

Employing single-cell RNA sequencing coupled with an array of bioinformatics approaches to ascertain the shared genetic characteristics between osteoporosis and obesity

From First Hospital of China Medical University, Shenyang, China

D. Liu,¹ F. Cao,¹ D. Liu,¹ H. Li,¹ L. Tao,¹ Y. Zhu¹

Department of Orthopedics, The First Hospital of China Medical University, Shenyang, China

Cite this article:

Bone Joint Res 2024;13(10): 573–587.

DOI: 10.1302/2046-3758.

1310.BJR-2023-0366.R1

Correspondence should be sent to Yue Zhu zhuyuedr@163.com

Aims

This study examined the relationship between obesity (OB) and osteoporosis (OP), aiming to identify shared genetic markers and molecular mechanisms to facilitate the development of therapies that target both conditions simultaneously.

Methods

Using weighted gene co-expression network analysis (WGCNA), we analyzed datasets from the Gene Expression Omnibus (GEO) database to identify co-expressed gene modules in OB and OP. These modules underwent Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment and protein-protein interaction analysis to discover Hub genes. Machine learning refined the gene selection, with further validation using additional datasets. Single-cell analysis emphasized specific cell subpopulations, and enzyme-linked immunosorbent assay (ELISA), protein blotting, and cellular staining were used to investigate key genes.

Results

WGCNA revealed critical gene modules for OB and OP, identifying the *Toll-like receptor (TLR)* signalling pathway as a common factor. *TLR2* was the most significant gene, with a pronounced expression in macrophages. Elevated *TLR2* expression correlated with increased adipose accumulation, inflammation, and osteoclast differentiation, linking it to OP development.

Conclusion

Our study underscores the pivotal role of *TLR2* in connecting OP and OB. It highlights the influence of *TLR2* in macrophages, driving both diseases through a pro-inflammatory mechanism. These insights propose *TLR2* as a potential dual therapeutic target for treating OP and OB.

Article focus

- What are the key genes shared between osteoporosis (OP) and obesity (OB)?
- Which cell subtype is most strongly associated with OP and OB?
- How do key genes and cell subtypes contribute to the progression of both diseases?

Key messages

- Toll-like receptor 2 (*TLR2*) acts as a cross-talk gene in the pathogenesis of both OP and OB.

- The inflammatory response induced by macrophages serves as a crucial bridge linking OP and OB.

Strengths and limitations

- This study suggests that the *TLR2* gene in macrophages is a potential target for the concurrent treatment of OP and OB, guiding the design of drugs that address both conditions through a single target.
- Our research primarily operates at the cellular level. Future work necessitates the development of animal models for

experimental validation to enhance the accuracy of the results.

Introduction

Osteoporosis (OP) is a metabolic bone disorder frequently encountered in clinical settings, characterized by diminished bone density, reduced mineral content, a decrease in trabecular number, and altered microarchitectural features of bone. Due to the increased brittleness of bones, the risk of fractures in patients correspondingly escalates, serving as a substantial risk factor for the occurrence of various postoperative complications.¹ With the ageing of the global population and increasing life expectancy, OP has emerged as a major health concern worldwide. Earlier research indicates a staggering 19.7% global prevalence rate of OP, predominantly affecting postmenopausal and perimenopausal women.^{2,3} The onset of this disorder is commonly attributed to profound alterations in bone remodelling. A pivotal factor for such changes is the imbalance caused by the diminished bone-forming capability of osteoblasts and the amplified bone-resorptive capacity of osteoclasts. Particularly, in the presence of ageing-related inflammation and hormonal influences, precursor osteoclast cells are stimulated by factors secreted from osteoblasts, T-cells, and macrophages, such as colony-stimulating factor-1 and receptor activator of nuclear factor κ B ligand (RANKL). This stimulation activates the β -catenin-dependent canonical Wnt pathway, fostering the differentiation of mature osteoclasts and amplifying bone resorption.^{4,5} Amid the myriad risk factors, inflammatory stimuli appear to be of paramount significance. OP manifests as a systemic inflammatory response leading to bone loss, typically accompanied by the prolific release of inflammatory mediators such as reactive oxygen species and inflammatory cytokines. Previous investigations have suggested that inflammatory cytokines such as interleukin (IL)-1, IL-6, and macrophage colony-stimulating factor (M-CSF) serve as precursors that stimulate the differentiation of osteoclasts via RANKL activators, either directly or indirectly augmenting bone resorption.⁶

Obesity (OB) is conventionally defined in the literature as a BMI exceeding 30 kg/m². As reported in 2021, nearly 200 million individuals globally are grappling with OB or overweight issues – specifically, in China the OB rates have reached 14.4% for women and 16.0% for men, with abdominal OB rates at 32.7% and 36.6%, respectively.⁷ Concurrently, the proportion of diseases and fatalities induced by OB has risen progressively. OB invariably precipitates functional aberrations across various bodily systems. Pathogeneses of numerous ailments, including type II diabetes, cardiovascular disorders, bone metabolic diseases, and rheumatoid autoimmune diseases, bear a close nexus with OB. The systemic metabolic imbalance induced by OB arises from its capability to incite systemic immune system activation and ensuing inflammatory responses. Indeed, OB is typified as a state of chronic low-grade inflammation, where the infiltration of bone marrow-derived inflammatory cells into the white adipose tissue (WAT) is a hallmark feature.⁸ This inflammatory state results from the hypertrophy of adipocytes and the activation of macrophages within the adipose tissue. These activated macrophages exude various inflammatory

cytokines such as tumour necrosis factor- α (TNF- α), monocyte chemoattractant protein-1 (MCP-1), and IL-1 β . The presence of these inflammatory cytokines actuates metabolic pathways affiliated with myriad diseases, thereby precipitating their onset. Interestingly, macrophages – responsible for antigen presentation, recruitment of peripheral immune cells, and secretion of inflammatory cytokines – have become the focal point of our attention. The polarization state of these macrophages undergoes alterations in varying microenvironments, manifesting as M1 (pro-inflammatory) and M2 (anti-inflammatory) types. It is this imbalance in the ratio of M1 to M2 cells that exacerbates inflammatory responses.⁹ The macrophage-mediated inflammatory response appears to be a pivotal nexus linking OB and OP.

Of note is the intricate relationship between OB and OP, both of which are metabolic disorders. A 2013 clinical study revealed that pre- and perimenopausal obese women displayed an inverse correlation between visceral and trunk fat content with bone density and trabecular volume. This suggests that obese women possess subpar bone formation capabilities and inferior bone quality.¹⁰ Another study, utilizing a mouse model of OB induced by a high-fat diet over 12 weeks, revealed simultaneous bone marrow adipose tissue (BMAT) expansion with bone loss and bone formation impediments.¹¹ Bone OB induces fat accumulation, where the WAT secretes a plethora of adipokines such as leptin and adiponectin, along with various inflammatory mediators, invoking an excessive inflammatory response that impacts bone formation and resorption. Lipidomic analysis of serum from obese model mice revealed that dysregulation in triglyceride and phospholipid levels, due to substantial alterations in lipid metabolic pathways, precipitates a decline in skeletal strength, thereby adversely affecting bone health.¹² Moreover, OB significantly hampers the skeletal development of adolescents. A survey by Goulding et al¹³ on 90 children and adolescents who repeatedly experienced forearm fractures demonstrated a pronounced negative correlation between BMI z-scores and the ultra-distal radius's bone mineral content (BMC) and bone mineral density (BMD), identifying these as crucial risk factors for fractures. In obese children and adolescents, diet, physical activity, and endocrine level alterations lead to ectopic deposition of BMAT, causing a physiological shift that disrupts the balance in mesenchymal stem cell (MSC) differentiation. This imbalance prioritizes the differentiation towards adipocytes at the expense of osteoblasts, consequently notably significantly increasing the risk and susceptibility to bone fragility and fractures.¹⁴ However, most of these studies have primarily probed from the perspective of clinical osteoporotic parameters and serological adipokines, without delving into the genetic level to uncover shared pathogenic genes and mechanisms underlying OP and OB.

With the advent of information technology, high-throughput sequencing and gene microarray techniques have garnered extensive utility in biomedicine, playing pivotal roles. Scholars can now measure the expression levels of a plethora of genes, consequently isolating those intimately associated with disease, thereby facilitating a profound exploration into the pathogenesis of illnesses. Employing an integrative bioinformatics approach, combining single-cell RNA sequencing (scRNA-seq) and bulk RNA-seq, we

