

Transcriptome-wide association study and eQTL colocalization identify potentially causal genes responsible for human bone mineral density GWAS associations

Basel Maher Al-Barghouthi^{1,2}, Will T Rosenow¹, Kang-Ping Du³, Jinho Heo⁴, Robert Maynard⁵, Larry Mesner^{1,6}, Gina Calabrese¹, Aaron Nakasone⁷, Bhavya Senwar⁸, Louis Gerstenfeld⁹, James Lerner³, Virginia Ferguson⁸, Cheryl Ackert-Bicknell⁵, Elise Morgan⁷, David Brautigan⁴, Charles R Farber^{1,2,6*}

¹Center for Public Health Genomics, School of Medicine, University of Virginia, Charlottesville, United States; ²Department of Biochemistry and Molecular Genetics, School of Medicine, University of Virginia, Charlottesville, United States; ³Department of Radiation Oncology, University of Virginia, Charlottesville, United States; ⁴Department of Microbiology, Immunology, and Cancer Biology, School of Medicine, University of Virginia, Charlottesville, United States; ⁵Department of Orthopedics, Anschutz Medical Campus, University of Colorado, Aurora, United States; ⁶Department of Public Health Sciences, School of Medicine, University of Virginia, Charlottesville, United States; ⁷Department of Mechanical Engineering, Boston University, Boston, United States; ⁸Department of Mechanical Engineering, University of Colorado Boulder, Boulder, United States; ⁹Department of Orthopaedic Surgery, Boston University Medical Center, Boston, United States

*For correspondence: crf2s@virginia.edu

Competing interest: [See page 17](#)

Funding: [See page 17](#)

Preprinted: [13 October 2021](#)

Received: 24 January 2022

Accepted: 16 November 2022

Published: 23 November 2022

Reviewing Editor: Nele Haelterman, Baylor College of Medicine, United States

© Copyright Al-Barghouthi et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract Genome-wide association studies (GWASs) for bone mineral density (BMD) in humans have identified over 1100 associations to date. However, identifying causal genes implicated by such studies has been challenging. Recent advances in the development of transcriptome reference datasets and computational approaches such as transcriptome-wide association studies (TWASs) and expression quantitative trait loci (eQTL) colocalization have proven to be informative in identifying putatively causal genes underlying GWAS associations. Here, we used TWAS/eQTL colocalization in conjunction with transcriptomic data from the Genotype-Tissue Expression (GTEx) project to identify potentially causal genes for the largest BMD GWAS performed to date. Using this approach, we identified 512 genes as significant using both TWAS and eQTL colocalization. This set of genes was enriched for regulators of BMD and members of bone relevant biological processes. To investigate the significance of our findings, we selected *PPP6R3*, the gene with the strongest support from our analysis which was not previously implicated in the regulation of BMD, for further investigation. We observed that *Ppp6r3* deletion in mice decreased BMD. In this work, we provide an updated resource of putatively causal BMD genes and demonstrate that *PPP6R3* is a putatively causal BMD GWAS gene. These data increase our understanding of the genetics of BMD and provide further evidence for the utility of combined TWAS/colocalization approaches in untangling the genetics of complex traits.

Editor's evaluation

Many GWAS studies have been done to understand the genetic contributions to bone density, but very few have managed to pinpoint the gene affected by a polymorphism that caused an observed difference. In this paper, your team shows how scientists can identify causative variants from GWAS studies.

Introduction

Osteoporosis, a disease characterized by low bone mineral density (BMD), decreased bone strength, and an increased risk of fracture, affects over 10 million individuals in the United States (**Black and Rosen, 2016; Burge et al., 2007**). BMD is the single strongest predictor of fracture and a highly heritable quantitative trait (**Miller et al., 1999; Ralston and Uitterlinden, 2010; Peacock et al., 2002**). Over the last decade, genome-wide association studies (GWASs) have identified over 1100 independent associations for BMD (**Morris et al., 2019; Estrada et al., 2012; Kemp et al., 2017**). However, despite the success of GWAS, few of the underlying causal genes have been identified (**Rocha-Braz and Ferraz-de-Souza, 2016; Sabik and Farber, 2017**).

One of the main difficulties in GWAS gene discovery is that the vast majority (>90%) of associations are driven by non-coding variation (**Giral et al., 2018; Edwards et al., 2013**). Over the last decade, approaches such as transcriptome-wide association studies (TWASs) and expression quantitative trait locus (eQTL) colocalization have been developed which leverage transcriptomic data in order to inform gene discovery by connecting non-coding disease-associated variants to changes in transcript levels (**Gusev et al., 2016; Barbeira et al., 2018; Pividori et al., 2020; Giambartolomei et al., 2014; Wen et al., 2017**). These approaches have proven successful for a wide array of diseases and disease-associated quantitative traits (**Pividori et al., 2020; Bhattacharya et al., 2020; Thom and Voight, 2020**). However, the osteoporosis field has lagged behind such efforts, due to the limited number of large-scale bone-related transcriptomic datasets.

In a TWAS, genetic predictors of gene expression (e.g., local eQTL – sets of genetic variants that influence the expression of a gene in close proximity, **Nica and Dermitzakis, 2013**) identified in a reference population (e.g., the Genotype-Tissue Expression [GTEx] project, **Consortium, 2013**) are used to impute gene expression in a GWAS cohort. Components of gene expression due to genetic variation are then associated with a disease or disease-associated quantitative trait. Genes identified by TWAS are often located in GWAS associations, suggesting that the genetic regulation of their expression is the mechanism underlying such associations. Several tools, for example, FUSION, PrediXcan, and MultiXcan (**Gusev et al., 2016; Gamazon et al., 2015; Barbeira et al., 2019**) have been developed to perform TWASs. Most of these tools use GWAS summary statistics, making TWAS widely applicable to large GWAS datasets. In contrast, eQTL colocalization is a statistical approach that determines if there is a shared genetic basis for two associations (e.g., a local eQTL and BMD GWAS locus). Recently, it has been demonstrated that prioritizing genes using both TWAS and eQTL colocalization provides a way to identify genes with the strongest support for causality (**Barbeira et al., 2018; Pividori et al., 2020**).

The GTEx project has generated RNA-seq data on over 50 tissues across hundreds of individuals (**Consortium, 2020**). Even though data on the tissues/cell types likely to be most relevant to BMD (bone or bone cells) were not included, the project demonstrated that many eQTL were shared across tissues (**Consortium, 2020; Battle et al., 2017**). Additionally, it is well known that effects in a wide-array of non-bone cell types and tissues can impact bone and BMD (**Fitzpatrick, 2002; Mirza and Canalis, 2015**). As a result, we sought to use the GTEx resource in conjunction with TWAS and eQTL colocalization to leverage non-bone gene expression data to identify putatively causal genes underlying BMD GWAS.

Here, we performed TWAS and eQTL colocalization using the GTEx resource and the largest BMD GWAS performed to date to identify potentially causal genes (**Morris et al., 2019**). Using this approach, we identified 512 genes significantly associated via TWAS with a significant colocalizing eQTL. To investigate the significance of our findings we selected Protein Phosphatase 6 Regulatory Subunit 3 (*PPP6R3*), the gene with the strongest support not previously implicated in the regulation of BMD, for further investigation. We demonstrate using mutant mice that *Ppp6r3* is a regulator of lumbar spine BMD. These results highlight the power of leveraging GTEx data, even in the absence

of data from the most relevant tissue/cell types, to increase our understanding of the genetic architecture of BMD.

Results

TWAS and eQTL colocalization identify potentially causal BMD GWAS genes

To identify potentially causal genes responsible for BMD GWAS associations, we combined TWAS and eQTL colocalization using GTEx data (**Figure 1A**). We began by performing a TWAS using reference gene expression predictions from GTEx (Version 8; 49 tissues) and the largest GWAS performed to date for heel estimated BMD (eBMD) (>1100 independent associations) (**Morris et al., 2019; Consortium, 2020**). The analysis was performed using S-MultiXcan, which allowed us to leverage information across all 49 GTEx tissues (**Barbeira et al., 2019**). Our analysis focused on protein-coding genes (excluded non-coding genes). A total of 2156 protein-coding genes were significantly (Bonferroni-adjusted p value ≤ 0.05) associated with eBMD (**Supplementary file 1a**).

Next, we identified colocalizing eQTL from each of the 49 tissues in GTEx using fastENLOC (**Pivori et al., 2020; Wen et al., 2017**). We identified 1182 colocalizing protein-coding genes with a regional colocalization probability (RCP) of 0.1 or greater (**Supplementary file 1b**). In total, 512 protein-coding genes were significant in both the TWAS and eQTL colocalization analyses (**Table 1** and **Supplementary file 1c**). Among the identified genes were many with well-known roles in the regulation of BMD, such as *RUNX2* (**Figure 1B**), *IGF1*, and *LRP6*, as well as novel genes such as *RERE* (**Figure 1C**). Overall, the identified genes had significantly colocalizing eQTL in all 49 GTEx tissues, with eQTL from cultured fibroblasts (132 genes), subcutaneous adipose tissue (117 genes), tibial artery (115 genes), and tibial nerve (114 genes) exhibiting the highest number of significant colocalizations (**Supplementary file 1d**). TWAS predictors were only generated for genes on autosomes and of the 1103 independent associations identified by **Morris et al., 2019**, 1097 were autosomal. For each of these, we defined a locus as the region consisting of ± 1 Mbp around each association. Of the 1097 loci, almost half (542; 49%) of the loci contained at least one of the 512 prioritized genes. Most loci overlapped one gene (mean = 1.7, median = 1); however, 184 loci overlapped multiple genes, including a locus on Chromosome (Chr.) 20 (lead SNP rs6142137) which contained 9 prioritized genes (**Figure 1D** and **Figure 1—figure supplement 1**).

Characterization of genes identified by TWAS/eQTL colocalization

To evaluate the ability of the combined TWAS/colocalization approach to identify genes previously implicated in the regulation of BMD, other bone traits, or the activity of bone cells, we queried the presence of 'known bone genes' within the list of the 512 prioritized protein-coding genes. To do so, we created a database-curated set of genes previously implicated in the regulation of bone processes (henceforth referred to as our 'known bone genes' list, $N = 1399$, **Supplementary file 1e**). Of the 512 genes identified above, 66 (12.9%) were known bone genes, representing a significant enrichment (odds ratio [OR] = 1.72; $p = 1.0 \times 10^{-4}$) over what would be expected by chance (**Supplementary file 1f**).

We also performed a Gene Ontology (GO) enrichment analysis of the 512 prioritized genes. We observed enrichments in several bone-relevant ontologies, such as 'ossification' ($p = 3.1 \times 10^{-6}$), 'skeletal system development' ($p = 2 \times 10^{-5}$), and 'regulation of osteoblast differentiation' ($p = 3.9 \times 10^{-5}$) (**Figure 2A** and **Supplementary file 1g**).

To compare our approach with the approach of prioritizing the closest genes to GWAS associations as potentially causal, we quantified the number of genes that were both the closest genes to eBMD GWAS associations and were members of the 'known bone gene' list. Of the 863 genes that were the closest genes to eBMD GWAS associations (**Supplementary file 1h**), 139 were members of the 'known bone gene list', representing a more significant enrichment of 'known bone genes' than our prioritization approach (OR = 2.56, $p = 6.37 \times 10^{-19}$). Of our 512 prioritized genes, 206 (40%) were also the closest genes to eBMD GWAS associations, with 27 of the remaining 306 prioritized genes (8.8%) being members of the 'known bone gene' list.

