Published in final edited form as:

Osteoporos Int. 2017 October ; 28(10): 3033–3042. doi:10.1007/s00198-017-4146-5.

Network based subcellular proteomics in monocyte membrane revealed novel candidate genes involved in osteoporosis

Y. Zeng^{1,2}, L. Zhang², W. Zhu^{2,3}, H. He², H. Sheng², Q. Tian², F.-Y. Deng^{2,4}, L.-S. Zhang¹, H.-G. Hu¹, and H.-W. Deng^{1,2}

¹College of Life Sciences and Bioengineering, Beijing Jiaotong University, Beijing 100044, China

²Center of Bioinformatics and Genomics, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA 70112, USA

³College of Life Sciences, Hunan Normal University, Changsha, Hunan 410081, China

⁴Laboratory of Proteins and Proteomics, Department of Epidemiology, Soochow University School of Public Health, Suzhou, Jiangsu 205123, China

Abstract

Summary—In this study, label-free-based quantitative subcellular proteomics integrated with network analysis highlighted several candidate genes including P4HB, ITGB1, CD36, and ACTN1 that may be involved in osteoporosis. All of them are predicted as significant membrane proteins with high confidence and enriched in bone-related biological process. The results were further verified in transcriptomic and genomic levels.

Introduction—Osteoporosis is a metabolic bone disease mainly characterized by low bone mineral density (BMD). As the precursors of osteoclasts, peripheral blood monocytes (PBMs) are supported to be important candidates for identifying genes related to osteoporosis. We performed subcellular proteomics study to identify significant membrane proteins that involved in osteoporosis.

Methods—To investigate the association between monocytes, membrane proteins, and osteoporosis, we performed label-free quantitative subcellular proteomics in 59 male subjects with discordant BMD levels, with 30 high vs. 29 low BMD subjects. Subsequently, we performed integrated gene enrichment analysis, functional annotation, and pathway and network analysis based on multiple bioinformatics tools.

Results—A total of 1070 membrane proteins were identified and quantified. By comparing the proteins' expression level, we found 36 proteins that were differentially expressed between high and low BMD groups. Protein localization prediction supported the notion that the differentially expressed proteins, P4HB ($p = 0.0021$), CD36 ($p = 0.0104$), ACTN1 ($p = 0.0381$), and ITGB1 ($p =$

Correspondence to: H.-W. Deng.

Compliance with ethical standards

Conflicts of interest

None.

Electronic supplementary material The online version of this article (doi:10.1007/s00198-017-4146-5) contains supplementary material, which is available to authorized users.

0.0385), are significant membrane proteins. Functional annotation and pathway and network analysis highlighted that P4HB, ITGB1, CD36, and ACTN1 are enriched in osteoporosis-related pathways and terms including “ECM-receptor interaction,” “Calcium ion binding,” “Leukocyte transendothelial migration,” and “reduction of cytosolic calcium levels.” Results from transcriptomic and genomic levels provided additional supporting evidences.

Conclusion—Our study strongly supports the significance of the genes P4HB, ITGB1, CD36, and ACTN1 to the etiology of osteoporosis risk.

Keywords

Network analysis; Osteoporosis; Peripheral blood monocytes; Quantitative subcellular proteomics

Introduction

Osteoporosis, a global public health problem with high heritability, is mainly characterized by low bone mineral density (BMD) [1]. Osteoporosis has become one of the most serious public health problems around the world, leading to millions of fractures annually [1, 2]. Among fractures at various skeletal sites, hip fracture is the most severe risk owing to its high prevalence, high morbidity and mortality, and excessive therapeutic cost [3, 4].

Monocytes are bone marrow-derived circulating leukocytes that can further differentiate into various cell types like macrophages and dendritic cells [5]. Peripheral blood monocytes constitute 5–10% of human leukocytes which can be easily isolated and purified from blood sample from individual subject [6]. In this study, we used the classical peripheral blood monocyte (PBM) subset (CD14⁺⁺CD16⁻) for our proteomics study. It is difficult (or nearly impossible) to collect fresh and sufficient osteoclasts from bone tissues in humans for an in vivo study like the present one, so we chose PBMs for osteoporosis study due to availability in large quantities and their important roles in osteoclastogenesis. As the sole source of osteoclasts in adult, human PBMs serve as precursors of osteoclasts and produce cytokines important for osteoclast differentiation, activation, and apoptosis and may also act as one of the most important target cells for sex hormones in bone metabolism [7, 8]. Therefore, PBMs are a major class of cells that are functionally relevant to the pathogenesis of osteoporosis, and they have been successfully used for bone-related etiology studies [8, 9].

Based on biochemical methods for isolation of subcellular components, integrated with sensitive LC-nano-ESI-MSE (liquid chromatograph-nano-electrospray ionization-mass spectrometry)-based quantitative proteomic analysis, subcellular proteome profiling can enhance coverage for each specific subcellular component while also provide more accurate localization of biological function for target proteins. Membrane proteins play significant roles in different human physiological processes including acting as signal receptors, enzymes, transport units, and cell adhesion molecules. Membrane proteins are directly or indirectly related to multiple human diseases including osteoporosis [10–12]. Missense mutations that disrupt membrane protein assembly often lead to diseases [13]. Transmembrane domains play central roles in disease-related pathways [14].

