ORIGINAL ARTICLE

JBMR®

Prioritization of Genes Relevant to Bone Fragility Through the Unbiased Integration of Aging Mouse Bone Transcriptomics and Human GWAS Analyses

Serra Kaya,[1] Charles A Schurman,[1,2] Neha S Dole,[1] Daniel S Evans,[3] and Tamara Alliston[1,2]

[1]Department of Orthopaedic Surgery, University of California, San Francisco, CA, USA
[2]UC Berkeley-UCSF Graduate Program in Bioengineering, San Francisco, CA, USA
[3]California Pacific Medical Center Research Institute, San Francisco, CA, USA

ABSTRACT

Identifying new genetic determinants of bone mineral density (BMD) and fracture promises to yield improved diagnostics and therapies for bone fragility. However, prioritizing candidate genes from genome-wide screens can be challenging. To overcome this challenge, we prioritized mouse genes that are differentially expressed in aging mouse bone based on whether their human homolog is associated with human BMD and/or fracture. Unbiased RNA-seq analysis of young and old male C57BL/6 mouse cortical bone identified 1499, 1685, and 5525 differentially expressed genes (DEGs) in 1, 2, and 2.5-year-old bone, relative to 2-month-old bone, respectively. Gene-based scores for heel ultrasound bone mineral density (eBMD) and fracture were estimated using published genome-wide association studies (GWAS) results of these traits in the UK Biobank. Enrichment analysis showed that mouse bone DEG sets for all three age groups, relative to young bone, are significantly enriched for eBMD, but only the oldest two DEG sets are enriched for fracture. Using gene-based scores, this approach prioritizes among thousands of DEGs by a factor of 5- to 100-fold, yielding 10 and 21 genes significantly associated with fracture in the two oldest groups of mouse DEGs. Though these genes were not the most differentially expressed, they included *Sost*, *Lrp5*, and others with well-established functions in bone. Several others have, as yet, unknown roles in the skeleton. Therefore, this study accelerates identification of new genetic determinants of bone fragility by prioritizing a clinically relevant and experimentally tractable number of candidate genes for functional analysis. Finally, we provide a website (www.mouse2human.org) to enable other researchers to easily apply our strategy. © 2022 American Society for Bone and Mineral Research (ASBMR).

KEY WORDS: AGING; GENETIC RESEARCH; EPIDEMIOLOGY; DISEASES AND DISORDERS OF/RELATED TO BONE; OSTEOCYTES

Introduction

Adult bone remodels continuously throughout life to maintain a healthy skeleton. However, aging causes imbalanced bone remodeling and progressive bone loss, leading to disrupted bone microarchitecture and fragility fractures. Men and women older than 65 years experience more than 9 million fractures annually worldwide,[1] severely impacting quality of life[2-4] and increasing risk of death.[5-7] Bone fragility results from deficits in bone mass and bone quality. Current methods to diagnose osteopenia and fracture risk primarily rely on measuring declining bone mineral density (BMD), which is a measure of bone mass and a strong predictor for incident fractures, but half of fragility fractures occur in individuals with clinically normal bone mass.[8] Although existing therapeutics are designed to increase bone mass, an unmet need remains for developing new diagnostics and therapies that target bone quality. Therefore, an improved understanding of the mechanisms responsible for age-related bone fractures is needed.

BMD is a highly heritable and polygenic trait.[9] Recent genome-wide association studies (GWAS) have identified many loci associated with BMD and fracture risk.[10-15] Establishing the relationship of these loci to biological mechanisms of disease requires identifying which of the variants and genes at an association locus play a causal role in bone fragility and understanding their biological function. The functional role of some GWAS-identified genes has been revealed through extensive molecular, cellular, and in vivo studies using genetically modified animals.[10,11,13,15,16] Nonetheless, using currently available experimental approaches, many years would be needed to functionally test the large number of GWAS-identified BMD and fracture candidate genes. New strategies to prioritize functional analyses, for example, by harnessing GWAS-based associations, could accelerate the discovery of clinically relevant mechanisms of bone fragility.

Received in original form February 6, 2021; revised form December 20, 2021; accepted January 17, 2022.
Address correspondence to: Tamara Alliston, PhD, University of California, San Francisco, Department of Orthopaedic Surgery, 513 Parnassus Avenue, S1155, San Francisco, CA 949041, USA. E-mail: tamara.alliston@ucsf.edu
Additional Supporting Information may be found in the online version of this article.
Journal of Bone and Mineral Research, Vol. 37, No. 4, April 2022, pp 804–817.
DOI: 10.1002/jbmr.4516

A large portion of age-related fragility fractures occur in individuals with clinically normal bone mass,[8] a feature significantly controlled by osteoblast and osteoclast activity.[17] Bone-embedded osteocytes orchestrate osteoblast and osteoclast function to regulate bone mass. Osteocytes also control bone quality, in part through the process of perilacunar/canalicular remodeling, which maintains the osteocyte lacunocanalicular network.[18-21] Several lines of evidence point to the importance of osteocyte function in age-related bone fragility, including the deterioration of the lacunocanalicular network, loss of bone mass and quality with age,[22-26] and the role of osteocyte cellular senescence in age-related bone loss.[27-31] However, the mechanisms that compromise osteocyte function in aging and the extent to which they contribute to bone fragility remain unclear.

As opposed to human populations, inbred laboratory mouse strains offer consistent environmental control and minimal genetic diversity, facilitating the analysis of genetic determinants of age-related bone fragility. Furthermore, skeletal changes in aging mouse cortical bone resemble changes in aging humans, including the increased periosteal diaphysis, decreased cortical bone mass, and increased cortical porosity.[32] Therefore, we examined the effect of age on gene expression in C57BL/6 mouse cortical bone, which is mainly populated by osteocytes. High-throughput RNA sequencing (RNA-seq) of mouse cortical bone at multiple ages afforded highly accurate and reproducible transcriptome profiling, yielding hundreds to thousands of differentially expressed genes (DEGs). However, prioritizing candidate genes for further analysis is difficult and relies on a priori knowledge.

To address the challenge of prioritizing among candidate genes, derived from either genome-wide mouse RNA-seq or large human GWAS, we integrated data generated by these two approaches. Genes that are differentially expressed with age in mouse cortical bone were prioritized using an unbiased computational approach based on their relevance to bone fragility in human GWAS. This approach identifies genes with well-known functional roles in human bone fragility and predicts new functional roles in the skeleton for genes that possess significant associations with bone fracture in humans.

Materials and Methods

Mice

Male C57BL/6 mice from the same colony were allowed to age with free access to water and Envigo/Teklad (Indianapolis, IN, USA) Global diet 2918 (irradiated) rodent chow throughout the study. Light/dark cycles were 12 hours long, from 6 a.m. to 6 p.m. The room temperature range was 70°F to 72°F with the humidity range of 30% to 70%. Animals were housed in standard cages with 5 mice/cage. Mice were euthanized at 2 months (2m), 11 months (1y), 23 months (2y), and 30 months (2.5y) ($n = 3$ for 2m; $n = 4$ for others) and tissues were collected from all mice simultaneously with approval from the Institutional Animal Care and Use Committee. The human equivalent of these mouse ages are listed below[33]:

Mouse age (months)	3–6	10–14	18–24	36
Human age equivalents (years)	20–30	38–47	56–69	94

RNA isolation and sequencing

Soft tissues, periosteum, and epiphyses were removed from dissected humeri before bone marrow removal by centrifugation. Cells within cortical bone prepared in this manner are primarily osteocytes, with minimal osteoblast contribution to RNA.[20,34] Prepared bones were immediately snap-frozen in liquid nitrogen. RNA was isolated as described.[35-37] Briefly, an Omni homogenizer was used to homogenize the frozen bones in QIAzol Lysis Reagent, and subsequently 10 to 15 μg total RNA was extracted and purified with the RNeasy Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. RNA quantity was assessed using a Nanodrop Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and RNA quality (260/280 ratio) was determined using a Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). For each sample that passed the Bioanalyzer quality-control threshold (RNA Integrity Number >7), 500 ng of total RNA were used for RNA library preparation using Illumina (San Diego, CA, USA) TruSeq Stranded mRNA sample preparation kit, according to the manufacturer's protocol. 50-bp single-end reads were sequenced using an Illumina HiSeq 4000 at the UCSF Functional Genomics Core (San Francisco, CA, USA). Sequencing yielded 745 million total reads with an average of 50 million read/sample. FastQC was used for quality control of raw RNA-seq data.[38] Reads containing adapters and low-quality reads were removed from the raw data. An average of 78% single-end reads were aligned uniquely to the Ensembl mouse GRCm38.78 reference genome using STAR (v. 2.5.2b).[39]

RNA sequencing analysis

Pairwise differential expression analysis

To determine up- or downregulated gene expression changes in 1y, 2y, and 2.5y bone to 2-month-old bone, differential expression analysis for three age comparison sets (2m-1y, 2m-2y, and 2m-2.5y) was performed by subsetting the read counts to the two age time points via the DESeq2 package (v.1.24.0)[40] in R (v.3.6.3). DESeq2 uses a negative binomial generalized linear model (GLM) for differential analysis and applies the Wald test for the significance of GLM coefficients. The Benjamini-Hochberg (BH) false discovery rate (FDR) method was used for p value adjustment.[41] Genes were considered differentially expressed according to the significance criteria of FDR < 0.1. For the principal component analysis, a set of 500 genes with the highest row variance was selected to assess overall similarity between the samples. ggplot2 R-package was used to plot the samples based on the first two principal components. Heatmaps were generated using the complexheatmap R-package to visualize the Z-scores of the gene expression patterns of DEGs in each pairwise comparison age group. Volcano plots were used to visualize the differential expression of genes. Downregulated, upregulated, and fracture-associated genes were specifically labeled in volcano plots using ggplot2 R-package.

Time course differential expression analysis

To identify genes that have changes in their expression in any direction across any of the time points, a likelihood ratio test (LRT) was performed using the DESeq2 package. The full model included an intercept term and age coded as a factor. The reduced model included an intercept term only. DEGs from LRT with BH adjusted p values (FDR) ≤ 0.05 were deemed

significant. To identify clusters of genes that have a common change of expression pattern over time, a clustering tool called *degPatterns* in the DEGreport package in R was used.[42] Normalized counts data of DEGs (FDR < 0.05) from LRT was entered into *degPatterns*. Briefly, a correlation test is performed using Kendall rank correlation coefficient to calculate all pairwise gene expression and to create a distance matrix. Then, the divisive hierarchical clustering, using the *diana* function in R, is performed to calculate gene expression profile similarities among samples. Finally, *degPatterns* produced nine different clustering groups, with expression changes over time illustrated in violin plots (Fig. 5*A*). Clusters with less than 15 genes are not included. Each gene belongs to only one clustering group. Moreover, the line in the graphs show the trend of the change in the gene's expression.

