

Research Article

Identifying Obstructive Sleep Apnea Syndrome-Associated Genes and Pathways through Weighted Gene Coexpression Network Analysis

Yan Li,¹ Li Li,² Hua Zhao,¹ Xiwen Gao ¹ and Shanqun Li ²

¹Department of Respiratory, Minhang Hospital, Fudan University, 170 Xinsong Road, Minhang District, Shanghai, China 201199

²Department of Pulmonary Medicine, Minhang Branch, Zhongshan Hospital, Fudan University, 180 Fenglin Rd, Shanghai, China 200032

Correspondence should be addressed to Xiwen Gao; xiwengao@sina.com and Shanqun Li; lisq1118@yeah.net

Received 18 November 2021; Accepted 5 January 2022; Published 29 January 2022

Academic Editor: Jianxin Shi

Copyright © 2022 Yan Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Obstructive sleep apnea syndrome (OSAS) is the most common type of sleep apnea disorder. The disease seriously affects the patient's respiratory system. At present, the prognosis of the disease is poor and there is a lack of effective treatments. Therefore, it is urgent to explore its pathogenesis and treatment methods. **Method.** We downloaded a set of expression profile data from GSE75097 related to OSAS based on the Gene Expression Omnibus (GEO) database and selected the representative differentially expressed genes (DEGs) from the sample of the GSE75097 dataset. WGCNA was used to find genes related to OSAS and obtain coexpression modules. The Gene Ontology (GO) function and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway were used to analyze genes from key modules. Finally, Cytoscape software was used to construct a protein-protein interaction (PPI) network and analyze the hub genes. **Result.** We obtained a total of 7565 DEGs. Through WGCNA, we got four coexpression modules and the modules most related to OSAS were green-yellow, magenta, purple, and turquoise, and we screened out eight hub genes (DDX46, RNF115, COPA, FBXO4, PA2G4, NHP2L1, CDC20, and PCNA). GO and KEGG analyses indicated that the key modules were mainly enriched in tRNA modification, nucleobase metabolic process, DNA ligation, regulation of cellular component movement, basal transcription factors, Huntington disease, and vitamin digestion and absorption. **Conclusion.** These pathways and hub genes can facilitate understanding the molecular mechanism of OSAS and provide a meaningful reference for finding biological targets of OSAS treatment.

1. Background

Obstructive sleep apnea syndrome (OSAS) is a sleep breathing disorder with unknown etiology, which is manifested as upper airway obstruction, apnea during sleep, hypoxia, snoring, and daytime sleepiness [1, 2]. The disease often occurs in the obese and the elderly, and it is also common in children with sleep-disordered breathing [3]. It is worth noting that OSAS can also affect the patient's cardiovascular system, respiratory system, cerebrovascular, central nervous system, endocrine, sexual function, kidney, blood system, and digestive system and even threaten life [4]. Currently, adenoid and tonsillectomy are the main treatments for most children with OSAS [5]. In addition, the Continuous Positive Airway

Pressure (CPAP) is also a device specially developed and manufactured for newborns used for respiratory support [6].

Genomics is aimed at systematically characterizing and quantifying all genes in an organism and studying the relationship among them and their impact on the organism [7]. Genome technology usually includes high-throughput sequencing technology, genome editing technology, and metagenomic technology [8]. Goldfield et al. believe that genomic technology can help to better understand disease subtypes and characteristics at the human level and provide individual-level diagnosis and personalized treatment [9]. Through comprehensive bioinformatics analysis, Gui et al. identified the candidate markers related to the pathology of Alzheimer's disease (AD) and provided comprehensive

insight into understanding AD's pathogenesis and potential new therapeutic targets [10]. With the further development of genome technology, the integration and reanalysis of data kept in public databases can give vital insights for a new study, as well as vast prospects for illness target research [11].

This time, we conducted a series of bioinformatics analyses on 48 samples in the GSE75097 dataset based on weighted gene coexpression network analysis (WGCNA) [12]. Firstly, differentially expressed genes (DEGs) were screened out from the samples, and on this basis, key modules highly related to OSAS were determined through WGCNA [13]. After that, Gene Ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed on DEGs in key modules [14], and protein-protein interaction (PPI) network construction and hub gene screening were performed again on modules with a large number of genes [15]. The above methods will help us find OSAS-related therapeutic targets.

2. Material and Methods

2.1. Acquisition of Microarray Dataset. The microarray dataset GSE75097 provided was downloaded from the Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) database. We analyzed whole-genome gene expression profiles of 48 patients with sleep-disordered breathing stratified into four groups: 6 primary snoring (PS) samples, 16 moderate to severe OSA (MSO) samples, 12 very severe OSA (VSO) samples, and 14 very severe OSA patients with long-term CPAP treatment (VSOC) samples. After that, we used Robust Multiarray Average (RMA) algorithm to process the above 48 samples and filtered the subsequent WGCNA data through nsFilter algorithm.

2.2. Construction of WGCNA and Identification of Key Modules. WGCNA is a freely accessible R package that identifies the gene set of interest based on the information of thousands of the most changing genes or all genes and performs significant association analysis with phenotypes to construct a gene coexpression network. In this study, we determined the best soft threshold based on this method to make the gene network follow the scale-free distribution. Then, we defined the gene coexpression correlation matrix, constructed a hierarchical clustering tree of gene modules according to the formed adjacency function, and finally determined the modules most relevant to OSAS based on the heat map for subsequent bioinformatics analysis.