Network analysis can identify interactive correlation and topology between different proteins. Pathways/modules generated by network analysis may reflect biological processes more comprehensively and objectively than single protein/gene analysis [15]. Based on the utilization of functional information and topological information, there are quite a few approaches to perform gene enrichment analysis, pathway analysis, functional cluster analysis, and network reconstruction analysis [16]. For instance, DAVID (Database for Annotation, Visualization, and Integrated Discovery) is a popular knowledgebase composed of multiple online tools which can provide abundant functional information from multiple databases for a list of genes. However, the topological features generated from DAVID may be quite limited [16, 17]. In contrast, as a frequently used platform for network analyses, the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database engine can provide more topological information [16, 18], including protein-protein interaction. Furthermore, Cytoscape, an open source software project for data integration, network analysis, and visualization [19], is a powerful tool with hundreds of comprehensive and up-to-date applications which can provide multiple functions of network analyses.

Based on label-free quantitative proteomics and integrative network analysis, our purpose in this study is to identify significant membrane proteins and the corresponding pathways/terms involved in osteoporosis. In order to gain additional evidences, we performed validation studies in multi-omics levels including proteomics, transcriptomics, and genomics levels.

Materials and methods

Human subjects

Our study was approved by the Institutional Review Boards of University of Missouri Kansas City and Tulane University. All subjects signed consent forms before being enrolled. A total of 59 unrelated Caucasian men from Kansas City, Missouri, and its surrounding areas were enrolled in the study. They were composed of 29 subjects with low hip BMD and 30 subjects with high hip BMD (Z-score mean $\pm$ standard deviation, -1.2 ± 0.4 vs. 1.6 ± 0.6 , respectively) (Table 1), corresponding to the top 6% (low BMD) vs. 16% (high BMD) of BMD distribution in Caucasian population [20]. The hip BMD (g/cm^2) (which is the combined value of three regions including femoral neck, trochanter, and intertrochanter) were measured using a Hologic 4500 W dual energy X-ray bone densitometer for each subject as we routinely do [9].

In order to minimize the influence of environmental factors that may affect bone metabolism, several chronic diseases and conditions were excluded. Detailed exclusion criteria have been described in previous studies [21] and mainly include chronic disorders of vital organs, autoimmune-related diseases, metabolic diseases, skeletal diseases, hematopoietic diseases, lymphoreticular diseases, and other diseases or any treatment that would affect gene/protein expression.

Peripheral blood monocyte isolation

One hundred twenty milliliters peripheral blood was drawn from each subject by a certificated phlebotomist. EDTA was used as anticoagulant, and fresh blood samples were processed immediately for PBM isolation following phlebotomy by experienced technicians. The PBMCs were isolated from whole blood using density gradient centrifugation with Histopaque-1077 (Sigma-Aldrich, Cat. No.10771). PBM isolation was performed with the Monocyte Isolation Kit II (Miltenyi Biotec, Cat. No. 130-091-153) following the manufacturer's recommendation. Flow cytometry analyses showed that the purity of isolated monocyte samples was more than 90%.

Subcellular proteome extraction

Utilizing different solubilities of certain subcellular compartments in the four selected reagents, four subcellular compositions (cytosolic fraction, membrane/organelle protein fraction, nucleic protein fraction, cytoskeletal fraction) were extracted from collected PBMs using the ProteoExtract® Subcellular Proteome Extraction Kit (Emd Millipore, Cat. NO. 539790) (S-PEK) following the manufacturer's instruction.

Proteome-wide protein expression profiling

PBM membrane proteome studies were profiled using a method of 2D-nanoLC-ESI-MS/MSE, through nanoAcquity ultra performance liquid chromatography coupled with Synapt high-definition mass spectrometry (HDMS) (Waters Corporation). The MassLynx 4.1 software (Waters) monitors proteome data acquisition. The protein digests (~500 ng) were injected into a 300um × 50 mm XBridge PST C18 NanoEase Column (Waters) and separated by solvent A (water with 0.1% FA) and solvent B (acetonitrile with 0.1% FA) at a flow rate of 0.3 µl/min using a gradient of 2 h as follows: 3% B initial, 10% B at 1.0 min, 30% B at 75 min, 40% B at 90 min, 95% B at 91 min, 95% B at 95 min, and 3% B at 96 min, equilibrate thereafter till 120 min. The eluate was analyzed by HDMS under positive ion V-mode. The following parameters were set for data acquisition: collision energy: 5 V for MS and ramp 15–40 V for MSE; scan time, 0.6 s per scan. The HDMS machine was calibrated daily to ensure high accuracy (2.0 ppm for lock mass of m/z 785.8426). LC-nano-ESI-MSE original datasets were acquired from each PBM proteome digest sample. Then, the acquired MSE data were processed with the ProteinLynx Global Server (PLGS) v2.4 (Waters) using the appropriate parameters recommended by Waters Corporation. Specific information regarding database searching and protein quantification based on PLGS can be found in our previous publications [21, 22]. Protein identification was achieved by searching an integrated human protein database including 153,078 protein entries from UniProt database, NCBI Reference Sequence Database (RefSeq database), Ensembl genome database, the Vertebrate Genome Annotation (VEGA) database, and H-Invitational (H-INV) database. Total ion counts of the top three matched peptides were used to quantify each protein. We added 100 femtomols of ADH1 in each prepared sample as internal references (recommended by Waters Corporation). By comparing the ion counts of each protein with ADH1, protein quantification level was exported in femtomol and nanogram.

Protein differential expression analyses

In order to detect the significant differentially expressed proteins (DEPs), the quantitative proteomics dataset was first normalized by the housekeeping protein beta actin. Then, protein expression values were adjusted using a linear regression model by weight and height to minimize their effects of individual differences between subjects. Then, using the normalized and adjusted data, we performed Student's *t* test in subjects with low vs. high BMD. A *p* value less than 0.05 was set as the nominal cut off for statistical significance.

