

Network-based Transcriptome-wide Expression Study for Postmenopausal Osteoporosis

Lan Zhang,¹ Tian-Liu Peng,² Le Wang,² Xiang-He Meng,³ Wei Zhu,¹ Yong Zeng,¹ Jia-Qiang Zhu,⁴ Yu Zhou,¹ Hong-Mei Xiao,² and Hong-Wen Deng¹

¹Center for Biomedical informatics and Genomics, Department of Medicine, Tulane University, New Orleans, Louisiana 70112; ²Institute of Reproduction and Stem Cell Engineering, School of Basic Medical Science, Central South University, Changsha, Hunan, 410013, China; ³Laboratory of Molecular and Statistical Genetics, College of Life Sciences, Hunan Normal University, Changsha, Hunan, 410081, China; and ⁴Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109

ORCID number: 0000-0002-0387-8818 (H.-W. Deng).

Purpose: Menopause is a crucial physiological transition during a woman's life, and it occurs with growing risks of health issues like osteoporosis. To identify postmenopausal osteoporosis-related genes, we performed transcriptome-wide expression analyses for human peripheral blood monocytes (PBMs) using Affymetrix 1.0 ST arrays in 40 Caucasian postmenopausal women with discordant bone mineral density (BMD) levels.

Methods: We performed multiscale embedded gene coexpression network analysis (MEGENA) to study functionally orchestrating clusters of differentially expressed genes in the form of functional networks. Gene sets net correlations analysis (GSNCA) was applied to assess how the coexpression structure of a predefined gene set differs in high and low BMD groups. Bayesian network (BN) analysis was used to identify important regulation patterns between potential risk genes for osteoporosis. A small interfering ribonucleic acid (siRNA)-based gene silencing in vitro experiment was performed to validate the findings from BN analysis.

Result: MEGENA showed that the "T cell receptor signaling pathway" and the "osteoclast differentiation pathway" were significantly enriched in the identified compact network, which is significantly correlated with BMD variation. GSNCA revealed that the coexpression structure of the "Signaling by TGF-beta receptor complex pathway" is significantly different between the 2 BMD discordant groups; the hub genes in the postmenopausal low and high BMD group are *FURIN* and *SMAD3* respectively. With siRNA in vitro experiments, we confirmed the regulation relationship of *TGFBR2-SMAD7* and *TGFBR1-SMURF2*.

Main Conclusion: The present study suggests that biological signals involved in monocyte recruitment, monocyte/macrophage lineage development, osteoclast formation, and osteoclast differentiation might function together in PBMs that contribute to the pathogenesis of postmenopausal osteoporosis. (*J Clin Endocrinol Metab* 105: 2678–2691, 2020)

Freeform/Key Words: postmenopausal osteoporosis, coexpression analysis, osteoclast, transcriptomics, siRNA

Abbreviations: BMD, bone mineral density; BN, Bayesian network; cDNA, complementary deoxyribonucleic acid; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; GEFOS, Genetic Factors for Osteoporosis Consortium; GO, Gene Ontology; GSNCA, gene sets net correlations analysis; GWAS, genome-wide association study; IL, interleukin; MCA, multiscale clustering analysis; M-CSF, macrophage colony-stimulating factor; MEGENA, multiscale embedded gene coexpression network analysis; MST, minimum spanning tree; PBM, peripheral blood monocyte; PFN, planar filtered network; siRNA, small interfering ribonucleic acid; SNP, single-nucleotide polymorphism; TCR, T Cell Receptor; TGF, transforming growth factor

An imbalance process between bone resorption (by osteoclasts) and bone formation (by osteoblasts) results in osteoporosis, which is a major health problem in postmenopausal women (1). It is well recognized that during the menopausal transition status, depletion of estrogen greatly affects osteoclastic bone resorption and the bone loss process (2, 3). Circulating monocytes can differentiate into a variety of cells (4), and mature peripheral blood monocytes (PBM) can be further categorized into 3 subpopulations in accordance with the surface expression of CD14/CD16 (5, 6). The classic subset of PBM (high level of CD14 and lack of CD16) can migrate to bone resorption sites, thus serving as latent osteoclast precursors. PBM play important roles in osteoclastogenesis by acting as osteoclast precursors and secreting osteoclastogenic factors (7–9). Therefore, PBM represent a class of cells that are of high functional relevance to osteoporosis development, and have been accepted for pathophysiology studies in the bone field (10).

Transcriptome-wide expression studies are a powerful tool for researchers to comprehensively perceive and quantify complex biological systems. Microarray studies have been generally applied in osteoporosis research in the identification of biomarkers crucial for osteoclasts and of functional pathways having profound effects for bone remodeling (11, 12).

Network-based analyses through expression data have become useful tools for systems biology. Multiscale embedded gene coexpression network analysis (MEGENA) is a method designed for constructing and analyzing large-scale coexpression networks (13). MEGENA first applies the network embedding technique on the topological sphere to construct unweighted coexpression networks and then identifies multiscale organizations of coherently coexpressed gene units (14, 15). The application of this method helped to reveal meaningful multiscale organizations of coexpressed gene clusters, and showed improved performance in identifying more coherent modules over well-established de novo network construction approaches such as weighted gene coexpression network analyses (13). Gene sets net correlations analysis (GSNCA) is a multivariate differential coexpression test for gene sets and it can detect significant changes in the coexpression structure of gene groups between 2 different biological conditions (16). Instead of considering differences in pairwise correlations between genes, the GSNCA method accounts for the global correlation structure of gene sets, and it works well on microarray datasets (17). As a probabilistic graphical model, Bayesian networks (BNs) has been applied in genetic related studies (18). Through a graphical separation process, the main role of BNs is to express conditional dependencies among variables.

Recently, a BN generated from gene coexpression modules was used to uncover regulatory driver genes affecting coronary artery disease (19), and it helped to recognize the causal network structures relevant to late-onset Alzheimer's disease as well (20). The directional relationships between nodes built by BNs could further help us to explore gene expression patterns from a mechanistic point of view.

In this study, we proposed a network-based platform for transcriptome-wide expression study of PBM from 40 postmenopausal women with high and low bone mineral density (BMD). The first goal was to identify individual genes contributing to postmenopausal osteoporosis pathogenesis. By employing advanced network analyses strategies, the second aim was to detect highly correlated gene groups (cofunctioning genes) and gene sets important for postmenopausal BMD variation. The innovative part of our study was that with a series of bioinformatics analyses, we not only found genes and pathways important for postmenopausal osteoporosis, but also validated regulation patterns between crucial gene pairs in PBM based on in vitro experiments. In addition, we applied small interfering ribonucleic acid (siRNA)-based gene silencing in vitro experiments to support regulatory relationships found by causal network inference.

Materials and Methods

Subjects for the microarray study

The study was approved by the Institutional Review Board, and signed consent documents were obtained from all study participants before being enrolled in the study. The detailed recruitment criteria and BMD measurement procedures have introduced in our previous studies (21). Hip BMD (g/cm^2) was measured using a Hologic dual-energy X-ray absorptiometer scanner (Hologic Corp., Waltham, MA). None of our subjects had any known conditions that might artificially increase BMD values, such as osteophytes and facet sclerosis. The obtained BMD value was transformed into a Z-score, which is the unit of standard deviation above or below the mean of a healthy, ethnic-, gender-, and age-matched reference population. Our discovery cohort includes 40 Caucasian postmenopausal females (GSE56815), 20 with high BMD and 20 with low BMD, and the 2 groups were matched for age and years of postmenopausal status. The workflow of this study is shown in Fig. 1.

Experiment procedures

PBM isolation, RNA extraction, and array procedure. As described in previous studies by our group, PBMs were isolated from whole blood samples with a monocyte-negative isolation kit (Miltenyi Biotec Inc., Auburn, CA) (22) that best represented the in vivo expression profiles in human subjects. The Qiagen RNeasy Mini kit (Qiagen, Inc., Valencia, CA) was used to extract total RNA from monocytes, and the Agilent

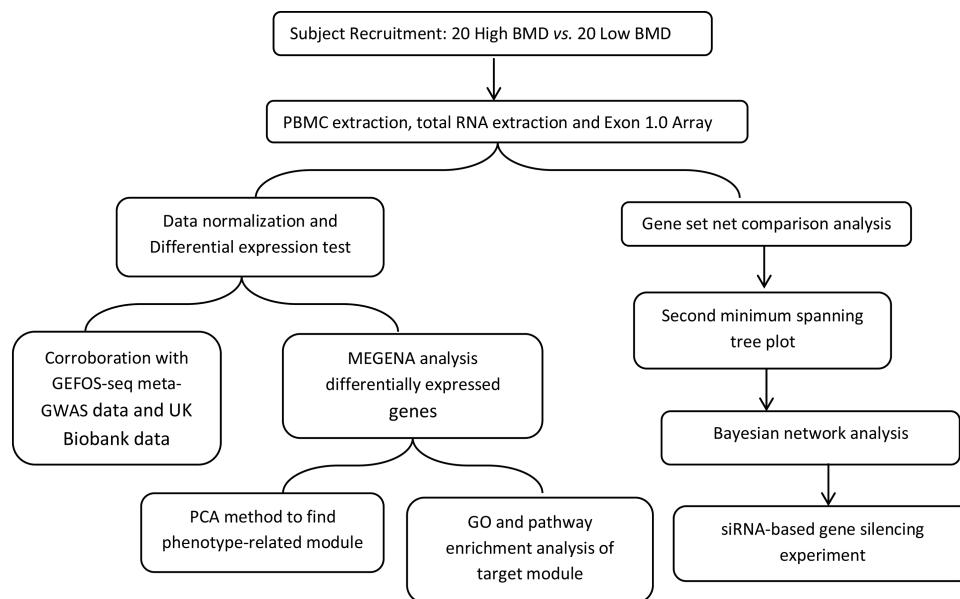


Figure 1. The workflow of the current study.

