Identification of inflammation-related biomarkers for osteoarthritis diagnosis and stratification through bulk and single-cell RNA-sequencing integration

JIE XIAO[1], WEIQING WANG[2], XIAOTIAN LI[1], SHIWEI XU[3], BI ZHANG[1] and XIN LIAO[1]

[1]Department of Orthopedics, The First People's Hospital of Jiande, Jiande, Zhejiang 311600, P.R. China;
[2]Department of Neurology, The First People's Hospital of Jiande, Jiande, Zhejiang 311600, P.R. China;
[3]Department of Equipment, The First People's Hospital of Jiande, Jiande, Zhejiang 311600, P.R. China

Received November 18, 2025; Accepted March 4, 2026

DOI: 10.3892/etm.2026.13164

Abstract. Osteoarthritis (OA) is a degenerative joint disease, which cannot be cured with present treatment methods. Increasing evidence implicates inflammation in OA pathogenesis, and this has led to investigations into inflammation-related biomarkers (TNF-a, IL-6 and IL-1b) that may guide diagnosis and targeted interventions. In the present study, a total of six Gene Expression Omnibus transcriptomic datasets, including 60 OA and 39 normal samples, were analyzed. Intersecting differentially expressed genes (DEGs) with an inflammatory response gene set defined inflammation-related DEGs (IRDGs). Consensus clustering, gene set variation analysis and estimation of stromal and immune cells in malignant tumor tissues using expression data/cell-type identification by estimating relative subsets of RNA transcripts immune profiling were also carried out. The features were reduced by least absolute shrinkage and selection operator (LASSO) and modeled using generalized linear models, random forests, support vector machines (SVM) and extreme gradient boosting. Single-cell data were subjected to Seurat clustering, SingleR annotation, Monocle pseudotime and Gene Ontology/Kyoto Encyclopedia of Genes and Genomes enrichment. Additionally, following modulation of low-density lipoprotein receptor (LDLR), adrenomedullin (ADM), MYC or NF-κB inhibitor-α (NFKBIA), the viability and apoptosis of ATDC5 cells were assessed. Subsequently, a total of 537 DEGs and 11 IRDGs were identified in the present study. In addition, two OA subtypes, cluster C1 and C2, were identified. Presenilin 1 expression was increased in cluster C2, while the expression of the other IRDGs was upregulated in cluster C1. Only the stromal scores differed significantly. LASSO and machine learning nominated four biomarkers, MYC, ADM, LDLR and NFKBIA, with SVM providing the best overall and robust external validation. Single-cell analysis of 1,464 chondrocytes revealed broad NFKBIA expression across nine subpopulations. Furthermore, it was demonstrated that downregulation of LDLR, ADM, MYC and NFKBIA reduced cell viability and induced apoptosis in ATDC5 chondrocytes. Integrative bulk single-cell transcriptomics and machine learning identified MYC, ADM, LDLR and NFKBIA as inflammation-associated OA biomarkers, revealing subtype-specific immune heterogeneity. Clinically, this signature may possibly enable earlier diagnosis, patient stratification and targeted interventions to slow cartilage degeneration.

Correspondence to: Dr Xin Liao, Department of Orthopedics, The First People's Hospital of Jiande, 599 Yanzhou Street, Jiande, Zhejiang 311600, P.R. China
E-mail: jdyyywb2021@163.com

Key words: biomarkers, osteoarthritis, single-cell RNA, inflammation, machine learning, sequencing

Introduction

Osteoarthritis (OA) is a common chronic degenerative joint disease that primarily affects adults, with incidence increasing with age and affecting >500 million individuals worldwide (1,2). The pathogenesis of OA is yet to be elucidated and consequently, no curative therapy is available. However, although a number of treatment options are available, postoperative rehabilitation following total joint arthroplasty poses persistent challenges, including muscle weakness and atrophy caused by prolonged immobility, which impair joint stability and function, as well as joint stiffness and reduced range of motion resulting from scar tissue formation or disuse (3). Among pharmacological options, oral and topical non-steroidal anti-inflammatory drugs (NSAID), including cyclooxygenase-2 inhibitors, are effective for alleviating pain but carry increased risks. NSAIDs may cause gastrointestinal ulcers, bleeding or perforation, particularly with long-term use or in patients with a history of peptic ulcer disease. NSAIDs may increase the risk of cardiovascular complications, including hypertension, heart failure exacerbation, myocardial infarction and stroke (4). Biomaterial-based targeted delivery systems release drugs in response to specific internal or external stimuli associated with the disease microenvironment, but an absence of standardized safety and biocompatibility guidelines limits

their use in the general population. For example, spherical lipid bilayer vesicles are capable of encapsulating both hydrophilic and hydrophobic drugs, thereby improving their stability and facilitating targeted delivery (5). Targeted strategies have focused on regulating pathways such as the PI3K/AKT/mTOR, EGFR, toll-like receptors (TLRs) and proinflammatory IL pathways in cartilage (6-8). Notably, intra-articular therapy holds potential for OA treatment as numerous cDNA-based interventions exhibit therapeutic effects in mouse models (9). Therefore, screening for effective biomarkers is important for guiding therapeutic development. Increasing evidence indicates the inflammatory nature of OA (10,11). Inflammation serves a key role in OA pathogenesis. Inflammatory mediators, such as IL-1b, IL-6 and TNF-a, derived from synovial tissues, cartilage and subchondral bone contribute to OA development and increase with age (12), which is consistent with the increased prevalence of OA in adults. Synovitis is commonly observed in OA-affected joints and is associated with radiographic progression and pain severity (13,14). In zymosan-induced arthritis, low-density lipoprotein-1 (LOX-1)-knockout mice exhibit decreases in inflammation, synovial hyperplasia, cartilage degeneration and MMP-3 expression levels (15). In a joint destabilization model of OA, LOX-1-knockout mice have attenuated disease with decreased levels of oxidized LDL (ox-LDL), runt-related transcription factor 2 (Runx2) and type X collagen, implicating LOX-1/ox-LDL-driven endochondral ossification (16). Cartilage inflammation is also associated with chloride-channel dysfunction (17). One study has revealed the involvement of pathophysiological inflammatory biomarkers in OA disease, including MMPs and TIMPs (18). Therefore, the present study suggested that inflammation-related biomarkers may represent possible avenues for OA diagnosis and treatment.

Bioinformatics is an important tool for integrating and interrogating biological data, which enables researchers to investigate novel research directions (19). Bulk RNA sequencing (RNA-seq) profiles average gene expression levels across mixed cell populations, enabling differential expression analysis but masking cell-type-specific signals (20). However, single-cell (sc)RNA-seq enables the characterization of functional cellular states of different cell types, cellular plasticity upon stimulation and cell differentiation or reprogramming trajectories (21). A number of previous studies use bulk RNA-seq (22-24) or scRNA-seq alone (25-27) to investigate the mechanisms underlying OA development and progression. However, integrating bulk and scRNA-seq provides a more comprehensive overview. Therefore, the present study combined these approaches in order to elucidate the molecular and cellular mechanisms driving OA progression. Furthermore, four machine learning (ML) models, namely, generalized linear models (GLM), random forests (RF), support vector machines (SVM) and extreme gradient boosting (XGB), were used to construct predictive models and possibly elucidate complex biological mechanisms. Therefore, the present study aimed to integrate bioinformatics and ML algorithms in order to highlight possible OA biomarkers and inflammation-related genes, followed by a scRNA-seq-based annotation of these biomarkers.

Materials and methods

Data collection. All human OA transcriptomic datasets, including bulk and scRNA-seq data, were obtained from the Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/). Transcriptome samples from the OA and control groups were screened using the search terms 'osteoarthritis' and 'human', respectively. The primary analysis incorporated six datasets, GSE12021 (28), GSE1919 (29), GSE36700 (30), GSE51588 (31), GSE55235 (32) and GSE55457 (32), comprising 99 samples. Although GSE12021 and GSE1919 were generated primarily for rheumatoid arthritis (RA) synovial studies, they also contain data on non-RA control samples, including OA and normal tissues. In the present study, OA and normal samples were selected and no RA samples were included in the present analyses. The validation group consisted of three datasets, GSE82107 (33), GSE117999 (34) and GSE169077 (35), containing a total of 52 samples. The scRNA-seq dataset used in the present study was GSE104782 (36), which was sourced from the GPL20301 platform (Illumina, Inc.). The patient clinical information is summarized in Table SI.

Identification of inflammation-related differentially expressed genes (DEGs). To integrate OA and control samples from the six datasets, the 'ComBat' function in the 'sva' R package (37) (version 3.52.0; Posit Software, PBC) was used to eliminate batch differences in gene expression across datasets. Briefly, 'ComBat' was applied with a batch set to the dataset source, included no additional covariates in the model in order to preserve biological variance and used the parametric empirical Bayes framework (38). The processed dataset contained 60 OA samples and 39 control samples. DEGs were identified using an absolute $\log_2$-fold change ≥1 and a false discovery rate <0.05. To characterize inflammation-related biology among these DEGs, the curated 'inflammatory response' gene set was retrieved from the 'c2.cp.kegg.symbols.gmt' collection in the Molecular Signatures Database (version 2023.2; Broad Institute, Inc.) and these genes were intersected with DEGs to obtain inflammation-related DEGs (IRDGs). The chromosomal locations of the intersecting genes and their protein-protein interaction (PPI) networks were mapped and analyzed.

Cluster analysis and immune characterization. On the basis of the expression profiles of the IRDGs, a consistency clustering analysis (39) was carried out using the 'ConsensusClusterPlus' R/Bioconductor package (version 1.68.0; Matt Wilkerson and Peter Waltman) (40,41). Cluster separation was visualized using heatmaps and assessed through principal component analysis (PCA) (42). Gene set variation analysis (GSVA) (43) was carried out to assess the pathway variability among different clusters. The cell-type identification by estimating relative subsets of RNA transcripts (CIBERSORT) R script (44) was used to deconvolute the gene expression matrix into the relative proportions of immune cell types. The estimation of stromal and immune cells in malignant tumor tissues using expression data R package (version 1.0.13; R Core Team) (45) was used to calculate immune infiltration scores (immune score), overall stromal content score (stromal scores) and composite scores (estimate scores) on the basis of gene expression data in OA

samples (46). Pearson's correlation coefficients were used to assess the associations between these scores and different types of immune cell populations (47). Differences in immune cell distribution, immune function and immune checkpoint expression among the different clusters were characterized based on immune infiltration results.

ML model construction and validation. Initial screening of IRDGs was carried out using least absolute shrinkage and selection operator (LASSO) regression with the 'glmnet()' function in the 'glmnet' R package (48) (version 4.1.8; CRAN; https://cran.r-project.org/package=glmnet). The penalty parameter was set to α=1 (L1 regularization). The optimal λ was determined by tenfold cross-validation and the final model used 'λ_min', which was the value yielding the lowest cross-validation error. The selected genes were used as inputs for four machine learning models, namely GLM, RF, SVM and XGB. After data partitioning, a prediction function was constructed to analyze disease-specific signature genes. Predictors were standardized using the 'glmnet' default settings (standardize=TRUE). Subsequently, the following models and parameters were used: SVM ['e1071' version 1.7-16, C=1.0, γ='scale' and 5-fold cross validation (CV)]; GLM/LASSO ('glmnet' version 4.1-8, family='binomial' and α=1); RF (version 4.7-1.1, repeated CV with 5 folds and savePredictions=TRUE); and XGB ('xgboost' version 1.7.8.1, booster='gbtree', max_depth=6 and learning_rate=0.1). The predictive accuracies of the four algorithms were compared by evaluating the residual box plots, reverse cumulative distribution of residuals and receiver operating characteristic (ROC) curves (49). The four most important genes were identified on the basis of their gene importance scores. Calibration curves and nomograms (50) were constructed for the best-performing models. To validate the prediction accuracy, the final model was validated using three independent datasets, GSE82107, GSE117999 and GSE169077, in which each inflammation-related gene was used as an individual predictor within the finalized model.

scRNA-seq data analysis. Single-cell gene attainment matrices were converted into Seurat object data structures using the 'Seurat' R package (version 4.4.0; CRAN). After integration and batch correction, gene expression was normalized for each cell using the Seurat 'LogNormalize' method (scaling factor=10,000). The Seurat 'ScaleData' was then applied to center and scale gene expression across all genes to support downstream analyses. Samples containing <3 cells and <50 genes were removed. Samples that met these thresholds were retained unless they failed additional quality control (QC) criteria. Strict quality control was applied to exclude low-quality cells, which were defined as cells with <200 detected genes or mitochondrial gene expression accounting for 10% or more of total transcripts (51). The data were batch corrected and normalized, and the top 3,000 genes with the highest intercellular coefficients of variation were extracted for further analysis. Following PCA and dimensionality reduction, t-distributed stochastic neighbor embedding (t-SNE) (52) was used to calculate the difference in neighbor distance for cell clustering. Marker genes for each subpopulation were identified and disease-specific signature gene expression was visualized. The 'SingleR' R package (version 2.6.0) (53) was used to annotate cell types. Single-cell differentiation trajectory analysis was carried out using the 'monocle' R package (54) to generate trajectory maps of dendritic cells, temporal changes, cell-type annotation and clustering (55).

