

# Classification of distinct osteoarthritis subtypes with different knee joint tissues by gene expression profiles

Cite this article:

*Bone Joint Res* 2023;12(12):702–711.

DOI: 10.1302/2046-3758.1212.BJR-2023-0021.R2

Correspondence should be sent to Jiaqian Wang  
1580443814@qq.com

Y. Xue,<sup>1</sup> L. Zhou,<sup>2</sup> J. Wang<sup>3</sup>

<sup>1</sup>Department of Orthopaedic, Wuxi Ninth People's Hospital of Soochow University, Wuxi, China

<sup>2</sup>Department of Orthopaedic, Lianshui County People's Hospital, Huai'an, China

<sup>3</sup>Department of Orthopaedic, Zhongshan Hospital, Fudan University, Shanghai, China

## Aims

Knee osteoarthritis (OA) involves a variety of tissues in the joint. Gene expression profiles in different tissues are of great importance in order to understand OA.

## Methods

First, we obtained gene expression profiles of cartilage, synovium, subchondral bone, and meniscus from the Gene Expression Omnibus (GEO). Several datasets were standardized by merging and removing batch effects. Then, we used unsupervised clustering to divide OA into three subtypes. The gene ontology and pathway enrichment of three subtypes were analyzed. CIBERSORT was used to evaluate the infiltration of immune cells in different subtypes. Finally, OA-related genes were obtained from the Molecular Signatures Database for validation, and diagnostic markers were screened according to clinical characteristics. Quantitative reverse transcription polymerase chain reaction (qRT-PCR) was used to verify the effectiveness of markers.

## Results

C1 subtype is mainly concentrated in the development of skeletal muscle organs, C2 lies in metabolic process and immune response, and C3 in pyroptosis and cell death process. Therefore, we divided OA into three subtypes: bone remodelling subtype (C1), immune metabolism subtype (C2), and cartilage degradation subtype (C3). The number of macrophage M0 and activated mast cells of C2 subtype was significantly higher than those of the other two subtypes. COL2A1 has significant differences in different subtypes. The expression of COL2A1 is related to age, and trafficking protein particle complex subunit 2 is related to the sex of OA patients.

## Conclusion

This study linked different tissues with gene expression profiles, revealing different molecular subtypes of patients with knee OA. The relationship between clinical characteristics and OA-related genes was also studied, which provides a new concept for the diagnosis and treatment of OA.

## Article focus

- This study linked different tissues with gene expression profiles, revealing different molecular subtypes of patients with knee osteoarthritis (OA).

- We also studied the relationship between clinical characteristics and OA-related genes.

## Key messages

- Through unsupervised clustering, we divided the OA gene expression profiles

of different tissues into three subtypes: bone remodelling subtype (C1), immune metabolism subtype (C2), and cartilage degradation subtype (C3).

### Strengths and limitations

- Unlike other unsupervised cluster analyses, our experiment covers almost all tissues of the knee joint, which is more comprehensive.
- We studied OA at the molecular level, and classified OA more accurately by combining clinical characteristics.
- A limitation of the study is that the tissue sample size is still too small, and further experiments are needed to verify the research.

### Introduction

Osteoarthritis (OA) is the most common disabling disease, within which knee OA has the highest incidence rate at present.<sup>1</sup> Patients tend to show pain, stiffness, dysfunction, and even deformity of the knee joint, which seriously affects quality of life.<sup>2</sup> At the time of writing, the aetiology of OA has not been fully clarified, and the main pathogenesis is the degeneration of articular cartilage.<sup>3</sup> However, cartilage wear caused by mechanical stress cannot explain all the problems. Genetic mechanism and immune inflammatory mechanism all play an important role in the occurrence and development of OA.<sup>4,5</sup> Components in the knee joint such as synovium, subchondral bone, and meniscus have their unique clinical manifestations. When the synovium of the knee joint is stimulated, it will produce inflammation, release various inflammatory factors, and cause joint effusion at the same time.<sup>6</sup> Subchondral bone is located in the deep layer of cartilage, and different patients show cystic changes or osteophyte formation.<sup>7</sup> Therefore, it is inappropriate to apply the same diagnosis and treatment to all patients. We believe that it is of great importance to classify OA subtypes based on knee components for early diagnosis and treatment.<sup>8</sup>

So far, clinical OA classification is mainly based on aetiology, clinical symptoms, imaging changes, and other factors. The Kellgren-Lawrence (KL)<sup>9</sup> classification is the most commonly used indicator, but this only focuses on the severity of knee damage shown on plain radiograph.<sup>10</sup> These indicators are not sensitive to early OA, and the first change in patients is often at the molecular level.<sup>11,12</sup> Studies have confirmed that OA involves changes in a variety of molecular pathways and markers, which are closely related to the functional changes of synovium, cartilage, and subchondral bone.<sup>13</sup> Inflammatory factor markers and cartilage metabolism markers are currently the two most widely studied indicators.<sup>13</sup> Interleukin 1 beta (IL-1 $\beta$ ) is an important inflammatory mediator that can stimulate synovium to secrete a variety of inflammatory factors and promote cartilage apoptosis. Detecting the content of IL-1 $\beta$  in patients' knee joints is helpful to judge the condition.<sup>14</sup> In addition, cartilage oligomeric matrix protein (COMP) is the degradation product of articular cartilage. In the early stages of OA, radiographs cannot highlight the damage of articular cartilage, yet the level of COMP has already increased.<sup>15</sup> However, the molecular communication within and between tissues is still unclear, and there is also crosstalk between these tissues in OA.<sup>16</sup> Hence, it is necessary to classify OA subtypes by combining gene expression information and knee joint tissue.<sup>17</sup>

Some studies have identified OA subtypes in individual tissues. Cao et al<sup>18</sup> discovered subtypes of lipid metabolism disorders and potential molecular markers ADCY7 in synovial tissue. In order to construct OA subtypes more comprehensively, we determined three subtypes by unsupervised clustering based on gene expression profiling of four tissues, defining these subtypes based on molecular characteristics and identifying the differences. Furthermore, we explored the relationship between OA markers and immune cells, patient age, and sex.

### Methods

#### Dataset collection of different tissues

We searched the relevant gene expression profiles from the Gene Expression Omnibus (GEO)<sup>19</sup> with the keyword "osteoarthritis". GEO is an international public repository that collects and organizes genomic data uploaded by researchers worldwide. The specific inclusion criteria are as follows: 1) Homo sapiens microarray analysis of OA; 2) knee joint tissue with sample size greater than 10; 3) the sample has complete clinical information; and 4) the KL grade of the patient's knee joint is the same. According to the inclusion criteria, five datasets of four different knee joint tissues were obtained. Cartilage sample dataset (GSE114007) of OA patients, synovial datasets (GSE55457, GSE12021), subchondral bone dataset (GSE1588), and meniscus dataset (GSE98918) were selected. Finally, 92 OA samples were obtained from the datasets, and the control samples were removed.