Bioanalyzer (Agilent, Santa Clara, CA) was applied to control RNA quality (23, 24). Preparation of complementary deoxyribonucleic acid (cDNA), hybridization, and scanning of the GeneChip Human Exon 1.0 ST array was performed based on the manufacturer's protocol for the Exon 1.0 array (Affymetrix, Santa Clara, CA).

Differential expression analysis

The robust multiarray average method (25) was applied to normalize the array signals through the Bioconductor Oligo package (26). After data normalization, we performed differential expression analysis through the Bioconductor LIMMA (linear models for microarray data) package (27).

To gain additional support for the differentially expressed genes in our discovery study, we checked the single-nucleotide polymorphisms (SNPs) of the top genes for their association signals in the GEFOS-seq (Genetic Factors for Osteoporosis Consortium) genome-wide association study (GWAS) datasets and UK Biobank GWAS datasets. In the bone field, GEFOS has developed consecutive meta-analyses of GWAS with the largest sample sizes, aiming to identify genetic variants related to BMD (28–30). The GEFOS-seq study meta-analyzed association results for BMD from 9 discovery cohorts ($n = 32\,965$), including a whole-genome sequencing in the UK10K project ($n = 2882$), a whole-exome sequencing study with 5 cohorts ($n = 3549$), and genotyped samples with deeply imputation using a combined UK10K/1000 Genomes reference panel ($n = 26\,534$) (30). Since GEFOS-seq did not use hip BMD as a phenotype, we checked the genetic association between SNPs with femoral neck BMD (FN-BMD) and lumbar spine BMD. The UK Biobank osteoporosis dataset contains estimated BMD (eBMD) GWAS summary statistics from 426 824 individuals. In this study, BMD was measured by quantitative ultrasound of the heel (a noninvasive eBMD that predicts fracture) (31). With stringent quality control of both eBMD and genome-wide genotypes, the study identified 518 genome-wide significant loci, explaining 20% of its variance (32). Since the UK Biobank osteoporosis study is the representative GWAS study that measures association between SNPs (and thus their

annotated genes) with eBMD, it could provide further partial support for our identified differential expressed genes.

MEGENA (with differential expressed genes)

MEGENA is an innovative coexpression network analysis method that can effectively and efficiently construct large-scale coexpression planar filtered networks (13). Parallelization of embedded network construction and multiscale clustering structure identification are the 2 major components of this method. It may uncover biologically meaningful gene orchestrating groups and key regulators of important biological processes. In the current study, we applied this method to construct the coexpression pattern among differential expressed genes.

Fast planar filtered network construction. The first step of MEGENA analysis is fast planar filtered network (PFN) construction, in which respective interaction strength (eg, Pearson's correlation) was used to rank gene pairs with greater similarity. Then the planarity was iteratively tested to build the embedded network including gene pairs with most similarities (33). Insignificant interactions were removed with $FDR > 0.05$.

Multiscale clustering analysis. The constructed PFNs were input to multiscale clustering analysis (MCA) to achieve the subsequent analyses. Three criteria, shortest path distances, local path index, and overall modularity, were incorporated in the MCA to identify locally compact clusters (34, 35). A hierarchical divisive approach is adopted by MCA to split complex interactions from PFN into more coherent gene groups through iterating processes (13).

Downstream analysis. Multiscale hub analysis was applied to identify significant hubs within each cluster (14, 35). To relate an established modular structure to phenotypes in our dataset, we calculated Pearson's correlation between eigengene (first principal component of cluster) and clinical traits; modules with

$P < .05$ were taken as phenotype-related modules. Interacting patterns for all members in the target module were searched through STRING version 9.1 (36). The GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis of compact gene clusters identified by MEGENA was performed with STRING software as well. The Bonferroni-adjusted P value ($<.05$) was used as a threshold for identifying significant functional clusters.

Gene sets net correlations analysis

GSNCA is a derivative method from the original Gene Sets Comparison Analysis. To address the limitation that traditional gene set comparison methods do not take into account the structure of the gene coexpression pattern, Rahmatallah et al. proposed the GSNCA method, which is based on the impression that biological systems tend to be more affected by changes in the activities of “high weight” genes (16). The GSNCA method introduces a weight (w_i) to each gene based on the gene’s cross-correlation with all the other genes to quantify their corresponding regulatory importance, and compares the normalized weight vectors for gene groups between 2 conditions. In this study, gene sets were taken from molecular signature database (MSigDB) C2 curated gene sets (37, 38), including 3272 metabolic, biochemical, and signaling pathways from the Reactome, BioCarta, and KEGG databases. The GSNCA analysis was implemented through the “GSAR” R Package. The detailed network features (average absolute cross-correlation, degree centrality of hub gene, normalized betweenness centrality of hub gene, average normalized betweenness centrality, normalized closeness centrality of hub gene, average normalized closeness centrality) were calculated for coexpression gene sets as well. MST2 (first and second minimum spanning trees) method was applied to present the coexpression network features in the low versus high BMD groups. It is believed that the union of the first and second MST, constructed from correlation distances, generates the minimal set of essential connections among genes (16).

Bayesian network analysis

BNs are graphical based models where nodes (vertex) represent random variables and the corresponding edges (arcs) represent probabilistic dependencies (18). BNs have been applied in numerous kinds of biological data like SNP data and gene expression profiles (39). In this study, we applied BN analysis to differentially coexpressed gene sets identified by the previous GSNCA. The gene expression data in the identified gene sets were discretized using Hartemink’s method (40), which was designed to preserve most pairwise dependencies. Then we performed constraint-based structure learning on the discretized data (39). After repeating the structure-learning procedure 100 times, gene network structures learned in the previous stage were averaged to construct a more optimal one using the model-averaging approach (41). Arcs present in at least 80% of the networks were used to generate the averaged network structure (18). The bootstrap resampling approach was performed to learn a set of 500 network structures for model averaging.

siRNA validation experiment for the identified *TGFBR2-SMAD7* and *TGFBR1-SMURF2* regulations

To provide general support for our approach and results, we chose to validate the causal regulatory relationships of

TGFBR2-SMAD7 and *TGFBR1-SMURF2* identified in the BN analysis within subjects with low BMD, by performing siRNA-based gene silencing in vitro experiments.

Cell culture. Since THP-1 (Human acute monocytic leukemia cell line) cells represent a valuable tool for investigating monocyte structure and function in both health and disease conditions (42), we used it in our experiment. THP-1 cells (Origin: Cell Biology Research Laboratory Modern Analytical Testing center, Central South University, Changsha, China) were cultured in RPMI 1640 medium (Gibco BRL, Rockville, MD) containing 10% heat-inactivated fetal bovine serum (Gibco, Thermo Fisher Scientific, Waltham, MA) and antibiotics (100 U/mL penicillin G and 100 µg/mL streptomycin) under a humidified atmosphere (37°C, 5% CO₂).

Transfection. siRNAs were generated against human TGFBR1 (sense strand 5'-GAACAGAAAGTTAAGGCCAA-3' and antisense strand 5'-UUGGCCUUACUUCUGUUC-3'), and TGFBR2 (sense strand 5'-CCATCATCCTGGAGATGA-3' and antisense strand 5'-UCAUCUCCAGGAUGAUGG-3') (Ribobio Company, Guangzhou, China). Cells were seeded at 1 × 10⁵ cells/well in 24-well plates. After 24 hours, cells were transfected with siRNA-TGFBR1 or siRNA-TGFBR2 in the presence of riboFECT™ CP transfection (Ribobio Company, Guangzhou, China) for 48 hours. Transfected cells were then collected for subsequent experiments. Flow cytometry was used to detect transfection efficiency.

Reverse transcription and real-time quantitative polymerase chain reaction. Total cellular RNA was isolated from the transfected THP-1 cells with Trizol (Invitrogen Corp, Carlsbad, CA) according to the manufacturer’s instructions. The extracted RNA was subjected to reverse transcription using the GoScript™ Reverse Transcription System kit (Promega Corp, Madison, WI). Real-time quantitative PCR (RT-PCR) was used to detect the relative mRNA level of SMAD7 and SMURF2. RT-PCR was conducted with cDNA converted from total RNA using amplification by 40 cycles of 95°C for 300 seconds, 95°C for 10 seconds, and 58°C for 10 seconds, 72°C for 30 seconds. The CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA) was used with MonAmp™ ChemoHS qPCR Mix (Monad Biotech Corp, Zhuhai, China). Expression of PCR products was normalized for that of glyceraldehyde 3-phosphate dehydrogenase (GAPDH). Primers for RT-PCR are shown in Table 1.

