A role for the *MEGF6* gene in predisposition to osteoporosis

Craig C. Teerlink PhD,[1,*] Michael J Jurynec PhD,[2,*] Rolando Hernandez BS,[3] Jeff Stevens MS,[1]

Dana C. Hughes PhD,[1,4] Cherie P. Brunker MD,[5,6] Kerry Rowe PhD,[6] David J. Grunwald PhD,[7]

Julio C. Facelli PhD,[3] Lisa A. Cannon-Albright PhD[1,8,9]

1. Genetic Epidemiology, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, 84132, USA

2. Department of Orthopaedics , University of Utah School of Medicine, Salt Lake City, 84108, USA

3. Department of Biomedical Informatics, University of Utah School of Medicine, Salt Lake City, 84108, USA

4. Rehabilitation Services, University of Utah Health, Salt Lake City, UT, 84132, USA

5. Geriatrics, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, 84132, USA

6. Intermountain Healthcare, Salt Lake City, UT, 84113, USA

7. Department of Human Genetics, University of Utah, Salt Lake City, UT, 84112, USA

8. George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, UT, 84148, USA

9. Huntsman Cancer Institute, Salt Lake City, UT, 84112, USA.

Corresponding author
Craig C. Teerlink, PhD
Phone: (801)587-9303, Fax: (801)581-6052, email: craig.teerlink@utah.edu

*The authors wish it to be known that, in their opinion, the first 2 authors should be regarded as joint First Authors.

ABSTRACT

Osteoporosis is a common skeletal disorder characterized by deterioration of bone tissue in later life. The set of genetic factors contributing to osteoporosis is not completely specified. High-risk osteoporosis pedigrees were analyzed to identify genes that may confer susceptibility to disease. Candidate predisposition variants were identified initially by whole exome sequencing of affected-relative-pairs, approximately cousins, from ten pedigrees. Variants were filtered on the basis of population frequency, concordance between pairs of cousins, affecting a gene associated with osteoporosis, and likelihood to have functionally damaging, pathogenic consequences. Subsequently variants were tested for segregation in 68 additional relatives of the index carriers. A rare variant in *MEGF6* (rs755467862) showed strong evidence of segregation with the disease phenotype. Predicted protein folding indicated the variant (Cys200Tyr) may disrupt structure of an EGF-like calcium-binding domain of MEGF6. Functional analyses demonstrated that complete loss of the paralogous genes *megf6a* and *megf6b* in zebrafish resulted in significant delay of cartilage and bone formation. Segregation analyses, in-silico protein structure modeling, and functional assays support a role for *MEGF6* in predisposition to osteoporosis.

INTRODUCTION

Osteoporosis is a common skeletal disorder characterized by deterioration of bone mineral density (BMD) and increased risk of fracture (1). Identification of genetic factors that have a strong influence on disease could enhance screening and knowledge concerning biological pathways involved. Genome-wide association studies (GWAS) have identified over 500 common single nucleotide polymorphisms (SNPs) associated with BMD (2), revealing several relevant biological pathways including genes involved in WNT signaling, the RANK/RANKL/OPG pathway, and genes involved in endochondral ossification. GWAS studies have revealed that osteoporosis is highly polygenic. SNPs identified thus far are estimated to account for about 20% of trait variation (3), whereas overall heritability estimates for BMD are higher, ranging from 50 to 90% (4, 5). The difference may in part be due to rare genetic variations whose effects are not typically detectable in GWAS. Family based designs present a complementary approach that may provide opportunities to discover rare, penetrant variants (6). Using a multi-generational pedigree resource, we performed exome sequencing on affected-relative-pairs with a deficiency in BMD from 10 high-risk osteoporosis pedigrees to identify rare, coding variants that may be contributing to osteoporosis risk in these families. Analysis of this resource identified a rare coding variant in the *MEGF6* (*multiple epidermal growth factor-like protein 6*, alternately known as *EGFL3*) gene that segregates to additional family members with osteoporosis. Computational and functional analyses indicate that the *MEGF6* variant likely disrupts wildtype (WT) protein function and that zebrafish paralogues contribute to bone and cartilage formation *in vivo*.

3

RESULTS

Variant prioritization

To identify genes associated with osteoporosis, multigenerational pedigrees were identified in which near-cousin pairs as well as additional individuals were confirmed to have a diagnosis of osteoporosis (see Materials and Methods). Exome sequencing of affected near-cousin pairs in each of ten pedigrees identified a total of ~45K variants. Among these, 548 coding variants in 480 genes were rare (MAF<0.005 in EXAC browser (7)) and concordant between a sequenced pair of affected cousins. Restricting candidates to identity-by-descent (IBD) regions for sequenced pairs as identified by shared genomic segments (SGS) analysis (8) further limited the selection to 125 rare, shared candidate variants in 111 genes. A search of PubMed for each of the 111 genes to identify previous reported association to osteoporosis further limited selection to variants in 12 candidate genes. In a fourth filtering step, variants that were scored as 'damaging' by at least 8 of 10 in-silico pathogenicity prediction algorithms included in ANNOVAR variant annotations were retained, resulting in 3 candidate genes (*MEGF6* (rs755467862), *PAX8* (rs1269341622), and *UMPS* (rs777350640)).

Confirmation of candidate variants via segregation

Six additional carriers of the *MEGF6* variant c.599 T>C; p.Cys200Tyr (rs755467862) were identified among 14 relatives of the index cousins that were assayed for the variant. Among the eight variant carriers identified in the pedigree, the sequenced index cousins as well as one relative had a physician-confirmed diagnosis of osteoporosis, four additional individuals had a physician confirmed diagnosis of osteopenia, and one individual was unaffected (screened in the

4

50th decade of life); four of the eight carriers had a medical history of bone fractures, and four had been treated previously for osteoporosis (Figure 1). Of the remaining six assayed relatives that did not carry the variant, none had a physician-confirmed diagnosis of osteoporosis. The de-identified pedigree segregating rs755467862 is depicted in Figure 1 and had significant evidence of segregation to affected relatives (osteoporosis or osteopenia, RVsharing p=0.0011) after correcting for 3 candidate variants (α=0.05/3=0.017) (9).

In silico functional evaluations of rs755467862

Variant rs755467862 is a non-synonymous variant substituting base T for C (c.599 T>C) in exon 5 that codes a Cysteine > Tyrosine amino acid substitution (p.Cys200Tyr). Population frequency has been estimated (MAF/number of alleles tested) as T=4e-5/243,578 (10). Variant rs755467862 was predicted to be damaging by almost all of the functional prediction software included in Annovar (11), including Polyphen (12), LRT_pred (13), MutationTaster (14), FATHMM (15), Radial_SVM (16), LR score (16), but not SIFT (17). The GERP score for the variant is 5.0, indicating strong evidence for conservation of the reference allele across species (18).

The candidate variant rs755467862 produces an amino acid substitution Cys200Tyr in the second EGF-like calcium binding domain of MEGF6. As MEGF6 is a large protein of 1541 amino acids, the entire protein is not amenable to existing 3D structure prediction methods. I-TASSER software (19) predicted the structure of the affected EGF-like calcium binding domain (residues 161-201) for both wild type and variant amino acid sequences, resulting in two distinct

5

candidate structures for the wild type and five structures for variant proteins. From the ten possible comparisons, five (50%) showed substantial differences between the wild type and variant predicted structures. An example of this is presented in Figure 2 where it is apparent that the amino acid substitution disrupts a disulfide bond between residues 187 and 200 in the second EGF-like calcium-binding domain, where a helix is formed instead of an open loop. Similar behavior was observed in the other comparisons. The main structure of EGF-like domains is a two-stranded β-sheet followed by a loop where the calcium ion binds (20). The collapse of this loop may eliminate this calcium binding site in MEGF6, which may decrease, or even eliminate, the functionality of the protein.

