BMC Oral Health

RESEARCH

Open Access

Identification of core genes related to exosomes and screening of potential targets in periodontitis using transcriptome profiling at the single-cell level

Wufanbieke Baheti[1,2], Diwen Dong[1], Congcong Li[1] and Xiaotao Chen[1*]

Abstract

Background The progression and severity of periodontitis (PD) are associated with the release of extracellular vesicles by periodontal tissue cells. However, the precise mechanisms through which exosome-related genes (ERGs) influence PD remain unclear. This study aimed to investigate the role and potential mechanisms of key exosome-related genes in PD using transcriptome profiling at the single-cell level.

Methods The current study cited GSE16134, GSE10334, GSE171213 datasets and 19,643 ERGs. Initially, differential expression analysis, three machine learning (ML) models, gene expression analysis and receiver operating characteristic (ROC) analysis were proceeded to identify core genes. Subsequently, a core gene-based artificial neural network (ANN) model was built to evaluate the predictive power of core genes for PD. Gene set enrichment analysis (GSEA) and immunoinfiltration analysis were conducted based on core genes. To pinpoint key cell types influencing the progression of periodontal at the single-cell level, a series of single-cell analyses covering pseudo-time series analysis were accomplished. The expression verification of core genes was performed through quantitative reverse transcription polymerase chain reaction (qRT-PCR).

Results CKAP2, IGLL5, MZB1, CXCL6, and AADACL2 served as core genes diagnosing PD. Four core gene were elevated in the PD group in addition to down-regulated AADACL2. The core gene-based-ANN model had AUC values of 0.909 in GSE16134 dataset, which exceeded AUC of each core gene, highlighting the accurately and credibly predictive performance of ANN model. GSEA revealed that ribosome was co-enriched by 5 core genes, manifesting the expression of these genes might be critical for protein structure or function. Immunoinfiltration analysis found that CKAP2, IGLL5, MZB1, and CXCL6 exhibited positive correlations with most discrepant immune cells/discrepant stromal cells, which were highly infiltrated in PD. B cells and T cells holding crucial parts in PD were identified as key cell types. Pseudo-time series analysis revealed that the expression of IGLL5 and MZB1 increased during T cell differentiation, increased and then decreased during B cell differentiation. The qRT-PCR proved the mRNA expression levels of CKAP2 and MZB1 were increased in the blood of PD patients compared to controls. But the mRNA expression

*Correspondence:
Xiaotao Chen
xiaotaochen@163.com

Full list of author information is available at the end of the article

levels of AADACL2 was decreased in the PD patients compared to controls. This is consistent with the trend in the amount of expression in the dataset.

Conclusion CKAP2, IGLL5, MZB1, CXCL6 and AADACL2 were identified as core genes associated with exosomes, helping us to understand the role of these genes in PD.

Keywords Periodontitis, Exosomes, ScRNA-seq, Machine learning, Core genes

Introduction

Periodontitis (PD) is considered to be one of the most prevalent diseases in humans [1] and ranks as the sixth most common global disease [2]. There is a noticeable upward trend in the prevalence of PD worldwide, with a 57.3% increase from 1990 to 2010 [3]. PD stands as the primary cause of adult tooth loss globally, impacting nutrient intake, quality of life, and self-esteem while also resulting in significant socio-economic implications and wastage of medical resources [4]. Dental calculus on the root surface and subgingival plaque are identified as crucial local factors contributing to the occurrence and progression of PD [5]. Furthermore, endotoxins produced by periodontal pathogens have been closely linked to systemic diseases. Research indicates that PD interacts with several systemic conditions including stroke, atherosclerosis, chronic kidney disease, cardiovascular disease, respiratory infections, and type 2 diabetes [6, 7]. Therefore, the treatment of PD is important for the prevention of these diseases. In addition, the loss of alveolar bone caused by periodontal pathogenic bacteria through host immune inflammation is usually difficult to repair naturally. How to address the inflammatory destruction and absorption of alveolar bone and promote reconstruction and repair is a challenge in the treatment of PD [8, 9].

In recent years, tissue engineering technology has focused on periodontal tissue regeneration. By means of various exogenous implantation of stem cells, new bone tissue can be generated in areas with existing bone defects to replace damaged tissues, thereby inhibiting progressive absorption of periodontal bone tissue and restoring healthy function of periodontal tissues [10]. Some studies have found that cell-to-cell communication greatly influences bone tissue repair. Exosomes can serve as mediators for intercellular communication and have been applied in bone tissue engineering [11]. Exosomes (EXOs) as important mediators of paracrine effects retain almost all advantages of source cells. Compared to stem cell therapy, EXOs have lower immunogenicity and better biocompatibility, resulting in lower post-transplantation immune-related adverse reactions [12–14], allowing them to mimic source cells to some extent and exert functions in modulating microenvironment. And the composition of the contents of EXOs from different types of cells is different, even for the same type of cells, their secreted EXOs can have a high degree of heterogeneity due to the different environments they are in. Exosomes carry various bioactive molecules such as miRNA, proteins, and growth factors. These molecules can regulate inflammation response and immune response in surrounding tissues, thereby affecting the development and severity of PD. ERG plays a pivotal role in regulating the exosome-mediated effects of these processes, rendering it a potential target for pharmacological interventions or therapies aimed at enhancing bone tissue repair or controlling periodontitis. Hence, it is imperative to investigate the mechanism underlying ERG's involvement in periodontitis progression by examining its specific alteration pattern during disease development and its impact on inflammation and immune response.However, the potential mechanisms of exosome-related genes in PD have not been fully elucidated.

This study employed a comprehensive array of bioinformatics approaches to elucidate the diagnostic significance of ERGs in PD. Furthermore, it investigated the molecular mechanisms, immune microenvironment, regulatory networks, and associated pharmacological agents pertaining to key genes influencing periodontitis. Additionally, single-cell analysis was utilized to pinpoint critical cell types that play pivotal roles in PD. These findings offer novel insights for therapeutic strategies targeting periodontitis.

Materials and methods

Data collection

Two PD-related transcriptome datasets, GSE16134 and GSE10334, from the GPL570 annotation platform were garnered from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) database. Specifically, the GSE16134 dataset encompassed microarray sequencing of from gingival tissue specimens from 120 patients with PD ($N_{\text{PD affected site (PD group)}}$: $N_{\text{Unaffected site (control)}}$=241:69). The GSE10334 dataset embraced microarray sequencing data of gingival tissue specimens from 90 patients with PD ($N_{\text{PD affected site}}$: $N_{\text{Unaffected site}}$ = 183:64). Single cell RNA sequencing (scRNA-seq) dataset GSE171213 (platform: GPL24676), collected from the GEO database as well, comprised high-throughput sequencing data of periodontal tissue samples from 5 patients with chronic PD and 4 healthy controls. A total of 19,643 exosome-related genes (ERGs) associated with blood sources were acquired from the exoRbase database

(http://www.exorbase.org/) (Table S1). In order to clearly demonstrate the complex research process of this study, we constructed a flowchart (Additional file 1) which had outlined the key aspects of the process from data collection to analysis of results.

Identification and functional analysis of candidate genes

The differentially expressed genes (DEGs) in the PD and control samples of the GSE16134 dataset were pinpointed through differential expression analysis with the R package 'limma' (Ver. 3.54.0) [15] ($|Log_2$fold-change (FC)$| > 2$, $p < 0.05$). A volcano plot and a expression heatmap illustrating the over-expressed and down-regulated DEGs, sorted in accordance with the value of log_2FC, were created applying R packages 'ggplot2' (Ver. 3.4.4) [16] and 'ComplexHeatmap' (Ver. 2.15.1) [17]. Next, candidate genes associated with exosome in PD were identified by intersecting DEGs with 19,643 ERGs through the R package 'ggvenn' (Ver. 0.1.9) [18]. To probe into the functional roles of candidate genes, Gene Ontology (GO) terms covering biological processes (BPs), molecular functions (MFs), and cellular components (CCs) as well as Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched by candidate genes were noted *via* the R package 'clusterProfiler' (Ver. 4.7.1.003) [19] based on the GO (https://geneontology.org/) and KEGG (https://www.kegg.jp/) databases ($p < 0.05$). Besides, the interactions of candidate genes at the protein level were investigated by establishing a protein-protein interaction (PPI) network with the help of the Search Tool for the Retrieval of Interacting Genes (STRING, https://string-db.org) database (Confidence level > 0.4).

Screening for core genes

In GSE16134 dataset, in sum of 3 machine learning (ML) model-based screening was proceeded to identify feature genes from candidate genes. Least absolute shrinkage and selection operator (LASSO) regression model was constructed by the R package 'glmnet' (Ver. 4.1-6) [20]. Candidate genes whose regression coefficients were not penalised to 0 when the model lambda value was at its minimum were deemed as LASSO-featured genes. Boruta algorithm with default parameters was performed with the help of the R package 'Boruta' (Ver. 8.0.0) [21], where candidate genes with low correlation were iteratively removed, and candidate genes with Z-values greater than the maximum Z-value of the shaded feature were identified as Boruta-featured genes. Additionally, the R package 'caret' (Ver. 6.0–93) [22] was proceeded to build a Support Vector Machine-Recursive Feature Elimination (SVM-RFE) model, which retained SVM-RFE-featured genes by removing features with the smallest scores in each iteration. Finally, the PD features that were retained in all 3 models were identified as the feature genes. In GSE16134 and GSE10334 datasets, discrepancies in the expression of feature genes between PD and control samples of the 2 datasets were assessed through Wilcoxon rank-sum test ($p < 0.05$). Feature genes with significantly discrepant expression and consistent expression trends in the PD and control samples of the 2 datasets were considered as candidate core genes. Receiver operating characteristic (ROC) analysis were subsequently proceeded to select core genes from candidate core genes. ROC curves corresponding to each feature genes were drawn through the R package 'pROC' (Ver. 1.18.5) [23] in the 2 datasets, and candidate core genes with area under curve (AUC) values above 0.7 within both datasets were selected as core genes.

Artificial neural network (ANN) modeling

Furthermore, an ANN, a type of ML model, was created to evaluate the predictive power of core genes for PD in GSE16134 dataset. First, each core gene was assigned an expression value of '0' or '1' based on its median expression, generating the high expression group and low expression group. The expression matrix was input into R package 'neuralnet' (Ver. 1.44.2) [24] to construct an ANN network in the GSE16134 dataset. Plotted ROC curves using the R package 'pROC' (Ver. 1.18.5) to assess the overall predictive power of the ANN model in the training set GSE16134 and validation set GSE10334.

Gene set enrichment analysis (GSEA)

To explore the potential functions played by the core genes in PD, GSEA was employed to clarify the KEGG pathways involved by core genes through referring to the c2.cp.kegg.v2023.2.Hs.symbols.gmt gene set, which was cited from the Molecular Signatures Database (MSigDB, https://dsigdb.tanlab.org/DSigDBv1.0/). Concretely, Spearman correlation analysis was conducted between the core genes and the remaining genes within the GSE16134 dataset using the R package 'psych' (Ver. 2.2.9) [25] to obtain correlation coefficients. These coefficients were sorted in descending order, and ranked genes were subsequently analyzed using the 'GSEA' function within the R package 'clusterProfiler' ($|$Normalized Enrichment Score (NES)$| > 1$, $p < 0.05$).

Immune infiltration analysis and drug prediction

To investigate the discrepancies between the immune microenvironment in periodontal samples from PD invasion and healthy periodontal samples of PD patients, the abundance of 28 immune cells [26] and 2 types of stromal cells (fibroblasts and endothelial cells) [27] was assessed in PD and control samples of the GSE16134 dataset adopting single-sample GSEA (ssGSEA) algorithm in 'GSVA' package (Ver. 1.42.0) [28]. Discrepancies in immune cell infiltration between PD and controls

were also compared by Wilcoxon rank-sum test, gaining discrepant immune cells and discrepant stromal cells ($p<0.05$). The correlation between core genes and discrepant immune cells/stromal cells was explored through Spearman correlation analysis ($|cor| > 0.3$, $p<0.05$). To develop latent therapeutics for PD, potential therapeutic compounds targeting core genes were obtained in the Drug Signatures Database (DSigDB, http://ctdbase.org/) and a drug-biomarker network was synthesized by means of Cytoscape software (Ver. 3.7.0) [29].