Cell culture and transfection. ATDC5 cells (The Cell Bank of Type Culture Collection of The Chinese Academy of Sciences) were cultured in a 1:1 mixture of DMEM and Ham's F-12 (Gibco; Thermo Fisher Scientific, Inc.) supplemented with 10% FBS (FuHeng Biology Co., Ltd.), 100 U/ml penicillin and 100 μg/ml streptomycin, and incubated at 37˚C in a humidified 5% CO_2 environment. Lentiviral short hairpin RNAs (shRNAs) targeting low-density lipoprotein receptor (LDLR), adrenomedullin (ADM), MYC and NF-κB inhibitor-α (NFKBIA) as well as a non-targeting shRNA control were purchased from Shanghai GenePharma Co., Ltd. In addition, shRNA oligonucleotides were cloned into the pLKO.1-puro lentiviral backbone (Sigma-Aldrich; Merck KGaA). The following shRNA sequences (in the 5'-3' direction) were used: shRNA control, TACTGTTAGGATCAGGAGG; LDLR shRNA#1, CGGAGGTGACCAACAATAGAA; LDLR shRNA#2, GATGTCATAAACGAAGCCATT; ADM shRNA#1, CCTCATTACTACTTGAACTTT; ADM shRNA#2, TCCAGACTCTTTAGGATATAG; MYC shRNA#1, CGACGAGGAAGAGAATTTCTA; MYC shRNA#2, GCTTCGAAACTCTGGTGCATA; NFKBIA shRNA#1, GAGTCAGAATTCACAGAGGAT; and NFKBIA shRNA#2, GGACGAGGAGTACGAGCAAAT. Lentiviral particles were produced in 293T cells using a 2nd-generation packaging system. Briefly, cells were co-transfected with the shRNA lentiviral vector together with psPAX2 and pMD2.G at a plasmid ratio of 4:3:1, with a total of 8 μg DNA per 10 cm dish, using Lipofectamine 2000. After incubation at 37˚C for 6-8 h, the transfection medium was replaced with fresh complete medium. Lentiviral supernatants were collected 72 h post-transfection and filtered through a 0.45 μm membrane. ATDC5 cells were infected with the lentiviral supernatants at a multiplicity of infection of 10 in the presence of 10 μg/ml polybrene (Shanghai GenePharma Co., Ltd.). At 48 h, the medium was replaced, and downstream experiments were conducted 72 h after infection. Stably transduced cells were selected and maintained in the presence of 2 μg/ml puromycin.

Reverse transcription-quantitative PCR (RT-qPCR) analysis. Total RNA from transfected ATDC5 cells was isolated using TRIzol reagent (Invitrogen; Thermo Fisher Scientific, Inc.) following the manufacturer's protocol. The mRNA was reverse transcribed into cDNA using the RevertAid RT Kit (Thermo Fisher Scientific, Inc.) according to the manufacturer's protocol. The target genes were amplified with an initial denaturation step at 95˚C for 10 min, followed by 40 amplification cycles consisting of 95˚C for 15 sec and 60˚C for 1 min. β-actin served as the internal reference for normalization. Relative mRNA expression levels were determined using the $2^{-\Delta\Delta Cq}$ method as described previously (56,57). The primer sequences used were: LDLR forward, 5'-TCCAATCAATTCAGCTGTGG-3'; LDLR reverse, 5'-GAGCCATCTAGGCAATCTCG-3'; GAPDH forward, 5'-AACGACCCCTTCATTGAC-3'; and GAPDH reverse 5'-TCCACGACATACTCAGCAC-3'.

Western blotting analysis. ATDC5 chondrocytes were lysed in RIPA buffer (Beyotime Biotechnology) and the lysates were quantified using the BCA method. Equal amounts of proteins (30 μg) were subjected to 10% of sodium dodecyl sulfate-polyacrylamide gel electrophoresis and PVDF transfer. The membranes were blocked at room temperature in 5% milk for 1 h and incubated with specific primary antibodies at 4˚C overnight, including anti-LDLR (1:1,000; cat. no. ab30532; Abcam), anti-MYC (1:1,500; cat. no. ab32072; Abcam), anti-ADM (1:1,500; cat. no. ab190819; Abcam), anti-NFKBIA (1:2,000; cat. no. ab7217; Abcam), anti-MMP13 (1:1,000; cat. no. ab39012; Abcam) and anti-GAPDH (1:5,000; cat. no. ab8245; Abcam) antibodies. The membranes were subsequently washed with TBST (TBS containing 0.1% Tween-20) and incubated with HRP-conjugated anti-mouse (1:5,000; cat. no. #7076; CST Biological Reagents Co., Ltd.) or anti-rabbit (1;5,000; cat. no #7074; CST Biological Reagents Co., Ltd.) secondary antibodies at room temperature for 1 h. The protein bands were detected using enhanced chemiluminescence substrate (Shanghai Yeasen Biotechnology Co., Ltd.) (58).

Cell Counting Kit-8 (CCK-8) assays. Proliferation of transfected ATDC5 chondrocytes was assessed using a CCK-8 assay. Briefly, 100 μl cell suspension ($4x10^3$ cells/well) was seeded into 96-well plates and incubated at 37˚C for 24, 48 and 72 h. Subsequently, 10 μl CCK-8 (Shanghai Yeasen Biotechnology Co., Ltd.) reagent was added to each well and incubated for 2 h at 37˚C, after which the absorbance at 450 nm was measured (59).

Apoptosis analysis. After centrifugation at 150 x g for 5 min at 4˚C, ATDC5 cells were resuspended in 500 μl of binding buffer [Multi Sciences (Lianke) Biotech, Co., Ltd.] at $1x10^6$ cells/ml. The cells were then stained with FITC-annexin V and propidium iodide (PI) by apoptosis detection kit [Multi Sciences (Lianke) Biotech, Co., Ltd.] according to the manufacturer's protocol and analyzed using FACSCalibur flow cytometry (BD Biosciences). Data were analyzed using FlowJo software (version 10.8.1; FlowJo LLC; BD Biosciences). Annexin V/PI quadrants were used to classify and quantify the viable, early apoptotic and late apoptotic/necrotic populations.

Statistical analysis. A total of six OA datasets were integrated and batch effects were removed to generate the final dataset that was used for analysis. IRDGs were extracted by comparing OA and control samples. Consensus clustering analysis was carried out on the basis of IRDG, followed by comparisons of gene expression, metabolic pathways, immune characteristics and drug sensitivity across different clusters. Combined LASSO regression analysis (60) and four ML algorithms were used to screen disease-associated biomarker genes, which were validated in three transcriptomic datasets. Target gene annotation was carried out using sc-transcriptomic data. All statistical analyses were carried out using the R programming language (version 4.1.3; Posit Software, PBC) (61) and Perl programming language (Strawberry Perl; version 5.30.0.1; Microsoft Corporation) (62). Perl scripts were primarily used for data integration and extraction, whereas R scripts were used for systematic data analysis and visualization. Experiments were repeated three times. Our data were presented as the mean ± SD. Differences between two groups were assessed using paired Student's t-tests, while comparisons among multiple groups were evaluated using one-way ANOVA with Tukey post-test. $P<0.05$ was considered to indicate a statistically significant difference.

Results

DEGs are identified across OA dataset from gene expression omnibus. Integration of the six datasets resulted in the identification of 8,238 co-expressed genes. Among these genes, 537 DEGs were identified between the OA and control samples, consisting of 204 genes with upregulated expression levels and 333 genes with downregulated expression levels. The six most significantly expressed DEGs were KLF transcription factor 9, MAF BZIP transcription factor F, JUND, GABA type A receptor associated protein like 1, heterogeneous nuclear ribonucleoprotein A1 and proline rich nuclear receptor coactivator 1 (Fig. 1A). By intersecting the 537 DEGs with 200 inflammation-associated genes, 11 IRDGs were identified (Fig. 1B), namely ADM, BTG anti-proliferation factor 2 (BTG2), cyclin dependent kinase inhibitor 1A (CDKN1A), IL-1 receptor type 1 (IL1R1), KLF transcription factor 6 (KLF6), LDLR, MYC, nicotinamide phosphoribosyltransferase (NAMPT), NFKBIA, presenilin 1 (PSEN1) and selectin E (SELE) (Fig. 1C). From the perspective of human chromosomes, BTG2 and SELE are located on chromosome 1. IL1R1, CDKN1A, NAMPT, MYC, KLF6, ADM and LDLR correspond to chromosomes 2, 6, 7, 8, 10, 11 and 19, respectively. PSEN1 and NFKBIA are located on chromosome 14 (Fig. 1D). Using the 11 IRDGs to construct a PPI network, five hub genes, NAMPT, NFKBIA, MYC, KLF6 and CDKN1A, were identified as central nodes in the network (Fig. 1E).

OA subtypes are identified based on differentially expressed inflammatory genes. Consistent clustering consensus matrix based on the expression of 11 IRDGs classified the OA samples into two distinct clusters (clusters C1 and C2; Fig. 2A). The clustering result using k=2 was determined to be optimal on the basis of the cumulative distribution function curve (Fig. 2B) and the δ area curve (Fig. 2C). Cluster C1 contained 25 samples and the remaining 35 were assigned to cluster C2. The expression levels of IRDGs in different clusters are shown in Fig. 2D. PSEN1 was highly expressed in cluster C2 but at low levels in cluster C1, whereas the other IRDGs showed upregulated expression in cluster C1 compared with cluster C2. PCA demonstrated the distinct classification of clusters C1 and C2 (Fig. 2E). GSVA revealed the top 10 upregulated ('prostate cancer', 'NOD like receptor signaling pathway', 'small cell lung cancer', 'pathways in cancer', 'adipocytokine signaling pathway', 'cytosolic DNA sensing pathway', 'nitrogen metabolism', 'apoptosis', 'JAK STAT signaling pathway', 'insulin signaling pathway') and downregulated pathways ('vasopressin regulated water reabsorption', 'long term potentiation', 'asthma', 'leishmania infection', 'biosynthesis of unsaturated fatty acids', 'gap junction', 'alpha linolenic acid metabolism', 'intestinal immune network for IgA production', 'ubiquitin mediated proteolysis', 'type 1 diabetes mellitus') in cluster C2 compared with those in cluster C1 (Fig. 2F).

