



The
University
Of
Sheffield.

Novel Molecular Mechanisms Leading to Osteogenesis Imperfecta

By:

Rebecca Charlotte Pollitt

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

The University of Sheffield

Faculty of Medicine, Dentistry and Health

Department of Oncology and Metabolism

April 2018

Table of Contents

i	SUMMARY	1
ii	INTRODUCTION	2
ii.i	Function of Bone:	2
ii.ii	Bone structure:	2
ii.iii	Osteoblasts.....	3
ii.iv	Osteocytes.....	3
ii.v	Osteoclasts.....	4
ii.vi	Bone Modelling.....	4
ii.vii	Bone Remodelling	4
ii.viii	Regulation of Bone Metabolism.....	5
ii.viii.i	Parathyroid Hormone (PTH)	5
ii.viii.ii	Calcitonin	6
ii.viii.iii	Vitamin D	6
ii.viii.iv	Other Factors	6
ii.ix	Local Regulation of Bone Remodelling.....	7
ii.x	Bone Matrix	8
ii.x.i	Collagen structure	9
ii.x.ii	Collagen Biosynthesis.....	9
ii.x.iii	Rough Endoplasmic Reticulum	10
ii.x.iv	Golgi Apparatus	11
ii.x.v	Extracellular Matrix	12
ii.xi	Osteogenesis Imperfecta.....	12
ii.xii	Genotype/Phenotype Correlation in OI.....	12
ii.xii.i	Defects in Collagen Molecule.....	12
ii.xii.ii	Defects in Collagen post translational Modification.....	15
ii.xii.iii	Defects in collagen processing/cleavage	17
ii.xii.iv	Defects in collagen cross-linking and folding	17
ii.xii.v	Defects in osteoblast development and function.....	19
ii.xii.vi	Defects in bone mineralisation	21
ii.xii.vii	22	
ii.xii.viii	ER Related Proteins	23
ii.xii.ix	OI-like phenotype.....	24
ii.xii.x	Classification of OI	25
ii.xiii	OI pathogenesis	28
ii.xiii.i	Bone characteristics.....	29
ii.xiii.ii	ER stress	29
ii.xiii.iii	TGF β signalling.....	31
ii.xiv	Treatment	32
ii.xv	Novel Molecular Mechanisms in OI	33
ii.xvi	Project Objective	38
ii.xvi.i	Objective 1.....	38
ii.xvi.ii	Objective 2.....	38
ii.xvi.iii	Objective 3.....	38
iii	METHODS AND MATERIALS.....	39
iii.i	Funding and Ethical Approval.....	39

iii.i.i	Targeted Exome Study.....	39
iii.i.ii	Whole Exome Sequencing.....	39
iii.i.iii	Deciphering Developmental Delay.....	39
iii.ii	Patient Recruitment.....	39
iii.ii.i	Participant Identification.....	39
iii.ii.ii	Exclusion Criteria.....	40
iii.ii.iii	Recruitment.....	40
iii.iii	DNA Extraction.....	40
iii.iv	Targeted Exome Sequencing.....	41
iii.iv.i	SureSelectXT Library preparation.....	41
iii.iv.ii	Nextera Library Preparation.....	46
iii.iv.iii	Sequencing on the Illumina MiSeq.....	47
iii.v	NGS analysis pipeline.....	48
iii.vi	Mouse Models - <i>TRAM2</i> Analysis.....	50
iii.vi.i	Primer Design.....	50
iii.vi.ii	PCR reaction.....	51
iii.vi.iii	Sanger Sequencing.....	52
iii.vi.iv	Analysis of Sanger Sequencing using Mutation Surveyor.....	53
iii.vii	DDD Complimentary Analysis Project.....	53
iii.viii	Whole Exome Sequencing.....	55
iii.viii.i	Personalis.....	55
iii.viii.ii	In house whole exome sequencing.....	55
iii.ix	Prioritisation of Sequence Variants from WES.....	57
iii.ix.i	Strategy 1: An inheritance hypothesis.....	57
iii.ix.ii	Strategy 2: Exomizer.....	58
iii.ix.iii	Strategy 3: Targeted Gene Approach.....	59
iii.x	Assessment of sequence variants.....	61
iii.x.i	Population Frequency.....	61
iii.x.ii	Web-based Literature Search.....	61
iii.x.iii	Variant Effect Predictors:.....	61
iii.x.iv	Potential Splice changes.....	62
iii.x.v	Severity of Amino Acid substitution.....	63
iii.xi	Method for Electron Microscopy.....	63
iv	RESULTS: TARGETED EXOME SEQUENCING.....	65
iv.i	Patient recruitment.....	65
iv.ii	Validation of Targeted Exome NGS Panel.....	66
iv.iii	SureSelect Targeted Exome Patient Results.....	69
iv.iii.i	Investigation of <i>COL1A1</i> variant.....	70
iv.iii.ii	Investigation of <i>BMP1</i> variants.....	72
v	RESULTS: MOUSE MODELS (<i>TRAM2</i> Gene Analysis).....	83
vi	RESULTS: WHOLE EXOME SEQUENCING (Personalis).....	87
vi.i	Patient A:.....	87
vi.i.i	MYH6 c.4136C>T,p.(Thr1379Met).....	87
vi.i.ii	ARID1A c.3660G>A,p.(Met1220Ile).....	87
vi.ii	Patient B.....	88
vi.ii.i	COL6A3 c.6427G>A, p.(Gly2143Arg).....	88
vi.ii.ii	SRCAP c.9029C>A,p.(Pro3010His).....	88

vi.ii.iii	TAPBP c.1169C>T,p.(Pro390Leu).....	97
vi.iii	Patient C.....	98
vi.iii.i	ARID1B c.1041_1043dupGGC;p.(Ala350dup).....	99
vi.iii.ii	SKIV2L c.3404T>C;p.(Ile1135Thr).....	99
vi.iii.iii	NBAS c.5741G>A;p.(Arg1914His) and c.3010C>T;p.(Arg1004*)	99
vii	RESULTS: WHOLE EXOME SEQUENCING (SDGS).....	107
vii.i	Patient D.....	109
vii.i.i	NBAS c.6971G>A, p.(Arg2324His) and c.4358G>A, p.(Cys1453Tyr).....	109
vii.ii	Patient E.....	110
vii.ii.i	P4HB c.1178A>G, p.(Tyr393Cys).....	110
vii.iii	Patient F.....	115
vii.iv	Patient G.....	115
vii.iv.i	P3H1 c.1224-80G>A.....	115
vii.v	Patient H.....	118
viii	RESULTS: DDD COMPLIMENTARY ANALYSIS.....	119
viii.i	Osteogenesis Imperfecta Patients.....	119
viii.ii	Non Osteogenesis Imperfecta Patients	121
viii.iii	Type I Collagen Genes.....	123
viii.iv	Ehlers Danlos Syndrome Genes	124
viii.v	Sodium Ion Channel Genes	127
viii.vi	SLC38A10	129
viii.vii	Complex Genotypes	131
ix	DISCUSSION.....	141
x	FUTURE PERSPECTIVE	152
xi	FUNDING AND ETHICAL APPROVAL.....	155
xii	ACKNOWLEDGEMENTS.....	156
xiii	ABBREVIATIONS	157
xiv	LIST OF TABLES AND FIGURES	161
xv	REFERENCES	165
xvi	APPENDIX.....	189

i **SUMMARY**

Osteogenesis imperfecta (OI) is an inherited disorder characterized by increased bone fragility, with severity ranging from very mild to lethal. Mutation in *COL1A1* and *COL1A2*, which code for type I collagen, are identified in the majority (~85%) of patients.

Research to identify the cause of OI in individuals without identified mutations has focused on severe and lethal forms of the disease. By focusing on patients not fulfilling the current definition of “severe” adopted by the NHS funded Highly Specialised OI Service for children with severe, complex and atypical OI, we aimed to contribute to the current knowledge of the molecular basis of OI.

We employed a number of sequencing strategies to achieve this: targeted exome, Sanger and whole exome analysis. Results in this thesis have established that these approaches, together with appropriate functional analysis, can identify novel causes of OI and bone fragility.

We identified four patients with variable presentation and mutations in *BMP1* and, importantly, highlight a risk of causing delayed healing, increased stiffness or atypical fractures by anti-resorptive treatment. We report the third occurrence of a c.1178A>G;p.Tyr393Cys *P4HB* mutation and describe what appears to be an emerging, distinctive radiological phenotype: meta-diaphyseal fractures with metaphyseal sclerosis.

We have expanded the clinical spectrum associated with *NBAS* mutations to include bone fragility that may present as atypical OI. Submission of a *UKGTN* gene dossier has ensured rapid transition of this research finding into the patient diagnostic pathway.

In addition, we have identified a number of novel genes and pathways that warrant further investigation, namely *SLC38A10*, *SRCAP*, *UGGT1*, *UBASH3B*, *SULF2-POSTN*, and voltage-gated sodium channel genes. Future work will focus on elucidating the significance of these findings and the development of tools to facilitate analysis and interpretation of genetic data as whole genome sequencing becomes more readily available for OI patients.

ii INTRODUCTION

ii.i Function of Bone:

Bone has several functions within the human body. The skeleton provides structural support, protects vital organs and allows movement. Bones provide an environment for the bone marrow to enable haemopoiesis and act as a reservoir for minerals such as calcium, phosphorus and magnesium.

Bone also acts as a reservoir for growth factors and cytokines such as insulin-like growth factors (IGF), transforming growth factor- β (TGF β) and bone morphogenetic proteins (BMP). Bone also helps maintain the body's acid-base balance.

Bone strength depends largely on bone mass (50-70% of bone strength), but material properties, geometry and microstructure are also important.

ii.ii Bone structure:

Bones consist of two tissue types: the cortical (compact) bone which is dense and tough and surrounds the marrow space and the trabecular (cancellous) bone which is composed of a honeycomb-like network of trabeculae. The boundary between the cortical and trabecular bone is generally difficult to define, particularly in the metaphyses where the trabeculae merge to form the cortex (Figure 1).

A dense connective tissue layer, the periosteum, surrounds the outer surface of cortical bone except at joints where bone is lined by cartilage. The periosteal attachment to bone is stronger towards the ends of the bone, around the metaphyses and growth plates, than it is in the diaphyseal (mid-shaft) region. In relation to OI, this difference in immediate local support may help explain why metaphyseal fractures are rare compared to those in the diaphysis, despite the bone being relatively fragile.

The periosteum is vascular, contains nerves, and is rich in osteoblasts and osteoclasts. It plays an important role in bone formation, growth and fracture repair.

Bone is composed of three cell types (osteoblasts, osteocytes and osteoclasts), collagen and noncollagenous structural proteins (including proteoglycans, sialoproteins, gla-containing proteins and the 2HS-glycoprotein, a plasma protein synthesized in the liver that is enriched in bone matrix), called osteoid, and inorganic mineral salts containing calcium and phosphate that are deposited within the bone matrix and give the bone its hardness and rigidity. There is also increasing evidence that haemopoietic lineage cells, in particular macrophages and lymphocytes act within the bone environment and cross-talk with cells that directly regulate bone mass and structure (Weitzmann 2017)

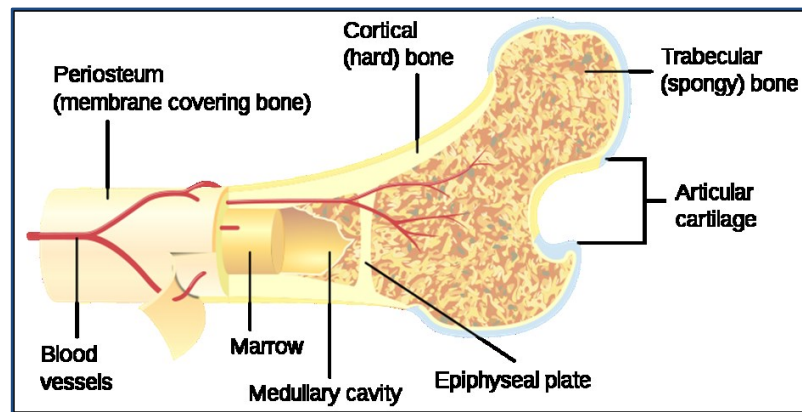


Figure 1 Bone cross-section diagram. (Pbroks13, CC BY 3.0)

<https://commons.wikimedia.org/w/index.php?curid=5188772>

ii.iii Osteoblasts

Osteoblasts originate from mesenchymal stem cells (MSC) and are responsible for bone matrix synthesis and subsequent mineralisation. Differentiation of MSCs to osteoblasts is regulated by the expression of a number of genes including Runt-related transcription factor 2 (*RUNX2*), distal-less homeobox 5 (*DLX5*) and Osterix (*OSX*). *RUNX2* is generally described as a master gene in osteoblast differentiation and has been shown to upregulate other osteoblast-related genes such as those encoding alkaline phosphatase (ALP), osteocalcin (OCN), bone sialoprotein (BSP) and type I collagen (*COL1A1/1A2*) (Bruderer, Richards et al. 2014). Expression of bone morphogenic proteins (BMPs) and Wingless (Wnt) signalling pathway proteins also regulate the commitment of MSCs towards osteoblasts.

Osteoblasts synthesise bone matrix in two main steps, the first step being secretion of an organic matrix consisting of collagens (mainly type I collagen), non-collagenous proteins (such as osteocalcin, osteonectin, bone sialoprotein II and osteopontin) and proteoglycans including decorin and biglycan. Subsequently, matrix mineralization by hydroxyapatite, a crystalline complex of phosphate and calcium ions, occurs. Osteoblasts that become embedded in the newly forming matrix give rise to osteocytes.

ii.iv Osteocytes

During the osteoblast/osteocyte transition morphological and ultra-structural changes occur. In addition, down-regulation of genes such as *BSP11*, *OCN*, *ALP* and *COL1A1/1A2* occurs. In contrast sclerostin (*SOST*) and dentine matrix protein (*DMP1*) are highly expressed.

In the osteocyte, organelles such as rough endoplasmic reticulum (ER) and Golgi apparatus are reduced in number and the nucleus to cytoplasmic ratio is increased when compared to osteoblasts. A network of plasma membrane extensions (filopodia) develop and radiate into microscopic canals in the ossified bone called canaliculi. These filopodia facilitate communication of osteocytes with each other and also with other cells types, such as osteoblasts, via the transport of small signalling molecules such as prostaglandins. Cell-cell communication is also achieved by the flow of interstitial fluid between the osteocytes filopodia and the bone canaliculi. It is through this process that

osteocytes are thought to act as mechanosensors and direct bone resorption or formation processes, although the exact mechanisms by which this is achieved have not been identified (Florencio-Silva, Sasso et al. 2015).

ii.v Osteoclasts

Osteoclasts are multinucleate cells that originate from haemopoetic stem cells. A critical osteoclastogenesis regulator is RANK ligand (RANKL). RANKL is secreted by osteoblasts, osteocytes and stromal cells and is itself regulated by osteoprotegerin (OPG). OPG binds RANKL and prevents it binding to its receptor, RANK, hence inhibiting osteoclastogenesis (Florencio-Silva, Sasso et al. 2015).

Osteoclasts function in the resorption of mineralised bone and are found in resorption cavities called Howship's lacunae. Here they secrete hydrogen ions, creating an acidic pH that dissolves the mineralized bone matrix. Hydrolytic enzymes such as cathepsin K and matrix metalloproteases are also released in order to digest the organic matrix components.

ii.vi Bone Modelling

Modelling is the process by which bones respond to physiologic or mechanical influences. The process is responsible for changes to bone size and shape and ensures the appropriate gain in skeletal mass during growth. For modelling to occur there needs to be independent action of osteoblasts and osteoclasts on separate bone surfaces. Bone modelling is less frequent than remodelling in adults.

ii.vii Bone Remodelling

This process ensures there is bone turnover while maintaining bone mass in the mature skeleton. It requires the coordination of bone resorption and formation by osteoblasts and osteoclasts respectively. Remodelling is influenced by many local and systemic factors including cytokines, hormones, chemokines and mechanical stimulation. It involves the continuous resorption of small compartments of old bone, replacement with newly synthesized matrix (osteoid), and subsequent mineralization of the matrix to form new bone (Figure 2). Known local control factors influencing remodelling include the RANK/RANKL regulation of osteoclast formation, sclerostin (SOST) production in osteocytes inhibiting bone formation, and Wnt signalling in bone and immune cells, including macrophages and lymphocytes (Martin 2017).

The remodelling process controls the repair of bone following injuries like fractures and also prevents accumulation of micro-damage that occurs during normal activity, possibly through changes in fluidic shear stress stimulating osteocyte signalling. Remodelling responses due to mechanical loading ensure that bone is added where needed and removed where it is not required, a process that helps to maintain bone strength as well as mineral homeostasis.

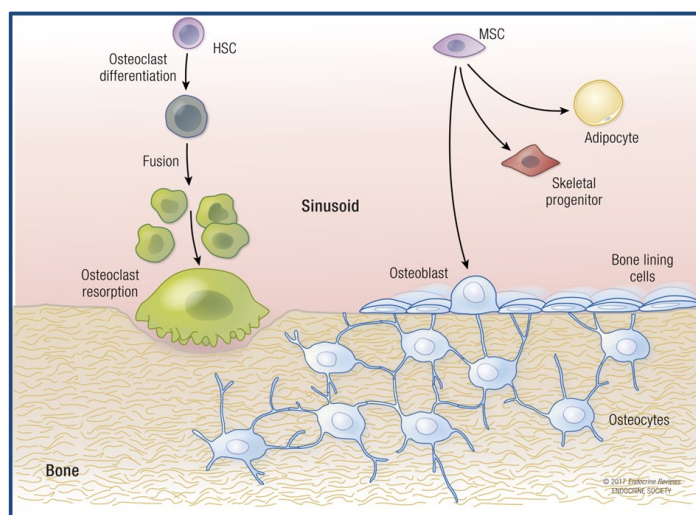


Figure 2 Bone remodelling.

Different cell types originate from either the hematopoietic or mesenchymal lineages. Osteocytes regulate both resorption by osteoclasts and formation by osteoblasts. Bone-lining cells and mesenchymal stromal cells can differentiate into osteoblasts or adipocytes. HSC, hematopoietic stem cell; MSC, mesenchymal stromal cell. Figure and legend adapted from Lee et al (Lee, Guntur et al. 2017)

ii.viii Regulation of Bone Metabolism

Ninety nine percent of the body's calcium and 85% of phosphorus is contained within the bones. Regulation of bone mineral metabolism results from the interaction of parathyroid hormone (PTH), calcitonin and vitamin D in the gastrointestinal (GI) tract, the kidneys and in bone by osteoblasts and osteoclasts.

ii.viii.i Parathyroid Hormone (PTH)

Parathyroid hormone, an 82 amino acid protein, is secreted from the parathyroid glands in response to low levels of serum calcium which is monitored through cell surface calcium sensing receptors (CSR). PTH regulates both serum calcium and phosphorus levels through its action on bone, kidneys and the GI tract. In bone, PTH exhibits its effects through osteoblasts and thus indirectly the stimulation of osteoclastic bone resorption via the RANK-OPG pathway, releasing calcium (and phosphate) from mineralised bone tissue.

In the kidneys, PTH increases the reabsorption of calcium and inhibits reabsorption of phosphate. PTH also increases the absorption of calcium in the GI tract by increasing production of activated vitamin D in the kidneys through up-regulation of 25-hydroxyvitamin D 1-alpha-hydroxylase and subsequent absorption of calcium via calbindin.

ii.viii.ii Calcitonin

Calcitonin is secreted from the thyroid gland in response to hypercalcaemia. Its main function is to inhibit osteoclast-mediated bone resorption and also results in reduced resorption of calcium and phosphorus in the kidneys.

ii.viii.iii Vitamin D

The active metabolite of Vitamin D, 1,25 dihydroxyvitamin D, increases calcium and phosphate absorption in the GI tract, providing mineral substrate for the mineralization of osteoid. It also acts with PTH to increase osteoclast activity. In addition vitamin D also regulates transcription of several bone proteins, most notably osteocalcin, type I collagen and alkaline phosphatase.

ii.viii.iv Other Factors

In addition to the hormones described above, many other hormones play a role in bone metabolism. These include gonadal steroids, such as oestrogen and testosterone that influence skeletal development and help maintain bone mass; growth hormones, insulin and androgens that promote skeletal growth and maturation; and thyroid hormones that can stimulate both synthesis and mineralisation of osteoid by osteoblasts and increase the number of osteoclasts.

	Stimulate bone formation	Decrease Bone formation	Stimulate bone resorption	Inhibit bone resorption
Growth factors	BMP-2, BMP-4, BMP-6, BMP-7, IGF-I, IGF-II, TGF- β , FGF, and PDGF		TNF, EGF, PDGF, FGF, M-CSF, and GM-CSF	
Cytokines	IL-4, IL-13, IFN, and OPG		IL-1, IL-6, IL-8, IL-11, PGE 2, PGE1, PGG2, PGI2, and PGH2	IFN- γ IL-4
Hormones	Growth hormone Vitamin D metabolites Androgens Insulin Low-dose PTH/PTHrP Progestogens	Glucocorticoids	PTH/PTHrP Glucocorticoids thyroid hormones High-dose vitamin D	Calcitonin Estrogens

Table 1 Hormones, growth factors and cytokines that influence bone metabolism and remodelling.

ii.ix Local Regulation of Bone Remodelling

Bones also contain many growth factors, including BMP and TGF- β , which are not only vital for bone development but play an important role in the local regulation of bone remodelling. These additional regulators of bone metabolism are summarised in Table 1.

The major signalling pathways in bone formation are:

Insulin-like Growth Factor I and II (IGF1 & IGFII)

Insulin-like Growth Factor I and II (IGF1 & IGFII) are polypeptides that are synthesised in the liver and osteoblasts. They increase collagen production in the bone by increasing osteoblast numbers and function. IGFII in particular is important during embryogenesis. The expression of IGF1 and IGFII is regulated by growth hormone, oestrogen, progesterone (which increase expression) and glucocorticoids (which inhibits expression). IGF1 & IGFII are linked to their respective binding proteins (IGFBP) which themselves can influence bone remodelling activity (Conover 2008).

Transforming growth factor beta (TGF β) and Bone morphogenic protein (BMP)

TGF β /BMP signalling has a fundamental role in both embryonic skeletal development and in bone homeostasis. TGF β /BMP signalling inhibits bone degrading enzymes such as metalloproteinase (MM), reduces osteoclast differentiation and activity, and strongly promotes osteoblast differentiation (Wu, Chen et al. 2016)

Platelet derived growth factor (PDGF)

PDGF stimulates osteoblast or osteoprogenitor cell (MSC) activity and is known to be up regulated following bone fracture. The current view is that PDGF contributes to bone repair by enabling osteoclasts to control osteoblast chemotaxis, however the exact mechanism by which this occurs remains to be elucidated (Caplan and Correa 2011).

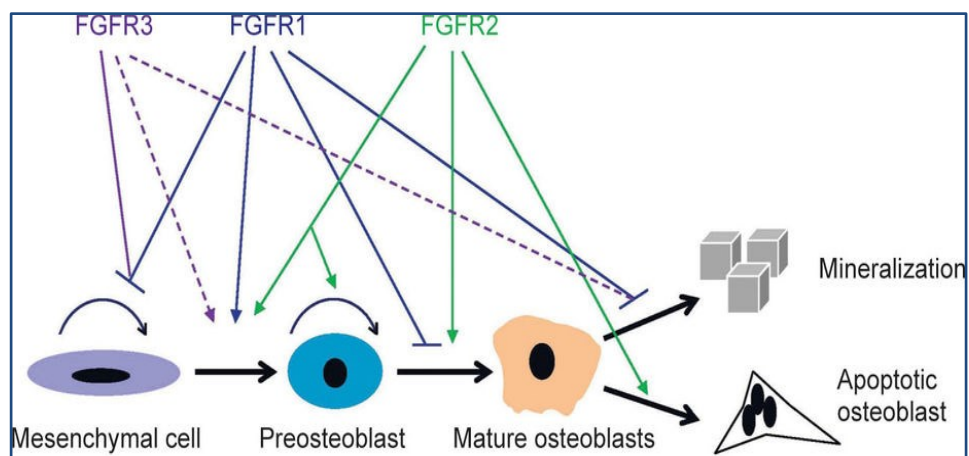


Figure 3 The role of FGFR1-3 in osteoblast differentiation.

The effect of FGFR3 on osteogenesis is unclear (marked by dotted lines). Figure adapted from (Su, Jin et al. 2014)

Fibroblastic growth factor (FGF)

The FGF/fibroblast growth factor receptor (FGFR) signalling pathway is essential for bone development, maintenance of homeostasis and regulation of phosphate levels. Multiple FGFs and FGFRs are also expressed during fracture repair (Su, Jin et al. 2014).

FGFR1 is believed to promote the differentiation of mesenchymal progenitors into pre-osteoblasts, but inhibit the proliferation of MSCs. It also inhibits the maturation and mineralization of osteoblasts. FGF2 is thought to positively modulate osteoblast differentiation and bone formation, whereas there is conflicting data concerning the effects of FGF3 signalling (Figure 3).

ii.x Bone Matrix

The bone matrix is an organized frame-work of organic and inorganic material that provides mechanical support and has a key role in the regulation of bone cell activity. It consists of 65-70% inorganic material (mineral) and around 5-10% water, the remainder being organic material such as type I collagen (90%) and non-collagenous proteins such as proteoglycan, glycoproteins and sialoproteins.

Type I collagen, the major protein component in bone, is secreted by osteoblasts and self-assembles into overlapping crosslinked groups of five. Each collagen molecule is ~ 1.5nm wide and 300nm in length (Streeter and de Leeuw 2011) and is staggered from its neighbour by approximately a quarter of its length, with a small gap between itself and the successive molecule. It is this gap that gives rise to the recognised banded appearance (known as D-periodic spacing, Figure 4) of collagen fibrils on electron microscopy (Fang, Goldstein et al. 2012). As the collagen molecules in each group coil around each other small pockets are created that act as binding sites for other matrix proteins and adhesion molecules and are important for mineralisation. Adhesion molecules, particularly integrins, are involved in the interaction between the matrix and bone cells.

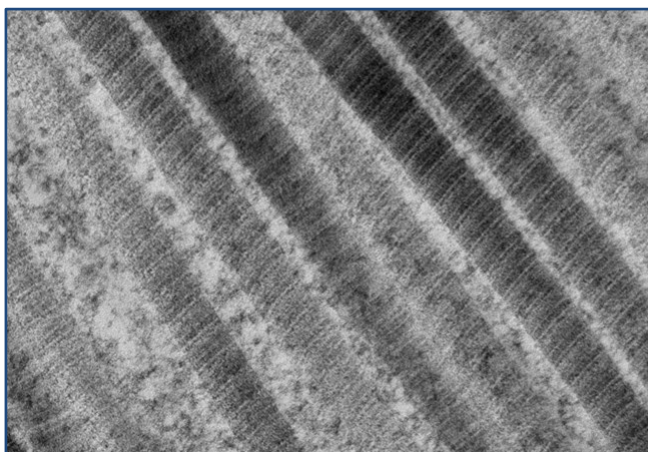


Figure 4 Electron microscopy of longitudinal cross section of dermal collagen.

Characteristic banding appearance (D-periodic spacing) caused by overlapping collagen molecules are seen (original magnification 12000X).

Collagen molecules are joined together by both enzymatic and non-enzymatic crosslinking. Lysyl oxidase and lysyl hydroxylase form enzymatic divalent crosslinks that in some mature to trivalent crosslinks. Measurement of pyridinoline and pyrrole crosslinks, which are cleaved during collagen degradation, can be used as an indication of bone turnover. Enzymatic crosslinks are considered to contribute positively to bone strength. In contrast non-enzymatic cross links, formed between sugars and exposed amino acids, are thought to have a negative effect on bone material properties, causing increased brittleness.

Crystalline hydroxyapatite is deposited along the collagen fibres in highly ordered stacks of mineral platelets. There is no current consensus as to how the location and spatial arrangement between these mineral platelets and the collagen fibres affects the material properties of bone in health and disease. However, it is known that changes in the bone matrix alter both the structure and function of bone tissue.

ii.x.i Collagen structure

Type I collagen is a heterotrimer consisting of two pro- α 1 chains and one pro- α 2 chain that are assembled into a triple helical structure. The pro- α 1 and pro- α 2 chains are encoded by the *COL1A1* and *COL1A2* genes respectively. Both pro- α 1 and pro- α 2 chains consist of a large central domain of 1014 amino acids containing repeated Gly-X-Y triplets, known as the 'triple helix' domain, which is flanked by short non-helical telopeptide domains and globular carboxyl and amino terminal propeptides (the C- and N- propeptides).

The Gly-X-Y triplets are key to the integrity of the type I collagen protein as glycine residues have the smallest side chain of any amino acid, consisting of a single hydrogen atom, allowing the three alpha chains to come together to form a tightly folded right-handed coil structure, the alpha helix. The X-position in the Gly-X-Y triplet is most often occupied by a proline residue, with hydroxyproline in the Y-position.

Along each alpha chain are regions that are important for interaction with other collagen molecules or extracellular matrix proteins, in particular proteoglycans. There are three of these regions, known as major ligand binding regions (MLBR) (Figure 5). Severe dominant OI mutations are thought to cluster in these regions, supporting the importance of proteoglycan-collagen interactions for normal bone homeostasis (Marini, Forlino et al. 2007).

Type I collagen is first synthesised as a soluble precursor protein called procollagen and undergoes several stages of processing and trafficking before fibrillogenesis occurs.

ii.x.ii Collagen Biosynthesis

The chains of type I collagen are first synthesized as procollagen molecules and the growing protein chains are translocated into the lumen of the rough endoplasmic reticulum (ER).

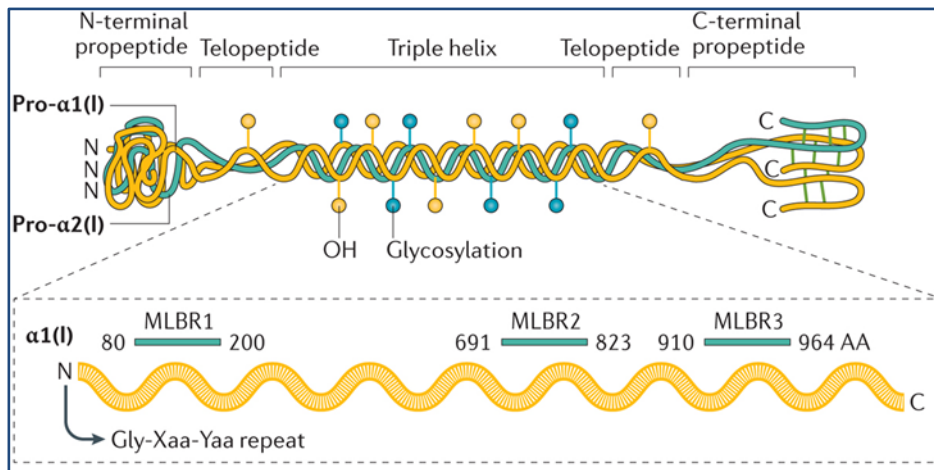


Figure 5 The structure of type I collagen.

Type I collagen consists of two pro- α 1 and pro- α 2 chains that each consist of a large central domain of 1014 amino acids containing repeated Gly-X-Y triplets, known as the 'triple helix' domain. This is flanked by short non-helical telopeptide domains and globular carboxyl and amino terminal propeptides (the C-terminal and N-terminal propeptides). The X-position in the Gly-X-Y triplet is most often occupied by a proline residue, with hydroxyproline in the Y-position. Along each alpha chain are regions that are important for interaction with other collagen molecules or extracellular matrix proteins, known as major ligand binding sites (MLBR). Figure adapted from (Marini, Forlino et al. 2017)

ii.x.iii Rough Endoplasmic Reticulum

Here the chain recognition sequence within the C-propeptide domains enable two pro- α 1 and one pro- α 2 chains to align (Khoshnoodi, Cartailier et al. 2006, Bourhis, Mariano et al. 2012). The alignment and stabilisation of the interactions between the three C-propeptides is facilitated by interchain disulphide bonds (Canty and Kadler 2005). The two pro- α 1 chains and one pro- α 2 chain then assemble in a carboxyl to amino direction leading to propagation of the triple helix. The telopeptides stabilise the newly synthesised collagen fibrils by means of intermolecular cross-links.

During chain assembly modification of lysine and proline residues by hydroxylation occurs. Specific 3'-hydroxylation of proline 1164 (also referred to as proline 986 using alternate nomenclature) is undertaken by the three protein complex formed by prolyl 3-hydroxylase 1 (P3H1), cartilage-associated protein (CRTAP), and cyclophilin B (CypB), (Ishikawa, Wirz et al. 2009) which are encoded by the *CRTAP*, *P3H1* and *PPIB* genes respectively. This 3'-hydroxylation is thought to allow binding of collagen chaperone molecules as well as having an effect on triple helix stability (Makareeva, Aviles et al. 2011).

Hydroxylation of lysine residues in the telopeptide is performed by lysyl hydroxylase 2 (LH2), encoded by the *PLOD2* gene (van der Slot, Zuurmond et al. 2003, Hyry, Lantto et al. 2009), whereas hydroxylation of two sites in the helical domain is undertaken by lysyl hydroxylase 1 (LH1), encoded by *PLOD1*.

In addition to hydroxylation processes, glycosylation of hydroxylysyl residues also occurs during helix formation. When propagation of the triple helix is complete this processing, known as post-translational modification, stops.

The ATP-independent heat shock protein HSP47, encoded by *SERPINH1*, is a further protein required for procollagen folding and is thought to help maintain triple helix stability (Christiansen, Schwarze et al. 2010). HSP47 is known to accompany procollagen from the ER to the golgi, where it dissociates and then returns to the ER for recycling (Canty and Kadler 2005). The precise function of HSP47 is still unknown although its pattern of expression suggests that it is collagen-specific rather than a general ER chaperone (Makareeva, Aviles et al. 2011).

A further ER-resident chaperone protein, FKBP65, encoded by *FKBP10*, also interacts with type I procollagen (Ishikawa, Vranka et al. 2008). FKBP65 is thought to be important for procollagen trafficking as well as molecular folding through interaction with the PLOD2 protein. In addition it forms a complex with CRTAP/P3H1/CyPB to facilitate hydroxylation of proline 1164 (Marini, Forlino et al. 2017). It has also been shown to co-operate with HSP47 during post translational modification (Duran, Nevarez et al. 2015).

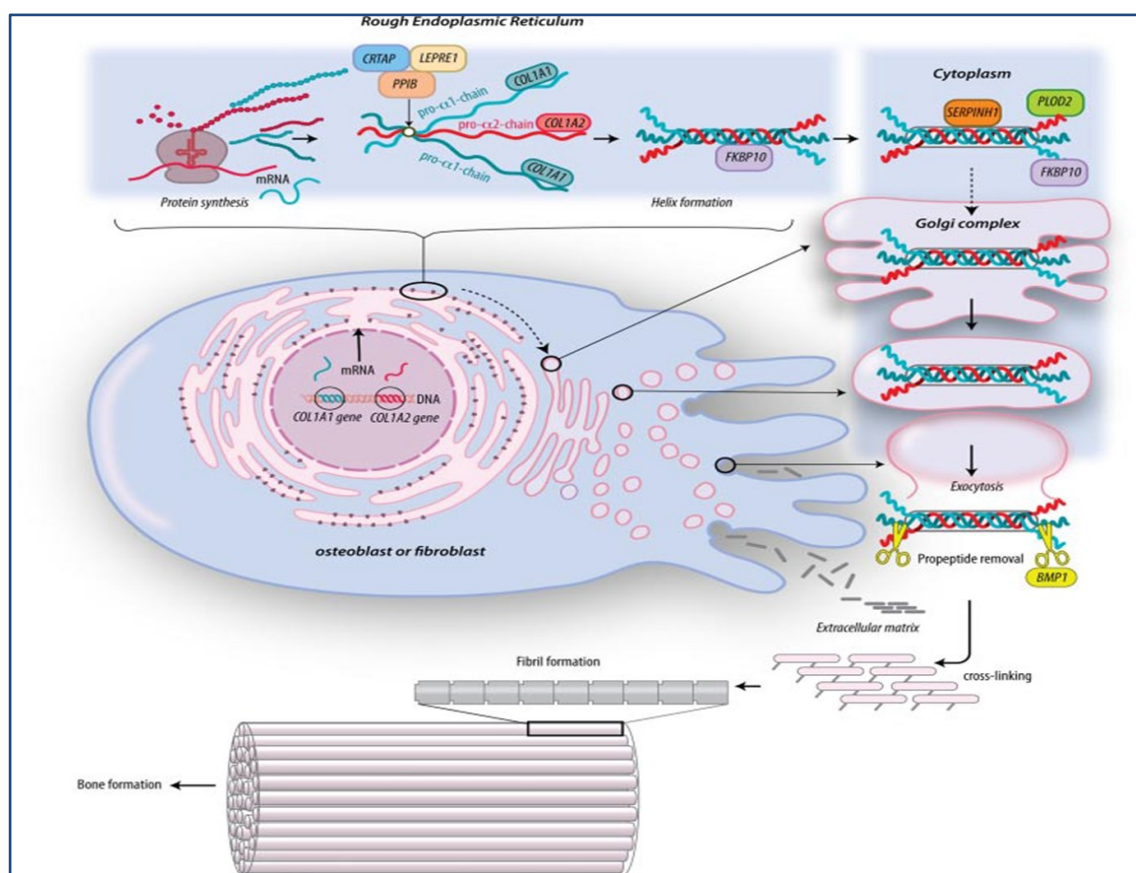


Figure 6 Steps in Type I Collagen Biosynthesis.

Figure modified from (Van Dijk and Sillence 2014)

ii.x.iv Golgi Apparatus

The type I procollagen protein exits the ER when it has achieved a stable folded configuration and is transported through the secretory pathway of the Golgi apparatus and plasma membrane into the extracellular matrix.

ii.x.v Extracellular Matrix

In the extracellular matrix procollagen is processed into collagen by the removal of the globular N and C propeptides by C-proteinases, such as bone morphogenetic protein 1, BMP-1, (Hartigan, Garrigue-Antar et al. 2003), and by the N-proteinase activity of the ADAMTS family of enzymes (Le Goff, Somerville et al. 2006). This results in spontaneous aggregation of collagen molecules to form fibrils, which is facilitated by the formation of intra- and inter-molecular crosslinks. These fibrils grow by lateral and end-to-end fusion to form type I collagen fibres (Figure 6).

ii.xi Osteogenesis Imperfecta

Osteogenesis Imperfecta (OI) is a rare inherited condition thought to affect approximately 6–7 in 100,000 individuals (Folkestad, Hald et al. 2016) and is characterised by low bone mass and an increased tendency to fracture, often with minimal or no apparent trauma. There is considerable variation in both severity and age of onset, which ranges from death in the perinatal period to very mildly affected individuals with few or no fractures.

The clinical features of OI can include fractures, short stature with bone deformities, hyperlaxity of joints and skin, and Wormian bones. Extraskelatal symptoms such as blue sclerae, progressive adult hearing loss and dentinogenesis imperfecta may also be evident. Other important features are bone pain and impaired mobility.

ii.xii Genotype/Phenotype Correlation in OI

Genetic characterisation of families affected with OI has shown that autosomal dominant mutations in the genes that encode the alpha chains of type I collagen, *COL1A1* and *COL1A2* can be identified in ~85% of affected individuals (Rohrbach and Giunta 2012). The recurrent autosomal dominant *IFITM5* mutation, c.-14C>T, is reported to account for an additional 4-5% of OI cases (Shapiro, Lietman et al. 2013).

Recessive mutations in a variety of genes are known to account for a further ~5% of cases (Forlino, Cabral et al. 2011). The contribution of each recessive gene is population dependent. In North American and European populations those encoding the components of the 3'-hydroxylation complex, *CRTAP*, *P3H1* and *PPIB* are the most common (Forlino, Cabral et al. 2011) whereas in Arabian populations *SERPINF1* and *FKBP10* are the more frequent (Shaheen, Alazami et al. 2012). In addition several recurrent mutations in *P3H1* have been identified, c.232delC;p.Gln78Serfs*30 in Irish travellers (Baldridge, Schwarze et al. 2008) and c.1080+1G>T in those of African descent (Bodian, Chan et al. 2009).

Considerable effort has been applied in an effort to define the relationship between genotype and phenotype in OI with the hope of improving prognostic information and genetic counselling/risk assessment for families with an affected individual.

ii.xii.i Defects in Collagen Molecule

ii.xii.i.i COL1A1 and COL1A2

Autosomal dominant mutations in *COL1A1* and *COL1A2* are described throughout the genes and there are currently over 1,200 unique OI variants reported (<http://www.le.ac.uk/genetics/collagen>). Mutations broadly fall into two categories, those resulting in a quantitative reduction in normal protein and those causing a qualitative defect.

Quantitative mutations causing a reduction in the amount of normal type I collagen (haploinsufficiency) are associated with milder OI phenotypes and are most commonly caused by heterozygous variants in *COL1A1* that lead to mRNA instability and nonsense mediated decay ('null' allele). Small deletions and duplications causing a frameshift and nonsense mutations are a common cause of 'null' alleles, although splice site changes resulting in intronic inclusion or activation of cryptic splice sites are also reported. Rarely whole gene deletions are also described (van Dijk, Huizer et al. 2010). Heterozygous 'null' alleles in *COL1A2* have yet to be described and are thought to result in a phenotype that cannot be distinguished from normal.

A rare autosomal recessive cardiac valvular form of Ehlers-Danlos syndrome results from homozygous 'null' mutations in the *COL1A2* gene (Schwarze, Hata et al. 2004), suggesting that type I collagen comprising of three pro- α 1 chains can assemble and maintain bone integrity in a near normal fashion in the absence of pro- α 2 chains. The lack of reported phenotypes for homozygous *COL1A1* 'null' mutations may indicate that this is incompatible with life.

Qualitative defects are associated with a wide phenotypic spectrum ranging from mild to lethal and are characterized by assembly of type I collagen comprising of mutant and normal alpha chains resulting in production of an abnormal protein (dominant-negative effect). Qualitative defects disrupt fibrillogenesis and generally have a greater phenotypic impact than quantitative changes due to their increased influence on the bone matrix.

The most frequent qualitative defect is the substitution of a glycine residue in a Gly-X-Y triplet in the triple helix domain that disrupts folding, slows chain propagation and results in over-modification of the protein. Splice site changes resulting in exon skipping, in frame deletions/insertion and changes in the C-terminal propeptide (Figure 8) also cause qualitative changes.

A key challenge in OI is the prediction of phenotypic severity from genotype. Early characterisation of glycine substitutions led to the development of a position-dependent gradient model for phenotypic severity for the pro- α 1 chain which was later refined to include the pro- α 2 chain (Lund, Aström et al. 1999). The model proposed that mutations affecting glycine residues towards the C-terminus are likely to be more clinically severe than those towards the N-terminus. This reflects the direction of chain propagation as mutations earlier in chain propagation are proposed to result in greater over-modification of the protein. However, it became apparent that this model required modification to incorporate the nature of the substituting amino acid, as those with larger and/or charged side chains have a greater impact on the folding of the triple helix.

Substitution of glycine residues with amino acids that have a charged side chain (aspartic acid, glutamic acid and arginine) as well as valine, which has a branched nonpolar side chain, leads to a severe phenotype more often than alanine which has smaller methyl group side chain (Figure 7).

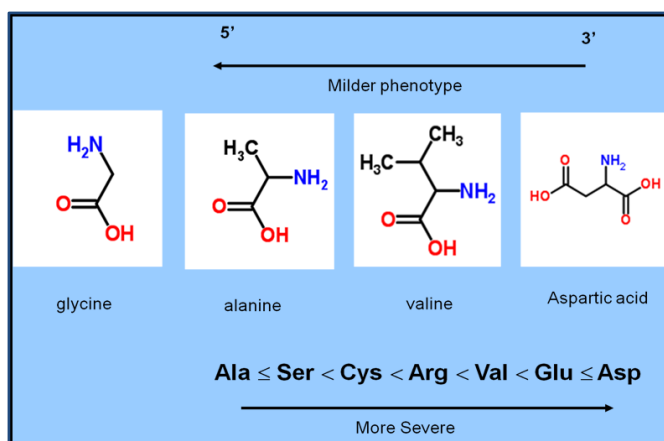


Figure 7 Gradient-dependent Severity Model for Glycine Substitutions in the Helical Domain of Type I collagen

Further refinement of this model was undertaken as the number of independent mutations grew (Marini, Forlino et al. 2007) resulting in the development of a regional model. There are two regions in the $\alpha 1$ chain where only lethal mutations have been described (amino acid positions 691-823 and 910-964) and these correspond to major ligand binding regions suggesting a role in intermolecular cross linking and fibril-chain interaction. Lethal regions in the $\alpha 2$ chain, of which there are eight, occur in clusters that correspond to the domains necessary for proteoglycan binding, further supporting the importance of fibril-matrix interaction (Bodian, Madhan et al. 2008).

In addition to glycine substitutions in the helical domain leading to disruption of chain assembly, splice site changes resulting in exon skipping and in frame deletions/insertion have also been shown to result in delayed chain assembly and subsequent over-modification of the procollagen trimer (Pace, Atkinson et al. 2001).

The phenotypic effect of non-glycine substitutions in the helical domain is more difficult to predict, and it can be hard to attribute pathogenicity to many changes. However, arginine residues when substituted by cysteine have been reported to be causative of classical Ehlers-Danlos Syndrome (EDS) (Nuytinck, Freund et al. 2000), infantile cortical hyperostosis (Caffey disease) (Cho, Moon et al. 2008) and EDS with a susceptibility to arterial rupture in early adulthood (Malfait, Symoens et al. 2007).

Mutations outside the helical domain have been described in a range of phenotypes (Figure 8). C-propeptide domain mutations have been demonstrated to delay chain assembly and result in reduced secretion of over modified but stable procollagen molecules. Mutations in this domain have been reported in mild OI (Pace, Kuslich et al. 2001) and in a variant of OI characterised by high bone mass (Takagi, Hori et al. 2011).

Retention of the C-propeptide due to mutations in either the pro- $\alpha 1$ or pro- $\alpha 2$ cleavage recognition sites have also been described to be associated with high bone mass (Lindahl, Barnes et al. 2011). The exact mechanism by which mutations in this region cause a high bone mass variant of OI is still to be determined. However, it has been hypothesised that both retention of the C-propeptide and the presence of overmodified molecules caused by mutations in this region support increased mineralisation density of the bone, which exceeds those of classical OI, because the C-propeptide

domains also act as signalling molecules within the extracellular matrix, hence influencing mineralisation. Alternatively, increase mineralization may be triggered by the physical effect of increased inter-molecular spacing due to retention of the globular C-propeptide moiety.

Mutations in the N-propeptide, other than those resulting in a 'null' allele, have not yet been described with the exception of a mutation affecting the signal peptide cleavage site (Pollitt, McMahon et al. 2006) and therefore the phenotype associated with mutations in this domain remains unknown. Mutations affecting the protease cleavage site located in exon 6, which result in reduced cleavage efficiency and retention of the N-propeptide, are associated with the arthrocholasic type of EDS. An overlapping EDS/OI phenotype is linked to mutations in the first 90 amino acid of the pro- α 1 chain due to their effect on both the secondary structure of the adjacent N-propeptide cleavage site and triple helix propagation. These patients exhibit bone fragility as well as symptoms of EDS (Cabral, Makareeva et al. 2005). Dermatosparatic EDS is caused by non-cleavage of both the pro- α 1 and pro- α 2 N-propeptides due to mutations in the N-proteinase enzyme gene, *ADAMTS2*.

The broad spectrum of phenotypes reported to be associated with mutations in the type I collagen genes illustrates the difficulty in predicting phenotypic outcome from genotype in this disorder.

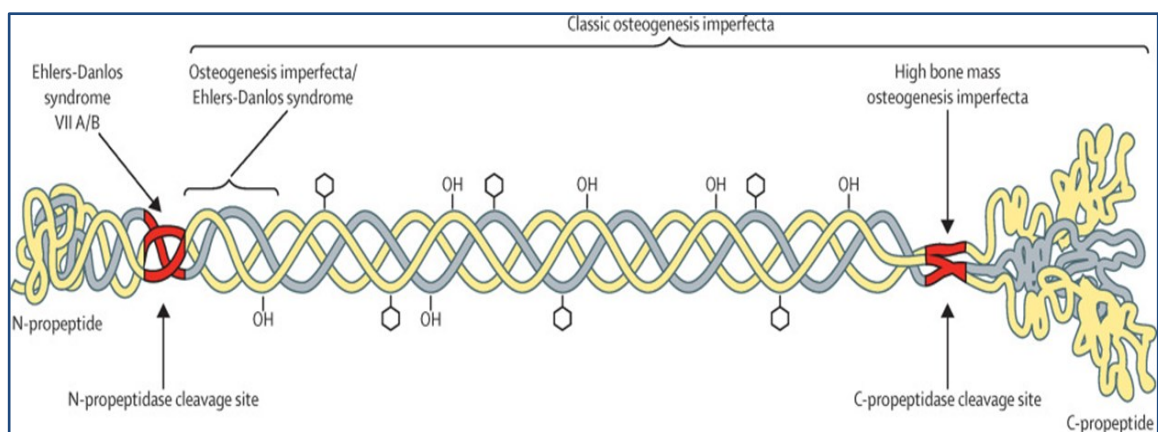


Figure 8 Different clinical phenotypes are linked to mutations at specific positions along the type I procollagen protein.

Those in the triple helical and C-propeptide domains are associated with Osteogenesis Imperfecta (OI). An OI/Ehlers-Danlos syndrome overlapping phenotypic variant is caused by amino acid substitutions in the N-anchor domain. Mutations in the N-propeptide cleavage site cause arthrocholasic Ehlers-Danlos (EDS VII). High bone mass variants of OI have been described in association with mutations at the C-propeptide cleavage site and for missense mutations within the C-propeptide domain.

Hexagons = hydroxyl lysine linked sugar molecules OH=hydroxyl group linked to proline or lysine residues.

Figure adapted from Forlino et al (Forlino and Marini 2016).

ii.xii.ii Defects in Collagen post translational Modification

The identification of autosomal recessive mutations in the prolyl 3-hydroxylation complex genes (*CRTAP/P3H1/PPIB*), responsible for the hydroxylation of proline 1164 in the α 1 chain, led the way in new gene discovery for OI. These mutations were thought to confirm that post-translational

modification and subsequent correct folding of type I collagen is essential to maintain bone strength. However, there is still uncertainty as to whether the lack of prolyl 3-hydroxylation or the loss of chaperone activity or both is the key mechanism. There is also evidence that the 3-hydroxylation site affects collagen protein interaction with SLRPs, such as biglycan and decorin, that are known to be important for fibrillogenesis (Kalamajski and Oldberg 2010).

ii.xii.ii.i CRTAP

CRTAP is a helper protein for the 3' hydroxylation complex. There is evidence that matrix disorganization as well as overmodification of the type I collagen protein contributes to the disease mechanism in CRTAP deficient OI, supporting an independent role as a matrix component for this protein (Valli, Barnes et al. 2012).

The description of a consanguineous family from Quebec with a distinct form of OI, termed type VII, (Ward, Rauch et al. 2002), and the subsequent identification of *CRTAP* as the causative gene, was key to expanding our knowledge of genes in OI (Morello, Bertin et al. 2006). The most prominent clinical features in this family were the presence of severe rhizomelia and coxa vara. No dentinogenesis imperfecta, ligamentous laxity or hearing impairment was noted and the sclerae were described as being minimally bluish. Affected individuals were confirmed to be homozygous for a c.472-1021C>G *CRTAP* variant in intron 1 that creates a cryptic splice site and the inclusion of a 73bp cryptic exon.

The majority of *CRTAP* mutations now reported result in a functional null allele with a consequent absence or significant reduction in levels of the protein. The spectrum of phenotypes associated with CRTAP deficiency range from severely deforming to lethal with rhizomelic shortening, bowing of long bones and grey/white sclerae. Skeletal abnormalities are often evident in the prenatal and early neonatal period. Relative macrocephaly may also be present (van Dijk, Nesbitt et al. 2009).

ii.xii.ii.ii P3H1

P3H1 encodes prolyl 3-hydroxylase1 (P3H1), the enzymatic component of the collagen 3-hydroxylation complex and is mutually stabilising with CRTAP (Chang, Barnes et al. 2010).

Mutations usually result in a null allele and are associated with moderately deforming to lethal phenotypes (van Dijk, Cobben et al. 2011) that overlap with CRTAP-related OI. There is usually severe under mineralisation of the bone, growth deficiency and extreme bone fragility.

A founder mutation, c.1080+1G>T in *P3H1* has been described at a frequency of 1.5% in West Africans and 0.4% in African Americans (Cabral, Barnes et al. 2012). The frequency of the common c.232delC;p.Gln78Serfs*30 mutation in Irish travellers is not reported.

ii.xii.ii.iii PPIB

PPIB encodes the cyclophilin B protein (CyPB), the third protein in the P3H1/CRTAP/CyPB complex. A limited number of mutations have been reported in this gene but have been described to result in decreased 3'-hydroxylation of proline 1164 and overmodification of type I collagen (van Dijk, Nesbitt et al. 2009).

Clinical presentation for *PPIB*-associated OI can range from perinatal lethality to moderate severity without rhizomelia or severe deformity of the long bones (Pyott, Pepin et al. 2011).

Interestingly, although mutations leading to OI have been identified in all three genes encoding the 3' hydroxylation complex, variants affecting proline 1164 itself have yet to be identified.

ii.xii.iii Defects in collagen processing/cleavage

ii.xii.iii.i BMP1

Bone morphogenetic protein 1 (BMP1) is a member of the astacin family of multidomain metalloprotease whose functions include activation of growth factors, degradation of polypeptides, and processing of extracellular proteins. The function of BMP1 includes the proteolytic removal of the C-propeptide from procollagen type I. This processing promotes self-assembly of collagen fibrils, thereby influencing fibre formation and matrix integrity. BMP1 is also described as influencing dorsal-ventral patterning through the indirect activation of some TGF β superfamily proteins (Asharani, Keupp et al. 2012).

BMP1/mTLD also removes the N-terminal peptides from the small leucine rich proteoglycans (SLRPs) biglycan, decorin and osteoglycan which themselves activate precursor forms of lysyl oxidase and lysyl oxidase-like, regulating intermolecular covalent cross-linking and influencing collagen fibril size and shape.

A mutation in the signal peptide of the *BMP1* gene has been reported in recessive OI presenting with bone fragility associated with an increase in bone mineral density. This mutation, p.Gly12Arg, causes markedly reduced post-translational N-glycosylation, impaired secretion and compromises BMP1 activity in zebra fish (Asharani, Keupp et al. 2012, Syx, Guillemyn et al. 2015).

In contrast, a missense mutation in the protease domain, p.(Phe249Leu), has been described as leading to mild OI with recurrent fractures, generalized bone deformity, osteopenia and Wormian bones. The mutant protein has been shown to have reduced proteolytic activity in-vivo (Martinez-Glez, Valencia et al. 2012).

It can be hypothesized that the described phenotypic variability between these reported cases may result from different functional consequences of the two mutations, particularly as the signal peptide mutation has a high bone mass phenotype similar to that reported in individuals with mutation in the type I collagen C-propeptide that cause delayed secretion. As further mutations are identified the relationship between mutations in this gene and those in the type I collagen C-propeptide, and their functional consequence, may become clearer.

ii.xii.iv Defects in collagen cross-linking and folding

ii.xii.iv.i SERPINH1

The ATP-independent heat shock protein HSP47, encoded by *SERPINH1*, is an ER resident multifunctional protein that acts downstream of the CRTAP/P3H1/CyPB 3-hydroxylation complex. HSP47 is believed to be required for procollagen folding, triple helix stability and quality control of the helix at the ER/Golgi boundary and is commonly referred to as a collagen 'chaperone'. Loss of function of the protein causes increased transit of type I collagen from ER to the Golgi (Van Dijk, Pals

et al. 2010). There is also evidence from an *Hsp47*^{-/-} mouse model that ER stress is increased leading to apoptosis, possibly contributing to the OI phenotype (Ito and Nagata 2017).

Two cases of recessive OI caused by missense mutations in *SERPINH1* have been reported with clinical features that include triangular face, relative macrocephaly, blue sclerae, micrognathia, short bowed limbs, thin ribs and multiple fractures (Christiansen, Schwarze et al. 2010, Duran, Nevarez et al. 2015). A third reported case of severe OI associated with a homozygous frameshift has been reported with the additional feature of hydranencephaly (Marshall, Lopez et al. 2016).

ii.xii.iv.ii PLOD2

Crucial for maintaining stability of the collagen protein in bone is the formation of lysyl and hydroxylysyl cross-links between molecules at the telopeptides and two sites located in the triple helix at position Lys87 and Lys930.

The enzyme lysyl hydroxylase 2 (LH2) is responsible for hydroxylation of telopeptide lysine and is encoded by *PLOD2*. This gene is alternatively spliced to produce two isoforms, LH2 long which is 21 amino acids longer than LH2 short. LH2 long is ubiquitously expressed whereas LH2 short has been detected only in spleen, cartilage, liver and placenta. Loss of function mutations affecting only the LH2 long isoform result in Bruck syndrome which is characterised by congenital contractures as well as susceptibility to low trauma fracture (Puig-Hervás, Temtamy et al. 2012).

Evidence that mutations in this gene can give rise to highly variable phenotypes, with and without contractures, and overlapping with mild to severe forms of recessive OI has been presented (Puig-Hervás, Temtamy et al. 2012).

ii.xii.iv.iii FKBP10

The *FKBP10* gene encodes the FKBP65 protein, an ER resident collagen chaperone, which is a member of the family of prolyl cis-trans isomerases. Recent evidence has been presented to indicate that FKBP65 also interacts with LH2, possibly resulting in misfolding of LH2 and thereby influencing telopeptide lysyl hydroxylation, which is reduced in individuals with *FKBP10* mutations (Schwarze, Cundy et al. 2013).

Novel mutations in the *FKBP10* gene were first identified by Alanay *et al* who described a cohort of five consanguineous Turkish families with moderately severe OI and a Mexican family with severe progressively deforming OI (Alanay, Avaygan et al. 2010). In contrast, those described by other groups are characteristic of Bruck syndrome (Kelley, Malfait et al. 2011, Schwarze, Cundy et al. 2013).

Marked intrafamilial variation has also been reported for mutations in this gene, with some members of the same family described as having Bruck syndrome and others OI (Shaheen, Al-Owain et al. 2011).

The broad inter- and intra-familial spectrum of phenotypes for *FKBP10* and *PLOD2* mutations, and the overlap between Bruck syndrome and OI phenotypes in these individuals supports that Bruck syndrome is a subtype of OI and does not represent a distinct clinical entity. In addition, the

overlapping phenotype between individuals with *FKBP10* and *PLOD2* mutations may be the consequence of their linked functional mechanism which influences telopeptide lysyl hydroxylation.

A further finding supporting a continuum of phenotypes is the p.(Tyr293del) founder mutation in Yup'ik Eskimos. This mutation causes an autosomal dominant congenital contracture disorder, Kuskokwin syndrome. In addition to contractures, which particularly affect the lower body and are progressive, affected individuals have short stature with osteopenic or osteoporotic vertebrae and many are reported to have low impact fractures. (Barnes, Duncan et al. 2013).

Of note is that mutations in *PLOD1*, affecting lysyl hydroxylation in the helical domain, are associated with kyphoscoliotic Ehlers Danlos Syndrome. This condition is characterized by hyperextensible and fragile skin, easy bruising, hypermobile joints, muscular hypotonia and progressive scoliosis. Scleral fragility with increased risk of optical globe rupture is also evident, although bone fragility is not present.

ii.xii.v Defects in osteoblast development and function

ii.xii.v.i *SP7*

The first gene discovered to cause recessive OI that is not involved in the type I collagen biosynthesis pathway was *SP7*, which encodes the Osterix protein (Osx). Osterix is a transcription factor expressed by osteoblasts and contains three Cys2-His2 zinc-finger DNA-binding domains. In mice, Osx has been shown to be essential for osteoblast differentiation and bone formation during development as well as postnatal bone growth and homeostasis. Mice that are deficient in Osx have perinatal lethal bone deformities similar to those seen in OI patients and show deficient osteoblast differentiation and proliferation (Nakashima, Zhou et al. 2002).

Until recently a single case of recessive OI had been reported in association with a homozygous frameshift mutation in *SP7* that results in loss of the last 81 amino acids, including the third zinc-finger motif of the protein. The affected individual is described as having moderately severe OI with fractures, mild bone deformities and white sclerae and it is speculated that the mutant protein must retain sufficient residual function to prevent this being a lethal phenotype (Lapunzina, Aglan et al. 2010).

The recent description of a c.946C>T;p.(Arg316Cys) in a sibship with short stature, osteoporosis and low impact fractures has suggested that hearing loss, that can be moderate to profound, could be a distinguishing feature of *SP7*- associated OI (Fiscaletti, Biggin et al. 2018).

ii.xii.v.ii *WNT1*

Recessive OI is now described in association with *WNT1*, a member of the WNT gene family. WNT proteins function as regulatory ligands for the frizzled transmembrane receptor family (FZD) and low density lipoprotein receptor-related proteins (LRP5/6), which trigger WNT signalling. WNT signalling is essential for many developmental and regulatory processes including cell proliferation, migration, differentiation and maintenance of bone, heart, muscle, and other tissues.

Historically WNT signalling has been divided into three main branches, the canonical β -catenin pathway, the non-canonical planar cell polarity pathway and the WNT-calcium pathway. The canonical β -catenin pathway, which has been identified as the dominant metabolic pathway regulating bone homeostasis, is activated by the binding of WNT ligands to the transmembrane receptor complex consisting of frizzled (FZD) and either LRP5 or LRP6 (Baron and Kneissel 2013). There is also increasing evidence that non-canonical WNT signalling has a role in controlling bone metabolism and interacts with other pathways affecting bone mass.

Mutations identified in WNT signalling proteins which affect bone mass have so far only been identified in those targeting the canonical WNT pathway (Figure 9), although genome wide association studies have identified common single nucleotide polymorphisms in both canonical and noncanonical WNT ligand genes as candidates for osteoporosis susceptibility.

Mutations in *WNT1* have been identified in 4 individuals from three families with a previous clinical diagnosis of OI type IV. The phenotype of these individuals includes numerous long bone fractures, multiple vertebral compression fractures and short stature. Severe abnormalities in brain development are described in the *Wnt1* knockout mouse and, although neurological problems were not observed in this cohort, two siblings from a consanguineous family were noted to have muscle hypotonia (Fahiminiya, Majewski et al. 2013).

Further *WNT1* mutations have been identified in seven affected individuals from four families described as having moderately severe, progressive phenotypes which are indistinguishable from OI type III. Affected individuals presented with recurrent fractures, bone deformity and short stature, although hearing and teeth formation are reported to be unaffected. In three of these families affected individuals are reported to have developmental delay, although whether this is entirely attributed to mutations in *WNT1* is unclear (Pyott, Tran et al. 2013).

Functional confirmation of the role of *WNT1* in the canonical WNT signalling pathway has been provided by Keupp and colleagues (Keupp, Beleggia et al. 2013) using expression vectors and dual luciferase reporter assays. Homozygous loss of function mutations, identical to those identified in individuals with moderate to progressively deforming OI, have been shown to result in almost complete failure of activation of the WNT- β -catenin signalling cascade.

The key role of *WNT1* in bone homeostasis is also supported by the recent studies of Joeng *et al* (Joeng, Lee et al. 2017) who developed both loss and gain of function mouse models. Their results demonstrated that loss of function mutations in mice results in reduced bone mass and low-impact fractures that are primarily associated with decreased osteoblast activity. In contrast, gain of function mutations (over expression) increases osteoblast numbers via activation of the mTORC1 signalling pathway.

Of particular interest is that autosomal dominant early-onset osteoporosis has been reported in a number of families with heterozygous *WNT1* loss of function mutations. Clinical features of affected individuals include low BMD and fragility fractures (Keupp, Beleggia et al. 2013, Laine, Joeng et al. 2013, Palomo, Al-Jallad et al. 2014).

This confirms that mutations in the WNT signalling pathway give rise to variable bone fragility phenotypes dependent on whether the mutation is inherited in a dominant or recessive manner. It

also emphasises that OI causative genes can have implications for other bone fragility disorders, such as osteoporosis.

ii.xii.vi Defects in bone mineralisation

ii.xii.vi.i SERPINF1

SERPINF1 encodes the 50-kDa protein pigment-epithelium-derived factor (PEDF), a multifunctional protein that is a strong inhibitor of angiogenesis as well as having involvement in bone formation and remodelling (Becker, Semler et al. 2011). The specific clinical phenotype associated with mutations in *SERPINF1*, causing loss of PEDF function, is OI type VI (Homan, Rauch et al. 2011).

Although the exact pathological mechanism underlying OI type VI is unknown, mouse studies have demonstrated that in normal mouse bone *Pedf* is localized to osteoblasts and osteocytes. The *Pedf* null mouse shows reduced trabecular bone volume, unmineralised bone matrix and increase bone fragility, mimicking findings in OI type VI and therefore providing a good model for investigation of the role of PEDF in this disorder (Bogan, Riddle et al. 2013).

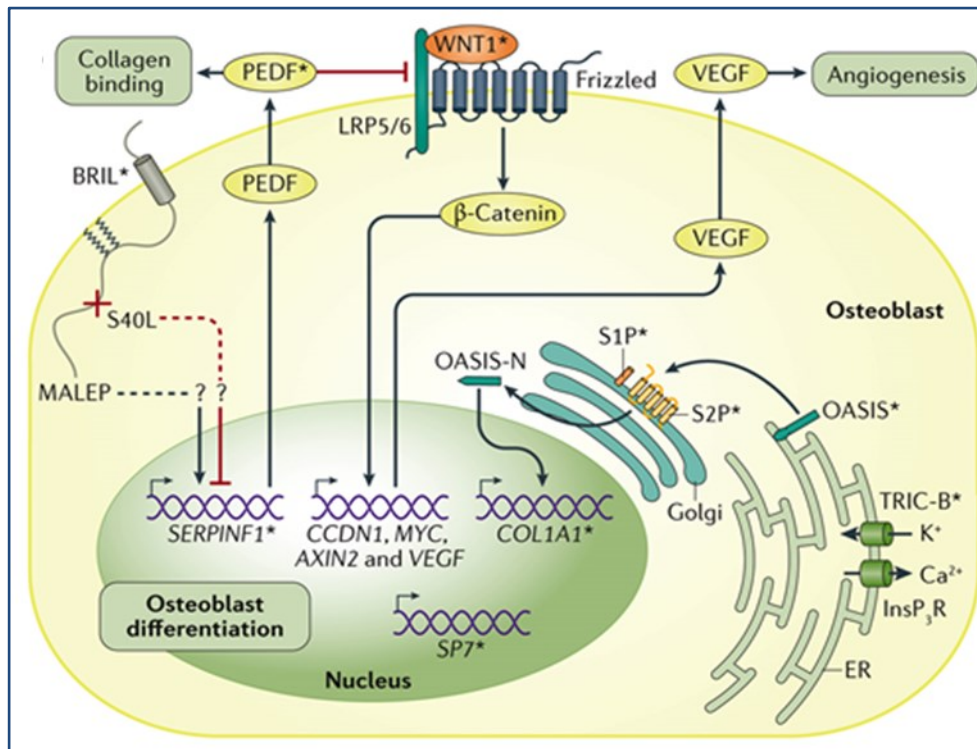


Figure 9 Regulation of bone formation and mineralization in osteogenesis imperfecta

WNT1 stimulates the transcription of genes involved in osteoblast differentiation via *LRP5/6*, *Frizzled* and the canonical β -catenin signalling pathway. The endopeptidase *S2P* (*MBTPS2*) is located in the Golgi and cleaves regulatory proteins, such as *OASIS*, via regulated intramembrane proteolysis (RIP) in times of ER stress. *TMEM38B/TRIC-B* regulates calcium flux within the ER.

Pigment epithelium-derived factor (*PEDF* encoded by *SERPINF1*), a strong inhibitor of angiogenesis, is involved in bone formation and remodelling. It cross talks with bone-restricted interferon-induced transmembrane protein-like protein (*BRIL*) to regulate mineralization.

Dashed lines represent unknown pathways. *Mutations identified in osteogenesis imperfecta. Figure adapted from Marini et al (Marini, Forlino et al. 2017)

Individuals with OI type VI appear healthy at birth with fractures of long bones, severe vertebral compression fractures and significantly low bone mineral density becoming evident between age 4 - 18 months (Glorieux, Ward et al. 2002, Wang, Liu et al. 2017). Characteristic histological findings are gross under mineralization of the osteoid and disorganized bone matrix with a distinct "fish-scale" appearance.

ii.xii.vi.ii *IFITM5*

Type V OI is characterised by hyperplastic callus formation following fracture or surgery, calcification of the interosseous membranes of the forearm and mesh-like appearance of lamellae under polarised light (Glorieux, Rauch et al. 2000). Affected individuals present with moderate to severe skeletal deformity (Lazarus, McInerney-Leo et al. 2014).

Inheritance was known to be autosomal dominant although the genetic aetiology remained unknown until linkage analysis and whole-exome sequencing identified a single causative mutation in the 5'-untranslated region of the *IFITM5* gene (alternatively known as Bone-restricted interferon induced transmembrane protein-like protein, BRIL). This single mutation, c.-14C>T, has been confirmed by transfection construct studies to create a novel in-frame start codon adding a string of five amino acids (Met-Ala-Leu-Glu-Pro) to the amino terminus of the BRIL protein. It is proposed that, as gene expression and protein localisation are restricted to skeletal tissue, this mutation presumably leads to dysregulation of bone formation (Cho, Lee et al. 2012, Semler, Garbes et al. 2012). In particular, a 'gain of function' is proposed, which leads to the formation of hyperplastic callus and ossification of membranes, although the exact mechanism requires further study.

A study of 42 individuals with OI type V reported considerable phenotypic variability even for this single mutation (Rauch, Moffatt et al. 2013). These findings have been confirmed in a further cohort of 17 individuals from 12 families where marked phenotypic variation was observed both within and between affected families (Shapiro, Lietman et al. 2013).

A second mutation in the *IFITM5* gene, c.119C>T, leading to a p.(Ser40Leu) substitution has been identified in association with atypical severe OI. Phenotypic and histological findings are suggestive of OI type VI, although *SERPINF1* analysis is negative. Protein expression studies carried out using patient osteoblasts suggest the p.(Ser40Leu) mutation stimulates intracellular *SERPINF1* regulation hence leading to the apparent impairment of mineralization (Farber, Reich et al. 2014). This supports the idea that BRIL and PEDF interact during osteoblast development, bone formation and mineralization (Figure 9).

ii.xii.vi.iii *SPARC*

Secreted protein, acidic, cysteine rich (SPARC, also known as Osteonectin) is a glycoprotein that binds extracellular matrix proteins including type I collagen within the extracellular space. It is expressed mainly in cells with high rates of extracellular matrix production, including osteoblasts in bone, and is required for collagen in bone to become calcified. Knock-out mice have defective bone formation leading to progressive osteoporosis along with intervertebral disc degeneration and reduced muscle mass (Delany, Amling et al. 2000).

Autosomal recessive OI caused by mutations in the *SPARC* gene have been identified by whole-exome sequencing in two female patients. Their phenotype is described as resembling OI type VI (associated with *SERPINF1* mutations) in that neither individual had any obvious skeletal abnormality at birth but developed progressive severe bone fragility (Mendoza-Londono, Fahiminiya et al. 2015). No further cases of *SPARC*-related OI have been reported to date.

ii.xii.vii ER Related Proteins

ii.xii.vii.i *CREB3L1*

Another gene exhibiting variable bone fragility dependent on mode of inheritance is *CREB3L1*. The gene encodes the ER-stress protein OASIS that is involved in the unfolded protein response (UPR). OASIS protein has a leucine zipper domain and a transmembrane domain that anchors it in the ER. During ER stress, the protein is released from the ER, transported to the Golgi and cleaved by S1P/MBTSP1 and S2P/MBTSP2 through regulated intramembrane proteolysis (RIP) (Figure 9). OASIS is also known to be a requirement for TGF β activation of genes involved in collagen matrix assembly (Chen, Lee et al. 2014) and OASIS $^{-/-}$ mice have severe osteopenia and are susceptible to spontaneous fractures (Murakami, Saito et al. 2009).

The first report of mutations in this gene being associated with OI is the description of two siblings with a homozygous deletion encompassing the entire *CREB3L1* gene (Symoens, Malfait et al. 2013). The siblings are described as being severely affected with intrauterine fractures.

More recently a c.934_936delAAG;p.(Lys312del) *CREB3L1* mutation has been identified in members of a consanguineous family. The mutation disrupts the DNA-binding site of the OASIS protein and prevents activation of SEC24D, a component of the COPII complex that moves secreted proteins between the ER and Golgi (Keller, Tran et al. 2018).

Individuals homozygous for the c.934_936delAAG;p.(Lys312del) mutation had multiple intrauterine fractures and reduced skull mineralisation whereas heterozygous individuals have blue sclerae, osteopenia and recurrent fractures.

ii.xii.vii.ii *MBTSP2*

A second gene involved in the RIP pathway, *MBTSP2*, has been identified in an X-linked form of moderate to severe OI in 2 extended pedigrees with missense mutations in the gene. Bone from affected individuals shows reduced hydroxylation of lysine 87 of α 1 and α 2 chains along with disorganised collagen crosslinking and reduced bone strength (Lindert, Cabral et al. 2016). The phenotype of obligate carrier females is not described but their LP/HP ratio is normal, suggesting minimal impact on bone metabolism.

The substrates of site-2 metalloprotease (S2P), encoded by *MBTSP2*, include activating transcription factor 6 (ATF6), a component of the unfolded protein response pathway and sterol regulatory element binding protein (SREBP), involved in cholesterol synthesis. S2P is located in the Golgi and, along with OASIS, cleaves these regulatory proteins during times of ER stress (Figure 9).

ii.xii.vii.iii TMEM38B

The identification of causative mutations in the *TMEM38B* gene in consanguineous Israeli Bedouin and Saudi families (Shaheen, Alazami et al. 2012, Volodarsky, Markus et al. 2013) further expanded the biochemical pathways associated with recessive OI.

The *TMEM38B* protein (TRIC-B) forms part of a trimeric cation channel which is known to regulate intracellular calcium levels in the ER (Figure 9) and is perinatally lethal in the knockout mouse model (Shaheen, Alazami et al. 2012, Volodarsky, Markus et al. 2013). Disruption of calcium flux caused by loss of *TMEM38B*/TRIC-B causes ER stress, increased BiP and dysregulation of collagen synthesis (Cabral, Ishikawa et al. 2016).

ii.xii.vii.iv SEC24D

The protein encoded by the *SEC24D* gene is part of the COPII complex. Proteins, including collagen molecules, are packaged into COPII transport vesicles as they exit the ER. Depletion of this COPII component has been shown to result in an autosomal recessive form of OI that has phenotypic overlap with Cole-Carpenter syndrome in some individuals (Garbes, Kim et al. 2015). It has been suggested that disturbed ossification of the skull and a specific facial appearance, that includes down-slanting palpebral fissures, ear dysplasia, and micrognathia, may be a characteristic feature of OI associated with *SEC24D* mutations (Zhang, Yue et al. 2017).

ii.xii.vii.v P4HB

A single autosomal dominant c.1178A>G;p.Tyr393Cys mutation in the *P4HB* gene has been identified in two individuals with Cole Carpenter Syndrome. The protein encoded by the gene, prolyl 4-hydroxylase beta subunit (also known as PDI), is abundant in the ER and acts as a chaperone preventing aggregation of the procollagen alpha chains and is involved in hydroxylation of prolyl residues in procollagen. The mechanism by which this specific mutation causes severe bone fragility remains to be confirmed but is proposed to result from interaction with two active site cysteines, Cys397 and Cys400, which normally form a reversible disulfide bond (Rauch, Fahiminiya et al. 2015).

ii.xii.viii OI-like phenotype

As well as the OI genes described above additional genes that cause bone fragility have been described and are sometimes referred to as causing an OI-like phenotype.

ii.xii.viii.i PLS3

The first report of X-linked inheritance of an OI-like phenotype was given by van Dijk *et al* who describe mutations in the actin-binding protein plastin 3 gene (*PLS3*) in five families with osteoporosis and fractures manifesting in childhood (van Dijk, Zillikens et al. 2013) . A further two families have been described where four males with a history of long bone fractures and mild vertebral compression fracture were found to have mutations in the *PLS3* gene. One of the carrier mothers was found to have low bone mass at the distal radius although without a history of fractures (Fahiminiya, Majewski et al. 2014).

The mechanism by which mutations in the *PLS3* gene cause osteoporosis and fractures is unknown although it is hypothesised that these mutations may lead to dysregulation of bone modelling/remodelling through mechano-sensing pathways in osteocytes.

Generally it is now considered that mutations in *PLS3* cause an X-linked form of osteoporosis rather than OI as it does not seem to be associated with generalized bone dysplasia, changes in bone shape or structure. Clinical presentation is mostly apparent in middle-aged adults, although onset in childhood with recurrent fractures, low (BMD) and vertebral compression fractures is also reported. Carrier females are reported to be mildly affected (Laine, Wessman et al. 2015).

ii.xii.viii.ii TAPT1

Exome sequencing following genome-wide homozygosity mapping has identified a homozygous splice site mutation in *TAPT1*, c.1108-1G>C, in a family with three affected individuals with a complex lethal osteochondrodysplasia syndrome. Multiple intrauterine fractures, incomplete ossification of the skull and severe undermineralisation of the skeleton are described along with abnormalities of the brain, kidney and lungs (Symoens, Barnes et al. 2015).

A further consanguineous family with three affected individuals is also described with multiple fractures and generalized undermineralisation of the entire skeleton. Additional findings in this family included cardiomegaly, hydrops fetalis, lung hypoplasia, and abnormalities of the umbilical arteries. The mutation in this family, c.1058A>T;p.(Asp353Val) lies within the same putative protein domain as c.1108-1G>C, namely the second extracellular/luminal loop of this transmembrane protein, although the function of this domain is still unclear.

TAPT1 encodes transmembrane anterior posterior transformation-1 protein and null mice show posterior to anterior transformation in the axial skeleton and do not survive. *TAPT1* mutant cells from one of the affected individuals have been shown to have delayed secretion of type I collagen which is overmodified. Knockdown of *tapt1b* in zebra fish causes craniofacial cartilage malformations, delayed ossification and aberrant differentiation of cranial neural crest cells (Symoens, Barnes et al. 2015).

ii.xii.ix Classification of OI

The classification of OI is based on the numerical scheme proposed by Sillence (Sillence, Senn et al. 1979) which suggested that there are four distinct syndromes within the OI spectrum:

- Type I mild disease with blue sclera and dominant inheritance
- Type II prenatally lethal with beaded ribs and crumpled long bones
- Type III progressively deforming with normal sclera and
- Type IV moderately deforming

Autosomal dominant inheritance was initially proposed for types I and IV and recessive inheritance for types II and III.

The subsequent identification of dominant mutations in all four types of OI in the genes encoding type I collagen (*COL1A1* and *COL1A2*) lead to the acceptance that OI is predominantly an autosomal

dominant disorder. Recurrence of severe OI (type II/III) in some families was attributed to germ line mosaicism and a risk of less than 10% was proposed (Pepin, Atkinson et al. 1997).

More recent evidence confirms that most cases of severe OI are caused by *de-novo* dominant mutations in either *COL1A1* or *COL1A2*, with somatic mosaicism reported in 16% of couples with an affected child. This is associated with a 27% recurrence risk for future pregnancies (Pyott, Pepin et al. 2011).

In the early 21st century the Sillence classification was extended to include three further distinct clinical groups of moderate to severe phenotypes, named as types V, VI and VII (Glorieux, Rauch et al. 2000, Glorieux, Ward et al. 2002, Ward, Rauch et al. 2002). These types were described in association with distinct clinical/histological findings and/or unknown genetic aetiology:

- Type V - moderate/severe, calcification of the interosseous membrane and hyperplastic callus formation
- Type VI - moderate/severe with accumulation of osteoid due to a mineralization defect
- Type VII - moderate/severe with autosomal recessive inheritance

The identification of homozygous loss of function mutations in the gene encoding cartilage-associated protein, *CRTAP*, (Barnes, Chang et al. 2006, Morello, Bertin et al. 2006) confirmed the recessive inheritance in OI type VII.

This triggered the subsequent identification of recessively inherited loss of function mutations in the other two proteins in the prolyl-3-hydroxylase complex, namely *P3H1* and *PPIB* (Cabral, Chang et al. 2007, van Dijk, Nesbitt et al. 2009, Barnes, Carter et al. 2010). OI associated with mutations in the *P3H1* and *PPIB* genes were assigned OI type VIII and IX respectively.

The number of genes to date associated with OI is seventeen and these now encompass additional genes involved in type I collagen biosynthesis as well those involved in mineralisation and osteoblast development and bone homeostasis (Table 2)

	OMIM number	Locus	Gene symbol	Sillence type	Main location	Bone deformity
Collagen synthesis						
Autosomal dominant	166200	17q21.33	<i>COL1A1</i>	I-IV	Extracellular matrix	Mild to very severe
Autosomal dominant	166200	7q21.3	<i>COL1A2</i>	I-IV	Extracellular matrix	Mild to very severe
Collagen processing/cleavage						
Autosomal recessive	614856	8p21.3	<i>BMP1</i>	XIII	Extracellular environment	Mild to severe
Collagen modification						
Autosomal recessive	610682	3p22.3	<i>CRTAP</i>	VII	Endoplasmic reticulum	Severe rhizomelia

Autosomal recessive	610915	1p34.2	<i>P3H1</i>	VIII	Endoplasmic reticulum	Severe rhizomelia
Autosomal recessive	259440	15q22.31	<i>PPIB</i>	IX	Endoplasmic reticulum	Severe
Collagen folding and cross-linking						
Autosomal recessive	613848	11q13.5	<i>SERPINH1</i>	X	Endoplasmic reticulum	Severe
Autosomal recessive	610968	17q21.2	<i>FKBP10</i>	XI	Endoplasmic reticulum	Mild to severe
Autosomal recessive	609220	3q24	<i>PLOD2</i>		Endoplasmic reticulum	Moderate to severe
Bone mineralisation						
Autosomal dominant	610967	11p15.5	<i>IFITM5</i>	V	Plasma membrane	Variable
Autosomal recessive	613982	17p13.3	<i>SERPINF1</i>	VI	Extracellular matrix	Moderate to severe
Autosomal recessive	182120	5q33.1	<i>SPARC</i>	XVII	Extracellular matrix	Moderate to severe
Osteoblast development						
Autosomal recessive	613849	12q13.13	<i>SP7</i>	XII	Nucleus	Severe
Autosomal recessive	615220	12q13.12	<i>WNT1</i>	XV	Extracellular matrix	Severe
ER-Related						
Autosomal recessive	616229	11p11.2	<i>CREB3L1</i>	XVI	Endoplasmic reticulum	Severe
X-linked	300294	Xp22.12	<i>MBTPS2</i>	XVIII	Endoplasmic reticulum	Moderate to severe
Autosomal recessive	615066	9q31.2	<i>TMEM38B</i>	XIV	Endoplasmic reticulum	Severe
Autosomal recessive	607186	4q26	<i>SEC24D</i>		Endoplasmic reticulum	Cole–Carpenter
Autosomal recessive	176790	17q25.3	<i>P4HB</i>		Endoplasmic reticulum	Cole–Carpenter

Table 2 Summary of Genes Associated with OI and their numerical classification.

OMIM=Online Mendelian Inheritance in Man

There is some debate in the literature concerning how to incorporate newly acquired genetic knowledge into the OI classification. One proposal (Barnes, Carter et al. 2010) is that these genes are added to the current classification as X, XI, and XII etc and this has been adopted by some groups including OMIM (Online Mendelian Inheritance in Man®). However, as the number of genes

increases, there is growing support for the classification to remain clinically focused, without reference to genetic aetiology although it could be argued that reference to genetic cause is vital if this has the potential to influence subsequent treatment or management of affected individuals.

A further difficulty in classification is posed by Bruck syndrome type 1 and type 2 which are characterised by bone fragility and joint contractures (Breslau-Siderius, Engelbert et al. 1998, van der Slot, Zuurmond et al. 2003). Autosomal recessive mutations in *FKBP10* and *PLOD2* have been identified in some affected individuals with these conditions (Ha-Vinh, Alanay et al. 2004, Hyry, Lantto et al. 2009, Setijowati, van Dijk et al. 2012). However, it is becoming apparent that the phenotype associated with *FKBP10* and *PLOD2* mutations overlaps with recessive OI (Puig-Hervás, Temtamy et al. 2012, Schwarze, Cundy et al. 2013). Indeed, there is some debate as to whether Bruck syndrome could be considered as a further OI variant, although it is not listed as such in the International Nomenclature group for Constitutional disorders of the Skeleton proposal (Warman, Cormier-Daire et al. 2011).

Similarly, Cole Carpenter syndrome is a skeletal dysplasia characterized by features of OI, such as frequent fractures and bone deformity, along with a distinctive facial appearance, craniosynostosis, ocular proptosis and hydrocephalus. Mutations in *P4HB* (Rauch, Fahiminiya et al. 2015), *SEC24D* (Garbes, Kim et al. 2015, Zhang, Yue et al. 2017) and *CRTAP* (Balasubramanian, Pollitt et al. 2015) have all been described to be causative.

The most recent proposal for OI nomenclature (Van Dijk and Sillence 2014) includes current genetic knowledge but is focused on clinical description. The new classification has five main categories:

- Type 1 Non-deforming OI with blue sclerae
- Type 2 Perinatally lethal OI
- Type 3 Progressively deforming
- Type 4 Common variable OI with normal sclerae
- Type 5 OI with calcification in interosseous membranes

The nomenclature is supported by a Severity Grading Scale designed to allow international agreed criteria to be established for classification of affected individuals. It is hoped this will help to inform clinical utility of treatments, development of new therapies and provide valuable information of likely prognosis for families. However, it seems reasonable to speculate that reference to genetic aetiology will be a requirement as personalised treatments become available.

ii.xiii OI pathogenesis

Identifying the genetic cause of OI has enabled considerable advances in our understanding of OI. However, there is still uncertainty about the underlying pathophysiology and the contribution that each proposed mechanism makes to clinical severity.

ii.xiii.i Bone characteristics

The main feature of OI is low bone mass and increased bone fragility. Bone tissue hypermineralisation, that leads to the characteristic 'brittleness', is a key feature of OI and differentiates it from other early onset bone fragility disorders.

At the macro level, OI bone from individuals with collagen defects shows cortical thinning and decreased trabecular bone. Increased osteoblasts and osteoclasts and a subsequent increase in bone turnover is observed, although at the osteoblast level deposited bone is reduced and is not compensated for by increased osteoblast numbers, resulting in reduced bone mass.

At the fibril level, increased or variable fibril diameter, as a result of defects in collagen assembly, and increased susceptibility to fibril 'kinking' leads to a disorganised matrix that is associated with increased mineral and water content. Mineral platelets, that are usually smaller in OI bone, are closely packed, overall resulting in hypermineralization. At the molecular level an increase in non-enzymatic cross linking of collagen molecules, adding to tissue stiffness, is also observed (Bishop 2016).

These multi level changes affect the ability of the OI bone to absorb normal mechanical stresses and to dissipate impacts that may result in fracture. This in turn will impact on the presence of growth factors and cytokines influencing the proliferation and differentiation of bone cells and matrix mineralisation.

ii.xiii.ii ER stress

As well as changes in bone architecture, a further proposed mechanism of OI pathogenesis is through the activation of ER stress mechanisms caused by misfolded protein molecules (Boot-Handford and Briggs 2010). This ER stress response activates synthesis of chaperone proteins designed either to help in protein folding or increase mutated protein degradation. There are three known pathways that are referred to as the unfolded protein response (UPR).

Increased unfolded protein in the ER activates three membrane receptors: protein kinase R (PKR)-like ER kinase (PERK), inositol-requiring enzyme 1 (IRE1) and activating transcription factor 6 (ATF6). These are normally bound to an HSP70 molecular chaperone located in the ER called binding immunoglobulin protein (BiP), which suppresses their activity. Disassociation of BiP from PERK, IRE1 and ATF6, when BiP is needed to promote protein folding, triggers these ER stress receptors and activates the downstream ER stress pathways (Figure 10).

Activation of PERK causes dimerisation and phosphorylation of eukaryotic translation initiation factor- α (eIF2 α), which induces more effective assembly of functional 80S ribosomes and translation of the transcription factor ATF4. This in turn upregulates transcription of genes involved in the UPR, including the transcription factor CCAAT/enhancer-binding protein homologous protein (CHOP) that induces apoptotic cell death.

In response to ER stress, the IRE1 protein degrades mRNA in the ER in a process called regulated IRE1-dependent decay (RIDD) that decreases protein biosynthesis. IRE1 also induces unconventional splicing of the mRNA that encodes X-box-binding protein 1 (XBP1), upregulating

genes encoding mediators for protein folding and those for controlling ER-associated degradation (ERAD).

Activation of ATF6, a transcription factor anchored in the ER membrane, causes its translocation to the Golgi. Here it is cleaved by the sequential action of site 1 protease (S1P) and site-2 proteases (S2P), releasing the cytoplasmic 50-kDa domain (ATF6-50). This enables translocation to the nucleus and upregulation of UPR target genes, including components of the ER-associated degradation (ERAD) pathway.

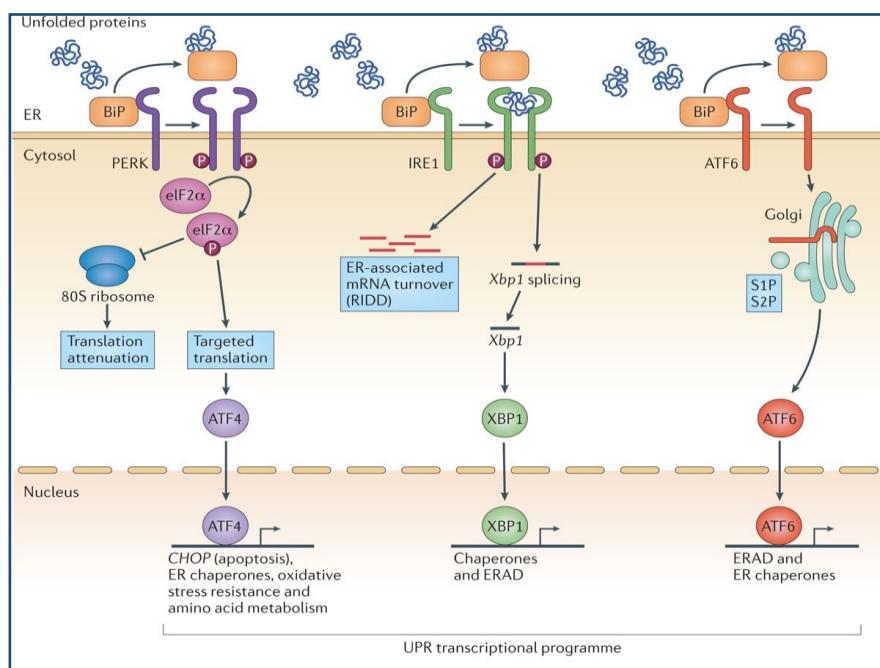


Figure 10 The unfolded protein response (UPR).

Unfolded proteins trigger disassociation of BiP from PERK, IRE1 and ATF6 triggering downstream ER stress responses. These activate synthesis of chaperone proteins designed either to help in protein folding, increase mutated protein degradation or induce apoptotic cell death. Figure adapted from (Celli and Tsolis 2015).

Mutations in the procollagen C-propeptide, that inhibit collagen chain assembly and cause propeptide misfolding, has been shown to activate the conventional unfolded protein response (UPR) leading to upregulation of BiP and subsequent ERAD (Chessler and Byers 1993, Makareeva, Aviles et al. 2011). A mouse model with a mutation in the C-propeptide domain of *COL1A1*, that causes exon skipping, provided evidence for ER stress-associated apoptosis of osteoblasts as a key component of OI pathogenesis (Lisse, Thiele et al. 2008).

In contrast, procollagen with triple helix mutations that cause misfolding does not trigger the conventional UPR but is removed by autophagy and lysosomal degradation, mainly through the PERK pathway (Besio, Iula et al. 2018). This pathway is not yet fully elucidated but is often referred to as the aggregated protein response (APR). Aggregates that accumulate due to lack of HSP47 (*SERPINH1*) also trigger APR.

In the $\alpha 2(I)$ -G610C mouse model misfolded procollagen has been shown to cause altered expression of ER chaperones αB crystalline and HSP47, as well as upregulation of the general stress response protein CHOP. Significantly, deficient osteoblast differentiation and maturation as well as abnormal response to signalling molecules such as WNT and TGF β and deposition of abnormal bone matrix was also demonstrated (Mirigian, Makareeva et al. 2016).

The ER-stress mechanism of pathogenesis is further supported by the identification of severe OI in individuals with loss of OASIS, a tissue-specific ER-stress transducer that alters gene transcription during mild ER-stress. Mutations in the *MBTPS2* gene, encoding S2P that cleaves regulatory proteins during times of ER stress have also been reported in individuals with OI.

ii.xiii.iii TGF β signalling

Transforming growth factor-beta (TGF β) signalling is known to be fundamental to both skeleton development and bone homeostasis, coupling resorption and formation to maintain bone mass (Lim, Grafe et al. 2017). TGF β is synthesised as a noncovalently bound complex with latency-associated protein (LAP). Osteoclast bone resorption results in the cleavage of LAP, releasing active TGF β inducing Smad-dependent and Smad-independent signalling (Figure 11).

In Smad-dependent signalling, TGF β binds to membrane-bound serine/threonine receptors (T β RI and T β RII) and, in a sequential phosphorylation cascade, leads to phosphorylated R-Smad (Smad2 or 3). R-Smads dissociate from their receptors and complex with Smad4, migrating to the nuclei where they regulate gene expression.

The Smad-dependent signalling pathway is negatively regulated by Smad7 that prevents R-Smad phosphorylation. Smad 7 itself is targeted by Arkadia with the resulting upregulation of Smad-dependent signalling.

TGF β –Smad signalling can locally attract osteoblast precursor cells, stimulating proliferation and differentiation. However, it also inhibits osteoblast maturation and transition into osteocytes as well as inhibiting osteoclast differentiation by reducing RANKL/OPG secretion.

In the non-Smad-dependent pathway, phosphorylated TAK1 recruits TAB1 initiating the MKK-p38 MAPK or MKK–ERK1/2 signalling cascade. The downstream target of both the Smad-dependent and the Smad-independent pathways is RUNX2, a key transcription factor associated with osteoblast differentiation. Continuous TGF β signalling can result in inhibition of osteoblast function, increasing the relative density of osteocytes and resulting in low bone mass.

Increased TGF- β signalling has been demonstrated in both the *Crtap*^{-/-} and *Col1a2*^{tm1.1Mcb} mouse models of OI. Smad2 phosphorylation and increased TGF- β target gene expression were also observed in both mice. (Grafe, Yang et al. 2014) In addition, treatment with 1D11, a TGF- β antibody, resulted in normalization of osteocyte numbers and improved bone strength.

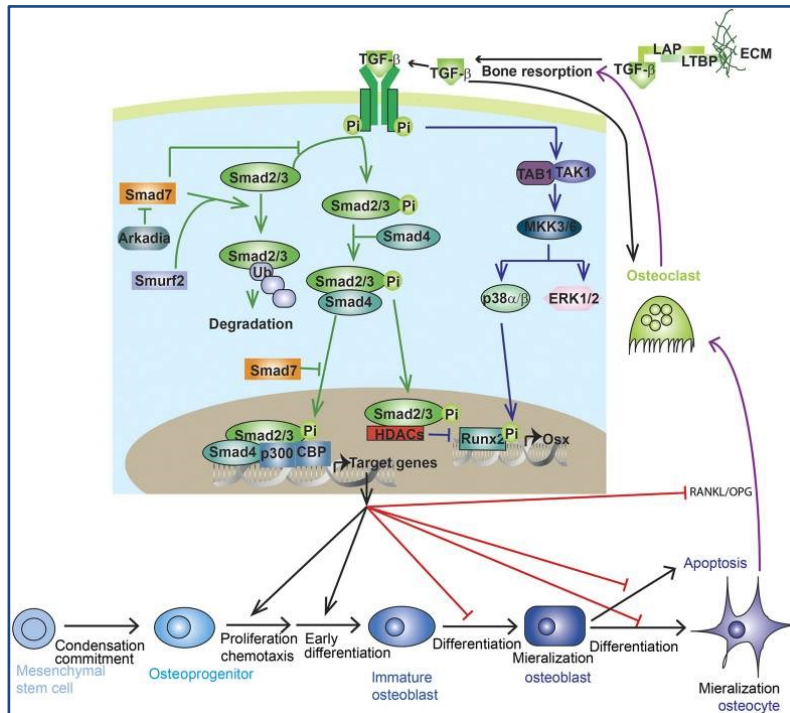


Figure 11 TGFβ signalling in bone.

In Smad-dependent signalling, TGFβ binds to TβRI and TβRII receptors leading to phosphorylated R-Smad (Smad2 or 3). R-Smads dissociate from their receptors and complex with Smad4, regulating gene expression. This pathway is negatively regulated by Smad7 that prevents R-Smad phosphorylation. Smad 7 itself is targeted by Arkadia with the resulting upregulation of Smad-dependent signalling. TGFβ–Smad signalling stimulates proliferation and differentiation of osteoblast precursor. However, it also inhibits osteoblast maturation and transition into osteocytes as well as inhibiting osteoclast differentiation by reducing RANKL/OPG secretion. In the non-Smad-dependent pathway, phosphorylated TAK1 recruits TAB1 initiating the MKK–p38 MAPK or MKK–ERK1/2 signalling cascade. The downstream target of both the Smad-dependent and the Smad-independent pathways is RUNX2, a key transcription factor associated with osteoblast differentiation. Figure modified from (Wu, Chen et al. 2016)

Similarly altered TGF-β signalling was also observed in *Brtl*^{+/-} mice, where the degree of TGF-β upregulation correlates with phenotypic severity (Bianchi, Gagliardi et al. 2015). These findings suggest that increased TGF-β signalling may be a common mechanism in both dominant and recessive forms of OI.

ii.xiv Treatment

Treatment for OI is not curative and requires a multidisciplinary approach, including rehabilitation (physical and occupational therapy), surgical interventions and medical treatment (Marr, Seasman et al. 2017)

The main pharmacological treatment for OI is bisphosphonates, antiresorptive agents that can improve bone architecture and increase bone mass by reducing osteoclastic bone resorption. However, the efficacy of this treatment in terms of fracture rate reduction and bone pain is not completely clear (Marini, Forlino et al. 2017).

A further treatment option is denosumab, a RANKL antibody that also inhibits osteoclast maturation. A limited number of studies have reported an increase in lumbar spine aBMD but without subsequent mobility improvements in children treated with denosumab (Hoyer-Kuhn, Franklin et al. 2016). Teriparatide, a human recombinant parathyroid hormone with known osteoanabolic effect, has also been shown to be effective at increasing BMD in mildly affected adults (Orwoll, Shapiro et al. 2014) without a clear effect on fracture risk.

Emerging therapies:

Sclerostin, is a negative regulator of bone formation and acts by inhibition of Wnt signalling, thus down-regulating osteoblast development and function. Treatment of post-menopausal osteoporosis by romosozumab, an anti-sclerostin antibody, results in increases in bone mineral density and reduced fracture risk (Cosman, Crittenden et al. 2016). Studies in *Wnt1sw/sw* mice showed significant improvements in bone quality and quantity when treated with anti-sclerostin antibody and suggest that this could be an effective gene specific treatment for *WNT1*-related OI (Joeng, Lee et al. 2017). Expanded studies in different mouse models would be useful to inform the likely efficacy in other genotypes.