Establishment of core gene-related networks

The GeneMANIA platform (http://genemania.org/), was leveraged to construct a gene-gene interaction (GGI) network, which aided in elucidating the functional linkages between genes that maintained similar functions to the core genes. In order to understand the underlying regulatory mechanism of the role of core genes in PD. The upstream transcription factors (TFs) and miRNAs targeting core genes were predicted by the miRNet database (https://www.mirnet.ca/miRNet/home.xhtml), and then the Cytoscape software was used to integrate the TFs and miRNAs, creating a TF-mRNA-miRNA network. In-depth characterisation of the RBP-RNA interactome will contribute to a better understanding of disease pathogenesis. Here, therefore, RBPs interacting with core genes were predicted using the ENCORI database (https://starbase.sysu.edu.cn/) and the core gene-RBP regulatory network was visualised through Cytoscape software.

Tissue gene expression analysis and gene pathway analysis

The expression levels of the core genes in different tissue types of the human body were analysed by the GTEx, BioGPS, and SAGE platforms in the GeneCards database (https://www.genecards.org/), which helped to shed light on the pathogenesis of PD. In order to further investigate the association between the core genes and their enriched top 1 KEGG pathway. Based on all PD affected samples from the GSE16134 dataset, Pearson correlation analysis was conducted *via* the 'cor' function in the R package (Ver. 4.2.2) to estimate the correlations between each core gene and their enriched top 1 pathway. The correlations between core genes and the genes in top 1 pathway were explored through Spearman correlation analysis *via* the 'cor' function. Scatter plots and bubble plots were drawn through R packages 'ggstatsplot' (Ver. 0.11.0) [30] and 'ggplot' (Ver. 3.4.4) [31] to show the results.

ScRNA-seq data preprocessing and cell annotation

To pinpoint key cell types influencing the progression of PD in patients at the single-cell level, the scRNA-seq dataset, GSE171213, were opted for a comprehensive series of single-cell analyses. In order to normalise the raw data of the GSE171213 dataset, low-quality cells and features were eliminated by quality control (QC) through the 'CreateSeuratObject' function of the R package 'Seurat' (Ver. 5.0.1) [32]. The percentage of mitochondrial genes in a single cell was calculated by the PercentageFeatureSet function. Specifically, cells with more than 20% mitochondrial genes and more than 5% erythroid genes were filtered. Cells with the number of RNA features detected in a single cell ranging from 100–6000, with a total number of RNA molecules less than 10,000 and with a percentage of mitochondrial genes less than 20% were retained. After the data were normalized by the 'NormalizeData' and 'FindVariableFeatures' functions, the first 2,500 highly variable genes (HVGs) were singled out with the help of the 'FindVariableFeatures' function. A scatter plot showing the standardized variance of these HVGs was created by the 'LabelPoints' function. The data were further standardized by 'ScaleData' function to ensure that the cells analysed belonged to different cell clusters. In compliance with the compliance with the first 2,500 HVGs, principal component analysis (PCA) was proceeded through the 'RunPCA' function to ensure consistency in the distribution of cells across samples and check for any obvious outliers. Subsequently, the optimal principal components (PCs) were ascertained through the 'JackStrawPlot' function ($p<0.05$), and the results were displayed by an elbowplot mapped via 'EIbowPlot' function. Unsupervised cluster analysis was performed utilizing Uniform Manifold Approximation and Projection (UMAP) by 'NormalizeData' and 'FindVariableFeatures' functions with a set resolution parameter of 1 to identify cell clusters. The marker genes corresponding to each cluster were gained *via* 'FindAllMarkers' function. Differentially expressed marker genes (DEMGs) in discrepant cell clusters were identified by Wilcoxon rank-sum test ($|\log_2 FC|>1$, adj.$p<0.05$). Cell types were obtained by annotating different cell clusters *via* R package 'SingleR' (Ver. 2.0.0) [32] in accordance with the expression of marker genes. A histogram was utilized to graphically represent the percentage of cell types in both PD and control groups.

Identification of key cell types and analysis of expression levels of core genes in cell types

The proportion of cell types in PD and control groups of GSE171213 dataset was estimated, and differences in proportion of each cell type between two groups were compared adopting the chi-square test, gaining discrepant cell types. Discrepant cell types that had been reported to hold a key part in PD were deemed as key cell types ($p<0.05$). The expression of core genes in cell types was examined, as well as in cell types of PD and controls.

Secondary clustering of key cell types and pseudo-time series analysis

To further explore the heterogeneity of the key cell types, these key cell types were clustered by UMAP for secondary dimensionality reduction clustering in the dataset GSE171213, acquiring molecular subtypes corresponding to key cell types. In order to study the differentiation trajectories of key cell types and the expression dynamics of core genes during their differentiation, pseudo-time series analysis was proceeded by the R package 'Monocle 3' (Ver. 1.0.0) [33] based on molecular subtypes corresponding to key cell types.

Expression verification of core genes

The expression of core genes was verified in 6 PD samples and 6 control samples by quantitative reverse transcription polymerase chain reaction (qRT-PCR). These samples collected with the approval of the patients and the Ethics Committee of People's Hospital of Xinjiang Uygur Autonomous Region, China. All participants were given informed consent. This study gets the approval of the People's Hospital of Xinjiang Uygur Autonomous Region, China ethics committee (approval number: KY20240312022). Total RNA was extracted from blood using Trizol reagent (Thermo Scientific Ambion, China). SureScript-First-strand-cDNA-synthesis-kit (Servicebio, Wuhan, China) was used to reverse mRNA transcription into cDNA. Reverse transcription-quantitative real-time polymerase chain reaction (RT-qPCR) was performed on the mRNA expression levels of *CKAP2*, *MZB1*, *AADACL2*, *IGLL5*, and *CXCL6*. The amplification reaction conditions were as follows: predenaturation at 95 °C for 60 s, followed by 40 cycles of denaturation at 95 °C for 20 s, 55 °C for 20 s, and 72 °C for 30 s. The relative expression levels of mRNA were calculated by the $2^{-\Delta\Delta CT}$ method with the internal reference values of GAPDH. The primer sequences used in the experiment were displayed in (Table S2).

Statistical analysis

In this study, we conducted bioinformatics analyses applying R software (Ver. 4.2.2). Inter-group discrepancies were evaluated with the Wilcoxon rank-sum test, identifying statistical significance at p-values below 0.05.

Results

The biological functions of candidate genes might be related to immune responses and inflammatory processes

Differential expression analysis of the GSE16134 microarray sequencing data yielded a number of 23 DEGs between PD and controls, characterized by 21 over-expressed and 2 down-regulated DEGs in PD (Fig. 1A-B). They were overlapped with 19,643 ERGs, obtaining 9 candidate genes associated with exosome and PD, namely CKAP2, IGLL5, MZB1, SPAG4, TNFRSF17, IRF4, CXCL6, AADACL2, and DSC1 (Fig. 1C). These 9 candidate genes mainly participated in 103 GO-BPs, 13 GO-CCs and 5 GO-MFs. The first 5 entries of BP, CC, and MF were presented in Fig. 1D. Of note, IGLL5I and CXCL6 were involved in identical BPs such as defense response to bacterium, leukocyte mediated immunity and humoral immune response. The identical BPs enriched by IGLL5 and MZB1 encompassed regulation of B cell activation and B cell activation (Table S3). In total, 8 KEGG pathways were identified for the candidate genes. Five of them were enriched by CXCL6, including IL-17 signaling pathway, TNF signaling pathway and rheumatoid arthritis, pertussis, viral protein interaction with cytokine and cytokine receptor. Cytokine-cytokine receptor interaction was co-enriched by TNFRSF17 and CXCL6. In addition, the pathways enriched for TNFRSF17 and IRF4 were intestinal immune network for IgA production and Th17 cell differentiation, respectively (Fig. 1E, Table S4). Summarily, the biological functions of candidate genes might be related to immune responses and inflammatory processes, such as cytokine signalling, immune cell activation, antibody production, and defence responses to infectious agents. A PPI network of 9 candidate genes was synthesised, the network demonstrated protein-level interactions of 4 candidate genes in addition to the discrete proteins encoded by 5 genes, forming 6 interaction pairs. MZB1, as an internal gene, maintained a close association with IRF4, TNFRSF17, and IGLL5 (Fig. 1F).

The core gene-based-ANN model could distinguish PD from controls accurately

The precision of LASSO model achieved the highest when minimum Lambda value was 0.08445074 and 5 LASSO-featured genes whose regression coefficients were not 0 were retained (CKAP2, IGLL5, MZB1, CXCL6, and AADACL2) (Fig. 2A). Boruta analysis found that the Z-values of all candidate genes exceeded the maximum Z-value of the shaded feature, so these 9 genes served as Boruta-featured genes (PCXCL6, CKAP2, IGLL5, SPAG4, IRF4, MZB1, AADACL2, DSC1, and TNFRSF17) (Fig. 2B). Similarly, through the iterative screening of the SVM-RFE model, none of the 9 candidate genes were eliminated, transforming into SVM-RFE-featured genes (Fig. 2C). Hence, CKAP2, IGLL5, MZB1, CXCL6, and AADACL2, noted as feature genes, were unquestionably deemed as PD features (Fig. 2D). Gene expression analysis manifested that CKAP2, IGLL5, MZB1 and CXCL6 were elevated in the PD group in addition to down-regulated AADACL2 and they kept fully consistent expression trends in GSE16134 and GSE10334 datasets ($p < 0.0001$). As a result, all 5 feature genes could be passed as candidate core genes (Fig. 2E, Fig.S1A). ROC analysis ascertained candidate core genes as core genes

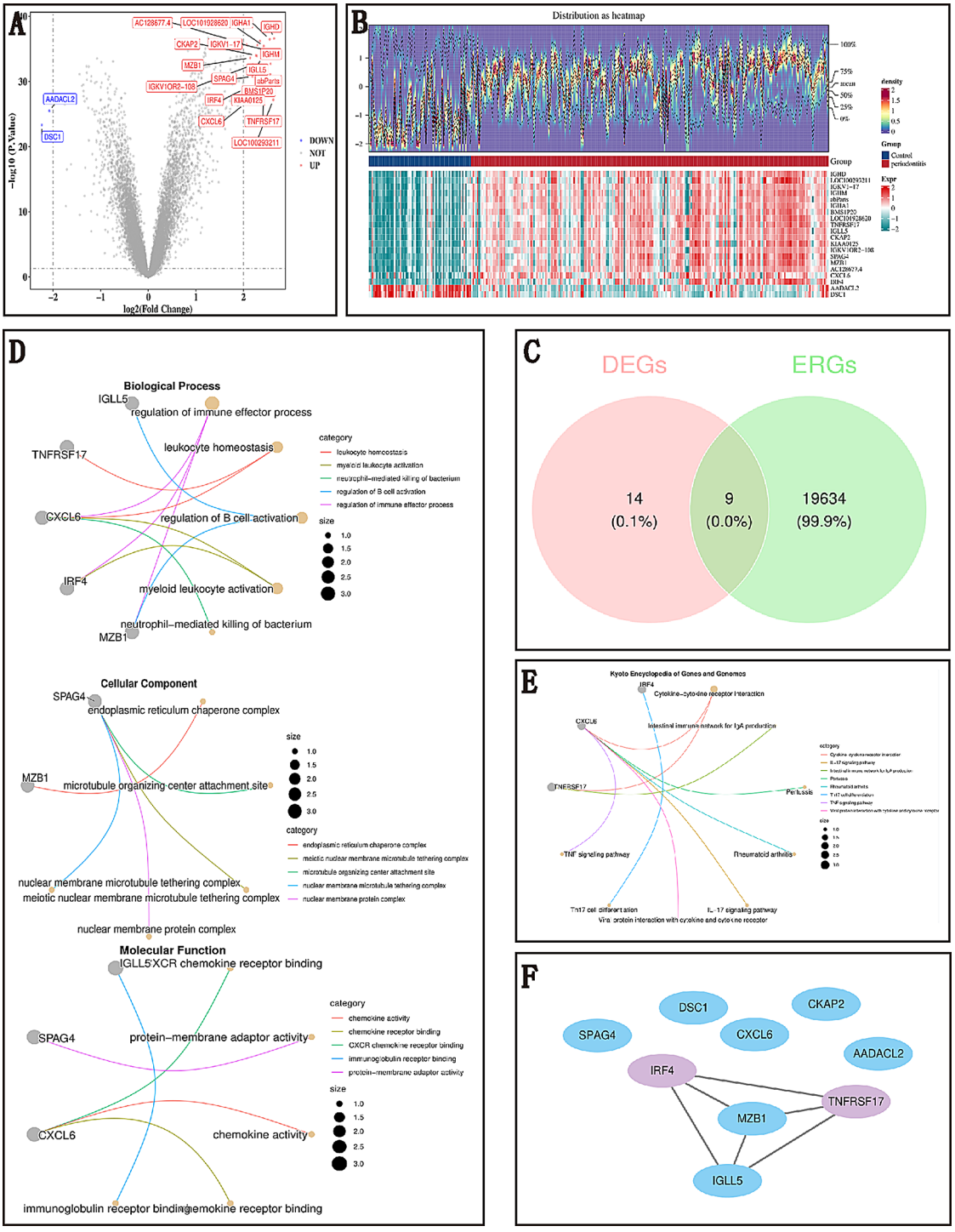

Fig. 1 Differential expression analysis. (**A**) Volcano map of differentially expressed genes (Red dots represent up-regulated genes, blue dots represent down-regulated genes; The 10 genes with the most significant up-down-regulation differences showed gene names). (**B**) Heat map of differentially expressed genes (The picture consists of two parts: The upper part is the expression density heat map of down-regulated top10 genes on the sample, showing the five quantiles and the average Lines of value; The next part is the heat map of down-regulated top10 genes on the sample). The data in this Fig. 1A, B were obtained by differential expression analysis of the dataset GSE16134. (**C**) Identification of candidate genes. (**D**) GO enrichment results (Gray dots represent genes, yellow dots represent different pathway names, and different color lines represent pathways enriched by genes).This figure was obtained by GO enrichment analysis of differentially expressed genes (**E**) KEGG enrichment results (Gray dots represent genes, yellow dots represent different pathway names, and different color lines represent pathways enriched by genes). This figure was obtained by KEGG enrichment analysis of differentially expressed genes. (**F**) PPI network diagram