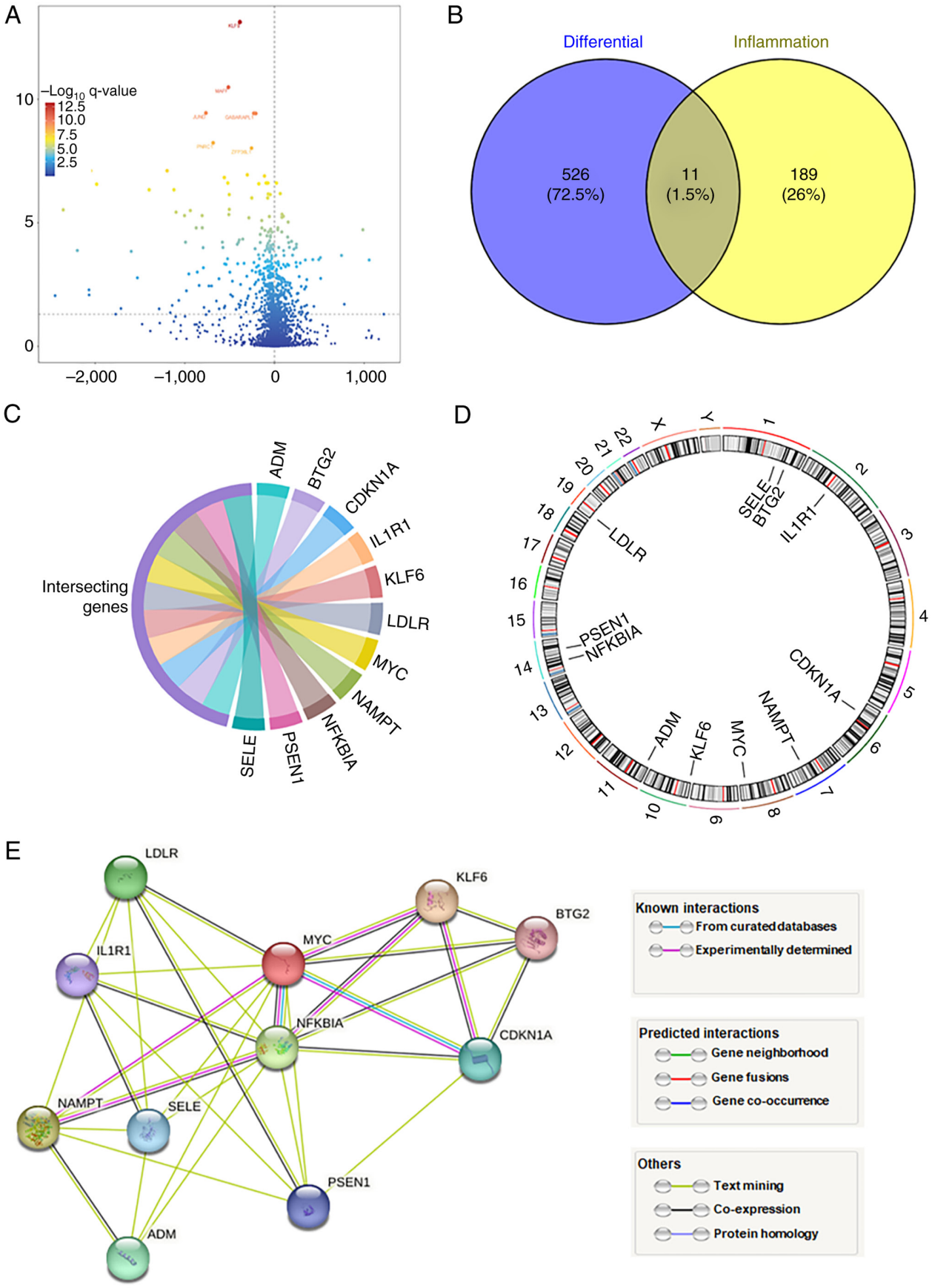

Figure 1. Differential gene expression analysis across Gene Expression Omnibus OA datasets. (A) Volcano plot of DEGs between patients with OA and controls. (B) Intersection of DEGs and inflammation-related genes. (C) IRDGs in patients with OA. (D) Chromosomal localization of IRDGs on 23 human chromosomes. (E) Protein-protein interaction networks constructed for the 11 IRDGs. DEGs, differentially expressed genes; IRDG, inflammation-related differentially expressed genes; OA, osteoarthritis. ADM, adrenomedullin; BTG2, B-cell translocation gene 2; CDKN1A, cyclin dependent kinase inhibitor 1A; IL1R1, interleukin 1 receptor type 1; KLF6, kruppel-like factor 6; LDLR, low density lipoprotein receptor; MYC, MYC proto-oncogene, NAMPT, nicotinamide phosphoribosyltransferase; NFKBIA, NFKB inhibitor alpha; SELE, selectin E.

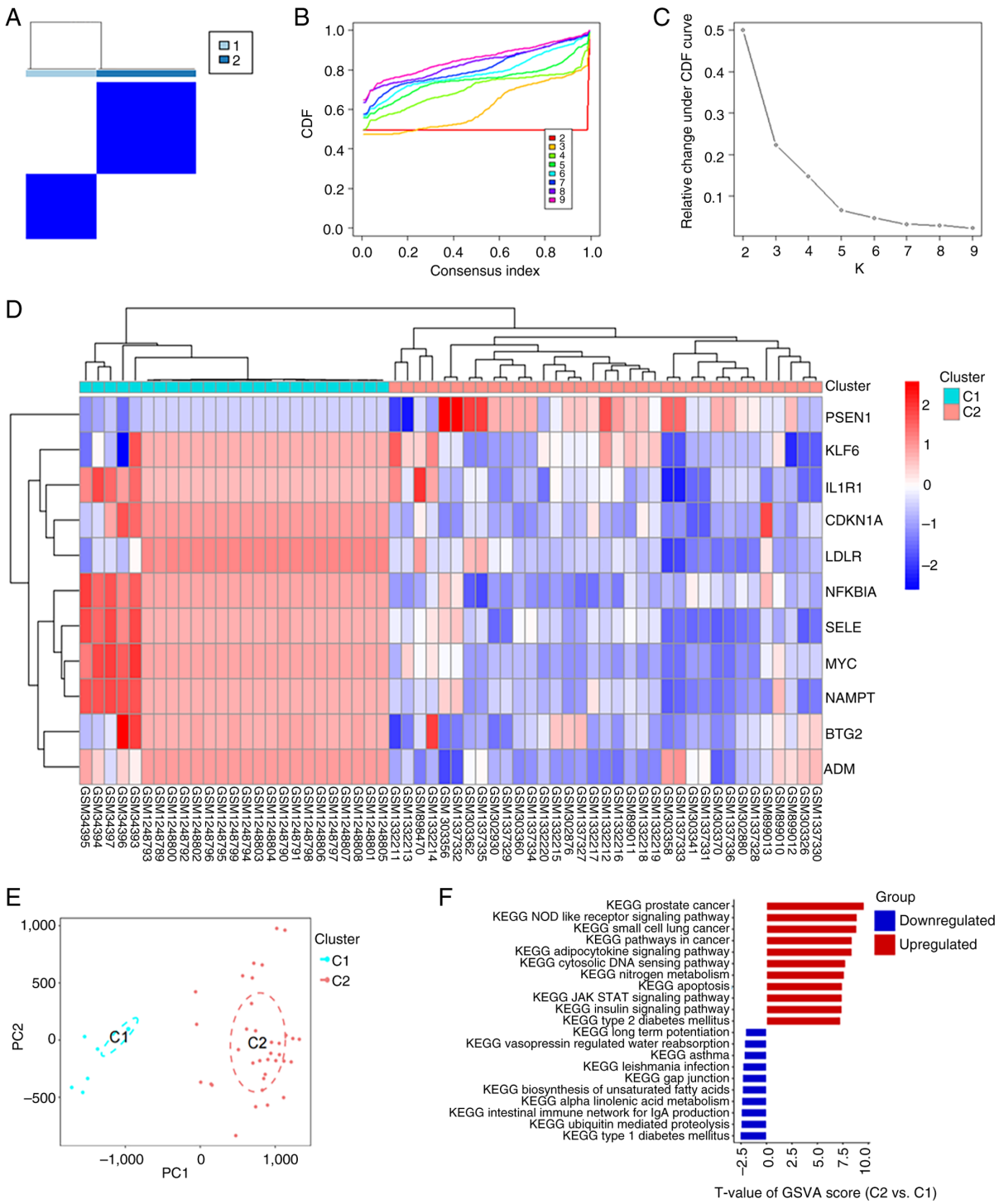

Figure 2. OA subtype classification based on differentially expressed inflammatory genes. (A) Consensus clustering matrix heatmap defining two clusters (k=2) and their correlation areas. (B) Consensus CDF curve demonstrating optimal clustering. (C) δ-area curve of consensus clustering in OA, indicating the relative change in the area under the CDF curve for each category number k compared with k-1. The horizontal axis represents the category number k, whereas the vertical axis represents the relative change in the area under the CDF curve. (D) Heatmap of the expression profiles of 11 inflammation-related differentially expressed genes in the two clusters. (E) PC analysis illustrating distinct immune subtypes in clusters C1 and C2. (F) GSVA pathway analysis comparing the top upregulated (red) and downregulated (blue) pathways between two distinct clusters. OA, osteoarthritis; CDF, cumulative distribution function; GSVA, gene set variation analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes; PC, principal component. ADM, adrenomedullin; BTG2, B-cell translocation gene 2; CDKN1A, cyclin dependent kinase inhibitor 1A; IL1R1, interleukin 1 receptor type 1; JAK, Janus kinase; KLF6, kruppel-like factor 6; LDLR, low density lipoprotein receptor; MYC, MYC proto-oncogene, NAMPT, nicotinamide phosphoribosyltransferase; NFKBIA, NFKB inhibitor alpha; STAT, signal transducer and activator of transcription; SELE, selectin E.

Different OA subtypes exhibit distinct immune characteristics. Only the stromal score differed among the immune, stromal and estimated scores between the two clusters (P=0.0011; Fig. 3A). Correlation analysis of the 22 immune cell types revealed primarily negative correlations, with neutrophils, activated mast cells and M2 macrophages exhibiting the

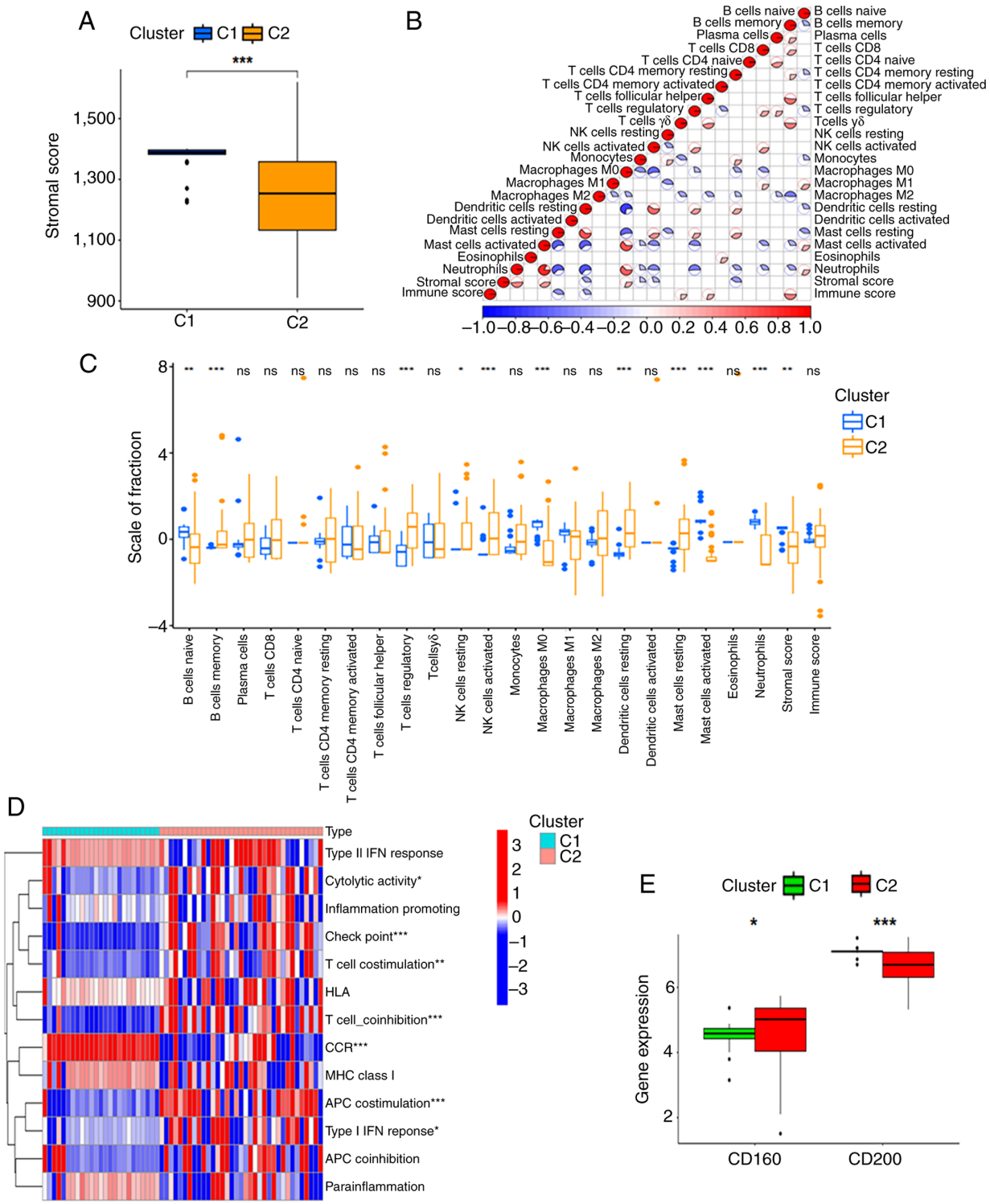

Figure 3. Immune characteristics of different OA subtypes. (A) Defining two clusters (k=2) matrix score boxplots. (B) Correlation matrix of 22 immune cell types, stromal scores and immune scores in OA. Red indicates a positive correlation and blue indicates a negative correlation. (C) Differentially expressed immune cells between clusters C1 and C2 in OA. (D) Differentially expressed immune functions in clusters C1 and C2. (E) Expression levels of immune checkpoints in clusters C1 and C2. *P<0.05, **P<0.01 and ***P<0.001. OA, osteoarthritis; ns, not significant; NK, natural killer; HLA, Human leukocyte antigen; CCR, CC chemokine receptor; MHC, major histocompatibility complex; APC, antigen-presenting cell.

