Bone Research (2013) 4: 301-310.
www.boneresearch.org

REVIEW

Contemporary Approaches for Identifying Rare Bone Disease Causing Genes

Charles R. Farber[1,2], Thomas L. Clemens[3,4]

[1]Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA; [2]Departments of Public Health Sciences and Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA, USA; [3]Department of Orthopaedic Surgery, Johns Hopkins School of Medicine, Baltimore, MD, USA; [4]Baltimore Veterans Administration Medical Center, Baltimore, Maryland, USA

Recent improvements in the speed and accuracy of DNA sequencing, together with increasingly sophisticated mathematical approaches for annotating gene networks, have revolutionized the field of human genetics and made these once time consuming approaches assessable to most investigators. In the field of bone research, a particularly active area of gene discovery has occurred in patients with rare bone disorders such as osteogenesis imperfecta (OI) that are caused by mutations in single genes. In this perspective, we highlight some of these technological advances and describe how they have been used to identify the genetic determinants underlying two previously unexplained cases of OI. The widespread availability of advanced methods for DNA sequencing and bioinformatics analysis can be expected to greatly facilitate identification of novel gene networks that normally function to control bone formation and maintenance.

***Keywords*:** exome sequencing; genome sequencing; bioinformatics; osteogenesis imperfect; pigment epithelium-derived factor (PEDF); bone restricted ifitm-like protein (Bril)

Bone Research (2013) **4**: 301-310. doi: 10.4248/BR201304001

Introduction

Our current understanding of many fundamental mechanisms underlying mammalian physiology has been profoundly influenced by discoveries in human genetics. Pioneers in human genetics such as Victor McKusick at Johns Hopkins and Charles Dent at University College in London spent their entire careers developing theoretical and technological methods for tracking the transmission of inheritable traits and linking them to gene defects impacting disease. In turn, these discoveries informed on how normal pathways controlled complex physiological processes. Over the last few decades, exponential improvements in the speed and accuracy of DNA sequencing, coupled with increasingly sophisticated mathematical approaches for annotating gene networks, have revolutionized the field of human genetics and made these once time consuming approaches assessable to most investigators. Indeed, at the time we are writing this article, high quality DNA sequences of entire genomes can be obtained commercially in a just a few days for ~$3 500 US dollars (http://www.genome.gov). Consequently, disease causing gene defects can be identified in a matter of a few weeks by most biomedical researchers provided that they can access appropriate bioinformatics expertise. Moreover, parallel advances in pathway analysis algorithms, together with open access to both chemical and pharmaceutical libraries, have for the first time enabled informed therapeutic targeting of disease causing gene pathways. In this perspective, we discuss some of these technological advances and describe how they have enabled the identification of the molecular defects underlying two rare bone diseases.

*Correspondence: Charles R. Farber ; Thomas L. Clemens
E-mail: crf2s@eservices.virginia.edu; tclemen5@jhmi.edu
Received 12 September 2013; Accepted 20 October 2013

NextGen genome sequencing

Few advances have impacted biology more than inno-

vations in DNA sequencing. In 2001, the cost of a human genome sequence was ~$100 million and took years to complete. Today, a human genome can be sequenced for ~$3 500 in a matter of days (http://www.genome.gov). The remarkable acceleration in our ability to sequence DNA has made genome/exome (the 1.5% of the genome consisting of protein-coding exons) sequencing routine and significantly increased our ability to identify variants responsible for Mendelian diseases. This is especially evident in the bone field where over the last three years Next-Generation Sequencing (NGS) has been used to identify new genes for osteogenesis imperfecta (OI) [as examples (1-5)], early-onset osteoporosis (6-7) and other skeletal dysplasias (8-10).

In addition to generating a reference genome sequence, the Human Genome Project created immense interest and investment in the creation of NGS platforms to sequence DNA faster and at a lower cost (11). NGS platforms have enabled the transition from sequencing DNA fragments one at a time to sequencing fragments in a massively parallel manner. To illustrate this idea it is useful to provide an overview of Illumina sequencing by synthesis, one of the most widely used NGS platforms (12). Illumina sequencing works by generating hundreds of millions of clonal DNA populations attached to a solid surface. Fluorescently labeled nucleotides are then incorporated one base at a time to the clonal fragments within a cluster. The addition of a particular base to a subset of clusters is recorded by high-resolution imaging. A different base is then added and imaged. This is repeated for multiple cycles until "reads" of 100-200 bases of sequence are generated for each of the millions of fragment clusters. Currently, the Illumina HiSeq2500 sequencer can generate 600 Gigabase pairs (Gbp) of sequence (200 human genome equivalents) every 11 days. There are also a number of other commercial or soon to be commercial NGS platforms. Examples include Ion Torrent semiconductor sequencing (13), Pacific Biosciences' single molecule real time (SMRT) sequencing (14) and Nanopore-based DNA sequencing by Oxford Nanpore Technologies (15). While these technologies are less mature than Illumina, they each have the potential to result in improvements in sequence quality, throughput and cost.

Disease gene identification and prioritization using exome sequencing

Mutations underlying single-gene disorders have traditionally been identified using candidate gene screening or linkage mapping/positional cloning strategies in pedigrees (16). However, these approaches require prior biological information regarding the disease or large families. In contrast, whole-genome or exome sequencing is capable of discovering disease genes in an unbiased manner (17). For single-gene disorders, exome sequencing is more efficient than whole genome sequencing as it focuses on regions where variants are the most likely to have functional significance (1.5% of the genome), thereby permitting the high coverage (>60X) that is needed to confidently identify variants. Also, the majority (>85%) of single-gene disorders are due to coding mutations (18). In an exome sequencing experiment, the coding regions of a genome are "captured" by hybridizing fragmented genomic DNA to a library of exonic DNA oligos. The captured sequences are then sequenced using a NGS platform (17).