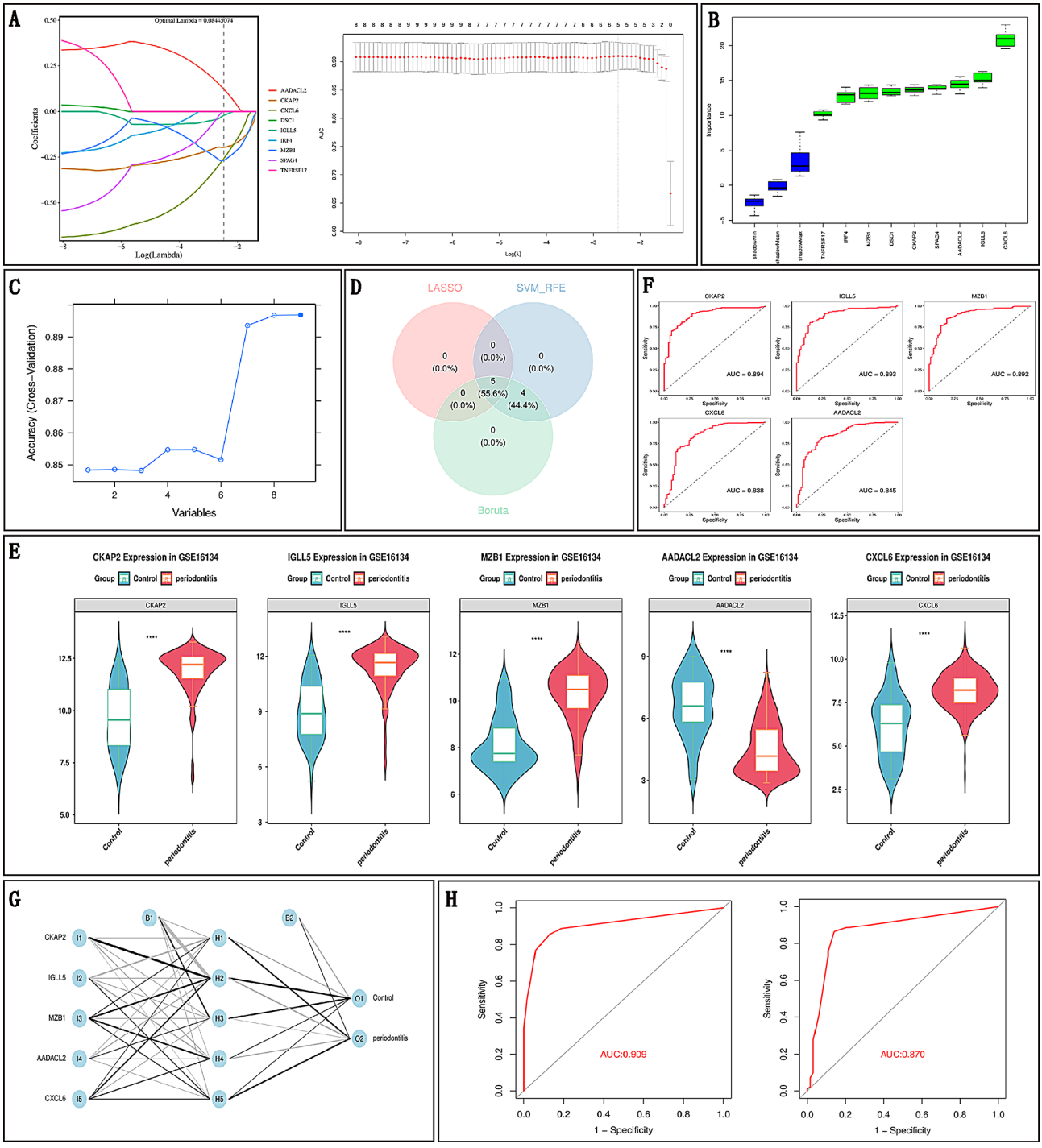

Fig. 2 Machine learning screens for core genes. (**A**). LASSO coefficient spectrum and Cross-validation of LASSO regression analysis. (**B**) Boruta algorithm identifies candidate core genes. (**C**) SVM-RFE screening candidate biomarkers. (**D**) The intersection of three algorithms is used to obtain the feature gene. (**E**) Expression of key feature genes in GSE16134 data sets. The ordinate is gene expression, blue is Control group, red is RIF group; ns is not significant; *, P.value < 0.05; **, P. value < 0.01; ***, P. value < 0.001; ****, P. value < 0.0001). (**F**) ROC curves of core genes in the GSE16134 datasets. (**G**) Artificial neural network diagram of core genes in training set. (**H**) ROC curve of artificial neural network in training set (GSE16134 datasets) and validation set (GSE10334 datasets) (AUC is the area under the ROC curve, and the closer the value of AUC is to 1, the better the prediction performance of the model)

because of their AUCs of greater than 0.80 in GSE16134 and GSE10334 datasets (Fig. 2F, Fig.S1B). The core gene-based-ANN model was developed in GSE16134 datasets, it could specifically distinguish between PD and control samples (Fig. 2G), the AUCs of the model in the GSE16134 and GSE10334 datasets were 0.909 and 0.870 correspondingly, which exceeded AUC of each core gene, highlighting that the predictive performance of ANN model was accurate and credible (Fig. 2H).

Identification of core gene function

GSEA revealed 66, 56, 54, 75, and 60 KEGG pathways involved by CKAP2, IGLL5, MZB1, AADACL2, and CXCL6 5 core genes respectively. In top 5 pathways, CKAP2, IGLL5, and AADACL2 might play a role in a selfsame process like intestinal immune network for IgA production. In addition to IGLL5, the remaining 4 core genes might also be involved in cell adhesion molecules (CAMs). Ribosome was co-enriched by 5 core genes, manifesting the expression of these genes might be critical for protein structure or function (Fig. 3A-E, Table S5).

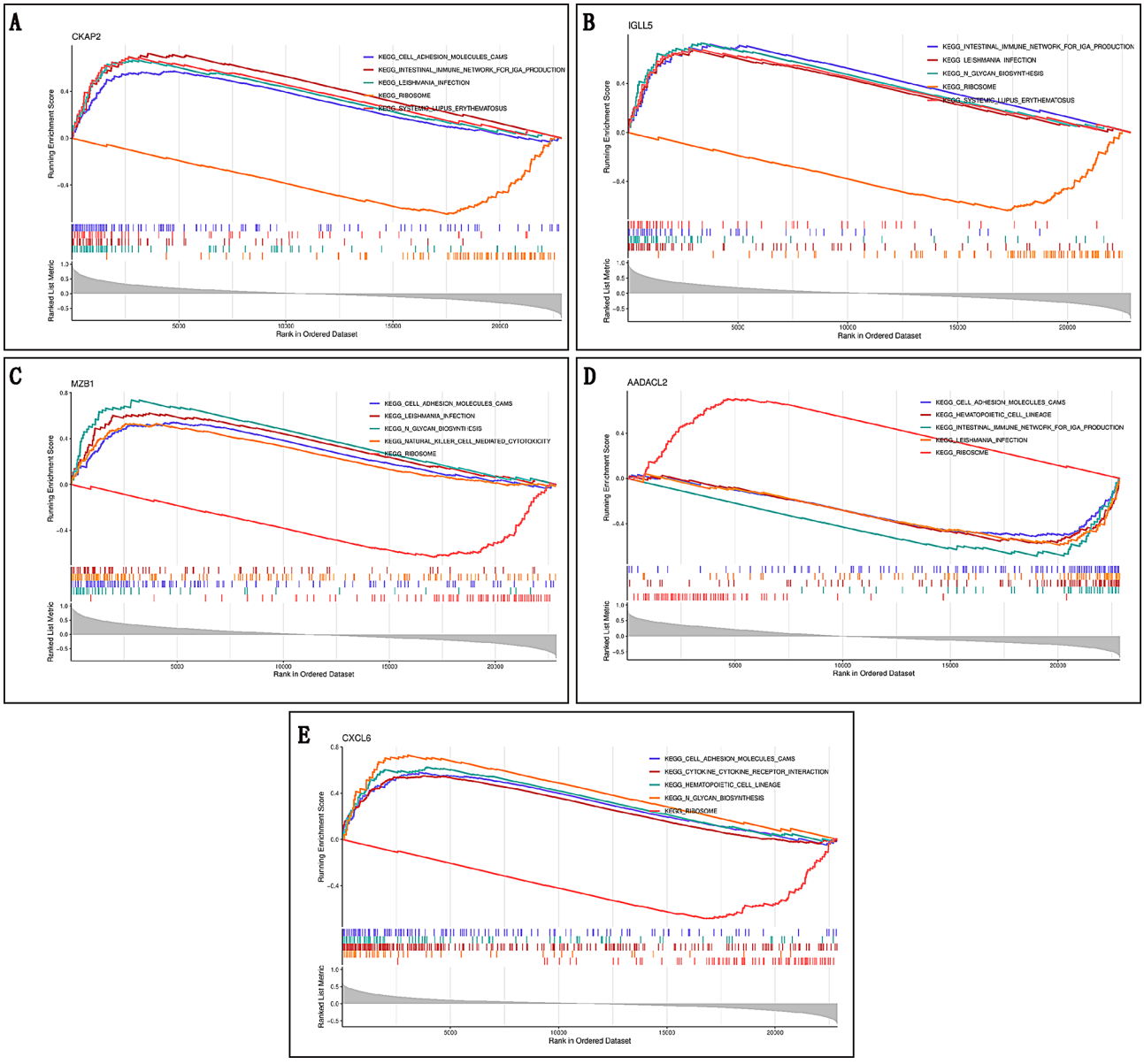

Fig. 3 GSEA analysis. (**A**) CKAP2 enrichment results. (**B**) IGLL5 enrichment results. (**C**) MZB1 enrichment results. (**D**) AADACL2 enrichment results (**E**) CXCL6 enrichment results. This figure was obtained by GSEA analysis of key genes (CKAP2, IGLL5, MZB1, AADACL2, CXCL6)

CKAP2, IGLL5, MZB1, and CXCL6 exhibited positive correlations with most discrepant immune cells/discrepant stromal cells

In addition to 1 adaptive immune cell (memory B cell), 3 types of innate immune cells (T helper Type 2 (Th2) cells, CD56bright natural killer (NK) cells, and CD56dim NK cells), totally 24 kinds of immune cells (discrepant immune cells) and 2 stromal cells (discrepant stromal cells) varied in abundance in the PD and control samples of the GSE16134 dataset. Among these, 23 kinds of discrepant immune cells and 2 kinds of discrepant stromal cells accounted for a relatively high proportion in PD except effector memeory CD4 T cells ($p<0.05$) (Fig. 4A). AADACL2 showed a negative correlation with the majority of discrepant immune cells/discrepant stromal cells, whereas CKAP2, IGLL5, MZB1, and CXCL6 exhibited positive correlations with most discrepant immune cells/discrepant stromal cells ($|cor| > 0.3$, $p<0.05$) Activated B cells exhibited the strongest positive association and the strongest negative association with MZB1 ($cor=0.89$, $p<0.05$) and AADACL2 ($cor = -0.50$, $p<0.05$) (Fig. 4B). By means of DSigDB, in sum of 13, 2 and 23 drugs interacting with MZB1, CKAP2 and CXCL6, respectively were predicted. IGLL5 and AADACL2 were not predicted to corresponding drugs. As a consequence, the drug-biomarker network was constituted of MZB1,