Raloxifene, a selective oestrogen receptor modulator (SERM) is also used to reduce fracture risk in post-menopausal osteoporosis. The mechanism of action is thought to be via a combination of reduction in bone resorption and altered skeletal hydration. These findings have been duplicated in bone from the *oim* mouse that also had reduced fracture rate following treatment (Berman, Wallace et al. 2016).

Another potential emerging therapy is anti TGF β antibodies that have been shown to increase bone maximum load and strength in recessive (*Crtap*^{-/-}) and dominant (*Col1a2*^{+/*P*.G610C}) mice. Response in these models is genotype dependent with the *Crtap*^{-/-} model exhibiting greater response to 1D11, a pan TGF β antibody. However this treatment did not appear to rescue the reduced collagen or increased brittleness in either of these OI models (Bi, Grafe et al. 2017). A study with fresolimumab, a human polyclonal antiTGF β antibody, is underway in adults with OI in the USA.

Mutant allele specific silencing has been achieved in fibroblasts from *Brtl* mouse using small interfering RNAs targeting the 3'UTR of the gene opening the possibility of bespoke mutation specific treatment (Lindahl, Kindmark et al. 2013, Rousseau, Gioia et al. 2014).

Fetal stem cell transplantation has been reported in two human cases of OI; efficacy is hard to assess due to variable assessment of severity at baseline (Le Blanc, Götherström et al. 2005, Götherström, Westgren et al. 2014).

ii.xv Novel Molecular Mechanisms in OI

Understanding both the genetics and underlying mechanism of OI pathogenesis is vital for the development of new and improved treatments.

Despite the recent proliferation of genes associated with OI, as a consequence of the recent advances in gene discovery technology, there remain ~5% of OI individuals in whom a causative

mutation has not been identified. Within this group is a cohort of families with mild/moderate OI without a mutation in the type I collagen genes. However, research to identify new OI genes has largely focused on severely affected individuals, often from large consanguineous families, where recessive inheritance is implicated.

Historically a number of strategies have been employed to identify causal genes in inherited disorders. Linkage analysis, using families with multiple affected members, and positional cloning were traditionally used to support gene identification, the most well known examples being the identification of the loci for Huntington disease (Gusella, Wexler et al. 1983) and discovery of the cystic fibrosis gene (Riordan, Rommens et al. 1989). Linkage analysis was traditionally employed in OI to identify which type I collagen gene was causal within an individual family.

More recently, the emergence of 'next generation' high through-put sequencing platforms and high density arrays has enabled rapid identification of more than 100 genes associated with Mendelian disorders (Rabbani, Mahdiah et al. 2012, Study 2015), including many of the newly described OI genes. High density single nucleotide variant (SN) arrays to identify homozygosity regions in three consanguineous OI families followed by targeted exome sequencing lead to the identification of the *TMEM38B* gene (Shaheen, Alazami et al. 2012, Volodarsky, Markus et al. 2013). Similarly, *SP7* was identified as the causative gene in a single affected individual from a consanguineous family using a combination of homozygosity mapping and a subsequent candidate gene sequencing approach (Lapunzina, Aglan et al. 2010). Linkage analysis and whole exome sequencing in a four-generation family has also been employed to identify the dominant mutation in *IFITM5* responsible for OI type V (Cho, Lee et al. 2012).

However, for non-consanguineous families where the number of affected individuals is small, or for apparently *de-novo* cases, the identification of the causal gene can be more problematic, particularly in disorders with genetic and phenotypic heterogeneity such as OI. The number of variants identified by whole exome sequencing can be large, averaging more than 80,000, dependent on ethnic origin (Belkadi, Bolze et al. 2015). Filtering and analyzing this data is challenging and interpretation of whole exome sequencing can therefore be difficult.

Even when data is filtered to remove common polymorphisms, several hundred potentially pathogenic variants can remain, dependent on study design and filtering strategy employed. For sporadic cases analysis of trios (proband and parents) may help to detect *de-novo* variants, although the presence of a *de-novo* variant is not considered sufficient to confirm causality as there is an estimated *de-novo* mutation rate of one in 10^8 bases per haploid genome per generation (Lynch 2010). Commonly, variants in genes already reported to be associated with the condition are analysed as a first step in exome data analysis.

The use of a targeted candidate gene approach to identify causative mutations, rather than whole exome sequencing, has the advantage that it is focused on existing knowledge of the condition and its possible biological origin. Candidate genes are generally those whose function directly or indirectly influences the developmental or regulatory processes of the trait under investigation. In OI this would be biological pathways known to be involved in skeletal development and homeostasis. Variants identified in an appropriate pathway can be more confidently interpreted than those outside, although ideally functional studies are required to confirm the effect of any variant on protein function.

Mutations in genes involved in pathways influencing bone development and homeostasis are known to be associated with a variety of conditions and some show inheritance in both an autosomal recessive and autosomal dominant manner for example *LRP5* and *ALPL*.

ii.xv.i.i LRP5

The *LRP5* gene encodes the low density lipoprotein receptor-related protein 5, a transmembrane co-receptor that participates in the WNT- β -catenin signalling pathway with Frizzled (*FZD4* gene) and LRP6. As previously discussed WNT signalling is involved in the development and maintenance of several tissues, influencing cell proliferation, adhesion and migration, and in particular plays a key role in skeletal homeostasis and eye development and function. Mutations in *LRP5* give rise to several conditions, which reflect the involvement of the protein in the affected tissue, and can be inherited in either a dominant or recessive manner.

Heterozygous loss of function mutations in *LRP5* result in familial exudative vitreoretinopathy (FEVR) a condition that is characterised by disruption to early eye development, abnormal blood supply to the retina and resulting loss of vision. As *LRP5* also plays a role in bone formation, individuals with FEVR may also have reduced bone mineral density (Qin, Hayashi et al. 2005).

Homozygous loss of function mutations are associated with the severe osteoporotic disorder Osteoporosis Pseudoglioma (OPPG) where severe low bone mineral density results in multiple fractures. Individuals with OPPG also have eye abnormalities. *LRP5* has also been linked to Idiopathic Juvenile Osteoporosis (IJO) and reduced bone mass in the heterozygous carriers of OPPG (Ferrari, Deutsch et al. 2005).

In contrast heterozygous gain of function mutations are associated with high bone mass conditions such as autosomal dominant osteopetrosis type 1 and endosteal hyperostosis (Little, Carulli et al. 2002).

ii.xv.i.ii ALPL

Hypophosphatasia is caused by loss or reduced function of tissue non-specific alkaline phosphatase (TNSALP). The functional mechanism of the TNSALP enzyme is not completely understood, but it is known to be essential for mineralization of bone and teeth and has been shown to increase concentrations of phosphate, which promotes mineralization, and decreased pyrophosphate, which is an inhibitor of mineral formation (Hessle, Johnson et al. 2002).

Hypophosphatasia is characterized by defective mineralization of bone and/or teeth caused by mutations in the *ALPL* gene which encodes TNSALP. In a similar manner to OI, there is considerable variation of clinical severity ranging from stillbirth without mineralized bone to mildly affected adults with early loss of dentition and few fractures. There are six recognized forms of the condition based on severity and age of onset (Table 3).

Type	Inheritance ¹	Cardinal Features	Dental Features	Clinical Diagnosis
Perinatal (lethal)	AR	Hypomineralisation, osteochondral spurs	N/A	Radiographs, prenatal ultrasound examination
Perinatal (benign)	AR or AD	Long-bone bowing, benign postnatal course	±	Prenatal ultrasound examination, clinical course
Infantile ²	AR	Craniosynostosis, Hypomineralisation, rachitic ribs, hypercalciuria	Premature loss, deciduous teeth	Clinical course, radiographs, laboratory findings
Childhood	AR or AD	Short stature, skeletal deformity, bone pain/fractures	Premature loss, deciduous teeth (incisors)	Clinical course, radiographs, laboratory findings
Adult ³	AR or AD	Stress fractures: metatarsal, tibia; chondrocalcinosis	±	Clinical course, radiographs, laboratory findings
Odontohypophosphatasia	AR or AD	Alveolar bone loss	Exfoliation (incisors), dental caries	Clinical course, dental panorex, laboratory findings

Table 3 Classification of Hypophosphatasia

Modified from National Center for Biotechnology Information (NCBI; <http://ncbi.nlm.nih.gov>).

The inheritance of the most severe forms, perinatal and infantile, is in an autosomal recessive manner whereas milder forms may be inherited in an autosomal recessive or autosomal dominant. Parents of infants with the recessive perinatal form of the disease may, on investigation, be shown to have undiagnosed mild symptoms. The differing inheritance patterns and marked phenotypic heterogeneity in this condition is related to the variable levels of residual enzyme activity caused by missense mutations that predominantly affect the functional domains of the TNSALP protein (Mornet 2000).

The examples of *LRP5* and *ALPL* show that genes encoding proteins participating in pathways influencing skeletal development may be associated with marked phenotypic variability and/or different inheritance patterns. This is further supported by the reports of heterozygous dominant *WNT1* and *CREB3L1* mutations in families with early onset osteoporosis and fractures as well as in families with moderately severe recessive OI (Keupp, Beleggia et al. 2013, Laine, Joeng et al. 2013, Palomo, Al-Jallad et al. 2014, Keller, Tran et al. 2018)

The prevalence and range of phenotypes associated with mutations in many of the severe autosomal recessive OI genes is currently unclear and the phenotype of heterozygous carriers is not well documented. These genes are therefore good candidates for further investigation.

ii.xvi Project Objective

The overall aim of this research was to contribute to the current knowledge of the molecular basis of Osteogenesis Imperfecta (OI) by investigation of patients with a mild/moderate clinical presentation without a causative mutation in the type I collagen genes being identified. This is important to be able to deliver personalised medicine for these patients in the future.

ii.xvi.i Objective 1.

The first objective of this study was to identify a cohort of families with mild/moderate OI without a mutation in the type I collagen genes.

This involved a comprehensive review of the molecular and phenotypic data obtained in a large cohort of OI patients referred to the Sheffield Diagnostic Genetics Service connective tissue disorder service. Those without a likely pathogenic mutation in the type I collagen genes were invited to participate.

ii.xvi.ii Objective 2.

The second objective was to elucidate the genetic aetiology in this cohort. A number of different strategies have been employed to achieve this:

- An initial study where all participants were screened for variants in genes previously describe in severe OI using a custom designed targeted Next Generation Sequencing exome panel (NGS) was undertaken.
- All participants without a pathogenic mutation in any of these genes were then screened by Sanger sequencing for mutations in the *TRAM2* gene. This gene was identified by the Sanger mouse phenotyping consortium as a strong candidate gene for OI.
- A small subset of participants who screened negative for all known OI genes and *TRAM2* were investigated using whole exome sequencing.
- Follow on studies to investigate potential pathogenicity and mechanisms of novel sequence variants identified using additional techniques such as examination of dermal collagen and fibroblasts, methylation studies and protein expression analysis, where appropriate.

ii.xvi.iii Objective 3

The third objective of this study was to explore a broader dataset of variants from patients who have clinical features relating to bone health, including fractures. This part of the study involves analysis of data from whole exome sequencing and array CGH analysis through a collaborative complementary analysis project (CAP) with the 'Deciphering Developmental Disorders' (DDD) study.

iii METHODS AND MATERIALS

iii.i Funding and Ethical Approval

iii.i.i Targeted Exome Study

The first element of this research project was to study autosomal recessive genes in a cohort of patients previously tested negative for the type I collagen genes. To achieve this I wrote an application to obtain funding from the Sheffield Children's Charity (Protocol Number: SCH/11/062). The application included development of the study design (Figure 12), preparation of age appropriate patient information sheets, consent forms and patient eligibility criteria (Appendix pages 277, 283 and 298).

Following the successful funding application, I gained ethical approval for the project from the Yorkshire and Humber NRES committee (Reference Number: 12/YH/0021). R&D approval was obtained from the Clinical Research Facility at Sheffield Children's NHS Foundation Trust and at each host organisation prior to the start of patient recruitment, in accordance with NHS research governance procedures.

iii.i.ii Whole Exome Sequencing

A subsequent application for funding to support whole exome sequencing in a smaller cohort of patient was developed in collaboration with Dr Balasubramanian, Dr Dalton and Prof Bishop to The Children's Hospital Charity (TCHC) grant number CA15001.

iii.i.iii Deciphering Developmental Delay

In collaboration with Dr Balasubramanian, a project application was submitted and successfully registered with Deciphering Developmental Delay study (CAP12, Appendix page 296)

iii.ii Patient Recruitment

iii.ii.i Participant Identification

Potential participants were initially identified for this study by a review of all patient records held at Sheffield Diagnostic Genetics Service to identify patients referred for analysis of the type I collagen genes. A total of 655 patients were identified. Patients with either a confirmed pathogenic mutation or a variant of uncertain significance in these genes were excluded. Presenting phenotypic details were reviewed for all remaining patients. Further exclusion criteria were subsequently applied as follows.

iii.ii.ii Exclusion Criteria

Individuals without a clear clinical phenotype were excluded as were those where analysis had been performed to exclude a diagnosis of OI in cases of suspected non-accidental injury (medico-legal testing).

Patients with a lethal (type II) phenotype or fulfilling the current definition of “severe” adopted by the NHS-funded Highly Specialised OI Service for children with severe, complex and atypical OI were excluded.

The “severe” definition criteria are:

- Neonates with OI who have multiple long bone fractures
- Children/infants with OI who have:
 - ✓ Six or more vertebrae crush-fractured/deformed (more likely to progress to scoliosis according to published data)
 - ✓ Multiple limb deformities resulting in corrective surgery and attendant frequent OT/PT input
 - ✓ Intractable bone pain despite intravenous bisphosphonate therapy

iii.ii.iii Recruitment

Consultant Clinical Geneticists who had referred individuals eligible for the study were contacted with information on the study. Those willing to support the study were asked to discuss the study with relevant patients and carers, either directly or through their Genetic counselling team. Patients interested in participation were sent an invitation letter and age appropriate patient information sheet describing the study and its objective. Recruitment involved obtaining consent for genetic analysis for the purpose of research.

The majority of participants had DNA stored at the Sheffield Diagnostic Genetics Department (SDGS). Where further DNA was required a fresh sample of blood was obtained and automated DNA extraction was performed.

iii.iii DNA Extraction

The Chemagen Magnetic Separation Module is used for automated extraction of genomic DNA from whole blood using the Chemagic DNA blood kit. The extraction is carried out in two stages, the first involves mixing and lysing of white blood cells to release DNA, the second involves DNA being bound to magnetic beads, from which the DNA is subsequently eluted. The kit contains magnetic beads, lysis buffer, binding buffer, wash buffers and elution buffer. The eluted DNA is stable at 4°C or -20°C.

A peripheral blood sample of 1-3ml (in EDTA) is required. The expected DNA concentration is ≥ 40 ng/ μ l in an elution volume of 0.3 ml (in 1 x TE buffer) which is sufficient for subsequent sequencing by either Sanger or Next Generation Sequencing methodologies.

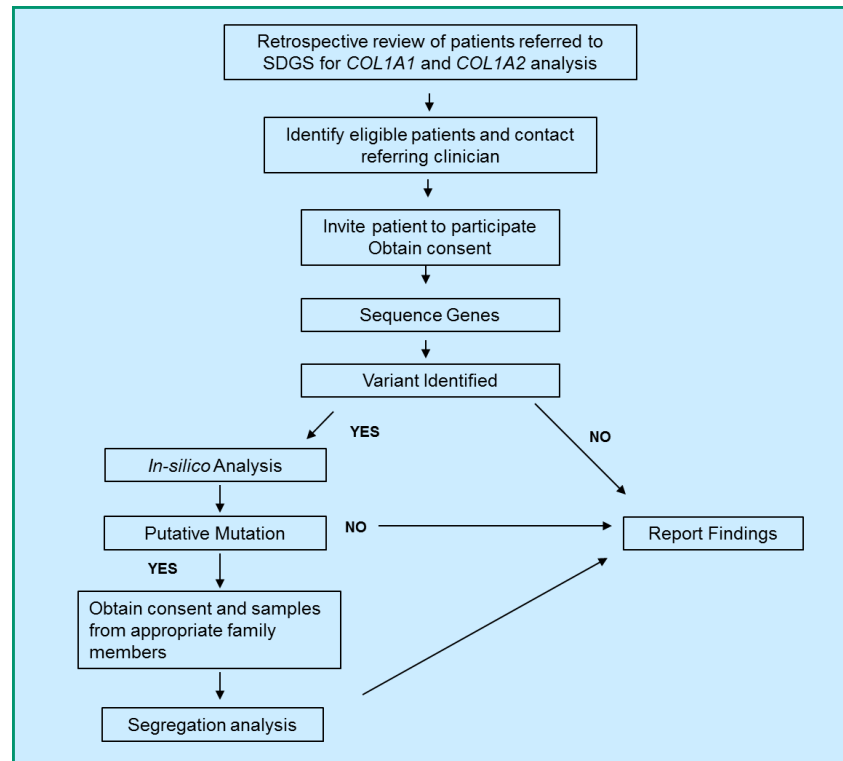


Figure 12 Study Workflow

Study Workflow for targeted exome sequencing showing patient recruitment, sample analysis and reporting.

iii.iv Targeted Exome Sequencing

Next Generation Sequencing (NGS) is a broad term used to describe the simultaneous sequencing of multiple fragments of DNA. This technology allows multiple genes of interest, or even the whole genome, from an individual to be sequenced in parallel.

The first part of this research was to use a candidate gene approach focused on recently identified genes leading to severe OI. To achieve this objective a ‘targeted’ next generation sequencing strategy was developed whereby only fragments of DNA from the specific genes of interest were ‘selected’ for sequencing. This process is referred to as targeted exome sequencing.

Two methodologies for targeted exome sequencing were assessed; SureSelect^{XT} (Agilent Technologies) and Nextera (Illumina). These methods are used to ‘capture’ or ‘enrich’ the DNA fragments of interest, in a process called library preparation, prior to sequencing. An overview of these two library preparation methods is shown in Figure 13. An Illumina MiSeq sequencing platform was then used to sequence the libraries generated for each participant.

iii.iv.i SureSelect^{XT} Library preparation

The SureSelect^{XT} system was used to capture genomic regions of interest using a custom designed probe set.

Custom Probe Design

A custom probe set was designed using the Agilent web based SureDesign software (<https://earray.chem.agilent.com/suredesign/>) to cover all coding regions and intron/exon boundaries of known OI genes (Table 4).

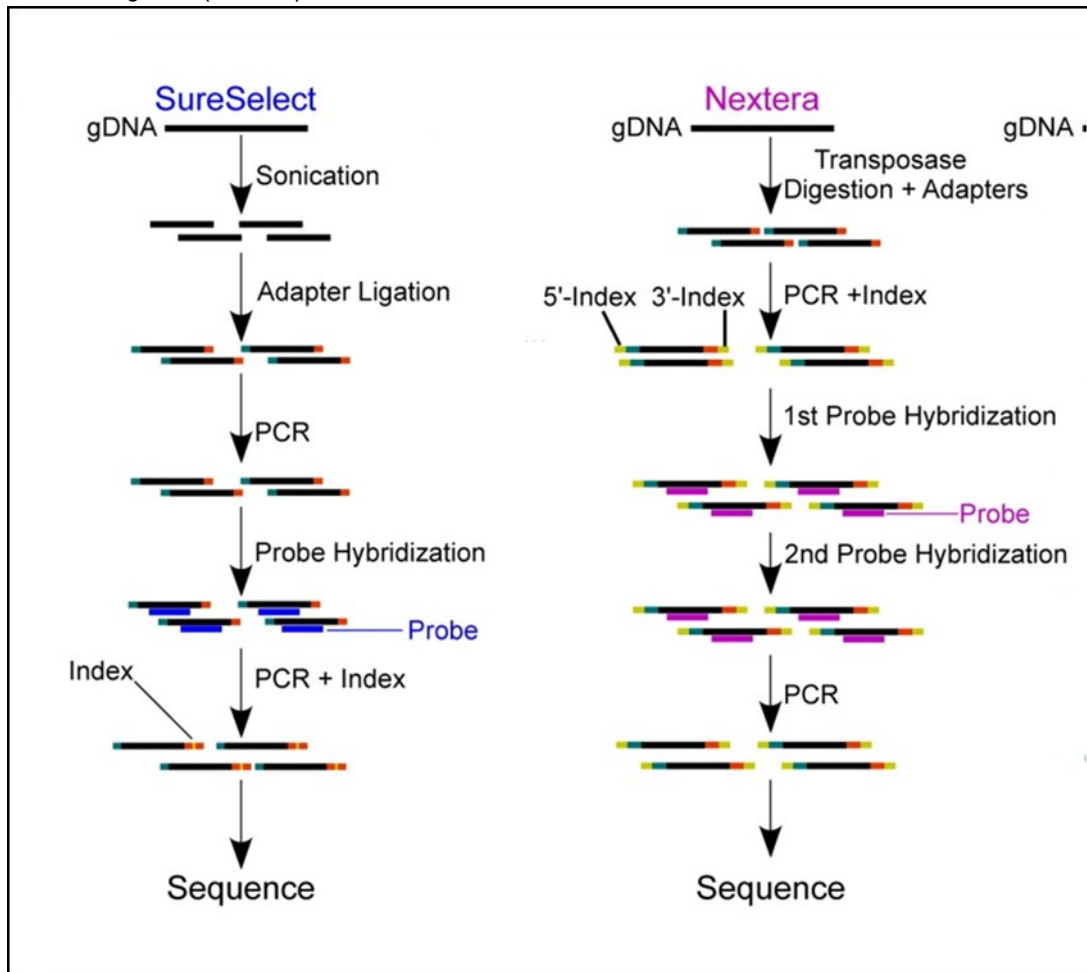


Figure 13 Comparison of SureSelect^{XT} and Nextera Library Capture Methods

Overview of SureSelect and Nextera Library Capture methods. Figure based on (Samorodnitsky, Datta et al. 2015) and is published under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License (CC BY NC ND).

The *COL1A1* and *COL1A2* genes were included in the design to confirm the absence of mutations in these genes in this cohort. The process of probe design involves creating a file (.BED format) containing a list of genomic co-ordinates corresponding to each coding region of the genes of interest. This file was uploaded to the SureDesign application and the parameters for probe design selected as follows: Flanking bases: 25, Density: X5, Masking: Moderately Stringent, Boosting: Max Performance.

Once the probe design was complete the resulting probe set was downloaded and inspected to identify any missed regions, the location of the probes and the gene structure.

Gene	Coding cDNA length	Coding exons	Reference Sequence	Chromosome location
Autosomal Dominant				
<i>COL1A1</i>	4392	52	NM_000088.3	17q21.33
<i>COL1A2</i>	4098	52	NM_000089.3	7q21.3
<i>IFITM5</i>	396	2	NM_001025295.1	11p15.5
Autosomal Recessive				
<i>BMP1</i>	2961	20	NM_006129.4	8p21.3
<i>BMP1</i>	85	1	NM_007799.3	8p21.3
<i>CREB3L1</i>	1557	12	NM_052854.3	11q11
<i>CRTAP</i>	1206	7	NM_006371.4	3p22.3
<i>FKBP10</i>	1749	10	NM_021939.3	17q21.2
<i>OSX/SP7</i>	1296	2	NM_152860.1	12q13.13
<i>P3H1</i>	2211	15	NM_22356.3	1p34.2
<i>PLOD2</i>	2277	19	NM_182943.2	3q24
<i>PPIB</i>	651	5	NM_000942.4	15q22.31
<i>SERPINF1</i>	1254	8	NM_002615.5	17p13.3
<i>SERPINH1</i>	1257	4	NM_001235.2	11q13.5
<i>TAPT1</i>	1701	14	NM_153365.2	4p15.32
<i>TMEM38B</i>	873	6	NM_018112.1	9q31.2
<i>WNT1</i>	1110	4	NM_005430.3	12q13.12
X-linked				
<i>PLS3</i>	1890	16	NM_005032.6	Xq23

Table 4 Location, size and reference sequence for genes in the Custom Probe Design

Any regions not covered by the initial design were processed through SureDesign without masking. This provided a further set of probes but with the potential to anneal to additional locations within the genome. Web Blat was used to assess how many times and how well each set of probes anneals to the genome. The best probe sets were manually selected and added to the initial probe design before finalising in SureDesign. For areas of 'difficult sequence', such as those with high GC content, the number of probe replicates was boosted in order to maximise the likelihood of probe capture of those regions. A summary of the SureSelect probe design work flow is given in Figure 14

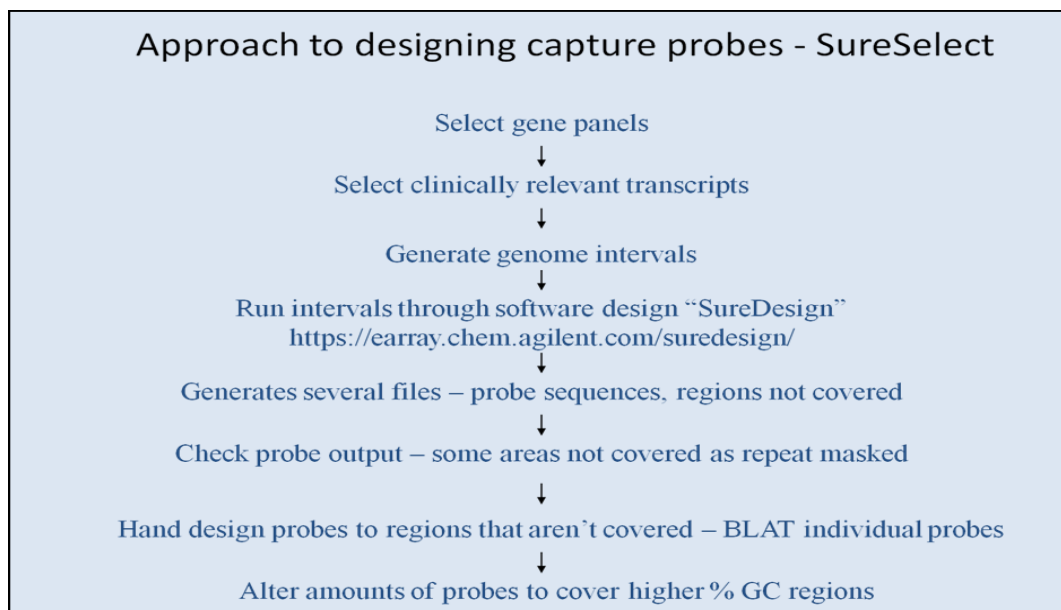


Figure 14 SureSelect Custom Probe Design Workflow

Summary of the workflow for custom probe design for Sureselect library preparation showing stepwise refinement to include regions of high GC content and regions masked during initial design process.

A completed probe design is listed in a tab delimited text file format as illustrated in Table 5 below.

TargetID	ProbelD	Sequence	Strand	Coordinates
chr11:27695606-27695831	BA_134634_000028	TTACTCACCTT...	+	chr11:27695599-27695718
chr11:27695606-27695831	BA_134634_000029	TCAGGGATCAG...	+	chr11:27695623-27695742
chr11:27695606-27695831	BA_134634_000030	ACTGAAACGTG...	+	chr11:27695647-27695766
chr11:27695606-27695831	BA_134634_000031	CAGCCCTTCTT...	+	chr11:27695671-27695790
chr11:27695606-27695831	BA_134634_000032	TCATGCCATAC...	+	chr11:27695695-27695814
chr11:27695606-27695831	BA_134634_000033	TAGACGCCAA...	+	chr11:27695719-27695838
chr11:27720928-27720952	BA_134634_000034	TGCCCGTCAG...	+	chr11:27720868-27720987
chr11:27720928-27720952	BA_134634_000035	ATCCTCAGCTA...	+	chr11:27720892-27721011
chr11:27722517-27722583	BA_134634_000036	AGCACCCAAG...	+	chr11:27722466-27722585
chr11:27722517-27722583	BA_134634_000037	GCGTGCGTTT...	+	chr11:27722490-27722609

Table 5 Example of a completed probe design for SureSelect library preparation.

SureSelect^{XT} Library Construction

DNA (>3 μ g) from each participant was 'normalised' to a concentration of 26.7ng/ μ l. The normalised DNA was then sonically sheared using a Covaris E220 sonicator to produce fragments between 150-200bp. Lastly end repair, adapter ligation and amplification were performed.

Successful library preparation was confirmed by electrophoresis to confirm a band of ~300bp and a smooth peak of fragments sizes was seen on scanning (Figure 15 and Figure 16).

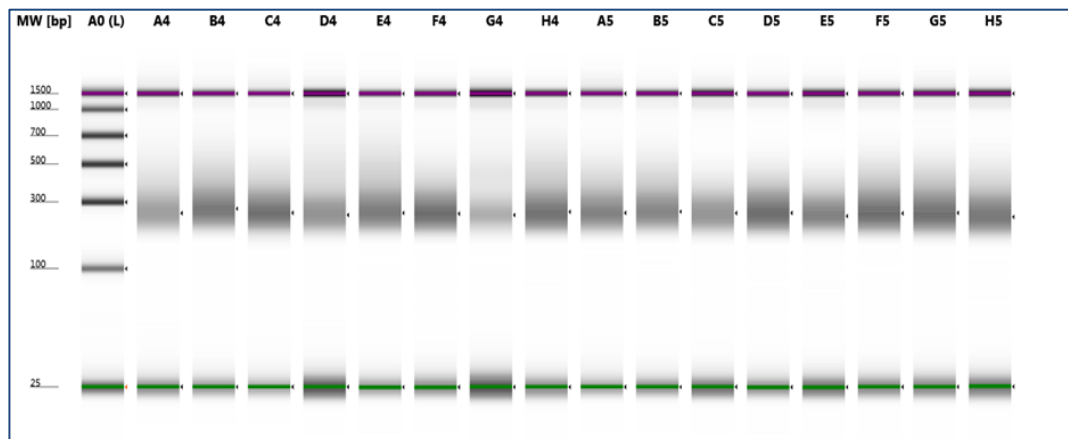


Figure 15 Gel image of DNA samples following library preparation

Each lane, marked A4-H4 and A5-H5, represents an individual patient sample. A size ladder is positioned in lane A0. Two standard markers are included at 1500bp (purple) and 25bp (green). Successful library preparation should produce a clear band of DNA at ~300bp.

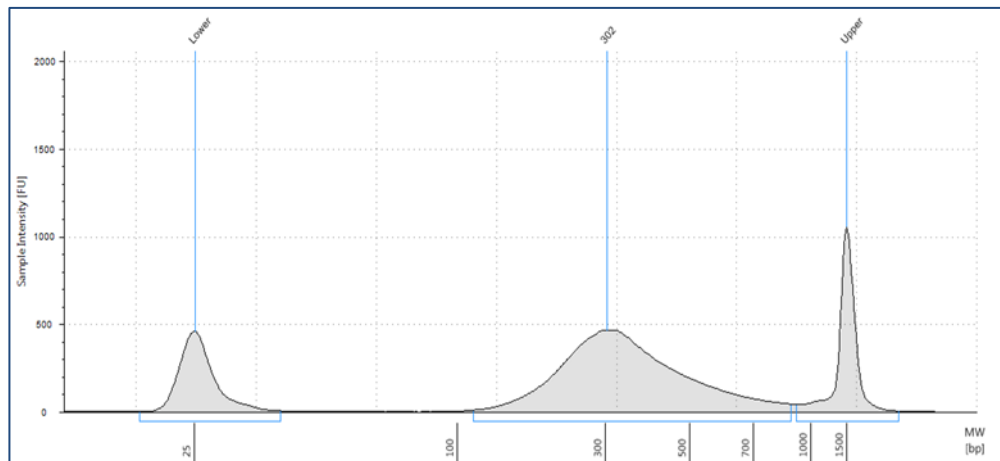


Figure 16 Scanned image of individual patient DNA following library preparation

The DNA was fragmented to ~300bp and standard markers are shown at 25bp and 1500bp

An average yield of half the input amount, eg for 2 μ g you would expect 1 μ g in total, was achieved. A minimum of 750ng was used for hybridisation/capture.

SureSelect^{XT} Hybridisation / capture

The SureSelect^{XT} target enrichment protocol combines the prepared DNA library with hybridisation reagents and the SureSelect Capture library consisting of the custom designed probes for the specific genes of interest. The samples were hybridised for 20 hours, then captured using magnetic beads. The final enriched library contained 10000-20000 pmol/l DNA with a maximum fragment size of ~300bp.

iii.iv.ii Nextera Library Preparation

Custom Probe Design

A custom probe set was developed using the Illumina DesignStudio Sequencing Assay Designer. In a similar process to the SureSelect design, genomic coordinates of each region of interest (coding regions and intron/exon boundaries) were used by the DesignStudio algorithm to optimize probe set design. Regions of poor coverage or those with high GC content are highlighted and these were manually reviewed and refined. An illustration of the completed design, as a tab delineated file, is shown in Table 6 below.

Probe ID	Chr	Start	Stop	Labels	Sequence	Design Warnings
720478	17	1674370	1674449	SERPINF1_4	GACTTGATCAGCAGCCCAGACATCCATGGTACC TATAAGGAGCTCCTTGACACGGTCACTGCCCCC CAGAAGAACCTCAA	
722847	17	1678493	1678572	SERPINF1_6	AGGTCTGTAGGGATAGGGGCAGGGTGGGGGGT GGATGGAGGGAGAGGATAGAGAAGCAAACAG GGTAGTGGGAATAAAA	Low Specificity
722082	17	1678393	1678472	SERPINF1_6	GAGGATTTCTACTTGGATGAAGAGAGGACCGTG AGGGTCCCCTATGATGTCGGACCCTAAGGCTGTT TTACGCTATGGCTT	
720243	8	22022896	22022975	BMP1_15UTR	TCCCTCGCCGCGCCCGCCAGCATGCCCGGC GTGGCCCGCCTGCCGCTGCTCGGGCTGCT GCTGCTCCCGCTCCC	Poor GC Content
722793	8	22022836	22022915	BMP1_15UTR	TCGCGATCGAGCAAGCAAGCGGGCGAGAGGAC GCCCTCCCCTGGCCTCCAGTGCGCCGCTTCCCT CGCCGCGCCCGCC	
723648	8	22031091	22031170	BMP1_2	CCTGGCTTCTTTTCTTTAGCTGCCTTTCTT GGGGACATTGCCCTGGACGAAGAGGACCTGAG GGCCTTCAGGTA	

Table 6 Completed probe design for part of the *SERPINF1* and *BMP1* genes for Nextera Library preparation

Nextera Library Preparation

Extracted DNA (>3µg) from each participant was cleaned (Ampure DNA clean up) and titrated down to a concentration of 5ng/µl. The library preparation consists of a fragmentation reaction, PCR amplification using known dual indices, library hybridisation to the custom designed oligos for the targeted regions and a second PCR amplification, with bead clean up between steps. Tagmentation is where the Nextera transposome fragments and tags the DNA with adapter sequences simultaneously, ready for the following PCR steps.

Following tagmentation there should be a broad distribution of DNA fragment size, ideally between 200-400bp. This was checked on an Agilent Technologies 2200 Tapestation using a High Sensitivity D1K tape. An example of successful tagmentation is given below in Figure 17:

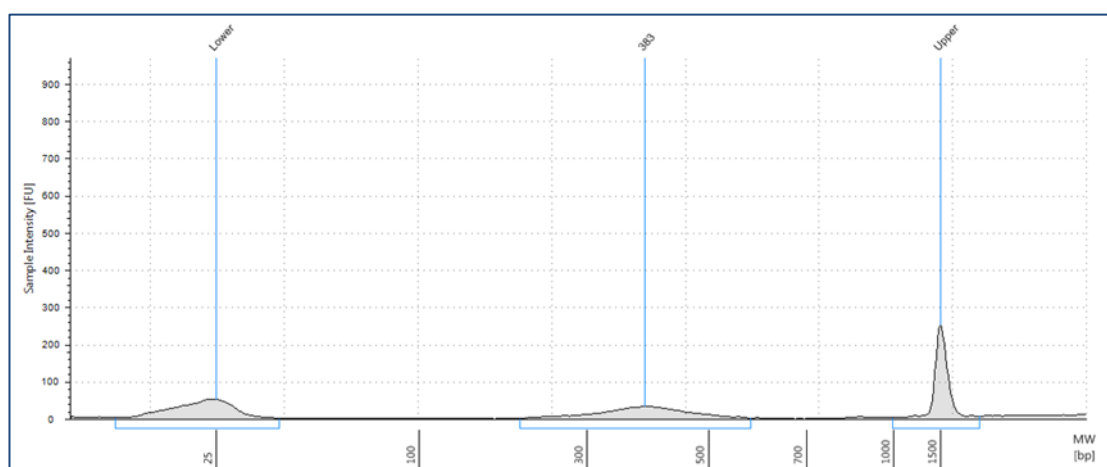


Figure 17 Tapestation analysis of DNA library preparation

The size of the tagmented DNA template should be between 200-400bp. An optimal profile would have a wide peak (as seen in above picture). A narrow peak would indicate that the DNA has been overly tagmented. Sample intensity appears low in comparison to SureSelect library preparation but this increases after subsequent PCR amplification.

Nextera Hybridization

The prepared DNA library was mixed with the custom designed capture probes and incubated at 58°C for between 90 minutes and 24hours. Streptavidin beads were then used to capture probes hybridized to the targeted regions of interest. This enriched library is washed and a second round of hybridization performed, which ensures high specificity for the captured regions. Following a sample clean up, the captured library was amplified via a 12 cycle PCR process. The library insert size was checked using the Tapestation to ensure it lies between 250-300bp.

iii.iv.iii Sequencing on the Illumina MiSeq.

Sequencing was performed on an Illumina MiSeq™ platform using the MiSeq™ Reagent Kit v2 performing 150bp paired-end reads. Sequencing using paired-end reads (sequencing both ends of a

DNA molecule) improves alignment to the genome particularly over repetitive regions. The detection of insertions or duplications is also improved.

The 'sequencing by synthesis' method used by Illumina uses fluorescently labelled dNTPs that are incorporated during each sequencing cycle. Imaging after each cycle allows base calling of the DNA molecule.

iii.v NGS analysis pipeline

Raw DNA sequence files obtained from the Illumina MiSeq™ platform were analysed through the pipeline developed at Sheffield Diagnostic Genetics Service (SDGS). This pipeline broadly adheres to the 'Best Practices' workflow by the Broad Institute (<http://www.broadinstitute.org/gatk/guide/best-practices>), but with some additional analysis and data-filtering steps added, some of which have been customised to accommodate local needs at SDGS.

The pipeline aligns reads to human genome build hg19, performs various QC checks, including the analysis of percent aligned reads and reads on target (i.e. reads that contain sequence that overlap the region of interest), and also depth of coverage. A minimum threshold of 30-fold read depth was set for exonic sequences and intronic sequences up to and including 5bp from the intron/exon boundary. A minimum threshold of 18-fold read depth was set for intronic sequences from 6bp to 25bp from the exon.

Variants were subsequently identified using HaplotypeCaller and annotated with Human Genome Variation Society (HGVS) nomenclature. Variants were filtered against a common polymorphism list which was compiled using NCBI dbSNP (Single Nucleotide Polymorphism Database), which contains data from large sequencing projects such as the NHLBI Exome Sequencing Project and 1000 Genome Project. In house reference files such as the polymorphism lists and preferred transcript list were also used.

The SDGS NGS analysis pipeline (v3.1.3) was used to take fastq files generated by the Illumina HiSeq through accepted best practices and produce filtered annotated variants for review, in the form of a Variant Call Format (VCF), a text file format containing a header followed by variant calls. The header contains information about the variant data and relevant reference sources (e.g. genome build version etc.) and defines the annotations used in the variant calls contained in the VCF file. In addition, the pipeline produces a number of QC files that allow assessment of the quality of the data.

The pipeline was contained within one script which handles the calling of other software required to process and analyse the sequencing data (sdgs_ngs_gene_panel_analysis.sh). The pipeline was run as a 'job' in the sun-grid-engine high performance computing environment using hg19 from NCBI as the reference genome.

The following steps were taken by the pipeline for analysis of patient's sequencing data.

- Check input arguments for validity
- Make output directory and start analysis log
- Run bwa align, sample, and sort into 2 bam files which are then merged

Clip overlapping reads
 Make regions of interest bed file
 Count total and failed reads
 Picard mark duplicates
 Calculate Duplicate stats
 GATK Indel realignment restricted to bed file
 GATK Haplotype Caller
 Filter VCF by bed file
 SnpSift annotation from dbSNP, COSMIC & HGVS
 Format the files required for annovar
 Run annovar (annotate variants with respect to functional consequences)
 Combine annotations
 Remove low quality calls and polymorphisms

A number of files were created by the pipeline, both results and QC files, which were used to check the quality of the run.

The bioinformatics pipeline uses open source software to align and call variants within the sequencing data. The versions of open source software used are listed in Table 7 below.

Software	Use	Version
bwa	Alignment	v0.7.15
samtools	Bam/Sam processing	v1.3.1
GATK	Indel realignment & Variant calling	v3.6
picard	Marking duplicates	v1.101
sambamba	Coverage calculations	v0.6.3
bedtools	Processing & manipulating bed files	v2.17.0
annovar	Annotating variants	2013-08-23 11:32:41 -0700
SnpSift	Annotating variants	3.3h (build 2013-08-20)

Table 7 Versions of open source software used in the bioinformatics pipeline (v3.1.3).

The SDGS bioinformatics pipeline is used to call and annotate sequence variants prior to analysis

Quality Control

A quality control report was generated by taking the *temp_run_report.txt* file that was generated for each sample and combining with others in the run to produce a run report. This file was used in determining the quality of the data for all samples in the run. The report contains the following columns for each sample:

tot_reads – Total number of reads
%_run – Percent of the run that is from that sample
Fail – Number of failed reads
PCR_dups & %PCR_dups – Total and percentage PCR duplicates
opt_dups & %opt_dups – Total and percentage Optical duplicates

lib_size – Estimated library size
med_read_length & 1st_quart_read_length & 3rd_quart_read_length – Median 1st and 3rd read length of non-failed non dups
Mapped & %mapped - Total mapped reads and percent mapped
proper_pairs & %proper_pairs - Total proper pairs and percent proper pairs
correct_orient & %correct_orient – Total correct orientation and percent correct orientation
med_insert & 1st_quart_insert & 3rd_quart_insert – Median insert size & 1st & 3rd quartiles.
small_insert & %small_insert - Number and % of read pairs with insert size below the expected insert size minus 50. The expected insert size for SureSelect libraries is 200 and for TruSight is 300.
large_insert & %large_insert - Number and % of read pairs with insert size above the expected insert size plus 50
broad_panel_on_target & %on_target - Number and % of not PCR duplicate reads mapping in the broad panel (strictly this is the number of not failed, not PCR duplicate mapped reads minus the number of not PCR duplicate reads that mapped off target) (% of reads in mapped category)
broad_panel_%5x ..10x ..15x ..20x ..25x ..30x ..35x ..40x ..45x ..%50x ..100x - % of broad panel covered by not failed not PCR duplicate reads at range of depths

A sample was flagged as poor quality if it failed any of the following criteria:

%mapped	≥80
%proper_pairs	≥70
%on_target	≥40
%30x	=100

Samples defined as poor quality were either rerun to obtain more sequence data or, if unlikely to generate sufficient sequencing data, a repeat blood sample and DNA extraction was performed.

iii.vi Mouse Models - TRAM2 Analysis

iii.vi.i Primer Design

Primer pairs for PCR amplification of coding regions of the *TRAM2* gene were designed using Primer3 software (<http://frodo.wi.mit.edu/primer3/>) (Table 8 below). All primers were checked to ensure that there are no common single nucleotide polymorphisms under the primer binding site, which can prevent primer annealing, using the Manchester National Genetics Reference Laboratory tool provided at URL <https://ngri.manchester.ac.uk/SNPCheckV3/snpcheck.htm>.

M13-tails were added to the 5' end of each primer to allow PCR by a standard sequencing protocol:

M13F1 (forward) - CAC GAC GTT GTA AAA CGA C

M13R (reverse) - CAG GAA ACA GCT ATG ACC

Exon	Forward Primer	Reverse Primer2	GC content
1	CCAGCGGCTAGAGGGTTAAG	AGGGGTACAGTGCACAAAGC	73.04%
2	ATGTAAAAGGGGCGGTTAGC	TGTACACAGATGGGGAAAAGG	44.21%
3	CTGAGGCTATAGCAAACGGG	GAGCTGCAGTTTCCTCCAAG	50.00%
4	ATGTTCAACCTTCTGTGGG	GGGTCCTTTGCAGTCATTTG	44.63%
5	TTTCTGTCAGGTGGGGACTG	CATCCACAATAGGCTCAGGG	59.41%
6	AGCTCCTCACAGGCAGAGTC	AGGAAGGGAGGTACCCCTG	55.76%
7	TGTGTGTTAAAGCAGCCCAG	TGCTAGGCACACAGAGGATG	48.01%
8	AAAGTGGGAGTGGAAAGCAG	AAAATAACGGGAGGGCACC	58.77%
9	CCCTTGAATCCTCTTCCCAC	GGCTGCTATCTCTTCTCCCC	52.83%
10	CTTGTTCTCCCCTGCACAAC	CACCCATCCCAGACAGACTT	63.00%
11	TCCATGTATAGCAGCTGGGC	CCCATTGCAAGACAGGTTTC	58.33%

Table 8 Primers designed for Sanger sequencing of the *TRAM2* gene

iii.vi.ii PCR reaction

PCR of *TRAM2* was undertaken using OneTaq® Quick-Load® 2X Mastermix which contains taq polymerase, dNTPs, MgCl₂ and buffers. There are two versions of the mastermix containing either standard or GC buffer. Standard buffer was used where the GC content of the primer was below 55%. For GC content 55-65% the GC buffer was used and for higher GC content an enhancer was added to improve annealing and amplification. The GC content of the *TRAM2* primers is listed in Table 8

	Standard exons	YELLOW exons	RED exons
	Vol (µl)		
OneTaq + Standard buffer	10		
OneTaq + GC buffer		10	10
GC Enhancer			3
Water	7	7	4
Primer @ 5pmol/µl	1	1	1
DNA @ 10ng/ µl	2	2	2

Table 9 Reagents used for *TRAM2* amplification by PCR

Each DNA sample was diluted to a concentration of 10ng/µl in water. Reagents were prepared for PCR amplification of *TRAM2* depending on GC content of primer sequences (Table 9). PCR cycling was undertaken according to Table 10.

Temp (°C)	Time (min)	Cycles
94	1	x1
94	0.5	x33
60	0.5	
68	1	
68	5	x1
15	10	x1

Table 10 PCR Cycling Conditions for *TRAM2* amplification by PCR

PCR products were visualised using on a 1.5% agarose gel using 5µl OneTaq PCR product run at 120V for 15 minutes.

The *TRAM2* primers were initially validated using two control DNA samples to ensure that PCR was successful and produced fragments of the expected size. A negative control, with all reagents but no DNA, was also included (Figure 18).

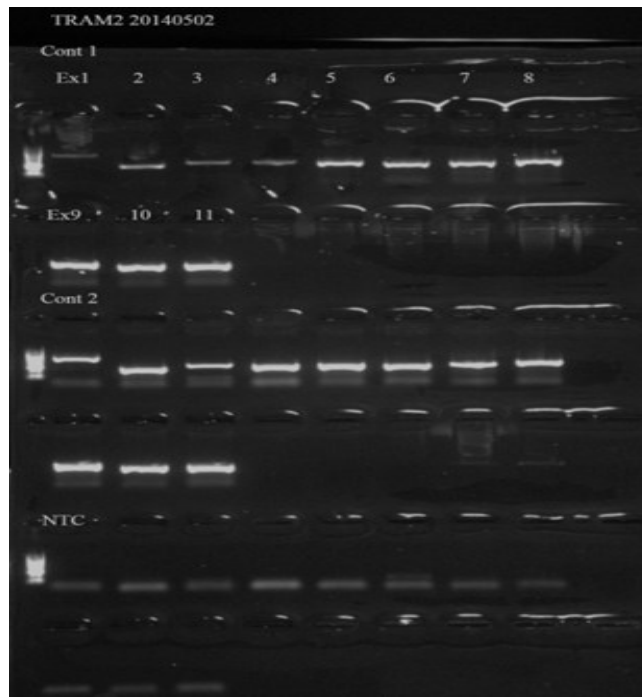


Figure 18 Electrophoresis of PCR products for the *TRAM2* gene

PCR products from normal control DNA (Cont 1 and Cont 2) and water blank (NTC) are shown.

iii.vi.iii Sanger Sequencing

DNA sequencing of PCR-amplified fragments of *TRAM2* was carried out using an ABI 3730 automated sequencer and Big Dye Terminator Sequencing protocol (Applied Biosystems, Foster City, USA; <http://www.appliedbiosystems.com>) (Table 11)

	Mastermix Reagents	PCR Cycling conditions															
BigDye Ready React v1.1 Terminator Mix	1µl	<table border="1"> <thead> <tr> <th>Temp.</th> <th>Time</th> <th>Cycles</th> </tr> </thead> <tbody> <tr> <td>95°C</td> <td>1 min</td> <td>1 cycle</td> </tr> <tr> <td>95°C</td> <td>15s</td> <td rowspan="3">25 cycles</td> </tr> <tr> <td>55°C</td> <td>10s</td> </tr> <tr> <td>60°C</td> <td>4min</td> </tr> </tbody> </table>			Temp.	Time	Cycles	95°C	1 min	1 cycle	95°C	15s	25 cycles	55°C	10s	60°C	4min
Temp.	Time				Cycles												
95°C	1 min				1 cycle												
95°C	15s				25 cycles												
55°C	10s																
60°C	4min																
Better Buffer	7µl																
H ₂ O	3.8µl																
Sequencing primer 1pmol/µl	3.2µl																
Total	15µl																
PCR product	5µl																

Table 11 Reagents and cycling condition for Sanger Sequencing of *TRAM2* PCR products

iii.vi.iv Analysis of Sanger Sequencing using Mutation Surveyor

Mutation Surveyor®, a DNA Sequencing software tool, was used for the analysis of Sanger sequencing generated by the ABI3730.

The software provides base calling, sequence alignment and assembly. The package includes automated download of annotated GenBank reference files and looks for non-correlation between the input sequence and the reference sequence allowing variant identification and annotation using Human Genome Variation Society nomenclature (HGVS).

iii.vii DDD Complimentary Analysis Project

The required subset of patient data was identified by searching the DDD database using Human Phenotype Ontology (HPO) terms. HPO provides a standard vocabulary to describe phenotypic information. The terms are developed with reference to Online Mendelian Inheritance in Man (OMIM), Orphanet, DECIPHER (Database of genomic variation and Phenotype in Humans using Ensembl Resources) and medical literature. There are currently more than 11,000 terms.

The following terms and their HPO numbers were used in this study:

- Abnormal susceptibility to fractures; Increased susceptibility to fractures (HP:0002659)
- Fractures of the long bones (HP:0003084)
- Fractures of vertebral bodies; Vertebral compression fractures (HP:0002953)
- Increased fracture rate; Recurrent fractures; Spontaneous fractures; Generalized osteopenia with pathologic fractures (HP:0002757)
- Pathologic fractures (HP:0002756)
- Generalized osteopenia; Osteopenia (HP:0000938)
- Severe osteoporosis (HP:0005897)
- Generalised osteoporosis with pathological fracture (HP:0005744)

The data returned for each patient included a list of potential candidate variants as identified by the DDD study, a complete list of unfiltered variants (VCF) and details of family structure and phenotypes (Table 12). In addition, the referring clinician details were provided for each patient.

File type	Description
CAP FILE	latest version of the proposal submission form.
SEARCH FILE:	list of updated HPO terms used to perform the searches
FAMILY ID FILE	maps DECIPHER IDs on to the person stable IDs for the proband, mother and father.
FAMILY RELATIONSHIPS FILE:	IDS for the proband, father and mother
CLINICAL FILE	all of the clinical observations, growth parameters and ontology term annotations that are available for each patient.
VARIANT FILES	VCF files for each patient and their parents (where applicable).
CANDIDATE VARIANTS:	high quality candidate variants in the proband across all genes in the relevant patient subset.
REPORTED SNV & INDEL FILE:	a flat text file containing SNVs and indels reported back to clinicians via DECIPHER
a flat text file containing SNVs and indels reported back to clinicians via DECIPHER	flat text file containing CNVs reported back to clinicians via DECIPHER
RESEARCH TRACK VARIANTS:	flat file containing variants of unknown significance, may contain additional variants of interest not present in the candidate file, such as low quality variants

Table 12 Data fields returned by the Deciphering Developmental Delay study for Complimentary Analysis Projects.

The candidate and research variants for each individual were first assessed at the gene level by looking at the function of the protein with emphasis on a potential role in bone disease. The sources of information interrogated for each gene is listed in Table 13.

Those with function related to bone metabolism were further assessed along with the provided clinical phenotypes of the patient and their family members. Finally, the full list of variants supplied in the VCF for each proband was interrogated. Any variants identified were assessed using the prioritisation strategies listed in section iii.ix below.

Source	Feature
GeneAtlas	Protein – function, cellular processes and pathways
	Associated disorders (if any)
	Cellular expression (tissue)
UniProt	Protein function
	GO terms for molecular and biological function
	Pathology and Biotech for involvement in disease
OMIM	Disease phenotype (if any)
	Gene description, function, molecular genetics and animal models.

Table 13 Source of terms used to assess the function of candidate genes.

Terms used to target genes with a potential role in bone metabolism

iii.viii Whole Exome Sequencing

iii.viii.i Personalis

Three participants were commercially sequenced by Personalis using their ACE Exome™ Assay. The Personalis pipeline uses the Burrows-Wheeler Aligner (BWA, version number 0.7.5a-r405) for alignment, the Broad Institute Genome Analysis Tool Kit (GATK) UnifiedGenotyper for variant calling and Variant Quality Score Recalibration (VQSR) to provide a quality score. AnnoL, a Personalis tool, is used to annotate variants. The annotation includes population frequencies from 1000 genomes and the NHLBI GO Exome Sequencing Project (ESP), those with a frequency of >1% in any of the populations are excluded. Further annotation includes data from dbSNP; predicted mutational impact using tools such as including SIFT and Polyphen2; known disease associations from databases such as OMIM and ClinVar; and data from pathway and network tools including Reactome (a database of reactions, pathways and biological processes) and the Molecular INTeraction database (MINT). Clinical information for each patient was provided to Personalis and this was used to rank variants, taking into account the similarity between presenting clinical features and known disease-gene associations.

Personalis returned a full data set of variants and this was further investigated at SDGS. Quality control was carried out, as recommended by the GATK best practice guidelines, using the following hard filters:

Values used for:	SNVs	Indels
QualByDepth (QD)	< 2	< 2
RMSMappingQuality (MQ)	< 40	
FisherStrand (FS)	> 60	> 200
HaplotypeScore	> 13	
MQRankSum	< -12.5	
ReadPosRankSum	< -8	< -20

Table 14 Quality Control filter settings used in analysis of Personalis sequence data

Population frequencies from the Exome Aggregation Consortium (gnomAD, version 0.3) were used in filtering and those with a frequency >5% in any subpopulation (excluding the 'other' population), or of >1% with a total allele count >1000 were removed. Variants with a reported frequency >1% in any of the 1000 genomes, NHLBI ESP or UK10K cohort populations were also excluded. After frequency filtering, variants were prioritised using the methods detailed in section iii.ix.

iii.viii.ii In house whole exome sequencing

The remaining participants, were sequenced at SDGS using QXT library preparation method and SureSelect Human All Exon V6 baits for target enrichment. The laboratory method for this is identical to that used for targeted exome sequencing of candidate genes using the custom designed probe set (section iii.iv.i and iii.iv.iii). Sequencing data from all exome participants and, when available, both parents was generated using a HiSeq 2500, using paired-end sequencing 2x100bp. The Human All Exon V6 probe set is designed to capture all protein coding regions of the genome and covers 99% of

sequence listed in the RefSeq, CCDS, GENCODE, HGMD_cds and OMIM_cds databases at an average coverage depth of 20x. Bcl2fastq (v1.8.4) was used to convert sequencing data to fastq files ready for mapping to the human genome. Fastq files were uploaded to Sapienta (v1.5.0) for mapping and variant calling.

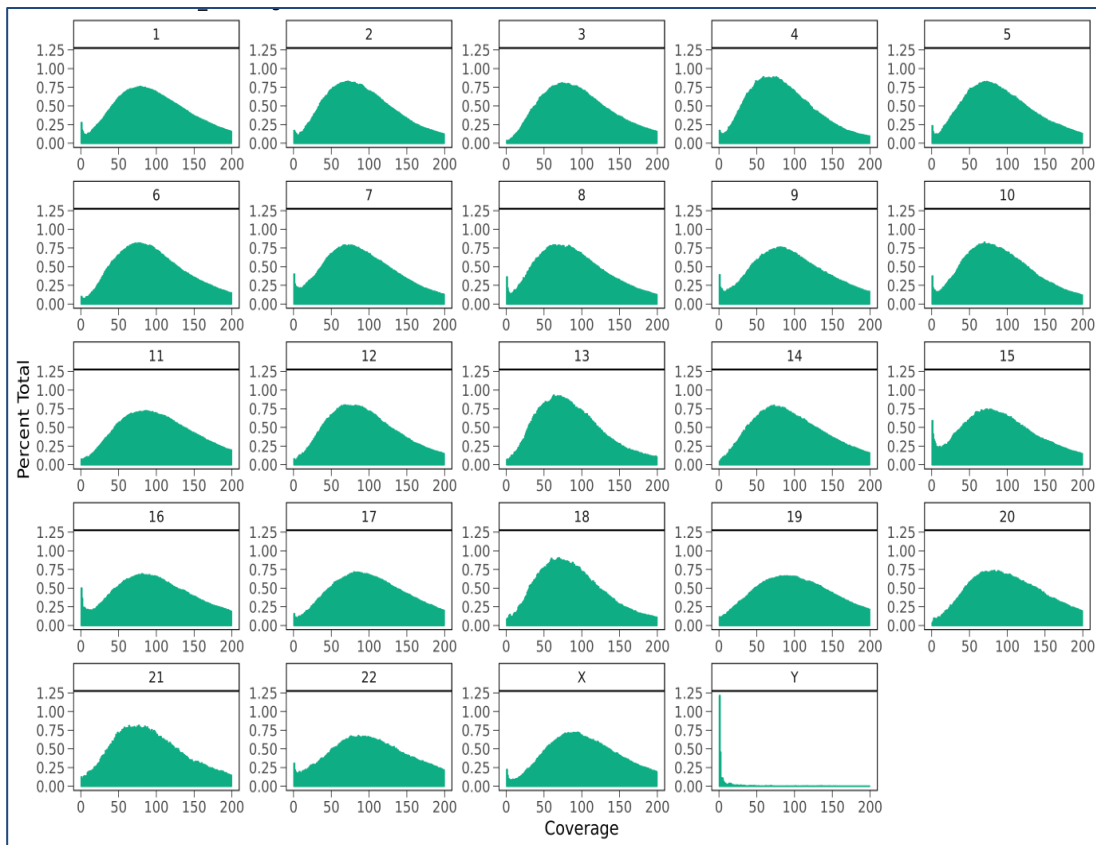


Figure 19 Depth of sequence coverage across all chromosomes for patient E using SureSelect Human All Exons V6r2

An average depth of 103x (standard deviation 62.1) and a median of 94x was obtained. The patient is female.

Sapienta does not provide a whole exome coverage summary; however, after mapping with BWA-MEM (a BWA algorithm recommended for high quality sequence queries, version 0.7.15) to GRCh37 a mean exonic coverage was obtained. Bases contribute to coverage if their quality is ≥ 30 and mapping quality is at least 10 for the read within which they are contained. Duplicate reads are excluded. “SureSelect Human All Exon V6 r2” design files were used to restrict coverage calculations. The coverage across each chromosome was assessed to ensure read depth was sufficient to allow confident variant calling, an example is given in Figure 19.

iii.ix Prioritisation of Sequence Variants from WES

The number of variants identified by whole exome sequencing can be large, averaging approximately 80,000, dependent on ethnic origin (Belkadi, Bolze et al. 2015). Filtering, analysis and interpretation of this data are challenging. An initial step for variant filtering in rare diseases was to exclude those found with significant frequency in the general population. In this study, information available in public databases from the following projects was used to remove variants with a minor allele frequency (MAF) above 1%:

- Exome Aggregation Consortium (gnomAD, <http://gnomAD.broadinstitute.org/>)
- NHLBI Exome Sequencing Project (ESP, <http://evs.gs.washington.edu/EVS/>)
- Genome of the Netherlands (GoNL, <http://www.nlgenome.nl/>)

Even when data was filtered to remove common polymorphisms, several hundred potentially pathogenic variants remained. Three different but overlapping strategies have been employed in this study to further filter and prioritise variants to identify the most interesting candidates.

iii.ix.i Strategy 1: An inheritance hypothesis

Variants were filtered using an inheritance hypothesis-based approach; this could be an autosomal recessive, autosomal dominant, X-linked or Y-linked model. Analysis of trios (proband and parents) was particularly useful for this approach and allowed detection of *de-novo* variants, which if present in a gene of interest can be indicative of pathogenicity.

Filtered variants were then ordered by predicted effect on protein function. Loss of function, including nonsense variants and shifts in the protein coding reading frame (frameshift), are in general more likely to be damaging than missense variants. Variants that affect the conserved splice donor or splice acceptor sites are also likely to have an effect on protein function. For small insertions or deletions, these can be ordered on whether they cause splicing disruption, in-frame protein changes or a frameshift.

The resulting list of variants was manually reviewed taking into account published literature, any known disease-gene interactions, metabolic and signalling pathways as well as any available gene function information and annotations (Figure 20).

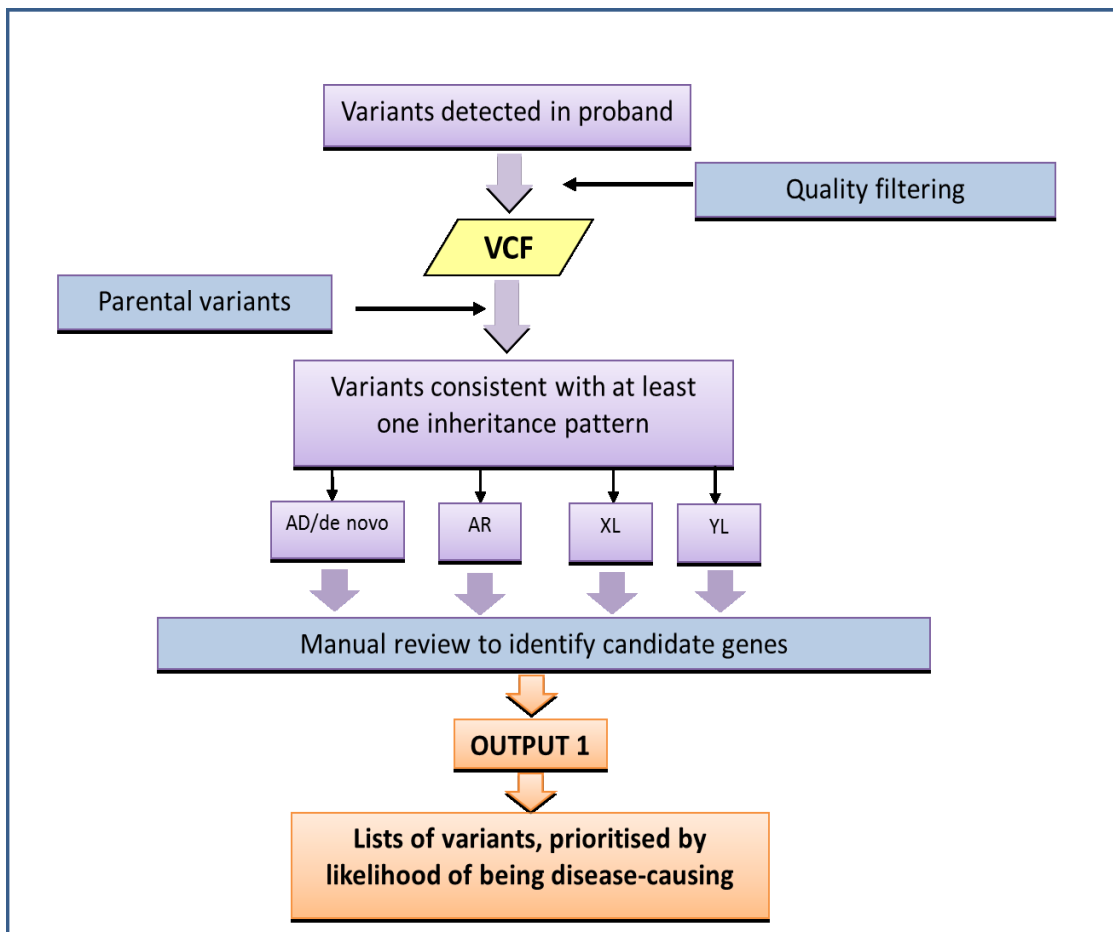


Figure 20 Variant prioritisation Strategy 1

Variants identified by exome sequencing are filtered on an inheritance based hypothesis model.

iii.ix.ii **Strategy 2: Exomizer**

Inheritance driven analysis of trio data as described in strategy 1 is a powerful tool to prioritise variants but there may still be many candidate variants remaining. For different pedigree structures, such as duos or singletons, inheritance filtering strategies are less useful. Each genome carries approximately 100 loss of function variants (MacArthur, Balasubramanian et al. 2012) requiring the use of many different approaches to identify disease-associated genes. One such approach is Exomizer, a phenotype driven tool that uses clinical data (in the form of HPO terms), model organism phenotype data and random-walk analysis (that uses direct and indirect protein interactions in biological networks, protein expression and biological similarities) to prioritise variants generated from whole exome analysis (Smedley, Jacobsen et al. 2015).

Exomizer is a Java program that uses VCF files to first filter then prioritise variants. The filtering steps include variant annotation relative to the University of California Santa Cruz (UCSC) reference genome hg19, removal of variants above a user defined MAF, removal of variants below a specified quality score and pedigree filtering if appropriate. Prioritisation is undertaken by a number of different methods including protein-protein interaction, comparison of patient phenotype and those in human disease databases and model organisms, and predicted effect on protein structure. Exomizer calculates both variant and gene based scores and combines these to generate a final combined

score that is used to rank the variants (Figure 21). The top 20 variants ranked by Exomizer were subsequently manually reviewed.

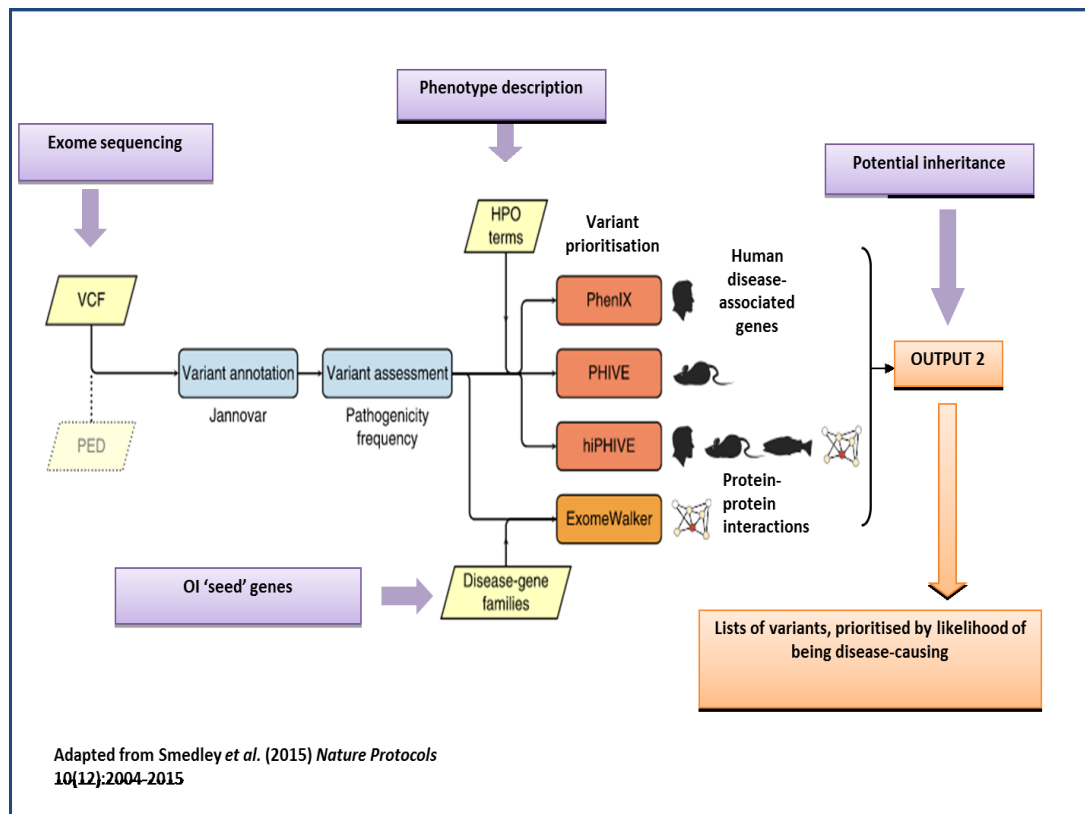


Figure 21 Variant Prioritisation Strategy 2

Strategy 2 uses Exomizer, a phenotype driven tool that uses clinical data (in the form of HPO terms), model organism phenotype data and random-walk analysis of protein-protein interactions to prioritise variants generated from whole exome analysis

iii.ix.iii **Strategy 3: Targeted Gene Approach**

The hypothesis for this strategy was that new pathways can be established in OI by identifying novel variants in genes already established in known diseases or bone phenotypes (Figure 22). A manually curated candidate gene list was developed by extensive database and literature review using the following search criteria:

- All genes known to be associated with OI and OI-like phenotypes
- All genes known to be associated with skeletal dysplasia and/or short stature
- Genes within BMD loci with genome-level significance detected in Genome Wide Association Studies
- Mouse models with reduced BMD, brittle bones, abnormal mineralization or reduced mineralization.

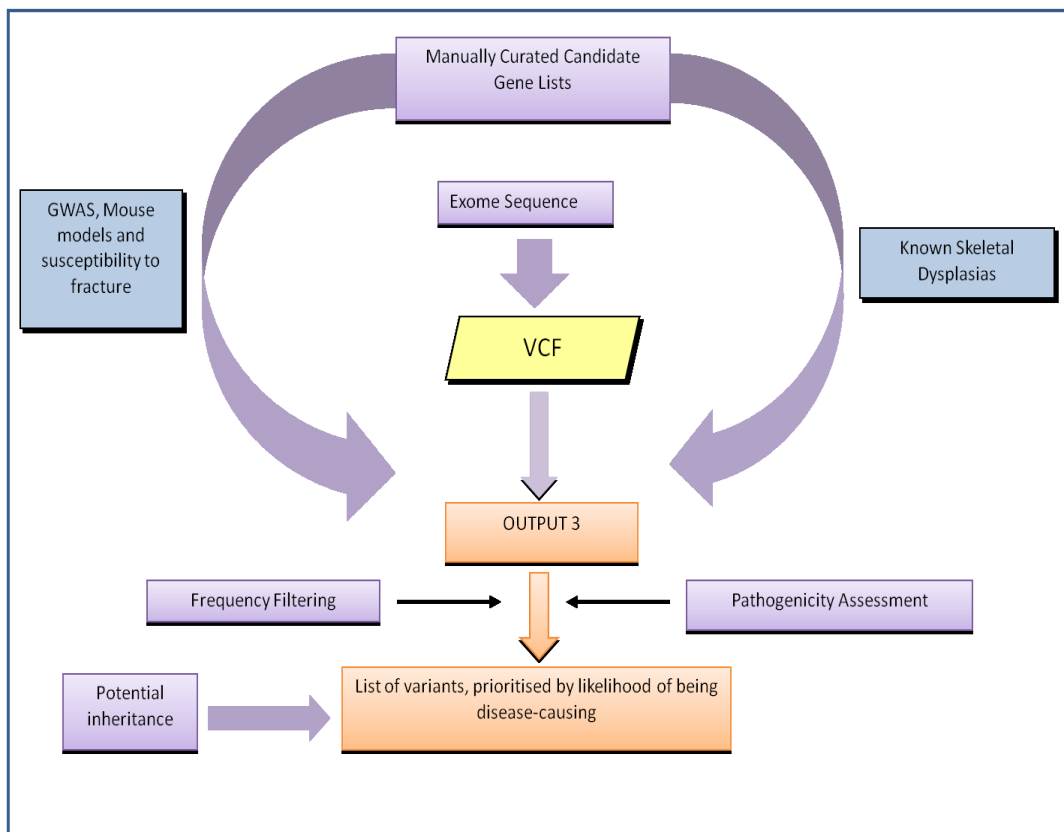


Figure 22 Variant prioritisation strategy 3

An extended list of candidate genes was developed using a combination of those described in association with skeletal disorders in man, bone changes in mouse models and those association with reduced bone mineral density in GWAS studies

Finally, genes identified using HPO terms related to bone fragility and increased fractures (Table 15) were also added. The final list of 635 targeted genes is given in the appendix, page 250.

HP:0003023	Bowing of limbs due to multiple fractures
HP:0005744	Generalized osteoporosis with pathological fracture
HP:0002757	Recurrent fractures
HP:0002659	Increased susceptibility to fractures

Table 15 HPO terms used to generate an extended target gene list used in variant prioritisation

The lists generated from strategies 1-3 above were combined to give a final list of prioritised variants. Variants in HLA and Mucin genes were removed from the variant list and, finally, the whole cohort dataset was reviewed, alongside data from the DDD study, to identify potential disease-related genes or variants enriched in this cohort in comparison to control populations.

iii.x Assessment of sequence variants

Sequence variants identified in all parts of this project were assessed for potential significance using a number of web based resources that explore possible modes of pathogenicity and sources of information where the variant may have been previously documented. The assessment tools used include:

iii.x.i Population Frequency

The population frequency in the following datasets was interrogated.

- Exome Aggregation Consortium (gnomAD) <http://gnomAD.broadinstitute.org/>
- NCBI dbSNP build 144 through Ensembl e83 <https://www.ncbi.nlm.nih.gov/projects/SNP/>
- NCBI ClinVar <https://www.ncbi.nlm.nih.gov/clinvar/>
- NHLBI GO Exome Sequencing Project (ESP) <http://evs.gs.washington.edu/EVS/>
- The Genome of the Netherlands Consortium (GoNL) <http://www.nlgenome.nl/>

Variants listed as present in >1% of any population were excluded from further analysis

iii.x.ii Web-based Literature Search

Resources used for the literature search included Google, Google Scholar, PubMed and the following Locus Specific Databases:

- The Human Gene Mutation Database (HGMD) (<http://www.hgmd.cf.ac.uk/ac/index.php>)
- Osteogenesis Imperfecta and Ehlers Danlos Syndrome Variant Databases <http://www.le.ac.uk/ge/collagen/>

Where a variant was identified as being previously reported, the available evidence for that classification was reviewed. This included assessment of *in-silico* variant predictions, any supplied pedigree/segregation data, population frequency data and any available functional studies.

iii.x.iii Variant Effect Predictors:

There are a number of tools available that predict the effect of missense changes on protein structure and function. These can be broadly assigned to one of three categories depending on the strategies used for prediction: machine learning, evolutionary conservation and protein sequence and structure. An outline of the tools used from each category is as follows.

Machine learning tools:

Machine Learning methods – learn patterns that can be applied to new scenarios. Some combine predictions from different sources and match the variant to the best fitting category. The quality of

their predictions is dependent on the information that the programs learn from and can be a source of false pattern recognition particularly if the program over trains.

SNPs&Go (<http://snps.biofold.org/snps-and-go/snps-and-go.html>) Trained on the effect of any substitution in association with adjacent residues and uses functional gene ontology.

MutPred (<http://mutpred.mutdb.org/>) This tool trains on HGMD and Swiss-Prot data. It gives an overall probability of a variant being deleterious based on 14 properties relating to protein structure and dynamics along with predicted functional properties alongside evolutionary conservation score.

HANSA (<http://www.cfd.org.in/HANSA/>) Trained on combination of position specific, structural and amino acid features (10 properties)

Evolutionary Conservation based tools:

These use multiple sequence alignments to assess the conservation of specific amino acid residues across species. Analysis against at least 5 mammals from different orders, plus chicken, frog and fish (zebra/fugu) was used where available.

Panther (<http://www.pantherdb.org/tools/csnpScoreForm.jsp>) Aligns sequence to protein subfamilies within its own library collection and calculated a probability of being pathogenic based on variation with the alignment

PROVEAN http://provean.jcvi.org/genome_submit_2.php?species=human retrieves scores from a pre-computed database that clusters sequences to give a predicted effect on protein translation and effect on function.

SIFT http://provean.jcvi.org/genome_submit_2.php?species=human Uses multiple sequence alignments generated from a BLAST search for similar sequences to calculate a probability score.

Mutation Assessor <http://mutationassessor.org/r3/> Uses multiple sequence alignments based on functional specificity.

Protein sequence and structure based tools:

The predictions from these tools are based on the severity of the amino acid substitution and the potential structural implications based on the surrounding residues.

PolyPhen2 <http://genetics.bwh.harvard.edu/pph2/> This is a supervised learning tool that assess evolutionary conservation and impact on functional region. The HumVar output is recommended in assessment for Mendelian diseases as this algorithm is trained on the differences between human disease-causing and neutral non-synonymous variants.

iii.x.iv Potential Splice changes

Intronic, synonymous and missense variants were assessed for their potential effect on splicing using the following tools:

- SpliceSiteFinder (<http://www.umd.be/HSF/>)
- MaxEntScan (<http://genes.mit.edu/burgelab/maxent/>)
- NNSplice (http://www.fruitfly.org/seq_tools/splice.html)
- GeneSplicer (<http://cbcb.umd.edu/software/GeneSplicer>)
- Human Splicing Finder (<http://www.umd.be/HSF/>)

Interpretation was based upon the consensus of at least three of the above programs and generally needs these to show a change of +/-10% in splicing efficiency to be considered to be of potential significance.

iii.x.v Severity of Amino Acid substitution

For missense variants data from Russell, the Grantham distance and BLOSUM scores from Alamut were used to give an overall idea of the difference in properties between the wildtype and mutant amino acids. Those with greater differences are most likely to be pathogenic.

Although there are a number of tools available to aid interpretation of variants, it was important to note that these are only predictions. Further studies, such as segregation and/or functional analysis may be required to draw a definite conclusion as to the significance of the variant.

iii.xi Method for Electron Microscopy

Punch biopsies of non-lesional skin were obtained from the upper inner arm. Each biopsy was cut into 4 columns and immediately fixed using 3% glutaraldehyde in 0.1M phosphate buffer at pH7.4. Samples were post-fixed in 2% aqueous osmium tetroxide, dehydrated through alcohol and acetone and then embedded in 812 epoxy resin in a flat bottomed BEEN® specimen embedding capsule.

Semi-thin sections (0.6µm) were produced on an Ultracut E ultramicrotome using a Diatome Histoknife. These were stained with 1% Toluidine blue in 1% sodium tetraborate at 120°C. Mid gold interference coloured sections were produced using a Diatome ultraknife and staining with uranyl acetate and Reynold's lead citrate.

Sections were photographed at step magnifications on a Philips 400 transmission electron microscope using an AMT 16 megapixel mid mount camera. Each section was screened for the presence of collagen 'flowers'. Collagen fibril diameters in the reticular dermis were measured manually (20000X magnification) using the camera software and mouse (Figure 23).

The mean diameter of 200 fibres was compared to the expected measurement for the age of the individual (Table 16). Fibroblasts and elastic fibre morphology were initially assessed at low magnification. The presence or absence of ER dilation, a feature providing evidence of stress, was assessed for each fibroblast identified (up to 20,000X magnification).

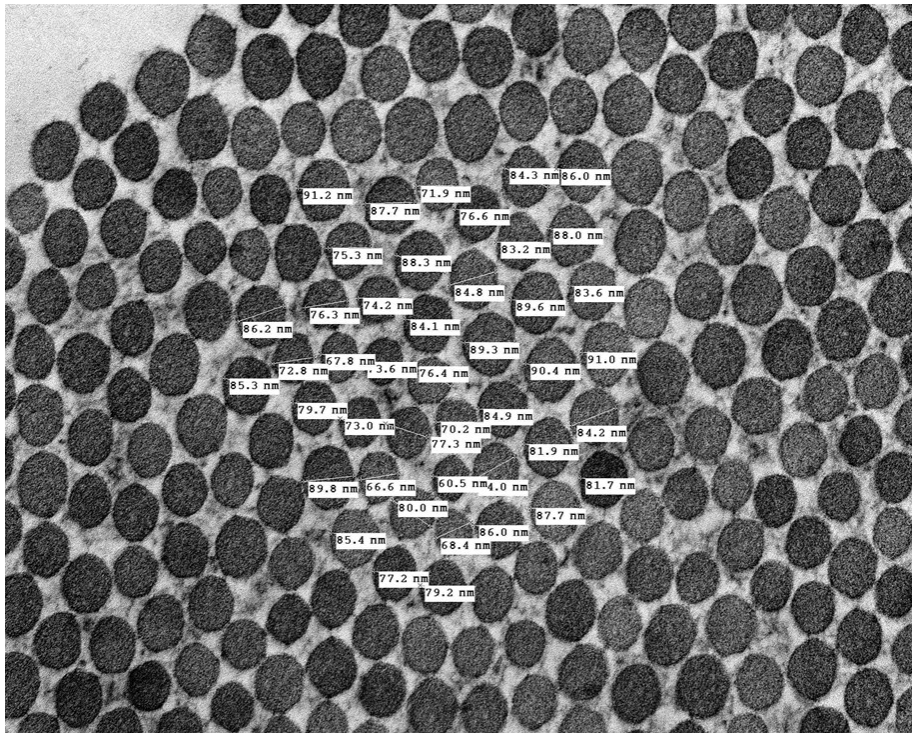


Figure 23 Electron microscopy of med-reticular dermis (20,000X magnification)

The mean diameter of collagen fibres (nm) was assessed by measuring 200 fibres at a magnification of 20,000.

Life Stage	Expected collagen fibre diameter (nm)
Fetus	70
Child/Adult	80-90
Elderly	70

Table 16 Expected diameter of collagen fibres in each life stage

Figures taken from Stewart et al (Stewart 1995)

iv RESULTS: TARGETED EXOME SEQUENCING

iv.i Patient recruitment

Eighteen families were recruited for targeted exome analysis (participant ID 001-018). Eight further families were recruited for whole exome sequencing (participant ID A-H). An additional cohort of *BMP1* patients were identified during the course of this study (patients I-K) following introduction of *BMP1* gene analysis into the diagnostic pathway following results for the NGS target panel part of this project. The presenting clinical details for each recruited participant are listed in Table 17

Participant ID	Clinical History
Targeted Exome	
001	Crush fracture of spine and generalised osteopenia
002	Family history of early onset osteopenia and osteoarthritis. Hypermobile
003	Multiple fractures and osteopenia
004	Dislocations, fractures, hearing loss in one ear. ?Blue sclerae. Other family members clinically affected.
005	Two fractures, one with no obvious trauma
006	Skull fracture, blue sclerae, Wormian bones.
007	Severe vertebral fracture and collapse, short stature, facial features suggestive of OI. Mother has similar features.
008	Numerous scars, history of >10 fractures
009	Hypermobile, fractures.
010	Clinically has OI type I, >30 fractures
011	Fractures, osteopenia
012	Antenatal fractures, autism, severe developmental delay
013	Atypical OI, severe scoliosis, grey sclerae, very hypermobile, no tissue fragility.
014	Recurrent fractures and osteopenia, characteristic facies of OI
015	Hypermobile. Metabolic bone disease. ? Osteopetrosis
016	OI type I
017	Fractures. Dysmorphic, developmental delay,
018	Multiple dislocations, easy/spontaneous bruising, blue sclerae, fractures, dental problems.
Whole Exome	
A	Cystic hygroma/increased nuchal translucency, moderate intellectual disability/mental retardation, fractures, flat or prominent occiput, dysmorphic facies, low rotated ears, downslanting palpebral fissures, thin upper lip, squint, atrial septal defect, ventricular septal defect. Wormian bones.
B	Multiple fractures, proximal shortening, scoliosis, possible contractures
C	Prematurity - 32 weeks, severe intellectual disability/developmental delay, fine motor delay, gross motor delay, speech delay, autism spectrum

	disorder/autistic features, scoliosis, failure to thrive, fractures - vertebral and tibial fractures, frontal bossing/prominent forehead, dysmorphic facies, low/posteriorly rotated ears, abnormality of hairline, hypertelorism, upslanting palpebral fissures, myopia, optic atrophy, hair hypoplasia, low immunoglobulin levels, short stature, bilateral optic pallor.
D	Dysmorphic, developmental delay, hypotonic, muscle weakness, short stature, osteopenia, fractures, high arched palate, frontal bossing, bilateral club feet, low posteriorly rotated ears.
E	Sparse falling out hair, blue sclerae, hypertelorism, delayed fontanel closure, club foot, crush fracture vertebrae, rib fracture, bone fragility.
F	Multiple fractures, blue sclerae, dentinogenesis imperfecta, hypermobile joints. Mother is hypermobile with fractures following insignificant trauma.
G	Failure to thrive, macrocephaly, persistent large anterior fontanelle, multiple fractures, dysmorphic features. Heterozygous mutation in <i>P3H1</i> c.1080+1G>T
H	Clinically osteogenesis imperfect with crush fracture vertebrae and multiple long bone fractures, scoliosis
BMP1	
I	Ten low trauma long bone fractures, white sclerae, normal hearing. No evidence of dentinogenesis imperfecta.
J	Three low impact fractures, grey sclerae, hypermobile fingers and no evidence of dentinogenesis imperfecta
K	High bone mass phenotype

Table 17 Clinical Details provided for recruited individuals

Eighteen families were recruited for targeted exome analysis (participant ID 001-018). Eight families were recruited for whole exome sequencing (participant ID A-H). A further cohort of BMP1 patients were identified during the course of this study (patients I-K)

There have been no participant withdrawals during this research project.

iv.ii Validation of Targeted Exome NGS Panel

To assess the relative merits of SureSelect vs Nextera panel design and library capture methodologies for targeted exome sequencing, the sequence coverage for each region of interest was assessed for each method. To achieve this, five DNA samples were run on both the SureSelect and Nextera panels. The number of region without adequate sequence for each method is shown in Table 18

	Number of regions without adequate sequence coverage.	
Sample Number	SureSelect	Nextera
1	2	17
2	0	14
3	0	13
4	0	5
5	0	15

Table 18 Comparison of regions without adequate sequence coverage for SureSelect vs Nextera library preparation in five control DNA samples.

Overall coverage for the Nextera panel was considerably lower than SureSelect for all samples. Visualisation of coverage for each method was undertaken and clearly demonstrated poorer sequence coverage across the entire panel. An example is shown in Figure 24.

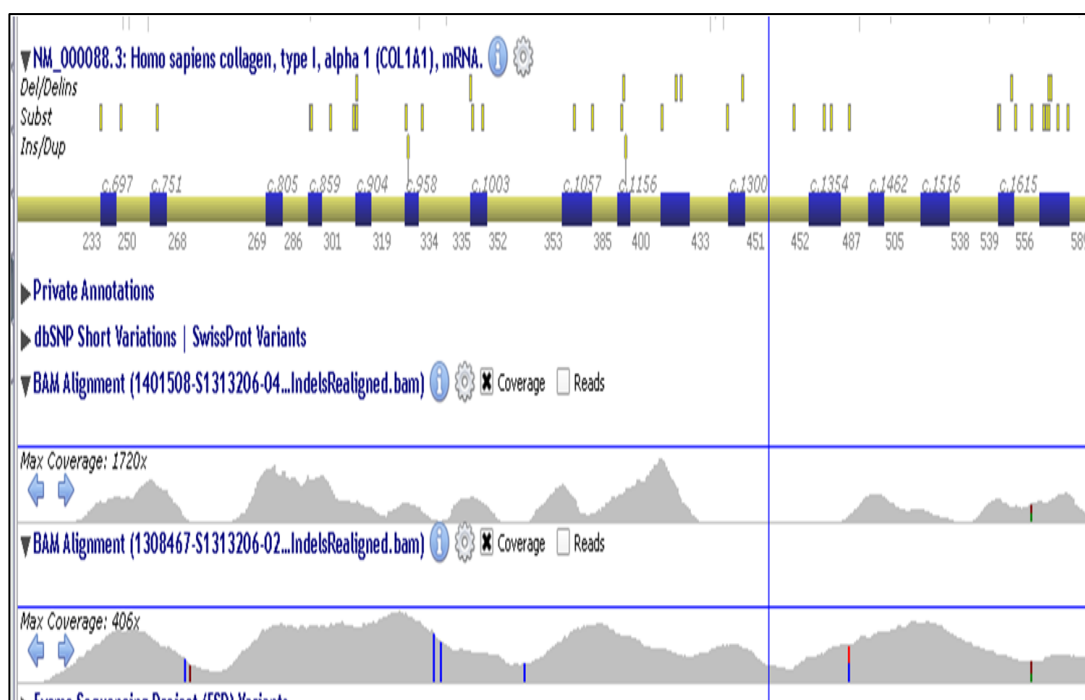


Figure 24 Visualisation of sequence coverage for SureSelect and Nextera methodologies

The number of sequence reads for the COL1A1 gene for a single control sample is shown. The top panel indicates the gene with exons coloured blue. The middle and bottom panels show the number of reads in each location combined to form peaks (coloured grey). The middle panel shows coverage with Nextera, the bottom with SureSelect.

One explanation for the reduced coverage of the Nextera panel in comparison to SureSelect is the difference in the number of probes produced for each region by the two design methodologies. SureSelect creates a number of overlapping probes for each region, increasing the likelihood that this

region will be captured whereas in comparison, the number of probes designed by Nextera is more limited, as illustrated in Figure 25.

The conclusion was that SureSelect gave superior coverage for the genes of interest in this study and was chosen as the method to evaluate further. The next step was to look at the ability of the SureSelect panel and bioinformatics analysis pipeline to correctly identify variants within the sequence generated.

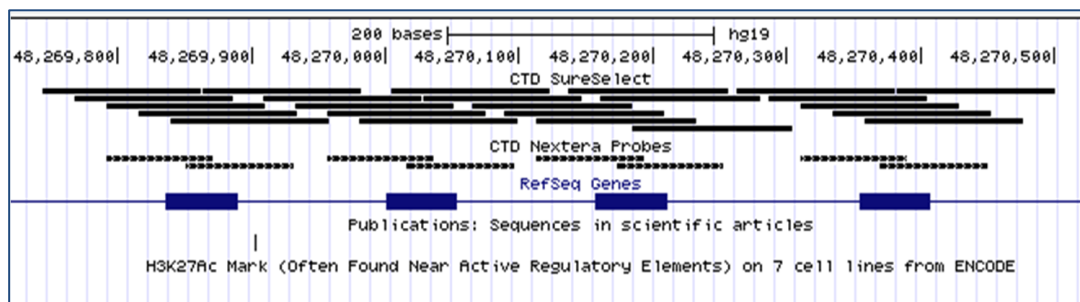


Figure 25 Comparison of the number of probes designed by SureSelect and Nextera

Coverage of four exons of the COL1A1 gene is shown.

SureSelect probe= black solid line Nextera Probe =black dashed line COL1A1 exon = blue rectangle

Variant lists from seven control samples that had been previously analysed using Sanger sequencing, the 'gold standard' method of sequencing, were compared for 14 genes. A total of 164 variants were available and the number in each gene are summarised in Table 19 below.

Gene	Variants	Gene	Variants	Gene	Variants
COL5A1	49	ALPL	14	PLOD1	1
COL1A1	21	CRTAP	5	COL5A2	14
COL1A2	24	COL3A1	12	SERPINH1	4
SP7	2	PPIB	1	P3H1	5
SERPINF1	11	FKBP10	1		

Table 19 List of genes and number of variants used to validate SureSelect targeted exome panel.

A wide region of interest (ROI) of 150bp+/- from each exon was employed to maximise the number of variants included in the process. All variants were identified in the targeted exome NGS results. No false positives or false negatives were identified. The SureSelect and MiSeq platform were considered to be performing well compared to Sanger sequencing and were therefore chosen as methods for the targeted exome sequencing of patient samples.

iv.iii SureSelect Targeted Exome Patient Results

All participants were screened for known OI genes using the targeted exome panel. Each identified variant was compared against the manually developed and curated database containing details of known common benign polymorphisms. Variants listed in the database are considered unlikely to cause OI, primarily due to their frequency in the general population (allele frequency >1%). Those listed in the database were excluded from further analysis.

Additional manual assessment of variants to remove further polymorphisms not listed in the database was undertaken. Regions of homopolymer, that can give false positive variant calls, were also removed. A total of 21 variants requiring pathogenicity assessment remained (Table 20).

Patient ID	Gene	Variant	Zygoty	dbSNP ref	Read count	Conclusion
1	<i>TMEM38B</i>	c.113-7A>G	het	rs186864213	1997	SNV
1	<i>COL1A1</i>	c.1350G>T;p.(E450D)	het	.	15	False Positive
1	<i>CREB3L1</i>	c.1194C>T;p.(S398S)	het	rs199639528	6929	SNV
2	<i>TAPT1</i>	c.837G>T;p.(M279I)	het	.	20	False Positive
3	<i>FKBP10</i>	c.590A>G;p.(K197R)	het	rs34764749	4526	SNV
3	<i>P3H1</i>	c.1322A>G;p.(D441G)	het	rs113593896	10705	SNV
3	<i>TAPT1</i>	c.613-3delT	het	.	33	False Positive
3	<i>BMP1</i>	c.2134G>A;p.(G712S)	het	rs117159093	6824	? UV
4	<i>COL1A1</i>	c.4070T>C;p.(L1357P)	het		5515	? UV
4	<i>CREB3L1</i>	c.962+18G>A	het	rs2288249	4077	SNV
4	<i>CREB3L1</i>	c.1231G>A;p.(A411T)	het	rs35652107	8155	SNV
5	<i>IFITM5</i>	c.80G>C;p.(G27A)	het	rs57285449	1309	SNV
5	<i>COL1A2</i>	c.71-17dupT	hom	rs144776919	2147	SNV
6	<i>P3H1</i>	c.1720+3G>A	het		1404	No splice effect predicted
8	<i>BMP1</i>	c.1112G>A;p.(R371H)	hom	rs145284541	5480	SNV
10	<i>SERPINF1</i>	c.85-13T>G	hom	rs199735427	1428	No splice effect predicted
11	<i>PPIB</i>	c.170T>C;p.(V57A)	het		1770	Likely SNV
12	<i>COL1A1</i>	c.1111G>T;p.(G371C)	het		4427	N:#
13	<i>CRTAP</i>	c.471+1G>A	het		46	False Positive
15	<i>BMP1</i>	c.1148G>A;p.(R383Q)	het		3888	?UV
15	<i>BMP1</i>	c.1293C>G;p.(Y431*)	het		2021	?UV

Table 20 Variants identified in participants 001-018 using targeted exome analysis

Each variant was manually assessed for likely pathogenicity and the results of this analysis are listed in the final conclusion column of the table.

SNV=single nucleotide variant (likely benign); UV=unclassified variant; N,#=likely pathogenic mutation; Het=heterozygous; Hom=homozygous

Four variants, *COL1A1* c.1350G>T;p.(E450D), *TAPT1* c.613-3Tdel, *TAPT1* c.837G>T;p.(M279I), and *CRTAP* c.471+1G>A were considered to be false positive calls due to their low read count, which indicates poor quality sequencing. Nine were classified as likely polymorphisms following evaluation as described in section iii.x due to their frequency in the general population and/or the lack of *in-silico* evidence to support a deleterious effect on protein structure or function.

A highly likely pathogenic mutation, c.1111G>T;p.(Gly371Cys) in the *COL1A1* gene, was identified in patient 12. This mutation affects a glycine residue in a Gly-X-Y triplet within the helical domain of the protein, a mutational mechanism well reported to be causative in OI. A different mutation affecting the same amino acid, c.1111G>A;p.(Gly371Ser), has previously been reported in a patient with type III OI (Marini, Forlino et al. 2007). No other variants of interested were identified in patient 12. It is likely that the c.1111G>T;p.(Gly371Cys) mutation was not detected by the original analysis of this patient due to the presence of a rare variant under a primer binding site, a known disadvantage of Sanger sequencing.

Four variants of particular interested remained: *BMP1* c.2134G>A;p.(Gly712Ser); *COL1A1* c.4070T>C;p.(Leu1357Pro); *BMP1* c.1293C>G;p.(Tyr431*) and *BMP1* c.1148G>A;p.(Arg383Gln). These variants were all confirmed by Sanger sequencing (Figure 26 and Figure 29).

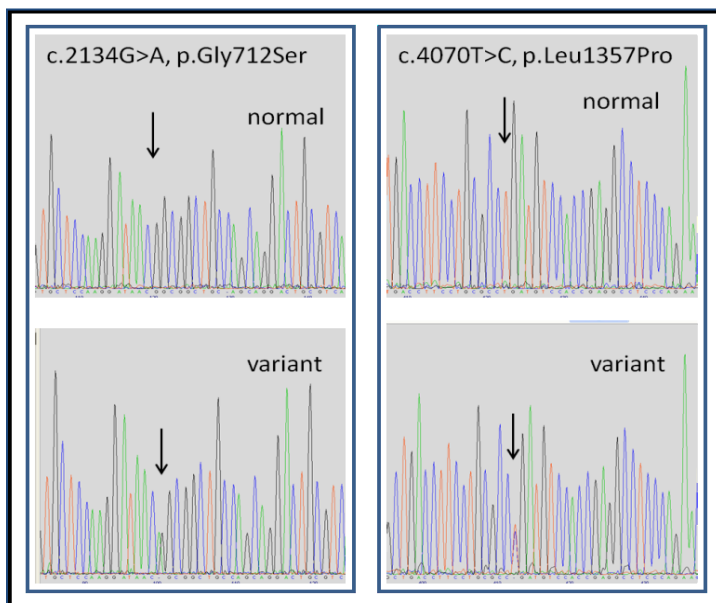


Figure 26 Sanger sequencing confirmation of *BMP1* mutations.

c.2134G>A;p.(Gly712Ser), participant 003 and COL1A1 c.4070T>C;p.(Leu1357Pro) participant 004.

iv.iii.i Investigation of *COL1A1* variant

The *COL1A1* p.(Leu1357Pro) variant is located in the C propeptide domain of the alpha 1 chain of type I collagen, a region that is essential for chain to chain recognition and subsequent assembly of the protein. Although this particular change has not been reported in the literature, three other variants affecting leucine residues in the C-propeptide, p.(Leu1464Pro), p.(Leu1437Gln) and p.(Leu1388Arg) have previously been published (Chessler, Wallis et al. 1993, Oliver, Thompson et al.

1996, Rugolotto, Monti et al. 2007) and are reported to be pathogenic. The p.(Leu1464Pro) variant affects the final amino acid before the termination codon in the protein and appeared *de-novo* in the affected individual. Analysis of collagen secretion from cultured fibroblasts from this individual showed a poorly secreted and overmodified protein supporting pathogenicity. Similarly, the p.(Leu1388Arg) appears to be *de-novo* in an individual with lethal OI and delayed secretion and extensive over modification of the type I collagen protein were again observed. No data is provided in the literature report in support of a pathogenic role for the p.(Leu1437Gln) change.