Functional annotation and gene enrichment analysis

The DAVID knowledgebase agglomerated tens of millions of gene/protein identifiers from a variety of public genomic resources; these identifiers improve cross-reference capability, particularly across NCBI and UniProt systems. More than 40 publicly available functional annotation sources have been comprehensively integrated and centralized by the DAVID gene clusters [17]. In order to maximize the coverage of important biological information, we imported all the unique gene symbols into this online tool for functional annotation and enrichment analysis. Comprehensive functional information for all of these genes enriched in specific terms was obtained. The results were classified as “functional annotation clustering,” “functional annotation chart,” and “functional annotation table.” We selected several terms/pathways to report based on three reasons as follow: firstly, all of them have significant *p* value ($p < 0.05$). Secondly, all of them are supported to be related with bone metabolism by extensive publications. Thirdly, each of them should contain at least one significant gene.

Pathway and network analysis and visualization based on DEPs

In order to observe the interactions between DEPs from subcellular proteomics profiling, 36 DEPs were imported into the online tool STRING for pathway and protein-protein interaction analysis. It is well-known that the interactions between proteins represent a crucial component for modern biology including proteomics studies. User-friendly and comprehensive, the STRING knowledge database has been designed to assemble, evaluate, and disseminate protein-protein association information [18]. In addition, all DEPs were imported into Cytoscape for network construction and visualization. ClueGO in conjunction with CluePedia, two easy-to-use applications based on Cytoscape_v3.2.0, greatly improved biological interpretation of large lists of genes, consequently providing a comprehensive view from different perspectives [23, 24]. The exported network gives a visualized summary of the interaction between protein and pathway/terms.

Validation studies: transcriptomics and genomics levels

To gain additional support for the candidate genes identified in our proteomics study, we assessed associations with BMD at different levels of previously published cohorts including RNA expression datasets and genome-wide association study (GWAS) dataset. At the transcriptomic level, four independent datasets were used. Cohort I contains 19 Caucasian females, including 10 high and 9 low hip BMD subjects (Affymetrix Human Genome U133A Array; GDS1287) [8]. Cohort II contains 26 Chinese females including 14 high and 12 low hip BMD subjects (Affymetrix Human Genome U133 Plus 2.0 Array; GSE7158)

[25]. Cohort III contains 73 Caucasian females including 42 high and 31 low hip BMD subjects (Affymetrix 1.0 ST arrays; GSE56814) [26]. Cohort IV contains 80 Caucasian females, including 40 high and 40 low hip BMD subjects (Affymetrix Human Genome U133A Array; GSE56815) [27]. Genes with p values less than 0.05 are purported to be DEPs between high and low BMD groups and thus are significantly associated with osteoporosis. A previous paper published by our group gave a specific introduction to these datasets [26]. At the genomic level, a previously released whole-genome sequencing, whole-exome sequencing dataset from the Genetic Factors for Osteoporosis (GEFOS) Consortium (<http://www.gefos.org/>) was used for validation studies [28]. The meta-analysis dataset contains 9 GWASs, a total of 32,965 subjects and approximately 10 million SNPs.

In this study, we performed a gene-based integrative analysis to evaluate the association between gene and hip BMD in multi-omics levels. Firstly, we determine gene-phenotype association scores separately for single omics data. For gene g , we computed q_g^k ($k = 1, \dots, D$, where D is the number of omics data) to capture the relationship between gene and phenotype in different omics levels. If several datasets were used at a single omics level (k th omics), we will first integrate the information from different datasets for each kind of omics data by a meta-analysis-based approach [29]:

$$q_g^k = -2 \sum_{j=1}^T \ln(P_g^{kj}) \quad (1)$$

where $j = 1, \dots, T$, where T is the number of cohorts for k th omics data and p_g^{kj} is the p value that assesses the association between gene g and phenotype in j th cohort in k th omics data.

q_g^k follows a chi-squared distribution with $2T$ degrees of freedom, if there is no association between phenotype and any cohort data for gene g . Thus, the p value (p_g^k) of the combined statistic q_g^k can be calculated. Secondly, we integrated the gene-phenotype association measurements ($p_g^1, \dots, p_g^D$) across D omics data into an overall score (S_{g-meta}):

$$S_{g-meta} = -2 \sum_{k=1}^D \ln(p_g^k) \quad (2)$$

S_{g-meta} has a chi-squared distribution with $2D$ degrees of freedom, if there is no association between phenotype and any omics data for gene g . Finally, the p value for the overall score S_{g-meta} was adjusted using false discovery rate (FDR) in order to correct for multiple hypothesis testing [30]. Note that if there are more than one signals (e.g., probes or SNPs) for gene g in one cohort data, the signal with most significant association with the phenotype will be chosen as a representation for gene g .

Results

Proteomics profiling

In this study, based on monocytes membrane component, among a total of 59 subjects, 1070 proteins corresponding to 516 unique genes were detected in five or more subjects by LC-nano-ESI-MSE, which were used for further analyses [21]. Among the 1070 proteins, 97.76% proteins were identified in both high and low BMD groups. Furthermore, a global normalization method was applied for data quality control in order to minimize the variability due to sample preparation or equipment conditions. We applied log2 transformation of our data for further analyses.

