



OPEN Identification of a novel disulfidptosis-related gene signature in osteoarthritis using bioinformatics analysis and experimental validation

Mingjie Wei, Xiangwen Shi, Wenbao Tang, Qian Lv, Yipeng Wu & Yongqing Xu✉

Osteoarthritis (OA) is a degenerative bone disease characterized by the destruction of joint cartilage and synovial inflammation, involving intricate immune regulation processes. Disulfidptosis, a novel form of programmed cell death, has recently been identified; however, the effects and roles of disulfidptosis-related genes (DR-DEGs) in OA remain unclear. We obtained six OA datasets from the GEO database, using four as training sets and two as validation sets. Differential expression analysis was employed to identify DR-DEGs, and unique molecular subtypes of OA were constructed based on these DR-DEGs. Subsequently, the immune microenvironment of OA patients was comprehensively analyzed using the “CIBERSORT” algorithm for immune infiltration. Various machine learning algorithms were utilized to screen characteristic DR-DEGs, and nomogram models and ROC curves were built based on these genes. The scRNA dataset (GSE169454) was used to classify chondrocytes in OA samples into distinct cell types, further exploring the gene distribution and correlation of characteristic DR-DEGs with specific cell subpopulations. Moreover, the expression levels of four characteristic DR-DEGs were validated through OA cell models and rat models. In our study, we identified 10 DR-DEGs with significant differences in expression within OA samples. Based on these DR-DEGs, two distinct molecular subtypes were recognized (cluster 1 and 2). ZNF484 and NDUFS1 were found to be significantly overexpressed in subtype 1, while the infiltration abundance of activated mast cells was markedly elevated in subtype 2. Moreover, significant differences were observed in the infiltration proportions of 11 immune cell types between OA and control samples, with 9 DR-DEGs demonstrating substantial correlations with immune cell infiltration levels. Further analysis of the scRNA dataset revealed that SLC3A2 and NDUFC1 were predominantly expressed in the preHTC subpopulation. All 10 DR-DEGs exhibited notably higher expression in the EC subpopulation across various cell types. The proportion of EC subgroups with high SLC3A2 expression increased, mainly enriching pathways related to inflammation, such as the IL-17 signaling pathway and TGF-beta signaling pathway. Using machine learning, we identified four characteristic DR-DEGs, which, in combination with the nomogram and ROC models, demonstrated promising performance in the diagnosis of OA. Additionally, *in vivo* validation confirmed a significant elevation of PPM1F expression in OA models. This study identified DR-DEGs as potential biomarkers for the diagnosis and classification of OA and provided a preliminary understanding of their role in the immune microenvironment. However, further experimental and clinical studies are required to validate their diagnostic value and therapeutic potential.

Keywords Osteoarthritis, Disulfidptosis, Chondrocyte, Single-cell sequencing, Diagnosis

Osteoarthritis (OA), a chronic degenerative bone disease, is characterized by the destruction of joint cartilage, pain in surrounding tissues, and stiffness caused by proliferative changes¹. OA is a key cause of chronic pain and disability, with the prevalence of pain and disability rates expected to rise annually due to an aging population^{2,3}.

Department of Orthopedic Surgery, 920th Hospital of Joint Logistics Support Force of PLA, Kunming, China.
✉email: xuyongqingkm@163.net

However, there are currently no definitive methods to delay or reverse disease progression in OA patients, which limits treatment options to early-stage drug interventions or late-stage surgical procedures⁴. Regrettably, early-stage OA patients also struggle to receive accurate and timely diagnoses. The most commonly used method for diagnosing OA is the Kellgren-Lawrence (KL) radiographic grading system, which differentiates OA levels based on the progressive narrowing and sclerotic appearance of joint spaces⁵. Additionally, magnetic resonance imaging (MRI) may be used to visualize microstructures in the joint to assist in OA diagnosis. However, not all patients with radiologically defined OA exhibit significant clinical symptoms, and not all patients with joint symptoms show radiographic specificity, making early diagnosis of OA extremely challenging⁶. At present, there is a lack of a specific diagnostic biomarker capable of achieving the goal of early OA diagnosis.

Recently, Liu et al.⁷ were the first to discover a novel form of cell death defined as “disulfidptosis”. Specifically, the actin cytoskeleton is susceptible to disulfide stress caused by the excessive accumulation of intracellular disulfide bond molecules, which in turn induces actin network collapse and cell death⁸. This unique mode of cell death presents interesting directions for targeted disease treatments. Previous studies have shown that disulfide death-related genes have been reported for the diagnosis and prediction of various types of cancer⁹, including lung adenocarcinoma¹⁰, hepatocellular carcinoma¹¹, and endometrial cancer¹². Multiple forms of programmed cell death, such as apoptosis¹³ and ferroptosis¹⁴, have been found to hold promising diagnostic value for OA patients. However, despite the growing interest in understanding the pathogenesis of OA, no study has reported the potential link between disulfide death-related genes and OA. Consequently, the discovery of this connection could further contribute to our understanding of OA progression and facilitate the development of more targeted treatment approaches.

This study identified disulfidptosis-related genes (DR-DEGs) through the integration of bulk RNA-seq data and machine learning algorithms, exploring their value in the diagnosis, and typing of OA patients. Additionally, using single-cell datasets, the study examined their distribution, expression, and potential biological functions across different cell types, and established an *in vivo* OA model to validate the expression of key DR-DEGs (Fig. 1). The aim of this study is to investigate the role and function of DR-DEGs as potential biomarkers for OA.

Materials and methods

920th Hospital of Joint Logistics Support Force Committee on Ethics approved this study and consented to participate (2024-53-01), and all experiments were performed in accordance with ARRIVE guidelines (<https://arriveguidelines.org>). All methods were performed in accordance with the relevant guidelines and regulations.

Data collection and procession

We obtained six transcriptome datasets of OA patients and healthy controls from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>), including GSE51588, GSE55235, GSE55457, GSE117999, and GSE169077. The GSE51588 dataset comprises 10 control samples and 40 OA patients; GSE55235 includes 10 control samples and 20 OA patients; GSE55457 consists of 10 control samples and 23 OA patients; GSE117999 contains 12 control samples and 12 OA patients; and GSE169077 consists of 5 control samples and 6 OA patients. Four datasets (GSE51588, GSE55235, GSE55457, and GSE77298) served as the training group, while the remaining two datasets (GSE117999 and GSE169077) were used as the validation group. Subsequently, gene expression data were extracted using Perl software, and gene names were converted into gene symbols. To reduce the errors caused by data analysis, the batch effects were eliminated using the “combat” function of the “sva” package¹⁵ after merging the files. Furthermore, the data consistency pre- and post-batch correction was demonstrated for the training group and control group through principal component analysis (PCA) analysis¹⁶.

Identification of disulfidptosis-related differentially expressed genes (DR-DEGs)

To extract DR-DEGs in OA patients, we utilized the “limma” package¹⁷ to perform differential expression analysis of genes between the OA and control groups, which included upregulated and downregulated genes. The DR-DEGs were visualized using heatmaps and box plots. The filtering criteria were set as p -value < 0.05 and $|\log_2\text{FC}| \geq 0.585$. Subsequently, a circos plot was used to display the chromosomal start and end positions of the DR-DEGs. Expression correlations between each DR-DEG were presented through correlation analysis. Lastly, the “ClusterProfiler” package¹⁸ was employed to conduct GO function and KEGG pathway enrichment analyses¹⁹ on DR-DEGs, further exploring their potential molecular mechanisms. Additionally, we used the GeneMANIA website (<http://genemania.org>) to predict interacting genes of DR-DEGs in OA patients.

Immune infiltration analysis

“CIBERSORT”²⁰ is a computational method commonly used for assessing immune cell abundance, which functions based on the gene expression dataset of 22 immune cell types. To evaluate changes in immune microenvironments in OA patients, immune scores were assigned to each OA and control sample. Subsequently, the differences in infiltration proportions of the 22 immune cell types between OA patients and control samples were comprehensively compared. Moreover, we further assessed the correlations between the expression of DR-DEGs and the infiltration levels of 22 immune cells, where $p < 0.05$ and $|R| > 0.30$ were considered to have significant correlations.

Identification of disulfidptosis-related subtyping patterns

As DR-DEGs have been demonstrated to be significantly associated with the immune microenvironment of OA patients, further investigation of the subtyping value of these DR-DEGs may prove beneficial in guiding stratified management and immunotherapy for OA. Initially, consensus clustering analysis was performed on OA samples using the “ConsensusClusterPlus” package²¹, and the stability of the clusters was assessed. Notably, an optimal K

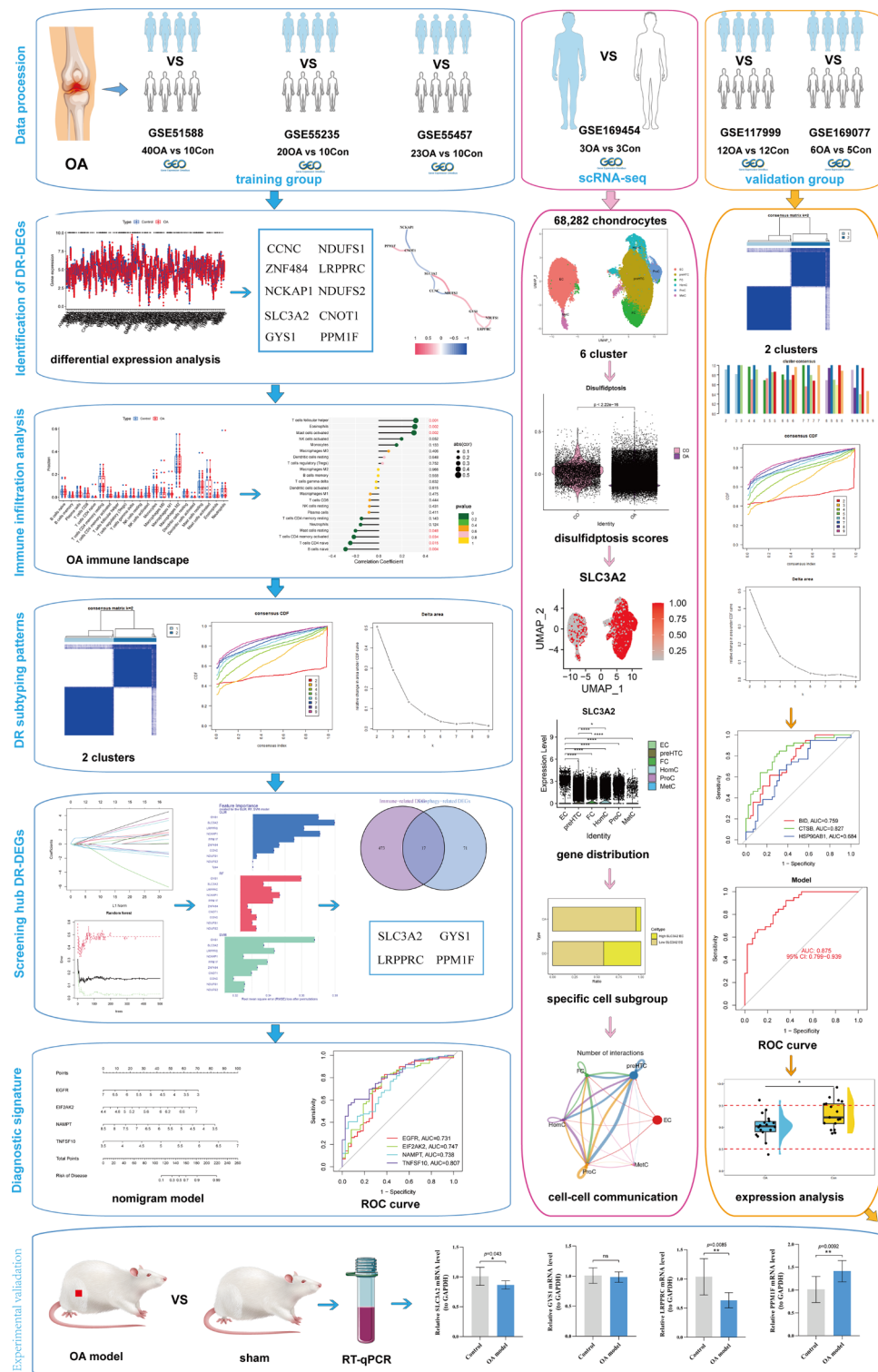


Fig. 1. Flowchart of this study.

value was determined through a combination of cluster heatmaps, consensus scores, and cumulative distribution function plots, with the range of K values set from 1 to 9. Subsequently, the discriminative capacity of different OA clusters was observed via principal component analysis, and the stability of disulfidptosis-related subtyping patterns was evaluated through an independent validation dataset. The expression differences of DR-DEGs in various OA subgroups were then assessed. Finally, the infiltration proportion differences of 22 immune cells across different OA subgroups were further evaluated based on immune scores established by “CIBERSORT”. Through this comprehensive analysis, insight into the potential subtyping value of DR-DEGs was provided,

which could contribute to personalized therapeutic strategies for OA patients based on their specific immune microenvironment and molecular characteristics.