The International Mouse Phenotype Consortium (IMPC) has recently measured whole body BMD in hundreds of mouse knockouts (**Dickinson et al., 2016; Swan et al., 2020**). We searched the IMPC

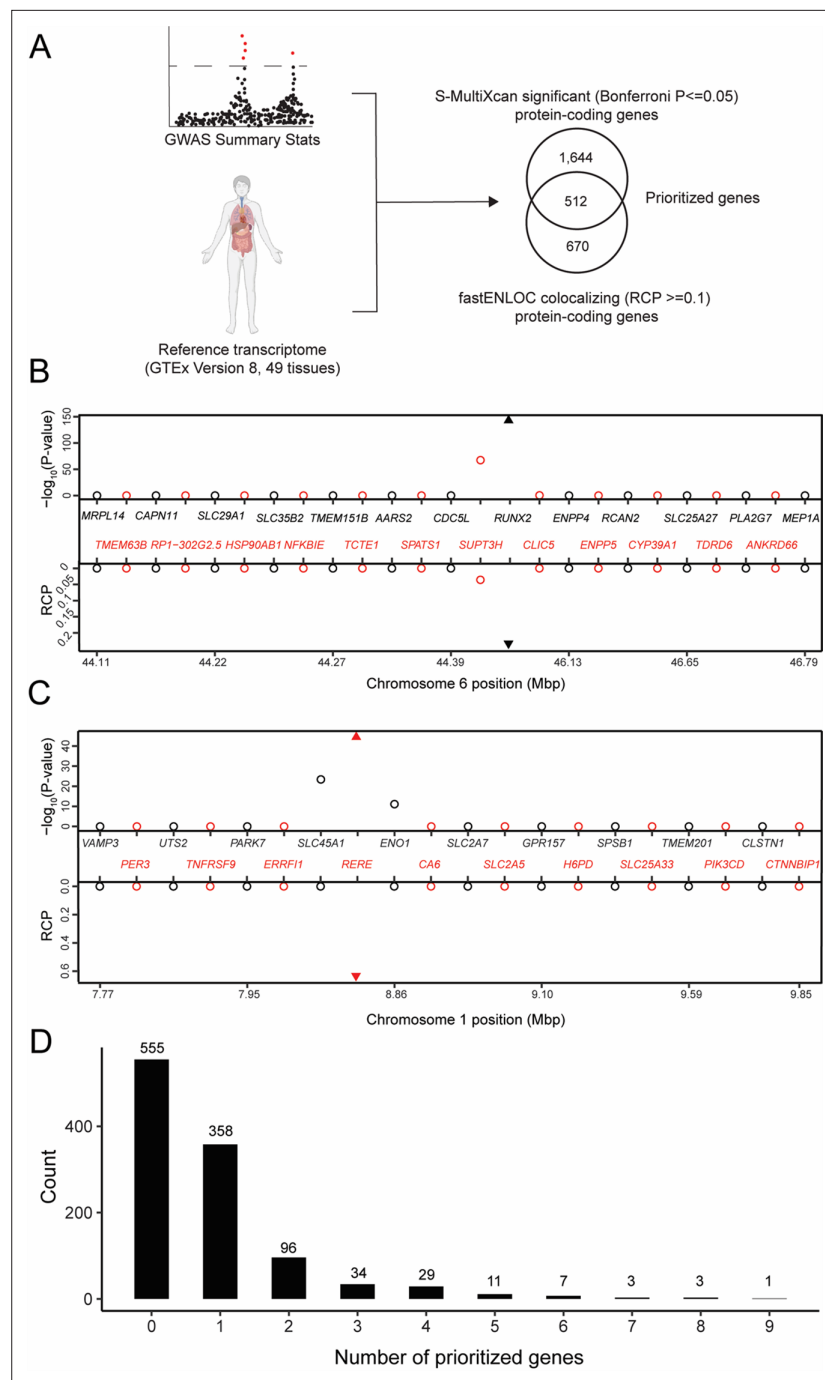


Figure 1. Transcriptome-wide association study (TWAS) and expression quantitative trait loci (eQTL) colocalization identify potentially causal bone mineral density (BMD) genome-wide association study (GWAS) genes. **(A)** Overview of the analysis. The human image was obtained from [BioRender.com](https://www.biorender.com). TWAS/colocalization plot for genes in the locus around *RUNX2* **(B)** and *RERE* **(C)**. The $-\log_{10}$ Bonferroni-adjusted p values from the TWAS analysis (top panel) and the maximum regional colocalization probabilities (RCPs) from the colocalization analyses (bottom panel). Genes alternate in color for visual clarity. Triangles represent *RUNX2* **(B)** and *RERE* **(C)**. **(D)** Distribution of prioritized genes within estimated BMD (eBMD) GWAS loci.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Chromosome 20 contains a locus with nine prioritized genes.

Table 1. Top 10 protein-coding genes significant by colocalization (RCP ≥ 0.1) and TWAS, sorted by TWAS p value.

Gene	Tissue with greatest RCP	Max. RCP	TWAS p value (Bonferroni)
<i>SPTBN1</i>	Cells_Cultured_Fibroblasts	0.9469	$<5 \times 10^{-324}$
<i>CCDC170</i>	Spleen	0.6582	$<5 \times 10^{-324}$
<i>FAM3C</i>	Artery_Tibial	0.4917	$<5 \times 10^{-324}$
<i>SEPT5</i>	Skin_Sun_Exposed	0.4868	2.26×10^{-286}
<i>FGFRL1</i>	Cells_Cultured_Fibroblasts	0.1611	5.31×10^{-272}
<i>GREM2</i>	Cells_Cultured_Fibroblasts	0.9998	4.31×10^{-257}
<i>GPATCH1</i>	Whole_Blood	0.3564	3.44×10^{-226}
<i>RHPN2</i>	Pituitary	0.2181	8.71×10^{-221}
<i>BMP4</i>	Brain_Cortex	0.5468	5.49×10^{-169}
<i>RUNX2</i>	Esophagus_Gastroesophageal_Junction	0.2372	1.99×10^{-146}

database for any of the 512 genes identified by TWAS and eQTL colocalization. Of the 512, 142 (27.7%) had been tested by the IMPC and 64 (12.5% of the 512 prioritized genes, 45% of the 142 IMPC-tested genes) had a nominally significant ($p \leq 0.05$) alteration of whole-body BMD in knockout/knockdown mice, compared to controls. Of the 64, 49 (76.5%) were not members of the ‘known bone gene’ list (**Figure 2—figure supplement 1A, B**).

An example of one of the 64 genes is *GPATCH1*, located within a GWAS association on human chromosome 19q13.11. *GPATCH1* is part of the catalytic step 2 spliceosome, and may be involved in mRNA splicing, and is predicted to enable RNA-binding activity. It is also expressed in bone cells in mouse (**Alliance of genome resources, 2022; Lattin et al., 2008**). Of all the genes in the region, *GPATCH1* had the strongest TWAS association ($p = 3.44 \times 10^{-226}$) (**Figure 2B**) and the strongest eQTL colocalization (whole blood, RCP = 0.36) (**Figure 2B–D**). The eQTL and BMD GWAS allele effects for the top SNPs were in the same direction, suggesting that decreasing the expression of *GPATCH1* would lead to decreased BMD. BMD data from the IMPC showed that female mice heterozygous for a *Gpatch1* null allele had decreased BMD ($p = 2.17 \times 10^{-8}$) (**Figure 2E**). Together, these data suggest that many of the genes identified by the combined TWAS/colocalization approach are likely causal BMD GWAS genes.

PPP6R3 is a candidate causal gene for a GWAS association on Chr. 11

To identify novel candidate genes for functional validation, we focused on genes with the strongest evidence of being causal. To do so, we increased the colocalization RCP threshold to 0.5, and then sorted genes based on TWAS Bonferroni-adjusted p values. Furthermore, we constrained the list of candidates for functional validation to genes which were not members of the ‘known bone gene’ list or genes with a nominal ($p \leq 0.05$) alteration in whole-body BMD as determined by the IMPC. This yielded 137 putatively causal BMD genes (**Table 2, Supplementary file 1i, and Figure 2—figure supplement 1C**).

Though it was not on the ‘known bone gene’ list, the first gene ranked by TWAS p value, *SPTBN1*, has been demonstrated to play a role in the regulation of BMD (**Calabrese et al., 2017**). The second, *PPP6R3*, has not been previously implicated in the regulation of BMD. *PPP6R3*, a regulatory subunit of protein phosphatase 6, which shows ubiquitous expression across tissues in humans (**Cristiano, 2020**), is located on human Chr. 11 within 1 Mbp of seven independent eBMD GWAS SNPs identified by **Morris et al., 2019** (subsequently referred to as ‘eBMD lead SNPs’) (**Figure 3A**). Of all the protein-coding genes ($N = 29$) in the ~1.8 Mbp region surrounding *PPP6R3*, its expression was the most significantly associated with eBMD by TWAS (Bonferroni = 5.7×10^{-93}) (**Figure 3B**). Furthermore, *PPP6R3* was the only gene in the region with eQTL (in four GTEx tissues, thyroid, ovary, brain_putamen_basal_ganglia, and stomach with RCPs = 0.53, 0.50, 0.28, and 0.14, respectively) that colocalized with at least one of the eBMD associations (**Figure 3B**). Based on these data, we chose to further investigate *PPP6R3* as a potentially causal BMD gene.

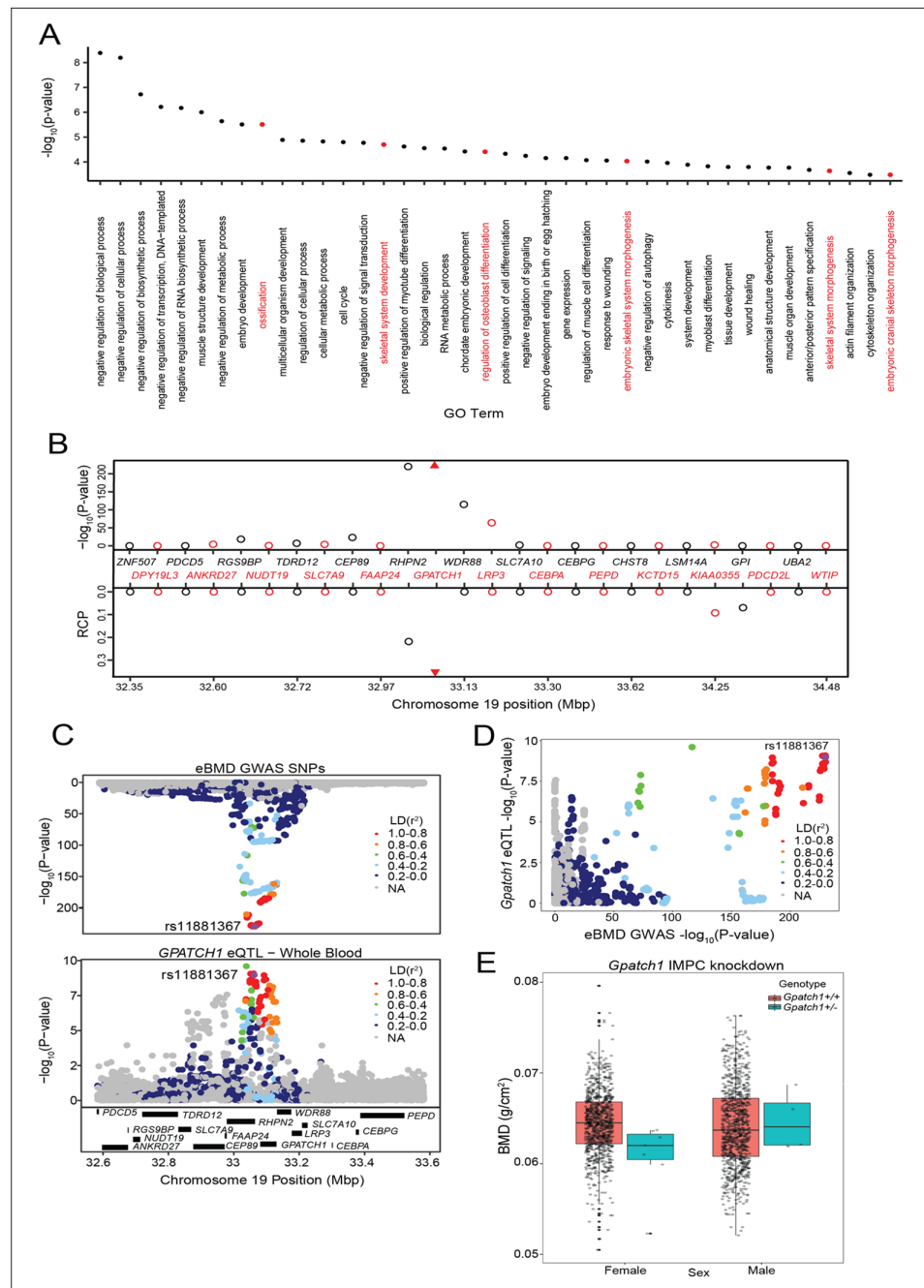


Figure 2. Transcriptome-wide association study (TWAS) and expression quantitative trait loci (eQTL) colocalization identify *Gpatch1* a potentially causal bone mineral density (BMD) genome-wide association study (GWAS) gene. (A) The top 40 terms from a Gene Ontology analysis of the 512 potentially causal BMD genes identified by our analysis. Terms with clear relevance to bone are highlighted in red. Only terms from the ‘Biological Process’ (BP) subontology are listed, and similar terms were removed for clarity. (B) TWAS/colocalization plot for genes in the locus around *GPATCH1* (± 1.5 Mbp). The $-\log_{10}$ Bonferroni-adjusted p values from the TWAS analysis (top panel) and the maximum regional colocalization probabilities (RCPs) from the colocalization analyses (bottom panel). Genes alternate in color for visual clarity. Triangles represent *GPATCH1*. (C) Mirrorplot showing estimated BMD (eBMD) GWAS locus (top panel) and colocalizing *GPATCH1* eQTL in whole blood (bottom panel). SNPs are colored by their linkage disequilibrium (LD) with rs11881367 (purple), the most significant GWAS SNP in the locus. (D) Scatterplot of $-\log_{10}$ p values for *GPATCH1* eQTL versus eBMD GWAS SNPs. SNPs are colored by their LD with rs11881367 (purple). (E) BMD in *Gpatch1* knockdown mice. $N = 7$ females and $N = 4$ males for *Gpatch1*^{-/-} mice, N

Figure 2 continued on next page

Figure 2 continued

= 880 females and $N = 906$ males for *Gpatch1^{+/+}* mice. Boxplots indicate the median (middle line), the 25th and 75th percentiles (box) and the whiskers extend to $1.5 * IQR$.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Schematic of prioritization pipeline.

