



UNIVERSITAT DE  
BARCELONA

## Identification and functional characterization of genetic *loci* involved in osteoporosis and atypical femoral fracture

Neus Roca Ayats

**ADVERTIMENT.** La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX ([www.tdx.cat](http://www.tdx.cat)) i a través del Dipòsit Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

**ADVERTENCIA.** La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR ([www.tdx.cat](http://www.tdx.cat)) y a través del Repositorio Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**WARNING.** On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX ([www.tdx.cat](http://www.tdx.cat)) service and by the UB Digital Repository ([diposit.ub.edu](http://diposit.ub.edu)) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Tesi Doctoral  
Universitat de Barcelona

**Identification and functional characterization of  
genetic *loci* involved in osteoporosis and atypical  
femoral fracture**

Memòria presentada per  
**Neus Roca Ayats**

per optar al grau de  
**Doctora per la Universitat de Barcelona**  
Programa de Genètica  
Departament de Genètica, Microbiologia i Estadística  
Curs 2018-2019

Tesi dirigida pel **Dr. Daniel Grinberg Vaisman** i la **Dra. Susanna Balcells Comas** al Departament de Genètica, Microbiologia i Estadística de la Facultat de Biologia de la Universitat de Barcelona

Dr. Daniel Grinberg Vaisman

Dra. Susanna Balcells Comas

Neus Roca Ayats  
Barcelona, 2019



Als meus pares i la meva germana



# AGRAÏMENTS

Vull donar les gràcies a totes aquelles persones que, d'una manera o altra, han fet possible aquesta tesi.

Breument:

A en Dani i la Susanna, per la confiança, paciència i comprensió. I per tot el que m'han ensenyat.

A l'Adolfo i la Natàlia, per haver-me obert les portes de les fractures atípiques i tot el que ha comportat. També a la resta de gent del grup de l'IMIM.

A totes les companyes i companys que han passat pels laboratoris 1 i 2 de la 3a planta durant aquests anys, especialment a aquelles amb qui he compartit més temps. He après molt de totes i cadascuna d'elles, tant dins com fora del laboratori. A la Núria, amb qui he treballat colze a colze, i a la Mònica, per la seva ajuda inestimable.

A la resta de membres del departament (ara secció) de Genètica, perquè sempre hi he trobat el cop de mà, consell o suport que he necessitat.

A tots els col·laboradors d'aquesta recerca. In particular, to Darío and the members of the Mundlos' group at MPI für molekulare Genetik in Berlin, for kindly teaching me so much.

A les propietàries de les molècules d'ADN que han passat per les meves mans, per haver-les cedit perquè xafardegessim una mica el seu manual d'instruccions.

Al Ministerio de Educación, Cultura y Deporte, per la concessió d'un contracte predoctoral i al MINECO, la Generalitat de Catalunya, la FEIOMM i el CIBERER per haver finançat aquesta recerca i la meva formació.

Als meus amics i la meva família, pel seu interès desinteressat per aquesta tesi i per la meva persona. A la meva germana i als meus pares, per haver-me fet costat sempre.

Moltes gràcies.



## ABSTRACT

Osteoporosis is a complex disease characterized by low bone mass, microarchitectural deterioration and increased fracture risk. Many genes/variants associated with osteoporosis have been identified but the underlying mechanisms are poorly understood. Hence, it is necessary to identify new variants/genes and to functionally characterize them. Nitrogen-containing bisphosphonates (N-BPs) are first-line treatment for osteoporosis that prevent osteoclast function. Very rarely, atypical femoral fractures (AFFs) occur after a long-term therapy. The pathogenic mechanisms underlying AFF remain unknown.

This PhD thesis contributed to the elucidation of the genetic determinants of osteoporosis and AFF.

On one side, we dissected the association signal in *C7ORF76* (7q21.3) in the BARCOS cohort (postmenopausal women) and functionally characterized the associated variants and regulatory elements within the *locus*. We identified 2 variants associated with BMD and osteoporotic fracture and showed that they are *cis*-eQTL for the neighbouring gene *SLC25A13* in primary osteoblasts. An upstream putative regulatory element (UPE) contained one of the variants and was functionally studied. Its regulatory capacity was demonstrated and it was shown to interact with a lncRNA and other regulatory elements within the region.

We also studied a previously described mouse *Dlx5/6* enhancer (eDlx#18) within the *locus*. It activated transcription in an osteoblastic context and it interacted with the *DLX5* promoter and with other *DLX5/6* enhancers. A SNP within eDlx#18 was shown to be a *cis*-eQTL for *DLX6* in primary osteoblasts. Finally, the homozygous deletion of eDlx#18 in mice resulted in reduced viability, decreased *Dlx5* expression in otic vesicle and branchial arches in E11.5 embryos, and a smaller dentary and several ossification defects in E17.5 embryos.

On the other side, a small cohort of N-BP-associated AFF patients was analysed by whole exome sequencing. We found 37 rare mutations in 34 genes shared by 3 sisters, including mutations in *GGPS1* and *CYP1A1*, also mutated in one unrelated patient. We functionally demonstrated that the p.Asp188Tyr mutation in *GGPS1* affects oligomerization of the enzyme and leads to a severe reduction in enzyme activity. *GGPS1* depletion in osteoblasts resulted in a strong mineralization reduction and a decreased expression of some osteoblastic markers. The depletion in osteoclast precursors led to increased osteoclast numbers but with reduced resorption activity.



---

# TABLE OF CONTENTS

## ABBREVIATIONS

<b>INTRODUCTION.....</b>	<b>1</b>
1. The bone tissue .....	3
1.1. Definition and function.....	3
1.2. Structure and types of bone .....	3
1.3. Constituents of bone .....	3
1.3.1. Cells .....	4
1.3.2. Extracellular matrix.....	6
1.4. Bone formation and development.....	7
1.5. Bone remodeling and homeostasis .....	8
2. Osteoporosis.....	10
2.1. Definition and classification .....	10
2.2. Diagnosis: bone mineral density.....	11
2.3. Epidemiology .....	13
2.4. Risk factors .....	13
2.5. Prevention and treatment.....	15
2.5.1. Bisphosphonates.....	18
3. Genetics of osteoporosis.....	23
3.1. Heritability of bone properties.....	23
3.2. Linkage analyses .....	24
3.3. Genetic association studies.....	24
3.3.1. Candidate genes association studies .....	25
3.3.2. Genome-wide association studies .....	26
3.4. Epigenetics of osteoporosis .....	34
3.5. Functional studies .....	35
3.5.1. Functional genomic integrative analysis .....	36
3.5.2. Gene expression studies and eQTLs.....	36
3.5.3. Chromatin conformation analysis .....	37
3.5.4. Other functional assays.....	38

## TABLE OF CONTENTS

---

3.5.5. Genome editing and animal models .....	39
3.6. Biological pathways underlying osteoporosis .....	40
3.6.1. Wnt/ $\beta$ -catenin signalling .....	40
3.6.2. OPG-RANK-RANKL signalling .....	41
3.6.3. NOTCH signalling.....	41
3.6.4. TGF- $\beta$ /BMP signalling .....	42
3.6.5. Ephrin signalling .....	42
3.6.6. Endochondral ossification and MSCs differentiation .....	43
4. Atypical femoral fracture .....	44
4.1. Definition and diagnosis .....	44
4.2. Epidemiology and risk factors.....	45
4.3. Pathogenesis .....	46
4.4. Genetics.....	47
4.4.1. Genetic studies in small AFF cohorts .....	48
4.4.2. Genetic studies in AFF patients with other monogenic bone diseases.	50
4.5. Prevention and management of AFF.....	53
<b>OBJECTIVES.....</b>	<b>55</b>
<b>RESULTS.....</b>	<b>59</b>
Thesis supervisors' report on the contribution of the PhD candidate to the articles included in this doctoral thesis .....	61
CHAPTER 1: Functional characterization of the <i>C7ORF76</i> locus, a GWAS signal for BMD and osteoporotic fracture .....	65
Article 1: "Functional characterization of the <i>C7ORF76</i> genomic region, a prominent GWAS signal for osteoporosis in 7q21.3" .....	65
Article 2: "A <i>DLX5/6</i> enhancer in the <i>C7ORF76</i> locus: Characterization of its role in development and in bone".....	79
CHAPTER 2: Identification and characterization of genetic susceptibility to N-BP-associated AFF .....	107
Article 3: " <i>GGPS1</i> Mutation and Atypical Femoral Fractures with Bisphosphonates" .....	107
Article 4: "Functional Characterization of a GGPPS Variant Identified in Atypical Femoral Fracture Patients and Delineation of the Role of GGPPS in Bone-Relevant Cell Types".....	111
Annex to Chapter 2.....	129

<b>DISCUSSION</b> .....	<b>139</b>
1. Homogeneity of phenotype .....	141
2. Next generation sequencing .....	143
3. Association studies .....	145
4. Replication of identified variants.....	152
5. Functional studies of variants and candidate genes.....	153
6. Epistasis and gene-environment interaction: pharmacogenetics and personalized medicine .....	165
7. Future perspectives .....	169
<b>CONCLUSIONS</b> .....	<b>171</b>
<b>REFERENCES</b> .....	<b>175</b>
<b>ANNEX</b> .....	<b>207</b>



**ABBREVIATIONS**

<b>AFF</b>	Atypical femoral fracture
<b>BMD</b>	Bone mineral density
<b>BP</b>	Bisphosphonate
<b>BSP</b>	Bone sialoprotein
<b>CNV</b>	Copy number variation
<b>CT</b>	Computed tomography
<b>CTSK</b>	Cathepsin K
<b>CRISPR</b>	Clustered regularly interspaced short palindromic repeats
<b>DLX</b>	Distal-less homeobox
<b>DXA</b>	Dual-energy X-ray absorptiometry
<b>ECM</b>	Extracellular matrix
<b>eQTL</b>	Expression quantitative trait loci
<b>FGF</b>	Fibroblast growth factor
<b>FN</b>	Femoral neck
<b>FPP</b>	Farnesyl pyrophosphate
<b>FPPS</b>	Farnesyl pyrophosphate synthase
<b>FRAX</b>	Fracture Risk Assessment tool
<b>GC</b>	Glucocorticoid
<b>GEFOS</b>	Genetic Factors for Osteoporosis
<b>GENOMOS</b>	Genetic Markers for Osteoporosis
<b>GGPP</b>	Geranylgeranyl pyrophosphate
<b>GGPPS</b>	Geranylgeranyl pyrophosphate synthase
<b>GWAS</b>	Genome-wide association study
<b>HBM</b>	High bone mass
<b>HH</b>	Hedgehog
<b>HOX</b>	Homeobox
<b>HPP</b>	Hypophosphatasia
<b>HSC</b>	Hematopoietic stem cell
<b>HWE</b>	Hardy-Weinberg equilibrium

## ABBREVIATIONS

---

<b>IPP</b>	Isopentenyl pyrophosphate
<b>LD</b>	Linkage disequilibrium
<b>lncRNA</b>	Long non-coding RNA
<b>LS</b>	Lumbar spine
<b>MAF</b>	Minor allele frequency
<b>M-CSF</b>	Macrophage colony stimulating factor
<b>MMP</b>	Matrix metalloprotease
<b>MPRA</b>	Massive parallel reporter assays
<b>MRI</b>	Magnetic resonance imaging
<b>MSC</b>	Mesenchymal stem cell
<b>NGS</b>	Next generation sequencing
<b>OB</b>	Osteoblast
<b>OC</b>	Osteoclast
<b>OCN</b>	Osteocalcin
<b>OCy</b>	Osteocyte
<b>OF</b>	Osteoporotic fracture
<b>OI</b>	Osteogenesis imperfecta
<b>ONJ</b>	Osteonecrosis of the jaw
<b>OPC</b>	Osteoclast precursor cell
<b>OPG</b>	Osteoprotegerin
<b>OPN</b>	Osteopontin
<b>OPPG</b>	Osteoporosis pseudoglioma syndrome
<b>OSX</b>	Osterix
<b>PBM</b>	Peak bone mass
<b>PP<sub>i</sub></b>	Inorganic pyrophosphate
<b>PTH</b>	Parathyroid hormone
<b>PYCD</b>	Pycnodysostosis
<b>QUS</b>	Quantitative ultrasound
<b>RANK</b>	Receptor activator of NF-κB
<b>RANKL</b>	Receptor activator of NF-κB ligand
<b>RUNX2</b>	Runt-related transcription factor 2
<b>SERM</b>	Selective estrogen receptor modulator
<b>SNP</b>	Single nucleotide polymorphism
<b>TAD</b>	Topologically associating domain

<b>TBS</b>	Trabecular bone score
<b>TF</b>	Transcription factor
<b>TGF-<math>\beta</math></b>	Transforming growth factor $\beta$
<b>TRAP</b>	Tartrate-resistant acid phosphatase
<b>TSS</b>	Transcription start site
<b>UTR</b>	Untranslated transcribed region
<b>WES</b>	Whole exome sequencing
<b>WGS</b>	Whole genome sequencing
<b>WHO</b>	World Health Organization
<b>XLH</b>	X-linked hyperphosphatemia



# **INTRODUCTION**



# **1. THE BONE TISSUE**

## **1.1. Definition and function**

Bone is a highly specialized dynamic mineralized connective tissue that, together with cartilage, constitutes the vertebrate skeletal system. It is a rigid tissue with a high resistance to both traction and compression.

It exerts a wide range of functions, including a mechanical function, providing structural support and facilitating movement; a protective function of all the internal systems, including vital organs and bone marrow; and a metabolic and endocrine function, regulating calcium and phosphorous homeostasis and energy metabolism (reviewed in Harada & Rodan, 2003; Lieben *et al.*, 2009). In addition, it provides the environment for hematopoiesis within the bone marrow.

## **1.2. Structure and types of bone**

Bone is normally formed in a lamellar pattern, in which collagen fibrils are laid down in alternating orientations, thus conferring a significant strength.

According to histology, bone tissue can be classified into cortical or compact bone and trabecular, cancellous or spongy bone. Cortical bone is dense, hard, with low porosity and it is mostly calcified. It has a low turnover rate and is responsible for mechanical activity. It is located mainly in the diaphysis of long bones (Box 1) and external part of all bones. Trabecular bone has a porous bone matrix and a high turnover rate. It is responsible for metabolic functions and harbours red bone matrix. It mainly constitutes the epiphysis of long bones and the interior of flat bones (reviewed in Clarke, 2008).

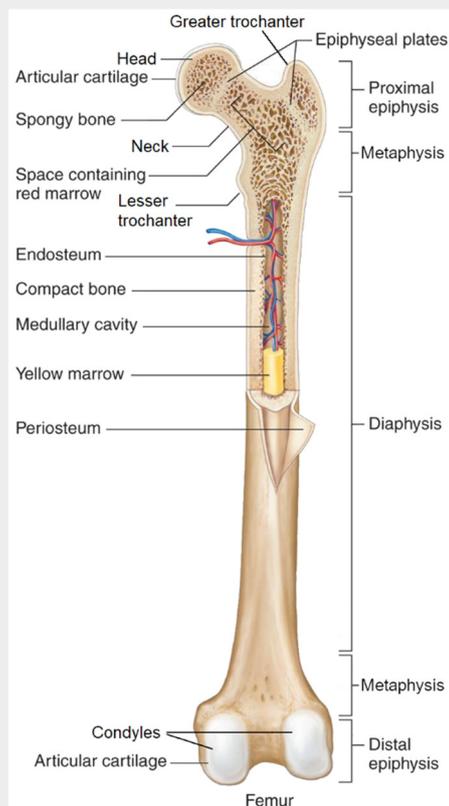
## **1.3. Constituents of bone**

Bone is formed by different cell types and a mineralized extracellular matrix.

**Box 1. Structure of long bones**

Long bones are composed of three parts: a shaft or diaphysis, two wide rounded ends or epiphyses, and two flared cone-shaped metaphyses between the diaphysis and the epiphyses. Metaphyses limit with epiphyseal or growth plates, made of cartilage, that allow the bone to progressively increase its length. Once the growth is complete, this cartilage is substituted by bone tissue.

Bones are covered by the periosteum, except for the articular regions, covered by hyaline cartilage. In the internal part of the bone, there is the medullary cavity, containing the yellow bone marrow. It is coated by the endosteum, as they are the cavities of the trabecular bone. Periosteum and endosteum are fibrous layers of connective tissue that contain osteoprogenitor cells and are responsible for bone growth, remodelling and fracture repair (reviewed in Clarke, 2008).



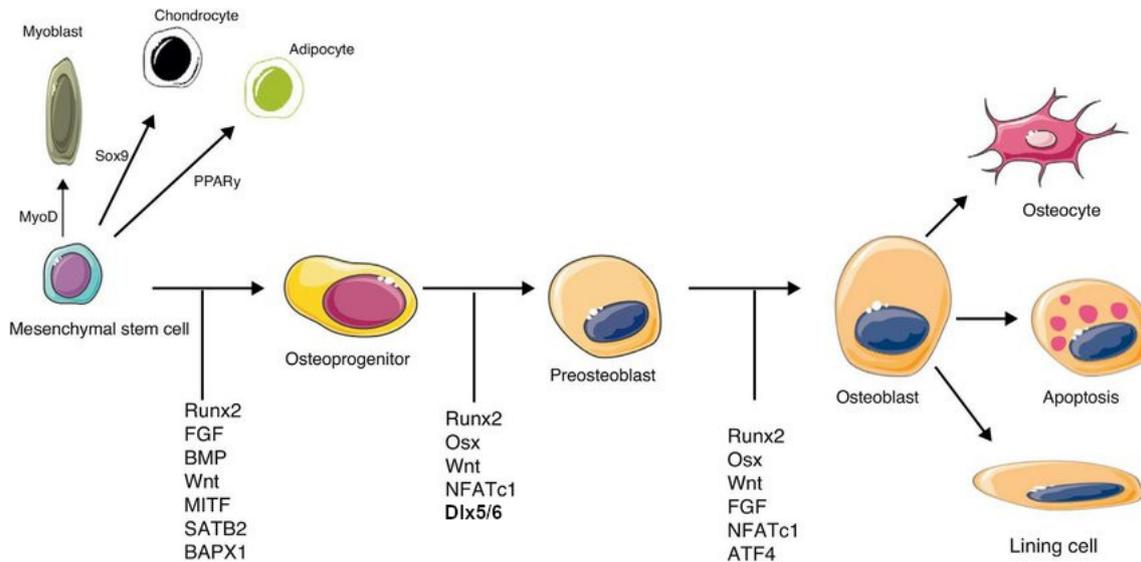
Modified from Shier *et al.*, 2016

**1.3.1. Cells**

There are 4 different cell types in bone: osteoblasts (OBs), osteoclasts (OCs), osteocytes (OCys) and lining cells.

Osteoblasts are bone forming cells that secrete a non-mineralized bone matrix, called osteoid, and are also involved in its mineralization. They have a cuboidal morphology, with a large nucleus and an abundant cytoplasm with many ribosomes and prominent rough endoplasmic reticulum and Golgi apparatus, due to the large amount of proteins they synthesize. OBs are distributed along the bone surface in a monolayer and are connected among them by gap junctions (reviewed in Florencio-Silva *et al.*, 2015).

Osteoblasts are derived from mesenchymal stem cells (MSCs) from bone marrow in the presence of specific growth factors, transcription factors and hormones, such as runt-related transcription factor 2 (RUNX2), osterix (OSX), Distal-less homeobox 5 (DLX5), bone morphogenetic proteins (BMPs), Hedgehogs (HH), Wnt/ $\beta$ -catenin or NOTCH signalling proteins, estrogens or parathyroid hormone (PTH) (Figure 1). Once differentiated, osteoblasts may undergo apoptosis or become osteocytes or lining cells.



**Figure 1.** Osteoblast differentiation from MSCs. Molecular signals involved in the regulation of this process are shown. Modified from Arboleya & Castañeda, 2013.

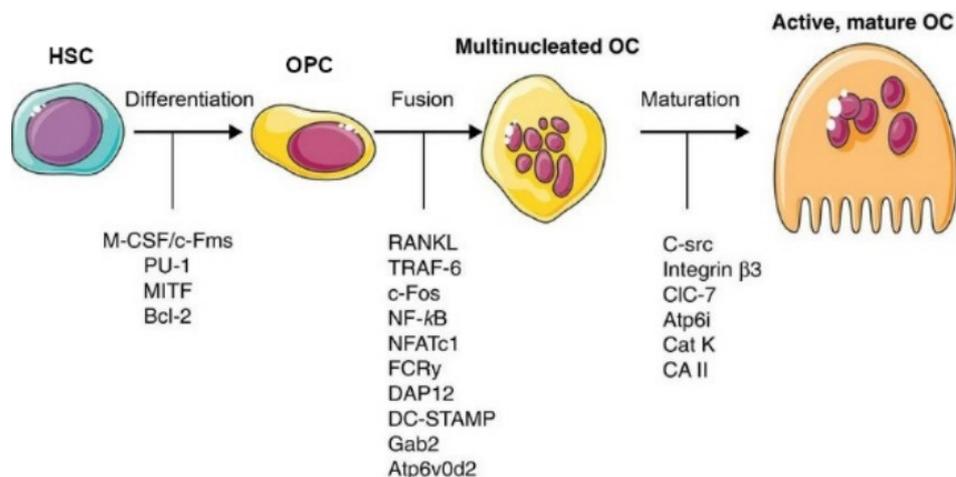
Osteocytes, the most abundant cell type in bone, are terminally differentiated osteoblasts that remain embedded in small *lacunae* of mineralized bone matrix. They have a flattened dendritic morphology, with small rough endoplasmic reticulum and Golgi apparatus, and they develop filipodial cellular processes that allow communication and small exchange of molecules with neighbour osteocytes and osteoblasts in the bone surface. OCys are responsible for bone matrix maintenance and bone remodelling control (revised in Prideaux *et al.*, 2016). They are also essential for mechanosensing and transduction into biochemical signals, leading to the response of bone to different mechanic stimuli (reviewed in Capulli *et al.*, 2014).

As OBs transition to OCys, many of the expressed OB markers, including collagen type I and alkaline phosphatase are downregulated, while OCy markers, such as matrix extracellular phosphoglycoprotein (MEPE), sclerostin (SOST), dentin matrix protein 1 (DMP-1) and phosphate regulating gene with homologies to endopeptidases on the X chromosome (PHEX) are upregulated.

Bone lining cells are flattened, elongated, quiescent cells that cover bone surfaces where neither bone resorption nor bone formation occurs. They have a metabolic function and the capacity of re-differentiate to osteoblasts (reviewed in Florencio-Silva *et al.*, 2015).

Osteoclasts are big multinucleated phagocytic cells that are responsible for bone resorption. They are generated by fusion of mononuclear osteoclast precursor cells (OPC) of the monocyte/macrophage lineage. These precursors are originated from

hematopoietic stem cells under the influence of several factors (Figure 2). Among them, the cytokines macrophage colony-stimulating factor (M-CSF), receptor activator of NF- $\kappa$ B ligand (RANKL), and interleukins (ILs).



**Figure 2.** Osteoclast differentiation from hematopoietic stem cells. Molecular signals involved in the regulation of this process are shown. HSC: hematopoietic stem cell; OPC: osteoclast precursor cell; OC: osteoclast. Modified from Arboleya & Castañeda, 2013.

Mature OCs are located in resorbed cavities, or Howship lacunae, and are activated by signals that promote a reorganization of the cytoskeleton and cellular adhesions. During bone resorption, osteoclasts are tightly bound to bone matrix surface through the sealing zone, generating a compartment beneath them where bone is resorbed. The ruffled border membrane of OCs is formed of microvilli that increase its surface where substance transport takes place. On the one side, there is a  $H^+$  and  $Cl^-$  flux through specific channels in order to acidify the extracellular bone matrix beneath. On the other site, there is a vesicle transport system where degrading enzymes, such as matrix metalloproteases (MMPs), tartrate resistant acid phosphatase (TRAP), and Cathepsin K (CTSK), are exocytosed and matrix degradation products, such as collagen fragments and minerals, are endocytosed and transcytosed to the basal membrane where they are secreted, contributing to calcium and phosphorous homeostasis (reviewed in Soysa & Alles, 2016).

### 1.3.2. Extracellular matrix

Bone extracellular matrix (ECM) is mainly synthesized by OBs and is formed by an organic fraction (30% of weight), that confers elasticity and flexibility, and a mineral fraction (70% of weight), which consists predominantly of calcium and phosphorous that

form hydroxyapatite crystals  $[Ca_{10}(PO_4)_6(OH)_2]$ . These crystals lay in between collagen fibrils, conferring resistance and stiffness to bone tissue (reviewed in Murshed, 2018).

Collagen type I is the principal component of the organic fraction of ECM. The rest includes proteoglycans and other non-collagenous proteins, such as osteocalcin (OCN), osteonectin, osteopontin (OPN), bone sialoprotein (BSP), and fibronectin. They are synthesized and secreted by osteoblasts and they exert multiple functions, including regulation of ECM mineralization and turnover and regulation of bone cell proliferation and activity (reviewed in Gentili & Cancedda, 2009; Florencio-Silva *et al.*, 2015).

#### **1.4. Bone formation and development**

During embryonic development, bones are formed through two major mechanisms: intramembranous ossification and endochondral ossification (Karaplis, 2008).

In intramembranous ossification, bones form directly from condensation of MSCs that differentiate to OBs. Osteoblasts secrete osteoid that slowly becomes mineralized. Flat bones, such as those of the cranium roof and mandible and maxillary are formed via intramembranous ossification.

Endochondral ossification involves the formation of a hyaline cartilage mould that is then substituted by bone tissue. Long bones, vertebrae, pelvic bones and bones of the base of the skull are formed by endochondral ossification. In this process, MSCs first differentiate to chondrocytes that generate a mould resembling the shape of the future bone. Later, chondrocytes become hypertrophic and undergo apoptosis, allowing the infiltration of blood vessels and OBs precursors that differentiate and ossify the structure. Endochondral ossification is the process by which long bones increase in their length throughout childhood and adolescence, since a thin layer of hyaline cartilage remains between the diaphysis and the epiphysis, known as the growth or epiphyseal plate (see Box 1). The cartilage is replaced by bone from one side, while it proliferates from the other side. In adulthood, once growth is complete, the epiphyseal plate becomes completely ossified.

Osteogenesis is tightly controlled at a molecular level by several growth factors and transcription factors, including RUNX2, vascular endothelial growth factor (VEGF), members of the homeobox (HOX), distal-less homeobox (DLX), Msh homeobox (MSX)

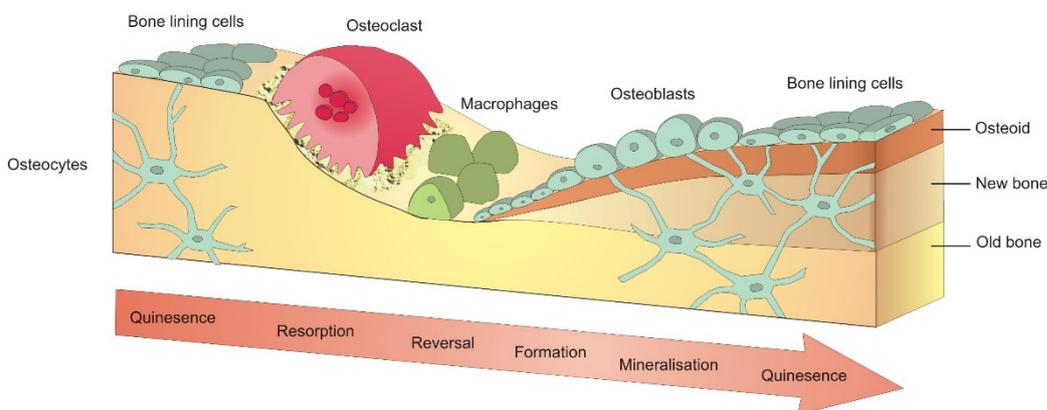
and paired-box (PAX) families, and fibroblast growth factors (FGFs), WNT, NOTCH, HH and BMP pathways (reviewed in Berendsen & Olsen, 2015; Runyan & Gabrick, 2017).

### 1.5. Bone remodelling and homeostasis

Bone remodelling or turnover is the cyclic process by which old bone is substituted by new bone in response to alterations in the physical activity, dietary calcium levels, hormonal changes, bone lesion and local paracrine signals within the bone microenvironment. It occurs throughout life and preserves the mineral homeostasis and the biomechanical properties of bone.

Bone remodelling is carried out in basic multicellular units, which include different cell types that are spatially and temporally coordinated. It consists of four steps: activation, resorption, reversal, and formation and mineralization.

Initially, pre-osteoclasts migrate and differentiate to mature osteoclasts that anchor to a bone surface free of lining cells. Osteoclasts begin to resorb bone by acidifying the ECM and releasing proteolytic enzymes, such as cathepsin K and MMPs. Thereafter, OCs undergo apoptosis and macrophages colonize the lacunae that are generated by bone resorption. Macrophages degrade the collagen remains and deposit proteoglycans to form a foundation line, that will cohesion the old bone with the new bone. They also release growth factors to stimulate MSC differentiation into OBs. Afterwards, OBs are situated on the foundation line and start to secrete new osteoid that will subsequently become mineralized (Figure 3).



**Figure 3.** The bone remodelling process and its phases. Extracted from <https://www.york.ac.uk/res/bonefromblood/background/boneremodelling.html>

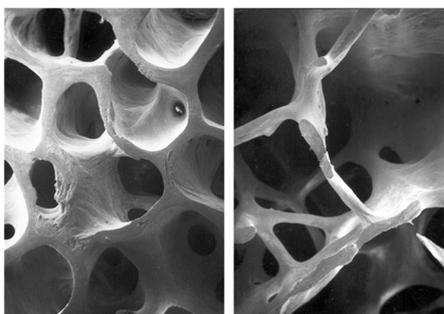
Bone remodelling is a tightly regulated process, both by local and systemic factors, and its imbalance leads to pathological situations, such as osteopetrosis, osteosclerosis or osteoporosis. The maintenance of bone homeostasis is largely dependent upon cellular communication between OCs and OBs, as well as the involvement of OCs (reviewed in Kenkre & Bassett, 2018; Kim & Koh, 2019).

Local coupling mechanisms include the RANK-RANKL-OPG system, the semaphorins, the ephrins and bone matrix-released molecules such as BMPs, IGFs and TGF- $\beta$ . Endocrine regulation of bone remodelling is mediated by the PTH, vitamin D, calcitonin, sex hormones, glucocorticoids (GCs), growth hormone and thyroid hormone.

## 2. OSTEOPOROSIS

### 2.1. Definition and classification

Osteoporosis is a systemic skeletal disorder characterized by low bone mass and microarchitectural deterioration of bone tissue (Figure 4), with a consequent decrease of bone strength and increase in bone fragility and fracture risk (NIH, 2001). It is a major worldwide public health concern.



**Figure 4.** Scanning electron micrographs of normal (left) and osteoporotic (right) bones. A loss of bone internal structure can be observed in osteoporotic bone. Extracted from Marx, 2004.

Osteoporotic or fragility fractures (OFs) are the major clinical outcome of osteoporosis. They occur mainly at the hip (proximal femur or femoral neck), vertebrae and distal forearm (Colles' fracture) and are associated with substantial morbidity, mortality, loss of independence and reduced quality of life (Sözen *et al.*, 2017). In many cases, osteoporosis is only diagnosed following a fragility fracture, since bone loss is asymptomatic. OFs can arise with minimal trauma, and clinical manifestations in patients are significant pain, disability and deformity (Eastell, 2017). Hip

fracture is the most serious OF, whereas vertebral fracture, the most prevalent, is often asymptomatic and usually does not require hospitalization (Schousboe, 2016).

Osteoporosis is etiologically classified as primary or secondary. Primary osteoporosis is the most common form and it can be divided into two subtypes. Type I primary osteoporosis, commonly known as postmenopausal osteoporosis, occurs as a consequence of a decrease in estrogen levels in postmenopausal women. It is characterized by a rapid bone loss and vertebral fracture is the most common outcome. Type II primary osteoporosis, or senile osteoporosis, affects both men and women older than 70-75 years and occurs as a consequence of hormonal (such as PTH) and metabolic changes associated with the normal process of ageing. Age-related bone loss results mainly in hip fractures (reviewed in Akkawi & Zmerly, 2018; Raisz, 2005).

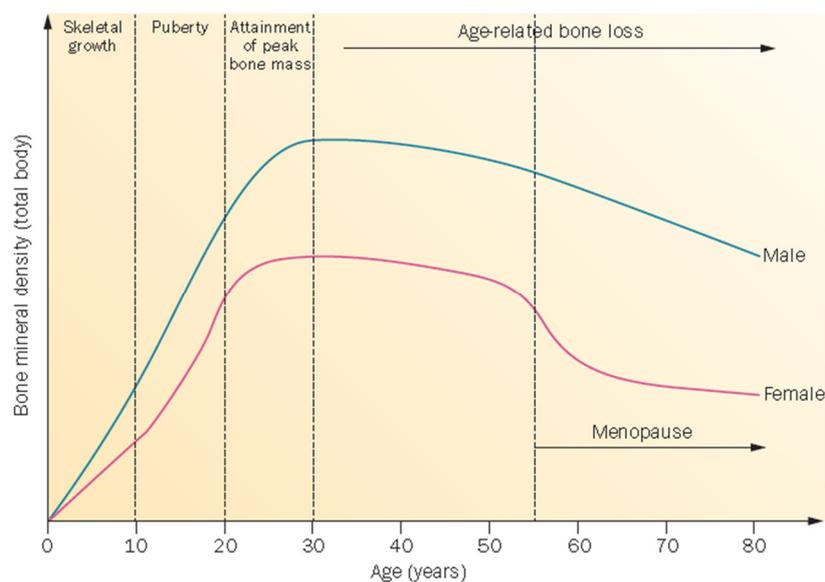
Secondary osteoporosis is caused by endocrine, rheumatic, hematologic, nephrologic or gastrointestinal pathological conditions that impair the normal development of bone mineral density or precipitate an excessive loss of bone mass, such as rheumatoid

arthritis, type I diabetes, hyperparathyroidism, hypogonadism, or intestinal calcium malabsorption (Emkey & Epstein, 2014). It can also ensue after pharmacological treatment (e.g. glucocorticoids) or extended periods of inactivity or immobilization (Alexandre & Vico, 2011; Briot & Roux, 2015). Glucocorticoid-induced osteoporosis is the most common secondary cause to osteoporosis.

The research presented in this thesis relates to postmenopausal osteoporosis.

## 2.2. Diagnosis: bone mineral density

Osteoporosis may be considered a consequence of an imbalanced bone remodelling, with a negative bone balance. Typically, in a healthy individual, bone balance is positive until the age of 25-30, when peak bone mass (PBM) is attained (Figure 5). Afterwards, bone mass gradually and asymptotically decreases throughout the lifetime. In women, PBM is lower than in men and markedly bone loss occurs in the first years after menopause due to a significant reduction of estrogen levels, that have a protective effect on bone (reviewed in Farr & Khosla, 2015; Hendrickx *et al.*, 2015).



**Figure 5.** Overview of BMD values during life in men and women. Extracted from Hendrickx *et al.*, 2015.

Bone mineral density (BMD) is used as a measure of bone mass. It is expressed in  $\text{g}/\text{cm}^2$  and it is measured by densitometric non-invasive techniques, such as dual-energy X-ray absorptiometry (DXA). BMD is the bone parameter clinically used for diagnosis of osteoporosis and fracture risk assessment, as well as for monitoring patients under

pharmacological treatment. Due to the strong influence of age, gender and ethnicity to BMD, two statistical parameters are used in the clinical practice, the T-score and the Z-score (Box 2). According to the World Health Organization (WHO), osteoporosis is defined as a T-score  $\leq -2.5$ , whereas osteopenia is defined as a T-score between  $-1$  and  $-2.5$  (Table 1) (World Health Organization, 1994).

### Box 2. The T-score and the Z-score

The T-score is the number of standard deviations (SDs) by which the BMD of an individual differs from the mean value observed in young healthy adults (25-30 years old) from the same gender and ethnicity.

The Z-score is the number of SDs by which the BMD of an individual differs from the mean value expected for the same age, gender and ethnicity (Kanis *et al.*, 2013).

**Table 1.** WHO's definition of osteoporosis

Diagnostic category	Criteria
Severe (or established) osteoporosis	T-score $\leq -2.5$ with one or more fractures
Osteoporosis	T-score $\leq -2.5$
Osteopenia	T-score $-1$ to $-2.5$
Normal	T-score $> -1.0$

In general, BMD is a good biomarker capturing intrinsic properties of bone biology. Nonetheless, its utility as a clinical indicator of osteoporosis is in some way limited and an increasing interest in bone quality has arisen. Hence, high-resolution non-invasive imaging techniques that are capable of assessing bone structure and strength have been developed, including magnetic resonance imaging (MRI), trabecular bone score (TBS) evaluation, quantitative ultrasound (QUS) and computed tomography (CT) (Link & Heilmeyer, 2016; Lorentzon & Cummings, 2015). Recently, microindentation was also introduced to measure bone material strength (Diez-Perez *et al.*, 2010). Quantification of biochemical bone turnover markers, including resorption markers, namely serum C-terminal cross-linked telopeptide of type-I collagen (s-CTX) and urinary N-terminal cross-linked telopeptide of type-I collagen (NTX), and formation markers, namely serum procollagen type-I N-terminal propeptide (s-PINP) and serum OCN, is also useful to determine the extent of bone deterioration and fracture risk and monitor treatment (Sözen *et al.*, 2017; Vasikaran *et al.*, 2011).

### **2.3. Epidemiology**

Osteoporosis is the most common bone metabolic disorder affecting around 200 million people all over the world and its prevalence will increase as life expectancy continues to rise and the population ages (Akkawi & Zmerly, 2018).

Osteoporosis is 3 times more prevalent in women than in men, due to their lower PBM and to their faster loss of bone mass, and it is more frequent in people of Caucasian ancestry. Nowadays, a 20-30% of Caucasian women over 50 years old and around 10% of Caucasian men over the age of 50 have osteoporosis, while the prevalence increases up to 50% in Caucasian women over 70 years old (Hernlund *et al.*, 2013; Wade *et al.*, 2014). Furthermore, about 30-50% of women and 15-30% of men aged 50 will sustain an OF throughout the remaining of their lives. It is estimated that OFs account for approximately 9 million fractures annually (Cauley, 2017; Morin *et al.*, 2013), which suppose a huge economic burden.

The incidence of OFs increases exponentially with age. Moreover, OF rates vary by geographic location, socioeconomic status, and ethnicity (Curtis *et al.*, 2016; Dhanwal *et al.*, 2011).

Hip fractures, although less common than other OF, account for the majority of mortality, morbidity and costs associated with osteoporosis (Compston *et al.*, 2019; Hernlund *et al.*, 2013). It has been determined that hip fractures are associated with an up to 36% excess mortality within 1 year, with a higher mortality in men than in women (Haentjens *et al.*, 2010). Notably, patients experiencing OFs are at considerable risk for subsequent OFs, with an even increased mortality risk (Bliuc *et al.*, 2009).

### **2.4. Risk factors**

Osteoporosis is a multifactorial or complex disease, in which both genetic and environmental factors play a fundamental role, as well as the interaction among them. The main risk factors for osteoporosis and OF are summarized in Table 2. Genetic susceptibility to osteoporosis will be discussed in section 3 of this introduction.

**Table 2.** Risk factors involved in osteoporosis and OF

<b>Risk factor</b>	<b>Comments</b>
Advanced age	It is one of the most important risk factors (Curtis <i>et al.</i> , 2016; Pouresmaeili <i>et al.</i> , 2018)
Female gender	It modulates both PBM acquisition and bone loss during postmenopause (Nieves, 2013)
Caucasian ethnicity	(Cauley, 2011; Lei <i>et al.</i> , 2006)
Early menopause	(Eastell, 2017; Gallagher, 2007)
Late menarche	(Guo <i>et al.</i> , 2005; Zhang <i>et al.</i> , 2018)
Bone geometry	Hip geometry measures, particularly longer hip axis length, increase susceptibility of hip fracture (Bouxsein & Karasik, 2006; Leslie <i>et al.</i> , 2015)
Family history (genetic factors)	It is an important risk factor, since osteoporosis and OFs have a high genetic load (Kanis <i>et al.</i> , 2004a; Compston <i>et al.</i> , 2017)
Previous low-trauma fractures	It is an important risk for future fractures (Johnell <i>et al.</i> , 2004; Kanis <i>et al.</i> , 2004b)
Low body mass index (BMI)	Especially thin build or small stature. A rapid weight loss is also correlated with a decrease in bone mass (Compston <i>et al.</i> , 2017; De Laet <i>et al.</i> , 2005)
Hormonal status	Sex-steroids are indispensable for PBM attainment and bone homeostasis (Pouresmaeili <i>et al.</i> , 2018; Riggs <i>et al.</i> , 2002)
Nutritional deficiency	Insufficient supply of calcium and vitamin D impairs bone formation and mineralization, and increases bone resorption (Bonjour <i>et al.</i> , 2013; Christodoulou <i>et al.</i> , 2013)
Low physical activity	Physical activity, especially weight-bearing exercise, stimulates osteocytes to trigger bone remodelling. Exercise reduces oxidative stress (Leeuwenburgh & Heinecke, 2001; Ozcivici <i>et al.</i> , 2010; Yuan <i>et al.</i> , 2016)
Cigarette smoking	Smoking is associated with a decrease in circulating vitamin D levels and an increase in serum PTH (Kanis <i>et al.</i> , 2005; Ward & Klesges, 2001)
Alcohol consumption	It affects bone mass in a dose-dependent manner (Cheraghi <i>et al.</i> , 2019; Ronis <i>et al.</i> , 2011)
Falls and factors that increase falling risk	For example, visual defects, dementia, muscular weakness, etc. (Compston <i>et al.</i> , 2017, 2019; Pouresmaeili <i>et al.</i> , 2018)
Medication	Some medication, such as GCs, aromatase inhibitors, anticonvulsants or antidepressants, are associated with secondary osteoporosis in a dose- and time-dependent manner (see 2.1; Emkey & Epstein, 2014)
Other diseases	see 2.1 (Emkey & Epstein, 2014)

Fracture risk assessment is largely based on BMD since OFs are highly related to low BMD values: for each SD decrease in BMD there is a 1.4- to 2.9-fold increase in fracture risk (Johnell *et al.*, 2005). However, in some cases, OFs occur in patients with BMD levels that do not fall within the osteoporotic range (Sornay-Rendu *et al.*, 2005; Unnanuntana *et al.*, 2010), due to other risk factors summarized in Table 2. In this regard,

the Fracture Risk Assessment tool (FRAX) was developed in order to predict the individual risk of OF on the basis of clinical settings, as well as BMD and bone turnover markers (Kanis *et al.*, 2007, 2017).

## 2.5. Prevention and treatment

The best treatment for osteoporosis is prevention. Prevention aims both at optimizing PBM and reducing bone loss rate. Thus, healthy lifestyle habits such as regular physical exercise, a balanced diet with an adequate calcium intake, sufficient sun exposure (essential to produce vitamin D), avoidance of smoking and reduction of alcohol consumption are fundamental (Compston *et al.*, 2017; Eastell, 2017).

Pharmacological treatment intends primarily to prevent OFs, as well as increase BMD levels and relieve osteoporosis symptoms. It can be divided in anti-resorptive and anabolic therapies (Table 3). Antiresorptive therapies are directed to inhibit osteoclastic bone resorption, whereas anabolic therapies aim at stimulating bone formation. Notably, not all pharmacological agents decrease the risk of OFs at all sites (Crandall *et al.*, 2014). Furthermore, antiresorptive therapies increase the degree and homogeneity of mineralization. Choice of drug should be based on site of diminished BMD and/or fracture, any secondary benefits, and contraindications. Supplementation with calcium and vitamin D is often advocated as an adjunct to other treatments.

The elucidation of the molecular mechanisms underlying the regulation of bone remodelling and osteoporosis pathogenesis has uncovered a number of new potential therapeutic targets for osteoporosis (Awasthi *et al.*, 2018). In an effort to specifically inhibit the resorption action of OCs, several cathepsin K inhibitors have been developed and clinically evaluated, the most promising one was odanacatib. Despite showing good results improving BMD levels and reducing OFs, they have not been pursued due to safety concerns, such as an increased risk of stroke (Duong *et al.*, 2016). MicroRNAs (miRNAs) are important regulators for bone homeostasis and have also emerged as promising targets for treating osteoporosis (Feng *et al.*, 2018). Another developing strategy consists in cell-based replacement therapy via the use of MSCs that can promote new bone formation by their differentiation into bone-forming cells or by acting in a paracrine manner through MSCs-derived exosomes (Li *et al.*, 2018; Phetfong *et al.*, 2016).

Table 3. Characteristics of osteoporosis therapies

Therapies	Mechanism of action	Efficacy <sup>a</sup>		Side effects	Comments	References
		BMD increase	Fracture risk reduction			
<b>Antiresorptive</b>						
<b>Bisphosphonates</b> (see 2.5.1)	Analogues of inorganic pyrophosphate with strong affinity for hydroxyapatite They inhibit OCs function and promote OCs apoptosis N-BPs inhibit FPPS	LS: 5.4-8.8% Hip: 3.1-5.9%	Vertebral: 40-70% Hip: 40-50% Non-vertebral: 25-50%	Gastrointestinal symptoms, hypocalcemia, musculoskeletal pain, flu-like symptoms Rare occurrence of ONJ and AFF	First-line therapy Low cost	Drake et al., 2008; Khan & Cheung, 2017; Russell et al., 2008
<b>Denosumab</b>	Human monoclonal antibody that binds to and inhibits RANKL, hindering OCs formation, function and survival	LS: 9.2% FN: 4.8%	Vertebral: 70% Hip: 40% Non-vertebral: 20%	Skin rash and infection, hypocalcemia, mild gastrointestinal symptoms, musculoskeletal pain Rare occurrence of ONJ and AFF	Upon cessation of therapy, rapid bone loss and loss of vertebral fracture protection occurs	Bone et al., 2011; Cummings et al., 2009, 2018; Reid et al., 2010
<b>Selective estrogen receptor modulators (SERMs)</b>	Ligands for the ER They can act as agonist or antagonist, depending on the target tissue Raloxifene inhibits OCs formation and activity, mimicking estrogen	LS: 2.7% FN: 2.4%	Vertebral: 40% Hip and non-vertebral: not proven	Leg cramps, hot flushes, edema and vasomotor symptoms Increased risk of thromboembolic events Substantially decreased risk of breast cancer	When discontinued, bone loss occurs fairly quickly No longer considered as first-line therapy	Chen et al., 2019; Ettinger et al., 1999; O'Connor, 2016; Qaseem et al., 2017
<b>Anabolic</b>						
<b>Teriparatide [rhPTH (1-34)]</b>	Intermittently administered, it enhances osteoblastogenesis through	LS: 9% FN: 3%	Vertebral: 70% Hip: no evidence Non-vertebral: 50%	Headache, mild upper gastrointestinal	Duration of therapy limited to 24 months because high PTH	Eastell & Walsh, 2017; Leder, 2017;

reducing OCy-derived SOST and increasing Wnt10b It decreases OB apoptosis	Amelioration of bone micro-architecture	symptoms, hypercalciuria, hypercalcemia Increased risk of osteosarcoma	signaling leads to bone resorption Very expensive	Neer et al., 2001
<b>Abaloparatide</b> Highly selective, high affinity PTHrP synthetic analogue The molecular mechanisms are still being defined	LS: 10.4% FN: 4% Amelioration of bone micro-architecture	Headache, dizziness, nausea, hypercalcemia, increased heart rate and palpitation	Approved by the American FDA in 2018; Leder, 2017; Liu et al., 2019; Miller et al., 2016	Bilezikian et al., 2018; Leder, 2017; Liu et al., 2019; Miller et al., 2016
<b>Romosozumab</b> Humanized antibody that inhibits OCy-derived SOST, preventing its inhibition on WNT signaling in OB and stimulating bone formation. Simultaneously, it inhibits RANKL release, suppressing OCs function and reducing bone resorption	LS: 11.3% FN: 3.7% Amelioration of bone micro-architecture	Constipation, back pain, headache, nasopharyngitis Increased cardiovascular events Rare occurrence of ONJ and AFF	Recently approved in USA and Japan	Bandeira & Bilezikian, 2017; Cosman et al., 2016, 2018; McClung et al., 2014; Ominsky et al., 2014

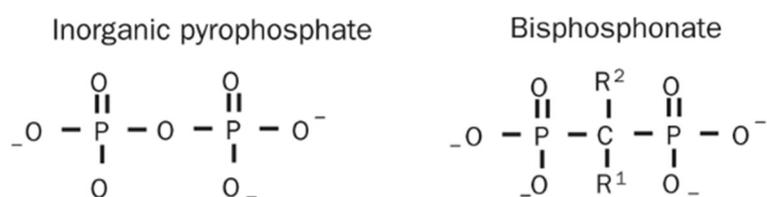
<sup>a</sup>The efficacy data are at 3 years for all antiresorptive agents, at 18 months for teriparatide and abaloparatide, and at 12 months for romosozumab  
AFF: atypical femoral fracture; BMD: bone mineral density; ER: estrogen receptor; FDA: food and drug administration; FN: femoral neck; FPPS: farnesyl pyrophosphate synthase; LS: lumbar spine; OB: osteoblast; OC: osteoclast; ONJ: osteonecrosis of the jaw; PTHrP: parathyroid hormone-related protein; rhPTH: recombinant human parathyroid hormone

The sequential use or the combination of anabolic and antiresorptive therapies have also been evaluated in an attempt to achieve higher bone mass and strength outcome than the resulted from monotherapy (McClung, 2017). In general, combination therapy shows no meaningful clinical benefit compared to monotherapy, the only exception being the simultaneous use of teriparatide and denosumab (Tsai *et al.*, 2013). Conversely, sequential therapy is often appropriate, specifically the use of an anabolic agent followed by an anti-resorptive drug. However, the skeletal responses differ depending upon skeletal site measurement, timing of administration and the specific sequence of drugs used (Cosman *et al.*, 2017; Shen *et al.*, 2017).

### 2.5.1. Bisphosphonates

Bisphosphonates (BPs) are cost-effective pharmacological agents that inhibit bone resorption and are used as first-line drugs for osteoporosis treatment, as well as in several rare bone diseases, such as Paget's disease of bone. Interestingly, BPs also possess antitumor and antiangiogenic properties, making them good candidates for cancer therapy (Giger *et al.*, 2013). Moreover, BPs are currently being explored for use in other non-skeletal applications, such as neurodegenerative diseases (Zameer *et al.*, 2018).

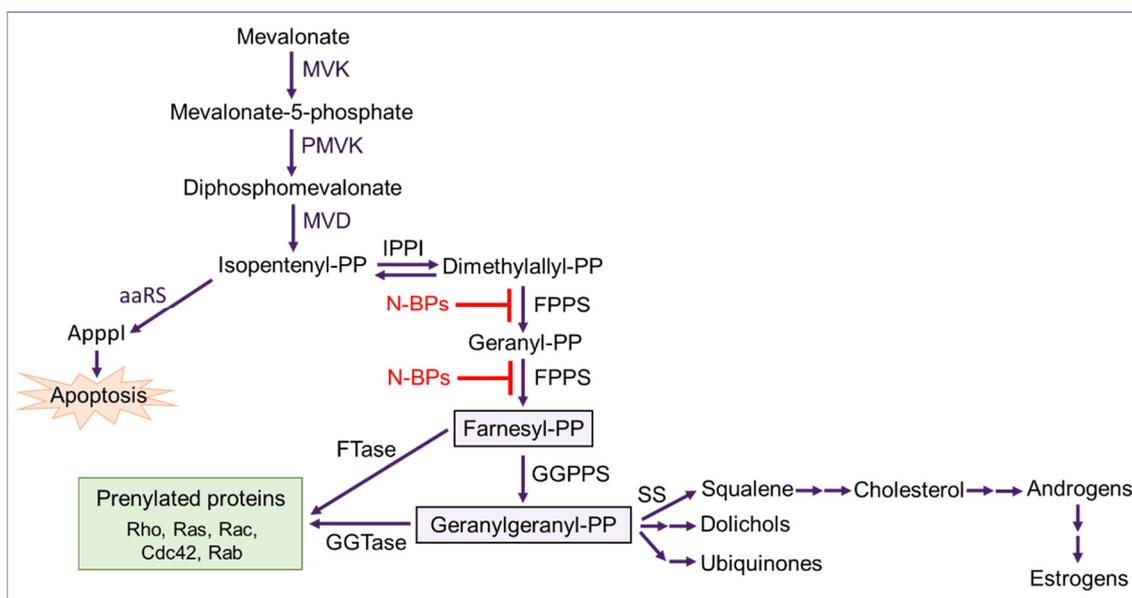
Chemically, BPs are stable synthetic analogues of naturally-occurring inorganic pyrophosphate (PP<sub>i</sub>; Figure 6). The P-C-P backbone structure is resistant to enzymatic and chemical hydrolysis and, therefore, BPs are not metabolized. The side chains (R<sup>1</sup> and R<sup>2</sup>) bound to the central C determine the binding affinity and the antiresorptive potency of each specific BP (Nancollas *et al.*, 2006; Russell *et al.*, 2008). Depending on the nature of the side chains, BPs can be classified in non-nitrogen containing BPs, no longer used, and nitrogenous bisphosphonates (N-BPs), including alendronate, ibandronate, risedronate and zoledronate, that are more effective.



**Figure 6.** Chemical structures of Inorganic pyrophosphate and bisphosphonates. Extracted from Drake *et al.*, 2008

BPs bind to hydroxyapatite with high affinity, conferring an extreme tissue selectivity, and remain attached to the mineralized bone for more than 5 years. They have multiple effects on hydroxyapatite, including the prevention of calcification, inhibition of aggregation of crystals and of hydroxyapatite dissolution (Russell *et al.*, 2008).

BPs are preferentially incorporated into sites of active bone remodelling. Thus, they come into close contact with osteoclasts that endocytose them. N-BPs inhibit the activity of farnesyl pyrophosphate synthase (FPPS), a key enzyme of the mevalonate pathway, by binding to its ligand pocket (Figure 7; (Kavanagh *et al.*, 2006b; van Beek *et al.*, 1999)). N-BPs are also able to inhibit other enzymes of this pathway, such as geranylgeranyl pyrophosphate synthase (GGPPS) and squalene synthase, albeit to a much lesser extent (Amin *et al.*, 1992; Kavanagh *et al.*, 2006a). The primary function of the mevalonate pathway is the production of cholesterol, as well as the synthesis of isoprenoid lipids, including farnesyl pyrophosphate (FPP) and geranylgeranyl pyrophosphate (GGPP), which are required for the post-translational modification (prenylation) of some proteins, such as small GTPases (Goldstein & Brown, 1990). Small GTPases (e.g. Rab, Rac, Ras, Rho and Cdc42) play central roles in the regulation of core osteoclast cellular activities including cell morphology, cytoskeletal arrangement, membrane ruffling, trafficking of vesicles, and apoptosis (Coxon & Rogers, 2003).



**Figure 7.** Mevalonate pathway with the inhibition by N-BPs in red.

MVK: mevalonate kinase; PMVK: phosphomevalonate kinase; MVD: mevalonate decarboxylase; IPPI: isopentenyl pyrophosphate isomerase; FPPS: farnesyl pyrophosphate synthase; GGPPS: geranylgeranyl pyrophosphate synthase; FTase: farnesyl transferase; GGTase: geranylgeranyl transferase; SS: squalene synthase; aaRS: aminoacyl tRNA synthetase.

FPPS catalyses the successive condensation of isopentenyl pyrophosphate with dimethylallyl pyrophosphate and geranyl pyrophosphate, generating farnesyl pyrophosphate (FPP; Figure 7). The inhibition of FPPS by N-BPs causes the cytosolic accumulation of non-prenylated small GTPases, that are not able to anchor to cellular membrane and to participate in protein-protein interactions to orchestrate bone resorption (Rogers *et al.*, 2011). It is also suggested that the antiresorptive activity may be mediated by the accumulation of small GTPases in their active state in the cytosol, that might cause the inappropriate activation of downstream signalling (Dunford *et al.*, 2006). In addition, isopentenyl pyrophosphate (IPP) also accumulates in the cytosol of OCs and reacts with adenosine monophosphate (AMP), generating Apppl, a cytotoxic molecule that triggers OCs' apoptosis (Mönkkönen *et al.*, 2006). N-BPs also prevent the generation of mature OCs, by suppressing the fusion of OC precursors due to the inhibition of GGPP biosynthesis (Tsubaki *et al.*, 2014). All in all, N-BPs alter OCs' gene expression program (Box 3).

It has been suggested that N-BPs have effects on OBs and OCs as well, limiting their apoptosis through connexin43 signalling (Bellido & Plotkin, 2011; Plotkin *et al.*, 2008), and enhance MSCs proliferation and initiation of osteoblastic differentiation (von Knoch *et al.*, 2005), producing changes in gene expression in these cell types (Box 3).

### **Box 3. Effects of N-BPs on gene expression**

#### *Osteoclasts*

Yuen *et al.* (2014) analysed the expression profile of human OCs treated with alendronate or risedronate during their differentiation. They developed a combined N-BPs gene signature, consisting of 6 up-regulated (such as *RGR*, *CAV3* or *ANGPTL3*) and 7-down-regulated genes (such as *CALD1*, *RUNX2*, *RGS6* and *COL14A1*), and assessed the N-BP-associated pathways. Among the enriched pathways they identified monoterpenoid and chondroitin sulfate biosynthesis, gap and tight junctions, SNARE interactions in vesicular transport, mTOR signalling pathway, metabolism of xenobiotics by cytochrome P450, apoptosis, and oxidative phosphorylation.

#### *Osteoblasts*

Wang & Stern (2011) studied the effects of different risedronate concentrations and duration of treatment on gene expression of a rat osteoblastic cell line. They found several genes related to cell differentiation, apoptosis, angiogenesis and metastasis whose expression was altered, including *Comp*, *Bmpr1a*, *Birc1b* and *Flt1* down-regulated and *Alpl*, *Cdk2*, *Col2a1*, *Col4a1*, *Ctsk*, *Faslg*, *Fgf2*, *Fos*, *Hk2*, *Jun*, *Pparg* and *Vegfa* up-regulated. Some other genes, such as

*Bmp2*, *Cdh1*, *Mmp10* and *Smad3*, were up- or down-regulated depending on risedronate concentration and/or duration of treatment.

#### Osteocytes

Bivi *et al.* (2009) studied the effects of alendronate and risedronate on gene expression of a murine osteocytic cell line. Among the gene ontology categories over-represented in differentially expressed genes, they identified zinc ion binding, transmembrane receptor protein Ser/Thr kinase signalling pathway, regulation of transcription, ATPase activity, and intracellular protein transport. Some genes found regulated by N-BPs were *Atp6v0b*, *Vps26b*, and *Ii17rc*.

#### MSCs

Ribeiro *et al.* (2014) showed that in human MSCs cultures, alendronate and zoledronate inhibited *VEGF* expression and up-regulated expression of osteogenic genes such as *ALPL*, *BMP-2*, *OPG* and *BGLAP* (osteocalcin).

The majority of N-BPs (e.g. alendronate, risedronate and ibandronate) are administered orally, while others (e.g. zoledronate) are administered intravenously. Oral N-BPs have a low intestinal absorption and hence, they should be taken after an overnight fast and 30-60 min before eating or drinking and without stretching out.

N-BPs have been demonstrated to reduce vertebral, hip and non-vertebral fracture incidence, although not all of them are effective at all body locations (Table 4). Furthermore, they are able to improve BMD at different skeletal sites and preserve bone microarchitecture (Russell *et al.*, 2008).

**Table 4.** Efficacy of N-BPs

N-BP	Fracture risk reduction (%)			BMD increase (%)		References
	Vertebral	Hip	Non-vertebral	Vertebral	Hip	
<b>Alendronate</b>	50%	51%	50%	8.8%	5.9%	(Liberman <i>et al.</i> , 1995)
<b>Risedronate</b>	41%	40%	36%	5.4%	3.1%	(Harris <i>et al.</i> , 1999)
<b>Ibandronate</b>	62%	NS	Only in high-risk population	6.5%	3.4%	(Chesnut III <i>et al.</i> , 2004)
<b>Zoledronate</b>	70%	41%	25%	6.7%	5.1%	(Black <i>et al.</i> , 2007)

\*All the data refers to 3-years treatment

NS: not significant

Nonetheless, N-BPs have some secondary adverse effects. The commonest include mild gastrointestinal symptoms, such as esophageal inflammation and gastric ulcers, especially in orally administered N-BPs, and acute-phase reaction with flu-like symptoms in intravenously administered N-BPs. Uncommon side effects include musculoskeletal pain, headache, hypocalcaemia and ocular effects. Very rarely, osteonecrosis of the jaw (ONJ) and atypical femoral fractures (AFFs; see section 4 of this Introduction) can occur after long-term use of N-BPs (Khan & Cheung, 2017; Reyes *et al.*, 2016).

N-BPs efficacy has been shown with up to 10 years of use (Bone *et al.*, 2004). However, given that prolonged use of N-BPs may lead to adverse events, some recommendations suggest considering a drug holiday after 3-5 years of treatment in individuals who are not at high risk of fracture (Adler *et al.*, 2016; Compston *et al.*, 2017). Due to the long half-lives of N-BPs in bone, their antiresorptive effect will persist for some time after discontinuation of treatment while reducing the risk of secondary adverse effects, such as AFF. Nevertheless, BMD levels and fracture risk should be reassessed some time after withdrawal and subsequent recommencement of treatment, reconsidered (Gatti *et al.*, 2015).

### 3. GENETICS OF OSTEOPOROSIS

Osteoporosis is a complex disease with a multifactorial etiology, in which both genetic and environmental factors, as well as their interplay, determine the phenotype. One of the most important risk factors for osteoporosis is a positive family history, which emphasizes the crucial role of genetics in the pathogenesis of the disease. Like in many other common diseases, many genes, most of which of small effect, contribute to the overall phenotype (reviewed in Clark & Duncan, 2015; Ralston & Uitterlinden, 2010).

#### 3.1. Heritability of bone properties

Heritability ( $h^2$ ) is the proportion of variance of a trait due to genetic variation and can be estimated in twin and family studies. As said, BMD is the bone parameter clinically used for the diagnosis of osteoporosis and fracture risk assessment. It is a highly heritable trait with an estimated heritability of 50-85%, as reviewed by Boudin *et al.* (2016). BMD heritability varies depending on ethnicity, gender and skeletal site studied, possibly reflecting different relative contributions of genetic and environmental influences. For example, males tend to have higher levels of heritability than females. Regarding skeletal site, spine BMD has a higher heritability than femoral BMD but the greatest degree of BMD heritability was found at the head ( $h^2 > 95\%$ ) (Tse *et al.*, 2009).

Likewise, other bone parameters have been shown to be highly heritable: bone geometry,  $h^2=30-70\%$  (Demissie *et al.*, 2007); bone turnover markers,  $h^2=30-75\%$  (Hunter *et al.*, 2001); bone ultrasound measures,  $h^2=40-50\%$  (Arden *et al.*, 1996); and measures of bone microarchitecture,  $h^2=20-80\%$  (Karasik *et al.*, 2017). OF has a heritability of 54-68% in peri-menopausal women but rapidly decreases with age, being around 3% after 79 years old (Richards *et al.*, 2012). Again, heritability depends on skeletal site and type of fracture, being higher for hip fractures than for wrist fractures. These data show that, as reviewed in section 2.4 of this Introduction, there are other factors besides BMD that may influence OF, such as propensity to falls.

Most of the genetic studies for osteoporosis have been based on BMD. However, other bone parameters, such as microarchitecture measures or hip geometry have been used.

### 3.2. Linkage analyses

Linkage analyses, based on the co-segregation of genetic markers (traditionally microsatellites and, afterwards, single nucleotide polymorphisms [SNPs]) with the disease within a family, have been successfully used to map *loci* and genes involved in a large number of Mendelian diseases, including bone monogenic disorders (e.g. osteopetrosis, high bone mass (HBM), osteoporosis pseudoglioma syndrome (OPPG), osteogenesis imperfecta (OI) and sclerosteosis; reviewed in Alonso & Ralston, 2014).

In addition, non-parametric linkage analyses, in which it is not required that the model of inheritance of the disease is defined, have been also used in complex diseases such as osteoporosis.

Some *loci* related to BMD have been identified by non-parametric linkage analysis but few have reached significance and there has been limited replication among studies (Alonso & Ralston, 2014; Ioannidis *et al.*, 2007). The main reason is probably lack of power, since a huge number of families would be needed to detect the likely small effects of each individual quantitative trait *locus* (QTLs) affecting BMD. Some examples are 1p36, 1q21-23 or 20p12 (Ioannidis *et al.*, 2007; Styrkarsdottir *et al.*, 2003; Zhang *et al.*, 2009). In these *loci*, 2 genes with SNPs significantly associated with BMD and/or OF were identified: *RERE*, a gene that encodes a protein of the atrophin family which, when overexpressed, triggers apoptosis, in 1p36 (Zhang *et al.*, 2009) and *BMP2*, encoding Bone Morphogenetic Protein 2, a member of the TGF $\beta$  superfamily known to play a role in bone and cartilage development, in 20p12.3 (Styrkarsdottir *et al.*, 2003). This study showed evidence to suggest that a rare protein coding variant in *BMP2* was linked to and associated with osteoporosis in North-Europeans. However, this association could not be replicated in other populations, suggesting that the association observed might be specific to this population (Medici *et al.*, 2006; Richards *et al.*, 2009).

The linkage approach highlighted gender-specific, age-specific and skeletal site-specific effects in genetic *loci* related to osteoporosis.

### 3.3. Genetic association studies

Association studies have been widely used in the genetics of complex diseases. They consist on detecting a statistical correlation between genetic markers and a quantitative or qualitative trait related to the phenotype of interest (Cordell & Clayton, 2005).

Depending on the sample of individuals used, association analyses can be classified as family-based or unrelated population-based. Among the latter, case-control studies compare the genotypic or allelic frequencies of a genetic marker between a group of affected individuals and a group of healthy individuals from the same population. They are used for binary or categorical traits, such as OF. For quantitative traits, the mean value of the quantitative variable is compared among the groups of individuals bearing each of the different genotypes (Simundic, 2010). In osteoporosis studies, BMD in different skeletal sites is the quantitative trait most frequently used.

Association studies are relatively easy to perform and useful to detect small effects from the selected variants. The main limitation is small sample size, which can lead to false negative results or false positive findings that cannot be replicated (Ioannidis, 2005). In order to address this issue, large scale association studies have been performed by large consortia which allowed the detection of risk alleles with modest effect size. In the field of osteoporosis, the Genetic Markers for Osteoporosis (GENOMOS) and the Genetic Factors for Osteoporosis (GEFOS) consortiums were created.

In genetic association studies, two different approaches can be distinguished: candidate genes association studies or genome-wide association studies (GWASs).

### **3.3.1. Candidate genes association studies**

In candidate genes association studies, the analysed polymorphisms are located near or within candidate genes suggested to be relevant for the disease of interest either because there is prior knowledge of their biological function and phenotype when mutated (functional candidate genes; e.g. genes mutated in monogenic bone disorders), because they are located within an area identified by linkage analysis (positional candidate genes) or because they show a change of expression levels (expressional candidate genes).

The identification of osteoporosis susceptibility genes by candidate gene association studies began in the early 90s by Morrison *et al.*, (1994), who described associations between polymorphisms in the vitamin D receptor gene (*VDR*) and BMD. Afterwards, approximately 200 candidate genes have been explored for their potential association with BMD or fractures (Yuan *et al.*, 2019). However, many of the studies were

inadequately powered and resulted in conflicting and frequently irreproducible results (Ioannidis, 2005).

Large-scale candidate genes studies involving 20,000-45,000 individuals were performed by the GENOMOS Consortium, showing consistent evidence for association of some historical candidate osteoporosis genes with BMD and/or fracture risk: *ESR1* (Ioannidis *et al.*, 2004), *COL1A1* (Ralston *et al.*, 2006), and *LRP5* (van Meurs *et al.*, 2008). The strongest and most significant associations were observed in *LRP5* variants, achieving  $p < 5 \times 10^{-8}$ . No significant association was observed for *TGF- $\beta$*  (Langdahl *et al.*, 2008) or *VDR* (Uitterlinden *et al.*, 2006). A large collaborative meta-analysis was performed using data from 19,000 subjects and 36,000 SNPs within 150 candidate genes chosen based on at least one previous study of this gene in osteoporosis (Richards *et al.*, 2009). Several SNPs from 9 genes (*ESR1*, *LRP4*, *LRP5*, *ITGA1*, *SOST*, *SPP1*, *TNFRSF11A* [RANK], *TNFRSF11B* [OPG], and *TNFSF11* [RANKL]) showed robust evidence of association with BMD at either the femoral neck (FN) or lumbar spine (LS), with SNPs from 4 genes (*LRP5*, *SOST*, *SPP1* and *TNFRSF11A*) significantly associated with fracture risk.

### 3.3.2. Genome-wide association studies

Genome-wide association studies (GWASs) are unbiased hypothesis-free approaches that explore up to millions of polymorphic genetic markers (generally, SNPs) distributed evenly across the genome in thousands of individuals, thanks to advances in high-throughput genomic technologies and the availability of large biobank studies. GWAS allow the identification of novel genes and pathways related to the phenotype of interest. Large sample sizes are required due to the high number of statistical tests carried out for all the variants assessed and to achieve a sufficient statistical power to detect associations of small-effect. Meta-analyses can be applied to maximize statistical power and obtain more accurate estimations of the effect size of individual genetic variants (Duncan & Brown, 2013; Visscher *et al.*, 2017; Wu *et al.*, 2013).

The first GWAS in the osteoporosis field was carried out by Kiel *et al.* in 2007. In this study, around 71,000 autosomal SNPs were assessed in 1,141 individuals and tested against a variety of phenotypes including BMD (FN-BMD, LS-BMD and trochanter BMD [TR-BMD]), fracture risk and QUS of the calcaneus. Despite finding nominal association with several SNPs in genes such as *MTHFR*, *ESR1*, *LRP5* and *COL1A1*, none of them achieved genome-wide significance ( $p < 5 \times 10^{-8}$ ) due to the small number of genotyped

SNPs and lack of power because of a limited sample size. In 2008, two simultaneously published GWASs identified some *loci* (i.e. *LRP5*, *TNFRSF11B*, *ESR1*, *TNFSF11*, *ZBTB40* and the major histocompatibility complex [MHC] *loci*) with SNPs associated with BMD at genome-wide significant level in the general population (Richards *et al.*, 2008; Styrkársdóttir *et al.*, 2008). The *LRP5*, *ZBTB40*, *TNFRSF11B* and MHC *loci* were also associated with OF, as well as the *SPTBN1*, *LRP4* and *TNFRSF11A loci*.

From then on, more than 40 GWAS and meta-analyses have been carried out and more than 500 candidate genes showing association to different bone-related traits have been identified (Table 5).

The first large-scale GWAS meta-analysis was conducted by the GEFOS consortium in around 19,000 individuals of 5 Northern European populations (Rivadeneira *et al.*, 2009). Thirteen novel *loci* associated with BMD containing 15 candidate osteoporosis susceptibility genes were identified. Moreover, they confirmed the association of 7 previously identified *loci*, although *SOST*, *MARK3* and MHC *loci* failed to achieve genome-wide significance. A second larger multi-ethnic GEFOS meta-analysis, with a total of 83,894 individuals from 17 GWASs, identified 56 *loci* including 32 additional novel *loci* that reached genome-wide significance with either LS-BMD, FN-BMD or both (Estrada *et al.*, 2012). Furthermore, 14 of the BMD-associated *loci* were also found significantly associated with OF, of which 6 reached  $p < 5 \times 10^{-8}$  (*FAM210A*, *SLC25A13*, *LRP5*, *MEPE*, *SPTBN1* and *DKK1*). Notably, no marker in genes of the RANK-RANKL-OPG pathway was found associated with fracture risk. Interestingly, it was the first study to examine the X chromosome in order to identify sex-specific effects.

An alternative approach to conventional DXA-measured BMD has been used in some studies: the measure of heel bone properties through QUS. Moayyeri *et al.* (2014) and Mullin *et al.* (2017) performed GWAS meta-analysis to assess the genetic determinants of heel broadband ultrasound attenuation (BUA) and velocity of sound (VOS) and described genome-wide significant associations in previously reported *loci* for DXA-BMD, as well as in 4 novel *loci* (*TMEM135*, *PPP1R3B*, *LOC387810*, *SEPT5/TBX1*). More recently, 3 studies have performed GWASs of estimated BMD (eBMD) from heel QUS in UK Biobank individuals. Kemp *et al.* (2017) identified 307 conditionally independent SNPs at 203 *loci* associated with eBMD, of which 153 were not reported previously. In 2018, Kim identified 1,362 independent SNPs clustered into 899 *loci* with a genome-wide significant association to eBMD. Of the 899 *loci*, 613 were novel. In 2019, Morris *et al.* published an study identifying 518 genome-wide significant *loci*, of which 301 were not previously described. In addition, they reported 13 *loci* associated with OF.

**Table 5.** GWASs studies for different bone characteristics and osteoporotic fractures

<b>Study</b>	<b>Traits</b>	<b>Sample size</b>	<b>N° GWS loci</b>	<b>N° novel loci<sup>a</sup></b>	<b>Key genes<sup>b</sup></b>
Richards et al. (2008)	LS-BMD, FN-BMD, OF	8,557	2	1	<i>LRP5, TNFRSF11B (OPG)</i>
Styrkarsdottir et al. (2008)	LS-BMD, FN-BMD, OF	13,786	8	6	<i>TNFRSF11B, TNFSF11 (RANKL), ESR1, ZBTB40, MHC locus, SPTBN1, LRP4, TNFRSF11A (RANK)</i>
Yang et al. (2008) CNV GWAS	HF	1,499	1	1	<i>UGT2B17</i>
Styrkarsdottir et al. (2009)	LS-BMD, FN-BMD, OF	15,375	9	4	<i>MARK3, SOST, SP7 (osterix), TNFRSF11A<sup>c</sup>, MHC locus, ZBTB40, TNFRSF11B, TNFSF11, ESR1/CCDC170</i>
Xiong et al. (2009)	LS-BMD, hip BMD, HF	9,828	2	2	<i>ADAMTS18, TGFBFR3</i>
Timpson et al. (2009)	TB-BMD	5,275 children	1	0	<i>SP7</i>
Rivadeneira et al. (2009) Meta-analysis	LS-BMD, FN-BMD	19,195	20	13	<i>WLS, SPTBN1, CTNBN1, MEPE, STARD3NL, FLJ42280, LRP4, DCDC5, SOX6, FOXL1, HDAC5, CRHR1, MEF2C</i>
Guo et al. (2010)	HF, TH-BMD	11,568	1	1	<i>ALDH7A1</i>
Zhao et al. (2010)	CT, BR, HF	5,293	1	1	<i>RTP3</i>
Kung et al. (2010) Meta-analysis	LS-BMD, FN-BMD, OF	18,898	1	1	<i>JAG1</i>
Deng et al. (2010) CNV GWAS	LS-BMD, FN-BMD, TH-BMD, CT	1,000	1	1	<i>VPS13B</i>
Hsu et al. (2010) Meta-analysis	LS-BMD, FN-BMD, FNSA, NNW, FNL	11,290	4	3	<i>TNFRSF11B, RAP1A, TBC1D8, OSBPL1A</i>
Paternoster et al. (2010)	cor-vBMD	5,789	1	0	<i>TNFSF11</i>
Kou et al. (2011)	LS-BMD, FN-BMD	6,953	1	1	<i>FONG</i>
Duncan et al. (2011)	Extreme high or low TH-BMD	21,789	23	2	<i>GALNT3, RSPO3, LRP4/ARHGAP1, MARK3, SOX6, MEPE, FAM3C, SOST, JAG1, FOXL1, WLS, HDAC5, FLJ42280</i>

Estrada et al. (2012) Meta-analysis	LS-BMD, FN-BMD, OF	83,894	56	32	<b>CDKAL1/SOX4, CPED1, WNT16, MBL2/DKK1, AXIN1, RPS6KA5, ERC1/WNT5B, FAM210A, FAM9B/KAL1, SOX9, KLHDC5/PTHLH, IDUA, NTAN1, SFRP4, SUPT3H/RUNX2</b>
Medina-Gomez et al. (2012) Meta-analysis	TB-BMD, skull BMD	13,712 Children	1	0	WNT16/FAM3C/CPED1
Zheng et al. (2012) Meta-analysis	CT, FA-BMD, FAF	5,878	2	0	WNT16/FAM3C, TNFSF11
Paternoster et al. (2013) Meta-analysis	cor-vBMD, trab-vBMD, CP, TBF, OF	6,930	5	2	TNFSF11, <b>LOC285735</b> , TNFRSF11B, ESR1/CCDC170, <b>FMN2/GREM2</b>
Koller et al. (2013) Meta-analysis	LS-BMD, FN-BMD	9,658	2	0	WNT16, ESR1/CCDC170
Hwang et al. (2013) Meta-analysis	OF	4,563	1	1	MECOM
Styrkarsdottir et al. (2013) WGS-based GWAS	LS-BMD, hip BMD, TB-BMD, OF	73,965	2	1	AKAP11, <b>LGR4</b>
Zheng et al. (2013) Meta-analysis	DR-BMD	6,584	2	0	MEF2C, WNT16
Oei et al. (2014a) Meta-analysis	Radiographic VF	2,995	1	1	<b>16q24</b>
Oei et al. (2014b) CNV GWAS meta-analysis	OF	5,178	1	1	<b>6p25.1</b>
Zhang et al. (2014) Meta-analysis	FN-BMD, LS-BMD, TH-BMD	27,061	15	2	<b>SMOC1, CLDN14, ZBTB40, FGFR1, MEPE, MEF2C, WLS, CCDC170/ESR1, FLJ42280/SHFM1, FAM3C/WNT16, TNFRSF11B, SOX6, LRP5, AKAP11, FOXL1</b>
Moayyeri et al. (2014) Meta-analysis	BUA, VOS, heel BMD	58,878	7	1	<b>TMEM135</b> , ESR1, SPTBN1, RSPO3, WNT16, DKK1/MBL2, GPATCH1
Kemp et al. (2014) Meta-analysis	TB-BMD, skull BMD, UL-BMD, LL-BMD	9,395 children	13	1	CPED1/WNT16/FAM3C, <b>RIN3</b> , WNT4, TNFRSF11A, LINTC, TNFSF11, GALNT3, FUBP3, KLHDC5, CENPW, COLEC10, PPP6R3, EYA4

Chesi et al. (2015)	DR-BMD, BMC	1,885 children	2	1	CPED1, <b>MIR31HG/MTAP</b>
Zheng et al. (2015)	FN-BMD, LS-BMD, FA-BMD, OF	508,253	36	1	<b>EN1, CPED1, FAM3C, WNT4/ZBTB40, WLS, RHPN2, JAG1, CSRN3/GALNT3, CTNNA1, MEF2C, CCDC170, SHFM1, COLEC10, AXIN1, SFRP4, SOX6, TNFRSF11A, HOXC10</b>
Styrkarsdottir et al. (2016a)	LS-BMD, hip BMD, OF	209,379	3	1	<b>COL1A2, LGR4, TNFSF11</b>
WGS-based GWAS					
Styrkarsdottir et al. (2016b)	LS-BMD, hip BMD, OF	274,941	14	1	<b>PTCH1, EN1, RSPO3, AXIN1, SOST</b>
WGS-based GWAS					
Mullin et al. (2016)	LS-BMD	6,696	1	0	WLS
Meta-analysis					
Pei et al. (2016a)	WT-BMD, hTR-BMD, hTR-BMD, FN-BMD, LS-BMD, FA-BMD	47,052	5	2	<b>FMN2, NAB1, MEF2C, ESR1, SHFM1</b>
Meta-analysis					
Pei et al. (2016b)	hTR-BMD, hTR-BMD	9,174	3	1	<b>RP11-384F7.1, CTNNA1, TNFRSF11B</b>
Nielson et al. (2016)	LS-vBMD, clinical VF radiographic VF	38,717	5	1	WNT4/ZBTB40, TNFRSF11B, AKAP11, FMN2/GREM2, <b>SLC1A3/RANBP3L</b>
Meta-analysis					
Chesi et al. (2017)	DR-BMD, LS-BMD, FN-BMD, TH-BMD	1,419 children	5	4	CPED1, <b>IZUMO3, RBFOX1, SPBT, TBPL2</b>
Mullin et al. (2017)	BUA, VOS, OF	16,627	8	3	<b>PPP1R3B, LOC387810, SEPT5/TBX1, SPTBN1, RSPO3, CCDC170/ESR1, WNT16, TMEM135</b>
Meta-analysis					
Kemp et al. (2017)	Heel eBMD, OF	142,487	203	153	<b>ARID1A, PKN2, TBX15, NGEF, SUSDS5, ERC2, BMP2, PLXDC2, BMP5, MEOX2, CREB5, AQP1, CADM1, EMP1, NFATC1, TMEM92, GPC6, BMP4, SMAD3, BMPR2, AXIN2, SLC8A1, PLCL1, SMAD9, ADAMTS5, TOM1L2, TCF7L1, APC, DUSP5, CD44, CCND1, CYP19A1, MAFB, RUNX1, RAI1, ZSCAN25, GRB10, DRG2, ETS2, PSMD13, CSF1</b>
Meta-analysis					
Medina-Gomez et al. (2018)	TB-BMD	66,628 Life-course	80	36	
Meta-analysis					
Alonso et al. (2018)	Clinical VF	10,683	1	1	2q13

Pei et al. (2018) Joint study of meta-analysis	LS-BMD, FN-BMD	40,449	9	2	<b>MACROD2, OSBPL2, ARHGAP25, ZSCAN25, TMCO5A, PLEKHA1, TMEM135, LOC338758, SMAD3</b>
Kim (2018)	Heel eBMD, OF	394,929	899	613	<b>WNT1, RSPO3, ESR1, SPTBN1, MEPE, BMP4, RP1L1, HDAC5, PRSS55, IDUA, MAPT, GPATCH1, SMG6</b>
Trajanoska et al. (2018)	OF	562,258	15	4	<b>STARD3NL, RSPO3, ESR1, WNT16/CPED1, FUBP3, ETS2, C7ORF76/SHFM1, MBL2/DKK1, GRB10/COBL, SPTBN1, RPS6KA5, SOST, CTNNB1, FAM210A/RNMT, LRP5</b>
Gregson et al. (2018)	Extreme high or low TH-BMD and LS-BMD	2,195	4	2	<b>NPR3, SPON1, MEF2C, WNT4</b>
Baird et al. (2019) Meta-analysis	DXA-derived hip shape modes	15,934	8	6	<b>ASTN2, SOX9, PTHLH, RUNX1, NKX3-2, FGFR4, GSC/DICER1, HHIP</b>
Morris et al. (2019)	Heel eBMD, OF	426,824	518	301	<b>DAAM2, COL11A1, SERPINC1, SEMA3D, PRKCE, HDAC4, HOXD11, BCL11A, SOX5, TGFB3, MMP16, IRS1, EPHA4, MSH6, SEPT11, LRRRC1, ADH1B, CTPS1, DNMT3A, MEIS1</b>
Hsu et al. (2019) Meta-analysis	FNL, FNSA, NNW, FNSM, BMD, OF	27,053	4	1	<b>IRX1/ADAMTS16, FGFR4, CCDC91, PPP6R3/LRP5</b>
Styrkarsdottir et al. (2019)	Hip BA, LS-BA, BMD, OF	50,336	13	6	<b>GDF5, ADAMTSL3, BCKDHB, COL11A1, MIR196A2/HOXC, SOX9, HHIP, CHRDL2, WNT4, DYM, ERC2, TBX4, CTDSP2</b>

<sup>a</sup>In relation with the studied phenotype

<sup>b</sup>In boldface, novel reported genes. Due to the large numbers of genes identified, a curated list of novel genes or functional validated genes are presented

<sup>c</sup>In Styrkarsdottir et al. (2008) TNFRSF11A was associated with OF, whereas in Styrkarsdottir et al. (2009) it was associated with BMD

BA: bone area; BMC: bone mineral content; BMD: bone mineral density; BR: buckling ratio; BUA: broadband ultrasound attenuation; cor-vBMD: cortical volumetric BMD; CP: cortical porosity; CT: cortical thickness; DR-BMD: distal radius BMD; DXA: dual energy X-ray absorptiometry; eBMD: estimated BMD; FAF: forearm fracture; FA-BMD: forearm BMD; FN-BMD: femoral neck BMD; FNL: femoral neck length; FNSA: femoral neck-shaft angle; FNSM: femoral neck section modulus; HF: hip fracture; hTR-BMD: hip trochanteric BMD; hTR-BMD: hip intertrochanteric BMD; LS-BA: lumbar spine bone area; LS-BMD: lumbar spine BMD; LS-vBMD: lumbar spine volumetric BMD; NNW: narrow neck width; OF: osteoporotic fracture; TB-BMD: total body BMD; TBF: trabecular bone fraction; TH-BMD: total hip BMD; trab-vBMD: trabecular volumetric BMD; UL-BMD: upper limb BMD; VF: vertebral fracture; VOS: velocity of sound; WT-BMD: Ward's triangle BMD

Furthermore, several additional GWAS for other bone parameters have been performed. Paternoster *et al.* published two studies (2010, 2013) in which they carried out GWASs of cortical and trabecular volumetric BMD as measured by peripheral quantitative computed tomography (pQCT). They replicated previously described associated *loci* in cortical BMD, including *TNFSF11*, which was also found to be associated with cortical porosity, and identified a novel bone-related *locus* (*FMN2/GREM2*). In addition, Zheng *et al.* (2012), showed association of *WNT16* with cortical bone thickness. GWAS studying hip structure parameters, such as femoral neck-shaft angle (FNSA), femoral neck length (FNL), femoral neck section modulus (FNSM) and narrow neck width (NNW), have also been carried out. For instance, Hsu *et al.* (2010, 2019) and Baird *et al.* (2019) identified 15 *loci* associated with hip geometry. Notably, the results of Hsu *et al.* (2019) showed an overlap with BMD in several signals, including *LRP5*.

Regarding OFs, most of the associations have been found by testing known GWAS BMD *loci*. However, some case-control GWASs have been carried out to elucidate the genetic determinants of OFs. Guo *et al.* (2010) performed the first GWAS for non-vertebral OFs in Chinese Han subjects and reported one associated *locus* containing the *ALDH7A1* gene. In 2018, Alonso *et al.* conducted a GWAS meta-analysis, in which they found that the SNP rs10190845 on chromosome 2q13 was genome-side significantly associated with clinical vertebral fractures with a large effect size. This *locus* had never been associated with OFs or BMD before, suggesting that the underlying mechanisms for this association might be independent of BMD. In the same year, Trajanoska *et al.* (2018) performed the largest GWAS on OFs at any skeletal site to date, comprising ~38,000 cases and ~227,000 controls. They identified 15 *loci* associated with OFs with modest effects, all of them being known BMD-associated *loci*, reinforcing the relationship between BMD and OF risk.

The majority of GWASs have tested common variants (minor allele frequency [MAF]≥5%) and the identified variants collectively explain a small proportion of the genetic variance of bone-related phenotypes. In this regard, some efforts have been done to identify low frequency or rare variants that might have greater effects. One of the explored approaches is to focus on individuals with extreme BMD. Gregson *et al.* (2018) performed the most comprehensive extreme BMD study to date and reported 2 new *loci*, *NPR3* and *SPON1*, associated with LS-BMD and total hip BMD, respectively. Another successfully widely used approach to identify rare variants is whole-genome sequencing (WGS). A sequencing-based study in Icelandic individuals reported a rare novel nonsense variant within *LGR4* associated with BMD and fracture risk

(Styrkarsdottir *et al.*, 2013). Another extremely powerful meta-analysis using WGS, whole-exome sequencing (WES) and genotype imputation identified 2 novel low-frequency non-coding variants with large effects on BMD and OFs, mapping near *EN1* and *WNT16* (Zheng *et al.*, 2015). Notably, *EN1* had never been associated with BMD before.