Protein interaction network

As variation in *MEGF6* expression or function has not been shown to be associated with bone development or homeostasis, we analyzed whether its protein product might have potential interactions with proteins more directly related to bone formation. Analysis with the STRING database tool (21) showed MEGF6 is related to pathways involved in various biological processes (Figure 3A) including angiogenesis (*EMILIN1*, a protein involved in vessel assembly (22), and most notably, osteoporosis (*ZBTB40* (23, 24)). Figure 3B contains a STRING graph re-centered on *ZBTB40* that shows connections with *WNT16*, which has been previously associated with BMD (25).

Analysis of zebrafish *megf6a* and *megf6b* gene expression and function

Our analyses indicated that a coding variant of *MEGF6* segregated with the occurrence of osteoporosis, that the variant was likely to be damaging to gene function, and that there was support for a functional link between the protein product of *MEGF6* and pathways known to contribute to osteoporosis. To uncover biological functions of *MEGF6*, we examined the consequences of complete loss of gene function in the zebrafish.

The zebrafish genome contains two *megf6* paralogues, *megf6a* and *megf6b*, which are orthologous to the single human *MEGF6* gene. Each orthologue has similar degree of identity (*megf6a* - 56% and *megf6b* – 54%) and similarity (*megf6a* - 68% and *megf6b* – 64%) to the human orthologue. The expression patterns of *megf6a* and *megf6b* were determined by whole-mount *in situ* hybridization (WISH) at 24, 36, and 48 hours post fertilization (hpf). At 24 hpf *megf6a* is broadly expressed in the head (Supplemental Figure 1A). At 36 hpf it is localized to the apical ectodermal ridge (AER) of the pectoral fin (Supplemental Figure 1B), the atrioventricular canal (Supplemental Figure 1C), and the notochord (Supplemental Figure 1D), and it is broadly expressed in the head. *megf6b* is expressed in the same domains as *megf6a* at 24 hpf, but it is also expressed in the developing somites (Supplemental Figure 1E). The expression pattern of *megf6b* is maintained in the fin fold and head at 36 hpf and is also expressed in the AER of the pectoral fin and in the notochord (Supplemental Figure 1F). Both *megf6a* and *megf6b* continue to be expressed in the notochord of 48 hpf larvae (data not shown). In the head of 48 hpf larvae, *megf6b* is expressed specifically in the mouth and pharyngeal region (Supplemental Figure 1G). These data indicate that *megf6a* and *megf6b* have mostly overlapping expression domains during the first 48 hpf of development and are expressed in regions of the embryo that give rise to cartilage and bone.

To generate zebrafish lacking both *megf6a* and *megf6b* gene function, we first generated a germline deletion of *megf6b*. Both *megf6a* and *megf6b* are large genes containing 34 and 38 exons, respectively. To make a large deletion in *megf6b*, we injected CRISPR/Cas9 ribonucleoprotein (RNP) complexes targeting exons 6 and 35 into one-cell stage embryos (26). We identified one allele, *z48* - a 51,860 bp deletion with a 53 bp insertion, which stably transmitted through the germline (Supplemental Figure 2). *megf6b* mutants are homozygous viable and the only overt phenotype is ectopic migration of a small number of melanocytes in the trunk and tail of 48 hpf larvae (Figure 4A-B').

Paralogous genes in zebrafish can have redundant biological functions (27). As *megf6a* and *megf6b* are both expressed in the fin fold, head, and notochord, we hypothesized they might have partially redundant functions. Therefore, we generated embryos lacking both genes by injection of *megf6b* mutant embryos with CRISPR-Cas9 RNPs that disrupt *megf6a* function. Our previous work demonstrated this CRISPR approach completely eliminates function of the targeted gene in the injected embryo and generates F0 embryos that faithfully recapitulate loss-of-function mutants (26). We injected CRISPR-Cas9 RNPs targeting exons 6 and 28 of *megf6a* into one-cell stage embryos generated from an intercross between *megf6b*$^{+/-}$ adults, providing both control wildtype and *megf6b* mutant embryos that lacked *megf6a* function.

 Whereas three-quarters of the intercross embryos appeared phenotypically wildtype following CRIPSR-Cas9 mutagenesis of *megf6a* (Figure 4C), roughly one-quarter of the larvae (21.7%, (n

=115)) exhibited a severe disruption of the fin fold with aberrant migration of melanocytes into the fin fold area (Figure 4 D-F). The aberrant embryos were genotyped, and all were found to be homozygous for the *megf6b* mutation, indicating that mutagenesis of *megf6a* only has measurable effect on fin fold development in the absence of *megf6b* function (p = 0.0056). These results indicate *megf6a* and *megf6b* have redundant functions during embryogenesis: loss of either gene had minimal effect on development whereas larvae lacking both *megf6a* and *megf6b* gene functions had severe defects on fin development.

Cartilage and bone form in an anterioposterior wave of development in the zebrafish larva, so that at any single time point posterior structures are developmentally younger than anterior ones. Therefore, to examine cartilage and bone formation in the head and in the vertebral column, we analyzed mutant larvae at two developmentally distinct time points, 10 and 14 days post fertilization (dpf). To determine if *megf6a* and *megf6b* have a function in cartilage and bone development in the head, we analyzed 10 dpf larvae that developed from $megf6b^{+/-}$ intercross eggs injected with *megf6a* RNPs. Cartilage and bone formation were visualized after staining larvae with Alcian Blue (cartilage) and Alizarin Red (bone) (28). Injection of *megf6a* RNPs into WT or $megf6b^{+/-}$ larvae had no measurable effect on bone and cartilage formation (Figure 5A and 5A'). However, twenty percent of the injected intercross larvae (n = 110) exhibited consistent delays in jaw development, including formation of the pharyngeal arches and development of both bones (5th ceratobranchyal (5CB) or cleithrum (CL)) and cartilaginous structures of the jaw (Figure 5B and 5B'). Genotyping revealed that all the injected embryos with jaw defects were $megf6b^{-/-}$, indicating that loss of both *megf6* genes was required (p = 0.0062) to produce significant delays in jaw development.

9

Given that *megf6a* and *megf6b* have overlapping expression domains in the notochord (the embryonic version of the vertebral column) and that the two genes have overlapping functions in fin fold development (Figure 4D-F), we analyzed the roles of the genes in bone and cartilage formation in the vertebral column and caudal fin of 14 dpf larvae. Progeny of a $megf6b^{+/-}$ intercross mating were injected with *megf6a* RNPs. Most larvae exhibited normal bone and cartilage formation in the vertebral column and caudal fin (Figure 5C). However, 17.5% of the larvae (n=80) exhibited a severe delay in the ossification of the posterior vertebrae and a delay in the formation of the cartilaginous structures and bony rays of the caudal fin (Figure 5D-F). All the abnormal larvae proved to be $megf6b^{-/-}$ upon genotyping, indicating that only embryos lacking both *megf6b* and *megf6a* gene functions had an altered phenotype (p < 0.007). The functions of both genes are required for proper formation of both cartilaginous and bony structures in the zebrafish larvae.

To determine if zebrafish lacking *megf6a* and *megf6b* function have an adult phenotype, we separated 48 hpf larvae that developed from $megf6b^{+/-}$ intercross eggs injected with *megf6a* RNPs (Figure 4) into two groups of 25 animals, those with normal fin folds ($megf6b^{-/?}$) and those with fin fold defects ($megf6b^{-/-}$). Despite the developmental delay observed in doubly mutant larvae, 100% of the individuals of each group developed to adulthood. At 4 months of age, we measured the weight (g) and length (cm) of individual male and female adult zebrafish. The males in the abnormal fin fold group weighed less (mean ± SD - 0.4098g ± 0.014g) than the males with normal fin folds (0.4556g ± 0.014g), while there was no significant difference in length between the two groups (abnormal fin folds – 3.671cm ± 0.032cm vs normal fin folds –

3.721cm ± 0.035cm). In contrast to the males, females in the abnormal fin fold group weighed

more (abnormal fin folds - 0.6616g ± 0.021g vs normal fin folds 0.5332g and ± 0.020g) and

were longer (abnormal fin folds – 3.971cm ± 0.035cm vs normal fin folds – 3.809cm ±

0.034cm) than females with normal fin folds (Supplemental Figure 3).

