GMR The Original

Comprehensive microRNA Profiling Reveals Novel Regulatory Networks in Human Osteoarthritis Articular Cartilage MiRNA Regulatory Networks in Osteoarthritis

J. Rao[1], S. Yan[1], J. Wei[1], Y. Wang[2] and L. Meng[1,2]*

1Guangxi Orthopedic Hospital, Nanning, Guangxi 530012, China.

2 Chongqing Medical University, Chongqing 401331, China.

Corresponding Author: Lin Meng; Email: drmenglin@yeah.net

ABSTRACT

Knee osteoarthritis (KOA) is a highly prevalent degenerative joint disease. To investigate its molecular mechanisms, we performed high-throughput sequencing to compare miRNA expression profiles in articular cartilage from KOA patients and healthy controls. We identified differentially expressed miRNAs, predicted their target genes, and conducted functional enrichment analysis. Total RNA was extracted from both sample groups, and miRNA libraries were constructed for PE-150 sequencing on the HiSeq/MiSeq platform. Raw data were subjected to fastq quality control and kcUID deduplication, after which 18-30 nt short RNAs were retained for alignment against the human reference genome, known miRNAs were annotated via MiRBase, novel miRNAs were predicted using mREvo and mirdeep2, differential expression analysis was performed with DESeq2, target genes were predicted by combining miRanda and RNAhybrid, and subsequent GO and KEGG enrichment analyses were carried out. The sequencing data showed high quality with a clean read coverage over 90% and a Q30 score above 85%; a total of 701–783 known miRNAs and 64–85 novel miRNAs were annotated, among which 131 differentially expressed miRNAs were identified, and the target genes of these differentially expressed miRNAs displayed significant enrichment in GO functional categories and KEGG pathways. These miRNAs play a potential role in KOA progression.

Keywords: High-throughput sequencing, Knee Osteoarthritis; Regulatory networks, Articular Cartilage

INTRODUCTION

Knee osteoarthritis (KOA) is a highly prevalent degenerative joint disease with primary clinical features of recurrent knee pain and joint stiffness; as the condition progresses, it may cause joint swelling, deformity, and functional impairment, and severe cases can even lead to disability [1]. The pathological essence of KOA involves multiple biological processes such as the imbalance between chondrocyte proliferation and apoptosis, degradation of the extracellular matrix, and inflammatory responses.

Epidemiological data reveal a persistent global increase in the incidence of knee osteoarthritis (KOA), with elderly populations being particularly vulnerable. Specifically, the prevalence of KOA exceeds 50% in individuals over 60 years old and rises to over 80% among those aged 75 and above. Notably, in Western countries, KOA has become one of the leading causes of chronic disability in middle-aged and elderly groups [2]. The prevalence of KOA in China is equally severe, with not only a yearly increase in incidence but also a trend toward younger onset, and the disease imposes multiple burdens on patients, families, and society: patients often experience psychological issues like anxiety and depression, families bear heavy financial burdens for medical care, and society faces significant pressures from labor force loss and healthcare resource consumption. In-depth exploration of KOA pathogenesis and the development of effective prevention and treatment strategies have become a critical research priority.

High-throughput sequencing (HTS) technology is a genomics research method based on large-scale parallel sequencing with core steps including sample library preparation, single-molecule-level fragment amplification, high-throughput sequencing, and bioinformatics analysis; this technology enables comprehensive and systematic elucidation of molecular changes during disease progression and provides crucial evidence for studying disease mechanisms and developing precision therapies [3].*MicroRNAs* (*miRNAs*) are endogenous short non-coding RNAs that regulate gene expression through post-transcriptional mechanisms, and in the progression of knee osteoarthritis (KOA), they play a crucial role by influencing biological processes such as chondrocyte proliferation, differentiation, apoptosis, and metabolism. Existing studies are largely confined to small-sample validation of individual *miRNAs* and lack systematic comparative analysis of *miRNA* expression profiles between knee osteoarthritis (KOA) patients and healthy controls, so the reliability and comprehensiveness of these findings require further enhancement.

This study aims to systematically screen differentially expressed *miRNAs* between knee osteoarthritis (KOA) patients and healthy controls through high-throughput *miRNA* sequencing of articular cartilage tissue samples. By further predicting their target genes and conducting functional enrichment analysis, this study seeks to elucidate the regulatory networks and potential mechanisms of *miRNAs* involved in knee osteoarthritis (KOA) pathogenesis, and this research provides novel experimental evidence and theoretical support for screening early diagnostic biomarkers and developing targeted therapeutic strategies for KOA.

MATERIALS AND METHODS

Sample Collection.

All articular cartilage tissue samples used in this study were obtained from patients undergoing knee surgery at Guangxi Orthopedic Hospital. The knee osteoarthritis (KOA) group comprised 4 patients, all meeting the diagnostic criteria outlined in the Chinese Association of Traditional Chinese Medicine's "Diagnostic and Treatment Standards for Knee Osteoarthritis Combining Traditional Chinese and Western Medicine." Macroscopic evaluation during surgery confirmed that all KOA donor sites displayed overt degeneration, corresponding to Outerbridge grade III (characterized by fissuring and fragmentation). In contrast, cartilage tissues harvested from the trauma control group were confirmed to be normal and displayed no signs of degeneration (Outerbridge grade I). The control group (CK group) was composed of patients undergoing surgery for traumatic knee injuries, with all cases confirmed to be

free of osteoarthritis. No statistically significant differences existed between the two groups in baseline characteristics such as age and gender (P>0.05), ensuring comparability. The study protocol was reviewed and approved by the Ethics Committee of Guangxi Orthopedic Hospital. All patients and their families signed written informed consent forms before sample collection.

Small RNA Library Construction and Sequencing

Total RNA was extracted from the KOA group and CK group samples, which served as the experimental materials. Systematic quality control of the total RNA samples was performed, including: 1% agarose gel electrophoresis to assess RNA degradation and contamination, Nanodrop analysis for RNA purity (requiring OD260/280 ratio between 1.8 and 2.1), Qubit for precise quantification of RNA concentration, and Bioptic Qsep100 analysis for RNA integrity (requiring RIN ≥ 7.0) [4]. Qualified samples were used for small RNA library construction, following this workflow: Illumina small RNA sequencing-specific 3′ and 5′ adapters were ligated to RNA molecules, followed by reverse transcription to synthesize cDNA, and PCR amplification to enrich small RNA fragments. 8% denaturing PAGE electrophoresis was used to separate and recover 140-160 bp target fragments. Post-assembly quality control: Initial quantification via Qubit 3.0, library dilution to 1 ng/μl, insert size analysis using Agilent 2100 Bioanalyzer (requiring main peak within 140-160 bp). Upon confirming the expected distribution, precise library concentration determination (>2 nM) via qPCR. Libraries passing quality control were submitted to Wuhan Kangce Technology Co., Ltd. for PE-150 sequencing on the Illumina HiSeq/MiSeq platform (e.g., HiSeq/MiSeq series). A target data volume of 6 GB was set for selected samples. The resulting raw data were utilized for subsequent *miRNA* expression analysis.