Besides SNPs, other genetic variants have also been studied in GWASs. Copy number variations (CNVs; DNA segments of 1 kb or larger present at a variable copy number in the population) have been associated with OF, BMD and hip geometry in some studies, identifying CNVs affecting several *loci*, such as the *UGT2B17*, *VPS13B131*, and 6p25.1 *loci* (Deng *et al.*, 2010; Oei *et al.*, 2014b; Yang *et al.*, 2008).

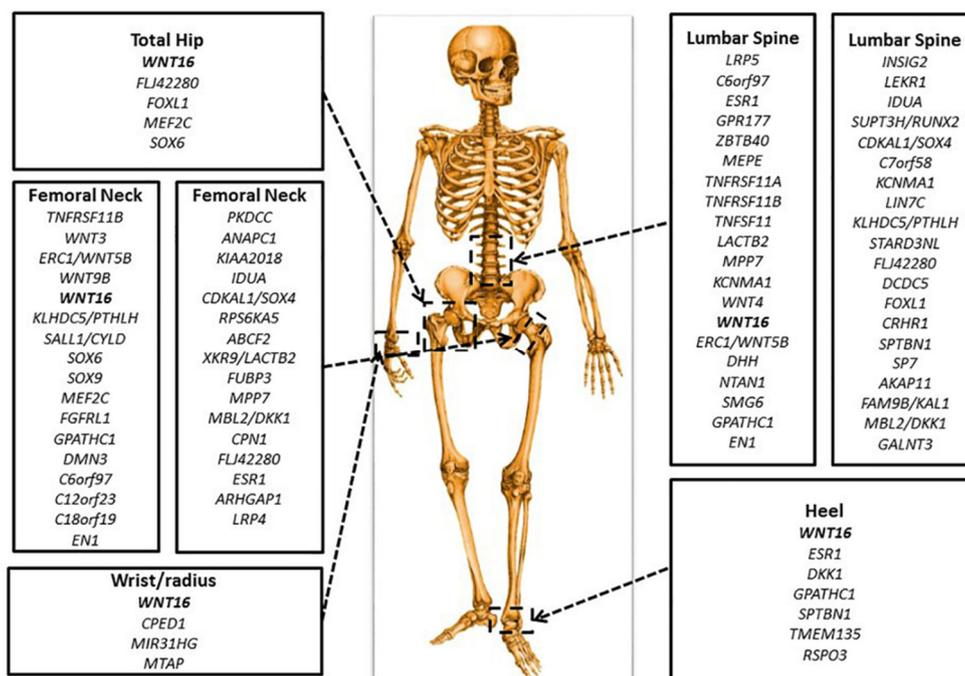
In an attempt to identify age-specific *loci*, some GWASs of BMD in children have also been performed. In 2009, Timpson *et al.* published the first one in which they identified four SNPs associated with total-body BMD (TB-BMD) in the *SP7 locus*. Medina-Gómez *et al.* (2012) identified variants in the *WNT16 locus* (also including *CPED1*) that showed genome-wide association with skull and TB-BMD in children. Moreover, in 2018, Medina-Gómez *et al.* performed a multi-ethnic life-course meta-analysis of TB-BMD in which they described variants in 80 *loci*, 36 of which had not been previously identified. In addition, they showed that variants in 2 *loci* displayed a clear age-specific effect, including variants in *ESR1* and in close proximity to *TNFSF11*. These results suggest that most of the genetic variants identified in GWAS regulate BMD early in life (i.e. PBM accrual) and their effect can be observed many years later.

As observed in linkage studies, GWASs have described a few sex-specific *loci*. For example, Zhang *et al.*, (2014) performed a meta-analysis for FN-BMD, LS-BMD and hip BMD and identified 2 novel *loci*, one (*CLDN14*) in a female-specific sample. In children, Chesi *et al.* (2017) reported 2 new sex-specific *loci* (*SPTB* and *IZUMO3*) associated with BMD at different skeletal sites. Estrada *et al.* (2012) performed sex-specific association analyses and identified only one *locus* at X chromosome (*FAM9B/KAL1*), which was male-specific.

To explore the potential ethnic specificity of osteoporosis *loci*, multi-ethnic GWAS have been performed. The first one was published by Xiong *et al.* (2009) and used a Caucasian cohort from USA as the discovery sample followed by replication in independent East Asian, African and Caucasian populations. They identified 2 novel genes (*ADAMTS18* and *TGFBR3*) associated with BMD and hip fracture. Kung *et al.* (2010) described associated variants in *JAG1*, using a Chinese discovery sample and European and Asian populations to replicate the findings. Besides, some replication

studies in different populations have been carried out, such as the one published by Stykarsdottir *et al.* (2010) in which they showed that 14 *loci* (out of 23) found associated with BMD in European populations are also associated in East-Asians.

Finally, GWASs have further supported skeletal site-specific effects of some *loci* (Figure 8). Kemp *et al.* (2014) clearly showed it assessing BMD at several axial and appendicular skeletal sites and reporting that variants at *CPED1* exerted a larger influence on skull and upper limb BMD when compared with lower limb BMD, whilst variants at *WNT16* were more strongly associated with upper limb BMD than with skull or lower limb BMD.



**Figure 8.** Genetic loci associated with BMD at various skeletal sites identified in GWASs. From Yuan *et al.*, 2019.

### 3.4. Epigenetics of osteoporosis

Epigenetics refers to heritable phenotype changes due to mechanisms other than the changes in the underlying DNA sequence. Epigenetic mechanisms are cell- and tissue-specific and are dependent on the interaction between the genome and the environment.

Most studies evaluating the association of epigenetic changes and osteoporosis have focused on miRNAs. miRNAs are endogenous small single-strand non-coding RNA molecules that post-transcriptionally regulate gene expression by targeting mRNAs and

inhibiting their translation or promoting their degradation. At the moment, numerous miRNAs that regulate bone remodelling, including differentiation and proliferation of OBs and OCs have been identified (reviewed in Bellavia *et al.*, 2019; Jing *et al.*, 2015). In addition, some studies have studied the expression profile of miRNAs in blood and bone samples from osteoporotic and non-osteoporotic individuals (reviewed in Letarouilly *et al.*, 2018). Besides, variants in miRNAs have been described as determinants of bone mass and BMD (De-Ugarte *et al.*, 2017; Dole & Delany, 2016).

Another epigenetic mechanism is histone modification, including methylation, acetylation or phosphorylation. These modifications are regulated by different types of enzymes, such as histone deacetylases, methyltransferases, or acetyltransferases. Several enzymes can influence bone remodelling by regulating genes involved in OB and OC differentiation. For instance, sirtuine 1 (encoded by *SIRT1*) is a histone deacetylase that regulates the *SOST* promoter, reducing its expression and increasing the Wnt/ $\beta$ -catenin signalling and, thus, bone formation (Cohen-Kfir *et al.*, 2011). Additionally, other studies have been carried out on the effect of histone modifications on the regulation of bone mass and the involvement in osteoporosis (reviewed in Vrtačnik *et al.*, 2014).

Finally, DNA methylation is a reversible modification of a cytosine residue located 5' to a guanosine residue (CpG). DNA methylation primarily represses gene expression by modulating the binding of proteins to DNA. Changes in DNA methylation are associated with aging and related diseases (Jung & Pfeifer, 2015). Some studies have assessed the role of DNA methylation in osteoporosis pathogenesis. For example, epigenome-wide association studies have been performed, in which the association between BMD or OF and methylation at multiple CpG sites has been tested (reviewed in Michou, 2018). Interestingly, several regions showing differential methylation overlap with the genes with variants associated with BMD or other bone parameters in GWASs.

### **3.5. Functional studies**

Functional studies are crucial for validating genetic associations and uncovering new genes/variants involved in a phenotype.

The association of a genetic marker with a trait of interest can be due to a direct causal relationship or to an indirect association, in which the associated polymorphism is in linkage disequilibrium (LD) with the causal variant. For this reason, it is necessary to

prove its functionality and to understand the biological mechanisms underlying the association (Gallagher & Chen-Plotkin, 2018).

### **3.5.1. Functional genomic integrative analysis**

The vast majority of the associated variants are located in non-coding regions of the genome, making plausible that they are regulatory variants. *In silico* integrative analyses in relevant cell types or tissues are used to map the variants to functional regions, characterized by chromatin states and histone modifications, binding of transcription factors (TFs), etc. In order to prioritize variants, Morris *et al.* (2019) surveyed chromatin accessibility of *loci* containing associated SNPs by generating ATAC-seq (assay for transposase-accessible chromatin using sequencing) maps in the human osteosarcoma cell line Saos-2 and using publicly available DNase I hypersensitive site maps in primary human osteoblasts (hOBs) from ENCODE. They found that SNPs were enriched for these genomic signatures of function.

Furthermore, integrative genomic analyses have also been used to map newly discovered associations with bone phenotypes. Guo *et al.* (2016) analysed some associated genes found in GWASs for their enrichment or depletion in epigenomic elements and found 4 TF binding sites, 27 histone marks, and 21 chromatin states segmentation types. Afterwards, they used this epigenomic signature to predict new candidate genes, which they tested for association with BMD and OF. Through this approach, they identified the *BDNF* gene. Qiu *et al.* (2019) prioritize putative enhancer SNPs (based on publicly available chromatin segmentation data from the Roadmap Epigenomics Project) and performed a GWAS meta-analysis for BMD. They identified 15 novel enhancer SNPs associated to BMD, 5 of which mapped to novel genes.

### **3.5.2. Gene expression studies and eQTLs**

Gene expression studies have been widely used to validate and identify genes involved in complex diseases. Differences in gene expression have been explored in individuals presenting or lacking the trait of interest. Recently, Ma *et al.* (2016) compared gene expression in B cell samples of postmenopausal women with high or low BMD and identified 308 differentially expressed genes, enriched in intracellular signalling cascade (e.g. *STAT5B*, *MAP2K5*), structural constituents of cytoskeleton (e.g. *CYLC2*, *TUBA1B*), membrane-enclosed lumen (e.g. *CCNE1*, *INTS5*) and purine biosynthesis and

metabolism (e.g. *ATP2C1*, *HPRT1*). Besides, Li *et al.* (2016) screened microarray data for differentially expressed genes between patients with osteoporosis and normal controls in peripheral blood monocytes. They found 373 up-regulated genes (e.g. *IRAK3*, *IFT52*, *NRIP1*) and 752 down-regulated genes (e.g. *SEMA4F*, *GATA6*, *GFOD2*) enriched in many osteoporosis-related signalling pathways, such as calcium signalling or androgen receptor binding.

Moreover, the emergence of studies on the allele-specific effect of a variant on gene expression at a cellular level (expression QTLs, eQTLs) has provided further insights on mechanisms underlying genetic association and disease pathophysiology. Genetic variants might affect gene expression through effects on transcription, splicing, or mRNA stability. Although eQTL data from primary bone cells is limited and human bone material is scarce, many studies have used this approach. For instance, Grundberg *et al.* (2009) carried out an eQTL study in primary hOBs from 95 Swedish unrelated donors and converged the SNPs identified as *cis*-eQTLs with BMD-associated SNPs reported in GWASs. They identified a potential osteoporosis candidate gene (*SRR*) comprising a strong *cis*-eQTL that was found nominally associated with BMD in the original GWAS and thus did not meet the criteria for follow-up studies. In 2018, Mullin *et al.* performed a *cis*-eQTL study in human OCs from 158 donors and found 24 BMD-associated variants from a GWAS meta-analysis significantly associated with the expression of 32 genes, such as *CYP19A1*, *CTNNB1*, *COL6A3* and *IQGAP1*. Finally, Hsu *et al.* (2019) used whole bone transcriptome data to evaluate the *cis*-eQTL capacity of the variants they found associated with different proximal femur geometry phenotypes in a GWAS meta-analysis. They discovered a variant near *PPP6R3* and *LRP5* that influenced *PPP6R3* expression and a variant near *FGFR4* that influenced *PDLIM7* expression. In addition, they also assessed the expression of candidate genes during cell differentiation in mouse calvarial osteoblasts.

### 3.5.3. Chromatin conformation analysis

Physical contact between a regulatory element and its target gene is crucial. Thus, evaluating physical interactions of candidate regions might help to understand their functionality. Chromatin conformation capture technologies (i.e. 3C, 4C, 5C, Hi-C, Capture-C) have been widely used to characterize GWAS-associated *loci* and variants, as well as to define topologically associating domains (TADs) in which interactions are more likely to occur and association signals are more likely to exert their effect.

In the bone field, many studies have taken advantage of these technologies, together with other approaches, to characterize previously described risk *loci* (e.g. Chen *et al.*, 2018; Zhu *et al.*, 2018). Recently, a study aiming at identifying candidate genes for complex traits using TAD data was published (Way *et al.*, 2017). They developed a method that prioritized genes within TAD boundaries including a GWAS signal based on Gene Ontology (GO) analysis. They tested it in BMD GWASs and identified candidate genes involved in bone pathways that, in many cases, were not the nearest gene to the lead signal (e.g. the *ACP2* gene, regulator of OBs metabolism, was implicated near the *ARHGAP1* locus). Chesni *et al.* (2019) performed a high-resolution genome-wide promoter-focused Capture C assay in primary human MSCs-derived OBs and combined it with ATAC-seq to detect BMD GWAS variants in open chromatin interacting with putative target gene promoters. Several novel genes were discovered, among which, *ING3* and *EPDR1*, that were verified by further functional analyses showing strong effects on osteoblastic and adipogenic differentiation.

#### **3.5.4. Other functional assays**

Other functional assays to test variants or putative regulatory regions are reporter assays, in which a region of interest is cloned upstream of a reporter gene in a pertinent cell type and the activity of the region and alternative alleles can be tested. Several regions can be tested at the same time by massive parallel reporter assays (MPRAs; Inoue & Ahituv, 2015). Moreover, regulatory variants can affect the binding of TFs, which can be *in silico* predicted and further validated *in vitro* by electrophoretic mobility assays (EMSAs) or chromatin immunoprecipitation followed by qPCR (ChIP-qPCR) using allele-specific probes (Hellman & Fried, 2007). As an example, in 2009, Xiong *et al.* performed EMSAs to demonstrate that the allele change of a SNP found associated with hip BMD and fracture in *ADAMTS18* generated a binding site for the TEL2 factor, as predicted by bioinformatic analyses.

All these experiments, however, do not test the region/variant in its genomic context. In this sense, gene editing experiments represent more physiologically-relevant methods to confirm the functionality of the gene/regulatory region/variant of interest.

### 3.5.5. Genome editing and animal models

Genome editing has been widely used to study disease-associated candidate genes, since it allows the modification of the gene or variant of interest with efficiency and precision. Specifically, the clustered regularly interspaced short palindromic repeats (CRISPR)-based systems have revolutionized the field due to their higher precision and flexibility (Gaj *et al.*, 2016). Analogous to MPRA mentioned above, high-throughput CRISPR screens to identify functional genes or noncoding regulatory regions have also been developed.

In the osteoporosis field, genome editing has been performed at a cellular and animal (both global and tissue specific) levels, analysing the molecular effects of the deletion of regions of interest and screening for genes or variants involved in the phenotype. For instance, Chen *et al.* (2018) and Zhu *et al.* (2018) deleted two putative enhancers by CRISPR-Cas9 in hFOB 1.19 and U2OS cells, respectively, and measured the expression of their putative target genes. Zheng *et al.* (2012) generated homozygous mice with targeted disruption of 2 GWAS osteoporosis candidate genes (*WNT16*, *FAM3C*) by homologous recombination techniques and showed that *Wnt16*<sup>-/-</sup> mouse had reduced cortical thickness and bone strength. Interestingly, the International Mouse Knock-out Consortium (IMKC) aims at generating knock-out mice of each of the known protein-coding genes in C57BL/6 mice and, as part of the International Mouse Phenotyping Consortium (IMPC), the Origins of Bone and Cartilage Disease project aims at identifying mutants with skeletal phenotypes (Freudenthal *et al.*, 2016). In 2012, Bassett *et al.* identified 9 new genetic determinants of bone mass and strength. Three of these knock-out strains (*Bbx*, *Cadm1*, *Fam73b*) presented weak but flexible bones with low mineral content, similar to those in postmenopausal osteoporotic individuals.

Another method to identify new regions (coding and non-coding) involved in bone phenotypes is random mutagenesis induced by chemicals such as N-ethyl-N-nitrosourea followed by skeletal phenotypic screening (Barbaric *et al.*, 2008; Mohan *et al.*, 2007).

Apart from mice, other animal models are used to study the genetic determinants of osteoporosis, such as rats, chicken, zebrafish and large animals (Karasik *et al.*, 2016). The main advantages of using animal models are the greater ability in environmental control, reproducibility, easier access to trait-relevant tissues and genetic manipulation. Considering the particular characteristics of each animal model, caution should be taken in translating the findings in human populations.

The most popular model for postmenopausal osteoporosis is generated in mouse, rat, sheep and non-human primates by ovariectomy, which causes a drastic reduction of estrogen levels, leading to a high bone turnover (reviewed in Komori, 2015). Zebrafish is another model system with great potential for functional studies. Osteoporosis-like phenotype can be induced by prednisolone treatment, which is associated with altered expression of several genes with a role in osteoblastogenesis and osteoclastogenesis (de Vrieze *et al.*, 2014).

### **3.6. Biological pathways underlying osteoporosis**

A large proportion of osteoporosis candidate genes discovered by the different approaches described in this section are involved in well-known crucial bone pathways.

#### **3.6.1. Wnt/ $\beta$ -catenin signalling**

Wnt signalling is the major bone anabolic pathway and it is critical for bone development during embryogenesis and for bone formation, resorption and coupling in postnatal bone, since it is involved in differentiation, proliferation and apoptosis of bone cells (Baron & Kneissel, 2013). WNT proteins are secreted glycoproteins that bind to the Frizzled membrane receptors and the low-density lipoprotein (LDL) receptor-related protein 5/6 (LRP5/6) co-receptors. The activation of the pathway results in the stabilization of  $\beta$ -catenin that accumulates and subsequently translocates to the nucleus, where it binds to the TCF/LEF1 TFs and initiates the transcription of the target genes promoting bone formation. The Wnt pathway is inhibited by DKK1 and SOST, two proteins secreted by OCys that bind to LRP5/6, exerting an antagonizing effect (Angers & Moon, 2009).

Furthermore, the Wnt pathway directly interacts with other important bone pathways. For example, JAG1 is a Wnt/ $\beta$ -catenin target but also an important component of the NOTCH pathway (Katoh & Katoh, 2006), and  $\beta$ -catenin up-regulates the expression of OPG in osteoblasts (Sato *et al.*, 2009).

Many genes from the Wnt pathway have been identified as osteoporosis susceptibility genes: *LRP5*, *SOST*, *WNT1*, *LRP4*, *AXIN1*, *CTNNB1*, *DKK1*, *MEF2C*, *PTHLH*, *RSPO3*, *SFRP4*, *WLS*, *EN1*, *WNT4*, *WNT5B*, *WNT16*. Many of these genes have been involved

in bone monogenic diseases and several functional studies to clarify their involvement in bone homeostasis have been carried out (reviewed in Mafi Golchin *et al.*, 2016).

### **3.6.2. OPG-RANK-RANKL signalling**

The OPG-RANK-RANKL signalling pathway is essential for regulating coupling between OB and OC activity. RANK, RANKL and OPG are members of the tumour necrosis factor (TNF)-related transmembrane cytokine superfamily, encoded by the *TNFRSF11A*, *TNFRSF11* and *TNFRSF11B* genes, respectively. RANKL is a soluble factor secreted mainly by OCys, and also by OBs, that binds to its receptor, RANK, on the cell surface of monocytes stimulating OC recruitment, differentiation and activation in the presence of monocyte colony stimulating factor M-CSF. Upon binding, NF- $\kappa$ B is activated and translocated to the nucleus, where transcription of osteoclastogenic genes is triggered (Boyce & Xing, 2007). OPG is a soluble decoy receptor of RANKL also secreted by OBs. It competes with RANK for binding to RANKL, preventing OC induction and, thus, bone resorption. Therefore, bone resorption is regulated by the ratio RANKL/OPG.

Recently, a RANKL reverse signalling has been described, in which RANK is secreted in vesicles by maturing OCs and binds to osteoblastic RANKL, inducing bone formation via *RUNX2* activation (Ikebuchi *et al.*, 2018).

RANK, RANKL and OPG have been shown to have other functions beyond regulating bone remodelling and coupling, including potential roles in other diseases, namely vascular calcification, diabetes and cancer (Harper *et al.*, 2016).

*TNFRSF11A*, *TNFRSF11* and *TNFRSF11B* have been repeatedly found in many GWASs and meta-analyses for BMD and OF, as reviewed in section 3.3.2 of this Introduction. Functional studies have highlighted the importance of these genes in bone physiology and mutations have been found in several skeletal dysplasias such as osteopetrosis or Paget's diseases of bone (Whyte, 2006).

### **3.6.3. NOTCH signalling**

NOTCH is a family of 4 transmembrane proteins (NOTCH1-4) that require cell-to-cell contact for activation through several ligands, such as JAG1/2 and DII1/3/4. The ligand-mediated activation induces a proteolytic cleavage releasing the NOTCH intracellular

domain, which translocates to the nucleus where transcription of target genes begins (Kopan & Ilagan, 2009). OBs and OCs require NOTCH signalling for differentiation and correct function, but the specific roles of NOTCH depend on the differentiation status of the cell (Regan & Long, 2013).

Several osteoporosis candidate genes are related to the NOTCH pathway, including *JAG1*, *MAPT*, and *NOTCH2*.

### **3.6.4. TGF- $\beta$ /BMP signalling**

Transforming growth factor  $\beta$  (TGF- $\beta$ ) and bone morphogenetic proteins (BMPs) are two families of the TGF- $\beta$  superfamily of proteins, involved in the control of cell proliferation, differentiation and other functions in many cell types. TGF- $\beta$  ligands bind as dimers to transmembrane receptors complexes that are comprised of two Serine-Threonine kinases and co-receptors. Upon binding, the SMAD TFs are phosphorylated and translocate into the nucleus to activate target gene expression.

In bone, these families of proteins play critical roles in development and tissue homeostasis (Wu *et al.*, 2016). For example, TGF- $\beta$ 1 is thought to be a coupling factor between bone formation and bone resorption. Several genes related to the TGF- $\beta$ /BMP signalling pathways have been identified as osteoporosis susceptibility genes, including *TGFBR3*, *BMP2*, *BMP4*, *SMAD3*, *SMAD9* and *BMPR2*.

### **3.6.5. Ephrin signalling**

Ephrin/Eph signalling is involved in adult tissue homeostasis and developmental processes, including bone, as well as fracture repair and skeletal response to PTH. Ephrins are the membrane-bound ligands of the Eph family of receptor tyrosine kinases. When ephrins bind to Eph on the neighbouring cell, a bidirectional signalling is activated.

In bone, Ephrin-B2 is expressed by OCs and binds to Eph-B4 in OBs, enhancing osteogenic differentiation and inhibiting osteoclastogenesis by reverse signalling suppressing the cFos-NFATc1 pathway (Pasquale, 2008). Some genes related in the ephrin signalling have been identified and osteoporosis susceptibility genes, including *EPHB2*, *EPHA4* and *NFATC1* (Morris *et al.*, 2019; Nielson *et al.*, 2016).

### **3.6.6. Endochondral ossification and MSCs differentiation**

As described in section 1.4 of this Introduction, the majority of bones in the human skeleton are formed through endochondral ossification. GWASs have identified many genes involved in this process, including genes related to cartilage development and ossification and OB differentiation. Some examples are: *IBSP* (bone sialoprotein 2), *PTH1H*, *RUNX2*, *SOX6*, *SOX9*, *SPP1* (osteopontin), *SOX4*, *FAM3C* and *SP7* (osterix). The roles of the individual genes in the different parts of these processes are reviewed by Mafi Golchin *et al.* (2016) and Richards *et al.* (2012). Notably, *RUNX2* is an essential TF for pre-OBs differentiation and homozygous knock-out mice show complete absence of bone, dying perinatally owing to a softened cartilaginous ribcage unable to support respiration (Komori *et al.*, 1997).

## 4. ATYPICAL FEMORAL FRACTURE

### 4.1. Definition and diagnosis

Atypical femoral fractures (AFFs) are a very rare type of fractures that occur at the subtrochanteric region or the femoral diaphysis of long bones (Figure 9; see Box 1). They were first described in 2005 by Odvina *et al.* in a series of patients on alendronate and with an over-suppression of bone turnover. In 2010, the American Society for Bone and Mineral Research (ASBMR) established the case definition for AFFs (Shane *et al.*, 2010), which was updated in 2014 (Shane *et al.*, 2014). AFFs have distinctive characteristics that are shown in Table 6. The diagnosis of AFF is based on femoral location (from just distal to the lesser trochanter to just proximal to the supracondylar flare) and the presence of at least 4 of 5 major features. Minor features, despite being commonly associated with AFF, are not required for the diagnosis.



**Figure 9.** Complete diaphyseal AFF. Note the fracture line running perpendicular at the long axis of the femur and becoming oblique as it progresses. Note the general thickness of the lateral cortex. Black arrow: endosteal callus reaction. Horizontal white arrow: periosteal callus reaction. Oblique white arrow: medial spike. Extracted from Schilcher, 2013.

**Table 6.** AFF case definition

Major features
The fracture is associated with minimal or no trauma, as in a fall from a standing height or less
The fracture line originates at the lateral cortex and is substantially transverse in its orientation, although it may become oblique as it progresses medially across the femur
Complete fractures extend through both cortices and may be associated with a medial spike; incomplete fractures involve only the lateral cortex
The fracture is noncomminuted or minimally comminuted
Localized periosteal or endosteal thickening of the lateral cortex is present at the fracture site (“beaking” or “flaring”)
Minor features
Generalized increase in cortical thickness of the femoral diaphyses
Unilateral or bilateral prodromal symptoms such as dull or aching pain in the groin or thigh
Bilateral incomplete or complete femoral diaphysis fractures
Delayed fracture healing

Extracted from Shane *et al.*, 2014

## 4.2. Epidemiology and risk factors

AFFs are very rare events, with an overall incidence in the general population of 3.0-9.8 per 100,000 persons-year (Khow *et al.*, 2017; Meier *et al.*, 2012b; Meling *et al.*, 2014), which represents a 3.5-5.7% of total subtrochanteric or femoral shaft fractures and a 0.2-0.8% of total hip fractures (Khow *et al.*, 2017; Saita *et al.*, 2015). Bilaterality has been reported in 20-60% of patients and, usually, both fractures occur at the same location of the contralateral side (Lim *et al.*, 2018; Probyn *et al.*, 2015). Besides, up to 70% of patients reported to have prodromal pain and in a 25-45% fracture healing was delayed (Black *et al.*, 2019; Shane *et al.*, 2014).

Several studies have shown a strong association of AFFs with N-BPs, with more than 80% of AFFs occurring in patients on N-BPs and only around a 10% in N-BP-naïve patients (Kharwadkar *et al.*, 2017; Mahjoub *et al.*, 2016; Schilcher *et al.*, 2014; Silverman *et al.*, 2018). In addition, AFFs have been also described in patients on denosumab or other anti-osteoporotic drugs (Black *et al.*, 2019; Bone *et al.*, 2017; Cosman *et al.*, 2016), as well as on GCs (Koh *et al.*, 2017). Likewise, very similar -if not clinically indistinguishable- fractures occur in other monogenic skeletal dysplasias, such as hypophosphatasia, pycnodysostosis or osteogenesis imperfecta (Meier *et al.*, 2012a; Sutton *et al.*, 2012; Yates *et al.*, 2011).

The incidence of N-BP-associated AFFs increases with the duration of the treatment, especially after 3 years, being 1.8 per 100,000 persons-year at 2 years of treatment, 38.9 per 100,000 persons-year at 6-8 years of treatment, and 113.1 per 100,000 persons-year at 10 years of treatment (Brown, 2017; Gedmintas *et al.*, 2013; Shane *et al.*, 2014). Thus, the overall relative risk of AFF for any BP use is 1.70 but it increases by an odds ratio of 2.74 for more than 5 years of N-BP therapy (Gedmintas *et al.*, 2013; Park-Wyllie *et al.*, 2011). Notably, following cessation of N-BPs, the risk diminishes by 70% per year (Schilcher *et al.*, 2015b; Silverman *et al.*, 2018). All in all, the benefits of N-BPs treatment far outweigh any AFF risks.

It has been reported that women have a 3-fold higher risk than men (Schilcher *et al.*, 2015b), probably due to the increased occurrence of osteoporosis and N-BP use in women. Interestingly, patients who develop AFFs are somewhat younger than those who develop non-AFF proximal femoral fractures, with a mean age range of 66-75 years versus 75-89 years, respectively (Khow *et al.*, 2017).

Contrarily to osteoporosis, Asian ethnic background increases the age-adjusted relative hazard of AFF by 6.6-fold compared to Caucasian women (Lo *et al.*, 2016),

possibly due to their differences in femur geometry. Schilcher *et al* (2015a) reported different localization patterns of AFFs in Singapore and Sweden.

In contrast to OFs, mortality rates for AFFs have been reported to be similar to those in the general population (Kharazmi *et al.*, 2016).

Risk factors for AFF may partly overlap with those for osteoporosis. Apart from the already mentioned (e.g. N-BPs, and especially treatment duration, other drugs, previous stress fracture of the contralateral femur, gender, Asian ethnicity, and age), other potential risk factors have been proposed. Among them, hip and femoral geometry (see section 4.3), physical activity, other comorbid conditions and high BMI (Black *et al.*, 2019; Koh *et al.*, 2017). Importantly, age is less strongly predictive of AFF than it is of OFs.

### **4.3. Pathogenesis**

The pathogenesis of AFF remains largely unknown but its epidemiological association with antiresorptive drugs (and, in particular, with N-BPs) led to several proposed mechanisms (Compston, 2011; Ettinger *et al.*, 2013; Ng *et al.*, 2014). However, given that AFFs also occur in patients not exposed to these drugs, some authors have suggested osteoporosis itself as a possible etiology of AFF (Adler, 2018). In addition, the AFF cases in patients with other bone disorders related to defects on bone mineralization, remodelling and collagen synthesis and structure provide further insight into the possible pathophysiology of AFFs.

AFFs are considered insufficiency or stress fractures because they develop over time (as manifested by prodromal pain), appear to start in locations of stress of the lateral femur and show a periosteal callus (Black *et al.*, 2019; Shane *et al.*, 2014). Thus, it might be useful to consider the etiology of stress fractures in relation to AFF development.

Some, but not all, of the reported cases of AFF presented with a severely reduced bone turnover (Odvina *et al.*, 2005; Qiu *et al.*, 2017; Visekruna *et al.*, 2008). Therefore, it has been posited as an underlying mechanism for AFF by which the mean age of bone increase, bone composition and mechanical properties are altered and strength and fracture resistance are reduced (Larsen & Schmal, 2018; Lloyd *et al.*, 2017). In this regard, it has been suggested that long-term N-BP therapy may result in accumulation of microcracks that may not be repaired and propagate until the AFF occurs (Allen & Burr, 2007; Shane *et al.*, 2014; Starr *et al.*, 2018).

On the one side, N-BP therapy causes an increase and uniformity of bone mineralization that makes the bone more rigid but brittle and enables microcrack initiation and propagation more rapidly (Donnelly *et al.*, 2012; Güerri-Fernández *et al.*, 2013). On the other side, the reduction of bone remodelling has been shown to alter collagen maturity and increase oxidative non-enzymatic collagen cross-linking, associated to an accumulation of advanced glycation end-products (AGEs), that reduce bone plasticity and toughness and increase bone brittleness and risk of fracture (Tang *et al.*, 2007; Vashishth *et al.*, 2001).

Besides, N-BPs may impair microcrack and AFF repair since they inhibit bone remodelling and have anti-angiogenic effects, hindering vascularization of the fracture zone required for healing, as capillars are a source of OC and OB precursors (Compston, 2011; Li *et al.*, 2001).

Following the consideration of AFF as stress fractures, mechanical loading has been postulated as a contributor to AFF pathophysiology and the geometry of the femur has been suggested as another important factor, since it influences femoral strain patterns and, thus, AFF development and location (Mahjoub *et al.*, 2016; Oh *et al.*, 2017). Indeed, femoral bowing and *coxa vara* have been associated with AFF and may confer increased relative hazard of AFF in the Asian population (Hagen *et al.*, 2014; Koh *et al.*, 2017; Oh *et al.*, 2014).

All in all, none of the mechanisms described are solely responsible for AFF. On the contrary, the current evidence suggests that the physiopathology is complex and AFF occurrence requires a “perfect storm” of subject-specific factors, such as response to N-BPs, femoral geometry and bone composition and microarchitecture. Moreover, genetic factors might also be involved in AFF pathogenesis (see section 4.4).

#### **4.4. Genetics**

The rare occurrence of AFFs, even in N-BPs users, together with some evidences such as the higher prevalence in Asian women or the higher propensity conferred by a certain femoral geometry, have raised the hypothesis that genetic factors predispose to AFFs (Nguyen *et al.*, 2018). In addition, the identification of 2 families (one of which studied in this thesis; Lau *et al.*, 2017) with multiple family members affected is also suggestive of an underlying genetic background.

Genetic studies on AFF susceptibility can be divided in those based on small cohorts of individuals with AFF and those based on patients with other monogenic bone diseases.

#### 4.4.1. Genetic studies in small AFF cohorts

Few studies with small cohorts of N-BP-associated AFF patients have been carried out, using different approaches, and they have identified mutations in different genes (Table 7).

**Table 7.** Genes mutated in AFF patients in cohort studies

Gene	Mutations	ExAC freq.	N° cases mut/ N° cases studied	Years N-BPs	Patients charact.	Genetic analysis	Ref.
<b>ALPL</b>	c.648+1G>A Heteroz.	8.24x10 <sup>-6</sup>	1 F / 11	NA <sup>a</sup>	HPP <sup>b</sup>	Gene seq.	Sum <i>et al.</i> , 2013
<b>COL1A2</b>	p.Arg708Gln Heteroz.	0.0008	1 F / 5	>5	No OI features	Gene seq.	Funck-Brentano <i>et al.</i> , 2017
<b>CTSK</b>	c.784+3A>C Homoz.	5.77x10 <sup>-5</sup>	2 consanguineous sisters / 11	0	No PYCD features	WES	Lau <i>et al.</i> , 2017
<b>PPEF2</b>	p.Arg388Gln	0.001	5 alleles / 26 alleles (13 F) <sup>c</sup>	1-10	-	Exon array	Pérez-Núñez <i>et al.</i> , 2015

<sup>a</sup>Duration of treatment not specified but the analysis was carried out during N-BP treatment

<sup>b</sup>HPP was diagnosed after the mutation was found

<sup>c</sup>The number of AFF cases bearing the mutation is not described, the authors give frequency of the mutation in the pool of 13 cases

F: female; HPP: hypophosphatasia; OI: osteogenesis imperfecta; PYCD: pycnodysostosis; WES: whole-exome sequencing

Three studies have searched for variants in candidate genes (Bhattacharyya *et al.*, 2016; Funck-Brentano *et al.*, 2017; Sum *et al.*, 2013). *ALPL*, the gene encoding for the tissue non-specific alkaline phosphatase (TNSALP), was the only one analysed in the 3 studies. TNSALP is the enzyme responsible for PP<sub>i</sub> hydrolysis and loss-of-function mutations in *ALPL* cause hypophosphatasia (HPP), due to extracellular accumulation of PP<sub>i</sub>, which inhibits bone mineralization (Whyte, 2016). Since N-BPs are analogues of PP<sub>i</sub> resistant to TNSALP activity and femoral fractures with atypical features occur in cases

of HPP without prior anti-resorptive therapy, it has been hypothesized that *ALPL* mutations can be a genetic risk factor for AFFs.

Sum *et al.* (2013) carried out a prospective *ALPL* mutation analysis of 11 patients with N-BP-associated AFFs in which they sequenced all coding exons and adjacent splice sites. In one patient, a single heterozygous mutation was found affecting the donor splice site in intron 6. This mutation was reported in lethal infantile HPP when associated with a second missense mutation on the other copy of the gene (Sergi *et al.*, 2001). Serum levels of alkaline phosphatase (ALP) in this AFF patient were low, although she was never diagnosed with HPP before.

In 2016, Bhattacharyya *et al.* conducted a retrospective case-control study to investigate the possible role of HPP as a risk factor for AFF. They analysed 10 patients who sustained N-BP-associated AFF with 13 controls, with a mean N-BP use of 9 years in both groups and they did not find any coding mutation in the *ALPL* gene in either AFF patients or controls. Additionally, no differences in ALP serum levels between the two groups were observed.

In a study by Funck-Brentano *et al.* (2017), the targeted sequencing of the *ALPL*, *COL1A1*, *COL1A2*, and *SOX9* genes was performed in 4 females and 1 male with N-BP-associated AFF. A heterozygous rare missense variant in *COL1A2* was identified in one patient. This gene encodes the pro- $\alpha$ 2 chain of type 1 collagen and the mutation found caused alterations in collagen fibrillogenesis (Vomund *et al.*, 2004). Notably, mutations in *COL1A2* cause osteogenesis imperfecta (OI), albeit no specific physical features of OI were identified in this patient, apart from short stature. No mutations were found in the other genes.

One study (Lau *et al.*, 2017) carried out a whole-exome sequencing (WES) study in a consanguineous family in whom 3 siblings (2 females and 1 male) sustained bilateral AFFs without previous N-BP exposure. WES of the 2 affected sisters unveiled a very rare homozygous mutation in the splice site of intron 6 of the *CTSK* gene (encoding for cathepsin K, essential for OCs-mediated bone resorption). Mutations in *CTSK* are associated with pycnodysostosis (PYCD), although the patients did not present any clinical feature of this disease, apart from short stature and high bone mass. OC culture from peripheral blood monocytes of affected patients exhibited a reduced bone resorptive activity. Moreover, Lau *et al.* sequenced the *CTSK* gene in 10 further cases with AFF and no mutation was found.

Finally, 2 studies performed genome-wide association analysis. Pérez-Núñez *et al.* (2015) conducted a pilot study in 13 AFF patients and 268 controls (87 healthy women and 181 patients with postmenopausal osteoporosis without AFFs). They explored the association of up to 300,000 genome-wide non-synonymous coding variants (with a minor allele frequency <0.03) with AFF by using an exon array. Twenty-one variants were found over-represented in the AFF group, although only one remained statistically significant after correction for multiple testing, due to small sample size. It is a missense variant in the *PPEF2* gene, which has no known function in bone metabolism. In addition, pathways analysis did not reveal any enriched pathway. Interestingly, AFF patients tended to accumulate a greater number of “risk variants”, suggesting that AFF might have a polygenic background.

More recently, and posterior to the work presented in this thesis, Kharazmi *et al.* (2019) published the largest case-control GWAS to date to determine whether common genetic variants contribute to risk of N-BP-associated AFFs. They compared 51 cases with two sets of controls: 4891 population controls or 324 matched controls that had been prescribed N-BPs due to osteoporosis but who did not have a diagnosis of cancer. They found 4 isolated SNPs associated with AFF when comparing with the general population controls. However, no statistically significant association was found when using the N-BP-treated controls, suggesting that either they were false positives, or they were related to the underlying phenotype that led to treatment indication. They also performed candidate gene analysis for 29 genes previously implicated in AFF or related bone diseases in other patients, but no statistically significant association was revealed when comparing AFF cases with either of the two control groups. They concluded that no evidence of a common genetic predisposition for N-BP-associated AFFs was found.

#### **4.4.2. Genetic studies in AFF patients with other monogenic bone diseases**

AFFs were found in individuals with 7 monogenic bone disorders affecting mineralization, bone remodelling, collagen synthesis and structure or OCy function. In some patients, the mutation underlying the disorder was described (Table 8).

Four cases of AFF occurring in adult HPP have been reported (Doshi *et al.*, 2009; Gagnon *et al.*, 2010; Maman *et al.*, 2016; Sutton *et al.*, 2012). In 3 of them, heterozygous or compound heterozygous mutations in *ALPL* were described and the genetic condition was unmasked after the occurrence of the AFF. Of those, only Sutton *et al.* (2012) reported a case with N-BP therapy after a misdiagnosis of postmenopausal osteoporosis.

Table 8. AFF patients with monogenic bone disorders

Monogenic bone disorder	Genes mut. in AFF patients	Mutations	N° AFF cases with mut/ total n° cases	BPs-treated cases (years)	Genetic analysis	References
<b>HPP</b>	ALPL	p.Ala176Thr and p.Val423Ala; p.Thr100Met and p.Glu191Lys; p.Arg71His	3 F / 4 F	1 (4 years)	Sanger seq.	Gagnon et al., 2010; Maman et al., 2016; Sutton et al., 2012; Doshi et al., 2009
<b>XLH</b>	NA	NA	0 / 1 M	0	NA	Whyte, 2009
<b>PYCD</b>	CTSK	p.Phe142Leufs*19 and p.Ser252Asn; exact mutation not reported for 2 cases	2 F, 1 M / 4 F, 3 M	0	Gene seq.	Nakase et al., 2007; Song et al., 2017; Hashem et al., 2015; Kundu et al., 2004; Yates et al., 2011; Yuasa et al., 2015
<b>Osteopetrosis</b>	NA	NA	0 / 4 F	0	NA	Amit et al., 2010; Kubaraci et al., 2013; Birmingham & McHale, 2008;
<b>OPPG</b>	LRP5	p.Arg752Trp and Trp79Arg	1 M / 1 M	0	Sanger seq.	Alonso et al., 2015
<b>OI</b>	COL1A2	p.Gly1102Ala	1 F / 4 F, 1 M	5 (5 years)	NA	Vasanwala et al., 2016; Holm et al., 2014; Meier et al., 2012a; Manolopoulos et al., 2013; Etxebarria-Foronda & Carpintero, 2015
<b>X-linked osteoporosis</b>	PLS3	p.Tyr79Ilefs*6	1 M	1 (9 years)	X-linked WES	van de Laarschot & Zillikens, 2016

F: female; HPP: hypophosphatasia; M: male; OI: osteogenesis imperfecta; OPPG: osteoporosis pseudoglioma syndrome; PYCD: pycnodysostosis; WES: whole-exome sequencing; XLH: X-linked hypophosphatemia

X-linked hypophosphatemia (XLH) has been also related to AFF since Whyte (2009) reported pseudo-fractures in the lateral cortex of the femoral shaft similar to AFF in a young N-BP-naïve male. XLH is caused by loss-of-function mutations of the *PHEX* gene (Fuente & Hernández, 2017). However, in the case reported no mutational analysis was performed.

AFFs have been described in 7 cases of PYCD (Hashem *et al.*, 2015; Kundu *et al.*, 2004; Nakase *et al.*, 2007; Song *et al.*, 2017; Yates *et al.*, 2011; Yuasa *et al.*, 2015). In 3 of them, the disease was unmasked after the AFF and no N-BP history was known for any of them. Mutations in *CTSK* were found in 3 N-BP-naïve patients (Nakase *et al.*, 2007; Song *et al.*, 2017). On the one hand, Song *et al.* (2017) reported a patient with AFF who had an underlying sclerosing bone disease. They target-sequenced 10 candidate genes by NGS to perform a differential molecular diagnosis and they found 2 heterozygous mutations in the *CTSK* gene. Nakase *et al.* (2007), on the other hand, presented the outcomes of surgical treatment of fractures of several patients already diagnosed with PYCD by genetic analysis, 2 of which were AFFs.

Four cases of AFFs occurring in N-BP-naïve individuals with osteopetrosis have been described (Amit *et al.*, 2010; Birmingham & McHale, 2008; Kumbaraci *et al.*, 2013). Osteopetrosis is caused by mutations in 8 genes, including *TNFSF11* (RANKL), *TNFRSF11A* (RANK), *CLCN7* and *OSTM1* (Sobacchi *et al.*, 2013). However, the underlying mutated genes were not described in these reports.

A single case report of AFF in a N-BP-naïve male with osteoporosis pseudoglioma syndrome (OPPG) has been described by Alonso *et al.* (2015). The patient had multiple fragility fractures and evidence of low bone turnover and carried two novel loss-of-function mutations in *LRP5*. Importantly, this is the only report of AFF occurring in a genetic condition with primary osteoblast dysfunction.

OI has been also related to N-BP-associated AFFs. On the one hand, 4 case reports in adults have been published to date, all with more than 3 years of N-BPs (Etxebarria-Foronda & Carpintero, 2015; Holm *et al.*, 2014; Manolopoulos *et al.*, 2013; Meier *et al.*, 2012a). However, in none of them the underlying gene mutation was described. On the other hand, Vasanwala *et al.* (2016) reported the only case of N-BP-associated AFF in a pediatric patient with OI type IV, who presented a heterozygous mutation in the *COL1A2* gene. OI is most often caused by defects in type 1 collagen synthesis (encoded by *COL1A1* and *COL1A2*) and structure, that leads to abnormal composition and organization of bone matrix, increased bone microdamage, stiffness and brittleness (Forlino & Marini, 2016). N-BPs may aggravate the situation by suppressing bone

remodelling and impeding microcrack repair. Indeed, a retrospective study demonstrated that a different pattern of femoral shaft fractures occurred in patients with OI treated with N-BPs compared to those not treated (Nicolaou *et al.*, 2012).

Finally, one case report of an AFF occurring in a patient with X-linked osteoporosis who had been treated with N-BPs was published (van de Laarschot & Zillikens, 2016). The patient presented a mutation in *PLS3*, the gene encoding for plastin 3 and responsible for X-linked osteoporosis, a juvenile form of osteoporosis that is thought to be due to a decreased mechanosensing by OCys (van Dijk *et al.*, 2013).

#### **4.5. Prevention and management of AFF**

As already stated in section 2.5.1 of this Introduction, the most extended measure to prevent N-BP-associated AFFs is to consider a drug holiday after 3-5 years of treatment in patients who are not at high risk of OF (Adler *et al.*, 2016; Compston *et al.*, 2017). In addition, femur imaging may be useful to early detect incomplete asymptomatic AFFs and avoid further progress to complete AFFs (van de Laarschot *et al.*, 2017).

In the case of complete AFFs, the first-line intervention is surgical fixation of bone, although the characteristic healing delay may hinder the recovery and increase morbidity. The management of incomplete AFFs depends on many factors, such as symptoms and radiographs. For painful AFFs, prophylactic surgery is recommended to prevent complete fracture. Otherwise, avoiding weight-bearing activity and surveillance is advised (Dell & Greene, 2018; Starr *et al.*, 2018).

Upon an AFF, N-BPs or other anti-resorptive agents should be discontinued and calcium and vitamin D supplementation should be considered, as well as hrPTH(1-34) treatment, that may improve fracture healing and mechanical strength, although the response has been variable (Im & Lee, 2015; Watts *et al.*, 2017).



# **OBJECTIVES**



The two main objectives of this thesis were to elucidate the causality and molecular mechanisms underlying the association of a GWAS signal for bone mineral density and osteoporotic fracture and to identify and characterize the genetic determinants of bisphosphonate-associated atypical femoral fracture.

To address them, the following specific objectives were proposed:

1. Functional characterization of the *C7ORF76 locus*, a GWAS signal for bone mineral density and osteoporotic fracture.
  - 1.1. To deeply re-sequence the *C7ORF76 locus* in a truncate selection of the BARCOS cohort in order to identify candidate variants, and to perform an association study with BMD and OF in the complete BARCOS cohort.
  - 1.2. To assess the possible role of the associated variants as *cis*-eQTLs in human primary osteoblasts.
  - 1.3. To characterize an upstream putative regulatory element (UPE) by reporter gene analysis and to identify its possible targets by chromatin conformation capture in osteoblastic cell types.
  - 1.4. To characterize an enhancer (eDlx#18) located in intron 2 of *C7ORF76* by reporter gene assays and to identify its targets by chromatin conformation capture in osteoblastic cell lines and mouse developing humeri.
  - 1.5. To assess the possible *cis*-eQTL function of 2 variants lying within the eDlx#18 enhancer in human primary osteoblasts.
  - 1.6. To generate a knock-out mouse model by CRISPR-Cas9 and to evaluate the expression of *Dlx5* in E11.5 embryos and the skeletal defects in E17.5 embryos.
2. Identification and characterization of genetic susceptibility to bisphosphonate-associated atypical femoral fracture
  - 2.1. To identify rare coding mutations in 3 sisters and 3 unrelated patients who sustained bisphosphonate-associated atypical femoral fracture by whole-exome sequencing.
  - 2.2. To evaluate the effect of the p.Asp188Tyr GGPPS mutation on its enzyme activity and structure.
  - 2.3. To delineate the role of GGPPS in bone cell types.



# **RESULTS**



---

## THESIS SUPERVISORS' REPORT ON THE CONTRIBUTION OF THE PHD CANDIDATE TO THE ARTICLES INCLUDED IN THIS DOCTORAL THESIS

**Thesis title:** "Identification and functional characterization of genetic *loci* involved in osteoporosis and atypical femoral fracture"

**Author:** Neus Roca Ayats

**Supervisors:** Dr. Daniel Grinberg Vaisman and Dra. Susanna Balcells Comas

**CHAPTER 1:** Functional characterization of the *C7ORF76 locus*, a GWAS signal for BMD and osteoporotic fracture

### Article 1

**Title:** Functional characterization of the *C7ORF76* genomic region, a prominent GWAS signal for osteoporosis in 7q21.3

**Authors:** Neus Roca-Ayats, Núria Martínez-Gil, Mónica Cozar, Marina Gerousi, Natàlia Garcia-Giralt, Diana Ovejero, Leonardo Mellibovsky, Xavier Nogués, Adolfo Díez-Pérez, Daniel Grinberg, Susanna Balcells

**Journal:** Bone      **Number:** 123 (2019)      **Pages:** 39-47

**Impact Factor (2018 JCR Science Edition):** 4.360

**Contribution of the PhD candidate:** Neus Roca performed most of the experiments in this study. In particular, she did the filtering, *in silico* analyses and experimental validation of the variants identified in the resequencing of the *C7ORF76* genomic region. She also performed the selection of the variants to be genotyped, the linkage disequilibrium analyses, and the association study with BMD and osteoporotic fracture. Next, she obtained and cultured the human primary osteoblasts used for transcriptomic and eQTL assays and extracted DNA and RNA for them. With those, she genotyped the relevant SNPs, and performed RT-qPCR and did the statistical analysis. Later, she designed luciferase reporter assays for selected regions, and carried out the statistical analysis. She also designed the 4C-seq experiments and performed cell culture and chromatin

fixation of the cell lines used. Finally, she performed the analysis of chromatin interaction results and TADs. In preparation of the MS, she performed the data analysis and interpretation, the elaboration of figures and tables and drafted the manuscript. After circulating among co-authors, she participated in the last revision and final editing of the manuscript.

### Article 2

**Title:** A *DLX5/6* enhancer in the *C7ORF76* locus: Characterization of its role in development and in bone

**Authors:** Neus Roca-Ayats, Núria Martínez-Gil, Mónica Cozar, Natàlia Garcia-Giralt, Xavier Nogués, Adolfo Díez-Pérez, Aleix Gavaldà-Navarro, Jordi Garcia-Fernández, Daniel Grinberg\*, Darío G. Lupiáñez\*, Susanna Balcells\*

**Journal:** To be submitted

**Contribution of the PhD candidate:** As in the previous article, Neus Roca was main author of this article. She was involved in the design of the study. Next, she did the *in silico* analyses of the enhancer, designed the luciferase reporter assays and performed the statistical analysis. Later, she cultured the human primary osteoblasts, performed DNA and RNA extraction from them, genotyped the relevant SNPs, did RT-qPCR of the relevant genes and performed the eQTL statistical analysis. She also designed the 4C-seq experiments, generated the libraries and analysed the chromatin interaction results and TADs. Finally, she generated knock-out mouse ESCs by CRISPR-Cas9 and validated the clones. She also performed the whole-mount *in situ* hybridization of E11.5 embryos, the skeletal preparation of E17.5 embryos and the assessment and measurements of bone parameters. She did the overall data analysis and interpretation, elaboration of figures and tables and drafting of the manuscript and was involved in the revision and final editing of the manuscript.

---

**CHAPTER 2:** Identification and characterization of genetic susceptibility to N-BP-associated AFF

### Article 3

**Title:** GGPS1 Mutation and Atypical Femoral Fractures with Bisphosphonates

**Authors:** Neus Roca-Ayats, Susana Balcells, Natàlia Garcia-Giralt, Maite Falcó-Mascaró, Núria Martínez-Gil, Josep F. Abril, Roser Urreizti, Joaquín Dopazo, José M. Quesada-Gómez, Xavier Nogués, Leonardo Mellibovsky, Daniel Prieto-Alhambra, James E. Dunford, Muhammad K. Javaid, R. Graham Russell, Daniel Grinberg, Adolfo Díez-Pérez

**Journal:** The New England Journal of Medicine                      **Number:** 376(18) (2017)  
**Pages:** 1794-1795

**Impact Factor (2017 JCR Science Edition):** 79.258

**Contribution of the PhD candidate:** Neus Roca performed the Sanger sequencing validation of the variants obtained through WES. She also performed the *in silico* analyses and pathway enrichment analysis of the validated variants. She participated in the data analysis and interpretation, and contributed to the drafting and revision of the manuscript.

### Article 4

**Title:** Functional Characterization of a GGPPS Variant Identified in Atypical Femoral Fracture Patients and Delineation of the Role of GGPPS in Bone-Relevant Cell Types

**Authors:** Neus Roca-Ayats\*, Pei Ying Ng\*, Natàlia Garcia-Giralt, Maite Falcó-Mascaró, Mónica Cozar, Josep Francesc Abril, José Manuel Quesada Gómez, Daniel Prieto-Alhambra, Xavier Nogués, James E Dunford, R Graham Russell, Roland Baron, Daniel Grinberg, Susana Balcells, Adolfo Díez-Pérez

\*co-first authors

**Journal:** Journal of Bone and Mineral Research                      **Number:** 33(12) (2018)  
**Pages:** 2091-2098

**Impact Factor (2018 JCR Science Edition):** 5.711

**Contribution of the PhD candidate:** Neus Roca performed the validation of the variants by Sanger sequencing and the *in silico* analysis of the variants. She also cloned, produced and purified both forms of GGPPS protein, the wild-type and the mutant one. Neus Roca was also responsible for the data analysis and their interpretation, the elaboration of the figures and tables and the drafting the manuscript. She participated in the revision and final editing of the manuscript.

Barcelona, 15<sup>th</sup> July 2019

Supervisors' signature

Dr. Daniel Grinberg Vaisman

Dra. Susanna Balcells Comas

## CHAPTER 1: FUNCTIONAL CHARACTERIZATION OF THE *C7ORF76* LOCUS, A GWAS SIGNAL FOR BMD AND OSTEOPOROTIC FRACTURE

### Article 1

Functional characterization of the *C7ORF76* genomic region, a prominent GWAS signal for osteoporosis in 7q21.3

#### Summary:

Genome-wide association studies (GWAS) have repeatedly identified genetic variants associated with bone mineral density (BMD) and osteoporotic fracture in non-coding regions of *C7ORF76*, a poorly studied gene of unknown function. The aim of the present study was to elucidate the causality and molecular mechanisms underlying the association. We re-sequenced the genomic region in two extreme BMD groups from the BARCOS cohort of postmenopausal women to search for functionally relevant variants. Eight selected variants were tested for association in the complete cohort and 2 of them (rs4342521 and rs10085588) were found significantly associated with lumbar spine BMD and nominally associated with osteoporotic fracture. *cis*-eQTL analyses of these 2 SNPs, together with SNP rs4727338 (GWAS lead SNP in Estrada *et al.*, *Nat Genet.* 44:491–501, 2012), performed in human primary osteoblasts, disclosed a statistically significant influence on the expression of the proximal neighbouring gene *SLC25A13* and a tendency on the distal *SHFM1*. We then studied the functionality of a putative upstream regulatory element (UPE), containing rs10085588. Luciferase reporter assays showed transactivation capability with a strong allele-dependent effect. Finally, 4C-seq experiments in osteoblastic cell lines showed that the UPE interacted with different tissue-specific enhancers and a lncRNA (*LOC100506136*) in the region.

In summary, this study is the first one to analyse in depth the functionality of *C7ORF76* genomic region. We provide functional regulatory evidence for the rs10085588, which may be a causal SNP within the 7q21.3 GWAS signal for osteoporosis.

#### Reference:

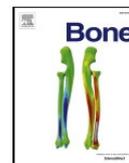
Neus Roca-Ayats, Núria Martínez-Gil, Mónica Cozar, Marina Gerousi, Natàlia Garcia-Giralt, Diana Ovejero, Leonardo Mellibovsky, Xavier Nogués, Adolfo Díez-Pérez, Daniel Grinberg, Susanna Balcells. Functional characterization of the *C7ORF76* genomic

region, a prominent GWAS signal for osteoporosis in 7q21.3. *Bone*. 2019;123:39-47. doi: 10.1016/j.bone.2019.03.014



Contents lists available at ScienceDirect

Bone

journal homepage: [www.elsevier.com/locate/bone](http://www.elsevier.com/locate/bone)

Full Length Article

## Functional characterization of the *C7ORF76* genomic region, a prominent GWAS signal for osteoporosis in 7q21.3



Neus Roca-Ayats<sup>a</sup>, Núria Martínez-Gil<sup>a</sup>, Mónica Cozar<sup>a</sup>, Marina Gerousi<sup>a</sup>, Natàlia Garcia-Giralt<sup>b</sup>, Diana Ovejero<sup>c</sup>, Leonardo Mellibovsky<sup>b</sup>, Xavier Nogués<sup>b</sup>, Adolfo Díez-Pérez<sup>b</sup>, Daniel Grinberg<sup>a,1</sup>, Susanna Balcells<sup>a,\*,1</sup>

<sup>a</sup> Department of Genetics, Microbiology and Statistics, Facultat de Biologia, Universitat de Barcelona, Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), ISCIII, IBUB, IRSJD, Barcelona, Catalonia, Spain

<sup>b</sup> Musculoskeletal Research Group, IMIM (Hospital del Mar Medical Research Institute), Centro de Investigación Biomédica en Red en Fragilidad y Envejecimiento Saludable (CIBERFES), ISCIII, Barcelona, Catalonia, Spain

<sup>c</sup> National Research Council, Institute of Clinical Physiology, Lecce, Italy

### ARTICLE INFO

#### Keywords:

*C7ORF76*  
Osteoporosis  
GWAS signal  
Non-coding regulatory variant  
Enhancer  
4C

### ABSTRACT

Genome-wide association studies (GWAS) have repeatedly identified genetic variants associated with bone mineral density (BMD) and osteoporotic fracture in non-coding regions of *C7ORF76*, a poorly studied gene of unknown function. The aim of the present study was to elucidate the causality and molecular mechanisms underlying the association. We re-sequenced the genomic region in two extreme BMD groups from the BARCOS cohort of postmenopausal women to search for functionally relevant variants. Eight selected variants were tested for association in the complete cohort and 2 of them (rs4342521 and rs10085588) were found significantly associated with lumbar spine BMD and nominally associated with osteoporotic fracture. *cis*-eQTL analyses of these 2 SNPs, together with SNP rs4727338 (GWAS lead SNP in Estrada et al., *Nat Genet.* 44:491–501, 2012), performed in human primary osteoblasts, disclosed a statistically significant influence on the expression of the proximal neighbouring gene *SLC25A13* and a tendency on the distal *SHFM1*. We then studied the functionality of a putative upstream regulatory element (UPE), containing rs10085588. Luciferase reporter assays showed transactivation capability with a strong allele-dependent effect. Finally, 4C-seq experiments in osteoblastic cell lines showed that the UPE interacted with different tissue-specific enhancers and a lncRNA (*LOC100506136*) in the region.

In summary, this study is the first one to analyse in depth the functionality of *C7ORF76* genomic region. We provide functional regulatory evidence for the rs10085588, which may be a causal SNP within the 7q21.3 GWAS signal for osteoporosis.

### 1. Introduction

Genome-wide association studies (GWAS) have been successfully used to identify genetic variants associated with complex traits and diseases, such as osteoporosis. In a few cases, the associated SNPs are located within a coding region of a gene, facilitating its functional evaluation. However, the vast majority of associated SNPs lie in non-coding regions, which make it challenging to understand the functional mechanisms underlying the association [1,2]. In addition, it is highly probable that the associated SNPs are in linkage disequilibrium (LD)

with the causal variant.

To date, many GWAS have been performed to find genetic association with bone mineral density (BMD) and osteoporotic fracture [3–12]. BMD is a genetically determined, extensively measured quantitative trait (heritability of 0.5–0.85) and, therefore, a good marker for bone status. Low-trauma fracture, the clinical outcome of osteoporosis, is also heritable, albeit to a lesser extent (heritability of 0.54–0.68) [13]. These GWAS have identified > 500 candidate *loci* [12], although the causal variants remain largely unknown. In addition, all the GWAS findings together only explain a small proportion (~20%) of the total

\* Corresponding author at: Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona. Av. Diagonal, 643. 08028 Barcelona, Catalonia, Spain.

E-mail address: [sbalcells@ub.edu](mailto:sbalcells@ub.edu) (S. Balcells).

<sup>1</sup> Last co-authors.

<https://doi.org/10.1016/j.bone.2019.03.014>

Received 29 January 2019; Received in revised form 4 March 2019; Accepted 12 March 2019

Available online 14 March 2019

8756-3282/ © 2019 Elsevier Inc. All rights reserved.

genetic impact on BMD [12]. Some associated *loci* contain genes not previously known to play a role in bone biology, which is the case of *C7ORF76* (AKA *FLJ42280*, currently annotated as transcript variant 6 of gene *SEM1* in GRCh38), in the 7q21.3 genomic region. Several SNPs within this region have been found significantly associated with both lumbar spine (LS) and femoral neck (FN) BMD as well as with osteoporotic fracture in different GWAS and meta-analyses [4,6–8,14–16]. In particular, in the largest meta-analysis carried out so far [7], the lead SNP of this *locus* (rs4727338) was one of the genome-wide top-associated signals. Yet, *C7ORF76* is a poorly studied gene of unknown function and there are other genes within this region that could also be responsible for the associations observed.

In this study, we aimed at elucidating the causality and molecular mechanisms underlying the strong association identified in the *C7ORF76* genomic region. We have deeply re-sequenced the genomic region in two extreme BMD groups of postmenopausal women of the BARCOS cohort and selected some variants to analyse their association in the full cohort. Through a combination of several *in silico* and experimental approaches, we studied a possibly causal SNP located in a regulatory region and demonstrated its functionality.

## 2. Material and methods

### 2.1. Study cohort

The BARCOS cohort consisted of 1490 postmenopausal women of Spanish descent from the Barcelona area, monitored at the Hospital del Mar (Barcelona, Spain). Exclusion criteria were any history of bone diseases, metabolic or endocrine disorders, hormone-replacement therapy, or use of drugs that could affect bone mass. BMD of all participants was measured at LS and FN by dual energy X-ray absorptiometry (DXA). The following data were also recorded: age, age of menarche and menopause, number of fractures and anthropometric measures such as weight and height. DNA is available from all samples of the cohort. Details of the cohort and DNA extraction have been described previously [17,18]. Written informed consent was obtained from all patients in accordance with the regulations of the Clinical Research Ethics Committee of Parc de Salut Mar, which approved the study. All experiments and protocols were approved by the Bioethics Committee of Universitat de Barcelona (IRB00003099).

For the re-sequencing of the *C7ORF76 locus*, two extreme LS-BMD groups were selected from the BARCOS cohort, using the statistical Z-score. The 50 women with the highest Z-score values (from 2.98 to 0.73) were included in the 50-H group and the 50 women with the lowest Z-score values (from –2.41 to –4.26) were included in the 50-L group.

### 2.2. Re-sequencing

Re-sequencing of the *C7ORF76* (ENSG00000197851; ENST00000356686.1) genomic region was performed in the 100 individuals of the two extreme subgroups of the BARCOS cohort, according to the Z-score. A 28 kb region (chr7:96,108,695–96,136,619; GRCh37), including the *C7ORF76* gene and the 3.8 kb upstream and 2 kb downstream regions of the gene, was amplified in 7 overlapping fragments by Long Range-PCR (Supplementary Table 1). All amplicons were purified and quantified using the Quant-iT PicoGreen dsDNA Reagent and Kit (Life Technologies, Thermo Fisher) before pooling them equimolarly into two groups, one of HBM and another of LBM. Both pools were tagged with a MID adaptor and an emulsion-PCR was carried out prior to massive parallel sequencing at 3600× coverage per pool with Roche's 454 GS Junior System. The massive parallel sequencing was carried out in the Genomics facilities of the Universitat de Barcelona. The raw data obtained were processed to trim the MIDs, using a custom pipeline, and were mapped against the reference genome (GRCh37), using the GS Mapper software (Roche). Mapped

reads were filtered, sorted and indexed using SAMtools [19]. Single Nucleotide Variants (SNVs) and indels were identified using GATK standard hard filtering parameters [20]. The variants were filtered according to the following criteria: coverage  $\geq 1,000$  reads, variants present in  $\geq 1\%$  of the reads per pool and low strand bias. The number of reads of a variant was normalised with its coverage and the variants were classified according to minor allele frequency (MAF): Common (MAF  $\geq 5\%$ ), and lower frequency variants (MAF  $< 5\%$ ). The variants were validated either by differential digestion with restriction enzymes or by high resolution melting, using the Light Cycler® 480 ResoLight Dye (Roche).

### 2.3. In silico functional analyses and motif analysis

*In silico* functional analyses consisted in annotating the European and Iberian MAF of the variants obtained from dbSNP and 1000 Genomes, when available; predicting the pathogenicity of exonic variants, using SIFT [21], PolyPhen [22] and Mutation Taster [23]; and, for the intronic variants, analysing the DNase I hypersensitivity, histone modifications, transcription factor binding, miRNAs binding, etc. of the regions of interest. All the *in silico* data was obtained from ENCODE [24], International Human Epigenome Consortium [25], The Roadmap Epigenomics Project, FANTOM5 [26], HaploReg [27], RegulomeDB [28], miRTarBase [29], miRdSNP [30], MirSNP [31], BioMart, and Ensembl and UCSC Genome Browser. All variants were analysed with the Variant Effect Predictor from Ensembl, the Variant Annotation Integrator from UCSC, and FuncPred from the National Institute of Environmental Health Sciences. Transcription factor binding sites prediction considering the different alleles of the variants was done using MatInspector [32] and the Bioconductor “motifbreakR” package (<https://bioconductor.org/packages/release/bioc/html/motifbreakR.html>) [33] using default method settings (weighted sum) and a p-value cut-off at  $5 \times 10^{-5}$  for SNPs. Super-enhancer data was obtained from a catalogue of super-enhancers in 86 human cell and tissue samples [34] and from SEdb [35].

### 2.4. SNV genotyping

Genotyping of 8 selected variants in the complete BARCOS cohort was carried out at LGC Genomics (Hoddesdon, UK). In addition, as BARCOS was included in the replication phase of the meta-analysis by Estrada et al. [7], the genotyping results of the SNP rs4727338 were also available. The genotyping of 6% of the samples was performed in duplicate, as a genotyping quality control, and showed a concordance above 99%.

### 2.5. Linkage disequilibrium analysis

The Haploview software [36] was used to calculate and represent the degree of linkage disequilibrium between the genotyped common variants using the default parameters.

### 2.6. Cell culture

The human osteosarcoma cell line Saos-2 was used for luciferase reporter assays and 4C-seq assays. It was obtained from the American Type Culture Collection (ATCC® HTB-85™) and grown in Dulbecco's Modified Eagle Medium (DMEM; Sigma-Aldrich), supplemented with 10% Fetal Bovine Serum (FBS; Gibco, Life Technologies) and 1% penicillin/streptomycin (Gibco, Life Technologies), at 37 °C and 5% of CO<sub>2</sub>. Human fetal osteoblasts (hFOB) 1.19 cells were used for 4C-seq assays. They were obtained from ATCC (ATCC® CRL-11372™) and grown in DMEM:F12 (1:1) medium without phenol red (Gibco, Life Technologies), supplemented with 10% FBS and 0.3 mg/ml Geneticin (Gibco, Life Technologies), at 34 °C and 5% of CO<sub>2</sub>. Human medulla-derived mesenchymal stem cells (MSCs) were also used for 4C-seq

assays. They were kindly provided by Dr. José Manuel Quesada Gómez, from Instituto Maimónides de Investigación Biomédica, Hospital Universitario Reina Sofía, Córdoba, Spain. They were grown in alpha-MEM medium (Gibco, Life Technologies), supplemented with 10% FBS, 1% penicillin/streptomycin and  $1 \times$  Glutamax (Gibco, Life Technologies), at 37 °C and 5% of CO<sub>2</sub>. Human primary osteoblasts (hOB) were used for eQTL assays. They were obtained from trabecular bone of women who underwent knee replacement due to osteoarthritis and who did not have any other pathology that could affect the bone status. Bony tissue was cut up into small pieces, washed in phosphate buffered saline (PBS; Gibco, Life technologies) to remove non-adherent cells, and placed on a 140 mm culture plate. Samples were cultured in DMEM supplemented with 10% FBS, 100 U/ml penicillin/streptomycin, 0.4% fungizone (Gibco, Life Technologies) and 100 µg/ml ascorbic acid (Sigma-Aldrich). DNA and RNA extractions were performed at maximum passage 2. HeLa and HEK293 cell lines were obtained from ATCC (ATCC® CCL-2™ and ATCC® CRL-1573™, respectively) and grown in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin at 37 °C and 5% CO<sub>2</sub>.

### 2.7. Human primary osteoblasts DNA extraction and sequencing

DNA was extracted from cultured hOBs using the Wizard® Genomic DNA Purification Kit (Promega), according to manufacturer's instructions. The concentration of the purified DNA was analysed in a spectrophotometer (Nanodrop). Genotypes for rs4727338, rs10429035, rs12674052, rs4342521, and rs10085588 were assessed by Sanger sequencing using the BigDye® Terminator v3.1 (Applied Biosystems) in the Genomics facilities of Universitat de Barcelona. Primers (Invitrogen, Thermo Fisher) were designed using the Primer3 Input 0.4.0 (Supplementary Table 2).

### 2.8. Human primary osteoblasts RNA extraction, retrotranscription and qPCR

RNA was extracted from cultured hOBs using the High Pure RNA Isolation kit (Roche), according to manufacturer's instructions. RNA was quantified using a Nanodrop spectrophotometer and retrotranscribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Thermo Fisher), according to the specifications of the manufacturer. qPCR was performed using UPL Probes (Roche) on a LightCycler 480 Instrument II (Roche). Expression of *HMBS* was used as a normalizing control, and fold changes (FC) were calculated by relative quantification, using the 2nd derivative method. Primers used to amplify the neighbouring genes of the *C7ORF76* locus are summarised in Supplementary Table 3.

### 2.9. Amplification of an UPstream regulatory Element (UPE)-derived transcript

UPE-derived transcript was amplified from HeLa, HEK293, human primary osteoblasts and Saos-2 cDNA (200 ng) by PCR, using the following primers: 5'-CACTTTTCAAATCCCACCTG-3' and 5'-TGAGAGCTGCTTAGAAATGGAA-3'. PCR products were run in a 2% agarose gel.

### 2.10. Luciferase reporter constructs and site-directed mutagenesis

The 750 bp-fragment containing the UPE was PCR-amplified from human genomic DNA using the following primers: 5'-CACTTTTCAAATCCCACCTG-3' and 5'-TGAGAGCTGCTTAGAAATGGAA-3', and cloned in both orientations using *XhoI* and *KpnI* restriction enzymes in the pGL3-Basic vector (Promega). The minor allele of the rs10085588 (A) was introduced with the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent), following the manufacturer instructions. All the plasmids were validated by Sanger sequencing.

### 2.11. In vitro luciferase assay

Saos-2 cells were seeded at a density of  $3.0 \times 10^5$  cells per well in a 6-well plate. After 24 h, they were transfected with 2.2 µg of total DNA per well using FuGENE HD reagent (Promega), according to manufacturer's instructions. Two plasmids were cotransfected in each well: the pGL3-Basic empty or with the UPE fragment cloned upstream of the Firefly Luciferase coding region and the pRL-TK plasmid, containing the Renilla Luciferase gene, in a proportion of 1/10. Forty-eight hours after transfection, cells were rinsed with PBS and lysed. The luciferase activity was measured using a Glomax Multi+ luminometer (Promega), with the Dual-Luciferase® Reporter Assay System reagents (Promega). Each experiment was performed in two biological replicates and was repeated 5 times.

### 2.12. 4C-seq

4C-seq was carried out at the Functional Genomics Service of the Centro Andaluz de Biología del Desarrollo (Sevilla, Spain). 4C-seq libraries were generated from Saos-2, hFOB 1.19 and hMSCs lines as described previously [37,38]. 4-bp cutters were used as primary (*DpnII*) and secondary (*Csp6I*) restriction enzymes. For each cell line, a total of 1.6 µg of library was amplified by PCR (primers used: CTGGAAGAGTCCCAGGGATC and AATGGAAGAGTGGAGATTCAGG; chr7:96,137,244-96,137,535). Samples were sequenced with Illumina Hi-Seq technology according to standard protocols at the Genomics Service of the Centro Nacional de Investigaciones Cardiovasculares (Madrid, Spain). 4C-seq data were analysed as described previously [39]. Briefly, raw sequencing data were demultiplexed and mapped to the corresponding reference genome (GRCh37). Reads located in fragments flanked by two restriction sites of the same enzyme, in fragments smaller than 40 bp or within a window of 10 kb around the viewpoint were filtered out. 4C-seq data were normalised by the total weight of reads within  $\pm 2$  Mb around the viewpoint. The experiments were carried out in one biological replicate.

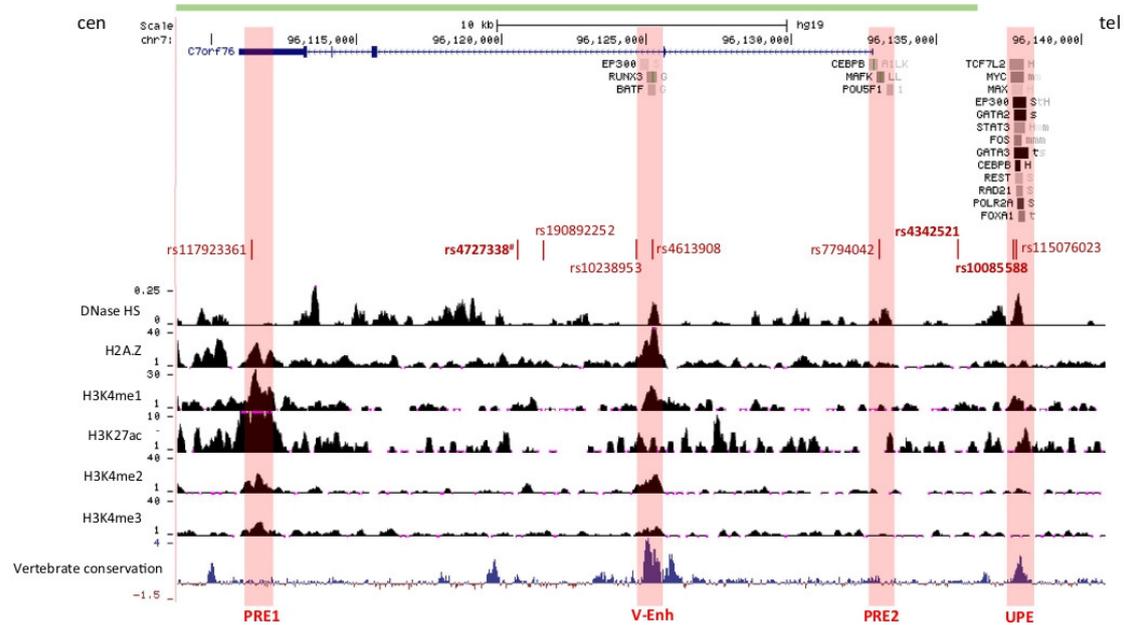
### 2.13. Topologically associating domain (TAD) analysis

TAD data on different cell types from Dixon et al. [40] was collected from the 3D Genome Browser (<http://promoter.bx.psu.edu/hi-c/>) [41] and displayed using the UCSC Genome Browser. The 3D Genome Browser was also used to visualise published Hi-C data.

### 2.14. Statistical methods

Statistical analyses were performed using the R software version 3.4.1. Hardy-Weinberg equilibrium (HWE) was calculated using Chi-square test and p-values < 0.01 were considered significant. Fisher's exact test with Bonferroni correction for multiple testing was used to statistically compare the differences of the genotype frequency of the common variants in each extreme group (HBM and LBM). Linear regression analysis adjusted by years since menopause (YSM) was performed to determine the association between BMD and the genotype of each SNV in the complete BARCOS cohort. Linear regression was also used to assess the association between gene expression levels and genotypes (*cis*-eQTL) in primary osteoblasts. All analyses were performed using the *SNPassoc* package testing the additive, recessive and dominant models. p-values < 0.05 were considered significant. Correction for multiple testing was performed using the Bonferroni's method for the number of SNPs tested.

Relative luciferase units (RLU, i.e. the ratio of the firefly luciferase activity over the Renilla luciferase activity) were calculated for each individual measurement and a one-way blocked ANOVA with TukeyHSD post-hoc test was performed. All the data was ascertained for normality, homoscedasticity and atypical data points and p-values < 0.05 were considered significant.



**Fig. 1.** Schema of *C7ORF76* locus. In green, region re-sequenced in the present study. In dark red, SNPs genotyped in the BARCOS cohort and in bold, SNPs found associated with BMD; # SNP previously genotyped (Estrada et al. [7]). Transcription factor ChIP-seq, DNase HS, H2A.Z, H3K4me1, H3K27ac, H3K4me2 and H3K4me3 ChIP-seq from osteoblasts (ENCODE), and vertebrate conservation are shown. The red boxes represent putative regulatory elements, according to epigenetic marks as well as transcription factors binding. Note that V-Enh corresponds to an enhancer (hs2311) described in VISTA Enhancers database [42]. PRE1 = putative regulatory element 1; PRE2 = putative regulatory element 2; V-Enh = VISTA enhancer hs2311; UPE = 4 kb upstream putative regulatory element.

**3. Results**

**3.1. Re-sequencing of *C7ORF76* locus in extreme BMD groups of the BARCOS cohort and variant functional annotation**

We re-sequenced 28 kb of the *C7ORF76* genomic region (chr7:96,108,695-96,136,619; GRCh37) in the 50 women with the highest and 50 women with the lowest LS-BMD of the BARCOS cohort (Fig. 1). Total number and location of single nucleotide variants (SNVs) detected before and after filtering and validating are shown in Table 1. Fifty-one common variants (MAF above 5%) and 59 lower frequency variants (MAF below 5%) were identified. The lower frequency variants were equally distributed between both extreme groups. To better assess the importance of the variants, we explored their functionality using publicly available *in silico* data (transcription factor binding, DNase I hypersensitivity, conservation, miRNAs binding, and histone marks). Twenty-eight variants were found in putative regulatory regions, 4 of which were located in osteoblast regulatory regions (rs9785005, rs10238953, rs4613908 and rs117923361) and 1 of them (rs4613908) was found in an active osteoblast enhancer. One missense variant was also identified but it was predicted to be tolerated by SIFT, to be benign by Polyphen and to be a polymorphism according to MutationTaster. None of the variants was predicted to affect miRNAs binding. Only 1

variant (rs4342521) showed nominal significance between the genotype frequencies of the two extreme groups (Fisher's exact test,  $p = 0.0382$ ). This SNP is located 3 kb upstream of the *C7ORF76* gene and the minor allele (T) was overrepresented in the LBM group.

**3.2. Association of variants with BMD in the complete cohort**

We selected 8 interesting variants, taking into consideration both frequency and functionality (Fig. 1), including 2 located in a putative regulatory region 4 kb upstream of the *C7ORF76* transcription start site (TSS), which we named UPE, that was not included in the re-sequencing. These 8 variants were genotyped in the complete BARCOS cohort ( $n = 1490$ ) to test their association with BMD and osteoporotic fracture. For all of them, we obtained MAF values similar to those found in the 1000 Genomes database for the European or Iberian populations (Supplementary Table 4). Significant differences were obtained with two of the 5' upstream SNPs, rs4342521 and rs10085588 (Fig. 1), for LS-BMD under additive and recessive models and nominal differences were obtained with the same SNPs for osteoporotic fracture, under additive and dominant models (Table 2). In all cases, the minor allele (T and A, respectively) had a BMD-lowering effect on LS and had a damaging effect on osteoporotic fracture. The two associated SNPs were found to be in linkage disequilibrium with the GWAS hit (rs4727338)

**Table 1**  
Number and location of single nucleotide variants found in this study.

	Raw	Selected by filtering and validating	Coding	Regulatory regions <sup>a</sup>	Putative osteoblast regulatory elements	Active enhancer
Common variants	96	51	0	12	3	1
LFV	24,243	59	1	16	1	0
Total	24,339	110	1	28	4	1

<sup>a</sup> Either present in flanking regions, 5'UTR, 3'UTR or introns; LFV: lower frequency variants (MAF < 0.05).

**Table 2**  
Association between common and LFV variants of *C7ORF76* and LS-BMD, FN-BMD, FN-BMD or osteoporotic fracture.

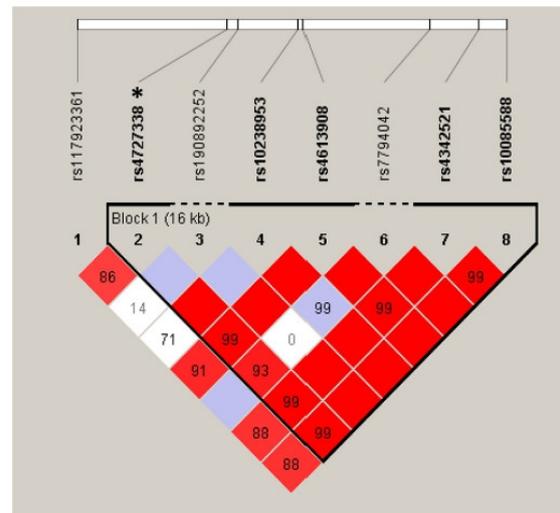
SNP	Genomic position (GRCh37)	Type of variant	p-value LS-BMD		Effect size (β coeff; 95% CI)		p-value FN-BMD		Effect size (β coeff; 95% CI)		p-value osteoporotic fracture		Effect size (OR; 95% CI)	
			A	R	A	R	A	R	A	R	A	R	A	R
rs115076023 <sup>a</sup>	7:96137731 G > A	5'UP	-	-	-0.0324 (-0.0541, 0.0108)	0.05032	-	-	-	-	0.08554	0.32373	0.02242	1.51 (1.05, 2.17)
rs10085588	7:96137674 G > A	5'UP	<b>0.00361</b>	<b>0.00343</b>	-0.0332 (-0.0547, -0.0117)	0.05233	0.12419	0.11909	0.30848	0.03911	0.39100	0.01974	1.53 (1.06, 2.19)	
rs4342521	7:96136005 G > T	5'UP	<b>0.00314</b>	<b>0.00251</b>	0.05233	0.05233	0.12453	0.09115	0.36111	0.03911	0.39100	0.01974	1.53 (1.06, 2.19)	
rs7794042	7:96132999 C > A	5'UP	0.68070	0.65857	0.63405	0.63405	0.85483	0.66330	0.89686	0.55917	1.00000	0.30987	0.30987	
rs4613908	7:96125315 G > A	I	0.06164	0.08889	0.15109	0.15109	0.55060	0.30356	0.86063	0.58472	0.27134	0.18380	0.18380	
rs10238953	7:96124975 T > C	I	0.57370	0.49214	0.69433	0.69433	0.71522	0.57609	0.83130	0.14074	0.48140	0.15002	0.15002	
rs190892252	7:96121343 A > G	I	0.21722	-	-	-	0.24861	-	-	0.63625	-	-	-	
rs4727338 <sup>b</sup>	7:96120675 G > C	I	<b>0.00141</b>	<b>0.00207</b>	0.02430	0.02430	0.06943	0.04540	0.27708	0.07618	0.56459	0.03270	1.50 (1.03, 2.18)	
rs117923361	7:96111486 C > A	3'UTR	0.82861	0.68376	0.87501	0.87501	0.08217	0.62988	0.08427	0.30507	1.00000	0.22534	0.22534	

Values in bold indicate statistical significance; values in italics indicate nominal significance.

A: Additive model; R: Recessive model; D: Dominant model.

<sup>a</sup> SNP found monomorphic in the BARCOS cohort.

<sup>b</sup> SNP previously genotyped in the BARCOS cohort (Estrada et al. [7]).



**Fig. 2.** Linkage disequilibrium plot of *C7ORF76* genotyped variants. \*GWAS hit from Estrada et al. [7]. The numbers within the squares and the colour scale both refer to D'/LOD values (with bright red: D' = 1 and LOD ≥ 2; white: D' < 1 and LOD < 2; blue: D' = 1 and LOD < 2 and shades of pink/red: D' < 1 and LOD ≥ 2).

from Estrada et al. [7] (Fig. 2), which was also significantly associated with LS-BMD in the BARCOS cohort (Table 2).

### 3.3. cis-eQTL analyses

To evaluate the functionality of the associated variants (rs4727338 –GWAS hit, rs4342521 and rs10085588) we first tested them as cis-eQTLs in human primary osteoblasts, a cell type unavailable in GTEx. Of note, we failed to detect *C7ORF76* mRNA expression in these cells. None of the SNPs were found to be eQTLs for *DLX5*, *DLX6*, *DLX6-AS1*, or *SHFM1* (*SEMI* transcript variant 5 in GRCh38), although a tendency was observed for the latter (Table 3). On the contrary, the 3 SNPs were found nominally associated with *SLC25A13* gene expression, where the minor alleles were associated with decreased gene expression. We also tested another SNP of the region (rs10429035) found associated with BMD in a previous GWAS meta-analysis [8] and described in GTEx to be eQTL for *SHFM1* in tibial artery. However, we failed to detect an association of this SNP with gene expression of any of the nearby genes (data not shown).

### 3.4. Evaluation of the regulatory capability of the UPE

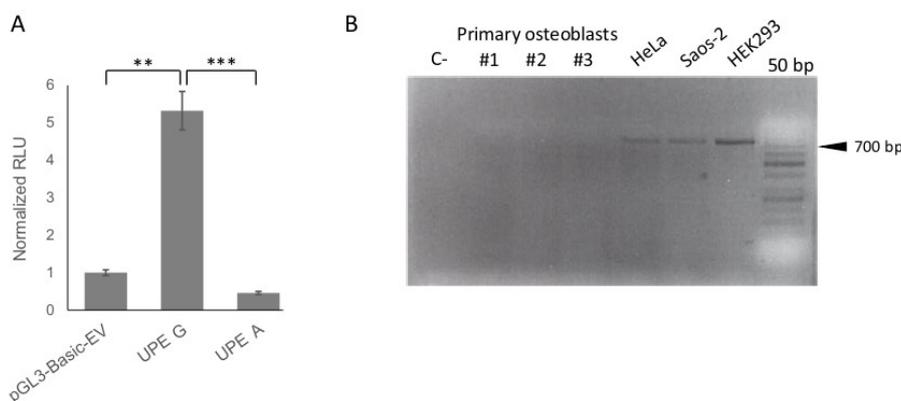
Next, as the associated SNP rs10085588 is located in a putative regulatory region 4 kb upstream of the *C7ORF76* gene (UPE; Fig. 1), we assessed the functionality of the UPE. We performed luciferase reporter

**Table 3**  
cis-eQTL analysis of 3 *C7ORF76* variants associated with BMD.

SNP	p-values				
	<i>DLX5</i>	<i>DLX6</i>	<i>DLX6-AS1</i>	<i>SHFM1</i>	<i>SLC25A13</i>
rs10085588	0.14207	0.53849	0.59125	0.05270	<i>0.03350</i>
rs4342521 <sup>a</sup>	0.09227	0.34167	0.55062	0.05725	<i>0.01442</i>
rs4727338 <sup>b</sup>	0.09227	0.34167	0.55062	0.05725	<i>0.01442</i>

Values in italics indicate nominal significance.

