International Journal of General Medicine

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Bioinformatic Screening of Key Regulatory Molecules and Mechanisms in Osteoporosis Based on Tryptophan Metabolism-Related Genes

Jie Ji, Yu Qin, Chanling Li, Yingjie Fu, Jing Cui, Qihui Cao

Department of Orthopedics, Beijing Jishuitan Hospital Guizhou Hospital, Guiyang, 550001, People's Republic of China

Correspondence: Qihui Cao, Email 50504561@qq.com

Background and Objective: Osteoporosis (OP) is a metabolic disease characterized by reduced bone mass and increased fracture risk. Tryptophan metabolism may play a crucial role in its pathogenesis, although the underlying mechanisms remain unclear. This study aimed to identify key regulatory genes and elucidate molecular mechanisms through integrated transcriptomic analysis.

Methods: OP-related datasets from Gene Expression Omnibus were analyzed using differential expression analysis, weighted gene co-expression network analysis, and tryptophan metabolism gene sets. Machine learning algorithms were applied to screen key genes and construct diagnostic models. Regulatory networks including protein-protein interaction, competing endogenous RNA, and transcriptional regulation were established. Key findings were validated through independent datasets, Mendelian Randomization, single-cell analysis, and RT-qPCR.

Results: Two key genes were identified: DVL1 (significantly downregulated) and RBM39 (significantly upregulated) in OP patients. DVL1 negatively correlated with resting memory CD4+ T cells and eosinophils, while RBM39 positively correlated with memory B cells and M2 macrophages. DVL1 (AUC = 0.75) and RBM39 (AUC = 0.91) demonstrated consistent expression patterns and excellent diagnostic performance. Functional enrichment revealed significant involvement in apoptosis, autophagy, and signaling pathways. Transcription factor YY1 was identified as a key regulator in the molecular network.

Conclusion: This bioinformatic study reveals the potential role of tryptophan metabolism in OP pathogenesis and identifies DVL1 and RBM39 as candidate diagnostic biomarkers and therapeutic targets through computational analysis. These findings are inferential based on transcriptomic data mining and require further validation through experimental approaches. Future studies should include functional characterization in cellular and animal models, prospective clinical cohort validation, and mechanistic investigations to confirm the diagnostic and therapeutic value of these candidates. If validated, these findings could provide a molecular foundation for developing non-invasive diagnostic tools and precision medicine approaches in OP management, potentially improving early detection and personalized treatment strategies for patients at risk of osteoporotic fractures.

Keywords: osteoporosis, tryptophan metabolism, diagnostic biomarkers, machine learning, DVL1, RBM39, Mendelian randomization

Introduction

Osteoporosis is characterized as a common systemic skeletal metabolic disease in which bone mass is reduced, bone microstructure is deteriorated, and bone fragility is increased, resulting in a significantly elevated fracture risk.[1] With the acceleration of global population aging, osteoporosis has emerged as a major challenge to global public health. It is estimated that approximately 200 million individuals worldwide are affected by osteoporosis, with one osteoporotic fracture occurring every three seconds.[2] In China, the prevalence of osteoporosis among the population over 50 years of age has been reported as 14.4% in males and 27.9% in females, with an annually increasing trend.[3] Osteoporotic fractures not only severely impact patients' quality of life but also impose substantial burdens on healthcare systems. The annual economic burden associated with osteoporosis-related fractures worldwide is estimated to exceed 100 billion US dollars.[4] The pathogenesis of osteoporosis has not been fully elucidated. Traditionally, it has been closely associated with age-

https://doi.org/10.2147/IJGM.S578987
Received: 5 November 2025
Accepted: 12 February 2026
Published: 18 February 2026

related bone remodeling imbalance, hormonal changes (particularly decreased estrogen levels), genetic factors, and lifestyle factors.[5] Bone tissue is subjected to continuous remodeling processes, and the development of osteoporosis is primarily attributed to the disruption of the balance between osteoclast-mediated bone resorption and osteoblast-mediated bone formation.[6] Recent studies have demonstrated complex interactions between the immune system and bone metabolism, termed "osteoimmunology", an emerging field that has revealed the significant roles of immune cells and inflammatory factors in the pathogenesis of osteoporosis.[7,8]

Metabolomic studies have indicated that abnormalities in multiple metabolic pathways play important roles in osteoporosis pathogenesis.[9] Tryptophan, as one of the essential amino acids in humans, has metabolic pathways that are critically involved in immune regulation, inflammatory response, neurotransmitter synthesis, and energy metabolism.[10] Tryptophan metabolism is primarily conducted through the kynurenine pathway and the 5-hydroxytryptamine (5-HT) pathway, with products from both pathways being closely related to bone metabolism.[11] Studies have suggested that indoleamine 2,3-dioxygenase (IDO), a key enzyme in the tryptophan metabolic pathway, exhibits important functions in bone metabolism, with its metabolite kynurenine capable of inhibiting osteoclast differentiation and influencing the progression of osteoporosis.[12,13] Related research has revealed that increased IDO activity can alleviate inflammation-related bone loss by suppressing T cell activation and enhancing regulatory T cell function.[14] Additionally, 5-hydroxytryptamine (5-HT), another metabolite of tryptophan, has been confirmed to affect the balance between bone formation and bone resorption.[15] It has been demonstrated that 5-HT produced in the intestine can reach bone tissue through the circulation, inhibiting osteoblast proliferation and activity, thereby resulting in decreased bone mass.[16]

Despite evidence linking individual tryptophan metabolites to bone metabolism, three critical gaps persist: (1) the complete tryptophan metabolic network in osteoporosis remains uncharacterized through integrated transcriptomic analysis;[17,18] (2) causal relationships between tryptophan metabolism-related gene expression and osteoporosis risk have not been established using genetic instrumental variables; and (3) the multi-layered regulatory mechanisms—including ceRNA networks and transcriptional control—governing these genes in osteoporosis are unknown. With the development of high-throughput sequencing technologies and bioinformatic analysis methods, transcriptomic research has provided important tools for understanding the molecular mechanisms of complex diseases. Weighted gene co-expression network analysis (WGCNA), as a systems biology approach, can be employed to identify disease-related gene modules and key regulatory genes.[19] Machine learning algorithms can assist in selecting biomarkers with high diagnostic value from numerous candidate genes.[20] Recent advances in machine learning have revolutionized biomarker discovery in osteoporosis research.[21,22] Liu et al used LASSO and Random Forest to identify CST3 as an osteoporosis biomarker with animal validation.[23] Furthermore, Random Forest and SVM-RFE were used to identify three senescence-related diagnostic biomarkers (PDPK1, TRIM28, WWP1) for osteoporosis with clinical validation.[24] These studies demonstrate that machine learning not only improves biomarker identification efficiency but also enhances the clinical translatability of discovered genes through rigorous feature selection and cross-validation strategies.

Mendelian Randomization (MR) is a method in which genetic variations are utilized as instrumental variables to infer causal relationships between exposures and outcomes.[25] This approach can effectively avoid the influence of confounding factors and reverse causality in traditional observational studies, providing a reliable pathway for exploring causal relationships between gene expression and disease risk.[26] Recent research using MR methods has discovered a causal relationship between RANKL gene expression and osteoporosis risk, providing strong support for therapeutic strategies targeting RANKL.[27] However, MR has not been applied to examine whether genetically determined expression of tryptophan metabolism-related genes causally influences osteoporosis risk.

The competing endogenous RNA (ceRNA) network represents an important mechanism for non-coding RNA regulation of gene expression, involving interactions among long non-coding RNAs (lncRNAs), microRNAs (miRNAs), and messenger RNAs (mRNAs).[28] Recent studies have shown that ceRNA networks play important roles in bone metabolism regulation.[29] For example, lncRNA DANCR can regulate CTNNB1 expression by competitively binding miR-320a, thereby influencing osteoclast differentiation.[30] However, ceRNA regulatory networks centered on tryptophan metabolism-related genes in osteoporosis have not been constructed or characterized.

Based on these knowledge gaps, this study tests three core hypotheses: (1) tryptophan metabolism-related genes form disease-specific co-expression modules in osteoporotic bone tissue identifiable through WGCNA; (2) genetically determined expression of key tryptophan metabolism genes causally affects osteoporosis risk; and (3) these genes are regulated through multi-layered mechanisms including ceRNA networks and transcription factors. We integrated transcriptomic analysis, WGCNA, machine learning, and Mendelian Randomization to identify key genes, characterize their regulatory networks (protein-protein interaction (PPI), ceRNA, transcription factors), and establish causal relationships with osteoporosis risk. This approach aims to reveal tryptophan metabolism's regulatory architecture in osteoporosis and identify diagnostic biomarkers and therapeutic targets.

Methods

Data Download and Preprocessing

OP-related transcriptome datasets GSE56815 (training; GPL96; PBMC; 40 OP vs 40 controls) and GSE13850 (validation; GPL96; 20 OP vs 20 controls) were obtained from GEO (https://www.ncbi.nlm.nih.gov/geoprofiles/). Preprocessing used the R package inSilicoMerging for background correction, gene symbol harmonization, and normalization. Single-cell RNA-seq dataset GSE172410 (3 young control vs 3 aged OP mice; GPL21103) was also downloaded from GEO. A total of 4688 tryptophan metabolism-related genes were retrieved from GeneCards (Supplementary Table 1). GWAS summary data for OP were sourced from FinnGen R12 (finngen_R12_M13_OSTEOPOROSIS; 956 cases, 345,118 controls), and GWAS for key gene exposures were obtained from OpenGWAS (https://opengwas.io/).

Differential Expression Analysis

Using the training set, DEGs were identified with limma. The thresholds of $|\log_2$ fold change (FC)$| > 1$ and $P < 0.05$ were applied. The $|\log_2 FC| > 1$ threshold was chosen to identify genes with biologically meaningful expression changes (>2-fold difference), while $P < 0.05$ ensures statistical significance in detecting differential expression patterns (Adjust the P-values using the Benjamini-Hochberg method). A volcano plot and heatmaps of the top 15 up- and downregulated genes were generated to visualize expression patterns.

WGCNA

Soft-threshold power β was selected via pickSoftThreshold. An adjacency matrix and TOM were built; dynamic tree cutting (minimum module size = 30 to ensure modules contained sufficient genes for robust functional interpretation; cut height = 0.25 to balance module specificity and gene coverage) identified modules. Module eigengenes were correlated with OP status (Pearson, $P < 0.05$). Hub genes were defined by $|MM| > 0.5$ and $|GS| > 0.5$, thresholds that identify genes strongly connected within modules and significantly associated with the disease phenotype.

Functional Enrichment Analysis

DAVID was used for KEGG and GO (BP/CC/MF) enrichment (Fisher's exact test, $P < 0.05$, FDR < 0.05). GSEA assessed KEGG pathway enrichment differences between OP and controls.

Key Gene Selection, Model Validation, and Enrichment

Random Forest (Gini importance; randomForest), LASSO (glmnet), and XGBoost (xgboost) were applied to candidate genes; intersecting genes across all three were defined as key genes. Model performance was evaluated using 5-fold cross-validation on the training set with mean AUC, accuracy, and F1 score as metrics. Reproducibility was ensured via the R package "caret" (seed=123). Hyperparameters were optimized through grid search: RF (ntree=300, min.node.size=2), LASSO (lambda=0.0032, 10-fold CV), and XGBoost (eta=0.1, max_depth=3, nrounds=100). Model stability was confirmed by permutation testing (1000 iterations; original AUC=0.78±0.03 vs permuted AUC=0.52±0.04, P<0.001) and minimal AUC difference (<0.05) between training and validation sets. Algorithms were implemented using "randomForest", "glmnet", and "xgboost" packages, with the intersection of screened genes defined as key genes. A logistic regression nomogram was constructed based on key genes, validated by calibration curves, and assessed for

clinical benefit via DCA. Expression differences were verified using Wilcoxon rank-sum test, diagnostic accuracy was evaluated by ROC analysis (AUC values), and biological functions were explored through single-gene GSEA.