#### Removal of batch effects and merging of datasets

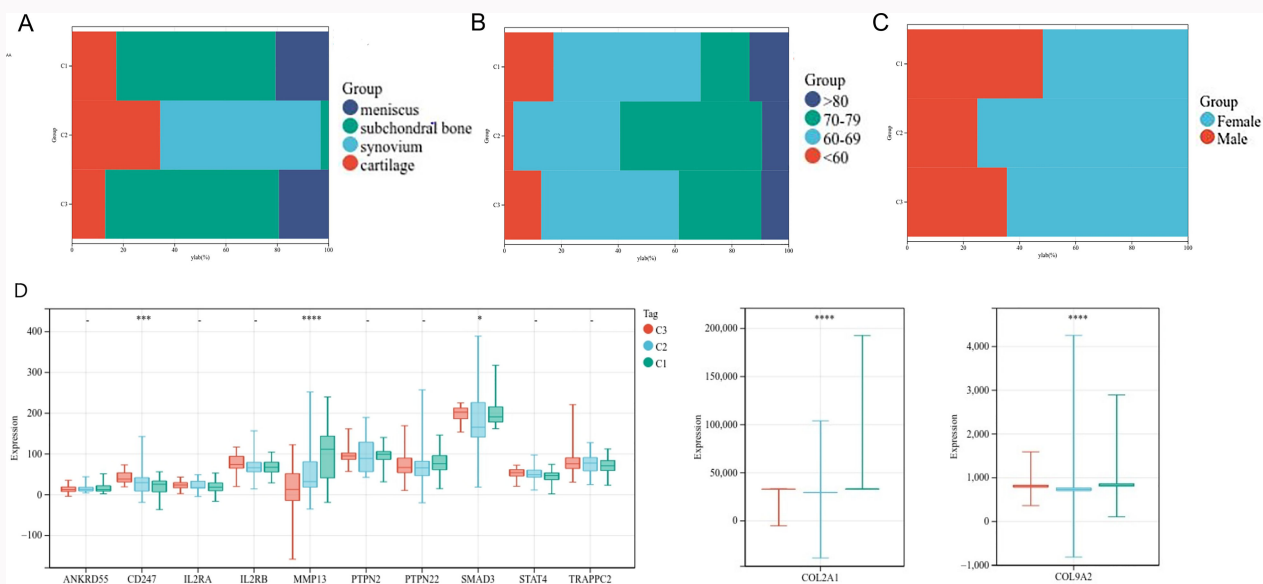
All R packages are implemented through RStudio (v4.1.3; USA). The raw data of five datasets were obtained from the series of matrix files and supplementary files. The probe expression matrix was transformed into gene expression matrix, and the rows and columns with the missing value ratio greater than 50% were deleted. The missing values were completed with "impute" package. After using the R package "inSilicoMerging" to merge five datasets, the "ComBat" algorithm was applied to remove the batch effect. Finally, we obtained a merging normalized gene expression matrix to enhance the reliability of the results.

#### Unsupervised cluster analysis

In order to clarify the different molecular subtypes of OA patients, the combined datasets were analyzed by unsupervised cluster analysis. The consistency clustering method is to perform subsampling from sample data and determine the clusters with a specific cluster count (k). Then, two items with the same number of occurrences in the same subsample have the same clustering, which is calculated and stored in the symmetric consistent matrix of each k. R package 'ConsenseClusterPlus' based on the 'pam' method was used, and resampled 80% of the samples for ten repetitions. The optimal number of clusters was determined according to the empirical cumulative distribution function plot and the tSNE diagram constructed by R software package 'Rtsne'. Finally, all tissue samples were divided into three subtypes.

#### Screening and enrichment analysis of DEGs

The differential expression analysis of different OA subtypes was carried out by the Linear models for microarray data



**Fig. 1** Clinical features and osteoarthritis (OA)-related gene validation. a) Tissue composition of different subtypes. b) Age composition of different subtypes. c) Sex composition of different subtypes. d) Expression of OA-related genes in different subtypes. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ , Mann-Whitney U test.

‘limma’ package. The differentially expressed genes (DEGs) were shown by volcano map, and  $p$ -value  $< 0.05$  and  $|\log_2FC| > 1$  were considered to be statistically significant. The ten up-regulated and ten down-regulated genes with the largest differential expression were shown by heatmap.

When we obtain DEG and wish to observe the function or pathway of these genes, enrichment analysis is the most convenient method. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of DEGs are performed by R package ‘clusterProfiler’. GO functional enrichment analysis includes cellular component (CC), molecular function (MF), and biological process (BP). KEGG is a pathway enrichment database. Ten significantly different GO terms and signal pathways were screened by  $p$ -value  $< 0.05$  and  $q$  value  $< 0.05$ . Gene set enrichment analysis (GSEA) compares the differential expression of all genes in two subtypes to improve the accuracy of enrichment analysis.

### Construction of PPI network

The protein–protein interaction (PPI) network was predicted using online tool STRING (ELIXIR, UK, and Global Biodata Coalition). By analyzing the interaction between candidate proteins from laboratory data, other databases, text mining, and predictive bioinformatics data, we can deeply understand the mechanism of disease occurrence or development. Interactions with a composite score greater than 0.4 were considered statistically significant. We downloaded the PPI network and used Cytoscape software (v3.8.0; National Institutes of Health, USA) for better visualization.

### Evaluation of immune cell infiltration

Based on our gene expression profile, the CIBERSORT algorithm was used to compare the infiltration of immune cells between different subtypes. CIBERSORT transformed the normalized gene expression matrix into the composition of

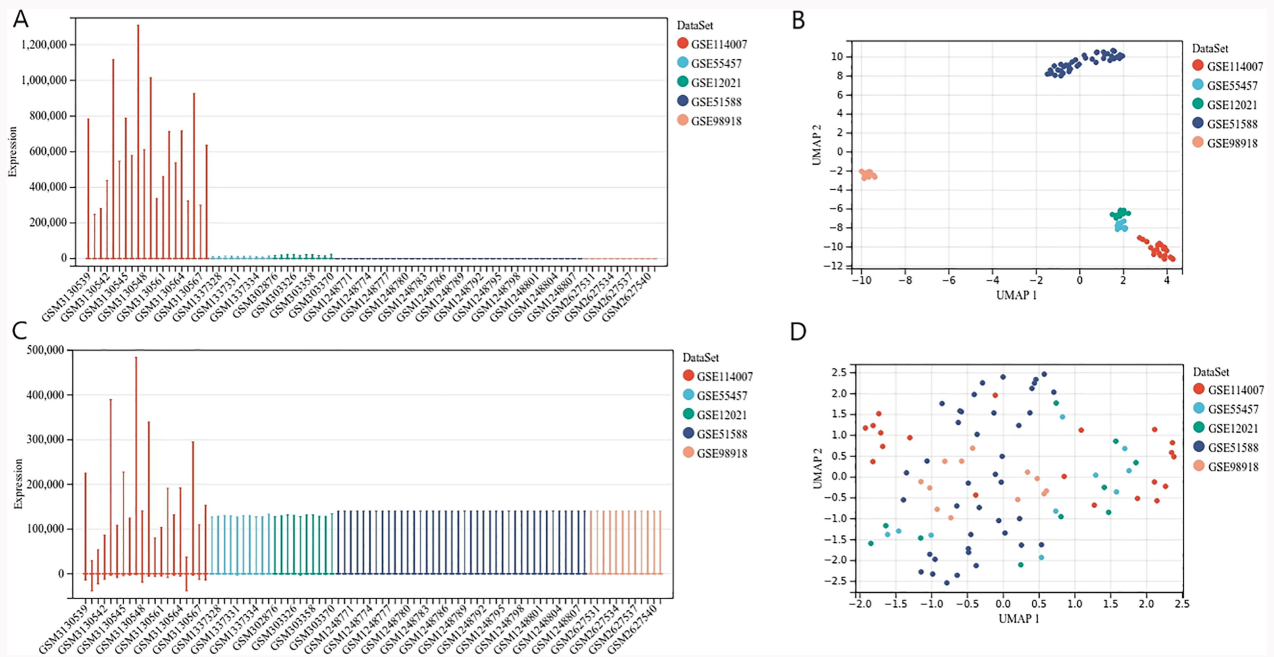
immune cells using the LM22 characteristic matrix (Stanford University, USA). A violin diagram compared the differences of immune cells between OA patients with different clinical characteristics. The relationship between immune cells and OA related genes was shown by correlation heatmap.