We first determined which of the seven associations colocalized with the *PPP6R3* eQTL (**Figure 3C**). The most significant *PPP6R3* eQTL SNP in thyroid tissue (the tissue with the highest RCP) was rs10047483 (Chr. 11, 68.464237 Mbp) (*PPP6R3* eQTL $p = 6.99 \times 10^{-8}$, eBMD GWAS $p = 1.2 \times 10^{-100}$) located in intron 1 of *PPP6R3*. The most significant eBMD lead SNP in the locus was rs11228240 (Chr. 11, 68.450822 Mbp, eBMD GWAS $p = 6.6 \times 10^{-101}$, *PPP6R3* eQTL $p = 3.7 \times 10^{-6}$), located upstream of *PPP6R3*. Consistent with the colocalization analysis, these two variants are in high LD ($r^2 = 0.941$) and rs10047483 does not exhibit strong LD ($r^2 < 0.104$) with any of the other six eBMD lead SNPs in the locus. The eQTL and BMD GWAS allele effects for rs10047483 were opposing, suggesting that a decrease in the expression of *PPP6R3* would lead to an increase in BMD.

A recent fracture GWAS identified 14 significant associations, one of which was located in the *PPP6R3* region (rs35989399, Chr. 11, 68.622433 Mbp) (**Morris et al., 2019**). We analyzed the fracture GWAS in the same manner as we did above for eBMD. We found that *PPP6R3* expression when analyzed by TWAS was significant for fracture (TWAS Bonferroni-pval = 6.0×10^{-3}) and the same *PPP6R3* eQTL colocalized with the fracture association (RCP = 0.49 in ovary, RCP = 0.36 in thyroid) (**Figure 3D**). Together, these data highlight *PPP6R3* as a strong candidate for one of the seven eBMD/fracture associations in this region.

***PPP6R3* is a regulator of femoral geometry, BMD, and vertebral microarchitecture**

To assess the effects of *PPP6R3* expression on bone phenotypes, we characterized mice harboring a gene-trap allele (*Ppp6r3^{tm1a(KOMP)Wtsi}*) (**Figure 4A**). We intercrossed mice heterozygous for the mutant allele to generate mice of all three genotypes (wild-type (WT, $N = 26$ females, 21 males), heterozygous (HET, $N = 33$ females, 36 males), and mutant (MUT, $N = 20$ females, 21 males)). The absence of *PPP6R3* protein in MUT mice was confirmed through western blotting (**Figure 4B**). Furthermore, we sequenced RNA extracted from the L5 vertebrae of wild-type and mutant mice, and found that *PPP6R3* was differentially expressed between the mutant and wild-type mice ($p_{\text{adj}} = 3.67 \times 10^{-104}$, \log_2 fold change = -1.63). We did not observe changes in gene expression for any other gene in the locus, including *Lrp5* ($p_{\text{adj}} = 0.99$, \log_2 fold change = -0.068), a well-known regulator of bone mass just upstream of *Ppp6r3* (**Supplementary file 1j**).

Table 2. Top 10 novel protein-coding genes significant by colocalization (RCP ≥ 0.5) and TWAS, sorted by TWAS p value.

Gene	Tissue with greatest RCP	Max. RCP	TWAS p value (Bonferroni)	# GTEx tissues with RCP ≥ 0.5
<i>SPTBN1</i>	Cells_Cultured_fibroblasts	0.9469	$< 5 \times 10^{-324}$	2
<i>PPP6R3</i>	Thyroid	0.5291	5.7×10^{-93}	1
<i>BARX1</i>	Colon_Transverse	0.7764	6.36×10^{-63}	1
<i>MEOX2</i>	Brain_Nucleus_accumbens_basal_ganglia	0.6286	3.21×10^{-53}	3
<i>RERE</i>	Adipose_Subcutaneous	0.6431	6.95×10^{-46}	4
<i>SIPA1</i>	Nerve_Tibial	0.9981	4.26×10^{-41}	1
<i>CAPZB</i>	Testis	0.6716	3.64×10^{-33}	1
<i>B4GALNT3</i>	Artery_Aorta	0.9241	2.67×10^{-33}	4
<i>TRPC4AP</i>	Breast_Mammary_Tissue	0.5577	8.62×10^{-31}	3
<i>AXL</i>	Minor_Salivary_Gland	0.6205	9.74×10^{-31}	3

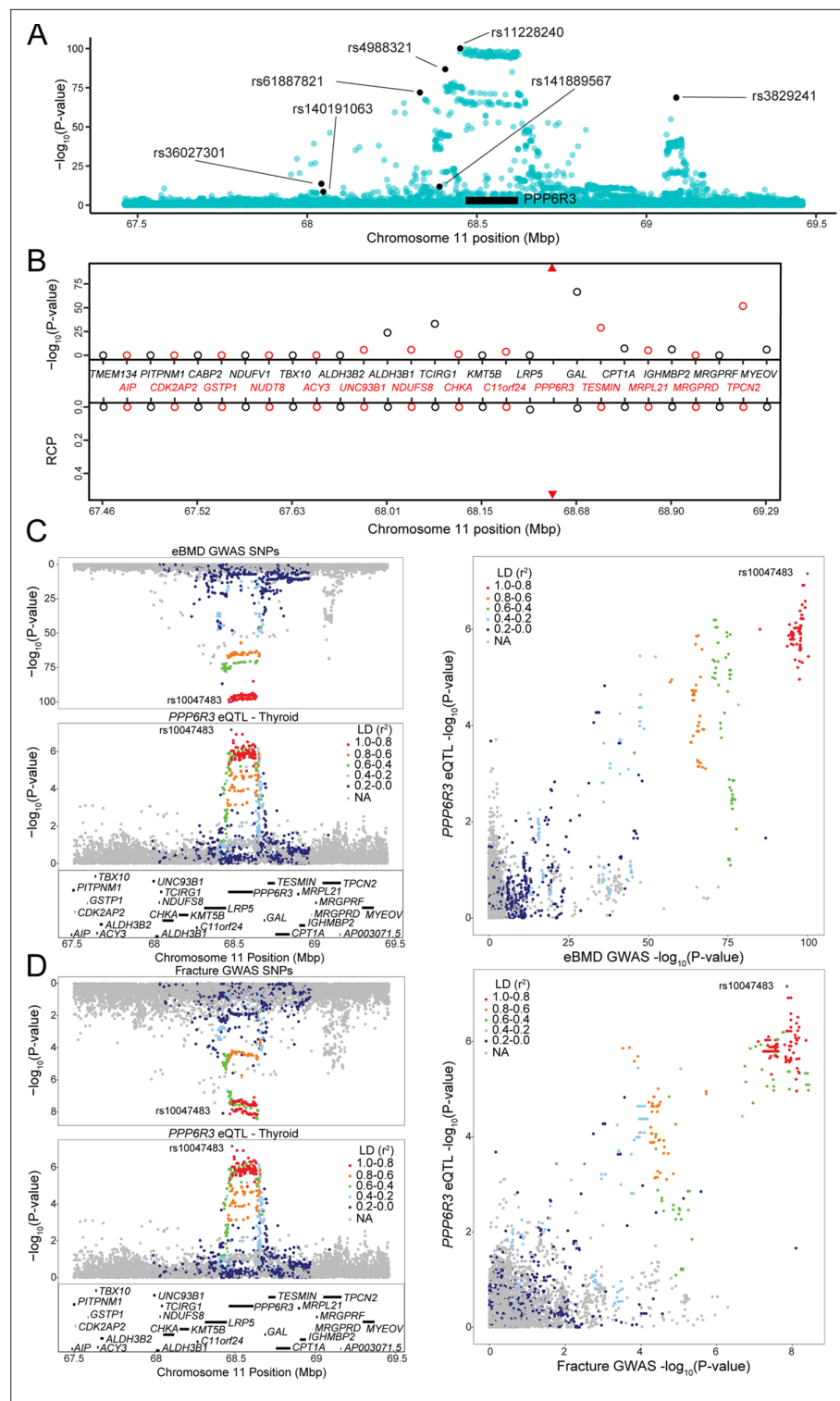


Figure 3. *PPP6R3* is a top 10 novel estimated bone mineral density (eBMD) gene. **(A)** eBMD genome-wide association study (GWAS) SNPs around the *PPP6R3* locus (± 1 Mbp). The y-axis represents $-\log_{10}$ eBMD GWAS p values. Highlighted SNPs (black) are the seven lead eBMD GWAS SNPs in the locus. **(B)** Transcriptome-wide association study (TWAS)/colocalization plot for genes in the locus around *PPP6R3* (± 1 Mbp). The $-\log_{10}$ Bonferroni-adjusted p values from the TWAS analysis (top panel) and the maximum regional colocalization probabilities (RCPs) from the colocalization analyses (bottom panel). Genes alternate in color for visual clarity. Triangles represent *PPP6R3*. Mirrorplot of the eBMD locus **(C)** and *PPP6R3* expression quantitative trait loci (eQTL)

Figure 3 continued on next page

Figure 3 continued

in thyroid, and fracture locus and *PPP6R3* eQTL in thyroid (D). The panels on the right are scatterplots of $-\log_{10}$ p values for *PPP6R3* eQTL and eBMD GWAS SNPs (C) and the *PPP6R3* eQTL and fracture GWAS SNPs (D). SNPs are colored by their linkage disequilibrium (LD) with rs10047483 (purple), the most significant *PPP6R3* eQTL in the locus. Not all genes are shown.

The BMD analyses presented above used heel eBMD GWAS data. We used these data because they represent the largest, most well-powered BMD GWAS to date (Estrada et al., 2012). However, to determine whether perturbation of *Ppp6r3* would be expected to impact femoral or lumbar spine BMD in a similar manner, we turned to a smaller GWAS to look at both of these traits. In a GWAS by Estrada et al., 2012, a total of 56 loci were identified for femoral neck (FNBMD) and lumbar spine (LSBMD) BMD. One of the 56 loci corresponded to the same SNPs associated with the *PPP6R3* eQTL. The locus was significant for LSBMD; however, it did not reach genome-wide significance for FNBMD (Figure 4—figure supplement 1).

We evaluated BMD at both the femur and the lumbar spine in *Ppp6r3^{tm1a(KOMP)Wtsi}* mice, with the expectation, based on the above data, that perturbation of *Ppp6r3* would have a stronger impact on BMD at the lumbar spine. At approximately 9 weeks of age, we measured areal BMD (aBMD) at the femur and lumbar spine using dual X-ray absorptiometry (DXA). First, we observed no change in body weight at 9 weeks that might impact bone phenotypes (Figure 4—figure supplement 2A). As the above analysis predicted, we observed a significant effect of *Ppp6r3* genotype on aBMD at the lumbar spine (WT vs. MUT $p = 0.01$, Figure 4—figure supplement 2C), but not the femur (WT vs. MUT $p = 0.26$, Figure 4—figure supplement 2D). It should also be noted, however, that we observed significantly decreased femoral width, but not length, in *Ppp6r3* mutant mice (anterior–posterior [AP] femoral width, WT vs. MUT $p = 0.02$; medial–lateral [ML] femoral width, WT vs. MUT $p = 2.2 \times 10^{-6}$, Figure 4—figure supplement 2B–D).

We further characterized the effects of *Ppp6r3* genotype on microarchitectural phenotypes, in spine and femur, via micro-computed tomography (μ CT). We observed significant ($p \leq 0.05$) decreases in trabecular bone volume fraction (BV/TV, WT vs. MUT $p = 0.015$, Figure 4—figure supplement 2E, F) and volumetric BMD (vBMD, WT vs. MUT $p = 0.015$, Figure 4—figure supplement 2G) of the lumbar spine as a function of *Ppp6r3* genotype, but found no significant changes in tissue mineral density (TMD, Figure 4—figure supplement 2E), trabecular separation (TbSp), trabecular thickness (TbTh), or trabecular number (TbN) (Figure 4—figure supplement 2F–H). In the femoral midshaft, we observed significant decreases in total area (Tot.Ar, WT vs. MUT $p = 0.00724$, Figure 4—figure supplement 2H) and medullary area (Ma.Ar, WT vs. MUT $p = 0.00903$, Figure 4—figure supplement 2I), which are concordant with the aforementioned observed decreases in femoral width in *Ppp6r3* mutant mice. We did not observe any changes in trabecular bone parameters in the distal femur.