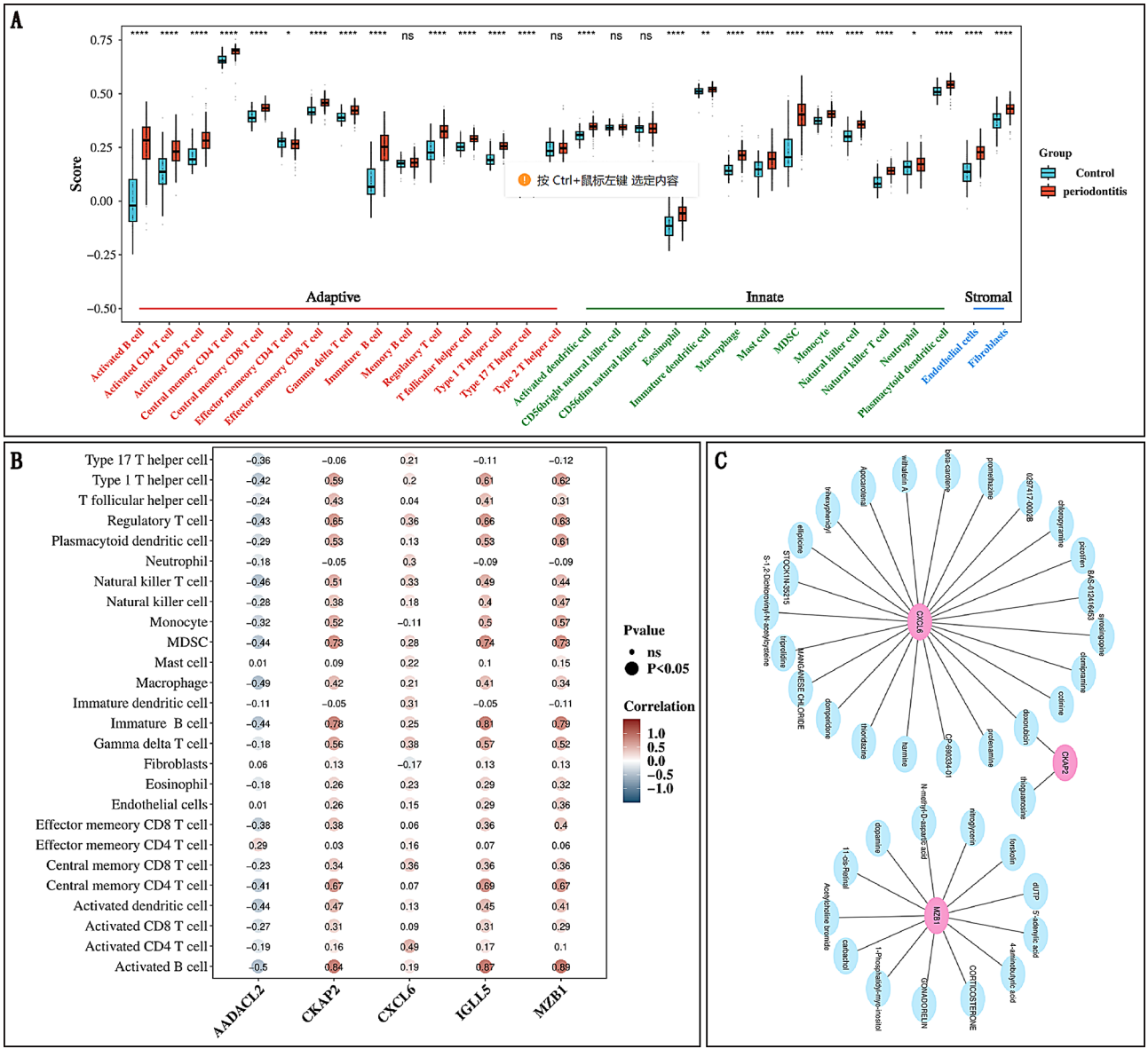

Fig. 4 Immune infiltration analysis and drug prediction network construction. (**A**) The difference of immune infiltrating cells in different tissue samples of the training set. This figure was evaluated in GSE16134 datasets using single-sample gene set enrichment analysis, and then obtained using the Wilcoxon test. (**B**) Correlation analysis between core genes and differential immune cells. (**C**) Key gene-drug interaction network

CKAP2, CXCL6 and their corresponding drugs. Of note, doxorubicin (DXR) could interact with CKAP2 and CXCL6 (Fig. 4C).

Multiple genes, miRNAs, TFs, and RBPs, could interact with core genes

A search of the GeneMANIA database uncovered 20 genes with biological functions akin to our core genes, the majority of which exhibited intricate interactions. Notably, CXCL6 and CXCL1 were identified for their potential analogous roles, particularly in cytokine activity (Fig. 5A). A TF-mRNA-miRNA network constituting of 5 core genes, 67 miRNAs, and 40 TFs was established to reveal the regulatory mechanisms underlying core gene function. hsa-mir-185-5p, MYC, CNOT3, TP53 and TRIM28 might regulate the expression of MZB1 and CKAP2. The expression of CKAP2 and CXCL6 might also be affected by hsa-mir-129-2-3p, hsa-mir-130a-3p, hsa-mir-27a-3p, hsa-mir 106a-5p, hsa-mir-941 and SOX2 (Fig. 5B). We further revealed the RBPs interacting with core genes with the help of ENCORI database, and a total of 131 RBPs were predicted, which interacted with their respective targeted core genes, forming 144 interaction pairs. As presented in the gene-RBP regulatory network, the same RBP could interact with different core genes, e.g. RBM4 could interact with MZB1 and CKAP2, RBPMS could interact with CKAP2 and CXCL6 (Fig. 5C).

Core genes actively expressed in most of the tissues

The expression levels of the core genes in different tissue types of the human body were examined through databases. According to the GTEx database, the core genes actively expressed in most of the tissues. CKAP2 expressed in all the tissues (Fig. 6A). IGLL5 and AADACL2 are highly expressed in lymph node, small intestine, spleen, adrenal gland, etc. (Fig. 6B-C). MZB1 highly expressed in skin tissue. MZB1 primarily expressed in skin tissue, and CXCL6 was mainly expressed in spleen and lung tissues (Fig. 6D-E). These 5 core genes were positively correlated with the co-enriched top 1 pathway, ribosome (Fig. 6F). CKAP2, IGLL5 and MZB1 exhibited negative associations with genes in ribosome, yet AADACL2 was positively correlated with these genes (Fig. 6G).

B cells and T cells holding crucial parts in PD progress were identified as key cell types

Post-QC, a single-cell dataset encompassing 24,540 features across 22,126 cells was obtained. Violin plots illustrated the distribution changes in nFeature_RNA, nCount_RNA, and percent.mt pre- and post-QC (Fig.

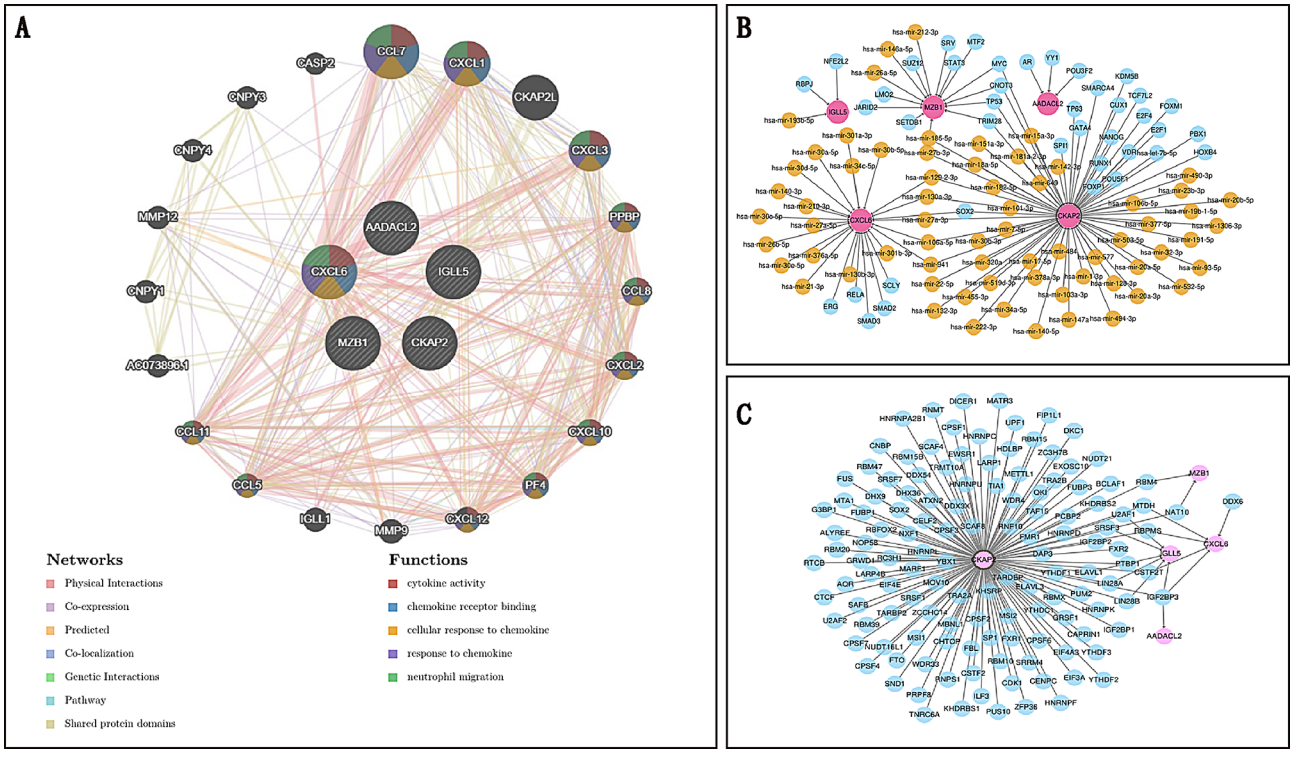

Fig. 5 The GeneMANIA Network. (**A**) GeneMANIA database analysis network diagram (The big circle in the middle is the key gene, and the small circle on the outside is the gene related to the key gene. The network lines on the right represent each other from top to bottom Action, co-expression, prediction, co-localization, genetic interaction, pathway, shared protein domain). (**B**) Molecular regulatory network of core genes (Rose red is the key gene; Yellow is miRNA; Blue is TF). (C) Key gene -RBP regulatory network (Pink is the key gene; Blue is RBP)