Initially four family members were available for analysis of the *COL1A1* c.4070T>C;p.(Leu1357Pro) variant in our family (participant 004). The variant was not detected in the proband's reportedly affected aunt or cousin but was present in the heterozygous state in her reportedly unaffected father. These findings suggested that this variant was not segregating with the features of OI in this family (Figure 27).

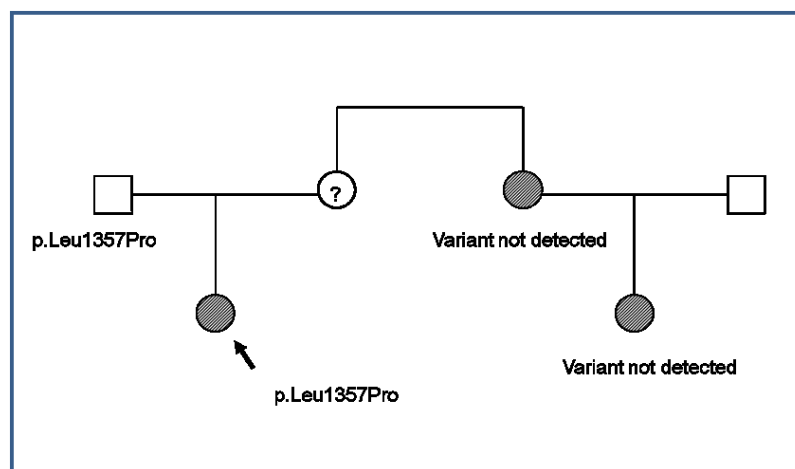


Figure 27 Segregation analysis for the *COL1A1* c.4070T>C;p.(Leu1357Pro) variant in family 004.

The filled shapes represent family members who were reported to be clinically affected. Females are represented by a circle, males by a square. The arrow indicates the family member initially recruited to the study.

However, the reportedly unaffected father was subsequently reviewed by a different consultant following the birth of a clinically affected second child from a different relationship. His clinical history was revealed to include numerous fractures during childhood and was highly suggestive of a diagnosis of OI. The second affected child was confirmed to be heterozygous for the c.4070T>C variant in *COL1A1*. Clinical review of the proband's aunt and cousin only identified features that would be consistent with normal bone mineralisation. This new information suggests that this variant is segregating with the OI symptoms in this family.

Each alpha chain of fibrillar collagens contains a unique sequence within the C-propeptide domain that allows correct recognition and assembly of the collagen molecule. The p.(Leu1357Pro) variant lies within the chain-chain recognition region of the alpha 1 chain but does not affect one of the chain specific amino acids (Figure 28). Comprehensive *in-silico* analysis of the c.4070T>C;p.(Leu1357Pro) variant, along with segregation with clinical features of OI and literature review, indicates likely pathogenicity and suggests that the physical properties of proline differ from leucine sufficiently to

interfere with chain recognition. This is likely to prevent or delay chain alignment and assembly resulting in over modification of the collagen molecule.

<p>Alpha 1 (I) GGQGSDPADVAI_QLTFLRLM_STE Alpha 2 (I) NVWGVTSKEMAT_QLAFMRLL_ANY Alpha 1 (III) BNPELPFDVLDV_QLAFLRLL_SSR</p>
--

Figure 28 Chain-chain recognition sequences for type I and III collagen molecules

The blue highlighted amino acids indicate the unique sequence for each chain that allow correct recognition and assembly of the collagen molecules. Leucine 1357 is highlighted in red.

The c.4070T>C;p.(Leu1357Pro) *COL1A1* variant is therefore likely to be an OI-causing mutation and adds to the number of C-propeptide domain mutations reported. Interestingly mutations in this region of the gene have been described in both mild OI (Pace, Kuslich et al. 2001) and a high bone mass variant (Takagi, Hori et al. 2011). None of the affected members of this family are reported to have features of high bone mass.

These results highlight the importance of phenotypic data in the interpretation of genetic findings.

iv.iii.ii Investigation of *BMP1* variants

The *BMP1* gene (OMIM 112264) is alternatively transcribed to produce two proteins, mammalian Tolloid (mTLD; NM_006129) and its shorter isoform bone morphogenic protein-1 (*BMP1*; NM_001199). The functions of the shorter *BMP1* protein include the proteolytic removal of the carboxyl-terminal propeptide from procollagen type I, II and III and the amino-terminal propeptide from types V and XI procollagen. It also influences dorsal-ventral patterning through the indirect activation of some TGF β superfamily proteins (Asharani, Keupp et al. 2012) and has a role in extracellular matrix assembly by proteolytic activation of lysyl oxidase (LOX). This is known to be essential for collagen cross-linking and processing of small leucine rich proteoglycans (SLRPs) (von Marschall and Fisher 2010).

mTLD has an alternative exon 16 to *BMP1* and a subsequent additional four exons encoding 256 amino acids. This isoform circulates at increased levels in the plasma during acute bone fracture and is known to have an important role in bone repair (Grgurevic, Macek et al. 2011).

Studies in *BMP1*/mTLD deficient patients with OI have demonstrated delayed cleavage of type I collagen C-propeptide (Martinez-Glez, Valencia et al. 2012, Valencia, Caparrós-Martin et al. 2014) and disorganization of type I/IV collagen fibrils as well as impaired processing of the SLRP procollagen (Syx, Guillemyn et al. 2015).

In vivo study of the *BMP1*/mTLD protein to elucidate the mechanisms by which mutations in the *BMP1* gene give rise to an OI phenotype has been hampered by the early lethality of *Bmp1*/TII1 knockout mice. However, postnatal ablation of the gene using floxed *Bmp1* and TII1 alleles have

produced models with reduced processing of procollagen and dentin matrix protein 1, increased bone turnover, decreased expression of sclerostin and induction of canonical Wnt signalling (Muir, Ren et al. 2014). These conditional knockout mice have weak brittle bones that are susceptible to spontaneous fractures. Studies in the *frilly fins* zebra fish *Bmp1a* mutant have shown delayed larval ossification progressing to relatively high bone mineral content in adults with malformation and evidence of fracture that has been shown to be independent of osteoblast maturation (Asharani, Keupp et al. 2012).

At the time of this study less than 20 individuals with OI have been identified with mutations in the *BMP1* gene. These patients have been described as having highly variable clinical presentations ranging from mild to severe progressively deforming OI. The majority, but not all, have presented with bone fragility associated with an increase in bone mineral density. It has been hypothesized that the described phenotypic variability between reported cases may result from different functional consequences of the *BMP1* mutations, although no clear genotype/phenotype correlation has yet emerged.

***BMP1* c.2134G>A;p.(Gly712Ser) : Participant 003**

The heterozygous *BMP1* c.2134G>A;p.(Gly712Ser) variant in participant 003 changes protein coding only in the longer *BMP1*/mTLD transcript, (mTLD; NM_006129). It affects a highly conserved nucleotide and amino acid and is situated in an EGF-like calcium-binding domain of the protein. In the shorter transcript (NM_001199) this variant, c.*626G>A, is downstream of the 3'UTR that finishes at c.*280.

In-silico protein conservation analysis of the p.(Gly712Ser) change using SIFT, PolyPhen2, Align GVD, SNPs3D, Panther and Mutation Taster all suggest that this variant may have a deleterious effect on protein structure and/or function, although the severity of the amino acid substitution is considered to be neutral.

However, the available allele frequency data for this change states that it is present in the heterozygous state at a frequency of 0.16% in European (Non-Finnish) alleles in the gnomAD database (sample size 63042); no homozygous individuals are recorded. This data suggests that, in the heterozygous state, this variant is unlikely to be responsible for the symptoms of OI in this patient. In addition, there are no literature reports that suggest that variants only affecting protein coding in the longer mTLD isoform are associated with increased bone fragility, although a role in bone repair cannot be completely excluded.

***BMP1* c.1293C>G;p.(Tyr431*) and c.1148G>A;p.(Arg383Gln) : Participant 015**

Two variants in the *BMP1* gene were identified in patient 015, c.1293C>G;p.(Tyr431*) and c.1148G>A;p.(Arg383Gln), that affect both isoforms of the protein. Segregation analysis showed that the c.1293C>G;p.(Tyr431*) is present in the mother and c.1148G>A;p.(Arg383Gln) in the father, confirming that these changes are on different alleles in participant 015.

The c.1148G>A;p.(Arg383Gln) missense variant is found within the highly conserved CUB1 domain of the protein and SIFT, PolyPhen, Align GVD and Mutation Taster analysis all support pathogenicity, although the severity of the amino acid substitution is favoured. There is very low frequency data

within the gnomAD population database, four alleles out of 121,292. No homozygous individuals are reported.

Investigation of the nonsense change identified in this individual, 1293C>G;p.(Tyr431*), indicates that variants leading to a 'null' allele within *BMP1* are associated with OI phenotypes. Two such changes, c.925del;p.(Asp309fs) and c.1839del;p.(Asn614fs) are recorded on the LOVD database (https://oi.gene.le.ac.uk/home.php?select_db=BMP1) associated with OI type III and type IV respectively. The 1293C>G;p.(Tyr431*) variant is not reported in the general population. Taken together, the available evidence for the two variants identified in this individual, suggest that they are highly likely pathogenic. So far, similar compound heterozygous changes that result in a 'null' allele and a mutation in a CUB domain have been reported to be associated with severe OI phenotypes.

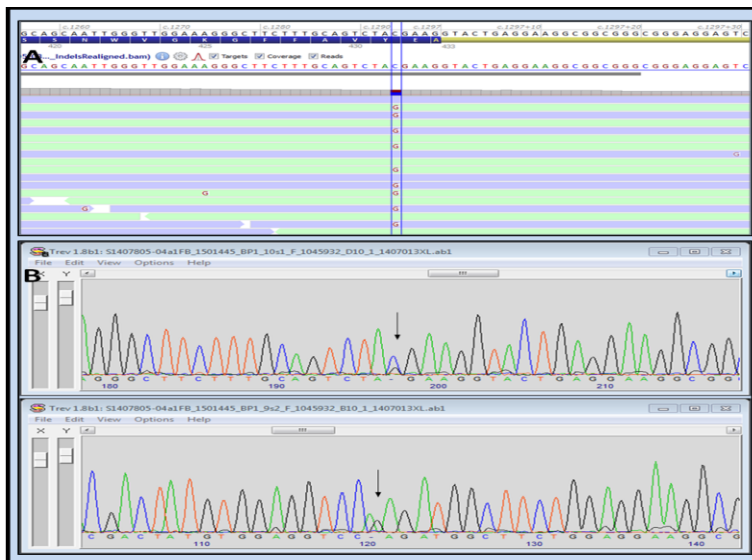


Figure 29 Targeted exome results for patient 015

A: NGS sequence data highlighting the c.1293C>G;p.(Tyr431) mutation in BMP1. B. Sanger sequence confirmation of the c.1293C>G;p.(Tyr431*) and c.1148G>A;p.(Arg383Gln) BMP1 mutations*

Patient 015 is a 7 year old female, the only child of healthy non-consanguineous parents of North European origin. There is no family history suggestive of a connective tissue disorder. She was born following IVF treatment at 39 weeks gestation with a birth weight of 2.976 kg. She was considered to be well immediately after birth and her early developmental assessment was normal. At age 17 months, she was noted to be non-weight bearing and was referred for physiotherapy. She was diagnosed as being hypermobile and provided with orthopaedic footwear. At 20 months of age, she was found to have congenital bilateral hip dislocation. Corrective surgery on both hips was undertaken, initially unsuccessful on the right side and further pinning was required.

The patient began to walk unaided at 2 years 10 months and her first fracture, of the right fibula, occurred a month later. She has sustained further fractures including a spiral fracture of the tibia and

three metatarsal fractures in the left foot (Figure 30). Following her fracture history, she was commenced on treatment with Pamidronate at 3.8 years of age. She remained on treatment for 12 months before raised BMD was identified by DXA resulting in cessation of treatment (Figure 30). A provisional diagnosis of osteopetrosis was suggested.

At her most recent clinical review aged 7 years, her weight was 24.35kg (>50th centile) and height 122cm (50-70th centiles). On examination, she has white sclerae and normal teeth, hearing and spine. She has a bossed forehead and has previously been noted to have a mild left sided ptosis. Scarring was minimal but haemosiderin deposits were noted on her lower limbs.



Figure 30 Patient 015: X-Ray of lower leg and sequential BMD measurements.

See inserted captions for more detailed description. BMD – Bone mineral density, LS – lumbar spine, BMAD - bone mineral apparent density

***BMP1* c.2188dupC;p.(Gln730fs): participants I and J**

Following identification of the *BMP1* likely pathogenic mutations in patient 015, the *BMP1* gene was added to the routine diagnostic NGS panel at SDGS and a further two individuals with OI associated with *BMP1* mutations (patients I and J) were identified. The clinical features of these patients were assessed and available BMD data gathered to produce the manuscript (Appendix page 226):

Phenotypic variability in patient with osteogenesis imperfecta caused by *BMP1* mutations. Pollitt RC *et al* Am J Med Genet A. 2016 ;170(12):3150-3156.

In addition to the phenotypic variability described in the manuscript, an important finding that is highlighted is the potential risk of causing delayed healing, increased stiffness, atypical fractures or even iatrogenic osteopetrosis by treatment with anti-resorptive therapy in *BMP1*-related OI. In particular, careful monitoring of response to bisphosphonate therapy is indicated in these patients.

***BMP1* c.355C>T;p.(Arg119Trp) : participant K**

Following publication of the manuscript, a further patient (K), again with a high bone mass phenotype, and a novel homozygous c.355C>T;p.(Arg119Trp) mutation in the pro-domain of the *BMP1* gene was identified.

It is not clear why some mutations are associated with increased BMD whereas others are reduced. Published functional studies have largely focused on the C-propeptide cleavage activity of *BMP1* mutations but *BMP1*/mTLD is also involved in processing of additional extra cellular matrix components. In particular *BMP1* deficiency has been shown to influence the processing of the SLRP prodecorin by impaired removal of the prodomain (Syx, Guillemyn *et al.* 2015).

Decorin is known to influence both collagen assembly and matrix mineralization (Mochida, Parisuthiman *et al.* 2009). Decorin also participates in intracellular and extracellular signalling including binding and inhibition of TGF- β . Reduced decorin levels cause increased levels of 'free' TGF- β that bind to the TGF- β RI and TGF- β RII receptors, inducing Smad dependent and Smad independent signalling (Markmann, Hausser *et al.* 2000, Lim, Grafe *et al.* 2017). As excessive TGF- β signalling is thought to be a common mechanism in OI pathogenesis (Grafe, Yang *et al.* 2014), it seems reasonable to propose that reduced decorin prodomain removal is contributing to the increased mineralization in these patients, by influencing downstream TGF- β signalling.

One possible explanation for the variable phenotypic presentation in this small cohort of *BMP1* OI patients is that the degree of increased mineralization is a direct result of the 'severity' of the mutation, the level of residual C-propeptide cleavage activity that remains and the resulting disruption to collagen assembly and matrix mineralisation.

To explore this idea further, collaboration with Dr. Cecilia Giunta, (University of Zurich) was established to examine collagen processing in dermal fibroblasts from patient 015, J, and K to see if variability in C-propeptide cleavage activity could be demonstrated.

Collagen Biochemical Analysis in patients 015, J and K.

To look at collagen biochemical processing, dermal fibroblasts are cultured in medium containing ^{14}C -proline which labels procollagen and collagen proteins. These proteins are subsequently isolated and analysed using 5% sodium-dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Gels are dried, fixed and exposed to an X-ray film before visualisation.

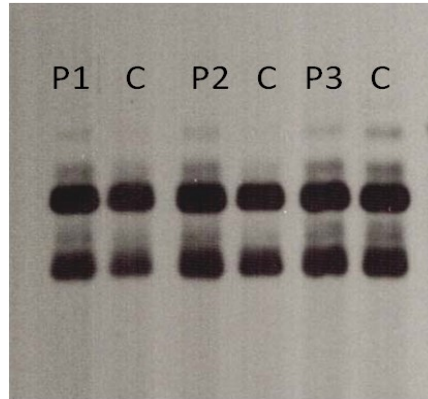


Figure 31 SDS-PAGE of the pepsinized procollagens in the medium.

Normal controls (C) and patients 015, J and K (P1, P2 and P3 respectively). No evidence of abnormal procollagens in the patient's samples was detected.

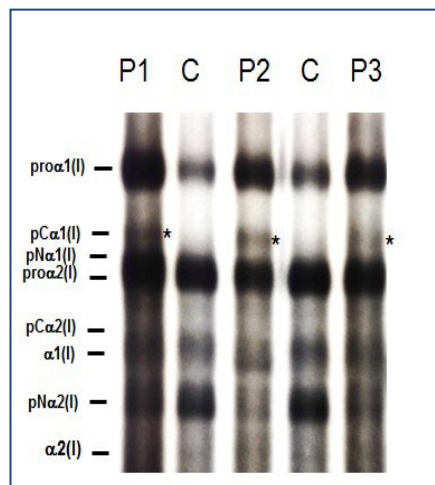


Figure 32 SDS-PAGE of the untreated procollagen in the presence of 0.05M dithiodthreitol, (DTT)

Patients 015, J and K (P1, P2 and P3 respectively).

** Accumulation of pCα1(I) procollagen species (those retaining the C-terminal propeptide) in patient sample when compared to normal controls.*

The isolated procollagen can be treated with pepsin to cleave the terminal non-helical regions allowing more specific examination of the triple helical portion of the protein. The SDS-PAGE of the pepsinized procollagens present in the medium is normal for all three patients (Figure 31). This

suggests that a defect in the type I collagen triple helical protein structure or processing is not responsible for the symptoms in these individuals SDS-PAGE of the untreated procollagen in the presence of dithiodthreitol, (DTT, a reducing agent) clearly shows the accumulation of pC α 1(I) procollagen species in all three patients (those retaining the C-terminal propeptide) and a decreased production of pN α 1(I) and pN α 2(I), which is more evident on the pN α 2(I) band (Figure 32). These results support that the *BMP1* mutations in these patients result in reduced cleavage activity. However, there is no indication that any specific mutation is more detrimental to collagen processing than the others. The cause of the variable phenotypic severity in these patients remains unclear.

Electron Microscopy of reticular dermis in patients 015, J and K

Electron microscopy (EM) of skin in OI patients with type I collagen mutations has shown increased variation in collagen fibril diameter, variation from the expected mean diameter and the presence of collagen 'flowers' (Balasubramanian, Sobey et al. 2016). A limited number of EM studies in *BMP1* positive patients have also shown variable collagen fibril diameter and irregular fibril outlines although collagen flowers have not been reported (Syx, Guillemyn et al. 2015). The morphology of dermal fibroblasts in *BMP1* mutation positive patients has not been described.

To further explore the functional effect of the *BMP1* mutations identified in this study, examination of dermal collagen fibrils and fibroblasts from patients 015, J and K was undertaken. These were compared against a normal control and a known pathogenic *COL1A2* mutation.

Normal Control

Collagen bundle packing was regular and collagen fibrils were within the expected range for life stage (85nm). Elastic fibres were peripherally irregular in outline but did not appear to be increased or decreased in number or diameter. Collagen fibrils were regular with only an occasional fibril with an irregular outline. No collagen 'flowers' or variability in fibril diameter were evident (Figure 33).

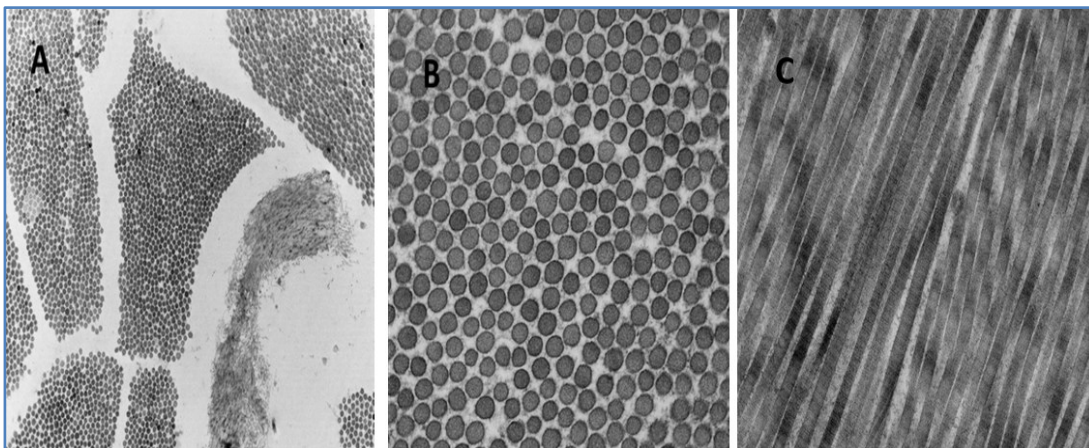


Figure 33 Normal control collagen and elastic fibres

A. Elastic fibres are peripherally irregular in outline but do not appear, to be increased or decreased in number or diameter. Collagen bundle packing is normal (9200X). B. Collagen fibrils are regular with only occasional fibrils with an irregular profile. No collagen 'flowers' or variability of fibril diameter was observed (20000X). C. Longitudinal section of collagen fibrils showing the classical D-period banding pattern.(9200X)

Examination of dermal fibroblasts showed normal cell morphology without any expansion of the ER or any protein 'aggregates' (Figure 34).

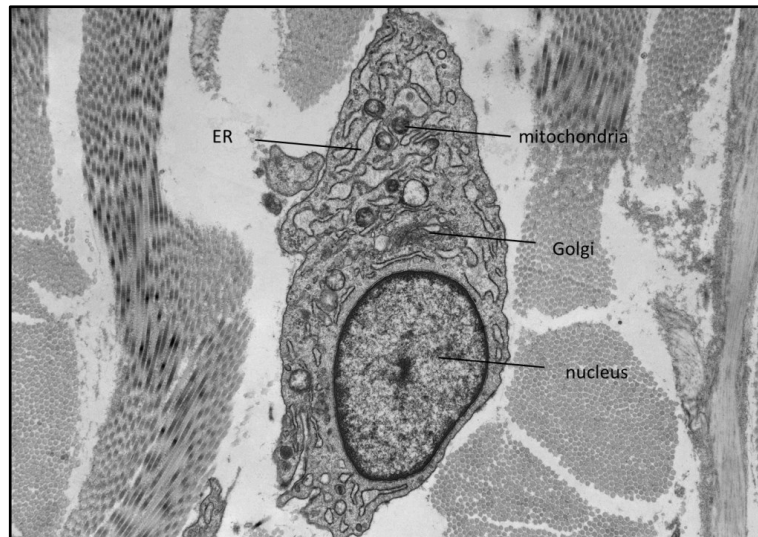


Figure 34 Normal control dermal fibroblast

The nucleus, mitochondria, golgi and ER are as labelled. The ER is not expanded (5500X).

Positive control

Fibroblasts from a patient with OI (type III) and a c.3269G>A;p.(Gly1090Asp) pathogenic mutation in *COL1A2* were examined as a positive control. These fibroblasts had grossly expanded ER that was protein-filled, supporting aberrant processing of the collagen molecule and subsequent ER stress (Figure 35). The presence of collagen flowers and variability in fibril diameter are known features of OI (Balasubramanian, Sobey et al. 2016) and were confirmed in this positive control sample.

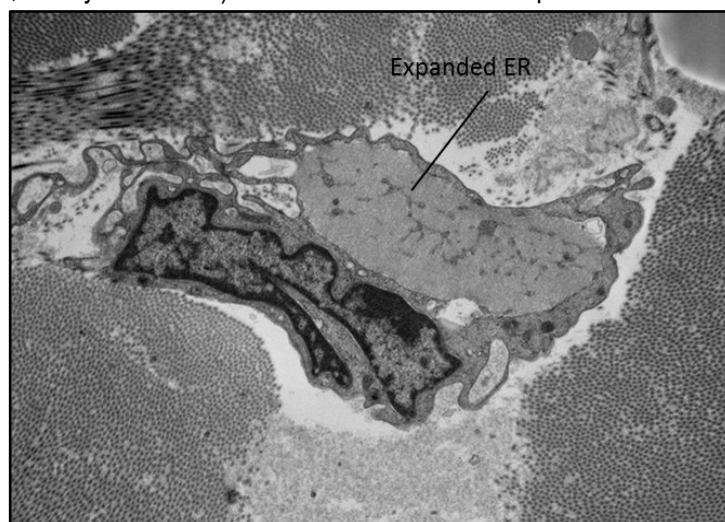


Figure 35 Ultra structure of dermal fibroblast from a patient with OI type III with a c.3269G>A;p.(Gly1090Asp) *COL1A2* pathogenic mutation.

The fibroblasts showed grossly expanded ER that was protein filled suggesting aberrant processing and ER stress.(2600X)

Patient 015

Some fibrils showed irregular contours similar to those previously described (Syx, Guillemyn et al. 2015). Variation in collagen fibril diameter is within normal limits, with a standard deviation of 8nm. Occasional small collagen flowers were also identified. Longitudinal sections showed the presence of twisted and disorganised fibrils (Figure 36). Collagen bundle packing appeared normal.

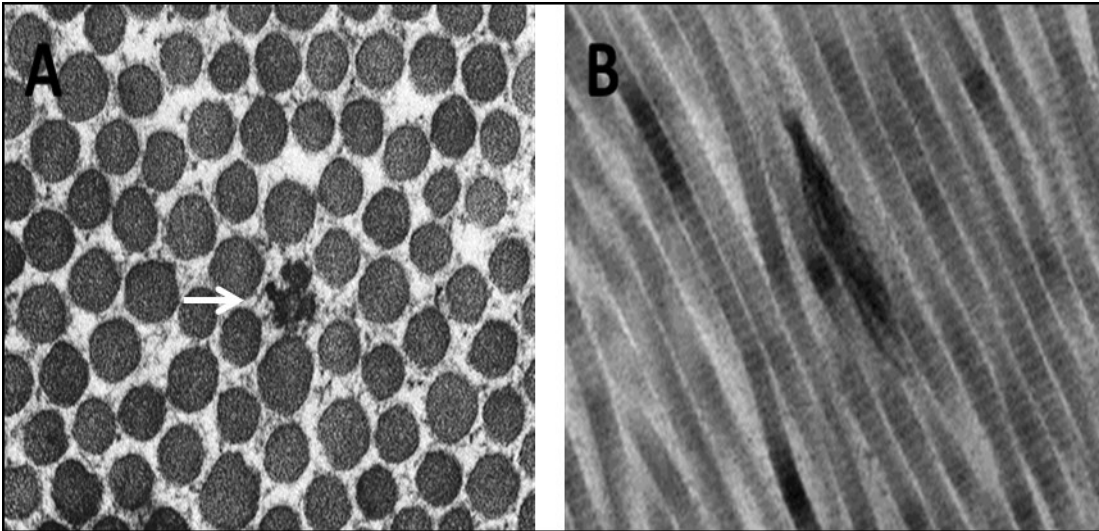


Figure 36 Ultra structure of collagen fibrils in reticular dermis from patient 015

A. Some fibrils show irregular contours similar to those previously described (Syx, Guillemyn et al. 2015) and there is mild fibril diameter variability. An occasional small cauliflower-like collagen fibrils was also identified (arrowed, 20000X) B. Longitudinal sections show the presence of twisted and disorganised fibrils (12000X)

Examination of fibroblasts identified only mildly expanded ER. Elastic fibres were unremarkable (Figure 37).

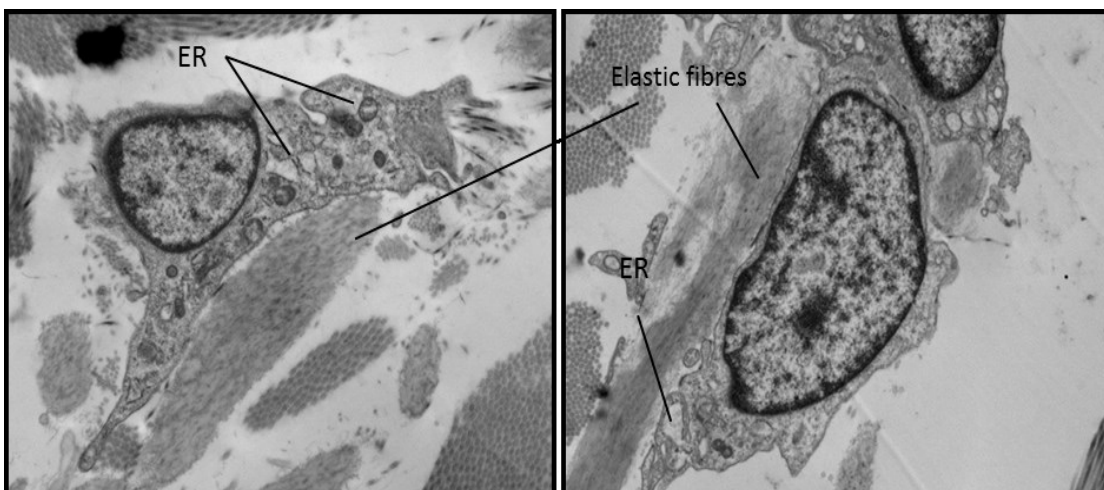


Figure 37 Ultrastructure of dermal fibroblasts from patient 015.

Gross morphology is normal with only mildly expanded ER. Elastic fibres unremarkable. (3300X)

Patient J

Collagen bundle packing was normal. A single fibroblast was identified with moderately expanded and protein-filled rough endoplasmic reticulum (Figure 38). Fibroblasts were otherwise unremarkable.

Collagen fibril diameter variation from the mean ranged from 5.8 to 9.4nm. Mean collagen fibril diameter, at 96nm, was in excess of expected for age (85nm). No collagen flowers were observed. Elastic fibre ultra structure appeared normal.

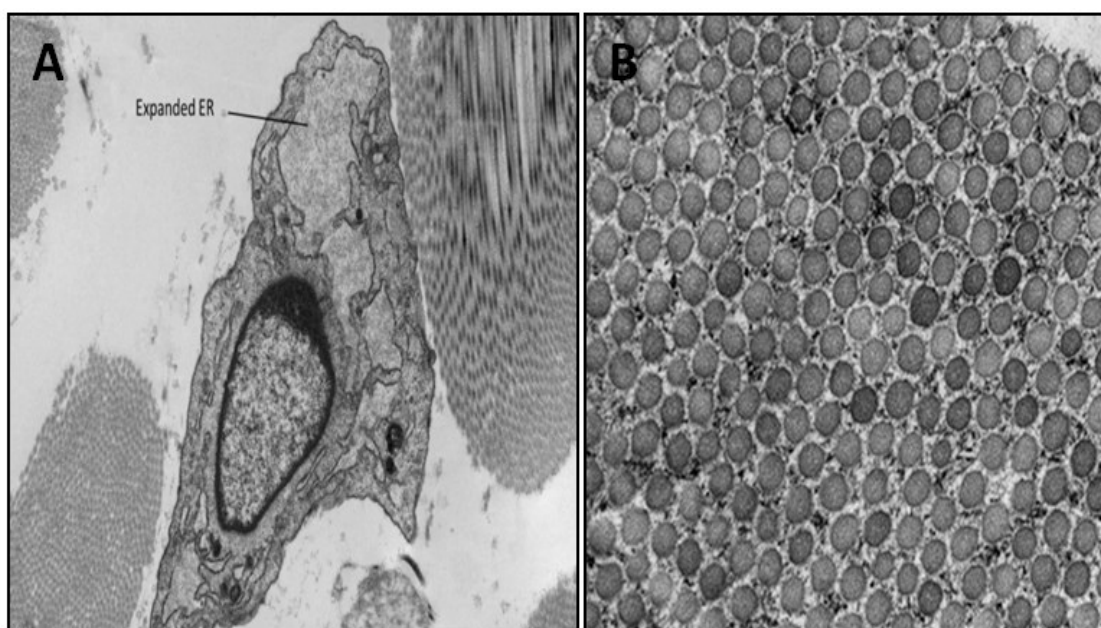


Figure 38 Ultra structure of collagen and fibroblast from reticular dermis of Patient J

A: A single fibroblast was identified with mildly expanded ER. All other fibroblasts were unremarkable (3300X) B: Mean collagen fibril diameter (96nm) was in excess of that expected for age (85nm). No collagen flowers were observed (20000X).

Patient K

Collagen bundle packing was unremarkable. Mean collagen fibril diameter at 97nm was slightly in excess of expected for age (87nm). Variation in collagen fibril diameter was not excessive at a root mean squared deviation of 5.6 to 12.1nm. A few collagen fibrils with an irregular outline and an occasional collagen flower were observed (Figure 39).

The rough endoplasmic reticulum was very slightly expanded and protein filled in occasional fibroblasts (Figure 39). Elastic fibres appeared unremarkable.

Cleavage of the type I collagen C-propeptide is extracellular and therefore it would not be anticipated that fibroblasts from *BMP1* mutation positive patients, where C-propeptide cleavage is disrupted, would show significant signs of ER stress. Rather, it might be expected that collagen fibrils would show features suggestive of disrupted fibril assembly.

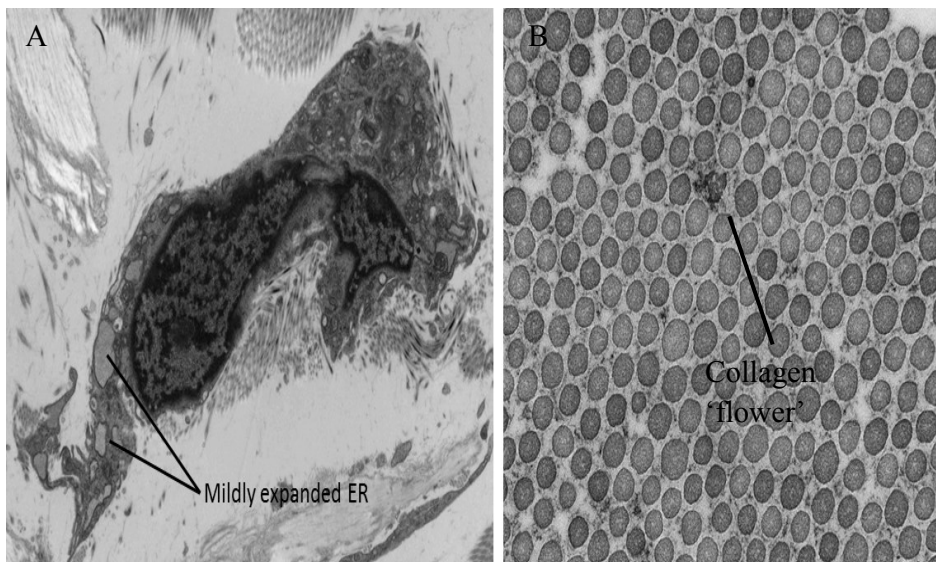


Figure 39 Ultra structure of collagen and fibroblast from reticular dermis of Patient K

A. Occasional fibroblast with mildly expanded and protein filled ER was observed (3300X). B. A few collagen fibrils with an irregular outline and an occasional collagen flower were observed (20000X)

The observations in our patient cohort support this idea and are concordant with previously described collagen fibril architecture (Syx, Guillemyn et al. 2015), but with the addition feature of an occasional collagen flower in some patients (015 and K).

Fibroblast morphology was generally unremarkable in this cohort. The mildly expanded ER identified in an occasional fibroblast in all three patients was in contrast to the gross expansion seen in the *COL1A2* mutation positive control. The significance of the mild expansion is unclear.

Summary

This group of patients demonstrated the importance of inclusion of *BMP1* gene analysis in the diagnostic investigation of patients with a clinical diagnosis of OI. This gene analysis should not be confined to patients with a noticeable high bone mass phenotype as bone mass in *BMP1*-related OI has been demonstrated to be highly variable (Pollitt, Saraff et al. 2016).

In addition, careful consideration should be given to anti-resorptive therapy in this ultra rare form of OI and close monitoring of response to treatment is recommended.

Functional exploration of the patients in this study by biochemical collagen analysis, and characterisation of collagen fibril and fibroblast ultra structural features has given no clues as to the mechanism responsible for the variable phenotypic severity in this group. The exact mechanism by which *BMP1* mutations cause bone fragility remains to be elucidated.

v **RESULTS: MOUSE MODELS (*TRAM2* Gene Analysis)**

The advances in genetic technology have considerably improved our ability to sequence and annotate the human genome. However, despite this, the function of many genes and their role in disease is still poorly understood. The mouse is considered the mammalian model of choice to study the function of genes and a large scale comprehensive mouse phenotyping consortium has been established (International Mouse Phenotyping Consortium, IMPC, www.mousephenotype.org) to try and address the questions about gene function (Ayadi, Birling et al. 2012). Some of these mouse models have already been valuable in assessing candidate genes associated with OI, for example in *TAPT1*, *SP7* and *SPARC*.

The IMPC consortium is focused on generating mutations in protein-coding genes of the mouse genome. Mutant mouse strains undergo comprehensive high-throughput phenotyping and the data generated from this phenotyping pipeline is made available through the web portal www.knockoutmouse.org. This functional gene annotation resource allows us to target investigation of genes of interest in rare diseases such as OI.

The phenotypic assessment performed by the Mouse Genetics Project (MGP) includes measurement of DEXA scores to provide an indication of bone mineral density (BMD) in their knock-out mouse models. BMD is currently used as a predictor of fracture risk, although correlation is poor. Approximately 160 knock-out mice are assessed per year, with an approximate 4% hit rate for DEXA. This data is useful in targeting further genes for analysis in our cohort, although mouse models can only point towards genes where knock-out, rather than gain of function or dominant-negative, mutations are causative.

Nine genes affecting bone strength have been published by the consortium so far (Bassett, Gogakos et al. 2012). Three of these, *Prpsap2*, *Slc38a10* and *Sparc*, are reported to result in bones that are described as weak and brittle. The *Prpsap2* and *Slc38a10* genes are thought to be important in bone development, whereas *Sparc* is a glycoprotein found in the extracellular matrix and is thought to have a role in bone mineralization. *SPARC* mutations have subsequently been identified in autosomal recessive OI, as described in section ii.xii.vi.iii, confirming the potential of mouse models to identify genes associated with bone fragility. The human homologue gene, *SLC38A10*, will be discussed further in section viii.vi.

Collaboration with the Sanger Mouse Genetics Project (Sanger MGP), a partner in the IMPC, highlighted a novel gene associated with skeletal abnormalities, the *Tram2* gene (Figure 42). Primary data in a *Tram2* *-/-* knockout mouse indicated reduced viability of homozygous pups (38/274 pups = 14%, expected 25%), with normal weight at birth but with post-weaning growth retardation (Figure 40).

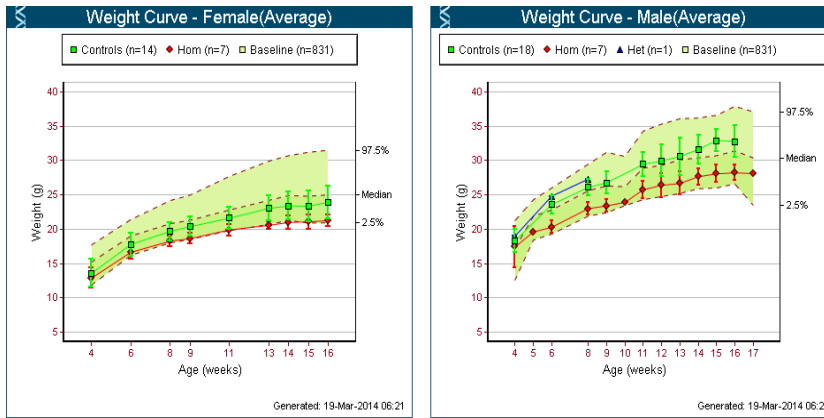


Figure 40 Male and Female weight curves for the *Tram2*^{-/-} mouse model.

Both sexes show reduced bone weight when compared with wildtype age and sex matched controls

Reduced bone mineralisation parameters were also recorded, including reduced bone mineral density in both males and females (Figure 41).

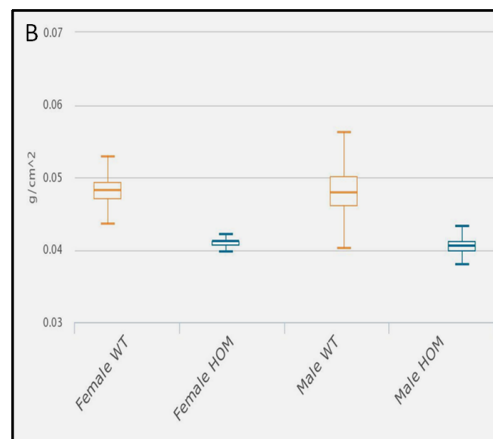


Figure 41 Measure bone mineral density in *Tram2*^{-/-} knockout mouse using the DEXA analyser (g/cm²).

Data provided by Chris Lelliot, Sanger MGP

Tram2^{-/-} knockout mouse also have substantial hearing impairment demonstrated by abnormal response to auditory stimulus (click box) measured by signal detection threshold in brainstem response. These mice are not completely deaf but require “louder” noise to produce a detectible electrical trace.

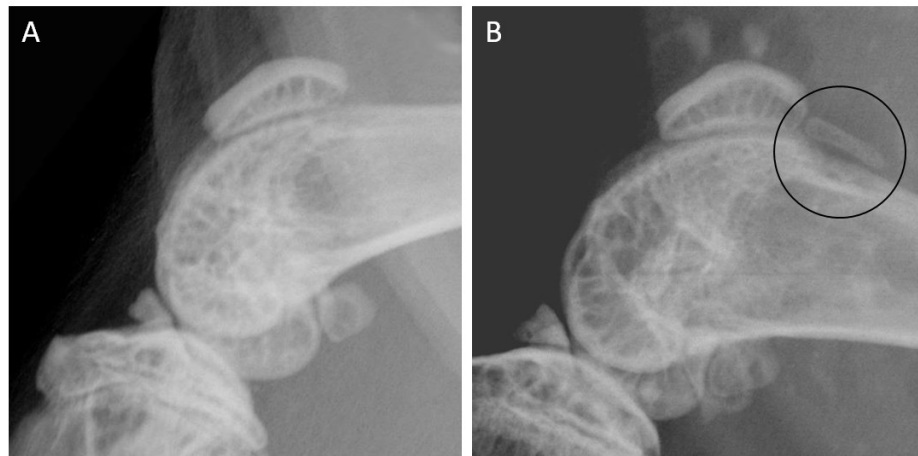


Figure 42 X-ray of adult mouse knee joint at 16wks.

A Normal control B Tram2 -/- knockout mouse with abnormal patella morphology highlighted. Images kindly provided by Chris Lelliot Sanger MGP

The translocation-associated membrane protein 2 gene (*TRAM2*) is cytogenetically located at 6p12.2 and encodes an ER multi-pass membrane protein. The function of *TRAM2* in humans is not completely clear at present. It has been shown to interact with *RUNX2*, a major transcription factor in osteoblasts, in a BMP-dependent manner (Pregizer, Barski et al. 2007) and is known to be required for type I collagen synthesis, most probably by interaction with *SERCA2b*, the calcium pump of the ER. This interaction may increase Ca^{2+} concentration at sites of collagen synthesis and facilitate the functioning of chaperones involved in collagen folding (Stefanovic, Stefanovic et al. 2004).

TRAM2 is also thought to mediate translocation of M4SF20 (transmembrane 4 L6 family 20), a polytopic membrane protein that inhibits proteolytic processing of *CREB3L1/OASIS*, in the absence of ceramide (Chen, Denard et al. 2016).

Taken together, the phenotype of the *Tram2* $-/-$ knockout mouse and the proposed protein function, suggested that *TRAM2* is a candidate gene for bone fragility. To explore this further the *TRAM2* gene was analysed by either Sanger sequencing or by interrogation of whole exome data in all patients in our cohort.

Sanger sequencing covered all coding sequence and intron/exon boundaries up to ± 25 bp. Where deeply intronic variants were identified in exome data these were discounted if they were described with an allele frequency of $>1\%$ in population datasets. Analysis of the remaining variants identified two recurring changes c.627-5C>T and c.470+5_470+6 dup (Table 21 and Figure 43).

The c.627-5C>T variant is listed on NCBI dbSNP with an average heterozygosity of 0.17. In the gnomAD cohort (121,004 chromosomes) 320 individuals are reported as being homozygous. This variant can therefore be classed as a benign polymorphism and is highly unlikely to result in symptoms of OI.

Participant ID	Variant	Zygoty	dbSNP ID
001	c.627-5C>T	het	rs9688915
002	c.627-5C>T	het	rs9688915
004	c.470+5_470+6 dup	het	rs57120044
007	c.627-5C>T	het	rs9688915
008	c.470+5_470+6 dup	het	rs57120044
014	c.470+5_470+6 dup	het	rs57120044
A	c.470+5_470+6 dup	homo	rs57120044

Table 21 Variants identified in the *TRAM2* gene in participants 001-018 and A-H

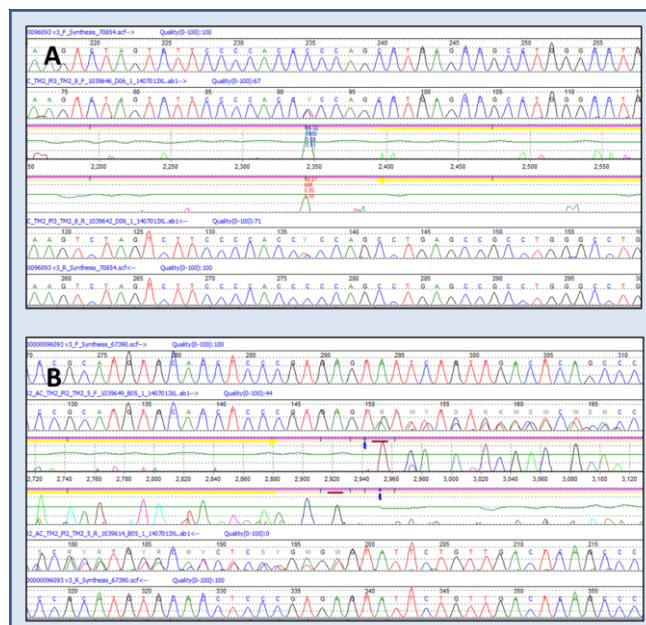


Figure 43 Sanger Sequence of two common variants in the *TRAM2* gene identified in our cohort.

A c.627-5C>T and B c.470+5_470+6 dup

Frequency data for the c.470+5_470+6 dup is comprehensive. Average heterozygosity is stated as 0.264 in NCBI dbSNP, although there is no data for the number of individuals that this data is derived from. However, gnomAD, ESP and GoNL datasets all show that this variant is common in the general population, with allele frequencies ranging from 6.24% to 49.41% depending on ethnicity. This variant is therefore also highly unlikely to result in symptoms of OI.

The results from this analysis suggest that sequencing single genes of interest is a poor strategy for identifying novel genes in OI cohorts.

vi **RESULTS: WHOLE EXOME SEQUENCING (Personalis)**

Three patients (participant ID: A, B and C) were commercially sequenced by Personalis® using their ACE Exome™ Assay. Personalis® returned all raw sequence data as well as a report highlighting 'significant' findings. This report used the HPO terms submitted for each patient to target genes potentially responsible for the patient's clinical symptoms.

Parental samples were also submitted for sequencing for patient B.

vi.i **Patient A:**

Two variants of uncertain significance were reported by Personalis: c.4136C>T;p.(Thr1379Met) in the *MYH6* gene and c.3660G>A;p.(Met1220Ile) in the *ARID1A* gene. These findings were assessed at the gene and variant level to clarify if they were likely to be responsible for the phenotypic features in this individual.

vi.i.i **MYH6 c.4136C>T;p.(Thr1379Met)**

The *MYH6* gene has been described in association with familial hypertrophic cardiomyopathy. This particular variant is present in multiple healthy genome datasets including NHLI ESP and gnomAD, It is recorded in 72 heterozygous individuals (in 66714 alleles) in the gnomAD European (non-Finnish) population and at a frequency of 0.2% in the Dutch (GoNL) population. This variant is therefore likely benign, although a contributing role in the cardiac abnormality in this patient (ventricular septal defect) cannot be completely excluded.

vi.i.ii **ARID1A c.3660G>A;p.(Met1220Ile)**

Nonsense, frameshift and microduplication mutations in the *ARID1A* gene are reported in association with an autosomal dominant mental retardation syndrome, Coffin-Siris type 2 (Kosho, Okamoto et al. 2014, Bidart, El Atifi et al. 2017). The c.3660G>A missense variant is novel in that it is not recorded in any population genome datasets. *In-silico* conservation analysis is inconclusive with multiple packages, including SIFT and Polyphen, producing conflicting predictions as to likely effect. In addition, although the amino acid is conserved among vertebrate species, the amino acid substitution is predicted to be neutral. Without further supporting evidence the clinical significance of this finding is unclear. However, although this variant has the potential to be contributing to this patient's symptoms of intellectual disability, it seems unlikely to be responsible for his bone fragility.

Further exploration of the Personalis sequence data was undertaken. In the absence of parental samples we searched for all LOF variants in this patient's dataset. A total of 123 high quality variants were identified and each one was manually assessed for likely contribution to patient's phenotype. Strategy 2 (exomizer) and strategy 3 (target gene list) variants were also explored. No OI or bone fragility candidate variants/genes were identified in this individual.

Segregation analysis in the parents of this individual would help to further explore the sequence data for this individual using inheritance patterns (strategy 1). This may also clarify the significance of the

MYH6 variant identified in this individual. The patient and his parents have been enrolled in the DDD project to facilitate this analysis.

vi.ii Patient B

A heterozygous variant, c.6427G>A;p.(Gly2143Arg) in the *COL6A3* gene was reported by Personalis.

vi.ii.i COL6A3 c.6427G>A;p.(Gly2143Arg)

This gene is associated with autosomal dominant and autosomal recessive collagen VI related myopathy including Bethlem and Ullrich myopathies and is reported at very low frequency in the general population (2/121290 alleles). Glycine 2143 lies within the triple helical domain of the protein and similar heterozygous variants in this region have been reported in association with Bethlem myopathy (Lampe and Bushby 2005). However, as collagen VI is a non-fibrillar collagen this variant cannot be presumed to be pathogenic, although *in-silico* analysis would support this. The variant is inherited from this individual's unaffected mother, and this could indicate non-penetrance, reduced expressivity or recessive inheritance. However, neither the child nor his mother has clinical features suggestive of a type VI related myopathy, indicating that this is most likely a rare benign sequence variant.

vi.ii.ii SRCAP c.9029C>A;p.(Pro3010His)

Exploration of this patient's sequencing data using strategies 1, 2 and 3 identified a *de-novo* c.9029C>A;p.(Pro3010His) variant in the *SRCAP* gene (NM_006662.2) that was not reported by Personalis. These findings were confirmed by Sanger sequencing (Figure 44).

This *SRCAP* gene maps to 16p11.2 and encodes the multifunctional SNF2-related chromatin-remodelling factor. The SRCAP protein is the catalytic component of a complex that mediates the ATP-dependent exchange of histone H2AZ/H2B dimers for nucleosomal H2A/H2B leading to transcriptional regulation of genes by chromatin remodelling.

Independently to its role in chromatin remodelling, SRCAP is a coactivator for the CREB-binding protein (CREBBP, also known as CBP). CREBBP regulates gene expression via mediating the interaction between transcription factors and transcription machinery (Monroy, Ruhl et al. 2001). In particular CREBBP acts with EP300 via the transcription factor cAMP-responsive element-binding protein (CREB) to regulate the transcription of a number of bone-related genes (Gordon, Stein et al. 2015). EP300 itself is associated with Rubinstein-Taybi syndrome 2 where patients have characteristic facial features, short stature, and moderate to severe learning disability (Gordon, Stein et al. 2015).

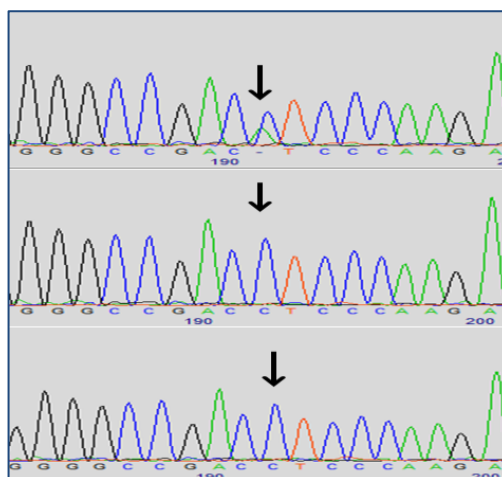


Figure 44 Sanger sequence confirmation of *de-novo* c.9029C>A;p.(Pro3010His) SRCAP variant in patient B.

Top = patient B, middle = maternal, lower = paternal

Truncating mutations in *SRCAP* are associated with Floating Harbor Syndrome (FHS), a condition characterized by short stature, delayed bone age and delayed speech (Messina, Atterrato et al. 2016). A characteristic facial appearance that includes, deep set eyes, short philtrum, thin lips, wide columella and long eyelashes is described.

The c.9029C>A variant in our patient lies in exon 34 of the *SRCAP* gene, a region that is known to be a 'hotspot' for FHS mutations (Hood, Lines et al. 2012). These heterozygous truncating mutations are predicted to abolish the three C-terminal AT-hook DNA-binding motifs while leaving the CBP-binding and ATPase domains intact (Figure 45). It is proposed that these mutations act in a dominant negative manner either by competing with wildtype protein in binding to its partners in the *SRCAP* chromatin remodelling complex and CREBBP or alternatively by interacting directly with wildtype *SRCAP* protein to form inactive heterodimers that are unable to bind DNA targets or chromatin partners (Messina, Atterrato et al. 2016).

A single case has been described of a *de-novo* 186kb microdeletion on chromosome 16p11.2 that encompasses the *SRCAP* gene. The patient is reported to have global developmental delay, speech impairment and behavioural problems (ADHD). Deep set eyes, bulbous nose with broad nasal bridge, and thin upper lip were also identified which could be considered as features that overlap with FHS, however it was felt that a diagnosis based on facial gestalt was not possible (Gerundino, Marseglia et al. 2014). The authors suggest that this patient might represent the first reported case of haploinsufficiency of *SRCAP* in association with a clinical phenotype, although acknowledge that further evidence is required. It is worth noting that this individual is not reported to have a skeletal phenotype.

To date, no missense variants in *SRCAP* have been reported in association with a clinical phenotype.

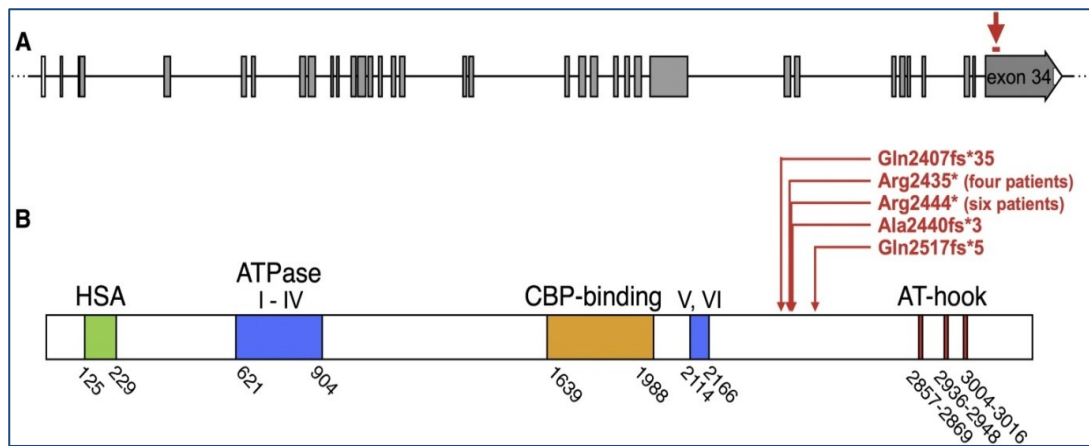


Figure 45 Locations of Floating Harbour Syndrome-Causing Mutations within SRCAP.

(A) Intron-exon structure of SRCAP. The mutation cluster in exon 34 is indicated by a red arrow.

(B) SRCAP domain structure. The amino acid positions of recognized domains are given below. The positions of heterozygous FHS mutations are shown in red.

Key: HSA = Helicase-SANT-associated domain. Modified from Hood et al. (Hood, Lines et al. 2012)

The c.9029C>A missense variant lies within the AT hook 3 domain and affects a moderately conserved amino acid. The physiochemical difference between proline and histidine are moderate and the amino acid substitution is predicted to be disfavoured (Unclassified Variant evaluation form, Appendix page 190). Review of variants listed in the dbSNP database did not identify any missense changes in the AT hook domains of the SRCAP gene with any frequency data in the normal population. In addition, search of the decipher database did not identify any further missense changes in this region of the gene. GnomAD constraint metrics for this gene are pLI=1.00 (pLI>=0.9 indicates an extreme loss of function intolerant gene) and the Z score for missense mutations is 2.23 (a positive Z score indicates increased constraint/intolerance).

Methylation Status

One mechanism for the potential pathogenicity of this variant could be via changes in chromatin remodelling and subsequent changes in DNA methylation status. A recent study looking at the methylation profile associated with FHS suggests that truncating mutations of SRCAP result in locus specific DNA methylation changes which in turn may regulate the expression of specific genes (Hood, Schenkel et al. 2016). It is possible that the c.9029C>A variant also has an impact on methylation status of genes, particularly those associated with bone metabolism.

To explore this hypothesis Prof Sadikovic (Western University, London, Canada) kindly undertook DNA methylation status analysis for patient B. They used an EPIC array (Illumina) that detects cytosine methylation at CpG islands and allows interrogation of methylation patterns at the genome-wide level, covering more than 850,000 methylation sites.

The results for patient B were compared against a large range of control samples from individuals of different ages (to counteract any bias in age related methylation status) and a cohort of patients with FHS. Multiple approaches from hierarchical clustering to support vector machine were employed. Patient B clustered with the normal control cohort (Figure 46 and Figure 47).

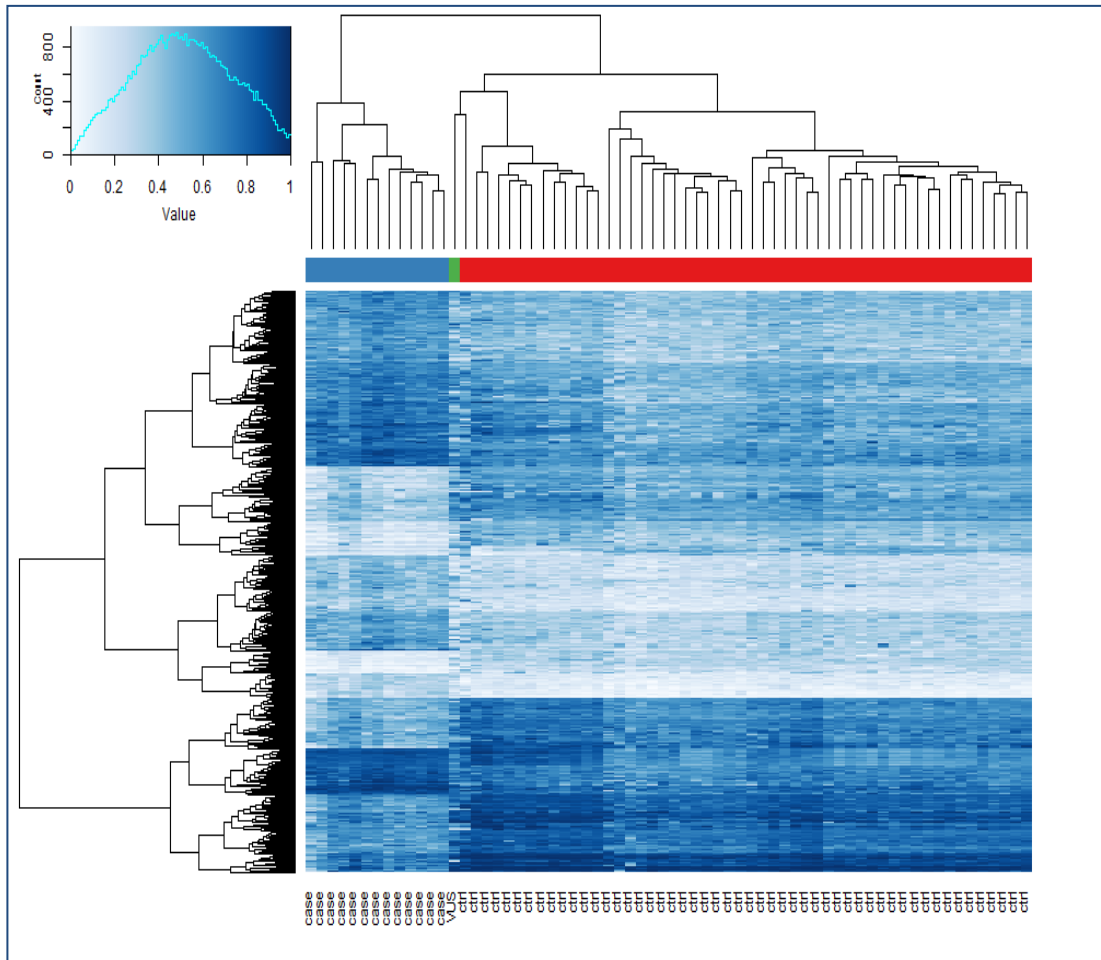


Figure 46 Hierarchical clustering of methylation status.

Patient B (green), patients with Floating Harbour syndrome (blue) and normal controls (red) using an Illumina EPIC methylation array. Patient X clusters with the normal control cohort.

In addition, multiple genomic regions with altered methylation specific to patients with *SRCAP* nonsense mutations were also examined and clearly demonstrated that this patient does not have the characteristic methylation profile previously identified in patients with FHS.

These results confirm that this patient does not have a phenotypic variant of FHS but does not exclude his *SRCAP* variant being associated with his clinical presentation. To further explore the methylation data from our patient, bump-hunter, a bioinformatic tool that removes batch effects and detects regions of interest and attaches statistical uncertainty (Jaffe, Murakami et al. 2012) was used to look for regions with differential methylation from the normal control cohort. A maximum Gap was set to 500bp and the patient was compared with 30 controls. Chromosomes X and Y were excluded.

Regions with more than 15% methylation difference were reported back along with images of each region.

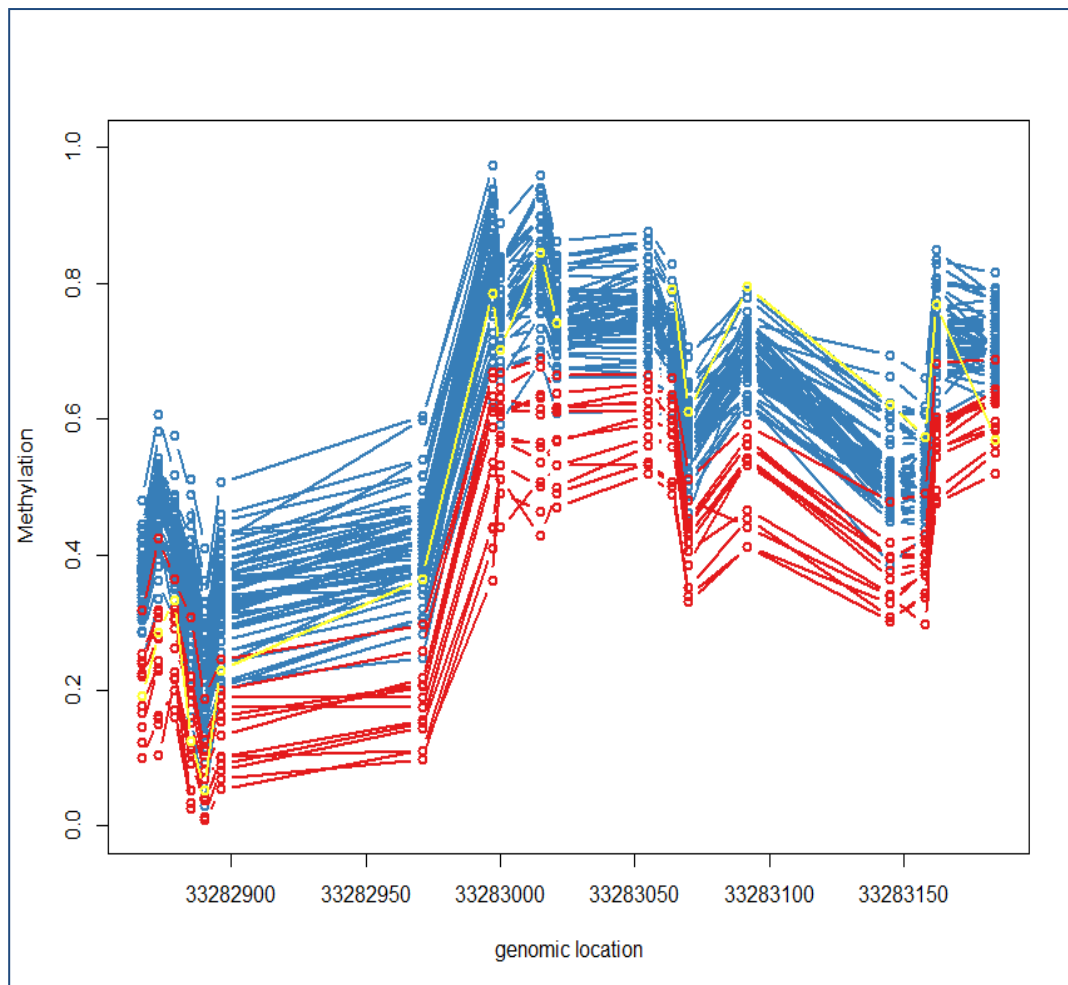


Figure 47 Methylated CpG island at chr6:33282867-33283184

Red:= FHS patients with SRCAP mutations, Blue:= Controls; Yellow = : Patient B c.9029C>A;p.(Pro3010His)

A total of 171 regions with altered methylation status were identified and manually assessed in turn. Several of the reported regions did not appear to be significant as there was overlap between our patient and control patients; these regions were discounted from further analysis.

Patient B appears to have regions that are both hyper and hypo methylated in comparison to controls (for examples see Figure 48 and Figure 49 respectively). Each of these genomic regions was interrogated for genes with a function that could be associated with the patient's phenotype.

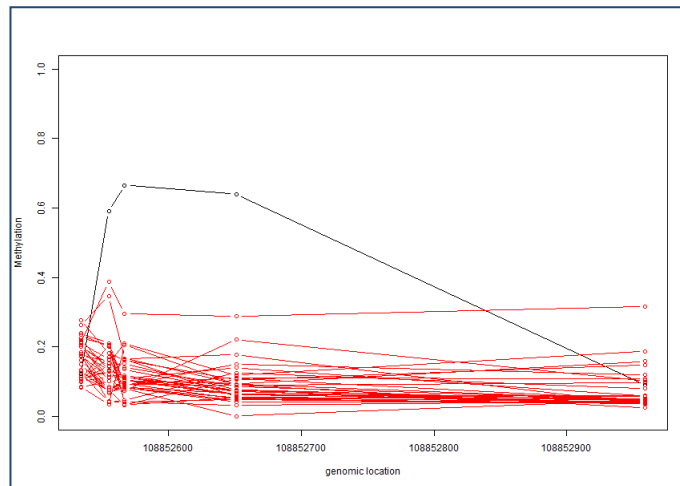


Figure 48 An example of a hypermethylation region plot.

Chromosome 4 genomic coordinates 108852555-108852651 (P value = $7.84E-06$). Black = patient, Red = controls.

This genomic region is within the promoter region of the *CYP2U1* gene. The protein catalyzes the hydroxylation of arachidonic acid, docosahexaenoic acid and other long chain fatty acids. Mutations in this gene are known to be associated with Spastic paraplegia 56 (SPG56), an autosomal recessive neurodegenerative disorder characterized by progressive weakness and spasticity of the lower limbs but with no skeletal features reported.

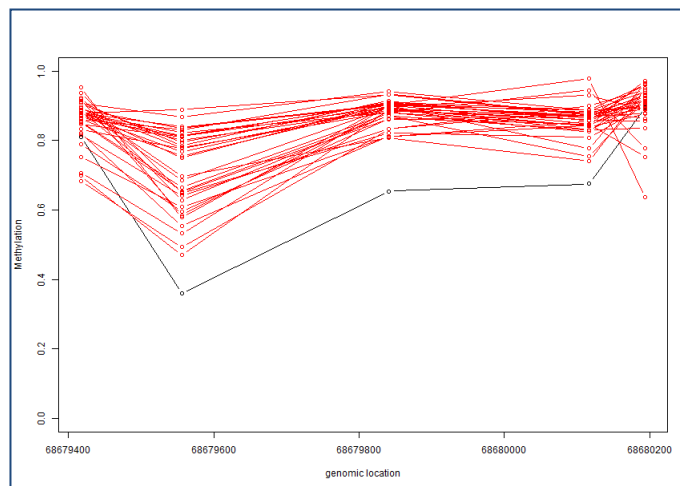


Figure 49 An example of a hypomethylation region plot.

Chromosome 16 genomic coordinates 68679556-68680117 (P value = 0.001248939). Black = patient, Red = controls

This genomic region covers of the 5' end of *Cadherin 3* (*CDH3*), a gene associated with autosomal recessive ectodermal dysplasia, ectrodactyly and macular dystrophy (EEMS and hypotrichosis congenital with juvenile macular dystrophy (HJMD)). The protein is a calcium dependant cell-cell adhesion glycoprotein and not usually expressed in blood.

Two genes of interest were identified, *PCOLCE* and *LOXL3*, both of which are hypomethylated in patient B compared to normal controls ($P= 0.006932885$ and $P=0.011544351$ respectively).

PCOLCE encodes a protein that binds the C-terminal propeptide of type I procollagen and enhances procollagen C-proteinase activity. There is no known phenotype associated with mutations in this

gene. Methylation of the gene promoter region was reduced in patient B (Figure 50), perhaps representing upregulation of the gene.

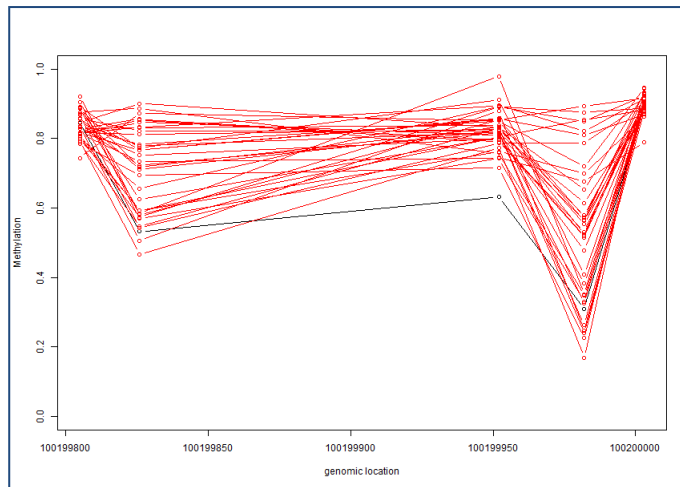


Figure 50 Methylation pattern for the *PCOLCE* gene region (chr7 100199826-100199982).

Patient B compared to normal controls. Black = patient, Red = controls

This region is hypomethylated when compared to controls (P= 0.006932885)

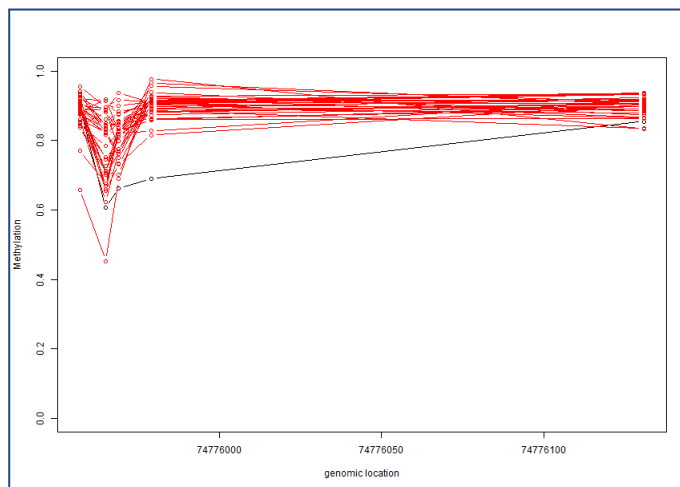


Figure 51 Methylation pattern for the *LOXL3* gene region (chr2 74775965-74775979).

Patient B compared to normal controls. Black = patient, Red = controls

This region is hypomethylated when compared to controls (P=0.011544351)

The *LOXL3* gene encodes lysyl oxidase-like 3. The lysyl oxidase family are known to have both enzymatic and non-enzymatic functions in bone (Trackman 2016). *LOXL3* null mice exhibit perinatal lethality and have cleft palate and spinal deformities (Zhang, Yang et al. 2015) and a homozygous

missense mutation has been described in a family with autosomal recessive Stickler syndrome. (Alzahrani, Al Hazzaa et al. 2015). *LOXL3* is hypomethylated in patient B (Figure 51).

It is not clear whether the altered methylation status of *PCOLCE* and *LOXL3* is a response to, rather than being the primary cause of, altered bone metabolism in this patient.

Activation of CREB-mediated transcription

Another mechanism for the potential pathogenicity of the c.9029C>A variant is through SRCAPs interaction with CREBBP/EP300 and activation of CREB-mediated transcription. Potential pathways include cAMP/CREB signalling pathways in osteoblasts such as mediation of PTH and/or BMP2 signalling. Osteoblast specific deficiency of CREB in mice causes reduced bone mass and decreases BMP2 expression, supporting that BMP2 is a critical transcriptional target for CREB in osteoblasts (Zhang, Edwards et al. 2011).

To investigate this hypothesis further, downstream targets of BMP2 were explored (Figure 52). One target of particular interest is *RUNX2*, a gene that is essential for osteoblast differentiation and bone development. Haploinsufficiency of *RUNX2* is associated with Cleidocranial dysplasia (CCD), an autosomal dominant skeletal dysplasia. Clinical features in CCD are hypoplastic or aplastic clavicles, dental abnormalities and delayed closure of the fontanel. A small subset of patients with severe CCD is reported to have severe osteopenia, absence of the clavicles, marked calvarial hypomineralization and progressive kyphoscoliosis (El-Gharbawy, Peeden et al. 2010).

Clinical presentation of our patient was reviewed to try to identify any features that would be consistent with this mechanism. Severe skeletal dysplasia was identified antenatally and he was born with limb deformities and fractures of the femurs and ribs. At 8 months of age he had developed multiple crush fractures of the vertebrae and went on to develop mild scoliosis by age 15 months. On X-ray he was noted to have 'dimpling' over both thighs and upper arm. A possibility of joint contractures was raised.

At age 18 months he was again noted to have deep pits (dimples) in his lateral thigh compartments but there was no evidence of contractures. It was thought that he was rhizomelic in appearance. Mild hypogammaglobulinaemia was identified and he was non-responsive to pneumococcal vaccination. Facially he has a wide midface but no other features consistent with FHS. No bone biopsy is available for review. He has no radiographic features to support cleidocranial dysplasia. These findings suggest that the *SRCAP* variant is not acting via CREBBP/EP300 to reduce downstream transcription of target genes such as BMP2.

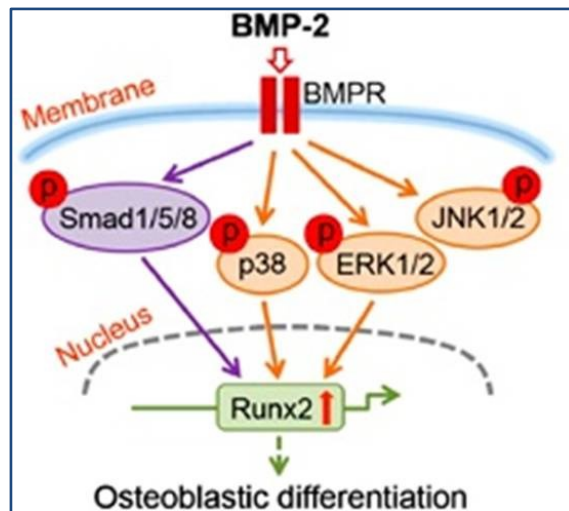


Figure 52 Downstream targets of BMP2.

BMP2 activates the canonical BMP/Smad signalling pathway and non-canonical MAPK pathways to regulate osteoblast differentiation of progenitor cells. Figure modified from Huang et al. (Huang, Yuan et al. 2014).

Electron Microscopy Ultra Structure Evaluation

In the search for further clues as to the cause of patient's B symptoms, electron microscopy examination of a skin biopsy was undertaken. Collagen bundle packing appeared normal, fibre diameter was as expected for age and there was minimal diameter variation. No collagen flowers were seen.

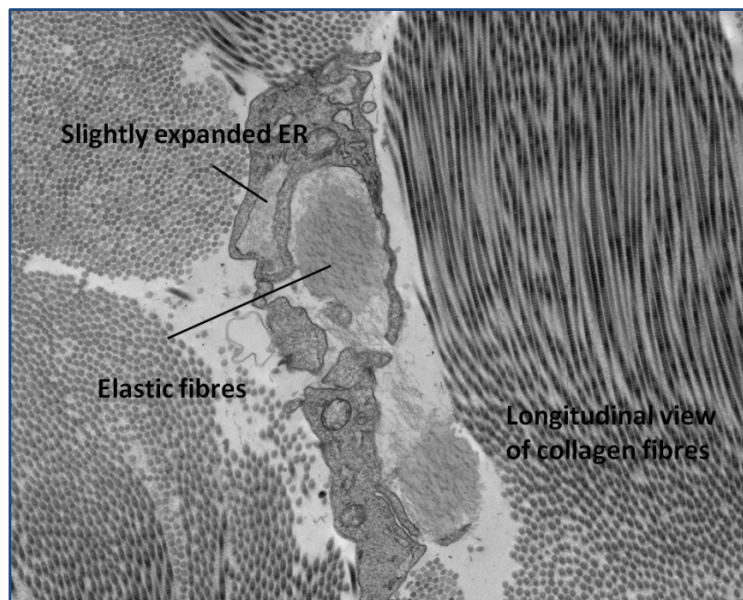


Figure 53 Ultra structure of collagen and fibroblast from reticular dermis of Patient B

Cross section of a fibroblast showing only slightly expanded ER. Collagen fibrils appeared normal and no collagen flowers were seen in this patient (3300X).

The majority of fibroblasts appeared normal. A single fibroblast with slightly expanded ER was identified (Figure 54).

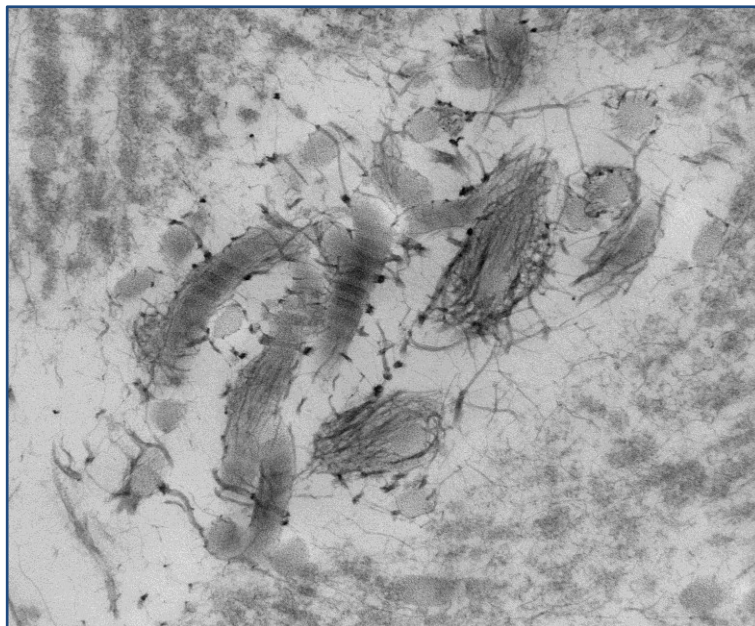


Figure 54 Proteoglycan wrapping around collagen fibres in dermis from patient B (20000X).

The significance of this observation is unclear

An interesting feature was the presence of proteoglycan wrapping around collagen fibres (Figure 54). The significance of this is unclear but this was also identified in patient E, who has a mutation in the *P4HB* gene (see Figure 68). In light of this finding Patient B's sequence data was further interrogated for variants in *P4HB* and other genes in the prolyl 4-hydroxylase complex. No candidate variants were identified.

vi.ii.iii **TAPBP c.1169C>T;p.(Pro390Leu)**

An additional finding of interest, given the mild hypogammaglobulinaemia and non-response to pneumococcal vaccination in this individual, is a *de-novo* c.1169C>T;p.(Pro390Leu) variant in the *TAPBP* gene (NM_003190) identified by exploration of this patient's sequence dataset. This gene encodes the Tapsin protein that mediates the association of the MHC class I with the transporter associated with antigen processing protein complex (TAP). Tapsin is highly expressed in neutrophils and is essential for the transport of antigenic peptides across the ER (Figure 55).

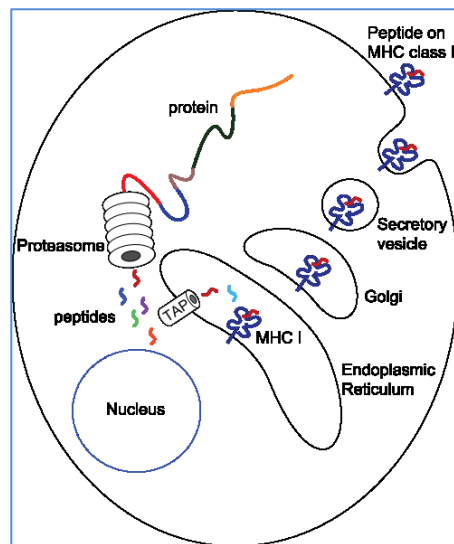


Figure 55 Transport of antigenic peptides across the ER.

Simplified diagram of cytoplasmic protein degradation by the proteasome, transport into endoplasmic reticulum by TAP complex, loading on MHC class I, and transport to the surface for presentation.

Figure By Scray, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=6251017>

Bi-allelic loss of function mutations in *TAPBP* are associated with Type I Bare Lymphocyte Syndrome, a condition characterised by chronic bacterial infection that is usually confined to the respiratory tract and manifests in the first decade of life. The c.1169C>T;p.(Pro390Leu) variant is predicted to be possibly damaging by Polyphen but tolerated by SIFT. The probability of this gene being loss of function intolerant in gnomAD is given as pLI=0.00 and therefore without a second likely pathogenic variant being identified in this patient the clinical significance of this finding is unclear.

This patient was subsequently entered into the DDD study (DDD patient number 136) and the *de-novo SRCAP* variant was reported back to the clinician as a candidate gene for the symptoms in this individual. The *TAPBP* variant was confirmed in the DDD data and listed in their candidate gene list but was not reported back to the clinician.

However, further evidence for the putative causative role of this *SRCAP* missense variant is needed. It should be noted that a reported frequency of *de-novo* point mutations per generation is in the region of 70-175, of which approximately 3 would be expected to cause changes in protein coding (Nachman and Crowell 2000, Conrad, Keebler et al. 2011). It therefore follows that a *de-novo* missense variant may not necessarily be associated with a deleterious clinical phenotype.

vi.iii Patient C

Personalis reported 4 heterozygous variants that were potentially related to the clinical presentation of patient C: c.1041_1043dupGGC;p.(Ala350dup) in *ARID1B*, c.3404T>C;p.(Ile1135Thr) in *SKIV2L*, and c.5741G>A;p.(Arg1914His) and c.3010C>T;p.(Arg1004*) both in the *NBAS* gene.

vi.iii.i ARID1B c.1041 1043dupGGC;p.(Ala350dup)

ARID1B (NM_020732.3) is associated with autosomal dominant Coffin-Siris syndrome (CSS) characterized by varying degrees of developmental delay, hypotonia, hirsutism/hypertrichosis, and sparse scalp hair. Congenital anomalies may include cardiac malformations, gastrointestinal, genitourinary, and central nervous systems abnormalities. Patient C is described as having severe intellectual disability, abnormality of the hair line and hair hypoplasia, symptoms which overlap with those described in CSS.

Previously reported pathogenic mutations in *ARID1B* included nonsense, frameshift and microdeletion (Hoyer, Ekici et al. 2012) mutations predicted to result in haploinsufficiency. The variant in patient C results in the addition of one extra alanine within the 11 Ala repeat region of exon 1 of the gene. Variants in this polyalanine tract have not been reported in association with disease and this particular copy number variant is reported in 1.17% of individuals in HGVD dataset. This variant therefore is most likely benign.

vi.iii.ii SKIV2L c.3404T>C;p.(Ile1135Thr)

Nonsense mutation in the *SKIV2L* gene have been reported in association with autosomal recessive trichohepatoenteric syndrome, a condition characterised by intractable diarrhoea, facial dysmorphism, intrauterine growth retardation, immunodeficiency, and hair abnormalities (Fabre, Charroux et al. 2012). *In-silico* conservation analysis of this variant did not support a likely deleterious effect on protein function and it was considered unlikely to be contributing to the symptoms of patient C.

vi.iii.iii NBAS c.5741G>A;p.(Arg1914His) and c.3010C>T;p.(Arg1004*)

The *NBAS* gene is reported in an autosomal recessive condition characterised by growth retardation, abnormal shaped neutrophil nuclei and optic atrophy with resulting loss of visual acuity. This is known as Short stature, optic nerve atrophy, and Pelger-Huet anomaly syndrome (SOPH) and was first described in an isolated Yakut population in Siberia (Maksimova, Hara et al. 2010). Additional clinical features include fine hair, facial asymmetry, long face with small features and narrow forehead, hypermobility of small joints and muscle hypotonia.

Compound heterozygous mutations in *NBAS* have more recently been described in acute liver failure (ALF)(Haack, Staufner et al. 2015) and further reports have suggested that this represents a multi-system disorder (Staufner, Haack et al. 2016) with phenotypic variability ranging from isolated ALF to a broader spectrum including short stature, osteopenia, optic atrophy, immunological abnormalities, and minor syndromic features such as hypotelorism. Many of these features were consistent with the clinical presentation in patient C. However, bone fragility severe enough to need bisphosphonate therapy, as was the case for patient C, had not been reported.

To explore this finding further the full set of variants reported by Personalis was investigated, primarily to try to identify any unreported variants that could be responsible for the bone fragility in this patient. Firstly, the sequence coverage was assessed to identify any potential regions of low coverage. The average depth in the target region was 115X which was considered adequate. The total number of variants was 189,298. After quality and frequency filtering 19,772 remained.

After filtering, loss of functions variants, i.e. those annotated as frameshift, splice site acceptor, splice site donor, or stop gained were extracted. There were 89 of these variants, and after manual review, none were considered to be in genes with an apparent function in bone. Finally, a larger group of missense variants identified using strategy 2 and 3 (exomizer and targeted gene panel) was compiled. There were 23 of these targeted missense variants all of which were considered to be likely benign following *in-silico* conservation analysis and manual review.

The presence of the *NBAS* variants in the patient was confirmed by Sanger sequencing. Subsequent sequencing of the parents showed that c.3010C>T;p.(Arg1004*) is present in the mother and c.5741G>A;p.(Arg1914His) in the father (Figure 56).

The c.5741G>A;p.(Arg1914His) variant has previously been described in the homozygous state in patients with SOPH syndrome (Maksimova, Hara et al. 2010) and in a single case with an overlapping SOPH and ALF phenotype who is heterozygous for this variant and a p.(Glu943*) nonsense mutation (Kortum, Marquardt et al. 2017).

The c.3010C>T;p.(Arg1004*) variant has been reported in an individual compound heterozygous for this variant and a c.3164T>C;p.(Leu1055Pro) variant in association with acute liver failure (Haack, Staufner et al. 2015, Staufner, Haack et al. 2016).

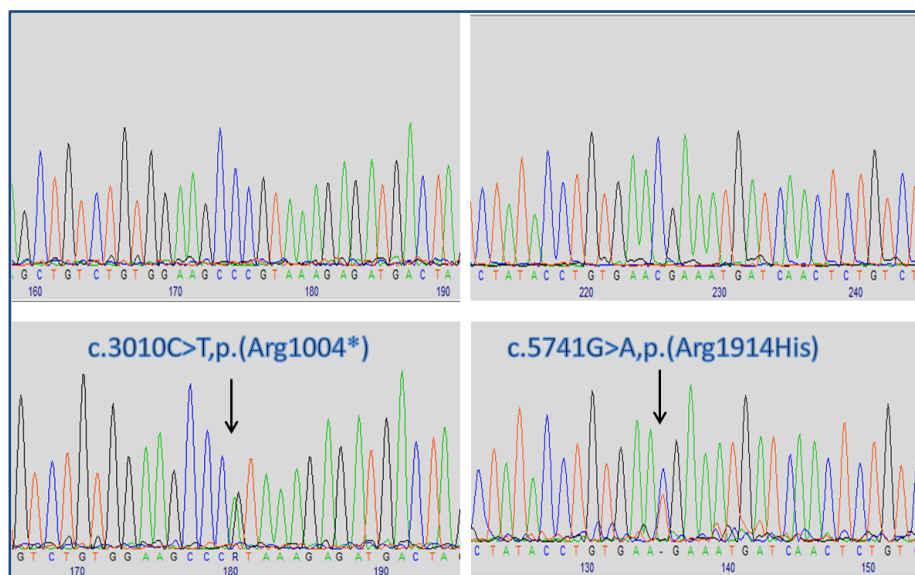


Figure 56 Sanger sequence confirmation of heterozygous *NBAS* variants in the parents of patient C.

The c.3010C>T;p.(Arg1004*) is present in the mother (bottom left) and c.5741G>A;p.(Arg1914His) in the father (bottom right). Normal control sequences are shown top right and top left respectively.

Western blot analysis of *NBAS* protein levels (performed by Dr S Brown, Sheffield RNAi Screening Facility, University of Sheffield) showed reduced levels in the patient's cultured fibroblasts compared to normal controls (Figure 57) and examination of peripheral blood film confirmed the presence of Pelger Huët anomaly (Figure 58) both further supporting that these two variants are pathogenic.

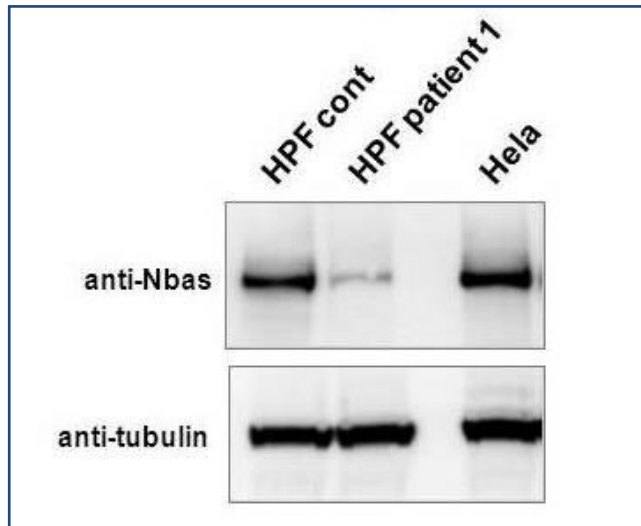


Figure 57 Western blot of cultured fibroblasts showing reduced NBAS protein levels compared to controls.

HPF: human primary fibroblasts

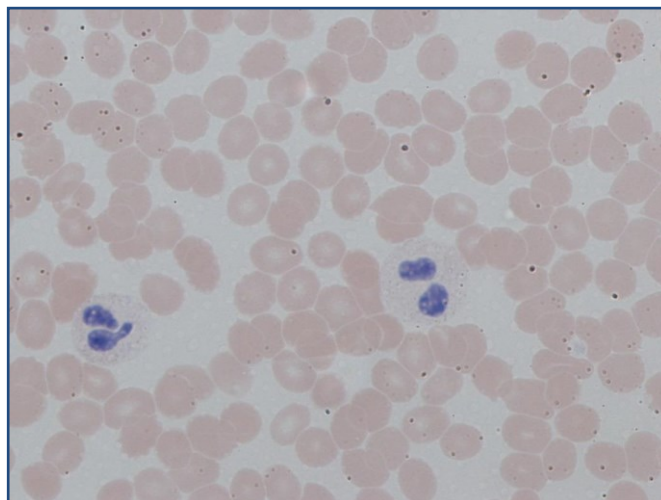


Figure 58 Peripheral blood film showing the presence of Pelger Huët anomaly in patient C.

Stained using May Grunwald Giemsa.

Electron Microscopy Ultra structure Evaluation

Electron microscopy showed normal collagen bundle packing. The mean collagen fibril diameter (75nm) was very mildly reduced for age (normal range 80-90nm (Stewart 1995)) There was no marked increase in diameter variability, apart from a small area near a vessel that showed variable diameter, the significance of which is unclear (Figure 59). Occasional medium sized collagen flowers (~200nm in diameter) were seen (Figure 60).

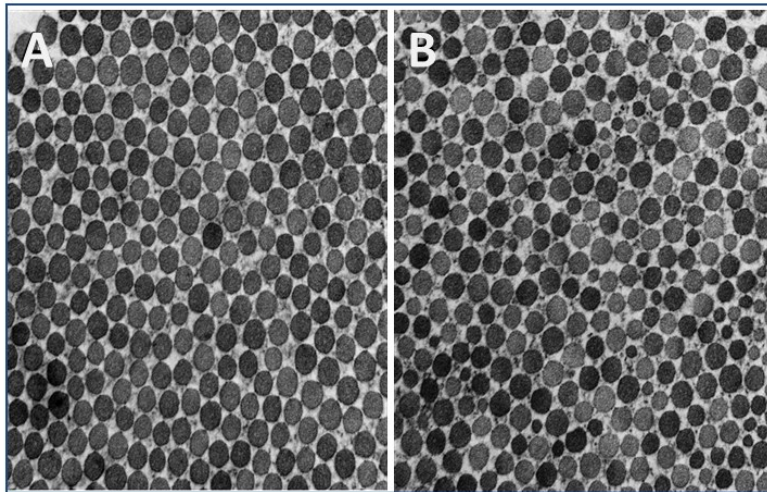


Figure 59 Electron microscopy of collagen fibrils in patient C

Electron microscopy showed normal collagen bundle packing. Collagen fibre diameter was within expected range for age (A) apart from a small area near a blood vessel that showed variable diameter (B) the significance of which is unclear (20,000X magnification).

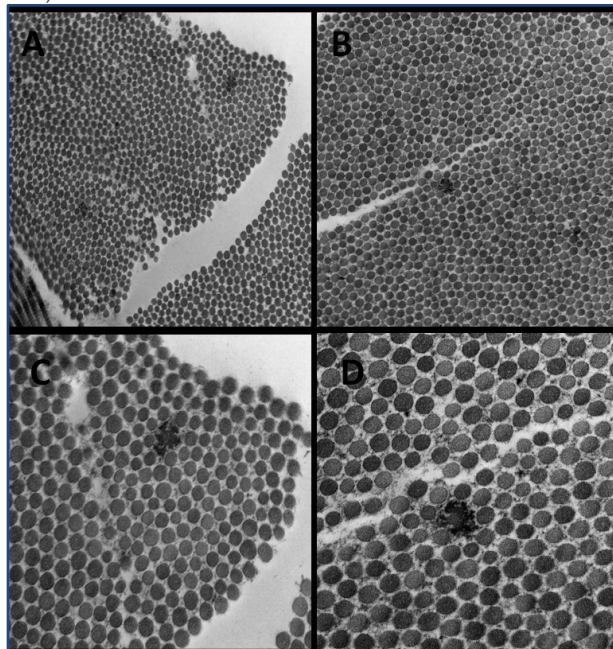


Figure 60 Electron microscopy of collagen fibrils showing occasional collagen flowers in patient C (~200nm in diameter).

A-B 9200X magnification, C-D 20,000X magnification

Ultra structure examination of fibroblasts showed grossly expanded protein filled ER (Figure 61). This protein has formed 'aggregates' within the ER of some fibroblasts and the golgi appears swollen (Figure 62).

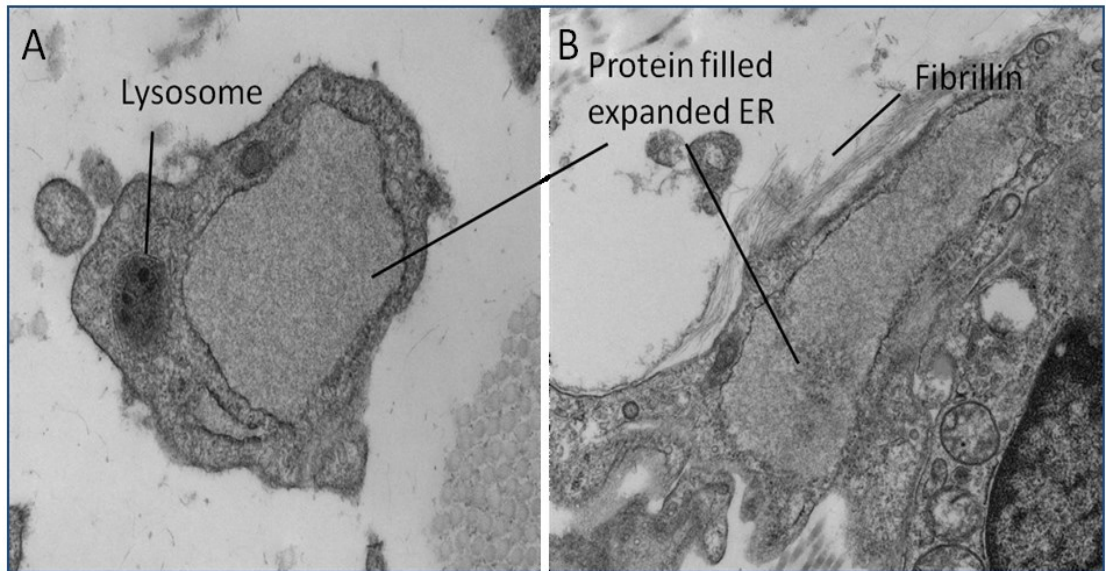


Figure 61 Electron microscopy of dermal fibroblasts from patient C showing markedly expanded protein filled ER.

A: 15600X magnification, and B 9200X magnification

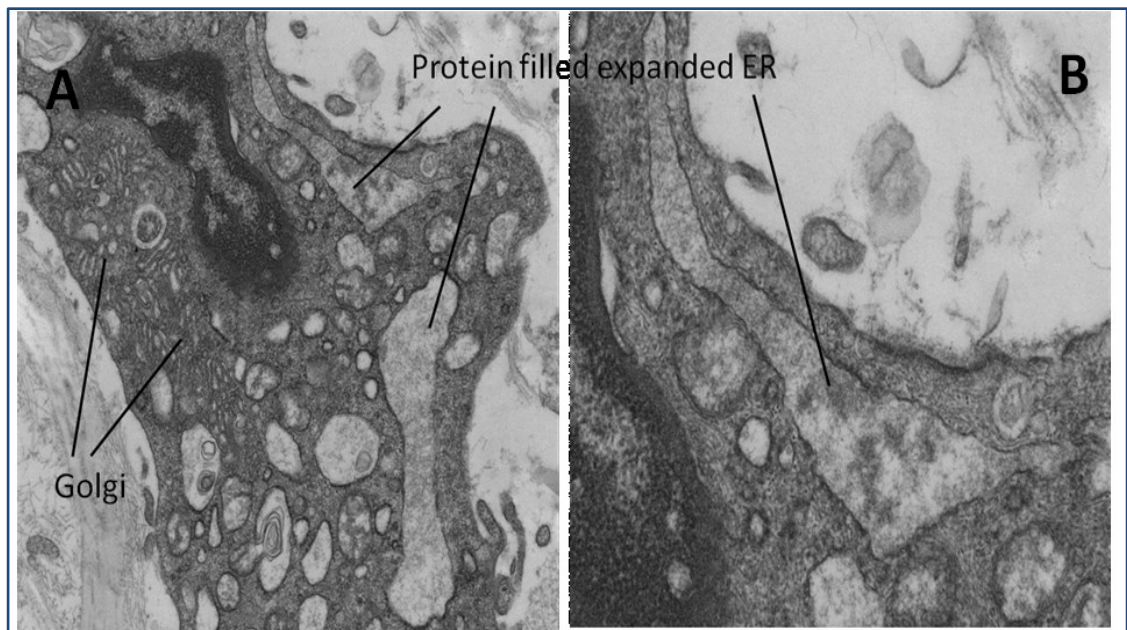


Figure 62 Electron microscopy of dermal fibroblast from patient C showing the presence of protein aggregates within the expanded ER.