### Clinical features and OA-related gene validation

The clinical characteristics of 92 tissue samples corresponding to patients were constructed by histogram. OA-related genes were downloaded from MSigDB database.<sup>20</sup> This database constructs a gene collection from multiple perspectives such as location, function, metabolic pathways, and targets. A violin diagram visualizes the relationship between different subtypes, different tissues, different ages, different sexes, and OA-related genes (Figure 1). All samples and corresponding clinical information are listed in Supplementary Table i. The diagnostic value of OA-related genes was evaluated by the operating characteristic curve (ROC) of subjects. ROC analysis was performed by the ‘pROC’ software package to obtain the area under the curve (AUC). An AUC greater than 0.75 indicates that biomarkers have good diagnostic value.

### qRT-PCR validation

In order to further verify the relationship between clinical characteristics and genetic markers, samples were obtained from 16 OA patients of Chinese ethnicity who underwent knee arthroplasty. They were grouped according to the type of tissues, sex, and age, as shown in Supplementary Table ii. All participants gave written informed consent. Quantitative real-time polymerase chain reaction (qRT-PCR) was carried out as the method previously described.<sup>21</sup> Simply put, total RNA was extracted from tissues using TRIzol reagent (Beyotime, China) and reverse transcribed. qRT-PCR was performed on CFX connect real-time PCR detection system (Bio-Rad, USA). The relative gene expressions were calculated by the  $2^{-\Delta\Delta Ct}$  method.



**Fig. 2** Remove batch effects and merge datasets. a) Box plot of five osteoarthritis (OA) datasets before batch correction. b) Five OA datasets before batch correction, as calculated using uniform manifold approximation and projection (UMAP). c) Box plot after batch correction. d) UMAP after batch correction.

Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was selected to normalize the expression levels of the target genes. All experiments were performed independently in triplicate. Specific primer sequences are shown in Supplementary Table iii. All methods were performed in accordance with the relevant guidelines and regulations.

### Statistical analysis

Paired *t*-test and Mann-Whitney U test were used to assess the significant differences between two groups. GraphPad Prism software (version 5.01; USA) was used for statistical analyses. Any *p*-values < 0.05 were considered significant.

## Results

### Identification of OA molecular subtypes

The flowchart of this study is shown in Supplementary Figure a. Five microarray raw datasets were used in total, including 20 cartilage samples from OA patients, 20 synovial samples from the same platform, 40 subchondral bone samples, and 12 meniscus samples (Figure 2a and 2b). The data after batch correction are presented in Figure 2c and 2d, which indicate that the batch effect of the merged data has been successfully eliminated. According to the mean consistency evaluation within the cluster group, the consistency is better when the specific cluster count  $k = 2$  to 4 (Figure 3a and 3b). When  $k = 3$ , tSNE analysis showed that there were differences among subtypes, especially cluster 1 and cluster 3 (Figure 3c and 3d). Thus, 92 OA tissue samples were divided into three subtypes: Cluster 1 ( $n = 29$ ), Cluster 2 ( $n = 32$ ), and Cluster 3 ( $n = 31$ ) (Figure 3e).

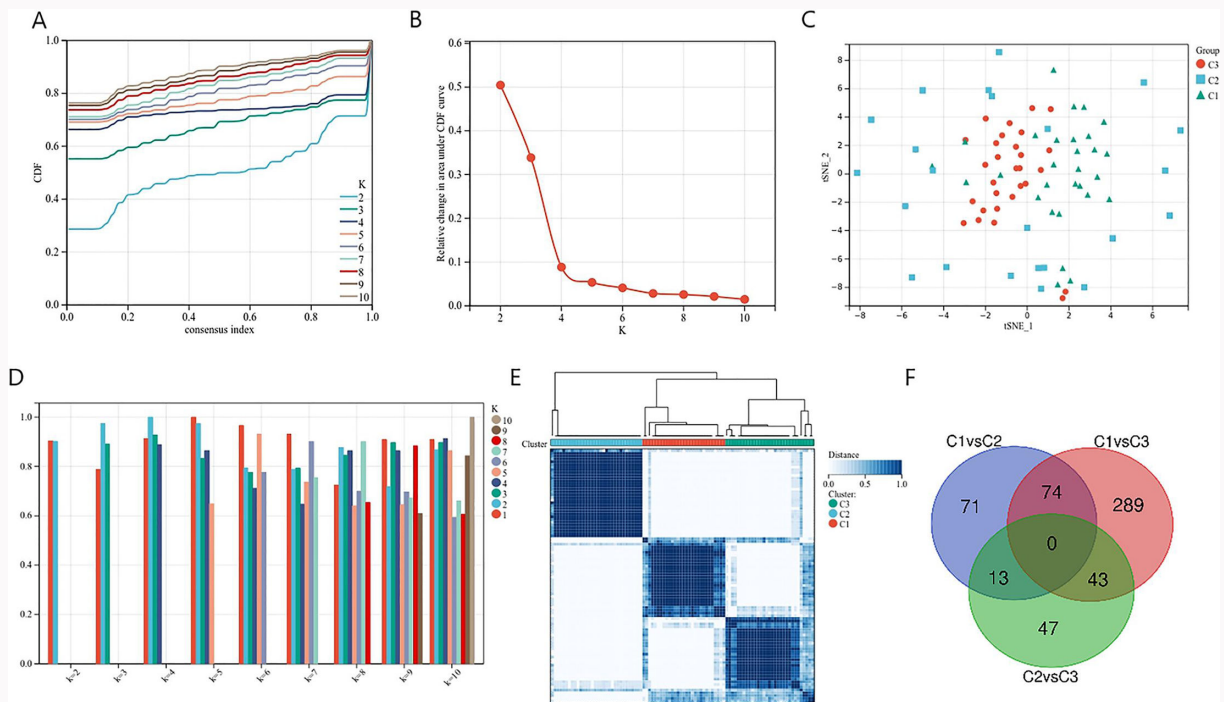
### DEGs and enrichment analysis of different subtypes

Compared with C1 and C2 subtypes, we found 128 up-regulated DEGs and 30 down-regulated DEGs (Figure 4a). Between

C1 and C3 subtypes, we found 239 up-regulated DEGs and 167 down-regulated DEGs (Figure 4b). Overall, 35 up-regulated DEGs and 68 down-regulated DEGs were obtained between C2 and C3 subtypes (Figure 4c). No common DEGs were found among the three groups of DEGs (Figure 3f). The most significant ten up-regulated DEGs and ten down-regulated DEGs were visualized by heatmap (Figure 4d and 4f). The protein-protein interaction (PPI) network diagram of DEGs is shown in Supplementary Figure b. The DEGs of C1 versus C2 are marked in blue, C1 versus C3 in red, and C2 versus C3 in green.

GO enrichment analysis showed that the differences between C1 and C2 subtypes were concentrated in the plasma membrane part, transporter activity, and organ development (Figure 5a). KEGG pathway analysis shows that the difference between the two lies in systemic lupus erythematosus, tyrosine metabolism, and other reactions (Figure 5d). Figure 5e compares the GO and KEGG differences between C1 and C3 subtypes. GO analysis focuses specifically on the plasma membrane part and condensed chromosome region. The functional enrichment of C2 and C3 DEGs lies in the metabolic process and filament sliding (Figure 5c). Cell adhesion molecules, systemic lupus erythematosus, and some other immune responses are the main enrichment pathway (Figure 5f).