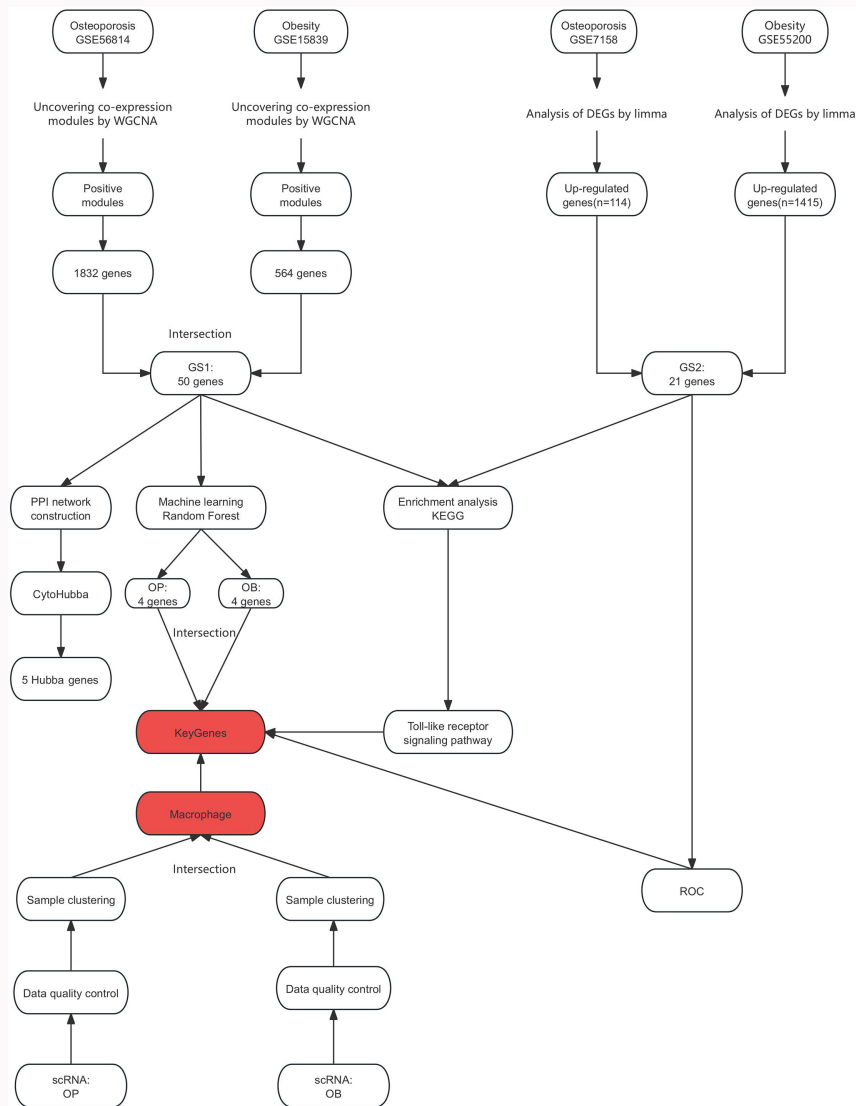


Fig. 1

Study flowchart. DEG, differentially expressed gene; GS1, gene set 1; GS2, gene set 2; GSE, Gene Expression Omnibus; KEGG, Kyoto Encyclopedia of Genes and Genomes; limma, linear models for microarray data; OB, obesity; OP, osteoporosis; PPI, protein-protein interaction; ROC, receiver operating characteristic; WGCNA, weighted gene co-expression network analysis.

first identified associated genes from bulk RNA-seq datasets. Leveraging machine-learning algorithms for statistical screening, we pinpointed the critical gene, Toll-like receptor 2 (*TLR2*). Utilizing single-cell technology and immune infiltration methods, we discerned that *TLR2* exhibits maximal expression in macrophages. Concurrently, we constructed a competitive endogenous RNA (ceRNA) network, seeking and analyzing the modulatory mechanisms of associated genes at the post-transcriptional level. Subsequently, through experimental methodologies, we corroborated the association of *TLR2* between two distinct diseases at the cellular level, heralding the future prospect of targeting this locus for the concurrent treatment of OB and OP.

Methods

Data download and processing

In Figure 1, the study flowchart is shown. We retrieved datasets related to OP and OB by searching the Gene Expression Omnibus (GEO) database with the keywords

“Obesity” and “Osteoporosis”. The selected tissues for sequencing included peripheral blood mononuclear cells (PBMCs) and adipose tissue. The datasets came from both Affymetrix (USA) and Illumina (USA) platforms. Using R software (version 4.1.3; R Foundation for Statistical Computing, Austria), we processed the raw data with the *affy* and *limma* packages.^{15,16} Data quality was assessed using the *arrayQualityMetrics* package,¹⁷ and normalization was performed using *gcrma* and *limiExpresso* functions.^{15,18} We then used the *ComBat* function from the *sva* package to reduce batch effects,^{19,20} resulting in a refined gene expression matrix. Finally, we annotated the matrix further using the *clusterProfiler* package’s *bitr* function,¹⁶ preparing it for subsequent analysis.

WGCNA and module gene selection

To explore the interrelationship between genes, we used datasets GSE151839 and GSE56814, applying the weighted gene co-expression network analysis (WGCNA) algorithm. In

the preprocessing stage, we selected the top 5,000 genes with the highest standard deviation, and eliminated genes with multiple missing values using the `goodSamplesGenes` function.²¹ For network construction, a soft-thresholding power was chosen based on the results from the `pickSoft-Threshold` function.²² In the module identification phase, the `blockwiseModules` function²¹ was used, setting parameters like `TOMType = "unsigned"`. The `plotDendroAndColors` function²¹ was employed to create a dendrogram representing the modules. Module feature analysis involved calculating the correlation between modules and traits using the `module-TraitCor` function,²¹ and a heatmap was generated using the `labeledHeatmap` function.²¹ Finally, key genes were identified through intramodular analysis, laying the groundwork for further visualization in Cytoscape (Institute for Systems Biology, USA), facilitated by the `exportNetworkToCytoscape` function.²¹

Enrichment analysis of module intersection genes

We focused on modules highly correlated with OB and OP. From these modules, shared genes were identified and named Gene Set 1 (GS1). In R, using the `clusterProfiler` package, Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was conducted ($p < 0.05$, hypergeometric test) (Supplementary Table i),^{23,24} and results were visualized using the `ComplexHeatmap` package.²⁵ Subsequently, KEGG results were displayed as a network of functional groups using the `ClueGO` plugin in Cytoscape (minimum interaction score threshold of 0.4).²⁶

Protein-protein interaction network construction and Hub gene selection

Using the STRING database (version 11.5),²⁷ we constructed a protein-protein interaction (PPI) network for GS1 with a minimum interaction score of 0.4.²⁸ The `cytoHubba` plugin in Cytoscape was employed to identify and analyze Hub genes within the network, which are closely associated with diseases and significantly influence experimental outcomes.^{29,30}

Machine learning

Random forest (RF) is an ensemble learning method that enhances the accuracy and stability of classification or regression by building multiple decision trees and using their vote or mean predictions. To reduce the number of feature genes, we employed this classifier for gene feature selection. We extracted and analyzed the expression matrices of GS1 in OB and bone pathology (OP), using the `randomForest` package,³¹ identifying two sets of disease-related genes. The intersection of these sets provided genes for further analysis.³² Before using RF, we set a seed to ensure reproducibility of our results. We then assessed the error rate of the model based on the dataset, with the error type being standard error (SE). This involved comparing each tree's classification results with the actual label to calculate errors.

Analyzing shared genes with a validation set

Enrichment analysis of differential genes in the validation set: Using R software and the `Limma` package, differential gene analysis was performed on the validation set with a cutoff of $\text{abs}(\log_2) > 0.3$. A p -value < 0.05 was considered statistically significant. This approach identified overlapping

differential genes between the two datasets, termed Gene Set 2 (GS2). In R, the `clusterProfiler` package was used, applying the `enrichKEGG` function to GS2 for KEGG enrichment analysis. Results were visualized using the `ComplexHeatmap` package. Finally, using the `ClueGO` plugin within Cytoscape, KEGG results were presented as a functional network, analyzing results from both test and validation sets to explore pathways associated with the two diseases.

In differential analysis of genes in the validation set, the intersection of key genes filtered by RF and Hub genes identified by `CytoHubba` provided genes most correlated with diseases, termed diagnostic genes or KeyGenes. The `limma` package was used to analyze differences between disease and normal groups in the validation set, with visualization using `boxplot`.

Using receiver operating characteristic (ROC) curves, to evaluate the diagnostic accuracy of key genes identified by machine-learning models and determine the optimal diagnostic threshold, the validation set was used as training data. The `pROC` package was employed to validate KeyGenes, calculating the area under the curve (AUC) and the 95% CI to predict the accuracy of distinguishing between OB and OP diseases using these KeyGenes. An AUC greater than 0.7 was considered to represent optimal diagnostic value.³³

Immune infiltration analysis

`CIBERSORT` analysis³⁴ is a tool designed to estimate the relative abundance of various cell types in mixed cell populations based on gene expression data. It is grounded in linear support vector regression, using an expression matrix of 22 immune cell subtypes as a reference to deconvolute the gene expression data of mixed cell populations.³⁴ We extracted the expression matrices of two diseases from the GS1 test set, and applied the `CIBERSORT` algorithm³⁵ for immune infiltration analysis to determine the relative abundance of immune cells. The correlation between the 22 immune cell types and KeyGenes was analyzed using the `limma` package, with visualization performed using `ggpubr` (Alboukadel Kassambara, France) and `ggExtra` (Dean Attali, Canada) packages.

Quality control and data processing in single-cell analysis

While bulk RNA-seq is unable to precisely pinpoint gene expression in specific cellular subpopulations, to reveal the cell clusters expressing TLR2 and for cellular experimental validation, we turned to single-cell analysis. We used R packages including `Seurat`, `tidyverse`, `dplyr`, and `patchwork`,³⁶ commonly used in scRNA-seq analysis. We constructed a `Seurat` object using the `CreateSeuratObject` function and conducted quality control through the `PercentageFeatureSet` function.³⁶ By employing the `logNormalize` function, we normalized RNA molecule counts for each cell, and then identified 3,000 highly variable genes using the `FindVariableFeatures` function and the 'vst' method.³⁶ We evaluated and corrected cell cycle effects using the `CaseMatch` and `CellCycleScoring` functions.³⁶ Principal component analysis (PCA) was performed, followed by cell clustering using the `Louvain` algorithm.³⁶ Finally, we used the `Harmony` package for batch effect correction and re-clustered and dimensionally reduced the data.³⁷

Annotation and visualization

Utilizing UMAP methodology and the Garnett package, we visualized cellular differences and performed annotations.³⁸ Data preprocessing involved normalization, scaling, and PCA, executed through the `preprocess_cds` function.³⁹ A cell classifier was trained using the marker file, classifying the dataset and updating the Seurat object. Finally, classification results and gene expression densities were displayed using the `DimPlot` function and the `plot_density` function from the `Nebulosa` package.⁴⁰

Construction of ceRNA network

In this research, we spotlighted TLR2 messenger RNA (mRNA), earmarked as the core of the ceRNA network. To predict miRNA and lncRNA targeting TLR2 mRNA, we used three tools: miRanda, miRDB, and TargetScan. Their interaction networks were visualized using Cytoscape software (v3.9.1).

