Electronic Journal of Biotechnology 65 (2023) 24–44

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

journal homepage: www.elsevier.com/locate/ejbt

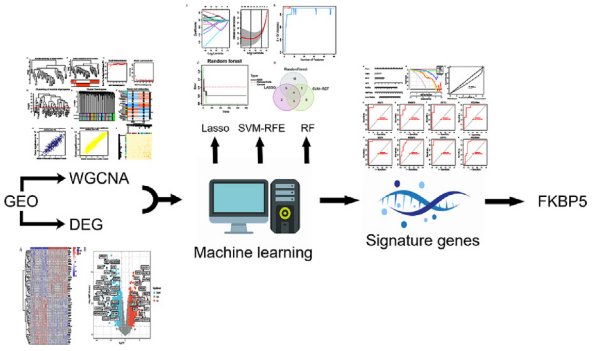
Research Article

Integrated analysis reveals potential significance of FKBP5 in the prognosis and immunity of osteoarthritis and pan-cancer

Yueying Xiao [a], Yanan Wang [b], Xiao Xu [c], Xiejia Jiao [a], Yanqing Huo [a,*]

[a] Department of Spine Surgery, The Second Hospital of Shandong University, 250033, Jinan, China
[b] Key Laboratory of Experimental Teratology of Ministry of Education, Institute of Medical Sciences, The Second Hospital of Shandong University, 250033, Jinan, China
[c] Sterile Supply Department, The First People Hospital of Jinan, 250033, Jinan, China

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 24 February 2023
Accepted 25 May 2023
Available online 11 July 2023

Keywords:
Diagnostic markers
Differentially expressed genes (DEGs)
Drug response
EPYC
FKBP5
Immune infiltration
Immune
Machine learning
Osteoarthritis
Pan-cancer

ABSTRACT

Background: The high disability rate of osteoarthritis (OA), a joint disease with an insidious onset and widespread effects, places a heavy financial burden on patients, families, and society. Traditional diagnostic approaches, including radiology and physical examination, cannot achieve early-stage screening of OA and thus, miss early intervention for patients. Therefore, the need of biomarkers for the early diagnosis of OA is crucial.

Results: A total of 390 differentially expressed genes (DEGs) were identified from the training set, and 1077 key module genes were found by constructing a weighted gene co-expression network, and 161 key genes were obtained as a result. Four diagnostic marker genes highly associated with OA were screened for key genes using machine learning algorithms, and the resulting nomogram model showed excellent predictive power and clinical value. After further background studies, immune infiltration and functional enrichment analysis, we found that FKBP5 may play an important role in the prognosis and immune infiltration of multiple cancers, and this hypothesis was verified by pan-cancer analysis.

Conclusions: We screened four diagnostic marker genes (FKBP5, EPYC, KLF9 and PDZRN4) that are highly associated with OA. And this led to a diagnostic model, which was assessed to have good predictive power and clinical value. FKBP5 may be a potential intervention target for human diseases such as osteoarthritis and tumors.

Peer review under responsibility of Pontificia Universidad Católica de Valparaíso

* Corresponding author.
E-mail address: 200462011187@email.sdu.edu.cn (Y. Huo).

https://doi.org/10.1016/j.ejbt.2023.05.002

How to cite: Xiao Y, Wang Y, Xu X, et al. Integrated analysis reveals potential significance of FKBP5 in the prognosis and immunity of osteoarthritis and pan-cancer. Electron J Biotechnol 2023;65. https://doi.org/10.1016/j.ejbt.2023.05.002.

1. Introduction

Osteoarthritis (OA) is a common joint disease characterized by changes in bone structure and loss of articular cartilage. The pathological changes mainly include degeneration of the articular cartilage, subchondral osteophytes, synovial inflammation, and hypertrophy of the meniscus and joint capsule. Previously, OA was referred to as a degenerative joint disease, but growing research suggests that this is an inaccurate term because OA is not only due to wear and tear of the joint but also due to many inflammatory mediators that cause tissue remodeling within the joint [1]. Risk factors of OA can be divided into two categories: systemic factors (age, sex, obesity, genetics, and diet) and factors associated with OA progression due to the joint itself (injury, poor joint alignment, and abnormal joint loading) [2]. With an aging and increasingly obese population, OA is becoming a growing personal and societal health burden. Approximately 250 million people are thought to be affected by this illness worldwide, and in this context, most patients do not receive appropriate management and treatment [3]. The current diagnostic criteria for OA are based on clinical symptoms (pain, transient pain <30 min, and swelling), physical examination (popping, restricted movement, joint compression, and osteophytes), and imaging results of the bone and joint, which have limitations and lags; irreversible damage has already occurred by the time many diagnoses are made. Damage to joint structures detected by X-ray and magnetic resonance imaging (MRI) precedes the development of pain [4], and the advent of MRI has led to the inclusion of statements of nonchondral tissue of joints in many patients with OA [5]. However, MRI is not used for the early detection of OA; therefore, there is an urgent need to identify more sensitive modalities capable of early diagnosis.

In the development of OA, metabolic changes in joints, cartilage, and synovium precede morphological changes in joint structures. Biomarkers that reflect metabolic changes in joints, cartilage, and synovium are gaining attention, including collagen and noncollagen biomarkers, and inflammatory and anti-inflammatory biomarkers, which are expressed in biological fluids, such as serum, synovial fluid, and urine [6]. The expression levels of these markers can directly reflect the pathophysiological changes in OA and can be used for disease assessment and treatment efficacy determination.

Immune mechanisms are key drivers of OA progression and have become a hot research topic in early diagnostic markers of OA in recent years [7]. Among them, macrophages are the most important cells in the synovial tissue [6], producing glycosylases, matrix metalloproteinases, and other destructive mediators that activate innate immunity in OA, suggesting that macrophages play an important role in the pathology of OA [8,9]. Macrophages and other immune cells can activate and regulate the complement system [10], which, as part of inflammation and the immune system, promote the immune response by enhancing antibodies and immune cells against antigens [10], responsible for catabolism and anabolism in the synovial joints. Furthermore, growing evidence indicates that T helper (Th) cells play a key role in the pathogenesis of OA [11], particularly in the inflammatory response of joints triggered by synovial infiltration; however, the clinical relevance of synovial and serum Th infiltration remains unclear [12]. Previous bioinformatics studies have demonstrated how immune cell infiltration and immunological-related pathways contribute to the development of OA [13,14]. These results highlight the crucial role of immunological processes in OA; however, the fundamental immunological molecular pathways remain poorly understood.

Gene coexpression networks can synthesize multiple levels and functional units in biological systems while studying their interactions, providing a more systematic and systemic research paradigm for biological studies [15]. In particular, the use of weighted gene coexpression network analysis (WGCNA) on high-throughput microarray data has produced significant results in the genetic analysis of numerous species, including humans and mice. Currently, machine-learning algorithms have been applied to many real-world bioinformatics problems, and many effective learning algorithms have been developed to identify specific disease-causing genes based on gene chip sequences [16].

With the advent of new technologies in biomedicine, there has been an exponential increase in high-throughput disease-related data. However, accurate prediction of disease progression is one of the most challenging tasks in the medical community. Machine-learning algorithms have become an important analytical tool in the medical field because of their ability to identify the key factors among a vast amount of data. Machine learning applications in the medical field include medical image analysis, disease prediction and genomics [17,18]. These applications require processing large amounts of complex data, including medical images, genomic data, patient history data, and more. Machine learning can help doctors better understand and analyze these data to improve disease prediction and treatment accuracy and efficiency. Among them, medical image analysis can help doctors automatically detect and diagnose diseases; disease prediction can be based on a large amount of patient data to predict the occurrence and progression of diseases and analyze the risk of patients; and genomics can be used to identify genes and genetic variants associated with diseases. Different machine learning algorithms are applicable to different data types and problem types, so a medical problem may require multiple machine learning algorithms to solve. Therefore, we used the three most widely used machine-learning algorithms in the biomedical field: least absolute shrinkage and selection operator (LASSO), random forest, and support vector machine recursive feature elimination (SVM-RFE). The LASSO algorithm has a strong feature selection capability, which can select important features and suppress the influence of irrelevant features to improve the accuracy and interpretability of the model. The RF algorithm has good generalization performance, can handle a large number of features, and can handle nonlinear and high-dimensional datasets. The SVM-RFE algorithm obtains the best subset of features by gradually eliminating unimportant features. The SVM-RFE algorithm improves the robustness and generalization of the model and can be applied to high-dimensional and nonlinear datasets. The three machine learning algorithms have their advantages and applications in specific problems and scenarios, and the combined use of the three algorithms can improve the accuracy and generalization ability of the model. We used microarray data from normal people and patients with OA to screen for

signature genes associated with the early diagnosis of OA and to lay the foundation for identifying potential biomarkers.

2. Materials and Methods

2.1. Data download and preprocessing

Gene expression profiles for GSE55235 and GSE12021 [19,20] were downloaded from the gene expression omnibus (GEO), and data for 39 human groups, including the normal group (n = 19) and the OA group (n = 20), were available. Gene expression profiles for GSE55457 [21] were also downloaded from GEO. Twenty samples of data comprising the human normal group (n = 10) and human OA group (n = 10) were used for subsequent model validation. Owing to the uniformity of the specimen source, study design, and platform source (GPL96) between the GSE55235 and GSE12021 datasets, the expression profiles of these two datasets were merged. The combat function of the sva package in R was used to eliminate the batch effect [22], where the probe name of each gene was changed to its symbol name, and the probes corresponding to multiple gene symbols were removed. The cancer dataset was derived from human tissue sequencing data from The Cancer Genome Atlas (TCGA); tumor cell line sequencing data were derived from the Cancer Cell Line Encyclopedia (CCLE) of the Broad Institute; and the Genotype Tissue Expression (GTEx) project was used to collect data from normal tissues. We used the XENA (https://xenabrowser.net) [23] database to obtain bioinformatic data including The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx), and performed a pan-cancer bioinformatic analysis.

2.2. Differential expression analysis

The limma package in R was used to analyze the combined expression matrix [24], and the results of the differentially expressed genes (DEGs) were displayed in a volcano plot. DEGs were screened in osteoarthritic and normal samples, and genes with |fold change (FC)| < 0.5 and multiple test correction values (q-values) <0.05 were considered to be differentially expressed in OA.

2.3. Weighted Gene Coexpression Network Analysis (WGCNA)

The combined GSE55235 and GSE12021 datasets were used to uncover strongly associated gene coexpression modules using the WGCNA algorithm. The distance between each gene was first calculated using the Pearson correlation coefficient, and then, a weighted gene coexpression network was constructed using the WGCNA package in R [25,26] with a soft threshold set to 6. To investigate the modules highly correlated with OA, we calculated the correlation coefficients between the module feature genes and OA. The higher the correlation coefficient, the more strongly the modules and OA were correlated. We then selected the two gene modules with the largest and smallest correlation coefficients, selected all the genes within them, and intersected with DEGs. The intersected genes were used for subsequent feature gene screening. We measured the importance of genes in the module by calculating the gene significance (GS) and module membership (MM).

2.4. Functional enrichment analysis

Functional analyses, including gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses, were performed using the clusterProfiler package [27]. Significant enrichment in this gene collection was indicated by an adjusted p-value <0.05.

2.5. Selection of feature genes

Three machine-learning algorithms were used to screen the feature genes. LASSO is a dimensionality reduction analysis method for performing gene selection [28] that compresses the coefficients of the target genes and makes certain regression coefficients zero, leading to the screening of feature genes. In this study, the glmnet package was used to determine the penalty parameter lambda by 10 × cross-verification (10-fold cross-verification) [29], and to determine the optimal lambda value corresponding to the minimum value of the cross-verification error mean. Random forest [30] is an algorithm that integrates multiple decision trees through an integrated-learning approach. It uses the R package RandomForest to construct a random forest classifier and predict disease classification by feature genes; the top 30 genes in importance are considered as feature genes. SVM-RFE [31] is a sequence backward selection algorithm. The model was trained on samples, and after ranking each feature according to its score, the component with the lowest rating was eliminated. The model was then retrained with the remaining features in preparation for the subsequent iteration, and finally, the correct number of features was selected. To identify the most powerful genes associated with OA diagnosis, appropriate feature genes were selected using the SVM-RFE method. The diagnostic effect was assessed using the receiver operating characteristic (ROC) curve and area under the curve (AUC). The combined GSE55235 and GSE12021 gene sets were used as the training set, and the GSE55457 and GSE12807 gene sets were used as the validation set.

2.6. Gene set enrichment analysis (GSEA)

Based on the median expression levels of the four feature genes, OA tissues were divided into two groups using GSEA, and the gene sets within the two groups were enriched for certain gene functions. The "C2.KEGG pathway gene set" was chosen as the reference gene collection. To obtain standardized enrichment scores for each analysis, 1000 permutation tests were performed. Statistical significance was defined as p < 0.05.

2.7. ROC curve plotting

The "pROC" package in R was applied to draw ROC curves for four disease signature genes, and the diagnostic performance of the disease signature genes was determined based on the AUC values.

2.8. Establishment of a nomogram

Feature genes were merged using the rms package to create the nomogram. Calibration curves were used to assess nomogram accuracy. The clinical utility of Norman plots was assessed using decision curves.

2.9. Single sample GSEA (ssGSEA)

Using 45 gene sets collected from the molecular signatures database [32] as the reference set, ssGSEA was performed using the gene set variation analysis (GSVA) R package [33] to obtain GSVA scores for each gene set. The difference in GSVA scores for each gene set between the osteoarthritic and normal tissues was compared using the limma package, with GSVA values indicating the absolute enrichment of each gene set. We calculated the spearman's correlation coefficients between GSVA scores and gene

expression using R and applied the "ggplot2" package to demonstrate them.

2.10. Prediction and signature gene correlation of immune infiltrating cells

The percentage of tumor-infiltrating immune-related cells in each sample was predicted using the CIBERSORT algorithm. Differential analysis of immune cells was performed using the R software packages "corrplot" and "vioplot." Correlations between immune cells and diagnostic features were analyzed using the spearman's test, and the results were visualized using R software. We used data from 39 samples from GSE55235 and GSE12021 in performing immune infiltrating cell prediction and signature gene correlation analysis, containing 19 samples from the normal group and 20 samples from the osteoarthritis group.

2.11. Relationship between survival analysis and clinical stage

The GEPIA database is an online platform that contains RNA sequencing data for tumor and normal tissues from the TCGA and GTEx datasets [34]. Using the survival plot module, the database was used to investigate the association between FKBP5 expression and tumor prognosis.

2.12. Cox regression analysis and survival analysis

The normalized TCGA dataset, including expression and prognostic information, was downloaded from the UCSC (https://xenabrowser.net/) database. The gene expression data for FKBP5 were further extracted and log2(x+1) was transformed to build a Cox proportional hazards regression model [35] to examine the association between gene expression and prognosis in each tumor (including parameters of overall patient survival and disease-free survival). Patients were divided into high- and low-expression groups according to the best cut-off values, and the prognostic differences between the two groups were analyzed using the survfit function of the R software survival package. The significance of the differences was assessed using the log-rank test method, and the p-value <0.05 was considered statistically significant.

2.13. Multiple algorithms for immunoinfiltration analysis

We integrated five algorithms, TIMER [36], xCell [37], EPIC [38], MCPcounter [39], and quanTIseq [40], to calculate the immune cell infiltration scores for several major immune cell types to determine whether there is a link between FKBP5 expression and immune infiltration.

3. Results

3.1. Data preprocessing and screening of DEGs

In this study, the GSE55235 and GSE12021 datasets were batch-corrected using the sva package and the DEGs of the combined datasets were identified using the limma package. A total of 390 DEGs were screened out, including 184 upregulated genes and 206 downregulated genes, according to the critical criteria of |log2FC|>1 and p<0.05. The DEGs of the combined dataset are displayed with a volcano plot in Fig. 1B; the top 30 upregulated and downregulated DEGs are displayed on the heat map in Fig. 1A.