GSEA analyzes the entire gene dataset and will not ignore genes with insignificant differential expression. C1 subtype function is enriched in the differentiation and development of bone joints and chondrocytes, C2 is enriched in immune cell regulation and some metabolic processes, and C3 is involved in pyroptosis and programmed necrotic cell death (Supplementary Figures ca to cc). KEGG also showed similar results for further verification of the characteristics of each subtype (Supplementary Figures cd to cf).



**Fig. 3**

Clustering analysis based on gene expression profiles of knee joint tissue. a) Clustering cumulative distribution function (CDF) curve. b) Relative change in area under the CDF curve. c) tSNE diagram of subgroups. d) Samples clustering consistency, determine  $k = 3$ . e) Clustering heatmap. f) Venn diagram of differentially expressed genes.

### Differences in immune cell infiltration

As shown in **Figure 6a**, there are many immune cells involved in OA, such as B cells, plasma cells, T cells, natural killer (NK) cells, macrophages, and mast cells, and there is correlation between them. The violin diagram shows that macrophage M2 and resting mast cells are the main immune cells in the tissues of OA patients. There were significant differences in plasma cells, monocytes, and mast cells resting among different subtypes. The proportion of regulatory T cells (Tregs) of C2 subtype is higher than that of C3 subtype, and the proportion of mast cells activated is higher than that of C1 subtype (**Figure 6b**). CD247, IL2RB, and STAT4 have a stronger correlation with immune cells, and are positively correlated with T follicular helper cells, Tregs, and  $\gamma\delta$  T cells. Patient age was positively correlated with memory resting CD4 T cells (**Figure 6c**). There are more monocytes in cartilage and meniscus, and no difference in immune cells between them. There are more mast cells activated and fewer dendritic cells resting in synovium (**Figure 6d**). There was no significant difference in immune cells between patients of different ages (**Figure 6e**). Female patients had more M0 macrophages and fewer M2 macrophages (**Figure 6f**).

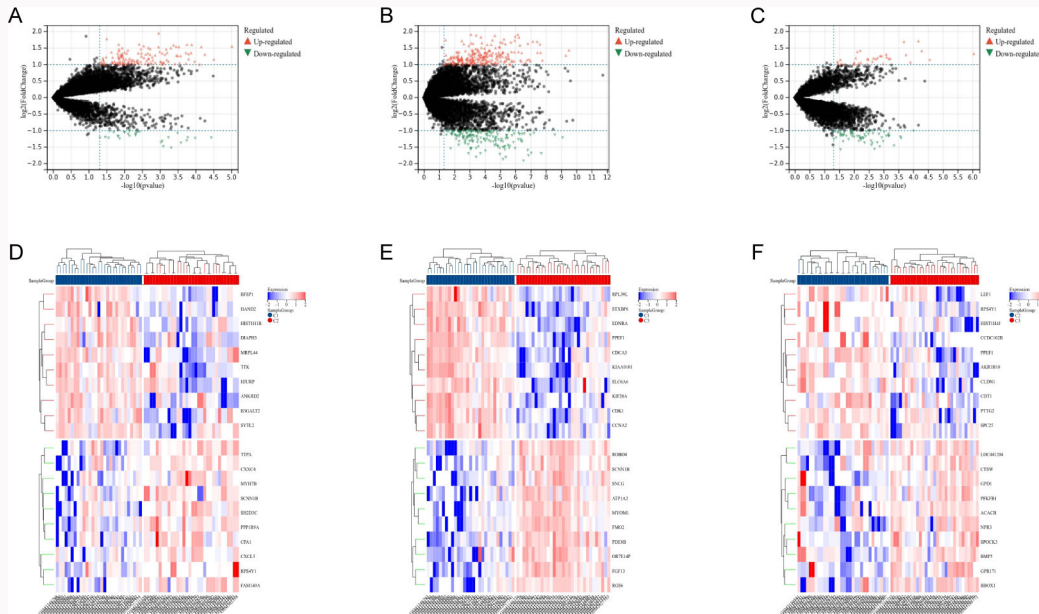
### Association between clinical features and OA-related genes

The C2 subtype contains all synovial tissue and no meniscus tissue (**Figure 1a**). There are more patients aged 70 to 79 years, and more female patients, in the C2 subtype group than the other two groups, but there is no significant difference between them (**Figure 1b and 1c**). There were, however, significant differences in CD247, matrix metalloproteinase (MMP)13, collagen type II alpha 1 chain (COL2A1), and COL9A2 among the three subtypes. MMP13, COL2A1,

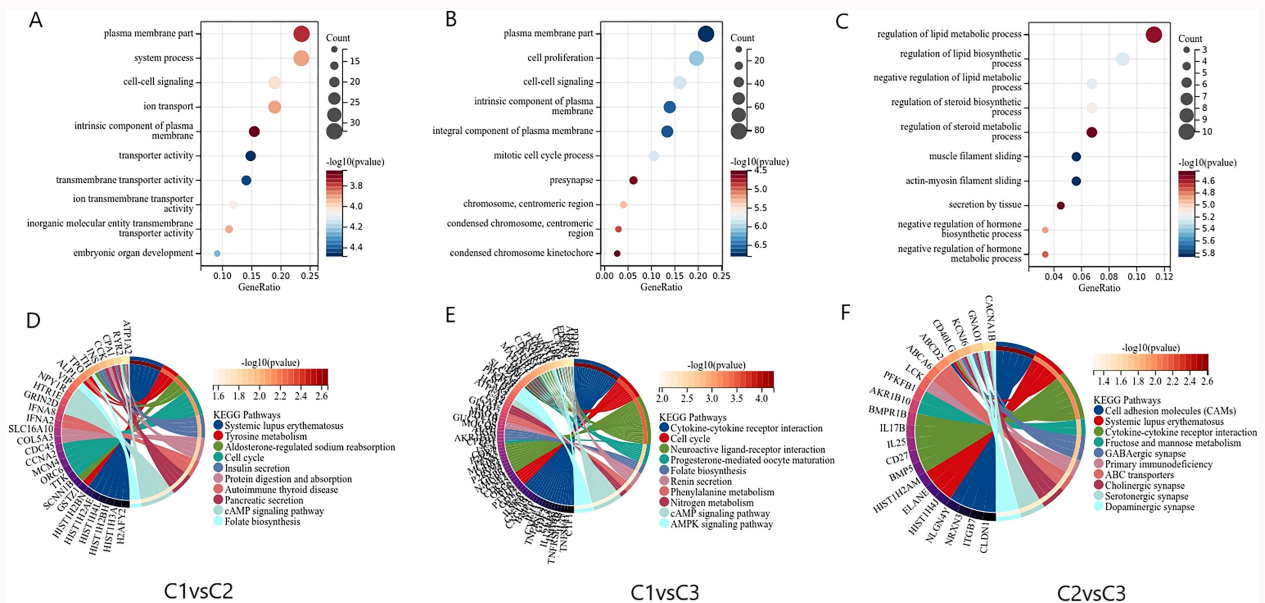
and COL9A2 are highly expressed in C1 subtype but less so in C3 subtype (**Figure 1d**). COL2A1 and COL9A2 have a relatively low expression in cartilage but are highly expressed in subchondral bone and meniscus (Supplementary Figure da). COL2A1 is highly expressed in young patients and less so in elderly patients. MMP13 expression differs according to age groups, but it is higher in both younger and older patients (Supplementary Figure db). In female patients, the expression of interleukin 2 receptor subunit beta (IL2RB) and trafficking protein particle complex subunit 2 (TRAPPC2) was higher, while the expression of protein tyrosine phosphatase non-receptor Type 2 (PTPN2) was lower (Supplementary Figure dc).

