Heliyon 10 (2024) e30335

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Multi-omics characterization of macrophage polarization-related features in osteoarthritis based on a machine learning computational framework

Ping Hu [a,b,1], Beining Li [a,b,1], Zhenyu Yin [a,b,1], Peng Peng [a,b], Jiangang Cao [c], Wanyu Xie [a,b], Liang Liu [d], Fujiang Cao [a,b,**], Bin Zhang [a,b,*]

[a] Department of Othopaedics, Tianjin Medical University General Hospital, 154 Anshan Road, Heping District, Tianjin, 300052, China
[b] International Science and Technology Cooperation Base of Spinal Cord Injury, Tianjin Key Laboratory of Spine and Spinal Cord Injury, Department of Orthopedics, Tianjin Medical University General Hospital, Tianjin, China
[c] Department of Sports Injury and Arthroscopy, Tianjin Hospital of Tianjin University, China
[d] Orthopaedic Center of Beijing Luhe Hospital, Capital Medical University, China

ARTICLE INFO

Keywords:
Osteoarthritis
Macrophage polarization
Bioinformatics
Machine learning
WGCNA

ABSTRACT

Background: OA imposes a heavy burden on patients and society in that its mechanism is still unclear, and there is a lack of effective targeted therapy other than surgery.
Methods: The osteoarthritis dataset GSE55235 was downloaded from the GEO database and analyzed for differential genes by limma package, followed by analysis of immune-related modules by xcell immune infiltration combined with the WGCNA method, and macrophage polarization-related genes were downloaded according to the Genecard database, and VennDiagram was used to determine their intersection. These genes were also subjected to gene ontology (GO), disease ontology (DO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analyses. Using machine learning, the key osteoarthritis genes were finally screened. Using single gene GSEA and GSVA, we examined the significance of these key gene functions in immune cell and macrophage pathways. Next, we confirmed the correctness of the hub gene expression profile using the GSE55457 dataset and the ROC curve. Finally, we projected TF, miRNA, and possible therapeutic drugs using the miRNet, TargetScanHuman, ENCOR, and NetworkAnalyst databases, as well as Enrichr.
Results: VennDiagram obtained 71 crossover genes for DEGs, WGCNA-immune modules, and Genecards; functional enrichment demonstrated NF-κB, IL-17 signaling pathway play an important role in osteoarthritis-macrophage polarization genes; machine learning finally identified CSF1R, CX3CR1, CEBPB, and TLR7 as hub genes; GSVA analysis showed that CSF1R, CEBPB play essential roles in immune infiltration and macrophage pathway; validation dataset GSE55457 analyzed hub genes were statistically different between osteoarthritis and healthy controls, and the AUC values of ROC for CSF1R, CX3CR1, CEBPB and TLR7 were more outstanding than 0.65.

* Corresponding author. Department of Othopaedics, Tianjin Medical University General Hospital, 154 Anshan Road, Heping District, Tianjin, 300052, China.
** Corresponding author. Department of Othopaedics, Tianjin Medical University General Hospital, 154 Anshan Road, Heping District, Tianjin, 300052, China.
E-mail addresses: fcao@tmu.edu.cn (F. Cao), philexchang@126.com (B. Zhang).
[1] Ping Hu, Beining Li, Zhenyu Yin contributed equally to this work and share first authorship.

https://doi.org/10.1016/j.heliyon.2024.e30335
Received 17 February 2024; Received in revised form 16 April 2024; Accepted 24 April 2024
Available online 27 April 2024

Conclusions: CSF1R, CEBPB, CX3CR1, and TLR7 are potential diagnostic biomarkers for osteoarthritis, and CSF1R and CEBPB play an important role in regulating macrophage polarization in osteoarthritis progression and are expected to be new drug targets.

1. Introduction

Osteoarthritis (OA), one of the most prevalent diseases affecting joint function, affects almost more than 240 million people worldwide [1–3], causes substantial economic losses and brings great inconvenience to the lives of patients [4,5].Most patients suffered from OA (59–87 %) have at least one other chronic disease, especially cardiometabolic conditions [6].According to a clinical trial, inhibitors of cathepsin K, inhibitors of Wnt, and anabolic growth factors may arrest structural progression, and nerve growth factor inhibitors may release OA pain. The treatment mentioned above methods are helpful for the symptoms and delaying the progression of OA. However, the research focusing on a single pathogenic gene has not fully elucidated its molecular changes and internal mechanism, in other words, was easy to neglect the bigger picture of OA itself. Focusing on one crucial tissue, creating a precise image of the molecular alterations and internal mechanism of OA, identifying the DEGs between OA patients and healthy individuals, and studying these molecular changes provide a unique perspective on OA [7].

As research progressed, OA was recognized as a disorder characterized by sustained low-grade inflammation [8].Macrophages from synovial act as immunocyte, engage in and be indispensable for the structural development of OA [9,10].Signaling pathways like mTOR, NF-κB, JNK, PI3K/Akt regulated activated macrophages and polarized macrophages into subtypes M1 or M2 in OA synovial tissues. In addition, synovitis also plays a crucial role in the progression of OA, where M1 macrophage-driven synovitis exacerbates the pathological process. The main pathological features of synovitis are abnormal synovial growth, inflammatory response, and vascular proliferation. Synovial inflammation increases the risk of OA progression and promotes the recruitment of monocytes, lymphocytes, and other white blood cells. Conversely, macrophages are critical mediators of synovial inflammatory activity and pathological cartilage and bone responses. Soluble mediators are working as highly accurate predictors of knee OA progression throughout the whole development, in which SF elastase and TGFβ1 are most representative [11–13].

The activated macrophages mainly show two subtypes, called M1 (Classical Activation) and M2 (Alternative Activation) [14]. For the synovium of OA sufferers, M1 macrophages account for the vast majority of macrophages. In contrast, the proportion M2 macrophages occupied is relatively tiny. When synovial macrophages were mainly polarized to M1 type, the severity of OA was aggravated. Researchers tried boosting the polarization of M2 macrophages or suppressing the polarization of M1 macrophages and gained similar results that osteoarthritis's progression was significantly slowing down [15].Above all, such consequences hint that inhibition of the polarization from macrophages to M1 subtype and stimulation of the transformation from M1 subtype to M2 were potentially effective strategies for the treatment of OA [16].

It can be seen that the biological treatment of immune infiltration or interfering with the polarization of macrophages are hot spots in the treatment of OA. However, scientific research has entered the era of omics. To analyze the role immunocytes played in the occurrence and progression of osteoarthritis, using bioinformatics tools, we screen potential biomarkers involved in the immune process and find drugs that intervene in macrophage polarization biomarkers, thereby delaying the progression of osteoarthritis, which can be applied to decision-making in future clinical practice. In this study, after obtaining DEGs and analyzing the macrophage polarization data set of Genecards, we used WGCAN algorithm and three different machine learning algorithms to screen further potential biomarkers perform expression level verification and functional analysis on them, including ROC diagnostic analysis, single gene GSVA analysis of GSEA [17,18], immune infiltration and macrophage polarization datasets. These data were used to study the function of these genes and the mechanism involved in macrophage polarization, and finally, identify the ideal drug target by using the drug database.

2. Materials and methods

2.1. Microarray data acquirement and data pretreatment

The microarray data of gene expression was downloaded from the database: Gene Expression Omnibus (GEO) [19]. For the dataset GSE55235, we detail the characteristics of 20 synovial tissues, comprising 10 normal and 10 osteoarthritis (OA) samples. We include patient demographics like age and gender, specify the disease stage, and outline sample collection methods. The exclusion of rheumatoid arthritis samples enhances the dataset's focus on OA, increasing its specificity for studying this condition. Similarly, GSE55457 follows the same protocol to ensure reproducibility. For GSE152805, we describe the demographics of three OA patients and explain the single-cell RNA sequencing methodology, including cell sorting techniques and sequencing depth, used in the analysis.

2.2. Identification of DEGs

Firstly, we treated the expression of GSE55235 more than 50 with logarithm before correcting and normalizing. The parameters | Log2fold change| >1 and $p < 0.05$ were regarded as the screening criterion for differentially expressed genes (DEGs) (Analyzed by R package "Limma" [20], genes satisfying $p < 0.05$ and $|log2FC|>1$ are considered as DEGs) As for the microarray data (GSE55457) for validating. After standardization, the expression of the hub gene was extracted for the final test. The R packages "ggplot2" and

"pheatmap" generated volcano maps and heat maps" to illustrate DEG results.

2.3. Immune cell infiltration analysis

We performed immune infiltration analysis using the R package "IOBR", which provides bulk analysis of the features and the correlation with clinical phenotypes, lncRNA analysis, genomic features, and features from scRNA-seq data in different cancer settings [21]. IOBR supports eight common immune infiltration algorithms, and we used the "xCell" algorithm for immune infiltration analysis. xCell, a novel method based on genetic features, was used to infer Sixty-four immune cell types [22].

2.4. Screening of target modules and genes based on WGCNA

In this research [23], WGCNA was used to identify the genes and modules most strongly associated with osteoarthritis immunity. Firstly, we removed the genes with a variance of 0, calculated each gene's mean standard deviation (MSD) separately, and deleted the genes with a proportion of MSD greater than 10 %. The samples were hierarchically grouped further to identify outliers and exclude anomalous data. The pearson correlation matrix and average correlation were calculated to perform all pairings of genes. A weighted adjacency matrix by the power function A mn = |C mn | ^ beta construct. Beta, a soft threshold value specified by the pickSoft function, may stress the strong correlation between genes. Topological overlap matrix (TOM) may quantify the network connection of a gene and determine the related difference degree (1-TOM). To categorize the genes with comparable expression characteristics into gene modules, we performed average relation hierarchical clustering by dissimilarity measurement method based on TOM, and the minimum size of the genome is 150. Merging modules, the threshold was set to 0.25. The R program "WGCNA" was then adopted to estimate the link between modules and differentially infiltrating immunocytes. Modules with high correlation coefficients were regarded as candidates to make the association with discrepantly infiltrating immunocytes and were chosen to be used in subsequent analyses.

2.5. Function and pathway enrichment analysis

The macrophage polarization-related gene sets were obtained from Gnencards [24] website (https://www.genecards.org/), and the intersection with DEGS and WGCNA module genes was taken to build a Venn diagram from the "Venn Diagram "R package. GO [25], KEGG [26], and DO [27] analyses were performed for the DEGs. The filtering criterion of enrichment analysis was identified as p-value <0.05, and the enriched consequences were significant when both were met. The R package 'org.Hs.eg.db' and the 'clusterProfiler' [28] were individually used for gene ID conversion and enrichment analysis results. The 'enrichplot' and 'ggplot2' packages were used to visualize results and bubble and bar mapping. Finally, all genes were analyzed for GSEA enrichment. For visualization, we saved and plotted the top five enrichment pathways.

2.6. Establishment of protein-protein interaction (PPI) networks

To reveal the protein interactions among genes, a PPI [29] net was constructed by STRING [30] (Search Tool for the Retrieval of Interacting Genes https://string-db.org/). We visualized such a network with a confidence score of 0.4 by Cytoscape [31].

2.7. Filtering and identification of potential biomarkers based on machine learning algorithms

We first used the R package "e1071" and the SVM-RFE [32] algorithm to screen for potential genes. Support Vector Machine (SVM) is a supervised learning model that accurately classifies data points by maximizing the margin between two hyperplanes. We used a linear kernel function and optimized the parameters cost (C) and gamma (γ) through 10-fold cross-validation, where C ranged from 2^(−5:5) and γ ranged from 2^(−5:5). The parameter combination with the highest average accuracy was selected to obtain the optimal SVM model. Next, we further screened potential genes using the LASSO logistic regression algorithm and the R package "glmnet". LASSO [33] performs gene selection and classification through regression analysis, achieving sparsity via L1 regularization to select the most relevant features. We used 10-fold cross-validation to determine the regularization parameter λ, ensuring that the bias term satisfies the minimum criterion. The optimal λ value was obtained using the "lambda.min" parameter of the cv. glmnet function. The screened differentially expressed genes (DEGs) were analyzed using the Random Forest function. Random Forest is an ensemble learning method that improves classification performance by constructing multiple decision trees and combining their predictions. We used 500 decision trees, with the number of features for each tree being the square root of the total number of features. The Gini index was used as the purity measure to select the best split. Feature importance was evaluated by calculating the Mean Decrease Gini for each feature across all trees.

2.8. Biological function and regulatory mechanism of hub gene

GSEA is used for identifying the potential functions of hub genes. We downloaded the external reference gene subset from the MSigDB [34] (Molecular Signature Database). The criterion for significant enrichment was $p < 0.05$. We downloaded C5: ontology gene sets from MSigDB and fifteen 25 macrophage-related gene sets were extracted. Also, the "GSVA" [35] R package was used to analyze the correlation between hub genes and macrophage regulation and presented in the form of heat maps drawn by R packages

"ggplot2" and "ggcor".