Identification of differentially expressed proteins

Based on adjusted protein expression value, we applied Student's *t* test to detect DEPs in low and high BMD groups. A total of 36 unique proteins were identified as DEPs (at a nominal significance level of $p < 0.05$). Among these 36 proteins, 11 proteins are upregulated and 25 proteins are down-regulated in H BMD group (Table 2).

Gene enrichment analysis and functional annotation

Corresponding to the total of 1070 proteins, 516 unique gene symbols were imported into the DAVID online tool for functional annotation and gene enrichment analysis. A total of 1405 terms in 13 categories were generated. As expected, most genes identified in this study were enriched in membrane related terms (Table 3). For the biological process and molecular function terms, with an integrated consideration of statistical results and physiological features, we highlighted several significant terms closely related to osteogenesis, bone metabolism, and skeletal physiology, such as “leukocyte transendothelial migration,” “calcium ion binding,” and “EF-hand type (calcium-binding motifs).” Individually, the DEPs included in selected terms were given (Table 4).

Pathway analysis, network reconstruction, and visualization

A total of 36 differentially expressed genes were imported into the STRING online tool for pathway and protein-protein interaction analysis. By evaluating the interaction between each differentially expressed gene, a visualized network was generated. We discarded those without interaction with other genes (Fig. 1). We highlighted several significant membrane genes such as CD36, ITGB1, ACTN1, and P4HB in this gene set (marked in Fig. 2 with red rectangles). All the DEPs were significantly enriched in 15 pathways ($FDR < 0.05$) based on the KEGG pathway database (Supplementary Table 1). In particular, those four genes play multiple and significant roles in bone-related pathways including the “ECM (extracellular matrix) receptor interaction pathway” and the “leukocyte transendothelial migration pathway.” Subsequently, network construction and visualization based on Cytoscape showed associations between DEPs and biological pathways/modules (Fig. 2). Specifically, they play central roles in numerous pathways/modules including response to elevated platelet cytosolic Ca^{2+} , the adherens junction pathway and the leukocyte transendothelial migration pathway which are thought to be related to osteogenesis, bone formation, and resorption. Integrative pathway and network analysis provided a comprehensive and intuitively visualized report of

the relationship between DEPs and biological pathways/terms, which allowed us to better understand the biological function of the DEPs.

Evidence from validation studies

To gain additional evidence for the candidate genes identified in our proteomics study, we performed validation studies at both transcriptomic and genomic levels, trying to seek validation across omics levels for those genes that are significant across sex. In RNA expression datasets, we found that CD36 is differentially expressed in the cohort I ($p = 0.0158$) and cohort II ($p = 0.027$). ACTN1 is differentially expressed in cohort I ($p = 1.48E-05$) and cohort IV ($p = 0.0001$). In the meta-analysis GWAS dataset, the most significant SNPs, rs144803881 ($p = 0.0348$), rs41272372 ($p = 0.0121$), rs150197367 ($p = 0.0089$), and rs4541004 ($p = 0.0053$) in genes P4HB, CD36, ITGB1, and ACTN1 are thought to be involved in osteoporosis (Supplementary Tables 2). Meta-analysis in multi-omics levels also indicates that the above four genes are significantly associated with osteoporosis. Since the proteomics study is based on male subjects, the validation studies contain subjects from both genders, we further verified that the above four genes are all located in euchromosome and have cross-sex effect in both genders. In conclusion, evidence from both transcriptomic and genomic levels further supported the contribution of these four genes to osteoporosis.

Discussion

Subcellular proteome profiling can enhance the coverage for each specific subcellular component, while also provide more accurate localization of biological function for target proteins. However, presenting an exact map of subcellular distribution of all proteins under different conditions is challenging, because protein subcellular localization is dynamic, without strict limitations or boundaries [31]. In particular, membrane or transmembrane proteins are often in extremely low abundance but play multiple and significant roles as signal receptors, enzymes, transport proteins, biomarkers, and cell adhesion molecules [32]. In this study, by using an efficient and reliable subcellular extraction kit in conjunction with high sensitivity and resolution LC-MS instruments, we performed comprehensive proteomics profiling in human monocyte membrane. In the cellular component analysis, as expected, most genes were enriched in membrane related terms (Table 3 and Fig. 2) which indicates that most of the corresponding proteins identified in our study are membrane proteins. Subsequently, integrated gene enrichment analysis, functional annotation, and pathway and network analysis based on multiple bioinformatics tools highlighted several significant membrane genes involved in osteoporosis including P4HB, ITGB1, ACTN1, and CD36. With an integrated consideration of statistical results and physiological features, we selected these four genes to report based on the following four reasons: firstly, these four genes are identified as differentially expressed genes, all of them have significant p value: P4HB ($p = 0.0021$), CD36 ($p = 0.0104$), ITGB1 ($p = 0.0385$), ACTN1 ($p = 0.0381$) (Table 2). Secondly, this study is a subcellular proteomics study based on membrane component of human PBMs, and these four genes are supported as significant membrane genes in cellular component analysis (Table 3). Thirdly, pathway and network analysis suggested that these four genes are involved in osteoporosis related terms and pathways such as “leukocyte

transendothelial migration,” “calcium ion binding,” and “ECM-receptor interaction” (Table 4 and Fig. 2). Fourthly, validation study at multi-omics levels indicated that these four genes are significantly associated with osteoporosis (Supplementary Table 2).