Statistical analysis. All the experiments were repeated 3 times to insure the reproducibility of the results. Differences between control and experimental groups were evaluated by the Kruskal-Wallis test with Steel’s post hoc test. A value of $P < .05$ was regarded as statistically significant.

Results

Differential expression analysis

Of the 40 samples, 22 283 probes were detected in total, and they could be well annotated to 12 434 genes, among which 716 genes were found significantly

Table 1. Primers for determining expression levels of *TGFBR1*, *TGFBR2*, *SMAD7*, *SMURF2*, and *GAPDH* genes.

Gene name	Forward primer	Reverse primer
<i>TGFBR1</i>	5'-TCAGCTGGTGGTGTAG-3'	5'-ATGTGAAGATGGGCAAGACC-3'
<i>TGFBR2</i>	5'-CAACATCAACCACAAACAGAG-3'	5'-CCGTCTCCGCTCCTAG-3'
<i>SMAD7</i>	5'-TTGCCTCGGACAGCTAAT-3'	5'-ATGGAGAACCGGGGAAACAC-3'
<i>SMURF2</i>	5'-CGCTTGATCAAAGTGGAAAT-3'	5'-GGTGATGGCATTGGAAAGA-3'
<i>GAPDH</i>	5'-GCACCGTCAAGGCTGAGAAC-3'	5'-TGGTGAAGACGCCAGTGG-3'

Table 2. Top significantly expressed genes in low vs high BMD postmenopausal Caucasians.

Gene	logFC	adj. P value	GEFOSFN.adjpval	GEFOSLS.adjpval	UKBiobank.adjpval
<i>ATP5H</i>	-0.24	2.27×10^{-4}	NA	4.29×10^{-3}	2.58×10^{-2}
<i>C1D</i>	-0.55	1.71×10^{-4}	NA	4.24×10^{-2}	NA
<i>CD320</i>	0.26	1.42×10^{-3}	NA	9.46×10^{-3}	6.26×10^{-3}
<i>DHX35</i>	0.17	3.82×10^{-3}	NA	3.44×10^{-2}	1.73×10^{-2}
<i>DICER1</i>	-0.58	2.99×10^{-3}	1.64×10^{-2}	7.41×10^{-3}	2.41×10^{-3}
<i>DIDO1</i>	0.19	5.62×10^{-4}	1.03×10^{-3}	7.95×10^{-5}	1.34×10^{-2}
<i>FAM222B</i>	0.22	3.94×10^{-3}	NA	1.81×10^{-2}	8.06×10^{-3}
<i>HIRA</i>	-0.61	1.88×10^{-3}	9.91×10^{-3}	4.08×10^{-3}	2.21×10^{-3}
<i>ITM2A</i>	0.23	1.52×10^{-3}	NA	NA	NA
<i>MYO9A</i>	0.26	3.94×10^{-3}	1.17×10^{-2}	1.31×10^{-2}	1.51×10^{-2}
<i>NDUFC1</i>	-0.28	2.63×10^{-3}	9.25×10^{-3}	NA	3.36×10^{-2}
<i>OSTF1</i>	0.34	1.25×10^{-3}	4.37×10^{-3}	4.24×10^{-2}	5.20×10^{-3}
<i>PDPK1</i>	0.18	4.46×10^{-3}	NA	NA	3.52×10^{-2}
<i>PFDN4</i>	-0.41	1.71×10^{-4}	3.35×10^{-3}	2.66×10^{-2}	2.52×10^{-2}
<i>RPL3</i>	0.21	3.10×10^{-4}	1.06×10^{-2}	1.11×10^{-2}	3.03×10^{-2}
<i>RPL35A</i>	-0.17	3.94×10^{-3}	NA	NA	1.33×10^{-3}
<i>SEC22B</i>	-0.57	1.02×10^{-3}	NA	NA	NA
<i>SET</i>	-0.63	2.27×10^{-4}	4.92×10^{-2}	NA	3.68×10^{-2}
<i>SLC35D2</i>	-0.28	1.00×10^{-5}	9.20×10^{-4}	1.32×10^{-3}	NA
<i>SNRPE</i>	-0.25	3.94×10^{-3}	4.16×10^{-2}	1.12×10^{-2}	5.32×10^{-4}
<i>TBC1D2</i>	-0.3	3.16×10^{-3}	5.64×10^{-3}	7.21×10^{-3}	1.75×10^{-2}
<i>TMEM258</i>	-0.32	1.82×10^{-3}	NA	1.90×10^{-2}	3.67×10^{-3}
<i>UBE2E1</i>	-0.3	1.25×10^{-3}	1.21×10^{-2}	2.86×10^{-3}	4.01×10^{-3}
<i>UBE3C</i>	-0.3	3.94×10^{-3}	8.98×10^{-3}	1.04×10^{-2}	9.63×10^{-3}
<i>XPO7</i>	0.28	7.87×10^{-4}	2.56×10^{-2}	2.58×10^{-2}	6.20×10^{-4}

Abbreviations: FN, femoral neck; LS, lumbar spine; LogFC, log of fold changes between high vs low BMD group; NA, the SNP of gene was not detected in the GEFOS-seq dataset or UK Biobank dataset.

Adjusted *P* value < .05 is used as thresholds for identifying significant genes.

differentially expressed with Benjamini and Hochberg adjusted *P* < .05. Table 2 shows the information for the top 25 differentially expressed genes and their corresponding *P* values in the GEFOS-seq database and the UK Biobank database.

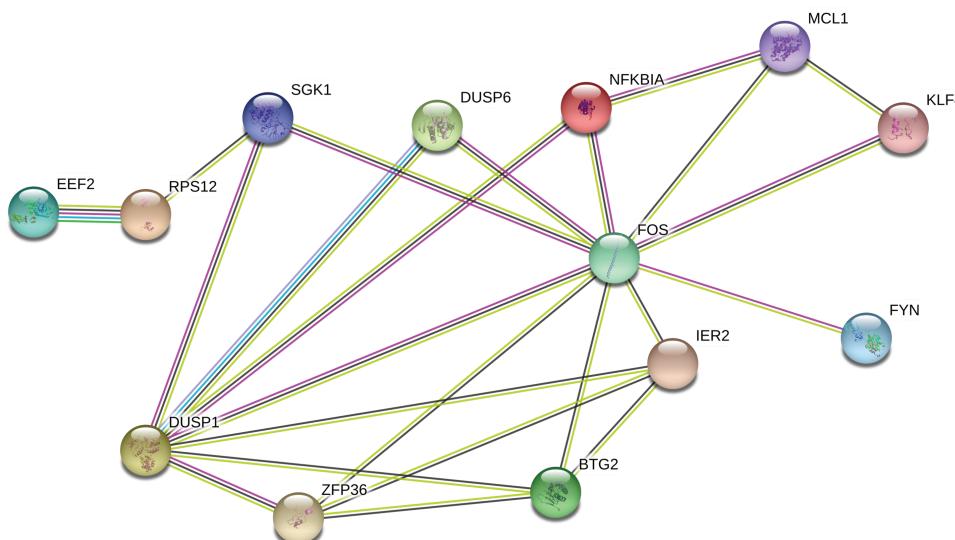
MEGENA analysis

The 716 differentially expressed genes were used in the subsequent MEGENA analysis. After parallelization, early termination and prior quality control steps, PFNs with 150 gene pairs entered into MCA to identify multiscale clusters. With MCA, there were in total 5 distinct scales identified: S1 (0.23 << α << 0.55), S2 (0.56 << α << 0.66), S3 (0.73 << α << 1.04), S4 (1.05 << α << 1.05), and S5 (1.07 << α << 1.61). In total 21 significant module clusters were constructed with module *P* < .05, with module size varying from 11 to 120.

By calculating Pearson's correlation between eigengenes (first component of the genes in the identified significant gene clusters) and phenotype, the gene cluster including *DUSP1*, *ZFP36*, *DUSP6*, *FOS*, *IER2*, *NFKBIA*, *KLF4*, *EEF2*, *FYN*, *MCL1*, *SGK1*, *RPS12*, and *BTG2* is significantly associated with BMD variation with *P* = .036. Based on GO enrichment analysis, 2 GO terms, “T cell receptor signaling pathway” and “Osteoclast differentiation” were significantly enriched. Detailed information is shown in Table 3. Based on computational predictions by STRING software, interacting relationships between genes were detected, as shown in Fig. 2. Detailed pairwise combined scores for the interacting network are shown in Table 4. The combined prediction score for pairwise association between *DUSP1*–*DUSP6*, *DUSP1*–*FYN*, *FYN*–*FOS*, and *EEF2*–*RPS12* reached the level of high confidence (combined prediction score > 0.7).

Table 3. Significant GO term in the BMD-related gene cluster by MEGENA analysis.

Term	P value	P value_fdr	P value_bonferroni
T cell receptor signaling pathway	2.94×10^{-5}	8.33×10^{-3}	8.45×10^{-3}
Osteoclast differentiation	5.81×10^{-5}	8.33×10^{-3}	1.67×10^{-2}

**Figure 2.** Interacting network view of genes in target BMD-related module. The network nodes represent genes, and lines represent the existence of evidence used in predicting the associations. Darker connecting lines between nodes represent higher combined score.**Table 4.** Interacting network identified in the BMD-related module by STRING.