Screening characteristic DR-DEGs based on multiple machine learning algorithms

To further screen for characteristic DR-DEGs, Random Forest (RF), Support Vector Machine (SVM), and Generalized Linear Models (GLM) were constructed using “DALEX”²², “randomForest”²³ and “kernlab” packages²⁴, and their performance was evaluated through ROC curves and residual values. Subsequently, the Least Absolute Shrinkage and Selection Operator (LASSO) regression model was established using the “glmnet” package²⁵, which resulted in a gene importance ranking. Additionally, a Venn diagram was utilized to display the intersecting genes obtained from the four machine learning algorithms, ultimately determining the characteristic DR-DEGs.

Single-gene GSEA (ssGSEA)

To investigate the potential molecular mechanisms of the characteristic DR-DEGs, OA samples were divided into groups based on the expression levels of each characteristic DR-DEG, resulting in high and low expression groups with their respective logFC values. Further KEGG-GSEA analysis was conducted according to the given KEGG gene set files. *p*-value of less than 0.05 indicated that the respective pathway was significantly enriched.

Processing of scRNA-seq data

The single-cell transcriptome data were obtained from the GSE169454 dataset, which included four knee OA patients and three healthy samples. Three OA patient cartilage samples (GSM5203392, GSM5203393, and GSM5203394) and three healthy cartilage samples (GSM5203389, GSM5203390, and GSM5203391) were chosen for further analysis. The processing of single-cell data was performed using the “Seurat” package²⁶, with the creation of six cartilage single-cell datasets as Seurat objects. Cells with fewer than 200 detected genes (min. features = 200) and genes covered by less than three cells (min. cells = 3) were filtered out. Additionally, cells with nFeatureRNA < 200 and > 10,000 were filtered to exclude situations involving cell death, empty droplets, or an overly large number of genes.

Next, the data were normalized using the “LogNormalize” method. The “FindVariableFeatures” function was employed to identify highly variable genes (HVGs) that exhibited substantial expression differences between cells. Moreover, the “vst” method was utilized to retain the top 2,000 HVGs for subsequent cell type identification. Prior to dimensionality reduction, the three OA samples and three control samples were integrated separately, and batch effects were removed using the “CCA” method. Then, all genes were normalized, and the top 1,500 HVGs were selected as the gene set for dimensionality reduction.

Cell classification and annotation

By comprehensively determining the PC values of the data using the JackStraw and Elbow methods, a *k*-nearest neighbors (KNN) graph based on Euclidean distance was constructed in the PCA space. The “FindNeighbors” and “FindClusters” functions were used for optimal cell clustering. After selecting the optimized resolution, the “UMAP” method was employed to visualize cell distribution in a low-dimensional space, which was divided into different cell clusters. Subsequently, manual annotation for different clusters was conducted based on cell type-specific expression markers²⁷, retaining only upregulated differentially expressed genes. The distribution of 18 disulfidptosis-related genes within cells was further investigated, and differential analysis of the expression of these genes across various cell types was performed.

Disulfidptosis gene set score and differential analysis

To investigate the role of disulfidptosis-related genes in OA patients and chondrocytes, the “AddModuleScore” function²⁸ of the “Seurat” package was used to calculate gene set scores based on 18 disulfidptosis regulators. Specifically, the average expression of the 18 disulfidptosis genes needed to be calculated first, followed by the calculation of the average expression of 100 randomly drawn control genes. The difference between the two averaged values constituted the disulfidptosis gene set score. Subsequently, the differences in the disulfidptosis scores between OA and control samples were examined, and the score differences between different cell types were further explored.

Cell-cell communication analysis

After creating a CellChat object²⁹, the ligand-receptor interaction database was set and imported. The expression data were preprocessed for use in cell communication analysis. Subsequently, the communication probabilities at the signaling pathway level were inferred by calculating the communication probabilities of all ligand-receptor interactions associated with each signaling pathway. The communication among different cell types was quantified through cell interaction counts and intensities.

Protein-protein and gene-miRNA interaction networks

We also investigated the interacting proteins and miRNAs with DR-DEGs to explore their potential molecular connections. First, a protein-protein interaction (PPI) network among DR-DEGs was established based on the “String” online database. In addition, a gene-miRNA network was constructed using the “NetworkAnalyst” online website (<https://www.networkanalyst.ca>). These network analyses help elucidate the potential molecular interactions and regulatory mechanisms involving DR-DEGs.

Diagnostic signature in OA

To evaluate the discriminative ability of the identified characteristic DR-DEGs in distinguishing between OA patients and normal samples, diagnostic and predictive models were established. First, a nomogram for characteristic DR-DEGs was constructed based on the “rms” and “rmda” packages³⁰, while clinical decision curves and calibration curves were employed to assess the value and fitting of the model. Next, ROC curves for each characteristic DR-DEG were generated using “pROC” package³¹ and the area under the curve (AUC) was calculated. Subsequently, a logistic regression model consisting of all characteristic DR-DEGs was built to assess the diagnostic performance. Additionally, an independent validation dataset was compiled, and ROC curves were created to evaluate diagnostic performance.

Expression analysis of characteristic DR-DEGs in the OA Validation Set

To validate the expression level differences of characteristic DR-DEGs identified in the bioinformatics results, we analyzed the expression levels of these characteristic DR-DEGs in an independent OA validation dataset (GSE117999 and GSE169077). The analysis allowed us to cross-check and confirm the relevance of these genes in distinguishing OA patients in independent datasets, strengthening the credibility of our findings.

OA animal model

To validate the expression of characteristic DR-DEGs in vivo, we established an OA rat model. A total of 12 male SD rats, 6–8 weeks old, were purchased from Kunming Chushang Co., Ltd. (Kunming, China), and randomly divided into 6 OA model rats and 6 control rats. The OA model was induced based on the anterior cruciate ligament transection (ACLT) surgery. Specifically, after anesthetizing both groups of rats with intraperitoneal injections of sodium pentobarbital, the left knee joint was fully exposed, followed by a small incision at the joint's skin along the patellar ligament. The control group's joint was washed with saline and then sutured layer by layer. In the OA model group, after exposing the cruciate ligament by peeling away muscle tissue, the cruciate ligament was cut, then washed with saline and sutured layer by layer. Four weeks later, the rats were euthanized using cervical dislocation, and the left knee joint cartilage tissue was collected for hematoxylin and eosin (HE) staining to verify the successful establishment of the OA model.

Reverse transcription-PCR (RT-PCR)

After the establishment of OA model, total RNA was extracted. The SweScrip RT I First Strand cDNA Synthesis Kit (Service Bio, Guangzhou, China) was used to reverse transcribe genes, completing the synthesis of the first cDNA strand. After a brief centrifugation, the reaction proceeded for 40 cycles under the following thermal cycling conditions using the CFX96 Real-Time PCR System (Bio-Rad, Hercules, CA, USA): pre-denaturation at 95 °C for 1 min, denaturation at 95 °C for 20 s, annealing at 55 °C for 20 s, and extension at 72 °C for 30 s. The primers were designed with the assistance of the NCBI Primer-BLAST Tool. The primer sequences can be found in supplementary table S1.

Western blot

Cartilage tissue was placed in a pre-chilled mortar, and an appropriate amount of liquid nitrogen was added. The tissue was ground into a fine powder. The powder was transferred to a centrifuge tube containing RIPA lysis buffer (with phosphatase inhibitors) and lysed on ice for 30 min. The lysed samples were centrifuged at 4 °C and 16,000 rpm for 15 min to remove cell debris. The supernatant (total protein) was collected and aliquoted into 80 µl samples. Next, the protein concentration of the supernatant was determined using a BCA protein assay kit. Following electrophoresis, the proteins were transferred from the gel to a PVDF membrane. The PVDF membrane was blocked with 5% (w/v) skim milk in TBST and incubated at room temperature (37 °C) for 60 min. Subsequently, the PVDF membrane was incubated overnight at 4 °C with a primary antibody diluted in 1% (w/v) skim milk in TBST. The membrane was then incubated with an HRP-conjugated secondary antibody at room temperature for 60 min. The antibody signal was detected using ECL chemiluminescent substrate. The grayscale values were analyzed using ImageJ software.

Statistical analysis

All data were processed and analyzed using R software (version 4.3.1). The Kruskal-Wallis's test was used to evaluate significant differences among two or more independent groups. The *t*-test or wilcoxon's rank-sum test was applied to examine significant differences between two independent groups. A *p*-value < 0.05 was considered statistically significant.