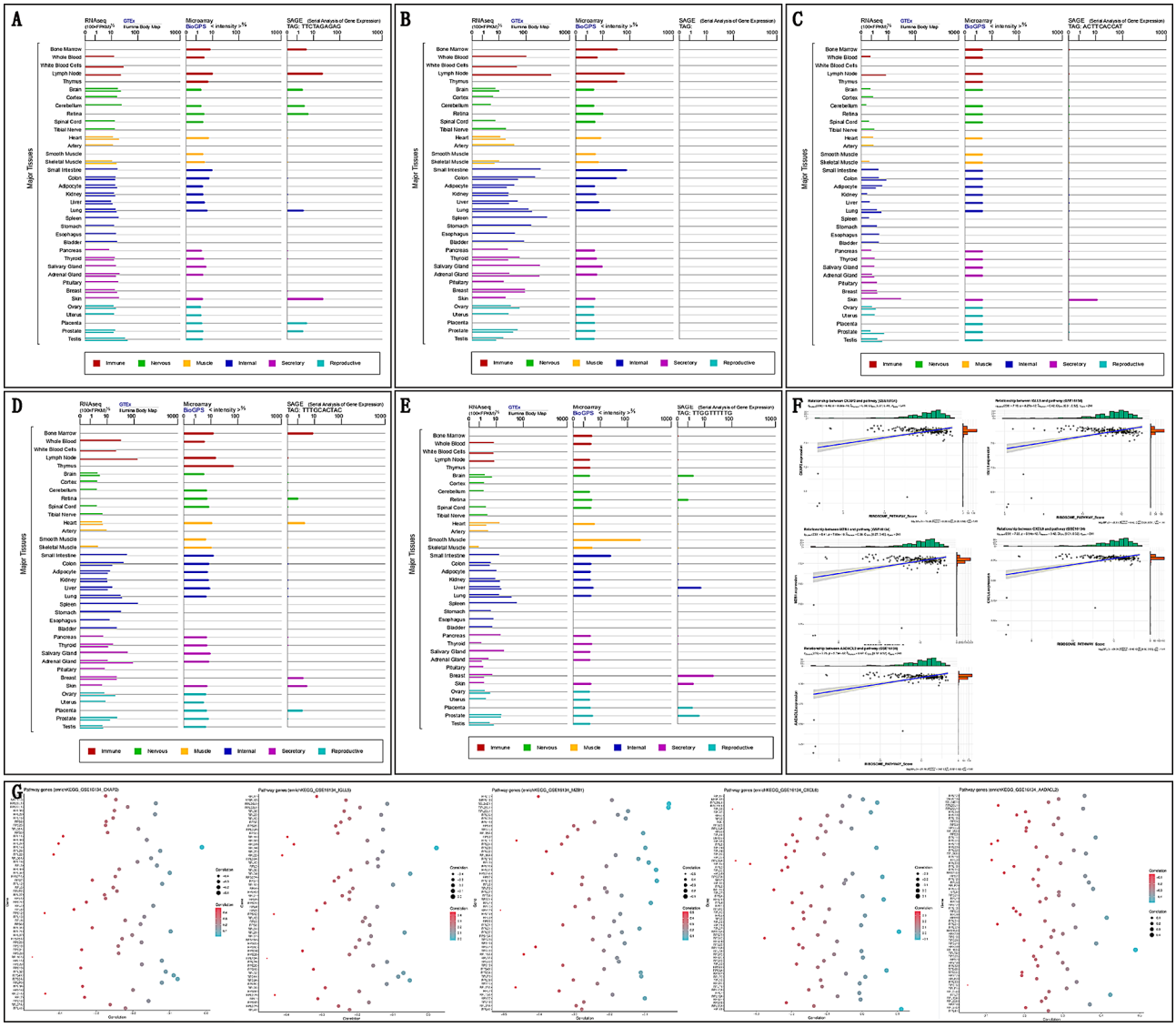

Fig. 6 Core genes actively expressed in most of the tissues. (**A-E**) Expression of CKAP2, IGLL5, MZB1, CXCL6 and AADACL2 in normal human tissues from GTEx, BioGPS, and SAGE. (**F**) Correlation analysis of key gene (CKAP2, IGLL5, MZB1, CXCL6 and AADACL2) and its enrichment TOP1 pathway score. (**G**) Correlation analysis of key gene (CKAP2) and its enrichment TOP1 pathway

S2 A-B). The top 30 HVGs were identified based on standardized variance (Fig.S2C). PCA confirmed the consistency of cell distributions across the GSE171213 dataset samples, indicating no distinct niche samples, with the dataset integrating well within the initial 30 PCs (Fig.S2D). The elbow plot demonstrated stabilization of standard deviation for these components, leading to the selection of the first 30 PCs for subsequent analysis (Fig. 7A). Utilizing the standard UMAP analysis in Seurat, 26 cell clusters were discerned and annotated into 11 distinct cell types B cells, chondrocytes, dendritic cells (DCs), endothelial cells, epithelial cells, monocytes, myelocytes, neutrophils, natural killer (NK) cells, T cells, tissue stem cells] according to marker gene expression (Fig. 7B-E). The percentage of endothelial cells and epithelial cells was much higher, yet the percentage of NK cells and myelocytes was much lower in PD samples compared to controls (Fig.S2 F-G). In addition to chondrocytes, other 10 cell types were identified as discrepant cell types between PD and control groups of GSE171213 dataset (Table 1).

Pseudo-time series analysis revealed the dynamic expression of core genes in B cells and T cells

B cells and T cells served as key cell types as a result of their crucial parts in PD progress [34–36]. AADACL2 was not detected in the GSE171213 dataset expression matrix. By observing the expression of core genes in various cell types, it was not difficult to find that MZB1 and IGLL5 actively expressed in B cells, CKAP2 highly

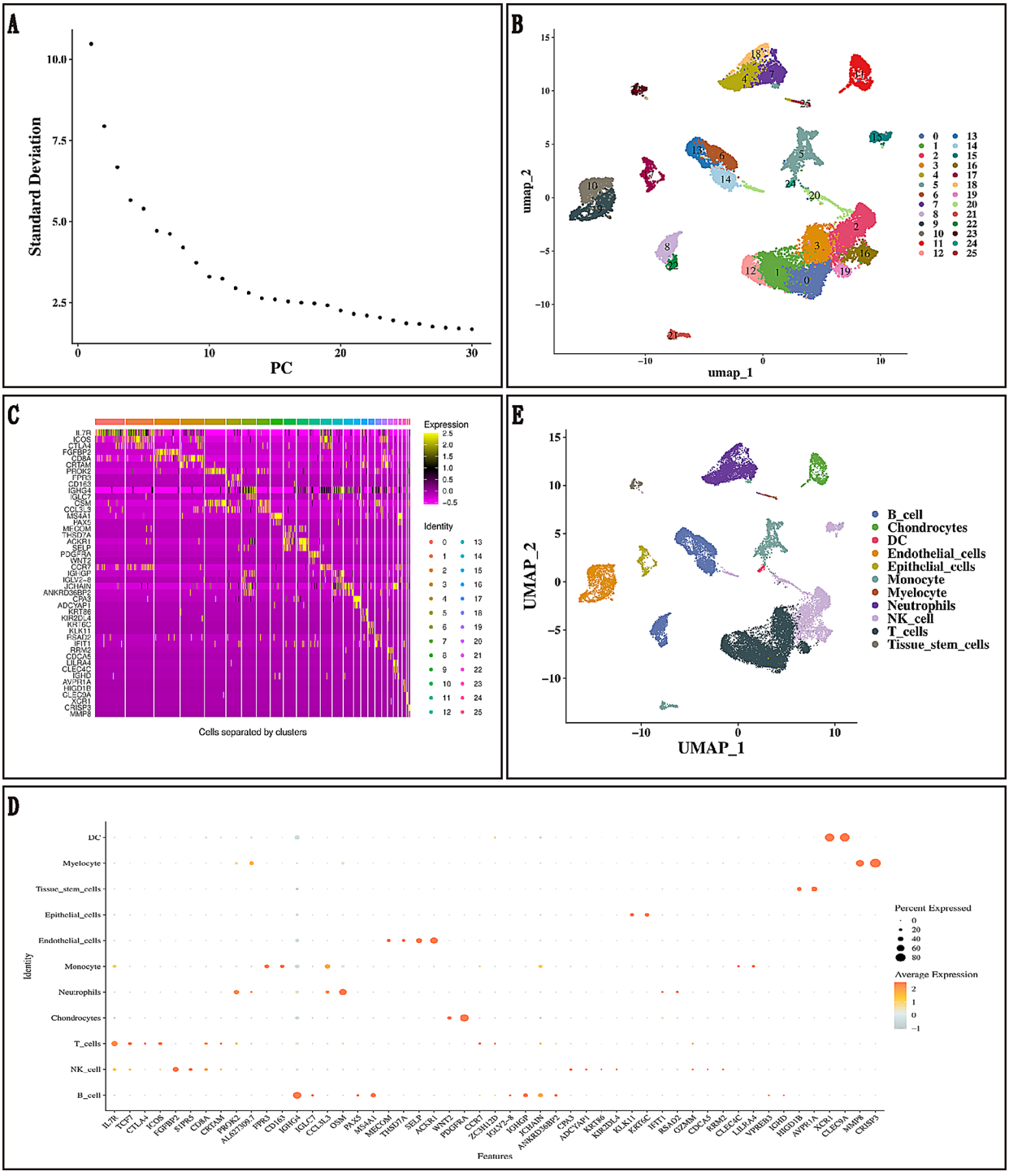

Fig. 7 B cells and T cells holding crucial parts in PD progress were identified **as key** cell types. (**A**) Single cell data dimension reduction cluster principal component lithotripsy map. (**B**) UMAP dimension reduction. (**C**) Heat maps of specific high-expression genes in different cell types. (**D**) Bubble map of maker gene expression in 11 cell types. (**E**) Cell type annotation diagram, different colors distinguish different subgroups of cell types. This figure was obtained by analysis in the GSE171213 single-cell dataset

expressed in T cells, and CXCL6 mainly expressed in epithelial cells and chondrocytes (Fig. 8A). In cell types of PD and controls, the expression of CKAP2 was lower in T cells in the PD group, while the expression of CXCL6 was higher in B cells in the PD group (Fig. 8B). Secondary clustering of B cells and T cells resulted in 13 and 9 molecular subtypes of T cells and B cells, respectively (Fig. 8C-D). Pseudo-time series analysis revealed that T

Table 1 Differences in abundance of each cell type in different groups of samples

Celltype	*P* Value	Statistic
Endothelial_cells	7.14E-260	1185.88147
NK_cell	4.17E-101	455.6903502
Epithelial_cells	7.43E-51	224.9750395
Tissue_stem_cells	3.95E-28	120.934683
Myelocyte	1.10E-25	109.7683071
Monocyte	5.18E-10	38.60708134
Neutrophils	6.27E-10	38.23418814
T_cells	2.67E-08	30.93291488
B_cell	1.02E-06	23.88797076
DC	0.001272592	10.38186276
Chondrocytes	0.164637021	1.931110326

and B cells underwent differentiation through distinct stages, with T cells showing 9 stages and B cells showing 5 stages. Stage 1 were the beginning stage of their differentiation, yet stage 1 and stage 4 were the end stage of T cell and B cell differentiation, respectively (Fig. 8E-F). During the development of T cells, the expression of CKAP2 elevated significantly and then decreased, while the expression of IGLL5 and MZB1 increased suddenly from a stable state (Fig. 8G). During B cell differentiation, the expression of CKAP2 decreased and then levelled off, whereas the expression of IGLL5 and MZB1 increased and then decreased, but was higher than at the beginning. The expression of CXCL6 was unaltered in T cells and in the progression of B cell development (Fig. 8G).