Interaction network	Score	Evidence
BTG2–FOS	0.45	Comentioned in PubMed abstracts, Coexpression
DUSP1–BTG2	0.40	Comentioned in PubMed abstracts, Coexpression
DUSP1–DUSP6	0.96	Comentioned in PubMed abstracts, Coexpression
DUSP1–FOS	0.65	Comentioned in PubMed abstracts, Coexpression, Experimental/Biochemical data
DUSP1–FYN	0.85	Comentioned in PubMed abstracts, Experimental/Biochemical data
DUSP1–SGK1	0.50	Comentioned in PubMed abstracts, Experimental/Biochemical data
DUSP1–ZFP36	0.46	Comentioned in PubMed abstracts, Experimental/Biochemical data
FYN–FOS	0.91	Comentioned in PubMed abstracts, Experimental/Biochemical data
KLF4–FOS	0.41	Comentioned in PubMed abstracts, Experimental/Biochemical data
MCL1–FOS	0.41	Comentioned in PubMed abstracts
NFKBIA–FOS	0.50	Comentioned in PubMed abstracts
SGK1–FOS	0.41	Comentioned in PubMed abstracts, Experimental/Biochemical data
ZFP36–FOS	0.56	Comentioned in PubMed abstracts, Coexpression
EEF2–RPS12	0.900	Comentioned in PubMed abstracts, Coexpression, Experimental/Biochemical data, Neighborhood in the Genome, Association in Curated Database

Gene set net correlation analysis

Based on GSNCA, the coexpression structure of 26 predefined pathways were significantly different between high and low BMD groups with $P < .05$ as shown in Table 5. The coexpression pattern of “Signaling by TGF-beta receptor complex” is significantly different between 2 discordant BMD groups with $P = .015$. In this pathway, 15 genes (UBE2D3, UBE2D1, TGFB1, TGFB2, SMAD2, FURIN, SMAD3, SKI, SMAD4, ZFYVE9, TGFB1, SMURF1, SMAD6, SMURF2, SMAD7) in our expression dataset were identified.

The hub gene with largest weight in the high BMD group is SMAD3, while the hub gene in the low BMD group is FURIN. Since TGFB1 plays an important role in this pathway, we compared some network features of TGFB1 in the high versus low BMD groups. In the low BMD group, the degree centrality and the normalized betweenness centrality of TGFB1 is 5 and 0.21, respectively. While in the high BMD group, the degree centrality and the normalized betweenness centrality of TGFB1 is 2 and 0, respectively. It suggests that TGFB1 has more connections and influence with

Table 5. Significant gene sets identified with gene set net comparison analysis.

Name	Adjusted P value
BIOCARTA_SODD_PATHWAY	.0060
KEGG_VIBRIO_CHOLERAES_INFECTION	.0080
REACTOME_GTP_HYDROLYSIS_AND_JOINING_OF_THE_60S_RIBOSOMAL_SUBUNIT	.0110
REACTOME_JNK_PHOSPHORYLATION_AND_ACTIVATION_MEDIATED_BY_ACTIVATED_HUMAN_TAK1	.0120
KEGG_RIBOSOME	.0130
BIOCARTA_ALK_PATHWAY	.0130
REACTOME_SIGNALING_BY_TGF_BETA	.0150
REACTOME_TRANSLATION	.0160
REACTOME_RNA_PolyMERASE_III_CHAIN_ELONGATION	.0160
KEGG_TAURINE_AND_HYPOTAURINE_METABOLISM	.0200
BIOCARTA_PLCE_PATHWAY	.0230
REACTOME_HORMONE_SENSITIVE_LIPASE_HSL_MEDIATED_TRIACYLGLYCEROL_HYDROLYSIS	.0280
KEGG_LYSOSOME	.0290
KEGG_RNA_PolyMERASE	.0340
REACTOME_PLC_BETA_MEDiated_EVENTS	.0340
KEGG_BASE_EXCISION_REPAIR	.0380
REACTOME_INFLUENZA_VIRAL_RNA_TRANSCRIPTION_AND_REPLICATION	.0390
REACTOME_BASE_EXCISION_REPAIR	.0400
REACTOME_DOPAMINE_NEUROTRANSMITTER_RELEASE_CYCLE	.0440
REACTOME_FORMATION_OF_A_POOL_OF_FREE_40S_SUBUNITS	.0450
REACTOME_CAM_PATHWAY	.0460
BIOCARTA_NOS1_PATHWAY	.0460
KEGG_GLYOXYLATE_AND_DICARBOXYLATE_METABOLISM	.0500
REACTOME_POST_TRANSLATIONAL_PROTEIN_MODIFICATION	.0500
REACTOME_COMPLEMENT CASCADE	.0500
REACTOME_ACTIVATION_OF_KAINATE_RECEPTEORS_UPON GLUTAMATE_BINDING	.0500

the gene network in the low BMD group. The MST2 plot of correlation patterns in the 2 BMD groups are shown in Fig. 3. And the network feature comparison is shown in Table 6. In addition, another 2 vital gene sets “REACTOME_CAM_PATHWAY” and “BIOCARTA_NOS1_PATHWAY” are found to be significantly different between the high and low BMD groups with $P < .05$. The MST2 plots of the correlation pattern of these 2 gene sets are shown in Figs. 4 and 5.

Bayesian network analysis

BN analysis was applied to the differentially coexpressed gene sets identified by the previous GSNCA. As shown in Table 7, in the “Signaling by TGF-beta” gene set, 4 significant directional gene pairs (strength >0.80 , direction >0.5) were identified in the high BMD group, while only 2 significant directional gene pairs were identified in the low BMD group. Strength measures the significance of arcs between nodes (genes). Strength >0.80 means that the arc appears in at least 80% of the networks with 500 times bootstrap resampling to Bayesian network structures used for model averaging; direction >0.5 indicates that the path flow of the 2 nodes appears most frequently. The arc with highest strength in the high BMD group is from *UBE2D3* to *UBE2D1* (0.98), while the highest strength in the low BMD group is from *TGFBR2* to *SMAD7* (0.82). It also worth noticing that the second highest strength in low BMD group is from *TGFBR1* to *SMURF2* (0.80).

siRNA gene silencing experiment

RT-PCR experiments showed that transfection with siRNA-TGFBR1 significantly decreases SMURF2 mRNA expression levels in THP1 cells ($P = .044$) when the silencing efficiency reaches 80% compared with siRNA control (Fig. 6). Transfection with siRNA-TGFBR2 significantly decreases SMAD7 mRNA expression levels in THP1 cells ($P = .018$) when the silencing efficiency reaches 70% to 80% compared with the siRNA control (Fig. 6).

Discussion

The present study applied both traditional differential expression analysis and state-of-the-art network analyses to transcriptome-wide gene expression data. We applied comprehensive network-based system biology analyses (de novo multiscale network construction, gene set comparison with predefined pathway information, and Bayesian network construction). The MEGENA method helps us build a coexpression network based on data expression features, while the coexpression network constructed by GSNCA incorporates more biological pathway information. Since MEGENA and GSNCA do not provide causal relationship between genes, BNs are a great complement, helping us explore regulatory function between genes. These 3 network