### Validation of OA-related genes

In order to further verify the effectiveness of gene diagnosis, ROC curve and qRT-PCR were used for verification. AUC among the different subtypes of COL1A2 was greater than 0.75: C1 vs C2 (AUC = 0.88), C2 vs C3 (AUC = 0.77), and C1 vs C3 (AUC = 0.84). MMP13 showed a good distinction (AUC > 0.75) between C1 vs C3 (AUC = 0.83), and COL9A2 has a good distinction between C2 vs C1 (AUC = 0.82) and C2 vs C3 (AUC = 0.80) (Supplementary Figure e). After further verification by qRT-PCR, no statistical differences were found in the GAPDH expression among OA patients of different tissues, ages, and sexes. COL2A1 and COL9A2 demonstrated a relatively low expression in cartilage, and a high expression in subchondral bone (**Figure 7a and 7b**). COL2A1 is highly expressed in OA patients under the age of 60 years, and less so in patients over the age of 80 years (**Figure 7c and 7d**). PTPN2 was significantly down-regulated in females with OA ( $p = 0.015$ , paired t-test),



**Fig. 4** Identification of differentially expressed genes (DEGs). a) Volcano map showing DEGs of C1 subtype vs C2 subtype. b) DEGs of C1 vs C3. c) DEGs of C2 vs C3. d) Heatmap shows the top ten up-regulated and down-regulated C1 vs C2 DEGs. e) Top ten up-regulated and down-regulated C1 vs C3 DEGs. f) Top ten up-regulated and down-regulated C2 vs C3 DEGs.



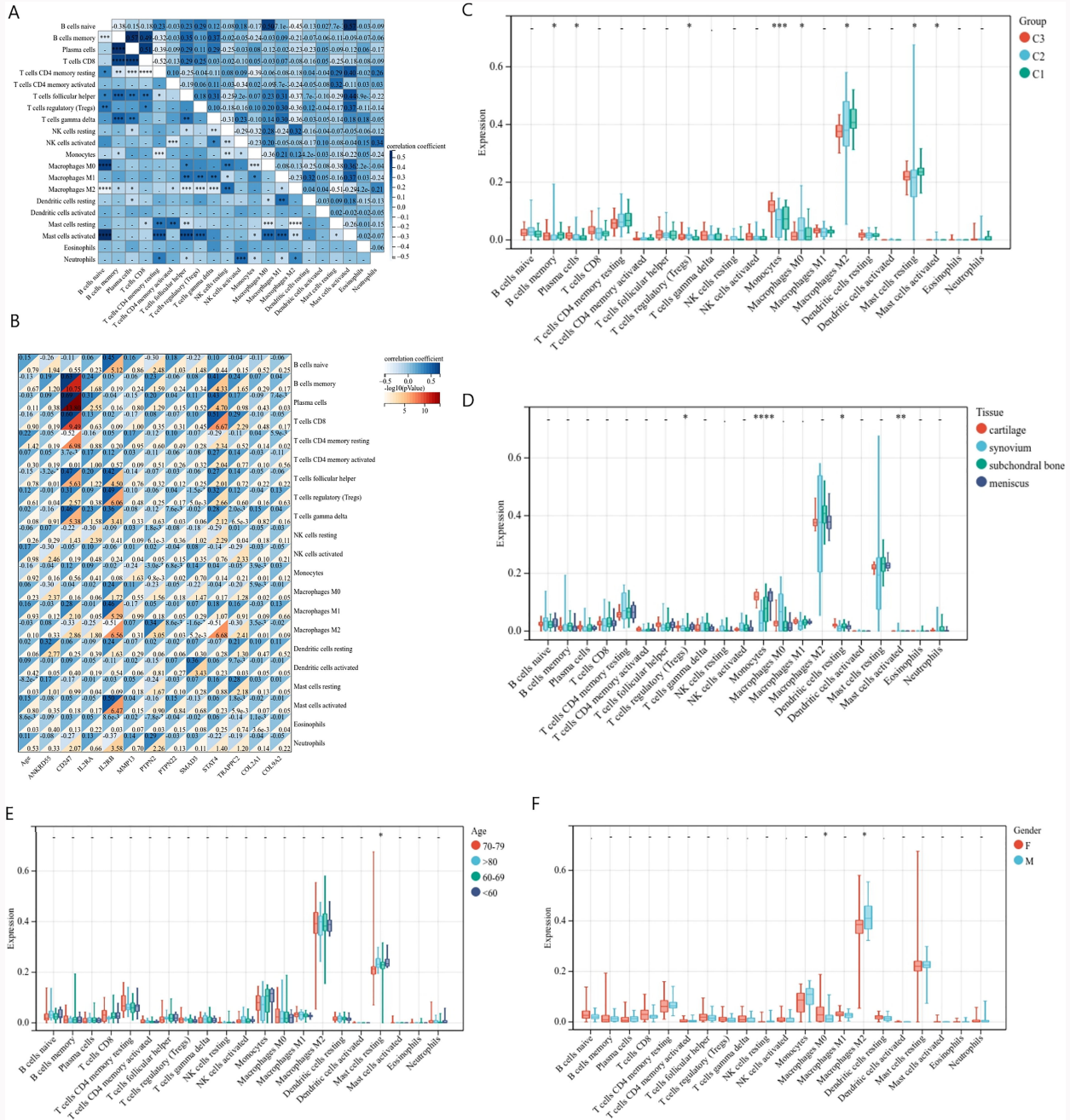
**Fig. 5** Function enrichment analysis of differentially expressed genes (DEGs). a) Gene Ontology (GO) enrichment analysis of C1 vs C2 DEGs. b) GO enrichment analysis of C1 vs C3 DEGs. c) GO enrichment analysis of C2 vs C3 DEGs. d) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of C1 vs C2 DEGs. e) KEGG pathway analysis of C1 vs C3 DEGs. f) KEGG pathway analysis of C2 vs C3 DEGs.

and TRAPPC2 was significantly up-regulated in females ( $p < 0.001$  paired  $t$ -test) (Figure 7e and 7f).

## Discussion

Previously, knee OA has been considered as a disease of articular cartilage, but with the deepening of research, knee OA is considered as a whole joint disease.<sup>22</sup> It includes the interaction between articular cartilage, synovium, subchondral bone, meniscus, tendon, and muscle. For example, cartilage debris caused by wear will stimulate synovium to

release inflammatory factors, which in turn will stimulate cartilage degeneration and apoptosis.<sup>23</sup> From the perspective of molecular biology, genes in cells have already begun to express abnormally in the asymptomatic and no-imaging stages of patients.<sup>24</sup> Based on microarray technology, even if some tissues have no significant changes, gene expression profiling can also be used to make an early diagnosis.<sup>25</sup> One study has even shown that knocking out specific genes in bone cells or using inhibitors can effectively salvage the structure of subchondral bone and alleviate cartilage

