate into bone, cartilage, fat, nerve, and other tissues. They can be extracted and separated from the autologous bone marrow of patients. As one of the ideal bone tissue engineering seed cells, they have been widely used in the construction of tissue engineering bone [9–11]. Compared with other stem cells, hBMSCs have shown greater advantages in osteogenesis differentiation, self-renewal ability, and secretion of growth factors, cytokines, and chemokines bioactive factors to regulate immune responses and enhance tissue regeneration [12–15]. Therefore, studying the osteogenic differentiation mechanism of hBMSCs will provide an important basis for the clinical application of bone tissue engineering.

Enhancers are DNA sequences that can significantly increase the frequency of gene transcription and contain multiple transcription factor binding sites. Cells control the transcription process primarily through binding transcription factors that control cell type-specific gene expression patterns [16–18]. SEs were proposed by Young et al. in 2013. They identified enhancers on mouse embryonic stem cells through genomic analysis of master transcription factors (mTFs) and mediator binding sites through ChIP-seq technology. Enhancers are generally divided into typical enhancers (TEs) and SEs based on ChIP-seq signal levels. SEs are rich in cell-specific key transcription factor binding motifs. These structural features enable it to bind to more transcription factors, cofactors, and transcriptional activity-related histone modification markers with higher density. Taken together, these characteristics give SEs stronger transcriptional activation ability and higher interference sensitivity, which drive and control the expression of cell identity genes and determine the specific expression pattern in different cell types, showing great application potential in developmental biology and studies related to pathogenic mechanisms in cancer and other diseases [14, 15].

SEs can regulate cell development and differentiation by driving the expression of cell identity genes. For example, SEs of progenitor B cells can bind to the transcription factor PU.1 at high density and activate transcription of Foxo1 and Inpp5d genes. Foxo1 is involved in determining the phenotype of B lymphocytes [19], whereas Inpp5d is involved in responses to B-cell antigen receptor signals [20]. Muscle tubule cell SE sequences can be combined with transcription factors MyoD, key for promoting transcription of the muscle cytoskeletal protein Myoglobin [21]. In T helper cells, SEs binds to the transcription factor T-bet to regulate Tcf7 gene, which is involved in the formation of T cells in the hematopoietic process [22]. SEs of macrophages are bound to transcription factor C/EBP-α, in which its downstream genes encode extracellular matrix glycoproteins, and thus participate in the recognition and clearance of apoptotic cells by macrophages [23]. In the process of preadipocyte early fat

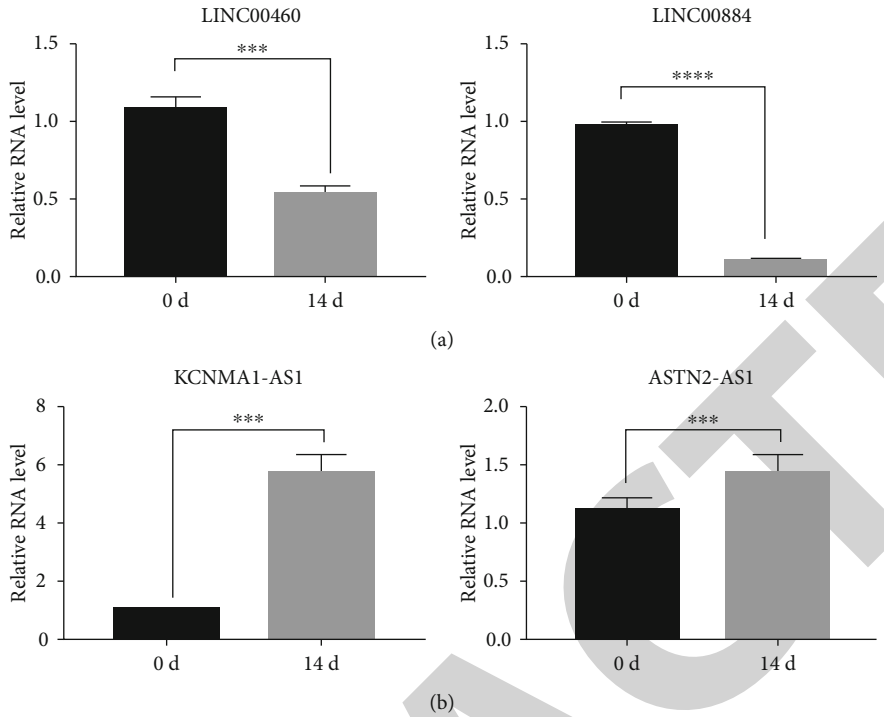


Figure 3: RT-qPCR validation of SE target genes during BMSC osteogenic differentiation. (a) SE target genes in the D0 group. (b) SE target genes in the D14 group. *** indicates $P < 0.001$; **** indicates $P < 0.0001$.

synthesis, Med1 and other important transcription factors in highly integrated hot areas (hotspots) constitute the key components of SEs, usually by interfering with hotspots and SE regional combination of transcription factors to control preadipocyte differentiation [24].