Cell culture and high-fat model establishment

RAW 264.7 cells, obtained from the Cell Bank of the Chinese Academy of Sciences, were resuspended in complete cell culture medium (Dulbecco's Modified Eagle Medium (DMEM; HyClone, USA) supplemented with 10% fetal bovine serum (FBS; HyClone)). Cells were uniformly distributed in cell culture dishes (Thermo Fisher Scientific, USA) and maintained in an environment with controlled temperature (37°C), carbon dioxide levels, and humidity. After 24 to 48 hours, cell adherence was examined under an inverted microscope. Non-adherent cells were discarded, and the culture medium was refreshed every two to three days. Second-generation cells were employed for subsequent experiments.

For osteoclast differentiation, second-generation RAW264.7 cells were treated with osteoclastogenic induction medium containing 50 ng/ml RANKL combined with DMEM (procured from R&D Systems, USA).

Palmitic acid (PA; Sigma-Aldrich, Merck, USA) was dissolved in 200 mM ethanol and combined with 10% fatty acid-free low endotoxin bovine serum albumin (BSA) (Solarbio, China) to yield a final concentration of 5 mM. This solution was then diluted with DMEM to obtain a 500 μM working solution for subsequent adipogenic accumulation assays.

Western blotting analysis

Mononuclear macrophages were initially treated with RIPA lysis buffer (Beyotime Biotechnology Institute, China). The cells were then subjected to ultrasonication for complete lysis, and the lysate was incubated on ice for 30 minutes. Following centrifugation at 4°C, the supernatant was transferred to EP tubes for bicinchoninic acid (BCA) protein quantification assay (Generay, China). Proteins (15 μg) were subjected to gel electrophoresis and transferred onto polyvinylidene fluoride (PVDF) membranes. Membranes loaded with proteins of varied molecular weights were blocked in serum-free protein solution, followed by an overnight incubation with primary antibodies at 4°C. The following day, the membranes were incubated with secondary antibodies at 4°C for an hour. After thorough washing, the bands were visualized using the chemiluminescence (ECL) system (Analytik Jena, Germany) with β-actin (molecular weight: 43 kDa) serving as the normalization control. Band intensity and relative protein expression levels were quantified using ImageJ software

(National Institutes of Health, USA). For osteoclast differentiation, the same protocol was employed. The following antibodies were used: TLR2, matrix metalloproteinase-9 (MMP9), Cathepsin K (CTSK), and Arginase 1 (ARG1) (all procured from Abcam, USA).

Transfection of siRNAs

TLR2 small interfering RNA (siRNA) was synthesized by GenScript Technologies, and cells were transfected using Lipofectamine 2000. Prior to transfection, three different sequences were evaluated for their inhibitory efficiency, with TLR2-si-3 (78.3% efficiency) being selected. According to the manufacturer's protocol, a concentration of 50 nM siRNA with Lipofectamine 2000 transfection reagent was introduced to RAW264.7 cells. After 72 hours at 37°C, these cells, along with untreated controls, were employed for subsequent assays. The siRNA sequences are listed in [Table I](#).

Crystal violet staining

RAW264.7 cells were seeded in six-well plates at a density of 2×10^6 cells/well, and treated with either 500 μM PA or PA + siRNA for 24 hours. Cells were then fixed with 4% paraformaldehyde (Biosharp, China) for 15 minutes, washed thrice with PBS, and nuclei were stained using a crystal violet staining kit (Beyotime Biotechnology Institute). Observations were subsequently made under a light microscope (Nikon Eclipse Ti; Nikon, Japan).

Enzyme-linked immunosorbent assay

Cell culture supernatants from the control group, PA group, and PA + siRNA group were collected, centrifuged at 4°C for 15 minutes to remove cellular debris, and then stored at -80°C for subsequent assays. Pro-inflammatory cytokines TNF-α, IL-1β, and IL-6 were quantified using commercial enzyme-linked immunosorbent assay (ELISA) kits (Sigma-Aldrich) as per the manufacturer's instructions, with readings taken on a microplate reader (Bio-Rad, USA).

Cellular immunohistochemical analysis

RAW264.7 cells were seeded in 24-well plates at a density of 2×10^4 cells per well, and treated with either 500 μM PA or PA + siRNA for 24 hours. Cells were then fixed with 4% paraformaldehyde (Biosharp) for 15 minutes, permeabilized using 0.1% Triton X-100 (Beyotime Institute of Biotechnology) for 20 minutes, and rinsed thrice with PBS. Blocking was performed with 1% BSA in PBS for 30 minutes. Primary antibodies against inducible Nitric Oxide Synthase (iNOS) and TLR2 were diluted at a ratio of 1:200, and cells were incubated with these antibodies overnight at 4°C in the dark. After washing thrice with Tris-Buffered Saline with Tween 20 (TBST), cells were incubated for one hour in the dark with the specific fluorescence-labelled secondary antibody Alexa Fluor 594 (catalogue number ab150080; Abcam). After additional washes with TBST, nuclei were stained using DAPI (Beyotime Institute of Biotechnology) for ten minutes at room temperature. Observations were made under a fluorescence inverted microscope (Nikon Eclipse Ti; Nikon, Japan).

TRAP staining

To assess the differentiation extent of osteoclasts, RAW264.7 cells were seeded in a 24-well plate at a density of 2×10^4 cells/

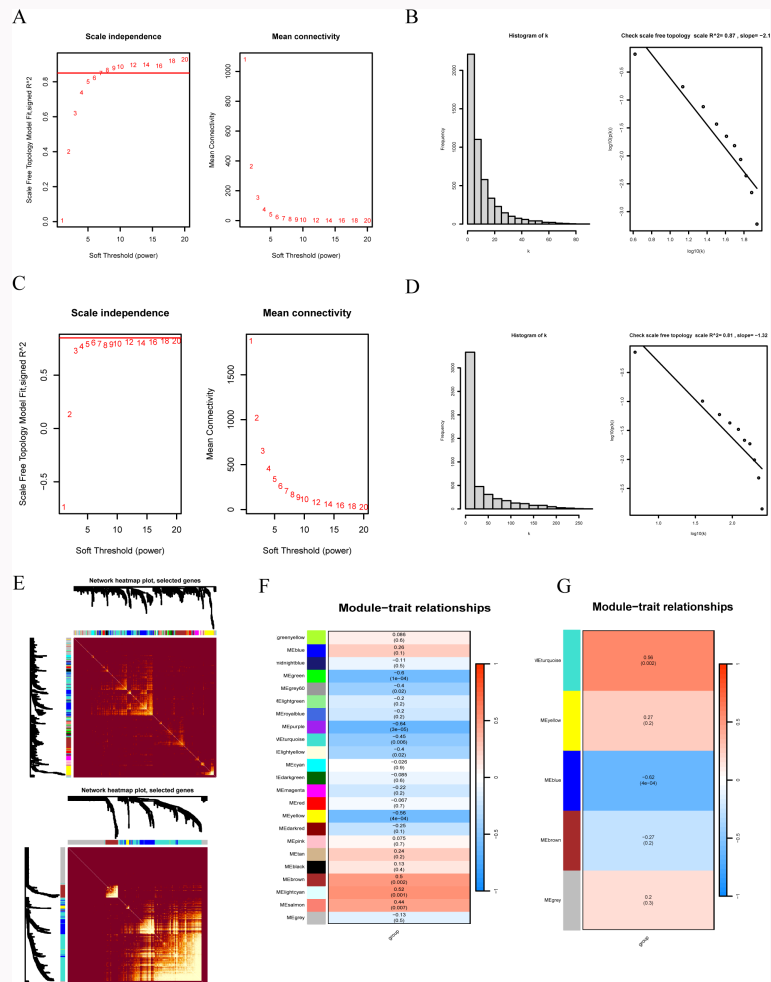


Fig. 2

Weighted gene co-expression network analysis (WGCNA) identified modules of genes associated with osteoporosis and obesity in the test set. a) Both graphs elucidate the scale-free properties of the gene co-expression network and changes in the mean node connectivity, opting for 7 as the soft-threshold value. b) Both graphs collectively showcase the verification of the scale-free nature of the gene co-expression network, where $R^2 = 0.87$ and the slope = -2.1 , aligning with the established criteria. c) Analogous to Figure 2a, albeit with different values; 20 is selected as the soft threshold. d) Mirroring Figure 2b, but with distinct values, $R^2 = 0.81$ and slope = -1.32 , both meeting the prerequisites. e) Topological overlap matrix (TOM) heatmap delineates the TOM of all genes examined. Paler hues signify low overlap, transitioning to deeper reds indicating increased overlap. Darkened blocks along the diagonal represent modules. The dendrogram of genes and module allocation is also presented on the left and top. The TOM plot facilitates a refined comprehension of inter-gene relationships. f) and g) The module clinical trait correlation heatmap, a visualization method within WGCNA analysis, manifests the associations between modules and clinical features. Within the heatmap, the depth of colour directly corresponds to the magnitude of correlation: red depicts positive correlation, while blue signifies inverse associations. Each cell enumerates the correlation and its statistical significance. All p-values were calculated with Pearson correlation coefficient.