An extremely active area of research in bioinformatics has focused on the processing and analysis of NGS sequence data (19). This has lead to a proliferation in the number of commercial and open-access software for sequence data analysis. A recent review surveyed 205 tools for various aspects of whole genome/exome sequence analysis, which gives a sense of the intense interest in this area (20). Here, we hope to provide a general overview of a typical exome sequence analysis pipeline (Figure 1).

As discussed above, DNA from individuals to be sequenced is subjected to exome capture. The samples are then sequenced, with a typical yield in the range of 50-90 million reads (100 bp) or 5-9 Gbp of sequence per sample. The first step in the analysis is to exclude low quality and duplicate reads. The next step is aligning these raw sequence reads to the human reference. In principle aligning reads is straightforward, but in reality there are a number of parameters to consider such as the number of mismatches to allow and whether to include reads that map equally well to multiple sites in the human genome. After excluding low quality reads and reads not mapping to the exome, a typical yield is 3-6 Gbp of sequence or 50-100 X coverage of the 50 megabase pairs (Mbp) exome. The end result is a list of sequence reads, their mapping locations and a quality score associated with the quality of the alignment.

The next step is to take the list of alignments and identify variants. In this step, variant positions in the genome, relative to a reference, are identified by comparing the "pileup" of base calls for every position in the exome. Homozygous variants are those in which all reads mapping to a position possess the non-reference allele, whereas heterozygous variants have the alternative allele in ~50% of reads. Most variants are single-nucleotide changes, although commonly used software tools also have the capacity to identify short insertions and

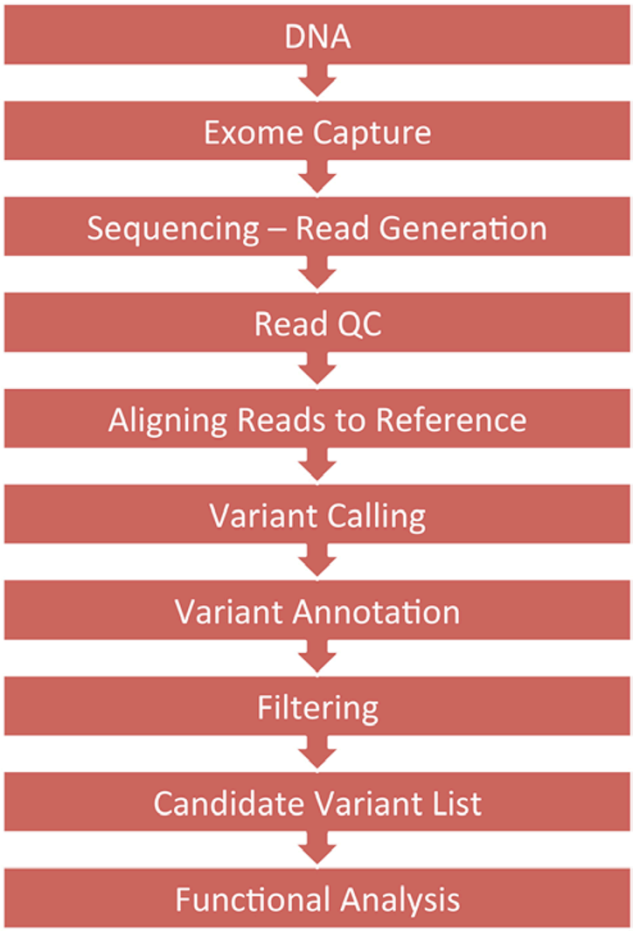

Figure 1 Typical exome sequence analysis pipeline.

deletions (INDELS). It has recently been demonstrated that comparing relative sequence depth (number of reads mapping to a location) between samples can also identify copy number changes across exons (due to larger scale INDELS or structural variants) (21). One of the challenges that arises in calling variants is systematic sequencing errors or misalignments that can result in false positive variant calls. However, this is becoming less of an issue as the quality of sequencing data and alignment algorithms improve.

Once variants are called, they next need to be annotated. Common annotation includes location in the human genome, functional class (e.g. missense or nonsense variant), frequency in large population based sequencing project data sets [such as the NHLBI exome sequencing project (22)], levels of evolutionary conservation, in silico functional predictions (e.g. SIFT (23) and Polyphen (24) for missense variants), among many others.

The exome of any given individual will possess several hundred putatively functional variants (25). Since the goal is identifying the single causal mutation, filtering strategies are critical in narrowing the list and can often go from hundreds of mutations to the disease-causing variants, although this is highly dependent on the amount and quality of data used for filtering. It is often useful to begin by filtering on annotation. For example, synonymous (coding variants that do not result in an amino acid substitution) and intronic variants (i.e. those intronic variants not affecting splicing; intronic variants are identified in exome sequencing because capture probes often extend into introns to ensure capture and high coverage of the beginning and ends of exons and mutations that affect splicing) are unlikely to be disease causing. Excluding common variants or variants observed in normal cohorts is also informative since variants causing rare Mendelian forms of disease will by definition be rare. In most cases, parental sequence data or data from other relatives provides the single most informative filter. For example, if evidence exists that the disease is recessive then parental variant information can be used to exclude variants that do not fit a recessive mode of inheritance or are homozygous in both the probands and an unaffected sibling. It has also been demonstrated by our group (see below) and others (26-27) that for genetically heterogeneous diseases, such as OI, using network-based approaches which functionally link novel genes to known disease genes is also an effective filtering strategy.

Identification of new genes causing osteogenesis imperfecta

The widespread availability and reduced cost of NGS and bioinformatics has armed the broader research community with the tools necessary to identify disease causing gene pathways. In bone research, a particularly active area of gene discovery has occurred in patients with rare bone disorders such as OI that are caused by mutations in single genes.

OI comprises a group of heritable connective tissue disorders that result in bone fragility and susceptibility to fracture, bone deformity and growth deficiency. Most cases of OI are due to dominant mutations in the type I collagen genes *COL1A1* or *COL1A2*, which alter the structure or quantity of type I collagen and cause a skeletal phenotype that ranges from subclinical to lethal. Other more rare forms of OI arise from gene mutations that affect proteins involved in collagen processing [i.e., prolyl hydroxylation, intracellular transport, or matrix incorporation (28)]. Indeed, the elucidation of the gene pathways underlying inherited forms of OI has uncovered the different molecular events that are required for

normal collagen processing (28). However, it is clear from many case reports in the literature that there are many inherited forms of OI in which the genetic defect is unknown. These cases represent a fertile area for the application of NGS. Here we describe two recent examples, which illustrate how these powerful approaches have been exploited to identify the genetic basis for rare cases of OI.