We also measured Procollagen 1 Intact N-Terminal Propeptide P1NP levels in plasma, and observed a significant increase in P1NP levels in mutant mice ($p = 0.00422$), suggesting that the observed decrease in bone mass is due to increased bone turnover (Figure 4—figure supplement 2I–K). Finally, to assess the effects of *Ppp6r3* genotype on bone matrix composition, we performed periosteal Raman spectroscopy on both the lumbar spines and femurs. We did not observe any significant ($p \leq 0.05$) effects of *Ppp6r3* genotype on bone matrix composition (Figure 4—figure supplements 3–6).

Discussion

BMD GWASs have identified over 1100 associations to date. However, identifying causal genes remains a challenge. To aid researchers in further dissecting the genetics of complex traits, reference transcriptomic datasets and computational methods have been developed for the prioritization and identification of causal genes underlying GWAS associations. In this work, our goal was to utilize these data and tools to prioritize putatively causal genes underlying BMD GWAS associations. Specifically, we used the GTEx eQTL reference dataset in 49 tissues to perform TWAS and eQTL colocalization on the largest BMD GWAS. Using this approach, we identified 512 putatively causal protein-coding genes that were significant in both the TWAS and colocalization approaches.

Our approach was inspired by a recent study that used the GTEx resource and a TWAS/eQTL colocalization approach similar to the one we employed. Pividori et al., 2020 recently combined

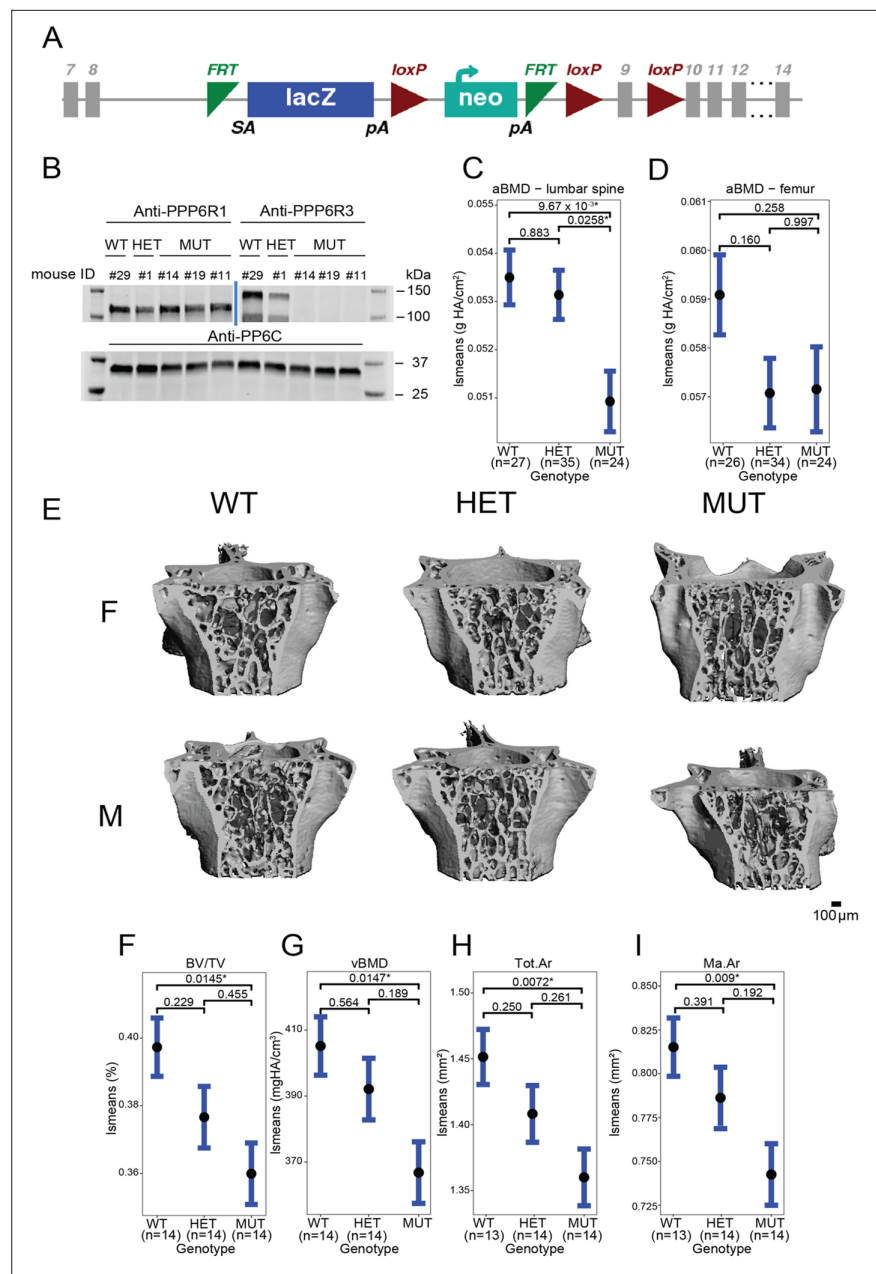


Figure 4. *Ppp6r3* functional validation shows an effect of genotype on bone mass. **(A)** Schematic of the *Ppp6r3* gene-trap allele (*Ppp6r3tm1a(KOMP)Wtsi*). Image obtained from the International Mouse Phenotype Consortium (IMPC). **(B)** Western blot of *Ppp6r3* in experimental mouse spleens. Top left panel shows that PPP6R1 protein (control) levels are not affected by the *Ppp6r3* gene-trap allele. Top right panel shows the effect of the gene-trap allele on PPP6R3 protein levels. The two bands are ostensibly due to different PPP6R3 isoforms. Bottom panel shows that PP6C protein (control) levels are not affected by the *Ppp6r3* gene-trap allele. For raw blots, refer to **Figure 4—source data 1**. Least-squares means for spinal **(C)** and femoral **(D)** areal BMD (aBMD) dual X-ray absorptiometry (DXA) in *Ppp6r3* wild-type (WT), heterozygous (HET), and mutant (MUT) mice. Contrast p values, adjusted for multiple comparisons are presented. * $p \leq 0.05$. **(E)** Representative images of vertebrae for the *Ppp6r3* experimental mice. Scale is shown on the bottom right. **(F, G)** Least-squares means for micro-computed tomography (μ CT) measurements in the lumbar spines of *Ppp6r3* WT, HET, and MUT mice. Contrast p values, adjusted for multiple comparisons are presented. * $p \leq 0.05$. **(H, I)** Least-squares means for μ CT measurements in the femoral midshaft of *Ppp6r3* WT, HET, and MUT mice. Contrast p values, adjusted for multiple comparisons are presented. * $p \leq 0.05$. Abbreviations: BV/TV – bone volume fraction, vBMD – volumetric bone mineral density, Tot.Ar – total area, Ma.Ar – medullary area. In panels **(C, D and F–I)**, the center points represent the least-square

Figure 4 continued on next page

Figure 4 continued

means, and the error bars represent confidence intervals at a confidence level of 0.9. Data presented here are from both male and female mice.

The online version of this article includes the following source data and figure supplement(s) for figure 4:

Source data 1. Raw, unedited, PPP6R3 western blots.

Figure supplement 1. LSBMD and FNBMD genome-wide association study (GWAS) SNPs in the PPP6R3 locus.

Figure supplement 2. *Ppp6r3* functional validation.

Figure supplement 3. Raman spectroscopy in femur, means.

Figure supplement 4. Raman spectroscopy in spines, means.

Figure supplement 5. Raman spectroscopy in femurs, standard deviation.

Figure supplement 6. Raman spectroscopy in spines, standard deviation.

TWAS and eQTL colocalization to GTEx and GWAS data on 4091 traits, including BMD, from the UK Biobank data. A total of 76 protein-coding genes were identified and of the 76, we identified 55 (72.4%) of the same genes in our implementation. There are several reasons for this discrepancy in the number of prioritized genes. First, both studies used a GWAS based on the UK Biobank (**Bycroft et al., 2018**); however, sample sizes were different. The PhenomeXcan project utilized GWAS data based on the analysis of ~207,000 individuals, whereas we used GWAS data based on the analysis of ~426,000 individuals (**Morris et al., 2019; Pividori et al., 2020**). Second, the two GWAS studies utilized different association models. Finally, due to the breadth of the PhenomeXcan project, they had a higher multiple-testing burden than we did, which led to different Bonferroni-adjusted p value thresholds ($p < 5.49 \times 10^{-10}$ vs. $p \leq 2.38 \times 10^{-6}$).

One of many novel genes identified in our study was *PPP6R3*, which was also identified in the PhenomeXcan project (**Pividori et al., 2020**). *PPP6R3* is a regulatory subunit of protein phosphatase 6 and has been implicated in several cancers (**Cristiano, 2020; Stefansson and Brautigan, 2006**). In humans, the *PPP6R3* protein shows ubiquitous expression across tissues, and may have an important role in maintaining immune self-tolerance (**Cristiano, 2020**). It is unclear how *PPP6R3* may be influencing BMD. However, protein phosphatase 6 has been shown to oppose activation of the nuclear factor kappa-light-chain enhancer of activated B cells (NF- κ B) pathway in lymphocytes (**Ziembik et al., 2017**). Since the NF- κ B signaling pathway is highly involved in osteoclastogenesis and bone resorption, it is possible that *PPP6R3* may be involved in the regulation of this pathway in osteoclasts (**Abu-Amer, 2013**). Further studies that characterize the role of *PPP6R3*, and the effects of its deletion, in bone cells are required to further elucidate its effect on BMD.

The *PPP6R3* locus demonstrated a high level of complexity, containing seven independent GWAS associations, at least one of which was also associated with fracture. Interestingly, just upstream of *PPP6R3* is *LRP5*, a WNT signaling co-receptor (**Mao et al., 2001**). *LRP5* is a well-known regulator of BMD and gain and loss of function mutations lead to high bone mass syndrome and osteoporosis pseudoglioma, respectively (**Mizuguchi et al., 2004; Boyden et al., 2002; Marques-Pinheiro et al., 2010; Gong et al., 2001**). *LRP5* expression was not significantly associated with eBMD by TWAS (Bonferroni $p = 1$), nor did it have a colocalizing eQTL in GTEx tissues (most significant RCP = 1.6×10^{-2} in pancreas). However, another eBMD lead SNP in the region, rs4988321, is a missense mutation in *LRP5* (Val667Met) that has been associated with BMD in multiple studies (**van Meurs et al., 2008; Brixen et al., 2007; Giroux et al., 2007**). While this variant represents an association that is independent of the rs10047483 association ($r^2 = 0.104$), it further highlights the complexity of this locus both in terms of the number of associations and target genes. We believe these data support a model of allelic heterogeneity involving multiple genes (at least *LRP5* and *PPP6R3*) at this locus, similar to other BMD loci identified by GWAS such as the 'Wnt16' locus (**Chesi et al., 2019**).

To determine the effect of *Ppp6r3* expression on bone, we characterized bone phenotypes in mice harboring a gene-trap allele (*Ppp6r3*^{tm1a(KOMP)Wtsi}). Consistent with the observation that the *PPP6R3* eQTL SNPs were significantly associated with lumbar spine, but not femoral neck BMD, we observed that *Ppp6r3* deletion had a significant effect on lumbar spine BMD, but not femoral BMD. Using μ CT, we further characterized the effect of *Ppp6r3* deletion on lumbar spine and femoral microarchitecture. We observed significant decreases in trabecular bone volume fraction (BV/TV) and vBMD of the lumbar spine as a function of *PPP6R3* genotype. While we did not observe significant effects of *Ppp6r3*

deletion on trabecular thickness or number, the direction of effects for those phenotypes suggests that the observed decrease in bone volume fraction and BMD may be explained by the cumulative but more subtle effects of *Ppp6r3* deletion on trabecular thickness and number. Consistent with our caliper-based observations that femoral width was decreased at the midshaft in *Ppp6r3* mutants, we observed a decrease in medullary and total area in the femoral midshaft. We also hypothesized that *Ppp6r3* deficiency might lead to defects in matrix composition; however, we did not observe any significant differences. The lack of differences might be due to the fact that we used relatively young mice for these assays.