Pathway analysis

To gain insights of the underlying biological processes altered by aging in bone, pathway over-representation analysis (ORA) was conducted using the *compareCluster* function in clusterProfiler package (v.3.12.0).[43] Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways that have BH-adjusted p values < 0.05 were considered significant.[44] To visualize the most significant KEGG pathways as a network, the *cnetplot* function was used. These network maps show genes that belong to each pathway, as well as the presence of common genes in multiple pathways. Supplemental Figs. S3 and S4 further illustrate the effect of age on Wnt signaling and parathyroid hormone (PTH) synthesis pathways, each of which was colored to show gene expression changes in the pairwise age comparisons using pathview.[45,46]

UK Biobank GWAS results of estimated bone mineral density and fracture

We made use of previously published GWAS results of bone mineral density estimated from heel ultrasounds (eBMD) and fracture based on the UK Biobank study available for download from the GEFOS website.[47] The UK Biobank study is a large, prospective study of ~500,000 participants in the UK of ages 40 to 69 years with phenotype data and single-nucleotide polymorphism (SNP) genotypes. Baseline measurements were recorded over 4 years from 2006 to 2010. The eBMD trait used a quantitative heel ultrasound to determine heel bone mineral density. The fracture outcome was based on fracture cases at any skeletal site that were identified from hospital-based fracture diagnosis and questionnaires (self-reported fractures). There were 426,824 UK Biobank participants with eBMD and 53,184/373,611 fracture cases/controls.[10] Genetic associations were available for 13,753,401 and 13,970,184 variants for eBMD and fracture, respectively.

Enrichment analysis of mouse DEGs using human UK Biobank GWAS result

Three steps were taken to perform gene set enrichment analysis using UK Biobank GWAS from human eBMD and fracture. First, human homologs were identified for all mouse genes. Second, gene-based scores were estimated from GWAS results using MAGMA (v.1.07b).[48] Third, gene set enrichment was performed using MAGMA to determine whether human homologs of mouse DEGs for each age comparison were enriched for significant gene-based scores.

Mouse-human homolog mapping

To analyze human orthologs of the experimentally identified mouse DEGs, all mouse Ensembl gene IDs were converted to mouse Entrez IDs using *org.Mm.eg.db* (v.3.10). Homology between mouse and human genes was determined based on NCBI's Homologene resource. Homologene IDs for mouse-human homologs were from the file *HOM_MouseHumanSequence.rpt* downloaded on April 28, 2020.[49] For situations with multiple human paralogs for a single mouse gene, we did not have a priori knowledge of which human paralog might be the most relevant, so we included all human paralogs within a homologene ID.

Gene-based scores

GWAS results provide single-variant associations, not gene-based scores. To generate gene-based scores from GWAS results, MAGMA software[48] was used. All gene-based scores were based on the most significant variant within 50 kb upstream and downstream of the gene region, resulting in 100 kb non-coding gene region in total, after adjusting for several potential confounders, such as gene size, gene density, and linkage disequilibrium (LD) between variants, as described.[48] Adjustment was performed using an adaptive permutation procedure. Thus, gene-based empirical p values can differ from GWAS reported p values because of permutation.

Gene set enrichment

MAGMA was then used for enrichment analysis specifically to determine if human homologs of aging mouse bone DEG sets (from pairwise and time course DE analysis) were enriched for associations with human eBMD or fracture.[48] MAGMA's gene set enrichment performs a one-sided test to determine whether gene-based scores in a gene set are greater than all gene-based scores across the genome. A probit transformation of the empirical gene-based p values generates Z-scores, and the Z-scores are used as the dependent variable in a generalized linear model. The generalized linear model includes covariate adjustment for inter-gene correlation using 1000 Genomes genotypes, gene size, and number of SNPs. After gene set enrichment analysis, the significance of gene-based scores within a gene set were examined and the Bonferroni method was used to control family-wise error rate. Gene-based scores with adjusted p value < 0.05 were deemed significant.

Statistical analysis

Statistical analysis was performed in R (v.3.6.3). Specific packages that were used to perform these analyses are described in their respective Materials and Methods section. MAGMA analysis software was used to perform gene-based associations and gene set enrichment analysis (https://ctg.cncr.nl/software/). Multiple hypothesis testing correction was used in differential expression analysis, KEGG pathway analysis (Benjamini-Hochberg), and gene-based score significance within a gene set (Bonferroni).

Data availability

RNA-seq data generated from mouse cortical bone predominantly populated by osteocytes are deposited in the Sequence Read Archive (SRA) database of the National Center for Biotechnology Information (NCBI) under the BioProject PRJNA695408.

Results

Unbiased identification of differentially expressed genes in aging mouse bone

To identify novel molecular determinants of age-associated bone fragility, we performed unbiased transcriptome-wide analysis of RNA isolated from osteocyte-enriched cortical bone of young (2 months) and aged mice (1 year, 2 years, and 2.5 years). Individual samples from each mouse clustered by age using the first two principal components (PCs) (Fig. 1*A*). The young (2 months) group clustered separately from all other aged groups along PC1, whereas the older groups, 1 year, 2 years, and 2.5 years, formed distinct clusters that spread along PC2, which shows an aging trajectory.

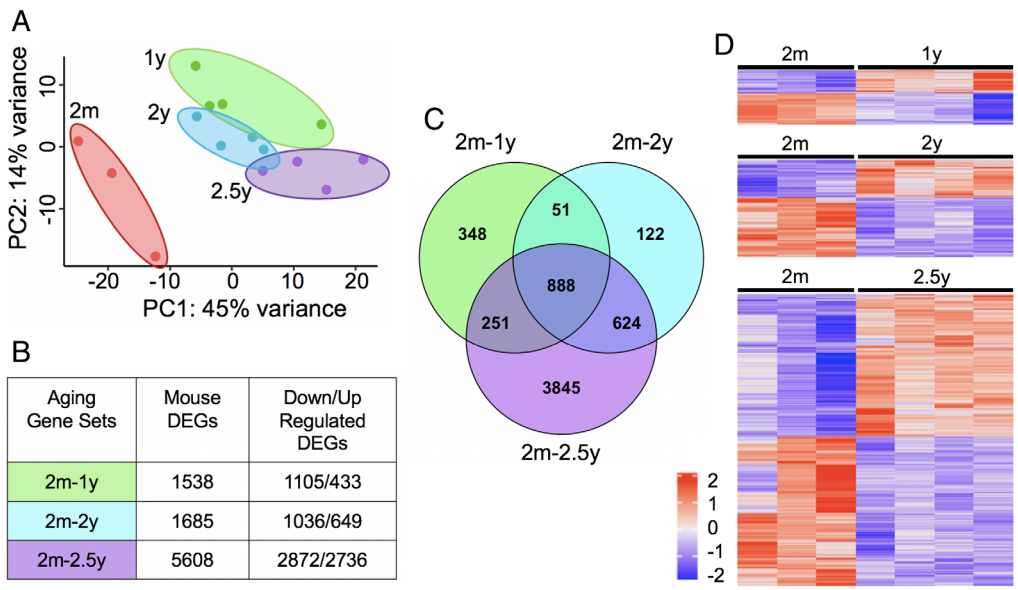
Analysis of gene expression changes across age, relative to 2-month-old bone, showed an increasing number of DEGs with age, such that comparison of 1-year- to 2-month-old bone (2m-1y) identified 1538 DEGs, 2-year- to 2-month-old bone (2m-2y) identified 1685 DEGs, and 2.5-year- to 2-month-old bone (2m-2.5y) identified 5608 DEGs (FDR < 0.1, Fig. 1*B*). A total of 14.5% of differentially expressed genes (888 DEGs) are common in all aged groups (Fig. 1*C*). The heatmap shows the segregation of the young and old bone gene expression (Fig. 1*D*). Although the number of downregulated DEGs increases with age, the proportion of downregulated versus upregulated DEGs decreases with age (1 year = 71.8%, 2 years = 61.5%, 2.5 years = 51.2%; Fig. 1*B*, *D*). The detailed results of the differential expression analysis of RNA-seq on mouse cortical bone at each age are available in Supplemental Table S1.

To focus our analysis of aging mouse bone DEGs, we performed pathway analysis. In spite of the increasingly divergent transcriptional profile in aging bone, pathway analysis revealed few differences in the age-regulated pathways at these three time points. DEGs in each aging gene set were most enriched for the same KEGG pathways, including those related to cell adhesion, ECM-receptor interactions, axon guidance, and protein digestion and absorption (BH-adjusted $p < 0.05$; Fig. 2, Supplemental Fig. S1, and Supplemental Table S2). Although bone fragility increases with age,[50] pathway analysis did not implicate a progressive age-related deregulation of a specific pathway in these comparisons. Moreover, a curated gene signaling pathway annotated for BMD or fracture does not exist; thus, we are not able to test for enrichment of such a pathway.

Aging mouse bone DEGs are enriched for human genetic associations with eBMD and fracture

In an effort to prioritize aging mouse bone DEGs for further study, we sought to determine whether their human homologs are associated with human eBMD and fracture. However, several challenges limit the use of mouse RNA-seq data to accurately mine GWAS data for genes that are clinically relevant to human bone mineral density, bone fracture, or bone fragility. Published GWAS results can include thousands of variant associations in a gene region, but typically gene-based scores are not provided. This can present a major obstacle for laboratory researchers attempting to assess whether a gene harbors sufficient evidence for association with a trait.

To address this need, we used an approach to easily identify human homologs of mouse genes, followed by determination of gene-based scores for eBMD and fracture using human GWAS



Aging Gene Sets	Mouse DEGs	Down/Up Regulated DEGs
2m-1y	1538	1105/433
2m-2y	1685	1036/649
2m-2.5y	5608	2872/2736

Fig. 1. Unbiased RNA sequencing analysis results for aging male mouse bone. (*A*) Principal component analysis shows the clustering of samples from each age group of 2m, 1y, 2y, and 2.5y bone, with respect to the first two principal components based on their similarity ($n = 3$ for 2m; $n = 4$ for 1y, 2y, and 2.5y). (*B*) Table shows the number of differentially expressed genes as a result of RNA-seq analysis (FDR < 0.1), as well as the number of downregulated and upregulated genes for each age group (1y, 2y, and 2.5y) compared with 2m. (*C*) The Venn diagram shows the number of common and individual differentially expressed genes for each of these aging gene sets 2m-1y, 2m-2y, and 2m-2.5y. (*D*) The heatmaps show the *Z*-scores of gene expression patterns for DEGs in each aging gene set; red and blue colors indicate upregulated and downregulated genes, respectively, compared with the 2-month group. Columns represent the samples and rows represent the genes.