To our knowledge, this study is the first to screen SEs involved in osteogenic differentiation of hBMSCs by ChIP-seq. To validate the sequencing results, four SE-related target genes were selected from cells before and after osteogenic differentiation for RT-qPCR verification. Although a large number of studies have shown that SEs can participate in cell development and differentiation, few studies have focused on osteogenic differentiation of hBMSCs and how SEs regulates osteogenic differentiation through their target genes. Previous research has confirmed that the Wnt/β-catenin, p38 mitogen-activated protein kinase (MAPK), Notch signal, nuclear factor-kappa B (NF-κB) signaling, BMP signaling, fibroblast growth factor (FGF) signaling, and ECM signaling pathways play important roles in the osteogenic differentiation process [25, 26]. For example, the Wnt/β-catenin pathway is a key regulator of bone homeostasis and one of the most common targets for interventional treatment of fracture patients [27]. It has long been accepted that β-catenin can activate osteogenic differentiation and osteogenesis, stimulating osteoblasts to produce osteogenic-specific transcription factors such as Runx2 and Osterix [28]. BMPs belong to the TGF-β superfamily, and numerous studies have shown that BMPs bind to receptor complexes consisting of heterotetramers of type I and type II Ser/Thr kinase receptors and activate SMADs (SMAD1, SMAD5, or SMAD8) through site-specific phosphorylation. Subsequently, phosphorylated SMADs form complexes with SMAD4 and enter the nucleus to regulate gene expression, thereby regulating osteogenic differentiation [29]. FGFs are a large family of proteins, most of which act by binding to Tyr kinase FGF receptors on the cell surface, triggering phosphorylation of a series of signaling proteins and activation of multiple signaling modules, including MAPK, phosphoinositide 3-kinase (PI3K), STAT1, and PKC [30]. These pathways are classical osteogenesis signaling pathways through which many target genes participate in the regulation of osteogenesis. In this study, bioinformatics was performed to conduct GO and KEGG analyses on SE target genes to determine the biological functions and pathways related to osteogenic differentiation. The top pathways with the highest enrichment scores in KEGG analysis included TGF-β, PI3K-Akt, and ECM receptor signaling pathways, which were closely related to osteogenic differentiation. The TGF-β signaling pathway is involved in many cellular processes and has multiple regulatory mechanisms. In the process of osteogenic differentiation of mesenchymal stem cells, TGF-β can interact with SMAD and BMP through the TGF-β type I receptor, thereby regulating osteogenic differentiation. Further, Li et al. [31] found that upregulation of miRNA-10b could promote osteogenic differentiation of mesenchymal stem cells, while downregulation could inhibit osteogenic differentiation. Target prediction and dual-luciferase reporter gene analysis identified SMAD2 as a potential target for miR-10b, demonstrating that it may regulate the osteogenic differentiation process through TGF-β signaling. Huang et al. [32] found that