Sequencing Data Processing

Raw sequencing data (raw reads) were subjected to quality control using fastp (version 0.23.0). Reads with 3′ and 5′ adapter sequences, reads shorter than 15 bp, low-quality reads (over 40% of bases below Q15 quality) and reads with more than 5 N bases were filtered to obtain high-quality clean data. The kcUID method was employed for molecular tag (UID) deduplication and error correction with identification and extraction of UID sequences within reads to eliminate redundancy induced by PCR amplification. Small RNAs (sRNAs) in the 18–30 nt length range were further filtered as targets for subsequent analysis.

miRNA Annotation and Novel miRNA Prediction

Quality-controlled sRNA sequences were aligned to the human reference genome (*Homo_sapiens GRC38*) using Bowtie to obtain their genomic positioning information. Subsequently, successfully aligned sRNAs were compared against the miRBase database to annotate known *miRNA* precursors (known_hairpin) and mature forms (known_mature) across samples and quantify their abundance. For sRNAs not matching known *miRNAs*, other non-coding RNA types (e.g., rRNA, tRNA, snRNA) and repetitive sequences were further excluded. Novel *miRNA* prediction was performed using mREvo and miRdeep2 software [5], with the number of predicted *miRNAs* counted and their length distribution and first base preferences analyzed.

Differentially Expressed miRNA Analysis.

To eliminate the impact of sequencing depth on expression level comparison, the expression level of *miRNA* was normalized using the Transcripts Per Million (TPM) method. The specific calculation

formula is TPM = (Readcount × 10^6) / Total *miRNA* reads[6], where Readcount represents the sequencing read count of a specific *miRNA* and Total *miRNA* reads denotes the sum of sequencing read counts of all *miRNAs* in the sample. Differential expression analysis was performed using DESeq2 software (version 1.20.0, R Foundation for Statistical Computing, Vienna, Austria). DESeq2 utilizes a negative binomial distribution model to model the count data and handle variability between biological replicates. The thresholds for statistically significant differential expression were prospectively set at $|\log_2(\text{Fold Change})| > 1$ and a False Discovery Rate (FDR) < 0.05. Results were visualized via a volcano plot, and the reliability of experimental findings was verified through inter-sample correlation analysis.

2.6 *Prediction of Differentially Expressed miRNA Target Genes.*

To reveal the potential regulatory functions of differentially expressed *miRNAs* in the system, this study adopted a "multi-tool combined prediction + intersection screening" strategy for target gene prediction. This study selected miRanda and RNA hybrid software, which are complementary in algorithmic principles and set their core discrimination thresholds, with miRanda using a binding free energy (Total Energy) ≤ -20kcal/mol as the screening criterion. RNA hybrid required its predicted results to meet statistical significance with a p-value < 0.01 and a minimum free energy (MFE) ≤ -28 kcal/mol. Finally, the intersection of the prediction results from the two tools was used to construct a high-confidence differential *miRNA* target gene set to enhance the reliability of subsequent functional analyses.

Enrichment Analysis of Differentially Expressed miRNA Target Genes

For the potential target gene sets of differentially expressed *miRNAs*, we performed Gene Ontology (GO) functional enrichment analysis and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis to systematically reveal the biological processes and signaling regulatory networks in which they participate. The GO enrichment analysis first mapped target genes to functional terms in the Gene Ontology database [7], counted the number of genes within each term and used the hypergeometric test formula [8]to compare differences in functional term distribution between the target gene set and the whole-genome background. This identified significantly enriched GO terms across the three major categories: biological process, cellular component, and molecular function.

For significantly enriched metabolic and signaling pathways among target genes, analysis was performed using the mainstream KEGG pathway database [9]. Similarly, the hypergeometric test formula was employed to assess enrichment significance of target genes within KEGG pathways, thereby inferring key pathophysiological processes potentially involved by differentially expressed *miRNAs*. All enrichment analysis results were evaluated using a false discovery rate (FDR) < 0.05 as the significance threshold.

RESULTS

Sequencing Data Analysis

Systematic quality control was conducted on RNA samples, and RNA quality analyses were performed via agarose gel electrophoresis, Nanodrop, Qubit, and Bioptic Qsep100. All samples passed the quality control and met the requirements for subsequent library preparation. Additionally, sequencing quality control processing was completed using fastp software, and the results are presented in Table 1. According to the data, all samples had an effective rate exceeding 90% [10]; Q20 and Q30 bases accounted

for over 90% and 85% of clean reads, respectively, and the GC content remained stable. Further analysis showed that the selected sRNAs were mainly distributed in the 18–30 nt length range with uniform base content distribution and no significant AT or GC segregation (exemplified by CK_1 with results shown in Figures 1 and 2). These findings confirm the high accuracy of sequencing data without obvious degradation or contamination. The total RNA extracted in this study has a reliable quality and fully meets the requirements of subsequent analysis.

TABLE 1: Summary Table of Sequencing Data and Quality Control

Sample	Raw Reads	Raw Bases (G)	Raw Q20 (%)	Raw Q30(%)	Raw GC (%)	Clean Reads	Clean Baes (G)	Clean Q20 (%)	Clean Q30 (%)	Clean GC (%)	Effective Rate (%)
CK8	35589780	5.34	91.63	86.98	58.14	35499596	1.92	99.1	97.02	48.64	99.75
CK_1	49927792	7.49	93.5	88.41	58.52	49848868	2.59	99.29	97.28	49.66	99.84
KOA_15	40858676	6.13	94.79	89.81	57.46	40798760	2.22	99.28	97.23	49.04	99.85
KOA_18	48247846	7.24	92.3	87.42	57.6	48168076	2.54	99.26	97.2	48.85	99.83
CK_5	51995388	7.8	94.4	89.39	58.65	51908402	2.65	99.25	97.19	49.47	99.83
KOA_16	45112588	6.77	90.35	85.91	59.04	45042722	2.32	99.27	97.24	49.38	99.85
KOA_17	46223212	6.93	95.03	90.16	57.99	46154002	2.35	99.29	97.29	49.56	99.85

Note：Sample: Sample Name.Raw Bases (G): Raw Off-machine Data Volume.Raw Q20 (%): Proportion of Bases with Quality Greater Than Q20 in Raw Off-machine Data.Raw Q30 (%): Proportion of Bases with Quality Greater Than Q30 in Raw Off-machine Data.Raw GC (%): Average Content of GC Bases in Raw Off-machine Data.Clean Reads: Quantity of Clean Reads After Quality Control.Clean Bases (G): Data Volume After Quality Control.Clean Q20 (%): Proportion of Bases with Quality Greater Than Q20 in Clean Reads.Clean Q30 (%): Proportion of Bases with Quality Greater Than Q30 in Clean Reads.Clean GC (%): Average Content of GC Bases in Clean Reads.Effective Rate (%): Proportion of Clean Reads in Raw Off-machine Reads.

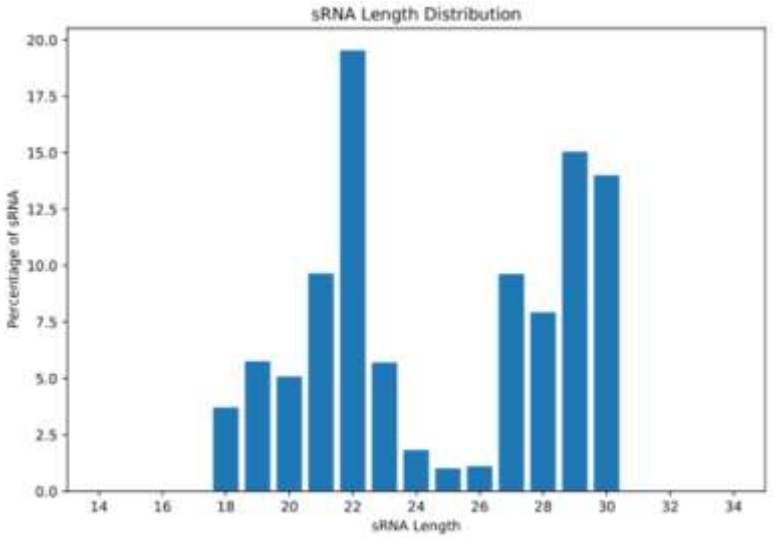

FIGURE1: Length Distribution of sRNA(CK_1)

Note：The horizontal axis represents the length of sRNA , and the vertical axis represents the percentage of sRNA with the corresponding length relative to the total sRNA.

