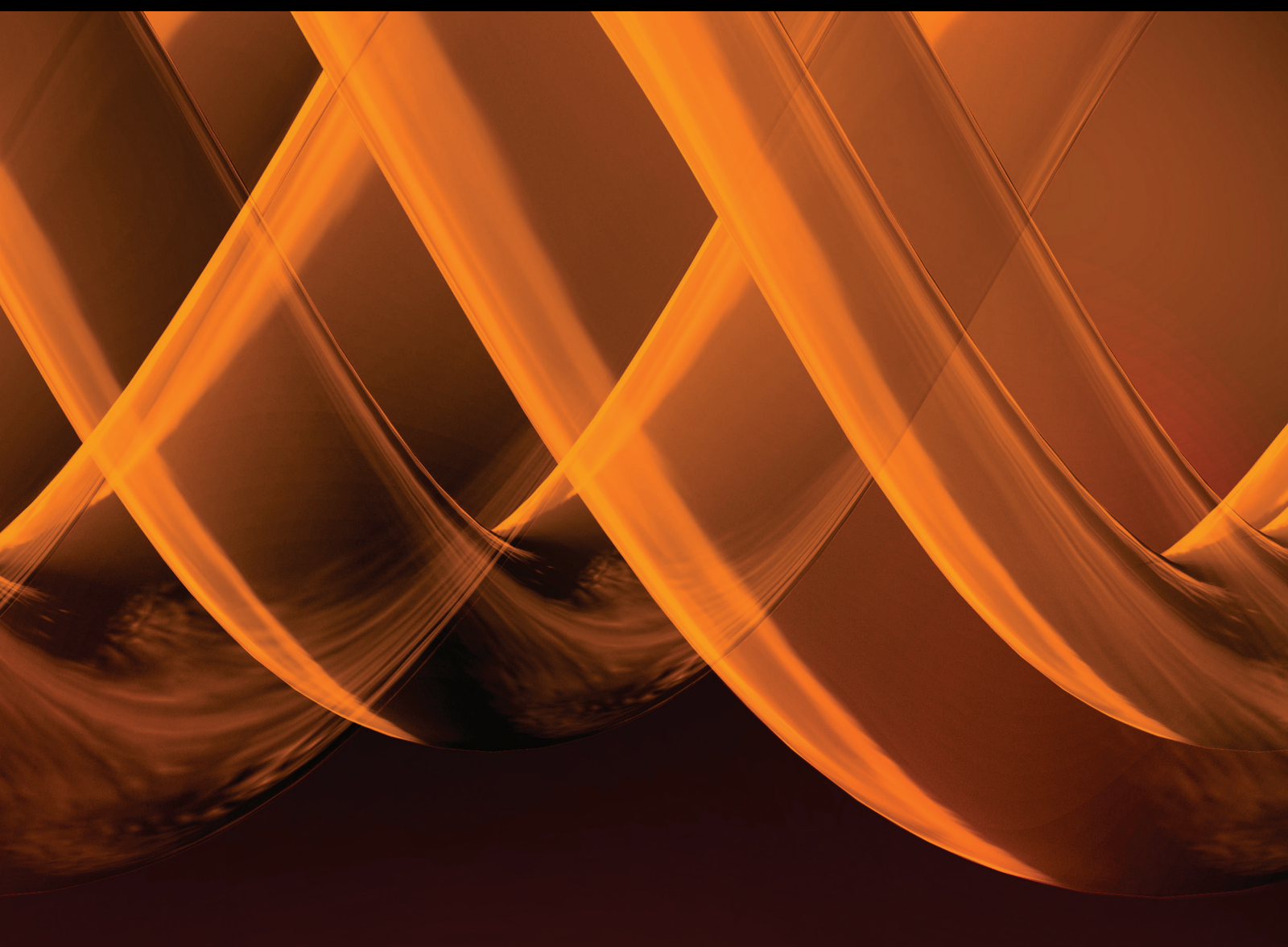
International Journal of Genomics

Noncoding RNAs and Base Modifications: Epigenomic Players Implicated in Neurological Disorders and Tumorigenesis

Lead Guest Editor: Yujing Li
Guest Editors: Luciano Vellón, Hong Jiang, Changwon Park, and Xuekun Li



Noncoding RNAs and Base Modifications: Epigenomic Players Implicated in Neurological Disorders and Tumorigenesis

International Journal of Genomics

Noncoding RNAs and Base Modifications: Epigenomic Players Implicated in Neurological Disorders and Tumorigenesis

Lead Guest Editor: Yujing Li
Guest Editors: Luciano Vellón, Hong Jiang, Changwon Park, and Xuekun Li

Editorial Board

Contents

Hindawi
International Journal of Genomics
Volume 2018, Article ID 9016018, 2 pages
https://doi.org/10.1155/2018/9016018

Editorial

Noncoding RNAs and Base Modifications: Epigenomic Players Implicated in Neurological Disorders and Tumorigenesis

Yujing Li[1], **Changwon Park**[1], **Hong Jiang**[2], **Xuekun Li**[3] and **Luciano Vellón**[4]

[1]*Emory University School of Medicine USA, USA*
[2]*Xiangya Hospital at the Central South University, China*
[3]*Zhejiang University School of Medicine, Hangzhou, Zhejiang, China 310029*
[4]*Institute of Biology and Experimental Medicine (IBYME-CONICET), Buenos Aires, Argentina*

Correspondence should be addressed to Yujing Li; yli29@emory.edu

Received 9 August 2018; Accepted 9 August 2018; Published 24 September 2018

Epigenetics is defined as the study of gene expression regulations by nongenetic but heritable other than DNA sequence variations. The epigenetic memories stored in the base modifications and histone modifications such as methylation, phosphorylation, and acetylation play important roles in regulation of all the metabolic pathways. Any aberrant alterations could lead to developmental abnormalities of diverse diseases including neurological disorders and cancers.

Before 2009, DNA methylation of 5-cytosine (5-mC) was the best characterized epigenetic marker. In recent years, functional significance of 5-hydroxycytosine (5-hmC), the sixth modified base, as an important epigenetic marker has been revealed. The 5-hmC, converted from 5mC and catalyzed by the ten-eleven translocation proteins (TETs), has become a hallmarker of cancer cells. The TET family consists of three members, TET1, TET2, and TET3, in mammals. It has been shown that TET1 functions as a tumor suppressor and tumor promoter in a cancer type-dependent manner. Deletion and inactivating mutations in the TET2 gene account for almost 30% of all myeloid malignancies, suggesting the essential roles of 5-hmC and TET2 in tumorigenesis. In addition, increasing accumulated lines of evidences are supporting the important roles of 5-hmC in almost all the known metabolic pathways that are involved in embryogenesis, cell reprogramming, stem cell self-renewal, proliferation and differentiation, central nervous system (CNS) development, and aging.

The regulatory mechanisms of 5-hmC in the metabolic pathways have been acknowledged to serve as the intermediate in the DNA demethylation process as well as the stable and independent marker to regulate gene expression by altering the epigenetic landscapes. Besides the base modifications, additional complexity of gene expression is added through dynamic changes on histone proteins and associated chromatin structural architectures. Although our understanding on the functions of epigenetic regulations is still incomplete, the accumulated data have suggested the link between aberrant alteration of base modification (such as 5-mC, 5-hmC, and m6A) levels and neurological diseases, cardiovascular diseases, and cancers. Additionally, noncoding RNAs (ncRNAs), particularly microRNAs (miRNAs) small RNAs and long noncoding RNAs (lncRNAs), have also been characterized as important epigenetic players in various pathophysiological events. Thus, it is of significance to publish a special issue focused on this topic. In this special issue, research and review articles are collected, covering the recent progress in the roles of the ncRNAS (particularly miRNAs and lncRNAs) and base modification epigenetic markers in pathogenesis of neurological disorders and cancers.

With the rapid development of high throughput sequencing technology and big data analysis, significant attentions have been paid to the identification of novel miRNA networks in regulating gene transcriptions involved in the pathogenesis of a certain type of disease such as cancer. Two research papers deal with the miRNA networks in breast

cancer and renal cell carcinoma, respectively. Anda-Jauregui et al. systematically compared the highly influential nonredundant miRNAs named Commodore miRNAs (Cdre-miRs) from the healthy and cancerous breast tissues and found the significant variations in the degree, clustering coefficient, and redundancy distributions for the Cdre-miRs and their target genes involved in the metabolic networks for tumorigenesis. Furthermore, the analysis revealed five Cdre-miRs, including miR-let7i, miR-292b, miR-511, and miR-141, and each of which targets the genes involved in particular functions closely relating to tumorigenesis and metastasis.

In a separate report, Sage et al. identified a subset of novel miRNAs in kidney tissues by analyzing small RNA transcriptomes generated from the clear cell renal cell carcinoma (ccRCC) vs. the normal renal tissues. The expression differences of these novel miRNAs between the ccRCC and the nonmalignant renal tissues suggest that they could serve as specific markers of the kidney cancers. Furthermore, the authors found the significant association of the newly identified miRNAs with patient survival rates, providing a new resource for the renal cancer study and addressing the significance of exploration for the uncharacterized epitranscriptomes.

Long noncoding RNAs (lncRNAs) represent another large class of ncRNAs that regulate gene expression at a transcriptional level by targeting transcriptional activators or repressors, as well as at posttranscriptional levels by enhancing pre-mRNA processing, splicing, transport, translation, and degradation. Functionally, lncRNAs are involved in epigenetic modifications as well, leading to chromosome remodeling. Thus far, only limited numbers of lncRNAs were identified and characterized. A research paper by Long et al. reports the identification of a novel lncRNA from the model mice and detection of the significant difference in expression between WT and the SCA3/MJD model mice at the transcriptional level, opening a new avenue for dissection of the SCA3/MJD pathogenesis.

Despite extensive studies on the base modifications, the majority of the research has focused on the genomic DNA, limiting the availability of the epitranscriptomic information. Wei et al. compared 5-mC distribution in the epitranscriptomes of a human and mouse and found that the 5-mC was highly enriched at 5′UTRs of human and mouse mRNAs, and an inverse correlation between mRNA and DNA 5-mC levels at the CpG sites was observed. Generally, it is well known that 5-mC levels in genomic DNA are negatively correlated with transcriptional levels. In contrast, elevated RNA 5-mC levels enhance RNA translation and increase RNA half-life, although the detailed mechanisms remain elusive. Additionally, the 5-mC levels of the mitochondrial RNAs are significantly elevated compared to that of nuclear genomic DNA. Altogether, this study uncovers important roles of the 5-mC-based epitranscriptomic landscapes in metabolism of RNA particularly mRNA and mitochondrial RNA.

Significant efforts paid to the studies on the epigenetic involvement to the pathogenesis of tumorigenesis lead to the dramatic achievements in elucidating the mechanisms, contributing to development of novel strategies for the epigenetic-based therapeutic treatments of cancers. Two separate articles by Feng et al. and He et al. highlighted the recent progress in epigenetic regulation of cervical cancer and ovarian cancer, respectively. They focused on the roles of three base modification markers including 5-mC, 5-hmC, and 6-mA in DNA and/or RNA in regulation of genes/pathways involved in tumorigenesis.

To sum up, this special issue specifically addresses the research progress in base modifications and ncRNA-mediated epigenetic regulation of cancers and neurological disorders by either presenting original research outcomes or highlighting the recent advances in the related research fields. The authors of the research and review articles in the special issue hold the hope that they make a shortcut for the readers to get the related information easily, who are dedicated enthusiastically in the research field of epigenetic regulation of tumorigenesis and neurological diseases.

Conflicts of Interest

The editors declare that they have no conflicts of interest regarding the publication of this special issue.

Yujing Li
Changwon Park
Hong Jiang
Xuekun Li
Luciano Vellón

Hindawi
International Journal of Genomics
Volume 2018, Article ID 6972397, 10 pages
https://doi.org/10.1155/2018/6972397

Research Article

Expanding the miRNA Transcriptome of Human Kidney and Renal Cell Carcinoma

Adam P. Sage, Brenda C. Minatel, Erin A. Marshall, Victor D. Martinez, Greg L. Stewart, Katey S. S. Enfield, and Wan L. Lam

Department of Integrative Oncology, British Columbia Cancer Research Centre, Vancouver, BC, Canada

Correspondence should be addressed to Adam P. Sage; asage@bccrc.ca

Received 2 March 2018; Revised 15 May 2018; Accepted 28 May 2018; Published 3 July 2018

Academic Editor: Yujing Li

Despite advancements in therapeutic strategies, diagnostic and prognostic molecular markers of kidney cancer remain scarce, particularly in patients who do not harbour well-defined driver mutations. Recent evidence suggests that a large proportion of the human noncoding transcriptome has escaped detection in early genomic explorations. Here, we undertake a large-scale analysis of small RNA-sequencing data from both clear cell renal cell carcinoma (ccRCC) and nonmalignant samples to generate a robust set of miRNAs that remain unannotated in kidney tissues. We find that these novel kidney miRNAs are also expressed in renal cancer cell lines. Moreover, these sequences are differentially expressed between ccRCC and matched nonmalignant tissues, implicating their involvement in ccRCC biology and potential utility as tumour-specific markers of disease. Indeed, we find some of these miRNAs to be significantly associated with patient survival. Finally, target prediction and subsequent pathway analysis reveals that miRNAs previously unannotated in kidney tissues may target genes involved in ccRCC tumourigenesis and disease biology. Taken together, our results represent a new resource for the study of kidney cancer and underscore the need to characterize the unexplored areas of the transcriptome.

1. Introduction

Despite recent advancements in the diagnosis and treatment of kidney cancer, patients are faced with a poor prognosis, especially when diagnosed at a later stage [1]. Kidney cancer is a heterogeneous disease with multiple subtypes, of which clear cell renal cell carcinoma (ccRCC) is the most frequently observed, accounting for 70–75% of cases [2]. While environmental risk factors including hypertension, smoking, obesity, and a history of chronic kidney disease may modulate an individual's susceptibility, ccRCC arises from molecular aberrations that can be both sporadic and inherited [2, 3]. Many of these alterations result from DNA copy number losses, mutations, and hypermethylation events, commonly affecting genes associated with cellular metabolism [4, 5]. The most frequently affected gene is the *von Hippel-Lindau* (VHL) tumour suppressor gene, while other molecular disruptions affecting multiple components of the PI3K-mTOR and AMPK signaling pathways have also been described [5, 6]. Considering the close association of metabolic reprogramming with ccRCC development and progression, remarkable advances have been made in the treatment of ccRCC patients with the use of antiangiogenic therapies [7]. Despite the increased treatment efficacy of antiangiogenic therapies, patient outcome is impaired by the lack of clinically relevant diagnostic or prognostic markers [8, 9].

The increased availability of next-generation sequencing has led to a dramatic increase in the understanding of noncoding RNAs (ncRNAs). Perhaps, the most well-studied type of ncRNA is microRNAs (miRNAs), short (~22 nt) transcripts that have emerged as critical regulators of gene expression. Since the discovery of *lin-4* in 1993, miRNAs have been found to regulate a multitude of transcripts and their subsequent cellular processes, including proliferation, metabolism, apoptosis, and development [10–12]. Moreover, the relatively long half-life of miRNAs makes them attractive candidates for biomarkers of disease [13]. miRNAs have been observed to be critical to kidney development, physiology, and pathology.

For instance, a cluster of miRNAs (miR-17~92) has been shown to regulate nephrogenesis; both miR-9 and miR-374 are observed to suppress claudin-14 and affect Ca^{2+} readsorption in the ascending limb of Henle; and recent studies have detected aberrant expression of miRNAs in kidney tumours [13–16]. However, despite the mounting evidence of a role for miRNAs in ccRCC, they have yet to be used in kidney cancer diagnostics.

miRNAs have been described to be genus- and tissue-specific; yet, initial characterizations of the human miRNA transcriptome have relied heavily on sequence abundance and conservation, while using relatively low-depth coverage techniques. In light of this, it has been hypothesized that the human genome likely encodes a markedly greater number of miRNAs than currently annotated, which may be able to be identified through a focus on individual tissues and cell lineages [17]. Indeed, recent genome-wide studies have uncovered miRNAs that have previously escaped detection and have observed that these newly detected miRNAs are highly tissue-specific [18]. Additionally, previously uncharacterized miRNAs may in fact represent novel regulators of tissue-specific biology and pathogenesis and may have utility in the clinic as disease markers. Thus, in this study, we use a large-scale analysis of high-throughput sequencing data to probe for novel miRNAs in human kidney tissue. Discovery of these previously unannotated miRNAs provides a new resource to delineate ccRCC pathogenesis.

2. Materials and Methods

2.1. Small RNA-Sequencing and Data Collection. A cohort of clear cell renal cell carcinoma (ccRCC) tumours with paired nonmalignant tissues ($n = 71$), as well as unpaired tumours ($n = 502$), was processed by The Cancer Genome Atlas (TCGA) Research Network (http://cancergenome.nih.gov/). Small RNA-sequencing data were generated on the Illumina HiSeq200 platform and were acquired from the Cancer Genomics Hub (cgHUB) Data Repository (dbgap Project ID: 6208) under the TCGA-KIRC data collection heading. All data analyzed in this experiment are available publically.

2.2. Preprocessing of Small RNA-Sequencing Data. All raw sequence data obtained from TCGA were processed using a previously published custom sequence analysis pipeline designed for small RNA sequence detection [19]. Raw BAM files from TCGA were first converted to FASTQ files of unaligned reads. Unaligned reads were trimmed based on their Phred quality score, which is required to be ≥20. The trimmed reads were then realigned to the current build of the human genome (hg38 annotation) using the Spliced Transcripts Alignment to a Reference (STAR) aligner.

2.3. Detection and Filtering of Novel miRNA Sequences. Through the OASIS online small-RNA-sequencing analysis platform, novel miRNA sequences were predicted using the miRDeep2 algorithm [20]. miRDeep2 takes both relative free energy and the p values associated with random folding to predict species with miRNA-like structure and to generate a miRDeep2 score reflective of the reliability of the prediction. To confirm the validity of these predictions, stringent manual assessments were performed to generate a robust set of previously unannotated miRNA sequences. Manual filtering was based on (1) an adequate number of sequencing reads covering each locus (≥10 reads); (2) no presence of rRNA/tRNA reads, based on the Rfam database [21]; (3) significant ($p < 0.05$) probability of miRNA-like secondary structure; and (4) removal of duplicate sequences. Standard nucleotide blast (BLASTn) was performed on all remaining predicted novel miRNA sequences using the BLAST+ command line application [22]. This step ensures that any sequences with homology to miRNAs annotated in miRBase v21 are not included in the final list of predicted sequences. Sequences with an expect (E) value of <0.1 were considered previously annotated and discarded from further analyses. The mean and standard deviation of the GC content of all predicted novel miRNA sequences were calculated in order to remove any transcripts with GC content ± 2 standard deviations from the mean. Expression levels of these newly detected sequences were assessed on a per-sample basis using the algorithm featureCounts v1.4.6 [23–25]. Quantification data were normalized according to the weighted trimmed mean of the log expression ratios (trimmed mean of M values, TMM method). Previously unannotated miRNA species were considered to be expressed if the sum of the sequencing reads across all samples was at least 10 reads. Finally, sequences were queried for their presence in five previous studies that have identified novel miRNAs in various tissues [17, 18, 26–28] (Supplemental Table 1).

2.4. Validation of miRNA Expression in Cell Lines. We have performed an in-depth analysis of the small noncoding RNA transcriptome of the National Cancer Institute's cancer cell line panel (NCI-60), which includes eight renal cancer cell lines (A498, CAKI-1, 786-0, TK-10, UO-31, ACHN, RXF393, and SN12C) [28]. Small RNA-sequencing data generated from these renal cancer cell lines were processed as described previously to detect and quantify the expression of the predicted novel miRNA sequences discovered in ccRCC tumours. Cell line characteristics and detailed sequencing information are available in Supplemental Table 2. An expression cutoff (TMM > 0.1) was used to ensure that the sequencing reads were expressed in at least one of the renal cancer cell lines. Predicted miRNA sequences that were expressed in data obtained from both TCGA patient samples and NCI-60 cell lines were considered to be validated and were assigned a unique ID consisting of Knm (kidney novel miRNA sequence), followed by the locus position (i.e., Knm22_2209).

2.5. Differential Expression of Previously Unannotated miRNAs in ccRCC and Nonmalignant Tissue. As performed in the cell line validation step, an expression cutoff of TMM > 0.1 in 10% of samples was used to determine newly detected miRNA loci that had detectable expression in both ccRCC tumour samples and paired nonmalignant tissue ($n = 71$). Fold change values were calculated as the ratio of expression of the newly detected miRNA loci in tumours to their expression in nonmalignant tissue. A Student t-test

was performed on the TMM expression values of the newly detected miRNAs in RStudio v3.3.3 to test for statistically significant differences in expression between ccRCC tumours and paired nonmalignant samples. Multiple correction analyses were performed using the Benjamini-Hochberg (BH) correction, to account for the large number of samples and probes being analyzed. Additionally, unsupervised hierarchical clustering analysis using average distance and Pearson correlation metrics of differentially expressed miRNAs was performed to visualize their expression on a per-sample basis (Supplemental Figure 2).

2.5.1. Cell Culture and Real-Time Quantitative PCR (RT-qPCR). The renal cancer cell line TK-10 and the immortalized nonmalignant embryonic kidney cell line HEK-293T were used to further explore the expression of the previously unannotated miRNAs and their deregulation in ccRCC tumours in vitro. TK-10 cells were cultured in RPMI 1640 + 10% FBS, while HEK-293T were cultured in DMEM + 10% FBS. Both cell lines were maintained in an incubator at 37°C and 5% CO_2. Once confluent, cells were harvested for RNA extraction using the Quick-RNA™ MiniPrep Kit (Zymo Research, Catalog number R1055), following manufacturer's guidelines. Custom reverse-transcription primers specific to the mature miRNA sequence were obtained for the novel miRNA candidates Knm3_1968 (GCAGAUUCC CAGAGUGGGACAG) and Knm17_1130 (UGAGGUGGA GGGUUGUGGGA) using the Custom TaqMan® Small RNA Assay Design Tool from Thermo Fisher. The cDNA conversions were performed with the TaqMan MicroRNA Reverse Transcription Kit according to manufacturer's instructions using 2 ng/μL RNA samples for both cell lines. Finally, RT-qPCR analyses using the custom primers generated from Thermo Fisher were performed in triplicate in an Applied Biosystems® 7500 Real-Time PCR System. Relative miRNA expression was calculated via the $2^{-\Delta\Delta Ct}$ method and normalized to the expression of U6 snRNA.

2.6. Survival Analyses of Previously Unannotated miRNAs. Phenotypic information for all ccRCC tumour samples was obtained from GDC (TCGA-KIRC) through UCSC Xena (http://xena.ucsc.edu/). Samples were sorted by high to low miRNA expression, and tertiles were defined. Patients were categorized by vital status and days to death/last follow-up. The Gehan-Breslow-Wilcoxon test was used to assess the significance of the associations between miRNA expression and patient outcome for each miRNA sequence examined. The log-rank test was also considered.

2.7. Protein-Coding Gene Target Prediction and Pathway Enrichment Analysis. To determine the potential target genes of the newly detected kidney miRNAs, they were queried against all human 3′ untranslated region (UTR) sequences, acquired from Ensembl through the BioMart tool (https://www.ensembl.org), using the miRanda v3.3a algorithm [29]. Predicted gene targets of the miRNAs were validated by individually running five separate scrambled sequences through the algorithm. Predicted targets that overlapped between any of the scrambled sequences and the true sequence were discarded. Strict parameters were used in the target prediction analysis, specifically an alignment score of ≥140 and an energy threshold of ≤−20 kcal/mol (Supplemental Table 3). Gene symbols identified by the miRanda algorithm and predicted to be targeted by at least 10% of the previously unannotated miRNAs were submitted to a comprehensive pathway enrichment analysis using pathDIP v.2.5.21.6, which assesses enrichment of the target genes in pathways obtained from 15 distinct public pathway resources (literature-curated (core) pathway memberships) [30]. In this study, we report all pathways enriched with corrected p values ≤ 0.05.

3. Results

3.1. Discovery of Previously Unannotated miRNAs in ccRCC and Normal Kidney Tissue. Small RNA-sequencing data of ccRCC tumours (unpaired ($n = 502$) and tumours with matched nonmalignant tissue ($n = 71$)) were obtained from cgHUB and processed, and data were analyzed using the miRDeep2 prediction algorithm through the OASIS platform [20, 23]. The raw output of this analysis predicted 96 and 280 unique, previously unannotated miRNAs in nonmalignant tissue and ccRCC tumours, respectively. Manual filtering based on read quality, likelihood of miRNA-like secondary structure, and significant folding values, followed by probing the degree of similarity with known miRNAs using the BLASTn algorithm, and removal of sequences with aberrant GC content, resulted in 40 nonmalignant and 143 tumour previously unannotated miRNAs (Figure 1, Supplemental Table 1).

To validate the occurrence of the previously unannotated miRNAs in kidney tissues, we assessed their expression in small RNA-sequencing data generated from the NCI-60 cell line panel [28]. A miRNA was considered to be expressed in the eight renal cell lines included in the panel if it had a normalized expression value of greater than 0.1 in at least one of the renal cancer cell lines. In these cell lines, 26 (65%) novel miRNAs detected in nonmalignant samples and 102 (71%) miRNAs detected in ccRCC tumours were also expressed, strengthening their confidence as true miRNA sequences in human renal tissues and suggesting possible relevance to kidney function and pathology. Thus, we sought to further examine the role that these unannotated and validated miRNAs expressed in kidney tissues may have in ccRCC tumourigenesis and their potential clinical relevance.

3.2. Previously Unannotated miRNAs Are Deregulated in ccRCC Tumours. Recent evidence suggests that widespread disruption of miRNA-coding gene regulatory networks is common in many cancer types. Perturbation of coding gene expression by miRNA-based regulation can be achieved in tumours through loss of a tumour-suppressive miRNA or the gain of an oncogenic miRNA. Thus, we analyzed the expression of these previously unannotated miRNA sequences in tumours and paired nonmalignant tissue to identify potentially novel oncogenic and tumour-suppressive miRNAs in ccRCC.

After filtering for expression in both groups of samples, 59 previously unannotated miRNA loci were considered to

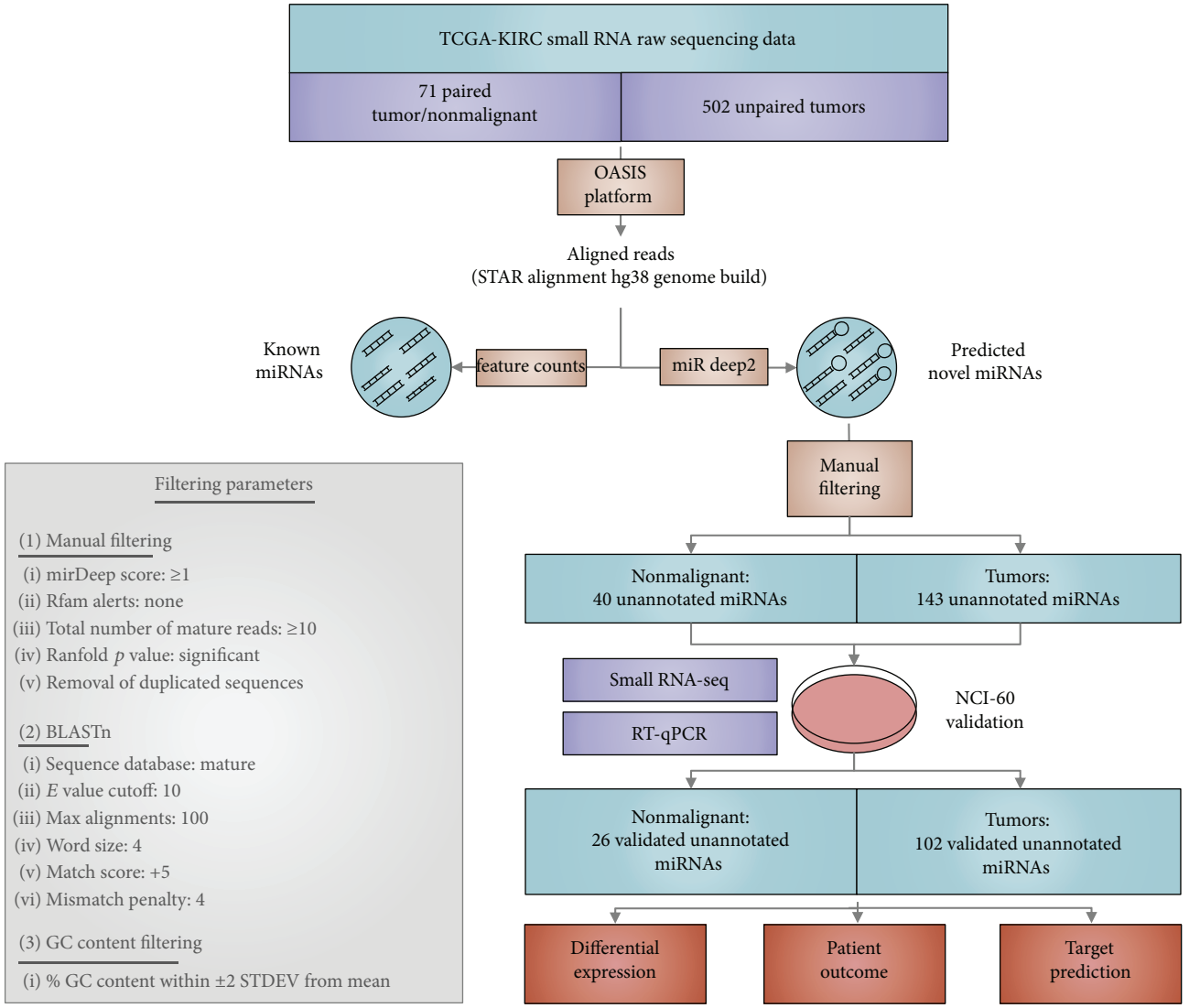
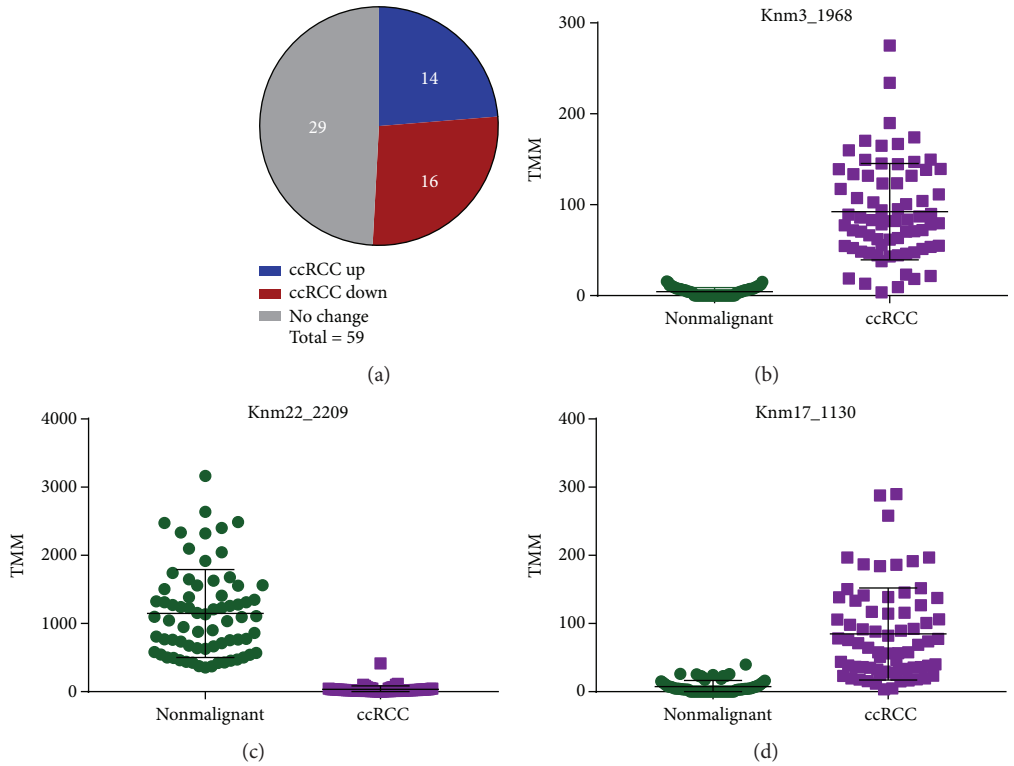


Figure 1: Experimental analysis pipeline. Detailed diagram of the analyses used to generate predicted and validated miRNA sequences previously unannotated in kidney tissues, along with subsequent explorations into their biological relevance.

be expressed in both ccRCC tumours and nonmalignant samples. A Student's t-test analysis revealed 30 of these 59 miRNAs to be significantly differentially expressed between the two groups (BH-$p < 0.05$). An analysis of the average fold change values between paired samples revealed that 14 miRNAs were significantly upregulated in ccRCC tumours, while the other 16 were downregulated (Figure 2). A subset of these miRNAs is of particular interest due to the magnitude of their expression differences between tumour and nonmalignant samples (Figure 2; Supplemental Figure 2). For example, the previously unannotated miRNA Knm22_2209 has an almost complete loss of expression in ccRCC tumours (28-fold downregulated in tumours, BH-$p = 5.3 \times 10^{-23}$), while previously unannotated miRNAs Knm3_1968 and Knm17_1130 show a 100- and 13-fold increase in expression in ccRCC samples, respectively (BH-$p = 1.4 \times 10^{-21}$, 2.6×10^{-14}). These findings serve to not only highlight the potential role of currently unannotated miRNAs in kidney cancer but also warrant investigation into their uses as biological markers of cancer onset.

3.3. Patient Outcome Predicted by Previously Unannotated miRNAs. The differential expression of these unannotated miRNAs in paired ccRCC tumour samples suggests their potential roles in kidney cancer and unexplored clinical utility. As such, we sought to examine whether any of the unannotated miRNA sequences deregulated in ccRCC tumours were associated with patient outcome. We examined differences in patient survival in those with high expression of a miRNA to those with low expression, defined by tertiles. Survival analysis was performed using the Gehan-Breslow-Wilcoxon test on phenotype data obtained from UCSC Xena and the expression profiles of the unannotated miRNA loci. Interestingly, two of the significantly differentially expressed unannotated miRNAs in paired ccRCC tumour samples displayed striking associations with patient survival (Figure 3). Again, the unannotated miRNA Knm22_2209 (28-fold downregulated in ccRCC, BH-$p = 5.3 \times 10^{-23}$) is particularly noteworthy due to the clear correlation between its low expression and poor overall survival (Figure 3(a), $p = 0.045$).



ccRCC ID	Sequence	BH-Cp	Fold-change	Direction
Knm22_2209	agcugccaguugaauaac	5.27181*e*-23	0.035173794	ccRCC down
Knm3_1968	gcagauucccagagugggacag	1.3798*e*-21	100.420012	ccRCC up
Knm12_604	cccagaaaucugguaugccag	1.3614*e*-14	0.268187489	ccRCC down
Knm17_1130	ugagguggaggguugguggga	2.62688*e*-14	13.36362758	ccRCC up
Knm2_1407	auccuagcuugccugagacugu	3.05493*e*-12	3.625000251	ccRCC up
Knm22_1790	gaugccugggaguugcgaucugc	8.92385*e*-11	1.995626467	ccRCC up
Knm18_1200	uuucagacaucagcugcaagc	2.87296*e*-10	2.058185167	ccRCC up
Knm6_2323	ccagggcagacucugagacaga	3.42007*e*-10	5.297669706	ccRCC up
Knm5_2934	gagggaagcauuggacugug	3.31128*e*-08	0.47579602	ccRCC down
Knm7_3340	aauuacagauugucucagaga	8.81302*e*-08	1.743335668	ccRCC up
Knm7_2512	aaaggauucagucggagcuc	7.12109*e*-07	0.762369557	ccRCC down
Knm21_2146	guuaagaugacagagccca	1.29437*e*-06	0.15751302	ccRCC down
Knm6_2419	ggcucuguggcuuaguuggc	1.64455*e*-05	5.505693906	ccRCC up
Knm7_2944	cggcucuggccuccugacccag	1.901*e*-05	0.779267252	ccRCC down
Knm1_272	uuggcugaugaauacugacu	5.32574*e*-05	1.050794948	ccRCC up
Knm21_2196	uagcuuaucgggcugaug	0.000409873	1.51365041	ccRCC up
Knm3_1987	agguuguaggaugcuaaacaga	0.000888669	0.830084461	ccRCC down
Knm22_1819	uugaggucggacaugguggcu	0.00648101	2.728823617	ccRCC up
Knm1_227	cuucagcugcagaacucucagu	0.007147135	0.472351886	ccRCC down
Knm1_317	uagacuggguugacucuggaga	0.008226141	0.702144236	ccRCC down
Knm20_1603	uuugauuggccaguggacagcg	0.008872931	1.758283114	ccRCC up
Knm5_2192	cugacucugaaaucuucuaaau	0.012211986	1.078366811	ccRCC up
Knm11_716	ucucuggccugucuccugcag	0.012975863	0.914637311	ccRCC down
Knm16_1165	ugagggagcaggaugaauagcu	0.015334661	1.032774555	ccRCC up
Knm1_77	cucaggacugggucauucuguu	0.016097135	0.659137709	ccRCC down
Knm11_484	uucgcugggaauucagccucugc	0.02117187	0.763540568	ccRCC down
Knm5_3166	cagauuuccugggcuuaguga	0.028946993	0.37194047	ccRCC down
Knm22_1773	ugguccaacgacaggaguag	0.037765777	0.698249344	ccRCC down
Knm10_427	uacagaucuggguucucuuaga	0.042221184	0.841755482	ccRCC down
Knm12_685	ugccgcacuucucccuacagg	0.046070777	0.234992371	ccRCC down

(e)

Figure 2: Differential expression of previously unannotated miRNAs between nonmalignant and ccRCC samples. (a) Pie chart representing the proportion of miRNAs that are significantly differentially expressed between the samples. (b) Normalized expression of the previously unannotated miRNA Knm3_1968 in individual nonmalignant (green) and ccRCC samples (purple). (c) Normalized expression of Knm22_2209 in individual nonmalignant (green) and ccRCC samples (purple). (d) Normalized expression of Knm17_1130 in individual nonmalignant (green) and ccRCC samples (purple). (e) Summary of differential expression results for each previously unannotated miRNA. BHC-p represents the corrected p value for the differential expression calculated using Student's t-test and Benjamini-Hochberg multiple correction analysis.

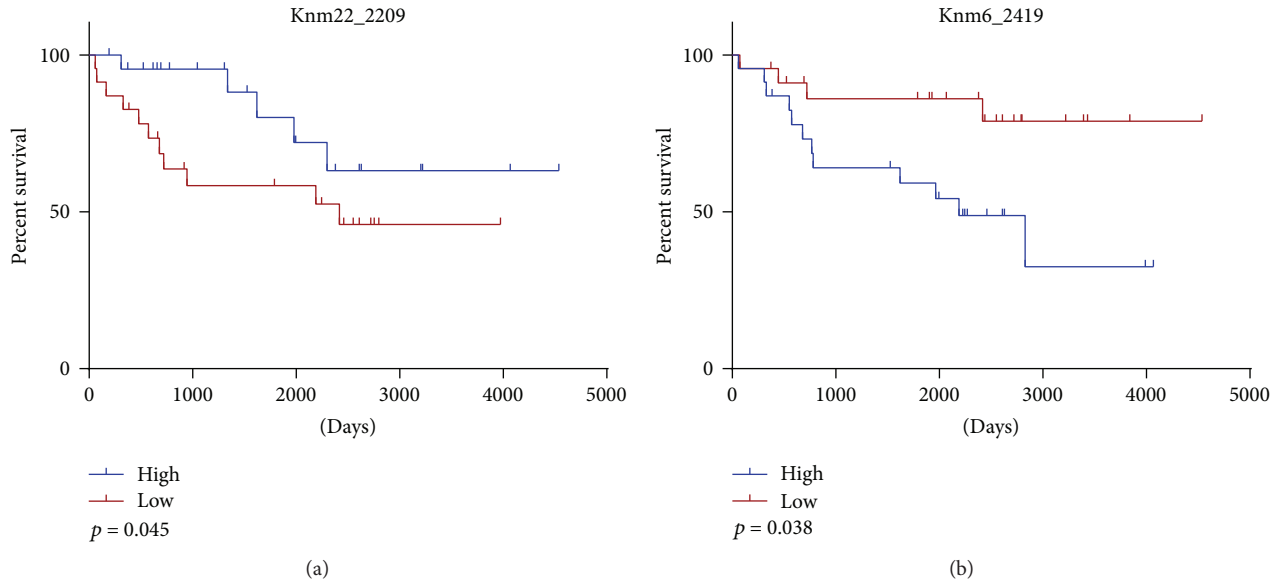


Figure 3: ccRCC patient overall survival predicted by previously unannotated miRNAs. Red lines represent patients with expression of the miRNA in the lower tertile of expression, while blue lines represent the upper tertile of expression. Calculation of p values for Knm22_2209 (a) and Knm6_2419 (b) was performed using the Gehan-Breslow-Wilcoxon test.

Alternatively, the Knm6_2419 miRNA is significantly upregulated in ccRCC tumours (FC = 5.5, BH-$p = 1.6 \times 10^{-5}$), and its high expression is significantly associated with a worsened overall patient survival (Figure 3(b), $p = 0.038$). Taken together, these results emphasize the clinical potential of previously undetected miRNAs.

3.4. Genes and Pathways Targeted by Newly Detected miRNAs. In order to gain deeper insight into the biological relevance of the newly detected miRNAs, target prediction analysis was performed using the miRanda v3.3a algorithm on the set of 30 miRNAs significantly differentially expressed in ccRCC. This algorithm reports potential gene targets for the analyzed sequences based on sequence complementarity and the thermodynamic stability of the complementary RNA duplexes [29]. Predicted target genes with an alignment score ≥ 140 and an energy threshold of ≤ -20 kcal/mol were considered as potential gene targets of our newly detected miRNAs in ccRCC. To generate a broad list of possible gene targets and their associated pathways that may be affected by the regulatory action of the 30 miRNAs, we examined genes that were targeted by at least 10% of these miRNAs (Supplementary Table 3), which were subsequently analyzed for pathway enrichment using pathDIP [30].

Pathway enrichment analysis revealed 63 significantly enriched core pathways (BH-$p \leq 0.05$). Interestingly, many of the enriched pathways are associated with cellular response to extracellular stimuli, organ development, and pathways indirectly associated with cellular metabolism, including the axon guidance pathway (110 gene targets), MAPK signaling cascade (58 gene targets), signaling mediated by FGFR and EGFR (70 and 69 gene targets, resp.), as well as the VEGF pathway (65 gene targets), and insulin-mediated signaling (63 gene targets) (Figure 4) [31, 32]. Considering that kidney cancers have been recognized to associate with altered hypoxic signaling, as well as cellular metabolism [4, 33], our results suggest that these previously unannotated miRNAs may play key roles in the regulation of ccRCC development and progression.

3.5. In Vitro Validation of Previously Unannotated miRNAs Deregulated in ccRCC. In order to experimentally confirm the existence of these transcripts and their consequent deregulation in tumour tissues, we performed RT-qPCR using custom primers specific to Knm3_1968 and Knm17_1130, previously unannotated miRNAs strongly overexpressed in ccRCC samples. We used the TK-10 cell line, which is representative of the ccRCC subtype of renal cancer, as well as HEK-293T, which is an immortalized nonmalignant embryonic kidney cell line commonly used as a control cell line to represent nonmalignant kidney tissue. RT-qPCR results were consistent with expression data from RNA-sequencing, wherein both Knm3_1968 and Knm17_1130 were found to be overexpressed in TK-10 cells relative to HEK-293T cells (average RQ = 8.56; 16.54, resp.; Supplemental Figure 3).

4. Discussion

In this study, we discovered miRNAs previously unannotated in kidney tissues using the miRDeep2 algorithm which represents an increase of 11.5% from the current number of kidney-related miRNAs annotated in miRBase v21 [24]. Collectively, our findings underscore the need to accurately define the landscape of miRNA transcription in human tissues, particularly due to their emerging roles in cell biology and disease.