P4HB is a protein coding gene which encodes the membrane protein beta subunit of prolyl 4-hydroxylase. The major biological function of P4HB is catalyzing the formation/rearrangement of disulfide bonds; thus, it is involved in homeostatic processes. It can modify proteins' structures both inside and outside the cell and may therefore play important roles in cell adhesion and cell migration [33]. Previous publications indicated that a heterozygous missense mutation in P4HB may cause a severe bone fragility disorder called Cole-Carpenter syndrome [34]. Another previous proteomics study in Chinese premenopausal females identified P4HB as a significant gene related to osteoporosis [35]. Meanwhile, a study on the pathogenesis of osteoporosis demonstrated that PD1 (an alias of P4HB) is downregulated in osteoblasts from osteoporosis patients [36]. In this study, P4HB is differentially expressed between high and low BMD groups with a highly significant p value ($p = 0.0021$). As the sole source of osteoclast precursors for the adult peripheral skeleton, PBMs can migrate from the bloodstream to bone and differentiate into osteoclasts which play crucial roles in bone maintenance and homeostasis [6]. All above, we can demonstrate with high confidence that P4HB may be involved in bone metabolism by regulating cell migration and homeostasis.

CD36 and ITGB1 are both involved in the ECM-receptor interaction pathway. The extracellular matrix (ECM) plays a significant role in cell behavior and migration during organ or tissue morphogenesis and homeostasis [37]. CD36 is also known as leukocyte differentiation antigen CD36, which may imply a role for CD36 in the process of monocyte differentiation and therefore regulating osteoclastogenesis. CD36 has been identified as a differentially expressed gene, and several gene expression studies have found that it is closely related to osteoporosis [38, 39]. CD36 binding factors may stimulate bone resorption by osteoclasts [40]. CD36 knockout experiments in mouse model revealed that CD36KO mice tend to have low bone mass compare with wild type mouse [41]. ITGB1 is also a key gene in the leukocyte transendothelial migration pathway. ITGB1 encodes the protein integrin, which is required during the proliferation and differentiation of osteoblasts [42]. Bone morphogenetic proteins-2 (BMP2) can regulate the expression level of integrin, and integrin may affect BMP2 function in osteoclasts. Crosstalk between integrin and BMP2 plays a crucial role in osteoclast activity [43]. Consistently, as significant membrane genes identified in our study, CD36 and ITGB1 were differentially expressed between high and low BMD groups. They play central roles in all DEPs and have a close relationship with well-known bone-related pathways. In addition, evidence from the transcriptomic and genomic levels further support the contribution of these two genes to osteoporosis.

The major biological functions of ACTN1 genes include regulation of actin cytoskeleton and calcium ion binding. Proteins included in the regulation of actin cytoskeleton pathway perform diverse effects on cell activity, including changes in cell shape, migration, proliferation, and survival in many cell types [44]. A previous study showed that osteoclasts undergo internal structural changes including rearrangements of actin cytoskeleton that prepare it to resorb bone [45]. Another study suggested that the status of the actin

cytoskeleton signaling network determines the size of osteoclast-like cells during cell fusion [46]. As one of the key driver genes, ACTN1 plays a central role in the “regulation of actin cytoskeleton” pathway. Furthermore, ACTN1 is an important transmembrane protein for calcium ion binding and transportation. Calcium is the basic component of bone matrix [47]. As a regulator of calcium ion binding and transporting, ACTN1 has been associated with osteoporosis in previous publications [39]. In our study, evidence from proteomics data, RNA expression data, and GWAS data support ACTN1 as a potential genetic risk for osteoporosis.

Altogether, in our study, label-free based quantitative subcellular proteomics integrated with network analysis highlighted several candidate genes including P4HB, ITGB1, CD36, and ACTN1 that may be involved in osteoporosis. All of them are predicted as significant membrane proteins with high confidence and enriched in bone-related biological process. Since Fisher’s meta-analysis is a little bit liberal and may have potential limitation, combined *p* value from meta-analysis for the above four genes should be treated with caution (the last two columns in Supplementary Table 2). However, all of them are validated in at least one individual dataset from single omics level.

Conclusions

In our study, label-free based quantitative subcellular proteomics integrated with network analysis highlighted several candidate genes including P4HB, ITGB1, CD36, and ACTN1 that may be involved in osteoporosis. All of them are predicted as significant membrane proteins with high confidence and enriched in bone-related biological process. The results were further verified in transcriptomic and genomic levels.

BMD, bone mineral density; PBMs, peripheral blood monocytes; LC-nano-ESI-MSE, liquid chromatograph-nano-electrospray ionization-mass spectrometry; DAVID, Database for Annotation, Visualization, and Integrated Discovery; STRING, Search Tool for the Retrieval of Interacting Genes/Proteins; PBMCs, peripheral blood mononuclear cells; HDMS, high-definition mass spectrometry; PLGS, ProteinLynx Global Server; DEPs, differentially expressed proteins; GWAS, genome-wide association study

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was partially supported by and/or benefited from grants from the National Institutes of Health [P50AR055081, R21AG27110, R01AR057049, R01AR059781] and Edward G. Schlieder Endowment to Tulane University.

This manuscript was not prepared in collaboration with investigators of the Framingham Heart Study (FHS) and does not necessarily reflect the opinions of the FHS, Boston University. The GWA meta-analysis study dataset which used for supportive study in this manuscript was integrated from some individual GWAS datasets. The FHS datasets were obtained from dbGaP at <http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap> through dbGaP accession phs000007.v14.p6. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, of the Genetic Determinants of Bone Fragility study was provided by the NIA Division of Geriatrics and Clinical Gerontology and the NIA Division of Aging Biology. Assistance with phenotype harmonization, SNP

selection, data cleaning, meta-analyses, data management, and dissemination and general study coordination was provided by the PAGE Coordinating Center (U01HG004801-01).