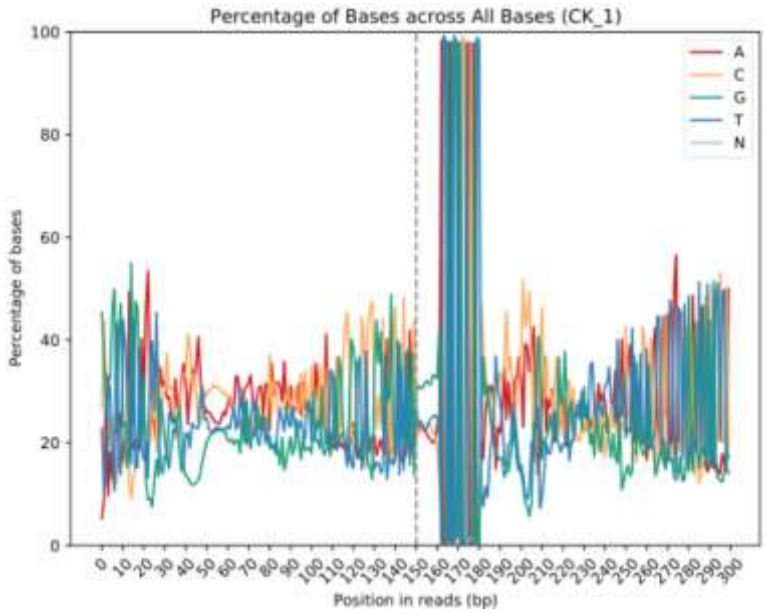

FIGURE 2: Base Content Distribution of Reads at Different Positions

Note: The horizontal axis represents the base positions on read1 and read2, and the vertical axis represents the proportion of the four bases (A, T, C, G) and the unknown base N at the corresponding position across all reads.

Prediction of Known and Novel miRNAs

Small RNA (sRNA) sequences preprocessed using fastp (v0.23.0) were aligned to the human reference genome via bowtie. The overall genome alignment efficiency ranged from 93.28% to 97.16% (Table 2), indicating high sequence align ability and good sample quality. Further alignment of genome-matched sRNAs against the miRBase database annotated known *miRNAs*. As shown in Table 3, the number of known *miRNA* precursors (known hairpin) ranged from 553 to 624 across samples, while known mature *miRNAs* (known mature) ranged from 701 to 783. The CK_5 sample contained the highest number of known *miRNAs*, while the KOA_15 sample contained the lowest. The total number of reads mapped to *miRNAs* (total_*miRNA*_reads) was generally higher in the control group than in the KOA group, exhibiting sample-specific expression patterns [11](Table 4). The distribution of first bases in known *miRNAs* showed distinct length-dependent preferences: for example, in CK_1, 18–19 nt *miRNAs* predominantly started with adenine (A), while 20–24 nt *miRNAs* predominantly featured uracil (U). At 25 nt, cytosine (C) became predominant and all 27 nt *miRNAs* exhibited adenine (A) as the first base (Figure 3). For sRNA sequences not matched to known *miRNAs*, novel *miRNA* prediction was performed using mREvo and miRDeep2 software. As shown in Table 5, the number of newly predicted mature *miRNAs* ranged from 64 to 85 across samples, with the highest count in CK_5 and the lowest in KOA_17. Expression patterns also exhibited sample-specificity (Table 6). The newly predicted *miRNAs* also exhibited significant first-base preferences: in the CK_1 sample, sequences of 18–21 nt predominantly featured cytosine (C), those of 22–23 nt predominantly featured uracil (U) and all 25nt sequences were entirely composed of uracil (U) [12](Figure 4). These base distribution patterns align with characteristics typical of Dicer enzyme processing products, indicating high reliability for the newly predicted *miRNAs*. Figure 5 uses the CK_1 sample to present the occurrence percentages of A, U, C, and G bases at each position of the novel *miRNA* from 1 to 22 nt, showing that the positional specificity of its base distribution conforms to the typical characteristics of *miRNA* and also verifies the reliability of the novel *miRNA* prediction results.

TABLE2: Summary Table of Alignment Information

Reads	CK_8	CK_1	KOA_15	KOA_18	CK_5	KOA_16	KOA_17
Clean_uniq_reads	284173	400929	285787	333713	469908	275240	373023
Clean_total_reads	4425556	7976907	7320258	8327389	7867758	7197726	7261297
Mapped_uniq_reads	235146	323240	239587	274276	387706	214069	292698
Mapped_total_reads	4300053	7608069	6828111	7920400	7511185	6886202	6876625
Mapping_rate	97.16%	95.38%	93.28%	95.11%	95.47%	95.67%	94.7%

Note: Clean_uniq_reads: Number of Unique Reads After Quality Control.Clean_total_reads: Total Number of Reads After Quality Control.Mapped_uniq_reads: Number of Mapped Unique Reads.Mapped_total_reads: Total Number of Reads Mapped to the Reference Genome.Mapping_rate: Percentage of Reads Mapped to the Reference Genome Relative to the Total Number of Reads.

TABLE 3: Summary Table of Known *miRNA* Alignment Results for Each Sample

Type	CK_8	CK_1	KOA_15	KOA_18	CK_5	KOA_16	KOA_17
known_hairpin	612	589	553	580	624	611	569
known_mature	773	748	701	722	783	764	718
total_*miRNA*_reads	2571516	2307070	1500918	2123767	2708760	1979767	1705021
uniq_*miRNA*_reads	3972	4182	3572	4020	4506	3939	3878

Note: known_hairpin: Refers to the *miRNA* precursors that are mapped. known mature: Refers to the mature *miRNAs* that are mapped. total_*miRNA*_reads: Refers to the number of small RNAs that are mapped.uniq_*miRNA*_reads: Refers to the types of small RNAs that are mapped.

TABLE4: Expression Profiles of Matched Known *miRNAs* (Partial)

miRNA	CK_8	CK_1	KOA_15	KOA_18	CK_5	KOA_16	KOA_17
hsa-miR-1468-5p	59	34	26	23	28	23	24
hsa-miR-502-5p	2	0	2	0	1	1	0
hsa-miR-5187-5p	22	6	5	1	5	2	6
hsa-miR-218-1-3p	0	0	0	2	1	1	1
hsa-miR-4712-5p	0	0	0	0	0	1	0

Note: Column 1: Mature *miRNA* ID.

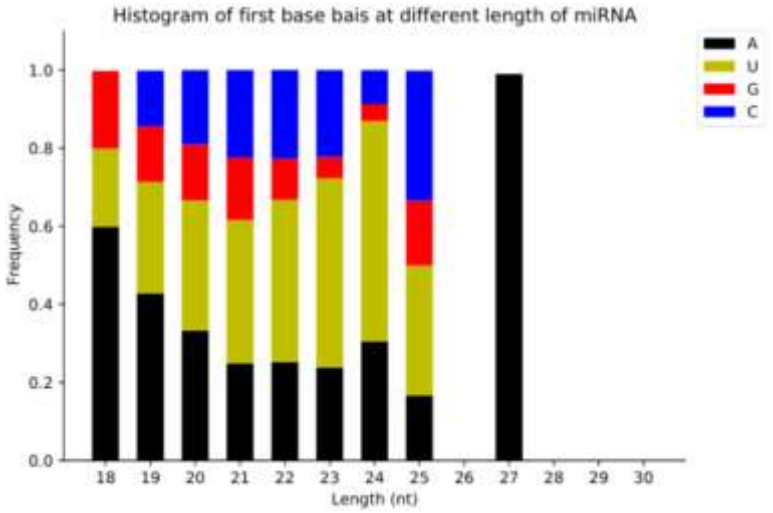

FIGURE 3: First-Base Preference of Known *miRNAs* with Lengths of 18~30 nt(CK_1)

Note: The horizontal axis represents the sequence length of *miRNA*, and the vertical axis represents the percentage of the occurrence of bases A/U/C/G at the first position in *miRNA*.