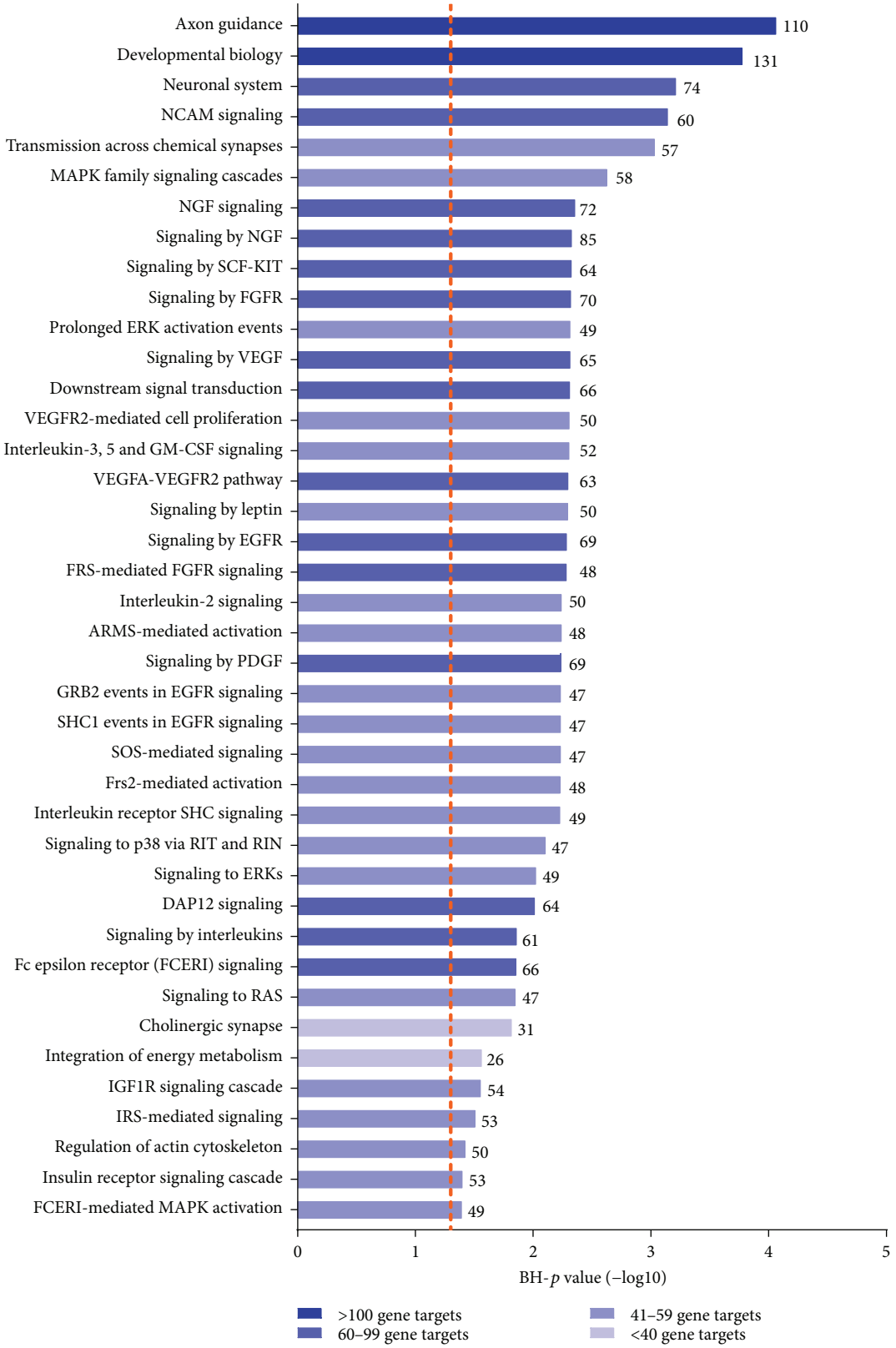


Figure 4: Pathways significantly enriched in protein-coding gene targets of previously unannotated miRNAs. Pathways are plotted according to the significance value associated with their enrichment in predicted miRNA gene targets. Individual genes in each pathway are predicted to be targeted by at least 10% of the previously unannotated miRNAs. Bars are coloured by the number of predicted gene targets involved in the pathway, with the total number of gene targets displayed on top of each bar. The orange dashed line represents the significance cutoff of BH-$p = 0.05$.

The miRDeep2 algorithm integrates artifacts of Dicer-based miRNA processing to generate a score for each predicted miRNA that is reflective of its likelihood of being a true positive [22]. Thus, we further processed the predicted sequences to eliminate sequences that (i) have low true positive detection rates (miRDeep2 score), (ii) share sequencing reads with other ncRNA species (tRNA/rRNA), (iii) do not meet the threshold of significance (ranfold p value), (iv) strongly differ from known miRNAs in G/C content, and (v) are already found in miRBase v21 (BLASTn). Together, these filtering steps enable us to generate a robust set of previously unannotated miRNAs that have low false-discovery rates.

To further validate these miRNAs in kidney tissues, we probed their expression in the NCI-60 panel of cell lines. Although the eight renal cancer cell lines do not reflect the spectrum of ccRCC in patients, they offer the opportunity to assess whether a fraction of the newly identified miRNAs occur in samples independent of TCGA. We found 128 miRNAs newly detected in kidney samples to be expressed in the eight renal cancer cell lines found in the panel. These findings, in tandem with our previous work characterizing the ncRNA transcriptome of the panel [28], serve to confirm the presence of the miRNAs detected in kidney tissues in our analysis and more broadly emphasize the extent of miRNA transcription that has previously escaped detection. Moreover, this highlights the utility of the NCI-60 cell line panel in future explorations of novel miRNA discovery and biology.

After validating their expression in cell lines, we examined the relevance of our panel of miRNAs newly detected in kidney tissues to ccRCC tumourigenesis and disease biology. Strikingly, we found a large proportion of these miRNAs to be significantly differentially expressed between ccRCC tumours and matched nonmalignant tissue. We further observed strongly deregulated miRNA sequences, both up- and downregulated in ccRCC samples. These cases—particularly Knm22_2209, Knm17_1130, and Knm3_1968—suggest that the manipulation of miRNA expression may be a factor in the tumourigenesis of ccRCC tumours. In addition, there are a number of our previously unannotated sequences that were not included in the differential expression analyses as their expression is only detected in either the nonmalignant or ccRCC paired samples. However, the expression of these sequences may be specific to either nonmalignant or cancerous kidney tissue (Supplemental Figure 2), suggesting that these miRNAs may represent exciting candidates for markers of ccRCC development or as therapeutic targets that may display limited off-target effects in normal cells.

The clinical applicability of a subset of the previously unannotated miRNAs is highlighted by our observations that several of the differentially expressed miRNAs are significantly associated with patient outcome. Specifically, the miRNAs Knm22_2209 and Knm6_2419 display significant relationships with the poor outcome of ccRCC patients. While preliminarily, these results highlight the potential impact that novel miRNA discovery can have on both cancer biology and clinical cancer intervention, warranting further characterization of the clinical relevance of previously undetected miRNAs in various pathologies.

The miRNAs discovered in our analyses are observed to target genes critical to both normal and diseased kidney biology. The axon guidance pathway mediates neuronal migration and positioning; in the kidney, this pathway has been shown to play key roles in organ development, in which deletion of critical pathway genes has been shown to lead to after-birth death due to kidney abnormalities [34]. Interestingly, 110 genes in this pathway are each targeted by at least 3 of the previously unannotated miRNAs. Studies have also implicated aberrations in genes from the axon guidance pathway with malignant cell growth and angiogenesis [35]. In fact, kidney cancers have been reported to rely on extensive metabolic reprogramming, in which most of the identified molecular drivers have been shown to participate in pathways related to cellular energy, nutrient metabolism, and oxygen sensing [5, 7]. The newly detected miRNAs in the kidney were found to target several components of the vascular endothelial growth factor (VEGF) pathway, a major regulator of angiogenesis, as well as signaling cascades in response to extracellular stimuli, such as the insulin receptor signaling cascades and FGRF and EGFR signaling.

Specifically, Knm3_1968 is found to target genes such as ACACA, which encodes acetyl-CoA carboxylase alpha, involved in fatty acid synthesis. Additionally, Knm22_2209 targets genes such as RASGRP2, which is associated with the MAPK signaling pathway. Thus, the observation that key pathways in ccRCC tumourigenesis are predicted to be targeted by the miRNAs discovered in our study, particularly axon guidance, metabolic reprogramming, and angiogenesis, emphasizes the potential regulatory role of these transcripts in kidney cancer biology.

Finally, both Knm3_1968 and Knm17_1130 were significantly overexpressed in ccRCC relative to nonmalignant cell lines, as shown by RT-qPCR (Supplemental Figure 3), which confirms the results observed from RNA-sequencing expression data. The expression of our newly detected sequences in cell lines not only confirms their presence in human kidney tissue but is also suggestive of their relevance to ccRCC biology. In light of these observations, further experiments may seek to elucidate the phenotypic consequences of previously unannotated miRNA deregulation.

Taken together, our results suggest that the discovery of previously unannotated miRNAs is an important next step in the exploration of the genomic landscape of tumourigenesis. While performed under the lens of kidney cancer, our findings have implications for cancers of all types, particularly in combination with previous studies that show these types of sequences to have a higher degree of specificity than those that are currently annotated. Although further validation and characterization of these types of sequences are required before these miRNAs can become clinically actionable, our findings highlight the untapped potential of the unexplored areas of the human transcriptome.

5. Conclusions

Here, we present an in-depth and large-scale discovery of previously unannotated miRNA sequences in human kidney

samples. We find that not only are these sequences indeed expressed in both patient samples and cell lines but also their expression may be relevant to normal and tumour biology in these tissues. Several of these newly detected miRNAs are both deregulated in ccRCC tumours and associated with poor patient outcome. Moreover, protein-coding genes and subsequent pathways predicted to be targeted by these sequences, such as the VEGF and EGFR pathways, are critical to kidney tumourigenesis. Taken together, our results provide a novel resource for studying kidney cancer biology and underline a need for further identification of miRNAs that have eluded previous detection methods. Through the discovery of previously unannotated sequences, future studies may uncover novel players in tumour biology that may result in better characterization of ccRCC molecular drivers and direct new diagnostic and treatment strategies in the clinic.

Data Availability

All data were accessed from the Genomic Data Commons platform through the National Institute of Health (https://gdc.cancer.gov/).

Conflicts of Interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by grants from the Canadian Institutes for Health Research (CIHR FRN-143345). Adam P. Sage, Brenda C. Minatel, Erin A. Marshall, Victor D. Martinez, and Katey S. S. Enfield were supported by scholarships from the University of British Columbia. Adam P. Sage, Erin A. Marshall, and Katey S. S. Enfield were also supported by scholarships from CIHR.

Supplementary Materials

Supplementary 1. Figure 1: genome-wide distribution of previously unannotated miRNA sequences. Purple bars represent $\log_2$ (total read counts) for previously unannotated miRNAs detected in ccRCC samples, while green bars indicate those detected in nonmalignant tissues. The outer circle represents all chromosomes from the human genome.

Supplementary 2. Figure 2: unsupervised hierarchical clustering analysis of differentially expressed previously unannotated miRNA sequences. Both samples and features were clustered using average linkage and Pearson correlation. Standardized expression is represented shown from high (red) to low (green), and samples are coloured according to their malignant status (blue represents nonmalignant and red represents tumour tissue).

Supplementary 3. Figure 3: fold change (RQ) values from RT-qPCR analysis of Knm3_1968 and Knm17_1130 in TK-10 ccRCC cells relative to nonmalignant HEK-293T cells.

Supplementary 4. Table 1: summary of OASIS output and subsequent filtering steps of predicted novel miRNAs detected in kidney tissues. This table lists the raw output of the computational analysis deducing previously undetected miRNAs from RNA-sequencing data of nonmalignant (NM) and clear cell renal cell carcinoma samples (TP) (columns A–E) and the results from further filtering steps (columns F–J) as described in the "Materials and Methods." Table 2: cell line characteristics and next-generation sequencing information for the NCI-60 renal cell lines used to validate the discovery of previously unannotated miRNA sequences. Table 3: predicted target genes of the newly detected miRNAs differentially expressed in clear cell renal cell carcinoma. This table lists the output values of the miRanda target prediction algorithm used to predict the protein-coding genes targeted by previously undetected miRNAs in kidney samples, as described in "Materials and Methods."

References

[1] Canadian Cancer Statistics Advisory Committee, *Canadian Cancer Statistics 2018*, Canadian Cancer Society, Toronto, ON, Canada, 2018.

[2] W. H. Chow, L. M. Dong, and S. S. Devesa, "Epidemiology and risk factors for kidney cancer," *Nature Reviews Urology*, vol. 7, no. 5, pp. 245–257, 2010.

[3] Y. Riazalhosseini and M. Lathrop, "Precision medicine from the renal cancer genome," *Nature Reviews Urology*, vol. 12, no. 11, pp. 655–666, 2016.

[4] W. M. Linehan, R. Srinivasan, and L. S. Schmidt, "The genetic basis of kidney cancer: a metabolic disease," *Nature Reviews Urology*, vol. 7, no. 5, pp. 277–285, 2010.

[5] The Cancer Genome Atlas Research Network, "Comprehensive molecular characterization of clear cell renal cell carcinoma," *Nature*, vol. 499, no. 7456, pp. 43–49, 2013.

[6] I. J. Frew and H. Moch, "A clearer view of the molecular complexity of clear cell renal cell carcinoma," *Annual Review of Pathology: Mechanisms of Disease*, vol. 10, no. 1, pp. 263–289, 2015.

[7] H. I. Wettersten, O. A. Aboud, P. N. Lara, and R. H. Weiss, "Metabolic reprogramming in clear cell renal cell carcinoma," *Nature Reviews Nephrology*, vol. 13, no. 7, pp. 410–419, 2017.

[8] R. Fisher, M. Gore, and J. Larkin, "Current and future systemic treatments for renal cell carcinoma," *Seminars in Cancer Biology*, vol. 23, no. 1, pp. 38–45, 2013.

[9] P. H. Tan, L. Cheng, N. Rioux-Leclercq et al., "Renal tumors: diagnostic and prognostic biomarkers," *The American Journal of Surgical Pathology*, vol. 37, no. 10, pp. 1518–1531, 2013.

[10] K. S. S. Enfield, L. A. Pikor, V. D. Martinez, and W. L. Lam, "Mechanistic roles of noncoding RNAs in lung cancer biology and their clinical implications," *Genetics Research International*, vol. 2012, Article ID 737416, 16 pages, 2012.

[11] R. C. Lee, R. L. Feinbaum, and V. Ambros, "The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14," *Cell*, vol. 75, no. 5, pp. 843–854, 1993.

[12] F. Calore, F. Lovat, and M. Garofalo, "Non-coding RNAs and cancer," *International Journal of Molecular Sciences*, vol. 14, no. 8, pp. 17085–17110, 2013.

[13] P. Trionfini, A. Benigni, and G. Remuzzi, "MicroRNAs in kidney physiology and disease," *Nature Reviews Nephrology*, vol. 11, no. 1, pp. 23–33, 2015.

[14] A. K. Marrone, D. B. Stolz, S. I. Bastacky, D. Kostka, A. J. Bodnar, and J. Ho, "*MicroRNA-17~92* is required for nephrogenesis and renal function," *Journal of the American Society of Nephrology*, vol. 25, no. 7, pp. 1440–1452, 2014.

[15] Y. Gong, V. Renigunta, N. Himmerkus et al., "Claudin-14 regulates renal Ca^{++} transport in response to CaSR signalling via a novel microRNA pathway," *The EMBO Journal*, vol. 31, no. 8, pp. 1999–2012, 2012.

[16] Q. Ma, Z. Peng, L. Wang et al., "miR-19a correlates with poor prognosis of clear cell renal cell carcinoma patients via promoting cell proliferation and suppressing PTEN/SMAD4 expression," *International Journal of Oncology*, vol. 49, no. 6, pp. 2589–2599, 2016.

[17] E. Londin, P. Loher, A. G. Telonis et al., "Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 10, pp. E1106–E1115, 2015.

[18] C. Backes, B. Meder, M. Hart et al., "Prioritizing and selecting likely novel miRNAs from NGS data," *Nucleic Acids Research*, vol. 44, no. 6, article e53, 2016.

[19] V. D. Martinez, E. A. Vucic, K. L. Thu et al., "Unique somatic and malignant expression patterns implicate PIWI-interacting RNAs in cancer-type specific biology," *Scientific Reports*, vol. 5, no. 1, article 10423, 2015.

[20] J. An, J. Lai, M. L. Lehman, and C. C. Nelson, "miRDeep* an integrated application tool for miRNA identification from RNA sequencing data," *Nucleic Acids Research*, vol. 41, no. 2, pp. 727–737, 2013.

[21] P. P. Gardner, J. Daub, J. G. Tate et al., "Rfam: updates to the RNA families database," *Nucleic Acids Research*, vol. 37, no. - Database, pp. D136–D140, 2009.

[22] C. Camacho, G. Coulouris, V. Avagyan et al., "BLAST+: architecture and applications," *BMC Bioinformatics*, vol. 10, no. 1, p. 421, 2009.

[23] V. Capece, J. C. Garcia Vizcaino, R. Vidal et al., "Oasis: online analysis of small RNA deep sequencing data," *Bioinformatics*, vol. 31, no. 13, pp. 2205–2207, 2015.

[24] A. Kozomara and S. Griffiths-Jones, "miRBase: annotating high confidence microRNAs using deep sequencing data," *Nucleic Acids Research*, vol. 42, no. D1, pp. D68–D73, 2014.

[25] Y. Liao, G. K. Smyth, and W. Shi, "featureCounts: an efficient general purpose program for assigning sequence reads to genomic features," *Bioinformatics*, vol. 30, no. 7, pp. 923–930, 2014.

[26] M. N. McCall, M.-S. Kim, M. Adil et al., "Toward the human cellular microRNAome," *Genome Research*, vol. 27, no. 10, pp. 1769–1781, 2017.

[27] C. Wake, A. Labadorf, A. Dumitriu et al., "Novel microRNA discovery using small RNA sequencing in post-mortem human brain," *BMC Genomics*, vol. 17, no. 1, p. 776, 2016.

[28] E. A. Marshall, A. P. Sage, K. W. Ng et al., "Small non-coding RNA transcriptome of the NCI-60 cell line panel," *Scientific Data*, vol. 4, article 170157, 2017.

[29] A. J. Enright, B. John, U. Gaul, T. Tuschl, C. Sander, and D. S. Marks, "MicroRNA targets in Drosophila," *Genome Biology*, vol. 5, no. 1, article R1, 2003.

[30] S. Rahmati, M. Abovsky, C. Pastrello, and I. Jurisica, "pathDIP: an annotated resource for known and predicted human gene-pathway associations and pathway enrichment analysis," *Nucleic Acids Research*, vol. 45, no. D1, pp. D419–D426, 2017.

[31] W. Zhang and H. T. Liu, "MAPK signal pathways in the regulation of cell proliferation in mammalian cells," *Cell Research*, vol. 12, no. 1, pp. 9–18, 2002.

[32] L. R. Harskamp, R. T. Gansevoort, H. van Goor, and E. Meijer, "The epidermal growth factor receptor pathway in chronic kidney diseases," *Nature Reviews Nephrology*, vol. 12, no. 8, pp. 496–506, 2016.

[33] C. H. Lee and R. J. Motzer, "Kidney cancer in 2016: the evolution of anti-angiogenic therapy for kidney cancer," *Nature Reviews Nephrology*, vol. 13, no. 2, pp. 69-70, 2017.

[34] U. Grieshammer, le Ma, A. S. Plump, F. Wang, M. Tessier-Lavigne, and G. R. Martin, "SLIT2-mediated ROBO2 signaling restricts kidney induction to a single site," *Developmental Cell*, vol. 6, no. 5, pp. 709–717, 2004.

[35] P. Mehlen, C. Delloye-Bourgeois, and A. Chedotal, "Novel roles for slits and netrins: axon guidance cues as anticancer targets?," *Nature Reviews Cancer*, vol. 11, no. 3, pp. 188–197, 2011.

Hindawi
International Journal of Genomics
Volume 2018, Article ID 5383517, 6 pages
https://doi.org/10.1155/2018/5383517

Research Article

Cerebellar lncRNA Expression Profile Analysis of SCA3/MJD Mice

Zhe Long,[1,2] Tianjiao Li,[1] Zhao Chen,[1] Yun Peng,[1] Chunrong Wang,[1] Xiaocan Hou,[1] Hongyu Yuan,[1] Puzhi Wang,[1] Yue Xie,[1] Lang He,[1] Xin Zhou,[1] Huirong Peng,[1] Rong Qiu,[3] Kun Xia,[4] Beisha Tang,[1,4,5,6,7,8,9] and Hong Jiang[1,4,6,10]

[1]*Department of Neurology, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China*
[2]*Sydney Medical School and the Brain & Mind Institute, The University of Sydney, 94 Mallett St, Camperdown, NSW 2050, Australia*
[3]*School of Information Science and Engineering, Central South University, Changsha, Hunan 410083, China*
[4]*Laboratory of Medical Genetics, Central South University, Changsha, Hunan 410078, China*
[5]*National Clinical Research Center for Geriatric Diseases, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China*
[6]*Key Laboratory of Hunan Province in Neurodegenerative Disorders, Central South University, Changsha, Hunan 410008, China*
[7]*Parkinson's Disease Center of Beijing Institute for Brain Disorders, Beijing 100069, China*
[8]*Collaborative Innovation Center for Brain Science, Shanghai 200032, China*
[9]*Collaborative Innovation Center for Genetics and Development, Shanghai 200433, China*
[10]*Xinjiang Medical University, Xinjiang 830011, China*

Correspondence should be addressed to Hong Jiang; jianghong73868@126.com

Received 1 March 2018; Revised 9 April 2018; Accepted 27 May 2018; Published 25 June 2018

Academic Editor: Elena Pasyukova

Spinocerebellar ataxia type 3 (SCA3) or Machado-Joseph disease (MJD) is the most common autosomal dominant spinocerebellar ataxia in China with highly clinical heterogeneity, such as progressive cerebellar ataxia, dysarthria, pyramidal signs, external ophthalmoplegia, dysphagia, and distal muscle atrophy. It is caused by the abnormal expansion of CAG repeats in a coding region of *ATXN3*. However, by focusing on the *ATXN3* itself cannot fully explain the heterogeneous clinical features of SCA3/MJD. With the discovery of the increasing number of long noncoding RNAs (lncRNAs) that are believed to be involved in spinocerebellar ataxia type 8 (SCA8) and Huntington disease (HD), we wonder whether the lncRNAs are differentially expressed in the SCA3/MJD patients compared to the nonpatients. As the first step, we used lncRNA-Seq to investigate differential expression of the lncRNAs in the SCA3/MJD mice. Two known lncRNAs, n297609 and n297477, and a novel lncRNA TCONS_00072962 have been identified in SCA3/MJD mice with abnormal expression. The first discovery of the novel lncRNA TCONS_00072962 enriched the lncRNA expression profile in the SCA3/MJD mouse model.

1. Introduction

PolyQ diseases is a group of disorders caused by CAG repeat expansions within the, respectively, responsible genes, including Huntington disease (HD), dentatorubral-pallidoluysian atrophy (DRPLA), spinocerebellar ataxias (SCA1, SCA2, SCA3/Machado-Joseph disease, SCA6, SCA7, and SCA17) [1–3], and the recently discovered Huntington disease-like 2 (HDL2) [4, 5]. Among these, the SCA3/MJD is an autosomal dominantly inherited disorder with high clinical heterogeneity, such as progressive cerebellar ataxia, dysarthria, pyramidal signs, external ophthalmoplegia, dysphagia, and distal muscle atrophy, with wide range of age of onset (AO) from 4 to 75 years old [6, 7].

The SCA3/MJD, the most common case, accounts for 62.64% of autosomal dominant spinocerebellar ataxia in China [8]. The abnormal expansion of CAG in the causative gene ATXN3 coding region causes SCA3/MJD. Healthy individuals usually have 12–40 CAG repeats, while SCA3/MJD patients over 51 repeats [9, 10]. The abnormally translated

polyQ tract leads to a conformational change in ATXN3, resulting in alternations of protein properties, including stability, subcellular location, and easier aggregation [11]. These alternations further lead to loss or gain of function and cause pathogenic effects. To explain the toxic effects, several hypotheses of pathogenic mechanisms, not mutually exclusive, have been presented, including aggregate formation [1, 11–13], disturbance of cellular protein and Ca2+ homeostasis [13–15], dysregulation of transcription [15, 16], axonal transport deficits [17, 18], impairment of mitochondrial function [15, 19, 20], and abnormal neuronal signalling [11].

Long noncoding RNA (lncRNA) is defined the nontranslatable RNA with the length of 200 nucleotides or above. The lncRNAs used to be regarded as the transcriptional "noise," the products of RNA polymerase II transcription, and did not have the biological function. However, the emerging evidence has proved their significant roles in the regulation of gene transcription, posttranscriptional regulation, and epigenetic regulation [21, 22]. Previous studies suggested that lncRNAs regulate the gene expression and transcriptional processes by several different functional mechanisms. Some show function as transcriptional regulation in cis or trains, some as an organization of nuclear domains, and others as regulation of proteins or RNA molecules. All the evidence indicated that lncRNAs have great potential to impact physiological and pathological processes. Furthermore, it has been found that some transcripts of lncRNA encode small proteins [23], making the noncoding inappropriate any longer to name this class of RNA.

In recent years, accumulating studies have found that lncRNAs are associated with neurodegenerative diseases. Spinocerebellar ataxia type 8 (SCA8), a kind of slowly progressive ataxia, is caused by the abnormal expansion of (CTG)n within the responsible gene ATXN8. A study proposes that the pathogenesis of SCA8 involves both protein and RNA gain-of-function mechanisms. (CTG)n-expanded ATXN8 encodes a pathogenic protein, and the antisense strand encodes CUG-enriched lncRNA ATXN8OS which is deposited in the nucleus and activates alternative splicing, resulting in an alternation of the expression of GABA-A transport factor 4 (GAT4/Gabt4) and finally loss of the GABAergic inhibition [24].

In a separate study, the expression of lncRNA was compared between the normal brain tissue and the brain tissue of patients with Huntington disease. A total of 35 upregulated and 146 downregulated lncRNA molecules were identified, and NEAT1 was selected by Bioinformatics. Based on the cell-level experiments, it was found that overexpression of NEAT1 was significantly resistant to H_2O_2-induced cellular damage, providing a new potential strategy for clinical treatment of the Huntington's disease [25].

To further explore the pathogenesis of SC3/MJD at RNA level, the lncRNAs specifically expressed in SCA3/MJD mice were investigated in this study.

2. Materials and Methods

2.1. SCA3/MJD Mice. SCA3/MJD mouse model (B6; CBA-Tg (ATXN3*) 84.2Cce/IbezJ; ID: 012075) from Jackson Laboratory was used, and the second generation was used in this study. The CAG repeats in the first generation mice are 84, and the *ATXN3* gene is widely expressed in various organs of the body, including the cerebellum, cerebral cortex, heart, lung, spleen, liver, and skeletal muscle [26, 27]. The SCA3/MJD adult mice (32 weeks old) of the second generation, in which carrying *ATXN3* positive rate is about 50%, and comparable age, number, and weight wild-type mice were used for experimental analysis. The study was approved by the Ethics Committee in Xiangya Hospital of Central South University.

2.2. Validation of Genotype of SCA3/MJD Mice. Validation of genotype was conducted in the second generation. Polymerase chain reaction (PCR), agarose gel electrophoresis, and capillary electrophoresis sequencing were used for genotype validation. Genomic DNA was extracted from mice tails. CAG repeats were amplified using a pair of primers 5′-CCAGTGACTACTTTGATTCG-3′ (forward) 5′-TGGCCTTTCACATGGATGTGAA-3′ (reverse). The amplification reactions contained 1 μL genomic DNA (50 ng/μL), 0.2 μL rTaq DNA polymerase (Takara, Japan), 0.2 μL dNTPs, 0.2 μL of each primer (100 ng/μL), 7.2 μL sterile water, and 1.0 μL 10x buffer (TaKaRa, Japan), for a total of 10 μL. The amplification was performed in Mastercyclers (Eppendorf AG, 22331 Hamburg, Germany) under the following conditions: initial denaturation at 95.0°C for 5 minutes, followed by 38 cycles of 95.0°C for 30 seconds, 59.0°C for 30 seconds, and 72.0°C for 30 seconds. PCR products were detected by 1% agarose gel electrophoresis (120 v, 30 min), and the results of PCR amplification were observed on imaging system after 15 minutes of ethidium bromide (EB) staining. Capillary electrophoresis sequencing was used for testing the repeats number of (CAG)n and performed on ABI 3730XL DNA Analyzer (Applied Biosystems, Foster City, CA, USA).

2.3. Validation of Phenotype of SCA3/MJD Mice. The phenotype was validated using the footprint and rotating tests. For footprint pattern analysis, the hind paws of mice were painted with black ink and the forepaws were painted with red ink. The mice walked along a narrow corridor paved with white paper. Pretraining was conducted for one week before the formal test. Mice were tested three times with 5-minute intervals. Stride length, hind paw width, front paw width, and front/hind footprint overlap were measured. For rotation, mice were placed on a rotating rod and must maintain its balance. The interval from the start of the rod rotating to the mice falling from the rotating rod was recorded. Mice were tested on separate trials at fixed speeds including 10 r/min and 20 r/min.

2.4. lncRNA-Seq. lncRNA-Seq, a high-throughput sequencing, was performed in BGI. After extracting the total RNA from mice cerebellum (three SCA3/MJD mice versus three wild-type mice), mRNA and noncoding RNAs are enriched by removing rRNA from the total RNA. By using the fragmentation buffer, the mRNAs and noncoding RNAs are fragmented into short fragments (about 200~500 nt), then

Table 1: The results of rotation test.

Rotation speed	Time (seconds)	Wild-type mice ($n = 6$)	SCA3/MJD mice ($n = 6$)	p value
10 r/min	Mean ± SD	250.067 ± 9.487	104.400 ± 8.902	0.023
20 r/min	Mean ± SD	196.467 ± 8.126	34.600 ± 6.710	0.002

The wild-type mice performed much better than the SCA3/MJD mice in the rotation test.

the first-strand cDNA is synthesized by random hexamer-primer using the fragments as templates, and dTTP is substituted by dUTP during the synthesis of the second strand. Short fragments are purified and resolved with EB buffer for end reparation and single nucleotide A (adenine) addition. After that, the short fragments are connected with adapters, and then the second strand is degraded using UNG (uracil-N-glycosylase) finally [28]. After agarose gel electrophoresis, the suitable fragments are selected for the PCR amplification as templates. During the QC steps, Agilent 2100 Bioanalyzer and ABI StepOnePlus real-time PCR system are used in quantification and qualification of the sample library. At last, the library could be sequenced using Illumina HiSeqTM 2000 or other sequencers when necessary. Differential expression analysis for both predicted novel lncRNA and lncRNA from the database has proceeded. It was compared between SCA3/MJD and wild-type mice (one by one, grouped randomly) through the Cuffdiff software to calculate the FPKM value of the gene or transcript in both samples and to detect the presence of differential expression. Also, group differential expression analysis was also performed by NOIseq method.

2.5. Quantitative Real-Time PCR. The total RNA from mice cerebellum (six SCA3/MJD mice versus six wild-type mice) was reversely transcribed to cDNA with a kit (Thermo Scientific, RevertAid First Strand cDNA #K1622). The 44 most differentially expressed lncRNAs were screened for further validation by qRT-PCR assays (Maxima SYBR Green qPCR Master Mix, CFX96, Bio-Rad, USA). β-Actin was used as an internal reference in the qRT-PCR analyses. The primers (see Table S1 in the Supplementary Material) were designed using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/). qRT-PCR assay was performed in triplicate in a volume of 20 μL containing 1 μL of cDNA. The relative expression level of lncRNA was calculated using the $2^{-\Delta\Delta Ct}$ method and $Ct > 35$ were excluded. A Wilcoxon rank sum test was used for statistical analyses, and $p < 0.05$ was considered statistical significance.

2.6. Bioinformatics Analysis. We conducted some biological analysis of the differentially expressed lncRNAs by search and comparison of databases such as NONCODE v5 (http://www.noncode.org/index.php), FANTOM5 (http://fantom.gsc.riken.jp/5/), STRING v.10.0 (http://version10.string-db.org/), and Gene Ontology Consortium (http://www.geneontology.org/) to predict the location, distribution, and function of these lncRNAs.

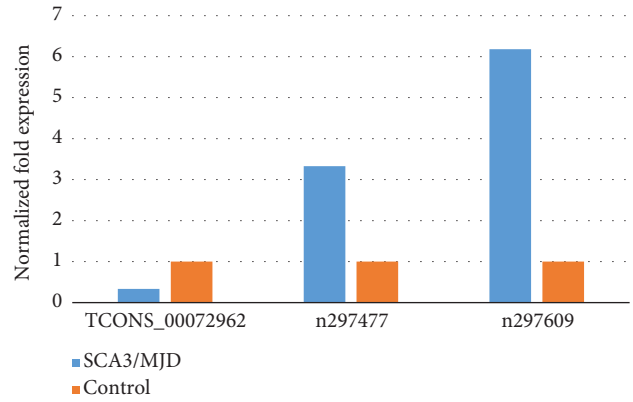


Figure 1: The expression level of the three lncRNAs.The expression level of three lncRNAs was different between wild-type and SCA3/MJD mice with statistical significance. The two known lncRNAs, n297477 and n297609, were upregulated in the cerebellum of SCA3/MJD mice. The expression level was increased by 3.329 times ($p = 0.016$) and 6.182 times ($p = 0.041$), respectively. The novel lncRNA TCONS_00072962 was downregulated ($p = 0.036$) in the cerebellum of SCA3/MJD mice, which was nearly one-third of that in wild-type mice.

3. Results

3.1. Genotyping and Phenotyping of the SCA3/MJD Mice. The SCA3/MJD mice that carry the ATXN3 were genotyped and phenotyped. As a result, the ATXN3 carrying rate was about 50% in the second generation of SCA3/MJD mice, by Jackson Laboratory's report. The number of CAG repeats was 84 or 83 in the second generation SCA3/MJD mice. In rotation test, the interval of mice keeping balance on the rotating rod was significantly different between the SCA3/MJD and wild-type mice no matter the rotation speed was 10 r/min or 20 r/min (see Table 1), indicating that the balance and motor abilities of the SCA3/MJD mice were much worse relative to the control mice. Morphologically, wider hind paw width was evident in SCA3/MJD mice.

Abnormal gait was observed by analyzing the footprint pattern. In contrast to the control mice's straight line with regular alternating gait, the SCA3/MJD mice showed unstable movements in a way that weaved from side to side when walked along the narrow corridor. Altogether suggest that the SCA3/MJD mice have similar clinical manifestations with the SCA3/MJD patients.

3.2. lncRNA-Seq. By using RNA-seq, 10,443 of novel and 13,395 of known lncRNAs were detected in 3 of the SCA3/MJD and 3 of the WT mice. One by one differential expression analysis showed total 2964 upregulated and 4376 downregulated lncRNAs, respectively, in the three experimental groups. More specifically, there were 745, 776, and 1250 of the upregulated lncRNAs in three groups, in contrast to the significantly increased numbers of 1285, 1065, and 1600, the downregulated lncRNAs, respectively, in three groups. Further group differential expression analysis found 193 of the upregulated and 467 of the downregulated lncRNAs in the SCA3/MJD mice. The lncRNAs with differential

Table 2: The summary of genetic association.

Gene name	Associated gene name	Proteins of associated gene
ATXN3	VCP	Valosin containing protein
	UBC	Ubiquitin C
	KCTD10	Potassium channel tetramerisation domain containing 10
	RAD23A	RAD23 homolog A
	RAD23B	RAD23 homolog B
	PARK2	Parkinson protein 2
	USP13	Ubiquitin specific peptidase 13
	UBE4B	Ubiquitination factor E4B
	STUB1	STIP1 homology and U-box containing protein 1
	SERPINC1	Serpin peptidase inhibitor, clade C (antithrombin), member 1
UBC	PSMD4	Proteasome (prosome, macropain) 26S subunit, non-ATPase 4
	HSP90AA1	Heat shock protein 90 kDa alpha (cytosolic), class A member 1
	HGS	Hepatocyte growth factor-regulated tyrosine kinase substrate
	PSMC2	Proteasome (prosome, macropain) 26S subunit, ATPase 2
	TSG101	Tumor susceptibility gene 101
	UBE2D2	Ubiquitin-conjugating enzyme E2D 2
	PSMD14	Proteasome (prosome, macropain) 26S subunit, non-ATPase 14
	CUL1	Cullin 1
	RPS27A	Ribosomal protein S27a (156 aa)

expression identified by the one by one differential expression analysis in two or three groups were chosen for further validation. The lncRNAs were sorted based on their FPKM and p values in the group differential expression analysis. For both differential expression analyses, a total of 44 lncRNAs were chosen for further experiments because they were statistically significant.

3.3. qRT-PCR. lncRNA number TCONS_00031478 was excluded due to Ct > 35. In the rest 43 of the chosen lncRNAs, 3 of the lncRNAs were validated to be differential expression between SCA3/MJD and WT mice cerebellum. It turned out that two of the three are the known lncRNAs number n297477 ($p = 0.016$) and number n297609 ($p = 0.041$) (NONCODE v5.0: http://www.noncode.org/), and the remaining one belonged to a novel lncRNA (number TCONS_00072962) ($p = 0.036$). $2^{-\triangle\triangle Ct}$ method was used for calculating the relative expression level of each lncRNA. When compared with wild-type mice, the two known lncRNAs, n297477 and n297609, were upregulated by 3.329-fold and 6.182-fold, respectively, in the cerebellum of SCA3/MJD mice (see Figure 1), while the expression level of novel lncRNA TCONS_00072962 was downregulated in the cerebellum of SCA3/MJD mice, which was nearly one-third of that in control mice.

4. Discussion

In this study, 2964 upregulated lncRNAs and 4376 downregulated lncRNAs were identified specifically in the SCA3/MJD mice using lncRNA-Seq analysis. Additionally, differentially expressed three lncRNAs, including one novel lncRNA and two known lncRNAs, were further characterized.

The lncRNA n297477 is transcribed from the antisense strand of the chr11: 6270375–6271530, which is highly expressed in mouse heart, hippocampus, liver, lung, spleen, and thymus. According to the records in database FANTOM5, n297477 is considered one of the transcripts of the *TMED4* gene whose promoter is located in the sense strand at position 142–392 (chr11: 6270517–6270767) and TATA box starts at 377 (chr11: 6270712) (http://www-bimas.cit.nih.gov/molbio/proscan/). According to the database STRING v.10.0, Tmed4 and ubiquitin C may be functional partners to each other, whereas the ubiquitin C might be a functional partner of ataxin3 encoded by the SCA3/MJD pathogenic gene *ATXN3* (Table 2; http://version10.string-db.org/cgi/network.pl?all_channels_on=1&block_structure_pics_in_bubbles=0&direct_neighbor=1&hide_disconnected_nodes=0&hide_node_labels=0&network_display_mode=svg&network_flavor=evidence&targetmode=proteins&identifier=9606.ENSP00000376965; http://version10.string-db.org/cgi/network.pl?all_channels_on=1&block_structure_pics_in_bubbles=0&direct_neighbor=1&hide_disconnected_nodes=0&hide_node_labels=0&network_display_mode=svg&network_flavor=evidence&targetmode=proteins&identifier=9606.ENSP00000404042).

Given that the ubiquitin-proteasome system (UPS) is involved in the pathogenesis of SCA3/MJD [29], it is plausible to speculate that lncRNA may participate in gene expression regulation when it is located close to the transcription start site of the promoter region. Since the lncRNA n297477 and TMED4 gene starting site are separated by 16 bp in the middle, and the gene promoter region is located in the lncRNA coding region, it is highly possible that the n297477 participates in TMED4 expression regulation. We assume that the altered expression of

the TMED4 affects its interaction with UBC, which affects the efficiency of ubiquitin C in the ubiquitin-proteasome system. Thus, the lncRNA n297477 may be involved in the regulation of ataxin-3 protein degradation by regulating UPS. The elevated expression of n297477 in the SCA3/MJD mice may serve as a response to the abnormal degradation of toxic ataxin-3.

The lncRNA n297609 is transcribed from the antisense strand of the chr9: 44047488–43856016, which is highly expressed in mouse heart, hippocampus, liver, lung, spleen, and thymus. According to the records in database FANTOM5, n297609 is considered one of the transcripts of the *THY1* gene which is highly homologous to humans. The lncRNA n297609 mainly participates in the protein phosphorylation (GO: 0006468), cell adhesion (GO: 0007155), cellular response to heat (GO: 0034605), positive regulation of transcription (GO: 0045893), and negative regulation of cell migration (GO: 0030336). Previous studies pointed out that protein casein kinase 2- (CK2-) dependent phosphorylation can control the stability, nuclear localization, and aggregation of ataxin-3 [30]. Thus, it is logical to assume that the lncRNA n297609 may be involved in biological processes, such as protein phosphorylation, and protein casein kinase 2- (CK2-) dependent phosphorylation plays a crucial role in SCA3/MJD pathophysiology [30]. However, this assumption requires further validation.

Moreover, we also found a novel lncRNA TCONS_00072962 with the genomic location at chr7:119737479–119737966, contributing to the expansion of lncRNA expression profiles of the mouse.

5. Conclusions

In summary, we used lncRNA-Seq to profile cerebellar expression in SCA3/MJD mice and identified three potential lncRNAs significantly associated with the disease. These identified lncRNAs will be beneficial for the further understanding of cerebellum gene coexpression network correlating with disease progression. Furthermore, investigation on the lncRNA-associated neuroprotective factors remains to be necessary for the elucidation of the therapeutic target implication.

Data Availability

The location of all lncRNAs is from the mm10 database.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Zhe Long and Tianjiao Li contributed equally to this work.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (nos. 81471156 and 81771231 to Hong Jiang; nos. 81430023 and 81130021 to Beisha Tang; no. 31401135 to Rong Qiu), National Key Research and Development Program of China (nos. 2016YFC0901504 and 2016YFC0905100 to Hong Jiang; no. 2016YFC1306000 to Beisha Tang), Clinical Research Funds of Xiangya Hospital (no. 2014L03 to Hong Jiang), the Clinical and Rehabilitation Fund of Peking University Weiming Biotech Group (no. xywm2015I10 to Hong Jiang), Youth Foundation of Xiangya Hospital (no. 2017Q03 to Zhao Chen), and Independent Exploration and Innovation Project of Graduate Student of Central South University (no. 1053320170177 to Tianjiao Li). The authors want to give the sincerest thanks to all those who helped in the study.

Supplementary Materials

Table S1: the supplementary material included 44 lncRNA locations, dysregulation (upregulated/downregulated), and primers used for qPCR amplification. *(Supplementary Materials)*

References

[1] A. Ashkenazi, C. F. Bento, T. Ricketts et al., "Polyglutamine tracts regulate beclin 1-dependent autophagy," *Nature*, vol. 545, no. 7652, pp. 108–111, 2017.

[2] S. Esteves, S. Duarte-Silva, and P. Maciel, "Discovery of therapeutic approaches for polyglutamine diseases: a summary of recent efforts," *Medicinal Research Reviews*, vol. 37, no. 4, pp. 860–906, 2016.

[3] H. L. Paulson, V. G. Shakkottai, H. B. Clark, and H. T. Orr, "Polyglutamine spinocerebellar ataxias - from genes to potential treatments," *Nature Reviews. Neuroscience*, vol. 18, no. 10, pp. 613–626, 2017.

[4] I. Paradisi, V. Ikonomu, and S. Arias, "Huntington disease-like 2 (HDL2) in Venezuela: frequency and ethnic origin," *Journal of Human Genetics*, vol. 58, no. 1, pp. 3–6, 2013.

[5] C. Wang, Z. Chen, F. Yang et al., "Analysis of the GGGGCC repeat expansions of the C9orf72 gene in SCA3/MJD patients from China," *PLoS One*, vol. 10, no. 6, article e0130336, 2015.

[6] X. Li, H. Liu, P. L. Fischhaber, and T. S. Tang, "Toward therapeutic targets for SCA3: insight into the role of Machado-Joseph disease protein ataxin-3 in misfolded proteins clearance," *Progress in Neurobiology*, vol. 132, pp. 34–58, 2015.

[7] H. Peng, C. Wang, Z. Chen et al., "The *APOE* $\varepsilon 2$ allele may decrease the age at onset in patients with spinocerebellar ataxia type 3 or Machado-Joseph disease from the Chinese Han Population," *Neurobiology of Aging*, vol. 35, no. 9, pp. 2179.e15–2179.e18, 2014.

[8] Z. Chen, P. Wang, C. Wang et al., "Updated frequency analysis of spinocerebellar ataxia in China," *Brain*, vol. 141, no. 4, article e22, 2018.

[9] C. Wang, H. Peng, J. Li et al., "Alteration of methylation status in the *ATXN3* gene promoter region is linked to the SCA3/MJD," *Neurobiology of Aging*, vol. 53, pp. 192.e5–192.e10, 2017.

[10] Z. Chen, C. Zheng, Z. Long et al., "(CAG)n loci as genetic modifiers of age-at-onset in patients with Machado-Joseph disease from mainland China," *Brain*, vol. 139, no. 8, article e41, 2016.

[11] C. Costa Mdo and H. L. Paulson, "Toward understanding Machado–Joseph disease," *Progress in Neurobiology*, vol. 97, no. 2, pp. 239–257, 2012.

[12] K. Seidel, S. Siswanto, M. Fredrich et al., "Polyglutamine aggregation in Huntington's disease and spinocerebellar ataxia type 3: similar mechanisms in aggregate formation," *Neuropathology and Applied Neurobiology*, vol. 42, no. 2, pp. 153–166, 2016.

[13] K. Rajamani, J. W. Liu, C. H. Wu et al., "N-Butylidenephthalide exhibits protection against neurotoxicity through regulation of tryptophan 2, 3 dioxygenase in spinocerebellar ataxia type 3," *Neuropharmacology*, vol. 117, pp. 434–446, 2017.

[14] M. Tada, M. Nishizawa, and O. Onodera, "Roles of inositol 1,4,5-trisphosphate receptors in spinocerebellar ataxias," *Neurochemistry International*, vol. 94, pp. 1–8, 2016.

[15] X. Chen, J. Wu, Y. Luo et al., "Expanded polyglutamine-binding peptoid as a novel therapeutic agent for treatment of Huntington's disease," *Chemistry & Biology*, vol. 18, no. 9, pp. 1113–1125, 2011.

[16] A. H. Chou, Y. L. Chen, C. C. Chiu et al., "T1-11 and JMF1907 ameliorate polyglutamine-expanded ataxin-3-induced neurodegeneration, transcriptional dysregulation and ataxic symptom in the SCA3 transgenic mouse," *Neuropharmacology*, vol. 99, pp. 308–317, 2015.

[17] S. Gunawardena, L.-S. Her, R. G. Brusch et al., "Disruption of axonal transport by loss of huntingtin or expression of pathogenic polyQ proteins in Drosophila," *Neuron*, vol. 40, no. 1, pp. 25–40, 2003.

[18] B. Malik, N. Nirmalananthan, L. G. Bilsland et al., "Absence of disturbed axonal transport in spinal and bulbar muscular atrophy," *Human Molecular Genetics*, vol. 20, no. 9, pp. 1776–1786, 2011.