Results

Identification of DR-DEGs in OA

After merging and eliminating the batch effects of four OA training set samples and two OA validation set samples separately, we obtained 30 control samples and 83 OA samples in the training set, and 17 control samples and 18 OA samples in the validation set (Fig. 2A and B). Differential expression analysis results suggested that 10 DR-DEGs exhibited significant differences between OA and control samples (Fig. 2C and D). Next, we determined the chromosomal positions of these DR-DEGs; apart from LRPPRC, NCKAP1, and NDUFS1 located on chromosome 2, the remaining DR-DEGs were situated on different chromosomes: NDUFS2, CCNC, ZNF484, SLC3A2, CNOT1, GYS1, and PPM1F were on chromosomes 1, 6, 9, 11, 16, 19, and 22, respectively (Fig. 2E). Correlation analysis results indicated significant correlations among 9 DR-DEGs, with the expression levels of GYS1, NCKAP1, and LRPPRC presenting significant positive correlations (Fig. 2F). Subsequently, the GO functional analysis results revealed that the 10 DR-DEGs were enriched in NADH dehydrogenase activity in Biological Process (BP), mitochondrial respiratory chain complex in Cellular Component (CC), and

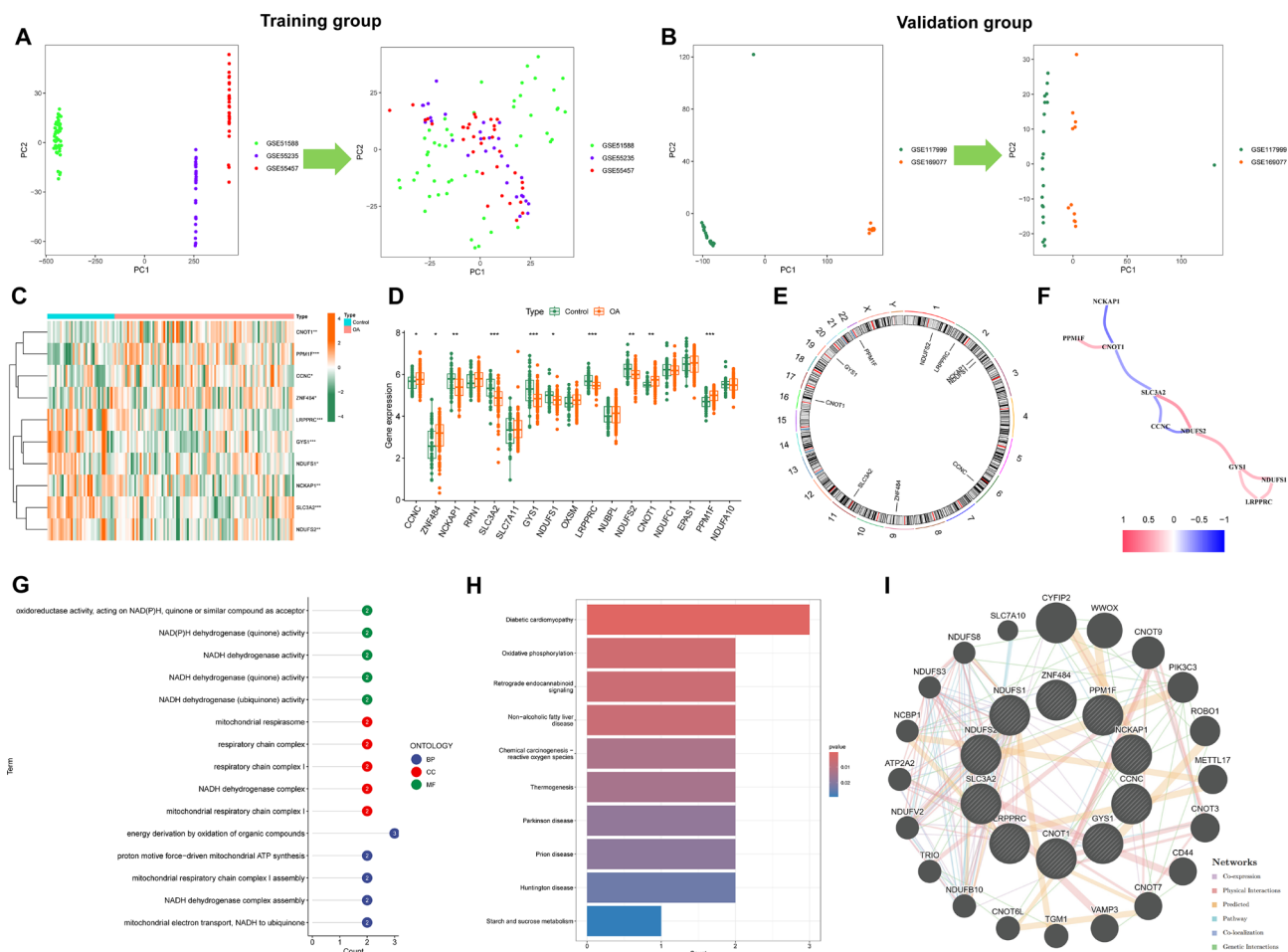


Fig. 2. Identification of disulfidptosis-related differentially expressed genes (DR-DEGs) in OA patients. (A) Principal component analysis (PCA) plots before and after batch correction in the training set. (B) PCA plots before and after batch correction in the validation set. (C) Heatmap of the 10 DR-DEGs. Green represents low expression genes in the samples, and orange represents high expression genes in the samples. (D) Differential expression analysis of the 17 DRGs. Green represents control samples, and orange represents OA samples. (E) Circos plot illustrating the chromosomal location distribution of the top 9 DR-DEGs. (F) Correlation analysis of the top 9 DR-DEGs. Blue indicates negative correlation, and red indicates positive correlation. (G) GO functional analysis of the 10 DR-DEGs, including BP, MF, and CC. (H) KEGG pathway analysis of the 10 DR-DEGs. (I) Gene-miRNAs interaction analysis of the 10 DR-DEGs. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

mitochondrial ATP synthesis in Molecular Function (MF) (Fig. 2G). These DR-DEGs were also involved in the regulation of pathways related to diabetic cardiomyopathy (Fig. 2H). Additionally, the 10 DR-DEGs had potential molecular connections with 20 genes, including co-expression, coexistence, and physical interactions (Fig. 2I).

Immune infiltration analysis in OA

Previous research has indicated that both cartilage and synovial tissues of OA patients are closely related to immune responses, and immune cells may be involved in the pathogenesis and progression of OA. Therefore, we calculated the immune scores for each OA and control sample based on the “CIBERSORT” algorithm and assessed the differences in the infiltration abundance of 22 immune cells between OA and control samples (Fig. 3A and B). Among the 22 immune cells, 5 showed significantly increased infiltration in OA samples, while 6 immune cell types exhibited significantly elevated abundance in the control samples. The correlation analysis of the 22 immune cells revealed that most immune cell types displayed negative associations with one another (Fig. 3C). Next, we further assessed the correlations between the 10 DR-DEGs and the 22 immune cells (Fig. 3D). The results demonstrated that 9 DR-DEGs, except for NDUFS2, were significantly correlated with immune cell infiltration levels (Fig. 3E-N). Specifically, ZNF484 was significantly correlated with 5 immune cells, SLC3A2 with 7 immune cells, PPM1F with 3 immune cells, NDUFS1 with 2 immune cells, NCKAP1 with 4 immune cells, LRPPRC with 5 immune cells, GYS1 with 8 immune cells, CNOT1 with 7 immune cells, and CCNC with 4 immune cells. GYS1, SLC3A2, and CNOT1 may play important roles in immune regulation in OA.

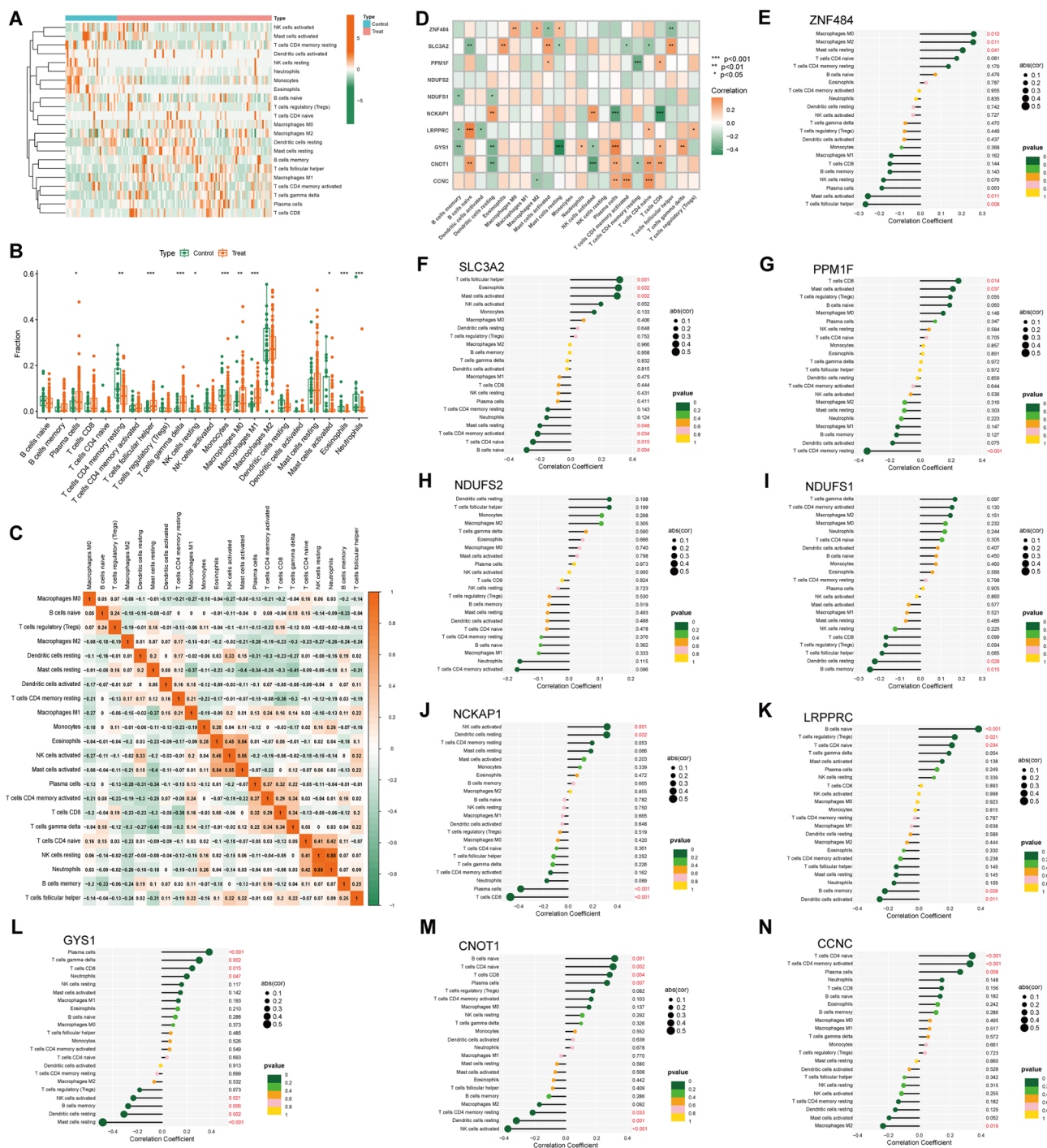


Fig. 3. Comprehensive immune landscape of OA patients. (A) Heatmap of the differences in 22 immune infiltrating cell types. Green represents low expression cells, and red represents high expression cells. (B) Box plot showing the expression differences of 22 different immune cells in OA and control samples. Green represents control samples, and red represents OA samples. (C) Correlation analysis of the infiltration levels of 22 immune cell types. Green indicates negative correlation, and red indicates positive correlation. (D) Correlation analysis of the expression levels of the 10 DR-DEGs with the infiltration levels of 22 immune cells. The correlation analysis of ZNF484 (E), SLC3A2 (F), PPM1F (G), NDUFS2 (H), NDUFS1 (I), NCKAP1 (J), LRPPRC (K), GYS1 (L), CNOT1 (M), and CCNC (N) with the infiltration levels of the 22 immune cells. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

Identification of DR clusters

The DR-DEGs demonstrated significant associations with the immune microenvironment of OA, prompting further investigation into the subtyping value of these ten DR-DEGs for OA patients. This could hold significant importance for personalized stratified management and immunotherapy in OA patients. First, through unsupervised clustering, OA patients were divided into distinct subgroups, and the optimal clustering pattern was determined with $K=2$ based on a combination of cluster heatmaps, consensus scores, CDF curves, and Delta area (Fig. 4A–D). The PCA plot showed clear discriminative capacity between subgroup 1 and subgroup 2 (Fig. 4E). Additionally, the subtyping pattern in the validation set based on the ten DR-DEGs also indicated that a clustering pattern with $K=2$ was optimal (Fig. 4F). Correlation analysis revealed significant differences for ZNF484 and NDUFS1 between the two OA subgroups, both of which exhibited significantly higher expression in subgroup 1 (Fig. 4G and H). The immune microenvironment analysis results of the two subgroups indicated that resting mast cells had significantly increased infiltration in OA subgroup 1, while activated mast cells showed substantially elevated infiltration abundance in subgroup 2 (Fig. 4I and J). This may suggest that mast cells are activated in OA subgroup 2, participating in immune regulation and immune responses.