Our hypothesis regarding the directions of effect of *Ppp6r3* expression on BMD based on the eQTL and eBMD/lumbar spine BMD GWAS were opposite to what we observed. There are several reasons that may explain this. First, our hypothesis was based on expression data in non-bone tissues and cell types. Recent studies have shown that the direction of eQTL effects can differ between different cells and tissues within humans (*Peters et al., 2016; Mizuno and Okada, 2019*). Second, our hypothesis was based on human data, while our functional experiments were performed in mice. Third, we globally deleted *Ppp6r3* in mice, as opposed to ablating it in a cell-type-specific manner. Future studies investigating which tissue/cell-type *Ppp6r3* is operative in and the generation of conditional *Ppp6r3* knockouts will allow us to unravel the precise role of this association and *Ppp6r3* in the regulation of bone mass.

As we and others have shown, the use of both TWAS and eQTL colocalization can prioritize putatively causal genes underlying GWAS associations. Here, we have shown the utility of this approach even in the absence of eQTL data from the most phenotype-relevant tissue. However, it is important to highlight the limitations of our analysis. While studies have shown that many eQTL are shared among tissues, the lack of eQTL data in bone and bone cells means that bone-specific eQTL were missed. For example, a study conducted by Mullin et al. performed eQTL colocalization and summary-based Mendelian randomization (SMR) by utilizing GWAS and expression data from osteoclast-like cells, and prioritized several eBMD genes (*Mullin et al., 2020*). Thirty-eight percent of the colocalizing eQTL and 19% of the SMR genes that they identified overlapped with our 512 prioritized genes, suggesting that we have missed many potential effector genes with eQTL specific to bone cells. In addition, the use of multiple non-bone tissues may have inflated the number of false positives based on coincidence of strong TWAS and eQTL colocalization signals that have no biological impact on bone. Furthermore, the lack of bone transcriptomic data may also explain the observed disparity between our hypothesized and observed direction-of-effect for *PPP6R3*. It is also important to note that due to the reliance of this approach on eQTL data, genes that affect BMD via non-expression-related mechanisms were not captured. Another limitation of our approach arises from the definition of loci based on linkage disequilibrium (LD). We used a set of previously defined approximately independent LD blocks, derived from a cohort of European individuals, in our fastENLOC analysis (*Berisa and Pickrell, 2016*). The inexact nature of these data may lead to spurious colocalizations due to mismatches in LD structure between the reference LD blocks and the GWAS/eQTL populations. Additionally, because the GWAS and eQTL data have mismatching LD structures, due to their being derived from cohorts with different ancestries, our analyses, particularly the colocalization analyses, may suffer from reduced power (*Hukku et al., 2021*). This also raises the related issue of the reduced generalizability of our results in non-European individuals, which brings further attention to the necessity of performing GWASs and providing reference data in diverse and underrepresented populations. Additionally, another issue arises when considering correlations in expression, and predicted expression, between genes in a locus, which may lead to spurious associations in TWAS analyses (*Wainberg et al., 2019*). Finally, as we show above, our method does not perform as well as prioritizing genes based on their proximity to GWAS associations. However, because our method utilizes systems genetics techniques and data, such as eQTL, we believe that our method prioritizes genes in a more biologically relevant manner. In fact, utilizing the closest gene method alone, *PPP6R3* would not have been prioritized as a bone-relevant gene. We suggest that future studies utilize both prioritization techniques, such as taking the closest genes to GWAS associations and cross referencing them with colocalizing and TWAS-associated genes, in order to provide further evidence for functional validation.

In summary, we applied a combined TWAS/colocalization approach using GTEx and identified 512 putatively causal BMD genes. We further investigated *PPP6R3* and demonstrated that it is a regulator of lumbar spine BMD. We believe this work provides a valuable resource for the bone genetics

community and may serve as a framework for prioritizing genes underlying GWAS associations using publicly available tools and data for a wide range of diseases.

Methods

fastENLOC colocalization

For each of the eBMD and fracture GWASs, we performed colocalization using fastENLOC, by following the tutorial and guidelines available at <https://github.com/xqwen/fastenloc> (Wen, 2022).

Briefly, for each GWAS, we converted variant coordinates to the hg38 human genome assembly, using the UCSC liftOver tool (minimum ratio of bases that must remap = 1; <https://genome.ucsc.edu/cgi-bin/hgLiftOver>). We calculated z-scores by dividing GWAS betas by standard errors. We then defined loci based on European LD blocks, as defined based on the results of *Berisa and Pickrell, 2016*.

Z-scores were then converted to posterior inclusion probabilities using torus (Wen, 2015). Finally, these data were colocalized with fastENLOC for all 49 GTEx V8 tissues, with the ‘-total_variants’ flag set to 14,000,000. Colocalization was performed using pre-computed GTEx multi-tissue annotations, obtained from <https://github.com/xqwen/fastenloc>; Wen, 2022. Finally, to identify protein-coding genes in the results, we utilized Ensembl’s ‘hsapiens_gene_ensembl’ dataset using biomaRt (version 2.45.8).

S-MultiXcan

We conducted a TWAS by integrating genome-wide SNP-level association summary statistics from an eBMD GWAS (Morris et al., 2019) with GTEx version 8 gene expression QTL data from 49 tissue types. We used the S-MultiXcan approach for this analysis, to correlate gene expression across tissues to increase power and identify candidate susceptibility genes (Barbeira et al., 2019). Default parameters were used, with the exception of the ‘--cutoff_condition_number’ parameter, which was set to 30. Bonferroni-correction of p values was performed on the resultant gene set (22,337 genes), using R’s ‘p.adjust’ function. This was followed by the removal of non-protein-coding genes. The analysis was also performed in the same manner using summary statistics from a fracture GWAS (Morris et al., 2019). Finally, to identify protein-coding genes in the results, we utilized Ensembl’s ‘hsapiens_gene_ensembl’ dataset using biomaRt (Durinck et al., 2009; Durinck et al., 2005).

Creation of the ‘known bone gene’ list

We generated a ‘known bone gene’ set as follows: First, we downloaded Gene Ontology IDs for the following terms: ‘osteo*’, ‘bone’, and ‘ossif*’ from AmiGO2 (version 2.5.13) (Carbon et al., 2009). After removal of non-bone-related terms, we extracted all mouse and human genes related to the GO terms, using biomaRt. From this list, we retained protein-coding genes.

We also used the ‘Human-Mouse: Disease Connection’ database available at the Mouse Genome Informatics website, to download human and mouse genes annotated with the terms ‘osteoporosis’, ‘bone mineral density’, ‘osteoblast’, ‘osteoclast’, and ‘osteocyte’. We used biomaRt to identify the gene biotypes, and retained protein-coding genes. We then used the MGI human-mouse homology table (http://www.informatics.jax.org/downloads/reports/HOM_MouseHumanSequence.rpt) to convert all mouse genes to their human homologs. Finally, we removed genes that were not interrogated in both the colocalization and the TWAS analyses.

GO enrichment analyses

GO analysis was performed for the set of protein-coding genes passing the colocalization threshold $RCP \geq 0.1$ and S-MultiXcan Bonferroni p value ≤ 0.05 , using the ‘topGO’ package (version 2.46.0) in R (Alexa and Rahnenfuhrer, 2021). Enrichment tests were performed for the ‘Molecular Function’, ‘Biological Process’, and ‘Cellular Component’ ontologies, using all protein-coding genes that were subjected to colocalization and MultiXcan analysis as background. Enrichment was performed using the ‘classic’ algorithm with Fisher’s exact test. p values were not adjusted for multiple testing.

LD calculations

LD between variants was calculated using the LDlinkR (version 1.0.2) R package, using the 'EUR' population (Myers *et al.*, 2020).

PPP6R3 knockout mouse generation

The study was carried out in strict accordance with NIH's Guide for the Care and Use of Laboratory Animals. Additionally, the University of Virginia Institutional Animal Care and Use Committee approved all animal procedures. *PPP6R3* gene-trap mice were generated using targeted embryonic stem cell clones heterozygous for the *PPP6R3*^{tm1a(KOMP)Wtsi} gene-trap allele obtained from the International Knockout Mouse Project (KOMP; <https://www.komp.org>). KOMP ES clones were karyotyped and injected using a XYClone Laser (Hamilton Thorne, Beverly, MA) into B6N-Tyr^{c-Brd}/BrdCrCrI (Charles River, Wilmington, MA) eight-cell stage embryos to create chimeric mice. Resultant chimeras were bred to B6N-Tyr^{c-Brd}/BrdCrCrI mice to obtain germline transmission of the *PPP6R3* gene-trap allele. From a breeding pair of two heterozygous mice, we generated our experimental population through HET × HET matings. Breeder mice were fed a breeder chow diet (Envigo Teklad S-2335 mouse breeder sterilizable diet, irradiated. Product # 7904), and experimental mice were fed a standard chow diet (Envigo Teklad LM-485 irradiated mouse/rat sterilizable diet. Product #7912).

Genotyping of PPP6R3 mice

DNA for genotyping was extracted from tail clips as follows: tail clips were incubated overnight at 55°C in a solution of 200 µl digestion/lysis buffer (Viagen Direct PCR [tail], Los Angeles, CA) and 1 mg/ml proteinase K (Viagen, Los Angeles, CA). After overnight incubation, tails were heated at 85°C for 45 min, and solutions were subsequently stored at 4°C.

For genotyping, PCR reactions were set up as follows. For each reaction, 1 µl of DNA was mixed with 24 µl of a master mix consisting of 19.5 µl nuclease-free H₂O, 2.5 µl 10× PCR reaction buffer (Invitrogen, Waltham, MA), 0.75 µl of MgCl₂ (Invitrogen, Waltham, MA), 0.5 µl of 10 mmol Quad dNTPs (Roche Diagnostics GmbH, Mannheim, Germany) 0.25 µl of Platinum Taq DNA polymerase (Invitrogen, Waltham, MA), and 0.25 µl of each primer, diluted to 20 µmol.

Primers: PCR primers were obtained from Integrated DNA Technologies, Coralville, IA.

Forward primer: 5'-CAC CTG GGT TGG TTA CAT CC-3'

Reverse primer: 5'-GAC CCT GCC TTA AAA CCA AA-3'

The following PCR settings were used:

- Initialization: 94°C, 120 s
- Denaturation: 94°C, 30 s (37 cycles)
- Annealing: 54°C, 30 s (37 cycles)
- Elongation: 72°C, 35 s (37 cycles)
- Final elongation: 72°C, 300 s

PCR products were run on a 2% agarose gel for 150 min at 60 V, to distinguish between wild-type, heterozygous and mutant *PPP6R3* mice.

PPP6R3 western blotting

Mouse spleens 20–40 mg in weight were suspended in 1% NP40 buffer (50 mM Tris [pH 8] 100 mM NaCl, 1% NP40, 1 mM EGTA(egtaic acid), 1 mM EDTA(ethylenediaminetetraacetic acid), Protease inhibitor cocktail (04-693-116-001, Roche), 1 mM PMSF(phenylmethylsulfonyl fluoride), 50 mM NaF, 0.2 mM sodium vanadate). The tissue was homogenized by RNase-free disposable pestles (Thermo Fisher #12-141-364) and incubated for 10 min on ice. After brief sonication, the sample was centrifuged for 10 min at 13,000 × rpm at 4°C. The protein concentration in the extract was measured by Bradford assay. 100 µg of sample protein was boiled 5 min in sodium dodecyl sulfate (SDS) sample buffer, loaded in each lane, resolved by gradient SDS–polyacrylamide gel electrophoresis (Bio-Rad #456-1085) and immunoblotted as described in Guergnon *et al.*, 2009. Primary antibodies were diluted 1:1000 (SAPS1 Ab: Thermo Fisher #PA5-44275, SAPS3 Ab: Thermo Fisher #PA5-58405, PP6C Ab: Sigma #HPA050940).

PPP6R3 functional validation

Experimental mice of both sexes were sacrificed at approximately 9 weeks of age (mean age = 61 days). At sacrifice, the right femurs were isolated, and femoral morphology (length and widths in anterior–posterior and medial–lateral orientations) was measured with digital calipers (Mitoyuto American, Aurora, IL).

Femurs were then wrapped in phosphate-buffered saline (PBS)-soaked gauze and stored at -20°C , until analysis. Lumbar vertebrae L3–L5 were also dissected at sacrifice and were wrapped in PBS-soaked gauze and frozen at -20°C . Given our prior experience in measuring bone geometry and microstructure, we used a minimum $N = 14$ for our analyses, ensuring 80% power to detect a statistically significant effect at an $\alpha \leq 0.05$.