WNT1

Wnt signaling is a well-established regulator of bone development and homeostasis (29). Much of the initial evidence implicating Wnt signaling and the skeleton came from discoveries that mutations in the WNT co-receptor LRP5 (30-32), and the WNT inhibitor SOST, caused altered bone mass (33-35). More recently, genome-wide association studies have identified common polymorphisms in or near a number of genes involved in WNT signaling (LRP5, SOST, CTTNB, WNT16, etc.) that are associated with bone mass in the general population (36). Now a series of recent NGS studies have implicated WNTs in OI through the discovery of disease-causing mutations in WNT1 (6-7, 37-38).

A total of four independent studies recently reported the discovery of multiple mutations in WNT1 in individuals with OI and early-onset osteoporosis. In common to all studies was the use of NGS to identify the initial mutations in WNT1. In the work by Fahiminiya*et al.*, 148 individuals with OI type IV were investigated (38). Sanger sequencing identified mutations in either COL1A1 or COL1A2 in 134 of the subjects. Mutations in other known OI genes were identified in 6 others. The other eight probands were subjected to exome sequencing. In four individuals from three families, homozygous mutations were identified in WNT1. In a similar study, Pyott*et al.* identified four OI families where the affected probands carried homozygous mutations in WNT1 (37).

Keupp *et al.* used exome sequencing to identify a 1 bp homozygous duplication in WNT1 in three affected individuals from a large consanguineous Turkish family (7). Follow-up capillary sequencing in 11 families identified four additional homozygous mutations in WNT1. In the same study, exome sequencing in a family with early-onset osteoporosis revealed a distinct heterozygous variant in WNT1 that segregated with the disease. Functional analysis revealed that WNT1 was expressed in differentiating osteoblasts and its expression increased as a function of osteoblast maturation. Also, in contrast to wild-type WNT1, overexpression of WNT1 constructs possessing three of the identified mutations did not induce canonical WNT signaling.

In the last study, Laine *et al.* investigated the basis of autosomal dominant early-onset osteoporosis in a large multigenerational pedigree (6). Linage analysis in 10 affected and 6 unaffected family members mapped the mutation to a 25.5 Mbp region on Chromosome 12. Targeted NGS of the region revealed a single novel variant in WNT1. Similar to the work of Keupp *et al.*, the authors also investigated the genetic basis of OI in a family with two affected individuals. Similar to the other studies, exome sequencing was used to discover a homozygous variant in WNT1. The authors also demonstrated that in contrast to wild-type WNT1, overexpression of the two mutants did not induce canonical WNT signaling as measured by reporter assays and the expression of WNT target genes. Moreover, overexpression of the mutant constructs in MC3T3 osteoblasts reduced the capacity of cells to form mineralized nodules relative to wild-type WNT1. The authors also used lineage tracing expressing with Wnt1-cre transgenic and $Rosa^{mT/mG}$ reporter mice to show that Wnt1 is expressed in a subset of osteocytes in subchondral and cortical bone.

PEDF and BRIL

Studies from the Clemens lab examining the role of vascularization of bone had identified several genes that impinged on angiogenic regulatory pathways (39). Among these was *Serpinf1*, which encodes the anti-angiogenic protein Pigment Epithelial-Derived Factor (PEDF) and was known to exert potent anti-angiogenic activity in several vascular beds (40-41). The protein was originally isolated from the epithelium of the developing retina where it is believed to coordinate proliferation and differentiation of the epithelial cells (40, 42-43). PEDF is ubiquitously expressed in both human and mouse (44) and has been implicated in cell cycle control (45), fat metabolism (46), and tumorgenicity (47).

In the course of our studies on the role of PEDF in bone, two studies described inactivating mutations in the *SERPINF1* gene as the cause of OI type VI (48-49). Type VI OI is distinct from other forms of the disease in that the afflicted subjects display an osteomalacia-like phenotype characterized by thickened osteoid and delayed mineralization. In addition, bisphosphonates are generally less effective in treating type VI than other subtypes of OI (50). The discovery of *SERPINF1* as the genetic basis for OI type VI came as a surprise to the field because PEDF had no obvious connection to bone. To establish a role for PEDF in bone we characterized the skeletal features of a mouse with unrestricted loss of PEDF (51). These mice exhibit skeletal features resembling those seen in patients with OI type VI including increased unmineralized osteoid.

We next searched for cases of unexplained OI with

familial inheritance in which PEDF production was compromised. In one case seen by Joan Marini at the NIH, a young girl born to normal parents presented with severe OI and was diagnosed on the basis of histomorphometry to have OI type VI. The proband has severe OI with relative macrocephaly, extreme short stature (at 25 years of age, her length is 50th percentile for a 28 month old girl), barrel chest and scoliosis. Histomorphometric analysis performed by Dr. Francis Glorieux revealed a histological picture diagnostic of OI Type VI including accumulation of unmineralized osteoid and a "fish scale" appearance of the matrix under polarized light. Osteoblasts isolated from the patient produced little if any PEDF as would be expected in OI VI (52).