strongest negative correlations with the other immune cell types (Fig. 3B). By contrast, memory B cells displayed the strongest positive correlation with the other immune cell types. Differential analysis of immune cell composition revealed that 10 immune cell types were differentially expressed (P<0.05) between clusters C1 and C2, with 8 exhibiting highly significant differences (P<0.001; Fig. 3C). In terms of immune functions, 7 immune functions were differentially enriched between the two clusters (P<0.05), with the most significantly altered immune functions being 'Check-point', 'T cell co-inhibition', chemokine receptor ('CCR') and antigen presenting cell '(APC) co-stimulation' (Fig. 3D). The expression levels of

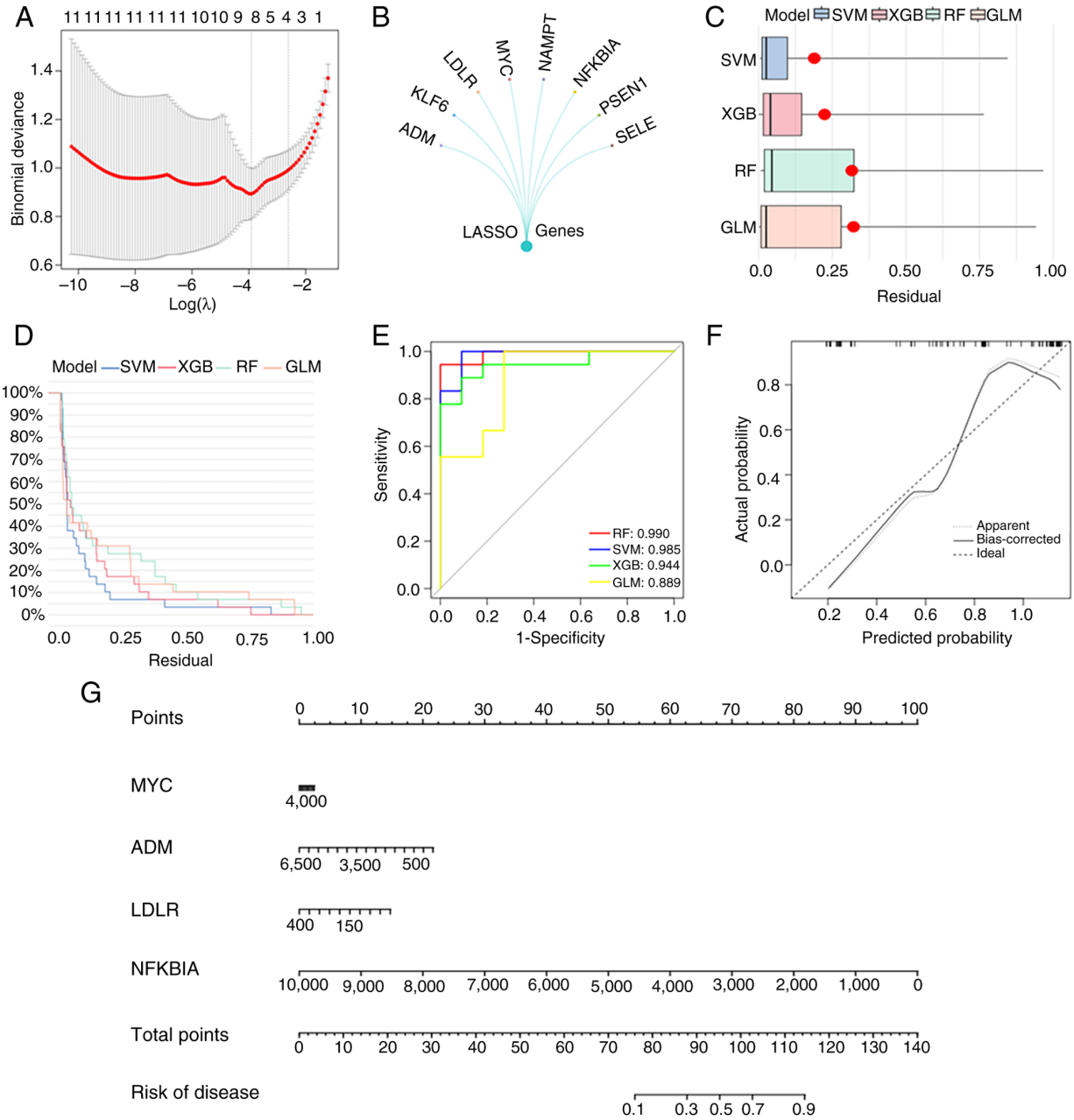

Figure 4. Machine-learning model construction. (A) LASSO regression analysis for preliminary screening of OA signature genes. (B) Selection of disease-related signature genes through LASSO regression. (C) Residual box line diagram for the GLM, RF, SVM and XGB models. The red dot indicates the root mean square of residuals. (D) Reverse cumulative distribution of residuals for different machine learning models. (E) Receiver operating curve analysis verifying model accuracy. (F) Calibration curve for assessing the accuracy of the nomogram. (G) Nomogram predicting OA incidence based on four signature genes. LASSO, least absolute shrinkage and selection operator; OA, osteoarthritis; GLM, generalized linear models; RF, random forests; SVM, support vector machines; XGB, extreme gradient boosting; ADM, adrenomedullin; LDLR, low-density lipoprotein receptor; NFKBIA, NF-κB inhibitor-α; KLF6, Krüppel-like factor 6; NAMPT, nicotinamide phosphoribosyltransferase; PSEN1, presenilin-1; SELE, selectin E.

immune checkpoints CD200 and CD160 significantly differed between the two clusters. Compared with cluster C1, cluster C2 exhibited increased CD160 expression but decreased CD200 expression (Fig. 3E).

A ML model is constructed to identify disease signature genes. Before the ML model was constructed, the 11 IRDGs were screened using LASSO regression analysis (Fig. 4A), which narrowed the selection to eight disease signature genes, namely ADM, KLF6, LDLR, MYC, NAMPT, NFKBIA, PSEN1 and SELE (Fig. 4B). The residual box-line plots indicated that the SVM model had the smallest root-mean-square residuals (Fig. 4C). The reverse cumulative residual distribution likewise favored the SVM model (Fig. 4D). ROC curve analysis revealed that the SVM model achieved 98.5% accuracy, which demonstrated that it was the best-performing model for disease signature gene prediction (Fig. 4E). Calibration curves revealed that the predicted values were highly consistent with the actual results, although a high fluctuation at a 90% prediction probability was noted, requiring further investigation (Fig. 4F). A nomogram was constructed based on four key disease signature genes (MYC, ADM, LDLR and

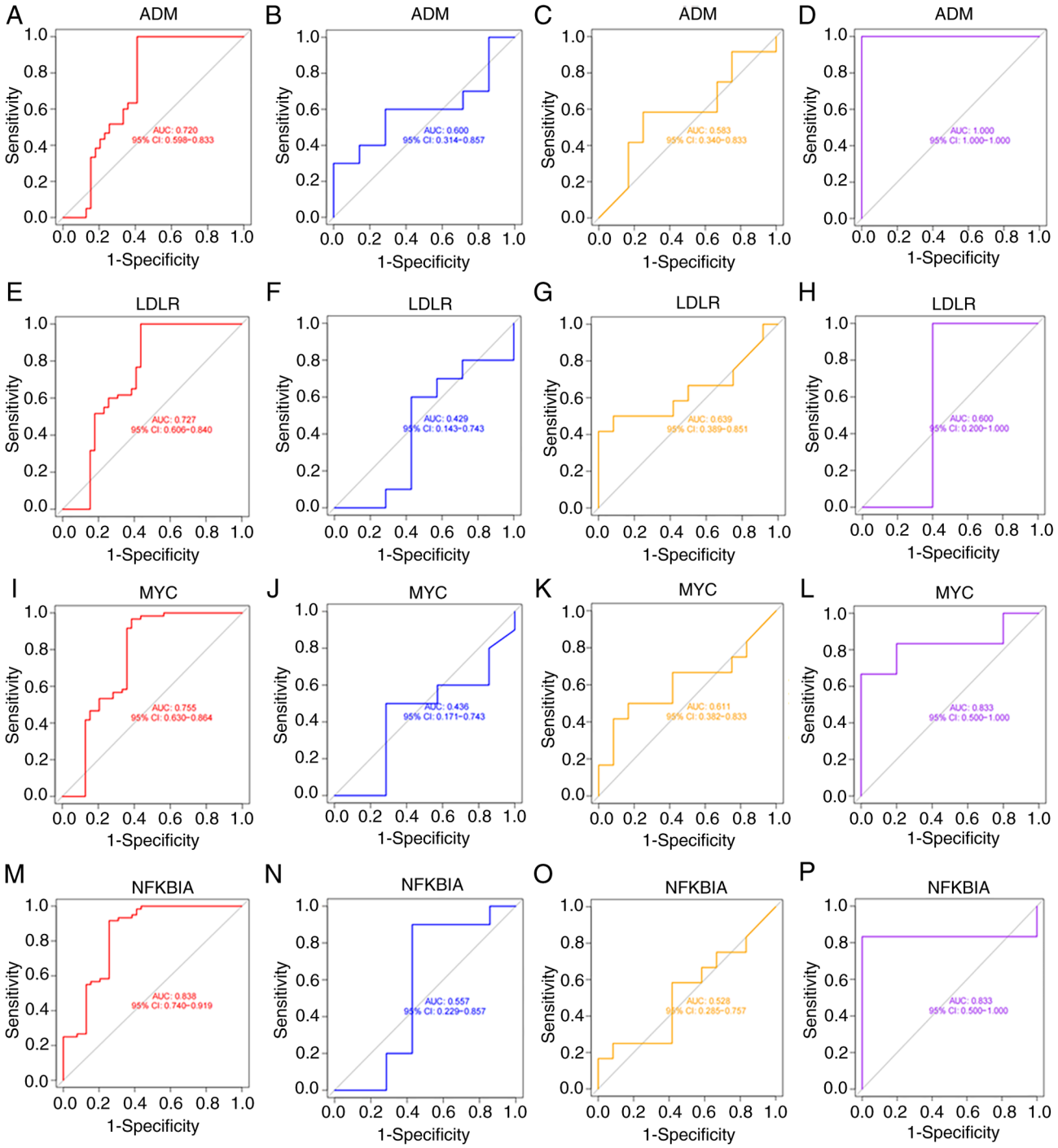

Figure 5. ROC curves of the four osteoarthritis disease signature genes in the training and test sets. ROC curves of ADM in the (A) training, (B) GSE82107, (C) GSE117999 and (D) GSE169077 datasets. ROC curves of LDLR in the (E) training, (F) GSE82107, (G) GSE117999 and (H) GSE169077 datasets. ROC curves of MYC in the (I) training, (J) GSE82107, (K) GSE117999 and (L) GSE169077 datasets. ROC curves of NFKBIA in the (M) training, (N) GSE82107, (O) GSE117999 and (P) GSE169077 datasets. ROC, receiver operating curve; ADM, adrenomedullin; LDLR, low-density lipoprotein receptor; NFKBIA, NF-κB inhibitor-α; AUC, area under the curve.

NFKBIA), which were selected by the SVM model, with each gene assigned a corresponding score to assess the risk of OA in individuals (Fig. 4G).

Model accuracy is validated. The classification performance of the ML model was evaluated in three validation datasets based on the four identified genes (MYC, ADM, LDLR, and NFKBIA). ADM had a 100% classification accuracy in the GSE169077 dataset (Fig. 5A-D). Compared with the validation datasets, LDLR had the highest classification accuracy (72.7%) in the training set (Fig. 5E-H). MYC had an 83.3% classification accuracy in the GSE169077 dataset (Fig. 5I-L). NFKBIA had an 83.8% accuracy in the training set and an 83.3% accuracy in the GSE169077 validation dataset (Fig. 5M-P). Collectively, the classification accuracy for all the genes in the training set exceeded 70%. Among the test datasets, GSE169077 demonstrated the best classification performance, followed by GSE117999, whereas GSE82107 yielded the

lowest accuracy. The differences in performance across datasets indicate that the classification ability of these genes is not completely uniform across cohorts. However, because the present study relied on retrospective public datasets with unavoidable differences in sample composition, sequencing platform, preprocessing workflow, and underlying biological heterogeneity, we could not determine the exact reason for the observed variation. Thus, our findings suggest potential predictive value of these genes, but further validation in larger, prospectively collected cohorts is required.