Immune Infiltration Analysis

To investigate immune microenvironment characteristics in osteoporosis tissues, CIBERSORT (https://cibersort.stanford.edu/) and ssGSEA were employed for immune cell infiltration analysis. Expression data were uploaded to CIBERSORT, and relative abundances of 22 immune cell types in OP and CON groups were estimated using the LM22 matrix and visualized via bar charts and heatmaps. Concurrently, ssGSEA was performed using "GSVA" to calculate enrichment scores based on marker gene sets. Pearson correlation assessed inter-cell relationships, while Wilcoxon test compared OP vs CON differences ($P < 0.05$). Pearson and Spearman analyses evaluated correlations between key genes (DVL1, RBM39) and immune infiltration, revealing regulatory relationships with the immune microenvironment.

PPI, ceRNA Network Construction, and TF Prediction

STRING (minimum interaction score = 0.4) generated a PPI network; Cytoscape enabled visualization and topological analysis, and key sub-networks centered on key genes were extracted. miRNet predicted key gene–miRNA links and upstream lncRNAs to build a ceRNA network, with Cytoscape used to identify central regulators. TRRUST provided transcription factor enrichment (hypergeometric test, $P < 0.05$) to construct TF–gene regulatory networks.

scRNA-Seq Analysis

Although cross-species comparison between human bulk transcriptomics and mouse single-cell data has inherent limitations due to species-specific differences in bone physiology (including continuous skeletal growth in mice, absence of osteonal remodeling in murine cortical bone, and divergent hormonal regulation), we utilized mouse scRNA-seq dataset GSE172410 to explore cellular heterogeneity and key gene expression patterns at single-cell resolution, recognizing that findings require validation in human samples.

Seurat (v5.1.0) performed QC and normalization: cells with 200–6000 detected genes and mitochondrial content <15% were retained; data were LogNormalize-transformed, variable features identified, scaled (ScaleData), and reduced by PCA, PCA retained the top 20 principal components. Clustering parameters were set at k.param=20 and resolution=0.6. Clustering used SNN and the Louvain algorithm; SingleR and literature markers aided cell-type annotation. Differences in cell-type proportions between OP and controls were compared, and key gene expression distributions were profiled at single-cell resolution.

Mendelian Randomization (MR)

Analysis followed three MR assumptions: ① IVs (SNPs) associate with exposures (DVL1, RBM39); ② IVs are independent of confounders; ③ IVs affect outcomes through exposures only. eQTL and osteoporosis data were processed using "gwasvcf", selecting SNPs at P<0.05 as IVs. SNPs with F>10 were retained based on R^2 and F-statistics. Data were harmonized via "gwasglue" and "TwoSampleMR". Five methods (IVW, MR-Egger, weighted median, weighted mode, simple mode) were applied, prioritizing IVW. Cochran's Q test determined model selection. MR-Egger intercept and MR-PRESSO assessed pleiotropy. ORs with 95% CIs were visualized via forest plots, scatter plots, and funnel plots. Leave-one-out analysis tested robustness. Steiger testing verified causal direction.

RT-qPCR

RT-qPCR validation using human samples was approved by the Ethics Committee of Beijing Jishuitan Hospital Guizhou Hospital (Approval Number: KT2024060504, Date: [2024–06-07]), and written informed consent was obtained from all participants. Total RNA was extracted from blood (MolPure® Blood RNA Kit), quantified by NanoDrop, and reverse-transcribed (PrimeScript™ RT; 37°C 15 min, 82°C 5s). qPCR (TB Green®) was run on a Gentier 96 system: 95°C 30s; 40 cycles of 95°C 5s and 60°C 30s; followed by melt-curve analysis. Reactions (20 μL) contained 10 μL TB Green mix, 0.4 μL each primer (0.2 μM), and 1 μL cDNA. All assays were in triplicate, and relative expression was calculated by the $2^{-\Delta\Delta Ct}$ method (primer sequences in Supplementary Table 2).

Results

Differential Expression Analysis of Transcriptomic Data

Differential expression analysis was performed on the training set data according to screening, and a total of 2535 genes were significantly differentially expressed in OP group, including 1715 upregulated genes and 820 downregulated genes (Figure 1A). Heatmaps displayed the expression patterns of the top 15 upregulated and downregulated genes (Figure 1B).

WGCNA Analysis

To identify key modules and hub genes of OP, WGCNA analysis was performed on the training dataset. First, the optimal soft threshold β was calculated as 12 through pickSoftThreshold function analysis (Figure 2A and B). Based on the optimal soft threshold, co-expression networks were constructed and gene clustering and module identification were performed, dividing genes with similar expression patterns into 10 modules (Figure 2C). Subsequently, correlations between module eigenvectors (module eigengene, ME) were calculated. Results showed that correlation distances between modules were all greater than 0.25, indicating good independence between modules (Figure 2D). Correlations between various modules and OP clinical phenotypes were further analyzed (Figure 2E). Results showed that 2 modules were significantly associated with OP development and progression ($P < 0.05$), namely yellow and tan modules. Finally, 457 hub genes were extracted from these 2 significantly correlated modules (|MM| > 0.5 and |GS| > 0.5), and intersected with previously identified DEGs to obtain 48 candidate genes (Figure 2F). These genes may play important regulatory roles in OP development and progression.

Functional Enrichment Analysis of Candidate Genes

GO and KEGG functional enrichment analysis was performed on the above 48 screened candidate genes, with the top 10 pathways visualized by significance ranking. In GO enrichment analysis, these genes were mainly enriched in acute inflammatory response, regulation of inflammatory response, killing of cells of other organisms, regulation of neuron

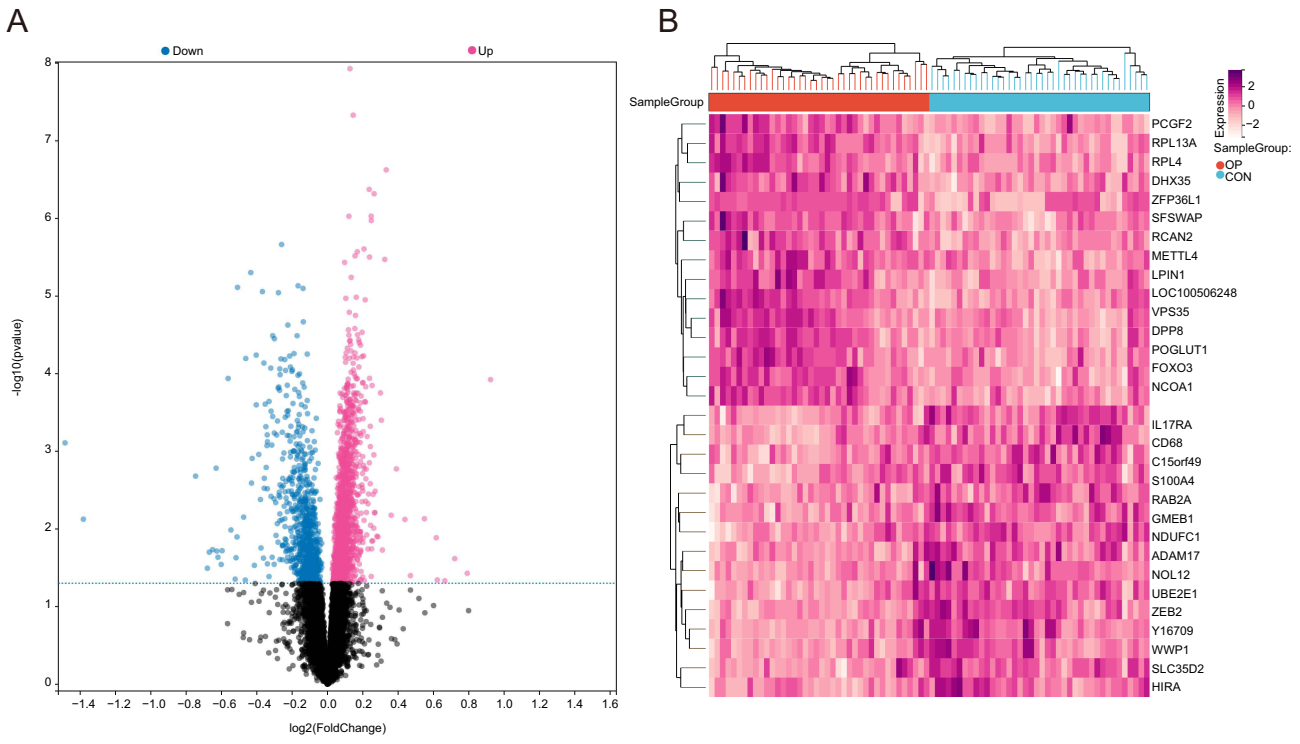

Figure 1 Results of differential expression analysis of transcriptomic data. (**A**) Volcano plot of differentially expressed genes in the training dataset. Rose-colored dots represent upregulated genes, blue dots represent downregulated genes, and black dots represent genes without significant differences. (**B**) Heatmap of differentially expressed genes in the training dataset, showing gene expression differences between disease group and normal group samples.

https://doi.org/10.2147/IJGM.S578987

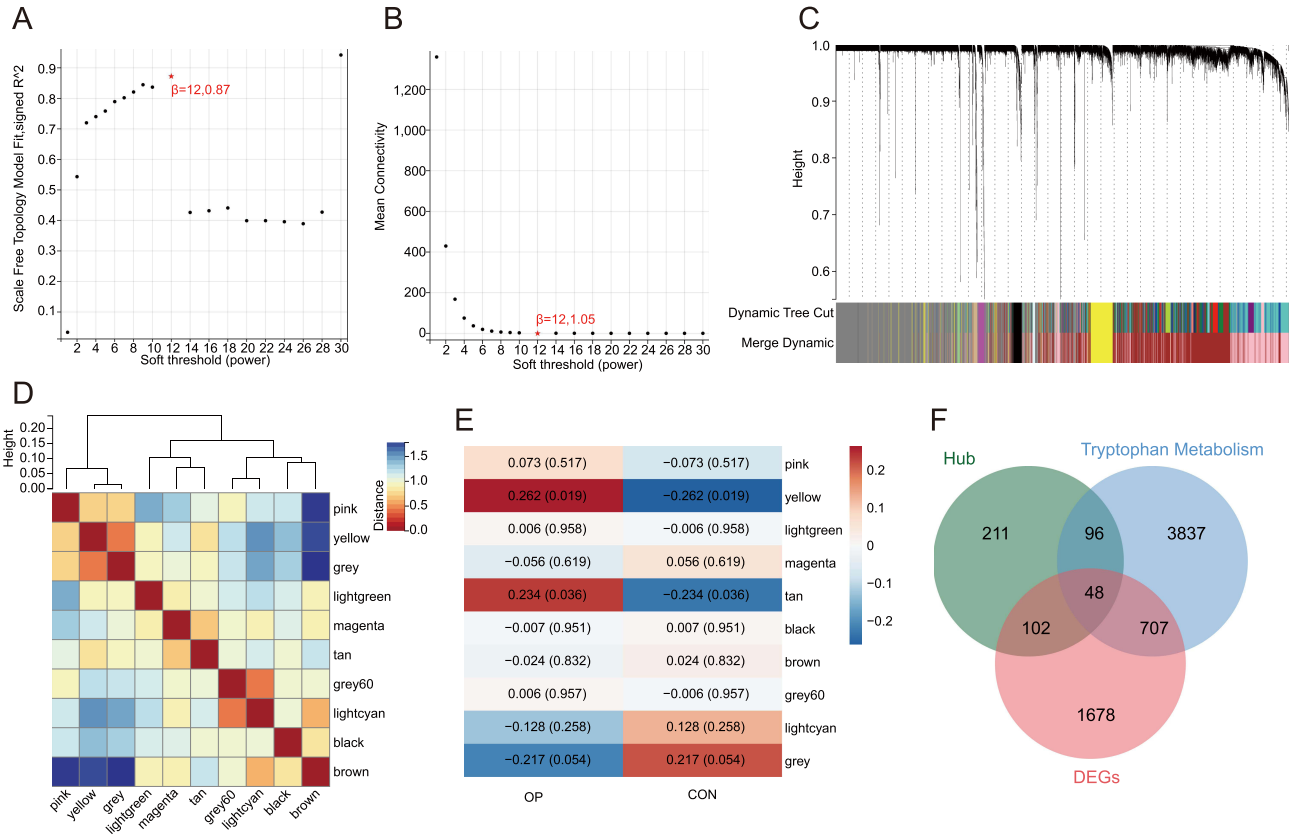

Figure 2 WGCNA analysis results. (**A**) Relationship plot between soft threshold power β and mean connectivity, used for screening the optimal soft threshold. (**B**) Relationship plot between soft threshold power β and scale-free network fit index (R^2). (**C**) Gene module clustering dendrogram based on Dynamic Tree Cut algorithm. (**D**) Correlation heatmap between module eigengenes (ME), with color intensity representing the strength of correlations between modules. (**E**) Correlation heatmap between different modules and OP clinical phenotypes (OP group vs CON group), with colors representing correlation magnitude and P values shown in parentheses. (**F**) Venn diagram of DEGs, Hub genes, and Tryptophan Metabolism-related genes.