11

DISCUSSION

Whole exome sequence data from pairs of related osteoporosis-affected individuals from 10 extended high-risk pedigrees were analyzed to identify a small set of rare, IBD-shared variants representing excellent candidates responsible for osteoporosis predisposition in the pedigrees studied. A rare *MEGF6* coding variant segregated with osteoporosis in the pedigree in which it was identified. *In silico* analyses of the predicted consequences of the amino acid substitution indicated the variant is likely to be damaging to normal protein function. STRING analysis indicated that MEGF6 may interact with proteins linked to osteoporosis. Loss-of-function analyses in the zebrafish revealed *megf6* gene function contributes to cartilage and bone formation. Together these support the interpretation that the very rare variant in *MEGF6* predisposes to osteoporosis.

The design of the current study focused on exome sequencing of pairs of related affected individuals selected from extended high-risk pedigrees. Selection of subjects focused on affected cousins, which has been shown to be ideal when selecting subjects to represent dense pedigrees (29). The main benefit of the affected-relative-pairs strategy is efficiency with respect to data capture, but at the potential cost of missing evidence of disease variants that may have been detectable in a design that sequences all pedigree members. Here the strategy was sufficiently robust to lead to identification of at least one variant in ten pedigrees. The reliance on SGS analysis to delineate regions of long-range sharing between target pair members provided a productive filter of potential candidate variants. The use of well-sampled high-risk pedigrees has the added advantage of being able to invoke a test of segregation among additional affected

12

family members to gain evidence of linkage to disease, which was shown here to be a particularly useful strategy for confirmation of the very rare variant rs755467862.

Consistent with our current finding of an association between *MEGF6* and osteoporosis, previous studies had observed an association between low BMD or osteoporosis and the chromosomal region containing the *MEGF6* locus (30, 31). A LOD score of +2.74 to 1p36.32 was reported in a resource of 1,270 subjects in 324 osteoporosis pedigrees (30). Independently, a LOD score of +3.07 to 1p36.3 was reported in a single Belgian pedigree with 34 subjects (31). In the latter study, it was reported that exon sequencing of *MEGF6* did not identify any coding variants among pedigree members.

In silico modeling of the WT and variant MEGF6 indicate that the Cys200Tyr may result in collapse of a loop that likely eliminates a calcium binding site in MEGF6, which may decrease, or even eliminate, the functionality of the protein. Similar arguments have been made by observation of mutations of EGF-like domains in other proteins. For instance, more than 60% of the mutations causing Marfan syndrome occur within fibrillin-1 cbEGF-like domains, emphasizing that correct folding of these domains is critical for molecular function (32). Mutations in cbEGF-like domains that affect cysteine residues are likely to alter disulfide bond formation, which may disrupt correct folding (32). Furthermore, mutations affecting residues in the calcium binding consensus sequence reduce calcium binding affinity, which may lead to structural destabilization (32).

13

The biological roles of MEGF6 are not well understood. MEGF6 appears to be a secreted calcium binding protein that regulates cell migration through TGF-β/SMAD mediated epithelial-mesenchymal transition (33). Depletion of *MEGF6* caused reduction of TGF-β1 expression and phosphorylation of SMAD2 and SMAD3, whereas *MEGF6* expression positively correlated with that of TGF-β1 expression, suggesting a role for *MEGF6* in TGF-β/SMAD signaling. In osteoporosis, TGF-β is proposed to regulate osteoclastogenesis through activation of osteoprotegerin (OPG) (34). OPG has been shown to inhibit the activity and survival of osteoclasts in vitro and bone resorption *in vivo* (23). It is possible that reduced *MEGF6* activity may alter TGF- β signaling and thus OPG expression, thereby contributing to osteoporosis. Furthermore, we identified another potential link between *MEGF6* and osteoporosis through a ZBTB40 -WNT16 interaction using the STRING database. Variants in *ZBTB40* and *WNT16* have been previously associated with osteoporosis (25, 35) and osteoblast-derived *WNT16* inhibits osteoclastogenesis to prevent bone fragility (36). Further studies are needed to determine if *MEGF6* regulates TGF- β signaling *in vivo* and if it interacts with *ZBTB40* or *WNT16*.

Our functional analyses demonstrate that the zebrafish orthologues of human *MEGF6*, *megf6a* and *megf6b*, are together required for proper development and mineralization of both cartilage and bone in the head, vertebrae, and caudal fin. In the head of 10 dpf larvae, we observed a delay in the formation of the cartilaginous pharyngeal arches and a delay in mineralization of the 5th ceratobranchyal (5CB) and the cleithrum (CL) bones, while there was no defect in the initial mineralization of the notochord (Figure 5A-B'). At 14 dpf there was a delay in the mineralization of the vertebral column and bones of the caudal fin, and a delay in the formation of cartilaginous structures of the caudal fin (Figure 5C-F). The mineralization of the 5CB, CL, and notochord

14

(vertebral column) represent distinct modes of bone ossification: endochondral, membranous, and perichordal bone ossification, respectively. Given there is a delay in mineralization of all of these structures, *megf6a* and *megf6b* have important functions in the distinct modes of bone ossification.

The bone phenotype observed in larvae lacking *megf6a* and *megf6b* function is strikingly similar to several previously reported zebrafish mutant phenotypes. A mutation in the *macrophage-stimulating protein* (*msp*) gene, or morpholino knockdown of its receptor, *ron-2*, display similar defects in the 5CB bone (37). The *msp* mutant phenotype can be rescued by exogenous calcium supplementation to the media, indicating that Ron-2 and Msp are necessary for bone mineralization and calcium homeostasis (37). The zebrafish *Chihuahua* (*Chi*) mutant disrupts *col1a1a* and is a model for classical osteogenesis imperfecta. Heterozygous *Chi* mutants have the same defects in the 5CB and CL bones and a delay in vertebral mineralization as are conspicuous in larvae with disrupted *mefg6a* and *megf6b* function (38). The zebrafish *frilly fins* (*frf*) mutant disrupts the *bmp1a* gene and is phenotypically similar to the *microwaved* mutant, which carries another mutant allele of *col1a1a* (39). BMP1 has a known function in the proteolytic processing of procollagen I C-propeptide to generate mature collagen type I (40), which may account for the phenotype similarities between the mutants. The *frf* mutant has a delay in the mineralization of the vertebrae reminiscent of larvae without *megf6a* and *megf6b* gene function (39). Given the similarity of these mutant phenotypes and with the perspective that *MEGF6* is a calcium binding protein, *MEGF6* may have a role in regulating calcium homeostasis and collagen processing during embryonic development.

15

A *Megf6*-deficient mouse was recently reported with no discernable phenotype (41). The apparent discrepancy between the mouse and zebrafish mutants may lay in the phenotypic analyses performed. There was no detailed histological analysis of embryonic skeletal development in the mouse, which may have overlooked subtle defects, and BMD was not analyzed in the adult. Skeletal phenotypes may be uncovered upon a more detailed analysis, since *Megf6* is expressed in the mouse musculoskeletal system during development (42), during osteoblast differentiation, and in mature bone (43). Furthermore, as the mouse mutation involved a coding alteration that resulted in no detectable *Megf6* transcript (41), the mutant transcript likely underwent non-sense mediated decay and may have triggered genetic compensation mechanisms that masked the null phenotype (44, 45).

