

# Identification of Key Glycolysis-Related Genes in Osteoarthritis and Their Correlation with Immune Infiltration Using Bioinformatics Analysis and Machine Learning

Yifang Zhu <sup>1</sup>, Lin Deng <sup>1</sup>, Junxiang Xia <sup>1</sup>, Jing Yang <sup>1</sup>, Dan Zhao <sup>1</sup>, Min Li <sup>2</sup>

<sup>1</sup>Clinical Laboratory, Sichuan Province Orthopedic Hospital, Chengdu, Sichuan, People's Republic of China; <sup>2</sup>Department of Rheumatoid Osteoarthropathy, Sichuan Province Orthopedic Hospital, Chengdu, Sichuan, People's Republic of China

Correspondence: Yifang Zhu, Email 53113072@qq.com

**Objective:** Osteoarthritis (OA) is a degenerative disorder associated with glycolysis. However, the precise mechanisms remain unclear. This study aimed to identify glycolysis-associated biomarkers and elucidate how glycolysis-related genes interact with the synovial immune microenvironment in OA progression.

**Methods:** Normal and OA synovial gene expression profile microarrays were obtained from the Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) were identified using limma package. Gene Ontology (GO) and KEGG enrichment analyses were conducted to explore biological functions. Weighted Gene Co-expression Network Analysis (WGCNA) was used to identify OA-associated genes, which were intersected with glycolysis genes from The Molecular Signatures Database (MSigDB) and DEGs to obtain key genes. Lasso regression and random forest models were employed to establish a risk model, and its predictive performance was evaluated using nomogram, Receiver Operating Characteristic (ROC) analysis, and Decision Curve Analysis (DCA). Gene Set Enrichment Analysis (GSEA) and Cibersort analysis were conducted to explore pathways and immune infiltration correlations.

**Results:** A total of 239 OA-associated genes were identified through WGCNA. Six hub genes were obtained by intersecting with glycolysis genes and DEGs. Four key glycolytic genes were selected by Lasso regression and random forest models. The nomogram showed that three genes (DDIT4, SLC16A7, SLC2A3) could predict OA risk accurately. The ROC analysis demonstrated an area under the curve (AUC) of 0.85, indicating good predictive performance. Distinct immune cell distribution patterns were observed in OA groups. Interaction networks were constructed for the key genes with related miRNAs, transcription factors (TFs), and small molecule drugs.

**Conclusion:** This study identified three key glycolysis-related genes (DDIT4, SLC16A7, SLC2A3) in OA, revealing their potential roles in disease progression and immune infiltration. These findings may provide new insights into the pathogenesis and therapeutic targets for OA, based on the identified genes and their interactions with the immune microenvironment.

**Keywords:** bioinformatics, DEGs, glycolysis, immune infiltration, machine learning, osteoarthritis

## Introduction

Osteoarthritis (OA) is the most prevalent degenerative joint disease and a leading cause of disability among the elderly, affecting more than 300 million individuals worldwide and incurring annual direct and indirect costs exceeding USD 300 billion.<sup>1</sup> The disease is characterized by the degeneration of articular cartilage, which leads to joint pain, stiffness, and reduced mobility.<sup>2</sup> Researchers found that OA exhibited considerable clinical and molecular heterogeneity,<sup>3</sup> highlighting the complexity of its pathogenesis. Therefore, despite its high prevalence and substantial disease burden, the precise molecular mechanisms governing OA initiation and progression remain incompletely understood. Current therapeutic approaches primarily focus on symptom management, encompassing exercise therapy, nonsteroidal anti-inflammatory drugs (NSAIDs), and intra-articular injections,

which offer limited disease-modifying effects.<sup>4</sup> For patients with end-state disease, joint replacement surgery provides significant symptomatic relief and functional improvement,<sup>4</sup> but the procedure carries inherent risks and potential complications associated with major surgery.

Cellular energy metabolism plays a fundamental role in maintaining tissue homeostasis. Glycolysis, a core metabolic pathway converting glucose to pyruvate to generate ATP and metabolic intermediates, is crucial for various cellular functions. Growing evidence indicates that metabolic reprogramming, particularly a pronounced shift towards away from oxidative phosphorylation,<sup>5,6</sup> is a key feature of OA pathology and represents a critical adaptation to the altered inflammatory microenvironment within the joint.<sup>5</sup> Specifically, this glycolytic shift is observed in critical joint tissues affected by OA, including chondrocytes and synovial cells.<sup>6,7</sup> This metabolic reprogramming is not merely a passive response but is also recognized as an active driver contributing to OA pathology.<sup>8</sup> It can significantly influence chondrocyte phenotypes, alter their subpopulations, promote extracellular matrix degradation, and ultimately drive disease progression.<sup>9</sup>

Furthermore, accumulating research underscores a profound link between altered cellular metabolism, particularly glycolysis, and immune cell infiltration in OA.<sup>10</sup> Metabolic changes occurring in resident joint cells, such as chondrocytes and synovial fibroblasts, actively modulate the local immune milieu.<sup>10</sup> This modulation significantly influences the recruitment, activation state, and functional behavior of various immune cell subsets (for example, macrophages, T cells) recruited into the synovium and potentially other joint tissues.<sup>11</sup> Importantly, distinct patterns of glycolytic activity within the joint tissues have been shown to correlate strongly with specific immune microenvironments or “inflamed” phenotypes in OA.<sup>10</sup> Therefore, deciphering the intricate interplay between dysregulated glycolysis and immune cell infiltration is essential not only for understanding OA pathogenesis but also for identifying novel diagnostic biomarkers and therapeutic targets for OA.

Bioinformatics and machine learning approaches have emerged as indispensable tools in dissecting the molecular complexity of multifactorial diseases like OA.<sup>2,7,12</sup> These computational methodologies facilitate the unbiased identification of key regulatory pathways, disease-associated gene signatures, and potential biomarkers that might be obscured in conventional analysis.<sup>13</sup> Given the established critical role of glycolysis in OA and its emerging strong correlation with immune dysregulation, the primary objective of the present study is to utilize comprehensive bioinformatics analysis of relevant transcriptomic data to identify and validate key glycolysis-related genes specifically implicated in human OA. Building upon this, we will investigate the correlation between the expression patterns of these identified glycolysis-related genes and the landscape of immune cell infiltration within OA tissues. Ultimately, this integrated approach aims to uncover crucial molecular players and networks, paving the way for the discovery of novel biomarkers with diagnostic or prognostic utility and actionable therapeutic targets for this debilitating disease.

## Materials and Methods

### Data Source

The osteoarthritis datasets (GSE55457, GSE55235) were filtered out from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>). GSE55457 and GSE55235 were sequenced on platform GPL96 and including 20 synovial tissues from healthy control and 20 synovial tissues from OA patients. The genes associated with glycolysis-related pathways were sourced from the Molecular Signatures Database (MSigDB) (<https://www.gsea-msigdb.org/gsea/msigdb/>).

### Identification of the DEGs

Microarray datasets were downloaded from the GEO database through the GEOquery package. The GSE55457 and GSE55235 datasets were merged. Considering the technical differences such as platform, probe, scanning parameters, experiment date often much greater than the biological differences, the direct merger will introduce a “batch effect”, resulting in false positives or masking the real difference. We then used the ComBat function from the sva package to remove batch effects, standardize the data, and annotate the probes. When multiple probes corresponding to the same molecule were happened, only the probe with the largest signal value was retained. The limma package was used to

analyze the difference between patients and control groups. The difference analysis results were filtered with  $|\log_2FC| \geq 0.58$  and  $P \text{ value} < 0.05$ . The DEGs were visualized by volcano diagram and heatmap with ggplot2 package.

## GO and KEGG Enrichment Analysis

To acquire the gene function of DEGs, we conducted Gene Ontology (GO) and Kyoto Encyclopedia of Gene and Genomes (KEGG) enrichment analysis with ClusterProfiler package, and the main enrichment results were presented with ggplot2 package. Adjust  $P < 0.05$  was the criteria considering as statistically significant.

## Weighted Gene Co-Expression Network Analysis (WGCNA)

WGCNA is a systems biology approach used to identify clusters (modules) of genes that exhibit correlated expression patterns and to assess the relationship between these modules and external traits, such as OA pathogenesis. The WGCNA R package was used to construct a gene co-expression network from the preprocessed GSE55457 and GSE55235 merged dataset. We first did the sample clustering and outlier detection, and then a sample-clustering tree was generated to identify and remove any outliers from the dataset. The adjacency matrix (AM) was transformed into a topological overlap measure (TOM) matrix to quantify the interconnectedness of genes. A soft thresholding power ( $\beta$ ) was chosen to ensure the network approached a scale-free topology. The DynamicTreeCut method was used to classify genes with similar expression profiles into distinct modules. The correlation between each module and OA pathogenesis was calculated. Modules with the absolute value of correlation coefficients  $\geq 0.5$  and  $p\text{-values} < 0.5$  were selected as relevant modules. Within the relevant modules, intra-modular important genes were identified based on their connectivity within the module. These genes were intersected with glycolysis genes retrieving from the MsigDB database and DEGs to obtain hub genes, which were considered as potential biomarkers for OA.