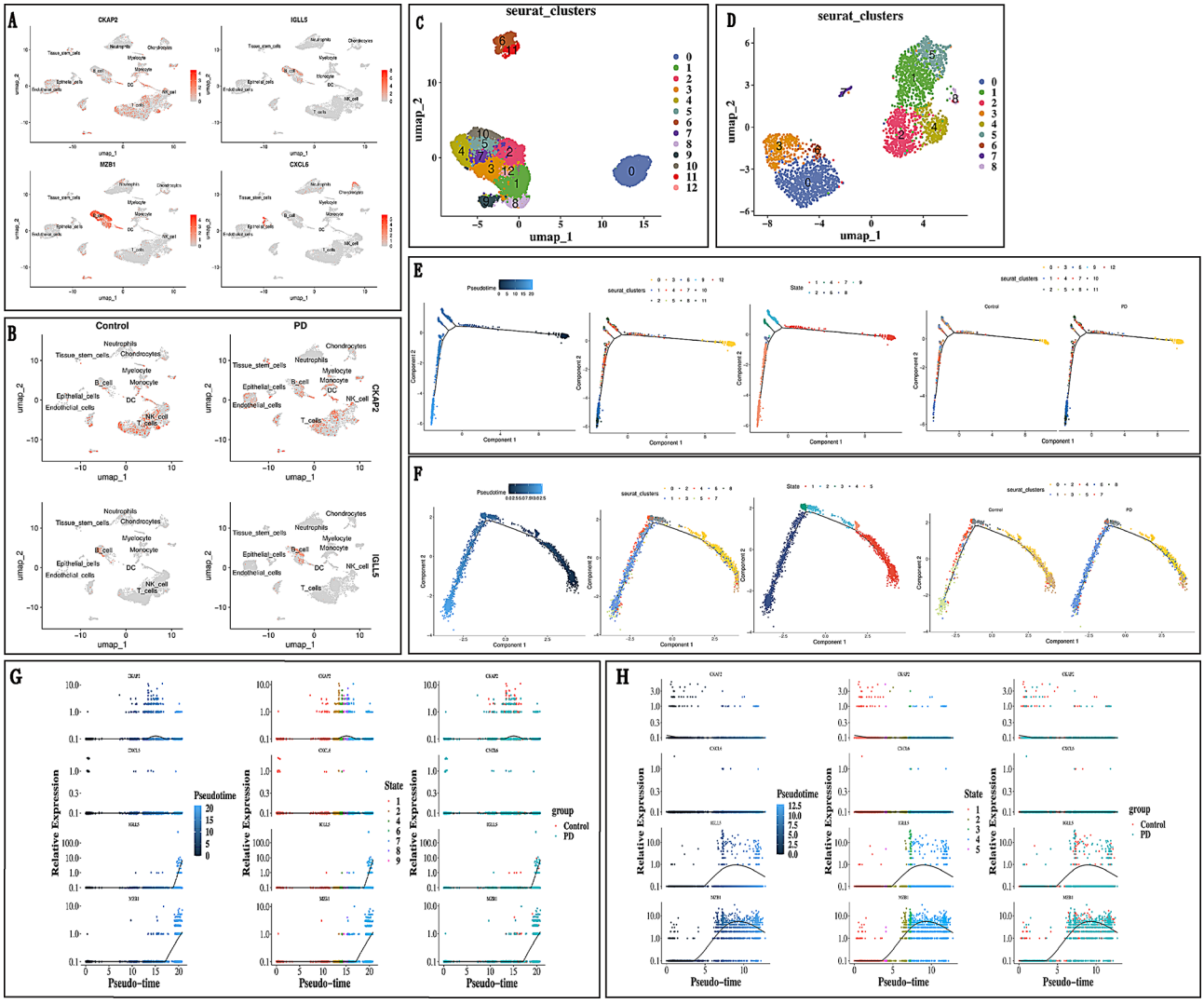

Fig. 8 The mRNA expression level of CKAP2, MZB1, AADACL2, IGLL5, and CXCL6. (**A**) Distribution of core genes in different cell types. (**B**) Distribution of core genes in different groups of samples. (**C**) The results of different subclusters of T cells. (**D**) The clustering results of different subpopulations of B cells. (**E-F**) T cell and B cell pseudo time series analysis. (**G-F**) Expression trends of core genes in T cells and B cell at different developmental stages.This figure was obtained from pseudo-time series analysis in the GSE171213 dataset

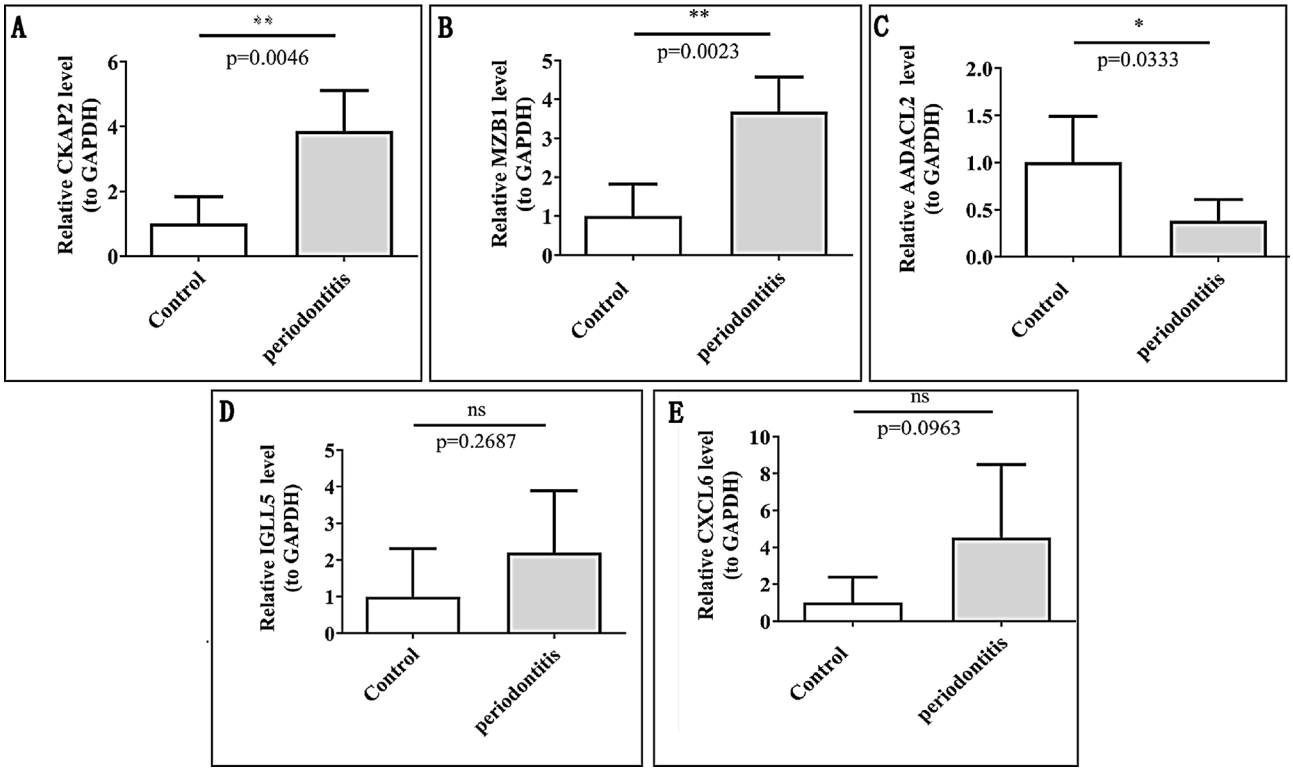

Fig. 9 The mRNA expression level of CKAP2, MZB1, AADACL2, IGLL5, and CXCL6. The expression validation results were obtained by RT-qPCR on 6 PD samples and 6 controls

The mRNA expression level of CKAP2, MZB1, AADACL2, IGLL5, and CXCL6

The mRNA expression levels of CKAP2 and MZB1 were increased in the blood of PD patients compared to controls ($p<0.05$) (Fig. 9A-B). And the mRNA expression levels of AADACL2 was decreased in the PD patients compared to controls ($p<0.05$) (Fig. 9C). The mRNA expression level of IGLL5 and CXCL6 were increased in PD patients, but there were no statistical significance ($p>0.05$) (Fig. 9D-E).

Discussion

PD is characterized by the colonization of diverse microorganisms that form biofilms on tooth surfaces. This condition adversely impacts the gums, alveolar bone, and periodontal ligament, ultimately resulting in tooth loss and systemic inflammation [37]. Notably, the progression of PD typically occurs with minimal external clinical manifestations; consequently, treatment is frequently postponed until advanced stages of the disease [38]. Furthermore, studies have demonstrated that PD is linked to a range of comorbidities, including diabetes mellitus [39], cardiovascular disorders [40], and rheumatoid arthritis [41]. Exosomes play a critical role in immune responses, viral pathogenicity, pregnancy-related processes, cardiovascular conditions, central nervous system disorders, and cancer progression. Through exosomes, proteins, metabolites, and nucleic acids are conveyed to target cells—effectively modulating their biological responses. This exosome-mediated interaction can either promote or inhibit disease progression; furthermore, the response to treatment may be influenced by comprehensive analyses of exosomal components. Therefore, investigating the roles of genes associated with exosomes in the progression of PD and their underlying mechanisms is imperative [42]. This study identified five core genes, CKAP2, IGLL5, MZB1, AADACL2, and CXCL6, through machine learning and expression validation. It further analyzed the molecular mechanisms, immune microenvironment, molecular regulatory networks, and related drugs of these core genes affecting PD. This provides new references for the prevention and treatment of PD.

MZB1 (Marginal Zone B- and B1-cell-specific protein) is a protein-coding gene recognized in the Human Gene Database of GeneCards® as being 'associated with IgM heavy and light chains. It facilitates the assembly and secretion of IgM, thereby contributing to the functional diversification of peripheral B cells through integrin activation [43]. Additionally, it functions as a hormone-regulated adipocytokine/pro-inflammatory cytokine linked to chronic inflammation, influencing cellular expansion. The MZB1 gene is implicated in several pathways: it positively regulates immunoglobulin biosynthesis, promotes apoptotic processes, enhances cellular proliferation, and

modulates B cell proliferation via integrin activation. Furthermore, MZB1 belongs to the family of endoplasmic reticulum (ER) chaperone proteins that regulate both surface expression and secretion of IgM [44–46]. MZB1 also seems to play a crucial role in the regulation of mitochondrial function, cell apoptosis, and calcium homeostasis [47, 48].Through the comprehensive integration and validation of transcriptomic and proteomic data, alongside immunohistochemical analyses, we propose that MZB1 is a protein integral to B cell development and antibody production. Its upregulation in PD indicates potential immune dysregulation in this context, suggesting that MZB1 may serve as a promising biomarker for PD [49].

Some researchers have indicated that CKAP2 may influence tumor development through the inflammation-associated NF-κB signaling pathway [50]. Moreover, the NF-κB system is pivotal in orchestrating inflammatory responses, adaptive immunity, cell differentiation and proliferation, as well as the survival of multicellular organisms [51]. Consequently, it is reasonable to hypothesize that CKAP2 modulates the canonical inflammatory NF-κB signaling pathway to enhance inflammation, which in turn affects the onset and progression of PD. however, further investigation is warranted.

Chemokines represent a distinct class of cytokines that facilitate the migration of specific immune cells to inflamed tissues, thereby playing an essential role in modulating the immune response. Dysregulation within the chemokine signaling pathway can have detrimental consequences, contributing to the pathogenesis of inflammatory diseases and cancer. C-X-C motif chemokine ligand 5 (CXCL5) and C-X-C motif chemokine ligand 6 (CXCL6) are notable members of this family, involved in various inflammatory conditions through their capacity to amplify immune regulatory mechanisms. Emerging evidence suggests that early inflammatory events—such as leukocyte proliferation and acute phase reactant production—can lead to heightened chemokine synthesis. Furthermore, research indicates that chemokines may exert regulatory influences on PD and bone remodeling under both physiological and pathological states [52]. Nevertheless, the precise relationship between CXCL5 and PD remains inadequately elucidated.

With the progression of PD severity, CXCL6-mediated neutrophil recruitment could synergistically enhance the recruitment pathway initiated by IL-8 [53]. Sixue Gao [54] conducted a comprehensive bioinformatics analysis revealing a significant association between CXCL5 and CXCL6 and the pathogenesis of PD. Subsequent immunohistochemical studies further substantiated the critical role of CXCL5 in PD development, indicating that both CXCL5 and CXCL6 may serve as valuable diagnostic biomarkers for this condition. Research suggests their predictive potential; moreover, integrating CXCL5 and CXCL6 with additional biomarkers may improve the accuracy of PD prediction.

This study data analysis has revealed significant reductions in the expression levels of specific genes, among which AADACL2 is a crucial protein associated with desmosomes, a vital cell-cell adhesion system. Esra et al. [55] conducted genomic studies on tissue samples from patients suffering from extensive aggressive PD (GAgP) and healthy controls to investigate molecular biomarkers relevant to PD. This was achieved through gene expression microarray analysis combined with network and pathway analyses to elucidate gene expression patterns. The most prominently upregulated genes included MZB1, TNFRSF17, PNOC, FCRL5, LAX1, BMS1P20, Igll5, MMP7, SPAG4, and MEI1; conversely, the most significantly downregulated genes were LOR, LAMB4, and AADACL2. Notably, the downregulation of AADACL2 aligns with our findings. However, there remains a lack of research exploring the relationship between AADACL2 and PD necessitating further investigation.

Immune cell infiltration represents a critical characteristic that can elucidate the defense mechanisms associated with various diseases. The extent of immune infiltration serves as a valuable reference for investigating the degree of immune cell presence in specific pathological conditions. The analysis of the correlation between core genes and immune cell populations in this study reveals that MZB1 exhibits the most pronounced positive correlation with activated B cells, whereas AADACL2 demonstrates the most significant negative correlation with these same cells. B lymphocytes (B cells) are part of the humoral component of the adaptive immune system and are specifically responsible for secreting antibodies. B cells can also present antigens and secrete cytokines. In mammals, B cells mature in the bone marrow, and the B cell receptors (BCRs) mature on their cell membranes, allowing B cells to bind to specific antigens and initiate an antibody response [56, 57]. Low levels of memory B cells in healthy periodontal tissue in vivo seem to be important for preventing bone loss due to subclinical inflammation. On the other hand, they can also exacerbate alveolar bone loss in a RANKL receptor activator (RANKL)-dependent manner and affect the severity of periodontitis [58]. PD is characterized by the concurrent loss of bone tissue, bacterial invasion, and host defense mechanisms. Research has demonstrated that specific pathogens play a role in forming a beneficial probiotic biofilm that actively contributes to the pathogenesis of PD [1]. This multifactorial condition is influenced not only by microbial factors but also by environmental variables and host characteristics, particularly immune defenses. Notably, leukocyte infiltration—especially B cell infiltration—is correlated with the severity of PD; indeed, distinct subpopulations

of infiltrating B cells are associated with inflammatory responses and bone resorption. During the inflammatory phase, neutrophils release a variety of inflammatory mediators, such as reactive oxygen species (ROS), to eliminate pathogens while potentially inducing damage to surrounding tissues [59]. In this phase, macrophages predominantly exhibit a pro-inflammatory phenotype (M1 type) and secrete pro-inflammatory factors like tumor necrosis factor-α (TNF-α) and interleukin-1β (IL-1β). The expression of genes associated with the inflammatory response is up-regulated.

