Department of Molecular Medicine and Surgery
Karolinska Institutet, Stockholm, Sweden

GENETIC CAUSES AND MOLECULAR MECHANISMS UNDERLYING RARE METABOLIC BONE DISEASES

Alice Costantini

Karolinska Institutet

Stockholm 2019

Cover art by Alessandro Berlese.
Other illustrations produced by Alice Costantini in collaboration with Alessandro Berlese.

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by E-print AB

ISBN 978-91-7831-471-3

Genetic causes and molecular mechanisms underlying rare metabolic bone diseases

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Alice Costantini

Principal Supervisor:
Professor Outi Mäkitie
Karolinska Institutet
Department of Molecular Medicine and Surgery

Co-supervisor(s):
Professor Magnus Nordenskjöld
Karolinska Institutet
Department of Molecular Medicine and Surgery

Associate professor Hong Jiao
Karolinska Institutet
Department of Biosciences and Nutrition

Assistant professor Fulya Taylan
Karolinska Institutet
Department of Molecular Medicine and Surgery

Professor Olle Kämpe
Karolinska Institutet
Department of Medicine

Opponent:
Professor Uwe Kornak
Charité - Universitätsmedizin Berlin
Institute of Medical Genetics and Human Genetics

Examination Board:
Associate professor Eva-Lena Stattin
Uppsala University
Department of Immunology, Genetics and Pathology

Professor Göran Andersson
Karolinska Institutet
Department of Laboratory Medicine

Professor Björn Andersson
Karolinska Institutet
Department of Cell and Molecular Biology

To my parents

“Only those who will risk going too far
can possibly find out how far one can go.”

T.S. Eliot

ABSTRACT

The skeletal system provides support for the body, enables movement and protects inner organs. Moreover, it supplies blood cells and acts as a reservoir for minerals and fat.

Several external factors, including nutrition and long-term illness, influence bone health but genetic factors also play an important role. More than 400 different rare skeletal diseases, collectively called skeletal dysplasias, have thus far been delineated and mutations in over 350 genes have been identified as underlying causes in these conditions. Although the recent evolution of the sequencing technologies and molecular methods has increased diagnostic yield of rare skeletal diseases, knowledge on the genetic and phenotypic features in some of these conditions is still limited and novel forms of skeletal dysplasia still remain to be characterized.

This thesis focused on rare skeletal diseases primarily affecting the major component of the skeleton, the bone. In **paper I** and **III** Sanger sequencing was used. In **paper I** this method excluded the presence of rare variants in *CRTAP*, encoding the cartilage associated protein, in patients with mild-to-severe skeletal fragility. In **paper III** two novel mutations in two components of the WNT signaling pathway, *LRP5* and *AMER1*, were identified in two patients affected by high bone mass. In **paper II** a custom designed high-resolution array-CGH, targeting all the genes thus far linked to skeletal diseases and the cilia genes, enabled the identification of two novel copy number variants (CNVs) affecting *COL1A2* and *PLS3* in two index patients with primary osteoporosis. Other rare CNVs in genes not yet related to bone homeostasis were detected and regarded as variants of unknown significance. In **papers IV** and **V** massively-parallel sequencing was applied. In **paper IV** five novel variants in the fibronectin gene (*FN1*), which was recently linked to spondylometaphyseal dysplasia with “corner fractures”, were revealed in five patients affected by this disease. Finally, in **paper V** two novel variants in the gene encoding the ribosomal protein L13, *RPL13*, were for the first time associated with a novel form of spondyloepimetaphyseal dysplasia.

Our findings expand the genetic and phenotypic spectrum of some known rare skeletal diseases. Moreover, a novel gene-disease association was identified but further studies are required to explore the pathomolecular mechanisms underlying this condition. Studying rare metabolic bone diseases is important not only for arriving at a specific diagnosis but also for understanding the pathogenesis of these conditions - only an increased understanding of the molecular mechanisms will enable the development of targeted therapies.

LIST OF SCIENTIFIC PAPERS

I. ***CRTAP* variants in early-onset osteoporosis and recurrent fractures**.
Costantini A, Vuorimies I, Makitie R, Mayranpaa MK, Becker J, Pekkinen M, Valta H, Netzer C, Kampe A, Taylan F, Jiao H, Makitie O. *Am J Med Genet A* 173(3) (2017) 806-808.

II. **Rare Copy Number Variants in Array-Based Comparative Genomic Hybridization in Early-Onset Skeletal Fragility**.
Costantini A, Skarp S, Kämpe A, Mäkitie RE, Pettersson M, Männikko M, Jiao H, Taylan F, Lindstrand A[#], Mäkitie O[#]. *Front Endocrinol* 9 (2018) 380.

III. **High bone mass due to novel *LRP5* and *AMER1* mutations**.
Costantini A, Kekäläinen P, Mäkitie RE, Mäkitie O. *Eur J Med Genet* 60(12) (2017) 675-679.

IV. **Novel fibronectin mutations and expansion of the phenotype in spondylometaphyseal dysplasia with "corner fractures"**.
Costantini A, Valta H, Baratang NV, Yap P, Bertola DR, Yamamoto GL, Kim CA, Chen J, Wierenga KJ, Fanning EA, Escobar L, McWalter K, McLaughlin H, Willaert R, Begtrup A, Alm JJ, Reinhardt DP, Mäkitie O[#], Campeau PM[#]. *Bone* 121 (2019) 163-171.

V. ***RPL13* variants in spondyloepimetaphyseal dysplasia**.
Costantini A, Alm JJ, Tonelli F[§], Valta H[§], Tran A, Chen S, Chagin A, Newton P, Daponte V, Kwon YU, Bae JY, Chung WY, Larsson O, Nishimura G, Näreoja T, Kim OH, Forlino A[#], Cho TJ[#], Mäkitie O[#]. *Manuscript* (2019)

[#] shared senior authorship

[§] shared third authorship

ADDITIONAL PUBLICATIONS

I. **New Insights Into Monogenic Causes of Osteoporosis**.
Mäkitie RE, Costantini A, Kämpe A, Alm JJ, Mäkitie O. *Front Endocrinol* 10 (2019) 70.

II. **A novel frameshift deletion in *PLS3* causing severe primary osteoporosis**.
Costantini A, Krallis P, Kämpe A, Karavitakis EM, Taylan F, Mäkitie O, Doulgeraki A. *J Hum Genet* 63(8) (2018) 923-926.

III. **Autosomal Recessive Osteogenesis Imperfecta Caused by a Novel Homozygous *COL1A2* Mutation**.
Costantini A, Tournis S, Kämpe A, Ul Ain N, Taylan F, Doulgeraki A, Mäkitie O. *Calcif Tissue Int* 103(3) (2018) 353-358.

IV. **Expansion of the clinical spectrum of frontometaphyseal dysplasia 2 caused by the recurrent mutation p.Pro485Leu in *MAP3K7***.
Costantini A, Wallgren-Pettersson C, Mäkitie O. *Eur J Med Genet* 61(10) (2018) 612-615.

V. **A novel *MYT1L* mutation in a patient with severe early-onset obesity and intellectual disability**.
Loid P, Mäkitie R, Costantini A, Viljakainen H, Pekkinen M, Mäkitie O. *Am J Med Genet A* 176(9) (2018) 1972-1975.

VI. ***PLS3* Deletions Lead to Severe Spinal Osteoporosis and Disturbed Bone Matrix Mineralization**.
Kampe AJ, Costantini A, Levy-Shraga Y, Zeitlin L, Roschger P, Taylan F, Lindstrand A, Paschalis EP, Gamsjaeger S, Raas-Rothschild A, Hovel M, Jiao H, Klaushofer K, Grasemann C, Makitie O. *J Bone Miner Res* 32(12) (2017) 2394-2404

VII. ***PLS3* sequencing in childhood-onset primary osteoporosis identifies two novel disease-causing variants**.
Kämpe AJ, Costantini A, Mäkitie RE, Jäntti N, Valta H, Mäyränpää M, Kröger H, Pekkinen M, Taylan F, Jiao H, Mäkitie O. *Osteoporos Int* 28(10) (2017) 3023-3032.

VIII. **Spondyloocular Syndrome: Novel Mutations in *XYLT2* Gene and Expansion of the Phenotypic Spectrum**.
Taylan F, Costantini A, Coles N, Pekkinen M, Héon E, Şiklar Z, Berberoğlu M, Kämpe A, Kiykim E, Grigelioniene G, Tüysüz B, Mäkitie O. *J Bone Miner Res* 31(8) (2016) 1577-85.

IX. **Value of rare low bone mass diseases for osteoporosis genetics**.
Costantini A, Mäkitie O. *Bonekey Rep* 5 (2016) 773.

PREFACE

My PhD journey started in January 2015 when I joined Outi Mäkitie's team in Clinical Genetics at Karolinska Institutet. During my undergraduate studies I became fascinated by the field of genetics but only during my PhD I became passionate about performing research on rare skeletal diseases. I feel grateful to have been a PhD student in medical genetics during these exciting times, in which the rapid development and application of massively-parallel sequencing technologies has tremendously revolutionized the way of approaching patients with rare monogenic diseases. My PhD project, which started with the use of Sanger sequencing to then move to high-resolution array-CGH and finally to high-throughput sequencing, is a proof of the shift from clinical genetics to clinical genomics. My PhD studies enabled me to learn different methods, to deepen my knowledge in clinical genetics, to acquire understanding of bone homeostasis and skeletal diseases as well as to develop myself as a researcher.

Stockholm, May 8th 2019

TABLE OF CONTENTS

1 INTRODUCTION 1
1.1 BONE STRUCTURE 1
1.1.1 Bone cells and extracellular matrix 2
1.1.2 Bone strength 3
1.2 SKELETAL DEVELOPMENT 4
1.2.1 Intramembranous ossification 4
1.2.2 Endochondral ossification 4
1.3 BONE HOMEOSTASIS 5
1.3.1 Bone modeling and remodeling 6
1.3.2 Calcium and phosphate homeostasis 7
1.3.3 Other regulators 8
1.4 THE HUMAN GENOME 9
1.4.1 Human genetic variation 9
1.4.2 Types of genetic variants 10
1.4.3 Rare monogenic diseases 11
1.4.4 Challenges in variant interpretation 12
1.5 RARE SKELETAL DISEASES 12
1.5.1 Diseases affecting growth 13
1.5.2 Diseases affecting bone homeostasis 16
1.5.3 Treatments 20
1.6 APPROACHES TO INVESTIGATE GENETIC SKELETAL DISEASES 22
1.6.1 From genetics to genomics 23
2 AIMS 25
3 PATIENTS AND METHODS 27
3.1 ETHICAL APPROVALS AND PATIENTS' CONSENTS 27
3.2 PATIENTS 27
3.3 METHODS 28
3.3.1 Genetic testing 28
3.3.2 Variant filtering, prioritization and interpretation 32
3.3.3 Variant validation 33
4 RESULTS AND DISCUSSION 37
4.1 SANGER SEQUENCING OF KNOWN DISEASE GENES 37
4.1.1 *CRTAP* screening in patients with skeletal fragility 37
4.1.2 Genes associated with HBM diseases 38
4.2 CUSTOM DESIGNED ARRAY-CGH TARGETING GENES RELATED TO SKELETAL FRAGILITY AND CILIARY FUNCTION 40
4.2.1 CNVs in known genes associated to skeletal fragility 40
4.2.2 CNVs in genes not yet associated with skeletal diseases 41
4.3 MPS TO SEARCH FOR DISEASE-CAUSING GENETIC DEFECTS IN RARE SKELETAL PHENOTYPES 43
4.3.1 Studies in SMD with "corner fractures" 43
4.3.2 Studies in SEMD 45
4.4 STUDY LIMITATIONS 49
4.5 ETHICAL CONSIDERATIONS 49
5 CONCLUSIONS 51
6 FUTURE PERSPECTIVES 55
7 ACKNOWLEDGMENTS 57
8 REFERENCES 61

LIST OF ABBREVIATIONS

$1,25(OH)_2$: 1,25-dihydroxyvitamin D (active vitamin D)

ACMG: American College of Medical Genetics

AMER1: adenomatous polyposis coli membrane recruitment 1

Array-CGH: comparative genomic hybridization arrays

BMD: bone mineral density

BMP: bone morphogenetic protein

BWA: Burrows-Wheeler Aligner

CADD: Combined Annotation-Dependent Depletion (CADD)

CNV: copy number variant

COL1A1/COL1A2: type I collagen chain 1/2

COMP: cartilage oligomeric matrix protein

CRTAP: cartilage associated protein

ddNTP: dideoxyribonucleotide triphosphate

DGV: Database of Genetic Variation

DKK: Dickkopf

dNTP: deoxyribonucleotide triphosphate

Dpf: days post-fertilization

DXA: dual-energy X-ray absorptiometry

eBMD: estimated BMD

ECM: extracellular matrix

ExAc: Exome Aggregation Consortium

F1: first filial

FGF: fibroblast growth factor

FZD: frizzled

GERP: Genomic Evolutionary Rate Profiling

gnomAD: Genome Aggregation Consortium

GOF: gain-of-function

GWAS: genome-wide association study

HBM: high bone mass

HPP: hypophosphatasia

HRP: horseradish peroxidase

HSC: hematopoietic stem cells

ICC: immunocytochemistry

IGV: Integrative Genomics Viewer

IHH: Indian hedgehog

LOF: loss-of-function

LRP: lipoprotein receptor related protein

MAF: minor allele frequency

MPS: massively-parallel sequencing

mRNA: messenger RNA

MSC: mesenchymal stem cell

NHEJ: non-homologous end joining

OCN: osteocalcin

OI: osteogenesis imperfecta

OMIM: Online Mendelian Inheritance in Man

OPG: osteoprotegerin

ORF: open reading frame

OSCS: osteopathia striata with cranial sclerosis

PCR: polymerase chain reaction

PLS: plastin

POC: primary ossification center

PTH: parathyroid hormone

PTHrP: parathyroid hormone-related protein

RANK: receptor activator of nuclear factor kappa-B

RANKL: receptor activator of nuclear factor kappa-B ligand

rER: rough endoplasmic reticulum

RP: ribosomal protein

RPL: L ribosomal protein

SD: standard deviation

SEMD: spondyloepimetaphyseal dysplasia

sgRNA: single guide RNA

SMD-CF: spondylometaphyseal dysplasia with “corner fractures”

SMD: spondylometaphyseal dysplasia

SNP: single nucleotide polymorphism

SNV: single nucleotide variant

SOC: secondary ossification center

SOST: sclerostin

SV: structural variant

TGF-β: transforming growth factor β

TNSALP: tissue-nonspecific alkaline phosphatase

VEP: Variant Effect Predictor

VUS: variant of uncertain significance

WB: Western blot

WES: whole-exome sequencing

WGS: whole-genome sequencing

WNT: wingless-type

XYLT: xylosyltransferase

1 INTRODUCTION

The skeletal system has several vital functions: it supports the body, enables movement and protects inner organs. Furthermore, it acts as a source of blood cells and a reservoir for minerals, mainly calcium and phosphate, as well as fat. The skeleton is composed of several tissues including bone, cartilage, tendons, ligaments and other connective tissues.

Several external factors influence bone health. These factors include nutrition, physical activity, potential long-term illnesses (e.g. inflammatory diseases) and medications such as glucocorticoid therapy. However, genetic factors also play an important role. This thesis will focus on rare skeletal diseases that are caused by a single genetic defect (monogenic, or Mendelian conditions) and primarily affecting the major component of the skeleton, the bone.

1.1 BONE STRUCTURE

The adult skeleton is composed of 206 skeletal elements, which can be subdivided into five groups based on their shape: long bones (e.g. tibia), short bones (e.g. phalanges), flat bones (e.g. scapula), irregular bones (e.g. vertebrae) and sesamoid bones (e.g. patella). Furthermore, two types of bone tissue have been characterized: the cortical bone and the cancellous bone. In long bones, the cortical bone forms the outer layer whereas the cancellous bone occupies the inner part that is in contact with the bone marrow (Fig. 1).

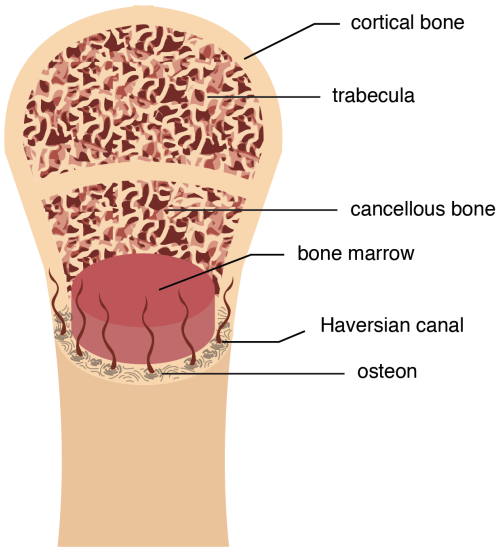


Figure 1. Schematic illustration of the inner structures of a long bone.

Osteons are the structural and functional units of the cortical bone. They consist of concentric bone layers (lamellae) surrounding a canal, known as Haversian canal, that supplies blood. In contrast, the cancellous bone is comprised of trabeculae that form a porous network (Fig. 1).

1.1.1 Bone cells and extracellular matrix

Bone is composed of cells and extracellular matrix (ECM) [Florencio-Silva et al., 2015]. Two types of cells, namely osteoblasts and osteocytes, derive from mesenchymal stem cells (MSCs) (Fig. 2A). Osteocytes are terminally-differentiated osteoblasts. Osteoclasts, which are embedded in the ECM, differentiate instead from hematopoietic stem cells (HSCs) (Fig. 2A). Osteocytes are the largest cell population (90-95%) in the adult skeleton, followed by osteoblasts (4-6%) and osteoclasts (approximately 1-2%). Each bone cell type expresses specific markers and captures stimuli sent by other cells or external factors to promptly regulate bone homeostasis.

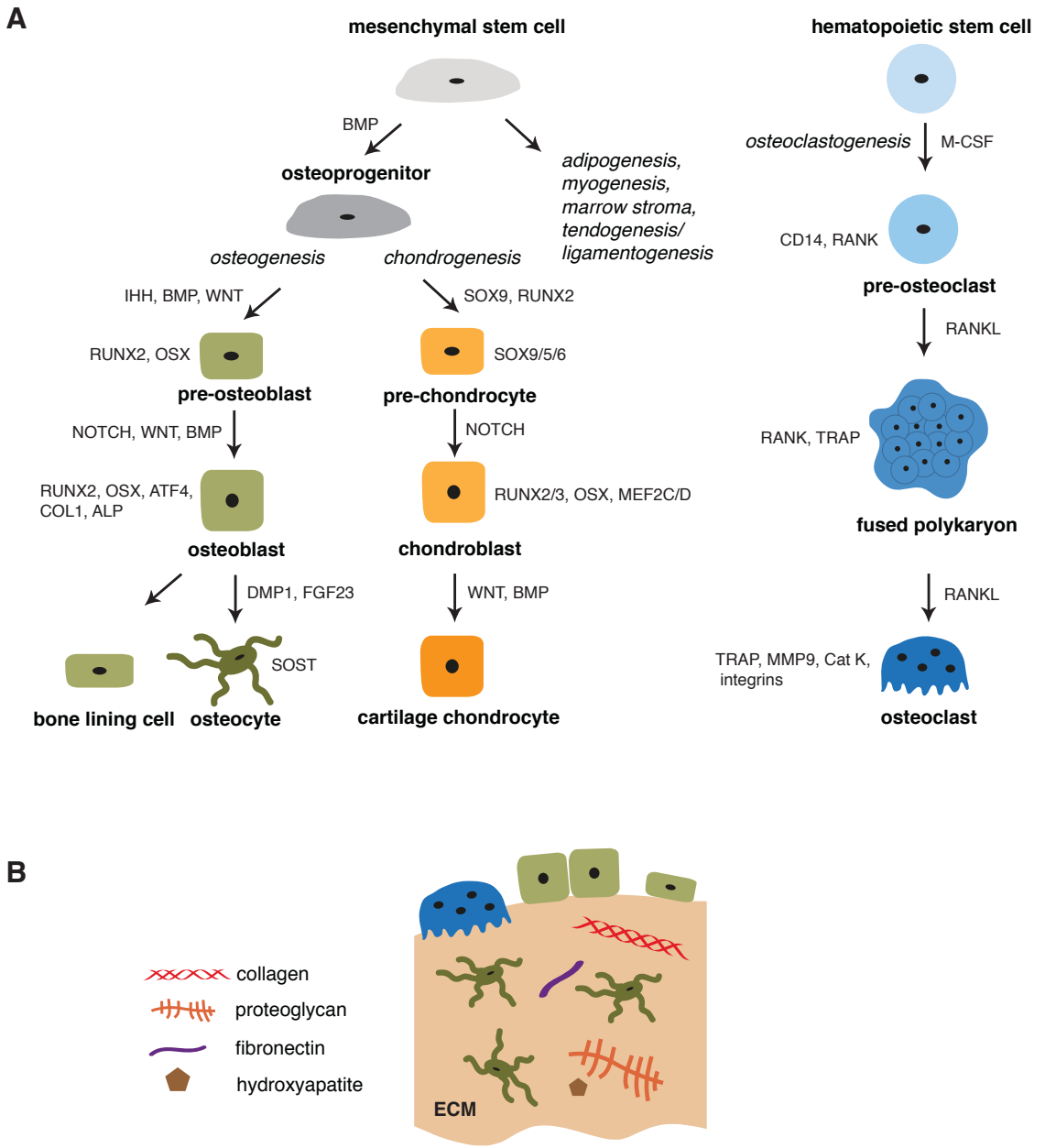


Figure 2. Differentiation of bone and cartilage cells from their stem cell precursors (A). Schematic representation of the bone cells and ECM (B).

Osteoblasts are the cells responsible for building up new bone. Since they are specialized in secreting a large amount of ECM, osteoblasts are characterized by an extensive Golgi complex, a large number of mitochondria as well as a dilated rough endoplasmic reticulum (rER) [Del Fattore et al., 2012]. In addition, they produce several factors to regulate cell-to-cell interactions in particular with the osteoclasts. Once osteoblasts reach maturity, they become bone lining cells (Fig. 2B).

Osteoclasts are large multinucleated macrophage-like cells located on the bone surface (Fig. 2B) and their role is to resorb bone. Cell membrane polarization together with extensive ion systems are required to dissolve the mineralized matrix [Vaananen et al., 2000]. Three different zones characterize the membrane of the osteoclasts: a sealing zone required to adhere to bone matrix, a ruffled border specialized to resorb mineralized matrix, and a functional secretory domain to exocytose the degraded matrix.

Osteocytes are stellar-shaped cells that reside within the bone matrix (Fig. 2B) [Dallas et al., 2013]. They communicate with each other as well as with osteoblasts and the bone marrow through their cytoplasmic extensions, named canaliculi.

Finally, the ECM is synthesized and secreted by the bone cells, primarily by the osteoblasts, and mainly composed of type I collagen (90%), other non-collagenous proteins (especially glycoproteins and proteoglycans), water, lipids and minerals (mostly hydroxyapatite) (Fig. 2B) [Young, 2003]. The ECM enables both adhesion and movement of the cells and it is also a source of growth factors and cytokines required for cell differentiation and signaling [Rozario and DeSimone, 2010]. Type I collagen is a protein comprised of three polypeptide chains, two alpha-1 and one alpha-2 chains, which are tightly packed together through hydrogen bonding. A recurrent sequence motif of three amino acids Xaa-Yaa-Gly, where every third position is occupied by glycine (Gly) and Xaa/Yaa can be any amino acid, characterizes the helical region of collagen. Moreover, the activity of several other proteins is required for the post-translational processing and folding of type I collagen as well as for fibril assembly. The assembly of a variety of other proteins besides collagen, such as fibronectin, determines the structure and organization of the ECM.

1.1.2 Bone strength

Bone strength, defined as the resistance of bone to fractures, is determined by the amount of, and ratio between, cortical and cancellous bone, and by the number, thickness and organization of the trabeculae within the cancellous component. Bone mineral density (BMD), corresponding to the mass per unit volume of mineralized bone, is measured by dual-energy X-ray absorptiometry (DXA) and it is commonly used as a measure for bone

strength. Twin studies have demonstrated that up to 85% of BMD variance can be explained by genetic variants [Pocock et al., 1987; Stewart and Ralston, 2000].

Bone strength is also impacted by bone quality, which is influenced by several factors such as the microarchitecture of cancellous bone, the shape of bones, and the mineralization and the molecular composition of the ECM.