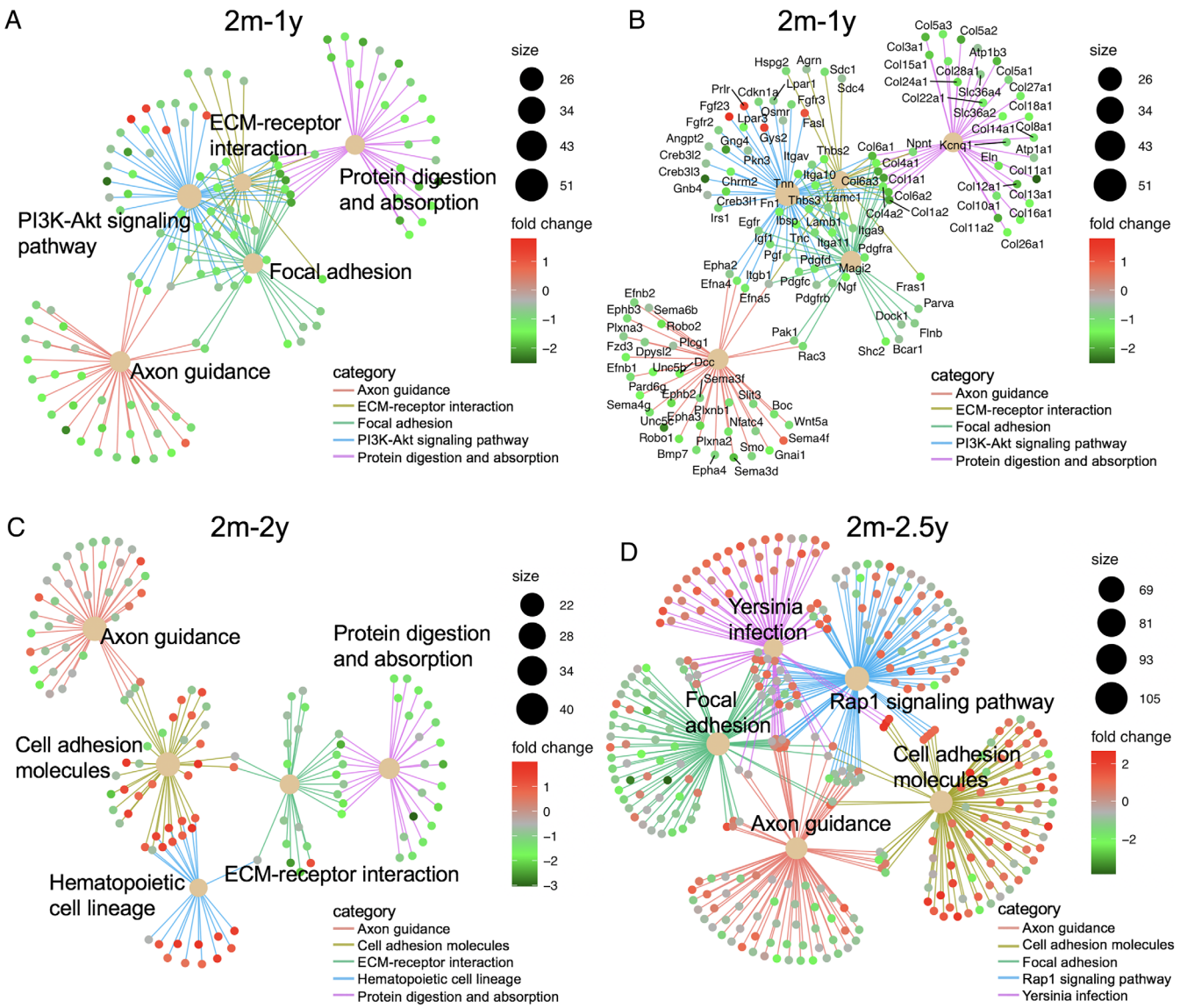


Fig. 2. Pathway analysis in aging bone. To prioritize specific differentially expressed genes (DEGs) for further mechanistic studies, we performed over-representation analysis using all DEGs identified in the three pairwise aging gene sets. The five most enriched mouse KEGG pathways are shown as beige nodes with edges connecting genes to corresponding pathways for 2m-1y (*A*), 2m-2y (*C*), and 2m-2.5y (*D*). For 2m-1y (*B*), gene names are also displayed. Many of the same pathways are significant in all three aging gene sets (Benjamini-Hochberg adjusted p value < 0.05).

results,[10] and finally gene set enrichment analysis. We tested whether human homologs of aged mouse bone DEGs are enriched for genome-wide associations with human eBMD or fracture, using the UK Biobank cohort of up to 426,824 participants, which is one of the biggest such analysis to date. For each aging mouse bone DEG, we identified the corresponding human homologs (Table 1 and Supplemental Table S3). Gene-based scores were calculated through MAGMA, with histograms of gene-based *Z*-scores for eBMD and fracture shown in Supplemental Fig. S2. Gene-set enrichment was then performed, also using MAGMA, to determine if any of the three aging gene sets was significantly enriched for gene-based scores with eBMD or fracture. Although all three gene sets were enriched for eBMD ($p < 0.005$ for all; Fig. 3*A* and Table 1), only the two oldest gene sets (2m-2y and 2m-2.5y) were significantly enriched for fracture ($p < 0.05$; Fig. 3*B* and Table 1). This is consistent with the age-dependent loss of bone quality with age,[50] in addition to the loss of bone mineral density.

Prioritization of aging mouse bone DEGs based on human eBMD and fracture associations

The significance of each gene-based score within the three aging gene sets was determined by correcting for multiple testing of the number of genes in each gene set (Table 2 and Supplemental Table S3). Using this approach, we identified that 20% of DEGs were significantly associated with eBMD. A much smaller fraction of these DEGs (<1%) was associated with fracture. All fracture-associated genes were also associated with eBMD (Table 2 and Supplemental Table S3). For the 2m-2.5y gene set of 5608 mouse

Table 1. MAGMA UK Biobank GWAS Enrichment Results for Human eBMD and Fracture

Aging gene sets	Mouse DEGs[a]	No. of genes in MAGMA analysis[b]	Bone mineral density[c]		Fracture[d]	
			MAGMA enrichment p value[e]	MAGMA significant genes[f]	MAGMA enrichment p value[e]	MAGMA significant genes[f]
2m-1y	1538	1275	3.6×10^{-4}	320	0.101	12
2m-2y	1685	1389	7.2×10^{-4}	337	0.030	10
2m-2.5y	5608	4671	3.1×10^{-3}	1092	0.031	21
Cluster 1	970	902	6.6×10^{-7}	256	0.024	10
Cluster 4	753	684	0.01	175	0.024	10

[a]Number of differentially expressed genes (DEGs) from mouse RNA-seq analysis (FDR < 0.1).
[b]Not all mouse DEGs have human homologs or gene-based GWAS scores.
[c]Bone mineral density is estimated from heel ultrasound (eBMD) from published UK Biobank GWAS.
[d]Published UK Biobank GWAS results for fracture.
[e]MAGMA enrichment is performed using human homologs of mouse DEGs in each aging gene set.
[f]Bonferroni correction is applied for each gene for each aging gene set. Significance level: adjusted p value <0.05.

DEGs, 1092 human homologs were significantly associated with eBMD, whereas only 21 were associated with fracture (Bonferroni adjusted p value < 0.05; Table 1).

For the oldest two aging gene sets (2m-2y and 2m-2.5y) that are both enriched for fracture (Fig. 3*B*), we evaluated the differential expression of individual mouse bone genes that were significantly associated with human fracture (adjusted p value < 0.05). Volcano plots reveal DEGs from mouse RNA-seq that are repressed (blue) or induced (red) with aging (Fig. 4*A*, *B*). The 10 genes from the 2m-2y gene set (Fig. 4*A*) and the 21 genes from the 2m-2.5y gene set (Fig. 4*B*) that are significantly associated with human fracture are highlighted in gold and labeled. *Wnt16, Fam3c, Cped1, Lrp5, Mindy4, Rnd1*, and *Glis2* were significant in both 2m-2y and 2m-2.5y sets, resulting in a total of 24 fracture-associated genes. The distribution of these fracture-associated genes shows that the clinically relevant genes are not simply those with the highest fold change or significant differences in RNA-seq analysis. The detailed information about the human fracture-associated genes for 2m-2y and 2m-2.5 gene sets are listed in Table 2, which includes MAGMA gene-based results, as well as UK Biobank GWAS p values and corresponding SNP rsIDs for each gene. We identified that the most significant fracture-associated SNPs are identical for several genes (rs2908007 for *Wnt16*, *Fam3c*, and *Cped1*; rs2430691 for *Mgp*, *Arhgdib*, and *Erp27*). These genes are located close to each other on the same chromosome, and it is highly possible that the same SNP contributes to the significance of multiple DEGs. Since *Wnt16* and *Mgp* have the smallest adjusted gene-based scores with established roles in the skeleton,[51,52] we chose not to further consider *Fam3c, Cped1, Arhgdib*, and *Erp27* (Table 2). All significant fracture-associated significant genes were also significantly associated with eBMD (adjusted p value < 0.05; Table 2).

The genes that are significantly associated with fracture in the 2m-2y and 2m-2.5y gene sets through our in-silico approach include *Wnt16*, *Lrp5*, and *Sost*, all of which have well-known functions in bone and clinical relevance to fracture.[23,51,53-55] On the other hand, several others have yet to be investigated in the skeleton, including *Ldlrad4*, *Atxn7l3*, *Mpp3*, and others. Although these genes are significantly associated with fracture at the whole-genome level (genome-wide p value < 5×10^{-8}; Table 2), their functional role in bone fragility remains to be determined. Importantly, this gene set enrichment analysis also identified novel genes that are associated with fractures that do not achieve genome-wide significance, including *Glis2*, *Smco3*, *Spon2*, and *Nbeal1* (Table 2). Therefore, our in silico approach identified novel genes associated with fracture in humans, some of which were not reported in the main table of published GWAS because of the stringent genome-wide significance threshold.

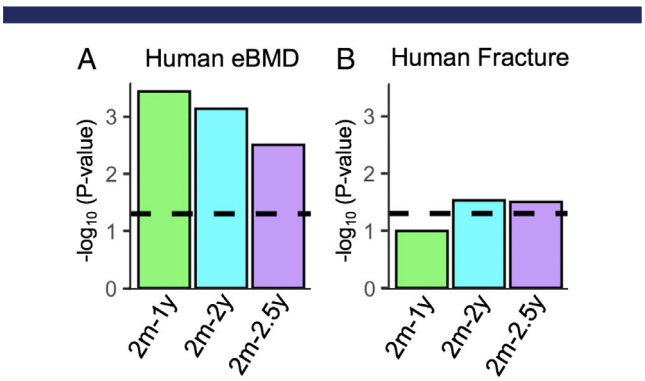


Fig. 3. Mouse differentially expressed genes (DEGs) enriched for associations with eBMD and fracture using human GWAS results. MAGMA was performed using human homologs of 1538, 1685, and 5608 DEGs in mouse RNA-seq for 2m-1y, 2m-2y, and 2m-2.5y, respectively, and GWAS results of human eBMD and fracture in the UK Biobank published at the genetic factors for osteoporosis database. The dashed line shows the significance threshold (p value < 0.05). Although all aging mouse gene sets were enriched for eBMD (*A*), only the 2m-2y and 2m-2.5y set were significantly enriched for fracture (*B*).