## Construction of OA Risk Model

We aimed to construct an OA risk model by employing a multi-step approach. Initially, the hub genes were subjected to LASSO regression with “glmnet” package and random forest models with “randomForest” package to identify the most relevant genes associated with OA. The intersection of the genes selected by both methods was then determined to obtain key genes. We further analyzed their expression patterns and the correlation coefficients among their expressions. Receiver operating characteristic (ROC) curves were generated to assess the diagnostic efficacy of the hub genes in distinguishing OA samples from normal controls with pROC 1.18.0. Subsequently, a nomogram plot was constructed based on the key genes and their expression-related features. The nomogram plot served as a graphical tool that integrated multiple variables into a single numerical score, facilitating the prediction of OA risk in a more intuitive and accessible manner. Finally, decision curve analysis (DCA) was performed to evaluate the clinical utility of the constructed OA risk model.

## GSEA Analysis

In this study, we performed Gene Set Enrichment Analysis (GSEA) using R (version 4.2.1) to analyze microarray datasets GSE55235 and GSE55457, which were downloaded from the Gene Expression Omnibus (GEO) database. We employed the clusterProfiler package (version 4.4.4) for the enrichment analysis on key genes. This approach provided insights into the underlying biological mechanisms of key genes participating in OA pathogenesis.

## Immune Cell Infiltration Analysis of OA

A comprehensive evaluation of immune cell infiltration was conducted using Cibersort algorithm (version 1.03) to further explore the role of immune cell infiltration in OA. Subsequently, the correlations between each infiltrated immune cell type were estimated, and significant correlations between each hub gene and the corresponding immune cells were also detected.

## miRNA–mRNA, TF-mRNA and Drug-mRNA Network Construction of Key Glycolytic Genes

The miRNA associated with key glycolytic genes was screened from the miRwalk website (<http://mirwalk.umm.uni-heidelberg.de/>) and the miRTabBase database ([https://awi.cuhk.edu.cn/~miRTarBase/miRTarBase\\_2025/php/index.php](https://awi.cuhk.edu.cn/~miRTarBase/miRTarBase_2025/php/index.php)) and

then obtained the common miRNA by intersection. The transcription factors that bind to the hub genes were identified from the TRRUST (version 2) database (<https://www.grnpedia.org/trrust/>), and the DGIdb (<https://www.dgldb.org/>) provided a simple interface for searching list of western medicines that had known or potential drug-gene interactions with hub genes, and these interactions were mined from DrugBank, PharmGKB, ChEMBL, Drug Target Commons, and other databases. We only preserve the drugs which had been authorized by the Food and Drug Administration (FDA). The interaction networks were constructed and visualized using Cytoscape 3.6.1 software.

## Statistical Analysis

All data manipulations and statistical analysis were conducted utilizing the R programming language (<https://www.r-project.org/>, version 4.1.0). When comparing continuous variables between two groups, the statistical significance for variables that followed a normal distribution was determined using the independent Student's *t*-test. In contrast, for variables that did not adhere to a normal distribution, the Mann–Whitney *U*-test (also known as the Wilcoxon rank-sum test) was employed to assess differences. All reported *p*-values were two-tailed, with a threshold of  $p < 0.05$  considered indicative of statistical significance.

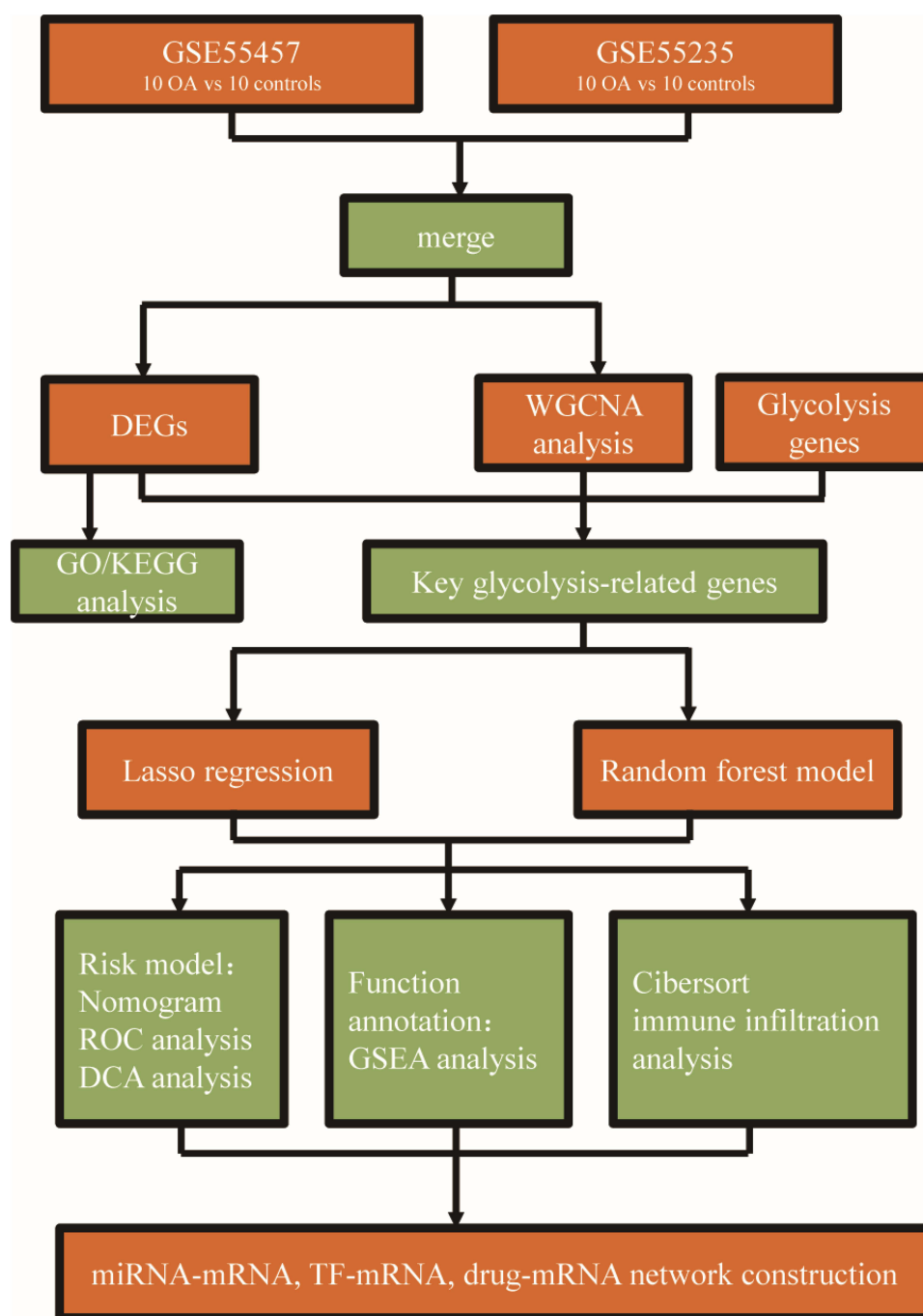
## Results

### Expression of DEGs and Functional Analysis of OA Patients

The workflow of this research is illustrated in Figure 1. The ComBat function of the *sva* package was used to remove batch differences of GSE55235 and GSE55457 chip datasets (Figure 2A). We provided Principal Component Analysis (PCA) plot to evaluation of correction efficacy (Figure 2B and C). The separation between the reference and test samples in Figure 2B indicated the presence of batch effects. After batch effects correction, the improved clustering of the samples in Figure 2C suggested that the batch effects have been effectively removed. DEGs were filtered out with the threshold of *P* value  $< 0.05$  and  $|\log_{2}FC| \geq 0.58$  and 1822 DEGs were filtered out, with 1026 upregulated genes and 796 downregulated genes (Figure 2D). The heatmap displayed the top 30 differential expression genes in OA patients (Figure 2E). To further understand the potential role of DEGs, GO and KEGG analysis were conducted. The biological process (BP) terms were focused on positive regulation of leukocyte activation, leukocyte activation, cytokine production and response to peptide, while those genes were mainly located at collagen-containing extracellular matrix, external side of plasma membrane, membrane microdomain, etc. (Figure 3A). The molecule function (MF) terms were concentrated mainly on signaling receptor activator activity, transcription activator activity, glycosaminoglycan binding, and cytokine receptor activity (Figure 3A). These DEGs were enriched in cytokine–cytokine receptor interaction, PI3K–Akt signaling pathway, MAPK signaling pathway, lipid and atherosclerosis, calcium signaling pathway, rheumatoid arthritis and IL-17 signaling pathway (Figure 3B).