Despite the dramatic developmental fin fold phenotype of *megf6b*$^{-/-}$ larvae injected with *megf6a* CRISPR/Cas9 reagents, adult zebrafish do not exhibit severe phenotypes (e.g., kyphosis), although there is a significant difference in the weight and length of these zebrafish. The adult males with abnormal fin folds weigh less than controls with normal fin folds without a reduction in overall length, which may reflect a loss in bone density. In contrast to the adult males, adult females with abnormal fin folds weigh more and are longer than controls with normal fin folds (Supplemental Figure 3). This phenotype may be due to factors unique to the teleost lineage, including sex differences and compensatory growth mechanisms. Zebrafish grow in size throughout their lifetime and have prodigious regenerative capacities (46, 47), which may contribute to the mild adult phenotype and the observed phenotypic differences between males and females. Further characterization of the zebrafish *megf6a;6b* mutant phenotype in aged

16

adults and generation of zebrafish with the human disease allele (48) will provide insight into the

genetic mechanisms of osteoporosis and may yield new targets for therapeutic intervention.

In sum, evidence of segregation, protein folding prediction, and *in vivo* functional evaluation

support a role for *MEGF6* in predisposition to development of osteoporosis.

MATERIALS AND METHODS

Utah osteoporosis pedigrees

Osteoporosis cases with a self-reported family history were ascertained and recruited from clinics associated with Intermountain Healthcare, the largest Utah health care provider, and local advertising in health care clinics; relatives of these cases were also recruited. All recruited individuals were offered a BMD scan for determination of osteoporosis phenotype. BMD was measured on a Hologic QDR 4500A fan-beam dual X-ray absorptiometry (DXA) bone densitometer by a single technician (49, 50). Quality control for the DXA was performed daily on a spine phantom and all scans were reviewed and evaluated by a single experienced physician. BMD was measured at the hip and spine (or at the wrist if the hip or spine could not be imaged). In addition to the anteroposterior (AP) view of the spine, lateral measurements were taken. Scans were excluded if their interpretation was confounded by poor positioning or abnormal anatomy.

A total of 1,871 individuals were consented, phenotyped, and sampled in 276 multi-generational pedigrees with 4–46 sampled individuals per pedigree. In addition to BMD measurements, a single physician reviewed a subject medical history questionnaire, and available medical records to assign a phenotype of osteoporosis, osteopenia, unaffected, or unknown. Ten pedigrees were selected for sequencing based on a physician diagnosis of osteoporosis in a pair of (approximately) cousins (29), and having additional sampled affected relatives for tests of segregation. This study was approved by the Institutional Review Boards of Intermountain Healthcare and the University of Utah.

18

Utah genomic data

Whole exome sequencing was performed on the 10 pairs of affected cousins at the Huntsman Cancer Institute's Genomics Core facility. DNA libraries were prepared from 3 micrograms of DNA using the Agilent SureSelect XT Human All Exon + UTR (v5) capture kit. Samples were run on the Illumina HiSeq 2000 instrument. Reads were mapped to the human genome GRCh37 reference using BWA-mem (51) for alignment and variants were called using Genome Analysis Toolkit (GATK) (52) software following Broad Institute Best Practices Guidelines. Exome capture resulted in an average of 85% of target bases being covered by greater than 10x coverage across the genome. Variants occurring outside the exon capture kit intended area of coverage were removed. Variants were annotated with Annovar, which contains predicted pathogenicity scores from 10 in-silico functional prediction algorithms (11). Samples were also genotyped on the Illumina OmniExpress high density SNP array (720,000 SNPs).

Candidate variant prioritization

Variants were filtered in four phases to achieve a small set of highly probable candidates. In the first phase, variants with population minor allele frequency (MAF) <0.005 that were observed in both index cases from a pedigree (affected cousins) were retained. In the second phase of variant filtration, the genomes of index pairs were restricted to regions exhibiting evidence of IBD transmission, or shared common ancestry. Focus on these IBD regions between sequenced cousin-pairs provides a useful filter that can rule out variants that are shared but were not inherited from a common ancestor and are inconsistent with a dominant mode of inheritance. SNP genotypes were used to estimate regions of long shared haplotypes between related

sequenced cases using SGS (8). SGS analysis identifies the set of contiguous markers at which genotyped individuals could share alleles; long runs of such markers indicate likely regions of IBD sharing from a common ancestor (53). Genomic regions were considered as IBD for a sequenced relative pair if the normalized shared-segment length was ≥20 standard deviations from the mean shared-segment length (54). This selection criteria was based on the expected amount of IBD sharing that should occur for a pair of cousins (~12.5%), and provided an equivalent reduction to the genomic search space. The third phase of variant filtration required that the gene had at least one published study connecting the gene with osteoporosis through some genetic study design, identified via PubMed search (55) with search terms 'osteoporosis' and gene name or alias provided by OMIM (Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), January 17, 2018: https://omim.org/). In the final filtering step, variants that were scored as 'damaging' by at least 8 of 10 in-silico pathogenicity prediction algorithms included in ANNOVAR variant annotations were retained.

Confirmation of candidate variants via segregation to affected relatives

The 10 pedigrees of the sequenced relative pairs had a total of 68 additional sampled relatives, ranging from 4 to 15 per family. To confirm segregation of the candidate predisposition variants in the pedigrees, each variant was assayed in the sampled relatives of sequenced cases using custom Taqman assays (Applied Biosystems, Carlsbad, CA, USA) according to the manufacturer's specifications on a Bio-Rad CFX96 Real-Time PCR instrument. Evidence of segregation to additional affected relatives was evaluated with RVsharing software (9) that

provides the probability of the observed configuration of sharing in a pedigree, expressed as an exact probability, assuming the shared variant is rare (<1%) and entered the pedigree only once.

Pathway analysis and protein structure prediction

The STRING database tool (21) with default settings was used to query a large number of known protein interactions concerning the candidate gene. Protein structure prediction software I-TASSER using full homology modeling was used to model the structure of critical domains assuming either wild-type or variant amino acid sequences (19, 56-58). Predicted structures were visualized and analyzed using Chimera software (59).

Zebrafish husbandry

Zebrafish (Danio rerio) were maintained in accordance with approved institutional protocols at the University of Utah. Adult zebrafish were maintained under standard conditions (60) and kept on a light-dark cycle of 14 hours in light and 10 hours in dark at 27°C. The Tu strain was used in all experiments.

megf6a and *megf6b* Mutant Generation

Mutations were generated with CRISPR-Cas9 reagents as described in (26). gRNA target sequences were as follows: *megf6a* exon 6 - CGCTGTCAGCATGGTGTTCT(TGG) and exon 28 - GCTCGCAGTGACCCTCTCAC(AGG); and *megf6b* exon 6 - AGCACACATGTGTGAACACT(AGG) and exon 35 - TCCTGTATCCGGATGACAGG(AGG). The PAM sequence is indicated in parentheses.

Target-specific Alt-R® crRNA and common Alt-R® tracrRNA were synthesized by IDT and dissolved in duplex buffer (IDT) as a 100μM stock solution. Equal volumes of the Alt-R® crRNA and Alt-R® tracrRNA stock solutions were mixed together and annealed in PCR machine using the following settings: 95°C, 5 min; cool at 0.1°C/sec to 25°C; 25°C, 5 min; 4°C. Cas9 protein (Alt-R® S.p. Cas9 nuclease, v.3, IDT, dissolved in 20mM HEPES-NaOH (pH 7.5), 350mM KCl, 20% glycerol) and crRNA:tracrRNA duplex mixed to generate a 5μM gRNA:Cas9 RNP complex (referred to as RNPs). Prior to microinjection, the RNP complex solution was incubated at 37°C, 5 min and then placed at room temperature. Approximately one nano-liter of 5μM RNP complex was injected into the cytoplasm of one-cell stage embryos.