Finally, we integrated the clinical features of 63 OA patients from two datasets, GSE51588 and GSE55457, which included gender and age (Fig. 4K). ZNF484 and NDUFS1 showed upregulated expression in OA cluster 1, while the expression levels of the 10 DR-DEGs exhibited no significant correlation with OA cluster, gender, or age.

Identification of characteristic DR-DEGs integrating machine learning algorithms

To further identify key DR-DEGs and analyze their potential roles, we screened for characteristic DR-DEGs through the integration of multiple machine learning algorithms. First, the importance scores of the 10 DR-DEGs were ranked using a random forest model (Fig. 5A and B). By integrating ROC curves, feature importance ranking, and residual values, three machine learning algorithms (RF, SVM, and GLM) were compared, with the top 5 importance-scored DR-DEGs being selected for each (Fig. 5C–F). Next, 8 DR-DEGs were screened using LASSO regression and 10-fold cross-validation (Fig. 5G and H). By identifying the intersection of the DR-DEGs selected by the four machine learning algorithms, four characteristic DR-DEGs (SLC3A2, GYS1, LRPPRC, and PPM1F) were ultimately obtained (Fig. 5I). GYS1 and LRPPRC showed a significant positive correlation (Fig. 5J).

Furthermore, we used ssGSEA to analyze the pathways that the four characteristic DR-DEGs might be enriched (Fig. 5K–N). The results suggested that SLC3A2, LRPPRC, and PPM1F were mainly enriched in obesity-related signaling pathways, while GYS1 was primarily enriched in cytokine-related signaling pathways.

Identification of OA chondrocyte cell types and disulfidptosis gene set scores

Single-cell data were obtained for three OA samples and three control samples from the GSE169454 dataset. After quality control and filtering, we obtained 68,282 chondrocytes from the cartilage samples of three OA patients and 6,455 chondrocytes from the three healthy control samples. The violin plots and correlation scatterplots for quality control demonstrated that the single-cell data filtering was reasonable. The top 2000 highly variable genes were identified for further analysis and visualization. After data normalization and PCA-based dimensionality reduction, an appropriate resolution was chosen, and seven cell clusters were identified, with UMAP visualization displaying the cellular distribution (Fig. 6A). Next, based on manual annotation, six distinct cell types were identified (Fig. 6B), and they were defined as follows. The heatmap of cell types shows the specifically expressed markers for each cell type, visualizing the significantly different marker genes in each cluster (Fig. 6C and D).

Using the “AddModuleScore” function to calculate the disulfidptosis gene set scores, differential expression analysis results revealed lower disulfidptosis scores in OA patients. In contrast, the EC subgroup had significantly higher disulfidptosis scores than the other five cell types (Fig. 6E and F). In the cellular distribution of 18 DRGs, SLC3A2 and NDUFC1 were primarily expressed in the preHTC subgroup (Fig. 6G). Further, in the differential expression analysis among different cell types, all 18 DRGs were significantly highly expressed in the EC subgroup (Fig. 7A–D), and apart from SLC3A2, which was expressed at lower levels in the OA group, the remaining 17 DRGs were expressed at significantly higher levels in the OA group (Fig. 7E).

Specific cell subgroup analysis and cell–cell communication patterns

Previous studies found that SLC3A2 might inhibit cartilage degeneration in OA patients, and it was significantly highly expressed in EC cells in this study. Therefore, we conducted further analysis on EC subgroups expressing SLC3A2. The proportion of EC subgroups with high SLC3A2 expression increased compared to control samples (Fig. 8A and B). Meanwhile, the expression levels of SLC3A2 were higher in different cell types in the OA group (Fig. 8C). KEGG enrichment analysis suggested that EC subgroups with high SLC3A2 expression were primarily enriched in the IL-17 signaling pathway and the TGF-beta signaling pathway (Fig. 8D). GO functional analysis indicated that EC subgroups with high SLC3A2 expression were significantly enriched in fibroblast proliferation in BP, extracellular matrix in CC, and ribosome composition in MF (Fig. 8E). Consistent with KEGG enrichment analysis results, GSEA enrichment analysis revealed that in the high SLC3A2 expression EC subgroup, inflammation-related pathways such as TGF-beta signaling pathway were significantly upregulated. Interestingly, myogenesis and angiogenesis-related signaling pathways were also significantly upregulated (Fig. 8F). These results suggest that SLC3A2 might regulate OA chondrocyte growth and development by modulating inflammatory responses, myogenesis, and angiogenesis.

Furthermore, cellular communication and interactions may affect cell differentiation processes, implying that communication between different cell types may influence the development and progression of OA. The statistics on the cellular interaction counts and intensities among the six cell types indicated that preHTC and EC have the highest cell numbers, with preHTC and proC demonstrating relatively strong communication intensity

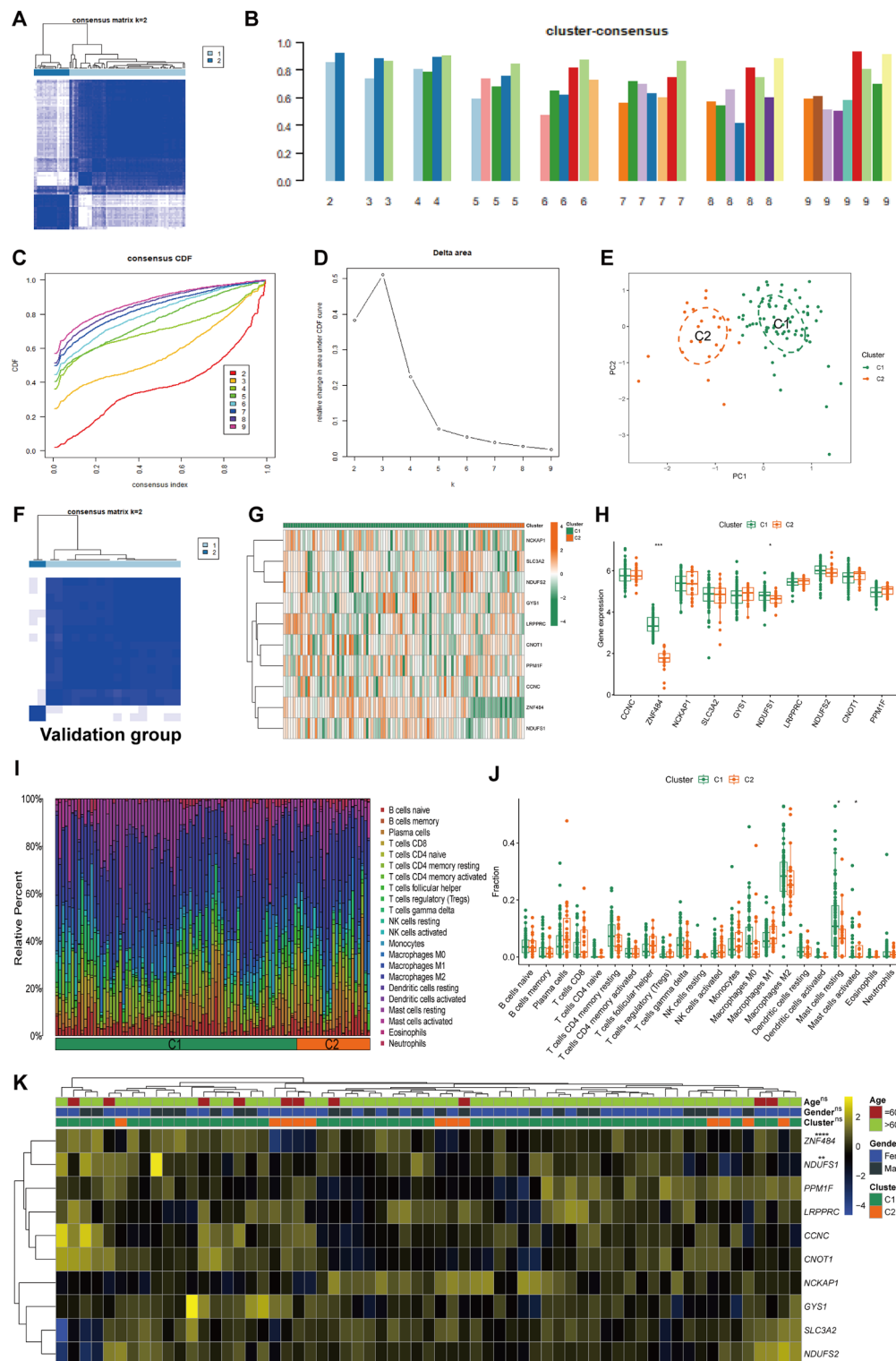


Fig. 4. Determining molecular subtypes of OA patients based on DR-DEGs. (A) Consensus matrix heatmap (cluster K = 2). (B) Consensus clustering scores (cluster K = 2–9). (C) Cumulative distribution functions of consistency (CDF) for the 2–9 clusters, illustrating the area under the curve. (D) Selecting an appropriate K value using delta area. (E) PCA analysis (cluster K = 2). (F) Consensus matrix heatmap in the OA validation cohort (cluster K = 2). (G) Heatmap of the expression of the 10 DR-DEGs in different samples. Green represents low expression genes, and orange represents high expression genes. (H) Expression differences of the 10 DR-DEGs between two different clusters. (I) Immune landscape of the two different clusters. (J) Differential infiltration proportions of 22 immune cells between the two different clusters. (K) Heatmap displays the relationship of 10 DR-DEGs and clinical features in OA patients (OA cluster, age and gender). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