A The ER is expanded and protein filled and the golgi appears swollen (9200X magnification). B The presence of protein aggregates within the ER is shown (20000X magnification).

Also noted within some fibroblasts were electron dense bodies in the lysosomes (Figure 63). While it is normal to accumulate these electron dense bodies with age, it is not expected in a child of this age. This may represent autophagy within the cell.

The *NBAS* gene encodes the neuroblastoma-amplified sequence protein and is a component of a soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) complex. SNAREs facilitate the docking and fusion of transport vesicles with target membranes, via membrane bound proteins on transport vesicles (v-SNARE) and target membranes (t-SNARE). The *NBAS* protein interacts with t-SNARE p31 and, along with other components, forms the syntaxin 18 complex that is involved in golgi to ER retrograde transport (Aoki, Ichimura et al. 2009). The disruption to this golgi-ER transport mechanism is supported by the EM findings in patient C.

Functional studies in fibroblasts from 14 patients with ALF and *NBAS* mutations demonstrated that decreased *NBAS* protein was mirrored by reduction in p31 supporting the role of *NBAS* in the SNARE complex (Staufner, Haack et al. 2016). Thermal instability of syntaxin 18, evidenced by compromised growth and further reduction of both *NBAS* and p31 in temperature-challenged fibroblasts, and subsequent ER stress induced apoptosis, has been suggested to be the fever-dependent mechanism by which ALF is triggered in *NBAS* patients (Staufner, Haack et al. 2016).

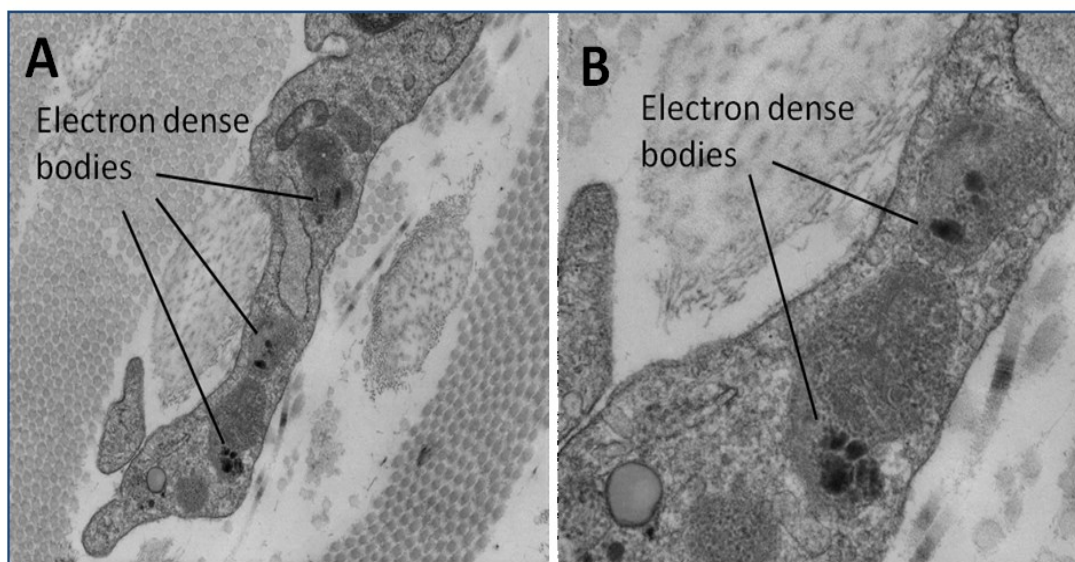


Figure 63 Electron microscopy of dermal fibroblast from patient C showing electron dense bodies within lysosomes.

A 2600X magnification. B 20000X magnification

Studies in zebrafish and *C.elegans* have also identified that the *NBAS* protein plays an important role in the nonsense mediated (NMD) decay pathway, a surveillance mechanism that targets mRNAs harbouring premature termination codons, preventing accumulation of shortened proteins that potentially interfere with normal cellular function. NMD is also known to regulate gene expression and *NBAS* has been shown to act, along with core NMD factors, to regulate expression of genes associated with bone mineralisation, osteoblast differentiation and bone development. In particular, *NBAS*-depletion has been demonstrated to result in upregulation of the *MGP* gene (Matrix Gla Protein) which acts as an inhibitor of bone formation (Longman, Hug et al. 2013).

As *NBAS* is proposed to function in both Golgi-ER transport and the NMD pathway, the mechanism by which the mutations identified in patient C result in bone fragility could be attributed to either pathway or to a combination of both. Interestingly, although our patient has had abnormal liver function tests since a young age, he has never had any episodes of ALF.

Patient C is the second child of healthy, non-consanguineous North European parents. He was born at 33 weeks gestation by spontaneous breech delivery following a normal pregnancy weighing 1.75kg (9th centile). He needed continuous positive airway pressure (CPAP) ventilation for 24 hours and phototherapy for jaundice. He had recurrent infections and significant feeding problems, requiring percutaneous gastrostomy insertion. He was diagnosed with a horizontal nystagmus, bilateral optic atrophy and myopia, but his hearing is normal.

His growth parameters remain below the 0.4th centile and at 7 years of age he weighed ~13.1kg (<0.4th centile), with height ~ 88cms (<0.4th centile) and head circumference ~ 49.5cms (0.4th-2nd centile). A skeletal survey showed multiple Wormian bones, slender tubular bones and osteopenia. Following recurrent fractures, a transiliac bone biopsy was taken and showed osteoporosis with high bone turnover with marked periosteal bone resorption, which was not consistent with a classical OI phenotype. He has moderate intellectual disability and a high-pitched voice. ArrayCGH (60K), FRAX, UPD7 and 11p15 methylation genetic testing were all negative.

On examination at 9-years of age, he was dysmorphic with proptosis, had a progeroid skin appearance, bilateral low-set ears, dentinogenesis imperfecta, and bilateral 5th finger clinodactyly. On lateral spine radiograph he has loss of vertebral height and a low lumbar bone BMAD Z-score of -3.5. Marked improvement of his bone health was achieved with Pamidronate treatment.

A second patient with biallelic *NBAS* mutations was highlighted by analysis of the CAP DDD data (Chapter viii) due to the HPO term 'Recurrent fractures' (HP:0002757). This patient had the same missense variant as patient C, c.5741G>A;p.(Arg1914His) inherited from the mother and a novel nonsense mutation, c.2032C>T;p.(Gln678*), inherited from the father. This patient was also described as having elevated hepatic transaminases in his DDD phenotypic data.

We concluded that mutations in *NBAS* result in a variable clinical presentation and that these two patients show that symptoms can include a skeletal phenotype with bone fragility. Vertebral and long bone fractures can result, potentially requiring treatment with bisphosphonates.

To facilitate rapid transition of *NBAS* gene testing into the diagnostic pathway for patients with appropriate clinical features, a gene dossier was prepared and submitted to the UK Genetic Testing Network (UKGTN). The UKGTN provides a mechanism for national evaluation of proposed new tests for NHS patients. The gene dossier covers characterises of the disorder/condition, testing strategies, clinical utility and testing criteria (see Appendix M, page 303). The gene dossier was accepted and is now published:

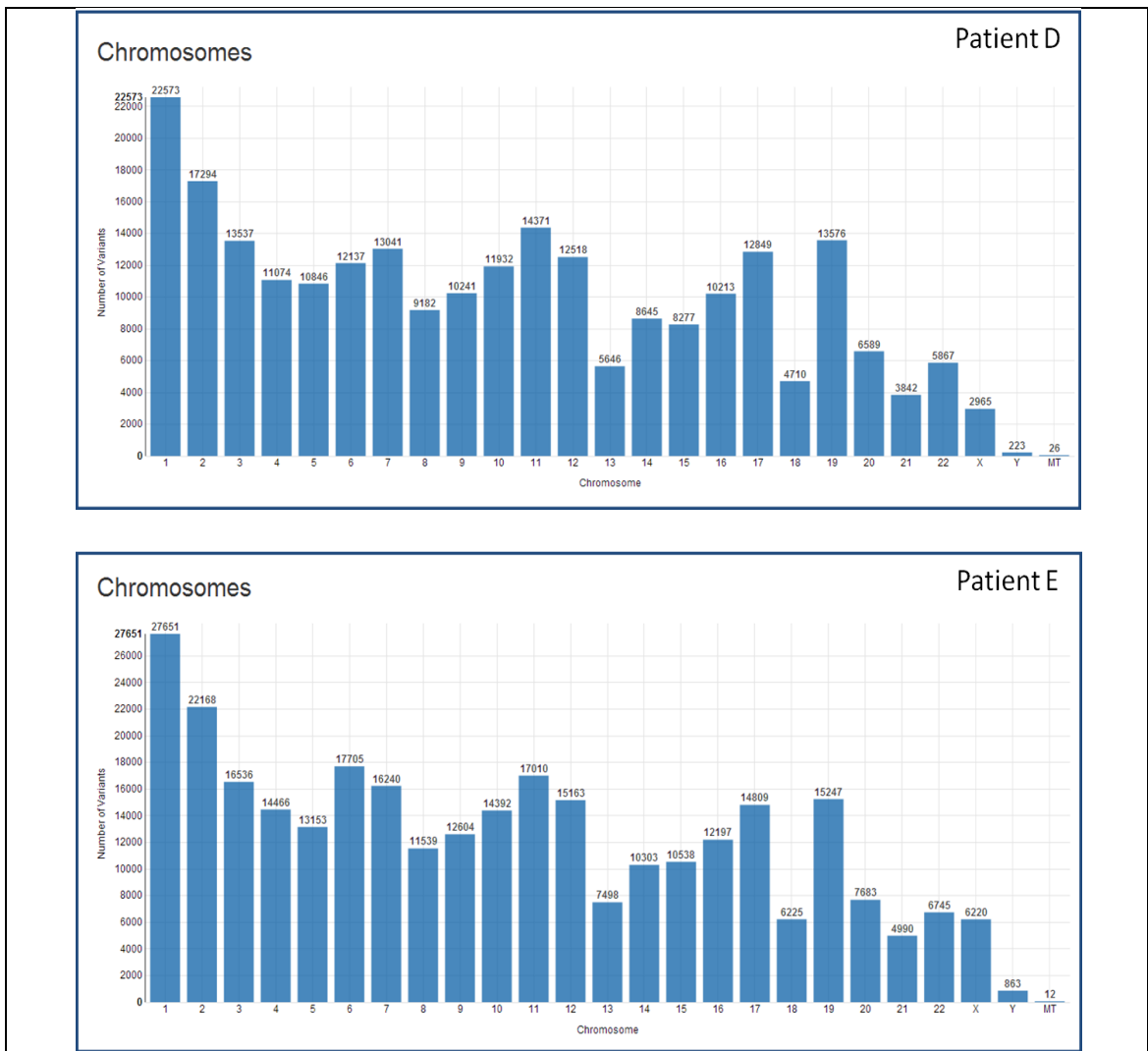
https://ukgtn.nhs.uk/uploads/tx_ukgtn/NBAS_related_Disorder_GD_Sept_2017.pdf

Sequence analysis of the *NBAS* gene has been introduced as a diagnostic test for inherited bone fragility at SDGS.

These two *NBAS* patients are presented in the co-authored manuscript 'Compound heterozygous variants in *NBAS* as a cause of atypical osteogenesis imperfecta' (Balasubramanian, Hurst et al. 2017), (Appendix G page 240)

vii RESULTS: WHOLE EXOME SEQUENCING (SDGS)

Five patients (participant ID D, E, F, G and H) were sequenced at SDGS using the SureSelect Human All Exon V6 baits for target enrichment. Parental samples were also sequenced with the exception of patient D where only maternal DNA was available. The total number of variants identified for each patient is illustrated in Figure 64 below.



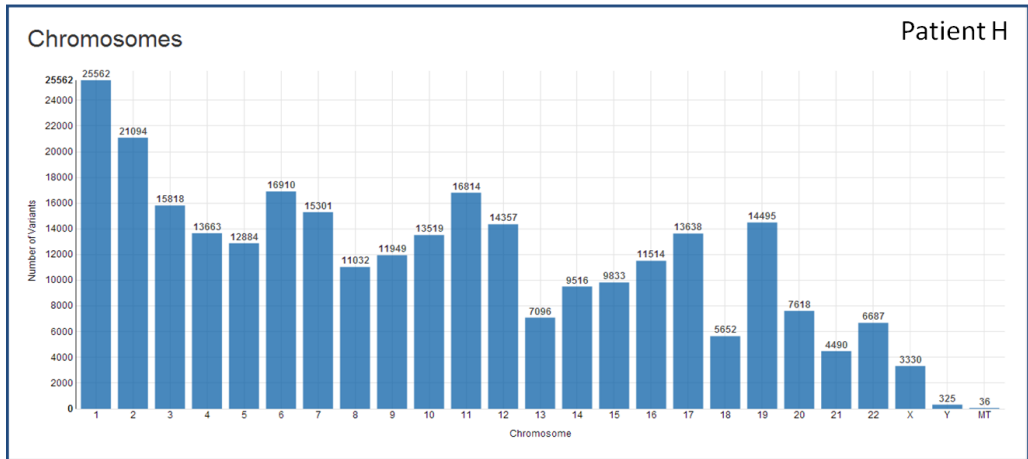
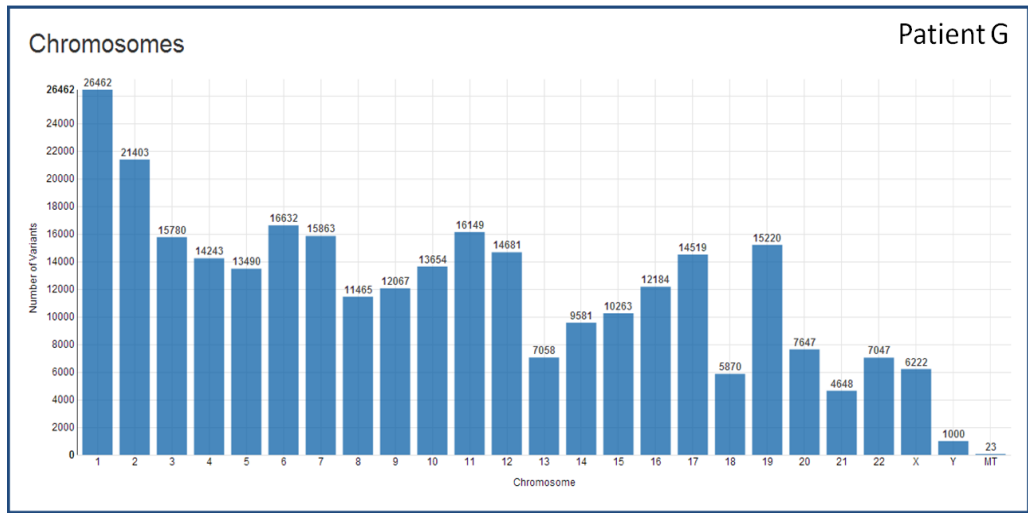
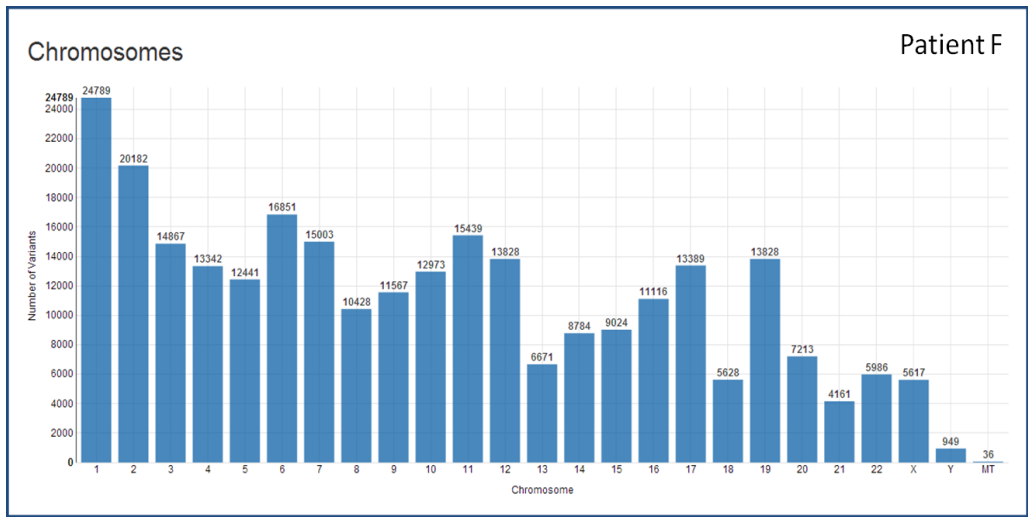


Figure 64 Distribution of variants identified in patients D-H using whole exome sequencing. Sequencing performed using the SureSelect Human All Exon V6 probe set and the Illumina HiSeq platform. The total number of variants identified in each chromosome is illustrated.

The total variants for each patient was first filtered against frequency in the general population, those with a minor allele frequency (MAF) above 1% were excluded. Variants were then filtered using inheritance (strategy 1, section iii.ix.i page 57) and using the candidate gene list (strategy 3, section iii.ix.iii page 59). A number of variants requiring manual assessment remained (Table 22). In addition, the total variants identified were prioritized using strategy 2 (Section iii.ix.ii, page 58) and the top twenty candidates were manually assessed.

Variant filtering	Patient D	Patient E	Patient F	Patient G	Patient H
Total variants	244836	306063	278239	297609	286512
<1%MAF	438	498	507	594	486
Inheritance (strategy 1)	167	–	68	61	77
Candidate genes (strategy 3)	2	23	3	2	2

Table 22 Variants identified in patient D-H.

Total variants and those remaining after filtering for minor allele frequency (MAF), mode of inheritance (strategy 1) and against the candidate genes list (strategy 3).

Filtering for inheritance was not possible for patient E as a paternal sample was not available.

vii.i Patient D

vii.i.i NBAS c.6971G>A;p.(Arg2324His) and c.4358G>A;p.(Cys1453Tyr)

The clinical features of patient D are described as being developmental delay, hypotonia, muscle weakness, short stature, osteopenia, fractures, high arched palate, frontal bossing, bilateral club feet and low posteriorly rotated ears. Interrogation of his whole exome data identified compound heterozygous variants in the NBAS gene: a paternal c.6971G>A;p.(Arg2324His) and a maternal c.4358G>A;p.(Cys1453Tyr).

The c.6971G>A;p.(Arg2324His) is located in exon 52 and is only 28 amino acids upstream from the final termination codon. It affects a highly conserved amino acid but there is a small physiochemical difference between arginine and histidine. The variant is reported at low frequency in the general population (17/246192 alleles; rs150030912) but no homozygous individuals are recorded. *In silico* conservation analysis is inconclusive.

The c.4358G>A;p.(Cys1453Tyr) is located in exon 37 and again affects a highly conserved amino acid. There is a large physiochemical difference between cysteine and tyrosine but this amino acid is not within a known functional domain of the protein. A population frequency of 146/276780 alleles is recorded in gnomAD (rs148644578) and a single homozygous individual from the Latino population (34412 alleles) is reported. *In-silico* analysis is inconclusive.

Apart from these two variants of unclear significance in the *NBAS* gene, no other candidate variants were identified. The patient and his family have been invited for further follow up investigation, including detailed phenotyping to include an FBC to look for PHA and liver function tests. To date, the family have declined this invitation.

vii.ii Patient E

vii.ii.i *P4HB* c.1178A>G;p.(Tyr393Cys)

Inheritance driven analysis of data from patient E (strategy 1) was limited as a sample of DNA from the patient's father was not available. Using strategies 2 and 3, a heterozygous c.1178A>G;p.(Tyr393Cys) variant in exon 9 of the *P4HB* gene (NM_000918.3) was identified. This variant is predicted to replace the tyrosine at position 393 with a cysteine. Sanger sequencing of the patient confirmed this result and the maternal sample showed no evidence of the variant (Figure 65).

This particular variant in *P4HB* has been described in Cole-Carpenter syndrome 1 (CLCRP1), an autosomal dominant condition characterized by bone fragility, craniosynostosis, ocular proptosis, hydrocephalus, and distinctive facial features (Rauch, Fahiminiya et al. 2015). Craniofacial features include frontal bossing, mid-facial hypoplasia and micrognathia.

Patient E is the first child of healthy, non-consanguineous parents with no significant family history. The pregnancy was unremarkable with no suggestions of antenatal bone fragility and she was born at term with a birth weight of 3710 grams (75th centile and Z score 0.6). She was diagnosed with positional talipes, for which she received physiotherapy. At aged 6 months, she presented with a fracture of the left distal tibia. On clinical examination she was noted to have a large head with a wide open anterior fontanelle, a broad face with small nose and flat nasal bridge, blue sclerae, long fingers, and broad thumbs and toes. Her weight Z score had reduced to -1.4.

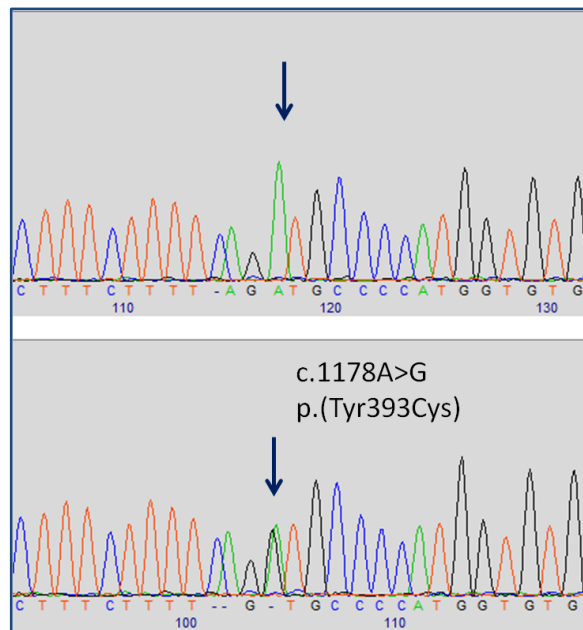


Figure 65 Sanger sequence of exon 9 of the *P4HB* gene in patient E.

Patient E (bottom) is heterozygous for the c.1178A>G;p.(Tyr393Cys) variant associated with Cole-Carpenter Syndrome. Her mother (top) showed no evidence of the variant.

A skeletal survey showed diffusely osteopenic bones with multiple vertebral body compression fractures in the thoracic and lumbar spine. Fractures of the anterior left 8th, 9th and 10th ribs were identified and historic fractures of both radial shafts were noted. Her long bones of both upper and lower limbs showed crumbling fractures with widening and sclerosis of the distal metaphyses. Wormian bones were seen in the skull. A clinical diagnosis of osteogenesis imperfecta (OI) was made and she commenced intravenous bisphosphonate (pamidronate) infusion therapy.

She continued to have non-traumatic fractures of upper and lower limb long bones until aged 2 years and 8 months of age when she had corrective tibial osteotomies and bilateral tibial roddings with Fassier–Duval lengthening rods. There has been a significant improvement of vertebral fractures with remodelling of all vertebral bodies while on bisphosphonate infusions, which she continues to receive.

At clinical review at aged 3 years, she had not sustained any fractures in the previous 18 months. She had a broad face with frontal bossing, flat nasal bridge, prominent eyes and bilateral low-set ears. Her growth parameters were, weight 12.2 kilograms (2nd-9th centile), height 77 cms (<0.4th centile) and head circumference 54 cms (99.6th centile). Bilateral limb deformities were evident and she is currently awaiting surgical correction and rodding procedures of her bilateral femur and upper limb bones. Although her gross motor development has been delayed due to limb deformities, her speech, language and social skills have progressed well and she attends a mainstream nursery.

P4HB encodes the beta subunit of prolyl 4-hydroxylase, an abundant multifunctional enzyme belonging to the protein disulfide isomerase (PDI) family that catalyzes the formation, breakage and rearrangement of disulfide bonds in polypeptide chains, therefore having a key role in protein folding.

The P4HB protein also functions as a chaperone that inhibits aggregation of procollagen alpha chains and forms a tetramer with the prolyl 4-hydroxylase alpha subunit, consisting of 2 α and 2 β -subunits. It is thought that the β -subunit is required to maintain the solubility of the α subunit and keep the complex within the ER where it hydroxylates many of the proline residues within the triple helical domain of type I collagen.

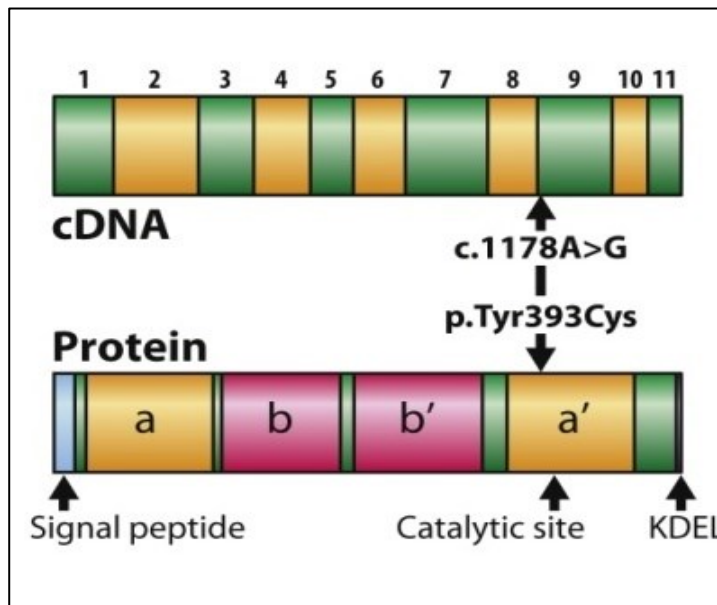


Figure 66 Schematic representation of the P4HB protein and the coding exons of *P4HB* cDNA (NM_000918.3).

The locations of the exons are aligned relative to the regions of the protein that each exon encodes. The domains of the protein are represented by a, b, b', a' with an ER motif (KDEL) present at the C terminus. The location of the c.1178A>G;p.(Tyr393Cys) mutation identified in patient E is indicated. Figure adapted from Rauch et al (Rauch, Fahiminiya et al. 2015)

The P4HB protein contains four TRX-like (thioredoxin-like) domains and a C-terminal extension domain. Two of these domains have disulphide isomerase activity, termed a and a' and are separated by the enzymatically inactive b and b' domains (Figure 66)

The cysteine residue created by the c.1178A>G substitution lies close to two key residues, p.Cys397 and p.Cys400, in the C terminal catalytic site (Figure 67). The p.Cys397 residue interacts with cysteines of many proteins and this is released by p.Cys400 (Wilkinson and Gilbert 2004). Immunoblotting experiments have suggested that p.Cys393 may interfere with this mechanism by forming more stable disulphide bridges (Rauch, Fahiminiya et al. 2015).

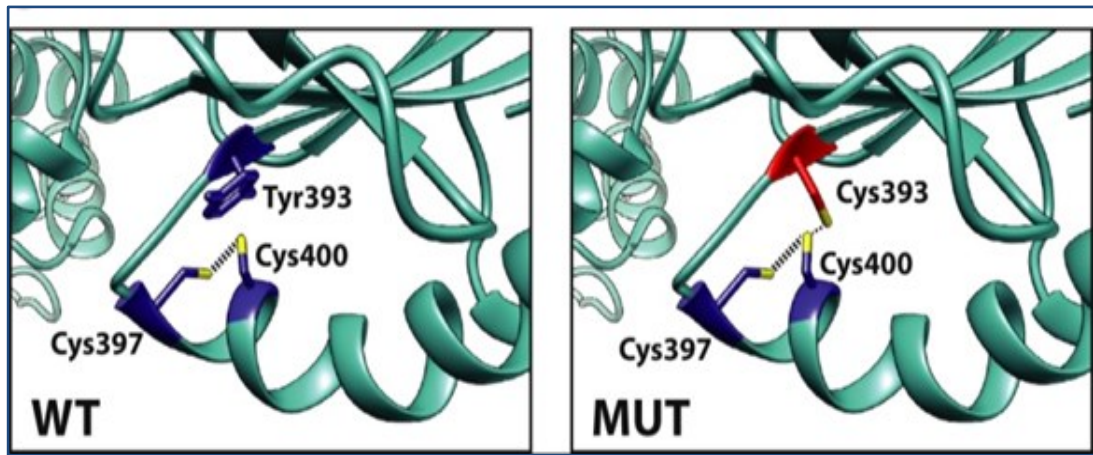


Figure 67 Protein structure model of P4HB indicating potential impact of the p.Tyr393Cys mutation in patient E.

The mutated p.Cys 393 is physically close to p.Cys 397 and p.Cys 400. These residues form the active site of the protein and normally form a disulfide bond that is disrupted by the mutated p.Cys 393 amino acid. Figure adapted from Rauch et al (Rauch, Fahiminiya et al. 2015)

As P4HB also plays a significant role in modification of type I collagen alpha chains it might be expected that at least part of the mechanism by which this variant acts would be via aberrant type I collagen processing. However, studies in fibroblasts from patients with Cole Carpenter have failed to demonstrate significant disruption to modification or secretion of type I procollagen. However, increased expression of heat shock protein 47, another collagen chaperone molecule, suggesting increased ER stress has been observed (Rauch, Fahiminiya et al. 2015). Immunofluorescence studies in patient E also demonstrated no major difference in extracellular type I collagen deposition in fibroblasts (data presented in our manuscript) (Balasubramanian, Padidela et al. 2017).

The potential for ER stress to be a pathogenic mechanism in this patient was investigated further by studying dermal fibroblasts. Elastic tissue, collagen bundle packing and fibroblasts appeared normal. The mean collagen fibril diameter is 74nm, with diameter variability ranging from deviation RMS 7.0 to 10.4nm. No collagen flowers were seen.

The EM data for our patient suggests that there is little to no effect on assembly of collagen fibres or significant fibroblast ER stress. This does not support the hypothesis of ER stress being the main mechanism by which the recurrent *P4HB* mutation causes bone fragility (Rauch, Fahiminiya et al. 2015).

We were also unable to demonstrate any major variation in extracellular type I collagen deposition in our patient's fibroblasts when compared to a normal control (immunofluorescence analysis performed by Dr J McCaughey and Dr D Stephens, School of Biochemistry, University of Bristol, Bristol). It is possible that aberrant PDI protein causes disorganisation of the ECM rather than directly affecting secretion of type I collagen. It would be interesting to see if we could demonstrate altered distribution of the PDI protein in our patient's fibroblasts or increased HSP47 protein as previously reported (Rauch, Fahiminiya et al. 2015). However, currently the mechanism by which the p.Tyr393Cys single amino acid substitution results in bone fragility is still to be resolved.

One previously undescribed observation is the presence of four small collections of individual collagen fibrils surrounded by proteoglycans (Figure 68); the significance of this feature is unclear. As discussed previously, we have also identified this feature in patient B (Section vi.ii, page 88).

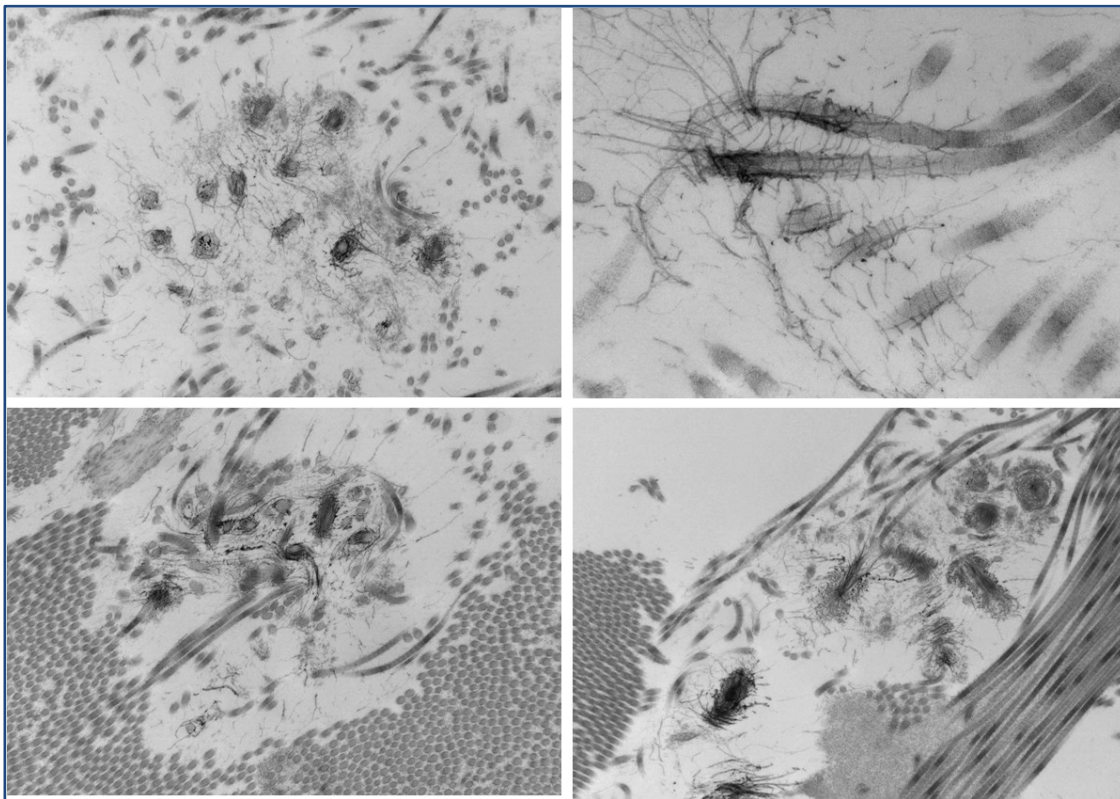


Figure 68 Electron microscopic ultra structure of collagen in patient E

Four small collections of individual collagen fibrils surrounded by proteoglycans were observed. The significance of this is unclear (9200X magnification)

Following confirmation of the molecular finding in patient E, reverse phenotyping confirmed that this patient has features that are consistent with a diagnosis of Cole-Carpenter syndrome. This patient also highlights what appears to be an emerging, distinctive radiological phenotype for these patients: meta-diaphyseal fractures with metaphyseal sclerosis. This finding is uncommon in patients with classical OI and should stimulate targeted mutation analysis in individuals with this radiological feature.

Patient E is the third case of Cole Carpenter Syndrome caused by a defect in *P4HB* and suggests the possibility of a recurrent *P4HB* mutation as a cause of this syndrome. This is an important finding considering the rare occurrence of such patients. Follow up of the patient after bisphosphonate therapy indicated a positive outcome on vertebral compression fractures supporting the current therapeutic approach of anti-resorptive therapy in patients with a proven *P4HB* mutation.

The patient has been presented in the following co-authored posters and manuscript:

Poster Presentation: International Skeletal Dysplasia Society 2017 **Cole Carpenter Syndrome Type 1 (P4HB Missense Mutation) has a Distinct Skeletal Phenotype**. Offiah C A, Padidela R, Pollitt C R, Bishop J N, Wagner E B, McCaughey J, Stephens J D, Balasubramanian M. (Appendix page 224)

Poster Presentation: International Conference on Childrens Bone Health 2017: **P4HB recurrent missense mutation causing ColeCarpenter syndrome: exploring the underlying mechanism**. M Balasubramanian, R Padidela, R Pollitt, N Bishop, Z Mughal, A Offiah, B Wagner, J Mccaughey, D Stephens. (Appendix page 225)

Manuscript: **P4HB recurrent missense mutation causing Cole-Carpenter syndrome** Meena Balasubramanian, Raja Padidela, Rebecca C Pollitt, Nicholas J Bishop, M Zulf Mughal, Amaka C Offiah, Bart E Wagner, Janine McCaughey, David J Stephens. J Med Genet. 2017 doi: 10.1136/jmedgenet-2017-104899. (Appendix page 232)

vii.iii Patient F

No candidate variants were identified in patient F whose clinical features are described as multiple fractures, blue sclerae, dentinogenesis imperfect and hypermobile joints. Interestingly the patient's mother is also described as being hypermobile with fractures following insignificant trauma. A maternally inherited variant was also sought in this patient but without success.

vii.iv Patient G

vii.iv.i P3H1 c.1224-80G>A

Patient G has failure to thrive, macrocephaly, persistent large anterior fontanelle, multiple fractures and dysmorphic features. She was enrolled in this study following identification of a heterozygous c.1080+1G>T splicing mutation in *P3H1* by diagnostic targeted exome analysis (Figure 69). No second likely mutation was identified in this gene and no mutation was identified in any other known OI gene.

The c.1080+1G>T is a common OI mutation in the African population and is reported in association with OI type II and III (Cabral, Chang et al. 2007, Baldrige, Schwarze et al. 2008, Pepin, Schwarze et al. 2013). Analysis of this patient's exome data confirmed the presence of the heterozygous c.1080+1G>T *P3H1* mutation and that this was inherited from her unaffected father.

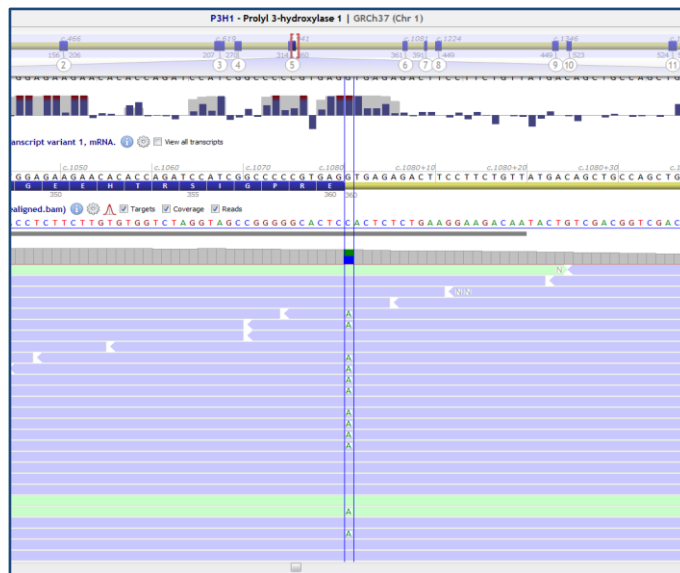


Figure 69 *P3H1* c.1080+1G>T splicing mutation in patient G identified by targeted exome sequencing.

The figure shows the structure of the *P3H1* gene (top) and the position of patient's sequence, with the 'A' substitution in green, below.

Additionally, a heterozygous c.1224-80G>A variant in intron 7 of the *P3H1* gene was identified in the exome data. The segregation of this mutation in this family was confirmed by Sanger sequencing to be present in the patient's unaffected mother (Figure 70). There is no evidence of the variant in her unaffected father (Figure 71).

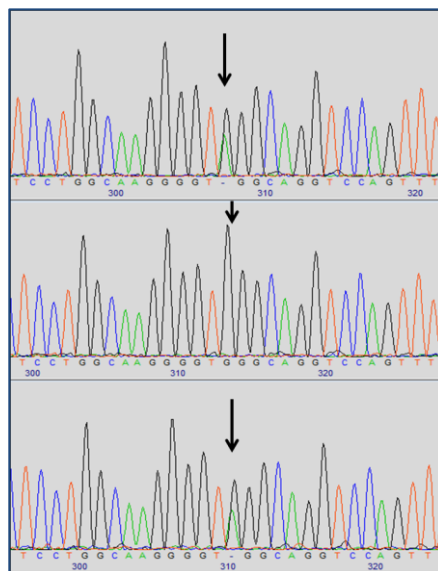


Figure 70 *P3H1* c.1224-80G>A variant identified by whole exome sequencing

Sanger sequence confirmation of c.1224-80G>A present in patient G (top) and her unaffected mother (bottom). There is no evidence of the variant in her unaffected father (middle.)

The c.1224-80G>A variant is not reported in the general population and splice site prediction analysis suggests that this may create a preferred cryptic donor splice site. Three prediction packages predicted the cryptic site to be favoured over the wildtype donor site: MaxEnt (5.30 vs 3.71), NNSplice (0.34 vs 0.0) and GeneSplicer (3.49 vs 0.64). Two prediction packages suggested similar splicing efficiency: SSF (72.13 vs 75.43) and HSF (82.87 vs 84.57). If the new cryptic splice site was used in preference to the wildtype donor site then this could lead to the addition of 92 base pairs to the mRNA sequence (Figure 71).

If this prediction is accurate then a frameshift would result, potentially leading to mRNA decay. Taken together with the previous heterozygous c.1080+1G>T *P3H1* mutation, this would lead to complete loss of *P3H1* protein function. Subsequent discussion with the referring clinician confirmed that this patient's clinical presentation would be consistent with *P3H1*-associated OI.



Figure 71 *P3H1* c.1244-80G>A splice site prediction analysis

Wildtype prediction (top) shows the presence of a weak cryptic splice site. The substitution of the 'A' for a 'G' at position c.1244-80 is indicated as a vertical blue line. The presence of an enhanced cryptic splice site caused by this substitution is indicated (bottom). The analysis was performed via Alamut using the prediction packages listed on the left, splice donor predictions are blue, splice acceptor predictions are green.

As the c.1080+1G>T mutation and the c.1224-80G>A variant are both predicted to result in a frameshift, and result in total loss of mRNA, it is not currently possible in our laboratory to confirm the RNA splice effect of the c.1224-80G>A variant. This is due to being unable to distinguish between total loss of the *P3H1* mRNA from any potential RNA splicing assay failure.

An option to explore this variant further would be to undertake collagen expression analysis in dermal fibroblasts. The identification of overmodified protein would support pathogenicity of the c.1224-80G>A variant. Fibroblasts could also be examined for signs of resulting ER stress. However, the patient has so far declined this analysis as they do not wish to undergo a skin biopsy.

This case clearly highlights the disadvantage of current targeted exome approaches to diagnosis of genetic disease. Generally only coding sequence is captured and important intronic findings may be missed.

vii.v Patient H

No candidate variants were identified in patient H who has a clinical diagnosis of osteogenesis imperfecta with crush fracture vertebrae, multiple long bone fractures and scoliosis.

viii RESULTS: DDD COMPLIMENTARY ANALYSIS

The Complementary Analysis Project (CAP12) established with the DDD project identified 49 individuals with clinical features relating to bone fragility (patient ID 101-149). Twenty seven of these are female and twenty two male.

A total of 939 candidate variants were identified in this cohort. This included three copy number variants: a 223Kb duplication in Chr 2, a 114kb deletion in chr X and a 658kb deletion in chr 1. The DDD project reported positive findings to the referring clinician in 29 individuals, with the number of variants ranging from one to four per individual. The majority of these reported variants are primarily associated with developmental disorders.

All variants were manually evaluated at the gene level focusing on reported function, as described in method section iii.x (page 61). After this initial assessment a total of 130 variants remained, with the number of variants for each individual ranging from none to nine (appendix R). These were evaluated at the variant level with reference to the clinical presentation and family structure. Each patient's complete VCF data was also interrogated as described in section iii.ix (page 57) to identify additional variants not highlighted by the DDD pipeline.

viii.i Osteogenesis Imperfecta Patients

Nine patients (7 females, 2 males) had a clinical diagnosis of Osteogenesis Imperfecta, one of which was noted as being atypical. The supplied clinical details of the nine OI patients, along with their results and family structure details are given in Table 23. No candidate variants have been identified in four of these individuals either by the DDD project or this study.

Patient ID	Sex	Family structure	Clinical Details	Results
101	F	Trio	Bowing of limbs due to multiple fractures, Fractured hand bones, Pathologic fracture	No candidate variant identified.
102	F	Trio	Abnormality of the radius, Blue sclerae, Multiple rib fractures, Pathologic fracture, Plagiocephaly, Thoracic kyphosis, Vertebral compression fractures	No candidate variant identified.
110	F	Singleton	Abnormality of the face, Asymmetry of spinal facet joints, Increased susceptibility to fractures, Joint hypermobility, Pes planus, Specific learning disability	<i>COL1A1</i> c.4006-1G>T
112	M	Trio	Decreased antibody level in blood, Elevated hepatic transaminases, Hypoglycemia, IgA deficiency, IgG deficiency, Increased susceptibility to fractures, Ligamentous	NBAS patient number 2 in our manuscript c.5741G>A;p.(Arg1914His)

			laxity, Nystagmus, Osteopenia, Progeroid facial appearance, Progressive cone degeneration, Recurrent fractures, Wide anterior fontanel	c.2032C>T;p.(Gln678*)
113	F	Singleton	Blue sclerae, Ligamentous laxity, Recurrent fractures, OI type V.	<i>IFITM5</i> c.-14C>T NOT reported.
136 (B)	M	Trio	Bowing of limbs due to multiple fractures, Congenital contracture, Decreased antibody level in blood, Multiple prenatal fractures	<i>SRCAP</i> <i>de novo</i> c.9029C>A;p.(Pro3010His)
137	F	Trio	Bowing of limbs due to multiple fractures, Congenital contracture, Decreased antibody level in blood, Multiple prenatal fractures	<i>COL1A2</i> c.280-2A>C NOT reported.
138	F	Trio	Disproportionate short stature, Hypoplasia of dental enamel, Limited elbow extension, Moderate generalized osteoporosis, Skeletal dysplasia	No candidate variant identified.
142	F	Trio	Bowing of limbs due to multiple fractures, Fractures of the long bones, Increased susceptibility to fractures, Osteopenia	No candidate variant identified.

Table 23 Clinical details, genetic results and family structure for DDD CAP 12 patients with a clinical diagnosis of Osteogenesis Imperfecta.

Results in black were reported to the referring clinician by the DDD project. Those in grey were not reported by the DDD project.

Of note is that, although DDD reported the pathogenic c.4066-1G>T *COL1A1* mutation in patient 110, they did not report the pathogenic c.280-2A>C *COL1A2* mutation in patient 137 even though it was listed in the supplied candidate gene list. This highlights a potential issue with currently available automated curation and reporting pipelines used in large cohort whole exome studies.

A second patient with a known pathogenic mutation (patient 113) was not reported by DDD. This patient had previously been screened at SDGS using the targeted exome NGS panel developed during the first part of this study. The c.-14C>T *IFITM5* autosomal dominant mutation known to be associated with OI type V had been identified and reported to the referring clinician.

The c.-14C>T mutation was not listed in the DDD candidate gene list for patient 113. Interrogation of the complete VCF identified that the c.-14C>T mutation had been identified by sequencing but was flagged as being of 'low quality' by the analysis pipeline and therefore excluded from further analysis. This highlights that regions of the genome where mutations are known to occur are not always reliably captured in exome sequencing and that 'false negatives' can result.

Interestingly an *SRCAP* c.3247C>T;p.Pro1083Ser variant was reported by DDD in this individual. This variant is reported in 12/28456 alleles in the South Asian population gnomAD dataset (rs776570251),

does not lie within a functional domain of the protein and *in-silico* conservation analysis supports this being a benign variant.

Patient 112 was reported to be compound heterozygous for a c.5741G>A;p.(Arg1914His) and a c.2032C>T;p.(Gln678*) in the *NBAS* gene and, along with patient C, is presented in the manuscript 'Compound heterozygous mutations in *NBAS* as a cause of atypical Osteogenesis Imperfecta' Bone. 2017 ;94:65-74 (Balasubramanian, Hurst et al. 2017) discussed previously (Section vi.iii, page 98).

The *de novo* c.9029C>A;p.(Pro3010His) *SRCAP* variant previously identified in patient B by exome sequencing was confirmed by the DDD project (patient 136) and reported back to the clinician. The c.1169C>T;p.(Pro390Leu) *TAPBP* variant also previously discussed (section vi.ii.iii page 97) was listed in the DDD candidate gene list but not reported to the clinician.

viii.ii Non Osteogenesis Imperfecta Patients

For patients without a stated clinical diagnosis of Osteogenesis Imperfecta a total of 19 candidate variants for bone fragility were identified in 12 patients, some of which are novel (Table 24)

ID	Clinical details	Gene	Variant	Inheritance
105	2-3 toe syndactyly, Abnormality of male external genitalia, Bilateral single transverse palmar creases, Cleft palate, Gastroesophageal reflux, Generalized hypotonia, Hemimegalencephaly, Hypermetropia, Lumbar hemivertebrae, Micrognathia, Mild short stature, Optic nerve hypoplasia, Osteopenia, Radioulnar synostosis, Ventriculomegaly, Vertebral compression fractures	<i>B4GALT7</i>	c.277dupC;p.(His93fs)	Paternal
		<i>B4GALT7</i>	c.641G>A;p.(Cys214Tyr)	Maternal
106	Abnormality of taste sensation, constipation, hyperhidrosis, hyperhidrosis, impaired pain sensation, lacrimation abnormality, muscle weakness, urticaria and painless fractures due to injury, has a <i>de novo</i> missense variant, , in <i>SCN9A</i>	<i>SCN9A</i>	c.4835T>G;p.(Leu1612Arg)	<i>De novo</i>
107	Abnormality of T cells, Bilateral cryptorchidism, Eczema, Glandular hypospadias, Hypoplasia of the corpus callosum, Intellectual	<i>ZEB2</i>	c.1212G>T;p.(Met404Ile).	Unknown
		<i>UBASH3B</i>	c.371Tdel;p.(Phe124fs)	Unknown

	disability, moderate, Low-set ears, Microcephaly, Micrognathia, Micropenis, Microtia, Osteopenia, Osteoporosis of vertebrae, Primary adrenal insufficiency, Severe short stature, Short foot, Stenosis of the external auditory canal	<i>SULF2</i>	c.2411A>G;p.(Asn804Ser)	Unknown
108	Ataxia, Blue sclerae, Delayed speech and language development, Global developmental delay, Horizontal nystagmus, Joint hypermobility, Poor coordination, Recurrent fractures	<i>COL1A1</i>	c.2644C>T;p.(Arg882*)	Maternal (recurrent fracture)
114	Cafe-au-lait spot, generalized dystonia, global developmental delay, osteopenia, reduced bone mineral density, seizures and status epilepticus.	<i>SCN8A</i>	c.5615G>A;p.(Arg1872Gly)	Unknown
117	Abnormal facial shape, Constipation, Global developmental delay, Increased susceptibility to fractures, Microcephaly, Polyhydramnios, Sleep disturbance, Unilateral deafness	<i>NBAS</i>	c.3776C>T;p.(Ser1259Phe)	Unknown
		<i>SULF2</i>	c.587T>C;p.(Ile196Thr)	Unknown
		<i>UGGT1</i>	c.371Tdel;p.(Phe124fs)	Unknown
127	Camptodactyly of finger, Dislocated wrist, Osteoporosis, Pes cavus, Scoliosis, Severe Myopia	<i>PLOD1</i>	c.153dup;p.(Asn52fs)	Paternal
128	Abnormality of blood circulation, aortic regurgitation, asthma, thin skin with reticulate pigmentation, generalized osteoporosis with pathologic fractures, genu valgum, glaucoma, migraine, myopia, otosclerosis, severe generalized osteoporosis, short foot, slender build, slender fingers and toes, and bronchiectasis	<i>SLC38A10</i>	c.1A>G;p.(Met1?)	Maternal (osteoporosis otosclerosis, abnormal kidney)
129	Dilation of lateral ventricles, generalized joint laxity, intellectual disability, moderate postnatal macrocephaly and recurrent fractures. His mother is hypermobile but no history of fractures is reported	<i>SCN9A</i>	c.5367_5368insGAGAACTC TAT;p.(Lys1790fs)	Maternal

140	Abnormality of the odontoid process, amelogenesis imperfecta, cerebral calcification, depressed nasal bridge, downslanted palpebral fissures, elevated circulating parathyroid hormone (PTH) level, hypophosphatemia, malar flattening, narrow mouth, osteoporosis, sagittal craniosynostosis, subglottic stenosis	<i>FAM20C</i>	c.1225C>T;p.(Arg409Cys)	Paternal+Maternal
		<i>LRP6</i>	c.4481dup;p.(Arg1495fs)	<i>De--novo</i>
141	Abnormality of mast cells, Autonomic dysregulation, Joint hypermobility, Orthostatic hypotension, Osteopenia, Seizures, Joint dislocations in young adult	<i>POSTN</i>	c.1108+2T>C;p.(?)	Maternal
		<i>BMP6</i>	c.334G>A;p.(Glu112Lys)	Maternal
149	2-3 toe syndactyly, Abnormality of dental enamel, Blue sclerae, Bruising susceptibility, Constipation, Delayed gross motor development, Delayed speech and language development, Downslanted palpebral fissures, Increased muscle fatiguability, Intellectual disability, severe, Joint hypermobility, Malar flattening, Recurrent fractures	<i>COL1A1</i>	c.1299+1G>A;p.(?)	<i>De novo</i>

Table 24 Candidate variants in CAP12 DDD patients following manual assessment of DDD supplied candidate list and interrogation of full VCF for patients 101-149.

All variants were identified in the heterozygous state with the exception of the FAM20C c.1225C>T;p.Arg409Cys variant which is homozygous in the proband. The inheritance of variants is stated as unknown if parental samples were unavailable.

viii.iii Type I Collagen Genes

Two patients without a stated clinical diagnosis of OI were found to have a pathogenic mutation in the type I collagen genes (Table 24)

Patient 108 *COL1A1* c.2644C>T;p.(Arg882*)

Patient 108 has a maternally inherited *COL1A1* c.2644C>T;p.(Arg882*) mutation that would be expected to result in a mild/moderate OI. The patient's mother is reported to have recurrent fractures which supports that this mutation is segregating with the symptoms of bone fragility in this family. This pathogenic mutation has been reported on nine occasions in the locus specific data base (<https://oi.gene.le.ac.uk>) in association with OI ranging in clinical severity of type I to type IV, which is consistent with the reported phenotype in the patient and her mother.

Patient 149 COL1A1 c.1299+1G>A

The second patient (149) has a *de-novo* COL1A1 c.1299+1G>A pathogenic mutation that has been reported 11 times in association with OI type 1 (<https://oi.gene.le.ac.uk>) and is also recorded on ClinVar and HGMD as pathogenic. This mutation is predicted to result in the complete loss of the wildtype donor site of intron 19 and may be expected to result in exon skipping and hence a moderate/severe phenotype. However, splice site analysis reveals that there is a cryptic splice site within exon 19 that is likely to be used as an alternate splice site, resulting in the loss of 8bp and a frameshift (Figure 72). Frameshift mutations usually result in a 'null allele' and are associated with a milder clinical phenotype; this is congruent with the phenotype of patient 149 and that of the previously reported individuals with type I OI.

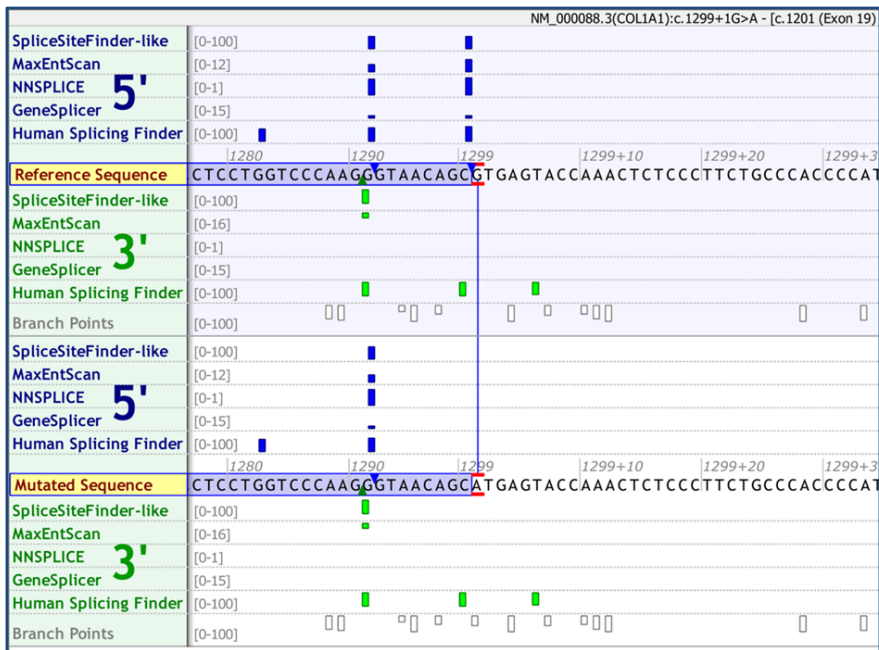


Figure 72 Predicted effect on splicing of the c.1229+1G>A COL1A1 pathogenic mutation identified in patient 149.

The prediction indicates the completed loss of wildtype donor site in intron 19 and shows the location of the cryptic donor site located upstream in exon 19 that is likely to be used as an alternative splice site.

Top = wildtype splice prediction Bottom = mutant splice prediction

viii.iv Ehlers Danlos Syndrome Genes

Two patients were identified with a potential diagnosis of Ehlers Danlos syndrome.

Patient 127 PLOD1 c.153dup;p.(Asn52fs)

No likely candidate genes were identified in the DDD variant list for patient 127, however interrogation of the entire VCF identified a paternally inherited heterozygous c.153dup;p.(Asn52fs) variant in the

procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1 gene (*PLOD1*). This gene is known to cause kyphoscoliotic Ehlers Danlos Syndrome (EDS).

Kyphoscoliotic EDS (previously known as EDS VI) is an autosomal recessive connective tissue disorder characterised by hyperextensible skin, atrophic scarring and generalised joint laxity. Scoliosis can be present at birth and is progressive. Muscle hypotonia and delay in attainment of gross motor milestones are also common features. Scleral fragility, rupture of the globe, marfanoid habitus and rupture of medium size arteries may also be present.

The majority of pathogenic *PLOD1* mutations are either small deletions/insertions that result in a frameshift or are substitutions that result in a nonsense mutations. However, a common larger duplication of 8.9kb, caused by homologous recombination of Alu sequences in introns 9 and 16, is reported to account for ~18% of reported mutations (Yeowell, Walker et al. 2005). The resulting deletion of exons 10 to 16 is not readily detectable by the exome sequencing performed in this project. However, analysis using Multiplex Ligation-dependent Probe Amplification (MLPA) is available as a diagnostic test at SDGD for this mutation (Figure 73).

The c.153dup *PLOD1* mutation has previously been reported (as c.153_154insC, using previous nomenclature) in an individual with kyphoscoliotic EDS (Heikkinen, Pousi et al. 1999) who is compound heterozygous for this mutation and a c.467-2delA mutation. Functional analysis confirmed loss of the lysyl hydroxylase protein in the patient's fibroblasts. No clinical details are provided for the affected individual in this case report.

The *PLOD1* protein forms hydroxylysine residues at collagen Xaa-Lys-Gly amino acid sequences (Yeowell, Allen et al. 2000, Walker, Overstreet et al. 2005) which are essential for stable collagen crosslinking. Deficiency of hydroxylysine-based pyridinoline cross-links in collagens can be detected in urine by an increased ratio of deoxypyridinoline to pyridinoline cross-links and is used as a diagnostic test in patients with suspected kyphoscoliotic EDS.

The clinical symptoms of patient 127 are reported as camptodactyly of the finger, dislocated wrist, osteoporosis, pes cavus, scoliosis and severe myopia. A diagnosis of Camptodactyly-arthropathy-coxa pericarditis syndrome (CACP) or Marfan syndrome was initially considered. Testing for mutations in the *FBN1* gene for Marfan syndrome was negative. Although the patient is described as having osteoporosis, that by ISCD definition requires one or more vertebral compression fractures or a clinically significant fracture history and BMD Z score of -2 or less. (<https://www.iscd.org/official-positions/2013-iscd-official-positions-pediatric/>), no details of any fracture history are provided

Interestingly camptodactyly, although not listed in the diagnostic criteria for kyphoscoliotic EDS, has been reported in individuals with mutations in *FKBP14* and *CHST14*, genes now identified in additional forms of kyphoscoliotic EDS (Malfait, Syx et al. 2010, Giunta, Baumann et al. 2018). Patients with EDS have been shown to have reduced BMD and an increased susceptibility to both vertebral and non-vertebral fractures (Unnanuntana, Rebolledo et al. 2011, Eller-Vainicher, Bassotti et al. 2016, Mazziotti, Dordoni et al. 2016).

The clinical features of our patient, along with the *PLOD1* mutation, suggest that a diagnosis of kyphoscoliotic EDS should be considered. The patient's clinician has been contacted and analysis of

a urine sample, to assess deoxyypyridinoline to pyridinoline ratio, has been suggested as an initial investigation, followed by MLPA analysis if the ratio is increased.

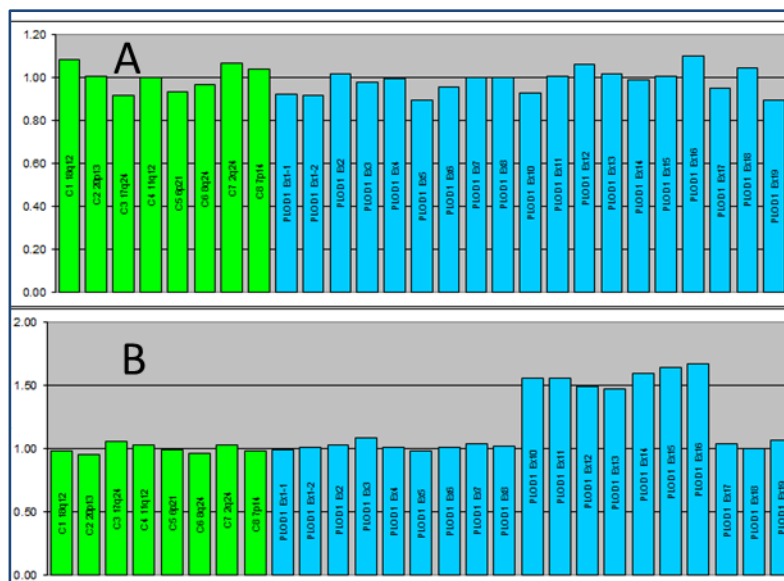


Figure 73 Multiplex Ligation-dependent Probe Amplification analysis of genomic DNA for the *PLOD1* gene.

A: Normal control. B: Heterozygous duplication of exons 10-16

Patient 105 *B4GALT7* c.277dupC;p.(His93fs) and c.641G>A;p.(Cys214Tyr)

A second patient (105) with pathogenic mutations in an EDS-associated gene was also identified. This individual is reported as having bilateral single transverse palmar creases, cleft palate, gastroesophageal reflux, generalized hypotonia, lumbar hemivertebrae, pes planus, micrognathia, mild short stature, optic nerve hypoplasia, osteopenia, radioulnar synostosis, ventriculomegaly, bowing of the long bones and vertebral compression fractures.

He is compound heterozygous for a paternally inherited c.277dupC;p.(His93fs) frameshift and a maternally inherited c.641G>A;p.(Cys214Tyr) missense variant in the *B4GALT7* gene. This gene was originally associated with 'progeroid' EDS (PEDS) but this has recently been reclassified as spondylodysplastic EDS (Brady, Demirdas et al. 2017).

The diagnostic criteria for spondylodysplastic EDS (spEDS) include short stature, skin hyperextensibility, facial dysmorphism, muscle hypotonia and joint laxity. Additional features reported include neurosensory hearing loss, radioulnar synostosis, hypermetropia, and varying degrees of developmental delay (Salter, Davies et al. 2016, Ritelli, Dordoni et al. 2017). Low bone mineral density has also been reported (Ritelli, Dordoni et al. 2017).

B4GALT7 encodes β 1,4-galactosyltransferase 7, a golgi resident enzyme involved in the synthesis of the glycosaminoglycan-protein (GAG) linkage region in proteoglycans. The enzyme catalyzes the transfer of the first galactose onto the xylose residue of the proteoglycan GAG linker region. Addition

of either N-acetylglucosamine or N-galactosylglucosamine to the linker region leads to the formation of heparan sulfate or chondroitin/dermatan sulfate respectively (Figure 74). Mutations in other linker protein genes (*XYLT1*, *XYLT2*, *B3GALT6*, and *B3GAT3*) are associated with skeletal diseases with overlapping clinical presentation. In particular *XYLT2* mutations are characterized severe childhood-onset primary osteoporosis (Taylan and Mäkitie 2016)

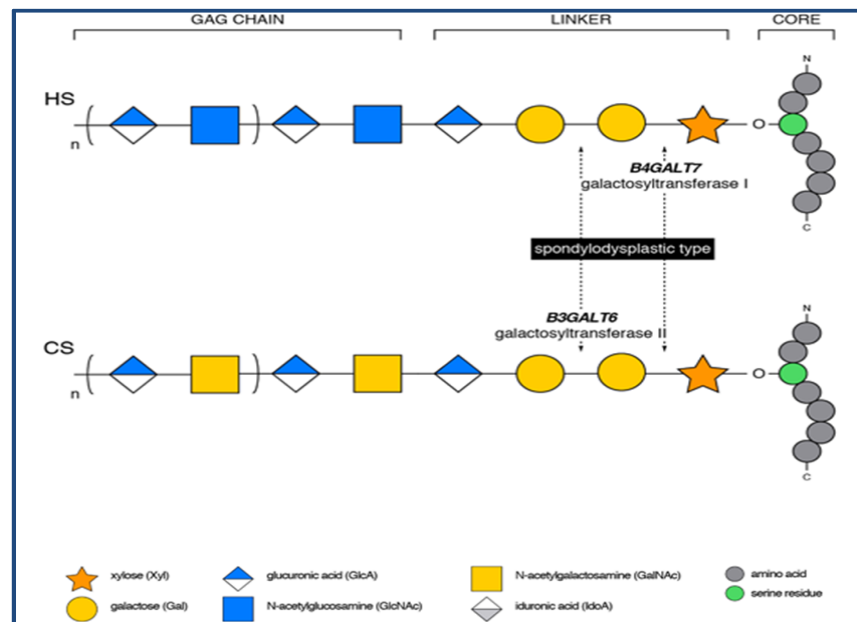


Figure 74 The synthesis of heparin sulphate and chondroitin/dermatan sulfate proteoglycans.

Sequential action of specific enzymes lead to synthesis of the linker region: Xylosyltransferase I/II (encoded by *XYLT1* and *XYLT2*), galactosyltransferase I (encoded by *B4GALT7*) and galactosyltransferase II (encoded by *B3GALT6*), and glucuronyltransferase I (encoded by *B3GAT3*). Following completion of the linker region, the addition of the following residue, either N-acetylglucosamine or N-galactosylglucosamine, determines which proteoglycan is synthesized. Figure modified from Brady et al (Brady, Demirdas et al. 2017)

Only a limited number of *B4GALT7* mutations have been reported to date, seven missense and two frameshift, with c.808C>T;p.(Arg270Cys) being the most common. The c.641G>A;p.(Cys214Tyr) variant in patient 105 affects a highly conserved amino acid within the galactosyltransferase functional domain of the protein and is predicted to be deleterious to function. The c.277dupC;p.(His93fs) is predicted to result in a premature termination codon 72 positions downstream and the mRNA is a likely target for nonsense mediated decay. Both the c.641G>A;p.(Cys214Tyr) and c.277dupC are reported at low frequency in the general population, 8/111480 European (non Finish) and 4/4260 African American alleles respectively.

The clinical features of this patient were considered to be consistent with biallelic mutations in the *B4GALT7* gene and he is presented by the referring centre in the article by Slater et al (Slater, Davies et al. 2016).

viii.v Sodium Ion Channel Genes

Defects in two voltage-gated sodium ion channels ($Na_v1.7$ and $Na_v1.9$) encoded by the *SCN9A* and *SCN11A* genes respectively, have been described in association with congenital insensitivity to pain

(CIP) and episodic/extreme pain syndromes (Bennett and Woods 2014). Frequent fractures are a known feature of insensitivity to pain disorders and lack of pain perception and injury awareness have been considered to be the reason for this.

However, although skeletal problems in CIP are well documented, there is little reported about mineral and skeletal homeostasis in affected individuals. Studies in one CIP family with an autosomal dominant gain of function mutation in *SCN11A* report normal biochemical parameters, no radiographic evidence of rickets or osteopenia and DXA lumbar spine and total hip BMD within the normal range. Bone histology was also unremarkable. However, it was noted that the fractures were often as a result of apparent low-trauma and the authors comment on the possibility of 'poor quality bone' being a contributing factor (Phatarakijirund, Mumm et al. 2016).

One recent report presents exome analysis in a consanguineous Egyptian family with a clinical presentation of "OI type IV", where a homozygous nonsense mutation c.570G>A;p.(Trp190*) in *SCN9A* in the proband and her affected aunt has been identified (Caparros-Martin, Aglan et al. 2017). The proband, who initially presented at 3 months with three fractures, is described as having greyish sclerae and with behaviours retrospectively assessed as being consistent with insensitivity to pain. The 14 year old aunt, also with insensitivity to pain, had four fractures during the first year of life with subsequent deformities and started to develop bowing of the long bone at age of 6-7 months old without fractures, suggestive of underlying bone pathology. Unfortunately, no bone parameters of the family are provided in the manuscript.

A further voltage-gated sodium channel, Na_v1.6 encoded by *SCN8A*, is associated with autosomal dominant early-infantile epileptic encephalopathy. Recent studies of bone mass and architecture in a child with a loss of function *SCN8A*-related epileptic encephalopathy using dual energy X-ray absorptiometry (DXA) and high-resolution peripheral quantitative computed tomography (HRpQCT) revealed altered trabecular and cortical bone parameters showing low bone mass, as well as an elevated serum bone resorption markers (Rolvien, Butscheidt et al. 2017). Studies in the *Scn8a*-deficient mouse model identified elevated bone resorption due to increased osteoclastogenesis as the underlying mechanism behind these altered bone parameters. A marked improvement in bone structure and no additional skeletal fractures following anti-resorptive therapy (Neridronate) is reported in an individual with *SCN8A*-related epileptic encephalopathy and severe bone loss with multiple fractures, supporting the remodelling pathology of bone fragility in this individual (Rolvien, Butscheidt et al. 2017).

Three patients with heterozygous variants in sodium ion channel genes were identified in the DDD data dataset:

Patient 129 *SCN9A* c.5367_5368insGAGAACTCTAT;p.(Lys1790fs)

Patient 129 has a maternally inherited c.5367_5368insGAGAACTCTAT;p.Lys1790fs variant in *SCN9A*. He is reported to have dilation of lateral ventricles, generalized joint laxity, intellectual disability, moderate postnatal macrocephaly and recurrent fractures. His mother is hypermobile but no history of fractures is reported. No second likely deleterious variant in this gene was identified.

Patient 106 *SCN9A* c.4835T>G;p.(Leu1612Arg)

A second male, patient 106, with a clinical phenotype of abnormality of taste sensation, constipation, hyperhidrosis, impaired pain sensation, lacrimation abnormality, muscle weakness, urticaria and painless fractures due to injury, has a *de-novo* missense variant, c.4835T>G;p.(Leu1612Arg), in *SCN9A*. Another missense mutation at this codon, p.(Leu1612Pro), has previously been associated with autosomal recessive paroxysmal extreme pain disorder.

Interestingly *SCN9A*, as well as being associated with autosomal recessive CIP, is also reported in a spectrum of autosomal dominant seizure disorders ranging from isolated febrile seizures to more generalized epilepsy.

Patient 114 *SCN8A* c.5615G>A;p.(Arg1872Gln)

Patient 114, a singleton, is reported as having Cafe-au-lait spot, generalized dystonia, global developmental delay, osteopenia, reduced bone mineral density, seizures and status epilepticus. He is heterozygous for a known pathogenic mutation in *SCN8A*, c.5615G>A;p.(Arg1872Gln). This mutation has previously been reported in an individual with repetitive language, macrocephaly, generalized hyper reflexia, clumsiness and autistic features (Larsen, Carvill et al. 2015). Other amino acid substitutions at this position, p.(Arg1872Leu) and p.(Arg1872Trp), are reported as pathogenic in early infantile epileptic encephalopathy type 13 (<https://portal.biobase-international.com/hgmd/pro/all.php>). It seems reasonable, given the previously described low bone mass in a child with a *SCN8A* loss of function mutation (Rolvien, Butscheidt et al. 2017), to consider the c.5615G>A;p.(Arg1872Gln) as a likely cause of the clinical symptoms and bone features in patient 114.

Taken together these findings suggest a potential pathway between pain sensation, epilepsy/seizures and bone parameters involving voltage-gated sodium channels, particularly *SCN9A*. However, a note of caution is that anti-convalescent medication and reduced mobility, which may be associated with epilepsy/seizures and developmental delay respectively, can both contribute to osteopenia.

viii.vi SLC38A10

Patient 128 *SLC38A10* c.1A>G;p.(Met1?)

Manual evaluation of data for patient 128, at both the gene and variant level, identified a heterozygous c.1A>G;p.(Met1?) variant of interest in the *SLC38A10* gene. This gene is located at 17q25.3 and is alternatively known as *SNAT11* or *PP1744*. The UniProt Go biological processes given for this gene include bone development. Frequency data for this variant in the general population is limited but it is listed in the genomAD dataset in 13 of 222,830 alleles with a maximum allele frequency in any population of 0.01%. No homozygous individuals are reported.

The Solute carrier (SLC) superfamily consists of 395 transporters divided into 52 different families with differing biochemical properties (Hediger, Cl  men  on et al. 2013). The SLC38 family has 11 members (A1-A11) that preferentially favour glutamine as substrate, with the exception of SLC38A4, which mainly transports alanine, asparagine, and cysteine (Sugawara, Nakanishi et al. 2000). SLC38

transporters are ubiquitously expressed within all cell types of the body and, in addition to their transport role, are thought to play an important role in amino acid signalling within many tissues including the liver, brain, kidney and intestine (Broer 2014).

Recently, the structure of the SLC38A10 protein, a previously orphan member of the group, has been modelled (Figure 75) predicting eleven transmembrane domains, an intracellular N-terminal domain and a 722 amino acid C-terminal domain that sits outside the membrane (Hellsten, Hägglund et al. 2017).

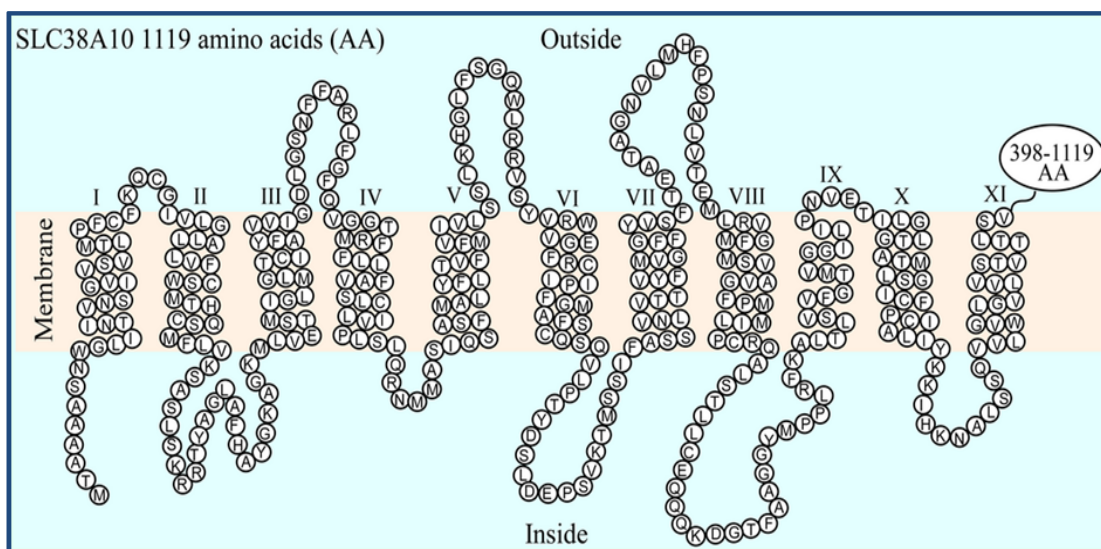


Figure 75 The predicted structure of the SLC38A10 protein.

Each amino acid is represented as a circle. The C-terminal domain (amino acids 398-1119) sit outside the membrane and are not depicted (Modified from Hellsten et al (Hellsten, Hägglund et al. 2017)).

Data supporting that the bidirectional transport of glutamine, glutamate, and aspartate by SLC38A10 is a key function, implicating the protein in neurotransmission, has also been published (Hellsten, Hägglund et al. 2017).

Glutamine is the most abundant amino acid and, together with glucose, is important in many metabolic and energy production pathways (Pochini, Scalise et al. 2014). It has emerged as an important regulator of osteoblasts, being actively metabolised and essential for matrix mineralization (Brown, Hutchison et al. 2011). Glutamine is an anaplerotic precursor (replenishing catalytic intermediates) that supplies part of the energy requirement of differentiating osteoblasts during bone formation (Karner, Esen et al. 2015). Glutamine anaplerosis into the citric acid cycle (TCA) is increased by raised glutaminase (Gls) activity in response to Wnt signalling. The resulting reduction in intracellular glutamine leads to activation of Gcn2 and Atf4 stimulating uptake and synthesis of amino acids, hence promoting protein synthesis from the target gene (Figure 76).

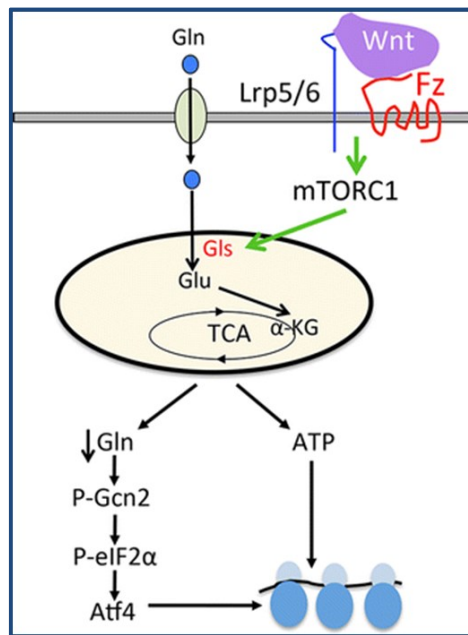


Figure 76 Glutamine signalling in osteoblasts in response to Wnt signalling

Wnt signalling stimulates glutamine metabolism through the citric acid cycle (TCA). Wnt signaling through Frizzled (Fz) and Lrp5/6 activates mTORC1; mTORC1 increases glutaminase (Gls) levels that stimulate glutamine oxidation for ATP production. Increased levels of glutamine consumption results in reduced intracellular glutamine, activates the Gcn2-Atf4 stress pathway and up-regulates gene transcription. Gln glutamine, Glu glutamate, α -KG α -ketoglutarate. Figure modified from Karner et al (Karner and Long 2017).

Slc38a10 knock out mice have a severe skeletal phenotype that includes weak and brittle bones, reduced body length and decreased bone mineral density, further supporting an important role for this gene in bone metabolism (Bassett, Gogakos et al. 2012).

The clinical features of patient 128 are given as: abnormality of blood circulation, aortic regurgitation, asthma, thin skin with reticulate pigmentation, generalized osteoporosis with pathologic fractures, genu valgum, glaucoma, migraine, myopia, otosclerosis, severe generalized osteoporosis, short foot, slender build, slender fingers and toes, and bronchiectasis.

The c.1A>G;p.(Met1?) *SLC38A10* variant in patient 128 is maternally inherited. The mother's clinical features are given as osteoporosis, otosclerosis and abnormal kidney. This supports a potential role for this variant in the bone phenotype in this family but is not conclusive. The patient and their family have been invited for further clinical review.

viii.vii Complex Genotypes

Patient 140 *FAM20C* c.1225C>T;p.Arg409Cys; *LRP6* c.4481dup;p.Arg1495fs

Two missense variants of interest have been identified in patient 140: a homozygous c.1225C>T;p.Arg409Cys in *FAM20C* and a *de-novo* heterozygous c.4481dup;p.Arg1495fs in *LRP6*. Only the *FAM20C* variant was reported back to the referring clinician by the DDD project.

Fam20C is a golgi serine/threonine protein kinase that phosphorylates Ser-X-Glu (S-X-E) motifs of secretory proteins in the extracellular matrix of bones and teeth, including dentin matrix protein 1 (DMP1) and osteopontin (OPN), and is essential for mineralization (Cui, Xiao et al. 2015, Tagliabracci, Wiley et al. 2015). Autosomal recessive mutation in *FAM20C* have been described in the osteosclerotic bone dysplasia Raine syndrome which has varying clinical presentation ranging from an aggressive onset and death during the first few hours or weeks of life (Simpson, Hsu et al. 2007) to survival into childhood or middle adulthood (Simpson, Scheuerle et al. 2009, Fradin, Stoetzel et al. 2011, Acevedo, Poulter et al. 2015).

Raine syndrome is characterised by generalized osteosclerosis, periosteal bone formation and craniofacial dysmorphia that includes midface hypoplasia, depressed nasal bridge and prominent forehead. Intracranial ossification is a common feature. Abnormal teeth (small with enamel dysplasia) are reported in surviving patients. Hypophosphataemia, caused by elevated levels of FGF23, is reported in patients with non-lethal Raine syndrome (Simpson, Scheuerle et al. 2009, Fradin, Stoetzel et al. 2011, Rafaelsen, Raeder et al. 2013).

The c.1225C>T;p.(Arg409Cys) *FAM20C* variant in patient 140 has been reported twice in the literature in the homozygous state, the first as part of a gene discovery study in neurogenetic disorders using whole-exome sequencing of consanguineous families (Alazami, Patel et al. 2015) and the second in a Raine syndrome case report (Seidahmed, Alazami et al. 2015). As both literature reports originate from the same research group, it is likely that this represents a single individual.

The skeletal features of the case report a patient, a male who died shortly after birth, include generalised increased density of all bone, particularly in the craniofacial bones. Failure of ossification is noted in the sacrum and lumbar vertebrae. The degree of osteosclerosis is noted to be relatively mild compared with previously reported Raine syndrome patients. The patient's chest is described as asymmetrical and bell shaped with thin ribs. (Seidahmed, Alazami et al. 2015).

The clinical features of patient 140 are given as abnormality of the odontoid process, amelogenesis imperfecta, cerebral calcification, depressed nasal bridge, downslanted palpebral fissures, elevated circulating parathyroid hormone (PTH) level, hypophosphatemia, malar flattening, narrow mouth. osteoporosis, sagittal craniosynostosis, subglottic stenosis. The majority of these are consistent with Raine Syndrome and were considered to offer a full explanation of phenotype by the referring clinician.

However, the *LRP6* *de-novo* c.4481dup;p.(Arg1495fs) variant was not reported to the referring clinician but may also be contributing to the phenotype in this patient. *LRP6* is a key transmembrane co-receptor in the canonical Wnt/ β -catenin signalling pathway, an important regulator of skeletal development and maintenance. Canonical Wnt signalling is triggered by the binding of a given Wnt ligand to a frizzled family (Fzd) seven-pass transmembrane receptor and a co-receptor of the arrow/Lrp family (such as LRP5 and LRP6). This results in the stabilisation and translocation of β -catenin, an important activator of gene transcription (Figure 77).

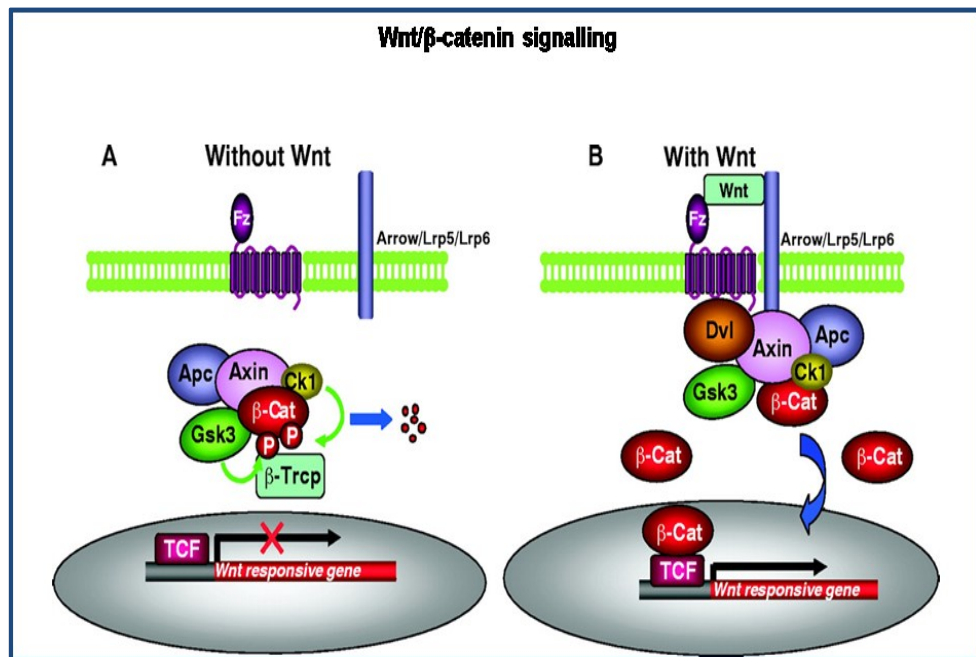


Figure 77 Simplified Wnt/β-catenin signaling.

(A) Without Wnt, the scaffolding protein Axin assembles a protein complex that contains Apc, Gsk3, Ck1 and β-catenin. In this complex, β-catenin is sequentially phosphorylated by Ck1 and Gsk3. Phosphorylated β-catenin is recognized by β-Trcp, which is a component of an ubiquitin-ligase complex that conjugates β-catenin with ubiquitin. Poly-ubiquitinated β-catenin is degraded by the proteasome. (B) In the presence of Wnt, β-catenin phosphorylation and degradation is inhibited. Accumulated β-catenin forms a nuclear complex with the DNA-bound TCF/LEF transcription factor, and together they activate Wnt-responsive genes. (Figure and legend modified from Xi He et al. (He, Semenov et al. 2004).

LRP6 has also been shown to be important in parathyroid hormone (PTH) signalling, which regulates osteoblast activity. PTH binds to its receptor, PTH1R, inducing the PTH-PTH1R-LRP6 protein complex, leading to increased bone formation in rat (Wan, Yang et al. 2008).

LRP5 and LRP6 have similar structures comprising of four YMTD β-propeller domains followed by EGF-like domains that bind the extracellular ligands, and five intracellular PPP(S/T)P domains that mediate downstream signalling (Williams and Insogna 2009). Studies in LRP5 and LRP6 deficient mouse models have shown that both LRP5 and LRP6 are required in osteoblasts for normal skeletal development and homeostasis (Riddle, Diegel et al. 2013).

As previously described (section ii.xv.i.i page 35) mutations in *LRP5* give rise to several conditions and can be inherited in either a dominant or recessive manner. Loss of function mutations can result in familial exudative vitreoretinopathy (FEVR), Osteoporosis Pseudoglioma (OPPG) and Idiopathic Juvenile Osteoporosis (IJO) whereas gain of function mutations are associated with high bone mass conditions such as autosomal dominant osteopetrosis type 1 and endosteal hyperostosis.

To date, heterozygous loss of function mutation in *LRP6* have been described in families with autosomal dominant oligodontia, a severe form of tooth agenesis, and combined tooth agenesis/orofacial clefting (Massink, Créton et al. 2015, Ockeloen, Khandelwal et al. 2016). Discussion with the authors (Massink, Créton et al. 2015) confirmed that although none of the

reported patients had generalized amelogenesis imperfecta, there were some teeth showing enamel defects. Moreover at least one patient showed clear ankylosis of deciduous teeth (without successors), and there were also patients who had tooth-shape anomalies. Skeletal assessment was not performed on their patients.

Very low bone mineral density (Z score of -3.4 at the femoral neck), with low-impact fracture, has been reported in an individual from a large family with early coronary disease (Mani, Radhakrishnan et al. 2007). Affected individuals were found to be homozygous for a putative partial loss of function mutation, p.(Arg611Cys) in *LRP6*. This mutation lies within the EGF-like domain of the LRP6 protein and functional studies confirmed reduced LRP6 signalling in NIH3T3 mutated cells. Two further affected members of this family also had early hip fracture and low BMD.

Studies in a mouse model, (ringelschwanz), who carry a homozygous p.(Arg886Trp) missense mutation in *Lrp6* also exhibit delayed ossification at birth and have a low bone mass phenotype in adults (Kokubu, Heinzmann et al. 2004).

The amelogenesis imperfecta and osteoporosis described in patient 140 is likely to be at least partly caused by his *de novo* loss of function *LRP6* variant, either in isolation or in combination with his *FAM20C* mutation. It would be interesting to explore this further as it seems reasonable to hypothesise that further skeletal consequences of mutations in *LRP6* remain undiscovered at present.

Patient 107 *ZEB2* c.1212G>T;p.(Met404Ile); *UBASH3B* c.371Tdel;p.(Phe124fs) and *SULF2* c.2411A>G;p.(Asn804Ser)

Patient 107 is described as having abnormality of T cells, bilateral cryptorchidism, eczema, glandular hypospadias, hypoplasia of the corpus callosum, intellectual disability, low-set ears, macrocephaly, micrognathia, micropenis, microtia, osteopenia, osteoporosis of vertebrae, primary adrenal insufficiency, severe short stature, short foot and stenosis of the external auditory canal.

A heterozygous c.1212G>T;p.(Met404Ile) missense variant in the Zinc finger E-box-binding homeobox 2 (*ZEB2*) gene (also known as *SMADIP1*, *ZFHX1B* and *SIP1*) was reported to the referring clinician by the DDD project. This gene is associated with autosomal dominant Mowat-Wilson syndrome (MOWS; OMIM 235730), a complex developmental delay disorder with moderate to severe intellectual disability, a distinctive facial gestalt, microcephaly, epilepsy, heart defects, urogenital malformations and Hirschsprung disease (HSCR).

Short stature is also a feature of this disorder but osteopenia, osteoporosis of the vertebrae and T cell abnormalities, as present in our patient, are not described. Also of note is that 70-75% of patients with MOWS have seizures, a further clinical feature that is not reported in patient 107. Interestingly a recent report of a cohort of 12 individuals suggests that the majority of MOWS patients also have features that resemble classical Ehlers Danlos syndrome, namely hyperextensible skin, joint hypermobility and atrophic scarring (Teraishi, Takaishi et al. 2017), features that are not described in patient 107.

The exact mechanism of action by which mutations in the *ZEB2* gene cause the range of clinical symptoms in MOWS patients is still unclear, although it is known that the *ZEB2* protein plays a vital role in neural crest cell migration during foetal development, most likely through TGF β signalling and interaction with nucleosome remodelling and histone deacetylation complex (NuRD).

Zeb2 knockout mice have changes in gene expression levels in dermal fibroblasts including many genes that are associated with connective tissue disorders, most noticeably the down regulation of collagenogenesis genes and up regulation of genes involved in collagenolysis. Electron microscopy of skin biopsy from MOWS patients has also shown reduced collagen fibre diameter supporting a role of *ZEB2* in dermal fibrillogenesis (Teraishi, Takaishi et al. 2017).

However, the majority of reported mutations in *ZEB2* (HGMD) are either truncating or larger gene deletions or rearrangements that result in haploinsufficiency. Missense mutations are rare (only 3 reported) and there are limited data available for the associated phenotype in these individuals. Severity of symptoms is described as mild (Ghoulid, Drevillon et al. 2013).

The c.1212G>T;p.(Met404Ile) has not been reported in association with disease or in the normal population. It affects a highly conserved nucleotide and amino acid, up to tetraodon (considering 12 species), and *in-silico* conservation prediction supports a deleterious effect. Another variant at this position, c.1211T>C;p.(Met404Thr), is reported in 1 in 246,072 normal alleles (gnomAD). It is therefore unclear whether the c.1212G>T;p.(Met404Ile) missense variant is pathogenic and could therefore be responsible for the clinical symptoms in patient 107.

In addition to the *ZEB2* variant above, patient 107 is also heterozygous for a c.371Tdel;p.(Phe124fs) frameshift variant in the *UBASH3B* gene and a c.2411A>G;p.(Asn804Ser) missense variant in the *SULF2* gene, neither of which were reported by DDD to the clinician.

UBASH3B (Ubiquitin Associated And SH3 Domain Containing B), also known as *KIAA1959*, *STS-1* or *TULA2*, is not associated with any known disorder in the OMIM database. The GO biological process for the protein includes participation in the collagen-activated tyrosine kinase receptor signalling pathway, negative regulation of bone resorption and negative osteoclast differentiation. The gene is highly expressed in bone, the immune system, reproductive organs and platelets.

In particular, *UBASH3B* is highly expressed in osteoclasts and is thought to be required for normal bone resorption. Knock-out mouse models have shown that osteoclast precursor numbers are increased by reduction in *UBASH3B* protein and exhibit an enhanced ability to resorb bone. Knock out mice also have an associated decrease in BMD (Back, Adapala et al. 2013). *UBASH3B* is also known to be a critical negative regulator of TCR signalling, regulating T-cell activation (Mikhailik, Ford et al. 2007, Newman, Liverani et al. 2014) and to influence receptor signalling in platelets (Reppschläger, Gosselin et al. 2016). Many of these functions would fit with the clinical presentation of our patient.

UBASH3B is described as being a loss of function intolerant gene with a pLI=0.99 in gnomAD. The c.371Tdel;p.(Phe124fs) variant is predicted to result in a premature termination codon 53 positions downstream and to result in nonsense-mediated mRNA decay. It has not been reported previously either in association with disease or in the general population.

However, this gene is located at 11q24.1, a region of chromosome 11 that is deleted in 8 patients in the Decipher database. Review of the supplied clinical features in these 8 patients did not identify any patient with either osteoporosis, osteopenia or increased susceptibility to fracture. It would be interesting to further explore the phenotype of these DDD patients and perhaps undertake platelet function analysis in patient 107 to further investigate the potential role of the *UBASH3B* gene. Another option would be to see how the development of osteoclasts from the patient's circulating mononuclear cells compares to that of normal controls (Fujikawa, Quinn et al. 1996).

The *SULF2* gene encodes a heparin sulphate 6-O-endosulfatase that regulates the release of growth factors from extracellular storage sites, activating multiple signalling pathways during growth, differentiation and repair of many tissues. It is vital for pathways that require heparin sulphate proteoglycan as a co-factor for ligand-receptor binding, including the TGF β 1/SMAD (Chen, Nakamura et al. 2017) and WNT-GLI1-CYCLIN D1 pathways (Nakamura, Fernandez-Barrena et al. 2013).

Of particular interest is that high levels of both *SULF1* and *SULF2* are expressed in osteoblasts actively forming bone (Zaman, Staines et al. 2016). Data suggesting a role of SULFs in Hedgehog signalling and within the extracellular matrix during fracture healing has also been presented, with the SULFs being proposed as a possible target for future treatment development to aid fracture healing (Zaman, Staines et al. 2016).

The c.2411A>G;p.(Asn804Ser) *SULF2* variant has been reported in only 1 in 240358 alleles in gnomAD. It lies within the highly conserved extracellular sulfatase domain of the protein and affects a highly conserved nucleotide and amino acid. However, *in-silico* conservation analysis is inconclusive and therefore this variant is of uncertain significance. Interestingly patient 117 also has a heterozygous missense variant in *SULF2*, a c.587T>C;p.(Ile196Thr) within the extracellular sulfatase domain.

Sequence data from both parents of this patient may clarify the inheritance pattern of these three variants and help to direct further investigation of their potential role in this patient's clinical symptoms.

Patient 117 *NBAS* c.3776C>T;p.(Ser1259Phe), *SULF2* c.587T>C;p.(Ile196Thr) and *UGGT1* c.371Tdel;p.(Phe124fs).

The clinical features of patient 117 are described as abnormal facial shape, constipation, global developmental delay, increased susceptibility to fractures, microcephaly, polyhydramnios, sleep disturbance and unilateral deafness. His developmental delay features are explained by a reported heterozygous CTCF (a chromatin binding factor gene) c.612_615delAAAG;p.(Lys206fs) mutation, a gene known to be associated with mental retardation, autosomal dominant 21. However, his symptoms of bone fragility are not explained by this finding.

Three heterozygous variants warranting further exploration were identified in patient 117: *NBAS* c.3776C>T;p.(Ser1259Phe), *SULF2* c.587T>C;p.(Ile196Thr) and *UGGT1* c.371Tdel;p.(Phe124fs).

As previously discussed we have identified the *NBAS* gene as causative of bone fragility in individuals with a complex OI/SOPH/Liver failure syndrome (Balasubramanian, Hurst et al. 2017). The

c.3776C>T; p.(Ser1259Phe) missense variant lies within the sec39 secretory domain, a region of the protein where causative missense mutations have now been described (Li, Qiu et al. 2017). However, interrogation of the complete VCF for this patient identified that this variant call is of very low quality and likely represents a false positive. No second candidate *NBAS* variant was identified in this patient and in addition the clinical features of the patient are not consistent with this gene being causative.

The second variant in patient 117 is a heterozygous c.371Tdel;p.(Phe124fs) variant in the *UGGT1* gene. UGGT is a glycan processing enzyme that acts as a folding sensor within the ER as part of the unfolded protein response (UPR). It acts by recognising unfolded glycoproteins, reglucosylating a single N-glycan near the misfolded part of the protein. This acts as a marker for sending the protein back for further folding attempts or for degradation (Figure 78), (Määttänen, Gehring et al. 2010, Izumi, Kuruma et al. 2017).

Proteoglycans are a known key component of bone tissue and have a broad function, organizing the extracellular matrix and participating in collagen fibrillogenesis (Lamoureux, Baud'huin et al. 2007). It could be speculated that pathogenic mutations in *UGGT1* could therefore affect bone structure and function, without a direct role in collagen I processing/trafficking being identified for the UGGT protein (Maiers, Kostallari et al. 2017).

The *UGGT1* gene is located at chromosome position 2q14.2 and contains 41 exons comprising 1555 amino acids, with the first 42 amino acids forming the signal peptide. The c.372del variant, in exon 4, is predicted to cause a frameshift, p.(Phe124fs), leading to a premature termination codon 13 positions downstream. It is likely to result in nonsense mediated decay of the mRNA. This variant has not been reported in the general population (gnomAD dataset) and the gene has not so far been associated with any human phenotype (OMIM).

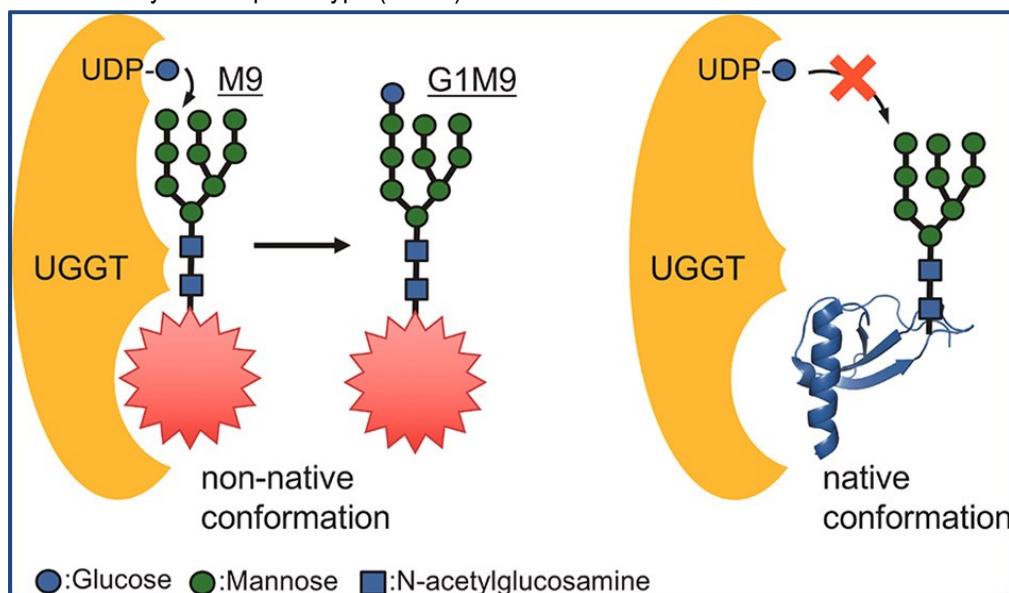


Figure 78 Discrimination of folding status of glycoprotein by folding sensor UDP-glucose:glycoprotein glucosyltransferase (UGGT) in the glycoprotein quality control system.