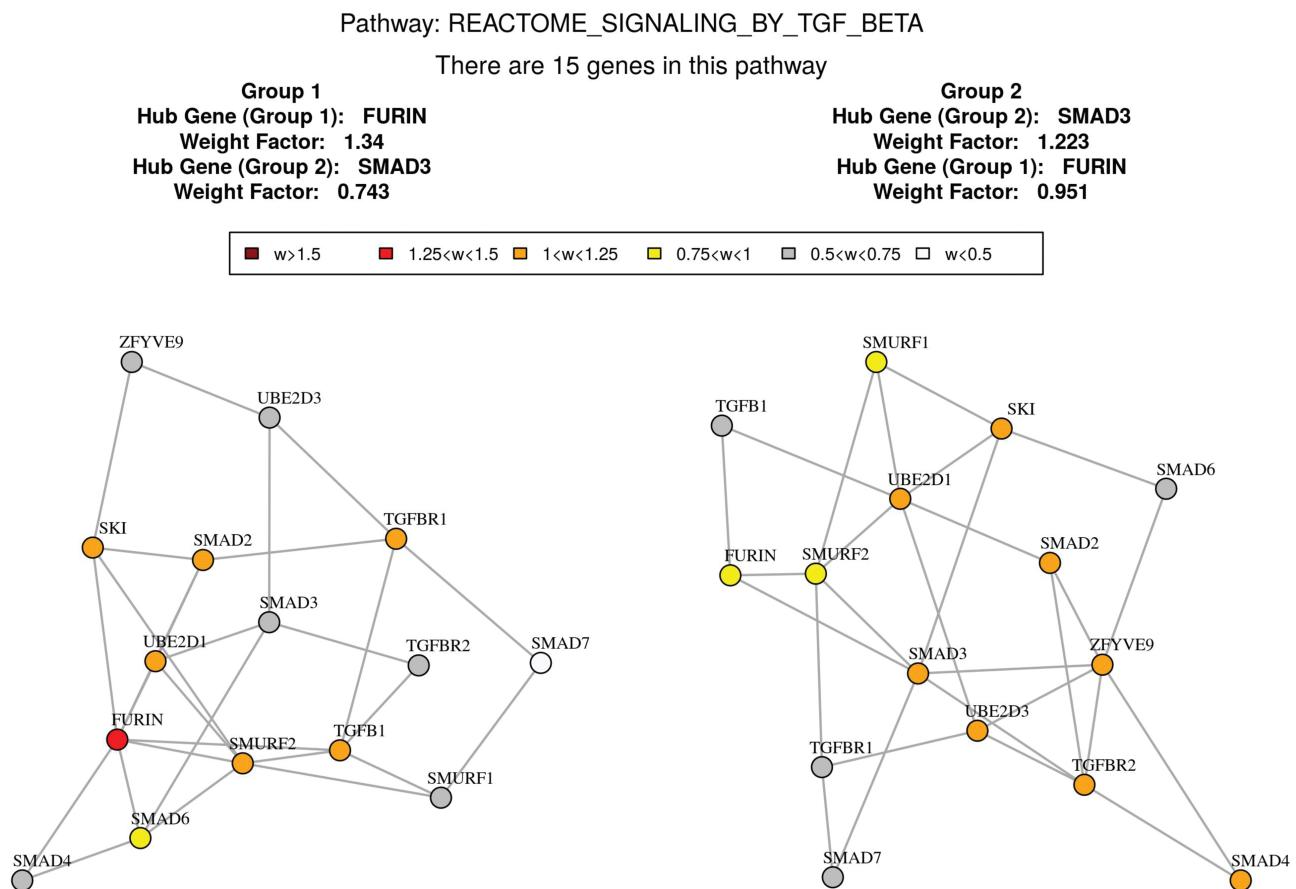


Figure 3. Co-expression network comparison between high and low BMD subjects in REACTOME_SIGNALING_BY_TGF_BETA by GSNCA.

Table 6. Coexpression network comparison between high and low BMD subjects in “REACTOME_SIGNALING_BY_TGF_BETA.”

	Low_BMD group	High_BMD group
Average absolute cross correlation	0.25	0.28
Degree centrality of hub gene	FURTIN (7)	SMAD3 (6)
Degree centrality of TGFB1	5	2
Normalized betweenness centrality of hub gene	FURTIN (0.26)	SMAD3 (0.29)
Normalized betweenness centrality of TGFB1	0.21	0.00
Normalized closeness centrality of hub gene	FURTIN (1.35)	SMAD (1.35)
Average normalized closeness centrality	1.0	1.1

FURTIN is a hub gene in the low BMD group and it connected to 7 genes in the coexpression structure; *SMAD* is a hub gene in the high BMD group and it has 6 neighbors. The connections of *TGFB1* in 2 BMD groups are 5 and 2 respectively. The normalized shortest paths that pass through *FURTIN* in the low BMD group is 0.26; the normalized shortest paths that pass through *SMAD3* in low the BMD group is 0.29. The normalized shortest paths that pass through *TGFB1* in 2 BMD groups are 0.21 and 0 respectively. The normalized mean distance from *FURTIN* to other vertices (genes) is 1.35 in the low BMD group, and the normalized mean distance from *SMAD3* to other vertices (genes) is 1.35 as well in the high BMD group.

analysis approaches together helped us discover transcriptome data more comprehensively. With a series of bioinformatics analyses, we not only found genes and pathways important for postmenopausal osteoporosis, but also validated regulation patterns between crucial gene pairs in PBMs based on in vitro experiments.

Based on differential expression analysis, 716 genes were significantly differentially expressed ($P_{\text{adjust}} < 0.05$) in subjects with extremely low versus high BMD. We checked the SNPs of the top 25 genes ($P_{\text{adjust}} < 0.005$) for

their association signals in the GEFOS dataset and the UK Biobank dataset. Since GEFOS study is a large-scale meta-analysis performed in the bone field and the UK Biobank osteoporosis study is the largest GWAS study that measures the association between SNPs with eBMD (crucial for bone fracture), the consistent results make our cohort study much more convincing.

Much previous research has provided further evidence for our findings. DICER1 ($P_{\text{val}} = 2.99 \times 10^{-3}$) is a conserved RNase III endonuclease that is essential

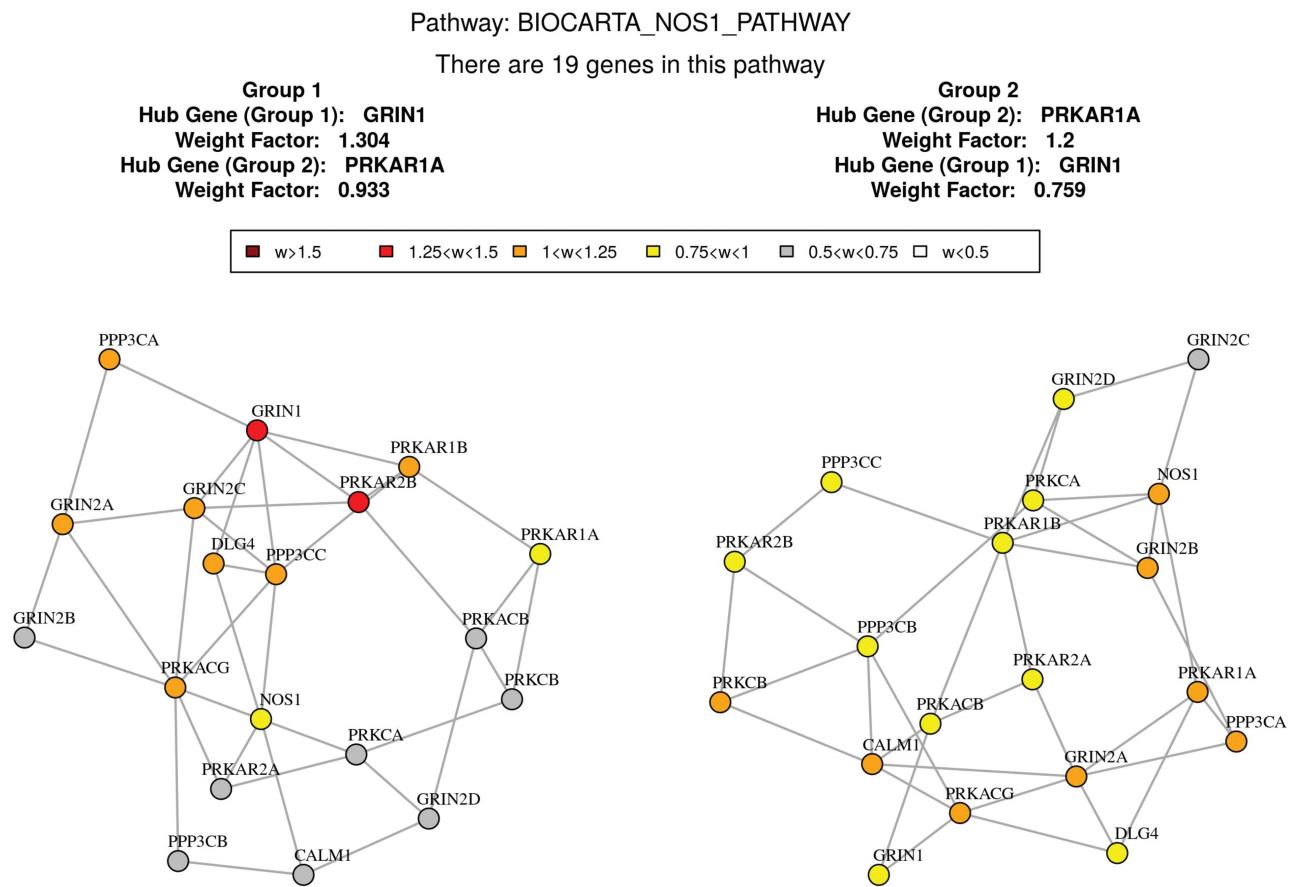


Figure 4. Co-expression network comparison between high and low BMD subjects in “BIOCARTA_NOS1_PATHWAY” by GSNCA.

for the processing of several classes of small RNAs. Downregulation of DICER1 expression inhibits osteogenic differentiation of several cells (43), and Dicer1 processed microRNAs may have an important role in the regulation of skeletal homeostasis (44). *HIRA* ($P_{\text{val}} = 1.88 \times 10^{-3}$) has been recorded downregulated in osteoporosis blood samples, and plays a significant role in histone modification in bone development in osteoporosis (45). *OSTF1* (Osteoclast stimulation factor) ($P_{\text{val}} = 1.25 \times 10^{-3}$) was initially identified as a factor involved in the indirect activation of osteoclasts, and it could indirectly enhance osteoclast formation and bone resorption activity in cell culture assays (46). *TBC1D2* ($P_{\text{val}} = 3.16 \times 10^{-3}$) is a protein-coding gene from the TBC1 domain family. It participates in the GTPase activator process, and regulation of *TBC1D2*-dependent activities plays a critical role in osteoclast function (47). In addition, we also found a novel gene *ITM2A* which was not detected in either of the meta-analyses. From previous research, *ITM2A* had been shown to be involved in bone growth, mineralization, and bone morphogenic protein metabolism (48). The significant change of *ITM2A* in our gene expression data suggests the importance of this gene in regulating the bone formation/resorption process.