Bulk RNA isolation, sequencing, and quantification

We isolated RNA from a randomly chosen subset ($n = 16$, 4/sex each of WT and MUT) of the available mice. Total RNA was isolated from L5 vertebrae, using the mirVana miRNA Isolation Kit (Life Technologies, Carlsbad, CA). Total RNA-seq libraries were constructed using Illumina TruSeq Stranded mRNA LT sample prep kits. Samples were sequenced to an average of 39 million 2×151 bp paired-end reads (total RNA-seq) on an Illumina NextSeq500 sequencer by Psomagen, Inc. A custom bioinformatics pipeline was used to quantify RNA-seq data. Briefly, RNA-seq FASTQ files were quality controlled using FASTQC (version 0.11.5) and MultiQC (version 1.11). Reads were trimmed using bbduk (bbmap package version 38.57). Trimmed reads were then aligned to the mm10 genome assembly with HISAT2 (version 2.1), and quantified with Stringtie (version 1.3.3). Read count information was then extracted with a Python script provided by the Stringtie website (prepDE.py).

Bulk RNA differential expression analyses

Using gene count matrices, differential expression was performed using DESeq2 (Love et al., 2014; Version 1.34.0) between wild-type and mutant samples. We used the 'DESeq' function from DESeq2, with a design formula of $\sim \text{sex} + \text{genotype}$. p values were adjusted using the 'p.value' function, using the 'BH' method.

Dual X-ray absorptiometry

Individual right femurs and the lumbar spine (L5 vertebrae) were isolated from surrounding soft tissues and frozen at -20°C in PBS. DXA was performed on the femurs and lumbar vertebrae using the Lunar Piximus II (GE Healthcare) as described previously by Beamer et al., 2011. In short, 10 isolated bones were placed in the detector field at a time and the samples were analyzed one by one, such that the region of interest (ROI) was set for one specimen at a time for data collection. The ROI for the femurs was on the entire isolated femur. For the spine, was on the entire isolated L5. Care was taken to ensure that the sample orientation was identical for all samples.

μCT and image analysis

All μCT analyses were carried out at the μCT Imaging Core Facility at Boston University using a Scanco Medical μCT 40 instrument (Brütisellen, Switzerland). The power, current, and integration time used for all scans were 70 kVp, 113 μA , and 200 ms, respectively. The L5 vertebrae and right femora were scanned at a resolution of 12 $\mu\text{m}/\text{voxel}$. Two volumes of interest (VOIs) in the L5 were selected for analysis: (1) the entire portion of the L5 vertebra extending from 60 μm caudal to the cranial growth plate in the vertebral body to 60 μm cranial to the caudal growth plate; and (2) only the trabecular centrum contained in the first VOI. Semi-automated-edge detection (Scanco Medical) was used to define the boundary between the trabecular centrum and cortical shell to produce the second VOI. Two VOIs were also analyzed for each femur: (1) a 0.3-mm-long segment of the diaphysis, centered at the mid-point of the bone; and (2) a 1.2-mm-long segment of the distal metaphyseal trabecular compartment. To define the location of the latter, the location of the distal femoral growth plate was determined, and the distal end of the VOI was set at 60 μm proximal to that growth plate. Gaussian filtering ($\sigma = 0.8$, support = 1) was used for partial background noise suppression. A scan of a potassium hydroxyapatite phantom allowed conversion of gray values to mineral density. For segmentation of bone tissue, the threshold was set at a 16-bit gray value of 7143 (521 mgHA/ccm), and this global threshold was applied to all of the samples. For each VOI, the following were calculated: total

volume (TV), bone volume (BV), bone volume fraction (BV/TV), BMD, and TMD. BMD was defined as the average density of all voxels in the VOI, whereas TMD was defined as the average density of all voxels in the VOI above the threshold (*Bouxsein et al., 2010*). For the second VOI, the following additional parameters were calculated: trabecular thickness (Tb.Th), trabecular separation (Tb.Sp), trabecular number (Tb.N), connectivity density (Conn.D), and structure model index (SMI) (*Bouxsein et al., 2010*).

For the femoral VOI, the following additional parameters were calculated instead: cortical thickness (Ct.Th), total area (Tot.Ar), marrow area (Ma.Ar), maximum and minimum moments of inertia, and polar moment of inertia (*Bouxsein et al., 2010*).

P1NP collection and quantification

Plasma was collected via submandibular bleeding from isoflurane anesthetized wild-type and mutant *PPP6R3* mice ($N = 10/\text{sex/genotype}$). Plasma P1NP levels were measured using commercially available kits from IDS (Gaithersburg, MD), according to the manufacturer's instructions. The assay sensitivity was 0.7 ng/ml. The intra-assay variation was 6.3%, and the inter-assay variation was 8.5%. All measurements were performed in duplicate.

Raman spectroscopy

Raman spectroscopy was performed using a Renishaw inVia Raman Microscope (Gloucestershire, UK) on each bone sample using a 785-nm-edge red incident laser. A rectangular filled map was created with 3 points in the x-axis and 20 points in the y-axis, for a total of 60 collected points. Each point was exposed 10 times for 6 s per exposure. A custom MATLAB script was used to evaluate the peak position, maximum intensity, peak width, full width at half maximum (FWHM), and the area under each peak. Peak area ratios were calculated for mineral:matrix, carbonate:phosphate, and crystallinity. Furthermore, the standard deviations of peak area ratios were calculated for each mouse, and were further used to evaluate the material heterogeneity in groups.

Statistical analyses

To calculate the enrichment of bone genes in prioritized genes, we performed Fisher's exact test, using R's 'fisher.test' function, with the alternative hypothesis set as 'greater'.

For the statistical analysis of the phenotyping results, we calculated least-squares means (lsmeans) using the 'emmeans' R package (version 1.5.2.1) (*Lenth, 2020*). Input for the lsmeans function was a linear model including terms for genotype, weight, and age in days. For sex-combined data, we also added a term for sex. For DXA phenotypes, we included a term for 'CenterRectX' and 'CenterRectY'. For the Raman spectroscopy data, weight and age were not included as terms in the linear model.

We used Tukey's HSD (honest significance test) test to test for significant differences in lsmeans, for each pair of genotype levels. Tukey's HSD also controls the family-wise error rate. We performed Welch two-sample *t*-test's, using R's 't.test' function, to quantify differences in P1NP levels between mutant and wild-type mice.

Analyses involving data from the International Mouse Phenotyping Consortium

For the IMPC data, we obtained data using their 'statistical-result' SOLR database, using the 'solrium' R package (version 1.1.4) (*Chamberlain, 2019*). We obtained experimental results using the 'Bone*Mineral*Density' parameter. We then pruned the resulting data to only include 'Successful' analyses, and removed experiments that included the skull. To generate the *Gpatch1* boxplot, we obtained raw data using from IMPC's 'statistical-raw-data' SOLR database for *Gpatch1*, and analyzed the data in the same manner as IMPC, using the 'OpenStats' R package (version 1.0.2), using the method = 'MM' and MM_BodyWeightIncluded = TRUE arguments (*Haselimashhadi et al., 2020*). Finally, mouse genes were converted to their human syntenic counterparts using Ensembl's 'hsapiens_gene_ensembl' and 'mmusculus_gene_ensembl' datasets through biomaRt.

PhenomeXcan data analysis

We obtained all significant PhenomeXcan gene-trait associations from their paper (<https://advances.sciencemag.org/content/6/37/eaba2083>), and used data for the

'3148_raw-Heel_bone_mineral_density_BMD' phenotype (Pividori et al., 2020). Furthermore, we constrained our search to only include genes that were annotated by the authors as 'protein_coding'.

LSBMD/FNBMD GWAS analysis

We obtained sex-combined LSBMD and FNBMD GWAS summary statistics from GEFOS (<http://www.gefos.org/?q=content/data-release-2012>), and then used a custom script that utilized the biomaRt R package to convert variants to their GRCh38 coordinates.

Data availability

eBMD and fracture GWAS summary statistics were obtained from GEFOS, as were the LSBMD and FNBMD GWAS summary statistics. GTEx eQTL data were obtained from the GTEx web portal. Data from the PhenomeXcan project were obtained from Pividori et al., 2020. Statistical data from the IMPC were obtained using an R interface to their SOLR database. *Ppp6r3* experimental data are provided on our GitHub (https://github.com/basel-maher/BMD_TWAS_colocalization; Al-Barghouthi, 2022). Mouse-Human homologs were obtained from MGI (http://www.informatics.jax.org/downloads/reports/HOM_MouseHumanSequence.rpt). We also obtained data from the MGI Human-Mouse:Disease Connection database (<http://www.informatics.jax.org/diseasePortal>). Gene Ontologies were obtained from AmiGO2 (<http://amigo.geneontology.org/amigo>).

Code availability

Analysis code and the raw data for our *Ppp6r3* functional validation analyses are available on GitHub (https://github.com/basel-maher/BMD_TWAS_colocalization, copy archived at [swh:1:rev:6aaa-8819c2e335013a665e76318dc98aeb9a52ce](https://www.swh.io/rev/6aaa-8819c2e335013a665e76318dc98aeb9a52ce); Al-Barghouthi, 2022).

Acknowledgements

Research reported in this publication was supported in part by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Number AR071657 to CRF, LCG, and EFM, and by the National Center for Research Resources of the National Institutes of Health under Award Number S10RR021072 to EFM. BMA-B. was supported in part by a National Institutes of Health, Biomedical Data Sciences Training Grant (5T32LM012416). The authors acknowledge Wenhao Xu (University of Virginia) and the Genetically Engineered Mouse Models (GEMM) core for their technical assistance in generating the *Ppp6r3* gene-trap mice. We thank Clifford J Rosen and Phuong T Le (MaineHealth Institute for Research) for performing the P1NP assay. We thank the IMPC for accessibility to BMD data in knockout mice (<https://www.mousephenotype.org/>). The data used for the analyses described in this manuscript were obtained from the IMPC SOLR database on 3/8/21. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 6/30/20.

Additional information

Competing interests

Cheryl Ackert-Bicknell: Reviewing editor, *eLife*. The other authors declare that no competing interests exist.

Funding

Funder	Grant reference number	Author
National Institute of Arthritis and Musculoskeletal and Skin Diseases	AR071657	Louis Gerstenfeld

Funder	Grant reference number	Author
National Center for Research Resources	S10RR021072	Jinho Heo Elise Morgan

The funders had no role in study design, data collection, and interpretation, or the decision to submit the work for publication.

Author contributions

Basel Maher Al-Barghouthi, Conceptualization, Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing – original draft, Project administration, Writing – review and editing; Will T Rosenow, Formal analysis; Kang-Ping Du, Resources, Investigation; Jinho Heo, Robert Maynard, Gina Calabrese, Aaron Nakasone, Bhavya Senwar, Investigation; Larry Mesner, Investigation, Writing – review and editing; Louis Gerstenfeld, Virginia Ferguson, Elise Morgan, Supervision, Funding acquisition, Writing – review and editing; James Larner, David Brautigan, Resources, Supervision; Cheryl Ackert-Bicknell, Supervision; Charles R Farber, Conceptualization, Resources, Supervision, Funding acquisition, Methodology, Writing – original draft, Project administration, Writing – review and editing

Author ORCIDs

Basel Maher Al-Barghouthi  <http://orcid.org/0000-0001-9816-8044>
Charles R Farber  <http://orcid.org/0000-0002-6748-4711>

Ethics

Our animal research protocol was reviewed and approved by the University of Virginia (UVA) Institutional Animal Care and Use Committee (IACUC), approved protocol #3741-12-20, titled 'Generation and characterization of mouse models of osteoporosis'. All experimental work was carried out under our approved UVA Institutional Biosafety Committee protocol #640-08.

Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.77285.sa1>

Author response <https://doi.org/10.7554/eLife.77285.sa2>

Additional files

Supplementary files

- Supplementary file 1. Supplementary data files associated with this manuscript. (a) 2156 protein-coding genes that are significant by S-MultiXcan transcriptome-wide association study (TWAS) analysis (Bonferroni p value ≤ 0.05). Columns are Ensembl ID, gene name, the nominal S-MultiXcan p value, the number of tissues available to S-MultiXcan, the number of independent components of variation among the tissues, and the Bonferroni-adjusted p value. (b) 1182 protein-coding genes that are significant in the fastEnloc colocalization analysis (regional colocalization probability [RCP] ≥ 0.1). Columns are Ensembl ID, signal cluster name from the expression quantitative trait loci (eQTL) analysis, the gene name, the RCP, and the colocalizing Genotype-Tissue Expression (GTEx) tissue. (c) The 512 protein-coding genes that are significant by both TWAS and colocalization (Bonferroni p value ≤ 0.05 and RCP ≥ 0.1). Columns are Ensembl ID, signal cluster name from the eQTL analysis, gene name, the RCP, the colocalizing GTEx tissue, the nominal S-MultiXcan p value, and the Bonferroni-adjusted p value. (d) Number of the 512 significantly colocalizing genes per GTEx tissue. Columns are the GTEx tissue and the number of unique colocalizing eQTL in the relevant tissue. (e) The 'known bone gene' list. Columns are gene name and Ensembl ID. (f) The 66 genes that are significant by both TWAS and colocalization (Bonferroni p value ≤ 0.05 and RCP ≥ 0.1), and are also members of the 'known bone gene' list. Columns are gene name and Ensembl ID. (g) Gene Ontology (GO) enrichments for the 512 protein-coding genes that are significant by both TWAS and colocalization (Bonferroni p value ≤ 0.05 and RCP ≥ 0.1). Only GO enrichments with a p value ≤ 0.05 are included. Columns are the GO IDs, GO terms, p values, the GO subontology (BP – Biological Process, CC – Cellular Component, MF – Molecular Function), and the Ensembl IDs for the genes that are members of the GO ontology. p values were calculated using a one-sided Fisher's exact test, and were not adjusted for multiple comparisons. (h) The 863 genes closest to estimated bone mineral density (eBMD) genome-wide association study (GWAS) associations. These data were obtained from *Morris et al., 2019*. (i) 137 novel putatively causal BMD genes, after increasing

RCP ≥ 0.5 , and removing genes that were members of the 'known bone gene' list and genes with a nominal ($p \leq 0.05$) alteration in BMD as determined by the International Mouse Phenotype Consortium (IMPC). Columns are Gene name, Ensembl ID, colocalization RCP, the tissue with the highest RCP, the S-Multixcan Bonferroni-adjusted p value, and the number of tissues that were significantly colocalizing. Note that TLN2 appears twice due to having the same RCP in two tissues. (j) Results of the differential expression analysis performed on RNA isolated from vertebrae. The analysis results show the difference in expression between mutant and wild-type mice. The columns are the Ensembl IDs, gene names, baseMean (the average of normalized count values, taken over all samples), log2 fold change, standard error of the log2 fold change estimate, the Wald statistic, the test p value, and the BH-adjusted p value.

- Transparent reporting form

Data availability

All data and source code are available on GitHub: https://github.com/basel-maher/BMD_TWAS_colocalization (copy archived at [swh:1:rev:6aaa8819c2e335013a665e76318dc98aeb9a52ce](https://swh.io/rev/6aaa8819c2e335013a665e76318dc98aeb9a52ce)).

The following previously published datasets were used:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Morris JA	2018	UK Biobank eBMD and Fracture GWAS Data Release 2018	http://www.gefos.org/?q=content/data-release-2018	GEFOS, release-2018
Estrada K	2012	Data Release 2012	http://www.gefos.org/?q=content/data-release-2012	GEFOS, release-2012
Pividori M	2020	PhenomeXcan	http://apps.hakyimlab.org/phenomexcan/	hakyimlab, phenomexcan
GTEC Consortium	2020	GTEC V8	https://gtexportal.org/home/datasets	gtexportal, V8
IMPC Consortium	2021	IMPC	https://www.mousephenotype.org/	mousephenotype, mousephenotype

References

- Abu-Amer Y.** 2013. Nf- κ b signaling and bone resorption. *Osteoporosis International* **24**:2377–2386. DOI: <https://doi.org/10.1007/s00198-013-2313-x>, PMID: 23468073
- Al-Barghouthi B.** 2022. BMD_TWAS_colocalization. [swh:1:rev:6aaa8819c2e335013a665e76318dc98aeb9a52ce](https://swh.io/rev/6aaa8819c2e335013a665e76318dc98aeb9a52ce). Software Heritage. https://archive.softwareheritage.org/swh:1:dir:868bf794273483babcafaaeab49f09b2207bfd17;origin=https://github.com/basel-maher/BMD_TWAS_colocalization;visit=swh:1:snp:4e5dc953f62a3e192cf830c4ba7f93a35284ce07;anchor=swh:1:rev:6aaa8819c2e335013a665e76318dc98aeb9a52ce
- Alexa A, Rahnenfuhrer J.** 2021. TopGO: enrichment analysis for gene ontology. 2.46.0. R Package. <https://bioconductor.org/packages/release/bioc/html/topGO.html>
- Alliance of genome resources.** 2022. Alliance of genome resources. 5.2.1. Alliancegenome. <https://www.alliancegenome.org/gene/hgnc:24658>
- Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, Torstenson ES, Shah KP, Garcia T, Edwards TL, Stahl EA, Huckins LM, GTEC Consortium, Nicolae DL, Cox NJ, Im HK.** 2018. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nature Communications* **9**:1825. DOI: <https://doi.org/10.1038/s41467-018-03621-1>, PMID: 29739930
- Barbeira AN, Pividori M, Zheng J, Wheeler HE, Nicolae DL, Im HK.** 2019. Integrating predicted transcriptome from multiple tissues improves association detection. *PLOS Genetics* **15**:e1007889. DOI: <https://doi.org/10.1371/journal.pgen.1007889>, PMID: 30668570
- Battle A, Brown CD, Engelhardt BE, Montgomery SB, GTEC Consortium, Laboratory, Data Analysis & Coordinating Center —Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEC groups, NIH Common Fund, NIH/NCI, NIH/NHGRI, NIH/NIMH, NIH/NIDA, Biospecimen Collection Source Site—NDRI, Biospecimen Collection Source Site—RPCI, Biospecimen Core Resource—VARI, Brain Bank Repository—University of Miami Brain Endowment Bank, Leidos Biomedical—Project Management, ELSI Study, Genome Browser Data Integration & Visualization—EBI, et al.** 2017. Genetic effects on gene expression across human tissues. *Nature* **550**:204–213. DOI: <https://doi.org/10.1038/nature24277>, PMID: 29022597

- Beamer WG**, Shultz KL, Coombs HF, DeMambro VE, Reinholdt LG, Ackert-Bicknell CL, Canalis E, Rosen CJ, Donahue LR. 2011. Bmd regulation on mouse distal chromosome 1, candidate genes, and response to ovariectomy or dietary fat. *Journal of Bone and Mineral Research* **26**:88–99. DOI: <https://doi.org/10.1002/jbmr.200>, PMID: 20687154
- Berisa T**, Pickrell JK. 2016. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics* **32**:283–285. DOI: <https://doi.org/10.1093/bioinformatics/btv546>, PMID: 26395773
- Bhattacharya A**, García-Closas M, Olshan AF, Perou CM, Troester MA, Love MI. 2020. A framework for transcriptome-wide association studies in breast cancer in diverse study populations. *Genome Biology* **21**:42. DOI: <https://doi.org/10.1186/s13059-020-1942-6>, PMID: 32079541
- Black DM**, Rosen CJ. 2016. Clinical practice: postmenopausal osteoporosis. *The New England Journal of Medicine* **374**:254–262. DOI: <https://doi.org/10.1056/NEJMc1513724>, PMID: 26789873
- Bouxsein ML**, Boyd SK, Christiansen BA, Guldberg RE, Jepsen KJ, Müller R. 2010. Guidelines for assessment of bone microstructure in rodents using micro-computed tomography. *Journal of Bone and Mineral Research* **25**:1468–1486. DOI: <https://doi.org/10.1002/jbmr.141>, PMID: 20533309
- Boyden LM**, Mao J, Belsky J, Mitzner L, Farhi A, Mitnick MA, Wu D, Insogna K, Lifton RP. 2002. High bone density due to a mutation in LDL-receptor-related protein 5. *The New England Journal of Medicine* **346**:1513–1521. DOI: <https://doi.org/10.1056/NEJMoa013444>, PMID: 12015390
- Brixen K**, Beckers S, Peeters A, Piters E, Balemans W, Nielsen TL, Wraae K, Bathum L, Brasen C, Hagen C, Andersen M, Van Hul W, Abrahamsen B. 2007. Polymorphisms in the low-density lipoprotein receptor-related protein 5 (LRP5) gene are associated with peak bone mass in non-sedentary men: results from the Odense androgen study. *Calcified Tissue International* **81**:421–429. DOI: <https://doi.org/10.1007/s00223-007-9088-z>, PMID: 18058054
- Burge R**, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A. 2007. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005–2025. *Journal of Bone and Mineral Research* **22**:465–475. DOI: <https://doi.org/10.1359/jbmr.061113>, PMID: 17144789
- Bycroft C**, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O’Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J. 2018. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**:203–209. DOI: <https://doi.org/10.1038/s41586-018-0579-z>, PMID: 30305743
- Calabrese GM**, Mesner LD, Stains JP, Tommasini SM, Horowitz MC, Rosen CJ, Farber CR. 2017. Integrating GWAS and co-expression network data identifies bone mineral density genes SPTBN1 and MARK3 and an osteoblast functional module. *Cell Systems* **4**:46–59. DOI: <https://doi.org/10.1016/j.cels.2016.10.014>, PMID: 27866947
- Carbon S**, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S, Hub A, Group WPW. 2009. AmiGO: online access to ontology and annotation data. *Bioinformatics* **25**:288–289. DOI: <https://doi.org/10.1093/bioinformatics/btn615>, PMID: 19033274
- Chamberlain S**. 2019. Solrium: general purpose R interface to “Solr. 1.1.4. R Package. <https://github.com/ropensci/solrium>
- Chesi A**, Wagley Y, Johnson ME, Manduchi E, Su C, Lu S, Leonard ME, Hodge KM, Pippin JA, Hankenson KD, Wells AD, Grant SFA. 2019. Genome-Scale capture C promoter interactions implicate effector genes at GWAS loci for bone mineral density. *Nature Communications* **10**:1260. DOI: <https://doi.org/10.1038/s41467-019-09302-x>, PMID: 30890710
- Consortium Gte**. 2013. The genotype-tissue expression (gtex) project. *Nature Genetics* **45**:580–585. DOI: <https://doi.org/10.1038/ng.2653>, PMID: 23715323
- Consortium Gte**. 2020. The gtex Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**:1318–1330. DOI: <https://doi.org/10.1126/science.aaz1776>, PMID: 32913098
- Cristiano L**. 2020. PPP6R3 (protein phosphatase 6 regulatory subunit 3). *Atlas of Genetics and Cytogenetics in Oncology and Haematology* **01**:70657. DOI: <https://doi.org/10.4267/2042/70657>
- Dickinson ME**, Flenniken AM, Ji X, Teboul L, Wong MD, White JK, Meehan TF, Weninger WJ, Westerberg H, Adissu H, Baker CN, Bower L, Brown JM, Caddle LB, Chiani F, Clary D, Cleak J, Daly MJ, Denegre JM, Doe B, et al. 2016. High-throughput discovery of novel developmental phenotypes. *Nature* **537**:508–514. DOI: <https://doi.org/10.1038/nature19356>, PMID: 27626380
- Durinck S**, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, Huber W. 2005. BioMart and bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**:3439–3440. DOI: <https://doi.org/10.1093/bioinformatics/bti525>, PMID: 16082012
- Durinck S**, Spellman PT, Birney E, Huber W. 2009. Mapping identifiers for the integration of genomic datasets with the R/bioconductor package biomart. *Nature Protocols* **4**:1184–1191. DOI: <https://doi.org/10.1038/nprot.2009.97>, PMID: 19617889
- Edwards SL**, Beesley J, French JD, Dunning AM. 2013. Beyond gwas: illuminating the dark road from association to function. *American Journal of Human Genetics* **93**:779–797. DOI: <https://doi.org/10.1016/j.ajhg.2013.10.012>, PMID: 24210251
- Estrada K**, Styrkarsdottir U, Evangelou E, Hsu Y-H, Duncan EL, Ntzani EE, Oei L, Albagha OME, Amin N, Kemp JP, Koller DL, Li G, Liu C-T, Minster RL, Moayyeri A, Vandenput L, Willner D, Xiao S-M, Yerges-Armstrong LM, Zheng H-F, et al. 2012. Genome-Wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nature Genetics* **44**:491–501. DOI: <https://doi.org/10.1038/ng.2249>, PMID: 22504420