Table 1. RNA interference sequences used for transfection.

Name	Forward (5'-3')	Reverse (5'-3')
tlr2-siRNA-1	UGGAGAAGGUGAAGCGAAUTT	AUUCGCUUCACCUUCUCCATT
tlr2-siRNA-2	GAGAUAGGAGAAUAGAUUTT	AAUCUAUUCUCCUAUCUUCTT
tlr2-siRNA-3	CAGCAGAAUCAUACAUAUTT	UAUUGUAUUGAUUCUGUGTT

siRNA, small interfering RNA; tlr, Toll-like receptor.

well and differentiated using 50 ng/ml RANKL in DMEM culture medium. Concurrently, supernatants previously collected from two groups were introduced to observe differentiation impact. The medium was refreshed every two days, and post six days, cells were fixed using 4% paraformaldehyde for 15 minutes.

Thereafter, tartrate-resistant acid phosphatase (TRAP) activity was gauged employing Leukocyte Acid Phosphatase Kits (Sigma-Aldrich). Cells bearing a nucleus count of three or more that tested positive for TRAP were identified as osteoclasts under a light microscope.

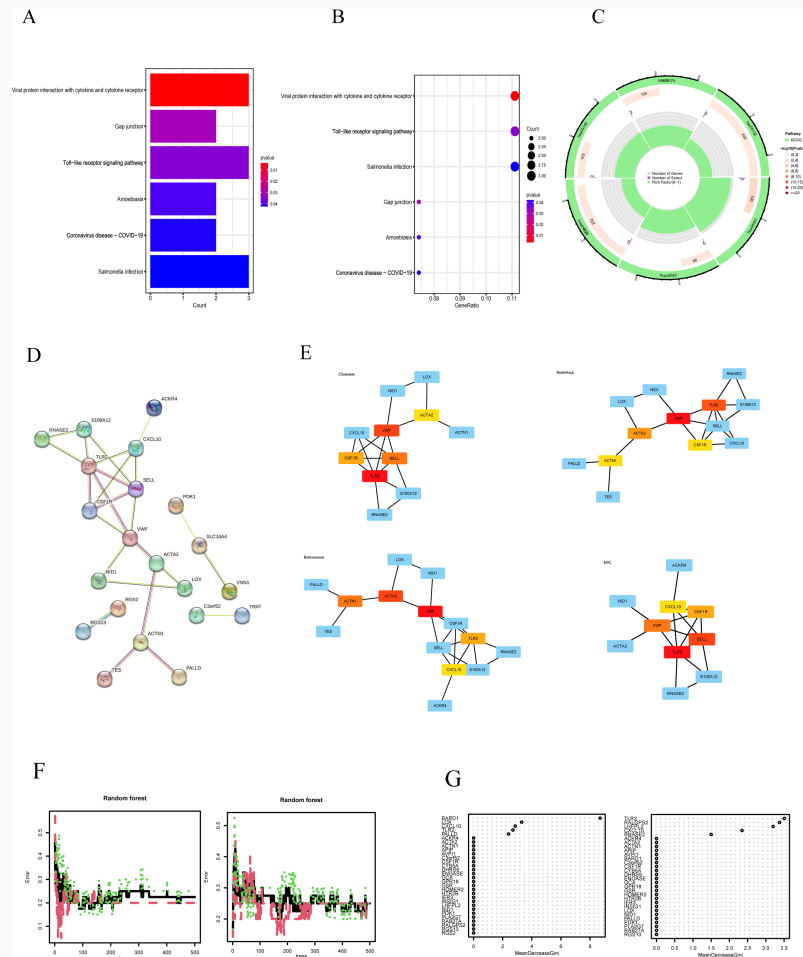


Fig. 3

Analysis of gene intersections within the test set module. a) to c) Bar charts are customarily employed to demonstrate the enrichment levels of each pathway, whereas bubble plots elucidate both the enrichment levels and the number of genes for each pathway. In these bubble plots, the colour of the bubble represents the p-value (or analogous statistical values), and the size signifies the number of overlapping genes (i.e. the genes provided that overlap with genes in a given pathway/Gene ontology (GO) term). Circular diagrams offer a vivid representation of the relationships among pathways. d) The protein-protein interaction (PPI) network diagram is depicted, where nodes symbolize proteins, and edges represent interactions between these proteins. Distinct colours and thicknesses of edges denote different sources of interaction. Sky-blue lines: known interactions from curated databases. Purple lines: known interactions from experimental studies. Blue lines: predicted interactions based on gene co-occurrence. Green lines: derived from the literature. e) The Hubgene network map, obtained using cytoHubba through diverse algorithms, assists in refining our comprehension of PPI dynamics and in unveiling pivotal biological processes and pathways. f) Random forest (RF) map of GSE151839 and GSE56814. g) RF graphs showing variable importance of GSE151839 and GSE56814. In Figure 3f, The RF map displays a model's error rate as the tree count increases, with the x-axis labelled "trees" (0 to 500) and the y-axis labelled "Error" (0 to 0.5). Three lines, each a different colour and style, show varying error trends: the black line represents the overall error rate, while the red and green lines represent the out-of-bag (OOB) error rates for the two factor levels of the group variable. Error rates tend to decrease as more trees are added. The RF graphs of variable importance in Figure 3g show the importance of variables in a RF, listing genes or variables on the y-axis and "MeanDecreaseGini" on the x-axis. This measure reflects the mean reduction in the model's Gini impurity caused by each variable, with larger values indicating greater importance for prediction in the model. All p-values were calculated with hypergeometric test. EPC, edge percolated component.

Statistical analysis

Data are presented as means and SDs. Statistical analysis was conducted using SPSS 22.0 (IBM, USA) on triplicate experiments. Independent-samples *t*-test, one-way analysis of variance (ANOVA), and Tukey's post hoc test were employed for mean comparison. The coefficient of determination (R^2) and Pearson correlation coefficient were used in the weighted gene co-expression network analysis (WGCNA) to assess the quality of the identified gene co-expression modules, with R^2 values closer to 1 indicating higher biological relevance. A hypergeometric test was used to assess the statistical significance of gene enrichment in specific pathways. In immune infiltration analysis and single-cell resolution studies,

all p-values were calculated using the permutation test. A p-value < 0.05 was deemed statistically significant.

Results

WGCNA analysis reveals co-expressed modules in OP and OB
Guided by scale independence and mean connectivity (Figures 2a to 2d), a soft-thresholding power of 8 ($R^2 = 0.87$) was selected for the GSE151839 dataset, while a power of 20 ($R^2 = 0.81$) was chosen for GSE56814. This revealed a higher number of gene clusters in the GSE151839 dataset (Figure 2e). Conversely, the majority of gene clusters in GSE56814 were concentrated within the turquoise, brown, and blue modules. Further examination of the module-clinical relationship heatmaps (Figures 2f and 2g) identified 23 modules in

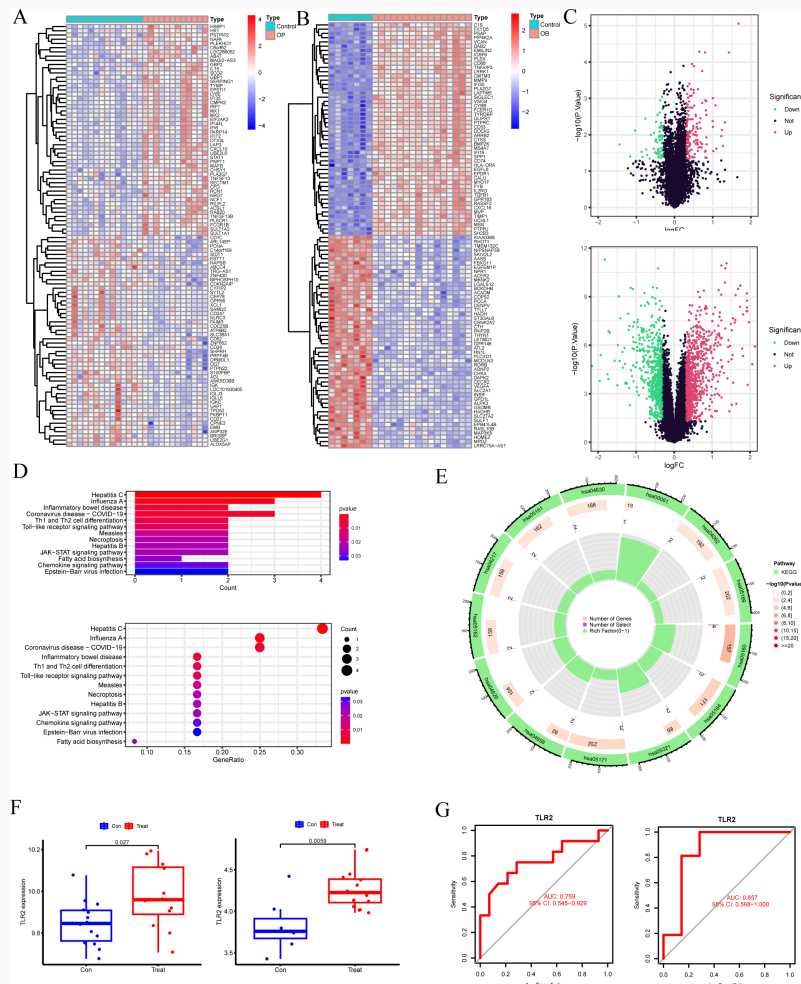


Fig. 4

Validation set: correlational analysis of differentially expressed genes (DEGs) associated with osteoporosis (OP) and obesity (OB). a) and b) Heatmap of DEGs, which visually displays the magnitude and clustering of gene expression levels. In the figure, a deeper colour denotes higher gene expression, while a lighter colour indicates lower expression. c) Volcano plot of DEGs combines statistical significance p-values with the fold change (logFC), facilitating a quick and intuitive identification of genes that are significantly altered and hold statistical relevance. d) and e) Bar graphs are commonly used to illustrate the enrichment level of each pathway, while bubble plots can showcase the enrichment levels alongside the number of genes. Within the bubble plot, the colour represents the p-value (or other metrics such as q-value), and the size signifies the number of genes, indicating the overlap between the submitted genes and those in a particular pathway/ Gene Ontology (GO) term. Circle diagrams offer a more vivid depiction of the relationships between pathways. f) Box plot elucidates high expression levels of the Toll-like receptor 2 (TLR2) gene in both the OP and OB groups in the validation set, underscoring its statistical significance in gene expression. g) Receiver operating characteristic (ROC) curve of a single gene serves to assess the sensitivity and specificity of that gene as a biomarker for survival. The TLR2 gene demonstrates promising performance in the validation set, as portrayed by its ROC curve. All p-values were calculated with independent-samples *t*-test.