death, neuron death, acute-phase response, regulation of apoptotic process, positive regulation of apoptotic process, apoptotic process, and cellular response to lipopolysaccharide (Figure 3A–C). In KEGG enrichment analysis, candidate genes were significantly enriched in Apoptosis, Autophagy-animal, Cholinergic synapse, EGFR tyrosine kinase inhibitor resistance, Neuroactive ligand-receptor interaction, cAMP signaling pathway, Calcium signaling pathway, TNF signaling pathway, NF-kappa B signaling pathway, and PI3K-Akt signaling pathway (Figure 3D and E). These results indicate that tryptophan metabolism-related candidate genes may be involved in OP pathogenesis through multiple biological processes and signaling pathways. GSEA analysis results showed that multiple KEGG pathways were significantly differentially enriched between OP and CON groups, with enriched pathways including Apoptosis, Cell cycle, DNA replication, Mismatch repair, and Base excision repair (Figure 3F).

Key Gene Screening, Model Validation and Functional Enrichment Analysis

To identify the most important diagnostic genes, three machine learning algorithms (Random Forest, LASSO, and XGBoost) were applied to the 48 candidate genes. Random Forest algorithm identified 9 genes with importance scores greater than 1.0 (Figure 4A). LASSO regression analysis identified 8 characteristic genes at the optimal lambda value (Figure 4B). XGBoost algorithm ranked gene importance and identified 10 key genes (Figure 4C). Taking the intersection of results from the three algorithms, 2 key genes were finally identified: DVL1 and RBM39 (Figure 4D). Based on these key genes, a nomogram risk prediction model was constructed (Figure 4E). Calibration curves demonstrated good accuracy of the nomogram prediction model (Figure 4F). Decision curve analysis (DCA) showed that the nomogram model had good clinical application value (Figure 4G). Through further analysis of key gene expression patterns in both OP and control group samples, we observed that DVL1 was significantly downregulated in the OP group while RBM39

https://doi.org/10.2147/IJGM.S578987

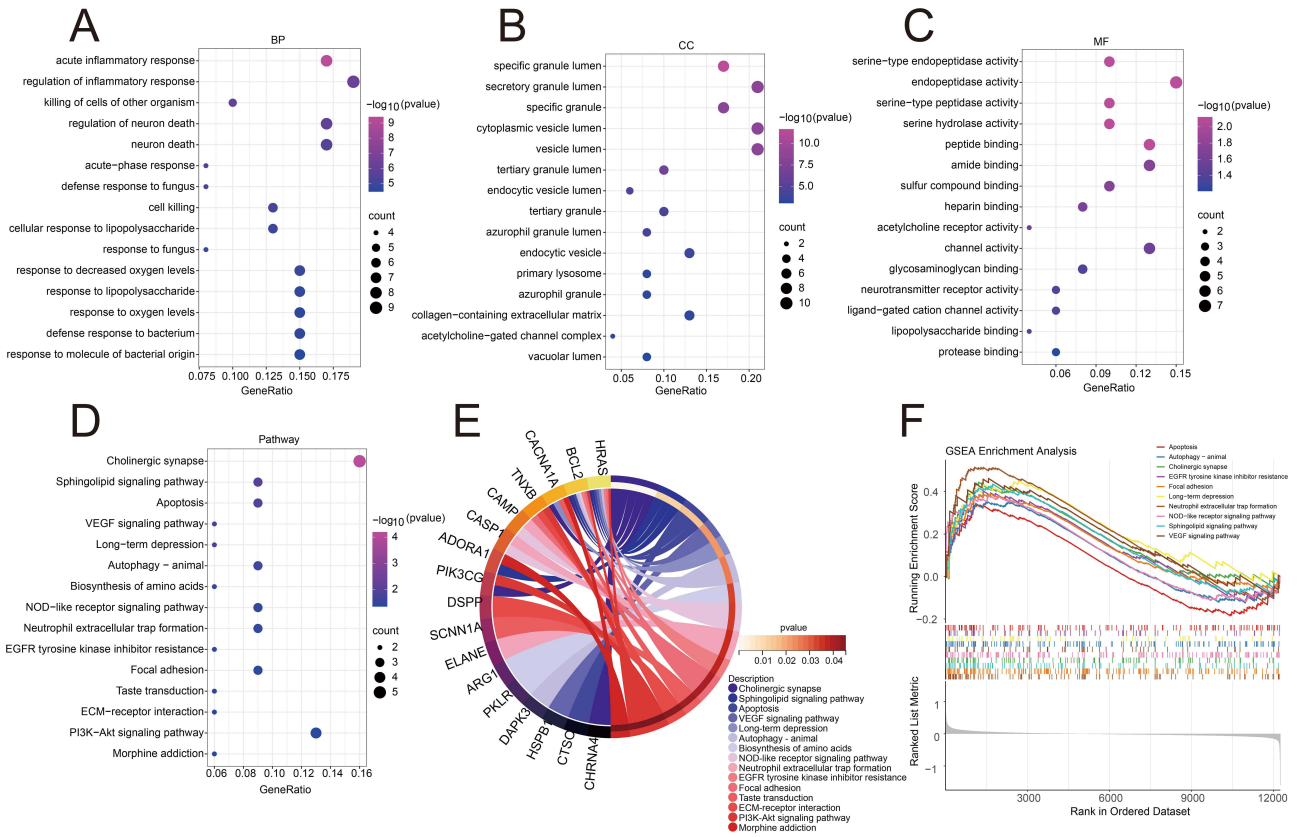

Figure 3 GO and KEGG functional enrichment analysis of candidate genes. (**A** and **B**) GO terms enriched by candidate genes for BP (**A**) and CC (**B**). (**C**) Bubble plot of GO-MF and KEGG pathway enrichment, with bubble size representing gene count, color indicating adjusted P-value, and the x-axis showing gene ratio. (**D**) Bar plot of significant KEGG pathways. (**E**) Chord diagram illustrating the relationships between candidate genes and their associated enriched functional pathways. (**F**) GSEA running enrichment score curves, with the x-axis representing the ranked gene list and the y-axis representing the running enrichment score (ES); vertical lines indicate the positions of gene set members in the ranked list.

was significantly upregulated, and ROC curve analysis demonstrated that both genes had good diagnostic performance in training and validation sets (Figure 4H and I). Single-gene GSEA functional enrichment analysis revealed the molecular mechanisms of the two key genes. DVL1 mainly upregulated hormone signaling pathways (such as estrogen signaling, parathyroid hormone synthesis, etc) and downregulated immune response-related pathways (such as autoimmune thyroid diseases, antigen processing and presentation, etc) (Figure 4J); while RBM39 mainly upregulated RNA metabolism and protein synthesis pathways (such as spliceosome, ribosome biogenesis in eukaryotes, etc) and downregulated tumor and inflammation-related pathways (Figure 4K). These results indicate that the two key genes participate in the pathogenesis of osteoporosis through different molecular mechanisms, providing important molecular basis for understanding its pathological mechanisms and developing therapeutic strategies.

Immune Infiltration Analysis

To elucidate immune microenvironment characteristics in osteoporosis tissues, CIBERSORT and ssGSEA analyzed immune cell infiltration in the training set. CIBERSORT identified 22 immune cell types in OP and CON samples, visualized via bar charts and heatmaps (Figure 5A). Correlation analysis showed strongest positive correlation between Mast cells resting and Eosinophils ($r = 0.502$, $P < 0.05$), and strongest negative correlation between T cells CD4 memory resting and T cells CD8 ($r = -0.711$, $P < 0.05$) (Figure 5B). T cells CD4 memory activated were significantly upregulated in OP ($P < 0.05$), while Monocytes, Eosinophils, and NK cells resting showed no significant differences (Figure 5C). DVL1 negatively correlated with T cells CD4 memory resting and Eosinophils; RBM39 positively correlated with B cells memory and Macrophages M2, negatively with T cells CD8 and Macrophages M0 (Figure 5D). Notably, the positive correlation between RBM39 and macrophages M2 warrants careful interpretation, as it may reflect maladaptive compensatory responses, context-dependent M2 dysfunction, or pathological