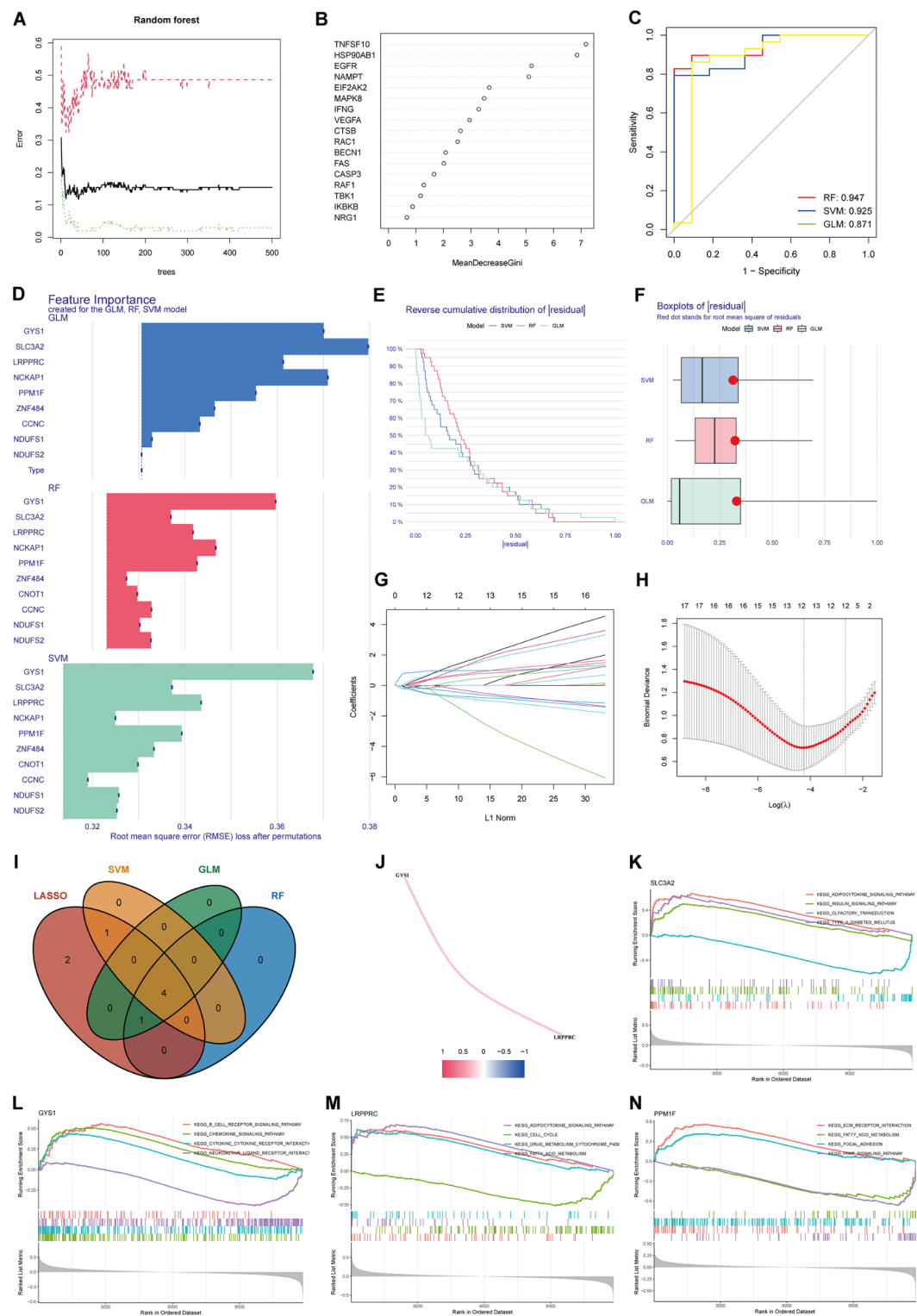


Fig. 5. Screening of key DR-DEGs using multiple machine learning algorithms. (A) Random forest (RF) model. (B) Ranking the importance of the 10 DR-DEGs based on MeanDecreaseGini values. (C) ROC curves for RF, SVM, and GLM. (D) Importance ranking of feature genes in RF, SVM, and GLM. (E) Reverse cumulative distribution of residual in RF, SVM, and GLM. (F) Box plot of residual in RF, SVM, and GLM. (G) Coefficient distribution plot of Lasso regression analysis. (H) Cross-validation curve. (I) Intersection of DR-DEGs from four machine learning algorithms. (J) Correlation analysis of GYS1 and LRPPRC. Single-gene GSEA in SLC3A2 (K), GYS1 (L), LRPPRC (M) and PPM1F (N).

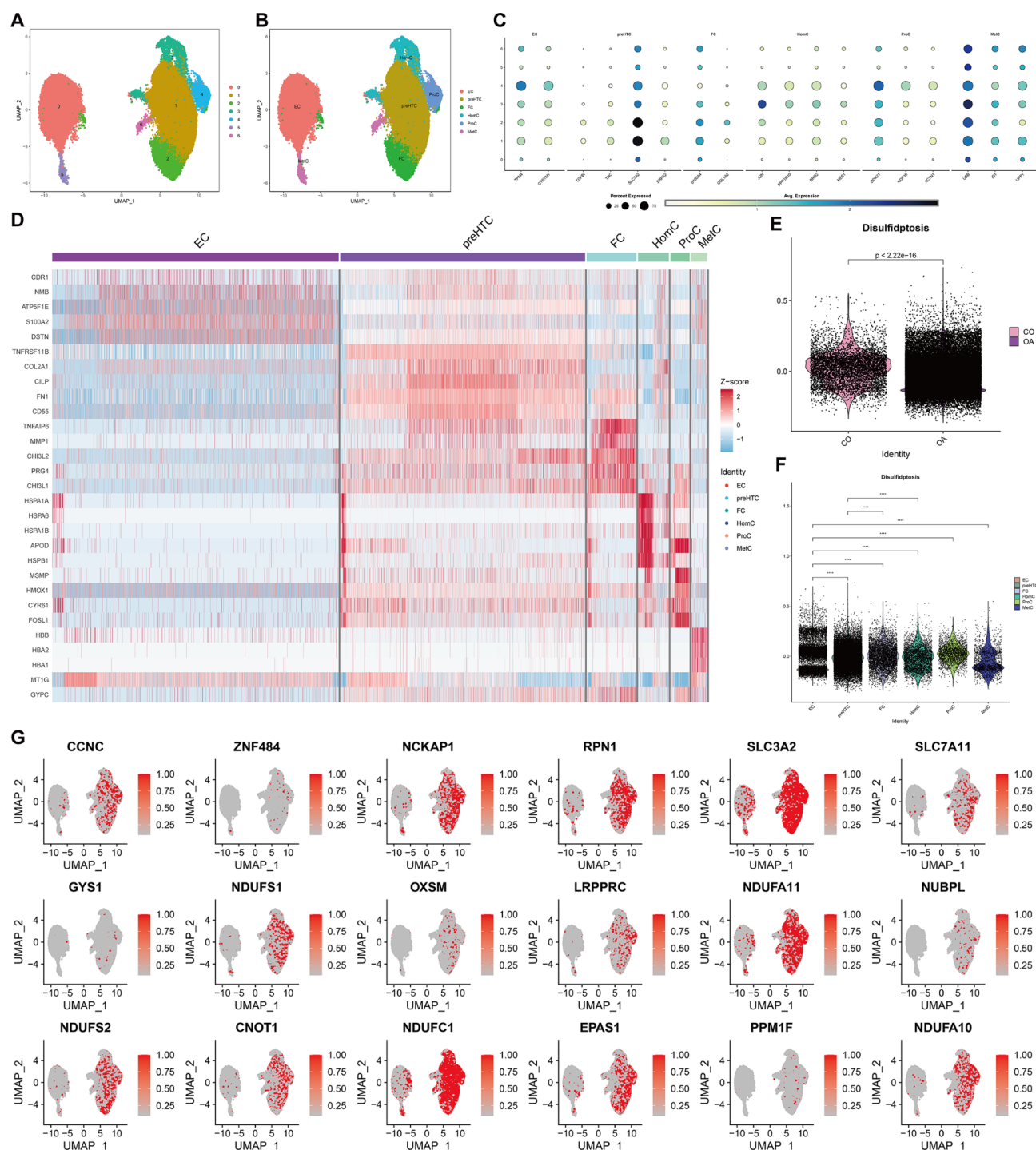


Fig. 6. Identification of chondrocyte subpopulations. (A) UMAP representation of the different 7 cell clusters (cluster = 0–6). (B) UMAP representation of the 6 distinct chondrocyte subpopulations with different features. (C) Specific markers for the 6 cell types. (D) Subtyping heatmap of the 6 chondrocyte subpopulations. (E) Differences in disulfidptosis gene set scores between OA and control samples. (F) Differences in disulfidptosis gene set scores among the 6 different chondrocyte cell types. (G) Violin plot showing the expression of the 18 DR-DEGs in OA chondrocytes. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

(Fig. 8G). The enriched signaling pathways between different cell types were further analyzed, with apoptosis-related signaling pathways being significantly upregulated in preHTC, HomC, and FC, but significantly downregulated in EC and MetC. Inflammation-related signaling pathways such as TGF- β -signaling pathway and TNF α -signaling-via-NF κ B were significantly upregulated in proC, preHTC, and HomC, while significantly downregulated in EC and MetC. Additionally, the inflammatory-response pathway was significantly enriched in

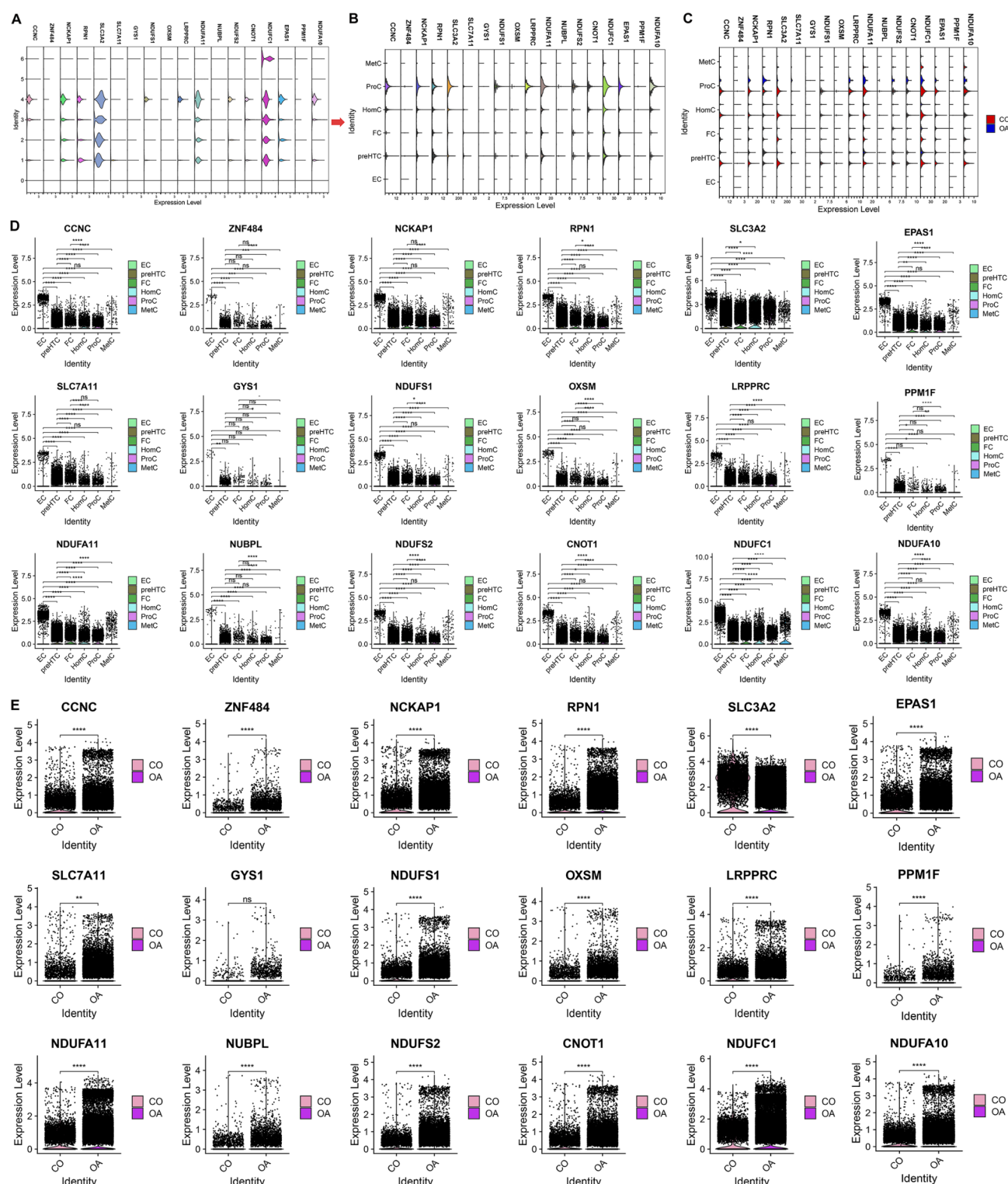


Fig. 7. Expression analysis of DR-DEGs in OA chondrocyte subpopulations. (A) Expression of the 18 DR-DEGs among the 7 cell clusters. (B) Expression of the 18 DR-DEGs among the 6 different OA chondrocyte subpopulations. (C) Composite violin plot showing the expression of the 18 DR-DEGs among the 6 different OA chondrocyte subpopulations in OA and control samples, respectively. (D) Expression differences of the 18 DR-DEGs among the different OA chondrocyte subpopulations. (E) Expression differences of the 18 DR-DEGs between OA and control samples. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