[19] J. C. Chang, S. L. Wu, F. Hoel et al., "Far-infrared radiation protects viability in a cell model of spinocerebellar ataxia by preventing polyQ protein accumulation and improving mitochondrial function," *Scientific Reports*, vol. 6, no. 1, article 30436, 2016.

[20] J. Y. Hsu, Y. L. Jhang, P. H. Cheng et al., "The truncated C-terminal fragment of mutant ATXN3 disrupts mitochondria dynamics in spinocerebellar ataxia type 3 models," *Frontiers in Molecular Neuroscience*, vol. 10, p. 196, 2017.

[21] F. P. Marchese, I. Raimondi, and M. Huarte, "The multidimensional mechanisms of long noncoding RNA function," *Genome Biology*, vol. 18, no. 1, p. 206, 2017.

[22] J. Beermann, M. T. Piccoli, J. Viereck, and T. Thum, "Noncoding RNAs in development and disease: background, mechanisms, and therapeutic approaches," *Physiological Reviews*, vol. 96, no. 4, pp. 1297–1325, 2016.

[23] F. Kopp and J. T. Mendell, "Functional classification and experimental dissection of long noncoding RNAs," *Cell*, vol. 172, no. 3, pp. 393–407, 2018.

[24] M. L. Moseley, T. Zu, Y. Ikeda et al., "Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8," *Nature Genetics*, vol. 38, no. 7, pp. 758–769, 2006.

[25] J. S. Sunwoo, S. T. Lee, W. Im et al., "Altered expression of the long noncoding RNA NEAT1 in Huntington's disease," *Molecular Neurobiology*, vol. 54, no. 2, pp. 1577–1586, 2017.

[26] C. K. Cemal, C. Huxley, and S. Chamberlain, "Insertion of expanded CAG trinucleotide repeat motifs into a yeast artificial chromosome containing the human Machado-Joseph disease gene," *Gene*, vol. 236, no. 1, pp. 53–61, 1999.

[27] C. K. Cemal, C. J. Carroll, L. Lawrence et al., "YAC transgenic mice carrying pathological alleles of the MJD1 locus exhibit a mild and slowly progressive cerebellar deficit," *Human Molecular Genetics*, vol. 11, no. 9, pp. 1075–1094, 2002.

[28] D. Parkhomchuk, T. Borodina, V. Amstislavskiy et al., "Transcriptome analysis by strand-specific sequencing of complementary DNA," *Nucleic Acids Research*, vol. 37, no. 18, article e123, 2009.

[29] F. Mori, M. Nishie, Y. S. Piao et al., "Accumulation of NEDD8 in neuronal and glial inclusions of neurodegenerative disorders," *Neuropathology and Applied Neurobiology*, vol. 31, no. 1, pp. 53–61, 2005.

[30] T. Mueller, P. Breuer, I. Schmitt, J. Walter, B. O. Evert, and U. Wüllner, "CK2-dependent phosphorylation determines cellular localization and stability of ataxin-3," *Human Molecular Genetics*, vol. 18, no. 17, pp. 3334–3343, 2009.

Hindawi
International Journal of Genomics
Volume 2018, Article ID 1351964, 19 pages
https://doi.org/10.1155/2018/1351964

Research Article

Topological Characterization of Human and Mouse m^5C Epitranscriptome Revealed by Bisulfite Sequencing

Zhen Wei[1,2] **Subbarayalu Panneerdoss**[3,4] **Santosh Timilsina**[3,4] **Jingting Zhu**[1,5] **Tabrez A. Mohammad**[3] **Zhi-Liang Lu**[1,5] **João Pedro de Magalhães**[2] **Yidong Chen**[3,6] **Rong Rong**[1,5] **Yufei Huang**[6,7] **Manjeet K. Rao**[3,4] **and Jia Meng**[1,5]

[1]*Department of Biological Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, China*
[2]*Integrative Genomics of Ageing Group, Institute of Ageing and Chronic Disease, University of Liverpool, L7 8TX Liverpool, UK*
[3]*Greehey Children's Cancer Research Institute, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA*
[4]*Department of Cellular Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA*
[5]*Institute of Integrative Biology, University of Liverpool, L7 8TX Liverpool, UK*
[6]*Department of Epidemiology and Biostatistics, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA*
[7]*Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX 78249, USA*

Correspondence should be addressed to Rong Rong; rong.rong@xjtlu.edu.cn, Yufei Huang; yhuang@utsa.edu, Manjeet K. Rao; raom@uthscsa.edu, and Jia Meng; jia.meng@xjtlu.edu.cn

Received 17 January 2018; Revised 14 April 2018; Accepted 17 April 2018; Published 13 June 2018

Academic Editor: Yujing Li

Background. Compared with the well-studied 5-methylcytosine (m^5C) in DNA, the role and topology of epitranscriptome m^5C remain insufficiently characterized. *Results*. Through analyzing transcriptome-wide m^5C distribution in human and mouse, we show that the m^5C modification is significantly enriched at 5′ untranslated regions (5′UTRs) of mRNA in human and mouse. With a comparative analysis of the mRNA and DNA methylome, we demonstrate that, like DNA methylation, transcriptome m^5C methylation exhibits a strong clustering effect. Surprisingly, an inverse correlation between mRNA and DNA m^5C methylation is observed at CpG sites. Further analysis reveals that RNA m^5C methylation level is positively correlated with both RNA expression and RNA half-life. We also observed that the methylation level of mitochondrial RNAs is significantly higher than RNAs transcribed from the nuclear genome. *Conclusions*. This study provides an in-depth topological characterization of transcriptome-wide m^5C modification by associating RNA m^5C methylation patterns with transcriptional expression, DNA methylations, RNA stabilities, and mitochondrial genome.

1. Introduction

DNA methylation is a well-established and extensively studied epigenetic phenomenon [1–4]. In contrast, mRNA methylation is still relatively an uncharted territory [5]. Although the presence of the chemical modifications to tRNA has been established in the 1970s [6–8], little is known about the epigenetic modifications to mRNA and other noncoding RNAs. Even less was known about their abundance, role, and mode of regulation until recently when several studies showed that N^6-methyladenosine (m^6A) is the most abundant messenger RNA (mRNA) modification in eukaryotes [9], and suggested to regulate a number of biological processes including translation efficiency [10], circadian clock [11], microRNA processing [12], RNA-protein interaction [13], RNA stability [14], heat shock response [15], and differentiation [16].

Compared to m^6A, even little is known about the abundance and role of transcriptome 5-methylcytosine (m^5C) modification. Existing studies of m^5C in cellular RNA have been largely confined to rRNA and tRNA [17]. For example,

RNA m^5C modification in plant rRNA and tRNA is reported to be conserved [18] and is shown to affect the stability of synthetic RNA [19, 20]. In the mammalian system, cytosine-5 methylation in tRNA has been shown to regulate Mg^{2+} binding, anticodon stem-loop conformation, and secondary structure stabilization [21, 22]. In addition, m^5C in tRNAs is reported to regulate protein translation in stress response, tissue differentiation, and neurodevelopment disorders [23–29]. In rRNA, m^5C is shown to regulate the translation process [30]. A recent study also showed that hm^5C, the intermediate of RNA m^5C demethylation, is enriched in poly(A)-tailed RNA and the coding sequences of the mRNA transcript, and it is associated with brain development and the active transcription of mRNA [11].

A recent advancement of the RNA bisulfite-sequencing (BS-Seq) technique [31–34] has enabled the transcriptome-wide m^5C profiling at single-base resolution and confirmed its widespread existence in the human transcriptome [34, 35]. Intriguing differences with respect to the degree of transcriptome m^5C methylation, functional classification, and position bias were reported with this technique [36], and it was recently shown that transcriptome m^5C promotes mRNA export through methyltransferase NSUN2 and reader ALYREF [37].

It is observed that m^5C modification may account for 20% of the total internal methylations on poly(A) RNA in the BHK21 cell line [38, 39]. However, it is not clear whether the transcriptome m^5C modification is differentially enriched in different cell types, and the topological relationship between RNA methylation and DNA methylation under the same cell lines has not been investigated.

In this study, using the BS-Seq approach, we identified transcriptome-wide mRNA m^5C methylome in mouse and human cells. Our results revealed that transcriptome m^5C is enriched and conserved at the 5′UTRs of target transcripts in both human and mouse cells. Interestingly, under all the examined cell lines, we observed a negative correlation of the methylation patterns between RNA m^5C methylation and DNA m^5C under the CpG context, and the RNA m^5C methylations are enriched on mitochondrial transcriptome.

2. Material and Methods

2.1. Sample Preparation and RNA Bisulfite Sequencing. MCF10A normal mammary epithelial cells and MDA-MB-468 breast cancer cells were obtained from ATCC. MCF10A cells were cultured and maintained in DMEM/F12 (Life Technologies, USA) supplemented with 5% horse serum, EGF (20 ng/ml), hydrocortisone (0.5 μg/ml), insulin (10 μg/ml), and anti-anti (Life Technologies, USA). Likewise, MDA-MB-468 cells were cultured in RPMI (Life Technologies, USA) supplemented with 10% FBS and anti-anti (Life Technologies, USA). For BS-Seq, total RNA was isolated from MCF10A and MDA-MB-468 cells and enriched for poly(A)+ RNA using poly(A) selection kits. The purified RNA is subjected to sodium bisulfite treatment at 60 degrees for 8 hours. The bisulfite-treated RNA was then reverse transcribed and subjected to deep sequencing using the Illumina RNA-Seq protocol. The data has been deposited under Gene Expression Ominous (GEO) with Accession Number GSE84230. To replenish the transcriptome BS-Seq data of the aforementioned human samples (MCF10A and MDA-MB-468), additional datasets are obtained from public resources, including DNA BS-Seq data from MCF10A (GEO GSM659628) [40], transcriptome m^5C methylation data from mouse embryo stem cells (ESCs) and mouse whole brain profiled by RNA BS-Seq (GEO GSE83432) [36], and mouse ESC DNA methylation data (GSM1873374) [7, 41].

2.2. Quality Control and Alignment of BS-Seq Data. The FASTQ files from BS-Seq samples are trimmed with Trim Galore [42], it removes low-quality 3′ ends with a Phred score threshold of 20, and it can remove potential adaptor contamination. Then, the reads are aligned to the reference genomes of mouse and human (mm10 and hg19) with MeRanGs in MeRanTK [43]. The methylation is called using MeRanCall, and regions of the 5′ ends and 3′ ends of the reads are ignored based on the threshold cutoff suggested by the M-bias plot generated by MeRanGs. The minimum read coverage for the methylation report was set at 10, and the minimum read base quality (Phred score) for methylation call is filtered at 30. The maximum read duplication level is set at 10 to prevent the PCR artefacts; the minimum nonconversion rate to report is set at 0 to include the nonmethylated sites as background control for further analysis.

For DNA bisulfite samples, the trimmed reads are aligned using Bismark under the following alignment setting: –score_min L,0,-0.6. The SAM files are filtered by Samtools using -F 1540 and -q 30 to remove reads that are duplicated and quality scores that are lower than 30. The methylation status of genome-wide cytosine sites is reported from the filtered SAM files with the Bismark methylation extractor using the following argument: –cytosine report. Also, the conversion rate biased ends are also ignored during methylation call based on the M-bias plots. The minimum read coverage was filtered at 10 as well.

2.3. Filtering False Positive m^5C Sites due to RNA Secondary Structure. It is known that secondary structures on RNAs prohibit bisulfite conversion and thus can result in false positive detection of transcriptome m^5C sites. As shown in Figure S1, the detected m^5C sites from MeRanTK are enriched with double-stranded regions of RNA, which are likely to be false positive errors due to a secondary structure. For this reason, an R package rBS2ndStructure was created to facilitate the elimination of the false positive methylation calls due to RNA secondary structures. Specifically, the RNA secondary structure is predicted with RNAfold from the Vienna RNA package [44] as it was performed by Amort et al. [36]. The transcriptome-wide full-length transcripts are extracted from UCSC gene annotation for both mm10 and hg19. Then, the double-stranded structures are predicted with the MEA method under alpha = 0.1. The folding temperature is set at 70 degrees, and the maximum pairing distance is set at 150 bp. For the mitochondrial chromosome and transcripts longer

than 8000 bp, the structures are predicted using sliding windows of 2000 bp and step size of 1000 bp. For both the RNA and the DNA methylation reports, the methylation sites overlapped with the predicted regions of secondary structures are filtered. Due to the lack of computational resources to predict structures on large intronic sequences, the cytosine sites that do not locate on the exons of known transcripts or the mitochondrial chromosome are filtered. The resulting methylation reports are then analyzed under the R environment using primarily GenomicFeatures [45], Guitar [46], and ggplot2 [47] packages.

The rBS2ndStructure package is publically available at Github (https://github.com/ZhenWei10/rBS2ndStructure) with precomputed RNA secondary structures of genome assembly mm10 and hg19 for convenient processing of RNA BS-Seq result.

2.4. Quantitative Analysis of Methylation Status. The methylation ratio (mRatio) of a specific cytosine site is calculated by

$$\text{Methylation ratio} = \frac{\#\text{of unconverted Cs}}{\#\text{of unconverted Cs} + \#\text{of converted Cs}}, \quad (1)$$

where "# of unconverted Cs" and "# of converted Cs" indicates the count of methylated (unconverted) Cs and unmodified Cs (converted Cs) at a specific cytosine site, respectively. The methylation rate is conceptually similar to the well-adapted concept of "beta value" in DNA methylation analysis [48], which indicates the percentage of methylated Cs among all Cs. Also, it is not difficult to show that

$$\begin{aligned}\text{Methylation ratio} &= \frac{\#\text{of unconverted C}}{\#\text{of unconverted C} + \#\text{of converted C}} \\ &= 1 - \frac{\#\text{of converted C}}{\#\text{of unconverted C} + \#\text{of converted C}} \\ &= 1 - \text{convertion rate},\end{aligned} \quad (2)$$

where the conversion rate has been previously defined in [35] and a smaller value suggests a higher percentage of RNA m^5C methylation.

To differentiate a set of statistically significantly methylated cytosine sites against potential technical randomness due to incomplete bisulfite conversion, the p values for the methylation state of both the DNA and RNA methylation are calculated by Fisher's exact test against the background conversion odds after the filtering of the sites mapped to introns and secondary structures. The adjusted p values (FDR) are then adjusted by the Benjamin & Hochberg method. The positive methylation states were decided when FDR < 0.05.

For the mouse samples containing 3 biological replicates, the methylated sites are judged as FDR < 0.05 among all 3 replicates. For other insignificant methylated sites to be kept in the analysis, the sites should be reproduced 3 times with coverage > 10. The converted reads and nonconverted reads are added on each site when combining the biological replicates.

The background bisulfite nonconversion rate is 2.75%, 2.74%, 1.18%, and 0.81% for MCF10A, MDA468, mouse ESC, and mouse brain samples, respectively (taking the average for samples with more than one biological replicate). The difference among nonconversion rates might be due to the biological difference of cell lines, batch variation, and different BS-Seq protocols.

2.5. Differential Methylation Analysis. The odds ratio (OR) or methylation fold change from differential analysis is defined as

$$\text{Odds ratio from differential methylation} = \frac{(\#\text{of unconverted Cs under cond\_1}/\#\text{of converted Cs under cond\_1})}{(\#\text{of unconverted Cs under cond\_2}/\#\text{of converted Cs under cond\_2})}. \quad (3)$$

Odds ratio (or methylation fold change) indicates whether the methylation is enriched under one condition compared with another condition. A value greater than 1 suggests increased methylation level, where as a value less than 1 suggests decreased methylation level. The statistical significance of the odds ratio is evaluated by the QNB method, which tests the homogeneity of association between methylated and unmodified molecules under two experimental conditions with the within-group variability assessed through 4 cross-linked negative binomial distributions [49].

Similar to the odds ratio from differential methylation analysis, the enrichment odds ratio of m^5C sites within a specific region can be defined as

$$\text{Enrichment odds ratio} = \frac{\left(\#\text{of } m^5\text{C sites within a region}/\#\text{of C sites within a region}\right)}{(\text{total}\#\text{of } m^5\text{C sites}/\text{total}\#\text{of C sites})}. \quad (4)$$

A value greater than 1 suggests that methylation sites are enriched within the tested region, and the statistical significance of enrichment can be evaluated by Fisher's exact test. Please note that, in this analysis, we used the total number of cytosine sites reported from MeRanTK rather than the total number of all 4 types of nucleotides.

2.6. Assessing the Distribution of m^5C Sites on mRNA. The distribution pattern of m^5C sites on mRNA is assessed with the Guitar R/Bioconductor [46]. Compared with other software tools and methods, the Guitar package provides an improved resolution by relying on only the mRNA transcripts that simultaneously have sufficient long (more than 100 bp) 5′ UTRs, CDSs, and 3′ UTRs. For instance, transcripts without annotated 5′UTRs will be excluded from the analysis. Additionally, Guitar does not rely on only the primary transcript (often defined as the longest transcript among all isoforms in practice) when solving an ambiguous association between a m^5C site and the isoform transcripts of a gene; instead, all ambiguous associations are considered with the weight of association evenly divided. For example, if a single m^5C site locates on the 3′UTR of a transcript and CDS of another isoform transcript of that gene, it is counted as if half of the m^5C site is located on the 3′ UTR and the other half located on 5′ UTR. In this way, the isoform information is largely retained. To our knowledge, the Guitar package should provide the most accurate assessment of a transcriptomic distribution pattern.

2.7. Differential Expression Analysis. Differential expression analysis was performed with the DESeq2 package [50] and the aligned RNA BS-Seq data.

2.8. Cell Culture and Viral Infection. Jurkat T lymphocytes were maintained in RPMI 1640 medium (Hyclone) supplemented with 5% (*v*/*v*) FBS (Gibco) and 100 U/ml penicillin/streptomycin (Hyclone). For infection, Jurkat cells were infected with known amounts (3×108 genome copies per 2×105 cells) of SRV for 18 hours at 37°C, followed by washing three times with PBS (Hyclone). Infected cells were incubated in completed culture medium for the indicated time. Successful infection was identified as the appearance of cytopathic effects in infected cells at 8 to 10 days postinfection.

2.9. Reverse Transcription and Real-Time PCR. SRV genome in culture medium was extracted by viral RNA extraction kit (TIANGEN) and reverse transcripted into cDNA by a reverse transcriptase PCR kit (TaKaRa). Cellular genome was extracted by a TIANamp Genomic DNA Kit (TaKaRa). Real-time PCR was performed in a 7500 Fast Real-Time PCR System (Applied Biosystems) by using a Premix Ex Taq (Probe qPCR) kit (TaKaRa). SRV genome positive control, primers, and probe, as well as GAPDH primers and probe were kindly provided by VRL China Ltd. [51].

2.10. Immunofluorescence Assay. Cells were seeded on poly-L-lysine (Sigma) coated slides, fixed with 4% paraformaldehyde for 15 minutes, permeabilized with precold pure methanol for 20 min at −20°C, and blocked with 5% BSA for 1 hour. Cells were then stained with the serum from an SRV-infected monkey (1 : 25 diluted in blocking buffer) overnight and visualized with DyLight™ 488-Labeled Anti-Human antibody (KPL). Cells were counterstained with Hoechst (Life Technologies) for 10 minutes and mounted on microscopy slides. Samples were imaged with a ZEISS LSM 880 Confocal Laser Scanning Microscope.

3. Results

3.1. Overview of mRNA m^5C Methylome Revealed by BS-Seq. After successful processing of the RNA BS-Seq datasets, a total of 3440 (0.40%), 1915 (0.29%), 35,246 (0.757%), and 25,301 (0.50%) RNA cytosine sites were identified as m^5C methylation (FDR < 0.05) sites in MCF10A, MDA-MB-468, mouse embryonic stem cell (ESC), and mouse whole brain, respectively. The overall transcriptome m^5C methylation level was much lower than the DNA m^5C methylation level (Figure S2). Importantly, we found that m^5C was widespread in different RNA families, where more than 50% of them were located on mRNA (Figure 1(a)). In MCF10A cells, 7131 protein-coding genes had sites reported after the filtering, of which 225 (3.15%) mRNAs contained m^5C sites. In MDA-MB-468 cells, 6320 protein-coding genes had reads aligned, of which 128 (2.06%) contained m^5C sites. In ESC and brain samples, the methylation status was available for 11,325 and 13,108 protein-coding genes, of which 3579 (31.6%) and 3065 (23.4%) contained m^5C sites. The difference in number of m^5C sites between different conditions is mostly due to different sequencing depth.

3.2. mRNA m^5C Is Enriched in 5′UTRs of Human and Mouse. To study the spatial organization of m^5C sites in the transcriptome, we first analyzed the relative enrichment (see Materials and Methods for more details) of m^5C sites on different types of RNA and at different regions (shown in Figure 1(b)) by compensating for the cytosine sites that do not carry m^5C modification. Our results showed that m^5C sites were consistently and significantly enriched at 5′UTRs in human and mouse with enrichment odds ratio of 3.138, 4.802, 2.744, and 1.601 (please see Table S1 for more details). The similar topology was already reported by previous studies [35, 36], and our observation further confirmed their conclusions. Also, we did observe a slight enrichment of m^5C sites in 3′UTR in mouse brain (enrichment odds ratio = 1.013 and $1.19E-02$), which is also reported in the study of Amort et al. [36]. 3′UTR enrichment was not observed in the other samples (odds ratio = 0.964, 0.971, and 0.617).

To further substantiate these findings, we plotted the distribution of the methylated and unmethylated cytosine sites located on mRNAs with the Guitar package [46]. In order to improve the resolution of this analysis and differentiate the distribution of m^5C sites on usually short 5′UTRs, only the mRNAs with a 5′UTR longer than 100 bp are used. As shown in Figure 2(a), the methylated cytosine sites were consistently enriched at 5′UTRs across all 4 samples when compared to unmethylated groups. Interestingly, this trend is also supported by the cytosine methylation sites reported

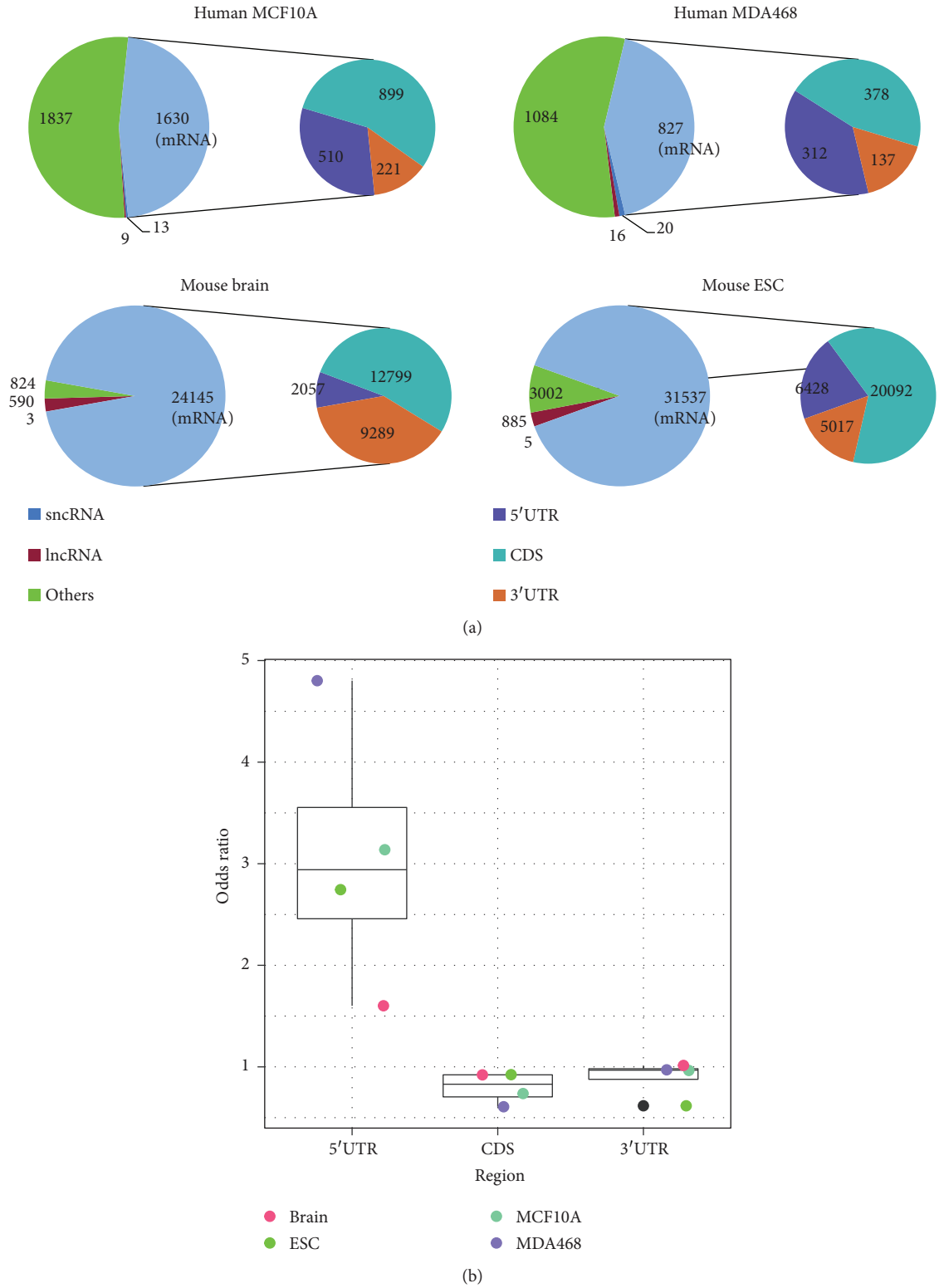
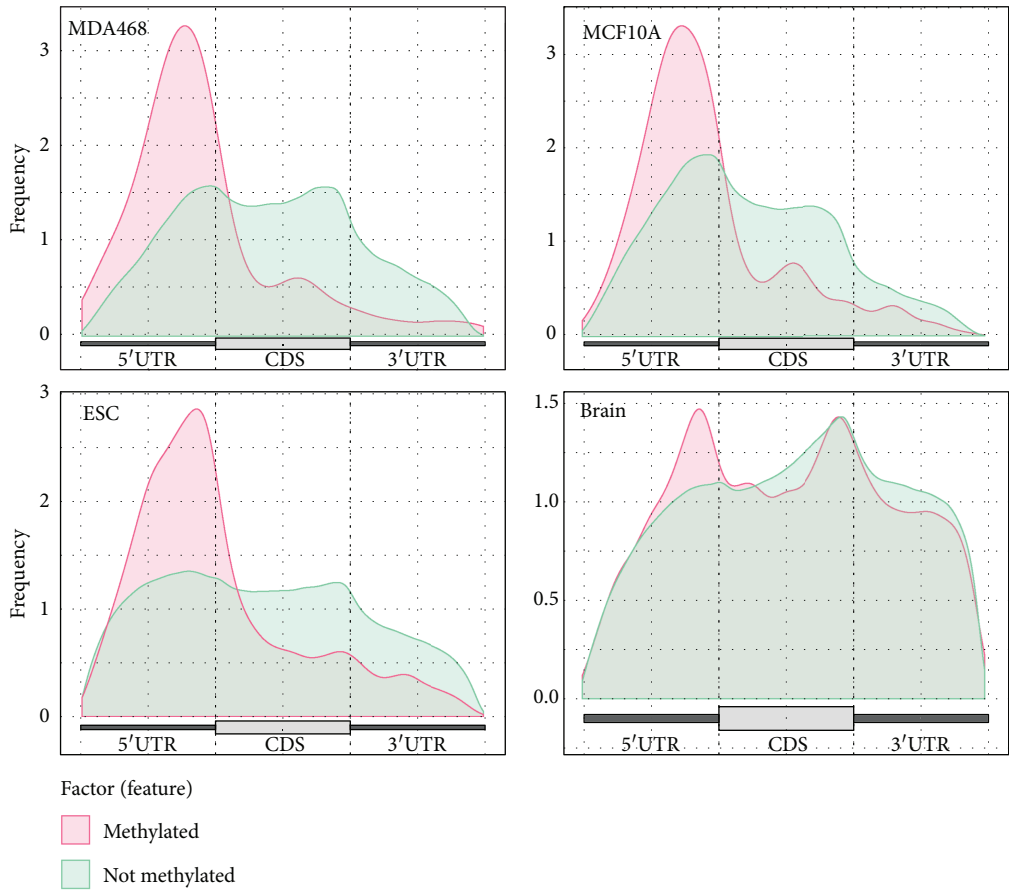
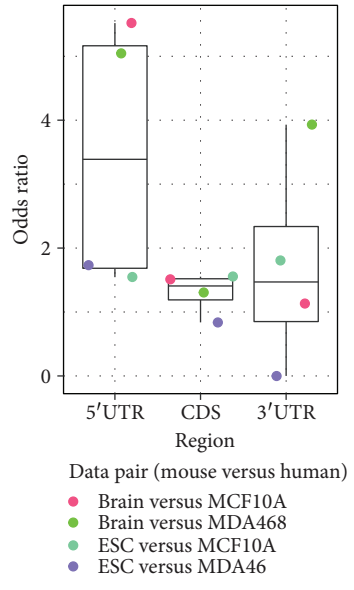


Figure 1: Distribution of transcriptome m^5C modification sites in human and mouse. (a) The pie chart shows transcriptome-wide distribution of m^5C sites in MCF10A, MDA-MB-468, mouse embryonic stem cell (ESC), and whole brain. The majority of the identified m^5C sites are located on mRNAs. (b) Graph showing status of m^5C frequency in different regions of mRNA. The result indicates that detected cytosine sites are consistently enriched at the 5′UTR on mRNA compared with the CDS and 3′UTR.



(a)



(b)

Figure 2: Conservation of m^5C in different mRNA regions. (b) Graph showing the status of m^5C frequency in different regions of the transcripts. We divided all the detected cytosine sites into 2 groups based on whether it is methylated. The result indicates that cytosine sites with significant methylation levels are consistently enriched at the 5′UTRs and near the start codon in all 4 samples. (b) A correlated methylation pattern is observed on 5′UTRs between different cell lines/tissues in human and mouse. The conserved cytosine residuals were retrieved with liftOver utility (http://genome.ucsc.edu/cgi-bin/hgLiftOver), and the correlation analysis is performed with Fisher's exact test. It is important to note that, although we failed to observe a correlated m^5C methylation pattern on CDSs and 3′UTRs of mRNA, it is possible that such pattern may emerge on strictly matched cell lines/tissues.

by Squires et al. [35], and there is no significant enrichment of m^5C sites observed in 3′UTR when all cytosine methylated sites were used as background (Figure S3).

When we further compared the methylation status of the conserved loci in human and mouse between different cell lines/tissues, we observed that, although the cell types/tissues we used were not strictly matched, a strong correlated methylation pattern was observed on the 5′UTR region (Figure 2(b) and Table S2). However, unlike the 5′UTR, the correlated pattern of m^5C sites were not consistently observed in CDSs or 3′UTRs in our study; the observed heterogeneity of the m^5C methylome in different transcript regions suggests that the m^5C mapped to the 5′UTR of the transcripts are more likely to be functionally important.

3.3. m^5C Site Exists under Different Nucleotide Contexts. Because RNA methyltransferase Dnmt2 shares strong sequence homology to all DNA DNMT methyltransferases [52], we reason that exploring the relationship between transcriptome and DNA m^5C methylation profiles may unravel interesting interplay between the two kinds of reversible chemical modifications. In mammalian cells, DNA methylation occurs mainly at CG dinucleotides (including ACG, CCG, TCG, and GCG, see Figure 3(b)). To study whether, like DNA methylation, transcriptome m^5C methylation also occurs at the similar position, we analyzed methylated cytosine in the transcriptome. For this purpose, we examined all the possible C-centered trinucleotide combinations. Unlike DNA, transcriptome m^5C occurs at all C-centered trinucleotides (Figure 3(a)) and was observed to be specifically enriched at GCA, ACG, CCG, GCG, CCC, and GCC. These results were found to be consistent within the same species (Pearson correlation = 0.96 and 0.92, Figure 3(c)) and between different species (Pearson correlation = 0.72, 0.75, 0.45, and 0.48, Figure 3(c)).

3.4. Negative Correlation in Methylation Level Is Observed between mRNA and the Corresponding Exonic Region of DNA. We next examined whether there exists any correlation between m^5C methylated/nonmethylated (m^5C methylation ratio) in the transcriptome and corresponding DNA exonic regions at each C-centered trinucleotide sites. Because DNA methylation occurs mainly at CG dinucleotides, as expected, we observed no strong correlation at non-CG trinucleotides. However, we observed significant negative correlation in methylation ratios between RNA and DNA at all four CG-containing trinucleotides. As a higher percentage of m^5C in mRNA is detected, the corresponding DNA exonic CG dinucleotide was less likely to be methylated (Figure 4(a)). Next, we grouped m^5C methylated at all CG sites according to their methylation ratio (methylated and unmethylated) and investigated their distributions in mRNA and the corresponding exonic regions of DNA. Consistent with our previous finding, we observed a significant negative correlation in both human and mouse cells. In particular, 5′UTR in mRNA showed a high methylation ratio, whereas the corresponding DNA region showed a significantly low methylation ratio (Figure 4(b)).

3.5. Transcriptome m^5C Sites Exhibit a Clustering Effect. In DNA methylation, it has been shown that the correlation of methylation rates between two CpG sites is related to the distance (see Figure S4), and the clustering effect can be as high as 0.7 for probes within 200 bp [53]. To address whether the mRNA m^5C methylation also exhibits a clustering effect, we examined the proportion of m^5C sites that are within 10 bp distance of other m^5C sites and compared this proportion with that from 1000 times of random permutation. Our analysis revealed that m^5C showed an obvious clustering effect in both mRNA and DNA (Figures 5(a) and 5(b)). In the ESC cell line, more than 76.7% of the mRNA methylation sites had at least one methylation site mapped within the 10 nt-flanked region, compared with 7.7% of such event by random permutation of methylation states on insignificant methylation sites of the methylated genes. In mouse ESC and brain cells, more than 43.02% and 30.06% of mRNA m^5C methylation sites existed within the m^5C-p-m^5C dimmers, compared with expected rate of 1.02% and 0.77% of such dimmers by the random permutation.

To further elucidate the clustering effect, we calculated the correlation of the methylation ratio between two cytosine sites with a specific distance. To our surprise, mRNA methylation exhibited a stronger clustering effect compared with DNA (Figure 5(c)). In addition, the correlation of the methylation ratio was consistently stronger within 1–3 nt distance as revealed by the higher correlation of the methylation ratio (0.76 in MCF10A and 0.79 in ESC). These results indicated that most CpC dimers are comethylated; the correlation of the methylation ratio can be as high as 0.58 in MCF10A and 0.47 in ESC for cytosine sites with a distance of 4–10 nt. Though the overall clustering effect of DNA methylation was not as strong as mRNA methylation, when only the CpG dinucleotide was considered, DNA methylation exhibited a stronger clustering effect than mRNA methylation (see Figure S4).

3.6. Transcriptome m^5C Is Strongly Enriched in Mitochondrial Transcripts. To further establish a physiological relevance of m^5C distribution, we examined the methylation level of RNAs encoded in different chromosomes. Surprisingly, m^5C modification was strongly enriched in RNAs transcribed specifically from mitochondrial DNA in normal and breast cancer cells as well as in mouse ESC and brain as revealed by enrichment odds ratios of 818.42949, 634.72723, 1028.52065, and 67.28553, respectively. In contrast, the enrichment odds ratios of RNA methylation for transcripts from other chromosomes were found to be roughly the same (Figure 6(a)). The RNA transcripts of all the major genes located on a mitochondrial chromosome were significantly methylated (Figure 6(b)). Previously, it was reported that methyltransferase NSUN5 can regulate mitochondrial gene expression [54], and we speculate that RNA m^5C may play a more vital regulatory role in mitochondria-related biological processes.

3.7. Dysregulation of RNA Methylome in Breast Cancer. Comparison of normal (MCF10A) and breast cancer (MDA-BM-468) m^5C epitranscriptomes identified 162 significant differential methylation sites (DMSs) located on

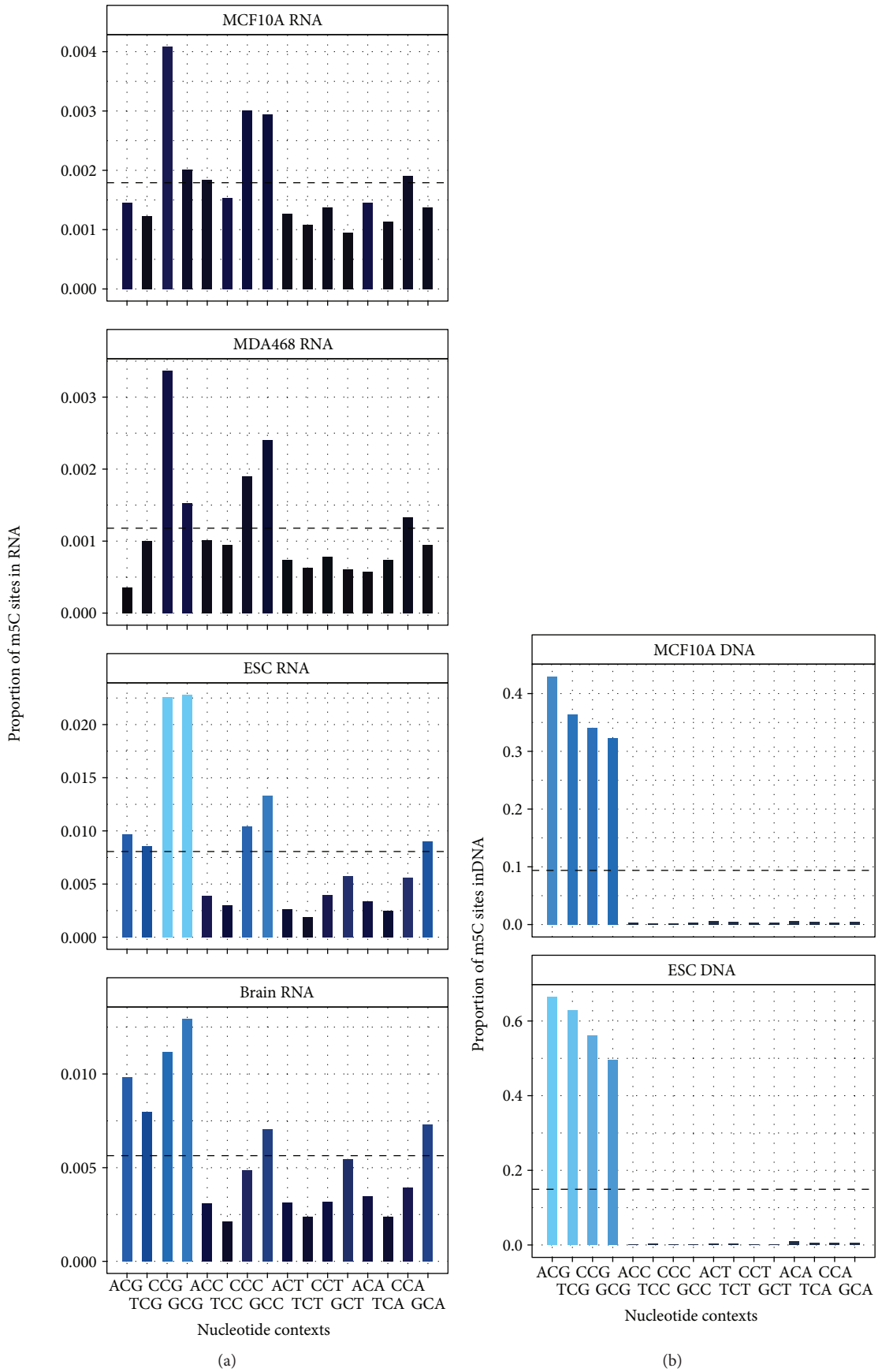
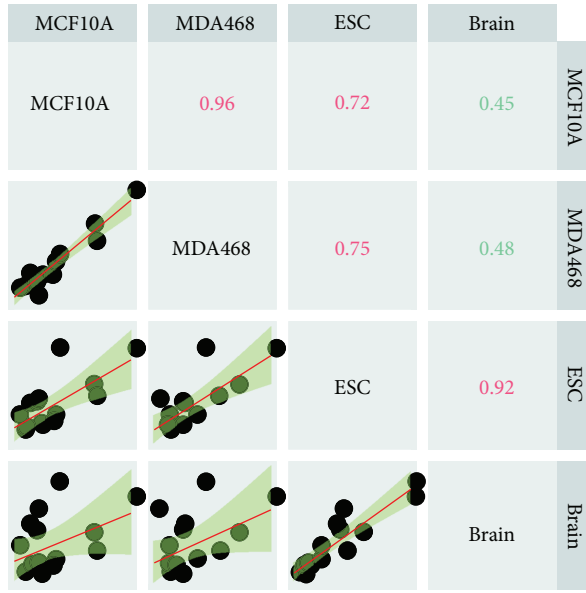


Figure 3: Continued.



(c)

Figure 3: Comparative distributions of mRNA and DNA m^5C methylation. (a) Bar graph shows the proportion of mRNA m^5C sites under different combinations of C-centered trinucleotides in mouse and human cells. The dotted line shows the average percentage of methylation under all trinucleotide contexts within the entire transcriptome. We observed that RNA m^5C occurs under all trinucleotide contexts and is slightly enriched in sequences containing CCG, GCG, GCC, GCU, and GCA. (b) Bar graph showing proportion of DNA m^5C sites in mouse and human cells. DNA cytosine sites were enriched exclusively in sequences containing CG dinucleotides (ACG, CCG, CCG, and TCG). (c) The coefficient of correlation between RNA methylation and trinucleotide sequences was found to be consistent between samples from the same species (Pearson correlation = 0.96 for human and 0.92 for mouse) and also between human and mouse cells (Pearson correlation = 0.72, 0.75, 0.45, and 0.48).

47 annotated genes at a significance level of 0.05. Among the 47 differentially methylated genes, 35 shows hypomethylation and 12 shows hypermethylation in cancer cells compared with the normal control cell line. The majority of the differential methylation sites show hypomethylation (Excel Sheet S1 and Figure 7(a)), and the m^5C hypomethylations are mostly located in the CDS and 3′UTR region of mRNA but not in the 5′UTR region (Figure 7(b)). We then investigated whether different m^5C mRNA methylation levels in normal and breast cancer cells have any functional correlation. We performed functional gene set enrichment analysis on genes containing DMS using the DAVID web server and found that many of the 47 differentially methylated genes are related to important biological functions of cancer, for example, regulation of apoptosis and programmed cell death with RTN4, NME2, CASP14, HSPB1, RPL11, and RPS3 differentially methylated (Excel Sheet S1).

Interestingly, like the difference between the breast cancer cell line MDA-MB-468 and the normal epicelial cell line MCF10A, similar mechanistic mouse stem cells [55] also exhibit dominant hypomethylation in the m^5C epitranscriptome when compared with mouse brain cells with 2513 genes hypomethylated and 767 genes hypermethylated (Figure 7(c) and Excel Sheet S2). Also similar to the previous case, the hypomethylations are mostly located in the CDS and 3′ UTR regions of mRNA, but not in the 5′UTR region (Figure 7(d)). Using DAVID, we found that hypermethylated genes in ESC cells are mostly enriched with the regulation of cell cycle (FZR1, E2F5, BOP1, TRRAP, CDK4, JUNB, etc.), cell death (SIVA1, MCL1, YPEL3, ARF6, UBQLN1, SHF, CIAPIN1, APLP1, GPX1, CASP3, etc.), and mRNA metabolic process (SCAF1, FIP1L1, STRAP, RBM15B, CWC15, XAB2, YBX1, AUH, SF3B2, APLP1, HNRNPL, etc.); the hypomethylated genes are enriched with functions related to ATP synthesis (ATP6V1F, ATP6V1C1, ATP6V0C, ATP6V1A, ATP6V0E, ATP6V1E1, ATP5C1, etc.) and mitochondrial ribosome (MRPL15, MRPL27, MRPL16, MRPL36, MRPL39, MRPL34, DAP3, etc.) (Excel Sheet S2). These results may suggest that the m^5C methylations are selectively methylate transcripts having functions.

3.8. Positive Correlation between m^5C mRNA Methylation and Expression Changes. In our data, as the gene expression is also estimated from RNA bisulfite-sequencing data, a direct comparison of expression and m^5C methylation changes may be problematic due to dependent noise. To eliminate the interference of dependent noise between expression and methylation data, the samples are further divided for different purposes. Specifically, the 3 biological replicates are divided into 2 groups, with 1 sample used for the estimation of expression changes and the other 2 samples for estimation of methylation changes. The expression changes and methylation changes are then compared. This procedure was repeated for 3 times using different grouping combinations.