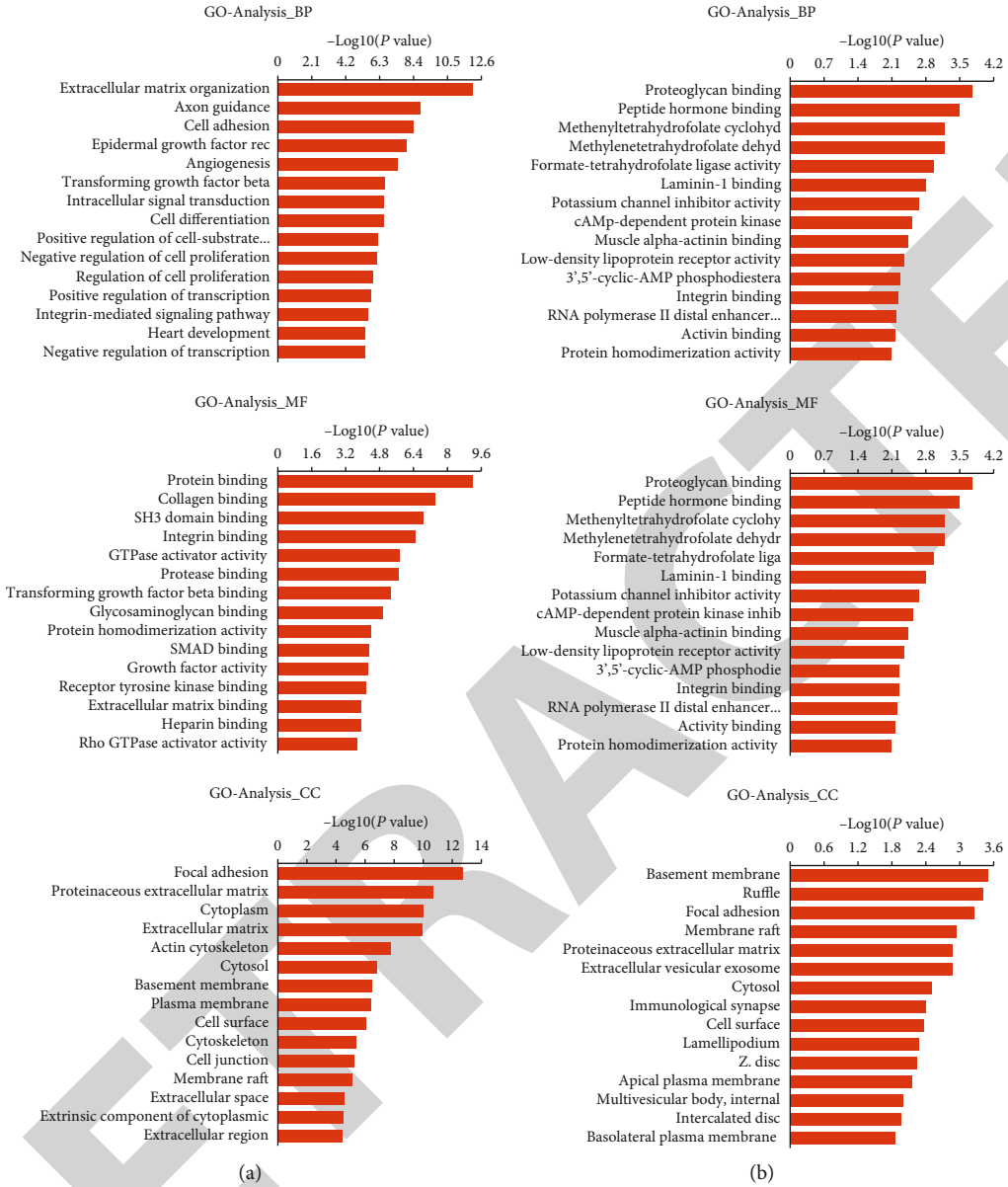


Figure 4: Gene ontology analyses of SE target genes. (a) Top 15 enrichment scores for biological processes, cellular components, and molecular functions among SE target genes of the D0 group. (b) Top 15 enrichment scores for biological processes, cellular components, and molecular functions among SE target genes of the D14 group.

LncRNA H19 regulates osteogenic differentiation of BMSCs through the TGF-β signaling pathway. Taken together, all these studies suggest that TGF-β signaling pathways play an important role in osteogenic differentiation in mesenchymal stem cells. Therefore, SEs may regulate the osteogenic differentiation of BMSCs by regulating the expression of target genes and, in turn, the TGF-β signaling pathway.

It is worthwhile to acknowledge the significance of the present study and its implications for future research. First, the epigenetic mechanisms that we have researched from the aspect of SE-LncRNAs provided a theoretical basis and a better understanding of the researchers in the area of BMSCs-osteodifferentiation. Second, the SE-LncRNAs biomarkers identified could provide the research directions for future experimental studies regarding the epigenetic mechanisms of BMSCs. Second, since many previous studies have shown that the epigenetic modification of BMSCs can obtain better osteodifferentiation capacity compared with BMSCs without any modification, the SE-LncRNAs identified in this study might be used for the epigenetic modification of BMSCs and further promote the osteogenic ability of BMSCs. The BMSCs with epigenetic manipulation have the potential to be used for the 3D-tissue engineered constructs for the treatment of periodontal defects, peri-implant bone defects, and maxillofacial bone defects in the dental area.