Analysis of change across all four time points prioritizes genes that are downregulated with age

To complement the pairwise analysis of young and old mice above, we investigated the behavior of mouse bone gene expression profiles over time using all four age groups of mice. Modeling gene expression across four time points using the likelihood ratio test identified 3306 genes that were differentially expressed using FDR-adjusted p value threshold of <0.05 (Supplemental Table S1). Divisive hierarchical cluster analysis identified nine distinct temporal patterns of gene expression (Fig. 5*A* and Supplemental Table S1), with the largest number of genes in clusters 1, 2, and 4 (970, 1025, and 753 genes, respectively).

Table 2. MAGMA-Identified Aging Mouse Bone DEGs with Variants Associated With Human Fracture Using UK Biobank GWAS Results for Pairwise Comparison Age Groups of 2m-2y (2y) and 2m-2.5y (2.5y) (Some Genes Are Shared in Both 2m-2y and 2m-2.5y)

	Human		Mouse		Fracture				Bone mineral density				
Age	Gene ID	Entrez ID	Gene ID	Entrez ID	MAGMA p value[a]	Adjusted p value[b]	GWAS p value[c]	GWAS SNP rsID[d]	MAGMA p value[a]	Adjusted p value[b]	GWAS p value[c]	GWAS SNP rsID[d]	Mouse gene expression[e]
2y 2.5y	WNT16[f*]	51384	Wnt16	93735	1.7×10^{-25}	2.4×10^{-22} 8.0×10^{-22}	7.2×10^{-50}	rs2908007	1.0×10^{-50}	1.4×10^{-47} 4.7×10^{-47}	$<5 \times 10^{-323}$	56 SNPs	Down
2.5y	RSPO3	84870	Rspo3	72780	2.5×10^{-25}	1.2×10^{-21}	2.5×10^{-25}	rs7741021	1.0×10^{-50}	4.7×10^{-47}	$<5 \times 10^{-323}$	32 SNPs	Down
2y 2.5y	FAM3C[f]	10447	Fam3c	27999	3.3×10^{-24}	4.6×10^{-21} 1.6×10^{-20}	7.2×10^{-50}	rs2908007	1.3×10^{-43}	1.8×10^{-40} 6.1×10^{-40}	$<5 \times 10^{-323}$	43 SNPs	Down
2y 2.5y	CPED1[f]	79974	Cped1	214642	5.9×10^{-16}	8.2×10^{-13} 2.8×10^{-12}	7.2×10^{-50}	rs2908007	1.0×10^{-50}	1.4×10^{-47} 4.7×10^{-47}	$<5 \times 10^{-323}$	111 SNPs	Down
2y	SOST	50964	Sost	74499	1.3×10^{-14}	1.8×10^{-11}	1.3×10^{-14}	rs4793022	1.0×10^{-50}	1.4×10^{-47}	3×10^{-114}	rs2741856	Up
2.5y	HROB	78995	BC030867	217216	4.9×10^{-12}	2.3×10^{-8}	3.0×10^{-12}	rs408282	1.5×10^{-27}	6.9×10^{-24}	2.1×10^{-68}	rs228746	Up
2.5y	MEOX1	4222	Meox1	17285	8.0×10^{-12}	3.8×10^{-8}	$4.6. \times 10^{-14}$	rs6503470	1.0×10^{-50}	4.7×10^{-47}	7.2×10^{-99}	rs9902329	Down
2.5y	ATXN7L3	56970	Atxn7l3	217218	4.5×10^{-11}	2.1×10^{-7}	1.9×10^{-11}	rs567375191	1.0×10^{-50}	4.7×10^{-47}	2.5×10^{-62}	rs2240226	Up
2.5y	FGFRL1	53834	Fgfrl1	116701	1.6×10^{-10}	7.5×10^{-7}	7.8×10^{-13}	rs78520297	1.8×10^{-21}	8.5×10^{-18}	5×10^{-193}	rs4505759	Down
2.5y	MPP3	4356	Mpp3	13384	4.1×10^{-10}	1.9×10^{-6}	9.2×10^{-14}	rs540771367	1.0×10^{-50}	4.7×10^{-47}	5.2×10^{-97}	rs78311416, rs185157719	Down
2.5y	LDLRAD4	753	Ldlrad4	52662	9.1×10^{-8}	4.3×10^{-4}	1.5×10^{-10}	rs4430817	1.7×10^{-18}	8.2×10^{-15}	2×10^{-155}	rs2040189	Down
2y 2.5y	LRP5	4041	Lrp5	16973	1.0×10^{-7}	1.4×10^{-4} 4.8×10^{-4}	7.5×10^{-9}	rs11228240	1.0×10^{-50}	1.4×10^{-47} 4.7×10^{-47}	6×10^{-101}	rs11228240	Down
2.5y	SFRP4	6424	Sfrp4	20379	1.1×10^{-6}	5.3×10^{-3}	4.9×10^{-9}	rs1047785	1.0×10^{-50}	4.7×10^{-47}	8×10^{-139}	rs1717740	Down
2y 2.5y	MINDY4	84182	Mindy4	330323	1.3×10^{-6}	1.8×10^{-3} 6.0×10^{-3}	1.9×10^{-9}	rs28362709	7.5×10^{-25}	1.1×10^{-21} 3.5×10^{-21}	3.6×10^{-88}	rs28362709	Down
2.5y	C11orf49	79096	1110051M20Rik	228356	1.4×10^{-6}	6.6×10^{-3}	2.5×10^{-8}	rs1055447	1.0×10^{-50}	4.7×10^{-47}	4.6×10^{-79}	rs7118404	Down
2.5y	EPDR1	54749	Epdr1	105298	1.7×10^{-6}	7.8×10^{-3}	4.9×10^{-9}	rs1047785	1.0×10^{-50}	4.7×10^{-47}	1×10^{-168}	rs939667	Down
2y 2.5y	RND1	27289	Rnd1	223881	4.1×10^{-6}	5.7×10^{-3} 0.019	4.2×10^{-8}	rs2926799	7.0×10^{-18}	9.8×10^{-15} 3.3×10^{-14}	3.3×10^{-76}	rs117962372	Up
2.5y	MGP[g*]	4256	Mgp	17313	4.3×10^{-6}	0.020	3.9×10^{-8}	rs2430691	2.9×10^{-18}	1.4×10^{-14}	1.8×10^{-18}	rs2430689	Down
2.5y	ARHGDIB[g]	397	Arhgdib	11857	5.6×10^{-6}	0.026	3.9×10^{-8}	rs2430691	1.5×10^{-17}	6.9×10^{-14}	1.8×10^{-18}	rs2430689	Up
2.5y	ERP27[g]	121506	Erp27	69187	7.1×10^{-6}	0.033	3.9×10^{-8}	rs2430691	4.1×10^{-14}	1.9×10^{-10}	1.8×10^{-18}	rs2430689	Up
2y 2.5y	GLIS2	84662	Glis2	83396	9.6×10^{-6}	0.013 0.045	1.6×10^{-7}	rs72766545	6.4×10^{-15}	8.9×10^{-12} 3.0×10^{-11}	6.4×10^{-15}	rs148092711	Down
2.5y	NBEAL1	65065	Nbeal1	269198	1.0×10^{-5}	0.048	1.0×10^{-7}	rs2351774	2.1×10^{-15}	1.0×10^{-47}	3.3×10^{-21}	rs189962749	Down
2y	SMCO3	440087	Smco3	654818	1.5×10^{-5}	0.020	4.0×10^{-7}	rs1861698	3.2×10^{-17}	4.4×10^{-14}	2.9×10^{-18}	rs2445411, rs10846065	Down
2y	SPON2	10417	Spon2	100689	1.7×10^{-5}	0.023	1.7×10^{-7}	rs4974591	1.0×10^{-50}	1.4×10^{-47}	8.3×10^{-80}	rs113727613	Down

[a]MAGMA gene-based score based on empirical min p value (Supplemental Table S3).
[b]Multiple testing correction for each mouse DEG using Bonferroni correction method. Significance level: adjusted p value < 0.05 (Supplemental Table S3).
[c]Published UK Biobank GWAS result of most significant SNP p value across a gene region. This p value is different from MAGMA empirical p value.
[d]rsID numbers for the most significant SNP identified in UK Biobank GWAS; in cases where multiple significant SNP exist, the number of SNPs are listed.
[e]Gene expression levels from mouse RNA-seq: down = gene expression is downregulated for old animals compared with the young ones (2 years versus 2 months); up = vice versa (Supplemental Table S1).
[f,g]These genes have the same most significant GWAS SNPs with WNT16([f*]) and MGP([g*]) and will not be considered for further mechanistic study.