2.3. Function Enrichment Analysis of Key Modules. The Database for Annotation, Visualization, and Integrated Discovery (DAVID, <https://david.ncifcrf.gov/>) is an online analysis tool released in 2003. Like other similar analysis tools, it associates the genes in the input list with the biological annotation term and then uses statistical methods to find the most significantly enriched biological annotations. GO is a database that defines and describes the functions of genes and proteins. It is divided into three categories: molec-

ular function (MF), biological process (BP), and cellular component (CC). The KEGG database has 17 subdatabases in 4 categories, including genome, chemistry, and system function information, and KEGG pathway is usually used to store information about gene pathways in different species. This time, we performed enrichment analysis on the genes in the key modules based on the DAVID database.

2.4. Acquisition of DEGs of Key Modules and Construction of PPI Networks. After performing functional enrichment analysis on the genes obtained in the key modules, the DEGs in the modules were screened by GEO2R (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>). Later, in order to find genes related to OSAS, we used the Search Tool for the Retrieval of Interacting Genes (STRING, <https://string-db.org/>) database and Cytoscape (<https://www.cytoscape.org/>) software to design the PPI network of the two modules with the largest number of DEGs. Finally, according to the Molecular Complex Detection (MCODE) method, the genes with the highest degree value were identified as the hub genes in this study.

3. Results

3.1. Data Processing. We downloaded the GSE75097 microarray dataset from the GEO database, which involved a total of 48 samples. The original files were background corrected, normalized, and summarized by the RMA algorithm of the R language, and a total of 7565 DEGs were obtained for subsequent WGCNA analysis (Figure 1).

3.2. Construction of WGCNA Related to OSAS Modules. After obtaining 7565 DEGs from 48 samples, WGCNA was constructed on this basis. First, we determined the optimal soft threshold ($\beta = 6$) that conformed to the scale-free topology network to define the adjacency matrix (Figure 2(a)). After that, we clustered the genes with the same expression into the same module to construct a modular gene tree diagram (Figure 2(b)). Then, we got 13 different gene modules, which were divided into two clusters, all related to OSAS (Figure 3(a)). Finally, we selected 9 modules to construct the eigengene adjacency heat map (Figure 3(b)). It should be noted that the gray module contained all genes that were not involved in clustering and would not be analyzed.

3.3. GO Enrichment Analysis of the Genes in Key Modules. According to Figures 4(a)–4(d), in GO, the genes of the green-yellow module were enriched in intra-Golgi vesicle-mediated transport, tRNA modification, RNA modification, negative regulation of transcription regulatory region DNA binding, positive regulation of smoothed signaling pathway and extrinsic apoptotic signaling pathway via death domain receptors, etc. The genes in the magenta module were enriched in the nucleobase metabolic process, pyrimidine nucleobase metabolic process, tRNA modification, negative regulation of JNK cascade and stress-activated MAPK cascade, low-density lipoprotein particle receptor catabolic process, chylomicron assembly, very-low-density lipoprotein particle assembly, etc. The genes in the purple module were

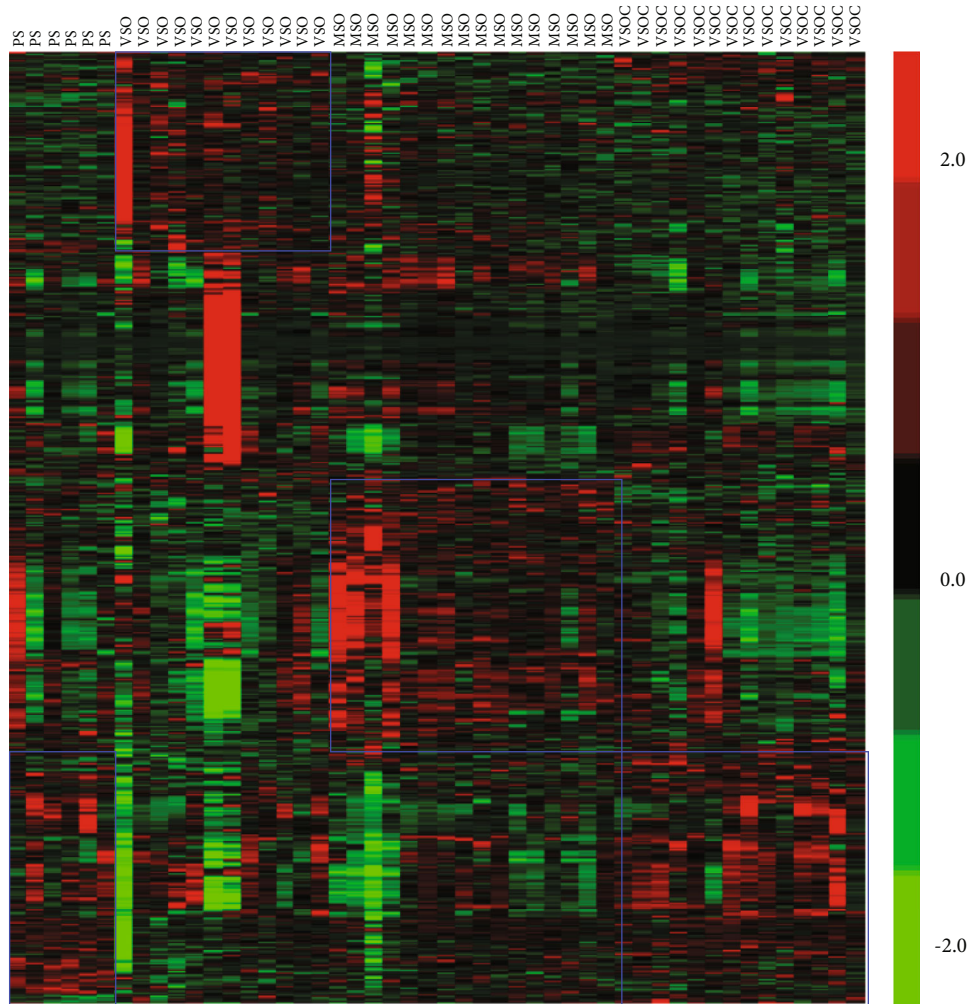


FIGURE 1: Heat map of the GSE75097 dataset. One column is a sample, and one row is a gene.