1.2 SKELETAL DEVELOPMENT

Skeletal patterning begins early during embryonic development and determines the precise location, shape and function of each skeletal element in the body [Olsen et al., 2000]. Aggregation of the MSCs into mesenchymal condensations is the first step in bone development [Hall and Miyake, 2000]. The skeletal elements are formed via two different processes: intramembranous ossification and endochondral ossification.

1.2.1 Intramembranous ossification

During intramembranous ossification, bone develops without the mediation of a cartilaginous phase [Opperman, 2000]. This process initiates when a group of MSCs aggregate and give rise to specialized cells. Some of these cells are responsible for tissue vascularization while others differentiate into osteoprogenitors and subsequently to osteoblasts, which locate in the ossification centers and produce unmineralized matrix (osteoid). Mineralization of osteoid causes trapping of osteoblasts in the matrix, leading to transformation of osteoblasts into osteocytes [Ornitz and Marie, 2002]. The deposition of osteoid around the capillaries constitutes the trabecular matrix whereas the osteoblasts on the surface form the periosteum, which is a layer of compact bone that protects the cancellous bone and the bone marrow. The craniofacial bones and clavicles are formed via intramembranous ossification whereas development of all other skeletal elements requires cartilage formation and are formed through endochondral ossification. [Long and Ornitz, 2013].

1.2.2 Endochondral ossification

Endochondral ossification is mediated by the epiphyseal plate, commonly known as the growth plate, which is a highly organized cartilaginous structure that allows longitudinal bones to elongate during childhood and adolescence. The growth plate can be subdivided into three zones occupied by chondrocytes with different functions and proliferation capacities: 1) resting zone, 2) proliferative zone and 3) hypertrophic zone (Fig. 3) [Brighton, 1978].

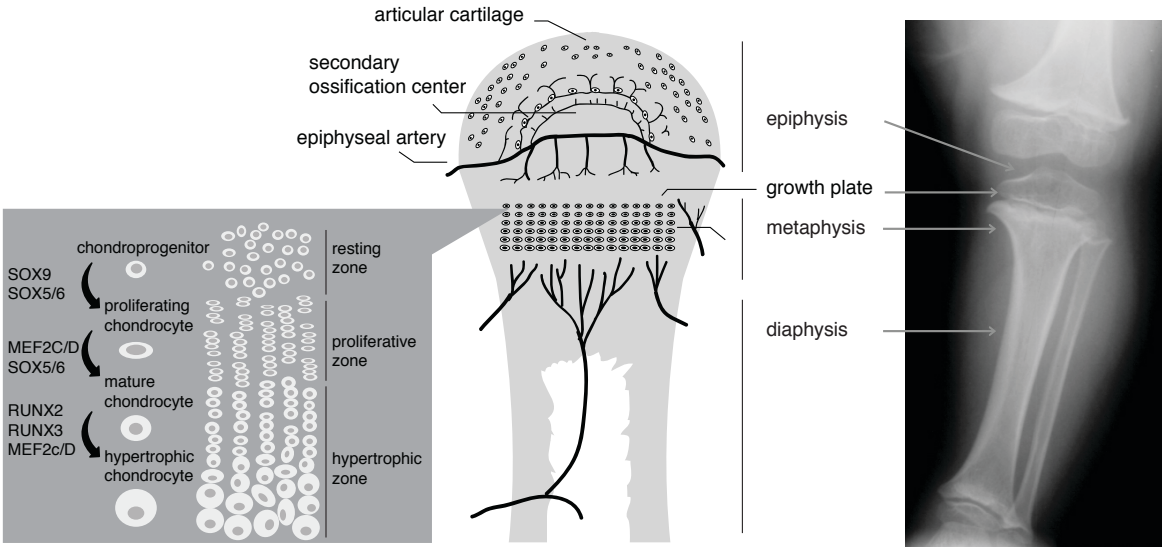


Figure 3. Structure of a long bone and details of the growth plate. On the left, different zones of the growth plate are represented. Chondrocytes in each zone have different properties. In the middle, a schematic picture of a tubular bone indicates the location of the growth plate and other bone structures. On the right, the X-ray shows the distal femur and the tibia and fibula of a child. The cartilaginous growth plates are open and not mineralized.

The function of the chondrocytes as well as their size and orientation differ in each zone and they are tightly regulated by several signaling pathways, including NOTCH, IHH, FGF, BMP, PTHrP, and WNT [Long and Ornitz, 2013]. Two other crucial structures are formed during cartilage growth: the primary ossification centers (POCs) and the secondary ossification centers (SOCs) [Olsen et al., 2000]. POCs are formed during prenatal development when hypertrophic chondrocytes are replaced by perichondrial osteoblasts that produce osteoid. At the same time, the cartilage matrix is degraded and this region becomes vascularized [Dao et al., 2012]. SOCs instead develop postnatally within the epiphysis and are also invaded by blood vessels (Fig. 3). At the end of puberty, the growth plate fuses with the epiphysis and the growth plate cartilage is replaced by bone. This culminates in the cessation of linear bone growth.

1.3 BONE HOMEOSTASIS

Bone is a dynamic organ in which old bone is cyclically resorbed by the osteoclasts and new bone is produced by the osteoblasts throughout life [Raisz, 1999]. A balance between bone formation and resorption is necessary to regulate mineral homeostasis and maintain skeletal integrity. To provide the proper amount of minerals to bone and, vice versa, to transfer minerals from bone to circulation and other organs, our skeleton is regulated by a complex endocrine network. In addition to hormones produced by endocrine glands, the bone itself produces hormones to regulate mineral homeostasis.

1.3.1 Bone modeling and remodeling

Osteoblasts are responsible for producing new bone while osteoclasts break down old mineralized bone. Osteocytes respond to mechanosensory signals and mediate the activity and communication between osteoblasts and osteoclasts by secreting soluble factors (Fig. 4) [Bonewald, 2011]. Bone formation and resorption are tightly coupled in order to maintain bone health, via two mechanisms known as bone modeling and remodeling. Bone modeling occurs during skeletal growth and allows the skeleton to adapt to loading by changing the size and shape of bones. Bone remodeling instead ensures mechanical integrity of the bone tissue. These mechanisms are necessary not only to respond to bone loading but also for fracture resistance, fracture healing and calcium and phosphate homeostasis [Florencio-Silva et al., 2015]. In this thesis, we refer only to bone remodeling as a general term to encompass bone formation and resorption during both skeletal growth and maturity.

Several signaling pathways and soluble factors secreted by other organs separate from the bone play a role in bone metabolism. The WNT signaling pathway plays a pivotal role in skeletal development and in bone remodeling, in particular by inducing bone formation [Clevers, 2006; Baron and Kneissel, 2013]. This pathway enhances the differentiation of MSCs into osteoblasts while it inhibits adipogenesis and chondrogenesis (Fig. 2A) [Day et al., 2005; Kennell and MacDougald, 2005]. Although the non-canonical pathway has an emerging role in bone homeostasis, the canonical pathway, also termed Wnt-β-catenin pathway, is the best understood [Liu et al., 2007]. This pathway is activated when a G-protein-coupled receptor protein, named Frizzled (FZD), binds to its co-receptor, either the low-density lipoprotein receptor related protein 5 or 6 (LRP5 and LRP6, respectively), to inactivate the cytosolic β-catenin “destruction complex”. This complex is composed of three proteins involved in the phosphorylation of β-catenin (glycogen synthase kinase 3, axin, and casein kinase 1) and by adenomatous polyposis coli [Baron and Kneissel, 2013]. Consequently, the degradation of the mediator β-catenin is prevented allowing for its translocation to the nucleus where it stimulates transcription of target genes. Sclerostin (SOST) and Dickkopf-related protein 1 (DKK1) are antagonists of the WNT pathway (Fig. 4). Mutations in some of the key participants of this pathway, such as *LRP5* and *SOST*, lead to diseases characterized by abnormal BMD.

The OPG-RANK-RANKL pathway is also required to regulate bone formation and resorption (Fig. 4) [Khosla, 2001]. Osteocytes secrete receptor activator of nuclear factor kappa-B ligand (RANKL) that binds to its receptor RANK on the osteoclasts to stimulate bone resorption. On the other hand, osteoblasts produce osteoprotegerin (OPG), which is an antagonist of this pathway and impedes RANK-RANKL interaction by binding RANKL.

In the recently identified RANKL reverse signaling, mature osteoclasts secrete vesicular RANK, which binds to RANKL on the surface of osteoblasts [Ikebuchi et al., 2018]. This close interaction and communication between the osteoblastic and osteoclastic lineages ensures balance between the opposing processes of bone resorption and bone formation.

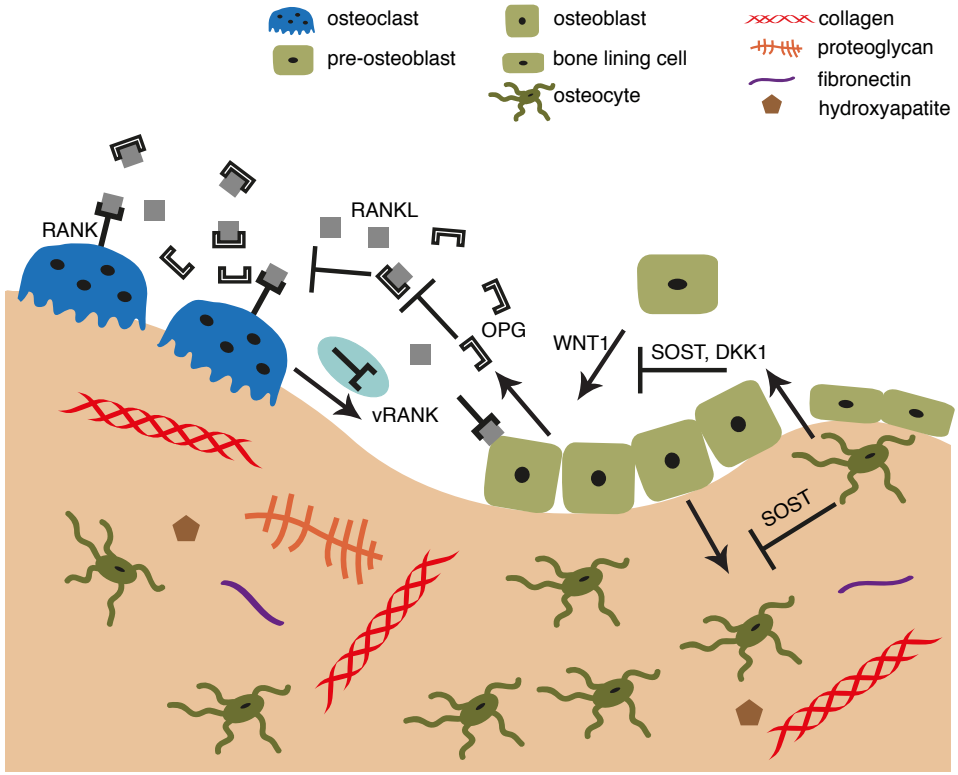


Figure 4. Simplified overview of bone homeostasis with main focus on the OPG-RANK-RANKL and the WNT signaling pathways. Osteoblasts, osteoclasts and osteocyte cooperate to regulate bone formation and resorption. *RANK= receptor activator of nuclear factor kappa-B; RANKL= receptor activator of nuclear factor kappa-B ligand; vRANK= vesicular RANK; OPG= osteoprotegerin; SOST= sclerostin; DKK1= Dickkopf-related protein 1; WNT1= Wnt Family Member 1.*

1.3.2 Calcium and phosphate homeostasis

Parathyroid hormone (PTH), vitamin D and fibroblast growth factor 23 (FGF23) are the major regulators of circulating calcium and phosphate levels; they influence the handling of these minerals in target organs, including bone (Fig. 5).

Reduced levels of ionized calcium in the circulation lead to a rapid increase in PTH secretion by the parathyroid glands (Fig. 5) [Bilezikian, 2019]. Consequently, high levels of PTH stimulate bone resorption and thus promote the release of calcium and phosphate from the hydroxyapatite crystals in bone. The increased PTH also promotes phosphaturia in the kidney, ensuring normal circulating phosphate concentration despite increased release from the bone.

Vitamin D, which is produced by the skin with UV radiation from sunlight and absorbed in the gut from diet, is further hydroxylated in the liver and the kidneys (Fig. 5). Once the

active form of vitamin D, 1,25-dihydroxyvitamin D ($1,25(OH)_2$) is produced by the kidneys, it stimulates calcium and phosphate absorption in the intestine and their release from bone, the net effect being an increase in circulating calcium and phosphate (Fig. 5) [Fraser and Kodicek, 1973]. Vitamin D deficiency is an acquired nutritional disease that leads to rickets, a condition characterized by impaired bone mineralization. Several genetic defects that impair vitamin D metabolism, like vitamin D-dependent rickets type I (MIM #264700) caused by mutations in the vitamin D 1α-hydroxylase gene, have also been identified.

FGF23, which is mainly produced by osteocytes, acts as a hormone to regulate phosphate homeostasis by modifying renal phosphate reabsorption and $1,25(OH)_2$ production (Fig. 5) [Shimada et al., 2004]. Several genetic conditions are linked to *FGF23* defects and can lead to either high or low levels of phosphate [Liu and Quarles, 2007]. Hypophosphatemia can lead to rickets (MIM #193100) and osteomalacia whereas hyperphosphatemia due to low FGF23 underlies tumoral calcinosis (MIM #617993).

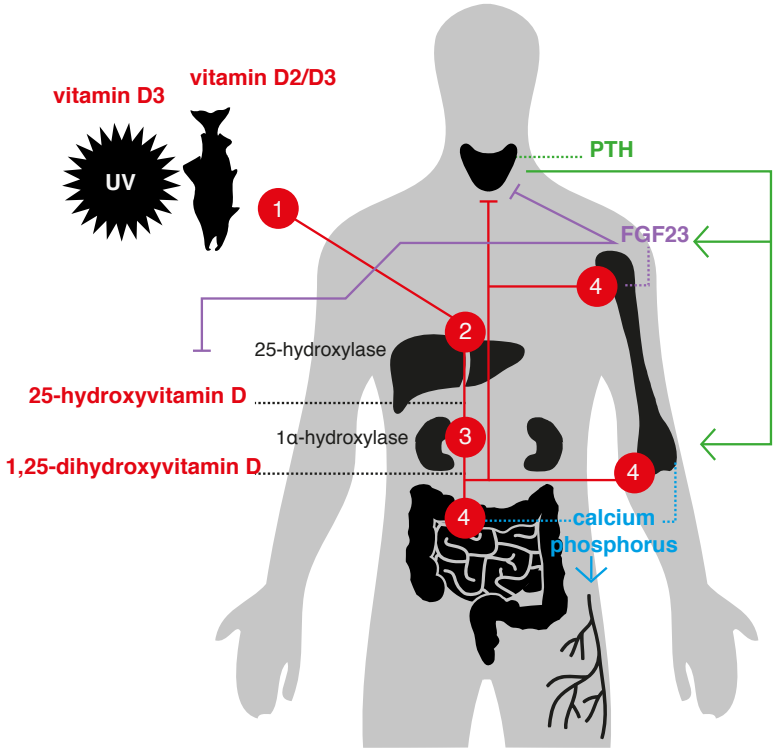


Figure 5. Role of vitamin D, PTH and FGF23 in regulating mineral homeostasis. PTH stimulates bone resorption when there is lack of calcium and phosphate in circulation. Active vitamin D is metabolized in the liver and the kidneys and the final form, 1,25-dihydroxvitamin D, leads to calcium and phosphate absorption from the intestine and release from the bone while inhibiting PTH. High levels of FGF23 inhibit both PTH and vitamin D. Dashed lines represent the organ from which each hormone and the vitamin D metabolite is released.

1.3.3 Other regulators

Several other hormones are important for controlling bone formation and resorption. Osteocalcin (OCN) is a protein produced by the osteoblasts and it participates in the

regulation of glucose metabolism [Lee et al., 2007; Clemens and Karsenty, 2011]. In mice OCN stimulates the pancreas to produce and secrete insulin, the adipocytes to produce adiponectin and the muscles to use glucose [Fukumoto and Martin, 2009]. Findings in human studies have however been inconclusive.

Calcitonin, which is secreted by the thyroid gland, lowers the calcium and phosphate levels in the blood by inhibiting bone resorption [Brown, 2007].

Finally, sex steroids play a major role in skeletal homeostasis beginning in early skeletal development. Estrogens and androgens are not only important for shaping bones differently in females and males but they also regulate skeletal growth. They reduce bone loss by maintaining a delicate balance between bone formation and resorption [van der Eerden et al., 2003; Bilezikian, 2019].

1.4 THE HUMAN GENOME

The human haploid genome is a sequence of 3.2 billion base pairs that are stored in 23 chromosomes (22 autosomes and one sex chromosome). Under physiological conditions, only our germ cells contain a single set of chromosomes whereas all other cells are diploid and host 23 pairs of chromosomes. Genes are segments of DNA composed of exons and introns that provide the template for producing RNA and polypeptides or proteins. Only exons, via a mechanism named RNA splicing, are used as a template to make a complementary RNA sequence (transcription), either a messenger RNA (mRNA) that will be translated into a protein or a mature non-coding RNA. Roughly 1% of the total DNA sequence is occupied by protein-coding genes, which are approximately 20,500 in total. Each gene might have different forms, known as alleles, but a normal individual has only two alleles for each gene, one inherited from the mother and the other inherited from the father. The more closely related the parents are, the more genomic sequence they share. A specific chromosomal location that defines the location of either an individual gene or a DNA sequence is named locus. A set of the alleles present at one or several loci is referred as the genotype while the phenotype comprises the observable characteristics (also known as traits) of a certain individual.

1.4.1 Human genetic variation

In the last quarter of the past century, the invention of recombinant DNA cloning and sequencing technologies enhanced the decryption of the genetic code of different species.

In 2001, the first draft of the human genome was sequenced and in 2003 it was fully completed by the Human Genome Project [Lander et al., 2001]. Soon after, the shift from

Sanger sequencing to massively-parallel sequencing (MPS) led to sequencing of thousands of human exomes and genomes and to discovery of a large inter-individual genetic variability. In fact, it is now known that a random human genome differs from the reference sequence at 4.1 million to 5 million sites [Genomes Project et al., 2015]. However, the majority of these variants are neutral, meaning that they do not have a notable impact on the phenotype.

In 2012, when the 1000 Genome Project was completed, a map of genetic variation from 1,092 human genomes from 14 different populations facilitated the distinction between common variants, shared among several individuals, and rare variants present in one or a few subjects [Genomes Project et al., 2010; Genomes Project et al., 2012]. Later, other genomic databases like the Exome Aggregation Consortium (ExAc) including over 60,000 exomes and subsequently expanded to the Genome Aggregation Consortium (gnomAD) containing 125,748 exome sequences and 15,708 whole-genome sequences improved genotype-phenotype associations [Lek et al., 2016].

Frequently, common variants at several loci determine one quantitative phenotype, such as vitamin D levels, which is then defined as a polygenic trait [Jiang et al., 2018]. On the other hand, some other traits are monogenic (determined by variation at a single locus) and can be inherited in a family according to four main inheritance patterns: autosomal dominant (attributed to a change in one copy of a gene on an autosomal chromosome), autosomal recessive (attributed to a change in both copies of a gene on an autosomal chromosome), X-linked dominant (attributed to a change in one copy of a gene on a sex chromosome), and X-linked recessive (attributed to a change in two copies of a gene on a sex chromosome). If an individual has two copies of the same allele the subject is homozygous for that trait and otherwise he or she is heterozygous.

Sometimes a subject might harbor a variant that is not present in either of the parents (*de novo* change). New variants arise in a germ cell of a parent or in the fertilized egg during early development.

1.4.2 Types of genetic variants

Genetic variants can be classified as small-scale variants if they affect one or a small number of nucleotides (< 50 base pairs, bp) and large-scale variants, or structural variants (SVs), if they involve over 50 bp [Tattini et al., 2015]. The small-scale variants include single base substitutions, also named single nucleotide variants (SNVs), and small insertions and deletions (indels). According to the mutation nomenclature proposed by the Human Genome Variation Society, SNVs are further subdivided to six groups, based on their effect at the peptide level: 1) synonymous variant – not altering the encoded amino

acid (silent change), 2) missense variant – substituting a certain amino acid into another amino acid, 3) start codon variant - altering the translation initiation codon (methionine 1), 4) nonsense variant - introducing an immediate stop codon, 5) no-stop change variant - removing the termination codon and generating an extension of the protein at the C-terminus, and 6) splicing variant – leading to abnormal splicing.

Concerning indels, they include all types of small changes that lead to a size change at a specific locus: duplications, deletions, insertions or combined insertions and deletions. Indels are classified as in-frame if they do not introduce a shift within the open reading frame (ORF; the coding sequence of triplets between the start codon and the stop codon) and as out-of-frame (or frameshift) when they shift the ORF and alter the C-terminal end of the protein by often introducing an early stop codon.

SVs can be either balanced changes if they do not lead to any gain or loss of genetic material, or unbalanced, if they affect the gene dosage. Copy number variants (CNVs) refer to large deletions or duplications (> 50 bp) that either decrease or increase the DNA content [Zarrei et al., 2015]. Translocations are rearrangements leading to an exchange in genetic material between chromosomes. They can be both balanced or unbalanced. Inversions are balanced changes in which the new sequence is the reverse-complement of the original sequence.

Genetic variants (or mutations) can be classified as loss-of-function (LOF) if the protein loses its activity and gain-of-function (GOF) if the protein acquires a new or enhanced activity. A genetic variant that is disease-causing is defined as pathogenic.

1.4.3 Rare monogenic diseases

To diagnose a rare monogenic disease, it is necessary to identify a single pathogenic variant (or a compound heterozygous variant) within millions of variants present in the genome.

The minor allele frequency (MAF) refers to frequency in which the less common allele occurs in a given population; this parameter needs to be considered when investigating a genetic condition. In fact, it is important to make a distinction between genetic variants that are relatively common among the general population (MAF > 1%), known as single nucleotide polymorphisms (SNPs), and rare variants shared by single or a few families. Theoretically, a genetic variant causing a rare monogenic disease has a MAF below 1%. However, most commonly, such a variant has either a MAF below 0.1% or it is completely absent from large genomic databases [Lek et al., 2016].

Although the genetic defects explaining a large number of monogenic diseases has already been identified, new mutations are continuously discovered. During recent years,

the American College of Medical Genetics and Genomics (ACMG) has made efforts to standardize the classification of pathogenicity of genetic variants. According to the ACMG guidelines, variants can be classified based on different parameters, including MAF, effect of the variant, *in silico* prediction scores (for missense and splicing variants) as well as co-segregation of the variant with the disease in families [Kearney et al., 2011; Richards et al., 2015; Jarvik and Browning, 2016].

To assess and/or validate the pathogenicity of a certain DNA variant it is often necessary to investigate its effect on the mRNA or protein structure and function. Furthermore, if a genetic defect is identified for the first time, functional studies are needed to understand the molecular mechanisms leading to disease.

1.4.4 Challenges in variant interpretation

Understanding the link between a particular variant and a phenotypic trait can sometimes be complicated by the presence of genetic mechanisms that render variant interpretation challenging.

In dominant conditions, haploinsufficiency arises when the expression of one wild-type copy of a gene is not sufficient to guarantee a normal phenotype. In some other cases, a defective copy of a gene interferes with the wild-type allele (dominant-negative effect) and confers a different function to the protein.

Non-penetrance can also be found in dominant conditions. This phenomenon occurs when an individual carries a pathogenic mutation without showing any sign of abnormal phenotype. Sometimes, variable expressivity leads to phenotypes marked by different levels of severity in patients with the same dominant mutation.

Both non-penetrance and variable expressivity might be due to the effect of other genes (genetic modifiers and epigenetic changes), to environmental factors or determined by pure chance.

1.5 RARE SKELETAL DISEASES

As previously described, bone formation begins early during embryonic development. Mutations in genes playing pivotal roles in bone development and homeostasis can reduce the capacity of bone to resist fractures or they may interfere with normal bone growth [Viguet-Carrin et al., 2006].

Rare skeletal diseases – skeletal dysplasias - comprise a group of over 400 conditions affecting the skeleton. These diseases are characterized by broad clinical and genetic heterogeneity [Bonafe et al., 2015]. As conventional radiographs are commonly used for

diagnosis, the nosology of genetic skeletal diseases has classified these conditions into 42 groups based on the most relevant criteria: clinical features, molecular mechanisms and radiological findings [Panda et al., 2014; Bonafe et al., 2015]. The severity of these disease ranges from perinatal and neonatal lethality to mild impairments, such as moderate growth delay [Kornak and Mundlos, 2003]. Although each single skeletal disease is rare and some conditions are significantly less frequent than others, the overall prevalence of these conditions is approximately 5 in 10,000 births [Panda et al., 2014]. Until now, mutations in approximately 350 genes have been identified as the underlying causes of these diseases [Bonafe et al., 2015].