GSE151839, with the light cyan and brown modules demonstrating positive correlations, bearing coefficients of 0.52 ($p = 0.001$) and 0.5 ($p = 0.002$, both Pearson correlation coefficient), encompassing 564 genes collectively. In GSE56814, five modules were discerned, with the turquoise module emerging as the sole positively correlated entity, exhibiting a correlation coefficient of 0.56 ($p = 0.002$) and comprising 1,832 genes.

Identify module intersection genes and conduct KEGG enrichment analysis

Intersecting positively correlated module genes from GSE151839 and GSE56814 yielded 50 overlapping genes, designated as GS1, which are thought to bear a significant association with OB and OP pathogenesis. This intersection elucidates the KEGG enrichment analysis for GS1 (Figures 3a to 3c), revealing its enrichment in pathways such as “Viral protein interaction with cytokine and cytokine receptor”, “Gap

junction”, and “Toll-like receptor signalling pathway”, with the TLR signalling pathway’s constituent genes suggesting its paramount importance in OB and OP.

Selection and validation of key genes

Utilizing the STRING database, we constructed a PPI network for GS1 to analyze protein functionality and biological system organization (Figure 3d), identifying potentially interactive nodal proteins. Four topological analyses were conducted on the GS1 PPI network with CytoHubba (Figure 3e), namely edge percolated component (EPC), BottleNeck, Closeness, and Betweenness, unveiling common genes such as SELL, TLR2, CXCL10, VWF, and CSF1R, all deemed pivotal nodes across the algorithms. Subsequently, leveraging machine learning, we targeted genes from a statistical standpoint. Initially, the expression matrix of GS1 in GSE151839 was extracted, and through a RF methodology, genes BARD1, LOX, CXCL10,

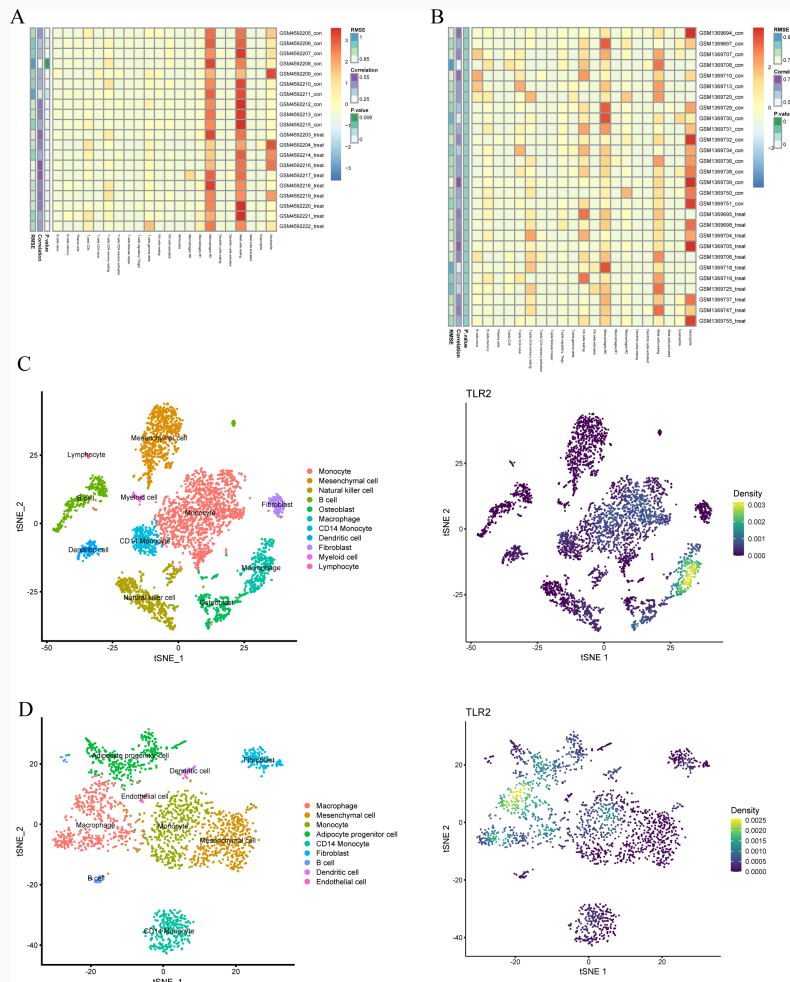


Fig. 5

Immune infiltration analysis and single-cell resolution studies. a) and b) Heatmaps delineating immune infiltration provide insights into the abundance of immunological cells. They provide a lucid visual representation of the density and clustering scenario of diverse immune cells within the tissues. In the heatmap, a darker hue signifies heightened immune cell abundance, whereas a lighter shade indicates reduced immune cell presence. c) t-distributed stochastic neighbor embedding (t-SNE) analysis of the osteoporosis (OP) single-cell RNA sequencing (scRNA-seq) data reveals 11 distinct cell subpopulations. Density plots highlight the pronounced expression of the Toll-like receptor 2 (TLR2) gene predominantly in macrophages. d) t-SNE examination of the OB scRNA-seq data uncovers nine unique cell subgroups. Density illustrations underscore the preeminent expression of the TLR2 gene mainly in macrophages. All p-values calculated with permutation test. RMSE, root mean square error.

TLR2, and PALLD emerged (seed = 226067). Analogously, for GSE56814, TLR2, RALGPS2, LHFPL2, and CXCL10 were discerned (seed = 849). An error analysis of results from both datasets was undertaken, with variable importance plots presented (Figures 3f and 3g). Utilizing 300 decision trees for GSE151839 and 200 for GSE56814, the mean Gini decrease value for the TLR2 gene conspicuously surpassed other genes, substantiating its preeminent correlation in both diseases.

Analyze shared genes using a validation set

Two datasets, GSE5520 and GSE7158, were selected as validation sets (Figures 4a to 4c). Differential analyses of these datasets revealed an intersecting gene set, GS2, containing 21 genes. Subsequent KEGG enrichment analysis of GS2 identified significant enrichment in pathways (Figures 4d and 4e) such as “Hepatitis C”, “Influenza A”, and notably, the “Toll-like receptor signaling pathway”, a pathway that was also enriched in GS1, reaffirming its association with the pathogenesis and progression of OP and OB. The gene TLR2, shared between CytoHubba and RF analyses, showed significant

differential expression between disease and control groups in both validation sets. Specifically, in the OP validation set from GSE7158 (Figure 4f), the expression of the TLR2 gene was higher in the disease group than in the control group ($p = 0.027$, independent-samples *t*-test). Similarly, in the OB validation set from GSE5520 (Figure 4f), TLR2 expression was again elevated in the disease group ($p = 0.006$, independent-samples *t*-test). ROC curves constructed with TLR2 as an input gene confirmed its differential expression in both OP and OB (Figure 4g). The AUC for the TLR2 ROC in the GSE7158 dataset for OP was 0.759 (95% CI 0.545 to 0.929), and for the GSE5520 dataset for OB (Figure 4g) it was 0.857 (95% CI 0.598 to 1.000), illustrating its high discriminatory potential for both conditions. Notably, TLR2 was also enriched in “Toll-like receptor signaling pathway”, emphasizing its pivotal role in both diseases.

