

Identification and functional analysis of genes mediating osteoclast-driven progression of osteoporosis

Science Progress

2024, Vol. 107(4) 1–20

© The Author(s) 2024

Article reuse guidelines:

sagepub.com/journals-permissionsDOI: [10.1177/00368504241300723](https://doi.org/10.1177/00368504241300723)journals.sagepub.com/home/sci

Qu Xu¹, Gangning Feng², Zhihai Zhang¹ ,
Jiangbo Yan¹, Zhiqun Tang¹, Rui Wang¹,
Penggang Ma¹, Ye Ma¹, Guang Zhu¹
and Qunhua Jin^{1,2}

¹The Third Ward of Orthopaedic Department, General Hospital of Ningxia Medical University, Yinchuan, China

²Institute of Osteoarthropathy, Ningxia Key Laboratory of Clinical and Pathogenic Microbiology, Institute of Medical Sciences, General Hospital of Ningxia Medical University, Yinchuan, China

Abstract

Objective The pathological mechanism of osteoporosis (OP) involves increased bone resorption mediated by osteoclasts and decreased bone formation mediated by osteoblasts, leading to an imbalance in bone homeostasis. Identifying key molecules in osteoclast-mediated OP progression is crucial for the prevention and treatment of OP.

Methods Differential expression analysis and weighted gene co-expression network analysis (WGCNA) were performed on the OP patient datasets from the GEO database. The results were intersected with the differential expression results from the osteoclast differentiation dataset to identify key genes. These key genes were then subjected to disease relevance analysis, and consensus clustering was performed on OP patient samples based on their expression profiles. The subgroups were analyzed for differences, followed by GO, KEGG, GSEA, and GSVA analyses, and immune infiltration. Finally, osteoclast differentiation model was constructed. After validating the success of the model using TRAP and F-actin staining, the differential expression of key genes was validated in vitro via Western blot.

Results CTRL, ARHGEF5, PPAP2C, VSIG2, and PBLD were identified as key genes. These genes exhibited strong disease relevance ($AUC > 0.9$). Functional enrichment results also indicated their close association with OP and osteoclast differentiation. In vitro differential expression validation showed that during osteoclast differentiation, CTRL was downregulated, while ARHGEF5,

Corresponding author:

Qunhua Jin, First Clinical Medical College, General Hospital of Ningxia Medical University, Yinchuan, China.
Email: 744070408@qq.com



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>)

which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

PPAP2C, VSIG2, and PBLD were upregulated, with all differences being statistically significant ($P < 0.05$).

Discussion Currently, there are no studies on the effects of these five genes on osteoclast differentiation. Therefore, it is meaningful to design *in vivo* and *in vitro* perturbation experiments to observe the impact of each gene on osteoclast differentiation and OP progression.

Conclusion CTRL, ARHGEF5, PPAP2C, VSIG2, and PBLD show high potential as molecular targets for basic and clinical research in osteoclast-mediated OP.

Keywords

Osteoporosis, osteoclast differentiation, WGCNA, consensus clustering, bioinformatics analysis

Introduction

Osteoporosis (OP) is a skeletal disorder characterized by decreased bone strength, significantly increasing the risk of fragility fractures.¹ Statistics show that individuals with OP have a lifetime fracture risk of 40%, with most fractures occurring in the spine, hip, or wrist.²⁻⁴ This condition severely impacts the quality of life for elderly patients and places a substantial burden on healthcare systems worldwide.⁵ Therefore, understanding the pathogenesis of OP and developing effective prevention and treatment strategies is crucial.

Current research indicates that the causes of OP include factors such as medication (e.g. glucocorticoids),⁶ aging,⁷ and estrogen deficiency.⁸ Pathologically, OP is primarily characterized by an imbalance in bone homeostasis, with increased bone resorption mediated by osteoclasts and decreased bone formation mediated by osteoblasts.⁹⁻¹³ Currently, treatments targeting osteoclast-mediated bone resorption in OP include drugs such as bisphosphonates (alendronate, risedronate, zoledronic acid, and ibandronate) or denosumab.¹⁴ However, treatment with bisphosphonates may lead to some serious side effects, such as osteonecrosis of the jaw, delayed dental eruption, atypical femoral fracture, and alterations to the bone growth system.¹⁵ The use of denosumab may result in multiple vertebral fractures related to rebound effects after discontinuation.¹⁶ Therefore, Identifying new molecular targets involved in osteoclast-mediated OP development is essential for providing reliable targets for future research and treatment.

In this study, we identified five key genes associated with osteoclast differentiation-mediated OP and explored their functions. First, using the weighted gene co-expression network analysis (WGCNA) algorithm, we screened a gene cluster most strongly correlated with OP from the dataset. Through differential analysis of OP and osteoclast differentiation datasets, we identified five key genes by intersecting the results. We then validated the association between these key genes and OP using nomograms and ROC curves (receiver operating characteristic curves). Through consensus clustering, we divided 12 patient samples into two subgroups, followed by differential analysis and functional enrichment analysis. Considering the important role of immune cell infiltration in the progression of OP, we examined the immune infiltration status of each sample and found that Cluster 2 had a higher proportion of activated mast cells, while plasma cells, $\gamma\delta$ T cells, activated NK cells, M1 and M2 macrophages, and resting mast cells were lower in proportion. Finally, we constructed a RANKL-induced osteoclast differentiation

model and validated the differential expression of the key genes through TRAP staining, F-actin detection, and Western blot analysis. The results indicate that these five key genes are reliable molecular targets for the diagnosis and treatment of OP.

Material and methods

Microarray data collection

Datasets GSE230665 and GSE225974 were obtained from the GEO database (GEO: <http://www.ncbi.nlm.nih.gov/geo>). The GSE230665 dataset includes 12 postmenopausal OP patients diagnosed according to the “Recommended Diagnostic Criteria for OP Patients in China (Second Draft)¹⁷” and 3 healthy postmenopausal women as the control group. The ages of the 15 participants range from 52 to 68 years, with an average of 62.35 ± 4.64 years. GSE225974 dataset consists of 8 PBMC and 8 osteoclast-like cell samples. Normalize the dataset using the “limma” package in R software (version 4.3.1).

Weighted gene co-expression network analysis

WGCNA is a bioinformatics technique employed to elucidate the correlation patterns among genes across different samples, facilitating the identification of gene modules with high co-expression. We used the “WGCNA” R package to build the gene co-expression network for OP. Initially, a sample clustering tree was plotted to identify outliers, which were subsequently removed. The adjacency matrix (AM) was then transformed into a topological overlap matrix (TOM). Following this, genes with similar functions were grouped into modules using the “DynamicTreeCut” method. Lastly, we determined the correlation between various modules and OP pathogenesis, choosing the module with the highest correlation score as the key gene set generated by WGCNA.

Identification of differentially expressed genes

Imported the GSE230665 and GSE225974 datasets into R software (version 4.3.1) and preprocessed them using the “limma” package. Differentially expressed genes (DEGs) were subsequently identified with the limma package, applying thresholds of $|l\log_2(\text{fold-change})| > 1$ and $P < 0.05$. To visualize all DEGs, volcano plots and DEG expression heatmaps were created using the “ggplot2” and “pheatmap” packages, respectively.

Validation of diagnostic efficiency of key genes

Initially, nomogram model to predict OP risk using the “rms” package. We utilized ROC curves to assess the model’s performance and the diagnostic efficiency of potential biomarkers. The area under the ROC curve (AUC) was employed to evaluate accuracy, with thresholds defined as follows: excellent accuracy ($0.9 \leq \text{AUC} < 1$), good accuracy ($0.8 \leq \text{AUC} < 0.9$), and non-informative accuracy ($\text{AUC} = 0.5$).

Consensus clustering of samples

Consensus clustering of all OP patient samples was carried out using the R package “ConsensusClusterPlus,” resulting in the division of the samples into two distinct groups. The clustering outcomes were visualized through Delta area plots and cumulative distribution function (CDF) plots. Heatmaps were created to depict gene expression profiles between these groups. Significant differences between the groups were evaluated using the Wilcoxon rank-sum test, with $P < 0.05$ deemed significant.