The differential expression analysis helps detection of target genes associated with the pathology of osteoporosis, providing us with some new genes that may serve as diagnostic biomarkers of postmenopausal BMD variation.

In MENEGA analysis, the 2 most significantly enriched pathways in the constructed BMD-associated gene cluster were T Cell Receptor signaling and aggregation and Osteoclast differentiation. T Cell Receptor (TCR) signaling could promote plenty of signaling cascades that ultimately determine cell fate, and TCR signaling provokes rapid increases in integrin function that facilitate T-cell activation (49). Integrin is central to physiological bone resorption, and it acts as a candidate antibone resorptive therapeutic target (50). Under certain pathological circumstances such as estrogen deficiency, the activation of T cells may disturb bone homeostasis and bring about bone destruction process (51). T cells work closely with RANK ligand or RANKL (receptor activator of nuclear factor- κ B ligand), essential stimulating signals for osteoclastogenesis, which is involved in activating mature osteoclasts (52). Moreover, T cells could secret various cytokines such as interleukin (IL)-1, IL-6, interferon- γ , or IL-4 that mediating osteoclastogenic signals in bone turnover (53). Osteoclast differentiation signaling has formerly been

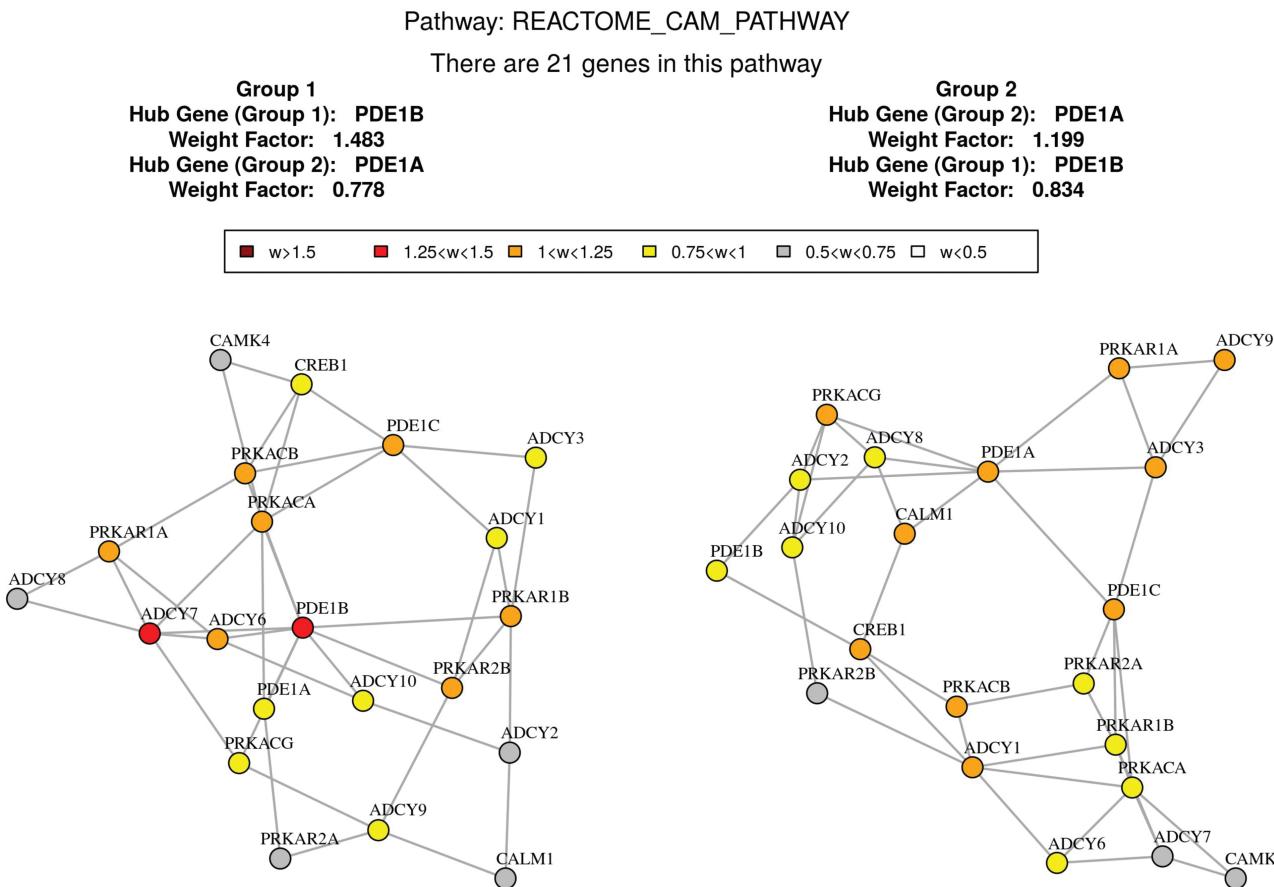


Figure 5. Co-expression network comparison between high and low BMD subjects in “REACTOME_CAM_PATHWAY” by GSNCA.

Table 7. Bayesian network analysis between high and low BMD subjects in “REACTOME_SIGNALING_BY_TGF_BETA.”

From-To (high BMD)	Strength	Direction	From-To (low BMD)	Strength	Direction
UBE2D3–UBE2D1	0.98	0.70	TGFBR2–SMAD7	0.82	0.73
UBE2D3–SMAD2	0.92	0.77	TGFBR1–SMURF2	0.80	0.55
UBE2D1–SMAD2	0.82	0.68			
FURIN–SMURF2	0.88	0.80			

Strength measures the significance of arcs between nodes (genes). Arcs are considered significant if they appear in at least 80% of the networks with bootstrap resampling to Bayesian network structures use for model averaging. Direction measures the direction of arcs learned by Bayesian network. Direction that appears most frequently (>0.5) is kept in the table.

revealed to cause postmenopausal osteoporosis. Estrogen deficiency stimulates osteoclast formation/differentiation by increasing hematopoietic progenitors and provides a larger recruited osteoclast progenitor pool (54), as well as increase the life span of osteoclasts (55). Furthermore, osteoclast differentiation signaling is controlled by 2 essential cytokines: macrophage colony-stimulating factor (M-CSF) and RANKL. M-CSF and RANKL bind to their respective receptors c-Fms/RANK to stimulate osteoclast differentiation through regulation of delicate signaling systems (55). Thus, the detected enriched “Osteoclast differentiation” pathway could provide further evidence that the activities of M-CSF-c-Fms signaling and

RANKL–RANK signaling might contribute to the pathogenesis of postmenopausal osteoporosis through their specific actions on osteoclastogenesis. Furthermore, based on the enrichment analysis in the target model, the combined prediction score for pairwise association between *DUSP1*–*DUSP6* and *DUSP1*–*FYN* reached the level of high confidence, which gives us a novel insight that the *DUSP* (Dual-specificity phosphatase) gene family is very active and important for osteoclastogenesis in PBMs.

In GSNCA, the gene coexpression pattern of REACTOME_SIGNALING_BY_TGF_BETA is significantly different in high versus low BMD subjects with $P < .05$. Transforming growth factor (TGF)- β signaling

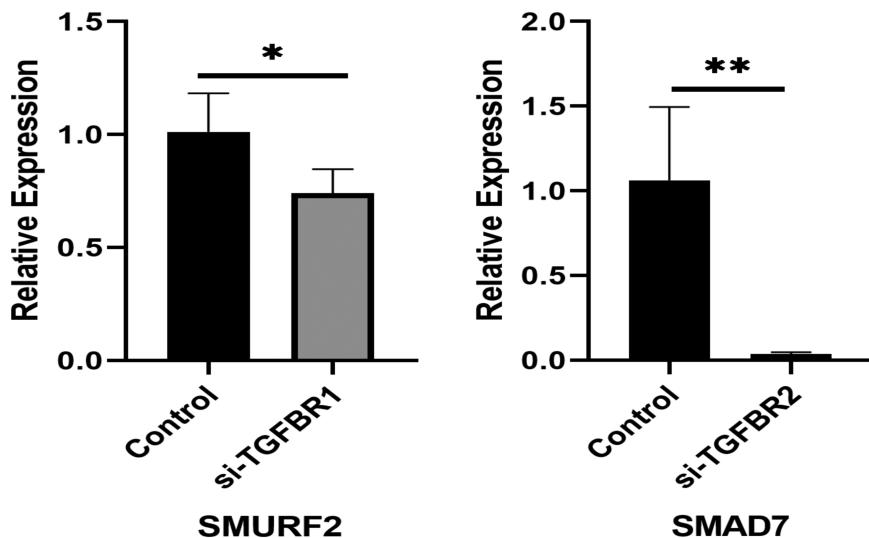


Figure 6. siRNA gene silencing effect of TGFBR2-SMAD7 and TGFBR1-SMURF2 in THP-1 cells.

is involved in a variety of cellular processes and plays an significant role in a wide range of biological systems (56). This signaling pathway incorporates various TGF- β /superfamily ligands and receptors as well as the receptor-regulated SMADs. TGF- β is known to regulate RANKL-induced osteoclast formation and bone resorbing activity and it has been demonstrated that in the presence of M-CSF, it could induce mononuclear phagocytes to differentiate into osteoclastic cells. Fox et al. suggested that the osteoclast/macrophage commitment could be efficiently induced by TGF- β -related cytokine signaling (57). In addition, Yan et al. have demonstrated that TGF- β -induced upregulation of RANK expression facilitated RANKL-induced osteoclastogenesis (58). From our analysis, the hub genes of REACTOME_SIGNALING_BY_TGF_BETA in the low BMD group and high BMD group are *FURIN* and *SMAD3* respectively, and their structure features are quite different. These have not been detected in any previous postmenopausal osteoporosis analysis. The result not only discloses different gene interacting patterns in PBMs between high and low BMD groups, but also suggests the importance of *FURIN* and *SMAD3* genes in PBMs during bone turnover.