Four osteoarthritis signature genes are identified. OA sc sequencing datasets contained data on 1,464 chondrocytes derived from 10 patients with OA who underwent knee replacement surgery, with a total of 34 samples. Based on the intercellular coefficient of variation, the top 3,000 genes with the greatest variation among the 15,880 genes were extracted and included in the subsequent analysis (Fig. S1A-F). Sc data were analyzed using PCA for dimensionality reduction and t-SNE analysis for clustering (Fig. S2A-L). The expression patterns of four OA disease signature genes (ADM, LDLR, MYC and NFKBIA) were identified by the SVM model and analyzed using a GSE104782 sc dataset. The expression levels in the nine chondrocyte clusters revealed that NFKBIA had the highest expression levels, followed by MYC, ADM and LDLR (Fig. 6A). This finding was supported by t-SNE scatter plots (Fig. 6B-E). Cartilage synovial cell expression analysis revealed that when comparing the expression levels of ADM, LDLR, MYC and NFKBIA, NFKBIA was the most highly expressed and widely distributed, whereas LDLR was expressed at the lowest level (Fig. 6F-J).

LDLR expression suppression inhibits cell viability and induces apoptosis. To investigate the function of LDLR in ATDC5 chondrocytes, the expression of LDLR was silenced using shRNA and overexpressed using LDLR cDNA. Knockdown using shRNA efficiently reduced LDLR expression levels at both the mRNA and protein levels compared with the control. Additionally, compared with the control, cDNA lentiviral delivery increased the expression levels of LDLR (Fig. 7A and B). Functionally, CCK-8 assays revealed that LDLR depletion decreased cell viability compared with the control. However, LDLR overexpression increased cell viability compared with the control (Fig. 7C). Furthermore, annexin V/PI staining revealed increased apoptosis after LDLR knockdown and reduced apoptosis after LDLR overexpression compared with the relevant controls (Fig. 7D and E). To investigate whether ADM, MYC and NFKBIA expression levels affected cell viability and apoptosis, their expression in ATDC5 cells was knocked down (Fig. S3A). Silencing ADM, MYC or NFKBIA reduced cell viability compared with the controls (Fig. S3B) and induced apoptosis in ATDC5 cells (Fig. S4). Collectively, these findings indicated that LDLR, ADM, MYC and NFKBIA may contribute to the regulation of ATDC5 cell viability and survival.

LDLR suppression increases the expression of MMP13 in ATDC5 cells. Loss of LDLR function may cause a decreased LDL uptake by chondrocytes and an increased cholesterol deposition, which may activate inflammatory pathways (63). Excess cholesterol can activate the NF-κB pathway, induce the expression of inflammatory cytokines (such as IL-1β and TNF-α) and upregulate MMP13 expression levels, promoting the degradation of the cartilage matrix (64). To determine whether LDLR regulates the expression of MMP13 in ATDC5 cells, MMP13 expression levels were quantified following the manipulation of LDLR expression levels. Knockdown of LDLR expression levels through shLDLR transfection increased MMP13 expression levels compared with the control. However, overexpression of LDLR reduced MMP13 expression levels compared with the control (Fig. 7F). These data suggested that LDLR may limit cartilage matrix degradation by suppressing the expression of MMP13.

Discussion

In recent years, numerous studies have investigated OA pathogenesis using integrative bulk and scRNA-seq techniques. For example, a study by Peng *et al* (65) identifies 58 oxidative stress-related DEGs and seven diagnostic genes [stanniocalcin 2, lymphocyte specific protein 1, collagen type VI α-1 chain, Fos proto-oncogene, selenoprotein N, tumor protein p53 and heat shock protein family A (Hsp70) member 8] in OA through integration of GEO bulk and scRNA-seq. Furthermore, by integrating bulk and sc-transcriptomics with ML, a study by Sun *et al* (66) identifies the dysregulated expression of the hub genes S100 calcium binding protein (S100)-A9, phorbol-12-myristate-13-acetate-induced protein 1, ectodysplasin A2 receptor and fatty acid synthase, revealing that programmed cell death is associated with OA. Additionally, another study identifies four m7G regulators, eukaryotic translation initiation factor 1 (EIF1), JUND, nudix hydrolase 16-like 1 and nuclear cap binding protein subunit 1, as diagnostic biomarkers for OA (67).

EIF1/JUND suppression induces MMP13 upregulation, collagen II downregulation and OA progression (67). An additional study mapped endoplasmic-reticulum-stress heterogeneity in osteoarthritic chondrocytes and identified insulin like growth factor binding protein 3 and S100A4 as key diagnostic regulators (68). Tryptophan metabolism, through the activation of the indoleamine 2,3-dioxygenase 1/tryptophan 2,3-dioxygenase-induced kynurenine pathway and the accumulation of 3-hydroxykynurenine/quinolinic acid, promotes cartilage degradation in OA chondrocytes (69). In the present study, 11 IRDGs were identified in the OA and normal samples. The regulatory network, composed of five hub genes (NAMPT, NFKBIA, MYC, KLF6 and CDKN1A), was highly coincident with the NF-κB/STAT3 inflammatory signaling axis and the specific expression of PSEN1 in cluster C2 suggested that it may exacerbate cartilage degradation through mitochondrial dysfunction. Notably, PSEN1 has been studied in association with Alzheimer's disease (AD) (70-72). Mutations in PSEN1 reduce barrier function, impair glucose metabolism and decrease drug efflux pump activity (73). Given that both OA and AD are degenerative diseases with a high prevalence in elderly individuals, future studies may benefit from investigating whether PSEN1 dysregulation contributes to OA pathogenesis.

In the inflammatory milieu, immune cell types were predominantly negatively correlated with one another. Significant differences were observed in numerous immune

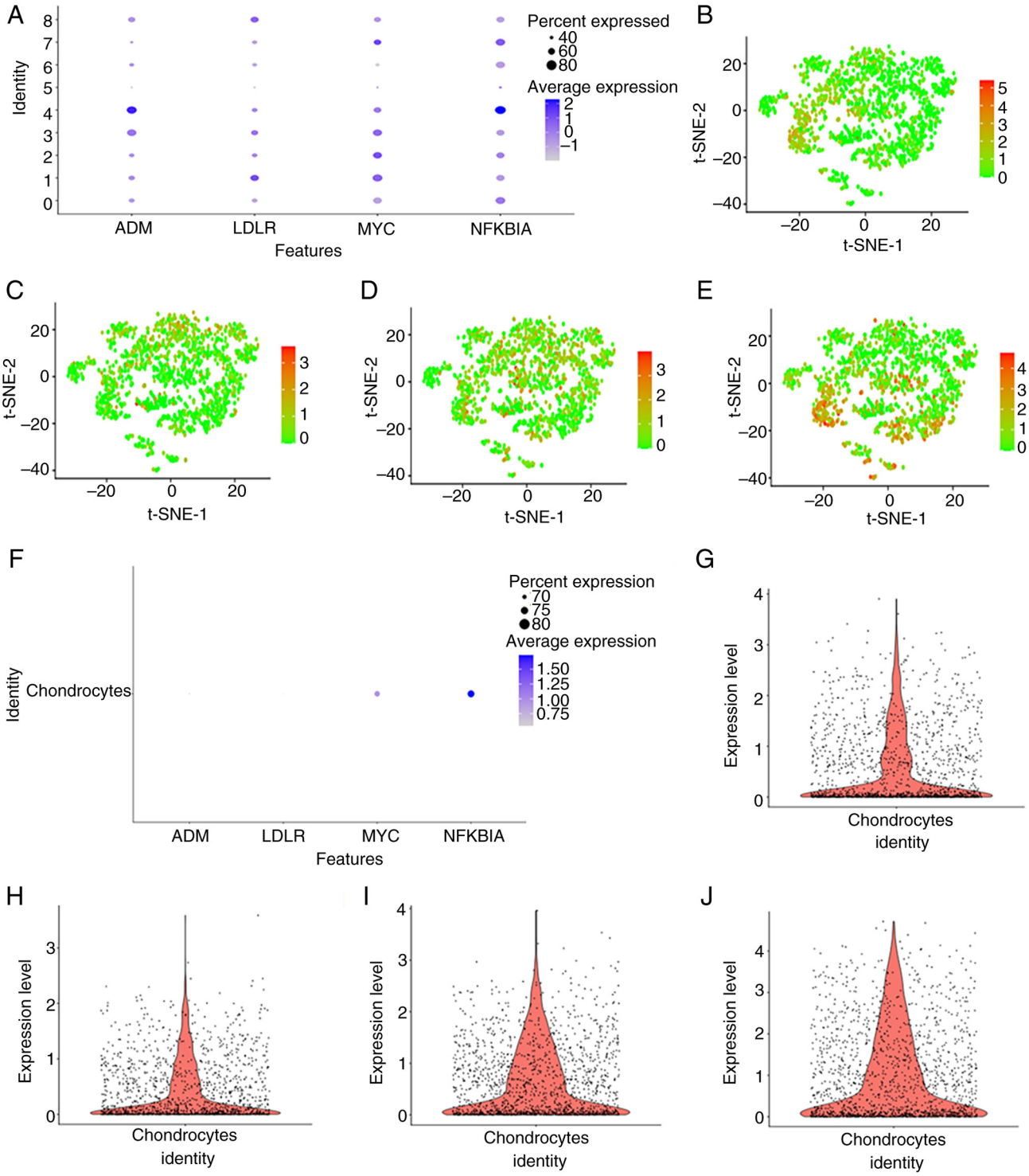

Figure 6. Expression profile analysis of four osteoarthritis signature genes in the single cell RNA-sequencing data. (A) Bubble map displaying the expression levels of ADM, LDLR, MYC and NFKBIA genes in the nine clusters. A darker blue color indicates a higher expression level. t-SNE scatter plots illustrating the expression of (B) ADM, (C) LDLR, (D) MYC and (E) NFKBIA in chondrocytes. Red represents upregulated genes and green indicates downregulated genes. (F) Bubble map visualizing the expression levels of four genes in chondrocytes. Violin diagram showing the differential expression of (G) ADM, (H) LDLR, (I) MYC and (J) NFKBIA in chondrocytes. ADM, adrenomedullin; LDLR, low-density lipoprotein receptor; NFKBIA, NF-κB inhibitor-α; t-SNE, t-distributed stochastic neighbor embedding.

cell populations (including memory B cells, T regulatory cells, activated natural killer cells, M0 macrophages, resting dendritic cells, resting mast cells, activated mast cells and neutrophils), four immune functions (checkpoint, T cell co-inhibition, CCR and APC co-stimulation) and one immune checkpoint (CD200). The present findings regarding immune cell infiltration and immune function align with those of a previous OA immune infiltration study (74). In patients with posttraumatic OA, expression of the cell surface glycoprotein CD200 receptor 1 is markedly increased and contributes to cartilage degeneration (75), suggesting that CD200 may be a key inflammation-related biomarker in patients with OA.