<sup>a</sup> p-values are identical as a reflection of the high LD between the 2 SNPs.



**Fig. 3.** (A) Luciferase activity of different versions of UPE, containing the G or A allele for the SNP rs10085588, in a forward or inverted orientation, in Saos-2 cells. Results are expressed as mean  $\pm$  SD. \*\*p-value < 0.01; \*\*\*p-value < 0.001 (B) PCR amplification of UPE from cDNA of different cell types. Expected size: 750 bp.

assays in Saos-2 cells to test UPE activity in both orientations and the effects of the two alleles of rs10085588. As shown in Fig. 3A, the forward UPE construct bearing the major allele (G) showed significantly increased luciferase expression compared to the empty vector (FC: 5.3,  $p = 0.0097$ ), and compared to the construct bearing the minor allele (A) (FC: 11.6,  $p < 0.001$ ). No activity was detected when the UPE was tested in the inverse orientation (data not shown). Results of luciferase assays were consistent with eQTL analyses, in the sense that the minor allele (A) fails to activate transcription and is associated with lower expression of *SLC25A13* in osteoblasts (see above). We also evaluated whether the UPE was transcribed (as occurs in many regulatory regions) by performing RT-PCR in cDNA of human primary osteoblasts, Saos-2, HeLa and HEK293 cells. We were able to amplify the UPE sequence from cDNA of HeLa, HEK293 and Saos-2 cells but we failed to amplify it from different cDNA samples of human primary osteoblasts (Fig. 3B). We also interrogated the FANTOM5 Cap Analysis of Gene Expression (CAGE) dataset [26] for evidence of UPE transcription and we could observe a signal for TSS expression in smooth muscle cells.

### 3.5. Chromatin interactions from the UPE

Finally, we investigated the possible genomic targets of the UPE by examining the 3D chromatin interactions by 4C-seq in different cell types (MSCs, hFOB 1.19 and Saos-2). We detected interaction between UPE and the genomic region spanning approximately 750 kb on each side of it (Fig. 4), and no other interactions were detected elsewhere in the genome. Notably, we found higher interaction levels with most of the tissue-specific enhancers described in the VISTA browser [42,43], especially in the hFOB 1.19 cell line. We could also detect higher interaction with the long non-coding RNA gene *LOC100506136*, upstream of *C7ORF76*. In addition, we analysed the topologically associated domains (TADs) of the region on different cell types using available Hi-C data [40] and the 3D Genome Browser [41] and we observed that in many cell types, the TAD containing the gene *C7ORF76* spanned from approximately 50 kb upstream of *DYNC111* TSS to approximately 25 kb upstream of *ACN9* TSS (Fig. 4), consistent with the 4C-seq results.

## 4. Discussion

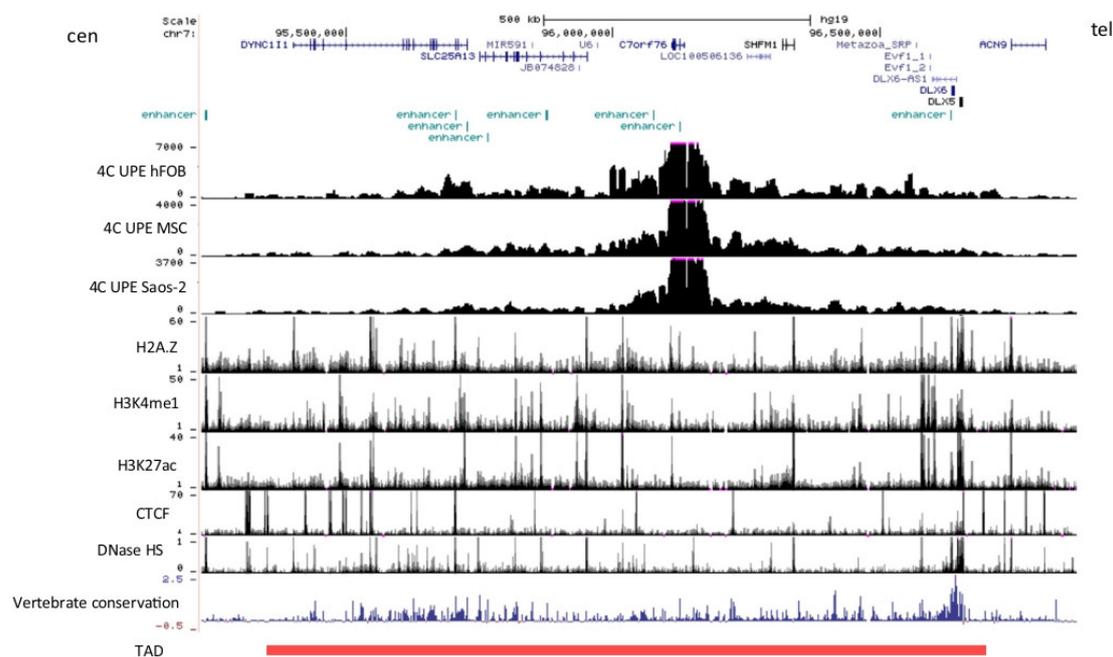
Different non-coding variants in the *C7ORF76* genomic region have been previously associated with BMD and osteoporotic fracture in different GWAS [4,6–8,14–16] and at this point, deciphering the functionality of this region would be the next logical step for understanding these associations. In this line, we have analysed the *C7ORF76* region in depth, including re-sequencing, testing variants for association in the

BARCOS cohort, and analysing the functionality of the variants and regulatory regions both *in silico* and experimentally. Two upstream variants (rs4342521 and rs10085588) were found significantly associated with LS-BMD and nominally associated with osteoporotic fracture. In addition, both SNPs have been identified as eQTLs of *SLC25A13* in human primary osteoblasts. Moreover, the SNP rs10085588 falls in a regulatory region (UPE) able to stimulate transcription in an allele-dependent manner. This UPE was found to interact with different tissue-specific enhancers and a lncRNA present in the nearby region.

The *C7ORF76* gene is an uncharacterised gene of unknown function without expression data in GTEx. In the current human genome assembly (GRCh38) it is annotated as alternative transcript 6 of *SEMI* (encoding a 26S proteasome complex subunit, whose alternative historical name is *SHFM1* for Split Hand and Foot Malformation 1). However, although RefSeq currently labels it as curated gene, surprisingly few data have been gathered and the real function of *C7ORF76* remains elusive. This is in contrast with the consistent finding of potent GWAS signals for osteoporosis within this gene. Of note, this gene seems to be not expressed in primary osteoblasts, Saos-2, HeLa and SH-SY5Y (data not shown) which suggests that either it might, indeed, not be an osteoblast gene, yet regulate BMD from a different cell type or organ, or, alternatively, other genes in the region could be causal for the association.

Most GWAS variants for complex diseases are located in non-coding regulatory regions (reviewed in Zhang et al. [44]) and several studies have pinpointed the importance of regulatory elements for the susceptibility to osteoporosis [45,46]. Moreover, for some common traits, it has been described that several causal variants exist in a single LD block, located in multiple enhancers that cooperatively influence gene expression (the so called super-enhancers) [34,47]. These enhancer clusters are highly cell type specific. In this respect, although publicly available datasets [34,35] did not consider the *C7ORF76* genomic region to be a super-enhancer, the functional annotation of this region revealed the presence of several putative regulatory elements containing SNPs prone to confer susceptibility to osteoporosis.

To elucidate the molecular mechanisms underlying the strong association, further exploration with high-coverage sequencing to prioritise potentially causal variants was necessary [48]. We performed an extreme-truncate selection of the BARCOS cohort as a discovery phase, prior to genotyping the selected variants in the complete cohort. The efficiency of extreme-truncate selection approach for quantitative trait association studies (e.g. BMD) has been widely proven, as well as its utility for detecting rare variants [6,49,50]. The 2 SNPs found associated in the BARCOS cohort in this study (rs4342521 and rs10085588) were previously found associated with LS-BMD, FN-BMD and



**Fig. 4.** 4C-seq using UPE as viewpoint in human fetal osteoblast (hFOB) 1.19, mesenchymal stem cells (MSC) and Saos-2 cell line. H2A.Z, H3K4me1, H3K27ac and CTCF ChIP-seq and DNase HS from osteoblasts (ENCODE data), and vertebrate conservation are shown. Experimentally validated active enhancers from VISTA Enhancer Browser [42] are shown in green. In red, topologically associated domain (TAD).

osteoporotic fracture in other GWA studies [4,9]. In contrast, we failed to find association in the rest of the variants interrogated, although other SNPs in this region have been found associated with LS-BMD, FN-BMD, and osteoporotic fracture [6,8,16], as well as with heel BMD [10,12] and total body BMD [11]. Taken together, these association data are consistent with the existence of a large LD block encompassing all the associated variants and the comparatively small sample size of the BARCOS cohort may preclude the detection of some of them.

To further delineate the role of the three associated variants (the two mentioned above and the GWAS SNP in Estrada et al. [7]), we performed *cis*-eQTL analyses in human primary osteoblasts and we detected a nominal association between the minor alleles of the three SNPs and decreased *SLC25A13* gene expression. We also detected a trend for association with decreased expression of *SHFM1*. These results reflect the LD among the SNPs. We failed to find an association between SNP rs10429035 and transcription levels of *SHFM1*, while in GTEx this SNP is described as eQTL for this gene in tibial artery. It could be that it is not eQTL in primary osteoblasts. Alternatively, our limited power ( $n = 45$ ) and the fact that primary cells are not as homogeneous as cell lines may have prevented us to detect it in primary osteoblasts [51].

Out of the 3 associated SNPs, only one lies within a regulatory region (UPE, see Fig. 1), while the others map to sequences lacking regulatory marks. We set out to experimentally study the UPE, which is a conserved region located approximately 4 kb upstream of *C7ORF76* TSS, with enhancer marks such as H3K27ac and H3K4me1, as well as a DNase hypersensitivity signal. According to ENCODE ChIP-seq data, many transcription factors bind to UPE, among which RAD21, MYC, POLR2A, and the P300 histone acetyltransferase, known to be involved in transcription regulation, initiation and elongation and in enhancer activity [52,53]. Our results, including luciferase assays, RT-PCR and 4C analyses, indicate that, indeed, the UPE acts as a regulatory element, able to activate transcription of a reporter gene. It is well known that many such elements can be transcribed [54–56], producing non-coding

RNAs (or eRNAs for enhancer-RNAs) and, in this sense, we found UPE to be transcribed in Saos-2 cells.

Our results also provide clear evidence that SNP rs10085588, within the UPE, is itself functional, since we showed that the minor allele (A) abolished luciferase activation, which was otherwise stimulated by the major allele G. We have performed an *in silico* analysis and found predictions that the histone deacetylase HDAC2 may bind the A allele more probably than the G (p-value: 5.30016e-06), a possible explanation for the reduced expression observed.

The 4C-seq experiments showed that, in the cells included in this study, the UPE interacted with several sites within the TAD where it belongs (see Fig. 4), and nowhere else in the genome. This suggests that the biological function of the GWAS signal should be limited to genes within this TAD, further supported by strong CTCF signals limiting the region. However, interactions with other genomic regions may exist in different tissues or differentiation stages, not tested by us. In this sense and interestingly, a computationally-based characterization of osteoporosis associated SNPs identified an interaction between the GWAS SNP rs4729260 in the *C7ORF76* region and the Xq12 genomic region [57], which contains the androgen receptor gene, known to be involved in bone metabolism [58].

Within the TAD, the UPE enhancer interacted with many other tissue-specific enhancers previously identified [42,43], suggesting that they may act cooperatively or redundantly in regulating gene expression. Several of these enhancers have been shown to affect *DLX5/6* gene expression [43]. However, we have not detected strong interaction signals with the *DLX5/6* region and, in our eQTL study in primary osteoblasts, we did not observe any effect of the associated SNPs on *DLX5/6* gene expression. Likewise, we detected modest interactions between the UPE and the *SHFM1* and the *SLC25A13* coding regions, and negligible signals with that of *DYNC111*. In contrast, we did detect a marked interaction with a lncRNA in the close vicinity of *C7ORF76*, namely *LOC100506136*. Interestingly, a recent study found a SNP

within *LOC100506136* to be one of the 2 out of approximately 55,000 genome-wide SNPs in lncRNAs associated with total hip BMD [59]. Another recent publication found a LS-BMD associated signal within *LOC100506136* in a Mexican-Mestizo cohort [60]. These studies suggest that this lncRNA could be involved in osteoporosis pathogenesis.

This work has some limitations including the sample size of the BARCOS cohort, and of the primary osteoblasts used for the eQTL analyses, both of which preclude the detection of variants with smaller effects. In addition, *C7ORF76* was not tested in other bone cells such as osteoclasts. However, these results provide interesting data to understand the functionality of this unexplored region, one of the most frequently found associated with osteoporosis in GWA studies. They also highlight the importance of regulatory variants in bone phenotypes and the usefulness of integrative approaches to uncover their functionality.

## 5. Conclusions

In summary, this study is the first one to analyse in depth the functionality of the *C7ORF76* genomic region, associated with BMD and osteoporotic fracture in many GWAS. We provide functional regulatory evidence for rs10085588, which may be a causal SNP within this 7q21.3 GWAS signal for osteoporosis.

## CRedit authorship contribution statement

**Neus Roca-Ayats:** Investigation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing. **Núria Martínez-Gil:** Investigation, Formal analysis, Writing - review & editing. **Mónica Cozar:** Investigation, Writing - review & editing. **Marina Gerousi:** Investigation, Writing - review & editing. **Natàlia Garcia-Giralt:** Conceptualization, Writing - review & editing. **Diana Ovejero:** Conceptualization, Writing - review & editing. **Leonardo Mellibovsky:** Conceptualization, Writing - review & editing. **Xavier Nogués:** Conceptualization, Writing - review & editing. **Adolfo Díez-Pérez:** Conceptualization, Writing - review & editing. **Daniel Grinberg:** Conceptualization, Funding acquisition, Formal analysis, Supervision, Writing - original draft, Writing - review & editing. **Susanna Balcells:** Conceptualization, Funding acquisition, Formal analysis, Supervision, Writing - original draft, Writing - review & editing.

## Acknowledgements

We thank E. Czwan for relevant technical assistance. This work was supported by the following grants: SAF2014-56562-R, SAF2016-75948-R (Spanish MINECO), 2014SGR932 (Generalitat de Catalunya), CIBERER (U720), and FEIOMM Investigación 2014. NRA is a recipient of an FPU predoctoral fellowship from the Spanish Ministerio de Educación, Cultura y Deporte; NMG is a recipient of a FI predoctoral fellowship from Generalitat de Catalunya.

## Declaration of interests

The authors declare they have no conflict of interests.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bone.2019.03.014>.

## References

- [1] M.T. Maurano, R. Humbert, E. Rynes, R.E. Thurman, E. Haugen, H. Wang, et al., Systematic localization of common disease-associated variation in regulatory DNA, *Science* 337 (2012) 1190–1195, <https://doi.org/10.1126/science.1222794>.
- [2] Y.G. Tak, P.J. Farnham, Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome, *Epigenetics Chromatin* 8 (2015) 57, <https://doi.org/10.1186/s13072-015-0050-4>.
- [3] J.B. Richards, F. Rivadeneira, M. Inouye, T.M. Pastinen, N. Soranzo, S.G. Wilson, et al., Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study, *Lancet* 371 (2008) 1505–1512, [https://doi.org/10.1016/S0140-6736\(08\)60599-1](https://doi.org/10.1016/S0140-6736(08)60599-1).
- [4] F. Rivadeneira, U. Styrkarsdóttir, K. Estrada, B.V. Halldórsson, Y.H. Hsu, J.B. Richards, et al., Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies, *Nat. Genet.* 41 (2009) 1199–1206, <https://doi.org/10.1038/ng.446>.
- [5] U. Styrkarsdóttir, B.V. Halldórsson, S. Gretarsdóttir, D.F. Gudbjartsson, G.B. Walters, T. Ingvarsson, et al., New sequence variants associated with bone mineral density, *Nat. Genet.* 41 (2009) 15–17, <https://doi.org/10.1038/ng.284>.
- [6] E.L. Duncan, P. Danoy, J.P. Kemp, P.J. Leo, E. McCloskey, G.C. Nicholson, et al., Genome-wide association study using extreme truncate selection identifies novel genes affecting bone mineral density and fracture risk, *PLoS Genet.* 7 (2011) e1001372, <https://doi.org/10.1371/journal.pgen.1001372>.
- [7] K. Estrada, U. Styrkarsdóttir, E. Evangelou, Y.H. Hsu, L.L. Duncan, E.E. Ntzani, et al., Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture, *Nat. Genet.* 44 (2012) 491–501, <https://doi.org/10.1038/ng.2249>.
- [8] L. Zhang, H.J. Choi, K. Estrada, P.J. Leo, J. Li, Y.-F. Pei, et al., Multistage genome-wide association meta-analyses identified two new loci for bone mineral density, *Hum. Mol. Genet.* 23 (2014) 1923–1933, <https://doi.org/10.1093/hmg/ddt575>.
- [9] H.F. Zheng, V. Forgetta, Y.H. Hsu, K. Estrada, A. Rosello-Diez, P.J. Leo, et al., Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture, *Nature* 526 (2015) 112–117, <https://doi.org/10.1038/nature14878>.
- [10] J.P. Kemp, J.A. Morris, C. Medina-Gomez, V. Forgetta, N.M. Warrington, S.E. Youlten, et al., Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis, *Nat. Genet.* 49 (2017) 1468–1475, <https://doi.org/10.1038/ng.3949>.
- [11] C. Medina-Gomez, J.P. Kemp, K. Trajanoska, J. Luan, A. Chesi, T.S. Ahluwalia, et al., Life-course genome-wide association study meta-analysis of total body BMD and assessment of age-specific effects, *Am. J. Hum. Genet.* 102 (2018) 88–102, <https://doi.org/10.1016/j.ajhg.2017.12.005>.
- [12] J.A. Morris, J.P. Kemp, S.E. Youlten, L. Laurent, J.G. Logan, R. Chai, et al., An atlas of genetic influences on osteoporosis in humans and mice, *Nat. Genet.* (2018), <https://doi.org/10.1038/s41588-018-0302-x>.
- [13] J.B. Richards, H.F. Zheng, T.D. Spector, Genetics of osteoporosis from genome-wide association studies: advances and challenges, *Nat. Rev. Genet.* 13 (2012) 576–588, <https://doi.org/10.1038/nrg3228>.
- [14] Y. Guo, J.-T. Wang, H. Liu, M. Li, T.-L. Yang, X.-W. Zhang, et al., Are bone mineral density loci associated with hip osteoporotic fractures? A validation study on previously reported genome-wide association loci in a Chinese population, *Genet. Mol. Res.* 11 (2012) 202–210, <https://doi.org/10.4238/2012.January.31.1>.
- [15] U. Styrkarsdóttir, B.V. Halldórsson, D.F. Gudbjartsson, N.L.S. Tang, J.M. Koh, S.M. Xiao, et al., European bone mineral density loci are also associated with BMD in East-Asian populations, *PLoS One* 5 (2010) e13217, <https://doi.org/10.1371/journal.pone.0013217>.
- [16] K. Trajanoska, J.A. Morris, L. Oei, H.-F. Zheng, D.M. Evans, D.P. Kiel, et al., Assessment of the genetic and clinical determinants of fracture risk: genome wide association and mendelian randomisation study, *Br. J. Med.* 362 (2018) k3225, <https://doi.org/10.1136/bmj.k3225>.
- [17] M. Bustamante, X. Nogués, L. Mellibovsky, L. Águeda, S. Jurado, E. Cáceres, et al., Polymorphisms in the interleukin-6 receptor gene are associated with bone mineral density and body mass index in Spanish postmenopausal women, *Eur. J. Endocrinol.* 157 (2007) 677–684, <https://doi.org/10.1530/EJE-07-0389>.
- [18] M. Bustamante, X. Nogués, L. Águeda, S. Jurado, A. Wesselius, E. Cáceres, et al., Promoter 2-1025 T/C polymorphism in the RUNX2 gene is associated with femoral neck BMD in Spanish postmenopausal women, *Calcif. Tissue Int.* 81 (2007) 327–332, <https://doi.org/10.1007/s00223-007-9069-2>.
- [19] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, et al., The Sequence Alignment/Map format and SAMtools, *Bioinformatics* 25 (2009) 2078–2079, <https://doi.org/10.1093/bioinformatics/btp352>.
- [20] M.A. DePristo, E. Banks, R. Poplin, K.V. Garimella, J.R. Maguire, C. Hart, et al., A framework for variation discovery and genotyping using next-generation DNA sequencing data, *Nat. Genet.* 43 (2011) 491–498, <https://doi.org/10.1038/ng.806>.
- [21] P. Kumar, S. Henikoff, P.C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm, *Nat. Protoc.* 4 (2009) 1073–1081, <https://doi.org/10.1038/nprot.2009.86>.
- [22] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, et al., A method and server for predicting damaging missense mutations, *Nat. Methods* 7 (2010) 248–249, <https://doi.org/10.1038/nmeth0410-248>.
- [23] J.M. Schwarz, D.N. Cooper, M. Schuelke, D. Seelow, MutationTaster2: mutation prediction for the deep-sequencing age, *Nat. Methods* 11 (2014) 361–362, <https://doi.org/10.1038/nmeth.2890>.
- [24] ENCODE Project Consortium, The ENCODE (ENCyclopedia Of DNA Elements) Project, *Science* 306 (2004) 636–640, <https://doi.org/10.1126/science.1105136>.
- [25] D. Bujold, D.A.de L. Morais, C. Gauthier, C. Côté, M. Caron, T. Kwan, et al., The International Human Epigenome Consortium Data Portal, *Cell Syst.* 3 (2016), <https://doi.org/10.1016/j.cels.2016.10.019> 496–499.e2.
- [26] A.R.R. Forrest, H. Kawaji, M. Rehli, J.K. Baillie, M.J.L. De Hoon, V. Haberler, et al., A promoter-level mammalian expression atlas, *Nature* 507 (2014) 462–470, <https://doi.org/10.1038/nature13182>.
- [27] L.D. Ward, M. Kellis, HaploReg: a resource for exploring chromatin states,

- conservation, and regulatory motif alterations within sets of genetically linked variants, *Nucleic Acids Res.* 40 (2012) D930–D934, <https://doi.org/10.1093/nar/gkr917>.
- [28] A.P. Boyle, E.L. Hong, M. Hariharan, Y. Cheng, M.A. Schaub, M. Kasowski, et al., Annotation of functional variation in personal genomes using RegulomeDB, *Genome Res.* 22 (2012) 1790–1797, <https://doi.org/10.1101/gr.137323.112>.
- [29] C.H. Chou, S. Shrestha, C.D. Yang, N.W. Chang, Y.L. Lin, K.W. Liao, et al., MiRTarBase update 2018: A resource for experimentally validated microRNA-target interactions, *Nucleic Acids Res.* 46 (2018) D296–D302, <https://doi.org/10.1093/nar/gkx1067>.
- [30] A.E. Bruno, L. Li, J.L. Kalabus, Y. Pan, A. Yu, Z. Hu, miRdSNP: a database of disease-associated SNPs and microRNA target sites on 3'UTRs of human genes, *BMC Genomics* 13 (2012) 44, <https://doi.org/10.1186/1471-2164-13-44>.
- [31] C. Liu, F. Zhang, T. Li, M. Lu, L. Wang, W. Yue, et al., MirSNP, a database of polymorphisms altering miRNA target sites, identifies miRNA-related SNPs in GWAS SNPs and eQTLs, *BMC Genomics* 13 (2012) 661, <https://doi.org/10.1186/1471-2164-13-661>.
- [32] K. Cartharius, K. Frech, K. Grote, B. Klocke, M. Haltmeier, A. Klingenhoff, et al., MatInspector and beyond: promoter analysis based on transcription factor binding sites, *Bioinformatics* 21 (2005) 2933–2942, <https://doi.org/10.1093/bioinformatics/bt1473>.
- [33] S.G. Coetzee, G.A. Coetzee, D.J. Hazelett, motifbreakR: an R/Bioconductor package for predicting variant effects at transcription factor binding sites, *Bioinformatics* 31 (2015) 3847–3849, <https://doi.org/10.1093/bioinformatics/btv470>.
- [34] D. Hnisz, B.J. Abraham, T.I. Lee, A. Lau, V. Saint-André, A.A. Sigova, et al., Super-enhancers in the control of cell identity and disease, *Cell* 155 (2013) 934–947, <https://doi.org/10.1016/j.cell.2013.09.053>.
- [35] Y. Jiang, F. Qian, X. Bai, Y. Liu, Q. Wang, B. Ai, et al., SEdb: a comprehensive human super-enhancer database, *Nucleic Acids Res.* 47 (2019) D235–D243, <https://doi.org/10.1093/nar/gky1025>.
- [36] J.C. Barrett, B. Fry, J. Maller, M.J. Daly, Haploview: analysis and visualization of LD and haplotype maps, *Bioinformatics* 21 (2005) 263–265, <https://doi.org/10.1093/bioinformatics/bth457>.
- [37] A. Fernández-Miñán, J. Bessa, J.J. Tena, J.L. Gómez-Skarmeta, Assay for transposase-accessible chromatin and circularized chromosome conformation capture, two methods to explore the regulatory landscapes of genes in zebrafish, in: H.W. Detrich, III, M. Westerfield, L.I. Zon (Eds.), *Methods Cell Biol.*, Elsevier Inc, 2016, pp. 413–430, <https://doi.org/10.1016/bs.mcb.2016.02.008>.
- [38] H.J.G. van de Werken, P.J.P. de Vree, E. Splinter, S.J.B. Holwerda, P. Klous, E. de Wit, et al., 4C technology: Protocols and data analysis, in: C. Wu, C.D. Allis (Eds.), *Methods Enzymol.*, Elsevier Inc, 2012, pp. 89–112, <https://doi.org/10.1016/B978-0-12-391938-0.00004-5>.
- [39] D. Noordermeer, M. Leleu, E. Splinter, J. Rougemont, W. De Laat, D. Duboule, The dynamic architecture of HOX gene clusters, *Science* 334 (2011) 222–225, <https://doi.org/10.1126/science.1207194>.
- [40] J.R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, et al., Topological domains in mammalian genomes identified by analysis of chromatin interactions, *Nature* 485 (2012) 376–380, <https://doi.org/10.1038/nature11082>.
- [41] Y. Wang, F. Song, B. Zhang, L. Zhang, J. Xu, D. Kuang, et al., The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions, *Genome Biol.* 19 (2018) 151, <https://doi.org/10.1186/s13059-018-1519-9>.
- [42] A. Visel, S. Minovitsky, I. Dubchak, L.A. Pennacchio, VISTA Enhancer Browser - a database of tissue-specific human enhancers, *Nucleic Acids Res.* 35 (2007) D88–D92, <https://doi.org/10.1093/nar/gkl822>.
- [43] R.Y. Birnbaum, D.B. Everman, K.K. Murphy, F. Gurrieri, C.E. Schwartz, N. Ahituv, Functional characterization of tissue-specific enhancers in the DLX5/6 locus, *Hum. Mol. Genet.* 21 (2012) 4930–4938, <https://doi.org/10.1093/hmg/dd336>.
- [44] F. Zhang, J.R. Lupski, Non-coding genetic variants in human disease, *Hum. Mol. Genet.* 24 (2015) R102–R110, <https://doi.org/10.1093/hmg/ddv259>.
- [45] S. Yao, Y. Guo, S.S. Dong, R.H. Hao, X.F. Chen, Y.X. Chen, et al., Regulatory element-based prediction identifies new susceptibility regulatory variants for osteoporosis, *Hum. Genet.* 136 (2017) 963–974, <https://doi.org/10.1007/s00439-017-1825-4>.
- [46] Y. Guo, S.S. Dong, X.F. Chen, Y.A. Jing, M. Yang, H. Yan, et al., Integrating epigenomic elements and GWASs identifies BDNF gene affecting bone mineral density and osteoporotic fracture risk, *Sci. Rep.* 6 (2016) 30558, <https://doi.org/10.1038/srep30558>.
- [47] O. Corradin, A. Saiakhova, B. Akhtar-Zaidi, L. Myeroff, J. Willis, R. Cowper-Sal lari, et al., Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits, *Genome Res.* 24 (2014) 1–13, <https://doi.org/10.1101/gr.164079.113>.
- [48] Y.H. Hsu, G. Li, C.T. Liu, J.A. Brody, D. Karasik, W.C. Chou, et al., Targeted sequencing of genome wide significant loci associated with bone mineral density (BMD) reveals significant novel and rare variants: the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) targeted sequencing study, *Hum. Mol. Genet.* 25 (2016) 5234–5243, <https://doi.org/10.1093/hmg/ddw289>.
- [49] G. Kang, D. Lin, H. Hakonarson, J. Chen, Two-stage extreme phenotype sequencing design for discovering and testing common and rare genetic variants: efficiency and power, *Hum. Hered.* 73 (2012) 139–147, <https://doi.org/10.1159/000337300>.
- [50] I.J. Barnett, S. Lee, X. Lin, Detecting rare variant effects using extreme phenotype sampling in sequencing association studies, *Genet. Epidemiol.* 37 (2013) 142–151, <https://doi.org/10.1002/gepi.21699>.
- [51] D.S. Paul, N. Soranzo, S. Beck, Functional interpretation of non-coding sequence variation: concepts and challenges, *BioEssays* 36 (2014) 191–199, <https://doi.org/10.1002/bies.201300126>.
- [52] P.B. Rahl, R.A. Young, MYC and transcription elongation, *Cold Spring Harb. Perspect. Med.* 4 (2014) a020990, <https://doi.org/10.1101/cshperspect.a020990>.
- [53] R. Raisner, S. Kharbanda, L. Jin, E. Jeng, E. Chan, M. Merchant, et al., Enhancer activity requires CBP/P300 bromodomain-dependent histone H3K27 acetylation, *Cell Rep.* 24 (2018) 1722–1729, <https://doi.org/10.1016/j.celrep.2018.07.041>.
- [54] R. Andersson, C. Gebhard, I. Miguel-Escalada, I. Hoof, J. Bornholdt, M. Boyd, et al., An atlas of active enhancers across human cell types and tissues, *Nature* 507 (2014) 455–461, <https://doi.org/10.1038/nature12787>.
- [55] E. Arner, C.O. Daub, K. Vitting-Seerup, R. Andersson, B. Lilje, F. Drablos, et al., Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells, *Science* 347 (2015) 1010–1014, <https://doi.org/10.1126/science.1259418>.
- [56] C.C. Hon, J.A. Ramilowski, J. Harshbarger, N. Bertin, O.J.L. Rackham, J. Gough, et al., An atlas of human long non-coding RNAs with accurate 5' ends, *Nature* 543 (2017) 199–204, <https://doi.org/10.1038/nature21374>.
- [57] L. Qin, Y. Liu, Y. Wang, G. Wu, J. Chen, W. Ye, et al., Computational characterization of osteoporosis associated SNPs and genes identified by genome-wide association studies, *PLoS One* 11 (2016) e0150070, <https://doi.org/10.1371/journal.pone.0150070>.
- [58] S.C. Manolagas, C.A. O'Brien, M. Almeida, The role of estrogen and androgen receptors in bone health and disease, *Nat. Rev. Endocrinol.* 9 (2013) 699–712, <https://doi.org/10.1038/nrendo.2013.179>.
- [59] Q. Zeng, K.H. Wu, K. Liu, Y. Hu, X.D. Chen, L. Zhang, et al., Genome-wide association study of lncRNA polymorphisms with bone mineral density, *Ann. Hum. Genet.* 82 (2018) 244–253, <https://doi.org/10.1111/ahg.12247>.
- [60] M. Villalobos-Comparán, R.F. Jiménez-Ortega, K. Estrada, A.Y. Parra-Torres, A. González-Mercado, N. Patiño, et al., A pilot genome-wide association study in postmenopausal Mexican-Mestizo women implicates the RMND1/CCDC170 locus as associated with bone mineral density, *Int. J. Genomics* 2017 (2017) 5831020, <https://doi.org/10.1155/2017/5831020>.

**Supplementary Information:****Supplementary Table 1.** Coordinates and primers for the amplification of the *C7ORF76* genomic region by LR-PCR

Fragment	Genomic coordinates (GRCh37)	Primer Fwd	Primer Rev	Product size (bp)
1	7:96,136,619-96,132,152	TTGACCTGAATACTGCCGC	GCCAAATGAATGTGGACAAG	4468
2	7:96,132,302-96,127,711	CACTGCTGGGTCTTAGATTGG	GCATGTGTGCATGATGTTGG	4592
3	7:96,127,863-96,123,028	TGCAAGTTTCCCTCAATTCATC	TCCCTCTCATCTGTGCAACAC	4836
4	7:96,123,158-96,118,311	TTAGGTGAGTAGAAAGCAATGGC	CTGGGTGGCTATAGACCTGAATAG	4848
5	7:96,118,477-96,114,227	GCGGCACTGTGAGAGTACATC	CCTGGTGAAATGGGAACA	4251
6	7:96,114,348-96,110,480	CTGACACTTTGGCAGCACC	GGGATTGTTGAAGCTGACCC	3869
7	7:96,110,702-96,108,695	CAACCATCACAACCCATAGAC	CCTGAGCAAGTCTCGTAAGTG	2008

**Supplementary Table 2.** Primers used for genotyping the selected SNPs

SNP	Primer Fwd	Primer Rev
rs4727338	CACATACACTTGACTGTGTTTGGT	GGATTCTGGCTTTGACATCC
rs10429035	TCTTTTGTGTTTGGAGGAAAGG	TCCTGTACGGAACCCTGACT
rs12674052	GGAAACCCTGTGTTATTTCAAGC	GGTTGCCCAAGTCACCAC
rs4342421	TAAATGTGACCTTTGTACTCAACA	AAATGTCAGAGGATGGTCCAG
rs10085588	TGTTCCAGATGCAAGATGATT	AGTGGAGATTCAGGGGGAAT

**Supplementary Table 3.** Primers used for qPCR assays

Gene	Primer Fwd	Primer Rev
<i>DLX5</i>	CTACAACCGCGTCCCAAG	GCCATTCACCATTCTCACCT
<i>DLX6</i>	ATATATTAGAGAAGAGCGAGGGAGAG	CCCTCTGCAGCCACCTTA
<i>AS-DLX6</i>	TGATTCCTGTATGTATGGCAGCTA	GGTTTTCTTTGTCTCAGCAAT
<i>SLC25A13</i>	AGATGGTTCGGTCCCCTT	GCAAACGGATCTTGACGATT
<i>SHFM1</i>	GACGACGAGTTTGAAGAGTTCC	CCCAATTATCCTCCAGACA
<i>HMBS</i>	TGCCCTGGAGAAGAATGAAG	CAGCATCATGAGGGTTTTCC

**Supplementary Table 4.** Minor allele frequency (MAF) from 1000 Genomes project: total population (ALL), European population (EUR) and Iberian population in Spain (IBS) and MAF of complete BARCOS cohort.

SNP	Minor allele	ALL	EUR	IBS	BARCOS
rs115076023	A	0.011	0.000	0.000	0.000
rs10085588	A	0.223	0.359	0.360	0.392
rs4342521	T	0.221	0.360	0.360	0.396
rs7794042	A	0.009	0.014	0.005	0.031
rs4613908	A	0.386	0.363	0.332	0.321
rs10238953	C	0.116	0.156	0.220	0.167
rs190892252	G	0.004	0.008	0.009	0.012
rs4727338	C	0.221	0.360	0.360	0.390
rs117923361	A	0.016	0.040	0.014	0.035



## Article 2

**Title:** A *DLX5/6* enhancer in the *C7ORF76* locus: Characterization of its role in development and in bone

### Summary:

Enhancers play important roles in precise spatiotemporal regulation of gene expression, essential for defining cell identity during development. *DLX5/6* are two transcriptional regulators involved in the development of branchial arches, inner ear and skeleton, among others, that have been related to split hand and foot malformation 1. Several tissue-specific enhancers thought to regulate *DLX5/6* have been described. The aim of this work was to functionally characterise one such enhancer (eDlx#18), present within *C7ORF76* (a locus repeatedly associated with bone mineral density and osteoporotic fracture in genome-wide associated studies), both in embryonic development and in a bone context. eDlx#18 displayed transactivation capacity in an osteoblastic cell line when tested using a reporter gene assay, and a SNP within eDlx#18 (rs10238953) was nominally associated with transcript levels of *DLX6* in human primary osteoblasts. In addition, 4C-seq in osteoblastic cell lines demonstrated interactions between eDlx#18 and the *DLX5* promoter, as well as with different *DLX5/6* tissue-specific enhancers described in the nearby region. Finally, a homozygous deletion of eDlx#18 caused a reduced survival in mouse embryos and several defects including decreased *Dlx5* expression in otic vesicle and branchial arches in E11.5 embryos, and a slightly smaller dentary, a deficient ossification of supraoccipital bone, vertebral bodies, sternum and pelvic bones, and minor affectations in the ribs in E17.5 embryos, while no limb malformations were observed. These phenotypes partly recapitulate the *Dlx5*<sup>-/-</sup> phenotype.

In summary, this is the first study to analyse in depth the functionality of the eDlx#18 enhancer. We provide functional evidence *in vivo* that this enhancer may regulate *DLX5/6* in different body locations during development and it may have an effect on osteogenesis.

### Reference:

Neus Roca-Ayats, Núria Martínez-Gil, Mónica Cozar, Natàlia Garcia-Giralt, Xavier Nogués, Adolfo Díez-Pérez, Aleix Gavaldà-Navarro, Jordi Garcia-Fernandez, Daniel Grinberg, Darío G. Lupiáñez, Susanna Balcells. A *DLX5/6* enhancer in the *C7ORF76* locus: Characterization of its role in development and in bone. To be submitted



## **A *DLX5/6* enhancer in the *C7ORF76* locus: Characterization of its role in development and in bone**

Neus Roca-Ayats<sup>1</sup>, Núria Martínez-Gil<sup>1</sup>, Mónica Cozar<sup>1</sup>, Natàlia Garcia-Giralt<sup>2</sup>, Xavier Nogués<sup>2</sup>, Adolfo Díez-Pérez<sup>2</sup>, Aleix Gavaldà-Navarro<sup>3</sup>, Jordi Garcia-Fernandez<sup>4</sup>, Daniel Grinberg<sup>1,\*</sup>, Dario Lupiañez<sup>5,\*</sup>, Susanna Balcells<sup>1,\*</sup>

<sup>1</sup>Department of Genetics, Microbiology and Statistics, Faculty of Biology, Universitat de Barcelona, Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), ISCIII, IBUB, IRSJD, Barcelona, Catalonia, Spain.

<sup>2</sup>Musculoskeletal Research Group, IMIM (Hospital del Mar Medical Research Institute), Centro de Investigación Biomédica en Red en Fragilidad y Envejecimiento Saludable (CIBERFES), ISCIII, Barcelona, Catalonia, Spain.

<sup>3</sup>Department of Biochemistry and Molecular Biomedicine, Faculty of Biology, Universitat de Barcelona, Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad i Nutrición (CIBEROBN), ISCIII, IBUB, IRSJD, Barcelona, Catalonia, Spain

<sup>4</sup>Department of Genetics, Microbiology and Statistics, Faculty of Biology, Universitat de Barcelona, IBUB, Barcelona, Catalonia, Spain

<sup>5</sup>Epigenetics and Sex Development Group, Berlin Institute for Medical Systems Biology, Max-Delbrück Center for Molecular Medicine, Berlin, Germany

\*Last co-authors

Corresponding author:

Susanna Balcells: [sbalcells@ub.edu](mailto:sbalcells@ub.edu)

Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona. Av. Diagonal, 643. 08028 Barcelona (Catalonia), Spain

### **Abstract**

Enhancers play important roles in precise spatiotemporal regulation of gene expression, essential for defining cell identity during development. *DLX5/6* are two transcriptional regulators involved in the development of branchial arches, inner ear and skeleton, among others, that have been related to split hand and foot malformation 1. Several tissue-specific enhancers thought to regulate *DLX5/6* have been described. The aim of

this work was to functionally characterise one such enhancer (eDlx#18), present within *C7ORF76* (a *locus* repeatedly associated with bone mineral density and osteoporotic fracture in genome-wide associated studies), both in embryonic development and in a bone context. eDlx#18 displayed transactivation capacity in an osteoblastic cell line when tested using a reporter gene assay, and a SNP within eDlx#18 (rs10238953) was nominally associated with transcript levels of *DLX6* in human primary osteoblasts. In addition, 4C-seq in osteoblastic cell lines demonstrated interactions between eDlx#18 and the *DLX5* promoter, as well as with different *DLX5/6* tissue-specific enhancers described in the nearby region. Finally, a homozygous deletion of eDlx#18 caused a reduced survival in mouse embryos and several defects including decreased *Dlx5* expression in otic vesicle and branchial arches in E11.5 embryos, and a slightly smaller dentary, a deficient ossification of supraoccipital bone, vertebral bodies, sternum and pelvic bones, and minor affectations in the ribs in E17.5 embryos, while no limb malformations were observed. These phenotypes partly recapitulate the *Dlx5*<sup>-/-</sup> phenotype.

In summary, this is the first study to analyse in depth the functionality of the eDlx#18 enhancer. We provide functional evidence *in vivo* that this enhancer may regulate *DLX5/6* in different body locations during development and it may have an effect on osteogenesis.

## Introduction

Projects aiming at deciphering the functional genome, such as ENCODE<sup>1</sup> and Roadmap Epigenomics<sup>2</sup>, have revealed a high abundance of *cis*-regulatory elements (i.e. enhancers) in the genome that play a central role in determining precise spatiotemporal gene expression patterns, essential for defining cell identity during development<sup>3-6</sup>. In addition, several human diseases caused by enhancer disruptions and mutations, as well as by disturbing enhancer-promoter interactions due to chromosomal rearrangements, have been identified<sup>7-9</sup>.

The distal-less homeobox genes *DLX5* and *DLX6* encode two paralogous transcriptional regulators important for skeletal, branchial arches, forebrain, olfactory placode and inner ear development<sup>10-15</sup>. Several evolutionary conserved *cis*-acting enhancers thought to regulate the *DLX5/6 locus* in a tissue-specific manner have been described in the 7q21.3 genomic region (Figure 1A)<sup>16</sup>. Disruptions of these enhancers are thought to cause isolated split-hand/-foot malformation 1 (SHFM1; OMIM #183600) or syndromic SHFM1 (OMIM #220600), in combination with the characteristic ectrodactyly, hearing loss (HL),

craniofacial anomalies (CF), and/or intellectual disability (ID)<sup>17–19</sup>. Rasmussen *et al.*<sup>17</sup>, proposed three phenotypic 7q21.3 subregions, and correlated them with the corresponding tissue-specific *DLX5/6* enhancers (Figure 1B). In addition, *DLX5* and *DLX6* intragenic mutations have been found in a few SHFM1 patients, of which only one presented hearing loss<sup>20–22</sup>.

Otherwise, the 7q21.3 genomic region has been repeatedly found associated with bone mineral density (BMD) and osteoporotic fracture in genome-wide association studies (GWASs) and meta-analyses<sup>23–25</sup>. In particular, SNPs within the *C7ORF76* locus have been found among the genome-wide top-associated signals. This locus has been studied in detail in a previous work by us<sup>26</sup>, highlighting the importance of non-coding regulatory regions. One of the previously described tissue-specific *DLX5/6* enhancers (eDlx#18) lies within intron 2 of *C7ORF76* (Figure 1C). eDlx#18 was found to be active in branchial arches of zebrafish embryos (72h post-fertilization, hfp) and mouse embryos (embryonic day E11.5) in transgenic enhancer assays<sup>16</sup> and, interestingly, it is marked as an enhancer active in osteoblasts by ENCODE<sup>1</sup>.

In this study, we aimed at characterising the eDlx#18 enhancer present in the *C7ORF76* locus, both in embryonic development and in a bone context. Through a combination of different experimental approaches, we studied its transcription activation capacity and long-range interactions, and we showed the effects of its deletion in mouse embryos.

## Materials and methods

### *In silico* functional annotation

Epigenetic regulatory features, such as DNase I hypersensitivity, histone modifications, conservation and miRNAs binding, were annotated for the region of interest using data from ENCODE<sup>1</sup>, International Human Epigenome Consortium<sup>27</sup>, The Roadmap Epigenomics Project, FANTOM5<sup>28</sup>, miRTarBase (<http://mirtarbase.mbc.nctu.edu.tw>), miRdSNP (<http://mirdsnp.ccr.buffalo.edu>), MirSNP (<http://bioinfo.bjmu.edu.cn/mirsnp/search>), RegulomeDB (<http://www.regulomedb.org>), HaploReg (<https://pubs.broadinstitute.org/mammals/haploreg>), BioMart, and Ensembl and UCSC Genome Browser.

### Cell culture

The human osteosarcoma cell line Saos-2 was used for luciferase reporter assays and 4C-seq assays. It was obtained from the American Type Culture Collection (ATCC® HTB-

85<sup>TM</sup>) and grown in Dulbecco's Modified Eagle Medium (DMEM; Sigma-Aldrich), supplemented with 10% Fetal Bovine Serum (FBS; Gibco, Life Technologies) and 1% penicillin/streptomycin (Gibco, Life Technologies), at 37°C and 5% of CO<sub>2</sub>. Human fetal osteoblasts (hFOB) 1.19 cells were used for 4C-seq assays. They were obtained from ATCC (ATCC<sup>®</sup> CRL-11372<sup>TM</sup>) and grown in DMEM:F12 (1:1) medium without phenol red (Gibco, Life Technologies), supplemented with 10% FBS and 0.3 mg/ml Geneticin (Gibco, Life Technologies), at 34°C and 5% of CO<sub>2</sub>. Human medulla-derived mesenchymal stem cells (MSCs) were used for 4C-seq assays. They were kindly provided by Dr. José Manuel Quesada Gómez, from Instituto Maimónides de Investigación Biomédica, Hospital Universitario Reina Sofía, Córdoba, Spain. They were grown in alpha-MEM medium (Gibco, Life Technologies), supplemented with 10% FBS, 1% penicillin/streptomycin and 1x Glutamax (Gibco, Life Technologies), at 37°C and 5% of CO<sub>2</sub>. Human primary osteoblasts (hOB) were used for eQTL assays. They were obtained from trabecular bone of patients who underwent knee replacement due to osteoarthritis. Bony tissue was cut up into small pieces, washed in phosphate buffered saline (PBS; Gibco, Life technologies) to remove non-adherent cells, and placed on a 140 mm culture plate. Samples were cultured in DMEM supplemented with 10% FBS, 100 U/ml penicillin/streptomycin, 0.4% fungizone (Gibco, Life Technologies) and 100 µg/ml ascorbic acid (Sigma-Aldrich). DNA and RNA extractions were performed at maximum passage 2. HeLa and HEK293 cell lines were obtained from ATCC (ATCC<sup>®</sup> CCL-2<sup>TM</sup> and ATCC<sup>®</sup> CRL-1573<sup>TM</sup>, respectively) and grown in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin at 37°C and 5% CO<sub>2</sub>.

#### *In vitro* luciferase assay

Different constructs of the enhancer eDlx#18 (box in Figure 2A and Supplementary Table 1) were cloned upstream of the SV40 promoter of the pGL3-Promoter vector (Promega). DNA was PCR-amplified from human genomic DNA and cloned using XhoI and KpnI restriction enzymes.

Saos-2 cells were seeded at a density of  $3.0 \times 10^5$  cells per well in a 6-well plate. After 24h, they were transfected with 2.2 µg of total DNA per well using FuGENE HD reagent, according to manufacturer instructions (Promega). Two plasmids were cotransfected in each well: the pGL3-Promoter empty or with the corresponding construct cloned upstream of the SV40 promoter and the pRL-TK plasmid, containing the Renilla Luciferase gene, in a proportion of 1/10. Forty-eight hours after transfection, cells were rinsed with PBS and lysed. The luciferase activity was measured using a Glomax Multi+ luminometer (Promega), with the Dual-Luciferase<sup>®</sup> Reporter Assay System reagents

(Promega). Each experiment was performed in two biological replicates and was repeated 5 times.

#### Human primary osteoblasts DNA extraction and sequencing

DNA was extracted from cultured hOBs using the Wizard® Genomic DNA Purification Kit (Promega), according to manufacturer's instructions. The concentration of the purified DNA was analysed in a spectrophotometer (Nanodrop). Genotypes for rs4613908 and rs10238953 were assessed by Sanger sequencing using the BigDye® Terminator v3.1 (Applied Biosystems) in the genomic services of Universitat de Barcelona. Primers (Invitrogen, Thermo Fisher) were designed using the Primer3 Input (v. 0.4.0) and the UCSC Genome Browser: for rs4613908: Fwd 5'-CCTACACACATACACCACCT-3' and Rev 5'-GTACAATGAAATGACAGCAAAC-3'; and for rs10238953: Fwd 5'-CTGTCTGTCAACCAAGCCAG-3' and Rev 5'-TGAAGGTCTTGTGGAGAGGC-3'.

#### Human primary osteoblasts RNA extraction, retrotranscription and qPCR

RNA was extracted from cultured hOBs using the High Pure RNA Isolation kit (Roche), according to manufacturer's instructions. RNA was quantified using a spectrophotometer (Nanodrop) and retrotranscribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Thermo Fisher), according to the specifications of the manufacturer. qPCR was performed using UPL Probes (Roche) on a LightCycler 480 Instrument II (Roche). The *HMBS* gene was used as an internal control, and fold changes were calculated by relative quantification, using the 2<sup>nd</sup> derivative method. Primers used are summarised in Supplementary Table 2.

#### 4C-seq

4C-seq libraries were generated from microdissected E14.5 mouse developing humeri as described previously<sup>29</sup>. 4-bp cutters were used as primary (Csp6I) and secondary (DpnII) restriction enzymes. For each viewpoint (eDlx#18 and *Dlx5* promoter), a total of 1.6 µg of library was amplified by PCR (primer sequences in Supplementary Table 3). Samples were sequenced with Illumina Hi-Seq technology according to standard protocols. For 4C-seq data analysis, reads were pre-processed and mapped to a corresponding reference (mm9) using BWA-MEM<sup>30</sup> and coverage was normalised as reported previously<sup>8</sup>. The viewpoint and adjacent fragments 1.5 kb up- and downstream were removed, and a window of two fragments was chosen to normalise the data per million mapped reads (RPM). 4C-seq experiments were carried out in two biological replicates. 4C-seq assays from human MSCs, and Saos-2 and hFOB 1.19 cell lines were performed similarly at the Functional Genomics Service of the Centro Andaluz de

Biología del Desarrollo (Sevilla, Spain). Primer sequences are listed in Supplementary Table 3.

#### Generation of enhancer knockout (KO) mice by CRISPR-Cas9

Two sgRNAs were designed flanking the region to delete, using the web designing tool <http://crispr.mit.edu>. To minimise off-target effects, guide sequences were chosen to have a quality score above 85%. Target region and guide sequences are listed in Supplementary Table 4. A pair of complementary oligos for every guide was annealed, phosphorylated and cloned into the BbsI site of pX459 CRISPR/Cas vector (Addgene). G4 ES cells (129/SvxC57BL/6 F1 hybrid) were cultured on CD1 MEF feeder layers under standard ES cell culture conditions and were transfected with 8 µg of each CRISPR construct using FuGENE HD reagent (Promega). After 12 hours, cells were re-plated on DR4 puro-resistant feeders and, after 1 day, they were selected with puromycin (2 µg/ml) for 2 days. Clones were then allowed to recover for 4-6 days, isolated, expanded and genotyped by PCR analysis (Supplementary Table 3) and Sanger sequencing. Mice were generated by tetraploid complementation<sup>31</sup>. Four pseudopregnant mothers per condition (KO and wild-type; WT) were used for obtaining embryos as well as adult mice. Embryos were collected at embryonic days E11.5 and E17.5. and were genotyped by Sanger sequencing. All animal procedures were in accordance with institutional, state and government regulations (LAGeSo Berlin).

#### Whole-mount *in situ* hybridization

ISH for *Dlx5* was carried out on WT and KO E11.5 embryos (n=3 for each group), as previously described<sup>32</sup>. Briefly, the DIG RNA probes were synthesised from a plasmid<sup>33</sup> kindly provided by Dr. Birnbaum (Ben-Gurion University of the Negev, Israel), using the T7 and T3 polymerases, for the antisense and sense probes, respectively. The probes were precipitated with EtOH with LiCl 100 mM. Embryos were collected and fixed overnight in PFA 4% at 4°C. Then, they were permeabilised with Proteinase K (10 µg/ml) in PBST for 30 min at RT and post-fixed for 20 min in 0.1% glutaraldehyde/4% PFA in PBS at RT. Embryos were incubated in hybridization buffer for 6h at 70 °C and in hybridization buffer plus riboprobe (800 ng/ml) overnight at 70°C. Afterwards, embryos were washed and blocked with 10% sheep serum and 1% BSA in TBST for 3h at RT and they were incubated with anti-DIG antibody tagged with alkaline phosphatase (1:3500) in blocking solution overnight at 4°C. The development was carried out by incubating the embryos in the chromogenic solution (0.45 µl/ml NBT and 3.5 µl/ml BCIP in NTMT) at RT protected from the light. Finally, embryos were re-fixed in 4% PFA for 30 min and stored at 4°C in 75% glycerol.

### Skeletal preparations

Mouse embryonic skeletons at E17.5 (WT: n=6; mutant: n=4) were stained with Alcian blue and Alizarin red, according to standard protocols<sup>34</sup>. Briefly, embryos were fixed for 12h with EtOH at RT, after removing the viscera and skin. Then, they were incubated in Alcian blue (150 mg/l in 80% EtOH, 20% acetic acid) for 20h at RT and post-fixed overnight with EtOH. Afterwards, embryos were incubated in 2% KOH for 6h at RT and in Alizarin red S (50 mg/l in 2% KOH) for 3h at RT. Finally, they were incubated in 2% KO at RT until the soft tissues were digested and the skeletal elements visible. Embryos were preserved in 25% glycerol. For comparison of limb skeletons from enhancer KO and WT embryos, general parameters such as bone number, shape, length, position or mineralization were assessed. Measurements of the ossified portions of humerus and femur (stylopodial elements) were normalised by those of the corresponding ulna and tibia (related zeugopodial elements), respectively.

### Statistical methods

Relative luciferase units (RLU, i.e. the ratio of the firefly luciferase activity over the Renilla luciferase activity) were calculated for each individual measurement and a one-way blocked ANOVA was performed. A TukeyHSD test was performed as a post-hoc test for multiple comparisons. The analyses were performed using R software version 3.4.1 and p-values<0.05 were considered significant. All the data was ascertained for normality, homoscedasticity and atypical data points.

Hardy-Weinberg equilibrium (HWE) was calculated using the chi-square test. Linear regression was used to assess the association between gene expression levels and genotypes in primary osteoblasts (*cis*-eQTL). Log-additive, dominant and recessive models were tested for each gene analysis. Correlation analyses were performed using R software version 3.4.1 with the *SNPassoc* package. All analyses were two-tailed and p-values<0.05 were considered significant. Correction for multiple testing was performed using the Bonferroni's method for the number of SNPs tested. A two-tailed unpaired Student t-test was used for assessing the differences in bone length between WT and mutant KO embryos.

## **Results**

### Functional annotation of the eDlx#18 enhancer

We explored the functional annotation of eDlx#18 enhancer (hs2311 element from VISTA Enhancer Browser<sup>35</sup>; chr7:96,124,919-96,125,415 GRCh37) using publicly available *in silico* data from different cell types. eDlx#18 is a highly conserved 497 bp-element located in intron 2 of *C7ORF76* at 7q21.3, enriched for typical enhancer marks, including H3K27ac, H3K4me1, and H2A.Z, as well as DNase hypersensitivity, in several cell types, among which osteoblasts, medulla-derived mesenchymal stem cells, human embryonic stem cells, hepatocytes, brain germinal matrix, and macrophages (Figure1C).

#### Evaluation of transactivation capability of the eDlx#18 enhancer *in vitro*

eDlx#18 was shown to have gene enhancer activity assessed by reporter gene expression in transgenic zebrafish and mice at embryonic stages (72 hpf and E11.5, respectively)<sup>16</sup>. Specifically, Birnbaum *et al.* showed that eDlx#18 drove gene expression in the branchial arches (in mice, mandibular branchial arch). To assess whether eDlx#18 was able to activate transcription in an osteoblastic context, we performed luciferase reporter assays in Saos-2 cells. We tested three different constructs of the enhancer (box in Figure 2A, Supplementary Table 1) cloned upstream of the SV40 promoter, in the pGL3-Promoter vector: the entire eDlx#18 enhancer, a larger fragment including 270 additional bp at each side (eDlx#18ext) and a smaller central fragment (292 bp), corresponding to the enhancer marks in some cell types (eDlx#18core). As shown in Figure 2A, the eDlx#18 construct showed a decreased luciferase expression compared to the empty vector (FC: 0.45,  $p < 0.001$ ). However, the eDlx#18ext construct showed no significant differences with the empty vector, while the eDlx#18core construct showed a significant increase of luciferase activity (FC: 2.6,  $p = 0.0059$ ), suggesting that the eDlx#18 might contain some repressor elements in its 5' and/or 3' ends and an activator region in its core that are functional, at least in osteoblastic cells.

We also evaluated if the eDlx#18 enhancer was transcribed (as occurs in many active enhancers) by performing RT-PCR in cDNA of human primary osteoblasts, Saos-2, HeLa and HEK293 cells. We were able to amplify the eDlx#18 sequence from cDNA of Saos-2, HeLa and HEK293 but failed to amplify it from different human primary osteoblast cDNA samples (Figure 2B). We also interrogated the FANTOM5 Cap Analysis of Gene Expression (CAGE) dataset<sup>28</sup> for evidence of eDlx#18 transcription and we could observe a signal for transcription start site (TSS) expression in several cell types, such as small cell lung carcinoma, occipital cortex, cervix carcinoma, Saos-2 cells, gastrointestinal carcinoma, mesenchymal stem cells, and Merkel cell carcinoma.

#### *cis*-eQTL analyses of SNPs within the eDlx#18 enhancer

Next, we evaluated whether the two SNPs present in the eDlx#18 enhancer (rs10238953 and rs4613908; Figure 1C) could be *cis*-eQTLs in human primary osteoblasts, a cell type not available in GTEx. Interestingly, rs10238953 showed a nominal association with *DLX6* gene expression ( $p=0.04667$ ), where the minor allele (C) was associated with a decreased gene expression. On the other hand, neither rs10238953 nor rs4613908 were found to be eQTLs for *DLX5*, *DLX6-AS1*, *SLC25A13* or *SHFM1*, although a tendency was observed for *SHFM1* and *SLC25A13* in the case of rs4613908 and for *SHFM1* in the case of rs10238953 (Table 1). Notably, we failed to detect *C7ORF76* expression in this cell type.

#### Long-range chromatin interactions between the eDlx#18 enhancer and *DLX5* promoter

We then wanted to analyse the 3D chromatin interactions of eDlx#18 by 4C-seq in different human cell types (MSCs, hFOB 1.19 and Saos-2) and in mouse developing humeri (embryonic day E14.5). We obtained similar interaction profiles in the 4 samples, consistent with the conserved nature of this enhancer. We detected interaction between eDlx#18 and the region spanning 750 kb on each side of it (Figure 3), and no other interactions were detected elsewhere in the genome. These results showed a clear topologically associated domain (TAD) within which the interactions among different regions would take place. Consistently, publicly available Hi-C data from many different human cell types (obtained from the 3D Genome Browser<sup>36,37</sup>) showed a TAD in this genomic region spanning from approximately 50 kb upstream of *DYNC111* TSS to approximately 25 kb upstream of *ACN9* TSS (Figure 3A). As expected, since eDlx#18 has been described as a *DLX5/6* regulatory element<sup>16</sup>, we found interaction between the enhancer and the *DLX5/6* region, although it was not evident in hFOB 1.19 cells. Notably, we detected interaction with many of the active enhancers described in the VISTA Browser<sup>16,35</sup>. We also found interaction between eDlx#18 and *SLC25A13* and between eDlx#18 and *SHFM1*, although not to the same extent in the different samples. Besides, an interaction between eDlx#18 and a regulatory region upstream of the *C7ORF76* TSS that we described recently<sup>26</sup> was also observed in the three human cell lines.

Then, we investigated the chromatin interactions from the *DLX5* promoter in the same samples: human Saos-2, hFOB 1.19 and MSCs, and mouse developing humeri (E14.5). Again, the interaction profiles were similar in the different cell types and we could only detect interactions within the TAD described above. As expected, we found higher interaction levels with the tissue-specific enhancers described in Birnbaum *et al.*<sup>16</sup>, and in the VISTA Enhancer Browser<sup>35</sup>, among which eDlx#18, further confirming that these enhancers regulate *DLX5/6* gene expression. In addition, we detected a strong

interaction with the *SHFM1* gene body, especially in the human cell types, and with the human long non-coding RNA gene *LOC100506136*.

#### Generation and phenotype of a targeted eDlx#18 knockout mutation

In order to elucidate the biological role of the eDlx#18 enhancer *in vivo*, we used CRISPR-Cas9 to homozygously delete a 12-kb fragment containing eDlx#18 in mouse ES cells, and then generated embryos by chimeric tetraploid aggregation. We selected the E11.5 embryonic stage to assess *Dlx5* expression by whole-mount RNA *in situ* hybridization. The E11.5 KO embryos showed normal morphology. *Dlx5* expression in WT embryos, was observed in the branchial arches, limb bud, otic vesicle, olfactory placode, genital tubercle and forebrain, as previously reported<sup>12,16</sup>. In contrast, in the homozygous KO embryos, expression in the otic vesicle was much reduced and expression in branchial arches was slightly reduced (Figure 4).

At E17.5 stage, selected to assess the skeletal morphology, we obtained four KO and six WT embryos all of similar body size. Of note, KO embryos did not have a good aspect, with two of them having a whitish colour and showing exencephaly, one of which presenting a hole at the interparietal zone (Figure 5A). We analysed the skeleton of the mutant embryos and observed that all KO embryos had deficient ossification of the supraoccipital bone (Figure 5B) and the dentary was slightly smaller with a reduced condylar process (Figure 5C). Inner ears of KO embryos seemed to present a normal morphology. No other craniofacial malformation was observed. Moreover, all the KO embryos presented several axial skeletal defects. An incomplete fusion and development of sternum was observed in all of them (Figure 6A), as well as curved irregular ribs and ectopic cartilage fusing ribs, the latter in 2/4 embryos (Figure 6A-B). In addition, ossification centres of vertebral bodies appeared much reduced (Figure 6C). Regarding the appendicular skeleton, no limb malformations were observed and there were not significant differences in long bone ossification. However, pelvic girdle bones were ossified to a lesser extent in the KO embryos compared to WT (Figure 6D).

Finally, no KO offspring was born (while WT offspring did), suggesting that the deletion might decrease embryonic viability.

## **Discussion**

Many developmental genes are regulated by a landscape of *cis*-acting enhancers<sup>5</sup>. Such is the case of *DLX5/6*, in the 7q21.3 genomic region, usually found associated with BMD

and osteoporotic fracture in GWAS<sup>16,23–25</sup>. The specific function of individual enhancers is not well-known. In this line, and since *DLX5* plays a central role in osteoblastogenesis<sup>14,38</sup>, we studied the eDlx#18 enhancer (chr7:96,124,919-96,125,415; GRCh37), located within intron 2 of the *C7ORF76* gene and previously described as a branchial arches enhancer by reporter gene assay in zebrafish and mouse<sup>16,35</sup>. We specially analysed the bone-related function of this enhancer by assessing its transcriptional activation capacity in an osteoblastic cell line, describing its interactions in different bone-related cell types and evaluating the skeletal affectations of a homozygous enhancer deletion. We found that eDlx#18 was able to activate gene transcription in Saos-2 cells and showed that SNP rs10238953 within eDlx#18 was nominally associated with *DLX6* gene expression. We observed that in bone lineage cells eDlx#18 interacted with the *DLX5* promoter and with the different *DLX5/6* tissue-specific enhancers described in the nearby region. Finally, we showed that the deletion of this enhancer caused several ossification defects and a reduced survival in mouse embryos.

*DLX5* and *DLX6* genes are organized in a convergently transcribed bigene cluster, show similar patterns of expression and display a partly redundant function<sup>10,39–41</sup>. They are involved in many developmental processes, such as sensory organ morphogenesis<sup>11,15</sup>, neurogenesis and forebrain development<sup>42,43</sup>, branchial arches<sup>11</sup>, craniofacial and limb development<sup>12,44</sup>, including chondrocyte differentiation<sup>45</sup> and osteoblastogenesis<sup>14,38</sup>. The regulation of developmental genes such as *DLX5/6* is tightly controlled in order to achieve the spatiotemporally precise and robust gene expression essential for establishing cell fate, lineage commitment, complex body plan and organogenesis<sup>46</sup>. This highly regulated gene expression is orchestrated by multiple enhancers that lie in regulatory landscapes that can span over hundreds of kilobases<sup>47,48</sup>. In the 7q21.3 region, there are 11 tissue-specific enhancers along ~1 Mb that presumably regulate *DLX5/6* expression, among which eDlx#18, that were identified by comparative genome analysis as non-coding evolutionary highly conserved regions, with a 70% of identity between human and frog, and tested for *in vivo* enhancer activity by reporter gene assays in zebrafish and mouse<sup>16,35</sup>. Indeed, eDlx#18 is enriched for typical enhancer marks (i.e. open chromatin, H3K4me1, H3K27ac, etc.) in several cell types, including osteoblasts. In addition, according to ENCODE ChIP-seq data, the EP300 histone acetyltransferase, known to function as a coactivator in active enhancers<sup>49</sup>, and RUNX3 bind to eDlx#18. Interestingly, RUNX3 is involved in chondrocyte maturation as well as osteoblast-lineage differentiation of MSCs<sup>50,51</sup>.

Luciferase assays showed that the core of the enhancer is able to activate transcription of a reporter gene. However, the full-length enhancer had a repressor activity in Saos-2 cells, which was neutralised when we tested an extended version of the enhancer, suggesting that there might be a repressor element out of the core of the enhancer that is functional at least in this cell type or that reductionist approaches such as luciferase reporter assays do not capture the complexity of the native genomic environment, especially important in cases where there might be an interplay among different regulatory elements in the region. A recent study in the multipartite enhancer cluster regulating *Ihh*, demonstrated the importance of analysing the enhancers in their native context<sup>52</sup>. Furthermore, eDlx#18 was found to be transcribed in Saos-2 cells, in accordance with the fact that many enhancers can be transcribed, producing non-coding RNA molecules, known as enhancer-RNAs (eRNAs), and that enhancer transcription is an indicator of enhancer activity<sup>53–56</sup>.

Enhancers need to contact somehow the promoters they regulate. Our 4C-seq experiments showed that, in the cells tested in this study, eDlx#18 interacted with several sites within the TAD where it belongs, including the *DLX5/6* region and many of the other tissue-specific *DLX5/6* enhancers described by Birnbaum *et al.*<sup>16</sup> and in the VISTA Enhancer Browser<sup>35</sup>. No other interaction elsewhere in the genome was found. Moreover, the *DLX5* promoter was found to interact with eDlx#18, further supporting that eDlx#18 regulates *DLX5* in bone cell types. However, the *cis*-eQTL analyses in human primary osteoblasts for 2 SNPs lying within the enhancer, showed a nominal association between rs10238953 and *DLX6* gene expression, suggesting that the enhancer might also regulate *DLX6* gene expression. The *DLX5* promoter also interacted with the other enhancers in this region, as previously shown for a few of them in limb embryonic tissue<sup>57,58</sup>. These results might be indicative of some kind of enhancer crosstalk, which would entail a complex regulation of the 3D structure of this genomic region. Little is known about the contribution of each individual enhancer to the regulation of *DLX5/6* expression. They might have distinct specific spatiotemporal activities or might overlap, displaying functional redundancy, presenting additive and/or synergistic effects and, thus, conferring robustness, flexibility, diversity, precision and complexity to the gene expression repertoire<sup>5,52,59,60</sup>. Interestingly, eDlx#18 also showed interaction with the neighbouring *SLC25A13* and *SHFM1* gene bodies. These results, together with the trend for association between the SNP rs10238953 and *SHFM1* gene expression and between rs4613908 and both *SLC25A13* and *SHFM1* gene expression, raise the possibility that different genes included in the same TAD are co-regulated by the same group of long-range enhancers, as previously reported by Nora *et al.*<sup>61</sup> and Gómez-Marín *et al.*<sup>62</sup>. In

addition, *Slc25a13*, *Shfm1*, *Dlx5* and *Dlx6* are similarly expressed in branchial arches, limb bud and otic vesicle of E11.5 mouse embryos<sup>16</sup>.

To analyse the function of eDlx#18, we generated knock-out mice with a homozygous 12 kb deletion containing it. Knock-out E11.5 embryos did not have any overt abnormality, 50% of E17.5 embryos presented exencephaly, and no offspring was born, suggesting that the deletion might decrease embryonic viability. Consistently, previous studies showed that *Dlx5*<sup>-/-</sup>, *Dlx6*<sup>-/-</sup> and *Dlx5/6*<sup>-/-</sup> mice presented perinatal lethality, with a variable percentage of exencephaly or anencephaly<sup>11,12,39</sup>. When we analysed *Dlx5 in situ* gene expression in E11.5 embryos, we detected a considerable reduction in otic placodes and a slight reduction in branchial arches. These results are coherent with the fact that eDlx#18 was described as a branchial arch enhancer in E11.5 mice<sup>16</sup> and that it is included in the SHFM and hearing loss phenotypic subregion, according to Rasmussen *et al.* (Figure 1B)<sup>17</sup>. However, E17.5 embryos did not present any major inner ear abnormality. It would have been interesting to test whether the new-born mice had some deafness problem. Recently, a deletion of 123.6 kb including part of the *SLC25A13* gene as well as two *DLX5/6* enhancers in mice was reported to be associated with a highly reduced expression of *Dlx5* in otic vesicles (E9.5 and E10.5), as well as severe inner ear dysmorphologies and deafness in adult mice<sup>63</sup>. However, when they deleted the enhancers individually, they did not observe any phenotypic abnormality. The reduced *Dlx5* expression in the branchial arches at E11.5 was translated in a slightly smaller dentary with hypoplastic condylar process in E17.5 embryos, similar to the jaws of *Dlx5*<sup>-/-</sup>, *Dlx6*<sup>-/-</sup> and *Dlx5/6*<sup>+/-</sup> mice<sup>64</sup>. Both *Dlx5* and *Dlx6* have been shown to be important for mandibular arch derivatives, as targeted double inactivation of these genes results in the homeotic transformation of the lower jaw into upper jaw<sup>65</sup>. In addition, a haploinsufficiency model has been proposed in which a threshold level of *Dlx5/6* activity is required for the WT morphology of mandibular arch derivatives<sup>39,41</sup>. In this sense, the disruption of eDlx#18 would decrease *Dlx5/6* gene expression to an extent that only causes slight defects in dentary morphology. *Dlx5*<sup>-/-</sup> mice also showed many craniofacial dysmorphic and ossification defects and minor affectations in the ribs<sup>11,44</sup>. Similarly, in our E17.5 embryos, a deficient ossification of supraoccipital bone was observed and several other skeletal defects were detected: a reduced ossification of the vertebral bodies, pelvic bones, and sternum, an incomplete development of the sternum, and dysmorphologies and ectopic cartilage in the ribs. However, ossification problems were much less severe than in *Dlx5*<sup>-/-</sup> mice, consistent with the presence of different enhancers that might regulate *Dlx5/6* gene expression redundantly. Recently, Osterwalder *et al.*<sup>60</sup>, reported several cases of limb-specific enhancers near the same genes that did not

cause any change in limb morphology when deleted individually but caused limb abnormalities when deleted by pairs. In our E17.5 embryos, no limb malformations were observed, consistent with the *Dlx5*<sup>-/-</sup> phenotype. It has been suggested that other *Dlx* genes, such as *Dlx6* or *Dlx2*, compensate for the absence of *Dlx5*, as they are also able to stimulate chondrogenesis and osteogenesis, and compound *Dlx* mutants *Dlx2/5*<sup>-/-</sup> and *Dlx5/6*<sup>-/-</sup> showed severe malformations of the distal limb<sup>12,38,66</sup>. It is worth noting that the *Dlx5/6*<sup>-/-</sup> mice recapitulate the SHFM phenotype seen in humans<sup>12</sup>. Due to the embryonic lethality, we were unable to measure the BMD and other bone parameters in adult mice to evaluate a possible osteopenic phenotype, as it has been observed in *Dlx5*<sup>+/-</sup> mice<sup>67</sup>. Conditional KO of the enhancer in an osteoblastic lineage may be necessary to study its BMD effect in adult animals.

This work has some limitations including the reduced number of embryos analysed and the sample size of the primary osteoblasts used for the eQTL analyses, that may have precluded the detection of small effects. In addition, possible effects of the eDlx#18 enhancer in BMD determination could not be determined since we did not obtain adult mice.

In summary, this study is the first one to analyse in depth the functionality of the eDlx#18 enhancer, within a region associated with BMD and osteoporotic fracture in many GWAS. We provide functional evidence *in vivo* that this enhancer may regulate *DLX5* and *DLX6* in different body locations and it may have an effect on osteogenesis.

#### Acknowledgements

We thank Dr. R. Birnbaum (Ben-Gurion University of the Negev, Israel) and Dr. J.M. Quesada-Gómez (Instituto Maimónides de Investigación Biomédica, Hospital Universitario Reina Sofía, Spain) for kindly providing the *Dlx5* probe plasmid and medulla-derived MSCs, respectively. This work was supported by the following grants: SAF2014-56562-R, SAF2016-75948-R (Spanish MINECO), 2014SGR932 (Generalitat de Catalunya), CIBERER (U720), and FEIOMM Investigación 2014. NRA was a recipient of an FPU predoctoral fellowship from the Spanish Ministerio de Educación, Cultura y Deporte; NMG is a recipient of a FI predoctoral fellowship from Generalitat de Catalunya.

#### Declaration of interests

The authors declare they have no conflict of interests.

## References

1. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
2. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
3. Coppola, C. J., Ramaker, R. C. & Mendenhall, E. M. Identification and function of enhancers in the human genome. *Hum. Mol. Genet.* **25**, R190–R197 (2016).
4. Shlyueva, D., Stampfel, G. & Stark, A. Transcriptional enhancers: From properties to genome-wide predictions. *Nat. Rev. Genet.* **15**, 272–286 (2014).
5. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell* **167**, 1170–1187 (2016).
6. Shen, Y. *et al.* A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116–120 (2012).
7. Kleinjan, D. A. & van Heyningen, V. Long-Range Control of Gene Expression: Emerging Mechanisms and Disruption in Disease. *Am. J. Hum. Genet.* **76**, 8–32 (2005).
8. Lupiáñez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
9. Rickels, R. & Shilatifard, A. Enhancer Logic and Mechanics in Development and Disease. *Trends Cell Biol.* **28**, 608–630 (2018).
10. Merlo, G. R. *et al.* Multiple functions of Dlx genes. *Int. J. Dev. Biol.* **44**, 619–626 (2000).
11. Depew, M. J. *et al.* Dlx5 regulates regional development of the branchial arches and sensory capsules. *Development* **126**, 3831–3846 (1999).
12. Robledo, R. F., Rajan, L., Li, X. & Lufkin, T. The Dlx5 and Dlx6 homeobox genes are essential for craniofacial, axial, and appendicular skeletal development. *Genes Dev.* **16**, 1089–1101 (2002).
13. Sajan, S. A., Rubenstein, J. L. R., Warchol, M. E. & Lovett, M. Identification of direct downstream targets of Dlx5 during early inner ear development. *Hum. Mol. Genet.* **20**, 1262–1273 (2011).
14. Samee, N. *et al.* Dlx5, a positive regulator of osteoblastogenesis, is essential for osteoblast-osteoclast coupling. *Am. J. Pathol.* **173**, 773–780 (2008).
15. Long, J. E., Garel, S., Depew, M. J., Tobet, S. & Rubenstein, J. L. R. DLX5 regulates development of peripheral and central components of the olfactory system. *J. Neurosci.* **23**, 568–78 (2003).

16. Birnbaum, R. Y. *et al.* Functional characterization of tissue-specific enhancers in the DLX5/6 locus. *Hum. Mol. Genet.* **21**, 4930–4938 (2012).
17. Rasmussen, M. B. *et al.* Phenotypic subregions within the split-hand/foot malformation 1 locus. *Hum. Genet.* **135**, 345–357 (2016).
18. Kouwenhoven, E. N. *et al.* Genome-wide profiling of p63 DNA-binding sites identifies an element that regulates gene expression during limb development in the 7q21 shfm1 locus. *PLoS Genet.* **6**, (2010).
19. Tayebi, N. *et al.* Deletions of exons with regulatory activity at the DYNC111 locus are associated with split-hand/split-foot malformation: Array CGH screening of 134 unrelated families. *Orphanet J. Rare Dis.* **9**, 108 (2014).
20. Sowińska-Seidler, A., Badura-Stronka, M., Latos-Bieleńska, A., Stronka, M. & Jamsheer, A. Heterozygous DLX5 Nonsense Mutation Associated with Isolated Split-Hand/Foot Malformation with Reduced Penetrance and Variable Expressivity in two Unrelated Families. *Birth Defects Res. Part A - Clin. Mol. Teratol.* **100**, 764–771 (2014).
21. Ullah, A., Hammid, A., Umair, M. & Ahmad, W. A Novel Heterozygous Intragenic Sequence Variant in DLX6 Probably Underlies First Case of Autosomal Dominant Split-Hand/Foot Malformation Type 1. *Mol. Syndromol.* **8**, 79–84 (2017).
22. Shamseldin, H. E., Faden, M. A., Alashram, W. & Alkuraya, F. S. Identification of a novel DLX5 mutation in a family with autosomal recessive split hand and foot malformation. *J. Med. Genet.* **49**, 16–20 (2012).
23. Estrada, K. *et al.* Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat. Genet.* **44**, 491–501 (2012).
24. Morris, J. A. *et al.* An atlas of genetic influences on osteoporosis in humans and mice. *Nat. Genet.* (2018). doi:doi: 10.1038/s41588-018-0302-x
25. Zheng, H. F. *et al.* Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature* **526**, 112–117 (2015).
26. Roca-Ayats, N. *et al.* Functional characterization of the C7ORF76 genomic region, a prominent GWAS signal for osteoporosis in 7q21.3. *Bone* **123**, 39–47 (2019).
27. Bujold, D. *et al.* The International Human Epigenome Consortium Data Portal. *Cell Syst.* **3**, 496–499.e2 (2016).
28. Forrest, A. R. R. *et al.* A promoter-level mammalian expression atlas. *Nature* **507**, 462–470 (2014).
29. van de Werken, H. J. G. *et al.* in *Methods in Enzymology* (eds. Wu, C. & Allis, C. D.) **513**, 89–112 (Elsevier Inc., 2012).
30. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler

- transform. *Bioinformatics* **25**, 1754–1760 (2009).
31. Artus, J. & Hadjantonakis, A.-K. in *Transgenic Mouse Methods and Protocols. Methods in Molecular Biology* **693**, 37–56 (2011).
  32. Ferran, J. L. *et al.* in *In Situ Hybridization Methods. Neuromethods* (ed. Hauptmann, G.) **99**, 61–82 (Humana Press, 2015).
  33. Zerucha, T. *et al.* A highly conserved enhancer in the *Dlx5/Dlx6* intergenic region is the site of cross-regulatory interactions between *Dlx* genes in the embryonic forebrain. *J. Neurosci.* **20**, 709–721 (2000).
  34. Rocha, P. P., Bleiss, W. & Schrewe, H. Mosaic expression of *Med12* in female mice leads to exencephaly, spina bifida, and craniorachischisis. *Birth Defects Res. Part A - Clin. Mol. Teratol.* **88**, 626–632 (2010).
  35. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA Enhancer Browser - A database of tissue-specific human enhancers. *Nucleic Acids Res.* **35**, D88-92 (2007).
  36. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
  37. Wang, Y. *et al.* The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome Biol.* **19**, 151 (2018).
  38. Li, H. *et al.* Expression and function of *Dlx* genes in the osteoblast lineage. *Dev. Biol.* **316**, 458–470 (2008).
  39. Bendall, A. J. Direct evidence of allele equivalency at the *Dlx5/6* locus. *Genesis* **54**, 272–276 (2016).
  40. Panganiban, G. & Rubenstein, J. L. R. Developmental functions of the *Distal-less/Dlx* homeobox genes. *Development* **129**, 4371–86 (2002).
  41. Quach, A., MacKenzie, R. K. & Bendall, A. J. Measuring inputs to a common function: The case of *Dlx5* and *Dlx6*. *Biochem. Biophys. Res. Commun.* **478**, 371–377 (2016).
  42. Perera, M. *et al.* Defective neuronogenesis in the absence of *Dlx5*. *Mol. Cell. Neurosci.* **25**, 153–161 (2004).
  43. Fazel Darbandi, S. *et al.* Functional consequences of *I56ii* *Dlx* enhancer deletion in the developing mouse forebrain. *Dev. Biol.* **420**, 32–42 (2016).
  44. Acampora, D. *et al.* Craniofacial, vestibular and bone defects in mice lacking the *Distal-less* related gene *Dlx5*. *Development* **126**, 3795–3809 (1999).
  45. Bendall, A. J., Hu, G., Levi, G. & Abate-Shen, C. *Dlx5* regulates chondrocyte differentiation at multiple stages. *Int. J. Dev. Biol.* **47**, 335–344 (2003).
  46. Andrey, G. & Mundlos, S. The three-dimensional genome: regulating gene

- expression during pluripotency and development. *Development* **144**, 3646–3658 (2017).
47. De Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499–506 (2013).
  48. Montavon, T. *et al.* A Regulatory Archipelago Controls Hox Genes Transcription in Digits. *Cell* **147**, 1132–1145 (2011).
  49. Raisner, R. *et al.* Enhancer Activity Requires CBP/P300 Bromodomain-Dependent Histone H3K27 Acetylation. *Cell Rep.* **24**, 1722–1729 (2018).
  50. Wang, Y. *et al.* RUNX3 plays an important role in mediating the BMP9-induced osteogenic differentiation of mesenchymal stem cells. *Int. J. Mol. Med.* **40**, 1991–1999 (2017).
  51. Yoshida, C. A. *et al.* Runx2 and Runx3 are essential for chondrocyte maturation, and Runx2 regulates limb growth through induction of Indian hedgehog. *Genes Dev.* **18**, 952–963 (2004).
  52. Will, A. J. *et al.* Composition and dosage of a multipartite enhancer cluster control developmental expression of *Ihh* (Indian hedgehog). *Nat. Genet.* **49**, 1539–1545 (2017).
  53. Mikhaylichenko, O. *et al.* The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes Dev.* **32**, 42–57 (2018).
  54. Andersson, R. *et al.* An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014).
  55. Hon, C. C. *et al.* An atlas of human long non-coding RNAs with accurate 5' ends. *Nature* **543**, 199–204 (2017).
  56. Arner, E. *et al.* Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. *Science* **347**, 1010–1014 (2015).
  57. Birnbaum, R. Y. *et al.* Coding exons function as tissue-specific enhancers of nearby genes. *Genome Res.* **22**, 1059–1068 (2012).
  58. Verzi, M. P. *et al.* The Transcription Factor MEF2C Is Required for Craniofacial Development. *Dev. Cell* **12**, 645–652 (2007).
  59. Ahituv, N. *et al.* Deletion of ultraconserved elements yields viable mice. *PLoS Biol.* **5**, 1906–1911 (2007).
  60. Osterwalder, M. *et al.* Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**, 239–243 (2018).
  61. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
  62. Gómez-Marín, C. *et al.* Evolutionary comparison reveals that diverging CTCF sites

- are signatures of ancestral topological associating domains borders. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7542–7547 (2015).
63. Johnson, K. R. *et al.* Deletion of a long-range *Dlx5* enhancer disrupts inner ear development in mice. *Genetics* **208**, 1165–1179 (2018).
  64. Jeong, J. *et al.* *Dlx* genes pattern mammalian jaw primordium by regulating both lower jaw-specific and upper jaw-specific genetic programs. *Development* **135**, 2905–2916 (2008).
  65. Beverdam, A. *et al.* Jaw transformation with gain of symmetry after *Dlx5/Dlx6* inactivation: Mirror of the past? *Genesis* **34**, 221–227 (2002).
  66. Depew, M. J., Simpson, C. A., Morasso, M. & Rubenstein, J. L. R. Reassessing the *Dlx* code: The genetic regulation of branchial arch skeletal pattern and development. *J. Anat.* **207**, 501–561 (2005).
  67. Samee, N. *et al.* Increased bone resorption and osteopenia in *Dlx5* heterozygous mice. *J. Cell. Biochem.* **107**, 865–872 (2009).