Functional enrichment analysis and visualization

To investigate the biological functional heterogeneity between the two groups, we conducted a differential analysis with thresholds of $|log_2(\text{fold-change})| > 1$ and $P < 0.05$, followed by enrichment analysis of the DEGs. Gene ontology (GO) enrichment analysis was employed for a broad functional analysis of gene sets. The KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis was used for a systematic analysis of gene functions, linking genomic data with functional information and high-throughput data. Gene set enrichment analysis (GSEA) was carried out using predefined gene sets from “c2.all.v7.0.symbols.gmt,” utilizing the GSEA method from the R package “clusterProfiler.” Gene set variation analysis (GSVA) was performed to convert the gene expression matrix across samples into gene set expression profiles, assessing pathway enrichment across samples. GSVA used gene sets “c2.all.v7.0.symbols.gmt” and “c5.all.v7.0.symbols.gmt” obtained from MSigDB (<https://www.gsea-msigdb.org/gsea/msigdb/>) to determine functional enrichment differences between the two groups.

Immune infiltration analysis

We utilized CIBERSORT to examine the expression matrix of immune cell subtypes in each sample, employing a deconvolution algorithm based on linear support vector regression. We employed the “CIBERSORT” R package to estimate the quantities of 22 immune cell types in OP patients, with visualization of results facilitated through rainbow plots. The Wilcoxon test was employed to assess differences in immune cell proportions between subgroups, and results were quantified using ssGSEA, with significance determined at $P < 0.05$. Spearman correlation was utilized to determine the correlation between immune cell expressions across all OP patient samples, and results were depicted using the R package “ggcorrplot.” Additionally, Spearman correlation was used to calculate the correlation between key gene expressions and immune cell expressions in OP patient samples, with visualization accomplished using ggcorrplot.

Cell culture

RAW264.7 cells were sourced from the American Type Culture Collection (Manassas, VA, USA) and grown in DMEM supplemented with 10% FBS, 100 µg/mL penicillin, and 100 units/mL streptomycin at 37 °C in a humidified 5% CO₂ environment.

Western blot analysis

RAW264.7 cells were seeded into 6-well plates at a density of 5×10^4 cells/well and incubated for 24 hours. Subsequently, the cells were either exposed to 50 ng/ml of RANKL or left untreated, and cultured for 7 days with media changed every 48 hours. Protein extraction was performed using a Total Protein Extraction Kit (Cat. No: KGB5303-100; Nanjing KeyGen Biotech Co., Ltd) according to the manufacturer's instructions. Protein concentration was determined using the BCA Protein Quantification Kit (Cat. No: KGB2101-500; Nanjing KeyGen Biotech Co., Ltd). For each sample, 20 µg of protein was separated by electrophoresis on a 10% sodium dodecyl sulfate–polyacrylamide gel (SDS–PAGE) and transferred onto polyvinylidene fluoride (PVDF) membranes. The membranes were blocked with 5% nonfat dry milk at room temperature for 60 minutes, and then incubated overnight at 4 °C with the following primary antibodies: CTRL (1:1000; Cat. No: 14856-1-AP), ARHGEF5 (1:1000; Cat. No: 11379-1-AP), PBLD (1:1000; Cat. No: 27891-1-AP), PPAP2C (1:1000; Cat. No: XY10451-1), VSIG2 (1:1000; Cat. No: BF8452), and β-actin (1:5000; Cat. No: 66009-1-IG). The next day, the membranes were washed five times with Tris-buffered saline with Tween-20 (TBST), each wash lasting 5 minutes. Then, the membranes were incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies (1:10,000; Cat. No: RGAM001/RGAR001) at room temperature for 40 minutes. Finally, the blots were visualized using an enhanced chemiluminescence (ECL) kit (ABclonal; Cat. No: RM00021), and the image data were analyzed using Image J software (Version 1.54q).¹⁸

In vitro osteoclastogenesis assay

RAW264.7 cells were plated onto 12-well plates with each well receiving 2×10^4 cells. Following a 24-hour incubation period, the culture medium was substituted with a fresh medium containing 50 ng/ml RANKL. The medium was refreshed every 48 hours until reaching day 7. At day 3, day 5, and day 7 time points, TRAP staining was executed according to the guidelines outlined by the manufacturer (Sigma; catalog no. 387A-1KT). Osteoclast maturation was assessed by identifying TRAP-positive cells containing three or more nuclei, and their relative abundance was determined either by area or by counting per field under a DMi8 optical microscope (Leica). Image processing was conducted using ImageJ software (Version 1.54q).¹⁸

Immunofluorescence assay for F-actin rings

RAW264.7 cells were introduced into 12-well plates with a seeding density set at 2×10^4 cells per well. Following an initial 24-hour incubation period, the culture medium underwent replacement with a fresh medium containing 50 ng/ml RANKL. This medium renewal process was repeated every 48 hours until reaching the 7th day. At the 3rd, 5th, and 7th day marks, the detection of F-actin rings within osteoclasts was executed as per the manufacturer's guidelines (Solarbio; catalog no. CA1610). Utilizing a Leica DMi8 confocal microscope, F-actin rings were visualized, and subsequent quantification of their quantity was performed utilizing ImageJ software (Version 1.54q).¹⁸

Statistical analysis

The data are presented as the mean \pm standard error of the mean (SEM). Statistical analysis was performed using GraphPad Prism 10.1.2 software. Comparisons between two groups with normally distributed data were conducted using Student's *t*-tests; otherwise, the nonparametric Mann–Whitney test was used. For comparisons of multiple groups of normally distributed data, one-way ANOVA followed by Tukey's test was applied; for data with a non-normal distribution, the Kruskal–Wallis test was used. Differences with a $P < 0.05$ were considered statistically significant.

Results

Data normalization

To offer a clearer portrayal of the research framework, we've condensed the structural outline of the article into a flowchart, showcased as in Figure 1. The OP dataset GSE230665 samples were divided into OP group (12 cases) and Control group (3 cases), while the osteoclast differentiation dataset GSE225974 samples were divided into OC group (8 cases) and PBMC group (8 cases). The data from both sets were standardized using the normalizeBetweenArrays function in the limma package. Figure 2A, C depicts the data structure of the OP dataset before standardization, presented as box plots and density plots, while Figure 2B, D shows the results after standardization. Similarly, Figure 2E, G illustrates the data structure of the osteoclast differentiation dataset before standardization, and Figure 2F, H displays the results after standardization. The results indicate that the standardization of both datasets produced satisfactory corrections.

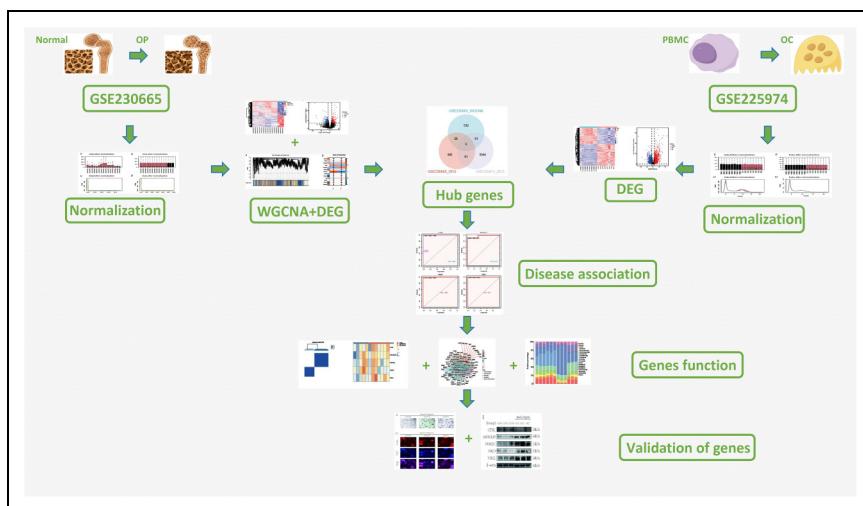


Figure 1. Flowchart.