2.9. Establishment of TF-hub-miRNA interactions network and potential therapeutic drugs prediction

The target miRNAs of hub genes were obtained from miRNet (https://www.mirnet.ca/), TargetScanHuman (https://www.targetscan.org/vert_80/), and ENCORI (https://starbase.sysu.edu.cn/), and the NetworkAnalyst [36] (https://www.networkanalyst.ca/) was used to predict the upstream TF (transcription factor). Finally, we constructed the TF-hub gene-miRNA Sankey map by R packages "ggplot2" and "ggalluvial". Subsequently, 5 potential biomarker-related drug prediction sites from DSigDB (the drug signature database) were searched in Enrichr, and Cytoscape constructed the hub gene-predicted drug interaction network.

2.10. Single cell sequencing analysis

We used the Seurat package (v.4.3.0) in R (R4.2.1) to process and visualize the scRNA-seq data. In the initial quality control step, we filtered out cells with less than 200 or more than 10,000 detected genes and genes detected in fewer than three cells. Then we

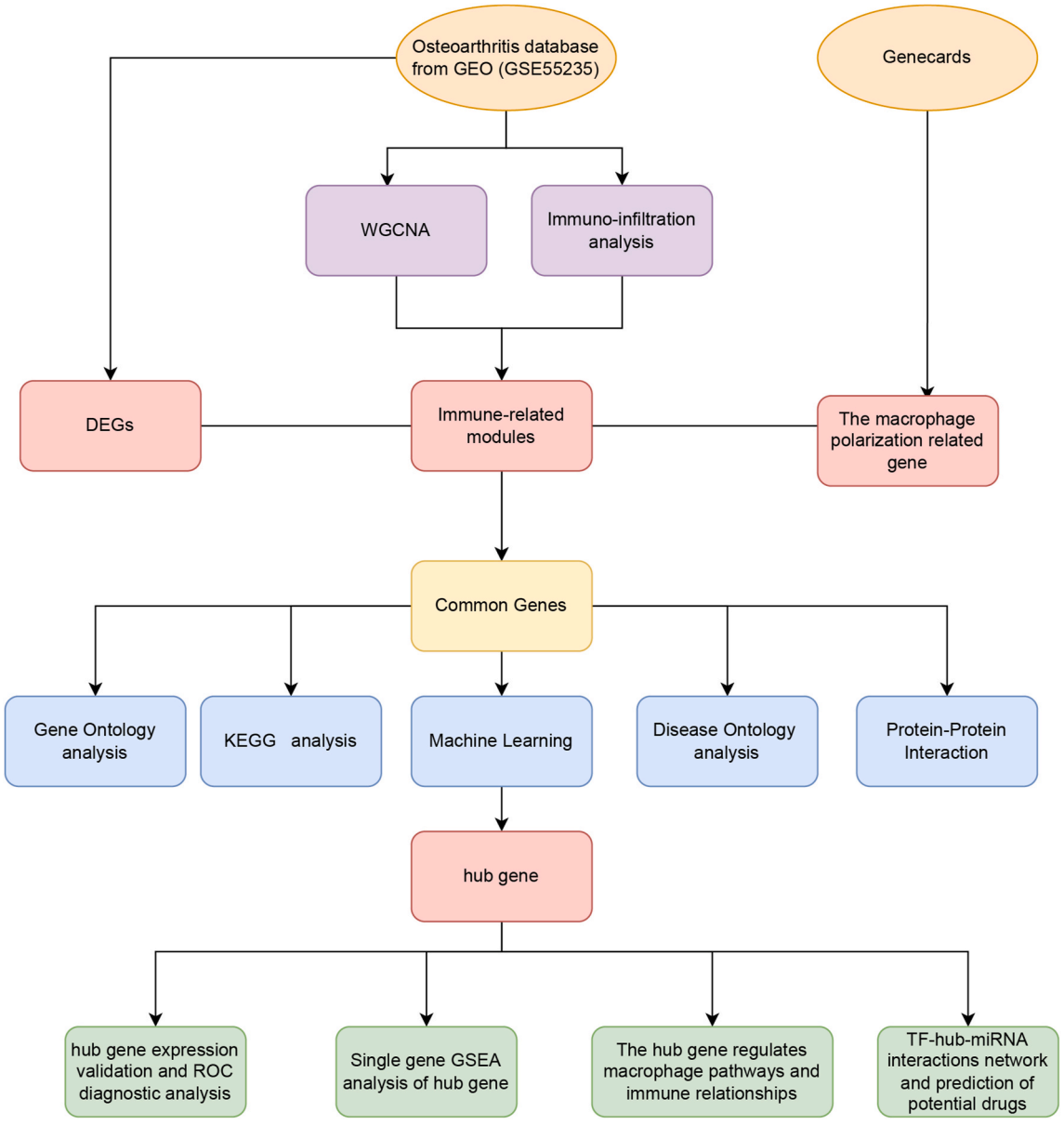

Fig. 1. Flowchart for research.

calculated the percentage of mitochondrial genes and hemoglobin genes for each cell as quality control indicators. We used violin plots to show the distribution of these indicators, as well as the number of detected genes and the total counts per cell. We removed cells with less than 200 or more than 10,000 detected genes, less than 100 or more than 30,000 total counts, more than 20 % mitochondrial genes, or more than 3 % hemoglobin genes. We normalized the data using "LogNormalize" and SCTransform and removed noise from mitochondrial genes. We scaled the data for downstream analysis such as PCA dimensionality reduction, setting the scale factor to 10,000. When filtering variable genes, we specified the top 2000 highly variable genes for downstream analysis. We used the FindClusters function to cluster the data. We used uniform manifold approximation and projection (UMAP) dimensionality reduction on the first 30 principal components to project the entire dataset into a two-dimensional space. For each cell cluster, we assigned cell type labels using statistical enrichment of marker gene sets and manual evaluation of gene expression of genes used by Ching-Heng Chou.

2.11. Statistical analysis

In this investigation, we compared gene expression in OA and healthy samples in the analyzed dataset GSE55235 and validation set GSE55457 using t-test, and validated its accuracy by performing Receiver Operating Characteristic (ROC) curve analysis of the screened hub genes with the R package "pROC" [37](21414208) and obtaining the corresponding (AUC) values. In addition, the expression levels of biomarkers were examined and box plots were created, and statistically significant differences were considered if *p < 0.05, **p < 0.01, ***p < 0.001.

3. Results

3.1. The identification of DEGs

Figs. 1 and 2 were the expression of each sample before and after correction. Using the R package "Limma" to perform variance analysis of the data sets. The criterion of DEGs were P < 0.05 and | log2FC | > 1. According to our standard, 1020 DEGs were discovered, in which the expression of 549 genes were up regulated and the left were down. The abscissa of Fig. 2C was log2FoldChange with the ordination of -log 10 (p-value). Red, blue, and gray nodes represent up-regulated DEGs, downregulated DEGs, and genes that are not notably differently expressed individually. Among them, genes with absolute value of log2FoldChange greater than 2 and p < 0.01 were annotated with gene names. As shown in Fig. 2 (D), We took | log2FC | of DEGs sorting, obtaining a significant

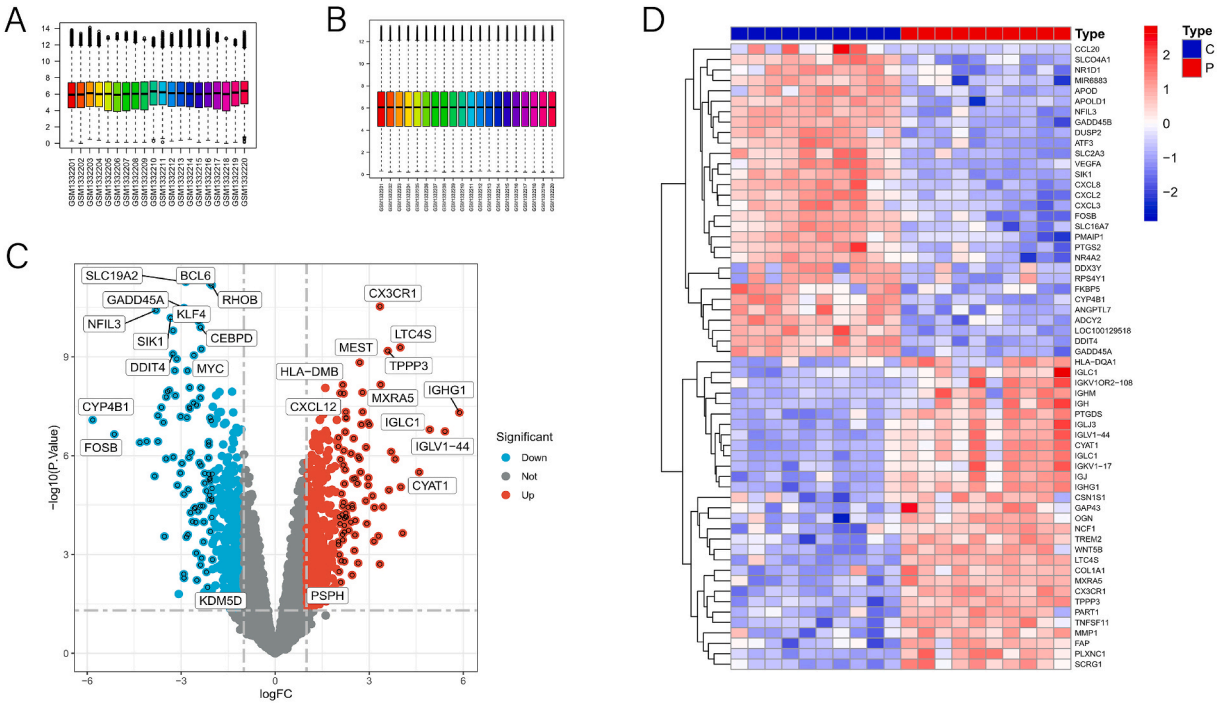

Fig. 2. Identification of DEGs in synovial membranes of patients with osteoarthritis. (A): Expression levels before correction; (B): Expression levels before correction; (C): Volcanic map of DEGs associated with OA. The abscissa is log2FoldChange and the ordinate is -log 10 (P value).; Red nodes, blue nodes and gray nodes represent up-regulated DEGs, downregulated DEGs, and genes that are not notably differentially expressed,individually; (D): Heat map of osteoarthritis-related DEG expression levels: Blue bars：normal control samples; Red bars：disease samples, red denotes high gene expression, while blue shows low gene expression.

increase 30 and 30 significantly lower genes to draw heat maps. We utilized a blue bar and red bar to separate the healthy group from the osteoarthritis group and used red representing high level of gene expression and blue representing low.

3.2. Immuno-infiltrating cell analysis

Fig. 3A shows the enriched abundance of immune infiltrating cells in OA and control groups, with the highest percentage of macrophages. Fig. 3b shows the correlation heat map between individual immune cells in the immune infiltrate, where red indicates positive correlation and blue indicates negative correlation. According to Fig. 3C, significant differences exist in immune cells between osteoarthritis and control samples ($P < 0.05$), in which B-cells, Basophils, $CD4^+$ naive T cells, $CD4^+$ Tcm, $CD4^+$ Tem, $CD8^+$ naive T-cells, $CD8^+$ T-cells, $CD8^+$ Tcm, $CD8^+$ Tem, cDC, Class-switched memory B-cells, CLP, CMP, Macrophages, Macrophages M1, Macrophages M2, Mast cells, MPP, pDC and Tgd cells are of remarkable. It is interesting to note that the different polarization states of macrophages are significantly different.