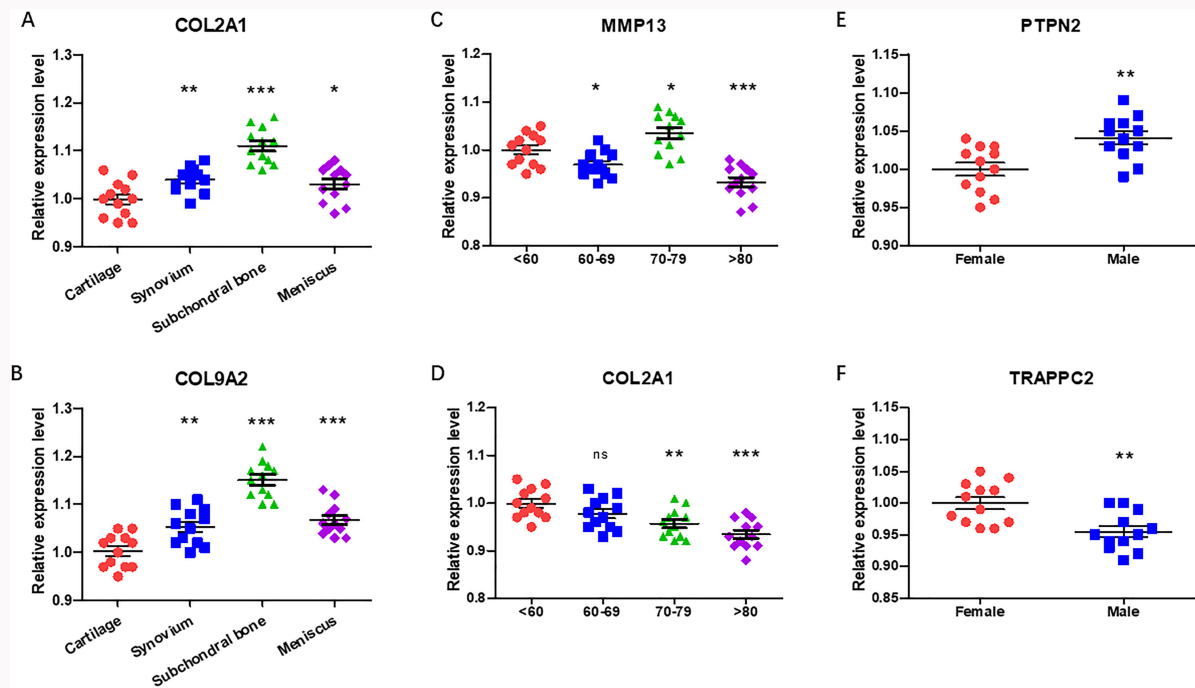


**Fig. 6** Analysis of immune cell infiltration. a) Correlation analysis of immune cells. b) Correlation between osteoarthritis (OA)-related genes and immune cells. c) Immune cell infiltration of different subtypes. d) Immune cell infiltration of different tissues. e) Immune cell infiltration of different ages. f) The difference in immune cell infiltration between male and female sex. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ , Mann-Whitney U test. NK, natural killer cells.

degeneration.<sup>26</sup> By combining gene expression profile data and clinical manifestations of different tissues, we finally divided OA into three subtypes: bone remodelling subtype (C1), immune metabolism subtype (C2), and cartilage degradation subtype (C3). This method not only establishes an effective classification model, but also an important means of finding gene markers and potential targets of drug therapy. It provides a specific direction for in-depth understanding of the pathogenesis and precision treatment of OA.

According to the enrichment analysis results of GSEA, C1 subtype function is enriched in skeletal joints, cartilage development, and glycosaminoglycan biosynthesis. This

subtype may represent the formation of osteophytes, and we describe it as a subtype of bone remodelling. C3 subtype is enriched in pyroptosis, cell death process, and vascular endothelial cell proliferation. It differs the most from the C1 subtype, and we classify it as a cartilage degeneration subtype. Both subtypes are caused by the imbalance between anabolism and catabolism.<sup>27</sup> Metabolic imbalance causes crosstalk between tissues, directly or indirectly affecting joint health and cartilage turnover, and altering the progression of OA.<sup>28</sup> A lot of evidence shows that in the early stage of OA, there is temporary bone loss in subchondral bone, and in the late stage, there is an increase in bone volume and



**Fig. 7**

Clinical features and osteoarthritis (OA)-related genes validated by quantitative real-time polymerase chain reaction (qRT-PCR). a) and b) qT-PCR of OA-related genes across different tissues. c) and d) qT-PCR of OA-related genes across different ages. e) and f) qPCR of OA-related genes across different sexes. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ , paired *t*-test. COL2A1, collagen type II alpha 1 chain; MMP13, matrix metalloproteinase 13; PTPN2, protein tyrosine phosphatase non-receptor type 2; TRAPPC2, trafficking protein particle complex subunit 2.

bone density in subchondral bone.<sup>29</sup> Therefore, we found that subchondral bone is mainly distributed in C1 and C3 subtypes. In addition, when OA occurs, the permeability of subchondral bone plate increases, vascular growth factor invades cartilage, and vascular proliferation occurs. Inflammatory factors can also enter cartilage through channels and interfere with proteoglycan metabolism of chondrocytes.<sup>30</sup> These previous research results are consistent with our conclusions, and also prove the effectiveness of our typing. We also found significant differences in the expression of CD247, MMP13, COL2A1, and COL9A2 in C1 and C3 subtypes. COL2A1 and MMP13 have a good differentiation effect on C1 and C3 subtypes, and the C1 vs C3 AUC values are both greater than 0.8. COL2A1 is expressed in normal cartilage, and its decrease is related to IL-1 $\beta$ -induced cartilage degeneration.<sup>31,32</sup> Deberg et al<sup>33</sup> found that COL2A1 was correlated with the severity of OA, and its one-year change was positively correlated with joint space stenosis. MMP13 acts as an extracellular matrix-degrading enzyme in OA, which can directly degrade type II collagen.<sup>34</sup> As one of the decomposition products of type II collagen, CTX-II can be detected in urine, blood, and joint fluid, and is expected to become an indicator for monitoring OA.<sup>35</sup> In summary, COL2A1 and MMP13 may be effective targets for the treatment of knee OA.<sup>36</sup>

It is worth noting that all synovial tissues are classified as C2 subtype. C2 subtype functions are enriched in various metabolic processes, and pathways are enriched in systemic lupus erythematosus and primary immunodeficiency. Based on the above results, we described C2 subtype as an immune metabolism subtype. Synovial inflammation is closely related to joint swelling, inflammatory pain, and other symptoms, representing the progress of OA.<sup>37</sup> However, in

the C2 subtype, the inflammation-related markers IL2RA and IL2RB did not show higher expression than other subtypes. Several treatment tests for low-grade inflammation also showed disappointing results.<sup>38</sup> Orange et al<sup>39</sup> also divided the synovial tissue of rheumatoid arthritis into a high-inflammation subtype characterized by extensive infiltration of leucocytes; a low inflammatory subtype characterized by enrichment in pathways including transforming growth factor  $\beta$ , glycoproteins, and neuronal genes; and a mixed subtype. We speculate that although the synovium of OA patients also exhibits histological changes similar to inflammatory arthritis, it is well known that OA is only a chronic low-grade inflammation.<sup>40</sup> Miller et al<sup>41</sup> found that OA inflammation is chronic, mainly mediated by the innate immune system, and that immune cells play an important role in the disease's development. In their analysis of immune cell infiltration, macrophages and mast cells were found to be the main immune cells involved in OA. Macrophage M0 and C2 activated mast cells were significantly higher than those of the other two subtypes. The role of macrophages in OA has always been a research hotspot: the infiltration and accumulation of macrophages in synovium are considered to be biomarkers of OA progress.<sup>42</sup> Targeting M2 macrophages and promoting their expression is an effective method for the treatment of inflammatory arthritis.<sup>43,44</sup> Tryptase, produced by mast cells, can induce chondrocyte inflammation and apoptosis.<sup>45</sup> Chen et al<sup>46</sup> also found that there are a high proportion of mast cells and macrophages, making them the main infiltrating cells in OA.