Genomic DNA extraction and genotyping

Genomic DNA was extracted from individual embryos at 24 hours post fertilization (hpf). Dechorionated embryos were incubated in 50 ul 50 mM NaOH at 95°C, 20 min. 1/10 volume of 1 M Tris-HCl (pH 8.0) was added to neutralize. Genome sequences containing CRISPR/Cas9 target sites were amplified with pairs of primers: *megf6a* ex6 HRMA F1 – GGAGATTGTCAATACCTGTGACT and *megf6a* ex6 HRMA R1 – AGGTCTTCGGCGAGCTGATA; *megf6a* ex28 HRMA F1 – GGTCAGGACTGTGCTGGAGTG and *megf6a* ex28 HRMA R1 – CCCACAGTCCCGTCCACGC; *megf6b* ex6 HRMA F1 – ATATTGACGAGTGCCAGGTTCAT and *megf6b* ex6 HRMA R1 – ATGCAGTCGAGAGCCGGCGTTA; *megf6b* ex35 HRMA F1 – GCTCTGGGTGTCAGCAGCAGT and *megf6b* ex35 HRMA R1 – GCTGACAGCGCGGGCCGT. To determine if individual gRNA:Cas9 RNPs produced

mutations at the desired target sites, high resolution melt analysis (HRMA) was performed on DNA isolated from 8 individual 24 hpf gRNA:Cas9 RNP-injected embryos using KAPA HRM FAST PCR Master Mix (61). To generate the *megf6b* deletion allele, a mixture of gRNA:Cas9 RNPs targeting exons 6 and 35 was injected into the cytoplasm of one-cell stage embryos. To detect deletion events, PCR was performed with *megf6b* ex6 HRMA F1 and *megf6b* ex35 HRMA R1 primers on DNA isolated from 8 individual 24 hpf G0 gRNA:Cas9 RNP injected embryos using KAPA 2G FAST PCR Master Mix. To remove *megf6a* gene function in G0 embryos, a mixture of gRNA:Cas9 RNPs targeting exons 6 and 28 was injected into the cytoplasm of one-cell stage embryos.

In Situ Hybridization

Whole mount in situ hybridization (WISH) was performed on embryos as described in (62). cDNA fragments used to generate riboprobes probes were amplified by RT-PCR using the following primers: *megf6a* ISH F – TGGCACCTGCAGCTGCCC, and *megf6a* ISH R – TCCAGCCGTTCAGACACGTGCA, or *megf6b* ISH F – TGAACAGACGTGTCCGCAGGG, and *megf6b* ISH R – TCACACTCGCACAGCAGAGAGC using KAPA 2G FAST Master Mix, sub-cloned into pGEM-T Easy, and subjected to Sanger sequencing for verification.

Cartilage and Bone Staining

Ten and 14 dpf zebrafish larvae were anesthetized with Tricaine methanesulfonate (ethyl 3-aminobenzoate methanesulfonate) and processed as described in (28) and https://wiki.zfin.org/pages/viewpage.action?pageId=13107375 with the following modifications.

23

Larvae were fixed in 2% paraformaldehyde for 1 hour, washed for 10 minutes in 50% EtOH, and then transferred to a solution containing 0.01% Alizarin Red and 0.04% Alcian Blue for 24 hours. Larvae were washed in 80% EtOH/10mM MgCl2 for 60 minutes, 50% EtOH for 30 minutes, 25% EtOH for 30 minutes, bleached in 3% H_2O_2/0.5% KOH for 15 minutes, washed in 2X 25% glycerol/0.1% KOH and then transferred to 50% glycerol/0.1% KOH for imaging.

Length and Weight Measurements

One-cell stage embryos from an intercross between $megf6b^{+/-}$ adults were injected with $megf6a$ RNPs and larvae were sorted at 48 hpf into two groups of 25 animals, those with normal fin folds ($megf6b^{-/?}$) and those with fin fold defects ($megf6b^{-/-}$). Animals were raised to 4 months of age. At 4 months of age adults were anesthetized using Tricaine methanesulfonate and the length (cm) of each zebrafish was determined by measuring from the anterior most portion of the head to the tip of the tail. Zebrafish were then weighed and allowed to recover from anesthesia.

Statistical analysis

One-cell stage embryos from an intercross between $megf6b^{+/-}$ adults were injected with $megf6a$ RNPs and larvae were sorted at 48 hpf based on whether or not they displayed the mutant fin fold phenotype. Embryos were then genotyped as describe above. For the cartilage and bone analysis, embryos were stained, sorted based on phenotype, and then genotyped. To determine whether phenotypically defective larvae were significantly enriched for the $megf6b^{-/-}$ genotype, Fisher's exact test with Bonferroni correction was used to determine statistical significance. A

Student's t-test was used to determine statistical significance for the weight and length comparisons.

CONFLICT OF INTEREST STATEMENT
No conflicts of interest are reported by the authors.

REFERENCES

1   Rocha-Braz, M.G. and Ferraz-de-Souza, B. (2016) Genetics of osteoporosis: searching for candidate genes for bone fragility. *Archives of endocrinology and metabolism*, **60**, 391-401.

2   Trajanoska, K. and Rivadeneira, F. (2019) The genetic architecture of osteoporosis and fracture risk. *Bone*, **126**, 2-10.

3   Morris, J.A., Kemp, J.P., Youlten, S.E., Laurent, L., Logan, J.G., Chai, R.C., Vulpescu, N.A., Forgetta, V., Kleinman, A., Mohanty, S.T. *et al.* (2019) An atlas of genetic influences on osteoporosis in humans and mice. *Nature genetics*, **51**, 258-266.

4   Duncan, E.L., Cardon, L.R., Sinsheimer, J.S., Wass, J.A. and Brown, M.A. (2003) Site and gender specificity of inheritance of bone mineral density. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research*, **18**, 1531-1538.

5   Francois, S., Benmalek, A., Guaydier-Souquieres, G., Sabatier, J.P. and Marcelli, C. (1999) Heritability of bone mineral density. *Revue du rhumatisme*, **66**, 146-151.

6   Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A. *et al.* (2009) Finding the missing heritability of complex diseases. *Nature*, **461**, 747-753.

7   Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B. *et al.* (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, **536**, 285-291.

8   Thomas, A., Camp, N.J., Farnham, J.M., Allen-Brady, K. and Cannon-Albright, L.A. (2008) Shared genomic segment analysis. Mapping disease predisposition genes in extended pedigrees using SNP genotype assays. *Annals of human genetics*, **72**, 279-287.

9   Bureau, A., Younkin, S.G., Parker, M.M., Bailey-Wilson, J.E., Marazita, M.L., Murray, J.C., Mangold, E., Albacha-Hejazi, H., Beaty, T.H. and Ruczinski, I. (2014) Inferring rare disease risk variants based on exact probabilities of sharing by multiple affected relatives. *Bioinformatics*, **30**, 2189-2196.

10   Genomes Project, C., Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T. and McVean, G.A. (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature*, **491**, 56-65.

11   Wang, K., Li, M. and Hakonarson, H. (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research*, **38**, e164.

12   Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S. and Sunyaev, S.R. (2010) A method and server for predicting damaging missense mutations. *Nature methods*, **7**, 248-249.

13   Liu, X., Jian, X. and Boerwinkle, E. (2013) dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. *Human mutation*, **34**, E2393-2402.

14   Schwarz, J.M., Cooper, D.N., Schuelke, M. and Seelow, D. (2014) MutationTaster2: mutation prediction for the deep-sequencing age. *Nature methods*, **11**, 361-362.

15   Shihab, H.A., Rogers, M.F., Gough, J., Mort, M., Cooper, D.N., Day, I.N., Gaunt, T.R. and Campbell, C. (2015) An integrative approach to predicting the functional effects of non-coding and coding sequence variation. *Bioinformatics*, **31**, 1536-1543.

16   Shihab, H.A., Gough, J., Cooper, D.N., Stenson, P.D., Barker, G.L., Edwards, K.J., Day, I.N. and Gaunt, T.R. (2013) Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Human mutation*, **34**, 57-65.

17   Ng, P.C. and Henikoff, S. (2003) SIFT: Predicting amino acid changes that affect protein function. *Nucleic acids research*, **31**, 3812-3814.