1.5.1 Diseases affecting growth

Short stature is a common feature for several skeletal diseases. As previously mentioned, the growth plate is the site where skeletal growth takes place and the chondrocyte function, proliferation and differentiation within the growth plate is regulated by several factors and hormones. Additionally, 90% of the volume of the cartilage is occupied by the ECM, which is rich in type II collagen, proteoglycans and glycoproteins. Rare genetic mutations in genes that play a pivotal role in chondrogenesis are likely to be a frequent cause of disproportionate short stature, which is a hallmark of several forms of skeletal dysplasias. Some rare forms of skeletal dysplasia are caused by defects in the primary cilium, a structure that is required for cell mechanosensing. These diseases are collectively named ciliopathies.

In this work, two particular types o, named spondylometaphyseal dysplasia (SMD) and spondyloepimetaphyseal dysplasia (SEMD), will be described.

Spondylometaphyseal dysplasia (SMD)

SMD affects mainly the spine (spondylo) and the metaphyses (metaphyseal) of tubular bones. The patients exhibit severe growth retardation, flat vertebrae (platyspondyly) and abnormal shape and maturation of metaphyses in different locations. Furthermore, a range of extra skeletal manifestations, including ocular impairment, respiratory problems and immune defects, can also be present in SMD (Table 1). To date, at least 9 different subtypes of SMD have been clinically characterized and a specific molecular defect has been identified in each of them (Table 1).

Table 1. Genetic defects causing SMD.

Type of SMD	MIM #°	Gene	Protein	I.P	Main clinical features
SMD with immune dysregulation	607944	*ACP5*	Tartrate-resistant acid phosphatase (TRAP)	AR	Impairment of the immune system
Odontochondrodysplasia (ODCD)	184260	*TRIP11*	Thyroid Hormone Receptor Interactor 11	AR	Joint laxity and dentinogenesis imperfecta
SMD with cone-rod dystrophy	608940	*PCYT1A*	Phosphate cytidylyltransferase 1	AR	Early-onset progressive visual impairment associated with a pigmentary maculopathy and cone-rod dysfunction
SMD with retinal degeneration, axial type	602271	Unknown	Unknown	AR	Impaired visual acuity, retinitis pigmentosa or pigmentary retinal degeneration
SMD Sedaghatian type	250220	*GPX4*	Glutathione peroxidase 4	AR	Severe hypotonia and cardiorespiratory problems; cardiac problems; half of the cases have central nervous system malformations
SMD Kozlowski type	184252	*TRPV4*	Transient receptor potential cation channel, subfamily V, member 4	AD	Narrow thorax, prominent joints and occasionally tail-like coccygeal appendage (caudal tail)
SMD axial	602271	*CFAP410*	Cilia And Flagella Associated Protein 410	AR	Impaired visual acuity and retinal impairment; mild to moderate respiratory problems in the neonatal period and later susceptibility to airway infection
SMD Megarbane-Dagher-Melike type	613320	*PAM16*	Presequence Translocase Associated Motor 16	AR	Any peculiar features; only two families have been described
SMD with corner fractures, SMD-CF*	184255	*FN1*	Fibronectin 1	AD	Irregular metaphyses with “corner fracture” appearance

** Previously known as SMD Sutcliffe type; °= phenotype MIM number; I.P= inheritance pattern; AD= autosomal dominant; AR= autosomal recessive.*

Spondyloepimetaphyseal dysplasia (SEMD)

SEMD is a subgroup of skeletal dysplasias of which hallmarks are severe short stature and skeletal impairments affecting the spine, metaphyses and epiphyses. Sometimes only a careful radiological investigation can clearly distinguish SEMD from SMD. Until today, over 20 different subtypes of SEMD have been identified (Table 2). Pseudoachondroplasia, caused by mutations in *COMP*, is one of the most common forms of SEMD. COMP is a cartilage ECM protein which, when mutated, is retained within the ER, thus compromising

chondrocyte function and increasing cell death [Acharya et al., 2014]. Recently, mutations in the gene (*TONSL*), encoding a protein involved in DNA repair, have been identified as the underlying cause of Sponastrime dysplasia [Chang et al., 2019].

Table 2. Genetic defects underlying SEMD.

Type of SEMD	MIM #°	Gene	Protein	I.P
Dyggve–Melchior–Clausen dysplasia (DMC)	223800 615222	DYM RAB33B	Dymeclin RAS-associated protein rab33b	AR
Smith–McCort dysplasia	607326	DYM	Dymeclin	AR
Immuno-osseous dysplasia (Schimke)	242900	SMARCAL 1	SWI/SNF-related regulator of chromatin subfamily A-like protein 1	AR
SED, Wolcott–Rallison type	226980	EIF2AK3	Translation initiation factor 2-alpha kinase-3	AR
SEMD, Matrilin type	608728	*MATN3*	Matrilin 3	AR
SEMD, short limb–abnormal calcification type	271665	*DDR2*	Discoidin domain receptor family, member 2	AR
SED tarda, X-linked (SED-XL)	313400	*SEDL*	Sedlin	XLR
Spondylodysplastic Ehlers–Danlos syndrome	612350	*SLC39A13*	Zinc transporter ZIP13	AR
Sponastrime dysplasia	271510	*TONSL*	Tonsoku Like, DNA Repair Protein	AR
Platyspondyly (brachyolmia) with amelogenesis imperfecta	601216	*unknown*	*unknown*	AR
CODAS syndrome	600373	*LONP1*	LON peptidase 1	AR
Opsismodysplasia	258480	*INPPL1*	Inositol polyphosphate phosphatase-like 1	AR
SEMD, Maroteaux type	184095	*TRPV4*	Transient receptor potential cation channel, subfamily V, member 4	AD
SEMD with joint laxity, type 2	603546	*KIF22*	Kinesin Family Member 22	AD
SEMD with joint laxity, type 1, with or without fractures	271640	*B3GALT6*	Beta-1,3-Galactosyltransferase 6	AR
SEMD Shohat type	602557	*DDRGK1*	DDRGK Domain Containing 1	AR
SEMD Faden-Alkuraya type	616723	*RSPRY1*	Ring Finger And SPRY Domain Containing 1	AR
SEMD Missouri type	602111	*MMP13*	Matrix Metallopeptidase 13	AD
SEMD Strudwick type	184250	*COL2A1*	Collagen type II alpha 1 chain	AD
SEMD X-linked	300106	*BGN*	Biglycan	XLR
SEMD Pakistani type	612847	*PAPSS2*	3'-phosphoadenosine 5'-phosphosulfate synthase 2	AR
SEMD Camera-Genevieve type	610442	*NANS*	N-acetylneuraminate synthase	AR

°= phenotype MIM number; I.P= inheritance pattern; AD= autosomal dominant; AR= autosomal recessive; XLR= X-linked recessive.

Ciliopathies with major skeletal involvement

Ciliopathies are rare recessive conditions that arise from cilia dysfunction. Cilia are small organelles consisting of microtubule filaments that in a few specialized cells enable movement. However, a single primary cilium can be found in the majority of the vertebrate cells and instead of being involved in cell's motility, it functions as a sensor of the cellular environment [Reiter and Leroux, 2017]. Altogether approximately 180 genes have already been linked to ciliopathies. However, mutations in another 240 genes playing a role in

ciliary structure and/or function might potentially also lead to ciliopathies [Anvarian et al., 2019].

A myriad of heterogeneous diseases affecting different tissues and organs arise from cilia impairments [Reiter and Leroux, 2017]. Some conditions are organ-specific (e.g. polycystic kidney disease) whereas some others like Bardet–Biedl syndrome affect multiple organs. A specific class of skeletal dysplasias are caused by ciliary defects [Huber and Cormier-Daire, 2012]. So far, approximately ten conditions with major skeletal impairment have been reported [Bonafe et al., 2015]. As one, the short rib–polydactyly syndrome types 1-3 are characterized by narrow chest, short ribs, short limbs and trident aspect of the acetabular roof [Huber and Cormier-Daire, 2012].

The molecular mechanisms linking mutations in cilia genes to skeletal impairments have yet to be fully elucidated. It is known that defective primary cilia affect bone growth due to impaired Hedgehog signaling pathways in some cases [Huber and Cormier-Daire, 2012]. For instance, the two genes encoding the EvC ciliary complex subunit 1-2, *EVC1* and *EVC2*, are mutated in chondroectodermal dysplasia (Ellis-van Creveld) [Ruiz-Perez et al., 2000]. The EVC protein, which is localized on the basal part of the primary cilium of chondrocytes, is only expressed during skeletal development in mice [Ruiz-Perez et al., 2007; Goetz and Anderson, 2010]. Furthermore, *Evc* knockout mice show reduced IHH only in the skeletal structures, thus demonstrating that this protein is required for normal transcriptional activation of IHH target genes specifically in chondrocytes.

1.5.2 Diseases affecting bone homeostasis

Several skeletal conditions arise from mutations in genes that encode proteins that are involved in bone remodeling. An impaired differentiation and/or function of osteoblasts, osteoclasts or osteocytes can lead to an imbalance between bone formation and resorption and cause an insufficient or an excessive accumulation of bone in the skeleton.

Osteoporosis developing later in adult life and especially post-menopausal osteoporosis are commonly due to hormonal imbalance and/or multiple deleterious genetic variants with small effect size. In contrast, primary osteoporosis in the young population is often caused by a single genetic variant with a large effect size [Makitie, 2013; Kampe et al., 2015; Costantini and Makitie, 2016]. In children a BMD value below -2.0 SD (Z-score) associated with increased fractures is indicative of osteoporosis [Makitie, 2013].

Osteogenesis imperfecta

Osteogenesis imperfecta (OI), also known as brittle bone disease, is the most common form of early-onset skeletal fragility. OI is a congenital disease with broad phenotypic

variability. Milder forms of the disease may show only low BMD and increased susceptibility to fractures whereas the severest forms are prenatally or postnatally lethal [Forlino and Marini, 2016; Marini et al., 2017]. Extra-skeletal impairments, including dentinogenesis imperfecta, blue sclerae and impaired hearing, might also be present in OI patients. Mutations in 20 different genes have thus far been linked to OI (Table 3) [Bonafe et al., 2015; Marini et al., 2017; Doyard et al., 2018; Makitie et al., 2019; Pekkinen et al., 2019].

Table 3. Genetic defects underlying OI and related bone fragility conditions.

MIM #°	Gene	Protein	Inheritance	Pathomolecular mechanism
166200; 166210; 259420; 166220	*COL1A1*	Collagen alpha-1(I) chain	AD	Defects in collagen type I synthesis, structure, folding, post-translational modification, processing and cross-linking
259420; 166210; 166220	*COL1A2*	Collagen alpha-2(I) chain	AD; AR*	
610682	*CRTAP*	Cartilage-associated protein	AR	
259440	*PPIB*	Peptidyl-prolyl cis-trans isomerase B; cyclophilin B	AR	
610915	*P3H1*	Prolyl 3-hydroxylase 1	AR	
610968	*FKBP10*	Peptidyl-prolyl cis-trans isomerase FKBP10	AR	
609220	*PLOD2*	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2	AR	
613848	*SERPINH1*	Serpin H1	AR	
614856	*BMP1*	Bone morphogenetic protein 1	AR	
616507	*SPARC*	SPARC; osteonectin	AR	Defects in other proteins leading to abnormal bone mineralization
613982	*SERPINF1*	Pigment epithelium-derived factor (PEDF)	AR	
610967	*IFITM5*	Interferon induced transmembrane protein 5	AD	
300910	*PLS3*	Plastin 3	XLD	
NA	*SGMS2*	Sphingomyelin Synthase 2	AD	
615066	*TMEM38B*	Trimeric intracellular cation channel type B	AR	Defects in osteoblast differentiation and function
615220	*WNT1*	Proto-oncogene Wnt-1	AR	
613849	*SP7*	Transcription factor Sp7; osterix	AR	
616229	*CREB3L1*	Cyclic AMP-responsive element-binding protein 3-like protein 1	AR	
301014	*MBTPS2*	Membrane-bound transcription factor site-2 protease	XLR	
617952	*TENT5A* (also known as *FAM46A*)	Terminal nucleotidyltransferase 5A	AR	Unknown

*°= phenotype MIM number; AD= autosomal dominant; AR= autosomal recessive; XLD/XLR= X-linked dominant/recessive; * Described only in a few consanguineous families; NA= not available.*

Up to 90% of OI cases can be explained by heterozygous mutations in one of the two genes encoding type I collagen (*COL1A1* and *COL1A2*). Rarely, biallelic *COL1A2* mutations have also been identified in consanguineous families [Costantini et al., 2018].

Glycine substitutions affecting the Xaa-Yaa-Gly pattern in type I collagen that unfold the helix (structural defect) are the most common cause of severe OI. Overall, mutations that impair the structure of the triple helix have more severe consequences than mutations leading to a reduced amount of protein, due to a dominant-negative effect [Makitie et al., 2019]. A particular form of OI characterized by high BMD derives from mutations in the C-terminus of collagen [Forlino and Marini, 2016].

Defects in proteins taking part in the processing and folding of type I collagen can also lead to OI. For instance, mutations in the molecular chaperone *FKBP10*, encoding immunophilin FKBP65, inhibit lysine hydroxylation and cause either OI or Bruck syndrome type 1, a condition characterized by congenital contractures and fractures (MIM #259450) [Alanay et al., 2010; Shaheen et al., 2010; Barnes et al., 2012]. Defects in the prolyl 3-hydroxylation complex, composed by three different proteins, the prolyl3-hydroxylase 1, the cartilage associated protein (CRTAP) and cyclophylin B, also lead to OI. As an example, biallelic *CRTAP* mutations lead to either lethal or severe OI (MIM #610682) [Cabral et al., 2007].

Some subtypes of OI arise from mineralization defects. For instance, mutations in *SPARC*, encoding the most abundant non-collagenous protein, osteonectin, affect the interaction of this protein with other components of the ECM, including collagen and hydroxyapatite, thus leading to bone fragility [Mendoza-Londono et al., 2015].

Defects in WNT signaling

WNT signaling is a major determinant of BMD and bone fragility. The significance of this pathway for bone health was discovered when biallelic loss-of-function mutations in *LRP5* were shown to cause two low BMD diseases: osteoporosis-pseudoglioma syndrome (MIM #259770) and early-onset osteoporosis (MIM #166710) [Gong et al., 2001; Hartikka et al., 2005].

Heterozygous mutations in *WNT1*, a secreted protein belonging to the canonical Wnt-β-catenin pathway, are responsible for an autosomal dominant form of osteoporosis (MIM #615221). This disease is characterized by defective modeling of long bones and low BMD in children and recurrent vertebral fractures, kyphosis and short stature in adults [Laine et al., 2013; Makitie et al., 2016]. On the other hand, biallelic *WNT1* mutations underlie severe autosomal recessive OI (MIM #615220) [Keupp et al., 2013; Laine et al., 2013; Pyott et al., 2013].

Recently, genetic defects in the frizzled-related protein 4 (*SFRP4*), an inhibitor of the Wnt pathway, have been detected in patients with Pyle's disease, which is characterized by bone fragility and cortical bone thinning (MIM #265900) [Simsek Kiper et al., 2016].

Mutations in *SGMS2*, encoding sphingomyelin synthetase 2, cause increased bone fragility and impaired bone mineralization which can be associated with severe skeletal dysplasia and neurological manifestations [Pekkinen et al., 2019].

PLS3 osteoporosis

In recent years, mutations in the plastin-3 gene (*PLS3*), have been found in patients with low BMD and recurrent fractures (MIM # 300910) [Van Dijk et al., 2013; Kampe et al., 2015; Kampe et al., 2017; Kampe et al., 2017]. Although the molecular function of PLS3 is still largely elusive, mutations in this Ca^{2+}-dependent actin-binding protein lead to reduced bone mineralization possibly due to impaired osteocytic mechanosensing [Kampe et al., 2017; Wesseling-Perry et al., 2017]. Since *PLS3* locates on the X chromosome, males are usually more severely affected than females.

Other forms of skeletal fragility

In addition to OI and osteoporosis-pseudoglioma syndrome, other syndromes featuring low BMD have been recognized. Bruck syndrome type 2 (MIM #609220), derives from biallelic mutations in a collagen-modifying enzyme, *PLOD2* [Ha-Vinh et al., 2004]. Spondylo-ocular syndrome (MIM #605822) is another autosomal recessive disease, characterized by low BMD, spinal compression fractures and cataract due to mutations in xylosyltransferase 2 (*XYLT2*) [Munns et al., 2015; Taylan et al., 2016]. This enzyme has a role in the biosynthesis of glycosaminoglycan chains and is essential for endochondral ossification [Taylan and Makitie, 2016].

High BMD diseases

Imbalances in bone remodeling can also lead to increased BMD. Two main types of high BMD diseases can be distinguished: 1) osteopetrosis, which is due to impaired osteoclast function and decreased bone resorption and 2) high bone mass (HBM) diseases caused by osteoblast dysfunction and increased bone formation.

Osteopetrosis is characterized by dense and at the same time fragile bones. Patients with osteopetrosis often feature also hearing and vision impairments. Some forms of osteopetrosis appear early after birth whereas in some milder forms patients receive a diagnosis only during adulthood. Mutations in over 10 different genes have been identified in osteopetrosis, including defects in the RANK/RANKL pathway [Bonafe et al., 2015]. Another major protein that can be mutated in osteopetrosis is the chloride channel called ClC-7 (CLNC7). *CLNC7* mutations impair the regulation of the relative acidity (pH) of osteoclasts and lead to reduced bone resorption [Kornak et al., 2001].

1.5.3 Treatments

Concerning pharmacological treatment of skeletal diseases, most of the achievements have been obtained for osteoporosis while no specific treatments are available for most of the skeletal dysplasias. Two major classes of drugs have been developed to increase BMD and decrease the risk of bone fractures: 1) anti-resorptive treatments that target osteoclasts and 2) anabolic treatments that act on osteoblasts.

Concerning anti-resorptive treatments, bisphosphonates, which are chemically stable derivatives of inorganic pyrophosphate, have been on the market for over four decades. Bisphosphonates suppress the activity of osteoclasts by promoting their apoptosis and consequently lead to increased bone mineralization and a rapid gain in BMD. Two commonly used bisphosphonates are alendronate and zoledronic acid. While per oral alendronate reduces the risk of sustaining vertebral fractures, this drug does not show any major impact on long bone fractures [Bone et al., 2004]. On the other hand, intravenous zoledronic acid given once a year efficiently reduces both vertebral fractures and long bone fractures [Black et al., 2007; Lorentzon, 2019]. Some side effects of bisphosphonate have been identified. The most severe consequence is osteonecrosis of the jaw, the risk of which increases with dental surgeries or removal of teeth. Negative effects on the gastrointestinal tract and kidneys as well as atypical femur fractures have also been identified [Pazianas et al., 2010; Lorentzon, 2019]. Another type of anti-resorptive drug is denosumab, which is a monoclonal antibody against RANKL. Even though this drug drastically decreases osteoclast activity, the effect on reduced bone resorption is rapidly reversed after only a few months [Bekker et al., 2004].

Among the anabolic treatments, teriparatide plays a major role. This drug is a recombinant protein form of the parathyroid hormone consisting of the first (N-terminus) 34 amino acids (PTH 1-34). When given intermittently, teriparatide acts on bone formation and leads to thickening of the cortex. The treatment decreases the risk of both vertebral and non-vertebral fractures. The gain in BMD is superior to anti-resorptive drugs but the high costs of the treatment limit its wide use [Lorentzon, 2019].

Romosozumab, an antibody targeting SOST, has been shown to increase bone formation (anabolic effect), to decrease bone resorption and also to reduce the risk of fractures in postmenopausal women [McClung et al., 2014; Cosman et al., 2016]. Recently, combined use of romosozumab and alendronate was shown to significantly reduce the risk of fractures compared to alendronate alone in over 4000 postmenopausal women with osteoporosis and fractures [Saag et al., 2017]. Concerns about the use of romosozumab have arisen due to side effects on cardiovascular and cerebrovascular systems [Saag et al., 2017].

Recently, an antibody against TGF-β has undergone animal testing but contrasting results have been obtained so far. While this drug rescues the phenotype in both *Crtap* knock-out mice and in mice with a heterozygous missense mutation p.Gly610Cys in *Col1a2*, the treatment is not effective in reducing the number of spontaneous fractures in a mouse model of autosomal dominant OI due to a *Col1a1* splicing mutation [Grafe et al., 2014; Tauer et al., 2019].

Although many pharmacological treatments have been tested and used in adults, less is known about the efficacy of these drugs for early-onset osteoporosis due to the limited number of clinical trials in young subjects [Makitie, 2013]. Bisphosphonates are the most widely applied treatment for children with OI even though this drug is not officially approved for pediatric use. Bisphosphonates lead to a gain in BMD, improve vertebral compression fractures and reduce pain in patients with moderate to severe OI but the results on long bone fractures are still inconclusive [Bishop et al., 2013; Hald et al., 2015; Palomo et al., 2015]. Despite the benefits seen in clinical practice, there are still unanswered questions regarding for example the appropriate drug dosage and treatment pausing, the quality of bone after treatment as well as the long-term safety of bisphosphonate treatment in young subjects.

Despite available pharmacotherapies, orthopedic surgeries are still often required to correct skeletal deformities and scoliosis or to prevent fractures by intramedullary rodding in patients with skeletal fragility. Furthermore, non-pharmacological interventions, including physical activity, physiotherapy, and calcium and vitamin D supplementation should be considered when treating these patients. A multidisciplinary team is needed to evaluate different aspects of the disease, such as major skeletal impairments, underlying genetic and molecular defects and potential extra-skeletal manifestations such as dental problems.

Regarding the skeletal dysplasias not featuring low BMD, no treatment is available for most of the conditions. One major problem is that skeletal dysplasias are an extremely heterogeneous class of diseases from both the phenotypic and genetic point of views. Often, impairments affecting tissues and organs other than the skeleton are found in patients with skeletal dysplasia. Recently, an efficient treatment has been developed for hypophosphatasia (HPP) [Whyte et al., 2012] and X-linked hypophosphatemic rickets [Carpenter et al., 2018]. HPP is a rare metabolic disease caused by mutations in the *ALPL* gene that lead to an impaired production of the tissue-nonspecific alkaline phosphatase (TNSALP) and defective mineralization. Forms of HPP that appear *in utero* are frequently associated with infantile and perinatal lethality due to impaired skeletal development and respiratory problems [Jelin et al., 2017]. A recombinant TNSALP enzyme, asfotase alfa, is given as a treatment to re-establish bone mineralization and to improve bone strength, thus leading to an increased perinatal and postnatal survival of HPP patients [Whyte et al.,

2012; Whyte et al., 2016]. In hypophosphatemic rickets, an antibody against FGF23 has been shown to prevent excessive renal phosphate loss and reverse radiographic changes of rickets [Carpenter et al., 2018].

1.6 APPROACHES TO INVESTIGATE GENETIC SKELETAL DISEASES

Skeletal diseases can be investigated using different genetic approaches. Family-based studies aim to identify genetic mutations having a large effect size in families with rare Mendelian skeletal diseases. On the other hand, oligogenic/polygenic conditions are investigated using genome-wide association studies (GWASs) by genotyping a large group of subjects to associate SNPs with small effect size to a certain trait (e.g. variance in BMD).

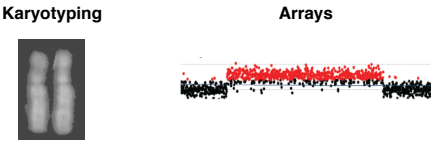
In the last decade, GWASs identified several genetic loci linked to BMD and the risk of fracture [Rivadeneira et al., 2009; Estrada et al., 2012; Paternoster et al., 2013]. In 2017, over 150 novel genetic loci associated with BMD were detected in a large GWAS [Kemp et al., 2017]. In the most recently published study, 518 loci (301 novel) were identified as being associated with estimated BMD (eBMD) in the largest GWAS ever performed [Morris et al., 2019]. These findings explain 20% of the eBMD variance [Morris et al., 2019]. In this work, the authors also identified the protein-coding genes in close proximity to the most significant SNPs and evaluated the skeletal phenotype of 126 knockout mice in which these genes were disrupted. This approach led to the identification of a novel candidate gene for osteoporosis involved in WNT signaling, *Daam2*: mice lacking this gene exhibited reduced bone strength. However, this is not the only example in which GWAS hits can potentially pinpoint genes linked to monogenic conditions. In fact, some variants that have been identified in previous GWASs are close to genes in which rare pathogenic variants have been detected in patients with rare skeletal diseases [Rivadeneira and Makitie, 2016].