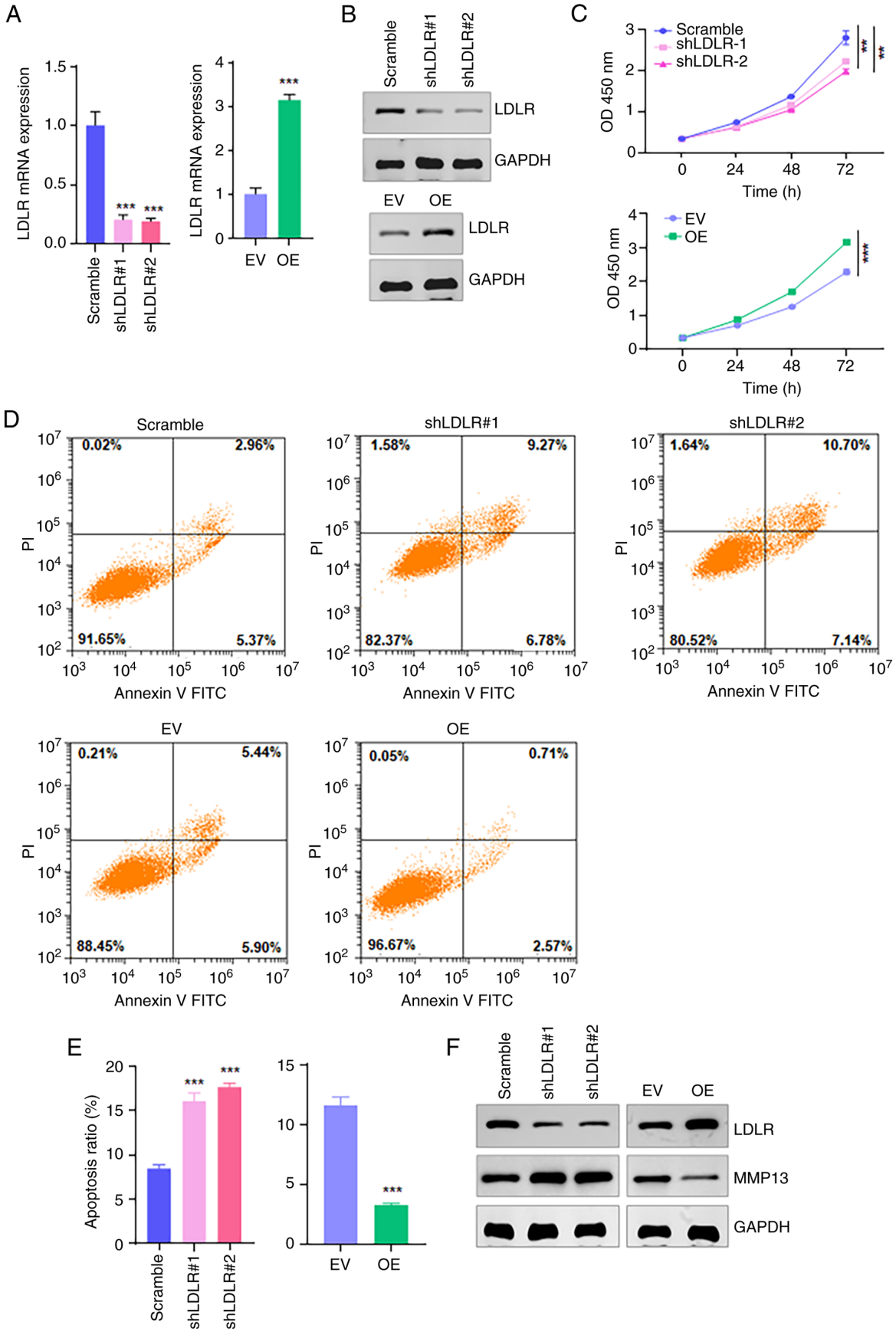

Figure 7. LDLR increases cell viability and suppresses apoptosis in ATDC5 chondrocytes. (A) Reverse transcription-quantitative PCR was carried out to measure the mRNA expression of LDLR in ATDC5 cells after LDLR modulation. (B) Western blotting was conducted to determine protein levels of LDLR in ATDC5 cells after LDLR modulation. (C) A Cell-Counting Kit 8 assay was carried out to evaluate ATDC5 cell viability after LDLR modulation. (D) Annexin V/PI flow cytometry was used to measure apoptosis in ATDC5 cells after LDLR modulation. (E) Quantitative results for apoptosis. (F) Western blotting was conducted to measure protein expression of MMP13 in ATDC5 cells after LDLR modulation. **P<0.01 and ***P<0.001. LDLR, low-density lipoprotein receptor; sh, short hairpin; EV, empty vector; OE, overexpression; OD, optical density.

Mechanistically, the Akt inhibitor MK-2206 has anti-autophagic, proapoptotic, prosenescent and procatabolic effects in human intervertebral discs. Conversely, Akt activation in intervertebral disc cells protects against inflammation through temsirolimus treatment (76). However, further research is needed to investigate the potential role of these drugs in OA treatment.

Using LASSO regression analysis in combination with four ML algorithms (GLM, RF, SVM and XGB), MYC, ADM, LDLR and NFKBIA were identified as inflammation-related biomarkers of OA. One study suggests that MYC mediates aberrant gene expression associated with OA and may serve as a potential therapeutic target (35). Bioinformatics of OA synovium shows elevated expression of MYC, JUN and VEGFA, which is associated with severity. Furthermore, MYC/VEGFA may induce p38-MAPK/JUN-mediated inflammation, proliferation and angiogenesis (77). Borojoa iridoid glycoside enhanced chondrocyte proliferation and reduced apoptosis through MYC upregulation (78). In ATDC5 cells, the miR-195-5p inhibitor restores proliferation and reduces apoptosis and inflammation through the activation of Wnt/β-catenin and NF-κB pathways, including MYC, cyclin D1 and p-p65 (79). ATDC5 cells lacking nuclear factor of activated T (NFATc)-1 and NFATc2 have an increased proliferation and an upregulated expression of MYC (80). In IL-1β-induced ATDC5 cells, overexpression of long non-coding RNA colorectal neoplasia differentially expressed enhances proliferation and reduces apoptosis, inflammation and extracellular matrix (ECM) degradation, while IL-1β activates Wnt/β-catenin signaling with increases in the expression of β-catenin, MYC and cyclin D1 (81). MYC overexpression enhances chondrocyte viability and suppresses ferroptosis. *In vivo*, MYC overexpression mitigates cartilage damage in OA mouse models (82). These findings suggested that MYC may promote chondrocyte viability and inhibit apoptosis. Consistently, the present data showed that MYC silencing reduced viability and increased apoptosis in ATDC5 cells.

A number of studies suggest that ADM may serve as a useful diagnostic biomarker for OA (83-86). For example, microarray profiling of OA bone identifies 150 DEGs, including ADM, implicating Wnt/TGF-β/bone morphogenetic protein and osteoclast pathways in altered bone remodeling and sex differences (87). Shear stress changed the transcriptional profile of osteoarthritic chondrocytes, with microarray analysis identifying 14 upregulated and 6 downregulated genes, including ADM, that may increase cell survival and matrix homeostasis (88). A transcriptome-wide association study, integrating genome-wide association studies and expression profiles, identifies shared OA genes and pathways, including ADM, FSTL1 genes and the PI3K/AKT pathway. ADM is a notable candidate, highlighting PI3K/AKT and immune-associated mechanisms, such as inflammatory signaling, antigen presentation, chemokine activity, and innate immune responses (89). Through integrating DEGs and weighted gene co-expression network analysis, a previous study identifies nuclear factor IL-3 regulated, ADM and osteoglycin as early OA biomarkers, associating MYC targets and NF-κB signaling with inflammatory pathways, including ROS and TNF-a (90). Using GSVA/single sample GSEA, hypoxia is implicated in OA and immune infiltration. A seven-gene index including ADM achieved high diagnostic accuracy (91). Consistently, the present study identified that ADM may be an inflammation-related OA biomarker. ADM prevents fibroblast-like synovial apoptosis in patients with RA by upregulating calcitonin receptor like receptor/receptor activity modifying protein 2 expression (92). In line with this finding, the present results showed that silencing ADM reduced cell viability and induced apoptosis in ATDC5 cells.

NFKBIA, an immune-related gene, has high predictive performance in OA diagnostic models (93-95). In the Han Chinese, SNPs in TLR10 and NFKBIA are associated with hip OA risk (96). Microarray analysis of joint fibroblast-like synoviocytes identifies 276 OA DEGs. The OA-specific hubs genes SELE, SERPINE1 and NFKBIA implicated that TNF signaling may contribute to OA inflammation by mediating the pro-inflammatory activity of fibroblast-like synoviocytes in the synovium (97).

Safflower yellow, a natural pigment extracted from safflower, protects chondrocytes and suppresses OA inflammation by inhibiting TNF-α-induced NF-κB activation, reducing MMP-13 expression levels and improving cartilage degeneration (98). Integrating OA synovium DEGs with aging genes, ML identifies activating transcription factor 3, KLF transcription factor 4, NFKBIA and superoxide dismutase 2 as downregulated diagnostic biomarkers associated with immune infiltration (99). In addition, tRNA-derived fragment 16 is upregulated in OA and worsens disease by targeting AlkB homolog 5 RNA demethylase, reducing NFKBIA mRNA stability, activating NF-κB and promoting inflammation and ECM degradation (100). Microarray and CIBERSORT analyses identifies macrophages as dominant infiltrates in OA, defines two immune-related OA subgroups and highlights hub genes including NFKBIA and MYC (101). The present study identified NFKBIA as a potential OA biomarker.

LDLR serves a pivotal role in OA pathogenesis. In a murine OA model, LDL accumulation through a cholesterol-rich diet or LDLR deficiency promotes synovial macrophage oxLDL uptake, increases S100A8 and TGF-β signaling and markedly increases osteophyte and enthesophyte formation (102). Through the use of drug-target Mendelian randomization, genetically proxied LDL reduction through 3-hydroxy-3-methylglutaryl-CoA reductase, proprotein convertase subtilisin/kexin type 9 and LDLR is associated with a reduced risk of OA (103). By integrating multi-tissue transcriptomics with ML and scRNA-seq, a previous study identifies five peripheral blood biomarkers, including LDLR, that may be used to predict OA, reveal age-specific expression profiles and B cell remodeling, as well as associations with immune infiltration and inflammatory cytokines (104). In the present study, it was revealed that LDLR increased cell viability, inhibited apoptosis and repressed the expression of MMP13 in ATDC5 cells, indicating a potential role for LDLR in OA pathogenesis. Intracellular cholesterol accumulation may be one of the upstream events involved in activating inflammatory pathways such as NF-κB, which is a key transcription factor driving the expression of matrix degrading enzymes such as MMP13. The results of the present study supported the concept that there may be a functional association between lipid metabolism disorders and the joint inflammation catabolic axis.

In conclusion, the present study integrated six GEO bulk transcriptomic cohorts with single-cell chondrocyte profiling and multi-algorithm ML to define inflammation-driven OA subtypes. A total of four key biomarkers were identified (ADM, LDLR, MYC and NFKBIA) and their expression was validated using scRNA-seq data. In addition, the present study added functional perturbation of ADM, LDLR, MYC and NFKBIA in ATDC5 cells. However, a number of limitations should be acknowledged. Firstly, the proposed mechanistic axis, such as the cholesterol/NF-κB/MMP13 axis, was primarily hypothesis-driven and was not directly investigated with causal experiments (such as pathway perturbation or rescue assays). Secondly, the biomarkers investigated in the present study relied on computational integration of public transcriptomic datasets, which may have introduced platform- and batch-related biases, residual confounding and model overfitting. Thirdly, the sample size was modest, which may limit the generalizability across diverse OA phenotypes and clinical stages. Furthermore, functional validation was carried out in ATDC5 cells, a mouse chondrogenic cell line, and potential differences compared with human OA chondrocytes and joint tissues were not evaluated. Subsequently, future work should prioritize validation of the present findings using multicenter human cohorts, including synovium and cartilage samples, longitudinal follow-up and treatment-response data. Mechanistic experiments using human primary chondrocytes, organoid models or mouse models with targeted perturbation of cholesterol handling and NF-κB activity should be carried out to determine causality. Lastly, the SVM-based model and nomogram may support earlier OA detection and immune-inflammation stratification from synovial or cartilage expression data, enabling risk scoring and cohort enrichment for trials. The LDLR and MMP13 could be targets and potential biomarkers for OA treatment and monitor response.

Acknowledgements

Not applicable.

Funding

The present study was supported by The Jiande City Science and Technology Development Project (grant nos. 2023SJZX13 and 2024SJZX08), The Zhejiang Provincial Medical and Health Science and Technology Project (grant no. 2022RC244) and The Hangzhou Agricultural and Social Development Project (grant no. 20241029Y206).

Availability of data and materials

The data generated in the present study may be requested from the corresponding author.