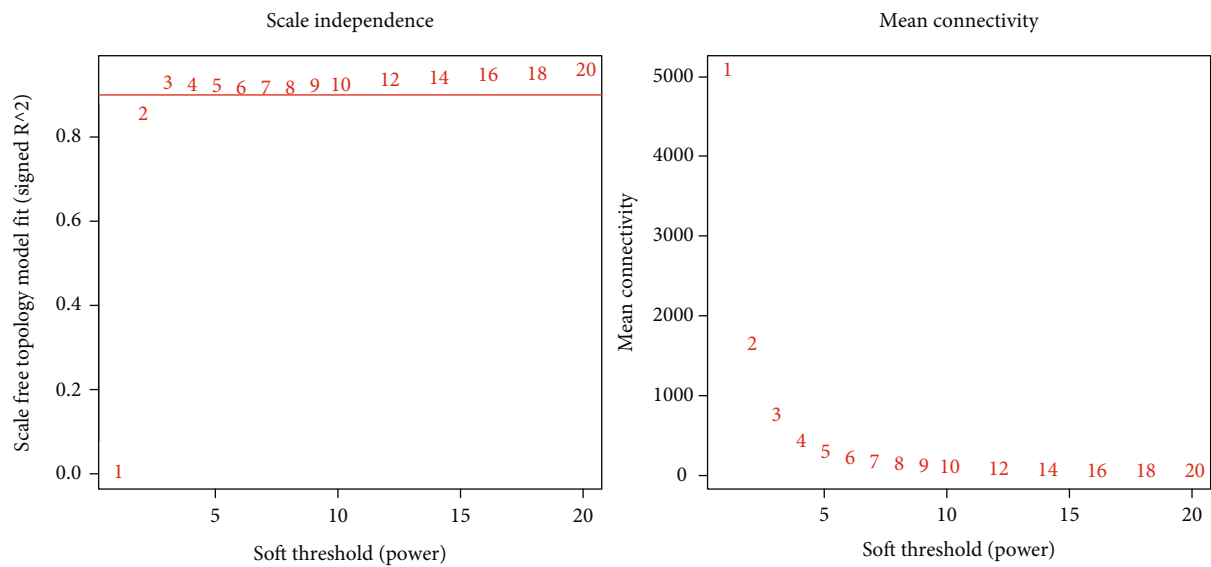
enriched in DNA ligation, pteridine-containing compound biosynthetic process, inositol lipid-mediated signaling, positive regulation of protein complex disassembly, phosphatidylinositol phosphorylation, regulation of telomerase RNA localization to Cajal body, acyl-CoA biosynthetic process, etc. In addition, genes in turquoise modules were also enriched in regulation of cellular component movement, mitochondrial ATP synthesis-coupled proton transport, protein localization to Golgi apparatus, regulation of microtubule-based process, cristae formation, ATP synthesis-coupled proton transport, purine ribonucleoside triphosphate biosynthetic process, etc.

3.4. KEGG Enrichment Analysis of the Genes in Key Modules.

According to the results of the enrichment of KEGG, it was not difficult to find that the genes in the green-yellow module were enriched in vasopressin-regulated water reabsorption, basal transcription factors, propanoate metabolism, selenocompound metabolism, thyroid cancer, other types of O-glycan biosynthesis, and other pathways (Figure 5(a)). The genes in the magenta module were enriched in Huntington's disease, pyrimidine metabolism, cardiac muscle contraction, oxidative phosphorylation, Parkinson's disease,