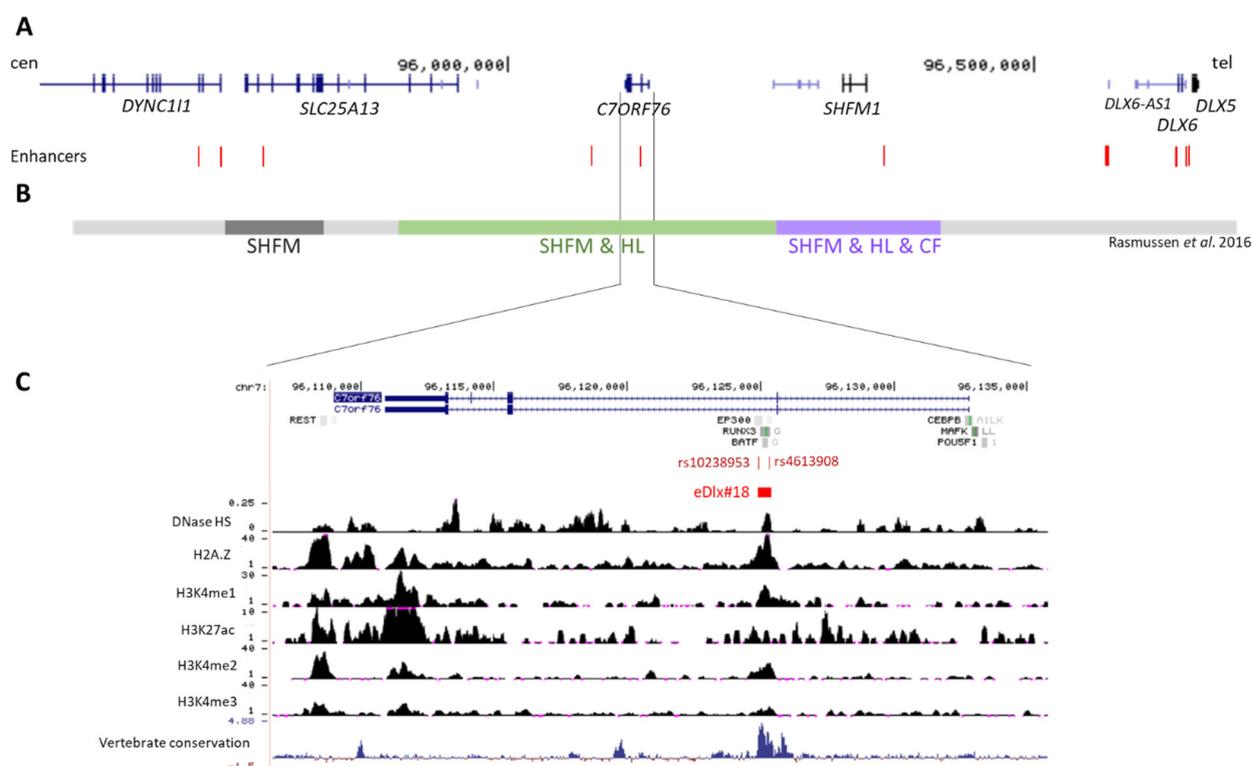
## Tables

**Table 1.** *cis*-eQTL analysis of 2 SNPs located within eDlx#18

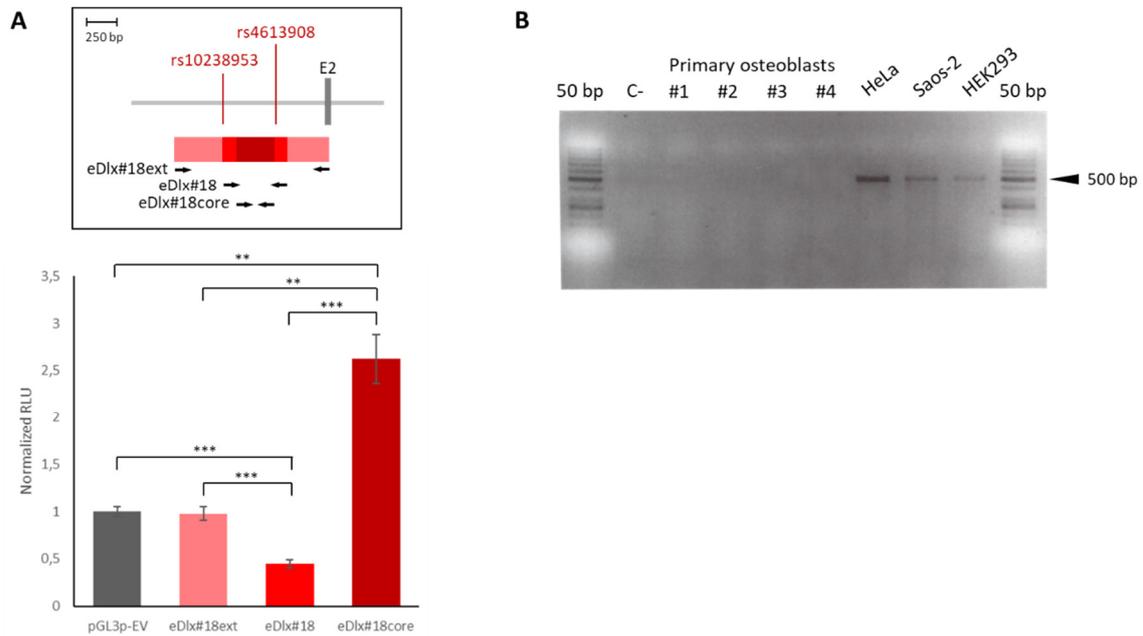
SNP	p-values				
	<i>DLX5</i>	<i>DLX6</i>	<i>AS-DLX6</i>	<i>SHFM1</i>	<i>SLC25A13</i>
rs4613908	0.29809	0.48637	0.47592	0.07315	0.07315
rs10238953	0.54728	<i>0.04667</i>	0.36131	0.07257	0.67308

Values in italics indicate nominal significance

## Figures

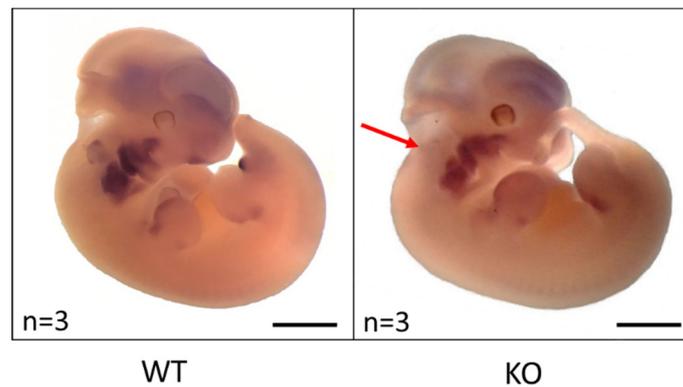


**Figure 1.** Genomic location of eDlx#18 enhancer. **A.** The SHFM locus of human chromosome 7. Position of the *DYNC111*, *SLC25A13*, *C7ORF76*, *SHFM1*, *DLX6-AS1*, *DLX6* and *DLX5* genes are shown. In red, previously reported tissue-specific enhancers within the region<sup>16</sup>. **B.** Phenotypic subregions within the SHFM locus established by Rasmussen *et al.*<sup>17</sup>. In grey, subregion associated with isolated split hand and foot malformation (SHFM); in green, subregion associated with SHFM with hearing loss (HL); and in purple, subregion associated with SHFM, HL and craniofacial anomalies (CF). **C.** eDlx#18 region in detail. In red, eDlx#18 location and in dark red, SNPs present in the enhancer. Transcription factor ChIP-seq, DNase HS, H2A.Z, H3K4me1, H3K27ac, H3K4me2 and H3K4me3 ChIP-seq from osteoblasts (ENCODE), and vertebrate conservation are shown.

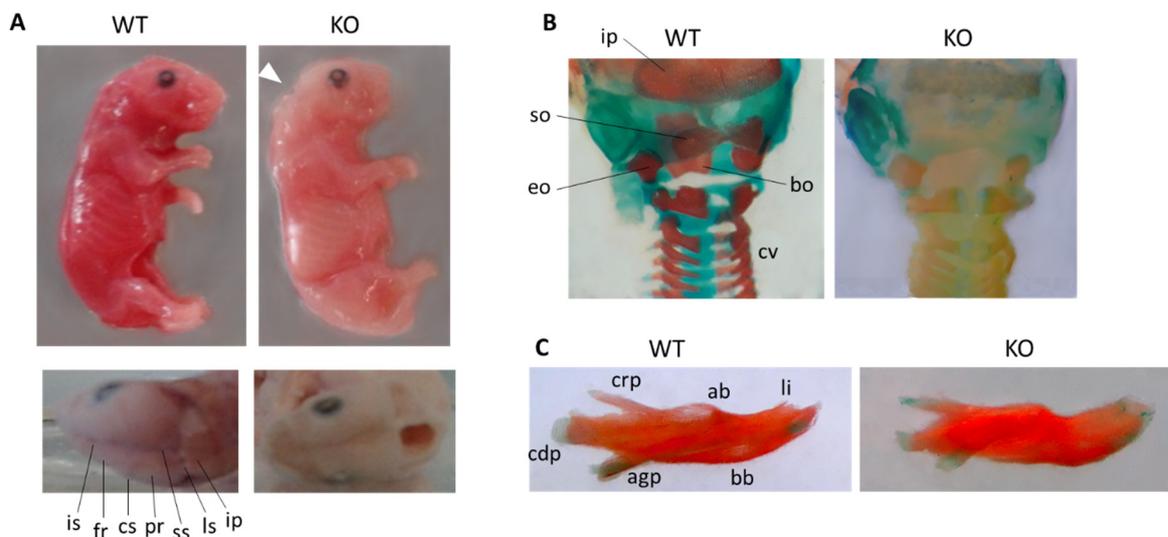


**Figure 2. A.** Luciferase activity of different constructs of eDlx#18 (shown in the box) cloned in pGL3-Promoter vector, upstream of a strong promoter (SV40), in Saos-2 cells. Results are expressed as mean $\pm$ SD. \*\* $p < 0.01$ ; \*\*\* $p < 0.001$  **B.** PCR amplification of eDlx#18 from cDNA of different cell types. Expected size: 480 bp.

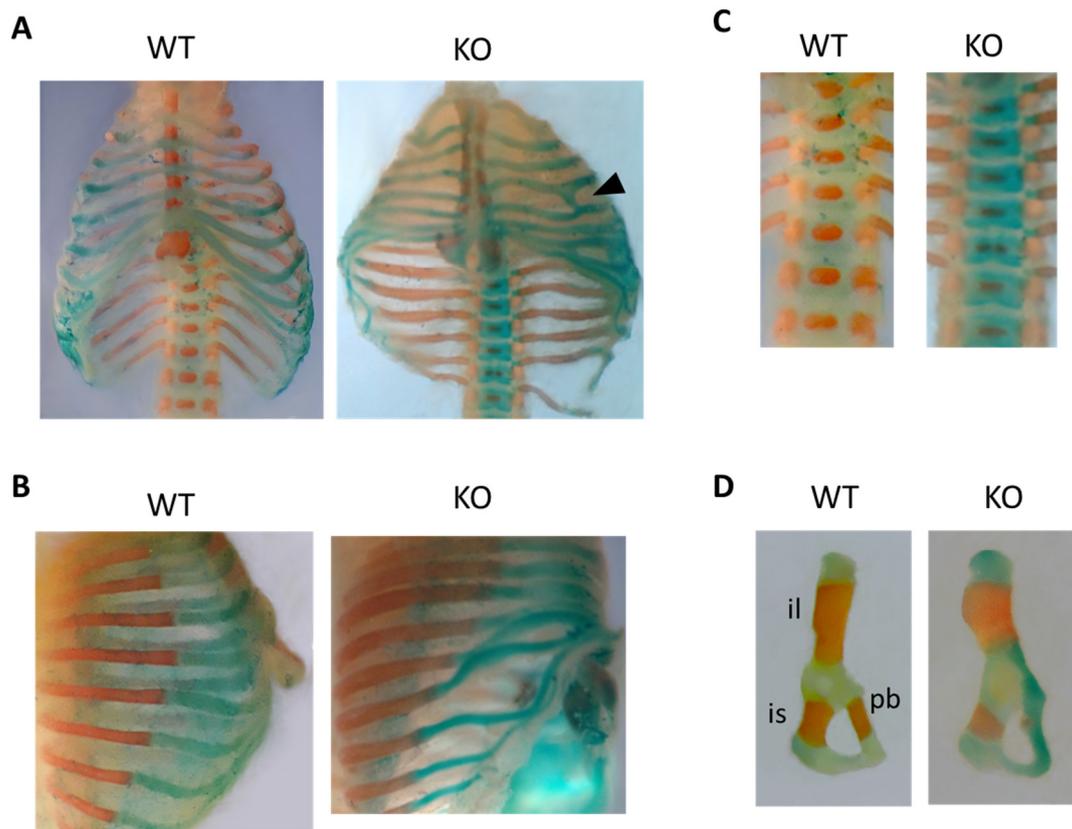




**Figure 4.** Whole-mount in situ hybridization for *Dlx5* in wild-type (left) and homozygous eDlx18 knock-out (right) E11.5 mouse embryos. WT embryos showed *Dlx5* expression in the branchial arches, genital tubercle, otic vesicle, forebrain, frontonasal prominence and limb buds (both in the AER and in the anterior limb bud). In the KO embryos, *Dlx5* expression appeared severely reduced in the otic vesicle (red arrow). Bars= 1 mm



**Figure 5.** General aspect and craniofacial affectations of E17.5 KO mice. **A.** General aspect of mice. KO mice presented a whitish colour (upper panels) and in 2/4 mice exencephaly was observed, one of which presenting a hole at the interparietal zone (white arrowhead and zenital view of EtOH-fixed mice at lower panels). **B.** Back view of mice skull, differentially stained for bone (alizarin red) and cartilage (alcian blue). KO mice had a deficient ossification of the supraoccipital bone. **C.** Lateral view of the mandible, differentially stained for bone and cartilage. Dentary of KO mice was slightly smaller with a reduced condylar process. ab, alveolar bone of mandible; agp, angular process of mandible; bb, basal bone of mandible; bo, basioccipital; cdp, condylar process of mandible; crp, coronoid process of mandible; cs, coronal suture; cv, cervical vertebrae; eo, exoccipital; fr, frontal; ip, interparietal; is, interfrontal suture; li, lower incisor; ls, lambdoid suture; pr, parietal; so, supraoccipital; ss, sagittal suture



**Figure 6.** Axial and appendicular skeletal defects of E17.5 KO mice, differentially stained for bone (alizarin red) and cartilage (alcian blue). **A.** Frontal view of rib cage. KO mice presented and incomplete fusion of the sternum and ectopic cartilage fusing ribs (black arrowhead). **B.** Lateral view of rib cage. KO mice showed abnormally curved irregular ribs. **C.** Vertebrae of KO mice had reduced ossification centres. **D.** Pelvic girdle of KO mice showed a reduced ossification. il, ilium; is, ischium; pb, pubis

**Supplementary Information:**

**Supplementary Table 1.** Primers used for cloning the different constructs of eDlx#18 in pGL3-Promoter vector

Construct	Primer Fwd	Primer Rev
eDlx#18ext	TGCTATGCAGCTCTGAAGGA	TCACTGCTCTTAGGGTGAAGTCAAA
eDlx#18	TCTCAGATTAAGAAACAACACC	CTGGCTTGGTTGACAGACAG
eDlx#18core	GCCAGCCATCTGTGTTTCATT	AGTTTCCGCGATCTTCCTTT

**Supplementary Table 2.** Primers used for qPCR assays

Gene	Primer Fwd	Primer Rev
<i>DLX5</i>	ctacaaccgctccaag	gccattcaccattctcacct
<i>DLX6</i>	atatattagagaagagcgaggagag	ccctctgcagccaccta
<i>AS-DLX6</i>	tgattcctgtatgatggcagcta	ggtttcctttgtctcagcaat
<i>SLC25A13</i>	agatgggtcgggtcccactt	gcaaacggatcttgacgatt
<i>SHFM1</i>	gacgacgagttgaagagttcc	ccaattatcctcccagaca
<i>HMBS</i>	tgccctggagaagaatgaag	cagcatcatgagggtttcc

**Supplementary Table 3.** Guide RNAs for CRISPR/Cas9 and genotyping primers

Construct	sgRNA	Genomic coordinates (mm9)	Genotyping primers
eDlx#18del	ACAAATAAACCCCTGACATCA	chr6:6,304,297-6,304,316	F: CACTGAAAAAGCCAGAGAAGA
	ATATTCCTACCAACCATGGT	chr6:6,316,607-6,316,626	R: CAGTTAGGCACTGTGGAAGC

**Supplementary Table 4.** Primers used for 4C-seq experiments

Viewpoint	1st primer	Genomic coordinates (GRCh37/mm9)	2nd primer	Genomic coordinates (GRCh37/mm9)
eDlx#18 (human)	GACAGCTTGTCAGGAAATGATC	chr7:96,124,920-96,124,941	GTTTATTCAAGGCCCTCTGG	chr7:96,124,878-96,124,897
<i>DLX5</i> promoter (human)	CCCGCAAAGGTGAATGGATC	chr7:96,655,042-96,655,061	ACAGAGCCTTGTGCTGTGG	chr7:96,654,432-96,654,452
eDlx#18 (mouse)	TTTCTGGCTGAGAACTGATT	chr6:6,316,501-6,316,520	ATGCAAGGAAGGGATAAACT	chr6:6,315,255-6,315,274
<i>Dlx5</i> promoter (mouse)	AAGAACCGCATCCTCTAAAC	chr6:6,832,546-6,832,565	AGTGTGCCTCCAGACCAAA	chr6:6,833,079-6,833,097

## CHAPTER 2: IDENTIFICATION AND CHARACTERIZATION OF GENETIC SUSCEPTIBILITY TO N-BP-ASSOCIATED AFF

### Article 3

*GGPS1* Mutation and Atypical Femoral Fractures with Bisphosphonates

**Summary:** -

**Reference:** Neus Roca-Ayats, Susana Balcells, Natàlia Garcia-Giralt, Maite Falcó-Mascaró, Núria Martínez-Gil, Josep F. Abril, Roser Urreizti, Joaquín Dopazo, José M. Quesada-Gómez, Xavier Nogués, Leonardo Mellibovsky, Daniel Prieto-Alhambra, James E. Dunford, Muhammad K. Javaid, R. Graham Russell, Daniel Grinberg, Adolfo Díez-Pérez. *GGPS1* Mutation and Atypical Femoral Fractures with Bisphosphonates. *The New England Journal of Medicine*. 2017;376(18):1794-1795. doi: 10.1056/NEJMc1612804.



## CORRESPONDENCE



## GGPS1 Mutation and Atypical Femoral Fractures with Bisphosphonates

**TO THE EDITOR:** Atypical femoral fractures have been associated with long-term bisphosphonate treatment.<sup>1,2</sup> However, the underlying mechanisms remain obscure. We studied three sisters who had atypical femoral fractures after receiving various oral bisphosphonates for 6 years. Two of the sisters had a single fracture (at the ages of 64 and 73 years), and one had bilateral fractures (one at the age of 60 years and the other at the age of 61 years). Given the low incidence of atypical femoral fractures in the general population (5.9 per 10,000 person-years),<sup>3</sup> we hypothesized that these sisters might have an underlying genetic background that contributed to these fractures.

We performed whole-exome sequencing to detect possible shared genetic variants involved in their apparent increased risk. In addition, we performed whole-exome sequencing in three unrelated patients with atypical femoral fractures who each had received bisphosphonates for more than 5 years. We prioritized rare nonsynonymous mutations in the variant filtering, and only mutations that were shared among the three sisters were considered. No mutation was found to be

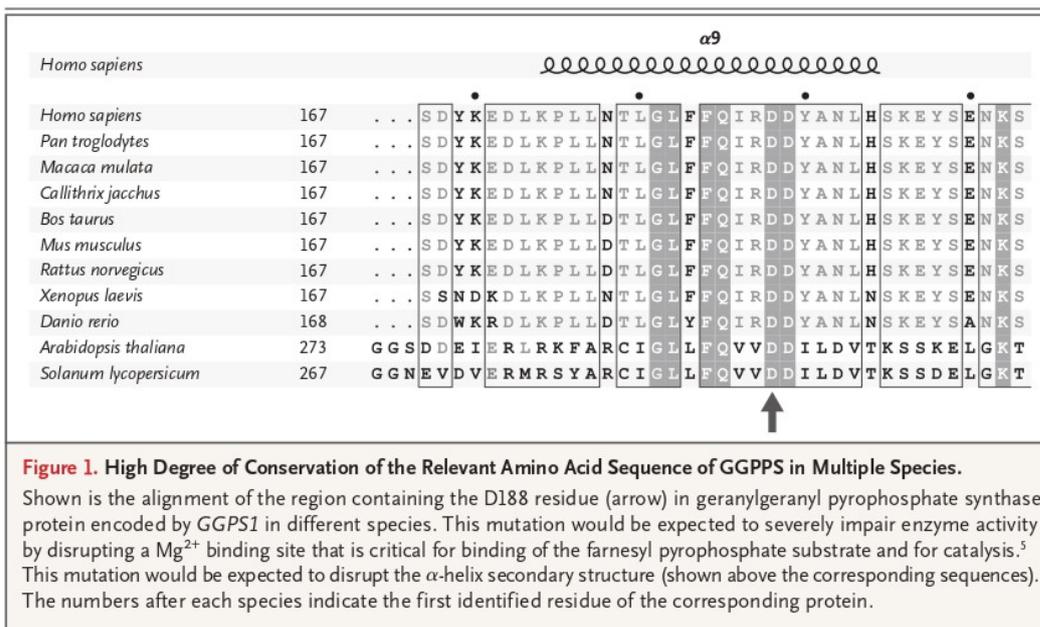
homozygous or in any gene containing mutations in both chromosomes (compound heterozygous). Assuming that a dominant model was involved, we detected 37 rare mutations (in 34 genes), among them a novel p.Asp188Tyr substitution in the enzyme geranylgeranyl pyrophosphate synthase (GGPPS), which is a site of inhibition by bisphosphonates in the mevalonate pathway.<sup>4</sup> The variant that is located in the genomic position g.235505746G→T on chromosome 1 (GRCh37/hg19) in *GGPS1* had the best conservation score and was not described in any of the available databases. This variant would be expected to severely impair the enzyme activity (Fig. 1). Furthermore, the gene encoding cytochrome P-450 family 1 subfamily A member 1 (*CYP1A1*), which is involved in steroid metabolism, was also mutated in all three sisters and in one of the unrelated patients, which suggests that it could be another potential susceptibility gene for bisphosphonate-related atypical femoral fractures. An additional mutation in the gene encoding mevalonate diphosphate decarboxylase (*MVD*) was detected in one unrelated patient.

Pathway analysis of the mutated genes showed enrichment of the isoprenoid biosynthetic pathway (GO:0008299), which includes *GGPS1*, *CYP1A1*, and *MVD* ( $P < 0.001$ ). We speculate that other variants that have been identified might also be involved in susceptibility to bisphosphonate-related atypical femoral fractures. Such variants include missense changes in the gene encoding fibronectin 1 (*FNI*) and in the genes encoding synapse defective Rho GTPase homolog 2 (*SYDE2*) and neuronal guanine nucleotide exchange factor (*NGEF*); the latter two proteins are regulators of small GTPases. We speculate that our results may support a model in which accumulation of susceptibility variants (including some in rele-

## THIS WEEK'S LETTERS

- 1794 **GGPS1 Mutation and Atypical Femoral Fractures with Bisphosphonates**
- 1795 **Crizanlizumab in Sickle Cell Disease**
- 1797 **Approach to Fever in the Returning Traveler**
- e37 **Bilateral versus Single Internal-Thoracic-Artery Grafts**
- e38 **A Highly Durable RNAi Therapeutic Inhibitor of PCSK9**

CORRESPONDENCE



**Figure 1. High Degree of Conservation of the Relevant Amino Acid Sequence of GGPPS in Multiple Species.**

Shown is the alignment of the region containing the D188 residue (arrow) in geranylgeranyl pyrophosphate synthase protein encoded by *GGPS1* in different species. This mutation would be expected to severely impair enzyme activity by disrupting a  $Mg^{2+}$  binding site that is critical for binding of the farnesyl pyrophosphate substrate and for catalysis.<sup>5</sup> This mutation would be expected to disrupt the  $\alpha$ -helix secondary structure (shown above the corresponding sequences). The numbers after each species indicate the first identified residue of the corresponding protein.

vant genes, notably *GGPS1*) may lead to a possible genetic component of predisposition to atypical femoral fractures.

Neus Roca-Ayats, M.Sc.  
 Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)  
 Barcelona, Spain

Susana Balcells, Ph.D.  
 University of Barcelona  
 Barcelona, Spain

Natàlia Garcia-Giralt, Ph.D.  
 Centro de Investigación Biomédica en Red de Fragilidad y Envejecimiento Saludable (CIBERFES)  
 Barcelona, Spain

Maite Falcó-Mascaró, M.Sc.  
 Núria Martínez-Gil, M.Sc.

Josep F. Abril, Ph.D.  
 University of Barcelona  
 Barcelona, Spain

Roser Urreizti, Ph.D.  
 CIBERER  
 Barcelona, Spain

Joaquín Dopazo, Ph.D.  
 CIBERER  
 Valencia, Spain

José M. Quesada-Gómez, M.D., Ph.D.  
 CIBERER  
 Cordoba, Spain

Xavier Nogués, M.D., Ph.D.  
 CIBERFES  
 Barcelona, Spain

Leonardo Mellibovsky, M.D., Ph.D.  
 Institut Hospital del Mar d'Investigacions Mèdiques  
 Barcelona, Spain

Daniel Prieto-Alhambra, M.D., Ph.D.  
 James E. Dunford, Ph.D.  
 Muhammad K. Javaid, M.B., B.S., Ph.D.  
 University of Oxford  
 Oxford, United Kingdom

R. Graham Russell, M.D., Ph.D.  
 University of Sheffield  
 Sheffield, United Kingdom  
 Daniel Grinberg, Ph.D.  
 CIBERER  
 Barcelona, Spain

Adolfo Díez-Pérez, M.D., Ph.D.  
 CIBERFES  
 Barcelona, Spain  
 adiez@parcadesalutmar.cat

Supported by grants (SAF2011-25431 and SAF2014-56562R) from the Spanish Ministry of Economy and Competitiveness, grants (2009SGR971 and 2014SGR932) from the Catalan government, a grant (FPU13/02066) from the Spanish Ministry of Education, Culture and Sports, a grant (U720) from CIBERER, and a grant (FIS PI13/00444) to CIBERFES from Instituto Carlos III, and by FEDER funds.

Disclosure forms provided by the authors are available with the full text of this letter at NEJM.org.

1. Shane E, Burr D, Abrahamsen B, et al. Atypical subtrochanteric and diaphyseal femoral fractures: second report of a task force of the American Society for Bone and Mineral Research. *J Bone Miner Res* 2014;29:1-23.
2. Schilcher J, Michaëlsson K, Aspenberg P. Bisphosphonate use and atypical fractures of the femoral shaft. *N Engl J Med* 2011;364:1728-37.
3. Feldstein AC, Black D, Perrin N, et al. Incidence and demography of femur fractures with and without atypical features. *J Bone Miner Res* 2012;27:977-86.
4. Russell RG. Bisphosphonates: the first 40 years. *Bone* 2011;49:2-19.
5. Kavanagh KL, Dunford JE, Bunkoczi G, Russell RGG, Oppermann U. The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a novel hexameric arrangement and inhibitory product binding. *J Biol Chem* 2006;281:22004-12.

DOI: 10.1056/NEJMc1612804

## Article 4

### Functional Characterization of a GGPPS Variant Identified in Atypical Femoral Fracture Patients and Delineation of the Role of GGPPS in Bone-Relevant Cell Types

#### Summary:

Atypical femoral fractures (AFFs) are a rare but potentially devastating event, often but not always linked to bisphosphonate (BP) therapy. The pathogenic mechanisms underlying AFFs remain obscure, and there are no tests available that might assist in identifying those at high risk of AFF. We previously used exome sequencing to explore the genetic background of three sisters with AFFs and three additional unrelated AFF cases, all previously treated with BPs. We detected 37 rare mutations (in 34 genes) shared by the three sisters. Notably, we found a p.Asp188Tyr mutation in the enzyme geranylgeranyl pyrophosphate synthase, a component of the mevalonate pathway, which is critical to osteoclast function and is inhibited by N-BPs. In addition, the CYP1A1 gene, responsible for the hydroxylation of 17 $\beta$ -estradiol, estrone, and vitamin D, was also mutated in all three sisters and one unrelated patient. Here we present a detailed list of the variants found and report functional analyses of the GGPS1 p.Asp188Tyr mutation, which showed a severe reduction in enzyme activity together with oligomerization defects. Unlike BP treatment, this genetic mutation will affect all cells in the carriers. RNAi knockdown of GGPS1 in osteoblasts produced a strong mineralization reduction and a reduced expression of osteocalcin, osterix, and RANKL, whereas in osteoclasts, it led to a lower resorption activity. Taken together, the impact of the mutated GGPPS and the relevance of the downstream effects in bone cells make it a strong candidate for AFF susceptibility. We speculate that other genes such as CYP1A1 might be involved in AFF pathogenesis, which remains to be functionally proved. The identification of the genetic background for AFFs provides new insights for future development of novel risk assessment tools.

#### Reference:

Neus Roca-Ayats\*, Pei Ying Ng\*, Natàlia Garcia-Giralt, Maite Falcó-Mascaró, Mónica Cozar, Josep Francesc Abril, José Manuel Quesada Gómez, Daniel Prieto-Alhambra, Xavier Nogués, James E Dunford, R Graham Russell, Roland Baron, Daniel Grinberg, Susana Balcells, Adolfo Díez-Pérez. Functional Characterization of a GGPPS Variant Identified in Atypical Femoral Fracture Patients and Delineation of the Role of GGPPS

in Bone-Relevant Cell Types. *Journal of Bone and Mineral Research*. 2018;33(12):2091-2098. doi: 10.1002/jbmr.3580

\*co-first authors



## Functional Characterization of a GGPS Variant Identified in Atypical Femoral Fracture Patients and Delineation of the Role of GGPS in Bone-Relevant Cell Types

Neus Roca-Ayats,<sup>1\*</sup> Pei Ying Ng,<sup>2\*</sup> Natàlia Garcia-Giralt,<sup>3</sup> Maite Falcó-Mascaró,<sup>1</sup> Mónica Cozar,<sup>1</sup> Josep Francesc Abril,<sup>4</sup> José Manuel Quesada Gómez,<sup>5</sup> Daniel Prieto-Alhambra,<sup>6,7</sup> Xavier Nogués,<sup>3</sup> James E Dunford,<sup>7</sup> R Graham Russell,<sup>7,8</sup> Roland Baron,<sup>2</sup> Daniel Grinberg,<sup>1\*\*</sup> Susana Balcells,<sup>1\*\*</sup> and Adolfo Díez-Pérez<sup>3\*\*</sup>

<sup>1</sup>Department of Genetics, Microbiology, and Statistics, Facultat de Biologia, Universitat de Barcelona, Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), ISCIII, IBUB, IRSJD, Barcelona, Spain

<sup>2</sup>Division of Bone and Mineral Research, Department of Oral Medicine, Harvard School of Dental Medicine, and Department of Medicine, Harvard Medical School, Boston, MA, USA

<sup>3</sup>Musculoskeletal Research Group, IMIM (Hospital del Mar Medical Research Institute), Centro de Investigación Biomédica en Red en Fragilidad y Envejecimiento Saludable (CIBERFES), ISCIII, Barcelona, Spain

<sup>4</sup>Department of Genetics, Microbiology, and Statistics, Facultat de Biologia, Universitat de Barcelona, IBUB, Barcelona, Spain

<sup>5</sup>Mineral Metabolism Unit, Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC), Hospital Universitario Reina Sofía, CIBERFES, ISCII, Córdoba, Spain

<sup>6</sup>GREMPAL (Grup de Recerca en Malalties Prevalents de l'Aparell Locomotor), Idiap Jordi Gol Primary Care Research Institute, CIBERFES, Autonomous University of Barcelona, Barcelona, Spain

<sup>7</sup>NIHR Musculoskeletal BRU and Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, Oxford, UK

<sup>8</sup>The Mellanby Centre for Bone Research, Department of Oncology and Metabolism, University of Sheffield, Sheffield, UK

### ABSTRACT

Atypical femoral fractures (AFFs) are a rare but potentially devastating event, often but not always linked to bisphosphonate (BP) therapy. The pathogenic mechanisms underlying AFFs remain obscure, and there are no tests available that might assist in identifying those at high risk of AFF. We previously used exome sequencing to explore the genetic background of three sisters with AFFs and three additional unrelated AFF cases, all previously treated with BPs. We detected 37 rare mutations (in 34 genes) shared by the three sisters. Notably, we found a p.Asp188Tyr mutation in the enzyme geranylgeranyl pyrophosphate synthase, a component of the mevalonate pathway, which is critical to osteoclast function and is inhibited by N-BPs. In addition, the *CYP11A1* gene, responsible for the hydroxylation of 17 $\beta$ -estradiol, estrone, and vitamin D, was also mutated in all three sisters and one unrelated patient. Here we present a detailed list of the variants found and report functional analyses of the *GGPS1* p.Asp188Tyr mutation, which showed a severe reduction in enzyme activity together with oligomerization defects. Unlike BP treatment, this genetic mutation will affect all cells in the carriers. RNAi knockdown of *GGPS1* in osteoblasts produced a strong mineralization reduction and a reduced expression of osteocalcin, osterix, and RANKL, whereas in osteoclasts, it led to a lower resorption activity. Taken together, the impact of the mutated GGPPS and the relevance of the downstream effects in bone cells make it a strong candidate for AFF susceptibility. We speculate that other genes such as *CYP11A1* might be involved in AFF pathogenesis, which remains to be functionally proved. The identification of the genetic background for AFFs provides new insights for future development of novel risk assessment tools. © 2018 American Society for Bone and Mineral Research.

**KEY WORDS:** ATYPICAL FEMORAL FRACTURES; BISPHOSPHONATES; *GGPS1*; WES

Received in original form January 25, 2018; revised form August 6, 2018; accepted August 26, 2018. Accepted manuscript online September 5, 2018.

Address correspondence to: Natalia Garcia-Giralt, PhD, IMIM-Hospital del Mar, C/ Dr. Aiguader 88, 08003 Barcelona, Spain. E-mail: ngarcia@imim.es

\*NR-A and PY are co-first authors.

\*\*DG, SB, and AD-P are co-last authors.

Additional Supporting Information may be found in the online version of this article.

Journal of Bone and Mineral Research, Vol. 33, No. 12, December 2018, pp 2091–2098

DOI: 10.1002/jbmr.3580

© 2018 American Society for Bone and Mineral Research

## Introduction

Osteoporosis with its associated fractures is the most common postmenopausal bone disorder, but it also affects older men and women of all ethnicities. Nitrogen-containing bisphosphonates (N-BPs) are currently the most commonly used treatments for osteoporotic disease in millions of patients worldwide. Although the clinically important antifracture efficacy of BPs and their overall safety have been robustly demonstrated in several clinical trials<sup>(1)</sup> and systematic reviews,<sup>(2,3)</sup> a number of uncommon adverse effects potentially associated with prolonged use of these drugs have also been described, among them atypical femoral fractures (AFFs).<sup>(4)</sup> These fractures, characterized by their location in the subtrochanteric region or femoral shaft, are distinct from classic osteoporotic fragility fractures.<sup>(5)</sup>

The pathogenic mechanisms underlying AFFs remain obscure, and there has been much speculation about the causes.<sup>(5)</sup> Given the low absolute incidence of AFFs, it may be hypothesized that rare underlying genetic causes may increase susceptibility to these fractures, which may then occur spontaneously or be triggered after additional interactions with bisphosphonates (BPs) or other antiresorptive drugs. Currently, there are no tests available, genetic or biochemical, that may assist in identifying those at high risk of AFFs. Identification of genetic determinants of AFF would therefore shed light on etiological mechanisms and lead not only to novel diagnostic and risk algorithms for the millions of patients taking bisphosphonates for either osteoporosis or cancer-related bone disease but also to possible therapeutic strategies for patients with delayed fracture or nonunion.

Previously, we identified 3 sisters who have been treated with BPs for more than 5 years and diagnosed with AFFs.<sup>(6)</sup> This observation suggested that a patient's genetic background may predispose the individual to AFF after long-term BP therapy. Accordingly, we performed whole-exome sequencing (WES) to identify potential AFF-related mutations in these three sisters and three other unrelated long-term BP-treated patients with AFFs. We identified several variants, which we list here. Among them, we identified the p.Asp188Tyr mutation in the geranylgeranyl diphosphate synthase (*GGPS1*) gene.<sup>(6)</sup> Given that this enzyme is a site of inhibition by bisphosphonates in the mevalonate pathway, we focused on the mutation found for further functional studies. We demonstrate that p.Asp188Tyr markedly reduced GGPP synthase activity. Using shRNA-mediated knockdown of *GGPS1* in both mouse calvarial and mouse macrophage cells lines, to recapitulate the global loss of synthase activity due to the p.Asp188Tyr mutation, we showed that loss of GGPPS function resulted in defective osteoblast and osteoclast activity. Therefore, this mutation may possibly explain the bone fragility in these patients, possibly exacerbated by BP treatment.

## Materials and Methods

### Subjects

For whole-exome sequencing analysis, 6 patients with AFFs who had received long-term (>5 years) treatment with BPs were recruited: 3 sisters from Hospital Universitario Reina Sofía (Córdoba, Spain) and 3 unrelated women from Hospital del Mar (Barcelona, Spain). Given that the clinical phenotype may be related in the majority of cases to exposure to BPs, we also

selected 3 women with more than 6 years of BP treatment but with no history of AFF. Baseline characteristics of AFF patients and controls are described in Supplemental Table S1. The 3 affected sisters, all with hypercholesterolemia, had been on statins and received regularly proton pump inhibitors (PPIs) but no glucocorticoids or other bone-acting agent except BPs. Their mother had a forearm fracture as well as 2 of the 3 sisters. Written informed consent was obtained from all patients in accordance with the regulations of the Clinical Research Ethics Committee of Parc de Salut Mar, which approved the study.

### Whole-exome sequencing

DNA of patients with AFF was extracted from peripheral blood with the Wizard Genomic DNA Purification Kit (Promega, Madrid, Spain) and used for whole-exome sequencing in the CNAG platform (Barcelona, Spain) using an Agilent capture kit and Illumina sequencing (Supplemental Methodology). The bioinformatics analysis is detailed in Supplemental Methodology. Genetic variants were filtered according to the following premises: 1) non-synonymous change; 2) not previously described or with a Minor Allele Frequency < 0.005 in NCBI dbSNP Human Build 135 (<http://www.ncbi.nlm.nih.gov/>), 1000 genomes project and ExAC database; 3) not present in NHLBI Go Exome Sequencing Project (ESP) (<http://evs.gs.washington.edu/EVS/>); and 4) not present in in-house exomes of individuals drawn from the general population ( $n=8$ ). Because of the small number of in-house exomes, variants were later searched for in the CSVS (Collaborative Spanish Variant Server), which at present includes data from 1644 Spanish individuals, most of them sequenced in the same facilities. SIFT,<sup>(7)</sup> PolyPhen,<sup>(8)</sup> Mutation Taster,<sup>(9)</sup> and conservation scores obtained from PhastCons<sup>(10)</sup> were used for prioritization sorting.

### Genetic variant validation

Filtered mutations were validated by polymerase chain reaction (PCR) and automatic Sanger sequencing. Sequencing was performed bidirectionally using BigDye v3.1 Terminator Cycle Sequencing Kit (Applied Biosystems, Carlsbad, CA, USA), according to the manufacturer's instructions. Relevant validated mutations were in silico analyzed (Supplemental Methodology) and screened in 3 samples from women without atypical fracture and long-term bisphosphonate use by Sanger sequencing.

### GGPPS enzyme activity and conformation

The cDNA for both wild-type and Asp188Tyr GGPPS were cloned into an inducible bacterial expression vector, and the resulting His-tagged proteins were expressed overnight in transformed *E. coli* BL-21 (DE3) cells. Protein extracts were obtained and correct expression was verified by Western blot using an anti-GGPPS antibody (sc-271680 Santa Cruz Biotechnology, Dallas, TX, USA). GGPPS was purified from extracts using Ni sepharose followed by gel filtration chromatography. Analysis of the oligomerization of the GGPPS monomers was undertaken using a Sephadex S300 gel filtration column. Enzyme activity was assayed using substrates, Farnesyl pyrophosphate, and C14-isopentenyl pyrophosphate (400 KBq/ $\mu$ Mol) at 20  $\mu$ M in buffer containing 100 mM HEPES pH7.5, 2 mM MgCl<sub>2</sub>, 0.1% Tween 20. Reactions were stopped after 10 minutes at 37 °C by the addition

of acidified methanol. Reaction products were extracted directly into water-immiscible scintillation fluid and quantified by scintillation counting.

#### Cell culture and transduction

MC3T3-E1 osteoblast/calvarial and RAW264.7 macrophage cell lines were cultured in complete  $\alpha$ -MEM (10% fetal bovine serum, 100 U/mL penicillin, and 100 U/mL streptomycin) and maintained in humidified conditions with 5% CO<sub>2</sub> at 37°C. To generate osteoblasts and macrophages depleted of *GGPS1* expression, MC3T3-E1 osteoblast/calvarial and RAW 264.7 macrophage cell lines were transduced with either five different *GGPS1* MISSION shRNAs or non-target shRNA control lentiviral transduction particles (MilliporeSigma, St. Louis, MO, USA). Stable cell lines for MC3T3-E1 were established through puromycin selection at 2  $\mu$ g/mL. For RAW 264.7 macrophages, successfully transduced cells were selected using puromycin at 6  $\mu$ g/mL for 7 to 8 days. Once single cell-derived colonies were observed, three individual cell colonies per shRNA were harvested using cell cloning cylinders and further expanded to form stable cell lines.

#### Mineralization assay and analysis

To assay for mineralization activity, stably transduced MC3T3-E1 cells were plated in 24-well plates and cultured in osteogenic media (complete  $\alpha$ -MEM with 2  $\mu$ g/mL of puromycin, 50  $\mu$ g/mL L-ascorbic acid, and 10 mM  $\beta$ -glycerophosphate). Osteogenic media was replaced every 3 days, and cells were fixed with 4% paraformaldehyde (PFA) after 21 days. Bone nodules were stained using Alizarin Red solution, and bone nodule area (mm<sup>2</sup>) were quantified using the Fiji software.

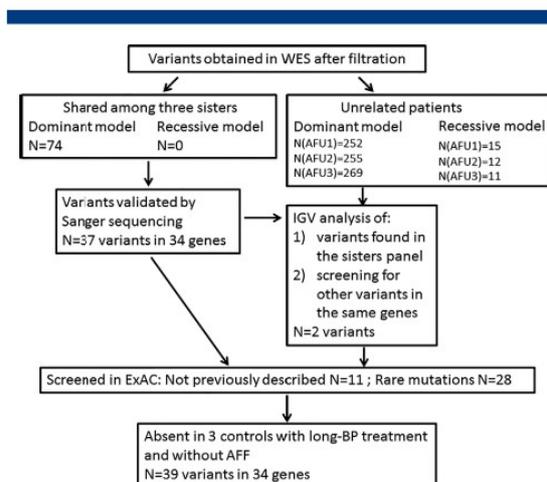
#### Osteoclast culture and resorption assay

Stably transduced RAW 264.7 macrophages were differentiated into osteoclasts in differentiation media (complete  $\alpha$ -MEM with 6  $\mu$ g/mL of puromycin, supplemented with 10 ng/mL recombinant mouse RANKL [R&D Systems, Minneapolis, MN, USA]). Media was changed every 48 hours, and after 4 days, cells were fixed with 4% PFA. Cells positive for TRAP activity and containing three or more nuclei were scored as mature osteoclasts. To assess resorptive activity, macrophages were plated onto 24-well Osteo Assay Surface plates (Corning, Lowell, MA, USA) and cultured in differentiation media for 7 days. Media was aspirated from the wells at the end of day 7, and cells were gently removed using a 10% bleach solution. The wells were washed with distilled water and dried well before a Von Kossa stain was performed to contrast between resorption pits formed and the surface coating. Six random fields per well were imaged using light microscopy, and the percentage of resorbed area was analyzed using the Fiji software.

## Results

#### Variants detected in WES

The 3 sisters (AFS1, AFS2, AFS3) and the 3 unrelated patients (AFU1, AFU2, AFU3) were distributed into two groups and analyzed separately. The workflow and number of identified variants are shown in Fig. 1. In a first step, only mutations shared by the 3 sisters were taken into account both in a dominant and a recessive model. No variants were identified in homozygosity, whereas 74 variants were identified in



**Fig. 1.** Flow chart of approach for detecting AFF-associated mutations. N(AFU1) = number of variants detected in the patient AFU1; N(AFU2) = number of variants in the patient AFU2; N(AFU3) = number of variants in the patient AFU3.

heterozygosity (consistent with a dominant model), 37 of which were validated by Sanger sequencing. In three genes (*FN1*, *BRAT1*, and *XAB2*), the sisters were found to carry two different mutations. Direct visualization of sequence reads with the IGV software as well as polymorphism analyses indicated that the variants were in phase in all cases, being double-mutant alleles rather than compound heterozygotes. The 37 coding variants shared by the 3 sisters, all missense except for one nonsense and one in-frame deletion, are listed in Supplemental Table S2, according to their conservation score. The first variant in the list, with the best conservation score, was in the *GGPS1* gene, as we previously described.<sup>(6)</sup>

In a second step, the genes with variants shared by the sisters were screened in the WES results of the unrelated patients using the IGV software. None of the variants was found in any of the unrelated patients. However, in *BRAT1* and *CYP1A1*, two other variants were found in AFU3 and AFU1, respectively (Supplemental Table S3). The *CYP1A1* variant present in AFU1 (p.Ser216Cys) is a change of a serine to a sulphur-containing amino acid next to the substrate binding site and is predicted to be very deleterious to its function. Likewise, the *CYP1A1* variant present in the sisters (p.Arg98Trp) is a very significant change of a basic (arginine) to an aromatic hydrophobic amino acid (tryptophan), lying in a hydrogen-bonded turn of the protein. Conversely, the three variants detected in the *BRAT1* gene (two of them in the 3 sisters, in a double-mutant allele, and one in patient AFU3) were predicted as unlikely to affect its function. None of the variants in Supplemental Tables S2 and S3 was found in 3 controls (long-term treated with BPs but without AFFs). A total of 11 mutations are not present either in the NCBI dbSNP or in ExAC. The other variants, without MAF in dbSNP, have allele frequencies  $\leq 2/10,000$  according to ExAC. Only 10 variants are present in the CSVS database, all but one (in *FN1*) with allele frequencies  $< 5/1000$ , in the Iberian population (Supplemental Table S2).

## Functional analyses of the GGPPS mutation

Asp188 is an active site residue of GGPP synthase, involved in the binding of the substrate via a magnesium salt bridge. Disruption of this residue is expected to lead to a vastly reduced rate of activity. To confirm this prediction, we produced mutant and wild-type recombinant GGPPS enzymes (Fig. 2A) and assayed their activity *in vitro*. As shown in Fig. 2B, the mutant displayed 5.7% of wild-type activity, with values of  $0.72 \pm 0.09$  cpm/ng/min for the wild-type and  $0.04 \pm 0.013$  cpm/ng/min for the mutant ( $n = 3$ ). Gel filtration experiments using a calibrated S300 column showed the wild-type enzyme as having a molecular weight in excess of 220 kDa, indicative of the expected hexameric conformation, in line with previous findings.<sup>(11)</sup> The mutant enzyme consistently showed two peaks corresponding to the hexamer and to the monomer (peak at approximately 38 kDa), suggesting the mutation has a destabilizing effect on the oligomerization of the enzyme (Fig. 2C).

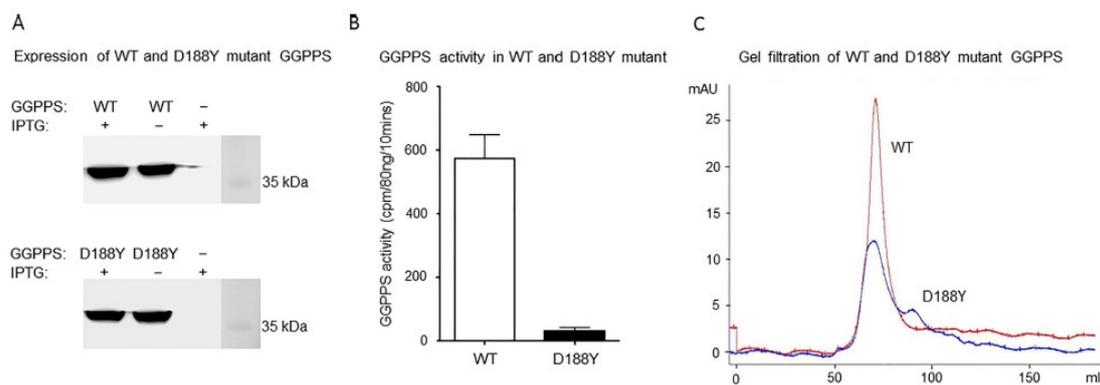
Next, we studied the effect of GGPPS depletion *in vitro* by utilizing shRNA-mediated knockdown of *GGPS1* in MC3T3-E1 and RAW 264.7 cells. To this end, five independent shRNAs against *GGPS1* (denoted #1 to #5) and a control non-targeting shRNA were initially screened for their efficacy in depleting *GGPS1* mRNA expression in MC3T3-E1 cells. mRNA expression levels were examined using RT-qPCR. Of the five shRNAs, only shRNAs #1 and #2 exhibited promising potential knockdown effects at the mRNA level in MC3T3-E1 cells (>50%) (Fig. 3A). However, when subjected to immunoblot analysis, only shRNA #1 showed a strong reduction of *GGPS1* at the protein level (Fig. 3B). As such, only shRNA #1 was used in further experiments.

Control and *GGPS1*-depleted MC3T3-E1 cells were cultured under mineralizing conditions and stained with alizarin red (Fig. 3C). Bone nodule formation *in vitro* was dramatically reduced after *GGPS1* inhibition (Fig. 3D). To assess whether the impaired mineralization activity of *GGPS1*-depleted MC3T3-E1 cells were a result of impaired osteoblast differentiation, we

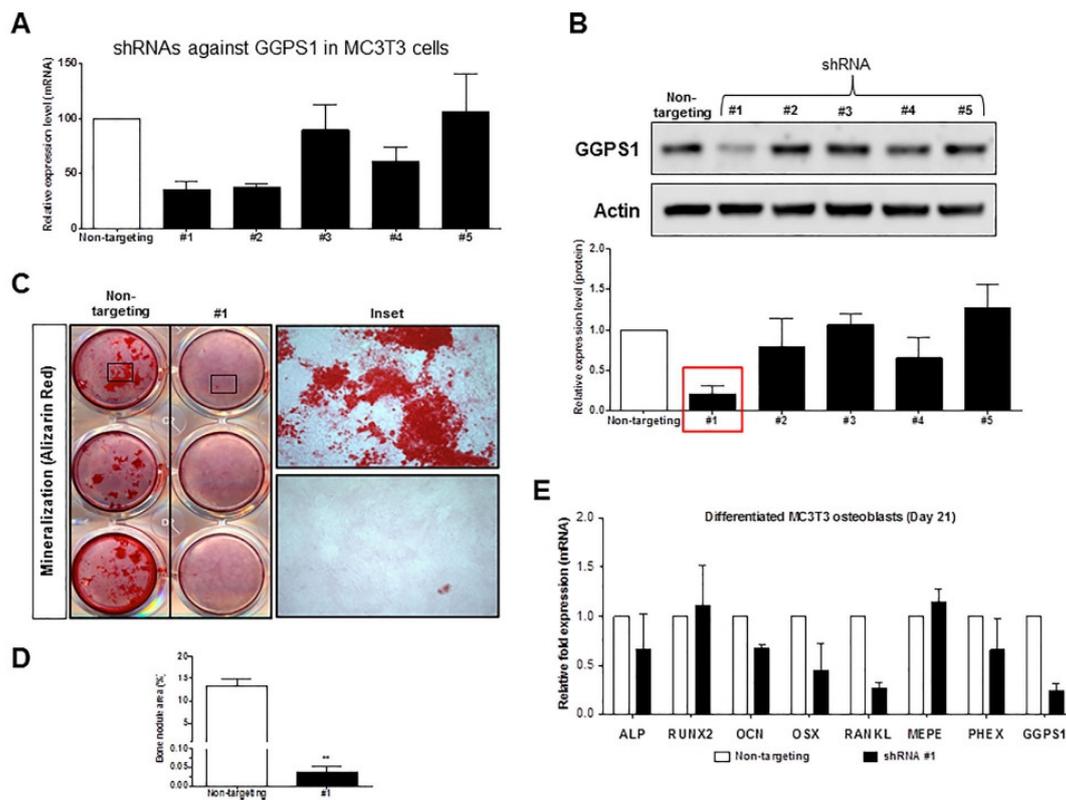
further analyzed the mRNA expression of key osteoblast markers using RT-qPCR. Interestingly, there were clear reductions in *RANKL*, *OSX*, and *OCN* mRNA expression in *GGPS1*-depleted cells (Fig. 3E), while no significant effects were observed for *RUNX2*, *ALPL*, *MEPE*, and *PHEX*.

Similarly, RAW 264.7 mouse macrophages were transduced with the same five shRNAs against *GGPS1* and a non-targeting shRNA control. Initial screening of the resultant five heterogeneous polyclonal pool of stable RAW 264.7 cells using RT-qPCR indicated that *GGPS1* knockdown efficiency was lower than expected (data not shown). As such, 2 to 3 monoclonal stables for each *GGPS1*-shRNA were generated (denoted shRNA #1A-C, #2A-B, #3A-C, #4A-C, and #5A-B). Using RT-qPCR, we again screened for the efficiency of *GGPS1* knockdown and found that monoclonal stable cell lines generated from shRNAs #1C, #2B, and #4B yielded the most potent effects, achieving consistent knockdown of *GGPS1* at the mRNA level (>65%) (Fig. 4A). At the protein level, however, only macrophages generated from *GGPS1* shRNA #4A exhibited a significantly decreased protein expression of *GGPS1* (Fig. 4B) and was therefore selected for further analyses.

To assess whether *GGPS1* was functionally required during osteoclast formation, control and *GGPS1* knockdown cells were plated in 24-well plates in triplicates and treated with RANKL every 48 hours over a course of 4 days. Cells were fixed with 4% PFA, stained for TRAP, and imaged using light microscopy (Fig. 4C). When quantitated, we found that loss of *GGPS1* expression increased osteoclast formation significantly (Fig. 4D). Lastly, to examine if *GGPS1* was necessary for resorptive activity, control and *GGPS1* knockdown cells were cultured on Osteo Assay Surface plates for a course of 7 days, supplemented with RANKL every 48 hours. Both TRAP activity and F-actin ring formation in *GGPS1* knockdown osteoclasts appeared indistinguishable from the control, and *GGPS1* knockdown osteoclasts also appeared to retain some resorptive abilities (Fig. 4E). However, when the resorptive pits were quantitated, we found that *GGPS1*-depleted osteoclasts had



**Fig. 2.** (A) Heterologous expression of WT and p.Asp188Tyr GGPPS, assessed by Western blot of 2.5  $\mu$ g of transformed non-induced or IPTG-induced *E. coli* extracts. (B) GGPPS enzyme activity in WT and p.Asp188Tyr mutant. p.Asp188Tyr GGPPS had 5.7% of the WT activity, measured by scintillation counting of [<sup>14</sup>C]Geranylgeranyl pyrophosphate. Results are expressed as mean  $\pm$  SD ( $n = 3$ ). \*\*\* $p < 0.001$ . (C) Gel filtration chromatograms for the WT GGPPS and the p.Asp188Tyr mutant. The WT enzyme appears to have a molecular weight of around 220 kDa, suggesting that it is present as a hexamer. The p.Asp188Tyr mutant enzyme consistently showed two peaks corresponding to the hexamer and the monomer (a peak around 38 kDa), suggesting that the mutant destabilizes the oligomerization of the enzyme.



**Fig. 3.** Effects of shRNA-mediated *GGPS1* depletion in MC3T3 cells. (A) qPCR analysis of shRNA-mediated knockdown of *GGPS1* expression. Data (mean  $\pm$  SD) are shown as percent of mRNA levels in the negative control sample. (B) *GGPS1* protein expression levels measured by Western blot. Data (means  $\pm$  SD) are shown as relative protein levels with respect to the non-targeting sample. (C) Alizarin Red staining of MC3T3 cells treated with the control shRNA or *GGPS1*-shRNA. (D) Quantification (%) of the area of mineralized nodule generated by MC3T3 cells treated with non-targeting shRNA or *GGPS1*-targeting shRNA. Results are expressed as mean  $\pm$  SD. \*\* $p < 0.01$ . (E) mRNA fold expression of osteoblast markers measured by RT-qPCR of control and *GGPS1*-shRNA-treated differentiated MC3T3 osteoblasts. Data (means  $\pm$  SD) are shown as relative mRNA fold expression with respect to the non-targeting sample.

decreased resorption area, although it did not reach significance (Fig. 4F).

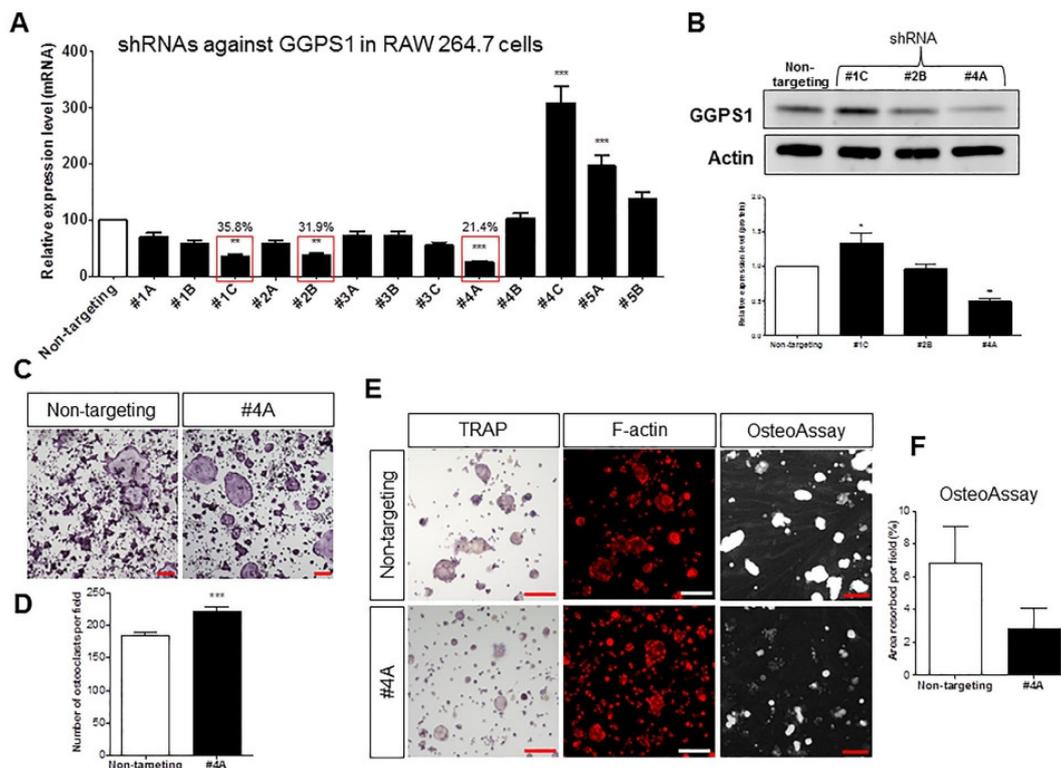
## Discussion

In the present study, we describe the list of rare variants identified by WES in 3 sisters affected with AFF. Because causality cannot be attributed to rare variants that segregate within a small family just because they are rare,<sup>(12,13)</sup> we have carefully analyzed the function of the most interesting variant, the p.Asp188Tyr mutation in *GGPS1*, which we recently reported elsewhere.<sup>(6)</sup> The results presented here provide functional evidence of pathogenicity of this *GGPS1* mutation and its role in regulating bone cells and their activities.

The *GGPS1* gene encodes the GGPPS enzyme involved in the mevalonate pathway (Supplemental Fig. S1), and along with farnesyl pyrophosphate synthase (FPPS), is known to be inhibited by a variety of N-BPs.<sup>(14)</sup> The primary function of the mevalonate pathway is the production of cholesterol, as well as

the synthesis of isoprenoid lipids, including farnesyl diphosphate (FPP) and geranylgeranyl diphosphate (GGPP),<sup>(15)</sup> which are required for the posttranslational modification (prenylation) of some proteins. The geranylgeranyl diphosphate synthase enzyme (GGPPS) catalyzes the synthesis of geranylgeranyl diphosphate (GGPP) from farnesyl diphosphate and isopentenyl diphosphate. GGPPS functions as a homohexamer, in which each monomer binds 3  $Mg^{2+}$  ions.<sup>(11)</sup>

We clearly show that the p.Asp188Tyr (D188Y) mutation severely impairs in vitro enzyme activity, consistent with the fact that it lies in the second aspartate-rich region, highly conserved across all GGPPS and FPPS, and involved in the binding of the substrates to the enzyme-active site via a  $Mg^{2+}$  salt bridge, which is essential for catalytic activity. It is well known that any disruption in this region results in an almost complete loss of activity.<sup>(16)</sup> We also show, by gel filtration experiments, that the p.D188Y mutation destabilizes the homohexameric conformation of the enzyme elucidated by Kavanagh and colleagues.<sup>(11)</sup> Taking all the data together, and according to the American College of Medical Genetics and Genomics criteria,<sup>(17)</sup> this



**Fig. 4.** Effects of shRNA-mediated *GGPS1* depletion in RAW264.7 macrophages after RANKL stimulation. (A) qPCR analysis of shRNA-mediated knockdown of *GGPS1* expression. Data (mean ± SD) are shown as percent of mRNA levels in the negative control sample. \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ . (B) *GGPS1* protein expression levels in the selected samples measured by Western blot. Data (means ± SD) are shown as relative protein levels with respect to the non-targeting sample. \* $p < 0.05$ ; \*\* $p < 0.01$ . (C) Tartrate-resistant acid phosphatase (TRAP) staining of cells treated with the control shRNA or *GGPS1*-shRNA. (D) Quantification of osteoclasts derived from RAW264.7 cells treated with non-targeting shRNA or *GGPS1*-targeting shRNA. Results are expressed as mean ± SD. \*\*\* $p < 0.001$ . (E) TRAP staining, F-actin staining, and resorptive pits generated after culturing cells on OsteoAssay surface plates of control and *GGPS1*-shRNA-treated cells. (F) Quantification (%) of the area of pits resorbed by control and *GGPS1*-shRNA-treated osteoclasts. Data are means ± SD.

mutation is classified as pathogenic, even though it has not been reported in any additional AFF patient so far.

To examine the functional role of GGPPS in bone cells, we performed in vitro studies in *GGPS1*-depleted osteoblasts or osteoclasts. GGPPS-depleted MC3T3 cells had reduced mineralization capacity and reduced gene expression of osteocalcin, osteonectin, and RANKL. These results are in agreement with those of Weivoda and Hohl,<sup>(18)</sup> in which GGPPS was inhibited with digeranyl bisphosphonate (DGBP) in MC3T3 cells. These authors suggested that the observed lack of mineralization and the decrease in *ALPL* and *OCN* gene expression were due to the accumulation of FPP and its subsequent activation of the glucocorticoid receptor,<sup>(19)</sup> which is known to inhibit osteoblast proliferation and bone formation and to increase osteoblast apoptosis.<sup>(20)</sup> In addition, the depletion of RANKL would lead to an aberrant osteoblast-osteoclast cross-talk. The effect of GGPPS depletion in osteoclasts was an increase in cell number and a slightly decreased activity. Disruption of *GGPS1* as depicted in our shRNA assay is predicted to reduce the synthesis of GGPP. Depletion of GGPP impairs the prenylation of GTPases such as

Rho, Rac, Rap1, and Rabs, which have been shown to play essential roles in both osteoclast formation and function.<sup>(21–25)</sup> Mouse models utilizing osteoclast targeted depletion or global depletion of these key GTPases have shown conflicting trends in the resulting osteoclast numbers, which is not well understood, and might stem from the different ages at which the different laboratory groups analyzed their mice specimen.<sup>(23,25)</sup> Interestingly, however, all of the mice models exhibited osteopetrotic phenotypes, indicating that osteoclast resorptive activity is highly dependent on geranylgeranylation.<sup>(23–26)</sup> Unlike these previous studies, our work in disrupting *GGPS1* does not specifically target any of the GTPases mentioned. Loss of prenylation and membrane localization of these GTPases after GGPPS depletion does not necessarily translate to inhibited GTPase function. In fact, it has been shown that unprenylated GTPases can remain in the GTP-bound form, accumulate in the cytosol, and retain partial functional activity such as inducing activation of the p38 MAPK,<sup>(27)</sup> which is an important signaling pathway for osteoclast differentiation and formation.<sup>(28)</sup> Therefore, despite the increased osteoclast numbers in our *GGPS1*

shRNA population, there is a slight decrease in bone-resorptive activity, consistent with previous studies showing that geranylgeranylation plays a pivotal role during osteoclast bone resorption due to alterations in vesicular traffic, a cellular function possibly less essential during osteoclast differentiation.

Of note, and unlike BPs, which preferentially target osteoclasts, the GGPPS mutation in the 3 sisters will affect all of their cells, including osteoblasts. Because the administration of bisphosphonate targets mainly FPP synthase, which is upstream of GGPPS, we speculate that the effect of bisphosphonates on the cell lines will be compounded because of the loss of both farnesylation and geranylgeranylation. However, although the relevant cell lines may reveal some answers, they may not fully replicate what happens in clinical cases, where in vivo osteoblast and osteoclast responses are intimately associated due to their coupling in bone remodeling. Furthermore, it appears that the onset of atypical femoral fractures usually occurs after prolonged bisphosphonate treatment, which is difficult to mimic in an in vitro environment. Developing an animal model strategy should provide more compelling evidence. The GGPS1-BP interaction is also supported by the finding of a common variant in the GGPS1 promoter, which was associated with lack of bone mineral density (BMD) improvement after BP therapy,<sup>(29)</sup> possibly indicating that the pathway was already impaired in these patients.

Another interesting potentially causative gene in our list is *CYP11A1*, which was found mutated in the 3 sisters, in 1 of our unrelated AFF patients, and also in another AFF patient reported elsewhere.<sup>(30)</sup> According to the American College of Medical Genetics and Genomics criteria,<sup>(17)</sup> these *CYP11A1* mutations may be classified as likely pathogenic. Functional studies needed to confirm their pathogenicity are underway. *CYP11A1* encodes a member of the cytochrome P450 superfamily, involved in the metabolism of drugs and xenobiotics and arises as a good AFF-susceptibility candidate because it is responsible for the hydroxylation of 17 $\beta$ -estradiol, estrone, and vitamin D in extrahepatic tissues.<sup>(31)</sup> Its role in bone biology is also supported by the association found between the *CYP11A1* C4887A polymorphism and a higher degree of estrogen catabolism and lower femoral BMD in postmenopausal women.<sup>(32)</sup>

The strengths of our study were the possibility to analyze 3 sisters with AFF and the choice of a hypothesis-free WES approach that allowed us to detect new variants not included in exon arrays, as previously performed.<sup>(33)</sup> On the other hand, the small number of AFF patients and controls studied here is an important limitation, and further WES of additional AFF cases are underway. Moreover, we could only analyze 3 sisters, who have an a priori chance of 1/8 of sharing any variant, which is above a conventional level of statistical significance. Another potential limitation of the study is the impact that hypercholesterolemia and statin treatment might have had in bone metabolism in the 3 sisters.

In summary, our results show the negative impact of the GGPPS p.Asp188Tyr mutation and the relevance of the downstream effects in bone cells, which makes it a candidate for AFF susceptibility. In addition, our data show other potential AFF contributory genes, although functional studies are needed to prove their involvement in the pathology. Further identification and/or replication of genetic variants will be necessary to detect at-risk individuals and to decide which patients are suitable for being treated with BPs with no risk of this side effect.

## Disclosures

All authors state that they have no conflicts of interest.

## Acknowledgments

We thank the patients for their kind participation. Funds for the study include grants SAF2014-56562R, SAF2016-75948-R (Spanish MINECO), PI12/02315 (FIS, ISCIII), 2014SGR932 (Catalan Government), and CIBERER (U720). This work was also supported by the Centro de Investigación Biomédica en Red en Fragilidad y Envejecimiento Saludable (CIBERFES; CB16/10/00245) and FEDER funds. JED was supported by the NIHR Biomedical Research Centre, Oxford, UK. NR is recipient of an FPU predoctoral fellowship from the Spanish Ministerio de Educación Cultura y Deporte. The work was also supported by a grant from the US government, NIH, NIAMS (R01 AR062054) to RB.

Authors' roles: Study design: AD-P, NG-G, SB, and DG. Study conduct: NR-A, PYN, NG-G, MF-M, MC, JFA, and JED. Patient and data collection: JMGG, XN, and DP-A. Data analysis: NR-A, PYN, NG-G, MF-M, JFA, GGR, RB, DG, SB, and AD-P. Data interpretation: all authors. Drafting manuscript: NR-A, PYN, NG-G, JED, GGR, RB, DG, SB, and AD-P. Revising manuscript content: all authors. Approving final version of manuscript: all authors. NR-A and PYN take responsibility for the integrity of the data analysis.

## References

1. Freemantle N, Cooper C, Díez-Pérez A, et al. Results of indirect and mixed treatment comparison of fracture efficacy for osteoporosis treatments: a meta-analysis. *Osteoporos Int.* 2013;24(1):209–17.
2. Zhou J, Ma X, Wang T, Zhai S. Comparative efficacy of bisphosphonates in short-term fracture prevention for primary osteoporosis: a systematic review with network meta-analyses. *Osteoporos Int.* 2016;27(11):3289–300.
3. Peng J, Liu Y, Chen L, et al. Bisphosphonates can prevent recurrent hip fracture and reduce the mortality in osteoporotic patient with hip fracture: a meta-analysis. *Pak J Med Sci.* 2016;32(2):499–504.
4. Kennel KA, Drake MT. Adverse effects of bisphosphonates: implications for osteoporosis management. *Mayo Clin Proc.* 2009;84(7):632–7; quiz 8.
5. Shane E, Burr D, Abrahamsen B, et al. Atypical subtrochanteric and diaphyseal femoral fractures: second report of a task force of the American Society for Bone and Mineral Research. *J Bone Miner Res.* 2014;29(1):1–23.
6. Roca-Ayats N, Balcells S, Garcia-Giral N, et al. GGPS1 mutation and atypical femoral fractures with bisphosphonates. *N Engl J Med.* 2017;376(18):1794–5.
7. 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(7):1073–81.
8. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7(4):248–9.
9. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods.* 2014;11(4):361–2.
10. Siepel A, Bejerano G, Pedersen JS, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* 2005;15(8):1034–50.
11. Kavanagh KL, Dunford JE, Bunkoczi G, Russell RG, Oppermann U. The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a novel hexameric arrangement and inhibitory product binding. *J Biol Chem.* 2006;281(31):22004–12.

12. MacArthur DG, Manolio TA, Dimmock DP, et al. Guidelines for investigating causality of sequence variants in human disease. *Nature*. 2014;508(7497):469–76.
13. Minikel EV, Vallabh SM, Lek M, et al. Quantifying prion disease penetrance using large population control cohorts. *Sci Transl Med*. 2016;8(322):322ra9.
14. Guo RT, Cao R, Liang PH, et al. Bisphosphonates target multiple sites in both cis- and trans-prenyltransferases. *Proc Natl Acad Sci U S A*. 2007;104(24):10022–7.
15. Goldstein JL, Brown MS. Regulation of the mevalonate pathway. *Nature*. 1990;343(6257):425–30.
16. Kavanagh KL, Guo K, Dunford JE, et al. The molecular mechanism of nitrogen-containing bisphosphonates as antiosteoporosis drugs. *Proc Natl Acad Sci U S A*. 2006;103(20):7829–34.
17. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405–24.
18. Weivoda MM, Hohl RJ. The effects of direct inhibition of geranylgeranyl pyrophosphate synthase on osteoblast differentiation. *J Cell Biochem*. 2011;112(6):1506–13.
19. Das S, Schapira M, Tomic-Canic M, Goyanka R, Cardozo T, Samuels HH. Farnesyl pyrophosphate is a novel transcriptional activator for a subset of nuclear hormone receptors. *Mol Endocrinol*. 2007;21(11):2672–86.
20. Weinstein RS, Jilka RL, Parfitt AM, Manolagas SC. Inhibition of osteoblastogenesis and promotion of apoptosis of osteoblasts and osteocytes by glucocorticoids. Potential mechanisms of their deleterious effects on bone. *J Clin Invest*. 1998;102(2):274–82.
21. Zou W, Izawa T, Zhu T, et al. Talin1 and Rap1 are critical for osteoclast function. *Mol Cell Biol*. 2013;33(4):830–44.
22. Pavlos NJ, Xu J, Riedel D, et al. Rab3D regulates a novel vesicular trafficking pathway that is required for osteoclastic bone resorption. *Mol Cell Biol*. 2005;25(12):5253–69.
23. Croke M, Ross FP, Korhonen M, Williams DA, Zou W, Teitelbaum SL. Rac deletion in osteoclasts causes severe osteopetrosis. *J Cell Sci*. 2011;124(Pt 22):3811–21.
24. Ito Y, Teitelbaum SL, Zou W, et al. Cdc42 regulates bone modeling and remodeling in mice by modulating RANKL/M-CSF signaling and osteoclast polarization. *J Clin Invest*. 2010;120(6):1981–93.
25. Itokawa T, Zhu ML, Troiano N, Bian J, Kawano T, Insogna K. Osteoclasts lacking Rac2 have defective chemotaxis and resorptive activity. *Calcif Tissue Int*. 2011;88(1):75–86.
26. Magalhaes JK, Grynaps MD, Willett TL, Glogauer M. Deleting Rac1 improves vertebral bone quality and resistance to fracture in a murine ovariectomy model. *Osteoporos Int*. 2011;22(5):1481–92.
27. Dunford JE, Rogers MJ, Ebetino FH, Phipps RJ, Coxon FP. Inhibition of protein prenylation by bisphosphonates causes sustained activation of Rac, Cdc42, and Rho GTPases. *J Bone Miner Res*. 2006;21(5):684–94.
28. Li X, Udagawa N, Itoh K, et al. p38 MAPK-mediated signals are required for inducing osteoclast differentiation but not for osteoclast function. *Endocrinology*. 2002;143(8):3105–13.
29. Choi HJ, Choi JY, Cho SW, et al. Genetic polymorphism of geranylgeranyl diphosphate synthase (GGSP1) predicts bone density response to bisphosphonate therapy in Korean women. *Yonsei Med J*. 2010;51(2):231–8.
30. Peris P, González-Roca E, Rodríguez-García SC, López-Cobo MM, Monegal A, Guañabens N. Incidence of mutations in the ALPL, GGPS1, and CYP1A1 genes in patients with atypical femoral fractures. *JBM Plus*. Epub 2018 May 30. DOI: 10.1002/10064.
31. Zhou SF, Liu JP, Chowbay B. Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metab Rev*. 2009;41(2):89–295.
32. Napoli N, Villareal DT, Mumm S, et al. Effect of CYP1A1 gene polymorphisms on estrogen metabolism and bone density. *J Bone Miner Res*. 2005;20(2):232–9.
33. Perez-Nunez I, Perez-Castrillon JL, Zarrabeitia MT, et al. Exon array analysis reveals genetic heterogeneity in atypical femoral fractures. A pilot study. *Mol Cell Biochem*. 2015;409(1–2):45–50.

## **Supplementary Information:**

### **Supplementary methodology**

#### Whole-Exome Sequencing (WES)

Library preparation for capture of selected DNA regions (SureSelect XT Human All Exon; cat: 5190-6208; Agilent Technologies) was performed according to Agilent's SureSelect protocol for Illumina paired-end sequencing. In brief, 3.0µg of genomic DNA was sheared on a Covaris™ E220 instrument. The fragment size (150-300 bp) and quantity were confirmed with the Agilent 2100 Bioanalyzer 7500 chip. The fragmented DNA was end-repaired, adenylated, and ligated to Agilent indexing-specific paired-end adaptors. The DNA with adaptor-modified ends was PCR-amplified (6 cycles, Herculase II fusion DNA polymerase from Agilent) with SureSelect Primer and SureSelect Pre-capture Reverse PCR primers (SureSelect XT Human All Exon), quality controlled on the DNA 7500 assay for the library size range of 250 to 450 bp, and hybridized for 24h at 65°C (Applied Biosystems 2720 Thermal Cycler). The hybridization mix was washed in the presence of magnetic beads (Dynabeads MyOne Streptavidin T1, Life Technologies) and the eluate was PCR-amplified (16 cycles) in order to add the index tags using SureSelectXT Indexes for Illumina. The final library size and concentration was determined on Agilent 2100 Bioanalyzer 7500 chip and sequenced on an Illumina HiSeq 2000 platform at a coverage of 40x with paired end runs of 2x76bp following the manufacturer's protocol. Images from the instrument were processed using the manufacturer's software to generate FASTQ sequence files.

#### WES data analysis

The Illumina RTA sequence analysis pipeline was used for base calling and quality control.

The data of the sequenced fragments, in FASTQ format, were aligned with the Burrows-Wheeler Aligner (1) free software (<http://bio-bwa.sourceforge.net/>) using the GRCh37 (hg19) build of the reference human genome. Mapped reads were filtered (leaving only those mapping in unique genomic positions with enough quality), sorted, and indexed with SAMtools (2). Mean mapping qualities were: 69.57 for AFS1, 69.96 for AFS2, 69.40 for AFS3, 69.82 for AFU1, 69.21 for AFU2 and 67.28 for AFU3. GATK (3) was then used to realign the reads as well as for the base quality score recalibration. Once a satisfactory alignment was achieved, single nucleotide variants and

indels were identified using GATK standard hard filtering parameters (4): Quality by Depth (QD) > 2.0, Fisher Strand (FS) < 60.0, Strand Odds Ratio (SOR) < 3.0, Root Mean Square Mapping Quality (MQ) > 40.0, Mapping Quality Rank Sum Test (MQRankSum) > -12.5 and Read Position Rank Sum Test (ReadPosRankSum) > -8.0.

For the final report of the exome-sequencing analysis, we used the VARIANT (5) annotation tool, which provides additional information on relevant variants for the final process of candidate gene selection. In particular, minor allele frequency (MAF) was obtained from dbSNP (6) and the 1000 Genomes project (<http://www.1000genomes.org>) (7) to help with the selection of new variants not reported in healthy populations to date. Finally, processed data were converted to BAM (binary equivalent SAM) format for variant detection and analysis using the Integrative Genomics Viewer (IGV) (<http://www.broadinstitute.org/igv>).

#### References

1. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009 Jul 15;25(14):1754-60.
2. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al*. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009 Aug 15;25(16):2078-9.
3. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, *et al*. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010 Sep;20(9):1297-303.
4. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, *et al*. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011 May;43(5):491-8.
5. Medina I, De Maria A, Bleda M, Salavert F, Alonso R, Gonzalez CY, *et al*. VARIANT: Command Line, Web service and Web interface for fast and accurate functional characterization of variants found by Next-Generation Sequencing. *Nucleic Acids Res*. 2012 Jul;40(Web Server issue):W54-8.
6. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, *et al*. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001 Jan 1;29(1):308-11.
7. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, *et al*. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012 Nov 1;491(7422):56-65.

*In silico* analysis

Mutations were located within the gene context using the UCSC Genome Browser (<https://genome.ucsc.edu/>) and the Ensemble Genome Browser (<http://www.ensembl.org/>). Gene information was extracted from GeneCards (<http://www.genecards.org/>) and BioGPS (<http://biogps.org/>). Information from other WES projects was extracted from the Exome Aggregation Consortium (ExAC) (<http://exac.broadinstitute.org>).

The *in silico* functional study of mutated proteins was performed using The Universal Protein Resource (UniProt) (<http://www.uniprot.org/>), RCSB Protein Data Bank (PDB) (<http://www.rcsb.org/pdb/>) and Pfam (<http://pfam.xfam.org/>). Protein alignments were performed using the UCSC Genome Browser or by on-line aligning of amino acidic sequences in FASTA format using Clustal Omega (<http://www.clustal.org/omega/>) and on-line ESPript (<http://esprict.ibcp.fr/>).

**Supplemental Table S1:** Patient characteristics

Patient	Atypical fracture	Age (years)	Weight (Kg)	T-score Lumbar spine	T-score total hip	Time on BP treatment (years)	Previous OP fractures
AFS1	unilateral	64	77	-1.1	-0.2	6	Colles
AFS2	unilateral	73	75	-2.5	-1.4	6	Colles
AFS3	bilateral	60/61	100	-0,3	bhpr	6	none
AFU1	bilateral	73/75	50.8	-1.9	-0.5	6	none
AFU2	unilateral	72	90	-2.0	-0.6	7	none
AFU3	unilateral	87	59.8	N/A	N/A	10	none
control 1		78	66.5	-2.5	-1.9	7	none
control 2		70	57.5	-1.2	-2.4	6	none
control 3		74	77.1	-1.5	-0.9	8	none

AFS = Atypical fracture sister; AFU = Atypical fracture unrelated; Age = Age at the time of fracture in AFF patients; bhpr = bilateral hip prosthesis replacement

**Supplemental Table S2.** Variants shared by the three sisters, found by exome sequencing

Gene	Protein	Variant <sup>1</sup>	Effect on the protein	dbSNP <sup>2</sup>	ExAC <sup>3</sup>	CSVs <sup>4</sup>	Mutation Taster <sup>5</sup>	Conservation <sup>6</sup>	Sift <sup>7</sup>	PolyPhen <sup>8</sup>
<i>GGPS1</i>	Geranylgeranyl diphosphate synthase	chr11.q.235505746G>T	p.D188Y				DC; 0.9999	700	0.000	1.000
<i>LRR1</i>	Leucine-rich repeat-containing 1	chr6.g.53707020G>A	p.R91Q		4.946e-05	0.0003	DC; 0.9999	685	0.050	0.746
<i>TUSC2</i>	Tumor suppressor candidate 2	chr3.g.50363807T>C	p.H83R		8.244e-06		DC; 0.8891	674	0.338	0.000
<i>SYDE2</i>	Synapse defective 1, Rho GTPase, homolog 2	chr1.g.85634903G>T	p.L893I		8.339e-06	0.0003	DC; 0.9999	639	0.018	0.997
<i>COG4</i>	Component of oligomeric golgi complex 4	chr16.g.70553552C>T	p.G85D				DC; 0.9999	627	0.150	0.735
<i>EML1</i>	Echinoderm microtubule associated protein like 1	chr14.g.100360993G>A	p.R211H		6.611e-05		DC; 0.9999	588	0.030	0.963
<i>KDMM4C</i>	Lysine(K)-specific demethylase 4C	chr9.g.6849579A>G	p.I170V	rs192832191 MAF=0.0004	2.471e-05		DC; 0.9999	584	0.000	0.509
<i>ERCC6L2</i>	Excision repair cross-complementation group 6 like 2	chr9.g.98718284A>T	p.I657L		8.278e-06		P; 0.9976	573	0.630	0.007
<i>PGRMC1</i>	Progesterone receptor membrane component 1	chrX.g.118377159C>A	p.P177H				DC; 0.9999	573	0.130	0.742
<i>FN1*</i>	Fibronectin 1	chr2.g.216235149C>T	p.V2241I		8.245e-06	0.0003	DC; 0.9997	551	0.009	0.045
<i>CYP11A1</i>	Cytochrome 450, family 1, subfamily A, polypeptide 1	chr15.g.75015147G>A	p.R98W		0.000108	0.0003	DC; 0.5242	540	0.000	0.998
<i>XAB2*</i>	XPA binding protein 2	chr19.g.7688142C>G	p.V385L		1.651e-05		DC; 0.9999	535	0.007	0.600
<i>GPR20</i>	G protein-coupled receptor 20	chr8.g.142367729C>T	p.D99N	rs200892677 MAF=0.0004	3.324e-05		DC; 0.9999	515	0.000	0.998
<i>TMEM25</i>	Transmembrane protein 25	chr11.g.118404174_118404176del	p.V239del				DC; 0.9999	510	N/A	N/A
<i>NGEF</i>	Guanine nucleotide exchange factor	chr2.g.233748153G>A	p.S542L		1.279e-05		DC; 0.9999	500	0.350	0.910
<i>NKAP</i>	NF-κB activating protein	chrX.g.119066123C>T	p.S265N	rs182030723 MAF=0.0006	6.847e-05	0.0003	DC; 0.9999	497	0.120	0.184
<i>NVL</i>	Nuclear-VCP like	chr1.g.224491450G>A	p.T312I		8.268e-06		DC; 0.9999	474	0.000	0.995
<i>FN1*</i>	Fibronectin 1	chr2.g.216251538G>A	p.R1496W	rs139078629 MAF=0.003	0.004904	0.0103	DC; 0.9999	466	0.005	0.998
<i>ATP6AP1</i>	ATPase, H <sup>+</sup> transporting, lysosomal accessory protein 1	chrX.g.153864043G>A	p.V407I		4.561e-05		DC; 0.9868	464	0.260	0.990
<i>LURAP1L</i>	Leucine rich adaptor protein 1-like	chr9.g.12821722G>A	p.R217H		4.948e-05		P; 0.9289	452	0.270	0.371
<i>HEPH1L1</i>	Hephaestin-like 1	chr11.g.93839224G>A	p.W991*				DC; 1	451	0.000	N/A
<i>NTPCR</i>	Nucleoside-triphosphatase, cancer-related	chr11.g.233091444G>A	p.R59Q		5.779e-05		DC; 0.9997	439	0.034	0.502
<i>XAB2*</i>	XPA binding protein 2	chr19.g.7688159G>C	p.T379R		1.652e-05		DC; 0.9999	420	0.059	0.200
<i>CHERP</i>	Calcium homeostasis endoplasmic reticulum protein	chr19.g.16631044C>T	p.R793H	rs202164310 MAF=0.0000	0.0001009		DC; 0.9371	366	0.12	0.716
<i>MEX3D</i>	Mex-3 homolog D	chr19.g.1555839G>C	p.T560R	rs538022731 MAF=0.0002			P; 0.8576	336	0.030	N/A
<i>BRAT1*</i>	BRCA1-associated ATM activator	chr7.g.2594007C>T	p.R20K	rs143390199 MAF=0.00002	1.651e-05		DC; 0.9349	333	0.192	0.010
<i>BRAT1*</i>	BRCA1-associated ATM activator	chr7.g.2580668G>A	p.T447M	rs368808380 MAF=0.0002	5.845e-05		P; 0.9999	333	0.110	0.275
<i>CUL9</i>	Cullin 9	chr6.g.43154714C>T	p.T423I				DC; 0.9979	251	0.000	0.993
<i>ALPK1</i>	Alpha-kinase 1	chr4.g.113353195A>C	p.D831A		0.0001255	0.0006	P; 0.9999	0	0.060	0.243
<i>CD37</i>	CD37 molecule	chr19.g.49840212C>G	p.I63M		2.476e-05	0.0003	P; 0.9999	0	0.040	0.028

Gene	Gene Description	Variant	Position	P-value	Conservation	Pathogenicity	Other
<b>IQCF6</b>	IQ motif containing F6	chr3:g.51812782G>A	p.R61W	P: 0.9999	0	0.010	N/A
<b>LFNG</b>	LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyl-transferase	chr7:g.25666829C>T	p.R375C	P: 0.9999	1.69e-05	0.020	0.772
<b>MGA</b>	MAX dimerization protein	chr15:g.41988923C>T	p.S572L	P: 0.9999	0	0.130	N/A
<b>POLJ</b>	Polymerase (DNA directed) iota	chr18:g.51820404T>C	p.V597A	P: 0.9999	0.00024	0.590	N/A
<b>SHC4</b>	SHC (Src homology 2 domain containing) family, member 4	chr15:g.49254675G>T	p.H180N	P: 0.9999	0.0003	1.000	0.000
<b>SMS</b>	Spermine synthase	chrX:g.21958982G>C	p.G14R	P: 0.9998	0	0.350	0.002
<b>SNAPC4</b>	Small nuclear RNA activating complex, polypeptide 4	chr9:g.139272279C>G	p.G1334R	P: 0.9999	2.675e-05	0.160	0.707

<sup>1</sup>Genomic position of the variant in the human reference genome GRCh37

<sup>2</sup>Reference SNP ID number (rs) and MAF (minor allele frequency) for the already described variants

<sup>3</sup>Allele frequency for the already described variants in ExAC database

<sup>4</sup>Allele frequency for the already described variants in Collaborative Spanish Variant Server (CSVS) database (<http://cvs.babelomics.org/>)

<sup>5</sup>Prediction of disease potential by Mutation Taster and probability of the prediction. DC=Disease Causing; P=Polymorphism (<http://www.mutationtaster.org/>)

<sup>6</sup>Conservation score from PhastCons (0 to 1000), being 1000 the most conserved locus and 0 a non-conserved locus

<sup>7</sup>Sift: 0-0.05 damaging (in bold); 0.051-1 tolerable (non-damaging)

<sup>8</sup>PolyPhen: 0-0.4 benign; 0.41-0.89 possibly damaging; 0.9-1 pathogenic (in bold)

\*Present in a double-mutant allele

**Supplemental Table S3.** Variants found in unrelated AFF patients in genes from Supplemental Table 2

Gene	Protein	Variant <sup>1</sup>	Effect on the protein	dbSNP <sup>2</sup>	ExAC <sup>3</sup>	CSVS <sup>4</sup>	Mutation Taster <sup>5</sup>	Conservation <sup>6</sup>	Sift <sup>7</sup>	PolyPhen <sup>8</sup>	AFF Patient
<b>BRAT1</b>	BRCA1-associated ATM activator	chr7:g.2580636C>T	p.E458K				P; 0.9999	333	0.568	0.000	AFU3
<b>CYP11A1</b>	Cytochrome 450, family 1, subfamily A, polypeptide 1	chr15:g.75014793T>A	p.S216C	rs146622566 MAF=0.0003	0.0001153		P; 0.9999	0	<b>0.004</b>	<b>0.987</b>	AFU1

<sup>1</sup>Genomic position of the variant in the human reference genome GRCh37

<sup>2</sup>Reference SNP ID number (rs) and MAF (minor allele frequency) for the already described variants

<sup>3</sup>Allele frequency for the already described variants in ExAC database

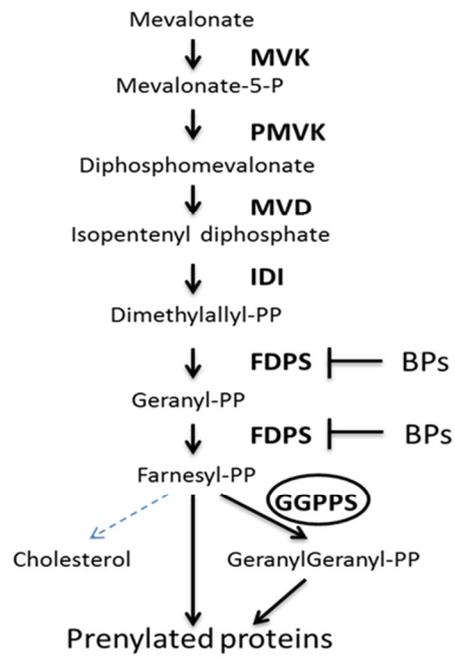
<sup>4</sup>Allele frequency for the already described variants in Collaborative Spanish Variant Server (CSVS) database (<http://csvs.babelomics.org/>)

<sup>5</sup>Prediction of disease potential by Mutation Taster and probability of the prediction. DC=Disease Causing; P=Polymorphism (<http://www.mutationtaster.org/>)

<sup>6</sup>Conservation score from PhastCons (0 to 1000), being 1000 the most conserved locus and 0 a non-conserved locus

<sup>7</sup>Sift: 0-0.05 damaging (in bold); 0.051-1 tolerable (non-damaging)

<sup>8</sup>PolyPhen: 0-0.4 benign; 0.41-0.89 possibly damaging; 0.9-1 pathogenic (in bold)



**Supplementary Figure S1.** GGPPS participate in the mevalonate pathway: bisphosphonates act by inhibiting the FPPS, thereby preventing prenylation and activation of small GTPases that are essential for the activity and survival of osteoclasts.