TABLE 5: Summary of Predicted Novel *miRNAs* & Per Sample sRNA Alignment to Them

Type	CK_8	CK_1	KOA_15	KOA_18	CK_5	KOA_16	KOA_17
mature_*miRNA*	69	67	66	69	85	65	64
total_*miRNA*	4555	7900	8295	51385	10732	4155	7543

Note: mature_*miRNA*: Refers to the number of predicted mature *miRNAs*. total_*miRNA*: Refers to the number of reads predicted to map to mature *miRNAs*.

TABLE 6: Expression Profiles of Novel *miRNAs* (Partial)

miRNA	CK_8	CK_1	KOA_15	KOA_18	CK_5	KOA_16	KOA_17
novel_1	350	330	259	458	302	324	365
novel 42	100	230	42	84	342	60	52
novel_29	31	21	20	27	40	16	22
novel_31	30	14	15	9	14	20	15

Note: Column 1: Predicted Mature *miRNA* ID.

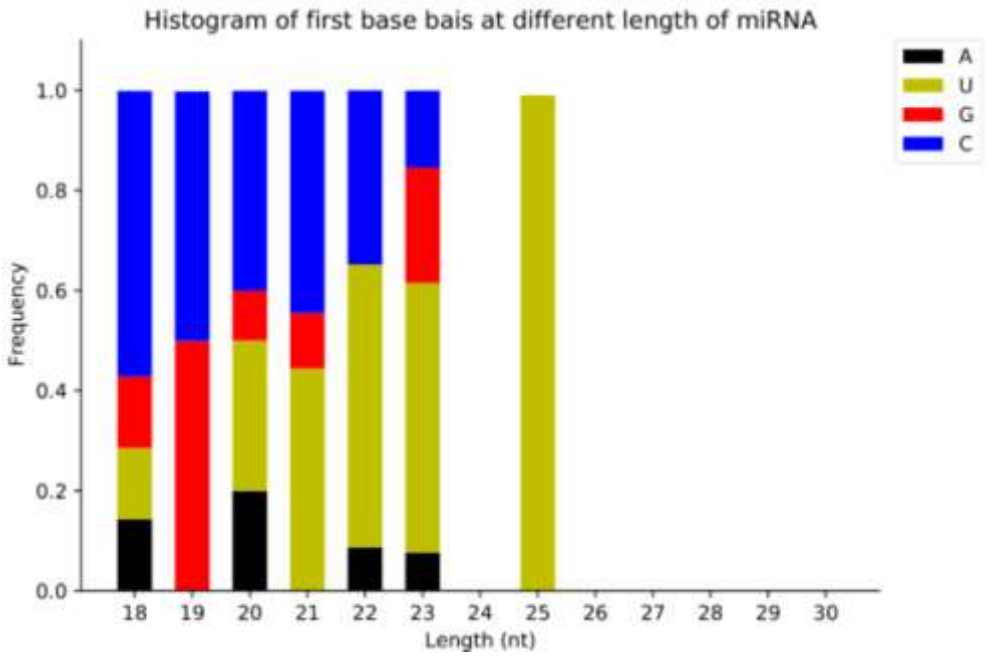

FIGURE4: First-Base Preference of Novel *miRNAs* with Lengths of 18~30 nt(CK_1)

Note: The horizontal axis represents the sequence length of *miRNA*, and the vertical axis represents the percentage of the occurrence of bases A/U/C/G at the first position in *miRNA*.

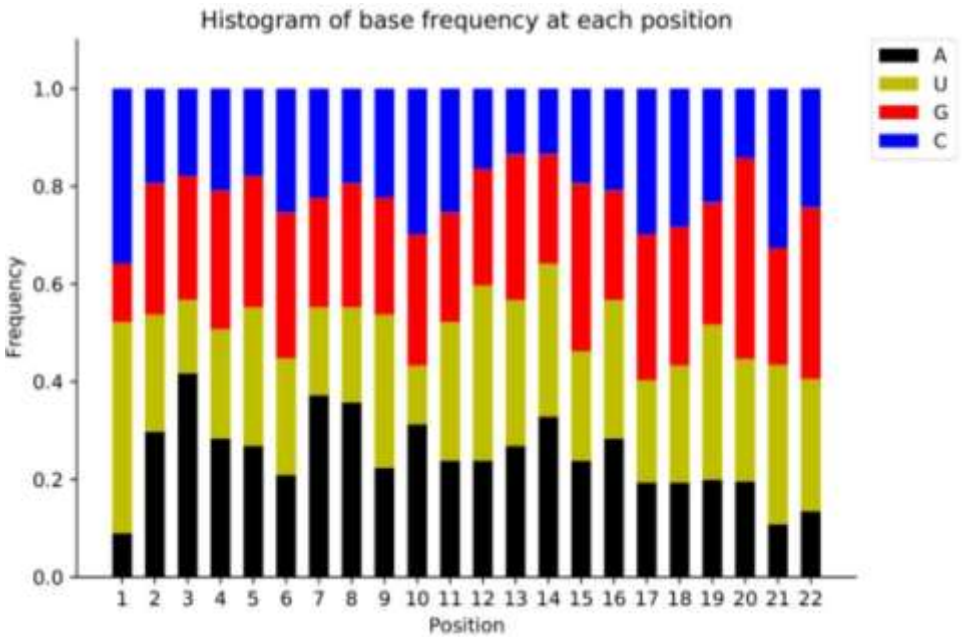

FIGURE5: Base Preference at Each Position of novel miRNA(CK_1)

Note: The horizontal axis represents the base position of miRNA, and the vertical axis represents the percentage of bases A/U/C/G appearing in the miRNA at that position.

Differential Expression miRNA Analysis

DESeq2 software was used to conduct differential expression analysis between the knee osteoarthritis (KOA) group and the control group based on TPM-normalized expression data of known and novel *miRNAs* [13]. A total of 131 miRNAs were identified as significantly differentially expressed, using the pre-specified thresholds of $|\log_2 FC| > 1$ and FDR < 0.05. Among these, 54 were significantly up-regulated and 77 were significantly down-regulated [14](Figure 6). The *miRNA* expression scatter plot of the comparison group shows that non-differential *miRNAs* are concentrated near the diagonal, while 77 down-regulated and 54 up-regulated differential *miRNAs* deviate from it. This pattern reflects the difference in *miRNA* expression between the KOA group and the control group. The results are presented in Figure 7. These findings indicate widespread dysregulation of *miRNAs* in knee osteoarthritis (KOA) articular cartilage tissue to further assess experimental reliability and intergroup consistency inter-sample correlation analysis was performed with results showing that correlation coefficients among control samples generally exceeded 0.90 while those in the KOA group ranged between 0.85 and 0.90.No significant outlier samples were identified (Figure 8) confirming excellent data quality and reliable reproducibility, which establishes a solid foundation for subsequent target gene identification and functional analysis. Hierarchical clustering analysis was performed using the TPM-normalized expression levels of 131 differential *miRNAs* between the KOA group and the control group. The results showed that samples from the control group and the KOA group formed independent clusters according to their groups with tight aggregation within the same group; The 131 differential *miRNAs* were divided into two clusters, where the "KOA high-expression cluster" showed warm-color high expression in the KOA group and cool-color low expression in the control group, while the "KOA low-expression cluster" showed the opposite expression trend. This result confirms the expression difference between the two groups and the repeatability of samples within the same group(Figure 9). Notably, the observed down-regulation of hsa-miR-1468-5p in KOA cartilage is of particular interest, as this *miRNA* is postulated to target *DUSP6*, a key modulator of the MAPK pathway. Since decreased *DUSP6* activity is linked to increased *MMP13* production, our data suggest that the loss of this *miRNA* contributes to cartilage degradation via the derepression of the *DUSP6*/*MMP13* pathway. We propose a novel model of cartilage degradation in KOA mediated by the *hsa-miR-1468-5p*/*DUSP6*/*MMP13* regulatory axis.