Exome sequencing was performed by Hudson Alpha, on DNA extracted from blood from the parents, proband (isolated OI case) and an unaffected sibling. The proband tested negative for mutations in known OI genes including *SERPINF1*. Exome sequencing of all four family members identified a total of 22 407 high-quality variants (SNPs and INDELS). Variants that were potentially causal were identified using a set of discrete filtering steps. We first identified variants with familial genotype patterns consistent with a recessive mode of inheritance or that were only observed in the proband (putative *de novo* variants). Variants with a frequency >1% in the population were also discarded, as these were unlikely to cause OI. Lastly, we manually inspected each variant. These analyses eliminated all but 18 variants (5 recessive and 13 *de novo*). Based on the observation that PEDF is secreted at very low levels in cultured osteoblasts of the proband (Figure 2A), we hypothesized that the mutant gene interacted with PEDF. It has been shown that genes that physically interact are often highly co-expressed. Therefore, we reasoned that one of the 18 variant genes would be connected to *SERPINF1* in a bone co-expression network. To test this prediction, we utilized a mouse co-expression network for bone generated in the Farber

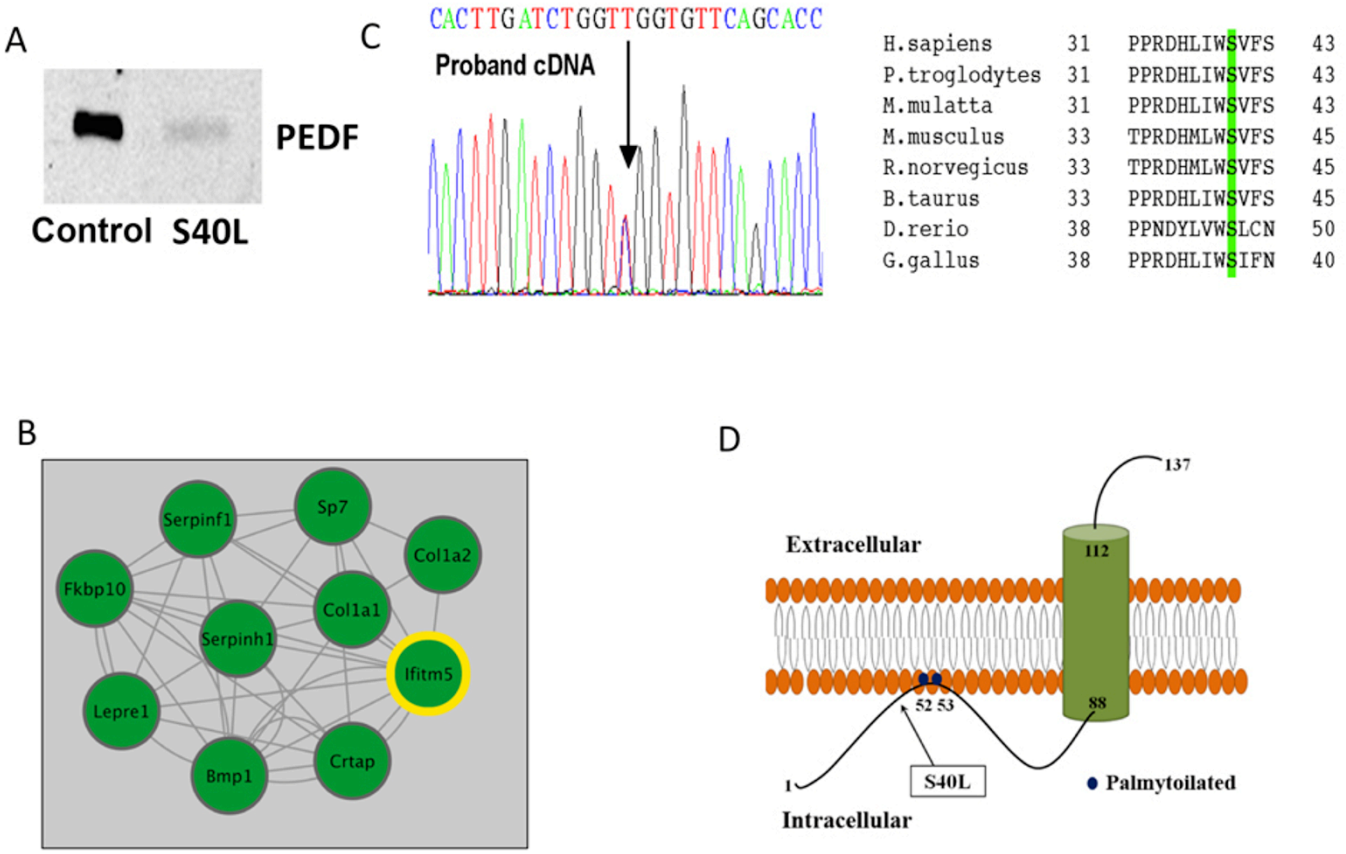

Figure 2 Identification of a dominant mutation in *IFITM5* in severe OI Type VI. A) Western blot for PEDF of conditioned medium from normal (Control) and patient (S40L) osteoblasts. B) Network illustrates the strong co-expression relationships between genes known to cause OI and *IFITM5* (outlined in yellow). C) Sanger sequencing confirmed the variant. The variant serine residue of BRIL (highlighted in green) is evolutionarily conserved. D) Predicted structure of the IFITM5/Bril protein (Pierre Moffatt, personal communication), which contains one transmembrane domains and an intracellular domain. The S40L mutation is located in the intracellular domain.

lab (53). Of the 18, *Ifitm5* was the only variant gene strongly co-expressed in bone with *Serpinf1* (Figure 2B), suggesting that it was causal.

The candidate causal *IFITM5* variant is located at 299 372 bp on chromosome 11. The variant is a C to T transition (c.119C>T) that causes a serine to leucine (p.S40L) substitution in bone restricted ifitm-like protein (BRIL; the protein encoded by *IFITM5*). Sanger sequencing confirmed that the mutation was *de novo* (Figure 2C). The variant serine residue is evolutionarily conserved (Figure 2C). *Ifitm5* also known as *fragilis4*, is a member of the mouse *fragilis* family, which consists of 5 genes and at least 3 pseudogenes (54). Ifitm proteins have evolved diverse roles, including the control of cell proliferation, promotion of homotypic cell adhesion, protection against viral infection, and facilitating germ cell development. Examination of the gene structures, chromosomal location and tissue distribution of the different members of this gene family indicates that *Ifitm5*/Bril has diverged more recently to serve different functions. As described above, Bril, is highly expressed by osteoblasts and enhances matrix mineralization by osteoblasts *in vitro* (55). However, precisely how Bril impacts bone mineralization and how the different Bril mutations impact bone in OI is still unclear.