### Screening Hub Genes Associated with OA Through WGCNA

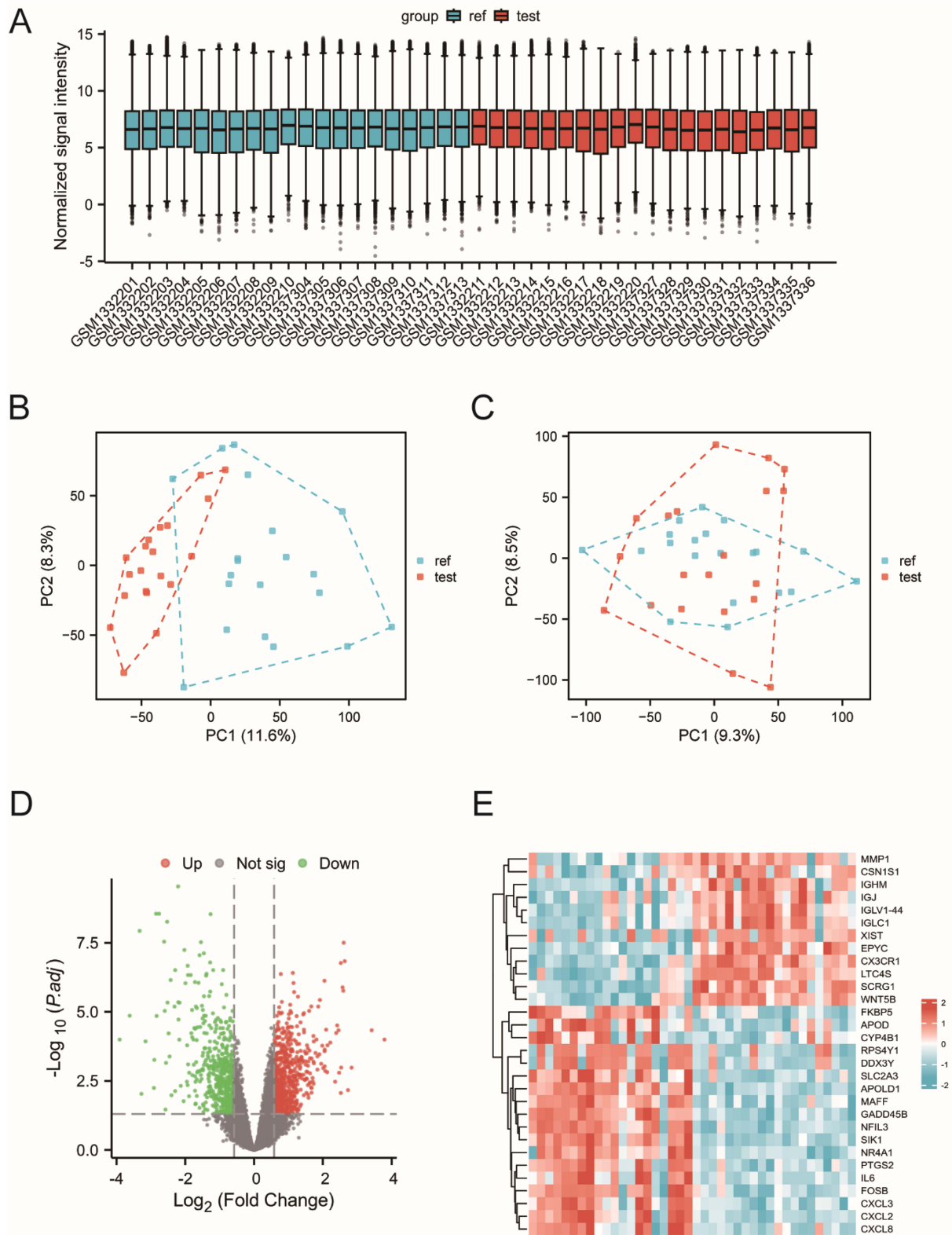
To pinpoint key genes correlated with OA, WGCNA was executed utilizing OA patient cohorts as well as control group. The sample dendrogram depicted the hierarchical clustering of samples based on gene expression patterns (Figure 4A). The curve crosses the horizontal guideline of  $R^2 = 0.85$  at  $k \approx 12$ . This indicated that at power 12 the gene co-expression network first satisfied the scale-free topology criterion ( $R^2 > 0.8$ ) (Figure 4B). Based on the above soft thresholds, we constructed the WGCNA module generation plot, which was cut into different modules, represented by distinct colors (Figure 4C). The WGCNA module-trait heatmap in Figure 4D displayed the correlations between different modules and specific traits. The brown module (correlation coefficient = 0.672, *p*-value =  $2.8 \times 10^{-6}$ ), yellow module (correlation coefficient = 0.742, *p*-value =  $6.4 \times 10^{-8}$ ) and red module (correlation coefficient = 0.656, *p*-value =  $5.8 \times 10^{-6}$ ) were found to be positively associated with OA. Inversely, the green module (correlation coefficient =  $-0.796$ , *p*-value =  $1.4 \times 10^{-9}$ ) and blue module (correlation coefficient =  $-0.63$ , *p*-value =  $1.7 \times 10^{-5}$ ) were negatively associated with OA (Figure 4D). From these modules, 239 genes were selected based on their significance and the absolute correlation coefficients exceeding 0.5 (Figure 4E–I), which were further intersected with 309 glycolysis genes retrieving from the MsigDB database and 1822 DEGs, and then 6 hub genes (DDIT4, VEGFA, HK3, FBP1, SLC16A7, SLC2A3) were obtained (Figure 4J), all of which participated in energy or glucose metabolism (Table 1).



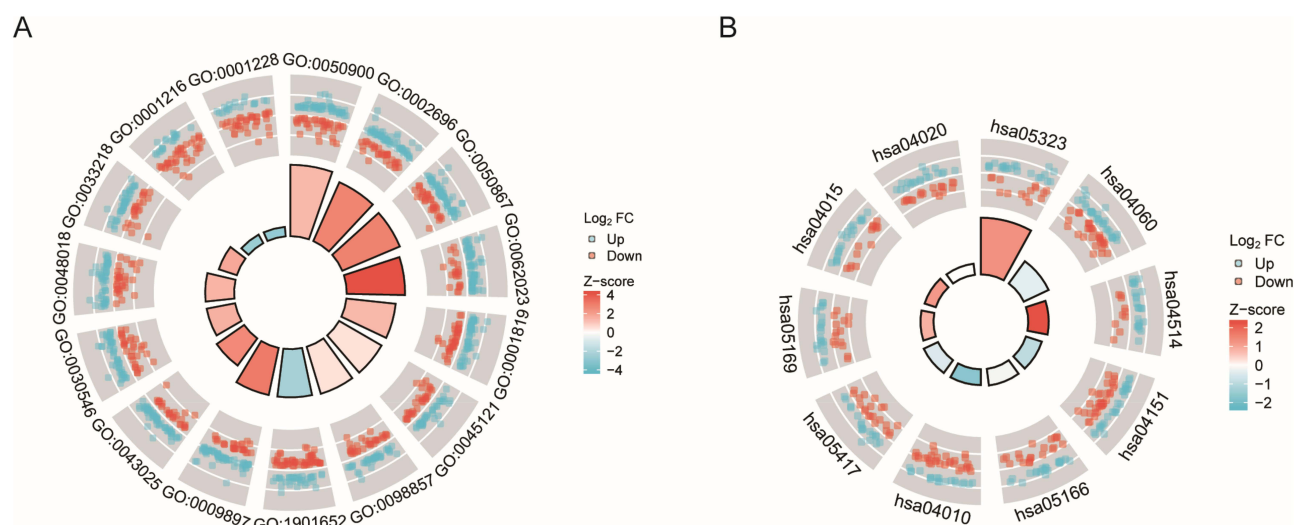
**Figure 1** The work flowchart of this research.

## Construction of OA Risk Model

These hub genes were further analyzed by LASSO regression algorithm, in which four key glycolytic genes (DDIT4, VEGFA, SLC16A7 and SLC2A3) were filtered out (Figure 5A and B). The six hub genes were imported into a random forest model and the above four genes were also screened out (Figure 5C), which were analyzed in subsequent research. Box plots illuminated substantial differences in the expression levels of four critical glycolytic genes between the control and OA groups (Figure 5D). Pearson correlation analysis indicated that VEGFA, SLC16A7 and SLC2A3 were positively correlated with each other, and only gene DDIT4 was negatively related with other genes (Figure 5E). These genes had perfect diagnostic value to distinguish OA patients from controls (Figure 5F–I). The relationship between the linear



**Figure 2** Differences between OA and control groups were analyzed based on GSE55235 and GSE55457 datasets. **(A)** The differential Boxplot was used to draw the corresponding data situation of each sample to view the sample correction results. **(B)** PCA plots before batch effect correction. **(C)** PCA plots after batch effect correction. **(D)** The differentially expressed genes were presented by volcano plot. The green, red, and gray dots represent genes that were down-regulated, up-regulated, and no differential expression genes, respectively. **(E)** The heatmap displayed the expression patterns of the top 30 differential expression genes.