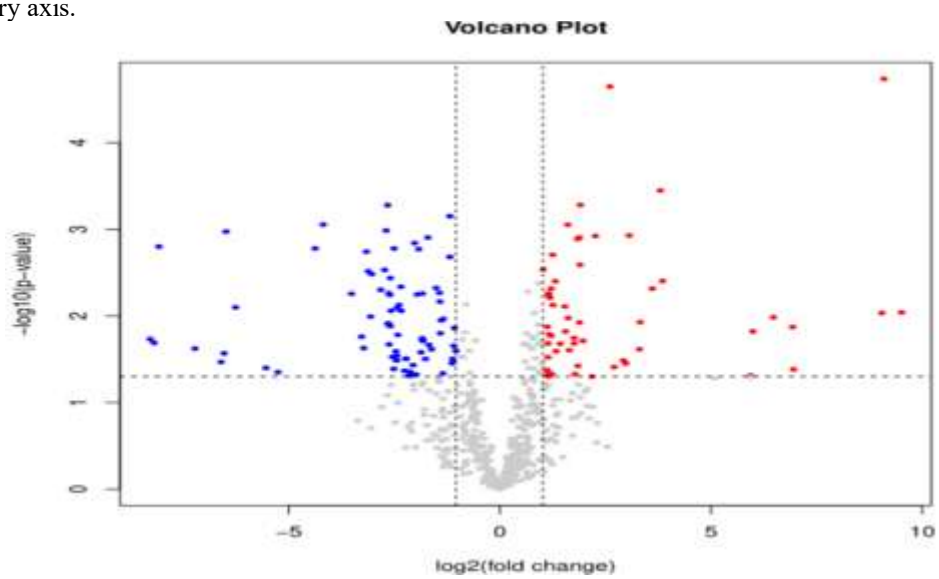

FIGURE 6: Volcano Plot of *miRNA* Expression

Note: Gray dots represent *miRNAs* with no significant differential expression, blue dots represent significantly down-regulated *miRNAs*, and red dots represent significantly up-regulated *miRNAs*.

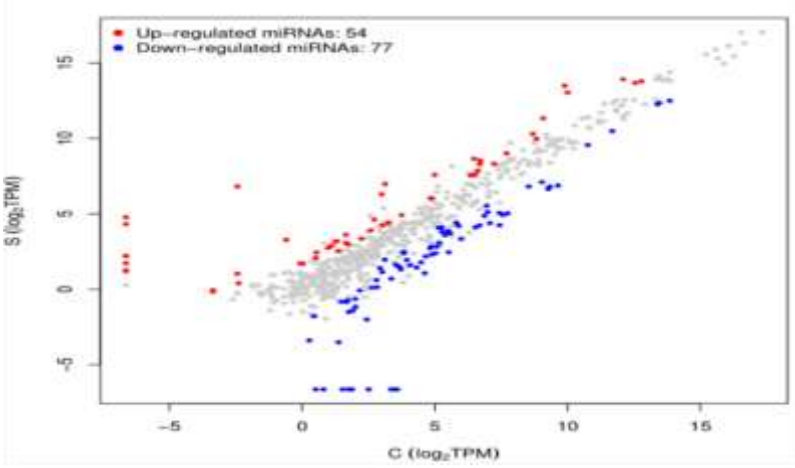

FIGURE 7: Expression Scatter Plot Between Comparison Groups
Note: Gray dots represent *miRNAs* with no significant differential expression.

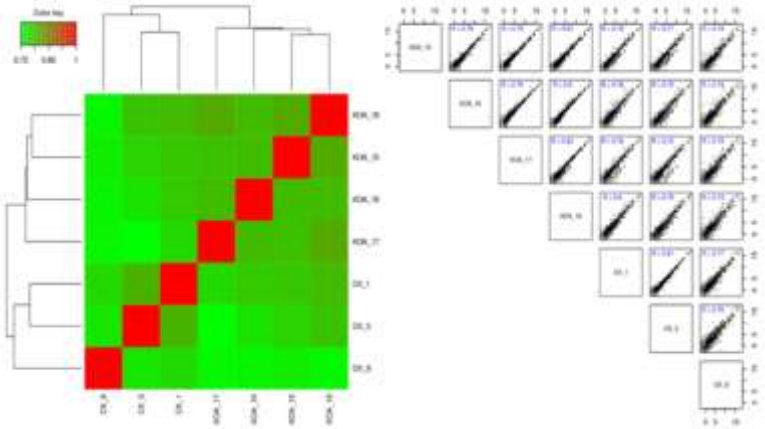

FIGURE 8: Correlation of *miRNA* Expression Levels between samples
Note: （A）Sample Correlation Heatmap.(B) Sample TPM Correlation.

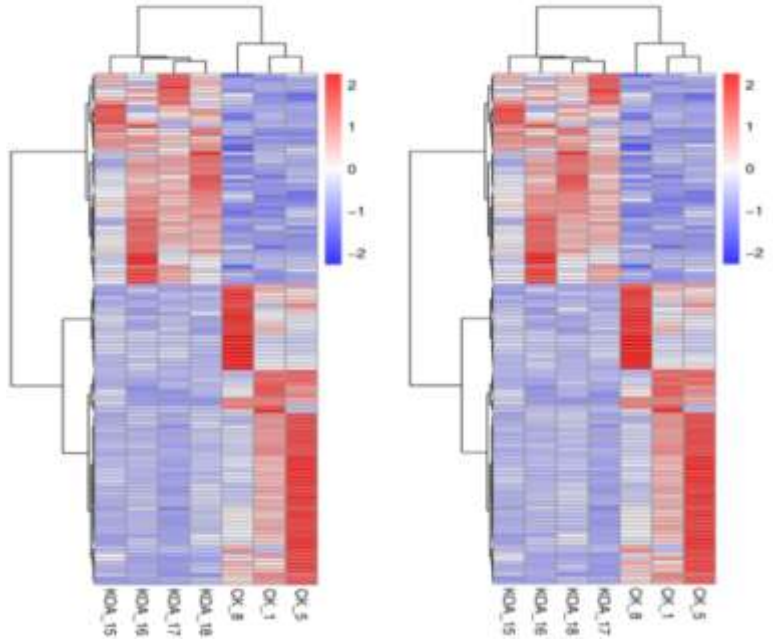

FIGURE 9: Clustering of differential *miRNA* expression levels

Note: （A）Clustered Heatmap of Differentially Expressed *miRNA* Expression in All Samples.（B）Clustered Heatmap of Differentially Expressed *miRNA* Expression between KOA and CK Groups.

miRNA Target Gene Prediction

To identify potential targets of differentially expressed *miRNAs,* this study used both miRanda and RNAhybrid for target gene prediction among 131 significantly differentially expressed *miRNAs*. miRanda prediction results (Table 7) revealed low binding free energies between differentially expressed *miRNAs* and target genes, indicating high binding stability [15]. Target gene lengths ranged from 2998 to 8082 bp, while *miRNA* lengths were concentrated between 21 and 22 nt, consistent with typical *miRNA*-target interaction characteristics [16].RNAhybrid predictions further indicated (Table 8) that most *miRNA*-target pairs exhibited binding free energies below -30 kcal/mol with statistical significance ($p < 0.01$).To enhance prediction reliability, intersection analysis of the two tools' results was performed. RNAhybrid predicted 1,394,043 potential target genes, while miRanda predicted 684,193. A total of 162,501 target genes were predicted by both tools (Figure 10). This intersecting target gene set underwent multi-algorithm validation, showing high credibility and providing a reliable data foundation for subsequent *miRNA* regulatory mechanism analysis and functional enrichment studies [17].