Identification of key genes in osteoclast-mediated osteoporosis progression

We began by performing WGCNA on the OP dataset to identify a set of genes most highly associated with OP. Initial clustering of all samples did not reveal any obvious outliers, so no samples were removed, and all samples were re-clustered accordingly (Figure 3A). The soft power threshold was determined to be 9 (scale-free topology fit index of 0.75) (Figure 3B and C). Using the WGCNA R package, we clustered the genes into modules, resulting in the construction of 13 gene co-expression modules (Figure 3D). These modules showed significant independence from each other. Figure 3E highlights the module most associated with OP, with the lightsteelblue module showing the most significant correlation (Figure 3E). This module contains 638 genes highly related to OP (Figure 3F). Next, we performed differential analysis on the two datasets using the limma package. For dataset GSE230665, with thresholds set at $|log_2(fold-change)| > 1$ and $P < 0.05$, we identified 596 DEGs, including 289 downregulated and 307 upregulated genes. These DEGs are visualized in a heatmap and volcano plot (Figure 3G and H). For dataset GSE225974, with the same thresholds, we identified 3521 DEGs, including 1532 downregulated and 1989 upregulated genes, also visualized in heatmap and volcano plot formats (Figure 3I and J). Finally, by intersecting the module genes identified by WGCNA with the DEGs from both datasets, we identified five key genes involved in osteoclast-mediated OP progression: CTRL, ARHGEF5, PPAP2C, VSIG2, and PBLD (Figure 3K).

Association of key genes with OP diagnosis

To begin, we illustrated the association between the expression levels of the five key genes and the diagnosis of OP using a nomogram plot (Figure 4A). The results revealed that at total scores of approximately 95 and 115, the probability of diagnosing OP was 1% and 99%, respectively. ROC validation demonstrated the high efficiency of the diagnostic model ($AUC = 1.000$) (Figure 4B). Subsequently, we plotted ROC curves for each key gene to individually validate their association with OP diagnosis. The results showed that, except for the gene ARHGEF5 with an AUC value of 0.972, the AUC values for the remaining four genes were all 1.000 (Figure 4C–G), indicating a robust association between these five key genes and OP diagnosis.

Functional analysis of key genes

To explore the functionality of the five key genes, we first utilized the ConsensusClusterPlus algorithm to cluster OP samples based on the expression levels of these key genes. Subsequently, we performed differential expression analysis on the resulting subgroups and further conducted functional enrichment analysis on the DEGs. The obtained results depict the biological functional differences between different OP subgroups defined by the five key genes. When $k = 2$, the clustering results exhibited better separation (Figure 5A). Additionally, considering the Delta area plot and CDF plot (Figure 5B and C), when $k = 2$, the slope of CDF descent and the change in AUC were low. Ultimately, two OP subtypes were identified (cluster1/cluster2) (Figure 5D), with cluster1 comprising 5 samples and cluster2 comprising 7 samples.

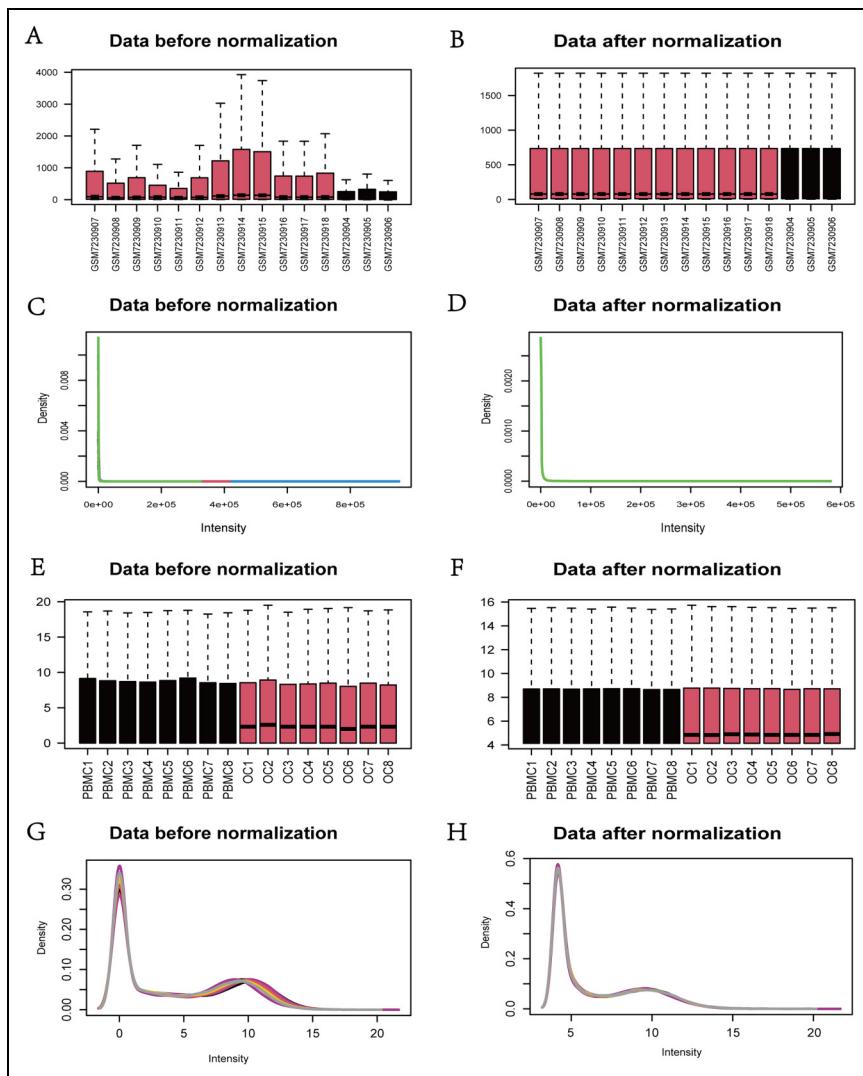


Figure 2. Data normalization. (A–D) Boxplots and density plots illustrate the differences before and after processing the osteoporosis dataset. (E–H) Boxplots show the differences before and after processing the osteoclast differentiation dataset.

Next, to explore the biological functional differences between the two OP subgroups, we initially conducted differential analysis on the two subgroups. Using stringent criteria of $\text{Ilog}_2(\text{fold-change}) > 1$ and $P < 0.05$, a comprehensive set of 1058 DEGs was pinpointed, comprising 333 downregulated genes and 725 upregulated genes (Figure 5E and F), elegantly depicted through both heatmap and volcano plot visualizations. Following this, the 1058 DEGs underwent meticulous scrutiny via functional enrichment