Authors' contributions

JX performed experiments, analyzed the data and wrote the manuscript. WQW, XTL, SWX and BZ performed experiments and analyzed the data. XL conceived and designed the study, drafted the manuscript and supervised the study. JX and XL confirm the authenticity of all the raw data. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Yao Q, Wu X, Tao C, Gong W, Chen M, Qu M, Zhong Y, He T, Chen S and Xiao G: Osteoarthritis: Pathogenic signaling pathways and therapeutic targets. Signal Transduct Target Ther 8: 56, 2023.
2. Tang S, Zhang C, Oo WM, Fu K, Risberg MA, Bierma-Zeinstra SM, Neogi T, Atukorala I, Malfait AM, Ding C and Hunter DJ: Osteoarthritis. Nat Rev Dis Primers 11: 10, 2025.
3. O'Connor D, Johnston RV, Brignardello-Petersen R, Poolman RW, Cyril S, Vandvik PO and Buchbinder R: Arthroscopic surgery for degenerative knee disease (osteoarthritis including degenerative meniscal tears). Cochrane Database Syst Rev 3: CD014328, 2022.
4. Richard MJ, Driban JB and McAlindon TE: Pharmaceutical treatment of osteoarthritis. Osteoarthritis Cartilage 31: 458-466, 2023.
5. Li J, Zhang H, Han Y, Hu Y, Geng Z and Su J: Targeted and responsive biomaterials in osteoarthritis. Theranostics 13: 931-954, 2023.
6. Hodgkinson T, Kelly DC, Curtin CM and O'Brien FJ: Mechanosignalling in cartilage: An emerging target for the treatment of osteoarthritis. Nat Rev Rheumatol 18: 67-84, 2022.
7. Miao MZ, Lee JS, Yamada KM and Loeser RF: Integrin signalling in joint development, homeostasis and osteoarthritis. Nat Rev Rheumatol 20: 492-509, 2024.
8. Zhang R, Zhang L, Tian B, Wang Y, Kang X and Zheng J: The gut-bone-cartilage triad: Microbial regulation of the Wnt/β-catenin signaling pathway in osteoarthritis joint remodeling (review). Mol Med Rep 33: 23, 2026.
9. Evans CH, Ghivizzani SC and Robbins PD: Osteoarthritis gene therapy in 2022. Curr Opin Rheumatol 35: 37-43, 2023.
10. Knights AJ, Redding SJ and Maerz T: Inflammation in osteoarthritis: The latest progress and ongoing challenges. Curr Opin Rheumatol 35: 128-134, 2023.
11. Rabie MA, Madry H, Cucchiarini M and El-Sayed NS: The brain-joint axis: Links between osteoarthritis and neurodegenerative disorders in aging. J Adv Res: October 21, 2025 (Epub ahead of print).
12. Motta F, Barone E, Sica A and Selmi C: Inflammaging and osteoarthritis. Clin Rev Allergy Immunol 64: 222-238, 2023.
13. Sanchez-Lopez E, Coras R, Torres A, Lane NE and Guma M: Synovial inflammation in osteoarthritis progression. Nat Rev Rheumatol 18: 258-275, 2022.
14. Wood MJ, Miller RE and Malfait AM: The genesis of pain in osteoarthritis: Inflammation as a mediator of osteoarthritis pain. Clin Geriatr Med 38: 221-238, 2022.
15. Hashimoto K, Oda Y, Nakagawa K, Ikeda T, Ohtani K and Akagi M: LOX-1 deficient mice show resistance to zymosan-induced arthritis. Eur J Histochem 62: 2847, 2018.
16. Hashimoto K, Mori S, Oda Y, Nakano A, Sawamura T and Akagi M: Lectin-like oxidized low density lipoprotein receptor 1-deficient mice show resistance to instability-induced osteoarthritis. Scand J Rheumatol 45: 412-422, 2016.
17. Lin Z, Deng Z, Liu J, Lin Z, Chen S, Deng Z and Li W: Chloride channel and inflammation-mediated pathogenesis of osteoarthritis. J Inflamm Res 15: 953-964, 2022.
18. Braaten JA, Banovetz MT, DePhillipo NN, Familiari F, Russo R, Kennedy NI and LaPrade RF: Biomarkers for osteoarthritis diseases. Life (Basel) 12: 1799, 2022.

19. Oliver GR, Hart SN and Klee EW: Bioinformatics for clinical next generation sequencing. Clin Chem 61: 124-135, 2015.
20. Thind AS, Monga I, Thakur PK, Kumari P, Dindhoria K, Krzak M, Ranson M and Ashford B: Demystifying emerging bulk RNA-Seq applications: The application and utility of bioinformatic methodology. Brief Bioinform 22: bbab259, 2021.
21. Vandereyken K, Sifrim A, Thienpont B and Voet T: Methods and applications for single-cell and spatial multi-omics. Nat Rev Genet 24: 494-515, 2023.
22. Qiao L, Gu J, Ni Y, Wu J, Zhang D and Gu Y: RNA-Seq Reveals the mRNAs, miRNAs, and lncRNAs expression profile of knee joint synovial tissue in osteoarthritis patients. J Clin Med 12: 1449, 2023.
23. Pemmari A, Leppänen T, Hämäläinen M, Moilanen T, Vuolteenaho K and Moilanen E: Widespread regulation of gene expression by glucocorticoids in chondrocytes from patients with osteoarthritis as determined by RNA-Seq. Arthritis Res Ther 22: 271, 2020.
24. Pemmari A, Tuure L, Hamalainen M, Leppänen T, Moilanen T and Moilanen E: Effects of ibuprofen on gene expression in chondrocytes from patients with osteoarthritis as determined by RNA-Seq. RMD Open 7: e001657, 2021.
25. Lv Z, Han J, Li J, Guo H, Fei Y, Sun Z, Dong J, Wang M, Fan C, Li W, *et al*: Single cell RNA-seq analysis identifies ferroptotic chondrocyte cluster and reveals TRPV1 as an anti-ferroptotic target in osteoarthritis. EBioMedicine 84: 104258, 2022.
26. Wang N, Liu K, Li JL, Pang WW, Zhang FR, Zeng Q, Deng Y, Qu XC, Chen XD, Deng HW and Tan LJ: Identification of osteoarthritis-related genes and potential drugs based on single cell RNA-seq data. Mol Med 32: 1, 2025.
27. Xu Q, Ma P, Wang R, Feng G and Jin Q: Prediction and validation based on scRNA-seq: ETS2 targets CEBPB to mediate osteoclast differentiation in osteoarthritis progression. Int J Biol Macromol 319: 145639, 2025.
28. Huber R, Hummert C, Gausmann U, Pohlers D, Koczan D, Guthke R and Kinne RW: Identification of intra-group, inter-individual, and gene-specific variances in mRNA expression profiles in the rheumatoid arthritis synovial membrane. Arthritis Res Ther 10: R98, 2008.
29. Ungethuem U, Haeupl T, Witt H, Koczan D, Krenn V, Huber H, von Helversen TM, Drungowski M, Seyfert C, Zacher J, *et al*: Molecular signatures and new candidates to target the pathogenesis of rheumatoid arthritis. Physiol Genomics 42A: 267-282, 2010.
30. Lauwerys BR, Hernández-Lobato D, Gramme P, Ducreux J, Dessy A, Focant I, Ambroise J, Bearzatto B, Nzeusseu Toukap A, Van den Eynde BJ, *et al*: Heterogeneity of synovial molecular patterns in patients with arthritis. PLoS One 10: e0122104, 2015.
31. Chou CH, Wu CC, Song IW, Chuang HP, Lu LS, Chang JH, Kuo SY, Lee CH, Wu JY, Chen YT, *et al*: Genome-wide expression profiles of subchondral bone in osteoarthritis. Arthritis Res Ther 15: R190, 2013.
32. Woetzel D, Huber R, Kupfer P, Pohlers D, Pfaff M, Driesch D, Häupl T, Koczan D, Stiehl P, Guthke R and Kinne RW: Identification of rheumatoid arthritis and osteoarthritis patients by transcriptome-based rule set generation. Arthritis Res Ther 16: R84, 2014.
33. Broeren MGA, de Vries M, Bennink MB, van Lent PLEM, van der Kraan PM, Koenders MI, Thurlings RM and van de Loo FAJ: Functional tissue analysis reveals successful cryopreservation of human osteoarthritic synovium. PLoS One 11: e0167076, 2016.
34. Brophy RH, Zhang B, Cai L, Wright RW, Sandell LJ and Rai MF: Transcriptome comparison of meniscus from patients with and without osteoarthritis. Osteoarthritis Cartilage 26: 422-432, 2018.
35. Fisch KM, Gamini R, Alvarez-Garcia O, Akagi R, Saito M, Muramatsu Y, Sasho T, Koziol JA, Su AI and Lotz MK: Identification of transcription factors responsible for dysregulated networks in human osteoarthritis cartilage by global gene expression analysis. Osteoarthritis Cartilage 26: 1531-1538, 2018.
36. Ji Q, Zheng Y, Zhang G, Hu Y, Fan X, Hou Y, Wen L, Li L, Xu Y, Wang Y and Tang F: Single-cell RNA-seq analysis reveals the progression of human osteoarthritis. Ann Rheum Dis 78: 100-110, 2019.
37. Leek JT, Johnson WE, Parker HS, Jaffe AE and Storey JD: The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 28: 882-883, 2012.
38. Zeidman P, Jafarian A, Corbin N, Seghier ML, Razi A, Price CJ and Friston KJ: A guide to group effective connectivity analysis, part 1: First level analysis with DCM for fMRI. Neuroimage 200: 174-190, 2019.
39. Liu H, Zhang X and Zhang X: Possible world based consistency learning model for clustering and classifying uncertain data. Neural Netw 102: 48-66, 2018.
40. Hu X, Ni S, Zhao K, Qian J and Duan Y: Bioinformatics-led discovery of osteoarthritis biomarkers and inflammatory infiltrates. Front Immunol 13: 871008, 2022.
41. Gentleman R, Carey V, Bates D, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, *et al*: Bioconductor: Open software development for computational biology and bioinformatic. Genome Biology 5, 2004.
42. Li P, Zhang W, Lu C, Zhang R and Li X: Robust kernel principal component analysis with optimal mean. Neural Netw 152: 347-352, 2022.
43. He Q, Fan B, Du P and Jin Y: Construction and validation of two hepatocellular carcinoma-progression prognostic scores based on gene set variation analysis. Front Cell Dev Biol 10: 806989, 2022.
44. Lu H, Wu J, Liang L, Wang X and Cai H: Identifying a novel defined pyroptosis-associated long noncoding RNA signature contributes to predicting prognosis and tumor microenvironment of bladder cancer. Front Immunol 13: 803355, 2022.
45. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria, 2025. https://www.R-project.org/.
46. Yue Y, Zhang Q and Sun Z: CX3CR1 acts as a protective biomarker in the tumor microenvironment of colorectal cancer. Front Immunol 12: 758040, 2022.
47. Li N, Li Y, Zheng P and Zhan X: Cancer stemness-based prognostic immune-related gene signatures in lung adenocarcinoma and lung squamous cell carcinoma. Front Endocrinol (Lausanne) 12: 755805, 2021.
48. Friedman J, Hastie T and Tibshirani R: Regularization paths for generalized linear models via coordinate descent. J Stat Softw 33: 1-22, 2010.
49. Parodi S, Verda D, Bagnasco F and Muselli M: The clinical meaning of the area under a receiver operating characteristic curve for the evaluation of the performance of disease markers. Epidemiol Health 44: e2022088, 2022.
50. Park SY: Nomogram: An analogue tool to deliver digital knowledge. J Thorac Cardiovasc Surg 155: 1793, 2018.
51. Jiang X, Tang F, Zhang J, He M, Xie T, Tang H, Liu J, Luo K, Lu S, Liu Y, *et al*: High GNG4 predicts adverse prognosis for osteosarcoma: Bioinformatics prediction and experimental verification. Front Oncol 13: 991483, 2023.
52. Cheng Y, Wang X and Xia Y: Supervised t-distributed stochastic neighbor embedding for data visualization and classification. INFORMS J Comput 33: 419-835, 2021.
53. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR, *et al*: Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol 20: 163-172, 2019.
54. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS and Rinn JL: The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 32: 381-386, 2014.
55. Li X, Liao Z, Deng Z, Chen N and Zhao L: Combining bulk and single-cell RNA-sequencing data to reveal gene expression pattern of chondrocytes in the osteoarthritic knee. Bioengineered 12: 997, 2021.
56. Xu X, Dai G, Liu CL, Yao Q, Cai X, Wang Y, Chen Z, Liu K, Zhu J, Ma J, *et al*: Fbxo2 suppresses prostate cancer progression by regulating YTHDF2 ubiquitination and degradation. Cell Death Dis 17: 153, 2025.
57. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25: 402-408, 2001.
58. Zeng R, Liu J, Lu F, Hong M, Lan T, Chen B, Pu Y, Tan Y, Wang P, Wang J and Wang W: Smurf2 enhances ubiquitin-mediated degradation of CASC3 and attenuates leukemia progression. iScience 28: 113411, 2025.
59. Hang S, Wang Q, Zhang J, Dong Y, Hu B, Wang P and Xu L: FBXO2 promotes hepatocellular carcinoma progression and sorafenib resistance by targeting USP49 for proteasomal degradation. Front Immunol 16: 1660034, 2025.
60. Kim G, Lim H, Kim Y, Kwon O and Choi JH: Intra-person multi-task learning method for chronic-disease prediction. Sci Rep 13: 1069, 2023.