3.2. Building coexpression modules through WGCNA

The expression matrix of the 39 patients in the combined dataset was read in R. weighted gene coexpression analysis was used for the two groups of samples, and the samples were clustered using the class average method (average) in the hclust function. The clustering results are shown i n Fig. 2A. The disease grouping data of the patients were imported, and the relationship between the sample grouping and the sample dendrogram was visualized, as shown in Fig. 2B. The construction of the weighted gene coexpression network required a soft threshold β to calculate the adjacency matrix weight parameter, and the optimal soft threshold was 6 using the pickSoftThreshold function; the selection process is shown in Fig. 2C, D, and E. Further construction of the weighted

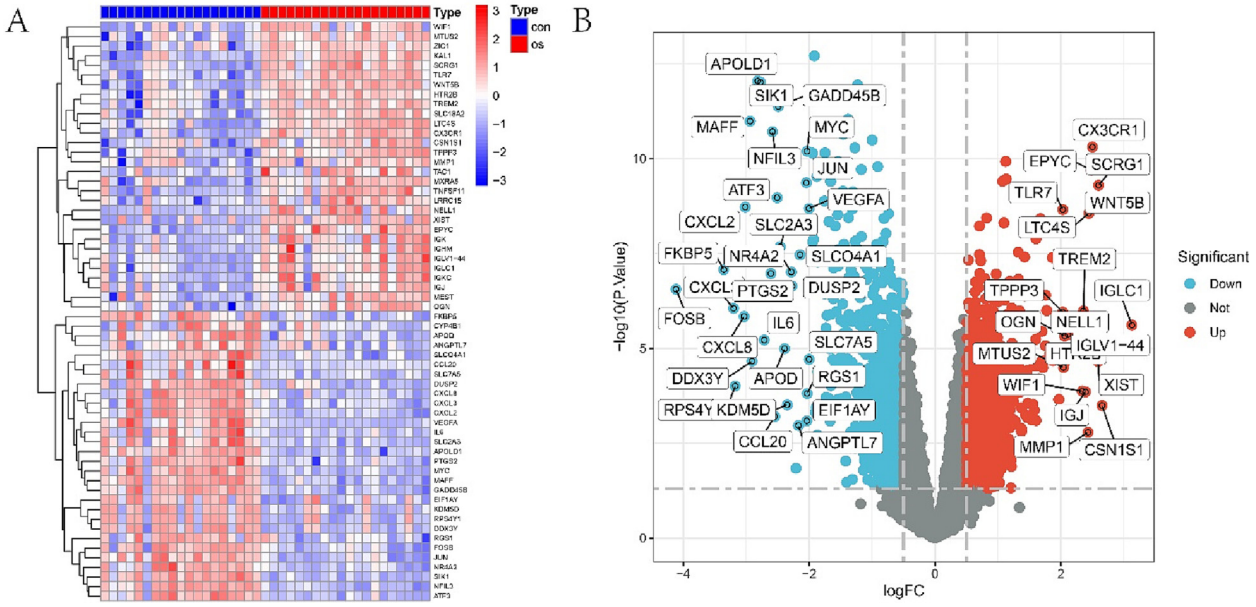


Fig. 1. Data processing of DEGs. (A) Heatmap of normal group and OA group. (B) The GSE55235 and GSE12021 datasets' distribution of all differentially expressed genes is shown in a volcano plot. The 186 genes that were down-regulated (light blue dots) and 184 that were up-regulated (red dots) are mapped. Black dots represent genes that have undergone no substantial alterations.

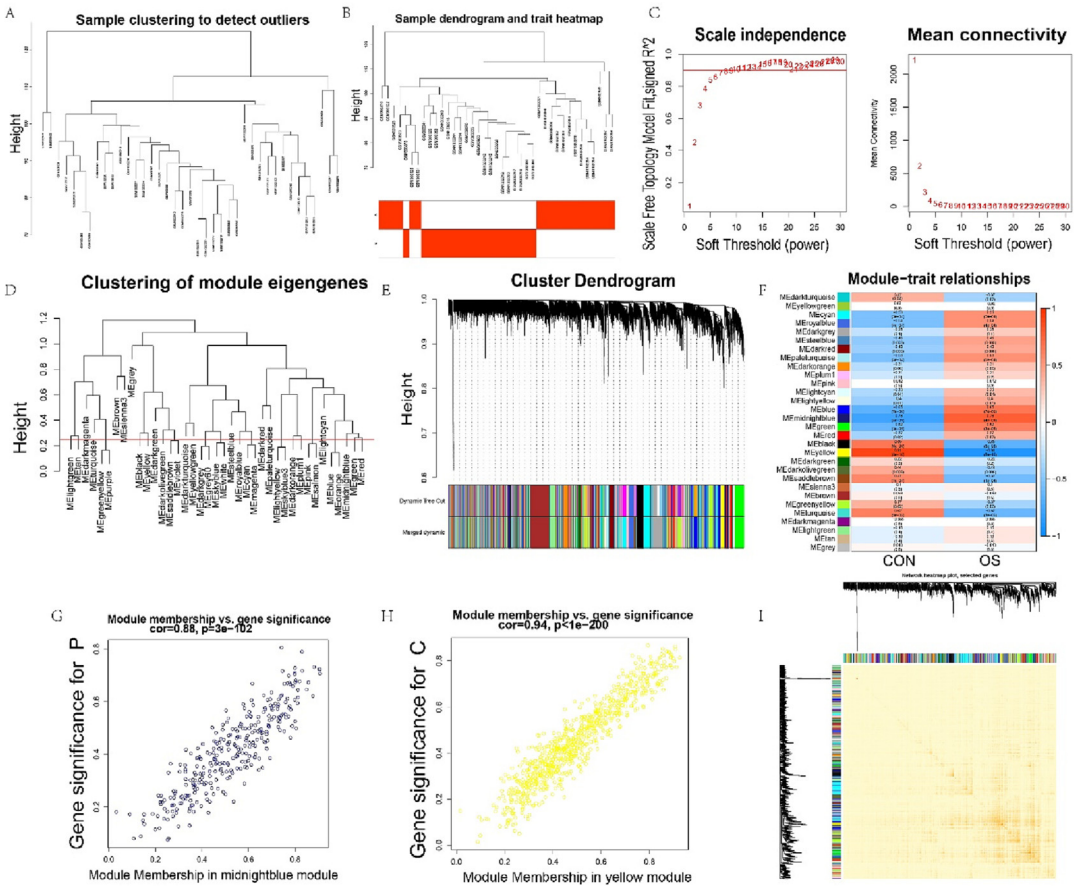


Fig. 2. Gene co-expression network construction by WGCNA. (A-B) Gene Module Clustering. (C) Selection of co-expressed gene network parameters. The optimal soft threshold is mainly referred to the left figure, which is the scale-free fit index (y-axis) under different soft thresholds (x-axis). The red line indicates the subjectively selected scale-free fit index, which is 0.9. From the left figure, when the scale-free fit index is 0.9, the minimum soft threshold for constructing scale-free network is 6, so we can choose 6 as the optimal soft threshold for subsequent analysis. The right figure shows the network connectivity with different soft thresholds. (D) Cut the clustering dendrogram at a height of 0.23 to find and combine related modules. (E) Delineation of gene co-expression networks. The top half is a hierarchical clustering tree diagram of genes, and the bottom half is the gene modules, or network modules. The top and bottom correspond to each other, and you can see that genes that are close together (clustered to the same branch) are divided into the same module. (F) Module-trait correlation heat map. The leftmost color block represents the module, and the rightmost color bar represents the range of correlations. In the middle part of the heat map, the darker the color the higher the correlation, red indicates positive correlation and green indicates negative correlation; the numbers in each cell indicate correlation and significance. (G-H) The relationship between the midnightblue module, yellow module and OA is shown in a scatter plot. The key modules midnightblue, yellow and the traits of interest were further mined to see if there was some correlation between the gene and module correlation (Module Membership, MM) and the gene and trait correlation (Gene Significance, GS). The above scatter plot shows that MM and GS are positively correlated, indicating that these genes, which are highly correlated with traits, also play a pivotal role in the key module. (I) Clustering dendrogram of module feature genes. Based on the topological overlap matrix, a heat map of the correlations between genes can be drawn. Where the darker the color, the stronger the interaction between genes.

gene network was based on TOM, the minimum module size was set to 50, and a total of 30 gene modules were detected. The module identification results are shown in Fig. 2F. Module-trait correlation analysis showed that, in terms of disease characteristics, the yellow module (765) showed the strongest negative correlation (cor = -0.91, 8e-16), and the midnight-blue module (312) showed the strongest positive correlation (cor = 0.78, 4e-09); the results are shown in Fig. 2G.

3.3. Acquisition of hub genes and GO functional enrichment analysis and KEGG pathway enrichment analysis

A total of 1,077 key module genes were intersected with 390 DEGs to obtain 161 hub genes, and a Venn diagram was drawn. As shown in Fig. 3A, KEGG pathway enrichment analysis (Table S1) and GO function enrichment analysis (Table S2) were performed on the hub genes, which enriched 810 biological processes, 62 molecular functions, 21 cellular components, and 42 pathways ($p < 0.05$). For BPs, the top 10 enrichment items were mainly involved in the response to peptides, cellular response to external stimuli, negative regulation of immune system processes, muscle cell proliferation, and other biological processes (Fig. 3B). The top 10 enrichment items for CCs and MFs are shown in Fig. 3B. According to the findings of the KEGG pathway enrichment study, the hub genes were primarily involved in cellular metabolic processes, including osteoblast differentiation and the signaling pathways of tumor necrosis factor (TNF), interleukin (IL)-17, p53, and JAK-STAT (Fig. 3C).

3.4. Screening for signature biomarkers through a comprehensive strategy

First, we applied the LASSO regression model for variable screening. As shown in Fig. 4A, each color curve in the first plot represents the trajectory of each independent variable coefficient. The later the coefficient is compressed to zero, the more important the variable becomes as the lambda value changes. The vertical coordinate in the second plot is the binomial deviance (dichotomous anomaly), which represents the magnitude of the error of the model. Our goal was to select the model with the lowest possible lambda characteristics and the smallest possible error. Lambda.min obtained the smallest mean value of the target covari-

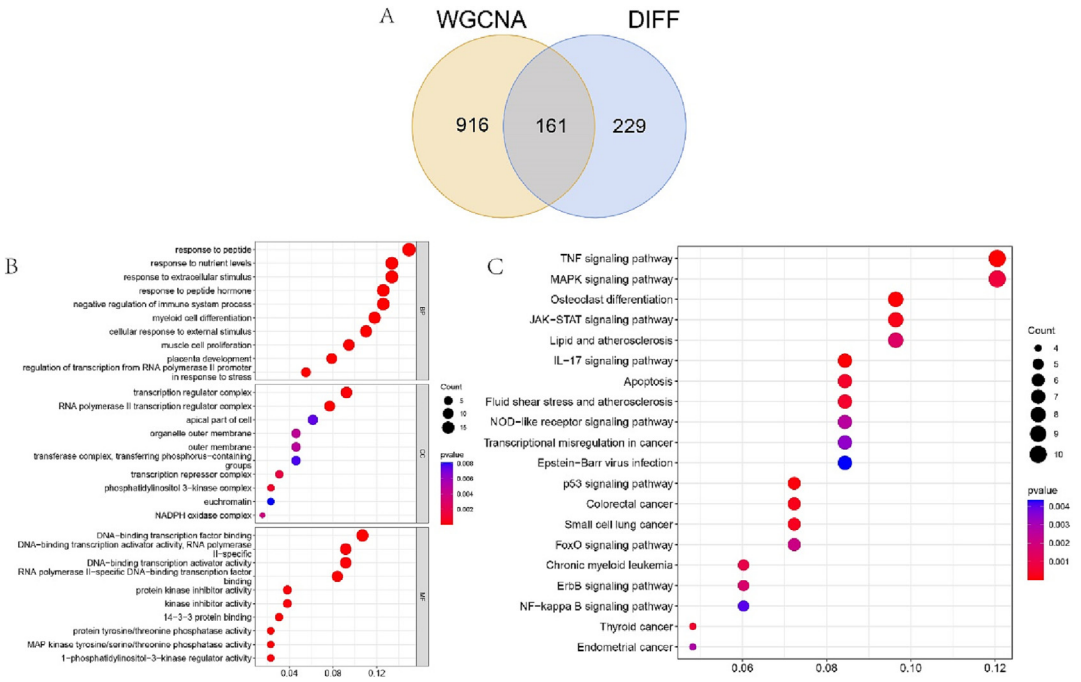


Fig. 3. Functional analysis of key module genes merged with DEGs. Venn diagram of key module genes and DEGs (A). GO (B) and KEGG (C) enrichment analysis of hub genes in the OA group.

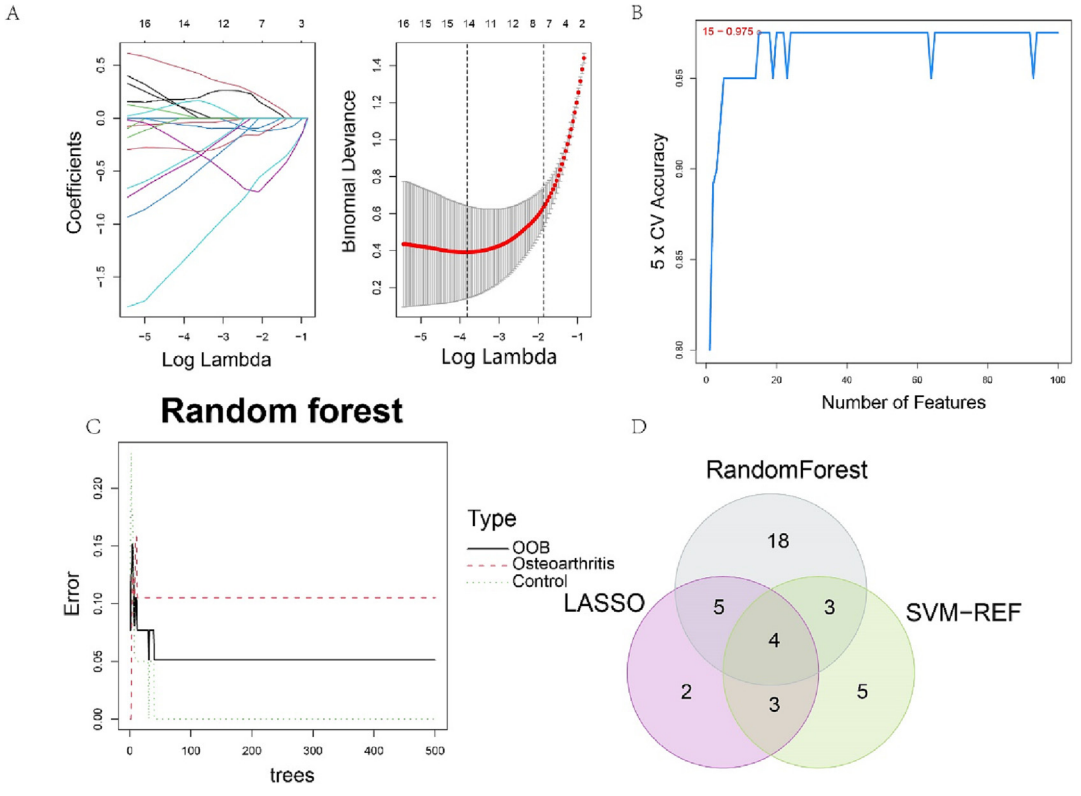


Fig. 4. Screening for signature genes. (A) Adjustment of parameter selection and LASSO coefficient curves for hub genes in the LASSO model. In the graph on the left, a line represents a gene, and the end of these genes will point to a vertical coordinate, which is the coefficient, and LASSO will count a coefficient for each gene. In the graph on the right, there are two dashed lines, and the parameter corresponding to the one on the left is called lambda.min. As you can see from the graph, the right graph corresponds to a value of 15 on the left, indicating that there are 15 coefficients for the genes that can stay. (B) Biological marker screening via the support vector machine recursive feature elimination (SVM-RFE) arithmetic. The horizontal coordinate Number of Features represents the number of features; then, the vertical coordinate 5 X CV Accuracy, which represents the accuracy of curve change after 5 times cross-validation. The Fig. 15−0.975 means there are 15 features with an accuracy of 0.975. The closer this accuracy is to 1, the higher the accuracy is. (C) randomForest error rate versus the number of classification trees. (D) The intersection genes obtained by the three algorithms. The horizontal coordinates represent the number of trees based on the random forest model, respectively, and the vertical coordinates represent the error rate of the corresponding model, using the number of trees corresponding to the point with the lowest error rate.

ates among all lambda values. The optimal lambda was determined to be 0.022 after 10-fold cross-validation. Fifteen genes were identified as feature genes of OA among the 161 target genes (Fig. 4A; Table S3). In the SVM-RFE algorithm, the classifier error was minimized when the number of features was 15 (Fig. 4B; Table S4). In addition, 161 target genes were used to construct the RF model (Fig. 4C; Table S5), and the top 30 genes were ranked according to their relative importance and identified as characteristic genes. Fig. 4C shows the performance of the random forest algorithm. By applying the gene set, we built the random forest model separately according to the different number of trees and calculated the error rate of this model. The number of trees corresponding to the point with the lowest error rate was adopted. The genes obtained by the three algorithms were intersected, and four final diagnostic signature genes were obtained: FKBP5, EPYC, KLF9, and PDZRN4 (Fig. 4D).