Soon after our findings, a different mutation in the *IFITM5* gene was described as the cause autosomal dominant OI type V (3-4). Patients with OI type V have skeletal abnormalities distinct from those seen in OI Type VI (56), which were found to be caused by a recurrent mutation in the 5'-UTR of *IFITM5* resulting in a mature Bril protein with 5 additional amino acids at its N terminus (Figure 2D). Thus in a matter on a few months two new proteins, which have no obvious connection to type 1 collagen synthesis or processing, were linked to two different forms of dominantly inherited OI. These exciting findings strongly suggest that both PEDF and Bril are critical components of a novel pathway required for normal bone matrix production and mineralization. Studies to test this possibility are currently underway.

The studies described above illustrate the power of NGS to identify mutations that underlie rare, heritable skeletal diseases. It is important to note that, with the exception of the large family segregating early-onset osteoporosis, traditional linkage mapping approaches would not have been useful because of the isolated nature and small size of each independent family. Thus, by identifying how the defective gene product causes genetic disorders, one can expeditiously identify critical molecular pathways that in some cases can be immediately targeted for therapeutic intervention. This is not a new idea but because of the vast improvements (speed) in the methodologies for finding defective genes, together with the exponential reduction in cost of sequencing, this strategy is now much more widely amenable to all investigators who have access to patient populations with rare genetic disorders.

The challenge of functional validation of genetic determinants

To understand the influence of a rare mutation on a skeletal phenotype it is usually necessary to perform functional studies *in vivo*. Such studies have firmly linked several OI associated gene mutations to a specific molecular process in collagen production, assembly or intracellular trafficking and secretion (28). However, although NGS has made mutation detection relatively straightforward, it is often unclear whether the identified sequence variant is the cause of the presenting phenotype. Although *in vitro* analysis with cell lines from patients can help in this regard, it is clear that a more comprehensive investigation at the tissue, organ or whole-organism level is required. One such effort is the International Knockout Mouse Consortium, which aims to mutate all protein-coding genes in the mouse and is providing resources to many laboratories that are studying the effects of loss-of-function alleles in different organs including the skeleton. As more gain-of-function and dominant-negative rare-disease-causing mutations are identified (e.g. *IFITM5*/BRIL), there will be a need for knock-in models to recapitulate these diseases. The development and wide adoption of genome-editing tools such as zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPERs) provide a way to rapidly generate knock-in and other precise genome editing events (57). In addition to providing information on pathways and modifying factors, many of these mouse models will facilitate investigation of treatment modalities for these disorders. Such large-scale phenotyping efforts will significantly challenge the research community to devise effective ways to capture, analyze and disseminate the genomic and phenotypic data.

Summary

In this perspective, we have attempted to illustrate how modern DNA sequencing and bioinformatics can be applied in a contemporary research lab setting to identify the genetic defects underlying rare bone disorders caused by mutations in single genes. In the cases illustrated here, the genes responsible for causing OI (*Wnt1*

and *Ifitm5*) were not previously known to function in the skeleton, but their linkage to a rare bone disorder prompts a new line of discovery research that will probe the normal function of the gene product in bone. Indeed, this type of approach has already led to discovery of *SOST* (ENSG00000167941) and *LRP5* (ENSG00000162337) as critical regulators of Wnt signaling in bone and resulted in development of new drugs to stimulate bone formation (58).

A recent estimate suggests that approximately 50% of predicted 7 000 rare monogenic diseases have already been identified, and that most of the remaining disease-causing genes will be identified within the next ten years. The immediate beneficiaries of these discoveries will be the patients and families impacted by rare disease causing genes. The availability of the new body of genetic information will vastly improve clinical diagnosis and optimize treatment of both rare and more common bone disorders. Parallel translational studies informed by improved knowledge of bone controlling gene networks will clarify fundamental biological mechanisms in bone biology. We anticipate that many investigators in the bone field will participate in the incipient decade of discovery.

References

1 Becker J, Semler O, Gilissen C, Li Y, Bolz HJ, Giunta C, Bergmann C, Rohrbach M, Koerber F, Zimmermann K, de Vries P, Wirth B, Schoenau E, Wollnik B, Veltman JA, Hoischen A, Netzer C. Exome sequencing identifies truncating mutations in human SERPINF1 in autosomal-recessive osteogenesis imperfecta. Am J Hum Genet. 2011;88:362-371.

2 Shaheen R, Alazami AM, Alshammari MJ, Faqeih E, Alhashmi N, Mousa N, Alsinani A, Ansari S, Alzahrani F, Al-Owain M, Alzayed ZS, Alkuraya FS. Study of autosomal recessive osteogenesis imperfecta in Arabia reveals a novel locus defined by TMEM38B mutation. J Med Genet. 2012;49:630-635.

3 Semler O, Garbes L, Keupp K, Swan D, Zimmermann K, Becker J, Iden S, Wirth B, Eysel P, Koerber F, Schoenau E, Bohlander SK, Wollnik B, Netzer C. A mutation in the 5'-UTR of IFITM5 creates an in-frame start codon and causes autosomal-dominant osteogenesis imperfecta type V with hyperplastic callus. Am J Hum Genet. 2012;91:349-357.

4 Cho TJ, Lee KE, Lee SK, Song SJ, Kim KJ, Jeon D, Lee G, Kim HN, Lee HR, Eom HH, Lee ZH, Kim OH, Park WY, Park SS, Ikegawa S, Yoo WJ, Choi IH, Kim JW. A single recurrent mutation in the 5'-UTR of IFITM5 causes osteogenesis imperfecta type V. Am J Hum Genet. 2012;91:343-348.