18   Davydov, E.V., Goode, D.L., Sirota, M., Cooper, G.M., Sidow, A. and Batzoglou, S. (2010) Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS computational biology*, **6**, e1001025.

19   Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. and Zhang, Y. (2015) The I-TASSER Suite: protein structure and function prediction. *Nature methods*, **12**, 7-8.

20   Rao, Z., Handford, P., Mayhew, M., Knott, V., Brownlee, G.G. and Stuart, D. (1995) The structure of a Ca(2+)-binding epidermal growth factor-like domain: its role in protein-protein interactions. *Cell*, **82**, 131-141.

21   Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P. *et al.* (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic acids research*, **45**, D362-D368.

22   Munjal, C., Opoka, A.M., Osinska, H., James, J.F., Bressan, G.M. and Hinton, R.B. (2014) TGF-beta mediates early angiogenesis and latent fibrosis in an Emilin1-deficient mouse model of aortic valve disease. *Disease models & mechanisms*, **7**, 987-996.

23   Boyle, W.J., Simonet, W.S. and Lacey, D.L. (2003) Osteoclast differentiation and activation. *Nature*, **423**, 337-342.

24   Chao, T.H., Yu, H.N., Huang, C.C., Liu, W.S. and Lu, K.H. (2012) Opposite associations of osteoprotegerin and ZBTB40 polymorphisms with bone mineral density of the hip in postmenopausal Taiwanese women. *Journal of the Chinese Medical Association : JCMA*, **75**, 335-340.

25   Garcia-Ibarbia, C., Delgado-Calle, J., Casafont, I., Velasco, J., Arozamena, J., Perez-Nunez, M.I., Alonso, M.A., Berciano, M.T., Ortiz, F., Perez-Castrillon, J.L. *et al.* (2013) Contribution of genetic and epigenetic mechanisms to Wnt pathway activity in prevalent skeletal disorders. *Gene*, **532**, 165-172.

26   Hoshijima, K., Jurynec, M.J., Klatt Shaw, D., Jacobi, A.M., Behlke, M.A. and Grunwald, D.J. (2019) Highly Efficient CRISPR-Cas9-Based Methods for Generating Deletion Mutations and F0 Embryos that Lack Gene Function in Zebrafish. *Developmental cell*, **51**, 645-657 e644.

27   Prince, V.E. and Pickett, F.B. (2002) Splitting pairs: the diverging fates of duplicated genes. *Nature reviews. Genetics*, **3**, 827-837.

28   Walker, M.B. and Kimmel, C.B. (2007) A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotechnic & histochemistry : official publication of the Biological Stain Commission*, **82**, 23-28.

29   Feng, B.J., Tavtigian, S.V., Southey, M.C. and Goldgar, D.E. (2011) Design considerations for massively parallel sequencing studies of complex human disease. *PloS one*, **6**, e23221.

30   Karasik, D., Myers, R.H., Hannan, M.T., Gagnon, D., McLean, R.R., Cupples, L.A. and Kiel, D.P. (2002) Mapping of quantitative ultrasound of the calcaneus bone to chromosome 1 by genome-wide linkage analysis. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA*, **13**, 796-802.

31   Willaert, A., Van Pottelbergh, I., Zmierczak, H., Goemaere, S., Kaufman, J.M., De Paepe, A. and Coucke, P. (2008) A genome-wide linkage scan for low spinal bone mineral density in a single extended family confirms linkage to 1p36.3. *European journal of human genetics : EJHG*, **16**, 970-976.

32   Kielty, C.M., Sherratt, M.J., Marson, A. and Baldock, C. (2005) Fibrillin microfibrils. *Advances in protein chemistry*, **70**, 405-436.

33   Hu, H., Wang, M., Wang, H., Liu, Z., Guan, X., Yang, R., Huang, R., Tang, Q., Zou, C., Wang, G. *et al.* (2018) MEGF6 Promotes the Epithelial-to-Mesenchymal Transition via the TGFbeta/SMAD Signaling Pathway in Colorectal Cancer Metastasis. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology*, **46**, 1895-1906.

34   Thirunavukkarasu, K., Miles, R.R., Halladay, D.L., Yang, X., Galvin, R.J., Chandrasekhar, S., Martin, T.J. and Onyia, J.E. (2001) Stimulation of osteoprotegerin (OPG) gene expression by transforming growth factor-beta (TGF-beta). Mapping of the OPG promoter region that mediates TGF-beta effects. *The Journal of biological chemistry*, **276**, 36241-36250.

35   Styrkarsdottir, U., Halldorsson, B.V., Gretarsdottir, S., Gudbjartsson, D.F., Walters, G.B., Ingvarsson, T., Jonsdottir, T., Saemundsdottir, J., Center, J.R., Nguyen, T.V. *et al.* (2008) Multiple genetic loci for bone mineral density and fractures. *The New England journal of medicine*, **358**, 2355-2365.

36   Moverare-Skrtic, S., Henning, P., Liu, X., Nagano, K., Saito, H., Borjesson, A.E., Sjogren, K., Windahl, S.H., Farman, H., Kindlund, B. *et al.* (2014) Osteoblast-derived WNT16 represses osteoclastogenesis and prevents cortical bone fragility fractures. *Nature medicine*, **20**, 1279-1288.

37   Huitema, L.F., Renn, J., Logister, I., Gray, J.K., Waltz, S.E., Flik, G. and Schulte-Merker, S. (2012) Macrophage-stimulating protein and calcium homeostasis in zebrafish. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, **26**, 4092-4101.

38   Gioia, R., Tonelli, F., Ceppi, I., Biggiogera, M., Leikin, S., Fisher, S., Tenedini, E., Yorgan, T.A., Schinke, T., Tian, K. *et al.* (2017) The chaperone activity of 4PBA ameliorates the skeletal phenotype of Chihuahua, a zebrafish model for dominant osteogenesis imperfecta. *Human molecular genetics*, **26**, 2897-2911.

39   Asharani, P.V., Keupp, K., Semler, O., Wang, W., Li, Y., Thiele, H., Yigit, G., Pohl, E., Becker, J., Frommolt, P. *et al.* (2012) Attenuated BMP1 function compromises osteogenesis, leading to bone fragility in humans and zebrafish. *American journal of human genetics*, **90**, 661-674.

40   Ge, G. and Greenspan, D.S. (2006) Developmental roles of the BMP1/TLD metalloproteinases. *Birth defects research. Part C, Embryo today : reviews*, **78**, 47-68.

41   Wang, Y., Song, H., Wang, W. and Zhang, Z. (2019) Generation and characterization of Megf6 null and Cre knock-in alleles. *Genesis*, **57**, e23262.

42   Diez-Roux, G., Banfi, S., Sultan, M., Geffers, L., Anand, S., Rozado, D., Magen, A., Canidio, E., Pagani, M., Peluso, I. *et al.* (2011) A high-resolution anatomical atlas of the transcriptome in the mouse embryo. *PLoS biology*, **9**, e1000582.

43   Chim, S.M., Qin, A., Tickner, J., Pavlos, N., Davey, T., Wang, H., Guo, Y., Zheng, M.H. and Xu, J. (2011) EGFL6 promotes endothelial cell migration and angiogenesis through the activation of extracellular signal-regulated kinase. *The Journal of biological chemistry*, **286**, 22035-22046.

44   El-Brolosy, M.A., Kontarakis, Z., Rossi, A., Kuenne, C., Gunther, S., Fukuda, N., Kikhi, K., Boezio, G.L.M., Takacs, C.M., Lai, S.L. *et al.* (2019) Genetic compensation triggered by mutant mRNA degradation. *Nature*, **568**, 193-197.

45   Ma, Z., Zhu, P., Shi, H., Guo, L., Zhang, Q., Chen, Y., Chen, S., Zhang, Z., Peng, J. and Chen, J. (2019) PTC-bearing mRNA elicits a genetic compensation response via Upf3a and COMPASS components. *Nature*, **568**, 259-263.