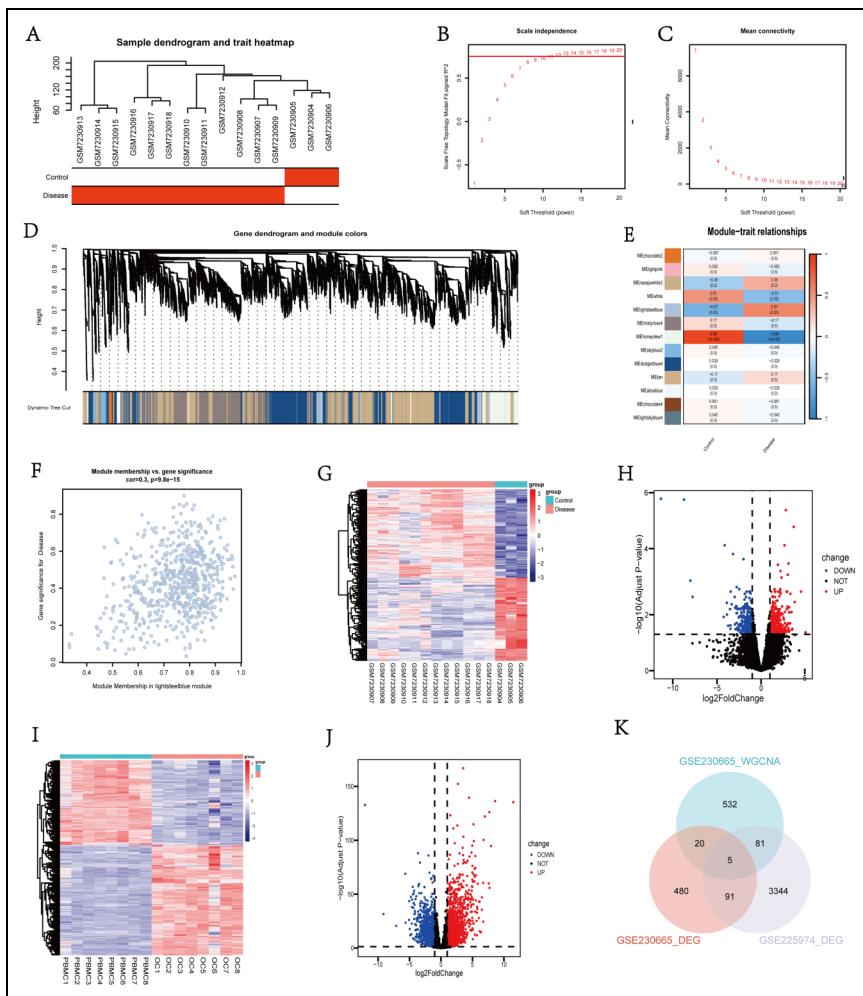


Figure 3. Identification of key genes in osteoclast-mediated osteoporosis progression. (A) Sample clustering, with all samples passing the cutoff threshold and located within clusters; (B, C) soft-thresholding power analysis was conducted to determine the scale-free fit index of network topology; (D) gene co-expression clusters, where modules in the gene co-expression network constructed by WGCNA are represented by different colors; (E) the relationship between two traits and 13 modules; (F) scatterplot illustrating the relationship between MM and GS in the lightsteelblue module; (G, H) heatmaps and volcano plots of DEGs in the GSE230665 dataset; (I, J) heatmaps and volcano plots of DEGs in the GSE225974 dataset; (K) Venn diagram showing the 5 overlapping hub genes.

analysis. The outcomes of GO enrichment analysis were elegantly displayed in both circle plot and heatmap formats (Figure 5G and I), unraveling notable enrichment in terms such as chromosome segregation, mitotic nuclear division, nuclear division,

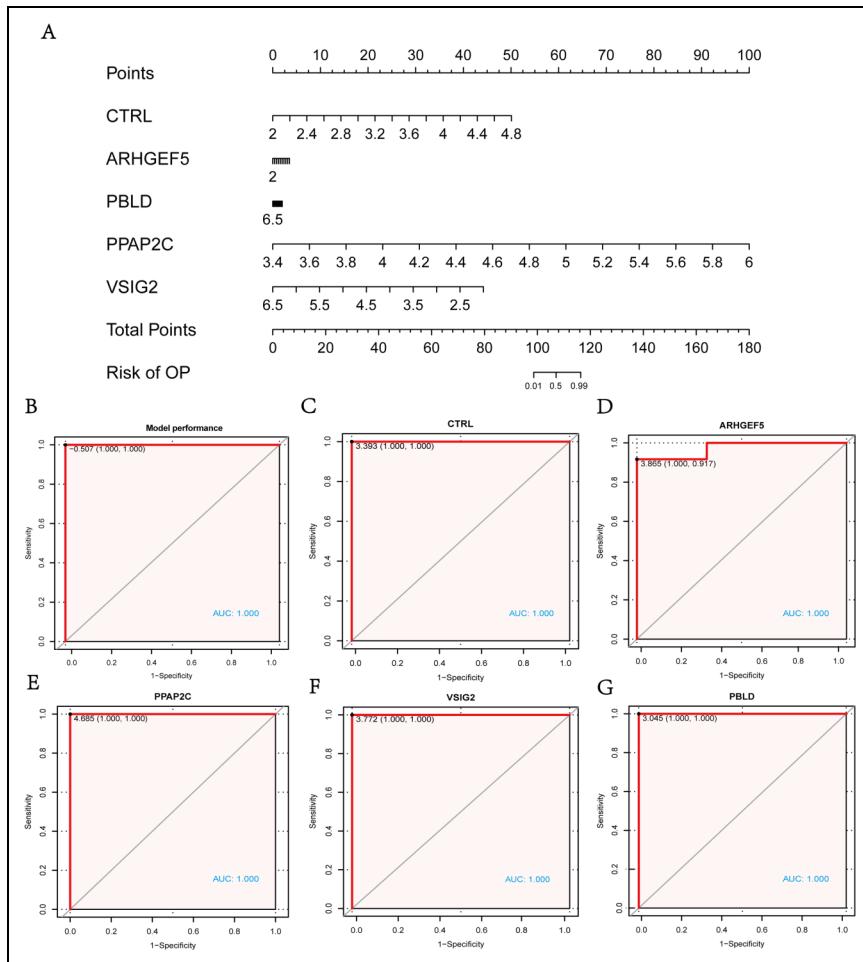


Figure 4. Association of key genes with OP diagnosis. (A) The nomogram illustrates the diagnostic efficacy of key genes for osteoporosis; (B) the ROC curve evaluates the effectiveness of the diagnostic model; (C–G) ROC curves demonstrate the association of the five key genes with OP diagnosis.

organelle fission, and regulation of mitotic nuclear division. Similarly, KEGG pathway enrichment analysis results were vividly portrayed in circle plot and heatmap formats (Figure 5H and J), indicating significant enrichment in pathways including Cell cycle, Human T-cell leukemia virus 1 infection, PPAR signaling pathway, and regulation of lipolysis in adipocytes. Furthermore, GSEA analysis of the DEGs illuminated intriguing patterns: in comparison to cluster1, the upregulated genes in cluster2 were notably enriched in Nakayama soft tissue tumors pca2 dn, Smid breast cancer luminal a up, whereas the downregulated genes exhibited predominant enrichment in graham cml dividing vs normal quiescent up, kobayashi egfr signaling 24hr dn (Figure 5K and L).