3.5. Diagnostic column line graphs and ROC curves

A nomogram (Fig. 5A) was constructed as a diagnostic tool for OA progression by combining four diagnostic signature genes (FKBP5, EPYC, KLF9, and PDZRN4). In the nomogram, a score was given to the expression of each gene, and a total score could be calculated for each patient. This total score corresponded to the probability of OA occurrence in each patient. Calibration curves and DCA curves were applied to the dataset (GSE55235, GSE12021, and GSE55457) to construct the predictive ability of the nomogram. The calibration curves (Fig. 5B) indicated that the nomogram could predict the occurrence of OA more accurately. DCA curves showed that the patients diagnosed with OA could benefit from the clinical treatment (Fig. 5C). The ROC curves of the genes (Fig. 5D) indicated that they all have the potential to be new diagnostic markers with AUC values of 0.989 (EPYC), 0.916 (FKBP5), 0.992 (KLF9), and 0.874 (PDZRN4). The validation set (GSE55457) could be similarly confirmed (Fig. 5E), with AUC values of 0.970 (EPYC), 0.940 (FKBP5), 0.980 (KLF9), and 0.930 (PDZRN4).

3.6. Signaling pathways associated with signature genes

We divided the dataset into high- and low-expression groups according to the median value of signature gene expression and performed GSEA on four signature genes. The results revealed that several signature gene high-expression groups were significantly enriched in immune response-related pathways, the IL-17 signaling pathway, allograft rejection, and the TNF signaling pathway (Fig. S1).

3.7. Hallmark Gene Set with ssGSEA

Further study of phenotypic differences between OA and normal samples using ssGSEA showed that in the set of all genes with significant differences, the OA group scored higher than controls in peroxisomes, and the OA group scored lower than controls in the gene sets for IL-2 STAT5 signaling, UV response up, p53 pathway, estrogen response early, apoptosis, TGF-β signaling, cholesterol homeostasis, hypoxia and TNF-α signaling via NF-κB (Fig. 6A). EPYC was significantly negatively correlated with UV response up, TNF-α signaling via NF-κB, TGF-β signaling, the p53 pathway, hypoxia, early estrogen response, and apoptosis, and significantly positively correlated with peroxisome and interferon-alpha responses. FKBP5 was significantly positively associated with UV response up, unfolded protein response, MYC targets V2, hypoxia,

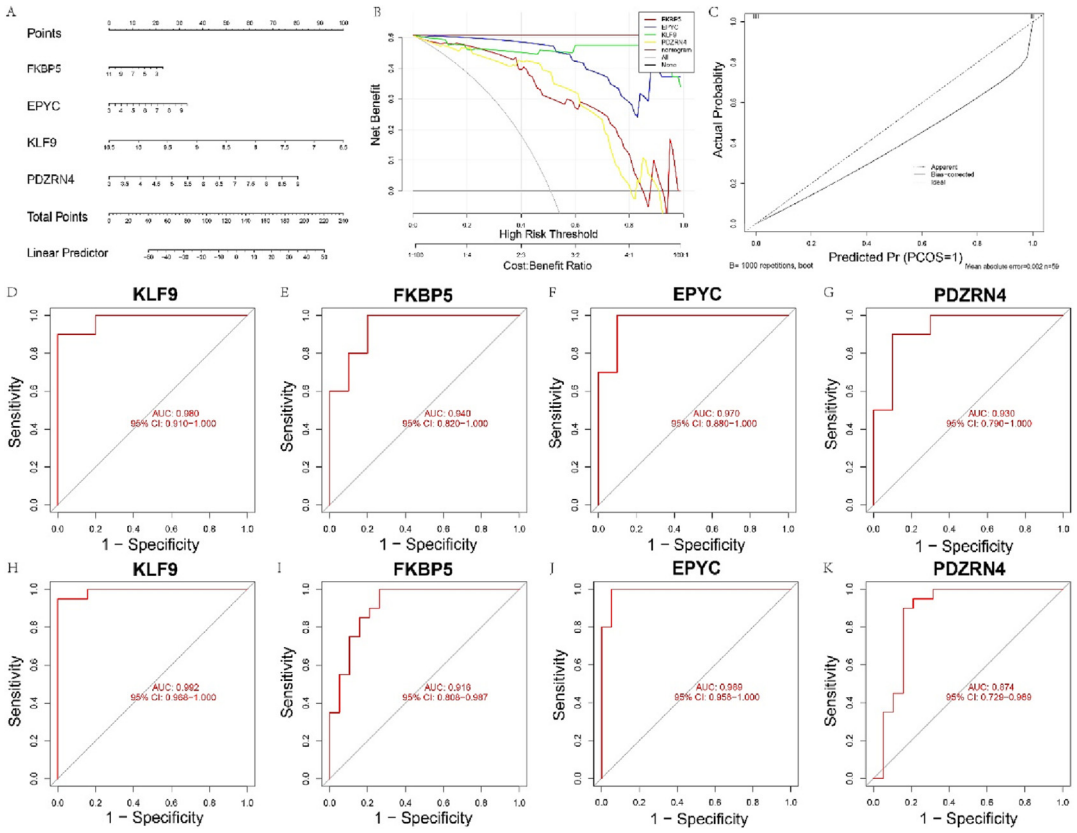


Fig. 5. Construction and validation of Nomogram diagnostic model; diagnostic significance of signature genes in OA. (A) Nomogram for predicting the diagnosis of OA. (B) Decision Curve Analysis (DCA) for evaluating clinical predictive model. (C) Calibration curves to evaluate consistency (D-G) ROC assay for signature genes of the training dataset. (H-K) ROC assay for signature genes of the validation dataset.

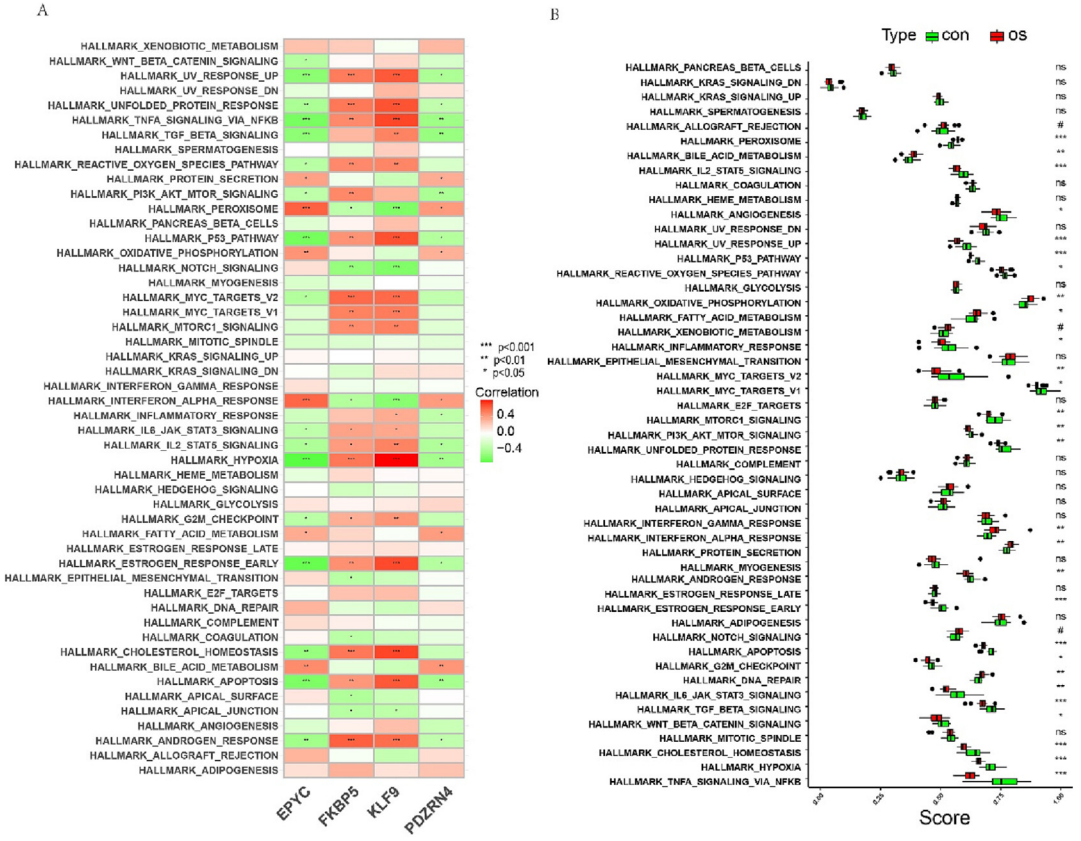


Fig. 6. Correlation of OA with Hallmark pathway. (A) Correlation of signature genes with Hallmark pathway. (B) Comparison of ssGSEA scores of Hallmark pathway in OA group and healthy controls. *P < 0.05, **P < 0.01, ***P < 0.001, NS not significant.

cholesterol homeostasis, and androgen response. KLF9 was significantly positively associated with UV response up, unfolded protein response, TNF-α signaling via NF-κB, P53 pathway, MYC target V2, MYC target V1, hypoxia, early estrogen response, cholesterol homeostasis, apoptosis, androgen response, and significantly negatively correlated with peroxisomal, notch signaling, and interferon-alpha responses (Fig. 6B).

3.8. Diagnostic signature gene correlation and expression validation

The expression of these four genes in OA was verified using the pooled dataset of GSE55235 and GSE12021 (Fig. 7A). FKBP5 and KLF9 were expressed at significantly lower levels in patients with OA, and EPYC and PDZRN4 were significantly highly expressed in patients with OA, which was confirmed in the GSE55457 dataset (Fig. 7B). Additionally, we calculated the Pearson correlation coefficient between diagnostic feature genes. As shown in Fig. 7C, there were significant interactions between the diagnostic features. Except for the significant positive correlations between FKBP5 and KLF9, EPYC, and PDZRN4, there were significant negative correlations between the remaining genes, suggesting that the four genes may not have mutually reinforcing relationships.

3.9. Immunoinfiltration analysis

The CIBERSORT algorithm was used to examine the fraction of 22 immune cells in the tumor microenvironment, and the results are shown in Fig. 8A. We did not detect any appreciable infiltration

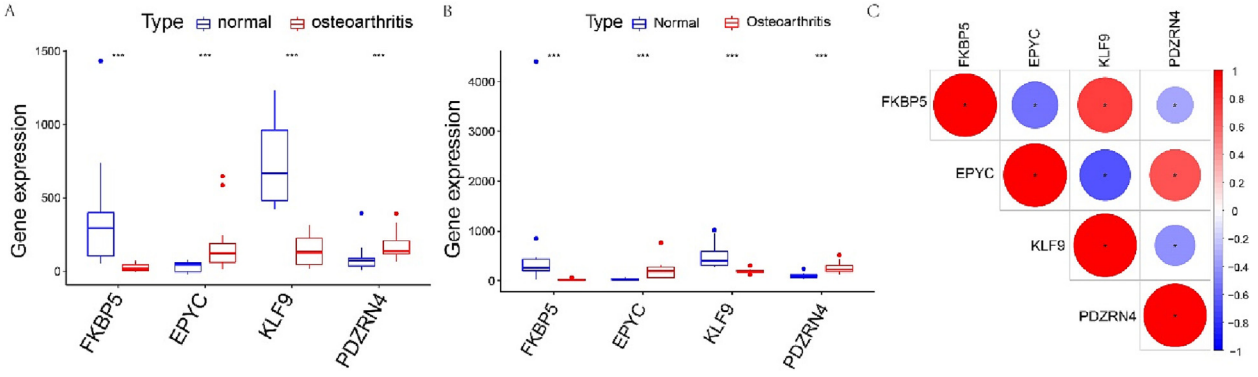


Fig. 7. Expression of signature genes and correlation analysis. (A) Differential analysis of signature genes in the training cohort. (B) Differential analysis of the signature genes in the validation cohort. (C) Correlation between signature genes.

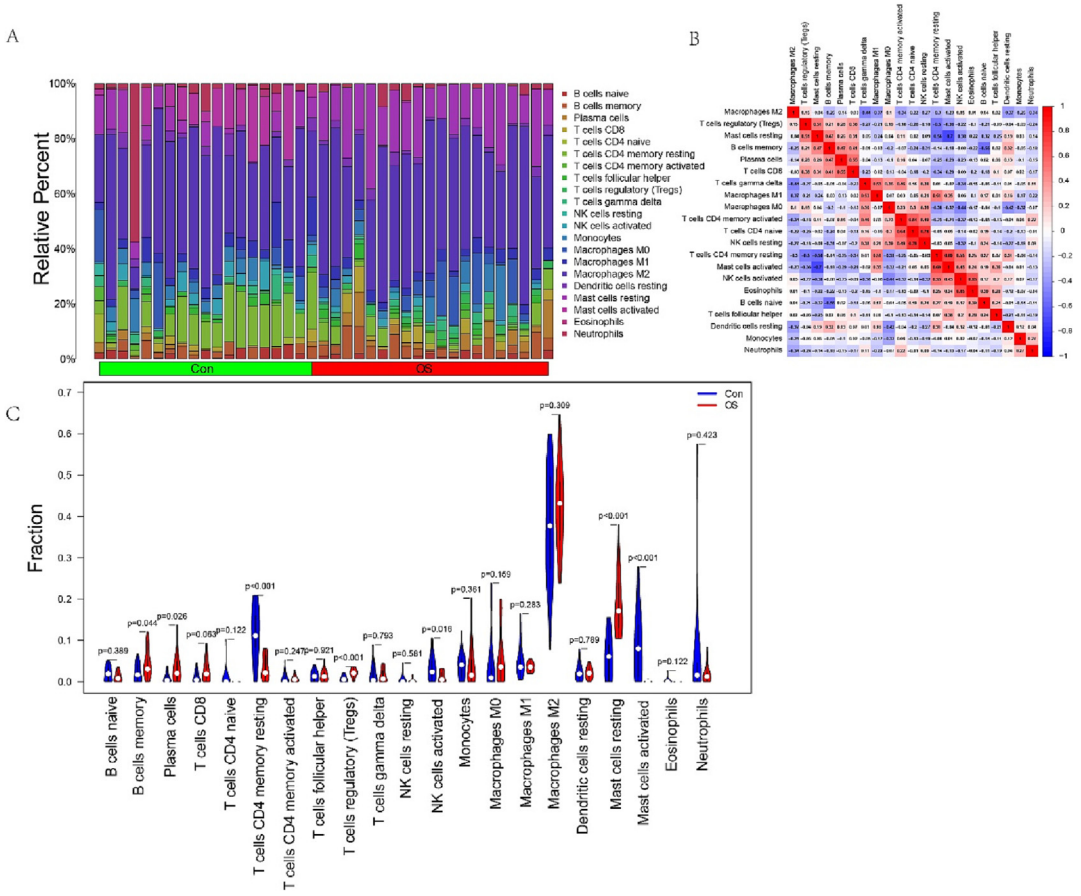


Fig. 8. Association of signature genes with immune infiltration in OA. (A) Application of the CIBERSORT algorithm to predict the percentage of 21 immune cell species. (B) After the removal of one non-differential immune cell, the graph shows the correlation heat map of 21 immune cells. (C) Violin plot showing differences in 21 immune infiltrating cells between the two groups.

of activated dendritic cells into the tumor microenvironment. Differential and correlation analyses of immune cells were performed to determine the correlation between OA and immune cell infiltration. In the differential analysis, seven immune cell types were statistically significant: memory B cells, plasma cells, resting memory CD4+ T cells, regulatory T cells (Tregs), activated natural-killer (NK) cells, resting mast cells, and activated mast cells (Fig. 8C). The heat map (Fig. 8B) demonstrates the Pearson correlation between immune cells, which showed that in patients with OA, naive CD4 + T cells were significantly and positively correlated with activated memory CD4+ T cells; resting NK cells were significantly and positively correlated with naive CD4+ T cells; activated mast cells were positively correlated with resting memory CD4+ T cells; eosinophils were positively correlated with activated NK cells; activated mast cells were negatively correlated with resting mast cells; naive B cells were negatively correlated with memory B cells; and resting mast cells were negatively correlated with resting memory CD4+ T cells, which confirmed the feasibility of correlation analysis to predict biomarkers.