High TLR2 expression in macrophages induces inflammation

Initial immune infiltration analysis was performed using CIBERSORT on the expression matrices of GSE151839 and

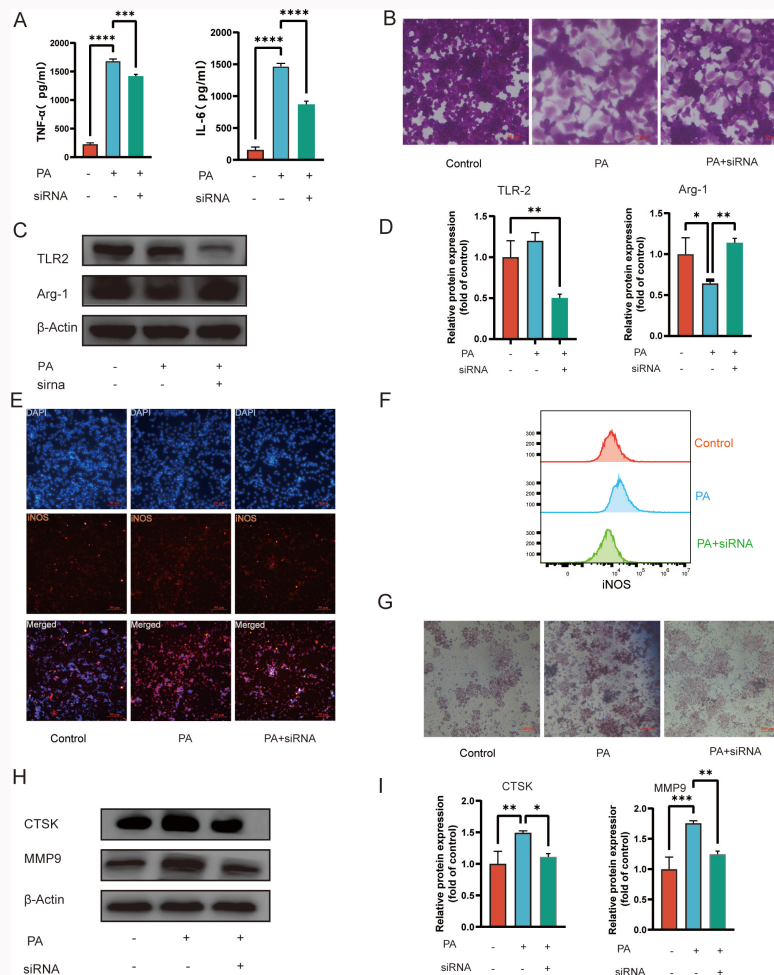


Fig. 6

Elevated expression of Toll-like receptor 2 (TLR2) in macrophages precipitates the onset of obesity and osteoporosis. a) Following transfection, the secretion levels of interleukin (IL)-6 and tumour necrosis factor- α (TNF- α) in the supernatant were quantified using enzyme-linked immunosorbent assay (ELISA). b) Alterations in macrophage morphology were observed before and after treatment with palmitic acid (PA) and small interfering RNA (siRNA)-TLR2. Produced with crystal violet staining, magnification 200 \times . c) At the protein level, evaluate the silencing efficiency of TLR2 and the expression levels of ARG1. d) The protein expression of TLR2/ARG1 compared with the control group was examined. e) After treating macrophages with PA and siRNA-TLR2, we observed that the localization and expression of iNOS was in a more aggregated state. f) Fluorescence intensity of iNOS was detected using flow cytometry. g) Upon staining cells with tartrate-resistant acid phosphatase (TRAP), the number of multinucleated cells increased in the PA-induced group. This change was reversed in the siRNA-TLR2 group. h) We employed Western blot analysis to detect the expression of proteins associated with osteoclast differentiation. i) The protein expression of Cathepsin K (CTSK)/matrix metalloproteinase-9 (MMP9) compared with the control group was examined. Data are shown as mean and SD. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$ compared with control cells, calculated with one-way analysis of variance. All representative data from three independent experiments are shown.

GSE56814 (Figure 5a). The GSE151839 dataset revealed relatively abundant immune cell types, including M2 macrophages, resting mast cells, and neutrophils. Following CIBERSORT analysis of the GSE56814 dataset and accounting for monocyte proportions, the dominant immune cell types were M0 macrophages, resting mast cells, and neutrophils (Figure 5b), possibly suggesting chronic inflammatory responses.

Subsequent single-cell analysis was undertaken to delineate specific cell subgroups. Employing dimensionality reduction and UMAP visualization techniques on single-cell datasets for OP and OB, 12 subgroups were identified in the OP dataset, including monocytes, mesenchymal cells, natural killer cells, B cells, osteoblasts, macrophages, CD14 monocytes, dendritic cells, fibroblasts, myeloid cells, and lymphocytes, with elevated TLR2 expression in macrophages (Figure 5c). In the OB dataset, nine cell subtypes were discerned, including

macrophages, mesenchymal cells, monocytes, adipocyte progenitor cells, CD14 monocytes, fibroblasts, B cells, dendritic cells, and endothelial cells, with the highest TLR2 expression observed in macrophages (Figure 5d). Based on these findings, it is postulated that heightened TLR2 expression may influence macrophage behaviour.

Macrophages in obesity induce M1 to M2 polarization and inflammatory response through high expression of TLR2

To investigate the role of the identified TLR2 gene in macrophage polarization and inflammation induction under a lipid accumulation environment, we incorporated 500 μ M PA into macrophage culture media to simulate the effects of lipid accumulation. Following 24-hour incubation with PA and post-gene knockdown macrophages, the supernatants from the aforementioned cultures, treated with both PA and siRNA, were harvested. Using ELISA kits, we measured the protein



Fig. 7 Competitive endogenous RNA (ceRNA) network diagram delineates the Toll-like receptor 2 (TLR2)-associated microRNAs (miRNAs) and long non-coding RNAs (lncRNAs).

levels of TNF- α and IL-6 in these supernatants. PA stimulation led to a significant increase in TNF- α and IL-6 protein levels in the RAW264.7 mouse macrophage culture media (Figure 6a). Following this, we employed crystal violet staining to observe macrophages post-PA treatment. We discerned an escalation in the count of irregularly shaped cells exhibiting protrusions and pseudopodia. However, in the siRNA-TLR2 group, a reversal in cellular morphological alterations was evident (Figure 6b). Moreover, the TLR2 knockdown samples showed lower levels compared to the non-knockdown groups. We then employed a Western blot assay to ascertain TLR2 protein levels and ARG1, a marker for M2 macrophages (Figures 6c and 6d). Remarkably, the TLR2 knockdown samples displayed elevated ARG1 expression compared to the group exposed solely to PA, while the iNOS levels were diminished. Utilizing a fluorescent inverted microscope to observe iNOS expression, we found the fluorescence intensity in the PA group to be notably augmented; however, the introduction of siRNA reversed this increase (Figures 6e and 6f). These findings provide substantial evidence that under lipid-accumulating conditions, TLR2 can direct macrophages towards M1 polarization, enhancing their inflammation-inducing potential.

Inflammation induced by macrophage polarization in osteoporosis alters osteoclast differentiation ability

To decipher the mechanism by which high expression of the TLR2 gene in macrophages influences the progression of OP, we proceeded to introduce supernatants from various groups into RAW264.7 cells and initiated differentiation with 50 ng/ml RANKL. Following differentiation, TRAP staining of the osteoclasts revealed that the TLR2 knockdown group could reverse the inflammation-induced promotional effect on osteoclast differentiation (Figure 6g). Furthermore, Western blot analysis revealed that the expression levels of CTSK

and MMP9 in the TLR2-knockdown treated supernatants were diminished, signifying reduced osteoclast differentiation capabilities (Figures 6h and 6i). Consequently, this affirms that elevated TLR2 expression within macrophages can augment inflammation-induced OP.

Construction of ceRNA network associated with TLR2

Within this network (Figure 7), TLR2 mRNA exhibited interactions with three distinct miRNAs, namely hsa-miR-561-3p, hsa-miR-654-3p, and hsa-miR-518d-5p. Conversely, we identified four long noncoding RNAs (lncRNAs) (RP11-231G3.1, CTC-273B12.5, CTC-548K16.6, and FAM230B) that shared interactions with at least one of these microRNAs (miRNAs). Specifically, RP11-231G3.1 was associated with hsa-miR-561-3p, both CTC-273B12.5 and CTC-548K16.6 with hsa-miR-518d-5p, and FAM230B with hsa-miR-654-3p. Ultimately, one mRNA, three miRNAs, and four lncRNAs were incorporated into the ceRNA network. This highlights the restricted control pathways of individual genes as well as their broad regulatory potential.

Discussion

To elucidate the intricate nexus between OB and OP, we conducted a WGCNA on selected datasets pertaining to both conditions. Following the identification of modular genes, a KEGG pathway enrichment analysis highlighted a significant association with the "Toll-like receptor signaling pathway". Previous literature denotes this pathway as pivotal in immunoregulation, playing a cardinal role in the host's pathogen recognition and immune response.⁴¹ Notably, its activation can amplify chronic inflammatory reactions which, when exacerbated, precipitate conditions such as OB and OP.^{42,43} Concurrently, this activated pathway can exacerbate insulin resistance, further potentiating OB.⁴⁴ Subsequent to this, leveraging the construction of a PPI network coupled with

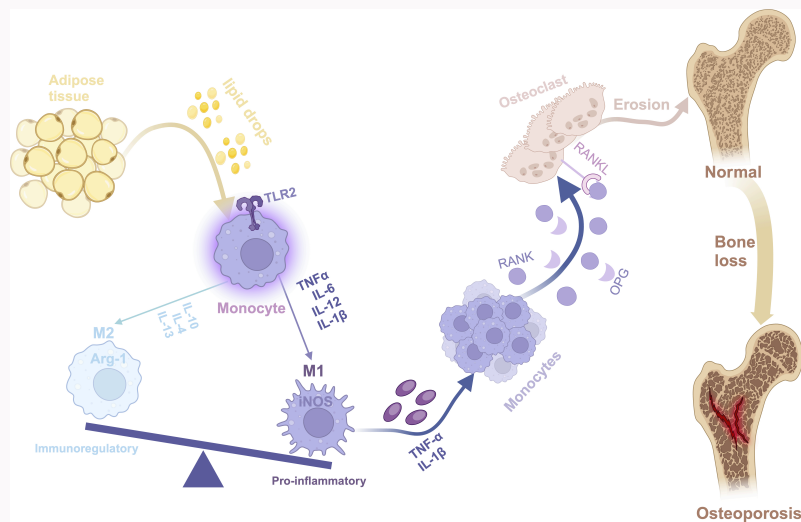


Fig. 8

Illustration of the signalling pathway. Arg-1, Arginase 1; IL-6, interleukin-6; iNOS, inducible nitric oxide synthase; OPG, osteoprotegerin; RANK, receptor activator of nuclear factor kappa B; RANKL, RANK ligand; TNF- α , tumour necrosis factor alpha.

machine learning-based filtering, we discerned two sets of genes – SELL, TLR2, CXCL10, VWF, CSF1R, and the combination of CXCL10 and TLR2 – from biological and statistical perspectives. Predominantly, these genes have been implicated in the activation and regulation of immune cells, chemotaxis, microbial recognition, and other inflammatory and immune response processes.^{45,46} After an exhaustive error analysis and correction, TLR2 emerged as the gene with the most significant correlation.