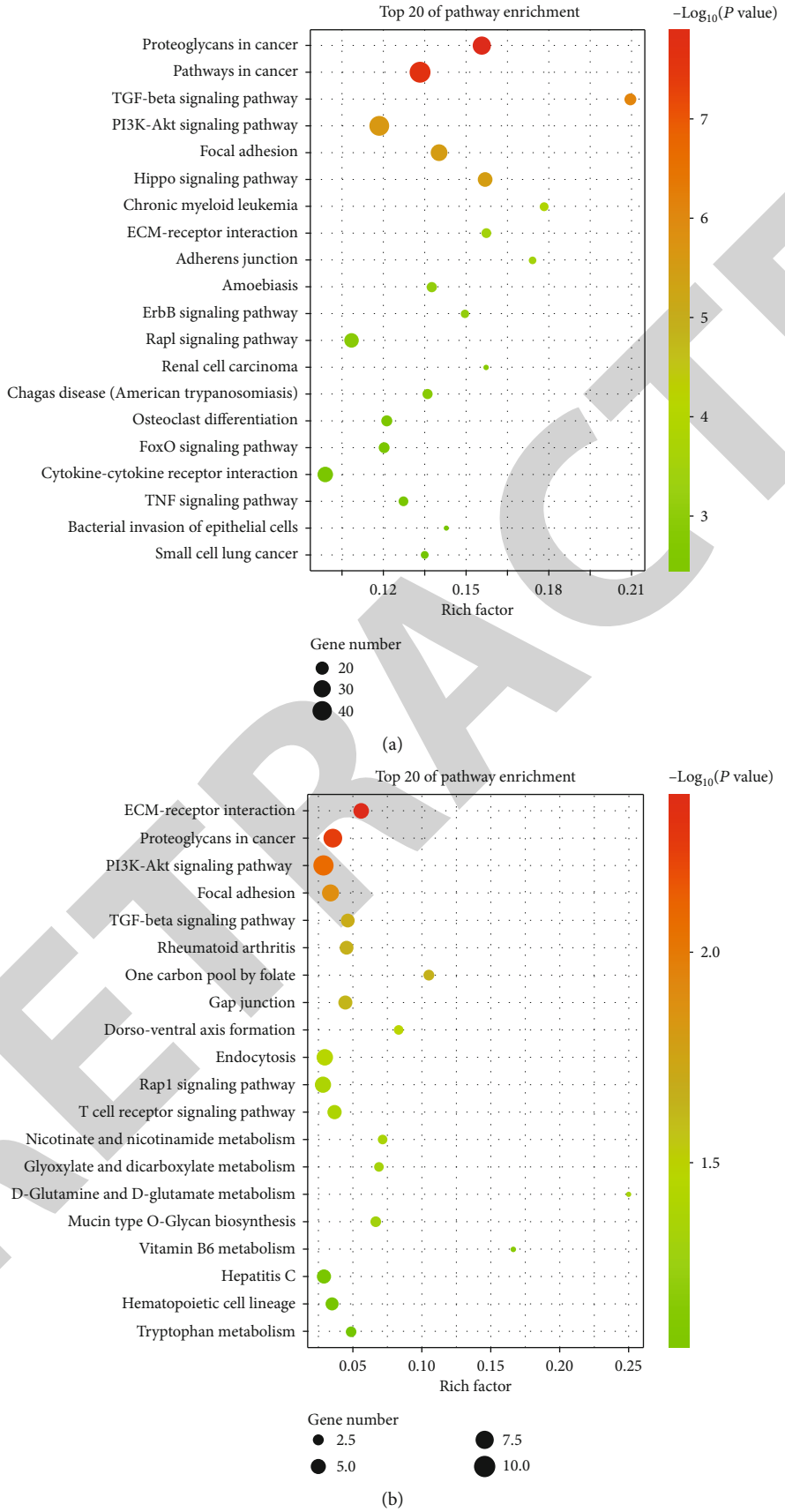


Figure 5: Kyoto Encyclopedia of Genes and Genomes analyses of SE target genes. (a) Top 15 enriched pathways in KEGG analysis for SE target genes in the D0 group. (b) Top 15 enriched pathways in KEGG analysis for SE target genes in the D14 group.

5. Conclusions

The current study suggests that SEs play an important role in the osteogenic differentiation of hBMSCs, providing new targets and theoretical basis for bone tissue engineering. Future studies will select 3–5 SE-related target genes for functional verification, to confirm their role in the osteogenic differentiation of hBMSCs, and to verify their mechanisms and pathways.