3.10. Analysis of the correlation between diagnostic signature genes and immune infiltrating cells

In the correlation analysis (Fig. 9), EPYC was positively correlated with resting mast cells, plasma cells, memory B cells, and Tregs and negatively correlated with activated NK cells, resting memory CD4+ T cells, and activated mast cells. FKBP5 was positively correlated with activated mast cells, resting memory CD4+ T cells, and eosinophils and negatively correlated with Tregs and resting mast cells. KLF9 was positively correlated with activated mast cells, resting memory CD4+ T cells, activated NK cells, and eosinophils, and negatively correlated with plasma cells, quiescent mast cells, and Tregs. PDZRN4 was positively correlated with resting mast cells, memory B cells, plasma cells, CD8+ T cells, and Tregs and negatively correlated with resting memory CD4+ T cells and activated mast cells.

3.11. Expression of FKBP5 in normal and tumor tissues

To further explore the interactions between diagnostic signature genes, we used Genemania to construct PPI networks. FKBP5 had the highest degree value; therefore, we included FKBP5 in a deeper analysis. As shown in Fig. S2, the PPI network centered on FKBP5 was significantly enriched in NF-κB signaling, toll-like receptor signaling pathway, threonine kinase activity, and pattern recognition receptor signaling pathways, all of which play an important role in tumorigenesis and chemoresistance; therefore, FKBP5 was included in further analysis. Taking the TCGA database as an example (Fig. 10A), among all 13 statistically significant tumor types, only two types of tumors (glioblastoma multiforme (GBM) and liver hepatocellular carcinoma (LIHC)) had significantly high expression of the FKBP5 gene (Fig. S3). After merging the TCGA database with the GTEx database (Fig. 10B), 18 of the 21 tumors had low FKBP5 expression, and the remaining three tumors, including brain lower grade glioma (LGG), GBM, and LIHC, had high FKBP5 expression. The expression of FKBP5 in these cell lines is shown in Fig. 10C.

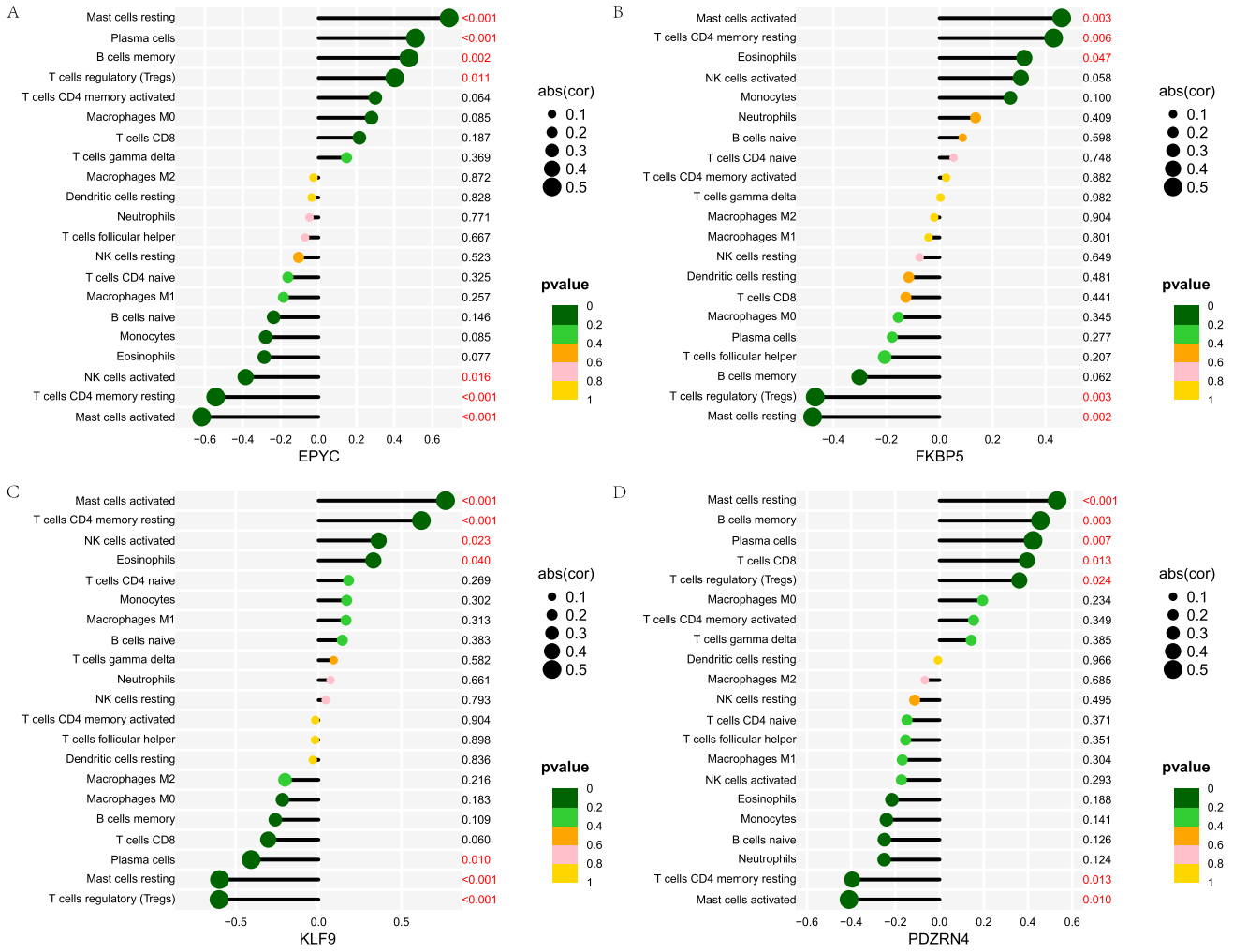


Fig. 9. Correlation of signature genes with infiltrating immune cells. (A) EPYC (B) FKBP5 (C) KLF9 (D) PDZRN4.

3.12. FKBP5 is associated with multiple tumor stages

The correlation between FKBP5 and pathological staging of different tumors was calculated using the online tool GEPIA. Lymph node diffuse large B-cell lymphoma (DLBC), head and neck squamous cell carcinoma (HNSC), kidney chromophobe, lung adenocarcinoma (LUAD), stomach adenocarcinoma (STAD), and uterine carcinosarcoma (UCS) were found to correlate significantly with tumor staging (Fig. 11), and the $p < 0.05$ was considered statistically significant.

3.13. The relationship between FKBP5 and tumor prognosis

To assess the predictive link between gene expression and prognosis in each tumor, as illustrated in Fig. 12, we created a Cox proportional hazards regression model. Prognostic significance was determined using a statistical test with the log-rank test. Finally, high expression in eight tumor types (GBMLGG, LGG, acute myeloid leukemia, LUAD, GBM, bladder urothelial carcinoma (BLCA), uveal melanoma (UVM), and STAD) was associated with a poor prognosis. Low expression in four tumor types (skin cutaneous melanoma (SKCM), skin cutaneous melanoma-metastasis, READ, and acute lymphoblastic leukemia) was associated with poor prognosis. Among these, FKBP5 had the highest risk for UVM. Patients were subsequently dichotomized according to the best cut-off values to obtain survival differences between the high and low FKBP5 expression groups in different cancer types. The results showed that the difference in survival was significant in all OS-related cancer types, while FKBP5 expression was not significantly correlated with OS in all TCGA tumor types (hazard ratio = 0.94).

As shown in Fig. 13, the correlation between gene expression and disease-free survival (DFS) in each tumor was analyzed by building a Cox proportional hazards regression model, and statistical tests were performed using the log-rank test. Unlike OS, there was a significant risk effect of FKBP5 expression in stomach and esophageal carcinoma, pheochromocytoma and paraganglioma, prostate adenocarcinoma (PRAD), and UCS, and FKBP5 expression levels were associated with DFS in a variety of tumors.

3.14. Correlation between FKBP5 expression and immune infiltration

We used the TIMER database to calculate the association between FKBP5 expression and multiple immune cell infiltration to determine whether gene expression affects the tumor immune environment. We found that infiltration scores for six immune cell types (B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, and dendritic cells) in the TIMER database were significantly associated with FKBP5 expression in multiple malignancies (Fig. 14A)

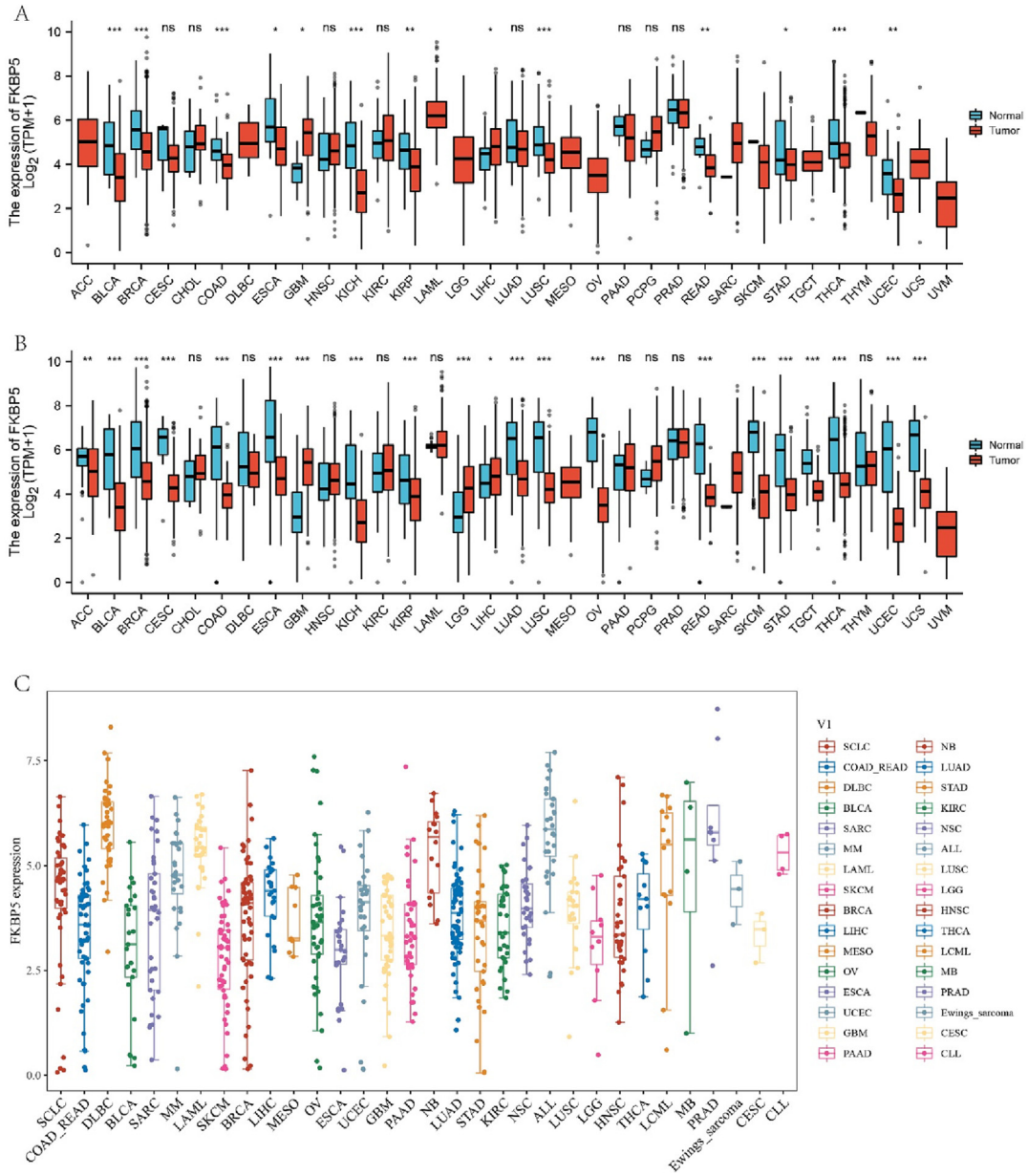


Fig. 10. FKBP5 expression in various cancers. (A) Differential expression of FKBP5 in tumor and normal tissues among 33 tumors in the TCGA database. (B) TCGA combined with data from normal tissues in the GTEx database, FKBP5 expression in pan-cancer. (C) Expression of FKBP5 in different tumor cell lines in the CCLE database.

and validated these using the algorithms xCELL (Fig. 14B), quanTIseq (Fig. 14C), MCPcounter (Fig. 14D), and EPIC (Fig. 14E).

3.15. Correlation of FKBP5 expression with cancer immune check-point gene expression

Several genes have been identified to be closely associated with immune responses and, therefore, act as immune checkpoints. We used cancer expression data from the TCGA database to examine the association between FKBP5 and immune check-points. In multiple cancer types, FKBP5 was significantly associated with multiple genes, including CD86, CD274, PDCD1LG2, CD28, CD48, CTLA4, TNFSF4, and bladder urothelial carcinoma CD200, FKBP5 was found to be coexpressed with more immune check-point genes in BLCA, SKCM, STAD, and testicular germ cell tumors (TGCTs), which may indicate that FKBP5 is involved in the tumor immune response through the above immune checkpoints in BLCA, SKCM, STAD, and TGCT (Figure 15A).

3.16. FKBP5 correlates with TMB and MSI in some cancers

In UCS, uterine corpus endometrial carcinoma (UCEC), and colon adenocarcinoma (COAD), FKBP5 expression was positively correlated with TMB, while in thymoma, thyroid carcinoma, STAD, PRAD, pancreatic adenocarcinoma, lung squamous cell carcinoma, and esophageal carcinoma (ESCA), it was negatively correlated with TMB. In UCEC, kidney renal papillary cell carcinoma, and COAD, FKBP5 was positively correlated with MSI, but in TGCT, STAD, PRAD, ovarian serous cystadenocarcinoma, HNSC, ESCA, DLBC, and breast invasive carcinoma, FKBP5 expression was negatively correlated with MSI (Fig. 15B and C).