Although GWASs have led to outstanding discoveries concerning the molecular mechanisms predisposing to bone fragility, these findings have had a limited impact on the clinical practice. Genetic testing of families with rare skeletal diseases offers several benefits, such as providing a diagnosis and a prognosis, enabling an appropriate management and surveillance of the disease as well as offering genetic counseling to the patients and their family members. During the last years, several rare monogenic forms of skeletal diseases have been characterized in family-based studies with MPS and subsequent functional validation of genetic findings.

1.6.1 From genetics to genomics

In the last four decades, the approaches to identify a novel disease-causing gene have changed with the evolution of sequencing technologies. Before the introduction of MPS, identification of novel genes underlying monogenic conditions was conducted by positional cloning and linkage analysis to identify regions to be prioritized for Sanger sequencing. These techniques were time-consuming and required large multi-case families.

From low-throughput technologies, like Sanger sequencing, where only small parts of DNA (up to 800 bp in length) could be amplified and sequenced at the same time, we have recently entered to an era of high-throughput technologies, where an entire exome (all the exons of the genome) or genomes (the complete set of DNA) of an individual can be sequenced in only few days. In the last 5-10 years, the cost of performing whole-exome sequencing (WES) and whole-genome sequencing (WGS) have considerably decreased with the outcome that the exome/genome of several thousands of individuals has been sequenced and several disease-causing genes have been identified. Gene panels, targeting a limited set of genes, or WES/WGS are appropriate methods for diagnosing conditions characterized by genetic heterogeneity, like OI [Bardai et al., 2016]. In the "post-genome era", Sanger sequencing is mainly applied to sequence single or a few candidate genes as well as to validate WES/WGS findings. The resolution of the genome-wide assays has increased substantially (Fig. 6) [Wright et al., 2018].



	Karyotyping	Arrays	WES/gene panels	WGS
Resolution	5-10 Mb	50-100 kb	1 bp	1 bp
Type of detected variants	Aneuploidy and polyploidy; large deletions, duplications, inversions, translocations (balanced/unbalanced)	Large unbalanced deletions and duplications (CNVs)	All variants in the exome/within coding regions of targeted genes; low coverage first exons of genes; low power for structural variants	All variants in the genome; low power for variants within repetitive regions
Number of detected variants	0-1	~10-30	~20,000	~4-5 millions
Diagnostic yield	Very low	Quite low	High	Very high
Risk for incidental findings	Low	Low	High	High

Figure 6. Evolutions of the genetic methods to investigate rare skeletal diseases. From karyotyping to the most recent and powerful method, WGS. Along with the increased resolution and the possibility to detect most types of genetic variants, MPS technologies also increment the risk of incidental findings. Figure adapted from Wright et al., 2018, with permission from the Publisher.

In the "pre-genome era", chromosomes were analyzed by karyotyping under the microscope. This method only allows for the identification of SVs involving at least 5 Mb of DNA. On the other hand, standard comparative genomic hybridization arrays (array-CGH) can identify CNVs up to 50 kb in size.

Nowadays, custom made array-CGHs are available to pinpoint smaller CNVs in targeted regions within the genome. SNP-arrays provide similar information to array-CGH but they also detect uniparental disomy and autozygosity, regions in which both copies of DNA derive from a common ancestor most likely due to consanguinity between the parents. However, balanced abnormalities, like inversions and translocations, can not be identified by arrays. Finally, WGS and WES can identify most types of genetic variants with a resolution of 1 bp.

Considering the number of variants detected by these different methods, usually no variants are detected by karyotyping because chromosomal abnormalities are rare among healthy individuals as well as in patients with skeletal diseases. On the other hand, as many as five million variants can be identified by WGS [Genomes Project et al., 2015].

Although MPS technologies offer several benefits, these methods generate a massive amount of data that have to be processed, analyzed and stored. Furthermore, both variant prioritization and interpretation are challenging, especially if the disease-causing variant reside outside the coding regions of the genome. Finally, MPS data might contain genetic information about conditions other than those affecting the skeleton (e.g. genetic variants that predispose to cancer). In this way, MPS can lead to a relatively high yield of incidental findings. Despite these challenges, MPS approach has completely revolutionized the field of rare skeletal diseases by leading to the identification of several novel disease mechanisms and by increasing the diagnostic rate.

2 AIMS

Although the use of MPS has considerably increased the diagnostic yield of rare skeletal diseases, both the phenotypic and the genetic spectrum of many of these entities still remain to be fully characterized. Furthermore, the genetic causes and molecular mechanisms underlying some extremely rare conditions have yet to be identified.

The overall goal of this thesis is to expand the genetic and pathomolecular background of rare metabolic bone diseases by studying patients or families affected by early-onset skeletal fragility, HBM as well as skeletal dysplasia (SMD and SEMD).

The specific aims are:

- To identify novel disease-causing gene variants and correlate them with the clinical phenotype of the patients (**Papers II-IV**)
- To discover potentially novel gene-disease associations (**Papers I**, **II** and **V**)
- To investigate novel genetic and molecular mechanisms leading to disease (**Paper V**)

3 PATIENTS AND METHODS

3.1 ETHICAL APPROVALS AND PATIENTS' CONSENTS

The majority of the patients included in this thesis are Finnish and were recruited at the Children's Hospital, Helsinki University Hospital, Finland. Due to the extreme rarity of some of the studied conditions, patients were also recruited from other clinics through an international collaborative network. Ethical approvals for our research studies were obtained from the Helsinki University Hospital Ethics Committee and the other Institutional Ethics Boards. All studies were carried out according to the World Medical Association Declaration of Helsinki. Prior to inclusion to our studies, a written informed consent was signed by each participant and/or the caregivers.

3.2 PATIENTS

The clinical evaluation of each patient was carried out as part of the patient's routine clinical care. The phenotype and skeletal features were carefully investigated by expert clinicians and radiologists. Biochemical parameters, including parameters of calcium metabolism and bone turnover markers, were measured in the majority of the patients. BMD was measured with DXA. BMD values were reported using the T-score in adults ≥ 30 years of age whereas the Z-score was adopted for subjects < 30 years of age. Furthermore, a preliminary genetic testing (with negative result) was performed in some of the patients before inclusion to our research, as described in the following paragraphs.

In **papers I-II** we studied a cohort of patients with mild to severe skeletal fragility. These patients were subdivided into two groups according to the inclusion criteria described in Fig. 7.

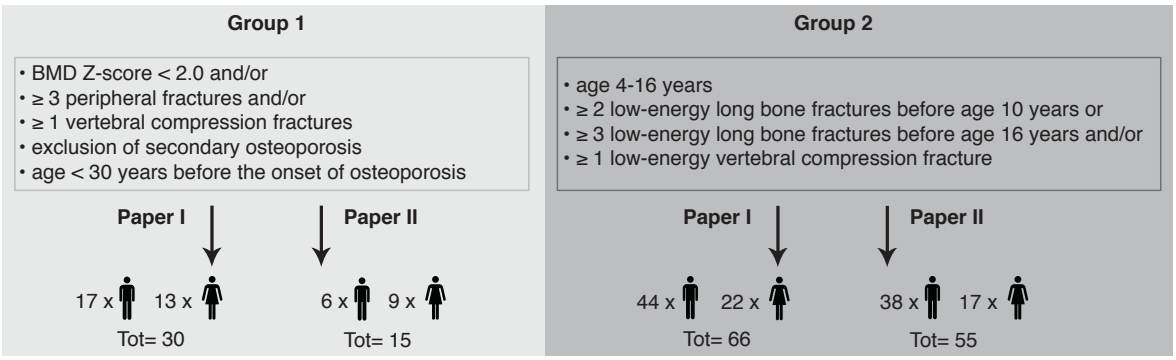


Figure 7. Inclusion criteria and number of patients with skeletal fragility included in paper I and II. *Tot= total*.

Group 1 encompassed children or young adults (N= 30 in **paper I**; N= 15 in **paper II**) fulfilling the diagnosis of early-onset osteoporosis. Sequencing of type I collagen genes

had been performed for some of the patients but no disease-causing variants had been identified. Group 2 involved children (N= 66 in **paper I**; N= 55 in **paper II**) who sustained multiple long bone and/or compression fractures but had usually normal BMD. Group 2 was carefully selected from an epidemiological study performed by Mäyränpää et al. [Mayranpaa et al., 2010; Mayranpaa et al., 2012]. No systematic genetic testing apart from *LRP5* screening was carried out for these patients before inclusion to our study [Saarinen et al., 2010].

In **paper III** two patients with HBM were recruited. Patient #1 (male, age 53 years) was enrolled to our study due to high BMD (T-score +10.1 for lumbar spine) and surgeries required to correct excessive bone formation. Patient #2 (female, age 17 years) was referred to our clinical care due to abnormal bone structure on radiographs and high BMD (Z-score 3.0 for lumbar spine). No preliminary genetic investigations were performed for Patients #1-2 prior to enrolment to this study.

In **paper IV** we recruited a Finnish trio (parents and their affected child) with a rare form of SMD. The most striking skeletal features identified in the index patient (Patient #1, male, 12.5 years), born from healthy non-consanguineous parents, were short stature and metaphyseal abnormalities with “corner fracture” appearance in long bones. No mutation was identified in any skeletal dysplasia-related gene included in the commercially available Comprehensive Skeletal Dysplasias and Disorders Panel by Blueprint Genetics at the time of recruitment. Four other patients with a similar skeletal phenotype (Patients #2-5) as well as some healthy and affected family members were recruited through the international collaboration.

In **paper V** a Finnish trio with a potentially novel form of SEMD was recruited. The index patient (Patient #1, male, 4.5 years), was born from healthy non-consanguineous parents. He featured severe and disproportionate short stature and on radiographs short ribs, vertebral changes and metaphyseal and epiphyseal abnormalities. The patient was mutation negative for all the genes included in the Comprehensive Skeletal Dysplasias and Disorders Panel (Blueprint Genetics). Two-unrelated Korean children (Patients #2-3) featuring the exact same abnormal clinical phenotype as the Finnish patient were also enrolled to our study along with other healthy and affected family members.

3.3 METHODS

3.3.1 Genetic testing

In order to identify the genetic cause of disease in patients with rare skeletal phenotypes, different genetic approaches and methods were used in this thesis (Table 4).

Table 4. Methods used in each paper.

	# targeted genes	Paper I	Paper II	Paper III	Paper IV	Paper V
Sanger Sequencing	1	Sequencing of *CRTAP* in Groups 1-2		Sequencing of *LRP5* in Patient #1	Sequencing of *FN1* in Patient #2	Sequencing of *RPL13* in Patients #2-3
Gene panel for osteopetrosis	21			Used for Patient #2		
Custom made array-CGH	> 1,150		Performed in Groups 1-2			
Whole-exome sequencing	~ 20,500				Performed in Patients #3-5	
Whole-genome sequencing	~ 20,500				Performed in Patient #1	Performed in Patient #1

Patients from Groups 1-2 in **papers I-II** were investigated using either Sanger sequencing or a custom designed array-CGH. The two patients with HBM described in **paper III** were analyzed using Sanger sequencing and a panel targeting 21 genes associated with osteopetrosis. Finally, WGS, WES as well as Sanger sequencing were carried out for patients included in **papers IV-V**.

Sanger sequencing

Sanger sequencing was used in this thesis when only one gene was to be analyzed. In **paper I** we used this method to investigate whether rare variants in the *CRTAP* gene, encoding the cartilage-associated protein, were present in patients with mild to severe skeletal fragility. The same method was also chosen for sequencing other candidate genes in **papers III-V**.

Briefly, in Sanger sequencing the regions of interest (up to 400 bp each) are amplified by polymerase chain reaction (PCR). Sanger sequencing reactions are performed individually for each amplified fragment using forward and reverse PCR primers, a DNA polymerase as well as the four deoxyribonucleotide triphosphates (dNTPs). Furthermore, low concentrations of labeled dideoxyribonucleotide analogs (ddNTPs), lacking two OH groups on two sugar carbons, are also added. Once a ddNTP is incorporated within the extending fragment DNA synthesis is terminated. Finally, all DNA fragments are separated by size and the fluorescent signal is detected on the semi-automated capillary electrophoresis instrument Applied Biosystems 3730 DNA Analyzer.

Custom designed array-CGH

A custom-made array-CGH with enriched probe density in >1,150 genes linked to skeletal diseases and/or associated with cilia function was designed to identify CNVs in patients

with different degrees of skeletal fragility (**paper II**). Our design had two arrays per slide with a total number of 400 k probes per sample (2 x 400 k). Out of 440 k oligonucleotide probes 180 k were evenly distributed throughout the genome whereas 220 k probes specifically targeted our genes of interest with an average coverage of one oligonucleotide probe (60 bp in length) per 100 bp and one probe per 500 bp in the non-coding regions.

The name array-CGH refers to the sets of DNA that are used in this method: a test DNA (patient DNA), a sex-matched reference DNA (control DNA) and a large set of DNA probes immobilized on the array. The array-CGH workflow consists of four main steps that are described in Fig. 8.

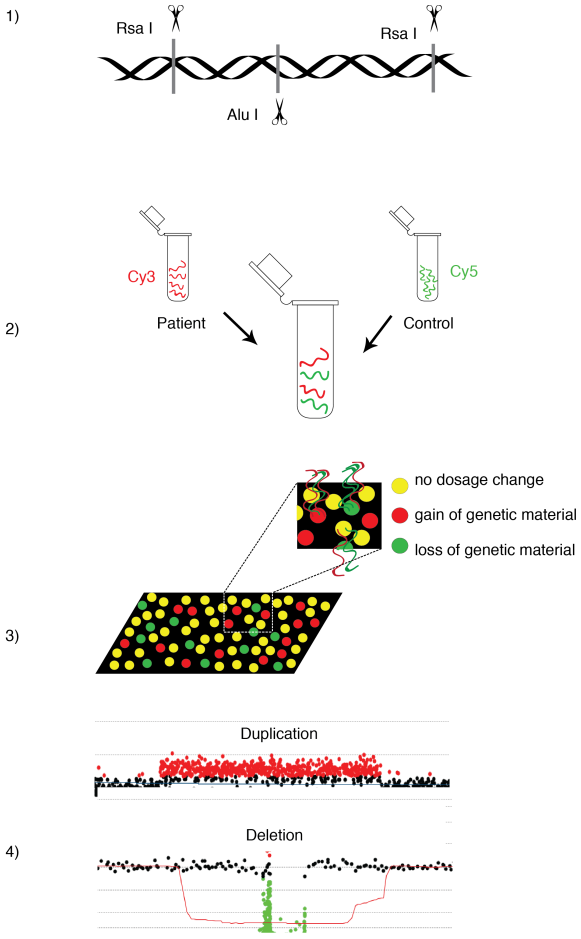


Figure 8. The array-CGH workflow consists of four main steps: 1) the same amount of test DNA and reference DNA are fragmented using restriction enzymes 2) test DNA fragments and reference fragments are first labeled using two different fluorescent dyes and then mixed together; 3) the DNA mix is hybridized on the array; the average color on the array corresponds to relative amount of test and reference samples; 4) the results are analyzed using a specific analysis software.

MPS

In order to identify the disease-causing variant in either a list of candidate genes (**paper III**) or in the entire exome/genome (**papers IV-V**) we used MPS. In the MPS field, several technologies with diverse sequencing chemistries have been developed by different companies. Among these, Illumina is the most widely chosen and it is also the technology that was used in this thesis.

The general workflow of MPS sequencing consists of three mains steps: DNA library construction, amplification of the separated DNA fragments (not always needed) and

sequencing (Fig. 9A). First, the DNA is randomly fragmented and two double-stranded adaptor oligonucleotides are ligated to each fragment. Each DNA fragment is then physically separated and amplified to give clusters of monoclonal DNA that will be used as templates for sequencing. Sequencing reactions are performed via sequencing-by-synthesis using a DNA polymerase and dNTPs, as in Sanger sequencing. In MPS sequencing the incorporation of nucleotides into growing DNA is recorded during the reaction (each base is detected by a specific light signal). In this thesis, sequencing was performed at both ends of the fragments to generate pair-end reads.

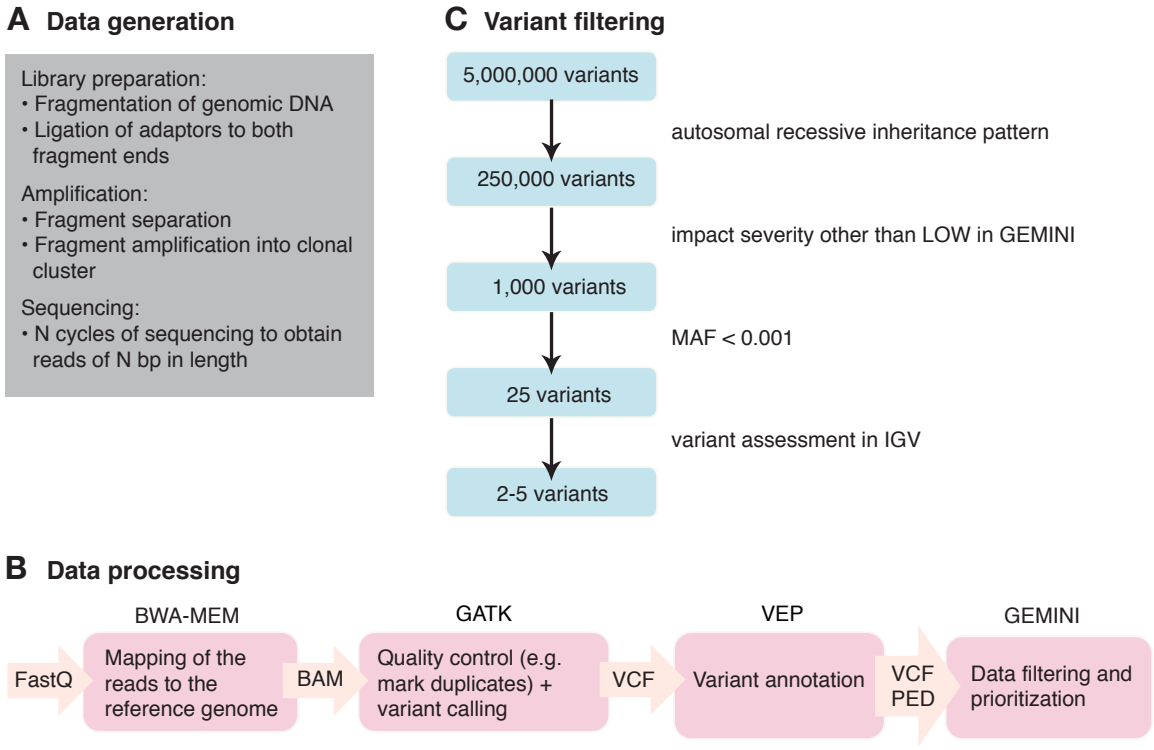


Figure 9. MPS workflow. Libraries are prepared, amplified and finally sequenced (A). The sequencing data are delivered from the sequencing facility in form of FastQ files (B). The sequencing reads are then mapped to reference genome and variants are called, annotated and loaded into GEMINI (B). Finally data filtering, prioritization and analysis take place (C); example based on a rare autosomal recessive condition. FastQ, BAM, VCF and PED are different file formats.

Once the sequencing part is accomplished, an enormous amount of raw data (sequencing reads) is produced and delivered in fastQ format. At this point, data pre-processing is performed (Fig. 9B). During this step, sequencing reads are mapped to the reference genome (GRCh37/UCSC hg19) and quality control is carried out to correct for technical biases. Afterwards, variant discovery takes place. Once all the positions at which the DNA sequence does not correspond to the reference genome are identified, the variants are annotated and filtered (Fig. 9B). Only after this step, the genetic analysis of the data takes place (Fig. 9C). SVs, which are not detectable using the pipeline designed for small scale

variants, were identified using the FindSV pipeline on GitHub (https://github.com/J35P312/FindSV) and manually assessed.

In order to identify the genetic cause of HBM in Patient #2 of **paper III** a MPS targeting the coding regions of 21 genes linked to osteopetrosis and HBM was performed at Bristol RGC, UK Genetic testing Network. The targeted regions were selected using Agilent Haloplex Method. Sequencing reads (2 x 150 bp) were generated using the Illumina MiSeq platform.

In **paper IV** Patients #3-5 were analyzed by WES. WES for Patient #3 was carried on at CEGH-CEL-Universitade de São Paulo. Library preparation was performed by capturing the exome with the Illumina TruSeq PCR-free and sequencing was performed on the HiSeq 2500 instrument. PCR-free methods offer a more comprehensive coverage in the challenging regions, such as GC rich and repetitive regions. Reads mapping was performed using Burrows-Wheeler Aligner (BWA-MEM) [Li, 2013], variants were called using Genome Analysis Toolkit (GATK) [Van der Auwera et al., 2013] and annotated using ANNOVAR [Wang et al., 2010]. WES for Patients #4-5 was conducted at GeneDx. Exome was captured with the Clincial Research Exome kit (Agilent Technologies, Santa Clara, CA) and sequenced on an Illumina system generating 2 x 100 bp pair-end reads. A detailed variant processing protocol has been described previously [Tanaka et al., 2015].

In **papers IV-V** we used WGS to identify the cause of disease in the two Finnish index patients (both named Patient #1 in the two papers). Both library preparation and sequencing were performed at the Science for Life laboratory (SciLifeLab) in Stockholm, Sweden. Library were prepared using the Illumina TruSeq PCR-free methods generating fragments of 350 bp. Afterwards the fragments were sequenced on the HiSeq X instrument as 2 x 150 bp pair-end reads with an average coverage of 30X. Reads mapping was carried out using BWA-MEM, data processing and variant calling were performed according to GATK best practice and variant annotated using Variant Effect Predictor (VEP) [McLaren et al., 2016].

3.3.2 Variant filtering, prioritization and interpretation

In **papers I-III** the genetic investigations led to a limited number of variants to be analyzed. In this way, variant interpretation could be done manually for each variant that was identified (including CNVs). This also applies to the samples that were investigated by Sanger sequencing in **papers IV-V**.

WES and WGS produce thousands to millions of variants that differ from the reference genome. In this way, after each DNA change is annotated, the variants have to be filtered (Fig. 9C) and prioritized. The Finnish families in **paper IV-V** were analyzed as trios,

meaning that the index patients plus their parents were sequenced. This strategy offers the possibility to filter the data according to the inheritance pattern thus considerably reducing the numbers of variants in GEMINI, a flexible tool used in this thesis to explore genetic variation (Fig. 9B) [Paila et al., 2013]. Furthermore, since we are studying rare diseases, a MAF < 0.001 in the gnomAD and SweGen [Ameur et al., 2017] databases was used (Fig. 9C). Moreover, by using the filter "impact severity" in GEMINI, it was possible to discard variants that are likely to have a small effect and/or variants of difficult interpretation (e.g. synonymous variants, intergenic variants as well as variants in the 3'/5'-UTR) while retaining variants affecting known non-coding RNAs or promoter regions (Fig. 9C) [Paila et al., 2013]. Once one or a few candidate variants are identified, they were manually evaluated using the program Integrated Genome Viewer (IGV). *In silico* prediction tools, like SIFT [Ng, 2003] and PolyPhen-2 [Adzhubei et al., 2013], give an hint about the effect of missense mutations on protein function by performing multiple sequence alignment in related proteins (paralogs in humans and orthologs between different species). Other *in silico* tools can instead predict the effect of splicing variants [Jian et al., 2014]. The Genomic Evolutionary Rate Profiling (GERP) score instead offers information about the conservation of a certain amino acid between different species. The information provided by most of the available *in silico* prediction tools are combined into a unique score, named Combined Annotation-Dependent Depletion (CADD) score [Kircher et al., 2014]. The higher CADD score a variant has, the more deleterious the change is likely to be.

Following this, other parameters have to be evaluated. First of all, it is important to investigate if a variant affects a gene that has already been linked to the studied disease. ClinVar is a database that contains all the genetic variants that have been linked to a certain phenotype. Some of these variants have been functionally validated whereas some others are only variant-phenotype observations. If the variant of interest has already been reported as causing a similar skeletal disease, there is strong evidence that the variant is pathogenic. However, if a variant has never been described before, segregation studies as well as experiments at mRNA and/or protein level might be needed. If a variant is identified in a gene that is not yet associated to a Mendelian disease, further investigations have to be carried out. For example, mRNA/protein expression levels in the relative tissue/cell type are a useful information that can be acquired from publicly available databases. Moreover, functional studies on cellular or animal models are always necessary when a new candidate gene is identified.

3.3.3 Variant validation

Different approaches to validate a certain genetic variant are used and they can be subdivided into two major groups: *in silico* validations and *in vivo* validations. In this thesis

several *in vitro* experiments at DNA and protein level were performed. In addition, a zebrafish model was generated.

In vitro validations

In order to confirm the presence of the candidate variants identified by MPS in the DNA Sanger sequencing was performed (**papers III-V)**.

In **paper II** breakpoint PCR and Sanger sequencing were performed to confirm two rare CNVs that were identified by array-CGH. Breakpoint PCR consists of performing PCR (or long-range PCR) and primer walking around the breakpoints detected by array-CGH. This strategy enabled us to pinpoint and sequence the exact breakpoints of one pathogenic deletion. Furthermore, by using this method it was possible to show that one likely pathogenic duplication was in tandem. Breakpoint PCR was also applied to screen for mutation in other family members.

Finally, WGS was used in **paper II** to exclude the presence of any likely pathogenic variant other than the identified CNV in the patient harboring the duplication.

Since germline mutations are present in all cells of the body, blood is a convenient source for extracting genomic DNA from patients with Mendelian diseases. However, protein expression is highly tissue-dependent. Since it was not possible to obtain a bone biopsy, we collected skin biopsies from patients and controls as a 'proxy-tissue' for investigating the cellular consequences of the mutation at the protein level in **paper V**. In order to validate the genetic findings in this paper, where a new candidate gene for SEMD dysplasia was identified, we collected a skin biopsy from each of the three index patients as well as from some healthy and affected family members. Biopsies from unrelated controls were also included. Primary dermal fibroblasts were isolated and cultured according to standard protocols as previously described [Pekkinen et al., 2019]. Expression of the target protein was investigated by Western blot (WB) and immunostaining.

WB was performed to evaluate if there was a difference in the expression level of a candidate ribosomal protein in fibroblasts from patients compared with controls. For WB, total protein lysate is extracted from fibroblasts cultures using standard procedures. According to the basic principles of WB, proteins are separated by gel electrophoresis and then transferred (blotted) to a membrane. In order to specifically detect the protein of interest, the membrane is blocked and incubated with a primary antibody that specifically targets the protein of interest. The detection of the protein-antibody interaction is detected by the use of a secondary antibody conjugated with horseradish peroxidase (HRP). By

adding the HRP substrate the HRP enzyme activity (and thus the amount of protein of interest) is detected by chemiluminescence.

Immunocytochemistry (ICC) experiments were carried out to investigate if the candidate ribosomal protein is localized to the same subcellular compartments in the patients' fibroblasts compared to controls. Furthermore, the co-localization of our protein of interest with other ribosomal proteins was investigated. Briefly, fibroblasts were cultured on cover slips for three days, fixed in 4% paraformaldehyde, permeabilized 0.1% triton-X in PBS, blocked in in 0.1% BSA in PBS, and then incubated with primary antibodies targeting the proteins of interest. Secondary fluorescently labeled antibodies were then added to allow signal detection by confocal imaging. Colocalization analysis was performed using the Colocalization Test plugin of ImageJ Fiji, where agreement in localization is expressed as Pearson's correlation coefficient [Dunn et al., 2011].

In vivo validations

Although cellular models are a useful tool to study some aspects of the molecular mechanisms leading to disease, animal models enable us to study the disease pathogenesis in the whole organism from embryo to adult stage. *In vivo* studies are necessary to investigate new gene-disease correlations.

In **paper V** we knocked out our gene of interest in zebrafish through CRISPR-Cas9 genome editing [Doudna and Charpentier, 2014]. Zebrafish (*Danio rerio*) was chosen because it has been shown to be a good model for investigating skeletal diseases [Witten et al., 2017]. By using this bony fish it is possible to study bone and cartilage formation as well as skeletal deformities since bone development and some basic skeletal components are highly conserved between teleost and humans [Witten et al., 2017]. In previous studies on OI and osteoporosis, this animal model has been used to understand the molecular mechanisms leading to disease as well as to perform drug testing [Van Dijk et al., 2013; Gistelinck et al., 2016; Gioia et al., 2017; Fiedler et al., 2018; Gistelinck et al., 2018]. Zebrafish has also been applied to validate the skeletal phenotype observed in patients with novel forms of skeletal dysplasia, such as a skeletal ciliopathy caused by *KIAA0753* mutations [Hammarsjö et al., 2017].

In general, zebrafish is widely used as a model organism because approximately 70% of the human genes have a orthologue in this species [Howe et al., 2013]. Furthermore, every week a couple of fish can produce hundreds of eggs that are externally fertilized and can be easily visualized and manipulated.

The CRISPR-Cas system is an adaptive system found in bacteria to protect them against viruses and plasmids. In our genome editing, a Cas9 endonucleases as well as a single

guide RNA (sgRNA) targeting a region just downstream of the mutated loci in the patients were injected into the zebrafish embryos. The sgRNA has a scaffold sequence that binds Cas9 as well as a 20 bp sequence that specifically targets our gene of interest specifically. In order for the Cas9 to bind the target, the sgRNA has to hybridize to a sequence that locates close to a NGG protospacer-associated motif (PAM sequence). Only in this way the Cas9 can perform a double strand break close to the targeted DNA sequence, which is consequently repaired by non-homologous end joining (NHEJ) resulting in the introduction of indels.

In order to obtain the knocked-out fish we inter-crossed two fish with the same heterozygous frameshift mutation in our gene of interest from first filial (F1) generation. Phenotypic characterization was performed by gross analysis of the larvae from 1 to 5 days post-fertilization (dpf) and by skeletal tissue staining upon fixation on 5 dpf larvae. Cartilage development and mineralization were evaluated using alizarin red and alcian blue staining. Cartilage deformities in the head were investigated based on the measurement of the angle between the left and the right ceratohyals.

4 RESULTS AND DISCUSSION

4.1 SANGER SEQUENCING OF KNOWN DISEASE GENES

In **papers I** and **III** we used Sanger sequencing to search for the genetic cause of disease in patients with early-onset skeletal fragility and HBM diseases.

4.1.1 *CRTAP* screening in patients with skeletal fragility

Previous studies have suggested that even carriers of mutations in genes causing severe recessive OI or other recessive syndromes characterized by low BMD may feature early-onset osteoporosis [Hartikka et al., 2005; Keupp et al., 2013; Laine et al., 2013]. In 2006, Morello et al. showed that the parents of index patients with *CRTAP* mutation-related OI were healthy carriers [Morello et al., 2006]. However, to the best of our knowledge, no larger studies investigating potential associations between *CRTAP* variants and low BMD and bone fractures have been performed.

To determine whether rare variants in the *CRTAP* exons or splicing boundaries could explain low BMD and increased bone fractures, in **Paper I** we sequenced *CRTAP* in two groups of patients with mild to severe skeletal fragility. While the phenotypes in the two groups were partly overlapping, overall, subjects in Group 1 more commonly had low BMD and vertebral fractures whereas Group 2 had multiple long-bone fractures and normal BMD. The screening of *CRTAP* detected no clearly pathogenic variants. We detected four synonymous SNPs in both groups of patients and one synonymous SNP in only one patient in Group 2 (Table 5). The patient harbouring this rare SNP was an 11-year-old boy who sustained two compression fractures but had normal BMD values for the lumbar spine and whole body. Although this SNP has never been reported in the Finnish population before, synonymous SNPs are unlikely to have a pathogenic effect. No parental samples were available for testing.

Regarding all identified SNPs in *CRTAP*, only one SNP, rs11558338, was less common in the fracture-prone children (Group 2) compared to the osteoporosis group (Group 1) and the general population (Table 5). However, even if the uncorrected p-value was statistically significant (p=0.0304), after correcting for multiple testing (Bonferroni correction) the p-value became insignificant suggesting that the finding is most likely random. As expected, the number of heterozygous individuals for each SNP was higher than the number of homozygous individuals in both groups of patients. A significant pathogenic role of a homozygous genotype for the identified SNPs can be excluded in light of the relatively high frequency of homozygous individuals in the general population.

Table 5. Genetic variants identified in the *CRTAP* gene.

SNP	DNA change *	Involved amino acid	MAF ExAc	MAF gnomAD	MAF Group 1	MAF Group 2	# patients Group 1		# patients Group 2	
rs11558338	c.213G>A	p.Leu71=	22%	15%	20%	8%	10		8	
							Het 8	Hom 2	Het 6	Hom 2
rs4076086	c.534C>T	p.Asp178=	17%	17%	22%	18%	11		21	
							Het 9	Hom 2	Het 18	Hom 3
rs35357409	c.558A>G	p.Ala186=	1%	1%	0%	1%	0		1	
							Het /	Hom /	Het 1	Hom /
rs1135127	c.1032T>G	p.Thr344=	38%	37%	38%	40%	20		42	
							Het 17	Hom 3	Het 31	Hom 11
rs1135128	c.1044G>A	p.Ser348=	38%	37%	38%	40%	20		42	
							Het 17	Hom 3	Het 31	Hom 11

** Reference sequence: NM_006371.4.*

In conclusion, we were aiming at searching for rare damaging *CRTAP* mutations as the cause of mild skeletal fragility in our two patient cohorts but we only identified synonymous changes that can also be detected at relatively high frequency in the general population. Our negative might indicate that there is no association between *CRTAP* variants and decreased bone mass and/or increased fractures in children and young adults when a typical phenotype of severe OI is not present. Since we only identified synonymous variants it is not possible to exclude that monoallelic variants with more severe consequences, like loss of function or certain missense mutations, may lead to decreased bone mass or increased risk of fractures. Moreover, other types of variants that are not detectable by Sanger sequencing might also affect *CRTAP* and contribute to disease.

4.1.2 Genes associated with HBM diseases

In **paper III** exons 2-4 of the *LRP5* gene were sequenced in a male patient with autosomal dominant HBM (Patient #1). Most of the HBM phenotype mutations in *LRP5* have previously been identified in these exons. We identified a novel missense mutation in exon 3, c.592A>T (p.Asn198Tyr) (reference sequence: NM_002335.3) (Fig. 10).

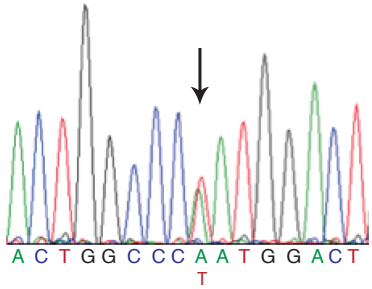


Figure 10. Sanger sequencing showing the missense variant, c.592A>T (p.Asn198Tyr), identified in *LRP5*.

Exons 2-4 encode the β-propeller 1, the extracellular domain that interacts with different proteins of the WNT pathway, such as the inhibitors SOST and DKK1. Since 71% of the residues, including Asn198, are conserved among the LRP5/6 receptors, Bourhis et al. used LRP6 as a surrogate for LRP5 to validate the effects of 9 mutations in *LRP5* identified in patients with HBM [Bourhis et al., 2011; Joiner et al., 2013]. Asparagine 198 strictly interacts with the asparagine of the 'Asn-Xaa-Ile' motif of the SOST/DKK1 inhibitors [Bourhis et al., 2011]. Biolayer interferometry showed that a mutation affecting the same codon as our mutation, p.Asn198Ser, causes a significant loss in affinity for DKK1 and SOST [Bourhis et al., 2011]. Furthermore, a structural model of the LRP5 protein clearly shows how the asparagine to serine change at position 198 leads to a loss of affinity due to the different amino acid structure [Gregson et al., 2014]. The mutation identified in our study substitutes asparagine for tyrosine and it is likely to be even more deleterious than the p.Asn198Ser change since an uncharged amino acid is substituted with a highly hydrophobic one. The identified gain-of-function mutation in *LRP5* is likely to affect the binding to SOST and DKK1, which are no longer capable to antagonize the association of WNT ligands to LRP5, leading to increased WNT signaling [Van Wesenbeeck et al., 2003; Semenov and He, 2006].

A female patient with HBM (Patient #2) was investigated using a commercially available gene panel for osteopetrosis and HBM. A novel herterozygous frameshift deletion in exon 2 of *AMER1* (reference sequence: NM_152424.3), c.655del (p.Glu219Argfs*6), was determined as the cause of osteopathia striata with cranial sclerosis (OSCS). Striations in the long bones, also present in our patient, are a hallmark of this disease in female patients [Jenkins et al., 2009]. *AMER1* locates on the X-chromosome and males with a hemizygous mutation are in general more severely affected than heterozygous females. AMER1 interacts with several proteins, including members of the β-catenin "destruction complex". Loss of function mutations in *AMER1* are likely to affect the ubiquitination of β-catenin, which is no longer degraded and leads to increased bone formation. Furthermore, AMER1 is a tumor suppressor and somatic *AMER1* mutations have been associated to Wilms tumors, a pediatric cancer. Our presently 17-year-old patient is carefully followed for potential extra-skeletal complications.

Although both gain-of-function mutations in *LRP5* and loss-of-function mutations in *AMER1* in **paper III** lead to an excessive bone accrual due to an increased activation of the WNT pathway, the phenotypes of the two studied patients were remarkably different. While cranial hyperostosis with no macrocephaly was evident in the patient with HBM, severe macrocephaly with small jaw was detected in the patient with OSCS. Moreover, the tubular bones of the two patients also appeared different: severe hyperostosis was seen in the

diaphyses of the HBM patient whereas typical striations and abnormally shaped long bones were noticed in the OSCS patient.

4.2 CUSTOM DESIGNED ARRAY-CGH TARGETING GENES RELATED TO SKELETAL FRAGILITY AND CILIARY FUNCTION

In **paper II** a custom designed array-CGH was used to search for rare CNVs in patients with skeletal fragility. The array-CGH was designed to especially cover genes that have already been linked to skeletal diseases and/or genes that are known or predicted to play a role in cilia. The specific focus was on CNVs because firstly, only a small number of large deletions and duplications have thus far been detected in patients with skeletal fragility. Secondly, this approach was regarded as a potential way of detecting novel genetic loci underlying skeletal fragility. Concerning the primary cilium, defects in this organelle have already been described in some types of skeletal dysplasia but no association with skeletal fragility has been determined so far. Since osteocytes are the mechanosensors of bone, our hypothesis was that abnormalities in the primary cilium could affect the way osteocytes perceive mechanical forces.

In order to identify rare CNVs we excluded variants listed in the Database of Genetic Variation (DGV) and variants that were present in a Finnish control group that had previously been described and screened using array-CGH [Viljakainen et al., 2015].

4.2.1 CNVs in known genes associated to skeletal fragility

Our array-CGH analysis led to the identification of two novel heterozygous CNVs in genes involved in skeletal fragility.

The first CNV is a novel deletion of around 4 kb, chr7: 94,024,366–94,028,364 (reference genome: GRCh37), identified in the N-propeptide of *COL1A2,* g.491_5060del (reference sequence: NM_000089.3) in a male patient affected by early-onset osteoporosis. The patient's phenotype was not as severe as in other patients with qualitative defects in type I collagen. *COL1A2* mutations, in general, have less severe consequences compared to *COL1A1* mutations. This is mostly due to the fact the collagen triple helix is formed by two α-1 chains and only one α-2 chain. Moreover, it has been shown that quantitative defects are better tolerated than qualitative defects affecting the Xaa-Yaa-Gly pattern [Forlino and Marini, 2016]. According to *in silico* predictions, the identified deletion is likely to lead to a frameshift in the ORF resulting in the introduction of an early stop codon in the mRNA that would lead to a truncated protein, p.Arg8Phefs∗14. In order to prevent the translation of a truncated protein, nonsense-mediated mRNA decay can act to remove the affected transcript. In this way, a reduced amount of normal type-I collagen, which is only produced

by wild-type allele, might lead to haploinsufficiency in this patient. This condition is less deleterious for the patient's phenotype than producing a truncated protein that interferes with the normal activity of wild-type collagen. Segregation analysis and genotyping confirmed the presence of the same CNV in the patient's affected brother and father, who are both characterized by a very similar skeletal phenotype. All patients have severe spinal changes but normal BMD at the proximal hips.

A tandem duplication of around 12.5 kb within *PLS3*, chrX: 114,848,381–114,860,880, the first duplication ever reported in this gene, was identified in a male patient with early-onset osteoporosis. The identified change, starting in intron 2 and ending in intron 3 (reference sequence: NM_005032.6), is likely to lead to a frameshift in the ORF. The CNV and ed with the disease in the proband's family, in which one affected brother and the affected mother also carried the same CNV. The two affected brothers featured severe skeletal impairments, characterized by multiple long bone and vertebral fractures as well as low BMD. On the other hand, the mother, who also has a wild-type copy of the gene, featured a milder skeletal phenotype. So far, *PLS3* mutations in females have been shown to give rise to phenotypes of variable severity. Although the majority have a milder phenotype, some individuals have a severe phenotype resembling that observed in males [Van Dijk et al., 2013; Kampe et al., 2017]. A skewed X-inactivation could determine the phenotypic severity in female patients with *PLS3* mutations.

Our investigation on skeletal fragility using a custom designed array-CGH revealed two novel and pathogenic or likely pathogenic CNVs in *COL1A2* and *PLS3* in two families with early-onset osteoporosis. Since from the literature a larger number of small-scale mutations than CNVs has been reported in patients with skeletal fragility, we expected to identify a small number of CNVs in genes associated to bone fragility in our groups. Moreover, patients in Group 2 were prone to bone fractures but were otherwise healthy. Therefore, it is not surprising that no pathogenic CNVs were detected in this group. It could be speculated that fractures in these patients are a polygenic trait resulting from multiple SNPs with low effect size rather than from a single gene defect with a major impact. Epigenetic changes might also contribute to the patients' phenotype in this group.

4.2.2 CNVs in genes not yet associated with skeletal diseases

Our second aim in **paper II** was to search for rare CNVs in genes playing a role in cilia. By performing array-CGH in Groups 1-2 we identified 12 rare heterozygous CNVs.

Seven out of these variants located in intronic regions of genes that have not been associated with skeletal diseases. Intronic variants are unlikely to affect the function of a protein and therefore these CNVs were classified as likely benign.

On the other hand, five heterozygous CNVs were detected in coding regions of genes not related to skeletal diseases. Four out of these CNVs affected genes that are involved neither in bone nor in cilia (*ETV1-DGKB, SCN4A, RPS6KL1-PGF and ATM*) whereas one variant was identified in a gene with ciliary function, *AGBL2*. All these CNVs were regarded as variants of uncertain significance (VUSs).

The largest deletion identified in our study (about 1.6 Mb) affected the entire *ETV1* gene and exons 21-25 of *DGKB* (reference sequence: NM_004080.2) (Fig. 11).

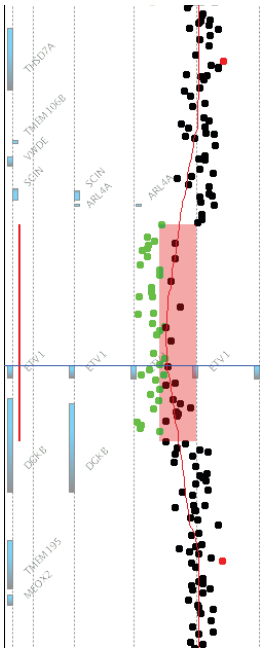


Figure 11. A large heterozygous deletion spanning the entire *ETV1* gene and partially *DKGB* identified by array-CGH.

ETV1 encodes a member of the ETS (E twenty-six) family of transcription factors. Gene fusions involving ETS transcription factors are responsible for different types of cancers, including a bone sarcoma named Ewing sarcoma (ETV1-EWS translocations) [Peter et al., 1997]. *DKGB* encodes a diacylglycerol kinase but its role in bone is not yet known. Although the detected deletion might potentially have an impact on bone fragility, no family samples were available for segregation analysis.

An intragenic deletion removing exons 15-17 of *SCN4A* (reference sequence: NM_000334.4) was identified in an 11-year-old male who had sustained three long bone fractures and had reduced BMD Z-score (-1.8 at lumbar spine) but did not fulfill the criteria for primary osteoporosis. Heterozygous *SCN4A* mutations are linked to myotonia congenita, which is characterized by painful muscle stiffness [Lerche et al., 1993]. Our

finding is unlikely to explain low BMD but muscle stiffness could potentially be the cause of the patient's falls and consequent bone fractures. For this reason, this variant has to be further investigated in the patient and his family.

The duplication that we identified in *ATM* (NM_000051.3), duplicating the last part of the gene starting from intron 62-63, was excluded as the cause of skeletal fragility in a fracture-prone child because homozygous *ATM* mutations underlie ataxia-telangiectasia [Savitsky et al., 1995; McConville et al., 1996].

Concerning the *RPS6KL1-PGF* deletion, removing exons 1-4 of *RPS6KL1* (NM_031464.4) and the entire *PGF* gene, we could not identify a direct link between the involved genes and bone homeostasis. *RPS6KL1* encodes the ribosomal protein S6 kinase. Despite the fact a GWAS hit has been found to BMD in another gene of the family of ribosomal protein kinases, *RPS6KA5* [Estrada et al., 2012], further studies are necessary to assess the significance of this CNV.

Finally, one deletion affected exons 5-7 of a gene belonging to the cilia proteome, *AGBL2* (NM_024783.3). Although we can not exclude a function of *AGBL2* in bone, a defect in tubulin deglutamylation was detected in the testes and sperm of a *Agbl2/Agbl3* double-knockout mice [Tort et al., 2014]. In this way, this deletion is unlikely to explain the skeletal phenotype of the studied patient.

Although the significance of these rare CNVs remains unclear and has to be further ascertained, the results indicate that custom designed array-CGH could potentially be a powerful method to detect novel gene-disease associations.

4.3 MPS TO SEARCH FOR DISEASE-CAUSING GENETIC DEFECTS IN RARE SKELETAL PHENOTYPES

In **papers IV-V** MPS was performed to find the genetic diagnosis in families with rare subtypes of skeletal dysplasia.

4.3.1 Studies in SMD with "corner fractures"

Paper IV started with the recruitment of a Finnish family with a rare form of SMD. The index patient (Patient #1), born from healthy non-consanguineous parents, featured short stature with severe metaphyseal changes in the long bones together with "corner fractures". "Corner fractures", which are not real fractures, are radiological findings appearing as lucent areas in the long bone metaphyses, especially in the tibias and femurs. At the time of initiation of our study, the causative gene in SMD with "corner fractures" (SMD-CF) was not known. In order to identify the genetic etiology of the disease

in the Finnish index family we carried out WGS. Through filtering variants according based on modes of inheritance, MAF < 0.001 and impact severity other than “LOW” in GEMINI one candidate variant was detected: a novel heterozygous variant in the fibronectin gene (*FN1*, reference sequence NM_212482.2), c.638G>A (p.Cys213Tyr), which was absent in the parents and thus determined as a *de novo* change (Fig. 12). *In silico* predictions classified this variant as probably damaging/deleterious.

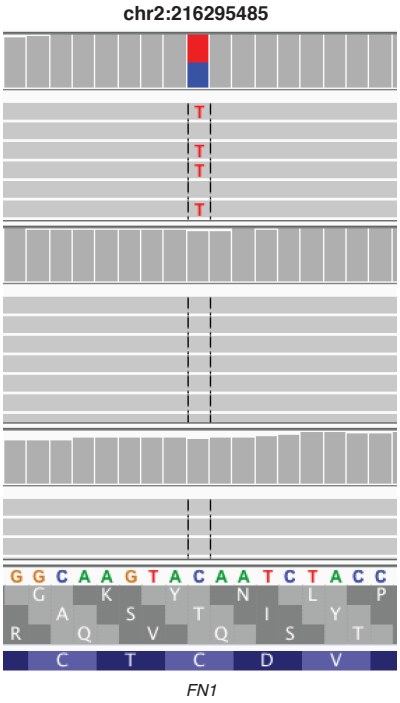


Figure 12. Snapshot from IGV showing the *de novo* c.638C>T (p.Cys213Tyr) variant in *FN1* identified in Patient #1.

In order to identify the genetic cause of disease in other four unrelated patients with a similar skeletal phenotype and “corner fractures” WES and Sanger sequencing were carried out in Patients #3-5 and Patient #2, respectively. Data analysis identified a novel *FN1* mutation in each patient: c.368G>A (p.Cys123Tyr), c.506G>A (p.Cys169Tyr), c.693C>G (p.Cys231Trp) and c.773G>A (p.Cys258Tyr). Half of these mutations were inherited from one affected parent whereas the other half were *de novo* variants.