A consistent and significantly positive correlation is observed (0.274, 0.303, and 0.254) between $\log_2$ expression fold change and $\log_2$ methylation fold change when comparing mouse embryo stem cells with brain cells (Figure S5),

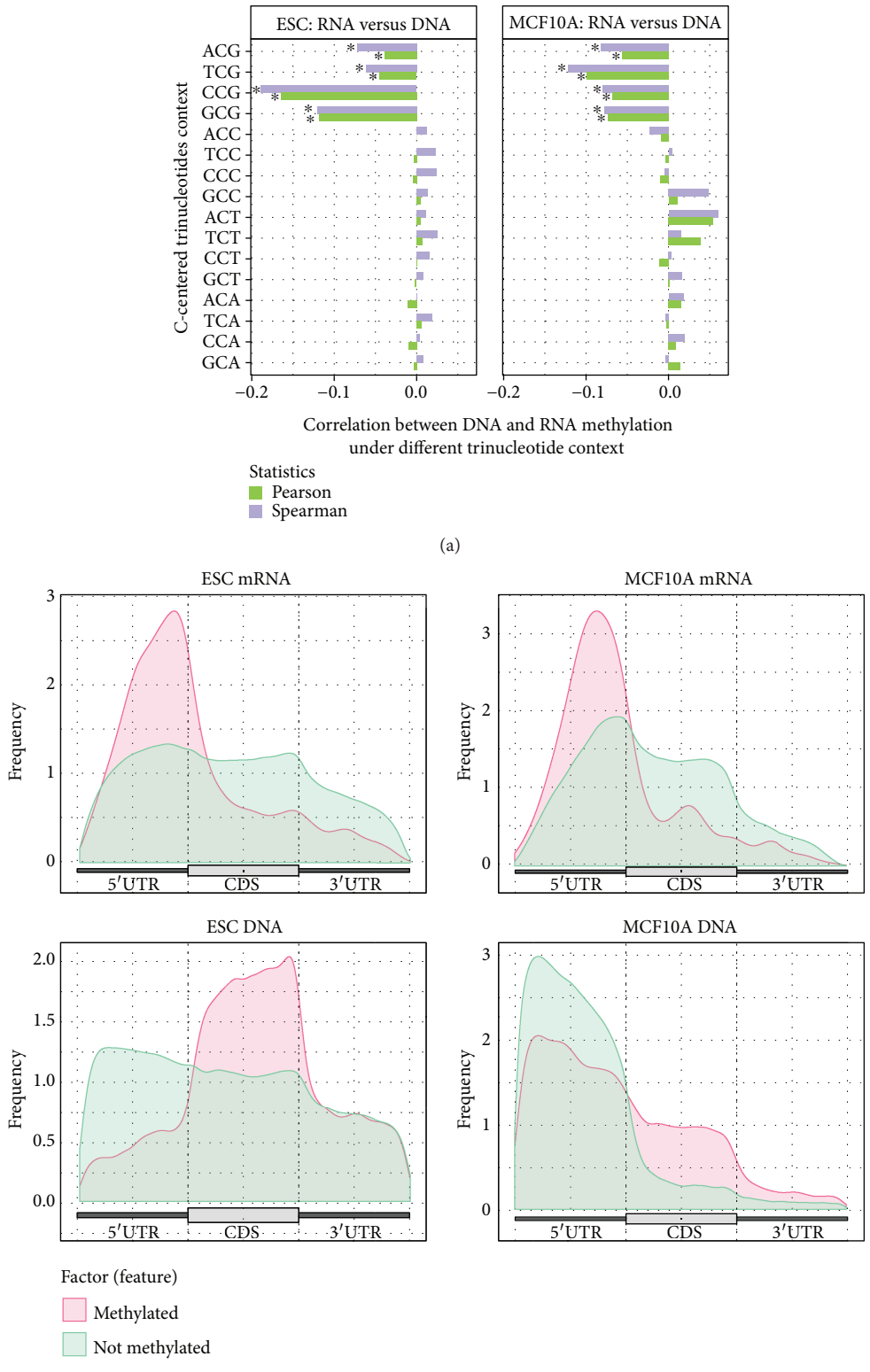


Figure 4: The methylation ratio of corresponding m^5C DNA and mRNA CpG islands shows negative correlation. (a) Negative correlation is observed between DNA and mRNA methylation ratio consistently under all four CG containing trinucleotides (ACG, TCG, CCG, and GCG) in both human and mouse, that is, if a specific CG dinucleotide in DNA is methylated, the corresponding dinucleotide in mRNA is significantly less likely to be methylated. *The top 4 nucleotide contexts under which the strongest correlation between DNA and RNA methylation level exists. (b) Comparative distributions of m^5C methylated CG sites in DNA and RNA show an enrichment of sites with a high methylation ratio in mRNA 5′UTR as opposed to an enrichment of low-methylation-ratio sites in DNA 5′UTR. The pattern is consistent in both the human MCF10A cell line and mouse embryo stem cells.

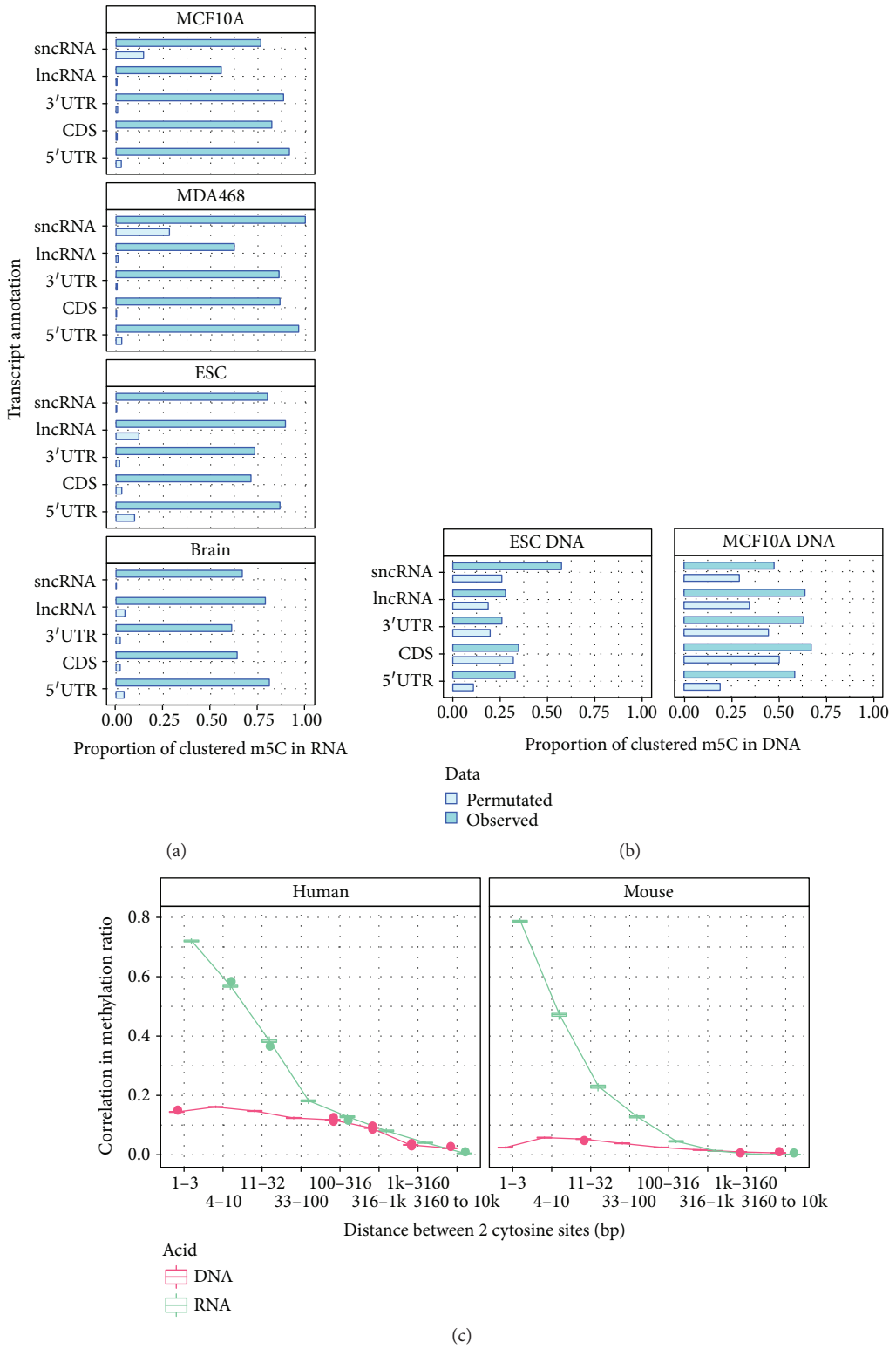


Figure 5: RNA m^5C modification exhibits a clustering effect. (a) Bar graph shows the proportion of clustered m^5C sites within 10 nt flanked regions. To evaluate the statistical significance, we generated 1000 permutated results as a comparison with the bars indicating a 99% confidence interval. Using these criteria, m^5C methylation showed a strong clustering effect consistently on different RNA families and on different regions of mRNA in human and mouse. Around 50% of the m^5C sites were clustered with each other within a 10 bp region. (b) DNA methylation also exhibited a clustering effect. However, the pattern is not that strong when all nucleotide contexts are considered. (c) Line graph showing correlation between RNA/DNA m^5C methylation and distance between cytosine sites. RNA m^5C methylation showed strong correlation with cytosine sites that are immediately close to each other. The clustering effect of DNA methylation is strong when only CpG context is considered (Figure S4).

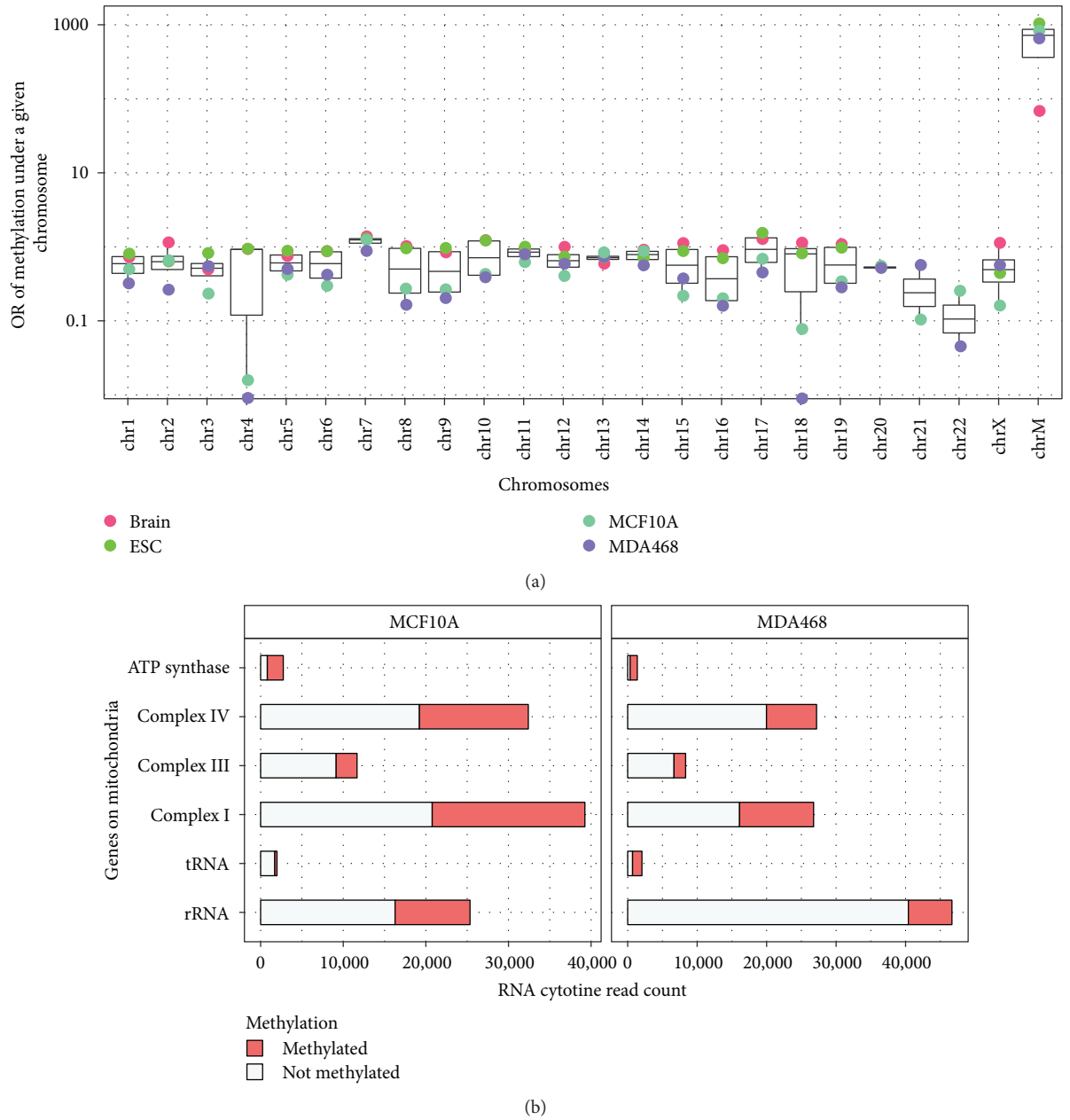


Figure 6: m^5C is enriched on mRNAs transcribed from mitochondrial DNA. (a) Bar graph depicting m^5C mRNA methylation sites on different chromosomes. RNAs transcribed from mitochondrial DNA (M) showed drastically increased frequency of m^5C sites (enrichment odds ratio of 818.42949, 634.72723, 1028.52065, and 67.28553). (b) Bar graph showing the number of methylated cytosine reads stacked with unmodified cytosine reads generated from 6 major classes of mitochondrial genes. The RNA transcripts of all the major genes located on a mitochondrial chromosome were significantly methylated.

suggesting that increased methylation level is likely to be associated with increased expression level. Although the specific molecular mechanism is not yet clear, the observed positive correlation between RNA m^5C and RNA expression confirmed our previous observed anticorrelation between DNA and RNA m^5C methylation (see Figure 4) from a different perspective.

To explain the positive correlation between expression and transcriptome m^5C methylation, we compared the methylation status of all the genes and their half-life, where the half-life of mouse genes were obtained from a previous study [56]. The mRNAs are classified into two groups based on whether they have at least one m^5C site or not. To exclude the confounding factor (effective size in methylation site calling), a generalized linear model of the binomial family was fitted to the half-life with both expression and methylation information. Our result suggests that there exists a significant positive correlation (p value $= 2.23e-12$) between the mRNA half-life and its m^5C methylation status in mouse embryo stem cells, and the positive association is also

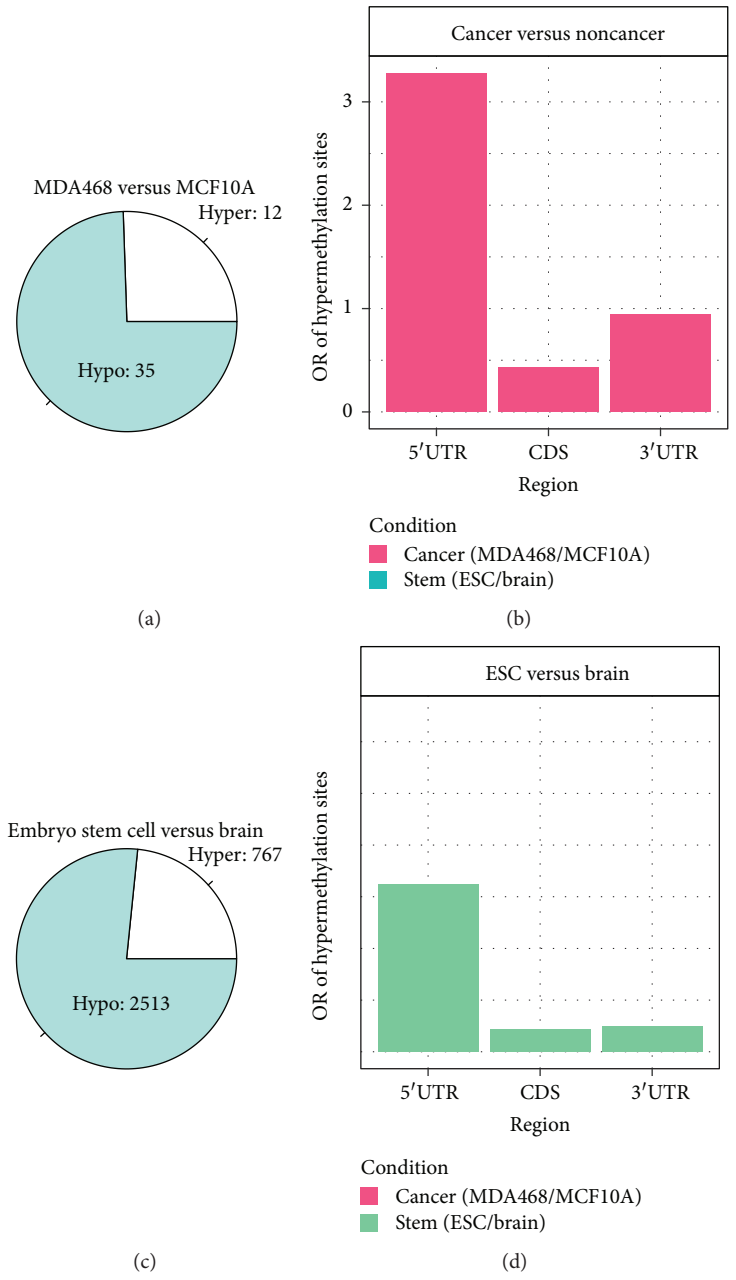


Figure 7: Differential m^5C mRNA methylation in different tissues. (a) Pie-diagram showing hypo- and hypermethylation in MDA468 when compared to MCF10A. A total of 47 differential methylated genes were identified between the breast cancer (MDA-MB-468) and normal control cell lines (MCF10A) with primary hypomethylation under cancer condition. (b) Bar graph showing odds ratio of hypermethylation sites with respect to all differentially methylated sites on different regions of mRNA. Hypermethylated sites were strongly enriched in 5′UTRs. (c) Pie diagram showing hypermethylation in mouse embryo stem cells when compared to whole brain cells. (d) Bar graph showing odds ratio of hypermethylation sites with respect to all differentially methylated sites on different regions of mRNA in the mouse experiment. Hypermethylated sites were strongly enriched in 5′UTRs.

confirmed on mouse whole brain dataset (p value = 0.0374). To further exclude the impact of mRNA expression in calling methylation status, we also extracted the genes whose $\log_2$ expression levels fall between 7 and 11, and then fit their mRNA half-life with a local regression. As shown in Figure 8, compared with the genes of a similar expression level but without an m^5C site, the half-life of the mRNAs that carry m^5C sites is clearly longer and the pattern is consistent in both mouse brain and ESC.

3.9. Dysregulation of RNA Methylome after Simian Retrovirus Infection. Simian retrovirus (SRV) infection of Jurkat T lymphocytes (Jurkat cells) was confirmed by syncytia formation, of which the membrane of the neighboring cells fused to one

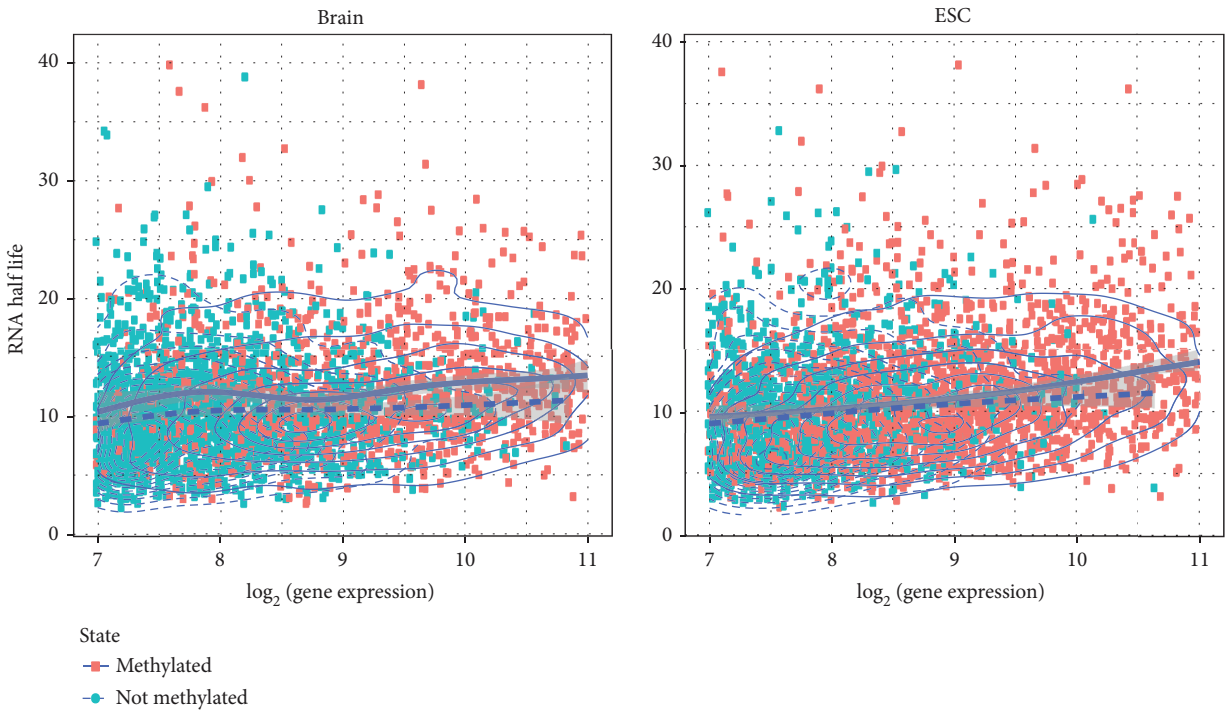


Figure 8: RNA m^5C status is positively correlated with RNA half-life. In the above figure, each red dot represents a gene that carries m^5C sites, and each blue dot represents a gene that does not carry an m^5C site. When comparing the methylated and unmethylated genes of similar expression, the genes that carry an m^5C site have longer RNA half-life than those that do not carry m^5C sites. (a) Positive correlation between RNA methylation status and RNA half-life is observed in mouse brain (p value = 0.0374, generalized linear model of binomial family). (b) Positive correlation between RNA methylation status and RNA half-life is observed in mouse embryo stem cells (p value = $2.23E-12$, generalized linear model of binomial family).

another. At 10 days postinfection, the formation of syncytium was observed among the Jurkat cells incubated with SRV (Figure 9(a)). The syncytium of Jurkat cells contains multiple nuclei and its size is dramatically larger than a single cell. SRV long terminal repeats (LTRs), which are reverse-transcribed from the RNA genome during the infection, contain critical sequences necessary for the integration, synthesis, and expression of viral DNA [1]. Therefore, the extent of SRV infection was assayed by monitoring SRV-LTR expression in Jurkat cells through quantitative real-time PCR. As shown in Figure 9(b), the copy number of SRV-LTR was gradually increased from 2 days to 10 days postinfection and then tended to be stable afterwards. Taken together, these results indicated that SRV was able to infect Jurkat cells and the infection reached maximum level after 10 days postinfection.

In order to investigate whether SRV could replicate in Jurkat cells, the SRV virions released in the culture medium were determined by measuring viral genome copy number through quantitative real-time PCR. As shown in Figure 9(c), the copy number of the SRV genome was gradually increased from 2 days to 14 days postinfection, suggesting that SRV was able to replicate in Jurkat cells.

We then measured the RNA methylome with bisulfite sequencing. A total of 2475 m^5C sites located on 517 genes are reported as differentially methylated 10 days postinfection of SRV with QNB p value < 0.05. Among them, 389 sites located on 158 genes are hypomethylated, while 2086 sites from 382 genes are hypermethylated. A gene ontology analysis using the DAVID website suggests that the differentially methylated genes are related to virus infection, specifically, hypermethylated genes are enriched with DNA replication (p value = $6.07E-5$), mitotic nuclear division (p value = $4.37E-4$), DNA replication initiation (p value = $3.48E-3$), autophagosome assembly (p value = $8.54E-3$), strand displacement (p value = $1.42E-2$), double-strand break repair via homologous recombination (p value = $3.42E-3$), and so on, while hypomethylated genes are enriched with the following biological processes including negative regulation of epidermal growth factor receptor signaling pathway (p value = $3.24E-3$), DNA damage checkpoint (p value = $2.54E-2$), cell migration (p value = $1.42E-2$), and so on (see Figure 9 and Excel Sheet S3). Similar to before, a positive correlation (0.07) is observed between RNA methylation level and expression level; however, as there are 23 genes that carry hyper- and hypomethylated sites simultaneously, it is expected that RNA m^5C carries more complicated biomolecular functions.

4. Discussion and Conclusion

The distribution of m^5C methylation in mRNA has been mysterious with inconsistent evidence reported from previous studies [35, 37]. Here, we profiled the human and mouse m^5C epitranscriptome using RNA BS-Seq data in human

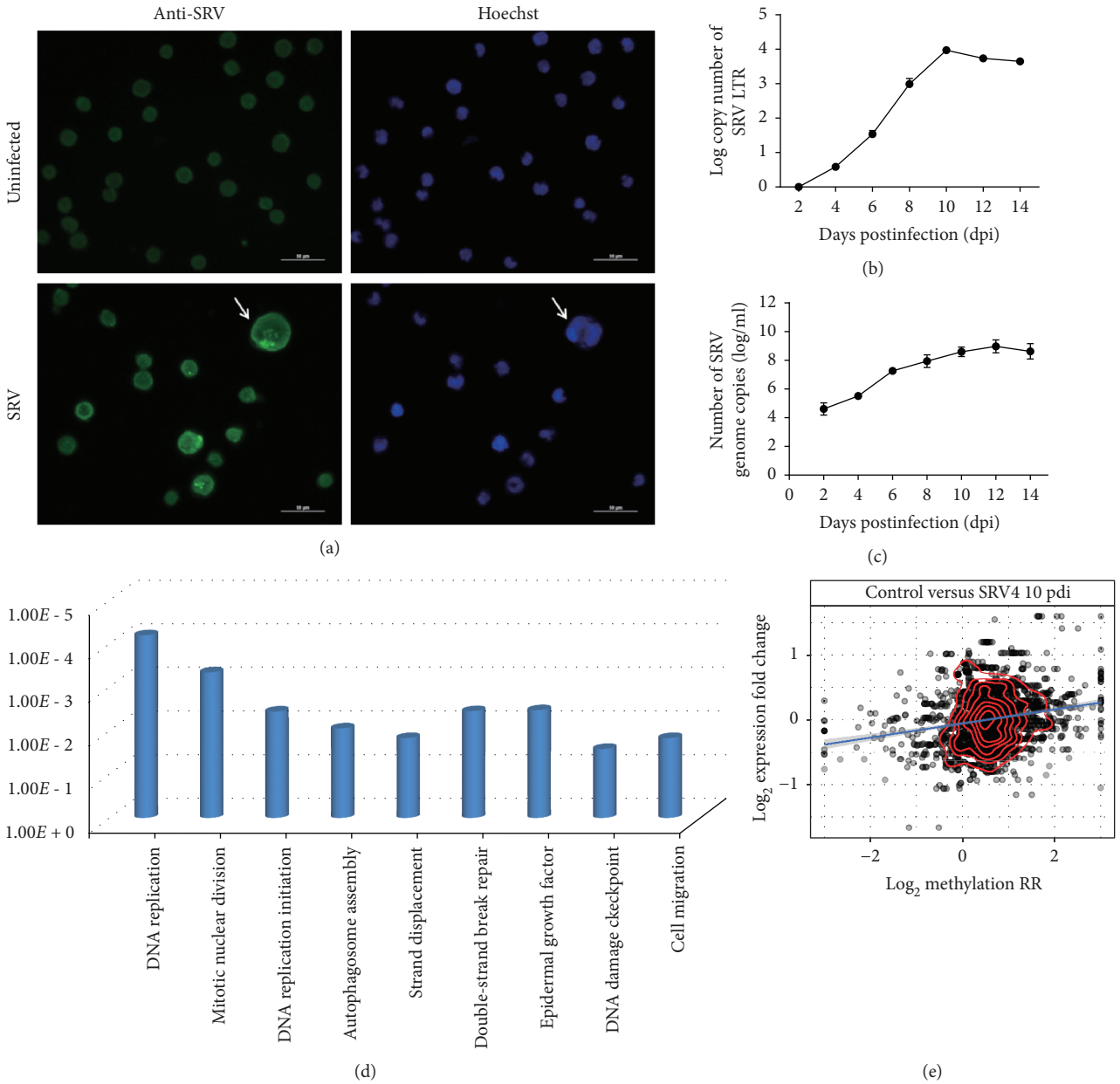


Figure 9: Dysregulation of RNA methylome after SRV infection of Jurkat cell. (a) At 10 days postinfection, uninfected or SRV-infected Jurkat cells were stained with SRV antibodies (green). Nuclei were visualized by Hoechst staining (blue). Arrows indicate the syncytium of infected cells. Scale bar: 50 μM. (b) The relative level of SRV-LTR in infected Jurkat cells was measured every two days by real-time PCR. GAPDH was used as the internal control. The relative level of SRV-LTR at each time point was normalized to the data at 2 dpi; mean ± SD, $n = 3$. (c) The absolute copy number of the SRV genome in culture medium was measured every two days by real-time PCR. SRV-LTR and SRV genome were not detected in all uninfected cells and culture medium, respectively; mean ± SD, $n = 3$. (d) The differentially methylated genes are enriched with the following functions including DNA replication (p value = $6.07E-5$), mitotic nuclear division (p value = $4.37E-4$), DNA replication initiation (p value = $3.48E-3$), autophagosome assembly (p value = $8.54E-3$), strand displacement (p value = $1.42E-2$), double-strand break repair via homologous recombination (p value = $3.42E-3$), and so on, while hypomethylated genes are enriched with the following biological processes including negative regulation of epidermal growth factor receptor signaling pathway (p value = $3.24E-3$), DNA damage checkpoint (p value = $2.54E-2$), cell migration (p value = $1.42E-2$), and so on (see Figure 9 and Excel Sheet S5). (e) A weak but positive correlation (Pearson correlation = 0.07) is observed between RNA methylation level and expression level, which is consistent with our previous result; however, there are 23 genes that carry hyper- and hypomethylated sites simultaneously, which implies that RNA m^5C carries more complicated biomolecular functions.

MCF10A, human MDA468, mouse ESC, and mouse whole brain cells. To eliminate the data sample bias, we employed a rigorous quality control procedure by filtering false positive m^5C sites due to the secondary structure and performed a comprehensive comparative analysis on cross-species conserved locus, cross-sample comparison of topological

transcriptome distributions of m^5C, and differential m^5C analysis. Our analysis clearly shows that m^5C is enriched at the 5′UTR in human and mouse cells, confirming the discovery of a few independent studies [35, 36, 46]. Additionally, an unambiguous correlated methylation pattern is observed on 5′UTRs, but not on CDS and 3′UTR, in different mouse and human cell lines/tissues, suggesting a more complex aggregation pattern of m^5C that may be further characterized. Together, these observations strongly imply the functional relevance of m^5C RNA methylation and 5′UTR of mRNA. It is important to note that, although we failed to observe a correlated m^5C methylation pattern on CDS and 3′UTR regions of mRNA, it is still possible that such pattern may emerge on strictly matched cell lines/tissues.

When comparing the DNA and RNA methylome in matched cell lines in human and mouse, a negative correlation in the methylation level is observed on matched locus on DNA and RNA, which is quite surprising given that the methyltransferase of DNA and RNA may share strong sequence homology [52]. This anticorrelation pattern is consistent at all four CG-containing trinucleotide contexts and ruled out the possibility of sample contamination or off-target effect, which should both lead to false positive correlation in data. It is possible that there exists an underlying biomolecular mechanism that functions on the matched locus of DNA and RNA in parallel to ensure their orchestrated methylation status.

Similar to DNA methylation, a clustering effect of m^5C on mRNA is also observed in both human and mouse. The local dependency, that is, the adjacent cytosine locus often exhibits a similar methylation status, has been widely used in DNA methylation data analysis for more robust and accurate quantification of epigenetics status [57–59]. It is reasonable to expect that similar statistical approaches may be carried over into the field of single-base resolution RNA methylation data to enhance the analysis of bisulfite RNA methylation sequencing data. It is worth mentioning that, around 30%–43% of m^5C residuals exist in pairs in our results after filtering potential secondary structures that may lead to incomplete conversion and false positive m^5C sites. The number may be over- or underestimated because of the unfiltered secondary structure, which leads to an overestimation of the clustering effect, and structured regions excluded from the analysis, which may affect the estimation in both directions. It is necessary to develop a more sensitive unbiased approach that can eliminate the impact of the RNA structure to more accurately assess the distribution of transcriptome m^5C modification.

Intriguingly, we observed a strong enrichment of m^5C methylation on mitochondrial transcripts with more than 50 folds of enrichment. Previously, it was reported that methyltransferase NSUN5 can regulate mitochondrial gene expression [54], and we speculate that RNA m^5C methylation may play a more vital regulatory role in mitochondria-related biological processes.

Additionally, in order to have a glimpse of the dynamics of m^5C on mRNA, differential RNA methylation analysis was performed between breast cancer cell line MDA-MB-468 and the control cell line MCF10A; a total of 47 genes are reported to be differentially methylated, including RTN4, NME2, CASP14, HSPB1, RPL11, and RPS3, which are related to apoptosis and programmed cell death. Although we showed previously that m^5C on mRNA are more likely to be linked to the 5′UTR function, it is observed that the differential methylation sites between the breast cancer cell line and normal control cell lines are mostly located on the CDS and 3′UTR. These observations together implied a profound role of m^5C methylation on different regions of mRNA and in cancer pathology.

Interestingly, an overall positive correlation between RNA m^5C methylation and RNA expression level is observed in our mouse and human datasets, which added to the growing importance of mRNA m^5C methylation in regulating gene expression. Although the specific molecular mechanism is not yet clear, the observed positive correlation between RNA m^5C and RNA expression echoes our previous observed anticorrelation between DNA and RNA m^5C methylation from a different perspective, because it has been well established that DNA methylation is anticorrelated with RNA expression. However, as it is known that the most abundant RNA modification m^6A methylation may enhance or reduce the stability of the RNA molecule through interaction with different m^6A readers [14, 60] or regulate RN-protein interaction [13], it is reasonable to assume that RNA m^5C may have versatile functionalities, and may get dominated by a distinct mechanism under a specific condition.

In summary, our study presented an in-depth topological characterization of the m^5C RNA methylome in human and mouse. There are interesting patterns depicted and quantified, which call for further studies to explain novel biomolecular mechanisms.

Abbreviations

m^6A:	N^6-Methyladenosine
hm^5C:	Hydroxymethylcytosine
Ψ:	Pseudouridine
m^1A:	N^1-Methyladenosine
m^5C:	5-Methylcytosine
BS-Seq:	Bisulfite sequencing
MDA-MB-468:	MDA468
MEF:	Mouse embryo fibroblast
MEF-Dnmt2$^-$:	Mouse embryo fibroblast with Dnmt2 knockdown
WGBS:	Whole genome bisulfite sequencing
RRBS:	Reduced representation bisulfite sequencing
mRatio:	Methylation ratio
OR:	Odds ratio
DMS:	Differential methylation site
lncRNA:	Long noncoding RNA
sncRNA:	Small noncoding RNA.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

Jia Meng, João Pedro de Magalhães, and Yufei Huang conceived the project. Zhen Wei and Jia Meng carried out the computational analysis. Subbarayalu Panneerdoss, Santosh Timilsina, and Manjeet K. Rao performed the wet experiments, and Tabrez A. Mohammad and Yidong Chen performed bioinformatics support for BS-Seq. All authors provided valuable discussion and helped draft the manuscript. All authors read and approved the final manuscript. Zhen Wei, Subbarayalu Panneerdoss, Santosh Timilsina, and Jingting Zhu contributed equally to this work and are co-first author.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61401370 and 31671373 to Jia Meng, 81373469 to Zhi-Liang Lu), the Jiangsu Natural Science Foundation (BK20140403 to Jia Meng), the Jiangsu University Natural Science Program (16KJB180027 to Jia Meng), the US National Institutes of Health (R01GM113245 to Yufei Huang), the IIMS Translational Technology Resource (TTR) Award to Manjeet K. Rao and Yidong Chen, and NCI Cancer Center Shared Resources NCI P30CA54174 to Yidong Chen. Part of the BS-Seq experiment was performed by the Genome Sequencing Facility of the Greehey Children's Cancer Research Institute, UTHSCSA. The authors thank the computational support from the UTSA Computational Systems Biology Core, funded by the National Institute on Minority Health and Health Disparities (G12MD007591) from the National Institutes of Health.

Supplementary Materials

File 1. SI_figures_and_tables: additional figures and statistical analysis. File 2. SI_sheets_S1: differential methylation information between MCF10A and MDA468. File 3. SI_sheets_S2: differential methylation information between mouse ESC and mouse brain. File 4. SI_sheets_S3: differential methylation information between SRV-infected and -uninfected Jurkat cells. *(Supplementary Materials)*

References

[1] M. M. Suzuki and A. Bird, "DNA methylation landscapes: provocative insights from epigenomics," *Nature Reviews Genetics*, vol. 9, no. 6, pp. 465–476, 2008.

[2] K. D. Robertson, "DNA methylation and human disease," *Nature Reviews Genetics*, vol. 6, no. 8, pp. 597–610, 2005.

[3] L. E. Reinius, N. Acevedo, M. Joerink et al., "Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility," *PLoS One*, vol. 7, no. 7, article e41361, 2012.

[4] W. Xie, C. L. Barr, A. Kim et al., "Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome," *Cell*, vol. 148, no. 4, pp. 816–831, 2012.

[5] C. He, "Grand challenge commentary: RNA epigenetics?," *Nature Chemical Biology*, vol. 6, no. 12, pp. 863–865, 2010.

[6] U. Schibler, D. E. Kelley, and R. P. Perry, "Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells," *Journal of Molecular Biology*, vol. 115, no. 4, pp. 695–714, 1977.

[7] R. Desrosiers, K. Friderici, and F. Rottman, "Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 71, no. 10, pp. 3971–3975, 1974.

[8] D. T. Dubin and R. H. Taylor, "The methylation state of poly A-containing-messenger RNA from cultured hamster cells," *Nucleic Acids Research*, vol. 2, no. 10, pp. 1653–1668, 1975.

[9] H. Grosjean, *Fine-Tuning of RNA Functions by Modification and Editing*, Springer, Berlin, Heidelberg, 2005.

[10] X. Wang, B. S. Zhao, I. A. Roundtree et al., "N^6-Methyladenosine modulates messenger RNA translation efficiency," *Cell*, vol. 161, no. 6, pp. 1388–1399, 2015.

[11] J. M. Fustin, M. Doi, Y. Yamaguchi et al., "RNA-methylation-dependent RNA processing controls the speed of the circadian clock," *Cell*, vol. 155, no. 4, pp. 793–806, 2013.

[12] C. R. Alarcon, H. Lee, H. Goodarzi, N. Halberg, and S. F. Tavazoie, "N^6-Methyladenosine marks primary microRNAs for processing," *Nature*, vol. 519, no. 7544, pp. 482–485, 2015.

[13] N. Liu, Q. Dai, G. Zheng, C. He, M. Parisien, and T. Pan, "N^6-Methyladenosine-dependent RNA structural switches regulate RNA-protein interactions," *Nature*, vol. 518, no. 7540, pp. 560–564, 2015.

[14] X. Wang, Z. Lu, A. Gomez et al., "N^6-Methyladenosine-dependent regulation of messenger RNA stability," *Nature*, vol. 505, no. 7481, pp. 117–120, 2014.

[15] J. Zhou, J. Wan, X. Gao, X. Zhang, S. R. Jaffrey, and S.-B. Qian, "Dynamic m^6A mRNA methylation directs translational control of heat shock response," *Nature*, vol. 526, no. 7574, pp. 591–594, 2015.

[16] S. Geula, S. Moshitch-Moshkovitz, D. Dominissini et al., "m^6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation," *Science*, vol. 347, no. 6225, pp. 1002–1006, 2015.

[17] S. Hussain, J. Aleksic, S. Blanco, S. Dietmann, and M. Frye, "Characterizing 5-methylcytosine in the mammalian epitranscriptome," *Genome Biology*, vol. 14, no. 11, p. 215, 2013.

[18] A. L. Burgess, R. David, and I. R. Searle, "Conservation of tRNA and rRNA 5-methylcytosine in the kingdom Plantae," *BMC Plant Biology*, vol. 15, no. 1, pp. 199–117, 2015.

[19] L. Warren, P. D. Manos, T. Ahfeldt et al., "Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA," *Cell Stem Cell*, vol. 7, no. 5, pp. 618–630, 2010.

[20] X. Zhang, Z. Liu, J. Yi et al., "The tRNA methyltransferase NSun2 stabilizes $p16^{INK4}$ mRNA by methylating the $3'$-untranslated region of p16," *Nature Communications*, vol. 3, no. 1, p. 712, 2012.

[21] Y. Chen, H. Sierzputowska-Gracz, R. Guenther, K. Everett, and P. F. Agris, "5-Methylcytidine is required for cooperative binding of magnesium(2+) and a conformational transition at the anticodon stem-loop of yeast phenylalanine tRNA," *Biochemistry*, vol. 32, no. 38, pp. 10249–10253, 1993.

[22] Y. Motorin and M. Helm, "tRNA stabilization by modified nucleotides," *Biochemistry*, vol. 49, no. 24, pp. 4934–4944, 2010.

[23] F. Tuorto, R. Liebers, T. Musch et al., "RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and

protein synthesis," *Nature Structural & Molecular Biology*, vol. 19, no. 9, pp. 900–905, 2012.

[24] M. Schaefer, T. Pollex, K. Hanna et al., "RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage," *Genes & Development*, vol. 24, no. 15, pp. 1590–1595, 2010.

[25] C. T. Y. Chan, Y. L. J. Pang, W. Deng et al., "Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins," *Nature Communications*, vol. 3, no. 1, p. 937, 2012.

[26] K. Rai, S. Chidester, C. V. Zavala et al., "Dnmt2 functions in the cytoplasm to promote liver, brain, and retina development in zebrafish," *Genes & Development*, vol. 21, no. 3, pp. 261–266, 2007.

[27] C.-M. Wei, A. Gershowitz, and B. Moss, "Methylated nucleotides block 5′ terminus of HeLa cell messenger RNA," *Cell*, vol. 4, no. 4, pp. 379–386, 1975.

[28] S. Hussain, F. Tuorto, S. Menon et al., "The mouse cytosine-5 RNA methyltransferase NSun2 is a component of the chromatoid body and required for testis differentiation," *Molecular and Cellular Biology*, vol. 33, no. 8, pp. 1561–1570, 2013.

[29] S. Blanco, S. Dietmann, J. V. Flores et al., "Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders," *The EMBO Journal*, vol. 33, no. 18, pp. 2020–2039, 2014.

[30] C. S. Chow, T. N. Lamichhane, and S. K. Mahto, "Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications," *ACS Chemical Biology*, vol. 2, no. 9, pp. 610–619, 2007.

[31] M. Schaefer, "Chapter fourteen. RNA 5-methylcytosine analysis by bisulfite sequencing," *Methods in Enzymology*, vol. 560, pp. 297–329, 2015.

[32] M. Schaefer, T. Pollex, K. Hanna, and F. Lyko, "RNA cytosine methylation analysis by bisulfite sequencing," *Nucleic Acids Research*, vol. 37, no. 2, article e12, 2009.

[33] Y. Motorin, F. Lyko, and M. Helm, "5-Methylcytosine in RNA: detection, enzymatic formation and biological functions," *Nucleic Acids Research*, vol. 38, no. 5, pp. 1415–1430, 2010.

[34] V. Khoddami and B. R. Cairns, "Identification of direct targets and modified bases of RNA cytosine methyltransferases," *Nature Biotechnology*, vol. 31, no. 5, pp. 458–464, 2013.

[35] J. E. Squires, H. R. Patel, M. Nousch et al., "Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA," *Nucleic Acids Research*, vol. 40, no. 11, pp. 5023–5033, 2012.

[36] T. Amort, D. Rieder, A. Wille et al., "Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain," *Genome Biology*, vol. 18, no. 1, p. 1, 2017.

[37] X. Yang, Y. Yang, B. F. Sun et al., "5-Methylcytosine promotes mRNA export—NSUN2 as the methyltransferase and ALYREF as an m^5C reader," *Cell Research*, vol. 27, no. 5, pp. 606–625, 2017.

[38] J. M. Adams and S. Cory, "Modified nucleosides and bizarre 5′-termini in mouse myeloma mRNA," *Nature*, vol. 255, no. 5503, pp. 28–33, 1975.

[39] M. Salditt-Georgieff, W. Jelinek, J. E. Darnell, Y. Furuichi, M. Morgan, and A. Shatkin, "Methyl labeling of HeLa cell hnRNA: a comparison with mRNA," *Cell*, vol. 7, no. 2, pp. 227–237, 1976.

[40] E.-J. Lee, L. Pei, G. Srivastava et al., "Targeted bisulfite sequencing by solution hybrid selection and massively parallel sequencing," *Nucleic Acids Research*, vol. 39, no. 19, article e127, 2011.

[41] F. Neri, S. Rapelli, A. Krepelova et al., "Intragenic DNA methylation prevents spurious transcription initiation," *Nature*, vol. 543, no. 7643, pp. 72–77, 2017.

[42] F. Krueger, "Trim Galore. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files," 2015.

[43] D. Rieder, T. Amort, E. Kugler, A. Lusser, and Z. Trajanoski, "meRanTK: methylated RNA analysis ToolKit," *Bioinformatics*, vol. 32, no. 5, pp. 782–785, 2016.

[44] I. L. Hofacker, "Vienna RNA secondary structure server," *Nucleic Acids Research*, vol. 31, no. 13, pp. 3429–3431, 2003.

[45] M. Lawrence, W. Huber, H. Pagès et al., "Software for computing and annotating genomic ranges," *PLoS Computational Biology*, vol. 9, no. 8, article e1003118, 2013.

[46] X. Cui, Z. Wei, L. Zhang et al., "Guitar: an R/Bioconductor package for gene annotation guided transcriptomic analysis of RNA-related genomic features," *BioMed Research International*, vol. 2016, Article ID 8367534, 8 pages, 2016.

[47] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*, Springer Science & Business Media, New York, NY, USA, 2009.

[48] P. Du, X. Zhang, C.-C. Huang et al., "Comparison of beta-value and M-value methods for quantifying methylation levels by microarray analysis," *BMC Bioinformatics*, vol. 11, no. 1, p. 587, 2010.

[49] L. Liu, S. W. Zhang, Y. Huang, and J. Meng, "QNB: differential RNA methylation analysis for count-based small-sample sequencing data with a quad-negative binomial model," *BMC Bioinformatics*, vol. 18, no. 1, p. 387, 2017.

[50] 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.