61. Wu B, Fu L, Guo X, Hu H, Li Y, Shi Y, Zhang Y, Han S, Lv C and Tian Y: Multi-omics profiling and digital image analysis reveal the potential prognostic and immunotherapeutic properties of CD93 in stomach adenocarcinoma. Front Immunol 14: 984816, 2023.
62. Moody J, Kouno T, Chang JC, Ando Y, Carninci P, Shin JW and Hon CC: SCAFE: A software suite for analysis of transcribed cis-regulatory elements in single cells. Bioinformatics 38: 5126-5128, 2022.
63. Kozijn AE, Gierman LM, van der Ham F, Mulder P, Morrison MC, Kühnast S, van der Heijden RA, Stavro PM, van Koppen A, Pieterman EJ, *et al*: Variable cartilage degradation in mice with diet-induced metabolic dysfunction: Food for thought. Osteoarthritis Cartilage 26: 95-107, 2018.
64. Weng Q, Hu T, Shen X, Han J, Zhang Y and Luo J: Ezetimibe prevents IL-1β-induced inflammatory reaction in mouse chondrocytes via modulating NF-κB and Nrf2/HO-1 signaling crosstalk. Curr Pharm Biotechnol 23: 1772-1780, 2022.
65. Peng J, Chen J, Gao D, Liang B, Huang Z, Xiong B, Zhou S, Tan G, Zhong Z and Zeng X: Integrative bulk and single-cell transcriptomic profiling reveals oxidative stress-related genes and potential therapeutic targets in osteoarthritis. Mediators Inflamm 2025: 1240226, 2025.
66. Sun Q, Zhong Y, Huang G and Lin Y: Characterizing programmed cell death features in osteoarthritis through integrative multiomics and machine learning analysis. PeerJ 13: e20104, 2025.
67. Sun H, Chen K, Xiong Z, Zhuang Y, Liu M, Ning X and Yang H: Integrated analysis of bulk RNA and single-cell RNA sequencing data reveals potential biomarkers and immune infiltrates associated with N7-methylguanosine in osteoarthritis. Int J Surg 112: 595-613, 2026.
68. Zhang Z, Cheng D, Dang J, Wang X, Fan H and Liu D: Decoding endoplasmic reticulum stress on chondrocyte driving osteoarthritis development through integrating single-cell and transcriptomic profiling. Int J Med Sci 22: 3906-3923, 2025.
69. Hou W, Chang Y, Zheng Z and Liang H: Tryptophan metabolic dysregulation drives immune activation and cartilage degradation in osteoarthritis: Functional validation integrated with transcriptomic and single-cell RNA sequencing. Biochem Biophys Res Commun 782: 152497, 2025.
70. Kwart D, Gregg A, Scheckel C, Murphy EA, Paquet D, Duffield M, Fak J, Olsen O, Darnell RB and Tessier-Lavigne M: A large panel of isogenic APP and PSEN1 mutant human iPSC neurons reveals shared endosomal abnormalities mediated by APP β-CTFs, Not Aβ. Neuron 104: 256-270.e5, 2019.
71. Lanoiselée HM, Nicolas G, Wallon D, Rovelet-Lecrux A, Lacour M, Rousseau S, Richard AC, Pasquier F, Rollin-Sillaire A, Martinaud O, *et al*: APP, PSEN1, and PSEN2 mutations in early-onset Alzheimer disease: A genetic screening study of familial and sporadic cases. PLoS Med 14: e1002270, 2017.
72. Arber C, Lovejoy C, Harris L, Willumsen N, Alatza A, Casey JM, Lines G, Kerins C, Mueller AK, Zetterberg H, *et al*: Familial Alzheimer's disease mutations in PSEN1 lead to premature human stem cell neurogenesis. Cell Rep 34: 108615, 2021.
73. Raut S, Patel R and Al-Ahmad AJ: Presence of a mutation in PSEN1 or PSEN2 gene is associated with an impaired brain endothelial cell phenotype in vitro. Fluids Barriers CNS 18: 3, 2021.
74. Wang W, Ou Z, Peng J, Wang N and Zhou Y: Bioinformatics-based analysis of potential candidates chromatin regulators for immune infiltration in osteoarthritis. BMC Musculoskelet Disord 23: 1123, 2022.
75. King JD, Rowland G, Villasante Tezanos AG, Warwick J, Kraus VB, Lattermann C and Jacobs CA: Joint fluid proteome after anterior cruciate ligament rupture reflects an acute posttraumatic inflammatory and chondrodegenerative state. Cartilage 11: 329-337, 2020.
76. Kakiuchi Y, Yurube T, Kakutani K, Takada T, Ito M, Takeoka Y, Kanda Y, Miyazaki S, Kuroda R and Nishida K: Pharmacological inhibition of mTORC1 but not mTORC2 protects against human disc cellular apoptosis, senescence, and extracellular matrix catabolism through Akt and autophagy induction. Osteoarthritis Cartilage 27: 965-976, 2019.
77. Rong G, Zhang Z, Zhan W, Chen M, Ruan J and Shen C: VEGFA, MYC, and JUN are abnormally elevated in the synovial tissue of patients with advanced osteoarthritis. Sci Rep 15: 2066, 2025.
78. Wang W, Mai H, Xu H, Jing B, Yu C, Li X, Chen D, Huang Y, Shao M and Pan T: 4,8-Dicarboxyl-8,9-iridoid-1-glycoside inhibits apoptosis in human osteoarthritis chondrocytes via enhanced c-MYC-mediated cholesterol metabolism in vitro. Arthritis Res Ther 25: 240, 2023.
79. Shu Y, Long J, Guo W and Ye W: MicroRNA-195-5p inhibitor prevents the development of osteoarthritis by targeting REGγ. Mol Med Rep 19: 4561-4568, 2019.
80. Ge X, Tsang K, He L, Garcia RA, Ermann J, Mizoguchi F, Zhang M, Zhou B, Zhou B and Aliprantis AO: NFAT restricts osteochondroma formation from entheseal progenitors. JCI Insight 1: e86254, 2016.
81. Zhang Z, Yang P, Wang C and Tian R: LncRNA CRNDE hinders the progression of osteoarthritis by epigenetic regulation of DACT1. Cell Mol Life Sci 79: 405, 2022.
82. Li S, Han J, Cao J, Han H, Lu B, Wen T and Bian W: ADORA2B, transcriptionally suppressing by MYC, promotes ferroptosis of chondrocytes via inhibition of the PI3K/Akt pathway in mice with osteoarthritis. Environ Toxicol 39: 2487-2501, 2024.
83. Zhang Z, Liu W, Xiong J, Chen T, Jiang L and Liu M: Candidate marker genes for diagnosis of osteoarthritis and prediction of their regulatory mechanisms. Folia Biol (Praha) 69: 22-33, 2023.
84. Liu X, Li G, Liu R, Yang L, Li L, Goswami A, Deng K, Dong L, Shi H and He X: Transcriptome combined with single cell to explore hypoxia-related biomarkers in osteoarthritis. J Chromatogr B Analyt Technol Biomed Life Sci 1246: 124274, 2024.
85. He H, Zhao X, Zhang B, Zhao S and Wu Y: Integrated multi-omics and machine learning reveals immune-metabolic signatures in osteoarthritis: From bulk RNA-seq to single-cell resolution. Front Immunol 16: 1599930, 2025.
86. Mao W, Bao Z, Zhang B and Wu L: Construction of a glycolysis-related diagnostic model for osteoarthritis through integrated bioinformatics analysis and machine learning. J Orthop Surg Res 20: 639, 2025.
87. Hopwood B, Tsykin A, Findlay DM and Fazzalari NL: Microarray gene expression profiling of osteoarthritic bone suggests altered bone remodelling, WNT and transforming growth factor-beta/bone morphogenic protein signalling. Arthritis Res Ther 9: R100, 2007.
88. Lee MS, Sun MT, Pang ST, Ueng SW, Chen SC, Hwang TL and Wang TH: Evaluation of differentially expressed genes by shear stress in human osteoarthritic chondrocytes in vitro. Chang Gung Med J 32: 42-50, 2009.
89. Qi X, Yu F, Wen Y, Li P, Cheng B, Ma M, Cheng S, Zhang L, Liang C, Liu L and Zhang F: Integration of transcriptome-wide association study and messenger RNA expression profile to identify genes associated with osteoarthritis. Bone Joint Res 9: 130-138, 2020.
90. Cao J, Ding H, Shang J, Ma L, Wang Q and Feng S: Weighted gene co-expression network analysis reveals specific modules and hub genes related to immune infiltration of osteoarthritis. Ann Transl Med 9: 1525, 2021.
91. Yao S, Deng M, Du X, Huang R and Chen Q: A novel hypoxia related marker in blood link to aid diagnosis and therapy in osteoarthritis. Genes (Basel) 13: 1501, 2022.
92. Uzan B, Ea HK, Launay JM, Garel JM, Champy R, Cressent M and Lioté F: A critical role for adrenomedullin-calcitonin receptor-like receptor in regulating rheumatoid fibroblast-like synoviocyte apoptosis. J Immunol 176: 5548-5558, 2006.
93. Zhang Q, Sun C, Liu X, Zhu C, Ma C and Feng R: Mechanism of immune infiltration in synovial tissue of osteoarthritis: A gene expression-based study. J Orthop Surg Res 18: 58, 2023.
94. Gao H, Di J, Yin M, He T, Wu D, Chen Z, Li S, He L and Rong L: Identification of chondrocyte subpopulations in osteoarthritis using single-cell sequencing analysis. Gene 852: 147063, 2023.
95. Zhou J, Huang J, Li Z, Song Q, Yang Z, Wang L and Meng Q: Identification of aging-related biomarkers and immune infiltration characteristics in osteoarthritis based on bioinformatics analysis and machine learning. Front Immunol 14: 1168780, 2023.
96. Tang H, Cheng Z, Ma W, Liu Y, Tong Z, Sun R and Liu H: TLR10 and NFKBIA contributed to the risk of hip osteoarthritis: Systematic evaluation based on Han Chinese population. Sci Rep 8: 10243, 2018.
97. Cai P, Jiang T, Li B, Qin X, Lu Z, Le Y, Shen C, Yang Y, Zheng L and Zhao J: Comparison of rheumatoid arthritis (RA) and osteoarthritis (OA) based on microarray profiles of human joint fibroblast-like synoviocytes. Cell Biochem Funct 37: 31-41, 2019.

98. Wang C, Gao Y, Zhang Z, Chi Q, Liu Y, Yang L and Xu K: Safflower yellow alleviates osteoarthritis and prevents inflammation by inhibiting PGE2 release and regulating NF-κB/SIRT1/AMPK signaling pathways. Phytomedicine 78: 153305, 2020.
99. Du J, Zhou T, Dong Y, Sun Y and Peng W: Identification and validation of aging related genes in osteoarthritis. Front Genet 16: 1561644, 2025.
100. Zhu C, Liu G, Zhou W, Lu Y and Chen W: tRF16 affects NFKBIA stability and promotes osteoarthritis progression by regulating ALKBH5 expression in m6A-dependent manner. Commun Biol 8: 868, 2025.
101. Jiang W, Li T, Tang Z, Li S and Lu M: Association of different types of osteoarthritis with different immune-related subtypes and hub genes. Altern Ther Health Med 30: 466-473, 2024.
102. de Munter W, Blom AB, Helsen MM, Walgreen B, van der Kraan PM, Joosten LAB, van den Berg WB and va Lent PLEM: Cholesterol accumulation caused by low density lipoprotein receptor deficiency or a cholesterol-rich diet results in ectopic bone formation during experimental osteoarthritis. Arthritis Res Ther 15: R178, 2013.
103. Ma W, Chen H, Zhang Z and Xiong Y: Association of lipid-lowering drugs with osteoarthritis outcomes from a drug-target Mendelian randomization study. PLoS One 19: e0293960, 2024.
104. Tu B, Fang R, Lu P, Liu M, Yang S, Hu D, Ruan R and Ning R: Multi-omic profiling reveals age-specific blood biomarkers and aging-driven B cell remodeling in osteoarthritis. Int J Surg 111: 7814-7828, 2025.