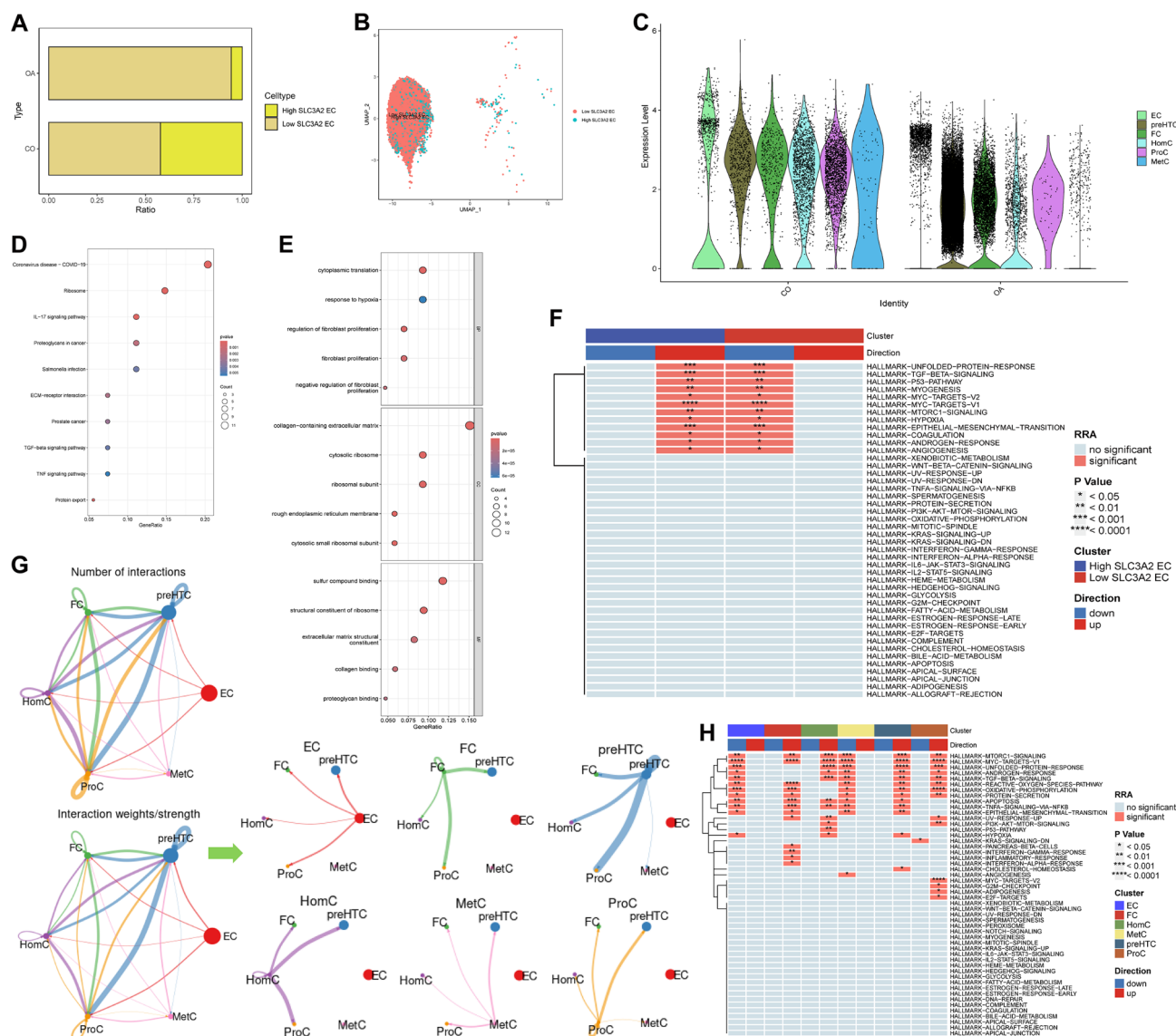


Fig. 8. Expression and functional analysis of SLC3A2 in ECs. (A) Proportion of High-TIMP1 EC and Low-TIMP1 EC groups in OA and control samples. (B) UMAP representation showing the gene distribution of High-TIMP1 EC and Low-TIMP1 EC groups. (C) Expression level of SLC3A2 among the 6 OA chondrocyte cell types in OA and control samples, respectively. (D) KEGG pathway analysis between High-TIMP1 EC and Low-TIMP1 EC groups. (E) GO functional analysis between High-TIMP1 EC and Low-TIMP1 EC groups. (F) Hallmark gene set enrichment analysis between High-TIMP1 EC and Low-TIMP1 EC groups. (G) Communication networks among different cell types. (H) Hallmark gene set enrichment analysis among different cell types. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

FC (Fig. 8H). These results suggest that cell apoptosis and inflammatory responses play essential roles in preHTC, HomC, and FC, whereas EC and MetC may participate in OA pathogenesis by downregulating inflammatory responses and inhibiting cell apoptosis.

PPI and gene-miRNA interaction networks

Based on the ten DR-DEGs, two protein interaction networks were obtained using the STRING online website (Fig. 9A). Further, the DR-DEGs-miRNA interaction networks were acquired through the “NetworkAnalyst” online website. Both CCNC and ZNF484 had interactions with two miRNAs, while the other eight DR-DEGs formed an interaction network with 270 miRNAs, comprising a total of 286 edges (Fig. 9B–D).

Diagnostic value of characteristic DR-DEGs in OA

OA patients lack specific early detection indicators, which significantly impacts their quality of life and prognosis. Therefore, we investigated the diagnostic and prognostic value of the four characteristic DR-DEGs for OA patients. Specifically, we first constructed a nomogram model, which displayed good fitting properties

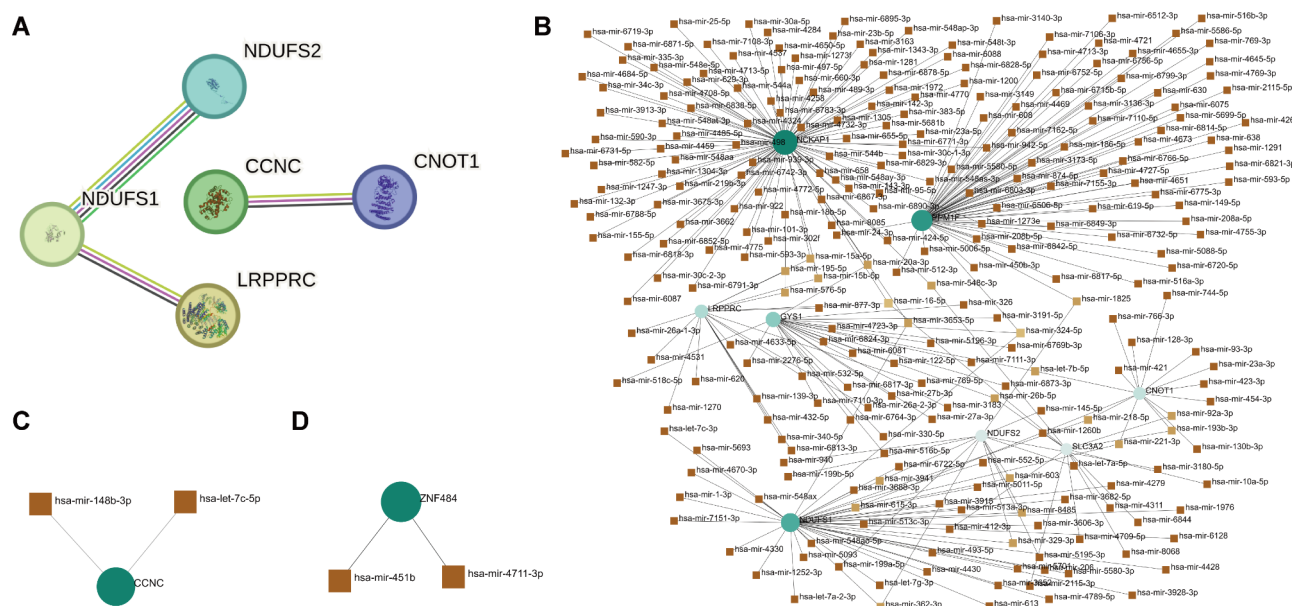


Fig. 9. Potential molecular mechanisms of key DR-DEGs. (A) Protein-protein interaction network. (B) Gene-TF-miRNA network based on key DR-DEGs. (C) ZNF484 and CCNC-miRNA interaction network.

and provided benefits for clinical patients (Fig. 10A–C). In the training set ROC curves, the AUC values of the four characteristic DR-DEGs were all greater than 0.700, and the AUC value for the logistic regression model was 0.820 (Fig. 10D and E). In the validation set ROC curves, the AUC values for SLC3A2 and LRP13C were both greater than 0.700, and the AUC value for the logistics regression model was 0.873 (Fig. 10F and G).

Validation of expression of characteristic DR-DEGs in validation set and OA animal models

Considering the significant value of four characteristic DR-DEGs in the diagnosis and molecular typing of OA, we utilized a validation cohort from the GEO database to explore the expression differences of these DR-DEGs between OA and healthy controls. Consistent with the training set results, SLC3A2 was highly expressed in the control group, while PPM1F was highly expressed in OA. Although there were no significant differences in the expression of GYS1 and LRP13C between the two groups, both tended to be higher in the control group (Fig. 10H). Further validation of the expression of the four characteristic DR-DEGs in OA tissues through RT-qPCR revealed that except for GYS1, whose mRNA levels showed no significant difference between groups, the other three DR-DEGs exhibited significant expression differences. SLC3A2 and LRP13C were downregulated, while PPM1F was overexpressed in the OA model (Fig. 11A), confirming the bioinformatics analysis results.

To verify the consistency of characteristic DR-DEGs expression in vivo, the protein expression levels in the OA model group and the control group were compared (Fig. 11B). The results of Western Blot analysis indicated that the protein level of SLC3A2 was significantly downregulated in the OA model group. Although the protein expression level of GYS1, LRP13C and PPM1F showed no significant difference between the two groups, it was consistent with the RT-qPCR results. These results suggest that the trends of mRNA and protein expression levels of GYS1, SLC3A2 and PPM1F in vivo are consistent, indicating their potential as therapeutic targets for OA.

Discussion

OA is a common degenerative bone disease characterized by joint cartilage destruction, synovitis, and subchondral bone sclerosis³². With the acceleration of population aging and increasing life expectancy, the prevalence of OA is expected to rise annually. A previous study suggested that by 2030, OA could become the most significant cause of disability worldwide³³. However, most early-stage OA patients exhibit subtle clinical symptoms, making traditional X-ray diagnosis rather challenging. MRI can provide detailed joint structure images, but its high cost limits its initial application^{34,35}. The diagnostic delay in early-stage OA patients leads to poor drug treatment outcomes or even the need for end-stage joint replacement surgery³⁶. Therefore, early diagnosis of OA is crucial for patient treatment and prognosis. Identifying reliable biomarkers for the early and rapid diagnosis of OA is of great significance, as it may help delay OA progression and improve quality of life. With the rapid development of bioinformatics and high-throughput sequencing technologies^{37,38}, elucidating the mechanisms of OA pathogenesis and progression from a molecular perspective holds significant potential for OA diagnosis and treatment.