UGGT transfers a glucose residue only to the M9 oligosaccharide on a glycoprotein having non-native conformation. Figure adapted from Izumi et al (Izumi, Kuruma et al. 2017).

Interrogation of the gnomAD dataset identified a further 8 'frameshift' variants in the gene. The majority of these have been identified in a single allele with the exceptions of c.1040_1041del;p.Lys347Argfs*3 (rs755827151) which is reported in 2/245270 alleles), c.1828dupT;p.(Ser610Phefs*4) identified in 4/276846 alleles (rs767476081) and c.4303_4304del;p.Arg1435Glyfs*11 present in 3/246254 alleles (rs75447246). As none of these are at a particularly high frequency in the general population it is difficult to assign significance to the *UGGT1* variant identified in this patient.

A further variant of interest was identified in patient 117, a heterozygous c.587T>C;p.(Ile196Thr) in *SULF2*. This is the second variant in *SULF2* identified in this cohort (patient 117, discussed previously). The c.587T>C;p.(Ile196Thr) variant has not been reported in the gnomAD database and *in silico* conservation analysis is inconclusive. However, it affects a highly conserved nucleotide and amino acid (up to fruitfly, considering 12 species) that, similarly to patient 117, lies within the sulfatase extracellular domain.

A next step in the investigation of the significance of the *UGGT1* and *SULF2* variants identified in this patient would be to test parental samples to ascertain if they are inherited or if they appear *de-novo*. This would give support to which, if either, are worth exploring further.

Patient 141: *BMP6* c.334G>A;p.(Glu112Lys) and *POSTN* c.1108+2T>C

Patient 141 is described as having an abnormality of mast cells, autonomic dysregulation, joint hypermobility, orthostatic hypotension, osteopenia, seizures and joint dislocations as a young adult. Many of these symptoms (autonomic dysregulation, joint hypermobility, orthostatic hypotension and joint dislocations) are commonly associated with hypermobile Ehlers Danlos Syndrome, a common connective tissue disorder with no identified genetic cause at present.

A heterozygous maternally-inherited c.334G>A;p.(Glu112Lys) variant in bone morphogenetic protein 6 (*BMP6*) was identified in this individual. *BMP6* is a major activator of the BMP-SMAD signalling pathway (Beederman, Lamplot et al. 2013) and has the ability to induce bone and cartilage bone mineralization and influence osteoblast differentiation (Lavery, Swain et al. 2008). It is also known to have a major role in iron metabolism and lack of *BMP6* results in massive iron overload (Meynard, Kautz et al. 2009).

The c.334G>A;p.(Glu112Lys) variant has not been previously reported but lies in the propeptide domain of the protein known to be crucial for *BMP6* processing and secretion. A similar variant affecting the same amino acid, c.334G>C;p.(Glu112Gln), has been proposed to cause iron overload (hereditary haemochromatosis) when in the heterozygous state (Piubelli, Castagna et al. 2017). A note of caution is that it is also reported in 0.30% of the Japanese population (HGVD).

In silico conservation analysis of c.334G>A;p.(Glu112Lys) does not support a pathogenic role and no symptoms of iron overload are reported in this individual. It seems likely therefore that the c.334G>A;p.(Glu112Lys) is a rare benign variant, although a role in bone metabolism in this patient cannot be completely excluded, particularly as it is maternally inherited and there is maternal hypermobility, osteopenia and autonomic dysregulation.

A second maternally inherited variant, c.1108+2T>C affecting the splice donor site of intron 8 of the *POSTN* gene, was also identified in patient 141. This variant is predicted to result in complete loss of the wildtype donor site (Figure 79) and potentially result in a skip of exon 8. Population frequency data shows it is present at very low level in the general population (3/243624 alleles in geNOMAD)

Splice prediction was undertaken using Alamut Visual 2.9 (Interactive Biosoftware, Rouen, France). All five prediction packages support complete loss of intron 8 wildtype donor site. *POSTN* encodes periostin, a multi-functional extracellular matrix protein with a role in tissue repair, inflammatory response, cardiovascular and respiratory systems as well as bone metabolism (Conway, Izuhara et al. 2014). Key to its function in connective tissue is that it forms an ECM assembly 'scaffold', contributing to tissue mechanical strength (Kudo and Kii 2018).

Periostin binds to type I collagen, tenascin C, lamin γ 2 and fibronectin and enhances BMP1 incorporation into the fibronectin matrix, thereby influencing downstream activation of lysyl oxidase (LOX) and subsequent collagen crosslinking (Trackman 2016). The formation of the periostin-BMP1-LOX complex is thought to underlie collagen fibre mechanical strength (Kudo and Kii 2018). A role in secretion of fibronectin from the ER has also been demonstrated (Kii, Nishiyama et al. 2016).



Figure 79 *In-silico* splice prediction of heterozygous *POSTN* c.1108+2T>C variant identified in patient 141.

Animal model studies have shown that collagen fibrillogenesis is disrupted in osteoblasts in periostin null mice although bone mineral density of cortical bone was indistinguishable from that of wildtype mice (Kii, Nishiyama et al. 2010). Increased serum periostin levels have also been observed in older women with osteoporotic hip fracture when compared to age matched controls (Yan, Liu et al. 2017), a finding that suggests periostin plays a role in fracture repair. Higher serum periostin levels were

also shown to be an independent risk factor for femoral neck BMD, suggesting that periostin plays a role in maintaining bone mass.

Of note is that periostin has been shown to be a key downstream target for SULF2-induced angiogenesis in hepatocellular carcinoma cells with upregulation of SULF2 resulting in both increased expression and secretion of periostin via the TGF β 1/SMAD signalling pathway (Chen, Nakamura et al. 2017).

It seems reasonable to consider *POSTN* as a candidate gene for at least some of the symptoms in this family. Further exploration of *POSTN* and *SULF2* variants within this cohort of patient would also be worth considering particularly as SULF2-POSTN has been proposed as a potential target for treatment to increase BMD and/or aid fracture healing. A first step could be to look at secretion of fibronectin from fibroblast ER in patient 141 and her mother in order to assess the functional effect of the *POSTN* variant.

NOTE: All genes of interest identified in the DDD study were subsequently analysed for all exome patients (IDs A-H). No additional candidate variants were identified in these genes.

ix DISCUSSION

This research aimed to contribute to the understanding of the molecular basis of Osteogenesis Imperfecta, a rare inherited condition characterised by low bone mass, increased bone tissue mineral density and an increased risk of fracture.

Genetic characterisation of families affected with OI has shown that autosomal dominant mutations in the genes that encode the alpha chains of type I collagen can be identified in ~85% of affected individuals (Rohrbach and Giunta 2012). Research to identify additional genes has largely focused on individuals with lethal or severe OI, often from large consanguineous families and has identified a further fifteen genes to date (Marini, Forlino et al. 2017).

It is clear that there are families with mild/moderate OI without an identified causative mutation in the type I collagen genes. We wanted to explore the genetic cause of OI in these individuals.

The **first objective** of this study was to identify a cohort of families with mild/moderate OI without a mutation in the type I collagen genes. We recruited eighteen families for targeted exome analysis and eight families for whole exome sequencing. An additional three patients were identified during the course of this study following introduction of *BMP1* gene analysis into the diagnostic pathway as a result of the targeted exome analysis findings.

The **second objective** was to elucidate the genetic aetiology of bone fragility in this cohort. We used a number of different sequencing strategies: NGS targeted exome sequencing looking at genes previously described in severe OI, single gene Sanger sequencing and NGS whole exome sequencing.

We identified the likely cause of OI in two patients using our targeted exome sequencing. The first patient is heterozygous for a c.4070T>C;p.(Leu1357Pro) *COL1A1* likely pathogenic mutation within the C-propeptide domain of the protein. This mutation is within the chain-chain recognition sequence, potentially disrupting chain alignment and assembly. Previously described mutations in this region of the gene have been associated with both mild OI and a high bone mass phenotypic variant. We did not identify any features of high bone mass in the three affected members of this family. In fact the symptoms in an adult member of this family were mild enough for him to be unaware that he is affected. This is concordant with the mild clinical severity as described by Pace et al (Pace, Kuslich et al. 2001). Originally the *COL1A1* mutation was thought not to be causative due to apparent non segregation. On clinical review of the family, including evaluation of additional individuals and revision of affected/unaffected status, it became apparent that this mutation does segregate with the features of OI in this family and is highly likely causative. This result highlights the importance of phenotypic data in the interpretation of genetic findings.

The second patient identified by targeted exome sequencing is compound heterozygous for two pathogenic mutations affecting the *BMP1* gene: a c.1293C>G;p.(Tyr431*) and a c.1148C>G;p.(Arg383Gln). At the time of this study less than 20 individuals with OI had been identified with mutations in the *BMP1* gene. The majority of these, but not all, have presented with bone fragility associated with an increase in bone mineral density (BMD). Our patient also presented

with increased BMD that become more evident following anti-resorptive treatment. We concluded that the c.1293C>G;p.(Tyr431*) and c.1148C>G;p.(Arg383Gln) changes were highly likely to be the cause of OI in this patient.

This result prompted the inclusion of *BMP1* in the OI diagnostic testing panel within SDGS. This led to the subsequent identification of two patients (siblings) who are homozygous for a c.2188dupC;p.(Gln730fs) highly likely pathogenic mutation in the *BMP1* gene. The clinical features of these siblings contrasts with the severe progressive form of OI described in patients who are compound heterozygous for this mutation and the recurring p.(Gly12Arg) signal peptide mutation (Syx, Guillemyn et al. 2015). We presented the phenotypic and BMD data for these three patients in our American Journal of Medical Genetics manuscript (Pollitt, Saraff et al. 2016). Following publication of the manuscript, a further patient with a high bone mass phenotype and a novel homozygous c.355C>T;p.(Arg119Trp) mutation in the pro-domain of the *BMP1* gene was identified.

One possible explanation for the variable phenotype in this small cohort of *BMP1* OI patients is that the degree of mineralization is a direct result of the 'severity' of the mutation. However, we were unable to demonstrate any difference in collagen C-propeptide cleavage activity by biochemical analysis in dermal fibroblasts, although accumulation of pCa1(I) procollagen species confirmed pathogenicity. Characterisation of collagen fibril and fibroblast ultra structural features also gave no clues as to the mechanism responsible for the variable phenotypic severity in this cohort.

Based on these results, it seems reasonable to suggest that mechanisms other than collagen processing play a part in determining the degree of mineralization in these patients. It would be interesting to measure decorin processing in order to confirm the previous observations by Syx et al who reported defective proteolytic procollagen processing in *BMP1* OI patients (Syx, Guillemyn et al. 2015). As decorin mediates TGF- β signalling, we can speculate that reduced decorin prodomain removal is contributing to the increased mineralization in these patients by influencing downstream TGF- β signalling. Assessment of TGF- β signalling in our cohort would help to determine the validity of this hypothesis.

An important conclusion from this part of the study is that caution is required for anti-resorptive therapy in *BMP1*-related OI due to the potential risk of delayed healing, increased bone stiffness, atypical fractures and iatrogenic osteopetrosis. In our manuscript we have stated that careful monitoring of response to bisphosphonate therapy is indicated in these patients.

Our data also clearly demonstrates the value of genetic diagnosis to inform treatment decisions in OI and the importance of inclusion of *BMP1* analysis in any diagnostic screening strategy. We also show that this analysis should not be confined to patients with a noticeable high bone mass phenotype as bone mass in *BMP1*-related OI can be highly variable.

It is worth noting that targeted exome sequencing, as used here, can only identify certain types of mutations, such as single base substitutions and small deletion/insertions. Larger changes and rearrangements are not detected. In addition, as probes are designed to capture specific regions of known genes and are usually focused on protein coding sequence, deeply intronic variants and those in regulatory regions will be missed. There is also very limited potential for novel gene discovery. However, an advantage of this strategy is that the identification of significant incidental findings is avoided.

A second strategy to identify the genetic cause of OI in this cohort was to screen all participants without an identified pathogenic mutation for mutations in the *TRAM2* gene using Sanger sequencing. Phenotypic data from a knockout mouse model suggested this to be a good candidate gene for OI. Only variants present at relatively high frequency in general population datasets were identified and we concluded that the diagnostic yield from this approach is likely to be extremely low and did not pursue this strategy further.

In the third strategy, we performed whole exome sequencing on eight patients who had screened negative for all known OI genes and *TRAM2*. Three of these patients, two singletons and a trio, were commercially sequenced by Personalis. Their complete datasets, along with a candidate gene report, were returned to our centre for confirmation and further analysis. The remaining five patients, four trios and one duo, were sequenced in-house using SureSelect Human All Exon V6 baits for target enrichment and an Illumina HiSeq platform. We identified likely causative mutations in three individuals and a further two that have variants that are currently of uncertain significance.

Causative mutations include compound heterozygous c.1224-80G>A and c.1080+1G>T in *P3H1*, a heterozygous c.1178A>G;p.(Tyr393Cys) *P4HB* mutation previously described in Cole Carpenter syndrome (CCS), and compound heterozygous c.5741G>A;p.(Arg1914His) and c.3010C>T;p.(Arg1004*) in the *NBAS* gene. The variants of uncertain significance are compound heterozygous c.6971G>A;p.(Arg2324His) and c.4358G>A;p.(Cys1453Tyr) in *NBAS* in one patient and *de-novo* c.9029C>A;p.(Pro3010His) *SRCAP* and c.1169C>T;p.(Pro390Leu) *TAPBP* variants in another patient.

Several observations can be made from these mutation data. The first is the advantage of extending sequence analysis to include non-coding regions of the gene. The *P3H1* c.1224-80G>A was identified as a result of exon/intron architecture in this region of the gene, in that the intron in which the mutation lies is relatively small (174bp). Therefore, targeting sequence capture to surrounding exons has captured intronic sequence by default. It follows that more deeply intronic variants, that may be pathogenic, are being missed using current exome sequencing strategies.

A further observation is that the diagnostic yield from our whole exome sequencing is 38%; this would increase to 63% if we can demonstrate causality in our two patients with variants of uncertain significance. This is concordant with reported diagnostic yield in larger exome sequencing projects such as UK10K and Decipher, and supports the clinical utility of this sequencing strategy in patients where the routine diagnostic testing is negative. Additional analysis for the patients where we have not identified causal mutations could include whole genome sequencing (WGS), array CGH, or multiplex ligation-dependent probe amplification (MLPA).

Our identification of the c.1178A>G;p.(Tyr393Cys) mutation represents only the third case of Cole Carpenter Syndrome caused by a defect in *P4HB* and suggests the possibility of this recurrent *P4HB* mutation being the cause of this syndrome. No other *P4HB* variants have been reported in association with any clinical phenotype. In addition, detailed phenotyping in our patient with the c.1178A>G;p.(Tyr393Cys) *P4HB* mutation highlights what appears to be an emerging, distinctive radiological phenotype for these CCS patients: meta-diaphyseal fractures with metaphyseal sclerosis, a feature that is uncommon in patients with classical OI. We have also identified that early treatment with anti-resorptive therapy in patients with a proven *P4HB* mutation can be beneficial in

reducing fracture incidence and vertebral compression fracture frequency (Balasubramanian, Padidela et al. 2017).

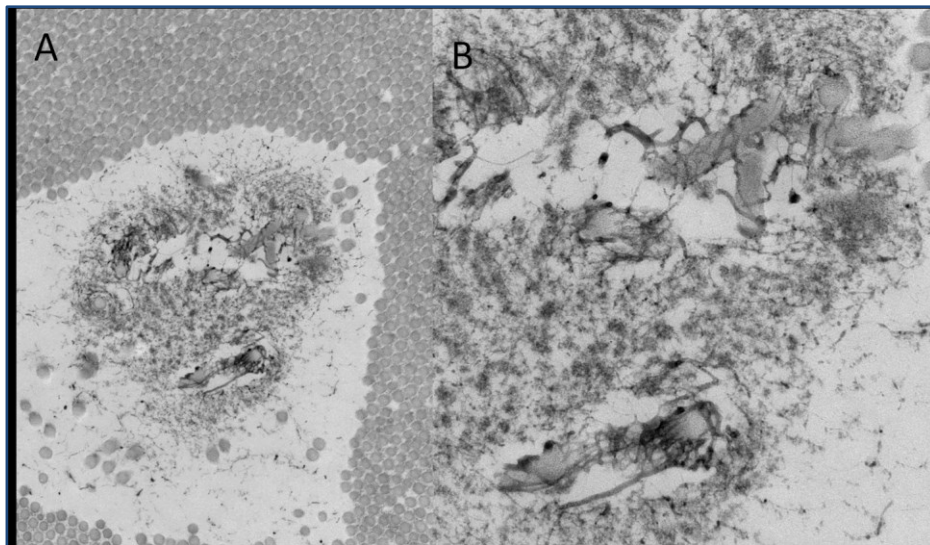


Figure 80 Electron microscopic ultra structure of dermal collagen in an OI patient with a c.838G>A, p.(Gly280Ser) COL1A2 mutation.

A small collection of individual collagen fibrils surrounded by proteoglycans were observed. The significance of this is unclear. Magnification: A=9200X, B=20000X.

It has previously been proposed that increased ER stress is the mechanism by which the *P4HB* mutation leads to CCS (Rauch, Fahiminiya et al. 2015). However, using morphological examination of reticular dermis by EM and protein expression analysis, we observed no major difference in extracellular type I collagen deposition by immunofluorescence studies, little effect on collagen fibril assembly and no sign of significant fibroblast ER stress. Our data therefore does not support this hypothesis. However, it is possible that the *P4HB* mutation may cause osteoblast ER stress and not fibroblast ER stress. It would be interesting to see if we could reproduce the altered distribution of the PDI protein in our patient's fibroblasts or increased HSP47 protein, in support of ER stress, as previously reported (Rauch, Fahiminiya et al. 2015). Additional analysis to look for upregulation of BiP (UPR marker) or autophagy, supporting an ER stress mechanism, may also be informative.

Our data allows us to speculate that it is the assembly, composition or architecture of the extracellular matrix that is disorganised in patients with this *P4HB* mutation rather than collagen secretion itself. An observation in our patient that may support this idea is the presence of four small collections of individual collagen fibrils surrounded by proteoglycans (Balasubramanian, Padidela et al. 2017). Functional analysis will be required to explore this further; a potential starting point could be to examine other collagen and proteoglycan function in these patients.

Interestingly, since publication of our manuscript describing this feature (Balasubramanian, Padidela et al. 2017), we have identified an OI patient with a c.838G>A;p.(Gly280Ser) pathogenic mutation in exon 17 of the *COL1A2* mutation that also has an area where proteoglycan is wrapping around

collagen fibrils (Figure 80). We have also identified this feature in our patient with a *de-novo* SRCAP variant (Figure 54, page 97). This raises the question as to whether this is an indication of disturbed extracellular matrix in OI or whether this is in fact a rare normal feature of dermal breakdown and repair. Further work on a broader range of dermal biopsies, that includes non-OI patients, is ongoing to try to clarify the significance of this finding. A note of caution here is that we cannot discount that this is a tissue specific finding and may not be reflected in other tissues, such as bone.

Our study has confirmed the potential of whole exome sequencing to identify novel causes of bone fragility by the identification of compound heterozygous c.5741G>A;p.(Arg1914His) and c.3010C>T;p.(Arg1004*) mutations in the NBAS gene. We were able to demonstrate reduced levels of the NBAS protein by Western blot analysis in patient cultured fibroblasts (performed in another laboratory) and the presence of Pelger Huët anomaly in the patient's neutrophils, confirming pathogenicity of the mutations.

The NBAS gene is known to cause short stature, optic nerve atrophy and Pelger-Huet anomaly syndrome (SOPH), and acute onset liver failure. More recently NBAS has been described as causing a multisystem disorder with a skeletal phenotype that can include early onset osteoporosis (Capo-Chichi, Mehawej et al. 2015, Segarra, Ballhausen et al. 2015, Staufner, Haack et al. 2016). Our data adds an important extension to the phenotype associated with this gene to include a presentation resembling 'atypical' OI (Balasubramanian, Hurst et al. 2017). It is worth noting that, although Segarra *et al* report a 'skeletal dysplasia suggesting a disturbance in bone mineralization' (Segarra, Ballhausen et al. 2015), bone data is currently not available from NBAS patients to support this observation.

The NBAS protein is a component of a soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) complex that is involved in golgi to ER retrograde transport (Aoki, Ichimura et al. 2009). We observed grossly expanded protein-filled ER by EM examination of patient fibroblasts, some of which had formed 'aggregates' within the ER. This strongly supports that golgi-ER transport is compromised in our patient.

NBAS is also involved in nonsense mediated decay (NMD) regulation of gene expression including genes associated with bone mineralisation, osteoblast differentiation and bone development. It has already been proposed that NBAS dysfunction in the NMD processes may explain the short stature and growth retardation described in SOPH syndrome. In particular, upregulation of matrix Gla protein (MGP), that acts as an inhibitor of cartilage calcification, has been confirmed in NBAS deficient zebrafish and *C. Elegans* (Longman, Hug et al. 2013)

We suggest that it is possibly a combination of Golgi-ER transport and NMD dysfunction that ultimately result in bone fragility in NBAS deficient patients (Balasubramanian, Hurst et al. 2017). We could also speculate that altered golgi-ER transport may indirectly affect NMD. To explore the role of the NMD pathway, we aim to perform RNA expression analysis in our patients, with a particular focus on genes associated with bone mineralisation, osteoblast differentiation and bone development.

We concluded that mutations in NBAS are a novel cause of heritable bone fragility and that this gene should be included in the diagnostic testing strategy for OI. The development of a gene dossier that includes testing criteria, and its acceptance and publication by UKGTN, is a significant advancement

in equitable access to *NBAS* testing for patients in the UK with associated implications for diagnosis, prognosis and to inform discussions around recurrence risk for these patients.

Using exome sequencing, we also identified a further patient with short stature, osteopenia and fractures who is compound heterozygous for a c.6971G>A;p.(Arg2324His) and a c.4358G>A;p.(Cys1453Tyr) in the *NBAS* gene. These two variants are of unclear significance at present. The patient and his family have been invited for further investigation; initially this would involve detailed phenotyping, an FBC to look for PHA and liver function tests. Follow on studies of *NBAS* expression in dermal fibroblasts and morphological examination may then be appropriate. To date, the family have declined our invitation.

In a further patient we identified two *de-novo* variants, c.9029C>A;p.(Pro3010His) in *SRCAP* and c.1169C>T;p.(Pro390Leu) in *TAPBP*. Truncating mutations in *SRCAP* are associated with Floating Harbour Syndrome (FHS), a condition characterized by short stature, delayed bone age and delayed speech (Messina, Attarrato et al. 2016). To date, no missense variants in *SRCAP* have been reported in association with a clinical phenotype.

The *SRCAP* protein is the catalytic component of a complex that mediates the ATP-dependent exchange of histone H2AZ/H2B dimers for nucleosomal H2A/H2B leading to transcriptional regulation of genes by chromatin remodelling. Independently *SRCAP* is a coactivator for the CREB-binding protein (CREBBP, also known as CBP). CREBBP acts with EP300 to regulate the transcription of a number of bone-related genes. We considered the c.9029C>A;p.(Pro3010His) variant to be a good candidate for a potential role in bone fragility.

One mechanism for the potential pathogenicity of this variant could be via changes in chromatin remodelling and subsequent changes in DNA methylation status. We analysed the DNA methylation status of our patient against normal and FHS control populations. The patient did not have the characteristic methylation status of FHS, confirming that he does not have a phenotypic variant of this disorder. A total of 171 regions with altered methylation status were identified when compared to normal controls. These regions contained two genes associated with bone metabolism, *POLCE* and *LOXL3*, both of which are hypomethylated. It is not clear whether the hypomethylation status of these genes is a response to, rather than being the primary cause of, altered bone metabolism. A note of caution is that methylation is often tissue specific and we have only examined status in peripheral blood. Further analysis in other tissues, particularly bone if available, may be helpful. However, it may still be difficult to distinguish causative changes to methylation status and, hence gene expression, from compensatory changes in response to other unidentified mechanisms.

We considered the interaction of *SRCAP* with CREBBP/EP300 and activation of CREB-mediated transcription as a further potential mechanism for this variant to cause bone fragility. Potential pathways include cAMP/CREB signalling pathways in osteoblasts such as mediation of PTH and/or BMP2 signalling. However, our patient has no radiographic features of the cleidocranial dysplasia associated with *RUNX2*, a downstream target of BMP2. Evidence is therefore lacking to support this as a pathogenic mechanism.

The *de novo* c.1169C>T;p.(Pro390Leu) in *TAPBP* highlights the dilemma of heterozygous findings in autosomal recessive genes that are potentially relevant to a patient's phenotype. The gene is described in Type I Bare Lymphocyte Syndrome, a condition characterised by chronic bacterial

infection, and our patient has mild gammaglobulinaemia and non response to pneumococcal vaccination. It is possible that either there is a second unidentified mutation elsewhere in the gene, that this represents a milder dominant form of the syndrome or that it is not relevant to the patient's clinical symptoms. Trying to reach a conclusion in these circumstances is challenging, particularly where the variant is *de-novo*. There is a reported frequency of *de-novo* point mutations per generation in the region of 70-175, of which approximately 3 would be expected to cause changes in protein coding sequence (Nachman and Crowell 2000, Conrad, Keebler et al. 2011). It therefore follows that a *de-novo* missense variant may not necessarily be associated with a deleterious clinical phenotype. Serological or flow cytometric analysis of HLA status, specifically the expressions of class I MHC molecules, would provide evidence to inform the clinical significance of this variant.

The **third objective** of this study was to explore a broader dataset of variants from additional patients by establishing a collaborative complementary analysis project with the 'Deciphering Developmental Disorders' (DDD) project. The required subset of patient data was identified by searching the DDD database using Human Phenotype Ontology (HPO) terms relating to reduced bone mass and potential bone fragility.

Forty nine patients were selected and a total of 939 candidate variants were assessed. Nine patients were reported to have a clinical diagnosis of OI and we identified causative variants in four of these. Our discovery of two mutations, *COL1A2* c.280-2A>C and *IFITM5* c.-14C>T, that were not reported by Decipher to the referring clinician, highlights a number of issues with current pipelines for large dataset analysis.

The first is that, even where a variant is of high sequence quality, causative mutations may not be reported as being relevant to the patient's clinical phenotype. In the case of *COL1A2* c.280-2A>C, the patient is stated as having a clinical diagnosis of OI and a known *COL1A2* mutation. This information is provided in the 'additional notes' data field rather than being in the HPO field. The HPO terms do however state 'bowing of limbs due to multiple fractures' and 'multiple prenatal fractures', both of which should have highlighted the *COL1A2* mutation. It is therefore unclear why this was not thought relevant to the patient's symptoms and leads us to suggest that caution is required in the interpretation of any negative results provided by automated pipelines.

The second issue is that, as discussed previously, intronic sequence may not be reliably captured by whole exome sequencing. In the case of the *IFITM5* c.-14C>T mutation, the sequence data had fallen below the acceptable quality threshold and was therefore disregarded. We could also speculate that, as this an intronic variant, it is unlikely to have been highlighted as potentially causative even if it was of good quality due to the limitations of splicing variant analysis in current analysis pipelines. It is worth stating that, for these two patients at least, standard diagnostic targeted panel analysis would have had an advantage over whole exome analysis as this includes specific probes to capture the *IFITM5* c.-14C>T mutation and all variants are manually assessed for likely causality.

For patients without a stated clinical diagnosis of Osteogenesis Imperfecta a total of 19 candidate variants for bone fragility were identified in 12 patients. Two patients without a stated clinical diagnosis of OI were found to have a pathogenic mutation in the type I collagen genes, *COL1A1* c.2644C>T;p.(Arg882*) and *COL1A1* c.1299+1G>A. This could suggest that milder forms of OI are under diagnosed.

Our data also highlighted the potential phenotypic overlap between OI and Ehlers Danlos syndrome. We identified compound heterozygous mutations in the *B4GALT7*, the gene associated with spondylodysplastic EDS (spEDS). The patient is reported as having mild short stature and generalized hypotonia, both of which are within the diagnostic criteria for spEDS. The additional clinical features of osteopenia, bowing of the long bones and vertebral compression fractures are more suggestive of OI. In our centre we have also identified patients with OI type I collagen mutations previously clinically diagnosed with classical EDS. We concluded that a broader OI/EDS testing strategy is clearly indicated in a diagnostic setting. This is further supported by the identification of a heterozygous *PLOD1* mutation, associated with kyphoscoliotic EDS, in our data.

As with the *TAPBP* *de novo* variant described earlier, the heterozygous *PLOD1* mutation in our data illustrates the difficulty with assessing the significance of these findings. Generally heterozygous variants, even if likely pathogenic, are not reported for autosomal recessive genes by exome analysis pipelines. Clearly important findings relevant to a patient's phenotype may therefore be missed and manual review of variants would be advantageous. Our data also illustrates the difficulty in distinguishing heterozygous carriers from patients with a 2nd mutation that is unidentified by current exome sequencing analysis and would confirm a diagnosis. A recurrent *PLOD1* mutation, a duplication of exons 10 to 16, would not be detected by the Decipher bioinformatics analysis pipeline and therefore the diagnosis in this patient is currently unresolved. We have contacted the patient's clinician and recommended that a urine sample for deoxyypyridinoline to pyridinoline ratio analysis is obtained.

We identified a number of pathways that may be important for the bone fragility in our DDD cohort. The first is voltage-gated sodium channels, where three patients with variants in the *SCN8A* or *SCN9A* suggest a potential pathway between pain sensation, epilepsy/seizures and bone parameters. This is supported by a recent report of a clinical phenotype "consistent with OI type IV" in association with a homozygous c.570G>A;p.(Trp190*) in *SCN9A* (Caparros-Martin, Aglan et al. 2017).

Data is also available that shows increased osteoclast numbers and bone resorption in *Scn8a*-deficient mouse along with disorganized collagen structure and increased mineralization heterogeneity. A patient with *SCN8A*-related epileptic encephalopathy and multiple fractures is reported with low bone mineral density (DXA Z score -2.9) and a decreased bone volume per tissue volume (BV/TV), as assessed by high-resolution peripheral quantitative computed tomography (HR-pQRT) in the tibia. Improvement in BV/TV (+46% at 25 months post treatment initiation) and no further fractures following anti-resorptive therapy (Neridronate) is described, supporting a bone remodelling/resorption pathology (Rolvien, Butscheidt et al. 2017).

We conclude that further clinical evaluation of our *SCN8A/SCN9A* patients, in particular their bone parameters, is warranted. However, a note of caution is that anti-convulsion medication and reduced mobility, which may be associated with epilepsy/seizures and developmental delay respectively, can both contribute to osteopenia.

A second candidate pathway is highlighted by the c.1A>G;p.(Met1?) variant in the glutamine, glutamate, and aspartate transporter gene, *SLC38A10* (Hellsten, Hägglund et al. 2017). Glutamine has emerged as an important regulator of osteoblasts, being actively metabolised and essential for matrix mineralization (Brown, Hutchison et al. 2011). Glutamine signalling, alongside Wnt signalling via Frizzled (Fz) and Lrp5/6 and activation of mTORC1, increases levels of key enzymes involved in

glucose and glutamine metabolism, notably glutamine anaplerosis into the citric acid cycle (TCA) (Figure 81).

Energy metabolism in osteoblasts has been proposed as a key mechanism in the osteoporosis associated with diabetes mellitus and anorexia nervosa (Lee, Guntur et al. 2017) and *Slc38a10* knockout mice have weak and brittle bones (Bassett, Gogakos et al. 2012). Anabolic treatment for osteoporosis is known to increase bone mass and reduce fracture risk possibly by stimulating osteoblastogenesis and/or altering osteoblast metabolic processes (Lee, Guntur et al. 2017). We propose that *SLC38A10* is therefore a good candidate gene for osteoporosis and increased fracture risk in man.

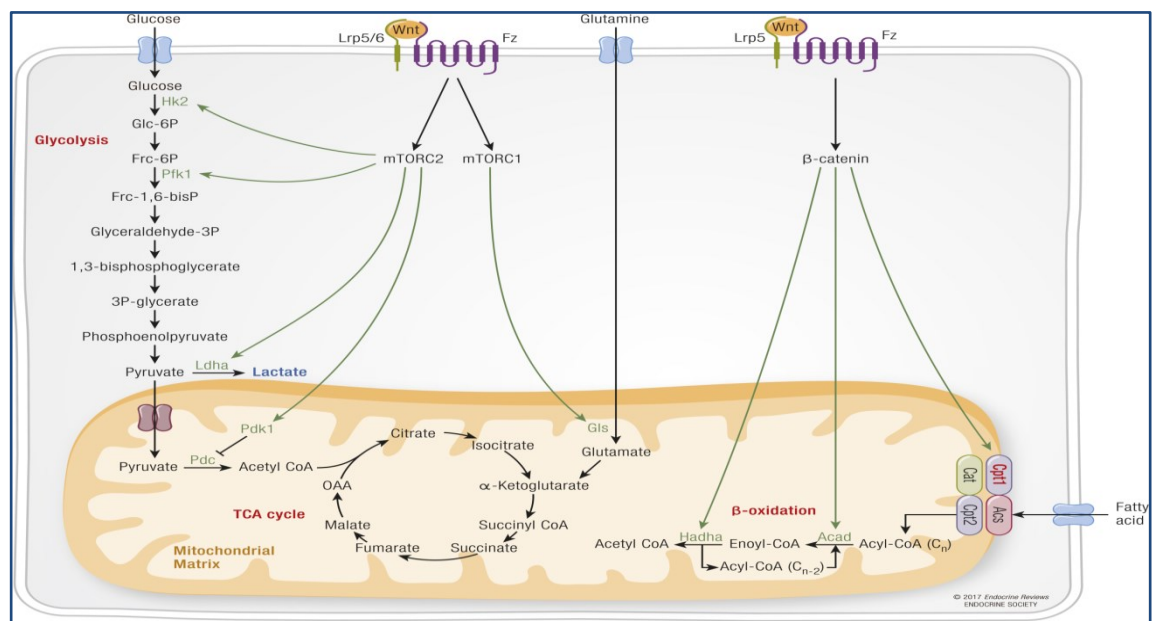


Figure 81 Energy Metabolism of Osteoblasts

Wnt signalling increases both glutamine anaplerosis into the citric acid cycle (TCA) and fatty acid oxidation. Through Frizzled (Fz) and Lrp5/6, *Wnt* signalling activates mTORC1, increasing levels of key enzymes involved in glucose and glutamine metabolism. *Wnt* signaling through β -catenin also increases expression of genes important for fatty acid oxidation. Figure modified from Lee et al (Lee, Guntur et al. 2017)

The clinical features of our patient include generalized osteoporosis with pathologic fractures, otosclerosis and severe generalized osteoporosis. The patient's mother, who is also heterozygous for the c.1A>G;p.(Met1?) variant is also reported to have osteoporosis and otosclerosis. This supports a potential role for this variant in the bone phenotype in this family but is not conclusive. The patient and their family have been invited for further clinical review.

With the identification of a *de-novo* loss of function *LRP6* variant in a patient who has a homozygous c.1225C>T;p.(Arg409Cys) mutation in the *FAM20C* gene, we have demonstrated that there is a risk of incorrectly attributing all phenotypic features to a single gene. It seems likely that both the amelogenesis imperfecta and osteoporosis described in our patient is likely to be at least partly caused by his *LRP6* variant, either in isolation or in combination with his *FAM20C* mutation. However, as the *LRP6* variant was not reported back to the clinician, it has been assumed that *FAM20C* fully explains the patient's whole phenotype.

Novel candidate genes identified in our cohort are *UBASH3B*, *UGGT1*, *SULF2* and *POSTN*. The inheritance of the variants in these genes is unknown, with the exception of *POSTN* which is maternally inherited. It will be important to confirm inheritance, and clinically assess any carrier parents, before embarking on any functional studies to clarify the role of these genes in bone fragility, particularly where patients have variants in more than one candidate gene.

UBASH3B is a loss of function intolerant gene (pLI=0.99 in gnomAD) that is highly expressed in osteoclasts, the immune system, reproductive organs and platelets, and is thought to be a requirement for normal bone resorption. In particular the *UBASH3B* protein negatively regulates osteoclast precursor numbers and in its absence osteoclasts exhibit an enhanced ability to resorb bone. This is supported by increased osteoclast precursors and decreased BMD in knock-out mice (Back, Adapala et al. 2013). *UBASH3B* is also known to be a critical negative regulator of TCR signalling, regulating T-cell activation. We identified a c.371Tdel;p.Phe124fs likely loss of function variant in a patient with, among other features, abnormality of T cells, osteopenia, osteoporosis of vertebrae and severe short stature. We concluded that these symptoms would fit with the reported function of *UBASH3B*.

However, the *UBASH3B* gene is located at 11q24.1, a region of chromosome 11 that is deleted in 8 patients in the Decipher database none of whom have reported osteoporosis, osteopenia or increased susceptibility to fracture. To explore the potential role of *UBASH3B* we plan to investigate the phenotype of these Decipher patients further. We would also like to undertake platelet function analysis and examine the development of osteoclasts from circulating mononuclear cells in our patient.

UGGT1, an ER resident enzyme, acts as a folding sensor reglucosylating unfolded glycoproteins as part of the unfolded protein response (UPR). The c.371Tdel;p.(Phe124fs) loss of function variant identified in our patient may have an effect either by causing ER stress or by release of misfolded proteoglycans that could disrupt extracellular matrix organization. There are several options available to explore the function of the variant in our patient: we could undertake enzymatic activity assays (Trombetta, Bosch et al. 1989); undertake gene and protein expression assays; examine a dermal biopsy for evidence of ER stress and matrix disorganisation; or measure markers of ER stress such as upregulation of BiP. We plan to obtain a dermal biopsy from our patient and assess collagen and fibroblast ultrastructure by EM as a first step.

Of particular interest is that we identified *SULF2* missense variants in both the *UBASH3B* and *UGGT1* patients discussed above (c.2411A>G;p.(Asn804Ser) and c.587T>C;p.(Ile196Thr) respectively). The *SULF2* protein regulates growth factor release from extracellular storage sites, activating multiple signalling pathways including TGFβ1/SMAD (Chen, Nakamura et al. 2017), WNT-GLI1-CYCLIN D1 pathways (Nakamura, Fernandez-Barrena et al. 2013) and Hedgehog/GLI1 (Zheng, Gai et al. 2013). It has been shown to be highly expressed in active osteoblasts, particularly during healing, and has been proposed as a possible future treatment target to aid fracture healing (Zaman, Staines et al. 2016).

We have also identified a c.1108+2T>C splicing variant within the *POSTN* gene, a key downstream target for *SULF2*, with upregulation of *SULF2* resulting in both increased expression and secretion of *POSTN* via the TGFβ1/SMAD signalling pathway (Chen, Nakamura et al. 2017). Periostin binds to type I collagen, tenascin C, lamin γ2 and fibronectin and enhances BMP1 incorporation into the

fibronectin matrix. This in turn influences downstream activation of lysyl oxidase (LOX) and subsequent collagen crosslinking (Trackman 2016). In a similar manner to SULF2, increased serum periostin levels have been observed in older women with osteoporotic hip fracture (Yan, Liu et al. 2017), a finding that suggests periostin may play a role in fracture repair. More recently periostin has been shown to play a protective role against mechanical stretch-induced apoptosis in osteoblast-like MG-63 cells (Yu, Yao et al. 2018).

It seems reasonable to consider *POSTN-SULF2* as a candidate pathway for at least some of the findings in these individuals. As periostin has a role in secretion of fibronectin from the ER (Kii, Nishiyama et al. 2016), a first step to investigate this further could be to assess secretion from fibroblast ER, using immunohistochemistry. Alternatively, we could measure the expression levels of both POSTN and SULF2 in osteoblasts. We are currently exploring what material is available from these patients.

In **conclusion**, we were able to address all research objectives formulated in this thesis and have identified novel genetic mechanisms leading to bone fragility and OI. We have also identified a number of novel genes and pathways that warrant further investigation as outlined above, namely *SLC38A10*, *SRCAP*, *UGGT1*, *UBASH3B*, *SULF2-POSTN*, and voltage-gated sodium channel genes.

x **FUTURE PERSPECTIVE**

During the course of this project there have been considerable advances in genetic technology. We have moved from sequential Sanger sequencing of individual genes, to routine NGS targeted exome panels through to whole exome and genome sequencing.

The launch of the Genomics England 100,000 Genomes Project, with the aim of sequencing patients with rare diseases and cancer, means there will be an ongoing expansion of phenotypic and genetic data available. This has the potential to improve identification of novel disease genes, phenotypic modifiers, factors influencing disease progression, and optimize development of novel therapies and treatment regimes.

Human Phenotypic Ontology (HPO) Term	
Abnormal heart morphology	Kyphosis
Abnormality of cardiac morphology	Low-set ears
Abnormality of the hip bone	Multiple prenatal fractures
Abnormality of the respiratory system	Multiple rib fractures
Abnormality of the teeth	Multiple small vertebral fractures
Abnormality of the vertebral column	Nephrolithiasis
Autistic behaviour	Osteoarthritis
Basilar impression	Osteopenia
Basilar invagination	Periosteal new bone (humeral diaphysis)
Blue sclerae	Periosteal new bone (humeral)
Bruising susceptibility	Periosteal new bone of humeral diaphysis
Calcification of the interosseus membrane of the forearm	Periosteal new bone of humerus
Coarse humeral trabeculae	Platybasia
Coarse trabeculae (humeral)	Platyspondyly
Contractures of the joints of the lower limbs	Pointed chin
Contractures of the joints of the upper limbs	Protrusio acetabuli
Corneal perforation	Scoliosis
Craniosynostosis	Sensorineural hearing impairment
Cutis laxa	Spondylolisthesis
Dentinogenesis imperfecta	Spondylolysis
Fractures of the long bones	Turricephaly
Gastric ulcer	Wormian bones
Gastroesophageal reflux	Global developmental delay
Joint dislocation	Joint hypermobility

Table 25 HPO data model developed by the 100,000 Genomes Project Musculoskeletal GeCIP, Bone Fragility SubDomain

Our study has highlighted some potential obstacles to gaining maximum benefit from the 100,000 Genomes Project data likely to become available for OI patients. The first is the need for a robust clinical dataset that allows deeper phenotyping, reducing the need for immediate clinical re-evaluation of the patient and their family if a variant is identified. The second is the need for improved annotation and filtering of sequencing data particularly splice site variants and non-coding regions, including regulatory elements, of the genome. Thirdly, analysis is currently focused on monogenic disease; there are significant interpretation challenges to identify polygenic associations and development of more sophisticated bioinformatics tools is indicated. Finally, even with the relatively small patient numbers in our study, we have identified a number of potential genes/pathways that warrant further analysis. As the number of candidate variants increases, during the 100,000 Genomes Project and beyond, there will be a growing need for more complex and functional analysis to inform variant causality.

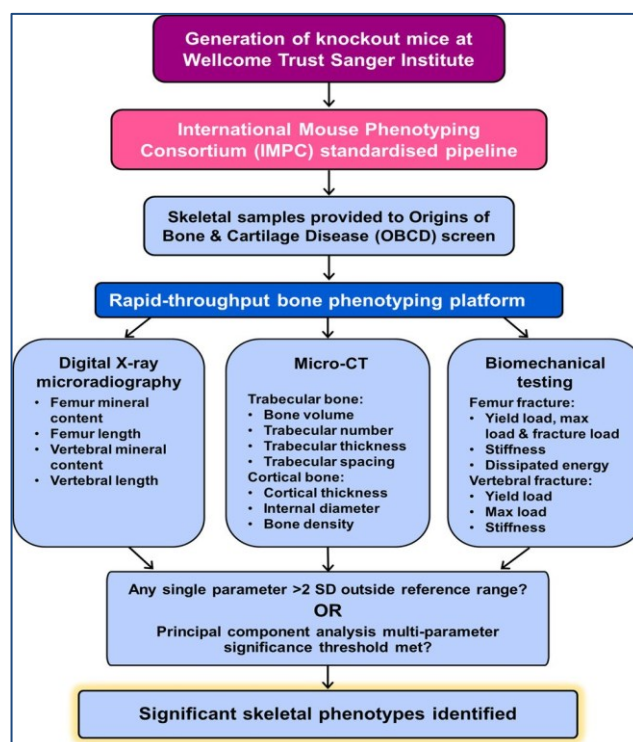


Figure 82 Flow chart showing the OBCD bone phenotyping platform.

The aim of this workflow is to identify significant abnormal skeletal phenotypes in mouse models, in conjunction with the IMPC standardised phenotyping project (Freudenthal, Logan et al. 2016)

To address some of these issues, a bone fragility sub-domain of the Musculoskeletal (MSK) Genomics England Clinical Interpretation (GeCIP) research group has been formed. I am the lead scientist for the MSK GeCIP. We have instigated a minimum data model for patient recruitment that provides more detailed skeletal phenotyping information than provided for the Decipher patients in this study (Table 25).

We have developed an initial research investigation pipeline using BAM and VCF files and summary data on SNVs, small indels and copy-number variant analyses. We will use a web-based utility (vcf2hgvs) to convert VCF variant calls into HGVS nomenclature-compliant variant descriptions.

These will first be interrogated using an inheritance based approach (Strategy iii.ix.i, page 57). Genes showing *de novo* variants and other very rare variant combinations will be analysed against other OI patients from our existing cohort and those of our GeCIP collaborators.

Data will be compared with candidates identified from our in-house exome studies and these candidates will be further explored by cell based studies through ongoing collaborations with colleagues in the Centre for Membrane Interaction and Dynamics (University of Sheffield). We plan to use zebra fish animal modelling as a screening tool for pathways/activity in both embryonic and adult developmental stages by collaboration with the Sheffield Zebrafish Screening Facility at the Bateson Centre.

Collaboration with the extended effort to phenotype mouse models by the Origins of Bone and Cartilage Disease (OBCD) project, part of the International Mouse Phenotyping Consortium, will be ongoing. They have a comprehensive rapid throughput phenotyping platform (Figure 82) designed to identify the monogenic cause of bone and cartilage disease; we will focus on those genes that may be important to skeletal fragility.

The data from our project has established that this approach to sequence analysis can identify novel causes of OI and bone fragility. An accurate genetic diagnosis can inform treatment decisions, lead to appropriate family counselling, improve our understanding of disease pathways and facilitate development of novel therapies. It also supports research into the assessment of response to therapy and our understanding of genotype/phenotype correlations.

xi FUNDING AND ETHICAL APPROVAL

I initiated this research project and wrote the application to obtain funding from the Sheffield Children's Charity (Protocol Number: SCH/11/062) for the targeted exome sequencing.

I completed the ethical approval application which was obtained following interview from the Yorkshire and Humber NRES committee (Reference Number: 12/YH/0021). R&D approval was obtained from the Clinical Research Facility at Sheffield Children's NHS Foundation Trust. Approval has also been obtained at each host organisation prior to the start of patient recruitment at that site, in accordance with NHS research governance procedures.

I co-authored the grant application for whole exome sequencing to The Sheffield Children's Hospital Charity (TCHC) grant number CA15001. I also co-authored the Complementary Analysis Project application to the Deciphering Developmental Delay (DDD) project (CAP12).

The DDD study presents independent research commissioned by the Health Innovation Challenge Fund [grant number HICF-1009-003], a parallel funding partnership between the Wellcome Trust and the Department of Health, and the Wellcome Trust Sanger Institute [grant number WT098051]. The views expressed in this publication are those of the author(s) and not necessarily those of the Wellcome Trust or the Department of Health. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and GEN/284/12 granted by the Republic of Ireland REC). The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network. This study makes use of DECIPHER (<http://decipher.sanger.ac.uk>), which is funded by the Wellcome Trust

xii ACKNOWLEDGEMENTS

I would like to express gratitude to the patients, their families and all the referring clinicians who participated in this research study.

I thank my supervisors Professor Nick Bishop and Professor Ann Dalton for their time, support and for providing me with the opportunity to work on this research project.

Also, special thanks to the team at Sheffield Diagnostic Genetics Service for their patience and willingness to help with this research, even when under considerable workload pressures. In particular, Emilie Jarrat for her assistance with the exome sequencing and Lucy Crooks, Matt Parker, Lizzy Sollars and Natalie Groves for doing amazing things with genetic data (aka bioinformatics).

I am very grateful to Dr Meena Balasubramanian and Dr Paul Arundel for their assistance with patient recruitment, their willingness to share clinical information, and their enthusiasm for collaborative research projects. I want to express thanks to all co-authors of manuscripts for sharing samples, data and for their helpful discussion.

I would also like to acknowledge Dr Bekim Sadikovic and Dr Erfan Aref-Eshghi at the London Health Sciences Centre, Ontario, Canada for undertaking methylation studies; Dr Chris Lelliot at the Wellcome Trust Sanger Institute, Cambridgeshire, UK for guiding us toward the *TRAM2* gene; and Dr Cecilia Giunta, University Children's Hospital, Zurich, Switzerland for performing collagen species analysis on our *BMP1* patient cohort.

I also must thank Bart Wagner, Sheffield Teaching Hospitals for helping to keep me enthusiastic about sitting in a darkened room on a Friday afternoon and sharing his vast knowledge of electron microscopy.

Finally, I give thanks to my family for having an on-going interest in my research and their unwavering support.

xiii ABBREVIATIONS

5'UTR – 5' Untranslated Region
AD – Autosomal-Dominant
ADAMTS – A Disintegrin and Metalloproteinase with Thrombospondin Motifs
ALF – Acute liver failure
ALP – Alkaline Phosphatase
ALPL – Alkaline Phosphatase, Liver/Bone/Kidney
Annol – A Personalis bioinformatics tool for variant annotation
AR – Autosomal-Recessive
ARG – Arginine
ARID1A – Gene encoding the AT-rich interactive domain-containing protein 1A
ARID1B – Gene encoding the AT-rich interactive domain-containing protein 1B
Array CGH – Array comparative genomic hybridization
ASP – Aspartic Acid or Aspartate
ATF6 – Activating transcription factor 6
B3GALT3 – Beta-1,3-N-Acetylgalactosaminyltransferase
B3GALT6 – Beta-1,3-Galactosyltransferase 6;
B4GALT7 – Beta-1,4-Galactosyltransferase 7
BGN – Biglycan
BiP – Binding Immunoglobulin Protein, also known as GRP78 and HSPA5.
BMD – Bone Mineral Density
BMP – Bone Morphogenetic Protein
BMP1 – Bone morphogenic protein 1
BMP2 – Bone morphogenic protein 2
BMP6 – Bone morphogenic protein 6
BRIL – Bone-Restricted Ifitm-Like Protein
BS – Bruck Syndrome
BSP – Bone sialoprotein
BWA - Burrows-Wheeler Aligner
BWA-MEM - a BWA algorithm recommended for high quality sequence queries
C – Carboxy
CCD – Cleidocranial dysplasia
CCS – Cole Carpenter Syndrome
cDNA – Complementary DNA
CHOP – CCAAT/enhancer binding protein
CHST14 – Carbohydrate Sulfotransferase 14
CNV – Copy Number Variation
COL1A1 – Type 1 collagen, alpha 1 chain gene
COL1A2 – Type 1 collagen, alpha 2 chain gene
COL6A3 – Alpha-3 chain of type VI collagen
COMP – Cartilage Oligomeric Matrix Protein
CREB – cAMP response element-binding protein
CREB3L1 – CAMP Responsive Element Binding Protein 3 Like 1
CREBBP – Creb-binding protein
CRTAP – Cartilage-Associated Protein
CSR – Calcium sensing receptor
CUB domain - An evolutionarily conserved protein domain of approximately 110 residues
CyPB – Cyclophilin B
DCN – Decorin
DDD – Deciphering Developmental Disorders
DECIPHER – DatabasE of genomIc variation and Phenotype in Humans using Ensembl Resources

DEL – Deletion
DEXA – Dual-energy X-ray absorptiometry
DI – Dentinogenesis Imperfecta
DLX5 – Distal-less homeobox 5
DMP1 – Dentine matrix protein
DNA – Deoxyribonucleic Acid
ECM – Extracellular Matrix
EDS – Ehlers-Danlos Syndrome
EDTA – Ethylenediaminetetraacetic acid
EM – Electron Microscopy
EP300 – Histone acetyltransferase p300
ER – Endoplasmic Reticulum
ERAD – ER associated degradation
FAM20C – Family With Sequence Similarity 20, Member C
FEVR – Familial exudative vitreoretinopathy
FGF – Fibroblast growth factor
FGFR – FGF/Fibroblast growth factor receptor
FHS – Floating Harbour Syndrome
FKBP10 – FK506-binding protein 10
FKBP14 – FK506 binding protein 14
FKBP65 – 65 kDa FK506 Binding Protein 10
FZD – Frizzled transmembrane receptor family
GAG – Glycosaminoglycan
GATK – Genome Analysis Tool Kit
GC-RICH – Guanine-Cytosine Rich
gDNA – Genomic DNA
GI – Gastrointestinal
Gly – Glycine
gnomAD – Genome Aggregation Database
GO – Gene Ontology
HGMD – Human Gene Mutation Database
HGVS – Human Genome Variation Society
HPO – Human Phenotype Ontology
HSP47 – Heat Shock Protein 47
IFITM5 – Interferon Induced Transmembrane Protein 5
IGF – Insulin-like growth factor
IGV – Integrative Genomics Viewer
IMPC – International Mouse Phenotyping Consortium, www.mousephenotype.org
Indel – Insertion-Deletion variant
IRE1 – Inositol-requiring enzyme 1
LAP – Latency Associated Protein
LH2 – Lysyl Hydroxylase 2
LOD – Logarithmic Odds Score
LOF – Loss of Function
LOVD – Leiden Open Variation Database
LOX – Lysyl Oxidase
LOXL3 – Lysyl Oxidase Like 3
LRP4 – Low-density lipoprotein receptor-related protein 4
LRP5 – Lipoprotein Receptor-Related Protein 5
LRP6 – Low-density lipoprotein receptor-related protein 6
MAF – Minor Allele Frequency
MBTPS1 – Membrane-Bound Transcription Factor Protease, Site 1, known as S1P
MBTPS2 – Membrane-Bound Transcription Factor Protease, Site 2, known as S2P

M-CSF – Macrophage Colony Stimulation Factor
MGP – Marix Gla Protein
MINT - Molecular INTeraction database
MLBR – Major Ligand Binding Region
MMP – Matrix Metalloproteinase
MOWS – Mowat Wilson Syndrome
mRNA – Messenger RNA
MSC – Mesenchymal Stem Cell
mTORC1 – Mammalian Target Of Rapamycin Complex 1
MYH6 – Myosin, Heavy Chain 6, Cardiac Muscle, Alpha
N – Amino
NBAS – Neuroblastoma Amplified Sequence gene
NCBI – National Centre for Biotechnology Information
NGS – Next Generation Sequencing
NMD – mRNA Nonsense Mediated Decay
OASIS – Old Astrocyte Specifically Induced Substance Protein encoded by *CREB3L1*
OCN – Osteocalcin
OI – Osteogenesis Imperfecta
OMIM – Online Mendelian Inheritance in Man
OPG – Osteoprotegerin
OPPG – Osteoporosis Pseudoglioma Syndrome
OSX – Osterix
P3H1 – Prolyl 3-Hydroxylase 1
P4HB – Procollagen-Proline, 2-Oxoglutarate-4-Dioxygenase, Beta Subunit
PCOLCE – Procollagen C-endopeptidase enhancer 1
PCR – Polymerase Chain Reaction
PDI – Protein Disulfide Isomerase
PEDF – Pigment Epithelium-Derived Factor
PERK – Protein Kinase R(PKR)-like ER kinase
PHA – Pelger Huet Anomaly
PLOD1 – Procollagen-Lysine,2-Oxoglutarate 5-Dioxygenase 1
PLOD2 – Procollagen-Lysine,2-Oxoglutarate 5-Dioxygenase 2
PLS3 – Plastin 3
POSTN – Gene encoding the Periostin protein.
PPIase – Peptidyl-Prolyl Isomerase
PPIB – Peptidyl-Prolyl Cis-Trans Isomerase B
PTC – Premature Termination codon
PTH1 – Parathyroid hormone 1
PTH1R – Parathyroid Hormone 1 Receptor
RANKL – Receptor Activator of NF- κ B Ligand
Reactome - (a database of reactions, pathways and biological processes
rER – Rough Endoplasmic Reticulum
RIDD – Regulated IRE1-dependent decay
RNA – Ribonucleic Acid
RUNX2 – Run-related Transcription Factor 2
Sanger MGP – Sanger Mouse Genetics Project
SCN11A Sodium Voltage-Gated Channel Alpha Subunit 11
SCN8A – Sodium Voltage-Gated Channel Alpha Subunit 8
SCN9A – Sodium voltage-gated channel alpha subunit 9
SDGS – Sheffield Diagnostic Genetics Service
SDS-PAGE – Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SEC24D – Sec24-Related Gene Family, Member D
SERPINF1 – Serpin Peptidase Inhibitor, Clade F, Member 1

SERPINH1 – Serpin Peptidase Inhibitor, Clade H, Member 1
SKIV2L – Superkiller Viralicidic Activity 2, *S. Cerevisiae*, Homolog-Like
SLC38A10 – Sodium-Coupled Neutral Amino Acid Transporter 10 also known as Solute Carrier Family 38, Member 10
SLRP – Small Leucine Rich Proteoglycans
SMAD – SMA and MAD Related Protein
SNARE – Soluble N-Ethylmaleimide-Sensitive Factor Activating Protein Receptor Protein
SNP – Single Nucleotide Polymorphism
SNV – Single Nucleotide Variant
SOPH – Short Stature, Optic Nerve And Pelguer Huet Anomaly Syndrome
SOST – Gene Encoding The Sclerostin Protein
SP7 – Gene encoding the Osterix protein
SPARC – Secreted Proteins Acidic Rich in Cysteine, also known as osteonectin
SRCAP – Snf2-Related Cbp Activator Protein
SREBP – Sterol Regulatory Element Binding Protein
SULF1 – Sulfatase 1
SULF2 – Sulfatase 2
TAPBP – Transporter Associated With Antigen Processing (TAP) Binding Protein
TAPT1 – Transmembrane Anterior Posterior Transformation 1 protein
TCA – The Citric Acid Cycle
TGF- β 1 – Transforming Growth Factor-Beta1
TMEM38B – Trimeric Intracellular Cation Channel Type B
TRAM2 – Translocation Associated Membrane Protein 2
TRIC – Monovalent Cation-Specific Transmembrane Channel
UGGT1 – UDP-Glucose Glycoprotein Glucosyltransferase 1
UKGTN – UK Genetic Testing Network
UPR – Unfolded Protein Response
VCF – Variant Call Format, text file
VQSR - Variant Quality Score Recalibration
WES – Whole Exome Sequencing
WGS – Whole Genome Sequencing
Wnt – Wingless
XYLT1 – Xylosyltransferase 1
XYLT2 – Xylosyltransferase 2
ZEB2 – Zinc Finger E-Box Binding Homeobox 2

xiv LIST OF TABLES AND FIGURES

Figure 1 Bone cross-section diagram. (Pbroks13, CC BY 3.0).....	3
Figure 2 Bone remodelling.....	5
Figure 3 The role of FGFR1-3 in osteoblast differentiation.....	7
Figure 4 Electron microscopy of longitudinal cross section of dermal collagen.	8
Figure 5 The structure of type I collagen.	10
Figure 6 Steps in Type I Collagen Biosynthesis.	11
Figure 7 Gradient-dependent Severity Model for Glycine Substitutions in the Helical Domain of Type I collagen	14
Figure 8 Different clinical phenotypes are linked to mutations at specific positions along the type I procollagen protein.	15
Figure 9 Regulation of bone formation and mineralization in osteogenesis imperfecta	21
Figure 10 The unfolded protein response (UPR).	30
Figure 11 TGF β signalling in bone.....	32
Figure 12 Study Workflow.....	41
Figure 13 Comparison of SureSelect ^{XT} and Nextera Library Capture Methods	42
Figure 14 SureSelect Custom Probe Design Workflow	44
Figure 15 Gel image of DNA samples following library preparation.....	45
Figure 16 Scanned image of individual patient DNA following library preparation.....	45
Figure 17 Tapestation analysis of DNA library preparation.....	47
Figure 18 Electrophoresis of PCR products for the TRAM2 gene	52
Figure 19 Depth of sequence coverage across all chromosomes for patient E using SureSelect Human All Exons V6r2.....	56
Figure 20 Variant prioritisation Strategy 1.....	58
Figure 21 Variant Prioritisation Strategy 2	59
Figure 22 Variant prioritisation strategy 3	60
Figure 23 Electron microscopy of med-reticular dermis (20,000X magnification)	64
Figure 24 Visualisation of sequence coverage for SureSelect and Nextera methodologies	67
Figure 25 Comparison of the number of probes designed by SureSelect and Nextera	68
Figure 26 Sanger sequencing confirmation of <i>BMP1</i> mutations.	70
Figure 27 Segregation analysis for the <i>COL1A1</i> c.4070T>C:p.(Leu1357Pro) variant in family 004.	71
Figure 28 Chain-chain recognition sequences for type I and III collagen molecules	72
Figure 29 Targeted exome results for patient 015	74
Figure 30 Patient 015: X-Ray of lower leg and sequential BMD measurements.	75
Figure 31 SDS-PAGE of the pepsinized procollagens in the medium.	77
Figure 32 SDS-PAGE of the untreated procollagen in the presence of 0.05M dithiodthreitol, (DTT).....	77
Figure 33 Normal control collagen and elastic fibres	78
Figure 34 Normal control dermal fibroblast.....	79
Figure 35 Ultra structure of dermal fibroblast from a patient with OI type III with a c.3269G>A,p.(Gly1090Asp) <i>COL1A2</i> pathogenic mutation.....	79
Figure 36 Ultra structure of collagen fibrils in reticular dermis from patient 015	80
Figure 37 Ultrastructure of dermal fibroblasts from patient 015.....	80

Figure 38 Ultra structure of collagen and fibroblast from reticular dermis of Patient J	81
Figure 39 Ultra structure of collagen and fibroblast from reticular dermis of Patient K	82
Figure 40 Male and Female weight curves for the <i>Tram2</i> ^{-/-} mouse model.	84
Figure 41 Measure bone mineral density in <i>Tram2</i> ^{-/-} knockout mouse using the DEXA analyser (g/cm ²).....	84
Figure 42 X-ray of adult mouse knee joint at 16wks.....	85
Figure 43 Sanger Sequence of two common variants in the <i>TRAM2</i> gene identified in our cohort.....	86
Figure 44 Sanger sequence confirmation of <i>de-novo</i> c.9029C>A,p.(Pro3010His) <i>SRCAP</i> variant in patient B.....	89
Figure 45 Locations of Floating Harbour Syndrome-Causing Mutations within <i>SRCAP</i>	90
Figure 46 Hierarchical clustering of methylation status.	91
Figure 47 Methylated CpG island at chr6:33282867-33283184.....	92
Figure 48 An example of a hypermethylation region plot.	93
Figure 49 An example of a hypomethylation region plot.....	93
Figure 50 Methylation pattern for the <i>PCOLCE</i> gene region (chr7 100199826-100199982).....	94
Figure 51 Methylation pattern for the <i>LOXL3</i> gene region (chr2 74775965-74775979).	94
Figure 52 Downstream targets of <i>BMP2</i>	96
Figure 53 Ultra structure of collagen and fibroblast from reticular dermis of Patient B.....	96
Figure 54 Proteoglycan wrapping around collagen fibres in dermis from patient B (20000X).....	97
Figure 55 Transport of antigenic peptides across the ER.	98
Figure 56 Sanger sequence confirmation of heterozygous <i>NBAS</i> variants in the parents of patient C.	100
Figure 57 Western blot of cultured fibroblasts showing reduced <i>NBAS</i> protein levels compared to controls.	101
Figure 58 Peripheral blood film showing the presence of Pelger Huët anomaly in patient C.	101
Figure 59 Electron microscopy of collagen fibrils in patient C	102
Figure 60 Electron microscopy of collagen fibrils showing occasional collagen flowers in patient C (~200nm in diameter).....	102
Figure 61 Electron microscopy of dermal fibroblasts from patient C showing markedly expanded protein filled ER.....	103
Figure 62 Electron microscopy of dermal fibroblast from patient C showing the presence of protein aggregates within the expanded ER.	103
Figure 63 Electron microscopy of dermal fibroblast from patient C showing electron dense bodies within lysosomes.....	104
Figure 64 Distribution of variants identified in patients D-H using whole exome sequencing.....	108
Figure 65 Sanger sequence of exon 9 of the <i>P4HB</i> gene in patient E.	111
Figure 66 Schematic representation of the <i>P4HB</i> protein and the coding exons of <i>P4HB</i> cDNA (NM_000918.3).	112
Figure 67 Protein structure model of <i>P4HB</i> indicating potential impact of the p.Tyr393Cys mutation in patient E.....	113
Figure 68 Electron microscopic ultra structure of collagen in patient E	114
Figure 69 <i>P3H1</i> c.1080+1G>T splicing mutation in patient G identified by targeted exome sequencing. ...	116
Figure 70 <i>P3H1</i> c.1224-80G>A variant identified by whole exome sequencing.....	116
Figure 71 <i>P3H1</i> c.1244-80G>A splice site prediction analysis	117
Figure 72 Predicted effect on splicing of the c.1229+1G>A <i>COL1A1</i> pathogenic mutation identified in patient 149.	124

Figure 73 Multiplex Ligation-dependent Probe Amplification analysis of genomic DNA for the <i>PLOD1</i> gene.....	126
Figure 74 The synthesis of heparin sulphate and chondroitin/dermatan sulfate proteoglycans.	127
Figure 75 The predicted structure of the SLC38A10 protein.....	130
Figure 76 Glutamine signalling in osteoblasts in response to Wnt signalling.....	131
Figure 77 Simplified Wnt/ β -catenin signaling.....	133
Figure 78 Discrimination of folding status of glycoprotein by folding sensor UDP-glucose:glycoprotein glucosyltransferase (UGGT) in the glycoprotein quality control system.....	137
Figure 79 <i>In-silico</i> splice prediction of heterozygous <i>POSTN</i> c.1108+2T>C variant identified in patient 141.	139
Figure 80 Electron microscopic ultra structure of dermal collagen in an OI patient with a c.838G>A, p.(Gly280Ser) <i>COL1A2</i> mutation.....	144
Figure 81 Energy Metabolism of Osteoblasts.....	149
Figure 82 Flow chart showing the OBCD bone phenotyping platform.	153
Table 1 Hormones, growth factors and cytokines that influence bone metabolism and remodelling.	6
Table 2 Summary of Genes Associated with OI and their numerical classification.....	27
Table 3 Classification of Hypophosphatasia	36
Table 4 Location, size and reference sequence for genes in the Custom Probe Design.....	43
Table 5 Example of a completed probe design for SureSelect library preparation.	44
Table 6 Completed probe design for part of the <i>SERPINF1</i> and <i>BMP1</i> genes for Nextera Library preparation	46
Table 7 Versions of open source software used in the bioinformatics pipeline (v3.1.3).....	49
Table 8 Primers designed for Sanger sequencing of the <i>TRAM2</i> gene.....	51
Table 9 Reagents used for <i>TRAM2</i> amplification by PCR	51
Table 10 PCR Cycling Conditions for <i>TRAM2</i> amplification by PCR.....	52
Table 11 Reagents and cycling condition for Sanger Sequencing of <i>TRAM2</i> PCR products	53
Table 12 Data fields returned by the Deciphering Developmental Delay study for Complimentary Analysis Projects.	54
Table 13 Source of terms used to assess the function of candidate genes.....	54
Table 14 Quality Control filter settings use in analysis of Personalis sequence data	55
Table 15 HPO terms used to generate an extended target gene list used in variant prioritisation.....	60
Table 16 Expected diameter of collagen fibres in each life stage	64
Table 17 Clinical Details provided for recruited individuals.....	66
Table 18 Comparison of regions without adequate sequence coverage for SureSelect vs Nextera library preparation in five control DNA samples.....	67
Table 19 List of genes and number of variants used to validate SureSelect targeted exome panel.....	68
Table 20 Variants identified in participants 001-018 using targeted exome analysis.....	69
Table 21 Variants identified in the <i>TRAM2</i> gene in participants 001-018 and A-H	86
Table 22 Variants identified in patient D-H.	109
Table 23 Clinical details, genetic results and family structure for DDD CAP 12 patients with a clinical diagnosis of Osteogenesis Imperfecta.....	120
Table 24 Candidate variants in CAP12 DDD patients following manual assessment of DDD supplied candidate list and interrogation of full VCF for patients 101-149.	123

Table 25 HPO data model developed by the 100K Genomes Project Musculoskeletal GeCIP, Bone
Fragility SubDomain 152

xv REFERENCES

Acevedo, A. C., J. A. Poulter, P. G. Alves, C. L. de Lima, L. C. Castro, P. M. Yamaguti, L. M. Paula, D. A. Parry, C. V. Logan, C. E. Smith, C. A. Johnson, C. F. Inglehearn and A. J. Mighell (2015). "Variability of systemic and oro-dental phenotype in two families with non-lethal Raine syndrome with FAM20C mutations." BMC Med Genet **16**: 8.

Alanay, Y., H. Avaygan, N. Camacho, G. E. Utine, K. Boduroglu, D. Aktas, M. Alikasifoglu, E. Tuncbilek, D. Orhan, F. T. Bakar, B. Zabel, A. Superti-Furga, L. Bruckner-Tuderman, C. J. Curry, S. Pyott, P. H. Byers, D. R. Eyre, D. Baldrige, B. Lee, A. E. Merrill, E. C. Davis, D. H. Cohn, N. Akarsu and D. Krakow (2010). Mutations in the gene encoding the RER protein FKBP65 cause autosomal-recessive osteogenesis imperfecta. Am J Hum Genet. **86**: 551-559.

Alazami, A. M., N. Patel, H. E. Shamseldin, S. Anazi, M. S. Al-Dosari, F. Alzahrani, H. Hijazi, M. Alshammari, M. A. Aldahmesh, M. A. Salih, E. Faqeih, A. Alhashem, F. A. Bashiri, M. Al-Owain, A. Y. Kentab, S. Sogaty, S. Al Tala, M. H. Temsah, M. Tulbah, R. F. Aljelaify, S. A. Alshahwan, M. Z. Seidahmed, A. A. Alhadid, H. Aldhalaan, F. AlQallaf, W. Kurdi, M. Alfadhel, Z. Babay, M. Alsogheer, N. Kaya, Z. N. Al-Hassnan, G. M. Abdel-Salam, N. Al-Sannaa, F. Al Mutairi, H. Y. El Khashab, S. Bohlega, X. Jia, H. C. Nguyen, R. Hammami, N. Adly, J. Y. Mohamed, F. Abdulwahab, N. Ibrahim, E. A. Naim, B. Al-Younes, B. F. Meyer, M. Hashem, R. Shaheen, Y. Xiong, M. Abouelhoda, A. A. Aldeeri, D. M. Monies and F. S. Alkuraya (2015). "Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families." Cell Rep **10**(2): 148-161.

Alzahrani, F., S. A. Al Hazzaa, H. Tayeb and F. S. Alkuraya (2015). "LOXL3, encoding lysyl oxidase-like 3, is mutated in a family with autosomal recessive Stickler syndrome." Hum Genet **134**(4): 451-453.

Aoki, T., S. Ichimura, A. Itoh, M. Kuramoto, T. Shinkawa, T. Isobe and M. Tagaya (2009). "Identification of the neuroblastoma-amplified gene product as a component of the syntaxin 18 complex implicated in Golgi-to-endoplasmic reticulum retrograde transport." Mol Biol Cell **20**(11): 2639-2649.

Asharani, P. V., K. Keupp, O. Semler, W. Wang, Y. Li, H. Thiele, G. Yigit, E. Pohl, J. Becker, P. Frommolt, C. Sonntag, J. Altmüller, K. Zimmermann, D. S. Greenspan, N. A. Akarsu, C. Netzer, E. Schönau, R. Wirth, M. Hammerschmidt, P. Nürnberg, B. Wollnik and T. J. Carney (2012). "Attenuated BMP1 function compromises osteogenesis, leading to bone fragility in humans and zebrafish." Am J Hum Genet **90**(4): 661-674.

Ayadi, A., M. C. Birling, J. Bottomley, J. Bussell, H. Fuchs, M. Fray, V. Gailus-Durner, S. Greenaway, R. Houghton, N. Karp, S. Leblanc, C. Lengger, H. Maier, A. M. Mallon, S. Marschall, D. Melvin, H. Morgan, G. Pavlovic, E. Ryder, W. C. Skarnes, M. Selloum, R. Ramirez-Solis, T. Sorg, L. Teboul, L. Vasseur, A. Walling, T. Weaver, S. Wells, J. K. White, A. Bradley, D. J. Adams, K. P. Steel, M. Hrabě de Angelis, S. D. Brown and Y. Hérault (2012). "Mouse large-scale phenotyping initiatives: overview of the European Mouse Disease Clinic (EUMODIC) and of the Wellcome Trust Sanger Institute Mouse Genetics Project." Mamm Genome **23**(9-10): 600-610.

Back, S. H., N. S. Adapala, M. F. Barbe, N. C. Carpino, A. Y. Tsygankov and A. Sanjay (2013). "TULA-2, a novel histidine phosphatase, regulates bone remodeling by modulating osteoclast function." Cell Mol Life Sci **70**(7): 1269-1284.

Balasubramanian, M., J. Hurst, S. Brown, N. J. Bishop, P. Arundel, C. DeVile, R. C. Pollitt, L. Crooks, D. Longman, J. F. Caceres, F. Shackley, S. Connolly, J. H. Payne, A. C. Offiah, D. Hughes, M. J. Parker, W. Hide, T. M. Skerry and D. Study (2017). "Compound heterozygous variants in NBAS as a cause of atypical osteogenesis imperfecta." Bone **94**: 65-74.

Balasubramanian, M., R. Padidela, R. C. Pollitt, N. J. Bishop, M. Z. Mughal, A. C. Offiah, B. E. Wagner, J. McCaughey and D. J. Stephens (2017). "P4HB recurrent missense mutation causing Cole-Carpenter syndrome." J Med Genet **55**:158-165.

Balasubramanian, M., R. C. Pollitt, K. E. Chandler, M. Z. Mughal, M. J. Parker, A. Dalton, P. Arundel, A. C. Offiah and N. J. Bishop (2015). "CRTAP mutation in a patient with Cole-Carpenter syndrome." Am J Med Genet A **167A**(3): 587-591.

Balasubramanian, M., G. J. Sobey, B. E. Wagner, L. C. Peres, J. Bowen, J. Bexon, M. K. Javaid, P. Arundel and N. J. Bishop (2016). "Osteogenesis imperfecta: Ultrastructural and histological findings on examination of skin revealing novel insights into genotype-phenotype correlation." Ultrastruct Pathol **40**(2): 71-76.

Baldrige, D., U. Schwarze, R. Morello, J. Lenington, T. K. Bertin, J. M. Pace, M. G. Pepin, M. Weis, D. R. Eyre, J. Walsh, D. Lambert, A. Green, H. Robinson, M. Michelson, G. Houge, C. Lindman, J. Martin, J. Ward, E. Lemyre, J. J. Mitchell, D. Krakow, D. L. Rimoim, D. H. Cohn, P. H. Byers and B. Lee (2008). "CRTAP and LEPRE1 mutations in recessive osteogenesis imperfecta." Hum Mutat **29**(12): 1435-1442.

Barnes, A. M., E. M. Carter, W. A. Cabral, M. Weis, W. Chang, E. Makareeva, S. Leikin, C. N. Rotimi, D. R. Eyre, C. L. Raggio and J. C. Marini (2010). Lack of cyclophilin B in osteogenesis imperfecta with normal collagen folding. N Engl J Med. **362**: 521-528.

Barnes, A. M., W. Chang, R. Morello, W. A. Cabral, M. Weis, D. R. Eyre, S. Leikin, E. Makareeva, N. Kuznetsova, T. E. Uveges, A. Ashok, A. W. Flor, J. J. Mulvihill, P. L. Wilson, U. T. Sundaram, B. Lee and J. C. Marini (2006). Deficiency of cartilage-associated protein in recessive lethal osteogenesis imperfecta. N Engl J Med. **355**: 2757-2764.

Barnes, A. M., G. Duncan, M. Weis, W. Paton, W. A. Cabral, E. L. Mertz, E. Makareeva, M. J. Gambello, F. L. Lacbawan, S. Leikin, A. Fertala, D. R. Eyre, S. J. Bale and J. C. Marini (2013). "Kuskokwim syndrome, a recessive congenital contracture disorder, extends the phenotype of FKBP10 mutations." Hum Mutat **34**(9): 1279-1288.

Baron, R. and M. Kneissel (2013). "WNT signaling in bone homeostasis and disease: from human mutations to treatments." Nat Med **19**(2): 179-192.

Bassett, J. H., A. Gogakos, J. K. White, H. Evans, R. M. Jacques, A. H. van der Spek, R. Ramirez-Solis, E. Ryder, D. Sunter, A. Boyde, M. J. Campbell, P. I. Croucher, G. R. Williams and S. M. G. Project (2012). "Rapid-throughput skeletal phenotyping of 100 knockout mice identifies 9 new genes that determine bone strength." PLoS Genet **8**(8): e1002858.

Beederman, M. J. D. Lamplot, G. Nan, J. Wang, X. Liu, L. Yin, R. Li, W. Shui, H. Zhang, S. H. Kim, W. Zhang, J. Zhang, Y. Kong, S. Denduluri, M. R. Rogers, A. Pratt,

R. C. Haydon, H. H. Luu, J. Angeles, L. L. Shi, L. L. and He, T. C. (2013). "BMP signaling in mesenchymal stem cell differentiation and bone formation." J Biomed Sci Eng **6**(8A): 32-52.

Belkadi, A., A. Bolze, Y. Itan, A. Cobat, Q. B. Vincent, A. Antipenko, L. Shang, B. Boisson, J. L. Casanova and L. Abel (2015). "Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants." Proc Natl Acad Sci U S A **112**(17): 5473-5478.

Bennett, D. L. and C. G. Woods (2014). "Painful and painless channelopathies." Lancet Neurol **13**(6): 587-599.

Berman, A. G., J. M. Wallace, Z. R. Bart and M. R. Allen (2016). "Raloxifene reduces skeletal fractures in an animal model of osteogenesis imperfecta." Matrix Biol **52-54**: 19-28.

Besio, R., G. Iula, N. Garibaldi, L. Cipolla, S. Sabbioneda, M. Biggiogera, J. C. Marini, A. Rossi and A. Forlino (2018). "4-PBA ameliorates cellular homeostasis in fibroblasts from osteogenesis imperfecta patients by enhancing autophagy and stimulating protein secretion." Biochim Biophys Acta **1864**(5 Pt A): 1642-1652.

Bi, X., I. Grafe, H. Ding, R. Flores, E. Munivez, M. M. Jiang, B. Dawson, B. Lee and C. G. Ambrose (2017). "Correlations Between Bone Mechanical Properties and Bone Composition Parameters in Mouse Models of Dominant and Recessive Osteogenesis Imperfecta and the Response to Anti-TGF- β Treatment." J Bone Miner Res **32**(2): 347-359.

Bianchi, L., A. Gagliardi, S. Maruelli, R. Besio, C. Landi, R. Gioia, K. M. Kozloff, B. M. Khoury, P. J. Coucke, S. Symoens, J. C. Marini, A. Rossi, L. Bini and A. Forlino (2015). "Altered cytoskeletal organization characterized lethal but not surviving *Brl+/-* mice: insight on phenotypic variability in osteogenesis imperfecta." Hum Mol Genet **24**(21): 6118-6133.

Bidart, M., M. El Atifi, S. Miladi, J. Rendu, V. Satre, P. F. Ray, C. Bosson, F. Devillard, D. Lehalle, V. Malan, J. Amiel, M. A. Mencarelli, M. Baldassarri, A. Renieri, J. Clayton-Smith, G. Vieville, J. Thevenon, F. Amblard, F. Berger, P. S. Jouk and C. Coutton (2017). "Microduplication of the ARID1A gene causes intellectual disability with recognizable syndromic features." Genet Med **19**(6): 701-710.

Bishop, N. (2016). "Bone Material Properties in Osteogenesis Imperfecta." J Bone Miner Res **31**(4): 699-708.

Bodian, D. L., T. F. Chan, A. Poon, U. Schwarze, K. Yang, P. H. Byers, P. Y. Kwok and T. E. Klein (2009). "Mutation and polymorphism spectrum in osteogenesis imperfecta type II: implications for genotype-phenotype relationships." Hum Mol Genet **18**(3): 463-471.

Bodian, D. L., B. Madhan, B. Brodsky and T. E. Klein (2008). "Predicting the clinical lethality of osteogenesis imperfecta from collagen glycine mutations." Biochemistry **47**(19): 5424-5432.

Bogan, R., R. C. Riddle, Z. Li, S. Kumar, A. Nandal, M. C. Faugere, A. Boskey, S. E. Crawford and T. L. Clemens (2013). "A mouse model for human osteogenesis imperfecta type VI." J Bone Miner Res **28**(7): 1531-1536.

- Boot-Handford, R. P. and M. D. Briggs (2010). "The unfolded protein response and its relevance to connective tissue diseases." Cell Tissue Res **339**(1): 197-211.
- Bourhis, J. M., N. Mariano, Y. Zhao, K. Harlos, J. Y. Exposito, E. Y. Jones, C. Moali, N. Aghajari and D. J. Hulmes (2012). "Structural basis of fibrillar collagen trimerization and related genetic disorders." Nat Struct Mol Biol **19**(10): 1031-1036.
- Brady, A. F., S. Demirdas, S. Fournel-Gigleux, N. Ghali, C. Giunta, I. Kapferer-Seebacher, T. Kosho, R. Mendoza-Londono, M. F. Pope, M. Rohrbach, T. Van Damme, A. Vandersteen, C. van Mourik, N. Voermans, J. Zschocke and F. Malfait (2017). "The Ehlers-Danlos syndromes, rare types." Am J Med Genet C Semin Med Genet **175**(1): 70-115.
- Breslau-Siderius, E. J., R. H. Engelbert, G. Pals and J. A. van der Sluijs (1998). "Bruck syndrome: a rare combination of bone fragility and multiple congenital joint contractures." J Pediatr Orthop B **7**(1): 35-38.
- Broer, S. (2014). "The SLC38 family of sodium-amino acid co-transporters." Pflugers Arch **466**(1): 155-172.
- Brown, P. M., J. D. Hutchison and J. C. Crockett (2011). "Absence of glutamine supplementation prevents differentiation of murine calvarial osteoblasts to a mineralizing phenotype." Calcif Tissue Int **89**(6): 472-482.
- Bruderer, M., R. G. Richards, M. Alini and M. J. Stoddart (2014). "Role and regulation of RUNX2 in osteogenesis." Eur Cell Mater **28**: 269-286.
- Cabral, W. A., A. M. Barnes, A. Adeyemo, K. Cushing, D. Chitayat, F. D. Porter, S. R. Panny, F. Gulamali-Majid, S. A. Tishkoff, T. R. Rebbeck, S. M. Gueye, J. E. Bailey-Wilson, L. C. Brody, C. N. Rotimi and J. C. Marini (2012). "A founder mutation in LEPRE1 carried by 1.5% of West Africans and 0.4% of African Americans causes lethal recessive osteogenesis imperfecta." Genet Med **14**(5): 543-551.
- Cabral, W. A., W. Chang, A. M. Barnes, M. Weis, M. A. Scott, S. Leikin, E. Makareeva, N. V. Kuznetsova, K. N. Rosenbaum, C. J. Tiffit, D. I. Bulas, C. Kozma, P. A. Smith, D. R. Eyre and J. C. Marini (2007). Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. Nat Genet **39**: 359-365.
- Cabral, W. A., M. Ishikawa, M. Garten, E. N. Makareeva, B. M. Sargent, M. Weis, A. M. Barnes, E. A. Webb, N. J. Shaw, L. Ala-Kokko, F. L. Lacbawan, W. Hogler, S. Leikin, P. S. Blank, J. Zimmerberg, D. R. Eyre, Y. Yamada and J. C. Marini (2016). "Absence of the ER Cation Channel TMEM38B/TRIC-B Disrupts Intracellular Calcium Homeostasis and Dysregulates Collagen Synthesis in Recessive Osteogenesis Imperfecta." PLoS Genet **12**(7): e1006156.
- Cabral, W. A., E. Makareeva, A. Colige, A. D. Letocha, J. M. Ty, H. N. Yeowell, G. Pals, S. Leikin and J. C. Marini (2005). "Mutations near amino end of alpha1(I) collagen cause combined osteogenesis imperfecta/Ehlers-Danlos syndrome by interference with N-propeptide processing." J Biol Chem **280**(19): 19259-19269.
- Canty, E. G. and K. E. Kadler (2005). "Procollagen trafficking, processing and fibrillogenesis." J Cell Sci **118**(Pt 7): 1341-1353.

Caparros-Martin, J. A., M. S. Aglan, S. Temtamy, G. A. Otaify, M. Valencia, J. Nevado, E. Vallespin, A. Del Pozo, C. Prior de Castro, L. Calatrava-Ferrerias, P. Gutierrez, A. M. Bueno, B. Sagastizabal, E. Guillen-Navarro, M. Ballesta-Martinez, V. Gonzalez, S. Y. Basaran, R. Buyukoglan, B. Sarikepe, C. Espinoza-Valdez, F. Cammarata-Scalisi, V. Martinez-Glez, K. E. Heath, P. Lapunzina and V. L. Ruiz-Perez (2017). "Molecular spectrum and differential diagnosis in patients referred with sporadic or autosomal recessive osteogenesis imperfecta." Mol Genet Genomic Med **5**(1): 28-39.

Caplan, A. I. and D. Correa (2011). "PDGF in bone formation and regeneration: new insights into a novel mechanism involving MSCs." J Orthop Res **29**(12): 1795-1803.

Capo-Chichi, J. M., C. Mehawej, V. Delague, C. Caillaud, I. Khneisser, F. F. Hamdan, J. L. Michaud, Z. Kibar and A. Mégarbané (2015). "Neuroblastoma Amplified Sequence (NBAS) mutation in recurrent acute liver failure: Confirmatory report in a sibship with very early onset, osteoporosis and developmental delay." Eur J Med Genet **58**(12): 637-641.

Celli, J. and R. M. Tsohis (2015). "Bacteria, the endoplasmic reticulum and the unfolded protein response: friends or foes?" Nat Rev Microbiol **13**(2): 71-82.

Chang, W., A. M. Barnes, W. A. Cabral, J. N. Bodurtha and J. C. Marini (2010). Prolyl 3-hydroxylase 1 and CRTAP are mutually stabilizing in the endoplasmic reticulum collagen prolyl 3-hydroxylation complex. Hum Mol Genet. England. **19**: 223-234.

Chen, G., I. Nakamura, R. Dhanasekaran, E. Iguchi, E. J. Tolosa, P. A. Romecin, R. E. Vera, L. L. Almada, A. G. Miamen, R. Chaiteerakij, M. Zhou, M. K. Asiedu, C. D. Moser, S. Han, C. Hu, B. A. Banini, A. M. Oseini, Y. Chen, Y. Fang, D. Yang, H. M. Shaleh, S. Wang, D. Wu, T. Song, J. S. Lee, S. S. Thorgeirsson, E. Chevet, V. H. Shah, M. E. Fernandez-Zapico and L. R. Roberts (2017). "Transcriptional Induction of Periostin by a Sulfatase 2-TGF β 1-SMAD Signaling Axis Mediates Tumor Angiogenesis in Hepatocellular Carcinoma." Cancer Res **77**(3): 632-645.

Chen, Q., B. Denard, C. E. Lee, S. Han, J. S. Ye and J. Ye (2016). "Inverting the Topology of a Transmembrane Protein by Regulating the Translocation of the First Transmembrane Helix." Mol Cell **63**(4): 567-578.

Chen, Q., C. E. Lee, B. Denard and J. Ye (2014). "Sustained induction of collagen synthesis by TGF-beta requires regulated intramembrane proteolysis of CREB3L1." PLoS One **9**(10): e108528.

Chessler, S. D. and P. H. Byers (1993). "BiP binds type I procollagen pro alpha chains with mutations in the carboxyl-terminal propeptide synthesized by cells from patients with osteogenesis imperfecta." J Biol Chem **268**(24): 18226-18233.

Chessler, S. D., G. A. Wallis and P. H. Byers (1993). "Mutations in the carboxyl-terminal propeptide of the pro alpha 1(I) chain of type I collagen result in defective chain association and produce lethal osteogenesis imperfecta." J Biol Chem **268**(24): 18218-18225.

Cho, T. J., K. E. Lee, S. K. Lee, S. J. Song, K. J. Kim, D. Jeon, G. Lee, H. N. Kim, H. R. Lee, H. H. Eom, Z. H. Lee, O. H. Kim, W. Y. Park, S. S. Park, S. Ikegawa, W. J. Yoo, I. H. Choi and J. W. Kim (2012). A single recurrent mutation in the 5'-UTR of IFITM5 causes osteogenesis imperfecta type V. Am J Hum Genet. **91**: 343-348.

Cho, T. J., H. J. Moon, D. Y. Cho, M. S. Park, D. Y. Lee, W. J. Yoo, C. Y. Chung and I. H. Choi (2008). "The c.3040C > T mutation in COL1A1 is recurrent in Korean patients with infantile cortical hyperostosis (Caffey disease)." J Hum Genet **53**(10): 947-949.

Christiansen, H. E., U. Schwarze, S. M. Pyott, A. AlSwaid, M. Al Balwi, S. Alrasheed, M. G. Pepin, M. A. Weis, D. R. Eyre and P. H. Byers (2010). Homozygosity for a missense mutation in SERPINH1, which encodes the collagen chaperone protein HSP47, results in severe recessive osteogenesis imperfecta. Am J Hum Genet. **86**: 389-398.

Conover, C. A. (2008). "Insulin-like growth factor-binding proteins and bone metabolism." Am J Physiol Endocrinol Metab **294**(1): E10-14.

Conrad, D. F., J. E. Keebler, M. A. DePristo, S. J. Lindsay, Y. Zhang, F. Casals, Y. Idaghdour, C. L. Hartl, C. Torroja, K. V. Garimella, M. Zilversmit, R. Cartwright, G. A. Rouleau, M. Daly, E. A. Stone, M. E. Hurles, P. Awadalla and G. Project (2011). "Variation in genome-wide mutation rates within and between human families." Nat Genet **43**(7): 712-714.

Conway, S. J., K. Izuhara, Y. Kudo, J. Litvin, R. Markwald, G. Ouyang, J. R. Arron, C. T. Holweg and A. Kudo (2014). "The role of periostin in tissue remodeling across health and disease." Cell Mol Life Sci **71**(7): 1279-1288.

Cosman, F., D. B. Crittenden, J. D. Adachi, N. Binkley, E. Czerwinski, S. Ferrari, L. C. Hofbauer, E. Lau, E. M. Lewiecki, A. Miyauchi, C. A. Zerbin, C. E. Milmont, L. Chen, J. Maddox, P. D. Meisner, C. Libanati and A. Grauer (2016). "Romosozumab Treatment in Postmenopausal Women with Osteoporosis." N Engl J Med **375**(16): 1532-1543.

Cui, J., J. Xiao, V. S. Tagliabracci, J. Wen, M. Rahdar and J. E. Dixon (2015). "A secretory kinase complex regulates extracellular protein phosphorylation." Elife **4**: e06120.

Delany, A. M., M. Amling, M. Priemel, C. Howe, R. Baron and E. Canalis (2000). "Osteopenia and decreased bone formation in osteonectin-deficient mice." J Clin Invest **105**(9): 1325.

Duran, I., L. Nevarez, A. Sarukhanov, S. Wu, K. Lee, P. Krejci, M. Weis, D. Eyre, D. Krakow and D. H. Cohn (2015). "HSP47 and FKBP65 cooperate in the synthesis of type I procollagen." Hum Mol Genet **24**(7): 1918-1928.

El-Gharbawy, A. H., J. N. Peeden, R. S. Lachman, J. M. Graham, S. R. Moore and D. L. Rimoin (2010). "Severe cleidocranial dysplasia and hypophosphatasia in a child with microdeletion of the C-terminal region of RUNX2." Am J Med Genet A **152A**(1): 169-174.

Eller-Vainicher, C., A. Bassotti, A. Imeraj, E. Cairoli, F. M. Olivieri, F. Cortini, M. Dubini, B. Marinelli, A. Spada and I. Chiodini (2016). "Bone involvement in adult patients affected with Ehlers-Danlos syndrome." Osteoporos Int **27**(8): 2525-2531.

Fabre, A., B. Charroux, C. Martinez-Vinson, B. Roquelaure, E. Odul, E. Sayar, H. Smith, V. Colomb, N. Andre, J. P. Hugot, O. Goulet, C. Lacoste, J. Sarles, J. Royet, N. Levy and C. Badens (2012). "SKIV2L mutations cause syndromic diarrhea, or trichohepatoenteric syndrome." Am J Hum Genet **90**(4): 689-692.

Fahiminiya, S., J. Majewski, H. Al-Jallad, P. Moffatt, J. Mort, F. H. Glorieux, P. Roschger, K. Klaushofer and F. Rauch (2014). "Osteoporosis Caused by Mutations in PLS3 - Clinical and Bone Tissue Characteristics." J Bone Miner Res **29**(8): 1805-1814.

Fahiminiya, S., J. Majewski, J. Mort, P. Moffatt, F. H. Glorieux and F. Rauch (2013). "Mutations in WNT1 are a cause of osteogenesis imperfecta." J Med Genet **50**(5): 345-348.

Fang, M., E. L. Goldstein, A. S. Turner, C. M. Les, B. G. Orr, G. J. Fisher, K. B. Welch, E. D. Rothman and M. M. Banaszak Holl (2012). "Type I collagen D-spacing in fibril bundles of dermis, tendon, and bone: bridging between nano- and micro-level tissue hierarchy." ACS Nano **6**(11): 9503-9514.

Farber, C. R., A. Reich, A. M. Barnes, P. Becerra, F. Rauch, W. A. Cabral, A. Bae, A. Quinlan, F. H. Glorieux, T. L. Clemens and J. C. Marini (2014). "A novel IFITM5 mutation in severe atypical osteogenesis imperfecta type VI impairs osteoblast production of pigment epithelium-derived factor." J Bone Miner Res **29**(6): 1402-1411.

Ferrari, S. L., S. Deutsch and S. E. Antonarakis (2005). "Pathogenic mutations and polymorphisms in the lipoprotein receptor-related protein 5 reveal a new biological pathway for the control of bone mass." Curr Opin Lipidol **16**(2): 207-214.

Fiscaletti, M., A. Biggin, B. Bennetts, K. Wong, J. Briody, V. Pacey, C. Birman and C. F. Munns (2018). "Novel variant in Sp7/Osx associated with recessive osteogenesis imperfecta with bone fragility and hearing impairment." Bone **110**: 66-75.

Florencio-Silva, R., G. R. Sasso, E. Sasso-Cerri, M. J. Simoes and P. S. Cerri (2015). "Biology of Bone Tissue: Structure, Function, and Factors That Influence Bone Cells." Biomed Res Int **2015**: 421746.

Folkestad, L., J. D. Hald, V. Canudas-Romo, J. Gram, A. P. Hermann, B. Langdahl, B. Abrahamsen and K. Brixen (2016). "Mortality and Causes of Death in Patients With Osteogenesis Imperfecta: A Register-Based Nationwide Cohort Study." J Bone Miner Res **31**(12): 2159-2166.

Forlino, A., W. A. Cabral, A. M. Barnes and J. C. Marini (2011). "New perspectives on osteogenesis imperfecta." Nat Rev Endocrinol **7**(9): 540-557.

Forlino, A. and J. C. Marini (2016). "Osteogenesis imperfecta." Lancet **387**(10028): 1657-1671.

Fradin, M., C. Stoetzel, J. Muller, M. Koob, D. Christmann, C. Debry, M. Kohler, M. Isnard, D. Astruc, P. Desprez, C. Zorres, E. Flori, H. Dollfus and B. Doray (2011). "Osteosclerotic bone dysplasia in siblings with a Fam20C mutation." Clin Genet **80**(2): 177-183.

Freudenthal, B., J. Logan, P. I. Croucher, G. R. Williams, J. H. Bassett and S. I. M. Pipelines (2016). "Rapid phenotyping of knockout mice to identify genetic determinants of bone strength." J Endocrinol **231**(1): R31-46.

Fujikawa, Y., J. M. Quinn, A. Sabokbar, J. O. McGee and N. A. Athanasou (1996). "The human osteoclast precursor circulates in the monocyte fraction." Endocrinology **137**(9): 4058-4060.

Garbes, L., K. Kim, A. Rieß, H. Hoyer-Kuhn, F. Beleggia, A. Bevot, M. J. Kim, Y. H. Huh, H. S. Kweon, R. Savarirayan, D. Amor, P. M. Kakadia, T. Lindig, K. O. Kagan, J. Becker, S. A. Boyadjiev, B. Wollnik, O. Semler, S. K. Bohlander, J. Kim and C. Netzer (2015). "Mutations in SEC24D, Encoding a Component of the COPII Machinery, Cause a Syndromic Form of Osteogenesis Imperfecta." Am J Hum Genet **96**(3): 432-439.

Gerundino, F., G. Marseglia, C. Pescucci, E. Pelo, M. Benelli, C. Giachini, B. Federighi, C. Antonelli and F. Torricelli (2014). "16p11.2 de novo microdeletion encompassing SRCAP gene in a patient with speech impairment, global developmental delay and behavioural problems." Eur J Med Genet **57**(11-12): 649-653.

Ghoumid, J., L. Drevillon, S. M. Alavi-Naini, N. Bondurand, M. Rio, A. Briand-Suleau, M. Nasser, L. Goodwin, P. Raymond, C. Yanicostas, M. Goossens, S. Lyonnet, D. Mowat, J. Amiel, N. Soussi-Yanicostas and I. Giurgea (2013). "ZEB2 zinc-finger missense mutations lead to hypomorphic alleles and a mild Mowat-Wilson syndrome." Hum Mol Genet **22**(13): 2652-2661.

Giunta, C., M. Baumann, C. Fauth, U. Lindert, E. M. Abdalla, A. F. Brady, J. Collins, J. Dastgir, S. Donkervoort, N. Ghali, D. S. Johnson, A. Kariminejad, J. Koch, M. Kraenzlin, N. Lahiri, B. Lozic, A. Y. Manzur, J. E. V. Morton, J. Pilch, R. C. Pollitt, G. Schreiber, N. L. Shannon, G. Sobey, A. Vandersteen, F. S. van Dijk, M. Witsch-Baumgartner, J. Zschocke, F. M. Pope, C. G. Bonnemann and M. Rohrbach (2018). "A cohort of 17 patients with kyphoscoliotic Ehlers-Danlos syndrome caused by biallelic mutations in FKBP14: expansion of the clinical and mutational spectrum and description of the natural history." Genet Med **20**(1): 42-54.