46   Gemberling, M., Bailey, T.J., Hyde, D.R. and Poss, K.D. (2013) The zebrafish as a model for complex tissue regeneration. *Trends in genetics : TIG*, **29**, 611-620.

47   Marques, I.J., Lupi, E. and Mercader, N. (2019) Model systems for regeneration: zebrafish. *Development*, **146**.

48   Hoshijima, K., Jurynec, M.J. and Grunwald, D.J. (2016) Precise Editing of the Zebrafish Genome Made Simple and Efficient. *Developmental cell*, **36**, 654-667.

49   Bansal, A., Hughes, D.C., Farnham, J.M., Bagi, C.M., O'Neil, G., Rowe, K., Shakib, J.H., Wood, G.C., Wyckoff, J.A. and Cannon-Albright, L.A. (2000) Impact of correlated factors on bone density in individuals with a family history of osteoporosis. *Journal of clinical densitometry : the official journal of the International Society for Clinical Densitometry*, **3**, 333-338.

50   Kanis, J.A., McCloskey, E.V., Johansson, H., Oden, A., Melton, L.J., 3rd and Khaltaev, N. (2008) A reference standard for the description of osteoporosis. *Bone*, **42**, 467-475.

51   Li, H. and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, **26**, 589-595.

52   McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. *et al.* (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, **20**, 1297-1303.

53   Thomas, A., Skolnick, M.H. and Lewis, C.M. (1994) Genomic mismatch scanning in pedigrees. *IMA journal of mathematics applied in medicine and biology*, **11**, 1-16.

54   Huff, C.D., Witherspoon, D.J., Simonson, T.S., Xing, J., Watkins, W.S., Zhang, Y., Tuohy, T.M., Neklason, D.W., Burt, R.W., Guthery, S.L. *et al.* (2011) Maximum-likelihood estimation of recent shared ancestry (ERSA). *Genome research*, **21**, 768-774.

55   Coordinators, N.R. (2017) Database Resources of the National Center for Biotechnology Information. *Nucleic acids research*, **45**, D12-D17.

56   Roy, A., Kucukural, A. and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nature protocols*, **5**, 725-738.

57   Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC bioinformatics*, **9**, 40.

58   Zhang, Y. (2009) I-TASSER: fully automated protein structure prediction in CASP8. *Proteins*, **77 Suppl 9**, 100-113.

59   Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry*, **25**, 1605-1612.

60   Westerfield, M. (2000) *The Zebrafish Book*. Univ. of Oregon Press, Eugene.

61   Dahlem, T.J., Hoshijima, K., Jurynec, M.J., Gunther, D., Starker, C.G., Locke, A.S., Weis, A.M., Voytas, D.F. and Grunwald, D.J. (2012) Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS genetics*, **8**, e1002861.

62   Jurynec, M.J., Xia, R., Mackrill, J.J., Gunther, D., Crawford, T., Flanigan, K.M., Abramson, J.J., Howard, M.T. and Grunwald, D.J. (2008) Selenoprotein N is required for ryanodine receptor calcium release channel activity in human and zebrafish muscle. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 12485-12490.

FIGURE LEGENDS

Figure 1. Osteoporosis pedigree segregating *MEGF6* variant. Physician confirmed diagnosis of osteoporosis is denoted by black fill and physician confirmed diagnosis of osteopenia is denoted by half-shading. Arrows indicate index sequenced subjects, '+' indicates confirmed carriage of the variant, 'F' indicates multiple fractures in the patient's medical history, and 'T' indicates previous treatment for osteoporosis.

Figure 2. Comparison of the structure of *MEGF6* wild type (tan) and variant (blue) in the region of the Cys200Tyr. The disulfide bond between residues 187 and 200 have been highlighted in the wild type and mutant structures in red and yellow, respectively. The elimination of a loop can be

visually observed, which is likely to adversely affect the calcium binding function of the first cbEGF-like domain in *MEGF6*.

Figure 3. Protein interaction network from STRING database centered on *MEGF6* (A) and *ZBTB40* (B). Red nodes indicate the query protein and first shell of interactors, white nodes indicate second shell of interactors. Empty nodes indicate proteins of unknown 3-dimensional structure and filled nodes indicate 3-dimensional structure is known or predicted. Teal connectors indicate known interactions from curated databases, purple connectors indicate experimentally determined interactions, black connectors indicate co-expression, and yellow connectors indicate that interactions originate from text-mining.

Figure 4. Larvae lacking *megf6a* and *megf6b* gene function have fin fold defects at 48 hpf. WT or *megf6b$^{-/+}$* larvae have no overt phenotypes (A and A'), while *megf6b$^{-/-}$* larvae have a few aberrant melanocytes (B', arrow) present in the caudal fin area (B and B'). A' and B' are high magnification views of dashed boxes in A and B. (C-F) Injection of *megf6a* RNPs into larvae from an intercross between *megf6b$^{+/-}$* adults. Loss of *megf6a* gene function has no effect in a WT or *megf6b$^{+/-}$* background (C), while loss of *megf6a* gene function in *megf6b$^{-/-}$* larvae disrupted fin fold development and led to aberrant migration of melanocytes into the caudal fin area (D-F). All abnormal larvae were genotyped and determined to be *megf6b$^{-/-}$*, indicating complete loss of *megf6b* function is required for the phenotype observed in the *megf6a* RNP-injected larvae (p < 0.0056, Fisher's exact test with Bonferroni correction). Panels D-F depict the spectrum of aberrant phenotypes observed. Lateral views with anterior to the left.

Figure 5. Cartilage and bone formation is delayed in larvae lacking both *megf6a* and *megf6b* gene function. Progeny of an *megf6b$^{-/+}$* intercross were injected with *megf6a* RNPs. (A-B') Cartilage and bone formation in the heads of RNP-injected WT, *megf6b$^{+/?}$* (A and A'), or
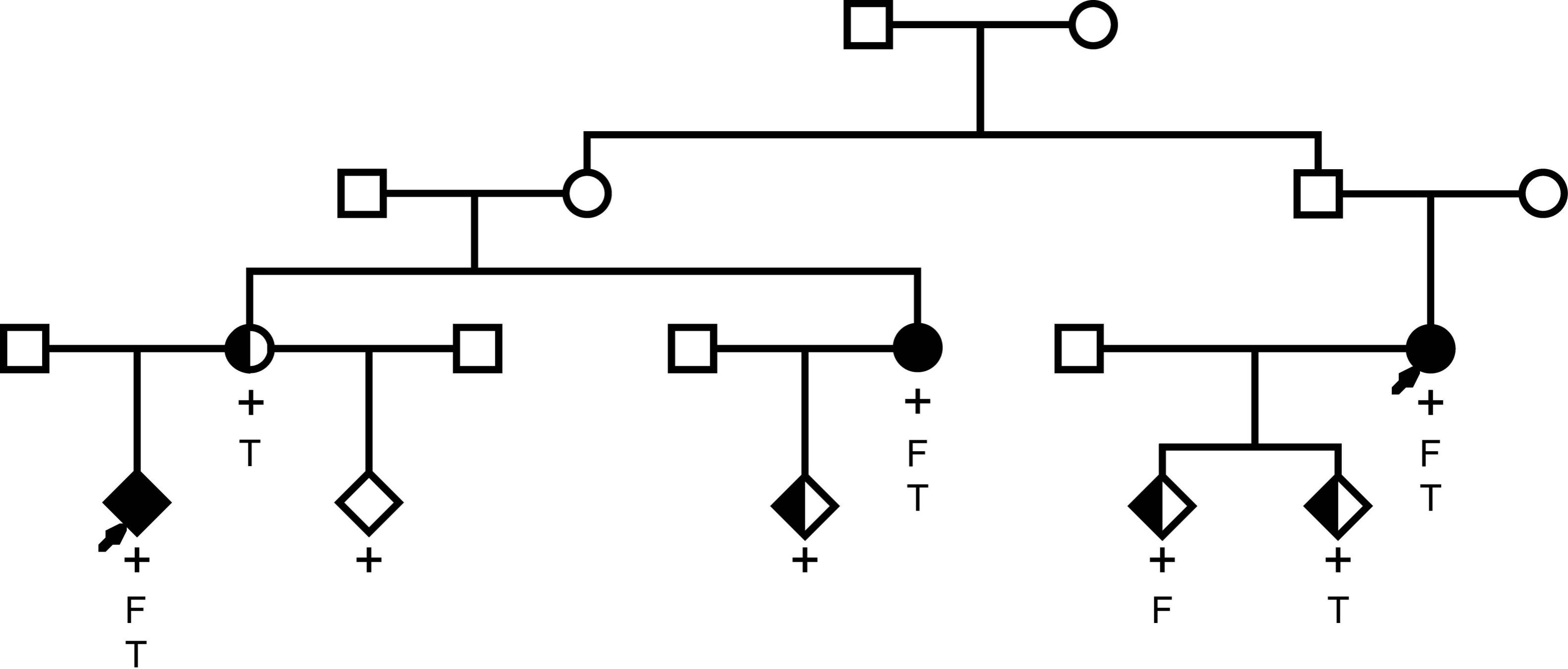
32