[51] C. L. Zao, J. A. Ward, L. Tomanek, A. Cooke, R. Berger, and K. Armstrong, "Virological and serological characterization of SRV-4 infection in cynomolgus macaques," *Archives of Virology*, vol. 156, no. 11, pp. 2053–2056, 2011.

[52] M. G. Goll, F. Kirpekar, K. A. Maggert et al., "Methylation of $tRNA^{Asp}$ by the DNA methyltransferase homolog Dnmt2," *Science*, vol. 311, no. 5759, pp. 395–398, 2006.

[53] P. Qiu and L. Zhang, "Identification of markers associated with global changes in DNA methylation regulation in cancers," *BMC Bioinformatics*, vol. 13, Supplement 13, p. S7, 2012.

[54] Y. Cámara, J. Asin-Cayuela, C. B. Park et al., "MTERF4 regulates translation by targeting the methyltransferase NSUN4 to the mammalian mitochondrial ribosome," *Cell Metabolism*, vol. 13, no. 5, pp. 527–539, 2011.

[55] R. Pardal, M. F. Clarke, and S. J. Morrison, "Applying the principles of stem-cell biology to cancer," *Nature Reviews Cancer*, vol. 3, no. 12, pp. 895–902, 2003.

[56] B. Schwanhausser, D. Busse, N. Li et al., "Global quantification of mammalian gene expression control," *Nature*, vol. 473, no. 7347, pp. 337–342, 2011.

[57] X. Wang, J. Gu, L. Hilakivi-Clarke, R. Clarke, and J. Xuan, "DM-BLD: differential methylation detection using a hierarchical Bayesian model exploiting local dependency," *Bioinformatics*, vol. 33, no. 2, pp. 161–168, 2017.

[58] H.-U. Klein and K. Hebestreit, "An evaluation of methods to test predefined genomic regions for differential methylation in bisulfite sequencing data," *Briefings in Bioinformatics*, vol. 17, no. 5, pp. 796–807, 2016.

[59] S. Li, F. E. Garrett-Bakelman, A. Akalin et al., "An optimized algorithm for detecting and annotating regional differential methylation," *BMC Bioinformatics*, vol. 14, Supplement 5, p. S10, 2013.

[60] H. Huang, H. Weng, W. Sun et al., "Recognition of RNA N^6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation," *Nature Cell Biology*, vol. 20, no. 3, pp. 285–295, 2018.

Hindawi
International Journal of Genomics
Volume 2018, Article ID 9585383, 10 pages
https://doi.org/10.1155/2018/9585383

Research Article

Nonredundant, Highly Connected MicroRNAs Control Functionality in Breast Cancer Networks

Guillermo de Anda-Jáuregui,[1] Jesús Espinal-Enríquez,[1,2] Diana Drago-García,[3] and Enrique Hernández-Lemus[1,2]

[1]*Computational Genomics Division, National Institute of Genomic Medicine, 14610 Mexico City, Mexico*
[2]*Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, 04510 Mexico City, Mexico*
[3]*Department of Biological Regulation, Weizmann Institute of Science, 7610001 Rehovot, Israel*

Correspondence should be addressed to Enrique Hernández-Lemus; ehernandez@inmegen.gob.mx

Received 2 March 2018; Accepted 19 April 2018; Published 29 May 2018

Academic Editor: Changwon Park

Alterations to transcriptional regulation are an important factor in breast cancer. Noncoding RNA, such as microRNA (miR), have very influential roles in the transcriptional regulation of genes. Transcriptional regulation can be successfully modeled and analyzed using complex network theory. Particularly, interactions between two distinct classes of biological elements, such as miR and genes, can be approached through the bipartite network formalism. Based on bipartite network properties, it is possible to identify highly influential miRs in the network, such as those that have a large number of connections indicating regulation of a large set of genes. Some miRs in a network are nonredundant, which indicates that they are solely responsible of the regulation of a particular set of genes, which in turn may be associated to a particular biological process. We hypothesize that highly influential, nonredundant miRs, which we call *Commodore miRs* (Cdre-miRs), have an important role on the control of biological functions through transcriptional networks. In this work, we analyze the regulation of gene expression by miRs in healthy and cancerous breast tissue using bipartite miR-gene networks inferred from the Cancer Genome Atlas (TCGA) expression data. We observe differences in the degree, clustering coefficient and redundancy distributions for miRs and genes in the network, indicating differences in the way that these elements interact with each other. Furthermore, we identify a small set of five Cdre-miRs in the breast cancer network: miR-190b, miR-let7i, miR-292-b, miR-511, and miR-141. The neighborhood of genes controlled by each of these miRs is involved in particular biological functions such as dynein structure-associated processes, immune response, angiogenesis, cytokine activity, and cell motility. We propose that these Cdre-miRs are important control elements of biological functions deregulated in breast cancer.

1. Introduction

Breast cancer is a highly frequent cancer and one leading cause of death for women worldwide [1]. Aside from these important epidemiological aspects (and likely behind them) is the fact that breast cancer is a highly heterogeneous disease, both in its molecular origins and in its clinical manifestations; a fact that calls loudly for an improved understanding of the molecular mechanisms behind breast cancer development. Gene expression studies have provided with unprecedented information to characterize biomolecular activities leading to (or at least associated with) tumorigenesis. Recently, next generation sequencing has allowed us to accurately measure not only mRNA transcripts, but also regulatory noncoding RNA molecules such as microRNAs (miRs) and long noncoding RNAs (lncRNAs). miRs in particular have received a growing deal of attention due to their regulatory effects that seem to be fundamental for breast tumor establishment and progression [2].

In particular, the regulatory role of miRs seems to be central to processes involving cellular homoeostasis through processes such as apoptosis, proliferation, and migration that when deregulated give rise to well-known hallmarks of cancer [3–5]. Specific families of miRs playing either the role of

"oncogenes" or "tumor suppressors" are commonly referred to as "oncomiRs" [6]. Specific miRNA regulation and coregulation patterns have been linked to oncogenic processes [7], specifically regarding breast cancer onset and evolution [2].

miR regulation at the transcriptional and posttranscriptional levels often occurs by inducing processes leading to mRNA destabilization [8]; transcriptome profiling is thus increasingly useful for the analysis of miR regulation in genome-wide settings. On the other hand, miRs have also been associated with other mechanisms of regulatory activity [9], in particular by widening target upregulation or downregulation [10]. These regulatory interactions are involved in mechanisms that may ensure biological robustness [11]. Regulatory relationships between miRs and genes (including transcription factors and other regulatory elements in the genome) are able to coexpress profiles that may be phenotype-inducing [12].

In order to improve our current knowledge on these matters, in particular regarding the transcriptional relation between miRNAs and their target genes (mRNAs), we propose the use of regulatory networks, using expression data from primary breast cancer tissue and matched control tissue. In our previous effort [13], we identified differences in the regulatory interactions of genes by other genes and by miRs; therefore, we decided to further explore these relationships using the bipartite graph formalism.

Bipartite networks are graphs composed by two disjoint sets of nodes and a set of edges such that each edge is surrounded by a node of each class. The two disjoint sets of nodes can be thought of as a *top* and *bottom* layers of nodes. The nature of bipartite graphs allows for the identification of topological parameters that describe both sets of nodes that are uniquely defined in this context. [14]. In this work, we will focus on three parameters: degree distributions, clustering coefficient, and redundancy coefficient.

1.1. Degree. The degree of a node is the number of edges that connect a node to its neighbors; it is thus the simplest measure of node connectivity. The degree distribution of a network is the most basic descriptor of a network [15]. In a bipartite network, the degree distribution is more informative if calculated for each set of nodes separately, as the connectivity of nodes in each set may be different. Thus, the degree distribution for a given set can be defined as follows:

$$P(K_{\text{NodeSet}}) = \frac{nk}{n}. \tag{1}$$

1.2. Clustering Coefficient. The clustering coefficient of a node is a measure of the local density of connections in which a given node participates. The clustering coefficient of a network is thus a measure of the cohesiveness of the network [16]. The clustering coefficient in a bipartite node is defined as follows:

$$\text{CC}(u) = \frac{\sum_{v \in N(N(v))} (|N(u) \cap N(v)| / |N(u) \cup N(v)|)}{|N(N(u))|}. \tag{2}$$

1.3. Redundancy Coefficient. Nodes of the same type can only be connected by being connected through nodes of another type. The redundancy coefficient of a node is the fraction of pairs of neighbors of the said node that are both linked to other nodes [16]; it is a measure of the importance of a given node in a layer for the connection of the nodes in the other layer. Redundant nodes can be removed from their layer without causing a disconnection of the nodes in the other layer, as seen in Figure 1. The formal definition for the redundancy coefficient is as follows [17]:

$$\text{RC}(u) = \frac{\left| \left\{ \{u, w\} \subseteq N(v), \exists v' \neq v, \left(v', u\right) \in E, \left(v', w\right) \in E \right\} \right|}{|N(v)|(|N(v)| - 1)/2}. \tag{3}$$

Noncoding RNAs, such as miRs, have a regulatory biological role. Meanwhile, genes may be involved in several biological activities, as they are translated into proteins which can have various functional roles. Usually, biological functions involve the interaction of several biomolecules to produce physiologically observable phenomena [18]. By regulating a set of genes, a given miR can in fact control a number of biological functions in a particular phenotype. We should take into account the following two considerations based on topological properties of a bipartite network:

(i) miRs with small neighborhoods (i.e., with low degree) are influencing few genes, being less likely to have effect on a biological function.

(ii) miRs with high redundancy are less likely to be necessary for a given group of genes to be coregulated, as there are other miRs that will keep the group connected.

Therefore, we consider that highly connected, nonredundant miRs may have a major role in the control of biological functions through transcriptional networks: removal of these miRs would lead to the disconnection of a group of genes, which would involve a loss of the concerted regulation of this gene set, and in turn, of biological functionality. We have termed these miRs, *Commodore miRs* (Cdre-miRs).

2. Materials and Methods

2.1. miR-mRNA Bipartite Network Model. We constructed a bipartite network representing transcriptional regulation of genes by miRs in breast tissue, in the physiological and cancer states, hereafter referred to as the *healthy* and *cancer* groups.

We used gene and miR transcription profiles from samples from the Cancer Genome Atlas [19] breast cancer dataset [20]. Data processing and handling was performed following the pipeline previously implemented by our group [13] (Supplementary Material 1). The data processing pipeline included a differential expression analysis using DESeq2 [21] in order to identify the log fold change (LFC) of gene expression between the cancer and healthy groups.

Regulatory interactions between miRs and genes were inferred in terms of information theory. Mutual information (MI) [22] was calculated for each miR-gene pair, using the

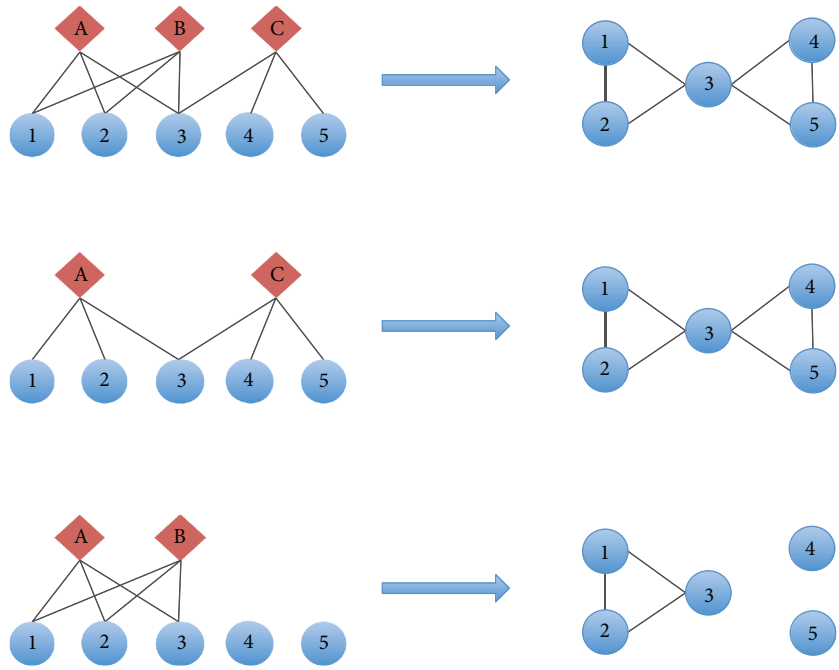


Figure 1: In this figure, we show a simple bipartite graph with 3 top nodes (A to C) and 5 bottom nodes (1 to 5). The projection of bottom nodes is shown on the right side. If node B is removed, no link in the projection is lost, as all bottom nodes connected through B are also connected through A; therefore, B is *redundant*. Meanwhile, removing node C causes the loss of connections of bottom nodes 4 and 5; therefore, C is *nonredundant*.

transcription profiles for either the healthy or cancer groups. Then, these interactions were filtered to conserve the 99.741%. MI calculation was performed using the engine of our in-house parallelized implementation of ARACNE [23].

Acknowledging that the canonical mode of action is that miRs regulate gene expression, we take this into account to interpret that mutual information interactions represent a directed relationship, from miR to gene; hence, higher values of mutual information indicate that the involved miR regulates directly the expression of the target gene.

2.2. Bipartite Network Analysis. Each network, healthy and cancer, was analyzed to identify general topological features, including number of miR and gene nodes, number of edges, and number of connected components. For each set of nodes, miR or genes, distributions were calculated for degree, clustering coefficient, and redundancy. Network analysis was performed using a combination of NetworkX for Python [16], igraph for R [24], and Cytoscape [25], as implemented previously [26].

2.3. Cdre-miR Identification. To identify highly influential, nonredundant Cdre-miRs, we separated all miRs in each network using two thresholds: a degree threshold of 100 and a redundancy threshold of 0.5.

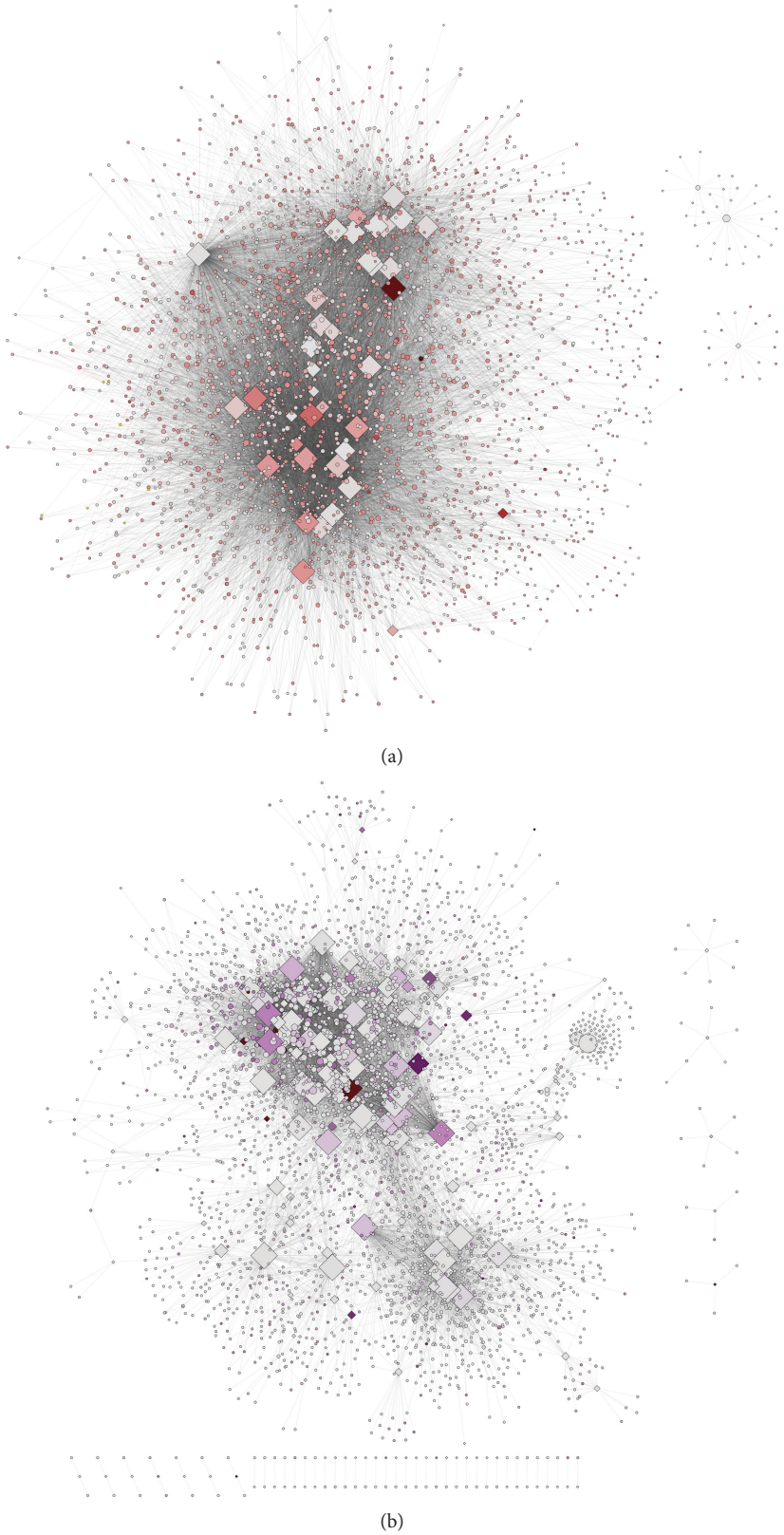
2.4. Enrichment Analysis. For each Cdre-miR identified, we performed an over-representation analysis (ORA) of Gene Ontology (GO) biological processes [27, 28]. We performed the analysis using the WebGestalt [29] portal, with default parameters (number of genes in category between 5 and 2000, Benjamini-Hochberg multiple test adjustment, top 10 most significant processes).

3. Results

3.1. miR-Gene Bipartite Networks in Health and Cancer. We constructed two comparable networks representing the two states: healthy and cancer. Figure 2 shows a visualization of these networks. Table 1 contains the basic descriptors for each network; these networks are also provided, as GML files, in Supplementary Material 2. Figure 3 shows the cumulative frequency distribution for redundancy values for miRs and genes in healthy and cancer networks. Other distribution plots including degree and clustering coefficient of gene nodes are provided in Supplementary Material 2.

3.2. Identification of Cdre-miRs. In Figure 4, we show a scatter plot of miRs, with the y-axis representing degree and the x-axis representing redundancy. We divide this plot in four sectors delimited by the thresholds selected to define high connectivity (degree larger than 100) and low redundancy (redundancy lower than 0.5). Five miRs were identified to be highly connected and nonredundant in the cancer setting: miR-190b, miR-let7i, miR-292-b, miR-511, and miR-141. Importantly, no Cdre-miRs were identified in healthy breast tissue.

Figure 4(a) shows the five Cdre-miRs and their first gene neighbors, colored by LFC. As it is expected, there is little overlap between the neighborhoods of each Cdre-miR; however, they are not completely isolated from each other, as there is a small subset of genes that are regulated by more than one Cdre-miR. Figure 4(b) shows the 3



(a)

(b)

Figure 2: Network visualizations. (a) Healthy breast tissue. (b) Breast cancer tissue. Node color intensity in both networks is proportional to expression levels.

Table 1: Bipartite network parameters.

Parameter	Healthy	Cancer
Nodes, miR	97	414
Nodes, gene	2967	3240
Edges	16,589	14,063
Average degree, miR	171.02	33.97
Average degree, gene	5.59	4.34
Average clustering coefficient, miR	0.28	0.18
Average clustering coefficient, gene	0.24	0.27

Gene Ontology biological processes more enriched in the gene neighborhoods of each Cdre-miR. Complete enrichment analysis results may be found in Supplementary Material 3. In this panel, we also include the enrichment of the neighborhood intersection of miR-511 and let-7i: the *innate immunity* process.

4. Discussion

Understanding the role of miR-gene regulation is crucial to unveil the mechanisms behind tumorigenesis and progression in breast cancer. Bipartite networks offer a powerful tool to analyze the behavior of these interactions and at the same time the structural and functional relevance of specific miRs in the transcriptional regulatory program.

In this work, we have shown that the bipartite networks representing miR-gene regulation in health and breast cancer exhibit notable structural differences. This can be observed in terms of structural parameters, including degree, clustering, and redundancy coefficients (as seen in Table 1, Figures 3 and 5, and Supplementary Material 4); the overall distribution of these parameters is qualitatively different in each network, which indicates that they have essentially different topologies. This result is in agreement with the widely understood notion that miR regulation is altered in cancer.

miRs are, on average, more connected in the health context than in cancer. They also happen to be more redundant, as seen in Figure 4. Biologically, this represents an abundance of joint miR regulation over sets of genes in health, which provides a mechanism that confers a structural robustness and a stronger maintenance of the transcriptional regulatory program in a normal tissue. Meanwhile, miRs in cancer tend to be less connected on average and also less redundant.

4.1. Cdre-miRs and Their Role in Breast Cancer. A major finding in this work is that the presence of highly connected, nonredundant miRs, which we define as *Commodores*, is a phenomenon only observed in the breast cancer network. Cdre-miRs emerge as important gene expression regulators, as each of them alone is responsible for the regulation of large groups of genes. Perhaps, more interesting is the fact that, through the regulation of these gene sets, each Cdre-miR is an important control element of specific biological processes.

For instance, miR-141 is a widely studied oncomiR [30–38]. Our group has previously observed that this miR is one of the most relevant mediators of structural processes in breast cancer and particularly involved in the epithelial-to-mesenchymal transition (EMT), as well as in the opposite mesenchymal-to-epithelial transition (MET) [13]. Here, we show that its functional importance is reflected by its crucial role in the miR-gene network. The processes that mir-141 controls through the regulation of its gene neighborhood, which includes genes such as VIM and ZEB1-2, are directly related to EMT-MET: motility, cell migration, and extracellular matrix (ECM) organization.

Another miR that has been previously studied in the context of breast cancer is miR-190b. Previously, it has been shown that this miR is associated with estrogen-positive breast cancer [39] and hormone therapy resistance [40]. However, its specific role in the pathogenesis of the disease has not been described. We identified mir-190b as a Cdre-miR that is associated to the regulation of dynein assembly, vitamin metabolism, and cell proliferation of mammary epithelial cells, a well-known hallmark of cancer. As a Cdre-miR, it is possible that deregulation of this miR is mechanistically involved in the acquisition of these distinctive cancer features.

Other important cancer features are controlled by miR-29b. This miR has been found in serum of patients with breast cancer [41] and overexpressed in cervical cancer [42]. In this work, we identify it as an important regulator of transport, angiogenesis, and epithelial cell migration, processes that are predominantly controlled by this miR.

4.2. Concurrent Regulation of Innate Immune Response by Cdre-miRs. So far, we have identified processes that are uniquely controlled by a single Cdre-miR. However, we also found processes that are controlled by two different Cdre-miRs, through the regulation of different gene neighborhoods. For instance, let-7i controls genes related to leukocyte cell-cell adhesion, adaptive immune response, and cell activation. Meanwhile, miR-511 regulates genes associated to cytokine production, cell activation, and adaptive immune response; that is, they each control related, but nonoverlapping, processes.

Importantly, the overlap between these two miRs is the largest observed between any pair of Cdre-miRs, but is less than 25% of the size of either miRs' neighborhood (53 out of 238 and 213, resp.). Perhaps more interestingly, by evaluating the enrichment of the intersection of these two Cdre-miRs, we identify that their jointly regulated genes highly overrepresent the innate immune response process, which is not significantly enriched by neither of neighborhood alone. This can be interpreted as an instance of a coregulated process that requires combined action of two different Cdre-miRs.

5. Conclusion

Using bipartite networks, we identified differences in the miR-gene regulation between healthy breast tissue and breast cancer. We christened those miRs that are highly connected and nonredundant as *Commodore miRs*. The emergence of Cdre-miRs is a network structural property that is only found in the context of breast cancer.

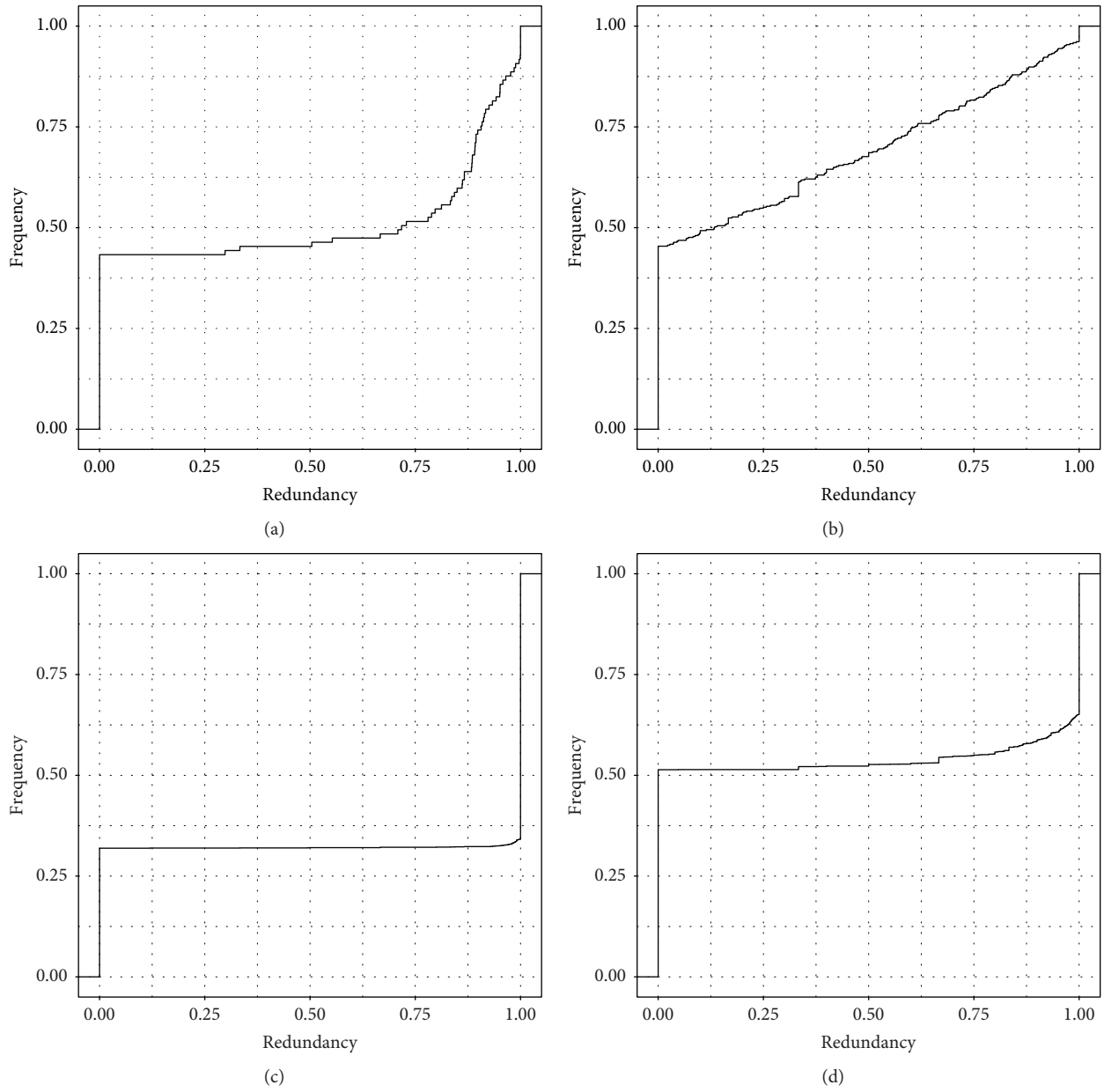


Figure 3: Redundancy coefficient cumulative frequency distribution plots. (a) miR, healthy. (b) miR, cancer. (c) gene, healthy. (d) gene, cancer. In each panel, x-axis represents redundancy coefficient values, and y-axis represents the normalized, accumulated frequency.

We identified five Cdre-miRs: miR-190, miR-29b, miR-141, miR-511, and let-7i that regulate specific biological processes. miR-190 is involved in cell structure, proliferation, and metabolism. miR-29b regulates cell transport, migration, and angiogenesis. miR-141 controls cell motility and extracellular matrix organization. miR-511 and let-7i independently regulate genes associated to cell adhesion and adaptive immune response through different gene sets. The latter shows that nonredundancy does not interfere with network robustness at a functional level.

Furthermore, innate immune response is a function which is not controlled by a single Cdre-miR, but it emerges in the context of the neighborhood overlap of miR-511 and let-7i. This is an example of how cooperation allows the emergence of new features that provide biological robustness to breast cancer. Identifying these elements is a step forward to identify actionable elements that provide controllability over the transcriptional networks.

Cdre-miRs may serve as novel biomarkers in breast cancer, which may be used to identify perturbation of biological functions in pathological phenotypes. Currently, therapeutic targeting of miRs in clinical oncology is limited; however, as technologies advance, the use of systems such as antagomiRs, microRNA mimics, and reporter systems will be available. In this context, we propose that Cdre-miRs may be attractive targets for the disruption of processes that are favorable for tumor growth and survival. With this in mind, Cdre-miRs may be an important

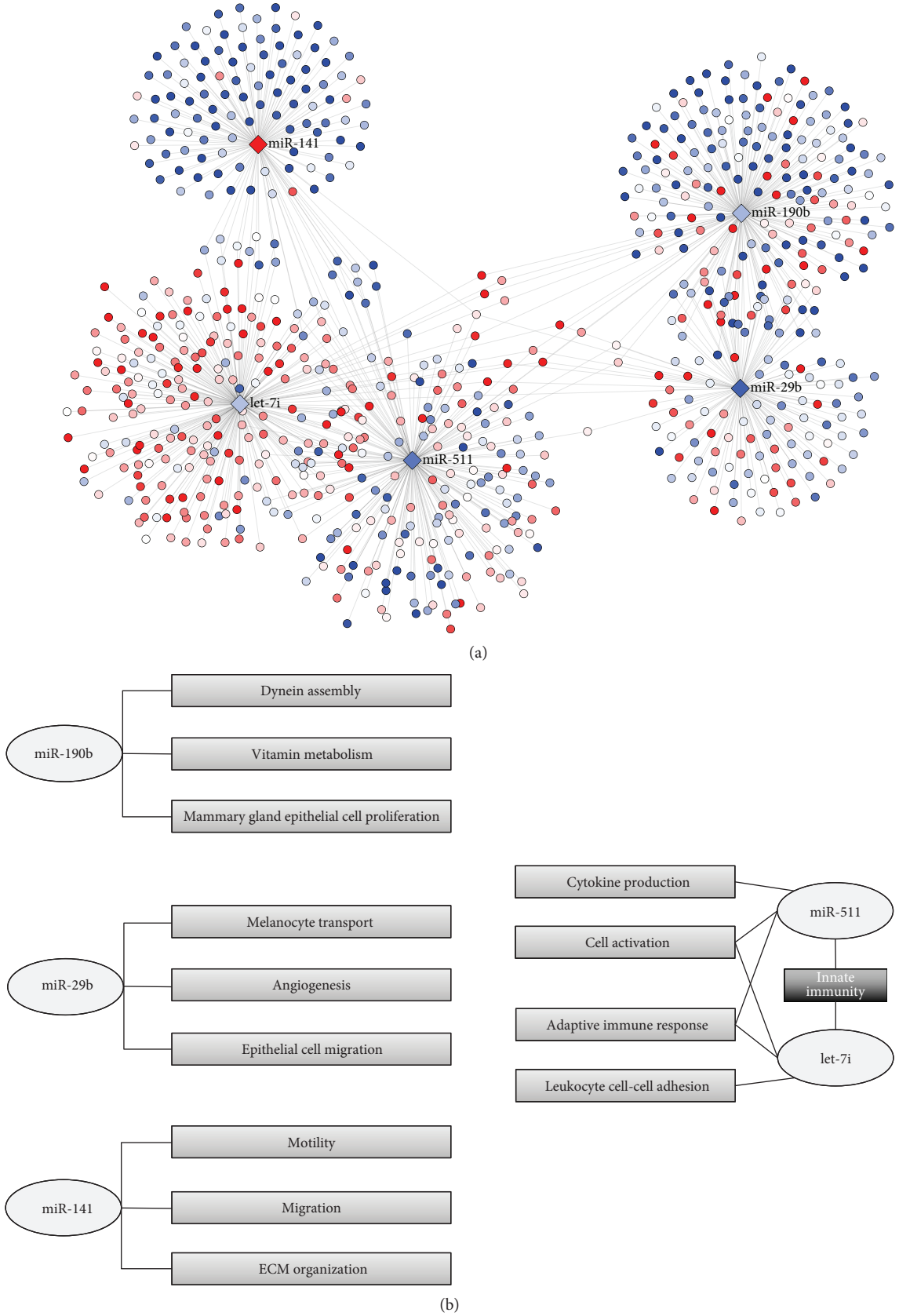


Figure 4: (a) shows the five Cdre-miRs and their gene neighbors. Nodes are colored in blue if subexpressed or red if overexpressed. (b) shows the GO biological processes enriched in each Cdre-miR neighborhood, in light grey. Additionally, the process enriched in the neighborhood intersection of mir-511 and let-7i, *innate immunity*, is shown in dark grey.

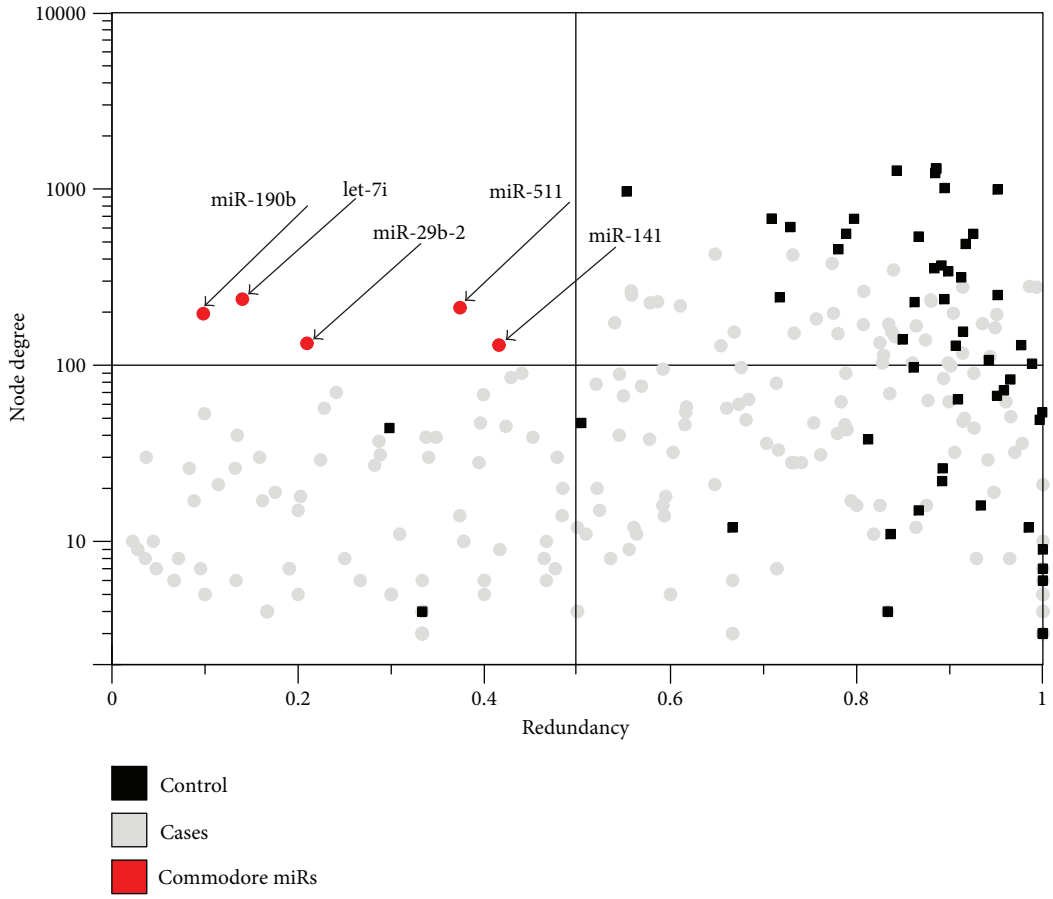


FIGURE 5: This scatterplot shows miR nodes in the healthy and cancer networks: grey dots for cancer and black squares for healthy. The *x*-axis represents the redundancy coefficient, and *y*-axis represents the degree of a miR. Only nodes with degrees larger than 2 are represented, as redundancy is not defined for nodes with a lower degree value. The plot is divided by lines representing thresholds for redundancy (0.5) and degree (100). The upper left sector of the plot is populated by five Cdre-miR nodes (red dots) with high connectivity and low redundancy.

element in the development of strategies for precision medicine in breast cancer. We strongly believe that further research in the role of Cdre-miRs in cancer shall provide a novel approach to understand the disease.

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 they have no conflicts of interest.

Authors' Contributions

Guillermo de Anda-Jáuregui and Jesús Espinal-Enríquez contributed equally to this work.

Acknowledgments

The research leading to these results has received funding from Grant no. 285544/2016 from the Consejo Nacional de Ciencia y Tecnología (CONACYT) as well as federal funding from the National Institute of Genomic Medicine (INMEGEN). One of the authors (Enrique Hernández-Lemus) also acknowledges support from the 2016 Marcos Moshinsky Research Chair in the Physical Sciences.

Supplementary Materials

Supplementary Material 1. Compressed (Zip) file containing the graphic pipeline followed in this work. We use raw counts from RNA-seq and miRNA-seq for breast tumors and adjacent control ("healthy") breast tissue from Genome Data Commons. MicroRNA preprocessing included mature count annotation, low-count filtering, and TMM normalization; RNA-seq preprocessing included mRNA annotation, RSEM normalization, and low-count filtering. This preprocessing is thoroughly described in [13]. These counts were integrated in an expression matrix which was used for a differential expression analysis (between cases and controls), and the inference of mutual information miR-gene networks. The final step is the identification of Cdre-miRs. Supplementary Material 2. Compressed (Zip) file containing GML representations of networks. Supplementary Material 3. Compressed

(Zip) file containing ORA results for each Cdre-miR gene neighborhood. Supplementary Material 4. Compressed (Zip) file containing degree and clustering coefficient distribution plots. *(Supplementary Materials)*

References

[1] J. Ferlay, I. Soerjomataram, R. Dikshit et al., "Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012," *International Journal of Cancer*, vol. 136, no. 5, pp. E359–E386, 2015.

[2] E. O'Day and A. Lal, "MicroRNAs and their target gene networks in breast cancer," *Breast Cancer Research*, vol. 12, no. 2, p. 201, 2010.

[3] D. Hanahan and R. A. Weinberg, "The hallmarks of cancer," *Cell*, vol. 100, no. 1, pp. 57–70, 2000.

[4] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: the next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011.

[5] M. Garofalo and C. M. Croce, "microRNAs: master regulators as potential therapeutics in cancer," *Annual Review of Pharmacology and Toxicology*, vol. 51, no. 1, pp. 25–43, 2011.

[6] W. C. S. Cho, "OncomiRs: the discovery and progress of microRNAs in cancers," *Molecular Cancer*, vol. 6, no. 1, p. 60, 2007.

[7] Y. Cai, X. Yu, S. Hu, and J. Yu, "A brief review on the mechanisms of miRNA regulation," *Genomics, Proteomics & Bioinformatics*, vol. 7, no. 4, pp. 147–154, 2009.

[8] D. Baek, J. Villén, C. Shin, F. D. Camargo, S. P. Gygi, and D. P. Bartel, "The impact of microRNAs on protein output," *Nature*, vol. 455, no. 7209, pp. 64–71, 2008.

[9] M. Selbach, B. Schwanhäusser, N. Thierfelder, Z. Fang, R. Khanin, and N. Rajewsky, "Widespread changes in protein synthesis induced by microRNAs," *Nature*, vol. 455, no. 7209, pp. 58–63, 2008.

[10] S. Vasudevan, Y. Tong, and J. A. Steitz, "Switching from repression to activation: microRNAs can up-regulate translation," *Science*, vol. 318, no. 5858, pp. 1931–1934, 2007.

[11] J. A. Vidigal and A. Ventura, "The biological functions of miRNAs: lessons from *in vivo* studies," *Trends in Cell Biology*, vol. 25, no. 3, pp. 137–147, 2015.

[12] G. Diaz, F. Zamboni, A. Tice, and P. Farci, "Integrated ordination of miRNA and mRNA expression profiles," *BMC Genomics*, vol. 16, no. 1, p. 767, 2015.

[13] D. Drago-García, J. Espinal-Enríquez, and E. Hernández-Lemus, "Network analysis of EMT and MET micro-RNA regulation in breast cancer," *Scientific Reports*, vol. 7, no. 1, p. 13534, 2017.

[14] J.-L. Guillaume and M. Latapy, "Bipartite graphs as models of complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 371, no. 2, pp. 795–813, 2006.

[15] M. Newman, *Networks: An Introduction*, Oxford University Press, 2010.

[16] A. A. Hagberg, D. A. Schult, and P. J. Swart, "Exploring network structure, dynamics, and function using NetworkX," in *Proceedings of the 7th Python in Science Conference*, pp. 11–15, Pasadena, CA USA, 2008.

[17] M. Latapy, C. Magnien, and N. Del Vecchio, "Basic notions for the analysis of large two-mode networks," *Social Networks*, vol. 30, no. 1, pp. 31–48, 2008.

[18] M. A. García-Campos, J. Espinal-Enríquez, and E. Hernández-Lemus, "Pathway analysis: state of the art," *Frontiers in Physiology*, vol. 6, 2015.

[19] The Cancer Genome Atlas Research Network, J. N. Weinstein, E. A. Collisson et al., "The cancer genome atlas pan-cancer analysis project," *Nature Genetics*, vol. 45, no. 10, pp. 1113–1120, 2013.

[20] C. X. Ma and M. J. Ellis, "The cancer genome atlas: clinical applications for breast cancer," *Oncology*, vol. 27, no. 12, pp. 1263–1269, 2013, 1274–9.

[21] 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.

[22] E. Hernández-Lemus and J. M. Siqueiros-García, "Information theoretical methods for complex network structure reconstruction," *Complex Adaptive Systems Modeling*, vol. 1, no. 1, p. 8, 2013.

[23] A. A. Margolin, I. Nemenman, K. Basso et al., "ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context," *BMC Bioinformatics*, vol. 7, Supplement 1, p. S7, 2006.

[24] G. Csardi and T. Nepusz, "The igraph software package for complex network research," *InterJournal, Complex Systems*, vol. 1695, 2006.

[25] P. Shannon, A. Markiel, O. Ozier et al., "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Research*, vol. 13, no. 11, pp. 2498–2504, 2003.

[26] G. de Anda-Jáuregui, K. Guo, B. A. McGregor, and J. Hur, "Exploration of the anti-inflammatory drug space through network pharmacology: applications for drug repurposing," *Frontiers in Physiology*, vol. 9, p. 151, 2018.

[27] 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.

[28] The Gene Ontology Consortium, "Expansion of the gene ontology knowledgebase and resources," *Nucleic Acids Research*, vol. 45, D1, pp. D331–D338, 2017.

[29] J. Wang, D. Duncan, Z. Shi, and B. Zhang, "WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013," *Nucleic Acids Research*, vol. 41, no. W1, pp. W77–W83, 2013.

[30] Y.-S. Yao, W.-S. Qiu, R.-Y. Yao et al., "miR-141 confers docetaxel chemoresistance of breast cancer cells via regulation of EIF4E expression," *Oncology Reports*, vol. 33, no. 5, pp. 2504–2512, 2015.

[31] H.-P. Tsai, S.-F. Huang, C.-F. Li, H.-T. Chien, and S.-C. Chen, "Differential microRNA expression in breast cancer with different onset age," *PLoS One*, vol. 13, no. 1, article e0191195, 2018.

[32] X. Li, S. Roslan, C. N. Johnstone et al., "MiR-200 can repress breast cancer metastasis through ZEB1-independent but moesin-dependent pathways," *Oncogene*, vol. 33, no. 31, pp. 4077–4088, 2014.

[33] P. Li, T. Xu, X. Zhou et al., "Downregulation of miRNA-141 in breast cancer cells is associated with cell migration and invasion: involvement of ANP32E targeting," *Cancer Medicine*, vol. 6, no. 3, pp. 662–672, 2017.

[34] J. Finlay-Schultz, D. M. Cittelly, P. Hendricks et al., "Progesterone downregulation of miR-141 contributes to expansion of stem-like breast cancer cells through maintenance of progesterone receptor and Stat5a," *Oncogene*, vol. 34, no. 28, pp. 3676–3687, 2015.

[35] B. G. Debeb, L. Lacerda, S. Anfossi et al., "miR-141-mediated regulation of brain metastasis from breast cancer," *Journal*

of the National Cancer Institute, vol. 108, no. 8, article djw026, 2016.

[36] V. Damiano, G. Brisotto, S. Borgna et al., "Epigenetic silencing of miR-200c in breast cancer is associated with aggressiveness and is modulated by ZEB1," *Genes, Chromosomes & Cancer*, vol. 56, no. 2, pp. 147–158, 2017.

[37] M. Alunni-Fabbroni, L. Majunke, E. K. Trapp et al., "Whole blood microRNAs as potential biomarkers in post-operative early breast cancer patients," *BMC Cancer*, vol. 18, no. 1, p. 141, 2018.

[38] N. Abedi, S. Mohammadi-Yeganeh, A. Koochaki, F. Karami, and M. Paryan, "miR-141 as potential suppressor of β-catenin in breast cancer," *Tumour Biology*, vol. 36, no. 12, pp. 9895–9901, 2015.

[39] G. Cizeron-Clairac, F. Lallemand, S. Vacher, R. Lidereau, I. Bieche, and C. Callens, "M*iR-190b*, the highest upregulated miRNA in ERα-positive compared to ERα-negative breast tumors, a new biomarker in breast cancers?," *BMC Cancer*, vol. 15, no. 1, p. 499, 2015.

[40] T. Joshi, D. Elias, J. Stenvang et al., "Integrative analysis of miRNA and gene expression reveals regulatory networks in tamoxifen-resistant breast cancer," *Oncotarget*, vol. 7, no. 35, pp. 57239–57253, 2016.