This study is the first to investigate the regulatory role of the novel cell death mode-disulfidptosis in OA, and further explore its value in OA subtyping and diagnostic efficacy. We identified 10 DR-DEGs with significant differences from 18 DRGs and used unsupervised clustering analysis to construct an OA subtyping model, revealing two distinct DR patterns. The expression levels of ZNF484 and NDUFS1 showed significant differences

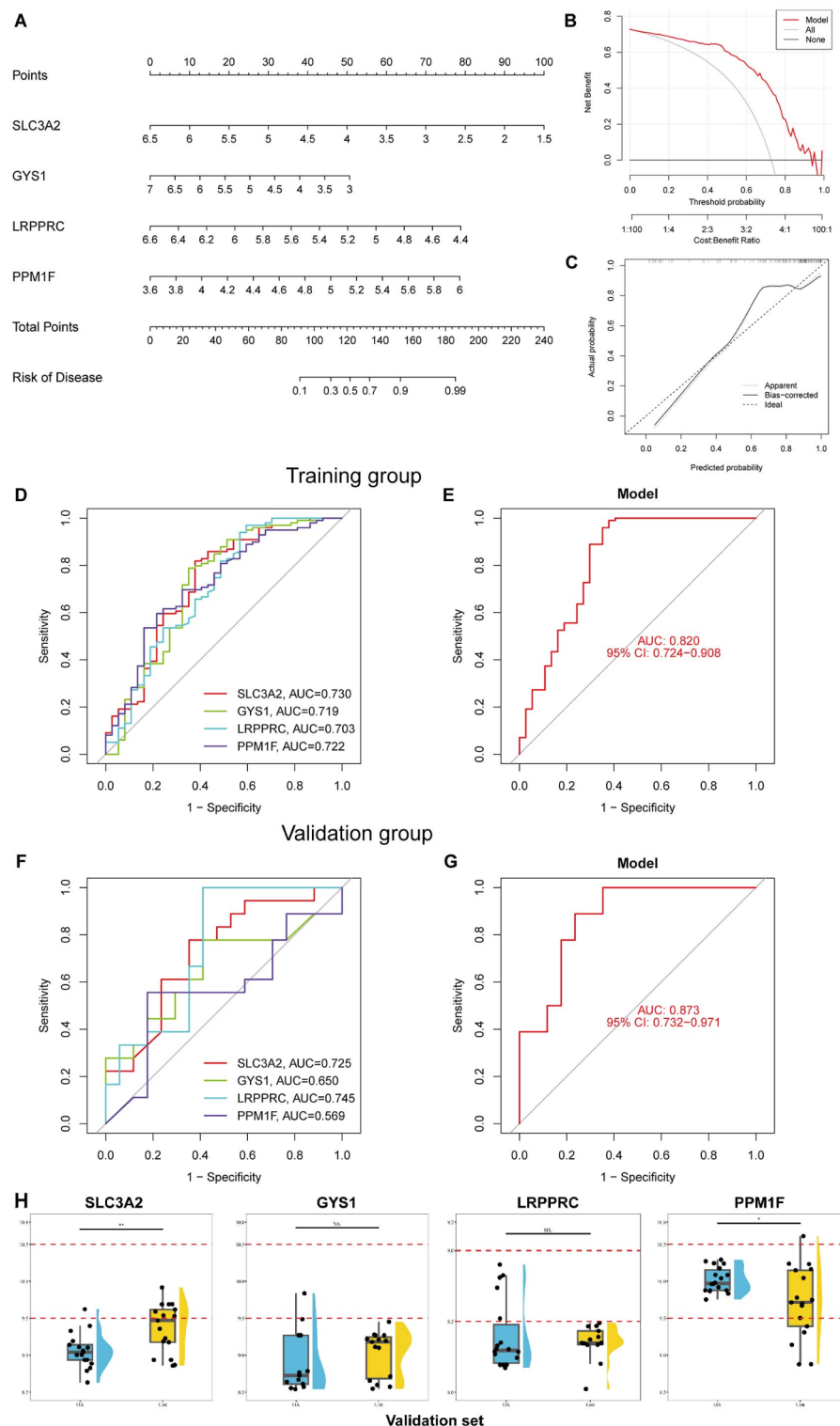


Fig. 10. Nomogram model and ROC curve based on characteristic DR-DEGs in OA training sets. (A) Nomogram model for OA patients. (B) Clinical decision curve for OA patients. (C) Fitted curves in OA training sets. (D) ROC curve of key DR-DEGs in OA training sets. (E) A Logistic regression model was constructed based on the OA training set. (F) ROC curve of key DR-DEGs in OA validation sets. (G) A Logistic regression model was constructed based on the OA validation set. (H) The expression levels of SLC3A2, GYS1, LRPPRC, and PPM1F in the OA validation set.

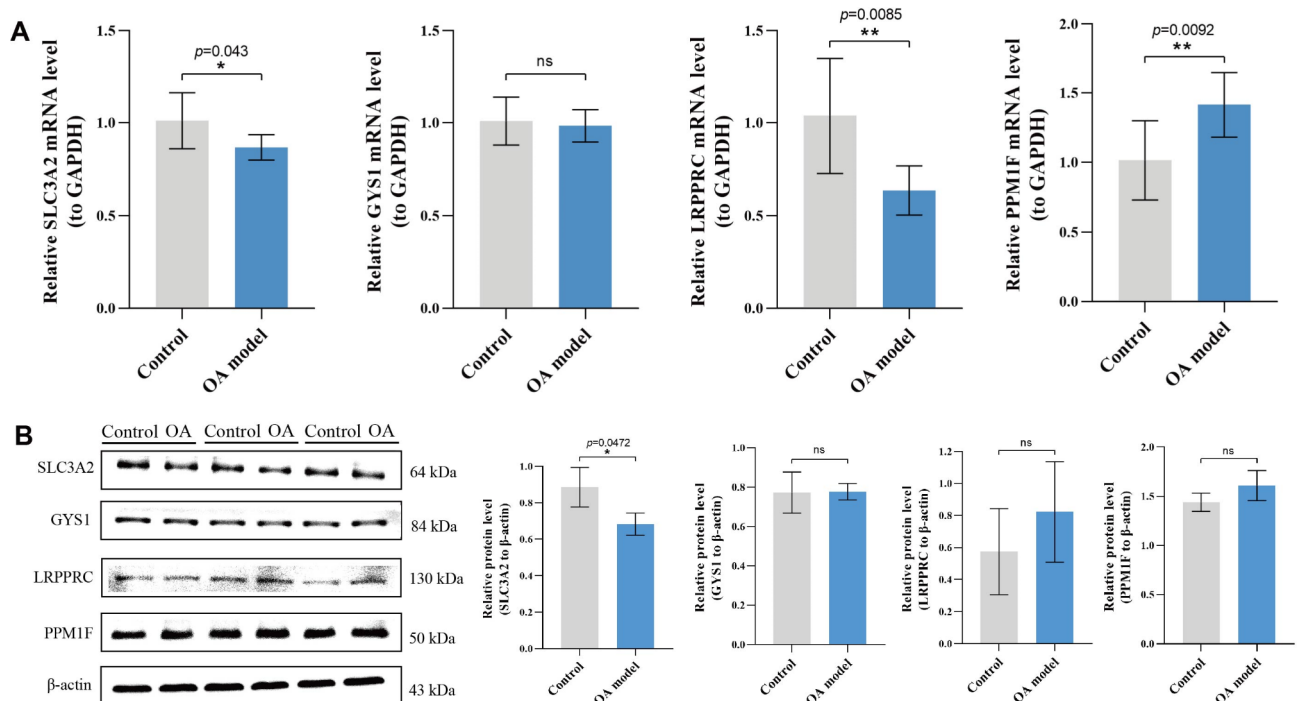


Fig. 11. RT-qPCR validation of key DR-DEGs. (A) The expression levels of SLC3A2, GYS1, LRPPRC, and PPM1F in the OA models. (B) The protein expression levels of SLC3A2, GYS1, LRPPRC, and PPM1F in the OA model group and the control group were compared.

between the two OA subgroups, with both being significantly overexpressed in subgroup (1) Immune-related analysis revealed that resting mast cells had significant infiltration in OA subgroup 1, while activated mast cells demonstrated substantially elevated infiltration abundance in subgroup (2) This may suggest mast cells were activated in OA subgroup 2, participating in immune regulation and immune responses. Previous studies have shown that synovium and cartilage of OA patients involve complex immune regulations and immune responses^{39,40}. Consistent with our results, we comprehensively analyzed the immune microenvironment of OA patients and found that the infiltration proportions of 11 immune cells changed in OA, with five immune cells significantly upregulated. Interestingly, we also found that the expression of ZNF484 was significantly positively correlated with macrophages M2 and resting mast cells, and significantly negatively correlated with activated mast cells. These results suggest that the abnormal expression of ZNF484 and NDUFS1 may be partly responsible for the disordered immune function in OA cartilage, and the low expression of ZNF484 in OA subgroup 2 may lead to the activation of mast cells. However, further mechanistic experiments are required for validation.

Additionally, we utilized scRNA data to explore the chondrocyte cell types in OA patients from a microscopic perspective. After comprehensive analysis of expressed markers, we identified six cell types: EC, preHTC, FC, HomC, ProC, and MetC. We further found that SLC3A2 was more widely distributed in EC and preHTC cells, prompting us to explore the cell cycle and biological function of SLC3A2. The proportion of EC subgroups with high SLC3A2 expression increased, primarily enriched in inflammation-related pathways such as the IL-17 signaling pathway and TGF- β signaling pathway. Previous studies have shown that TGF- β plays a critical role in chondrocyte proliferation and the synthesis and release of extracellular matrix proteins⁴¹. Downregulation of TGF- β receptor expression in mice promotes the progression of OA⁴². Additionally, IL-17 exacerbates joint damage by promoting inflammatory responses in OA patients⁴³. This suggests that high SLC3A2 expression might influence EC pathophysiological processes by regulating inflammatory responses. Liu et al.⁴⁴ integrated bioinformatics and clinical results, founding that SLC3A2 was downregulated in OA patients, further in vitro experiments showed that SLC3A2 knockdown could promote cartilage degeneration, making it a potential OA treatment target. In this study, we also found that SLC3A2 was significantly downregulated in OA samples, a result consistent in subsequent OA animal models. Interestingly, Karouzakis et al.⁴⁵ found that the expression of SLC3A2 is significantly higher in rheumatoid arthritis (RA) synovial fibroblasts compared to osteoarthritis (OA) synovial fibroblasts. Jiao et al.⁴⁶ further experimentally validated the significantly elevated expression of SLC3A2 in RA tissues, demonstrating that knocking down SLC3A2 inhibits the migration and proliferation of human fibroblast-like synoviocytes-RA (FLS-RA). These results suggest that SLC3A2 exhibits heterogeneity and different functions in various diseases such as OA and RA, highlighting its potential as a valuable therapeutic target. The specific mechanisms related to the function of SLC3A2 need to be further explored in disease models of OA and RA.