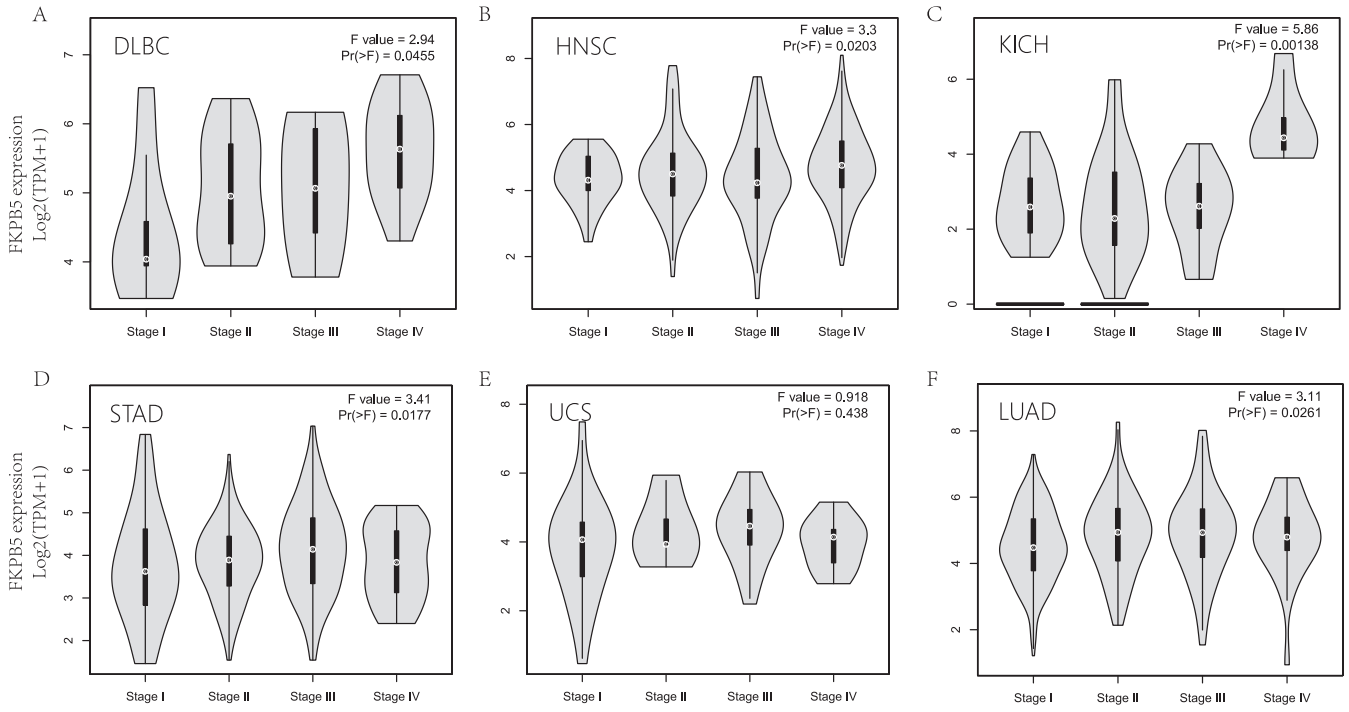


Fig. 11. Correlation between FKBP5 and pathological stage of cancers. (A) DLBC; (B) HNSC; (C) KICH; (D) STAD; (E) UCS (F) LUAD.

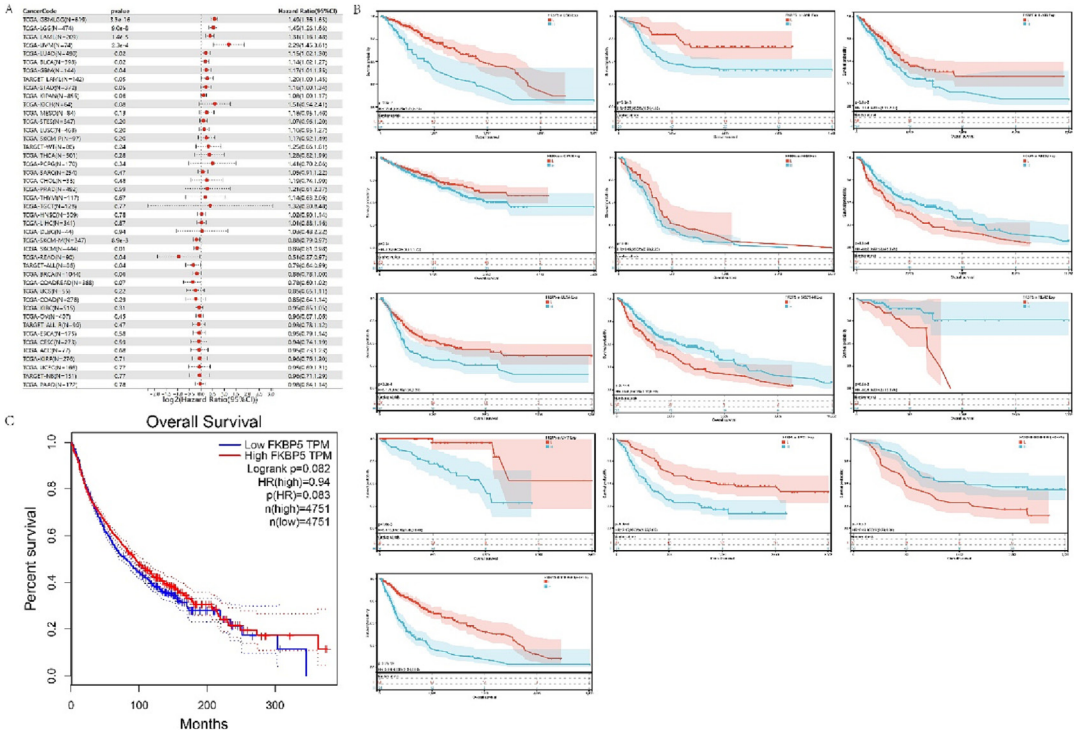


Fig. 12. Expression of FKBP5 in different tumors in relation to overall survival (OS). (A) Using information from the TCGA database, single variate Cox regression analysis was done to look into the connection between FKBP5 expression and overall survival (OS) in various tumor types. (B) In 13 tumor types, FKBP5 is correlated with overall survival (OS). (C) By classifying different cancers into high and low FKBP5 expression groups based on median FKBP5 expression levels, the link between overall survival (OS) and FKBP5 expression was examined.

3.17. FKBP5 and drug reactions

FKBP5 expression was positively associated with drug response in patients treated with hydroxyurea, fenretinide, fostamatinib, DMAPT, RH1, 8-chloroadenosine, chlorambucil, 3 -bromopyruvate (acid), imexon, nitrogen mustard, uracil mustard, parthenolide, chelerythrine, and vorinostat. Notably, there was a negative correlation between FKBP5 expression and kahalide F, a decapping peptide from the sea slug Elysia rufescens that targets alterations in lysosomal membrane function and is currently undergoing clinical

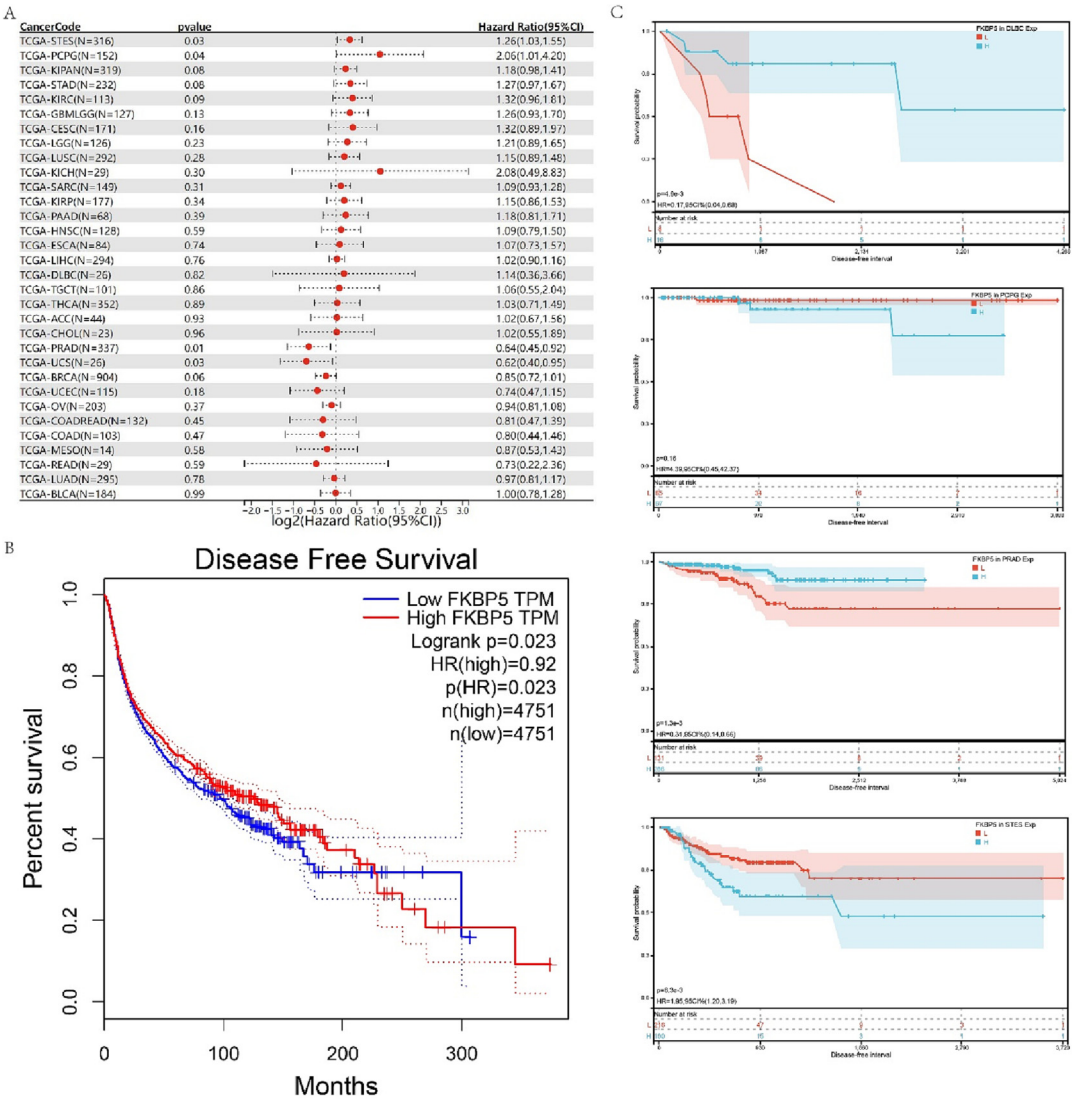


Fig. 13. Expression of FKBP5 in different tumors in relation to disease-free survival (DFS). (A) One variation using information from the TCGA database, Cox regression analysis was used to examine the connection between FKBP5 expression and DFS in various tumor types. (C) By splitting different tumors into groups with high and low FKBP5 expression based on median FKBP5 expression values, the connection between DFS and FKBP5 expression was examined. (B) In 4 tumor types, FKBP5 is correlated with DFS.

trials for a variety of oncology treatments. FKBP5 and the expected drug responses are shown in Fig. 16.

3.18. Cancer phenotypes and methylation of FKBP5

We applied the GEPIA database to identify the 94 genes that were most significantly associated with FKBP5 in pan-cancer. The gene list was extracted, and the functionally enriched analysis was performed. It was found that the genes significantly associated with FKBP5 were significantly enriched in the functions of steroid hormone signaling and intracellular substance transport (Fig. 17A). Subsequently, we constructed an FKBP5-centered interaction network using the BioGRID database (Fig. 17B). We found that FKBP5 physically interacts with NR3C1 and that NR3C1 has an important function in steroid hormone signaling (Fig. 17C). And the expression of FKBP5 had a significant positive correlation with NR3C1 in most tumors (Fig. 17D). Based on these results, we hypothesized that FKBP5 may play a tumor-promoting role in cancer by driving steroid hormone signaling and intracellular substance transport. DNA methylation directly affects cancer development and progression. We investigated the DNA methylation of FKBP5 using the UALCAN and TCGA databases. According to the UALCAN database, FKBP5 methylation levels were significantly decreased in READ, BLCA, KIRC, LUAD, LIHC, THCA, PCPG, TGCT, BRCA, PRAD, CESC, HNSC, LUSC, UCEC, COAD, PAAD and THYM tissues compared to normal tissues (Fig. 17E; Fig. S4).

3.19. Validation of the role of FKBP5 in immunotherapy in the TIGER and TISIDB databases

According to data in public databases, in glioblastoma groups, FKBP5 was low and statistically significant in the immunotherapy non-responsive group. In Melanoma immunotherapy groups, the gene was not consistently expressed in different data sets (Fig. S5). Detailed information can be found in Table S6.

4. Discussion

OA is a chronic disease with high global prevalence, and its diagnosis relies on clinical manifestations and imaging examinations. These ancillary tests are often lagging, making it difficult to predict and diagnose OA cases at an early stage; however, current

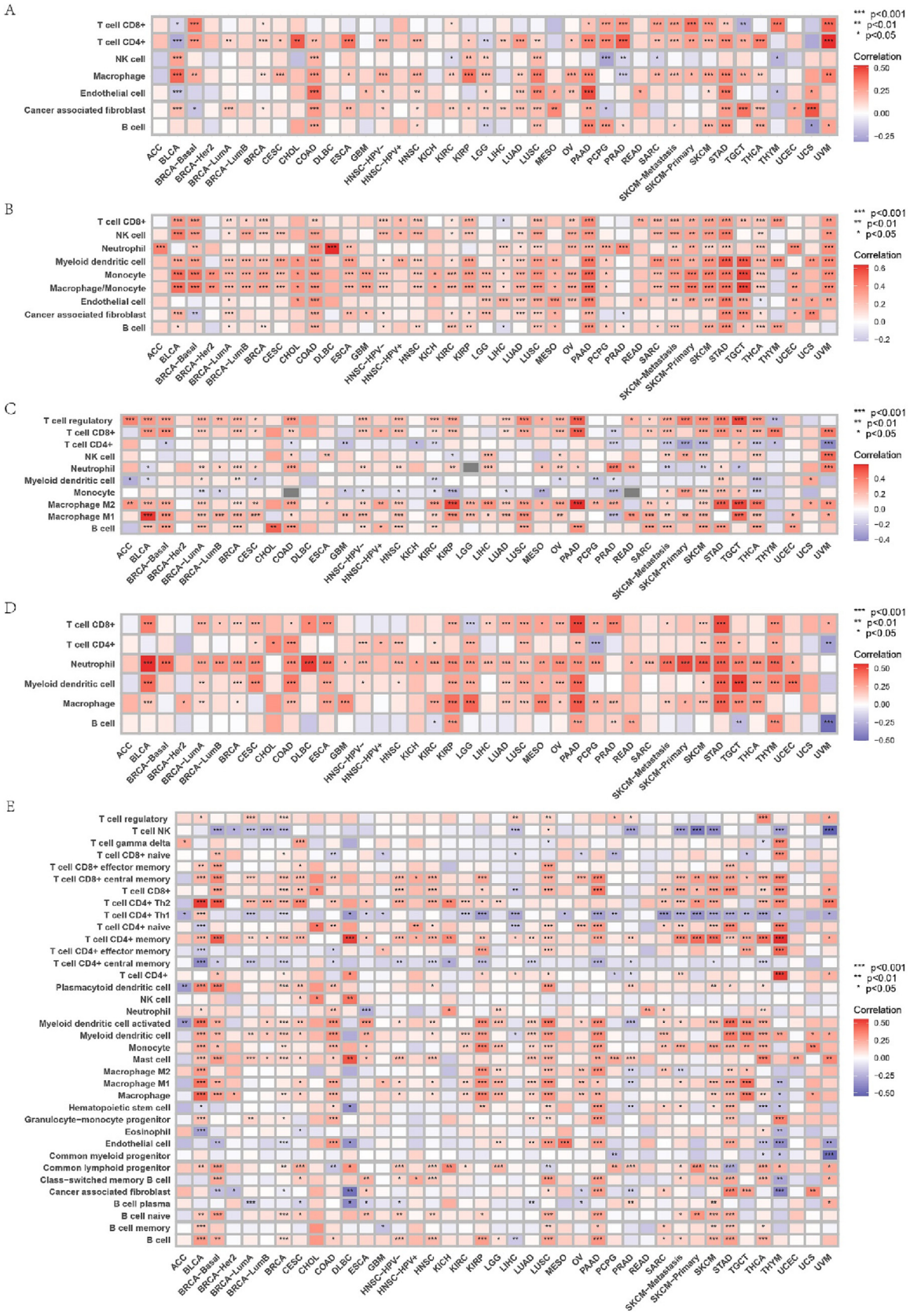


Fig. 14. Application of multiple tumor immune infiltration analysis tools to analyze the role of FKBP5 in tumor immunity. (A) XPIC; (B) MCPCOUNTER; (C) QUANTISEQ; (D) TIMER; (E) XCELL.