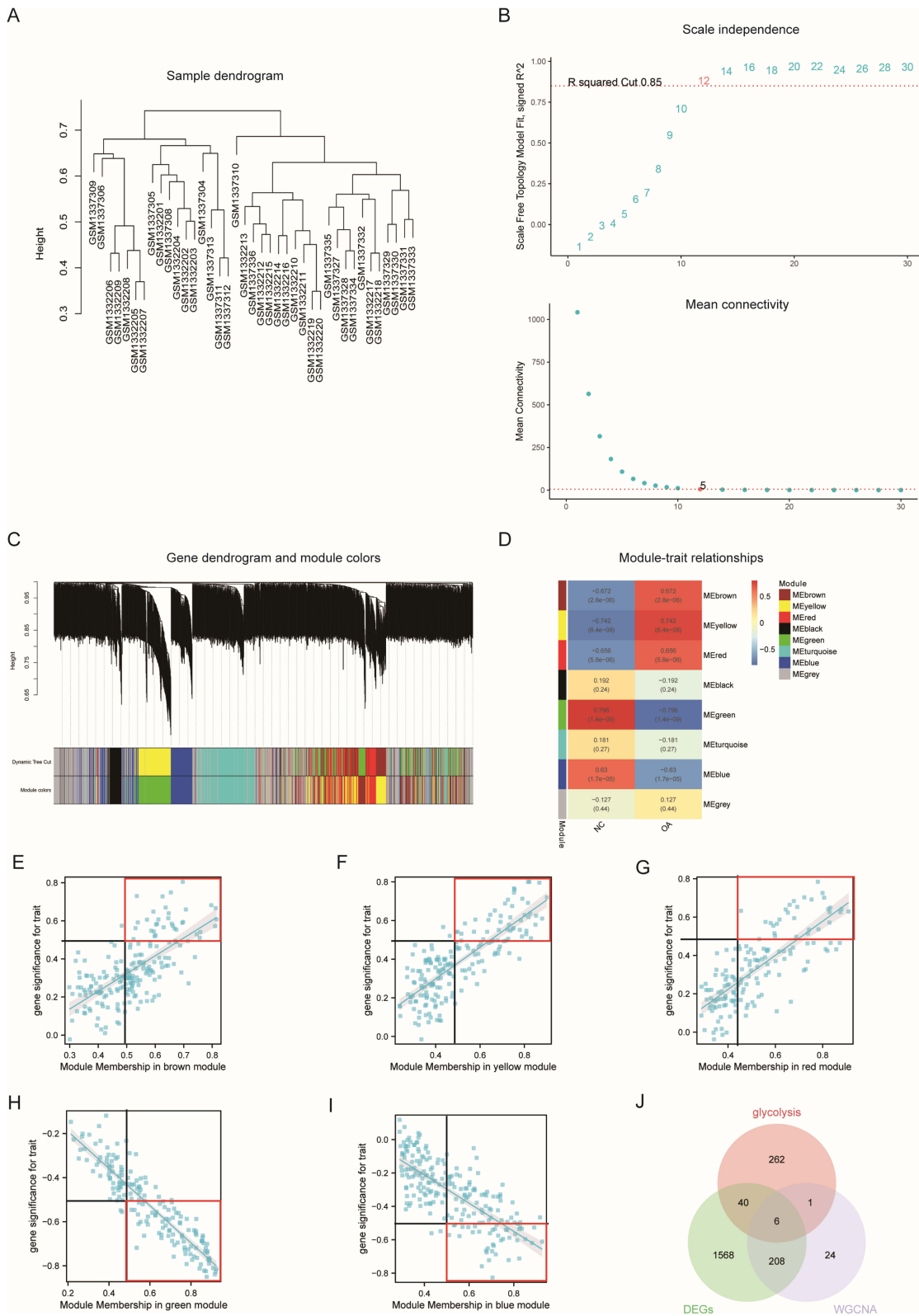


**Figure 3** GO and KEGG analysis of DEGs related with OA were conducted and displayed with circle map. **(A)** GO analysis results. **(B)** KEGG analysis results.

predictor and risk was delineated, including DDIT4 with a negative correlation, SLC16A7 and SLC2A3 showing a positive correlation (Figure 5J). As the linear predictor increased, so did the risk, suggesting a reliable risk assessment model. The calibration curve (Figure 5K) demonstrated good agreement between the predicted probabilities and actual outcomes. This indicated that the prediction of our model was well-calibrated, enhancing the credibility of our findings. GSEA enrichment analysis associated with DDIT4, SLC16A7 and SLC2A3 participated in oxidative phosphorylation pathways, lysosome, MAPK signaling pathway, cell adhesion molecules, adipocytokine signaling pathway, spliceosome, cytokine–cytokine receptor interaction, Nod-like receptor signaling pathway (Figure 6).

## Immune Infiltration Profile of OA

We further explored the immune cell composition and gene expression analysis in normal control and OA subjects. The heatmap illustrated the estimated proportions of various immune cell types across subjects categorized into control and OA groups (Figure 7A). Distinct patterns of immune cell distribution were observable between two groups. Notably, T cells CD4 memory resting and Mast cells activated exhibited higher proportions in control group, whereas Mast cells resting and Plasma cells showed elevated proportions in OA group. The boxplot provided a detailed comparison of immune cells composition between control and OA groups (Figure 7B). Significant differences in cell proportions were observed for several cell types, including T cells CD4 memory resting, Mast cells activated, Mast cells resting and Plasma cells, with p-values indicated for statistical significance (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ). The correlation matrix reveals the relationships between different immune cell types (Figure 7C). Strong positive correlations were evident among certain cell types, such as T cells CD4 memory resting and Mast cells activated, Macrophages M1 and T cells gamma delta. While negative correlations were observed between others, like Mast cells resting and Mast cells activated, T cells gamma delta and Macrophages M2, Macrophages M2 and T cells gamma delta, and the like. The heatmap displayed the expression levels of DDIT4, SLC16A7 and SLC2A3 in various immune cell types (Figure 7D). DDIT4 showed higher expression in T cells CD4 memory resting and Macrophage M1, but lower expression in Plasma cells. SLC16A7 was highly expressed in T cells CD4 memory resting, Mast cells activated, Macrophages M1, and lowly expressed in Mast cells resting, Plasma cells. SLC2A3 had higher expression in T cells CD4 memory resting, Mast cells activated, but lower expression in Mast cells resting and Macrophages M0.



**Figure 4** WGCNA analysis of gene expression and module relationships. **(A)** Sample dendrogram with the clustering of samples. **(B)** Scale Independence plot and Mean connectivity plot demonstrating the robustness of module identification across different soft-thresholding powers. **(C)** Gene dendrogram and module colors based on gene expression patterns and their association with specific gene modules. **(D)** Heatmap of gene expression correlation labeled with coefficients and corresponding p-values. **(E–I)** Scatter plot illustrating the correlation between gene significance for OA related trait and module membership. There was a positive correlation (in the upper right corner of the figure) or a negative correlation (in the lower right corner of the figure) among the genes in the different modules within the red box. **(J)** Venn diagram displaying the overlap between DEGs, hub genes identified by WGCNA and glycolysis related genes.