Glorieux, F. H., F. Rauch, H. Plotkin, L. Ward, R. Travers, P. Roughley, L. Lalic, D. F. Glorieux, F. Fassier and N. J. Bishop (2000). "Type V osteogenesis imperfecta: a new form of brittle bone disease." J Bone Miner Res **15**(9): 1650-1658.

Glorieux, F. H., L. M. Ward, F. Rauch, L. Lalic, P. J. Roughley and R. Travers (2002). "Osteogenesis imperfecta type VI: a form of brittle bone disease with a mineralization defect." J Bone Miner Res **17**(1): 30-38.

Gordon, J. A., J. L. Stein, J. J. Westendorf and A. J. van Wijnen (2015). "Chromatin modifiers and histone modifications in bone formation, regeneration, and therapeutic intervention for bone-related disease." Bone **81**: 739-745.

Grafe, I., T. Yang, S. Alexander, E. P. Homan, C. Lietman, M. M. Jiang, T. Bertin, E. Munivez, Y. Chen, B. Dawson, Y. Ishikawa, M. A. Weis, T. K. Sampath, C. Ambrose, D. Eyre, H. P. Bächinger and B. Lee (2014). "Excessive transforming growth factor- β signaling is a common mechanism in osteogenesis imperfecta." Nat Med **20**(6): 670-675.

Grgurevic, L., B. Macek, M. Mercep, M. Jelic, T. Smoljanovic, I. Erjavec, I. Dumic-Cule, S. Prgomet, D. Durdevic, D. Vnuk, M. Lipar, M. Stejskal, V. Kufner, J. Brkljacic, D. Maticic and S. Vukicevic (2011). "Bone morphogenetic protein (BMP)1-3 enhances bone repair." Biochem Biophys Res Commun **408**(1): 25-31.

Gusella, J. F., N. S. Wexler, P. M. Conneally, S. L. Naylor, M. A. Anderson, R. E. Tanzi, P. C. Watkins, K. Ottina, M. R. Wallace and A. Y. Sakaguchi (1983). "A polymorphic DNA marker genetically linked to Huntington's disease." Nature **306**(5940): 234-238.

Götherström, C., M. Westgren, S. W. Shaw, E. Aström, A. Biswas, P. H. Byers, C. N. Mattar, G. E. Graham, J. Taslimi, U. Ewald, N. M. Fisk, A. E. Yeoh, J. L. Lin, P. J. Cheng, M. Choolani, K. Le Blanc and J. K. Chan (2014). "Pre- and postnatal transplantation of fetal mesenchymal stem cells in osteogenesis imperfecta: a two-center experience." Stem Cells Transl Med **3**(2): 255-264.

Ha-Vinh, R., Y. Alanay, R. A. Bank, A. B. Campos-Xavier, A. Zankl, A. Superti-Furga and L. Bonafe (2004). "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): 115-120.

Haack, T. B., C. Staufner, M. G. Kopke, B. K. Straub, S. Kolker, C. Thiel, P. Freisinger, I. Baric, P. J. McKiernan, N. Dikow, I. Harting, F. Beisse, P. Burgard, U. Kotzaeridou, J. Kuhr, U. Himbert, R. W. Taylor, F. Distelmaier, J. Vockley, L. Ghaloul-Gonzalez, J. Zschocke, L. S. Kremer, E. Graf, T. Schwarzmayr, D. M. Bader, J. Gagneur, T. Wieland, C. Terrile, T. M. Strom, T. Meitinger, G. F. Hoffmann and H. Prokisch (2015). "Biallelic Mutations in NBAS Cause Recurrent Acute Liver Failure with Onset in Infancy." Am J Hum Genet **97**(1): 163-169.

Hartigan, N., L. Garrigue-Antar and K. E. Kadler (2003). "Bone morphogenetic protein-1 (BMP-1). Identification of the minimal domain structure for procollagen C-proteinase activity." J Biol Chem **278**(20): 18045-18049.

He, X., M. Semenov, K. Tamai and X. Zeng (2004). "LDL receptor-related proteins 5 and 6 in Wnt/beta-catenin signaling: arrows point the way." Development **131**(8): 1663-1677.

Hediger, M. A., B. Cléménçon, R. E. Burrier and E. A. Bruford (2013). "The ABCs of membrane transporters in health and disease (SLC series): introduction." Mol Aspects Med **34**(2-3): 95-107.

Heikkinen, J., B. Pousi, M. Pope and R. Myllylä (1999). "A null-mutated lysyl hydroxylase gene in a compound heterozygote British patient with Ehlers-Danlos syndrome type VI." Hum Mutat **14**(4): 351.

Hellsten, S. V., M. G. Häggglund, M. M. Eriksson and R. Fredriksson (2017). "The neuronal and astrocytic protein SLC38A10 transports glutamine, glutamate, and aspartate, suggesting a role in neurotransmission." FEBS Open Bio **7**(6): 730-746.

Hessle, L., K. A. Johnson, H. C. Anderson, S. Narisawa, A. Sali, J. W. Goding, R. Terkeltaub and J. L. Millan (2002). "Tissue-nonspecific alkaline phosphatase and plasma cell membrane glycoprotein-1 are central antagonistic regulators of bone mineralization." Proc Natl Acad Sci U S A **99**(14): 9445-9449.

Homan, E. P., F. Rauch, I. Grafe, C. Lietman, J. A. Doll, B. Dawson, T. Bertin, D. Napierala, R. Morello, R. Gibbs, L. White, R. Miki, D. H. Cohn, S. Crawford, R. Travers, F. H. Glorieux and B. Lee (2011). "Mutations in SERPIN1 cause osteogenesis imperfecta type VI." J Bone Miner Res **26**(12): 2798-2803.

Hood, R. L., M. A. Lines, S. M. Nikkel, J. Schwartzentruber, C. Beaulieu, M. J. Nowaczyk, J. Allanson, C. A. Kim, D. Wiczorek, J. S. Moilanen, D. Lacombe, G. Gillessen-Kaesbach, M. L. Whiteford, C. R. Quaid, I. Gomy, D. R. Bertola, B. Albrecht, K. Platzer, G. McGillivray, R. Zou, D. R. McLeod, A. E. Chudley, B. N. Chodirker, J. Marcadier, J. Majewski, D. E. Bulman, S. M. White, K. M. Boycott and F. C. Consortium

- (2012). "Mutations in SRCAP, encoding SNF2-related CREBBP activator protein, cause Floating-Harbor syndrome." Am J Hum Genet **90**(2): 308-313.
- Hood, R. L., L. C. Schenkel, S. M. Nikkel, P. J. Ainsworth, G. Pare, K. M. Boycott, D. E. Bulman and B. Sadikovic (2016). "The defining DNA methylation signature of Floating-Harbor Syndrome." Sci Rep **6**: 38803.
- Hoyer, J., A. B. Ekici, S. Ende, B. Popp, C. Zweier, A. Wiesener, E. Wohlleber, A. Dufke, E. Rossier, C. Petsch, M. Zweier, I. Göhring, A. M. Zink, G. Rappold, E. Schröck, D. Wiczorek, O. Riess, H. Engels, A. Rauch and A. Reis (2012). "Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability." Am J Hum Genet **90**(3): 565-572.
- Hoyer-Kuhn, H., J. Franklin, G. Allo, M. Kron, C. Netzer, P. Eysel, B. Hero, E. Schoenau and O. Semler (2016). "Safety and efficacy of denosumab in children with osteogenesis imperfecta—a first prospective trial." J Musculoskelet Neuronal Interact **16**(1): 24-32.
- Huang, R. L., Y. Yuan, J. Tu, G. M. Zou and Q. Li (2014). "Opposing TNF- α /IL-1 β - and BMP-2-activated MAPK signaling pathways converge on Runx2 to regulate BMP-2-induced osteoblastic differentiation." Cell Death Dis **5**: e1187.
- Hyry, M., J. Lantto and J. Myllyharju (2009). "Missense mutations that cause Bruck syndrome affect enzymatic activity, folding, and oligomerization of lysyl hydroxylase 2." J Biol Chem **284**(45): 30917-30924.
- Ishikawa, Y., J. Vranka, J. Wirz, K. Nagata and H. P. Bächinger (2008). "The rough endoplasmic reticulum-resident FK506-binding protein FKBP65 is a molecular chaperone that interacts with collagens." J Biol Chem **283**(46): 31584-31590.
- Ishikawa, Y., J. Wirz, J. A. Vranka, K. Nagata and H. P. Bächinger (2009). "Biochemical characterization of the prolyl 3-hydroxylase 1 cartilage-associated protein cyclophilin B complex." J Biol Chem **284**(26): 17641-17647.
- Ito, S. and K. Nagata (2017). "Biology of Hsp47 (Serpine H1), a collagen-specific molecular chaperone." Semin Cell Dev Biol **62**: 142-151.
- Izumi, M., R. Kuruma, R. Okamoto, A. Seko, Y. Ito and Y. Kajihara (2017). "Substrate Recognition of Glycoprotein Folding Sensor UGGT Analyzed by Site-Specifically (¹⁵N)-Labeled Glycopeptide and Small Glycopeptide Library Prepared by Parallel Native Chemical Ligation." J Am Chem Soc **139**(33): 11421-11426.
- Jaffe, A. E., P. Murakami, H. Lee, J. T. Leek, M. D. Fallin, A. P. Feinberg and R. A. Irizarry (2012). "Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies." Int J Epidemiol **41**(1): 200-209.
- Joeng, K. S., Y. C. Lee, J. Lim, Y. Chen, M. M. Jiang, E. Munivez, C. Ambrose and B. H. Lee (2017). "Osteocyte-specific WNT1 regulates osteoblast function during bone homeostasis." J Clin Invest **127**(7): 2678-2688.
- Kalamajski, S. and A. Oldberg (2010). "The role of small leucine-rich proteoglycans in collagen fibrillogenesis." Matrix Biol **29**(4): 248-253.

Karner, C. M., E. Esen, A. L. Okunade, B. W. Patterson and F. Long (2015). "Increased glutamine catabolism mediates bone anabolism in response to WNT signaling." J Clin Invest **125**(2): 551-562.

Karner, C. M. and F. Long (2017). "Wnt signaling and cellular metabolism in osteoblasts." Cell Mol Life Sci **74**(9): 1649-1657.

Keller, R. B., T. T. Tran, S. M. Pyott, M. G. Pepin, R. Savarirayan, G. McGillivray, D. A. Nickerson, M. J. Bamshad and P. H. Byers (2018). "Monoallelic and biallelic CREB3L1 variant causes mild and severe osteogenesis imperfecta, respectively." Genet Med **20**(4): 411-419.

Kelley, B. P., F. Malfait, L. Bonafe, D. Baldrige, E. Homan, S. Symoens, A. Willaert, N. Elcioglu, L. Van Maldergem, C. Verellen-Dumoulin, Y. Gillerot, D. Napierala, D. Krakow, P. Beighton, A. Superti-Furga, A. De Paepe and B. Lee (2011). "Mutations in FKBP10 cause recessive osteogenesis imperfecta and Bruck syndrome." J Bone Miner Res **26**(3): 666-672.

Keupp, K., F. Beleggia, H. Kayserili, A. M. Barnes, M. Steiner, O. Semler, B. Fischer, G. Yigit, C. Y. Janda, J. Becker, S. Breer, U. Altunoglu, J. Grunhagen, P. Krawitz, J. Hecht, T. Schinke, E. Makareeva, E. Lausch, T. Cankaya, J. A. Caparros-Martin, P. Lapunzina, S. Temtamy, M. Aglan, B. Zabel, P. Eysel, F. Koerber, S. Leikin, K. C. Garcia, C. Netzer, E. Schonau, V. L. Ruiz-Perez, S. Mundlos, M. Amling, U. Kornak, J. Marini and B. Wollnik (2013). "Mutations in WNT1 cause different forms of bone fragility." Am J Hum Genet **92**(4): 565-574.

Khoshnoodi, J., J. P. Cartailier, K. Alvares, A. Veis and B. G. Hudson (2006). "Molecular recognition in the assembly of collagens: terminal noncollagenous domains are key recognition modules in the formation of triple helical protomers." J Biol Chem **281**(50): 38117-38121.

Kii, I., T. Nishiyama and A. Kudo (2016). "Periostin promotes secretion of fibronectin from the endoplasmic reticulum." Biochem Biophys Res Commun **470**(4): 888-893.

Kii, I., T. Nishiyama, M. Li, K. Matsumoto, M. Saito, N. Amizuka and A. Kudo (2010). "Incorporation of tenascin-C into the extracellular matrix by periostin underlies an extracellular meshwork architecture." J Biol Chem **285**(3): 2028-2039.

Kokubu, C., U. Heinzmann, T. Kokubu, N. Sakai, T. Kubota, M. Kawai, M. B. Wahl, J. Galceran, R. Grosschedl, K. Ozono and K. Imai (2004). "Skeletal defects in ringelschwanz mutant mice reveal that Lrp6 is required for proper somitogenesis and osteogenesis." Development **131**(21): 5469-5480.

Kortum, F., I. Marquardt, M. Alawi, G. C. Korenke, S. Spranger, P. Meinecke and K. Kutsche (2017). "Acute Liver Failure Meets SOPH Syndrome: A Case Report on an Intermediate Phenotype." Pediatrics **139**(1): e20160550.

Kosho, T., N. Okamoto and C.-S. S. I. Collaborators (2014). "Genotype-phenotype correlation of Coffin-Siris syndrome caused by mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A." Am J Med Genet C Semin Med Genet **166C**(3): 262-275.

Kudo, A. and I. Kii (2018). "Periostin function in communication with extracellular matrices." J Cell Commun Signal **12**(1): 301-308.

Laine, C. M., K. S. Joeng, P. M. Campeau, R. Kiviranta, K. Tarkkonen, M. Grover, J. T. Lu, M. Pekkinen, M. Wessman, T. J. Heino, V. Nieminen-Pihala, M. Aronen, T. Laine, H. Kröger, W. G. Cole, A. E. Lehesjoki, L. Nevarez, D. Krakow, C. J. Curry, D. H. Cohn, R. A. Gibbs, B. H. Lee and O. Mäkitie (2013). "WNT1 mutations in early-onset osteoporosis and osteogenesis imperfecta." N Engl J Med **368**(19): 1809-1816.

Laine, C. M., M. Wessman, S. Toiviainen-Salo, M. A. Kaunisto, M. K. Mäyränpää, T. Laine, M. Pekkinen, H. Kröger, V. V. Välimäki, M. J. Välimäki, A. E. Lehesjoki and O. Mäkitie (2015). "A novel splice mutation in PLS3 causes X-linked early onset low-turnover osteoporosis." J Bone Miner Res **30**(3): 510-518.

Lamoureux, F., M. Baud'huin, L. Duplomb, D. Heymann and F. Rédini (2007). "Proteoglycans: key partners in bone cell biology." Bioessays **29**(8): 758-771.

Lampe, A. K. and K. M. Bushby (2005). "Collagen VI related muscle disorders." J Med Genet **42**(9): 673-685.

Lapunzina, P., M. Aglan, S. Temtamy, J. A. Caparros-Martin, M. Valencia, R. Leton, V. Martinez-Glez, R. Elhossini, K. Amr, N. Vilaboia and V. L. Ruiz-Perez (2010). Identification of a frameshift mutation in Osterix in a patient with recessive osteogenesis imperfecta. Am J Hum Genet. **87**: 110-114.

Larsen, J., G. L. Carvill, E. Gardella, G. Kluger, G. Schmiedel, N. Barisic, C. Depienne, E. Brilstra, Y. Mang, J. E. Nielsen, M. Kirkpatrick, D. Goudie, R. Goldman, J. A. Jähn, B. Jepsen, D. Gill, M. Döcker, S. Biskup, J. M. McMahon, B. Koeleman, M. Harris, K. Braun, C. G. de Kovel, C. Marini, N. Specchio, T. Djémié, S. Weckhuysen, N. Tommerup, M. Troncoso, L. Troncoso, A. Bevot, M. Wolff, H. Hjalgrim, R. Guerrini, I. E. Scheffer, H. C. Mefford, R. S. Møller and E. R. C. CRP (2015). "The phenotypic spectrum of SCN8A encephalopathy." Neurology **84**(5): 480-489.

Lavery, K., P. Swain, D. Falb and M. H. Alaoui-Ismaili (2008). "BMP-2/4 and BMP-6/7 differentially utilize cell surface receptors to induce osteoblastic differentiation of human bone marrow-derived mesenchymal stem cells." J Biol Chem **283**(30): 20948-20958.

Lazarus, S., A. M. McInerney-Leo, F. A. McKenzie, G. Baynam, S. Broley, B. V. Cavan, C. F. Munns, J. E. Pruijs, D. Sillence, P. A. Terhal, K. Pryce, M. A. Brown, A. Zankl, G. Thomas and E. L. Duncan (2014). "The IFITM5 mutation c.-14C > T results in an elongated transcript expressed in human bone; and causes varying phenotypic severity of osteogenesis imperfecta type V." BMC Musculoskelet Disord **15**: 107.

Le Blanc, K., C. Götherström, O. Ringdén, M. Hassan, R. McMahon, E. Horwitz, G. Anneren, O. Axelsson, J. Nunn, U. Ewald, S. Nordén-Lindeberg, M. Jansson, A. Dalton, E. Aström and M. Westgren (2005). "Fetal mesenchymal stem-cell engraftment in bone after in utero transplantation in a patient with severe osteogenesis imperfecta." Transplantation **79**(11): 1607-1614.

Le Goff, C., R. P. Somerville, F. Kesteloot, K. Powell, D. E. Birk, A. C. Colige and S. S. Apte (2006). "Regulation of procollagen amino-propeptide processing during mouse embryogenesis by specialization of homologous ADAMTS proteases: insights on collagen biosynthesis and dermatosparaxis." Development **133**(8): 1587-1596.

Lee, W. C., A. R. Guntur, F. Long and C. J. Rosen (2017). "Energy Metabolism of the Osteoblast: Implications for Osteoporosis." Endocr Rev **38**(3): 255-266.

Li, J. Q., Y. L. Qiu, J. Y. Gong, L. M. Dou, Y. Lu, A. S. Knisely, M. H. Zhang, W. S. Luan and J. S. Wang (2017). "Novel NBAS mutations and fever-related recurrent acute liver failure in Chinese children: a retrospective study." BMC Gastroenterol **17**(1): 77.

Lim, J., I. Grafe, S. Alexander and B. Lee (2017). "Genetic causes and mechanisms of Osteogenesis Imperfecta." Bone **102**: 40-49.

Lindahl, K., A. M. Barnes, N. Fratzi-Zelman, M. P. Whyte, T. E. Hefferan, E. Makareeva, M. Brusel, M. J. Yaszemski, C. J. Rubin, A. Kindmark, P. Roschger, K. Klaushofer, W. H. McAlister, S. Mumm, S. Leikin, E. Kessler, A. L. Boskey, O. Ljunggren and J. C. Marini (2011). "COL1 C-propeptide cleavage site mutations cause high bone mass osteogenesis imperfecta." Hum Mutat **32**(6): 598-609.

Lindahl, K., A. Kindmark, N. Laxman, E. Åström, C. J. Rubin and Ö. Ljunggren (2013). "Allele dependent silencing of collagen type I using small interfering RNAs targeting 3'UTR Indels - a novel therapeutic approach in osteogenesis imperfecta." Int J Med Sci **10**(10): 1333-1343.

Lindert, U., W. A. Cabral, S. Ausavarat, S. Tongkobpetch, K. Ludin, A. M. Barnes, P. Yeetong, M. Weis, B. Krabichler, C. Srichomthong, E. N. Makareeva, A. R. Janecke, S. Leikin, B. Röthlisberger, M. Rohrbach, I. Kennerknecht, D. R. Eyre, K. Suphapeetiporn, C. Giunta, J. C. Marini and V. Shotelersuk (2016). "MBTPS2 mutations cause defective regulated intramembrane proteolysis in X-linked osteogenesis imperfecta." Nat Commun **7**: 11920.

Lisse, T. S., F. Thiele, H. Fuchs, W. Hans, G. K. Przemeck, K. Abe, B. Rathkolb, L. Quintanilla-Martinez, G. Hoelzlwimmer, M. Helfrich, E. Wolf, S. H. Ralston and M. Hrabé de Angelis (2008). "ER stress-mediated apoptosis in a new mouse model of osteogenesis imperfecta." PLoS Genet **4**(2): e7.

Little, R. D., J. P. Carulli, R. G. Del Mastro, J. Dupuis, M. Osborne, C. Folz, S. P. Manning, P. M. Swain, S. C. Zhao, B. Eustace, M. M. Lappe, L. Spitzer, S. Zweier, K. Braunschweiger, Y. Benchekroun, X. Hu, R. Adair, L. Chee, M. G. FitzGerald, C. Tulig, A. Caruso, N. Tzellas, A. Bawa, B. Franklin, S. McGuire, X. Noguez, G. Gong, K. M. Allen, A. Anisowicz, A. J. Morales, P. T. Lomedico, S. M. Recker, P. Van Eerdewegh, R. R. Recker and M. L. Johnson (2002). "A mutation in the LDL receptor-related protein 5 gene results in the autosomal dominant high-bone-mass trait." Am J Hum Genet **70**(1): 11-19.

Longman, D., N. Hug, M. Keith, C. Anastasaki, E. E. Patton, G. Grimes and J. F. Caceres (2013). "DHX34 and NBAS form part of an autoregulatory NMD circuit that regulates endogenous RNA targets in human cells, zebrafish and *Caenorhabditis elegans*." Nucleic Acids Res **41**(17): 8319-8331.

Lund, A. M., E. Aström, S. Söderhäll, M. Schwartz and F. Skovby (1999). "Osteogenesis imperfecta: mosaicism and refinement of the genotype-phenotype map in OI type III. Mutations in brief no. 242. Online." Hum Mutat **13**(6): 503.

Lynch, M. (2010). "Rate, molecular spectrum, and consequences of human mutation." Proc Natl Acad Sci U S A **107**(3): 961-968.

MacArthur, D. G., S. Balasubramanian, A. Frankish, N. Huang, J. Morris, K. Walter, L. Jostins, L. Habegger, J. K. Pickrell, S. B. Montgomery, C. A. Albers, Z. D. Zhang, D. F. Conrad, G. Lunter, H. Zheng, Q. Ayub, M. A. DePristo, E. Banks, M. Hu, R. E.

Handsaker, J. A. Rosenfeld, M. Fromer, M. Jin, X. J. Mu, E. Khurana, K. Ye, M. Kay, G. I. Saunders, M. M. Suner, T. Hunt, I. H. Barnes, C. Amid, D. R. Carvalho-Silva, A. H. Bignell, C. Snow, B. Yngvadottir, S. Bumpstead, D. N. Cooper, Y. Xue, I. G. Romero, J. Wang, Y. Li, R. A. Gibbs, S. A. McCarroll, E. T. Dermitzakis, J. K. Pritchard, J. C. Barrett, J. Harrow, M. E. Hurles, M. B. Gerstein, C. Tyler-Smith and G. P. Consortium (2012). "A systematic survey of loss-of-function variants in human protein-coding genes." Science **335**(6070): 823-828.

Maiers, J. L., E. Kostallari, M. Mushref, T. M. deAssuncao, H. Li, N. Jalan-Sakrikar, R. C. Huebert, S. Cao, H. Malhi and V. H. Shah (2017). "The unfolded protein response mediates fibrogenesis and collagen I secretion through regulating TANGO1 in mice." Hepatology **65**(3): 983-998.

Makareeva, E., N. A. Aviles and S. Leikin (2011). "Chaperoning osteogenesis: new protein-folding disease paradigms." Trends Cell Biol **21**(3): 168-176.

Maksimova, N., K. Hara, I. Nikolaeva, T. Chun-Feng, T. Usui, M. Takagi, Y. Nishihira, A. Miyashita, H. Fujiwara, T. Oyama, A. Nogovicina, A. Sukhomyasova, S. Potapova, R. Kuwano, H. Takahashi, M. Nishizawa and O. Onodera (2010). "Neuroblastoma amplified sequence gene is associated with a novel short stature syndrome characterised by optic nerve atrophy and Pelger-Huët anomaly." J Med Genet **47**(8): 538-548.

Malfait, F., S. Symoens, J. De Backer, T. Hermanns-Le, N. Sakalihasan, C. M. Lapiere, P. Coucke and A. De Paepe (2007). "Three arginine to cysteine substitutions in the pro-alpha (I)-collagen chain cause Ehlers-Danlos syndrome with a propensity to arterial rupture in early adulthood." Hum Mutat **28**(4): 387-395.

Malfait, F., D. Syx, P. Vlumens, S. Symoens, S. Nampoothiri, T. Hermanns-Lê, L. Van Laer and A. De Paepe (2010). "Musculocontractural Ehlers-Danlos Syndrome (former EDS type VIB) and adducted thumb clubfoot syndrome (ATCS) represent a single clinical entity caused by mutations in the dermatan-4-sulfotransferase 1 encoding CHST14 gene." Hum Mutat **31**(11): 1233-1239.

Mani, A., J. Radhakrishnan, H. Wang, M. A. Mani, C. Nelson-Williams, K. S. Carew, S. Mane, H. Najmabadi, D. Wu and R. P. Lifton (2007). "LRP6 mutation in a family with early coronary disease and metabolic risk factors." Science **315**(5816): 1278-1282.

Marini, J. C., A. Forlino, H. P. Bächinger, N. J. Bishop, P. H. Byers, A. Paepe, F. Fassier, N. Fratzl-Zelman, K. M. Kozloff, D. Krakow, K. Montpetit and O. Semler (2017). "Osteogenesis imperfecta." Nat Rev Dis Primers **3**: 17052.

Marini, J. C., A. Forlino, W. A. Cabral, A. M. Barnes, J. D. San Antonio, S. Milgrom, J. C. Hyland, J. Körkkö, D. J. Prockop, A. De Paepe, P. Coucke, S. Symoens, F. H. Glorieux, P. J. Roughley, A. M. Lund, K. Kuurila-Svahn, H. Hartikka, D. H. Cohn, D. Krakow, M. Mottes, U. Schwarze, D. Chen, K. Yang, C. Kuslich, J. Troendle, R. Dalgleish and P. H. Byers (2007). "Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans." Hum Mutat **28**(3): 209-221.

Markmann, A., H. Hausser, E. Schönherr and H. Kresse (2000). "Influence of decorin expression on transforming growth factor-beta-mediated collagen gel retraction and biglycan induction." Matrix Biol **19**(7): 631-636.

- Marr, C., A. Seasman and N. Bishop (2017). "Managing the patient with osteogenesis imperfecta: a multidisciplinary approach." J Multidiscip Healthc **10**: 145-155.
- Marshall, C., J. Lopez, L. Crookes, R. C. Pollitt and M. Balasubramanian (2016). "A novel homozygous variant in SERPINH1 associated with a severe, lethal presentation of osteogenesis imperfecta with hydranencephaly." Gene **595**(1): 49-52.
- Martin, T. J. (2017). "Reflecting on Some Discoveries of 40 Years and Their Outcomes." J Bone Miner Res **32**(10): 1971-1976.
- Martinez-Glez, V., M. Valencia, J. A. Caparros-Martin, M. Aglan, S. Temtamy, J. Tenorio, V. Pulido, U. Lindert, M. Rohrbach, D. Eyre, C. Giunta, P. Lapunzina and V. L. Ruiz-Perez (2012). "Identification of a mutation causing deficient BMP1/mTLD proteolytic activity in autosomal recessive osteogenesis imperfecta." Hum Mutat **33**(2): 343-350.
- Massink, M. P., M. A. Créton, F. Spanevello, W. M. Fennis, M. S. Cune, S. M. Savelberg, I. J. Nijman, M. M. Maurice, M. J. van den Boogaard and G. van Haaften (2015). "Loss-of-Function Mutations in the WNT Co-receptor LRP6 Cause Autosomal-Dominant Oligodontia." Am J Hum Genet **97**(4): 621-626.
- Mazziotti, G., C. Dordoni, M. Doga, F. Galderisi, M. Venturini, P. Calzavara-Pinton, R. Maroldi, A. Giustina and M. Colombi (2016). "High prevalence of radiological vertebral fractures in adult patients with Ehlers-Danlos syndrome." Bone **84**: 88-92.
- Mendoza-Londono, R., S. Fahiminiya, J. Majewski, M. Tétreault, J. Nadaf, P. Kannu, E. Sochett, A. Howard, J. Stimec, L. Dupuis, P. Roschger, K. Klaushofer, T. Palomo, J. Ouellet, H. Al-Jallad, J. S. Mort, P. Moffatt, S. Boudko, H. P. Bächinger, F. Rauch and C. R. C. Consortium (2015). "Recessive Osteogenesis Imperfecta Caused by Missense Mutations in SPARC." Am J Hum Genet **96**(6): 979-985.
- Messina, G., M. T. Atterato and P. Dimitri (2016). "When chromatin organisation floats astray: the Srcap gene and Floating-Harbor syndrome." J Med Genet **53**(12): 793-797.
- Meynard, D., L. Kautz, V. Darnaud, F. Canonne-Hergaux, H. Coppin and M. P. Roth (2009). "Lack of the bone morphogenetic protein BMP6 induces massive iron overload." Nat Genet **41**(4): 478-481.
- Mikhailik, A., B. Ford, J. Keller, Y. Chen, N. Nassar and N. Carpino (2007). "A phosphatase activity of Sts-1 contributes to the suppression of TCR signaling." Mol Cell **27**(3): 486-497.
- Mirigian, L. S., E. Makareeva, E. L. Mertz, S. Omari, A. M. Roberts-Pilgrim, A. K. Oestreich, C. L. Phillips and S. Leikin (2016). "Osteoblast Malfunction Caused by Cell Stress Response to Procollagen Misfolding in $\alpha 2(I)$ -G610C Mouse Model of Osteogenesis Imperfecta." J Bone Miner Res **31**(8): 1608-1616.
- Mochida, Y., D. Parisuthiman, S. Pornprasertsuk-Damrongsri, P. Atsawasuwana, M. Sricholpech, A. L. Boskey and M. Yamauchi (2009). "Decorin modulates collagen matrix assembly and mineralization." Matrix Biol **28**(1): 44-52.
- Monroy, M. A., D. D. Ruhl, X. Xu, D. K. Granner, P. Yaciuk and J. C. Chrivia (2001). "Regulation of cAMP-responsive element-binding protein-mediated transcription by the SNF2/SWI-related protein, SRCAP." J Biol Chem **276**(44): 40721-40726.

Morello, R., T. K. Bertin, Y. Chen, J. Hicks, L. Tonachini, M. Monticone, P. Castagnola, F. Rauch, F. H. Glorieux, J. Vranka, H. P. Bachinger, J. M. Pace, U. Schwarze, P. H. Byers, M. Weis, R. J. Fernandes, D. R. Eyre, Z. Yao, B. F. Boyce and B. Lee (2006). CRTAP is required for prolyl 3- hydroxylation and mutations cause recessive osteogenesis imperfecta. Cell **127**: 291-304.

Mornet, E. (2000). Hypophosphatasia: the mutations in the tissue-nonspecific alkaline phosphatase gene. Hum Mutat **15**: 309-315.

Muir, A. M., Y. Ren, D. H. Butz, N. A. Davis, R. D. Blank, D. E. Birk, S. J. Lee, D. Rowe, J. Q. Feng and D. S. Greenspan (2014). "Induced ablation of Bmp1 and Tll1 produces osteogenesis imperfecta in mice." Hum Mol Genet **23**(12): 3085-3101.

Murakami, T., A. Saito, S. Hino, S. Kondo, S. Kanemoto, K. Chihara, H. Sekiya, K. Tsumagari, K. Ochiai, K. Yoshinaga, M. Saitoh, R. Nishimura, T. Yoneda, I. Kou, T. Furuichi, S. Ikegawa, M. Ikawa, M. Okabe, A. Wanaka and K. Imaizumi (2009). "Signalling mediated by the endoplasmic reticulum stress transducer OASIS is involved in bone formation." Nat Cell Biol **11**(10): 1205-1211.

Määttänen, P., K. Gehring, J. J. Bergeron and D. Y. Thomas (2010). "Protein quality control in the ER: the recognition of misfolded proteins." Semin Cell Dev Biol **21**(5): 500-511.

Nachman, M. W. and S. L. Crowell (2000). "Estimate of the mutation rate per nucleotide in humans." Genetics **156**(1): 297-304.

Nakamura, I., M. G. Fernandez-Barrena, M. C. Ortiz-Ruiz, L. L. Almada, C. Hu, S. F. Elsawa, L. D. Mills, P. A. Romecin, K. H. Gulaid, C. D. Moser, J. J. Han, A. Vrabel, E. A. Hanse, N. A. Akogyeram, J. H. Albrecht, S. P. Monga, S. O. Sanderson, J. Prieto, L. R. Roberts and M. E. Fernandez-Zapico (2013). "Activation of the transcription factor GLI1 by WNT signaling underlies the role of SULFATASE 2 as a regulator of tissue regeneration." J Biol Chem **288**(29): 21389-21398.

Nakashima, K., X. Zhou, G. Kunkel, Z. Zhang, J. M. Deng, R. R. Behringer and B. de Crombrughe (2002). "The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation." Cell **108**(1): 17-29.

Newman, T. N., E. Liverani, E. Ivanova, G. L. Russo, N. Carpino, D. Ganea, F. Safadi, S. P. Kunapuli and A. Y. Tsygankov (2014). "Members of the novel UBASH3/STS/TULA family of cellular regulators suppress T-cell-driven inflammatory responses in vivo." Immunol Cell Biol **92**(10): 837-850.

Nuytinck, L., M. Freund, L. Lagae, G. E. Pierard, T. Hermanns-Le and A. De Paepe (2000). Classical Ehlers-Danlos syndrome caused by a mutation in type I collagen. Am J Hum Genet **66**: 1398-1402.

Ockeloen, C. W., K. D. Khandelwal, K. Dreesen, K. U. Ludwig, R. Sullivan, I. A. L. M. van Rooij, M. Thonissen, S. Swinnen, M. Phan, F. Conte, N. Ishorst, C. Gilissen, L. RoaFuentes, M. van de Vorst, A. Henkes, M. Steehouwer, E. van Beusekom, M. Bloemen, B. Vankeirsbilck, S. Bergé, G. Hens, J. Schoenaers, V. V. Poorten, J. Roosenboom, A. Verdonck, K. Devriendt, N. Roeleveldt, S. N. Jhangiani, L. E. L. M. Vissers, J. R. Lupski, J. de Ligt, J. W. Von den Hoff, R. Pfundt, H. G. Brunner, H. Zhou, J. Dixon, E. Mangold, H. van Bokhoven, M. J. Dixon, T. Kleefstra, A. Hoischen and

- Carels, C. E. L. (2016). "Novel mutations in LRP6 highlight the role of WNT signaling in tooth agenesis." Genet Med **18**(11): 1158-1162.
- Oliver, J. E., E. M. Thompson, F. M. Pope and A. C. Nicholls (1996). "Mutation in the carboxy-terminal propeptide of the Pro alpha 1(I) chain of type I collagen in a child with severe osteogenesis imperfecta (OI type III): possible implications for protein folding." Hum Mutat **7**(4): 318-326.
- Orwoll, E. S., J. Shapiro, S. Veith, Y. Wang, J. Lapidus, C. Vanek, J. L. Reeder, T. M. Keaveny, D. C. Lee, M. A. Mullins, S. C. Nagamani and B. Lee (2014). "Evaluation of teriparatide treatment in adults with osteogenesis imperfecta." J Clin Invest **124**(2): 491-498.
- Pace, J. M., M. Atkinson, M. C. Willing, G. Wallis and P. H. Byers (2001). "Deletions and duplications of Gly-Xaa-Yaa triplet repeats in the triple helical domains of type I collagen chains disrupt helix formation and result in several types of osteogenesis imperfecta." Hum Mutat **18**(4): 319-326.
- Pace, J. M., C. D. Kuslich, M. C. Willing and P. H. Byers (2001). "Disruption of one intra-chain disulphide bond in the carboxyl-terminal propeptide of the proalpha1(I) chain of type I procollagen permits slow assembly and secretion of overmodified, but stable procollagen trimers and results in mild osteogenesis imperfecta." J Med Genet **38**(7): 443-449.
- Palomo, T., H. Al-Jallad, P. Moffatt, F. H. Glorieux, B. Lentle, P. Roschger, K. Klaushofer and F. Rauch (2014). "Skeletal characteristics associated with homozygous and heterozygous WNT1 mutations." Bone **67**: 63-70.
- Pepin, M., M. Atkinson, B. J. Starman and P. H. Byers (1997). "Strategies and outcomes of prenatal diagnosis for osteogenesis imperfecta: a review of biochemical and molecular studies completed in 129 pregnancies." Prenat Diagn **17**(6): 559-570.
- Pepin, M. G., U. Schwarze, V. Singh, M. Romana, A. Jones-Lecointe and P. H. Byers (2013). "Allelic background of LEPRE1 mutations that cause recessive forms of osteogenesis imperfecta in different populations." Mol Genet Genomic Med **1**(4): 194-205.
- Phatarakijirund, V., S. Mumm, W. H. McAlister, D. V. Novack, D. Wenkert, K. L. Clements and M. P. Whyte (2016). "Congenital insensitivity to pain: Fracturing without apparent skeletal pathobiology caused by an autosomal dominant, second mutation in SCN11A encoding voltage-gated sodium channel 1.9." Bone **84**: 289-298.
- Piubelli, C., A. Castagna, G. Marchi, M. Rizzi, F. Busti, S. Badar, M. Marchetti, M. De Gobbi, A. Roetto, L. Xumerle, E. Suku, A. Giorgetti, M. Delledonne, O. Olivieri and D. Girelli (2017). "Identification of new BMP6 pro-peptide mutations in patients with iron overload." Am J Hematol **92**(6): 562-568.
- Pochini, L., M. Scalise, M. Galluccio and C. Indiveri (2014). "Membrane transporters for the special amino acid glutamine: structure/function relationships and relevance to human health." Front Chem **2**: 61.
- Pollitt, R., R. McMahon, J. Nunn, R. Bamford, A. Afifi, N. Bishop and A. Dalton (2006). "Mutation analysis of COL1A1 and COL1A2 in patients diagnosed with osteogenesis imperfecta type I-IV." Hum Mutat **27**(7): 716.

Pollitt, R. C., V. Saraff, A. Dalton, E. A. Webb, N. J. Shaw, G. J. Sobey, M. Z. Mughal, E. Hobson, F. Ali, N. J. Bishop, P. Arundel, W. Höglér and M. Balasubramanian (2016). "Phenotypic variability in patients with osteogenesis imperfecta caused by BMP1 mutations." Am J Med Genet A **170**(12): 3150-3156.

Pregizer, S., A. Barski, C. A. Gersbach, A. J. García and B. Frenkel (2007). "Identification of novel Runx2 targets in osteoblasts: cell type-specific BMP-dependent regulation of Tram2." J Cell Biochem **102**(6): 1458-1471.

Puig-Hervás, M. T., S. Temtamy, M. Aglan, M. Valencia, V. Martínez-Glez, M. J. Ballesta-Martínez, V. López-González, A. M. Ashour, K. Amr, V. Pulido, E. Guillén-Navarro, P. Lapunzina, J. A. Caparrós-Martín and V. L. Ruiz-Perez (2012). "Mutations in PLOD2 cause autosomal-recessive connective tissue disorders within the Bruck syndrome--osteogenesis imperfecta phenotypic spectrum." Hum Mutat **33**(10): 1444-1449.

Pyott, S. M., M. G. Pepin, U. Schwarze, K. Yang, G. Smith and P. H. Byers (2011). "Recurrence of perinatal lethal osteogenesis imperfecta in sibships: parsing the risk between parental mosaicism for dominant mutations and autosomal recessive inheritance." Genet Med **13**(2): 125-130.

Pyott, S. M., T. T. Tran, D. F. Leistriz, M. G. Pepin, N. J. Mendelsohn, R. T. Temme, B. A. Fernandez, S. M. Elsayed, E. Elsobky, I. Verma, S. Nair, E. H. Turner, J. D. Smith, G. P. Jarvik and P. H. Byers (2013). "WNT1 mutations in families affected by moderately severe and progressive recessive osteogenesis imperfecta." Am J Hum Genet **92**(4): 590-597.

Qin, M., H. Hayashi, K. Oshima, T. Tahira, K. Hayashi and H. Kondo (2005). "Complexity of the genotype-phenotype correlation in familial exudative vitreoretinopathy with mutations in the LRP5 and/or FZD4 genes." Hum Mutat **26**(2): 104-112.

Rabbani, B., N. Mahdieh, K. Hosomichi, H. Nakaoka and I. Inoue (2012). "Next-generation sequencing: impact of exome sequencing in characterizing Mendelian disorders." J Hum Genet **57**(10): 621-632.

Rafaelsen, S. H., H. Raeder, A. K. Fagerheim, P. Knappskog, T. O. Carpenter, S. Johansson and R. Bjercknes (2013). "Exome sequencing reveals FAM20c mutations associated with fibroblast growth factor 23-related hypophosphatemia, dental anomalies, and ectopic calcification." J Bone Miner Res **28**(6): 1378-1385.

Rauch, F., S. Fahiminiya, J. Majewski, J. Carrot-Zhang, S. Boudko, F. Glorieux, J. S. Mort, H. P. Bächinger and P. Moffatt (2015). "Cole-Carpenter Syndrome Is Caused by a Heterozygous Missense Mutation in P4HB." Am J Hum Genet **96**(3): 425-431.

Reppschläger, K., J. Gosselin, C. A. Dangelmaier, D. H. Thomas, N. Carpino, S. E. McKenzie, S. P. Kunapuli and A. Y. Tsygankov (2016). "TULA-2 Protein Phosphatase Suppresses Activation of Syk through the GPVI Platelet Receptor for Collagen by Dephosphorylating Tyr(P)346, a Regulatory Site of Syk." J Biol Chem **291**(43): 22427-22441.

Riddle, R. C., C. R. Diegel, J. M. Leslie, K. K. Van Koeveering, M. C. Faugere, T. L. Clemens and B. O. Williams (2013). "Lrp5 and Lrp6 exert overlapping functions in osteoblasts during postnatal bone acquisition." PLoS One **8**(5): e63323.

Riordan, J. R., J. M. Rommens, B. Kerem, N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. Lok, N. Plavsic and J. L. Chou (1989). "Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA." Science **245**(4922): 1066-1073.

Ritelli, M., C. Dordoni, V. Cinquina, M. Venturini, P. Calzavara-Pinton and M. Colombi (2017). "Expanding the clinical and mutational spectrum of B4GALT7-spondylodysplastic Ehlers-Danlos syndrome." Orphanet J Rare Dis **12**(1): 153.

Rohrbach, M. and C. Giunta (2012). "Recessive osteogenesis imperfecta: clinical, radiological, and molecular findings." Am J Med Genet C Semin Med Genet **160C**(3): 175-189.

Rolvien, T., S. Butscheidt, A. Jeschke, A. Neu, J. Denecke, C. Kubisch, M. H. Meisler, K. Puschel, F. Barvencik, T. Yorgan, R. Oheim, T. Schinke and M. Amling (2017). "Severe bone loss and multiple fractures in SCN8A-related epileptic encephalopathy." Bone **103**: 136-143.

Rousseau, J., R. Gioia, P. Layrolle, B. Lieubeau, D. Heymann, A. Rossi, J. C. Marini, V. Trichet and A. Forlino (2014). "Allele-specific Col1a1 silencing reduces mutant collagen in fibroblasts from Brl mouse, a model for classical osteogenesis imperfecta." Eur J Hum Genet **22**(5): 667-674.

Rugolotto, S., E. Monti, M. Carli, A. Pietrobelli, F. Antoniazzi and L. Tato (2007). "Pulmonary function tests in an infant with osteogenesis imperfecta and early biphosphonate treatment." Acta Paediatr **96**(12): 1856-1857.

Salter, C. G., J. H. Davies, R. J. Moon, J. Fairhurst, D. Bunyan, N. Foulds and D. Study (2016). "Further defining the phenotypic spectrum of B4GALT7 mutations." Am J Med Genet A **170**(6): 1556-1563.

Schwarze, U., T. Cundy, S. M. Pyott, H. E. Christiansen, M. R. Hegde, R. A. Bank, G. Pals, A. Ankala, K. Conneely, L. Seaver, S. M. Yandow, E. Raney, D. Babovic-Vuksanovic, J. Stoler, Z. Ben-Neriah, R. Segel, S. Lieberman, L. Siderius, A. Al-Aqeel, M. Hannibal, L. Hudgins, E. McPherson, M. Clemens, M. D. Sussman, R. D. Steiner, J. Mahan, R. Smith, K. Anyane-Yeboa, J. Wynn, K. Chong, T. Uster, S. Aftimos, V. R. Sutton, E. C. Davis, L. S. Kim, M. A. Weis, D. Eyre and P. H. Byers (2013). "Mutations in FKBP10, which result in Bruck syndrome and recessive forms of osteogenesis imperfecta, inhibit the hydroxylation of telopeptide lysines in bone collagen." Hum Mol Genet **22**(1): 1-17.

Schwarze, U., R. Hata, V. A. McKusick, H. Shinkai, H. E. Hoyme, R. E. Pyeritz and P. H. Byers (2004). "Rare autosomal recessive cardiac valvular form of Ehlers-Danlos syndrome results from mutations in the COL1A2 gene that activate the nonsense-mediated RNA decay pathway." Am J Hum Genet **74**(5): 917-930.

Segarra, N. G., D. Ballhausen, H. Crawford, M. Perreau, B. Campos-Xavier, K. van Spaendonck-Zwarts, C. Vermeer, M. Russo, P. Y. Zambelli, B. Stevenson, B. Royer-Bertrand, C. Rivolta, F. Candotti, S. Unger, F. L. Munier, A. Superti-Furga and L. Bonafe (2015). "NBAS mutations cause a multisystem disorder involving bone, connective tissue, liver, immune system, and retina." Am J Med Genet A **167A**(12): 2902-2912.

Seidahmed, M. Z., A. M. Alazami, O. B. Abdelbasit, K. Al Hussein, A. M. Miqdad, O. Abu-Sa'da, T. Mustafa, S. Bahjat and F. S. Alkuraya (2015). "Report of a case of Raine syndrome and literature review." Am J Med Genet A **167A**(10): 2394-2398.

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

Setijowati, E. D., F. S. van Dijk, J. M. Cobben, R. R. van Rijn, E. A. Sistermans, S. M. Faradz, S. Kawiya and G. Pals (2012). "A novel homozygous 5 bp deletion in FKBP10 causes clinically Bruck syndrome in an Indonesian patient." Eur J Med Genet **55**(1): 17-21.

Shaheen, R., M. Al-Owain, E. Faqeih, N. Al-Hashmi, A. Awaji, Z. Al-Zayed and F. S. Alkuraya (2011). "Mutations in FKBP10 cause both Bruck syndrome and isolated osteogenesis imperfecta in humans." Am J Med Genet A **155A**(6): 1448-1452.

Shaheen, R., A. M. Alazami, M. J. Alshammari, E. Faqeih, N. Alhashmi, N. Mousa, A. Alsinani, S. Ansari, F. Alzahrani, M. Al-Owain, Z. S. Alzayed and F. S. Alkuraya (2012). "Study of autosomal recessive osteogenesis imperfecta in Arabia reveals a novel locus defined by TMEM38B mutation." J Med Genet **49**(10): 630-635.

Shapiro, J. R., C. Lietman, M. Grover, J. T. Lu, S. C. Nagamani, B. C. Dawson, D. M. Baldrige, M. N. Bainbridge, D. H. Cohn, M. Blazo, T. T. Roberts, F. S. Brennen, Y. Wu, R. A. Gibbs, P. Melvin, P. M. Campeau and B. H. Lee (2013). "Phenotypic variability of osteogenesis imperfecta type V caused by an IFITM5 mutation." J Bone Miner Res **28**(7): 1523-1530.

Sillence, D. O., A. Senn and D. M. Danks (1979). "Genetic heterogeneity in osteogenesis imperfecta." J Med Genet **16**(2): 101-116.

Simpson, M. A., R. Hsu, L. S. Keir, J. Hao, G. Sivapalan, L. M. Ernst, E. H. Zackai, L. I. Al-Gazali, G. Hulskamp, H. M. Kingston, T. E. Prescott, A. Ion, M. A. Patton, V. Murday, A. George and A. H. Crosby (2007). "Mutations in FAM20C are associated with lethal osteosclerotic bone dysplasia (Raine syndrome), highlighting a crucial molecule in bone development." Am J Hum Genet **81**(5): 906-912.

Simpson, M. A., A. Scheuerle, J. Hurst, M. A. Patton, H. Stewart and A. H. Crosby (2009). "Mutations in FAM20C also identified in non-lethal osteosclerotic bone dysplasia." Clin Genet **75**(3): 271-276.

Smedley, D., J. O. Jacobsen, M. Jager, S. Kohler, M. Holtgrewe, M. Schubach, E. Siragusa, T. Zemojtel, O. J. Buske, N. L. Washington, W. P. Bone, M. A. Haendel and P. N. Robinson (2015). "Next-generation diagnostics and disease-gene discovery with the Exomiser." Nat Protoc **10**(12): 2004-2015.

Staufner, C., T. B. Haack, M. G. Köpke, B. K. Straub, S. Kölker, C. Thiel, P. Freisinger, I. Baric, P. J. McKiernan, N. Dikow, I. Harting, F. Beisse, P. Burgard, U. Kotzaeridou, D. Lenz, J. Kühr, U. Himbert, R. W. Taylor, F. Distelmaier, J. Vockley, L. Ghaloul-Gonzalez, J. A. Ozolek, J. Zschocke, A. Kuster, A. Dick, A. M. Das, T. Wieland, C. Terrile, T. M. Strom, T. Meitinger, H. Prokisch and G. F. Hoffmann (2016). "Recurrent

- acute liver failure due to NBAS deficiency: phenotypic spectrum, disease mechanisms, and therapeutic concepts." J Inherit Metab Dis **39**(1): 3-16.
- Stefanovic, B., L. Stefanovic, B. Schnabl, R. Bataller and D. A. Brenner (2004). "TRAM2 protein interacts with endoplasmic reticulum Ca²⁺ pump Serca2b and is necessary for collagen type I synthesis." Mol Cell Biol **24**(4): 1758-1768.
- Stewart, K. J. (1995). "A quantitative ultrastructural study of collagen fibrils in human skin, normal scars, and hypertrophic scars." Clin Anat **8**(5): 334-338.
- Streeter, I. and N. H. de Leeuw (2011). "A molecular dynamics study of the interprotein interactions in collagen fibrils." Soft Matter **7**(7): 3373-3382.
- Study, D. D. D. (2015). "Large-scale discovery of novel genetic causes of developmental disorders." Nature **519**(7542): 223-228.
- Su, N., M. Jin and L. Chen (2014). "Role of FGF/FGFR signaling in skeletal development and homeostasis: learning from mouse models." Bone Res **2**: 14003.
- Sugawara, M., T. Nakanishi, Y. J. Fei, R. G. Martindale, M. E. Ganapathy, F. H. Leibach and V. Ganapathy (2000). "Structure and function of ATA3, a new subtype of amino acid transport system A, primarily expressed in the liver and skeletal muscle." Biochim Biophys Acta **1509**(1-2): 7-13.
- Symoens, S., A. M. Barnes, C. Gistelincx, F. Malfait, B. Guillemy, W. Steyaert, D. Syx, S. D'hondt, M. Biervliet, J. De Backer, E. P. Witten, S. Leikin, E. Makareeva, G. Gillissen-Kaesbach, A. Huysseune, K. Vleminckx, A. Willaert, A. De Paepe, J. C. Marini and P. J. Coucke (2015). "Genetic Defects in TAPT1 Disrupt Ciliogenesis and Cause a Complex Lethal Osteochondrodysplasia." Am J Hum Genet **97**(4): 521-534.
- Symoens, S., F. Malfait, S. D'hondt, B. Callewaert, A. Dheedene, W. Steyaert, H. P. Bächinger, A. De Paepe, H. Kayserili and P. J. Coucke (2013). "Deficiency for the ER-stress transducer OASIS causes severe recessive osteogenesis imperfecta in humans." Orphanet J Rare Dis **8**: 154.
- Syx, D., B. Guillemy, S. Symoens, A. B. Sousa, A. Medeira, M. Whiteford, T. Hermanns-Le, P. J. Coucke, A. De Paepe and F. Malfait (2015). "Defective Proteolytic Processing of Fibrillar Procollagens and Prodecorin Due to Biallelic BMP1 Mutations Results in a Severe, Progressive Form of Osteogenesis Imperfecta." J Bone Miner Res **30**(8): 1445-1456.
- Tagliabracci, V. S., S. E. Wiley, X. Guo, L. N. Kinch, E. Durrant, J. Wen, J. Xiao, J. Cui, K. B. Nguyen, J. L. Engel, J. J. Coon, N. Grishin, L. A. Pinna, D. J. Pagliarini and J. E. Dixon (2015). "A Single Kinase Generates the Majority of the Secreted Phosphoproteome." Cell **161**(7): 1619-1632.
- Takagi, M., N. Hori, Y. Chinen, K. Kurosawa, Y. Tanaka, K. Oku, H. Sakata, R. Fukuzawa, G. Nishimura, J. Spranger and T. Hasegawa (2011). "Heterozygous C-propeptide mutations in COL1A1: osteogenesis imperfecta type IIC and dense bone variant." Am J Med Genet A **155A**(9): 2269-2273.
- Taylan, F. and O. Mäkitie (2016). "Abnormal Proteoglycan Synthesis Due to Gene Defects Causes Skeletal Diseases with Overlapping Phenotypes." Horm Metab Res **48**(11): 745-754.

Teraishi, M., M. Takaishi, K. Nakajima, M. Ikeda, Y. Higashi, S. Shimoda, Y. Asada, A. Hijikata, O. Ohara, Y. Hiraki, S. Mizuno, T. Fukada, T. Furukawa, N. Wakamatsu and S. Sano (2017). "Critical involvement of ZEB2 in collagen fibrillogenesis: the molecular similarity between Mowat-Wilson syndrome and Ehlers-Danlos syndrome." Sci Rep **7**: 46565.

Trackman, P. C. (2016). "Enzymatic and non-enzymatic functions of the lysyl oxidase family in bone." Matrix Biol **52-54**: 7-18.

Trombetta, S. E., M. Bosch and A. J. Parodi (1989). "Glucosylation of glycoproteins by mammalian, plant, fungal, and trypanosomatid protozoa microsomal membranes." Biochemistry **28**(20): 8108-8116.

Unnanuntana, A., B. J. Rebolledo, M. M. Khair, E. F. DiCarlo and J. M. Lane (2011). "Diseases affecting bone quality: beyond osteoporosis." Clin Orthop Relat Res **469**(8): 2194-2206.

Valencia, M., J. A. Caparrós-Martin, M. S. Sirerol-Piquer, J. M. García-Verdugo, V. Martínez-Glez, P. Lapunzina, S. Temtamy, M. Aglan, A. M. Lund, P. G. Nikkels, V. L. Ruiz-Perez and E. Ostergaard (2014). "Report of a newly identified patient with mutations in BMP1 and underlying pathogenetic aspects." Am J Med Genet A **164A**(5): 1143-1150.

Valli, M., A. M. Barnes, A. Gallanti, W. A. Cabral, S. Viglio, M. A. Weis, E. Makareeva, D. Eyre, S. Leikin, F. Antoniazzi, J. C. Marini and M. Mottes (2012). "Deficiency of CRTAP in non-lethal recessive osteogenesis imperfecta reduces collagen deposition into matrix." Clin Genet **82**(5): 453-459.

van der Slot, A. J., A. M. Zuurmond, A. F. Bardoel, C. Wijmenga, H. E. Pruijs, D. O. Sillence, J. Brinckmann, D. J. Abraham, C. M. Black, N. Verzijl, J. DeGroot, R. Hanemaaijer, J. M. TeKoppele, T. W. Huizinga and R. A. Bank (2003). Identification of PLOD2 as telopeptide lysyl hydroxylase, an important enzyme in fibrosis. J Biol Chem **278**: 40967-40972.

van Dijk, F. S., J. M. Cobben, A. Kariminejad, A. Maugeri, P. G. Nikkels, R. R. van Rijn and G. Pals (2011). Osteogenesis Imperfecta: A Review with Clinical Examples. Mol Syndromol. **2**: 1-20.

van Dijk, F. S., M. Huizer, A. Kariminejad, C. L. Marcelis, A. S. Plomp, P. A. Terhal, H. Meijers-Heijboer, M. M. Weiss, R. R. van Rijn, J. M. Cobben and G. Pals (2010). "Complete COL1A1 allele deletions in osteogenesis imperfecta." Genet Med **12**(11): 736-741.

van Dijk, F. S., I. M. Nesbitt, E. H. Zwikstra, P. G. Nikkels, S. R. Piersma, S. A. Fratantoni, C. R. Jimenez, M. Huizer, A. C. Morsman, J. M. Cobben, M. H. van Roij, M. W. Elting, J. I. Verbeke, L. C. Wijnaendts, N. J. Shaw, W. Hogler, C. McKeown, E. A. Sistermans, A. Dalton, H. Meijers-Heijboer and G. Pals (2009). PPIB mutations cause severe osteogenesis imperfecta. Am J Hum Genet. United States. **85**: 521-527.

van Dijk, F. S., G. Pals, R. R. Van Rijn, P. G. Nikkels and J. M. Cobben (2010). Classification of Osteogenesis Imperfecta revisited. Eur J Med Genet. Netherlands, 2009 Elsevier Masson SAS. **53**: 1-5.

- van Dijk, F. S. and D. O. Sillence (2014). "Osteogenesis imperfecta: clinical diagnosis, nomenclature and severity assessment." Am J Med Genet A **164A**(6): 1470-1481.
- van Dijk, F. S., M. C. Zillikens, D. Micha, M. Riessland, C. L. Marcelis, C. E. de Die-Smulders, J. Milbradt, A. A. Franken, A. J. Harsevoort, K. D. Lichtenbelt, H. E. Pruijs, M. E. Rubio-Gozalbo, R. Zwertbroek, Y. Moutaouakil, J. Egthuijsen, M. Hammerschmidt, R. Bijman, C. M. Semeins, A. D. Bakker, V. Everts, J. Klein-Nulend, N. Campos-Obando, A. Hofman, G. J. te Meerman, A. J. Verkerk, A. G. Uitterlinden, A. Mageri, E. A. Sistermans, Q. Waisfisz, H. Meijers-Heijboer, B. Wirth, M. E. Simon and G. Pals (2013). "PLS3 mutations in X-linked osteoporosis with fractures." N Engl J Med **369**(16): 1529-1536.
- Volodarsky, M., B. Markus, I. Cohen, O. Staretz-Chacham, H. Flusser, D. Landau, I. Shelef, Y. Langer and O. S. Birk (2013). "A Deletion Mutation in TMEM38B Associated with Autosomal Recessive Osteogenesis Imperfecta." Hum Mutat **34**(4): 582-586.
- von Marschall, Z. and L. W. Fisher (2010). "Decorin is processed by three isoforms of bone morphogenetic protein-1 (BMP1)." Biochem Biophys Res Commun **391**(3): 1374-1378.
- Walker, L. C., M. A. Overstreet, A. Siddiqui, A. De Paepe, G. Ceylaner, F. Malfait, S. Symoens, P. Atsawasuwana, M. Yamauchi, S. Ceylaner, R. A. Bank and H. N. Yeowell (2005). "A novel mutation in the lysyl hydroxylase 1 gene causes decreased lysyl hydroxylase activity in an Ehlers-Danlos VIA patient." J Invest Dermatol **124**(5): 914-918.
- Wan, M., C. Yang, J. Li, X. Wu, H. Yuan, H. Ma, X. He, S. Nie, C. Chang and X. Cao (2008). "Parathyroid hormone signaling through low-density lipoprotein-related protein 6." Genes Dev **22**(21): 2968-2979.
- Wang, J. Y., Y. Liu, L. J. Song, F. Lv, X. J. Xu, A. San, J. Wang, H. M. Yang, Z. Y. Yang, Y. Jiang, O. Wang, W. B. Xia, X. P. Xing and M. Li (2017). "Novel Mutations in SERPINF1 Result in Rare Osteogenesis Imperfecta Type VI." Calcif Tissue Int **100**(1): 55-66.
- Ward, L. M., F. Rauch, R. Travers, G. Chabot, E. M. Azouz, L. Lalic, P. J. Roughley and F. H. Glorieux (2002). "Osteogenesis imperfecta type VII: an autosomal recessive form of brittle bone disease." Bone **31**(1): 12-18.
- Warman, M. L., V. Cormier-Daire, C. Hall, D. Krakow, R. Lachman, M. LeMerrer, G. Mortier, S. Mundlos, G. Nishimura, D. L. Rimoïn, S. Robertson, R. Savarirayan, D. Sillence, J. Spranger, S. Unger, B. Zabel and A. Superti-Furga (2011). "Nosology and classification of genetic skeletal disorders: 2010 revision." Am J Med Genet A **155A**(5): 943-968.
- Weitzmann, M. N. (2017). "Bone and the Immune System." Toxicol Pathol **45**(7): 911-924.
- Wilkinson, B. and H. F. Gilbert (2004). "Protein disulfide isomerase." Biochim Biophys Acta **1699**(1-2): 35-44.
- Williams, B. O. and K. L. Insogna (2009). "Where Wnts went: the exploding field of Lrp5 and Lrp6 signaling in bone." J Bone Miner Res **24**(2): 171-178.

- Wu, M., G. Chen and Y. P. Li (2016). "TGF- β and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease." Bone Res **4**: 16009.
- Yan, J., H. J. Liu, H. Li, L. Chen, Y. Q. Bian, B. Zhao, H. X. Han, S. Z. Han, L. R. Han, D. W. Wang and X. F. Yang (2017). "Circulating periostin levels increase in association with bone density loss and healing progression during the early phase of hip fracture in Chinese older women." Osteoporos Int **28**(8): 2335-2341.
- Yeowell, H. N., J. D. Allen, L. C. Walker, M. A. Overstreet, S. Murad and S. F. Thai (2000). "Deletion of cysteine 369 in lysyl hydroxylase 1 eliminates enzyme activity and causes Ehlers-Danlos syndrome type VI." Matrix Biol **19**(1): 37-46.
- Yeowell, H. N., L. C. Walker and L. M. Neumann (2005). "An Ehlers-Danlos syndrome type VIA patient with cystic malformations of the meninges." Eur J Dermatol **15**(5): 353-358.
- Yu, K. W., C. C. Yao, J. H. Jeng, H. Y. Shieh and Y. J. Chen (2018). "Periostin inhibits mechanical stretch-induced apoptosis in osteoblast-like MG-63 cells." J Formos Med Assoc **117**(4): 292-300.
- Zaman, G., K. A. Staines, C. Farquharson, P. T. Newton, J. Dudhia, C. Chenu, A. A. Pitsillides and G. K. Dhoot (2016). "Expression of Sulf1 and Sulf2 in cartilage, bone and endochondral fracture healing." Histochem Cell Biol **145**(1): 67-79.
- Zhang, H., H. Yue, C. Wang, J. Gu, J. He, W. Fu, W. Hu and Z. Zhang (2017). "Novel mutations in the SEC24D gene in Chinese families with autosomal recessive osteogenesis imperfecta." Osteoporos Int **28**(4): 1473-1480.
- Zhang, J., R. Yang, Z. Liu, C. Hou, W. Zong, A. Zhang, X. Sun and J. Gao (2015). "Loss of lysyl oxidase-like 3 causes cleft palate and spinal deformity in mice." Hum Mol Genet **24**(21): 6174-6185.
- Zhang, R., J. R. Edwards, S. Y. Ko, S. Dong, H. Liu, B. O. Oyajobi, C. Papasian, H. W. Deng and M. Zhao (2011). "Transcriptional regulation of BMP2 expression by the PTH-CREB signaling pathway in osteoblasts." PLoS One **6**(6): e20780.
- Zheng, X., X. Gai, S. Han, C. D. Moser, C. Hu, A. M. Shire, R. A. Floyd and L. R. Roberts (2013). "The human sulfatase 2 inhibitor 2,4-disulfonylphenyl-tert-butyl nitron (OKN-007) has an antitumor effect in hepatocellular carcinoma mediated via suppression of TGFB1/SMAD2 and Hedgehog/GLI1 signaling." Genes Chromosomes Cancer **52**(3): 225-236.

APPENDIX

A. SRCAP c.9029C>A,p.(Pro3010His) pathogenicity evaluation form	190
B. Patient phenotypes and variants evaluated in the DDD CAP12 study	197
C. Poster: International Skeletal Dysplasia Society (2017)	224
D. Poster: International Conference on Childrens Bone Health (2017).....	225
E. Manuscript: Phenotypic Variability in Patients with Osteogenesis Imperfecta caused by BMP1 Mutations.....	226
F. Manuscript: P4HB recurrent missense mutation causing Cole-Carpenter syndrome.....	232
G. Manuscript: Compound Heterozygous Vairants in NBAS as a cause of atypical osteogenesis imperfecta.....	240
H. List of Genes in Targeted Panel for Exome analysis.....	250
I. Patient information sheet for young people.....	277
J. Approved funding application for Targeted Exome Sequencing Project.....	283
K. DDD complementary research proposal submission form	296
L. Confirmation of Ethical Approval for Target Exomes.....	298
M. NBAS Gene Dossier Approved by UKGTN	303

A. SRCAP c.9029C>A.p.(Pro3010His) pathogenicity evaluation form

Unclassified Variant Evaluation

Checklist to help document lines of evidence in classifying a sequence variant

Disease	Floating-Harbor syndrome		Ref Seq No		Variant
Gene	SRCAP	cDNA	NM_006662.2	DNA	c.9029C>A
Exon	34			Protein	p.Pro3010His

Classification

Class 1 - Neutral polymorphism

Class 2 - Probably neutral polymorphism

3

Class 3 - Unclassified variant

Class 4 - Probably pathogenic mutation

Class 5 - Pathogenic mutation

CONCLUSION

Truncating mutations in this gene are associated with Floating-Harbor syndrome (FHS), inherited in an autosomal dominant manner. FHS clinical features = proportionate short stature, delayed bone age, delayed speech development, and dysmorphic facial features: triangular face with deep-set eyes, long eyelashes, bulbous nose, wide columella, short philtrum, and thin lips.

All the pathogenic FHS mutations reported to date are in the last 2 exons; exons 33 and 34.

The variant is not listed on dbSNP and is de-novo in the proband.

The variant is in a moderately conserved nucleotide and protein position and is within a functional domain of the protein.

In silico analysis - conflicting predictions.

This variant is of uncertain significance.

SUMMARY

		Support pathogenicity
	Extent of Analysis	N/I
	Population Frequency	YES
	Web-based Literature Search	N/I
	Frameshift Mutation	N/A
	In-frame Deletion or Insertion	N/A
	Nonsense Mutation	N/A
	Evolutionary Conservation	NO
	Severity of Amino Acid Substitution	YES
	Splice Site Prediction	NO
	De-Novo Variant	YES
FINAL CLASSIFICATION	1 = Neutral polymorphism 2 = Probably neutral polymorphism 3 = Unclassified variant 4 = Probably pathogenic mutation 5 = Pathogenic mutation	3

Extent of Analysis

Has the entire coding sequence been analysed?	blank
Methodology (sequencing or [which] mutation scanning method)	Sanger
MLPA for deletions/duplications	No
Have other unclassified variants been characterised in this gene in this patient?	No
Classification of any other variants	blank
Is there genetic heterogeneity?	Yes
What other gene(s) has/have been analysed?	All Exomes
Have any variants been characterised in this/these gene(s)?	No
List any other variants identified	N/A
Classification of variants in these other genes	N/A
<p>Comments – Variant identified from exome data as de-novo.</p> <p>Patient's clinical features are suggestive of Osteogenesis Imperfecta but are atypical. He has rhizomelia and mild immunodeficiency as well as multiple fractures. No likely causative mutation has been identified in genes known to be associated with OI.</p> <p>Heterozygous truncating variants in SRCAP gene are associated with Floating Harbor syndrome (autosomal dominant) which has a skeletal phenotype of short stature.</p>	
Does the extent of analysis support pathogenicity?	N/I

Population Frequency

Has the UV been tested in normal controls?	no
If yes, how many?	

Has the UV been tested in other screened patients?	no
If yes how many?	
Is the UV on NCBI dbSNP http://www.ncbi.nlm.nih.gov/pubmed	No
Is the UV on ENSEMBL www.ensembl.org gene variation report	No
Is the UV on the ALlele FREquency Database http://alfred.med.yale.edu/alfred/index.asp gene variation report?	No
RefSNP accession number (rs#)	
Allele frequency recorded on dbSNP	
Minor Allele Count	
Population studied :	
There is no report of either this variant or any other variant at this amino acid position	
Comments	
Does population frequency support pathogenicity?	YES

Web-based Literature Search

Search tool		Search Terms
PubMed http://www.ncbi.nlm.nih.gov/pubmed	<input checked="" type="checkbox"/>	"SRCAP" ("9029C>A" "9029C->A" "9029C-->A" "9029C/A" "Pro3010His" "P3010H")
Google Scholar http://scholar.google.co.uk/	<input checked="" type="checkbox"/>	As above
Google http://www.google.co.uk	<input checked="" type="checkbox"/>	As above
Locus specific database (specify)	<input type="checkbox"/>	
DMuDB http://www.ngrl.org.uk/Manchester/projects/informatics/dmudb	<input type="checkbox"/>	
HGMD http://www.biobase-international.com/product/hgmd	<input checked="" type="checkbox"/>	SRCAP mutations
OMIM http://www.omim.org/	<input type="checkbox"/>	
Other - specify	<input type="checkbox"/>	
References and comments		
<p>This mutation not listed in the HGMD database. There are 2 missense mutation that are listed as ?disease causing mutation and reported with an association with atrial septal defect but these are not in the same region as the p.Pro3010His. Other mutations in the gene are nonsense mutations reported in association with Floating-Harbor Syndrome and tightly clustered around the last 2 exons of the gene (Hood et al., 2012 Mutations in SRCAP, Encoding SNF2-Related CREBBP Activator Protein, Cause Floating-Harbor Syndrome. Am J Hum Genet). The p.Pro3010His is located in this region.</p>		
Does review of the literature support pathogenicity?		N/I

Frameshift Mutation

Does the mutation result in a shift in the reading frame?	no
Comments	
Does this support pathogenicity?	N/A

In-frame Deletion or Insertion Mutation

Does the mutation result in an altered protein that maintains the reading frame?	no
Comments	
Does this support pathogenicity?	N/A

Nonsense Mutation

Does the change result in a stop mutation?	no
Comments	
Does this support pathogenicity?	N/A

Evolutionary Conservation

Analysis of a variant should be performed using at least three different programmes to achieve a consensus			
Context http://pfam.xfam.org/ & http://www.uniprot.org/			
Is the UV in a conserved domain?			
Nucleotide	yes	Amino acid	yes
Is this a functional domain?.			yes
What function? AT hook, DNA binding domain and AT hook like			
Is the residue predicted to affect post-translational modification, disulphide bonding etc? No			
Comments Moderately conserved nucleotide and amino acid (considering 11 species)			
The following programs are grouped dependent on type of tool- please use at least one from each section (these are in order of NGRL recommendation)			
Machine-learning (SVM)			
SNPs&GO http://snps.biofold.org/snps-and-go/			
Method	Prediction	RI	Probability
PhD-SNP	Neutral	5	0.249
PANTHER	Disease	3	0.672
SNPs&GO	Neutral	4	0.311
Comments conflictng results and low reliability			
Supervised-learning			
MutPred http://mutpred.mutdb.org/			
Probability of being deleterious			0.481
Hypotheses			none
Comment			

Loss of catalytic residue at P3010 (P = 0.0081)	
Loss of glycosylation at P3010 (P = 0.014)	
Gain of MoRF binding (P = 0.0526)	
Loss of ubiquitination at K3012 (P = 0.1881)	
Loss of methylation at R3009 (P = 0.1902)	
Note : Only one supervised learning has returned a result.	
Sequence/evolutionary conservation	
PROVEAN http://provean.jcvi.org/genome_submit_2.php?species=human	
- N.B. this tool makes PROVEAN and SIFT predictions simultaneously	
Score	-2.53
PREDICTION	Deleterious
Number of Sequences used	82
Comments	
SIFT Analysis http://provean.jcvi.org/genome_submit_2.php?species=human	
- N.B. this tool makes PROVEAN and SIFT predictions simultaneously and gives a more confident prediction than SIFT through Alamut.	
SIFT score	0.000
SIFT median conservation score	4.32
Number of Sequences used	33
SIFT prediction	Damaging
Comments	
MutationAssessor http://mutationassessor.org/	
Functional Impact	Neutral
FI score	0.695
Comments 2/3 Sequence/evolutionary conservation tools suggest that this is a deleterious variant	
Protein sequence and structure based	
PolyPhen2 Analysis http://genetics.bwh.harvard.edu/pph2/	
HumDiv (For evaluating rare alleles)	Probably damaging
HumDiv Score	0.997
HumDiv Sensitivity and Specificity:	0.41 0.98
HumVar (For Mendelian disease)	Probably damaging
HumVar Score	0.781

HumVar Sensitivity and Specificity:	0.76 0.87
Alignment checked for validity	yes
3D Visualization present?	no
Comments	
Does evolutionary conservation analysis support pathogenicity?	
Comments	NO

Severity of amino Acid Substitution

Russell Analysis	
Effect according to Russell http://www.russelllab.org/aas/	Disfavoured
Degree of effect (numerical value)	-2
Amino acid comparison (Alamut)	
BLOSUM45	-2
BLOSUM62	-2
BLOSUM80	-4
Grantham distance	77
Comments : The function of the SRCAP protein is not fully characterised, Russel score has been taken from the 'all protein' type.	
Does severity of amino acid substitution support pathogenicity?	YES

Splice Site Prediction

Predicted to affect splicing by SpliceSiteFinder (like) Accessed through Alamut	no
Predicted to affect splicing by MaxEntScan ; Accessed through Alamut	no
http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html	
http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq_acc.html	
Predicted to affect splicing by NNSplice (FruitFly) ; Accessed through Alamut	no
http://www.fruitfly.org/seq_tools/splice.html	
Predicted to affect splicing by GeneSplicer ; Accessed through Alamut	no
http://www.cbcb.umd.edu/software/GeneSplicer/gene_spl.shtml	
Predicted to affect splicing by Human Splicing Finder ; Accessed through Alamut	no
http://www.umd.be/HSF3/	
Comments: No significant effect on splicing predicted	
Do predicted splicing effects support pathogenicity?	NO

De-Novo Variant

Is father unaffected?	Yes
Does father carry UV?	No
Has paternity been confirmed ?(PowerPlex 16 assay)	No
Is mother unaffected?	Yes
Does mother carry UV?	No
Has germline mosaicism been reported for this disorder?	No
Comments – Likely de-novo	
Do parental studies support pathogenicity	YES

B. Patient phenotypes and variants evaluated in the DDD CAP12 study

Highlighted variants (Pink) were reported by Decipher (DDD) to the patients referring clinician as potentially causative of the patient’s symptoms.

Patient Phenotype	ID	Decipher ID	Chr	Location	Gene	Notes	Consequence	ref/alt alleles	trio_genotype*
Bowing of limbs due to multiple fractures, Fractured hand bones, Pathologic fracture	101	258358	1	109271328	<i>FNDC7</i>	c.1451_1453del;p.(Asp484del). Also heterozygous for <i>GLI3</i> c.2179G>A;p.(Gly727Arg) variants reported as pathogenic and associated with Postaxial polydactyly A/B but present in up to 1% of population and described as benign.	In frame deletion	AATG/A	1/0/0
Abnormality of the radius, Blue sclerae, Multiple rib fractures, Pathologic fracture, Plagiocephaly, Thoracic kyphosis, Vertebral compression fractures	102	259467	2	86285812	86509436	CNV reported Ratio 0.504374. DUPLICATION. Genes in region: IMMT - mitochondrial protein" " <i>REEP1</i> - AD HSP" " <i>PTCD3</i> - mitochondrial protein" " <i>SNORD94</i> - small nuclear RNA" " <i>POLR1A</i> - AD Acrofacial dysostosis, Cincinnati type. A ribosomopathy characterized by a spectrum of mandibulofacial dysostosis phenotypes, with or without extra facial skeletal defects (Weaver et al 2015) including severe conductive hearing loss, short stature. congenital short, bowed femurs with metaphyseal flaring, dysplastic acetabula, and delayed or absent ossification of the capital	Large duplication		1/0/1

					femoral epiphyses" "MIR4779 - micro RNA" "MRPL35 - mitochondrial ribosomal RNA"				
			3	38739788	<i>SCN10A</i>	c.4923G>T;p.(Met164Ile). Gene associated with Brugada syndrome. Polymorphisms associated with atrial fibrillation.	Missense variant	C/A	1/0/1
Abnormality of the helix, Joint laxity, Low-set ears, Narrow foot, Narrow mouth, Plagiocephaly, Recurrent fractures, Scoliosis, Seizures, Slanting of the palpebral fissure	103	260858	16	89350233	<i>ANKRD11</i>	Reported by DDD. Gene associated with KBG syndrome= devdel, short stature and skeletal abnormalities. BUT not associated with recurrent fractures. LIKELY BENIGN, maternally inherited.	Missense variant	C/A	1/NA/NA
			7	40039066	<i>CDK13</i>	Reported by DDD. PATHOGENIC mutation reported in association with congenital heart disease, devdel, clinodactyl and hyper mobility. Sifrim et al Nat Genet 2016 – DDD publication. DE-NOVO	Missense variant	G/A	1/NA/NA
			15	51834626	<i>DMXL2</i>	0.003% in gnomAD. ? Associated with AD deafness. Not consistent with reported patient phenotype.	Missense variant	C/T	1/NA/NA

			7	154672650	<i>DPP6</i>	c.2131A>T; p.(Lys711*).Not reported. Disease - Mental retardation, autosomal dominant 33 - microcephaly and variable mental retardation. Loss of function mutations cause this (Liao, C., Fu, F., Li, R., Yang, W., Liao, H., Yan, J., Li, J., Li, S., Yang, X., Li, D. Loss-of-function variation in the DPP6 gene is associated with autosomal dominant microcephaly and mental retardation. <i>Europ. J. Med. Genet.</i> 56: 484-489, 2013).	Stop gained	A/T	1/NA/NA
			15	101566237	<i>LRRK1</i>	c.2300G>T;p.(Arg767Leu). GO biological processes included: bone resorption, intracellular signal transduction, negative regulation of peptidyl-tyrosine phosphorylation, osteoclast development, positive regulation of canonical Wnt signalling pathway, positive regulation of intracellular signal transduction. BUT present in 2/120712 alleles in gnomAD	Missense variant	G/T	1/NA/NA
			16	81213294	<i>PKD1L2</i>	PKD gene	Missense variant	G/C	1/NA/NA
			15	67457239	<i>SMAD3</i>	Reported by DDD, stop_gained. Loeyes Dietz 3 syndrome. May get early onset osteoarthritis. BENIGN as alternate transcript.	Stop gained	G/A	1/NA/NA
Disproportionate short stature, Hypoplasia of dental enamel, Limited elbow extension, Moderate generalized osteoporosis, Skeletal dysplasia	104	262876	17	27963096	<i>SSH2</i>	<i>De novo</i> c.2152G>A;p.(Val718Met). Protein tyrosine phosphatase that play a key role in the regulation of actin filaments. <i>SSH2</i> =exomizer gene.	Missense variant	C/T	1/0/0
2-3 toe syndactyly, Abnormality of male external genitalia, Bilateral single	105	263310	5	177031401	<i>B4GALT7</i>	Reported by DDD, c.277dupC;p.(His93fs). Paternally inherited	frameshift	G/GC	1/0/1

transverse palmar creases, Cleft palate, Gastroesophageal reflux, Generalized hypotonia, Hemimegalencephaly, Hypermetropia, Lumbar hemivertebrae, Micrognathia, Mild short stature, Optic nerve hypoplasia, Osteopenia, Radioulnar synostosis, Ventriculomegaly, Vertebral compression fractures			5	177035541	<i>B4GALT7</i>	Reported by DDD, c.641G>A;p.(Cys214Tyr). Maternally inherited	Missense variant	G/A	1/1/0
Abnormality of taste sensation, Constipation, Hyperhidrosis, Impaired pain sensation, Lacrimation abnormality, Muscle weakness, Painless fractures due to injury, Urticaria	106	263659	X	70360588	<i>MED12</i>	Reported by DDD, c.736_741del; p.(Gln249_Gln250del) inframe_deletion. Gene associated with Opitz-Kaveggia syndrome (305450), also known as FG syndrome, is an X-linked disorder characterized by mental retardation, relative macrocephaly, hypotonia, and constipation. ClinVar likely benign - in poly Gln tract.	In frame deletion	G/GAG CAGC	1/1/0
			2	167056281	<i>SCN9A</i>	c.4835T>G; p.(Leu1612Arg). <i>SCN9A</i> =HPO gene congenital insensitivity to pain - autosomal recessive, no other likely deleterious variants identified (p.Leu1612Pro reported in Paroxysmal extreme pain disorder). Mutation at this codon associated with paroxysmal extreme pain disorder. Gene also associated with congenital insensitivity to pain.	Missense variant	A/C	1/0/0
Abnormality of T cells, Bilateral cryptorchidism, Eczema, Glandular hypospadias, Hypoplasia of the corpus callosum, Intellectual disability, moderate, Low-set ears, Microcephaly,	107	264065	2	233399868	<i>CHRND</i>	Associated with lethal multiple pterguim syndrome characterised by arthrogyposis, fetal akensis and hydrops	Missense variant	G/A	2/NA/NA
			21	47532415	<i>COL6A2</i>	Bethlem myopathy : patient does not have and present in 0.013% AA popn.	Missense variant	G/A	1/NA/NA

Micrognathia, Micropenis, Microtia, Osteopenia, Osteoporosis of vertebrae, Primary adrenal insufficiency, Severe short stature, Short foot, Stenosis of the external auditory canal			20	46290600	<i>SULF2</i>	c.2411A>G;p.(Asn804Ser). GO biological process include: bone development, embryonic skeletal development, heparan sulphate proteoglycan metabolic process. GeneAtlas basic functions include:SULF1, SULF2 plays an important role in the development of the inner ear. gnomAD freq 1 in 240358 alleles. SULF2 variant also in patient 265826	Missense variant	T/C	1/NA/NA
			11	122647887	<i>UBASH3B</i>	c.371Tdel;p.(Phe124fs). GO biological process: collagen-activated tyrosine kinase receptor signalling pathway, negative regulation of bone resorption, negative osteoclast differentiation. Expressed in bone, immune system, reproductive and platelets. Gene atlas - suppression of T-cell receptor signalling. Critical regulator of the signalling pathways that regulate T-cell activation. pLI=0.99.We consider pLI >= 0.9 as an extremely LoF intolerant set of genes.	Frameshift variant	TC/T	1/NA/NA
			2	145157542	<i>ZEB2</i>	c.1212G>T;p.(Met404Ile). Reported by DDD, missense_variant. Associated with Mowat-Wilson syndrome. Complex dev del disorder, not assoc with bone fragility	Missense variant	C/A	1/NA/NA
Ataxia, Blue sclerae, Delayed speech and language development, Global developmental delay, Horizontal nystagmus, Joint hypermobility, Poor coordination, Recurrent fractures	108	264081	17	48267063	<i>COL1A1</i>	Reported by DDD- pathogenic mutation, stop gained. Inherited from mother. c.2644C>T;p.(Arg882*)	Stop gained	G/A	1/1/0
			2	166797557	<i>TTC21B</i>	AR Short-rib thoracic dysplasia 4 with or without polydactyly	Missense variant	C/G	1/1/0

Microcephaly, Osteopenia, Ovarian cyst, Specific learning disability, Tetralogy of Fallot	109	264182	X	41205786	DDX3X	c.1526A>T;p.(Asn509Ile). Reported de-novo missense-variant. XLD mental retardation s-linked type 102. Below average general intellectual functioning associated with impairments in adaptive behaviour and manifested during the developmental period. Additionally, patients manifest variable non-neurologic features such as joint hyperlaxity, skin pigmentary abnormalities, cleft lip and/or palate, hearing and visual impairment, and precocious puberty. Previously reported pathogenic (Snijders Blok (2015) Am J Hum Genet 97: 343)	Missense variant	A/T	1/0/0
Abnormality of the face, Asymmetry of spinal facet joints, Increased susceptibility to fractures, Joint hypermobility, Pes planus, Specific learning disability	110	264424	17	48263382	COL1A1	Reported by DDD, splice _variant. C.4006-1G>T	Splice variant	C/A	1/NA/NA
			17	80789837	ZNF750	c.494T>C;p.(Val165Ala). Reported by DDD, missense_variant. Seborrhea-like dermatitis with psoriasiform elements - inheritance not stated in OMIM. No skeletal phenotype reported. One frameshift mutation described in HGMD.	Missense variant	A/G	1/NA/NA
Abnormality of upper lip, Drooling, Generalized seizures, Global developmental delay, Intellectual disability, severe, Involuntary movements, Microcephaly, Nystagmus, Osteopenia, Postnatal microcephaly, Reduced bone mineral density, Thick upper lip vermilion	111	264568							

Decreased antibody level in blood, Elevated hepatic transaminases, Hypoglycaemia, IgA deficiency, IgG deficiency, Increased susceptibility to fractures, Ligamentous laxity, Nystagmus, Osteopenia, Progeroid facial appearance, Progressive cone degeneration, Recurrent fractures, Wide anterior fontanel	112	264693	2	15378794	NBAS	Reported, c.5741G>A;p.(Arg1914His)	Missense variant	C/T	1/1/0
			2	15607518	NBAS	Reported stop_gained, c.2032C>T;p.(Gln678*)	Stop gained	G/A	1/0/1
Blue sclerae, Ligamentous laxity, Recurrent fractures	113	264830	11	67258406	AIP	Gene causes a benign hormone-secreting adenoma - Clinical manifestations of Cushing syndrome include facial and obesity, abdominal striae, muscular weakness, osteoporosis, arterial hypertension, diabetes	Missense variant	G/A	1/NA/NA
			8	27657118	ESCO2	Gene associated with AR Roberts syndrome. Rare autosomal recessive disorder characterized by pre- and postnatal growth retardation, microcephaly, bilateral cleft lip and palate, and mesomelic symmetric limb reduction.	Missense variant	A/G	1/NA/NA
			1	22168854	HSPG2	Gene associated with Autosomal recessive: Dysegmental dysplasia, Silverman-Handmaker type and Schwartz-Jampel syndrome, type 1(lethal forms of neonatal short-limbed dwarfism)	Missense variant	C/T	1/NA/NA
			9	140347297	NSMF	Variant is intronic in major transcript with no predicted affect on splicing.	Missense variant	G/A	1/NA/NA
			16	30732293	SRCAP	Reported missense_variant. c.3247C>T;p.(Pro1083Ser). 12/28456 South Asian alleles, not in mutation hotspot region.	Missense variant	C/T	1/NA/NA

Cafe-au-lait spot, Generalized dystonia, Global developmental delay, Osteopenia, Reduced bone mineral density, Seizures, Status epilepticus	114	265503	5	156920055	<i>ADAM19</i>	Uni-Prot May be involved in osteoblast differentiation and/or osteoblast activity in bone. 4/126710 alleles in Europeans. pLI 0.00. Decipher has 12 individuals with loss of one copy of this gene, none with skeletal.	Stop gained	G/A	1/NA/NA
			6	75862094	<i>COL12A1</i>	c.3670C>T;p.(Arg1224Cys) Gene - Ullrich/Bethlem -AD, missense mutations reported as pathogenic. This variant present in 7/276644 alleles therefore unlikely pathogenic.	Missense variant	G/A	1/NA/NA
			6	70962012	<i>COL9A1</i>	c.1771A>G; p.(Thr591Ala). Present in 1 in 33552 Latino. Gene assoc with stickler syndrome (AR) - Bones are affected by slight platyspondylisis and large, often defective epiphyses. Also gene may be assoc with MED6 (AD) -Joint pain, joint deformity, waddling gait, and short stature are the main clinical signs and symptoms. Radiological examination of the skeleton shows delayed, irregular mineralization of the epiphyseal ossification centers and of the centers of the carpal and tarsal bones. NO second likely pathogenic mutation identified.	Missense variant	A/G	1/NA/NA
			3	33155582	<i>CRTAP</i>	Not reported. c.13C>A;p.(Arg5Ser). Monoallelic. Mainly LOF mutations described - unlikely to be causative in this individual. No likely 2nd mutation in VCF.	Missense variant	C/A	1/NA/NA
			X	48652207	<i>GATA1</i>	Disease - X-linked dyserythropoietic anaemia and thrombocytopenia	Missense variant	G/A	1/NA/NA

			12	52200885	SCN8A	Reported missense_variant. Disease=Epileptic encephalopathy, early infantile, 13 (EIEE13) a disorder characterized by markedly delayed cognitive and motor development, attention deficit disorder, and cerebella ataxia. c.5615G>A;p.(Arg1872Gln) reported pathogenic mutation in assoc with AD epileptic encephalopathy	Missense variant	G/A	1/NA/NA
Few cafe-au-lait spots, Gait ataxia, Generalized hypotonia, Global developmental delay, Pectus excavatum, Preaxial hand polydactyly, Recurrent fractures	115	265512	X	118975168	UPF3B	Reported de novo frameshift_variant. Assoc with X-linked recessive mental retardation	Frameshift variant	TTTTCT	2/0/0
Abnormality of eye movement, Agammaglobulinemia, Feeding difficulties in infancy, Gastroesophageal reflux, Generalized myoclonic seizures, Global developmental delay, Infantile axial hypotonia, Inguinal hernia, Multiple rib fractures, Optic disc pallor, Recurrent respiratory infections, Vertebral compression fractures, Wormian bones	116	265802	22	41573186	COL10A1	COL10A1 c.880G>A;p.(Gly294Arg) - in triple helix but this is not where mutations are reported in MED. Also other glycine substitutions reported in gnomAD data = doesn't support pathogenicity. COL10A1=skeletal dysplasia gene (panel app)	Missense variant	G/A	1/0/0
Abnormal facial shape, Constipation, Global developmental delay, Increased susceptibility to fractures, Microcephaly, Polyhydramnios, Sleep disturbance, Unilateral deafness	117	265826	17	80403807	C17orf62	c.231G>A, synonymous change in major transcript, no splice effect predicted. V low frequency 0.0048%.	Missense variant	C/T	1/NA/NA
			12	75698578	CAPS2	c.794_795del;p.(Met265Asnfs*6) - candidate gene for intellectual disability.	Frameshift variant	CA/C	1/NA/NA
			2	238285643	COL6A3	AD/AR Bethlem myopathy c.2842G>T;p.(Val948Phe). Only 0.0015% in	Missense variant	C/A	1/NA/NA

				Eur pop (1 allele).					
			16	67645346	<i>CTCF</i>	c.612_615del;p.(Lys206fs). Reported frameshift_variant, Mental retardation, autosomal dominant 21	frameshift	CAAAG/C	1/NA/NA
			12	49422934	<i>KMT2D</i>	c.1416C>T;p.(Arg4721Cys) Reported missense_variant. Gene associated with AD Kabuki syndrome = congenital mental retardation syndrome with additional features, including postnatal dwarfism, a peculiar facies characterized by long palpebral fissures with eversion of the lateral third of the lower eyelids, a broad and depressed nasal tip, large prominent earlobes, a cleft or high-arched palate, scoliosis, short fifth finger, persistence of finger pads, radiographic abnormalities of the vertebrae, hands, and hip joints, and recurrent otitis media in infancy	Missense variant	G/A	1/NA/NA
			2	15506745	<i>NBAS</i>	c.3776C>T; p.(Ser1259Phe), low quality score in VCF. ? False positive. No other likely 2nd mutation in VCF. Clinical features not suggestive of this gene being causative	Missense variant	G/A	1/NA/NA
			11	70798909	<i>SHANK2</i>	c.64G>T;p.(Ala22Ser). Reported missense_variant. Gene associated with autism susceptibility	Missense variant		
			20	46319020	<i>SULF2</i>	c.587T>C;p.(Ile196Thr). No frequency data. GO biological processes include: bone development, heparan sulfate proteoglycan metabolic process, positive regulation of WNT signalling. Look at GeneAtlas function info. Gene also in another patient.	Missense variant	A/G	1/NA/NA

						PolyPhen=possibly damaging(0.809), SIFT=deleterious(0.03)			
			2	128910450	UGGT1	c.371Tdel;p.(Phe124fs) From UniProt: Recognizes glycoproteins with minor folding defects. Reglucosylates single N-glycans near the misfolded part of the protein, thus providing quality control for protein folding in the endoplasmic reticulum	Frameshift variant	A/AC	1/NA/NA
Bilateral sensor neural hearing impairment, Bruising susceptibility, Cavernous hemangioma, Deeply set eye, Highly arched eyebrow, Intellectual disability, Mandibular prognathia, Mild postnatal growth retardation, Osteopenia, Petechiae, Short lingual frenulum	118	265937	22	41573186	EP300	c.5471A>C;p.(Gln1824Pro). Reported De-novo AD Rubinstein-Taybi syndrome 2 (RSTS2) craniofacial abnormalities, postnatal growth deficiency, broad thumbs, broad big toes, mental retardation and a propensity for development of malignancies. Decipher = likely pathogenic and partially explains clinical phenotype.	Missense variant	A/C	1/0/0
Absence seizures, Absent speech, Camptodactyly of finger, Episodic hypoventilation, Generalized tonic-clonic seizures, Global developmental delay, Hyperacusis, Intellectual disability, profound, Knee flexion contracture, Kyphosis, Motor delay, Narrow foot, Osteoporosis, Short foot	119	269622							
Cerebellar hypoplasia, Cortical visual impairment, Hip dislocation, Microcephaly, Recurrent pneumonia, Scoliosis, Seizures, Severe generalized osteoporosis, Severe global developmental delay	120	270466	X	41674114	41788517	CNV reported Ratio -1.783055. Loss. Gene is CASK, associated with XLD Mental retardation and microcephaly with pontine and cerebellar hypoplasia.	Large deletion		1/0/0

Adducted thumb, Cubitus valgus, Dysarthria, Gait ataxia, Gait disturbance, High palate, Hypocalcemia, Increased susceptibility to fractures, Long face, Long fingers, Long foot, Osteoporosis, Pes planus, Reduced bone mineral density, Scoliosis, Slender finger, Tall stature	121	270572	X	153999050	<i>DKC1</i>	c.932A>T;p.(Tyr311Phe). Reported - missense_variant, brother of DDDP107936. Dyskeratosis congenica, X-linked (DKCX). A rare, progressive bone marrow failure syndrome characterized by the triad of reticulated skin hyperpigmentation, nail dystrophy, and mucosal leukoplakia. Early mortality is often associated with bone marrow failure, infections, fatal pulmonary complications, or malignancy. Symptoms can include osteoporosis - HPO:0002659 - susceptibility to fractures.	Missense variant	A/T	2/1/0
			21	44486410	<i>CBS</i>	Also heterozygous for CBS c.394C>T;p.(Arg132Cys). ClinVar submission - likely pathogenic (updated 2nd May 2016 - now uncertain significance). Very low frequency data. AR homocysteineuria. Affected individuals are often tall and slender and have Marfanoid habitus. Prone to osteoporosis. Some controversy about whether carriers have an increased risk of cardiovascular disease.	Missense variant	C/T	1/0/1
			1	43213881	<i>P3H1</i>	Also heterozygous for <i>P3H1</i> c.1828C>T;p.(Arg610Cys). Only found in 2 alleles in 119078. Polyphen pathogenic. Brother does not have this variant.	Missense variant	C/T	1/1/0
			20	60892813	<i>LAMA5</i>	c.7261C>T;p.(Arg2421Trp). Present in 97/269372 alleles but no homozygotes. GeneAtlas: function includes: stimulates osteogenic gene expression in mesenchymal stem cells, in the absence of any other osteogenic stimulus (Klees 2007) • activates	Missense variant	G/A	1/0/1

						extracellular matrix production and osteogenic gene focusing in human mesenchymal stem cells (Klees 2007)			
			20	60901773	LAMA5	c.5258C>T;p.(Thr1753Met). Present in 52/276498 allele and one homozygote	Missense variant	G/A	1/1/0
Athetoid cerebral palsy, Gait disturbance, High palate, Increased susceptibility to fractures, Long foot, Macrotia, Moderate global developmental delay, Osteoarthritis, Osteopenia, Relative macrocephaly, Slender finger, Tall stature	122	270575	5	14751270	ANKH	Reported De-novo initiator_codon_variant UV c.595A>G;p.(Met199Val) Autosomal dominant inheritance of Craniometaphyseal dysplasia and chondrocalcinosis	Missense variant	T/C	1/0/0
			X	153999050	DKC1	missense_variant, brother of DDDP100275	Missense variant	A/T	2/1/0
			20	60892813	LAMA5	See brother's results above (patient 121).	Missense variant	G/A	1/0/1
			20	60901773	LAMA5	See brother's results above (patient 121).	Missense variant	G/A	1/1/0
Abnormality of the pinna, Alopecia of scalp, Congenital sensorineural hearing impairment, Fractures of the long bones, Premature ovarian failure, Short stature	123	271487	21	47783795	PCNT	Mutations in this gene cause Seckel syndrome-4 and microcephalic osteodysplastic primordial dwarfism type II.	In frame insertion	C/CCTT	1/1/0
			21	47783796	PCNT	AR inheritance - this region of gene has deletions reported in GnomAD, these changes both 0.028%. These two changes are called next to each other and probably represent one change that is miss called. Both in-frame. Unlikely pathogenic.	Inframe deletion	TGCG/T	1/1/1

Generalized seizures, Growth hormone deficiency, Hip dislocation, Moderate generalized osteoporosis, Severe global developmental delay, Short stature, Thoracolumbar scoliosis	124	274096	1	27835713	28494006	CNV reported. Ratio -0.764081 LOSS. Genes in region include AHDC1 associated with Xia-Gibbs syndrome - hypotonia, and delayed psychomotor development with absent or poor expressive language. Mild dysmorphic features were present, including low-set or protuberant ears, flat nasal bridge, hypertelorism, up- or down slanting palpebral fissures, and micrognathia. Also sleep apnoea.	Large deletion		
Astigmatism, Deeply set eye, Delayed speech and language development, Gastroesophageal reflux, Global developmental delay, High anterior hairline, Increased susceptibility to fractures, Large for gestational age, Narrow chest, Poor gross motor coordination, Seizures, Triangular face	125	274310	1	153800509	<i>GATAD2B</i>	c.314_315insA;p.(Asp106fs). Reported frameshift_variant. AD mental retardation with dysmorphic features including hypertelorism and narrow palpebral fissures. LoF mutations	Frameshift variant	C/CT	1/0/0
			3	168834157	<i>MECOM</i>	AD Radioulnar synostosis with amegakaryocytic thrombocytopenia 2	Missense variant	G/C	1/1/0
			3	168840462	<i>MECOM</i>	AD Radioulnar synostosis with amegakaryocytic thrombocytopenia 2	Missense variant	T/C	1/0/1
Autistic behaviour, Easy fatigability, Hemihypotrophy of lower limb, Narrow chest, Osteopenia, Rocker bottom foot, Short lower limbs, Tapered finger	126	274716							
Campodactyly of finger, Dislocated wrist, Osteoporosis, Pes cavus, Scoliosis, Severe Myopia	127	276443			<i>PLOD1</i>	Osteoporosis, scoliosis, pes cavus, severe myopia. No fractures. Heterozygous for the c.153dup;p.(Asn52fs). Known mutation associated with kyphoscoliotic EDS. Can we test for common duplication in this patient or do urine cross-links?	frameshift,	G/A	1/0/0