FN1 mutations were linked to SMD-CF (MIM #184255) for the first time in 2017, during the course of our study, and only seven mutations had been reported before our study [Lee et al., 2017; Cadoff et al., 2018]. Prior to understanding the genetic etiology, SMD-CF was described as a separate clinical entity in a dozen of patients who lacked a genetic diagnosis [Langer et al., 1990]. Fibronectin is a glycoprotein that assembles into a highly organized matrix and partakes in cell and protein interactions within the bone ECM. This protein is a dimer composed by three types of modules, which organize into functional domain I-III that interact with other proteins within the ECM, including collagen, integrins and glycosaminoglycans [Potts and Campbell, 1994]. Interestingly, all five *FN1* mutations identified in our study cluster to type-I domains involved in ECM assembly and four out of

the five affected residues partake in disulfide bonds. Disulfide bridges are formed between the free sulfhydryl groups (-SH) of two cysteine residues that are in close proximity and determine the structure of fibronectin. Since the majority of the *FN1* mutations affect disulfide bridges, it could be argued that these changes do not only affect the overall shape and stability of the protein but they might also impair the ability of fibronectin to bind other proteins. Previous studies have shown that the mutated fibronectin is preferentially retained within the ER instead of being secreted [Lee et al., 2017; Cadoff et al., 2018]. Although functional validations were not performed, the disease mechanism is likely to be the same in our patients.

The most common features among the patients were corner fractures and short stature. Scoliosis was detected in three out of four patients. For the first time, bilateral femoral fractures and osteopenia were described in a patient with SMD-CF (Patient #1). Furthermore, some of the skeletal changes at the proximal metaphyses in the same patient resolved over time, similar to the case described by Cadoff et al [Cadoff et al., 2018].

Our study not only expands the still limited mutation spectrum of SMD-CF but also highlights the phenotypic variability of this disease.

4.3.2 Studies in SEMD

In **paper V** we investigated the genetic and molecular mechanism underlying a potentially novel form of SEMD in a Finnish trio, in which the index patient, a 4.5-year old boy with severe skeletal dysplasia and growth failure since birth (Patient #1), was the only affected subject. WGS analysis identified two candidate variants: a rare heterozygous missense variant in the *UBC* gene (reference sequence: NM_021009.6), c.2045G>A, (p.Arg682Lys) and 2) a novel heterozygous missense variant in the *RPL13* gene (reference sequence NM_000977.3), c.533C>A (p.Ala178Asp). *UBC* encodes ubiquitin C while *RPL13* encodes the ribosomal protein L13. Since the *UBC* change was predicted to be likely benign according to both SIFT and PolyPhen-2, the *RPL13* variant was regarded to be the most likely cause of the patient's skeletal disease.

In order to increase the power of our study and to confirm the pathogenic role of *RPL13* variants in this specific form of SEMD, we recruited two unrelated index patients (Patients #2-3) featuring the same skeletal phenotype as the Finnish patient. These children were born to unrelated Korean families. Sanger sequencing of *RPL13* in the Korean families led to the identification of two likely damaging mutations: Patient #2 inherited the same c.533C>A (p.Ala178Asp) change detected in Patient #1 from her affected mother; Patient #3 harbored instead a *de novo* missense mutation c.553G>C (p.Ala185Pro). Both

mutations were absent from the gnomAD database but two changes affecting the same codons, p.Ala178Val and p.Ala185Val, were previously reported in one and five individuals, respectively. Both alanine and valine are hydrophobic amino acids.

RPL13 is a protein of the large 60S subunit of the ribosome but its specific molecular function is not yet known. Since the two mutations identified in our study clustered only a few amino acids apart each other, this region in RPL13 might be essential for the protein to function and/or to correctly assemble into the ribosome.

All three index patients feature severe growth retardation and abnormal metaphyseal and epiphyseal changes, mainly affecting the growth plate, and delayed ossification of the SOCs. Other family members in Family #2 carry the same mutation and feature impairments at the same skeletal sites. However, the skeletal changes in these relatives were milder than in the index patient. Interestingly, the index patient's aunt, who carries the same mutation, does not exhibit the skeletal disease. Apparently, both non-penetrance and variable expressivity of the disease seem to be present in this family. It could be speculated that protective modifiers, which are alleles and/or variants in other genes, might prevent the skeletal disease to manifest, or to be notably milder in some mutation carriers. The effect of *RPL13* mutations can also be possibly compensated by increased expression of the wild-type allele (allelic compensation). Furthermore, lifestyle as well as environmental factors might additionally be involved in these mechanisms.

Mutations in other ribosomal proteins (RPs), ribosomal RNA or other components playing a role in ribosomes have been identified in congenital diseases, such as Diamond-Blackfan anemia [Draptchinskaia et al., 1999] and Shwachman-Diamond syndrome [Boocock et al., 2003]. All diseases that are known or suspected to be caused by ribosome dysfunction are collectively called ribosomopathies. Even if ribosomes are organelles present in all types of cells, most of the time, ribosomopathies are tissue-specific diseases, predominantly affecting the bone marrow-derived cell lineages and the skeletal tissues [Mills and Green, 2017]. Our patients feature severe skeletal impairments but unlike other patients with RP mutations, they have not developed hematological or immunological abnormalities thus far. Moreover, they do not feature any remarkable extra-skeletal impairments. It could therefore be assumed that the predominantly affected cell type is the growth plate chondrocyte. Although intact ribosome function is essential for all cells, it could be hypothesized that chondrocytes are vulnerable to changes in global mRNA translation rates as chondrocytes produce a large amount of ECM and for this reason they are likely to require a large concentration of ribosomes. Furthermore, it is also possible that the production and/or function of a certain protein playing a pivotal role in the growth plate is specifically affected by the identified *RPL13* mutations. In addition, ribosomopathies

manifest a broad variability in clinical manifestations even between subjects with the same RP mutation, as it was noticed in Patient's #2 family [Narla and Ebert, 2010].

Different models have been proposed to explain ribosomal dysfunction, including ribosome stress due to p53 activation, reduced global or specific mRNA translational efficiency and defects in ribosome assembly, but the specific mechanisms underlying ribosome dysfunction remain still a matter of debate [Xue and Barna, 2012; Mills and Green, 2017].

In order to validate our genetic findings, we performed functional studies both *in vitro* and *in vivo*. Our *in vitro* studies showed that there is no significant difference in RPL13 expression in the fibroblasts of patients (N= 4) and healthy controls (N= 4) suggesting that the protein is not degraded. Moreover, no significant difference in RPL13 localization in fibroblasts of patients (N= 3) as compared to fibroblasts of sex-matched healthy controls (N= 3). RPL13 is primarily distributed in the perinuclear region, corresponding to the ER (Fig. 13).

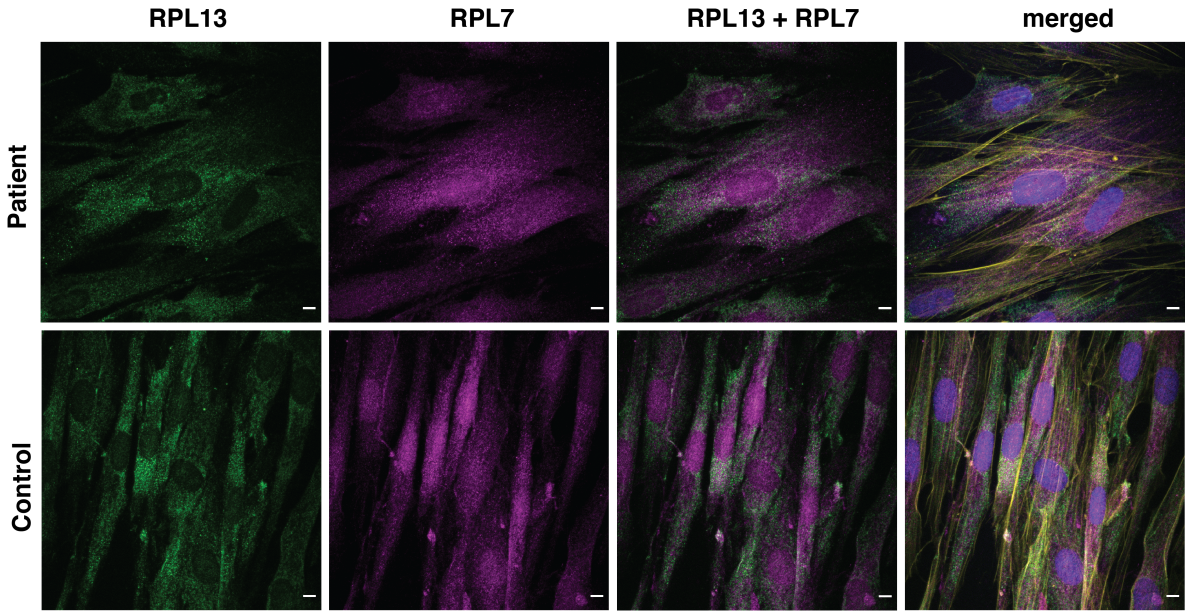


Figure 13. ICC of dermal fibroblasts from one patient and a healthy control showing staining for RPL13 (column 1), RPL7 (column 2), co-staining for RPL13 and RPL7 (column 3) and merged with structural dyes Hoechst and Phalloidin (column 4). Scale bars= 10 µm; patient= Patient #2, age 6 years; control 2= female, age 30 years.

Since all the identified mutations are missense, they are unlikely to lead to a reduced amount of protein expression due to proteasomal degradation and these results are therefore in line with what could be expected. In contrast, the mutations could impair the function of RPL13. In order to test the hypothesis that *RPL13* mutations affect the binding of the protein to the large subunit of the ribosome, we evaluated the colocalization of RPL13 with other two proteins of the 60S subunit, RPL7 and RPL28, in the cells of the

patients and controls. Surprisingly, we observed a significant increase in RPL13-RPL7 colocalization in the fibroblasts of patients compared to the fibroblasts of controls (Fig. 13).

To evaluate the effects of a *RPL13* mutations *in vivo*, a CRISPR/Cas9 mediated knock-out of the orthologue gene was generated in zebrafish, in order to target the region just downstream of the two mutations identified in our patients. Our preliminary results show that around 25% of larvae in F1 generation, derived from the mating of two heterozygous larvae harboring the same frameshift mutation c.571_577delCTTTTCG (p.Lys191Alafs*32) that is predicted to cause an elongation of the protein C terminal end of the protein, feature an abnormal phenotype. Although these larvae have yet to be genotyped, they are likely to correspond to the knocked out larvae. These abnormal larvae feature between 2 and 5 days post-fertilization (dpf) reduced body size, craniofacial defects and decreased pigmentation in the body. Whie pigmentation of the eyes is rescued at 5 dpf, the body still remains pale. Combined alizarin red and alcian blue staining show delayed cartilage formation and ossification in the head of the knockouts compared to wild-type fish at 5 dpf (Fig. 14). Concerning the reduced pigmentation, a previous study showed that RPL13 plays a role in melanocytes [Kardos et al., 2014]. Specifically, RPL13 silencing in melanoma cells inhibits cell viability [Kardos et al., 2014]. Rpl13 impairment could then explain the partial loss of pigmentation in our fish model.

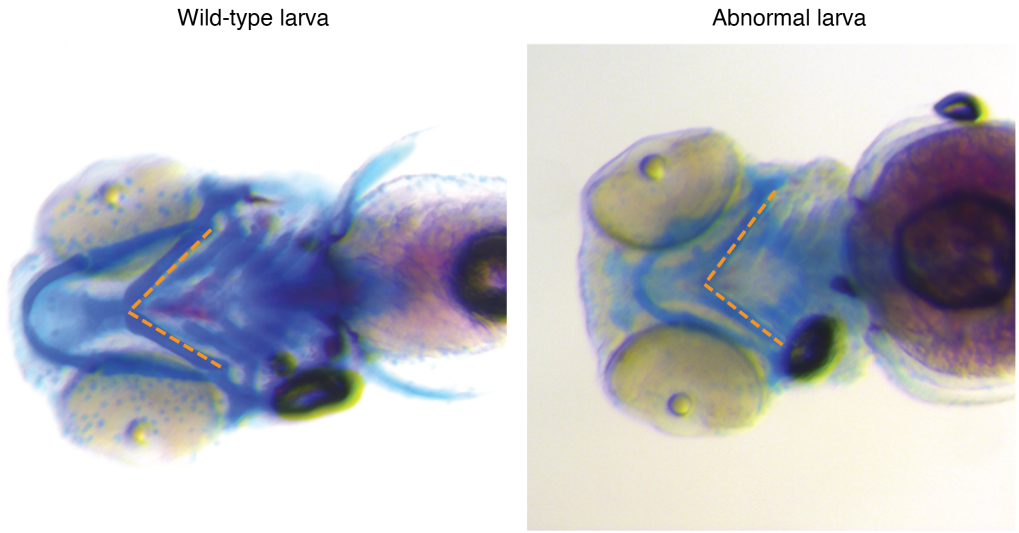


Figure 14. Head of a wild-type zebrafish larva (left) and an abnormal larva (right) at 5 dpf. Alcian blue staining as well as alizarin red staining are less intense in the abnormal fish, suggesting delayed cartilage formation and reduced mineralization. Craniofacial defects are also evident and the angle between the two ceratohyals (dashed lines) is wider than in the wild-type larva.

Although the functional studies exploring the specific molecular mechanisms leading to disease are still ongoing, our study is the first to report *RPL13* mutations to be associated with a novel subtype of SEMD.

4.4 STUDY LIMITATIONS

One limitation in this thesis concerns the small sample size of Groups 1-2 in **paper I**. A potential association between *CRTAP* variants and skeletal fragility might be missed due to the limited number of studied patients. Moreover, the disease in patients from Group 2, who are only mildly affected, might be a polygenic trait rather than a monogenic condition. In this way, a GWAS could be a more suitable approach to investigate genetic associations between SNPs and risk of fractures but would naturally require a significantly larger cohort of similarly affected subjects. Although most of the patients from Group 1-2 have also been investigated using array-CGH in **paper II**, small-scale mutations are overall more frequent than CNVs in patients with rare skeletal diseases. In order to increase the diagnostic rate, other types of mutations could be sought in these patients using different methods, preferentially MPS. Apart from the CNVs identified in the known OI or osteoporosis genes, the rare CNVs that were identified in our study have not been functionally validated. For this reason, the significance of most of these findings still remains unclear.

Concerning **paper III**, we identified novel genetic mutations in certain genes that are already known to play a role in skeletal diseases. Even if they can be classified as pathogenic according to the ACMG guidelines, functional studies might provide additional insights to the molecular mechanism underlying these diseases.

In **paper IV** a mutation in *FN1* was discovered in patient with SMD and “corner fractures” before even *FN1* was known as the causative gene for the disease. Despite the fact that some functional studies to demonstrate the pathogenicity of *FN1* mutations have been performed, experimental evidence showing how the different mutations contribute to disease in relevant cell types as well as *in vivo* is still lacking.

Functional studies in **paper V** are still ongoing. One of the major limitations in our study, as is usually the case in human bone diseases, is the lack of investigations on a human tissue that is predominantly affected. Although it is not possible to collect bone or cartilage / growth plate biopsies from the patients, it would be important to investigate the role of the two identified *RPL13* mutations in a human cell line/tissue other than the dermal fibroblasts.

4.5 ETHICAL CONSIDERATIONS

Performing research on patients with rare skeletal diseases offers the possibility to give a diagnosis to the families, to better understand the underlying pathogenesis and eventually develop efficient therapies.

Once the genetic cause of the disease is identified the parents can be informed about the risk of having another affected child. Estimating the recurrence risk for a genetic disease influences the clinical management of the condition as well as the reproductive choices of the parents of the affected child. Several pathogenic mutations identified in **papers III-V** are *de novo* variants and thus the recurrence risk given to the parents is significantly lower than for example when considering an autosomal recessive inheritance pattern. Although the recurrence risk for *de novo* mutations is in general low, the probability of having a second child with a new mutation can considerably increase depending on how the mutation has occurred. Most of *de novo* mutations have a paternal origin and the risk of new variants increase with paternal age. If a *de novo* mutation is transmitted by the mother or a young father the recurrence risk for autosomal dominant diseases is shown to be higher than the empirical risk of the disease [Campbell et al., 2014]. However, the parental origin of the *de novo* mutations identified in this thesis was not determined and for this reason an empirical risk might be given to the parents during genetic counseling.

Moreover, genetic investigations cannot always be carried out without any harm or potential risks to the studied subjects. For instance, in order to carefully evaluate the skeletal phenotype, DXA measurements were carried out in most of the patients included in this thesis. This investigation involves a very small dose of radiation. However, in most cases these evaluations were performed on clinical grounds and not solely for study purposes. Further, in addition to blood samples, small skin biopsies were collected from the families included in **paper V**. This is a minor procedure but more invasive than obtaining a blood sample and usually performed only for study purposes.

Concerning the risk of incidental findings, we did not identify any. However, **paper II** revealed some VUSs. Despite the fact that they might be benign, we can not exclude their contribution to the patients' phenotype.

Despite these disadvantages, several important scientific discoveries were made. From the clinical perspective, identifying the specific genetic cause of the child's condition enables a better management of the disease and offers the possibility of prenatal or even pre-implantation genetic diagnosis, which parents of a severely affected child may wish in future pregnancies.

5 CONCLUSIONS

The studies in this thesis have increased our knowledge of the phenotypic and genetic spectrum of some rare skeletal diseases. Novel genetic variants have been identified in patients with skeletal fragility, HBM and SMD. Moreover, a novel genetic form of SEMD caused by a defective ribosomal protein has been identified and is currently under further investigations. The findings identified in each paper are highlighted as follows:

Paper I

- Rare variants in *CRTAP* seem to be an unlikely cause of low BMD and/or fractures in young patients who lack other OI features

Paper II

- A novel deletion in *COL1A2* and a novel tandem duplication within *PLS3* were detected in two index patients with significant osteoporosis and their affected family members. These findings imply that screening of large scale mutations in genes related to skeletal fragility might be indicated if sequencing of the candidate genes does not reveal the genetic cause
- Five CNVs in coding regions of genes not previously associated with skeletal fragility (*ETV1-DGKB*, *AGBL2*, *ATM*, *RPS6KL1-PGF* and *SCN4A*) were identified in patients with susceptibility to fractures. These changes were determined as VUSs and their significance remains uncertain
- Our custom designed array-CGH might be a valid method to pinpoint novel candidate loci in early-onset skeletal fragility

Paper III

- Two novel mutations in *LRP5* and A*MER1* were identified as the underlying cause of HBM in two patients
- Although both mutations lead to HBM due to increased activation of the Wnt-β-catenin pathway the patients' clinical features appear different

Paper IV

- Five novel missense mutations in the fibronectin gene were detected in five patients with a rare subtype of SMD characterized by corner fractures
- All five mutations cluster in type-I domains and four out of five mutations affect disulfide bonds, suggesting that these regions might play an essential role in the pathogenesis of the disease

- Novel clinical features not previously associated with SMD-CF, including bilateral femoral fractures and osteopenia, were identified in one patient

Paper V

- Two novel mutations in the ribosomal protein L13 gene (*RPL13*) were identified to be associated with a novel subtype of SEMD in three unrelated families
- WB analysis and ICC experiments in dermal fibroblasts did not show a significant difference in RPL13 expression and localization in cells from patients compared to cells from healthy controls
- A significant increase in colocalization of RPL7 with RPL13 was detected in patient cells
- Our CRISPR-Cas9 knock-out zebrafish model showed that *rpl13* disruption is likely to cause an abnormal skeletal phenotype, characterized by reduced body length and delayed ossification

In addition, this thesis shows how the field of clinical genetics has moved from genetics to genomics together with the revolution in the sequencing and molecular technologies. From the screening of a single candidate gene by Sanger sequencing (**papers I** and **III**) our approach shifted to a genome-wide analysis using array-CGH (**paper II**) and MPS technologies (**papers IV-V**). MPS outperforms previous sequencing methods by offering the possibility of detecting small scale variants as well as CNVs in either a list of candidate genes (gene panels, **paper III**) or in the whole exome (**paper IV**) or genome (**papers IV-V**). Moreover, WGS also enables the detection of other structural changes, such as chromosomal translocations and inversions. Despite the fact that the cost of performing MPS has considerably decreased in the last years, it would still be too expensive to use this technology as a routine approach to investigate all patients with rare skeletal diseases.

Another aspect that is evident in this thesis is that research on rare skeletal diseases would not be possible without international collaborations. As it has been shown in **papers IV-V**, patients with the same extremely rare phenotype might be spread in different places around the world and only through an expert and collaborative research network it is possible to group them together and pinpoint the genetic defect underlying their condition. In fact, analyzing more than one family with the rare disease increases the likelihood of identifying the genetic cause of the disease and understanding its pathophysiology. No genetic studies on rare bone diseases could be successfully carried out without reliable

phenotyping and therefore collaborative networks between clinicians and basic scientists are essential.

In summary, this PhD thesis has led to the identification of novel mutations in genes already known to underlie rare bone diseases and identified a novel form of skeletal dysplasia. These outcomes expand our understanding of the genetic and molecular mechanisms underlying bone fragility and skeletal dysplasia. The delineated novel subtype of SEMD expands the genic scenario underlying skeletal dysplasias. Despite the fact that functional studies to understand the specific mechanisms connecting *RPL13* mutations to SEMD are still ongoing, this discovery is likely to give promising results.

Last but not least our findings will be of great importance during genetic counseling and for the management of the disease. Moreover, our research provides means to further investigations on rare skeletal diseases that might eventually lead to the development of targeted therapies for the patients affected by these conditions.

6 FUTURE PERSPECTIVES

Future studies will need to be performed to increase the diagnostic yield in patients with rare skeletal diseases and to better elucidate the pathomolecular mechanisms underlying these conditions.

First, we would like to use our custom made array-CGH in larger cohorts to identify novel CNVs associated to skeletal fragility and to proof the efficiency of this method. Moreover, the VUSs detected in our patients with skeletal fragility must be validated.

Concerning our findings on *FN1* mutations in SMD-CF, functional studies are needed to elucidate the role of fibronectin in bone. Since most of the identified mutations affect disulfide bonds, we will continue studies together with the collaborators to understand how the mutations lead to this specific phenotype.

Moreover, the investigation of the RPL13 involvement in SEMD will be carried on for better understanding the mechanisms leading to ribosome dysfunction and possibly diagnose the disease in other patients with rare *RPL13* mutations. The primary goal is to genotype the zebrafish larvae in order to confirm that the ones showing an abnormal phenotype are those carrying the *rpl13* frameshift mutation in homozygosity. In addition, the altered co-localization of RPL7 with RPL13 requires further investigations. We will continue exploring the pathomolecular mechanisms leading to disease by checking the translational efficiency by sucrose gradient and polysome profiling and by measuring the extent of new protein synthesis in fibroblasts of patients and controls, respectively. Moreover, cell cycle analysis will be carried out to understand if *RPL13* mutations cause cell cycle arrest due to p53 activation.

In addition, several interesting skeletal phenotypes still remain unexplained in our research cohorts. Using the existing clinical and research networks we have planned to recruit additional families and continue our search for novel genetic forms of skeletal dysplasia using MPS. The current success rate of WES/WGS in rare Mendelian diseases is approximately 25-50% [Cummings et al., 2017]. In order to improve the diagnostic yield we are aiming to increase our knowledge in interpreting the functional and clinical impact of variants in non-coding regions and variants of difficult interpretations, including e.g. synonymous variants causing deleterious effects on the splicing mechanisms, possibly by combining RNA-sequencing with WGS. It has already been shown that non-coding variants might be responsible for some monogenic skeletal diseases. As two examples, biallelic variants in a long non-coding RNA, named RMRP, have been found to underlie cartilage-hair hypoplasia whereas a neomorphic mutation in the microRNA-140 has

recently been associated to novel rare subtype of skeletal dysplasia [Ridanpaa et al., 2001; Grigelioniene et al., 2019].

Finally, due to the genetic and phenotypic heterogeneity of rare skeletal diseases, personalized treatments would be needed. The ultimate goal of our research on rare skeletal disease is to apply the gained knowledge to develop targeted strategies to make the future of our patients brighter.

7 ACKNOWLEDGMENTS

This journey could not have been as great as it was without the guidance and support of several people around me. I would like to start acknowledging my supervisors.

Outi Mäkitie, you have been a guide for me during my entire PhD. You have taught me many things and I grew a lot as a researcher thanks to you. You are a clear example of the fact that if you want to be a successful and worldwide respected researcher you must, first of all, be a good person. With your calm attitude you have always supported and motived me and made me passionate about doing research in the field of rare skeletal diseases. I feel very lucky to be part of your team and I could not wish for a better supervisor.