TABLE7: Summary Table of Predicted *miRNA* Target Genes Based on Miranda (Partial)

MiRNA	Target_ Gene	Tot_Score	Tot_E nergy	Max_S core	Max_Energy	*MiRNA* _length	Target_gene_length	Positions
hsa-miR-1468-5p	*AAK1*	151	-24.57	151	-24.57	21	8082	2444
hsa-miR-1468-5p	*FAM129C*	145	-20.14	145	-20.14	21	3373	1671
hsa-miR-1468-5p	*RIMBP2*	170	-23.97	170	-23.97	21	2998	1364
hsa-miR-1468-5p	*CINP*	161	-24.23	161	-24.23	21	3566	2137

Note:*miRNA*: *miRNA* ID. Tot_Score: Total Matching Score.Tot_Energy: Total Matching Free Energy.Max_Score: Maximum Matching Score.Max_Energy: Maximum Matching Free Energy.*miRNA*_length: *miRNA* Length.Target_gene_length: Target Gene Length.Positions: Matching Start Positions.

TABLE8: Summary Table of Predicted *miRNA* Target Genes Based on RNAhybrid (Partial)

Target _Gene	Gene_ length	*miRNA*	*miRNA*_length	MFE	pvalue	Positions
RAET1E	1112	*hsa-miR-1468-5p*	21	-31.9	0.03	95
BCAS3	818	*hsa-miR-1468-5p*	21	-33.2	0.01	422
TINAGL1	707	*hsa-miR-1468-5p*	21	-30.7	0.02	171
ATG4D	351	*hsa-miR-1468-5p*	21	-28.2	0.03	188

Note:Target_Gene: Target Gene.Gene_length: Target Gene Length.*miRNA*: *miRNA* ID.*miRNA*_length: *miRNA* Length.MFE: Minimum Free Energy.pvalue: Statistical Significance Level.Positions: Matching Start Positions.

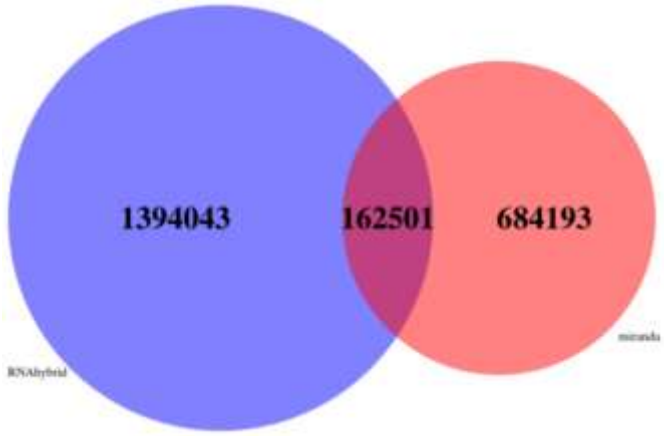

FIGURE 10: Venn Diagram of *miRNA* Predicted Target Genes

Note： The blue ellipse represents the number of target genes predicted by the RNAhybrid algorithm, and the red ellipse represents the number of target genes predicted by the Miranda algorithm.

Differential miRNA Target Gene Enrichment Analysis

To clarify the biological functions of differentially expressed *miRNAs* and their potential roles in the pathogenesis of knee osteoarthritis (KOA), this study performed Gene Ontology (GO) functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis on the selected high-confidence target gene sets, with the results visualized in Figure 11 (GO enrichment) and Figure 12 (KEGG pathways) [18]. Gene Ontology (GO) enrichment analysis revealed significant enrichment of target genes in three major categories: Biological Process (BP) involving fundamental life activities like cellular processes and biological regulation, Cellular Component (CC) featuring structures such as cell membranes and organelles, and Molecular Function (MF) encompassing functions including protein binding [19-20]. These findings indicate that differentially expressed *miRNAs* may participate in the occurrence and progression of knee osteoarthritis (KOA) by regulating the aforementioned fundamental cellular activities and molecular functions, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis further shows significant enrichment of target genes in pathways including the PI3K-Akt signaling pathway, mTOR signaling pathway, cancer pathways, signal transduction pathways, and neuroactive ligand-receptor interactions, suggesting these pathways may play crucial roles in the KOA process [21]. In summary, this

study identified potential target genes of KOA-associated differentially expressed *miRNAs* and their involvement in biological processes and signaling pathways via systematic GO and KEGG enrichment analyses, providing new insights into the molecular mechanisms underlying KOA.

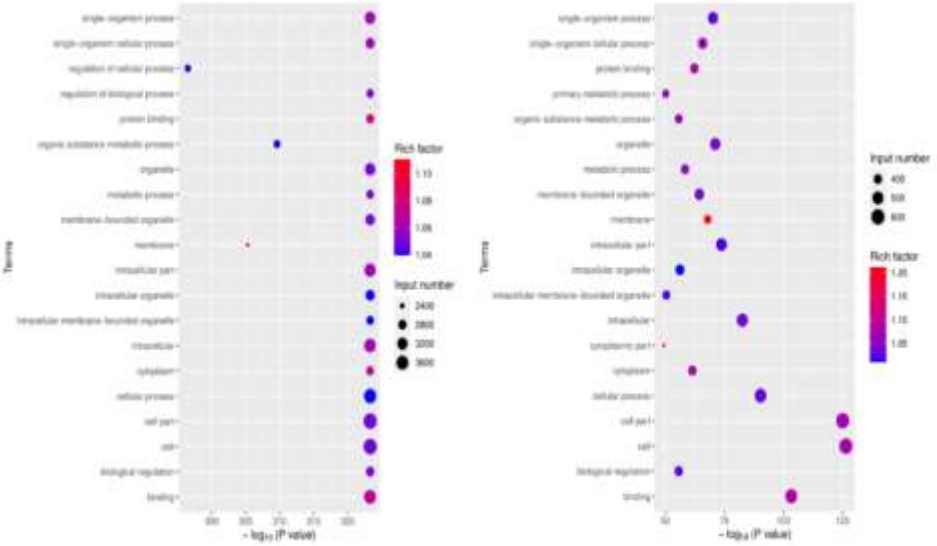

FIGURE 11: GO Enrichment Plot of Upregulated/Downregulated Genes
Note: The horizontal axis represents the significance of enrichment, and the vertical axis represents the enrichment GO terms. (A) GO enrichment analysis of down-regulated genes. (B) GO enrichment analysis of up-regulated genes.