5 Venturi G, Gandini A, Monti E, Dalle Carbonare L, Corradi M, Vincenzi M, Valenti MT, Valli M, Pelilli E, Boner A, Mottes M, Antoniazzi F. Lack of expression of SERPINF1, the gene coding for pigment epithelium-derived factor, causes progressively deforming osteogenesis imperfecta with normal type I collagen. J Bone Miner Res. 2012;27:723-728.

6 Laine CM, Joeng KS, Campeau PM, Kiviranta R, Tarkkonen K, Grover M, Lu JT, Pekkinen M, Wessman M, Heino TJ, Nieminen-Pihala V, Aronen M, Laine T, Kröger H, Cole WG, Lehesjoki AE, Nevarez L, Krakow D, Curry CJ, Cohn DH, Gibbs RA, Lee BH, Mäkitie O. WNT1 mutations in early-onset osteoporosis and osteogenesis imperfecta. N Engl J Med. 2013;368:1809-1816.

7 Keupp K, Beleggia F, Kayserili H, Barnes AM, Steiner M, Semler O, Fischer B, Yigit G, Janda CY, Becker J, Breer S, Altunoglu U, Grünhagen J, Krawitz P, Hecht J, Schinke T, Makareeva E, Lausch E, Cankaya T, Caparrós-Martín JA, Lapunzina P, Temtamy S, Aglan M, Zabel B, Eysel P, Koerber F, Leikin S, Garcia KC, Netzer C, Schönau E, Ruiz-Perez VL, Mundlos S, Amling M, Kornak U, Marini J, Wollnik B. Mutations in WNT1 cause different forms of bone fragility. Am J Hum Genet. 2013;92:565-574.

8 Glazov EA, Zankl A, Donskoi M, Kenna TJ, Thomas GP, Clark GR, Duncan EL, Brown MA. Whole-exome re-sequencing in a family quartet identifies POP1 mutations as the cause of a novel skeletal dysplasia. PLoS Genet. 2011;7:e1002027.

9 Min BJ, Kim N, Chung T, Kim OH, Nishimura G, Chung CY, Song HR, Kim HW, Lee HR, Kim J, Kang TH, Seo ME, Yang SD, Kim DH, Lee SB, Kim JI, Seo JS, Choi JY, Kang D, Kim D, Park WY, Cho TJ. Whole-exome sequencing identifies mutations of KIF22 in spondyloepimetaphyseal dysplasia with joint laxity, leptodactylic type. Am J Hum Genet. 2011;89:760-766.

10 Davis E, Savage J, Willer J, Jiang YH, Angrist M, Androutsopoulos A, Katsanis N. Whole exome sequencing and functional studies identify an intronic mutation in TRAPPC2 that causes SEDT. Clin Genet. 2013 May 8. [Epub ahead of print]

11 Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet. 2010;11:31-46.

12 Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara E Catenazzi M, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM,

Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Rogers J, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R, Smith AJ. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008;456:53-59.

13 Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schorn M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza JA, Namsaraev E, McKernan KJ, Williams A, Roth GT, Bustillo J. An integrated semiconductor device enabling non-optical genome sequencing. Nature. 2011;475:348-352.

14 Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. Real-time DNA sequencing from single polymerase molecules. Science. 2009;323:133-138.

15 Stoddart D, Heron AJ, Mikhailova E, Maglia G, Bayley H. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. Proc Natl Acad Sci U S A. 2009;106: 7702-7707.

16 Borecki IB, Province MA. Genetic and genomic discovery using family studies. Circulation. 2008;118:1057-1063.

17 Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet. 2011;12:745-755.

18 Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, Nayir A, Bakkaloğlu A, Ozen S, Sanjad S, Nelson-Williams C, Farhi A, Mane S, Lifton RP. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A. 2009;106:19096-19101.

19 Gogol-Döring A, Chen W. An overview of the analysis of next generation sequencing data. Methods Mol Biol. 2012;802:249-257.

20 Pabinger S, Dander A, Fischer M, Snajder R, Sperk M, Efremova M, Krabichler B, Speicher MR, Zschocke J, Trajanoski Z. A survey of tools for variant analysis of next-generation genome sequencing data. Brief Bioinform. 2013 Jan 21. [Epub ahead of print]

21 Krumm N, Sudmant PH, Ko A, O'Roak BJ, Malig M, Coe BP; NHLBI Exome Sequencing Project, Quinlan AR, Nickerson DA, Eichler EE. Copy number variation detection and genotyping from exome sequence data. Genome Res. 2012;22:1525-1532.

22 Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, Kang HM, Jordan D, Leal SM, Gabriel S, Rieder MJ, Abecasis G, Altshuler D, Nickerson DA, Boerwinkle E, Sunyaev S, Bustamante CD, Bamshad MJ, Akey JM; Broad GO; Seattle GO; NHLBI Exome Sequencing Project. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science. 2012;337:64-69.

23 Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4:1073-1081.

24 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7: 248-249.

25 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491:56-65.

26 Erlich Y, Edvardson S, Hodges E, Zenvirt S, Thekkat P, Shaag A, Dor T, Hannon GJ, Elpeleg O. Exome sequencing and disease-network analysis of a single family implicate a mutation in KIF1A in hereditary spastic paraparesis. Genome Res. 2011;21:658-664.

27 Zhu J, Cui L, Wang W, Hang XY, Xu AX, Yang SX, Dou JT, Mu YM, Zhang X, Gao JP. Whole exome sequencing identifies mutation of EDNRA involved in ACTH-independent macronodular adrenal hyperplasia. Fam Cancer. 2013;12:657-667.

28 Eyre DR, Weis MA. Bone collagen: new clues to its mineralization mechanism from recessive osteogenesis imperfecta. Calcif Tissue Int. 2013;93:338-347.

29 Regard JB, Zhong Z, Williams BO, Yang Y. Wnt signaling in bone development and disease: making stronger bone with Wnts. Cold Spring Harb Perspect Biol. 2012;4. pii: a007997.