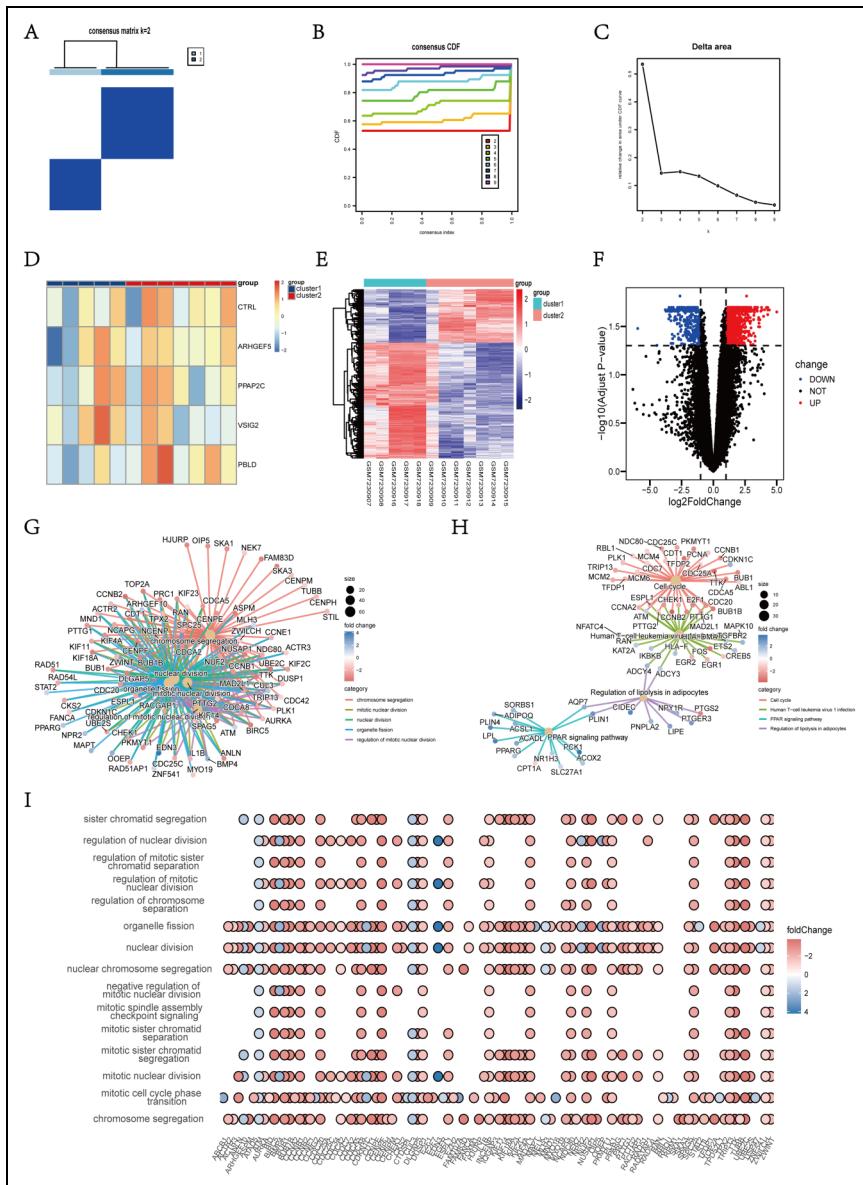


Figure 5. Functional analysis of key genes. (A) Consensus clustering results plot; (B) CDF plot, displaying the curve of CDF changes with the number of clusters; (C) delta area plot, showing the relationship between the AUC of CDF curve and the number of clusters; (D) heatmap of clustering for key genes, where red represents cluster 2 and blue represents cluster 1; (E, F) DEGs of cluster 1 and cluster 2 displayed as heatmaps and volcano plots; (G, I) GO enrichment results of DEGs in cluster 1 and cluster 2 displayed as sunburst plots and heatmaps; (H, J) KEGG enrichment results of DEGs in cluster 1 and cluster 2 displayed as sunburst plots and heatmaps; (K–N) GSEA and GSVA enrichment results of DEGs in cluster 1 and cluster 2.

(continued)

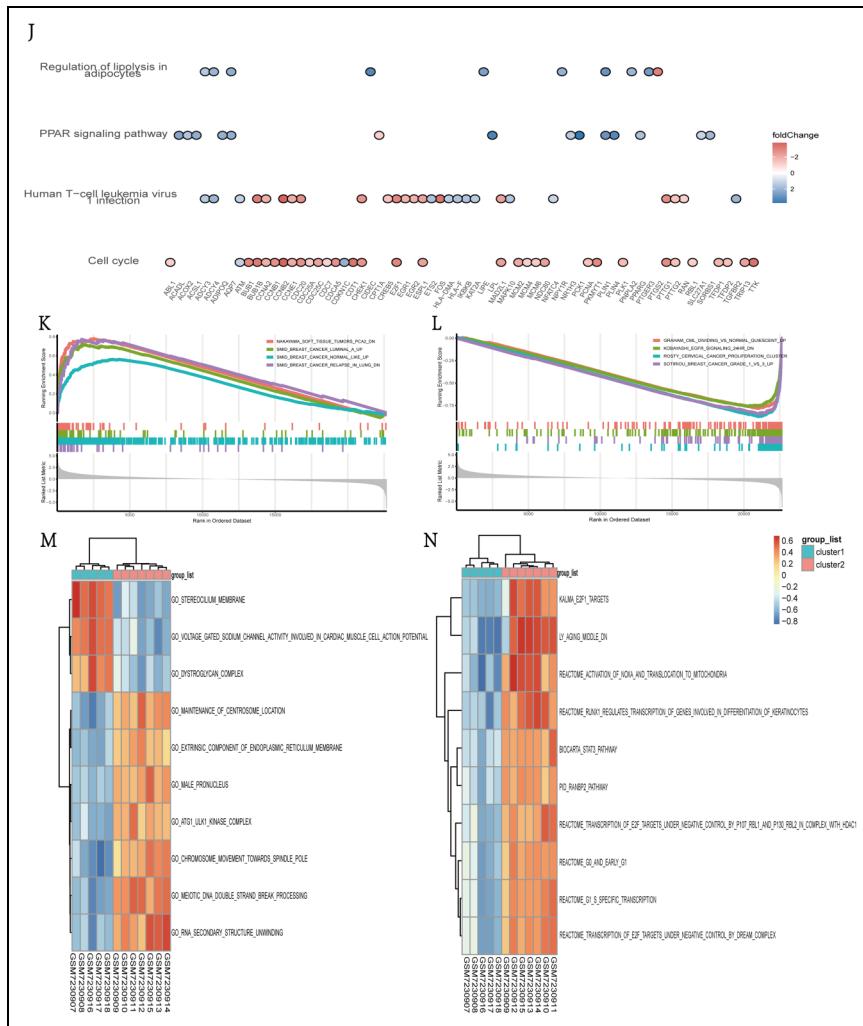


Figure 5. Continued.

Employing a more refined threshold of $|log_2(\text{fold-change})| > 0.8$ and $P < 0.05$, GSVA enrichment analysis uncovered activation of go stereocilium membrane in cluster1, juxtaposed with the activation of kalma e2f1 targets in cluster2 (Figure 5M and N).

Immunological analysis between OP subgroups

Before delving into the disparities in immunological analysis between the two distinct OP subgroups, our initial step was to compute the proportions of various immune cell populations within samples obtained from 12 OP patients using the CIBERSORT function

(Figure 6A). Additionally, correlation analysis was conducted to unveil the intricate associations among various immune cell types (Figure 6B). Following this, attention was directed towards unraveling the correlations between the five pivotal genes utilized as grouping criteria and the diverse immune cell subsets. The findings unveiled CTRL and ARHGEF5 as exhibiting robust correlations with T cells regulatory (Tregs), PPAP2C with T cells follicular helper, VSIG2 with Macrophages M2, and PBLD with T cells regulatory (Tregs) (Figure 6C). Further exploration entailed scrutinizing the discrepancies in immune infiltration between cluster1 and cluster2. The analyses revealed that, relative to cluster1, cluster2 demonstrated an elevated proportion of Mast cells activated, while the proportions of Plasma cells, T cells gamma delta, NK cells activated, Macrophages M1, Macrophages M2, and Mast cells resting were notably diminished (Figure 6D). Moreover, employing ssGSEA for quantifying immune infiltration between the two clusters uncovered that, in contrast to cluster1, cluster2 exhibited heightened proportions of Activated CD4 T cell, Effector memory CD4 T cell, Gamma delta T cell, Immature dendritic cell, and Plasmacytoid dendritic cell, alongside decreased proportions of Effector memory CD8 T cell and T follicular helper cell (Figure 6E).

Validation of differential expression of key genes

To validate the differential expression of the five key genes during osteoclast differentiation, we constructed an osteoclast differentiation model using RANKL. Firstly, the efficacy of the model construction was confirmed by TRAP staining and immunofluorescence detection of F-actin (Figure 7A and C). The results demonstrated that, as the cells were cultured for 7 days compared to day 3, day5, the number of mature osteoclasts occupying the field of view gradually increased, showing statistically significant differences (Figure 7B, D). Subsequently, under the same culture conditions, cell proteins were extracted and the differential expression of the five key genes during osteoclast differentiation was detected via Western blot (Figure 7E). The results revealed that CTRL was significantly downregulated during osteoclast differentiation (Figure 7F), while ARHGEF5, PPAP2C, PBLD, and VSIG2 were all upregulated during osteoclast differentiation (Figure 7G–J). This trend is consistent with what we observed in the osteoclast differentiation dataset (Figure 7K).