[41] O. Shaker, M. Maher, Y. Nassar, G. Morcos, and Z. Gad, "Role of microRNAs -29b-2, -155, -197 and -205 as diagnostic biomarkers in serum of breast cancer females," *Gene*, vol. 560, no. 1, pp. 77–82, 2015.

[42] L. S. Servín-González, A. J. Granados-López, and J. A. López, "Families of microRNAs expressed in clusters regulate cell signaling in cervical cancer," *International Journal of Molecular Sciences*, vol. 16, no. 6, pp. 12773–12790, 2015.

Hindawi
International Journal of Genomics
Volume 2018, Article ID 8260652, 11 pages
https://doi.org/10.1155/2018/8260652

Review Article

The Progress of Methylation Regulation in Gene Expression of Cervical Cancer

Chunyang Feng, Junxue Dong, Weiqin Chang, Manhua Cui, and Tianmin Xu

The Second Hospital of Jilin University, Jilin, Changchun 130041, China

Correspondence should be addressed to Tianmin Xu; xutianmin@126.com

Received 27 December 2017; Revised 2 March 2018; Accepted 27 March 2018; Published 16 April 2018

Academic Editor: Yujing Li

Cervical cancer is one of the most common gynecological tumors in females, which is closely related to high-rate HPV infection. Methylation alteration is a type of epigenetic decoration that regulates the expression of genes without changing the DNA sequence, and it is essential for the progression of cervical cancer in pathogenesis while reflecting the prognosis and therapeutic sensitivity in clinical practice. Hydroxymethylation has been discovered in recent years, thus making 5-hmC, the more stable marker, attract more attention in the field of methylation research. As markers of methylation, 5-hmC and 5-mC together with 5-foC and 5-caC draw the outline of the reversible cycle, and 6-mA takes part in the methylation of RNA, especially mRNA. Furthermore, methylation modification participates in ncRNA regulation and histone decoration. In this review, we focus on recent advances in the understanding of methylation regulation in the process of cervical cancer, as well as HPV and CIN, to identify the significant impact on the prospect of overcoming cervical cancer.

1. Introduction

Cervical cancer, which is one of the three most common gynecological tumors, has been the fourth leading cause of cancer-associated death among women worldwide, as well as becoming the second most commonly diagnosed cancer in developing countries. According to statistics, newly diagnosed cases and cervical cancer-associated deaths are approximately 520,000 and 260,000, respectively, every year, which affected youth trends more clearly [1]. It is widely recognized that persistent infection of high-risk-HPV (hr-HPV) accounts for the process from cervical intraepithelial neoplasia (CIN) to neoplasms, and vaccines of HPV and application of screening methods contribute a lot towards cervical carcinoma prevention. However, for established infections, vaccines have limited function and full-type coverage has not been achieved yet [1]. Additionally, as the 5-year survival rate is about 15% among advanced patients, the prognosis still remains unoptimistic in the late stages [2, 3]. Hence, it cries out for investigating the underlying molecular mechanisms on different biological expression levels to understand the genesis and progression of cervical cancer.

While gene mutation theory is incapable of providing reasonable explanations for many biological changes in tumor development, epigenetic alteration is drawing more attention, which involves modifications such as methylations of DNA and RNA, acetylations of histone, and regulations of ncRNA and aberrant chromatin. Methylated modification is extensively studied these years. DNA methylation mainly occurs at CpG islands where the methyltransferase DNMT family mediates the transfer of a methyl group to cytosines, generating 5-methylcytosine (5-mC), which can be oxidized into 5-hydroxymethylcytosine (5-hmC), 5-foC, and 5-caC by TET proteins step by step, so that methylation is achieved reversibly [4, 5]. Methylation decoration in RNAs is as common as it is in DNAs. M6A is one of the markers in mRNA methylation, and modifications take place in nascent pre-mRNAs predominantly [6]. Additionally, miR-RNAs and lnc-RNAs take part in epigenetic modifications themselves, and their biological functions are affected by the methylation state at the same time.

In this article, we summarize several recent studies of methylation regulation in the field of cervical cancer and discuss the potential of these molecular mechanisms in the

period of gene expression, to get some enlightenment in epigenetics to carry forward the prevention and treatment of cervical cancer.

2. Hydroxymethylation and Cervical Cancer

2.1. Hydroxymethylation and Its Regulations. In 1972, 5-hmC was initially found in bacteriophages and then in mammalian DNA. Currently, 5-hmC, a more stable epigenetic mark than 5-mC, plays an important role in epigenetics and works as an intermediate in demethylation [7]. It has been confirmed that the brain has the highest concentration of 5-hmC, while the rectum, liver, colon, and kidney are subordinate. In contrast, 5-hmC is at a low level in the lung, placenta, and breast [8]. The regulation of DNA hydroxymethylation is mediated by several factors, among which human ten-eleven translocation (TET) is identified as a dioxygenase for converting 5-mC to 5-hmC; meanwhile, αKG, Fe^{2+}, and ascorbate may activate the TET proteins as cofactors [9].

The TET protein family consists of TET1, TET2, and TET3, and their C-terminal catalytic domains come from a high degree of homology, which can be regulated by CXXC finger protein 1 (CFP1). Different CXXC domains have different functions; the CXXC5 domain of TET2 is able to downregulate TET2 with a 5-hmC decrease. But CXXC4 was found to be binding to the unmethylated DNA of TET1, TET2, and TET3, which then starts a caspase-dependent degradative process [10]. Some researchers found that in TET1-lacking cells, 5-hmC was reduced while 5-mC was increased. Moreover, TET1 can control 5-hmC by regulating hydroxylase activity to convert 5-mC to 5-hmC, which is HIF-1 dependent; at the same time, TET1 can also bind to CpG regions to stop some DNA methyltransferase activity [11]. It was demonstrated that TET3 is important for proper DNA repair, cell survival, and promotion of 5-hmC [12]. Besides, 5-hmC levels are also partly regulated by microRNAs. There are also some genes regulating 5-hmC, such as IDH1, IDH2, SDH, and FH [13]. Those factors are linked to the alteration of 5-hmC levels in cancer.

2.2. DNA Hydroxymethylation in Cervical Cancer and Other Cancers. To have cervical cancer treated and diagnosed precisely, many researches about 5-mC and other epigenetic modifications of cervical cancer aim to find treatment methods and diagnostic markers. But 5-hmC of cervical cancer is less researched, as there are only two articles about 5-hmC in cervical cancer.

Zhang et al. used immunohistochemistry to detect the expression of 5-hmC, 5-mC, and TET1/2/3 in 140 cervical squamous cell carcinoma (CSCC) tissues and 40 normal cervical tissues. They found that the expression of 5-hmC was an independent prognostic factor of squamous cell carcinoma, and compared with normal cervix tissues, the level of 5-mC was increased while 5-hmC was significantly decreased, which predicts poor prognosis of CSCC. Moreover, only the expression of TET2 was decreased in CSCC [14]. In contrast, Bhat et al. found that the 5-mC and 5-hmC levels were both significantly reduced in squamous cell carcinoma, but receiver operating characteristic curve analysis showed a significant difference in 5-mC and 5-hmC between normal and squamous cell carcinoma tissues. They also tested the promoter methylation of 33 genes; only PROX1, NNAT, ARHGAP6, HAND2, NKX2-2, PCDH10, DAPK1, RAB6C, and PITX2 could effectively tell the difference among the various stages of tumor with high sensitivity and specificity [15]. Expressions of 5-hmC and 5-mC in cervical cancer need further demonstrations, and these related results may serve as useful biomarkers for the early detection and accurate management of cervical cancer.

Although 5-hmC was studied little in cervical cancer, it is a noticeable part in other cancers; scientists have been making further studies for deeper mechanisms of 5-hmC as well.

It is demonstrated that TET1 and TET3 catalyze the conversion from 5-mC to 5-hmC by activating the TNFα-p38-MAPK signaling axis and inducing tumor malignancy and poor prognosis in breast cancer patients [16]. In prostate cancer, the androgen receptor decreases the expression of miR-29b which targets both TET2 and 5-hmC; 5-hmC represses FOXA1 activity, while its reduction activates the mTOR pathway and AR of prostate cancer [17]. In DLD1 cells, knockdown of TET1 will promote cancer cell growth, migration, invasion, and even epithelial-mesenchymal transition (EMT) which can also reduce UTX-1 but increase the EZH2 expression which can cause a loss of H3K27 methylation at the epithelial gene E-cadherin promoter [18]. In contrast, the levels of TETs are similar in colorectal tumor tissue and normal tissues. TET2 targets promoters marked by 5-hmC in normal tissue and turns it to colorectal cancer tissue [19].

3. DNA Methylation in Cervical Cancer and CIN

In cervical lesions, aberrant DNA methylation includes hypomethylation and hypermethylation. In cervical cancer and high-grade cervical intraepithelial neoplasia, most genes are hypermethylated; only three promoter regions are hypomethylated (Table 1).

3.1. Gene Hypomethylation in Cervical Cancer and CIN. Hypomethylation often occurs in the promoter region of genes, regardless if the gene is for a protein or RNA. The STK31 gene targets at oncogene E7 of HPV16. Its promoter/exon 1 is hypomethylated in HPV16/18-positive cervical cell lines, which induces an integration of HPV16E7/E6 [20]. The COL17A1 promoter is also hypomethylated in cervical cancer, and it precisely predicts both the increased invasive nature and patient outcome [21]. In CIN tissues, the rDNA promoter region reveals significant hypomethylation at cytosines in the context of CpG dinucleotides, which can result in an increase in rRNA synthesis in the development of human cervical cancer [22].

3.2. Gene Hypermethylation in Cervical Cancer and CIN

3.2.1. Genome-Wide Studies of Aberrant Gene Methylation. There are some genome-wide studies of aberrant gene expression and methylation profiles which reveal susceptibility genes and underlying mechanisms of cervical cancer. In one study, a total of 1357 DEGs as well as 666 cervical cancer-

Table 1: DNA methylation of CIN or cervical cancer in recent studies.

Name of gene	Methylation status	Methylation-variable position	Function/relevant pathway	Reference	Notes
STK31	Hypomethylation	Promoter/exon 1	HPV oncogene-E6/E7	[20]	CIN III and CCA
COL17A1	Hypomethylation	Promoter	Collagen XVII	[21]	CCA
Ribosomal DNA	Hypomethylation	Promoter	rRNA synthesis	[22]	CIN II-III, CCA
EDN3 and EDNRB	Hypermethylation	Promoter	MAPK signal pathway MITF-Wnt/β-catenin signal pathway	[20, 23]	
VIM	Hypermethylation	Promoter	Epithelial-mesenchymal transition and aggressiveness	[24]	Ib1 and IIa stages of CCA
AJAP1 and SOX17	Hypermethylation	Promoter	Wnt signal pathway	[25]	
SFRP1 and SFRP4	Hypermethylation	Promoter	Wnt/β-catenin signal pathway	[25]	
CDKN2A	Hypermethylation	Downstream region	p16(INK4A)/p14(ARF)	[26]	CIN and CCA
IFN-γ	Hypermethylation	Promoter	IFN-γ-cancer immunoediting	[27]	CIN II-III and CCA
SALL3	Hypermethylation	Promoter	hrHPV-induced immortalization and malignant transformation	[28]	HPV-infected
EPB41L3	Hypermethylation	Promoter	DAL-1 protein	[29]	CIN II-III
CADM1/MAL	Hypermethylation	Unmentioned	Lesion-specific	[30]	CIN II-III and CCA
PAX1	Hypermethylation	Promoter	Unclear yet	[32]	CIN and CCA
DAPK1	Hypermethylation	Promoter	Epithelial-mesenchymal transition	[34]	CIN III and CCA
Keap1	Hypermethylation	Promoter	NRF2	[35]	CCA
GPX3	Hypermethylation	Promoter	Repair oxidative damages and lymph node metastasis	[36]	CCA
LDOC1	Hypermethylation	Promoter	Nuclear transcription factor	[37]	CCA
RASSF	Hypermethylation	Promoter	Ras protein	[38, 42]	CCA or plasma of CCA
DOC2B	Hypermethylation	Promoter	AKT1 and ERK1/2 signal pathway	[40]	CIN and CCA
MEG3	Hypermethylation	Promoter	Proliferation and apoptosis	[41]	Plasma of CIN III and CCA

(CC-) related methylation sites were screened out and 26 DEGs with 35 CC-related methylation sites were identified; ACOX3, CYP39A1, and DPYS are potential risk markers in CC, which were significantly enriched in 25 subpathways of 6 major pathways. EDN3 and EDNRB might play important roles in the molecular mechanism of CC [23]. In another study, 32 genes that might be associated with prognosis in the stages between Ib1 and IIa cervical cancer are profiled, among which the VIM gene is frequently methylated in CSCC and VIM methylation might predict a favorable prognosis [24]. The 14 hypermethylated genes, including ADRA1D, AJAP1, COL6A2, EDN3, EPO, HS3ST2, MAGI2, POU4F3, PTGDR, SOX8, SOX17, ST6GAL2, SYT9, and ZNF614, are implicated in β-catenin signaling in cervical carcinogenesis [25].

3.2.2. Gene Hypermethylation Found in Cervical Cancer/CIN Tissue Cell Lines and Patients' Plasmas. Gene hypermethylation is found in CIN cervical cancer tissues, cervical cancer cells, and even cervical cancer patients' plasmas. The methylation rates of IFN-γ, FHIT, MGMT, CDKN2A, SALL3, and gene promoters were significantly higher in cervical cancer tissues than those in CIN and normal cervical tissues, which are related to the progression of cervical oncogenesis. CDKN2A methylation may lead to the development of malignant disease by increased p16(INK4A)/p14(ARF) expression [26–28]. LINE-1, HS3ST2, CCNA1, EPB41L3, EDNRB, LMX1, and DPYS were hypermethylated in cervical cancer tissues, CIN III and CIN II, versus normal tissues and CIN I, of which EPB41L3 seems to be the best marker. CADM1 is regulated by p53, and CADM1/MAL is hypermethylated in the HPV16/18-infected cell lines. The methylation status in cervical scrapes appears to represent the worst underlying lesion, particularly CIN III and cervical cancer. Results imply that hypermethylation of these genes may be highly associated with the development of cervical cancer [29–31]. Specific hypermethylated genes serve as the early prevention and prognostic prediction for cervical cancer. The different methylation statuses of all three genes PAX1, SOX1, and ZNF582 showed reasonable concordance in normal control samples as well as CIN I, CIN II, CIN III, and SCC samples [32, 33]. The promoter methylation statuses

of DAPK1, MGMT, and RARB were positively correlated with the cervical disease grades, respectively. DAPK1 combined with the other two showed a significantly positive correlation with cervical disease grade as well [34]. The promoter hypermethylation of Keap1 significantly increased nuclear NRF2 expression in cervical cancer tissues, which is a marker of poor prognosis in patients with cervical cancer [35]. And the promoter of GPX3 is significantly downregulated due to its promoter hypermethylation in cervical cancer tissues; at the same time, GPX3 expression plays a role in the development of cervical squamous cell carcinoma and is significantly related to lymph node metastasis and prognosis in cervical cancer patients [36]. Promoter methylation and the loss of LDOC1 expression are frequent events in cervical cancer and could be potential molecular markers in cervical cancer [37]. Hypermethylation of RASSF2A and TSLC1 downregulating the expression of RASSF2A and TSLC1 was detected, which predicts a greater risk of progressing towards invasive cervical cancer [38, 39]. Hypermethylation of DOC2B promotes colony formation and cell proliferation, induces cell cycle arrest, and represses cell migration and invasion deeply; the promoter region of the DOC2B gene inhibiting AKT1 and ERK1/2 signaling is hypermethylated in premalignant and malignant cervical tissues and cervical cancer cell lines [40]. Those gene promoter methylations may be correlated with clinical stage and tumor grade and play a crucial role in cervical cancer progression.

The level of MEG3 methylation is significantly higher in cervical cancer tissues and patients' plasmas than in adjacent normal tissues and plasmas of healthy participants, respectively [41]. Promoter hypermethylation of some other genes like MYOD1, CALCA, hTERT, and RASSF1A can also be detected in serum samples of cervical cancer patients and are related to lymph node metastasis and FIGO stage [42, 43]. In conclusion, the present studies clearly showed that MEG3, MYOD1, CALCA, hTERT, and RASSF1A methylation in plasma can serve as diagnostic and prognostic biomarkers for cervical cancer patients, providing useful information for clinical management.

3.2.3. Gene Hypermethylation Found in Different Ethnicities. The hypermethylation status of genes in cervical cancer patients is associated with different countries. In the North Indian population, methylation of the p16 gene promoter which induces loss of tumor-suppressing activity and promotes the development of cervical cancer is observed significantly in FIGO stage III [44, 45]. Meanwhile, correlated with clinical parameters, promoter hypermethylation and expression loss of PARK-2, RARβ, and FHIT are significantly higher in cervical cancer than in CINs and normal tissues, resulting in a significant association with tumor stage and histological grade [46, 47]. In Uighur women, increased methylation was detected at 13 CpG sites, and a high methylation level was associated with the risk of CIN2$^+$; the strongest related site was 6650 [48]. The methylation level of the ERp57 gene promoter is higher in CSCC than in CIN, and normal tissues in Uighur women. Hypermethylation occurs only in certain CpG islands and sites, such as CpG1, CpG5, and CpG7, and it differs significantly in CSCC, CIN, or control groups [49]. In Uygur and Han, aberrant methylation of TFPI2 is present in a higher proportion of invasive cervical carcinoma (ICC) clinical samples [50]. Apart from that, hypermethylation is related to different age groups as well. Hypermethylation of the CDKN2A gene promoter is a frequent epigenetic change in younger patients with cervical carcinoma and implies a significant epigenetic role in tumor development in this age group [51].

3.3. The Relationship between HPV and Aberrant DNA Methylation in Cervical Cancer/CIN (Figure 1). On the one hand, HPV and aberrant host gene methylation contribute to CIN and CCA, respectively, methylation of HPV can prevent itself from cleaning to keep the persistent infection state, and the host methylation level can also reflect the level of HPV-associated CCA. On the other hand, the HPV genome and host act on each other by methylated regulation. HPV takes part in the methylation of host genomes such as FAM19A4 and LHX1; the methylation of HPV itself can also work with the methylation of PAX1 and SOX1 in the host to enhance transcription, both of which induce bad outcomes of the host cervix.

3.3.1. Methylation Status of HPV Genome in Cervical Cancer/CIN. HPV genome epigenetic alterations play an important role in cervical cancer progression. Among them, methylation of CpG sites in the L1, L2, and LCR regions in different types of HPV is studied most, and several deep relationships between the methylation of those regions and cervical cancer/CIN have been found out. HPV L1 gene methylation was the risk factor to cervical and elevated levels. HPV16 L1 methylation affects E6/E7 mRNA levels and can detect high-grade cervical lesions (CIN2$^+$) [52, 53]. It also prolongs the cleaning of HPV infection and increases the risk of HPV cleaning failure in premalignant cervical lesion patients [54]. Besides, a panel of 12 HPV16 CpG sites which are methylated in L1, L2, and E5 can work as an informative biomarker for the triage of women positive for HPV16 infection and is correlated with the severity of cervical neoplasia, even cervical cancer [55]. But some other evidence shows HPV16 L1/L2 DNA methylation weakly associated with cervical disease grade in young women, which means HPV DNA methylation as a biomarker must take into account women's age [56]. The L1 and L2 regions of other types of HPVs are methylated in cervical cancer/CIN. Aberrant methylation of CpG sites in the L1 and L2 regions of HPV18 and other high-risk HPV types including HPV31, HPV33, HPV45, HPV52, HPV51, and HPV58 relates with the progression from early-stage CINs and may be considered as a biomarker of the progression of cervical neoplasia [57, 58]. Another research shows that the methylation of L1 in HPV16, HPV18, and HPV52 does not only play an important role in cervical cancer alone. The methylation of most HPV types except HPV52 also works together with the methylation of host genes including PAX1 and SOX1, which leads to a more significant result of cervical cancer/CIN [59]. Combining HPV methylation with PAX1 methylation improves the clustering for CIN2$^+$ and methylated CpG sites in HPV31 LCR,

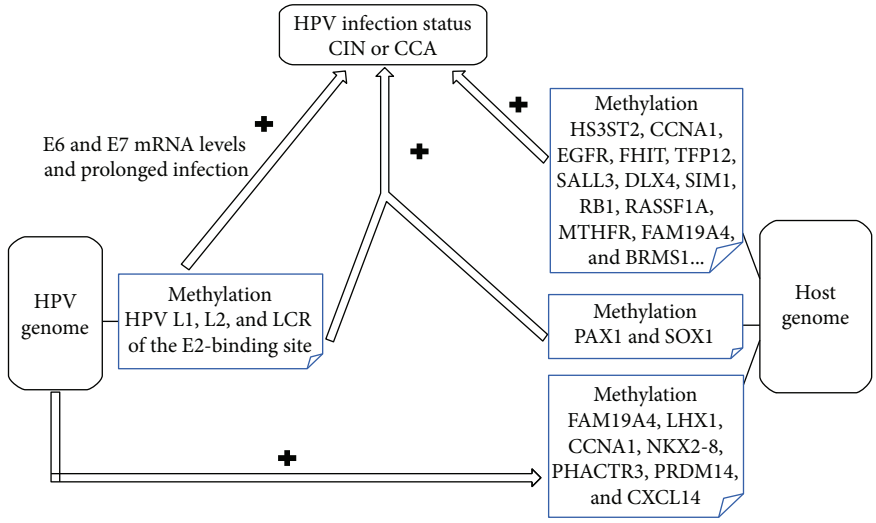


Figure 1: Methylation regulation between HPV and host genome in CIN or cervical cancer.

including position 7479 and/or 7485, which is the promoter distal E2-binding site, suggesting a potential regulatory mechanism for papillomavirus transcription [60].

3.3.2. The Interaction between HPV and Aberrant Methylation of Other Genes. In cervical cancer/CIN, methylated HPVs and other genes correlate with each other and serve as diagnostic and prognostic biomarkers for cervical cancer. Methylated HS3ST2, CCNA1, EGFR promoter, FHIT, TFPI2, CpG6, and CpG15 sites were associated with HPV16 infection in the progression of cervical cancer [61–63]. The results indicate that methylated genes may play important roles and be effective targets for the prevention and treatment of cervical cancer. HPV infection is also associated with hypermethylation of the promoter region of SALL3, DLX4, and SIM1 genes, which should be a significant progression marker for HPV infection in cervical cancer [64]. The methylation-mediated gene silencing of PRDM14, a regulator of NOXA and PUMA-mediated apoptosis, becomes an important factor in the development of hr-HPV-ICC (invasive cervical carcinoma) and offers a novel therapeutic target for HPV-induced cervical cancers [65]. In addition, that FAM19A4 promoter methylation even altered DNA methylation seems to be associated with HPV infection and high-risk types of HPV-induced carcinogenesis in the uterine cervix, $CIN3^+$, and may increase with disease progression [66]. Moreover, some of the methylated genes have been demonstrated as attractive markers for hr-HPV-positive women, with a high reassurance for the detection of cervical carcinoma and advanced CIN2/3 lesions, such as EPB41L3 and FAM19A4 [29, 67].

Not only can gene methylation affect HPV infection, but HPV also results in other genes' aberrant methylation. HPV can result in novel DNA methylation events, including FAM19A4, LHX1, NKX2–8, PHACTR3, and PRDM14 genes in cervical carcinogenesis [68]. Numerous pieces of evidence suggest that HPV16 E7 oncoprotein mediates DNA hypermethylation in the CCNA1 and CXCL14 promoter and suppresses gene expression. The data also shows that E7 induces CCNA1 methylation by forming a complex with Dnmt1 at the CCNA1 promoter [69, 70]. The potential carcinogenic mechanism of HPVs, including influencing the DNA methylation pathway to affect DNA methylation and mRNA expression levels of those genes, can be utilized not only as a biomarker for early detection, disease progression, diagnosis, and prognosis of cervical cancer but also to design effective therapeutic strategies.

3.3.3. Identification of Cervical Cancer by HPV and Gene Methylation Test. Currently, the HPV DNA test is one of the most vital tools to identify the risk of cervical cancer/CIN. Some studies show that detecting the methylation status of a few kinds of genes can also give evidence for diagnosing $CIN2^+$ or help the HPV test to improve the specificity and sensitivity in the detection of cervical cancer/CIN.

In an independent cohort test, the methylated PCDHA4 and PCDHA13 test is equally sensitive but more specific than the human papillomavirus (HPV) test in the diagnosis of $CIN2^+$ [71]. Combining the triage by MAL/miR-124-2 methylation analysis with threshold-80 and HPV16/18 genotyping can reach higher $CIN3^+$ sensitivity and identify women at the highest risk of cervical (pre)cancer [72, 73]. Combining parallel testing of PAX1, DAPK1, RARB, WIF1, and SLIT2 DNA methylation and HPV DNA increases specificity to identify cervical cancer and achieves better precision than single HPV DNA testing does [74–76]. Above all, methylation of some genes has a prospect to be an auxiliary biomarker for cervical cancer screening.

Now, cervical (pre)cancer is usually classified by histologic pathology, but cervical conization will lead to a high risk of premature delivery and abortion for patients. A quantitative measurement of HPV-type 16 L1/L2 DNA methylation has demonstrated its correlation with cervical disease grade. The best separation between normal and dyskaryotic samples is achieved by assessment of the L1/L2 CpGs at nucleotide positions 5600 and 5609 [77]. At the same time, CCNA1 promoter methylation serves as a potential marker for distinguishing between histologic LSIL (low-

grade squamous intraepithelial lesion)/negative and HSIL (high-grade squamous intraepithelial lesion)/positive [78].

4. Methylation-Related Regulations on Other Levels in Cervical Cancer/CIN

According to the central dogma of molecular biology, epigenetic modifications also occur in the process of genetic information expression, such as the DNA level mentioned above, RNA level including mRNA and ncRNA (noncoding RNA, miR-RNA, and lnc-RNA are included), and protein level involving common protein or histone.

4.1. Pervasive Gene Expression Adjustment of Cervical Cancer/CIN at RNA Level

4.1.1. m6A Induces Methylated Regulation in mRNA. mRNAs carry genetic information by encoding polypeptides or proteins; that m6A methylates mRNA is widespread in eukaryotic cells. N^6-Methyladenosine (m6A), which is an abundant and conservative RNA modification, is involved in a series of biological processes such as differentiation, metabolism, immune tolerance, and neuronal signaling by impacting on mRNA splicing, export, localization, translation, and stability [79]. As the UV cross-linking immunoprecipitation and single-nucleotide resolution show, the distribution of m6A is not random in mature transcripts but concentrates around the $3'$ untranslated regions (UTRs), stop codons, and is within internal long exons [80]. The reversibility of m6A is accomplished by the orchestrated action of a battery of enzymes or proteins: as readers, proteins YTHDF and hnRNP recognize m6A-containing mRNA; as writers, METTL3, METTL14, and the WTAP complex support RNA methylation; and as erasers, FTO and ALKBH5 prop up RNA demethylation [79, 81].

As investigations about relationships between various tumors and m6A deepen, some crucial targets of tumor biological processes are found. Theories of m6A are elucidated increasingly in GSC, AML, HCC, BRC, and so on. Inhibition of FTO not only suppresses growth and self-renewal but also prolongs the lifespan of grafted mice and restrains tumor progression additionally compared with overexpression of METTL3 [82]. However, research of m6A about cervical cancer is rarely covered.

4.1.2. ncRNAs Play an Important Role in Methylation Regulation. As genomics analysis shows, there are numerous transcripts being generated in the human body; just 1–2% transcripts own the function of encoding polypeptides or proteins, and the remaining 98% noncoding products play vital roles in many biological process, including proliferation, differentiation, and apoptosis [83]. They are all hot topics in the field of apparent genetics.

With NGS and qRT-PCR applied, levels of various miR-RNAs in cervical cancer are evaluated: most miR-RNAs are downregulated and relevant downstream signal pathways or target genes and proteins are reported, such as SOX2 of miR-145, TCF of miR-212, Bcl-2 of miR-187, and NF-κb of miR-429, performing significant relationships with FIGO stage, lymph node metastasis and prognosis of patients in clinic, and colony formation, tumor size, proliferation, differentiation, apoptosis, and invasion on the lab research *in vivo* and *in vitro* [84–87]. There are still some miR-RNAs upregulated in CCA, such as miR-9 [88]. However, Zhang et al. reported that miR-9 is downregulated in cervical cancer on account of hypermethylation of miR-9 precursor promoters, which weakens the inhibiting effect on activity of the IL-6/Jak/STAT3 pathway [89]. These different outcomes may be induced by the potentially different methylation status in the objects.

Impacts of miR-RNA on the progress from HPV infection to cervical cancer are nonnegligible. Morel et al. reported that miR-375 could destabilize HPV16 early viral mRNA and contribute to the regulation of E6/E7 expression, which indicated the role of miR-RNA in high-risk HPV-associated carcinogenesis [90]. Yeung et al. revealed that HPV16 E6 takes part in epigenetic regulation of host gene-associated cervical cancer development; HPV16 E6 methylates the promoter region of the host gene of miR-23b, C9, or f3; and downregulated miR-23b enhances c-MET pathway-induced apoptosis of cervical cancer cells [91].

lnc-RNA interacting with miR-RNA regulates cervical cancer biological activity. lnc-RNA MEG3 is negatively relevant with FIGO stages, tumor size, lymphatic metastasis, and HR-HPV infection, and downexpressed MEG3 in cervical cancer reduces the inhibition effect on miR-21-5p expression, which leads to less apoptosis and more proliferation of cancer cells [92]. There are some cases about interactions between lnc-RNA, miR-RNA, and histone. For example, Zhang et al. explained the regulatory mechanism of lnc-RNA PVT1, which is overexpressed in cervical cancer: PVT1 binds with EZH2 directly to activate EZH2 to increase the histone H3K27 trimethylation level of the miR-200b promoter so that downexpressed miR-200b enhanced proliferation, cycle progression, and migration [93].

4.2. Methylation Research Related to Cervical Cancer Therapy Applications.

Many mechanisms of methylation-associated regulations become the targets of therapy in the fields of chemotherapy and radiotherapy. It has been shown that cisplatin as well as 5-azacytidine touch off cytotoxic and growth inhibitory effects *in vitro* by demethylating the promoters of ESR1, BRCA1, RASSF1A, MLH1, MYOD1, hTERT, and DAPK1 to reexpress these tumor-associated genes [94]. Narayan et al. identified inactivation of decoy receptors TNFRSF10C and TNFRSF10D as major target genes at the 8p MDR region. On the one hand, the promoter hypermethylation of TNFRSF10C was an early event in cervical tumorigenesis; on the other hand, inactivation of decoy receptors induced extrinsic-apoptotic-pathway-dependent cell death in the cooperation of TRAIL and cisplatin in the presence of DNA-damaging drugs [95]. These covers above demonstrate that methylation-associated regulation offers an idea for developing new therapy targets.

Besides, methylation modifications impact on the sensitivity of chemotherapy and radiotherapy. Radiosensitization occurs when SiHa cells accept the therapeutic regimen combining DNA methylation inhibitor hydralazine with a

histone deacetylase inhibitor valproic acid; unexpectedly, the efficacy of cisplatin chemoradiation was increased under the use of two epigenetic drugs [96]. Furthermore, epigenetic modifications also participate in therapeutic resistance. A univariate and hierarchical cluster analysis uncovered that standard chemoradiation resistance contacts closely with lower ESR1 transcript levels as well as unmethylated ESR1, unmethylated MYOD1, and methylated hTERT promoter [97]. In an article about the suppressor of cytokine signaling (SOCS) family and cervical cancer, Kim et al. found that DNA methylation contributed to SOCS1 downregulation, and histone deacetylation may be the mechanism of SOCS1 and SOCS3 regulation; in the meantime, ectopic expression of SOCS1 or SOCS3 could induce radioresistance of HeLa cells [98]. Similarly, a research about type-I ribosome-inactivating protein trichosanthin reported that Smac demethylation was subdued and Twist was upregulated in TCS-resistance cervical cells, which indicated that aberrant mitochondrial methylation may be partly the reason for drug resistance [99].

5. Conclusion

Cervical cancer is likely to be the first tumor which can receive idealized prevention and cure depending on the vital status of HPV in the pathological process. In spite of the astounding advances of screening plans and HPV vaccines, cervical cancer is still threatening the physical and psychological health of females with the absence of effective treatment, surveillance indexes, and fundamentally unclear molecular mechanisms. Over the past decades, methylation modification has been identified as having a significant role in the generation of cervical cancer. With the development of methylation-detecting techniques, there may be more convenient choices to explore it, not limited to cells and tissues, but techniques like liquid biopsy to advanced clinical transformation. We believe that it can not only enrich the markers for the early diagnosis and prognosis evaluation with other biomarkers to improve sensitivity and specificity in the clinic but also provide targets for exploiting new drugs as well as modifying the sensitivity in radiotherapy and chemotherapy for cervical cancer. However, there are still some items to be investigated deeply. Firstly, studies about the relationships between 5-hmC or 6mA and cervical cancer are rare, especially the aspect of HPV infection. Secondly, some researches find that methylation modification does not act itself but correlates with other epigenetic forms such as ncRNA regulation and histone decoration; therefore, the effective application of methylation relies on the simplification of key points. Additionally, from HPV infection via CIN to cervical cancer, relevant researches of the dynamic pathogenesis are inconsecutive. Moreover, it is recognized that methylation regulation is reversible, which is the unique advantage of therapy; only by enabling the reversibility controllable can we make full use of the characteristic. In conclusion, 5-hmC of hydroxymethylation, 5-mC of methylation, and 6-mA of RNA methylation are typical mechanisms of the methylation modification in gene expression; some ncRNA and histone regulations are involved in methylation in the meantime, and these investigations have profound instructive significance in the process of overcoming cervical cancer.

Disclosure

Chunyang Feng and Junxue Dong should be considered as co-first authors.

Conflicts of Interest

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

Authors' Contributions

Chunyang Feng and Junxue Dong contributed to this work equally.

Acknowledgments

This work was supported by grants from the National Key R&D Program of China (2016YFC1302901), Jilin Province Science and Technology funds (20180201032YY, 20150204007YY, and 20140204022YY), Jilin Province Development and Reform Commission funds (2014G073 and 2016C046-2), and Education Department of Jilin Province funds (JJKH20170804KJ).

References

[1] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal, "Global cancer statistics, 2012," *CA: a Cancer Journal for Clinicians*, vol. 65, no. 2, pp. 87–108, 2015.

[2] Y. B. Zhao, J. H. Wang, X. X. Chen, Y. Z. Wu, and Q. Wu, "Values of three different preoperative regimens in comprehensive treatment for young patients with stage Ib2 cervical cancer," *Asian Pacific Journal of Cancer Prevention*, vol. 13, no. 4, pp. 1487–1489, 2012.

[3] W. Yang, L. Hong, X. Xu, Q. Wang, J. Huang, and L. Jiang, "LncRNA GAS5 suppresses the tumorigenesis of cervical cancer by downregulating miR-196a and miR-205," *Tumor Biology*, vol. 39, no. 7, 2017.

[4] K. Williams, J. Christensen, and K. Helin, "DNA methylation: TET proteins—guardians of CpG islands?," *EMBO Reports*, vol. 13, no. 1, pp. 28–35, 2011.

[5] B. F. Yuan, "5-Methylcytosine and its derivatives," *Advances in Clinical Chemistry*, vol. 67, pp. 151–187, 2014.

[6] S. Ke, A. Pandya-Jones, Y. Saito et al., "m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover," *Genes & Development*, vol. 31, no. 10, pp. 990–1006, 2017.

[7] O. L. Kantidze and S. V. Razin, "5-Hydroxymethylcytosine in DNA repair: a new player or a red herring?," *Cell Cycle*, vol. 16, no. 16, pp. 1499–1501, 2017.

[8] W. Li and M. Liu, "Distribution of 5-hydroxymethylcytosine in different human tissues," *Journal of Nucleic Acids*, vol. 2011, Article ID 870726, 5 pages, 2011.

[9] S. R. M. Kinney and S. Pradhan, "Ten eleven translocation enzymes and 5-hydroxymethylation in mammalian

development and cancer," *Advances in Experimental Medicine and Biology*, vol. 754, pp. 57–79, 2013.

[10] J. An, A. Rao, and M. Ko, "TET family dioxygenases and DNA demethylation in stem cells and cancers," *Experimental & Molecular Medicine*, vol. 49, no. 4, p. e323, 2017.

[11] C. J. Mariani, A. Vasanthakumar, J. Madzo et al., "TET1-mediated hydroxymethylation facilitates hypoxic gene induction in neuroblastoma," *Cell Reports*, vol. 7, no. 5, pp. 1343–1352, 2014.

[12] G. Ficz and J. G. Gribben, "Loss of 5-hydroxymethylcytosine in cancer: cause or consequence?," *Genomics*, vol. 104, no. 5, pp. 352–357, 2014.

[13] L. I. Kroeze, B. A. van der Reijden, and J. H. Jansen, "5-Hydroxymethylcytosine: an epigenetic mark frequently deregulated in cancer," *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, vol. 1855, no. 2, pp. 144–154, 2015.

[14] L. Y. Zhang, C. S. Han, P. L. Li, and X. C. Zhang, "5-Hydroxymethylcytosine expression is associated with poor survival in cervical squamous cell carcinoma," *Japanese Journal of Clinical Oncology*, vol. 46, no. 5, pp. 427–434, 2016.

[15] S. Bhat, S. P. Kabekkodu, V. K. Varghese et al., "Aberrant gene-specific DNA methylation signature analysis in cervical cancer," *Tumour Biology*, vol. 39, no. 3, pp. 1–16, 2017.

[16] M. Z. Wu, S. F. Chen, S. Nieh et al., "Hypoxia drives breast tumor malignancy through a TET–TNFα–p38–MAPK signaling axis," *Cancer Research*, vol. 75, no. 18, pp. 3912–3924, 2015.

[17] K. Takayama, A. Misawa, T. Suzuki et al., "Tet2 repression by androgen hormone regulates global hydroxymethylation status and prostate cancer progression," *Nature Communications*, vol. 6, no. 1, p. 8219, 2015.

[18] Z. Zhou, H.-S. Zhang, Y. Liu et al., "Loss of TET1 facilitates DLD1 colon cancer cell migration via H3K27me3-mediated down-regulation of E-cadherin," *Journal of Cellular Physiology*, vol. 233, no. 2, pp. 1359–1369, 2018.

[19] S. Uribe-Lewis, R. Stark, T. Carroll et al., "5-Hydroxymethylcytosine marks promoters in colon that resist DNA hypermethylation in cancer," *Genome Biology*, vol. 16, no. 1, p. 69, 2015.

[20] F. F. Yin, N. Wang, X. N. Bi et al., "Serine/threonine kinases 31(STK31) may be a novel cellular target gene for the HPV16 oncogene E7 with potential as a DNA hypomethylation biomarker in cervical cancer," *Virology Journal*, vol. 13, no. 1, p. 60, 2016.

[21] P. U. Thangavelu, T. Krenács, E. Dray, and P. H. G. Duijf, "In epithelial cancers, aberrant COL17A1 promoter methylation predicts its misexpression and increased invasion," *Clinical Epigenetics*, vol. 8, no. 1, p. 120, 2016.

[22] H. Zhou, Y. Wang, Q. Lv et al., "Overexpression of ribosomal RNA in the development of human cervical cancer is associated with rDNA promoter hypomethylation," *PLoS One*, vol. 11, no. 10, article e0163340, 2016.

[23] H. Lin, Y. Ma, Y. Wei, and H. Shang, "Genome-wide analysis of aberrant gene expression and methylation profiles reveals susceptibility genes and underlying mechanism of cervical cancer," *European Journal of Obstetrics, Gynecology, and Reproductive Biology*, vol. 207, pp. 147–152, 2016.

[24] M. K. Lee, E. M. Jeong, J. H. Kim, S. B. Rho, and E. J. Lee, "Aberrant methylation of the vim promoter in uterine cervical squamous cell carcinoma," *Oncology*, vol. 86, no. 5-6, pp. 359–368, 2014.

[25] Y. C. Chen, R. L. Huang, Y. K. Huang et al., "Methylomics analysis identifies epigenetically silenced genes and implies an activation of β-catenin signaling in cervical cancer," *International Journal of Cancer*, vol. 135, no. 1, pp. 117–127, 2014.

[26] N. A. Wijetunga, T. J. Belbin, R. D. Burk et al., "Novel epigenetic changes in CDKN2A are associated with progression of cervical intraepithelial neoplasia," *Gynecologic Oncology*, vol. 142, no. 3, pp. 566–573, 2016.

[27] D. Ma, C. Jiang, X. Hu et al., "Methylation patterns of the IFN-γ gene in cervical cancer tissues," *Scientific Reports*, vol. 4, no. 1, 2014.

[28] X. Wei, S. Zhang, D. Cao et al., "Aberrant hypermethylation of SALL3 with HPV involvement contributes to the carcinogenesis of cervical cancer," *PLoS One*, vol. 10, no. 12, article e0145700, 2015.

[29] N. Vasiljević, D. Scibior-Bentkowska, A. R. Brentnall, J. Cuzick, and A. T. Lorincz, "Credentialing of DNA methylation assays for human genes as diagnostic biomarkers of cervical intraepithelial neoplasia in high-risk HPV positive women," *Gynecologic Oncology*, vol. 132, no. 3, pp. 709–714, 2014.

[30] R. van Baars, J. van der Marel, P. J. F. Snijders et al., "CADM1 and MAL methylation status in cervical scrapes is representative of the most severe underlying lesion in women with multiple cervical biopsies," *International Journal of Cancer*, vol. 138, no. 2, pp. 463–471, 2016.

[31] H. J. Woo, S. J. Kim, K.-J. Song et al., "Hypermethylation of the tumor-suppressor cell adhesion molecule 1 in human papillomavirus-transformed cervical carcinoma cells," *International Journal of Oncology*, vol. 46, no. 6, pp. 2656–2662, 2015.

[32] J. Xu, L. Xu, B. Yang, L. Wang, X. Lin, and H. Tu, "Assessing methylation status of PAX1 in cervical scrapings, as a novel diagnostic and predictive biomarker, was closely related to screen cervical cancer," *International Journal of Clinical and Experimental Pathology*, vol. 8, no. 2, pp. 1674–1681, 2015.

[33] C. C. Chang, R. L. Huang, Y. P. Liao et al., "Concordance analysis of methylation biomarkers detection in self-collected and physician-collected samples in cervical neoplasm," *BMC Cancer*, vol. 15, no. 1, p. 418, 2015.

[34] Y. Sun, S. Li, K. Shen, S. Ye, D. Cao, and J. Yang, "DAPK1, MGMT and RARB promoter methylation as biomarkers for high-grade cervical lesions," *International Journal of Clinical and Experimental Pathology*, vol. 8, no. 11, pp. 14939–14945, 2015.

[35] J. Q. Ma, H. Tuersun, S. J. Jiao, J. H. Zheng, J. B. xiao, and A. Hasim, "Functional role of NRF2 in cervical carcinogenesis," *PLoS One*, vol. 10, no. 8, article e0133876, 2015.

[36] X. Zhang, Z. Zheng, S. Yingji et al., "Downregulation of glutathione peroxidase 3 is associated with lymph node metastasis and prognosis in cervical cancer," *Oncology Reports*, vol. 31, no. 6, pp. 2587–2592, 2014.

[37] M.-L. Buchholtz, J. Jückstock, E. Weber, I. Mylonas, D. Dian, and A. Brüning, "Loss of LDOC1 expression by promoter methylation in cervical cancer cells," *Cancer Investigation*, vol. 31, no. 9, pp. 571–577, 2013.

[38] Y. Lin, M. Cui, T. Xu, W. Yu, and L. Zhang, "Silencing of cyclooxygenase-2 inhibits the growth, invasion and migration of ovarian cancer cells," *Molecular Medicine Reports*, vol. 9, no. 6, pp. 2499–2504, 2014.

[39] X. Zhao, Y. Cui, Y. Li et al., "Significance of TSLC1 gene methylation and TSLC1 protein expression in the progression of cervical lesions," *Zhonghua Zhong Liu Za Zhi*, vol. 37, no. 5, pp. 356–360, 2015.

[40] S. P. Kabekkodu, S. Bhat, R. Radhakrishnan et al., "DNA promoter methylation-dependent transcription of the double C2-like domain β (DOC2B) gene regulates tumor growth in human cervical cancer," *The Journal of Biological Chemistry*, vol. 289, no. 15, pp. 10637–10649, 2014.

[41] J. Zhang, T. Yao, Z. Lin, and Y. Gao, "Aberrant methylation of MEG3 functions as a potential plasma-based biomarker for cervical cancer," *Scientific Reports*, vol. 7, no. 1, p. 6271, 2017.

[42] H. Wang, S. Y. Pan, Z. R. Pang et al., "Quantitative detection of APC/RASSF1A promoter methylation in the plasma of patients with cervical diseases," *Zhonghua Fu Chan Ke Za Zhi*, vol. 48, no. 12, pp. 929–934, 2013.

[43] A. K. Jha, V. Sharma, M. Nikbakht et al., "A comparative analysis of methylation status of tumor suppressor genes in paired biopsy and serum samples from cervical cancer patients among North Indian population," *Genetika*, vol. 52, no. 2, pp. 255–259, 2016.

[44] F. L. Wang, Y. Yang, Z. Y. Liu, Y. Qin, and T. Jin, "Correlation between methylation of the p16 promoter and cervical cancer incidence," *European Review for Medical and Pharmacological Sciences*, vol. 2017, no. 21, 6 pages, 2017.

[45] A. Gupta, M. K. Ahmad, A. A. Mahndi, R. Singh, and Y. Pradeep, "Promoter methylation and relative mRNA expression of the p16 gene in cervical cancer in North Indians," *Asian Pacific Journal of Cancer Prevention*, vol. 17, no. 8, pp. 4149–4154, 2016.