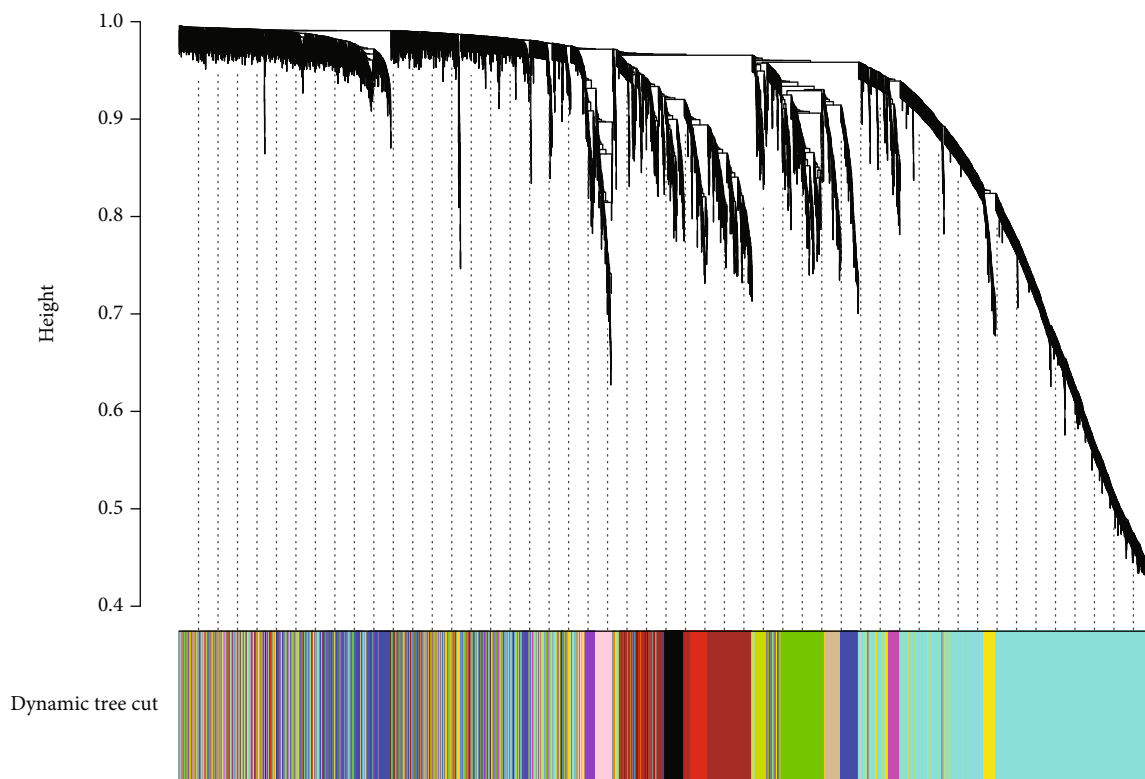
fat digestion and absorption, nonalcoholic fatty liver disease (NAFLD), and other pathways (Figure 5(b)). The genes in the purple module were enriched in vitamin digestion and absorption, folate biosynthesis, fatty acid biosynthesis, RNA polymerase, fatty acid degradation, base excision repair, etc. (Figure 5(c)). Genes in the turquoise module were enriched in alpha-linolenic acid metabolism, glycosylphosphatidylinositol- (GPI-) anchor biosynthesis, maturity onset diabetes of the young, collecting duct acid secretion, linoleic acid metabolism, citrate cycle (TCA cycle), etc. (Figure 5(d)).

3.5. DEGs and Hub Genes of Key Modules. Figure 6 shows the heat map of DEGs in the four key modules. Among them, the number of DEGs in the green-yellow module was 846, the number of DEGs in the magenta module was 206, that in the purple module was 2743, and that in the turquoise module was 152. Next, we selected the two modules with the largest number of DEGs to construct the PPI network. It could be seen from Figure 7(a) that the PPI network of the yellow-green module consisted of 30 nodes and 101 edges; DDX46, RNF115, COPA, and FBXO4 had the same degree value, all being 8. In addition, we also constructed three different PPI



(a)

Gene dendrogram and module colors



(b)

FIGURE 2: WGCNA construction. (a) The left picture is used to determine the optimal soft threshold, and the right picture shows the network connectivity under different soft thresholds. (b) Gene dendrogram and module colors.