3.3. Establishment of weighted co-expression networks and identification of immune-related modules

To identify the genes that are most linked with immunocytes in osteoarthritis, WGCNA was used. As shown in Fig. 4A, we clustered all the samples and correlated them with clinical traits. There were no outliers and most samples with the same clinical traits were clustered together. Notably, the control sample, GSM1332210, was clustered in the osteoarthritis group. Based on unsigned and average connectivity , $\beta = 7$ (scale free $R2 = 0.85$) was used as the optimal "soft" threshold (Fig. 4 B). In addition, we set 150 as the minimum number of genes in the module, combined module threshold of 0.25. Finally, we obtained 14 co-expression modules (Fig. 4C). Notably, gray modules were the gene sets that cannot be assigned to any module. According to correlation analysis of modules and traits (infiltrating immunocytes), the blue module was highly positive in correlation with Macrophages, Macrophages_M1, and Macrophages_M2 (Cor = 0.7, P = 6e-04). While the Pink was negatively correlated with Macrophages,

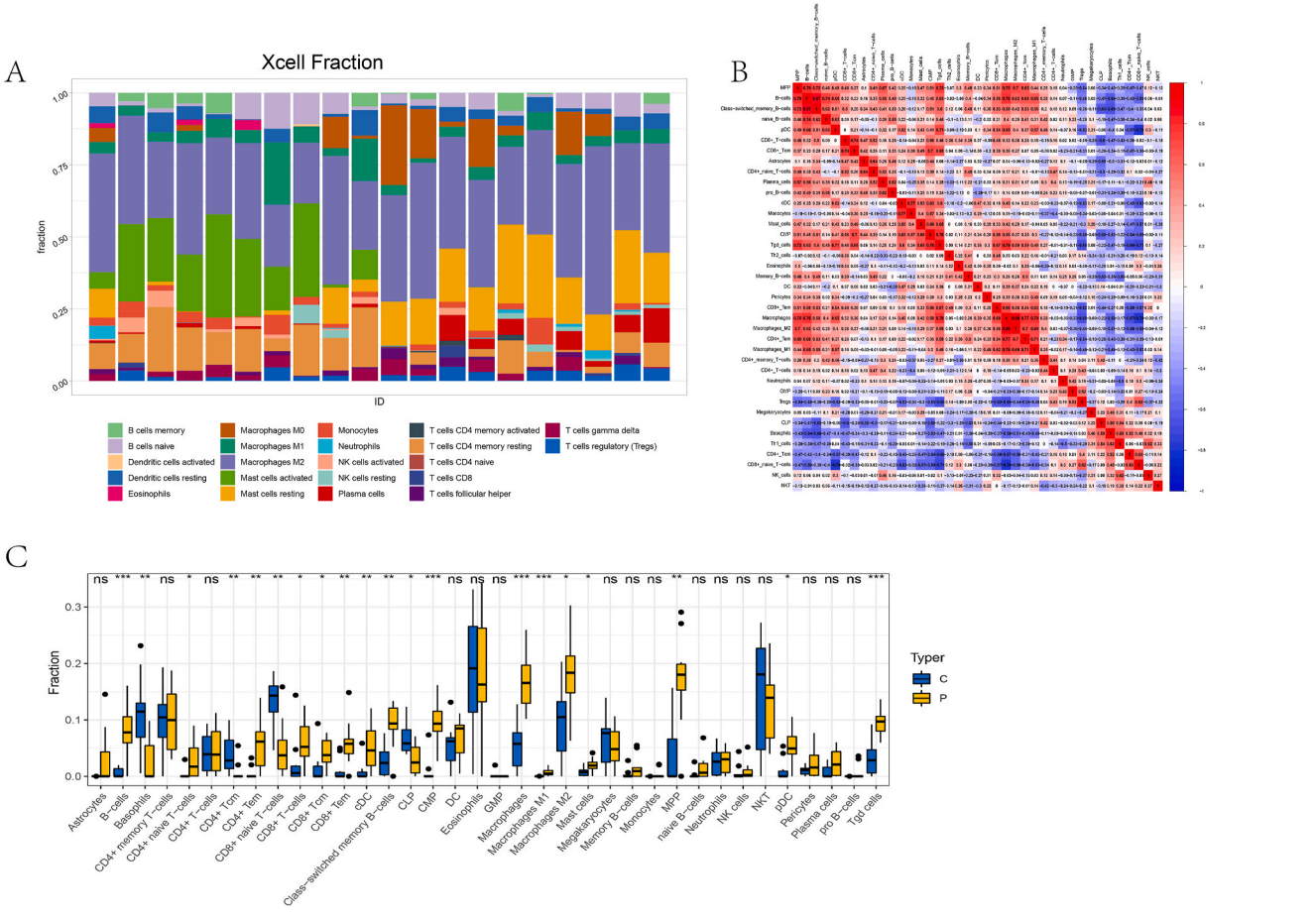

Fig. 3. Immune infiltration analysis
(A) : Immune infiltration analysis via xcell algorithm , the enrichment fraction diagram of infiltrating immunocytes; (B): Heatmap of inter-immune cell correlations; (C) : Box diagram of immune infiltrating cells in Control and OA samples. $p < 0.05$,"*"; $p < 0.01$,"**"; $p < 0.001$,"***".

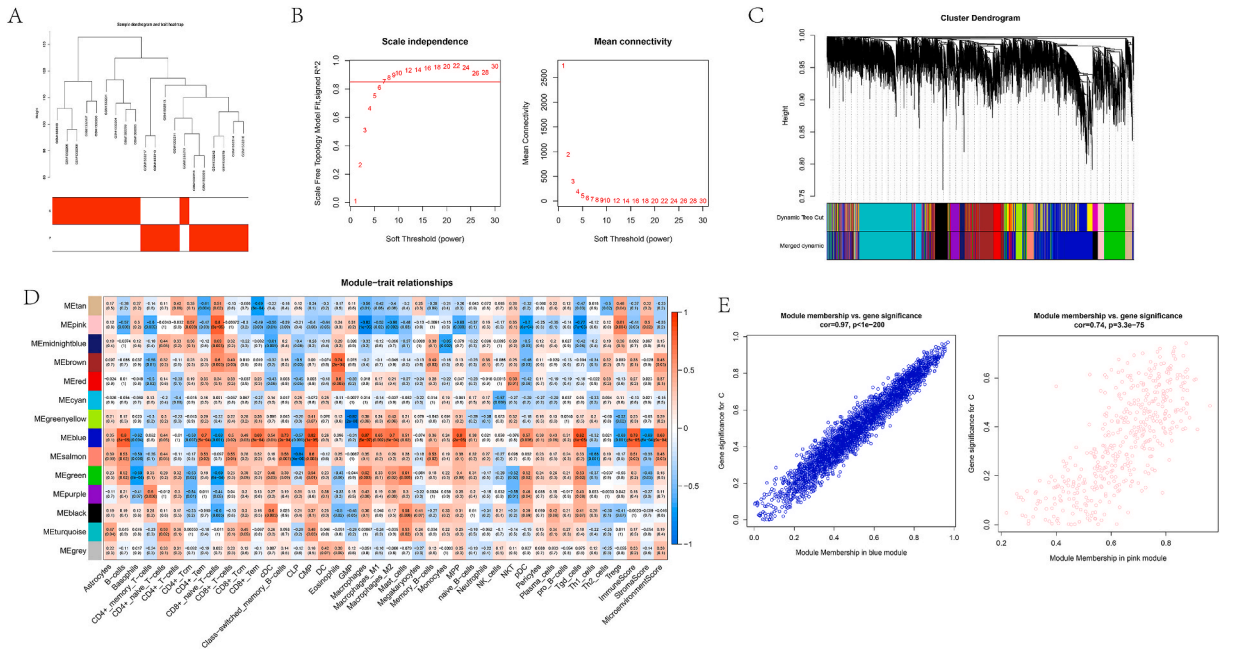

Fig. 4. WGCNA analysis to appraise osteoarthritis immune hub module.
(A): Clinical sample trait clustering; (B) : Analyze optimal soft thresholds power; (C) : Modules discovered by combining modules with feature factors larger than 0.25; (D) : Heatmap of correlation between WGCNA module and immune cells; (E): Scatter plot of key modules (left: blue module, right: pink module).

Macrophages_M1 and Macrophages_M2 (cor = −0.66, p = 0.002) (Fig. 4 D). Finally, the Blue (cor = 0.97, p < 1e-200) and Pink (0.74, p = 3.3e-75) modules were selected as the target modules in that they had the highest correlation with immunity (Fig. 4 E).

3.4. *Functional and pathway enrichment analysis*

From the previous Immune infiltration results, we found the blue and pink modules had the highest correlation with macrophages. We downloaded 500 genes related to polarized macrophages from Genecards (https://www.genecards.org/) . We identified 71 differentially expressed polarized macrophage-related genes and the intersection of genes related to DEGs- WGCNA-macrophage polarization is shown in Fig. 5 A.

To explore the role of 71 differentially expressed polarized macrophage genes in osteoarthritis. The terms GO, KEGG, and DO are shown in Fig. 5 (B-G) . The aspects of such results presented by GO enrichment analysis were as follows : BP (Biological processes): leukocyte cell-cell adhesion, leukocyte migration; CC (Cellular component): intracellular vesicles; MF (Molecular function): immune receptor activity, cytokine receptor binding (Fig. 5 b). By KEGG analysis, the results showed that most DEGs were enriched in rheumatoid arthritis, leishmaniasis, lipids and arteriosclerosis, osteoclast differentiation, and NF-kappa B signaling pathway (Fig. 5c). GSEA revealed several up-regulated pathways, including allograft rejection, asthma, degradation of other sugars, primary immune deficiency and type I diabetes. As for down regulated pathways, circadian rhythm, IL-17 signaling pathway, Legionnaires' disease, mineral absorption and cancer were discovered (Fig. 5*E*–F). The ridge diagram will show the expression distribution of the core enriched genes in GSEA enriched category, from which we can see that the up-regulated pathways mainly include: Phagosome, Lysosome, Cell adhesion molecules, Leukocyte transendothelial migration, etc. the down-regulated pathways mainly include: NF-kappa B and TNF(Fig. 5G).

3.5. *Machine learning-based hub genes screening*

To further identify potential biologically critical genes for osteoarthritis from the 71 target genes, we performed three different kinds of machine learning. Firstly, according to SVM-RFE analysis, the SVM model showed the best accuracy and error rate based on 19 characteristic genes (Fig. 6A and B). Therefore , we identified 19 genes CSF1R, CXCL12, MYC, AHR, JAK2, CX3CR1, CD4, CEBPB, TNFSF11, ITGAM, NFKBIA, VCAM1, OLR1, CXCL2, HLA-DPB1, CIITA, TLR7, IL1R1, EDN1. In addition, based on osteoarthritis and control samples, we determined the most appropriate log (λ) (=7) by 10x cross validation when constructing the LASSO model (Fig. 6C). Finally, with nonzero coefficients, 7 genes identified were as follows: CX3CR1, CXCL12, HLA-DPB1, TLR7, CSF1R, CEBPB, ITGAM. We then, used random forest to further filter down the 71 target genes. A recurring random forest categorization was conducted for all possible values of the 71 genes and the average error rate was analyzed in such model. We finally selected 14 genes with MeanDecreaseAccuracy greater than 0.014 (MYC, ZFP36, CEBPB, CX3CR1, KLF4, CXCL2, NFKBIA, TNFSF11, VEGFA, TREM2, TLR7,

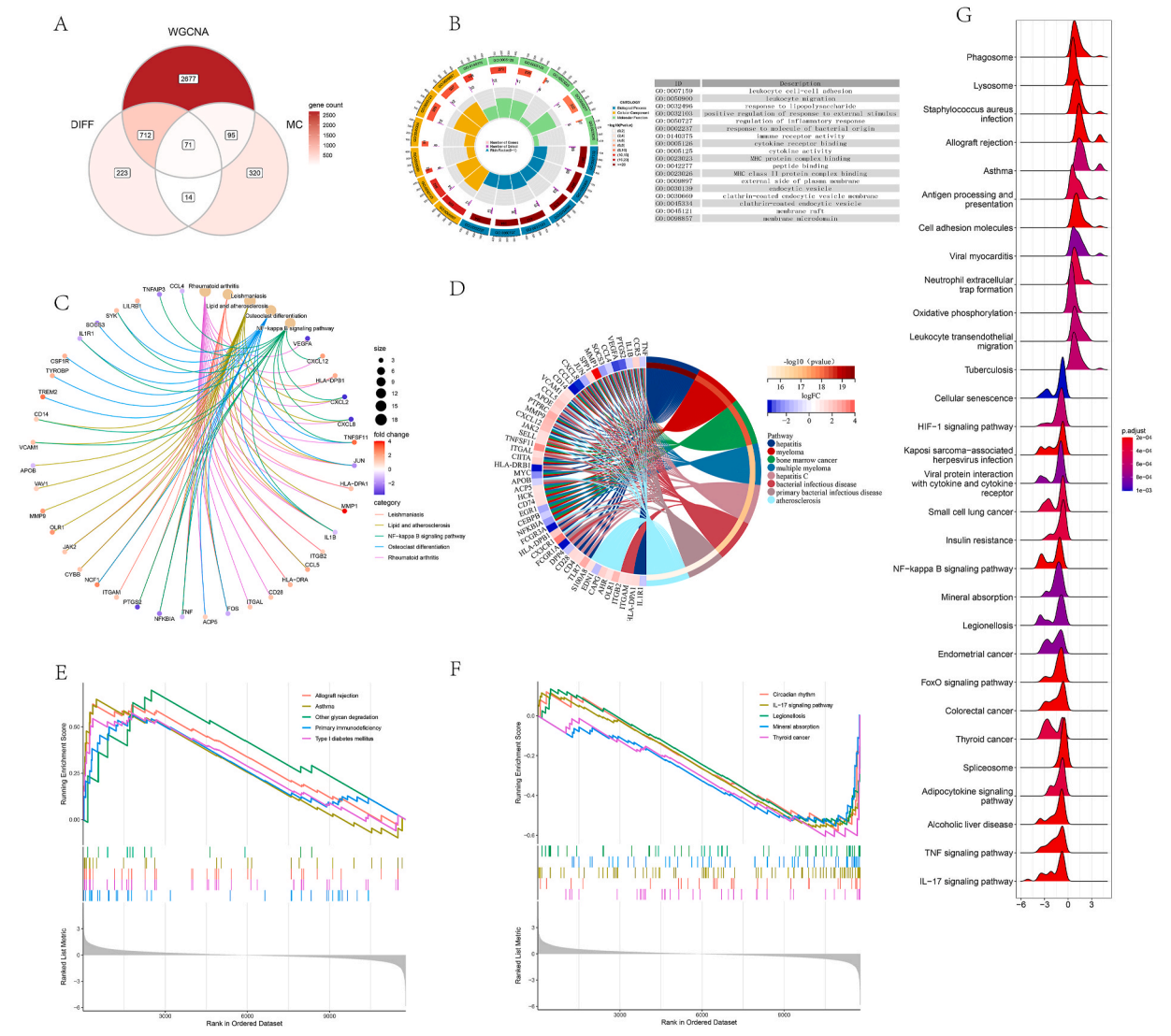

Fig. 5. Functional analysis of DEGs-WGCNA-macrophage polarization genes
(A): Venn diagram of genes associated with polarized macrophages; (B):GO analysis bar graph of polarized macrophages; (C):KEGG analyzes chordal diagrams of polarized macrophages; (D):DO analyzes circle diagram of polarized macrophages; (E), results of up-regulated pathway in GSEA analysis; (F):results of down-regulated pathway in GSEA analysis; (G):Ridge map showing GSEA enrichment pathway.