**Table 1** Gene Characteristics of 6 Hub Genes

Gene Symbol	Description	Biological Function
DDIT4	DNA Damage Inducible Transcript 4	Responses to cellular energy levels and cellular stress
VEGFA	Vascular Endothelial Growth Factor A	Protecting cells from hypoxia-mediated cell death
HK3	Hexokinase 3	Mediating the initial step of glycolysis
FBPI	Fructose-Bisphosphatase 1	Acting as a rate-limiting enzyme in gluconeogenesis.
SLC16A7	Solute Carrier Family 16 Member 7	Catalyzing the rapid transport across the plasma membrane of monocarboxylates
SLC2A3	Solute Carrier Family 2 Member 3	Facilitatively transporting glucose

## Screening of miRNA, TF and Small Molecule Drug Related with Key Genes and Interaction Networks Construction

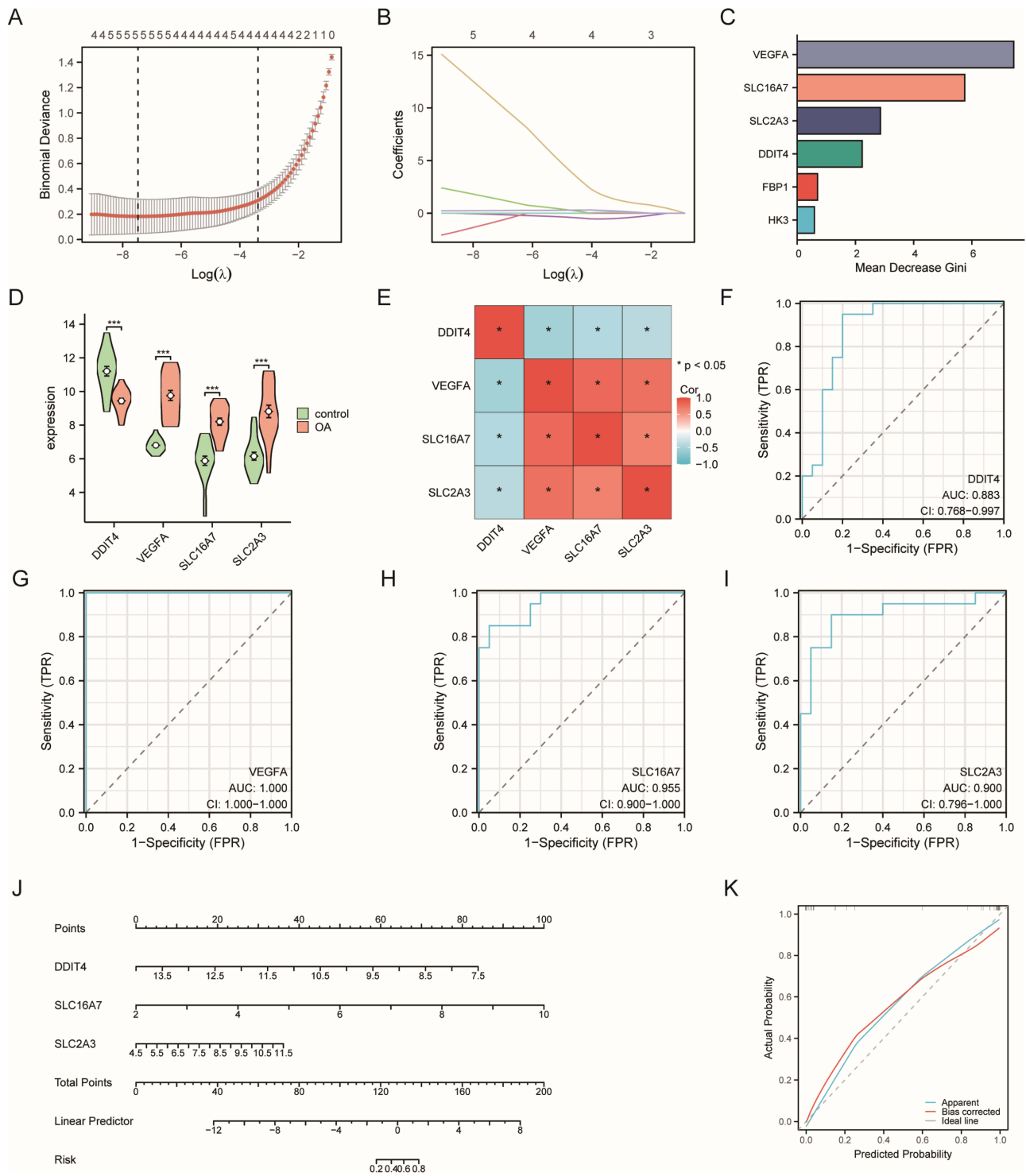
34 related miRNAs with three key glycolytic genes were obtained, in which 14 miRNAs were related with SLC2A3, 13 miRNAs with DDIT4 and 7 miRNAs with SLC16A7 (Figure 8A). Six transcription factors, which have been found participating in glycolysis process or OA pathophysiological process, were interacted with three key glycolytic genes (Figure 8B). Nine small molecule drugs were obtained which have known or potential drug–gene interactions with three key glycolytic genes (Figure 8C).

## Discussion

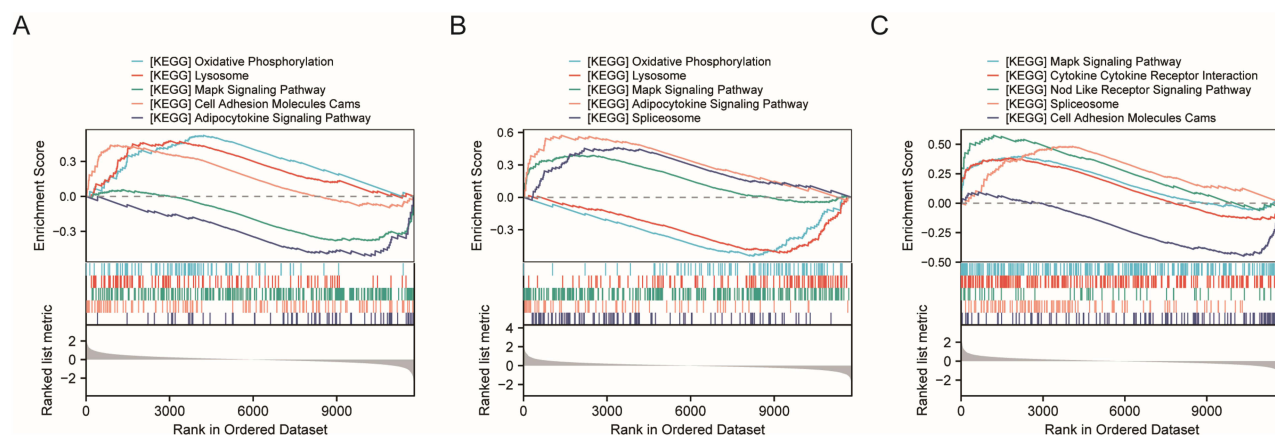
OA is a common degenerative joint disease primarily affecting the elderly, leading to joint pain, stiffness, and functional impairment, severely impacting patients' quality of life. With the global aging population, the incidence of osteoarthritis continues to rise, and it is estimated that by 2050, the number of osteoarthritis patients worldwide will exceed 100 million.<sup>10</sup> Currently, treatment methods for osteoarthritis mainly include medication, physical therapy, and surgical intervention; however, these methods often yield unsatisfactory results and are accompanied by varying degrees of side effects. There is an urgent need to explore new biomarkers and therapeutic targets to improve patient prognosis and quality of life.<sup>14</sup> In recent years, more and more studies have shown that energy metabolism, especially glycolysis, is closely related to the occurrence and development of OA.<sup>10</sup> Research has found that the level of glycolysis in synovial tissue is significantly increased in patients with OA, which may be due to changes in the inflammatory microenvironment.<sup>8,9,15,16</sup> However, until now, the interaction between glycolysis and immune infiltration in OA remains unexplored, yet warrants immediate and comprehensive investigation.

This study aims to identify key glycolytic genes associated with osteoarthritis through bioinformatics analysis and machine learning methods and to explore their relationship with immune cell infiltration. In this study, we integrated GEO datasets, WGCNA and MsigDB database to screen for glycolysis-related genes associated with OA. We further constructed a risk model using Lasso regression and random forest models, ultimately identifying three key genes (DDIT4, SLC16A7, and SLC2A3). The predictive performance of the risk model was evaluated using Nomogram, ROC analysis, and Decision Curve Analysis (DCA), demonstrating high clinical application value.