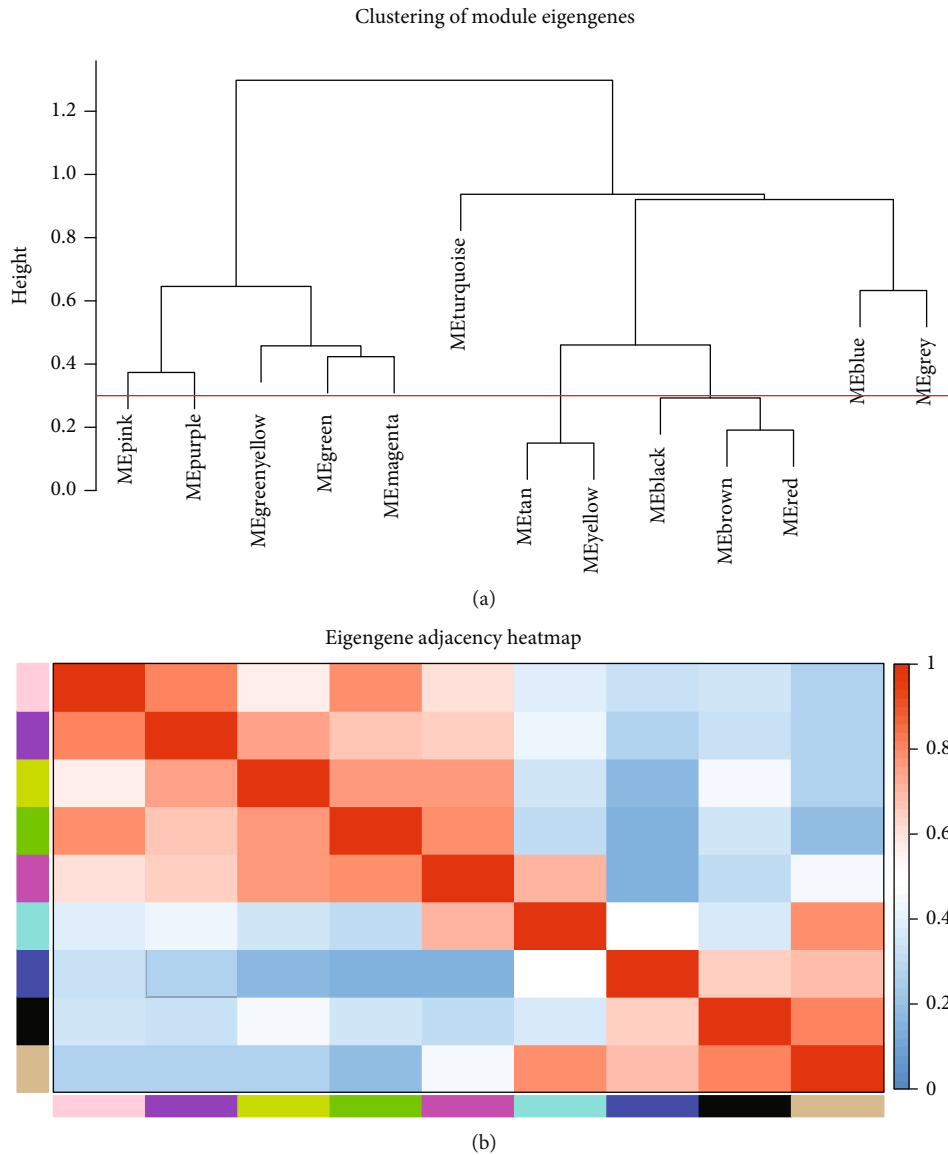


FIGURE 3: Modules in WGCNA. (a) Cluster map of 13 eigengene modules. (b) Eigengene adjacency heat map of 9 modules.

networks based on the DEGs in the purple module. The degree values of the hub genes (PA2G4, NHP2L1) in the PPI network (39 nodes and 681 edges) in Figure 7(b) were all 94, and the hub gene in the PPI network (56 nodes and 523 edges) in Figure 7(c) was CDC20 (degree = 33). The hub gene obtained in the PPI network (49 nodes and 398 edges) in Figure 7(d) was PCNA (degree = 23). The detailed information regarding the hub genes is shown in Table 1.

4. Discussion

Herein, we conduct a series of bioinformatics analyses on 48 samples in the GSE75097 dataset based on WGCNA. In the functional enrichment analysis of GO, the genes of these modules are mainly enriched in biological processes related to epigenetics, including RNA modification, DNA binding and tRNA, nucleobase metabolic process, pyrimidine nucleobase metabolic process, negative regulation of JNK

cascade, and tRNA modification. The so-called epigenetics refers to heritable phenotypic changes without those in DNA sequence [16]. In recent years, hypoxia-mediated epigenetic regulation has had a major function in the mechanism of OSAS [17]. Under chronic intermittent hypoxia-reoxygenation circumstances, the epigenetic process influences the adaptive potential and phenotypic variability, which contributes to the development of numerous deleterious effects of OSAS [18, 19].

In addition, chemical modification of nucleobases functions importantly in different levels of gene expression control. It includes modulating tRNA bases to regulate translation or cytosine methylation and demethylation in the promoter region to silence and reactivate genes. Chen et al. speculated that hypomethylation of the six gene promoter regions implicated in the NPR2 and SP140 pathways may play a crucial role in the establishment of the OSAS phenotype of excessive daytime sleepiness [20].



FIGURE 4: GO function enrichment results. (a) Green-yellow module. (b) Magenta module. (c) Purple module. (d) Turquoise module.