More interestingly, we found that the expression of COL2A1 gradually decreased with age, which may be related to the expression of MMP13, which behaves in the same



way. COL2A1 is only considered as a structural component of cartilage matrix. Recently, however, it was found to be an external signal molecule that can significantly inhibit chondrocyte hypertrophy.<sup>47</sup> The symptoms of OA can also be improved by stimulating the supplement of joint collagen. In addition, we also found that the expression of PTPN2 and TRAPPC2 may be related to the sex of OA patients. In synovial samples from patients with OA, PTPN2 can regulate IL-6 production, cell death, and autophagy.<sup>48</sup> In some immune diseases, the expression of PTPN2 in male patients is significantly increased.<sup>49</sup> Mutations in the TRAPPC2 gene were confirmed to be linked to X-linked recessive osteochondral dysplasia. The disease will cause joint pain, and degenerative osteoarthritis will occur early.<sup>50</sup> Yuan et al<sup>51</sup> also used RNA sequence datasets to perform unsupervised cluster analysis of gene expression profiles; however, unlike our study, they only classified the cartilage transcriptome, which is not as comprehensive as our experiment. They also combined clinical data, and the classification was similar to ours. A limitation of our study is that the tissue sample size is still too small, and further experiments are needed to verify this bioinformatics research.

In conclusion, through unsupervised clustering, we divided the OA gene expression profiles of different tissues into three subtypes: bone remodelling subtype (C1), immune metabolism subtype (C2), and cartilage degradation subtype (C3). The relationship between clinical characteristics of different subtypes and OA genes was studied, providing guidance for personalized treatment and more effective treatment targets for OA.

### Supplementary material

Research flowchart, protein-protein interaction network, gene set enrichment analysis, clinical features, and receiver operating characteristic curve.

### References

- Swain S, Sarmanova A, Mallen C, et al. Trends in incidence and prevalence of osteoarthritis in the United Kingdom: findings from the Clinical Practice Research Datalink (CPRD). *Osteoarthritis Cartilage*. 2020;28(6):792–801.
- Glyn-Jones S, Palmer AJR, Agricola R, et al. Osteoarthritis. *Lancet*. 2015;386(9991):376–387.
- Loeser RF, Collins JA, Diekman BO. Ageing and the pathogenesis of osteoarthritis. *Nat Rev Rheumatol*. 2016;12(7):412–420.
- van den Bosch MHJ, van Lent P, van der Kraan PM. Identifying effector molecules, cells, and cytokines of innate immunity in OA. *Osteoarthritis Cartilage*. 2020;28(5):532–543.
- He CP, Jiang XC, Chen C, et al. The function of lncRNAs in the pathogenesis of osteoarthritis. *Bone Joint Res*. 2021;10(2):122–133.
- Scanzello CR, Goldring SR. The role of synovitis in osteoarthritis pathogenesis. *Bone*. 2012;51(2):249–257.
- Goldring MB, Goldring SR. Articular cartilage and subchondral bone in the pathogenesis of osteoarthritis. *Ann NY Acad Sci*. 2010;1192:230–237.
- Junker S, Krumbholz G, Frommer KW, et al. Differentiation of osteophyte types in osteoarthritis - proposal of a histological classification. *Joint Bone Spine*. 2016;83(1):63–67.
- Menkes CJ. Radiographic criteria for classification of osteoarthritis. *J Rheumatol Suppl*. 1991;27:13–15.
- Abdelaziz H, Balde OM, Citak M, Gehrke T, Magan A, Haasper C. Kellgren-Lawrence scoring system underestimates cartilage damage when indicating TKA: preoperative radiograph versus intraoperative photograph. *Arch Orthop Trauma Surg*. 2019;139(9):1287–1292.
- Appleton CTG, Pitelka V, Henry J, Beier F. Global analyses of gene expression in early experimental osteoarthritis. *Arthritis Rheum*. 2007;56(6):1854–1868.
- Li B, Ding T, Chen H, et al. *CircStrn3* targeting *microRNA-9-5p* is involved in the regulation of cartilage degeneration and subchondral bone remodelling in osteoarthritis. *Bone Joint Res*. 2023;12(1):33–45.
- Kraus VB, Collins JE, Hargrove D, et al. Predictive validity of biochemical biomarkers in knee osteoarthritis: data from the FNIH OA Biomarkers Consortium. *Ann Rheum Dis*. 2017;76(1):186–195.
- Fleischmann RM, Bliddal H, Blanco FJ, et al. A phase II trial of Lutikizumab, an anti-interleukin-1 $\alpha/\beta$  dual variable domain immunoglobulin, in knee osteoarthritis patients with synovitis. *Arthritis Rheumatol*. 2019;71(7):1056–1069.
- Plsikova Matejova J, Spakova T, Harvanova D, et al. A preliminary study of combined detection of COMP, TIMP-1, and MMP-3 in synovial fluid: Potential indicators of osteoarthritis progression. *Cartilage*. 2021;13(2\_suppl):1421S–1430S.
- Wang M, Tan G, Jiang H, et al. Molecular crosstalk between articular cartilage, meniscus, synovium, and subchondral bone in osteoarthritis. *Bone Joint Res*. 2022;11(12):862–872.
- Ji Q, Zheng Y, Zhang G, et al. Single-cell RNA-seq analysis reveals the progression of human osteoarthritis. *Ann Rheum Dis*. 2019;78(1):100–110.
- Cao X, Cui Z, Ding Z, et al. An osteoarthritis subtype characterized by synovial lipid metabolism disorder and fibroblast-like synoviocyte dysfunction. *J Orthop Translat*. 2022;33:142–152.
- No authors listed. Gene Expression Omnibus. <https://www.ncbi.nlm.nih.gov/geo> (date last accessed 18 October 2023).
- No authors listed. <https://www.gsea-msigdb.org/gsea/index.jsp> (date last accessed 18 October 2023).
- Wang J, Xue Y, Zhou L. Comparison of immune cells and diagnostic markers between spondyloarthritis and rheumatoid arthritis by bioinformatics analysis. *J Transl Med*. 2022;20(1):196.
- Loeser RF, Goldring SR, Scanzello CR, Goldring MB. Osteoarthritis: A disease of the joint as an organ. *Arthritis Rheum*. 2012;64(6):1697–1707.
- Roebuck MM, Jamal J, Lane B, et al. Cartilage debris and osteoarthritis risk factors influence gene expression in the synovium in end stage osteoarthritis. *Knee*. 2022;37:47–59.
- Lv Z, Yang YX, Li J, et al. Molecular classification of knee osteoarthritis. *Front Cell Dev Biol*. 2021;9:725568.
- Fisch KM, Gamini R, Alvarez-Garcia O, et al. Identification of transcription factors responsible for dysregulated networks in human osteoarthritis cartilage by global gene expression analysis. *Osteoarthritis Cartilage*. 2018;26(11):1531–1538.
- Tu M, Yang M, Yu N, et al. Inhibition of cyclooxygenase-2 activity in subchondral bone modifies a subtype of osteoarthritis. *Bone Res*. 2019;7:29.
- Karsdal MA, Bay-Jensen AC, Lories RJ, et al. The coupling of bone and cartilage turnover in osteoarthritis: opportunities for bone antiresorptives and anabolics as potential treatments? *Ann Rheum Dis*. 2014;73(2):336–348.
- Bay-Jensen AC, Slagboom E, Chen-An P, et al. Role of hormones in cartilage and joint metabolism: understanding an unhealthy metabolic phenotype in osteoarthritis. *Menopause*. 2013;20(5):578–586.
- Mahjoub M, Berenbaum F, Houard X. Why subchondral bone in osteoarthritis? The importance of the cartilage bone interface in osteoarthritis. *Osteoporos Int*. 2012;23 Suppl 8:S841–6.
- Walsh DA, McWilliams DF, Turley MJ, et al. Angiogenesis and nerve growth factor at the osteochondral junction in rheumatoid arthritis and osteoarthritis. *Rheumatology (Oxford)*. 2010;49(10):1852–1861.
- Hick A-C, Fonck M, Costes B, et al. Serum levels of Coll2-1, a specific biomarker of cartilage degradation, are not affected by sampling conditions, circadian rhythm, and seasonality. *Cartilage*. 2021;13(1\_suppl):540S–549S.
- Cheng F, Hu H, Sun K, Yan F, Geng Y. miR-455-3p enhances chondrocytes apoptosis and inflammation by targeting COL2A1 in the *in vitro* osteoarthritis model. *Biosci Biotechnol Biochem*. 2020;84(4):695–702.
- Deberg MA, Labasse AH, Collette J, Seidel R, Reginster JY, Henrotin YE. One-year increase of Coll 2-1, a new marker of type II collagen degradation, in urine is highly predictive of radiological OA progression. *Osteoarthritis Cartilage*. 2005;13(12):1059–1065.
- Salerno A, Brady K, Rikkers M, et al. MMP13 and TIMP1 are functional markers for two different potential modes of action by mesenchymal