JUN, CXCL8, CSF1R) as the follow-up study the expected genes (as in Figure D). The Venn diagram showed the overlapping genes in LASSO, SVM and RF modules, and finally, four hub genes CSF1R, CX3CR1, CEBPB, TLR7 were identified (Fig. 6E).

3.6. PPI network construction

We build a PPI network further to dig into the protein interactions of the target genes. Supplementary Figure (S1) shows that the PPI network includes 71 nodes with 854 edges. We use red and green to distinguish between up-regulated and down-regulated genes. Such interaction maps of the first five go entries of the 71 target genes with the hub genes were conducted then (Fig. 7).

3.7. Biological functions and regulating mechanisms of the potential biomarkers

With the plan to deeply discover the biological roles and regulatory mechanisms of CSF1R, CX3CR1, CEBPB and TLR7, GSEA was performed based on their ordered gene expression matrices. We performed single-gene GSEA functional analysis of four potential biomarkers of osteoarthritis: we found that in addition to the previously enriched pathways, some new pathways such as: Rap 1 signaling pathway, mucin type o-glycan biosynthesis (Fig. 8A–D) were identified, suggesting possible involvement in macrophage migration and polarization. To analyze the correlation between hub genes, we then utilized the Pearson correlation coefficient

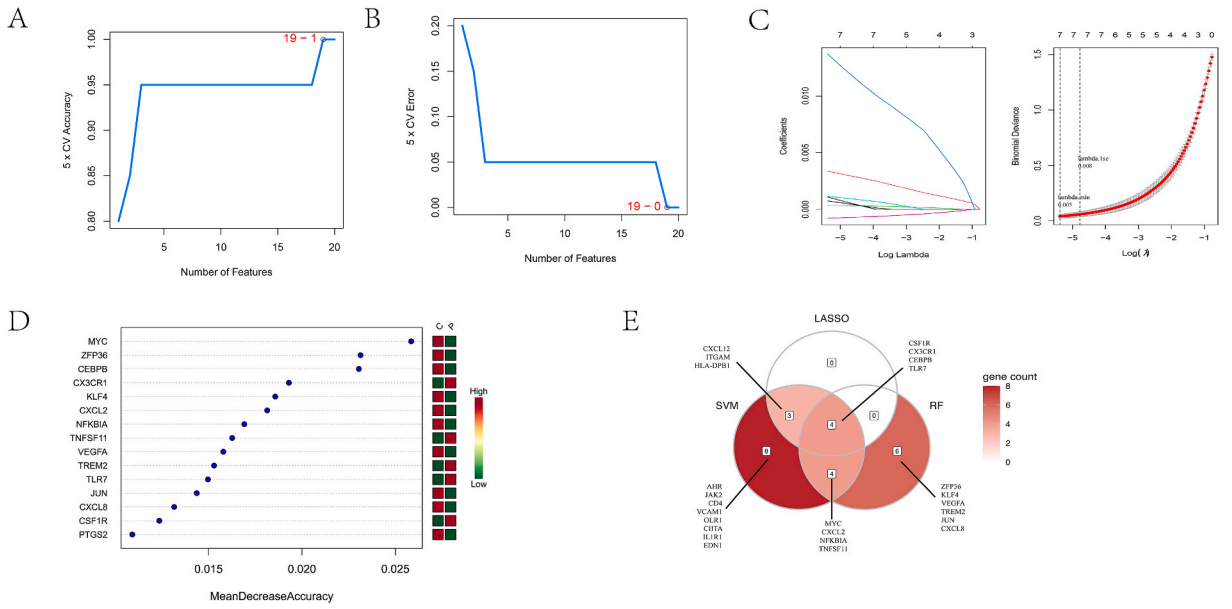

Fig. 6. Machine learning-based screening of possible biomarkers
(A-B) : Optimal accuracy and error rate of SVM model based on 19 feature genes; (C) : Logarithmic (Lambda) values of the three genes in the LASSO model (left), the most suitable logarithmic (Lambda) values in the LASSO model (right); (D) : Results of random forest analysis. x-axis shows genetic variables, y-axis shows decrease in mean accuracy; (E): Venn diagrams of three types of machine learning, showing their overlapping genes.

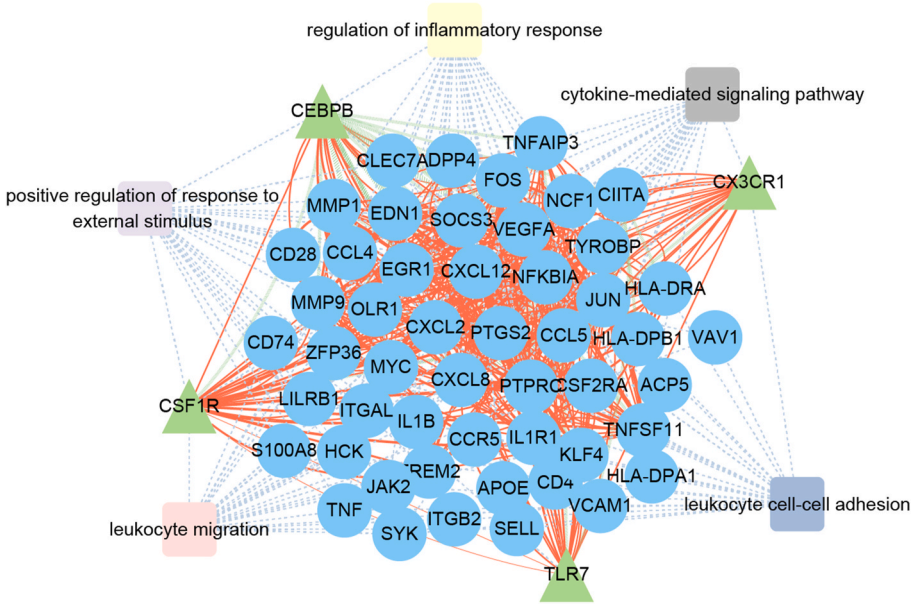

Fig. 7. Interaction networks of 71 target genes with hub genes and GO entries.

algorithm. The heat map showed a significant positive correlation between CEBPB and TLR7 (Supplementary Fig. S2).

3.8. The result relation from hub gene immune infiltration analysis and the regulation of macrophage biological process

We exacted 25 macrophage-related biological processes from msigdb's C5: ontology gene sets. The enrichment fraction of gene characteristics of CSF1R, CX3CR1, CEBPB, and TLR7 were calculated by R package GSVA. The heat maps of hub genes and macrophage-related biological processes were drawn by R packages "ggplot2" and "ggcor". In the outcome of immune infiltration, we found that CEBPB was positively correlated with macrophage M0, M1, M2 ($p < 0.01$) and $CD4^+$ memory T-cells ($p < 0.05$). CEBPB

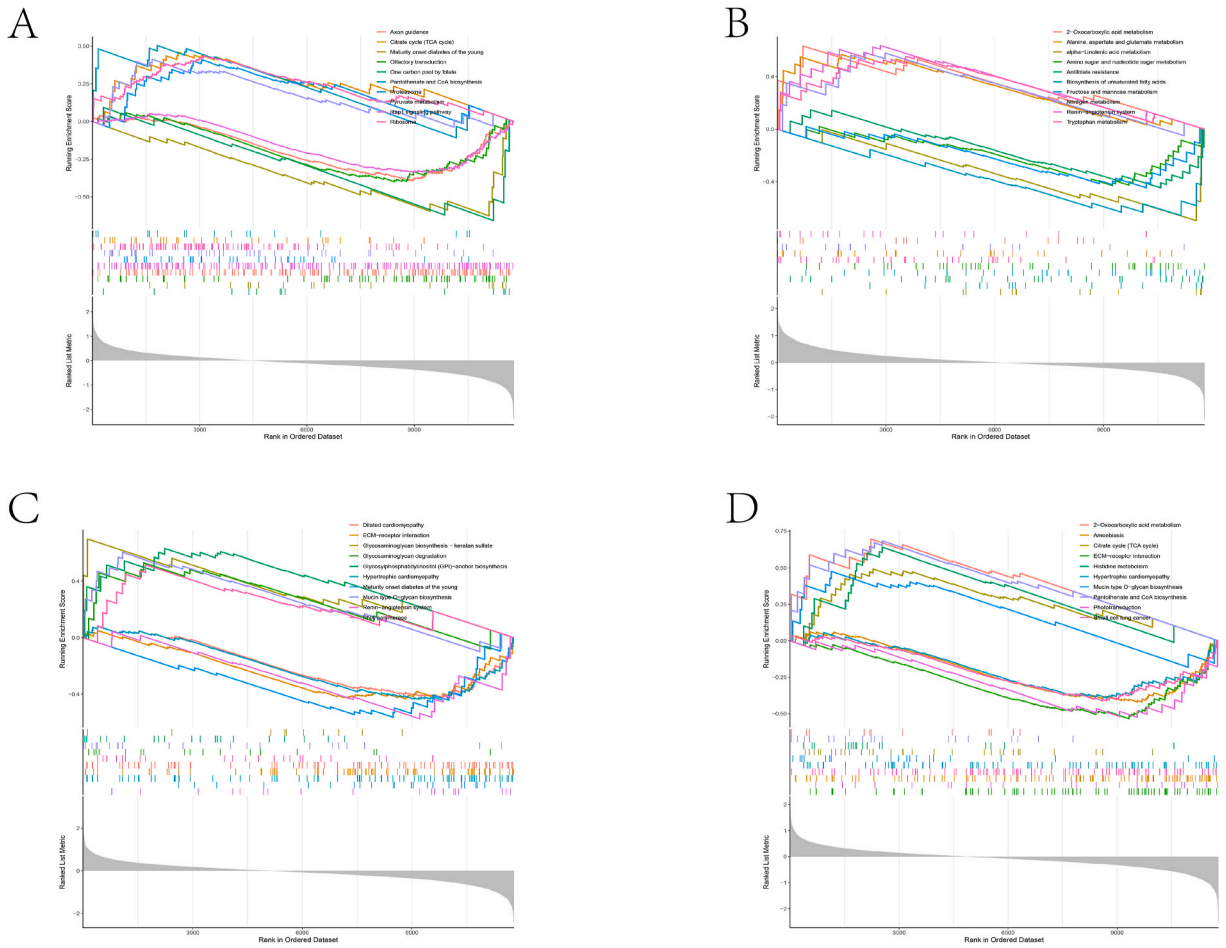

Fig. 8. Single genes GSEA analysis of hub gene. (A:CSF1R; B: CX3CR1; C：CEBPB; D：TLR7).

was positively correlated with biological processes like macrophage activation involved in immune response, macrophage chemotaxis, and macrophage inflammatory protein 1 alpha production (Fig. 9A).

From outcome of immune infiltration，CSF1R played a negative role in correlation with $CD4^+$ T cells and NK cells ($p < 0.05$) but positively correlated with $CD4^+$ Tem. CSF1R was positively correlated with biological processes like macrophage cytokine production, granulocyte-macrophage colony stimulating factor production, macrophage activation, regulation of macrophage activation, macrophage apoptotic process, rule of cellular response to macrophage colony stimulating factor stimulus, regulation of macrophage apoptotic process. ($p < 0.05$) (Fig. 9B).

CX3CR1, in a negative correlation with Neutrophils ($p < 0.05$) , acted in a positive correlation with Megakaryocytes instead ($p < 0.05$) . TLR7 was negatively correlated with CD4+Tcm ($p < 0.01$),GMP($p < 0.05$)but positively correlated with Macrophages ($p < 0.05$). However, Neither CX3CR1 nor TLR7 was significantly involved in macrophage biological processes (Fig. 9C–D).

3.9. Potential biological criteria expression and macrophage correlation

To further verify the correlation between macrophages and hub genes. To show the correlation between hub genes and macrophages directly, we then used the R package “ggExtra” to plot the marginal density according to the results of the previous immunoinfiltration analysis. The correlation coefficients of CEBPB with Macrophages and Macrophages M1 were −0.59 and −0.52, respectively (Fig. 10A–B); the correlation coefficients of CSF1R respectively were 0.69 and 0.6 with Macrophages M1 and Macrophages M2 (Fig. 10C–D). Additional correlations between other hub genes and macrophage polarization are shown in the Supplementary Figure (s3).