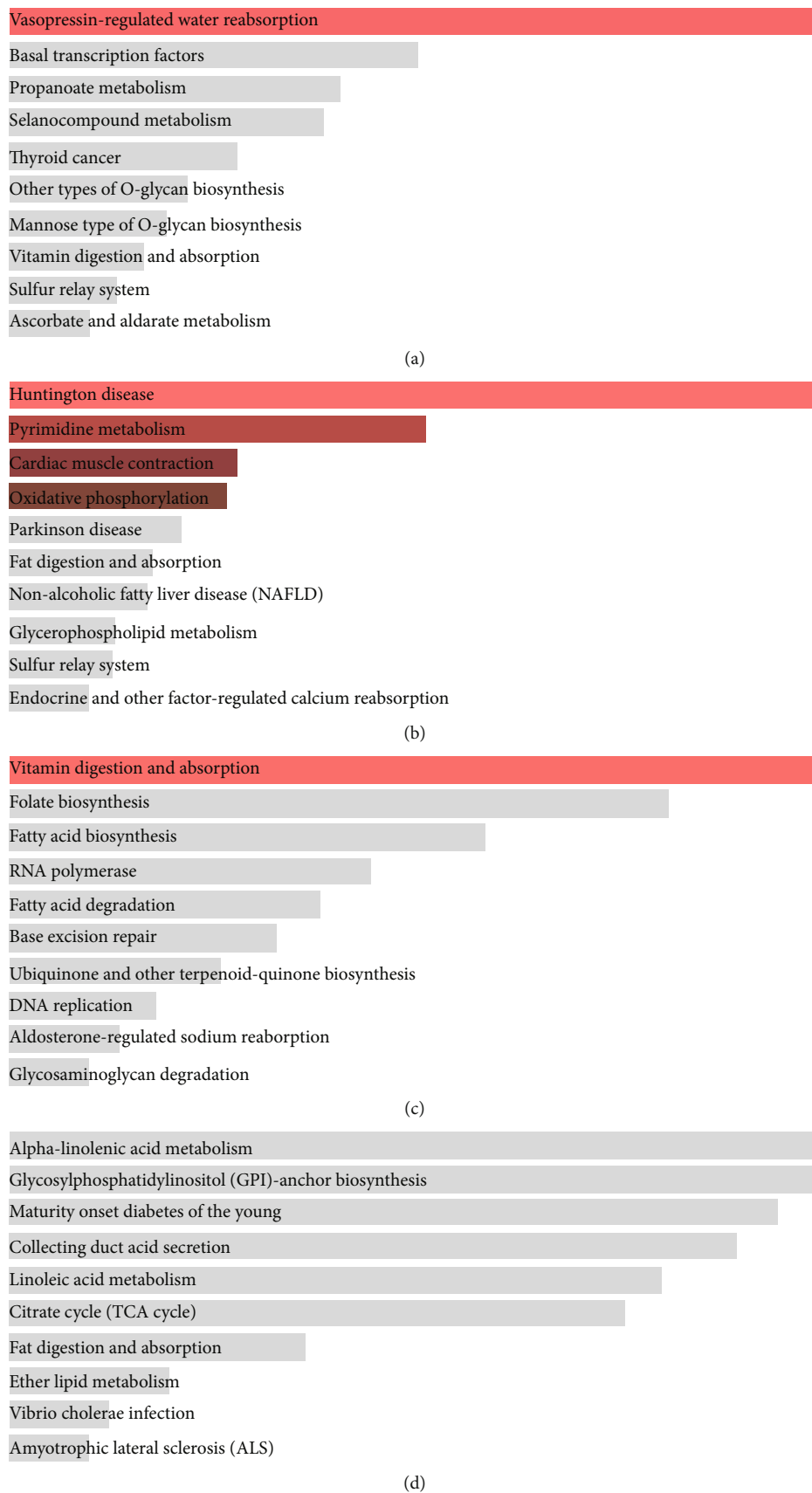
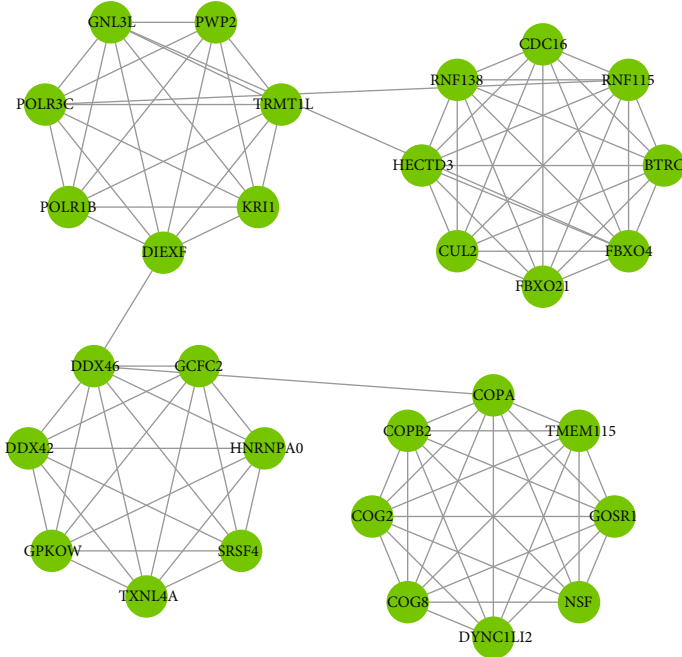
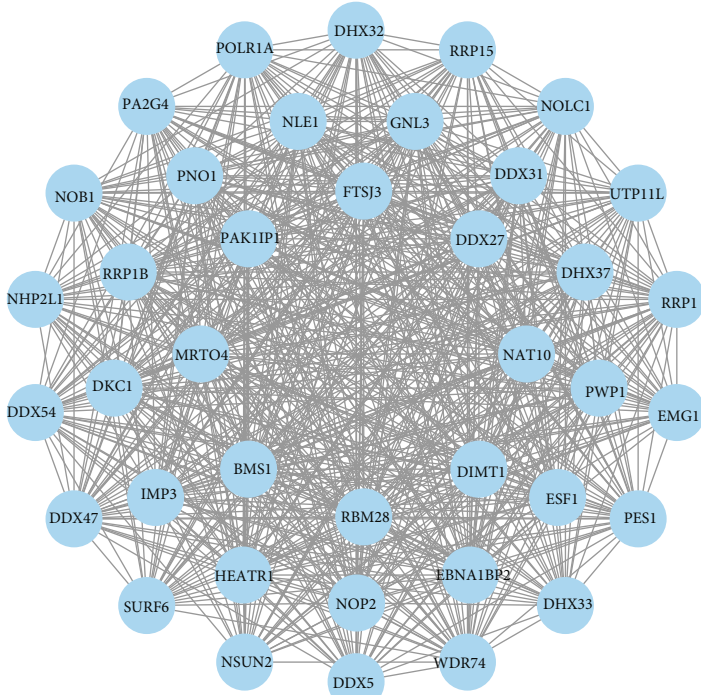


FIGURE 5: KEGG pathway enrichment results. (a) Green-yellow modules. (b) Magenta module. (c) Purple module. (d) Turquoise modules.



(a)



(b)

FIGURE 7: Continued.

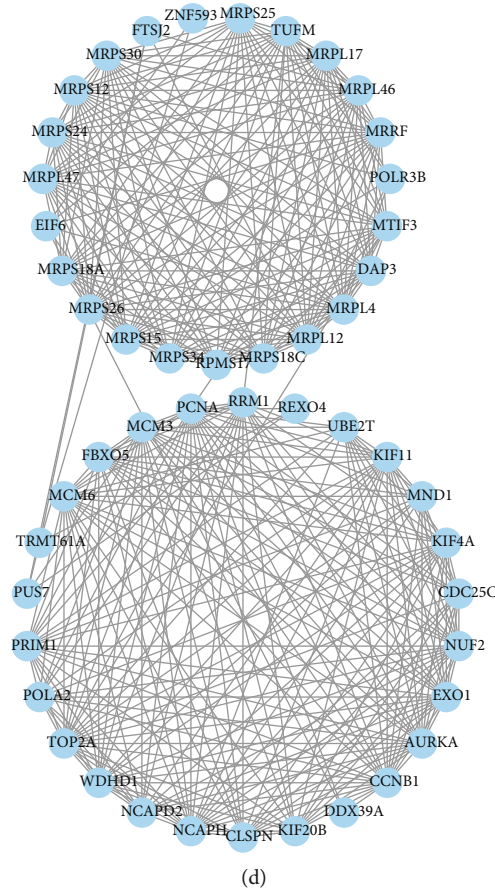


FIGURE 7: PPI network of green-yellow and purple modules. (a) Green-yellow module, hub genes are DDX46, RNF115, COPA, and FBXO4. (b–d) Purple module, hub genes are PA2G4, NHP2L1, CDC20, and PCNA.