TLR2, a member of the TLR family, has been clinically demonstrated to overexpress in obese patients, contingent on MyD88-induced inflammation.⁴⁷ As a conserved pattern recognition receptor, TLR2 is instrumental in inflammatory responses and host defense against pathogenic invasions, recognizing pathogen-associated molecular patterns (PAMPs) and interacting with myriad endogenous agonists pertinent to the body's immune response. Upon external stimuli, TLR2 forms heterodimers with TLR family members, TLR1 or TLR6, co-activating downstream pathways and promoting inflammation.⁴⁸ It is ubiquitously expressed on monocytes, macrophages, B cells, and natural killer (NK) cells, consistent with our single-cell analysis findings. Increasing evidence suggests that TLR2 can recognize a variety of lipopeptides and participate in the transduction process of inflammatory signals.⁴⁹ Howe et al.⁵⁰ extracted macrophages from mice chronically fed with high saturated fatty acids (PA) and stimulated them with the TLR2 agonist lipoteichoic acid (LTA), resulting in a significant increase in the inflammatory marker IL-6 in the supernatant. Our research underscores that the activation of TLR2, triggered by lipid accumulation-induced free fatty acids and adipocyte-secreted inflammatory cytokines, induces a shift in the polarization state of monocyte macrophages. This transition from ARG1-expressing anti-inflammatory M2 macrophages to iNOS-expressing pro-inflammatory M1 macrophages also entailed morphological changes. Pre-existing literature has identified M1 macrophages as pro-inflammatory, secreting an abundance of inflammatory cytokines such as TNF- α and IL-6, perpetuating the body's inflammatory cascade. To affirm TLR2's role in this transformation, we conducted a knock-out

study, which ascertained that these changes were reversed upon its deletion, reinforcing our premise that the identified gene indeed mediates the progression of OB through inflammatory responses.

Subsequently, we delved into the implications for OP. The onset of OP is intrinsically linked to an augmented differentiation of osteoclasts, leading to increased bone resorption. Consequently, inflammation appears to be one of the pivotal factors influencing the differentiation of osteoclasts and, in turn, the emergence of OP. Following stimulation with RANKL, we treated precursor osteoclasts with distinct conditioned media and discerned that the differentiation capacity of osteoclasts treated with fatty acids increased. This inference is substantiated by the upregulated expression of differentiation-associated proteins, such as CTSK, and the proliferation of multinucleated osteoclasts, as evidenced by an escalated proportion of TRAP-stained positive cells. Concurrently, abrogation of TLR2 led to a diminished osteoclast differentiation, aligning with the findings of Kassem et al.⁵¹ Intriguingly, Madel's study revealed that the probiotic yeast *Saccharomyces cerevisiae* CNCM I-745 specifically targets and inhibits TLR2 and Dectin-1, diminishing the differentiation capacity of inflammatory osteoclasts (iOC), which further underscores the crucial role of TLR2 in promoting osteoclast differentiation and pathological bone loss. TLR4, closely associated with TLR2 and a member of the TLR family, has been demonstrated to promote osteoclast differentiation leading to OP via the MyD88-TRAF6-NF- κ B signalling pathway.^{52,53} In addition, the TLR2/NF- κ B signalling pathway has been demonstrated to be subject to modulation by the circadian regulatory gene BMAL1, which alters the polarization status of macrophages in mouse bone marrow, thereby promoting osteogenesis by reducing macrophage pyroptosis.⁵⁴ We therefore postulate that TLR2 acts as a key molecule in this cascade (Figure 8).

Notwithstanding this, the precise mechanism of TLR2 in inflammatory OP remains elusive. Our pathway analysis suggests that the influence of the immune system serves as the nexus between OP and OB conditions.

Zhang and Ni,⁵⁵ through serological assessments, identified a positive correlation between systemic immune indices and OP, which was particularly attributed to the surge in neutrophils and macrophages, culminating in this immune activation response. Our subsequent immune infiltration analysis pinpointed pivotal cells within the immune system, such as macrophages, mast cells, and neutrophils. Single-cell analysis further confirmed the overexpression of the TLR2 gene in macrophages, which correlates with the onset of both conditions, aligning with their research outcomes. Hence, we hypothesize that an augmented immune response coupled with inflammation is the central mechanism underpinning the interplay between these two diseases. We subsequently constructed an intricate ceRNA network centred around TLR2, identifying miRNAs such as hsa-miR-561-3p, hsa-miR-654-3p, and hsa-miR-518d-5p, as well as lncRNAs including RP11-231G3.1, CTC-273B12.5, CTC-548K16.6, and FAM230B. While the therapeutic potential of these non-coding RNAs in disease contexts remains largely uncharted, they offer promising avenues for pinpointing therapeutic targets and drug design for concurrent treatment of both conditions.

At the time of writing, the concurrent incidence of OP and OB remains prevalent. The primary clinical treatments for these conditions include surgical interventions, predominantly joint arthroplasty, and pharmacotherapy primarily involving bisphosphonates. However, challenges such as poor target specificity and substantial side effects have led to suboptimal patient compliance. Consequently, there is an urgent need to identify more effective treatment strategies and medications.^{56,57} Notably, vitamin D is recognized as an osteoprotective agent and a therapeutic for OP, with recent studies confirming its role in regulating insulin synthesis and metabolism, thereby influencing OB.⁵⁸ Additionally, the trace element selenium has been shown to mitigate OB by adjusting liver fat accumulation and the resultant iron homeostasis imbalance. A recent meta-analysis revealed a positive correlation between selenium intake and BMD in volunteers, while inversely correlating with the incidence of OP, suggesting selenium's potential as a therapeutic agent for OP.^{59,60} Based on these findings, we hypothesize that targeting a single drug could address the clinical challenges presented by these comorbidities. Our research has identified the key gene TLR2 and its induced macrophage inflammatory response, unveiling the potential molecular mechanisms increasing the comorbidity of OB and OP. This discovery provides potential diagnostic markers for both diseases, and guides the future design and development of targeted therapies.

In conclusion, in our study we introduced a novel approach by employing bioinformatics analysis combined with machine learning to identify the shared pivotal gene, *TLR2*, between OP and OB conditions. Concurrently, through immunoinfiltration and single-cell analyses, we elucidated that the elevated expression of TLR2 within macrophages serves as a critical mechanism underlying the onset of both diseases. Centering on TLR2, we constructed a ceRNA network, laying a theoretical foundation for future pharmacological designs targeting a singular point to concurrently treat both ailments. However, further investigations are warranted to unravel the downstream signalling pathways and molecular mechanisms. Additionally, our validation experiments were confined to

cellular assays, without delving into the pathogenic roles of the gene at an animal model level.

References

1. Williamson TK, Passfall L, Ihejirika-Lomedico R, et al. Assessing the influence of modifiable patient-related factors on complication rates after adult spinal deformity surgery. *Bone Joint J.* 2022;104-B(11):1249–1255.
2. Xiao PL, Cui AY, Hsu CJ, et al. Global, regional prevalence, and risk factors of osteoporosis according to the World Health Organization diagnostic criteria: a systematic review and meta-analysis. *Osteoporos Int.* 2022;33(10):2137–2153.
3. Abe S, Kashii M, Shimada T, et al. Relationship between distal radius fracture severity and 25-hydroxyvitamin-D level among perimenopausal and postmenopausal women. *Bone Jt Open.* 2022;3(3):261–267.
4. Zupan J, Jeras M, Marc J. Osteoimmunology and the influence of pro-inflammatory cytokines on osteoclasts. *Biochem Med (Zagreb).* 2013; 23(1):43–63.
5. Maeda K, Takahashi N, Kobayashi Y. Roles of Wnt signals in bone resorption during physiological and pathological states. *J Mol Med (Berl).* 2013;91(1):15–23.
6. Braun T, Schett G. Pathways for bone loss in inflammatory disease. *Curr Osteoporos Rep.* 2012;10(2):101–108.
7. Mu L, Liu J, Zhou G, et al. Obesity prevalence and risks among Chinese adults: findings from the China PEACE Million Persons Project, 2014–2018. *Circ Cardiovasc Qual Outcomes.* 2021;14(6):e007292.
8. Kawai T, Autieri MV, Scalia R. Adipose tissue inflammation and metabolic dysfunction in obesity. *Am J Physiol Cell Physiol.* 2021;320(3): C375–C391.
9. Murray PJ. Macrophage polarization. *Annu Rev Physiol.* 2017;79:541–566.
10. Cohen A, Dempster DW, Recker RR, et al. Abdominal fat is associated with lower bone formation and inferior bone quality in healthy premenopausal women: a transiliac bone biopsy study. *J Clin Endocrinol Metab.* 2013;98(6):2562–2572.
11. Lu L, Tang M, Li J, et al. Gut microbiota and serum metabolic signatures of high-fat-induced bone loss in mice. *Front Cell Infect Microbiol.* 2021;11: 788576.
12. Dai X, Liu B, Hou Q, et al. Global and local fat effects on bone mass and quality in obesity. *Bone Joint Res.* 2023;12(9):580–589.
13. Goulding A, Grant AM, Williams SM. Bone and body composition of children and adolescents with repeated forearm fractures. *J Bone Miner Res.* 2005;20(12):2090–2096.
14. Fintini D, Cianfarani S, Cofini M, et al. The bones of children with obesity. *Front Endocrinol (Lausanne).* 2020;11:200.
15. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics.* 2008;24(13):1547–1548.
16. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics.* 2004;20(3):307–315.
17. Kauffmann A, Gentleman R, Huber W. arrayQualityMetrics—a bioconductor package for quality assessment of microarray data. *Bioinformatics.* 2009;25(3):415–416.
18. Lim WK, Wang K, Lefebvre C, Califano A. Comparative analysis of microarray normalization procedures: effects on reverse engineering gene networks. *Bioinformatics.* 2007;23(13):i282–8.
19. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics.* 2012;28(6):882–883.
20. Zhang Y, Parmigiani G, Johnson WE. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genom Bioinform.* 2020; 2(3):lqaa078.
21. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.
22. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5): 284–287.
23. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 2023;51(D1):D587–D592.