https://doi.org/10.2147/IJGM.S578987

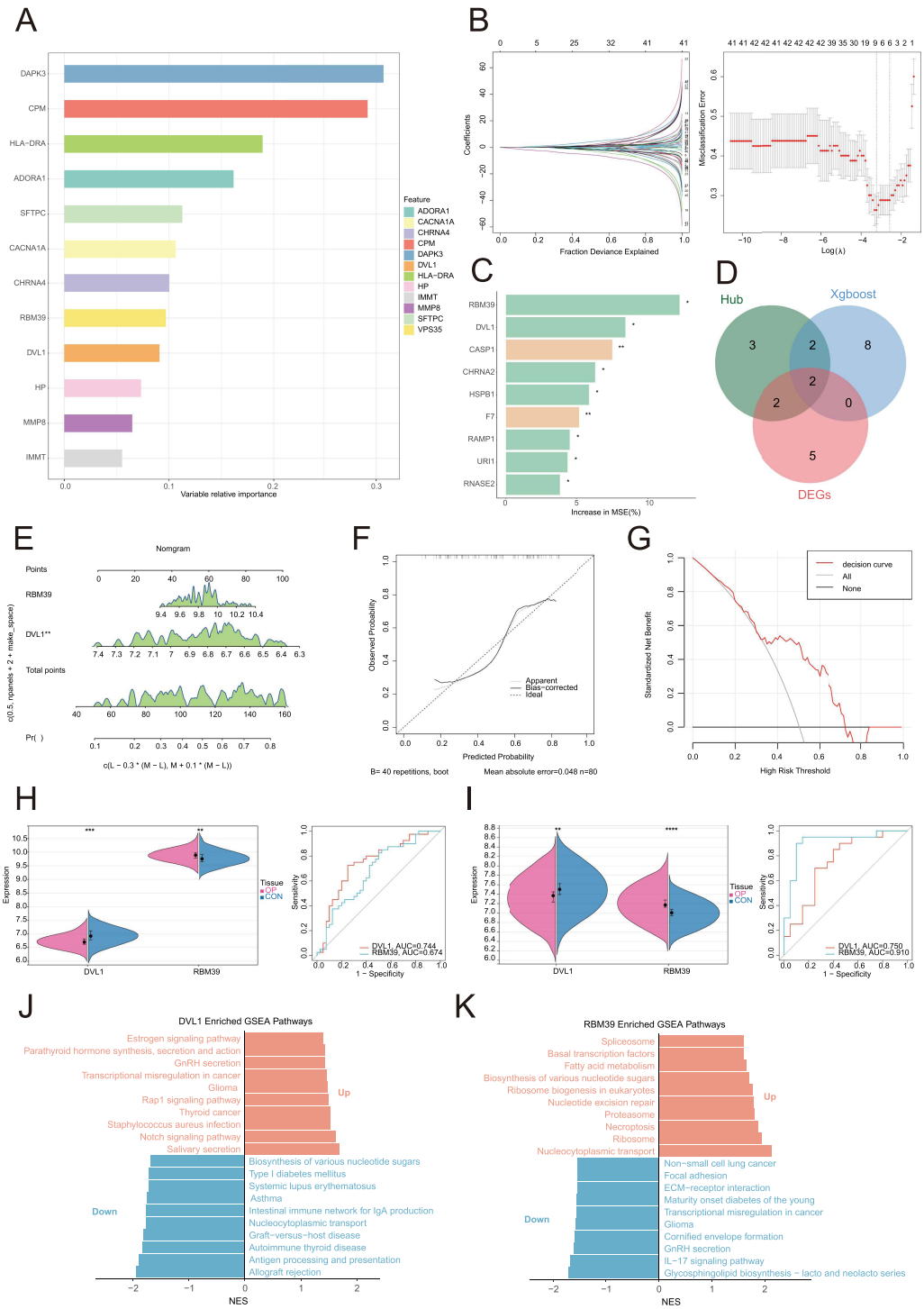

Figure 4 Key gene screening through machine learning algorithms. (**A**) Importance ranking plot of candidate genes by XGBoost algorithm. (**B**) LASSO regression model path plot showing the trend of gene coefficients with penalty parameter λ changes. (**C**) Importance ranking plot of candidate genes by Random Forest algorithm, with genes sorted by mean squared error increase percentage (MSE Increase %). (**D**) Venn diagram showing the intersection of feature genes screened by three machine learning algorithms (XGBoost, LASSO, and RF). (**E**) Nomogram risk prediction model based on key genes (DVL1 and RBM39). (**F**) Calibration curve for evaluating the accuracy of the nomogram prediction model. (**G**) Decision curve analysis (DCA) evaluating the clinical application value of the nomogram model. (**H**) Wilcoxon rank-sum test analysis of key genes DVL1 and RBM39 expression patterns in the training set; receiver operating characteristic (ROC) curve analysis of key genes DVL1 and RBM39 in the training set. (**I**) Wilcoxon rank-sum test analysis of key genes DVL1 and RBM39 expression patterns in the validation set; receiver operating characteristic (ROC) curve analysis of key genes DVL1 and RBM39 in the validation set. (**J** and **K**) GSEA enrichment analysis of DVL1 and RBM39 genes.

https://doi.org/10.2147/IJGM.S578987

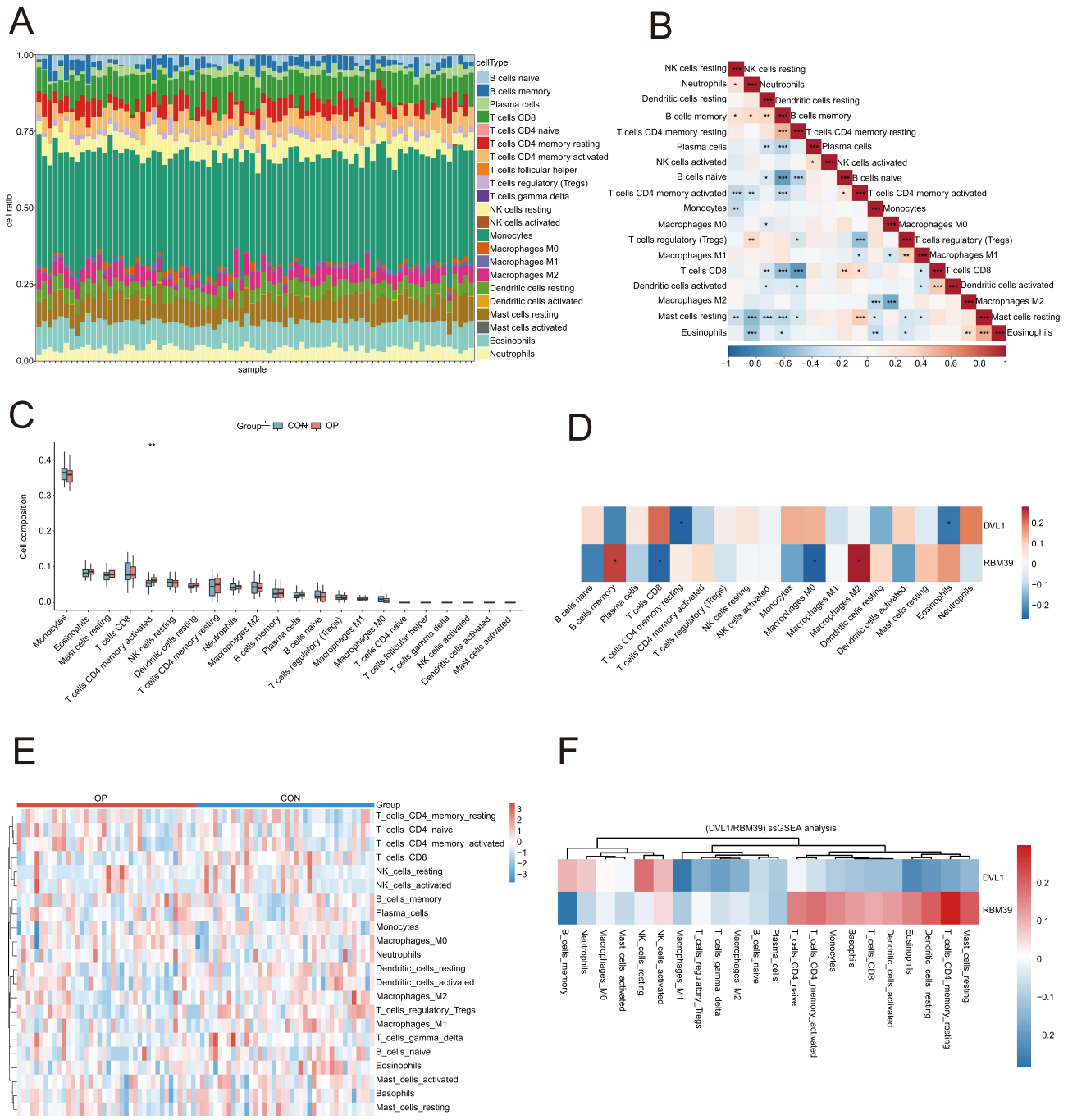

Figure 5 Immune infiltration analysis. (**A**) Bar chart of immune cell subpopulation proportions. (**B**) Correlation heatmap between immune cell subpopulations. Red indicates positive correlation, blue indicates negative correlation. (**C**) Cell differences between OP group and CON group. (**D**) Correlations between key genes DVL1 and RBM39 and immune cells. Red indicates positive correlation, blue indicates negative correlation, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. (**E**) Heatmap of immune cell enrichment scores analyzed by ssGSEA algorithm in OP and CON group samples; (**F**) Spearman correlation heatmap and clustering analysis between key genes DVL1 and RBM39 expression levels and immune cell enrichment scores.

polarization rather than protective mechanisms. ssGSEA results were highly consistent with CIBERSORT (Figure 5E and F), validating T cells CD4 memory activated upregulation in OP ($P < 0.05$) and DVL1/RBM39 correlation patterns.

PPI and ceRNA Network Construction and TF Prediction

To explore regulatory mechanisms of key genes in OP development and progression, analysis was performed from three aspects: protein-protein interactions, ceRNA regulation, and transcription factor regulation. The PPI network contained

https://doi.org/10.2147/IJGM.S578987

202 interaction relationships between 43 genes (Figure 6A). Additionally, PPI networks centered on key genes were constructed to explore regulatory effects of key genes on other candidate genes. Results showed that key genes had direct interaction relationships with HSPB, BOP1, HRAS, DSPP, and BCL2 genes (Figure 6B).

To explore molecular regulatory mechanisms of key genes, key gene-related miRNAs were predicted using the miRNet database, and upstream lncRNAs were predicted based on miRNAs. The two prediction results were then merged and visualized. The ceRNA network visualization showed a total of 2 key mRNAs, 86 miRNAs, 21 lncRNAs and

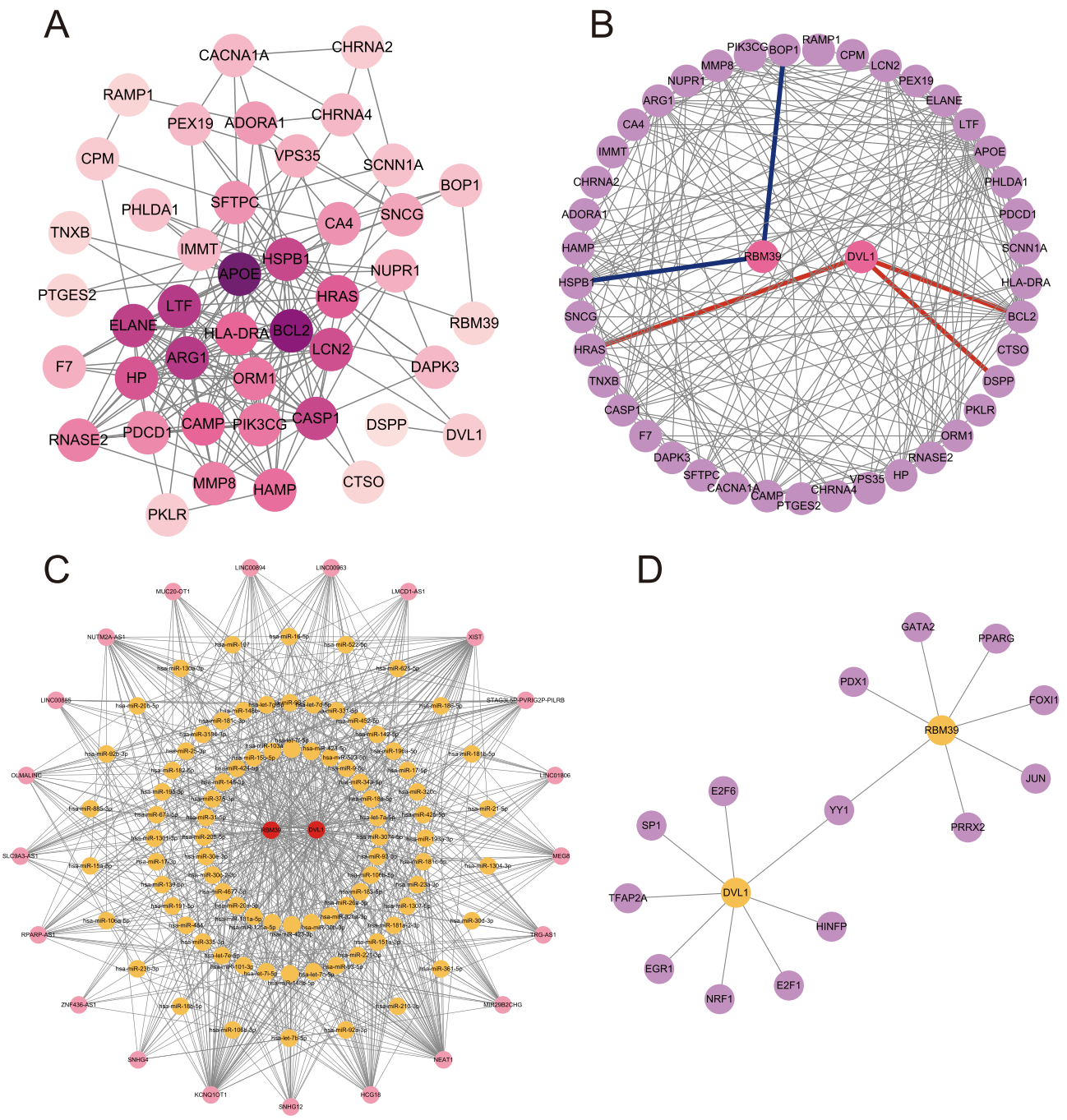

Figure 6 Network regulatory mechanisms of key genes. (**A**) PPI network of candidate genes. Darker colors indicate higher Degree values of genes in the PPI network. (**B**) PPI network related to key genes. Rose-colored nodes represent key genes, purple nodes represent genes interacting with key genes. (**C**) ceRNA network construction. Pink nodes represent lncRNAs, yellow nodes represent miRNAs, red nodes represent mRNAs. (**D**) Transcription factor prediction. Purple nodes represent transcription factors, yellow nodes represent key genes.