TABLE 1: The hub genes with higher degrees of PPI network of OSAS.

Gene symbol	Description	Degree	Module
DDX46	DEAD-box helicase 46	8	Green-yellow
RNF115	Ring finger protein 115	8	Green-yellow
COPA	Coatomer subunit α	8	Green-yellow
FBXO4	F-box protein 4	8	Green-yellow
PA2G4	Proliferation-associated 2G4	94	Purple
NHP2L1	U4/U6 small nuclear ribonucleoprotein	94	Purple
CDC20	Cell division cycle 20	33	Purple
PCNA	Proliferating cell nuclear antigen	23	Purple

indicated folic acid deficiency is substantially associated with obstructive sleep apnea in women [29].

The hub genes we selected from green-yellow module and purple module are reported to correlate with breath and disease related to the lung. There are reports on the relationship between RNF115 and lung adenocarcinoma (LAC). RNF115-mediated p53 ubiquitination can act as the marker of the prognosis of patients with LAC [30]. RNF115 induces a significant arrest of the G1 phase of LAC cells to achieve the function of inhibiting cell viability *in vitro*, further reducing tumor proliferation in xenograft models. These

findings suggest that RNF115 might be a novel predictive biomarker and that the RNF115-p53 axis could be a target for LAC therapy. The cuffed oropharyngeal airway (COPA) is a modified Guedel-type oral airway with a cuff on the end. Propofol and sevoflurane are equally effective in promoting the placement of COPA [31]. However, propofol usually causes apnea, which is complicated by the poor sealing effect of COPA on the airway. At the same time, sevoflurane allows spontaneous breathing to continue and provides sufficient pharynx seal immediately after that COPA is placed, so it may be advantageous when apnea is not required.

Propofol and sevoflurane can promote the placement of COPA to treat obstructive sleep-disordered breathing. Studies have confirmed that several tumor aggressiveness markers like PCNA have high expression, providing molecular evidence for the relationship between apnea and cancer. PCNA is a biomarker that identifies the relationship between apnea and cancer [32]. The results of this study need further experimental verification. In future research, we will explore the specific mechanism of the hub gene in the pathogenesis of OSAS. The correlation between the expression level of the hub gene and the clinical parameters of OSAS patients needs to be further explored.

In conclusion, our analysis results show that the genes in the selected key modules are mainly enriched in epigenetics, DNA methylation, and nucleobases. The hub genes (RNF115, COPA, and PCNA) selected from the key modules are biomarkers related to OSAS, which provides useful information for the therapies of diseases related to OSAS. However, five of these hub genes have not been studied in OSAS, but they are all important hub genes for understanding OSAS, predicting prognosis, and developing precise therapies.

Data Availability

All data analyzed during this study are obtained from published article or are available from the corresponding author on reasonable request.

Consent

Consent is not applicable.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

Yan Li, Li Li, and Shanqun Li designed the study. Yan Li, Xiwen Gao, Hua Zhao, and Shanqun Li performed the study and drafted the article. Yan Li, Li Li, and Shanqun Li conducted data acquisition, data analysis, and interpretation. All authors discussed the results and agreed to be accountable for all aspects of the work. All authors read and approved the final manuscript. Yan li and Li Li contributed equally to this work.

Acknowledgments

This work is supported by the Minhang District Committee of Science and Technology (Grant Number 2014MHZ074) (applicant: Yan Li).