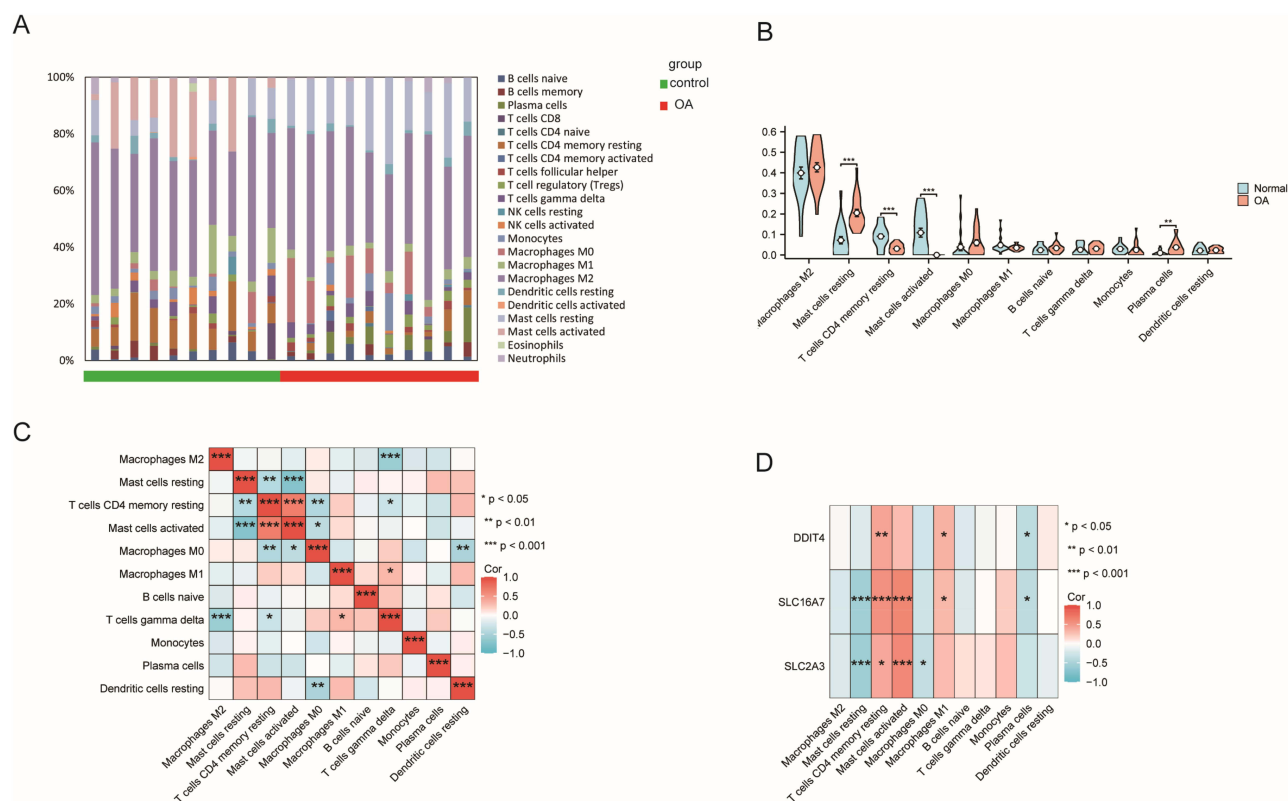
DDIT4, also known as REDD-1, is a protein that plays a crucial role in cellular stress responses and has been shown to be involved in various diseases, including OA. Yin et al have shown that DDIT4 is upregulated in the cartilage of OA patients and is correlated with the severity.<sup>17</sup> Another study had a different finding, which found that DDIT4 expression was significantly reduced in aged and OA cartilage,<sup>18</sup> and the deficiency of DDIT4 exacerbated the severity of experimental OA model, indicating its protective role in cartilage homeostasis.<sup>19</sup> In fact, our study found that DDIT4 expression was down-regulated in synovial tissues of OA patients. The reason why different studies reached different conclusions might be due to the different samples being tested. We will use clinical samples from different parts of OA patients to detect the expression level of DDIT4 or single cell sequencing to verify our hypothesis. SLC16A7, also named as MCT2, is a member of the monocarboxylate transporter family, which participating in transporting metabolites, such as lactate, pyruvate, and ketone bodies. SLC16A7 can efficiently transport lactic acid and pyruvate, the metabolites that play a key role in glycolysis, out of the cell, helping relieve the acidic environment within the cell, thereby maintaining the normal process of glycolysis.<sup>20</sup> However, the role of SLC16A7 in OA has not been extensively explored. In the



**Figure 5** The identification and diagnostic value analysis of key glycolytic genes in OA patients. **(A)** Plot of binomial distribution bias versus log ( $\lambda$ ) of LASSO regression. The plot showed the relationship between Binomial Deviance and the logarithm of  $\lambda$ . Points represented different coefficients and their corresponding deviance values. **(B)** Coefficients plot displayed the coefficients of HK3, FBP1, DDIT4, SLC2A3, SLC16A7, and VEGFA against the log ( $\lambda$ ). **(C)** Mean decrease Gini plot illustrated the mean decrease in Gini index for genes DDIT4, VEGFA, SLC2A3 and SLC16A7, indicating their importance in the model. **(D)** Gene expression boxplot of four glycolytic genes in OA and control samples. **(E)** The heatmap of correlation analysis between differentially expressed glycolytic genes, the color of each small square represents the correlation, with the red color representing the stronger positive correlation and the blue color representing the stronger negative correlation. Asterisks in the small squares indicate the significance of the statistical difference. (\*p<0.05). **(F–I)** ROC curves of four glycolytic genes. **(J)** Nomogram plot showed the relationship between the linear predictor and risk, with points representing different levels of risk. **(K)** This calibration curve compared the predicted probability with the actual probability, in which the ideal line representing perfect calibration.

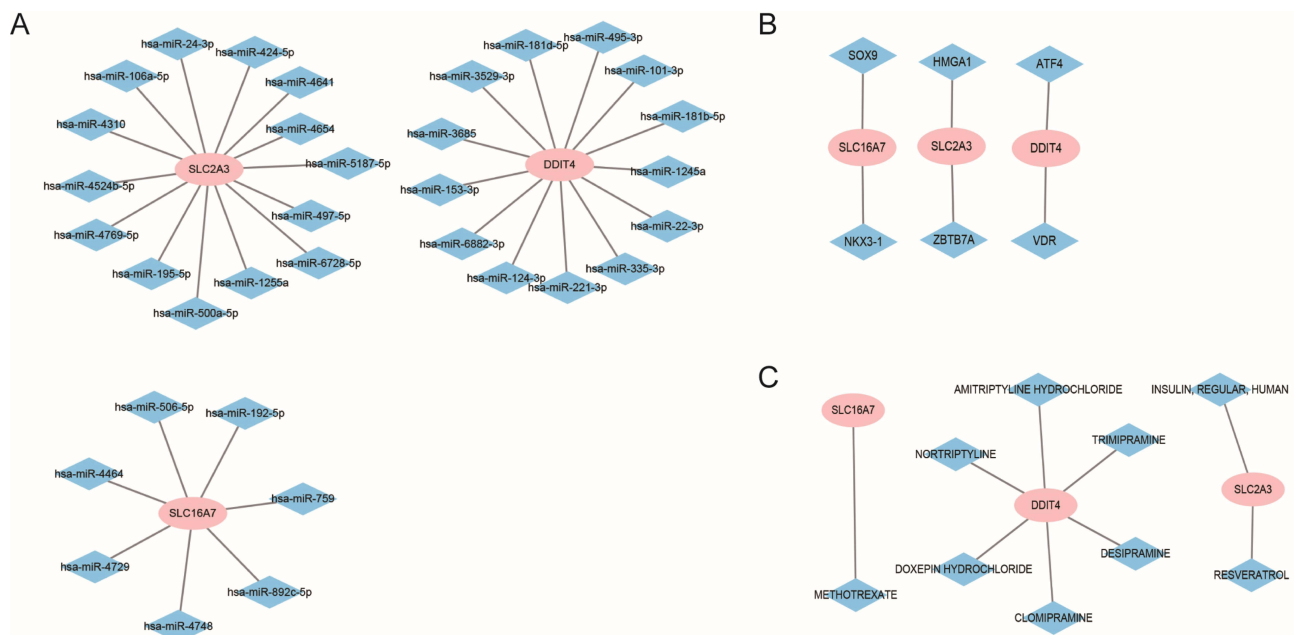


**Figure 6** GSEA plot of high and low expression groups with key glycolytic biomarkers screening from machine learning. **(A)** GSEA enrichment pathways associated with DDIT4 highlighted oxidative phosphorylation pathways, lysosome, MAPK signaling pathway, cell adhesion molecules, adipocytokine signaling pathway. **(B)** GSEA enrichment pathways associated with SLC16A7 highlighted oxidative phosphorylation, lysosome, MAPK signaling pathway, spliceosome and adipocytokine signaling pathway. **(C)** GSEA enrichment pathways associated with SLC2A3 were mainly concentrated on MAPK signaling pathway, cytokine–cytokine receptor interaction, Nod like receptor signaling pathway, spliceosome, cell adhesion molecules.



**Figure 7** Group comparisons with different immune cell subsets and their correlation with gene expressions. **(A)** The bar chart showed the proportion of immune cell subsets with control and OA groups. **(B)** The boxplot revealed main immune cells levels between normal control and OA patients with statistical significance denoted by asterisks. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ . **(C)** The correlation of specific immune cell subsets, including Macrophages M2, Mast cells resting, T cells CD4 memory resting, Mast cells activated, Macrophages M0, Macrophages M1, B cells naive, T cells gamma delta, Monocytes, Plasma cells, and Dendritic cells resting, with statistical significance denoted by asterisks. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ . **(D)** The correlation of gene expression (SLC2A3, SLC16A7, DDIT4) with specific immune cell subsets.

present study, we discovered that SLC16A7 is highly expressed in OA, suggesting its potential involvement in the pathogenesis of this disease. SLC2A3 is a key facilitative glucose transporter that plays a crucial role in glucose uptake and metabolism. Our study revealed that SLC2A3 expression was upregulated in OA and closely associated with glucose metabolism and immune infiltration, similar to what has been observed in carcinoma.<sup>21–23</sup> These genes had good



**Figure 8** The interaction network of miRNA (A), TF (B) and small molecule drug (C) with key glycolytic genes.

accuracy in diagnosing osteoarthritis, with the area under the ROC curve exceeding 0.85. These findings suggest that the glycolytic process may play an important role in the pathogenesis of OA and provide new perspectives for potential diagnosis and therapeutic targets and multi-gene combined diagnostic panel is worthy of further research.