705 interaction relationships. Among them, with NEAT1, XIST, and KCNQ1OT1 were lncRNAs with the highest Degree values in the ceRNA network; hsa-let-7b-5p, hsa-let-7e-5p, hsa-let-7f-5p, and hsa-let-7g-5p were miRNAs with the highest Degree values in the ceRNA network; RBM39 and DVL1 were mRNAs with the highest Degree values in the ceRNA network, with mutual regulatory relationships between them (Figure 6C), suggesting that these lncRNAs and miRNAs in ceRNA network mechanisms may influence OP disease processes by regulating key gene expression.

Next, transcription factor enrichment analysis was performed on the 2 key genes. Results showed that a total of 14 potential transcription factor binding sites were significantly enriched in key gene promoter regions. PDX1, GATA2, PPARG, FOXI1, JUN, PRRX2, and YY1 were transcription factors for RBM39; YY1, HINFP, E2F1, NRF1, EGR1, TFAP2A, SP1, and E2F6 were transcription factors for DVL1, with YY1 being a common transcription factor for both key genes DVL1 and RBM39, suggesting that these transcription factors may play important roles in regulating key gene transcriptional expression (Figure 6D).

scRNA-Seq Single-Cell Sequencing Analysis Reveals Cellular Heterogeneity and Key Gene Expression Patterns

Based on single-cell sequencing data, comprehensive single-cell analysis was performed on OP and CON group samples. Quality control results showed screening of high-quality cells based on gene detection numbers, expression abundances, and mitochondrial genes (Figure 7A). Dimensionality reduction visualization results indicated that different samples presented clear distribution patterns in t-SNE and UMAP spaces, with samples showing both similarities and maintaining individual characteristics, providing good resolution for cell subpopulation identification (Figure 7B). Clustering analysis successfully identified 11 cell clusters with different transcriptomic characteristics (Figure 7C). Based on marker gene expression profile analysis of each cell cluster, molecular characteristics of each cluster were systematically analyzed, revealing significantly different gene expression patterns between clusters (Figure 7D). Combined with known cell type-specific marker genes, 9 major cell types were successfully identified and annotated, including Myonuclei, Endothelial cells, Macrophages, Neuron cells, Pericytes, Satellite cells, Tenocytes, Fibro/adipogenic progenitors (FAPs), and Neutrophils, which have different biological functions in tissues (Figure 7E). Marker gene expression analysis further validated the accuracy of cell type identification (Figure 7F). Intergroup comparison analysis revealed significant differences in cellular composition between OP and CON groups, finding obvious changes in proportions of multiple cell types (FAPs, Tenocytes, Satellite cells) between groups, and such changes in cellular composition may be closely related to relevant pathophysiological processes (Figure 7G). Intercellular communication network analysis showed that endothelial cells exhibited extremely active communication characteristics with many other types of cells (Figure 7H). Furthermore, expression distributions of key genes DVL1 and RBM39 were analyzed at the single-cell level. Results showed that DVL1 gene was significantly downregulated in the OP group, while RBM39 gene was significantly upregulated in the OP group, suggesting that these two genes participate in OP pathological processes through different expression patterns, completely consistent with expression trends in transcriptomic data (Figure 7I).

MR Validation

DVL1 and RBM39 eQTL data were used as exposure factors, with OP as the outcome; five algorithms were combined for MR analysis (Inverse variance weighted test, MR-Egger, weighted median, Simple mode, simple median). Results were mainly referenced from IVW, with reference to subsequent heterogeneity analysis results of Cochran Q assessment; random-effects IVW was used when heterogeneity existed, and fixed-effects IVW was used when no heterogeneity existed.

F statistics of SNPs used for exposure factors were all greater than 10, indicating effective avoidance of bias from weak instrumental variables; when OP served as outcomes, 28 SNPs were used as instrumental variables for DVL1 and RBM39 factors in MR analysis. MR results for DVL1 and RBM39 factors are shown in Supplementary Figure 1, where based on the IVW method, DVL1 gene expression level was negatively correlated with OP risk (OR=0.94, 95% CI: 0.80–1.10, $P = 0.44$), while RBM39 was positively correlated with OP risk (OR=1.22, 95% CI: 1.10–1.36, $P < 0.05$). Furthermore, correlations between exposure factors and outcomes were assessed; when intercepts existed, confounding

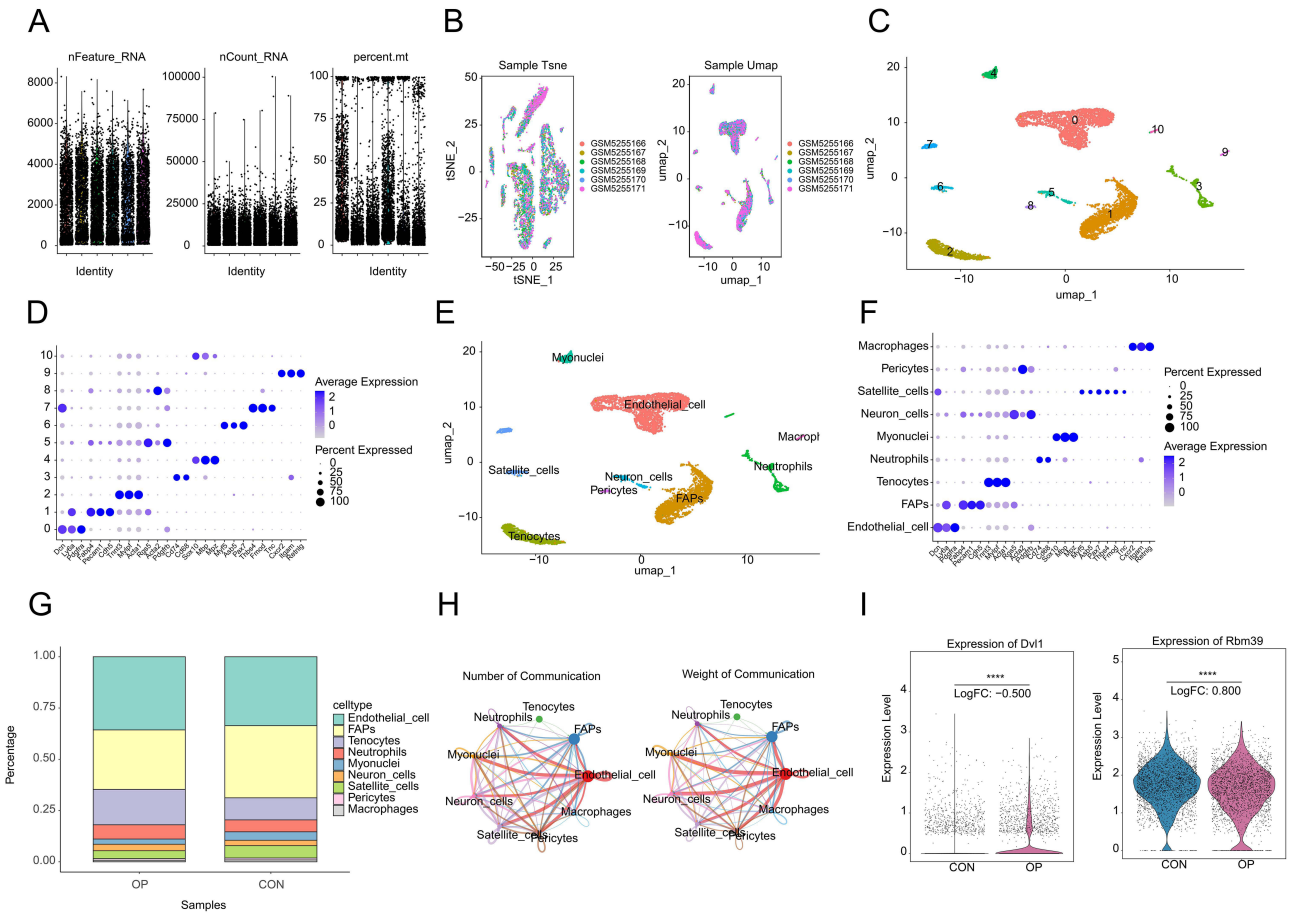

Figure 7 Single-cell RNA sequencing analysis reveals tissue cellular heterogeneity and intercellular communication networks. (**A**) Quality control indicator distribution plots showing the distribution characteristics of nFeature_RNA, nCount_RNA, and mitochondrial gene expression proportion (percent.mt) in filtered cells. (**B**) Sample distribution visualization: left panel shows t-SNE analysis displaying spatial distribution of different samples, right panel shows UMAP analysis demonstrating overall sample distribution patterns. (**C**) Cell clustering results based on UMAP algorithm, dividing cells into 11 different cell clusters (0–10). (**D**) Heatmap of marker gene expression for each cell cluster, showing molecular characteristics of different clusters. (**E**) Cell type annotation results showing spatial distribution of 9 major cell types. (**F**) Cell type-specific marker gene expression dot plot, with bubble size representing expression proportion and color intensity representing average expression level. (**G**) Stacked bar chart comparing cell type proportions between OP group and CON group. (**H** and **I**) Intercellular communication analysis network plots: (**H**) shows intercellular communication numbers, (**I**) Shows communication intensity, with line thickness and color intensity representing communication strength. (**J**) Dvl1 and Rbm39 gene expression distribution plot in single cells, ***P < 0.0001.

factors were indicated. As shown in Supplementary Figure 1A, with continued focus on the Inverse variance weighted method, factor intercepts were small, indicating minimal influence of confounding factors without affecting result reliability; DVL1 line slope was negative, indicating a protective factor; RBM39 line slope was positive, indicating a risk factor. For randomness assessment, if samples were distributed roughly symmetrically along both sides of the IVW line, MR conformed to Mendel's second law of random grouping (Supplementary Figure 1B). Furthermore, for assessing diagnostic efficacy of various SNP sites in predicting exposure factors for outcomes, solid points on the left indicated reduction, while solid points on the right indicated increase. DVL1 was on the left side of the reference line, indicating a protective factor; RBM39 was on the right side of the reference line, indicating a risk factor (Supplementary Figure 1C).