References

- [1] G. Gulotta et al., "Risk factors for obstructive sleep apnea syndrome in children: state of the art," *International Journal of Environmental Research and Public Health*, vol. 16, no. 18, 2019.
- [2] C. Maspero et al., "Obstructive sleep apnea syndrome: a literature review," *Minerva Stomatologica*, vol. 64, no. 2, pp. 97–109, 2015.
- [3] D. A. Dobrosielski et al., "Diet and exercise in the management of obstructive sleep apnoea and cardiovascular disease risk," *European Respiratory Review*, vol. 26, no. 144, 2017.
- [4] A. Lo Bue, A. Salvaggio, and G. Insalaco, "Obstructive sleep apnea in developmental age. A narrative review," *European Journal of Pediatrics*, vol. 179, no. 3, pp. 357–365, 2020.
- [5] V. Stavrou et al., "Brief review: ergospirometry in patients with obstructive sleep apnea syndrome," *Journal of Clinical Medicine*, vol. 7, no. 8, 2018.
- [6] S. Gupta and S. M. Donn, "Continuous positive airway pressure: physiology and comparison of devices," *Seminars in Fetal & Neonatal Medicine*, vol. 21, no. 3, pp. 204–211, 2016.
- [7] L. Del Giacco and C. Cattaneo, "Introduction to genomics," *Methods in Molecular Biology*, vol. 823, pp. 79–88, 2012.
- [8] R. Guigo and M. de Hoon, "Recent advances in functional genome analysis," *F1000Res*, vol. 7, p. 1968, 2018.
- [9] R. Goldfield, J. A. Bianculli, and R. W. Sager, "Analytical constants of the fixed oil from the seed of *Pyralia pubera*," *American Journal of Pharmacy and the Sciences Supporting Public Health*, vol. 126, no. 2, pp. 62–64, 1954.
- [10] H. Gui, Q. Gong, J. Jiang, M. Liu, and H. Li, "Identification of the hub genes in Alzheimer's disease," *Computational and Mathematical Methods in Medicine*, vol. 2021, 2021.
- [11] Y. J. Zhang et al., "Integrated bioinformatic analysis of differentially expressed genes and signaling pathways in plaque psoriasis," *Molecular Medicine Reports*, vol. 20, no. 1, pp. 225–235, 2019.
- [12] B. Zhang and S. Horvath, "A general framework for weighted gene co-expression network analysis," *Statistical Applications in Genetics and Molecular Biology*, vol. 4, p. 17, 2005.
- [13] S. A. Khan, H. Eggleston, K. M. Myles, and Z. N. Adelman, "Differentially and co-expressed genes in embryo, germ-line and somatic tissues of *Tribolium castaneum*," *G3 (Bethesda)*, vol. 9, no. 7, pp. 2363–2373, 2019.
- [14] Q. Tao et al., "Network pharmacology and molecular docking analysis on molecular targets and mechanisms of Huashi Baidu formula in the treatment of COVID-19," *Drug Development and Industrial Pharmacy*, vol. 46, no. 8, pp. 1345–1353, 2020.
- [15] A. Athanasios et al., "Protein-protein interaction (PPI) network: recent advances in drug discovery," *Current Drug Metabolism*, vol. 18, no. 1, pp. 5–10, 2017.
- [16] L. Zhang, Q. Lu, and C. Chang, "Epigenetics in health and disease," *Advances in Experimental Medicine and Biology*, vol. 1253, pp. 3–55, 2020.
- [17] O. H. Liu, M. Kiema, M. Beter, S. Ylä-Herttuala, J. P. Laakkonen, and M. U. Kaikkonen, "Hypoxia-mediated regulation of histone demethylases affects angiogenesis-associated functions in endothelial cells," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 40, no. 11, pp. 2665–2677, 2020.
- [18] F. F. de Lima, D. R. Mazzotti, S. Tufik, and L. Bittencourt, "The role inflammatory response genes in obstructive sleep apnea syndrome: a review," *Sleep & Breathing*, vol. 20, no. 1, pp. 331–338, 2016.
- [19] A. Rus et al., "Upregulation of cardiac NO/NOS system during short-term hypoxia and the subsequent reoxygenation period," *European Journal of Histochemistry*, vol. 55, no. 2, article e17, 2011.

- [20] Y. L. Chen et al., "GJA1 expression and left atrial remodeling in the incidence of atrial fibrillation in patients with obstructive sleep apnea syndrome," *Biomedicine*, vol. 9, no. 10, 2021.
- [21] P. Adjamian et al., "Neuroanatomical abnormalities in chronic tinnitus in the human brain," *Neuroscience and Biobehavioral Reviews*, vol. 45, pp. 119–133, 2014.
- [22] J. A. Dempsey, "Central sleep apnea: misunderstood and mistreated!," *F1000Res*, vol. 8, 2019.
- [23] R. J. Maughan and S. M. Shirreffs, "Muscle cramping during exercise: causes, solutions, and questions remaining," *Sports Medicine*, vol. 49, Suppl 2, pp. 115–124, 2019.
- [24] Y. C. Kuo et al., "Short-term expiratory muscle strength training attenuates sleep apnea and improves sleep quality in patients with obstructive sleep apnea," *Respiratory Physiology & Neurobiology*, vol. 243, pp. 86–91, 2017.
- [25] A. H. Hall, J. Saiers, and F. Baud, "Which cyanide antidote?," *Critical Reviews in Toxicology*, vol. 39, no. 7, pp. 541–552, 2009.
- [26] H. Kaur and P. Singh, "Rationally designed molecules for resurgence of cyanide mitigated cytochrome c oxidase activity," *Bioorganic Chemistry*, vol. 82, pp. 229–240, 2019.
- [27] J. L. Parker-Cote et al., "Challenges in the diagnosis of acute cyanide poisoning," *Clinical Toxicology (Philadelphia, Pa.)*, vol. 56, no. 7, pp. 609–617, 2018.
- [28] S. Papa et al., "The oxidative phosphorylation system in mammalian mitochondria," *Advances in Experimental Medicine and Biology*, vol. 942, pp. 3–37, 2012.
- [29] X. Zhang et al., "HFD and HFD-provoked hepatic hypoxia act as reciprocal causation for NAFLD via HIF-independent signaling," *BMC Gastroenterology*, vol. 20, no. 1, p. 366, 2020.
- [30] Z. Luo et al., "RNF115-mediated ubiquitination of p53 regulates lung adenocarcinoma proliferation," *Biochemical and Biophysical Research Communications*, vol. 530, no. 2, pp. 425–431, 2020.
- [31] J. M. van Vlymen et al., "Use of the cuffed oropharyngeal airway as an alternative to the laryngeal mask airway with positive-pressure ventilation," *Anesthesiology*, vol. 90, no. 5, pp. 1306–1310, 1999.
- [32] S. Perini, D. Martinez, C. C. Montanari, and C. Z. Fiori, "Enhanced expression of melanoma progression markers in mouse model of sleep apnea," *Revista Portuguesa de Pneumologia*, vol. 22, no. 4, pp. 209–213, 2016.