GSEA enrichment analysis indicated that three key glycolysis-related genes participated in oxidative phosphorylation pathways, lysosome, MAPK signaling pathway, and cytokine receptor interaction, further validating the key role of glycolysis in osteoarthritis. These enrichment results provide a theoretical basis for metabolic intervention, suggesting that in-depth exploration of the glycolytic pathway and its potential mechanisms of interaction with immune cells is an important direction for future research.<sup>24,25</sup>

Additionally, we found that these three genes were also closely correlated with immune infiltration. Immune cell infiltration manifested that DDIT4 had higher expression in T cells CD4 memory resting and Macrophage M1, but lower expression in Plasma cells, which suggested that DDIT4 may play a role in immune memory, cellular and humoral immunity and inflammation regulation. Studies have found that DDIT4 regulated the glycolysis process and participates in the excessive activation of fibroblast-like synoviocytes (FLSs) and cartilage damage induced by high glucose, indicating that overexpression of DDIT4 can inhibit the secretion of inflammatory factors and alleviate the pathological process of osteoarthritis.<sup>26</sup> This provides a reference for us to understand the role of DDIT4 in osteoarthritis from glycolysis and immune modulation. SLC16A7 was highly expressed in T cells CD4 memory resting, Mast cells activated, Macrophages M1, lowly expressed in Mast cells resting and Plasma cells, manifested the complex role of this gene in immune cell function and inflammatory response. Previous studies noted that SLC16A7 expression was upregulated in FLS of rheumatoid arthritis.<sup>27</sup> By regulating the transport of lactic acid and sugar metabolism, it affected the metabolic reprogramming of immune cells.<sup>27</sup> This indicates that SLC16A7 plays an important role in the inflammatory response and the regulation of immune cell functions and may be related to the pathological mechanism of osteoarthritis. SLC2A3 had higher expression in T cells CD4 memory resting, Mast cells activated, but lower expression in Mast cells resting and Macrophages M0, which merited further investigation into its metabolic regulatory role in immune cell function. Previous research found that SLC2A3 promotes the infiltration of macrophages in gastric cancer by reprogramming glycolysis,<sup>28</sup> indicating that SLC2A3 played a significant role in regulating the process of sugar metabolism and influencing the functions of immune cells which maybe play the similar role in osteoarthritis. This result indicates that glycolysis-related genes may influence the inflammatory response in osteoarthritis by regulating immune cell infiltration.

This is consistent with existing literature, suggesting that changes in glucose metabolism may have profound effects on the immune microenvironment, thereby affecting the disease progression.<sup>29,30</sup> The mechanism of glycolysis-immune cross-talk is of great significance for us to understand the underlying mechanism of genes in OA regulation.

Finally, we constructed interaction networks with miRNAs, transcription factors, and small molecule drugs. A total of 34 miRNAs were identified as related to three key glycolytic genes, with 14 linked to SLC2A3, 13 to DDIT4, and 7 to SLC16A7. These miRNAs have been reported to be involved in processes such as cartilage injury,<sup>31,32</sup> cartilage cell proliferation and apoptosis,<sup>33–36</sup> cartilage matrix degradation,<sup>37</sup> inflammation,<sup>33,35,38</sup> to regulate the occurrence and development of osteoarthritis. Six transcription factors may interact with these glycolytic genes to take part in glycolysis<sup>39</sup> or OA pathophysiology.<sup>40</sup> Additionally, nine small molecule drugs with known or potential interactions with the three key glycolytic genes were identified, which may be new therapy choices. Take resveratrol as an example, as a natural polyphenol, resveratrol may exert its effects through anti-inflammatory, antioxidant, promoting chondrocyte proliferation and inhibiting matrix-degrading enzymes.<sup>41</sup> However, we found resveratrol can target SLC2A3 to modulate the glycolysis process, which may be added with first-line therapies (such as diclofenac) to form a combined treatment plan.

While this study has several limitations: first, the bioinformatics analyses relied on publicly available transcriptomic datasets derived from bulk RNA sequencing, which may introduce biases due to sample heterogeneity, limited sample sizes, and variations in experimental protocols across dataset; second, although computational models demonstrated robust predictive performance, the lack of experimental validation *in vitro* or *in vivo* limits the ability to confirm the functional roles of these genes in glycolysis regulation, immune modulation, or OA progression; third, the study identified correlations between key genes and immune cell populations but did not establish causal relationships or molecular mechanisms linking glycolysis to immune dysregulation in OA; fourth, the miRNA, TF, and drug interaction networks were constructed based on predictive databases, which required to confirm by experimental validation; finally, the study focused exclusively on glycolysis-related genes, potentially overlooking crosstalk with other metabolic pathways (for example, oxidative phosphorylation, lipid metabolism) that may also contribute to OA pathogenesis.

## Conclusion

In summary, this study successfully identifies three key genes (DDIT4, SLC16A7, and SLC2A3) related to glycolysis and reveals their significant correlation with immune cell infiltration. These findings provide new insights into the pathogenesis of osteoarthritis and lay the foundation for the development of future potential therapeutic strategies. Future research should focus on these validating findings through experimental approaches and exploring their clinical application in OA diagnosis and treatment.

## Data Sharing Statement

The data in this paper come from GEO database (<https://www.ncbi.nlm.nih.gov/geo/>) GSE55235 and GSE55457.

## Ethical Approval

The study utilized publicly available data from the GEO database (GSE55457 and GSE55235), which have been anonymized and do not involve any identifiable personal information. This study was exempt from ethical review by our IRB based on Article 32 of the Measures for Ethical Review of Life Science and Medical Research Involving Human Subjects dated February 18, 2023, China.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This study was funded by the Research program of Chengdu Health Commission (2023308, to Yifang Zhu).

## Disclosure

The authors declare no competing interests in this work.