To assess analysis result reliability, three analyses were performed including heterogeneity tests, horizontal pleiotropy tests, and leave-one-out analysis. First, heterogeneity tests were performed; if P values analyzed by heterogeneity tests were less than 0.05, heterogeneity existed; as shown in Supplementary Table 3, DVL1 and RBM39 showed no heterogeneity ($P > 0.05$). Second, horizontal pleiotropy tests were performed: if exposure factor P values were greater than 0.05, no horizontal pleiotropy existed; otherwise, horizontal existed, as shown in Supplementary Table 4. The results showed that DVL1 exhibited evidence of horizontal pleiotropy (P = 0.03 < 0.05), while RBM39 showed no horizontal

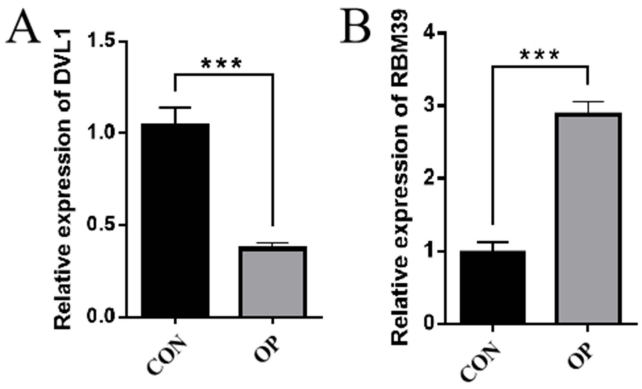

Figure 8 RT-qPCR detection of key gene expression levels. (**A** and **B**) Relative mRNA expression levels of DVL1 (**A**) and RBM39 (**B**) in OP and CON groups, as determined by RT-qPCR. Data are presented as mean ± SD (n = 5 per group). Individual data points represent biological replicates. Statistical significance was determined by unpaired Student's *t*-test. ****P < 0.0001.

pleiotropy (P = 0.25 > 0.05). Additionally, MR-PRESSO was also used for horizontal pleiotropy testing. As shown in Supplementary Table 5, the Global test for DVL1 showed P = 0.45 > 0.05, indicating no outlier SNPs detected by MR-PRESSO; in contrast, RBM39's Global test showed P < 0.05, suggesting the presence of potential outlier SNPs. However, comprehensive sensitivity analyses demonstrated the robustness of the causal relationship for RBM39: (1) leave-one-out analysis showed that no individual SNP had a substantial impact on the causal estimate (Supplementary Figure 1D); (2) MR-Egger intercept test showed no evidence of systematic horizontal pleiotropy (P = 0.25 > 0.05, Supplementary Table 4); and (3) heterogeneity test indicated consistent causal estimates across instrumental variables (P > 0.05, Supplementary Table 3). These findings collectively support the reliability of the causal relationship between RBM39 and osteoporosis. In contrast, DVL1 showed evidence of systematic pleiotropy in MR-Egger intercept test (P = 0.03 < 0.05, Supplementary Table 4), though MR-PRESSO analysis showed no outliers (P = 0.45 > 0.05, Supplementary Table 5). The discrepancy between MR-Egger (detecting pleiotropy) and MR-PRESSO (not detecting outliers) for DVL1 suggests that while horizontal pleiotropy exists, it may not be driven by outlier SNPs but rather reflects DVL1's biological complexity as a hub gene in Wnt signaling. This warrants cautious interpretation of its causal effect and highlights the importance of triangulating evidence across multiple sensitivity tests. Finally, Leave-one-out analysis was performed, showing reliable results with no individual SNP having significant impact on outcomes (Supplementary Figure 1D). The potential for reverse causality in the research process was considered. Reverse Mendelian randomization analysis yielded non-significant results ($P > 0.05$) (Supplementary Table 6). Additionally, the Steiger directionality test was applied to validate the correctness of the causal direction between exposure and outcome, confirming correct directionality ($P < 0.05$) (Supplementary Table 7). The reverse causal effect of osteoporosis on gene expression was thus excluded, and a unidirectional causal relationship was established with no evidence of bidirectionality.

RT-qPCR

RT-qPCR analysis revealed significant differential mRNA expression between CON and OP groups. DVL1 mRNA levels were significantly downregulated in OP group compared to CON group ($P < 0.001$), while RBM39 mRNA levels were significantly upregulated ($P < 0.001$) (Figure 8A and B). These findings confirmed that the identified key genes exhibit concordant changes at transcriptional levels in OP pathogenesis.

Discussion

This study systematically explored the key molecular mechanisms of tryptophan metabolism in the development and progression of osteoporosis through integrating transcriptomics, MR, and machine learning algorithms. The results demonstrated that DVL1 and RBM39, as core regulatory genes related to tryptophan metabolism, participate in the pathological processes of osteoporosis through complex molecular networks, providing new insights into the in-depth understanding of this disease's pathogenesis.

Dishevelled-1 (DVL1), as a key component of the Wnt signaling pathway, controls cell proliferation, survival, and differentiation during development.[31] The Wnt/β-catenin signaling pathway is widely recognized as one of the core pathways regulating bone formation and bone metabolism.[32,33] Subsequent extensive studies have further confirmed the crucial role of this pathway in adult bone remodeling and its potential as a therapeutic target for osteoporosis.[34,35] Additionally, activation of the Wnt signaling pathway during bone metastasis can promote cell proliferation, induce epithelial-mesenchymal transition, regulate extracellular matrix remodeling, stimulate angiogenesis and immune escape mechanisms, thereby enabling tumor cells to successfully colonize and continue growing in the bone tissue environment.[36] DVL1, as a core protein for intracellular signal transduction, exhibited a significantly downregulated expression pattern in this study, which has important biological significance. The decrease in its expression level may directly affect the transmission efficiency of Wnt signaling, resulting in weakened osteoblast differentiation and reduced bone formation capacity. Single-gene GSEA in this study showed that DVL1 mainly regulates hormone signaling pathways, particularly estrogen and parathyroid hormone-related pathways. This finding is highly consistent with clinical observations that postmenopausal women become a high-risk population for osteoporosis due to decreased estrogen levels.[37] Meanwhile, estrogen deficiency can increase fracture incidence.[38] Recent molecular mechanism studies have shown that estrogen can regulate Wnt signaling activity through multiple pathways.[39] The downregulation of DVL1 expression may disrupt this delicate hormonal regulatory balance, leading to bone metabolism tilting toward bone resorption.

As an important member of the RNA-binding protein family, RNA Binding Motif Protein 39 (RBM39) mainly participates in post-transcriptional regulatory processes, including pre-mRNA splicing, mRNA stability regulation, and translational control.[40] RBM39 is an important component of the spliceosome, and its abnormal expression may lead to widespread alternative splicing abnormalities, thereby affecting fundamental cellular biological functions.[41] RBM39 is dysregulated in many cancers, and its upregulation can enhance cancer cell proliferation.[42,43] In this study, RBM39 showed significantly upregulated expression, providing a new perspective for understanding the molecular mechanisms of osteoporosis. GSEA analysis in this study showed that RBM39 mainly upregulates RNA metabolism and protein synthesis-related pathways, including spliceosome and ribosome biogenesis. This finding echoes recent research on the role of RNA metabolism in bone diseases. Studies have found that abnormal alternative splicing is a common feature of various bone diseases, particularly affecting the correct expression and modification of bone matrix proteins.[44] In osteoporosis, overexpression of RBM39 may lead to abnormal splicing of key bone metabolism-related genes, thereby affecting protein function and stability. In summary, the downregulated expression of DVL1 may weaken bone formation capacity by inhibiting the Wnt signaling pathway and hormonal regulatory networks, while the upregulated expression of RBM39 may disrupt the fine regulation of bone metabolism by interfering with RNA metabolism and protein synthesis processes. These two factors may synergistically contribute to the occurrence and development of osteoporosis. An important consideration is that while our study began with a tryptophan metabolism gene set, the ultimately identified key genes—DVL1 (Wnt pathway component) and RBM39 (RNA splicing regulator)—are not classic tryptophan metabolic enzymes such as IDO, TDO, or kynurenine pathway enzymes. This apparent disconnect requires careful interpretation. We propose that DVL1 and RBM39 may represent indirect regulatory nodes linking tryptophan metabolism to bone homeostasis through several potential mechanisms: (1) Tryptophan metabolites (particularly kynurenine pathway products) can modulate Wnt signaling, and DVL1 as a key Wnt pathway component may mediate these effects on osteoblast differentiation; (2) RBM39, as an RNA splicing factor, may regulate alternative splicing of genes encoding tryptophan metabolic enzymes or their downstream effectors, thereby indirectly controlling metabolic flux; (3) Both genes may be co-regulated with tryptophan metabolic genes within broader transcriptional programs governing immune-metabolic homeostasis in osteoporosis. However, we acknowledge that the mechanistic links between DVL1/RBM39 and tryptophan metabolism remain speculative based on current data. Future research should employ targeted approaches including: (1) metabolomics analysis to determine whether DVL1/RBM39 expression correlates with tryptophan metabolite profiles; (2) functional experiments examining whether tryptophan metabolites regulate DVL1/RBM39 expression or vice versa; (3) investigation of whether DVL1/RBM39 modulation affects expression or activity of core tryptophan metabolic enzymes. Until such mechanistic studies are completed, our findings should be interpreted as

identifying potential regulatory genes associated with a tryptophan metabolism gene signature in osteoporosis, rather than demonstrating direct participation in tryptophan metabolic pathways.

The lncRNAs and miRNAs, as important epigenetic regulatory factors, play crucial regulatory roles in the pathogenesis of osteoporosis. Nuclear Enriched Abundant Transcript 1 (NEAT1), as a nuclear-enriched long non-coding RNA, has been recently discovered to have important functions in bone metabolism regulation. Studies have shown that NEAT1 can regulate the expression of osteoblast differentiation-related genes through ceRNA mechanisms, and its abnormal expression is closely related to the occurrence of osteoporosis.[45] X–inactive specific transcript (XIST), as a key regulatory factor for X-chromosome inactivation, can significantly inhibit osteoblast differentiation in bone marrow mesenchymal stem cells.[46,47] KCNQ1 opposite strand/antisense transcript 1 (KCNQ1OT1), as an antisense transcript of the imprinted gene KCNQ1, plays an important role in bone development and bone homeostasis maintenance. KCNQ1OT1 can enhance osteoblast proliferation and differentiation capacity by inhibiting miR-205-5p to promote RICTOR expression.[48] In postmenopausal osteoporosis patients, dysregulated expression of KCNQ1OT1 may lead to disruption of the osteoblast/osteoclast balance, thereby accelerating bone loss.[49] The let-7 family, as classic tumor suppressor miRNAs, also has important significance in bone metabolism regulation. In the pathological process of osteoporosis, the let-7 family can weaken osteoblast differentiation capacity, leading to bone metabolism tilting toward bone resorption.[50] Transcription factor enrichment analysis identified Yin Yang-1 (YY1) as a common regulatory factor for DVL1 and RBM39, which is of significant importance. YY1, as an important transcriptional regulatory factor, exerts bone protective effects in postmenopausal osteoporosis by regulating the YY1/TGFβ1 signaling axis. Studies have shown that upregulation of YY1 expression can inhibit osteoblast ferroptosis, promote osteoblast viability and mineralization function, while coordinately regulating the expression of key osteogenic differentiation factors such as Runx2 and BMP2.[51] From a therapeutic perspective, the regulatory networks identified in this study suggest potential intervention targets. The YY1-RBM39 regulatory axis, in particular, represents a promising avenue for drug development. Modulating YY1 transcriptional activity or RBM39 splicing function could potentially influence tryptophan metabolism pathways and bone homeostasis. Additionally, the Wnt signaling pathway involvement through DVL1 suggests that existing Wnt-targeted therapies might be repurposed or optimized based on tryptophan metabolic status. However, these therapeutic implications require extensive preclinical validation before clinical translation.

Immune infiltration analysis revealed significant upregulation of CD4+ memory activated T cells in osteoporosis patients, and significant correlations between multiple cell subpopulations and key genes, which may play important regulatory roles in the pathogenesis of osteoporosis. Studies have shown that CD4+ T cells can directly affect osteoclast and osteoblast functions by secreting cytokines.[52] Activated CD4+ memory T cells mainly secrete pro-inflammatory cytokines such as tumor necrosis factor-α (TNF-α), interleukin-17 (IL-17), and receptor activator of nuclear factor kappa-B ligand (RANKL), which can promote osteoclast differentiation and bone resorption activity.[53] The findings of this study are highly consistent with this mechanism, suggesting their role in promoting bone destruction in osteoporosis progression. Conversely, CD8+ T cells usually exert bone protective effects by secreting anti-inflammatory factors and inhibiting osteoclast activity to maintain bone homeostasis.[54] M2-type macrophages are generally considered to have anti-inflammatory and tissue repair functions, promoting osteoblast differentiation and bone formation during normal bone remodeling.[55] This study found positive correlations between RBM39 and both M2 macrophages and memory B cells. The RBM39-M2 association requires careful interpretation: while M2 macrophages typically exert bone-protective functions through IL-10 and TGF-β secretion, their co-elevation with RBM39 may reflect (1) maladaptive compensation with functional exhaustion, (2) context-dependent pro-fibrotic dysfunction, or (3) RBM39-driven pathological polarization rather than protective responses. Similarly, the positive correlation between memory B cells and RBM39 may reflect compensatory changes in B cell function, as studies have found significant B-cell dysfunction in postmenopausal osteoporosis patients leading to increased bone resorption.[56] These findings highlight the importance of distinguishing correlation from causation in immune-metabolic interactions. Future mechanistic validation through functional assays and conditional knockout models is essential to determine whether these associations represent therapeutic targets versus reactive biomarkers.