3.10. Expression validation of potential biomarkers

To further explore the role of hub genes CEBPB, CSF1R, CX3CR1, and TLR7 in osteoarthritis, we compared their expression levels in osteoarthritis patients and healthy groups using the analytical dataset GSE55235 and the validation dataset GSE55457, respectively.

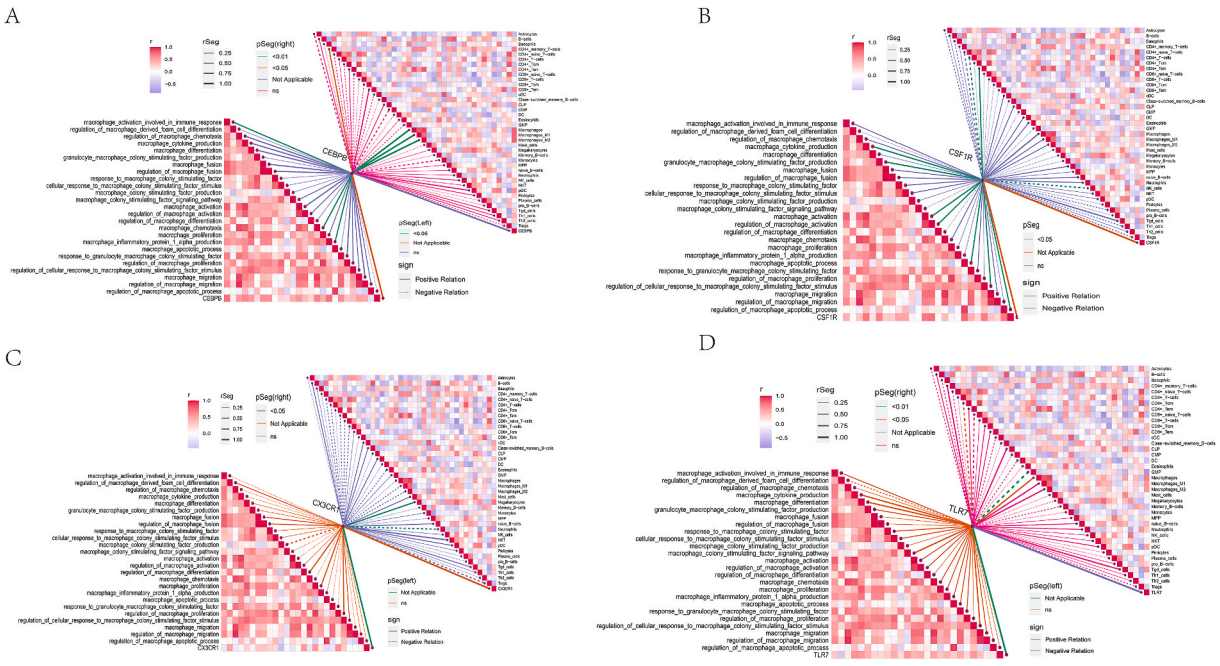

Fig. 9. Diagram of the relationship between immune infiltration and macrophage regulation for (A:CEBPB; B:CSF1R; C : CX3CR1; D : TLR7).

The AUC of the ROC of CEBPB, CSF1R, CX3CR1, and TLR7 for the analyzed dataset GSE55235 was 0.990, 1.000, 1.000, and 1.000, respectively (Fig. 11A–D); The p-values for the difference in expression of CSF1R, CX3CR1, CEBPB, and TLR7 between normal control and osteoarthritis groups were 1.1e-5, 1.1e-5, 2.2e-5, 1.1e-5 (Fig. 11E).

The AUC from the ROC curve of CEBPB, CSF1R, CX3CR1, and TLR7 in verification data set GSE55457 were 0.900, 0.780, 0.900, 0.930, individually (Fig. 12A–D). The p values of CSF1R, CX3CR1, CEBPB, and TLR7 expression differences between control group and OA group were 0.04, 1.5e-3, 1.5e-3, and 4.9e-4, respectively (Fig. 12E).

The hub genes CEBPB, CSF1R, CX3CR1, and TLR7 were same at the level of expression in the two data sets, and the ROC curves showed that the 4 genes can be used as diagnostic biomarkers for OA.

3.11. Gene regulatory networks targeted by TF and miRNA and potential therapeutic drugs prediction

To investigate the regulatory mechanisms of CEBPB, CSF1R, CX3CR1and TLR7, we predicted their TF and miRNA and constructed a TF-hub gene-miRNA sang-map, in which CSF1R was regulated by FOXC1, TFAP2A, hsa-miR-155-5p, and hsa-miR-34a-5p; CEBPB was regulated by NR2F1, PRRX2, hsa-miR-155-5p, and hsa-miR-374a-5p (Fig. 13). In addition, we predicted the possible potential drugs (Supplementary Fig. S4).

3.12. Single-cell sequencing analysis reveals the distribution of hub genes

We performed quality control and normalization on the data. Fig. 14A–B shows the quality control plots for single-cell data. The left panel of A shows the relationship between nCount_RNA and percent. mt, with a correlation coefficient of −0.2, indicating that the lower the percentage of mitochondrial DNA in the total DNA, the higher the total number of RNA molecules. The right panel of A shows the relationship between nCount_RNA and nFeature_RNA, with a correlation coefficient of 0.94, indicating that the higher the total number of RNA molecules, the more genes are expressed. Based on Figure B, we filtered out most doublets/dead cells/empty droplets. Then we named each cell based on the top genes of each cell cluster, namely sub-synovial fibroblasts (SSF), synovial fibroblasts (SIF), HLA-DRA + cells, smooth muscle cells (SMC), endothelial cells (EC), T cells, mast cells and proliferative immune cells (ProIC) (Fig. 14C–E). Finally, we used umap and bubble plots to show the expression of hub genes in synovial tissue, and found that they were mainly expressed in HLA-DRA + cells (Fig. 14D–F).

4. Discussion

In recent years, the incidence of osteoarthritis has been increasing [38]. For OA is difficult to diagnose in the early stage, and the late intervention method is mostly surgery, it often leads to poor prognosis of patients and brings heavy economic burden to patients and society. We used bioinformatics technology to compare OA patients with normal people to find OA biomarkers, which is helpful for disease diagnosis, discovery of potential target drugs, improvement of prognosis, etc. In previous bioinformatics studies, researchers

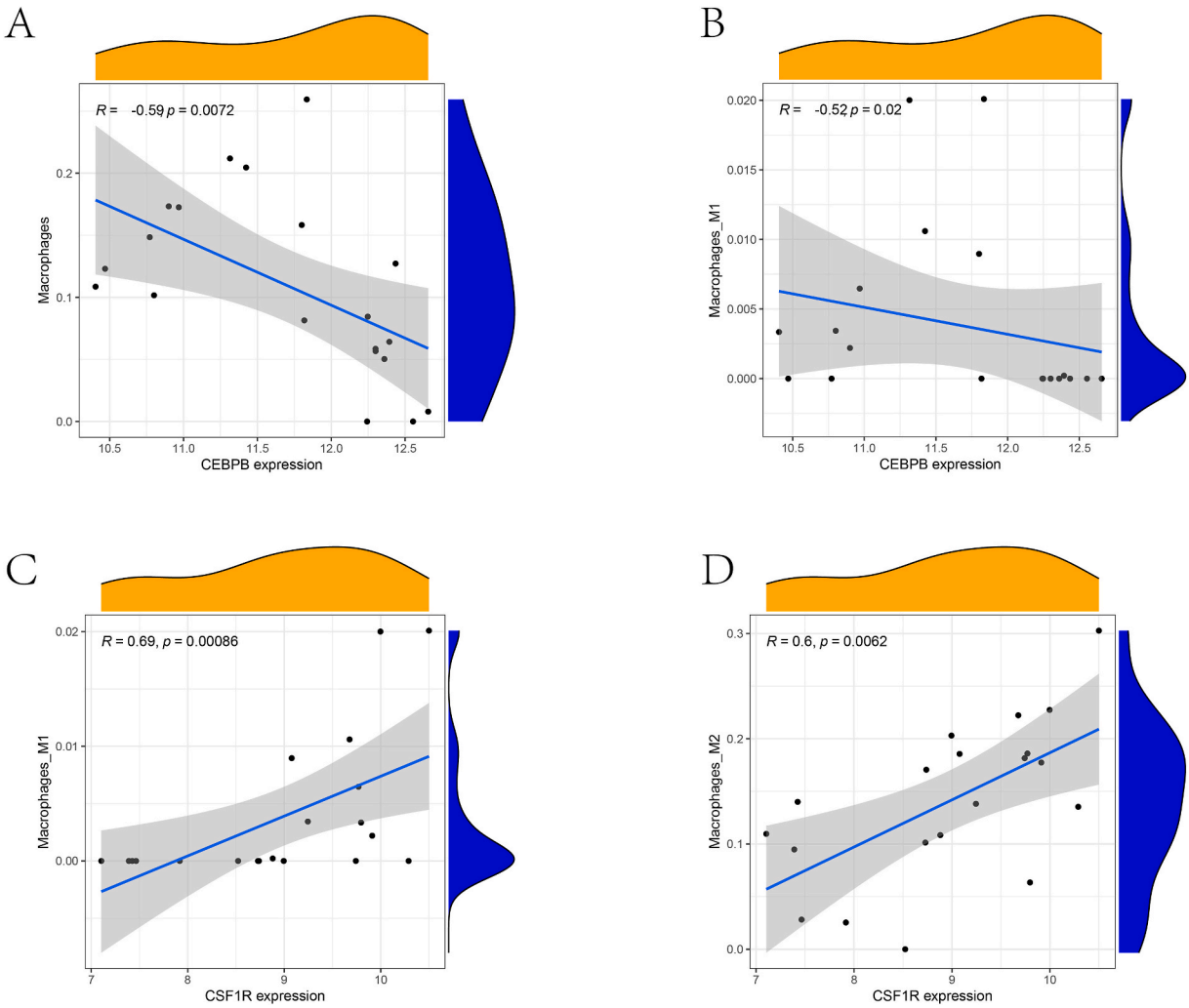

Fig. 10. Scatter plot of the expression of hub genes CEBPB and CSF1R with different polarization states of macrophages. (A): Scatter plot of CEBPB and Macrophages; (B): Scatter plot of CEBPB and Macrophages M1; (C): Scatter plot of CSF1R versus Macrophages M1; (D): Scatter plot of CSF1R versus Macrophages M2.

focused on the simple functional enrichment analyzes of differential genes, which only explored the expression of hub genes in diseases but lacking further functional analysis.

Firstly, we analyzed osteoarthritis sequencing data using two algorithms, limma and WGCNA. In addition, we downloaded macrophage polarization-related genes from the website GeneCards, and perform functional annotation on these common 71 genes. From GO analysis, we found these genes involved in biological processes mainly include leukocyte cell-cell adhesion and leukocyte migration. In addition, the therapeutic Wnt inhibition alleviates disease severity in traumatic OA by promoting the anticatalytic effect of chondrocytes and the antifibrotic effect on synovial fibroblasts. Any drugs that can inhibit Wnt singaling pathway may be promising for the treatment of OA [39]. Leukocyte recruitment is a vital factor in inducing the arthritic state. In synovial fluid and synovial tissue, leukocyte infiltration indicates inflammatory arthritis, and preventing leukocyte recruitment can be a promising way to relieve OA [40]. In addition, KEGG analysis includes NF-kappa B signaling pathway and Lipid and atherosclerosis. NF-κB (nuclear factor-κB) , working as a transcription factor, functions in the expression of genes that regulate cell proliferation and apoptosis and react to inflammation and immune responses. In osteoarthritis (OA) model, the degree of NF-κB activation was confirmed to modulate the promotion or inhibition of OA [41]. NF-κB signaling pathway impacts cartilage matrix reshaping, chondrocyte death, and synovial inflammation and also indirectly stimulates some downstream regulators of terminal chondrocyte differentiation [42]. The NF-κB signaling pathway acts as a vital factor in the onset and progress of OA. By activating Nrf2 and inhibiting the STING-dependent NF-κB pathway, some unfavorable events like inflammatory reaction, ECM degradation, and senescence of chondrocytes triggered by IL-1β may be alleviated. Regulating the synovial macrophages polarizing via the NF-κB axis does help to attenuate the course of OA [43]. DO analysis was similar to KEGG in that OA patients were more likely to suffer from atherosclerosis simultaneously. At the early stages of OA, lipid deposition in joints is observed before tissue changes [44]. On the other hand, different from the previous differential gene