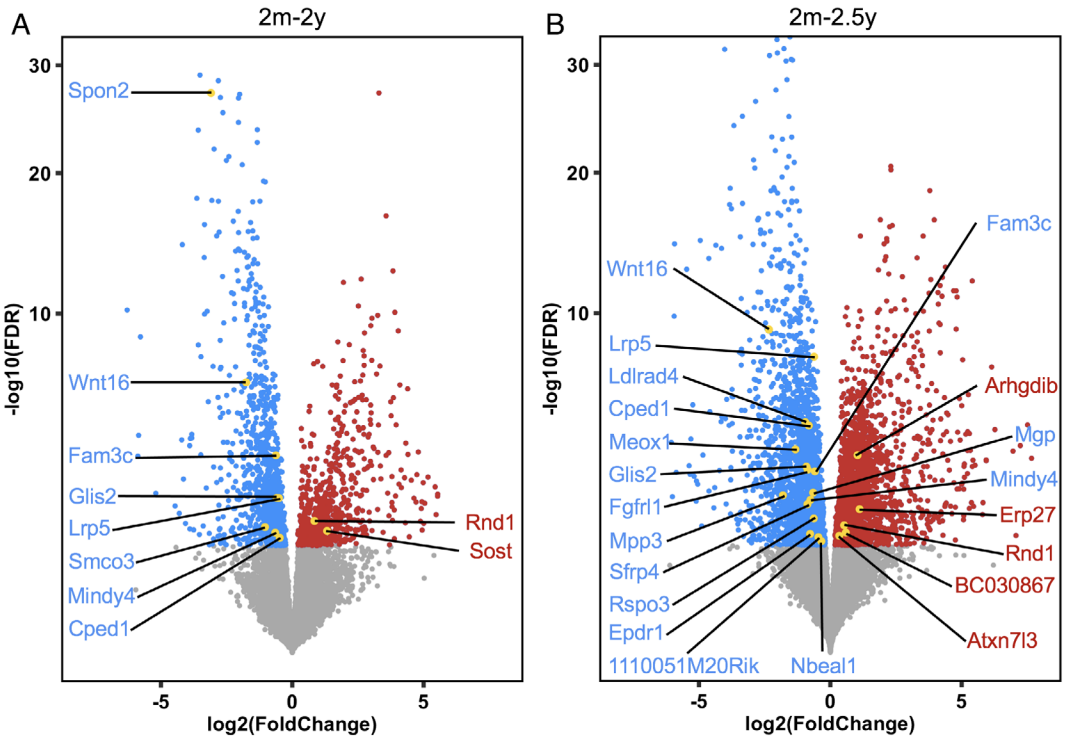


Fig. 4. Significant human fracture-associated gene expression levels and corresponding FDR values in mouse RNA-seq for 2m-2y and 2m-2.5y gene sets. Volcano plots show the DEGs as red (upregulated) and blue (downregulated) for aged groups compared with the 2m group (significance criteria: FDR < 0.1). Non-DEGs are colored as gray (FDR > 0.1). The oldest two aging gene sets of 2m-2y (*A*) and 2m-2.5y (*B*) that were enriched for fracture were further investigated for their individual adjusted gene-based scores (Bonferroni adjusted *p* value < 0.05). The mouse homologs of fracture-associated significant human genes were colored as gold dots in the plots and gene names labeled as blue for downregulated genes and red for upregulated genes. The square root transformation was used for the *y*-axis to visualize the fracture-associated genes in the volcano plots.

We next sought to determine if human homologs of the genes in any of these nine clusters were enriched for human eBMD and fracture, based on UK Biobank GWAS results. Only two of the nine clusters, clusters 1 and 4, were enriched for both eBMD and fracture (Fig. 5*B* and Table 1). Both clusters have a pattern of downregulated gene expression over time. Specifically, cluster 4 consists of genes that are consistently downregulated from 2 months until 2.5 years. Cluster 1, the second-largest cluster, contains genes that are downregulated from 2 months to 1 year but upregulated until 2.5 years. Although cluster 2 is the single largest cluster, this group of upregulated genes did not enrich for either eBMD or fracture. Therefore, genes that are downregulated with age have the most significant associations with human eBMD and fracture.

Whether KEGG pathway analysis is conducted on gene sets comparing two ages (Fig. 2) or on cluster gene sets 1 and 4 from the continuous analysis of gene expression over four ages, similar pathways emerge, including axon guidance, focal adhesions, ECM-receptor interactions (Fig. 6*A* and Supplemental Table S2). Given that most studies do not have access to four aging time points, we sought to determine which pairwise grouping best predicted the effects observed across time. Because cluster 1 and 4 genes are mostly downregulated, we repeated pathway analysis separately for the downregulated and upregulated DEGs from each pairwise group. Ten of the 14 pathways identified in the downregulated pairwise gene sets are shared with those identified for clusters 1 and 4 (Fig. 6*B*), but none of the pathways identified in the upregulated pairwise gene sets are common (Fig. 6*C*). Therefore, separate analysis of upregulated and downregulated genes may improve the ability to discriminate age-dependent effects on bone, even when comparing two age groups.

Of the human homologs of genes in clusters 1 and 4, 25% are significantly associated with human eBMD and 1% are significantly associated with fracture. Many of the fracture-associated genes were also identified in the pairwise analysis of two ages (Table 3). Several additional fracture-associated (as well as eBMD-associated) genes were identified in the continuous analysis of four ages, including *Tmem175, Cacnb3, Cyp20a1, Raph1*, and *C11orf24* from cluster 1 and *Mark3, Sox6, Samd5*, and *Wnt5b* from cluster 4. Except for *Tmem175* and *Cacnb3*, the others do not pass the stringent GWAS significance criteria for fracture but emerge as significant in this gene set enrichment analysis.

Website: www.mouse2human.org

A major challenge to laboratory researchers using unbiased genomic approaches is making informed decisions about which novel genes to pursue further in time- and resource-intensive laboratory-based analyses. Even after applying statistical methods to control for false discovery or pathway analyses, researchers must rely on prior knowledge as they sift through long lists of DEGs. To enable other investigators to identify and prioritize analysis of mouse genes based on their potential clinical relevance to human bone mineral density and fracture, we created a browsable and

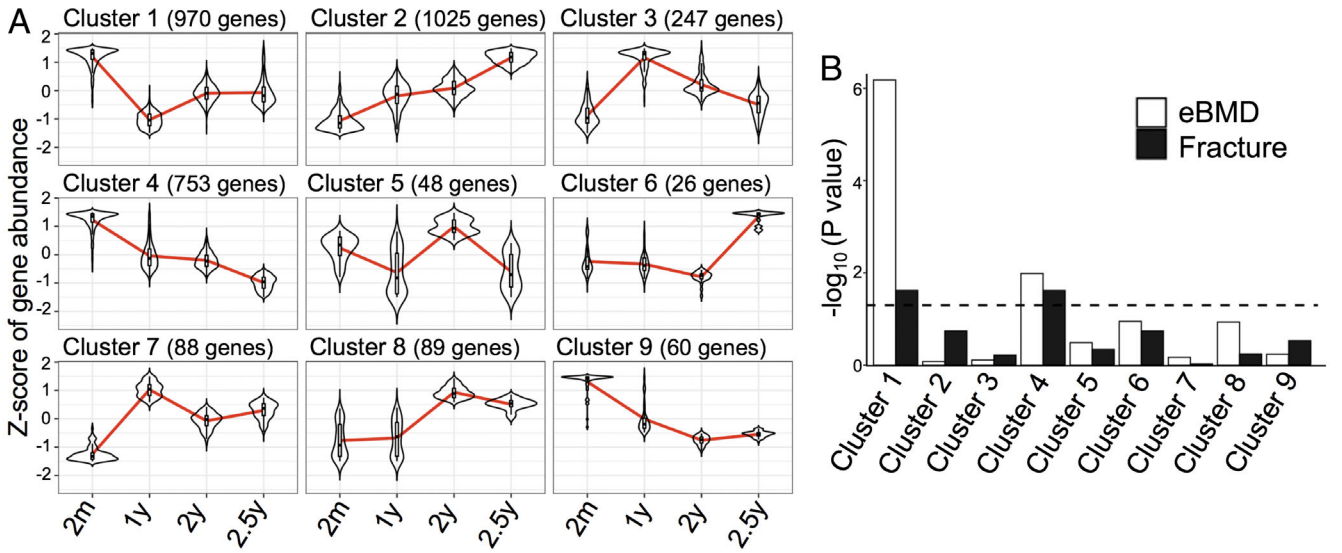


Fig. 5. Mouse differential gene expression across four age group clusters in nine distinct temporal patterns, two of which are enriched for human eBMD and fracture. (*A*) Likelihood ratio test using gene expression from mouse RNA-seq over all time points identifies 3306 differentially expressed genes (FDR <0.05). DEGs cluster into shared gene expression profiles using *degPatterns* function, with each gene belonging to one clustering group and clusters of fewer than 15 genes not included. (*B*) MAGMA gene set enrichment analysis for human eBMD and fracture associations using all nine clusters identifies significant associations for cluster 1 and cluster 4. The dashed line shows p value = 0.05.

interactive web-based tool: www.mouse2human.org. Users can search any mouse or human gene of interest to find its gene-based score for human eBMD and fracture in the UK Biobank, using the same gene scoring methods described in Materials and Methods. A batch query of a group of genes will generate p values adjusted for multiple testing using the Bonferroni and Benjamini-Hochberg (1995) FDR method. The code to perform gene set enrichment analysis, as shown in Figs. 3 and 5, is made available at this website. Though this analysis is too computationally intensive to operate on the website, users can perform enrichment analysis in their local computational environment to find enrichment for custom gene sets for eBMD or fracture.

Discussion

Using an unbiased approach, we prioritize study of clinically relevant genes that are differentially expressed in aging bone, based on their association with human eBMD or fracture. Combined use of aged mouse bone transcriptome and UK Biobank GWAS analyses prioritized 20% of DEGs with significant associations with eBMD and <1% DEGs with significant associations for fracture. This unbiased selection of <1% of DEGs in aging mouse bone motivates further mechanistic study to elucidate their functional relevance in bone fragility. Among the fracture-associated genes, several of them are well-known regulators of skeletal health such as *Sost* and *Wnt16*. Other genes possess strong associations with human fracture as well, such as *Mpp3* or *Mindy4*, but have no reported role in the skeleton. Given the therapeutic use of Sclerostin antibodies in osteoporosis, the other genes identified here are promising candidates for mechanistic study to identify new molecular participants, and potential drug targets, to treat age-related bone fragility. To make this resource publicly available, we generated a computational resource: www.mouse2human.org, so that others can prioritize candidate genes based on their relevance to human eBMD and fracture.

Aging studies typically rely on two time points, comparing young versus old. Here, we analyzed cortical bone gene expression across four time points, the longest of which is older than prior analyses.[56,57] Relative to pairwise analysis of two time points, the continuous analysis of all four time points increased statistical power in identifying differentially expressed genes and revealed nine distinct temporal patterns of gene expression. Reassuringly, the results of pathway analysis using the pairwise and continuous approaches were largely similar and also correspond to prior analysis of gene expression in 19-week- versus 19-month-old mouse cortical bone,[56] with an age-dependent decrease in Wnt signaling and increase in immunomodulatory cytokine signaling. Separation of up- and downregulated genes improved the concordance of results from pathway analysis of two ages to those acquired using a continuous analysis of four ages, suggesting that this strategy can improve the prioritization of genes even in studies with two time points.