Abbreviations

SEs: Super enhancers
hBMSCs: Human bone marrow mesenchymal stem cells
TGF-β: Transforming growth factor β
PI3K: Phosphoinositide 3-kinase
ECM: Extracellular matrix
RT-qPCR: Real-time fluorescence quantitative polymerase chain reaction
TEs: Typical enhancers
WB: Western blot
RUNX2: Runt-related transcription factor 2
OPN: Osteopontin
COL1A1: Collagen type I alpha 1
ChIP-seq: Chromatin immunoprecipitation-sequencing
MACS: Model-based analysis of ChIP-seq
GO: Gene ontology
DAVID: Database for annotation, visualization, and integrated discovery
KEGG: Kyoto encyclopedia of genes and genomes
MF: Molecular functions
CC: Cell composition
BP: Biological processes
mTFs: Master transcription factors
C/EBP-α: CCAAT enhancer-binding protein alpha
Med1: Mediator complex subunit 1
MAPK: Mitogen-activated protein kinase
NF-κB: Nuclear factor kappa B
FGF: Fibroblast growth factor
BMPs: Bone morphogenetic proteins
SMADs: Drosophila mothers against decapentaplegic protein
STAT1: Signal transducer and activator of transcription 1
PKC: Protein kinase C
LncRNA: Long noncoding RNA
miRNA: Micro RNA.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

All the authors declare no conflicts of interest.

Authors' Contributions

Zhijie Huang and Bo Jia contributed equally to this work.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant numbers 81670950), Science and Technology Project of Guangzhou Province (grant numbers 201802020018), China Postdoctoral Science Foundation (grant numbers 2019M652979), the Natural Science Foundation of Guangdong Province (grant numbers 2019A1515010408).We would like to thank Editage (http://www.editage.cn) for English language editing.

References

[1] S. Bhumiratana, J. C. Bernhard, D. M. Alfi et al., "Tissue-engineered autologous grafts for facial bone reconstruction," *Science Translational Medicine*, vol. 8, no. 343, pp. 343r–383r, 2016.

[2] W. L. Grayson, B. A. Bunnell, E. Martin, T. Frazier, B. P. Hung, and J. M. Gimble, "Stromal cells and stem cells in clinical bone regeneration," *Nature Reviews. Endocrinology*, vol. 11, no. 3, pp. 140–150, 2015.

[3] F. Kawecki, W. P. Clafshenkel, M. Fortin, F. A. Auger, and J. Fradette, "Biomimetic tissue-engineered bone substitutes for maxillofacial and craniofacial repair: the potential of cell sheet technologies," *Advanced Healthcare Materials*, vol. 7, no. 6, article 1700919, 2018.

[4] C. R. M. Black, V. Goriainov, D. Gibbs, J. Kanczler, R. S. Tare, and R. O. C. Oreffo, "Bone tissue engineering," *Current Molecular Biology Reports*, vol. 1, no. 3, pp. 132–140, 2015.

[5] E. Pintus, M. Baldassarri, L. Perazzo, S. Natali, D. Ghinelli, and R. Buda, "Stem cells in osteochondral tissue engineering," *Advances in Experimental Medicine and Biology*, vol. 1058, pp. 359–372, 2018.

[6] W. A. Whyte, D. A. Orlando, D. Hnisz et al., "Master transcription factors and mediator establish super-enhancers at key cell identity genes," *Cell*, vol. 153, no. 2, pp. 307–319, 2013.

[7] D. Hnisz, B. J. Abraham, T. I. Lee et al., "Super-enhancers in the control of cell identity and disease," *Cell*, vol. 155, no. 4, pp. 934–947, 2013.

[8] B. Langmead, C. Trapnell, M. Pop, and S. L. Salzberg, "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome," *Genome Biology*, vol. 10, no. 3, p. R25, 2009.

[9] E. Gentleman, R. J. Swain, N. D. Evans et al., "Comparative materials differences revealed in engineered bone as a function of cell-specific differentiation," *Nature Materials*, vol. 8, no. 9, pp. 763–770, 2009.

[10] S. Sart, A. Tsai, Y. Li, and T. Ma, "Three-dimensional aggregates of mesenchymal stem cells: cellular mechanisms, biological properties, and applications," *Tissue Engineering Part B: Reviews*, vol. 20, no. 5, pp. 365–380, 2014.

[11] S. Cai, Y.-P. Tsui, K.-W. Tam et al., "Directed differentiation of human bone marrow stromal cells to fate-committed Schwann cells," *Stem Cell Reports*, vol. 9, no. 4, pp. 1097–1108, 2017.

[12] X. Li, M. Wang, X. Jing et al., "Bone marrow- and adipose tissue-derived mesenchymal stem cells: characterization, differentiation, and applications in cartilage tissue engineering," *Critical Reviews in Eukaryotic Gene Expression*, vol. 28, no. 4, pp. 285–310, 2018.

[13] R. Vishnubalaji, M. Al-Nbaheen, B. Kadalmani, A. Aldahmash, and T. Ramesh, "Comparative investigation of the differentiation