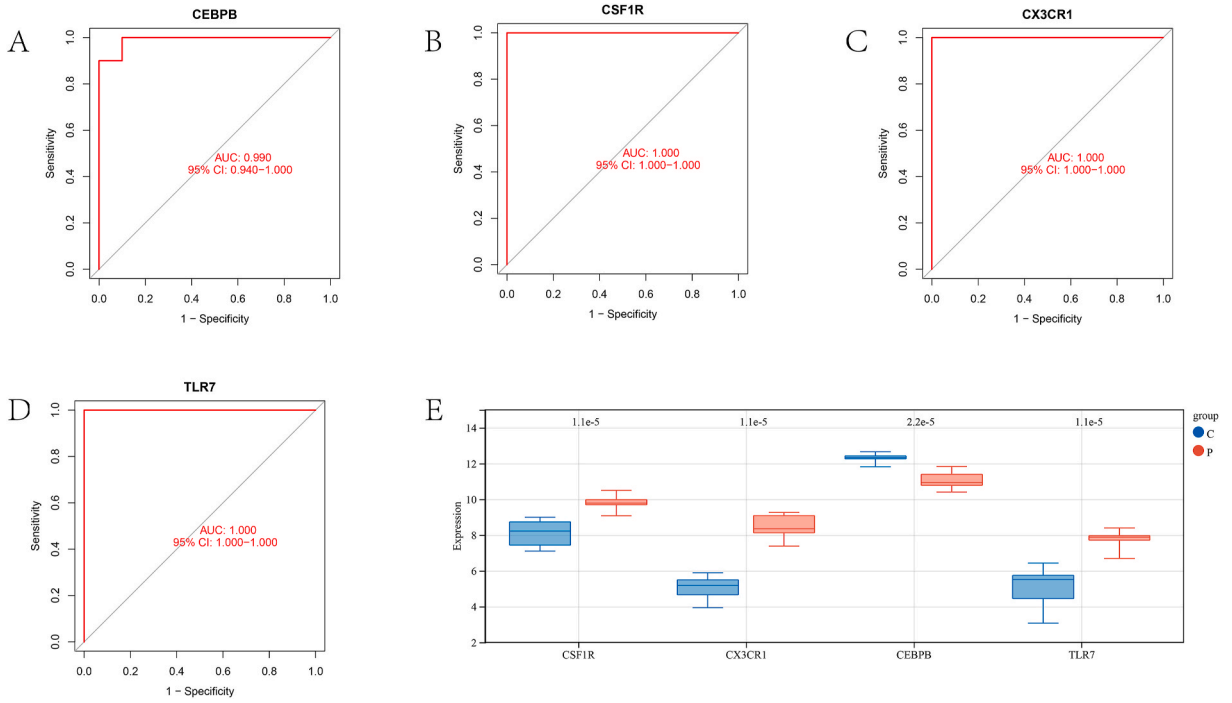

Fig. 11. ROC curves and box plots showing potential biomarkers for the analyzed data set GSE55235, (A–D): ROC curves of CEBPB, CSF1R, CX3CR1, and TLR7 for the analyzed data set GSE55235; (E): box plots of CEBPB, CSF1R, CX3CR1, and TLR7 for the analyzed data set GSE55235.

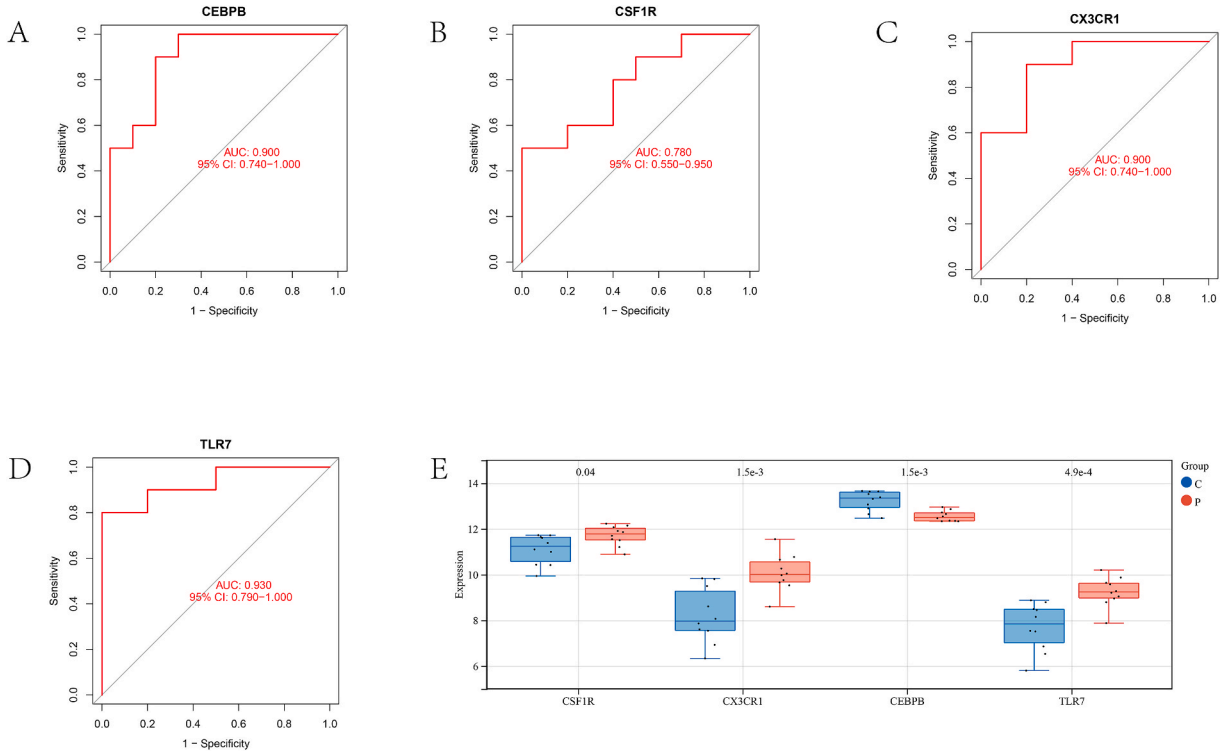

Fig. 12. ROC curves and box plots showing potential biomarkers for the Validation dataset GSE55457. (A–D): ROC curves of CEBPB, CSF1R, CX3CR1, and TLR7 for the Validation dataset GSE55457; (E): box plots of CEBPB, CSF1R, CX3CR1, and TLR7 for the Validation dataset GSE55457.

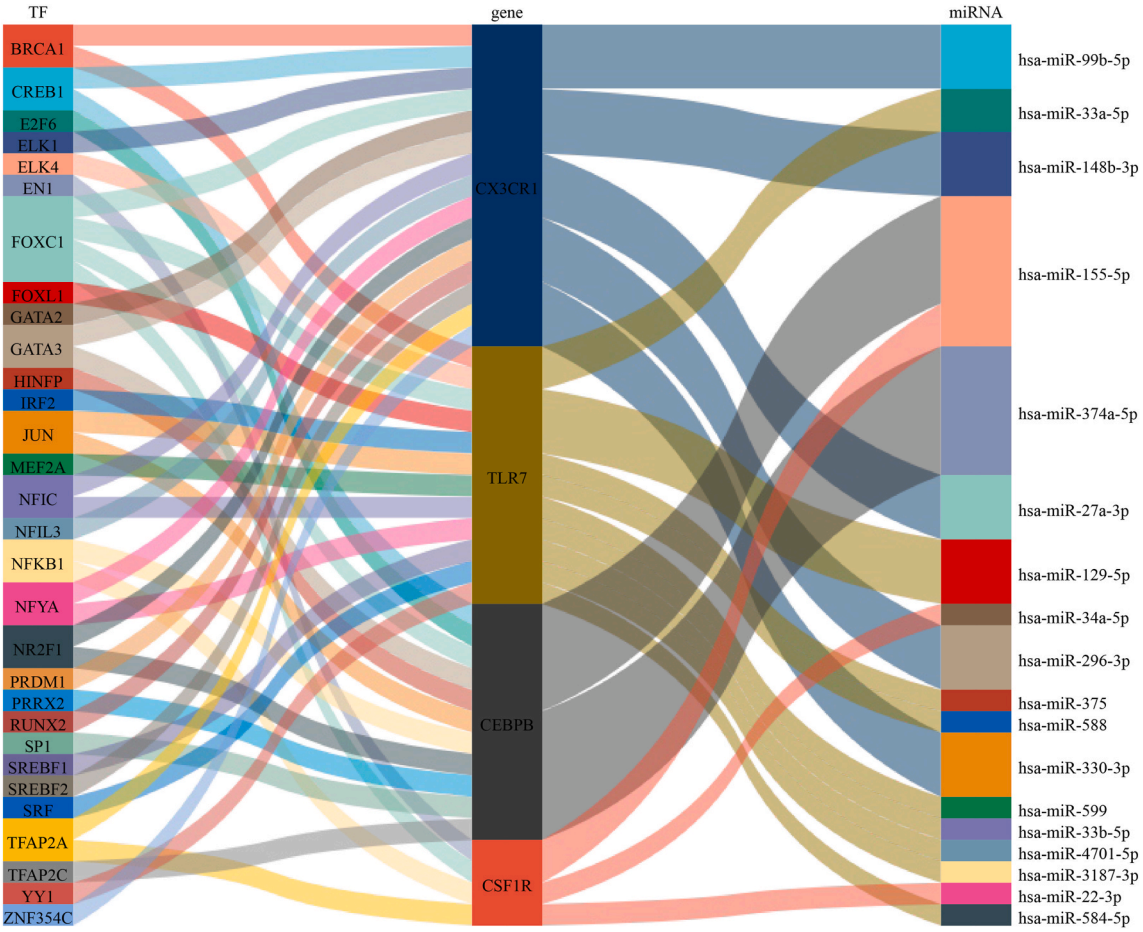

Fig. 13. TF—hub—miRNA sankitu.

enrichment analysis, GSEA can better reflect the pathways involved in biological disease processes. As a newfound T cell-derived cytokine, IL-17 was proved vital for the bone resorption of osteoclast [45]. Therapies targeting IL-17 can improve clinical manifestations of psoriatic arthritis but have not been discovered in OA patients [46,47]. In senescent cells, researchers find that intra-articular injection of IL-17 neutralizing antibody reduces joint degeneration and expression of Cdkn1a, a senescence marker.

Next, we combined LASSO logistic regression, SVM-RFE, and RF algorithms to identify CSF1R, CX3CR1, CEBPB, and TLR7 as potential biomarkers. To further verify the reliability of these 4 biomarkers, another synovial inflammation dataset was used for validation. The AUC value of the ROC curve from each hub gene in the analysis data set and validation set was more significant than 0.65, indicating that the hub genes we screened can be used as diagnostic biomarkers for osteoarthritis. In addition, we verified the expression of hub genes in the osteoarthritis group and the healthy control group, finding that the expression of each hub gene was consistent in the analysis data set and validation set. (The hub genes CSF1R, CX3CR1, and TLR7, are all up-regulated in osteoarthritis patients, and CEBPB is down-regulated in osteoarthritis patients.)

CSF1R, one of the most common pro-inflammatory cytokines, can lead to various inflammatory conditions. Colony-stimulating factor-1 (CSF-1), also known as macrophage colony-stimulating factor (MCSF), is the specific ligand of CSF1R [48]. Combination of the two triggers the differentiation of macrophages and cells proliferation from the monocyte/macrophage lineage. Also, CSF1R is thought to regulate monocyte migration and activation, which were considered as an essential factor in the pathogenesis of arthritis [49]. There was a comment that blocking the CSF-1 receptor can protect bone and cartilage, so that relieves OA progression [50]. Both CSF1 and IL-34 are crucial for the survival, proliferation, and differentiation of macrophages. CSF1R signaling not only drives macrophage differentiation but also plays a role in controlling tissue homeostasis, repair, and inflammation. However, developmental defects in knockout models for CSF1R, CSF1, or IL34 have limited the utility of these models in studying the receptor's function under steady-state and inflammatory conditions [51,52]. In the context of OA, the stimulation of CSF1R by CSF-1 or IL-34 promotes macrophage differentiation, activation, and osteoclastogenesis. Pharmacological inhibition of CSF1R has been beneficial in animal models of arthritis, suggesting that modulating CSF1R signaling could be a potential therapeutic strategy for OA. The relative contributions of CSF-1 and IL-34 to CSF1R signaling in rheumatoid arthritis (RA) have been investigated, providing insights into how these ligands may influence the disease process [53].