## Annex to Chapter 2

In the initial version of this article there was a candidate gene study in the 3 unrelated patients and a gene/protein network analysis that were finally published in an article in the Spanish journal *Revista de Osteoporosis and Metabolismo Mineral*, included in the Annex section of this thesis.

### Material and methods

#### Candidate gene analysis

Exome sequencing data from the 3 unrelated patients were analysed using the Integrative Genomics Viewer (IGV; <http://www.broadinstitute.org/igv>) with the aim of finding mutations in candidate genes (Table 1). Mutations were selected based on the same criteria used for the three sisters' exomes in Roca-Ayats *et al.*<sup>1</sup> and were tested in 3 controls (long-term BPs use but without AFF) and in the ExAC database.

#### Network construction

The Atypical Femoral Fractures Gene interaction Network (AFFGeNet) was constructed as in Boloc *et al.*<sup>2</sup> to identify genes or proteins that interacted with the 37 AFF driver genes (Supplemental Table S2 from Roca-Ayats *et al.*<sup>1</sup> and Table 2) considering the interactions directional and binary.

High-throughput interaction data were retrieved from BioGRID (version 3.4.133)<sup>3,4</sup> and STRING (Search Tool for the Retrieval of Interacting Genes/Proteins, version 10)<sup>5</sup>, with additional information from GeneOntology (<http://geneontology.org>), GeneCards, OMIM, UniProt, RefSeq, and UCSC. This whole human gene/protein interaction network included 26,934 nodes and 794,052 edges.

A Perl script was implemented to capture the interactions subnetwork using AFF genes to find all possible pair-wise shortest paths by applying the Dijkstra algorithm implemented in the Graph Perl module. The Graph::Directed module was used to define the whole network data structure as a directed graph, which simplified the calculations for the AFF subnetwork. Pair-wise connectivity was explored using Circos<sup>6</sup>. The script produced a skeleton graph stored in JSON format to make data available on the AFFGeNet web interface (<https://compgen.bio.ub.edu/AFFgenes/>, available upon request). This web interface was developed for user-friendly network exploration by

researchers. It was implemented via PHP scripts to process queries, integrate the data, and display the resulting network through the open-source Cytoscape JavaScript library for graph analysis and visualization<sup>7</sup>. The main web form provides one entry point that focuses on selected genes (similarly to other current gene/protein browsers). The web display facilitates interaction with the nodes by zooming, displacing, changing the graph layout (user can choose that from a list including grid, random, circle, breadthfirst, cose, concentric, and so on), adding or removing nodes, and retrieving information about AFF genes. The border colour of the nodes identifies them as drivers (purple), as downstream (green) or upstream (turquoise) partners of selected drivers, and as “other” (grey). The filling core of the nodes encodes bone-specific gene expression, which was retrieved from the Gene Expression Omnibus (GEO)<sup>8</sup>. The colour scale goes from intense yellow (underexpressed) to dark blue (overexpressed), with white indicating no change of expression. A bone-related GSE database was included in the network construction: Osteoclastic precursor cells treated or not with bisphosphonates (alendronate or risedronate) during their differentiation into mature osteoclasts<sup>9</sup> (GSE63009). For this specific task, a standard protocol based on the Bioconductor<sup>10</sup> limma R package was run.

### Pathway enrichment analysis

Functional Enrichment Analysis was assessed using the DAVID bioinformatics tool<sup>11</sup> (<https://david.ncifcrf.gov/>).

## **Results**

### Candidate gene analysis in three unrelated AFF patients

We used IGV to screen, in the exomes of the three unrelated patients, several candidate genes involved in bone metabolism, osteoclast function and the mevalonate pathway (listed in Table 1). Three variants were found and validated in *MMP9* (AFU3), *MVD* (AFU2) and *RUNX2* (AFU3) (Table 2). The mutation in *MMP9*, coding for type IV collagenase, involved the change of a hydrophobic amino acid within the catalytic domain to a hydrophilic residue. This variant was predicted to be damaging according to SIFT and PolyPhen scores. The *MVD* variant p.Arg97Gln, rs376949804, was predicted to be non-damaging by SIFT and PolyPhen. The mutation in *RUNX2* is a substitution of a cyclic amino acid for an aliphatic hydrophobic amino acid in a proline/serine/threonine-rich region. This change is described in dbSNP (NCBI) as rs201584115 with a minor allele frequency (MAF) of 0.0004 and is predicted to be probably damaging to its function.

Any of these mutations were found in 3 controls (long-term treated with BPs without AFF) and were present in ExAC with a MAF<0.0005.

### Gene/protein interaction network and pathway analysis

Gene/protein connections were constructed to investigate functional pathways related to the 37 mutated genes detected in the WES approach and to detect potential causative genes and the molecular mechanisms that might be involved in the generation of AFFs. Incoming and outgoing connections for all genes at distances 1 to 4 are summarized in Figure 1. *FN1* is the only gene connected with others at distance 1. At distance 2, more connectivity is observed. The majority of the pairwise shortest path connectivity for AFF driver genes is observed at distance 3. The only gene without any interaction at any level is *IQCF6*.

The network of gene/protein interactions shows that *GGPS1* and *CYP1A1*, two of the most relevant driver genes, are connected at distance 3, through *INS* and *IL6* (Figure 2A). Four other AFF driver genes (*RUNX2*, *MVD*, *MMP9* and *PGRMC1*) are connected to either of them at distance 2. Furthermore, *FN1* appears connected to *MMP9* at distance 1. Likewise, the driver genes *SYDE2* and *NGEF*, which are small GTPase activators, are interconnected at distance 2 through *RHOB* (Figure 2B).

Pathway interrogation with the DAVID web tool yielded the isoprenoid biosynthetic pathway (GO:0008299), containing the *GGPS1*, *MVD* and *CYP1A1* genes, as enriched among the 37 mutated genes (p-value 0.0006).

## **Discussion**

Several other genes with variants in the three sisters might also contribute to AFF susceptibility. *FN1* encodes an extracellular matrix protein necessary for the regulation of type I collagen deposition by osteoblasts, essential for matrix mineralization, and fibronectin levels have been shown to be affected by BP treatment<sup>12</sup>. *SYDE2* and *NGEF* encode two regulators of small GTPases. Their respective roles in activating RHO GTPases and in exchanging their guanine nucleotides constitute interesting clues to their putative effects on osteoclast function and responses to BPs. RHO GTPases are downstream targets of the BPs since they need to be prenylated for their cellular function. Additionally, our gene/protein interaction network shows how *NGEF* is tightly related to the ephrins and ephrin receptors (Figure 2B), which are key players in the coupling mechanism between osteoclasts and osteoblasts<sup>13</sup>. Another group of genes mutated in

the 3 sisters encode nuclear proteins with pleiotropic effects on gene expression and/or DNA repair (*KDM4C*, *XAB2*, *NVL*, *NKAP*, *ERCC6L2*). Notably, *KDM4C* encodes a JmjC-domain-containing lysine-specific demethylase recently found associated with age at menarche<sup>14</sup>, which is a biomarker for bone density. *PGRMC1*, which encodes progesterone receptor membrane component 1 and was mutated in the sisters, was previously reported to be involved in premature ovarian failure<sup>15</sup>. Finally, *COG4* (encoding subunit 4 of the conserved oligomeric Golgi Complex) and *EML1* (encoding echinoderm microtubule-associated protein-like 1) are of interest, given the importance of vesicle transport through the Golgi in osteoclasts<sup>16</sup> and of the primary cilium in osteocytes<sup>17</sup>, respectively.

Through a candidate gene approach, two crucial proteins for bone remodelling (RUNX2 and MMP9) and another enzyme of the mevalonate pathway (mevalonate diphosphate decarboxylase, MVD) were found mutated in 3 unrelated AFF patients. RUNX2 is a master transcription factor for osteoblastic differentiation<sup>18</sup>, while MMP9, a metalloproteinase expressed in osteoclasts, degrades the extracellular matrix of bone<sup>19</sup>. As a consequence, it influences the architecture of trabecular bone and the structure of cortical bone<sup>20</sup> both of which might be involved in AFF risk. RUNX2 is known to activate MMP9 gene expression<sup>21</sup> and this interaction may have synergistic effects on the biomechanical properties of bone in patient AFU3, who bears these two mutations (Note: This interaction is not shown in Figure 2A so that other interactions could be clearly displayed). Finally, an MVD missense mutation predicted as tolerable was confirmed in patient AFU2, adding a second mutated gene in the mevalonate pathway.

All in all, the functions and prior knowledge on these genes are commensurate with their possible involvement in the pathology, and especially the observed alterations in the mevalonate pathway. Taken together, all these rare variants may belong to a genetic pool that provides the background for the development of bone changes that give rise to AFFs and the possible negative interaction with BPs. It is likely that several genes with small additive effects, and their interactions, are involved in long-term BP-related AFFs. Furthermore, each individual patient would be a carrier of different specific genetic variants.

## References

1. Roca-Ayats N, Ng PY, Garcia-Giralt N, Falcó-Mascaró M, Cozar M, Abril JF, *et al.* Functional characterization of a GGPPS variant identified in atypical femoral fracture

patients and delineation of the role of GGPPS in bone-relevant cell types. *J Bone Miner Res.* 2018;33(12):2091–2098.

2. Boloc D, Castillo-Lara S, Marfany G, Gonzalez-Duarte R, Abril JF. Distilling a Visual Network of Retinitis Pigmentosa Gene-Protein Interactions to Uncover New Disease Candidates. *PLoS One.* 2015;10(8):e0135307

3. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* 2006;34(Database issue):D535-9.

4. Chatr-Aryamontri A, Breitkreutz BJ, Heinicke S, Boucher L, Winter A, Stark C, *et al.* The BioGRID interaction database: 2013 update. *Nucleic Acids Res.* 2013;41(Database issue):D816-23.

5. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 2013;41(Database issue):D808-15.

6. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, *et al.* Circos: an information aesthetic for comparative genomics. *Genome Research.* 2009;19(9):1639-45.

7. Lotia S, Montojo J, Dong Y, Bader GD, Pico AR. Cytoscape app store. *Bioinformatics.* 2013;29(10):1350-1.

8. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, *et al.* NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res.* 2013;41(Database issue):D991-5.

9. Yuen T, Stachnik A, Iqbal J, Sgobba M, Gupta Y, Lu P, *et al.* Bisphosphonates inactivate human EGFRs to exert antitumor actions. *Proceedings of the National Academy of Sciences of the United States of America.* 2014;111(50):17989-94.

10. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004;5(10):R80.

11. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44-57.

12. Insalaco L, Di Gaudio F, Terrasi M, Amodeo V, Caruso S, Corsini LR, *et al.* Analysis of molecular mechanisms and anti-tumoural effects of zoledronic acid in breast cancer cells. *J Cell Mol Med* 2012; 16(9):2186-2195.
13. Sims NA, Martin TJ. Coupling the activities of bone formation and resorption: a multitude of signals within the basic multicellular unit. *Bonekey Rep* 2014; 3:481.
14. Perry JR, Day F, Elks CE, Sulem P, Thompson DJ, Ferreira T, *et al.* Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature*. 2014;514(7520):92-7
15. Mansouri MR, Schuster J, Badhai J, Stattin EL, Losel R, Wehling M, *et al.* Alterations in the expression, structure and function of progesterone receptor membrane component-1 (PGRMC1) in premature ovarian failure. *Hum Mol Genet*. 2008;17(23):3776-83
16. Xia WF, Tang FL, Xiong L, Xiong S, Jung JU, Lee DH, *et al.* Vps35 loss promotes hyperresorptive osteoclastogenesis and osteoporosis via sustained RANKL signaling. *J Cell Biol*. 2013;200(6):821-37.
17. Nguyen AM, Jacobs CR. Emerging role of primary cilia as mechanosensors in osteocytes. *Bone*. 2013;54(2):196-204.
18. Ducy P, Zhang R, Geoffroy V, Ridall AL, Karsenty G. *Osf2/Cbfa1*: a transcriptional activator of osteoblast differentiation. *Cell* 1997; 89(5):747-754.
19. Okada Y, Naka K, Kawamura K, Matsumoto T, Nakanishi I, Fujimoto N, *et al.* Localization of matrix metalloproteinase 9 (92-kilodalton gelatinase/type IV collagenase = gelatinase B) in osteoclasts: implications for bone resorption. *Lab Invest* 1995; 72(3):311-322.
20. Nyman JS, Lynch CC, Perrien DS, Thiollou S, O'Quinn EC, Patil CA, *et al.* Differential effects between the loss of MMP-2 and MMP-9 on structural and tissue-level properties of bone. *J Bone Miner Res* 2011; 26(6):1252-1260.
21. Pratap J, Javed A, Languino LR, van Wijnen AJ, Stein JL, Stein GS, *et al.* The Runx2 osteogenic transcription factor regulates matrix metalloproteinase 9 in bone metastatic cancer cells and controls cell invasion. *Mol Cell Biol* 2005; 25(19):8581-8591.

Tables

**Table 1.** Candidate genes screened in unrelated AFF patients using IGV

<i>ACP5</i>	<i>CALCR</i>	<i>FDPS</i>	<i>OPG</i>	<i>SP7</i>
<i>ALPL</i>	<i>COL1A1</i>	<i>LRP5</i>	<i>OPN</i>	<i>TNFRSF11A</i>
<i>BGLAP</i>	<i>COL1A2</i>	<i>LRP6</i>	<i>OSTM1</i>	<i>TNFSF11</i>
<i>BSP</i>	<i>CTSK</i>	<i>MMP9</i>	<i>PTH</i>	<i>VDR</i>
<i>BMPs</i>	<i>DMP1</i>	<i>MVD</i>	<i>RUNX2</i>	<i>VTN</i>
<i>CALCA</i>	<i>ESR1</i>	<i>MVK</i>	<i>SOST</i>	<i>WNTs</i>

**Table 2.** Variants found in candidate genes in unrelated AFF patients

Gene	Protein	Variant <sup>1</sup>	Effect on the protein	dbSNP <sup>2</sup>	ExAC <sup>3</sup>	CSVS <sup>4</sup>	Mutation Taster <sup>5</sup>	Conservation <sup>6</sup>	Sift <sup>7</sup>	PolyPhen <sup>8</sup>	AFF Patient
<i>MMP9</i>	Matrix metalloproteinase 9	chr20:g.44641147 T>C	p.M419T		8.242e-06		DC; 0.9999	496	<b>0.000</b>	<b>1.000</b>	AFU3
<i>MVD</i>	Mevalonate diphosphate decarboxylase	chr16:g.88723957 C>T	p.R97Q	rs376949804 MAF= 3e-05	3.448e-05	0.0003	P; 0.9999	0	0.448	0.009	AFU2
<i>RUNX2</i>	Runt-related transcription factor 2	chr6:g.45480010 C>T	p.P296L	rs201584115 MAF=0.0004	0.0002066	0.0003	DC; 0.9999	642	<b>0.040</b>	<b>0.999</b>	AFU3

<sup>1</sup>Genomic position of the variant in the human reference genome GRCh37

<sup>2</sup>Reference SNP ID number (rs) and MAF (minor allele frequency) for the already described variants

<sup>3</sup>Allele frequency for the already described variants in ExAC database

<sup>4</sup>Allele frequency for the already described variants in Collaborative Spanish Variant Server (CSVS) database

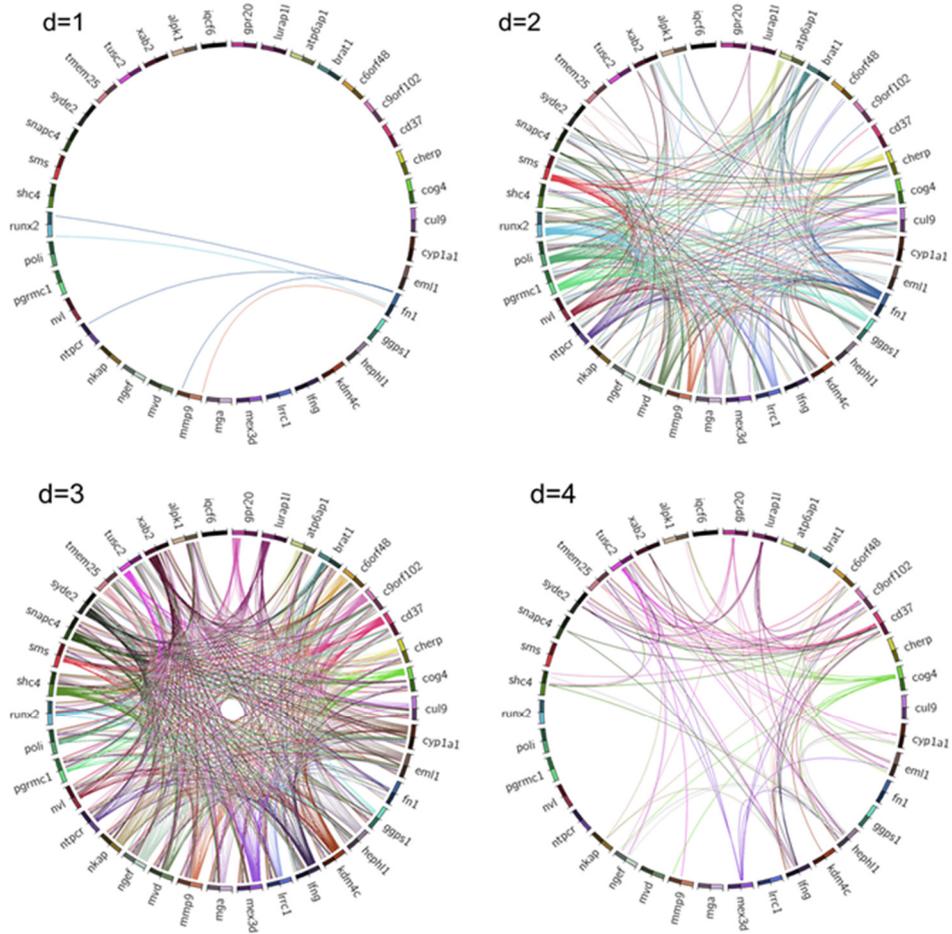
<sup>5</sup>Prediction of disease potential by Mutation Taster and probability of the prediction. DC=Disease Causing; P=Polymorphism

<sup>6</sup>Conservation score from PhastCons (0 to 1000), being 1000 the most conserved locus and 0 a non-conserved locus

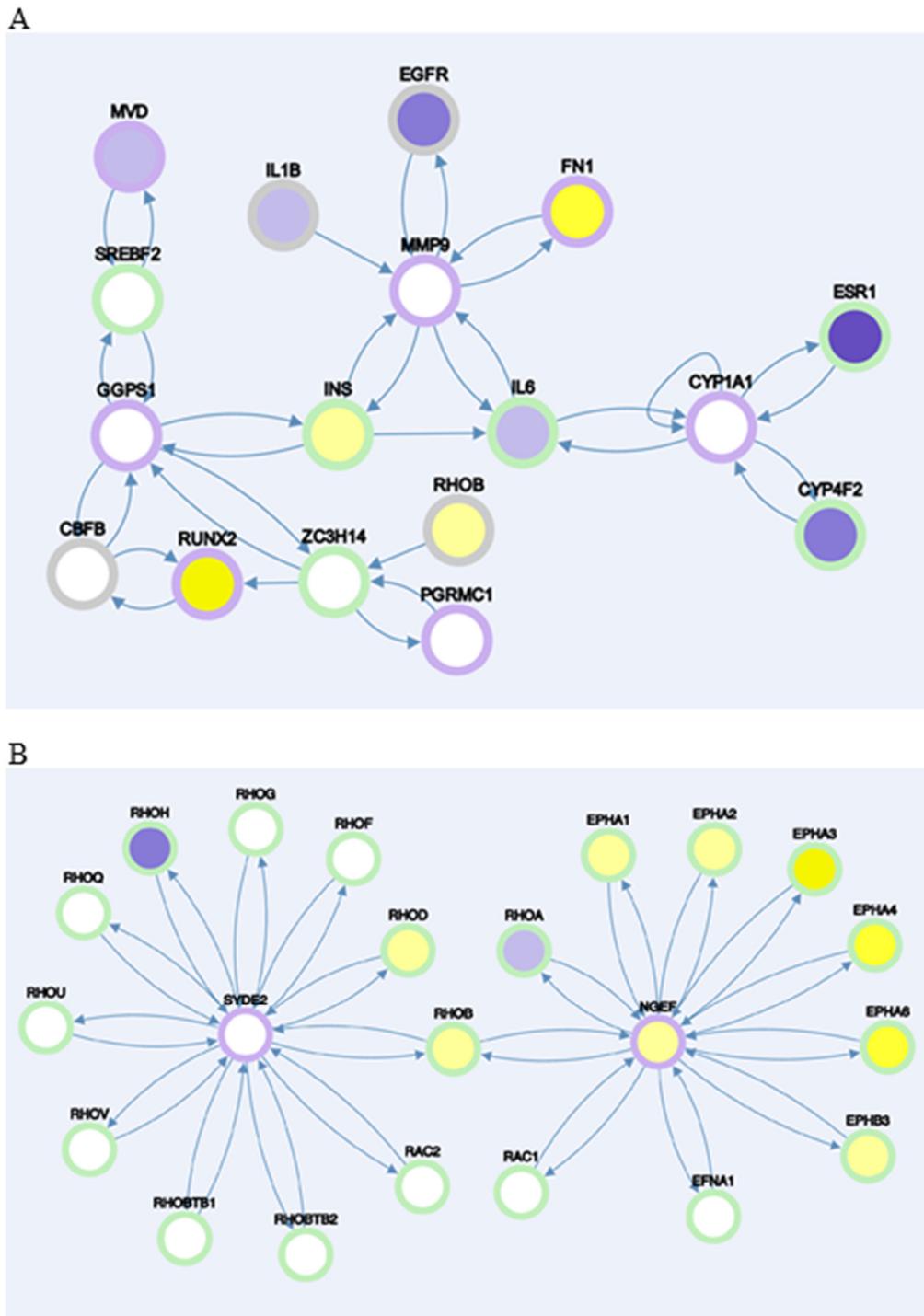
<sup>7</sup>Sift: 0-0.05 damaging (in bold); 0.051-1 tolerable (non-damaging)

<sup>8</sup>PolyPhen: 0-0.4 benign; 0.41-0.89 possibly damaging; 0.9-1 pathogenic (in bold)

Figures



**Figure 1:** Summary of the pair-wise connectivity of the genes found mutated in the AFF patients (MAP genes), showing incoming and outgoing connections. Standard symbols for the 37 MAP genes used to build the network are depicted on the outer ring of the diagram. Each driver gene that interacted with another MAP gene is shown by a distinct color box, while those for which there was no reported interaction are shown in black.



**Figure 2:** **A.** AFGeNet visualization focusing on some *GGPS1* and *CYP1A1* partners at distance 2 (and some of the *MMP9* partners at distance 1). **B.** AFGeNet visualization of some *SYDE2* and *NGEF* partners at distance 1.

Please note that many connections have been omitted for the sake of clarity. In particular, the nodes *RUNX2* and *FN1* have not been expanded to display all of their partners. Fill color is indicative of underexpression (yellow), overexpression (purple) and no change in expression (white) in osteoclasts treated with alendronate or risedronate (data from Yuen *et al.* 2014<sup>8</sup>). Outer color code is purple for genes found mutated in our AFF patients (MAP driver genes); green for genes downstream of the chosen node genes; grey for other genes.



# **DISCUSSION**



Osteoporosis is a common disorder determined by both genetic and environmental factors, as well as by their interaction. This PhD thesis aimed at contributing to the elucidation of the genetic determinants of osteoporosis and atypical femoral fracture (AFF), an extremely rare complication of nitrogen-containing bisphosphonates (N-BPs) therapy for osteoporosis. On the one side, a genome-wide association study (GWAS) signal has been deeply studied, including dissection of the association signal in a Spanish cohort and the functional characterization of the associated variants in their genomic context. On the other side, a small cohort of N-BP-associated AFF patients has been sequenced and the most interesting mutation found has been functionally characterized using molecular and cellular approaches.

To elucidate the molecular bases of diseases several aspects have to be considered and are discussed here.

## **1. Homogeneity of phenotype**

One of the most important aspect when studying the genetic component of a disease is the homogeneity and accuracy in the phenotyping. It is especially relevant when studying various individuals together, such as large cohorts of individuals used in association analysis for complex diseases, families used in linkage studies or small groups of individuals, and also when comparing or replicating studies. Phenotyping is also crucial to improve the understanding of disease pathogenesis. Otherwise, a bad phenotypic selection can prevent the identification of causal genetic variant/s (Vissers & Veltman, 2015).

In the case of osteoporosis, bone mineral density (BMD) and osteoporotic fracture (OF) are the phenotypes used for its diagnosis and they are mainly used in association studies. BMD is a quantitative variable with a clear hereditary component that can vary depending on the measure site or the technique used (Lorentzon & Cummings, 2015). For this reason, it is essential that all the individuals involved in an association study have the BMD measured at the same site and with the same technique (usually DXA or QUS). OF is a qualitative phenotype that is also widely used because the obtention of predictive models (such as FRAX) have a great impact in clinical practice. Again, OF can occur at different body sites and, thus, when used in association studies, the location has to be clearly defined. For example, some studies have performed a GWAS in hip fracture or vertebral fracture, while some others have used all type of fractures to identify

genetic variants predisposing to OF in general (Alonso *et al.*, 2018; Trajanoska *et al.*, 2018). Other bone parameters, such as size, porosity or cortical bone thickness, have been used in GWASs since they are also good measures of bone health.

In this thesis (Article 1, Chapter 1), a deep re-sequencing and an association study of a genomic region (7q21.3) previously found in several GWASs as associated with BMD and OF has been performed. We used the BARCOS cohort, which consisted of 1490 unrelated postmenopausal women of Spanish descent from the Barcelona area, monitored at the Hospital del Mar (Barcelona). Exclusion criteria were any history of bone diseases, metabolic or endocrine disorders, hormone-replacement therapy or use of drugs that could affect bone mass. Therefore, the cohort meets the homogeneity standards desirable for this kind of studies. Lumbar spine (LS) and femoral neck (FN) BMD measured by DXA and OF, including all skeletal sites, were used.

Interestingly, it has been observed that each phenotype (including OF or BMD at different skeletal sites) can have singular genetic determinants and, therefore, a particular gene or variant can give different results depending on the phenotype assessed. In this line, we genotyped 8 SNPs in the BARCOS cohort and we also included the rs4727338 SNP genotyped previously in Estrada *et al.* (2012) and 3 of them (rs10085588, rs4342521 and rs4727338) showed significant association with LS-BMD. However, only rs4727338 was nominally associated with FN-BMD, although it could also be due to the smaller sample size for FN-BMD association. Regarding OF, the 3 SNPs showed nominal association.

In the case of AFFs, an accurate case definition is also of paramount importance so that they can be clearly distinguished from ordinary osteoporotic femoral fractures. Moreover, since they are a very rare event whose pathogenesis and pathophysiology remain largely unknown, it is very important that all studies report on the same condition, so that we can understand more and more all aspects related to AFFs. To address this question, the American Society for Bone and Mineral Research (ASBMR) established the case definition in 2010 and updated it in 2014, on the base of newer evidence (see Table 6; Shane *et al.*, 2010, 2014). In this line, all the patients we studied in Chapter 2 (Articles 3 and 4), coming from 2 different hospitals (Hospital Universitario Reina Sofía, Córdoba, and Hospital del Mar, Barcelona), fulfilled the ASBMR criteria for AFF diagnosis. In addition, all of them had taken BPs for more than 5 years, excluding those cases of BP-independent AFF. Besides, individuals who had taken N-BPs for more than 5 years but who did not sustain an AFF were used as controls, in which the variants identified in the patients were not present. This approach allowed us to exclude some

variants putatively related to the underlying osteoporotic phenotype, rather than with the AFF. However, as osteoporosis is a complex disease where many small-effect variants contribute, the number of controls used (n=3) is clearly far too small to disregard all the osteoporosis-related variants.

Importantly, and as mentioned in the Introduction, very similar fractures can occur in other monogenic skeletal diseases, including disorders of mineralization (e.g. hypophosphatasia [HPP]), impaired bone remodelling (e.g. pycnodysostosis [PYCD]) and defects on collagen synthesis and structure (e.g. osteogenesis imperfecta [OI]). Furthermore, occasionally, AFF has been the presenting symptom of an unsuspected HPP. For example, both Sutton *et al.* (2012) and Peris *et al.* (2019) reported mutations in *ALPL* in AFF patients initially misdiagnosed with postmenopausal osteoporosis and treated with N-BPs. After the AFF, these patients presented an increase in ALP substrates and low serum ALP activity, consistent with the diagnosis of HPP. The occurrence of these fractures in patients with other monogenic bone disorders might be confusing when studying the genetics underlying AFF. Thus, we excluded patients presenting with other bone diseases, such as HPP, from our small cohort. Further studies to understand the putative differences among these fractures and AFFs occurring in osteoporotic patients are needed.

## 2. Next generation sequencing

Nowadays, several methods are used to identify the genetic bases of diseases. Different approaches are used depending on the biological hypothesis in question, including characteristics of the disease, such as its incidence and inheritance pattern, variant type of interest and size of the regions of interest.

The sequencing of the human genome and the advent of next generation sequencing technologies (NGS) boosted the discovery of new genes involved in many kinds of diseases and traits (Goldstein *et al.*, 2013). Currently, several NGS platforms that use different sequencing methods are available, with their own advantages and limitations (reviewed in Chakravorty & Hegde, 2017). Remarkably, read lengths are relatively short (35-700 bp) and error rates range from 0.1% to 15% and, hence, Sanger sequencing validation and/or deep coverage are needed.

NGS technologies have enabled a quick and cost-effective sequencing of genomes, exomes and gene panels or regions of interest. Whole genome sequencing (WGS) offers

the most comprehensive discovery of all type of genetic variants (coding and non-coding, common and rare, structural, etc.) across the genome. Targeted enrichment sequencing allows for selective sequencing of regions of interest, reducing the amount of data generated, as well as the cost. The most widely used is whole exome sequencing (WES), which sequences the protein-coding regions in the genome (around 1.2% of the genome) and has been proved to be successful in uncovering mutations causing Mendelian diseases. It has been estimated that exonic mutations cause the majority of monogenic diseases. However, this estimation may be biased by the difficulty of identifying disease-causing mutations in the non-coding genome (Petersen *et al.*, 2017; Sun *et al.*, 2015). Currently, one of the main challenges resides in interpreting NGS data: prioritizing, validating and functionally characterizing the variants (see section 5 of this Discussion)

In this thesis (Articles 3 and 4, Chapter 2), we took advantage of NGS to gain insight into the genetic architecture of N-BP-associated AFF. Before this thesis, no familial clustering had been reported and, therefore, no segregation analyses had investigated the architecture of AFF. Given the low incidence of N-BP-associated AFFs and the availability of 3 affected sisters, we hypothesized that there was a genetic predisposition in the form of shared rare genetic variants and performed WES of the 3 sisters and 3 unrelated patients, which allowed us to identify putative novel variants. After filtering non-synonymous and rare variants (MAF<0.005) and prioritizing them, we found 37 heterozygous variants in 34 genes shared by the 3 sisters. We validated the variants by Sanger sequencing and we further studied the most conserved one, a novel mutation in *GGPS1*. Afterwards, another WES study of consanguineous familial AFF without N-BP exposure was performed in which a mutation in the *CTSK* gene was detected, although the family did not present any clinical feature of PYCD (Lau *et al.*, 2017).

Furthermore, we deeply explored the genetic contribution of a candidate region that arose from BMD and OF GWASs (namely, *C7ORF76* at 7q21.3; Article 1, Chapter 1) by means of ultra-deep sequencing (3600x coverage) of 7 overlapping LR-PCR-amplified fragments in two extreme LS-BMD groups of women (n=50 per group) from the BARCOS cohort. We identified and compared the number and frequency of variants present in each group. The most interesting variants were selected and tested for association in the whole cohort. Afterwards, further functional analyses were performed. With this approach we aimed at identifying all the variation within this *locus*, including variants that were not captured in GWASs because of their low frequency, low effect size or structural characteristics (that is, CNVs, microsatellites, etc.). Targeted resequencing of *loci* emerging from GWASs in truncate selections of patients has been successfully

performed by other research groups, including a work on *loci* associated with BMD in which the authors identified new rare non-coding variants (Hsu *et al.*, 2016).

NGS technology is also used in functional genomic applications going beyond the identification of DNA genetic variants. For example, they are used in ChIP-seq, ATAC-seq, methyl-seq, 4C-seq, RNA-seq studies to functionally characterize the genome or mRNAs (e.g. DNA-protein interaction, histone modification, chromatin accessibility and interaction, DNA methylation, mRNA expression levels, alternative splicing or transcript discovery; Chakravorty & Hegde, 2017).

### 3. Association studies

Association studies are one of the most used methodologies to identify genetic susceptibility variants for complex diseases. In the osteoporosis field, and as described in the Introduction, many candidate gene association studies and GWASs have been performed. The main advantage of GWASs is the possibility of exploring, in a hypothesis-free way, the genetic variability scattered across the genome associated with a trait of interest. Therefore, many variants, genes or genomic regions associated with bone characteristics that do not have any known relationship with bone biology have been identified. Such is the case of the *C7ORF76 locus* in 7q21.3, which has been repeatedly found associated with LS-BMD, FN-BMD, OF, heel eBMD and total body BMD (Duncan *et al.*, 2011; Estrada *et al.*, 2012; Kemp *et al.*, 2017; Medina-Gomez *et al.*, 2018; Morris *et al.*, 2019; Rivadeneira *et al.*, 2009; Trajanoska *et al.*, 2018; Zhang *et al.*, 2014; Zheng *et al.*, 2015). As mentioned, in this thesis a two-stage approach was used, consisting in candidate region deep sequencing of extreme phenotypes as a discovery phase followed by association of selected variants with BMD and OF in a Spanish cohort to identify the causal variant for the association in the *C7ORF76 locus* (Article 1, Chapter 1).

The genetic bases of AFF have also been investigated through association studies elsewhere. As reported in the Introduction, two studies performed case-control GWASs, although using a very small number of patients. Pérez-Núñez *et al.* (2015) assessed rare known coding variants in 13 AFF patients and 268 controls and found an over-representation of risk variants in the case group. However, due to small sample size, only one variant was found significantly associated. Kharazmi *et al.* (2019) assessed common variants in 51 cases and 5215 controls and the small sample sizes precluded the identification of genome-wide significant associations. In addition, Kharazmi *et al.* (2019) also performed a candidate gene association analyses in 29 genes, again with

negative results. As already discussed in section 2 of this Discussion, we used a different approach to discover the genetic susceptibility to AFFs.

When performing association studies, several aspects have to be considered, including the homogeneity of the phenotype (already discussed in section 1 of this Discussion), and others discussed below.

In association studies sample size is pivotal, since it directly impacts the statistical power (i.e. the probability to detect a real association). Apart from sample size, there are several factors that influence the estimation of statistical power: disease prevalence, genetic heterogeneity, linkage disequilibrium between the genotyped polymorphism and the functional variant and their allelic frequencies, the effect or risk that the variant confers (i.e.  $\beta$  coefficient or odds ratio, OR) to the phenotype and its inheritance pattern (Nsengimana & Bishop, 2017).

In complex diseases, it is commonly believed that causal variants have small effects on the phenotype, correlating with the “common disease, common variant” hypothesis (Schork *et al.*, 2009). In this sense, osteoporosis is not an exception and association studies have identified a high number of genes and variants, mainly with a small effect (Morris *et al.*, 2019). Therefore, statistical power will tend to be low unless the sample size is increased considerably.

The sample size of the BARCOS cohort allowed the identification of common variants with moderate effects and/or in high linkage disequilibrium with the putative causal variants but precluded the identification of smaller effects. Notably, we failed to detect association in several of the variants interrogated, although some were found associated with BMD and OF in a large GWAS (Kim, 2018).

One way to increase sample size is creating large consortia. The GEFOS (Genetic Factors for Osteoporosis) consortium, which continues the work of the GENOMOS (Genetic Markers for Osteoporosis) consortium, was created with this goal and has carried out many large GWASs and meta-analyses for distinct bone properties and characteristics. Indeed, the BARCOS cohort participated in the replication stage in several of them.

Meta-analyses consist in analysing together data obtained from different cohorts, which provides an increased statistical power. However, the main problem they present is heterogeneity. The variants we found associated with BMD and OF were also found strongly associated in several meta-analyses (Kemp *et al.*, 2017; Medina-Gomez *et al.*, 2018; Morris *et al.*, 2019; Zheng *et al.*, 2015).

Another way of improving statistical power while avoiding a dramatic increase in sample size is using selected cohorts with minimal genetic heterogeneity, such as isolated populations (e.g. Icelanders) or extreme truncate selection of individuals for quantitative traits. For instance, one of the first GWASs in the osteoporosis field used an Icelandic discovery cohort and identified 5 genomic regions with variants significantly associated with BMD (Styrkarsdottir *et al.*, 2008). Sims *et al.* (2008) robustly identified Wnt pathway genes of relevant effect sizes involved in BMD variation by using a small cohort of postmenopausal women having either low or high hip BMD ( $n_{\text{total}}=344$ ). In 2011, Duncan *et al.* performed the first osteoporosis GWAS using an extreme-truncate selection design (1055 women with extreme high BMD and 900 women with extreme low BMD) and reported the replication of 21 of 26 known BMD-associated genes and the identification of 6 new *loci* with suggestive association. Extreme-truncate selection studies have been proven effective for detecting rare variants associated with complex traits, correlating with the “common disease, rare variant” hypothesis (Barnett *et al.*, 2013; Kang *et al.*, 2012).

In this thesis (Article 1, Chapter 1), we used an extreme-truncate selection approach with the 50 women with the highest Z-score (0.73 to 2.98) and the 50 women with the lowest Z-score (-4.26 to -2.41) of the BARCOS cohort to identify new putative causal variants in this region, including rare variants, and tested them for association in the whole cohort.

Association studies require genotyping of a certain number (ranging from a few in candidate genes studies to hundreds of thousands in GWASs) of polymorphisms, generally SNPs, in a high number of samples. In such a situation, genotyping errors may occur and affect the reliability of the results. Thus, a rigorous quality control has to be carried out. It is usual to re-genotype a certain percentage of samples or to include internal controls with previously known genotypes, as well as negative controls. In our study, the genotyping of 6% of the samples was performed in duplicate and showed a concordance above 99%.

Genotyping errors may be due to the low quality or quantity of the DNA samples or to the low reliability of the genotyping assay for a given variant. Therefore, it is important to establish the genotyping rate for individual and for variant and eliminate those individuals or variants that do not reach a specific threshold (i.e. 80-90%). Another way to detect systematic genotyping errors is testing the Hardy-Weinberg equilibrium (HWE), under the assumption that a high error rate generates disequilibrium (Pompanon *et al.*, 2005). In our study, all these quality controls have been accomplished.

Several confounding factors could lead to spurious results, either false negative or false positive. One of the most important is population stratification, mainly because of ethnicity mixture. Different ethnic groups might present a different disease prevalence, as it happens in osteoporosis, which is more prevalent in Caucasian individuals (Cauley, 2011), and/or different allelic frequency for certain polymorphisms. Therefore, it is of paramount importance that the cohort used in association studies is genetically homogeneous or, at least, individuals from different ethnicity are distributed homogeneously across the phenotype. In this thesis, all the postmenopausal women included in the cohort BARCOS are from the Barcelona area and of Spanish descent. In addition, in a previous work, the presence of population stratification was ruled out in the cohort (Agueda *et al.*, 2008).

Other confounding factors may influence the statistical association and, thus, have to be homogeneously distributed in the cohort or taken into consideration as a covariable in statistical analyses. In osteoporosis studies, some factors to take into account are gender, age, years since menopause, BMI, nutritional and hormonal status, etc. As mentioned, the BARCOS cohort was composed only by postmenopausal women and exclusion criteria included endocrine and metabolic bone disorders and certain therapies that influence bone status. Moreover, several data related to such variables were also recorded, including anthropometric measures, age, age of menarche and menopause, and number of children, among others. In our association study (Article 1, Chapter 1), we evaluated the influence of all the variables available on BMD or OF and, accordingly, used years since menopause as a covariable in the association analyses.

In association studies, a large number of variants is usually tested and, thus, the amount of statistical comparisons is also high. As a consequence, statistical significance thresholds have to be reconsidered to avoid increasing the incidence of false positives. Although there is not a universal method, several approaches have been developed. The most rigorous ones prevent the occurrence of false positives but also favour the emergence of false negatives, while the less rigorous ones produce the contrary effect.

Some commonly used multiple testing correction methods are Bonferroni's and False Discovery Rate (FDR), the first being more conservative than the latter. In our association studies (as well as in functional analyses) we applied the Bonferroni's method. Bonferroni's adjustment consists in dividing the traditional significance threshold (0.05) by the number of independent tests performed. It is easy to use, and it ensure a false positive incidence lower than 5%.

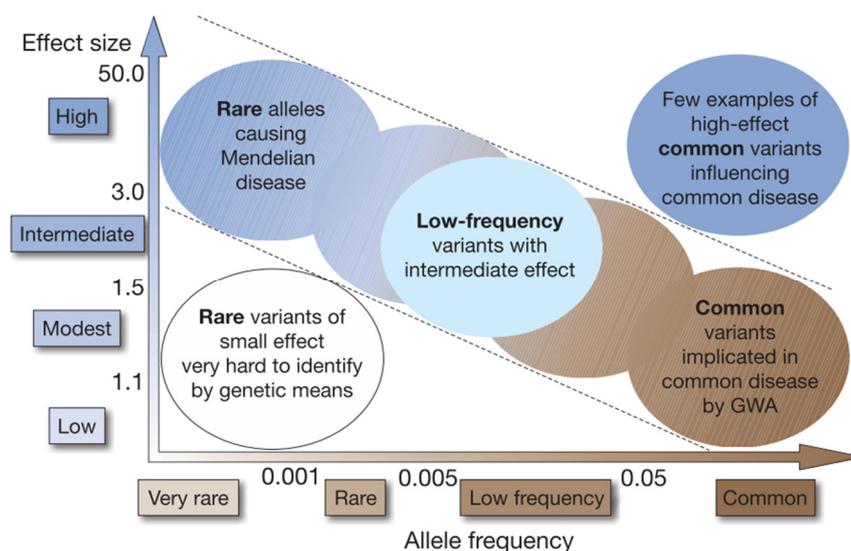
Osteoporosis is a complex disease with a heritable component. As described in the Introduction, BMD has an estimated heritability of 50-85% (reviewed in Boudin *et al.*, 2016) and other bone phenotypes have their own estimates of heritability. To date, the genetic variants identified in association studies collectively explain around 20% of the predicted genetic variance of bone-related phenotypes (Morris *et al.*, 2019). This has been referred to as the “missing heritability” problem, and it is a common finding for virtually all complex diseases (Manolio *et al.*, 2009). Therefore, there must be other variants or elements that explain this missing heritability and many options have been proposed.

Most of the association studies used SNPs as genetic markers because they are easy to genotype in a high-throughput and automatic way. However, other types of genetic variants may also contribute to complex traits. Structural variants (or CNVs), which usually have a higher impact on gene expression, might help to explain a proportion of the missing heritability of complex diseases (Nagao, 2015). As reported in the Introduction, some studies have assessed the contribution of CNVs to bone-related phenotypes (e.g. OF, BMD, hip geometry), identifying some associated *loci* (Deng *et al.*, 2010; Oei *et al.*, 2014; Yang *et al.*, 2008).

Another source of missing heritability that is being explored more and more is rare variation. Most of the association studies use common genetic variants (frequency >5%) present in the general population but with a higher frequency in affected individuals. This correlates with the “common disease, common variant” hypothesis that postulates that variability of common diseases is explained by common variants, each with small effect on the phenotype. However, the use of NGS technologies, that has allowed an in-depth sequencing of the genome, has brought up the possibility to analyse rare or low frequency variants (frequency <1% or 1-5% in the population, respectively) and their contribution to common diseases, testing the “common disease, rare variant” hypothesis that argues that rare or LF variants may have a more penetrant effect on the phenotype (Figure 10; Schork *et al.*, 2009). Notably, larger sample sizes are required for association studies with rare-low frequency variants and comparing the number or aggregated frequency of rare variants is usually useful (Zuk *et al.*, 2014).

In the field of osteoporosis, some rare and low frequency variants with large effects on BMD and OFs have been identified, in some cases using NGS-approaches (Kemp *et al.*, 2017; Morris *et al.*, 2019; Styrkarsdottir *et al.*, 2013; Zheng *et al.*, 2015). Morris *et al.* (2019) showed that the rare and low frequency variants they identified explained a 0.8% and 1.7% of the variance in eBMD, respectively. Notably, only a small amount of rare

and low frequency variants has been identified so far and, thus, their contribution to heritability is expected to increase.



**Figure 10.** Feasibility of identifying genetic variants by risk allele frequency and effect size (odds ratio). Extracted from (Manolio *et al.*, 2009).

Importantly, and as discussed in section 2 of this Discussion, deep re-sequencing of previously associated regions (using common variation) provide an opportunity to detect and assess rare variants and structural variants comprehensively. In Article 1 (Chapter 1) of this thesis we deeply-explored the genetic variation in the *C7ORF76 locus*, analysing both common and rare-low frequency variants (MAF<5%). We found several interesting low frequency variants but, due to the small sample size of the BARCOS cohort, we were not able to detect any association with BMD or OF. Rare-low frequency variants were grouped and compared together between the two extreme groups, but their distribution was balanced.

Some studies suggest that genetic variants whose associations failed to reach genome-wide significance ( $p < 5 \times 10^{-8}$ ) also contribute to phenotypic variation (Zhang *et al.*, 2012). Moreover, they argue that multiple variants at a single *locus* may jointly influence a trait (i.e. allelic heterogeneity). Yet, the dissection of allelic heterogeneity is complicated by the correlation between SNPs and an accurate estimation of allelic effects is needed to identify all the effective SNPs in a *locus* (Gusev *et al.*, 2013; Zhang *et al.*, 2012). In this sense, re-sequencing of known *loci* (including *loci* with lower significance levels) is again a good strategy to elucidate all the putative missing heritability buried in them.

Complex diseases are related to multiple intermediate phenotypes, involved in their pathogenesis. Sometimes, the genetic component of intermediate phenotypes is not detectable when studying the main complex disease but it also contributes to its heritability (Blanco-Gómez *et al.*, 2016).

It has also been proposed that part of the missing heritability may be explained by genetic interactions or epistasis, as well as gene-environment interaction (see section 6 of this Discussion). Briefly, different variants interact to regulate variation in the phenotype, giving a non-additively effect (Zuk *et al.*, 2012). Demonstrating and mapping epistasis and gene-environment interactions is challenging due to large sample size needed. Thus, the magnitude of the contribution of epistasis to the missing heritability of complex phenotypes is difficult to determine (Mackay, 2014).

Finally, epigenetic changes may also account for a proportion of missing heritability. DNA methylation and chromatin assembly states, as well as miRNAs, have been found involved in many complex phenotypes (Trerotola *et al.*, 2015). As described in the Introduction, many studies have assessed the role of epigenetics in osteoporosis involvement, contributing to the unmasking of further heritability (Letarouilly *et al.*, 2018; Michou, 2018; Vrtačnik *et al.*, 2014).

After identifying a *locus* statistically associated with a complex trait, the main challenge resides in the identification and characterization of the causal variant responsible for the association signal. The association can be due to a direct causal relationship between the trait and the genotyped variant or due to an indirect association, in which the genotyped variant is in LD with the truly causal variant. For this reason, it is necessary to prove its functionality, as well as to replicate the association in an independent population (Gallagher & Chen-Plotkin, 2018). Achieving confidence in the determination of causality between a gene or variant and a disease is a complicated task that requires various types of supportive data. In addition, the challenge also resides in evaluating the functional impact of the variants in the same large scale as they are being discovered in association studies (Visscher *et al.*, 2017).

In Chapter 1 of this thesis, we replicated and dissected a GWAS signal in *C7ORF76* to identify the causal variants. In addition, we functionally assessed the associated variants, as well as the regulatory elements within this *locus* to shed light on the association.

Both replication and functional analyses are also crucial in determining pathogenicity of mutations identified in sequencing-based studies (e.g. WES) aiming at elucidating the

genetic causes of a disease. In Chapter 2 of this thesis, we identified and functionally characterized a mutation in *GGPS1* involved in AFF.

These aspects will be discussed in the following sections.

#### **4. Replication of identified variants**

Replication is fundamental when studying the genetic background for complex and Mendelian diseases. Variants related to complex traits are identified by statistical associations. Therefore, replication in an independent sample is crucial to confirm that the positive result found is real and not a spurious result or false positive. Many association studies use a two-stage design: a discovery stage in which a high number of variants are genotyped and a replication stage where only those variants found significantly associated in the discovery stage are assessed in a second independent sample. In addition, independent replication studies are also carried out in which regions previously found associated with a complex trait of interest are studied in other samples, sometimes of different ethnicity.

Frequently, the association results are discordant and do not replicate. Some possible reasons are differences in ethnic composition of samples, a high phenotypic variability, small sample size, differences in environmental factors or epistasis (Greene *et al.*, 2009). This has not been the case of the *C7ORF76* locus, which was found associated in many GWASs and meta-analyses assessing BMD and OF in different skeletal sites, ethnicities genders and ages (Duncan *et al.*, 2011; Estrada *et al.*, 2012; Kemp *et al.*, 2017; Medina-Gomez *et al.*, 2018; Morris *et al.*, 2019; Rivadeneira *et al.*, 2009; Trajanoska *et al.*, 2018; Zhang *et al.*, 2014; Zheng *et al.*, 2015). In Article 1 (Chapter 1), we also replicated this association with LS- and FN-BMD and OF in our Spanish cohort of postmenopausal women. Specifically, the two SNPs we found associated (rs10085588 and rs4342521) were previously found associated with LS- and FN-BMD and OF in other GWASs, although they were not the lead SNPs (Rivadeneira *et al.*, 2009; Zheng *et al.*, 2015). As already mentioned, we failed to find association in the rest of the variants interrogated, although some of them and others in this locus not tested by us have been found associated with BMD and OF elsewhere.

In monogenic diseases, replication is a strong evidence of pathogenicity. Although our results and others' suggest that N-BP-associated AFF is not a monogenic disease, replication of our findings would have given more compelling evidence of causality.

Unfortunately, we did not find any mutation in *GGPS1* in the 3 unrelated patients, nor in other patients we have analysed lately. Consistently, Peris *et al.* (2019) also analysed the presence of mutations and polymorphisms in the *GGPS1* gene by Sanger sequencing in 17 women with N-BP-associated AFFs and no mutation was found. Furthermore, we did not replicate any of the mutations reported in the last years (Funck-Brentano *et al.*, 2017; Lau *et al.*, 2017; Pérez-Núñez *et al.*, 2015). Instead, we did identify a mutation in *CYP1A1* in one unrelated patient of our small cohort, in addition to the mutation present in the 3 sisters. Moreover, in the same study by Peris *et al.* (2019), mutations in *CYP1A1* were reported in 2 patients with N-BP-associated AFF after a glucocorticoid-induced osteoporosis. These results make *CYP1A1* a good candidate gene for AFF and functional analyses of the mutations are underway. Besides, because AFF incidence differs between ethnicities, it would also be interesting to corroborate these findings in other populations.

## 5. Functional studies of variants and candidate genes

Functional analyses are fundamental for determining causality or pathogenicity of genetic variants. Furthermore, they are also important to understand the molecular pathophysiology underlying a disease and to further apply this knowledge to therapeutic approaches.

Several layers of evidence are needed and are discussed below. Depending on the nature of the variants, the functional approaches will differ and may include *in silico* computational analyses, *in vitro* or *in vivo* analyses, using cellular and animal systems.

### Definition and annotation of functional elements in the genome

When searching for causative genetic variants and characterizing them, it is crucial to know (*a priori* or *a posteriori*) their location and in which kind of functional element of the genome they lie: protein-coding exon, untranslated transcribed region (UTR), intron, enhancer, promoter, etc. The first step of variant characterization is the *in silico* analysis of their genomic location since, in the end, it will determine the kind of approaches used to characterize them. Therefore, the definition and large-scale annotation of these elements is pivotal. In this line, the ENCODE Project made a huge effort to comprehensively annotate all the functional sequences in the human genome, using a variety of assays and methods (The ENCODE Project Consortium, 2012).

The concept of “gene” has evolved and become more complex since it was first proposed in 1909 by Johannsen, who coined the term to denote an abstract “unit of inheritance”, without specific material attributes. Afterwards and successively, it designated a dimensionless point on a chromosome, a linear segment within a chromosome and, at the early 1960s, a discrete sequence on a DNA molecule that encodes a polypeptide chain. In the late 1970s, the discovery of introns (sequences interrupting the coding sequence) further modified the definition of a gene (Portin & Wilkins, 2017).

During the first 50 years, some characteristics of the gene were described, such as hereditary transmission, genetic recombination, mutation and gene function. More recently, the advances in molecular genetics, and especially projects such as the ENCODE or FANTOM projects (The ENCODE Project Consortium, 2012; The FANTOM Consortium & RIKEN Genome Exploration Research Group, 2005), have uncovered many other characteristics that add more complexity to the concept of “gene”. For example, a single gene can produce more than one mRNA by the means of alternative splicing or alternative promoters, transcription start sites (TSSs) or polyA sites, generating overlapping transcripts (Raabe & Brosius, 2015). In addition, boundaries to transcription are far from clear, leading in some cases to gene fusions and chimeric transcripts (Parra *et al.*, 2006). All in all, it seems that the human genome is comprehensively transcribed from both DNA strands. Furthermore, many genes encode for a diversity of noncoding RNA molecules that are not translated to proteins but also exert some biological function, such as long non-coding RNAs (lncRNAs), microRNAs (miRNAs), transfer RNAs (tRNAs) or PIWI-associated small RNAs. Actually, there are more genes encoding regulatory RNAs than proteins in the human genome and have a role in the regulation of epigenetics processes, differentiation and development. lncRNAs are >200 nucleotides in length, can be intronic, antisense or intergenic and are dynamically expressed in a range of differentiating systems (reviewed in Morris & Mattick, 2014). Altogether, this leads to the need to redefine the concept of “gene”. Therefore, recently, Portin & Wilkins (2017) proposed a comprehensive molecular definition: “A gene is a DNA sequence (whose component segments do not necessarily need to be physically contiguous) that specifies one or more sequence-related RNAs/proteins that are both evoked by genetic regulatory networks (GRNs) and participate as elements in GRNs, often with indirect effects, or as outputs of GRNs, the latter yielding more direct phenotypic effects”.

The evolution of the definition of “gene” is a clear example of how concepts are linked to the techniques that allow their characterization and, at the same time, how the

availability of an accurate definition is crucial for identifying and annotating new elements (in this case, new genes). Several approaches to identify new genes are used, either experimental or computational. Importantly, sequencing technologies have boosted this area of research. For example, the sequencing and mapping of RNAs have allowed the identification of transcribed regions of the genome. The GTEx (Genotype-Tissue Expression) Project represents the most comprehensive study of tissue-specific gene expression (53 non-diseased tissue sites) to date and has allowed the annotation of different transcripts in different cell types, as well as the correlation of gene expression with genetic variation (eQTLs; The GTEx Consortium, 2013). In the computational side, the identification of most genes in sequenced genomes is based either on their homology to other known genes, or the statistically significant signature of a gene sequence, namely *ab initio* algorithm-based gene prediction (Mudge & Harrow, 2016).

In this thesis (Articles 1 and 2 from Chapter 1), we explored the *C7ORF76* genomic region at 7q21.3, which was repeatedly found to be associated with BMD and OF in many GWASs. The *C7ORF76* gene is an uncharacterized gene of unknown function without expression data in GTEx. During the last years, the annotation of this gene has changed. First, it was annotated as an independent gene with the names *C7ORF76* or *FLJ42280*. It was described by Ota *et al.* (2004), in a large-scale attempt to characterize all the full-length cDNAs in different human tissues. Then, it was described as an alternative downstream transcript of *SHFM1*, a gene involved in split hand and foot malformation type 1. Moreover, the exons of *C7ORF76* were included in several long transcripts coming from the principal *SHFM1* promoter. Currently, in the GRCh38 genome assembly, it is annotated as transcript variant 6 of gene *SEM1* (encoding a 26S proteasome complex subunit), the new name for *SHFM1*. Although RefSeq labels it as a curated gene, surprisingly few data have been gathered, the hypothetical protein has not been characterized and the real function of *C7ORF76* remains unknown. In addition, epigenetic marks do not clearly show either the presence of a promoter or gene body signature in the putative exons. Our attempts to detect gene expression in several cell types such as primary osteoblasts, Saos-2, HeLa, SH-SY5Y and lymphocytes, have not succeed. Considering that it seems that the genome is a continuum of constitutive transcription, sometimes without evident biological function, it is not clear the exact role of this putative gene. All of this, together with the identification of regulatory elements in this region, suggest that *C7ORF76* might not be the gene underlying the association with BMD and OF found in GWASs, or at least not in the cell types analysed here. Another example with an opposite result is that of Kou *et al.* (2011), who performed a GWAS on OF in Japanese population and found a SNP associated within a new *in silico* predicted

gene. Using several experimental approaches, they identified a promoter and two transcripts, one of which ubiquitously expressed in various tissues, including bone. By protein motif prediction, they identified a signal peptide and a formiminotransferase domain in its N-terminal and named the gene *FONG*.

An example of the importance of characterizing and accurately annotating the genes for biomedical research is WES, in which protein-coding exons are captured by hybridization to a set of probes. Notably, capture kits have been changing and improving over the time to include new annotated genes and to better cover all the coding regions of the genome.

More and more, genetic variants within regulatory regions have been identified as causative for both Mendelian and complex diseases, highlighting the importance of such regions (Ma *et al.*, 2015). Indeed, the vast majority of variants identified by GWASs lie in the noncoding genome, often close to DNase I hypersensitivity sites, indicative of regulatory potential, complicating their functional assessment (Maurano *et al.*, 2012). The 7q21.3 genomic region studied in Chapter 1 is a good example of it, since the variants identified in many GWASs and meta-analysis are repeatedly located in the non-coding genome. Therefore, the proper definition and annotation of these elements is crucial to understand the molecular bases of diseases.

Promoters and enhancers are non-coding regulatory elements of the genome that control gene expression. They are defined by a set of characteristics, comprising histone marks, DNA methylation, TF binding, chromatin accessibility and conservation.

Historically, promoters and enhancers have been considered as two distinct classes of regulatory elements. Promoters are defined as DNA sequences able to recruit RNA polymerase II and that regulate and initiate transcription at proximal TSSs. In recent years, they have been identified mainly by mapping an epigenetic signature (e.g. low H3K4me1:H3K4me3 ratio, H3K27ac) and by sequencing the 5' ends of RNAs (Andersson *et al.*, 2015).

Enhancers are genomic elements that regulate transcription of distantly located genes by binding to promoters. They are described to be bound by TFs and to function in an orientation-, position- and distance-independent manner. Enhancers tightly control gene expression in a cell-type and spatiotemporally specific manner and one gene can be regulated by multiple enhancers with overlapping or differing activities (*IHH* or *DLX5/6* are examples of genes regulated by many enhancers; Ong & Corces, 2011; Will *et al.*, 2017). Such condition-specific regulation requires a higher-order chromatin architecture

that places enhancers in close 3D proximity with the target promoters (Schoenfelder & Fraser, 2019).

Enhancers have been identified mainly by gene reporter assays and by enrichment of marks such as H3K27ac, a high H3K4me1:H3K4me3 ratio, H2A.Z, binding of Mediator, P300 and TFs, and DNase I hypersensitivity (Coppola *et al.*, 2016). However, not all the putative enhancers mapped by genome-wide assays (e.g. P300 or H3K27ac chromatin immunoprecipitation followed by sequencing [ChIP-seq]) function as enhancers *in vivo* (Catarino & Stark, 2018). Actually, only a 59% of mouse candidate enhancers and a 52% of human candidate enhancers tested in large-scale *in vivo* enhancer activity assays drive reporter gene expression in some tissues (Visel *et al.*, 2007). Such was the case of the *DLX5/6* enhancer within the *C7ORF76* locus (namely eDlx#18) studied in Article 2 of Chapter 1, identified by comparative genome analysis as a non-coding evolutionary highly conserved region and tested for *in vivo* enhancer activity by reporter gene assay in zebrafish and mouse (Birnbaum *et al.*, 2012). In these transgenic assays it was described to be active in branchial arches of zebrafish embryos (72h post-fertilization) and mouse embryos (embryonic day 11.5). In addition, eDlx#18 was marked as an active enhancer in osteoblasts by ENCODE data. In Article 2, we further characterized it during embryonic development and in a bone context. In Article 1 (Chapter 1), we also annotated 3 more putative regulatory regions (based on publicly available epigenetic marks, TF binding and open chromatin data) containing several variants identified in the re-sequencing stage. One of them, which we named UPE was located 4kb upstream of the *C7ORF76* TSS and contained one of the variants (rs10085588) we found associated with LS-, FN-BMD and OF in the BARCOS cohort. We further characterized the UPE, showing that it has regulatory activity and probably acts as an enhancer for the lncRNA gene *LOC100506136* or the neighbouring genes *SHFM1* and *SLC25A13*.

Lately, the historical distinction between enhancers and promoters has been challenged, since enhancers and promoters have been shown to have similar behavioural characteristics, adding more arguments to the similar histone modifications and structural properties they have (Andersson *et al.*, 2015; Kim & Shiekhattar, 2015). It has been described that some promoters can exert distal-acting enhancer functions (Diao *et al.*, 2017; Diao *et al.*, 2017) and that both promoters and active enhancers can bind RNA polymerase II and initiate transcription, (Andersson *et al.*, 2014; Chen *et al.*, 2017). However, the properties of the RNAs they produce may differ substantially. On the one side, promoters of protein-coding genes produce mRNAs that are generally multiexonic, highly abundant, polyadenylated and translated. On the other side, enhancers produce

noncoding RNAs, called eRNAs, that are generally shorter (0.1-1 kb), unspliced, non-polyadenylated, less stable, low in abundance and retained in the nucleus. In addition, they are often bidirectionally transcribed and the functional role of eRNAs remains unclear (Mikhaylichenko *et al.*, 2018). In this line, we demonstrated that both UPE and eDlx#18 were transcribed in bone cells.

#### Computational prediction of variant effects

One approach to characterize a variant is the *in silico* prediction of its effect. Regarding coding variants, the genetic code allows a rapid translation of the genetic variant into the protein effect, in terms of amino acid change. Missense variants can affect the protein structure and function to a different extent depending on the location of the variant within the protein and the type of change. For example, an amino acid change in the active site of an enzyme or in a protein-protein interacting domain will be highly deleterious and a change from a neutral amino acid to a charged one (e.g. Ala to Asp) will be more deleterious than a change to another neutral amino acid (e.g. Ala to Gly). Several computational predictors, such as SIFT (Kumar *et al.*, 2009), Polyphen (Adzhubei *et al.*, 2010) and MutationTaster (Schwarz *et al.*, 2014), use this information, as well as conservation and other parameters, to give a score of deleteriousness used to prioritize variants. Indels might affect the protein sequence and structure if they are not multiple of 3, causing a frameshift.

In our WES analysis of the 3 sisters (Chapter 2), we used these tools to prioritize the list of variants shared by them. The GGPPS p.Asp188Tyr mutation was the most conserved one, with the most deleterious scores in SIFT, Polyphen and MutationTaster. The mutation elicits the change of an acidic amino acid (aspartate) to an uncharged aromatic residue (tyrosine) close to the active site of the enzyme. Specifically, it is located in the highly conserved second aspartate-rich region, which is involved in the binding to the substrate through a Mg<sup>2+</sup> salt bridge, predicting a severe impairment of enzyme activity. In addition, an  $\alpha$ -helix secondary structure is also predicted to be disrupted by the mutation. To confirm all these predictions, we performed experimental functional analyses (see below).

As for non-coding variants, other kind predictions can be made in order to assess their functional effects. Predictors of miRNAs or TF binding and conservation scores are available and have been used in Article 1 (Chapter 1) to prioritize the variants for association studies, as well as to characterize the associated variants afterwards. None of the variants assessed was predicted to affect miRNAs binding. The minor A allele of

the rs10085588 SNP, located within UPE, was predicted to bind to HDAC2 more probably than the G allele, possibly explaining the results observed in the reporter gene assays (see next sub-section of this Discussion).

#### Experimental assessment of regulatory variants: reporter assays and eQTLs

As mentioned, the approaches used for assessing the functionality of variants depend on the characteristics of the variant. Apart from determining the molecular functions of the regulatory variants and elements, the genes affected by them have to be investigated, as well as the relationship between these genes and disease risk (Gallagher & Chen-Plotkin, 2018).

Regulatory-element activity is highly cell-type specific and, therefore, should be analyzed in the context of the disease-relevant tissue (Albert & Kruglyak, 2015). In the osteoporosis field, OBs and osteoblastic cell lines are the main cell types used for functional analyses, although other cell types have also been used, including OCs and MSCs. However, the availability of human bone material is limited and is not incorporated into consortium-based, large-scale studies, such as GTEx.

Many studies have shown an overlap between regulatory variants and eQTLs and others have integrated data from eQTLs and GWASs to identify complex trait-associated variants that influence transcript levels of putative target genes (Albert & Kruglyak, 2015; Zhu *et al.*, 2016). In this thesis, we assessed the *cis*-eQTL role of the variants within the *C7ORF76* locus associated with BMD and OF, as well as of the variants lying within the eDlx#18 enhancer (Articles 1 and 2, Chapter 1). We did so in a set of primary hOBs (n=45) and detected a nominal association between the minor alleles of the 3 associated SNPs and a decreased *SCL25A13* gene expression, as well as a trend for association with decreased expression of *SHFM1*. We also found that the rs10238953 SNP within eDlx#18 showed a nominal association with *DLX6* gene expression and a trend with *SHFM1*. The other eDlx#18 SNP (rs4613908) also showed a trend with *SHFM1* and *SLC25A13*. These results pointed at these genes as the putative target genes for these regulatory variants, although this strategy does not demonstrate how the SNP influences gene expression and other approaches are necessary to confirm its mechanistic relevance (see next section). In addition, the reduced sample size might have precluded the identification of some smaller effects on neighbouring genes.

In parallel, the putative regulatory elements have to be tested for their transactivation capability. Reporter gene assays in cellular systems are widely used to evaluate the regulatory activity of candidate DNA fragments since they are easy and fast to perform.

In addition, the different alleles of a SNP can also be tested to assess a possible allele-specific activity of the putative regulatory element. MPRA have been developed, in which thousands of variants can be tested in a single experiment (Inoue & Ahituv, 2015).

In Chapter 1 of this thesis, we performed luciferase reporter assays in Saos-2 cells to evaluate the transactivation capacity of the UPE element and showed that, indeed, it was able to activate transcription of the luciferase gene. Interestingly, the minor allele (A) of rs10085588 abolished luciferase activation, being consistent with eQTL analyses. In addition, we also performed luciferase assays with eDlx#18 in Saos-2 cells to assess its enhancer activity in an osteoblastic context. As mentioned, this enhancer was identified by transgenic assay in zebrafish and mouse embryos and was found active in branchial arches (Birnbaum *et al.*, 2012). However, due to the enhancer marks detected in osteoblasts in ENCODE data, we wanted to verify that eDlx#18 was also able to activate transcription in a bone context and we showed that the core of the enhancer harboured regulatory activity in Saos-2 cells.

Yet, gene reporter assays have some limitations. First, small differences in expression may be difficult to distinguish statistically due to the transcriptional noise that this technique can display and the differences in reporter activity that can result from small unavoidable differences in the molar amounts of each plasmid transfected. Most importantly, cell culture-based reporter assays do not test the transcriptional function of a candidate region/variant in its native genomic context, but in the context of plasmid DNA. Therefore, intricate relationships between DNA, histones, TFs, etc. are not considered. In light of this, gene editing represents a more physiologically-relevant method to confirm the function of a regulatory region/variant of interest.

### 3D genome organization

Since the genome is folded and spatially organized into the nucleus of the cells, functional assessment of variants and genomic regions requires the understanding of the three-dimensional implications.

Microscopy-based and chromosome conformation capture (i.e. 3C, 4C, 5C, Hi-C and Capture-C) techniques revealed that the genome is hierarchically compartmentalized into domains, such as topologically associating domains (TADs). TADs are relatively cell-type invariant domains with preferential intradomain interactions (Dixon *et al.*, 2012; Nora *et al.*, 2012). Importantly, it has been shown that this 3D organization plays an important role in genome function, including transcriptional control of genes by facilitating interactions between gene promoters and enhancers (Bonev & Cavalli, 2016;

Schoenfelder & Fraser, 2019). In addition, disruption of TADs by structural rearrangements or boundary deletions has been shown to result in diseases (Lupiáñez *et al.*, 2016).

One of the main challenges of regulatory variant characterization is the discovery of the target gene, that mediates the underlying biological mechanisms of the association. In articles from Chapter 1 (Articles 1 and 2) we took advantage of 4C-seq to associate regulatory elements with their putative target genes or interactors in different samples (medulla-derived mesenchymal stem cells, the human fetal osteoblasts 1.19 and the Saos-2 cell lines, and developing humeri from E14.5 mice, the latter only for eDlx#18). We described that both UPE and eDlx#18 interacted with several sites within the TAD they belong to, including a lncRNA (*LOC100506136*) in the case of UPE and the *DLX5/6* region and promoter in the case of eDlx#18. Many other studies have used similar approaches to map enhancer-promoter interactions, as well as to further assess eQTL analysis results, in order to shed light on the causal mechanisms of diseases (Javierre *et al.*, 2016). In the field of osteoporosis, and similarly to our work, two studies from the same group used publicly-available Hi-C data to establish the target genes for some regulatory SNPs identified in GWASs and the authors found that the SNPs were *cis*-eQTLs for the genes they interacted with (Chen *et al.*, 2018; Zhu *et al.*, 2018). Nonetheless, in the majority of cases it is mechanistically unclear how variants in putative regulatory elements contribute to gene expression changes that may underlie complex diseases. In Article 1 we described that the minor allele of rs10085588 (A) drastically reduced luciferase expression, functioned as *cis*-eQTL for *SLC25A13* (which also interacted modestly with UPE) and *in silico* analysis predicted the loss of binding of the histone deacetylase HDAC2, which could explain the other results. Notably, we did not evaluate the *LOC100506136* in our eQTL analysis and, hence, we could not determine if rs10085588 within UPE would influence its expression.

Apart from the interactions already mentioned, we also found that both UPE and eDlx#18 interacted with other enhancers described in the nearby region. These results are consistent with a large “spatial regulome”, as it has been observed in other occasions (Beagrie *et al.*, 2017; Will *et al.*, 2017), where multiple regulatory elements converge and may have functional redundancy or specific spatiotemporal activities that confer robustness, precision and flexibility to gene expression. Besides, eDlx#18 showed interaction with the neighbouring *SLC25A13* and *SHFM1* genes, correlating with the *cis*-eQTL analyses and suggesting that different genes within the same TAD might be co-regulated by the same group of enhancers. Supporting this notion, it has been observed that genes within the same TADs share coordinated gene expression profiles (Gómez-

Marín *et al.*, 2015; Nora *et al.*, 2012). Putative transcriptional co-regulation would imply that sequence variants affecting the activity of regulatory elements might ultimately affect multiple genes within a TAD.

### Assessment of enzymatic mutations: *GGPS1* as an example

Enzymes are proteins that catalyze chemical reactions within organisms. When a mutation is found which modifies an enzyme, the straightforward test to perform is an evaluation of the enzymatic activity of the WT and the mutant forms.

The most conserved and putatively deleterious mutation shared by the 3 AFF sisters was a missense mutation in the *GGPS1* gene (p.Asp188Tyr), encoding GGPPS, an enzyme of the mevalonate pathway that catalyzes the reaction just downstream of the reaction targeted by N-BPs. GGPPS catalyzes the synthesis of GGPP from FPP and IPP and is also known to be inhibited by N-BPs, although to a much lesser extent (Kavanagh *et al.*, 2006a). It functions as a homohexamer, in which each monomer binds 3 Mg<sup>2+</sup> ions.

As mentioned in a previous section of this Discussion, the p.Asp188Tyr mutation was predicted to severely impair enzyme activity, as well as to disrupt its secondary structure. Therefore, we performed *in vitro* functional analyses of the p.Asp188Tyr mutation, which showed a severe reduction in enzyme activity, consistent with previous work showing that disruption of the second aspartate-rich region results in an almost complete loss of enzyme activity (Kavanagh *et al.*, 2006b). Furthermore, we showed mild oligomerization defects, since the homohexameric conformation of the enzyme is destabilized by the mutation. Interestingly, another work on GGPPS p.Asp188Tyr mutation has been recently published (Lisnyansky *et al.*, 2018) and the authors also observed a decreased catalytic activity of the mutated GGPPS, consistent with our results. In addition, they showed that it is unable to support cross-species complementation. In contrast, they only observed the hexameric conformation of the enzyme in crystallographic experiments, although they saw a slight break of the tertiary symmetry, as well as a lower thermal stability of the mutated enzyme. In addition, the new tyrosine residue sterically interfered with substrate binding.

### Cellular and animal models

When characterizing genetic variants, and due to the high complexity of the genome, it is important to test their functionality in a biological context. Therefore, cellular and animal models are crucial. Nowadays, with the advent of CRISPR-Cas9 technologies,

genome editing is easier, faster and more efficient and has facilitated the study of the underlying biology pinpointed by the genetic discoveries.

Several human bone cell models are currently used to mimic bone metabolic processes and to allow for easy, fast, non-invasive, reproducible and representative analyses of many molecular processes that take place during differentiation and activity of OBs, OCs and OCys (Kartsogiannis & Ng, 2004). In this line, we generated *GGPS1* knock-down cell models by using shRNAs in OBs and OCs precursor cells properly stimulated for their differentiation. In OBs, *GGPS1* KD produced a dramatic reduction of bone nodule formation and mineralization and a reduced expression of the typical osteoblastic markers: osteocalcin, osterix and RANKL. In OCs, *GGPS1* KD led to an increase of the OC number but with a lower resorption activity. Nonetheless, these cell models do not fully replicate what happens in clinical cases, where the mutation is present in heterozygosis (so, the WT *GGPS1* expression is less reduced) and we have not tested if the mutant protein might have a dominant negative effect upon the WT form. In addition, *in vivo* OBs and OCs are tightly coupled in bone remodelling and, thus, their responses to *GGPS1* depletion or mutation should be intimately associated. To answer all these questions, an animal model would provide more compelling evidence. In any case, we propose that the excessive inhibition of OC activity achieved through the combination of the mutation and the N-BPs may lead to a reduced bone remodelling and toughness, which may increase AFF susceptibility (see also section 6 of this Discussion).

Animal models are a powerful tool to study the biological mechanisms underlying bone diseases, identifying novel pathways regulating bone development, maintenance and resilience, as well as to develop novel treatments. The complexity and dynamism of the bone microenvironment only occurs in live animals and cannot be modelled *ex vivo*. In addition, the use of animal models allows us to finely control experimental variables (Ackert-Bicknell & Karasik, 2013). The mouse is used extensively because of its high degree of genome similarity, numerous techniques for genetic manipulation, capacity to mimic human multifactorial disease phenotypes and easiness of manipulation. In addition, physiologically and anatomically, mice and humans are remarkably similar. In the bone field, many of the key molecules that regulate bone have the same functions in humans and mice and the human genetic disorders causing abnormalities of bone are recapitulated in mice. Similarly, endocrine and metabolic control of bone is preserved in mice (Bonucci & Ballanti, 2014). However, mice do not present the same bone structure (that is Haversian systems) that is present in humans and other mammals and, hence, they are not good models for studying certain aspects of bone biology. Other animal models, such as zebrafish or dogs, have also been used in bone biology studies (Karasik

*et al.*, 2016). Considering the particular characteristics of each animal model, caution should be taken in translating the findings to human populations.

In this thesis, we generated a mouse model knocked-out for eDlx#18 by CRISPR/Cas9. The deletion in homozygosis caused a reduced survival in mouse embryos, as well as decreased expression of *Dlx5* in otic vesicles and branchial arches of E11.5 embryos and several morphological and ossification defects in E17.5 embryos, recapitulating, to some extent, the phenotype seen in *Dlx5*<sup>-/-</sup>, *Dlx6*<sup>-/-</sup> and *Dlx5/6*<sup>+/-</sup> mice. Considering that many individual enhancer deletions cause subtle phenotypes at the organismal level (Osterwalder *et al.*, 2018), the effects of eDlx#18 knock-out in mice were somehow surprising. However, other studies have shown that enhancer relationships in landscapes might vary including, for instance, redundancy, exclusivity or synergy, and that might differ depending on location and time of gene regulation in order to confer robustness and flexibility to the gene expression repertoire (Long *et al.*, 2016; Will *et al.*, 2017). The perinatal lethality precluded the evaluation of the bone properties in adult mice, therefore, the putative involvement of eDlx#18 in BMD determination and bone strength could not be determined. Further work is advisable to better elucidate the role of eDlx#18 in the bone context, such as an OB-specific depletion of the enhancer.

Regarding AFF, animal models may provide insights into the pathophysiological mechanisms at cellular and tissue levels. In addition, they would be useful to evaluate the involvement of N-BPs in the phenotype, as well as their interaction with the genetic background (see next section). Animal models have been extensively used for drug evaluation in osteoporosis therapy, including N-BPs, since they allow a preliminary assessment of safety and tolerability (Russell *et al.*, 2008). N-BPs have been studied in mice, rabbits, dogs, pigs, sheep and monkey but one of the most used models are beagle dogs (Allen & Burr, 2007; Burr *et al.*, 2015), that recapitulate human response to drugs. Burr *et al.* (2015) studied the time-dependent effects of oral alendronate on dogs and reported a significant reduction in rib bone toughness with longer exposure to alendronate, showing a time-dependent deterioration of cortical bone. Yet, animals do not appear to fracture spontaneously, even following prolonged treatment with high-doses of N-BPs. Instead, they can be used to study alterations in the structural and material properties of the bone.

In our study, generating a heterozygous mutation in *GGPS1* by CRISPR-Cas9 and generating an animal model, would have been the best option to mimic the 3 sisters' condition, assessing the effect of the p.Asp188Tyr mutation and analyzing the impact of N-BPs in this context. A *GGPS1* homozygous knock-out mouse was generated by the

IMKC/IMPC and preweaning lethality was reported. In addition, heterozygous knock-out, showed impaired glucose tolerance (only in male mice) and cataracts. Vertebrae and digits showed a normal number, size and morphology. These phenotypes show that GGPPS and protein geranylgeranylation are important for other processes and it would be interesting to treat heterozygous knock-out mice with N-BPs to see their effect. Information about these aspects in our patients was not available.

## **6. Epistasis and gene-environment interaction: pharmacogenetics and personalized medicine**

Genetic interaction or epistasis has been demonstrated to be a common phenomenon, in which there exists a highly interconnected network of genes with related molecular functions that contribute to and regulate various cellular processes, resulting in a particular phenotype (Mackay, 2014). Epistasis is involved in complex traits, in which different variants contribute non-additively to the phenotype. It also occurs in Mendelian traits, where a major gene responsible for the phenotype can be influenced by modifier genes, giving a certain phenotypic variation. The interaction can be synergistic or suppressive depending on the sense of the modulation (Cole *et al.*, 2017).

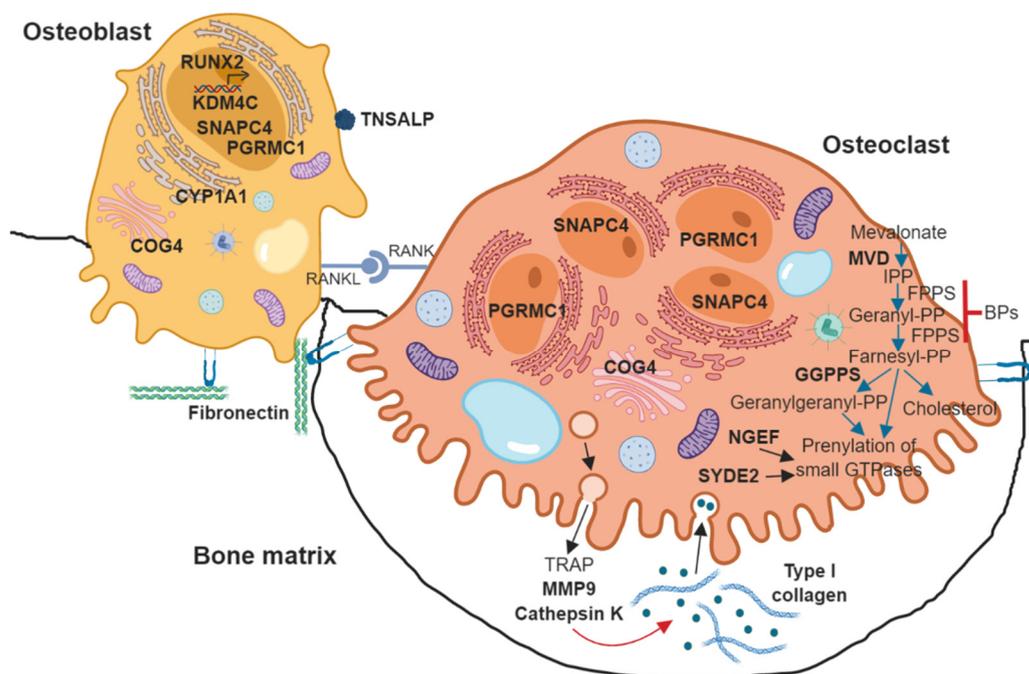
Some studies have identified gene-gene interactions related to osteoporosis. For instance, Yang *et al.* (2013) identified one interacting gene pair (*RBMS3* and *ZNF516*) with consistently significant effects on hip BMD while Wang *et al.* (2018) identified a significant interaction between a SNP in *P2X7R* and a polymorphism in *ESR1* increasing osteoporosis risk in Chinese postmenopausal women.

Studies aiming at delineating the genetic architecture of AFF (including Articles 3 and 4 from Chapter 2 of this thesis) have uncovered genetic heterogeneity, since each individual patient presented different specific genetic variants. In addition, in our WES-based genetic analysis of the 3 sisters, we found several rare mutations in heterozygosis, consistent with a dominant inheritance. Then, we functionally studied the most conserved one, which appeared to be in a gene of the same metabolic pathway where BPs act. However, other variants found in the 3 sisters might potentially contribute to the phenotype, although functional studies or replication are needed to prove their involvement in the pathology. Notably, the *CYP1A1* gene was found mutated in other AFF patients (see section 4 of this Discussion). We speculate that our results may support a genetic architecture model in which accumulation of susceptibility variants,

with different effect sizes, may give rise to the development of bone changes that lead to AFF (Figure 11).

The susceptibility variants may have an additive effect, but also a probable interaction among them would contribute to the phenotype. To further characterize the putative functional relationships among all the mutated genes and identify the pathways involved, we generated a gene/protein interaction network (Annex to Chapter 2) that would help to identify other putative candidate genes to AFF and better understand the molecular mechanisms underlying their pathophysiology. An example is the interaction of *NGEF*, a guanine nucleotide exchange factor, with ephrins and ephrin receptors, involved in OB-OC coupling (Kim & Koh, 2019). One limitation of the pathway is that it is based on curated interactions from literature and those genes that have not been studied will be under-represented in the network. In addition, any interesting interaction and its functionality in the context of N-BP-associated AFF should be assessed experimentally in a relevant system.

All in all, addressing gene-gene interactions is crucial to characterize N-BP-associated AFF, a trait involving complex pharmacologic mechanisms.



**Figure 11.** Proteins found mutated in AFF patients (in bold) in the context of bone tissue

Moreover, in our speculative model, the genetic background predisposing to AFF would also negatively interact with N-BPs. The effect of genetic variants on the response to treatment is the study object of pharmacogenetics (Meyer, 2004). Generally, variants

in drug transporters, drug targets or metabolizing enzymes contribute to drug efficacy and safety, involving drug pharmacokinetics and pharmacodynamics (Sadée, 2013). For example, polymorphisms in two genes encoding liver enzymes responsible for the metabolism of many commonly used drugs (i.e. *CYP2C19* and *CYP2D6*) regulate their expression and activity and, as a consequence, how people will respond to drugs. Pharmacogenetics opens the door to a personalized or precision medicine to predict drug treatment outcome and optimize treatment and dosing regimens.

Concerning osteoporosis, currently available drugs are non-effective in all treated patients and/or induce adverse effects (see Table 3, Introduction). Since all of anti-osteoporotic drugs act at the level of bone cells, interfering in different pathways that regulate their differentiation, proliferation and activity, it is likely that variants in genes involved in these pathways or their regulation are responsible for inter-individual differences in drug response. To identify such genetic variation, it is necessary to understand the molecular mechanism of drug action. For example, a CRISPRi-mediated genome-wide screening identified *SLC37A3* as a protein that, forming a complex with *ATRAID*, was required for the intracellular trafficking of N-BPs to their molecular target (i.e. *FPPS*) in mammalian cells (Yu *et al.*, 2018). However, the high complexity of bone metabolism hinders the identification of variants. In addition, a large number of variants and their interaction may be involved, exerting various degrees of influence. To date, few studies are available on the pharmacogenetics of osteoporosis, which investigated some major osteoporosis candidate genes (e.g. *VDR*, *ESR1*, *ESR2* and *COL1A1*) in relation to anti-resorptive drug responses evaluated in terms of BMD and bone turnover markers variation (reviewed in Marini & Brandi, 2014).