[46] R. Shu, J. He, C. Wu, and J. Gao, "The association between-RARβ and FHIT promoter methylation and the carcinogenesis of patients with cervical carcinoma: a meta-analysis," *Tumour Biology*, vol. 39, no. 6, 10 pages, 2017.

[47] A. Naseem, Z. I. Bhat, P. Kalaiarasan, B. Kumar, G. Gandhi, and M. M. A. Rizvi, "Genetic and epigenetic alterations affecting PARK-2 expression in cervical neoplasm among North Indian patients," *Tumour Biology*, vol. 39, no. 6, 2017.

[48] M. Niyazi, S. Sui, K. Zhu, L. Wang, Z. Jiao, and P. Lu, "Correlation between methylation of human papillomavirus-16 L1 gene and cervical carcinoma in Uyghur women," *Gynecologic and Obstetric Investigation*, vol. 82, no. 1, pp. 22–29, 2017.

[49] M. Abdula, G. Abudulajiang, R. Amiduo, A. Abudala, and A. Hasim, "Association of promoter methylation of ERp57 gene with the pathogenesis of cervical lesions in Uighur women," *Zhonghua Zhong Liu Za Zhi*, vol. 35, no. 8, pp. 600–603, 2013.

[50] Y. Dong, Q. Tan, L. Tao et al., "Hypermethylation of TFPI2 correlates with cervical cancer incidence in the Uygur and Han populations of Xinjiang, China," *International Journal of Clinical and Experimental Pathology*, vol. 8, no. 2, pp. 1844–1854, 2015.

[51] J. Y. Liau, S. L. Liao, C. H. Hsiao, M. C. Lin, H. C. Chang, and K. T. Kuo, "Hypermethylation of the CDKN2a gene promoter is a frequent epigenetic change in periocular sebaceous carcinoma and is associated with younger patient age," *Human Pathology*, vol. 45, no. 3, pp. 533–539, 2014.

[52] C. Qiu, Y. Zhi, Y. Shen, J. Gong, Y. Li, and X. Li, "High-resolution melting analysis of HPV-16L1 gene methylation: a promising method for prognosing cervical cancer," *Clinical Biochemistry*, vol. 48, no. 13-14, pp. 855–859, 2015.

[53] L. Mirabello, M. Schiffman, A. Ghosh et al., "Elevated methylation of HPV16 DNA is associated with the development of high grade cervical intraepithelial neoplasia," *International Journal of Cancer*, vol. 132, no. 6, pp. 1412–1422, 2013.

[54] F. Yang-Chun, C. Zhen-Zhen, H. Yan-Chun, and M. Xiu-Min, "Association between PD-L1 and HPV status and the prognostic value for HPV treatment in premalignant cervical lesion patients," *Medicine*, vol. 96, no. 25, article e7270, 2017.

[55] J. L. Brandsma, M. Harigopal, N. B. Kiviat et al., "Methylation of twelve CpGs in human papillomavirus type 16 (HPV16) as an informative biomarker for the triage of women positive for HPV16 infection," *Cancer Prevention Research*, vol. 7, no. 5, pp. 526–533, 2014.

[56] D. Bryant, S. Hibbitts, M. Almonte, A. Tristram, A. Fiander, and N. Powell, "Human papillomavirus type 16 L1/L2 DNA methylation shows weak association with cervical disease grade in young women," *Journal of Clinical Virology*, vol. 66, pp. 66–71, 2015.

[57] M. Kalantari, K. Osann, I. E. Calleja-Macias et al., "Methylation of human papillomavirus 16, 18, 31, and 45 L2 and L1 genes and the cellular DAPK gene: considerations for use as biomarkers of the progression of cervical neoplasia," *Virology*, vol. 448, pp. 314–321, 2014.

[58] V. Simanaviciene, V. Popendikyte, Z. Gudleviciene, and A. Zvirbliene, "Different DNA methylation pattern of HPV16, HPV18 and HPV51 genomes in asymptomatic HPV infection as compared to cervical neoplasia," *Virology*, vol. 484, pp. 227–233, 2015.

[59] Y. W. Hsu, R. L. Huang, P. H. Su et al., "Genotype-specific methylation of HPV in cervical intraepithelial neoplasia," *Journal of Gynecologic Oncology*, vol. 28, no. 4, p. e56, 2017.

[60] B. László, A. Ferenczi, L. Madar et al., "CpG methylation in human papillomavirus (HPV) type 31 long control region (LCR) in cervical infections associated with cytological abnormalities," *Virus Genes*, vol. 52, no. 4, pp. 552–555, 2016.

[61] Q. Zuo, W. Zheng, J. Zhang et al., "Methylation in the promoters of HS3ST2 and CCNA1 genes is associated with cervical cancer in Uygur women in Xinjiang," *The International Journal of Biological Markers*, vol. 29, no. 4, pp. e354–e362, 2014.

[62] W. Zhang, Y. Jiang, Q. Yu et al., "EGFR promoter methylation, EGFR mutation, and HPV infection in Chinese cervical squamous cell carcinoma," *Applied Immunohistochemistry & Molecular Morphology*, vol. 23, no. 9, pp. 661–666, 2015.

[63] L. X. Bai, J. T. Wang, L. Ding et al., "Folate deficiency and FHIT hypermethylation and HPV 16 infection promote cervical cancerization," *Asian Pacific Journal of Cancer Prevention*, vol. 15, no. 21, pp. 9313–9317, 2014.

[64] J. Sakane, K. Taniyama, K. Miyamoto et al., "Aberrant DNA methylation of DLX4 and SIM1 is a predictive marker for disease progression of uterine cervical low-grade squamous intraepithelial lesion," *Diagnostic Cytopathology*, vol. 43, no. 6, pp. 462–470, 2015.

[65] S. Snellenberg, S. A. G. M. Cillessen, W. Van Criekinge et al., "Methylation-mediated repression of PRDM14 contributes to apoptosis evasion in HPV-positive cancers," *Carcinogenesis*, vol. 35, no. 11, pp. 2611–2618, 2014.

[66] R. Luttmer, L. M. A. De Strooper, J. Berkhof et al., "Comparing the performance of FAM19A4 methylation analysis, cytology and HPV16/18 genotyping for the detection of cervical (pre)-cancer in high-risk HPV-positive women of a gynecologic

outpatient population (COMETH study)," *International Journal of Cancer*, vol. 138, no. 4, pp. 992–1002, 2016.

[67] L. M. A. De Strooper, C. J. L. M. Meijer, J. Berkhof et al., "Methylation analysis of the FAM19A4 gene in cervical scrapes is highly efficient in detecting cervical carcinomas and advanced CIN2/3 lesions," *Cancer Prevention Research*, vol. 7, no. 12, pp. 1251–1257, 2014.

[68] R. D. Steenbergen, M. Ongenaert, S. Snellenberg et al., "Methylation-specific digital karyotyping of HPV16E6E7-expressing human keratinocytes identifies novel methylation events in cervical carcinogenesis," *The Journal of Pathology*, vol. 231, no. 1, pp. 53–62, 2013.

[69] L. Cicchini, J. A. Westrich, T. Xu et al., "Suppression of antitumor immune responses by human papillomavirus through epigenetic downregulation of CXCL14," *MBio*, vol. 7, no. 3, pp. e00270–e00216, 2016.

[70] K. Chalertpet, W. Pakdeechaidan, V. Patel, A. Mutirangura, and P. Yanatatsaneejit, "Human papillomavirus type 16 E7 oncoprotein mediates CCNA1 promoter methylation," *Cancer Science*, vol. 106, no. 10, pp. 1333–1340, 2015.

[71] K. H. Wang, C. J. Lin, C. J. Liu et al., "Global methylation silencing of clustered proto-cadherin genes in cervical cancer: serving as diagnostic markers comparable to HPV," *Cancer Medicine*, vol. 4, no. 1, pp. 43–55, 2015.

[72] V. M. J. Verhoef, D. A. M. Heideman, F. J. van Kemenade et al., "Methylation marker analysis and hpv16/18 genotyping in high-risk HPV positive self-sampled specimens to identify women with high grade CIN or cervical cancer," *Gynecologic Oncology*, vol. 135, no. 1, pp. 58–63, 2014.

[73] V. M. J. Verhoef, R. P. Bosgraaf, F. J. van Kemenade et al., "Triage by methylation-marker testing versus cytology in women who test HPV-positive on self-collected cervicovaginal specimens (PROHTECT-3): a randomised controlled non-inferiority trial," *The Lancet Oncology*, vol. 15, no. 3, pp. 315–322, 2014.

[74] L.-Y. Kong, W. Du, L. Wang, Z. Yang, and H.-S. Zhang, "PAX1 methylation hallmarks promising accuracy for cervical cancer screening in Asians: results from a meta-analysis," *Clinical Laboratory*, vol. 61, no. 10, pp. 1471–1479, 2015.

[75] E. M. Siegel, B. M. Riggs, A. L. Delmas, A. Koch, A. Hakam, and K. D. Brown, "Quantitative DNA methylation analysis of candidate genes in cervical cancer," *PLoS One*, vol. 10, no. 3, article e0122495, 2015.

[76] Y. Chen, Z. Cui, Z. Xiao et al., "PAX1 and SOX1 methylation as an initial screening method for cervical cancer: a meta-analysis of individual studies in Asians," *Annals of Translational Medicine*, vol. 4, no. 19, pp. 365–365, 2016.

[77] D. Bryant, A. Tristram, T. Liloglou, S. Hibbitts, A. Fiander, and N. Powell, "Quantitative measurement of human papillomavirus type 16 L1/L2 DNA methylation correlates with cervical disease grade," *Journal of Clinical Virology*, vol. 59, no. 1, pp. 24–29, 2014.

[78] S. Chujan, N. Kitkumthorn, S. Siriangkul, and A. Mutirangura, "CCNA1 promoter methylation: a potential marker for grading Papanicolaou smear cervical squamous intraepithelial lesions," *Asian Pacific Journal of Cancer Prevention*, vol. 15, no. 18, pp. 7971–7975, 2014.

[79] A. Maity and B. Das, "N6-methyladenosine modification in mRNA: machinery, function and implications for health and diseases," *The FEBS Journal*, vol. 283, no. 9, pp. 1607–1630, 2016.

[80] S. Ke, E. A. Alemu, C. Mertens et al., "A majority of m6A residues are in the last exons, allowing the potential for 3′ UTR regulation," *Genes & Development*, vol. 29, no. 19, pp. 2037–2053, 2015.

[81] F. Li, S. Kennedy, T. Hajian et al., "A radioactivity-based assay for screening human m6A-RNA methyltransferase, METTL3-METTL14 complex, and demethylase ALKBH5," *Journal of Biomolecular Screening*, vol. 21, no. 3, pp. 290–297, 2016.

[82] Q. Cui, H. Shi, P. Ye et al., "m^6A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells," *Cell Reports*, vol. 18, no. 11, pp. 2622–2634, 2017.

[83] A. Granados López and J. López, "Multistep model of cervical cancer: participation of miRNAs and coding genes," *International Journal of Molecular Sciences*, vol. 15, no. 9, pp. 15700–15733, 2014.

[84] X. Zhou, Y. Yue, R. Wang, B. Gong, and Z. Duan, "MicroRNA-145 inhibits tumorigenesis and invasion of cervical cancer stem cells," *International Journal of Oncology*, vol. 50, no. 3, pp. 853–862, 2017.

[85] C. Zhou, D. M. Tan, L. Chen et al., "Effect of miR-212 targeting TCF7L2 on the proliferation and metastasis of cervical cancer," *European Review for Medical and Pharmacological Sciences*, vol. 21, no. 2, pp. 219–226, 2017.

[86] H. Li, Y. Sheng, Y. Zhang, N. Gao, X. Deng, and X. Sheng, "MicroRNA-138 is a potential biomarker and tumor suppressor in human cervical carcinoma by reversely correlated with TCF3 gene," *Gynecologic Oncology*, vol. 145, no. 3, pp. 569–576, 2017.

[87] W. Li, J. Liang, Z. Zhang et al., "MicroRNA-329-3p targets MAPK1 to suppress cell proliferation, migration and invasion in cervical cancer," *Oncology Reports*, vol. 37, no. 5, pp. 2743–2750, 2017.

[88] S. Azizmohammadi, A. Safari, S. Azizmohammadi et al., "Molecular identification of miR-145 and miR-9 expression level as prognostic biomarkers for early-stage cervical cancer detection," *QJM*, vol. 110, no. 1, pp. 11–15, 2017.

[89] J. Zhang, J. Jia, L. Zhao et al., "Down-regulation of microRNA-9 leads to activation of IL-6/Jak/STAT3 pathway through directly targeting IL-6 in HeLa cell," *Molecular Carcinogenesis*, vol. 55, no. 5, pp. 732–742, 2016.

[90] A. Morel, A. Baguet, J. Perrard et al., "5azadC treatment upregulates miR-375 level and represses HPV16 E6 expression," *Oncotarget*, vol. 8, no. 28, pp. 46163–46176, 2017.

[91] C. L. Yeung, T. Y. Tsang, P. L. Yau, and T. T. Kwok, "Human papillomavirus type 16 E6 suppresses microRNA-23b expression in human cervical cancer cells through DNA methylation of the host gene C9orf3," *Oncotarget*, vol. 8, no. 7, pp. 12158–12173, 2017.

[92] J. Zhang, T. Yao, Y. Wang, J. Yu, Y. Liu, and Z. Lin, "Long noncoding RNA MEG3 is downregulated in cervical cancer and affects cell proliferation and apoptosis by regulating miR-21," *Cancer Biology & Therapy*, vol. 17, no. 1, pp. 104–113, 2015.

[93] S. Zhang, G. Zhang, and J. Liu, "Long noncoding RNA PVT1 promotes cervical cancer progression through epigenetically silencing miR-200b," *APMIS*, vol. 124, no. 8, pp. 649–658, 2016.

[94] S. Sood and R. Srinivasan, "Alterations in gene promoter methylation and transcript expression induced by cisplatin in comparison to 5-azacytidine in HeLa and SiHa cervical cancer cell lines," *Molecular and Cellular Biochemistry*, vol. 404, no. 1-2, pp. 181–191, 2015.

[95] G. Narayan, D. Xie, G. Ishdorj et al., "Epigenetic inactivation of TRAIL decoy receptors at 8p12-21.3 commonly deleted region confers sensitivity to Apo2l/trail-cisplatin combination therapy in cervical cancer," *Genes, Chromosomes & Cancer*, vol. 55, no. 2, pp. 177–189, 2016.

[96] E. Mani, L. A. Medina, K. Isaac-Olivé, and A. Dueñas-González, "Radiosensitization of cervical cancer cells with epigenetic drugs hydralazine and valproate," *European Journal of Gynaecological Oncology*, vol. 35, no. 2, pp. 140–142, 2014.

[97] S. Sood, F. D. Patel, S. Ghosh, A. Arora, L. K. Dhaliwal, and R. Srinivasan, "Epigenetic alteration by DNA methylation of ESR1, MYOD1 and hTERT gene promoters is useful for prediction of response in patients of locally advanced invasive cervical carcinoma treated by chemoradiation," *Clinical Oncology*, vol. 27, no. 12, pp. 720–727, 2015.

[98] M. H. Kim, M. S. Kim, W. Kim et al., "Suppressor of cytokine signaling (SOCS) genes are silenced by DNA hypermethylation and histone deacetylation and regulate response to radiotherapy in cervical cancer cells," *PLoS One*, vol. 10, no. 4, article e0123133, 2015.

[99] L. Cui, J. Song, L. Wu et al., "Smac is another pathway in the anti-tumour activity of trichosanthin and reverses trichosanthin resistance in CaSki cervical cancer cells," *Biomedicine & Pharmacotherapy*, vol. 69, pp. 119–124, 2015.

Hindawi
International Journal of Genomics
Volume 2018, Article ID 1465348, 13 pages
https://doi.org/10.1155/2018/1465348

Review Article

Endometriosis Malignant Transformation: Epigenetics as a Probable Mechanism in Ovarian Tumorigenesis

Jiaxing He, Weiqin Chang, Chunyang Feng, Manhua Cui, and Tianmin Xu

The Second Hospital of Jilin University, Jilin, Changchun 130041, China

Correspondence should be addressed to Tianmin Xu; xutianmin@126.com

Received 27 December 2017; Accepted 1 March 2018; Published 27 March 2018

Academic Editor: Changwon Park

Endometriosis, defined as the presence of ectopic endometrial glands and stroma outside the uterine cavity, is a chronic, hormone-dependent gynecologic disease affecting millions of women across the world, with symptoms including chronic pelvic pain, dysmenorrhea, dyspareunia, dysuria, and subfertility. In addition, there is well-established evidence that, although endometriosis is considered benign, it is associated with an increased risk of malignant transformation, with the involvement of various mechanisms of development. More and more evidence reveals an important contribution of epigenetic modification not only in endometriosis but also in mechanisms of endometriosis malignant transformation, including DNA methylation and demethylation, histone modifications, and miRNA aberrant expressions. In this present review, we mainly summarize the research progress about the current knowledge regarding the epigenetic modifications of the relations between endometriosis malignant transformation and ovarian cancer in an effort to identify some risk factors probably associated with ectopic endometrium transformation.

1. Introduction

Endometriosis is a chronic and hormone-dependent disease, defined as the presence of ectopic endometrial glands and stroma outside the uterine cavity [1]. The prevalence of endometriosis is widely different for various ethnic groups [2–4], which is likely to be about 5–15% of reproductive-age women and 3 to 5% of postmenopausal women [3]. To date, it is well established that endometriosis is a chronic inflammatory disease and the chronic inflammation is associated with pain and infertility [1]. As a common disease in reproductive woman, ectopic endometrium is predominantly detected in the pelvic compartment like the utero-sacral ligament, Douglas cavity, and ovary; moreover, the ectopic endometrial tissue can attach to other tissues including the bladder and ureter as well as the lung. Though several hypotheses are reported in order to explain the pathogenesis of endometriosis, mainly including coelomic metaplasia, retrograde menstruation, and lymphatic and vascular dissemination, none of them can explain all the different types of endometriosis.

Despite the fact that endometriosis is considered a benign condition because of its normal histology, the cellular, histologic, and molecular data strongly demonstrate that endometriosis has neoplastic characteristics [5, 6]. There is strong evidence that endometriosis shares striking features with malignancy [5]. Similar to cancer, ectopic endometrial tissue can result in normal tissue dissemination, invasion, and organ damage, as well as neoangiogenesis. It is reported that endometriosis is associated with ovarian cancer in all aspects of research fields including epigenetics; the link between endometriosis and ovarian cancer was reported for the first time as early as 1925 [7]. In the last nine decades, epidemiological investigation has been accumulated that endometriosis may contribute to the development and progression of ovarian cancer. In a cohort study by Melin et al., where 63,630 eligible women diagnosed with endometriosis entered, the risk of ovarian cancer (SIR 1.37) was moderately increased as compared with that of the general population [8]. As technology develops, multiple mechanisms about the occurrence of ovarian cancers associated with malignant transformation of endometriosis have been studied for a long time, but they

still remain elusive. Currently, it is well demonstrated that epigenetic modifications contribute to ovarian tumorigenesis. Epigenetics is described as a heritable modification in gene expression without alteration of DNA sequence compared with gene mutation [9]. The epigenetic modifications so far involve DNA methylation, histone modifications, and noncoding microRNAs (miRNAs) [10–12]. In the present review, we mainly summarize the research progress regarding the epigenetic modifications of the relations between endometriosis malignant transformation and ovarian cancer in an effort to identify some risk factors probably associated with ectopic endometrium transformation.

2. DNA Methylation

DNA methylation, the most frequently studied epigenetic alteration, occurs at the carbon-5 position of cytosine residues, exclusively in CpG dinucleotide sequences, and inhibits gene transcription [13]. DNA methylation, referring to the addition of the methyl groups into the cytosines from *S*-adenosyl L-methionine, is mediated by a family of enzymes known as the DNA methyltransferases (DNMTs) including DNMT1, DNMT3a, and DNMT3b. DNA methylation is a heritable epigenetic occurrence that significantly regulates gene expression without changing DNA sequence [14]. Most CpG sites in the human genome are methylated. However, local CpG islands, the CpG-rich regions, founded in the promoter regions of widely expressed genes are in unmethylated conditions [15, 16]. It is well evidenced that hypermethylation of genes can result in inhibition of gene expression, whilst hypomethylation may give rise to increased transcription and protein activation. Furthermore, a considerable number of evidence have proved the positive relation between DNA methylation and tumor occurrence and progression. On the other hand, DNA hypomethylation also contributes to oncogenesis when previously inactivated oncogenes are transcriptionally activated [17].

2.1. Genes Involved in Endometriosis Malignant Transformation. A number of genes, which are silenced or activated by DNA methylation, have been investigated in malignant transformation of endometriosis. Moreover, some researches that were published demonstrate the common epigenetic alteration between endometriosis and ovarian cancers. It is testified that some major genes are actually involved in the malignant transformation of ovarian endometriosis; among these contributing genes, epigenetic inactivation of Runt-related transcription factor 3 (RUNX3) [18], human mutL homolog 1 (hMLH1) [19], E-cadherin (CDH1) [20], Ras-association domain family of gene 2 (RASSF2) [21], and P16 and phosphatase and tensin homolog deleted on chromosome 10 (PTEN) [22] by promoter hypermethylation was well observed; however, long interspersed nuclear element-1 (LINE-1) [23] and syncytin-1 [24] were hypomethylated and activated. An example of this is the study carried out by Guo et al. [18] in which RUNX3 promoter hypermethylation, which results in RUNX3 inactivation and decreased RUNX3 protein expression, has been identified in the 18 of 30 (60%) patients with endometriosis-associated ovarian carcinoma (EAOC). Besides, the degree of RUNX3 hypermethylation and decreased RUNX3 protein expression in the eutopic endometrium from the EAOC group was significantly higher than that in the endometriosis (EM) and control endometrium (CE) groups. It is probable that the tissue histology in the eutopic endometrium may appear normal and intrinsic molecular abnormalities have occurred. Furthermore, it is evidenced that patients with surgical stage IC EAOC have a higher degree of RUNX3 hypermethylation than those with stages IA and IB. This phenomenon suggests that RUNX3 is implicated in the progression of malignant transformation of ovarian EM. Therefore, RUNX3 gene hypermethylation is reputed to be an early event in the pathogenesis of EAOC. Another similar study by Ren et al. [19] exploring the relationship between hMLH1 hypermethylation and malignant transformation of ovarian endometriosis is consistent with the results of the aforementioned RUNX3 promoter hypermethylation research. hMLH1 is a member of the DNA mismatch repair (MMR) system which corrects errors in DNA replication during proliferation. Ren and his colleagues illustrated this point clearly that absence of hMLH1 protein expression that resulted from aberrant promoter methylation is associated with malignant evolution of ovarian endometrium.

Other gene methylation may play an equivalent crucial role in the malignant transformation of endometriosis similar to that mentioned above, although few researchers have been able to draw on systematic researches into DNA methylation conditions of those relevant genes. It is well demonstrated that some crucial gene hypermethylation are implicated in the pathogenesis of EAOC. A recent research by Ren et al. [21] screened differentially aberrant methylated candidate genes associated with the malignant transformation of ovarian endometriosis by MCA-RDA, and nine differentially methylated candidate genes emerged in the study of malignant transformation of ovarian endometriosis. Among these nine candidate genes, RASSF2, SPOCK2, and RUNX3 were proved in other researches; therefore, the remaining six candidate genes were further studied including GSTZ1, CYP2A, GBGT1, NDUFS1, and ADAM22, as well as TRIM36. On the basis of those gene functions, they may take part in the malignant evolution of ovarian endometriosis. For example, ARID1A, identified as a tumor suppressor gene, encodes BAF250a, a key component of the SWI-SNF chromatin remodeling complex. A large number of researches demonstrate that the loss of ARID1A expression has been noted in approximately 40% of endometriotic lesions [25]. This study identified mutations in the ARID1A gene in ovarian clear cell and endometrioid carcinomas; these results represent that mutations in ARID1A are an early event in the malignant transformation of endometriosis. In a similar study, Lakshminarasimhan and his colleagues discovered that the downregulation of ARID1A expression in an endometriosis cell line enhances colony formation capacity, cell adhesiveness, and invasiveness, suggesting that low ARID1A expression might be an early event in the malignant transformation of endometriosis to ovarian clear cell carcinoma (OCCC) [26]. Although the link of ARID1A expression and OCCC transformation is well established, whether DNA

methylation of ARID1A significantly matters remains elusive. In addition, a study about hypermethylation of ARID1A in breast cancer by Zhang et al. [27] demonstrated that the promoter hypermethylation in the ARID1A gene is strongly associated with ARID1A gene low mRNA expression.

2.2. Are Hormones Useful for Endometriosis Transformation? It is known that endometriosis is an estrogen-sensitive and progesterone-resistant disease [28]. Estrogens have a paramount influence on various physiological processes including cell growth, reproduction, and differentiation and, in the meantime, also on pathological processes such as cancer, metabolic disease, and inflammation. The association between estrogen and various cancers is well reviewed [29]. Lots of clinical studies show that estradiol (E2) plays a key role in endometriosis. The role of E2 is regulated via the estrogen receptors (ERs) including estrogen receptors α (ER-α) and β (ER-β), which are, respectively, encoded by estrogen receptor gene 1 (ESR1) and estrogen receptor gene 2 (ESR2). Several studies investigated the expression of the ERs in the normal and ectopic endometrium of patients with endometriosis; a study reported by Cavallini et al. [30] confirmed the downregulation of ER-α and upregulation of ER-β in ovarian endometriotic tissue compared with eutopic tissue. Whether epigenetics such as DNA methylation is responsible for the expression of ERs in endometriotic cells needs to be further studied. Xue et al. [31] confirmed that the ESR2 gene promoter is hypomethylated in stromal endometriotic cells, which could be related to the upregulation of ER-β. However, low expression of the ER-β gene via promoter hypermethylation in tumors was observed [32]. Meyer et al. found [33] that ESR1 promoters (both ESR1A and ESR1B) are methylated, but the study reported by Toderow et al. [34] indicated that ER-α is not regulated by methylation of the promotor region in endometriosis. Without a doubt, estrogen-relevant genes and ER signal pathways are involved in the development of ovarian cancer [35–37]. Yamaguchi et al. [38], using MS-PCR to identify clear cell carcinoma-specific gene methylation, showed that 64 specific genes were involved in ER-associated pathways among the 276 hypermethylated genes. Representative ER pathway genes including ESR1, BMP4, DKK1, SOX11, SNCG, and MOSC1 are downregulated by promoter methylation, which are in accordance with decreased expression of ER-α. In addition, WT1, as one of the representative genes regulated by the ER-α signaling pathway, is downregulated in patients with endometriosis, consistent with the loss of WT1 expression in ovarian clear cell carcinoma [39, 40]. Furthermore, Akahane et al. [41] showed that decreased expression of ER-α occurred with progression from endometriosis to OCCC and that disappearance of hormone dependency might be associated with malignant transformation to OCCC. In conclusion, estrogen-relevant genes and pathways actually contribute to the malignant transformation of endometriosis, and the inconsistency of the ER-β gene expression between endometriosis and ovarian cancer and associated molecular mechanisms needs to be further investigated.

In the normal endometrium, progesterone strongly interacts with the activation of inflammatory pathways, recruits an influx of various immune cells, and mediates local inflammation [42]. Through binding to the nuclear receptors progesterone receptor isoform A (PRA) and progesterone receptor isoform B (PRB), which are members of the superfamily of ligand-activated transcription factors, the progesterone responses are regulated by directly binding to DNA and regulating the expression of target genes [43]. Several researches supported lower levels of protein expression of PRA and PRB in the eutopic endometrium and the ectopic lesions of patients compared with the normal endometrium of the control group [44]. In addition, another study, that by Fazleabas [45] using experimental animals as disease models, showed that the progesterone receptor (PR) and relevant signaling regulators exerted their effects in the early stages of endometriosis; however, with disease progression, PR expression and some targets of PR lost contact in the eutopic endometrium and the ectopic lesions of endometriosis. In short, it is well established that PR resistance plays an essential role in the occurrence, development, and progression of endometriosis, but it remains unclear whether epigenetic modifications such as DNA methylation contribute to alteration of PR-involved components. In a literature reported by Nie et al. [46], they investigated epigenetic modifications of hormones in endometriosis, and the results revealed that promoter of PRB was hypermethylated; additionally, treatment with both trichostatin A (TSA) and 5-aza-2'-deoxycytidine (ADC) increased PRB gene and protein expression in ectopic endometrial stromal cells but reduced cell viability of ectopic endometrial stromal cells. Another important study by Li et al. [47] investigating the consequences of inhibition of DNA methylation further revealed that lesion growth was ameliorated and PR and PR-target gene expression were restored; the results indicated a potential association between epigenetic regulation and PR-target signal pathways in the pathogenesis of endometriosis. In addition, one of the PR-involved components, Gata2, has been previously evidenced by Böhm et al. [48] as balancing the transcriptional activity of nuclear receptors including PGR; moreover, the expression of Gata2 gene also is consistent with that of PR in the epithelium and stroma of the uterus [49].

As indicated above, PR is dynamically associated with the occurrence of endometriosis and the progression of advanced endometriosis. Whether ovarian cancer involved in endometriosis malignant transformation also is controlled and regulated by epigenetic modification of progesterone related signaling pathways until now remains ambiguous. A previous research using immunohistochemical methods to compare the different expressions of tissues of endometriosis and EAOC indicated that the expression of PR and ER in EAOC was statistical significantly lower than that in endometriosis [50]. In addition, estrogen and progesterone regulated normal endometrial cells in proliferation and differentiation via mediating Wnt/β-catenin signaling whose main components include WNT7a, DKK-1, β-catenin, and GSK-3β [51]. However, abnormal activation the WNT/β-catenin signaling pathway via multiple regulators is reputed to be associated with ovarian cancer with epigenetic modification [52–56]. For instance, significant downregulation of the

Wnt antagonist SFRP5 is observed through the promoter hypermethylation associated with overall survival in ovarian cancer; moreover, epigenetic silence of SFRP5 expression leads to activation of the Wnt pathway and promotes ovarian cancer progression [56]. As indicated above, progesterone, together with other factors, probably plays a key role in indirectly regulating the progression and development of ovarian cancer. Though elaborate mechanisms about whether hormones actually contribute to endometriosis malignant transformation, to date, still remain mysterious, we will be capable of discovering and explaining the correlation between them in the future.

2.3. Oxidative Stress. A large body of literature has investigated that reactive oxygen species- (ROS-) mediated oxidative stress enacts a significant role in the pathophysiology of endometriosis [57–59]. ROS are a group of oxygen including chemically reactive molecules containing superoxide (O_2^-), hydroxyl (OH^-), hydrogen peroxide (H_2O_2), nitrogen oxide (NO), and nitrogen dioxide (NO_2). ROS are intermediaries produced by normal oxygen metabolism, whilst during excess of ROS release, the balance between ROS and antioxidants is broken and induces cellular damage through a variety of mechanisms, finally leading to harmful effects [58]. Many theories associated with ROS-induced endometriosis progression have been elaborated so far as possibly an important factor involved in the progression of endometriosis and malignant transformation. Oxidative stress can eliminate or induce specific DNA and histone methylation by regulating corresponding enzymes such as DNMTs and ten-eleven translocations (TETs); elaborate details are reviewed by Ito et al. in a recent literature [60]. For example, a latest study carried out by Xie et al. [61] which investigated the mechanism of the correlation between oxidative stress and ARID1A gene expression illustrated that ROS decreased the expression level of ARID1A gene via regulating the methylation of its promoter. In addition, another epigenetic enzyme, the TET family of hydroxylases of encoding genes (TET1, TET2, and TET3), significantly downregulated in endometriosis [62]. TET-mediated DNA demethylation may act as a protection against oxidative stress, which also can prove the link between oxidative stress and DNA methylation of endometriosis. Moreover, oxidative stress was extensively studied and reported in the amount of mechanisms of cancer including ovarian cancer [63–67]. Oxidative stress plays an effective role in carcinogenesis through epigenetic alterations. ROS lead to tumorigenesis by inhibiting or silencing of tumor suppressor genes through promoter hypermethylation. The intestine-specific transcription factor caudal type homeobox-1 (CDX1) is downregulated with the treatment of hydrogen peroxide (H_2O_2) because CDX1 promoter is hypermethylated and treatment with 5-aza-dC reversed this effect [68]. In ovarian cancer, a study reported by Hou et al. [69] showed that H_2O_2 downregulated miR-29b by directly targeting its mRNA 3′-UTR in ovarian cancer cells. Additionally, there is a new research proved by Mahalingaiah et al. [70] which demonstrated that a low level of chronic oxidative stress results in the malignant transformation of human renal tubular epithelial cells, and the potential role is the aberrant expression of epigenetic regulatory genes involved in DNA methylation (DNMTs) as well as histone modifications (HDAC1, HAT1) in human renal tubular epithelial cells malignantly transformed by chronic oxidative stress. Given all that, oxidative stress-mediated ovarian cancer malignantly transformed by endometriosis seems to hold great promise.

3. Histone Modifications

Histone modification exerts an equivalent effect on epigenetic regulation the same as DNA methylation. Histones are proteins that make up nucleosomes, which are the fundamental unit of chromatin. Epigenetic modifications such as acetylation, methylation, phosphorylation, and ubiquitylation regulate chromatin structure and gene expression. Histone acetylation is mediated by a class of enzymes called histone acetyl transferases (HATs), which allow chromatin to be in a more unstable state to accomplish gene expression. In contrast, histone deacetylation is regulated by the histone deacetylases (HDACs) and converts chromatin to a more condensed or transcriptionally repressive state for inhibiting gene expression [71]. For example, acetylation at lysine 9 (K9) of H3 is implicated in the transcriptionally active condition of chromatin [16, 72]. Unlike histone acetylation, histone methylation seems to be more elusive. Histone methylation can be either stimulatory or inhibitory to the condensed state of chromatin depending on the particular lysine residue modified. Furthermore, the extent of the methylation status (mono-, di-, and trimethylation) also remains implicated. For example, trimethylation of H3K4 is involved in the transcriptionally active condition of chromatin [16, 73]; however, inverse results present in the mono-, di-, and trimethylation of H3K9, which take part in repressing gene expression [74, 75]. Besides, other types of histone modifications (phosphorylation or ubiquitination) are associated with chromatin condensation status and regulating gene expression, forming a network of sophisticated crosstalk.

3.1. Aberrant Enzyme Expression. There is evidence to support the theory that aberrant HDAC pathways promote cancer growth and metastasis including ovarian cancer [76–78]. HDACs play a crucial role in regulating important cell processes such as cell growth, differentiation, and apoptosis. For instance, sirtuin 1 (SIRT1) is a promising family member and a class III HDAC, which regulates histone acetylation levels as well as the DNA repair [79]. There is a study that demonstrated that SIRT1 expression was significantly increased in epithelial ovarian carcinomas (EOCs) compared to benign tumors [80]. However, another study by Xiaomeng et al. revealed that SIRT1 expression level decreased in eutopic endometrium [81]. Moreover, this is supported by a recent study indicating that the expression of nuclear HDAC1, HDAC2, and HDAC3 proteins was increased in carcinomas compared with benign tumors [77]. A study investigating HDAC expression of endometriosis showed that levels of gene and protein expressions of HDAC1 and HDAC2 were higher in ectopic endometrium than in normal endometrium [82]. Beside, another crucial histone methylation is mediated

by histone methyltransferases and histone demethylase. Abnormal expressions of corresponding enzymes are probable to promote oncogene expression and progression of malignant tumors. For example, H3-K27 methylation is regulated by the enhancer of zester homolog 2 (EZH2), which is a key histone methyltransferase that belongs to a subunit of polycomb repressive complex 2. The results published by Guo et al. [83] found that EZH2 expression was significantly higher in ovarian carcinoma than in benign and normal tissues. Another study proved by Kuang and coworkers [84] revealed that EZH2 expression was positively correlated to KDM2B, which controls gene expression by the demethylation of dimethyl histone H3 lysine 36 (H3K36me2) and trimethyl histone H3 lysine 4 (H3K4me3). Both of them play an important role in the development and progression of ovarian cancer.

3.2. Change of Relevant Genes and Signal Pathways. In a literature, it was indicated that global histone H4 acetylation and histone H3K4 methylation level decreased significantly in both eutopic and ectopic endometrium compared with controls. However, there was no difference in H3 acetylation between endometriosis patients and controls [81]. In addition, a study using a dominant-negative histone overexpression approach demonstrated that the tumor suppressor gene RASSF1 is directly downregulated by the methylation of H3-K27. Furthermore, the results suggest that targeted epigenetic therapies of H3-K27 methylation hold great promise [85]. Furthermore, it is well recognized that signal pathways are associated with various aspects of cancer progression. A study by Hurst et al. [86] investigated the molecular mechanism of G-protein-coupled receptor (GPCR) pathways. They discovered that in the regulator of G-protein signaling 2 (RGS2), as an inhibitor of GPCRs, its protein expression is downregulated in ovarian cancer progression. A relevant study by Cacan [87] showed that loss of histone acetylation at RGS2 promoter genes results in the loss of RGS2 expression and indicated that the downregulation of the RGS2 gene is partly due to accumulation of HDACs at the promoter region of RGS2 in chemoresistant ovarian cancer cells.

As mentioned above, we talked about the association between endometriosis and ovarian cancer, such as HDAC gene expression and the increase of proteins of both HDAC1 and HDAC2 in endometriosis and ovarian cancer. Therefore, HDAC expression may be responsible for the malignant transformation of endometriosis although the association between them is still ambiguous. On the other hand, more marked differences were observed in the results of two kinds of researches, such as SIRT1 upregulation in ovarian cancer and downregulation in endometriosis. In spite of those opposite results, the mechanisms of endometriosis malignant evolution will be completed in the future.

3.3. Inflammation and Immune Disorders Need to Be Further Investigated in Epigenetic Modification. Without a doubt, endometriosis is a complex, chronic inflammatory disease with variable symptoms in women. Inflammation and immune disorders absolutely play a key role in the pathobiology of endometriosis; therefore, inflammatory cells and inflammatory cytokines are regulated as target components in endometriosis patients; also, the immune system of women with endometriosis is also dysfunctional. The immune system contains a variety of immune cells, including macrophages, dendritic cells, natural killer cells, T helper cells, and B cells, which have been proved to be disordered in patients with endometriosis [88, 89]. Disorders of inflammatory cell populations in ectopic endometrium and their secretory products exert a harmful influence on normal microenvironment, inducing the development of the disease. Moreover, the immune system and endometrial cells secrete several cytokines and growth factors that promote invasion and growth of ectopic endometrium [90]; those cytokines and growth factors are representative inflammatory mediators, leading to inflammatory response and finally aggravating endometriosis, even causing ovarian cancer.

To date, the most robust pathogenic hypothesis involved in inflammation response is based on the so-called retrograde menstruation phenomenon. Through retrograde flow, viable endometrial fragments reach and implant onto the peritoneum and abdominal organs, leading to chronic inflammation with formation of adhesions and severe infertility. Chronic inflammation, in turn, also promotes proliferation and growth of ectopic endometrial tissue [28]. The presence of ectopic tissue is associated with secretion disorder of inflammatory cells and factors. Macrophages and associated signaling cascade are alternatively activated in patients with endometriosis, which were observed in the study reported by Mahdian et al. [91]. In addition, various inflammatory factors play different roles in infertility in patients with endometriosis [92, 93]. A study by Yang et al. [94] investigated the relations between exposure to pelvic microenvironments with overproduced inflammatory factors and structural or functional tissue abnormalities; their data showed that telocytes (TCs; interstitial Cajal-like cells (ICLCs)) were significantly decreased and interstitial fibrosis was observed, accompanied with an increased level of inducible nitric oxide synthase (iNOS), cyclooxygenase-2 (COX-2), lipid peroxide (LPO), and estradiol, which suggested that inflammation induced TC damage and fibrosis and dysmotility of the oviduct finally leading to subfertility or infertility. Yoshida et al. [95] indicated that interleukins 1 and 6 directly affect sperm mobility. In addition, Hosseini et al. indicated that epigenetic changes of CYP19A1 (aromatase) gene promoters may lead to poor oocyte and embryo condition by impairing follicular steroidogenesis in patients with endometriosis [96]. Tao et al., who studied the pathogenesis of endometriosis-associated infertility, confirmed the tight correlation between monocyte chemotactic protein- (MCP-) 1 and peritoneal leptin levels and infertility in the early stage of endometriosis [97]. Rathore et al. indicated that ghrelin and leptin might contribute to the pathophysiology of infertility, and leptin is associated with inflammatory factors such as IL-6 in patients with endometriosis [98] (Figure 1).

In recent years, cytokines caught the intense attention of researchers due to their involvement in the pathogenesis of endometriosis and cancer. Endometriosis is often accompanied by marked changes of inflammatory cytokines,

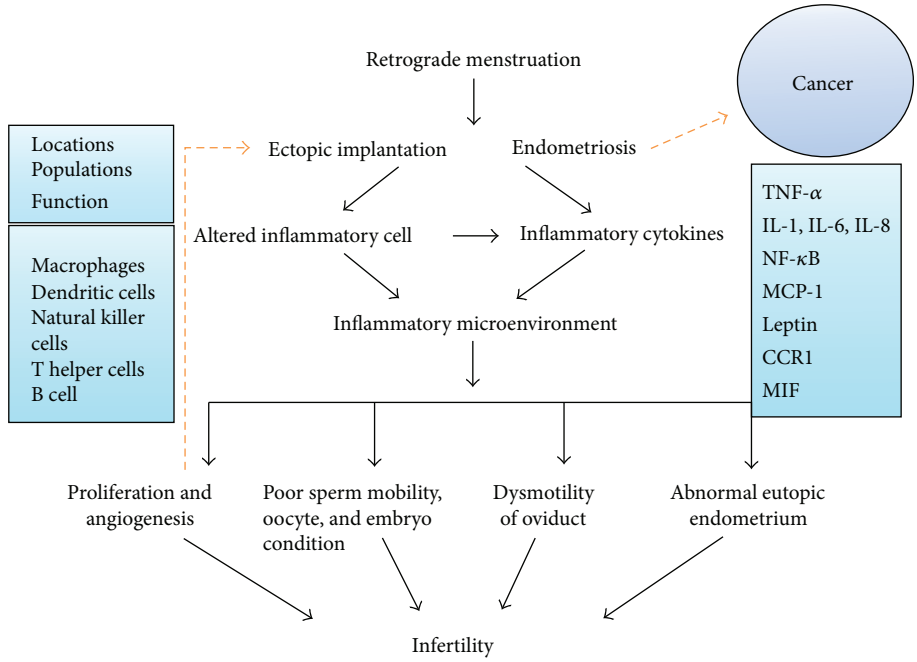


Figure 1: The potential inflammatory mechanisms between endometriosis and infertility. The figure indicates that various factors may result in infertility in patients with endometriosis. Inflammatory cytokines are secreted by inflammatory cells including TNF-α, IL-1, IL-6, IL-8, NF-κB, MCP-1, leptin, CCR1, MIF, and COX-2. Inflammatory responses depend on locations, populations, and functions of inflammatory cells, which include macrophages, dendritic cells, natural killer cells, T helper cells, and B cells. On the one hand, inflammatory responses alter microenvironment and influence various aspects of fertility; on the other hand, chronic exposure to microenvironments with overproduced inflammatory factors leads to ectopic implant proliferations and angiogenesis, which in turn promote growth and invasion of ectopic endometrium even in the development of cancer.

including epithelial cell-derived neutrophil-activating peptide-78 (ENA-78), macrophage migration inhibitory factor (MIF), high-sensitivity C-reactive protein (hs-CRP), tumor necrosis factors (TNF-α), interleukin-1β (IL-1β), IL-6, IL-8, interferon-induced protein 10 (IP-10), and chemokine receptor 1 (CCR1) [99, 100]. There is great expectations for one of them, IL-8, in the development and progression of endometriosis. The literature strongly suggests that IL-8 might play an important role in adhesion and growth of the endometrial implants [101, 102]. A literature by Ulukus et al. [103] showed higher epithelial IL-8 expression in eutopic endometrium of patients with endometriosis, as compared to normal women. In addition, the literature suggested that increased IL-8 expression levels in women with endometriosis might contribute to the development of endometriosis and finally progression of chronic inflammation, even probably malignant transformation [104, 105]. In ovarian cancer, epigenetic modifications are regarded as regulations and mediations of ovarian cancer development, and epigenetic therapies as inhibition of ovarian cancer cells. For instance, the study investigated the specific involvement of HDACs and HATs in the epigenetic regulation of IL-8 expression in ovarian cancer cells; the results indicated that the IL-8 expression in OC cells is regulated by CBP and might enhance effectiveness of HDAC inhibitors in OC treatment [106]. They previously showed that inhibition of histone deacetylase (HDAC) activity increased IL-8 expression in OC cells, resulting in their increased survival and proliferation [107].

The role of chemokine epigenetic regulations in endometriosis malignant transformation has so far remained ambiguous, but those researches suggest that chemokines might appear as a valuable tool to influence the correlation between endometriosis and ovarian cancers and perform early diagnosis of ovarian tumors.