Funding source The Framingham Heart Study is conducted and supported by the National Heart, Lung, and Blood Institute NHLBI in collaboration with Boston University (Contract No. N01-HC-25195). Funding for SHARe Affymetrix genotyping was provided by NHLBI Contract N02-HL-64278. SHARe Illumina genotyping was provided under an agreement between Illumina and Boston University. Funding support for the Framingham Bone Mineral Density datasets was provided by NIH grants R01AR/AG 41398. Funding support for the Genetic Determinants of Bone Fragility (the Indiana fragility study) was provided through the NIA Division of Geriatrics and Clinical Gerontology. Genetic Determinants of Bone Fragility is a genome-wide association studies funded as part of the NIA Division of Geriatrics and Clinical Gerontology. Support for the collection of datasets and samples were provided by the parent grant, Genetic Determinants of Bone Fragility (P01-AG018397). Funding support for the genotyping which was performed at the Johns Hopkins University Center for Inherited Diseases Research was provided by the NIH NIA. The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human Services through contracts N01WH22110, 24152, 32100-2, 32105-6, 32108-9, 32111-13, 32115, 32118, 32119, 32122, 42107-26, 42129-32, and 44221. WHI PAGE is funded through the NHGRI Population Architecture Using Genomics and Epidemiology (PAGE) network (Grant No. U01HG004790).

References

1. Zhao LJ, Liu XG, Liu YZ, et al. Genome-wide association study for femoral neck bone geometry. *J Bone Miner Res.* 2010; 25:320–329. [PubMed: 20175129]
2. Liu YJ, Shen H, Xiao P, Xiong DH, Li LH, Recker RR, Deng HW. Molecular genetic studies of gene identification for osteoporosis: a 2004 update. *J Bone Miner Res.* 2006; 21:1511–1535. [PubMed: 16995806]
3. Cummings SR, Melton LJ. Epidemiology and outcomes of osteoporotic fractures. *Lancet.* 2002; 359:1761–1767. [PubMed: 12049882]
4. Cauley JA, Thompson DE, Ensrud KC, Scott JC, Black D. Risk of mortality following clinical fractures. *Osteoporos Int.* 2000; 11:556–561. [PubMed: 11069188]
5. Geissmann F, Manz MG, Jung S, Sieweke MH, Merad M, Ley K. Development of monocytes, macrophages, and dendritic cells. *Science.* 2010; 327:656–661. [PubMed: 20133564]
6. Gordon S, Taylor PR. Monocyte and macrophage heterogeneity. *Nat Rev Immunol.* 2005; 5:953–964. [PubMed: 16322748]
7. Narni-Mancinelli E, Soudja SM, Crozat K, Dalod M, Gounon P, Geissmann F, Lauvau G. Inflammatory monocytes and neutrophils are licensed to kill during memory responses in vivo. *PLoS Pathog.* 2011; 7:e1002457. [PubMed: 22241983]
8. Liu YZ, Dvornyk V, Lu Y, Shen H, Lappe JM, Recker RR, Deng HW. A novel pathophysiological mechanism for osteoporosis suggested by an in vivo gene expression study of circulating monocytes. *J Biol Chem.* 2005; 280:29011–29016. [PubMed: 15965235]
9. 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. Uzer G, Fuchs RK, Rubin J, Thompson WR. Plasma and nuclear membranes convey mechanical information to regulate mesenchymal stem cell lineage. *Stem Cells.* 2016
11. Ma M, Chen X, Lu L, Yuan F, Zeng W, Luo S, Yin F, Cai J. Identification of crucial genes related to postmenopausal osteoporosis using gene expression profiling. *Aging Clin Exp Res.* 2015
12. Notomi T, Kuno M, Hiyama A, Ezura Y, Honma M, Ishizuka T, Ohura K, Yawo H, Noda M. Membrane depolarization regulates intracellular RANKL transport in non-excitable osteoblasts. *Bone.* 2015; 81:306–314. [PubMed: 26211991]
13. Ng DP, Poulsen BE, Deber CM. Membrane protein misassembly in disease. *Biochim Biophys Acta.* 2012; 1818:1115–1122. [PubMed: 21840297]
14. Logan CY, Nusse R. The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol.* 2004; 20:781–810. [PubMed: 15473860]
15. Swa HL, Shaik AA, Lim LH, Gunaratne J. Mass spectrometry based quantitative proteomics and integrative network analysis accentuates modulating roles of annexin-1 in mammary tumorigenesis. *Proteomics.* 2014
16. Wu X, Hasan MA, Chen JY. Pathway and network analysis in proteomics. *J Theor Biol.* 2014