megf6b$^{-/-}$ (B and B') larvae were analyzed at 10 dpf. Arrowhead indicates the 5th ceratobranchyal (5CB) and arrow indicates the cleithrum (CL). Twenty-two of 110 larvae exhibited delayed formation of jaw bones and pharyngeal arches, and all of these were *megf6b*$^{-/-}$, indicating loss of both genes is required for the developmental defect ($p < 0.0062$, Fisher's exact test with Bonferroni correction). A and B, lateral view and A' and B', ventral view. Asterisk indicates the notochord. (C-F) Cartilage and bone formation in in the tails of RNP-injected WT, *megf6b*$^{+/?}$, or *megf6*$^{-/-}$ larvae were analyzed at14 dpf. Mineralization in the vertebra and formation of cartilaginous structures of the caudal fin are significantly delayed in larvae lacking *megf6a* and *megf6b* function (17.5%, n=80) (D-F) compared to WT or *megf6b*$^{+/?}$ (C) larvae injected with *megf6a* RNPs. All abnormal larvae were genotyped and determined to be *megf6b*$^{-/-}$ indicating loss of both genes is required for the delayed phenotype ($p < 0.007$, Fisher's exact test with Bonferroni correction). Panels D-F depict the spectrum of phenotypes observed when *megf6a* RNPs are injected into *megf6b*$^{-/-}$ embryos. C-F, lateral views with anterior to the left.
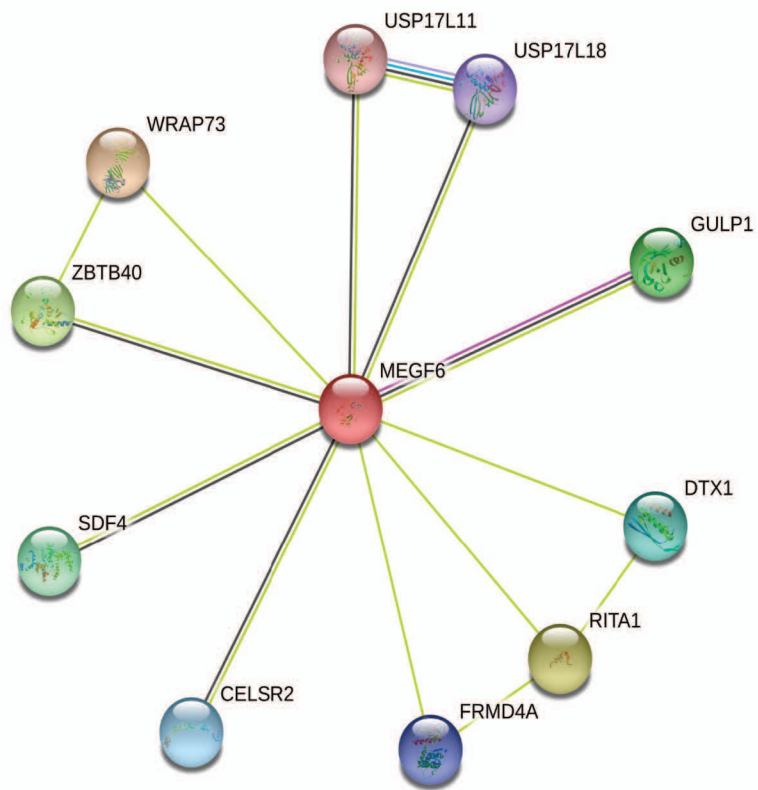
Supplemental Figure 1. *megf6a* and *megf6b* have both overlapping and unique expression domains. Whole mount *in situ* hybridization of *megf6a* (A-D) and *megf6b* (E-G) at 24 hpf (A and E), 36 hpf (B, C, D, and F), and 48 hpf (G). *megf6a* and *megf6b* are expressed in the fin fold (arrow in A) and in the head of 24 hpf embryos (A and E), while only *megf6b* is expressed in somitic tissue (arrow in E). At 36 hpf *megf6a* and *megf6b* are expressed in the fin fold (D and F), AER (arrowhead in B and F), atrioventricular canal (arrowhead in C), and notochord (arrow in D and F), and head. *megf6b* is expressed in the mouth (arrow in G) and pharyngeal region (arrowheads in G) in 48 hpf larvae. Lateral views with anterior to the left.

Supplemental Figure 2. Generation of a heritable *megf6b* deletion mutant. (A) RNPs were designed to target sites (lightning bolts) within exon 6 (ex 6 gRNA) and exon 35 (ex 35 gRNA)

of *megf6b*. Primer pairs used to amplify junction fragments joining target sites are indicated (red arrows). (B) Detailed view of the ex 6 gRNA and ex 35 gRNA sites with gRNA target sequence in red, PAM site in blue, and Cas9 cleavage site indicated by lightning bolts. *megf6bΔ ex 6/35* represents a conceptual precise joining of sequences flanking the presumed cleavage sites in target sites ex 6 and ex 35. *megf6b^{z48}* represents the deletion allele recovered in the germline. This allele contains a 51,860 bp deletion with a 53 bp insertion.

Supplemental Figure 3. Weight (g) and length (cm) of adult zebrafish lacking *megf6a* and *megf6b* function. The weight (A) and length (B) of individual male and female 4 month old zebrafish generated from *megf6b$^{+/-}$* intercross eggs injected with *megf6a* RNPs. Circles indicate embryos identified at 48 hpf with normal fin folds (*megf6b$^{+/?}$*) and squares indicate those with fin fold defects (*megf6b$^{-/-}$*). Males with abnormal fin folds weighed less than controls with normal fin folds (A), while there was no difference in length between the two groups (B). Females with abnormal fin folds weighed more and were longer when compared to controls with normal fin folds (A and B). n = 14 males and n = 11 females with normal fin folds and n= 14 males and n = 7 females with abnormal fin folds. A Student's t-test was used to determine statistical significance. Data are represented as means ± SD. *p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001.

A, A' WT

B, B' *megf6b*[-/-]

C *megf6b*[+/?]; *megf6a* RNPs

D *megf6b*[-/-]; *megf6a* RNPs

E *megf6b*[-/-]; *megf6a* RNPs

F *megf6b*[-/-]; *megf6a* RNPs

megf6b$^{+/?}$; megf6a RNPs

megf6b$^{-/-}$; megf6a RNPs

megf6b$^{+/?}$; megf6a RNPs

megf6b$^{-/-}$; megf6a RNPs

megf6b$^{-/-}$; megf6a RNPs

megf6b$^{-/-}$; megf6a RNPs