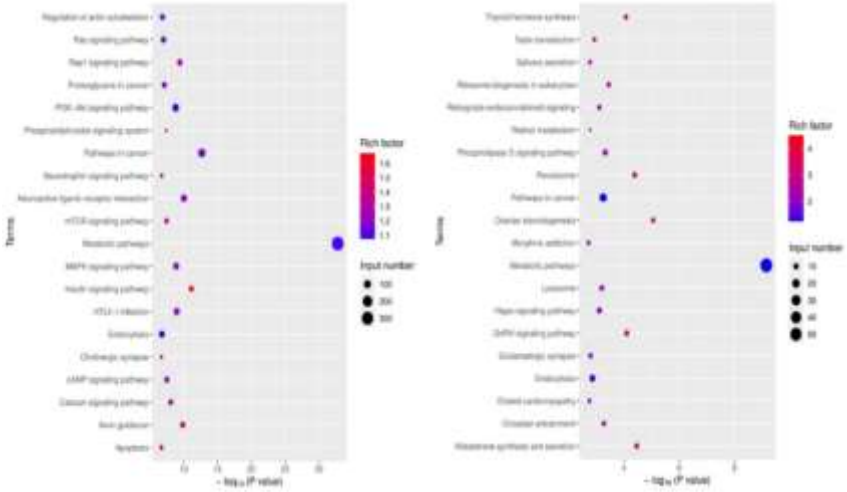

FIGURE 12: KEGG Metabolic Pathway Diagram of Upregulated/Downregulated Genes
Note: The horizontal axis represents the significance of enrichment, and the vertical axis represents the enriched KEGG pathways. （A）KEGG Metabolic Pathway Map of Downregulated Genes. （B）KEGG Metabolic Pathway Map of Upregulated Genes.

DISCUSSION

This study used human knee osteoarthritis (KOA) articular cartilage tissues and normal control articular cartilage tissues as research subjects, adopted the UID *miRNA*-seq library preparation strategy combined with PE-150 sequencing, obtained a minimum data volume of 6 GB per sample [22], and identified a total of

701–783 known *miRNAs* and 64–85 novel *miRNAs* through systematic analysis. This study obtained 162,501 high-confidence target genes via combined prediction with RNA hybrid and miRanda tools and subsequent cross-validation, and further GO and KEGG enrichment analysis of these target genes showed significant enrichment in biological processes closely related to KOA pathology, including cell proliferation regulation and inflammatory response. KEGG analysis showed these target genes were mainly enriched in pathways including the PI3K-Akt signaling pathway, mTOR signaling pathway, and metabolic pathways. These pathways are closely associated with chondrocyte survival, proliferation, autophagy, and inflammatory responses, thus forming the primary focus of discussion in this study. Multiple studies have indicated that the PI3K-Akt pathway, as a core signaling axis in the pathological process of KOA, exhibits abnormal activation. This abnormal activation is closely associated with cartilage degradation, bone dysfunction, and synovial inflammation, and the pathway itself also serves as a potential therapeutic target for KOA [23]. The findings of this study further confirm the crucial role of the mTOR signaling pathway in the occurrence and progression of KOA. Existing evidence reveals two key mechanisms: first, excessive activation of the mTOR pathway suppresses chondrocyte autophagy and promotes the expression of matrix-degrading enzymes (e.g., matrix metalloproteinase-3, MMP-3), thereby exacerbating cartilage destruction [24]; second, inhibiting the mTORC1 pathway reduces the secretion of matrix metalloproteinase-13 (MMP-13) and a disintegrin and metalloproteinase with thrombospondin motifs-5 *(ADAMTS-5*). Moreover, this inhibition can significantly delay the progression of joint structural damage in animal models by activating the autophagy pathway to clear abnormal protein aggregates [25]. The association between metabolic pathways and KOA is complex and multidimensional. Its primary mechanisms are manifested in that metabolic products can directly or indirectly induce changes in epigenetic modifications—including DNA methylation, histone acetylation, and methylation—to regulate the transcription of KOA-related genes. For instance, acetyl-CoA derived from fatty acid oxidation (FAO) can act as a substrate for histone acetylation, thereby activating the expression of catabolic genes such as *MMP13* [26-27]. Research has also revealed that aging and metabolic abnormalities may lead to mitochondrial dysfunction and reactive oxygen species (ROS) accumulation to further activate inflammasomes and exacerbate chondrocyte damage [28].

This study employed strict sequencing quality control procedures and a multi-software analysis strategy to identify differentially expressed *miRNAs* with potential functional roles and their target genes in knee osteoarthritis (KOA) cartilage tissue, preliminarily clarified the molecular pathways involved in these *miRNAs*, and provided theoretical foundations and data support for the development of diagnostic biomarkers and targeted therapies for KOA. The knee osteoarthritis (KOA) *miRNA* expression profile (comprising 131 *miRNAs*) revealed in this study indicates that the pathogenesis of KOA is driven by a complex regulatory network, updates the previous understanding of the role of a single molecule, and provides a valuable reservoir of candidate molecules for the development of novel diagnostic biomarkers and targeted therapeutic strategies.

This study has certain limitations: The most notable limitation is the relatively small sample size, which may constrain the statistical power and generalizability of the findings. The differential *miRNAs* identified in this study based on high-throughput sequencing have not been independently validated through quantitative real-time polymerase chain reaction (qRT-PCR) or functional experiments. This study did not provide visual quality control evidence like RNA electrophoresis images and library quality inspection graphs, although the clearly documented quality control standards indicated all samples and data met the required quality criteria. The results of this preliminary screening study provide high-confidence candidate molecules and a theoretical basis for subsequent verification. Future work will expand the sample size, incorporate biological replicates, and validate the regulatory relationships between differentially expressed *miRNAs* and their target genes through functional experiments. These efforts will deepen the understanding of their molecular functions and provide new strategies for the precision treatment of KOA.

CONCLUSION

This study constructs an *miRNA*-mediated regulatory network in knee osteoarthritis (KOA), which comprises 131 differentially expressed *miRNAs*, their target genes, and key signaling pathways such as PI3K-Akt and mTOR. This network reveals the potential regulatory mechanisms of *miRNAs* in knee osteoarthritis (KOA) (e.g., the *hsa-miR-1468-5p*/*DUSP6*/*MMP13* axis) and provides novel molecular targets and pathways as well as a theoretical basis for disease diagnosis and targeted therapy.

Acknowledgments

Not applicable

Data Availability Statement

The datasets generated and analyzed during the current study are not deposited in a public repository, but all key results are comprehensively illustrated in the figures of this manuscript. The corresponding author is available to provide additional details related to the data presented in the figures upon request.

Ethics approval and consent to participate

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Guangxi Orthopedic Hospital. Written informed consent was obtained from all individual participants included in the study. For participants who were minors or incapacitated adults, written informed consent was provided by their parents or legal guardians.

Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

Author Contributions

Lin Meng led the project's conceptualization, secured funding, provided supervision, and contributed to writing-review & editing. Junjie Rao performed formal analysis and wrote the original draft. Shidu Yan, Jianjie Wei, and Yangyang Wang were involved in data curation, investigation, methodology development, and visualization. All authors have read and approved the final version.

Funding

The work is supported by the General Program of Chongqing Natural Science Foundation（CSTB2022NSCQ-MSX0076）and the General Program of Guangxi Natural Science Foundation（2025GXNSFAA069213）.