30 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, Batch JA, Beighton P, Black GC, Boles RG, Boon LM, Borrone C, Brunner HG, Carle GF, Dallapiccola B, De Paepe A, Floege B, Halfhide ML, Hall B, Hennekam RC, Hirose T, Jans A, Jüppner H, Kim CA, Keppler-Noreuil K, Kohlschuetter A, LaCombe D, Lambert M, Lemyre E,

Letteboer T, Peltonen L, Ramesar RS, Romanengo M, Somer H, Steichen-Gersdorf E, Steinmann B, Sullivan B, Superti-Furga A, Swoboda W, van den Boogaard MJ, Van Hul W, Vikkula M, Votruba M, Zabel B, Garcia T, Baron R, Olsen BR, Warman ML; Osteoporosis-Pseudoglioma Syndrome Collaborative Group. LDL receptor-related protein 5 (LRP5) affects bone accrual and eye development. Cell. 2001;107:513-523.

31 Little RD, Carulli JP, Del Mastro RG, Dupuis J, Osborne M, Folz C, Manning SP, Swain PM, Zhao SC, Eustace B, Lappe MM, Spitzer L, Zweier S, Braunschweiger K, Benchekroun Y, Hu X, Adair R, Chee L, FitzGerald MG, Tulig C, Caruso A, Tzellas N, Bawa A, Franklin B, McGuire S, Nogues X, Gong G, Allen KM, Anisowicz A, Morales AJ, Lomedico PT, Recker SM, Van Eerdewegh P, Recker RR, Johnson ML. A mutation in the LDL receptor-related protein 5 gene results in the autosomal dominant high-bone-mass trait. Am J Hum Genet. 2002;70:11-19.

32 Boyden LM, Mao J, Belsky J, Mitzner L, Farhi A, Mitnick MA, Wu D, Insogna K, Lifton RP. High bone density due to a mutation in LDL-receptor-related protein 5. N Engl J Med. 2002;346:1513-1521.

33 Brunkow ME, Gardner JC, Van Ness J, Paeper BW, Kovacevich BR, Proll S, Skonier JE, Zhao L, Sabo PJ, Fu Y, Alisch RS, Gillett L, Colbert T, Tacconi P, Galas D, Hamersma H, Beighton P, Mulligan J. Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein. Am J Hum Genet. 2001;68:577-589.

34 Balemans W, Patel N, Ebeling M, Van Hul E, Wuyts W, Lacza C, Dioszegi M, Dikkers FG, Hildering P, Willems PJ, Verheij JB, Lindpaintner K, Vickery B, Foernzler D, Van Hul W. Identification of a 52 kb deletion downstream of the SOST gene in patients with van Buchem disease. J Med Genet. 2002;39:91-97.

35 Balemans W, Ebeling M, Patel N, Van Hul E, Olson P, Dioszegi M, Lacza C, Wuyts W, Van Den Ende J, Willems P, Paes-Alves AF, Hill S, Bueno M, Ramos FJ, Tacconi P, Dikkers FG, Stratakis C, Lindpaintner K, Vickery B, Foernzler D, Van Hul W. Increased bone density in sclerosteosis is due to the deficiency of a novel secreted protein (SOST). Hum Mol Genet. 2001;10:537-543.

36 Estrada K, Styrkarsdottir U, Evangelou E, Hsu YH, Duncan EL, Ntzani EE, Oei L, Albagha OM, Amin N, Kemp JP, Koller DL, Li G, Liu CT, Minster RL, Moayyeri A, Vandenput L, Willner D, Xiao SM, Yerges-Armstrong LM, Zheng HF, Alonso N, Eriksson J, Kammerer CM, Kaptoge SK, Leo PJ, Thorleifsson G, Wilson SG, Wilson JF, Aalto V, Alen M, Aragaki AK, Aspelund T, Center JR, Dailiana Z, Duggan DJ, Garcia M, Garcia-Giralt N, Giroux S, Hallmans G, Hocking LJ, Husted LB, Jameson KA, Khusainova R, Kim GS, Kooperberg C, Koromila T, Kruk M, Laaksonen M, Lacroix AZ, Lee SH, Leung PC, Lewis JR, Masi L, Mencej-Bedrac S, Nguyen TV, Nogues X, Patel MS, Prezelj J, Rose LM, Scollen S, Siggeirsdottir K, Smith AV, Svensson O, Trompet S, Trummer O, van Schoor NM, Woo J, Zhu K, Balcells S, Brandi ML, Buckley BM, Cheng S, Christiansen C, Cooper C, Dedoussis G, Ford I, Frost M, Goltzman D, González-Macías J, Kähönen M, Karlsson M, Khusnutdinova E, Koh JM, Kollia P, Langdahl BL, Leslie WD, Lips P, Ljunggren Ö, Lorenc RS, Marc J, Mellström D, Obermayer-Pietsch B, Olmos JM, Pettersson-Kymmer U, Reid DM, Riancho JA, Ridker PM, Rousseau F, Slagboom PE, Tang NL, Urreizti R, Van Hul W, Viikari J, Zarrabeitia MT, Aulchenko YS, Castano-Betancourt M, Grundberg E, Herrera L, Ingvarsson T, Johannsdottir H, Kwan T, Li R, Luben R, Medina-Gómez C, Palsson ST, Reppe S, Rotter JI, Sigurdsson G, van Meurs JB, Verlaan D, Williams FM, Wood AR, Zhou Y, Gautvik KM, Pastinen T, Raychaudhuri S, Cauley JA, Chasman DI, Clark GR, Cummings SR, Danoy P, Dennison EM, Eastell R, Eisman JA, Gudnason V, Hofman A, Jackson RD, Jones G, Jukema JW, Khaw KT, Lehtimäki T, Liu Y, Lorentzon M, McCloskey E, Mitchell BD, Nandakumar K, Nicholson GC, Oostra BA, Peacock M, Pols HA, Prince RL, Raitakari O, Reid IR, Robbins J, Sambrook PN, Sham PC, Shuldiner AR, Tylavsky FA, van Duijn CM, Wareham NJ, Cupples LA, Econs MJ, Evans DM, Harris TB, Kung AW, Psaty BM, Reeve J, Spector TD, Streeten EA, Zillikens MC, Thorsteinsdottir U, Ohlsson C, Karasik D, Richards JB, Brown MA, Stefansson K, Uitterlinden AG, Ralston SH, Ioannidis JP, Kiel DP, Rivadeneira F. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. Nat Genet. 2012;44:491-501.