Magnus Nordenskjöld, you are the master of clinical genetics. I am proud to be one of your PhD students. Many thanks for your guidance and humor during these four years. I also appreciate your effort of trying to encourage me to speak Svenska.

Fulya Taylan, you have been a good supervisor and also a good friend. I admire your broad and deep knowledge not only about science but also about general facts. I have learnt a lot from you.

Hong Jiao, you are the first person who guided me in the 'scary' field of bioinformatics. Thank you for teaching me a lot and for supporting me.

Olle Kämpe, my research benefited from your constructive comments during our meetings.

Now I would like to thank the other members of my group. **Anders**, you have an endless patience in explaining me bioinformatics-related matters. You have always been kind and willing to help me. **Jessica**, since you joined our group I felt that we became a perfect team. You have helped me so much with the lab work, with the writing and offered me moral support, in particular during these last months. I am very grateful to you. I would also like to acknowledge **Noor** and **Nadi** for their help with my research. A big thank-you goes also to our team in Helsinki, to **Riikka** in particular, for her invaluable contributions to my research projects and for always welcoming me in Helsinki, to **Helena Valta**, **Petra**, **Ilkka**, **Mervi**, **Sanna**, **Minna P** and **Mira**.

I owe a special thank-you also to other researchers that helped me during my PhD: **Richard Rosenquist Brandell**, for your great suggestions on my last project and **Anna Lindstrand**, for your precious help with my array-CGH study.

I am also grateful to our collaborators, **Antonella Forlino** in particular, and her research group, who hosted me for four months in Pavia, Italy: **Fra**, **Nadia**, **Laura**, **Rossella**, **Gabriella**, **Valentina** and **Chiara**. Many thanks also to **Anh** and **Tuomas**: my

immunofluorescence experiments would not have looked so excellent without your help. I would also like to thank **Sini**, **Minna Männikkö**, **Artemis Doulgeraki**, **Shan**, **Ola Larsson**, **Carina Wallgren-Pettersson**, **Andrei Chagin**, **Philippe M. Campeau** and **Tae-Joon Cho**.

The work environment would not have been so positive and friendly without the members of the Clinical Genetics and Rare Diseases groups. My great friends **Bianca** and **Raquel**, you have always been there for me and I would not have laugh as much as I did without you. **Nina**, thanks for always cheering me up and for offering your help whenever I needed it. **Giedre Grigelioniene**, I am grateful to you for first introducing me to the Clinical Genetics group at KI when I was an Erasmus student. **Anna H**, thanks for our funny chats. **Jesper**, your optimism is contagious and your help with bioinformatics priceless. **Nikos,** you are a good listener and a good friend. **Diego** you brought a good taste of Italian humor to the group. Many thanks also to all the other colleagues: **Jessada**, **Maria P**, **Samina**, **Isabel**, **Mastoureh**, **Ann Nordgren**, **Vasilios**, **Ellika**, **Clara**, **Sofia**, **Benedicte**, **Wolfgang**, **Aron**, **Ingegerd**, **Liselotte**, **Anaya** and all the ones I might have forgotten.

Thanks also to our ex-colleagues and friends at CMM especially: **Elisa** for your support, **Ankit**, **Bo**, **Eva**, **Helena**, **Winnie**, **Sissi** and **Kunal**.

During my PhD at KI I had the opportunity to make many friends. A big thank-you goes to my multicultural “Ciao bellas” group: **Mirco**, you are one of the best people I know and your delicious food made me feel less homesick. **Lorenzo** and **Gabriele**, thanks to you I did not forget my origins “venete”. Due to space limit I must just mention the rest of the amazing party & life companions: **Joep**, **Maria T**, **Marco**, **Harpa**, **Anna Huguet**, **Jemina**, **Nuria**, **Esther**, **Alek** and **Vojta**.

I am also very thankful to my Italian friends in Stockholm, especially to **Marta**. You are an amazing and caring friend. You support is essential for me and our friendship is invaluable. **Giulia**, **Claudio** and **Daniele** you are such good, down to earth and funny people. I really enjoy spending time with you.

It is now time to thank a group of people that had to deal with me for many hours per day, including evenings and weekends: my lovely flatmates/friends! **Miriam** you have always being on my side both in the lab and outside the lab. **Franci** (and **Fred**): thanks for the nice evenings and for always hosting me at your place. **Maria Peleli**, in the year we lived together we cultivated such a strong friendship that its roots still go all the way to Greece. I will always bring with me all our happy memories together.

My life in Sweden would not have been so fun without my “FyFa family”, who ‘adopted’ me since day one: **Shane**, we basically started and ended our PhD journey together. You are

such a funny guy and one of my dearest friend. Thanks also to the other FyFa and FyFa-related buddies: **Alessia**, **Chiara**, **Jana**, **Phil**, **Steph**, **Catherine**, **Marin**, **Hannah**, **Thibo**, **Marilena**, **Christa**, **Andre**, **Sabine**, **Henna**, **Luca**, **Giacomo**, **Björn**, **Philip** and **Rubin**. I am also grateful to **Mattias Cärlström** for being my PhD mentor.

My best friends from Italy, “collegiali spertesi”: **Elena**, **Gio**, **Stefy** and **Dally**. We have been friends for over 20 years (29, with some of you). We are proof that a real friendship can survive the long distance.

Finally, I would like to thank my family, who is my strength. My parents, **Lena** and **Lauro**, and my brother, **Mett**, for always believing in me. Despite the physical distance you have always been with me in this journey. Without your love, encouragement and sacrifice I would not be where I am now. All my cousins and relatives, **Katia**, **Paolo** and **Linda** in particular, who travelled all the way to Stockholm to be with me on my “big day”. Last but not least **Ale**, there are no proper words to describe how important you are for me. I can just say that you always supported my decisions and have been making me as happy as I could be since 10+ years.

8 REFERENCES

Acharya C, Yik JHN, Kishore A, Van Dinh V, Di Cesare PE, Haudenschild DR, Cartilage oligomeric matrix protein and its binding partners in the cartilage extracellular matrix: Interaction, regulation and role in chondrogenesis, Matrix Biol. 37 (2014) 102-111.

Adzhubei I, Jordan DM, Sunyaev SR, Predicting functional effect of human missense mutations using PolyPhen-2, Curr Protoc Hum Genet Chapter 7 (2013) Unit7 20.

Alanay Y, Avaygan H, Camacho N, Utine GE, Boduroglu K, Aktas D, Alikasifoglu M, Tuncbilek E, Orhan D, Bakar FT, Zabel B, Superti-Furga A, Bruckner-Tuderman L, Curry CJ, Pyott S, Byers PH, Eyre DR, Baldridge D, Lee B, Merrill AE, et al., Mutations in the gene encoding the RER protein FKBP65 cause autosomal-recessive osteogenesis imperfecta, Am J Hum Genet 86(4) (2010) 551-9.

Ameur A, Dahlberg J, Olason P, Vezzi F, Karlsson R, Martin M, Viklund J, Kahari AK, Lundin P, Che H, Thutkawkorapin J, Eisfeldt J, Lampa S, Dahlberg M, Hagberg J, Jareborg N, Liljedahl U, Jonasson I, Johansson A, Feuk L, et al., SweGen: a whole-genome data resource of genetic variability in a cross-section of the Swedish population, Eur J Hum Genet 25(11) (2017) 1253-1260.

Anvarian Z, Mykytyn K, Mukhopadhyay S, Pedersen LB, Christensen ST, Cellular signalling by primary cilia in development, organ function and disease, Nat Rev Nephrol (2019).

Bardai G, Moffatt P, Glorieux FH, Rauch F, DNA sequence analysis in 598 individuals with a clinical diagnosis of osteogenesis imperfecta: diagnostic yield and mutation spectrum, Osteoporos Int 27(12) (2016) 3607-3613.

Barnes AM, Cabral WA, Weis M, Makareeva E, Mertz EL, Leikin S, Eyre D, Trujillo C, Marini JC, Absence of FKBP10 in recessive type XI osteogenesis imperfecta leads to diminished collagen cross-linking and reduced collagen deposition in extracellular matrix, Hum Mutat 33(11) (2012) 1589-98.

Baron R, Kneissel M, WNT signaling in bone homeostasis and disease: from human mutations to treatments, Nat Med 19(2) (2013) 179-92.

Bekker PJ, Holloway DL, Rasmussen AS, Murphy R, Martin SW, Leese PT, Holmes GB, Dunstan CR, DePaoli AM, A Single-Dose Placebo-Controlled Study of AMG 162, a Fully Human Monoclonal Antibody to RANKL, in Postmenopausal Women, J Bone Miner Res 19(7) (2004) 1059-1066.

Bilezikian JP, Primer on the metabolic bone diseases and disorders of mineral metabolism, Ninth edition / editor-in-Chief, John P. Bilezikian, MD, PhD (hon) ; senior associate editors, Roger Bouillon, MD, PhD, FRCP, Tom Clemens, PhD, Juliet Compston, MD, FRCP, FRCPath, FMedSci ; associate editors, Doug Bauer, MD and nine others . ed., Wiley-Blackwell, Hoboken, NJ, 2019.

Bishop N, Adami S, Ahmed SF, Antón J, Arundel P, Burren CP, Devogelaer J-P, Hangartner T, Hosszú E, Lane JM, Lorenc R, Mäkitie O, Munns CF, Paredes A, Pavlov H, Plotkin H, Raggio CL, Reyes ML, Schoenau E, Semler O, et al., Risedronate in children with osteogenesis imperfecta: a randomised, double-blind, placebo-controlled trial, Lancet 382(9902) (2013) 1424-1432.

Black DM, Delmas PD, Eastell R, Reid IR, Boonen S, Cauley JA, Cosman F, Lakatos P, Leung PC, Man Z, Mautalen C, Mesenbrink P, Hu H, Caminis J, Tong K, Rosario-Jansen T, Krasnow J, Hue TF, Sellmeyer D, Eriksen EF, et al., Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis, N Engl J Med 356(18) (2007) 1809-22.

Bonafe L, Cormier-Daire V, Hall C, Lachman R, Mortier G, Mundlos S, Nishimura G, Sangiorgi L, Savarirayan R, Sillence D, Spranger J, Superti-Furga A, Warman M, Unger S, Nosology and classification of genetic skeletal disorders: 2015 revision, Am J Med Genet A 167A(12) (2015) 2869-92.

Bone HG, Hosking D, Devogelaer JP, Tucci JR, Emkey RD, Tonino RP, Rodriguez-Portales JA, Downs RW, Gupta J, Santora AC, Liberman UA, Alendronate Phase IIIOTSG, Ten years' experience with alendronate for osteoporosis in postmenopausal women, N Engl J Med 350(12) (2004) 1189-99.

Bonewald LF, The amazing osteocyte, J Bone Miner Res 26(2) (2011) 229-38.

Boocock GR, Morrison JA, Popovic M, Richards N, Ellis L, Durie PR, Rommens JM, Mutations in SBDS are associated with Shwachman-Diamond syndrome, Nat Genet 33(1) (2003) 97-101.

Bourhis E, Wang WR, Tam C, Hwang JY, Zhang YN, Spittler D, Huang OW, Gong Y, Estevez A, Zilberleyb I, Rouge L, Chiu C, Wu Y, Costa M, Hannoush RN, Franke Y, Cochran AG, Wnt Antagonists Bind through a Short Peptide to the First beta-Propeller Domain of LRP5/6, Structure 19(10) (2011) 1433-1442.

Brighton CT, Structure and function of the growth plate, Clin Orthop Relat Res (136) (1978) 22-32.

Brown EM, Clinical lessons from the calcium-sensing receptor, Nat Clin Pract Endocrinol Metab 3(2) (2007) 122-133.

Cabral WA, Chang W, Barnes AM, Weis M, Scott MA, Leikin S, Makareeva E, Kuznetsova NV, Rosenbaum KN, Tifft CJ, Bulas DI, Kozma C, Smith PA, Eyre DR, Marini JC, Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta, Nat Genet 39(3) (2007) 359-65.

Cadoff EB, Sheffer R, Wientroub S, Ovadia D, Meiner V, Schwarzbauer JE, Mechanistic insights into the cellular effects of a novel FN1 variant associated with a spondylometaphyseal dysplasia, Clin Genet 94(5) (2018) 429-437.

Campbell Ian M, Stewart Jonathan R, James Regis A, Lupski James R, Stankiewicz P, Olofsson P, Shaw Chad A, Parent of Origin, Mosaicism, and Recurrence Risk: Probabilistic Modeling Explains the Broken Symmetry of Transmission Genetics, Am J Hum Genet 95(4) (2014) 345-359.

Carpenter TO, Whyte MP, Imel EA, Boot AM, Hogler W, Linglart A, Padidela R, Van't Hoff W, Mao M, Chen CY, Skrinar A, Kakkis E, San Martin J, Portale AA, Burosumab Therapy in Children with X-Linked Hypophosphatemia, N Engl J Med 378(21) (2018) 1987-1998.

Chang HR, Cho SY, Lee JH, Lee E, Seo J, Lee HR, Cavalcanti DP, Makitie O, Valta H, Girisha KM, Lee C, Neethukrishna K, Bhavani GS, Shukla A, Nampoothiri S, Phadke SR, Park MJ, Ikegawa S, Wang Z, Higgs MR, et al., Hypomorphic Mutations in TONSL Cause SPONASTRIME Dysplasia, Am J Hum Genet 104(3) (2019) 439-453.

Clemens TL, Karsenty G, The osteoblast: an insulin target cell controlling glucose homeostasis, J Bone Miner Res 26(4) (2011) 677-80.

Clevers H, Wnt/β-Catenin Signaling in Development and Disease, Cell 127(3) (2006) 469-480.

Cosman F, Crittenden DB, Adachi JD, Binkley N, Czerwinski E, Ferrari S, Hofbauer LC, Lau E, Lewiecki EM, Miyauchi A, Zerbini CA, Milmont CE, Chen L, Maddox J, Meisner PD, Libanati C, Grauer A, Romosozumab Treatment in Postmenopausal Women with Osteoporosis, N Engl J Med 375(16) (2016) 1532-1543.

Costantini A, Makitie O, Value of rare low bone mass diseases for osteoporosis genetics, Bonekey Rep 5 (2016) 773.

Costantini A, Tournis S, Kampe A, Ul Ain N, Taylan F, Doulgeraki A, Makitie O, Autosomal Recessive Osteogenesis Imperfecta Caused by a Novel Homozygous COL1A2 Mutation, Calcif Tissue Int 103(3) (2018) 353-358.

Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, Bolduc V, Waddell LB, Sandaradura SA, O'Grady GL, Estrella E, Reddy HM, Zhao F, Weisburd B, Karczewski KJ, O'Donnell-Luria AH, Birnbaum D, Sarkozy A, Hu Y, Gonorazky H, et al., Improving genetic diagnosis in Mendelian disease with transcriptome sequencing, Sci Transl Med 9(386) (2017).

Dallas SL, Prideaux M, Bonewald LF, The osteocyte: an endocrine cell ... and more, Endocrine reviews 34(5) (2013) 658-90.

Dao DY, Jonason JH, Zhang Y, Hsu W, Chen D, Hilton MJ, O'Keefe RJ, Cartilage-specific beta-catenin signaling regulates chondrocyte maturation, generation of ossification centers, and perichondrial bone formation during skeletal development, J Bone Miner Res 27(8) (2012) 1680-94.

Day TF, Guo X, Garrett-Beal L, Yang Y, Wnt/beta-catenin signaling in mesenchymal progenitors controls osteoblast and chondrocyte differentiation during vertebrate skeletogenesis, Dev Cell 8(5) (2005) 739-50.

Del Fattore A, Teti A, Rucci N, Bone cells and the mechanisms of bone remodelling, Front Biosci (Elite Ed) 4 (2012) 2302-21.

Doudna JA, Charpentier E, Genome editing. The new frontier of genome engineering with CRISPR-Cas9, Science 346(6213) (2014) 1258096.

Doyard M, Bacrot S, Huber C, Di Rocco M, Goldenberg A, Aglan MS, Brunelle P, Temtamy S, Michot C, Otaify GA, Haudry C, Castanet M, Leroux J, Bonnefont JP, Munnich A, Baujat G, Lapunzina P, Monnot S, Ruiz-Perez VL, Cormier-Daire V, FAM46A mutations are responsible for autosomal recessive osteogenesis imperfecta, J Med Genet 55(4) (2018) 278-284.

Draptchinskaia N, Gustavsson P, Andersson B, Pettersson M, Willig TN, Dianzani I, Ball S, Tchernia G, Klar J, Matsson H, Tentler D, Mohandas N, Carlsson B, Dahl N, The gene encoding ribosomal protein S19 is mutated in Diamond-Blackfan anaemia, Nat Genet 21(2) (1999) 169-175.

Dunn KW, Kamocka MM, McDonald JH, A practical guide to evaluating colocalization in biological microscopy, Am J Physiol Cell Physiol 300(4) (2011) C723-42.

Estrada K, Styrkarsdottir U, Evangelou E, Hsu YH, Duncan EL, Ntzani EE, Oei L, Albagha OM, Amin N, Kemp JP, Koller DL, Li G, Liu CT, Minster RL, Moayyeri A, Vandenput L, Willner D, Xiao SM, Yerges-Armstrong LM, Zheng HF, et al., Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture, Nat Genet 44(5) (2012) 491-501.

Fiedler IAK, Schmidt FN, Wolfel EM, Plumeyer C, Milovanovic P, Gioia R, Tonelli F, Bale HA, Jahn K, Besio R, Forlino A, Busse B, Severely Impaired Bone Material Quality in Chihuahua Zebrafish Resembles Classical Dominant Human Osteogenesis Imperfecta, J Bone Miner Res 33(8) (2018) 1489-1499.

Florencio-Silva R, Sasso GRD, Sasso-Cerri E, Simoes MJ, Cerri PS, Biology of Bone Tissue: Structure, Function, and Factors That Influence Bone Cells, Biomed Res Int (2015).

Forlino A, Marini JC, Osteogenesis imperfecta, Lancet 387(10028) (2016) 1657-1671.

Fraser DR, Kodicek E, Regulation of 25-hydroxycholecalciferol-1-hydroxylase activity in kidney by parathyroid hormone, Nat New Biol 241(110) (1973) 163-6.

Fukumoto S, Martin TJ, Bone as an endocrine organ, Trends Endocrinol Metab 20(5) (2009) 230-6.

Genomes Project C, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA, A map of human genome variation from population-scale sequencing, Nature 467(7319) (2010) 1061-73.

Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA, An integrated map of genetic variation from 1,092 human genomes, Nature 491(7422) (2012) 56-65.

Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR, A global reference for human genetic variation, Nature 526(7571) (2015) 68-74.

Gioia R, Tonelli F, Ceppi I, Biggiogera M, Leikin S, Fisher S, Tenedini E, Yorgan TA, Schinke T, Tian K, Schwartz JM, Forte F, Wagener R, Villani S, Rossi A, Forlino A, The chaperone activity of 4PBA ameliorates the skeletal phenotype of Chihuahua, a zebrafish model for dominant osteogenesis imperfecta, Hum Mol Genet 26(15) (2017) 2897-2911.

Gistelinck C, Gioia R, Gagliardi A, Tonelli F, Marchese L, Bianchi L, Landi C, Bini L, Huysseune A, Witten PE, Staes A, Gevaert K, De Rocker N, Menten B, Malfait F, Leikin S, Carra S, Tenni R, Rossi A, De Paepe A, et al., Zebrafish Collagen Type I: Molecular and Biochemical Characterization of the Major Structural Protein in Bone and Skin, Sci Rep 6 (2016) 21540.

Gistelinck C, Kwon RY, Malfait F, Symoens S, Harris MP, Henke K, Hawkins MB, Fisher S, Sips P, Guillemyn B, Bek JW, Vermassen P, De Saffel H, Witten PE, Weis M, De Paepe A, Eyre DR, Willaert A, Coucke PJ, Zebrafish type I collagen mutants faithfully recapitulate human type I collagenopathies, Proc Natl Acad Sci U S A 115(34) (2018) E8037-E8046.

Goetz SC, Anderson KV, The primary cilium: a signalling centre during vertebrate development, Nat Rev Genet 11(5) (2010) 331-44.

Gong Y, Slee RB, Fukai N, Rawadi G, Roman-Roman S, Reginato AM, Wang H, Cundy T, Glorieux FH, Lev D, Zacharin M, Oexle K, Marcelino J, Suwairi W, Heeger S, Sabatakos G,

Apte S, Adkins WN, Allgrove J, Arslan-Kirchner M, et al., LDL receptor-related protein 5 (LRP5) affects bone accrual and eye development, Cell 107(4) (2001) 513-23.

Grafe I, Yang T, Alexander S, Homan EP, Lietman C, Jiang MM, Bertin T, Munivez E, Chen Y, Dawson B, Ishikawa Y, Weis MA, Sampath TK, Ambrose C, Eyre D, Bachinger HP, Lee B, Excessive transforming growth factor-beta signaling is a common mechanism in osteogenesis imperfecta, Nat Med 20(6) (2014) 670-5.

Gregson CL, Poole KE, McCloskey EV, Duncan EL, Rittweger J, Fraser WD, Smith GD, Tobias JH, Elevated circulating Sclerostin concentrations in individuals with high bone mass, with and without LRP5 mutations, J Clin Endocrinol Metab 99(8) (2014) 2897-907.

Grigelioniene G, Suzuki HI, Taylan F, Mirzamohammadi F, Borochowitz ZU, Ayturk UM, Tzur S, Horemuzova E, Lindstrand A, Weis MA, Grigelionis G, Hammarsjo A, Marsk E, Nordgren A, Nordenskjold M, Eyre DR, Warman ML, Nishimura G, Sharp PA, Kobayashi T, Gain-of-function mutation of microRNA-140 in human skeletal dysplasia, Nat Med (2019).

Ha-Vinh R, Alanay Y, Bank RA, Campos-Xavier AB, Zankl A, Superti-Furga A, Bonafe L, Phenotypic and molecular characterization of Bruck syndrome (osteogenesis imperfecta with contractures of the large joints) caused by a recessive mutation in PLOD2, Am J Med Genet A 131(2) (2004) 115-20.

Hald JD, Evangelou E, Langdahl BL, Ralston SH, Bisphosphonates for the prevention of fractures in osteogenesis imperfecta: meta-analysis of placebo-controlled trials, J Bone Miner Res 30(5) (2015) 929-33.

Hall BK, Miyake T, All for one and one for all: condensations and the initiation of skeletal development, Bioessays 22(2) (2000) 138-147.

Hammarsjö A, Wang Z, Vaz R, Taylan F, Sedghi M, Girisha KM, Chitayat D, Neethukrishna K, Shannon P, Godoy R, Gowrishankar K, Lindstrand A, Nasiri J, Baktashian M, Newton PT, Guo L, Hofmeister W, Pettersson M, Chagin AS, Nishimura G, et al., Novel KIAA0753 mutations extend the phenotype of skeletal ciliopathies, Sci Rep 7(1) (2017).

Hartikka H, Makitie O, Mannikko M, Doria AS, Daneman A, Cole WG, Ala-Kokko L, Sochett EB, Heterozygous mutations in the LDL receptor-related protein 5 (LRP5) gene are associated with primary osteoporosis in children, J Bone Miner Res 20(5) (2005) 783-9.

Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L, McLaren S, Sealy I, Caccamo M, Churcher C, Scott C, Barrett JC, Koch R, Rauch GJ, White S, Chow W, et al., The zebrafish reference genome sequence and its relationship to the human genome, Nature 496(7446) (2013) 498-503.

Huber C, Cormier-Daire V, Ciliary disorder of the skeleton, Am J Med Genet C Semin Med Genet 160C(3) (2012) 165-74.

Ikebuchi Y, Aoki S, Honma M, Hayashi M, Sugamori Y, Khan M, Kariya Y, Kato G, Tabata Y, Penninger JM, Udagawa N, Aoki K, Suzuki H, Coupling of bone resorption and formation by RANKL reverse signalling, Nature 561(7722) (2018) 195-200.

Jarvik GP, Browning BL, Consideration of Cosegregation in the Pathogenicity Classification of Genomic Variants, Am J Hum Genet 98(6) (2016) 1077-1081.

Jelin AC, O'Hare E, Blakemore K, Jelin EB, Valle D, Hoover-Fong J, Skeletal Dysplasias: Growing Therapy for Growing Bones, Front Pharmacol 8 (2017) 79.