GWAS typically focus on variant associations that are significant after correcting for multiple testing of genome-wide variants; however, it is widely acknowledged that this practice leads to a large number of false negatives.[58] A gene can contain thousands of variants, each with their own trait association. Although a single genome-wide significant (p value $< 5 \times 10^{-8}$) variant provides clear evidence of trait association, a look-up of genes within a predetermined gene set does not necessitate such a stringent multiple test correction. However, within a single gene region, the appearance of multiple correlated variants in linkage disequilibrium (LD) makes it difficult to accurately estimate the number of independent hypothesis tests being conducted and ultimately to correct for multiple testing. Extending this task to multiple genes requires additional statistical consideration. Diving deeper into specific genetic loci motivated by

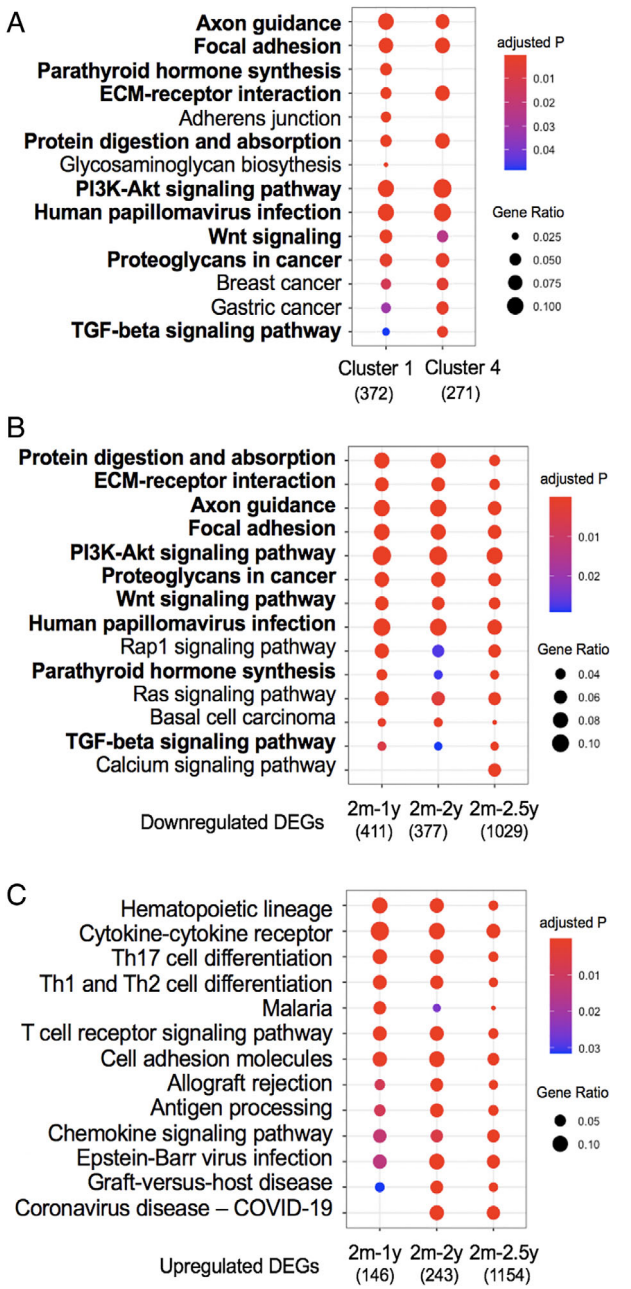


Fig. 6. Pathway analysis shows similar significant KEGG pathways between cluster groups 1 and 4 and downregulated genes from pairwise age comparisons. The *compareCluster* function performed over-representation analysis using all DEGs from clusters 1 and 4 (*A*), and downregulated (*B*) and upregulated (*C*) DEGs separately as identified in the three pairwise aging gene sets. The most enriched significant mouse KEGG pathways are listed. The clusters 1 and 4, which are dominated by downregulated genes, show similar significant pathways as downregulated DEGs in all three aging gene sets of pairwise comparisons. Pathways common to multiple gene sets are shown in bold font. The dot colors indicate the adjusted *p* values (Benjamini-Hochberg) and the dot sizes indicate the number of genes in the DEG list associated with the KEGG pathway. Gene ratio is a/b where a is the number of unique genes that are common in a specific KEGG pathway and the related DEG list, and b (given in parentheses under each DEG list name) is the number of unique genes that are common across all KEGG pathways and the related DEG list.

laboratory findings can potentially reveal novel important genes that might not have been prioritized by statistical evidence alone. Recently, the Musculoskeletal Knowledge Portal (MSK-KP) at mskkp.org was developed, which hosts human GWAS results for many traits, including eBMD and fracture.[59] This resource allows easy access to the GWAS of variants for a gene of interest, but challenges remain. Hundreds of variant associations may exist for each human gene, but a single summary gene-based score is not available. In addition, the corresponding mouse homolog is not provided, which can present a roadblock for mouse laboratory researchers.

Such in silico gene prioritization approaches are valuable in this study, and similar methods have recently been employed by others.[60,61] However, implementing these strategies requires significant computational, genetic, and epidemiologic expertise. The work herein has three important features. First, we provide easy-to-use mouse-to-human homolog matching that includes orthologs and paralogs, which is important when there is no a priori knowledge to select a particular paralog. Second, we precalculate the computationally intensive gene-based scores that require permutation, and we make these gene-based scores available on the www.mouse2human.org website, which provides an easy tool to determine one gene-based score that encompasses all variants for each gene. Knowing which genes have relevance to human eBMD and fracture will facilitate translation of experimental laboratory results toward clinical studies. Importantly, the value of this website does not depend on the aging mouse RNA-seq results used in the present study. Any mouse or human study result can be used to determine relevance of individual genes, or genes within a gene set, to human eBMD and fracture. Third, we performed an enrichment analysis to determine if the DEGs from young and old mouse bone RNA-seq are significantly associated with human eBMD or fracture. Although a gene can be significant for eBMD or fracture, a collection of genes as a whole might not be enriched for eBMD or fracture. The 2m-1y gene set of 1389 genes contain 12 genes that are significantly associated with fracture, but the whole gene set is not significantly enriched for fracture (Table 1 and Supplemental Table S3). This focused our attention on the genes in the other gene sets (2m-2y, 2m-2.5y) that were significantly enriched for fracture. These results confirm the previous experimental studies that demonstrate that aging impacts the quality of bones and, as a result, increases the chance of fracture.[50,62-64]

After reviewing each of the fracture-associated genes for the 2m-2y and 2m-2.5y gene sets, we identified *Wnt16, Sost, Lrp5*, and *Mgp* that have well-known functions in bone[51,52,55,65] and are causally implicated in osteoporosis.[23,51,54,66,67] In addition, *Sfrp4, Meox1*, and *Epdr1* have been shown in several functional studies to play important roles in osteoclasts or osteoblasts[68-70] and to participate in human diseases: *Sfrp4* deficiency causes Pyle's syndrome[71] and *Meox1*-deficient zebrafish resemble human Klippel-Feil Syndrome.[72] Notably, none of these fracture-associated genes, including *Sost* or *Lrp5*, belong to the most differentially expressed gene category according to the mouse RNA-seq results (Fig. 4*A*, *B*). Their false discovery rate values are very close to the significance level, which shows that without the enrichment study, these genes might have been overlooked in an unbiased analysis.

Given that these findings verify that this in silico approach successfully prioritizes mouse genes according to their relevance to human fracture, it is exciting to consider the potential skeletal roles of the other lesser-known genes identified here. Four

Table 3. MAGMA-Identified Aging Mouse Bone Degs With Variants Associated With Human Fracture Using UK Biobank GWAS Results for Cluster 1 and Cluster 4

	Human		Mouse		Fracture				Bone mineral density				
Cluster #	Gene ID	Entrez ID	Gene ID	Entrez ID	MAGMA *p* value[a]	Adjusted *p* value[b]	GWAS *p* value[c]	GWAS SNP rsID[d]	MAGMA *p* value[a]	Adjusted *p* value[b]	GWAS *p* value[c]	GWAS SNP rsID[d]	Common tables
Cluster 1	FAM3C[f]	10447	Fam3c	27999	3.3×10^{-24}	3.0×10^{-21}	7.2×10^{-50}	rs2908007	1.3×10^{-43}	1.2×10^{-40}	$<5 \times 10^{-323}$	43 SNPs	2y and 2.5y
	TMEM175	84286	Tmem175	72392	2.9×10^{-10}	2.6×10^{-7}	1.8×10^{-12}	rs76048788	5.7×10^{-35}	5.2×10^{-32}	2×10^{-170}	rs115134980	-
	LRP5	4041	Lrp5	16973	1.0×10^{-7}	9.1×10^{-5}	7.5×10^{-9}	rs11228240	1.0×10^{-50}	9.0×10^{-48}	6×10^{-101}	rs11228240	2y and 2.5y
	MINDY4	84182	Mindy4	330323	1.3×10^{-6}	1.2×10^{-3}	1.9×10^{-9}	rs28362709	7.5×10^{-25}	6.8×10^{-22}	3.6×10^{-88}	rs28362709	2y and 2.5y
	CACNB3	784	Cacnb3	12297	1.3×10^{-6}	1.2×10^{-3}	4.2×10^{-8}	rs2926799	1.0×10^{-50}	9.0×10^{-48}	5.0×10^{-77}	rs79626908	-
	CYP20A1	57404	Cyp20a1	77951	2.0×10^{-6}	1.8×10^{-3}	4.2×10^{-8}	rs11676242	3.3×10^{-21}	3.0×10^{-18}	3.3×10^{-21}	rs189962749	-
	RAPH1	65059	Raph1	77300	6.8×10^{-6}	6.1×10^{-3}	3.8×10^{-8}	rs55881728	1.1×10^{-13}	9.9×10^{-13}	1.3×10^{-13}	rs71427704	-
	GLIS2	84662	Glis2	83396	9.6×10^{-6}	8.7×10^{-3}	1.6×10^{-7}	rs72766545	6.4×10^{-15}	5.8×10^{-12}	6.4×10^{-15}	rs148092711	2y and 2.5y
	SMCO3	440087	Smco3	654818	1.5×10^{-5}	0.013	4.0×10^{-7}	rs1861698	3.2×10^{-17}	2.9×10^{-14}	2.9×10^{-18}	rs2445411, rs10846065	2y
	C11orf24	53838	1810055G02Rik	72056	4.4×10^{-5}	0.039	1.6×10^{-7}	rs74704614	3.6×10^{-39}	3.3×10^{-36}	6.3×10^{-66}	rs61887796	-
Cluster 4	WNT16[f*]	51384	Wnt16	93735	1.7×10^{-25}	1.1×10^{-22}	7.2×10^{-50}	rs2908007	1.0×10^{-50}	6.8×10^{-48}	$<5 \times 10^{-323}$	56 SNPs	2y and 2.5y
	CPED1[f]	79974	Cped1	214642	5.9×10^{-16}	4.0×10^{-13}	7.2×10^{-50}	rs2908007	1.0×10^{-50}	6.8×10^{-48}	$<5 \times 10^{-323}$	111 SNPs	2y and 2.5y
	MEOX1	4222	Meox1	17285	8.0×10^{-12}	5.4×10^{-9}	$4.6. \times 10^{-14}$	rs6503470	1.0×10^{-50}	6.8×10^{-48}	7.2×10^{-99}	rs9902329	2y
	FGFRL1	53834	Fgfrl1	116701	1.6×10^{-10}	1.1×10^{-7}	7.8×10^{-13}	rs78520297	1.8×10^{-21}	1.2×10^{-18}	5×10^{-193}	rs4505759	2y
	LDLRAD4	753	Ldlrad4	52662	9.1×10^{-8}	6.2×10^{-5}	1.5×10^{-10}	rs4430817	1.7×10^{-18}	1.2×10^{-15}	2×10^{-155}	rs2040189	2y
	SPON2	10417	Spon2	100689	1.7×10^{-5}	0.011	1.7×10^{-7}	rs4974591	1.0×10^{-50}	6.8×10^{-48}	8.3×10^{-80}	rs113727613	2.5y
	MARK3	4140	Mark3	17169	4.1×10^{-5}	0.028	1.9×10^{-7}	rs1013650	5.3×10^{-25}	3.6×10^{-22}	2×10^{-110}	rs11622356	-
	SOX6	55553	Sox6	20679	4.8×10^{-5}	0.033	9.2×10^{-8}	rs12796718 rs12787015	3.5×10^{-34}	2.4×10^{-31}	6.2×10^{-37}	rs4595506	-
	SAMD5	389432	Samd5	320825	5.3×10^{-5}	0.036	8.7×10^{-8}	rs553306937	0.022	0.04	3.5×10^{-5}	rs540384	-
	WNT5B	81029	Wnt5b	22419	6.9×10^{-5}	0.047	1.6×10^{-7}	rs9668959 rs12787015	1.0×10^{-50}	6.8×10^{-48}	1.5×10^{-68}	rs7959604	-