24. Xu Z, Yu Z, Li S, Tian Z, Yuan J, You F. Exploration of the core gene signatures and mechanisms between NAFLD and sarcopenia through transcriptomic level. *Front Endocrinol (Lausanne)*. 2023;14:1140804.
25. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32(18):2847–2849.
26. Bindea G, Mlecnik B, Hackl H, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*. 2009;25(8):1091–1093.
27. No authors listed. STRING: functional protein association networks. 2023. <https://cn.string-db.org/> (date last accessed 31 July 2024).
28. Kong X, Wang C, Wu Q, et al. Screening and identification of key biomarkers of depression using bioinformatics. *Sci Rep*. 2023;13(1):4180.
29. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014;8 Suppl 4(Suppl 4):S11.
30. Sang L, Wang XM, Xu DY, Zhao WJ. Bioinformatics analysis of aberrantly methylated-differentially expressed genes and pathways in hepatocellular carcinoma. *World J Gastroenterol*. 2018;24(24):2605–2616.
31. Simon SM, Glaum P, Valdovinos FS. Interpreting random forest analysis of ecological models to move from prediction to explanation. *Sci Rep*. 2023;13(1):3881.
32. Mogensen UB, Ishwaran H, Gerds TA. Evaluating random forests for survival analysis using prediction error curves. *J Stat Softw*. 2012;50(11):1–23.
33. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77.
34. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–457.
35. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol*. 2018;1711:243–259.
36. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573–3587.e29.
37. Korsunsky I, Millard N, Fan J, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods*. 2019;16(12):1289–1296.
38. Pliner HA, Shendure J, Trapnell C. Supervised classification enables rapid annotation of cell atlases. *Nat Methods*. 2019;16(10):983–986.
39. Qiu X, Hill A, Packer J, Lin D, Ma Y-A, Trapnell C. Single-cell mRNA quantification and differential analysis with Census. *Nat Methods*. 2017;14(3):309–315.
40. Alquicira-Hernandez J, Powell JE. Nebulosa recovers single-cell gene expression signals by kernel density estimation. *Bioinformatics*. 2021;37(16):2485–2487.
41. Lim K-H, Staudt LM. Toll-like receptor signaling. *Cold Spring Harb Perspect Biol*. 2013;5(1):a011247.
42. AlQranei MS, Senbanjo LT, Aljohani H, Hamza T, Chellaiah MA. Lipopolysaccharide- TLR-4 axis regulates osteoclastogenesis independent of RANKL/RANK signaling. *BMC Immunol*. 2021;22(1):23.
43. Jialal I, Kaur H, Devaraj S. Toll-like receptor status in obesity and metabolic syndrome: a translational perspective. *J Clin Endocrinol Metab*. 2014;99(1):39–48.
44. Kan B, Zhao Q, Wang L, Xue S, Cai H, Yang S. Association between lipid biomarkers and osteoporosis: a cross-sectional study. *BMC Musculoskelet Disord*. 2021;22(1):759.
45. Arabpour M, Saghazadeh A, Rezaei N. Anti-inflammatory and M2 macrophage polarization-promoting effect of mesenchymal stem cell-derived exosomes. *Int Immunopharmacol*. 2021;97:107823.
46. Xiang C, Li H, Tang W. Targeting CSF-1R represents an effective strategy in modulating inflammatory diseases. *Pharmacol Res*. 2023;187:106566.
47. Guerrero-Romero F, Castellanos-Juárez FX, Salas-Pacheco JM, Morales-Gurrola FG, Salas-Leal AC, Simental-Mendía LE. Association between the expression of TLR4, TLR2, and MyD88 with low-grade chronic inflammation in individuals with metabolically healthy obesity. *Mol Biol Rep*. 2023;50(5):4723–4728.
48. Colleselli K, Stierschneider A, Wiesner C. An update on toll-like receptor 2, its function and dimerization in pro- and anti-inflammatory processes. *Int J Mol Sci*. 2023;24(15):12464.
49. Cao D, Wang W, Li S, et al. TLR2-deficiency promotes prenatal LPS exposure-induced offspring hyperlipidemia. *Front Physiol*. 2019;10:1102.
50. Howe AM, Burke S, O'Reilly ME, McGillicuddy FC, Costello DA. Palmitic acid and oleic acid differently modulate TLR2-mediated inflammatory responses in microglia and macrophages. *Mol Neurobiol*. 2022;59(4):2348–2362.
51. Kassem A, Lindholm C, Lerner UH. Toll-like receptor 2 stimulation of osteoblasts mediates Staphylococcus aureus induced bone resorption and osteoclastogenesis through enhanced RANKL. *PLoS One*. 2016;11(6):e0156708.
52. Madel MB, Halper J, Ibáñez L, et al. Specific targeting of inflammatory osteoclastogenesis by the probiotic yeast *S. boulardii* CNCM I-745 reduces bone loss in osteoporosis. *Elife*. 2023;12:e82037.
53. Lu Q, Jiang C, Hou J, et al. Patchouli alcohol modulates the pregnancy X receptor/toll-like receptor 4/nuclear factor kappa B axis to suppress osteoclastogenesis. *Front Pharmacol*. 2021;12:684976.
54. Chen L, Yu C, Xu W, et al. Dual-targeted nanodiscs revealing the cross-talk between osteogenic differentiation of mesenchymal stem cells and macrophages. *ACS Nano*. 2023;17(3):3153–3167.
55. Zhang S, Ni W. High systemic immune-inflammation index is relevant to osteoporosis among middle-aged and older people: a cross-sectional study. *Immun Inflamm Dis*. 2023;11(8):e992.
56. Bukowski BR, Sandhu KP, Bernatz JT, et al. CT required to perform robotic-assisted total hip arthroplasty can identify previously undiagnosed osteoporosis and guide femoral fixation strategy. *Bone Joint J*. 2023;105-B(3):254–260.
57. Watanabe N, Miyatake K, Takada R, et al. The prevalence and treatment of osteoporosis in patients undergoing total hip arthroplasty and the levels of biochemical markers of bone turnover. *Bone Joint Res*. 2022;11(12):873–880.
58. Cândido FG, Bressan J. Vitamin D: link between osteoporosis, obesity, and diabetes? *Int J Mol Sci*. 2014;15(4):6569–6591.
59. Xie H, Wang N, He H, et al. The association between selenium and bone health: a meta-analysis. *Bone Joint Res*. 2023;12(7):423–432.
60. Liu G, Li J, Pang B, et al. Potential role of selenium in alleviating obesity-related iron dyshomeostasis. *Crit Rev Food Sci Nutr*. 2023;63(29):10032–10046.

Author information

D. Liu, PhD, Orthopedic Researcher
 F. Cao, PhD, Orthopedic Researcher
 D. Liu, PhD, Orthopedic Researcher
 H. Li, MSc, Orthopedic Researcher
 L. Tao, PhD, Professor of Orthopedics
 Y. Zhu, PhD, Professor of Orthopedics
 Department of Orthopedics, The First Hospital of China Medical University, Shenyang, China.

Author contributions

D. Liu: Data curation, Formal analysis, Investigation, Software, Writing – original draft.
 F. Cao: Formal analysis, Software, Writing – original draft.
 D. Liu: Conceptualization, Validation.

H. Li: Data curation, Investigation.

L. Tao: Conceptualization, Resources, Writing – review & editing.
 Y. Zhu: Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding statement

The authors disclose receipt of the following financial or material support for the research, authorship, and/or publication of this article: funding from the National Natural Science Foundation of China (32200943, 82072387) and the Shenyang Young and Middle-aged Innovative Talent Project (RC210171).

ICMJE COI statement

All authors report funding from the National Natural Science Foundation of China (32200943, 82072387) and the Shenyang

Young and Middle-aged Innovative Talent Project (RC210171), related to this study. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data sharing

The data that support the findings for this study are available to other researchers from the corresponding author upon reasonable request.

Open access funding

The authors report that the open access funding for their manuscript was self-funded.

© 2024 Liu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (CC BY-NC-ND 4.0) licence, which permits the copying and redistribution of the work only, and provided the original author and source are credited. See <https://creativecommons.org/licenses/by-nc-nd/4.0/>