Another two important signaling pathways detected in GSNCA analysis are BIOCARTA_NOS1_PATHWAY and REACTOME_CAM_PATHWAY. In the NOS1 pathway, glutamate stimulates NMDA-mediated calcium influx, promoting nitric oxide (NO) synthesis, thus activating guanylate cyclase. It has been found that 17 β -estradiol protects osteocytes against apoptosis via the NO/cGMP signaling pathway (59), and NO is capable of inducing inflammatory cell apoptosis including direct downregulation of leukocyte functions.

In addition, it has been shown that the exogenous NO could stimulate apoptosis in cell types such as monocytes, monocyte-derived macrophages, neutrophils, and eosinophils (60). The significant difference of NOS1 signaling between high and low BMD group may suggest different apoptosis activities of monocytes, and our analysis provides further support that NOS1 signaling plays an important role in postmenopausal osteoporosis. Ca²⁺/CAM-mediated cell signaling is ubiquitous and functions at all levels of the cell life cycle including cell genesis, cell function/activity and cell death. It has been reported that CAM pathways participates in the regulation of both osteoclasts differentiation and function (61). Therefore, our gene set comparison analysis of the transcriptome data not only helped to identify several vital genes that may be ignored by traditional differential expression analyses, but also suggested several important signaling process in postmenopausal osteoporosis.

In the BN analysis, the highest strength between gene pairs in the low BMD group of the SIGNALING_BY_TGF-BETA pathway is from *TGFBR2* to *SMAD7* (strength = 0.82, direction = 0.73). The result indicates the arc between *TGFBR2* and *SMAD7* appears in at least 82% of the networks with 500 times bootstrap resampling to BN structures used for model averaging, and the most frequent path flow direction is from *TGFBR2* to *SMAD7*. Through functional validations, when the silencing efficiency reaches 70% to 80%, the expression of *SMAD7* mRNA was significantly downregulated ($P = .018$). As transmembrane kinase, *TGFBR2* could induce apoptosis and differentiation processes activated by TGF- β (62). In previous studies, it has been demonstrated that the binding process between TGF- β and *TGFBR2* could

initiate phosphorylation of SMAD families and regulate the transcription of several well-known TGFBR2 target genes like SMAD7 (63). Additionally, SMAD7 is an essential factor for the regulation loop of the TGF- β /Smad signaling pathway (64). Our functional validation confirms the regulatory information flow from TGFBR2 to SMAD7 in PBM that we derived *in vivo* in humans, and their regulation pattern might have some further influence on those signaling pathways that are important for postmenopausal osteoclastogenesis. The gene pair with second highest strength in the BN analysis within the low BMD group is TGFBR1-SMURF2 (strength = 0.79, direction = 0.55). The arc between TGFBR1 and SMURF2 appears in at least 79% of the networks with 500 times bootstrap resampling to BNs structures used for model averaging, and the probability of direction between these 2 genes is 0.55. The result indicates interaction between TGFBR1 and SMURF2 in TGF- β signaling is very active in the PBM of low BMD group. Through functional validations, we found that when the silencing efficiency reaches 80%, the expression of SMURF2 mRNA was significantly downregulated ($P = .044$). TGFBR1 is an important transmembrane protein for TGF- β signals, and it could be regulated by SMAD ubiquitination regulatory factors (Smurf) including SMURF2 (65). In addition, TGFBR1 could interact with SMURF2 to affect ubiquitin-dependent degradation (66). Most of previous studies indicated the regulation pattern of SMURF2 to TGFBR1 (67), but our BN analysis provides an innovative insight that the feedback signal from TGFBR1 to SMURF2 is predominant for monocyte activities within low BMD group. Since SMURF2 could act as an essential regulator of the molecular communication between osteoblasts and osteoclasts (67). Our result support a new hypothesis that in TGF- β pathway, the regulation effect of TGFBR1 to SMURF2 may influence the bone resorption activities in PBMs within low BMD group, and the information flow from TGFBR1 to SMURF2 is predominant.

Integrative network-based analysis with transcriptome-wide expression data holds great promise for illuminating the genetic basis of complex diseases. The present study demonstrated that MEGENA, GSNCA, and BN analyses were useful for identifying postmenopausal osteoporosis-related biologically meaningful gene groups and signaling pathways. Since genes do not usually work independently, the significant coexpression pattern could reveal how genes function within PBMs. This integrative analysis strategy not only helps to reveal disease-associated genes through transcriptomics data, but also shed light on potential therapeutic strategies for postmenopausal osteoporosis.

Acknowledgments

Financial Support: This study was partially supported or benefited by grants from the National Institutes of Health [R01AR059781, P20GM109036, R01MH107354, R01MH104680, R01GM109068, R01AR069055, and U19AG055373].

Additional Information

Correspondence and Reprint Requests: Hong-Wen Deng, Center for Biomedical Informatics and Genomics, Department of Medicine, School of Medicine, Tulane University, New Orleans, LA 70112, USA. E-mail: hdeng@tulane.edu.

Disclosure Summary: All authors approved the final content of the manuscript and declared no conflict of interest related to this manuscript.

Data Availability: All data generated or analyzed during this study are included in this published article or in the data repositories listed in References.

References

1. Lerner UH. Bone remodeling in post-menopausal osteoporosis. *J Dent Res.* 2006;85(7):584-595.
2. Aggarwal N, Raveendran A, Khandelwal N, et al. Prevalence and related risk factors of osteoporosis in peri- and postmenopausal Indian women. *J Midlife Health.* 2011;2(2):81-85.
3. Sioka C, Fotopoulos A, Georgiou A, Xouria X, Papadopoulos A, Kafele-Ezra JA. Age at menarche, age at menopause and duration of fertility as risk factors for osteoporosis. *Climacteric.* 2010;13(1):63-71.
4. Geissmann F, Manz MG, Jung S, Sieweke MH, Merad M, Ley K. Development of monocytes, macrophages, and dendritic cells. *Science.* 2010;327(5966):656-661.
5. Soudja SM, Ruiz AL, Marie JC, Lauvau G. Inflammatory monocytes activate memory CD8(+) T and innate NK lymphocytes independent of cognate antigen during microbial pathogen invasion. *Immunity.* 2012;37(3):549-562.
6. Auffray C, Sieweke MH, Geissmann F. Blood monocytes: development, heterogeneity, and relationship with dendritic cells. *Annu Rev Immunol.* 2009;27:669-692.
7. Braun T, Zwerina J. Positive regulators of osteoclastogenesis and bone resorption in rheumatoid arthritis. *Arthritis Res Ther.* 2011;13(4):235.
8. Amarasekara DS, Yun H, Kim S, Lee N, Kim H, Rho J. Regulation of osteoclast differentiation by cytokine networks. *Immune Netw.* 2018;18(1):e8.
9. Zupan J, Jeras M, Marc J. Osteoimmunology and the influence of pro-inflammatory cytokines on osteoclasts. *Biochem Med (Zagreb).* 2013;23(1):43-63.
10. Zhou Y, Deng HW, Shen H. Circulating monocytes: an appropriate model for bone-related study. *Osteoporos Int.* 2015;26(11):2561-2572.
11. Grundberg E, Brändström H, Lam KC, et al. Systematic assessment of the human osteoblast transcriptome in resting and induced primary cells. *Physiol Genomics.* 2008;33(3):301-311.
12. Liu YZ, Zhou Y, Zhang L, et al. Attenuated monocyte apoptosis, a new mechanism for osteoporosis suggested by a transcriptome-wide expression study of monocytes. *PLoS One.* 2015;10(2):e0116792.