Despite inherent limitations of using mouse models—including species differences in skeletal growth patterns, cortical bone remodeling mechanisms, and hormonal profiles compared to humans—single-cell transcriptomic analysis

provided valuable insights into cellular heterogeneity. Single-cell transcriptomic analysis revealed that multiple cell subpopulations play complex and precise regulatory roles in the pathogenesis of osteoporosis. Endothelial cells play a central role in bone-vascular coupling, and their functional abnormalities directly affect bone remodeling processes. Studies have found that endothelial cells regulate osteoblast proliferation and differentiation by secreting vascular endothelial growth factor (VEGF). Furthermore, the decline in endothelial cell function leads to decreased bone vascularization, thereby affecting the supply of nutrients and oxygen and accelerating bone loss in osteoporosis.[57] Notably, the active communication characteristics exhibited by endothelial cells in the cell-cell communication network found in this study further confirm the central position of the vascular system in the bone microenvironment. Endothelial cells not only affect adjacent bone cells through direct paracrine signals but may also indirectly affect the metabolic status of the entire bone microenvironment by regulating local oxygen and nutrient supply.[58] The dysfunction of this complex cell-cell interaction network may be a key link in the pathogenesis of osteoporosis, providing an important cellular basis for understanding the systemic characteristics of this disease.

This study confirmed the key roles of DVL1 and RBM39 in osteoporosis through multiple levels of validation strategies. MR analysis provided strong causal evidence, showing that DVL1 expression levels were negatively correlated with osteoporosis risk (OR=0.94), while RBM39 was positively correlated (OR=1.22). The sensitivity analyses revealed different patterns of evidence strength for the two MR-identified genes. DVL1 showed no outlier SNPs in MR-PRESSO but exhibited systematic horizontal pleiotropy in MR-Egger intercept test. In contrast, RBM39's MR-PRESSO detected potential outliers, yet demonstrated no systematic pleiotropy in MR-Egger and maintained stable estimates in leave-one-out analysis. For RBM39, the discordance between MR-PRESSO and MR-Egger suggests that individual SNP deviations do not introduce systematic bias through pleiotropic pathways. Combined with the absence of heterogeneity, significant IVW result, and correct causal directionality in Steiger test, these findings support a robust causal relationship. This is further strengthened by RBM39's direct involvement in bone metabolism through regulating alternative splicing of osteogenic genes and bone matrix proteins. For DVL1, the systematic pleiotropy detected by MR-Egger, despite the absence of outliers in MR-PRESSO, suggests that its pleiotropic effects may reflect its biological role as a central hub in Wnt signaling regulating multiple cellular processes. The non-significant IVW result further limits the strength of causal inference. This qualitative difference underscores the importance of triangulating evidence from multiple sensitivity analyses. Based on comprehensive evaluation, RBM39 provides more robust causal evidence and was therefore prioritized for experimental validation. RT-qPCR experimental validation further confirmed the reliability of transcriptomic analysis results, with RBM39 significantly upregulated and DVL1 significantly downregulated in osteoporosis patients, completely consistent with bioinformatics analysis results. Single-cell transcriptomic analysis further validated the differential expression patterns of these genes at the cell subpopulation level, revealing their specific expression changes in different bone cell types. This multi-dimensional validation strategy from population genetics, tissue transcriptomics to single-cell levels not only ensures the scientific rigor of research results but also provides a solid experimental foundation for the clinical translational application of DVL1 and RBM39 as potential therapeutic targets and biomarkers for osteoporosis. However, the diagnostic model based on DVL1 and RBM39 showed moderate performance (AUC > 0.6), though direct comparison with traditional imaging methods or bone turnover markers requires prospective clinical studies.

Currently, dual-energy X-ray absorptiometry (DXA) remains the gold standard for osteoporosis diagnosis; however, DXA primarily captures static bone mineral density and cannot reflect dynamic molecular changes underlying early-stage bone metabolic dysregulation. In contrast, peripheral blood mRNA expression levels of DVL1 and RBM39 may serve as complementary indicators, potentially enabling detection of pathological processes before clinically significant bone loss becomes apparent on DXA.

The clinical feasibility of DVL1 and RBM39 as biomarkers requires evaluation across technical, population, and economic dimensions. Both genes can be quantified via standard RT-qPCR from peripheral blood mononuclear cells, offering advantages over invasive bone biopsies. RNA stabilization systems (PAXgene, Tempus) ensure sample integrity during collection and transport, facilitating multi-center studies. However, validation across diverse populations with varying ethnic backgrounds, comorbidities, and medication histories remains essential, as current transcriptomic datasets primarily represent limited demographic groups and exhibit racial disparities in screening and diagnosis. Large-scale

prospective studies with longitudinal follow-up are needed to establish standardized cutoff values and sensitivity/specificity thresholds. Cost-effectiveness analyses comparing molecular biomarker panels with current diagnostic standards are crucial for healthcare system adoption, though early evidence suggests molecular markers may enhance fracture risk assessment tools like FRAX, particularly for early-stage high-risk identification.

Although this study provided new insights into understanding the molecular mechanisms of osteoporosis through multi-level bioinformatics analysis, some limitations still exist. The use of mouse-derived single-cell RNA-sequencing data (GSE172410) introduces significant translational limitations. While this dataset provides insights into cell type-specific expression patterns, direct extrapolation to human bone biology is problematic due to fundamental species differences in skeletal physiology, bone cell biology, and regulatory mechanisms. Furthermore, this dataset originates from normal mouse skeletal tissue rather than an osteoporosis model, limiting its relevance to disease pathogenesis. The identification of cell types such as myonuclei and tenocytes reflects the anatomical sampling (femur and tibia with associated soft tissues) but does not represent typical human PBMC or bone marrow cell populations analyzed in the bulk transcriptomic studies. Ideally, validation should employ human bone tissue single-cell RNA-sequencing from osteoporosis patients and matched controls, which would enable direct assessment of DVL1 and RBM39 expression in human osteoblasts, osteoclasts, osteocytes, and bone marrow stromal cells. Such data would provide definitive evidence regarding cell type-specific roles of these genes in human osteoporosis pathogenesis. Critically, the core transcriptomic data (GSE56815) derives from peripheral blood mononuclear cells (PBMCs) rather than bone tissue. While PBMC gene expression may reflect systemic immune-metabolic status, it cannot be directly equated with molecular events in the bone microenvironment. Therefore, extrapolation of conclusions requires caution, and validation in bone tissue samples is essential. First, transcriptomic data can only reflect changes in gene expression levels, while changes in protein levels and functional activities require further research. Second, single-cell data came from mouse models, and their applicability in humans needs further validation. Third, modest sample sizes limit statistical power and generalizability. Cross-species comparison between human bulk and mouse single-cell transcriptomics introduces inherent limitations, as mice exhibit continuous skeletal growth, lack osteonal remodeling in cortical bone, and possess different hormonal profiles compared to humans. These species-specific differences in bone metabolism, cellular composition, and regulatory networks may affect the translatability of molecular mechanisms. Reliance on publicly available datasets restricts control over experimental variables, tissue processing protocols, and clinical phenotyping depth. These limitations necessitate large-scale, prospectively designed human studies with matched bulk and single-cell transcriptomic profiling from the same individuals, and validation across diverse populations and osteoporosis subtypes (postmenopausal, senile, secondary).

Future research can validate the specific functions of DVL1 and RBM39 in bone metabolism regulation through in vitro cell experiments and in vivo animal models; second, conduct large-scale prospective cohort studies to validate the clinical diagnostic value of these biomarkers; third, develop a translational pathway for therapeutic interventions by: (1) screening and optimizing small molecule modulators targeting DVL1 upregulation or RBM39 inhibition through high-throughput drug screening platforms; (2) designing gene therapy strategies such as adeno-associated virus (AAV)-mediated DVL1 overexpression or CRISPR-based RBM39 knockdown in preclinical models; (3) engineering mesenchymal stem cells or osteoblast precursors with modified DVL1/RBM39 expression for autologous cell therapy; (4) conducting pharmacokinetics, toxicology, and efficacy studies in animal models to establish dose-response relationships and safety profiles; (5) progressing promising candidates through Phase I/II clinical trials to assess safety, tolerability, and preliminary efficacy in osteoporosis patients; finally, combine multi-omics data (proteomics, metabolomics, etc) to further improve the molecular atlas of osteoporosis and identify synergistic therapeutic targets that can enhance treatment efficacy when combined with DVL1/RBM39-based interventions.

Conclusion

This study systematically elucidated the mechanisms of tryptophan metabolism in osteoporosis, identifying DVL1 and RBM39 as key regulatory genes with potential as diagnostic biomarkers and therapeutic targets. DVL1 and RBM39 offer multiple therapeutic opportunities including direct molecular targeting through small molecule modulators or biologics, metabolic intervention via tryptophan pathway regulation, personalized treatment guided by expression profiling for patient stratification and treatment prediction, drug repurposing of existing compounds, and early prevention through risk

identification before significant bone loss occurs. These findings provide a molecular foundation for precision diagnosis and individualized treatment of osteoporosis. Future validation should include functional characterization through experimental models, prospective clinical cohort studies, and development of targeted therapeutic strategies. Current limitations include reliance on publicly available datasets and incomplete mechanistic understanding. Integrating multi-omics data and conducting mechanistic studies will further advance effective diagnostic and therapeutic strategies for osteoporosis.

Data Sharing Statement

All relevant data supporting the findings of this research are provided within the manuscript. The datasets utilized in this analysis were sourced from publicly accessible databases, specifically the GEO and GeneCards repositories. The processed data generated during this research can be made available to interested researchers upon reasonable request through correspondence with the corresponding author via the provided Email contact.

Ethics Approval and Consent to Participate

This study was approved by the Ethics Committee of Beijing Jishuitan Hospital Guizhou Hospital (Approval Number: KT2024060504, Date: [2024-06-07]). All procedures were conducted in accordance with the Declaration of Helsinki and relevant institutional guidelines.

All participants or their legal guardians provided written informed consent prior to sample collection. All samples were anonymized, and patient confidentiality was strictly maintained throughout the study.

Acknowledgments

We would like to express our sincere gratitude to all the participants for their contributions to this study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

There is no funding to report.

Disclosure

The authors declare no conflicts of interest related to this study.

References

1. Camacho PM, Petak SM, Binkley N, et al. American Association of Clinical Endocrinologists/American College of Endocrinology clinical practice guidelines for the diagnosis and treatment of postmenopausal osteoporosis-2020 update. *Endocr Pract*. 2020;26(Suppl 1):1–46. doi:10.4158/GL-2020-0524SUPPL
2. Sözen T, Özışık L, Başaran N. An overview and management of osteoporosis. *Eur J Rheumatol*. 2017;4(1):46–56. doi:10.5152/eurjrheum.2016.048
3. Chen P, Li Z, Hu Y. Prevalence of osteoporosis in China: a meta-analysis and systematic review. *BMC Public Health*. 2016;16(1):1039. doi:10.1186/s12889-016-3712-7
4. Borgström F, Karlsson L, Ortsäter G, et al. Fragility fractures in Europe: burden, management and opportunities. *Arch Osteoporos*. 2020;15(1):59. doi:10.1007/s11657-020-0706-y
5. Morin SN, Feldman S, Funnell L, et al. Clinical practice guideline for management of osteoporosis and fracture prevention in Canada: 2023 update. *Cmaj*. 2023;195(39):E1333–e1348. doi:10.1503/cmaj.221647
6. Veis DJ, O'Brien CA. Osteoclasts, master sculptors of bone. *Annu Rev Pathol*. 2023;18:257–281. doi:10.1146/annurev-pathmechdis-031521-040919
7. Tsukasaki M, Takayanagi H. Osteoimmunology: evolving concepts in bone-immune interactions in health and disease. *Nat Rev Immunol*. 2019;19(10):626–642. doi:10.1038/s41577-019-0178-8
8. Rao CM, Singh P, Maikap D, et al. Musculoskeletal disorders in chronic obstructive airway diseases: a neglected clinical entity. *Mediterr J Rheumatol*. 2021;32(2):118–123. doi:10.31138/mjr.32.2.118