- Fitzpatrick LA.** 2002. Secondary causes of osteoporosis. *Mayo Clinic Proceedings* **77**:453–468. DOI: <https://doi.org/10.4065/77.5.453>, PMID: 12004995
- Gamazon ER,** Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, Eyer AE, Denny JC, GTEx Consortium, Nicolae DL, Cox NJ, Im HK. 2015. A gene-based association method for mapping traits using reference transcriptome data. *Nature Genetics* **47**:1091–1098. DOI: <https://doi.org/10.1038/ng.3367>, PMID: 26258848
- Giambartolomei C,** Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V. 2014. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLOS Genetics* **10**:e1004383. DOI: <https://doi.org/10.1371/journal.pgen.1004383>, PMID: 24830394
- Giral H,** Landmesser U, Kratzer A. 2018. Into the wild: GWAS exploration of non-coding rnas. *Frontiers in Cardiovascular Medicine* **5**:181. DOI: <https://doi.org/10.3389/fcvm.2018.00181>, PMID: 30619888
- Giroux S,** Elfassih L, Cardinal G, Laflamme N, Rousseau F. 2007. Lrp5 coding polymorphisms influence the variation of peak bone mass in a normal population of French-Canadian women. *Bone* **40**:1299–1307. DOI: <https://doi.org/10.1016/j.bone.2007.01.004>, PMID: 17307038
- Gong Y,** Slee RB, Fukai N, Rawadi G, Roman-Roman S, Reginato AM, Wang H, Cundy T, Glorieux FH, Lev D, Zacharin M, Oexle K, Marcelino J, Suwairi W, Heeger S, Sabatakos G, Apte S, Adkins WN, Allgrove J, Arslan-Kirchner M, et al. 2001. Ldl receptor-related protein 5 (LRP5) affects bone accrual and eye development. *Cell* **107**:513–523. DOI: [https://doi.org/10.1016/s0092-8674\(01\)00571-2](https://doi.org/10.1016/s0092-8674(01)00571-2), PMID: 11719191
- Guergnon J,** Derewenda U, Edelson JR, Brautigan DL. 2009. Mapping of protein phosphatase-6 association with its SAPs domain regulatory subunit using a model of helical repeats. *BMC Biochemistry* **10**:24. DOI: <https://doi.org/10.1186/1471-2091-10-24>, PMID: 19835610
- Gusev A,** Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, Jansen R, de Geus EJC, Boomsma DI, Wright FA, Sullivan PF, Nikkola E, Alvarez M, Civelek M, Lusi AJ, Lehtimäki T, Raitoharju E, Kähönen M, Seppälä I, Raitakari OT, et al. 2016. Integrative approaches for large-scale transcriptome-wide association studies. *Nature Genetics* **48**:245–252. DOI: <https://doi.org/10.1038/ng.3506>, PMID: 26854917
- Haselimashhadi H,** Mason JC, Mallon A-M, Smedley D, Meehan TF, Parkinson H. 2020. OpenStats: a robust and scalable software package for reproducible analysis of high-throughput phenotypic data. *PLOS ONE* **15**:e0242933. DOI: <https://doi.org/10.1371/journal.pone.0242933>, PMID: 33378393
- Hukku A,** Pividori M, Luca F, Pique-Regi R, Im HK, Wen X. 2021. Probabilistic colocalization of genetic variants from complex and molecular traits: promise and limitations. *American Journal of Human Genetics* **108**:25–35. DOI: <https://doi.org/10.1016/j.ajhg.2020.11.012>, PMID: 33308443
- Kemp JP,** Morris JA, Medina-Gomez C, Forgetta V, Warrington NM, Youten SE, Zheng J, Gregson CL, Grundberg E, Trajanoska K, Logan JG, Pollard AS, Sparkes PC, Ghirardello EJ, Allen R, Leitch VD, Butterfield NC, Komla-Ebri D, Adoum A-T, Curry KF, et al. 2017. Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nature Genetics* **49**:1468–1475. DOI: <https://doi.org/10.1038/ng.3949>, PMID: 28869591
- Lattin JE,** Schroder K, Su AI, Walker JR, Zhang J, Wiltshire T, Saijo K, Glass CK, Hume DA, Kellie S, Sweet MJ. 2008. Expression analysis of G protein-coupled receptors in mouse macrophages. *Immunome Research* **4**:5. DOI: <https://doi.org/10.1186/1745-7580-4-5>, PMID: 18442421
- Lenth R.** 2020. Emmeans: estimated marginal means, aka least-squares means. 1.5.2.1. R Package. <https://github.com/rvnlenth/emmeans>
- Love MI,** Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-Seq data with *deseq2*. *Genome Biology* **15**:550. DOI: <https://doi.org/10.1186/s13059-014-0550-8>, PMID: 25516281
- Mao J,** Wang J, Liu B, Pan W, Farr GH, Flynn C, Yuan H, Takada S, Kimelman D, Li L, Wu D. 2001. Low-density lipoprotein receptor-related protein-5 binds to axin and regulates the canonical wnt signaling pathway. *Molecular Cell* **7**:801–809. DOI: [https://doi.org/10.1016/s1097-2765\(01\)00224-6](https://doi.org/10.1016/s1097-2765(01)00224-6), PMID: 11336703
- Marques-Pinheiro A,** Levasseur R, Cormier C, Bonneau J, Boileau C, Varret M, Abifadel M, Allanore Y. 2010. Novel LRP5 gene mutation in a patient with osteoporosis-pseudoglioma syndrome. *Joint Bone Spine* **77**:151–153. DOI: <https://doi.org/10.1016/j.jbspin.2009.11.013>, PMID: 20096619
- Miller PD,** Zapalowski C, Kulak CA, Bilezikian JP. 1999. Bone densitometry: the best way to detect osteoporosis and to monitor therapy. *The Journal of Clinical Endocrinology and Metabolism* **84**:1867–1871. DOI: <https://doi.org/10.1210/jcem.84.6.5710>, PMID: 10372677
- Mirza F,** Canalis E. 2015. Management of endocrine disease: secondary osteoporosis: pathophysiology and management. *European Journal of Endocrinology* **173**:R131–R151. DOI: <https://doi.org/10.1530/EJE-15-0118>, PMID: 25971649
- Mizuguchi T,** Furuta I, Watanabe Y, Tsukamoto K, Tomita H, Tsujihata M, Ohta T, Kishino T, Matsumoto N, Minakami H, Niikawa N, Yoshiura K-I. 2004. Lrp5, low-density-lipoprotein-receptor-related protein 5, is a determinant for bone mineral density. *Journal of Human Genetics* **49**:80–86. DOI: <https://doi.org/10.1007/s10038-003-0111-6>, PMID: 14727154
- Mizuno A,** Okada Y. 2019. Biological characterization of expression quantitative trait loci (eQTLs) showing tissue-specific opposite directional effects. *European Journal of Human Genetics* **27**:1745–1756. DOI: <https://doi.org/10.1038/s41431-019-0468-4>, PMID: 31296926
- Morris JA,** Kemp JP, Youten SE, Laurent L, Logan JG, Chai RC, Vulpescu NA, Forgetta V, Kleinman A, Mohanty ST, Sergio CM, Quinn J, Nguyen-Yamamoto L, Luco A-L, Vijay J, Simon M-M, Pramatarova A, Medina-Gomez C, Trajanoska K, Ghirardello EJ, et al. 2019. An atlas of genetic influences on osteoporosis in humans and mice. *Nature Genetics* **51**:258–266. DOI: <https://doi.org/10.1038/s41588-018-0302-x>, PMID: 30598549

- Mullin BH**, Tickner J, Zhu K, Kenny J, Mullin S, Brown SJ, Dudbridge F, Pavlos NJ, Mocarski ES, Walsh JP, Xu J, Wilson SG. 2020. Characterisation of genetic regulatory effects for osteoporosis risk variants in human osteoclasts. *Genome Biology* **21**:80. DOI: <https://doi.org/10.1186/s13059-020-01997-2>, PMID: 32216834
- Myers TA**, Chanock SJ, Machiela MJ. 2020. ldlr: an R package for rapidly calculating linkage disequilibrium statistics in diverse populations. *Frontiers in Genetics* **11**:157. DOI: <https://doi.org/10.3389/fgene.2020.00157>, PMID: 32180801
- Nica AC**, Dermitzakis ET. 2013. Expression quantitative trait loci: present and future. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **368**:20120362. DOI: <https://doi.org/10.1098/rstb.2012.0362>, PMID: 23650636
- Peacock M**, Turner CH, Econs MJ, Foroud T. 2002. Genetics of osteoporosis. *Endocrine Reviews* **23**:303–326. DOI: <https://doi.org/10.1210/edrv.23.3.0464>, PMID: 12050122
- Peters JE**, Lyons PA, Lee JC, Richard AC, Fortune MD, Newcombe PJ, Richardson S, Smith KGC. 2016. Insight into genotype-phenotype associations through eQTL mapping in multiple cell types in health and immune-mediated disease. *PLOS Genetics* **12**:e1005908. DOI: <https://doi.org/10.1371/journal.pgen.1005908>, PMID: 27015630
- Pividori M**, Rajagopal PS, Barbeira A, Liang Y, Melia O, Bastarache L, Park Y, Consortium Gte, Wen X, Im HK. 2020. PhenomeXcan: mapping the genome to the phenome through the transcriptome. *Science Advances* **6**:eaba2083. DOI: <https://doi.org/10.1126/sciadv.aba2083>, PMID: 32917697
- Ralston SH**, Uitterlinden AG. 2010. Genetics of osteoporosis. *Endocrine Reviews* **31**:629–662. DOI: <https://doi.org/10.1210/er.2009-0044>, PMID: 20431112
- Rocha-Braz MGM**, Ferraz-de-Souza B. 2016. Genetics of osteoporosis: searching for candidate genes for bone fragility. *Archives of Endocrinology and Metabolism* **60**:391–401. DOI: <https://doi.org/10.1590/2359-3997000000178>, PMID: 27533615
- Sabik OL**, Farber CR. 2017. Using GWAS to identify novel therapeutic targets for osteoporosis. *Translational Research* **181**:15–26. DOI: <https://doi.org/10.1016/j.trsl.2016.10.009>, PMID: 27837649
- Stefansson B**, Brautigan DL. 2006. Protein phosphatase 6 subunit with conserved Sit4-associated protein domain targets I κ B ϵ . *The Journal of Biological Chemistry* **281**:22624–22634. DOI: <https://doi.org/10.1074/jbc.M601772200>, PMID: 16769727
- Swan AL**, Schütt C, Rozman J, Del Mar Muñiz Moreno M, Brandmaier S, Simon M, Leuchtenberger S, Griffiths M, Brommage R, Keskivali-Bond P, Grallert H, Werner T, Teperino R, Becker L, Miller G, Moshiri A, Seavitt JR, Cissell DD, Meehan TF, Acar EF, et al. 2020. Mouse mutant phenotyping at scale reveals novel genes controlling bone mineral density. *PLOS Genetics* **16**:e1009190. DOI: <https://doi.org/10.1371/journal.pgen.1009190>, PMID: 33370286
- Thom CS**, Voight BF. 2020. Genetic colocalization atlas points to common regulatory sites and genes for hematopoietic traits and hematopoietic contributions to disease phenotypes. *BMC Medical Genomics* **13**:89. DOI: <https://doi.org/10.1186/s12920-020-00742-9>, PMID: 32600345
- van Meurs JBJ**, Trikalinos TA, Ralston SH, Balcells S, Brandi ML, Brixen K, Kiel DP, Langdahl BL, Lips P, Ljunggren O, Lorenc R, Obermayer-Pietsch B, Ohlsson C, Pettersson U, Reid DM, Rousseau F, Scollen S, Van Hul W, Agueda L, Akesson K, et al. 2008. Large-Scale analysis of association between LRP5 and LRP6 variants and osteoporosis. *JAMA* **299**:1277–1290. DOI: <https://doi.org/10.1001/jama.299.11.1277>, PMID: 18349089
- Wainberg M**, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, Ermel R, Ruusalepp A, Quertermous T, Hao K, Björkegren JLM, Im HK, Pasaniuc B, Rivas MA, Kundaje A. 2019. Opportunities and challenges for transcriptome-wide association studies. *Nature Genetics* **51**:592–599. DOI: <https://doi.org/10.1038/s41588-019-0385-z>, PMID: 30926968
- Wen X**. 2015. Effective QTL Discovery Incorporating Genomic Annotations. *Genetics*. DOI: <https://doi.org/10.1101/032003>
- Wen X**, Pique-Regi R, Luca F. 2017. Integrating molecular QTL data into genome-wide genetic association analysis: probabilistic assessment of enrichment and colocalization. *PLOS Genetics* **13**:e1006646. DOI: <https://doi.org/10.1371/journal.pgen.1006646>, PMID: 28278150
- Wen X**. 2022. fastENLOC: fast enrichment estimation aided colocalization analysis. GitHub. <https://github.com/xqwen/fastenloc>
- Ziembik MA**, Bender TP, Lerner JM, Brautigan DL. 2017. Functions of protein phosphatase-6 in NF- κ B signaling and in lymphocytes. *Biochemical Society Transactions* **45**:693–701. DOI: <https://doi.org/10.1042/BST20160169>, PMID: 28620030