37 Pyott SM, Tran TT, Leistritz DF, Pepin MG, Mendelsohn NJ, Temme RT, Fernandez BA, Elsayed SM, Elsobky E, Verma I, Nair S, Turner EH, Smith JD, Jarvik GP, Byers PH. WNT1 mutations in families affected by moderately severe and progressive recessive osteogenesis imperfecta. Am J Hum Genet. 2013;92:590-597.

38 Fahiminiya S, Majewski J, Mort J, Moffatt P, Glorieux FH, Rauch F. Mutations in WNT1 are a cause of osteogenesis imperfecta. J Med Genet. 2013;50:345-348.

39 Wang Y, Wan C, Deng L, Liu X, Cao X, Gilbert SR, Bouxsein ML, Faugere MC, Guldberg RE, Gerstenfeld LC, Haase VH, Johnson RS, Schipani E, Clemens TL. The hypoxia-inducible factor alpha pathway couples angiogenesis to osteogenesis during skeletal development. J Clin Invest. 2007;117:1616-1626.

40 Tombran-Tink J, Barnstable CJ. PEDF: a multifaceted neurotrophic factor. Nat Rev Neurosci. 2003;4:628-636.

41 Crawford SE, Fitchev P, Veliceasa D, Volpert OV. The many facets of PEDF in drug discovery and disease: a diamond in the rough or split personality disorder? Expert Opin Drug Discov. 2013;8:769-792.

42 Doll JA, Stellmach VM, Bouck NP, Bergh AR, Lee C, Abramson LP, Cornwell ML, Pins MR, Borensztajn J, Crawford SE. Pigment epithelium-derived factor regulates the vasculature and mass of the prostate and pancreas. Nat Med. 2003;9:774-780.

43 Dawson DW, Volpert OV, Gillis P, Crawford SE, Xu H, Benedict W, Bouck NP. Pigment epithelium-derived factor: a potent inhibitor of angiogenesis. Science. 1999;285:245-248.

44 Ek ET, Dass CR, Choong PF. PEDF: a potential molecular therapeutic target with multiple anti-cancer activities. Trends Mol Med. 2006;12:497-502.

45 Filleur S, Nelius T, de Riese W, Kennedy RC. Characterization of PEDF: a multi-functional serpin family protein. J Cell Biochem. 2009;106:769-775.

46 Chung C, Doll JA, Gattu AK, Shugrue C, Cornwell M, Fitchev P, Crawford SE. Anti-angiogenic pigment epithelium-derived factor regulates hepatocyte triglyceride content through adipose triglyceride lipase (ATGL). J Hepatol. 2008;48:471-478.

47 Ho TC, Chen SL, Shih SC, Chang SJ, Yang SL, Hsieh JW, Cheng HC, Chen LJ, Tsao YP. Pigment epithelium-derived factor (PEDF) promotes tumor cell death by inducing macrophage membrane tumor necrosis factor-related apoptosis-inducing ligand (TRAIL). J Biol Chem. 2011;286:35943-35954.

48 Homan EP, Rauch F, Grafe I, Lietman C, Doll JA, Dawson B, Bertin T, Napierala D, Morello R, Gibbs R, White L, Miki R, Cohn DH, Crawford S, Travers R, Glorieux FH, Lee B. Mutations in SERPINF1 cause osteogenesis imperfecta type VI. J Bone Miner Res. 2011;26:2798-2803.

49 Tucker T, Nelson T, Sirrs S, Roughley P, Glorieux FH, Moffatt P, Schlade-Bartusiak K, Brown L, Rauch F. A co-occurrence of osteogenesis imperfecta type VI and cystinosis. Am J Med Genet A. 2012;158A:1422-1426.

50 Land C, Rauch F, Travers R, Glorieux FH. Osteogenesis imperfecta type VI in childhood and adolescence: effects of cyclical intravenous pamidronate treatment. Bone. 2007;40:638-644.

51 Bogan R, Riddle RC, Li Z, Kumar S, Nandal A, Faugere MC, Boskey A, Crawford SE, Clemens TL. A mouse model for human osteogenesis imperfecta type VI. J Bone Miner Res. 2013;28:1531-1536.

52 Rauch F, Husseini A, Roughley P, Glorieux FH, Moffatt P. Lack of circulating pigment epithelium-derived factor is a marker of osteogenesis imperfecta type VI. J Clin Endocrinol Metab. 2012;97:E1550-E1556.

53 Calabrese G, Bennett BJ, Orozco L, Kang HM, Eskin E, Dombret C, De Backer O, Lusis AJ, Farber CR. Systems genetic analysis of osteoblast-lineage cells. PLoS Genet. 2012;8:e1003150.

54 Lange UC, Saitou M, Western PS, Barton SC, Surani MA. The fragilis interferon-inducible gene family of transmembrane proteins is associated with germ cell specification in mice. BMC Dev Biol. 2003;3:1.

55 Moffatt P, Gaumond MH, Salois P, Sellin K, Bessette MC, Godin E, de Oliveira PT, Atkins GJ, Nanci A, Thomas G. Bril: a novel bone-specific modulator of mineralization. J Bone Miner Res. 2008;23:1497-1508.

56 Rauch F, Moffatt P, Cheung M, Roughley P, Lalic L, Lund AM, Ramirez N, Fahiminiya S, Majewski J, Glorieux FH. Osteogenesis imperfecta type V: marked phenotypic variability despite the presence of the IFITM5 c.-14C>T mutation in all patients. J Med Genet. 2013;50:21-24.

57 Gaj T, Gersbach CA, Barbas CF 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 2013;31:397-405.

58 Paszty C, Turner CH, Robinson MK. Sclerostin: a gem from the genome leads to bone-building antibodies. J Bone Miner Res. 2010;25:1897-1904.