17. Sherman BT, Huang da W, Tan Q, Guo Y, Bour S, Liu D, Stephens R, Baseler MW, Lane HC, Lempicki RA. DAVID knowledgebase: a gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis. *BMC Bioinf.* 2007; 8:426.
18. Franceschini A, Szklarczyk D, Frankild S, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 2013; 41:D808–D815. [PubMed: 23203871]
19. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003; 13:2498–2504. [PubMed: 14597658]
20. Hoiberg M, Nielsen TL, Wraae K, Abrahamsen B, Hagen C, Andersen M, Brixen K. Population-based reference values for bone mineral density in young men. *Osteoporos Int.* 2007; 18:1507–1514. [PubMed: 17534538]
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(M111):011700. [PubMed: 21817168]
22. Zhang L, Liu YZ, Zeng Y, et al. Network-based proteomic analysis for postmenopausal osteoporosis in Caucasian females. *Proteomics.* 2015
23. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z, Galon J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics.* 2009; 25:1091–1093. [PubMed: 19237447]
24. Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics.* 2013; 29:661–663. [PubMed: 23325622]
25. Lei SF, Wu S, Li LM, et al. An in vivo genome wide gene expression study of circulating monocytes suggested GBP1, STAT1 and CXCL10 as novel risk genes for the differentiation of peak bone mass. *Bone.* 2009; 44:1010–1014. [PubMed: 19223260]
26. Liu YZ, Zhou Y, Zhang L, Li J, Tian Q, Zhang JG, Deng HW. Attenuated monocyte apoptosis, a new mechanism for osteoporosis suggested by a transcriptome-wide expression study of monocytes. *PLoS One.* 2015; 10:e0116792. [PubMed: 25659073]
27. Chen XD, Xiao P, Lei SF, et al. Gene expression profiling in monocytes and SNP association suggest the importance of the STAT1 gene for osteoporosis in both Chinese and Caucasians. *J Bone Miner Res.* 2010; 25:339–355. [PubMed: 19594299]
28. Zheng HF, Forgetta V, Hsu YH, et al. Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature.* 2015; 526:112–117. [PubMed: 26367794]
29. Tyekucheva S, Marchionni L, Karchin R, Parmigiani G. Integrating diverse genomic data using gene sets. *Genome Biol.* 2011; 12:R105. [PubMed: 22018358]
30. Benjamini Y, Hochberg Y. Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J R Stat Soc B Met.* 1995; 57:289–300.
31. Drissi R, Dubois ML, Boisvert FM. Proteomics methods for subcellular proteome analysis. *FEBS J.* 2013; 280:5626–5634. [PubMed: 24034475]
32. Qiu X, Zhang H, Lai YR. Quantitative targeted proteomics for membrane transporter proteins: method and application. *AAPS J.* 2014; 16:714–726. [PubMed: 24830943]
33. Bi S, Hong PW, Lee B, Baum LG. Galectin-9 binding to cell surface protein disulfide isomerase regulates the redox environment to enhance T-cell migration and HIV entry. *Proc Natl Acad Sci U S A.* 2011; 108:10650–10655. [PubMed: 21670307]
34. Rauch F, Fahiminiya S, Majewski J, Carrot-Zhang J, Boudko S, Glorieux F, Mort JS, Bachinger HP, Moffatt P. Cole-Carpenter syndrome is caused by a heterozygous missense mutation in P4HB. *Am J Hum Genet.* 2015; 96:425–431. [PubMed: 25683117]
35. Deng FY, Liu YZ, Li LM, et al. Proteomic analysis of circulating monocytes in Chinese premenopausal females with extremely discordant bone mineral density. *Proteomics.* 2008; 8:4259–4272. [PubMed: 18924182]
36. Hino S, Kondo S, Yoshinaga K, et al. Regulation of ER molecular chaperone prevents bone loss in a murine model for osteoporosis. *J Bone Miner Metab.* 2010; 28:131–138. [PubMed: 19760141]

37. Lukashev ME, Werb Z. ECM signalling: orchestrating cell behaviour and misbehaviour. *Trends Cell Biol.* 1998; 8:437–441. [PubMed: 9854310]
38. Balla B, Kosa JP, Kiss J, et al. Different gene expression patterns in the bone tissue of aging postmenopausal osteoporotic and non-osteoporotic women. *Calcified Tissue Int.* 2008; 82:12–26.
39. Trost Z, Trebse R, Prezelj J, Komadina R, Logar DB, Marc J. A microarray based identification of osteoporosis-related genes in primary culture of human osteoblasts. *Bone.* 2010; 46:72–80. [PubMed: 19781675]
40. Carron JA, Wagstaff SC, Gallagher JA, Bowler WB. A CD36-binding peptide from thrombospondin-1 can stimulate resorption by osteoclasts in vitro. *Biochem Biophys Res Commun.* 2000; 270:1124–1127. [PubMed: 10772961]
41. Kevorkova O, Martineau C, Martin-Falstra L, Sanchez-Dardon J, Brissette L, Moreau R. Low-bone-mass phenotype of deficient mice for the cluster of differentiation 36 (CD36). *PLoS One.* 2013; 8
42. Sununliganon L, Singhatanadgit W. Highly osteogenic PDL stem cell clones specifically express elevated levels of ICAM1, ITGB1 and TERT. *Cytotechnology.* 2012; 64:53–63. [PubMed: 21866310]
43. Lai CF, Cheng SL. Alphavbeta integrins play an essential role in BMP-2 induction of osteoblast differentiation. *J Bone Miner Res.* 2005; 20:330–340. [PubMed: 15647827]
44. Yamaguchi H, Condeelis J. Regulation of the actin cytoskeleton in cancer cell migration and invasion. *BBA-Mol Cell Res.* 2007; 1773:642–652.
45. Boyle WJ, Simonet WS, Lacey DL. Osteoclast differentiation and activation. *Nature.* 2003; 423:337–342. [PubMed: 12748652]
46. Takito J, Otsuka H, Yanagisawa N, Arai H, Shiga M, Inoue M, Nonaka N, Nakamura M. Regulation of osteoclast multinucleation by the actin cytoskeleton signaling network. *J Cell Physiol.* 2014
47. Fleisch H. Physiology and biochemistry of bone formation. *Klin Wochenschr.* 1966; 44:360–363. [PubMed: 4861945]

Fig. 1.