## References

- Sun Y, Fang Y, Li X, et al. A static magnetic field enhances the repair of osteoarthritic cartilage by promoting the migration of stem cells and chondrogenesis. *J Orthopaedic Translation*. 2023;39:43–54. doi:10.1016/j.jot.2022.11.007
- Hu K, Ou Y, Xiao L, et al. Identification and construction of a disulfidptosis-mediated diagnostic model and associated immune microenvironment of osteoarthritis from the perspective of PPPM. *J Inflamm Res*. 2024;17:3753–3770. doi:10.2147/jir.S462179
- Jin Z, Chang B, Wei Y, et al. Curcumin exerts chondroprotective effects against osteoarthritis by promoting AMPK/PINK1/Parkin-mediated mitophagy. *Biomed Pharmacother*. 2022;151:113092. doi:10.1016/j.biopha.2022.113092
- Katz JN, Arant KR, Loeser RF. Diagnosis and treatment of hip and knee osteoarthritis: a review. *JAMA*. 2021;325(6):568–578. doi:10.1001/jama.2020.22171
- Adam MS, Zhuang H, Ren X, et al. The metabolic characteristics and changes of chondrocytes in vivo and in vitro in osteoarthritis. *Front Endocrinol*. 2024;15:1393550. doi:10.3389/fendo.2024.1393550
- Zheng L, Zhang Z, Sheng P, et al. The role of metabolism in chondrocyte dysfunction and the progression of osteoarthritis. *Ageing Res Rev*. 2021;66:101249. doi:10.1016/j.arr.2020.101249
- Huang X, Meng H, Shou Z, et al. Identification of basement membrane-related biomarkers associated with the diagnosis of osteoarthritis based on machine learning. *BMC Med Genomics*. 2023;16(1):198. doi:10.1186/s12920-023-01601-z
- Jiang D, Guo J, Liu Y, et al. Glycolysis: an emerging regulator of osteoarthritis. *Front Immunol*. 2023;14:1327852. doi:10.3389/fimmu.2023.1327852
- Wen ZH, Sung CS, Lin SC, et al. Intra-articular lactate dehydrogenase a inhibitor oxamate reduces experimental osteoarthritis and nociception in rats via possible alteration of glycolysis-related protein expression in cartilage tissue. *Int J Mol Sci*. 2023;24(13):10770. doi:10.3390/ijms241310770
- Chen Z, Hua Y. Gene signature based on glycolysis is closely related to immune infiltration of patients with osteoarthritis. *Cytokine*. 2023;171:156377. doi:10.1016/j.cyto.2023.156377
- Henry ÓC, O'Neill LAJ. Metabolic reprogramming in stromal and immune cells in rheumatoid arthritis and osteoarthritis: therapeutic possibilities. *European J Immunol*. 2025;55(4):e202451381. doi:10.1002/eji.202451381
- Liang Y, Lin F, Huang Y. Identification of biomarkers associated with diagnosis of osteoarthritis patients based on bioinformatics and machine learning. *J Immunol Res*. 2022;2022:5600190. doi:10.1155/2022/5600190
- Fu Q, Zhang S. The role of mitochondrial dysfunction in the pathogenesis of atherosclerosis: a new exploration from bioinformatics analysis. *Medicine*. 2025;104(22):e42601. doi:10.1097/md.00000000000042601
- Huang J, Zhou J, Xue X, et al. Identification of aging-related genes in diagnosing osteoarthritis via integrating bioinformatics analysis and machine learning. *Ageing*. 2024;16(1):153–168. doi:10.18632/aging.205357
- Zhang Q, Sun C, Liu X, et al. Mechanism of immune infiltration in synovial tissue of osteoarthritis: a gene expression-based study. *J Orthopaedic Surg Res*. 2023;18(1):58. doi:10.1186/s13018-023-03541-x
- Yang J, Li S, Li Z, et al. Targeting YAP1-regulated glycolysis in fibroblast-like synoviocytes impairs macrophage infiltration to ameliorate diabetic osteoarthritis progression. *Adv Sci*. 2024;11(5):e2304617. doi:10.1002/advs.202304617
- Yin H, Zhang Y, Wang K, et al. The involvement of regulated in development and DNA damage response 1 (REDD1) in the pathogenesis of intervertebral disc degeneration. *Exp Cell Res*. 2018;372(2):188–197. doi:10.1016/j.yexcr.2018.10.001
- Alvarez-Garcia O, Olmer M, Akagi R, et al. Suppression of REDD1 in osteoarthritis cartilage, a novel mechanism for dysregulated mTOR signaling and defective autophagy. *Osteoarthritis Cartilage*. 2016;24(9):1639–1647. doi:10.1016/j.joca.2016.04.015
- Alvarez-Garcia O, Matsuzaki T, Olmer M, et al. Regulated in development and DNA damage response 1 deficiency impairs autophagy and mitochondrial biogenesis in articular cartilage and increases the severity of experimental osteoarthritis. *Arthritis Rheumatol*. 2017;69(7):1418–1428. doi:10.1002/art.40104
- Zhang B, Jin Q, Xu L, et al. Cooperative transport mechanism of human monocarboxylate transporter 2. *Nat Commun*. 2020;11(1):2429. doi:10.1038/s41467-020-16334-1
- Chu M, Zheng K, Li X, et al. Comprehensive analysis of the role of SLC2A3 on prognosis and immune infiltration in head and neck squamous cell carcinoma. *Analytical Cellular Pathol*. 2022;2022:2371057. doi:10.1155/2022/2371057
- Liu Y, Jin A, Quan X, et al. miR-590-5p/Tiam1-mediated glucose metabolism promotes malignant evolution of pancreatic cancer by regulating SLC2A3 stability. *Can Cell Inter*. 2023;23(1):301. doi:10.1186/s12935-023-03159-3
- Jacquier V, Gitenay D, Fritsch S, et al. RIP140 inhibits glycolysis-dependent proliferation of breast cancer cells by regulating GLUT3 expression through transcriptional crosstalk between hypoxia induced factor and p53. *Cell Mol Life Sci*. 2022;79(5):270. doi:10.1007/s00018-022-04277-3
- Wang S, Zhang Z, Liang J, et al. Identification of several inflammation-related genes based on bioinformatics and experiments. *Int Immunopharmacol*. 2023;121:110409. doi:10.1016/j.intimp.2023.110409
- Du J, Zhou T, Zhang W, et al. Developing the new diagnostic model by integrating bioinformatics and machine learning for osteoarthritis. *J Orthopaedic Surg Res*. 2024;19(1):832. doi:10.1186/s13018-024-05340-4
- Qiang S, Cheng C, Dong Y, et al. DDIT4 participates in high glucose-induced fibroblast-like synoviocytes overactivation and cartilage injury by regulating glycolysis. *Regener Ther*. 2025;29:51–59. doi:10.1016/j.reth.2025.02.017
- Torres A, Pedersen B, Guma M. Solute carrier nutrient transporters in rheumatoid arthritis fibroblast-like synoviocytes. *Front Immunol*. 2022;13:984408. doi:10.3389/fimmu.2022.984408
- Yao X, He Z, Qin C, et al. SLC2A3 promotes macrophage infiltration by glycolysis reprogramming in gastric cancer. *Can Cell Inter*. 2020;20(1):503. doi:10.1186/s12935-020-01599-9
- Nie Q, Cao H, Yang J, et al. Integration RNA bulk and single cell RNA sequencing to explore the change of glycolysis-related immune microenvironment and construct prognostic signature in head and neck squamous cell carcinoma. *Transl Oncol*. 2024;46:102021. doi:10.1016/j.tranon.2024.102021

30. Xie W, Guo H, Zhang J, et al. Comprehensive analysis of the relationship between metabolic reprogramming and immune function in prostate cancer. *Onco Targets Ther.* 2021;14:3251–3266. doi:10.2147/ott.S304298
31. Zhang C, He W. Circ\_0020014 mediates CTSB expression and participates in IL-1 $\beta$ -prompted chondrocyte injury via interacting with miR-24-3p. *J Orthopaedic Surg Res.* 2023;18(1):877. doi:10.1186/s13018-023-04370-8
32. Xu J, Qian X, Ding R. MiR-24-3p attenuates IL-1 $\beta$ -induced chondrocyte injury associated with osteoarthritis by targeting BCL2L12. *J Orthopaedic Surg Res.* 2021;16(1):371. doi:10.1186/s13018-021-02378-6
33. Zou H, Lu C, Qiu J. Long non-coding RNA LINC00265 promotes proliferation, apoptosis, and inflammation of chondrocytes in osteoarthritis by sponging miR-101-3p. *Autoimmunity.* 2021;54(8):526–538. doi:10.1080/08916934.2021.1978432
34. Zhang Y, Lu R, Huang X, et al. Circular RNA MELK promotes chondrocyte apoptosis and inhibits autophagy in osteoarthritis by regulating MYD88/NF- $\kappa$ B signaling axis through MicroRNA-497-5p. *Contrast Media Mol Imaging.* 2022;2022(1):7614497. doi:10.1155/2022/7614497
35. Fan G, Liu J, Zhang Y, et al. LINC00473 exacerbates osteoarthritis development by promoting chondrocyte apoptosis and proinflammatory cytokine production through the miR-424-5p/LY6E axis. *Exp Ther Med.* 2021;22(5):1247. doi:10.3892/etm.2021.10682
36. Zhang X, Ni X, Ni X, et al. Long non-coding RNA H19 modulates proliferation and apoptosis in osteoarthritis via regulating miR-106a-5p. *J Biosci.* 2019;44(2):44.
37. Hou L, Shi H, Wang M, et al. MicroRNA-497-5p attenuates IL-1 $\beta$ -induced cartilage matrix degradation in chondrocytes via Wnt/ $\beta$ -catenin signal pathway. *Int J Clin Exp Pathol.* 2019;12(8):3108–3118.
38. Xiang Q, Wang J, Wang T, et al. Combination of baicalein and miR-106a-5p mimics significantly alleviates IL-1 $\beta$ -induced inflammatory injury in CHON-001 cells. *Exp Ther Med.* 2021;21(4):345. doi:10.3892/etm.2021.9776
39. Zhou Z, Ye S, Chen J, et al. ATF4 promotes glutaminolysis and glycolysis in colorectal cancer by transcriptionally inducing SLC1A5. *Acta Biochim Biophys Sin.* 2024;10(3724/abbs.2024226). doi:10.3724/abbs.2024226
40. Ai J, Zhao F, Zhou X. HMGA1 aggravates oxidative stress injury and inflammatory responses in IL-1 $\beta$ -induced primary chondrocytes through the JMJD3/ZEB1 axis. *Int Arch Allergy Immunol.* 2023;184(3):279–290. doi:10.1159/000526680
41. Zou X, Xu H, Qian W. The role and current research status of resveratrol in the treatment of osteoarthritis and its mechanisms: a narrative review. *Drug Metab Rev.* 2024;56(4):399–412. doi:10.1080/03602532.2024.2402751

## Open Access Rheumatology: Research and Reviews

### Publish your work in this journal

Open Access Rheumatology Research and Reviews is an international, peer-reviewed, open access journal publishing original research, reports, editorials, reviews and commentaries on all aspects of clinical and experimental rheumatology in the clinic and laboratory including the following topics: Pathology, pathophysiology of rheumatological diseases; Investigation, treatment and management of rheumatological diseases; Clinical trials and novel pharmacological approaches for the treatment of rheumatological disorders. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/open-access-rheumatology-research-and-reviews-journal>

**Dovepress**  
Taylor & Francis Group