9. Rohatgi N, Zou W, Li Y, et al. BAP1 promotes osteoclast function by metabolic reprogramming. *Nat Commun*. 2023;14(1):5923. doi:10.1038/s41467-023-41629-4
10. Cervenka I, Agudelo LZ, Ruas JL. Kynurenines: tryptophan's metabolites in exercise, inflammation, and mental health. *Science*. 2017;357(6349). doi:10.1126/science.aaf9794
11. Badawy AA. Tryptophan metabolism: a versatile area providing multiple targets for pharmacological intervention. *Egypt J Basic Clin Pharmacol*. 2019;9. doi:10.32527/2019/101415
12. Vidal C, Li W, Santner-Nanan B, et al. The kynurenine pathway of tryptophan degradation is activated during osteoblastogenesis. *Stem Cells*. 2015;33(1):111–121. doi:10.1002/stem.1836
13. Chen Y, Yang C, Deng Z, et al. Gut microbially produced tryptophan metabolite melatonin ameliorates osteoporosis via modulating SCFA and TMAO metabolism. *J Pineal Res*. 2024;76(3):e12954. doi:10.1111/jpi.12954
14. He J, Song R, Xiao F, et al. Cu(3)P/1-MT nanocomposites potentiated photothermal-immunotherapy. *Int J Nanomed*. 2023;18:3021–3033. doi:10.2147/IJN.S414117
15. Chabbi-Achengli Y, Coudert AE, Callebert J, et al. Decreased osteoclastogenesis in serotonin-deficient mice. *Proc Natl Acad Sci U S A*. 2012;109(7):2567–2572. doi:10.1073/pnas.1117792109
16. Yadav VK, Ryu JH, Suda N, et al. Lrp5 controls bone formation by inhibiting serotonin synthesis in the duodenum. *Cell*. 2008;135(5):825–837. doi:10.1016/j.cell.2008.09.059
17. Chen X, Wang Z, Duan N, et al. Osteoblast-osteoclast interactions. *Connect Tissue Res*. 2018;59(2):99–107. doi:10.1080/03008207.2017.1290085
18. Wang J, Zhang H, Cao Y, et al. circ_0134120: a new frontier in understanding postmenopausal osteoporosis pathogenesis. *J Mol Endocrinol*. 2024;73(1):13. doi:10.1530/JME-23-0140
19. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf*. 2008;9:559. doi:10.1186/1471-2105-9-559
20. Swan AL, Mobasheri A, Allaway D, et al. Application of machine learning to proteomics data: classification and biomarker identification in postgenomics biology. *Omics*. 2013;17(12):595–610. doi:10.1089/omi.2013.0017
21. Carvalho FR, Gavaia PJ. Enhancing osteoporosis risk prediction using machine learning: a holistic approach integrating biomarkers and clinical data. *Comput Biol Med*. 2025;192(Pt B):110289. doi:10.1016/j.compbiomed.2025.110289
22. Carvalho FR, Gavaia PJ, Brito Camacho A. OrthoMortPred: predicting one-year mortality following orthopedic hospitalization. *Int J Med Inform*. 2024;192:105657. doi:10.1016/j.ijmedinf.2024.105657
23. Liu H, Feng Y, Lin B, et al. Identification of Cystatin 3 as a potential diagnostic biomarker for osteoporosis using machine learning. *J Inflamm Res*. 2025;18:16805–16823. doi:10.2147/JIR.S569594
24. Liu T, Fan J, Fang J, et al. Identification and validation of the cellular senescence-associated molecular pattern and diagnostic markers for osteoporosis. *BMC Med Genomics*. 2025;18(1):140. doi:10.1186/s12920-025-02205-5
25. Ho J, Mak CCH, Sharma V, et al. Mendelian randomization studies of lifestyle-related risk factors for osteoarthritis: a PRISMA review and meta-analysis. *Int J Mol Sci*. 2022;23(19):11906. doi:10.3390/ijms231911906
26. Zheng J, Baird D, Borges MC, et al. Recent developments in mendelian randomization studies. *Curr Epidemiol Rep*. 2017;4(4):330–345. doi:10.1007/s40471-017-0128-6
27. Kemp JP, Morris JA, Medina-Gomez C, et al. Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nat Genet*. 2017;49(10):1468–1475. doi:10.1038/ng.3949
28. Salmena L, Poliseno L, Tay Y, et al. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell*. 2011;146(3):353–358. doi:10.1016/j.cell.2011.07.014
29. Li T, Yuan J, Xu P, et al. PMAIP1, a novel diagnostic and potential therapeutic biomarker in osteoporosis. *Aging*. 2024;16(4):3694–3715. doi:10.18632/aging.205553
30. Wang Y, Zeng X, Wang N, et al. Long noncoding RNA DANCR, working as a competitive endogenous RNA, promotes ROCK1-mediated proliferation and metastasis via decoying of miR-335-5p and miR-1972 in osteosarcoma. *Mol Cancer*. 2018;17(1):89. doi:10.1186/s12943-018-0837-6
31. Vamadevan V, Chaudhary N, Maddika S. Ubiquitin-assisted phase separation of dishevelled-2 promotes Wnt signalling. *J Cell Sci*. 2022;135(24). doi:10.1242/jcs.260284
32. Hu L, Chen W, Qian A, et al. Wnt/beta-catenin signaling components and mechanisms in bone formation, homeostasis, and disease. *Bone Res*. 2024;12(1):39.
33. Vlashi R, Zhang X, Wu M, et al. Wnt signaling: essential roles in osteoblast differentiation, bone metabolism and therapeutic implications for bone and skeletal disorders. *Genes Dis*. 2023;10(4):1291–1317. doi:10.1016/j.gendis.2022.07.011
34. Gao Y, Chen N, Fu Z, et al. Progress of Wnt signaling pathway in osteoporosis. *Biomolecules*. 2023;13(3):483. doi:10.3390/biom13030483
35. Zhang Y, Zhao Y, Xie Z, et al. Activating Wnt/beta-Catenin signaling in osteocytes promotes osteogenic differentiation of BMSCs through BMP-7. *Int J Mol Sci*. 2022;23(24):16045. doi:10.3390/ijms232416045
36. Ben-Ghedalia-Peled N, Vago R. Wnt signaling in the development of bone metastasis. *Cells*. 2022;11(23):3934. doi:10.3390/cells11233934
37. Charde SH, Joshi A, Raut J. A comprehensive review on postmenopausal osteoporosis in women. *Cureus*. 2023;15(11):e48582. doi:10.7759/cureus.48582
38. Kim H, Kang MJ, Baek JK, et al. Relationship between menopausal hormone therapy and incidence of fractures in postmenopausal women. *Climacteric*. 2024;27(2):165–170. doi:10.1080/13697137.2023.2273528
39. Padilla-Banks E, Jefferson WN, Papas BN, et al. Developmental estrogen exposure in mice disrupts uterine epithelial cell differentiation and causes adenocarcinoma via Wnt/beta-catenin and PI3K/AKT signaling. *PLoS Biol*. 2023;21(10):e3002334. doi:10.1371/journal.pbio.3002334
40. Eleouet M, Lu C, Zhou Y, et al. Insights on the biological functions and diverse regulation of RNA-binding protein 39 and their implication in human diseases. *Biochim Biophys Acta Gene Regul Mech*. 2023;1866(1):194902. doi:10.1016/j.bbagrm.2022.194902
41. Campagne S, Jutzi D, Malard F, et al. Molecular basis of RNA-binding and autoregulation by the cancer-associated splicing factor RBM39. *Nat Commun*. 2023;14(1):5366. doi:10.1038/s41467-023-40254-5
42. Lu CP, Li JB, Li DB, et al. RNA-binding motif protein RBM39 enhances the proliferation of gastric cancer cells by facilitating an oncogenic splicing switch in MRPL33. *Acta Pharmacol Sin*. 2025;46(4):1068–1081. doi:10.1038/s41401-024-01431-4

43. Ji T, Yang Y, Yu J, et al. Targeting RBM39 through indisulam induced mis-splicing of mRNA to exert anti-cancer effects in T-cell acute lymphoblastic leukemia. *J Exp Clin Cancer Res*. 2024;43(1):205. doi:10.1186/s13046-024-03130-8
44. Penna MS, Hu RC, Rodney GG, et al. The role of Limch1 alternative splicing in skeletal muscle function. *Life Sci Alliance*. 2023;6(6):e202201868. doi:10.26508/lsa.202201868
45. Zhao X, Zhao D, Geng B, et al. A novel ceRNA regulatory network involving the long noncoding NEAT1, miRNA-466f-3p and its mRNA target in osteoblast autophagy and osteoporosis. *J Mol Med*. 2022;100(11):1629–1646. doi:10.1007/s00109-022-02255-7
46. Morey C, Rougeulle C, Ouimette JF. Unleashing XIST from X-chromosome inactivation. *Curr Opin Cell Biol*. 2025;92:102446. doi:10.1016/j.ceb.2024.102446
47. Chen X, Yang L, Ge D, et al. Long non-coding RNA XIST promotes osteoporosis through inhibiting bone marrow mesenchymal stem cell differentiation. *Exp Ther Med*. 2019;17(1):803–811. doi:10.3892/etm.2018.7033
48. Yang JJ, Peng WX, Zhang MB. LncRNA KCNQ1OT1 promotes osteogenic differentiation via miR-205-5p/RICTOR axis. *Exp Cell Res*. 2022;415(1):113119. doi:10.1016/j.yexcr.2022.113119
49. Wang Z, Zhang H, Li Q, et al. Long non-coding RNA KCNQ1OT1 alleviates postmenopausal osteoporosis by modulating miR-421-3p/mTOR axis. *Sci Rep*. 2023;13(1):2333. doi:10.1038/s41598-023-29546-4
50. Chen R, Liao X, Chen F, et al. Circulating microRNAs, miR-10b-5p, miR-328-3p, miR-100 and let-7, are associated with osteoblast differentiation in osteoporosis. *Int J Clin Exp Pathol*. 2018;11(3):1383–1390.
51. Wang H, Qian J. Picroside II regulates the YY1/TGFβ1 axis by inhibiting osteoblast ferroptosis to alleviate symptoms of postmenopausal osteoporosis. *Cytotechnology*. 2025;77(3):114. doi:10.1007/s10616-025-00783-x
52. Okamoto K, Takayanagi H. Effect of T cells on bone. *Bone*. 2023;168:116675. doi:10.1016/j.bone.2023.116675
53. Zhang R, Peng S, Zhu G. The role of secreted osteoclastogenic factor of activated T cells in bone remodeling. *Jpn Dent Sci Rev*. 2022;58:227–232. doi:10.1016/j.jdsr.2022.07.001
54. Avery D, Morandini L, Sheakley L, et al. CD4(+) and CD8(+) T cells reduce inflammation and promote bone healing in response to titanium implants. *Acta Biomater*. 2024;179:385–397. doi:10.1016/j.actbio.2024.03.022
55. Lu D, Zhang Y, Liang S, et al. M2 macrophages guide periosteal stromal cell recruitment and initiate bone injury regeneration. *Biomedicines*. 2024;12(6):1205. doi:10.3390/biomedicines12061205
56. Sapra L, Saini C, Mishra PK, et al. Compromised anti-osteoclastogenic and immunomodulatory functions of regulatory B cells (Bregs) aggravate inflammatory bone loss in post-menopausal osteoporosis. *Biochim Biophys Acta Mol Basis Dis*. 2025;1871(3):167675. doi:10.1016/j.bbadis.2025.167675
57. Shi H, Zhao Z, Jiang W, et al. A review into the insights of the role of endothelial progenitor cells on bone biology. *Front Cell Dev Biol*. 2022;10:878697. doi:10.3389/fcell.2022.878697
58. Petrillo S, Genova T, Chinigo G, et al. Endothelial cells promote osteogenesis by establishing a functional and metabolic coupling with human mesenchymal stem cells. *Front Physiol*. 2021;12:813547. doi:10.3389/fphys.2021.813547