Regarding N-BPs, polymorphisms in genes related to the Wnt pathway and other bone candidate genes have been associated with the response to N-BPs. For instance, Palomba *et al.* (2003) showed that postmenopausal women on alendronate carrying the *b* allele of the *VDR* Bsm-I polymorphism had a greater increase in LS-BMD than carriers of the *B* allele. Wang *et al.* (2018b) showed that common SNPs in *SOST* contribute to the bone response at femoral neck to alendronate treatment in Chinese osteoporotic and osteopenic women. In addition, studies exploring the influence of genetic variants in the mevalonate pathway on N-BPs response have been carried out. Polymorphisms of *FPPS*, the main target of N-BPs, were found to be associated with the response to N-BP therapy, as measured by changes in BMD and bone turnover markers, in Caucasian women (Marini *et al.*, 2008; Olmos *et al.*, 2012). However, this association was not observed in Asian women, suggesting an ethnic-specificity (Choi *et al.*, 2010; Yi *et al.*, 2014). Interestingly, Choi *et al.* (2010) also reported the influence of a polymorphism in

the promoter region of the *GGPS1* gene on N-BPs response, assessed as changes in LS- and FN-BMD after 1 year treatment, in Korean osteoporotic women. Women with a homozygous deletion (c.-8188delA) showed a lower BMD improvement and had a 7-fold higher risk of non-response to BPs. In our study, this deletion was present with the same frequency in patients and controls (data not shown).

Although some studies have demonstrated the interaction of N-BPs with variants in genes of the mevalonate pathway, we did not demonstrate such an interaction with the *GGPS1* mutation in our study (Article 4, Chapter 2). Nevertheless, considering that N-BPs are able to bind and inhibit GGPPS (Kavanagh *et al.*, 2006a), Lisnyansky *et al.* (2018) determined the affinity of zoledronate for GGPPS (both WT and p.Asp188Tyr), showing that the mutant exhibited a 3-fold reduction in the binding affinity of zoledronate, although it could still be inhibited by the drug. In any case, we speculate that the effect of N-BPs will be compounded due to the loss of both farnesylation and geranylgeranylation of proteins. Moreover, the heterozygous AFF patients would keep a residual GGPPS activity, sufficient to support physiologic cellular function in normal conditions but reduced below a crucial threshold upon N-BP treatment, leading to impaired bone remodelling and increased AFF susceptibility.

Another interesting gene found mutated in the 3 sisters and in 1 unrelated patient in our study, as well as in Peris *et al.* (2019) was *CYP1A1*, which encodes a member of the cytochrome P450 superfamily of enzymes involved in the metabolism of drugs and xenobiotics. *CYP1A1* is an aryl hydrocarbon hydroxylase and its potential endogenous substrates include eicosanoids and steroid hormones, such as 17  $\beta$ -estradiol, estrone and vitamin D (Zhou *et al.*, 2009). *CYP1A1* has been involved in bone biology and osteoporosis, as supported by Napoli *et al.* (2005). They showed that a c.4887C>A (p.Thr461Asn) polymorphism was related to a significantly higher degree of estrogen catabolism and lower femoral BMD in postmenopausal women. Therefore, *CYP1A1* would appear to be another potential susceptibility gene for AFF, but the exact mechanism is open to speculation. Functional studies are underway to study the effect of the mutations in *CYP1A1* and their putative interaction with N-BPs in bone cells.

A pharmacogenomic study investigating the role of genetic variation in the risk of developing an adverse effect, such as AFF, in response to N-BPs was performed recently (Kharazmi *et al.*, 2019). They used two sets of controls: general population-based controls and matched N-BP-treated controls to exclude the associations due the underlying diseases that lead to drug prescription. Using the second set of controls they did not find any common variant significantly associated with N-BP-related AFF.

Probably, the low statistical power of this GWAS precluded the identification of putative variants. In addition, considering the very low incidence of N-BP-associated AFF, our hypothesis has been that, rather than common variants, rare variants are more probable to be responsible for the genetic predisposition to AFF.

Apart from N-BPs, other environmental factors might interact with genetic variants in AFF determination, some of which affecting also osteoporosis risk (e.g. diet or physical activity). For example, calcium and vitamin D interactions with genetic variants in several genes on bone phenotypes have been identified (reviewed in Ackert-Bicknell & Karasik, 2013)

All in all, we propose a mechanism in which susceptibility variants from different pathways, together with the interaction among them and with N-BPs, as well as other possible comorbid conditions, give rise to AFF, in what we might call “the perfect storm”. In this context, it is highly important to replicate and characterize the variants involved in AFF, as well as to identify new variants, in order to develop prediction tests to detect at-risk individuals and decide which patients are suitable for being treated with BPs with no risk of this side effect.

## **7. Future perspectives**

Osteoporosis is a major worldwide public health concern, which a huge economic and social impact. Currently, treatments are not totally effective in all treated patients or skeletal sites. In addition, there are concerns about side effects and long-term safety, which have reduced the adherence of patients to them.

Knowledge derived from large-scale sequencing efforts and comprehensive GWASs is increasing, providing novel insight on the key regulatory mechanisms that control skeletal physiology. Such abundance of genetic discoveries has to be functionally interpreted in order to be able to translate them into clinical applications, including a better definition and diagnosis of diseases at a molecular level, the identification of biomarkers or the development of new therapies. In this line, it has been shown that selecting genetically supported targets could double the success rate of new drugs (Nelson *et al.*, 2015).

As illustrated in this thesis, an aspect of biomedical genetics research that is still challenging is the functional characterization of the genetic discoveries, including the integration of different bioinformatic and experimental strategies. In this sense,

systematic large-scale functional screening is demanding and collaborative strategies are arising.

In the field of N-BP-associated AFF, the main prospect is the better definition of the disease at genetic and molecular levels, allowing a better understanding of the pathophysiological mechanisms underlying the disease and providing new insights for future development of novel risk assessment tools. Therefore, a collaborative effort to collect and share samples from multiple unrelated individuals with such a rare phenotype is needed.

# **CONCLUSIONS**



---

Conclusions from the characterization of the *C7ORF76* locus:

- Two SNPs (rs10085588 and rs4342521) located upstream of the *C7ORF76* gene are significantly associated with LS-BMD and nominally associated with OF in the BARCOS cohort of Spanish postmenopausal women.
- The minor allele of the two associated variants (rs10085588 and rs4342521), together with the minor allele of the SNP rs4727338, are nominally associated with a decreased expression of the proximal neighbouring gene *SLC25A13* in human primary osteoblasts, therefore acting as *cis*-eQTLs.
- A conserved putative regulatory element located upstream of *C7ORF76* (UPE), containing rs10085588, has been identified based on DNase hypersensitivity signal, enhancer marks (H3K4me1 and H3K27ac) and TF binding data.
- Reporter gene assays have shown that UPE is able to activate luciferase expression with a strong allele-dependent effect, in which the minor allele (A) abolishes luciferase activity. This is evidence that rs10085588 is itself functional.
- The UPE is transcribed in several cell types, including Saos-2 cells.
- Chromatin conformation capture experiments in different osteoblastic cells have shown that UPE interacts with several sites within the TAD where it belongs and nowhere else in the genome: it interacts with a lncRNA (*LOC100506136*) and different tissue-specific enhancers.
- A 292-bp central part of eDlx#18 harbours regulatory activity in an osteoblastic context since it is able to activate luciferase gene expression in Saos-2 cells.
- eDlx#18 is transcribed in several cell types, including Saos-2 cells.
- eDlx#18 interacts with several sites within the TAD it belongs to and nowhere else in the genome, as detected by 4C-seq in human osteoblastic cells and in mouse E14.5 developing humeri. It interacts with the *DLX5/6* region and with many of the other *DLX5/6* tissue-specific enhancers, as well as with UPE.
- The *DLX5* promoter interacts with several sites within the TAD and nowhere else in the genome, as detected by 4C-seq in the same samples. High interaction levels have been detected with the tissue-specific enhancers in the region, among which eDlx#18.
- An SNP within eDlx#18 (rs10238953) is nominally associated with *DLX6* gene expression in human primary osteoblasts, acting as *cis*-eQTL.
- The homozygous deletion of eDlx#18 by CRISPR/Cas9 results in a decreased survival in mouse embryos and in a reduced *Dlx5* expression in otic vesicle and

## CONCLUSIONS

---

branchial arches of E11.5 mouse embryos, as seen by whole-mount RNA *in situ* hybridization

- In the E17.5 mouse embryo skeleton, the homozygous deletion of *eDlx#18* results in smaller dentary, a deficient ossification of the supraoccipital bone (with 50% of exencephaly), vertebral bodies, sternum and pelvic bones and minor affectations in the ribs. No limb malformations are observed.

Conclusions from the identification and characterization of genetic susceptibility to N-BP-associated AFF:

- Thirty-seven rare non-synonymous mutations in 34 genes shared by 3 sisters with N-BP-associated AFF have been identified by WES, including novel missense mutations in *GGPS1* and in *CYP1A1*.
- The *BRAT1* and *CYP1A1* genes, mutated in the 3 sisters, present another mutation in another unrelated AFF patient (AFU3 and AFU1, respectively)
- Two unrelated patients present mutations in candidate genes: AFU2 has a mutation in *MVD* and AFU3 has a mutation in *MMP9* and a mutation in *RUNX2*.
- Pathway enrichment analysis of the mutated genes has shown an enrichment of the isoprenoid biosynthetic pathway, including *GGPS1*, *MVD* and *CYP1A1*.
- The p.Asp188Tyr mutation in *GGPS1* results in a severe reduction in enzyme activity, exhibiting a 5.7% of the WT activity. It also results in oligomerization defects, as seen by molecular exclusion chromatography.
- The depletion of *GGPS1* in MC3T3-E1 cells by shRNA produces a strong mineralization reduction and a decreased expression of bone formation/bone turnover markers *BGLAP* (osteocalcin), *SP7* (osterix) and *TNFSF11* (*RANKL*), while *RUNX2*, *ALPL*, *MEPE* and *PHEX* are unaffected.
- The depletion of *GGPS1* in RAW 264.7 macrophages, followed by osteoclastogenic differentiation, leads to an increased osteoclast formation. However, these osteoclasts seem to have lower resorption activity.

## **REFERENCES**



- Ackert-Bicknell, C. L., & Karasik, D. 2013. Impact of the environment on the skeleton: Is it modulated by genetic factors? *Curr. Osteoporos. Rep.* 11(3):219–228.
- Adler, R. A. 2018. MANAGEMENT OF ENDOCRINE DISEASE: Atypical femoral fractures: risks and benefits of long-term treatment of osteoporosis with anti-resorptive therapy. *Eur. J. Endocrinol.* 178(3):R81–R87.
- Adler, R. A., El-Hajj Fuleihan, G., Bauer, D. C., Camacho, P. M., Clarke, B. L., Clines, G. A., *et al.* 2016. Managing Osteoporosis in Patients on Long-Term Bisphosphonate Treatment: Report of a Task Force of the American Society for Bone and Mineral Research. *J. Bone Miner. Res.* 31(1):16–35.
- Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., *et al.* 2010. A method and server for predicting damaging missense mutations. *Nat. Methods* .
- Agueda, L., Bustamante, M., Jurado, S., Garcia-Giralt, N., Ciria, M., Saló, G., *et al.* 2008. A haplotype-based analysis of the LRP5 gene in relation to osteoporosis phenotypes in Spanish postmenopausal women. *J. Bone Miner. Res.* 23(12):1954–1963.
- Akkawi, I., & Zmerly, H. 2018. Osteoporosis: Current Concepts. *Joints* 6(2):122–127.
- Albert, F. W., & Kruglyak, L. 2015. The role of regulatory variation in complex traits and disease. *Nat. Rev. Genet.* 16(4):197–212.
- Alexandre, C., & Vico, L. 2011. Pathophysiology of bone loss in disuse osteoporosis. *Jt. Bone Spine* 78(6):572–576.
- Allen, M. R., & Burr, D. B. 2007. Three years of alendronate treatment results in similar levels of vertebral microdamage as after one year of treatment. *J. Bone Miner. Res.* 22(11):1759–1765.
- Alonso, N., Estrada, K., Albagha, O. M. E., Herrera, L., Reppe, S., Olstad, O. K., *et al.* 2018. Identification of a novel locus on chromosome 2q13, which predisposes to clinical vertebral fractures independently of bone density. *Ann. Rheum. Dis.* 77(3):378–385.
- Alonso, N., & Ralston, S. H. 2014. Unveiling the mysteries of the genetics of osteoporosis. *J. Endocrinol. Invest.* 37(10):925–934.
- Alonso, N., Soares, D. C., V McCloskey, E., Summers, G. D., Ralston, S. H., & Gregson, C. L. 2015. Atypical femoral fracture in osteoporosis pseudoglioma syndrome associated with two novel compound heterozygous mutations in LRP5. *J. Bone Miner. Res.* 30(4):615–620.
- Amin, D., Cornell, S. A., Gustafson, S. K., Needle, S. J., Ullrich, J. W., Bilder, G. E., *et al.* 1992. Bisphosphonates used for the treatment of bone disorders inhibit squalene synthase and cholesterol biosynthesis. *J. Lipid Res.* 33(11):1657–1663.
- Amit, S., Shekhar, A., Vivek, M., Shekhar, S., & Biren, N. 2010. Fixation of Subtrochanteric Fractures in Two Patients with Osteopetrosis Using a Distal Femoral Locking Compression Plate of the Contralateral Side. *Eur. J. Trauma Emerg. Surg.* 36(3):263–269.

## REFERENCES

---

- Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., *et al.* 2014. An atlas of active enhancers across human cell types and tissues. *Nature* 507(7493):455–461.
- Andersson, R., Sandelin, A., & Danko, C. G. 2015. A unified architecture of transcriptional regulatory elements. *Trends Genet.* 31(8):426–433.
- Angers, S., & Moon, R. T. 2009. Proximal events in Wnt signal transduction. *Nat. Rev. Mol. Cell Biol.* 10(7):468–477.
- Arboleya, L., & Castañeda, S. 2013. Osteoimmunology: the study of the relationship between the immune system and bone tissue. *Reumatol. Clínica* 9(5):303–315.
- Arden, N. K., Baker, J., Hogg, C., Baan, K., & Spector, T. D. 1996. The Heritability of Bone Mineral Density, Ultrasound of the Calcaneus and Hip Axis Length: A Study of Postmenopausal Twins. *J. Bone Miner. Res.* 11(4):530–534.
- Awasthi, H., Mani, D., Singh, D., & Gupta, A. 2018. The underlying pathophysiology and therapeutic approaches for osteoporosis. *Med. Res. Rev.* 38(6):2024–2057.
- Baird, D. A., Evans, D. S., Kamanu, F. K., Gregory, J. S., Saunders, F. R., Giuraniuc, C. V., *et al.* 2019. Identification of Novel Loci Associated With Hip Shape: A Meta-Analysis of Genomewide Association Studies. *J. Bone Miner. Res.* 34(2):241–251.
- Bandeira, L., & Bilezikian, J. P. 2017. Novel Therapies for Postmenopausal Osteoporosis. *Endocrinol. Metab. Clin. North Am.* 46(1):207–219.
- Barbaric, I., Perry, M. J., Dear, T. N., Rodrigues Da Costa, A., Salopek, D., Marusic, A., *et al.* 2008. An ENU-induced mutation in the *Ankrd11* gene results in an osteopenia-like phenotype in the mouse mutant Yoda. *Physiol. Genomics* 32(3):311–321.
- Barnett, I. J., Lee, S., & Lin, X. 2013. Detecting Rare Variant Effects Using Extreme Phenotype Sampling in Sequencing Association Studies. *Genet. Epidemiol.* 37(2):142–151.
- Baron, R., & Kneissel, M. 2013. WNT signaling in bone homeostasis and disease: from human mutations to treatments. *Nat. Med.* 19(2):179–192.
- Bassett, J. H. D., Gogakos, A., White, J. K., Evans, H., Jacques, R. M., van der Spek, A. H., *et al.* 2012. Rapid-Throughput Skeletal Phenotyping of 100 Knockout Mice Identifies 9 New Genes That Determine Bone Strength. *PLoS Genet.* 8(8):e1002858.
- Beagrie, R. A., Scialdone, A., Schueler, M., Kraemer, D. C. A., Chotalia, M., Xie, S. Q., *et al.* 2017. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* 543(7646):519–524.
- Bellavia, D., De Luca, A., Carina, V., Costa, V., Raimondi, L., Salamanna, F., *et al.* 2019. Deregulated miRNAs in bone health: Epigenetic roles in osteoporosis. *Bone* 122:52–75.
- Bellido, T., & Plotkin, L. I. 2011. Novel actions of bisphosphonates in bone: Preservation of osteoblast and osteocyte viability. *Bone* 49(1):50–55.
- Berendsen, A. D., & Olsen, B. R. 2015. Bone development. *Bone* 80:14–18.

- Bhattacharyya, T., Jha, S., Wang, H., Kastner, D. L., & Remmers, E. F. 2016. Hypophosphatasia and the risk of atypical femur fractures: a case-control study. *BMC Musculoskelet. Disord.* 17:332.
- Bilezikian, J. P., Hattersley, G., Fitzpatrick, L. A., Harris, A. G., Shevroja, E., Banks, K., *et al.* 2018. Abaloparatide-SC improves trabecular microarchitecture as assessed by trabecular bone score (TBS): a 24-week randomized clinical trial. *Osteoporos. Int.* 29(2):323–328.
- Birmingham, P., & McHale, K. 2008. Case Reports: Treatment of Subtrochanteric and Ipsilateral Femoral Neck Fractures in an Adult with Osteopetrosis. *Clin. Orthop. Relat. Res.* 466(8):2002–2008.
- Bivi, N., Bereszczyk, J. Z., Romanello, M., Zeef, L. A. H., Delneri, D., Quadrioglio, F., *et al.* 2009. Transcriptome and Proteome Analysis of Osteocytes Treated with Nitrogen-Containing Bisphosphonates. *J. Proteome Res.* 8(3):1131–1142.
- Black, D. M., Abrahamsen, B., Bouxsein, M. L., Einhorn, T., & Napoli, N. 2019. Atypical Femur Fractures – Review of epidemiology, relationship to bisphosphonates, prevention and clinical management. *Endocr. Rev.* 40(2):333–368.
- Black, D. M., Delmas, P. D., Eastell, R., Reid, I. R., Boonen, S., Cauley, J. A., *et al.* 2007. Once-Yearly Zoledronic Acid for Treatment of Postmenopausal Osteoporosis. *N. Engl. J. Med.* 356(18):1809–1822.
- Blanco-Gómez, A., Castillo-Lluva, S., del Mar Sáez-Freire, M., Hontecillas-Prieto, L., Mao, J. H., Castellanos-Martín, A., *et al.* 2016. Missing heritability of complex diseases: Enlightenment by genetic variants from intermediate phenotypes. *BioEssays* 38(7):664–673.
- Bliuc, D., Nguyen, N. D., Milch, V. E., Nguyen, T. V., Eisman, J. A., & Center, J. R. 2009. Mortality Risk Associated With Low-Trauma Osteoporotic Fracture and Subsequent Fracture in Men and Women. *JAMA - J. Am. Med. Assoc.* 301(5):513–521.
- Bone, H. G., Bolognese, M. A., Yuen, C. K., Kendler, D. L., Miller, P. D., Yang, Y.-C., *et al.* 2011. Effects of Denosumab Treatment and Discontinuation on Bone Mineral Density and Bone Turnover Markers in Postmenopausal Women with Low Bone Mass. *J. Clin. Endocrinol. Metab.* 96(4):972–980.
- Bone, H. G., Hosking, D., Devogelaer, J.-P., Tucci, J. R., Emkey, R. D., Tonino, R. P., *et al.* 2004. Ten Years' Experience with Alendronate for Osteoporosis in Postmenopausal Women. *N. Engl. J. Med.* 350(12):1189–1199.
- Bone, H. G., Wagman, R. B., Brandi, M. L., Brown, J. P., Chapurlat, R., Cummings, S. R., *et al.* 2017. 10 years of denosumab treatment in postmenopausal women with osteoporosis: results from the phase 3 randomised FREEDOM trial and open-label extension. *Lancet Diabetes Endocrinol.* 5(17):513–523.
- Bonev, B., & Cavalli, G. 2016. Organization and function of the 3D genome. *Nat. Rev. Genet.* 17(11):661–678.
- Bonjour, J.-P., Kraenzlin, M., Lévassieur, R., Warren, M., & Whiting, S. 2013. Dairy in Adulthood: From Foods to Nutrient Interactions on Bone and Skeletal Muscle Health. *J. Am. Coll. Nutr.* 32(4):251–263.
- Bonucci, E., & Ballanti, P. 2014. Osteoporosis — Bone Remodeling and Animal Models. *Toxicol. Pathol.* 42(6):957–969.

## REFERENCES

---

- Boudin, E., Fijalkowski, I., Hendrickx, G., & Van Hul, W. 2016. Genetic control of bone mass. *Mol. Cell. Endocrinol.* 432:3–13.
- Bouxsein, M. L., & Karasik, D. 2006. Bone Geometry and Skeletal Fragility. *Curr. Osteoporos. Rep.* 4(2):49–56.
- Boyce, B. F., & Xing, L. 2007. Biology of RANK, RANKL, and osteoprotegerin. *Arthritis Res. Ther.* 9(Suppl1):S1.
- Briot, K., & Roux, C. 2015. Glucocorticoid-induced osteoporosis. *RMD Open* 1(1):e000014.
- Brown, J. P. 2017. Antiresorptives: Safety Concerns—Clinical Perspective. *Toxicol. Pathol.* 45(7):859–863.
- Burr, D. B., Liu, Z., & Allen, M. R. 2015. Duration-dependent effects of clinically relevant oral alendronate doses on cortical bone toughness in beagle dogs. *Bone* 71:58–62.
- Capulli, M., Paone, R., & Rucci, N. 2014. Osteoblast and osteocyte: Games without frontiers. *Arch. Biochem. Biophys.* 561:3–12.
- Catarino, R. R., & Stark, A. 2018. Assessing sufficiency and necessity of enhancer activities for gene expression and the mechanisms of transcription activation. *Genes Dev.* 32(3–4):202–223.
- Cauley, J. A. 2011. Defining Ethnic and Racial Differences in Osteoporosis and Fragility Fractures. *Clin. Orthop. Relat. Res.* 469(7):1891–1899.
- Cauley, J. A. 2017. Osteoporosis: fracture epidemiology update 2016. *Curr. Opin. Rheumatol.* 29(2):150–156.
- Chakravorty, S., & Hegde, M. 2017. Gene and Variant Annotation for Mendelian Disorders in the Era of Advanced Sequencing Technologies. *Annu. Rev. Genomics Hum. Genet.* 18:229–256.
- Chen, H., Du, G., Song, X., & Li, L. 2017. Non-coding Transcripts from Enhancers: New Insights into Enhancer Activity and Gene Expression Regulation. *Genomics, Proteomics Bioinforma.* 15(3):201–207.
- Chen, L.-R., Ko, N.-Y., & Chen, K.-H. 2019. Medical Treatment for Osteoporosis: From Molecular to Clinical Opinions. *Int. J. Mol. Sci.* 20(9):E2213.
- Chen, X.-F., Zhu, D.-L., Yang, M., Hu, W.-X., Duan, Y.-Y., Lu, B.-J., *et al.* 2018. An Osteoporosis Risk SNP at 1p36.12 Acts as an Allele-Specific Enhancer to Modulate LINC00339 Expression via Long-Range Loop Formation. *Am. J. Hum. Genet.* 102(5):776–793.
- Cheraghi, Z., Doosti-Irani, A., Almasi-Hashiani, A., Baigi, V., Mansournia, N., Etminan, M., *et al.* 2019. The effect of alcohol on osteoporosis: A systematic review and meta-analysis. *Drug Alcohol Depend.* 197:197–202.
- Chesi, A., Mitchell, J. A., Kalkwarf, H. J., Bradfield, J. P., Lappe, J. M., Cousminer, D. L., *et al.* 2017. A Genomewide Association Study Identifies Two Sex-Specific Loci, at SPTB and IZUMO3, Influencing Pediatric Bone Mineral Density at Multiple Skeletal Sites. *J. Bone Miner. Res.* 32(6):1274–1281.

- Chesi, A., Mitchell, J. A., Kalkwarf, H. J., Bradfield, J. P., Lappe, J. M., McCormack, S. E., *et al.* 2015. A trans-ethnic genome-wide association study identifies gender-specific loci influencing pediatric aBMD and BMC at the distal radius. *Hum. Mol. Genet.* 24(17):5053–5059.
- Chesi, A., Wagley, Y., Johnson, M. E., Manduchi, E., Su, C., Lu, S., *et al.* 2019. Genome-scale Capture C promoter interaction analysis implicates novel effector genes at GWAS loci for bone mineral density. *Nat. Commun.* 10(1):1260.
- Chesnut III, C. H., Skag, A., Christiansen, C., Recker, R., Stakkestad, J. A., Hoiseth, A., *et al.* 2004. Effects of Oral Ibandronate Administered Daily or Intermittently on Fracture Risk in Postmenopausal Osteoporosis. *J. Bone Miner. Res.* 19(8):1241–1249.
- Choi, H. J., Choi, J. Y., Cho, S. W., Kang, D., Han, K. O., Kim, S. W., *et al.* 2010. Genetic polymorphism of geranylgeranyl diphosphate synthase (GGSP1) predicts bone density response to bisphosphonate therapy in Korean women. *Yonsei Med. J.* 51(2):231–238.
- Christodoulou, S., Goula, T., Ververidis, A., & Drosos, G. 2013. Vitamin D and Bone Disease. *Biomed Res. Int.* 2013:396541.
- Clark, G. R., & Duncan, E. L. 2015. The genetics of osteoporosis. *Br. Med. Bull.* 113(1):73–81.
- Clarke, B. 2008. Normal bone anatomy and physiology. *Clin. J. Am. Soc. Nephrol.* 3(Suppl 3):S131–S139.
- Cohen-Kfir, E., Artsi, H., Levin, A., Abramowitz, E., Bajayo, A., Gurt, I., *et al.* 2011. Sirt1 is a regulator of bone mass and a repressor of Sost encoding for sclerostin, a bone formation inhibitor. *Endocrinology* 152(12):4514–4524.
- Cole, B. S., Hall, M. A., Urbanowicz, R. J., Gilbert-Diamond, D., & Moore, J. H. 2017. Analysis of Gene-Gene Interactions. *Curr. Protoc. Hum. Genet.* 95(1):1.14.1–1.14.10.
- Compston, J. 2011. Pathophysiology of atypical femoral fractures and osteonecrosis of the jaw. *Osteoporos. Int.* 22(12):2951–2961.
- Compston, J., Cooper, A., Cooper, C., Gittoes, N., Gregson, C., Harvey, N., *et al.* 2017. UK clinical guideline for the prevention and treatment of osteoporosis. *Arch. Osteoporos.* 12(1):43.
- Compston, J. E., McClung, M. R., & Leslie, W. D. 2019. Osteoporosis. *Lancet* 393(10169):364–376.
- Coppola, C. J., Ramaker, R. C., & Mendenhall, E. M. 2016. Identification and function of enhancers in the human genome. *Hum. Mol. Genet.* 25(R2):R190–R197.
- Cordell, H. J., & Clayton, D. G. 2005. Genetic association studies. *Lancet* 366(9491):1121–1131.
- Cosman, F., Crittenden, D. B., Adachi, J. D., Binkley, N., Czerwinski, E., Ferrari, S., *et al.* 2016. Romosozumab Treatment in Postmenopausal Women with Osteoporosis. *N. Engl. J. Med.* 375(16):1532–1543.
- Cosman, F., Crittenden, D. B., Ferrari, S., Lewiecki, E. M., Jaller-Raad, J., Zerbini, C., *et al.* 2018. Romosozumab FRAME Study: A Post Hoc Analysis of the Role of

## REFERENCES

---

- Regional Background Fracture Risk on Nonvertebral Fracture Outcome. *J. Bone Miner. Res.* 33(8):1407–1416.
- Cosman, F., Nieves, J. W., & Dempster, D. W. 2017. Treatment Sequence Matters: Anabolic and Antiresorptive Therapy for Osteoporosis. *J. Bone Miner. Res.* 32(2):198–202.
- Coxon, F. P., & Rogers, M. J. 2003. The Role of Prenylated Small GTP-Binding Proteins in the Regulation of Osteoclast Function. *Calcif. Tissue Int.* 72(1):80–84.
- Crandall, C. J., Newberry, S. J., Diamant, A., Lim, Y.-W., Gellad, W. F., Booth, M. J., *et al.* 2014. Comparative effectiveness of pharmacologic treatments to prevent fractures: an updated systematic review. *Ann. Intern. Med.* 161(10):711–723.
- Cummings, S. R., Ferrari, S., Eastell, R., Gilchrist, N., Beck Jensen, J.-E., McClung, M., *et al.* 2018. Vertebral Fractures After Discontinuation of Denosumab: A Post Hoc Analysis of the Randomized Placebo-Controlled FREEDOM Trial and Its Extension. *J. Bone Miner. Res.* 33(2):190–198.
- Cummings, S. R., San Martin, J., McClung, M., Siris, E. S., Eastell, R., Reid, I. R., *et al.* 2009. Denosumab for Prevention of Fractures in Postmenopausal Women with Osteoporosis. *N. Engl. J. Med.* 361(8):756–765.
- Curtis, E. M., van der Velde, R., Moon, R. J., van den Bergh, J. P. W., Geusens, P., de Vries, F., *et al.* 2016. Epidemiology of fractures in the United Kingdom 1988 – 2012: Variation with age, sex, geography, ethnicity and socioeconomic status. *Bone* 87:19–26.
- Dao, L. T. M., Galindo-Albarrán, A. O., Castro-Mondragon, J. A., Andrieu-Soler, C., Medina-Rivera, A., Souaid, C., *et al.* 2017. Genome-wide characterization of mammalian promoters with distal enhancer functions. *Nat. Genet.* 49(7):1073–1081.
- De-Ugarte, L., Caro-Molina, E., Rodríguez-Sanz, M., Garcíá-Pérez, M. A., Olmos, J. M., Sosa-Henríquez, M., *et al.* 2017. SNPs in bone-related miRNAs are associated with the osteoporotic phenotype. *Sci. Rep.* 7(1):516.
- De Laet, C., Kanis, J. A., Odén, A., Johanson, H., Johnell, O., Delmas, P., *et al.* 2005. Body mass index as a predictor of fracture risk: A meta-analysis. *Osteoporos. Int.* 16(11):1330–1338.
- de Vrieze, E., van Kessel, M. A. H. J., Peters, H. M., Spanings, F. A. T., Filk, G., & Metz, J. R. 2014. Prednisolone induces osteoporosis-like phenotype in regenerating zebrafish scales. *Osteoporos. Int.* 25(2):567–578.
- Dell, R., & Greene, D. 2018. A proposal for an atypical femur fracture treatment and prevention clinical practice guideline. *Osteoporos. Int.* 29(6):1277–1283.
- Demissie, S., Dupuis, J., Cupples, L. A., Beck, T. J., Kiel, D. P., & Karasik, D. 2007. Proximal hip geometry is linked to several chromosomal regions: Genome-wide linkage results from the Framingham Osteoporosis Study. *Bone* 40(3):743–750.
- Deng, F.-Y., Zhao, L.-J., Pei, Y.-F., Sha, B.-Y., Liu, X.-G., Yan, H., *et al.* 2010. Genome-wide copy number variation association study suggested VPS13B gene for osteoporosis in Caucasians. *Osteoporos. Int.* 21(4):579–587.
- Dhanwal, D. K., Dennison, E. M., Harvey, N. C., & Cooper, C. 2011. Epidemiology of hip fracture: Worldwide geographic variation. *Indian J. Orthop.* 45(1):15–22.

- Diao, Y., Fang, R., Li, B., Meng, Z., Yu, J., Qiu, Y., *et al.* 2017. A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. *Nat. Methods* 14(6):629–635.
- Diez-Perez, A., Güerri, R., Nogues, X., Cáceres, E., Peña, M. J., Mellibovsky, L., *et al.* 2010. Microindentation for In Vivo Measurement of Bone Tissue Mechanical Properties in Humans. *J. Bone Miner. Res.* 25(8):1877–1885.
- Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., *et al.* 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485(7398):376–380.
- Dole, N. S., & Delany, A. M. 2016. MicroRNA variants as genetic determinants of bone mass. *Bone* 84:57–68.
- Donnelly, E., Meredith, D. S., Nguyen, J. T., Gladnick, B. P., Rebolledo, B. J., Shaffer, A. D., *et al.* 2012. Reduced cortical bone compositional heterogeneity with bisphosphonate treatment in postmenopausal women with intertrochanteric and subtrochanteric fractures. *J. Bone Miner. Res.* 27(3):672–678.
- Doshi, K. B., Hamrahian, A. H., & Licata, A. A. 2009. Teriparatide treatment in adult hypophosphatasia in a patient exposed to bisphosphonate: a case report. *Clin. Cases Miner. Bone Metab.* 6(3):266–269.
- Drake, M. T., Clarke, B. L., & Khosla, S. 2008. Bisphosphonates: mechanism of action and role in clinical practice. *Mayo Clin. Proc.* 83(9):1032–45.
- Duncan, E. L., & Brown, M. A. 2013. Genome-wide association studies. In R. V. Thakker, M. P. Whyte, J. A. Eisman, & T. Igarashi (Eds.), *Genet. Bone Biol. Skelet. Dis.* (pp. 93–100). Elsevier Inc.
- Duncan, E. L., Danoy, P., Kemp, J. P., Leo, P. J., McCloskey, E., Nicholson, G. C., *et al.* 2011. Genome-wide association study using extreme truncate selection identifies novel genes affecting bone mineral density and fracture risk. *PLoS Genet.* 7(4):e1001372.
- Dunford, J. E., Rogers, M. J., Ebetino, F. H., Phipps, R. J., & Coxon, F. P. 2006. Inhibition of protein prenylation by bisphosphonates causes sustained activation of Rac, Cdc42, and Rho GTPases. *J. Bone Miner. Res.* 21(5):684–694.
- Duong, L. T., Leung, A. T., & Langdahl, B. 2016. Cathepsin K Inhibition: A New Mechanism for the Treatment of Osteoporosis. *Calcif. Tissue Int.* 98(4):381–397.
- Eastell, R. 2017. Prevention and management of osteoporosis. *Medicine (Baltimore)*. 45(9):565–569.
- Eastell, R., & Walsh, J. S. 2017. Anabolic treatment for osteoporosis: teriparatide. *Clin. Cases Miner. Bone Metab.* 14(2):173–178.
- Emkey, G. R., & Epstein, S. 2014. Secondary osteoporosis: Pathophysiology & diagnosis. *Best Pract. Res. Clin. Endocrinol. Metab.* 28(6):911–935.
- Estrada, K., Styrkarsdottir, U., Evangelou, E., Hsu, Y. H., Duncan, E. L., Ntzani, E. E., *et al.* 2012. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat. Genet.* 44(5):491–501.
- Ettinger, B., Black, D. M., Mitlak, B. H., Knickerbocker, R. K., Nickelsen, T., Genant, H. K., *et al.* 1999. Reduction of Vertebral Fracture Risk in Postmenopausal Women

## REFERENCES

---

- With Osteoporosis Treated With Raloxifene: Results From a 3-Year Randomized Clinical Trial. *JAMA - J. Am. Med. Assoc.* 282(7):637–645.
- Ettinger, B., Burr, D. B., & Ritchie, R. O. 2013. Proposed pathogenesis for atypical femoral fractures: Lessons from materials research. *Bone* 55(2):495–500.
- Etxebarria-Foronda, I., & Carpintero, P. 2015. An atypical fracture in male patient with osteogenesis imperfecta. *Clin. Cases Miner. Bone Metab.* 12(3):278–281.
- Farr, J. N., & Khosla, S. 2015. Skeletal changes through the lifespan - From growth to senescence. *Nat. Rev. Endocrinol.* 11(9):513–521.
- Feng, Q., Zheng, S., & Zheng, J. 2018. The emerging role of microRNAs in bone remodeling and its therapeutic implications for osteoporosis. *Biosci. Rep.* 38(3):pii:BSR20180453.
- Florencio-Silva, R., Rodrigues, G., Sasso-Cerri, E., Simões, M. J., & Cerri, P. S. 2015. Biology of Bone Tissue: Structure, Function and Factors That Influence Bone Cells. *Biomed Res. Int.* 2015:421746.
- Forlino, A., & Marini, J. C. 2016. Osteogenesis imperfecta. *Lancet* 387(10028):1657–1671.
- Freudenthal, B., Logan, J., Sanger Institute Mouse Pipelines, Croucher, P. I., Williams, G. R., & Duncan Bassett, J. H. 2016. Rapid phenotyping of knockout mice to identify genetic determinants of bone strength. *J. Endocrinol.* 231(1):R31-46.
- Fuente, R., & Hernández, O. 2017. X-linked hypophosphatemia and growth. *Rev. Endocr. Metab. Disord.* 18(1):107–115.
- Funck-Brentano, T., Ostertag, A., Debais, F., Fardellone, P., Collet, C., Mornet, E., *et al.* 2017. Identification of a p.Arg708Gln variant in COL1A2 in atypical femoral fractures. *Jt. Bone Spine* 84(6):715–718.
- Gagnon, C., Sims, N. A., Mumm, S., McAuley, S. A., Jung, C., Poulton, I. J., *et al.* 2010. Lack of Sustained Response to Teriparatide in a Patient with Adult Hypophosphatasia. *J. Clin. Endocrinol. Metab.* 95(3):1007–1012.
- Gaj, T., Sirk, S. J., Shui, S., & Liu, J. 2016. Genome-Editing Technologies: Principles and Applications. *Cold Spring Harb. Perspect. Biol.* 8(12):a023754.
- Gallagher, J. C. 2007. Effect of early menopause on bone mineral density and fractures. *Menopause* 14(3):567–571.
- Gallagher, M. D., & Chen-Plotkin, A. S. 2018. The Post-GWAS Era: From Association to Function. *Am. J. Hum. Genet.* 102(5):717–730.
- Gatti, D., Adami, S., Viapiana, O., & Rossini, M. 2015. The use of bisphosphonates in women: when to use and when to stop. *Expert Opin. Pharmacother.* 16(16):2409–2421.
- Gedmintas, L., Solomon, D. H., & Kim, S. C. 2013. Bisphosphonates and risk of subtrochanteric, femoral shaft, and atypical femur fracture: A systematic review and meta-analysis. *J. Bone Miner. Res.* 28(8):1729–1737.
- Gentili, C., & Cancedda, R. 2009. Cartilage and Bone Extracellular Matrix. *Curr. Pharm. Des.* 15(12):1334–1348.

- Giger, E. V, Castagner, B., & Leroux, J.-C. 2013. Biomedical applications of bisphosphonates. *J. Control. Release* 167(2):175–188.
- Goldstein, D. B., Allen, A., Keebler, J., Margulies, E. H., Petrou, S., Petrovski, S., *et al.* 2013. Sequencing studies in human genetics: Design and interpretation. *Nat. Rev. Genet.* 14(7):460–470.
- Goldstein, J. L., & Brown, M. S. 1990. Regulation of the mevalonate pathway. *Nature* 343(6257):425–430.
- Gómez-Marín, C., Tena, J. J., Acemel, R. D., López-Mayorga, M., Naranjo, S., de la Calle-Mustienes, E., *et al.* 2015. Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *Proc. Natl. Acad. Sci. U. S. A.* 112(24):7542–7547.
- Greene, C. S., Penrod, N. M., Williams, S. M., & Moore, J. H. 2009. Failure to Replicate a Genetic Association May Provide Important Clues About Genetic Architecture. *PlosOne* 4(6):e5639.
- Gregson, C. L., Newell, F., Leo, P. J., Clark, G. R., Paternoster, L., Marshall, M., *et al.* 2018. Genome-wide association study of extreme high bone mass: Contribution of common genetic variation to extreme BMD phenotypes and potential novel BMD-associated genes. *Bone* 114:62–71.
- Grundberg, E., Kwan, T., Ge, B., Lam, K. C. L., Koka, V., Kindmark, A., *et al.* 2009. Population genomics in a disease targeted primary cell model. *Genome Res.* 19(11):1942–1952.
- Güerri-Fernández, R. C., Nogués, X., Quesada Gómez, J. M., Torres Del Pliego, E., Puig, L., Garcia-Giralt, N., *et al.* 2013. Microindentation for in vivo measurement of bone tissue material properties in atypical femoral fracture patients and controls. *J. Bone Miner. Res.* 28(1):162–168.
- Guo, Y., Dong, S. S., Chen, X. F., Jing, Y. A., Yang, M., Yan, H., *et al.* 2016. Integrating Epigenomic Elements and GWASs Identifies BDNF Gene Affecting Bone Mineral Density and Osteoporotic Fracture Risk. *Sci. Rep.* 6:30558.
- Guo, Y., Tan, L.-J., Lei, S.-F., Yang, T.-L., Chen, X.-D., Zhang, F., *et al.* 2010. Genome-wide association study identifies ALDH7A1 as a novel susceptibility gene for osteoporosis. *PLoS Genet.* 6(1):e1000806.
- Guo, Y., Zhao, L.-J., Shen, H., Guo, Y., & Deng, H.-W. 2005. Genetic and Environmental Correlations between Age at Menarche and Bone Mineral Density at Different Skeletal Sites. *Calcif. Tissue Int.* 77(6):356–360.
- Gusev, A., Bhatia, G., Zaitlen, N., Vilhjalmsón, B. J., Diogo, D., Stahl, E. A., *et al.* 2013. Quantifying Missing Heritability at Known GWAS Loci. *PLoS Genet.* 9(12):e1003993.
- Haentjens, P., Magaziner, J., Colón-Emeric, C. S., Vanderschueren, D., Milisen, K., Velkeniers, B., *et al.* 2010. Meta-analysis: Excess Mortality After Hip Fracture Among Older Women and Men. *Ann. Intern. Med.* 152(6):380–390.
- Hagen, J. E., Miller, A. N., Ott, S. M., Gardner, M., Morshed, S., Jeray, K., *et al.* 2014. Association of Atypical Femoral Fractures with Bisphosphonate Use by Patients with Varus Hip Geometry. *J. Bone Jt. Surg. - Am. Vol.* 96(22):1905–1909.

## REFERENCES

---

- Harada, S., & Rodan, G. A. 2003. Control of osteoblast function and regulation of bone mass. *Nature* 423(6937):349–355.
- Harper, E., Forde, H., Davenport, C., Rochfort, K. D., Smith, D., & Cummins, P. M. 2016. Vascular calcification in type-2 diabetes and cardiovascular disease: Integrative roles for OPG, RANKL and TRAIL. *Vascul. Pharmacol.* 82:30–40.
- Harris, S. T., Watts, N. B., Genant, H. K., McKeever, C. D., Hangartner, T., Keller, M., *et al.* 1999. Effects of Risedronate Treatment on Vertebral and Nonvertebral Fractures in Women With Postmenopausal Osteoporosis: A Randomized Controlled Trial. Vertebral Efficacy With Risedronate Therapy (VERT) Study Group. *JAMA - J. Am. Med. Assoc.* 282(14):1344–1352.
- Hashem, J., Krochak, R., Culbertson, M. D., Mileto, C., & Goodman, H. 2015. Atypical femur fractures in a patient with pycnodysostosis: a case report. *Osteoporos. Int.* 26(8):2209–2212.
- Hellman, L. M., & Fried, M. G. 2007. Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nat. Protoc.* 2(8):1849–1861.
- Hendrickx, G., Boudin, E., & Van Hul, W. 2015. A look behind the scenes: The risk and pathogenesis of primary osteoporosis. *Nat. Rev. Rheumatol.* 11(8):462–474.
- Hernlund, E., Svedbom, A., Ivergard, M., Compston, J., Cooper, C., Stenmark, J., *et al.* 2013. Osteoporosis in the European Union: medical management, epidemiology and economic burden A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). *Arch. Osteoporos.* 8:136.
- Holm, J., Eiken, P., Hyldstrup, L., & Jensen, J.-E. B. 2014. Atypical femoral fracture in an osteogenesis imperfecta patient successfully treated with teriparatide. *Endocr. Pract.* 20(10):e187-190.
- Hsu, Y.-H., Estrada, K., Evangelou, E., Ackert-Bicknell, C., Akesson, K., Beck, T., *et al.* 2019. Meta-Analysis of Genomewide Association Studies Reveals Genetic Variants for Hip Bone Geometry. *J. Bone Miner. Res.* e3698.
- Hsu, Y. H., Li, G., Liu, C. T., Brody, J. A., Karasik, D., Chou, W. C., *et al.* 2016. Targeted sequencing of genome wide significant loci associated with bone mineral density (BMD) reveals significant novel and rare variants: The Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) targeted sequencing study. *Hum. Mol. Genet.* 25(23):5234–5243.
- Hsu, Y. H., Zillikens, M. C., Wilson, S. G., Farber, C. R., Demissie, S., Soranzo, N., *et al.* 2010. An integration of genome-wide association study and gene expression profiling to prioritize the discovery of novel susceptibility loci for osteoporosis-related traits. *PLoS Genet.* 6(6):e1000977.
- Hunter, D., De Lange, M., Sneider, H., MacGregor, A. J., Swaminathan, R., Thakker, R. V., *et al.* 2001. Genetic Contribution to Bone Metabolism, Calcium Excretion, and Vitamin D and Parathyroid Hormone Regulation. *J. Bone Miner. Res.* 16(2):371–378.
- Hwang, J.-Y., Lee, S. H., Go, M. J., Kim, B.-J., Kou, I., Ikegawa, S., *et al.* 2013. Meta-analysis identifies a MECOM gene as a novel predisposing factor of osteoporotic fracture. *J. Med. Genet.* 50(4):212–219.

- Ikebuchi, Y., Aoki, S., Honma, M., Hayashi, M., Sugamori, Y., Khan, M., *et al.* 2018. Coupling of bone resorption and formation by RANKL reverse signalling. *Nature* 561(7722):195–200.
- Im, G.-I., & Lee, S.-H. 2015. Effect of Teriparatide on Healing of Atypical Femoral Fractures: A Systemic Review. *J. Bone Metab.* 22(4):183–189.
- Inoue, F., & Ahituv, N. 2015. Decoding enhancers using massively parallel reporter assays. *Genomics* 106(3):159–164.
- Ioannidis, J. P. A. 2005. Why Most Published Research Findings Are False. *PLoS Med.* 2(8):e124.
- Ioannidis, J. P. A., Ng, M. Y., Sham, P. C., Zintzaras, E., Lewis, C. M., Deng, H.-W., *et al.* 2007. Meta-Analysis of Genome-Wide Scans Provides Evidence for Sex- and Site-Specific Regulation of Bone Mass. *J. Bone Miner. Res.* 22(2):173–183.
- Ioannidis, J. P. A., Ralston, S. H., Bennett, S. T., Brandi, M. L., Grinberg, D., Karassa, F. B., *et al.* 2004. Differential Genetic Effects of ESR1 Gene Polymorphisms on Osteoporosis Outcomes. *JAMA - J. Am. Med. Assoc.* 292(17):2105–2114.
- Javierre, B. M., Burren, O. S., Wilder, S. P., Kreuzhuber, R., Hill, S. M., Sewitz, S., *et al.* 2016. Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. *Cell* 167(5):1369–1384.
- Jing, D., Hao, J., Shen, Y., Tang, G., Li, M.-L., Huang, S.-H., *et al.* 2015. The role of microRNAs in bone remodeling. *Int. J. Oral Sci.* 7(3):131–143.
- Johnell, O., Kanis, J. A., Oden, A., Johansson, H., De Laet, C., Delmas, P., *et al.* 2005. Predictive value of BMD for hip and other fractures. *J. Bone Miner. Res.* 20(7):1185–1194.
- Johnell, O., Kanis, J. A., Odén, A., Sernbo, I., Redlund-Johnell, I., Petterson, C., *et al.* 2004. Fracture risk following an osteoporotic fracture. *Osteoporos. Int.* 15(3):175–179.
- Jung, M., & Pfeifer, G. P. 2015. Aging and DNA methylation. *BMC Biol.* 13:7.
- Kang, G., Lin, D., Hakonarson, H., & Chen, J. 2012. Two-stage extreme phenotype sequencing design for discovering and testing common and rare genetic variants: Efficiency and power. *Hum. Hered.* 73(3):139–147.
- Kanis, J. A., Harvey, N. C., Johansson, H., Odén, A., Leslie, W. D., & McCloskey, E. V. 2017. FRAX Update. *J. Clin. Densitom.* 20(3):360–367.
- Kanis, J. A., Johansson, H., Oden, A., Johnell, O., De Laet, C., Eisman, J. A., *et al.* 2004a. A family history of fracture and fracture risk: a meta-analysis. *Bone* 35(5):1029–1037.
- Kanis, J. A., Johnell, O., De Laet, C., Johansson, H., Oden, A., Delmas, P., *et al.* 2004b. A meta-analysis of previous fracture and subsequent fracture risk. *Bone* 35(2):375–382.
- Kanis, J. A., Johnell, O., Oden, A., Johansson, H., De Laet, C., Eisman, J. A., *et al.* 2005. Smoking and fracture risk: a meta-analysis. *Osteoporos. Int.* 16(2):155–162.

## REFERENCES

---

- Kanis, J. A., McCloskey, E. V., Johansson, H., Cooper, C., Rizzoli, R., & Reginster, J.-Y. 2013. European guidance for the diagnosis and management of osteoporosis in postmenopausal women. *Osteoporos. Int.* 24(1):23–57.
- Kanis, J. A., Oden, A., Johnell, O., Johansson, H., De Laet, C., Brown, J., *et al.* 2007. The use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. *Osteoporos. Int.* 18(8):1033–1046.
- Karaplis, A. C. 2008. Embryonic Development of Bone and Regulation of Intramembranous and Endochondral Bone Formation. In J. P. Bilezikian, L. G. Raisz, & T. J. Martin (Eds.), *Princ. Bone Biol.* (3rd ed., pp. 53–84). Elsevier Inc.
- Karasik, D., Demissie, S., Zhou, Y., Lu, D., Broe, K. E., Bouxsein, M. L., *et al.* 2017. Heritability and Genetic Correlations for Bone Microarchitecture: The Framingham Study Families. *J. Bone Miner. Res.* 32(1):106–114.
- Karasik, D., Rivadeneira, F., & Johnson, M. L. 2016. The genetics of bone mass and susceptibility to bone diseases. *Nat. Rev. Rheumatol.* 12(6):323–334.
- Kartsogiannis, V., & Ng, K. 2004. Cell lines and primary cell cultures in the study of bone cell biology. *Mol. Cell. Endocrinol.* 228(1–2):79–102.
- Katoh, M., & Katoh, M. 2006. Notch ligand, JAG1, is evolutionarily conserved target of canonical WNT signaling pathway in progenitor cells. *Int. J. Mol. Med.* 17(4):681–685.
- Kavanagh, K. L., Dunford, J. E., Bunkoczi, G., Russell, R. G. G., & Oppermann, U. 2006a. The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a novel hexameric arrangement and inhibitory product binding. *J. Biol. Chem.* 281(31):22004–22012.
- Kavanagh, K. L., Guo, K., Dunford, J. E., Wu, X., Knapp, S., Ebetino, F. H., *et al.* 2006b. The molecular mechanism of nitrogen-containing bisphosphonates as antiosteoporosis drugs. *Proc. Natl. Acad. Sci.* 103(20):7829–7834.
- Kemp, J. P., Medina-Gomez, C., Estrada, K., St Pourcain, B., Heppe, D. H. M., Warrington, N. M., *et al.* 2014. Phenotypic Dissection of Bone Mineral Density Reveals Skeletal Site Specificity and Facilitates the Identification of Novel Loci in the Genetic Regulation of Bone Mass Attainment. *PLoS Genet.* 10(6):e1004423.
- Kemp, J. P., Morris, J. A., Medina-Gomez, C., Forgetta, V., Warrington, N. M., Youlten, S. E., *et al.* 2017. Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nat. Genet.* 49(10):1468–1475.
- Kenkre, J. S., & Bassett, J. H. D. 2018. The bone remodelling cycle. *Ann. Clin. Biochem.* 55(3):308–327.
- Khan, M., & Cheung, A. M. 2017. Drug-Related Adverse Events of Osteoporosis Therapy. *Endocrinol. Metab. Clin. North Am.* 46(1):181–192.
- Kharazmi, M., Hallberg, P., Schilcher, J., Aspenberg, P., & Michaëlsson, K. 2016. Mortality after atypical femoral fractures: a cohort study. *J. Bone Miner. Res.* 31(3):491–497.

- Kharazmi, M., Michaëlsson, K., Schilcher, J., Eriksson, N., Melhus, H., Wadelius, M., *et al.* 2019. A Genome-Wide Association Study of Bisphosphonate-Associated Atypical Femoral Fracture. *Calcif. Tissue Int.* .
- Kharwadkar, N., Mayne, B., Lawrence, J. E., & Khanduja, V. 2017. Bisphosphonates and atypical subtrochanteric fractures of the femur. *Bone Jt. Res.* 6(3):144–153.
- Khow, K. S. F., Shibu, P., Yu, S. C. Y., Chehade, M. J., & Visvanathan, R. 2017. Epidemiology and postoperative outcomes of atypical femoral fractures in older adults: A systematic review. *J. Nutr. Heal. Aging* 21(1):83–91.
- Kiel, D. P., Demissie, S., Dupuis, J., Lunetta, K. L., Murabito, J. M., & Karasik, D. 2007. Genome-wide association with bone mass and geometry in the Framingham Heart Study. *BMC Med. Genet.* 8(Suppl 1):S14.
- Kim, B.-J., & Koh, J.-M. 2019. Coupling factors involved in preserving bone balance. *Cell. Mol. Life Sci.* 76(7):1243–1253.
- Kim, S. K. 2018. Identification of 613 new loci associated with heel bone mineral density and a polygenic risk score for bone mineral density, osteoporosis and fracture. (B. O. Williams, Ed.) *PLoS One* 13(7):e0200785.
- Kim, T.-K., & Shiekhatar, R. 2015. Architectural and Functional Commonalities between Enhancers and Promoters. *Cell* 162(5):948–959.
- Koh, J. H., Myong, J. P., Yoo, J., Lim, Y.-W., Lee, J., Kwok, S.-K., *et al.* 2017. Predisposing factors associated with atypical femur fracture among postmenopausal Korean women receiving bisphosphonate therapy: 8 years' experience in a single center. *Osteoporos. Int.* 28(11):3251–3259.
- Koller, D. L., Zheng, H.-F., Karasik, D., Yerges-Armstrong, L., Liu, C.-T., McGuigan, F., *et al.* 2013. Meta-Analysis of Genome-Wide Studies Identifies WNT16 and ESR1 SNPs Associated With Bone Mineral Density in Premenopausal Women. *J. Bone Miner. Res.* 28(3):547–558.
- Komori, T. 2015. Animal models for osteoporosis. *Eur. J. Pharmacol.* 759:287–294.
- Komori, T., Yagi, H., Nomura, S., Yamaguchi, A., Sasaki, K., Deguchi, K., *et al.* 1997. Targeted Disruption of *Cbfa1* Results in a Complete Lack of Bone Formation owing to Maturational Arrest of Osteoblasts. *Cell* 89(5):755–764.
- Kopan, R., & Ilagan, M. X. G. 2009. The Canonical Notch Signaling Pathway: Unfolding the Activation Mechanism. *Cell* 137(2):216–233.
- Kou, I., Takahashi, A., Urano, T., Fukui, N., Ito, H., Ozaki, K., *et al.* 2011. Common variants in a novel gene, FONG on chromosome 2q33.1 confer risk of osteoporosis in Japanese. *PLoS One* 6(5):e19641.
- Kumar, P., Henikoff, S., & Ng, P. C. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* 4(7):1073–1081.
- Kumbaraci, M., Karapinar, L., Incesu, M., & Kaya, A. 2013. Treatment of bilateral simultaneous subtrochanteric femur fractures with proximal femoral nail antirotation (PFNA) in a patient with osteopetrosis: case report and review of the literature. *J. Orthop. Sci.* 18(3):486–489.

## REFERENCES

---

- Kundu, Z. S., Marya, K. M., Devgan, A., Yadav, V., & Rohilla, S. 2004. Subtrochanteric fracture managed by intramedullary nail in a patient with pycnodysostosis. *Jt. Bone Spine* 71(2):154–156.
- Kung, A. W. C., Xiao, S. M., Cherny, S., Li, G. H. Y., Gao, Y., Tso, G., *et al.* 2010. Association of JAG1 with Bone Mineral Density and Osteoporotic Fractures: A Genome-wide Association Study and Follow-up Replication Studies. *Am. J. Hum. Genet.* 86(2):229–239.
- Langdahl, B. L., Uitterlinden, A. G., Ralston, S. H., Trikalinos, T. A., Balcells, S., Brandi, M. L., *et al.* 2008. Large-scale analysis of association between polymorphisms in the transforming growth factor beta 1 gene (TGFB1) and osteoporosis: The GENOMOS study. *Bone* 42(5):969–981.
- Larsen, M. S., & Schmal, H. 2018. The enigma of atypical femoral fractures: a summary of current knowledge. *EFORT Open Rev.* 3(9):494–500.
- Lau, S.-L., McInerney-Leo, A., Landoa-Bassonga, E., & Duncan, E. 2017. Homozygous variant in Cathepsin K in a family with multiple cases of bilateral atypical femoral fracture but without clinical features of pyknodysostosis. *JBMR Plus* 1(S1):S1–S11.
- Leder, B. Z. 2017. Parathyroid Hormone and Parathyroid Hormone-Related Protein Analogs in Osteoporosis Therapy. *Curr. Osteoporos. Rep.* 15(2):110–119.
- Leeuwenburgh, C., & Heinecke, J. W. 2001. Oxidative Stress and Antioxidants in Exercise. *Curr. Med. Chem.* 8(7):829–838.
- Lei, S. F., Chen, Y., Xiong, D. H., Li, L. M., & Deng, H. W. 2006. Ethnic difference in osteoporosis-related phenotypes and its potential underlying genetic determination. *J. Musculoskelet. Neuronal Interact.* 6(1):36–46.
- Leslie, W. D., Lix, L. M., Morin, S. N., Johansson, H., Odén, A., McCloskey, E. V., *et al.* 2015. Hip Axis Length Is a FRAX- and Bone Density- Independent Risk Factor for Hip Fracture in Women. *J. Clin. Endocrinol. Metab.* 100(5):2063–2070.
- Letarouilly, J.-G., Broux, O., & Clabaut, A. 2018. New insights into the epigenetics of osteoporosis. *Genomics* .
- Li, J. J., Wang, B. Q., Fei, Q., Yang, Y., & Li, D. 2016. Identification of candidate genes in osteoporosis by integrated microarray analysis. *Bone Jt. Res.* 5(12):594–601.
- Li, J., Mashiba, T., & Burr, D. B. 2001. Bisphosphonate treatment suppresses not only stochastic remodeling but also the targeted repair of microdamage. *Calcif. Tissue Int.* 69(5):281–286.
- Li, Y., Jin, D., Xie, W., Wen, L., Chen, W., Xu, J., *et al.* 2018. Mesenchymal Stem Cells-Derived Exosomes: A Possible Therapeutic Strategy for Osteoporosis. *Curr. Stem Cell Res. Ther.* 13(5):362–368.
- Liberman, U. A., Weiss, S. R., Bröll, J., Minne, H. W., Quan, H., Bell, N. H., *et al.* 1995. Effect of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. *N. Engl. J. Med.* 333(22):1437–1444.
- Lieben, L., Callewaert, F., & Bouillon, R. 2009. Bone and metabolism: A complex crosstalk. *Horm. Res.* 71(suppl 1):134–138.

- Lim, S.-J., Yeo, I., Yoon, P.-W., Yoo, J. J., Rhyu, K.-H., Han, S.-B., *et al.* 2018. Incidence, risk factors, and fracture healing of atypical femoral fractures: a multicenter case-control study. *Osteoporos. Int.* 29(11):2427–2435.
- Link, T. M., & Heilmeyer, U. 2016. Bone Quality — Beyond Bone Mineral Density. *Semin. Musculoskelet. Radiol.* 20(3):269–278.
- Lisnyansky, M., Kapelushnik, N., Ben-Bassat, A., Marom, M., Loewenstein, A., Khananshvil, D., *et al.* 2018. Reduced Activity of Geranylgeranyl Diphosphate Synthase Mutant Is Involved in Bisphosphonate-Induced Atypical Fractures. *Mol. Pharmacol.* 94(6):1391–1400.
- Liu, J., Curtis, E. M., Cooper, C., & Harvey, N. C. 2019. State of the art in osteoporosis risk assessment and treatment. *J. Endocrinol. Invest.* .
- Lloyd, A. A., Gludovatz, B., Riedel, C., Luengo, E. A., Saiyed, R., Marty, E., *et al.* 2017. Atypical fracture with long-term bisphosphonate therapy is associated with altered cortical composition and reduced fracture resistance. *Proc. Natl. Acad. Sci.* 114(33):201704460.
- Lo, J. C., Hui, R. L., Grimsrud, C. D., Chandra, M., Neugebauer, R. S., Gonzalez, J. R., *et al.* 2016. The association of race/ethnicity and risk of atypical femur fracture among older women receiving oral bisphosphonate therapy. *Bone* 85:142–147.
- Long, H. K., Prescott, S. L., & Wysocka, J. 2016. Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell* 167(5):1170–1187.
- Lorentzon, M., & Cummings, S. R. 2015. Osteoporosis: the evolution of a diagnosis. *J. Intern. Med.* 277(6):650–661.
- Lupiáñez, D. G., Spielmann, M., & Mundlos, S. 2016. Breaking TADs: How Alterations of Chromatin Domains Result in Disease. *Trends Genet.* 32(4):225–237.
- Ma, M., Chen, X., Lu, L., Yuan, F., Zeng, W., Luo, S., *et al.* 2016. Identification of crucial genes related to postmenopausal osteoporosis using gene expression profiling. *Aging Clin. Exp. Res.* 28(6):1067–1074.
- Ma, M., Ru, Y., Chuang, L.-S., Hsu, N.-Y., Shi, L.-S., Hakenberg, J., *et al.* 2015. Disease-associated variants in different categories of disease located in distinct regulatory elements. *BMC Genomics* 16(Suppl 8):S3.
- Mackay, T. F. C. 2014. Epistasis and quantitative traits: Using model organisms to study gene-gene interactions. *Nat. Rev. Genet.* 15(1):22–33.
- Mafi Golchin, M., Heidari, L., Ghaderian, S. M. H., & Akhavan-Niaki, H. 2016. Osteoporosis: A Silent Disease with Complex Genetic Contribution. *J. Genet. Genomics* 43(2):49–61.
- Mahjoub, Z., Jean, S., Leclerc, J. T., Brown, J. P., Boulet, D., Pelet, S., *et al.* 2016. Incidence and Characteristics of Atypical Femoral Fractures: Clinical and Geometrical Data. *J. Bone Miner. Res.* 31(4):767–776.
- Maman, E., Briot, K., & Roux, C. 2016. Atypical femoral fracture in a 51-year-old woman: Revealing a hypophosphatasia. *Jt. Bone Spine* 83(3):346–348.

## REFERENCES

---

- Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., *et al.* 2009. Finding the missing heritability of complex diseases. *Nature* 461(7265):747–753.
- Manolopoulos, K. N., West, A., & Gittoes, N. 2013. The Paradox of Prevention — Bilateral Atypical Subtrochanteric Fractures due to Bisphosphonates in Osteogenesis Imperfecta. *J. Clin. Endocrinol. Metab.* 98(3):871–872.
- Marini, F., & Brandi, M. L. 2014. Pharmacogenetics of osteoporosis. *Best Pract. Res. Clin. Endocrinol. Metab.* 28(6):783–793.
- Marini, F., Falchetti, A., Silvestri, S., Bagger, Y., Luzi, E., Tanini, A., *et al.* 2008. Modulatory effect of farnesyl pyrophosphate synthase (FDPS) rs2297480 polymorphism on the response to long-term amino-bisphosphonate treatment in postmenopausal osteoporosis. *Curr. Med. Res. Opin.* 24(9):2609–2615.
- Marx, J. 2004. Coming to Grips With Bone Loss. *Science* 305(5689):1420–1422.
- Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., *et al.* 2012. Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337(6099):1190–1195.
- McClung, M. R. 2017. Using Osteoporosis Therapies in Combination. *Curr. Osteoporos. Rep.* 15(4):343–352.
- McClung, M. R., Grauer, A., Boonen, S., Bolognese, M. A., Brown, J. P., Diez-Perez, A., *et al.* 2014. Romosozumab in Postmenopausal Women with Low Bone Mineral Density. *N. Engl. J. Med.* 370(5):412–420.
- Medici, M., van Meurs, J. B. J., Rivadeneira, F., Zhao, H., Arp, P. P., Hofman, A., *et al.* 2006. BMP-2 Gene Polymorphisms and Osteoporosis: The Rotterdam Study. *J. Bone Miner. Res.* 21(6):845–854.
- Medina-Gomez, C., Kemp, J. P., Estrada, K., Eriksson, J., Liu, J., Reppe, S., *et al.* 2012. Meta-analysis of genome-wide scans for total body BMD in children and adults reveals allelic heterogeneity and age-specific effects at the WNT16 locus. *PLoS Genet.* 8(7):e1002718.
- Medina-Gomez, C., Kemp, J. P., Trajanoska, K., Luan, J., Chesi, A., Ahluwalia, T. S., *et al.* 2018. Life-Course Genome-wide Association Study Meta-analysis of Total Body BMD and Assessment of Age-Specific Effects. *Am. J. Hum. Genet.* 102(1):88–102.
- Meier, R. P. H., Lorenzini, K. I., Uebelhart, B., Stern, R., Peter, R. E., & Rizzoli, R. 2012a. Atypical femoral fracture following bisphosphonate treatment in a woman with osteogenesis imperfecta—a case report. *Acta Orthop.* 83(5):548–550.
- Meier, R. P. H., Perneger, T. V., & Stern, R. 2012b. Increasing occurrence of atypical femoral fractures associated with bisphosphonate use. *Arch. Intern. Med.* 172(12):930–936.
- Meling, T., Nawab, A., Harboe, K., & Fosse, L. 2014. Atypical femoral fractures in elderly women: A fracture registry-based cohort study. *Bone Jt. J.* 96-B(8):1035–1040.
- Meyer, U. A. 2004. Pharmacogenetics – five decades of therapeutic lessons from genetic diversity. *Nat. Rev. Genet.* 5(9):669–676.
- Michou, L. 2018. Epigenetics of bone diseases. *Jt. Bone Spine* 85(6):701–707.

- Mikhaylichenko, O., Bondarenko, V., Harnett, D., Schor, I. E., Males, M., Viales, R. R., *et al.* 2018. The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes Dev.* 32(1):42–57.
- Miller, P. D., Hattersley, G., Riis, B. J., Williams, G. C., Lau, E., Russo, L. A., *et al.* 2016. Effect of Abaloparatide vs Placebo on New Vertebral Fractures in Postmenopausal Women With Osteoporosis. A Randomized Clinical Trial. *JAMA - J. Am. Med. Assoc.* 316(7):722–733.
- Moayeri, A., Hsu, Y. H., Karasik, D., Estrada, K., Xiao, S. M., Nielson, C., *et al.* 2014. Genetic determinants of heel bone properties: Genome-wide association meta-analysis and replication in the GEFOS/GENOMOS consortium. *Hum. Mol. Genet.* 23(11):3054–3068.
- Mohan, S., Chest, V., Chadwick, R. B., Wergedal, J. E., & Srivastava, A. K. 2007. Chemical mutagenesis induced two high bone density mouse mutants map to a concordant distal chromosome 4 locus. *Bone* 41(5):860–868.
- Mönkkönen, H., Auriola, S., Lehenkari, P., Kellinsalmi, M., Hassinen, I. E., Vepsäläinen, J., *et al.* 2006. A new endogenous ATP analog (Apppl) inhibits the mitochondrial adenine nucleotide translocase (ANT) and is responsible for the apoptosis induced by nitrogen-containing bisphosphonates. *Br. J. Pharmacol.* 147(4):437–445.
- Morin, S. N., Lix, L. M., Majumdar, S. R., & Leslie, W. D. 2013. Temporal Trends in the Incidence of Osteoporotic Fractures. *Curr. Osteoporos. Rep.* 11(4):263–269.
- Morris, J. A., Kemp, J. P., Youtten, S. E., Laurent, L., Logan, J. G., Chai, R., *et al.* 2019. An atlas of genetic influences on osteoporosis in humans and mice. *Nat. Genet.* 51(2):258–266.
- Morris, K. V., & Mattick, J. S. 2014. The rise of regulatory RNA. *Nat. Rev. Genet.* 15(6):423–437.
- Morrison, N. A., Qi, J. C., Tokita, A., Kelly, P. J., Crofts, L., Nguyen, T. V., *et al.* 1994. Prediction of bone density from vitamin D receptor alleles. *Nature* 367(6460):284–287.
- Mudge, J. M., & Harrow, J. 2016. The state of play in higher eukaryote gene annotation. *Nat. Rev. Genet.* 17(12):758–772.
- Mullin, B. H., Walsh, J. P., Zheng, H. F., Brown, S. J., Surdulescu, G. L., Curtis, C., *et al.* 2016. Genome-wide association study using family-based cohorts identifies the WLS and CCDC170/ESR1 loci as associated with bone mineral density. *BMC Genomics* 17:136.
- Mullin, B. H., Zhao, J. H., Brown, S. J., Perry, J. R. B., Luan, J., Zheng, H.-F., *et al.* 2017. Genome-wide association study meta-analysis for quantitative ultrasound parameters of bone identifies five novel loci for broadband ultrasound attenuation. *Hum. Mol. Genet.* 26(14):2791–2802.
- Mullin, B. H., Zhu, K., Xu, J., Brown, S. J., Mullin, S., Tickner, J., *et al.* 2018. Expression Quantitative Trait Locus Study of Bone Mineral Density GWAS Variants in Human Osteoclasts. *J. Bone Miner. Res.* 33(6):1044–1051.
- Murshed, M. 2018. Mechanism of Bone Mineralization. *Cold Spring Harb. Perspect. Med.* 8(12):pii: a031229.

## REFERENCES

---

- Nagao, Y. 2015. Copy number variations play important roles in heredity of common diseases: a novel method to calculate heritability of a polymorphism. *Sci. Rep.* 5(1):17156.
- Nakase, T., Yasui, N., Hiroshima, K., Ohzono, K., Higuchi, C., Shimizu, N., *et al.* 2007. Surgical outcomes after treatment of fractures in femur and tibia in pycnodysostosis. *Arch. Orthop. Trauma Surg.* 127(3):161–165.
- Nancollas, G. H., Tang, R., Phipps, R. J., Henneman, Z., Gulde, S., Wu, W., *et al.* 2006. Novel insights into actions of bisphosphonates on bone: Differences in interactions with hydroxyapatite. *Bone* 38(5):617–627.
- Napoli, N., Villareal, D. T., Mumm, S., Halstead, L., Sheikh, S., Cagaanan, M., *et al.* 2005. Effect of CYP1A1 gene polymorphisms on estrogen metabolism and bone density. *J. Bone Miner. Res.* 20(2):232–239.
- Neer, R. M., Arnaud, C. D., Zanchetta, J. R., Prince, R., Gaich, G. A., Reginster, J.-Y., *et al.* 2001. Effect of Parathyroid Hormone (1-34) on fractures and bone mineral density in postmenopausal women with osteoporosis. *N. Engl. J. Med.* 344(19):1434–1441.
- Nelson, M. R., Tipney, H., Painter, J. L., Shen, J., Nicoletti, P., Shen, Y., *et al.* 2015. The support of human genetic evidence for approved drug indications. *Nat. Genet.* 47(8):856–860.
- Ng, A. C., Png, M. A., Chua, D. T., Koh, J. S. B., & Howe, T. Sen. 2014. Review: Epidemiology and pathophysiology of atypical femur fractures. *Curr. Osteoporos. Rep.* 12(1):65–73.
- Nguyen, H. H., van de Laarschot, D. M., Verkerk, A. J., Milat, F., Zillikens, M. C., & Ebeling, P. R. 2018. Genetic risk factors for atypical femoral fractures (AFFs): a systematic review. *JBMR Plus* 2(1):1–11.
- Nicolaou, N., Agrawal, Y., Padman, M., Fernandes, J. A., & Bell, M. J. 2012. Changing pattern of femoral fractures in osteogenesis imperfecta with prolonged use of bisphosphonates. *J. Child. Orthop.* 6(1):21–27.
- Nielson, C. M., Liu, C.-T., Smith, A. V., Ackert-Bicknell, C. L., Reppe, S., Jakobsdottir, J., *et al.* 2016. Novel Genetic Variants Associated With Increased Vertebral Volumetric BMD, Reduced Vertebral Fracture Risk, and Increased Expression of SLC1A3 and EPHB2. *J. Bone Miner. Res.* 31(12):2085–2097.
- Nieves, J. W. 2013. Nonskeletal Risk Factors for Osteoporosis and Fractures. In R. Marcus, D. Feldman, D. W. Dempster, M. Luckey, & J. A. Cauley (Eds.), *Osteoporosis* (Fourth Edi., pp. 817–839). Elsevier Inc.
- NIH, C. D. P. on O. P. D. and T. 2001. Osteoporosis Prevention, Diagnosis, and Therapy. *JAMA - J. Am. Med. Assoc.* 285(6):785–795.
- Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., *et al.* 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485(7398):381–385.
- Nsengimana, J., & Bishop, D. T. 2017. Design Considerations for Genetic Linkage and Association Studies. In R. C. Elston (Ed.), *Stat. Hum. Genet. Methods Mol. Biol.* (Vol. 1666, pp. 257–281). New York: Humana Press.

- O'Connor, K. M. 2016. Evaluation and Treatment of Osteoporosis. *Med. Clin. North Am.* 100(4):807–826.
- Odvina, C. V., Zerwekh, J. E., Rao, D. S., Maalouf, N., Gottschalk, F. A., & Pak, C. Y. C. 2005. Severely suppressed bone turnover: a potential complication of alendronate therapy. *J. Clin. Endocrinol. Metab.* 90(3):1294–1301.
- Oei, L., Estrada, K., Duncan, E. L., Christiansen, C., Liu, C. T., Langdahl, B. L., *et al.* 2014a. Genome-wide association study for radiographic vertebral fractures: A potential role for the 16q24 BMD locus. *Bone* 59:20–27.
- Oei, L., Hsu, Y.-H., Stykarsdottir, U., Eussen, B. H., de Klein, A., Peters, M. J., *et al.* 2014b. A genome-wide copy number association study of osteoporotic fractures points to the 6p25.1 locus. *J. Med. Genet.* 51(2):122–131.
- Oh, Y., Fujita, K., Wakabayashi, Y., Kurosa, Y., & Okawa, A. 2017. Location of atypical femoral fracture can be determined by tensile stress distribution influenced by femoral bowing and neck-shaft angle: a CT-based nonlinear finite element analysis model for the assessment of femoral shaft loading stress. *Injury* 48(12):2736–2743.
- Oh, Y., Wakabayashi, Y., Kurosa, Y., Ishizuki, M., & Okawa, A. 2014. Stress fracture of the bowed femoral shaft is another cause of atypical femoral fracture in elderly Japanese: a case series. *J. Orthop. Sci.* 19(4):579–586.
- Olmos, J. M., Zarrabeitia, M. T., Hernández, J. L., Sañudo, C., González-Macías, J., & Riancho, J. A. 2012. Common allelic variants of the farnesyl diphosphate synthase gene influence the response of osteoporotic women to bisphosphonates. *Pharmacogenomics J.* 12(3):227–232.
- Ominsky, M. S., Niu, Q.-T., Li, C., Li, X., & Ke, H. Z. 2014. Tissue-Level Mechanisms Responsible for the Increase in Bone Formation and Bone Volume by Sclerostin Antibody. *J. Bone Miner. Res.* 29(6):1424–1430.
- Ong, C.-T., & Corces, V. G. 2011. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat. Rev. Genet.* 12(4):283–293.
- Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B. J., Afzal, S. Y., *et al.* 2018. Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* 554(7691):239–243.
- Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., *et al.* 2004. Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat. Genet.* 36(1):40–45.
- Ozcvici, E., Luu, Y. K., Adler, B., Qin, Y.-X., Rubin, J., Judex, S., *et al.* 2010. Mechanical signals as anabolic agents in bone. *Nat. Rev. Rheumatol.* 6(1):50–59.
- Palomba, S., Numis, F. G., Mossetti, G., Rendina, D., Vuotto, P., Russo, T., *et al.* 2003. Effectiveness of alendronate treatment in postmenopausal women with osteoporosis: relationship with Bsm1 vitamin D receptor genotypes. *Clin. Endocrinol. (Oxf)*. 58(3):365–371.
- Park-Wyllie, L. Y., Mamdani, M. M., Gunraj, N., Austin, P. C., & Whelan, D. B. 2011. Bisphosphonate Use and the Risk of Subtrochanteric or Femoral Shaft Fractures in Older Women. *J. Am. Med. Assoc.* 305(8):783–789.

## REFERENCES

---

- Parra, G., Reymond, A., Dabbouseh, N., Dermitzakis, E. T., Castelo, R., Thomson, T. M., *et al.* 2006. Tandem chimerism as a means to increase protein complexity in the human genome. *Genome Res.* 16(1):37–44.
- Pasquale, E. B. 2008. Eph-Ephrin Bidirectional Signaling in Physiology and Disease. *Cell* 133(1):38–52.
- Paternoster, L., Lorentzon, M., Lehtimäki, T., Eriksson, J., Kähönen, M., Raitakari, O., *et al.* 2013. Genetic Determinants of Trabecular and Cortical Volumetric Bone Mineral Densities and Bone Microstructure. *PLoS Genet.* 9(2):e1003247.
- Paternoster, L., Lorentzon, M., Vandenput, L., Karlsson, M. K., Ljunggren, Ö., Kindmark, A., *et al.* 2010. Genome-wide association meta-analysis of cortical bone mineral density unravels allelic heterogeneity at the RANKL locus and potential pleiotropic effects on bone. *PLoS Genet.* 6(11):e1001217.
- Pei, Y. F., Hu, W. Z., Hai, R., Wang, X. Y., Ran, S., Lin, Y., *et al.* 2016a. Genome-wide association meta-analyses identified 1q43 and 2q32.2 for hip Ward's triangle areal bone mineral density. *Bone* 91:1–10.
- Pei, Y. F., Hu, W. Z., Yan, M. W., Li, C. W., Liu, L., Yang, X. L., *et al.* 2018. Joint study of two genome-wide association meta-analyses identified 20p12.1 and 20q13.33 for bone mineral density. *Bone* 110:378–385.
- Pei, Y. F., Xie, Z. G., Wang, X. Y., Hu, W. Z., Li, L. B., Ran, S., *et al.* 2016b. Association of 3q13.32 variants with hip trochanter and intertrochanter bone mineral density identified by a genome-wide association study. *Osteoporos. Int.* 27(11):3343–3354.
- Pérez-Núñez, I., Pérez-Castrillón, J. L., Zarrabeitia, M. T., García-Ibarbia, C., Martínez-Calvo, L., Olmos, J. M., *et al.* 2015. Exon array analysis reveals genetic heterogeneity in atypical femoral fractures. A pilot study. *Mol. Cell. Biochem.* 409(1–2):45–50.
- Peris, P., González-Roca, E., Rodríguez-García, S. C., López-Cobo, M. del M., Monegal, A., & Guañabens, N. 2019. Incidence of Mutations in the ALPL, GGPS1, and CYP1A1. *JBMR Plus* 3(1):29–36.
- Petersen, B.-S., Fredrich, B., Hoepfner, M. P., Ellinghaus, D., & Franke, A. 2017. Opportunities and challenges of whole-genome and -exome sequencing. *BMC Genet.* 18(1):14.
- Phetfong, J., Sanvoranart, T., Nartprayut, K., Nimsanor, N., Seenprachawong, K., Prachayasittikul, V., *et al.* 2016. Osteoporosis: the current status of mesenchymal stem cell-based therapy. *Cell. Mol. Biol. Lett.* 21:12.
- Plotkin, L. I., Lezcano, V., Thostenson, J., Weinstein, R. S., Manolagas, S. C., & Bellido, T. 2008. Connexin 43 Is Required for the Anti-Apoptotic Effect of Bisphosphonates on Osteocytes and Osteoblasts In Vivo. *J. Bone Miner. Res.* 23(11):1712–1721.
- Pompanon, F., Bonin, A., Bellemain, E., & Taberlet, P. 2005. Genotyping errors: causes, consequences and solutions. *Nat. Rev. Genet.* 6(11):847–859.
- Portin, P., & Wilkins, A. 2017. The Evolving Definition of the Term “Gene.” *Genetics* 205(4):1353–1364.

- Pouresmaeili, F., Kamalidehghan, B., Kamarehei, M., & Goh, Y. M. 2018. A comprehensive overview on osteoporosis and its risk factors. *Ther. Clin. Risk Manag.* 14:2029–2049.
- Prideaux, M., Findlay, D. M., & Atkins, G. J. 2016. Osteocytes: The master cells in bone remodelling. *Curr. Opin. Pharmacol.* 28:24–30.
- Probyn, L., Cheung, A. M., Lang, C., Lenchik, L., Adachi, J. D., Khan, A., *et al.* 2015. Bilateral atypical femoral fractures: how much symmetry is there on imaging? *Skeletal Radiol.* 44(11):1579–1584.
- Qaseem, A., Forciea, M. A., McLean, R. M., & Denberg, T. D. 2017. Treatment of Low Bone Density or Osteoporosis to Prevent Fractures in Men and Women: A Clinical Practice Guideline Update from the American College of Physicians. *Ann. Intern. Med.* 166(11):818–839.
- Qiu, C., Shen, H., Fu, X., Xu, C., Tian, Q., & Deng, H. 2019. Integrative genomic analysis predicts novel functional enhancer-SNPs for bone mineral density. *Hum. Genet.* 138(2):167–185.
- Qiu, S., Divine, G. W., Palnitkar, S., Kulkarni, P., Guthrie, T. S., Honasoge, M., *et al.* 2017. Bone Structure and Turnover Status in Postmenopausal Women with Atypical Femur Fracture After Prolonged Bisphosphonate Therapy. *Calcif. Tissue Int.* 100(3):235–243.
- Raabe, C. A., & Brosius, J. 2015. Does every transcript originate from a gene? *Ann. N. Y. Acad. Sci.* 1341:136–148.
- Raisz, L. G. 2005. Pathogenesis of osteoporosis: concepts, conflicts, and prospects. *J. Clin. Invest.* 115(12):3318–3325.
- Ralston, S. H., & Uitterlinden, A. G. 2010. Genetics of osteoporosis. *Endocr. Rev.* 31(5):629–662.
- Ralston, S. H., Uitterlinden, A. G., Brandi, M. L., Balcells, S., Langdahl, B. L., Lips, P., *et al.* 2006. Large-Scale Evidence for the Effect of the COL1A1 Sp1 Polymorphism on Osteoporosis Outcomes: The GENOMOS Study. *PLoS Med.* 3(4):e90.
- Regan, J., & Long, F. 2013. Notch signaling and bone remodeling. *Curr. Osteoporos. Rep.* 11(2):126–129.
- Reid, I. R., Miller, P. D., Brown, J. P., Kendler, D. L., Fahrleitner-Pammer, A., Valter, I., *et al.* 2010. Effects of Denosumab on Bone Histomorphometry: The FREEDOM and STAND Studies. *J. Bone Miner. Res.* 25(10):2256–2265.
- Reyes, C., Hitz, M., Prieto-Alhambra, D., & Abrahamsen, B. 2016. Risks and Benefits of Bisphosphonate Therapies. *J. Cell. Biochem.* 117(1):20–28.
- Ribeiro, V., Garcia, M., Oliveira, R., Gomes, P. S., Colaço, B., & Fernandes, M. H. 2014. Bisphosphonates induce the osteogenic gene expression in co-cultured human endothelial and mesenchymal stem cells. *J. Cell. Mol. Med.* 18(1):27–37.
- Richards, J. B., Kavvoura, F. K., Rivadeneira, F., Styrkársdóttir, U., Estrada, K., Halldórsson, B. V., *et al.* 2009. Collaborative Meta-analysis: Associations of 150 Candidate Genes With Osteoporosis and Osteoporotic Fracture. *Ann. Intern. Med.* 151(8):528–537.

## REFERENCES

---

- Richards, J. B., Rivadeneira, F., Inouye, M., Pastinen, T. M., Soranzo, N., Wilson, S. G., *et al.* 2008. Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet* 371(9623):1505–1512.
- Richards, J. B., Zheng, H. F., & Spector, T. D. 2012. Genetics of osteoporosis from genome-wide association studies: Advances and challenges. *Nat. Rev. Genet.* 13(8):576–588.
- Riggs, B. L., Khosla, S., & Melton, L. J. 2002. Sex Steroids and the Construction and Conservation of the Adult Skeleton. *Endocr. Rev.* 23(3):279–302.
- Rivadeneira, F., Styrkársdóttir, U., Estrada, K., Halldórsson, B. V., Hsu, Y. H., Richards, J. B., *et al.* 2009. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat. Genet.* 41(11):1199–1206.
- Rogers, M. J., Crockett, J. C., Coxon, F. P., & Mönkkönen, J. 2011. Biochemical and molecular mechanisms of action of bisphosphonates. *Bone* 49(1):34–41.
- Ronis, M. J. J., Mercer, K., & Chen, J.-R. 2011. Effects of Nutrition and Alcohol Consumption on Bone Loss. *Curr. Osteoporos. Rep.* 9(2):53–59.
- Runyan, C. M., & Gabrick, K. S. 2017. Biology of Bone Formation, Fracture Healing, and Distraction Osteogenesis. *J. Craniofac. Surg.* 28(5):1380–1389.
- Russell, R. G. G., Watts, N. B., Ebtino, F. H., & Rogers, M. J. 2008. Mechanisms of action of bisphosphonates: similarities and differences and their potential influence on clinical efficacy. *Osteoporos. Int.* 19(6):733–759.
- Sadee, W. 2013. Gene-gene-environment interactions between drugs, transporters, receptors, and metabolizing enzymes: Statins, SLCO1B1, and CYP3A4 as an example. *J. Pharm. Sci.* 102(9):2924–2929.
- Saita, Y., Ishijima, M., Mogami, A., Kubota, M., Baba, T., Kaketa, T., *et al.* 2015. The incidence of and risk factors for developing atypical femoral fractures in Japan. *J. Bone Miner. Metab.* 33(3):311–318.
- Sato, M. M., Nakashima, A., Nashimoto, M., Yawaka, Y., & Tamura, M. 2009. Bone morphogenetic protein-2 enhances Wnt/ $\beta$ -catenin signaling-induced osteoprotegerin expression. *Genes to Cells* 14(2):141–153.
- Shier, D., Butler, J., & Lewis, R. 2016. Hole's Human Anatomy and Physiology (14th ed.). McGraw-Hill.
- Schilcher, J. 2013. Epidemiology, radiology and histology of atypical femoral fractures. *Acta Orthop.* 84(352):1–26.
- Schilcher, J., Howe, T. Sen, Png, M. A., Aspenberg, P., & Koh, J. S. B. 2015a. Atypical Fractures are Mainly Subtrochanteric in Singapore and Diaphyseal in Sweden: A Cross-Sectional Study. *J. Bone Miner. Res.* 30(11):2127–2132.
- Schilcher, J., Koeppen, V., Aspenberg, P., & Michaëlsson, K. 2014. Risk of atypical femoral fracture during and after bisphosphonate use. *N. Engl. J. Med.* 371(10):974–976.
- Schilcher, J., Koeppen, V., Aspenberg, P., & Michaëlsson, K. 2015b. Risk of atypical femoral fracture during and after bisphosphonate use. *Acta Orthop.* 86(1):100–107.