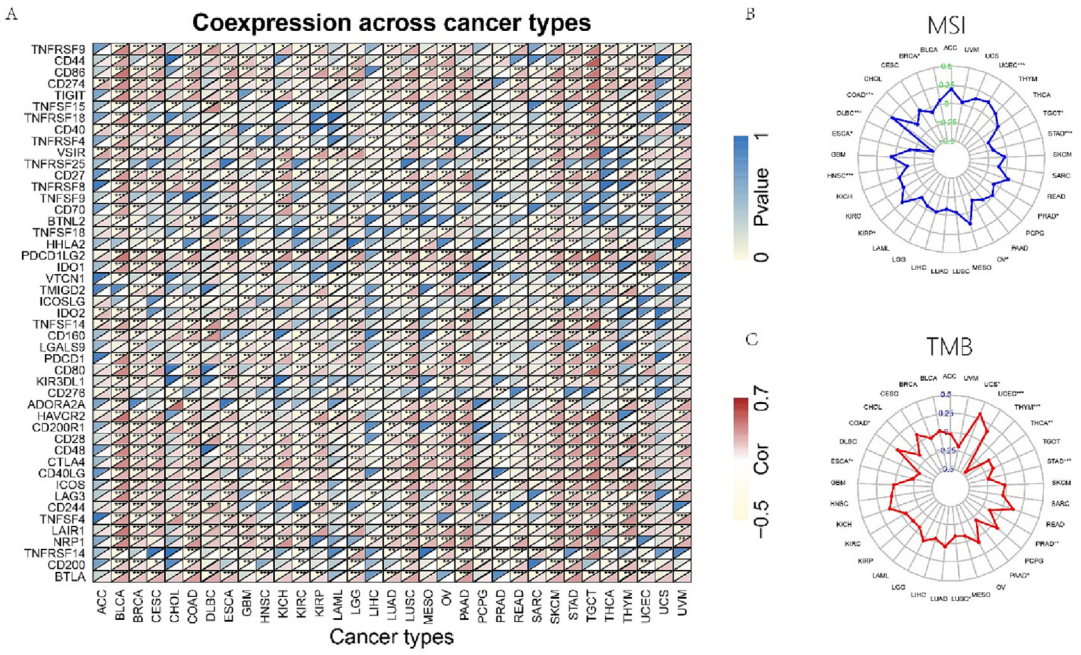


Fig. 15. (A) FKBP5 expression significantly correlates with most immune checkpoints in multiple tumors. (B) FKBP5 correlates with Microsatellite Instability (MSI) in a variety of tumors. (C) FKBP5 correlates with tumor mutation burden (TMB) in a variety of tumors.

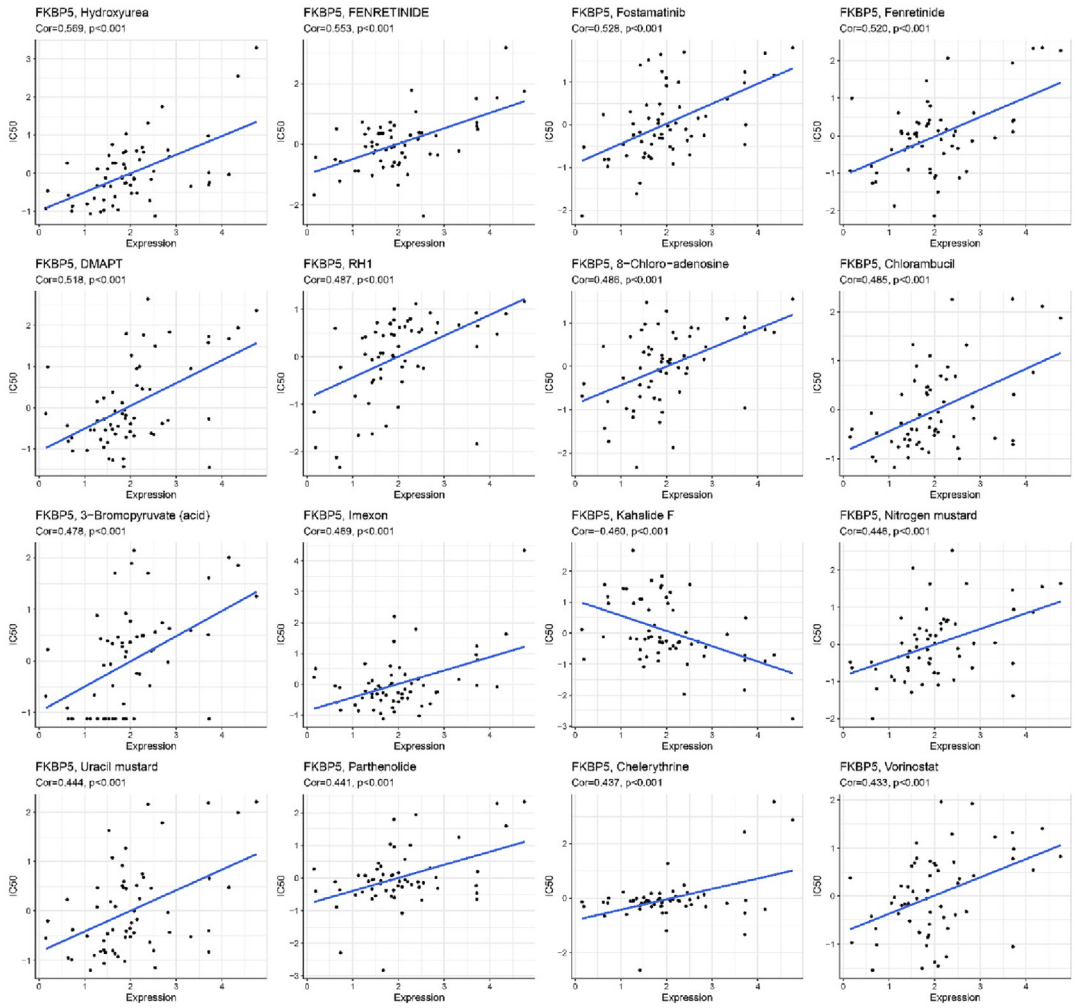


Fig. 16. Prediction of the relationship between FKBP5 expression and expected drug response.

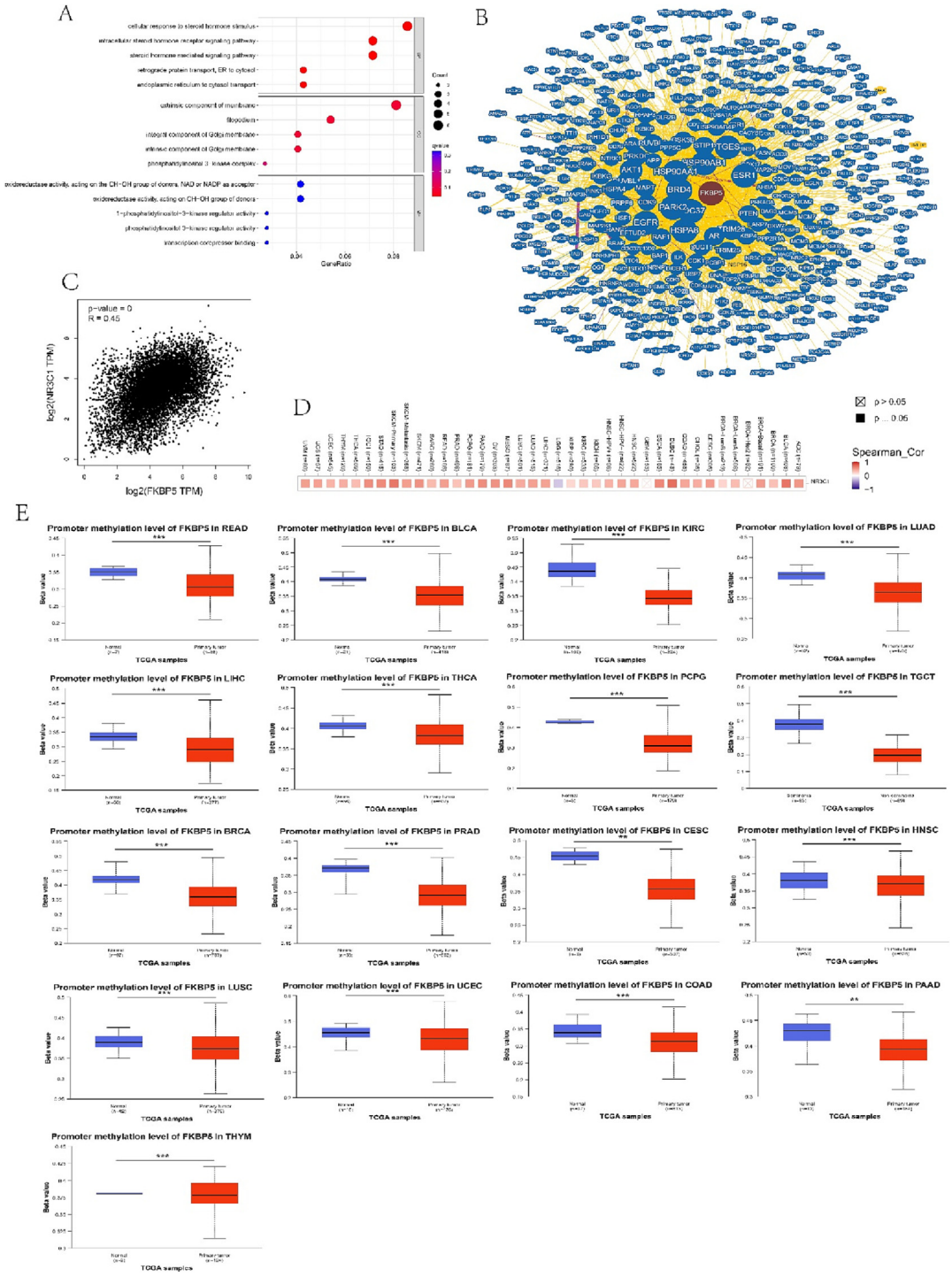


Fig. 17. (A) Functional enrichment analysis of FKBP5-related genes. (B) Interacting network of FKBP5. (C) Correlation analysis of FKBP5 with NR3C1. (D) Correlation analysis of FKBP5 and NR3C1 in multiple tumors. (E) Methylation differences of FKBP5 in multiple tumor samples.

technology does not yet allow for a cure for advanced OA. In addition, OA as a multifactorial disease has recently received more attention regarding the role of inflammation and immune cells in its pathophysiology, and the targets of these pathways bring new possibilities for future OA immunotherapy [41]. However, many challenges remain in identifying OA-related biomarkers and targets for drug therapy from current studies. As high-throughput data continue to accumulate, new biological techniques and applications are needed to identify specific molecules as biomarkers for early diagnosis and treatment, and to further explore the different molecular mechanisms involved in the pathogenesis and progression of OA. Therefore, the search for characteristic diagnostic markers and clarification of the mechanisms of immune infiltration is important for the prognosis of OA.

In the coexpressed gene set, 390 DEGs were identified, including 184 upregulated genes and 206 downregulated genes, among which the set of genes located in the same biological pathway or exhibiting the same function under certain conditions often have some connection. A gene coexpression network has been constructed [42]. By constructing a linear model of the relationship

between genes within a module and the relationship between gene traits within a module (MM & GS), the modules that are better associated with clinical traits are screened. The intersecting genes of the module genes and differentially expressed genes were used as key (hub) genes for subsequent analysis. A total of 161 hub genes were screened out, and GO enrichment analysis showed that hub genes were mainly involved in cellular responses, such as external stimulus, response to peptide, and myeloid cell differentiation. KEGG enrichment showed that hub genes were associated with the JAK-STAT, p53, TNF, and IL-17 signaling pathways, as well as osteoclast differentiation.

The results of this study demonstrate a strong association between multiple ILs and cytokines in OA. Previous studies have shown that ILs, such as IL-1 β, IL-6, IL-15, IL-17, IL-18, and TNFs, such as TNF-α, are the main inflammatory factors involved in OA. These inflammatory factors lead to degradation of the extracellular matrix and articular cartilage damage by affecting the release of other inflammatory mediators and activating downstream signaling pathways [43]. IL-1, IL-6, and TNF-α are involved in OA through the NF and JAK/STAT pathways, which play important roles in the pathogenesis of rheumatoid arthritis and OA [44], and inhibitors targeting IL-1, IL-6, and TNF-α have been clinically beneficial [44]. It has been shown that the p53 gene inhibits osteoblast proliferation and enhances osteoblast apoptosis, thereby reducing bone formation and acting as a bone-remodeling agent [45,46]. OA as a bone resorption disease is closely related to osteoclast bioactivity, and abnormal osteoclast function leads to disruption of subchondral bone remodeling. The above enrichment results validate the relationship between OA-related genes and functional molecular signaling pathways, which reflects the correctness of the hub gene screen.

Through our comprehensive strategy and multiple validations, FKBP5, EPYC, KLF9, and PDZRN4 were found to have high diagnostic value and could be potential diagnostic and therapeutic targets for OA. LASSO is a statistical approach for compressed estimation that creates a penalty function to determine the best diagnostic model. The model preserves certain regression coefficients that are zero, while compressing the regression coefficients such that the total absolute value of the regression coefficients employed is less than a fixed value. Subset shrinkage is an advantage of this approach [31]. A support vector machine (SVM) is a general machine-learning method that uses the principle of minimizing structural risk while minimizing empirical error, which can better solve the problems characterized by high latitude, small samples, and nonlinearity. It is an algorithm that combines SVM with recursive feature filtering (RFE) to train the model by multiple iterations to remove unnecessary features that reduce the spatial dimensionality [47]. Random forest is an integrated algorithm consisting of decision trees [48], whose models have the advantages of high randomness, noise immunity, and high interpretability, but at the same time, their models tend to be too general and do not have the ability to correctly handle difficult samples. All three algorithms have their own characteristics, and this study integrated each. The ROC curves of the test group confirmed that the integration of the three algorithms improved the accuracy while reducing the number of critical genes. Previous studies have identified biomarkers of OA using machine-learning methods. For example, Liang et al. [49] used SVM-RFE and Lasso algorithms to screen APOLD1 and EPYC as diagnostic genes for OA; Zhang et al. [50] used similar screening methods to select EPYC and KLF9 as diagnostic genes, which may indicate that differences in logFC values, gene set selection, and parameter selection affect the screening of candidate genes. However, they did not validate the expression of diagnostic genes based on intra-articular cavity samples, while the EPYC and KLF9 genes were identified as diagnostic markers for OA, which also indicated the feasibility of our analysis strategy.