Jenkins ZA, van Kogelenberg M, Morgan T, Jeffs A, Fukuzawa R, Pearl E, Thaller C, Hing AV, Porteous ME, Garcia-Minaur S, Bohring A, Lacombe D, Stewart F, Fiskerstrand T, Bindoff L, Berland S, Ades LC, Tchan M, David A, Wilson LC, et al., Germline mutations in WTX cause a sclerosing skeletal dysplasia but do not predispose to tumorigenesis, Nat Genet 41(1) (2009) 95-100.

Jian X, Boerwinkle E, Liu X, In silico tools for splicing defect prediction: a survey from the viewpoint of end users, Genet Med 16(7) (2014) 497-503.

Jiang X, O'Reilly PF, Aschard H, Hsu YH, Richards JB, Dupuis J, Ingelsson E, Karasik D, Pilz S, Berry D, Kestenbaum B, Zheng J, Luan J, Sofianopoulou E, Streeten EA, Albanes D, Lutsey PL, Yao L, Tang W, Econs MJ, et al., Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels, Nat Commun 9(1) (2018) 260.

Joiner DM, Ke J, Zhong Z, Xu HE, Williams BO, LRP5 and LRP6 in development and disease, Trends Endocrinol Metab 24(1) (2013) 31-9.

Kampe AJ, Costantini A, Levy-Shraga Y, Zeitlin L, Roschger P, Taylan F, Lindstrand A, Paschalis EP, Gamsjaeger S, Raas-Rothschild A, Hovel M, Jiao H, Klaushofer K, Grasemann C, Makitie O, PLS3 Deletions Lead to Severe Spinal Osteoporosis and Disturbed Bone Matrix Mineralization, J Bone Miner Res 32(12) (2017) 2394-2404.

Kampe AJ, Costantini A, Makitie RE, Jantti N, Valta H, Mayranpaa M, Kroger H, Pekkinen M, Taylan F, Jiao H, Makitie O, PLS3 sequencing in childhood-onset primary osteoporosis identifies two novel disease-causing variants, Osteoporos Int 28(10) (2017) 3023-3032.

Kampe AJ, Makitie RE, Makitie O, New Genetic Forms of Childhood-Onset Primary Osteoporosis, Horm Res Paediatr 84(6) (2015) 361-9.

Kardos GR, Dai MS, Robertson GP, Growth inhibitory effects of large subunit ribosomal proteins in melanoma, Pigment Cell Melanoma Res 27(5) (2014) 801-12.

Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST, Working Group of the American College of Medical Genetics Laboratory Quality Assurance C, American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants, Genet Med 13(7) (2011) 680-5.

Kemp JP, Morris JA, Medina-Gomez C, Forgetta V, Warrington NM, Youlten SE, Zheng J, Gregson CL, Grundberg E, Trajanoska K, Logan JG, Pollard AS, Sparkes PC, Ghirardello EJ, Allen R, Leitch VD, Butterfield NC, Komla-Ebri D, Adoum AT, Curry KF, et al., Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis, Nat Genet 49(10) (2017) 1468-1475.

Kennell JA, MacDougald OA, Wnt signaling inhibits adipogenesis through beta-catenin-dependent and -independent mechanisms, J Biol Chem 280(25) (2005) 24004-10.

Keupp K, Beleggia F, Kayserili H, Barnes AM, Steiner M, Semler O, Fischer B, Yigit G, Janda CY, Becker J, Breer S, Altunoglu U, Grunhagen J, Krawitz P, Hecht J, Schinke T, Makareeva E, Lausch E, Cankaya T, Caparros-Martin JA, et al., Mutations in WNT1 cause different forms of bone fragility, Am J Hum Genet 92(4) (2013) 565-74.

Khosla S, Minireview: the OPG/RANKL/RANK system, Endocrinology 142(12) (2001) 5050-5.

Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J, A general framework for estimating the relative pathogenicity of human genetic variants, Nat Genet 46(3) (2014) 310-5.

Kornak U, Kasper D, Bosl MR, Kaiser E, Schweizer M, Schulz A, Friedrich W, Delling G, Jentsch TJ, Loss of the ClC-7 chloride channel leads to osteopetrosis in mice and man, Cell 104(2) (2001) 205-15.

Kornak U, Mundlos S, Genetic disorders of the skeleton: a developmental approach, Am J Hum Genet 73(3) (2003) 447-74.

Laine CM, Joeng KS, Campeau PM, Kiviranta R, Tarkkonen K, Grover M, Lu JT, Pekkinen M, Wessman M, Heino TJ, Nieminen-Pihala V, Aronen M, Laine T, Kroger H, Cole WG, Lehesjoki AE, Nevarez L, Krakow D, Curry CJ, Cohn DH, et al., WNT1 mutations in early-onset osteoporosis and osteogenesis imperfecta, N Engl J Med 368(19) (2013) 1809-16.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, et al., Initial sequencing and analysis of the human genome, Nature 409(6822) (2001) 860-921.

Langer LO, Jr., Brill PW, Ozonoff MB, Pauli RM, Wilson WG, Alford BA, Pavlov H, Drake DG, Spondylometaphyseal dysplasia, corner fracture type: a heritable condition associated with coxa vara, Radiology 175(3) (1990) 761-6.

Lee CS, Fu H, Baratang N, Rousseau J, Kumra H, Sutton VR, Niceta M, Ciolfi A, Yamamoto G, Bertola D, Marcelis CL, Lugtenberg D, Bartuli A, Kim C, Hoover-Fong J, Sobreira N, Pauli R, Bacino C, Krakow D, Parboosingh J, et al., Mutations in Fibronectin Cause a Subtype of Spondylometaphyseal Dysplasia with “Corner Fractures”, Am J Hum Genet 101(5) (2017) 815-823.

Lee NK, Sowa H, Hinoi E, Ferron M, Ahn JD, Confavreux C, Dacquin R, Mee PJ, McKee MD, Jung DY, Zhang Z, Kim JK, Mauvais-Jarvis F, Ducy P, Karsenty G, Endocrine regulation of energy metabolism by the skeleton, Cell 130(3) (2007) 456-69.

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, et al., Analysis of protein-coding genetic variation in 60,706 humans, Nature 536(7616) (2016) 285-91.

Lerche H, Heine R, Pika U, George AL, Jr., Mitrovic N, Browatzki M, Weiss T, Rivet-Bastide M, Franke C, Lomonaco M, et al., Human sodium channel myotonia: slowed channel inactivation due to substitutions for a glycine within the III-IV linker, J Physiol 470 (1993) 13-22.

Li H, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM, arXiv: 1303.3997v1 [q-bio.GN] (2013).

Liu S, Quarles LD, How fibroblast growth factor 23 works, J Am Soc Nephrol 18(6) (2007) 1637-47.

Liu Y, Bhat RA, Seestaller-Wehr LM, Fukayama S, Mangine A, Moran RA, Komm BS, Bodine PV, Billiard J, The orphan receptor tyrosine kinase Ror2 promotes osteoblast differentiation and enhances ex vivo bone formation, Mol Endocrinol 21(2) (2007) 376-87.

Long F, Ornitz DM, Development of the endochondral skeleton, Cold Spring Harb Perspect Biol 5(1) (2013) a008334.

Lorentzon M, Treating osteoporosis to prevent fractures: current concepts and future developments, J. Intern. Med (2019).

Makitie O, Causes, mechanisms and management of paediatric osteoporosis, Nat Rev Rheumatol 9(8) (2013) 465-75.

Makitie RE, Costantini A, Kampe A, Alm JJ, Makitie O, New Insights Into Monogenic Causes of Osteoporosis, Front Endocrinol (Lausanne) 10 (2019) 70.

Makitie RE, Haanpaa M, Valta H, Pekkinen M, Laine CM, Lehesjoki AE, Schalin-Jantti C, Makitie O, Skeletal Characteristics of WNT1 Osteoporosis in Children and Young Adults, J Bone Miner Res 31(9) (2016) 1734-42.

Marini JC, Forlino A, Bachinger HP, Bishop NJ, Byers PH, Paepe A, Fassier F, Fratzl-Zelman N, Kozloff KM, Krakow D, Montpetit K, Semler O, Osteogenesis imperfecta, Nat Rev Dis 3 (2017) 17052.

Mayranpaa MK, Makitie O, Kallio PE, Decreasing incidence and changing pattern of childhood fractures: A population-based study, J Bone Miner Res 25(12) (2010) 2752-9.

Mayranpaa MK, Viljakainen HT, Toiviainen-Salo S, Kallio PE, Makitie O, Impaired bone health and asymptomatic vertebral compressions in fracture-prone children: a case-control study, J Bone Miner Res 27(6) (2012) 1413-24.

McClung MR, Grauer A, Boonen S, Bolognese MA, Brown JP, Diez-Perez A, Langdahl BL, Reginster JY, Zanchetta JR, Wasserman SM, Katz L, Maddox J, Yang YC, Libanati C, Bone HG, Romosozumab in postmenopausal women with low bone mineral density, N Engl J Med 370(5) (2014) 412-20.

McConville CM, Stankovic T, Byrd PJ, McGuire GM, Yao QY, Lennox GG, Taylor MR, Mutations associated with variant phenotypes in ataxia-telangiectasia, Am J Hum Genet 59(2) (1996) 320-30.

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F, The Ensembl Variant Effect Predictor, Genome Biol 17(1) (2016) 122.

Mendoza-Londono R, Fahiminiya S, Majewski J, Care4Rare Canada C, Tetreault M, Nadaf J, Kannu P, Sochett E, Howard A, Stimec J, Dupuis L, Roschger P, Klaushofer K, Palomo T, Ouellet J, Al-Jallad H, Mort JS, Moffatt P, Boudko S, Bachinger HP, et al., Recessive osteogenesis imperfecta caused by missense mutations in SPARC, Am J Hum Genet 96(6) (2015) 979-85.

Mills EW, Green R, Ribosomopathies: There's strength in numbers, Science 358(6363) (2017).

Morello R, Bertin TK, Chen Y, Hicks J, Tonachini L, Monticone M, Castagnola P, Rauch F, Glorieux FH, Vranka J, Bachinger HP, Pace JM, Schwarze U, Byers PH, Weis M, Fernandes RJ, Eyre DR, Yao Z, Boyce BF, Lee B, CRTAP is required for prolyl 3-

hydroxylation and mutations cause recessive osteogenesis imperfecta, Cell 127(2) (2006) 291-304.

Morris JA, Kemp JP, Youlten SE, Laurent L, Logan JG, Chai RC, Vulpescu NA, Forgetta V, Kleinman A, Mohanty ST, Sergio CM, Quinn J, Nguyen-Yamamoto L, Luco AL, Vijay J, Simon MM, Pramatarova A, Medina-Gomez C, Trajanoska K, Ghirardello EJ, et al., An atlas of genetic influences on osteoporosis in humans and mice, Nat Genet 51(2) (2019) 258-266.

Munns CF, Fahiminiya S, Poudel N, Munteanu MC, Majewski J, Sillence DO, Metcalf JP, Biggin A, Glorieux F, Fassier F, Rauch F, Hinsdale ME, Homozygosity for frameshift mutations in XYLT2 result in a spondylo-ocular syndrome with bone fragility, cataracts, and hearing defects, Am J Hum Genet 96(6) (2015) 971-8.

Narla A, Ebert BL, Ribosomopathies: human disorders of ribosome dysfunction, Blood 115(16) (2010) 3196-205.

Ng PC, SIFT: predicting amino acid changes that affect protein function, Nucleic Acids Res 31(13) (2003) 3812-3814.

Olsen BR, Reginato AM, Wang W, Bone development, Annu Rev Cell Dev Biol 16 (2000) 191-220.

Opperman LA, Cranial sutures as intramembranous bone growth sites, Dev Dynam 219(4) (2000) 472-485.

Ornitz DM, Marie PJ, FGF signaling pathways in endochondral and intramembranous bone development and human genetic disease, Genes Dev 16(12) (2002) 1446-65.

Paila U, Chapman BA, Kirchner R, Quinlan AR, GEMINI: integrative exploration of genetic variation and genome annotations, PLoS Comput Biol 9(7) (2013) e1003153.

Palomo T, Fassier F, Ouellet J, Sato A, Montpetit K, Glorieux FH, Rauch F, Intravenous Bisphosphonate Therapy of Young Children With Osteogenesis Imperfecta: Skeletal Findings During Follow Up Throughout the Growing Years, J Bone Miner Res 30(12) (2015) 2150-2157.

Panda A, Gamanagatti S, Jana M, Gupta AK, Skeletal dysplasias: A radiographic approach and review of common non-lethal skeletal dysplasias, World J Radiol 6(10) (2014) 808-25.

Paternoster L, Lorentzon M, Lehtimaki T, Eriksson J, Kahonen M, Raitakari O, Laaksonen M, Sievanen H, Viikari J, Lyytikainen LP, Mellstrom D, Karlsson M, Ljunggren O, Grundberg E, Kemp JP, Sayers A, Nethander M, Evans DM, Vandenput L, Tobias JH, et al., Genetic determinants of trabecular and cortical volumetric bone mineral densities and bone microstructure, PLoS Genet 9(2) (2013) e1003247.

Pazianas M, Cooper C, Ebetino FH, Russell RGG, Long-term treatment with bisphosphonates and their safety in postmenopausal osteoporosis, Ther Clin Risk Manag 6 (2010) 325-43.

Pekkinen M, Terhal PA, Botto LD, Henning P, Makitie RE, Roschger P, Jain A, Kol M, Kjellberg MA, Paschalis EP, van Gassen K, Murray M, Bayrak-Toydemir P, Magnusson MK, Jans J, Kausar M, Carey JC, Somerharju P, Lerner UH, Olkkonen VM, et al., Osteoporosis and skeletal dysplasia caused by pathogenic variants in SGMS2, JCI Insight 4(7) (2019).

Peter M, Couturier J, Pacquement H, Michon J, Thomas G, Magdelenat H, Delattre O, A new member of the ETS family fused to EWS in Ewing tumors, Oncogene 14 (1997) 1159-1164.

Pocock NA, Eisman JA, Hopper JL, Yeates MG, Sambrook PN, Eberl S, Genetic determinants of bone mass in adults. A twin study, J Clin Invest 80(3) (1987) 706-10.

Potts JR, Campbell ID, Fibronectin structure and assembly, Curr Opin Cell Biol 6(5) (1994) 648-55.

Pyott SM, Tran TT, Leistritz DF, Pepin MG, Mendelsohn NJ, Temme RT, Fernandez BA, Elsayed SM, Elsobky E, Verma I, Nair S, Turner EH, Smith JD, Jarvik GP, Byers PH, WNT1 mutations in families affected by moderately severe and progressive recessive osteogenesis imperfecta, Am J Hum Genet 92(4) (2013) 590-7.

Raisz LG, Physiology and pathophysiology of bone remodeling, Clin Chem 45(8 Pt 2) (1999) 1353-8.

Reiter JF, Leroux MR, Genes and molecular pathways underpinning ciliopathies, Nat Rev Mol Cell Biol 18(9) (2017) 533-547.

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, Committee ALQA, 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 17(5) (2015) 405-24.

Ridanpaa M, van Eenennaam H, Pelin K, Chadwick R, Johnson C, Yuan B, vanVenrooij W, Pruijn G, Salmela R, Rockas S, Makitie O, Kaitila I, de la Chapelle A, Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia, Cell 104(2) (2001) 195-203.

Rivadeneira F, Makitie O, Osteoporosis and Bone Mass Disorders: From Gene Pathways to Treatments, Trends Endocrinol Metab 27(5) (2016) 262-81.

Rivadeneira F, Styrkarsdottir U, Estrada K, Halldorsson BV, Hsu YH, Richards JB, Zillikens MC, Kavvoura FK, Amin N, Aulchenko YS, Cupples LA, Deloukas P, Demissie S, Grundberg E, Hofman A, Kong A, Karasik D, van Meurs JB, Oostra B, Pastinen T, et al., Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies, Nat Genet 41(11) (2009) 1199-206.

Rozario T, DeSimone DW, The extracellular matrix in development and morphogenesis: a dynamic view, Dev Biol 341(1) (2010) 126-40.

Ruiz-Perez VL, Blair HJ, Rodriguez-Andres ME, Blanco MJ, Wilson A, Liu YN, Miles C, Peters H, Goodship JA, Evc is a positive mediator of Ihh-regulated bone growth that localises at the base of chondrocyte cilia, Development 134(16) (2007) 2903-12.

Ruiz-Perez VL, Ide SE, Strom TM, Lorenz B, Wilson D, Woods K, King L, Francomano C, Freisinger P, Spranger S, Marino B, Dallapiccola B, Wright M, Meitinger T, Polymeropoulos MH, Goodship J, Mutations in a new gene in Ellis-van Creveld syndrome and Weyers acrodental dysostosis, Nat Genet 24(3) (2000) 283-6.

Saag KG, Petersen J, Brandi ML, Karaplis AC, Lorentzon M, Thomas T, Maddox J, Fan M, Meisner PD, Grauer A, Romosozumab or Alendronate for Fracture Prevention in Women with Osteoporosis, N Engl J Med 377(15) (2017) 1417-1427.

Saarinen A, Mayranpaa MK, Lehesjoki AE, Makitie O, Low-density lipoprotein receptor-related protein 5 (LRP5) variation in fracture prone children, Bone 46(4) (2010) 940-5.

Savitsky K, Sfez S, Tagle DA, Ziv Y, Sartiel A, Collins FS, Shiloh Y, Rotman G, The complete sequence of the coding region of the ATM gene reveals similarity to cell cycle regulators in different species, Hum Mol Genet 4(11) (1995) 2025-32.

Semenov MV, He X, LRP5 mutations linked to high bone mass diseases cause reduced LRP5 binding and inhibition by SOST, J. Biol. Chem. 281(50) (2006) 38276-84.

Shaheen R, Al-Owain M, Sakati N, Alzayed ZS, Alkuraya FS, FKBP10 and Bruck syndrome: phenotypic heterogeneity or call for reclassification?, Am J Hum Genet 87(2) (2010) 306-7; author reply 308.

Shimada T, Kakitani M, Yamazaki Y, Hasegawa H, Takeuchi Y, Fujita T, Fukumoto S, Tomizuka K, Yamashita T, Targeted ablation of Fgf23 demonstrates an essential physiological role of FGF23 in phosphate and vitamin D metabolism, J Clin Invest 113(4) (2004) 561-8.

Simsek Kiper PO, Saito H, Gori F, Unger S, Hesse E, Yamana K, Kiviranta R, Solban N, Liu J, Brommage R, Boduroglu K, Bonafe L, Campos-Xavier B, Dikoglu E, Eastell R, Gossiel F, Harshman K, Nishimura G, Girisha KM, Stevenson BJ, et al., Cortical-Bone Fragility--Insights from sFRP4 Deficiency in Pyle's Disease, N Engl J Med 374(26) (2016) 2553-62.

Stewart TL, Ralston SH, Role of genetic factors in the pathogenesis of osteoporosis, J Endocrinol 166(2) (2000) 235-45.

Tanaka AJ, Cho MT, Millan F, Juusola J, Retterer K, Joshi C, Niyazov D, Garnica A, Gratz E, Deardorff M, Wilkins A, Ortiz-Gonzalez X, Mathews K, Panzer K, Brilstra E, van Gassen KL, Volker-Touw CM, van Binsbergen E, Sobreira N, Hamosh A, et al., Mutations in SPATA5 Are Associated with Microcephaly, Intellectual Disability, Seizures, and Hearing Loss, Am J Hum Genet 97(3) (2015) 457-64.

Tattini L, D'Aurizio R, Magi A, Detection of Genomic Structural Variants from Next-Generation Sequencing Data, Front Bioeng Biotechnol 3 (2015) 92.

Tauer JT, Abdullah S, Rauch F, Effect of Anti-TGF-beta Treatment in a Mouse Model of Severe Osteogenesis Imperfecta, J Bone Miner Res 34(2) (2019) 207-214.

Taylan F, Costantini A, Coles N, Pekkinen M, Heon E, Siklar Z, Berberoglu M, Kampe A, Kiykim E, Grigelioniene G, Tuysuz B, Makitie O, Spondyloocular Syndrome: Novel Mutations in XYLT2 Gene and Expansion of the Phenotypic Spectrum, J Bone Miner Res 31(8) (2016) 1577-85.

Taylan F, Makitie O, Abnormal Proteoglycan Synthesis Due to Gene Defects Causes Skeletal Diseases with Overlapping Phenotypes, Horm Res Paediatr 48(11) (2016) 745-754.

Tort O, Tanco S, Rocha C, Bieche I, Seixas C, Bosc C, Andrieux A, Moutin MJ, Aviles FX, Lorenzo J, Janke C, The cytosolic carboxypeptidases CCP2 and CCP3 catalyze posttranslational removal of acidic amino acids, Mol Biol Cell 25(19) (2014) 3017-3027.

Vaananen HK, Zhao H, Mulari M, Halleen JM, The cell biology of osteoclast function, J Cell Sci 113 (Pt 3) (2000) 377-81.

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA, From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline, Curr Protoc Bioinformatics 43 (2013) 11 10 1-33.

van der Eerden BC, Karperien M, Wit JM, Systemic and local regulation of the growth plate, Endocr Rev 24(6) (2003) 782-801.

Van Dijk FS, Zillikens MC, Micha D, Riessland M, Marcelis CL, de Die-Smulders CE, Milbradt J, Franken AA, Harsevoort AJ, Lichtenbelt KD, Pruijs HE, Rubio-Gozalbo ME, Zwertbroek R, Moutaouakil Y, Egthuijsen J, Hammerschmidt M, Bijman R, Semeins CM, Bakker AD, Everts V, et al., PLS3 mutations in X-linked osteoporosis with fractures, N Engl J Med 369(16) (2013) 1529-36.

Van Wesenbeeck L, Cleiren E, Gram J, Rodney KB, Bénichou O, Scopelliti D, Key L, Renton T, Bartels C, Gong Y, Warman ML, de Vernejoul M, Bollerslev J, Van Hul W, Six Novel Missense Mutations in the LDL Receptor-Related Protein 5 (LRP5) Gene in Different Conditions with an Increased Bone Density, Am J Hum Genet 72 (2003) 763–771.

Viguet-Carrin S, Garnero P, Delmas PD, The role of collagen in bone strength, Osteoporos Int 17(3) (2006) 319-36.

Viljakainen H, Andersson-Assarsson JC, Armenio M, Pekkinen M, Pettersson M, Valta H, Lipsanen-Nyman M, Makitie O, Lindstrand A, Low Copy Number of the AMY1 Locus Is Associated with Early-Onset Female Obesity in Finland, PloS one 10(7) (2015) e0131883.

Wang K, Li M, Hakonarson H, ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data, Nucleic Acids Res 38(16) (2010) e164.

Wesseling-Perry K, Makitie RE, Valimaki VV, Laine T, Laine CM, Valimaki MJ, Pereira RC, Makitie O, Osteocyte Protein Expression Is Altered in Low-Turnover Osteoporosis Caused by Mutations in WNT1 and PLS3, J Clin Endocrinol Metab 102(7) (2017) 2340-2348.

Whyte MP, Greenberg CR, Salman NJ, Bober MB, McAlister WH, Wenkert D, Van Sickle BJ, Simmons JH, Edgar TS, Bauer ML, Hamdan MA, Bishop N, Lutz RE, McGinn M, Craig S, Moore JN, Taylor JW, Cleveland RH, Cranley WR, Lim R, et al., Enzyme-replacement therapy in life-threatening hypophosphatasia, N Engl J Med 366(10) (2012) 904-13.

Whyte MP, Madson KL, Phillips D, Reeves AL, McAlister WH, Yakimoski A, Mack KE, Hamilton K, Kagan K, Fujita KP, Thompson DD, Moseley S, Odrljin T, Rockman-Greenberg C, Asfotase alfa therapy for children with hypophosphatasia, JCI Insight 1(9) (2016) e85971.

Witten PE, Harris MP, Huysseune A, Winkler C, Small teleost fish provide new insights into human skeletal diseases, Methods Cell Biol 138 (2017) 321-346.

Wright CF, FitzPatrick DR, Firth HV, Paediatric genomics: diagnosing rare disease in children, Nat Rev Genet 19(5) (2018) 253-268.

Xue S, Barna M, Specialized ribosomes: a new frontier in gene regulation and organismal biology, Nat Rev Mol Cell Biol 13(6) (2012) 355-369.

Young MF, Bone matrix proteins: their function, regulation, and relationship to osteoporosis, Osteoporos Int 14(S3) (2003) 35-42.

Zarrei M, MacDonald JR, Merico D, Scherer SW, A copy number variation map of the human genome, Nat Rev Genet 16(3) (2015) 172-83.