[a]MAGMA gene-based score based on empirical min *p* value (Supplemental Table S3).
[b]Multiple testing correction for each mouse DEG using Bonferroni correction method. Significance level: adjusted *p* value < 0.05 (Supplemental Table S3).
[c]Published UK Biobank GWAS result of most significant SNP *p* value across a gene region. This *p* value is different from MAGMA empirical *p* value.
[d]rsID numbers for the most significant SNP identified in UK Biobank GWAS; in cases where multiple significant SNP exist, the number of SNPs are listed.

genes, *Glis2, Nbeal1, Smco3*, and *Spon2*, are significantly associated with fracture but do not pass the stringent GWAS significance level (Table 2). Although their function in bone has not been reported, *Spon2* is one of the most significantly downregulated genes in aging, which makes *Spon2* a possible candidate for further studies (Fig. 4*A*). A previous yeast study indicates that *Glis2* is a negative regulator of the Wnt/β-catenin signaling pathway.[73] Other human and mouse studies in non-skeletal tissues demonstrate that *Glis2* enhances BMP signaling and induces Tgfβ1 gene expression.[74,75] Furthermore, among the remaining significant fracture-associated genes that were identified in previous GWAS analyses, *Ldlrad4* and *Atxn7l3* do not have detailed experimental studies in bone. The International Mouse Phenotyping Consortium (IMPC) has some evidence that these genes have associated skeletal phenotypes.[76-78] Additionally, in human liver cancer cell line studies, both genes were implicated in the control of TGFβ signaling,[79,80] a well-known regulator of bone mass and quality.[20,81-84] Moreover, a recent study listed *Glis2* and *Ldlrad4* as osteocyte transcriptome signature genes.[61] Further deletion of Ldlrad4 in mice had abnormal structural and functional skeletal phenotypes.[61] Therefore, especially the six genes mentioned above (*Glis2, Nbeal1, Smco3*, *Spon2*, *Ldlrad4*, and *Atxn7l3*), each of which we find is differentially expressed in aging mouse bone, are significantly associated with human fracture and represent compelling candidates for further experimental study. Moreover, focusing on genes associated with fracture, in addition to eBMD, has the potential to efficiently advance discovery of new mechanisms required for the control of bone quality.

This study has limitations as well. Although GWAS takes into account ethnic differences among populations while identifying genetic variants associated with a trait, it cannot eliminate the influence of environment and lifestyle on the trait, which complicates identification of true causal genes. Additionally, UK Biobank GWAS of fracture[10] was adjusted for age; this could have reduced the number of mouse genes that were significantly associated with human traits. Moreover, we could not unveil sex-specific differences, fracture site, or the origin of the fractures, and participants range from 40 to 69 years old. Furthermore, despite studies confirming the concordance of heel ultrasound and DXA results,[11] UK Biobank uses eBMD values estimated from quantitative heel ultrasounds. We used only the UK Biobank GWAS results for eBMD and fracture because it has half a million participants, but this analysis could be repeated in other populations. Although the cortical bone preparation used here enriches for osteocytes, age-related changes in mouse gene expression could also reflect contributions of vascular, neuronal, or other cell populations. Nonetheless, computational integration of widely available mouse RNA-seq data and large-scale human GWAS studies has the potential to accelerate discovery of clinically relevant mechanisms of disease, in and beyond the skeleton.

Disclosures

All authors state that they have no conflicts of interest.

Acknowledgments

We thank S Melov and Buck Institute for providing aged animals; T Lang for helpful scientific discussions; and M Sirota and I Paranjpe for their assistance with RNA-seq analysis. This study benefited from resources and expertise provided by the NIH/NIAMS P30AR070155-supported UCSF PREMIER (Precision Medicine in Rheumatology), and the NIH/NIAMS P30AR075055-supported UCSF Core Center for Musculoskeletal Biology and Medicine, which also provided pilot funding through the Tools and Technology grant program. Additional funding was provided by NIH R01DE019284, NIH U24AG051129 (DSE), the Read Research Foundation, and the Sandler Program for Breakthrough Biomedical Research (TA).

Author Contributions

Serra Kaya: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; software; supervision; validation; visualization; writing – original draft; writing – review and editing. **Charles A Schurman:** Investigation; writing – review and editing. **Neha S Dole:** Conceptualization; writing – review and editing. **Daniel S Evans:** Conceptualization; data curation; formal analysis; investigation; software; validation; visualization; writing – review and editing. **Tamara Alliston:** Conceptualization; funding acquisition; methodology; project administration; resources; supervision; writing – review and editing.

Data Availability Statement

The data that supports the findings of this study are available in the supplementary material of this article.

References

1. Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. Osteoporos Int. 2006; 17:1726-1733.
2. Abimanyi-Ochom J, Watts JJ, Borgström F, et al. Changes in quality of life associated with fragility fractures: Australian arm of the International Cost and Utility Related to Osteoporotic Fractures Study (AusICUROS). Osteoporos Int. 2015;26:1781-1790.
3. Lips P, van Schoor NM. Quality of life in patients with osteoporosis. Osteoporos Int. 2005;16:447-455.
4. Gold DT, Williams SA, Weiss RJ, et al. Impact of fractures on quality of life in patients with osteoporosis: a US cross-sectional survey. J Drug Assess. 2019;8:175-183.
5. Cauley JA, Thompson DE, Ensrud KC, Scott JC, Black D. Risk of mortality following clinical fractures. Osteoporos Int. 2000;11:556-561.
6. Brown JP, Adachi JD, Schemitsch E, et al. Mortality in older adults following a fragility fracture: real-world retrospective matched-cohort study in Ontario. BMC Musculoskelet Disord. 2021;22:105.
7. Guzon-Illescas O, Fernandez EP, Villarias NC, et al. Mortality after osteoporotic hip fracture: incidence, trends, and associated factors. J Orthop Surg Res. 2019;14:203.
8. Wainwright SA, Marshall LM, Ensrud KE, et al. Hip fracture in women without osteoporosis. J Clin Endocrinol Metab. 2005;90:2787-2793.
9. Krall EA, Dawson-Hughes B. Heritable and life-style determinants of bone mineral density. J Bone Miner Res. 2009;8:1-9. https://doi.org/10.1002/jbmr.5650080102
10. Morris JA, Kemp JP, Youlten SE, et al. An atlas of genetic influences on osteoporosis in humans and mice. Nat Genet. 2019;51:258-266.
11. Kemp JP, Morris JA, Medina-Gomez C, et al. Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. Nat Genet. 2017;49:1468-1475.
12. Estrada K, Styrkarsdottir 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;44:491-501.
13. Medina-Gomez C, Kemp JP, Trajanoska K, et al. Life-course genome-wide association study meta-analysis of total body BMD and assessment of age-specific effects. Am J Hum Genet. 2018;102:88-102.