During the regression phase, macrophages initiated polarization towards the anti-inflammatory phenotype known as M2 type. Consequently, M2 macrophages secreted crucial anti-inflammatory factors such as interleukin-10 (IL-10) and transforming growth factor-β (TGF-β), thereby facilitating the reparative process of periodontal tissue [60, 61].

The above studies suggest that B cells play an important role in the development of PD. And MZB1 and AADACL2 were significantly correlated. This further suggests that MZB1 and AADACL2 may also influence the development of PD to some extent.

A recent study revealed elevated expression levels of MZB1 in nasal polyp tissues (CRSwNP) compared to healthy controls. Single-cell transcriptomic analyses and epitope mapping indicate an aberrantly high expression of the MZB1 gene within a specific B cell population present in polyp tissues. Furthermore, ex vivo stimulation with MZB1 significantly upregulates IgE mRNA expression, suggesting that MZB1 is predominantly expressed in plasma cells and mature B cells located within the nasal mucosa. Therefore, we propose that MZB1 may facilitate IgE production by B cells in middle periodontal tissues during periods of inflammation related to PD [55]. However, there remains a lack of literature indicating any negative correlation between AADACL2 and B cells. Thus further investigation into this relationship is warranted.

Exosomes derived from the host play a significant role in the pathogenesis of PD. Periodontal ligament fibroblasts (PDLFs) represent the primary cell population interacting with pathogenic microorganisms during the initial stages of PD. Upon stimulation by lipopolysaccharide, PDLF-derived exosomes enhance the expression of IL-6 and TNF-α in osteocytes while concurrently reducing collagen type I, bone morphogenetic proteins, and alkaline phosphatase activity, leading to increased bone resorption and destruction within the alveolar bone [62]. Under conditions of circulating tensile stress, PDLF-derived exosomes can inhibit the NF-κB (nuclear factor kappa-B) signaling pathway to downregulate IL-1β expression in lipopolysaccharide-stimulated macrophages, thereby impeding disease progression [63]. Periodontal ligament stem cells (PDLSCs), which are specialized mesenchymal stem cells capable of self-renewal when exposed to an inflammatory microenvironment, also contribute significantly to this process [64]. Exosomes secreted by PDLSCs upon lipopolysaccharide stimulation exhibit elevated levels of miR-155-5p and its downstream target Sirtuin-1; these molecules collectively modulate Th17 cell populations downward while promoting regulatory T cell (Treg) numbers. This regulation occurs through a network involving T17/Treg/miR-155-5p/Sirtuin-1 that effectively suppresses cellular inflammation and delays disease progression [65]. Molecular regulation results revealed that hsa-mir-185-5p, MYC, CNOT3, TP53, and TRIM28 potentially modulate the expression of MZB1 and CKAP2. Furthermore, it has been discovered that hsa-mir-185-5p and hsa-miR-130a-3 are associated with periodontitis-related inflammation [66]. Stimulation of inflammatory factors can alter the expression levels of these miRNAs, subsequently impacting the expression of core genes and promoting inflammation, immune cell recruitment, tissue repair, and ultimately influencing periodontitis development [67]. CXCL6 and CKAP2 suggest that doxorubicin is a potential treatment for oral mucositis by regulating intracellular signal transduction molecules associated with cell proliferation, apoptosis or inflammation.

This study employed a comprehensive array of bioinformatics methodologies to elucidate the diagnostic significance of ERGs in PD. Five promising therapeutic targets (CKAP2, IGLL5, MZB1, AADACL2, CXCL6) were identified for potential intervention in PD treatment. Additionally, an examination of the molecular mechanisms, immune microenvironment, molecular regulatory networks, and associated pharmacological agents influencing the progression of PD was conducted for core genes, thereby offering novel insights into therapeutic strategies for this condition. However, the current study still has limitations. Firstly, in terms of sample selection, it is challenging to determine whether the identified exosome-related genes represent characteristics of the inflammatory activity phase or the regression phase due to a lack of distinction between different inflammatory states. Secondly, there may be heterogeneity among the samples. The current experiments are limited to PCR, making it difficult for us to ascertain if these exosome-related genes can serve as effective intervention targets and thus limiting the generalizability and practical application value of the study results. In light of this, future research should focus on several aspects for improvement. On one hand, a more refined sample grouping method should be employed to independently analyze samples from both the inflammatory phase and regression phase in order to determine specific representation of exosome-related genes during different periods.

On the other hand, further experimental verification is necessary for identified exosome-related genes including functional experiments (such as gene knockout and overexpression experiments), cell model experiments, and animal model experiments in order to establish more accurate gene function and mechanisms of action. Additionally, deeper bioinformatics analysis is needed by integrating state-of-the-art algorithms and databases to reevaluate gene relationships and their association with diseases so as to provide a more valuable foundation for disease diagnosis, treatment, and prevention.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12903-024-05409-w.

Supplementary Material 1

Acknowledgements
We would like to thank the Scientific and Technological Assistance Programs of Xinjiang province for their assistance and guidance in this research.

Author contributions
Wufanbieke·Baheti and Xiaotao Chen searched the literature and conceived the study. Diwen Dong and Congcong Li were involved in protocol development, obtaining ethical approval, patient recruitment, and data analysis. Wufanbieke·Baheti wrote the first draft of the man-uscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Funding
This work was supported by the Science and Technology Department of Xinjiang Uygur Autonomous Region [grant number 2023D01C78] and National Natural Science Foundation of China [grant number 82260195].

Data availability
All data generated or analysed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate
This research project was reviewed and approved by the Ethics Committee of People's Hospital of Xinjiang Autonomous Region. The approval process ensured that the study adhered to ethical principles and guidelines, including the protection of the rights, welfare, and privacy of the participants. All participants were given informed consent.

Clinical trial number
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Stomatology, People's Hospital of Xinjiang Autonomous Region, Urumqi City, China
[2]The Affiliated Hospital of Stomatology, School of Stomatology, Zhejiang University School of Medicine, Hangzhou, China

Received: 8 October 2024 / Accepted: 31 December 2024
Published online: 06 January 2025

References

1. Kassebaum NJ, Bernabe E, Dahiya M, Bhandari B, Murray CJ, Marcenes W. Global burden of severe periodontitis in 1990–2010: a systematic review and meta-regression. J Dent Res. 2014;93(11):1045–53. https://doi.org/10.1177/0022034514552491.
2. Collaborators GS. Measuring the health-related Sustainable Development Goals in 188 countries: a baseline analysis from the global burden of Disease Study 2015. Lancet. 2016;388(10053):1813–50. https://doi.org/10.1016/S0140-6736(16)31467-2.
3. Jin LJ, Lamster IB, Greenspan JS, Pitts NB, Scully C, Warnakulasuriya S. Global burden of oral diseases: emerging concepts, management and interplay with systemic health. Oral Dis. 2016;22(7):609–19. https://doi.org/10.1111/odi.12428.
4. Chapple IL, Van der Weijden F, Doerfer C, Herrera D, Shapira L, Polak D, et al. Primary prevention of periodontitis: managing gingivitis. J Clin Periodontol. 2015;42(Suppl 16):S71–6. https://doi.org/10.1111/jcpe.12366.
5. Marcano R, Rojo M, Cordoba-Diaz D et al. Pathological and Therapeutic Approach to Endotoxin-Secreting Bacteria Involved in Periodontitis. Toxins (Basel). 1970;80.
6. Bui FQ, Almeida-da-Silva CLC, Huynh B, Trinh A, Liu J, Woodward J, et al. Association between periodontal pathogens and systemic disease. Biomed J. 2019;42(1):27–35. https://doi.org/10.1016/j.bj.2018.12.001.
7. Kramer A, Splieth C. Health promotion through structured oral hygiene and good tooth alignment. GMS Hyg Infect Control. 2022;17:Doc08. https://doi.org/10.3205/dgkh000411.
8. Zheng DX, Kang XN, Wang YX, Huang YN, Pang CF, Chen YX, et al. Periodontal disease and emotional disorders: a meta-analysis. J Clin Periodontol. 2021;48(2):180–204. https://doi.org/10.1111/jcpe.13395.
9. Herrera D, Sanz M, Kebschull M, et al. Treatment of stage IV periodontitis: the EFP S3 level clinical practice guideline. J Clin Periodontol. 2022;49(Suppl 24):4–71. https://doi.org/10.1111/jcpe.13639.
10. Tomokiyo A, Wada N, Maeda H. Periodontal Ligament Stem cells: regenerative potency in Periodontium. Stem Cells Dev. 2019;28(15):974–85. https://doi.org/10.1089/scd.2019.0031.
11. Mianehsaz E, Mirzaei HR, Mahjoubin-Tehran M, Rezaee A, Sahebnasagh R, Pourhanifeh MH, et al. Mesenchymal stem cell-derived exosomes: a new therapeutic approach to osteoarthritis? Stem Cell Res Ther. 2019;10(1):340. https://doi.org/10.1186/s13287-019-1445-0.
12. Garbern JC, Lee RT. Cardiac stem cell therapy and the promise of heart regeneration. Cell Stem Cell. 2013;12(6):689–98. https://doi.org/10.1016/j.stem.2013.05.008.
13. Yan XZ, Yang F, Jansen JA, de Vries RB, van den Beucken JJ. Cell-based approaches in Periodontal Regeneration: a systematic review and Meta-analysis of Periodontal defect models in animal experimental work. Tissue Eng Part B Rev. 2015;21(5):411–26. https://doi.org/10.1089/ten.TEB.2015.0049.
14. Liang X, Ding Y, Zhang Y, Tse HF, Lian Q. Paracrine mechanisms of mesenchymal stem cell-based therapy: current status and perspectives. Cell Transpl. 2014;23(9):1045–59. https://doi.org/10.3727/096368913X667709.
15. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 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.
16. Liu S, Xie X, Lei H, Zou B, Xie L. Identification of key circRNAs/lncRNAs / miRNAs/mRNAs and pathways in Preeclampsia using Bioinformatics Analysis. Med Sci Monit. 2019;25:1679–93. https://doi.org/10.12659/MSM.912801.
17. Li J, Wang H, Cao F, Cheng Y. A bioinformatics analysis for diagnostic roles of the E2F family in esophageal cancer. J Gastrointest Oncol. 2022;13(5):2115–31. https://doi.org/10.21037/jgo-22-855.
18. Zheng Y, Gao W, Zhang Q, Cheng X, Liu Y, Qi Z, Li T. Ferroptosis and autophagy-related genes in the pathogenesis of ischemic cardiomyopathy. Front Cardiovasc Med. 2022;9:906753. https://doi.org/10.3389/fcvm.2022.906753.
19. Yu G, Wang LG, Han Y, He QY. 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.
20. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized Linear models via Coordinate Descent. J Stat Softw. 2010;33(1). https://doi.org/10.18637/jss.v033.i01.
21. Shubin NJ, Navalkar K, Sampson D, Yager TD, Cermelli S, Seldon T, et al. Serum protein changes in Pediatric Sepsis patients identified with an aptamer-based multiplexed Proteomic Approach. Crit Care Med. 2020;48(1):e48–57. https://doi.org/10.1097/CCM.0000000000004083.
22. Lopez-Diaz JOM, Mendez-Gonzalez J, Lopez-Serrano PM, Sanchez-Perez FJ, Mendez-Encina FM, Mendieta-Oviedo R, et al. Dummy regression to predict

dry fiber in Agave lechuguilla Torr. In two large-scale bioclimatic regions in Mexico. PLoS ONE. 2022;17(9):e0274641. https://doi.org/10.1371/journal.pone.0274641.
23. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. pROC: an open-source package for R and S + to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77. https://doi.org/10.1186/1471-2105-12-77.
24. Li S, Que Y, Yang R, He P, Xu S, Hu Y. Construction of Osteosarcoma diagnosis model by Random Forest and Artificial neural network. J Pers Med. 2023;13(3). https://doi.org/10.3390/jpm13030447.
25. Goncharova IA, Nazarenko MS, Babushkina NP, Markov AV, Pecherina TB, Kashtalap VV, et al. [Genetic predisposition to early myocardial infarction]. Mol Biol (Mosk). 2020;54(2):224–32. https://doi.org/10.31857/S0026898420020044.
26. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. Cell Rep. 2017;18(1):248–62. https://doi.org/10.1016/j.celrep.2016.12.019.
27. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, 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.
28. Hanzelmann S, Castelo R, Guinney J. BMC Bioinformatics. 2013;14:7. https://doi.org/10.1186/1471-2105-14-7. GSVA: gene set variation analysis for microarray and RNA-seq data.
29. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014;8(Suppl 4):S11. https://doi.org/10.1186/1752-0509-8-S4-S11.
30. Shi Y, Wang Y, Dong H, Niu K, Zhang W, Feng K, et al. Crosstalk of ferroptosis regulators and tumor immunity in pancreatic adenocarcinoma: novel perspective to mRNA vaccines and personalized immunotherapy. Apoptosis. 2023;28(9–10):1423–35. https://doi.org/10.1007/s10495-023-01868-8.
31. Kim SG, Menapace DC, Mims MM, Shockley WW, Clark JM. Age-related histologic and biochemical changes in Auricular and Nasal Cartilages. Laryngoscope. 2024;134(3):1220–6. https://doi.org/10.1002/lary.30990.
32. Yu L, Shen N, Shi Y, Shi X, Fu X, Li S, et al. Characterization of cancer-related fibroblasts (CAF) in hepatocellular carcinoma and construction of CAF-based risk signature based on single-cell RNA-seq and bulk RNA-seq data. Front Immunol. 2022;13:1009789. https://doi.org/10.3389/fimmu.2022.1009789.
33. Ni L, Yang H, Wu X, Zhou K, Wang S. The expression and prognostic value of disulfidptosis progress in lung adenocarcinoma. Aging. 2023;15(15):7741–59. https://doi.org/10.18632/aging.204938.
34. Han Y, Yu C, Yu Y. Galla Turcica alleviates gingiva inflammation and alveolar bone resorption via regulating Th1/Th17 in a mouse model of periodontitis. Adv Traditional Med. 2024;24(3):881–90. https://doi.org/10.1007/s13596-024-00745-2.
35. Alarcon-Sanchez MA, Guerrero-Velazquez C, Becerra-Ruiz JS, Rodriguez- Montano R, Avetisyan A, Heboyan A. IL-23/IL-17 axis levels in gingival crevicular fluid of subjects with periodontal disease: a systematic review. BMC Oral Health. 2024;24(1):302. https://doi.org/10.1186/s12903-024-04077-0.
36. Danielsen AK, Damgaard C, Massarenti L, Ostrup P, Riis Hansen P, Holmstrup P, Nielsen CH. B-cell cytokine responses to Porphyromonas gingivalis in patients with periodontitis and healthy controls. J Periodontol. 2023;94(8):997–1007. https://doi.org/10.1002/JPER.22-0438.
37. Semedo-Lemsaddek T, Tavares M, Sao Braz B, Tavares L, Oliveira M. Enterococcal infective endocarditis following Periodontal Disease in Dogs. PLoS ONE. 2016;11(1):e0146860. https://doi.org/10.1371/journal.pone.0146860.
38. Niemiec BA. Periodontal disease. Top Companion Anim Med. 2008;23(2):72–80. https://doi.org/10.1053/j.tcam.2008.02.003.
39. Bascones-Martínez A, González-Febles J, Sanz-Esporrín J. Diabetes and periodontal disease. Review of the literature. Am J Dent. 2014;27(2):63–7.
40. Merchant AT, Virani SS. Evaluating Periodontal Treatment to Prevent Cardiovascular Disease: challenges and possible solutions. Curr Atheroscler Rep. 2017;19(1):4. https://doi.org/10.1007/s11883-017-0640-7.
41. Marotte H. Non-surgical periodontal disease: a new treatment for rheumatoid arthritis? Joint Bone Spine. 2020;87(1):1–3. https://doi.org/10.1016/j.jbspin.2019.05.002.
42. Zhang J, Li S, Li L, Li M, Guo C, Yao J, Mi S. Exosome and exosomal microRNA: trafficking, sorting, and function. Genomics Proteom Bioinf. 2015;13(1):17–24. https://doi.org/10.1016/j.gpb.2015.02.001.
43. Li D, Zhu Y, Zhang L, Shi L, Deng L, Ding Z, et al. MZB1 targeted by mir-185-5p inhibits the migration of human periodontal ligament cells through NF-kappaB signaling and promotes alveolar bone loss. J Periodontal Res. 2022;57(4):811–23. https://doi.org/10.1111/jre.13014.
44. Kanda M, Tanaka C, Kobayashi D, Tanaka H, Shimizu D, Shibata M, et al. Epigenetic suppression of the immunoregulator MZB1 is associated with the malignant phenotype of gastric cancer. Int J Cancer. 2016;139(10):2290–8. https://doi.org/10.1002/ijc.30286.
45. Flach H, Rosenbaum M, Duchniewicz M, Kim S, Zhang SL, Cahalan MD, et al. Mzb1 protein regulates calcium homeostasis, antibody secretion, and integrin activation in innate-like B cells. Immunity. 2010;33(5):723–35. https://doi.org/10.1016/j.immuni.2010.11.013.
46. Rosenbaum M, Andreani V, Kapoor T, Herp S, Flach H, Duchniewicz M, Grosschedl R. MZB1 is a GRP94 cochaperone that enables proper immunoglobulin heavy chain biosynthesis upon ER stress. Genes Dev. 2014;28(11):1165–78. https://doi.org/10.1101/gad.240762.114.
47. Zhang L, Wang YN, Ju JM, Shabanova A, Li Y, Fang RN, et al. Mzb1 protects against myocardial infarction injury in mice via modulating mitochondrial function and alleviating inflammation. Acta Pharmacol Sin. 2021;42(5):691–700. https://doi.org/10.1038/s41401-020-0489-0.
48. Geary B, Sun B, Tilvawala RR, et al. Peptidylarginine deiminase 2 citrullinates MZB1 and promotes the secretion of IgM and IgA. Front Immunol. 2024;14:1290585. https://doi.org/10.3389/fimmu.2023.1290585.
49. Sunnetci-Akkoyunlu D, Guzeldemir-Akcakanat E, Alkan B, Gurel B, Balta-Uysal VM, Akgun E, et al. Altered expression of MZB1 in periodontitis: a possible link to disease pathogenesis. J Periodontol. 2023;94(11):1285–94. https://doi.org/10.1002/JPER.23-0224.
50. Mitchell S, Vargas J, Hoffmann A. Signaling via the NFkappaB system. Wiley Interdiscip Rev Syst Biol Med. 2016;8(3):227–41. https://doi.org/10.1002/wsbm.1331.
51. Gachpazan M, Habbibirad S, Kashani H, Jamialahmadi T, Rahimi HR, Sahebkar A. Targeting Nuclear factor-Kappa B Signaling Pathway by Curcumin: implications for the treatment of multiple sclerosis. Adv Exp Med Biol. 2024;1291:41–53. https://doi.org/10.1007/978-3-030-56153-6_3. PMID:34331683.
52. Plemmenos G, Evangeliou E, Polizogopoulos N, Chalazias A, Deligianni M, Piperi C. Central Regulatory Role of Cytokines in Periodontitis and Targeting options. Curr Med Chem. 2021;28(15):3032–58. https://doi.org/10.2174/0929867327666200824112732.
53. Kebschull M, Demmer R, Behle JH, Pollreisz A, Heidemann J, Belusko PB, et al. Granulocyte chemotactic protein 2 (gcp-2/cxcl6) complements interleukin-8 in periodontal disease. J Periodontal Res. 2009;44(4):465–71. https://doi.org/10.1111/j.1600-0765.2008.01134.x.
54. Chen Y, Wang H, Yang Q, Zhao W, Chen Y, Ni Q, Li W, Shi J, Zhang W, Li L, Xu Y, Zhang H, Miao D, Xing L, Sun W. Single-cell RNA landscape of the osteoimmunology microenvironment in periodontitis. Theranostics. 2024;12(3):1074–96. https://doi.org/10.7150/thno.65694.
55. Guzeldemir-Akcakanat E, Sunnetci-Akkoyunlu D, Orucguney B, Cine N, Kan B, Yilmaz EB, et al. Gene-expression profiles in generalized aggressive periodontitis: a Gene Network-based microarray analysis. J Periodontol. 2016;87(1):58–65. https://doi.org/10.1902/jop.2015.150175.
56. Cerutti A, Puga I, Cols M. New helping friends for B cells. Eur J Immunol. 2012;42(8):1956–68. https://doi.org/10.1002/eji.201242594.
57. Bonilla FA, Oettgen HC. Adaptive immunity. J Allergy Clin Immunol. 2010;125(2 Suppl 2):S33–40. https://doi.org/10.1016/j.jaci.2009.09.017.
58. Figueredo CM, Lira-Junior R, Love RM. T and B cells in Periodontal Disease: New functions in a Complex scenario. Int J Mol Sci. 2019;20(16). https://doi.org/10.3390/ijms20163949.
59. Liu J, Han X, Zhang T, Tian K, Li Z, Luo F. Reactive oxygen species (ROS) scavenging biomaterials for anti-inflammatory diseases: from mechanism to therapy. J Hematol Oncol. 2023;16(1):116. https://doi.org/10.1186/s13045-023-01512-7.
60. Liu F, Qiu H, Xue M, Zhang S, Zhang X, Xu J, et al. MSC-secreted TGF-beta regulates lipopolysaccharide-stimulated macrophage M2-like polarization via the Akt/FoxO1 pathway. Stem Cell Res Ther. 2019;10(1):345. https://doi.org/10.1186/s13287-019-1447-y.
61. Tao W, Zhang G, Liu C, Jin L, Li X, Yang S. Low-dose LPS alleviates early brain injury after SAH by modulating microglial M1/M2 polarization via USP19/FOXO1/IL-10/IL-10R1 signaling. Redox Biol. 2023;66:102863. https://doi.org/10.1016/j.redox.2023.102863.
62. Zhao M, Dai W, Wang H, Xue C, Feng J, He Y, et al. Periodontal ligament fibroblasts regulate osteoblasts by exosome secretion induced by inflammatory stimuli. Arch Oral Biol. 2019;105:27–34. https://doi.org/10.1016/j.archoralbio.2019.06.002.

63. Wang Z, Maruyama K, Sakisaka Y, Suzuki S, Tada H, Suto M, et al. Cyclic Stretch Force induces Periodontal Ligament cells to Secrete exosomes that suppress IL-1beta production through the inhibition of the NF-kappaB signaling pathway in macrophages. Front Immunol. 2019;10:1310. https://doi.org/10.3389/fimmu.2019.01310.
64. Fawzy El-Sayed KM, Elahmady M, Adawi Z, Aboushadi N, Elnaggar A, Eid M, et al. The periodontal stem/progenitor cell inflammatory-regenerative cross talk: a new perspective. J Periodontal Res. 2019;54(2):81–94. https://doi.org/10.1111/jre.12616.
65. Zheng Y, Dong C, Yang J, Jin Y, Zheng W, Zhou Q, et al. Exosomal microRNA-155-5p from PDLSCs regulated Th17/Treg balance by targeting sirtuin-1 in chronic periodontitis. J Cell Physiol. 2019;234(11):20662–74. https://doi.org/10.1002/jcp.28671.
66. Zheng L, Chopra A, Weiner J 3rd, Beule D, Dommisch H, Schaefer AS. miRNAs from Inflamed Gingiva Link Gene Signaling to increased MET expression. J Dent Res. 2023;102(13):1488–97. https://doi.org/10.1177/00220345231197984.
67. Aradya A, Kiran PK, Raghavendra Swamy KN, Doddawad VG, Ranganatha N, Sravani K. Oral risk factors in patients with Cancer Undergoing Chemotherapy - A Pilot Study. Indian J Dent Res. 2024;35(2):126–30. https://doi.org/10.4103/ijdr.ijdr_516_23.

Publisher's note