stem/stromal cells when treating osteoarthritis. *Stem Cells*. 2020;38(11):1438–1453.

35. Henrotin Y, Chevalier X, Deberg M, et al. Early decrease of serum biomarkers of type II collagen degradation (Coll2-1) and joint inflammation (Coll2-1 NO<sub>2</sub>) by hyaluronic acid intra-articular injections in patients with knee osteoarthritis: A research study part of the Biovisco study. *J Orthop Res*. 2013;31(6):901–907.
36. Wang M, Sampson ER, Jin H, et al. MMP13 is a critical target gene during the progression of osteoarthritis. *Arthritis Res Ther*. 2013;15(1):R5.
37. Sellam J, Berenbaum F. The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis. *Nat Rev Rheumatol*. 2010;6(11):625–635.
38. Robinson WH, Lepus CM, Wang Q, et al. Low-grade inflammation as a key mediator of the pathogenesis of osteoarthritis. *Nat Rev Rheumatol*. 2016;12(10):580–592.
39. Orange DE, Agius P, DiCarlo EF, et al. Identification of three rheumatoid arthritis disease subtypes by machine learning integration of synovial histologic features and RNA sequencing data. *Arthritis Rheumatol*. 2018;70(5):690–701.
40. Smith MD, O'Donnell J, Highton J, Palmer DG, Rozenbils M, Roberts-Thomson PJ. Immunohistochemical analysis of synovial membranes from inflammatory and non-inflammatory arthritides: scarcity of CD5 positive B cells and IL2 receptor bearing T cells. *Pathology*. 1992;24(1):19–26.
41. Miller RJ, Malfait AM, Miller RE. The innate immune response as a mediator of osteoarthritis pain. *Osteoarthritis Cartilage*. 2020;28(5):562–571.
42. Bondeson J, Blom AB, Wainwright S, Hughes C, Caterson B, van den Berg WB. The role of synovial macrophages and macrophage-produced mediators in driving inflammatory and destructive responses in osteoarthritis. *Arthritis Rheum*. 2010;62(3):647–657.
43. Ma Y, Yang H, Zong X, et al. Artificial M2 macrophages for disease-modifying osteoarthritis therapeutics. *Biomaterials*. 2021;274:120865.
44. Mo H, Wang Z, He Z, et al. Decreased Peli1 expression attenuates osteoarthritis by protecting chondrocytes and inhibiting M1-polarization of macrophages. *Bone Joint Res*. 2023;12(2):121–132.
45. Kulkarni P, Harsulkar A, Mårtson A-G, Suutre S, Mårtson A, Koks S. Mast cells differentiated in synovial fluid and resident in osteophytes exalt the inflammatory pathology of osteoarthritis. *Int J Mol Sci*. 2022;23(1):541.
46. Chen Z, Ma Y, Li X, Deng Z, Zheng M, Zheng Q. The immune cell landscape in different anatomical structures of knee in osteoarthritis: A gene expression-based study. *Biomed Res Int*. 2020;2020:9647072.
47. Sun X, Huang H, Pan X, et al. EGR1 promotes the cartilage degeneration and hypertrophy by activating the Krüppel-like factor 5 and  $\beta$ -catenin signaling. *Biochim Biophys Acta Mol Basis Dis*. 2019;1865(9):2490–2503.
48. Aradi B, Kato M, Filkova M, et al. Protein tyrosine phosphatase nonreceptor type 2: an important regulator of Interleukin-6 production in rheumatoid arthritis synovial fibroblasts. *Arthritis Rheumatol*. 2015;67(10):2624–2633.
49. Zhang Q, Li H, Hou S, et al. Association of genetic variations in PTPN2 and CD122 with ocular Behcet's disease. *Br J Ophthalmol*. 2018;102(7):996–1002.
50. Zhang C, Du C, Ye J, et al. A novel deletion variant in TRAPPC2 causes spondyloepiphyseal dysplasia tarda in a five-generation Chinese family. *BMC Med Genet*. 2020;21(1):117.
51. Yuan C, Pan Z, Zhao K, et al. Classification of four distinct osteoarthritis subtypes with a knee joint tissue transcriptome atlas. *Bone Res*. 2020;8(1):38.

### Author information

Y. Xue, MMed, Orthopaedic Surgeon, Department of Orthopaedic, Wuxi Ninth People's Hospital of Soochow University, Wuxi, China.

L. Zhou, MD, Orthopaedic Surgeon, Department of Orthopaedic, Lianshui County People's Hospital, Huai'an, China.

J. Wang, MMed, Doctoral Student, Department of Orthopaedic, Zhongshan Hospital, Fudan University, Shanghai, China.

### Author contributions

Y. Xue: Investigation, Writing – original draft.

L. Zhou: Writing – review & editing.

J. Wang: Conceptualization, Investigation, Writing – original draft.

Y. Xue and L. Zhou contributed equally to this work.

### Funding statement

The authors disclose receipt of the following financial or material support for the research, authorship, and/or publication of this article: this work was supported by Wuxi Municipal Health and Family Planning Commission Research Youth Project (Q202130) and Wuxi Top medical expert team of "Taihu Talent Progra".

### ICMJE COI statement

The authors disclose receipt of the following financial or material support for the research, authorship, and/or publication of this article: this work was supported by Wuxi Municipal Health and Family Planning Commission Research Youth Project (Q202130) and Wuxi Top medical expert team of "Taihu Talent Progra".

### Data sharing

The datasets analysed during the current study are available in the GEO repository with accession number (GSE114007, GSE55457, GSE12021, GSE51588, and GSE98918).

### Acknowledgements

We would like to thank Bo Guo for communicating data and information.

### Ethical review statement

This study was approved by the research ethics committee of Lianshui County People's Hospital. Informed consent was obtained from all individual participants included in the study.

### Open access funding

The authors confirm that the open access fee for this study was self-funded.

© 2023 Wang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (CC BY-NC-ND 4.0) licence, which permits the copying and redistribution of the work only, and provided the original author and source are credited. See <https://creativecommons.org/licenses/by-nc-nd/4.0/>