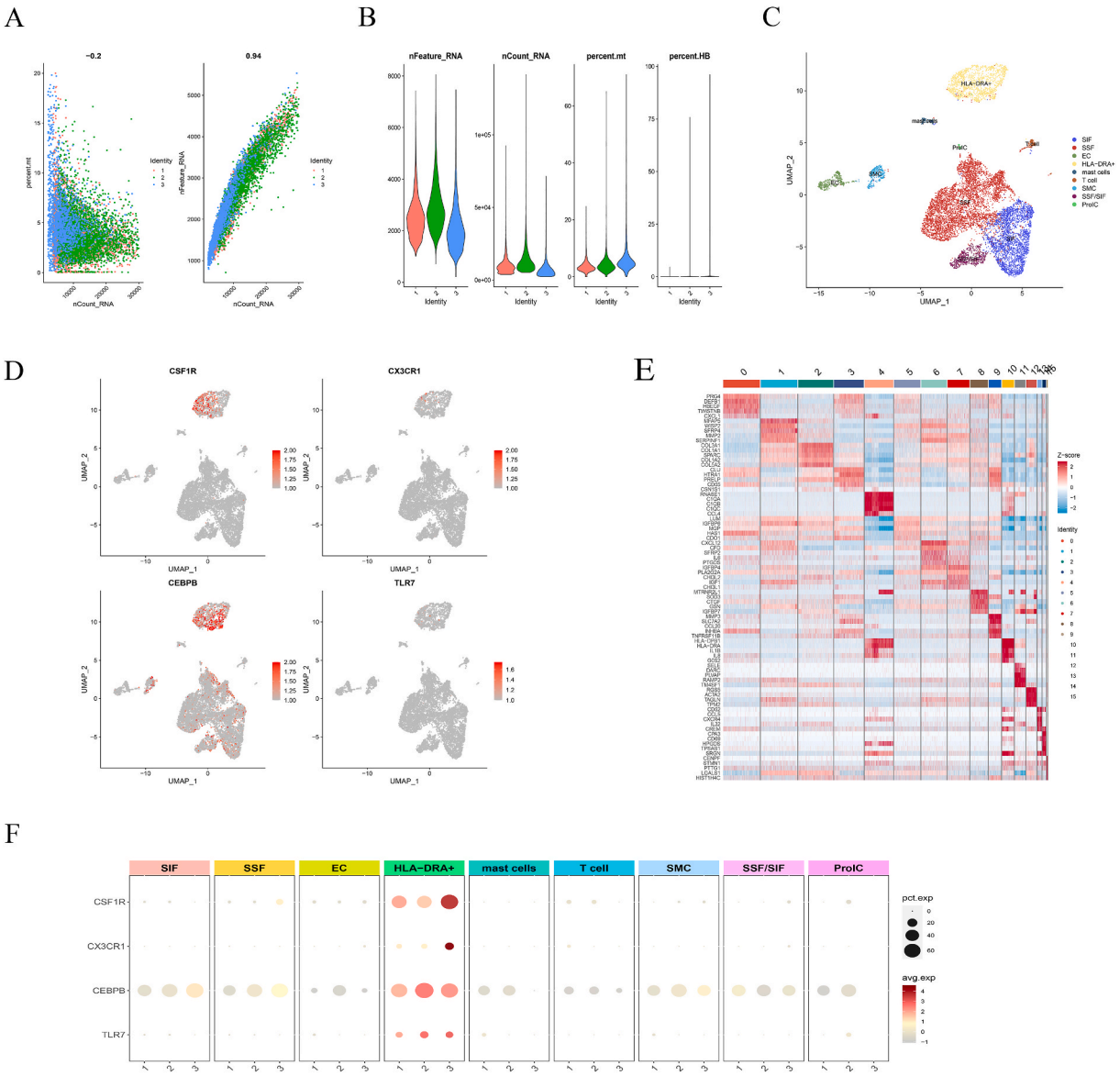

Fig. 14. Visualization of single-cell results.
A:Scatter plot and correlation coefficient of the four features in the dataset
B:violin plot showing the distribution of the four features in the dataset
C:umap plot showing the clustering results and distribution of the cells in the dataset
D:hub gene of the cells separately
E:heat map showing the expression levels of cell clusters on the five genes with the largest mean log fold change in each cluster
F:bubble plot showing the distribution of hub genes in cells as well as inside the sample.

CX3CR1 motif chemokine receptor 1 (CX3CR1) comes from the G protein-coupled receptor 1 (GPCR1) family. As the sole known member of the CX3C chemokine receptor subfamily till now, CX3CR1 is also the fractalkine receptor or G-protein coupled receptor 13 (GPR13) [54]. The CX3CL1/CX3CR1 axis is involved in leukocyte recruitment and chemokines associated with many inflammatory diseases [54–56]. According to Culemann [57], CX3CR1+ synovial macrophages are situated in the topmost synovial layer atop the fibroblast-like synoviocytes. To limit inflammatory reactions and preserve the intra-articular structures, the thick physical barrier created by CX3CR1+ macrophages separates the articular cavity from the surrounding synovium [57]. In OA, $CX3CR1^{+}$ macrophages form a tight-junction-mediated barrier at the synovial lining, which is thought to protect the joint [57]. Additionally, a subset of macrophages characterized as $TREM2^{+}CX3CR1^{+}FOLR2^{+}$has been identified, suggesting a complex heterogeneity within the joint macrophage population that could be relevant for understanding the pathophysiology of OA and designing targeted therapies [58].

The CCAAT/enhancer-binding protein β (C/EBPβ), encoded by the CEBPB gene, is a transcription factor implicated in the

regulation of various biological processes, including immune responses and endochondral ossification. In the context of OA, C/EBPβ has been identified as a potent regulator, with computational predictions and motif-reporter assays indicating interactions with RUNX2, a key player in skeletal development and OA progression [59]. Notably, C/EBPβ has been found to regulate the signaling pathway between β-adrenergic stimulation and the expression of Arg1, a gene associated with the M2 macrophage phenotype. β-adrenergic signaling induces Cebpb gene expression, which subsequently leads to Arg1 gene expression, suggesting a potential mechanism through which C/EBPβ influences macrophage polarization [60]. This finding is particularly significant as it links macrophage polarization with a known molecular pathway, providing a potential target for therapeutic intervention. Furthermore, the role of C/EBPβ extends beyond the regulation of macrophage polarization. It has been shown to interact with other signaling molecules and transcription factors, such as NF-κB, that are involved in immune responses, including emergency myelopoiesis, which can be relevant in the context of inflammation and OA [61].

TLR7 is a member of the Toll-like receptor family, which plays a pivotal role in the innate immune response by recognizing pathogen-associated molecular patterns. In the context of OA, TLR7 activation can lead to the production of pro-inflammatory cytokines and contribute to joint inflammation and damage. The link between TLR7 and OA has been strengthened by the identification of an X-linked TLR7 polymorphism, suggesting a genetic predisposition to the disease [62]. The interaction between macrophage polarization and TLR7 signaling is complex and can significantly influence the inflammatory environment within the OA joint. For instance, the activation of TLR7 can promote the polarization of macrophages towards the M1 phenotype, as evidenced by studies showing an increase in M1 macrophages upon TLR7/8 activation [63]. This suggests that TLR7 may serve as a therapeutic target to modulate macrophage polarization and, in turn, ameliorate OA symptoms.

In OA models, the synovial tissue release miR-21, which causes knee pain by TLR7 in surgical rats. Extracellular miRNAs in joints were performed to be a potentially reasonable therapeutic strategy for the relief of pain induced by OA [64]. Hypoxia-inducible factor 2α (HIF-2α) may function as a regulator of OA by enhancing endochondral ossification. Induced by HIF-2α and targeting MMP-13, C/EBPβ and RUNX2 govern cartilage degeneration. Comparative research about this molecular network in chondrocytes may provide a meaningful reference in clinical treatment for OA patients [59].

While this study identifies potential biomarkers and therapeutic targets for osteoarthritis, it is important to consider its limitations. The reliance on bioinformatics and publicly available datasets introduces potential biases from variations in sample collection and patient characteristics. Future research should validate these findings through independent cohorts and experimental models to solidify the conclusions. Additionally, exploring interactions between macrophages and other cell types like chondrocytes could provide deeper insights into osteoarthritis pathogenesis. The roles of identified targets like CSF1R and CEBPB need further experimental validation to confirm their efficacy and safety as therapeutic options. Integrating multi-omics data and conducting functional studies on the identified genes could enhance our understanding of the molecular mechanisms driving the disease.

5. Conclusion

In summary, we screened 71 target genes from OA patients' synovial tissue and identified CEBPB, CSF1R, CX3CR1, and TLR7 as potential diagnostic biomarkers for osteoarthritis through machine learning. The 4 genes may get involved in macrophage activation. Among them, CEBPB and CSF1R were involved in various macrophage-related pathways, which may be 2 new drug targets in OA.

Data availability statement

The datasets in this study can all be found in public repositories. The names and login numbers of the repositories can be found in the article/supplementary material.

Ethics statement

The human participant data we used were downloaded from a public database site and the study did not require ethical review or approval, as required by local legislation and agencies.

Funding

This research was supported by the Wu Jieping Medical Foundation Clinical Research Special Fund (Grant No. 320.6750.2020-03-13).

CRediT authorship contribution statement

Ping Hu: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Beining Li:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Zhenyu Yin:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Peng Peng:** Writing – review & editing, Validation, Methodology, Investigation. **Jiangang Cao:** Methodology, Investigation, Funding acquisition. **Wanyu Xie:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Liang Liu:** Methodology, Investigation. **Fujiang Cao:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Data curation. **Bin Zhang:** Writing – original draft, Supervision, Project

administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Thanks to GEO database for sharing the data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e30335.

References

[1] J.N. Katz, K.R. Arant, R.F. Loeser, Diagnosis and treatment of hip and knee osteoarthritis: a review, JAMA 325 (6) (2021) 568–578.
[2] S. Glyn-Jones, A.J.R. Palmer, R. Agricola, A.J. Price, T.L. Vincent, H. Weinans, A.J. Carr, Osteoarthritis. Lancet (London, England) 386 (9991) (2015) 376–387.
[3] J. Martel-Pelletier, A.J. Barr, F.M. Cicuttini, P.G. Conaghan, C. Cooper, M.B. Goldring, S.R. Goldring, G. Jones, A.J. Teichtahl, J.-P. Pelletier, Osteoarthritis, Nat. Rev. Dis. Prim. 2 (2016) 16072.
[4] D.J. Hunter, S. Bierma-Zeinstra, Osteoarthritis. Lancet 393 (10182) (2019) 1745–1759.
[5] Y. Wang, L. Chen, F. Li, M. Bao, J. Zeng, J. Xiang, H. Luo, J. Li, L. Tang, TLR4 rs41426344 increases susceptibility of rheumatoid arthritis (RA) and juvenile idiopathic arthritis (JIA) in a central south Chinese Han population, Pediatric rheumatology online journal 15 (1) (2017) 12.
[6] G.A. Hawker, Osteoarthritis is a serious disease, Clin. Exp. Rheumatol. 37 (5) (2019) 3–6. Suppl 120.
[7] S. Su, J. He, C. Wang, F. Gao, D. Zhong, P. Lei, A new dressing system reduces the number of dressing changes in the primary total knee arthroplasty: a randomized controlled trial, Frontiers in surgery 9 (2022) 800850.
[8] Z. Lv, X. Xu, Z. Sun, Y.X. Yang, H. Guo, J. Li, K. Sun, R. Wu, J. Xu, Q. Jiang, S. Ikegawa, D. Shi, TRPV1 alleviates osteoarthritis by inhibiting M1 macrophage polarization via Ca(2+)/CaMKII/Nrf2 signaling pathway, Cell Death Dis. 12 (6) (2021) 504.
[9] Y. Wu, Z. Wang, X. Fu, Z. Lin, K. Yu, Geraniol-mediated osteoarthritis improvement by down-regulating PI3K/Akt/NF-κB and MAPK signals: in vivo and in vitro studies, Int. Immunopharm. 86 (2020) 106713.
[10] S.H. Chang, D. Mori, H. Kobayashi, Y. Mori, H. Nakamoto, K. Okada, Y. Taniguchi, S. Sugita, F. Yano, U.-I. Chung, J.-R. Kim-Kaneyama, M. Yanagita, A. Economides, E. Canalis, D. Chen, S. Tanaka, T. Saito, Excessive mechanical loading promotes osteoarthritis through the gremlin-1-NF-κB pathway, Nat. Commun. 10 (1) (2019) 1442.
[11] M.F. Hsueh, X. Zhang, S.S. Wellman, M.P. Bolognesi, V.B. Kraus, Synergistic roles of macrophages and Neutrophils in osteoarthritis progression, Arthritis Rheumatol. 73 (1) (2021) 89–99.
[12] T.M. Griffin, C.R. Scanzello, Innate inflammation and synovial macrophages in osteoarthritis pathophysiology, Clin. Exp. Rheumatol. 37 (5) (2019) 57–63. Suppl 120.
[13] E. Sanchez-Lopez, R. Coras, A. Torres, N.E. Lane, M. Guma, Synovial inflammation in osteoarthritis progression, Nat. Rev. Rheumatol. 18 (5) (2022) 258–275.
[14] H. Zhang, D. Cai, X. Bai, Macrophages regulate the progression of osteoarthritis, Osteoarthritis Cartilage 28 (5) (2020) 555–561.
[15] H. Zhang, C. Lin, C. Zeng, Z. Wang, H. Wang, J. Lu, X. Liu, Y. Shao, C. Zhao, J. Pan, S. Xu, Y. Zhang, D. Xie, D. Cai, X. Bai, Synovial macrophage M1 polarisation exacerbates experimental osteoarthritis partially through R-spondin-2, Ann. Rheum. Dis. 77 (10) (2018) 1524–1534.
[16] T.L. Fernandes, A.H. Gomoll, C. Lattermann, A.J. Hernandez, D.F. Bueno, M.T. Amano, Macrophage: a potential target on cartilage regeneration, Front. Immunol. 11 (2020) 111.
[17] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. U. S. A. 102 (43) (2005) 15545–15550.
[18] A. Subramanian, H. Kuehn, J. Gould, P. Tamayo, J.P. Mesirov, GSEA-P: a desktop application for gene set enrichment analysis, Bioinformatics 23 (23) (2007) 3251–3253.
[19] T. Barrett, S.E. Wilhite, P. Ledoux, C. Evangelista, I.F. Kim, M. Tomashevsky, K.A. Marshall, K.H. Phillippy, P.M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C.L. Robertson, N. Serova, S. Davis, A. Soboleva, NCBI GEO: archive for functional genomics data sets–update, Nucleic Acids Res. 41 (Database issue) (2013) D991–D995.
[20] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47.
[21] D. Zeng, Z. Ye, R. Shen, G. Yu, J. Wu, Y. Xiong, R. Zhou, W. Qiu, N. Huang, L. Sun, X. Li, J. Bin, Y. Liao, M. Shi, W. Liao, IOBR: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures, Front. Immunol. 12 (2021) 687975.
[22] D. Aran, Z. Hu, A.J. Butte, xCell: digitally portraying the tissue cellular heterogeneity landscape, Genome Biol. 18 (1) (2017) 220.
[23] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinf. 9 (2008) 559.
[24] S. Fishilevich, S. Zimmerman, A. Kohn, T. Iny Stein, T. Olender, E. Kolker, M. Safran, D. Lancet, Genic insights from integrated human proteomics in GeneCards, Database 2016 (2016).
[25] C. Gene Ontology, Gene ontology consortium: going forward, Nucleic Acids Res. 43 (Database issue) (2015) D1049–D1056.
[26] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, K. Morishima, KEGG: new perspectives on genomes, pathways, diseases and drugs, Nucleic Acids Res. 45 (D1) (2017) D353–D361.
[27] L.M. Schriml, J.B. Munro, M. Schor, D. Olley, C. McCracken, V. Felix, J.A. Baron, R. Jackson, S.M. Bello, C. Bearer, R. Lichenstein, K. Bisordi, N.C. Dialo, M. Giglio, C. Greene, The human disease ontology 2022 update, Nucleic Acids Res. 50 (D1) (2022) D1255–D1261.
[28] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16 (5) (2012) 284–287.
[29] T. Berggard, S. Linse, P. James, Methods for the detection and analysis of protein-protein interactions, Proteomics 7 (16) (2007) 2833–2842.
[30] D. Szklarczyk, A.L. Gable, K.C. Nastou, D. Lyon, R. Kirsch, S. Pyysalo, N.T. Doncheva, M. Legeay, T. Fang, P. Bork, L.J. Jensen, C. von Mering, The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets, Nucleic Acids Res. 49 (D1) (2021) D605–D612.