References

1. Pan J, Xie Z, Shen H, Huang Z, Zhang X, Liao B. The effect of mild to moderate knee osteoarthritis on gait and three-dimensional biomechanical alterations. *Front Bioeng Biotechnol.* 2025 ; 13:1562936. doi: 10.3389/fbioe.2025.1562936.
2. Wang R, Xiang J, Ren M, Lin J. Temporal trends of knee osteoarthritis prevalence over a 7-year period in Chinese adults: findings from the CHARLS study 2011-2018. *Front Public Health*. 2025;13:1593859. doi: 10.3389/fpubh.2025.1593859.
3. Füllgrabe J, Gosal WS, Creed P, Harding NJ, Gandelman O, Golder P, et al. Simultaneous sequencing of genetic and epigenetic bases in DNA. *Nat Biotechnol*. 2023;41(10):1457-1464. doi: 10.1038/s41587-022-01652-0.
4. Ziegenhain C, Hendriks GJ, Hagemann-Jensen M, Sandberg R. Molecular spikes: a gold standard for single-cell RNA counting. *Nat Methods*. 2022;19(5):560-566. doi: 10.1038/s41592-022-01446-x.
5. Wen M, Shen Y, Shi S, Tang T. miREvo: an integrative microRNA evolutionary analysis platform for next-generation sequencing experiments. *BMC Bioinformatics*. 2012;13:140. doi: 10.1186/1471-2105-13-140.

6. Abrams ZB, Johnson TS, Huang K, Payne PRO, Coombes K. A protocol to evaluate RNA sequencing normalization methods. *BMC Bioinformatics*. 2019 ; 20(Suppl 24):679. doi: 10.1186/s12859-019-3247-x.
7. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*. 2021 ; 2(3):100141. doi: 10.1016/j.xinn.2021.100141.
8. Cao J, Zhang S. A Bayesian extension of the hypergeometric test for functional enrichment analysis. *Biometrics*. 2014 ; 70(1):84-94. doi: 10.1111/biom.12122. Epub 2013 Dec 9.
9. Liu X, Wang W, Zhang X, Liang J, Feng D, Li Y, et al. Metabolism pathway-based subtyping in endometrial cancer: An integrated study by multi-omics analysis and machine learning algorithms. *Mol Ther Nucleic Acids*. 2024 ; 35(2):102155. doi: 10.1016/j.omtn.2024.102155.
10. Chen S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *Imeta*. 2023 ; 2(2):e107. doi: 10.1002/imt2.107.
11. Giraldez MD, Spengler RM, Etheridge A, Godoy PM, Barczak AJ, Srinivasan S, et al. Comprehensive multi-center assessment of small RNA-seq methods for quantitative miRNA profiling. *Nat Biotechnol*. 2018 ; 36(8):746-57. doi: 10.1038/nbt.4183.
12. Lee YY, Kim H, Kim VN. Sequence determinant of small RNA production by DICER. *Nature*. 2023 ; 615(7951):323-30. doi: 10.1038/s41586-023-05722-4.
13. Ali SA, Espin-Garcia O, Wong AK, Potla P, Pastrello C, McIntyre M, et al. Circulating microRNAs differentiate fast-progressing from slow-progressing and non-progressing knee osteoarthritis in the Osteoarthritis Initiative cohort. *Ther Adv Musculoskelet Dis*. 2022 ; 14:1759720X221082917. doi: 10.1177/1759720X221082917.
14. Burke H, Cellura D, Freeman A, Hicks A, Ostridge K, Watson A, et al. Pulmonary EV miRNA profiles identify disease and distinct inflammatory endotypes in COPD. *Front Med (Lausanne)*. 2022 ; 9:1039702. doi: 10.3389/fmed.2022.1039702.
15. Jacinta-Fernandes A, Xavier JM, Magno R, Lage JG, Maia AT. Allele-specific miRNA-binding analysis identifies candidate target genes for breast cancer risk. *NPJ Genom Med*. 2020 ; 5:4. doi: 10.1038/s41525-019-0112-9.
16. Chen W, Wu C, Li Y, Wang T, Huang M, Wang M, et al. Mir-483-5p-mediated activating of the IGF2/H19 enhancer up-regulates IGF2/H19 expression via chromatin loops to promote the malignant progression of hepatocellular carcinoma. *Mol Cancer*. 2025 ; 24(1):10. doi: 10.1186/s12943-024-02204-7.
17. Zhu H, Chang M, Wang Q, Chen J, Liu D, He W. Identifying the Potential of miRNAs in Houttuynia cordata-Derived Exosome-Like Nanoparticles Against Respiratory RNA Viruses. *Int J Nanomedicine*. 2023 ; 18:5983-6000. doi: 10.2147/IJN.S425173.
18. Bose B, Moravec M, Bozdag S. Computing microRNA-gene interaction networks in pan-cancer using miRDriver. *Sci Rep*. 2022;12(1):3717. doi: 10.1038/s41598-022-07628-z.
19. Xu W, Liu J, Qi H, Si R, Zhao Z, Tao Z, et al. A lineage-resolved cartography of microRNA promoter activity in C. elegans empowers multidimensional developmental analysis. *Nat Commun*. 2024;15(1):2783. doi: 10.1038/s41467-024-47055-4.
20. Li J, Ma X, Lin H, Zhao S, Li B, Huang Y, et al. MHIF-MSEA: a novel model of miRNA set enrichment analysis based on multi-source heterogeneous information fusion. *Front Genet*. 2024;15:1375148. doi: 10.3389/fgene.2024.1375148.
21. Cai J, Chen T, Jiang Z, Yan J, Ye Z, Ruan Y, et al. Bulk and single-cell transcriptome profiling reveal extracellular matrix mechanical regulation of lipid metabolism reprogramming through *YAP*/*TEAD4*/*ACADL* axis in hepatocellular carcinoma. *Int J Biol Sci*. 2023;19(7):2114-31. doi: 10.7150/ijbs.82177.
22. Ning Y, Zhang F, Li S, Wang C, Wu Y, Chen S, et al. Integrative analysis of miRNA in cartilage-derived extracellular vesicles and single-cell RNA-seq profiles in knee osteoarthritis. *Arch Biochem Biophys*. 2023;748:109785. doi: 10.1016/j.abb.2023.109785.

23. Liu DR, Liao TY, Wei YB, Fang Y, Wang PM, Mao J, et al. Study on the mechanism of Yiceng Tietie in alleviating TGF-β1-induced synovial fibrosis in rats with knee osteoarthritis based on the PI3K/AKT/HIF-1α signaling pathway.*J Nanjing Univ Tradit Chin Med.* 2023; 39(8): 738-745. DOI:10.14148/j.issn.1672-0482.2023.0738.
24. Wang H, Zhang H, Sun Q, Yang J, Zeng C, Ding C, et al. Chondrocyte mTORC1 activation stimulates miR-483-5p via HDAC4 in osteoarthritis progression. *J Cell Physiol*. 2019;234(3):2730-40.doi: 10.1002/jcp.27088.
25. Caramés B, Hasegawa A, Taniguchi N, Miyaki S, Blanco FJ, Lotz M. Autophagy activation by rapamycin reduces the severity of experimental osteoarthritis. *Ann Rheum Dis*. 2012; 71(4):575-81. doi: 10.1136/annrheumdis-2011-200557.
26. Ziogas A, Novakovic B, Ventriglia L, Galang N, Tran KA, Li W, et al. Long-term histone lactylation connects metabolic and epigenetic rewiring in innate immune memory. *Cell*. 2025 ; 188(11):2992-3012.e16.doi: 10.1016/j.cell.2025.03.048.
27. Mei Z, Yilamu K, Ni W, Shen P, Pan N, Chen H, et al. Chondrocyte fatty acid oxidation drives osteoarthritis via *SOX9* degradation and epigenetic regulation. *Nat Commun*. 2025;16(1):4892.doi: 10.1038/s41467-025-60037-4.
28. Shao Y, Zhang H, Guan H, Wu C, Qi W, Yang L, et al. *PDZK1* protects against mechanical overload-induced chondrocyte senescence and osteoarthritis by targeting mitochondrial function. *Bone Res*. 2024;12(1):41.doi: 10.1038/s41413-024-00344-6