13. Song WM, Zhang B. Multiscale embedded gene co-expression network analysis. *PLoS Comput Biol.* 2015;11(11):e1004574.
14. Song WM, Di Matteo T, Aste T. Building complex networks with Platonic solids. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2012;85(4 Pt 2):046115.
15. Song WM, Di Matteo T, Aste T. Hierarchical information clustering by means of topologically embedded graphs. *PLoS One.* 2012;7(3):e31929.
16. Rahmatallah Y, Emmert-Streib F, Glazko G. Gene Sets Net Correlations Analysis (GSNCA): a multivariate differential coexpression test for gene sets. *Bioinformatics.* 2014;30(3):360-368.
17. van Dam S, Võsa U, van der Graaf A, Franke L, de Magalhães JP. Gene co-expression analysis for functional classification and gene-disease predictions. *Brief Bioinform.* 2018;19(4):575-592.
18. Friedman N, Linial M, Nachman I, Pe'er D. Using Bayesian networks to analyze expression data. *J Comput Biol.* 2000;7(3-4):601-620.
19. Mäkinen VP, Civelek M, Meng Q, et al.; Coronary ARtery Disease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Consortium. Integrative genomics reveals novel molecular pathways and gene networks for coronary artery disease. *PLoS Genet.* 2014;10(7):e1004502.
20. Zhang B, Gaiteri C, Bodea LG, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell.* 2013;153(3):707-720.
21. Deng FY, Lei SF, Zhang Y, et al. Peripheral blood monocyte-expressed ANXA2 gene is involved in pathogenesis of osteoporosis in humans. *Mol Cell Proteomics.* 2011;10(11):M111.011700.
22. Strauss-Ayali D, Conrad SM, Mosser DM. Monocyte subpopulations and their differentiation patterns during infection. *J Leukoc Biol.* 2007;82(2):244-252.
23. Schroeder A, Mueller O, Stocker S, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol.* 2006;7(1):3.
24. Kiewe P, Gueller S, Komor M, Stroux A, Thiel E, Hofmann WK. Prediction of qualitative outcome of oligonucleotide microarray hybridization by measurement of RNA integrity using the 2100 Bioanalyzer capillary electrophoresis system. *Ann Hematol.* 2009;88(12):1177-1183.
25. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* 2003;31(4):e15.
26. Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. *Bioinformatics.* 2010;26(19):2363-2367.
27. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.
28. Rivadeneira F, Styrkársdóttir U, Estrada K, et al.; Genetic Factors for Osteoporosis (GEFOS) Consortium. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat Genet.* 2009;41(11):1199-1206.
29. Estrada K, Styrkársdóttir U, Evangelou E, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet.* 2012;44(5):491-501.
30. Zheng HF, Forgetta V, Hsu YH, et al.; AOGC Consortium; UK10K Consortium. Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature.* 2015;526(7571):112-117.
31. Bauer DC, Ewing SK, Cauley JA, Ensrud KE, Cummings SR, Orwoll ES; Osteoporotic Fractures in Men (MrOS) Research Group. Quantitative ultrasound predicts hip and non-spine fracture in men: the MrOS study. *Osteoporos Int.* 2007;18(6):771-777.
32. Morris JA, Kemp JP, Youlten SE, et al.; 23andMe Research Team. An atlas of genetic influences on osteoporosis in humans and mice. *Nat Genet.* 2019;51(2):258-266.
33. Tumminello M, Aste T, Di Matteo T, Mantegna RN. A tool for filtering information in complex systems. *Proc Natl Acad Sci U S A.* 2005;102(30):10421-10426.
34. Albert R, Barabasi AL. Statistical mechanics of complex networks. *Rev Mod Phys.* 2002;74(1):47-97.
35. Newman ME. Modularity and community structure in networks. *Proc Natl Acad Sci U S A.* 2006;103(23):8577-8582.
36. Snel B, Lehmann G, Bork P, Huynen MA. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res.* 2000;28(18):3442-3444.
37. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst.* 2015;1(6):417-425.
38. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics.* 2011;27(12):1739-1740.
39. Dawson JA, Kendziora C. An empirical Bayesian approach for identifying differential coexpression in high-throughput experiments. *Biometrics.* 2012;68(2):455-465.
40. Chu T, Glymour C, Scheines R, Spirtes P. A statistical problem for inference to regulatory structure from associations of gene expression measurements with microarrays. *Bioinformatics.* 2003;19(9):1147-1152.
41. Claeskens G, Hjort NL. *Model Selection and Model Averaging.* Cambridge, UK: Cambridge University Press; 2008.
42. Bosshart H, Heinzelmann M. THP-1 cells as a model for human monocytes. *Ann Transl Med.* 2016;4(21):438.
43. Kim DS, Lee SY, Lee JH, Bae YC, Jung JS. MicroRNA-103a-3p controls proliferation and osteogenic differentiation of human adipose tissue-derived stromal cells. *Exp Mol Med.* 2015;47(7):e172.
44. Bendre A, Moritz N, Vääränen V, Määttä JA. Dicer1 ablation in osterix positive bone forming cells affects cortical bone homeostasis. *Bone.* 2018;106:139-147.
45. Xia B, Li Y, Zhou J, Tian B, Feng L. Identification of potential pathogenic genes associated with osteoporosis. *Bone Joint Res.* 2017;6(12):640-648.
46. Vermeren M, Lyraiki R, Wani S, et al. Osteoclast stimulation factor 1 (Ostf1) KNOCKOUT increases trabecular bone mass in mice. *Mamm Genome.* 2017;28(11-12):498-514.
47. Xing WR, Goodluck H, Zeng C, Mohan S. Role and mechanism of action of leucine-rich repeat kinase 1 in bone. *Bone Res.* 2017;5(1):17003.
48. Camirand A, Goltzman D, Gupta A, Kaouass M, Panda D, Karaplis A. The role of parathyroid hormone-related protein (PTHrP) in osteoblast response to microgravity: mechanistic implications for osteoporosis development. *PLoS One.* 2016;11(7):e0160034.
49. Burbach BJ, Medeiros RB, Mueller KL, Shimizu Y. T-cell receptor signaling to integrins. *Immunol Rev.* 2007;218(1):65-81.
50. Teitelbaum SL. Osteoporosis and integrins. *J Clin Endocrinol Metab.* 2005;90(4):2466-2468.
51. Weitzmann MN, Pacifici R. Estrogen deficiency and bone loss: an inflammatory tale. *J Clin Invest.* 2006;116(5):1186-1194.
52. Zhao W, Liu Y, Cahill CM, Yang W, Rogers JT, Huang X. The role of T cells in osteoporosis, an update. *Int J Clin Exp Pathol.* 2009;2(6):544-552.
53. D'Amico L, Roato I. Cross-talk between T cells and osteoclasts in bone resorption. *Bonekey Rep.* 2012;1:82.
54. Jilka RL, Hangoc G, Girasole G, et al. Increased osteoclast development after estrogen loss: mediation by interleukin-6. *Science.* 1992;257(5066):88-91.
55. Zhao R. Immune regulation of osteoclast function in postmenopausal osteoporosis: a critical interdisciplinary perspective. *Int J Med Sci.* 2012;9(9):825-832.
56. Guo X, Wang XF. Signaling cross-talk between TGF-beta/BMP and other pathways. *Cell Res.* 2009;19(1):71-88.

57. Fox SW, Evans KE, Lovibond AC. Transforming growth factor-beta enables NFATc1 expression during osteoclastogenesis. *Biochem Biophys Res Commun.* 2008;366(1):123-128.
58. Yan T, Riggs BL, Boyle WJ, Khosla S. Regulation of osteoclastogenesis and RANK expression by TGF-beta1. *J Cell Biochem.* 2001;83(2):320-325.
59. Joshua J, Kalyanaraman H, Marathe N, Pilz RB. Nitric oxide as a mediator of estrogen effects in osteocytes. *Vitam Horm.* 2014;96:247-263.
60. Taylor EL, Megson IL, Haslett C, Rossi AG. Nitric oxide: a key regulator of myeloid inflammatory cell apoptosis. *Cell Death Differ.* 2003;10(4):418-430.
61. Wu X, Ahn EY, McKenna MA, Yeo H, McDonald JM. Fas binding to calmodulin regulates apoptosis in osteoclasts. *J Biol Chem.* 2005;280(33):29964-29970.
62. Kim SJ, Im YH, Markowitz SD, Bang YJ. Molecular mechanisms of inactivation of TGF-beta receptors during carcinogenesis. *Cytokine Growth Factor Rev.* 2000;11(1-2):159-168.
63. Lee J, Ballikaya S, Schönig K, et al. Transforming growth factor beta receptor 2 (TGFBR2) changes sialylation in the microsatellite unstable (MSI) Colorectal cancer cell line HCT116. *PLoS One.* 2013;8(2):e57074.
64. Zhang H, Ta N. Effect of isoporsoralen on Smad7 in osteoblastic MC3T3-E1 cells. *Exp Ther Med.* 2017;14(2):1561-1567.
65. Bizet AA, Tran-Khanh N, Saksena A, Liu K, Buschmann MD, Philip A. CD109-mediated degradation of TGF- β receptors and inhibition of TGF- β responses involve regulation of SMAD7 and Smurf2 localization and function. *J Cell Biochem.* 2012;113(1):238-246.
66. Cai Y, Zhou CH, Fu D, Shen XZ. Overexpression of Smad ubiquitin regulatory factor 2 suppresses transforming growth factor- β mediated liver fibrosis. *J Dig Dis.* 2012;13(6):327-334.
67. Xu Z, Greenblatt MB, Yan G, et al. SMURF2 regulates bone homeostasis by disrupting SMAD3 interaction with vitamin D receptor in osteoblasts. *Nat Commun.* 2017;8(1):14570.