It is worth noting that the diagnostic value of immune and cell death-related genes in OA has been widely studied. Ren et al.³⁹ constructed a diagnostic model for OA patients based on four immunogenic cell death-related genes, with AUC values all greater than 0.700. We identified four characteristic DR-DEGs through

machine learning, and ROC results indicated that the four DR-DEGs (SLC3A2, GYS1, LRPPRC, and PPM1F) have good diagnostic efficacy for OA in the training set (AUC > 0.700), with an AUC value of 0.820 for the logistics regression model of the four DR-DEGs. In the validation cohort, the AUC value was as high as 0.873. This suggests that the four DR-DEGs can effectively differentiate OA from healthy samples, providing new molecular targets and insights for early diagnosis of OA patients. While the models exhibit strong predictive performance in both the training and validation cohorts, the lack of large-scale, independent clinical datasets limits their translational potential. Future studies should validate these models in diverse patient populations, ideally using prospective clinical cohorts and additional tissue samples such as synovium or serum.

Despite these findings, there are some limitations to this study. Firstly, the limited number of OA datasets in the GEO database constrains the reliability of the diagnostic model, and this should be interpreted cautiously. Future studies should aim to include diverse patient cohorts and integrate prospective clinical data to assess the robustness of the identified diagnostic markers. Secondly, we conducted a preliminary exploration of the immunological relevance and subtyping value of SLC3A2, but the underlying molecular mechanisms still require further experimental validation. Future studies should focus on functional validation of DR-DEGs, such as using gene knockdown or overexpression models to explore their effects on inflammatory signaling pathways and immune cell interactions in OA. Finally, in terms of immune heterogeneity in synovium and cartilage, a larger single-cell dataset is needed for support.

Conclusion

This study systematically explored the diagnostic and classification value of disulfidptosis-related genes in OA cartilage samples and identified four key genes (SLC3A2, GYS1, LRPPRC, and PPM1F) as diagnostic markers for OA patients. However, the connection between disulfidptosis-related genes and OA immune and inflammatory responses still requires further mechanistic experimental validation.

Data availability

The microarray data used to support the findings of this study can be downloaded from the GSE51588, GSE55235, GSE55457, GSE117999, GSE167077 and GSE169454 datasets (<https://www.ncbi.nlm.nih.gov/geo>). The processed data are available from the corresponding author upon request.

Received: 26 April 2024; Accepted: 3 January 2025

Published online: 08 January 2025

References

- Perruccio, A. V. et al. Osteoarthritis year in review 2023: epidemiology & therapy. *Osteoarthr. Cartil.* (2023).
- Jang, S., Lee, K. & Ju, J. H. Recent updates of diagnosis, pathophysiology, and treatment on Osteoarthritis of the knee. *Int. J. Mol. Sci.* **22**, 5 (2021).
- Mandl, L. A. Osteoarthritis year in review 2018: clinical. *Osteoarthr. Cartil.* **27**(3), 359–364 (2019).
- Sharma, L. et al. Knee tissue lesions and prediction of incident knee osteoarthritis over 7 years in a cohort of persons at higher risk. *Osteoarthr. Cartil.* **25**(7), 1068–1075 (2017).
- Zhang, Y. & Jordan, J. M. Epidemiology of osteoarthritis. *Clin. Geriatr. Med.* **26**(3), 355–369 (2010).
- Hannan, M. T., Felson, D. T. & Pincus, T. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. *J. Rheumatol.* **27**(6), 1513–1517 (2000).
- Liu, X. et al. Actin cytoskeleton vulnerability to disulfide stress mediates disulfidptosis. *Nat. Cell. Biol.* **25**(3), 404–414 (2023).
- Liu, X., Zhuang, L. & Gan, B. Disulfidptosis: disulfide stress-induced cell death. *Trends Cell. Biol.* (2023).
- Zhao, D. et al. Molecular landmarks of tumor disulfidptosis across cancer types to promote disulfidptosis-target therapy. *Redox Biol.* **68**, 102966 (2023).
- Ma, X. et al. Leveraging a disulfidptosis/ferroptosis-based signature to predict the prognosis of lung adenocarcinoma. *Cancer Cell. Int.* **23**(1), 267 (2023).
- Zhang, C. et al. Development and experimental validation of a machine learning-based disulfidptosis-related ferroptosis score for hepatocellular carcinoma. *Apoptosis* (2023).
- Shi, S., Tang, X. & Liu, H. Disulfidptosis-related lncRNA for the establishment of Novel Prognostic Signature and therapeutic response prediction to Endometrial Cancer. *Reprod. Sci.* (2023).
- Chen, Y., Zhang, Y., Ge, Y. & Ren, H. Integrated single-cell and bulk RNA sequencing analysis identified pyroptosis-related signature for diagnosis and prognosis in osteoarthritis. *Sci. Rep.* **13**(1), 17757 (2023).
- Bi, S., Han, B., Fan, H., Liu, Y. & Cui, X. Mitochondria-Related Gene MAOB is a key biomarker of Osteoarthritis and Inhibition of its expression reduces LPS-induced chondrocyte damage. *Biochem. Genet.* (2023).
- Zhang, Y., Parmigiani, G. & Johnson, W. E. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genom. Bioinform.* **2**(3), lqaa078 (2020).
- David, C. C. & Jacobs, D. J. Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol. Biol.* **1084**, 193–226 (2014).
- Ritchie, M. E. et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**(7), e47 (2015).
- Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* **16**(5), 284–287 (2012).
- Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51**(D1), D587–d92 (2023).
- Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M. & Alizadeh, A. A. Profiling Tumor infiltrating Immune cells with CIBERSORT. *Methods Mol. Biol.* **1711**, 243–259 (2018).
- Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**(12), 1572–1573 (2010).
- Wei, Z. et al. Using machine learning approach to predict depression and anxiety among patients with epilepsy in China: a cross-sectional study. *J. Affect. Disord.* **336**, 1–8 (2023).
- Wang, J. & Shi, L. Prediction of medical expenditures of diagnosed diabetics and the assessment of its related factors using a random forest model, MEPS 2000–2015. *Int. J. Qual. Health Care* **32**(2), 99–112 (2020).

24. Scharl, T., Grü, B. & Leisch, F. Mixtures of regression models for time course gene expression data: evaluation of initialization and random effects. *Bioinformatics* **26**(3), 370–377 (2010).
25. Li, Y., Lu, F. & Yin, Y. Applying logistic LASSO regression for the diagnosis of atypical Crohn's disease. *Sci. Rep.* **12**(1), 11340 (2022).
26. Yip, S. H., Sham, P. C. & Wang, J. Evaluation of tools for highly variable gene discovery from single-cell RNA-seq data. *Brief. Bioinform.* **20**(4), 1583–1589 (2019).
27. Ji, Q. et al. Single-cell RNA-seq analysis reveals the progression of human osteoarthritis. *Ann. Rheum. Dis.* **78**(1), 100–110 (2019).
28. Liu, J., Shi, Y. & Zhang, Y. Multi-omics identification of an immunogenic cell death-related signature for clear cell renal cell carcinoma in the context of 3P medicine and based on a 101-combination machine learning computational framework. *Epma J.* **14**(2), 275–305 (2023).
29. Almet, A. A., Cang, Z., Jin, S. & Nie, Q. The landscape of cell-cell communication through single-cell transcriptomics. *Curr. Opin. Syst. Biol.* **26**, 12–23 (2021).
30. Zhang, Z. et al. Decision curve analysis: a technical note. *Ann. Transl. Med.* **6**(15), 308 (2018).
31. Robin, X. et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **12**, 77 (2011).
32. Zeng, D. et al. Cartilage organoids and osteoarthritis research: a narrative review. *Front. Bioeng. Biotechnol.* **11**, 1278692 (2023).
33. Thomas, E., Peat, G. & Croft, P. Defining and mapping the person with osteoarthritis for population studies and public health. *Rheumatol. (Oxf.)* **53**(2), 338–345 (2014).
34. Fatema, K. et al. Development of an automated optimal distance feature-based decision system for diagnosing knee osteoarthritis using segmented X-ray images. *Heliyon* **9**(11), e21703 (2023).
35. Piccolo, C. L., Mallio, C. A., Vaccarino, F., Grasso, R. F. & Zobel, B. B. Imaging of knee osteoarthritis: a review of multimodal diagnostic approach. *Quant. Imaging Med. Surg.* **13**(11), 7582–7595 (2023).
36. Roelofs, A. J. & De Bari, C. Osteoarthritis year in review 2023. *Biol. Osteoarthr. Cartil.* (2023).
37. Conte, M. I., Fuentes-Trillo, A. & Domínguez Conde, C. Opportunities and tradeoffs in single-cell transcriptomic technologies. *Trends Genet.* (2023).
38. Yadav, D. et al. Next-generation sequencing transforming clinical practice and precision medicine. *Clin. Chim. Acta* **551**, 117568 (2023).
39. Ren, T., Yin, N., Du, L., Pan, M. & Ding, L. Identification and validation of FPR1, FPR2, IL17RA and TLR7 as immunogenic cell death related genes in osteoarthritis. *Sci. Rep.* **13**(1), 16872 (2023).
40. Sengprasert, P., Kamenkit, O., Tanavalee, A. & Reantragoon, R. The immunological facets of chondrocytes in osteoarthritis: a narrative review. *J. Rheumatol.* (2023).
41. Xia, C. et al. TGF- β /Smad2 signalling regulates enchondral bone formation of Gli1(+) periosteal cells during fracture healing. *Cell. Prolif.* **53**(11), e12904 (2020).
42. Shen, J. et al. Deletion of the transforming growth factor β receptor type II gene in articular chondrocytes leads to a progressive osteoarthritis-like phenotype in mice. *Arthritis Rheum.* **65**(12), 3107–3119 (2013).
43. Kamiab, Z. et al. Role of Interleukin-17 family cytokines in disease severity of patients with knee osteoarthritis. *Adv. Rheumatol.* **64**(1), 11 (2024).
44. Liu, H. et al. Identification of SLC3A2 as a potential therapeutic target of Osteoarthritis involved in Ferroptosis by Integrating Bioinformatics, clinical factors and experiments. *Cells* **11**, 21 (2022).
45. Karouzakis, E., Gay, R. E., Gay, S. & Neidhart, M. Increased recycling of polyamines is associated with global DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. *Arthritis Rheum.* **64**(6), 1809–1817 (2012).
46. Jiao, W. et al. Anti-proliferation and anti-migration effects of Yishen Tongbi decoction in experimental rheumatoid arthritis by suppressing SLC3A2/integrin β 3 signaling pathways. *Phytomedicine* **114**, 154741 (2023).

Acknowledgements

We are grateful to the authors who gave the GEO public dataset.

Author contributions

M W and Y X conceived the study and wrote the manuscript. X S and W T carried out the data collection and data analysis. Q L and Y W contributed to the data curation, methodology, and validation. All authors reviewed the results and approved the final version of the manuscript. All authors agreed to the publication.

Funding

This study was supported by the grants from Yunnan Orthopedics and Sports Rehabilitation Clinical Medicine Research Center (Grant No. 202102AA310068), Yunnan Clinical Medical Center for Traumatology and Orthopedics (Grant No. ZX20191001), and the Scientific Research Fund Project Department of Education of Yunnan Province (Grant No. 2024Y251).

Competing interests

The authors declare no competing interests.

Informed consent

920th Hospital of Joint Logistics Support Force Committee on Ethics approved this study and consented to participate (2024-053-01), and all methods are reported in accordance with ARRIVE guidelines.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-85569-z>.

Correspondence and requests for materials should be addressed to Y.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025