Protein-protein interaction between DEPs. Each note represents one DEP. *Different line colors* represent different types of evidences for the associations. Notes with no interaction with others were discarded. From the network, we highlighted several significant membrane genes such as CD36, ITGB1, ACTN1, and P4HB as hub genes in this gene set

Fig. 2. Network construction and visualization based on DEPs. *Each dot* represents one DEP. *Each circle* represents a pathway/term. All the pathways/terms were functionally grouped and highlighted with different colors. The membrane DEPs of interest (ITGB1, ACTN1, and CD36) are marked with *red rectangles*

Table 1

Basic characteristics of subjects

Group	Caucasian male with low BMD	Caucasian male with high BMD
Sample size	29	30
Age (years)	40.3 (7.6)	41.1 (7.5)
Height (cm)	173.7 (7.3)	178.2 (7.1)
Weight (kg)	75.1 (15.7)	103.7 (23.2)
Hip BMD Z-score	−1.2 (0.4)	1.6 (0.6)

Data were presented as mean (SD). Hip BMD is presented as Z-score. Z-score is defined as the number of standard deviations a subject's BMD from the average value of the reference group (with the same age, gender, and ethnicity)

Differentially expressed proteins

Table 2

UniProt ID	Gene symbols	p value	Ratio H/L	UniProt ID	Gene symbols	p value	Ratio H/L
Q93084	ATP2A3	0.0008	1.1	P14770	GP9	0.0226	0.5
B3KTQ9	P4HB	0.0021	1.8	P42166	TMPO	0.0231	0.5
P54819	AK2	0.0030	0.6	P09486	SPARC	0.0240	0.3
Q5TFQ8	SIRPB1	0.0034	1.1	P04899	GNAI2	0.0344	0.7
Q9NZ08	ERAP1	0.0036	2.7	P13073	COX4I1	0.0359	1.1
B3KS31	TUBB6	0.0059	0.3	O00264	PGRMC1	0.0367	0.5
P63261	ACTG1	0.0065	0.5	P12814	ACTN1	0.0381	0.7
Q5JTV8	TOR1AIP1	0.0084	2.0	P05556	ITGB1	0.0385	0.4
E9PLT1	CD36	0.0104	1.4	O43491	EPB41L2	0.0391	0.1
A6QL64	ANKRD36	0.0116	0.2	Q92928	RAB1C	0.0403	1.7
B1PS43	MYH11	0.0127	0.2	Q15365	PCBP1	0.0409	0.5
P21333	FLNA	0.0141	0.3	C9JW96	PHB	0.0411	1.3
P60660	MYL6	0.0168	0.5	P61026	RAB10	0.0434	1.2
P10515	DLAT	0.0175	0.7	P27824	CANX	0.0448	0.2
Q13349	ITGAD	0.0187	6.6	Q3LAZ2	HLA	0.0451	0.4
P46940	IQGAP1	0.0211	0.7	P61006	RAB8A	0.0456	0.5
E9PK25	CFL1	0.0218	0.4	P08567	PLEK	0.0462	0.5
Q4JCP9	ACTN3	0.0224	0.3	B4DQU5	RAB11A	0.0473	0.3

With normalized expression data, Student's *t* test was applied to identify DEPs. A *p* value less than 0.05 was set as the nominal cutoff for significance. 36 proteins reached the nominal significant statistical level. The ratio H/L represents the fold change of protein expression level between low and high BMD groups, which reflected that proteins were upregulated (>1) or downregulated (<1) in the high BMD group

Table 3

Major membrane-related terms from cellular component analysis

Terms	Number of gene	<i>p</i> value	DEPs enriched
GO:0005886-plasma membrane	166	0.0003	P4HB, RAB8A, GNAI2, PLEK, PHB, RAB1C, ACTN1, ACTN3, ITGB1, FLNA, IQGAP1, GP9, EPB41L2, CD36, RAB11A, ITGAD, RAB10
GO:0044459-plasma membrane part	109	7.94E-05	RAB8A, CD36, PLEK, PHB, RAB1C, ACTN1, ITGAD, ACTN3, RAB10, ITGB1, GP9
GO:0005624-membrane fraction	46	0.0012	P4HB, CD36, PLEK, GNAI2, PGRMC1, ITGB1
SP_PIR_KEYWORDS: membrane	178	0.0248	P4HB, RAB8A, PHB, TOR1AIP1, RAB1C, ITGB1, CANX, IQGAP1, GP9, CD36, ATP2A3, PGRMC1, ERAP1, RAB11A, ITGAD, TMPO, RAB10

The first column lists major membrane-related terms from cellular component analysis; the second column lists the number of genes enriched in each specific term; the third column shows the corresponding *p* value of each term; the last column lists the DEPs enriched in each specific term

Table 4

Biological process analysis highlighted DEPs enriched in osteoporosis related terms

Term	<i>p</i> value	DEPs
Leukocyte transendothelial migration	1.19E-11	ITGB1, ACTN1, ACTG1, GNAI2
Calcium ion binding	0.0061	MYL6, ATP2A3, ACTN1, ITGAD, ACTN3, SPARC, CANX
EF-hand type	0.0132	MYL6, ACTN1, ACTN3, SPARC
ECM-receptor interaction	0.0031	CD36, ITGB1, GP9

The first column lists the specific terms related to osteoporosis; the second column lists the *p* value of each term; the third column lists those DEPs enriched in each term

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript