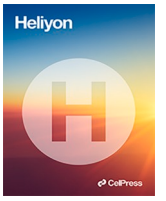
Heliyon 10 (2024) e23675

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



# Identification of the transcriptome signatures and immune-inflammatory responses in postmenopausal osteoporosis

Pan Gao [a,b,c], Xiaoguang Pan [d], Shang Wang [f], Sijia Guo [c], Zhanying Dong [c], Zhefeng Wang [f], Xue Liang [d], Yan Chen [c], Fang Fang [c], Ling Yang [c], Jinrong Huang [c,e], Chenxi Zhang [c], Conghui Li [c,d,e], Yonglun Luo [a,c,d], Songlin Peng [f,**], Fengping Xu [a,b,c,*]

[a] College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
[b] BGI Cell, Shenzhen 518083, China
[c] BGI Research, Shenzhen 518083, China
[d] Qingdao-Europe Advanced Institute for Life Sciences, BGI-Shenzhen, Qingdao 266555, China
[e] Lars Bolund Institute of Regenerative Medicine, BGI-Qingdao, BGI-Shenzhen, Qingdao 266555, China
[f] Department of Spine Surgery, Shenzhen People's Hospital, The First Affiliated Hospital, Southern University of Science and Technology, Shenzhen 518020, Guangdong, China

ARTICLE INFO

*Keywords:*
Postmenopausal osteoporosis
Transcriptome signatures
RNA sequencing
Immune-inflammatory responses
Risk prediction model

ABSTRACT

Postmenopausal osteoporosis is the most common type of osteoporosis in women. To date, little is known about their transcriptome signatures, although biomarkers from peripheral blood mononuclear cells are attractive for postmenopausal osteoporosis diagnoses. Here, we performed bulk RNA sequencing of 206 samples (124 postmenopausal osteoporosis and 82 normal samples) and described the clinical phenotypic characteristics of postmenopausal women. We then highlighted the gene set enrichment analyses between the extreme T-score group and the heathy control group, revealing that some immune-inflammatory responses were enhanced in postmenopausal osteoporosis, with representative pathways including the mitogen-activated protein kinase (NES = 1.6, FDR <0.11) pathway and B_CELL_RECEPTOR (NES = 1.69, FDR <0.15) pathway. Finally, we developed a combined risk prediction model based on lasso-logistic regression to predict postmenopausal osteoporosis, which combined eleven genes (*PTGS2*, *CXCL16*, *NECAP1*, *RPS23*, *SSR3*, *CD74*, *IL4R*, *BTBD2*, *PIGS*, *LILRA2*, *MAP3K11*) and three pieces of clinical information (age, procollagen I N-terminal propeptide, β isomer of C-terminal telopeptide of type I) and provided the best prediction ability (AUC = 0.97). Taken together, this study filled a gap in the large-scale transcriptome signature profiles and revealed the close relationship between immune-inflammatory responses and postmenopausal osteoporosis, providing a unique perspective for understanding the occurrence and development of postmenopausal osteoporosis.

* Corresponding author. College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China.
** Corresponding author. Department of Spine Surgery, Shenzhen People's Hospital, The First Affiliated Hospital, Southern University of Science and Technology, Shenzhen 518020, Guangdong, China.
*E-mail addresses:* songlin824@gmail.com (S. Peng), xufengping@genomics.cn (F. Xu).

https://doi.org/10.1016/j.heliyon.2023.e23675
Received 26 November 2022; Received in revised form 25 November 2023; Accepted 9 December 2023
Available online 12 December 2023

## 1. Introduction

Osteoporosis (OP) is a systemic bone disease characterized by a damaged bone microstructure and decreasing bone mass, leading to an increased risk of fragility fractures [1,2]. OP is more common in women and is largely related to the rapid decrease in ovarian estrogen secretion that could cause 3–5% bone loss in postmenopausal women within 5–10 years, ultimately resulting in postmenopausal osteoporosis (PMOP) [3,4]. Although the incidence of fractures displays variability across different countries, nearly 50 % of women with an average age exceeding 50 years exhibit susceptibility to fractures [5].

Understanding the molecular mechanisms of OP holds crucial significance for the advancement of therapeutic interventions [6,7]. Bone homeostasis is mainly maintained by the balance of osteoblasts and osteoclasts, which are the basis of the stability of the skeletal system [8]. Previous studies have shown that estrogen deficiency affects bone metabolism via the osteoprotegerin (OPG)-receptor activator of the NF-kB ligand (RANKL)-receptor activator of the NF-kB (RANK) axis, resulting in enhanced RANKL signaling and osteoclast differentiation while also inhibiting bone formation by reducing the induction of OPG [9,10]. In addition, several inflammatory and immune pathways are closely related to OP pathogenesis. The reduction in estrogen stimulates the production of inflammatory cytokines, i.e., IL-1, IL-6, and IL-7, which can act on osteoblasts and cause bone loss [11]. Estrogen receptors on the surface of T cells, osteoblasts, and osteoclasts can be directly activated through a classical receptor pathway [12]. However, most of these findings are based on ovariectomy models of inbred mouse strains and need to be carefully validated in humans.

PMOP is the most common form of OP and skeletal disease in women. A better understanding of the transcriptome alterations associated with PMOP might provide new clues for the development of prevention and treatment strategies. To date, very few RNA sequencing (RNA-seq) studies have been conducted. According to the PubMed database, only 26 studies using the keywords “RNA sequencing” or “RNA-seq” and “postmenopausal osteoporosis” have been documented. In contrast, there are over 16,000 documented publications containing the keywords “RNA sequencing” or “RNA-seq” and "cancer". In addition, some studies of transcriptome changes in PMOP only included a small sample size to explore the expression pattern [13,14]. Di et al. [15] later identified the differential expression profiles of multiple transcripts of a few peripheral blood mononuclear cell (PBMC) samples by RNA sequencing. These studies indicated that the immune system plays an important role in PMOP. Recent developments in massively parallel sequencing tools [16] make it possible to conduct a profound study of genome-wide transcriptome profiles in PBMC samples of PMOP patients in a large cohort. In particular, the relationship between the immune response and PMOP requires further investigation. Relevant research has also shown that larger improvements in BMD are associated with greater reductions in fracture risk, particularly for vertebral and hip fractures [17]. Vladyslav et al. found that the osteoporotic fracture probability in 50-year-old women with an extreme T-score (less than −3.5) was 13 %, much higher than that in women with a T-score $\leq -2.5$ [18]. Dargent et al. used 6933 EPIDOS study data points and found that elderly women with extreme T-scores (less than −3.5) have a risk of hip fracture more than two times higher than women of the same age [19]. However, there is no systematic study on the transcriptome signatures of extreme T-score patients. Due to the possibility of severe clinical outcomes in these patients, we believe that conducting transcriptome research on them is also meaningful.

In this study, we performed a transcriptome analysis on PMOP and healthy controls (CTL) in a large-scale data set using RNA-seq tools. We explored differentially expressed genes (DEGs) in all samples and focused on revealing the transcriptome signatures for the extreme T-score PMOP group. We also identified 137 DEGs that are closely associated with PMOP and are involved in some osteoporosis-related pathways. Interestingly, we further found that the immune and inflammatory pathways were enhanced in PMOP. Finally, we built a risk prediction model based on eleven risk genes and three pieces of clinical information, which exhibited high prediction ability (AUC = 0.97). This study further deepened our understanding of the potential mechanism of PMOP and provided effective research resources for the further development of therapeutic drugs or novel diagnostic methods.

## 2. Materials and methods

### 2.1. Participants and samples

This prospective study was conducted utilizing a cohort of 206 women aged 50 and older. This cohort comprised 124 postmenopausal osteoporosis PBMC samples (T-score $\leq -2.5$), and 82 normal bone mass PBMC samples serving as the controls (T-score $\geq -1$). All participants were treatment-naive, and their clinical information is summarized in the supplementary information.

### 2.2. RNA extraction and library construction

Total RNA was extracted from PBMCs using TRIzol (Ambion) according to the manual. DNase I was used to digest double-stranded and single-stranded DNA in total RNAs. Next, magnetic beads were purified to recover the reaction products, and the RNase H method was used to remove the rRNA. Through cDNA synthesis, end repair, A addition, adaptor ligation, and several rounds of PCR amplification with PCR mix were performed to enrich the cDNA fragments. Then, we heated and denatured the recovered double-stranded PCR products and cyclized them with the splint oligo sequence(5′-ACGTACTGAGAGGCATGGCGACCT-3′). The single-stranded circular DNAs were formatted as the final library. We also used an Agilent 2100 Bioanalyzer (Agilent DNA 1000 Agents) to detect the fragment size and RNA integrity (RIN >7). The final library was amplified with phi29 (Thermo) to make DNA nanoballs, which were loaded into the patterned nanoarray, and single-end 50-base reads were generated on the BGISEQ-500 platform. See more details: Supplementary Information 3–4.

### 2.3. Data processing

Paired-end sequencing was carried out, and initial quality control of raw reads was conducted using fastp v0.23.1 software [20]. This encompassed filtering out low-quality bases (Q20 < 0.85), removal of adaptors and base correction. Subsequently, the clean reads were directly quantified by salmon v1.9.0 software [21] based on Homo_sapiens.GRCh38.100.gtf.gz. This method could ensure the fast quantification of transcripts. The study cohort was uniformly of a single gender, and the age ranged from 50 to 80 years old. To enhance the reliability of the results, we selected the higher genes that were expressed in at least 20 % of all samples, each with an average read count exceeding 1. Genes with lower expression were subsequently excluded from further analysis.

### 2.4. Gene set enrichment analysis

Gene set enrichment analysis was conducted using the clusterProfiler R package [22]. It involved preranked gene set enrichment analysis based on log2-fold change ($log_2FC$). In addition, the functional pathways were determined through KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis applied to the candidate DEGs. The KEGG pathways exhibiting a P (adjusted) < 0.05 were considered to be significantly enriched.

### 2.5. GEO data source

The gene microarray expression data from GSE56815 were procured, including the gene expression data of blood monocytes originating from 20 postmenopausal women with a high BMD and 20 postmenopausal women with a low BMD. Additionally, the study incorporated data from GSE149938, comprising single-cell RNA sequencing data spanning 32 human blood cells sourced from 21 healthy donors. These datasets were downloaded from the Gene Expression Omnibus (GEO) database, and our application complied with the guidelines and policies of each database.

### 2.6. Risk prediction model

To further evaluate the proportional impact of risk genes on PMOP, these genes were used as dependent variables in constructing a

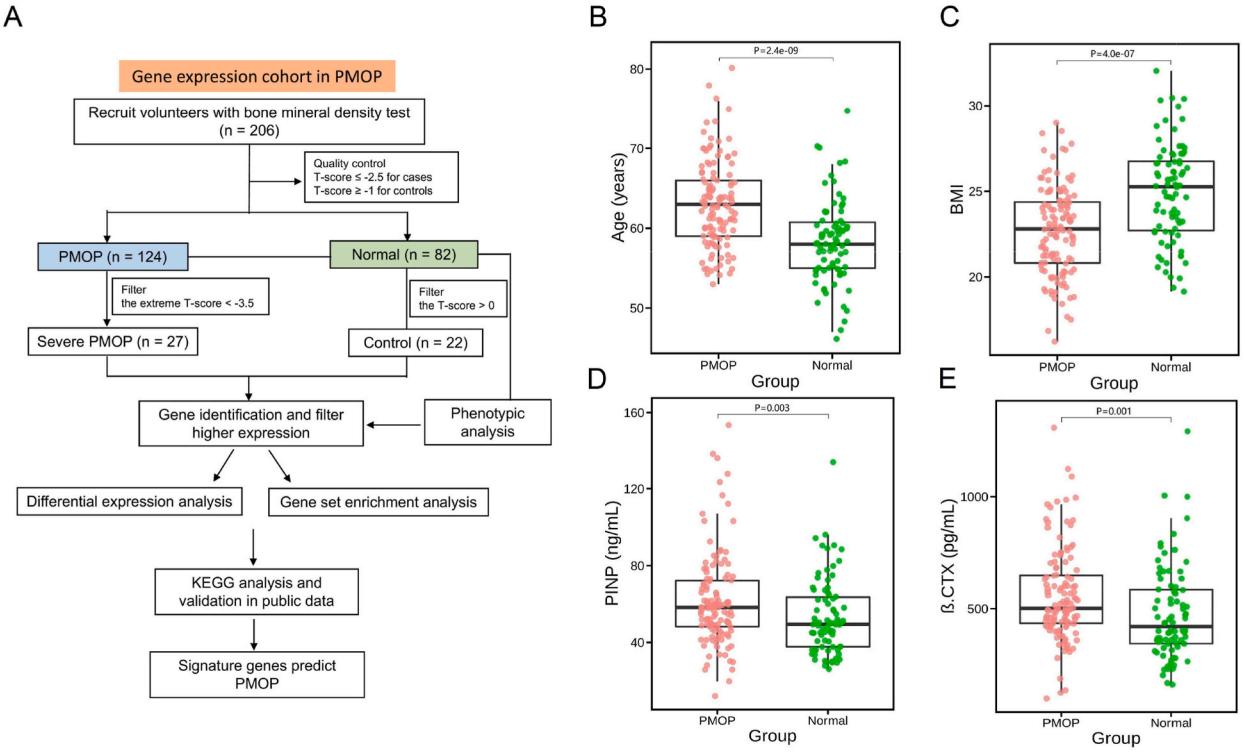



**Fig. 1.** Description of the postmenopausal osteoporosis (PMOP) transcriptome cohort and characteristics of clinical information. (A) Workflow for transcriptome profiling in the PMOP cohort. (B) Boxplot of age comparison between the PMOP and normal groups (with median and quartile). (C) Boxplot of body mass index (BMI) comparison between the PMOP and normal groups (with median and quartile). (D) Boxplot of Procollagen I N-Terminal propeptide (PINP) comparison between the PMOP and normal groups (with median and quartile). (E) Boxplot of β isomer of C-terminal telopeptide of type I (β.CTX) comparison between the PMOP and normal groups (with median and quartile).

lasso-logistical regression model, adhering to the minimalist model principle. The optimal variables were identified using the lasso R package. Subsequently, a risk prediction model was formulated, incorporating a linear combination of the risk gene expression levels and the corresponding regression coefficients (β) determined from the lasso-logistical regression model. The formula for the risk score calculation was as follows: risk score = expression of gene1 × β1(gene1) + expression of gene2 × β2(gene2) + … expression of gene n × βn (gene n) [23]. Finally, the risk prediction model was established, and the predictive accuracy was evaluated through the construction of a receiver operating characteristic curve (ROC) using the ROCR package within R software.

### 2.7. Statistics analysis

All statistical analyses were performed using R version 4.0.3. The Wilcoxon test was predominantly employed to analyze differences in gene expression, and the false discovery rate (FDR) was rigorously controlled for multiple comparisons. Gene read counts were standardized as transcripts per kilobase of exon model per million mapped reads (TPM). The DEseq2 package (v1.36.0) was used to identify DEGs [24], with a threshold of $P < 0.05$ (corrected P value by FDR) and $|\log_2 FC| > 1$. Univariate logistical analysis was conducted to assess the correlation between gene predictors and PMOP within our cohort. Odds ratios exceeding 1 were indicative of risk genes, while OR values below 1 indicated protective genes. The diagnostic accuracy of potential biomarkers was evaluated by ROC curves, with the corresponding area under the curve (AUC) computed. ComplexHeatmap and ggplot2 packages were used for visualization of the results.

## 3. Results

### 3.1. Description of the PMOP transcriptome cohort and phenotypic analysis

The discovery cohort consisted of 206 women ranging in age from 45 to 80, including 124 PMOP patients (T-score $\leq -2.5$) and 82 normal individuals (T-score $\geq -1$), based on the traditional T-score standard (Fig. 1A; Supplementary table 1). We collected PBMC samples and basic clinical information (age and BMI), and we measured the T-score of the lumbar vertebra (L1-L4), calcium ($Ca^{2+}$), magnesium ($Mg^{2+}$), phosphate (Pi), cholesterol (CHOL), triglyceride (TG), high-density lipoprotein (HDL), and low-density lipoprotein (LDL) for each participa nt (Supplementary table 2). We also analyzed two plasma auxiliary biomarkers: procollagen I N-terminal propeptide (PINP) and the β isomer of C-terminal telopeptide of type I (β.CTX), which are important biomarkers reflecting the level of bone turnover [25,26].

Clinical phenotypic information revealed clear phenotypic differences between the PMOP and normal groups (Table 1). The average age of the PMOP group was significantly higher than that of the normal group (Wilcoxon test, $p = 2.4e\text{-}09$) (Fig. 1B), and we also found that the PMOP group had a lower BMI value (Wilcoxon test, $p = 4.0e\text{-}07$) (Fig. 1C). Interestingly, the two bone turnover biomarkers (PINP and β.CTX) were significantly different in the two groups of middle-aged and elderly women. The changes were consistent with the PMOP group showing higher values than the normal group (Wilcoxon test, $p = 0.003$ for PINP, $p = 0.001$ for β.CTX) (Fig. 1D and E). The T-scores of different parts of the lumbar spine (L1-L4) and cervical spine were different, and the T-score of the PMOP group was significantly lower than that of the control group. In addition, we also found that the plasma $Ca^{2+}$, $Mg^{2+}$, Pi, CHOL, and other indicators in the PMOP population were not significantly changed compared with the controls, which may indicate that bone metabolism activity has little effect on these plasma clinical indicators and that their clinical diagnostic significance is not clear.

**Table 1**
Clinical characteristics of patients in the cohort.

| Clinical features | The postmenopausal osteoporosis transcriptome cohort (n = 206, mean ± SD) | | |
|---|---|---|---|
| | PMOP (n = 124) | Control (n = 82) | *P* value (PMOP vs CTL) |
| Age (years) | 63.00 ± 5.35 | 58.28 ± 5.18 | 2.391e-09 |
| BMI (kg/m$^2$) | 22.68 ± 2.59 | 24.64 ± 3.95 | 3.969e-07 |
| PINP (ng/mL) | 61.81 ± 24.00 | 53.29 ± 9.71 | 0.003 |
| β.CTX (pg/mL) | 557.7 ± 206.4 | 476.6 ± 207.9 | 0.001 |
| Lumbar vertebra | | | |
| L1 | −2.41 ± 1.07 | 0.00 ± 0.71 | 2.2e-16 |
| L2 | −2.88 ± 0.65 | 0.19 ± 0.77 | 2.2e-16 |
| L3 | −2.61 ± 0.88 | 0.49 ± 0.88 | 2.2e-16 |
| L4 | −2.41 ± 0.91 | 0.71 ± 1.02 | 2.2e-16 |
| Cervical vertebra | −1.97 ± 0.76 | −0.01 ± 0.65 | 2.2e-16 |
| T-score | −3.20 ± 0.50 | −0.38 ± 0.50 | 2.2e-16 |
| $Ca^{2+}$ (mmol/mL) | 2.32 ± 0.1 | 2.32 ± 0.10 | 0.71 |
| $Mg^{2+}$ (mmol/mL) | 0.87 ± 0.07 | 0.86 ± 0.06 | 0.91 |
| $P^{3+}$ (mmol/mL) | 1.16 ± 0.12 | 1.13 ± 0.12 | 0.05 |
| CHOL | 5.40 ± 1.08 | 5.46 ± 1.06 | 0.59 |
| TG | 1.44 ± 0.89 | 1.51 ± 0.88 | 0.44 |
| HDL | 1.44 ± 0.33 | 1.37 ± 0.28 | 0.16 |
| LDL | 2.94 ± 0.85 | 3.03 ± 0.75 | 0.22 |

### 3.2. Transcriptome profiling in extreme T-score PMOP and control groups

In this study, we first performed DEG analysis comprising all PMOP and controls. Age batch effects were removed, but no significant DEGs were identified, and the sample PCA showed that the first two principal components could only explain the lower variatio n (Supplementary fig. 1). Reasoning that the transcription level in PBMCs may be interfered with by many factors, such as individual heterogeneity and physiological and biochemical indicators beyond PMOP [27], it was not unexpected that a base-control comparison failed to capture any DEGs. Considering the clinical significance of the extreme T-score group, we next focused on the transcriptome signatures in the extreme T-score group based on the above facts of increased clinical fracture risk.

To enrich PMOP transcriptome signatures, we screened 27 extreme PMOP samples from 124 PMOP samples according to the extreme T-score standard (T-score $< -3.5$), which were compared with the 22 normal samples (T-score$>0$) for 82 controls (Supplementary table 3). The PCA showed that the first two principal components could explain 44.8 % of the variation, which means that the extreme T-score had good performance (Supplementary fig. 2). We performed a differentially expressed gene analysis and identified 137 DEGs in extreme T-score PMOP compared to the controls (Fig. 2A; Supplementary table 4). A gene ontology analysis with these 137 DEGs identified 56 significantly upregulated and 3 downregulated pathways, of which 3 upregulated pathways are known to be associated with PMOP, including the osteoclast differentiation pathway, IL-17 signaling pathway, and mitogen-activated protein kinase (MAPK) signaling pathway (Fig. 2B; Supplementary table 5) [28–30]. Interleukin-17 (IL-17) is a typical cytokine that plays an important role in resistance to extracellular bacterial and fungal infections and the pathogenesis of various autoimmune diseases [31]. In addition, several immune-associated pathways, such as B-cell receptor signaling, Th17 cell differentiation, and Toll-like receptor signaling, were also regulated in PMOP.

Significant changes in transcriptome profiles between the extreme T-score group and the control group accounted for these pathways, which may be closely related to the progression of PMOP. For the five representative genes with significant differences between the three pathways, gene expression levels tended to be upregulated in PMOP (Fig. 2C–E), which indicates that these genes may enhance osteoclastic differentiation or weaken osteogenic differentiation, such as significant and common PMOP-associated genes. To validate how the JUN gene changes during PMOP progression, we performed an integrated microarray analysis using public GEO data (GSE56815, n = 40). JUN was upregulated in cases of PMOP progression, which was consistent with our bulk RNA-seq

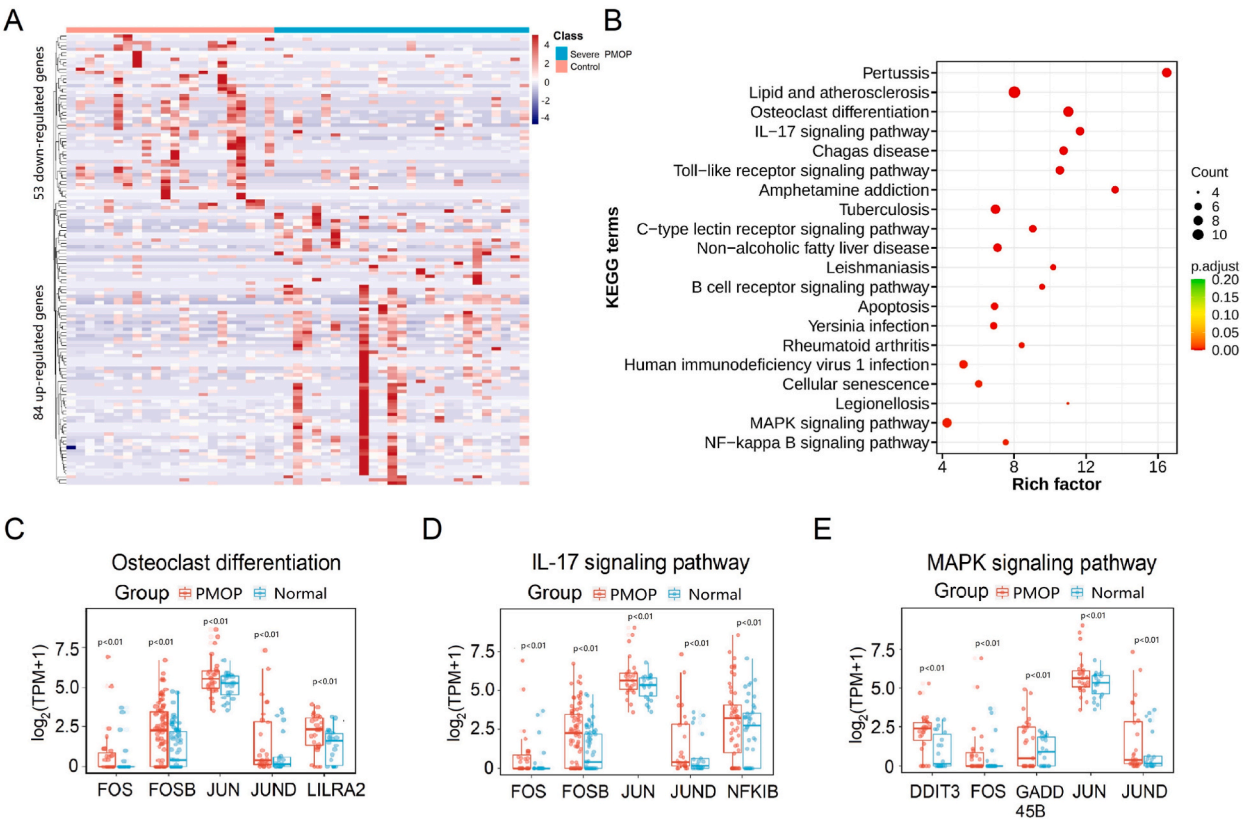



**Fig. 2.** Genome-wide transcriptome profiling and functional enrichment analysis in the extreme T-score PMOP and normal groups (n = 49). (A) The heatmap representing heterarchical clusters based on differential expression genes. (B) The top 20 upregulated KEGG pathways based on differential genes. (C) Comparison of expression levels of five representative genes in the osteoclast differentiation pathway. (D) Comparison of expression levels of five representative genes in the IL-17 signaling pathway. (E) Comparison of expression levels of five representative genes in the MAPK signaling pathway.

results (Supplementary fig. 3). In addition, we also explored the expression patterns of the JUN gene in the blood scRNA-seq data (GSE149938). The results showed that the JUN gene was highly expressed in all subclusters of monocytes compared with other cell types, as well as higher expression in classical monocytes and interM (intermediate) (Supplementary fig. 4).

### 3.3. Gene set enrichment analysis highlights immune and inflammatory pathways in PMOP

To further gain a system level understanding of the relationship between gene sets and the extreme T-score PMOP, we applied gene set enrichment analysis (GSEA) to identify differential gene sets (Fig. 3A). We identified 1402 gene sets that were enriched in cases of T-score PMOP at a false discovery rate (FDR) < 0.25. Thirty-eight eligible gene sets were obtained from KEGG databases (Fig. 3B). The results show that many active genes in the controls were downregulated in the extreme T-score PMOP group and vice versa. The most highly enriched terms were inflammatory and immune gene sets, concordant with their promotion of PMOP. A similar pattern was also observed for other gene sets, such as FC_EPSILON_RI (NES = 1.65, FDR<0.12), NOD_LIKE_ RECEPTOR (NES = 1.58, FDR<0.11), and CHEMOKINE (NES = 1.57, FDR<0.11), which may mediate inflammatory responses. This suggests that the changed phenotype observed in the extreme T-score PMOP group and the controls was related to inflammatory effects (Fig. 3C–E). The most highly enriched inflammatory terms in extreme T-score PMOP indicate that the occurrence of PMOP is accompanied by the enhancement of

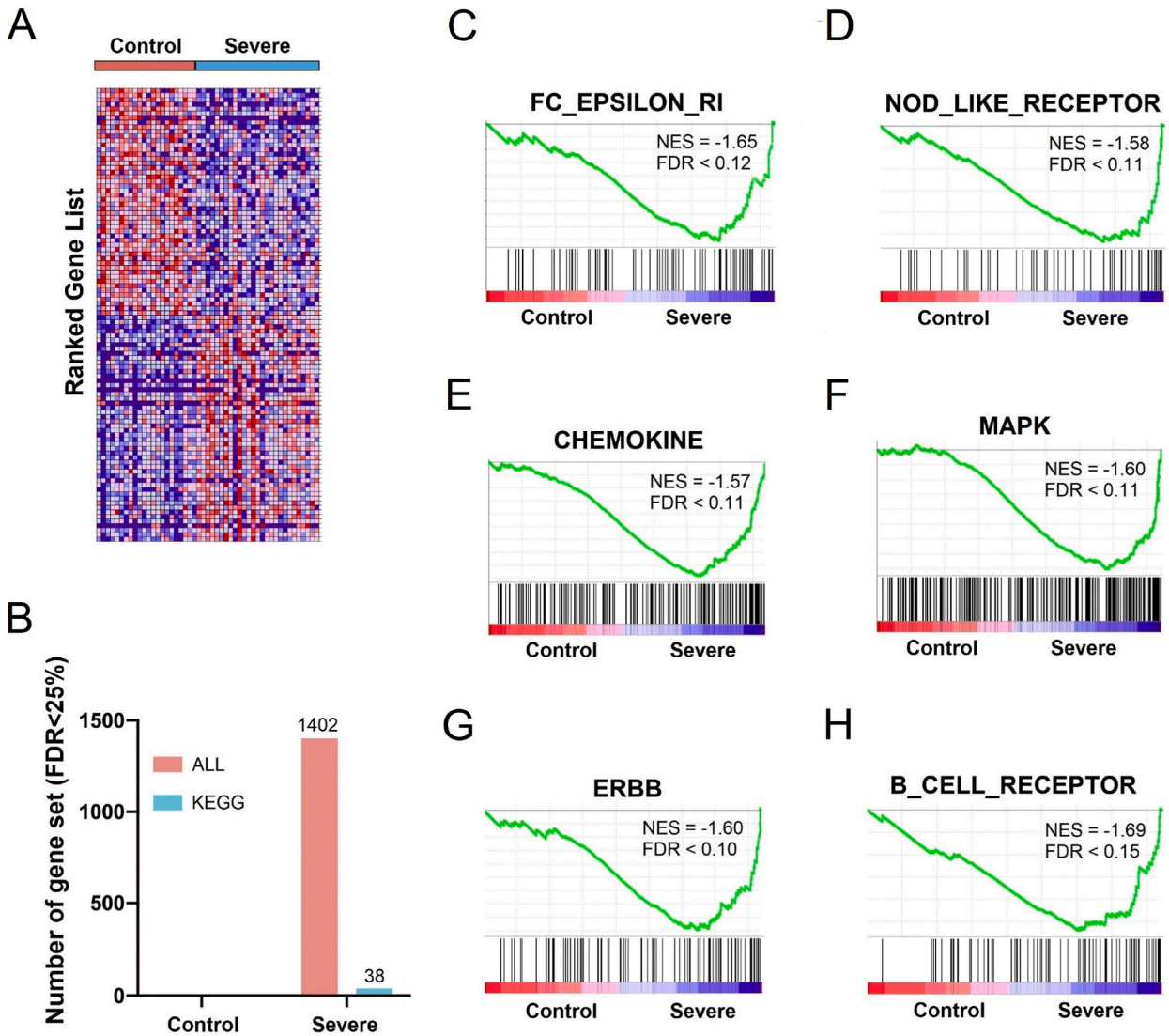



**Fig. 3.** Gene set enrichment analysis based on the extreme T-score PMOP and the controls (n = 49). (A) Heatmap of ranked gene list between the extreme T-score PMOP and normal samples. (B) Number of total key gene sets in the extreme T-score group and the controls (FDR <0.25). (C–G) GSEA of up-regulated genes in the extreme T-score group for FC-EPSILON_RI, NON_LIKE_RECEPTOR, CHEMOKIN, MAPK and ERBB related to inflammatory response. (H) GSEA of up-regulated genes in the extreme T-score PMOP for B_CELL_ RECEPTOR related to immunity response.

inflammatory responses.

MAPK (NES = 1.6, FDR<0.11) and ERBB (NES = 1.6, FDR<0.10) were also affected (Fig. 3F and G). The deletion of MAPK-, p38a-, or p38b-encoding genes of MAPK pathway members results in a degree of bone loss due to osteoblast differentiation defects in mice [32]. Our results show that the MAPK pathway was upregulated in extreme T-score cases, which is consistent with the results of previous pathway enrichment analyses. Our study further revealed a positive effect of the MAPK cell differentiation function on extreme T-score PMOP. It was interesting to find that the B_CELL_RECEPTOR (NES = 1.69, FDR<0.15)-related gene set members were upregulated in cases (Fig. 3H), implying that osteoporosis in postmenopausal women may lead to the significant expansion of the B-cell population. This is in agreement with early research and further supports the concept that humoral immunity plays a positive role in PMOP [33]. Our study may provide additional evidence that B cells positively promote estrogen-deficient osteoporosis based on large-scale RNA-seq data.

### 3.4. Assessment of immune cell ratios between the PMOP and normal groups

We analyzed the infiltration of immune cells between the PMOP and normal groups. The CIBERSORT algorithm was used to assess the data, and we drew a proportional component diagram of 22 representative immune cells in all the samples (Fig. 4A). The results showed that the largest proportions of the three cell types were NK cells, monocytes, and T cells, which basically conformed to the composition features of PBMCs. Furthermore, we also compared the proportion of different cell types in different samples (Fig. 4B),

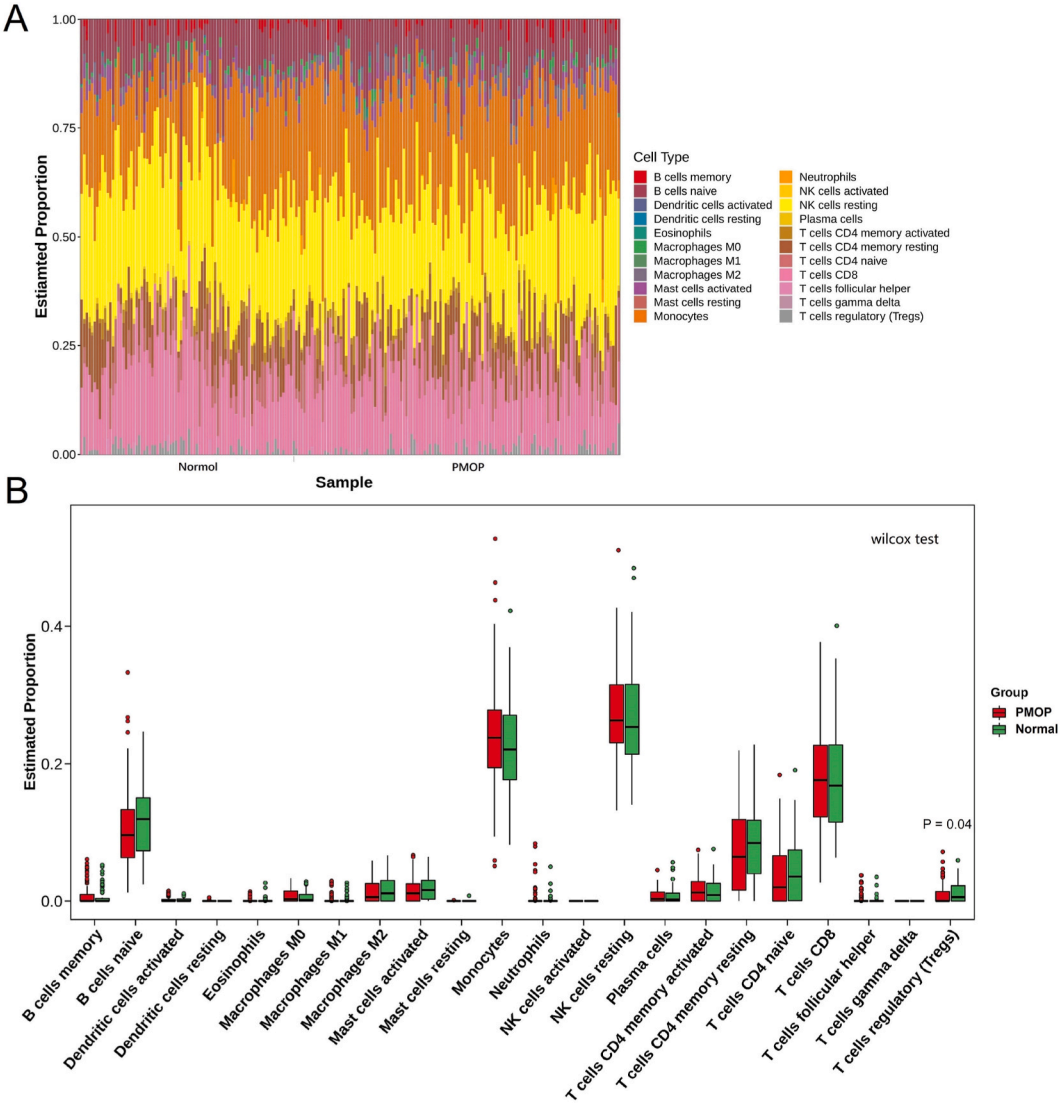



**Fig. 4.** Profiling peripheral blood mononuclear cells (PBMCs) with CIBERSORT based on all samples (n = 206). (A) Composition of 22 immune cell types in the PMOP and normal samples. (B) Comparison of the proportion of 22 cell types in different samples.

and we found a significant difference in the proportion of Treg cells between the normal group and PMOP group (P = 0.04). There was no significant change in the proportion of other T-cell types from the normal group to the PMOP group. In addition, there was no significant difference in the proportion of other cell types between the PMOP group and the normal group.

### 3.5. A risk prediction model with a good effect for PMOP

To generate a better PMOP risk model, we started with the top 15 differential genes (between extreme T-score PMOP and the controls), which were selected as the potential disease predictors in our cohort. In addition, we also included representative genes of the MAPK, B_CELL_RECEPTOR, IL-17, and NOD_LIKE_RECEPTOR pathways in our regression equation. A univariate logistical regression and cross validation analysis determined that eleven genes were significantly associated with PMOP: *PTGS2* (OR = 1.3, CI = 1.06–1.72, $p = 0.03$), *CXCL16* (OR = 1.32, CI = 1.08–1.71, $p = 0.02$), *NECAP1* (OR = 1.66, CI = 1.17–2.58, $p = 0.01$), *RPS23* (OR = 0.42, CI = 0.15–0.79, $p = 0.03$), *SSR3* (OR = 0.78, CI = 0.61–0.97, $p = 0.03$), *CD74* (OR = 1.03, CI = 1.01–1.07, $p = 0.01$), *IL4R* (OR = 1.38, CI = 1.08–1.89, $p = 0.02$), *BTBD2* (OR = 1.16, CI = 1.03–1.33, $p = 0.03$), *PIGS* (OR = 1.94, CI = 1.18–4.62, $p = 0.04$), *LILRA2* (OR = 1.34, CI = 1.08–1.75, $p = 0.02$), *MAP3K11* (OR = 7.12, CI = 1.23–3409.1, $p = 0.09$) (Fig. 5A). These eleven genes were regarded as potential risk genes related to PMOP. To further evaluate the relative contribution of risk genes affecting PMOP, the lasso regression algorithm was applied to analyze these risk genes. The eleven remarkable risk genes were chosen as dependent variables to construct a risk predictive model based on the minimum criteria. The regression coefficient (β) represented the hazard contribution to disease and was used to calculate the risk score for each sample. The risk score was calculated using the following formula: risk score = (0.013 × *PTGS2*) + (0.039 × *CXCL16*) + (0.065 × *NECAP1*) + (−0.154 × *RPS23*) + (−0.056 × *SSR3*) + (0.001 × *CD74*) + (0.005 × *IL4R*) + (0.020 × *BTBD2*) + (0.064 × *PIGS*) + (0.002 × *LILRA2*) + (0.001 × *MAP3K11*).

Next, univariate and multivariate logistical regression analyses were used to determine whether the risk score was an independent indicator as well as the role of multiple clinical information in the diagnosis of PMOP. The univariate logistical analysis showed that age (OR = 1.32, CI = 1.14–1.62; $p < 0.01$), PINP (OR = 1.04, CI = 1.01–1.08, $p = 0.009$), β.CTX (OR = 1.01, CI = 1.0–1.01; $p = 0.004$),

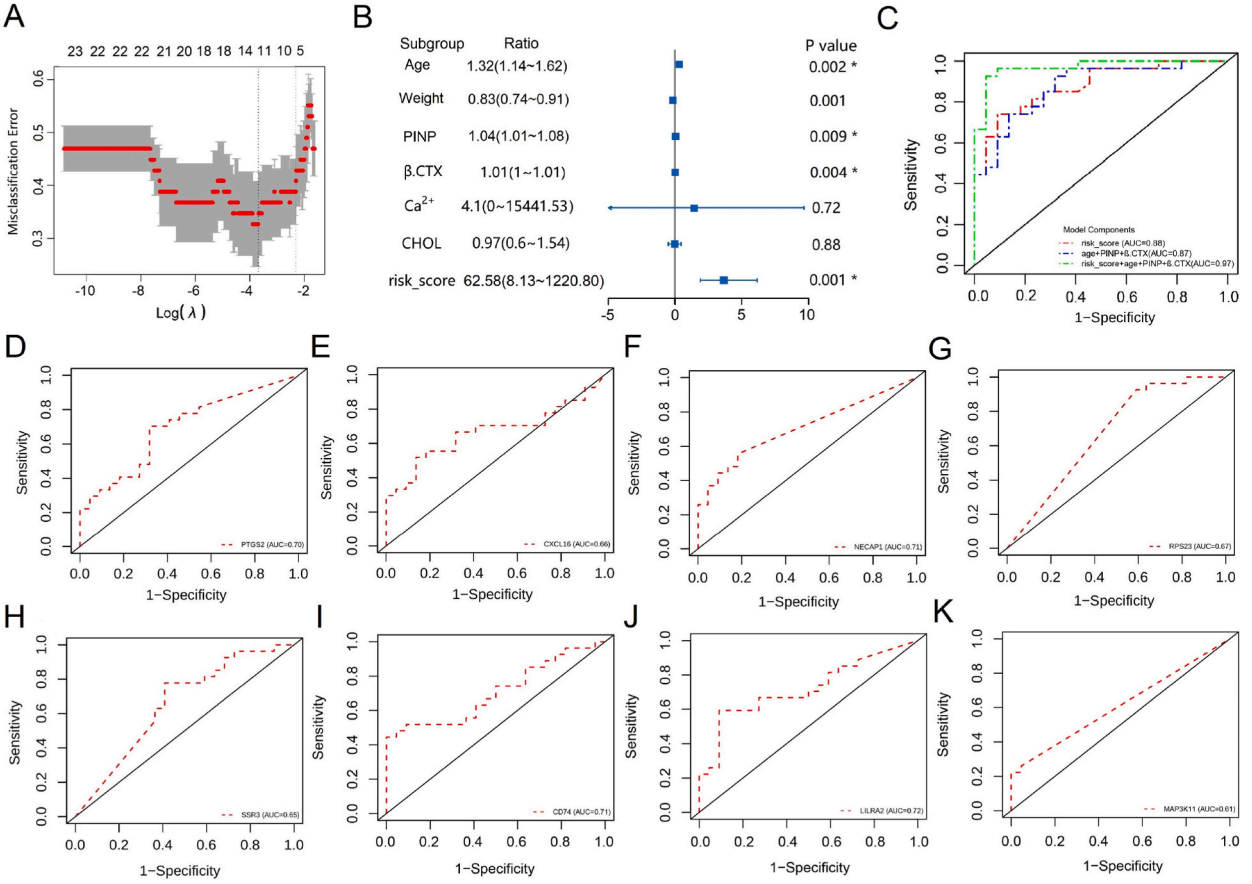



**Fig. 5.** The transcriptome signatures based on deferential genes could predict PMOP. (A) The cross validation for selecting the best genes. (B) Univariate logistical regression analyses of the associations between clinical factors and PMOP. (C) ROC derived from three models incorporating different numbers of predictive variables. The three model components include risk score, clinical information (age, PINP, β.CTX), and a combined risk score (risk score, age, PINP, β.CTX), respectively. (D–K) The independent prediction effect of eleven risk genes.

and risk score (OR = 62.58, CI = 8.13–1220.80, $p = 0.001$) were significantly correlated with PMOP (Fig. 5B). Therefore, multiple predictors were included in the multivariate logistical regression model, which included three types of model components: risk score, clinical information (age + PINP + β.CTX), and a combined risk score (risk score + age + PINP + β.CTX). Then, the prediction effects (sensitivity and specificity) of these models were described by ROC. The AUCs of the independent risk score and the clinical information were 0.88 and 0.87, respectively. However, the AUC of the combined risk model, including the transcriptome signature and clinical information, was 0.97 (Fig. 5C). In addition, we also explored the prediction effects of the representative genes alone, such as *PTGS2* (AUC = 0.70), *CXCL16* (AUC = 0.66), *NECAP1* (AUC = 0.71), *RPS23* (AUC = 0.67), *SSR3* (AUC = 0.65), *CD74* (AUC = 0.71), *LILRA2* (AUC = 0.72) and *MAP3K11* (AUC = 0.61) (Fig. 5D–K). The results of the risk prediction model showed that the transcriptome signatures had a good predictive effect on PMOP and that the predictive power was close to that of classical clinical diagnostic information. Furthermore, the combined risk model performed better, which can further improve prediction accuracy compared with the risk score model component using eleven genes.

## 4. Discussion

Normal bone remodeling maintains a constant bone mass by breaking the dynamic balance between the destruction of bones by osteoclasts and rebuilding by osteoblasts [34]. The disorder of bone metabolism causes bone-related diseases such as PMOP, which is mainly caused by estrogen deficiency. Since PMOP is common, it attracts widespread attention [35,36]. The dysregulation of mRNAs has been described in many bone metabolism diseases [37]. However, many transcriptome studies in osteoporosis have mainly focused on wet experiments or a few sequencing data, which cannot show the panorama of transcriptome changes on a large scale [38]. In this study, the transcriptome profiles of PBMCs in PMOP were determined using bulk RNA-seq data. We studied gene expression changes between the extreme T-score cases and the controls. Moreover, comparative analyses and annotations of gene set levels were completed and further corroborated immune and inflammatory responses in PMOP.

Vitamin D deficiency and reduced calcium absorption are common in elderly individuals, and osteoporosis is an age-related bone disease [39]. Age-related osteoporosis often coexists with multiple other comorbidities (e.g., atherosclerosis, diabetes) that share aging as a major risk factor [40]. In this study, it was also found that the average age of postmenopausal osteoporosis patients was also higher than that of the control group, and we can speculate that aging is also one of the reasons for the disease in women. In addition, chronic complex diseases are controlled by multiple genetic and environmental factors, and several minor effect genes may work together to contribute to common diseases [41]. No significant differentially expressed genes were found based on the differential expression analysis under traditional grouping conditions for all samples. Given that PMOP with an extreme T-score has a higher probability of fracture, we provided unique and comprehensive transcriptome profiles for the extreme T-score group.

Numerous studies have established a correlation between immune-inflammatory responses and multiple complex diseases, including osteoporosis [42]. Th17 cell-produced IL-17 is involved in the pathogenesis of bone loss, and blocking IL-17 action in a mouse model of rheumatoid arthritis reduces disease symptoms [43]. Prior research has also highlighted that inflammation is closely related to bone metabolism and that its environment contributes to the reduction in BMD [44]. For example, Melda et al. clarified that mice lacking RANKL in B lymphocytes were partially protected from the bone loss caused by ovariectomy [45]. Kim et al. demonstrated that TNF-α in T lymphocytes could promote bone loss in conditions such as inflammatory osteolysis and postmenopausal osteoporosis [46]. In this study, we found that some pathways associated with immune responses were also associated with PMOP, such as B-cell receptors, Th17 cell differentiation, and Toll-like receptor signaling. This may reflect the functional changes in humoral immunity, autoimmune function, and nonspecific immunity, which may contribute to PMOP. Furthermore, we show that the B-cell receptor signaling pathway was upregulated by enrichment analysis, which may indicate that the extreme T-score group is accompanied by the enhancement of humoral immunity. Our large-scale bulk RNA-seq data provided evidence that the immune response is related to PMOP, which can also provide abundant data resources for future studies. In the future, researchers can use these data in combination with other omics data to find more biological mechanisms.

Predictive methods may contribute to therapeutic strategies. For many years, the aberrant metabolism of estrogen has been considered to be one of the causes of PMOP. In the current circumstances where PMOP is a chronic progressive disease that easily causes fractures and other risks, early intervention and prevention may be the other approach for PMOP treatment. However, circulating risk biomarkers targeting this disease are still unsuccessful. In this study, we explored the predictive effect of transcriptome signatures and clinical information in PMOP through a Lasso regression model. The eleven genes constructed by our risk prediction model increased predictive sensitivity and specificity in PMOP, which can also provide reference value for clinical auxiliary diagnosis. The findings may provide a low-cost, noninvasive predictive method for PMOP screening in a general population of older women, contributing to disease prevention.

## 5. Limitations

This study had some limitations that should be noted. First, we did not conduct further experiments to investigate how immune-inflammatory responses mediate PMOP. Second, the sample size enrolled in this study was relatively small, and a larger sample size is needed to validate the predictive effect of these transcriptome signatures. Third, our current study is limited to the transcriptional level and has not yet collected other omics data. In the future, multi-omics data analysis will help us further explore the genetic characteristics of PMOP.

## 6. Conclusions

In summary, the results of this study suggest that the immune-inflammatory response pathways are enhanced in PMOP, revealing potential new evidence of the relationship between immunity and osteoporosis. In addition, a panel of eleven transcriptome signatures may be a potential blood biomarker for PMOP, and interestingly, the combination of eleven transcriptome signatures and clinical information in the risk model can further improve the prediction effect of PMOP. This study provides a unique resource for the comprehensive transcriptome expression profile of PBMCs in PMOP, laying a new foundation for understanding the pathogenesis and development of the disease.

## Ethics statement

Each of the studies was approved by the BGI ethical committee and Shenzhen People's Hospital ethical committee (reference number BGI-IRB 17081-T2).

## Funding

This research was supported by the Science, Technology, and Innovation Commission of Shenzhen Municipality (Grant no. JCYJ20160301151248779 and Grant no. JCYJ20200109150410232), Basic and Applied Basic Research Foundation of Guangdong Province (Grant no. 2019B1515120034).

## Informed consent statement

Informed consent was obtained from all subjects and consent was also obtained to publish clinical data of the patients involved in this study.

## Data Availability Statement

Data that support the findings of this study have been deposited into the CNGB Sequence Archive (CNSA) of the China National GeneBank Database (CNGBdb) with accession number CNP0001951.

## CRediT authorship contribution statement

**Pan Gao:** Conceptualization, Data curation, Visualization, Writing – original draft, Formal analysis. **Xiaoguang Pan:** Data curation, Formal analysis, Methodology. **Shang Wang:** Investigation, Resources. **Sijia Guo:** Software. **Zhanying Dong:** Investigation, Resources. **Zhefeng Wang:** Resources. **Xue Liang:** Software. **Yan Chen:** Resources, Software. **Fang Fang:** Resources. **Ling Yang:** Project administration, Resources. **Jinrong Huang:** Project administration, Resources. **Chengxi Zhang:** Resources, Data curation. **Conghui Li:** Methodology. **Yonglun Luo:** Conceptualization, Supervision, Writing – review & editing. **Songlin Peng:** Conceptualization, Resources, Supervision. **Fengping Xu:** Conceptualization, Project administration, Resources, Supervision, Writing – review & editing.

## Declaration of competing interest

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

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e23675.

## Supplementary Data

Supplementary data to this article can be found in supplementary information.

## Abbreviations

| | |
|---|---|
| PMOP | postmenopausal osteoporosis |
| PINP | procollagen I N-terminal propeptide |
| β.CTX | β isomer of C-terminal telopeptide of type I |
| OP | osteoporosis |
| OPG | osteoprotegerin |
| RANKL: | receptor activator of the NF-kB ligand |

| | |
|---|---|
| RANK | receptor activator of the NF-kB |
| PBMC | peripheral blood mononuclear cell |
| DEG | differentially expressed genes |
| BMD | bone mineral density |
| FDR | the false discovery rate |
| TPM | transcripts per kilobase of exon model per million mapped reads |
| OR | odds ratios |
| ROC: | receiver operating characteristic curve |
| AUC | the area under curve |
| BMI | body mass index |
| MAPK | mitogen-activated protein kinase |