- Schoenfelder, S., & Fraser, P. 2019. Long-range enhancer–promoter contacts in gene expression control. *Nat. Rev. Genet.* .
- Schork, N. J., Murray, S. S., Frazer, K. A., & Topol, E. J. 2009. Common vs. rare allele hypotheses for complex diseases. *Curr. Opin. Genet. Dev.* 19(3):212–219.
- Schousboe, J. T. 2016. Epidemiology of Vertebral Fractures. *J. Clin. Densitom.* 19(1):8–22.
- Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. 2014. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* 11(4):361–362.
- Sergi, C., Mornet, E., Troeger, J., & Voightlaender, T. 2001. Perinatal Hypophosphatasia: Radiology, Pathology and Molecular Biology Studies in a Family Harboring a Splicing Mutation (648+1A) and a Novel Missense Mutation (N400S) in the Tissue-Nonspecific Alkaline Phosphatase (TNSALP) Gene. *Am. J. Med. Genet. Part A* 103(3):235–240.
- Shane, E., Burr, D., Abrahamsen, B., Adler, R. A., Brown, T. D., Cheung, A. M., *et al.* 2014. Atypical subtrochanteric and diaphyseal femoral fractures: Second report of a task force of the American society for bone and mineral research. *J. Bone Miner. Res.* 29(1):1–23.
- Shane, E., Burr, D., Ebeling, P. R., Abrahamsen, B., Adler, R. A., Brown, T. D., *et al.* 2010. Atypical Subtrochanteric and Diaphyseal Femoral Society for Bone and Mineral Research. *J. Bone Miner. Res.* 25(11):2267–2294.
- Shen, Y., Gray, D. L., & Martinez, D. S. 2017. Combined Pharmacologic Therapy in Postmenopausal Osteoporosis. *Endocrinol. Metab. Clin. North Am.* 46(1):193–206.
- Silverman, S., Kupperman, E., & Bukata, S. 2018. Bisphosphonate-related atypical femoral fracture: Managing a rare but serious complication. *Cleve. Clin. J. Med.* 85(11):885–893.
- Sims, A.-M., Shephard, N., Carter, K., Doan, T., Dowling, A., Duncan, E. L., *et al.* 2008. Genetic Analyses in a Sample of Individuals With High or Low BMD Shows Association With Multiple Wnt Pathway Genes. *J. Bone Miner. Res.* 23(4):499–506.
- Simundic, A.-M. 2010. Methodological issues of genetic association studies. *Clin. Chem. Lab. Med.* 48(Suppl 1):S115–S118.
- Sobacchi, C., Schulz, A., Coxon, F. P., Villa, A., & Helfrich, M. H. 2013. Osteopetrosis: genetics, treatment and new insights into osteoclast function. *Nat. Rev. Endocrinol.* 9(9):522–536.
- Song, H. K., Sohn, Y. B., Choi, Y. J., Chung, Y.-S., & Jang, J.-H. 2017. A case report of pycnodysostosis with atypical femur fracture diagnosed by next-generation sequencing of candidate genes. *Medicine (Baltimore).* 96(12):e6367.
- Sornay-Rendu, E., Munoz, F., Garnero, P., Duboeuf, F., & Delmas, P. D. 2005. Identification of Osteopenic Women at High Risk of Fracture: The OFELY Study. *J. Bone Miner. Res.* 20(10):1813–1819.
- Soysa, N. S., & Alles, N. 2016. Osteoclast function and bone-resorbing activity: An overview. *Biochem. Biophys. Res. Commun.* 476(3):115–120.

## REFERENCES

---

- Sözen, T., Özişik, L., & Başaran, N. Ç. 2017. An overview and management of osteoporosis. *Eur. J. Rheumatol.* 4(1):46–56.
- Starr, J., Tay, Y. K. D., & Shane, E. 2018. Current Understanding of Epidemiology, Pathophysiology, and Management of Atypical Femur Fractures. *Curr. Osteoporos. Rep.* 16(4):519–529.
- Styrkarsdottir, U., Cazier, J.-B., Kong, A., Rolfsson, O., Larsen, H., Bjarnadottir, E., *et al.* 2003. Linkage of Osteoporosis to Chromosome 20p12 and Association to BMP2. *PLoS Biol.* 1(3):e69.
- Styrkarsdottir, U., Halldorsson, B. V., Gretarsdottir, S., Gudbjartsson, D. F., Walters, G. B., Ingvarsson, T., *et al.* 2008. Multiple Genetic Loci for Bone Mineral Density and Fractures. *N. Engl. J. Med.* 358(22):2355–2365.
- Styrkarsdottir, U., Halldorsson, B. V., Gretarsdottir, S., Gudbjartsson, D. F., Walters, G. B., Ingvarsson, T., *et al.* 2009. New sequence variants associated with bone mineral density. *Nat. Genet.* 41(1):15–17.
- Styrkarsdottir, U., Halldorsson, B. V., Gudbjartsson, D. F., Tang, N. L. S., Koh, J. M., Xiao, S. M., *et al.* 2010. European bone mineral density loci are also associated with BMD in East-Asian populations. *PLoS One* 5(10):e13217.
- Styrkarsdottir, U., Stefansson, O. A., Gunnarsdottir, K., Thorleifsson, G., Lund, S. H., Stefansdottir, L., *et al.* 2019. GWAS of bone size yields twelve loci that also affect height, BMD, osteoarthritis or fractures. *Nat. Commun.* 10(1):2054.
- Styrkarsdottir, U., Thorleifsson, G., Eiriksdottir, B., Gudjonsson, S. A., Ingvarsson, T., Center, J. R., *et al.* 2016a. Two Rare Mutations in the COL1A2 Gene Associate With Low Bone Mineral Density and Fractures in Iceland. *J. Bone Miner. Res.* 31(1):173–179.
- Styrkarsdottir, U., Thorleifsson, G., Gudjonsson, S. A., Sigurdsson, A., Center, J. R., Lee, S. H., *et al.* 2016b. Sequence variants in the PTCH1 gene associate with spine bone mineral density and osteoporotic fractures. *Nat. Commun.* 7:10129.
- Styrkarsdottir, U., Thorleifsson, G., Sulem, P., Gudbjartsson, D. F., Sigurdsson, A., Jonasdottir, A., *et al.* 2013. Nonsense mutation in the LGR4 gene is associated with several human diseases and other traits. *Nature* 497(7450):517–520.
- Sum, M., Huskey, M., Diemer, K., Civitelli, R., Gardner, M., McAndrew, C., *et al.* 2013. TNSALP mutation analysis in women with atypical femoral fracture and bisphosphonate therapy for osteoporosis. *J. Bone Miner. Res.* 28(S1):S295.
- Sun, Y., Ruivenkamp, C. A. L., Hoffer, J. V., Vrijenhoek, T., Kriek, M., van Asperen, C. J., *et al.* 2015. Next-Generation Diagnostics: Gene Panel, Exome, or Whole Genome? *Hum. Mutat.* 36(6):648–655.
- Sutton, R. A. L., Mumm, S., Coburn, S. P., Ericson, K. L., & Whyte, M. P. 2012. “Atypical femoral fractures” during bisphosphonate exposure in adult hypophosphatasia. *J. Bone Miner. Res.* 27(5):987–994.
- Tang, S. Y., Zeenath, U., & Vashishth, D. 2007. Effects of non-enzymatic glycation on cancellous bone fragility. *Bone* 40(4):1144–1151.
- The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489(7414):57–74.

- The FANTOM Consortium, & RIKEN Genome Exploration Research Group, R. G. E. R. 2005. The Transcriptional Landscape of the Mammalian Genome. *Science* 309(5740):1559–1563.
- The GTEx Consortium. 2013. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* 45(6):580–585.
- Timpson, N. J., Tobias, J. H., Richards, J. B., Soranzo, N., Duncan, E. L., Sims, A. M., *et al.* 2009. Common variants in the region around Osterix are associated with bone mineral density and growth in childhood. *Hum. Mol. Genet.* 18(8):1510–1517.
- Trajanoska, K., Morris, J. A., Oei, L., Zheng, H.-F., Evans, D. M., Kiel, D. P., *et al.* 2018. Assessment of the genetic and clinical determinants of fracture risk: genome wide association and mendelian randomisation study. *Br. Med. J.* 362:k3225.
- Trerotola, M., Relli, V., Simeone, P., & Alberti, S. 2015. Epigenetic inheritance and the missing heritability. *Hum. Genomics* 9:17.
- Tsai, J. N., Uihlein, A. V, Lee, H., Kumbhani, R., Siwila-Sackman, E., McKay, E. A., *et al.* 2013. Teriparatide and denosumab, alone or combined, in women with postmenopausal osteoporosis: the DATA study randomised trial. *Lancet* 382(9886):50–56.
- Tse, K. Y., Macias, B. R., Meyer, R. S., & Hargens, A. R. 2009. Heritability of Bone Density: Regional and Gender Differences in Monozygotic Twins. *J. Orthop. Res.* 27(2):150–154.
- Tsubaki, M., Komai, M., Itoh, T., Imano, M., Sakamoto, K., Shimaoka, H., *et al.* 2014. Nitrogen-containing bisphosphonates inhibit RANKL- and M-CSF-induced osteoclast formation through the inhibition of ERK1/2 and Akt activation. *J. Biomed. Sci.* 21:10.
- Uitterlinden, A. G., Ralston, S. H., Brandi, M. L., Carey, A. H., Grinberg, D., Langdahl, B. L., *et al.* 2006. The Association between Common Vitamin D Receptor Gene Variations and Osteoporosis: A Participant-Level Meta-Analysis. *Ann. Intern. Med.* 145(4):255–264.
- Unnanuntana, A., Gladnick, B. P., Donnelly, E., & Lane, J. M. 2010. The Assessment of Fracture Risk. *J. Bone Jt. Surg.* 92(3):743–753.
- van Beek, E., Pieterman, E., Cohen, L., Löwik, C., & Papapoulos, S. 1999. Farnesyl pyrophosphate synthase is the molecular target of nitrogen-containing bisphosphonates. *Biochem. Biophys. Res. Commun.* 264(1):108–111.
- van de Laarschot, D. M., Smits, A. A. A., Buitendijk, S. K. C., Stegenga, M. T., & Zillikens, M. C. 2017. Screening for Atypical Femur Fractures Using Extended Femur Scans by DXA. *J. Bone Miner. Res.* 32(8):1632–1639.
- van de Laarschot, D. M., & Zillikens, M. C. 2016. Atypical femur fracture in an adolescent boy treated with bisphosphonates for X-linked osteoporosis based on PLS3 mutation. *Bone* 91:148–151.
- van Dijk, F. S., Zillikens, M. C., Micha, D., Riessland, M., Marcelis, C. L. M., de Die-Smulders, C. E., *et al.* 2013. PLS3 Mutations in X-Linked Osteoporosis with Fractures. *N. Engl. J. Med.* 369(19):1529–1536.

## REFERENCES

---

- van Meurs, J. B. J., Trikalinos, T. A., Ralston, S. H., Balcells, S., Brandi, M. L., Brixen, K., *et al.* 2008. Large-Scale Analysis of Association Between LRP5 and LRP6 Variants and Osteoporosis. *JAMA - J. Am. Med. Assoc.* 299(11):1277–1290.
- Vasanwala, R. F., Sanghrajka, A., Bishop, N. J., & Högl, W. 2016. Recurrent Proximal Femur Fractures in a Teenager With Osteogenesis Imperfecta on Continuous Bisphosphonate Therapy: Are We Overtreating? *J. Bone Miner. Res.* 31(7):1449–1454.
- Vashishth, D., Gibson, G. J., Khoury, J. I., Schaffler, M. B., Kimura, J., & Fyhrie, D. P. 2001. Influence of nonenzymatic glycation on biomechanical properties of cortical bone. *Bone* 28(2):195–201.
- Vasikaran, S., Eastell, R., Bruyère, O., Foldes, A. J., Garnero, P., Griesmacher, A., *et al.* 2011. Markers of bone turnover for the prediction of fracture risk and monitoring of osteoporosis treatment: a need for international reference standards. *Osteoporos. Int.* 22(2):391–420.
- Visekruna, M., Wilson, D., & McKiernan, F. E. 2008. Severely suppressed bone turnover and atypical skeletal fragility. *J. Clin. Endocrinol. Metab.* 93(8):2948–2952.
- Visel, A., Minovitsky, S., Dubchak, I., & Pennacchio, L. A. 2007. VISTA Enhancer Browser - A database of tissue-specific human enhancers. *Nucleic Acids Res.* 35(Database issue):D88-92.
- Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., *et al.* 2017. 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am. J. Hum. Genet.* 101(1):5–22.
- Vissers, L. E. L. M., & Veltman, J. A. 2015. Standardized phenotyping enhances Mendelian disease gene identification. *Nat. Genet.* 47(11):1222–1224.
- Vomund, A. N., Braddock, S. R., Krause, G. F., & Phillips, C. L. 2004. Potential modifier role of the R618Q variant of pro $\alpha$ 2(I)collagen in type I collagen fibrillogenesis: in vitro assembly analysis. *Mol. Genet. Metab.* 82(2):144–153.
- von Knoch, F., Jaquiere, C., Kowalsky, M., Schaeren, S., Alabre, C., Martin, I., *et al.* 2005. Effects of bisphosphonates on proliferation and osteoblast differentiation of human bone marrow stromal cells. *Biomaterials* 26(34):6941–6949.
- Vrtačnik, P., Marc, J., & Ostanek, B. 2014. Epigenetic mechanisms in bone. *Clin. Chem. Lab. Med.* 52(5):589–608.
- Wade, S. W., Strader, C., Firtzpatrick, L. A., Anthony, M. S., & O'Malley, C. D. 2014. Estimating prevalence of osteoporosis: examples from industrialized countries. *Arch. Osteoporos.* 9:182.
- Wang, H., Gong, C., Liu, X., Rao, S., Li, T., He, L., *et al.* 2018a. Genetic interaction of purinergic P2X7 receptor and ER- $\alpha$  polymorphisms in susceptibility to osteoporosis in Chinese postmenopausal women. *J. Bone Miner. Metab.* 36(4):488–497.
- Wang, J., & Stern, P. H. 2011. Dose-dependent differential effects of risedronate on gene expression in osteoblasts. *Biochem. Pharmacol.* 81(8):1036–1042.
- Wang, W.-J., Fu, W.-Z., He, J.-W., Wang, C., & Zhang, Z.-L. 2018b. Association between SOST gene polymorphisms and response to alendronate treatment in

- postmenopausal Chinese women with low bone mineral density. *Pharmacogenomics J.* .
- Ward, K. D., & Klesges, R. C. 2001. A Meta-Analysis of the Effects of Cigarette Smoking on Bone Mineral Density. *Calcif. Tissue Int.* 68(5):259–270.
- Watts, N. B., Aggers, D., McCarthy, E. F., Savage, T., Martinez, S., Patterson, R., *et al.* 2017. Responses to treatment with teriparatide in patients with atypical femur fractures previously treated with bisphosphonates. *J. Bone Miner. Res.* 32(5):1027–1033.
- Way, G. P., Youngstrom, D. W., Hankenson, K. D., Greene, C. S., & Grant, S. F. A. 2017. Implicating candidate genes at GWAS signals by leveraging topologically associating domains. *Eur. J. Hum. Genet.* 25(11):1286–1289.
- Whyte, M. P. 2006. Paget's disease of bone and genetic disorders of RANKL/OPG/RANK/NF- $\kappa$ B signaling. *Ann. N. Y. Acad. Sci.* 1068(1):143–164.
- Whyte, M. P. 2009. Atypical femoral fractures, bisphosphonates, and adult hypophosphatasia. *J. Bone Miner. Res.* 24(6):1132–1134.
- Whyte, M. P. 2016. Hypophosphatasia-aetiology, nosology, pathogenesis, diagnosis and treatment. *Nat. Rev. Endocrinol.* 12(4):233–246.
- Will, A. J., Cova, G., Osterwalder, M., Chan, W. L., Wittler, L., Brieske, N., *et al.* 2017. Composition and dosage of a multipartite enhancer cluster control developmental expression of *Ihh* (Indian hedgehog). *Nat. Genet.* 49(10):1539–1545.
- World Health Organization. 1994. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. World Health Organ. Tech. Rep. Ser. 843:1–129.
- Wu, M., Chen, G., & Li, Y.-P. 2016. TGF- $\beta$  and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease. *Bone Res.* 4:16009.
- Wu, S., Liu, Y., Zhang, L., Han, Y., Lin, Y., & Deng, H.-W. 2013. Genome-wide approaches for identifying genetic risk factors for osteoporosis. *Genome Med.* 5(5):44.
- Xiong, D.-H., Liu, X.-G., Guo, Y.-F., Tan, L.-J., Wang, L., Sha, B.-Y., *et al.* 2009. Genome-wide Association and Follow-Up Replication Studies Identified ADAMTS18 and TGFBR3 as Bone Mass Candidate Genes in Different Ethnic Groups. *Am. J. Hum. Genet.* 84(3):388–398.
- Yang, T.-L., Chen, X.-D., Guo, Y., Lei, S.-F., Wang, J.-T., Zhou, Q., *et al.* 2008. Genome-wide Copy-Number-Variation Study Identified a Susceptibility Gene, UGT2B17, for Osteoporosis. *Am. J. Hum. Genet.* 83(6):663–674.
- Yang, T.-L., Guo, Y., Li, J., Zhang, L., Shen, H., Li, S. M., *et al.* 2013. Gene-Gene Interaction Between RBMS3 and ZNF516 Influences Bone Mineral Density. *J. Bone Miner. Res.* 28(4):828–837.
- Yates, C. J., Bartlett, M. J., & Ebeling, P. R. 2011. An atypical subtrochanteric femoral fracture from pycnodysostosis: A lesson from nature. *J. Bone Miner. Res.* 26(6):1377–1379.

## REFERENCES

---

- Yi, L., Haijuan, L., Mei, L., Peiran, Z., Xiaoping, X., Weibo, X., *et al.* 2014. Association of farnesyl diphosphate synthase polymorphisms and response to alendronate treatment in Chinese postmenopausal women with osteoporosis. *Chin. Med. J. (Engl)*. 127(4):662–668.
- Yu, Z., Surface, L. E., Park, C. Y., Horlbeck, M. A., Wyant, G. A., Abu-Remaileh, M., *et al.* 2018. Identification of a transporter complex responsible for the cytosolic entry of nitrogen-containing bisphosphonates. *Elife* 7:e36620.
- Yuan, J., Tickner, J., Mullin, B. H., Zhao, J., Zeng, Z., Morahan, G., *et al.* 2019. Advanced Genetic Approaches in Discovery and Characterization of Genes Involved With Osteoporosis in Mouse and Human. *Front. Genet.* 10:288.
- Yuan, Y., Chen, X., Zhang, L., Wu, J., Guo, J., Zou, D., *et al.* 2016. The roles of exercise in bone remodeling and in prevention and treatment of osteoporosis. *Prog. Biophys. Mol. Biol.* 122(2):122–130.
- Yuasa, T., Maeda, K., Kaneko, K., & Yoshikata, K. 2015. Total Hip Arthroplasty after Treatment of an Atypical Subtrochanteric Femoral Fracture in a Patient with Pycnodysostosis. *Case Rep. Orthop.* 2015:731910.
- Yuen, T., Stachnik, A., Davies, T. F., Bian, Z., Sun, L., Lu, P., *et al.* 2014. Bisphosphonates inactivate human EGFRs to exert antitumor actions. *Proc. Natl. Acad. Sci.* 111(50):17989–17994.
- Zameer, S., Najmi, A. K., Vohora, D., & Akhtar, M. 2018. Bisphosphonates: Future perspective for neurological disorders. *Pharmacol. Reports* 70(5):900–907.
- Zhang, G., Karns, R., Sun, G., Indugula, S. R., Cheng, H., Havas-Augustin, D., *et al.* 2012. Finding Missing Heritability in Less Significant Loci and Allelic Heterogeneity: Genetic Variation in Human Height. *PLoS One* 7(12):e51211.
- Zhang, H., Sol-Church, K., Rydbeck, H., Stabley, D., Spotila, L. D., & Devoto, M. 2009. High resolution linkage and linkage disequilibrium analyses of chromosome 1p36 SNPs identify new positional candidate genes for low bone mineral density. *Osteoporos. Int.* 20(2):341–346.
- Zhang, L., Choi, H. J., Estrada, K., Leo, P. J., Li, J., Pei, Y.-F., *et al.* 2014. Multistage genome-wide association meta-analyses identified two new loci for bone mineral density. *Hum. Mol. Genet.* 23(7):1923–1933.
- Zhang, Q., Greenbaum, J., Zhang, W.-D., Sun, C.-Q., & Deng, H.-W. 2018. Age at menarche and osteoporosis: A Mendelian randomization study. *Bone* 117:91–97.
- Zhao, L.-J., Liu, X.-G., Liu, Y.-Z., Liu, Y.-J., Papiasian, C. J., Sha, B.-Y., *et al.* 2010. Genome-Wide Association Study for Femoral Neck Bone Geometry. *J. Bone Miner. Researc* 25(2):320–329.
- Zheng, H. F., Duncan, E. L., Yerges-Armstrong, L. M., Eriksson, J., Bergström, U., Leo, P. J., *et al.* 2013. Meta-analysis of genome-wide studies identifies MEF2C SNPs associated with bone mineral density at forearm. *J. Med. Genet.* 50(7):473–478.
- Zheng, H. F., Forgetta, V., Hsu, Y. H., Estrada, K., Rosello-Diez, A., Leo, P. J., *et al.* 2015. Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature* 526(7571):112–117.

- 
- Zheng, H. F., Tobias, J. H., Duncan, E., Evans, D. M., Eriksson, J., Paternoster, L., *et al.* 2012. WNT16 influences bone mineral density, cortical bone thickness, bone strength, and osteoporotic fracture risk. (G. Gibson, Ed.)*PLoS Genet.* 8(7):e1002745.
- Zhou, S.-F., Liu, J.-P., & Chowbay, B. 2009. Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metab. Rev.* 41(2):89–295.
- Zhu, D. L., Chen, X. F., Hu, W. X., Dong, S. S., Lu, B. J., Rong, Y., *et al.* 2018. Multiple Functional Variants at 13q14 Risk Locus for Osteoporosis Regulate RANKL Expression Through Long-Range Super-Enhancer. *J. Bone Miner. Res.* 33(7):1335–1346.
- Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M. R., Powell, J. E., *et al.* 2016. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* 48(5):481–487.
- Zuk, O., Hechter, E., Sunyaev, S. R., & Lander, E. S. 2012. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc. Natl. Acad. Sci.* 109(4):1193–1198.
- Zuk, O., Schaffner, S. F., Samocha, K., Do, R., Hechter, E., Kathiresan, S., *et al.* 2014. Searching for missing heritability: Designing rare variant association studies. *Proc. Natl. Acad. Sci.* 111(4):E455–E464.



**ANNEX**



## Other articles related to this Thesis

- Neus Roca-Ayats, Mónica Cozar Morillo, Marina Gerousi, Esteban Czwan, Roser Urreizti, Martínez-Gil Núria, Garcia-Giralt, Natàlia, Mellibovsky Leonardo, Xavier Nogués, Adolfo Díez-Pérez, Susana Balcells, Daniel Grinberg. Identificación de variantes genéticas asociadas con la densidad mineral ósea (DMO) en el gen *FLJ42280*. *Revista de Osteoporosis y Metabolismo Mineral*. 2017;9(1):28-34. doi: 10.4321/S1889-836X2017000100005
- Neus Roca-Ayats, Maite Falcó-Mascaró, Natàlia Garcia-Giralt, Mónica Cozar, Josep Francesc Abril, José Manuel Quesada-Gómez, Daniel Prieto-Alhambra, Xavier Nogués, Leonardo Mellibovsky, Adolfo Díez-Pérez, Daniel Grinberg, Susanna Balcells. Estudio genético de la fractura femoral atípica mediante la secuenciación del exoma en tres hermanas afectas y tres pacientes no relacionadas. *Revista de Osteoporosis y Metabolismo Mineral*. 2018;10(4):108-118. doi: 10.4321/S1889-836X2018000400002
- Jenny Serra-Vinardell, Neus Roca-Ayats, Laura De-Ugarte, Lluïsa Vilageliu, Susanna Balcells, Daniel Grinberg. Bone development and remodeling in metabolic disorders. *Journal of Inherited Metabolic Disease*. 2019 Epub ahead of print. doi: 10.1002/jimd.12097



**Roca-Ayats N<sup>1</sup>, Cozar Morillo M<sup>1</sup>, Gerousi M<sup>1</sup>, Czwan E<sup>2</sup>, Urreiziti R<sup>1</sup>, Martínez-Gil N<sup>1</sup>, García-Giralt N<sup>3</sup>, Mellibovsky L<sup>3</sup>, Nogués X<sup>3</sup>, Díez-Pérez A<sup>3</sup>, Balcells S<sup>1</sup>, Grinberg D<sup>1</sup>**

<sup>1</sup> Departamento de Genética, Microbiología y Estadística - Facultad de Biología - Universidad de Barcelona - CIBERER (Centro de Investigación en Red de Enfermedades Raras) - IBUB (Instituto de Biomedicina de la Universidad de Barcelona) - Barcelona (España)

<sup>2</sup> Laboratorios Roche Diagnostics Deutschland GmbH - Mannheim (Alemania)

<sup>3</sup> Unidad de Investigación de Patología Ósea y Articular (URFOA) - IMIM (Instituto Hospital del Mar de Investigaciones Médicas) - Parque de Salud Mar - RETICEF (Red Temática de Investigación Cooperativa en Envejecimiento y Fragilidad) - Barcelona (España)

## Identificación de variantes genéticas asociadas con la densidad mineral ósea (DMO) en el gen *FLJ42280*

DOI: <http://dx.doi.org/10.4321/S1889-836X2017000100005>

Correspondencia: Daniel Grinberg - Departamento de Genética, Microbiología y Estadística - Facultad de Biología - Avda. Diagonal, 643 - 08028 Barcelona (España)  
Correo electrónico: [dgrinberg@ub.edu](mailto:dgrinberg@ub.edu)

Fecha de recepción: 15/10/2016

Fecha de aceptación: 15/12/2016

*Trabajo premiado con una beca de Investigación FEIOMM 2014.*

### Resumen

*FLJ42280* es un posible gen de susceptibilidad a la osteoporosis. Distintos estudios de GWAs han identificado 4 SNPs no-codificantes en este gen que se asocian a la densidad mineral ósea (DMO) y el riesgo de fractura.

Para descubrir la causa de la asociación entre estos SNPs y la osteoporosis, se realizó una búsqueda de variantes genéticas mediante resecuenciación de 28 kb que contienen el gen, en una selección truncada de mujeres con DMO muy baja (n=50) o muy alta (n=50) de la cohorte BARCOS (Barcelona Cohorte Osteoporosis, cohorte de mujeres postmenopáusicas de Barcelona). Las variantes encontradas se filtraron y se analizó su frecuencia en cada grupo. Se analizó el solapamiento de las variantes con elementos funcionales del proyecto ENCODE y también se calculó el desequilibrio de ligamiento entre los SNPs de la región. Finalmente, se hizo un análisis de eQTL de los 4 SNPs no-codificantes respecto a los niveles de expresión de genes cercanos a *FLJ42280* en linfoblastos.

Se seleccionaron 110 variantes. Las diferencias de sus frecuencias entre los dos grupos estuvieron por debajo del poder estadístico del diseño experimental. Sin embargo, 3 variantes solaparon con posibles *enhancers* y una solapó con un *enhancer* activo en osteoblastos (rs4613908). Se observó un fuerte desequilibrio de ligamiento entre los 4 SNPs no-codificantes y el SNP rs4613908, que pertenecen a un bloque que abarca el gen casi por completo. Ninguno de los SNPs no-codificantes mostró asociación con los niveles de expresión de genes cercanos a *FLJ42280*.

En conclusión, el SNP rs4613908 podría estar implicado funcionalmente en la determinación de la DMO. Serán necesarios experimentos concretos para confirmarlo.

**Palabras clave:** *FLJ42280*, densidad mineral ósea, variantes genéticas, eQTLs, enhancers.

## Identification of genetic variants associated with bone mineral density (BMD) in the *FLJ42280* gene

### Summary

*FLJ42280* is a possible gene for susceptibility to osteoporosis. Different studies of GWAs have identified 4 non-coding SNPs in this gene associated with bone mineral density (BMD) and fracture risk.

In order to ascertain the cause of the association between these SNPs and osteoporosis, we searched for genetic variants by resequencing the 28-kb gene, in a truncated selection of women with very low (n=50) or very high BMD (N=50) of the BARCOS cohort (Barcelona Cohort Osteoporosis, cohort of postmenopausal women in Barcelona). The variants found were filtered and their frequency analyzed in each group.

The overlap of the variants with functional elements of the ENCODE project was calculated. Finally, an eQTL analysis of the 4 SNPs-coding was performed on the expression levels of *FLJ42280* neighbor genes in lymphoblasts.

In all, 110 variants were selected. The differences in their frequencies between the two groups were below the statistical power of the experimental design. However, three variants overlapped with possible enhancers and one overlapped with an active enhancer in osteoblasts (rs4613908). A strong linkage disequilibrium was observed between the 4 non-coding SNPs and the SNP rs4613908, which belong to a block spanning the gene almost completely. None of the non-coding SNPs showed association with the expression levels of *FLJ42280* neighbor genes.

In conclusion, the SNP rs4613908 could be involved functionally in determining BMD. Tangible experiments will be required to confirm this.

**Key words:** *FLJ42280*, bone mineral density, genetic variants, eQTLs, enhancers.

### Introducción

La osteoporosis es una enfermedad compleja caracterizada por baja masa ósea y deterioro de la microarquitectura del tejido óseo que conduce a un mayor riesgo de fractura. Por ejemplo, en EE.UU. se producen 1,5 millones de nuevos casos de fracturas cada año, lo que representa una enorme carga económica para los sistemas de atención de la salud. La osteoporosis se define clínicamente a través de la medición de la densidad mineral ósea (DMO), que sigue siendo el mejor predictor de fractura<sup>1,2</sup>. Estudios de heredabilidad utilizando gemelos o familias han demostrado que el 50-85% de la variación en la densidad mineral ósea está determinada genéticamente<sup>3</sup>. Las fracturas osteoporóticas también muestran heredabilidad independiente de la densidad mineral ósea<sup>4</sup>.

Los estudios de asociación en genoma completo (GWAs) han ampliado muchísimo la comprensión de la arquitectura genética de enfermedades comunes y complejas<sup>5</sup>. Esta aproximación genómica está proporcionando información clave sobre los mecanismos de la enfermedad, con perspectivas para el diseño de estrategias más eficaces de evaluación del riesgo de enfermedad y para el desarrollo de nuevas intervenciones terapéuticas<sup>6</sup>. Sin embargo, las variantes genéticas que se identifican en los GWAs se encuentran con frecuencia en las zonas no-codificantes del genoma, cuya posible función es menos conocida y en muchos casos estas señales pueden estar en desequilibrio de ligamiento con variantes causales que no han sido genotipadas. El meta-análisis de GWAs para DMO y fractura osteoporótica de Estrada *et al.*<sup>7</sup> identificó hasta 56 *loci* genómicos asociados con la DMO, 14 de los cuales también estaban asociados con las fracturas osteoporóticas. Uno

de los SNPs cuya asociación con ambos fenotipos mostró una significación más sólida (rs4727338) se encuentra en una región intrónica del gen *FLJ42280*, señalándolo como un *locus* de susceptibilidad para la osteoporosis (Figura 1). Otros trabajos de GWA demostraron que otros SNPs intrónicos del mismo gen (rs7781370, rs10429035 y rs4729260) también se asociaban con la DMO<sup>8,9</sup>. *FLJ42280* es un gen muy poco estudiado, cuya relación con la biología del hueso se desconoce.

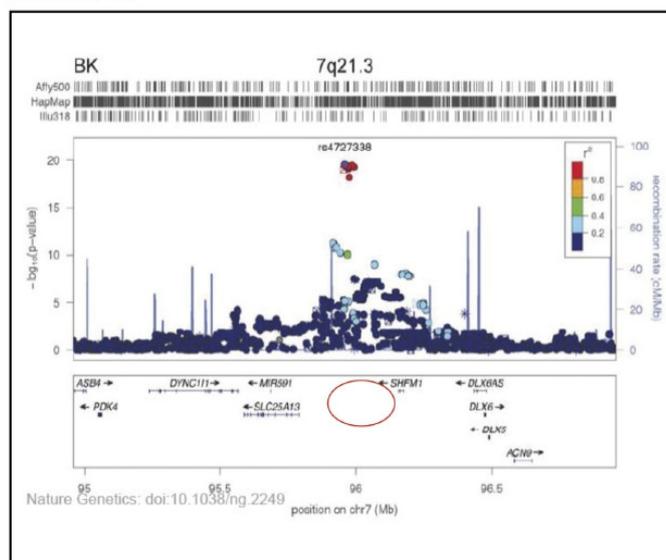
En este contexto, el objetivo de este trabajo fue dar sentido a esta asociación mediante la determinación de cuál es la variante causal. ¿Es rs4727338 el SNP causal o hay otro SNP en desequilibrio de ligamiento con él que sea el verdadero SNP funcional? Para ello, hemos explorado la variabilidad genética de la región genómica donde se encuentra el gen *FLJ42280* y hemos abordado la funcionalidad de estas variantes por enfoques diferentes. En primer lugar, por resecuenciación de la región en mujeres con DMO extremadamente alta o extremadamente baja para buscar variantes con una distribución desequilibrada entre los dos grupos; en segundo lugar, estudio bioinformático de la superposición de las variantes halladas con señales funcionales definidas en el proyecto ENCODE (*The Encyclopedia of DNA Elements*) y finalmente evaluación del posible papel como eQTLs de algunas de las variantes halladas.

### Material y métodos

#### Selección de la muestra de estudio

La muestra de este estudio está formada por 100 mujeres de la cohorte BARCOS<sup>10</sup>. Esta cohorte está compuesta por unas 1.500 mujeres postmenopáu-

Figura 1. Región genómica 7q21.3 con las señales de asociación a DMO de los SNPs testados por Estrada *et al.*<sup>7</sup>. La coordenada entre la posición genómica del SNP rs4727338 (eje x) y el  $-\log_{10}$  del valor de  $p$  (eje y) de su asociación con la DMO (eje y) está marcada con un cuadro de línea roja. Dicho SNP muestra la mayor significación en esta región. Los puntos de colores son las coordenadas del resto de SNPs estudiados en la región. Cada color indica un grado distinto de desequilibrio de ligamiento entre cada SNP y el SNP rs4727338. El gen *FLJ42280* no está mostrado porque, en el momento en que se realizó el meta-análisis de GWAs, este gen todavía no estaba anotado en el genoma. Su localización entre *SLC25A13* y *SHFM1* se indica con un óvalo rojo. Esta figura es una modificación de la que se presenta en Estrada *et al.*<sup>7</sup>



sicas españolas monitorizadas en el Hospital del Mar de Barcelona. Las mujeres diagnosticadas de osteomalacia, de enfermedad de Paget, de algún trastorno metabólico o endocrino, o que estuvieran siguiendo una terapia de sustitución hormonal o tratadas con fármacos que pudieran afectar la masa ósea, fueron excluidas de la cohorte. Las mujeres con una menopausia temprana (antes de los 40 años) también fueron excluidas. La información recogida para cada muestra fue la DMO, la edad, la edad de menarquia, la edad de menopausia, los años desde la menopausia, el peso y la estatura. De cada paciente se obtuvieron muestras de sangre y consentimientos informados escritos, según las regulaciones del Comité Ético de Investigación Clínica del Parque de Salud Mar. La DMO ( $\text{g}/\text{cm}^2$ ) fue medida en el cuello del fémur y en la columna lumbar. Se utilizó un densitómetro de rayos X de energía dual para realizar las medidas.

Se seleccionaron dos grupos de 50 muestras con valores de DMO extremos, según el valor del *Z-score*. Concretamente, los grupos consistieron en las 50 muestras con el *Z-score* más alto (rango: de 2,98 a 0,73) y las 50 muestras con el *Z-score* más bajo (rango: de -2,41 a -4,26) de la cohorte BARCOS.

### Preparación de las muestras genómicas

El ADN de cada mujer se extrajo a partir de muestras de sangre periférica. La concentración y la calidad de las muestras de ADN (*ratios* 260/280 y 260/230) se midieron por espectrofotometría en un aparato NanoDrop ND-1000 (NanoDrop Products). Para determinar la integridad del ADN, se analizaron 5  $\mu\text{l}$  de cada muestra mediante electroforesis en gel de agarosa al 1%. Finalmente, las muestras se normalizaron a una concentración de 100 ng/ $\mu\text{l}$ .

### Long-Range PCR (LR-PCR)

Se dividió una región genómica de 28 kb (que contiene el gen *FLJ42280* [22 kb] junto con 3,8 kb de región flanqueante a 5' y 2 kb de región flanqueante a 3') en 7 fragmentos solapantes (Figura 2). Los tamaños y las coordenadas de estos 7 fragmentos y los pares de cebadores utilizados para amplificarlos, se muestran en la tabla 1. Los fragmentos, de entre 2 y 5 kb, se amplificaron mediante LR-PCR. Cada reacción de LR-PCR incluyó: 100 ng de ADN genómico, 5  $\mu\text{l}$  de tampón *Ex Taq* "Magnesium +" (20 mM  $\text{Mg}^{2+}$ ; Takara) x10, mezcla de dNTPs (a 2,5 mM cada uno), *Ex Taq* polimerasa (5 U/ $\mu\text{l}$ ) y cebadores (20  $\mu\text{M}$ ), en un volumen final de 50  $\mu\text{l}$ .

Las reacciones se llevaron a cabo en un termociclador GeneAmp® PCR System 2700 (Applied Biosystems). Cada fragmento requirió unas condiciones de tiempo de elongación y temperatura de hibridación distintas. El número total de amplicones fue de 700 (100 muestras x 7 fragmentos). Se comprobó la cantidad y calidad de todos los amplicones mediante electroforesis en gel de agarosa al 1% p/v en tampón TBE x1.

### Purificación y cuantificación de las muestras

Para eliminar los restos de los reactivos de la PCR, los productos de PCR se purificaron utilizando placas de filtro de 96 pocillos con un tamaño de poro adecuado (Pall Corporation). Se aplicó el vacío (Vacuum Manifold, Merck Millipore) y el ADN retenido en el filtro se resuspendió en 35  $\mu\text{l}$  de agua milliQ. A continuación, los productos de PCR se cuantificaron mediante Quant-iT PicoGreen dsDNA Reagent and Kit (Life Technologies), según las instrucciones del fabricante. Brevemente, se construyó una curva estándar de concentraciones mediante medidas de emisión de fluorescencia a 520 nm después de haber excitado el ADN a 480 nm. La curva se utilizó posteriormente para calcular la concentración del ADN de las muestras.

### Normalización de las concentraciones de las muestras y pooling

Las muestras en las placas se normalizaron a una concentración de 5 ng/μl y, a continuación, se mezcló en un solo tubo 5 μl de cada muestra de una placa (un tubo por placa) mediante el The epMotion® 5075 Liquid Handling Workstation (Eppendorf). Así, se obtuvieron 14 tubos con 250 μl cada uno, dos para cada fragmento de PCR (DMO alta y DMO baja). Los 14 pools se concentraron hasta 5 veces utilizando el Genevac EZ-2 evaporator (Genevac SP Scientific) y se cuantificó cada tubo mediante Qubit® 2.0 Fluorometer (Life Technologies). Finalmente, los fragmentos de PCR se mezclaron equimolarmente en dos tubos, uno para la DMO alta y otro para la DMO baja.

### Secuenciación masiva en paralelo

La secuenciación masiva de las muestras se llevó a cabo en el Servicio de Genómica de los Centros Científicos y Tecnológicos de la Universidad de Barcelona, utilizando el GS 454 Junior System (Roche). Brevemente, se fragmentó el ADN por nebulización, se prepararon dos librerías marcadas con adaptadores (secuencias de 10 nucleótidos) distintos, una para cada grupo, que se mezclaron en un solo tubo. Seguidamente, se amplificó la mezcla por PCR de emulsión y la librería final se cargó en una placa *picotiter* donde se llevó a cabo la pirosecuenciación. Se realizaron 4 carreras de secuenciación, correspondientes a 140 Mb de datos finales (35 Mb/carrera). Este volumen de datos proporciona una cobertura teórica de 40x para cada muestra inicial.

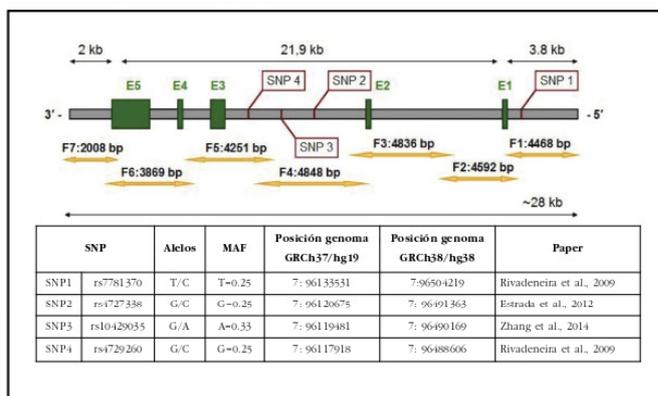
### Procesamiento de los datos de secuenciación y selección de variantes

Las lecturas obtenidas de la secuenciación fueron preprocesadas en base a su calidad y se alinearon contra el genoma de referencia (GRCh37) utilizando el programa GS Mapper (Roche). Las lecturas se indexaron y filtraron utilizando SAMtools. Se detectaron las variantes presentes en los dos grupos mediante GATK utilizando parámetros de filtrado estándar<sup>11</sup>. Las variantes encontradas se priorizaron según los siguientes criterios: se seleccionaron las variantes con una cobertura de al menos 1.000 lecturas, presentes en un 1% de las lecturas y con un *strand bias* bajo. El número de lecturas de las variantes que pasaron los filtros fue normalizado por la cobertura y las variantes fueron clasificadas entre comunes (con una frecuencia mayor al 5%) y raras o de baja frecuencia (con una frecuencia menor al 5%).

### Análisis funcional y estadístico de las variantes

Se compararon las frecuencias de cada variante entre los dos grupos mediante un test exacto de

Figura 2. En la región genómica de *FLJ42280* estudios previos de GWA han identificado 4 SNPs que muestran asociación con la DMO, cuyos detalles se muestran en la tabla incluida en la figura. Para conocer mejor la variabilidad de este *locus*, la región se subdividió en 7 fragmentos solapantes (flechas amarillas) para su resecuenciación en mujeres con muy alta o muy baja DMO



Fisher, aplicando la corrección de Bonferroni para comparaciones múltiples. El análisis funcional de las variantes consistió en mirar si estaban descritas en bases de datos como dbSNP y 1000 Genomas y, en caso afirmativo, buscar su MAF en la población europea e ibérica. Además, para las variantes exónicas se observó qué cambio de aminoácido suponían y su severidad predicha por SIFT, PolyPhen y Provean. Para las variantes intrónicas, se analizó la región que contiene la variante: lugares de hipersensibilidad a la DNasa, unión de factores de transcripción, metilación del ADN, modificaciones de histonas, regiones reguladoras, etc. Todos estos datos fueron obtenidos de bases de datos y repositorios como Ensembl, UCSC Genome Browser, ENCODE, BioMart, MatInspector. También se utilizó HaploReg para buscar anotaciones de regulación. Finalmente, todas las variantes encontradas fueron analizadas con el *Variant Effect Predictor* de Ensembl y UCSC y con el *SNP function prediction* del Instituto Nacional de Ciencias de la Salud Ambiental (*National Institute of Environmental Health Sciences*) de los EE.UU.

### Análisis del desequilibrio de ligamiento

Para calcular el desequilibrio de ligamiento entre todas las variantes de la región genómica de *FLJ42280* se utilizaron los genotipos de los SNPs presentes en la región y los haplotipos de los individuos de HapMap fase 3. Para calcular tal desequilibrio y generar un gráfico se utilizó el software HaploView.

### Análisis de eQTLs

Los SNPs que dieron significativos en los distintos GWAs y el SNP rs4613908 fueron evaluados como posibles eQTLs mediante dos aproximaciones: utilizando el portal del proyecto GTEx y utilizando los genotipos de esos SNPs en individuos de HapMap y los niveles de expresión de genes en *cis* en los

Tabla 1. Amplicones utilizados para secuenciar la región *FLJ42280*

Frag	Primers	Tamaño (pb)	Coordenadas genómicas
1	<b>1F</b> TTGACCTGAATACTGCGC <b>1R</b> GCCAAATGAATGTGGACAAG	4.468	7:96136619-96132152
2	<b>2F</b> CACTGCTGGGTCTTAGATTGG <b>2R</b> GCATGTGTGCATGATGTTGG	4.592	7:96132302-96127711
3	<b>3F</b> TGCAAGTTTCCCTCAATTCATC <b>3R</b> TCCCTCTCATCTGTGCAACAC	4.836	7:96127863-96123028
4	<b>4F</b> TTAGGTGAGTAGAAAGCAATGGC <b>4R</b> CTGGGTGGCTATAGACCTGAATAG	4.848	7:96123158-96118311
5	<b>5F</b> GCGGCACTGTGAGAGTACATC <b>5R</b> CCTGGTGAAAATGGGAACA	4.251	7:96118477-96114227
6	<b>6F</b> CTGACACTTTGGCAGCACC <b>6R</b> GGGATTGTTGAAGCTGACCC	3.869	7:96114348-96110480
7	<b>7F</b> CAACCATCACAACCCATAGAC <b>7R</b> CCTGAGCAAGTCTCGTAAAGTG	2.008	7:96110702-96108695

mismos individuos. Concretamente, se obtuvieron los genotipos de los SNPs de 210 individuos no emparentados de la fase 1 y 2 de HapMap y los niveles de expresión de los genes *SHFM1*, *SLC25A13* y *DLX5* de una línea celular linfoblastoide de los mismos individuos obtenidos.

## Resultados

### Variantes halladas y pistas sobre su función

La región genómica de *FLJ42280* (28 kb) se secuenció masivamente en dos *pools* de DNA correspondientes a las 50 mujeres con mayor DMO y las 50 mujeres con menor DMO de la cohorte BARCOS (ver detalles en Material y Métodos), a una alta profundidad (alrededor 3.600x por grupo). Se comparó el número y la frecuencia de las variantes que se encontraron en cada grupo. Se identificó un total de 110 variantes, de las cuales 18 eran nuevas y 59 fueron variantes raras o de baja frecuencia (Tabla 2). Se observó que el número de variantes de baja frecuencia entre los dos grupos extremos era equilibrado. Así mismo, se observó que las diferencias de frecuencia de todas las variantes estaban por debajo de la potencia estadística del diseño, aunque 9 mostraron una tendencia.

Para cada variante, se analizó su superposición con elementos funcionales anotados en el genoma por el proyecto ENCODE. Cuatro de las variantes solaparon con posibles secuencias potenciadoras de la transcripción (o *enhancers*) de osteoblastos y de ellas una [SNP rs4613908; MAF(CEU)=0,39] solapó con un enhancer activo en osteoblastos (Figura 3).

### Análisis de desequilibrio de ligamiento

También se estudió el desequilibrio de ligamiento entre todas las variantes comunes en esta región. Se creó un gráfico de desequilibrio de ligamiento (LD)

utilizando HaploView e información de haplotipos del proyecto HapMap (Figura 4) y se observó que hay un gran bloque de LD que incluye casi todo el gen (a excepción de la región 3' UTR) y que por la parte *upstream* del gen se extiende 5 kb más allá de la región resecuenciada. También se constató que los SNPs rs4613908 y rs4727338 (meta-análisis de GWAs de Estrada *et al.*) presentan una gran desequilibrio de ligamiento entre ellos.

### Análisis de eQTLs

Para completar el análisis funcional, se realizó un análisis de eQTLs. Disponiendo de los genotipos de los cuatro SNPs asociados a la DMO, y del SNP rs4613908 de 210 individuos del proyecto HapMap y de los niveles de expresión génica de un *array* genómico en líneas linfoblastoides de estos mismos individuos, se determinó si los diferentes alelos o genotipos de los SNPs correlacionaban con los niveles de expresión génica de los genes situados en la región genómica de *FLJ42280*. Ninguno de los SNPs mostró influencia sobre los niveles de expresión de los genes *SHFM1*, *SLC25A13* o *DLX5* (en el *array* no hay información de niveles de expresión de *FLJ42280*). También se accedió a la base de datos GTEx para recabar información de eQTL para los mismos SNPs y el resultado fue negativo para todos ellos. Finalmente, se realizaron búsquedas de anotaciones de regulación en HaploReg. Este último análisis confirmó que la secuencia que rodea el SNP rs4613908 está altamente conservada entre los mamíferos y que en varios tipos celulares, incluyendo los osteoblastos primarios, contiene marcas de cromatina típicas de secuencias potenciadoras (H3K4me1, H3K27ac). Por otra parte, HaploReg destacó la alteración de motivos reguladores de este SNP y de rs10429035, pero no mostró ningún efecto de estos SNPs sobre la expresión génica.

Tabla 2. Número y localización de variantes de un solo nucleótido halladas en este estudio

	Crudo	Filtrado	Codificante	Región reguladora*	Putativo enhancer de osteoblasto	Enhancer activo en osteoblasto
Variantes comunes	96	51	0	12	3	1
LFV	24.243	59	1	16	1	0
Total	24.339	110	1	28	4	1

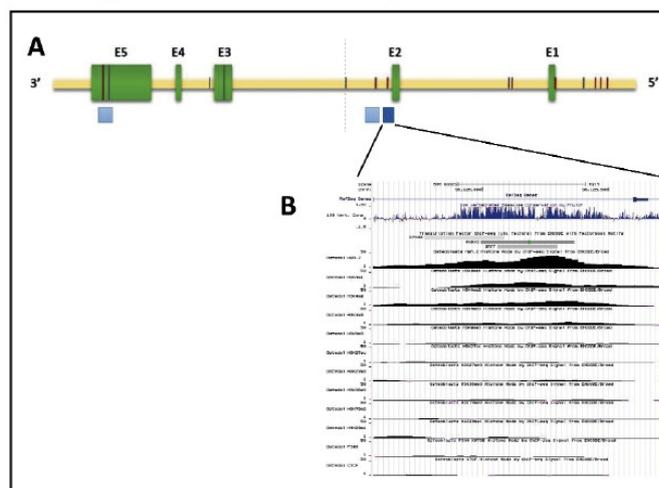
\*incluye regiones flanqueantes, 5'UTR, 3'UTR e intrones; LFV: variante de baja frecuencia (MAF<5%).

### Discusión

Se ha realizado un barrido exhaustivo de una región genómica (28 kb) en 7q21.3 que contiene varias señales muy fuertes de asociación entre 4 SNPs y la densidad mineral ósea<sup>7-9</sup>. Se ha querido conocer todas las variantes puntuales presentes en regiones codificantes (exones del gen *FLJ42280*) y no codificantes (intrones, 3'UTR, 5'UTR y flancos del gen) y evaluar el potencial funcional de estas variantes para pronosticar cuáles de ellas podrían ser las responsables de la asociación con la DMO. Se ha observado que la variante rs4613908 solapa con un potenciador génico (*enhancer*) activo en osteoblastos contenido en una secuencia con elevada conservación evolutiva. Dicho SNP (con sus dos variantes alélicas) podría estar afectando a la DMO por el hecho de alterar este potenciador génico. Queda por determinar cuál es el gen diana del mencionado potenciador.

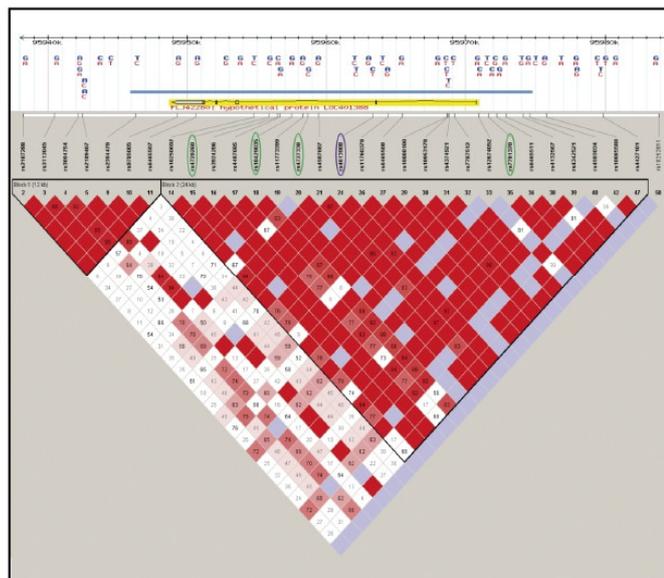
Hasta la fecha, no nos constan otros trabajos de otros autores que hayan abordado la base funcional de la asociación con DMO de los SNPs situados en regiones no codificantes del gen *FLJ42280*. De hecho, este gen ha sido anotado recientemente en el genoma humano, de modo que cuando se detectó la asociación de los SNPs de la región, el gen todavía no constaba en el mapa de 7q21.3 y los SNPs quedaban entre los genes *SLC25A13* y *SHFM1* (Figura 1). Por ello, Estrada *et al.*<sup>7</sup> propusieron que la funcionalidad de la asociación podía estar relacionada con *SLC25A13*. Actualmente, *FLJ42280* sigue siendo un gen anotado, con muy pocos datos experimentales que lo confirmen. Es pues muy probable que la función de los SNPs asociados a la DMO esté relacionada con otros genes. En este sentido, el gen *SHFM1* se ha asociado a algunos casos hereditarios de malformación de mano hendida-pie hendido (*Split hand and foot malformation 1*; OMIM #183600) y el gen *DLX5*, situado a continuación, es de hecho el gen res-

ponsable de dicha enfermedad, dado que existen pacientes con mutaciones puntuales en *DLX5* que cosegregan con la enfermedad<sup>12</sup>. Se han descrito una serie de potenciadores que afectan a la expresión de *DLX5* en distintos tejidos y estadios del desarrollo y que se distribuyen a lo largo de varios cientos de kilobases. Estudios realizados en ratones y pez cebra han caracterizado estos potenciadores y han mostrado que funcionan durante el desarrollo<sup>13,14</sup>. Algunos de ellos muestran especificidad de tejido y correlacionan con determinados fenotipos presentes en pacientes con malformación de mano hendida-pie hendido portadores de varias anomalías cromosómicas (deleciones o translocaciones) que afectan a los potenciadores mencionados. Al colocar estos potenciadores de *DLX5* sobre el mapa de la región 7q21.3, hemos visto con sorpresa que el SNP rs4613908, que acabamos de comentar como buen candidato funcional, se



ponsable de dicha enfermedad, dado que existen pacientes con mutaciones puntuales en *DLX5* que cosegregan con la enfermedad<sup>12</sup>. Se han descrito una serie de potenciadores que afectan a la expresión de *DLX5* en distintos tejidos y estadios del desarrollo y que se distribuyen a lo largo de varios cientos de kilobases. Estudios realizados en ratones y pez cebra han caracterizado estos potenciadores y han mostrado que funcionan durante el desarrollo<sup>13,14</sup>. Algunos de ellos muestran especificidad de tejido y correlacionan con determinados fenotipos presentes en pacientes con malformación de mano hendida-pie hendido portadores de varias anomalías cromosómicas (deleciones o translocaciones) que afectan a los potenciadores mencionados. Al colocar estos potenciadores de *DLX5* sobre el mapa de la región 7q21.3, hemos visto con sorpresa que el SNP rs4613908, que acabamos de comentar como buen candidato funcional, se

Figura 4. Gráfico del desequilibrio de ligamiento presente en la región genómica de *FLJ42280*. Se observa un bloque grande (24 kb) que incluye gran parte del gen excepto su región 3'UTR. Los 4 SNPs que se han asociado a DMO en los distintos GWAs están señalados con óvalos verdes. El SNP que solapa con el potenciador activo en osteoblastos está señalado con un óvalo morado. La línea horizontal azul indica la región genómica resecuenciada en este estudio



encuentra en uno de estos potenciadores (eDLX#18), situado a unas 500 kb de *DLX5*. El potenciador eDLX#18 se ha descrito como activo en los arcos branquiales en estadios embrionarios<sup>13</sup>.

Existen evidencias de que *DLX5* está involucrado en la determinación de la DMO<sup>15</sup>, lo que nos hace proponer la hipótesis de que el potenciador eDLX#18 también es activo como un potenciador para *DLX5* en osteoblastos de adultos y que nuestro SNP de interés es un eQTL en osteoblastos. Será muy necesario comprobar esta hipótesis mediante análisis de expresión de *DLX5* en osteoblastos primarios y genotipación de rs4613908 de los mismos.

**Conflicto de intereses:** Los autores declaran no tener conflicto de intereses en relación con este trabajo.

**Agradecimientos:** Este trabajo se ha llevado a cabo gracias a la ayuda económica brindada por una beca FEIOMM de 2014.

## Bibliografía

- Johnell O, Kanis JA, Oden A, Johansson H, De Laet C, Delmas P, et al. Predictive Value of BMD for Hip and Other Fractures. *J Bone Miner Res.* 2005;20:1185-94.
- Kanis JA, Oden A, Johnell O, Johansson H, De Laet C, Brown J, et al. The use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. *Osteoporos Int.* 2007;18:1033-46.
- Peacock M, Turner CH, Econs MJ, Foroud T. Genetics of Osteoporosis. *Endocr Rev.* 2002;23:303-26.
- Deng HW, Chen WM, Recker S, Stegman MR, Li JL, Davies KM, et al. Genetic determination of Colles' fracture and differential bone mass in women with and without Colles' fracture. *J Bone Miner Res.* 2000;15:1243-52.
- Hardy J, Singleton A. Genomewide Association Studies and Human Disease. *N Engl J Med.* 2009;360:1759-68.
- Manolio TA. Genomewide Association Studies and Assessment of the Risk of Disease. *N Engl J Med.* 2010;363:166-76.
- Estrada K, Styrkarsdottir U, Evangelou E, Hsu YH, Duncan EL, Ntzani E, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet.* 2012;44:491-501.
- Rivadeneira F, Styrkarsdottir U, Estrada K, Halldórsson BV, Hsu YH, Richards JB, et al. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat Genet.* 2009;41:1199-206.
- Zhang LL, Choi HJ, Estrada K, Leo PJ, Li J, Pei YF, et al. Multistage genome-wide association meta-analyses identified two new loci for bone mineral density. *Hum Mol Genet.* 2014;23:1923-33.
- Bustamante M, Nogués X, Enjuanes A, Elosua R, García-Giralt N, Pérez-Edo L, et al. COL1A1, ESR1, VDR and TGFBI polymorphisms and haplotypes in relation to BMD in Spanish postmenopausal women. *Osteoporos Int.* 2007;18:235-43.
- DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43:491-8.
- Shamseldin HE, Faden MA, Alashram W, Alkuraya FS. Identification of a novel *DLX5* mutation in a family with autosomal recessive split hand and foot malformation. *J Med Genet.* 2012;49:16-20.
- Birnbaum RY, Everman DB, Murphy KK, Gurrieri F, Schwartz CE, Ahituv N. Functional characterization of tissue-specific enhancers in the *DLX5/6* locus. *Hum Mol Genet.* 2012;21:4930-8.
- Rasmussen MB, Kreiborg S, Jensen P, Bak M, Mang Y, Lodahl M, et al. Phenotypic subregions within the split-hand/foot malformation 1 locus. *Hum Genet.* 2016;135:345-57.
- Ulsamer A, Ortuño MJ, Ruiz S, Susperregui AR, Osses N, Rosa JL, et al. BMP-2 Induces Osterix Expression through Up-regulation of *Dlx5* and Its Phosphorylation by p38. *J Biol Chem.* 2008;283:3816-26.



**Roca-Ayats N<sup>1</sup>, Falcó-Mascaró M<sup>1</sup>, García-Giralt N<sup>2</sup>, Cozar M<sup>1</sup>, Abril JF<sup>3</sup>, Quesada-Gómez JM<sup>4</sup>, Prieto-Alhambra D<sup>5,6</sup>, Nogués X<sup>2</sup>, Mellibovsky L<sup>2</sup>, Díez-Pérez A<sup>2</sup>, Grinberg D<sup>1</sup>, Balcells S<sup>1</sup>**

<sup>1</sup> Departamento de Genética, Microbiología y Estadística - Facultad de Biología - Universidad de Barcelona - Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER) - Instituto de Salud Carlos III (ISCIII) - Instituto de Biomedicina de la Universidad de Barcelona (IBUB) - Instituto de Investigación Sant Joan de Déu (IRSJD) - Barcelona (España)

<sup>2</sup> Unidad de Investigación en Fisiopatología Ósea y Articular (URFOA); Instituto Hospital del Mar de Investigaciones Médicas (IMIM) - Parque de Salud Mar - Centro de Investigación Biomédica en Red de Fragilidad y Envejecimiento Saludable (CIBERFES); Instituto de Salud Carlos III (ISCIII) - Barcelona (España)

<sup>3</sup> Departamento de Genética, Microbiología y Estadística; Facultad de Biología; Universidad de Barcelona - Instituto de Biomedicina de la Universidad de Barcelona (IBUB) - Barcelona (España)

<sup>4</sup> Unidad de Metabolismo Mineral; Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC); Hospital Universitario Reina Sofía - Centro de Investigación Biomédica en Red de Fragilidad y Envejecimiento Saludable (CIBERFES); Instituto de Salud Carlos III (ISCIII) - Córdoba (España)

<sup>5</sup> Grupo de Investigación en Enfermedades Prevalentes del Aparato Locomotor (GREMPAL) - Instituto de Investigación en Atención Primaria (IDIAP) Jordi Gol - Centro de Investigación Biomédica en Red de Fragilidad y Envejecimiento Saludable (CIBERFES) - Universidad Autónoma de Barcelona (UAB) - Barcelona (España)

<sup>6</sup> Instituto Nacional para la Investigación de la Salud (NIHR); Unidad de Investigación Biomédica (BRU) Musculoquelética y Centro de Investigación Botnar; Departamento de Ortopedia, Reumatología y Ciencias Musculoqueléticas Nuffield - Universidad de Oxford - Oxford (Reino Unido)

## Estudio genético de la fractura femoral atípica mediante la secuenciación del exoma en tres hermanas afectas y tres pacientes no relacionadas

DOI: <http://dx.doi.org/10.4321/S1889-836X2018000400002>

Correspondencia: Neus Roca-Ayats - Departamento de Genética, Microbiología y Estadística - Facultad de Biología - Universidad de Barcelona - Avda. Diagonal, 643 - 08028 Barcelona (España)  
Correo electrónico: [neus.roca@ub.edu](mailto:neus.roca@ub.edu)

Fecha de recepción: 12/06/2018

Fecha de aceptación: 24/09/2018

*Trabajo becado asistir al 39º Congreso de la ASBMR (Denver, 2017).*

### Resumen

**Objetivos:** Las fracturas atípicas de fémur (FAF) son un tipo de fracturas poco frecuentes, a menudo relacionadas con un tratamiento prolongado con bisfosfonatos (BPs). Actualmente no se conocen con exactitud sus mecanismos patogénicos y no hay pruebas para identificar aquellos pacientes con un alto riesgo de sufrir una FAF. El objetivo de este trabajo es investigar las bases genéticas de las FAFs.

**Material y métodos:** Se secuenció el exoma completo de 3 hermanas y de 3 pacientes adicionales no relacionadas, todas tratadas con BPs durante más de 5 años. Se seleccionaron variantes compartidas por las hermanas, de baja frecuencia y potencialmente patogénicas, y se construyó una red de interacciones de genes y proteínas con los datos hallados.

**Resultados:** Identificamos 37 variantes raras (en 34 genes) compartidas por las 3 hermanas, algunas de ellas no descritas anteriormente. La variante más llamativa fue la mutación p.Asp188Tyr en el enzima geranilgeranil pirofosfato sintasa (codificada por el gen *GGPS1*), de la vía del mevalonato y esencial para la función del osteoclasto. Otro hallazgo interesante fueron dos mutaciones (una en las 3 hermanas y una en una paciente no relacionada) en el gen *CYP11A1*, implicado en el metabolismo de los esteroides. Identificamos otras variantes que también podrían estar involucradas en la susceptibilidad a las FAFs o en el fenotipo osteoporótico subyacente, tales como las presentes en los genes *SYDE2*, *NGEF*, *COG4* y la *FNI*.

**Conclusiones:** Nuestros datos son compatibles con un modelo donde la acumulación de variantes de susceptibilidad podría participar en la base genética de las FAFs.

**Palabras clave:** *fractura atípica de fémur, bisfosfonatos, GGPS1, CYP11A1, secuenciación completa del exoma.*

## Genetic study of atypical femoral fractures using exome sequencing in three affected sisters and three unrelated patients

### Summary

**Objectives:** Atypical femoral fractures (AFF) are rare, often related to long-term bisphosphonate (BPs) treatment. Their pathogenic mechanisms are not precisely known and there is no evidence to identify patients with a high risk of AFF. The aim of this work is to study the genetic bases of AFFs.

**Material and methods:** Whole-exome sequencing was carried out on 3 sisters and 3 unrelated additional patients, all treated with BPs for more than 5 years. Low frequency, potentially pathogenic variants shared by the 3 sisters, were selected and a network of gene and protein interactions was constructed with the data found.

**Results:** We identified 37 rare variants (in 34 genes) shared by the 3 sisters, some not previously described. The most striking variant was the p.Asp188Tyr mutation in the enzyme geranylgeranyl pyrophosphate synthase (encoded by the *GGPS1* gene), from the mevalonate pathway and essential for osteoclast function. Another noteworthy finding was two mutations (one in the 3 sisters and one in an unrelated patient) in the *CYP11A1* gene, involved in the metabolism of steroids. We identified other variants that could also be involved in the susceptibility to AFFs or in the underlying osteoporotic phenotype, such as those present in the *SYDE2*, *NGEF*, *COG4* and *FN1* genes.

**Conclusions:** Our data are compatible with a model where the accumulation of susceptibility variants could participate in the genetic basis of AFFs.

**Key words:** atypical femoral fractures, bisphosphonates, *GGPS1*, *CYP11A1*, whole-exome sequencing.

### Introducción

La osteoporosis y sus fracturas asociadas son el problema óseo postmenopáusico más común, y afecta a mujeres y hombres de todas las etnias. Los bisfosfonatos nitrogenados (N-BPs), incluyendo alendronato, risendronato, ibandronato y zolendronato, son el tratamiento más utilizado para la osteoporosis en millones de pacientes en todo el mundo. A pesar de la importante eficacia anti-fractura de los BPs, ampliamente demostrada en varios ensayos clínicos<sup>1</sup> y revisiones sistemáticas<sup>2</sup>, se han descrito algunos efectos adversos poco frecuentes potencialmente asociados a su uso prolongado, entre ellos las fracturas atípicas de fémur (FAFs)<sup>3</sup>. Estas fracturas son no-traumáticas y están caracterizadas por su localización subtrocantérica o en la diáfisis del fémur, y frecuentemente son bilaterales<sup>4</sup>.

Los mecanismos patogénicos de las FAFs no son del todo conocidos, y se ha especulado mucho sobre sus causas. Se ha propuesto que una supresión excesiva de la resorción ósea por parte de los N-BPs podría contribuir a desencadenar una FAF pero su fisiopatología es compleja y se cree que hay otros factores importantes involucrados. Algunos factores de riesgo propuestos son el grosor cortical y la geometría pélvica<sup>5</sup>. Además, se han descrito casos de FAF en pacientes afectados por otras enfermedades óseas monogénicas, como la hipofosfatasa<sup>6</sup>, la osteogenesis imperfecta<sup>7</sup> o el síndrome de osteoporosis pseudoglioma<sup>8</sup>.

Dada la baja incidencia de las FAFs en la población general (5,9 casos por 100.000 personas/año), podemos hipotetizar que hay unas causas genéticas raras subyacentes que pueden incrementar la susceptibilidad a las FAFs, y que pueden ocurrir espontáneamente o desencadenarse después de la interacción con los BPs. Actualmente

no hay pruebas genéticas o bioquímicas que puedan ayudar a identificar los pacientes con un elevado riesgo a sufrir una FAF. La identificación de los determinantes genéticos de las FAFs ayudaría a esclarecer los mecanismos etiológicos, al desarrollo de herramientas de diagnóstico y de evaluación del riesgo de sufrir una FAF, y a posibles estrategias terapéuticas.

Anteriormente, identificamos 3 hermanas diagnosticadas con FAF que fueron tratadas con BPs durante más de 5 años<sup>9</sup>. Esta observación nos sugirió que podría haber un trasfondo genético que predispusiera a las FAFs relacionadas al uso prolongado de BPs. En consecuencia, llevamos a cabo la secuenciación del exoma completo de las 3 hermanas y de otras 3 pacientes no relacionadas para identificar mutaciones potencialmente relacionadas con las FAFs en estas pacientes. Identificamos 37 variantes raras compartidas por las 3 hermanas, una de las cuales se estudió en detalle<sup>9</sup>. En el presente trabajo describimos el conjunto de variantes encontradas y su posible interacción.

### Material y métodos

#### Pacientes

Se estudiaron seis pacientes con FAFs y que habían sido tratadas durante más de 5 años con BPs: 3 hermanas visitadas en el Hospital Universitario Reina Sofía (Córdoba, España) y 3 pacientes no relacionadas visitadas en el Hospital del Mar (Barcelona, España). Como controles, se estudiaron 3 pacientes tratadas con BPs por más de 6 años pero sin FAFs. Las características de pacientes y controles están descritas en la tabla 1. Las 3 hermanas afectas fueron tratadas con estatinas y recibían regularmente PPIs pero no habían sido tratadas con glucocorticoides ni ningún otro compuesto que afecte al

hueso, aparte de los BPs. En el caso de las fracturas unilaterales, se realizaron pruebas radiológicas y RMN que descartaban la fractura contralateral. Se obtuvo consentimiento informado escrito de todas las pacientes, de acuerdo con la regulación del Comité Ético de Investigación Clínica del Parque de Salud Mar, que aprobó el estudio.

#### Secuenciación del exoma completo

Se extrajo ADN de sangre periférica de las pacientes con el kit Wizard Genomic DNA Purification (Promega) y se utilizó para secuenciar el exoma completo en el Centro Nacional de Análisis Genómico (CNAG) (Barcelona). Las librerías se generaron con el kit de captura de exones SureSelect XT Human All Exon; cat:5190-6208 (Agilent Technologies), después de haber fragmentado el ADN y ligado los adaptadores específicos de Agilent. La secuenciación *paired-end* (2x76 pb) se realizó en la plataforma Illumina HiSeq2000. Las imágenes del instrumento se procesaron utilizando el programa del fabricante para generar archivos de secuencia FASTQ.

El análisis bioinformático se llevó a cabo en la plataforma de Bioinformática para Enfermedades Raras (Bier) del CIBERER, en Valencia. Los archivos FASTQ se alinearon con el programa libre Burrows-Wheeler Aligner<sup>10</sup> (<http://bio-bwa.sourceforge.net/>) utilizando el ensamblado del genoma humano de referencia GRCh37 (hg19)<sup>11</sup>. Las variantes de un solo nucleótido y los indels se identificaron utilizando el programa GATK<sup>12</sup>. Finalmente, para añadir a las variantes información sobre la frecuencia del alelo minoritario (*minor allele frequency*; MAF) proveniente de dbSNP y del proyecto 1000 Genomas ([\[www.1000genomes.org\]\(http://www.1000genomes.org\)\)<sup>13</sup>, se utilizó la herramienta de anotación VARIANT<sup>14</sup>. Los datos se convirtieron al formato BAM \(\*binary equivalent SAM\*\) y se visualizaron mediante el programa Integrative Genomics Viewer \(IGV\) \(<http://www.broadinstitute.org/igv>\).](http://</a></p>
</div>
<div data-bbox=)

Las variantes genéticas se filtraron según las siguientes premisas: a) variante no-sinónima, b) no descrita previamente o con una MAF <0,005 en dbSNP y en el proyecto 1000 Genomas, c) no presente en NHLBI Go Exome Sequencing Project (ESP) (<http://evs.gs.washington.edu/EVS/>), y d) no presente en 8 exomas de individuos de la población general, obtenidos en nuestro laboratorio.

Inicialmente sólo se tuvieron en cuenta las mutaciones compartidas por las tres hermanas, tanto en un modelo de herencia dominante como recesivo. Después se priorizaron mutaciones en genes candidatos en las otras tres pacientes. Las puntuaciones de SIFT<sup>15</sup>, PolyPhen<sup>16</sup> y de conservación evolutiva obtenidas de PhastCons<sup>17</sup> se utilizaron para priorizar las variantes.

#### Validación de las variantes genéticas

Las mutaciones encontradas se validaron mediante PCR y secuenciación Sanger, que fue llevada a cabo bidireccionalmente utilizando el kit BigDye™ v3.1 Terminator Cycle Sequencing (Applied Biosystems), según las instrucciones del fabricante. Los cebadores utilizados para la validación se diseñaron utilizando el programa OligoEvaluator (Sigma-Aldrich). Finalmente, las mutaciones validadas se buscaron en el Exome Aggregation Consortium (ExAC) para obtener sus frecuencias poblacionales, y se analizaron mediante secuenciación Sanger en las 3 mujeres controles.

Tabla 1. Características de pacientes y controles

Paciente	Fractura atípica	Edad <sup>a</sup> (años)	Peso (Kg)	T-score columna vertebral	T-score cadera	Tiempo de tratamiento con BPs (años)	Fracturas osteoporóticas previas
AFS1	Unilateral; medio-diafisaria <sup>b</sup>	64	77	-1,1	-0,2	6	Colles
AFS2	Unilateral; medio-diafisaria <sup>b</sup>	73	75	-2,5	-1,4	6	Colles
AFS3	Bilateral; medio-diafisaria <sup>b</sup>	60/61	100	-0,3	Rbpc <sup>c</sup>	6	Ninguna
AFU1	Bilateral; medio-diafisaria	73/75	50,8	-1,9	-0,5	6	Ninguna
AFU2	Unilateral; medio-diafisaria	72	90	-2,0	-0,6	7	Ninguna
AFU3	Unilateral; subtrocantérica	87	59,8	N/A	N/A	10	Ninguna
Control 1		78	66,5	-2,5	-1,9	7	Ninguna
Control 2		70	57,5	-1,2	-2,4	6	Ninguna
Control 3		74	77,1	-1,5	-0,9	8	Ninguna

AFS: hermanas con FAF; AFU: pacientes con FAF no relacionadas, (°): edad al momento de la fractura atípica; (°): fracturas localizadas aproximadamente en el mismo sitio; (°): reemplazo bilateral de prótesis cadera.

### Análisis *in silico*

Las mutaciones se localizaron en su contexto genético utilizando el UCSC Genome Browser (<https://genome.ucsc.edu/>) y el Ensembl Genome Browser (<http://www.ensembl.org/>) y se extrajo información de los genes de GeneCards (<http://www.genecards.org/>) y BioGPS (<http://biogps.org/>). Se realizó un análisis de enriquecimiento funcional utilizando la herramienta bioinformática DAVID<sup>18</sup> (<https://david.ncifcrf.gov/>).

El estudio funcional *in silico* de las proteínas mutadas se realizó utilizando Uniprot (<http://uniprot.org/>), RCSB Protein Data Bank (PDB) (<http://www.rcsb.org/pdb/>) y Pfam (<http://pfam.xfam.org/>). Los alineamientos de proteínas se realizaron utilizando el UCSC Genome Browser y los programas Clustal Omega (<http://www.clustal.org/omega>) y ESPript (<http://esprpt.ibcp.fr>).

### Construcción de la red

La red de interacción de los genes FAF (AFFGeNet) se construyó según Boloc *et al.*<sup>19</sup> para identificar genes o proteínas que interactúan con los 37 genes FAF, considerados como genes *driver* (Tablas 2a y 2b), teniendo en cuenta las interacciones binarias y direccionales. Los datos de interacción *high-throughput* se obtuvieron de BioGRID (versión 3.4.133)<sup>20</sup> y STRING [Search Tool for the Retrieval of Interacting Genes/Proteins] versión 10<sup>21</sup> y la red se enriqueció con información adicional de GeneOntology (<http://geneontology.org>), GeneCards, OMIM, UniProt, RefSeq, y UCSC.

Se implementó un *script* de Perl para capturar la sub-red de interacción utilizando los genes FAF para encontrar todos los caminos más cortos entre dos genes aplicando el algoritmo Dijkstra. La conectividad por parejas se analizó utilizando Circos<sup>22</sup>. El *script* produjo un gráfico esqueleto en formato JSON para poder visualizar los datos en la interficie web AFFGeNet (<https://compgen.bio.ub.edu/AFFGenes>, disponible bajo demanda). El formulario web contiene una entrada que se centra en los genes seleccionados, y la visualización de la red permite añadir o quitar nodos y mostrar información de los genes FAF. El color del borde identifica los nodos como *drivers* (lila), parejas *upstream* (verde) o *downstream* (azul) de los *drivers* seleccionados, y otros (gris). El color del interior de los nodos representa la expresión génica específica del hueso, que se obtuvo del Gene Expression Omnibus (GEO)<sup>23</sup>, concretamente de un estudio sobre células precursoras de osteoclastos tratadas o no tratadas con BPs (alendronato o risendronato) durante su diferenciación a osteoclasto maduro<sup>24</sup> (GSE63009). La escala de colores va de amarillo intenso (subexpresado) a azul oscuro (sobrexpresado), siendo el blanco indicativo de ningún cambio de expresión.

### Resultados

#### Variantes detectadas en la secuenciación del exoma completo en las 3 hermanas

Las tres hermanas (AFS1, AFS2, AFS3) y las 3 pacientes no relacionadas (AFU1, AFU2, AFU3) se analizaron separadamente.

Los exomas de las 3 hermanas se interseccionaron y no se identificó ninguna variante en homocigosis en común. Por el contrario, se identificaron 74 variantes en heterocigosis compartidas (coherentes con un modelo de herencia dominante), 37 de las cuales se validaron por secuenciación Sanger. En 3 de los genes (*FNI*, *BRATI* y *XAB2*), se encontraron 2 mutaciones diferentes. En los tres casos se pudo determinar que las variantes se encontraban en fase, siendo alelos doble-mutantes y no heterocigotos compuestos, mediante la visualización de los *reads* con el programa IGV y el análisis de polimorfismos intragénicos. Las 37 variantes compartidas por las 3 hermanas, todas ellas codificantes, se muestran en la tabla 2a, ordenadas según su puntuación de conservación. Se trata de variantes de cambio de sentido (n=35), una variante truncante y una delección en fase. La primera variante de la lista, con la mejor puntuación de conservación y predicha como deletérea, se encuentra en el gen *GGPS1*, tal y como describimos anteriormente<sup>9</sup>.

#### Análisis de los genes mutados en las 3 pacientes no relacionadas

Los genes con variantes compartidas por las 3 hermanas (Tabla 2a) se analizaron en los exomas de las pacientes no relacionadas utilizando el programa IGV. Ninguna de las variantes de la Tabla 2a se encontró en las pacientes no relacionadas. No obstante, se encontraron otras dos variantes en los genes *BRATI* y *CYP11A1*, en las pacientes AFU3 y AFU1, respectivamente (Tabla 2b).

La variante de *CYP11A1* presente en la paciente AFU1 (p.Ser216Cys) supone el cambio de una serina a una cisteína, en una posición cercana al sitio de unión al sustrato. Los predictores de patogenicidad sugirieron que este cambio es muy deletéreo para la función de la proteína. Igualmente, la variante de *CYP11A1* presente en las tres hermanas (p.Arg98Trp) supone el cambio de un aminoácido básico (arginina) a un aminoácido aromático hidrofóbico (triptófano), en un giro de la proteína con puentes de hidrógeno. Por el contrario, las tres variantes encontradas en el gen *BRATI* (dos en las tres hermanas, en un alelo doble mutante, y una en la paciente AFU3) no afectan a la función de la proteína, según los predictores.

#### Análisis de genes candidatos en 3 pacientes no relacionadas

A continuación, se utilizó el programa IGV para analizar, en los exomas de las tres pacientes no relacionadas, distintos genes involucrados en el metabolismo óseo, la función osteoclastica y la vía del mevalonato. Se encontraron variantes en los genes *MMP9* (AFU3), *MVD* (AFU2) y *RUNX2* (AFU3), que se validaron por secuenciación Sanger (Tabla 2b). La mutación en el gen *MMP9*, que codifica la colagenasa de tipo IV, implica el cambio de una metionina (un aminoácido hidrofóbico con un grupo que contiene azufre) a una treonina (aminoácido hidrofílico) en la posición 419, dentro del dominio catalítico. Esta variante aparece en la base de datos ExAC,

con una frecuencia alélica muy baja ( $8,2e-06$ ), y SIFT y PolyPhen predijeron que probablemente perjudica su función. El gen *MVD* codifica la enzima mevalonato 5-difosfato decarboxilasa, de la vía del mevalonato. La variante encontrada (p.Arg97Gln; rs376949804) supone el cambio de un aminoácido básico a un aminoácido neutro y está presente en la base de datos ExAC, también con una frecuencia alélica muy baja ( $3,4e-05$ ). Se trata de un cambio no

perjudicial para la función de la proteína, según SIFT y PolyPhen. La mutación en *RUNX2* es una substitución de una prolina, un aminoácido cíclico, por una leucina, un aminoácido alifático hidrofóbico, en la posición 296, dentro de una región rica en prolinas, serinas y treoninas. Este cambio, descrito en dbSNP (rs20184115), tiene una MAF= $0,0004$  y probablemente afecta la función de la proteína, según los predictores.

Tabla 2a. Variantes compartidas por las 3 hermanas, encontradas en la secuenciación del exoma

Gen	Proteína	Variante <sup>a</sup>	Efecto en la proteína	dbSNP <sup>b</sup>	ExAC <sup>c</sup>	Conser-vación <sup>d</sup>	SIFT <sup>e</sup>	PolyPhen <sup>f</sup>
<i>GGPS1</i>	Geranilgeranil difosfato sintasa	chr1:g.235505746G>T	p.D188Y			700	<b>0,000</b>	<b>1,000</b>
<i>LRRC1</i>	Proteína con repeticiones ricas en leucinas 1	chr6:g.53707020G>A	p.R91Q		4,946e-05	685	<b>0,050</b>	0,746
<i>TUSC2</i>	Candidato supresor de tumores 2	chr3:g.50363807T>C	p.H83R		8,244e-06	674	0,338	0,000
<i>SYDE2</i>	Proteína activadora de GTPasa Rho	chr1:g.85634903G>T	p.L893I		8,339e-06	639	<b>0,018</b>	<b>0,997</b>
<i>COG4</i>	Subunidad 4 del complejo oligomérico conservado del Golgi	chr16:g.70553552C>T	p.G85D			627	0,150	0,735
<i>EML1</i>	Proteína asociada a microtúbulos	chr14:g.100360993G>A	p.R211H		6,611e-05	588	<b>0,030</b>	<b>0,963</b>
<i>KDM4C</i>	Demetilasa específica de lisinas(K) 4C	chr9:g.6849579A>G	p.I170V	rs192832191 MAF=0,0004	2,471e-05	584	<b>0,000</b>	0,509
<i>ERCC6L2</i>	Proteína de reparación por escisión del DNA	chr9:g.98718284A>T	p.I657L		8,278e-06	573	0,630	0,007
<i>PGRMC1</i>	Componente de membrana 1 del receptor de progesterona	chrX:g.118377159C>A	p.P177H			573	0,130	0,742
<i>FN1</i> *	Fibronectina	chr2:g.216235149C>T	p.V2241I		8,245e-06	551	<b>0,009</b>	0,045
<i>CYP11A1</i>	Citocromo P450 11A1	chr15:g.75015147G>A	p.R98W		0,000108	540	<b>0,000</b>	<b>0,998</b>
<i>XAB2</i> *	Proteína de unión a XPA 2	chr19:g.7688142C>G	p.V385L		1,651e-05	535	<b>0,007</b>	0,600
<i>GPR20</i>	Receptor acoplado a proteína G 20	chr8:g.142367729C>T	p.D99N	rs200892677 MAF=0,0004	3,324e-05	515	<b>0,000</b>	<b>0,998</b>
<i>TMEM25</i>	Proteína transmembrana 25	chr11:g.118404174_118404176del	p.V239del			510	N/A	N/A
<i>NGEF</i>	Factor intercanviador de nucleótidos de guanina	chr2:g.233748153G>A	p.S542L		1,279e-05	500	0,350	<b>0,910</b>
<i>NKAP</i>	Proteína activadora de NFκB	chrX:g.119066123C>T	p.S265N	rs182030723 MAF=0,0006	6,847e-05	497	0,120	0,184
<i>NVL</i>	Proteína nuclear que contiene valosina	chr1:g.224491450G>A	p.T312I		8,268e-06	474	<b>0,000</b>	<b>0,995</b>
<i>FN1</i> *	Fibronectina	chr2:g.216251538G>A	p.R1496W	rs139078629 MAF=0,003	0,004904	466	<b>0,005</b>	<b>0,998</b>
<i>ATP6AP1</i>	Subunidad S1 de ATPasa de protones vacuolar	chrX:g.153664043G>A	p.V407I		4,561e-05	464	0,260	<b>0,990</b>
<i>LURAP1L</i>	Proteína adaptadora rica en leucinas 1	chr9:g.12821722G>A	p.R217H		4,948e-05	452	0,270	0,371
<i>HEPFL1</i>	Proteína similar a la hefestina	chr11:g.93839224G>A	p.W991*			451	<b>0,000</b>	N/A

Tabla 2a. (cont.)