FKBP5, a member of the immunophilic protein family, was first identified 20 years ago by Baughman et al. and has now been associated with a variety of specific diseases, such as stressor exposure leading to multiple psychiatric disorders, asthma, obesity, and type 2 diabetes [51,52]. It has been established that age- and stress-related epigenetic characteristics increase the FKBP5 response to NF-κB through a positive feedback loop, and that FKBP5 overexpression promotes NF-κB-associated peripheral inflammation [53]. Many studies now support a central role for NF-κB signaling in the intrachondrocytic inflammatory response [54]. Therefore, we hypothesized that FKBP5 is involved in the regulation of the chondrocyte inflammatory response. In addition, FK506-binding protein 51 (FKBP5/FKBP51) is an important regulator of the stress response that acts as a cochaperone to alter the folding and activity of other proteins involved in regulating the sensitivity of the glucocorticoid receptor (GR) [55], while GR-mediated signaling has been shown to be involved in the metabolism of OA, affecting intraarticular behavior and communication of multiple cells in healthy and abnormal states [56].

EPYC is a member of the small leucine-rich proteoglycan (SLRP) family, which regulates protofibrillogenesis by interacting with collagen fibers and other extracellular matrix proteins. According to previous reports, SLRPs may be crucial to the pathophysiology of OA [57,58]. SLRP fragmentation is increased in OA joints compared to normal cartilage [59], whereas it has been shown that chondrocytes increase the production of SLRPs in advanced arthritic joints [60]. Therefore, it is reasonable to infer that SLRPs are essential for maintaining normal chondrocyte activity and cartilage tissue integrity. In addition, many SLRP metabolic fragments have been reported as possible biomarkers for assessing cartilage metabolism or OA progression [59,60], and some investigators have advocated SLRP fragmentation as a promising biomarker for assessing OA status [61]. Although the role of EPYC as a biomarker in OA is unknown, it has a high potential for OA diagnosis based on its background and raw letter analysis.

KLF9 is a protein-coding gene that encodes a transcription factor. This encoded protein binds to the GC box of the promoter and regulates mRNA transcription. Reactive oxygen species (ROS) are by-products of cellular respiration and metabolism in the mitochondria, and small amounts of ROS are actively produced by cells to act as signaling molecules. Under oxidative stress conditions, cellular metabolism and expression products maintain the cellular redox state by activating NF-E2-related factor 2 (Nrf2). Nrf2 induces KLF9 expression, while KLF9 amplifies oxidative stress by inhibiting Txnrd2 [62], increasing intracellular ROS, which then participates in the pathophysiological progression of OA through complex signaling pathways [63]. The results of ssGSEA showed a significant correlation between KLF9 and hypoxia and apoptosis; therefore, the mechanism of apoptosis induced by ROS in patients with OA and the potential role of KLF9 in this process need to be further investigated.

PDZRN4, a member of the PDZRN family, is more likely to function as an E3 ubiquitin ligase because of its multiple PDZ and RING structural domains [64]. Current studies suggest that it may act as a tumor suppressor in a variety of cancers [65,66,67]. In addition, PDZRN4 has been proposed as a novel biomarker for multiple sclerosis; however, its exact role of PDZRN4 in OA requires further investigation.

There is growing evidence to support the key role of the innate and acquired immune systems in low-grade inflammation associated with OA. Macrophages are the most common immune cell type in osteoarthritic synovial infiltrates [68] and have been classified into two subclasses: “classically activated” (M1) and “alternatively activated” (M2) [69]. M1 macrophages are activated by inflammatory mediators (e.g., TNF- and interferon γ (IFNγ)) or pathogen-associated molecular patterns, and secrete proinflamma-

tory cytokines (e.g., IL-1, IL-6, IL-12, and TNF-α) and other tissue damage signals. In contrast, M2 macrophages are activated by another pathway, releasing growth and angiogenic factors and regulating T cell function, which are involved in the downregulation of inflammation and tissue remodeling. However, a simple classification of macrophages does not reflect the complexity of cellular mechanisms, and altered polarization states of both may play a key role in the coordination of inflammation and regeneration, which requires further in-depth studies. T cells are second only to macrophages in the proportion of synovial infiltrating immune cells [70], and CD4+ T cells are more abundantly aggregated in the synovial membranes. Specifically, synovial tissue is dominated by the presence of Th1, Th17, and cytotoxic T cells, and these cell types produce catabolic substances including IL-2, IL-17, IL-23, IFN-γ, and TNF-α, which promote infiltration of immune cells and stimulate cartilage matrix protease destruction, thereby exacerbating cartilage damage [11,71]. Therefore, regulation of the local immune microenvironment and tissue regeneration microenvironment should receive more attention in the treatment of OA. In a study on OA and rheumatoid arthritis (RA), Tregs were increased in OA and RA compared to normal synovium, and Tregs present in RA and OA synovium were very similar [72]. In the presence of TGF-β, naive T cells differentiate into Treg cells, and Tregs produce cytokines, such as TGF-β, IL-10, and IL-35, which suppress inflammation and maintain immune tolerance. The production of Tregs is one of the main targets of immunotherapy and transplantation in autoimmune diseases, and Treg infiltration is significantly higher in patients with OA than in normal patients, as confirmed by the CIBERSORT algorithm. This suggests that the induction of Tregs in early OA might be a new research frontier for OA diagnosis and treatment. Immune infiltration analysis also revealed that mast cell infiltration was significantly higher in patients with OA than in the normal group because of the storage of multiple mediators, including cytokines and enzymes. These have been shown to regulate bone homeostasis and participate in a variety of bone metabolic diseases, and may regulate bone formation and healing tissue remodeling by influencing multiple pathological responses of osteoclasts in bone injury [46]. A clinical study found that the use of H1 antihistamine therapy was associated with a decrease in the prevalence of OA, which supports the important role of MC in the pathogenesis of OA, thus demonstrating that mast cells can be a new therapeutic target for OA [73].

A relatively large number of synovial infiltrates were observed in the OA synovium. In OA, cartilage destruction caused by multiple factors can activate humoral immunity, and multiple antibodies against cartilage breakdown products, such as bone bridge protein, cartilage intermediate layer protein, YKL-39, fibulin, and collagen have been detected [74,75]. In addition, immune cells infiltrating the OA synovium are oligoclonal, confirming that intraarticular B-cell infiltration is antigen-driven [76]. Antibodies play an important role in many diseases. When risk factors are present, humoral immunity activates the complement system and forms the terminal complement complex (TCC). However, the absence of TCC might negatively affect fracture healing [77].

Based on high-throughput data from TCGA and CCLE databases, we found that FKBP5 was differentially expressed in multiple cancer tissues. In addition, prognostic analysis revealed that this molecule has a poor prognosis in multiple tumors. In combination with ssGSEA, FKBP5 was associated with more metabolic dysfunction, which plays an important role in cancer progression, than other genes. A search for cancers associated with FKBP5 revealed that this molecule has not been systematically elaborated in pan-cancer; therefore, we elevated FKBP5 to a new dimension for analysis.

We found that several existing literatures have demonstrated the role of FKBP5 in inflammation and immune regulation. It has been shown that stress- and age-induced epigenetic alterations upregulate FKBP5 expression, and that FKBP5 can regulate inflammatory responses by enhancing TNF-induced NF-κB activation. In contrast, counteracting FKBP5 by genetic deletion (CRISPR/Cas5-mediated) or selective pharmacological inhibition blocked the effect on NF-κB [51,78]. In addition, several studies have found that the expression level of FKBP5 is regulated by TNF-α and other inflammatory factors, which also suggests that FKBP5 is closely related to the TNF pathway in the regulation of inflammation and immune responses. Although the link between FKBP5 and OA and cancer is not fully understood, previous studies have suggested that FKBP5 may play a role in both diseases by participating in inflammatory and cellular signaling pathways. Therefore, we believe that our investigations may provide new insights into the pathogenesis of these diseases and may identify common therapeutic targets.

Some studies have reported a correlation between OA and the development of several types of cancer. First, OA has been shown to be associated with an increased risk of cancer, which greatly increases the risk of leukaemia subtypes, myelodysplastic syndromes and essential thrombocytosis [17]. Second, drugs used to treat OA may be associated with the observed risk of cancer, and studies have shown that NSAIDs [18] and hyaluronic acid compounds [23] taken by patients with osteoarthritis promote the development of cancer. Therefore, we further explored the role of FKBP5, a gene characteristic of osteoarthritis, in pancancer. FKBP5 has been shown to be differentially expressed in a variety of tumors, including melanoma, glioma, colon cancer, and prostate cancer [79]. Among them, FKBP5 expression was associated with metastatic lesions in melanoma [80] and correlated with staging and overall survival in glioma [81], which was confirmed in our study and identified additional malignancies associated with FKBP5 prognosis and staging. FKBP5 has been shown to regulate different signaling pathways under specific conditions affecting tumor metabolism, including the steroid hormone-receptor signaling pathway, NF-κB signaling pathway, and AKT-PHLPP pathway, which have been shown to be involved in tumorigenesis and tumor resistance responses [82]. For example, FKBP5 transcription can be induced in prostate cancer cells by activating steroid hormone receptors, thereby stimulating cell growth, suggesting that FKBP5 can act as a drug target to disrupt AR-mediated signaling in prostate cancer [83]. This AR-mediated signaling acts as a marker for several functional tumor signaling pathways, as well as in response to antitumor drugs [84]. Notably, this may correlate with the use of multiple survival methods by tumor cells to evade immunity; therefore, we retrieved and compiled a list of 47 immune checkpoint genes to predict their correlation with FKBP5 expression and used multiple algorithms to estimate TMB in the tumor. We hope that this information can be used to predict the effect of immunotherapy with FKBP5 or to develop new targeted therapeutic options.

However, despite the use of multiple algorithms to screen genes and integrate information from multiple databases, there were still some limitations to this study. First, although bioinformatics analysis provided us with important information about FKBP5 in malignancies, molecular biology approaches or additional in vitro and in vivo experiments are needed to validate the potential role of FKBP5 in cancer and tumor immunity. Although we integrated data from OA samples from all populations, the limited clinical information in the database made it difficult to avoid conclusions with limitations and bias.

5. Conclusions

In this study, FKBP5, EPYC, KLF9, and PDZRN4 were identified as potential biomarkers of OA, thereby expanding the known poten-

tial molecular mechanisms of OA lesions and predicting the development of immune cells in OA. In addition, FKBP5, a core gene associated with OA and pancancer, is a promising therapeutic target in several cancers and is a marker of immune infiltration and poor prognosis for OA.

Author contributions

- Study conception and design: Y Xiao; Y Huo
- Data collection: Y Wang
- Analysis and interpretation of results: Y Xiao; X Jiao; X Xu
- Draft manuscript preparation: Y Xiao; Y Huo
- Revision of the results and approval of the final version of the manuscript: Y Xiao; Y Huo

Financial support

This research is funded by the Shengli Oilfield Central Hospital (The contract no: 6010220027).

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgments

We thank The Department of Orthopaedic Trauma, The Second Hospital, Cheeloo College of Medicine, Shandong University and Key Laboratory of Experimental Teratology of Ministry of Education, Institute of Medical Sciences/Department of Neurology, The Second Hospital, Cheeloo College of Medicine, Shandong University.

Preprint statement

The paper is published in preprint at https://www.research-square.com; https://doi.org/10.21203/rs.3.rs-2471831/v1.

Supplementary material

https://doi.org/10.1016/j.ejbt.2023.05.002.

References

[1] Loeser RF, Goldring SR, Scanzello CR, et al. Osteoarthritis: a disease of the joint as an organ. Arthritis Rheum 2012;64(6):1697–707. https://doi.org/10.1002/art.34453. PMid: 22392533.
[2] Palazzo C, Nguyen C, Lefevre-Colau MM, et al. Risk factors and burden of osteoarthritis. Ann Phys Rehabil Med. 2016;59(3):134–8. https://doi.org/10.1016/j.rehab.2016.01.006. PMid: 26904959.
[3] Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. Lancet 2019;393 (10182):1745–59. https://doi.org/10.1016/S0140-6736(19)30417-9. PMid: 31034380.
[4] Martel-Pelletier J, Barr AJ, Cicuttini FM, et al. Osteoarthritis. Nat Rev Dis Primers. 2016;2:16072. https://doi.org/10.1038/nrdp.2016.72. PMid: 27734845.
[5] Conaghan PG, Felson D, Gold G, et al. MRI and non-cartilaginous structures in knee osteoarthritis. Osteoarthritis Cartilage. 2006;14(Suppl A):A87–94. https://doi.org/10.1016/j.joca.2006.02.028. PMid: 16713722.
[6] Kumavat R, Kumar V, Malhotra R, et al. Biomarkers of joint damage in osteoarthritis: current status and future directions. Mediators Inflamm 2021;2021:5574582. https://doi.org/10.1155/2021/5574582. PMid: 33776572.
[7] Rosshirt N, Hagmann S, Tripel E, et al. A predominant Th1 polarization is present in synovial fluid of end-stage osteoarthritic knee joints: analysis of peripheral blood, synovial fluid and synovial membrane. Clin Exp Immunol 2019;195(3):395–406. https://doi.org/10.1111/cei.13230. PMid: 30368774.
[8] Zhang H, Cai D, Bai X. Macrophages regulate the progression of osteoarthritis. Osteoarthritis Cartilage 2020;28(5):555–61. https://doi.org/10.1016/j.joca.2020.01.007. PMid: 31982565.
[9] Bondeson J, Blom AB, Wainwright S, et al. The role of synovial macrophages and macrophage-produced mediators in driving inflammatory and destructive responses in osteoarthritis. Arthritis Rheum 2010;62(3):647–57. https://doi.org/10.1002/art.27290. PMid: 20187160.
[10] Silawal S., Triebel J., Bertsch T., et al. Osteoarthritis and the complement cascade. Clin Med Insights Arthritis Musculoskelet Disord 2018;11:1179544117751430. http://dx.doi.org/10.1177/1179544117751430. PMid: 29434479.
[11] Li YS, Luo W, Zhu SA, et al. T cells in osteoarthritis: alterations and beyond. Front Immunol. 2017;8:356. https://doi.org/10.3389/fimmu.2017.00356. PMid: 28424692.
[12] Nees TA, Rosshirt N, Zhang JA, et al. t helper cell infiltration in osteoarthritis-related knee pain and disability. J Clin Med 2020;9(8):2423. https://doi.org/10.3390/jcm9082423. PMid: 32751139.
[13] Zhao C. Identifying the hub gene and immune infiltration of osteoarthritis by bioinformatical methods. Clin Rheumatol 2021;40(3):1027–37. https://doi.org/10.1007/s10067-020-05311-0. PMid: 32785809.
[14] Yuan WH, Xie QQ, Wang KP, et al. Screening of osteoarthritis diagnostic markers based on immune-related genes and immune infiltration. Sci Rep. 2021;11(1):7032. https://doi.org/10.1038/s41598-021-86319-7. PMid: 33782454.
[15] Arnone MI, Davidson EH. The hardwiring of development: Organization and function of genomic regulatory systems. Development 1997;124 (10):1851–64. https://doi.org/10.1242/dev.124.10.1851. PMid: 9169833.
[16] Molla M, Waddell M, Page D, et al. Using machine learning to design and interpret gene-expression microarrays. AI Magazine 2004;25(1):23. https://doi.org/10.1609/aimag.v25i1.1745.
[17] Cai Y, Fan X, Zhao L, et al. Comparing machine learning-derived MRI-based and blood-based neurodegeneration biomarkers in predicting syndromal conversion in early AD. Alzheimers Dement 2023. https://doi.org/10.1002/alz.13083. PMid: 37087687.
[18] Lu YT, Plets M, Morrison G, et al. Cell-free DNA methylation as a predictive biomarker of response to neoadjuvant chemotherapy for patients with muscle-invasive bladder cancer in SWOG S1314. Eur Urol Oncol 2023. https://doi.org/10.1016/j.euo.2023.03.008. PMid: 37087309.
[19] Huber R, Hummert C, Gausmann U, et al. Identification of intra-group, inter-individual, and gene-specific variances in mRNA expression profiles in the rheumatoid arthritis synovial membrane. Arthritis Res Ther 2008;10(4):R98. https://doi.org/10.1186/ar2485. PMid: 18721452.
[20] Woetzel D, Huber R, Kupfer P, et al. Identification of rheumatoid arthritis and osteoarthritis patients by transcriptome-based rule set generation. Arthritis Res Ther. 2014;16(2):R84. https://doi.org/10.1186/ar4526. PMid: 24690414.
[21] Bogan RL, Murphy MJ, Hennebold JD. Dynamic changes in gene expression that occur during the period of spontaneous functional regression in the rhesus macaque corpus luteum. Endocrinology 2009;150(3):1521–9. https://doi.org/10.1210/en.2008-1201. PMid: 18948396.
[22] Leek JT, Johnson WE, Parker HS, et al. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;28(6):882–3. https://doi.org/10.1093/bioinformatics/bts034. PMid: 22257669.
[23] Goldman MJ, Craft B, Hastie M, et al. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol. 2020;38(6):675–8. https://doi.org/10.1038/s41587-020-0546-8. PMid: 3244485.
[24] Ritchie ME, Phipson B, Wu D, et al. *limma* powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43(7):e47. https://doi.org/10.1093/nar/gkv007. PMid: 25605792.
[25] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinf 2008;9:559. https://doi.org/10.1186/1471-2105-9-559. PMid: 19114008.
[26] Cingiz MÖ, Biricik G, Diri B. The performance comparison of gene co-expression networks of breast and prostate cancer using different selection criteria. Interdiscip Sci. 2021;13(3):500–10. https://doi.org/10.1007/s12539-021-00440-9. PMid: 34003445.
[27] Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16(5):284–7. https://doi.org/10.1089/omi.2011.0118. PMid: 22455463.
[28] Mullah MAS, Hanley JA, Benedetti A. LASSO type penalized spline regression for binary data. BMC Med Res Methodol. 2021;21(1):83. https://doi.org/10.1186/s12874-021-01234-9. PMid: 33894761.
[29] Engebretsen S, Bohlin J. Statistical predictions with glmnet. Clin. Epigenetics 2019;11(1):123. https://doi.org/10.1186/s13148-019-0730-1. PMid: 31443682.
[30] Sun J, Zhong G, Huang K, et al. Banzhaf random forests: Cooperative game theory based random forests with consistency. Neural Netw. 2018;106:20–9. https://doi.org/10.1016/j.neunet.2018.06.006. PMid: 30007122.
[31] Youssef Ali Amer A. Global-local least-squares support vector machine (GLocal-LS-SVM). PLoS One. 2023;18(4):e0285131. PMid: 37104506.
[32] Liberzon A, Birger C, Thorvaldsdóttir H, et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015;1(6):417–25. https://doi.org/10.1016/j.cels.2015.12.004. PMid: 26771021.
[33] Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinf 2013;14:7. https://doi.org/10.1186/1471-2105-14-7. PMid: 23323831.
[34] Wu H, Chen Z, Wu Y, et al. Integrating protein-protein interaction networks and somatic mutation data to detect driver modules in pan-cancer. Interdiscip