4. MicroRNA Alteration

Over the past 20 years, another epigenetic regulation of gene expression has been discovered and well established, which participates in posttranscriptional gene downregulation mediated by small, non-protein-coding RNA molecules named microRNAs (miRNAs) [108, 109]. Since their initial discovery in 1993 [110], small noncoding RNAs or microRNAs have been intensely investigated across almost all biomedical fields including tumors. MicroRNAs, a class of single strand, are endogenously expressed, and noncoding RNAs, approximately 22 nucleotides in length, were found to mediate a series of essential biological processes including cell cycle, differentiation, development, and apoptosis as well as metabolism [111–114]. Plenty of work has been absorbed in investigating the biogenesis of miRNAs.

miRNA host gene is transcribed by RNA polymerase II, and transcription products are so-called pri-miRNA [115]. The pri-miRNA has to undergo two processing steps in order to become a mature miRNA. First, pri-miRNA is precisely recognized and cleaved by the enzyme Drosha which interacted with an RNA binding protein DGCR8, forming the

so-called "microprocessor." The following processing step occurs in cytoplasm: the predominant enzyme called Dicer, also associated with an RNA binding protein named TRBP [116], cleaves the pri-miRNA to a short RNA duplex approximately 21–25 nucleotides in length, depending on the type of Dicer and miRNA [117]. The two strands of the duplex have their own independent influence: one strand of duplex is incorporated into RNA-induced silencing complex (RISC) as a mature miRNA exerting its function, and the remaining strand is usually degraded. However, both strands of some miRNAs are likely to be selected into RISC. The RISC or microRNAs are able to prevent mRNA translation and induce mRNA degradation by matching $3'$ untranslated regions of target mRNAs; this phenomenon is known as RNA interference. MicroRNAs are strictly controlled in normal cells. Once microRNAs become deregulated, those aberrant productions can lead to the occurrence and progression of diseases. Researches showed that microRNAs may be associated with the pathogenesis of various human cancers.

4.1. Contributing MicroRNAs in Ovarian Cancer and Endometriosis. In ovarian cancer, the role of miRNAs is present in different biological processes including cell cycle, apoptosis, proliferation, invasion, and metastasis, and even chemoresistance. miRNAs are mediated by Drosha and Dicer; several reports showed that Dicer and Drosha mRNA expression levels and the corresponding proteins were decreased in the majority of ovarian cancers compared with normal tissues [118, 119]. Moreover, ovarian cancers were found to significantly upregulate four members of miR-200 family of miRNAs containing miR-200a, miR-141, miR-200c, and miR-200b, whereas miR-199a, miR-140, miR-145, and miR-125b1 were downregulated among most miRNAs [120]. On the other hand, alterations in the expression levels of different members of miR-200 family are differently associated with the distinct histotypes of ovarian carcinomas. miR-200a and miR-200c overexpressions occur in all the three histotypes including serous and endometrioid as well as clear cell, whereas miR-200b and miR-141 upmodulation exists in endometrioid and serous histotypes [120]. Additionally, another family of miRNAs, the let-7 (lethal-7) family, as tumor suppressor miRNAs, also gets widespread attention in multiple human tumors [121]. Remarkably reduced expressions of the let-7i were observed in tumors of cancer patients with poor survival [122, 123]. In the ovarian cancer, let-7i significantly reduced expression in chemotherapy-resistant patients as reported in a study; moreover, reduced let-7i expression significantly increased the resistance of ovarian cancer cells to the chemotherapy drug *cis*-platinum [123]. Several studies also unravel other miRNA expressions associated with ovarian cancers; for instance, miR-16, miR-20a, miR-21, miR-23a, miR-23b, miR-27a, miR-93, miR-30c and miR-30d, and miR-30e-3p were found to be overregulated, whereas miR-10b, miR-26a, miR-29a, miR-99a, miR-100, miR-125a, miR-125b, miR-143, miR-145, miR-199a, miR-214, miR-22, and miR-519a have opposite expressions [124–126]. Consequently, different miRNAs exert their influence on ovarian cancer development and progression.

Only few researches about malignant transformation were published. Several recent studies investigated the mechanisms of malignant transformation of endometriosis. Tissue inhibitor of metalloproteinases 3 (TIMP3), a proapoptotic protein [127], is proved by Qin and coworkers as a direct target of miR-191 [128]. The data of the study revealed that miR-191 expression was significantly higher in both endometriosis and EAOC and that TIMP3 expression was negatively correlated with miR-191 expression [129]. Additionally, another study, discussing whether cancer-associated miRNA single nucleotide polymorphisms (miRSNPs) accelerate endometriosis development and progression, demonstrated that MIR196A2 and MIR100 influenced endometriosis development and related clinical phenotypes [130].

4.2. EMT Affects Early Stage of the Oncogenic Transformation. Epithelial-mesenchymal transition (EMT) is a highly conserved cellular process that converts immotile and polarized epithelial cells to motile mesenchymal cells, which occurs in embryonic development, fibrosis, and wound healing; meantime, cancer development and progression that resemble embryonic development are regulated and controlled by EMT [131]. As usual, whether EMT occurs or not is detected by corresponding protein expression marks; E-cadherin and cytokeratins are the most common markers for the epithelial phenotypes and N-cadherin and vimentin for the mesenchymal [131]. Multiple noncoding RNAs are reputed to govern EMT; two major regulatory networks are demonstrated to be considered as the core regulatory machinery—the miR-34-SNAI1 and miR-200-ZEB1 axes—which are also controlled by various mediators.

EMT processes enhance migration and invasion of cells, which are prerequisites for the implantation of endometriotic lesions. A previous study reported by Bartley et al. [132] showed that the expression levels of N-cadherin, Twist and Snail, were significantly higher in endometriosis than in endometrium. However, in endometriosis, the expression of E-cadherin was inversely decreased in comparison with that in endometrium. Another study also proved that EMT-related processes might be involved in the pathogenesis of pelvic endometriosis [133]. Filigheddu et al. [134] showed the downregulation of miR-200b expression in the ectopic endometrium compared with the eutopic endometrium of endometriosis patients, together with enhancing of EMT. Eggers et al. [135], who investigated whether miR-200b expression contributes to EMT and invasive growth in endometriosis, indicated that upregulation of miR-200b reverts EMT and inhibits migration and invasion of cells of the endometriotic cell.

The significance of EMT during cancer progression has been commonly recognized, and EMT processes are thought to take part in many cancer cell metastases and progression. For instance, EMT occurs at the invasive front and single mesenchymal-like cells are detected to lose E-cadherin expression in colon carcinoma [136]. Knockdown of Linc-ROR in ovarian cancer cell lines prevents EMT processes through the repression of Wnt/β-catenin signaling; the results suggest that EMT can be an important phenomenon

in the invasion and metastasis of ovarian cancer [137]. Another study showed that inhibition of miR-23a reduced the TGF-β1-induced EMT, invasion, and metastasis in breast cancer cells though directly targeting CDH1 that activated induced Wnt/β-catenin signaling [138]. This phenomenon is also discovered in other tumors, which reveals an EMT expression profile and shows increased vimentin and loss of E-cadherin.

As mentioned above, EMT is reputed to be an invasive behavior and enables normal cells to be more aggressive, further having a potential malignant tendency. EMT may exert its momentous transitional effects on endometriosis malignant transformation; furthermore, the occurrence of EMT is probable in an early stage event of endometriosis malignant ovarian cancer in the future.

5. Conclusion

The overall aim of this review was to summarize the epigenetic modifications of the relations between endometriosis malignant transformation and ovarian cancer; moreover, due to the constraints of research progress and attention, we talked more about the potential correlations between them by relevant literatures. Beyond the abovementioned epigenetic modifications, other epigenetics such as long noncoding RNA (lncRNA) and posttranslational modifications (PTMs) are associated with the pathogenesis of endometriosis and ovarian cancer [139–143]. As complex gynecologic disease is closely related to the cancer, enigmatic etiology of endometriosis and mechanism of endometriosis malignant transformation are absolute worth correctly uncovering and elucidating in the future. In recent years, researchers have made significant strides in understanding the disease-specific molecular pathways governing the development of endometriosis in ectopic locations by studying the blood, peritoneal fluid, and eutopic endometrium of women with the disease [144–147]. Vicente-Muñoz et al. [147] identified the plasma metabolites of endometriosis patients and found higher concentration of valine, fucose, choline-containing metabolites, lysine/arginine, and lipoproteins and lower concentration of creatinine than in healthy women, which can help to get a better understanding of the molecular mechanisms of endometriosis. Studying epigenetic modifications of endometriosis, as well as investigating the correlation between endometriosis and ovarian cancer, will propel our understanding of the pathogenesis of endometriosis malignant transformation, with the potential for early diagnostic interventions and new effective therapies.

Conflicts of Interest

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

Authors' Contributions

Jiaxing He and Weiqin Chang contributed to this work equally and should be considered as co-first authors.

Acknowledgments

This study was supported by grants from the National Natural Science Foundation of China (81302242), Department of Science and Technology of Jilin Province (20150204007YY and 20140204022YY), Development and Reform Commission of Jilin Province (2014G073 and 2016C046-2), and Department of Education of Jilin Province (JJKH20170804KJ).

References

[1] P. Vercellini, P. Viganò, E. Somigliana, and L. Fedele, "Endometriosis: pathogenesis and treatment," *Nature Reviews Endocrinology*, vol. 10, no. 5, pp. 261–275, 2014.

[2] P. Viganò, F. Parazzini, E. Somigliana, and P. Vercellini, "Endometriosis: epidemiology and aetiological factors," *Best Practice & Research Clinical Obstetrics & Gynaecology*, vol. 18, no. 2, pp. 177–200, 2004.

[3] H. S. Kim, T. H. Kim, H. H. Chung, and Y. S. Song, "Risk and prognosis of ovarian cancer in women with endometriosis: a meta-analysis," *British Journal of Cancer*, vol. 110, no. 7, pp. 1878–1890, 2014.

[4] V. H. Eisenberg, C. Weil, G. Chodick, and V. Shalev, "Epidemiology of endometriosis: a large population-based database study from a healthcare provider with 2 million members," *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 125, no. 1, pp. 55–62, 2018.

[5] J. Siufi Neto, R. M. Kho, D. F. dos Santos Siufi, E. C. Baracat, K. S. Anderson, and M. S. Abrão, "Cellular, histologic, and molecular changes associated with endometriosis and ovarian cancer," *The Journal of Minimally Invasive Gynecology*, vol. 21, no. 1, pp. 55–63, 2014.

[6] P. S. Munksgaard and J. Blaakaer, "The association between endometriosis and ovarian cancer: a review of histological, genetic and molecular alterations," *Gynecologic Oncology*, vol. 124, no. 1, pp. 164–169, 2012.

[7] J. A. Sampson, "Endometrial carcinoma of the ovary, arising in endometrial tissue in that organ," *Archives of Surgery*, vol. 10, no. 1, pp. 1–72, 1925.

[8] A. Melin, P. Sparen, and A. Bergqvist, "The risk of cancer and the role of parity among women with endometriosis," *Human Reproduction*, vol. 22, no. 11, pp. 3021–3026, 2007.

[9] C. A. Barton, N. F. Hacker, S. J. Clark, and P. M. O'Brien, "DNA methylation changes in ovarian cancer: implications for early diagnosis, prognosis and treatment," *Gynecologic Oncology*, vol. 109, no. 1, pp. 129–139, 2008.

[10] F. W. Grimstad and A. Decherney, "A review of the epigenetic contributions to endometriosis," *Clinical Obstetrics and Gynecology*, vol. 60, no. 3, pp. 467–476, 2017.

[11] B. Borghese, K. T. Zondervan, M. S. Abrao, C. Chapron, and D. Vaiman, "Recent insights on the genetics and epigenetics of endometriosis," *Clinical Genetics*, vol. 91, no. 2, pp. 254–264, 2017.

[12] S. Gordts, P. Koninckx, and I. Brosens, "Pathogenesis of deep endometriosis," *Fertility and Sterility*, vol. 108, no. 6, pp. 872–885.e1, 2017.

[13] H. J. Smith, J. M. Straughn, D. J. Buchsbaum, and R. C. Arend, "Epigenetic therapy for the treatment of epithelial ovarian cancer: a clinical review," *Gynecologic Oncology Reports*, vol. 20, pp. 81–86, 2017.

[14] O. Koukoura, S. Sifakis, and D. A. Spandidos, "DNA methylation in endometriosis (review)," *Molecular Medicine Reports*, vol. 13, no. 4, pp. 2939–2948, 2016.

[15] J. Sandoval and M. Esteller, "Cancer epigenomics: beyond genomics," *Current Opinion in Genetics & Development*, vol. 22, no. 1, pp. 50–55, 2012.

[16] L. Maldonado and M. O. Hoque, "Epigenomics and ovarian carcinoma," *Biomarkers in Medicine*, vol. 4, no. 4, pp. 543–570, 2010.

[17] H. T. Nguyen, G. Tian, and M. M. Murph, "Molecular epigenetics in the management of ovarian cancer: are we investigating a rational clinical promise?," *Frontiers in Oncology*, vol. 4, p. 71, 2014.

[18] C. Guo, F. Ren, D. Wang et al., "RUNX3 is inactivated by promoter hypermethylation in malignant transformation of ovarian endometriosis," *Oncology Reports*, vol. 32, no. 6, pp. 2580–2588, 2014.

[19] F. Ren, D. Wang, Y. Jiang, and F. Ren, "Epigenetic inactivation of hMLH1 in the malignant transformation of ovarian endometriosis," *Archives of Gynecology and Obstetrics*, vol. 285, no. 1, pp. 215–221, 2012.

[20] Y. Li, D. An, Y. X. Guan, and S. Kang, "Aberrant methylation of the E-cadherin gene promoter region in endometrium and ovarian endometriotic cysts of patients with ovarian endometriosis," *Gynecologic and Obstetric Investigation*, vol. 82, no. 1, pp. 78–85, 2017.

[21] F. Ren, D. B. Wang, T. Li, Y. H. Chen, and Y. Li, "Identification of differentially methylated genes in the malignant transformation of ovarian endometriosis," *Journal of Ovarian Research*, vol. 7, no. 1, p. 73, 2014.

[22] M. Martini, M. Ciccarone, G. Garganese et al., "Possible involvement of *hMLH1*, $p16^{INK4a}$ and *PTEN* in the malignant transformation of endometriosis," *International Journal of Cancer*, vol. 102, no. 4, pp. 398–406, 2002.

[23] A. Senthong, N. Kitkumthorn, P. Rattanatanyong, N. Khemapech, S. Triratanachart, and A. Mutirangura, "Differences in LINE-1 methylation between endometriotic ovarian cyst and endometriosis-associated ovarian cancer," *International Journal of Gynecological Cancer*, vol. 24, no. 1, pp. 36–42, 2014.

[24] H. Zhou, J. Li, K. Podratz et al., "Hypomethylation and activation of syncytin-1 gene in endometriotic tissue," *Current Pharmaceutical Design*, vol. 20, no. 11, pp. 1786–1795, 2014.

[25] K. C. Wiegand, S. P. Shah, O. M. al-Agha et al., "ARID1A mutations in endometriosis-associated ovarian carcinomas," *The New England Journal of Medicine*, vol. 363, no. 16, pp. 1532–1543, 2010.

[26] R. Lakshminarasimhan, C. Andreu-Vieyra, K. Lawrenson et al., "Down-regulation of ARID1A is sufficient to initiate neoplastic transformation along with epigenetic reprogramming in non-tumorigenic endometriotic cells," *Cancer Letters*, vol. 401, pp. 11–19, 2017.

[27] X. Zhang, Q. Sun, M. Shan et al., "Promoter hypermethylation of ARID1A gene is responsible for its low mRNA expression in many invasive breast cancers," *PLoS One*, vol. 8, no. 1, article e53931, 2013.

[28] L. C. Giudice and L. C. Kao, "Endometriosis," *The Lancet*, vol. 364, no. 9447, pp. 1789–1799, 2004.

[29] M. Jia, K. Dahlman-Wright, and J.-Å. Gustafsson, "Estrogen receptor alpha and beta in health and disease," *Best Practice & Research Clinical Endocrinology & Metabolism*, vol. 29, no. 4, pp. 557–568, 2015.

[30] A. Cavallini, L. Resta, A. M. Caringella, E. Dinaro, C. Lippolis, and G. Loverro, "Involvement of estrogen receptor-related receptors in human ovarian endometriosis," *Fertility and Sterility*, vol. 96, no. 1, pp. 102–106, 2011.

[31] Q. Xue, Z. Lin, Y. H. Cheng et al., "Promoter methylation regulates estrogen receptor 2 in human endometrium and endometriosis," *Biology of Reproduction*, vol. 77, no. 4, pp. 681–687, 2007.

[32] I. Kyriakidis and P. Papaioannidou, "Estrogen receptor beta and ovarian cancer: a key to pathogenesis and response to therapy," *Archives of Gynecology and Obstetrics*, vol. 293, no. 6, pp. 1161–1168, 2016.

[33] J. L. Meyer, D. Zimbardi, S. Podgaec, R. L. Amorim, M. S. Abrão, and C. A. Rainho, "DNA methylation patterns of steroid receptor genes ESR1, ESR2 and PGR in deep endometriosis compromising the rectum," *International Journal of Molecular Medicine*, vol. 33, no. 4, pp. 897–904, 2014.

[34] V. Toderow, M. Rahmeh, S. Hofmann et al., "Promotor analysis of ESR1 in endometrial cancer cell lines, endometrial and endometriotic tissue," *Archives of Gynecology and Obstetrics*, vol. 296, no. 2, pp. 269–276, 2017.

[35] A. Bardin, N. Boulle, G. Lazennec, F. Vignon, and P. Pujol, "Loss of ERβ expression as a common step in estrogen-dependent tumor progression," *Endocrine-Related Cancer*, vol. 11, no. 3, pp. 537–551, 2004.

[36] S. Y. Jeon, K. A. Hwang, and K. C. Choi, "Effect of steroid hormones, estrogen and progesterone, on epithelial mesenchymal transition in ovarian cancer development," *The Journal of Steroid Biochemistry and Molecular Biology*, vol. 158, pp. 1–8, 2016.

[37] A. Ciucci, G. F. Zannoni, M. Buttarelli et al., "Multiple direct and indirect mechanisms drive estrogen-induced tumor growth in high grade serous ovarian cancers," *Oncotarget*, vol. 7, no. 7, pp. 8155–8171, 2016.

[38] K. Yamaguchi, Z. Huang, N. Matsumura et al., "Epigenetic determinants of ovarian clear cell carcinoma biology," *International Journal of Cancer*, vol. 135, no. 3, pp. 585–597, 2014.

[39] M. J. Worley Jr., S. Liu, Y. Hua et al., "Molecular changes in endometriosis-associated ovarian clear cell carcinoma," *European Journal of Cancer*, vol. 51, no. 13, pp. 1831–1842, 2015.

[40] S. Matsuzaki, M. Canis, C. Darcha, P. J. Déchelotte, J. L. Pouly, and G. Mage, "Expression of WT1 is down-regulated in eutopic endometrium obtained during the midsecretory phase from patients with endometriosis," *Fertility and Sterility*, vol. 86, no. 3, pp. 554–558, 2006.

[41] T. Akahane, A. Sekizawa, T. Okuda, M. Kushima, H. Saito, and T. Okai, "Disappearance of steroid hormone dependency during malignant transformation of ovarian clear cell cancer," *International Journal of Gynecological Pathology*, vol. 24, no. 4, pp. 369–376, 2005.

[42] J. J. Brosens, M. S. C. Wilson, and E. W.-F. Lam, "FOXO transcription factors: from cell fate decisions to regulation of human female reproduction," *Advances in Experimental Medicine and Biology*, vol. 665, pp. 227–241, 2009.

[43] M. Al-Sabbagh, E. W.-F. Lam, and J. J. Brosens, "Mechanisms of endometrial progesterone resistance," *Molecular and Cellular Endocrinology*, vol. 358, no. 2, pp. 208–215, 2012.

[44] A. Hayashi, A. Tanabe, S. Kawabe et al., "Dienogest increases the progesterone receptor isoform B/A ratio in patients with ovarian endometriosis," *Journal of Ovarian Research*, vol. 5, no. 1, p. 31, 2012.

[45] A. Fazleabas, "Progesterone resistance in a baboon model of endometriosis," *Seminars in Reproductive Medicine*, vol. 28, no. 1, pp. 075–080, 2010.

[46] J. Nie, X. Liu, and S.-W. Guo, "Promoter hypermethylation of progesterone receptor isoform B (PR-B) in adenomyosis and its rectification by a histone deacetylase inhibitor and a demethylation agent," *Reproductive Sciences*, vol. 17, no. 11, pp. 995–1005, 2010.

[47] Y. Li, M. K. Adur, A. Kannan et al., "Progesterone alleviates endometriosis via inhibition of uterine cell proliferation, inflammation and angiogenesis in an immunocompetent mouse model," *PLoS One*, vol. 11, no. 10, article e0165347, 2016.

[48] M. Böhm, W. J. Locke, R. L. Sutherland, J. G. Kench, and S. M. Henshall, "A role for GATA-2 in transition to an aggressive phenotype in prostate cancer through modulation of key androgen-regulated genes," *Oncogene*, vol. 28, no. 43, pp. 3847–3856, 2009.

[49] C. A. Rubel, H. L. Franco, J. W. Jeong, J. P. Lydon, and F. J. DeMayo, "GATA2 is expressed at critical times in the mouse uterus during pregnancy," *Gene Expression Patterns*, vol. 12, no. 5-6, pp. 196–203, 2012.

[50] M. G. del Carmen, A. E. Smith Sehdev, A. N. Fader et al., "Endometriosis-associated ovarian carcinoma: differential expression of vascular endothelial growth factor and estrogen/progesterone receptors," *Cancer*, vol. 98, no. 8, pp. 1658–1663, 2003.

[51] A. Pazhohan, F. Amidi, F. Akbari-Asbagh et al., "The Wnt/β-catenin signaling in endometriosis, the expression of total and active forms of β-catenin, total and inactive forms of glycogen synthase kinase-3β, WNT7a and DICKKOPF-1," *European Journal of Obstetrics & Gynecology and Reproductive Biology*, vol. 220, pp. 1–5, 2018.

[52] J. A. MacLean, M. L. King, H. Okuda, and K. Hayashi, "WNT7A regulation by miR-15b in ovarian cancer," *PLoS One*, vol. 11, no. 5, article e0156109, 2016.

[53] S. Yoshioka, M. L. King, S. Ran et al., "WNT7A regulates tumor growth and progression in ovarian cancer through the WNT/β-catenin pathway," *Molecular Cancer Research*, vol. 10, no. 3, pp. 469–482, 2012.

[54] M. L. King, M. E. Lindberg, G. R. Stodden et al., "WNT7A/β-catenin signaling induces FGF1 and influences sensitivity to niclosamide in ovarian cancer," *Oncogene*, vol. 34, no. 26, pp. 3452–3462, 2015.

[55] Z. Deng, L. Wang, H. Hou, J. Zhou, and X. Li, "Epigenetic regulation of IQGAP2 promotes ovarian cancer progression via activating Wnt/β-catenin signaling," *International Journal of Oncology*, vol. 48, no. 1, pp. 153–160, 2016.

[56] H. Y. Su, H. C. Lai, Y. W. Lin et al., "Epigenetic silencing of SFRP5 is related to malignant phenotype and chemoresistance of ovarian cancer through Wnt signaling pathway," *International Journal of Cancer*, vol. 127, no. 3, pp. 555–567, 2010.

[57] M. Alizadeh, S. Mahjoub, S. Esmaelzadeh, K. Hajian, Z. Basirat, and M. Ghasemi, "Evaluation of oxidative stress in endometriosis: a case-control study," *Caspian Journal of Internal Medicine*, vol. 6, no. 1, pp. 25–29, 2015.

[58] J. Donnez, M. M. Binda, O. Donnez, and M. M. Dolmans, "Oxidative stress in the pelvic cavity and its role in the pathogenesis of endometriosis," *Fertility and Sterility*, vol. 106, no. 5, pp. 1011–1017, 2016.

[59] A. Van Langendonckt, F. Casanas-Roux, and J. Donnez, "Oxidative stress and peritoneal endometriosis," *Fertility and Sterility*, vol. 77, no. 5, pp. 861–870, 2002.

[60] F. Ito, Y. Yamada, A. Shigemitsu et al., "Role of oxidative stress in epigenetic modification in endometriosis," *Reproductive Sciences*, vol. 24, no. 11, pp. 1493–1502, 2017.

[61] H. Xie, P. Chen, H. W. Huang, L. P. Liu, and F. Zhao, "Reactive oxygen species downregulate ARID1A expression via its promoter methylation during the pathogenesis of endometriosis," *European Review for Medical and Pharmacological Sciences*, vol. 21, no. 20, pp. 4509–4515, 2017.

[62] F. J. Roca, H. A. Loomans, A. T. Wittman, C. J. Creighton, and S. M. Hawkins, "Ten-eleven translocation genes are downregulated in endometriosis," *Current Molecular Medicine*, vol. 16, no. 3, pp. 288–298, 2016.

[63] A. V. Bhat, S. Hora, A. Pal, S. Jha, and R. Taneja, "Stressing the (epi)genome: dealing with reactive oxygen species in cancer," *Antioxidants & Redox Signaling*, 2017.

[64] J. Zhong, L. Ji, H. Chen et al., "Acetylation of hMOF modulates H4K16ac to regulate DNA repair genes in response to oxidative stress," *International Journal of Biological Sciences*, vol. 13, no. 7, pp. 923–934, 2017.

[65] X. Gao, X. Gào, Y. Zhang, L. P. Breitling, B. Schöttker, and H. Brenner, "Associations of self-reported smoking, cotinine levels and epigenetic smoking indicators with oxidative stress among older adults: a population-based study," *European Journal of Epidemiology*, vol. 32, no. 5, pp. 443–456, 2017.

[66] T. Maser, M. Rich, D. Hayes et al., "Tolcapone induces oxidative stress leading to apoptosis and inhibition of tumor growth in neuroblastoma," *Cancer Medicine*, vol. 6, no. 6, pp. 1341–1352, 2017.

[67] W. Li, Y. Guo, C. Zhang et al., "Dietary phytochemicals and cancer chemoprevention: a perspective on oxidative stress, inflammation, and epigenetics," *Chemical Research in Toxicology*, vol. 29, no. 12, pp. 2071–2095, 2016.

[68] R. Zhang, K. A. Kang, K. C. Kim et al., "Oxidative stress causes epigenetic alteration of CDX1 expression in colorectal cancer cells," *Gene*, vol. 524, no. 2, pp. 214–219, 2013.

[69] M. Hou, X. Zuo, C. Li, Y. Zhang, and Y. Teng, "miR-29b regulates oxidative stress by targeting SIRT1 in ovarian cancer cells," *Cellular Physiology and Biochemistry*, vol. 43, no. 5, pp. 1767–1776, 2017.

[70] P. K. Mahalingaiah, L. Ponnusamy, and K. P. Singh, "Oxidative stress-induced epigenetic changes associated with malignant transformation of human kidney epithelial cells," *Oncotarget*, vol. 8, no. 7, pp. 11127–11143, 2017.

[71] M. A. Dawson and T. Kouzarides, "Cancer epigenetics: from mechanism to therapy," *Cell*, vol. 150, no. 1, pp. 12–27, 2012.

[72] M. Esteller, "Cancer epigenomics: DNA methylomes and histone-modification maps," *Nature Reviews Genetics*, vol. 8, no. 4, pp. 286–298, 2007.

[73] D. Schübeler, D. M. MacAlpine, D. Scalzo et al., "The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote," *Genes & Development*, vol. 18, no. 11, pp. 1263–1271, 2004.

[74] C. Martin and Y. Zhang, "The diverse functions of histone lysine methylation," *Nature Reviews Molecular Cell Biology*, vol. 6, no. 11, pp. 838–849, 2005.

[75] H. A. LaVoie, "Epigenetic control of ovarian function: the emerging role of histone modifications," *Molecular and Cellular Endocrinology*, vol. 243, no. 1-2, pp. 12–18, 2005.

[76] D. J. Marsh, J. S. Shah, and A. J. Cole, "Histones and their modifications in ovarian cancer—drivers of disease and therapeutic targets," *Frontiers in Oncology*, vol. 4, p. 144, 2014.

[77] A. Hayashi, A. Horiuchi, N. Kikuchi et al., "Type-specific roles of histone deacetylase (HDAC) overexpression in ovarian carcinoma: HDAC1 enhances cell proliferation and HDAC3 stimulates cell migration with downregulation of e-cadherin," *International Journal of Cancer*, vol. 127, no. 6, pp. 1332–1346, 2010.

[78] W. Weichert, "HDAC expression and clinical prognosis in human malignancies," *Cancer Letters*, vol. 280, no. 2, pp. 168–176, 2009.

[79] S. Ropero and M. Esteller, "The role of histone deacetylases (HDACs) in human cancer," *Molecular Oncology*, vol. 1, no. 1, pp. 19–25, 2007.

[80] K. Y. Jang, K. S. Kim, S. H. Hwang et al., "Expression and prognostic significance of SIRT1 in ovarian epithelial tumours," *Pathology*, vol. 41, no. 4, pp. 366–371, 2009.

[81] X. Xiaomeng, Z. Ming, M. Jiezhi, and F. Xiaoling, "Aberrant histone acetylation and methylation levels in woman with endometriosis," *Archives of Gynecology and Obstetrics*, vol. 287, no. 3, pp. 487–494, 2013.

[82] M. Colón-Díaz, P. Báez-Vega, M. García et al., "HDAC1 and HDAC2 are differentially expressed in endometriosis," *Reproductive Sciences*, vol. 19, no. 5, pp. 483–492, 2012.

[83] J. Guo, J. Cai, L. Yu, H. Tang, C. Chen, and Z. Wang, "EZH2 regulates expression of p57 and contributes to progression of ovarian cancer *in vitro* and *in vivo*," *Cancer Science*, vol. 102, no. 3, pp. 530–539, 2011.

[84] Y. Kuang, F. Lu, J. Guo et al., "Histone demethylase KDM2B upregulates histone methyltransferase EZH2 expression and contributes to the progression of ovarian cancer in vitro and in vivo," *OncoTargets and Therapy*, vol. 10, pp. 3131–3144, 2017.

[85] P. H. Abbosh, J. S. Montgomery, J. A. Starkey et al., "Dominant-negative histone H3 lysine 27 mutant derepresses silenced tumor suppressor genes and reverses the drug-resistant phenotype in cancer cells," *Cancer Research*, vol. 66, no. 11, pp. 5582–5591, 2006.

[86] J. H. Hurst, N. Mendpara, and S. B. Hooks, "Regulator of G-protein signalling expression and function in ovarian cancer cell lines," *Cellular and Molecular Biology Letters*, vol. 14, no. 1, pp. 153–174, 2009.

[87] E. Cacan, "Epigenetic regulation of RGS2 (regulator of G-protein signaling 2) in chemoresistant ovarian cancer cells," *Journal of Chemotherapy*, vol. 29, no. 3, pp. 173–178, 2017.

[88] L. Schulke, M. Berbic, F. Manconi, N. Tokushige, R. Markham, and I. S. Fraser, "Dendritic cell populations in the eutopic and ectopic endometrium of women with endometriosis," *Human Reproduction*, vol. 24, no. 7, pp. 1695–1703, 2009.

[89] I. Jeung, K. Cheon, and M. R. Kim, "Decreased cytotoxicity of peripheral and peritoneal natural killer cell in endometriosis," *BioMed Research International*, vol. 2016, Article ID 2916070, 6 pages, 2016.

[90] E. Oral, D. L. Olive, and A. Arici, "The peritoneal environment in endometriosis," *Human Reproduction Update*, vol. 2, no. 5, pp. 385–398, 1996.

[91] S. Mahdian, R. Aflatoonian, R. S. Yazdi et al., "Macrophage migration inhibitory factor as a potential biomarker of endometriosis," *Fertility and Sterility*, vol. 103, no. 1, pp. 153–159.e3, 2015.

[92] C. M. Kyama, L. Overbergh, A. Mihalyi et al., "Endometrial and peritoneal expression of aromatase, cytokines, and adhesion factors in women with endometriosis," *Fertility and Sterility*, vol. 89, no. 2, pp. 301–310, 2008.

[93] D. Hornung, I. P. Ryan, V. A. Chao, J. L. Vigne, E. D. Schriock, and R. N. Taylor, "Immunolocalization and regulation of the chemokine RANTES in human endometrial and endometriosis tissues and cells," *The Journal of Clinical Endocrinology & Metabolism*, vol. 82, no. 5, pp. 1621–1628, 1997.

[94] X. J. Yang, J. Yang, Z. Liu, G. Yang, and Z. J. Shen, "Telocytes damage in endometriosis-affected rat oviduct and potential impact on fertility," *Journal of Cellular and Molecular Medicine*, vol. 19, no. 2, pp. 452–462, 2015.

[95] S. Yoshida, T. Harada, T. Iwabe et al., "A combination of interleukin-6 and its soluble receptor impairs sperm motility: implications in infertility associated with endometriosis," *Human Reproduction*, vol. 19, no. 8, pp. 1821–1825, 2004.

[96] E. Hosseini, F. Mehraein, M. Shahhoseini et al., "Epigenetic alterations of *CYP19A1* gene in *Cumulus* cells and its relevance to infertility in endometriosis," *Journal of Assisted Reproduction and Genetics*, vol. 33, no. 8, pp. 1105–1113, 2016.

[97] Y. Tao, Q. Zhang, W. Huang, H. L. Zhu, D. Zhang, and W. Luo, "The peritoneal leptin, MCP-1 and TNF-α in the pathogenesis of endometriosis-associated infertility," *American Journal of Reproductive Immunology*, vol. 65, no. 4, pp. 403–406, 2011.

[98] N. Rathore, A. Kriplani, R. K. Yadav, U. Jaiswal, and R. Netam, "Distinct peritoneal fluid ghrelin and leptin in infertile women with endometriosis and their correlation with interleukin-6 and vascular endothelial growth factor," *Gynecological Endocrinology*, vol. 30, no. 9, pp. 671–675, 2014.

[99] H. Kobayashi, Y. Higashiura, H. Shigetomi, and H. Kajihara, "Pathogenesis of endometriosis: the role of initial infection and subsequent sterile inflammation (review)," *Molecular Medicine Reports*, vol. 9, no. 1, pp. 9–15, 2014.

[100] B. D. McKinnon, D. Bertschi, N. A. Bersinger, and M. D. Mueller, "Inflammation and nerve fiber interaction in endometriotic pain," *Trends in Endocrinology & Metabolism*, vol. 26, no. 1, pp. 1–10, 2015.

[101] J. Sikora, M. Smycz-Kubańska, A. Mielczarek-Palacz, and Z. Kondera-Anasz, "Abnormal peritoneal regulation of chemokine activation—the role of IL-8 in pathogenesis of endometriosis," *American Journal of Reproductive Immunology*, vol. 77, no. 4, 2017.

[102] N. Malhotra, D. Karmakar, V. Tripathi, K. Luthra, and S. Kumar, "Correlation of angiogenic cytokines-leptin and IL-8 in stage, type and presentation of endometriosis," *Gynecological Endocrinology*, vol. 28, no. 3, pp. 224–227, 2012.

[103] M. Ulukus, E. C. Ulukus, E. N. Tavmergen Goker, E. Tavmergen, W. Zheng, and A. Arici, "Expression of interleukin-8 and monocyte chemotactic protein 1 in women with endometriosis," *Fertility and Sterility*, vol. 91, no. 3, pp. 687–693, 2009.

[104] J. L. Herington, K. L. Bruner-Tran, J. A. Lucas, and K. G. Osteen, "Immune interactions in endometriosis," *Expert Review of Clinical Immunology*, vol. 7, no. 5, pp. 611–626, 2011.

[105] H. Cakmak, O. Guzeloglu-Kayisli, U. A. Kayisli, and A. Arici, "Immune-endocrine interactions in endometriosis," *Frontiers in Bioscience (Elite Edition)*, vol. 1, pp. 429–443, 2009.

[106] H. R. Gatla, Y. Zou, M. M. Uddin, and I. Vancurova, "Epigenetic regulation of interleukin-8 expression by class I HDAC and CBP in ovarian cancer cells," *Oncotarget*, vol. 8, no. 41, pp. 70798–70810, 2017.

[107] H. R. Gatla, Y. Zou, M. M. Uddin et al., "Histone deacetylase (HDAC) inhibition induces IκB kinase (IKK)-dependent interleukin-8/CXCL8 expression in ovarian cancer cells," *Journal of Biological Chemistry*, vol. 292, no. 12, pp. 5043–5054, 2017.

[108] R. Schickel, B. Boyerinas, S. M. Park, and M. E. Peter, "MicroRNAs: key players in the immune system, differentiation, tumorigenesis and cell death," *Oncogene*, vol. 27, no. 45, pp. 5959–5974, 2008.

[109] J. D. Kuhlmann, J. Rasch, P. Wimberger, and S. Kasimir-Bauer, "MicroRNA and the pathogenesis of ovarian cancer—a new horizon for molecular diagnostics and treatment?," *Clinical Chemistry and Laboratory Medicine*, vol. 50, no. 4, pp. 601–615, 2012.

[110] R. C. Lee, R. L. Feinbaum, and V. Ambros, "The *C. elegans* heterochronic gene *lin*-4 encodes small RNAs with antisense complementarity to *lin*-14," *Cell*, vol. 75, no. 5, pp. 843–854, 1993.

[111] Y. Wang and R. Blelloch, "Cell cycle regulation by microRNAs in stem cells," *Results and Problems in Cell Differentiation*, vol. 53, pp. 459–472, 2011.

[112] M. Carleton, M. A. Cleary, and P. S. Linsley, "MicroRNAs and cell cycle regulation," *Cell Cycle*, vol. 6, no. 17, pp. 2127–2132, 2007.

[113] S. Oliveto, M. Mancino, N. Manfrini, and S. Biffo, "Role of microRNAs in translation regulation and cancer," *World Journal of Biological Chemistry*, vol. 8, no. 1, pp. 45–56, 2017.

[114] N. Li, B. Long, W. Han, S. Yuan, and K. Wang, "MicroRNAs: important regulators of stem cells," *Stem Cell Research & Therapy*, vol. 8, no. 1, p. 110, 2017.

[115] Y. Lee, M. Kim, J. Han et al., "MicroRNA genes are transcribed by RNA polymerase II," *The EMBO Journal*, vol. 23, no. 20, pp. 4051–4060, 2004.

[116] S. M. Hammond, "An overview of microRNAs," *Advanced Drug Delivery Reviews*, vol. 87, pp. 3–14, 2015.

[117] M. Ha and V. N. Kim, "Regulation of microRNA biogenesis," *Nature Reviews Molecular Cell Biology*, vol. 15, no. 8, pp. 509–524, 2014.

[118] G. Pampalakis, E. P. Diamandis, D. Katsaros, and G. Sotiropoulou, "Down-regulation of dicer expression in ovarian cancer tissues," *Clinical Biochemistry*, vol. 43, no. 3, pp. 324–327, 2010.

[119] W. M. Merritt, Y. G. Lin, L. Y. Han et al., "Dicer, Drosha, and outcomes in patients with ovarian cancer," *The New England Journal of Medicine*, vol. 359, no. 25, pp. 2641–2650, 2008.

[120] M. V. Iorio, R. Visone, G. di Leva et al., "MicroRNA signatures in human ovarian cancer," *Cancer Research*, vol. 67, no. 18, pp. 8699–8707, 2007.

[121] A. E. Pasquinelli, B. J. Reinhart, F. Slack et al., "Conservation of the sequence and temporal expression of *let*-7 heterochronic regulatory RNA," *Nature*, vol. 408, no. 6808, pp. 86–89, 2000.

[122] S. Shell, S. M. Park, A. R. Radjabi et al., "Let-7 expression defines two differentiation stages of cancer," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 27, pp. 11400–11405, 2007.

[123] N. Yang, S. Kaur, S. Volinia et al., "MicroRNA microarray identifies *Let-7i* as a novel biomarker and therapeutic target in human epithelial ovarian cancer," *Cancer Research*, vol. 68, no. 24, pp. 10307–10314, 2008.

[124] H. Lee, C. Park, G. Deftereos et al., "MicroRNA expression in ovarian carcinoma and its correlation with clinicopathological features," *World Journal of Surgical Oncology*, vol. 10, no. 1, p. 174, 2012.

[125] T. H. Kim, Y. K. Kim, Y. Kwon et al., "Deregulation of miR-519a, 153, and 485-5p and its clinicopathological relevance in ovarian epithelial tumours," *Histopathology*, vol. 57, no. 5, pp. 734–743, 2010.

[126] A. Mahdian-shakib, R. Dorostkar, M. Tat, M. S. Hashemzadeh, and N. Saidi, "Differential role of microRNAs in prognosis, diagnosis, and therapy of ovarian cancer," *Biomedicine & Pharmacotherapy*, vol. 84, pp. 592–600, 2016.

[127] B. Shen, Y. Jiang, Y.-R. Chen et al., "Expression and inhibitory role of TIMP-3 in hepatocellular carcinoma," *Oncology Reports*, vol. 36, no. 1, pp. 494–502, 2016.

[128] S. Qin, Y. Zhu, F. Ai et al., "MicroRNA-191 correlates with poor prognosis of colorectal carcinoma and plays multiple roles by targeting tissue inhibitor of metalloprotease 3," *Neoplasma*, vol. 61, no. 1, pp. 27–34, 2014.

[129] M. Dong, P. Yang, and F. Hua, "miR-191 modulates malignant transformation of endometriosis through regulating TIMP3," *Medical Science Monitor*, vol. 21, pp. 915–920, 2015.

[130] C. Y. Chang, M. T. Lai, Y. Chen et al., "Up-regulation of ribosome biogenesis by *MIR196A2* genetic variation promotes endometriosis development and progression," *Oncotarget*, vol. 7, no. 47, pp. 76713–76725, 2016.

[131] M. A. Nieto, R. Y.-J. Huang, R. A. Jackson, and J. P. Thiery, "EMT: 2016," *Cell*, vol. 166, no. 1, pp. 21–45, 2016.

[132] J. Bartley, A. Jülicher, B. Hotz, S. Mechsner, and H. Hotz, "Epithelial to mesenchymal transition (EMT) seems to be regulated differently in endometriosis and the endometrium," *Archives of Gynecology and Obstetrics*, vol. 289, no. 4, pp. 871–881, 2014.

[133] S. Matsuzaki and C. Darcha, "Epithelial to mesenchymal transition-like and mesenchymal to epithelial transition-like processes might be involved in the pathogenesis of pelvic endometriosis," *Human Reproduction*, vol. 27, no. 3, pp. 712–721, 2012.

[134] N. Filigheddu, I. Gregnanin, P. E. Porporato et al., "Differential expression of microRNAs between eutopic and ectopic endometrium in ovarian endometriosis," *Journal of Biomedicine and Biotechnology*, vol. 2010, Article ID 369549, 29 pages, 2010.

[135] J. C. Eggers, V. Martino, R. Reinbold et al., "MicroRNA miR-200b affects proliferation, invasiveness and stemness of endometriotic cells by targeting ZEB1, ZEB2 and KLF4," *Reproductive Biomedicine Online*, vol. 32, no. 4, pp. 434–445, 2016.

[136] T. Brabletz, A. Jung, S. Reu et al., "Variable β-catenin expression in colorectal cancers indicates tumor progression driven

by the tumor environment," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 18, pp. 10356–10361, 2001.

[137] Y. Lou, H. Jiang, Z. Cui, L. Wang, X. Wang, and T. Tian, "Linc-ROR induces epithelial-to-mesenchymal transition in ovarian cancer by increasing Wnt/β-catenin signaling," *Oncotarget*, vol. 8, no. 41, pp. 69983–69994, 2017.

[138] F. Ma, W. Li, C. Liu et al., "miR-23a promotes TGF-β1-induced EMT and tumor metastasis in breast cancer cells by directly targeting CDH1 and activating Wnt/β-catenin signaling," *Oncotarget*, vol. 8, no. 41, pp. 69538–69550, 2017.

[139] J.-J. Qiu, L.-C. Ye, J.-X. Ding et al., "Expression and clinical significance of estrogen-regulated long non-coding RNAs in estrogen receptor α-positive ovarian cancer progression," *Oncology Reports*, vol. 31, no. 4, pp. 1613–1622, 2014.

[140] Y. Wang, Y. Chen, and J. Fang, "Post-transcriptional and post-translational regulation of central carbon metabolic enzymes in cancer," *Anti-Cancer Agents in Medicinal Chemistry*, vol. 17, no. 11, pp. 1456–1465, 2017.

[141] R. Asadollahi, C. A. C. Hyde, and X. Y. Zhong, "Epigenetics of ovarian cancer: from the lab to the clinic," *Gynecologic Oncology*, vol. 118, no. 1, pp. 81–87, 2010.

[142] P.-R. Sun, S.-Z. Jia, H. Lin, J.-H. Leng, and J.-H. Lang, "Genome-wide profiling of long noncoding ribonucleic acid expression patterns in ovarian endometriosis by microarray," *Fertility and Sterility*, vol. 101, no. 4, pp. 1038–1046.e7, 2014.

[143] C. Zhou, T. Zhang, F. Liu et al., "The differential expression of mRNAs and long noncoding RNAs between ectopic and eutopic endometria provides new insights into adenomyosis," *Molecular BioSystems*, vol. 12, no. 2, pp. 362–370, 2016.

[144] H. Liu and J. H. Lang, "Is abnormal eutopic endometrium the cause of endometriosis? The role of eutopic endometrium in pathogenesis of endometriosis," *Medical Science Monitor*, vol. 17, no. 4, pp. RA92–RA99, 2011.

[145] M. Zhou, J. Fu, L. Xiao et al., "miR-196a overexpression activates the MEK/ERK signal and represses the progesterone receptor and decidualization in eutopic endometrium from women with endometriosis," *Human Reproduction*, vol. 31, no. 11, pp. 2598–2608, 2016.

[146] H. Jørgensen, A. S. Hill, M. T. Beste et al., "Peritoneal fluid cytokines related to endometriosis in patients evaluated for infertility," *Fertility and Sterility*, vol. 107, no. 5, pp. 1191–1199.e2, 2017.

[147] S. Vicente-Muñoz, I. Morcillo, L. Puchades-Carrasco, V. Payá, A. Pellicer, and A. Pineda-Lucena, "Pathophysiologic processes have an impact on the plasma metabolomic signature of endometriosis patients," *Fertility and Sterility*, vol. 106, no. 7, pp. 1733–1741.e1, 2016.