Gen	Proteína	Variante <sup>a</sup>	Efecto en la proteína	dbSNP <sup>b</sup>	ExAC <sup>c</sup>	Conser-vación <sup>d</sup>	SIFT <sup>e</sup>	PolyPhen <sup>f</sup>
<i>NTPCR</i>	Trifosfatasa de nucleósidos relacionada con cáncer	chr1:g.233091444G>A	p.R59Q		5,779e-05	439	<b>0,034</b>	0,502
<i>XAB2</i> *	Proteína de unión a XPA 2	chr19:g.7688159G>C	p.T379R		1,652e-05	420	0,059	0,200
<i>CHERP</i>	Proteína del retículo endoplasmático de homeostasis del calcio	chr19:g.16631044C>T	p.R793H	rs202164310 MAF=0,0000	0,0001009	366	0,120	0,716
<i>MEX3D</i>	Proteína de unión a RNA MEX3D	chr19:g.1555839G>C	p.T560R	rs538022731 MAF=0,0002		366	<b>0,030</b>	N/A
<i>BRAT1</i> *	Activador de ATM asociado a BRACA1	chr7:g.2594007C>T	p.R20K	rs143390199 MAF=2e-05	1,651e-05	333	0,192	0,010
<i>BRAT1</i> *	Activador de ATM asociado a BRACA1	chr7:g.2580668G>A	p.T447M	rs368808380 MAF=0,0002	5,845e-05	333	0,110	0,275
<i>CUL9</i>	Culina 9	chr6:g.43154714C>T	p.T423I			251	<b>0,000</b>	<b>0,993</b>
<i>ALPK1</i>	α-quinasa 1	chr4:g.113353195A>C	p.D831A		0,0001255	0	0,060	0,243
<i>CD37</i>	Antígeno de leucocito CD37	chr19:g.49840212C>G	p.I63M		2,476e-05	0	<b>0,040</b>	0,028
<i>IQCF6</i>	Proteína F7 que contiene motivos IQ	chr3:g.51812782G>A	p.R61W			0	<b>0,010</b>	N/A
<i>LFNG</i>	Péptido O-fucosil 3-β-N-acetilglucosaminil-transferasa	chr7:g.2566829C>T	p.R375C		1,69e-05	0	<b>0,020</b>	0,772
<i>MGA</i>	Proteína asociada al gen MAX	chr15:g.41988923C>T	p.S571L			0	0,130	N/A
<i>POLI</i>	Polimerasa de ADN iota	chr18:g.51820404T>C	p.V597A	rs543509008 MAF=0,0002	0,00024	0	0,590	N/A
<i>SHC4</i>	Proteína 4 transformadora de SHC	chr15:g.49254675G>T	p.H180N			0	1,000	0,000
<i>SMS</i>	Esperrina sintasa	chrX:g.21958982G>C	p.G14R			0	0,350	0,002
<i>SNAPC4</i>	Polipéptido 4 del complejo activador de snRNAs	chr9:g.139272279C>G	p.G1334R		2,675e-05	0	0,160	0,707

Tabla 2b. Otras variantes encontradas en las pacientes no relacionadas

Gen	Proteína	Variante <sup>a</sup>	Efecto en la proteína	dbSNP <sup>b</sup>	ExAC <sup>c</sup>	Conser-vación <sup>d</sup>	SIFT <sup>e</sup>	Poly Phen <sup>f</sup>	Paciente FAF
<i>BRAT1</i>	Activador de ATM asociado a BRACA1	chr7:g.2580636C>T	p.E458L			333	0,568	0,000	AFU3
<i>CYP11A1</i>	Citocromo P450 11A1	chr15:g.75014793T>A	p.S216C	rs146622566 MAF=0,0003	0,0001153	0	<b>0,004</b>	<b>0,987</b>	AFU1
<i>MMP9</i>	Metalopeptidasa de matriz 9	chr20:g.44641147T>C	p.M419T		8,242e-06	496	<b>0,000</b>	<b>1,000</b>	AFU3
<i>MVD</i>	Mevalonato difosfato decarboxilasa	chr16:g.88723957C>T	p.R97Q	rs376949804 MAF=3e-05	3,448e-05	0	0,448	0,009	AFU2
<i>RUNX2</i>	Factor de transcripción 2 relacionado con Runt	chr6:g.45480010C>T	p.P296L	rs201584115 MAF=0,0004	0,0002066	642	<b>0,040</b>	<b>0,999</b>	AFU3

(<sup>a</sup>): posición genómica de la variante en el genoma de referencia humano GRCh37; (<sup>b</sup>): número de identificador de referencia del SNP (rs) y MAF (frecuencia del alelo minoritario) de las variantes descritas; (<sup>c</sup>): frecuencia alélica de las variantes descritas en la base de datos ExAC; (<sup>d</sup>): puntuación de conservación del PhastCons (0 a 1.000), siendo 1.000 el *locus* más conservado y 0 un *locus* no conservado; (<sup>e</sup>): SIFT: 0-0,05 perjudicial (en negrita); 0,051-1 tolerable; (°): PolyPhen: 0-0,4 benigno; 0,41-0,89 posiblemente perjudicial; 0,9-1 patológico (en negrita); (°): presente en un alelo doble mutante.

### Análisis de las variantes en individuos controles y en la población general

Ninguna variante de las tablas 2a y 2b fue encontrada en 3 controles (pacientes tratadas con BPs durante un período largo de tiempo pero sin FAFs). Todas las variantes detectadas en las pacientes con FAF se buscaron en la base de datos ExAC para determinar si se trataba de variantes nuevas o muy raras (MAF <0,005). En ese sentido, once mutaciones no se encontraron ni en dbSNP ni en ExAC (*GGPS1*: p.D188Y; *COG4*: p.G85D; *PGRMC1*: p.P177H; *TMEM25*: p.V239del; *HEPH1*: p.W991\*; *CUL9*: p.T423I; *IQCF6*: p.R61W; *MGA*: p.S571L; *SHC4*: p.H180N; *SMS*: p.G14R; *BRAT1*: p.E458L). Las otras variantes tienen frecuencias  $\leq 1/10000$ , según ExAC.

### Red de interacción génica/proteica y enriquecimiento de vías

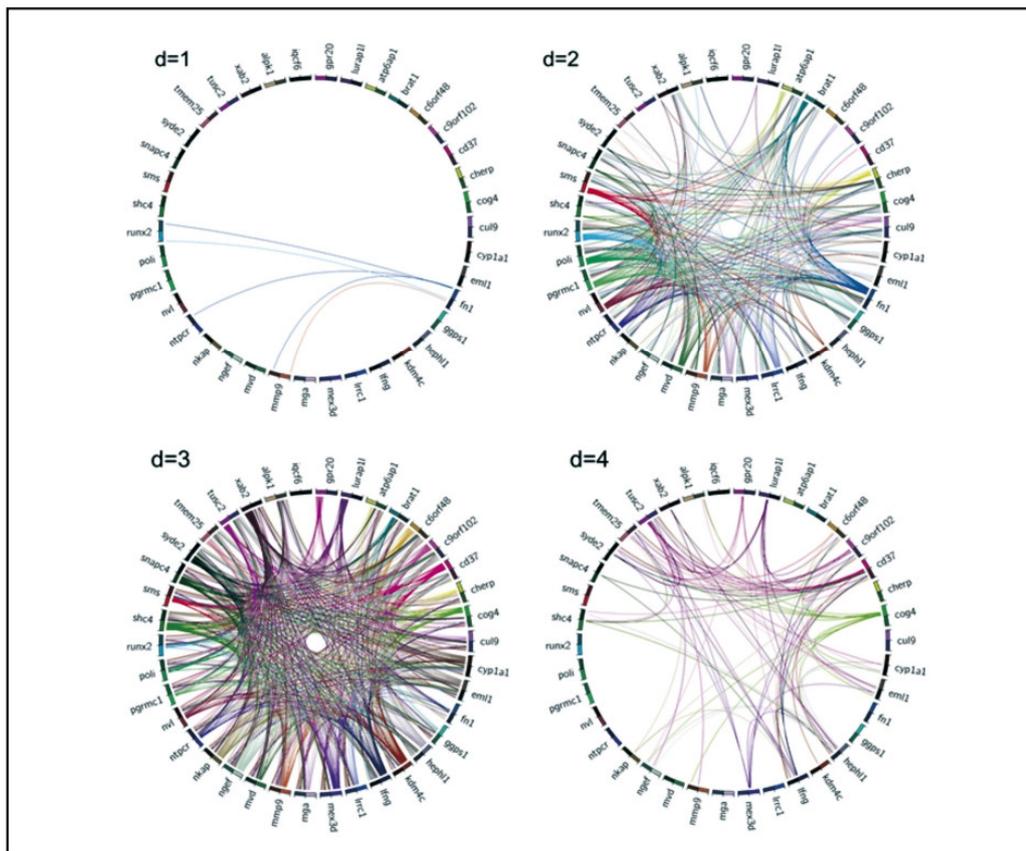
Se construyó una red de interacciones entre genes y/o proteínas para investigar las vías funcionales relacionadas con los 37 genes mutados encontrados en la secuenciación de los exomas y detectar otros genes potencialmente causales, así como mecanismos moleculares que puedan estar implicados en la generación de las FAFs. La figura 1

muestra la conectividad entre parejas de genes. En distintos círculos, se muestran las conexiones de entrada y de salida para los 37 genes a distancias 1 a 4, respectivamente. A distancia 1 casi no hay interacciones, siendo *FN1* el único gen conectado con otros. A distancia 2 se observa más conectividad. La mayoría de la conectividad entre parejas de genes se observa a distancia 3. El único gen que no presenta ninguna interacción a ningún nivel es *IQCF6*.

La red de interacciones de genes/proteínas muestra que *GGPS1* y *CYP11A1*, dos de los genes *driver* más relevantes, se conectan a distancia 3, a través de *INS* y *IL6* (Figura 2a). Otros 4 genes *driver* (*RUNX2*, *MVD*, *MMP9* y *PGRMC1*) están conectados con *GGPS1* a distancia 2. *MMP9* también está a distancia 2 de *CYP11A1*. Además, *FN1* y *MMP9* están conectados a distancia 1. De manera similar, los genes *driver* *SYDE2* y *NGEF* están interconectados a distancia 2, a través de *RHOB* (Figura 2b).

El análisis de enriquecimiento de vías en los 37 genes mutados, realizado con la herramienta DAVID, dio como resultado la vía de biosíntesis de los isoprenoides (GO:0008299) ( $p=0,0006$ ), que contiene los genes *GGPS1*, *MVD* y *CYP11A1*.

Figura 1. Esquema de la conectividad entre parejas de genes a distancias 1 a 4. En los círculos se muestran los símbolos de los 37 genes FAF encontrados en este estudio y sus conexiones de entrada y de salida



## Discusión

En este trabajo hemos estudiado el trasfondo genético de 3 hermanas con FAF y 3 pacientes adicionales, no relacionadas, a través de la secuenciación masiva del exoma para identificar posibles genes de susceptibilidad a la patología. Hemos identificado 37 variantes raras (en 34 genes) compartidas por las 3 hermanas, algunas de ellas no descritas anteriormente y consideradas dañinas por los predictores. La variante más llamativa fue la mutación p.Asp188Tyr en el gen *GGPS1*, que presentó la mejor puntuación de conservación, y que ya hemos descrito en un trabajo previo<sup>9</sup>. Otro hallazgo interesante fueron las dos mutaciones en el gen *CYP11A1*, una encontrada en las tres hermanas y la otra en una paciente no relacionada. Sin embargo, hay otras variantes que también podrían estar involucradas, en distintos grados, en la susceptibilidad a las FAFs asociadas a BPs o en el fenotipo osteoporótico subyacente, de modo que nuestros datos serían compatibles con un modelo en el cual la acumulación de variantes de susceptibilidad podría contribuir a la base genética de las FAFs.

Los estudios epidemiológicos sugieren que existe una relación entre las FAFs y un tratamiento prolongado con BPs. Shane *et al.*, describieron períodos de tratamiento de una mediana de 7 años<sup>4</sup>. El riesgo absoluto de sufrir una FAF asociada al tratamiento con BPs se encuentra entre 2 casos por 100.000 pacientes/año a los 2 años de tratamiento y 78 casos por 100.000 pacientes/año a los 8 años de tratamiento<sup>25</sup>. Estos datos sugieren que la duración de la terapia con BPs influiría positivamente en el riesgo de sufrir estas fracturas. En nuestro estudio, los casos de 6 pacientes con FAF después de un tratamiento a largo plazo con BPs son consistentes con esta asociación. Además, la ocurrencia de las FAFs en las 3 hermanas sugiere una predisposición genética con un papel determinante en la patología. Este estudio ha sido el primer análisis de exoma de pacientes de FAF. Hemos priorizado mutaciones raras, no-sinónimas, compartidas por las 3 hermanas. No se encontró ninguna mutación en homocigosis o heterocigosis compuesta en ningún gen. Estos hallazgos van en contra de un patrón de herencia recesivo para estos casos y son consistentes con el hecho que la FAF no es una enfermedad genética severa que ocurra durante las primeras etapas de la vida. No obstante, en el modelo dominante, se encontraron 34 genes mutados, algunos muy importantes para el metabolismo óseo. En un trabajo anterior que tenía por objetivo descubrir las causas genéticas de las FAFs, se utilizó un chip de exoma con >300.000 variantes codificantes ya conocidas y se encontraron 21 variantes raras sobrerrepresentadas en 13 pacientes de FAF<sup>26</sup>. Sin embargo, ninguno de estos alelos de riesgo se encontró en los pacientes analizados en nuestro estudio. En concreto, no se encontraron variantes en el gen *PPEF2*, el único con un cambio asociado significativamente con el fenotipo en el estudio de Pérez-Núñez *et al.*<sup>26</sup> Esto apunta a una base genética heterogénea para las FAFs. En todo caso, es importante señalar que nuestra aproximación metodológica difiere de la del estudio mencionado en tanto que analizamos toda

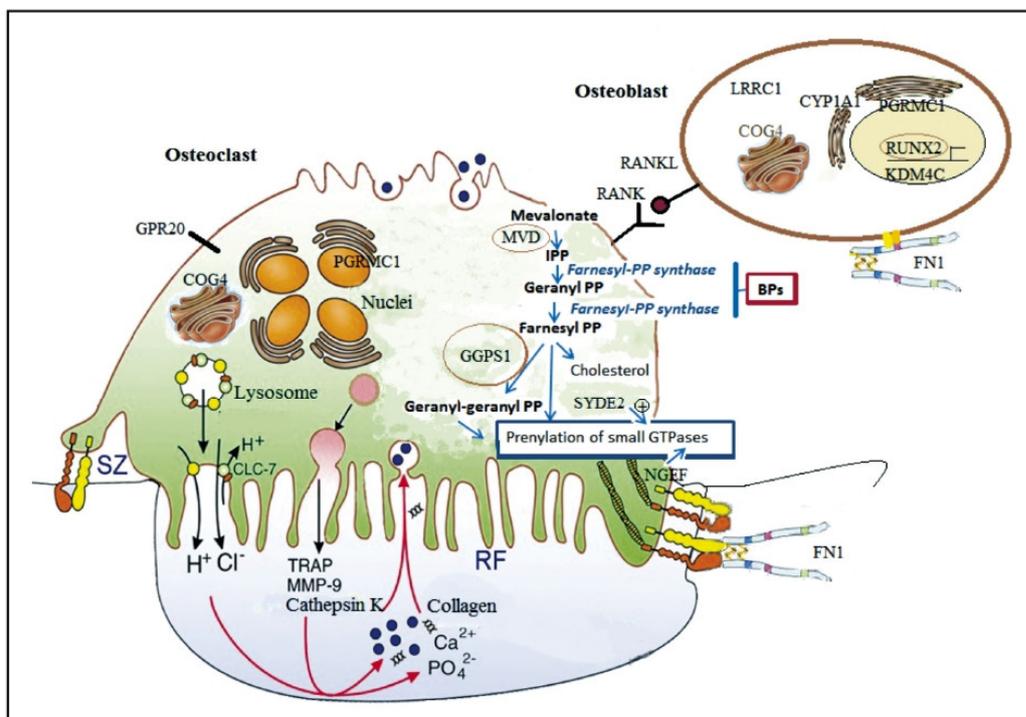
la secuencia del exoma, cosa que nos permitió encontrar variantes no descritas anteriormente.

En el presente estudio, el único gen con mutaciones en las 3 hermanas y en pacientes no relacionados fue *CYP11A1*. Recientemente, Peris *et al.*<sup>27</sup> secuenciaron este gen en 17 pacientes de FAF y encontraron otra mutación en una de ellas. El gen *CYP11A1* codifica la enzima citocromo P450 1A1 que está involucrada en el metabolismo de fármacos y xenobióticos. Se trata de una hidroxilasa de hidrocarburos arilos y sus sustratos exógenos potenciales incluyen hidrocarburos aromáticos policíclicos, y está implicada en la formación de distintos tipos de cáncer humanos. Sus sustratos endógenos incluyen eicosanoides, que pueden generar productos biológicamente activos que actúan en el sistema vascular, entre otros. Este gen también es responsable de la hidroxilación del 17β-estradiol, la estrona y la vitamina D en tejidos extrahepáticos<sup>28</sup>. Esto es coherente con su papel en la biología ósea, una idea apoyada por Napoli *et al.*<sup>29</sup>, quienes demostraron que el polimorfismo C4887A estaba relacionado con un aumento significativo del catabolismo de los estrógenos y con una densidad mineral ósea (DMO) femoral baja en mujeres postmenopáusicas. Por lo tanto, *CYP11A1* se presenta como otro gen de susceptibilidad potencial a las FAFs, aunque el mecanismo exacto de su acción en el metabolismo óseo todavía es desconocido y más estudios son necesarios para elucidarlo.

Entre los otros genes con variantes en las tres hermanas, *FNI* codifica la fibronectina, una proteína de la matriz extracelular necesaria para la regulación de la deposición del colágeno de tipo I por parte de los osteoblastos, esencial para la mineralización de la matriz extracelular, y cuyos niveles se han visto afectados por el tratamiento con BPs<sup>30</sup>. Encontramos que las tres hermanas eran portadoras de un alelo doble mutante (p.V224I y p.R1496W) en *FNI*, donde las dos mutaciones fueron consideradas como dañinas por los predictores de patogenicidad. Esta fibronectina alterada podría afectar la mineralización ósea y/o la respuesta a los BPs y estar relacionada con el riesgo a sufrir una FAF en estas mujeres. También encontramos mutados 2 reguladores de GTPasas pequeñas: *SYDE2* y *NGEF*. Sus funciones respectivas (activación de las GTPasas RHO y de intercambio de sus nucleótidos de guanina) constituyen pistas sobre posibles efectos en la función osteodástica y en la respuesta a los BPs. Las RHO GTPasas están en la vía del mevalonato en una posición por debajo del sitio de acción de los BPs, ya que tienen que ser preniladas (famesiladas o geranilgeraniladas) para su correcta función celular. Por otra parte, nuestra red de interacción de genes/proteínas muestra como *NGEF* está muy relacionado con las efrinas y los receptores de efrinas (Figura 2b), que tienen un papel clave en el mecanismo de acoplamiento entre osteoclastos y osteoblastos<sup>31</sup>. Otro grupo de genes mutados en las 3 hermanas codifican proteínas nucleares con efectos pleiotrópicos sobre la expresión génica y/o la reparación del DNA (*KDM4C*, *XAB2*, *NVL*, *NKAP*, *ERCC6L2*). De ellos destacamos el gen *KDM4C*, que codifica una demetilasa lisina-específica que contiene un dominio JmjC, que ha sido previamente asociado con la edad de menarquia<sup>32</sup>, un biomarcador para la densidad ósea.



Figura 3. Proteínas codificadas por los genes mutados en las pacientes de FAF de este estudio y relacionadas con la función ósea



**Conflicto de intereses:** Los autores declaran no tener conflicto de interés.

### Bibliografía

- Freemantle N, Cooper C, Diez-Perez A, Gitlin M, Radcliffe H, Shepherd S, et al. Results of indirect and mixed treatment comparison of fracture efficacy for osteoporosis treatments: A meta-analysis. *Osteoporos Int.* 2013;24(1):209-17.
- Peng J, Liu Y, Chen L, Peng K, Xu Z, Zhang D, et al. Bisphosphonates can prevent recurrent hip fracture and reduce the mortality in osteoporotic patient with hip fracture: A meta-analysis. *Pakistan J Med Sci.* 2016;32(2):499-504.
- Kennel KA, Drake MT. Adverse effects of bisphosphonates: Implications for osteoporosis management. *Mayo Clinic Proceedings.* 2009;84(7):632-7.
- Shane E, Burr D, Abrahamsen B, Adler RA, Brown TD, Cheung AM, et al. Atypical subtrochanteric and diaphyseal femoral fractures: second report of a task force of the American Society for Bone and Mineral Research. *J Bone Miner Res.* 2014;29(1):1-23.
- Taormina DP, Marcano AI, Karia R, Egol KA, Tejwani NC. Symptomatic atypical femoral fractures are related to underlying hip geometry. *Bone.* 2014;63:1-6.
- Sutton RAL, Mumm S, Coburn SP, Ericson KL, Whyte MP. "Atypical femoral fractures" during bisphosphonate exposure in adult hypophosphatasia. *J Bone Miner Res.* 2012;27(5):987-94.
- Meier RPH, Lorenzini KI, Uebelhart B, Stern R, Peter RE, Rizzoli R. Atypical femoral fracture following bisphosphonate treatment in a woman with osteogenesis imperfecta case report. *Acta Orthop.* 2012;83(5):548-50.
- Alonso N, Soares DC, V McCloskey E, Summers GD, Ralston SH, Gregson CL. Atypical femoral fracture in osteoporosis pseudoglioma syndrome associated with two novel compound heterozygous mutations in LRP5. *J Bone Miner Res.* 2015;30(4):615-20.
- Roca-Ayats N, Balcells S, Garcia-Giralt N, Falcó-Mascaró M, Martínez-Gil N, Abril JF, et al. GGPS1 mutation and atypical femoral fractures with bisphosphonates. *N Engl J Med.* 2017;376(18):1794-5.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754-60.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-9.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297-303.
- Consortium T 1000 GP. An integrated map of genetic variation from 1,092 human genomes. *Nature.* 2012;491:56.
- Medina I, De Maria A, Bleda M, Salavert F, Alonso R, Gonzalez CY, et al. VARIANT: Command Line, Web service and Web interface for fast and accurate functional characterization of variants found by Next-Generation Sequencing. *Nucleic Acids Res.* 2012;40:W54-8.
- Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* 2009;4:1073-81.
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nature Methods.* 2010;7(4):248-9.
- Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* 2005;15(8):1034-50.

18. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4:44-57.
19. Boloc D, Castillo-Lara S, Marfany G, González-Duarte R, Abril JF. Distilling a Visual Network of Retinitis Pigmentosa Gene-Protein Interactions to Uncover New Disease Candidates. *PLoS One.* 2015;10(8):e0135307.
20. Chatr-Aryamontri A, Oughtred R, Boucher L, Rust J, Chang C, Kolas NK, et al. The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* 2017;45(D1):D369-79.
21. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, et al. STRING v9.1: Protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 2013;41(D1):D808-15.
22. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: An information aesthetic for comparative genomics. *Genome Res.* 2009;19(9):1639-45.
23. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: Archive for functional genomics data sets - Update. *Nucleic Acids Res.* 2013;41(D1):D991-5.
24. Yuen T, Stachnik A, Iqbal J, Sgobba M, Gupta Y, Lu P, et al. Bisphosphonates inactivate human EGFRs to exert antitumor actions. *Proc Natl Acad Sci U S A.* 2014;111(50):17989-94.
25. Brown JP, Morin S, Leslie W, Papaioannou A, Cheung AM, Davison KS, et al. Bisphosphonates for treatment of osteoporosis: expected benefits, potential harms, and drug holidays. *Can Fam Physician.* 2014;60(4):324-33.
26. Pérez-Núñez I, Pérez-Castrillón JL, Zarrabeitia MT, García-Ibarbia C, Martínez-Calvo L, Olmos JM, et al. Exon array analysis reveals genetic heterogeneity in atypical femoral fractures. A pilot study. *Mol Cell Biochem.* 2015;409:45-50.
27. Peris P, González-Roca E, Rodríguez-García SC, del Mar López-Cobo M, Monegal A, Guañabens N. Incidence of mutations in the ALPL, GGPS1 and CYP1A1 genes in patients with atypical femoral fractures. *JBM Plus.* 2018; [Epub ahead of print].
28. Zhou SF, Liu J-P, Chowbay B. Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metab Rev.* 2009;41(2):89-295.
29. Napoli N, Villareal DT, Mumm S, Halstead L, Sheikh S, Cagaanan M, et al. Effect of CYP1A1 gene polymorphisms on estrogen metabolism and bone density. *J Bone Miner Res.* 2005;20(2):232-9.
30. Insalaco L, Gaudio F Di, Terrasi M, Amodeo V, Caruso S, Corsini LR, et al. Analysis of molecular mechanisms and anti-tumoural effects of zoledronic acid in breast cancer cells. *J Cell Mol Med.* 2012;16(9):2186-95.
31. Sims NA, Martin TJ. Coupling the activities of bone formation and resorption: a multitude of signals within the basic multicellular unit. *Bonekey Rep.* 2014;3:481.
32. Perry JRB, Day F, Elks CE, Sulem P, Thompson DJ, Ferreira T, et al. Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature.* 2014;514(7520):92-7.
33. Mansouri MR, Schuster J, Badhai J, Stattin EL, Lösel R, Wehling M, et al. Alterations in the expression, structure and function of progesterone receptor membrane component-1 (PGRMC1) in premature ovarian failure. *Hum Mol Genet.* 2008;17(23):3776-83.
34. Xia WF, Tang FL, Xiong L, Xiong S, Jung JU, Lee DH, et al. Vps35 loss promotes hyperresorptive osteoclastogenesis and osteoporosis via sustained RANKL signaling. *J Cell Biol.* 2013;200(6):821-37.
35. Nguyen AM, Jacobs CR. Emerging role of primary cilia as mechanosensors in osteocytes. *Bone.* 2013;54(2):196-204.
36. Ducy P, Zhang R, Geoffroy V, Ridall AL, Karsenty G. Osf2/Cbfa1: A transcriptional activator of osteoblast differentiation. *Cell.* 1997;89(5):747-54.
37. Okada Y, Naka K, Kawamura K, Matsumoto T, Nakanishi I, Fujimoto N, et al. Localization of matrix metalloproteinase 9 (92-kilodalton gelatinase/type IV collagenase = gelatinase B) in osteoclasts: Implications for bone resorption. *Lab Investig.* 1995;72(3):311-22.
38. Nyman JS, Lynch CC, Perrien DS, Thiollay S, O'Quinn EC, Patil CA, et al. Differential effects between the loss of MMP-2 and MMP-9 on structural and tissue-level properties of bone. *J Bone Miner Res.* 2011;26(6):1252-60.
39. Pratap J, Javed A, Languino LR, van Wijnen AJ, Stein JL, Stein GS, et al. The Runx2 osteogenic transcription factor regulates matrix metalloproteinase 9 in bone metastatic cancer cells and controls cell invasion. *Mol Cell Biol.* 2005;25(19):8581-91.





## INVITED ARTICLE



WILEY

## Bone development and remodeling in metabolic disorders

Jenny Serra-Vinardell<sup>1,2</sup> | Neus Roca-Ayats<sup>1</sup> | Laura De-Ugarte<sup>3,4</sup> | Lluïsa Vilageliu<sup>1</sup> | Susanna Balcells<sup>1</sup> | Daniel Grinberg<sup>1</sup>

<sup>1</sup>Department of Genetics, Microbiology and Statistics, Faculty of Biology, Universitat de Barcelona, CIBERER, IBUB, IRSJD, Barcelona, Spain

<sup>2</sup>Section of Human Biochemical Genetics, Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland

<sup>3</sup>Department of Anatomy and Cell Biology, Indiana University School of Medicine, Indianapolis, Indiana

<sup>4</sup>Department of Anatomy and Cell Biology, Indiana Center for Musculoskeletal Health, Indianapolis, Indiana

### Correspondence

Daniel Grinberg, Department of Genetics, Microbiology and Statistics, Faculty of Biology, Universitat de Barcelona, CIBERER, IBUB, IRSJD, Barcelona, Spain. Email: dgrinberg@ub.edu

**Communicating Editor:** Robin Lachmann

### Funding information

CIBERER, Grant/Award Number: U720; Generalitat de Catalunya, Grant/Award Number: 2014SGR932; Spanish MINECO, Grant/Award Number: SAF2016-75948-R

### Abstract

There are many metabolic disorders that present with bone phenotypes. In some cases, the pathological bone symptoms are the main features of the disease whereas in others they are a secondary characteristic. In general, the generation of the bone problems in these disorders is not well understood and the therapeutic options for them are scarce. Bone development occurs in the early stages of embryonic development where the bone formation, or osteogenesis, takes place. This osteogenesis can be produced through the direct transformation of the pre-existing mesenchymal cells into bone tissue (intramembranous ossification) or by the replacement of the cartilage by bone (endochondral ossification). In contrast, bone remodeling takes place during the bone's growth, after the bone development, and continues throughout the whole life. The remodeling involves the removal of mineralized bone by osteoclasts followed by the formation of bone matrix by the osteoblasts, which subsequently becomes mineralized. In some metabolic diseases, bone pathological features are associated with bone development problems but in others they are associated with bone remodeling. Here, we describe three examples of impaired bone development or remodeling in metabolic diseases, including work by others and the results from our research. In particular, we will focus on hereditary multiple exostosis (or osteochondromatosis), Gaucher disease, and the susceptibility to atypical femoral fracture in patients treated with bisphosphonates for several years.

### KEYWORDS

atypical femoral fracture, bone development, bone remodeling, CYP1A1, EXT2, Gaucher disease, GGPPS, multiple hereditary exostosis

## 1 | OSTEOCHONDROMATOSIS OR MULTIPLE HEREDITARY EXOSTOSIS

Osteochondromatosis is characterized by the growth of multiple benign tumors mainly in long bones. The disease belongs to the group of congenital disorders of glycosylation. In particular, it is a defect of O-glycosylation. It is inherited as an autosomal dominant disease and it is caused by monoallelic mutations either in the *EXT1* or in the *EXT2* genes (reviewed in Wuyts and van Hul<sup>1</sup>). As a consequence of this, the alternative name for the disorder is EXT1/

EXT2-CDG. The EXT1 (exostosin 1) and EXT2 (extosin 2) proteins form a copolymerase involved in heparan sulfate biosynthesis.

The disease is related to bone development. Endochondral ossification is one of the two essential processes during fetal development of the mammalian skeletal system, by which bone tissue is created through the replacement of growing cartilage by bone (the other process is intramembranous ossification). In long bones, there is a region called the growth plate, in which chondrocytes proliferate and are replaced by osteoblasts. In this process, several signaling molecules play key roles in the regulation and

direction of bone growth.<sup>2</sup> Extracellular heparan sulfate makes a barrier that regulates the flux of these molecules. Individuals with an inherited mutation in one allele of *EXT1* or *EXT2* can suffer a second (somatic) mutation in the wild-type (WT) allele. This will lead to a local lack of heparan sulfate and an impaired regulation of bone growth, giving rise to osteochondromas.<sup>3</sup> The role of heparan sulfate in the disease has been recently reviewed by Maurizio Pacifici.<sup>4</sup> The decrease in exostosin and heparan sulfate levels caused by the second hit (the somatic mutation in the WT allele of the *EXT* gene with a germline mutation in the other allele) causes a decrease in signaling molecules such as fibroblast growth factor (FGF), MAPK/ERK Kinase (MEK), extracellular-signal-regulated kinase (ERK), Noggin, and Gremlin, and an increase of bone morphogenetic protein (BMP) and hedgehog (Hh) signaling and of heparanase, which are involved in the abnormal growing of bone. For this reason, this disease can be considered as an example of abnormal bone development. The most severe complication of osteochondromatosis is malignant transformation of an osteochondroma into a peripheral secondary chondrosarcoma. The estimated lifelong risk varies among different studies from 1% to 25%.<sup>5</sup> Czajka and DiCaprio<sup>6</sup> studied a large international, heterogeneous cohort of around 800 patients with multiple hereditary exostoses and reported a proportion of 2.7% of malignant degeneration to chondrosarcoma. The mechanisms for this malignant change are not clear. Different authors reported genetic mutations in genes different from *EXT1* and *EXT2* during chondrosarcoma progression and it is assumed that these are necessary to progress into malignancy (Reference<sup>7</sup> and references therein). Musso et al<sup>7</sup> described a surprising case in which they observed a loss of the *EXT2* mutant allele in the peripheral secondary chondrosarcoma, instead of the expected loss of the *EXT2* WT allele, suggesting a different cell of origin for osteochondromas and chondrosarcomas. However, this is a topic that is still open.

Our group has studied the mutations present in Spanish<sup>8</sup> and Latino American<sup>9-12</sup> patients. In general, we sequenced polymerase chain reaction-amplified exons and flanking regions of the *EXT1* and *EXT2* genes, followed by multiplex ligation-dependent probe amplification (MLPA) analysis if the sequencing of the exons gave negative results. In some cases, we observed mosaicism in the first affected individual of the family.<sup>8</sup> Occasionally, one of the parents of the proband, with a very mild phenotype, learned about his/her disease after our analysis of the family. In one particular case, a patient came to us after having obtained negative results both of sequencing and MLPA analyses by a private molecular diagnosis company. We thought that she could either bear a hidden mutation in *EXT1* or *EXT2*, or bear a mutation in a hypothetical *EXT3* gene, suggested to exist but never

found. We resequenced the *EXT1* and *EXT2* genes and with some of the single nucleotide polymorphisms (SNPs) found in the resequencing, we performed a segregation analysis in the patient's family to see if we were able to discard the involvement of these two genes in their disease. The result was the opposite: the segregation analysis showed an apparent lack of heredity of *EXT2*, only consistent with a deletion (Figure 1A,B), which we demonstrated afterward by MLPA (Figure 1C). It was a serious mistake of the private company.

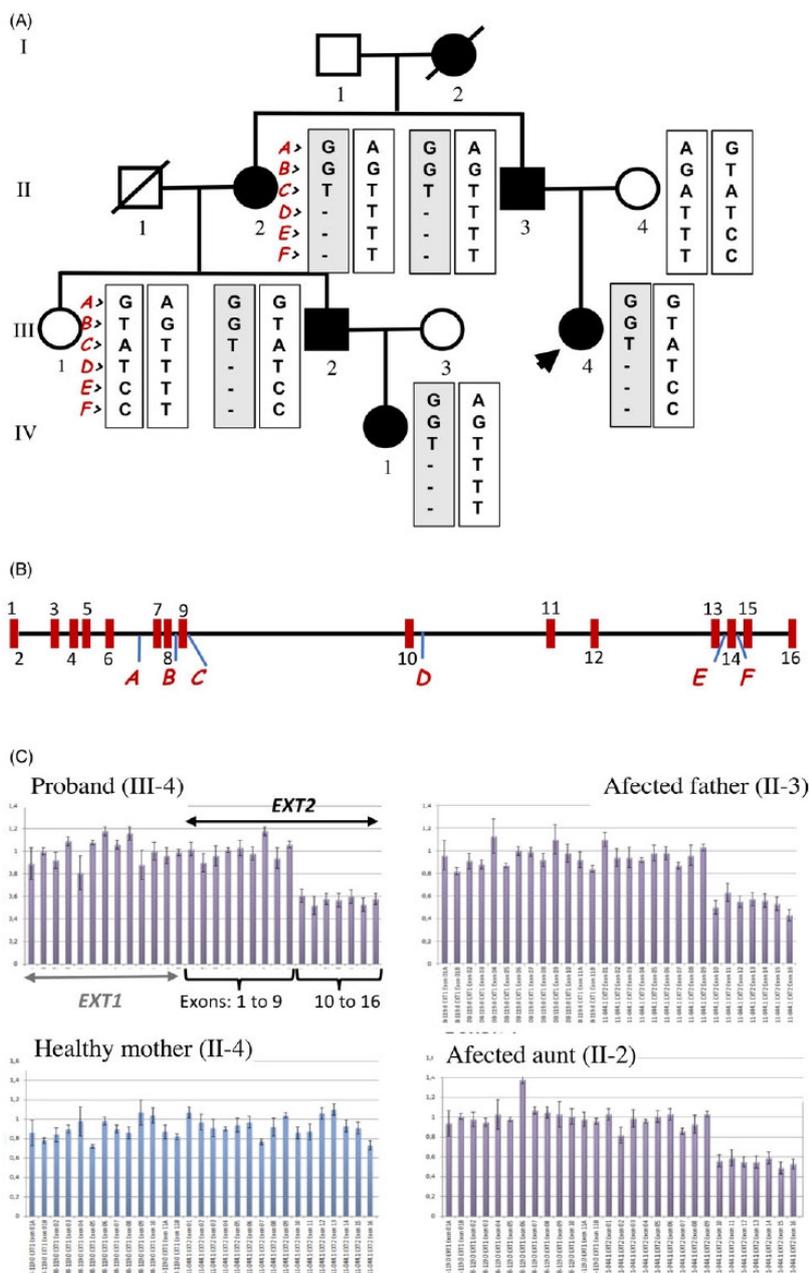
As mentioned above, osteochondromatosis is inherited as an autosomal dominant disease. Thus, only a mutation in one allele of either *EXT1* or *EXT2* is present in the germline. Until a few years ago, no case with germline mutations in the two alleles of these genes was known. However, two cases with homozygous germline missense mutations in the *EXT2* gene have been recently published.<sup>13,14</sup> Surprisingly, the phenotype is characterized by seizures and developmental disorders without exostoses. The only bone phenotype was osteopenia, present in one of the patients. The proposed name for this novel syndrome is autosomal recessive *EXT2*-related syndrome (AREXT2). Notably, these cases show that the consequences of these homozygous missense mutations are critical for brain development, while not affecting bone. The role of heparan sulfate in the brain is not well understood but very recently, it was shown that it organizes neuronal synapses through neuroligin partnerships.<sup>15</sup> Further research will be necessary to better understand the dual roles of *EXT2* in bone and brain.

We have also studied the family of *EXT* genes from a different point of view. In Sanfilippo disease (mucopolysaccharidosis III), there is an accumulation of heparan sulfate in the lysosomes due to an impaired function of one of several lysosomal enzymes. We have assayed the inhibition of the *EXTL2* and *EXTL3* genes, as a substrate reduction therapy for Sanfilippo C disease in patients' fibroblasts<sup>16</sup> and we are currently performing a similar approach on Sanfilippo C neurons, derived from induced pluripotent stem cells (iPSC) generated by our group.<sup>17</sup>

## 2 | GAUCHER DISEASE

Lysosomal storage diseases (LSDs) are a group of more than 50 disorders, which mainly result from the deficient activity of specific lysosomal enzymes. This deficiency produces a progressive accumulation of specific substrates affecting different biochemical or cellular pathways, which subsequently will cause the tissue pathology.<sup>18</sup> Gaucher disease (GD), the most common LSD, is caused by mutations in the *GBA1* gene (MIM# 606463) that produce a defective activity of glucocerebrosidase (EC 3.2.1.45; *GBA1*), the lysosomal

**FIGURE 1** Analysis of a pedigree with a mutation in the *EXT2* gene. (A) Segregation analysis of *EXT2* polymorphic markers. A-F refer to positions in the map shown in panel B. (B) Position of the polymorphic markers, A-F, in relation to *EXT2* exons. (C) MLPA analysis of the DNA of the proband, her parents and her affected aunt, showing the heterozygous deletion of *EXT2* exons 10 to 16 in the affected individuals



enzyme responsible for the hydrolysis of glucosylceramide (GlcCer) into glucose and ceramide. As a result of this autosomal recessive genetic defect, GlcCer and glucosylsphingosine (GlcSph) accumulate in the lysosomes of macrophages (revised by Sidransky<sup>19</sup>) generating the typical “Gaucher cells,” the hallmark of the disease. Based on the absence or presence and severity of neuronopathic involvement, GD has been classified into three clinical phenotypes, non-

neuronopathic (GD1), acute neuronopathic (GD2), and chronic or subacute neuronopathic (GD3).<sup>20</sup> An extreme phenotypic variability has been reported for GD, within each of the clinical types, and even among patients with the same mutations. This variability is likely due to a multitude of factors such as genetic background, environment, and epigenetic status.<sup>19</sup> Davidson et al<sup>21</sup> reviewed genetic modifiers that influence the phenotypic outcome of GD.

GD1 is the most frequent form of the disease and it is characterized by heterogeneous manifestations including visceral (hepatosplenomegaly), hematological, and skeletal symptoms. Bone involvement affects up to 90% of GD1 patients and it is for them the most debilitating feature because it has a major impact on their life quality.<sup>22</sup> Skeletal manifestations include erlenmeyer flask deformity, fractures due to osteopenia or osteoporosis, osteosclerosis, osteonecrosis, bone pain, bone crisis, growth retardation during childhood and, rarely, acute osteomyelitis.<sup>23</sup> The majority of these manifestations could be explained by the disruption of the balance between osteoblastic bone formation and osteoclastic bone resorption. The markers of bone metabolism are useful to measure changes in the activities of osteoblasts and osteoclasts. However, controversial results on the alteration of bone formation and of bone resorption markers in GD and their response to enzyme therapy have been reported (revised by van Dussen et al<sup>24</sup>). Furthermore, we do not have a conclusive statement from previous studies where mouse or cell models were used to understand bone pathology in GD. Basically, the lack of consistence between studies is an evidence that the bone pathology in GD is complex and, based on the wide phenotypic spectrum in patients, may be a pleiotropic disease.

Campeau et al<sup>25</sup> revealed that mesenchymal stromal cells (MSCs) from a GD1 patient (N370S/L444P genotype) displayed an altered secretome that may contribute to the skeletal and immune problems in GD. Afterward, Mistry et al<sup>26</sup> provided an important model, using a conditional *GBA1* knockout in hematopoietic and mesenchymal cell lineages. They reported that the model recapitulated the main features of GD1, such as visceral and hematologic diseases together with a profound osteopenia. In addition, they provided evidence that whereas the mouse model had a normal osteoclast formation (osteoclastogenesis), the bone formation (osteoblastogenesis) appeared to be defective. They suggested that the osteoblastogenesis was inhibited by the accumulation of the lipids GlcCer and GlcSph, and their consequent interaction with the protein kinase C (PKC). Later on, the same authors proposed that the extralysosomal glucocerebrosidase *GBA2* transformed the increased levels in serum of GlcCer and, mainly, GlcSph in sphingosine, which inhibited osteoblast survival.<sup>27</sup> Despite this previous evidence, other studies confirmed a correlation between bone features and the osteoclast number in GD.<sup>28–32</sup>

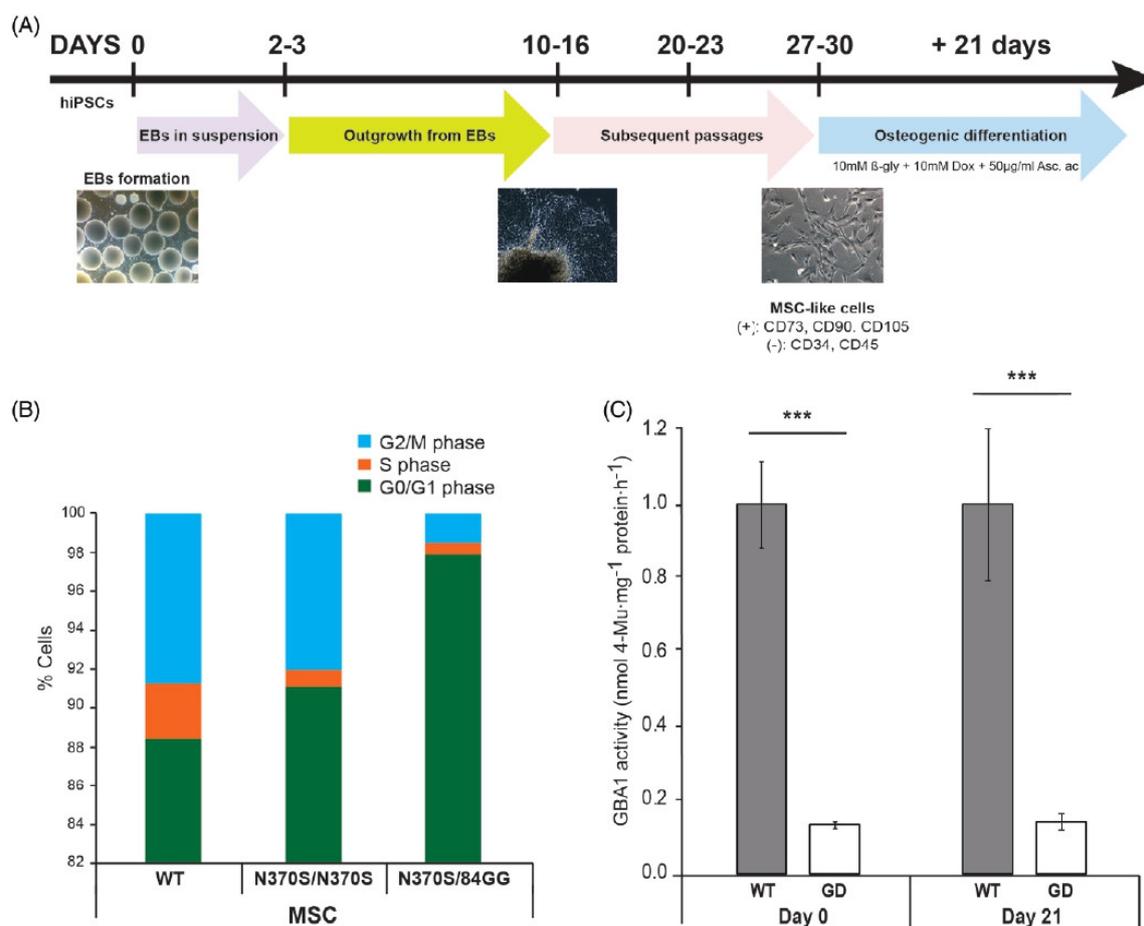
Several groups decided to use human induced Pluripotent Stem Cells (iPSC)-derived osteoblasts to address the unknown of the bone pathology in GD. The main study so far is the one recently published by Ricardo Feldman's group.<sup>33</sup> They showed that GD iPSC-derived osteoblasts had developmental and lysosomal defects that impaired bone matrix deposition. Moreover, they showed that the canonical Wnt pathway was affected. In concordance, many studies

showed the importance of this pathway in bone metabolism, including ours.<sup>34–36</sup>

The clinical bone manifestations in GD are included in the bone remodeling problems category, because bone problems appear as a postnatal trait. Whether the osteoblasts are downregulated or the osteoclasts are upregulated is one of the main questions that needs to be answered in order to look for an effective treatment.

Our aim was to create osteoblasts derived from WT human iPSC generated in our group<sup>17</sup> and compare them with those generated from GD patients (N370S/N370S, a gift from Ricardo Feldman, G202R/L444P, a gift from Gustavo Tiscomia,<sup>37</sup> and N370S/84GG<sup>38</sup>). For the MSC-like cell induction, we used a multistep culture method, summarized in Figure 2A. For each MSC induction (WT, N370S/N370S, G202R/L444P, and N370S/84GG), the expression of specific MSC surface markers (CD73, CD90, and CD105) and the absence of expression of hematopoietic markers (CD34 and CD45) were verified (data not shown). This qualitative step was required to continue with the subsequent experiments. From now on, the MSC-like cells will be mentioned as MSC.

The MSCs with G202R/L444P and N370S/84GG genotypes presented lack of proliferation capacity and, therefore, could not be used in subsequent experiments. These cells were larger and flatter compared to the WT and N370S/N370S MSC (spindle-shaped cells in both cases). The lack of self-renewal of these large flattened cells has been previously described in the literature.<sup>39,40</sup> Because the relationship between cell size, morphology, and senescence is well known,<sup>41,42</sup> we evaluated the cell cycle profile in the MSC stage (in WT, N370S/N370S and N370S/84GG genotypes) by flow cytometry as an attempt to investigate the possible alteration of the cell cycle in GD cells. As shown in Figure 2B, there was a remarkable proportion (97.9%) of GD N370S/84GG cells arrested in the G0/G1 phase compared to WT cells (88.4%). Concomitant with this, there was a reduction in the number of replicative cells (phase S, N370S/84GG MSC = 0.6% compared to WT MSC = 2.9%) and in the G2/M subpopulation (N370S/84GG MSC = 1.5% compared to WT = 8.7%). Regarding the other GD genotype (N370S/N370S), the number of cells in the S phase had also notably decreased in comparison to WT cells (N370S/N370S MSC = 0.9%, WT MSC = 2.9%). However, N370S/N370S MSC in G0/G1 (91.1%) and G2/M (8.0%) phase appeared to be similar to MSC WT (G0/G1 = 88.4% and G2/M = 8.7%). The classic features characterizing the senescence phenotype of MSCs include growth arrest in the G1 phase of the cell cycle, enlarged or flattened morphology and increased expression of senesce-associated  $\beta$ -galactosidase (not evaluated because the  $\beta$ -galactosidase activity in LSD is deregulated and could



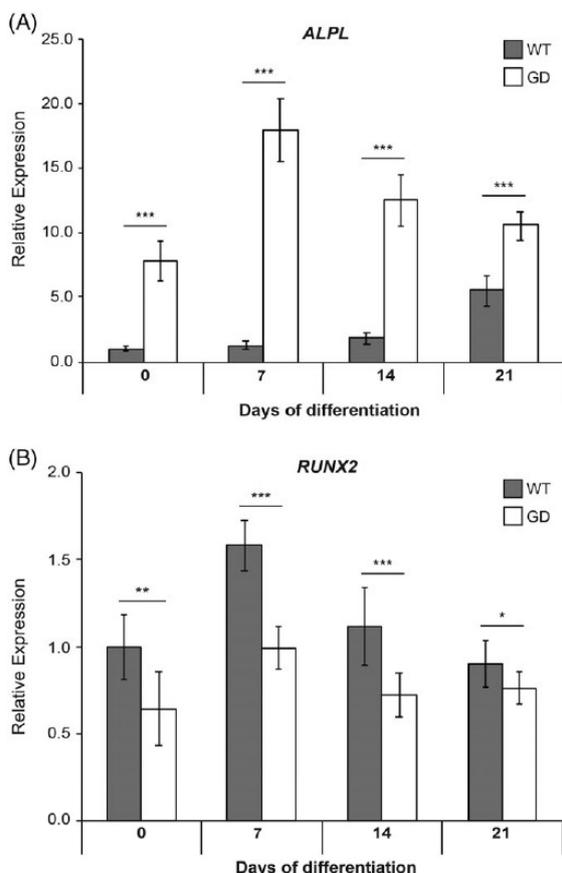
**FIGURE 2** (A) Schematic diagram of the induction of mesenchymal stromal cells (MSC) including four main steps: embryonic body (EB) formation, cell outgrowth from EBs on Matrigel-coated dishes, cell dissociation ( $\times 3$ ) and monolayer culture. (B) Percentage of MSC in each phase of the cell cycle for the indicated genotypes. (C) GBA1 activity in cells of both N370S/N370S and wild-type (WT) genotypes, at the indicated time points of the osteoblastic differentiation process. Results are shown as mean of GBA1 activity  $\pm$  SD. GBA1 activity is expressed in nanomoles 4-Methylumbelliferone (MU)  $\times$  per milligram protein  $\times$  per hour. Statistical differences (\*\*\*)  $P$ -value  $< .001$ , using Student's  $t$  test. See Supporting Information, Data S1 for further details

drive to misinterpretation). Thus, only GD cells with the N370S/N370S genotype were used in further experiments.

To check the GBA1 activity before and after the differentiation process, we analyzed the enzyme activity in the MSC stage (considered as day 0) and after 21 days of the osteogenic differentiation. We showed that the activity of GBA1, both on day 0 and day 21, in GD cells was less than 15% (GD day 0 = 13%, GD day 21 = 14%) compared to that of the WT cells. Moreover, the difference between GBA1 activity in WT and GD cells did not change during the differentiation process (Figure 2C).

We aimed to evaluate in vitro the potential impact of GBA1 deficiency on two representative genes of the osteogenic development and mineralization process, runt-related

transcription factor 2 (*RUNX2*) and tissue-nonspecific alkaline phosphatase (*ALPL*), by real-time polymerase chain reaction (RT-PCR). The expression of *ALPL* during the differentiation was significantly higher in GD compared to WT cells (Figure 3A). Although this is in disagreement with other authors who found the expression of this gene lower in GD compared to WT,<sup>26,33</sup> other studies have reported no significant differences in the mRNA levels of this gene between MSCs of GD and WT.<sup>43</sup> In the same direction, Lecourt et al<sup>44</sup> showed that the activity of this enzyme did not change in condroitin B epoxide-treated MSCs, indicating that the inhibition of the GBA1 enzyme does not impact in the expression of *ALPL*. Regarding *RUNX2* expression along differentiation, higher levels were observed at day 7 in



**FIGURE 3** Relative mRNA expression of (A) *ALPL* and (B) *RUNX2* genes during the osteoblastic induction (time points: 0 [MSC], 7, 14, and 21 days) in wild-type (WT) and Gaucher disease (GD) cells. Hypoxanthine phosphoribosyltransferase (*HPRT*) was used as a housekeeping gene. Data represent the mean values (nine replicates) normalized by WT day 0 (=WT MSC)  $\pm$  SD. Statistical differences (\*)  $P$ -value < .05, (\*\*)  $P$ -value < .01, (\*\*\*)  $P$ -value < .001 and not significant (n.s) using Student's  $t$  test. See Data S1 for further details

both genotypes (GD and WT), corresponding to the MSC to osteoblasts transition. At all time-points, the expression of *RUNX2* was higher in WT cells (Figure 3B).

To evaluate the osteoblast functionality and, therefore, the mineralization of the extracellular matrix, we confirmed the existence of hydroxyapatite crystals (the bone mineral content) by the staining of calcium (Figure 4A) and phosphate (Figure 4B). Both calcium and phosphate were observed from day 14 in WT and GD cells.

The results of these experiments indicate that GD MSC seem to have a normal behavior, and they may differentiate properly to functional osteoblasts, in agreement with the model generated by Lecourt et al.<sup>44</sup> We can assume that the

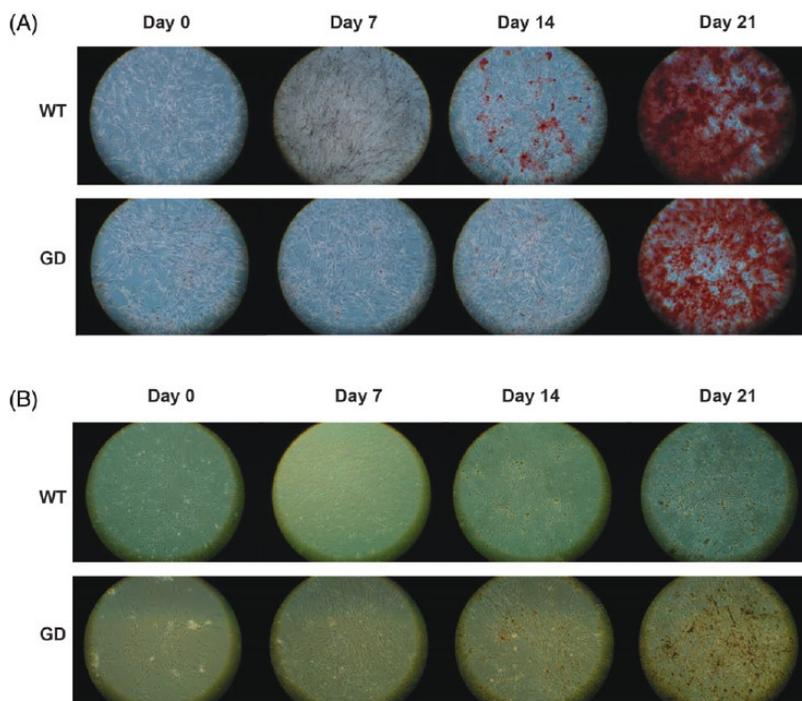
GD cells have a normal bone development, as expected, because bone problems in GD are due to defects in bone remodeling and not in bone development. The differences between WT and GD in the expression of *RUNX2* and *ALPL* seem not to affect the function of GD osteoblasts.

Because osteoblast function was unaffected, GD bone pathology could be due to an increase in osteoclast formation (osteoclastogenesis) or function. Osteoclasts are multinucleated, tartrate resistant acid phosphatase (TRAP) positive, and bone-resorbing giant cells derived from the differentiation and fusion of mononuclear hematopoietic progenitors' cells in the monocyte/macrophage lineage. The osteoblast-lineage cells (MSC, pre-osteoblasts, osteoblasts, and osteocytes) play an essential role in the regulation of osteoclastogenesis.

Osteoblast lineage cells and osteoclasts have a permanent cross talk through molecules such as the receptor activator of the nuclear factor kappa B (RANK) and its ligand (RANKL). Osteoblast lineage cells express in their surface and release RANKL that binds to the receptor RANK in the pre-osteoclasts' surface and promotes osteoclasts' differentiation, activation, and survival. We performed coculturing of MSC or osteoblasts derived either from WT or GD iPSC with monocytes from healthy female donors (Figure 5A). The objective was to assess whether the osteoclastogenesis, promoted by the WT or GD iPSC-derived MSC/osteoblasts, was similar. With 4',6-diamidino-2-phenylindole (DAPI) and TRITC-phalloidin, we were able to compare the nuclei number and the size of the osteoclasts generated by the two osteoblastic lineages (Figure 5B). The cells need to have at least three nuclei to be considered mature osteoclasts. The TRAP staining of these cells revealed the osteoclast identity (Figure 5C). The osteoclasts generated by GD iPSC-derived MSC/osteoblasts were larger than the ones generated by WT MSC/osteoblasts (Figure 5D). However, there were no differences in the number of nuclei per osteoclasts (Figure 5E), indicating that the larger osteoclasts are probably due to an increased spreading, rather than to an increment of cell fusion. We also analyze TRAP released into the culture media because it has been reported that it is in good correlation with the number of mature osteoclasts generated.<sup>45</sup> However, we did not find consistent differences between the values under both conditions, assuming that the number of osteoclasts generated by either WT or GD iPSC-derived MSC/osteoblasts was similar (data not shown).

In summary, the results presented here show that we were able to generate models of MSC and osteoblasts from GD- (and WT)-iPSC. However, it seems that these models do not reproduce faithfully the bone pathology in GD. Maybe the complexity of the bone pathology in this disease cannot be explained with cell-autonomous models. Similar to what had happened with previous cellular models, our model did not

**FIGURE 4** Histochemical assays at 0, 7, 14, and 21 days of the osteoblastic differentiation process in both, wild-type (WT) and Gaucher disease (GD) cells. (A) Alizarin red staining for calcium deposits. (B) von Kossa staining for phosphate deposits. See Data S1 for further details



allow the evaluation of the direct cross-talking with other cells and other molecules such as the cytokines, which are known to be very important in the bone remodeling.

### 3 | ATYPICAL FEMORAL FRACTURE

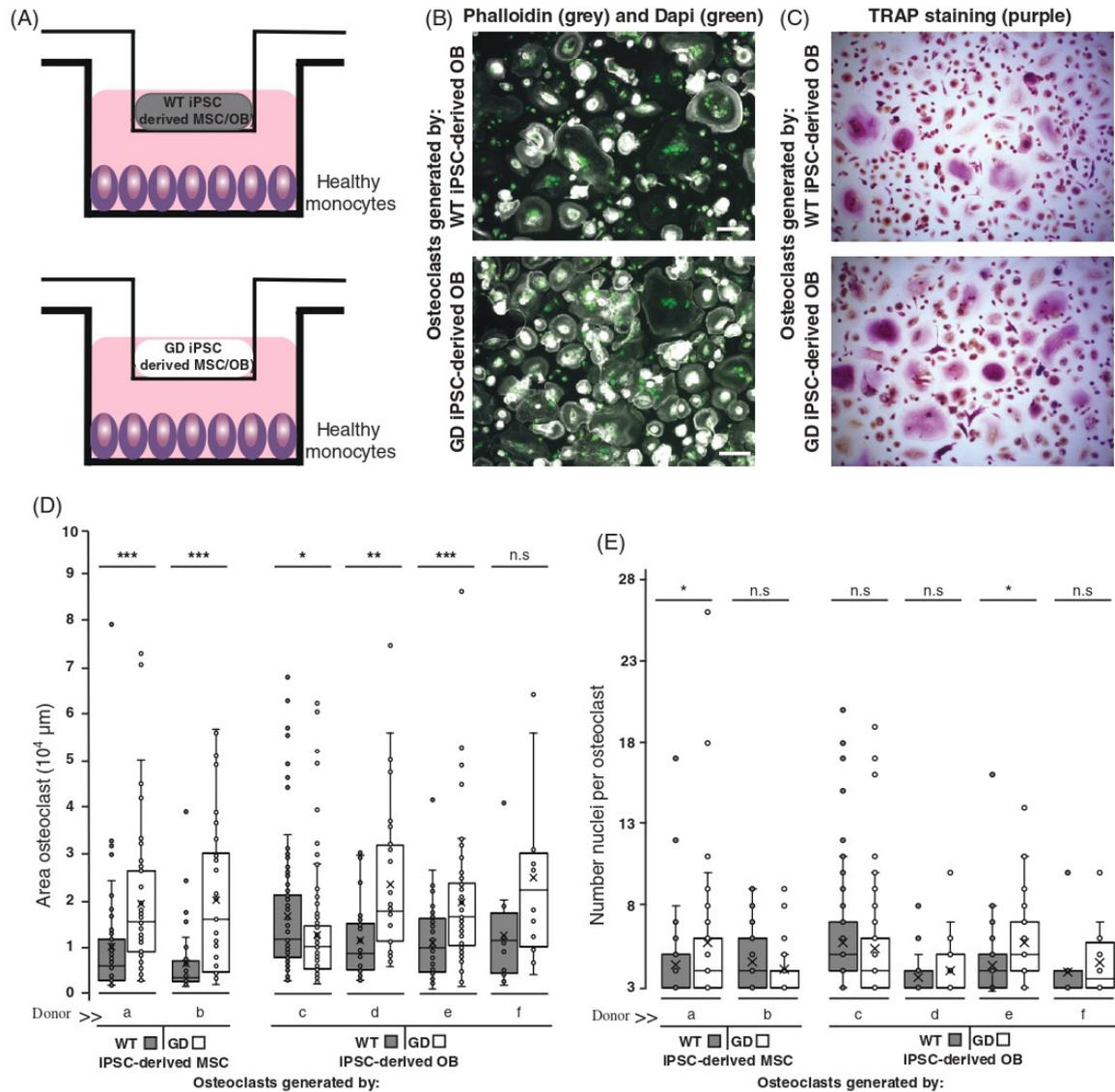
The last example is that of a particular type of bone fracture, known as atypical femoral fracture (AFF). AFFs were defined as atraumatic or low-trauma fractures located in the subtrochanteric region or femoral shaft. The diagnosis of AFF specifically excludes high trauma fractures, fractures of the femoral neck, intertrochanteric fractures with spiral subtrochanteric extension, pathological fractures associated with primary or metastatic bone tumors, and periprosthetic fractures. The fractures are usually not comminuted. Other characteristic radiographic features of AFFs (Figure 6A) include a transverse fracture line at the point of origination in the lateral cortex. As the fracture propagates across the diaphysis to the medial cortex, the orientation may become more oblique and when it becomes complete, a prominent medial “spike” may be present. There may be a focal or diffuse periosteal reaction of the lateral cortex surrounding the region where the fracture initiated. This reaction may appear as cortical “beaking” or “flaring” adjacent to a discrete transverse lucent fracture line or as focal thickening of the lateral cortex. Focal and diffuse endosteal reactions near the fracture

site have been reported more recently. This focal cortical thickening represents cortical hypertrophy and may be unilateral or bilateral. There may also be generalized cortical thickening.<sup>46</sup> Some epidemiological studies have suggested a relationship between AFFs and long-term bisphosphonates (BPs) therapy, the main treatment for osteoporosis.<sup>47,48</sup> Denosumab treatment has also been related to AFFs (reviewed in Anastasilakis et al<sup>49</sup>), although this relationship is not so clear. In any case, the pathogenesis of AFFs has not yet been elucidated.

We studied three sisters who had AFFs after receiving various oral BPs for 6 years.<sup>50</sup> Two of the sisters had a single fracture and one had bilateral fractures. Given the low incidence of AFFs in the general population (3.0–9.8 cases per 100 000 person-year,<sup>51</sup>), we hypothesized that these sisters might have an underlying genetic background that contributed to these fractures.

We performed whole-exome sequencing in the three sisters and in three unrelated patients with AFFs who each had received BPs for more than 5 years. We prioritized rare non-synonymous mutations in the variant filtering, and only mutations that were shared among the three sisters were considered. No mutation was found to be homozygous or in compound heterozygosity.

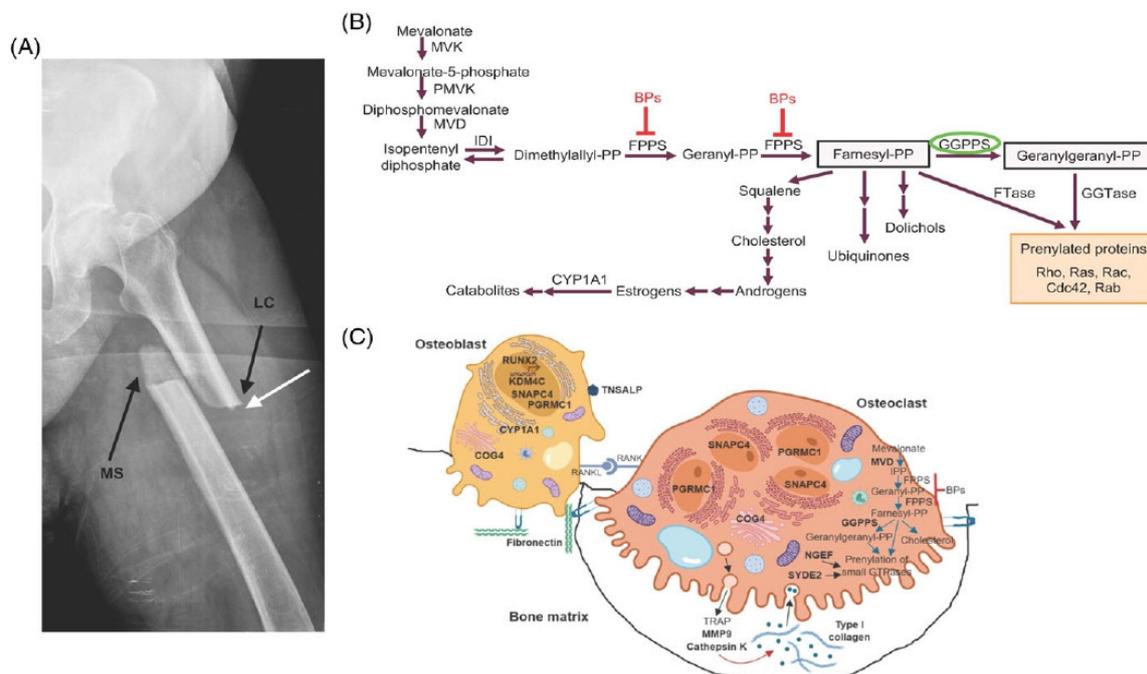
Assuming that a dominant model was involved, we detected 37 rare mutations (in 34 genes), among them a novel p.Asp188Tyr substitution in the enzyme



**FIGURE 5** (A) Schematic representation of the coculture experiment where the osteoclastogenesis of healthy monocytes induced by wild-type (WT) or Gaucher disease (GD) induced Pluripotent Stem Cells (iPSC)-derived MSC/osteoblasts can be evaluated. A membrane separates the well in two compartments. The upper one contains the WT or GD iPSC-derived mesenchymal stromal cells (MSC) or osteoblasts (OB) and the bottom one contains the peripheral blood CD14<sup>+</sup> monocytes from the same healthy female donor. Molecules, such as RANKL, but not cells can cross the membrane. (B) TRITC-Phalloidin staining (in gray) for the actin cytoskeleton and 4',6-diamidino-2-phenylindole (DAPI) staining (in green) for the nuclei Scale bar: 100 $\mu\text{m}$ . (C) Tartrate resistant acid phosphatase (TRAP) staining (in purple/pink) as a specific marker for osteoclasts. (D) Size and (E) number of nuclei per osteoclast analyzed. The letters in the x axis: a, b, c, d, e, and f correspond to different healthy female donors, whose monocytes were distributed in two identical parts, one to assess osteoclastogenesis generated by the WT cells and the other for the generated by the GD cells. Results are showed as median  $\pm$  confidence interval. Statistical differences (\*)  $P$ -value  $< .05$ , (\*\*)  $P$ -value  $< .01$ , and (\*\*\*)  $P$ -value  $< .001$  and not significant (n.s), using Mann-Whitney-Wilcoxon test. See Data S1 for further details

geranylgeranyl pyrophosphate synthase (GGPPS). This variant had the best conservation score and was not described in any of the available population databases. Interestingly, GGPPS is a homohexameric enzyme<sup>52</sup> that participates in

the mevalonate (or isoprenoid) pathway (Figure 6B), catalyzing a reaction just downstream of the main site of inhibition by BPs.<sup>53</sup> The mevalonate pathway leads to the production of cholesterol and isoprenoid lipids, such as



**FIGURE 6** (A) Radiograph of an atypical femoral fracture. Note the transverse fracture line that becomes oblique as it progresses (white arrow). LC: focal thickening of the lateral cortex; MS: medial spike. (B) Mevalonate pathway, indicating the steps affected by bisphosphonates (BPs, in red) and the position of GGPPS (circled in green). (C) Proteins found mutated in AFF patients (in bold) in the context of bonetissue

farnesyl diphosphate (FPP) and geranylgeranyl diphosphate (GGPP), required for the posttranslational prenylation of some proteins, including small GTPases. Prenylation of small GTPases is necessary for their activation and essential for osteoclast function and survival.<sup>54</sup>

We performed *in vitro* functional analyses of the GGPPS p.Asp188Tyr mutation, which showed a severe reduction in enzyme activity together with mild oligomerization defects.<sup>55</sup> Interestingly, another work on GGPPS p.Asp188Tyr mutation has been recently published<sup>56</sup> and the authors also observed a decreased catalytic activity of the mutated GGPPS, consistent with our results. In addition, they showed that it is unable to support cross-species complementation. In contrast, they only observed hexameric conformation of the enzyme in crystallographic experiments, although they saw a slight break of the tertiary symmetry, as well as a lower thermal stability of the mutated enzyme. In addition, the new tyrosine residue sterically interferes with substrate binding. Moreover, and considering that GGPPS can also be inhibited by BPs, although to a lesser extent,<sup>57</sup> they studied the affinity of the GGPPS-D188Y for zoledronate, a commonly used BP, demonstrating that it exhibited a reduction in the binding affinity although it could still be inhibited by zoledronate.

We also performed cellular functional assays, such as RNAi knockdown of the *GGPS1* gene in osteoblasts and in osteoclasts. In osteoblasts, RNAi produced a strong mineralization reduction and a reduced expression of the typical osteoblastic markers osteocalcin, osteonin, and RANKL, whereas in osteoclasts, it led to an increase of the osteoclast number with a lower resorption activity.<sup>55</sup>

We and others have shown that GGPPS p.Asp188Tyr mutation has an impact on protein function and relevant effects on bone cells, making it a strong candidate for AFF susceptibility. We propose that excessive inhibition of osteoclastic activity by the mutation plus BPs may lead to reduced bone remodeling and toughness, which may increase AFF susceptibility.

Few studies aiming to elucidate the genetics underlying AFF have been carried out, identifying some genes involved in AFF (Table 1). One of the most interesting genes is *CYP1A1*, an enzyme involved in the metabolism of drugs and xenobiotics and responsible for the hydroxylation of steroid hormones, such as estrogens<sup>63</sup> (Figure 6B). Mutations in *CYP1A1* were found in the three sisters and one unrelated patient in our study<sup>50,55</sup> and also reported elsewhere in two patients with AFF and glucocorticoid-induced osteoporosis.<sup>58</sup> Another study identified a missense variant in the *PPEF2* gene significantly associated with AFF by exon

TABLE 1 Genes mutated in AFF patients in cohort studies

Gene	Mutations	ExAC freq.	N° AFF cases	Years BPs	Patients characteristics	Genetic analysis	Ref.
<i>GGPS1</i>	p.Asp188Tyr(Het.)	Novel	3 sisters	6	–	WES	50,55
<i>CYP11A1</i>	p.Arg98Trp(Het.)	0.0001	3 sisters	6	–	WES	55
<i>CYP11A1</i>	p.Ser216Cys(Het.)	0.0001	1 unrelated patient	6	–	WES	55
<i>CYP11A1</i>	p.Arg136His(Het.)	2.14e-5	1 patient	3	GCC-OP	Sanger seq	58
<i>CYP11A1</i>	p.Val409Ile(Het.)	0.0003 <sup>a</sup>	1 patient	12	GCC-OP; RA	Sanger seq	58
<i>PPEF2</i>	p.Arg388Gln	0.001	13 patients	1-10	–	Exon array	59
<i>COL1A2</i>	p.Arg708Gln(Het.)	0.0008	1 patient	>5	No OI features	Sanger seq.	60
<i>CTSK</i>	c.784+3A>C(Homozyg.)	5.77e-5	2 consanguineous sisters	0	No pycnodysostosis features	WES	61
<i>ALPL</i>	p.Gly288Ala(Het.)	Novel	1 patient	8	HPP	Sanger seq.	58
<i>ALPL</i>	c.648+1G>A(Het.)	8.24e-6	1 patient	NA <sup>b</sup>	HPP	Sanger seq.	62

Abbreviations: AR, rheumatoid arthritis; GCC-OP, glucocorticoid-induced osteoporosis; HPP, hypophosphatasia unmasked after AFF.

<sup>a</sup>Not present in ExAC, frequency from ALSPAC cohort.

<sup>b</sup>Duration of treatment not specified but the analysis was carried out during BP treatment.

array analysis of a small cohort.<sup>59</sup> To date, this gene has no known function in bone metabolism. Some studies have identified genes previously known to be involved in monogenetic bone diseases. Recently, Funck-Brentano et al<sup>60</sup> identified a heterozygous mutation in *COL1A2*, coding for type 1 collagen, a major component of the bone extracellular matrix. Mutations in *COL1A2* cause osteogenesis imperfecta (OI). However, no specific physical features of OI were identified in this patient, apart from short stature. In addition, Lau et al<sup>61</sup> identified a homozygous mutation in *CTSK* gene in a consanguineous family, encoding for cathepsin K and related to pycnodysostosis, although the patient identified had no clinical features of this disease. Finally, some cohort studies identified AFFs in patients of hypophosphatasia initially misdiagnosed with postmenopausal osteoporosis and treated with BPs.<sup>58,62</sup> Heterozygous mutations in tissue non-specific alkaline phosphatase (TNSALP; *ALPL* gene) were identified. In these cases, it might be that mild, unrecognized forms of some monogenetic bone diseases underlie the etiology of AFFs. Taken together, all the genetic studies carried out in AFF patients advocate a heterogeneous genetic component of predisposition to AFF, in which each individual patient would be a carrier of different specific genetic variant/s (Figure 6C). In addition, in our study, we identified some other interesting rare variants in the three sisters and unrelated patients (eg, *MMP9*, *MVD*, *RUNX2*, *NGEF*, *SYDE2*, *FNI*)<sup>64</sup>, suggesting a polygenic or complex background involving, to different extent, several additional variants in the susceptibility to AFF. All in all, we speculate that an accumulation of a few susceptibility variants from different pathways and their interactions constitute the genetic predisposition that, together with BPs and/or other comorbid

conditions, give rise to AFF, in what we might call “the perfect storm.” Further identification and/or replication of genetic variants, as well as functional studies of the identified variants, are needed to detect at-risk individuals for clinical decision-making.

## ACKNOWLEDGMENTS

The authors are grateful to M. Cozar for technical assistance and to patient DR and her family for collaboration and permanent support. J.S.-V. was a recipient of an FI fellowship from the Catalan government and NRA was a recipient of an FPU fellowship from the Spanish government. This work was supported by the following grants: SAF2016-75948-R (Spanish MINECO), 2014SGR932 (Generalitat de Catalunya), and CIBERER (U720).

## CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

## ORCID

Daniel Grinberg  <https://orcid.org/0000-0001-9859-2590>

## REFERENCES

- Wuyts W, van Hul W. Molecular basis of multiple exostoses: mutations in the EXT1 and EXT2 genes. *Hum Mutat.* 2000;15(3):220-227.
- Stickens D, Evans GA. A sugar fix for bone tumours? *Nat Genet.* 1998;19(2):110-111.

3. Bovee JVMG. EXTra hit for mouse osteochondroma. *Proc Natl Acad Sci USA*. 2010;107(5):1813-1814.
4. Pacifici M. The pathogenic roles of heparan sulfate deficiency in hereditary multiple exostoses. *Matrix Biol*. 2018;71-72:28-39.
5. Beltrami G, Ristori G, Scoccianti G, Tamburini A, Capanna R. Hereditary multiple exostoses: a review of clinical appearance and metabolic pattern. *Clin Cases Miner Bone Metab*. 2016;13(2):110-118.
6. Czajka CM, DiCaprio MR. What is the proportion of patients with multiple hereditary exostoses who undergo malignant degeneration? *Clin Orthop Relat Res*. 2015;473(7):2355-2361.
7. Musso N, Caronia FP, Castorina S, Lo Monte AI, Barresi V, Condorelli DF. Somatic loss of an EXT2 gene mutation during malignant progression in a patient with hereditary multiple osteochondromas. *Cancer Genet*. 2015;208(3):62-67.
8. Sarrion P, Sangorrin A, Urreiziti R, et al. Mutations in the EXT1 and EXT2 genes in Spanish patients with multiple osteochondromas. *Sci Rep*. 2013;3:1346.
9. Cammarata-Scalisi F, Cozar M, Grinberg D, et al. Double mutant alleles in the EXT1 gene not previously reported in a teenager with hereditary multiple exostoses. *Arch Argent Pediatr*. 2015;113(2):e109-e112.
10. Cammarata-Scalisi F, Stock F, Avendaño A, Cozar M, Balcells S, Grinberg D. Clinical and molecular study in a family with multiple osteochondromatosis. *Acta Ortop Mex*. 2018;32(2):108-111.
11. Delgado MA, Martínez-Domenech G, Sarrion P, et al. A broad spectrum of genomic changes in latinamerican patients with EXT1/EXT2-CDG. *Sci Rep*. 2014;4:6407.
12. Delgado MA, Sarrion P, Azar N, et al. A novel nonsense mutation of the EXT1 gene in an Argentinian patient with multiple hereditary exostoses: a case report. *J Bone Joint Surg Am*. 2012;94(11):e76.
13. El-Bazzal L, Atkinson A, Gillart AC, et al. A novel EXT2 mutation in a consanguineous family with severe developmental delay, microcephaly, seizures, feeding difficulties, and osteopenia extends the phenotypic spectrum of autosomal recessive EXT2-related syndrome (AREXT2). *Eur J Med Genet*. 2019;63(4):259-264.
14. Farhan SM, Wang J, Robinson JF, et al. Old gene, new phenotype: mutations in heparan sulfate synthesis enzyme, EXT2 leads to seizure and developmental disorder, no exostoses. *J Med Genet*. 2015;52(10):666-675.
15. Zhang P, Lu H, Peixoto RT, et al. Heparan sulfate organizes neuronal synapses through neuroligin partnerships. *Cell*. 2018;174(6):1450-1464.
16. Canals I, Benetó N, Cozar M, Vilageliu L, Grinberg D. EXTL2 and EXTL3 inhibition with siRNAs as a promising substrate reduction therapy for Sanfilippo C syndrome. *Sci Rep*. 2015a;5:13654.
17. Canals I, Soriano J, Orlandi JG, et al. Activity and high-order effective connectivity alterations in Sanfilippo C patient-specific neuronal networks. *Stem Cell Reports*. 2015b;5(4):546-557.
18. Futerman AH, van Meer G. The cell biology of lysosomal storage disorders. *Nat Rev Mol Cell Biol*. 2004;5(7):554-565.
19. Sidransky E. Gaucher disease: complexity in a 'simple' disorder. *Mol Genet Metab*. 2004;83(1-2):6-15.
20. Beutler E, Grabowski GA. Gaucher disease. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The metabolic and molecular bases of inherited disease*. 7th ed. New York, NY: McGraw-Hill; 1995:2641-2669.
21. Davidson BA, Hassan S, Garcia EJ, Tayebi N, Sidransky E. Exploring genetic modifiers of Gaucher disease: the next horizon. *Hum Mutat*. 2018;39(12):1739-1751.
22. Giraldo P, Solano V, Pérez-Calvo JI, Giralto M, Rubio-Félix D, Spanish Group on Gaucher disease. Quality of life related to type 1 Gaucher disease: Spanish experience. *Qual Life Res*. 2005;14(2):453-462.
23. Mikosch P, Hughes D. An overview on bone manifestations in Gaucher disease. *Wien Med Wochenschr*. 2010;160(23-24):609-624.
24. van Dussen L, Lips P, Everts VE, et al. Markers of bone turnover in Gaucher disease: modeling the evolution of bone disease. *J Clin Endocrinol Metab*. 2011;96(7):2194-2205.
25. Campeau PM, Rafei M, Boivin MN, Sun Y, Grabowski GA, Galipeau J. Characterization of Gaucher disease bone marrow mesenchymal stromal cells reveals an altered inflammatory secretome. *Blood*. 2009;114(15):3181-3190.
26. Mistry PK, Liu J, Yang M, et al. Glucocerebrosidase gene-deficient mouse recapitulates Gaucher disease displaying cellular and molecular dysregulation beyond the macrophage. *Proc Natl Acad Sci USA*. 2010;107(45):19473-19478.
27. Mistry PK, Liu J, Sun L, et al. Glucocerebrosidase 2 gene deletion rescues type 1 Gaucher disease. *Proc Natl Acad Sci USA*. 2014;111(13):4934-4939.
28. Bondar C, Mucci J, Crivaro A, et al. In vitro osteoclastogenesis from Gaucher patients' cells correlates with bone mineral density but not with Chitotriosidase. *Bone*. 2017;103:262-269.
29. Mucci JM, Scian R, De Francesco PN, et al. Induction of osteoclastogenesis in an in vitro model of Gaucher disease is mediated by T cells via TNF- $\alpha$ . *Gene*. 2012;509(1):51-59.
30. Mucci JM, Suqueli García F, de Francesco PN, et al. Uncoupling of osteoblast-osteoclast regulation in a chemical murine model of Gaucher disease. *Gene*. 2013;532(2):186-191.
31. Reed M, Baker RJ, Mehta AB, et al. Enhanced differentiation of osteoclasts from mononuclear precursors in patients with Gaucher disease. *Blood Cells Mol Dis*. 2013;51(3):185-194.
32. Reed MC, Bauernfreund Y, Cunningham N, Beaton B, Mehta AB, Hughes DA. Generation of osteoclasts from type 1 Gaucher patients and correlation with clinical and genetic features of disease. *Gene*. 2018;678:196-206.
33. Panicker LM, Srikanth MP, Castro-Gomes T, Miller D, Andrews NW, Feldman RA. Gaucher disease iPSC-derived osteoblasts have developmental and lysosomal defects that impair bone matrix deposition. *Hum Mol Genet*. 2018;27(5):811-822.
34. Estrada K, Styrkarsdottir U, Evangelou E, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet*. 2012;44(5):491-501.
35. Martínez-Gil N, Roca-Ayats N, Monistrol-Mula A, et al. Common and rare variants of WNT16, DKK1 and SOST and their relationship with bone mineral density. *Sci Rep*. 2018;8(1):10951.
36. Zheng HF, Forgetta V, Hsu YH, et al. Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature*. 2015;526(7571):112-117.
37. Tiscornia G, Vivas EL, Izpisua Belmonte JC. Diseases in a dish: modeling human genetic disorders using induced pluripotent cells. *Nat Med*. 2011;17(12):1570-1576.
38. Park IH, Arora N, Huo H, et al. Disease-specific induced pluripotent stem cells. *Cell*. 2008;134(5):877-886.

39. Colter DC, Class R, DiGirolamo CM, et al. Rapid expansion of recycling stem cells in cultures of plastic-adherent cells from human bone marrow. *Proc Natl Acad Sci USA*. 2000;97(7):3213-3218.
40. DiGirolamo CM, Stokes D, Colter D, et al. Propagation and senescence of human marrow stromal cells in culture: a simple colony-forming assay identifies samples with the greatest potential to propagate and differentiate. *Br J Haematol*. 1999;107(2):275-281.
41. Bayreuther K, Rodemann HP, Hommel R, Dittmann K, Albiez M, Francz PI. Human skin fibroblasts in vitro differentiate along a terminal cell lineage. *Proc Natl Acad Sci USA*. 1988;85(14):5112-5116.
42. von Zglinicki T, Petrie J, Kirkwood TB. Telomere-driven replicative senescence is a stress response. *Nat Biotechnol*. 2003;21(3):229-230.
43. Campeau PM, Rafei M, Boivin MN, et al. Characterization of Gaucher disease bone marrow mesenchymal stromal cells reveals an altered inflammatory secretome. *Blood*. 2009;114(15):3181-3190.
44. Lecourt S, Vanneaux V, Cras A, et al. Bone marrow microenvironment in an in vitro model of Gaucher disease: consequences of glucocerebrosidase deficiency. *Stem Cells Dev*. 2012;21(2):239-248.
45. Alatalo SL, Halleen JM, Hentunen TA, Mönkkönen J, Väinänen HK. Rapid screening method for osteoclast differentiation in vitro that measures tartrate-resistant acid phosphatase 5b activity secreted into the culture medium. *Clin Chem*. 2000;46(11):1751-1754.
46. Shane E, Burr D, Abrahamsen B, et al. Atypical subtrochanteric and diaphyseal femoral fractures: second report of a task force of the American Society for Bone and Mineral Research. *J Bone Miner Res*. 2014;29(1):1-23.
47. Gedmintas L, Solomon DH, Kim SC. Bisphosphonates and risk of subtrochanteric, femoral shaft, and atypical femur fracture: a systematic review and meta-analysis. *J Bone Miner Res*. 2013;28:1729-1737.
48. Liu L, Li C, Yang P, et al. Association between alendronate and atypical femur fractures: a meta-analysis. *Endocr Connect*. 2015;4:58-64.
49. Anastasilakis AD, Polyzos SA, Makras P. Therapy of endocrine disease: Denosumab vs bisphosphonates for the treatment of postmenopausal osteoporosis. *Eur J Endocrinol*. 2018;179(1):R31-R45.
50. Roca-Ayats N, Balcells S, Garcia-Giralt N, et al. GGPS1 mutation and atypical femoral fractures with bisphosphonates. *N Engl J Med*. 2017;376(18):1794-1795.
51. Khow KSF, Shibu P, Yu SCY, Chehade MJ, Visvanathan R. Epidemiology and postoperative outcomes of atypical femoral fractures in older adults: a systematic review. *J Nutr Health Aging*. 2017;21:83-91.
52. Kavanagh KL, Dunford JE, Bunkoczi G, Russell RGG, Oppermann U. The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a novel hexameric arrangement and inhibitory product binding. *J Biol Chem*. 2006a;281:22004-22012.
53. Kavanagh KL, Guo K, Dunford JE, et al. The molecular mechanism of nitrogen-containing bisphosphonates as antiosteoporosis drugs. *Proc Natl Acad Sci USA*. 2006b;103:7829-7834.
54. Itzstein C, Coxon FP, Rogers MJ. The regulation of osteoclast function and bone resorption by small GTPases. *Small GTPases*. 2011;2:117-130.
55. Roca-Ayats N, Ng PY, Garcia-Giralt N, et al. Functional characterization of a GGPPS variant identified in atypical femoral fracture patients and delineation of the role of GGPPS in bone-relevant cell types. *J Bone Miner Res*. 2018b;33(12):2091-2098.
56. Lisnyansky M, Kapelushnik N, Ben-Bassat A, et al. Reduced activity of geranylgeranyl diphosphate synthase mutant is involved in bisphosphonate-induced atypical fractures. *Mol Pharmacol*. 2018;94:1391-1400.
57. Guo R-T, Cao R, Liang P-H, et al. Bisphosphonates target multiple sites in both cis- and trans-prenyltransferases. *Proc Natl Acad Sci USA*. 2007;104:10022-10027.
58. Peris P, González-Roca E, Rodríguez-García SC, et al. (2019) Incidence of mutations in the ALPL, GGPS1, and CYP11A1. *JBMR Plus* 3(1):29-36
59. Pérez-Núñez I, Pérez-Castrillón JL, Zarrabeitia MT, et al. Exon array analysis reveals genetic heterogeneity in atypical femoral fractures. A pilot study. *Mol Cell Biochem*. 2015;409:45-50.
60. Funck-Brentano T, Ostertag A, Debais F, et al. Identification of a p.Arg708Gln variant in COL1A2 in atypical femoral fractures. *Joint Bone Spine*. 2017;84:715-718.
61. Lau S-L, McInemey-Leo AM, Lanoda-Bassonga E, et al. Homozygous variant in Cathepsin K in a family with multiple cases of bilateral atypical femoral fracture but without clinical features of pyknodysostosis. *JBMR Plus*. 2017;1(S1):S1-S11.
62. Sum M, Huskey M, Diemer K, et al. TNSALP mutation analysis in women with atypical femoral fracture and bisphosphonate therapy for osteoporosis. *J Bone Miner Res*. 2013;28(S1):S295.
63. Zhou S-F, Liu J-P, Chowbay B. Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metab Rev*. 2009;41:89-295.
64. Roca-Ayats N, Falcó-Mascaró M, Garcia-Giralt N, et al. Genetic study of atypical femoral fractures using exome sequencing in three affected sisters and three unrelated patients. *Rev Osteoporos Metab Miner*. 2018a;10(4):108-118.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Serra-Vinardell J, Roca-Ayats N, De-Ugarte L, Vilageliu L, Balcells S, Grinberg D. Bone development and remodeling in metabolic disorders. *J Inherit Metab Dis*. 2019;1–12. <https://doi.org/10.1002/jimd.12097>