## References


[1] S.R. Cummings, L.J. Melton, Epidemiology and outcomes of osteoporotic fractures, Lancet 359 (9319) (2002) 1761–1767.
[2] T.D. Rachner, S. Khosla, L.C. Hofbauer, Osteoporosis: now and the future, Lancet 377 (12) (2011) 1276–1287.
[3] R. Eastell, T.W. O'Neill, L.C. Hofbauer, B. Langdahl, I.R. Reid, et al., Postmenopausal osteoporosis, Nat. Rev. Dis. Prim. 2 (2016), 16069.
[4] S.W. Wade, C. Strader, L.A. Fitzpatrick, M.S. Anthony, C.D. Malley, Estimating prevalence of osteoporosis: examples from industrialized countries, Arch. Osteoporosis 9 (2014) 182.
[5] J.A. Kanis, A. Odén, E.V. McCloskey, H. Johansson, D.A. Wahl, et al., Systematic review of hip fracture incidence and probability of fracture worldwide, Osteoporos. Int. 23 (2012) 2239–2256.
[6] J. Luo, Z. Yang, Y. Ma, Z. Yue, H. Lin, et al., LGR4 is a receptor for RANKL and negatively regulates osteoclast differentiation and bone resorption, Nat. Med. 22 (5) (2016) 539–546.
[7] M. Wu, G. Chen, Y.P. Li, TGF-B and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease, Bone. Res. 4 (2016), 16009.
[8] T. Nakashima, M. Hayashi, T. Fukunaga, K. Kurata, M. Oh-hora, et al., Evidence for osteocyte regulation of bone homeostasis through RANKL expression, Nat. Med. 17 (2011) 1231–1234.
[9] H. Jessop, G. Zaman, K. Lee, R. Suswillo, L.E. Lanyon, Endocrinology: bone adaptation requires oestrogen receptor-alpha, Nature 424 (2003), 389-389.
[10] L.C. Hofbauer, M. Schoppet, Clinical implications of the osteoprotegerin/RANKL/RANK system for bone and vascular diseases, JAMA 292 (4) (2004) 490–495.
[11] R. Zhao, Immune regulation of osteoclast function in postmenopausal osteoporosis: a critical interdisciplinary perspective, Int. J. Med. Sci. 9 (9) (2012) 825–832.
[12] J.A. Clowes, B.L. Riggs, S. Khosla, The role of the immune system in the pathophysiology of osteoporosis, Immunol. Rev. 208 (2005) 207–227.
[13] Q. Wang, Y. Li, Y. Zhang, L. Ma, L. Lin, et al., LncRNA MEG3 inhibited osteogenic differentiation of bone marrow mesenchymal stem cells from postmenopausal osteoporosis by targeting miR-133a-3p, Biomed. Pharmacother. 89 (1) (2017) 1178–1186.
[14] L. Li, X. Wang, X. Liu, R. Guo, R. Zhang, Integrative analysis reveals key mRNAs and lncRNAs in monocytes of osteoporotic patients, Math. Biosci. Eng. 16 (5) (2019) 5947–5970.
[15] J. Di, X. Wu, H. Yu, L. Jiang, Y. Zhang, Systematic analysis of lncRNAs, mRNAs, circRNAs and miRNAs in patients with postmenopausal osteoporosis, Am. J. Transl. Res. 10 (5) (2018) 1498–1510.
[16] O. Govaere, S. Cokell, D. Tiniakos, R. Queen, R. Younes, et al., Transcriptomic profiling across the nonalcoholic fatty liver disease spectrum reveals gene signatures for steatohepatitis and fibrosis, Sci. Transl. Med. 12 (572) (2020), eaba4448.
[17] M.L. Bouxsein, R. Eastell, L.Y. Lui, L.A. Wu, A.E.D. Papp, et al., Change in bone density and reduction in fracture risk: a meta-regression of published trials, J. Bone Miner. Res. 34 (4) (2019) 632–642.
[18] V. Vladyslav Povoroznyuk, V. Nataliia Grygorieva, V. Eugene McCloskey, H. Johansson, A. John, Kanis, Application of FRAX to determine the risk of osteoporotic fractures in the Ukrainian population, International Journal of Osteoporosis and Metabolic Disorders 11 (2018) 7–13.
[19] P. Dargent-Molina, M.N. Douchin, C. Cormier, P.J. Meunierg, Use of clinical risk factors in elderly women with low bone mineral density to identify women at higher risk of hip fracture: the EPIDOS prospective study, Osteoporosis Int. 13 (7) (2002) 593–599.
[20] S. Chen, Y.Q. Zhou, Y. Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor, Bioinformatics 34 (17) (2018) i884–i890.
[21] R. Patro, G. Duggal, M.I. Love, R.A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression, Nat. Methods 14 (4) (2017) 417–419.
[22] G. Yu, L. Wang, Y. Han, Q. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16 (5) (2012) 284–287.
[23] X. Chen, L. Liu, M. Chen, J. Xiang, Y. Wan, et al., Five-gene risk score model for predicting the prognosis of multiple myeloma patients based on gene expression profiles, Front. Genet. 12 (2021), 785330.
[24] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (12) (2014) 550.
[25] J.H. Krege, N. E Lane, J.M. Harris, PINP as a biological response marker during teriparatide treatment for osteoporosis, Osteoporos. Int. 25 (2014) 2159–2171.
[26] Q.Q. Kong, T.W. Sun, Q.Y. Dou, F. Li, Q. Tang, et al., β-CTX and ICTP act as indicators of skeletal metastasis status in male patients with non-small cell lung cancer, Int. J. Biol. Marker. 22 (3) (2007) 214–220.
[27] X. Lin, H. Yu, C. Zhao, Y. Qian, D. Hong, et al., The peripheral blood mononuclear cell count is associated with bone health in elderly men: a cross-sectional population-based study, Medicine 95 (15) (2016) e3357.
[28] B. Jones, Regulating osteoclast differentiation to prevent bone loss, Nat. Rev. Rheumatol. 10 (66) (2014), 206-206.
[29] C.J. DeSelm, Y. Takahata, J. Warren, J.C. Chappel, T. Khan, et al., IL-17 mediates estrogen-deficient osteoporosis in an act1-dependent manner, J. Cell. Biochem. 113 (9) (2012) 2895–2902.
[30] W. Wang, J. Bai, W. Zhang, et al., Protective effects of punicalagin on osteoporosis by inhibiting osteoclastogenesis and inflammation via the NF-κB and MAPK pathways, Front. Pharmacol. 11 (2020).
[31] M.J. Mcgeachy, D.J. Cua, S.L. Gaffen, The IL-17 family of cytokines in health and disease, Immunity 50 (4) (2019) 892–906.
[32] M.B. Greenblatt, J.H. Shim, W. Zou, D. Sitara, M. Schweitzer, et al., The p38 MAPK pathway is essential for skeletogenesis and bone homeostasis in mice, J. Clin. Invest. 120 (7) (2010) 2457–2473.
[33] S. Xu, Y. Zhang, B. Liu, K. Li, B. Huang, et al., Activation of mTORC1 in B lymphocytes promotes osteoclast formation via regulation of β-catenin and RANKL/OPG, J. Bone Miner. Res. 31 (7) (2016) 1320–1333.
[34] M. Zaidi, Skeletal remodeling in health and disease, Nat. Med. 13 (7) (2007) 791–801.
[35] B. Langdahl, S. Ferrari, D.W. Dempster, Bone modeling and remodeling: potential as therapeutic targets for the treatment of osteoporosis, Ther. Adv. Musculoskelet. Dis. 8 (6) (2016) 225–235.
[36] P.N. Mesquita, J.M.C. Maia, F. Bandeira, Postmenopausal osteoporosis, Endocrinology and Diabetes (2014) 305–321.

[37] K. Wang, Y. Wang, Z. Hu, L. Zhang, G. Li, et al., Bone-targeted lncRNA OGRU alleviates unloading-induced bone loss via miR-320-3p/Hoxa10 axis, Cell Death Dis. 11 (2020) 382.
[38] J.N. Farr, M.M. Roforth, K. Fujita, K.M. Nicks, J.M. Cunningham, et al., Effects of age and estrogen on skeletal gene expression in humans as assessed by RNA sequencing, PLoS One 10 (9) (2015), e0138347.
[39] F. Jakob, L. Seefried, M. Schwab, Age and osteoporosis, Internist 55 (2014) 755–761.
[40] S. Khosla, et al., Inhibiting cellular senescence: a new therapeutic paradigm for age-related osteoporosis, The Journal of Clinical Endocrinology & Metabolism 103 (4) (2018) 1282–1290.
[41] J. Flint, M. Munafò, Genesis of a complex disease, Nature 511 (2014) 412–413.
[42] P. Pietschmann, D. Mechtcheriakova, A. Meshcheryakova, Immunology of osteoporosis: a mini-review, Gerontology 62 (2) (2016) 128–137.
[43] A.M. Tyagi, K. Srivastava, M.N. Mansoori, R. Trivedi, N. Chattopadhyay, et al., Estrogen deficiency induces the differentiation of IL-17 secreting Th17 cells: a new candidate in the pathogenesis of osteoporosis, PLoS One 7 (9) (2012), e44552.
[44] A.A. Pérez, E.F. Trepat, M.G. Fresco, A.J. Mora, V. López, et al., Role of toll-like receptor 4 on osteoblast metabolism and function, Front. Physiol. 9 (2018) 504.
[45] M. Onal, J. Xiong, X. Chen, J.D. Thostenson, M. Almeida, et al., Receptor activator of nuclear factor κB ligand (RANKL) protein expression by B lymphocytes contributes to ovariectomy-induced bone loss, J. Biol. Chem. 287 (35) (2012) 29851–29860.
[46] M.T. Gillespie, Impact of cytokines and T lymphocytes upon osteoclast differentiation and function, Arthritis Res. Ther. (2007) 9–103.