Discussion

Studies have shown that osteoclast-mediated bone resorption plays a key role in OP.^{19–21} In this study, to explore the molecular mechanisms, we used WGCNA and differential analysis to identify five key genes (CTRL, PPAP2C, ARHGEF5, VSIG2, PBLD) related to OP and osteoclast differentiation. Using ConsensusClusterPlus, we divided patients into two groups, compared their biological functions and immune infiltration differences, and validated the differential expression of these five genes during osteoclast differentiation via Western blot. To date, there have been no studies on the impact of these five genes on osteoclast differentiation function. Therefore, for research directions aimed at exploring the mechanism or diagnosis and treatment of OP centered on

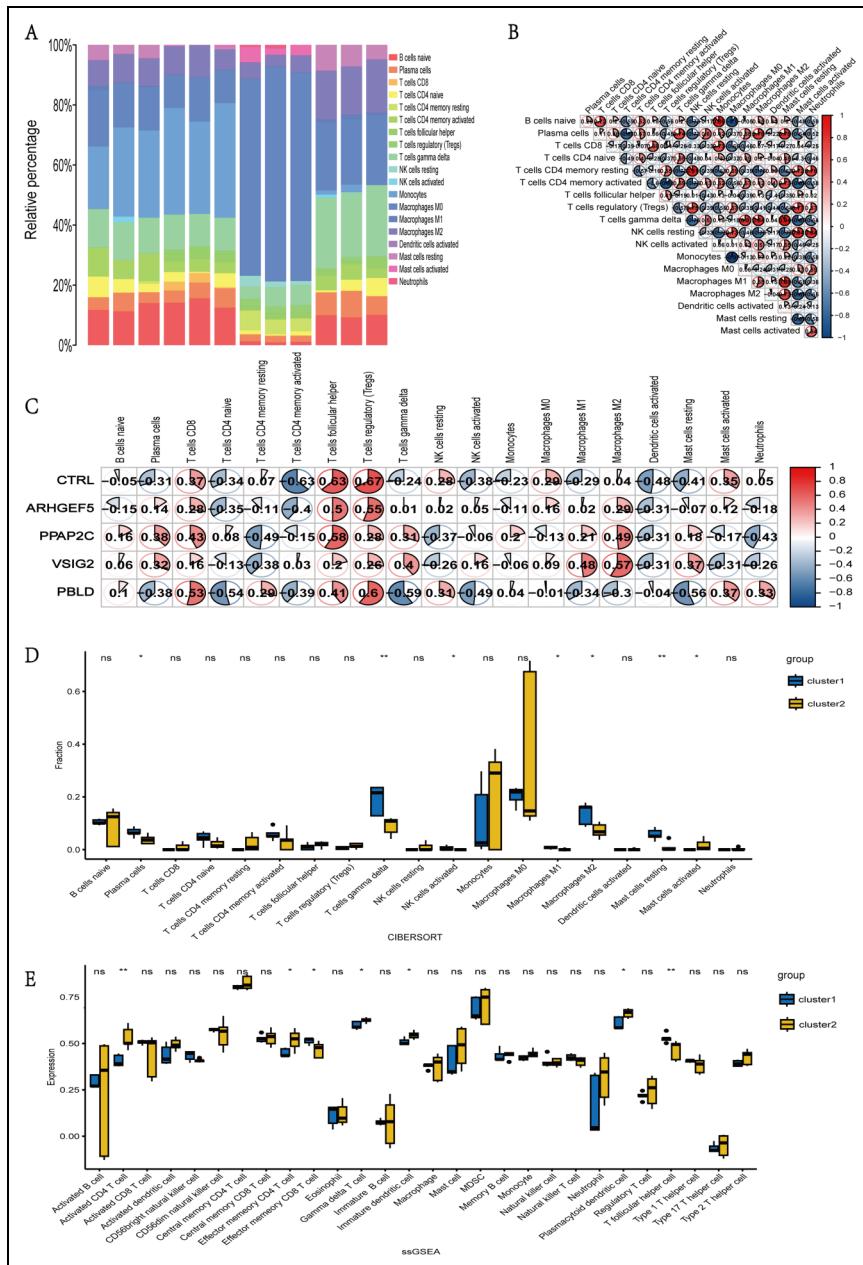


Figure 6. Immunological analysis between OP subgroups. (A) Rainbow plot illustrating immune infiltration in OP samples; (B) correlation results among various immune cells; (C) correlation results between five key genes and various immune cells; (D) CIBERSORT analysis between OP subgroups; (E, F) quantification of OP subgroup immune infiltration using ssGSEA.

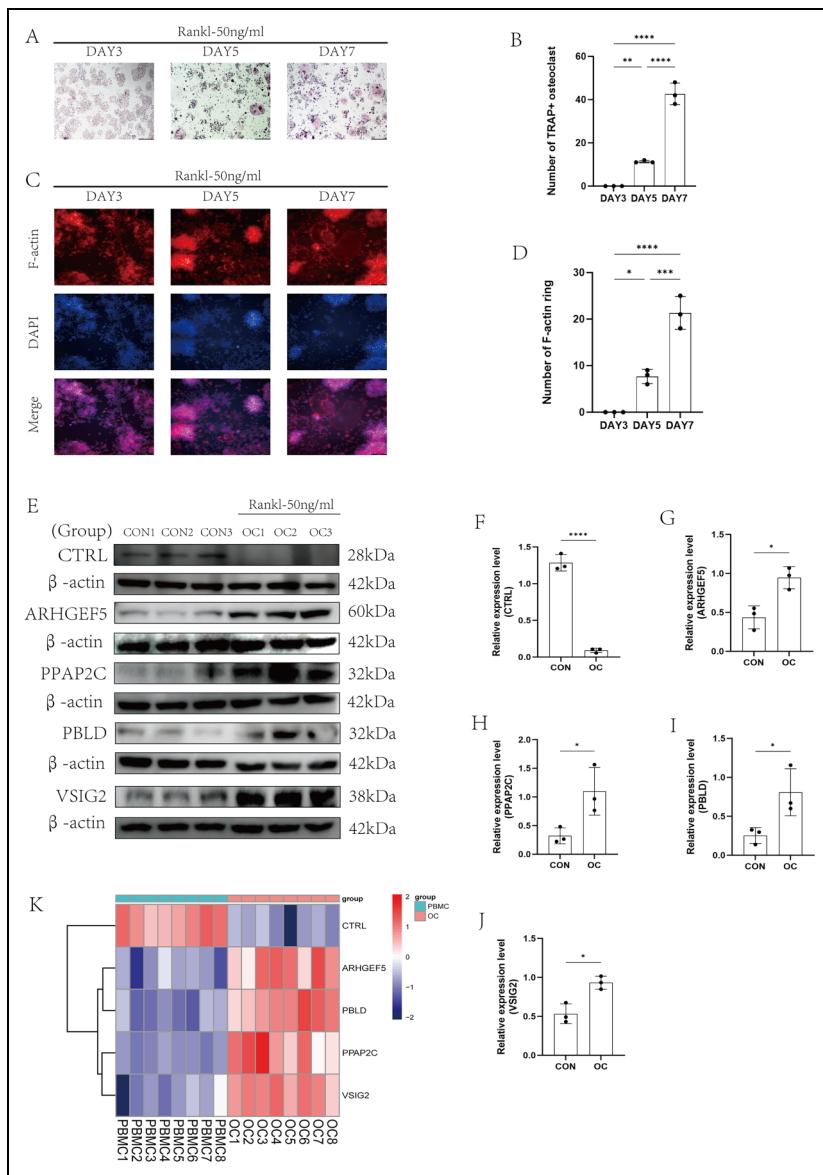


Figure 7. Immunological analysis between OP subgroups. (A) TRAP staining; (B) statistical results of the number of TRAP+ OC at different time points, with $n = 3$ in each group, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, and $****P < 0.0001$. Scale bar, 200 μ m; (C) immune infiltration analysis for F-actin; (D) Statistical results of the number of F-actin ring at different time points, with $n = 3$ in each group, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, and $****P < 0.0001$. Scale bar, 100 μ m; (E–J) Western blot validation results of the five key genes. Osteoclast (OC) group compared to the Control (CON) group, with $n = 3$, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, and $****P < 0.0001$; (K) expression clustering heatmap of the five key genes in the osteoclast differentiation dataset.

osteoclasts, the genes screened and their validated differential expression results in this study may provide a reliable starting point.

Previous studies have shown that ARHGEF5 can strongly activate RhoA,^{22–24} and the star molecule RhoA has been extensively demonstrated to promote osteoclast differentiation.^{25–27} Therefore, we have reason to believe that ARHGEF5 can promote osteoclast differentiation, which is consistent with our cellular level validation results (increased expression during osteoclast differentiation). VSIG2 belongs to the immunoglobulin superfamily (IgSF) and can activate the mTOR signaling pathway,²⁸ while the mTOR/NFATc1 or ERK/mTOR/Beclin1 signaling pathway can promote osteoclast differentiation.^{25,29} Based on this, we also have reason to believe that VSIG2 can promote osteoclast differentiation, which is similarly consistent with our cellular level validation results (increased expression during osteoclast differentiation). Currently, there is limited research on PBLD, but studies have shown that PBLD can inhibit NF- κ B signaling in epithelial cells,³⁰ which can be proved to attenuate osteoclast differentiation.^{31–33} Therefore, PBLD may weaken osteoclast differentiation by inhibiting the NF- κ B signaling pathway. Our cellular level validation results show that PBLD expression is increased during osteoclast differentiation, indicating that theoretically PBLD can promote osteoclast differentiation. This difference may be due to differences in cell types, or PBLD may promote osteoclast differentiation through non-NF- κ B-dependent mechanisms, and the specific mechanism can be explored further through more in-depth cellular experiments. Currently, there is limited research on CTRL and PPAP2C in other cell types and diseases, and gene functions need to be further explored. Cellular level validation results show that, contrary to the expression trend of PPAP2C, CTRL is downregulated during osteoclast differentiation. Therefore, we speculate that CTRL may have the function of inhibiting osteoclast differentiation, serving as a protective factor for OP.

Furthermore, we divided OP patient samples into two subgroups based on the expression characteristics of these five genes. First, we analyzed the biological functional differences between the subgroups. GO enrichment analysis results: chromosome segregation, mitotic nuclear division, nuclear division indicate significant differences in cell cycle and cell metabolism between the subgroups, and bone metabolism is closely related to the occurrence and development of OP.^{34,35} KEGG pathway enrichment analysis results: PPAR signaling pathway, Regulation of lipolysis in adipocytes. The PPAR signaling pathway has also been shown to regulate hepatic fatty acid metabolism,³⁶ indicating significant differences in lipid metabolism between the subgroups. Scholars such as Haemin Kim, Bo Wang, and Li Tian have proposed that lipids in bone and bone marrow play an important role in regulating the differentiation and function of bone cells.,^{37–39} but the mechanism needs further study. Based on these results, it can be inferred that under the influence of lipid metabolism in bone and bone marrow, the degree of osteoclast differentiation between the subgroups will also differ, thereby affecting the progression of both OP groups. GSEA enrichment analysis results: Kobayashi EGFR signaling 24hr dn. EGFR is believed to regulate the proliferation and differentiation of osteoblasts, chondrocytes, and osteoclasts in bone physiology and pathophysiology,⁴⁰ affecting the balance of bone metabolism and thus influencing the development of OP. GSVA enrichment analysis results: kalma e2f1 targets are activated in cluster2. E2F1, as a transcription factor, has been confirmed by multiple studies to

affect osteoclast differentiation, thereby affecting bone loss in OVX-induced mice.^{41,42} Therefore, the degree of osteoclast activation and the progression of OP may differ between the subgroups. There are currently many studies indicating that the interaction between bone and immune cells has an important impact on the progression of OP.^{43,44} Therefore, we further analyzed the differences in immune infiltration between the subgroups. The results showed that compared to cluster1, the proportion of Mast cells activated was higher in cluster2, but the proportions of Plasma cells, T cells gamma delta, NK cells activated, Macrophages M1, Macrophages M2, and Mast cells resting were all lower. Mast cells activated and NK cell activation are involved in nonspecific immunity, while Plasma cells, gamma delta T cells, Macrophages M1, and Macrophages M2 are involved in specific immunity, indicating differences in immune environments between the subgroups. Different types of immunity affect the progression of OP through different pathways and mechanisms.⁴⁵ The above GO, KEGG, GSEA, GSVA, immune infiltration results, and interpretations of the results all indicate that there are differences in the progression of OP or osteoclast activation between the two subgroups. This also indirectly indicates that the subgroups obtained using these five genes as grouping criteria have good discriminative ability.

However, this study also has many limitations. First, the study is of small and not very well-defined OP and control samples. It is not at all clear that this will have general usefulness. Secondly, this study screened out key genes through a series of bioinformatics methods and conducted disease association analysis, functional analysis, and cellular level differential expression validation. However, the impact of each gene on osteoclast differentiation and OP progression was not thoroughly discussed. Finally, only the differential expression of target genes has been validated *in vitro*, and further immunohistochemical validation *in vivo* has not been performed. Of course, these are also the focus of our subsequent research topics. In the future, we will observe the effects of each gene on osteoclast differentiation and OP progression and their specific mechanisms through intervention experiments *in vitro* and *in vivo*.

Conclusion

In summary, we identified five key genes involved in osteoclast-mediated OP progression through a series of bioinformatics methods: CTRL, ARHGEF5, PPAP2C, VSIG2, and PBLD. These genes not only exhibit strong disease relevance but also show significant associations with OP and osteoclast differentiation according to functional enrichment results. Differential expression results from *in vitro* experiments further suggest that these five key genes are highly credible molecular targets for both basic and clinical research on osteoclast-mediated OP.

Acknowledgements

The experiments in this study were conducted at the Ningxia Institute of Bone and Joint Diseases. The authors would like to express their gratitude to the laboratory for providing the experimental facilities and equipment.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical and consent statements

This paper does not involve any ethical issues as we did not study any human or animal subjects, nor did we collect any personal information or sensitive data.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China, Key R&D Project of Autonomous Region, Ningxia Medical University General Hospital “Medical Engineering Special,” Autonomous Region Major Scientific and Technological Achievements Transformation Project (grant numbers 82160433, No.U22A20285, No.2023BEG02018, No. NYZYYG-001, No. 2023CJE09037).

ORCID iDs

Zhihai Zhang  <https://orcid.org/0009-0004-1119-5154>

Qunhua Jin  <https://orcid.org/0009-0008-9911-3211>

References

1. Adejuigbe B, Kallini J, Chiou D, et al. Osteoporosis: molecular pathology, diagnostics, and therapeutics. *Int J Mol Sci* 2023; 24: 2–16.
2. Foessl I, Dimai HP and Obermayer-Pietsch B. Long-term and sequential treatment for osteoporosis. *Nat Rev Endocrinol* 2023; 19: 520–533.
3. Zhang W, Zhao W, Li W, et al. The imbalance of cytokines and lower levels of tregs in elderly male primary osteoporosis. *Front Endocrinol (Lausanne)* 2022; 13: 779264.
4. Tabatabaei-Malazy O, Salari P, Khashayar P, et al. New horizons in treatment of osteoporosis. *Daru* 2017; 25: 2.
5. Agrawal AC and Garg AK. Epidemiology of osteoporosis. *Indian J Orthop* 2023; 57: 45–48.
6. Chen M, Fu W, Xu H, et al. Pathogenic mechanisms of glucocorticoid-induced osteoporosis. *Cytokine Growth Factor Rev* 2023; 70: 54–66.
7. Johnston CB and Dagar M. Osteoporosis in older adults. *Med Clin North Am* 2020; 104: 873–884.
8. Arceo-Mendoza RM and Camacho PM. Postmenopausal osteoporosis: latest guidelines. *Endocrinol Metab Clin North Am* 2021; 50: 167–178.
9. Li Z, Li D, Chen R, et al. Cell death regulation: a new way for natural products to treat osteoporosis. *Pharmacol Res* 2023; 187: 106635.
10. Zhivodernikov IV, Kirichenko TV, Markina YV, et al. Molecular and cellular mechanisms of osteoporosis. *Int J Mol Sci* 2023; 24: 121–264.

11. Kaur M, Nagpal M and Singh M. Osteoblast-n-osteoclast: making headway to osteoporosis treatment. *Curr Drug Targets* 2020; 21: 1640–1651.
12. Yan C, Shi Y, Yuan L, et al. Mitochondrial quality control and its role in osteoporosis. *Front Endocrinol (Lausanne)* 2023; 14: 1077058.
13. Liang B, Burley G, Lin S, et al. Osteoporosis pathogenesis and treatment: existing and emerging avenues. *Cell Mol Biol Lett* 2022; 27: 72.
14. Brown JP. Long-term treatment of postmenopausal osteoporosis. *Endocrinol Metab (Seoul)* 2021; 36: 544–552.
15. Vargas-Franco JW, Castaneda B, Redini F, et al. Paradoxical side effects of bisphosphonates on the skeleton: what do we know and what can we do? *J Cell Physiol* 2018; 233: 5696–5715.
16. Lamy O, Stoll D, Aubry-Rozier B, et al. Stopping denosumab. *Curr Osteoporos Rep* 2019; 17: 8–15.
17. Xie L, Feng E, Li S, et al. Comparisons of gene expression between peripheral blood mono-nuclear cells and bone tissue in osteoporosis. *Medicine (Baltimore)* 2023; 102: e33829.
18. Schneider CA, Rasband WS and Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012; 9: 671–675.
19. Da W, Tao L and Zhu Y. The role of osteoclast energy metabolism in the occurrence and development of osteoporosis. *Front Endocrinol (Lausanne)* 2021; 12: 675385.
20. Deng M, Luo J, Cao H, et al. METTL14 represses osteoclast formation to ameliorate osteoporosis via enhancing GPX4 mRNA stability. *Environ Toxicol* 2023; 38: 2057–2068.
21. Liang QL, Xu HG, Yu L, et al. Binding-induced fibrillogenesis peptide inhibits RANKL-mediated osteoclast activation against osteoporosis. *Biomaterials* 2023; 302: 122331.
22. Wang Z, Kumamoto Y, Wang P, et al. Regulation of immature dendritic cell migration by RhoA guanine nucleotide exchange factor Arhgef5. *J Biol Chem* 2009; 284: 28599–28606.
23. Lee HK, Ji S, Park SJ, et al. Odontogenic ameloblast-associated protein (ODAM) mediates junctional epithelium attachment to teeth via integrin-ODAM-rho guanine nucleotide exchange factor 5 (ARHGEF5)-RhoA signaling. *J Biol Chem* 2015; 290: 14740–14753.
24. Kuroiwa M, Oneyama C, Nada S, et al. The guanine nucleotide exchange factor Arhgef5 plays crucial roles in Src-induced podosome formation. *J Cell Sci* 2011; 124: 1726–1738.
25. Wang J, Xu C, Zhang J, et al. Rhoa promotes osteoclastogenesis and regulates bone remodeling through mTOR-NFATc1 signaling. *Mol Med* 2023; 29: 49.
26. Kim H, Takegahara N and Choi Y. Protocadherin-7 regulates osteoclast differentiation through intracellular SET-binding domain-mediated RhoA and Rac1 activation. *Int J Mol Sci* 2021; 22: 31–56.
27. Zhang J, Jiang T, Zhang Y, et al. Phillygenin prevents osteoclast differentiation and bone loss by targeting RhoA. *Phytother Res* 2024; 38: 1863–1881.
28. Xu J, Quan G, Huang W, et al. VSIG2 promotes malignant progression of pancreatic ductal adenocarcinoma by enhancing LAMTOR2-mediated mTOR activation. *Cell Commun Signal* 2023; 21: 223.
29. Tang H, Zhu S, Chen K, et al. IL-17A regulates autophagy and promotes osteoclast differentiation through the ERK/mTOR/Beclin1 pathway. *PLoS One* 2023; 18: e0281845.
30. Chen S, Liu H, Li Z, et al. Epithelial PBLD attenuates intestinal inflammatory response and improves intestinal barrier function by inhibiting NF-kappaB signaling. *Cell Death Dis* 2021; 12: 563.
31. Feng X, Zhu S, Qiao J, et al. CX3CL1 Promotes M1 macrophage polarization and osteoclast differentiation through NF-kappaB signaling pathway in ankylosing spondylitis in vitro. *J Transl Med* 2023; 21: 573.
32. Wu L, Luo Z, Liu Y, et al. Aspirin inhibits RANKL-induced osteoclast differentiation in dendritic cells by suppressing NF-kappaB and NFATc1 activation. *Stem Cell Res Ther* 2019; 10: 375.

33. Wei R, Cao Y, Wu H, et al. Inhibition of VCP modulates NF-kappaB signaling pathway to suppress multiple myeloma cell proliferation and osteoclast differentiation. *Aging (Albany NY)* 2023; 15: 8220–8236.
34. Li J, Chen X, Lu L, et al. The relationship between bone marrow adipose tissue and bone metabolism in postmenopausal osteoporosis. *Cytokine Growth Factor Rev* 2020; 52: 88–98.
35. Li J, Wei G, Liu G, et al. Regulating type H vessel formation and bone metabolism via bone-targeting oral micro/nano-hydrogel microspheres to prevent bone loss. *Adv Sci (Weinh)* 2023; 10: e2207381.
36. Mao Z, Feng M, Li Z, et al. ETV5 regulates hepatic fatty acid metabolism through PPAR signaling pathway. *Diabetes* 2021; 70: 214–226.
37. Kim H, Oh B and Park-Min KH. Regulation of osteoclast differentiation and activity by lipid metabolism. *Cells* 2021; 10: 201–315.
38. Wang B, Wang H, Li Y, et al. Lipid metabolism within the bone micro-environment is closely associated with bone metabolism in physiological and pathophysiological stages. *Lipids Health Dis* 2022; 21: 5.
39. Tian L and Yu X. Lipid metabolism disorders and bone dysfunction—interrelated and mutually regulated (review). *Mol Med Rep* 2015; 12: 783–794.
40. Schneider MR, Sibilia M and Erben RG. The EGFR network in bone biology and pathology. *Trends Endocrinol Metab* 2009; 20: 517–524.
41. Doi K, Murata K, Ito S, et al. Role of lysine-specific demethylase 1 in metabolically integrating osteoclast differentiation and inflammatory bone resorption through hypoxia-inducible factor 1alpha and E2F1. *Arthritis Rheumatol* 2022; 74: 948–960.
42. Xie X, Hu L, Mi B, et al. Metformin alleviates bone loss in ovariectomized mice through inhibition of autophagy of osteoclast precursors mediated by E2F1. *Cell Commun Signal* 2022; 20: 165.
43. Fischer V and Haffner-Luntzer M. Interaction between bone and immune cells: implications for postmenopausal osteoporosis. *Semin Cell Dev Biol* 2022; 123: 14–21.
44. Wu D, Cline-Smith A, Shashkova E, et al. T-cell mediated inflammation in postmenopausal osteoporosis. *Front Immunol* 2021; 12: 687551.
45. Saxena Y, Routh S and Mukhopadhyaya A. Immunoporosis: role of innate immune cells in osteoporosis. *Front Immunol* 2021; 12: 687037.