[31] D. Otasek, J.H. Morris, J. Boucas, A.R. Pico, B. Demchak, Cytoscape Automation: empowering workflow-based network analysis, Genome Biol. 20 (1) (2019) 185.
[32] X. Zhou, D.P. Tuck, MSVM-RFE: extensions of SVM-RFE for multiclass gene selection on DNA microarray data, Bioinformatics 23 (9) (2007) 1106–1114.
[33] R. Alhamzawi, H.T.M. Ali, The Bayesian adaptive lasso regression, Math. Biosci. 303 (2018) 75–82.
[34] A. Liberzon, C. Birger, H. Thorvaldsdottir, M. Ghandi, J.P. Mesirov, P. Tamayo, The Molecular Signatures Database (MSigDB) hallmark gene set collection, Cell Syst 1 (6) (2015) 417–425.
[35] S. Hanzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7.
[36] G. Zhou, O. Soufan, J. Ewald, R.E.W. Hancock, N. Basu, J. Xia, NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis, Nucleic Acids Res. 47 (W1) (2019) W234–W241.
[37] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, M. Müller, pROC: an open-source package for R and S+ to analyze and compare ROC curves, BMC Bioinf. 12 (2011) 77.
[38] L.A. Mandl, Osteoarthritis year in review 2018: clinical, Osteoarthritis Cartilage 27 (3) (2019) 359–364.
[39] C. Lietman, B. Wu, S. Lechner, A. Shinar, M. Sehgal, E. Rossomacha, P. Datta, A. Sharma, R. Gandhi, M. Kapoor, P.P. Young, Inhibition of Wnt/beta-catenin signaling ameliorates osteoarthritis in a murine model of experimental osteoarthritis, JCI Insight 3 (3) (2018).
[40] T. Hu, Z. Zhang, C. Deng, X. Ma, X. Liu, Effects of beta 2 integrins on osteoclasts, macrophages, chondrocytes, and synovial fibroblasts in osteoarthritis, Biomolecules 12 (11) (2022).
[41] E. Jimi, H. Fei, C. Nakatomi, NF-kappaB signaling regulates physiological and pathological chondrogenesis, Int. J. Mol. Sci. 20 (24) (2019).
[42] P. Lepetsos, K.A. Papavassiliou, A.G. Papavassiliou, Redox and NF-kappaB signaling in osteoarthritis, Free Radic. Biol. Med. 132 (2019) 90–100.
[43] L. Ni, Z. Lin, S. Hu, Y. Shi, Z. Jiang, J. Zhao, Y. Zhou, Y. Wu, N. Tian, L. Sun, A. Wu, Z. Pan, X. Zhang, X. Wang, Itaconate attenuates osteoarthritis by inhibiting STING/NF-kappaB axis in chondrocytes and promoting M2 polarization in macrophages, Biochem. Pharmacol. 198 (2022) 114935.
[44] V. Gkretsi, T. Simopoulou, A. Tsezou, Lipid metabolism and osteoarthritis: lessons from atherosclerosis, Prog. Lipid Res. 50 (2) (2011) 133–140.
[45] S. Kotake, N. Udagawa, N. Takahashi, K. Matsuzaki, K. Itoh, S. Ishiyama, S. Saito, K. Inoue, N. Kamatani, M.T. Gillespie, T.J. Martin, T. Suda, IL-17 in synovial fluids from patients with rheumatoid arthritis is a potent stimulator of osteoclastogenesis, J. Clin. Invest. 103 (9) (1999) 1345–1352.
[46] P.J. Mease, Inhibition of interleukin-17, interleukin-23 and the TH17 cell pathway in the treatment of psoriatic arthritis and psoriasis, Curr. Opin. Rheumatol. 27 (2) (2015) 127–133.
[47] S.K. Raychaudhuri, A. Saxena, S.P. Raychaudhuri, Role of IL-17 in the pathogenesis of psoriatic arthritis and axial spondyloarthritis, Clin. Rheumatol. 34 (6) (2015) 1019–1023.
[48] F.J. Pixley, E.R. Stanley, CSF-1 regulation of the wandering macrophage: complexity in action, Trends Cell Biol. 14 (11) (2004) 628–638.
[49] M.L. Toh, J.Y. Bonnefoy, N. Accart, S. Cochin, S. Pohle, H. Haegel, M. De Meyer, C. Zemmour, X. Preville, C. Guillen, C. Thioudellet, P. Ancian, A. Lux, B. Sehnert, F. Nimmerjahn, R.E. Voll, G. Schett, Bone- and cartilage-protective effects of a monoclonal antibody against colony-stimulating factor 1 receptor in experimental arthritis, Arthritis Rheumatol. 66 (11) (2014) 2989–3000.
[50] S. Onuora, Experimental arthritis: antibody against CSF-1 receptor protects bone and cartilage, Nat. Rev. Rheumatol. 10 (5) (2014) 260.
[51] W. Lin, D. Xu, C.D. Austin, P. Caplazi, K. Senger, Y. Sun, S. Jeet, J. Young, D. Delarosa, E. Suto, Z. Huang, J. Zhang, D. Yan, C. Corzo, K. Barck, S. Rajan, C. Looney, V. Gandham, J. Lesch, W.C. Liang, E. Mai, H. Ngu, N. Ratti, Y. Chen, D. Misner, T. Lin, D. Danilenko, P. Katavolos, E. Doudemont, H. Uppal, J. Eastham, J. Mak, P.E. de Almeida, K. Bao, A. Hadadianpour, M. Keir, R.A.D. Carano, L. Diehl, M. Xu, Y. Wu, R.M. Weimer, J. DeVoss, W.P. Lee, M. Balazs, K. Walsh, K.R. Alatsis, F. Martin, A.A. Zarrin, Function of CSF1 and IL34 in macrophage homeostasis, inflammation, and cancer, Front. Immunol. 10 (2019) 2019.
[52] S.H. Mun, P.S.U. Park, K.H. Park-Min, The M-CSF receptor in osteoclasts and beyond, Exp. Mol. Med. 52 (8) (2020) 1239–1254.
[53] S. Garcia, L.M. Hartkamp, B. Malvar-Fernandez, I.E. van Es, H. Lin, J. Wong, L. Long, J.A. Zanghi, A.L. Rankin, E.L. Masteller, B.R. Wong, T.R. Radstake, P. P. Tak, K.A. Reedquist, Colony-stimulating factor (CSF) 1 receptor blockade reduces inflammation in human and murine models of rheumatoid arthritis, Arthritis Res. Ther. 18 (2016) 75.
[54] E. Ferretti, V. Pistoia, A. Corcione, Role of fractalkine/CX3CL1 and its receptor in the pathogenesis of inflammatory and malignant diseases with emphasis on B cell malignancies, Mediat. Inflamm. 2014 (2014) 480941.
[55] T. Imai, N. Yasuda, Therapeutic intervention of inflammatory/immune diseases by inhibition of the fractalkine (CX3CL1)-CX3CR1 pathway, Inflamm. Regen. 36 (2016) 9.
[56] W.H. Yuan, Q.Q. Xie, K.P. Wang, W. Shen, X.F. Feng, Z. Liu, J.T. Shi, X.B. Zhang, K. Zhang, Y.J. Deng, H.Y. Zhou, Screening of osteoarthritis diagnostic markers based on immune-related genes and immune infiltration, Sci. Rep. 11 (1) (2021) 7032.
[57] S. Culemann, A. Gruneboom, J.A. Nicolas-Avila, D. Weidner, K.F. Lammle, T. Rothe, J.A. Quintana, P. Kirchner, B. Krljanac, M. Eberhardt, F. Ferrazzi, E. Kretzschmar, M. Schicht, K. Fischer, K. Gelse, M. Faas, R. Pfeifle, J.A. Ackermann, M. Pachowsky, N. Renner, D. Simon, R.F. Haseloff, A.B. Ekici, T. Bauerle, I. E. Blasig, J. Vera, D. Voehringer, A. Kleyer, F. Paulsen, G. Schett, A. Hidalgo, G. Kronke, Locally renewing resident synovial macrophages provide a protective barrier for the joint, Nature 572 (7771) (2019) 670–675.
[58] M. Kurowska-Stolarska, S. Alivernini, Synovial tissue macrophages in joint homeostasis, rheumatoid arthritis and disease remission, Nat. Rev. Rheumatol. 18 (7) (2022) 384–397.
[59] M. Hirata, F. Kugimiya, A. Fukai, T. Saito, F. Yano, T. Ikeda, A. Mabuchi, B.R. Sapkota, T. Akune, N. Nishida, N. Yoshimura, T. Nakagawa, K. Tokunaga, K. Nakamura, U.I. Chung, H. Kawaguchi, C/EBPbeta and RUNX2 cooperate to degrade cartilage with MMP-13 as the target and HIF-2 alpha as the inducer in chondrocytes, Hum. Mol. Genet. 21 (5) (2012) 1111–1123.
[60] D.M. Lamkin, S. Srivastava, K.P. Bradshaw, J.E. Betz, K.B. Muy, A.M. Wiese, S.K. Yee, R.M. Waggoner, J.M.G. Arevalo, A.J. Yoon, K.F. Faull, E.K. Sloan, S. W. Cole, C/EBPβ regulates the M2 transcriptome in β-adrenergic-stimulated macrophages, Brain Behav. Immun. 80 (2019) 839–848.
[61] G. Li, Y. Sun, I. Kwok, L. Yang, W. Wen, P. Huang, M. Wu, J. Li, Z. Huang, Z. Liu, S. He, W. Peng, J.X. Bei, F. Ginhoux, L.G. Ng, Y. Zhang, Cebp1 and Cebpβ transcriptional axis controls eosinophilopoiesis in zebrafish, Nat. Commun. 15 (1) (2024) 811.
[62] X. Xi, A. Mehmood, P. Niu, J. Yang, Y. Wang, H. Zhou, X. Han, L. Ma, S. Jin, Y. Wu, Association of X-linked TLR-7 gene polymorphism with the risk of knee osteoarthritis: a case-control study, Sci. Rep. 12 (1) (2022) 7243.
[63] Y. Zhang, Z. Feng, J. Liu, H. Li, Q. Su, J. Zhang, P. Huang, W. Wang, J. Liu, Polarization of tumor-associated macrophages by TLR7/8 conjugated radiosensitive peptide hydrogel for overcoming tumor radioresistance, Bioact. Mater. 16 (2022) 359–371.
[64] N. Hoshikawa, A. Sakai, S. Takai, H. Suzuki, Targeting extracellular miR-21-TLR7 signaling provides long-lasting analgesia in osteoarthritis, Mol. Ther. Nucleic Acids 19 (2020) 199–207.