14. Rivadeneira F, Styrkársdottir U, Estrada K, et al. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. Nat Genet. 2009;41:1199-1206.
15. Richards J, Rivadeneira F, Inouye M, et al. Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. Lancet. 2008;371:1505-1512.
16. Guan Y, Ackert-Bicknell CL, Kell B, Troyanskaya OG, Hibbs MA. Functional genomics complements quantitative genetics in identifying disease-gene associations. PLoS Comput Biol. 2010;6:1000991.
17. Razi H, Birkhold AI, Weinkamer R, Duda GN, Willie BM, Checa S. Aging leads to a dysregulation in mechanically driven bone formation and resorption. J Bone Miner Res. 2015;30:1864-1873.
18. Qing H, Ardeshirpour L, Pajevic PD, et al. Demonstration of osteocytic perilacunar/canalicular remodeling in mice during lactation. J Bone Miner Res. 2102;27:1018-1029. https://doi.org/10.1002/jbmr.1567
19. Tang SY, Herber RP, Ho SP, Alliston T. Matrix metalloproteinase-13 is required for osteocytic perilacunar remodeling and maintains bone fracture resistance. J Bone Miner Res. 2012;27:1936-1950.
20. Dole NS, Mazur CM, Acevedo C, et al. Osteocyte-intrinsic TGF-β signaling regulates bone quality through perilacunar/canalicular remodeling. Cell Rep. 2017;21:2585-2596.
21. Kaya S, Basta-Pljakic J, Seref-Ferlengez Z, et al. Lactation induced changes in the volume of osteocyte lacunar-canalicular space alter mechanical properties in cortical bone tissue. J Bone Miner Res. 2017;32:688-697.
22. Milovanovic P, Zimmermann EA, Riedel C, et al. Multi-level characterization of human femoral cortices and their underlying osteocyte network reveal trends in quality of young, aged, osteoporotic and antiresorptive-treated bone. Biomaterials. 2015;45:46-55.
23. Suen PK, Qin L. Sclerostin, an emerging therapeutic target for treating osteoporosis and osteoporotic fracture: a general review. J Orthop Transl. 2016;4:1-13.
24. Hemmatian H, Bakker AD, Klein-Nulend J, van Lenthe GH. Aging, osteocytes, and mechanotransduction. Curr Osteoporos Rep. 2017;15:401-411.
25. Tiede-Lewis LM, Xie Y, Hulbert MA, et al. Degeneration of the osteocyte network in the C57BL/6 mouse model of aging. Aging (Albany NY). 2017;9:2190-2208.
26. Schurman C, Verbruggen S, Alliston T. Degenerated lacunocanalicular networks, mass transport and osteocyte pericellular fluid flow in bone with aging and disrupted TGF-β signaling. PNAS. 2021;118:e2023999118.
27. Farr JN, Kaur J, Doolittle ML, Khosla S. Osteocyte cellular senescence. Curr Osteoporos Rep. 2020;18:559-567.
28. Jilka RL, O'Brien CA, Roberson PK, Bonewald LF, Weinstein RS, Manolagas SC. Dysapoptosis of osteoblasts and osteocytes increases cancellous bone formation but exaggerates cortical porosity with age. J Bone Miner Res. 2014;29:103-117.
29. Jilka RL, O'Brien CA. The role of osteocytes in age-related bone loss. Curr Osteoporos Rep. 2016;14:16-25.
30. Farr JN, Fraser DG, Wang H, et al. Identification of senescent cells in the bone microenvironment. J Bone Miner Res. 2016;31:1920-1929.
31. Kim H-N, Xiong J, MacLeod RS, et al. Osteocyte RANKL is required for cortical bone loss with age and is induced by senescence. JCI Insight. 2020;5:e138815.
32. Jilka RL. The relevance of mouse models for investigating age-related bone loss in humans. J Gerontol Ser. A Biol Sci Med Sci. 2013;68:1209-1217.
33. Flurkey K, Mcurrer J, Harrison D. Mouse models in aging research. In Fox JG, Davisson MT, Quimby FW, Barthold SW, Newcomer CE, Smith AL, eds. The mouse in biomedical research, vol. 3. 2nd ed. Cambridge, MA: Elsevier; 2007 pp 637-672. https://doi.org/10.1016/B978-012369454-6/50074-1
34. Halleux C, Kramer I, Allard C, Kneissel M. Isolation of mouse osteocytes using cell fractionation for gene expression analysis. Methods Mol Biol. 2012;816:55-66.
35. Kelly NH, Schimenti JC, Patrick Ross F, van der Meulen MCH. A method for isolating high quality RNA from mouse cortical and cancellous bone. Bone. 2014;68:1-5.
36. Mazur CM, Woo JJ, Yee CS, et al. Osteocyte dysfunction promotes osteoarthritis through MMP13-dependent suppression of subchondral bone homeostasis. Bone Res. 2019;7:34.
37. Fowler TW, Acevedo C, Mazur CM, et al. Glucocorticoid suppression of osteocyte perilacunar remodeling is associated with subchondral bone degeneration in osteonecrosis. Sci Rep. 2017;7:44618.
38. Andrews S. Babraham Bioinformatics – FastQC a quality control tool for high throughput sequence data. 2018. Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
39. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15-21.
40. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2 moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol. 2014;15:521-550.
41. Pollard KS, Dudoit S, van der Laan MJ. Multiple testing procedures: the multtest package and applications to genomics. In: Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S. eds. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Statistics for Biology and Health*. New York: Springer; 2005: 249-271.
42. Pantano L. DEGreport: report of DEG analysis. R package version 1.30.0. 2021. Available at: https://lpantano.github.io/DEGreport/.
43. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. Omi A J Integr Biol. 2012;16:284-287.
44. Storey JD. A direct approach to false discovery rates. J R Stat Soc Ser B Stat Methodol. 2002;64:479-498.
45. Luo W, Pant G, Bhavnasi YK, Blanchard SG, Brouwer C. Pathview web: user friendly pathway visualization and data integration. Nucleic Acids Res. 2017;45:W501-W508.
46. Luo W, Brouwer C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics. 2013;29:1830-1831.
47. UK Biobank eBMD and Fracture GWAS Data Release 2018 – GEFOS. 2018. Available at: http://www.gefos.org/?q=content/data-release-2018.
48. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol. 2015;11:1-19.
49. Human and Mouse Homology Classes with Sequence information (tab-delimited). 2020. Available at: http://www.informatics.jax.org/downloads/reports/index.html#homology.
50. Zimmermann EA, Schaible E, Bale H, et al. Age-related changes in the plasticity and toughness of human cortical bone at multiple length scales. Proc Natl Acad Sci U S A. 2011;108:14416-14421.
51. Movérare-Skrtic S, Henning P, Liu X, et al. Osteoblast-derived WNT16 represses osteoclastogenesis and prevents cortical bone fragility fractures. Nat Med. 2014;20:1279-1288.
52. Zhang J, Ma Z, Yan K, Wang Y, Yang Y, Wu X. Matrix Gla protein promotes the bone formation by up-regulating Wnt/β-catenin signaling pathway. Front Endocrinol (Lausanne). 2019;10:891.
53. Gori F, Lerner U, Ohlsson C, Baron R. A new WNT on the bone: WNT16, cortical bone thickness, porosity and fractures. Bonekey Rep. 2015;4:669.
54. Cui Y, Niziolek PJ, MacDonald BT, et al. Lrp5 functions in bone to regulate bone mass. Nat Med. 2011;17:684-691.
55. Williams BO. LRP5: from bedside to bench to bone. Bone. 2017;102:26-30.
56. Galea GL, Meakin LB, Harris MA. Old age and the associated impairment of bones' adaptation to loading are associated with transcriptomic changes in cellular metabolism, cell-matrix interactions and the cell cycle. Gene. 2017;599:36-52.
57. Wei Y, Fu J, Wu W, Wu J. Comparative profiles of DNA methylation and differential gene expression in osteocytic areas from aged and young mice. Cell Biochem Funct. 2020;38:721-732.
58. Ioannidis JPA, Tarone R, McLaughlin JK. The false-positive to false-negative ratio in epidemiologic studies. Epidemiology. 2011;22:450-456.

59. Kiel DP, Kemp JP, Rivadeneira F, et al. The musculoskeletal knowledge portal: making omics data useful to the broader scientific community. J Bone Miner Res. 2020;35:1626-1633.
60. McDonald MM, Khoo WH, Ng PY, et al. Osteoclasts recycle via osteomorphs during RANKL-stimulated bone resorption. Cell. 2021;184:1330-1347.e13.
61. Youlten SE, Kemp JP, Logan JG, et al. Osteocyte transcriptome mapping identifies a molecular landscape controlling skeletal homeostasis and susceptibility to skeletal disease. Nat Commun. 2021;12:2444.
62. Akkus O, Adar F, Schaffler MB. Age-related changes in physicochemical properties of mineral crystals are related to impaired mechanical function of cortical bone. Bone. 2004;34:443-453.
63. Willinghamm MD, Brodt MD, Lee KL, Stephens AL, Ye J, Silva MJ. Age-related changes in bone structure and strength in female and male BALB/c mice. Calcif Tissue Int. 2010;86:470-483.
64. Creecy A, Uppuganti S, Girard MR, et al. The age-related decrease in material properties of BALB/c mouse long bones involves alterations to the extracellular matrix. Bone. 2020;130:115126.
65. Robling AG, Niziolek PJ, Baldridge LA, et al. Mechanical stimulation of bone in vivo reduces osteocyte expression of Sost/sclerostin. J Biol Chem. 2008;283:5866-5875.
66. Chang MK, Kramer I, Keller H. Reversing LRP5-dependent osteoporosis and SOST deficiency-induced sclerosing bone disorders by altering WNT signaling activity. J Bone Miner Res. 2014;29:29-42.
67. Sun Y, Wang X, Chen R, et al. Osteoking downregulates Mgp in an osteoporotic fracture rat model. J Tradit Chin Med. 2020;40:422-431.
68. Haraguchi R, Kitazawa R, Mori K, et al. sFRP4-dependent Wnt signal modulation is critical for bone remodeling during postnatal development and age-related bone loss. Sci Rep. 2016;6:25198.
69. Skuntz S, Mankoo B, Nguyen M-TT, et al. Lack of the mesodermal homeodomain protein MEOX1 disrupts sclerotome polarity and leads to a remodeling of the cranio-cervical joints of the axial skeleton. Dev Biol. 2009;332:383-395.
70. Chesi A, Wagley Y, Johnson ME, et al. Genome-scale capture C promoter interactions implicate effector genes at GWAS loci for bone mineral density. Nat Commun. 2019;10:1260.
71. Simsek Kiper PO, Saito H, Gori F, et al. Cortical-bone fragility — insights from sFRP4 deficiency in Pyle's disease. N Engl J Med. 2016;374:2553-2562.
72. Dauer MVP, Currie PD, Berger J. Skeletal malformations of Meox1-deficient zebrafish resemble human Klippel–Feil syndrome. J Anat. 2018;233:687-695.
73. Kim Y-S, Kang HS, Jetten AM. The Krüppel-like zinc finger protein Glis2 functions as a negative modulator of the Wnt/β-catenin signaling pathway. FEBS Lett. 2007;581:858-864. https://doi.org/10.1016/j.febslet.2007.01.058
74. Gruber TA, Gedman AL, Zhang J, et al. An Inv(16)(p13.3q24.3)-encoded CBFA2T3-GLIS2 fusion protein defines an aggressive subtype of pediatric acute megakaryoblastic leukemia. Cancer Cell. 2012;22:683-697.
75. Attanasio M, Uhlenhaut NH, Sousa VH, et al. Loss of GLIS2 causes nephronophthisis in humans and mice by increased apoptosis and fibrosis. Nat Genet. 2007;39:1018-1024.
76. Ldlrad4 mouse gene details. 2019. Available at: https://www.mousephenotype.org/data/genes/MGI:1277150.
77. Epdr1 mouse gene details. 2019. Available at: https://www.mousephenotype.org/data/genes/MGI:2145369.
78. Atxn7l3 mouse gene details. 2019. Available at: https://www.mousephenotype.org/data/genes/MGI:3036270.
79. Nakano N, Maeyama K, Sakata N, et al. C18 orf1, a novel negative regulator of transforming growth factor-β signaling. J Biol Chem. 2014;289:12680-12692.
80. Sun N, Zhong X, Wang S, et al. ATXN7L3 positively regulates SMAD7 transcription in hepatocellular carcinoma with growth inhibitory function. EBioMedicine. 2020;62:103108.
81. Balooch G, Balooch M, Nalla RK, et al. TGF-beta regulates the mechanical properties and composition of bone matrix. Proc Natl Acad Sci U S A. 2005;102:18813-18818.
82. Edwards JR, Nyman JS, Lwin ST, et al. Inhibition of TGF-β signaling by 1D11 antibody treatment increases bone mass and quality in vivo. J Bone Miner Res. 2010;25:2419-2426. https://doi.org/10.1002/jbmr.139
83. Tang SY, Alliston T. Regulation of postnatal bone homeostasis by TGFβ. Bonekey Rep. 2013;2:255.
84. Mohammad KS, Chen CG, Balooch G, et al. Pharmacologic inhibition of the TGF-β type I receptor kinase has anabolic and anti-catabolic effects on bone. PLoS One. 2009;4:e5275.