Abnormality of blood circulation, Aortic regurgitation, Asthma, Bronchiectasis, Fine, reticulate skin pigmentation, Generalized osteoporosis with pathologic fractures, Genu valgum, Glaucoma, Migraine, Myopia, Otosclerosis, Severe generalized osteoporosis, Short foot, Slender build, Slender finger, Slender toe, Thin skin	128	276614	9	123808662	C5	Splice donor variant c.66-1 G>T NM_001735.2 inherited from mother. Complement component 5 deficiency (C5D) autosomal dominant. A rare defect of the complement classical pathway associated with susceptibility to severe recurrent infections, predominantly by Neisseria gonorrhoea or Neisseria meningitidis. NO report of any bone phenotype - probably incidental finding.	splice_variant	C/A	1/1/0
			16	89259925	CDH15	c.1904del;p.(Gly636fs) Reported by DDD missense_variant. AD autism gene.	Frameshift	CG/C	1/1/0
			14	21896152	CHD8	Gene associated with Mental Retardation, Autosomal Dominant 3; MRD3	Missense variant	G/T	1/1/0
			11	103026129	DYNC2H1	AR inheritance - short-rib thoracic dysplasia-3 with or without polydactyly (SRTD3; 613091).	Missense variant	C/T	1/0/1
			21	46021438	KRTAP10-7	Gene known as TSPEAR. Involvement in disease Deafness, autosomal recessive, 98 (DFNB98)1 Publication A form of non-syndromic sensorineural hearing loss with prelingual onset. Sensorineural deafness results from damage to the neural receptors of the inner ear, the nerve pathways to the brain, or the area of the brain that receives sound information. NOT listed as associated with bone phenotype.	Missense variant	C/T	1/1/0
			17	79268721	SLC38A10	? Candidate gene UniProt Go biological process: bone development.c.1A>G;p.(Met1?) SLC38A10=Mice weak + brittle gene. GnomAD allele frequency 4 in 118,638 but pLI= 0.00 The	Initiator codon variant	T/C	1/1/0

						closer pLI is to one, the more LoF intolerant the gene. Maternally inherited, mum has osteoporosis, otosclerosis, abnormal kidney			
			2	96955598	<i>SNRNP200</i>	c.2879C>T;p.(Ala960Val). UniProt GO biological process=osteoblast differentiation. Gene associated with Retinitis Pigmentas 33 (autosomal dominant). This gene 1-2%of AD RP. Missense mutations in HGMD - this variant listed as unknown significance (Ellingford et al J Med Genet. 2016 Nov; 53(11): 761–767). Our patient has some optic clinical features but variant inherited from mum who doesn't have any optical features.	Missense variant	G/A	1/1/0
Dilation of lateral ventricles, Generalized joint laxity, Intellectual disability, moderate, Postnatal macrocephaly, Recurrent fractures	129	277040	19	50436333	<i>ATF5</i>	c.833G>A;p.(Arg278Lys). UniProt: Function: May be involved in osteogenic differentiation (PubMed:22442021). Inherited from mum who has no symptoms.	Missense variant	G/A	1/0/1
			1	9804609	<i>CLSTN1</i>	Induces KLC1 association with vesicles and functions as a cargo in axonal transport	Missense variant	C/T	1/0/1
			22	41546158	<i>EP300</i>	Freq data =benign	Missense variant	C/A	1/0/1
			22	41574510	<i>EP300</i>	Freq data =benign	Intron variant	TCAG/T	1/1/0
			7	50514031	<i>FIGNL1</i>	c.954del;p.(Ser319fs). UniProt function: May regulate osteoblast proliferation and differentiation. Inherited pLI=0.00	Frameshift variant	AC/A	1/1/0

			11	102707247	<i>MMP3</i>	GO function: Can degrade fibronectin, laminin, gelatins of type I, III, IV, and V; collagens III, IV, X, and IX, and cartilage proteoglycans. Activates procollagenase. Polymorphism in gene associated with increased risk of coronary artery disease. BUT this a deeply intronic change with allele frequency 0.0038%	Regulatory region variant	G/A	1/1/0
			2	167055749	<i>SCN9A</i>	c.5367_5368insGAGAACTCTAT;p.(Lys1790fs).Pain disorders (AR); epilepsy with febrile seizures (AD); primary erythralgia (AD). Inherited from mum - ? LOF mutations not pathogenic. No 2nd likely mutation identified.	Frameshift variant	A/AGAGAACTCTAT	1/1/0
			16	3642722	<i>SLX4</i>	c.2305C>G. Fanconi anaemia (P) Some individuals affected by Fanconi anaemia of complementation group P have skeletal anomalies. Clinical features of patient don't indicate Fanconi's phenotype. 267/276130 alleles, one homozygotes.	Missense variant	C/G	1/0/1
			16	3645688	<i>SLX4</i>	c.1931A>C;p.(Asp644Ala) no frequency data.	Missense variant	T/G	1/1/0
Aplasia/Hypoplasia of the patella, Fractures of the long bones, Limited elbow flexion/extension, Proximal placement of thumb, Radial head subluxation	130	277078			<i>MYO7</i>	c.731G>A; p.(Arg244His) associated with non-syndromal deafness and retinosa pigmentosa (RP). None of these symptoms reported in this patient.	Missense variant	C/T	2/1/1
			8	145740620	<i>RECQL4</i>	c.1379C>T;p.(Pro466Leu). Reported missense_variant. Rothmund–Thomson (RTS), RAPADILINO and Baller–Gerold (BGS), are characterized by growth retardation and radial defects, but RAPADILINO syndrome lacks the main dermal manifestation, poikiloderma that is	Missense variant	G/A	1/0/1

						a hallmark feature in both RTS and BGS			
			8	145741453	RECQL4	c.1048_1049del;p.(Arg350fs). Reported frameshift_variant. Compound heterozygote - considered to fully explain patients phenotype in Decipher.	Frameshift	CCT/C	1/1/0
Broad-based gait, Cone/cone-rod dystrophy, Global developmental delay, Horizontal pendular nystagmus, Open mouth, Osteoporosis, Reduced visual acuity, Seizures, Short philtrum, Ventriculomegaly, Wide nose	131	277867	15	89382022	ACAN	c.199G>A;p.(Ala67Thr). Reported probable SNP 0.19% AD Osteochondritis dissecans, short stature, and early-onset osteoarthritis. Clinical phenotype doesn't fit.	Missense variant	G/A	1/0/1
			15	89386826	ACAN	c.998A>G;p.(Asn333Ser) AD Spondyloepimetaphyseal dysplasia, aggrecan type and Kimberley type. Inherited from unaffected parent. 0.0056% in European. Skeletal dysplasia characterized by shortening of the trunk and limbs	Missense variant	A/G	1/1/0
Anteverted nares, Aplasia/Hypoplasia of fingers, Aplasia/Hypoplasia of the distal phalanges of the toes, Atrioventricular canal defect, Cone-shaped epiphysis, Delayed speech and language development, Gastrostomy tube feeding in infancy, Global developmental delay, Growth abnormality, High anterior hairline, Intrauterine growth retardation, Long philtrum, Nasogastric tube feeding in	132	278587	20	57478742	GNAS	c.2257A>G;p.(Met753Val). Reported missense_variant. McCune-Albright syndrome = symptoms include: Pathologic fracture or bone deformity may be presenting manifestations and pseudarthrosis occurs frequently. Deafness and blindness can result from impingement of the bony process on the cranial foramina.	Missense variant	A/G	1/0/0

infancy, Osteopenia, Prenatal movement abnormality, Recurrent respiratory infections, Short stature, Synophrys									
Abnormal emotion/affect behaviour, Autistic behaviour, Blue sclerae, Fractures of the long bones, Muscular hypotonia, Short stature, Sleep disturbance	133	281522	19	18893921	COMP	c.2170G>A;p.(Val724Ile0. Reported missense_variant. Gene associated with MED1. COMP protein is a pentameric extracellular matrix protein that catalyzes the assembly of collagens and promotes formation of well-defined fibrils. Present in 0.006% of AA population in gnomAD. Phenotype includes: skeletal dysplasias pseudoachondroplasia and multiple epiphyseal dysplasia	Missense variant	C/T	1/1/0
			22	41573804	EP300	c.6089A>G;p.(Gln2030Arg). Reported missense_variant. No frequency data.	Missense variant	A/G	1/1/0
			4	5720973	EVC	c.175-2A>G. Reported as disease causing mutation ins Ellis-van Creveld syndrome (EVC)(Valencia (2009) Hum Mutat 30: 1667). An autosomal recessive condition characterized by the clinical tetrad of chondrodystrophy, polydactyly, ectodermal dysplasia and cardiac anomalies. Patients manifest short-limb dwarfism, short ribs, postaxial polydactyly, and dysplastic nails and teeth. Congenital heart defects, most commonly an atrioventricular septal defect, are observed in 60% of affected individuals. None of these features lists in patient phenotype. No	Splice variant	A/G	1/1/0

					2nd likely mutation in VCF.				
			17	36891739	PCGF2	c.770_771del;p.(Glu57fs). Reported frameshift_variant. The Deciphering Developmental Disorders Study (2015) identified 2 unrelated patients with intellectual disability who shared the same heterozygous missense mutation in the PCGF2 gene (this patient and a second). No indication of skeletal phenotype in the second patient	Frameshift	ACT/A	1/1/0
			9	2039799	SMARCA2	Reported missense_variant. AD Nicolaides-Baraitser syndrome. A rare disorder characterized by severe mental retardation with absent or limited speech, seizures, short stature, sparse hair, typical facial characteristics, brachydactyly, prominent finger joints and broad distal phalanges	Missense variant	AGC/C GC	1/1/0
Abnormal renal morphology, Abnormality of the liver, Bladder diverticulum, Congenital hip dislocation, Hydronephrosis, Motor delay, Osteopenia, Pulmonic stenosis, Short stature	134	282638							

Abnormality of dentin, Absent speech, Autistic behavior, Cognitive impairment, Drooling, Fractures of the long bones, Microcephaly, Myopathic facies, Neonatal hypotonia, Open mouth, Osteopenia, Reduced bone mineral density, Seizures, Strabismus	135	283516							
Bowling of limbs due to multiple fractures, Congenital contracture, Decreased antibody level in blood, Multiple prenatal fractures	136	283921	16	30750390	<i>SRCAP</i>	Reported missense_variant. C.9029C>A;p.(Pro3010His)	Missense variant	C/A	1/0/0
			6	33272115	<i>TAPBP</i>	Bi-allelic loss of function mutations in <i>TAPBP</i> are associated with Type I Bare Lymphocyte Syndrome, a condition characterised by chronic bacterial infection.	Missense variant	G/A	1/0/0
Bowling of limbs due to multiple fractures, Congenital contracture, Decreased antibody level in blood, Multiple prenatal fractures	137	285317	7	94033866	<i>COL1A2</i>	Not reported by DDD c.280-2A>C. Not reported in Decipher database. Inherited from mum who is clinically affected.	Splice variant	A/C	1/1/0
			7	31011710	<i>GHRHR</i>	c.597T>C;p.(Thr199Thr), synonymous change. Gene associated with isolated growth hormone deficiency.	Missense variant	T/C	1/1/0
			2	220434949	<i>OBSL1</i>	3m syndrome - recessive.	Missense variant	C/A	1/0/0
			14	57269042	<i>OTX2</i>	c.305G>A;p.(Arg102His). Reported missense_variant. Retinal Dystrophy with or without Pituitary Dysfunction. Gene probably reported due to 'blue sclerae' being listed under 'eye abnormality' in DDD.	Missense variant	G/A	1/1/0

			5	86564716	RASA1	c.448C>T;p.(Leu150Phe). Reported missense_variant. Capillary malformation-arteriovenous malformation. Parkes Weber syndrome is characterized by a cutaneous flush with underlying multiple micro-AVFs (arteriovenous fistulas), in association with soft tissue and skeletal hypertrophy of the affected limb. Low level frequency data in gnomAD.	Missense variant	C/T	1/1/0
Abnormal aggressive, impulsive or violent behaviour, Fractures of the long bones, Severe expressive language delay, Smooth philtrum, Tetralogy of Fallot, Thin vermilion border	138	286790							
Macroglossia, Moderate global developmental delay, Multiple exostoses, Neonatal respiratory distress, Osteopenia, Postnatal macrocephaly, Seizures, Sparse scalp hair	139	287325	X	39934294	BCOR	c.305G>A;p.(Arg102Gln) Reported missense_variant. X-linked Microphthalmia, syndromic 2. Multiple congenital anomaly syndrome characterized by eye anomalies, facial abnormalities, cardiac anomalies and dental abnormalities.	Missense variant	G/A	1/1/0
			10	28824551	WAC	c.139C>T;p.(Arg47*) Reported stop_gained. Previously reported ?cause of DD	Stop gained	C/T	1/0/0
Phenotype: Abnormality of the odontoid process, Amelogenesis imperfecta, Cerebral calcification, Depressed nasal bridge, Down slanted palpebral fissures, Elevated circulating parathyroid hormone (PTH) level, Hypophosphatemia, Malar flattening, Narrow mouth, Osteoporosis, Sagittal craniosynostosis, Subglottic stenosis	140	294005	7	295967	FAM20C	c.1225C>T;p.(Arg409Cys). Reported missense_variant. Protein = plays a key role in biomineralization of bones and teeth. Raine syndrome (RNS)= Autosomal recessive osteosclerotic bone dysplasia with neonatal lethal outcome. Clinical features include generalized osteosclerosis, craniofacial dysplasia and microcephaly. Reported pathogenic mutation in Alazami et al 2015 but	Missense variant	C/T	2/1/1

						this is an exome sequencing project and no functional evidence provided. Full explanation of phenotype reported in Decipher.			
			16	86602355	FOXC2	Disorder characterized by primary limb lymphedema associated with distichiasis (double rows of eyelashes, with extra eyelashes growing from the Meibomian gland orifices)	Upstream variant	A/G	2/1/1
			14	58934462	KIAA0586	Reported missense_variant. Gene associated with Joubert syndrome 23 (phenotype was relatively mild and homogeneous, consisting mainly of neurologic features, such as delayed development, abnormal eye movements, and the molar tooth sign on brain imaging) and short rib-polydactyly syndrome, type IX (severe hydrocephaly, polydactyly, and skeletal abnormalities) which is usually lethal due to thoracic insufficiency. Likely benign in Decipher.	Missense variant	A/G	2/1/1
			12	12277566	LRP6	c.4481dup;p.(Arg1495fs).Cell-surface coreceptor of Wnt/beta-catenin signalling, which plays a pivotal role in bone formation. Gene associated with tooth agenesis, selective, 7 (STHAG7), autosomal dominant form of selective tooth agenesis, a common anomaly characterized by the congenital absence of one or more teeth.	Frameshift variant	C/CT	1/0/0
Abnormality of mast cells, Autonomic dysregulation, Joint hypermobility, Orthostatic hypotension, Osteopenia, Seizures, obsolete Joint dislocations in	141	294027	6	7727522	BMP6	c.334G>A;p.(Glu112Lys) .Go biological processes - bone mineralization, osteoblast differentiation. Likely polymorphic variant as in del/ins in dbSNP	Missense variant	GAGCA GCAGC /AAGCA GCAGC	1/1/0

young adult			5	60224690	<i>ERCC8</i>	AR Cockayne syndrome = rare disorder characterized by cutaneous sensitivity to sunlight, abnormal and slow growth, cachectic dwarfism, progeroid appearance, progressive pigmentary retinopathy and sensorineural deafness	Splice variant	C/G	1/1/0
			12	49418641	<i>KMT2D</i>	c.1587G>C;p.(Glu5291Asp). Reported - missense_variant. Kabuki syndrome = congenital mental retardation syndrome with additional features, including postnatal dwarfism, a peculiar facies characterized by long palpebral fissures with eversion of the lateral third of the lower eyelids, a broad and depressed nasal tip, large prominent earlobes, a cleft or high-arched palate, scoliosis, short fifth finger, persistence of fingerpads, radiographic abnormalities of the vertebrae, hands, and hip joints, and recurrent otitis media in infancy. Likely benign in Decipher.	Missense variant	G/C	1/1/0
			13	38158851	<i>POSTN</i>	Variant c.1108+2T> C. Gene = periostin, osteoblast specific factor. Involved in bone metabolism. Good candidate gene. UniProt - Enhances incorporation of BMP1 in the fibronectin matrix of connective tissues, and subsequent proteolytic activation of lysyl oxidase LOX. Variant c.1108+2T> C, 0.0027% in gnomAD. Inherited from mum who has Joint hypermobility, Osteopenia, Autonomic dysregulation	Splice variant	A/G	1/1/0
Bowing of limbs due to multiple fractures, Fractures of the long bones,	142	294357							

Increased susceptibility to fractures, Osteopenia									
Aplasia/Hypoplasia involving the musculature of the upper limbs, Aplasia/Hypoplasia of the thumb, Delayed gross motor development, Pathologic fracture, Protruding ear, Recurrent bacterial infections	143	295131							
High palate, Multiple small vertebral fractures, Osteoporosis, Proportionate tall stature	144	295328							
Flexion contracture of toe, Fractures of the long bones, High pitched voice, Increased hepatic glycogen content, Intellectual disability, moderate, Low-set ears, Moderately short stature, Patellar dislocation, Pelvic kidney, Prominent epicanthal folds, Scoliosis, Short foot, Upslanted palpebral fissure, Vesicoureteral reflux	145	301078	2	216226780	<i>FN1</i>	c.7274G>A;p.(Arg2452His). GeneAtlas Basic function, multiple but includes: upregulates osteoclast activity despite inhibiting osteoclast formation and these effects involve nitric oxide and IL1B signalling. Present in 0.2% of GoNL pop. No homozygotes. Gene associated with AD plasma fibronectin deficiency.	Missense variant	C/T	1/0/1
			2	216240047	<i>FN1</i>	c.6047C>T; p.(Pro2016Leu). 0.06% in gnomAD, no homozygotes	Missense variant	G/A	1/1/0
Apnea, Autistic behavior, Clinodactyly of the 5th finger, Feeding difficulties, Focal seizures, Gait imbalance, Generalized tonic-clonic seizures, Impaired pain sensation, Increased susceptibility to fractures, Intellectual disability, severe, Long philtrum, Micrognathia, Moderate global	146	301415	15	93467724	<i>CHD2</i>	c.236T>C;p.(Leu79Pro). Reported by DDD missense_variant. A severe form of epilepsy characterized by onset of multiple seizure types in the first few years of life and associated with poor prognosis. Affected individuals have cognitive regression and intellectual disability	Missense variant	T/C	1/NA/NA
			21	38877655	<i>DYRK1A</i>	c.1309C>T;p.(Arg437Ter). Reported stop_gained. AD mental retardation type 7.	Missense variant	C/T	1/NA/NA

developmental delay, Optic nerve hypoplasia, Painless fractures due to injury, Posteriorly rotated ears, Sensorineural hearing impairment, Severe expressive language delay, Short nose, Stereotypic behaviour, Typical absence seizures			6	85446617	<i>TBX18</i>	c.1610C>G;p.(Ser537Cys). Reported missense_variant. Gene associated with AD congenital anomalies of the kidney and urinary tract. No skeletal phenotype listed associated with this condition. Likely benign in Decipher.	Missense variant	C/G	1/NA/NA
			21	38877655	<i>DYRK1A</i>	c.1309C>T;p.(Arg437Ter). Reported stop_gained. AD mental retardation type 7.	Stop gained	C/T	1/NA/NA
			11	76868371	<i>MYO7A</i>	c.782G>A; p.(Gly261Asp). AR Usher syndrome. profound congenital hearing impairment with unintelligible speech, early retinitis pigmentosa (usually evident within the first decade), and constant vestibular dysfunction	Missense variant	G/A	1/NA/NA
Abnormality of the temporomandibular joint, Absent distal interphalangeal creases, Camptodactyly of finger, Hashimoto thyroiditis, Increased susceptibility to fractures, Lateral displacement of patellae, Myopia, Short neck	147	301797							
Abnormality of the teeth, Bone pain, Delayed eruption of teeth, Fractures of the long bones, High palate, Joint hypermobility, Large posterior fontanelle, Narrow chest, Wide anterior fontanel	148	304402	3	176756186	<i>TBL1XR1</i>	c.962C>G;p.(Thr321Ser). Reported missense_variant. AD mental retardation type 41. A form of mental retardation, a disorder characterized by significantly below average general intellectual functioning associated with impairments in adaptive behaviour and manifested during the developmental period. No skeletal manifestations.	Missense variant	G/C	1/0/0

2-3 toe syndactyly, Abnormality of dental enamel, Blue sclerae, Bruising susceptibility, Constipation, Delayed gross motor development, Delayed speech and language development, Down slanted palpebral fissures, Increased muscle fatigability, Intellectual disability, severe, Joint hypermobility, Malar flattening, Recurrent fractures	149	304572	17	48272592	COL1A1	Reported splice_variant c.1299+1G>A	Splice variant	G/A	1/0/0
--	-----	--------	----	----------	--------	-------------------------------------	----------------	-----	-------

* Trio Genotype column indicates the presence of the variant in the proband/mother/ father, where:

1 = present on one allele

2 = present on both alleles

0 = not present.

NA indicates that the parent was not available for analysis

C. Poster: International Skeletal Dysplasia Society (2017)

DOES COLE-CARPENTER TYPE 1 HAVE A DISTINCT SKELETAL PHENOTYPE?

Balasubramanian M¹, Padidela R², Pollitt RC^{1,3}, Bishop NJ^{1,3}, Wagner BE⁴, McCaughey J⁵, Stephens JD⁵, Offiah AC^{1,3}

1. Sheffield Children's Hospital NHS Foundation Trust 2. Royal Manchester Children's Hospital 3. University of Sheffield 4. Sheffield Teaching Hospitals NHS Foundation Trust 5. University of Bristol



BACKGROUND (1)

- Cole & Carpenter¹ first described a unique phenotype in two unrelated infants including:
 - Normal intelligence
 - Growth failure
 - Characteristic facies, blue sclerae, proptosis, micrognathia
 - Craniosynostosis, hydrocephalus
 - "Distinctive metaphyseal lucencies/compression fractures"
 - Osteopaenia and diaphyseal fractures and deformity consistent with severe osteogenesis imperfecta (OI)
- Since the paper by Cole & Carpenter, there have been three further published cases with almost identical clinical phenotype: Amor et al², Balasubramanian et al³, Garbes et al⁴
- Since the paper by Cole & Carpenter, there have been four further published cases with atypical clinical phenotype: Garbes et al⁴, Marwaha et al⁵, MacDermot et al⁶
- Table 1 summarises the 9 cases reported to date, including the abnormal gene [References 3, 4, 7]

AIM

- To present a tenth case of Cole-Carpenter syndrome and to demonstrate what we believe to be the characteristic radiological phenotype of the *P4HB* mutation

CASE PRESENTATION (1)

- The proband is the first child of healthy non-consanguineous Caucasian parents
- Pregnancy was unremarkable with no antenatal signs of bone fragility
- At 6 months of age, she presented with left tibial deformity
- Examination revealed failure to thrive, macrocephaly with a wide open anterior fontanelle, blue sclerae, small nose and a flat nasal bridge (Fig. 1)

FIG. 1: PROBAND AT 6 MONTHS (A) AND 3 YEARS (B) OF AGE



- A skeletal survey showed generalised osteopenia with Wormian bones and multiple vertebral and rib fractures. Her long bones were slender with unusual metadiaphyseal compression fractures of upper (Fig. 2) and lower limbs with widening and sclerosis of the distal metaphyses
- By age 2.5 years, there was considerable bowing deformity with persistent (but less pronounced) metaphyseal sclerosis and irregularity (Fig. 3)
- Targeted OI panel testing (all published dominant & recessive OI genes) and dosage analysis of *COL1A1/A2* genes using *MLPA* was negative
- Whole exome sequencing confirmed a *P4HB* mutation

DISCUSSION

- This is only the tenth reported case of Cole-Carpenter syndrome (CCS), a rare form of OI with characteristic clinical features, craniosynostosis and hydrocephalus
- The first 2 reported cases were unrelated and in addition to the characteristic clinical features, multiple diaphyseal fractures and progressive long bone deformity, also had metaphyseal lucencies and compression fractures. Both patients had the same heterozygous missense mutation in exon 9 of *P4HB* (NM_000918.3:c.1178A>G [p.Tyr393Cys]), the gene that encodes protein disulfide isomerase⁷. This is now known as CCS 1 (CLCRP1 OMIM #11240)
- Two other patients with the characteristic CCS clinical phenotype but NOT the metaphyseal compression fractures have had *CRTAP*³ and *SEC24D*² mutations
- SEC24D* mutations have also been confirmed in two other fetuses (female siblings) with similar CCS features but uncharacteristic severe calvarial ossification defects and no metadiaphyseal compression fractures in the one fetus that was imaged⁴. This is now known as CCS 2 (CLCRP2 OMIM #616294)
- As far as we know, causative mutations have not been reported in the 3 other published CCS patients. The patient reported by Amor et al² has the typical phenotype and irregular widened metaphyses of CCS1, but no metadiaphyseal impact fractures. The other two have atypical features; small hands and feet, no fractures⁵ and severe hydrops fetalis, no Wormian bones and normal metaphyses⁶
- The case presented here, is only the third to have BOTH the characteristic phenotype of CCS1 and the metadiaphyseal compression fractures described in the original cases. She is also the only other reported case with the same missense mutation in *P4HB*

CONCLUSION

- We have not seen similar metadiaphyseal fractures or *P4HB* mutations amongst our large cohort of OI patients (xx with typical OI and yy with atypical OI)
- We therefore suggest that CCS1 (due to a recurrent *P4HB* missense mutation) has a distinct radiological phenotype
- More cases are required for confirmation

REFERENCES

- Cole DEC, Carpenter TO Bone fragility, craniosynostosis, ocular proptosis, hydrocephalus and distinctive facial features: A newly recognized type of osteogenesis imperfecta J Pediatr 1987;110:76-80
- Amor DJ, Sivarirayan R, Schneider AS, Bankier A New case of Cole-Carpenter syndrome Am J Med Genet 2000;92:273-277
- Balasubramanian M, Pollitt RC, Chandler KE et al *CRTAP* mutation in a patient with Cole-Carpenter syndrome Am J Med Genet 2015;167A:587-591
- Garbes L, Kim K, Riess A et al Mutations in *SEC24D*, Encoding a component of the COPII machinery, cause a syndromic form of osteogenesis imperfecta AM J Hum Genet 2015;96:432-439
- Marwaha RK, Sarkar B, Kataria S, Jaysree K Cole-Carpenter's syndrome Ind J Pediatr 1993;60:305-308
- MacDermot KD, Buckley B, Van Someren V Osteopenia, abnormal dentition, hydrops fetalis and communicating hydrocephalus Clin Genet 1995;48:217-220
- Rauch F, Fahiminiya S, Majewski J et al Cole-Carpenter syndrome is caused by a heterozygous missense mutation in *P4HB* Am J Hum Genet 2015;96:425-431

BACKGROUND (2)

TABLE 1: SUMMARY OF 9 REPORTED CASES

Author	Typical clinical phenotype	MCF*	Gene
Cole (1987)	Yes	Yes	<i>P4HB</i> ¹
Cole (1987)	Yes	Yes	<i>P4HB</i> ¹
Marwaha (1993)	No	No	?
MacDermott (1995)	No	No	?
Amor (2000)	Yes	No	?
Balasubramanian (2015)	Yes	No	<i>CRTAP</i>
Garbes (2015)	Yes	No	<i>SEC24D</i>
Garbes (2015)	No (Fetus)	No	<i>SEC24D</i>
Garbes (2015)	No (Fetus)	?	<i>SEC24D</i>

* MCF - metaphyseal compression fracture(s)

CASE PRESENTATION (2)

FIG. 2: FOREARMS (6 MONTHS)



FIG. 3: RIGHT LEG (31 MONTHS)



Fig. 2: There is generalised reduction in bone density with relative flaring of the metaphyses. The long bones are slender. There are slightly angulated and displaced transverse compression fractures at the junction of distal metaphysis and diaphysis of both radii and ulnae. Note the irregular metaphyseal sclerosis at the fracture sites and of distal humerus and proximal radii and ulnae, where there are no fractures. There is a further minimally displaced fracture of the shaft of the left radius (arrow)

Fig. 3: There is significant bowing deformity of the leg bones at sites of previous metadiaphyseal fractures. Increased metaphyseal sclerosis persists, although not as obvious as on the earlier radiographs. The bones are dense, presumably due to bisphosphonate therapy (note bisphosphonate-related dense periphery of the tarsal bones)

D. Poster: International Conference on Childrens Bone Health (2017)

P4HB recurrent missense mutation causing Cole-Carpenter syndrome: exploring the underlying mechanism

Meena Balasubramanian^{1,2}, Raja Padidela³, Rebecca C Pollitt^{4,5}, Nicholas J Bishop⁴, M Zulf Mughal³, Amaka C Offiah⁴, Bart E Wagner⁶, Janine McCaughey⁷, David J Stephens⁷

¹Sheffield Clinical Genetics Service, Sheffield Children's NHS Foundation Trust; ²Highly Specialised Service for Severe, Complex and Atypical OI Service, Sheffield Children's NHS Foundation Trust, Sheffield; ³Department of Paediatric Endocrinology, Royal Manchester Children's Hospital, Central Manchester University Hospitals NHS Foundation Trust; ⁴Academic Unit of Child Health, University of Sheffield; ⁵Sheffield Diagnostic Genetics Service, Sheffield Children's NHS Foundation Trust; ⁶Department of Histopathology, Royal Hallamshire Hospital, Sheffield; ⁷School of Biochemistry, University of Bristol, Bristol

Introduction

- In 1986, Cole and Carpenter⁽¹⁾ reported two unrelated infants with multiple fractures and deformities of bone, with a skeletal phenotype similar to severe osteogenesis imperfecta (OI). In addition, these patients also had proptosis, blue sclerae and a distinct facial gestalt.
- Radiologically, these patients had characteristic skeletal manifestations including craniosynostosis, in addition to deformities seen in severe progressive OI.
- Since the first description, there was only one further report of a patient with a similar phenotype⁽²⁾ until Balasubramanian *et al.*, 2015 described a 12-year old patient with phenotypic CCS who had homozygous mutations in *CRTAP* and described CCS as a variant of recessive OI⁽³⁾.
- In 2015, two reports were published describing the aetiology of CCS^(4,5):
- Rauch *et al.*, 2015 described a heterozygous missense mutation c.1178A>G, p.Tyr393Cys in *P4HB* in the original cohort of patients; this came to be known as Cole-Carpenter syndrome 1 (OMIM #11240).
- Garbes *et al.*, 2015 reported a German boy with phenotypic CCS and 2 fetuses with antenatal presentation of a severe bone fragility in a separate German family. All patients in this cohort were found to have compound heterozygous mutations in *SEC24D*. This has now come to be known as Cole-Carpenter syndrome 2 (OMIM #616294).

Clinical Features

- 1st child of healthy, non-consanguineous White, European parents with no significant family history.
- No suggestions antenatally of bone fragility. She was born in a good condition at term with a birth weight of 3710 grams (Z-score 0.6).
- At 6 months of age, she presented with a left distal tibial fracture.
- O/E: failure to thrive (weight Z-score reduced to -1.4), relatively large head with wide open fontanelle, blue sclera, small nose, flat nasal bridge, right upturned ear lobe, a broad face, long fingers, broad thumbs and great toes.
- Started on treatment with bisphosphonates at 7-months of age.
- Skeletal survey: diffusely osteopenic bones with multiple vertebral body compression fractures; fractures of anterior left 8th, 9th & 10th ribs; Wormian bones over sagittal and lambdoid sutures. All long bones showed "crumpling" metaphyseal fractures with widening, sclerosis, and irregularity of the distal metaphyses.
- Currently, 3-years of age, has normal intelligence and mobile in a wheel-chair.

Phenotype

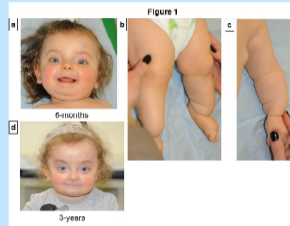


Figure 1: Photographs of patient (frontal and profile) demonstrating frontal bossing, blueish-grey sclerae, flat nasal bridge, limb deformities aged 6-months and 4 years

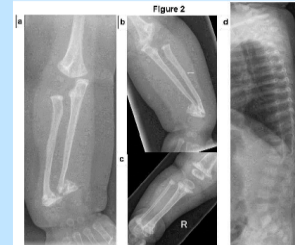


Figure 2: Selected radiographs from a skeletal survey (aged 6-months)



Note the irregular metaphyseal sclerosis at the fracture sites, but also of distal humerus and proximal radius and ulna, where there are no fractures. Symmetrical metaphyseal fractures of this nature are uncommon in OI

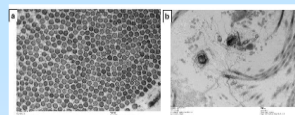
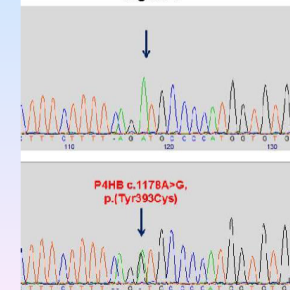


Figure 3: Electron microscopy demonstrating scattered small collagen fibrils within a bundle of unremarkable fibrous collagen. Original magnification x20,000; 3b: Proteoglycan fibres wrapping around collagen fibrils. Original magnification x12,000

Figure 4



Results

- Exome sequencing identified a heterozygous c.1178A>G, p.Tyr393Cys pathogenic mutation in exon 9 of *P4HB* (NM_000918.3) (Figure 4).
- This mutation is predicted to replace tyrosine at position 393 with a cysteine; *in silico* analysis supports its likely pathogenicity.

Functional Studies

- Electron microscopy showed elastic tissue and fibroblasts were unremarkable with no collagen flowers seen. The mean collagen fibril diameter was 74nm, with a range from 39- 95nm.
- Collagen deposited by patient fibroblasts in culture showed no statistically detectable differences (with a p-value of 0.76 for the quantified mid-intensity and 0.34 for high-intensity deposition) compared to controls⁽⁶⁾.

Discussion

- P4HB* (Prolyl 4-hydroxylase, beta subunit), also known as 'Procollagen-proline 2-oxoglutarate-4-dioxygenase, beta subunit (OMIM #176790) encodes for PDI (Protein Disulfide isomerase); in nucleated cells, in its tetrameric form and assists in the formation of 4-hydroxyproline in collagen.
- It is likely that the recurrent missense mutation in *P4HB* results in a very specific form of bone fragility by interfering with collagen formation, which could conceivably be associated with additional changes in deposition and/or assembly of extracellular matrix (ECM).
- The clinical phenotype appears consistent in CCS but there also appears to be an emerging, distinctive radiological phenotype with a definitive phenotypic clue to genotype.
- Metadiaphyseal fractures with metaphyseal sclerosis in long tubular bones are uncommon in OI and may be a pointer to this specific mutation.

Conclusions

- PDI proteins have a role in haemostasis, infectious disease, lipid homeostasis, neurodegeneration, cancer and infertility⁽⁷⁾; we now know that it causes bone disease (OI).
- Metaphyseal 'crumpling' fractures with metaphyseal sclerosis may be a clue to the underlying genotype.
- The recurrent mutation in *P4HB*, c.1178A>G, p.Tyr393Cys causes Cole Carpenter syndrome.

References

- Cole DE and Carpenter TO. Bone fragility, craniosynostosis, ocular proptosis, hydrocephalus, and distinctive facial features: a newly recognized type of osteogenesis imperfecta. *J Pediatr*. 1987; 110:78-88.
- Amad JJ *et al.*. New cases of Cole-Carpenter syndrome. *Am J Med Genet*. 2005;20(4):273-7.
- Balasubramanian M, *et al.* *CRTAP* mutation in a patient with Cole-Carpenter syndrome. *Am J Med Genet A*. 2015; 167(9):1697-91.
- Rauch F *et al.* Cole-Carpenter syndrome is caused by a heterozygous missense mutation in *P4HB*. *Am J Hum Genet*. 2015;96(2):425-31.
- Garbes L *et al.* Mutations in *SEC24D*, encoding a component of the COP11 machinery, cause a syndromic form of osteogenesis imperfecta. *Am J Hum Genet*. 2015;96(3):452-6.
- McCaughy *et al.* TFG Promotes Organization of Transitional ER and Efficient Collagen Secretion. *Cell Rep*. 2016; 16:1946-1956.
- Bertram AM. The protein disulfide isomerase family: key players in health and disease. *Antioxid. Redox Signal*. 2012; 16: 784-798.

E. Manuscript: Phenotypic Variability in Patients with Osteogenesis Imperfecta caused by BMP1 Mutations

ORIGINAL ARTICLE

AMERICAN JOURNAL OF
medical genetics PART
A

Phenotypic Variability in Patients With Osteogenesis Imperfecta Caused by *BMP1* Mutations

Rebecca C. Pollitt,^{1,2*} Vrinda Saraff,³ Ann Dalton,¹ Emma A. Webb,^{3,4} Nick J. Shaw,^{3,4} Glenda J. Sobey,⁵ M. Zulf Mughal,⁶ Emma Hobson,⁷ Farhan Ali,⁸ Nicholas J. Bishop,^{2,6} Paul Arundel,⁹ Wolfgang Högler,^{3,4} and Meena Balasubramanian^{9,10}

¹Sheffield Diagnostic Genetics Service, Sheffield Children's NHS Foundation Trust, Sheffield, UK

²Academic Unit of Child Health, Department of Oncology and Metabolism, University of Sheffield, Sheffield, UK

³Department of Endocrinology and Diabetes, Birmingham Children's Hospital, Birmingham, UK

⁴Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

⁵National EDS Service, Sheffield Children's NHS Foundation Trust, Sheffield, UK

⁶Department of Paediatric Endocrinology, Royal Manchester Children's Hospital, Central Manchester University Hospitals, Manchester, UK

⁷Department of Clinical Genetics, Chapel Allerton Hospital, Leeds, UK

⁸Department of Paediatric Orthopaedic Surgery, Royal Manchester Children's Hospital, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

⁹Highly Specialised Severe, Complex and Atypical OI Service, Sheffield Children's NHS Foundation Trust, Sheffield, UK

¹⁰Sheffield Clinical Genetics Service, Sheffield Children's NHS Foundation Trust, Sheffield, UK

Manuscript Received: 6 January 2016; Manuscript Accepted: 18 August 2016

Osteogenesis Imperfecta (OI) is an inherited bone fragility disorder most commonly associated with autosomal dominant mutations in the type I collagen genes. Autosomal recessive mutations in a number of genes have also been described, including the *BMP1* gene that encodes the mammalian Tollid (mTLD) and its shorter isoform bone morphogenic protein-1 (*BMP1*). To date, less than 20 individuals with OI have been identified with *BMP1* mutations, with skeletal phenotypes ranging from mild to severe and progressively deforming. In the majority of patients, bone fragility was associated with increased bone mineral density (BMD); however, the full range of phenotypes associated with *BMP1* remains unclear. Here, we describe three children with mutations in *BMP1* associated with a highly variable phenotype: a sibship homozygous for the c.2188delC mutation that affects only the shorter *BMP1* isoform and a further patient who is compound heterozygous for a c.1293C>G nonsense mutation and a c.1148G>A missense

How to Cite this Article:

Pollitt RC, Saraff V, Dalton A, Webb EA, Shaw NJ, Sobey GJ, Mughal MZ, Hobson E, Ali F, Bishop NJ, Arundel P, Högler W, Balasubramanian M. 2016. Phenotypic variability in patients with osteogenesis imperfecta caused by *BMP1* mutations.

Am J Med Genet Part A 9999A:1–7.

mutation in the CUB1 domain. These individuals had recurrent fractures from early childhood, are hypermobile and have no evidence of dentinogenesis imperfecta. The homozygous siblings with OI had normal areal BMD by dual energy X-ray

Wolfgang Högler and Meena Balasubramanian are the joint senior authors.

Conflicts of interest: None.

Grant sponsor: The Children's Hospital Charity, Sheffield, Research Fund; Grant number: CA11004.

*Correspondence to:

Rebecca C. Pollitt, Sheffield Diagnostic Genetics Service, Sheffield NHS Foundation Trust, Western Bank, Sheffield S10 2TH. UK.

E-mail: rebecca.pollitt@sch.nhs.uk

Article first published online in Wiley Online Library

(wileyonlinelibrary.com): 00 Month 2016

DOI 10.1002/ajmg.a.37958

absorptiometry whereas the third patient presented with a high bone mass phenotype. Intravenous bisphosphonate therapy was started in all patients, but discontinued in two patients and reduced in another due to concerns about increasing bone stiffness leading to chalk-stick fractures. Given the association of *BMP1*-related OI with very high bone material density, concerns remain whether anti-resorptive therapy is indicated in this ultra-rare form of OI. © 2016 Wiley Periodicals, Inc.

Key words: Osteogenesis imperfecta; bone fragility; *BMP1*; bone morphogenic protein-1; high bone mass

INTRODUCTION

Osteogenesis Imperfecta (OI) is a rare inherited connective tissue disorder characterised by an increased tendency to fracture, often with minimal or no apparent trauma. Extra-skeletal features can include short stature, skin and joint hyper-extensibility, blue sclerae, deafness, and dentinogenesis imperfecta. Other features are bone pain, deformities, scoliosis, and impaired mobility.

Genetic characterisation of families affected with OI has shown that autosomal dominant mutations in the genes that encode the alpha chains of type I collagen (*COL1A1* and *COL1A2*) can be identified in approximately 85–90% of affected individuals [Forlino and Marini, 2016]. Mutations in a variety of other genes encoding proteins involved in type I collagen biosynthesis, bone cell differentiation, bone formation, and bone remodeling are known to result in rare forms of autosomal recessive OI [Mendoza-Londono et al., 2015]. A hallmark of OI at the tissue level is increased bone mineralisation density [Rauch and Glorieux, 2004].

Mutations in the *BMP1* gene have been described in a small number of individuals with OI. The *BMP1* gene (OMIM 112264) is alternatively transcribed to produce two proteins, Mammalian Tolloid (mTLD) and its shorter isoform bone morphogenic protein-1 (BMP1). The BMP1/mTLD protein acts as an astacin metalloprotease, whose functions include the proteolytic removal of the carboxyl-terminal propeptide from procollagen type I, II, and III and the amino-terminal propeptide from types V and XI procollagen. Studies in BMP1/mTLD deficient patients with OI have demonstrated delayed cleavage of type I collagen C-propeptide [Valencia et al., 2014] and disorganization of type I/V collagen fibrils as well as impaired processing of the small leucine rich proteoglycan (SLRP) decorin [Syx et al., 2015].

The OI phenotype of individuals with *BMP1* mutations has been described as recurrent fractures, generalized bone deformity, osteopenia, and Wormian bones [Martinez-Glez et al., 2012], and also as bone fragility associated with an increase in areal bone mineral density (BMD) as measured by dual-energy X-ray absorptiometry (DXA) [Asharani et al., 2012], similar to mutations that affect the C-propeptide cleavage site of type I collagen [Lindahl et al., 2011]. At the tissue level, bone of one OI patient with *BMP1* mutations was found to be even more hypermineralized than in OI caused by collagen mutations [Hoyer-Kuhn et al., 2013]. Patients with mutations affecting the C-propeptide cleavage site have also been reported to have bone mineral content that exceeds that of classical OI [Lindahl et al., 2011].

However, less than 20 individuals have been described with *BMP1*-related OI and therefore the full range of phenotypes associated with mutations in this gene is not well established. Here, we describe three patients that further expand the phenotypic spectrum of *BMP1*-related OI; a sibship presenting with OI and normal BMD and a further patient presenting with a high bone mass phenotype.

MATERIALS AND METHODS

Clinical Information and DNA Extraction

Clinical information was obtained from the patients' medical records. The patients and their parents provided informed consent. Total genomic DNA was isolated from 2 to 5 ml peripheral blood taken from patients and parents using standard extraction methods.

DNA Sequencing and Mutation Analysis

Targeted exome sequencing using SureSelect XT (Agilent Technologies) and Illumina MiSeq platform was used to sequence all coding regions and intron/exon boundaries of genes previously described in OI. The variants identified in the *BMP1*/mTLD gene were compared to reference sequences NM_001199.3 and NM_006129.4 and their pathogenicity assessed using Alamut Visual version 2.6 (Interactive Biosoftware, Rouen, France).

RESULTS

Clinical Characteristics

Patient 1, the first child of consanguineous parents of Asian origin was born at 37 weeks gestation weighing 2.5 kg. At 6 years of age, he was referred to the metabolic bone clinic for investigation having sustained 8–10 fractures in total including both tibiae, forearms, and a finger and toe. His first fracture, of the right wrist, occurred at age 12 months. He had no other significant medical conditions, normal hearing and cognitive development. Due to non-union of a transverse right tibia fracture, he had sustained at age 4 years, he mobilised using a wheelchair. On clinical examination, there were no dysmorphic features. He had white sclerae, hypermobile fingers, and no evidence of dentinogenesis imperfecta but severe dental decay requiring multiple tooth extractions. He had normal serum calcium and phosphate, alkaline phosphatase, parathyroid hormone, and vitamin D levels. His lumbar spine BMD Z score, measured by DXA was –0.9. A diagnosis of OI was thought to be likely and intravenous pamidronate treatment (1.5 mg/kg/day over 2 days 3 monthly) was commenced on the basis of persistent vertebral compression fracture at L1 (Fig. 1) on a background of multiple long bone fractures. The non-union of the long-standing right tibial fracture required intramedullary nailing which allowed better mobility but never united subsequently. Further intramedullary nailing of left tibia and femur was required because of mid-shaft fractures. Pamidronate treatment was paused after 1 year to promote healing of osteotomies. From age 9 years, he received zoledronic acid (0.05 mg/kg/day single dose 6 monthly). His bone density increased in response to bisphosphonate therapy (Table I).



FIG. 1. Patient 1 X-rays. (A) Compression fracture at L1 at age 8.5 years. (B) Left low impact sub-trochanteric femur fracture at age 11 years (after 4.5 years on bisphosphonate therapy).

However, he continued to have long bone fractures, including a new vertebral fracture (T10), an oblique subtrochanteric left femur fracture age 11 years (Fig. 1) and a chalk-stick like mid-shaft femur fracture age 13 years, following 7.5 years on bisphosphonate therapy (Fig. 2), which was subsequently discontinued.

He is currently growing along the 10th centile for height and less than 4th centile for weight. His current head circumference is 52.2 cm (age 14 years, between -1 to -2 SD).

His sister, Patient 2, was born at 36 weeks gestation weighing 2.42 kg. She presented at 7 months of age with bilateral ulnar and radial fractures following a fall down the stairs while in the arms of her older sister. Over the following 3 years, she had three low impact tibia fractures necessitating right tibial rodding. Zoledronic acid infusions (0.035 mg/kg/day single dose 4 monthly) were started at 3 years of age. On clinical examination, she had gray sclerae, hypermobile fingers, and no evidence of dentinogenesis

TABLE I. Response in Lumbar Spine (LS) and Total Body Less Head (TBLH) Bone Density Z-Scores (by DXA) During Intravenous Bisphosphonate Therapy in Patient 1

Age in years	6 Pre-treatment	7 Treatment paused	8 Treatment recommenced	9	10	12	14
LS BMD	-0.9	-0.8	-1.1	-0.8	-0.4	0.3	0.3
LS BMAD [L1-L4]	-0.6	-0.6	-0.6	0.4	1.3	2.1	1.8
TBLH BMD	NA	-0.5	-1.7	-0.4	-0.8	0.3	-0.4

LS BMD, lumbar spine bone mineral density; LS BMAD, Lumbar spine bone mineral apparent density; TBLH, total body less head. Treatment was paused at age 7 due to rodding surgery.

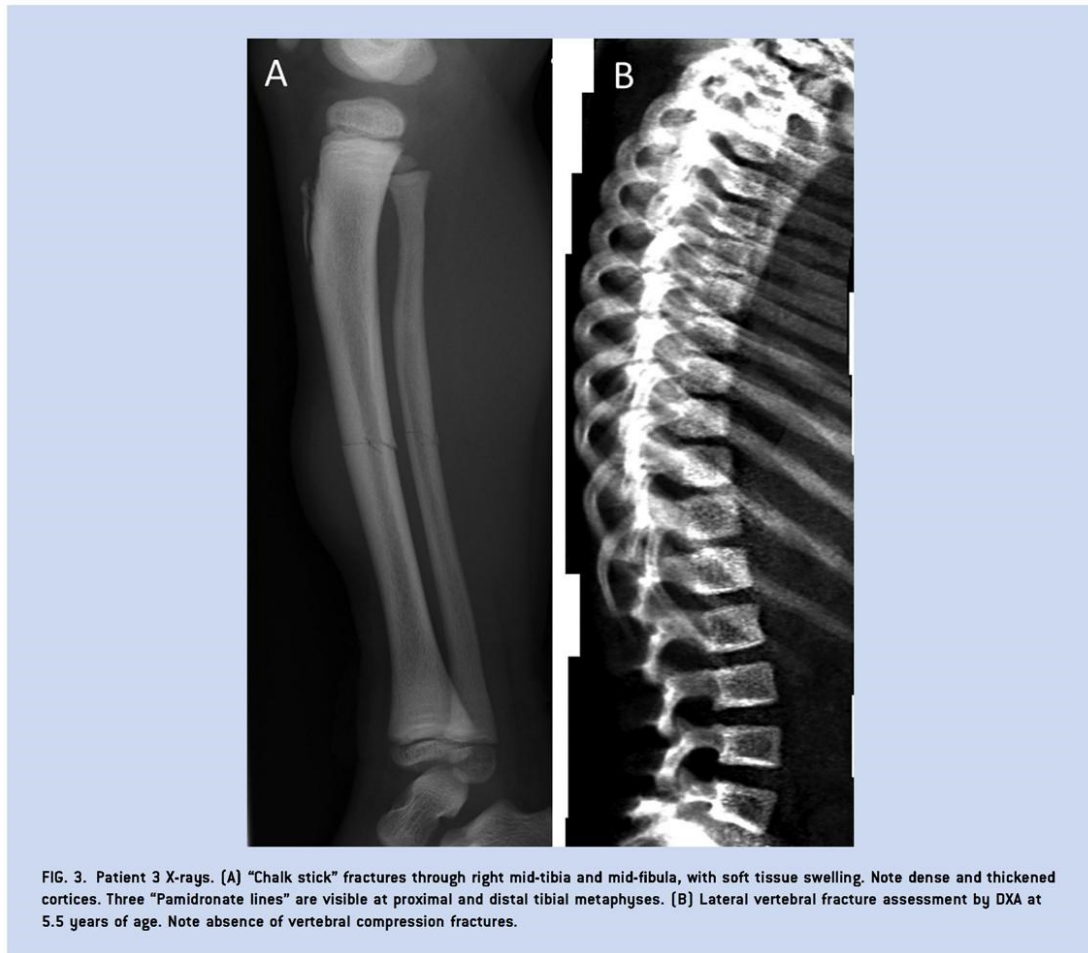


FIG. 2. Patient 1 X-rays, age 13 years: Left "chalk-stick" mid-shaft femur fracture following a fall at school (after 7.5 years on bisphosphonate therapy). Note the mild Erlenmeyer-shape deformity of distal femur with "bisphosphonate" lines.

imperfecta. She had no other medical conditions, normal hearing and cognitive and gross motor development. Her height and weight are on the 3rd centile for age. Her bone profile and vitamin D level at the time of diagnosis were normal. DXA scanning prior to treatment was not performed due to lack of paediatric bone mineral apparent density (BMAD) reference data and cooperation of children under 5 years [Crabtree et al., 2014]. At the age of 5.2 years, her lumbar spine BMAD Z score was 0.9 and total body less head (TBLH) BMD Z score was 1.5, following 2 years of zoledronate treatment during which she sustained one further tibia fracture. Lateral vertebral assessment showed no evidence of vertebral compression fractures. Her head circumference currently is 47.5 cm (age 5.2 years, -2 SD). In light of her increasing BMAD and experience with her brother, zoledronate therapy was reduced in frequency to a once yearly infusion (0.05 mg/kg).

Patient 3, the only child of healthy non-consanguineous parents of North European origin. She was born following IVF treatment at 39 weeks gestation with a birth weight of 2.976 kg and her early developmental assessments were normal with the exception of gross motor development which was delayed; she sat up at 1 year of age and walked around 17 months of age. She was diagnosed with bilateral dislocated hips at

22 months of age for which she had surgery twice, and was immobilised in the hip spica. The patient's first fracture, of the right fibula, occurred at the age of 2 years and 11 months. She was reviewed at the age of 3 years and 2 months, having sustained separate fractures of her right fibula and left tibia. Her lumbar spine (L1-4) BMD was 0.726 g/cm^2 (BMAD Z score $+4.2$, calculated retrospectively) and did not influence management because lateral thoracic and lumbar spine radiographs done at the same time did not show any vertebral deformity, osteopenia, or clear radiological evidence of increased bone density. Her lower-limb fragility fractures were attributed to prolonged immobilisation in hip spica. She was subsequently seen in another centre having had further fractures, including a spiral fracture of the tibia and three metatarsal fractures in the left foot and was empirically treated with pamidronate at age $3\frac{3}{4}$ years (1 mg/kg on three consecutive days, 3 monthly). She received 4 cycles (total dose 12 mg/kg) before treatment was discontinued as her long bones now appeared abnormally dense on radiographs, and she started to suffer apparent "chalk stick fractures" of her tibiae and fibulae (Fig. 3). At the age of $5\frac{1}{2}$ years, bone mineral density measurements were undertaken using various imaging techniques (Table II). The girl's distal radial total and trabecular



volumetric BMD Z score, measured by peripheral computed tomography, were markedly elevated being +8.7 and +9.2, respectively. At the lumbar spine (LS), her BMAD Z score (+4.3), measured by DXA was elevated but surprisingly not the volumetric trabecular BMD Z score (+0.3) measured by QCT. This apparent discrepancy suggests that the trabecular compartment in the LS is less affected than that at the distal radius; however, different reference data used to calculate Z scores by these two techniques in the presence of high cortical bone mass may be a contributing factor. Radiograph of the spine and lateral vertebral fracture assessment by DXA (Fig. 3) did not reveal vertebral fractures. A provisional diagnosis of a mild form of osteopetrosis was suggested; however, genetic testing for a panel of 21 genes, including *CLCN7* and *LRP5*, was negative. She has continued to have long bone fractures; the most recent are

“chalk-stick” like mid-shaft fractures of the left tibia and fibula at age 7 $\frac{3}{4}$ years, 3 years after discontinuation of bisphosphonate therapy.

On clinical examination at aged 7 years, she had white sclerae and normal teeth, hearing and spine. She has a bossed forehead and mild left sided ptosis. She has generalised hypermobility with a Beighton score of 8/9 with soft, velvety, and very stretchy skin.

Identification of *BMP1* Mutations

Mutations in the *COL1A1* and *COL1A2* genes were excluded in all three patients. Targeted exome sequencing for a panel of additional genes associated with OI revealed that patient 1 and 2 were homozygous for the previously described c.2188dupC mutation [Syx et al., 2015]. The parents were confirmed to be heterozygous carriers.

TABLE II. Volumetric Bone Mineral Density Z-Scores of Patient 3 Measured by Peripheral Quantitative Computed Tomography (Distal Radius), DXA (Lumbar Spine), and Quantitative Computer Tomography (Lumbar Spine)

Age in years	3.2	5.5	6.5	7
Distal radial total volumetric BMD	Not measured	+8.7	+7.1	+7.2
Distal radial trabecular volumetric BMD	Not measured	+9.2	+6.9	+6.6
LS BMAD [L1–L4]	+4.2	+4.3	+3.1	+3.4
LS volumetric trabecular BMD [L1–L3]	Not measured	+0.35	Not measured	Not measured

BMD, Bone mineral density; LS, lumbar spine; BMAD, bone mineral apparent density.

Pamidronate was started at 3.8 years of age and she remained on treatment for 12 months before identification of raised BMD led to treatment discontinuation.

Patient 3 was compound heterozygous for two novel mutations, a c.1293C>G;p. (Tyr431*) non-sense mutation and a c.1148G>A;p. (Arg383Gln) missense mutation in the CUB1 domain of *BMP1*. Parental testing demonstrated that the c.1293C>G;p. (Tyr431*) was present in the mother and c.1148G>A;p. (Arg383Gln) in the father.

DISCUSSION

To date the majority, but not all, of individuals described with *BMP1*-associated OI have presented with bone fragility associated with increased BMD although no clear genotype–phenotype correlation has yet emerged.

The c.2188dupC identified in patients 1 and 2 is predicted to have different outcomes dependent on the gene transcript. In the shorter *BMP1* transcript, this mutation would lead to the creation of an extended protein (p.Gln730Profs*294), whereas in the longer mTLD transcript this mutation is predicted to result in an intronic duplication (c.2108-605dupC). Two individuals who are compound heterozygous for this change and the recurring signal peptide mutation, p.(Gly12Arg), have previously been described [Syx et al., 2015]. These individuals are reported to have a severe progressive form of OI. Patient 1 and patient 2 presented with a phenotype suggesting that the mutant protein may have residual C-propeptide cleavage activity and the c.2188dupC may therefore represent a relatively “mild” mutation.

The markedly increased bone mass and “chalk-stick” pattern of long-bone fractures of patient 3 initially suggested a diagnosis of osteopetrosis. To date, similar compound heterozygous changes that result in a “null” allele and a mutation in a CUB domain have been associated with severe OI phenotypes. Interestingly, a patient with a c.925delC frameshift mutation and a p.(Gly498Arg) substitution in the CUB2 domain is reported to have severe rhizomelic deformities, short stature, and to have sustained over 100 fractures [Syx et al., 2015]. Unfortunately, no data is available for the associated BMD in this patient but, in contrast to patient 3, extensive skeletal surveys showed generalized undermineralization of long bones.

It remains unclear why some *BMP1* mutations are associated with increased BMD and others with normal or reduced BMD. Areal BMD measurements provide a composite value for bone mass within a given area, and do not reflect tissue mineralisation density—the combination of increased bone material density with reduced bone mass (as is typical in OI) can give values for BMD that

sit within the normal range for age in children. However, this is clearly not the case for patient 3, where size corrected LS BMD (BMAD) and distal radius volumetric BMD are elevated (Table II). Mineral crystals in OI patients are known to be smaller, have high calcium content and are more densely packed than in normal bone. Tissue mineralisation density may be a reflection of the degree of matrix disorganisation; some of the highest values are in type VI OI, where patients have a severely disrupted lamellar structure [Land et al., 2007]. The multiple potential effects on matrix organisation resulting from mutations in *BMP1* could be similarly disruptive.

Bone tissue analysis of trabecular and cortical bone from an individual homozygous for the *BMP1* p.(Gly12Arg) signal peptide mutation demonstrated increased regions of unmineralised matrix at sites of new bone formation, possibly caused by a delay in matrix maturation necessary for mineralisation. In contrast, hypermineralisation was observed at older bone sites in the same bone sample, hypothesised to result from an increase in matrix space caused by retention of the C-propeptide in collagen fibrils which is subsequently filled by mineral crystals [Hoyer-Kuhn et al., 2013].

Functional studies in patients with *BMP1* mutations have largely focused on C-propeptide cleavage activity. However, *BMP1*/mTLD is also involved in processing of additional extra cellular matrix components, in particular the processing of the SLRP prodecorin by removal of the prodomain, which has been shown to be delayed in patients with *BMP1* mutations [Syx et al., 2015]. Decorin is known to influence both collagen assembly and regulate matrix mineralization [Mochida et al., 2009]. The CUB domains of *BMP1*/mTLD are essential for C-proteinase activity; thus, mutations in different CUB domains may also be contributing to the variation in mineralization seen in these patients through their interaction with SLRPs. Potentially, this may explain why our patients with the c.2188dup, where all the CUB domains are intact, did not present with a high bone mass phenotype.

Intravenous bisphosphonate therapy was started in all patients, but discontinued in two patients and reduced in the younger sibling due to concerns about increasing bone stiffness contributing to occurrence of chalk-stick fractures. While bisphosphonates are not known to increase bone material density or stiffness in OI caused by collagen gene mutations [Weber et al., 2006], they impair bone remodeling/repair and healing, possibly allow accumulation of microdamage [Chapurlat and Delmas, 2009] and are linked to atypical fractures in adults [Shane et al., 2014]. Given the very high bone material density associated with *BMP1*-related OI and the

F. Manuscript: P4HB recurrent missense mutation causing Cole-Carpenter syndrome

Downloaded from <http://jmg.bmj.com/> on January 17, 2018 - Published by group.bmj.com
JMG Online First, published on December 20, 2017 as 10.1136/jmedgenet-2017-104899
Genotype-phenotype correlations

ORIGINAL ARTICLE

P4HB recurrent missense mutation causing Cole-Carpenter syndrome

Meena Balasubramanian,^{1,2} Raja Padidela,³ Rebecca C Pollitt,^{4,5} Nicholas J Bishop,⁴ M Zulf Mughal,³ Amaka C Offiah,⁴ Bart E Wagner,⁶ Janine McCaughey,⁷ David J Stephens⁷

¹Sheffield Clinical Genetics Service, Sheffield Children's NHS Foundation Trust, Sheffield, UK

²Highly Specialised Service for Severe, Complex and Atypical OI Service, Sheffield Children's NHS Foundation Trust, Sheffield, UK

³Department of Paediatric Endocrinology, Royal

Manchester Children's Hospital, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

⁴Academic Unit of Child Health, University of Sheffield, Sheffield, UK

⁵Sheffield Diagnostic Genetics Service, Sheffield Children's NHS Foundation Trust, Sheffield, UK

⁶Department of Histopathology, Royal Hallamshire Hospital, Sheffield, UK

⁷Cell Biology Laboratories, School of Biochemistry, University of Bristol, Bristol, UK

Correspondence to

Dr Meena Balasubramanian, Sheffield Clinical Genetics Service, Sheffield Children's NHS Foundation Trust, Sheffield S10 2TH, UK; meena.balasubramanian@sch.nhs.uk

Received 29 June 2017

Revised 7 November 2017

Accepted 15 November 2017



CrossMark

To cite: Balasubramanian M, Padidela R, Pollitt RC, et al. *J Med Genet* Published Online First: [please include Day Month Year]. doi:10.1136/jmedgenet-2017-104899

ABSTRACT

Background Cole-Carpenter syndrome (CCS) is commonly classified as a rare Osteogenesis Imperfecta (OI) disorder. This was following the description of two unrelated patients with very similar phenotypes who were subsequently shown to have a heterozygous missense mutation in *P4HB*.

Objectives Here, we report a 3-year old female patient with severe OI who on exome sequencing was found to carry the same missense mutation in *P4HB* as reported in the original cohort. We discuss the genetic heterogeneity of CCS and underlying mechanism of *P4HB* in collagen production.

Methods We undertook detailed clinical, radiological and molecular phenotyping in addition, to analysis of collagen in cultured fibroblasts and electron microscopic examination in the patient reported here.

Results The clinical phenotype appears consistent in patients reported so far but interestingly, there also appears to be a definitive phenotypic clue (crumpling metaphyseal fractures of the long tubular bones with metaphyseal sclerosis which are findings that are uncommon in OI) to the underlying genotype (*P4HB* variant).

Discussion *P4HB* (Prolyl 4-hydroxylase, betasubunit) encodes for PDI (Protein Disulfide isomerase) and in cells, in its tetrameric form, catalyses formation of 4-hydroxyproline in collagen. The recurrent variant in *P4HB*, c.1178A>G, p.Tyr393Cys, sits in the C-terminal reactive centre and is said to interfere with disulphide isomerase function of the C-terminal reactive centre. *P4HB* catalyses the hydroxylation of proline residues within the X-Pro-Gly repeats in the procollagen helical domain. Given the inter-dependence of extracellular matrix (ECM) components in assembly of a functional matrix, our data suggest that it is the organisation and assembly of the functional ECM that is perturbed rather than the secretion of collagen type I per se.

Conclusions We provide additional evidence of *P4HB* as a cause of a specific form of OI-CCS and expand on response to treatment with bisphosphonates in this rare disorder.

INTRODUCTION

In 1987, Cole and Carpenter reported two unrelated infants with multiple fractures and deformities of bone, with a skeletal phenotype similar to severe osteogenesis imperfecta (OI). In addition, these patients also had proptosis, blue sclerae, hydrocephalus and a distinct facial gestalt. They were reported

to be of normal intelligence.¹ Radiologically, these patients had characteristic skeletal manifestations including craniosynostosis, in addition to the deformities seen in severe progressive OI.

Since the first description, there was only one further report of a patient with a similar phenotype² until Balasubramanian *et al*³ described a 12-year-old patient with phenotypic Cole-Carpenter syndrome (CCS) who had homozygous mutations in *CRTAP* and described CCS as a variant of recessive OI. This patient had the characteristic facial features, ocular proptosis, craniosynostosis (corrected in childhood) and severe bone fragility (needing treatment with bisphosphonates). In 2015, two reports were published describing the aetiology of CCS^{4,5}: Rauch *et al*, described a heterozygous missense mutation in prolyl 4-hydroxylase, beta subunit (*P4HB*) in the original cohort of patients; this came to be known as CCS1 (CLCRP1 OMIM #11240); and Garbes *et al* reported a German boy with phenotypic CCS and two fetuses from terminated pregnancies with antenatal presentation of a severe bone fragility in a separate German family. All patients in this cohort were found to have compound heterozygous mutations in *SEC24D*. The authors described their presentation to be overlapping with cranio-lenticulo-sutural dysplasia (CLSD OMIM #607812) caused by mutations in a related COPII gene, *SEC23A*, but summarised the phenotype to be more like CCS. This has now come to be known as CCS2 (CLCRP2 OMIM #616294). Table 1 provides an overview of the clinical, radiological and molecular diagnosis of patients with CCS reported in the literature and this patient.

Therefore, it appears that CCS is a genetically heterogeneous condition with a strikingly similar phenotype, common features being severe bone fragility and distinct facial dysmorphism with/without craniosynostosis. The facial dysmorphism mainly comprises relative macrocephaly with wide-open fontanelle, blue sclera, small nose, flat nasal bridge and a broad face with ocular proptosis. This is perhaps slightly different to the facial features characteristically seen in OI due to other genetic aetiology.

Here we describe a 3-year-old patient with severe bone fragility who, on exome sequencing, was shown to have the same missense mutation as described by Rauch *et al*. Reverse phenotyping

BMJ

Balasubramanian M, et al. *J Med Genet* 2017;0:1–8. doi:10.1136/jmedgenet-2017-104899

Copyright Article author (or their employer) 2017. Produced by BMJ Publishing Group Ltd under licence.

Genotype-phenotype correlations

Table 1 Comparison of clinical, radiological and molecular diagnosis in patients reported with CCS with patient reported here

	This patient	Cole and Carpenter ¹	Patient 1	Cole and Carpenter ¹	Patient 2	Amor ²	Balasubramanian ³	Garbes ⁴	Patient 1	Garbes ⁵	Garbes ⁵
	3 years	4 years	18 months	30 months	12 years	7 years	Fetus 1	Fetus 2			
Age	3 years	4 years	18 months	30 months	12 years	7 years	Fetus 1	Fetus 2			
Sex	F	M	M	F	F	M	F	F			
Ethnicity	White	White	White	White	Asian	German	German	German			
Consanguinity	No	No	No	No	Yes	No	No	No			
Molecular diagnosis	<i>P44H</i> c.1178A>G, p.Tyr393Cys	<i>P44H</i> c.1178A>G, p.Tyr393Cys	<i>P44H</i> c.1178A>G, p.Tyr393Cys	NK	<i>CRTAP</i> c.118G>T, p.Glu40*	<i>SEC24D</i> c.613C>T, p.Gln205*, c.3044C>T, p.Ser1015Phe	<i>SEC24D</i> c.2933A>C, p.Gln978Pro (mat); c.3044C>T, p.Ser1015Phe (pat)	<i>SEC24D</i> c.2933A>C, p.Gln978Pro (mat); c.3044C>T, p.Ser1015Phe (pat)			
Prenatal fractures	-	-	-	-	-	+	+	++			
Postnatal fractures	++	++	++	+	++	+	+	NA			
Craniosynostosis	-	+	-	+	+	+	+	NA			
Vertebral fractures/scoliosis	++	++	++	+	++	+	+	NA			
MCF	+	+	+	-	-	-	-	NA			
Wormian bones	+	+	-	+	+	+	+	NA			
BP therapy	+	+?	+	+	+	+	+	NA			
Ocular proptosis	+	+	+	+	+	+	+	NA			
Blue sclerae	+	+	+	+	+	+	+	NA			
Macrocephaly	+	+	+	+	+	+	+	NA			
Hydrocephalus	-	+	+	-	-	+	+	NA			
DI	-	+	-	+	+	-	-	NA			
Intelligence	N	N	N	N	N	N	N	NA			

+?, increase in BMD, no clinical response; BMD, bone mineral density; BP, bisphosphonates; CCS, Cole-Carpenter syndrome; DI, dentinogenesis imperfecta; MCF, metaphyseal compression fractures; NA, not available/applicable; NK, not known yet.

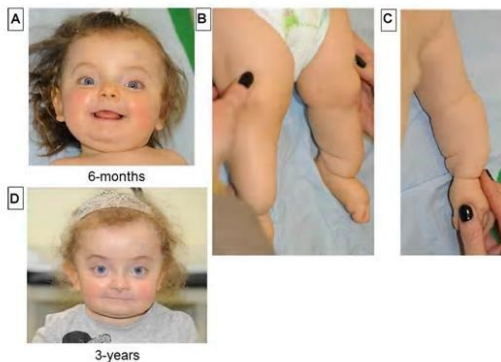


Figure 1 Photographs of patient (frontal and profile) demonstrating frontal bossing, blueish grey sclerae, flat nasal bridge and prominent veins; limb deformities aged 6 months (A–C) and aged 3 years (D).

suggested that this patient was likely to be in the CCS spectrum, and functional studies were undertaken in human primary fibroblasts.

CLINICAL REPORT

The proband is the first child of healthy, non-consanguineous white European parents with no significant family history. She has an older maternal half-brother who is fit and well. The pregnancy was unremarkable with no suggestions antenatally of bone fragility. She was born in a good condition at term with a birth weight of 3710 g (z score 0.6). From 1 week of age, her oral intake reduced, and she was noted to be irritable. She had multiple admissions for suspected urinary tract infection; however, her ultrasound scans of the renal tract, micturating cystourethrogram and dimercaptosuccinic acid (DMSA) scan were all normal. She was diagnosed with positional talipes for which she received physiotherapy. At 6 months of age, she presented with a left distal tibial fracture. On clinical examination, she was noted to have failure to thrive (weight Z score reduced to -1.4), relatively large head with wide-open fontanelle, blue sclera, small nose, flat nasal bridge, right upturned ear lobe, a broad face, long fingers, relatively broad thumbs and great toes (figure 1A–C). A skeletal survey showed diffusely osteopaenic bones with multiple vertebral body compression fractures in the thoracic and lumbar spine and fractures of the anterior left 8th, 9th and 10th ribs. She had Wormian bones over sagittal and lambdoid sutures noted on her skull radiograph and old fractures of both distal radial shafts. All long bones showed ‘crumpling’ metaphyseal fractures with widening, sclerosis and irregularity of the distal metaphyses (figure 2A–E).

A clinical diagnosis of OI was made, and she was commenced on intravenous bisphosphonate (pamidronate) infusions from 7 months of age, which she continues to receive at the current age of 3 years and 8 months. Until 2 years and 8 months of age, she had multiple non-traumatic fractures of long bones of her upper and lower limbs. Her vertebral fractures have, however, significantly improved with remodelling of all vertebral bodies. She had corrective tibial osteotomies and bilateral tibial rodding's with Fassier-Duval lengthening rods at 2 years and 8 months of age, and she is currently awaiting the surgical correction of deformities of her both femora and upper limb bones. Developmentally,

she has progressed well with her speech and language and social skills. Her gross motor development has been delayed because of her limb bone deformities.

At a recent genetics review at 3 years of age, she was noted to be in a mainstream nursery and had not sustained any fractures in the last 18 months. She was wearing glasses for a squint, and her teeth were noted to be brittle but with no evidence of dentinogenesis imperfecta. Her growth parameters were: weight: 12.5 kg (Z score -1.84), height: 77 cm (Z score -4.99) and head circumference: 54 cm (Z score $+2.5$); she was noted to have frontal bossing with a broad face, flat nasal bridge, prominent eyes with bilateral low-set ears (figure 1D) and bilateral limb deformities.

Genetic testing prior to whole exome sequencing included targeted OI panel testing, which at that time included all genes so far published in dominant and recessive OI (October 2014); dosage analysis of *COL1A1/A2* genes was reported negative using Multiplex Ligation-dependent Probe Amplification (MLPA). She was subsequently recruited to a research study being undertaken in children with atypical OI with appropriate parental consent.

MATERIALS AND METHODS

This patient was recruited into a research project to study atypical forms of OI to establish genotype-phenotype correlations. Funding was obtained from the Children's Hospital Charity.

DNA sequencing and exome sequencing

Whole exome sequencing was performed using SureSelect QXT library preparation method and SureSelect Human All Exon V6 baits for target enrichment. Libraries were sequenced on a HiSeq 2500 using paired-end sequencing 2×100 bp. Bcl2fastq (V1.8.4) was used to convert sequencing data to fastq files. Fastq files were uploaded to Sapianta (V1.5.0) for mapping and variant calling. Sapianta does not provide a whole exome coverage summary; however, after mapping with bwa mem (0.7.15) to GRCh37, a mean exonic coverage of $103 \times$ (SD 62.1) and a median of $94 \times$ was obtained. Bases contribute to coverage if their quality is ≥ 30 and mapping quality is at least 10 for the read within which they are contained. Duplicate reads are excluded. ‘SureSelect Human All Exon V6 r2’ design files were used to restrict coverage calculations; here we state values for the ‘Covered’ bed file.

Variants reported with a frequency of $>1\%$ in the National Heart, Lung, and Blood Institute (NHLBI) GO Exome Sequencing Project (ESP) or Exome Aggregation Consortium (ExAC) cohort populations were excluded. Further filtering of variants to look for potential loss of function changes (frameshift, splice site acceptor, splice site donor or stop gain) was undertaken. Finally, a larger group of variants within a targeted gene list was assessed. The gene list consisted of those reported in connection with bone dysplasias, identified in Genome-Wide Association Studies (GWAS) studies of change in bone density and implicated in bone metabolism in mouse models, and genes with human phenotype ontology terms relating to increased susceptibility to fracture, totalling approximately 600 genes.

Electron Microscopy (EM) from skin biopsy

EM of the skin ellipse biopsy specimen was performed by: fixing in 3% phosphate buffered glutaraldehyde, postfixation in 2% aqueous osmium tetroxide, processing to epoxy resin and at embedding, orientation to display sections of epidermis and full thickness dermis. Toluidine blue stained semi-thin sections ($0.6 \mu\text{m}$) were assessed by light microscopy, and thin sections (85 nm) were cut on an ultramicrotome using a diamond knife,

Genotype-phenotype correlations

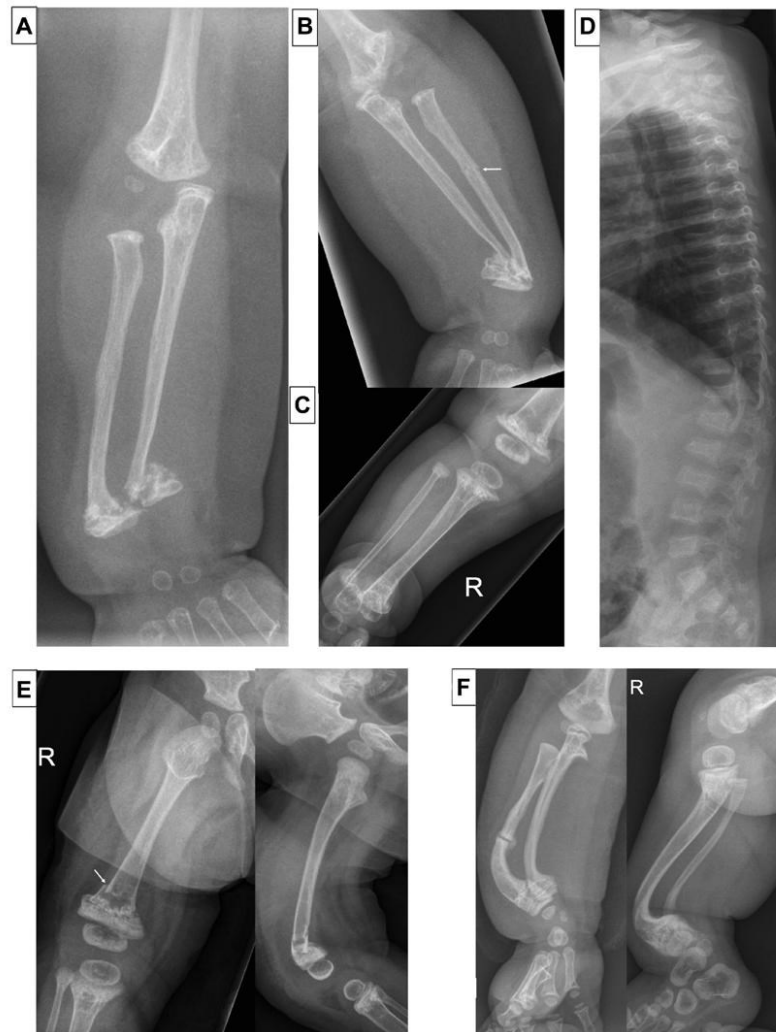


Figure 2 Selected radiographs from a skeletal survey (aged 6 months). (A) AP right forearm. (B) AP left forearm. (C) AP right leg. There is generalised reduction in bone density with relative flaring of the metaphyses. There are slightly angulated and displaced transverse fractures at the junction of metaphyses and diaphyses of all imaged bones except the humeri. Note the irregular metaphyseal sclerosis at the fracture sites and also of distal humerus and proximal radius and ulna, where there are no fractures. Symmetrical metadiaphyseal fractures of this nature are uncommon in OI. There is a further minimally displaced fracture of the shaft of the left radius (arrow). (D) Lateral spine: there are multiple vertebral compression fractures. (E) AP and lateral right femur (aged 10 months). Osteopaenia, flared metaphyses and an acute metadiaphyseal junction fracture in a slightly more proximal position than the previous fracture. As in A, note the irregular sclerosis around the healing fracture. Sclerosis (but no fractures) is also seen around the proximal tibial and fibular metaphyses. (F) Lateral right forearm, lateral right leg (aged 31 months). There is significant bowing deformity of the forearm and leg bones. The bones are dense, presumably due to bisphosphonate therapy (note bisphosphonate lines at distal radial and ulnar metaphyses and dense periphery of the tarsal bones). There are healing fractures of the right midradial and distal humeral shafts. OI, osteogenesis imperfecta.

mounted on a 200 thin bar hexagonal mesh copper grid. The material was stained with uranyl acetate and lead citrate, and seven grid meshes were examined using a Philips 400 transmission EM. Image capture and collagen fibril diameter quantitations were performed using an AMT 16-megapixel midmount

digital camera. The diameters of 301 unselected collagen fibrils, sectioned transversely, from five collagen bundles, at original magnification $\times 20,000$, in the midreticular dermis, were measured. The diameters of 318 collagen bundles were measured at original magnification $\times 550$. The mean of the mean

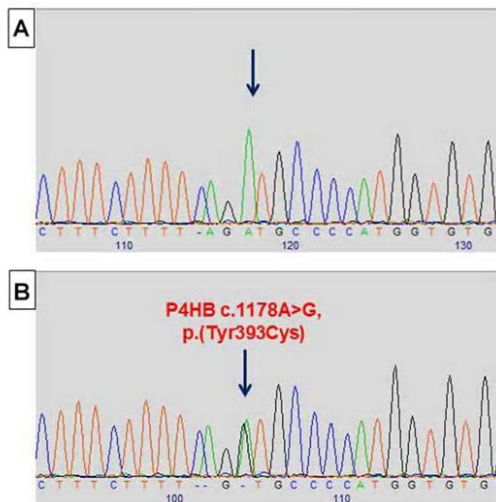


Figure 3 Sequencing electropherograms demonstrating a heterozygous c.1178A>G mutation predicted to result in a p.(Tyr393Cys) amino acid change was identified in the *P4HB* gene (NM_000918.3) (B) and normal trace in mother (A).

diameters, range and deviation root mean squared (RMS) were calculated. In addition, collagen fibrils were checked for the presence of collagen flowers, and fibroblasts were examined for expanded rough Endoplasmic Reticulum (rER) and elastic fibres for structural abnormalities.

Collagen analysis on cultured fibroblasts

Primary patient-derived fibroblasts were cultured in Ham's F10 supplemented with 12% fetal calf serum. Cells were grown for 72 hours on glass coverslips and incubated in the presence of 167 μ M (50 μ g per mL) ascorbic acid for 0.5 or 24 hours, respectively, prior to fixation. Cells were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) and processed for immunofluorescence using anticollagen I α I (Novus Biologicals NB600-408) and imaged by confocal microscopy as described previously using an Alexa-Fluor-568-conjugated secondary antibody (Thermo Fisher Scientific), Prolong Diamond with DAPI for mounting and a Leica SP5 for analysis and image acquisition.⁶ Images were compared using automated segmentation of collagen fluorescence intensity using Volocity (V.6.3, Perkin

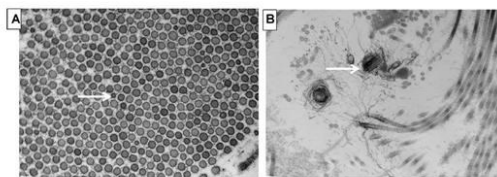


Figure 4 (A) Electron microscopy demonstrating scattered small collagen fibrils within a bundle of unremarkable fibrous collagen. Original magnification $\times 20\,000$; (B) proteoglycan fibres wrapping around collagen fibrils. Original magnification $\times 12\,000$ (indicated by arrows).

Elmer). Statistical analysis was performed using an unpaired Student's t-test (P values).

For semi-quantitative analysis of COL1a1 levels in control and patient cells, cells were seeded confluent and grown for 24 hours. Subsequently, cells were incubated in serum free Ham's F10 medium supplemented with or without 167 μ M (50 μ g per mL) ascorbic acid for 24 hours. The medium was collected, and the cells lysed for 15 min in buffer containing 50 mM Tris-HCl, 150 mM NaCl, 1% (v/v) Triton X-100 and 1% (v/v) protease inhibitor cocktail (Calbiochem) at pH 7.4 on ice. Protein fractions of medium and lysate were centrifuged at 13 500 rpm at 4°C for 10 min. The cell pellet was discarded. The supernatant was denatured and run under reducing conditions on a 3%–8% Tris-Acetate precast gel (NuPAGE) for 135 min at 100V in Tris-Acetate running buffer supplemented with antioxidant. Transfer of protein bands onto a nitrocellulose membrane was performed at 15V overnight. The membrane was blocked using 5% (w/v) milk powder in Tris Buffered Saline with Tween (TBST) for 30 min at room temperature (RT) and incubated with antibodies against collagen I α I (Novus Biologicals NB600-408; rabbit, 1:500) and glyceraldehyde 3-phosphate dehydrogenase (GAPDH; Thermo Fisher Scientific AM4300, mouse, 3:1000) as loading control for 1.5 hours at RT. After repeated rinsing with TBST, the membrane was incubated for 1.5 hours at RT with horseradish peroxidase (HRP)-conjugated antibodies diluted in the blocking solution (1:5000) against mouse and rabbit, respectively. The wash step was repeated, and detection was performed using Promega WB-ECL reaction reagents and autoradiography films with overnight exposure and subsequent development.

RESULTS

Genetic analysis

Exome sequencing identified a heterozygous c.1178A>G, p.Tyr393Cys pathogenic mutation in exon 9 of *P4HB* (NM_000918.3) (figure 3). This mutation is predicted to replace tyrosine at position 393 with a cysteine; in silico analysis supports its likely pathogenicity. This result was confirmed by Sanger sequencing. DNA was only available from the mother, and she tested negative for this mutation; the father's sample was not available for analysis.

Electron microscopy

Patient elastic tissue and fibroblasts were unremarkable with no collagen flowers seen (figure 4A). The mean collagen fibril diameter was 74 nm, with a range from 39 to 95 nm (RMS deviation 7–10 nm) with a mean collagen bundle diameter of 3.3 μ m. These parameters were within the normal range for patient age. Adjacent to the sweat gland tubules, there were areas of small collections of individual collagen fibrils with proteoglycans wrapping around them (figure 4B), the significance of which is uncertain.

Collagen analysis

Collagen deposited by patient fibroblasts in culture showed no statistically detectable differences (with a P value of 0.76 for the quantified midintensity and 0.34 for high-intensity deposition) compared with controls (figure 5). While there was a slight trend towards a steady-state decrease in collagen secretion, at short times of incubation in the presence of ascorbic acid, we detected no differences in gross organisation or secretion of type 1 collagen.

Genotype-phenotype correlations

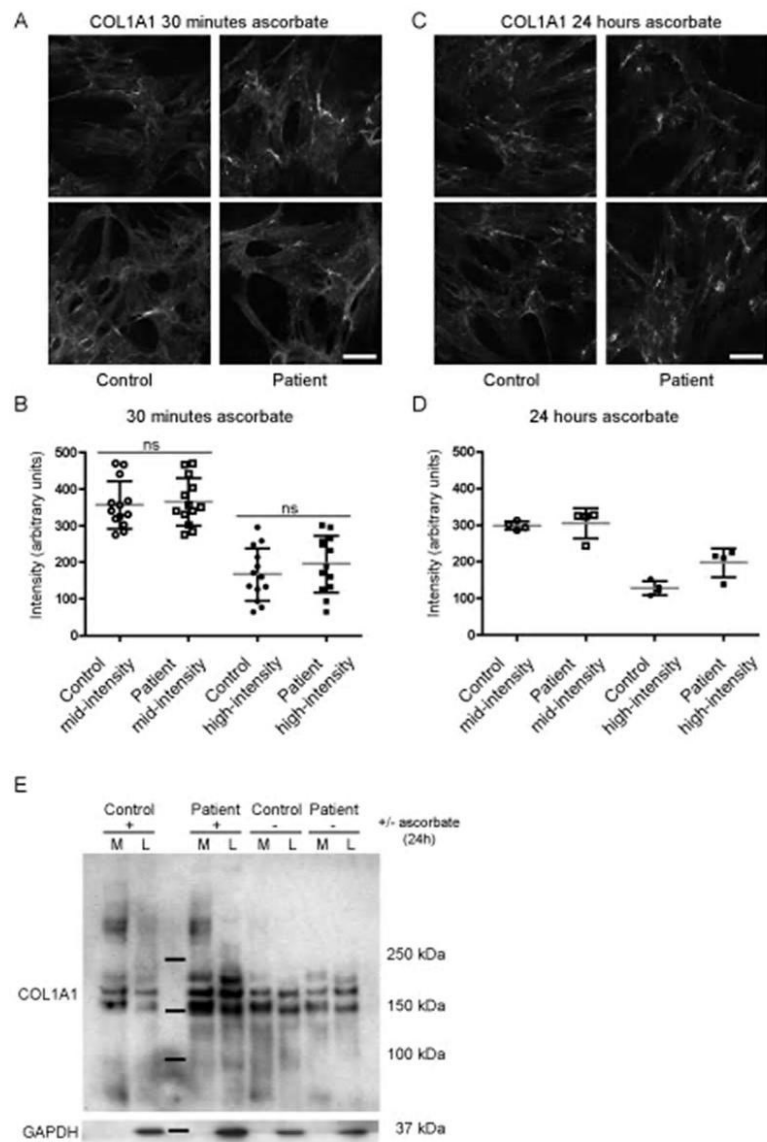


Figure 5 Immunofluorescence and subsequent quantification of extracellular collagen I of control and patient-derived skin fibroblasts after 30 min (A,B) and 24 hours (C,D) of ascorbic acid-2-phosphate. 5. (A–C) Maximum projection of Z-stacks consisting of 24 slices with $\Delta z=0.08\mu\text{m}$. Scale bar indicates $50\mu\text{m}$. (C,D) Quantification of A and C for midintensity values (16%–62% or 40–159; empty shapes) and high-intensity values (63%–100% or 160–255; filled shapes) for control (circles) and patient (squares). The mean is displayed in grey. Error bars show the SD.(E) Western blot showing COL1a1 levels in control and patient primary fibroblasts. Lanes show fractions of medium (M) and lysate (L) of control and patient cells incubated with (+), or without ascorbate acid (–) for 24 hours. Expected band size for COL1a1 monomer 140 kDa, dimer with COL1a2 270 kDa, trimer 400 kDa.

DISCUSSION

P4HB, also known as ‘Procollagen-proline 2-oxoglutarate-4-dioxygenase, beta subunit (OMIM $\#176790$)’ maps to chromosome 17q25.3 and spans 18Kb with 10 exons (Pajunen *et al*).

P4HB encodes for protein disulfide isomerase (PDI) and, in nucleated cells, in its tetrameric form, assists in the formation of 4-hydroxyproline in collagen by *P4HA* prolyl hydroxylase. It is likely that the recurrent missense mutation in *P4HB* results in a

very specific form of bone fragility by interfering with collagen formation, which could conceivably be associated with additional changes in deposition and/or assembly of extracellular matrix (ECM).

CCS (OMIM 112240) is mainly characterised by bone fragility, craniosynostosis, ocular proptosis, hydrocephalus and a distinctive facial appearance. Intelligence is reported to be normal. In their original report, Cole and Carpenter described two patients¹: patient 1 had multiple fractures, frontal bossing, proptotic appearance and progressive hydrocephalus. Skeletal survey at 4 months of age showed numerous metaphyseal irregularities and compression fractures in the long bones. Progressive craniosynostosis was noted at 9 months of age. His voice was noted to be high pitched, but words were reported to be well formed. Patient 2 had hypotonia and failure-to-thrive identified at 2 months of age and facial dysmorphism was noted including micrognathia, frontal bossing and proptotic appearance. A skeletal survey showed multiple metaphyseal compression fractures. In the report, the authors concluded that both patients had a variant form of OI but with additional findings of multiple metaphyseal irregularities, craniosynostosis, hydrocephalus, ocular proptosis and a distinctive face that was not characteristic of any known syndromes with craniosynostosis or type of OI. Both patients were reported to have normal neurological development (when reported patient 1 was 4 years and patient 2 was 19 months of age).

Rauch *et al* described re-evaluation of these children as adults and reported both patients 1 and 2 as using a wheelchair with severe scoliosis and marked limb deformities; treatment with intravenous pamidronate over the previous 12 months in patient 1 and over 3 years in patient two had not altered their clinical status.⁴ The radiographs in that paper clearly show the additional effects of prolonged immobility in association with the severe bone disease, notably gracile long bones with popcorn epiphyses. Whole exome sequencing (WES) on genomic DNA was performed on both patients, which revealed the same heterozygous variant c.1178A>G, p.Tyr393Cys in exon 9 of *P4HB*. The mechanism pertaining to gain-of-function of mutant PDI protein was proposed and tested, which has not been replicated in the current study.

In this report, we identified the same missense mutation in another patient with severe OI who has clinical overlap with CCS and how this variant results in severe bone fragility. The obvious difference in our patient clinically is the age at which pamidronate was commenced (7 months), which appears to have ameliorated the phenotype with reduced fracture frequency, vertebral fracture incidence and reduced distress, which she suffered in the first few months of life. This may be supportive of early benefits of commencing treatment in patients with CCS.

The clinical phenotype appears consistent in patients reported so far with CCS but interestingly, there also appears to be an emerging, distinctive radiological phenotype with what appears to be a definitive phenotypic clue to the underlying genotype (*P4HB* mutation). Figure 2 demonstrates metadiaphyseal 'crumpling' fractures with metaphyseal sclerosis in the long tubular bones, findings that are uncommon in OI and may be a pointer to the role in disease of this specific mutation. Patient 1 described by Cole and Carpenter also has the same metadiaphyseal compression fractures with irregular sclerosis and lucencies and subsequent bowing deformity (figure 1 in Cole and Carpenter paper). Therefore, we propose that the metaphyseal 'crumpling' fractures with metaphyseal sclerosis may be a clue to the underlying genotype. This was dissimilar to the findings of osteopaenia and bowing with fractures in radiographs of patient

reported by Amor *et al*,² who did not share similar radiological findings as described above and tested negative for the *P4HB* recurrent mutation. Interestingly, reviewing the radiographs of the patient reported by Balasubramanian *et al*³ with phenotypic CCS who has homozygous variants in *CRTAP*, the metaphyseal compression fractures and sclerosis were not present, thus adding further evidence to this observation.

P4HB encodes the beta subunit of prolyl 4-hydroxylase, which is a highly abundant multifunctional enzyme belonging to the PDI family. The PDI family comprises PDI and PDI-like proteins with more than 20 members having been studied in humans.⁷ PDI is usually found in the endoplasmic reticulum (ER) but can also be released to function at the ECM or cell surface. PDI is the most abundant family member comprising 0.8% of total cellular protein in yeast and mammalian cells⁸ and is critical for cell viability in yeast.⁹ It functions as a redox catalyst and as a chaperone by preventing protein aggregation in the ER or by retaining proteins within the ER where necessary.¹⁰

P4HB comprises four TRX-like (thioredoxin-like) domains termed a, b, b', a', a linker (x) and a C-terminal extension domain organised in the order of abb'xa'c. However, only two of these four domains (a and a') have disulphide isomerase activity and are separated from each other by the enzymatically inactive b and b' domains. P4H exists as a 2:2 stoichiometric heterotetramer and is generally called the α -subunit existing as three isoforms (α (I), α (II) and α (III)). All three α subunits require a β -subunit, that is, PDI to form an enzymatically active complex. The PDI β -subunit is required to maintain the solubility of the P4H α subunit and keep the complex within the ER. As well as the disulphide isomerase activity, the tetrameric form is active in hydroxylation of proline residues within procollagen. Therefore, PDI as a stand-alone disulphide isomerase and as a P4H heterotetramer is crucial for procollagen biosynthesis.¹¹

Hence, PDI family members function as molecular chaperones and as disulphide oxidoreductase/isomerases. As a result, they can make, break or rearrange disulphide bonds in client proteins, including collagen and ECM proteins. These disulphide bonds are essential to ensure the protein has structural stability and also to keep multimeric complexes together. Cellular function relies on the ability of proteins to adopt their correct folds, and it is very well recognised that misfolded proteins are known to result in disease; this is especially true in OI. However, it should be noted that the recurrent *P4HB* mutation reported here does not directly alter the catalytic CGHC motif of PDI. Rather, the mutation results in the gain of a cysteine residue, which may lead to a 'gain of function' phenotype. In this scenario, PDI could either misoxidise, fail to interact appropriately with P4HA or interact inappropriately with other proteins through spurious intermolecular disulphide bonds.

The recurrent mutation in *P4HB*, c.1178A>G, p.Tyr393Cys, sits towards the C-terminus and is said to interfere with disulphide isomerase function. *P4HB* catalyses the hydroxylation of proline residues within the X-Pro-Gly repeats in the procollagen helical domain. However, our data indicate no major difference in extracellular collagen type I deposition by patient fibroblasts, as judged by immunofluorescence. While there was a minor trend towards an increase in extracellular collagen type I compared with the control, this was not statistically significant. Our experiments do not define whether this collagen is triple helical and we have not examined other collagen types or proteoglycans. Given the interdependence of ECM components in assembly of a functional matrix, our data suggest that it is possible that it is

Genotype-phenotype correlations

the organisation and assembly of the functional ECM that is perturbed rather than the secretion of collagen type I per se. We postulate that the defect might arise from overall changes in matrix composition or architecture. However, this will require further functional analysis of whole ECM, which is ongoing.

PDI proteins have already been implicated in human disease and are said to have roles in haemostasis, infectious disease, lipid homeostasis, neurodegeneration, cancer and infertility^{12–16}; we now know that it is also involved in bone disease (OI).

In this report, we demonstrate that CCS is caused by a recurrent *P4HB* mutation with a very specific clinical and radiological phenotype and provide further clues to the underlying genotype. We have shown that possibly *P4HB* acts on collagen synthesis not by affecting the global secretion of procollagen type I from cells but in maintaining a functional ECM.

Acknowledgements We thank this family for their participation in this report; Professor David J Amor, University of Melbourne, for providing us with input regarding his patient and Dr Adam Benham, University of Durham, for critical review of the manuscript. We would also like to thank Matthew Parker and Emilie Jarratt at Sheffield Diagnostic Genetics Service for their assistance with the exome sequencing.

Contributors All authors contributed to preparation and critical review of manuscript; MB: study design, writing up the manuscript, recruitment of patient and phenotyping; RP, NJB and MZM: patient phenotyping; ACO: radiology input; RCP: exome sequencing; BEW: EM studies; JM and DJS: functional studies on patient fibroblasts.

Funding This research was supported by The Sheffield Children's Hospital Charity (TCHC) grant number CA15001 and a postgraduate scholarship from the University of Bristol to JM.

Competing interests None declared.

Ethics approval Yorkshire and Humber REC.

Provenance and peer review Not commissioned; externally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- 1 Cole DE, Carpenter TO. Bone fragility, craniostylosis, ocular proptosis, hydrocephalus, and distinctive facial features: a newly recognized type of osteogenesis imperfecta. *J Pediatr* 1987;110:76–80.
- 2 Amor DJ, Savarirayan R, Schneider AS, Bankier A. New case of Cole-Carpenter syndrome. *Am J Med Genet* 2000;92:273–7.
- 3 Balasubramanian M, Pollitt RC, Chandler KE, Mughal MZ, Parker MJ, Dalton A, Arundel P, Offiah AC, Bishop NJ. CRTAP mutation in a patient with Cole-Carpenter syndrome. *Am J Med Genet A* 2015;167A:587–91.
- 4 Rauch F, Fahiminiya S, Majewski J, Carrot-Zhang J, Boudko S, Glorieux F, Mort JS, Bächinger HP, Moffatt P. Cole-Carpenter syndrome is caused by a heterozygous missense mutation in *P4HB*. *Am J Hum Genet* 2015;96:425–31.
- 5 Garbes L, Kim K, RieB A, Hoyer-Kuhn H, Beleggia F, Bevoit A, Kim MJ, Huh YH, Kweon HS, Savarirayan R, Amor D, Kakadia PM, Lindig T, Kagan KO, Becker J, Boyadjiev SA, Wollnik B, Semler O, Bohlender SK, Kim J, Netzer C. Mutations in *SEC24D*, encoding a component of the COPII machinery, cause a syndromic form of osteogenesis imperfecta. *Am J Hum Genet* 2015;96:432–9.
- 6 McCaughey J, Miller VJ, Stevenson NL, Brown AK, Budnik A, Heesom KJ, Alibhai D, Stephens DJ. TFG promotes organization of transitional ER and efficient collagen secretion. *Cell Rep* 2016;15:1648–59.
- 7 Benham AM. The protein disulfide isomerase family: key players in health and disease. *Antioxid Redox Signal* 2012;16:781–9.
- 8 Hatahet F, Ruddock LW. Protein disulfide isomerase: a critical evaluation of its function in disulfide bond formation. *Antioxid Redox Signal* 2009;11:2807–50.
- 9 LaMantia M, Miura T, Tachikawa H, Kaplan HA, Lennarz WJ, Mizunaga T. Glycosylation site binding protein and protein disulfide isomerase are identical and essential for cell viability in yeast. *Proc Natl Acad Sci U S A* 1991;88:4453–7.
- 10 Kosuri P, Alegre-Cebollada J, Feng J, Kaplan A, Inglés-Prieto A, Badilla CL, Stockwell BR, Sanchez-Ruiz JM, Holmgren A, Fernández JM. Protein folding drives disulfide formation. *Cell* 2012;151:794–806.
- 11 Ishikawa Y, Bächinger HP. A molecular ensemble in the rER for procollagen maturation. *Biochim Biophys Acta* 2013;1833:2479–91.
- 12 Gruber CW, Cemazar M, Heras B, Martin JL, Craik DJ. Protein disulfide isomerase: the structure of oxidative folding. *Trends Biochem Sci* 2006;31:455–64.
- 13 Pajunen L, Jones TA, Goddard A, Sheer D, Solomon E, Pihlajaniemi T, Kivirikko KI. Regional assignment of the human gene coding for a multifunctional polypeptide (*P4HB*) acting as the beta-subunit of prolyl 4-hydroxylase and the enzyme protein disulfide isomerase to 17q25. *Cytogenet Cell Genet* 1991;56:165–8.
- 14 Soares Moretti AI, Martins Laurindo FR. Protein disulfide isomerases: Redox connections in and out of the endoplasmic reticulum. *Arch Biochem Biophys* 2017;617:106–19.
- 15 van Anken E, Braakman I. Endoplasmic reticulum stress and the making of a professional secretory cell. *Crit Rev Biochem Mol Biol* 2005;40:269–83.
- 16 Wilkinson B, Gilbert HF. Isomerase Pdisulfide. *Biochim Biophys Acta* 2004;1699:35–44.