Sci. 2022;14(1):151–67. https://doi.org/10.1007/s12539-021-00475-y. PMid: 34491536.
[35] Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. Ann Statistics 1982;10(4):1100–20. https://doi.org/10.1214/aos/1176345976.
[36] Li T, Fu J, Zeng Z, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res 2020;48(W1):W509–14. https://doi.org/10.1093/nar/gkaa407. PMid: 32442275.
[37] Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol 2017;18(1):220. https://doi.org/10.1186/s13059-017-1349-1. PMid: 29141660.
[38] Racle J, de Jonge K, Baumgaertner P, et al. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. Elife. 2017;6: e26476. https://doi.org/10.7554/eLife.26476. PMid: 29130882.
[39] Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol 2016;17(1):218. https://doi.org/10.1186/s13059-016-1070-5.
[40] Finotello F, Mayer C, Plattner C, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. Genome Med 2019;11:34. https://doi.org/10.1186/s13073-019-0638-6.
[41] Lopes EBP, Filiberti A, Husain SA, et al. Immune contributions to osteoarthritis. Curr Osteoporos Rep. 2017;15(6):593–600. https://doi.org/10.1007/s11914-017-0411-y. PMid: 29098574.
[42] Hirose O, Yoshida R, Imoto S, et al. Statistical inference of transcriptional module-based gene networks from time course gene expression profiles by using state space models. Bioinformatics 2008;24(7):932–42. https://doi.org/10.1093/bioinformatics/btm639. PMid: 18292116.
[43] Ansari MY, Ahmad N, Haqqi TM. Oxidative stress and inflammation in osteoarthritis pathogenesis: role of polyphenols. Biomed Pharmacother. 2020;129:. https://doi.org/10.1016/j.biopha.2020.110452. PMid: 32768946110452.
[44] Nishimura R, Hata K, Takahata Y, et al. Role of signal transduction pathways and transcription factors in cartilage and joint diseases. Int J Mol Sci 2020;21 (4):1340. https://doi.org/10.3390/ijms21041340. PMid 32079226.
[45] Liu H, Li B. p53 control of bone remodeling. J Cell Biochem. 2010;111 (3):529–34. https://doi.org/10.1002/jcb.22749. PMid: 20589754.
[46] Komori T. Cell death in chondrocytes, osteoblasts, and osteocytes. Int J Mol Sci. 2016;17(12):2045. https://doi.org/10.3390/ijms17122045. PMid: 27929439.
[47] Sain SR. The nature of statistical learning theory. Technometrics 1996;38 (4):409. https://doi.org/10.1080/00401706.1996.10484565.
[48] Statnikov A, Wang L, Aliferis CF. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. BMC Bioinf 2008;9:319. https://doi.org/10.1186/1471-2105-9-319. PMid: 18647401.
[49] Liang Y, Lin F, Huang Y. Identification of biomarkers associated with diagnosis of osteoarthritis patients based on bioinformatics and machine learning. J Immunol Res. 2022;2022:5600190. https://doi.org/10.1155/2022/5600190. PMid: 35733917.
[50] Zhang J, Zhang S, Zhou Y. KLF9 and EPYC acting as feature genes for osteoarthritis and their association with immune infiltration. J Orthop Surg Res 2022;17(1):365. https://doi.org/10.1186/s13018-022-03247-6. PMid: 35902862.
[51] Zannas AS, Wiechmann T, Gassen NC, et al. Gene-stress-epigenetic regulation of *FKBP5:* Clinical and translational implications. Neuropsychopharmacology. 2016;41(1):261–74. https://doi.org/10.1038/npp.2015.235. PMid: 26250598.
[52] Wang CC, Shen WJ, Anuraga G, et al. Penetrating exploration of prognostic correlations of the FKBP gene family with lung adenocarcinoma. J Pers Med. 2022;13(1):49. https://doi.org/10.3390/jpm13010049. PMid: 36675710.
[53] Zannas AS, Jia M, Hafner K, et al. Epigenetic upregulation of FKBP5 by aging and stress contributes to NF-κB-driven inflammation and cardiovascular risk. Proc Natl Acad Sci USA 2019;116(23):11370–9. https://doi.org/10.1073/pnas.1816847116. PMid: 31113877.
[54] Olivotto E, Otero M, Marcu KB, et al. Pathophysiology of osteoarthritis: Canonical NF-κB/IKKβ-dependent and kinase-independent effects of IKKα in cartilage degradation and chondrocyte differentiation. RMD Open 2015;1 (Suppl 1):e000061. https://doi.org/10.1136/rmdopen-2015-000061. PMid: 26557379.
[55] Schiene-Fischer C, Yu C. Receptor accessory folding helper enzymes: the functional role of peptidyl prolyl *cis/trans* isomerases. FEBS Lett. 2001;495(1–2):1–6. https://doi.org/10.1016/S0014-5793(01)02326-2. PMid: 11322937.
[56] Savvidou O, Milonaki M, Goumenos S, et al. Glucocorticoid signaling and osteoarthritis. Mol Cell Endocrinol 2019;480:153–66. https://doi.org/10.1016/j.mce.2018.11.001. PMid: 30445185.
[57] Hanada Y, Nakamura Y, Ishida Y, et al. Epiphycan is specifically expressed in cochlear supporting cells and is necessary for normal hearing. Biochem Biophys Res Commun 2017;492(3):379–85. https://doi.org/10.1016/j.bbrc.2017.08.092. PMid: 28864419.
[58] Ni GX, Li Z, Zhou YZ. The role of small leucine-rich proteoglycans in osteoarthritis pathogenesis. Osteoarthritis Cartilage 2014;22(7):896–903. https://doi.org/10.1016/j.joca.2014.04.026 PMid: 24795272.
[59] Melrose J, Fuller ES, Roughley PJ, et al. Fragmentation of decorin, biglycan, lumican and keratocan is elevated in degenerate human meniscus, knee and hip articular cartilages compared with age-matched macroscopically normal and control tissues. Arthritis Res Ther 2008;10(4):R79. https://doi.org/10.1186/ar2453. PMid: 18620607.
[60] Aigner T, Fundel K, Saas V, et al. Large-scale gene expression profiling reveals major pathogenetic pathways of cartilage degeneration in osteoarthritis. Arthritis Rheum 2006;54(11):3533–44. https://doi.org/10.1002/art.22174. PMid: 17075858.
[61] Young AA, Smith MM, Smith SM, et al. Regional assessment of articular cartilage gene expression and small proteoglycan metabolism in an animal model of osteoarthritis. Arthritis Res Ther 2005;7(4):R852. https://doi.org/10.1186/ar1756. PMid: 15987487.
[62] Zucker SN, Fink EE, Bagati A, et al. Nrf2 amplifies oxidative stress via induction of Klf9. Mol Cell. 2014;53(6):916–28. https://doi.org/10.1016/j.molcel.2014.01.033. PMid: 24613345.
[63] Lepetsos P, Papavassiliou AG. ROS/oxidative stress signaling in osteoarthritis. Biochim Biophys Acta. 2016;1862(4):576–91. https://doi.org/10.1016/j.bbadis.2016.01.003. PMid: 26769361.
[64] Hong J, Won M, Ro H. The molecular and pathophysiological functions of members of the LNX/PDZRN E3 Ubiquitin Ligase Family. Molecules 2020;25(24):5938. https://doi.org/10.3390/molecules25245938 PMid: 33333989.
[65] Hu T, Yang H, Han ZG. PDZRN4 acts as a suppressor of cell proliferation in human liver cancer cell lines. Cell Biochem Funct. 2015;33(7):443–9. https://doi.org/10.1002/cbf.3130 PMid: 26486104.
[66] Lu YL, Yang X, Liu YK. Reduced PDZRN4 promotes breast cancer progression and predicts poor prognosis. Int J Clin Exp Pathol 2019;12(1):142–53. PMid: 31933728.
[67] Hua Y, Ma X, Liu X, et al. Abnormal expression of mRNA, microRNA alteration and aberrant DNA methylation patterns in rectal adenocarcinoma. PLoS One 2017;12(3):e0174461. PMid: 28350845.
[68] Fernandes TL, Gomoll AH, Lattermann C, et al. Macrophage: a potential target on cartilage regeneration. Front Immunol 2020;11:111. https://doi.org/10.3389/fimmu.2020.00111. PMid: 32117263.
[69] Pawłowska J, Mikosik A, Soroczynska-Cybula M, et al. Different distribution of CD4 and CD8 T cells in synovial membrane and peripheral blood of rheumatoid arthritis and osteoarthritis patients. Folia Histochem Cytobiol. 2009;47(4):627–32. https://doi.org/10.2478/v10042-009-0117-9 PMid: 20430731.
[70] Pessler F, Chen LX, Dai L, et al. A histomorphometric analysis of synovial biopsies from individuals with Gulf War Veterans' Illness and joint pain compared to normal and osteoarthritis synovium. Clin Rheumatol. 2008;27(9):1127–34. https://doi.org/10.1007/s10067-008-0878-0 PMid: 18414968.
[71] Bommireddy R, Doetschman T. TGFβ1 and T_{reg} cells: alliance for tolerance. Trends Mol Med. 2007;13(11):492–501. https://doi.org/10.1016/j.molmed.2007.08.005 PMid: 17977791.
[72] Sakkas LI, Koussidis G, Avgerinos E, et al. Decreased expression of the CD3ζ chain in T cells infiltrating the synovial membrane of patients with osteoarthritis. Clin Diagn Lab Immunol. 2004;11(1):195–202. https://doi.org/10.1128/CDLI.11.1.195-202.2004. PMid: 14715568.
[73] Shirinsky I, Shirinsky V. H_1-antihistamines are associated with lower prevalence of radiographic knee osteoarthritis: a cross-sectional analysis of the Osteoarthritis Initiative data. Arthritis Res Ther. 2018;20(1):116. https://doi.org/10.1186/s13075-018-1619-7. PMid: 29880063.
[74] Ozeki N, Muneta T, Koga H, et al. Not single but periodic injections of synovial mesenchymal stem cells maintain viable cells in knees and inhibit osteoarthritis progression in rats. Osteoarthritis Cartilage 2016;24 (6):1061–70. https://doi.org/10.1016/j.joca.2015.12.018. PMid: 26880531.
[75] Haseeb A, Haqqi TM. Immunopathogenesis of osteoarthritis. Clin Immunol. 2013;146(3):185–96. https://doi.org/10.1016/j.clim.2012.12.011. PMid: 23360836.
[76] Shiokawa S, Matsumoto N, Nishimura J. Clonal analysis of B cells in the osteoarthritis synovium. Ann Rheum Dis 2001;60(8):802–5. https://doi.org/10.1136/ard.60.8.802. PMid: 11454647.
[77] Mödinger Y, Rapp AE, Vikman A, et al. Reduced terminal complement complex formation in mice manifests in low bone mass and impaired fracture healing. Am J Pathol 2019;189(1):147–61. https://doi.org/10.1016/j.ajpath.2018.09.011. PMid: 30339839.
[78] Klengel T, Binder EB. *FKBP5* allele-specific epigenetic modification in gene by environment interaction. Neuropsychopharmacology. 2015;40(1):244–6. https://doi.org/10.1038/npp.2014.208. PMid: 25482174.
[79] Smith E, Palethorpe HM, Ruszkiewicz AR, et al. Androgen receptor and androgen-responsive gene FKBP5 are independent prognostic indicators for esophageal adenocarcinoma. Dig Dis Sci 2016;61(2):433–43. https://doi.org/10.1007/s10620-015-3909-0. PMid: 26467701.
[80] Romano S, Staibano S, Greco A, et al. FK506 binding protein 51 positively regulates melanoma stemness and metastatic potential. Cell Death Dis 2013;4 (4):e578. PMid: 23559012.
[81] Jiang W, Cazacu S, Xiang C, et al. FK506 binding protein mediates glioma cell growth and sensitivity to rapamycin treatment by regulating NF-κB signaling pathway. Neoplasia 2008;10(3):235–43. https://doi.org/10.1593/neo.07929. PMid: 18320068.
[82] Li L, Lou Z, Wang L. The role of FKBP5 in cancer aetiology and chemoresistance. Br J Cancer 2011;104(1):19–23. https://doi.org/10.1038/sj.bjc.6606014. PMid: 21119664.

[83] Mfotie Njoya E, Eloff JN, McGaw LJ. *Croton gratissimus* leaf extracts inhibit cancer cell growth by inducing caspase 3/7 activation with additional anti-inflammatory and antioxidant activities. BMC Complement Altern Med 2018;18(1):305. https://doi.org/10.1186/s12906-018-2372-9. PMid: 30428879.

[84] Pei H, Li L, Fridley BL, et al. FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt. Cancer Cell 2009;16(3):259–66. https://doi.org/10.1016/j.ccr.2009.07.016. PMid: 19732725.