G. Manuscript: Compound Heterozygous Variants in NBAS as a cause of atypical osteogenesis imperfecta

Bone 94 (2017) 65–74



Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone



Full Length Article

Compound heterozygous variants in NBAS as a cause of atypical osteogenesis imperfecta



M. Balasubramanian^{a,b,*}, J. Hurst^c, S. Brown^d, N.J. Bishop^{b,e}, P. Arundel^b, C. DeVile^b, R.C. Pollitt^{e,f}, L. Crooks^{f,g}, D. Longman^h, J.F. Caceres^h, F. Shackleyⁱ, S. Connolly^j, J.H. Payne^k, A.C. Offiah^{b,e}, D. Hughes^l, DDD Study^m, M.J. Parker^a, W. Hideⁿ, T.M. Skerry^o

^a Sheffield Clinical Genetics Service, Sheffield Children's NHS Foundation Trust, UK

^b Highly Specialised Service for Severe, Complex and Atypical OI, UK

^c NE Thames Clinical Genetics Service, Great Ormond Street Hospital, UK

^d Sheffield RNAi Screening Facility, Department of Biomedical Sciences, University of Sheffield, UK

^e Academic Unit of Child Health, University of Sheffield, UK

^f Sheffield Diagnostic Genetics Service, Sheffield Children's NHS Foundation Trust, UK

^g Department of Biosciences and Chemistry, Sheffield Hallam University, UK

^h MRC Human Genetics Unit, IGMM, University of Edinburgh, UK

ⁱ Department of Paediatric Immunology, Sheffield Children's NHS Foundation Trust, UK

^j Department of Paediatric Hepatology, Sheffield Children's NHS Foundation Trust, UK

^k Department of Paediatric Haematology, Sheffield Children's NHS Foundation Trust, UK

^l Department of Histopathology, Sheffield Teaching Hospitals NHS Foundation Trust, UK

^m DDD Study, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK

ⁿ Centre for Computational Biology, Sheffield Institute of Translational Neuroscience, University of Sheffield, UK

^o Mellanby Bone Research Centre, Department of Oncology & Metabolism, University of Sheffield, UK

ARTICLE INFO

Article history:

Received 7 September 2016

Revised 17 October 2016

Accepted 21 October 2016

Available online 24 October 2016

Keywords:

Bone

Fragility

Osteogenesis imperfecta

NBAS

Nonsense mediated decay (NMD)

Secretory pathway

Collagen expression

ABSTRACT

Background: Osteogenesis imperfecta (OI), the commonest inherited bone fragility disorder, affects 1 in 15,000 live births resulting in frequent fractures and reduced mobility, with significant impact on quality of life. Early diagnosis is important, as therapeutic advances can lead to improved clinical outcome and patient benefit.

Report: Whole exome sequencing in patients with OI identified, in two patients with a multi-system phenotype, compound heterozygous variants in NBAS (neuroblastoma amplified sequence). Patient 1: NBAS c.5741G>A p.(Arg1914His); c.3010C>T p.(Arg1004*) in a 10-year old boy with significant short stature, bone fragility requiring treatment with bisphosphonates, developmental delay and immunodeficiency. Patient 2: NBAS c.5741G>A p.(Arg1914His); c.2032C>T p.(Gln678*) in a 5-year old boy with similar presenting features, bone fragility, mild developmental delay, abnormal liver function tests and immunodeficiency.

Discussion: Homozygous missense NBAS variants cause SOPH syndrome (short stature; optic atrophy; Pelger-Huet anomaly), the same missense variant was found in our patients on one allele and a nonsense variant in the other allele. Recent literature suggests a multi-system phenotype. In this study, patient fibroblasts have shown reduced collagen expression, compared to control cells and RNAseq studies, in bone cells show that NBAS is expressed in osteoblasts and osteocytes of rodents and primates. These findings provide proof-of-concept that NBAS mutations have mechanistic effects in bone, and that NBAS variants are a novel cause of bone fragility, which is distinguishable from 'Classical' OI.

Conclusions: Here we report on variants in NBAS, as a cause of bone fragility in humans, and expand the phenotypic spectrum associated with NBAS. We explore the mechanism underlying NBAS and the striking skeletal phenotype in our patients.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

NBAS (neuroblastoma-amplified sequence), also previously referred to as NAG (neuroblastoma-amplified gene) contains 52 exons, spans 420 kb and is mapped to chromosome 2p24.3 [1]. It was initially identified as a gene co-amplified with the N-myc (MYCN) gene in

* Corresponding author at: Sheffield Clinical Genetics Service, Sheffield Children's NHS Foundation Trust; Western Bank, Sheffield S10 2TH, UK.

E-mail address: meena.balasubramanian@nhs.net (M. Balasubramanian).

neuroblastoma cell lines [2]. *NBAS* was initially described as a novel factor involved in the nonsense mediated (NMD) decay pathway in human cells, in zebrafish and in nematodes [3,4]. It was shown that *NBAS* acts in concert with core NMD factors to co-regulate a large number of endogenous RNA targets [5]. Subsequently, *NBAS* was also identified as a component of the syntaxin 18 complex, which is involved in Golgi-to-endoplasmic reticulum (ER) retrograde transport [6,7]. *NBAS* is also said to be an important component of the ER tethering complex [8]. Using whole exome sequencing, we identified compound heterozygous variants in *NBAS* in two patients with bone fragility: Patient 1 presented with significant short stature, bone fragility requiring treatment with bisphosphonates, developmental delay and immunodeficiency and was independently investigated for a novel cause of bone fragility (having tested negative for all published variants in OI). Patient 2 was recruited to the Deciphering Developmental Disorders (DDD) study and underwent trio whole exome sequencing.

Homozygous missense variants in *NBAS* have been implicated in a hereditary short stature syndrome referred to as SOPH syndrome (short stature; optic atrophy; Pelger-Huet anomaly) observed in the Yakut Siberian population isolate [9]. Compound heterozygous mutations in *NBAS* have also been described in acute onset liver failure [10]. Further reports have suggested a multi-system phenotype [11,12,13]. Bone fragility severe enough to need bisphosphonate therapy has not reported as a feature, so far, in association with *NBAS*. Therefore, we propose that compound heterozygous mutations in *NBAS* can account for a significant form of bone fragility and needs to be considered in the differential diagnosis of osteogenesis imperfecta (OI).

2. Clinical report

2.1. Patient 1

This 10-year old boy is the second child of healthy, non-consanguineous parents, of North European origin, with no significant family history. He was born by spontaneous breech delivery at 33 weeks gestation following a normal pregnancy, with a birth weight of 1.75 kg (9th centile). He required continuous positive airway pressure ventilation

(CPAP) for 24 h and phototherapy for jaundice. He had recurrent infections and significant problems with his feeding, requiring percutaneous gastrostomy insertion to maintain adequate nutrition. He had recurrent admissions to hospital with infections including severe recurrent shingles with worsening of liver function during infectious episodes and consequent progressive lymphopenia and hypogammaglobulinaemia requiring immunoglobulin replacement. He was diagnosed with a horizontal nystagmus, bilateral optic atrophy and myopia, needing corrective glasses, but his hearing was reported as normal. He has moderate intellectual disability and growth parameters remain well below the 0.4th centile. In early childhood, he went on to sustain fractures of his tibia and metatarsals, following minimal trauma. He attends an integrated play school, as he has been diagnosed with an autism spectrum disorder.

On examination at 9-years of age, he was dysmorphic, with proptosis, progeric appearance to his skin, bilateral low-set ears, dentinogenesis imperfecta, and bilateral 5th finger clinodactyly with bulbous tips to his fingers and toes (Fig. 1a–e). He has a high-pitched voice and growth parameters at 7 years of age were: weight – 13.1 kg (<0.4th centile), height – 88 cm (≪0.4th centile), head circumference – 49.5 cm (0.4th–2nd centile), with no evidence of asymmetry. Repeat cranial ultrasounds and MRI-brain scans did not show any evidence of ventricular dilatation despite initial concerns regarding a large anterior fontanelle. A dysplasia skeletal survey showed multiple Wormian bones, slender tubular bones and osteopenia (Fig. 2a–d). A transiliac bone biopsy at 7-years of age, following recurrent low-trauma fractures, demonstrated osteoporosis with high bone turnover with marked periosteal bone resorption, which was different to appearances in classical OI (Fig. 3a–b). In terms of his skeletal phenotype, he had a tender back with loss of vertebral height on lateral spine radiograph and a low lumbar bone mineral areal density (BMAD) with a Z-score of – 3.5 at 9-years of age. He was commenced on pamidronate with remarkable improvement to his bone health. Collagen species analysis on SDS-Page gel was normal with no obvious shift, but electron microscopy showed multiple small collagen flowers, variable shaped collagen fibrils with a mean collagen fibril diameter (CFD) of 75 nm (Fig. 3c–d). Genetic testing so far included: normal 60 K ArrayCGH, FRAX, UPD7 and 11p15

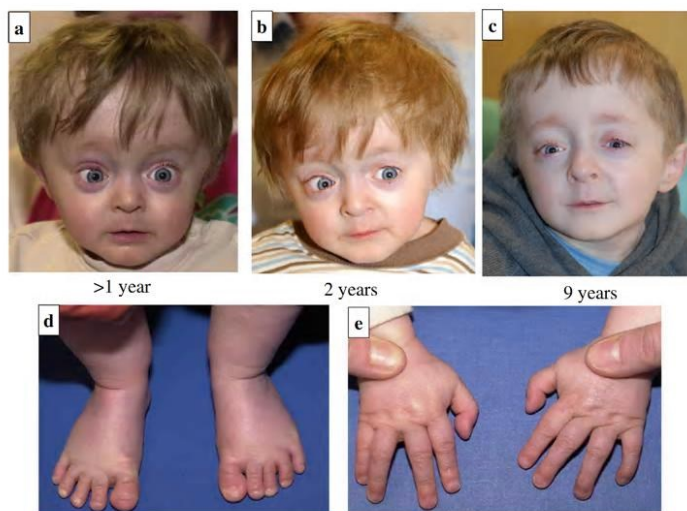


Fig. 1a–e. Panels a–c: Facial features as an infant and aged > 1, 2 and 9 years showing grey sclerae, broad forehead, bilateral low-set ears, proptosis and progeric appearance; panels d–e: Hands and feet at 2-years of age.

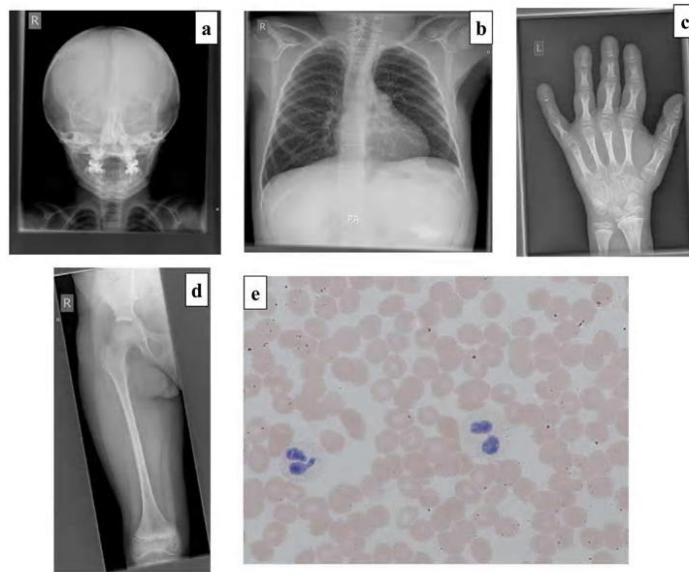


Fig. 2a–d. Radiographs demonstrating slender ribs, tubular long bones with thin cortices and osteopenia consistent with a diagnosis of OI. 2a: AP skull radiograph (aged 6 years). There are multiple Wormian bones; the clavicles are slender and the erupted teeth are relatively dense. The anterior fontanelle remains open. 2b–d: Selected images from a full dysplasia skeletal survey (aged 9 years 8 months). 2b: Left hand. There is significant periarticular osteopenia (see panel c). The metacarpals (and less marked) the proximal phalanges are overmodelled and there is an ivory epiphysis of the terminal phalanx of the fifth finger. The terminal tufts are prominent. 2c: AP chest. The ribs and clavicles are slender; however, there are no fractures and vertebral body height is preserved. Note the presence of a gastrostomy. 2d: AP right femur. Overmodelled with slender diaphysis and relatively flared distal metaphysis. Periarticular osteopenia is again noted (see panel a). 2e: Peripheral blood film in Patient 1 demonstrating a hypolobulated neutrophil (left) and a Pelger-Huet cell (right).

methylation testing, targeted exome panel including all published genes known to cause bone fragility, all of which were reported as negative.

2.2. Patient 2

This 6-year old boy is the first child of healthy, non-consanguineous parents (mother is of Northern-Spanish origin whilst father is of Italian origin) with no significant family history. He was born at 38 weeks gestation, by spontaneous delivery following a normal pregnancy with a birth weight of 2.44 kg (2nd centile). He did not need to go to Special Care Baby Unit but went on to develop a spiral fracture of his left femur whilst being positioned for a feed at 3-months of age. A skeletal survey performed at the time showed osteopenic slender bones, thin skull vault with a large anterior fontanelle, but no Wormian bones. These features were thought to be consistent with but not completely diagnostic of type IV OI. He had recurrent infections, problems with his feeding and poor weight gain. He also went on to develop recurrent episodes of ketotic hypoglycaemia, which resolved spontaneously and hypogammaglobulinaemia, needing 2-weekly immunoglobulin infusions. He was diagnosed with abnormal liver function when he had several episodes of temporarily elevated transaminases triggered by infections. He was diagnosed with a horizontal nystagmus and bilateral optic atrophy, but his hearing was reported to be normal. He has mild intellectual disability, managing in mainstream school with 1-to-1 help. His growth parameters remain below the 0.4th centile, with some preservation of head circumference (0.4th–2nd centile).

On examination at 5-years of age, he was dysmorphic, with brachycephaly, proptosis, progeric appearance to his skin, bilateral low-set ears, greyish sclerae, and bilateral 5th finger clinodactyly with bulbous tips to his fingers and toes (Fig. 4a–c). He has a high-pitched, distinctive voice and growth parameters were: weight ~14.4 kg (<0.4th centile),

height ~96.9 cm (<0.4th centile), head circumference ~49.6 cm (0.4th–2nd centile), with no evidence of asymmetry. Repeat cranial ultrasounds, in view of persistent large anterior fontanelle, and a MRI-brain scan at one year of age showed non-specific white matter changes, but otherwise normal appearances. In terms of his skeletal phenotype, he has gone on to sustain stress fractures of his feet, diagnosed on bone scan and a low lumbar bone mineral areal density (BMAD) with a Z-score of -4.01 at 5-years of age. He has recently been started on pamidronate with a good response to therapy. A dysplasia skeletal survey subsequently confirmed similar features as reported previously (Fig. 4d–e). Genetic testing so far included: normal 60 K ArrayCGH, targeted exome panel including genes known to cause OI, all of which were reported as negative. He was subsequently recruited to the DDD Study (Decipher Patient ID: 264693).

3. Materials and methods

Patient 1 was recruited into a research project to study atypical forms of OI, to establish genotype: phenotype correlations therein. Ethical approval was obtained from the local regional ethics committee (LREC) to undertake phenotyping and genetic work-up in this group of patients. Patient 2 was recruited to the DDD Study.

3.1. DNA sequencing and mutation analysis in Patient 1

Total genomic DNA was isolated from 2 to 5 ml peripheral blood using standard extraction methods. Whole exome sequencing was performed by Personalis using their ACE Exome™ Assay. The data were analysed with the Personalis ACE pipeline; this uses BWA (version number 0.7.5a-r405) for alignment, GATK's UnifiedGenotyper for variant calling and GATK's VQSR to provide a site quality score using the single patient sample. The average depth in the target region was $115\times$.

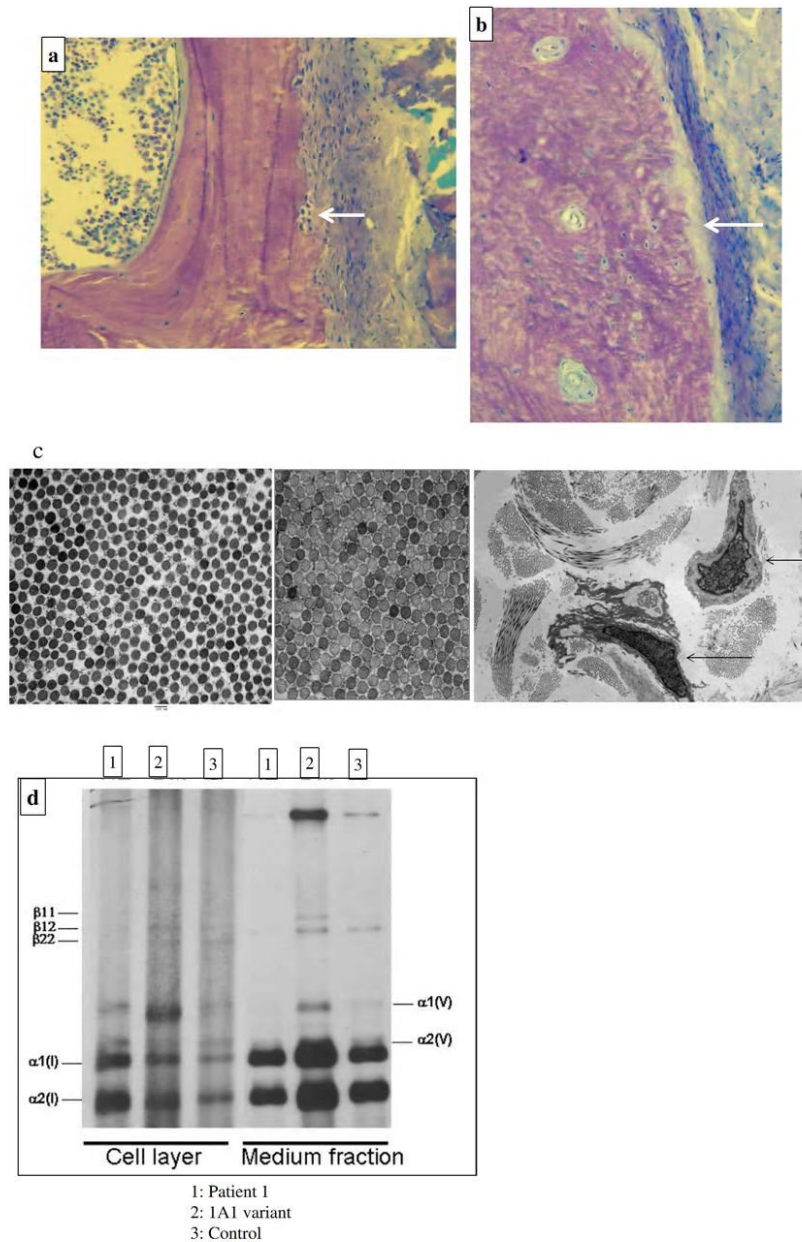


Fig. 3a–b. Panel a: Toluidine blue-stained section of an undecalcified trans-iliac bone biopsy, original magnification 400× demonstrating cortex, with periosteum on the right showing high turnover osteopenia with marked sub-periosteal bone resorption (arrow) and normal lamellar bone matrix structure in Patient-1, aged 9 years; panel b: appearance in 'Classical OI' with abnormal matrix pattern and increased periosteal bone formation surface (arrow), Toluidine blue; Original magnification of 400×. c–d: Panel c: Electron microscopy of skin biopsy from Patient 1 (left image) compared to normal control (middle image) showing normal collagen in mid-reticular dermis with mildly reduced mean collagen fibril diameter (CFD); original magnification of 20,000×; lower magnification (right image) with arrows indicating deep reticular dermal fibroblast with expanded protein filled rough endoplasmic reticulum; original magnification ×2600. Panel d: Normal collagen species analysis with no deviation from control (1 and 3) in comparison to patient with a *COL1A1* variant (2).

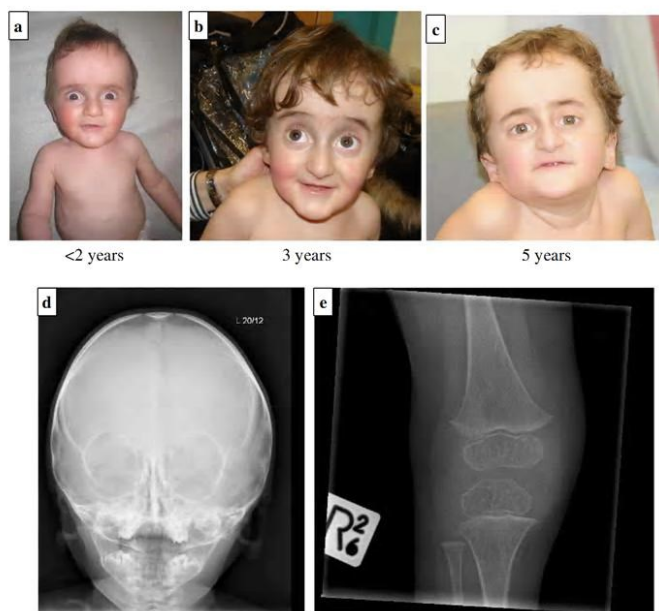


Fig. 4a–e. Panels a–c: Facial appearance of Patient 2 aged <2, 3 and 5 years demonstrating similar facial dysmorphism to Patient 1 with progeric appearance, grey sclerae, broad forehead, bilateral low-set ears, proptosis; panels d–e: Skull and right lower limb X-rays demonstrating thin skull vault, slender bones and osteopenia.

Variants were extensively annotated with AnnotL, a Personalis tool. The annotation includes frequencies by subpopulation from 1000 genomes and the Exome Sequencing Project; information from dbNSFP; predicted mutational impacts from tools including SIFT, Polyphen2 and MutationTaster; known associations from disease databases such as OMIM, ClinVar and the GWAS Catalog; and data from pathway and network tools including Reactome and MINT. Calls were quality filtered by Personalis and variants reported at >1% in any of the subpopulations were excluded. Personalis was supplied with clinical features of the patient and used these in ranking variants, taking into account the similarity in features with reported disease-causing mutations in the same gene. Evidence for the top ranking variants was manually reviewed by Personalis to reduce the list to potentially causative variants.

Additionally, the full set of variants reported by Personalis was investigated using their annotation. Quality control was carried out using the following hard filters, as recommended by the GATK best practice guidelines when there are exome sequences for <30 samples plus an extra filter on genotype quality: for SNVs, $QD < 2$, $MQ < 40$, $FS > 60$, $HaplotypeScore > 13$, $MQRankSum < -12.5$, $ReadPosRankSum < -8$; for Indels, $QD < 2$, $ReadPosRankSum < -20$, $FS > 200$. Population frequencies from the Exome Aggregation Consortium (ExAC, version 0.3) were used in filtering. Only passed ExAC variants were used and alleles at multiallelic sites were left aligned to ensure they could be matched with the patient data. The number of ExAC individuals in a sub-population that are genotyped at a site can be low, meaning that the frequency is estimated with a large error. Therefore, from the ExAC set only variants that had a frequency >5% in any subpopulation (excluding the 'other' population), or of >1% with a total allele count >1000 were retained. These variants were filtered out from the Personalis dataset. Variants reported with a frequency >1% in any of the 1000 genomes, NHLBI ESP or UK10K cohort populations were also excluded. After filtering, those variants annotated as frameshift, splice site acceptor, splice site donor, or stop gained were

extracted and are described as loss of function (LoF) variants. Finally, a larger group of variants annotated within a targeted gene list was compiled, which are described as targeted missense variants. The targeted gene list consisted of genes reported in connection with skeletal dysplasia, identified in GWAS studies to be associated with a change in bone density, implicated in bone metabolism in mouse models, and genes with human phenotype ontology terms relating to increased susceptibility to fracture (HPO:0002659), and totalled approximately 600 genes. The LoF and targeted missense variants were manually reviewed, including looking for compound heterozygotes, to assess whether they might contribute to the patient's phenotype.

3.2. DNA sequencing and mutation analysis in Patient 2

Trio-based exome sequencing was performed for Patient 2 and his parents as part of the DDD Study, as previously described [14,15]. Putative *de novo* and inherited variants were identified from exome data using DeNovoGear software [15,16] and were validated using targeted Sanger sequencing.

3.3. Western blot and microscopic analysis of collagen human primary fibroblasts (HPF) on Patient 1

HPF was grown in fully supplemented AmnioMAX C-100 media (Gibco) in hypoxic conditions. Prior to lysis, cells were washed twice with PBS and lysed for 20 min at 4 °C in IP buffer (10 mM Tris-HCl [pH 8], 150 mM NaCl, 1 mM EGTA, 1% NP-40, 0.2% Na-deoxycholate, complete protease inhibitor (Roche), 1 mM DTT). Proteins were resolved by SDS-PAGE using 3%–8% tris-acetate gel (Life Technologies) and analysed for the presence of NBAS by probing with NBAS antibody (Abcam). Uniform protein loading was confirmed by probing with tubulin antibody (TUB 2.1, Sigma-Aldrich).

NBAS Patient 1 cultured fibroblasts and control sample were grown for 3 days in 96 well plates, fixed and stained with anti-Col1A1 antibody (green) and Hoechst (blue) and imaged using a high content microscope.

4. Results

4.1. Patient 1

Personalis reported four top ranking variants. Two were heterozygous variants in the NBAS, c.5741G>A p.(Arg1914His) and c.3010C>T

p.(Arg1004*) (Fig. 5a). The impact of the variants was determined for transcript NM_015909.3. The potential pathogenic effect of the missense variant and nonsense variant were assessed using Alamut Visual version 2.6 (Interactive Biosoftware, Rouen, France) and the Association of Clinical Genetic Science best practice guidelines for the evaluation of pathogenicity and the reporting of sequence variants in clinical molecular genetics.

The other two variants did not appear contributory to the phenotype and were of uncertain significance; a novel heterozygous variant of uncertain significance in *ARID1B*, c.1041_1043dupGGC (p.Ala350dup), and another heterozygous novel variant of uncertain significance in *SKIV2L*,

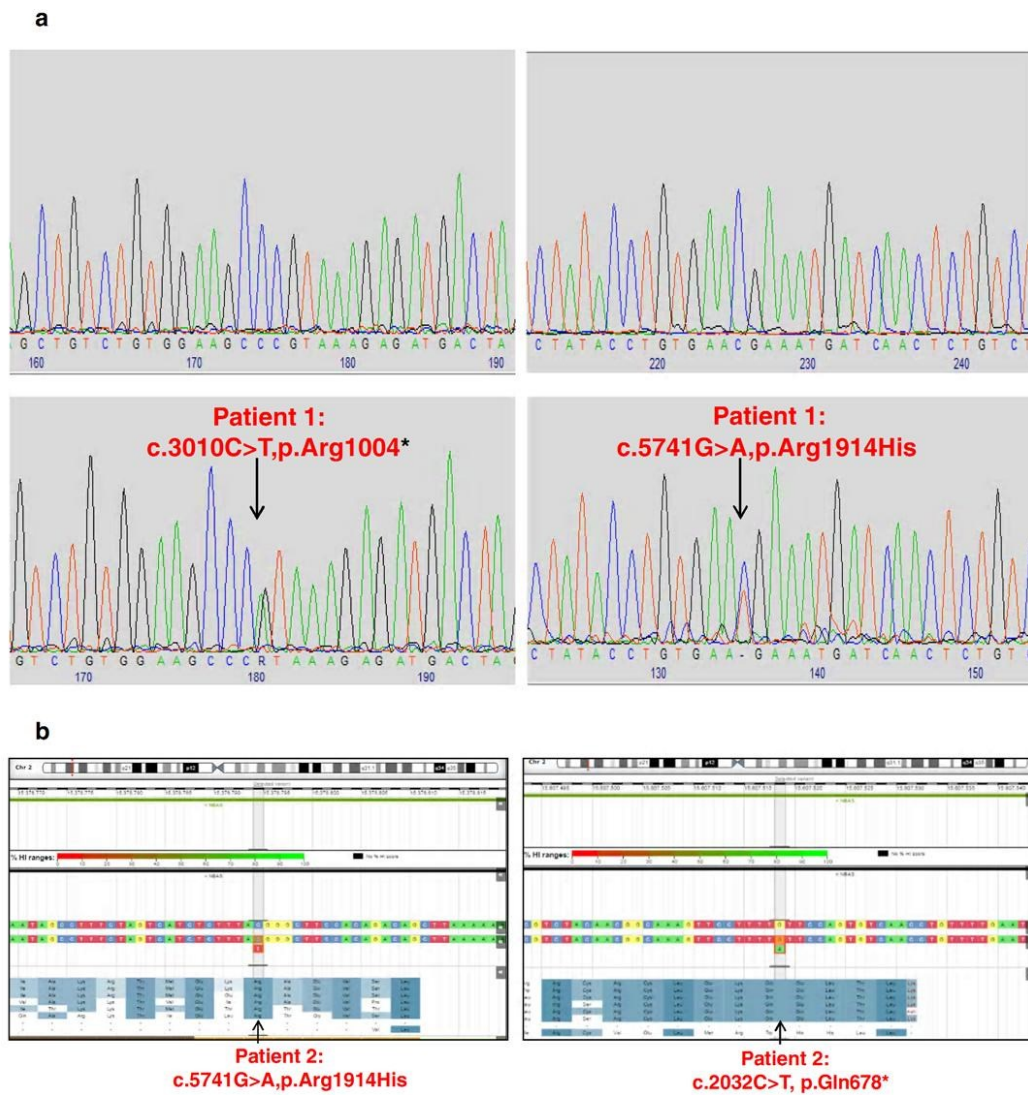


Fig. 5a and b. Electropherograms (forward sequence) demonstrating NBAS variants in Patient 1 compared to normal control and Patient 2 (sequence variant plot).

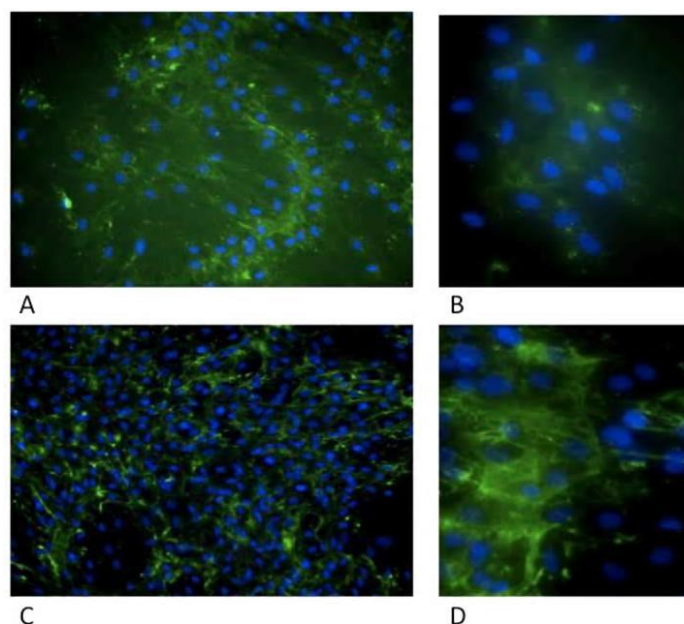


Fig. 6. NBAS patient 1 cultured fibroblasts (A, B) and control sample (C, D) were grown for 3 days in 96 well plates, fixed and stained with anti-Col1A1 antibody (green) and Hoechst (blue) and imaged using a high content microscope. (A, B) show increased diffuse cytoplasmic staining. Collagen bundles from control (C, D). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

c.3404 T>C (p.Ile1135Thr). Previously reported pathogenic variants in *ARID1B* associated with autosomal dominant mental retardation type 12 include nonsense and frameshifting insertions/deletions which result in haploinsufficiency [17]. c.1041_1043dupGGC (p.Ala350dup), is an in-frame duplication variant in exon 1 of the *ARID1B* gene which results in the addition of one extra alanine to a stretch of 11 consecutive alanine residues. The product of the *ARID1B* gene plays a role in chromatin remodeling for transcriptional activation and repression of select genes. To our knowledge, variants in this polyalanine tract have not been reported in association with disease. Variants in *SKIV2L* have been previously reported in the literature in association with trichoshepatoenteric syndrome-2 (THES2), an autosomal recessive disease [18]. However, both the variants are unlikely to account for Patient 1's phenotype.

The total number of detected variants was 189,298. After quality and frequency filtering, 19,772 remained. There were 89 LoF variants, none of which were in genes with an apparent function in bone. There were 23 targeted missense variants that were predicted to be likely benign following *in-silico* analysis.

The presence of the *NBAS* variants was confirmed by Sanger sequencing of the proband. Sequencing of the parents showed that c.3010C>T p.(Arg1004*) is present in the mother and c.5741G>A p.(Arg1914His) in the father. The c.5741G>A is the same missense variant which was described in homozygous form in patients with SOPH syndrome [5] c.3010C>T was reported by Haack et al. in a compound heterozygous form to be associated with acute liver failure [10].

Pilot studies of collagen expression and transport in *NBAS* cells cultured from patients described in this study, show that collagen secretion appears reduced and collagen bundles appear more diffuse, as compared with control cells consistent with interference with trafficking and secretion (Fig. 6). Furthermore, Western blot analysis shows reduced level of *NBAS* protein in patients, as compared to control cells

(Fig. 7). This implies that the compound heterozygous mutation reported in this patient, compromises the stability of the protein, resulting in a dramatic reduction in *NBAS* protein levels.

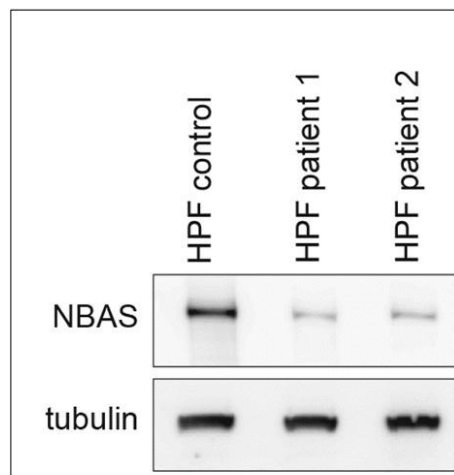


Fig. 7. Western blot on Patient 1 and 2 cultured fibroblasts showing reduced *NBAS* protein levels compared to controls; HPF: human primary fibroblasts.

4.2. Patient 2

Two heterozygous variants in the *NBAS* gene were identified, c.5741G>A p.(Arg1914His); c.2032C>T p.(Gln678*) (Fig. 5b) and segregation analysis showed that c.5741G>A p.(Arg1914His) is present in the mother and c.2032C>T p.(Gln678*) is present in the father. The c.5741G>A is the same missense variant which was described in homozygous form in patients with SOPH syndrome [5]. The c.2032>T is a novel variant and has not been reported in HGMD or present in the Exome Aggregation Consortium (ExAC) dataset. Western blot analysis of human primary fibroblasts (HPF) cultured from patients showed reduced level of *NBAS* protein in patients, as compared to control cells (Fig. 7).

5. Discussion

Osteogenesis imperfecta (OI) is a disease encompassing a group of disorders mainly characterised by bone fragility. There is a broad spectrum of clinical severity in OI, ranging from multiple fractures *in-utero* and perinatal death, to near-normal adult stature and low fracture incidence. Facial dysmorphism has been noted [19]. Silience et al., in 1979 provided the clinical classification, which has been expanded [20,21]. Defects in genes encoding type I collagen (*COL1A1/A2*) can be identified in 85% of patients with a clinical diagnosis of OI [21,22]. So far, several other genes have been implicated in rare forms of heritable bone fragility including autosomal dominant type V OI (*IFTM5*), X-linked osteoporosis (*PLS3* and *MBTPS2*), autosomal recessive forms (*BMP1/CREB3L1/CRTAP/FKBP10/P3H1/P4HB/PLD2/PP1B/SEC24D/SERPINF1/SERPINH1/SP7/SPARC/TMEM38B*) and heterozygous mutations in *WNT1/LRP5*.

NBAS (the human neuroblastoma amplified sequence gene) was first isolated using genome scanning techniques in neuroblastoma cell lines [1]. The *NBAS* protein contains WD40 repeats (β -propeller domain; PFAM domain PF00400) in the N-terminal part of the protein and a SEC39 domain (PFAM domain PF08314), involved in the secretory pathway (Fig. 8). Syntaxins play an important role in membrane fusion of transport vesicles with the acceptor compartment [23]. Although, the exact function of *NBAS* has not been clearly elucidated, it has been shown that a peripheral protein encoded by this gene is a sub-unit of the Syntaxin 18 complex [8] involved in Golgi-to-endoplasmic reticulum retrograde transport. [6].

There are precedents for defects in the core secretory machinery resulting in human disease. Proteins exiting the ER are packaged into COPII transport vesicles which are made up of three protein complexes: SAR1, SEC23/SEC24 and SEC13/SEC31 [24] Mammalian COPII proteins have been increasingly implicated in human disease [25,26,27]. In

human phenotypes, the underlying pathogenesis appears to be due to failure to recruit the specific COPII protein, resulting in a large reduction in the packaging of specific cargo proteins *in vitro*, accompanied by swelling of the ER with untransported cargo *in vivo* in a specific tissue [28].

In the context of bone fragility, SEC24 is mainly responsible for sorting cargo molecules through interactions during COPII vesicle assembly. Mutations in *SEC24D*, a tissue specific isoform of the COPII coat is associated with a recessive form of OI [29]. Evidence to-date suggests that similarly *NBAS* variants have a cargo-selective, tissue-specific phenotype. It has been shown to be localised to the ER; required for Golgi-ER retrograde transport and its loss is associated with defects in protein glycosylation. Collagen is a large extracellular matrix protein, synthesized in ER as a rigid rod precursor (procollagen which is approximately 300 nm in length) and packaged into COPII transport vesicles (which are typically 70–100 nm in size). This packaging of collagen requires the transmembrane (TM) protein, TANGO1 and the enzymatic activity of CUL3–KLHL12. Jin et al. showed that modification of a COPII protein allows the formation of transport vesicles large enough to hold a bulky cargo like procollagen [30,31].

Evidence also suggests that *NBAS* plays a role in nonsense mediated mRNA decay (NMD) [5]. The NMD pathway is a highly-conserved surveillance mechanism that selectively degrades mRNAs harbouring premature termination codons (PTC), acting to prevent the accumulation of truncated proteins that may interfere with cellular function [32,33]. NMD also has an important role in controlling the expression of many naturally occurring transcripts [34]. *NBAS* is a bona-fide NMD factor that acts together with core NMD factors to regulate expression of a large number of endogenous RNA targets [5]. Interestingly, *NBAS*-regulated genes harbour sequence features associated with protein trafficking and ER-coupled protein modifications. Multiple targets for *NBAS* appear to have a role in regulation of bone mineralisation, osteoblast differentiation and bone development. The target with strongest up-regulation upon *NBAS*-depletion was *MGP* gene (matrix Gla protein), which acts as an inhibitor of bone formation [5]. *MGP* variants are associated with Keutel syndrome which shows abnormal cartilage calcification [35]. RNAseq studies in bone cells show that *NBAS* is expressed in osteoblasts and osteocytes of rodents and primates.

CRISPR-Cas9 technology has been used to generate stable knockout *NBAS* cell lines in human SAOS2 osteoblast cells. Further studies are underway to understand the precise mechanism of action of *NBAS* in NMD and secretion, resulting in a multi-system phenotype. Since *NBAS* has been proposed to function in the NMD pathway and Golgi-ER transport, the effect on bone fragility could be attributed to either pathway in isolation, or alternatively to a combination of both. We propose three

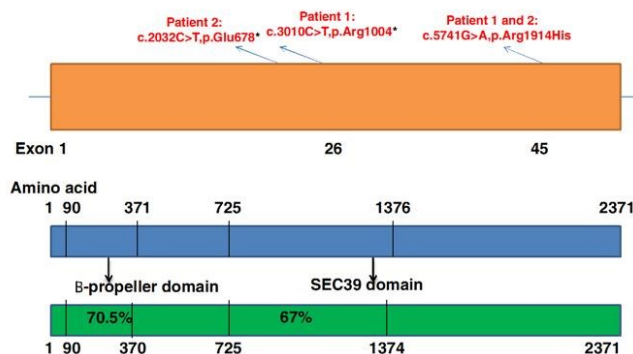


Fig. 8. Schematic representation of *NBAS* structure with known protein domains in human (blue) and zebrafish (green) protein (shaded boxes represent the regions of sequence conservation within proteins; the level of conservation is indicated in percentage). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

possible models to explain how *NBAS* variants give rise to phenotypes and human disease based on the cell studies (WB and high content microscopy) and EM findings; 1) due to a compromised Golgi-ER retrograde transport, 2) due to an NMD phenotype, 3) a dual role *i.e.* a compromised Golgi-ER transport indirectly affecting NMD.

Homozygous missense variants c.5741G>A(p.Arg1914His) in *NBAS* were reported as being associated with SOPH syndrome (short stature with optic atrophy and Pelger-Huet anomaly) [9]. Both patients reported here have features of SOPH *i.e.* short stature, granulocyte left shift with Pelger Huet cells but the facial dysmorphism is distinct as explained by the different ethnicity of our patients. Mutations in *NBAS* were reported in individuals with early-onset recurrent acute liver failure, but none of the individuals reported had a skeletal phenotype [10]. More recent literature is suggesting a multi-system phenotype with a skeletal phenotype and early-onset osteoporosis [11,12,13]. Interestingly, our patients also have abnormal liver function tests of unexplained cause, since a young age, but have never had any episodes of acute onset liver failure (ALF). In addition, our patients also have a predominant skeletal phenotype with bone fragility, multiple vertebral (Patient 1) and long bone fractures (Patient 1 and 2) needing treatment with Pamidronate. Hence, mutations in *NBAS* are likely to be a novel cause of heritable bone fragility and should be included in the targeted gene panel testing for OI that is currently offered in diagnostic genetic testing, in order to clarify diagnosis, inform prognosis and discussions around recurrence risk (up to 25%).

6. Summary

Bone fragility, severe enough to need therapy in childhood, has not been previously reported as a feature associated with variants in *NBAS*. Therefore, we hypothesise that compound heterozygous variants in *NBAS* accounts for bone fragility and is a novel cause of osteogenesis imperfecta (OI). The increasing evidence pointing to a role for *NBAS* in liver, immune and connective tissue coupled with its extreme phenotypic variability make understanding *NBAS* function important.

Acknowledgments

We thank the families for their participation in this report.

Funding

Patient 1: This research was supported by The Sheffield Children's Hospital Charity (TCHC) grant number CA15001.

Patient 2: The Deciphering Developmental Disorders (DDD) study presents independent research commissioned by the Health Innovation Challenge Fund [grant number HICF-1009-003], a parallel funding partnership between the Wellcome Trust and the Department of Health, and the Wellcome Trust Sanger Institute [grant number WT098051]. The views expressed in this publication are those of the author(s) and not necessarily those of the Wellcome Trust or the Department of Health. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and GEN/284/12 granted by the Republic of Ireland REC). The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network.

Competing interests

No competing interest to declare.

References

[1] D.K. Scott, J.R. Board, X. Lu, A.D. Pearson, R.M. Kenyon, J. Lunec, The neuroblastoma amplified gene, *NAG*: genomic structure and characterisation of the 7.3 kb transcript predominantly expressed in neuroblastoma, *Gene* 307 (2003) 1–11.

[2] W. Ki, Z. XX, L. BJ, R. Kuick, A. PF, H. Kovar, D. Thoraval, S. Motyka, J.R. Alberts, H. SM, Co-amplification of a novel gene, *NAG*, with the *N-myc* gene in neuroblastoma, *Oncogene* 18 (1) (1999) 233–238.

[3] D. Longman, R.H. Plasterk, I.L. Johnstone, J.F. Cáceres, Mechanistic insights and identification of two novel factors in the *C. elegans* NMD pathway, *Genes Dev.* 21 (9) (2007) 1075–1085.

[4] C. Anastasaki, D. Longman, A. Capper, E.E. Patton, J.F. Cáceres, *Dhx34* and *Nbas* function in the NMD pathway and are required for embryonic development in zebrafish, *Nucleic Acids Res.* 39 (9) (2011 May) 3686–3694.

[5] D. Longman, N. Hug, M. Keith, C. Anastasaki, E.E. Patton, G. Grimes, Cáceres J.F. *DHX34* and *NBAS* form part of an autoregulatory NMD circuit that regulates endogenous RNA targets in human cells, zebrafish and *Caenorhabditis elegans*, *Nucleic Acids Res.* 41 (17) (2013) 8319–8331.

[6] T. Aoki, S. Ichimura, A. Itoh, M. Kuramoto, T. Shinkawa, T. Isobe, M. Tagaya, Identification of the neuroblastoma-amplified gene product as a component of the syntaxin 18 complex implicated in Golgi-to-endoplasmic reticulum retrograde transport, *Mol. Biol. Cell* 20 (11) (2009) 2639–2649.

[7] A. Spang, Retrograde traffic from the Golgi to the endoplasmic reticulum, *Cold Spring Harb. Perspect. Biol.* 1 (5(6)) (2013 Jun) (Review).

[8] W. Hong, S. Lev, Tethering the assembly of SNARE complexes, *Trends Cell Biol.* 24 (1) (2014 Jan) 35–43.

[9] N. Maksimova, K. Hara, I. Nikolaeva, T. Chun-Feng, T. Usui, M. Takagi, Y. Nishihira, A. Miyashita, H. Fujiwara, T. Oyama, A. Nogovicina, A. Sukhomyasova, S. Potapova, R. Kuwano, H. Takahashi, M. Nishizawa, O. Onodera, Neuroblastoma amplified sequence gene is associated with a novel short stature syndrome characterised by optic nerve atrophy and Pelger-Huët anomaly, *J. Med. Genet.* 47 (8) (2010) 538–548.

[10] T.B. Haack, C. Stauffer, M.G. Köpke, B.K. Straub, S. Kölker, C. Thiel, P. Freisinger, I. Baric, P.J. McKiernan, N. Dikow, I. Harting, F. Beisse, P. Burgard, U. Kotzaeridou, J. Kühr, U. Himmert, R.W. Taylor, F. Distelmaier, J. Vockley, L. Ghaloul-Gonzalez, J. Zschocke, L.S. Kremer, E. Graf, T. Schwarzmayr, D.M. Bader, J. Gagneur, T. Wieland, C. Terrile, T.M. Strom, T. Meitinger, G.F. Hoffmann, H. Prokisch, Biallelic mutations in *NBAS* cause recurrent acute liver failure with onset in infancy, *Am. J. Hum. Genet.* 97 (1) (2015) 163–169.

[11] N. Garcia Segarra, D. Ballhausen, H. Crawford, M. Perreau, B. Campos-Xavier, K. van Spaendonck-Zwarts, C. Vermeer, M. Russo, P.-Y. Zambelli, B. Stevenson, B. Royer-Bertrand, C. Rivolta, F. Candotti, S. Unger, F. Munier, A. Superti-Furga, L. Bonaf, e, *NBAS* mutations cause a multisystem disorder involving bone, connective tissue, liver, immune system, and retina, *Am. J. Med. Genet. Part A* 167A (12) (2015) 2902–2912.

[12] J.M. Capo-Chichi, C. Mehawej, V. Delague, C. Caillaud, I. Khneisser, F.F. Hamdan, J.L. Michaud, Z. Kibar, A. Mégarbané, Neuroblastoma amplified sequence (*NBAS*) mutation in recurrent acute liver failure: confirmatory report in a sibship with very early onset, osteoporosis and developmental delay, *Eur. J. Med. Genet.* 11 (58(12)) (2015) 637–641.

[13] C. Stauffer, H. TB, K. MG, S. BK, S. Kölker, C. Thiel, P. Freisinger, I. Baric, M.K. PJ, N. Dikow, I. Harting, F. Beisse, P. Burgard, U. Kotzaeridou, D. Lenz, J. Kühr, U. Himmert, T. RW, F. Distelmaier, J. Vockley, L. Ghaloul-Gonzalez, O. JA, J. Zschocke, A. Kuster, A. Dick, D. AM, T. Wieland, C. Terrile, S. TM, T. Meitinger, H. Prokisch, H. GF, Recurrent acute liver failure due to *NBAS* deficiency: phenotypic spectrum, disease mechanisms, and therapeutic concepts, *J. Inher. Metab. Dis.* 39 (1) (2016) 3–16.

[14] W. CF, F. TW, J. WD, S. Clayton, M.R. JF, M. van Kogelenberg, K. DA, K. Ambridge, B. DM, T. Bayzina, B. AP, E. Bragin, C. EA, S. Gribble, P. Jones, N. Krishnappa, M. LE, R. Miller, M. KI, V. Parthiban, E. Prigmore, D. Rajan, A. Sifrim, S. GJ, T. AR, A. Middleton, M. Parker, C. NP, B. JC, H. ME, F.P. DR, F. HV, DDD Study, Genetic diagnosis of developmental disorders in the DDD Study: a scalable analysis of genome-wide research data, *Lancet* 385 (9975) (2015) 1305–1314.

[15] Deciphering Developmental Disorders Study, Large-scale discovery of novel genetic causes of developmental disorders, *Nature* 519 (7542) (2015 Mar 12) 223–228.

[16] A. Ramu, M.J. Noordam, R.S. Schwartz, A. Wuster, M.E. Hurler, R.A. Cartwright, D.F. Conrad, DeNovoGear: de novo indel and point mutation discovery and phasing, *Nat. Methods* 10 (10) (2013) 985–987.

[17] J. Hoyer, A.B. Ekiçi, S. Endebe, B. Popp, C. Zweier, A. Wiesener, E. Wohlleber, A. Dufke, E. Rossier, C. Petsch, M. Zweier, I. Göhring, A.M. Zink, G. Rappold, E. Schröck, D. Wiczorek, O. Riess, H. Engels, A. Rauch, A. Reis, Haploinsufficiency of *ARID1B*, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability, *Am. J. Hum. Genet.* 90 (3) (2012 Mar 9) 565–572.

[18] A. Fabre, B. Charroux, C. Martinez-Vinson, B. Roquelaura, E. Odul, E. Sayar, H. Smith, V. Colomb, N. Andre, J.P. Hugot, O. Goulet, C. Lacoste, J. Sarles, J. Royet, N. Levy, C. Badens, *SKIV2L* mutations cause syndromic diarrhea, or trichohepatoenteric syndrome, *Am. J. Hum. Genet.* 90 (4) (2012 Apr 6) 689–692.

[19] R.J. Gorlin, M.M. Cohen, R.C.M. Hennekam, Syndromes affecting bone: the osteogenesis imperfecta, Syndromes of the Head and Neck, Oxford Monographs on Medical Genetics, 2001.

[20] D.O. Silience, A. Senn, D.M. Danks, Genetic heterogeneity in osteogenesis imperfecta, *J. Med. Genet.* 16 (2) (1979) 101–116.

[21] A. Forlino, J.C. Marini, Osteogenesis imperfecta, *Lancet* 387 (10028) (2016) 1657–1671.

[22] J.C. Marini, A.R. Blissett, New genes in bone development: what's new in osteogenesis imperfecta, *J. Clin. Endocrinol. Metab.* 98 (8) (2013) 3095–3103.

[23] R. Jahn, R.H. Scheller, SNAREs—engines for membrane fusion, *Nat. Rev. Mol. Cell Biol.* 7 (9) (2006 Sep) 631–643.

[24] D. Jensen, R. Schekman, COPII-mediated vesicle formation at a glance, *J. Cell Sci.* 124 (Pt 1) (2011) 1–4.

[25] B. Jones, E.L. Jones, S.A. Bonney, H.N. Patel, A.F. Mensenkamp, S. Eichenbaum-Voline, M. Rudling, U. Myrdal, G. Annesi, S. Naik, N. Meadows, A. Quattrone, S.A. Islam, R.P.

- Naoumova, B. Angelin, R. Infante, E. Levy, C.C. Roy, P.S. Freemont, J. Scott, C.C. Shoulders, Mutations in a Sar1 GTPase of COPII vesicles are associated with lipid absorption disorders, *Nat. Genet.* 34 (1) (2003) 29–31.
- [26] S.A. Boyadjiev, J.C. Fromme, J. Ben, S.S. Chong, C. Nauta, D.J. Hur, G. Zhang, S. Hamamoto, R. Schekman, M. Ravazzola, L. Orci, W. Eyaid, Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking, *Nat. Genet.* 38 (10) (2006) 1192–1197.
- [27] B. Singleton, D. Bansal, N. Varma, R. Das, S. Naseem, U.N. Saikia, P. Malhotra, S. Varma, R.K. Marwaha, M.J. King, M. Ahmed, Homozygosity mapping reveals founder SEC23B-Y462C mutations in Indian congenital dyserythropoietic anemia type II, *Clin. Genet.* 88 (2) (2015) 195–197.
- [28] J.C. Fromme, M. Ravazzola, S. Hamamoto, M. Al-Balwi, W. Eyaid, S.A. Boyadjiev, P. Cosson, R. Schekman, L. Orci, The genetic basis of a craniofacial disease provides insight into COPII coat assembly, *Dev. Cell* 13 (5) (2007) 623–634.
- [29] L. Garbes, K. Kim, A. Rieß, H. Hoyer-Kuhn, F. Beleggia, A. Bevot, M.J. Kim, Y.H. Huh, H.S. Kweon, R. Savarirayan, D. Amor, P.M. Kakadia, T. Lindig, K.O. Kagan, J. Becker, S.A. Boyadjiev, B. Wollnik, O. Semler, S.K. Bohlander, J. Kim, C. Netzer, Mutations in SEC24D, encoding a component of the COPII machinery, cause a syndromic form of osteogenesis imperfecta, *Am. J. Hum. Genet.* 5 (96(3)) (2015) 432–439.
- [30] L. Jin, K.B. Pahuja, K.E. Wickliffe, A. Gorur, C. Baumgärtel, R. Schekman, M. Rape, Ubiquitin-dependent regulation of COPII coat size and function, *Nature* 482 (7386) (2012) 495–500.
- [31] D.J. Stephens, Cell biology: Collagen secretion explained, *Nature* 482 (7386) (2012) 474–475.
- [32] N. Hug, D. Longman, J.F. Cáceres, Mechanism and regulation of the nonsense-mediated decay pathway, *Nucleic Acids Res.* 44 (4) (2016) 1483–1495.
- [33] Y.F. Chang, J.S. Imam, M.F. Wilkinson, The nonsense-mediated decay RNA surveillance pathway, *Annu. Rev. Biochem.* 76 (2007) 51–74 (Review).
- [34] O. Isken, L.E. Maquat, The multiple lives of NMD factors: balancing roles in gene and genome regulation, *Nat. Rev. Genet.* 9 (9) (2008 Sep) 699–712 (Review).
- [35] P.B. Munroe, R.O. Olgunturk, J.P. Fryns, L. Van Maldergem, F. Ziereisen, B. Yuksel, R.M. Gardiner, E. Chung, Mutations in the gene encoding the human matrix Gla protein cause Keutel syndrome, *Nat. Genet.* 21 (1) (1999) 142–144.

H. List of Genes in Targeted Panel for Exome analysis

Source Key :

A = Skeletal dysplasia monogenic disease

B = HPO susceptibility to fractures

C = Mouse model abnormal mineralization

D = Mouse model decreased bone mineral density

E = Collagen prolyl hydroxylase family

F = TGF β signalling

G = Participates in osteoblastogenic processes

H = Genome wide association studies (GWAS)

I = Personal communication

Gene symbol	Source	OMIM/HGN C Number	Chromosome Location	Gene name
MARCH9	C	613336	12q14.1	MEMBRANE ASSOCIATED RING-CH-TYPE FINGER 9
ACAN	A	155760	15q26	AGGRECAN
ACP2	C,D	171650	11p11.2	ACID PHOSPHATASE 2, LYSOSOMAL
ACP5	A	171640	19p13.2	TARTRATE-RESISTANT ACID PHOSPHATASE (TRAP)
ACSBG2	C	614363	19p13.3	ACYL-COA SYNTHETASE, BUBBLEGUM FAMILY MEMBER 2
ADAR	D	146920	1q21.3	ADENOSINE DEAMINASE, RNA-SPECIFIC
ADGRD1	D	613639	12q24.33	ADHESION G PROTEIN-COUPLED RECEPTOR D1
ADGRF5	D	H:19030	6p12.3	ADHESION G PROTEIN-COUPLED RECEPTOR F5
AGA	B	613228	4q34.3	ASPARTYL-GLUCOSAMINIDASE
AGPAT2	B	603100	9q34.3	1-ACYLGLYCEROL-3-PHOSPHATE O-ACYLTRANSFERASE 2
AGPAT3	D	614794	21q22.3	1-ACYLGLYCEROL-3-PHOSPHATE O-ACYLTRANSFERASE 3
AGXT	B	604285	2q37.3	ALANINE-GLYOXYLATE AMINOTRANSFERASE
AIP	B	605555	11q13.2	ARYL HYDROCARBON RECEPTOR INTERACTING PROTEIN
AJAP1	D	610972	1p36.32	ADHERENS JUNCTION-ASSOCIATED PROTEIN 1

AKR1A1	D	103830	1p34.1	ALDO-KETO REDUCTASE FAMILY 1, MEMBER A1
AKT2	C,D	164731	19q13.2	V-AKT MURINE THYMOMA VIRAL ONCOGENE HOMOLOG 2
AKTIP	D	608483	16q12.2	AKT-INTERACTING PROTEIN
ALDH18A1	A	138250	10q24.1	ALDEHYDE DEHYDROGENASE 18 FAMILY, MEMBER A1
ALDH2	C,D	100650	12q24.12	ALDEHYDE DEHYDROGENASE 2 FAMILY
ALDH3A1	D	100660	17p11.2	ALDEHYDE DEHYDROGENASE, FAMILY 3, SUBFAMILY A, MEMBER 1
ALOX15	G,H	152392	17p13.2	ARACHIDONATE 15-LIPOXYGENASE
ALPL	A,D	171760	1p36.1-p34	ALKALINE PHOSPHATASE, TISSUE NON-SPECIFIC (TNSALP)
ANKH	A	605145	5p15.2-14.2	HOMOLOG OF MOUSE ANK (ANKYLOSIS) GENE
ANKRD11	C,D	611192	16q24.3	ANKYRIN REPEAT DOMAIN-CONTAINING PROTEIN 11
ANO5	A.	608662	11p14.3	ANOCTAMIN 5
ANOS1 (KAL1)	B	300836	Xp22.31	ANOSMIN 1
ANTXR2	B	608041	4q21	ANTHRAX TOXIN RECEPTOR 2
AP4E1	D	607244	15q21.2	ADAPTOR-RELATED PROTEIN COMPLEX 4, EPSILON-1 SUBUNIT
AQP3	D	600170	9p13.3	AQUAPORIN 3
ARHGAP31	A	610911	3q13.33	RHO GTPASE-ACTIVATING PROTEIN 31
ARHGEF4	D	605216	2q21.1	RHO GUANINE NUCLEOTIDE EXCHANGE FACTOR 4
ARL4D	D	600732	17q21.31	ADP-RIBOSYLATION FACTOR-LIKE 4D
ARRDC1	D	H:28633	9q34.3	ARRESTIN DOMAIN CONTAINING 1
ASXL1	D	612990	20q11.21	ADDITIONAL SEX COMBS-LIKE 1
ATF3	D	603148	1q32.2	ACTIVATING TRANSCRIPTION FACTOR 3;
ATP6V0A2	A	611716	12q24.3	ATPASE, H+ TRANSPORTING, LYSOSOMAL, V0 SUBUNIT A2
ATP7A	B	300011	Xq21.1	ATPASE, CU(2+)-TRANSPORTING, ALPHA POLYPEPTIDE
ATRX	A	300032	Xq21.1	ATR-X
AXIN1	H	603816	16p13.3	AXIS INHIBITOR 1
B3GALT6	A	615291	1p36.33	UDP-GAL:BETA-GAL BETA-1,3-GALACTOSYLTRANSFERASE POLYPEPTIDE 6

B3GAT3	A	606374	11q12.3	BETA-1,3-GLUCURONYLTRANSFERASE 3
B4GALT7	A	604327	5q35	XYLOSYLPROTEIN 4-BETA-GALACTOSYLTRANSFERASE DEFICIENCY
BACH2	D	605394	6q15	BTB AND CNC HOMOLOGY 2
BBS5	D	603650	2q31.1	BBS5 GENE
BBX	D	H:14422	3q13.12	BBX, HMG-BOX CONTAINING
BLM	A	604610	15q26.1	RECQ PROTEIN-LIKE 3
BMP1	A	112264	8p21.3	BONE MORPHOGENETIC PROTEIN 1
BMP2	D	112261	20p12.3	BONE MORPHOGENETIC PROTEIN TYPE 2
BMP4	G	112262	14q22.2	BONE MORPHOGENETIC PROTEIN 4
BMP7	G	112267	20q13.31	BONE MORPHOGENETIC PROTEIN 7
BMPR1B	A	603248	4q23	BONE MORPHOGENETIC PROTEIN RECEPTOR 1B
BOC	D	608708	3q13.2	BOC CELL ADHESION ASSOCIATED, ONCOGENE REGULATED
BRPF1	D	602410	3p25.3	BROMODOMAIN- AND PHD FINGER-CONTAINING PROTEIN
BSCL2	B	606158	11q12.3	BSCL2 GENE
BTK	A	300300	X22.1	BRUTON AGAMMAGLOBULINEMIA TYROSINE KINASE
BTRC	D	603482	10q24.32	BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN
C17ORF53	H	H:28460	17q21.31	CHROMOSOME 17 OPEN READING FRAME 53
C17ORF62	D	H:28672	17q25.3	CHROMOSOME 17 OPEN READING FRAME 62
C1ORF112	C	H:25565	1q24.2	CHROMOSOME 1 OPEN READING FRAME 112
C6ORF97 (CCDC170)	H	H:21177	6q25.1	COILED-COIL DOMAIN CONTAINING 170
CA2	B	611492	8q22	CARBONIC ANHYDRASE 2
CANT1	A	613165	17q25.3	CALCIUM-ACTIVATED NUCLEOTIDASE 1
CASR	A,D	601199	3q13.3-21	CALCIUM-SENSING RECEPTOR
CAST	D	114090	5q15	CALPASTATIN
CAV1	B	601047	7q31.2	CAVEOLIN 1

CBS	B	613381	21q22.3	CYSTATHIONINE BETA-SYNTHASE
CC2D2A	A	612013	4p15	COILED-COIL AND C2 DOMAINS-CONTAINING PROTEIN 2A
CCDC8	A	614145	19q13.32	COILED-COIL DOMAIN CONTAINING 8
CDC6	A	602627	17q21.2	CELL DIVISION CYCLE 6, S. CEREVISIAE, HOMOLOG OF
CDH3	A	114021	16q22	CADHERIN 3
CDK5RAP2	D	608201	9q33.2	CDK5 REGULATORY SUBUNIT-ASSOCIATED PROTEIN 2
CDKN1C	A	600856	11p15.4	CYCLIN-DEPENDENT KINASE INHIBITOR 1C
CDT1	A	605525	16q24.3	CHROMATIN LICENSING AND DNA REPLICATION FACTOR 1
CENPJ	D	609279	13q12.12- q12.13	CENTROMERIC PROTEIN J
CEP290	A	610142	12q21.32	CENTROSOMAL PROTEIN
CHD2	D	602119	15q26.1	CHROMODOMAIN HELICASE DNA-BINDING PROTEIN 2
CHD7	B	608892	8q12.2	CHROMODOMAIN HELICASE DNA-BINDING PROTEIN 7
CHRNA1	B	100690	2q31.1	CHOLINERGIC RECEPTOR, NICOTINIC, ALPHA POLYPEPTIDE 1
CHRND	B	100720	2q37.1	CHOLINERGIC RECEPTOR, NICOTINIC, DELTA POLYPEPTIDE
CHRNG	B	100730	2q37.1	CHOLINERGIC RECEPTOR, NICOTINIC, GAMMA POLYPEPTIDE
CHST14	A	608429	15q14	CARBOHYDRATE SULFOTRANSFERASE 14; ERMATAN 4-SULFOTRANSFERASE
CHST3	A	603799	10q22.1	CARBOHYDRATE SULFOTRANSFERASE 3; CHONDROITIN 6-SULFOTRANSFERASE
CHSY1	A	608183	15q26.3	CHONDROITIN SULFATE SYNTHASE 1
CIDEC	D	612120	3p25.3	CELL DEATH-INDUCING DFFA-LIKE EFFECTOR C
CISD2	C	611507	4q24	CDGSH IRON SULFUR DOMAIN PROTEIN 2
CLCN5	A	300008	Xp11.22	CHLORIDE CHANNEL 5
CLCN7	A,H	602727	16p13	CHLORIDE CHANNEL 7
CLDN14	H	605608	21q22.13	CLAUDIN 14
CLK1	C	601951	2q33.1	CDC-LIKE KINASE 1

CLSTN1	D	611321	1p36.22	CALSYNTENIN 1
CLSTN3	D	611324	12p13.31	CALSYNTENIN 3
CLVS1	D	611292	8q12.2- q12.3	CLAVESIN 1
COL10A1	A	120110	6q21–22.3	COLLAGEN 10 ALPHA-1 CHAIN
COL11A1	A	120280	1p21	TYPE 11 COLLAGEN ALPHA-1 CHAIN
COL11A2	A	120290	6p21.3	TYPE 11 COLLAGEN ALPHA-2 CHAIN
COL12A1	A	120320	6q13-q14	COLLAGEN, TYPE XII, ALPHA-1
COL14A1	D	120324	8q24.12	COLLAGEN TYPE XIV ALPHA 1 CHAIN
COL1A1	A	120150	17q21–22	COLLAGEN 1, ALPHA-1 CHAIN
COL1A2	A	120160	7q21.3	COLLAGEN 1, ALPHA-2 CHAIN
COL2A1	A	120140	12q13.1	TYPE 2 COLLAGEN
COL5A1	A	120215	9q34.3	COLLAGEN 5 ALPHA-1 CHAIN
COL9A1	A	120210	6q13	COLLAGEN 9 ALPHA-1 CHAIN
COL9A2	A	120260	1p32.2–33	COLLAGEN 9 ALPHA-2 CHAIN
COL9A3	A	120270	20q13.3	COLLAGEN 9 ALPHA-3 CHAIN
COMP	A	600310	19p13.1	COMP
CREB3L1	D	616215	11p11.2	CAMP RESPONSE ELEMENT-BINDING PROTEIN 3-LIKE 1
CREBBP	A	600140	16p13.3	CREB-BINDING PROTEIN
CRTAP	A	605497	3p22.3	CARTILAGE-ASSOCIATED PROTEIN
CSF1	D	120420	1p13.3	COLONY-STIMULATING FACTOR, MACROPHAGE-SPECIFIC
CSF1R	D	164770	5q32	COLONY-STIMULATING FACTOR 1 RECEPTOR
CSRP2BP(KAT14)	D	617501	20p11.23	LYSINE ACETYLTRANSFERASE 14
CTC1	B	613129	17p13.1	CONSERVED TELOMERE MAINTENANCE COMPONENT 1
CTNNB1	C,H	116806	3p22.1	CADHERIN-ASSOCIATED PROTEIN, BETA
CTNND2	B	604275	5p15.2	CATENIN, DELTA-2

CTSK	A,D	601105	1q21	CATHEPSIN K
CUL7	A,D	609577	6p21.1	CULLIN 7
CYB561	D	600019	19q23.3	CYTOCHROME B561
CYLD	H	605018	16q12.1	CYLD GENE
CYP26B1	A	605207	2p13.2	CYTOCHROME P450, SUBFAMILY XXVIB, POLYPEPTIDE 1
CYP27B1	B,D	609506	12q14.1	CYTOCHROME P450, SUBFAMILY XXVIIB, POLYPEPTIDE 1
CYP2R1	B	608713	11p15.2	CYTOCHROME P450, SUBFAMILY IIR, POLYPEPTIDE 1
DAGLA	D	614015	11q12.2	DIACYLGLYCEROL LIPASE, ALPHA
DCDC5	H	612321	11p14.1	DOUBLECORTIN DOMAIN-CONTAINING PROTEIN 5
DDR2	A	191311	1q23	DISCOIDIN DOMAIN RECEPTOR FAMILY, MEMBER 2
DDX52	C	612500	17q12	DEAD BOX POLYPEPTIDE 52
DDX58	H	609631	9p21.1	DEAD BOX POLYPEPTIDE 58
DEF6	D	610094	6p21.1	DEF6, MOUSE, HOMOLOG OF
DHCR24	A	606418	1p33-31.1	3-BETA-HYDROXYSTEROL DELTA-24-REDUCTASE
DHCR7	A	602858	11q13.4	7-DEHYDROCHOLESTEROL REDUCTASE
DHX40	C	607570	17q23.1	DEAH BOX POLYPEPTIDE 40
DKC1	B	300126	Xq28	DYSKERIN
DKK1	H	605189	10q21.1	DICKKOPF, XENOPUS, HOMOLOG OF, 1
DLG4	D	602887	17p13.1	DISCS LARGE, DROSOPHILA, HOMOLOG OF, 4
DLL3	A	602768	19q13	DELTA-LIKE 3
DLX3	A	600525	17q21	DISTAL-LESS HOMEBOX 3
DMP1	A	600980	4q21	DENTIN MATRIX ACIDIC PHOSPHOPROTEIN 1
DMXL2	D	612186	15q21.2	DMX-LIKE 2
DNAJB3	D	H:32397	2q37.1	DNAJ HEAT SHOCK PROTEIN FAMILY (HSP40) MEMBER B3
DNASE1L2	D	602622	16p13.3	DEOXYRIBONUCLEASE I-LIKE 2
DNASE2B	D	608057	1p31.1-	DEOXYRIBONUCLEASE II BETA

p22.3				
DSPP	A	125485	4q22.1	DENTIN SIALOPHOSPHOPROTEIN
DUSP6	B	602748	12121.33	DUAL-SPECIFICITY PHOSPHATASE 6
DYM	A	607461	18q12–21.1	DYMECLIN
DYNC2H1	A	603297	11q22.3	DYNEIN, CYTOPLASMIC 2, HEAVY CHAIN 1
EBP	A	300205	Xp11.23	EMOPAMIL-BINDING PROTEIN
EFEMP2	B	604633	11q13.1	EGF-CONTAINING FIBULIN-LIKE EXTRACELLULAR MATRIX PROTEIN 2
EFNB1	A	300035	Xq13.1	EPHRIN B1
EFTUD2	A	603892	17q21.31	ELONGATION FACTOR TU GTP-BINDING DOMAIN-CONTAINING 2
EIF2AK3	A	604032	2p12	TRANSLATION INITIATION FACTOR 2-ALPHA KINASE-3
ELMO1	D	606420	7p14.2- p14.1	ENGULFMENT AND CELL MOTILITY GENE 1
ENPP1	A	173335	6q23	ECTONUCLEOTIDE PYROPHOSPHATASE/PHOSPHODIESTERASE 1
ENTPD6	D	603160	20p11.21	ECTONUCLEOSIDE TRIPHOSPHATE DIPHOSPHOHYDROLASE 6
EOMES	D	604615	3p24.1	EOMESODERMIN, XENOPUS, HOMOLOG OF
EP300	A	602700	22q13	E1A-BINDING PROTEIN, 300-KDA
EPC1	C	610999	10p11.2	ENHANCER OF POLYCOMB, DROSOPHILA, HOMOLOG OF, 1
ERCC6	A	609413	10q11.23	EXCISION REPAIR CROSS-COMPLEMENTING, GROUP 6
ERCC8	A	609412	5q12.1	EXCISION REPAIR CROSS-COMPLEMENTING, GROUP 8
ERLIN2	D	611605	8p11.23	ENDOPLASMIC RETICULUM LIPID RAFT-ASSOCIATED PROTEIN 2
ESCO2	A	609353	8p21.1	HOMOLOG OF ESTABLISHMENT OF COHESION—2
ESR1	H	133430	6q25.1- q25.2	ESTROGEN RECEPTOR 1
EVC	A	604831	4p16	ELLIS–VAN CREVELD 1 PROTEIN
EVC2	A	607261	4p16	EVC GENE 2
EXT1	A	608177	8q23–24.1	EXOSTOSIN-1

EXT2	A	608210	11p12–11	EXOSTOSIN-2
FADD	D	602457	11q13.3	FAS-ASSOCIATED VIA DEATH DOMAIN
FAM111A	A	615292	11q12.1	FAMILY WITH SEQUENCE SIMILARITY 111, MEMBER A
FAM134B	B	613114	5p15.1	FAMILY WITH SEQUENCE SIMILARITY 134, MEMBER B
FAM134C	D	616498	17q21.2	FAMILY WITH SEQUENCE SIMILARITY 134
FAM20C	A,D	611061	7p22	FAMILY WITH SEQUENCE SIMILARITY 20, MEMBER C
FAM210A	H	H:28346	18p11.21	FAMILY WITH SEQUENCE SIMILARITY 210 MEMBER A
FAM73B	D	616774	9q34.11	FAMILY WITH SEQUENCE SIMILARITY 73, MEMBER B
FAM83G	D	615886	17p11.2	FAMILY WITH SEQUENCE SIMILARITY 83, MEMBER G
FARSA	D	602918	19p13.2	PHENYLALANINE-TRNA SYNTHETASE, ALPHA SUBUNIT
FBLN1	A	135820	22q13.3	FIBULIN 1
FBLN5	B	604580	14q32.12	FIBULIN 5
FBN1	A	134797	15q21.1	FIBRILLIN 1
FBXO11	C,D	607871	2p16.3	F-BOX ONLY PROTEIN 11
FBXW4	A	608071	10q24	DACTYLIN
FERMT3 (KIND3)	A	607901	11q12	FERMITIN 3 (KINDLIN 3)
FEZF1	B	613301	7q31.32	FEZ FAMILY ZINC FINGER PROTEIN 1
FGD1	A	300546	Xp11.22	FYVE, RHOGEF, AND PH DOMAIN-CONTAINING PROTEIN 1
FGF10	A	602115	5p13–p12	FIBROBLAST GROWTH FACTOR 10
FGF17	B	603725	8p21.3	FIBROBLAST GROWTH FACTOR 17
FGF23	A,D	605380	12p13.3	FIBROBLAST GROWTH FACTOR 23
FGF7	D	148180	15q21.2	FIBROBLAST GROWTH FACTOR 7
FGF8	B	600483	10q24.32	FIBROBLAST GROWTH FACTOR 8
FGF9	A	600921	13q11–q12	FIBROBLAST GROWTH FACTOR 9
FGFR1	A,D	136350	8p11.23- p11.22	FIBROBLAST GROWTH FACTOR RECEPTOR 1

FGFR2	A	176943	10q26.12	FIBROBLAST GROWTH FACTOR RECEPTOR 2	
FGFR3	A	134934	4p16.3	FIBROBLAST GROWTH FACTOR RECEPTOR 3	
FGFR4	D	134935	5q35.2	FIBROBLAST GROWTH FACTOR RECEPTOR 4	
FGFRL1	H	605830	4p16.3	FIBROBLAST GROWTH FACTOR RECEPTOR-LIKE 1	
FGL1	C	605776	8p22	FIBRINOGEN-LIKE 1	
FIG4	B	609390	6q21	FIG4, S. CEREVISIAE, HOMOLOG OF	
FKBP10	A	607063	17q21.2	FK506-BINDING PROTEIN 10	
FLJ42280 (SEM1)	H	601285	7q21.3	SEM1, 26S PROTEASOME COMPLEX SUBUNIT	
FLNA	A	300017	Xq28	FILAMIN A	
FLNB	A,D	603381	3p14.3	FILAMIN B	
FLRT3	B	604808	20p12.1	FIBRONECTIN-LIKE DOMAIN-CONTAINING LEUCINE-RICH TRANSMEMBRANE PROTEIN 3	
FOS	B	164810	14q24.3	V-FOS FBJ MURINE OSTEOSARCOMA VIRAL ONCOGENE HOMOLOG	
FOSB	G	164772	19q13.32	V-FOS FBJ MURINE OSTEOSARCOMA VIRAL ONCOGENE HOMOLOG B	
FOSL1	G	136515	11q13.1	FOS-LIKE ANTIGEN 1	
FOSL2	G	601575	2p23.2	FOS-LIKE ANTIGEN 2	
FOXC1	A	601090	6p25.3	FORKHEAD BOX C1	
FOXC2	H	602402	16q24.1	FORKHEAD BOX C2	
FOXL1	H	603252	16q24.1	FORKHEAD BOX L1	
FOXO3	D	602681	6q21	FORKHEAD BOX O3A	
FRS2	D	607743	12q15	FIBROBLAST GROWTH FACTOR RECEPTOR SUBSTRATE 2	
FSHR	D	136435	2p16.3	FOLLICLE-STIMULATING HORMONE RECEPTOR	
FUBP3	H	603536	9p34.11- q34.12	FAR UPSTREAM ELEMENT-BINDING PROTEIN 3	
FZD4	B	604579	11q14.2	FRIZZLED, DROSOPHILA, HOMOLOG OF, 4	
GALNT3	A	601756	2q24-q31	UDP-N-ACETYL-ALPHA-D-GALACTOSAMINE:POLYPEPTIDE	N-

				ACETYL GALACTOSAMINYL TRANSFERASE 3
GATA1	B	305371	Xp11.23	GATA-BINDING PROTEIN 1
GBA	A	606463	1q22	GLUCOSIDASE, BETA, ACID
GDF5	A	601146	20q11.2	GROWTH AND DIFFERENTIATION FACTOR 5
GH1	A	139250	17q23.3	GROWTH HORMONE 1
GHR	A	600946	5p13-p12	GROWTH HORMONE RECEPTOR
GHRHR	A,D	139191	7p14.3	GROWTH HORMONE-RELEASING HORMONE RECEPTOR
GHSR	A	601898	3q26.31	GROWTH HORMONE SECRETAGOGUE RECEPTOR
GK	B	300474	Xp21.2	GLYCEROL KINASE
GLE1	B	603371	9q34.11	GLE1, RNA EXPORT MEDIATOR
GLI2	A	165230	2q14.2	GLI-KRUPPEL FAMILY MEMBER 2
GLI3	A	165240	7p13	GLI-KRUPPEL FAMILY MEMBER 3
GLRB	D	138492	4q32.1	GLYCINE RECEPTOR, BETA SUBUNIT
GNA11	C	139313	19p13.3	GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-11
GNAS	A,C	139320	20q13	GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-STIMULATING ACTIVITY SUBUNIT 1
GNPTAB	A	607840	4q21-23	N-ACETYLGLUCOSAMINE 1-PHOSPHOTRANSFERASE, ALPHA/BETA SUBUNITS
GORAB	A	607983	1q24.2	SCYL1-BINDING PROTEIN 1
GPATCH1	H	H:24658	19q13.12	G PATCH DOMAIN CONTAINING 1
GPC6	A	604404	13q31-q32	GLYPICAN 6
GPR116 (ADGRF5)	D	H:19030	6p12.3	ADHESION G PROTEIN-COUPLED RECEPTOR F5
GPR133	H	613639	12q24.33	ADHESION G PROTEIN-COUPLED RECEPTOR D1
GPR152	D	H:23622	11q13.2	G PROTEIN-COUPLED RECEPTOR 152
GPR177	H	611514	1p31.3	WNTLESS, DROSOPHILA, HOMOLOG OF
GPX4	A	138322	19p13.3	GLUTATHIONE PEROXIDASE 4
GREM1	A	603054	15q13-q14	GREMLIN 1, FORMIN 1

GRIN2D	D	602717	19q13.33	GLUTAMATE RECEPTOR, IONOTROPIC, N-METHYL-D-ASPARTATE, SUBUNIT 2D
GSC	A	138890	14q32.13	GOOSECOID HOMEobox
GSE1	C	616886	16q24.1	GSE1 COILED-COIL PROTEIN
GULOP	D	240400	8p21.1	HYPOASCORBEMIA
HACE1	D	610876	6q16.3	HECT DOMAIN- AND ANKYRIN REPEAT-CONTAINING E3 UBIQUITIN PROTEIN LIGASE 1
HAX1	D	605998	1q21.3	HCLS1-ASSOCIATED PROTEIN X1
HBS1L	D	612450	6q23.3	HBS1-LIKE PROTEIN
HDAC4	A	605314	2q37.3	HISTONE DEACETYLASE 4
HDAC5	H	605315	17q21.3	HISTONE DEACETYLASE 5
HDAC8	D	300269	Xq13.1	HISTONE DEACETYLASE 8
HECTD2	D	H:26736	10q23.32	HECT DOMAIN CONTAINING 2
HERC4	D	609248	10q22	HECT DOMAIN AND RCC1-LIKE DOMAIN 4
HESX1	A	601802	3p14.3	HOMEobox GENE EXPRESSED IN ES CELLS
HOOK3	D	607825	8p11.21	HOOK, DROSOPHILA, HOMOLOG OF, 3
HOXA11	A	142958	7p15-14.2	HOMEobox A11
HOXD13	A	142989	2q31	HOMEobox D13
HPGD	A	601688	4q34-35	15-ALPHA-HYDROXYPROSTAGLANDIN DEHYDROGENASE
HRAS	B	190020	11p15.5	V-HA-RAS HARVEY RAT SARCOMA VIRAL ONCOGENE HOMOLOG
HS6ST1	B	604846	2q14.3	HEPARAN SULFATE 6-O-SULFOTRANSFERASE 1
HSPG2	A	142461	1q36-34	PERLECAN
HTR3B	D	604654	11q23.2	5-HYDROXYTRYPTAMINE RECEPTOR 3B
IBSP	H	147563	4q22.1	INTEGRIN-BINDING SIALOPROTEIN
ICK	A	612325	6p12.3	INTESTINAL CELL KINASE
ID1	D	600349	20q11.21	INHIBITOR OF DIFFERENTIATION 1
IDH1	B	147700	2q34	ISOCITRATE DEHYDROGENASE 1

IDH2	B	147650	15q26.1	ISOCITRATE DEHYDROGENASE 2
IFIH1	H	606951	2q24.2	INTERFERON-INDUCED HELICASE C DOMAIN-CONTAINING PROTEIN 1
IFITM5	A	614757	11p15.5	INTERFERON-INDUCED TRANSMEMBRANE PROTEIN 5
IFT122	A	606045	3q21	INTRAFLAGELLAR TRANSPORT 122 (CHLAMYDOMONAS, HOMOLOG OF)
IFT140	A	614620	16p13.3	INTRAFLAGELLAR TRANSPORT 140, CHLAMYDOMONAS, HOMOLOG OF
IFT172	A	607386	2p23.3	INTRAFLAGELLAR TRANSPORT 172, CHLAMYDOMONAS, HOMOLOG OF
IFT43	A	614068	14q24.3	INTRAFLAGELLAR TRANSPORT 43, CHLAMYDOMONAS, HOMOLOG OF
IFT80	A	611177	3q25.33	INTRAFLAGELLAR TRANSPORT 80, CHLAMYDOMONAS, HOMOLOG OF
IGF1	A	147440	12q23.2	INSULIN-LIKE GROWTH FACTOR I
IGF1R	A	147370	15q26.3	INSULIN-LIKE GROWTH FACTOR I RECEPTOR
IHH	A	600726	2q35	INDIAN HEDGEHOG
IKBKAP	B	603722	9q31.3	INHIBITOR OF KAPPA LIGHT POLYPEPTIDE GENE ENHANCER IN B CELLS, KINASE COMPLEX-ASSOCIATED PROTEIN
IKBKG (NEMO)	A	300248	Xq28	INHIBITOR OF KAPPA LIGHT POLYPEPTIDE GENE ENHANCER, KINASE OF
IL17RD	B	606807	3p14.3	INTERLEUKIN 17 RECEPTOR D
IL1RN	A	147679	2q14.2	INTERLEUKIN 1 RECEPTOR ANTAGONIST
INPPL1	D	600829	11q13.4	INOSITOL POLYPHOSPHATE PHOSPHATASE-LIKE 1
INSR	A	147670	19p13.2	INSULIN RECEPTOR
IQSEC3	D	612118	12p13.33	IQ MOTIF- AND SEC7 DOMAIN-CONTAINING PROTEIN 3
ITGA5	F	135620	12q13.13	INTEGRIN SUBUNIT ALPHA 5
JAG1	H	601920	20p12.2	JAGGED 1
JUN	G	165160	1p32.1	V-JUN AVIAN SARCOMA VIRUS 17 ONCOGENE HOMOLOG
KAT6B	A	605880	10q22.2	LYSINE ACETYLTRANSFERASE 6B
KCNJ16	C	605722	17q24.3	POTASSIUM CHANNEL, INWARDLY RECTIFYING, SUBFAMILY J, MEMBER 16
KCNMA1	H	600150	10q22.3	POTASSIUM CHANNEL, CALCIUM-ACTIVATED, LARGE CONDUCTANCE, SUBFAMILY M, ALPHA MEMBER 1

KDM6A	A	300128	Xp11.3	LYSINE-SPECIFIC DEMETHYLASE 6A
KIF1A	B	601255	2q37.3	KINESIN FAMILY MEMBER 1A
KIF22	A	603213	16p11.2	KINESIN FAMILY MEMBER 22
KIF3B	D	603754	20q11.21	KINESIN FAMILY MEMBER 3B
KIF7	A	611254	15q26.1	KINESIN FAMILY MEMBER 7
KISS1R	B,D	604161	19p13.3	KISS1 RECEPTOR
KIT	B	164920	4q12	V-KIT HARDY-ZUCKERMAN 4 FELINE SARCOMA VIRAL ONCOGENE HOMOLOG
KMT2D	A	602113	12q13.12	LYSINE-SPECIFIC METHYLTRANSFERASE 2D
KRAS	A	190070	12p12.1	V-KI-RAS2 KIRSTEN RAT SARCOMA VIRAL ONCOGENE HOMOLOG
L3MBTL2	D	611865	22q13.2	L3MBT-LIKE 2
LACTB2	H	H:18512	8q13.3	LACTAMASE, BETA 2
LBR	B	600024	1q42.1	LAMIN B RECEPTOR, 3-BETA-HYDROXYSTEROL DELTA (14)-REDUCTASE
LEMD3	A,B	607844	12q14	LEM DOMAIN-CONTAINING 3
LEPRE1(P3H1)	A	610339	1p34.2	LEUCINE- AND PROLINE-ENRICHED PROTEOGLYCAN 1
LFNG	A	602576	7p22	LUNATIC FRINGE
LHX3	A	600577	9q34.3	LIM HOMEBOX GENE 3
LIFR	A	151443	5p13.1	LEUKEMIA INHIBITORY FACTOR RECEPTOR
LMBR1	A	605522	7q36	PUTATIVE RECEPTOR PROTEIN
LMNA	A	150330	1q21.2	LAMIN A/C
LMX1B	A	602575	9q34.1	LIM HOMEBOX TRANSCRIPTION FACTOR 1
LPIN2	B	605519	18p11.3	LIPIN 2
LRP4	H	604270	11p11.2	LOW DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 4
LRP5	A	603506	11q13.2	LOW DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 5
LRP6	H	603507	12p13.2	LOW DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 6
LTBP1	D	150390	2p22.3	LATENT TRANSFORMING GROWTH FACTOR-BETA-BINDING PROTEIN 1
MAFB	A	608968	20q12	V-MAF MUSCULOAPONEUROTIC FIBROSARCOMA ONCOGENE FAMILY, PROTEIN

				B
MAMSTR	D	610349	19q13.33	MEF2-ACTIVATING SAP TRANSCRIPTIONAL REGULATOR
MAP3K10	D	600137	19q13.2	MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 10
MAP3K7	F	602614	6q15	MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 7
MAPKAPK2	D	602006	1q32.1	MITOGEN-ACTIVATED PROTEIN KINASE-ACTIVATED PROTEIN KINASE 2
MAPKBP1	B,D	616786	15q15.1	MITOGEN-ACTIVATED PROTEIN KINASE-BINDING PROTEIN 1
MARHC09	A	613336	12q14.1	MEMBRANE-ASSOCIATED RING-CH FINGER PROTEIN 9
MATN3 (MBL1)	H	602109	2p23-24	MATRILIN 3
MBL2	H	154545	10q21.1	LECTIN, MANNOSE-BINDING, SOLUBLE, 2
MBTPS1	G	603355	16q23.3- q24.1	MEMBRANE-BOUND TRANSCRIPTION FACTOR PROTEASE, SITE 1
MBTPS2	A	300294	Xp22.12	MEMBRANE-BOUND TRANSCRIPTION FACTOR PROTEASE, SITE 2
MCPH1	D	607117	8p23.1	MICROCEPHALIN
MDK	D	162096	11p11.2	MIDKINE
MECOM	C	165215	3q26.2	MDS1 AND EVI1 COMPLEX LOCUS
MEF2C	H	600662	5q14.3	MADS BOX TRANSCRIPTION ENHANCER FACTOR 2, POLYPEPTIDE C
MEPE	H	605912	4q22.1	MATRIX, EXTRACELLULAR, PHOSPHOGLYCOPROTEIN
MESP2	A	605195	15q26	MESODERM POSTERIOR (EXPRESSED IN) 2
MGP	A	154870	12p13.1- 12.3	MATRIX GLA PROTEIN
MIR96	C	611606	7q32.2	MICRO RNA 96
MKS1	A	609883	17q23	MKS1 GENE
MMP13	A	600108	11q22.2	MATRIX METALLOPROTEINASE 13
MMP2	A,F	120360	16q13	MATRIX METALLOPROTEINASE 2
MMP9	A,F	120361	20q13.12	MATRIX METALLOPROTEINASE 9
MOGS	B	601336	2p13.1	MANNOSYL-OLIGOSACCHARIDE GLYCOSIDASE

MPI	C	154550	15q24.1	MANNOSEPHOSPHATE ISOMERASE
MPP7	H	610973	10p12.1	MEMBRANE PROTEIN, PALMITOYLATED 7
MPV17	B	137960	2p23.3	MPV17, MOUSE, HOMOLOG OF
MRAP2	D	615410	6q14.2	MELANOCORTIN 2 RECEPTOR ACCESSORY PROTEIN 2
MRPS35	C	611995	12p11.22	MITOCHONDRIAL RIBOSOMAL PROTEIN S35
MSX2	A	123101	5q35.2	MUSCLE SEGMENT HOMEBOX 2
MTA1	C	603526	14q32.33	METASTASIS-ASSOCIATED GENE 1
MTAP	B	156540	9p21.3	METHYLTHIOADENOSINE PHOSPHORYLASE
MYCN	A	164840	2p24.3	V-MYC AVIAN MYELOCYTOMATOSIS VIRAL-RELATED ONCOGENE, NEUROBLASTOMA-DERIVED
MYO7A	C	276903	11q13.5	MYOSIN VIIA
MYOG	D	159980	1q32.1	MYOGENIN
MYSM1	C	612176	1p32.1	MYB-LIKE, SWIRM, AND MPN DOMAINS-CONTAINING PROTEIN 1
NAB2	D	602381	12q13.3	NGFIA-BINDING PROTEIN 2
NBAS	A	608025	2p24.3	NEUROBLASTOMA-AMPLIFIED SEQUENCE
NBEAL2	D	614169	3p21.31	NEUROBEACHIN-LIKE 2
NBN	A	602667	8q21.3	NIBRIN
NBR1	D	166945	17q21.31	NEIGHBOR OF BRCA1 GENE 1
NCALD	D	606722	8q22.3	NEUROCALCIN, DELTA
NDP	B	300658	Xp11.3	NORRIN
NEK1	A	604588	4q33	NIMA RELATED KINASE 1
NF1	D	162200	17q11.2	NEUROFIBROMIN
NFIX	B	164005	19p13.3	NUCLEAR FACTOR IX
NGF	B	162030	1p13.2	NERVE GROWTH FACTOR
NGFR	D	162010	17q21.33	NERVE GROWTH FACTOR RECEPTOR
NHLH2	D	162361	1p13.1	NESCIENT HELIX LOOP HELIX 2

NHP2 (NOLA2)	B	606470	5q35.3	NUCLEOLAR PROTEIN FAMILY A, MEMBER 2; NOLA2
NIPBL	A	608667	5p13.1	NIPPED-B-LIKE
NKX3-2	A	602183	4p16.1	NK3 HOMEODOMAIN 2
NLRP3	A	606416	1q44	NLR FAMILY, PYRIN DOMAIN-CONTAINING 3
NMNAT1	C	608700	1p36.22	NICOTINAMIDE NUCLEOTIDE ADENYLYLTRANSFERASE 1
NOG	A	602991	17q22	NOGGIN
NOP10	B	606471	15q14	NUCLEOLAR PROTEIN FAMILY A, MEMBER 3; NOLA3
NOTCH2	A	600275	1p12-p11	NOTCH, DROSOPHILA, HOMOLOG OF, 2
NPR2	A	108961	9p13-12	NATRIURETIC PEPTIDE RECEPTOR 2
NRAS	B	164790	1p13.2	NEUROBLASTOMA RAS VIRAL ONCOGENE HOMOLOG
NSDHL	A	300275	Xp11	NAD(P)H STEROID DEHYDROGENASE-LIKE PROTEIN
NSMF	B	608137	9q34.3	NMDA RECEPTOR SYNAPTONUCLEAR SIGNALING AND NEURONAL MIGRATION FACTOR
NTRK1	A	191315	1q23.1	NEUROTROPHIC TYROSINE KINASE, RECEPTOR, TYPE 1
NXN	D	612895	17p13.3	NUCLEOREDOXIN
OBSL1	A	610991	2q35	OBSCURIN-LIKE 1
OCRL	B	300535	Xq26.1	OCRL GENE
OFD1	A	300170	Xp22.2	CHR. X OPEN READING FRAME 5
ORC1	A	601902	1p32.3	ORIGIN RECOGNITION COMPLEX, SUBUNIT 1, S. CEREVISIAE, HOMOLOG OF
ORC4	A	603056	2q23.1	ORIGIN RECOGNITION COMPLEX, SUBUNIT 4, S. CEREVISIAE, HOMOLOG OF
ORC6	A	607213	16q11.2	ORIGIN RECOGNITION COMPLEX, SUBUNIT 6, S. CEREVISIAE, HOMOLOG OF
OSTM1	A	607649	6q21	GRAY LETHAL/OSTEOPETROSIS ASSOCIATED TRANSMEMBRANE PROTEIN
P3H1/LEPRE1	A	610339	1p34.2	PROLYL 3-HYDROXYLASE 1
P4HA1	E	176710	10q22.1	PROCOLLAGEN-PROLINE, 2-OXOGLUTARATE-4-DIOXYGENASE, ALPHA SUBUNIT, ISOFORM 1
P4HA2	E	600608	5q31.1	PROCOLLAGEN-PROLINE, 2-OXOGLUTARATE-4-DIOXYGENASE, ALPHA SUBUNIT,

				ISOFORM 2
P4HA3	E	608987	11q13.4	PROCOLLAGEN-PROLINE, 2-OXOGLUTARATE-4-DIOXYGENASE, ALPHA SUBUNIT, ISOFORM 3
P4HB	A	176790	17q25.3	PROCOLLAGEN-PROLINE, 2-OXOGLUTARATE-4-DIOXYGENASE, BETA SUBUNIT
PAK1	D	602590	11q13.5- q14.1	P21 PROTEIN-ACTIVATED KINASE 1
PAPSS2	A	603005	10q23–q24	PAPS-SYNTHETASE 2
PARL	D	607858	3q27.1	PRESENILIN-ASSOCIATED RHOMBOID-LIKE PROTEIN
PARN	B	604212	16p13.12	POLYADENYLATE-SPECIFIC RIBONUCLEASE
PARVB	C	608121	22q13.31	PARVIN, BETA
PAX3	A	606597	2q36.1	PAIRED BOX GENE 3
PCNT	A	605925	21q22.3	PERICENTRIN
PDE11A	B	604961	2q31.2	PHOSPHODIESTERASE 11A
PDE8B	B	603390	5q13.3	PHOSPHODIESTERASE 8B
PER2	C	603426	2q37.3	PERIOD, DROSOPHILA, HOMOLOG OF, 2
PEX7	A	601757	6q22–24	PEROXISOMAL PTS2 RECEPTOR
PFN1	D	176610	17p13.2	PROFILIN 1
PGD	C	172200	1p36.22	6-PHOSPHOGLUCONATE DEHYDROGENASE, ERYTHROCYTE
PHEX	A,C	300550	Xp22	X-LINKED HYPOPHOSPHATEMIA MEMBRANE PROTEASE
PHOSPHO1	D	H:16815	17q21.32	PHOSPHATASE, ORPHAN 1
PIAS2	D	603567	18q12.1- q12.3	PROTEIN INHIBITOR OF ACTIVATED STAT2
PIGV	A	610274	1p36.11	PHOSPHATIDYLINOSITOL-GLYCAN BIOSYNTHESIS CLASS V PROTEIN (GPI MANNOSYLTRANSFERASE 2)
PITX1	A	602149	5q31.1	PAIRED-LIKE HOMEODOMAIN TRANSCRIPTION FACTOR 1 (PITUITARY HOMEobox 1)
PITX2	A	601542	4q25	PAIRED-LIKE HOMEODOMAIN TRANSCRIPTION FACTOR 2

PITX3	D	602669	10q24.32	PAIRED-LIKE HOMEODOMAIN TRANSCRIPTION FACTOR 3
PJA2	D	H:17481	5q21.3	PRAJA 2
PLCG2	C	600220	16q23.3	PHOSPHOLIPASE C, GAMMA-2
PLD5	D	H:17481	1q43	PHOSPHOLIPASE D FAMILY, MEMBER 5
PLEKHM1	A	611466	17q21.3	PLECKSTRIN HOMOLOGY DOMAIN-CONTAINING PROTEIN, FAMILY M, MEMBER 1
PLOD1	D	153454	1p36.22	PROCOLLAGEN-LYSINE, 2-OXOGLUTARATE 5-DIOXYGENASE
PLOD2	A	601865	3q24	PROCOLLAGEN-LYSINE, 2-OXOGLUTARATE 5-DIOXYGENASE 2
PLOD3	B	603066	7q22.1	PROCOLLAGEN-LYSINE, 2-OXOGLUTARATE 5-DIOXYGENASE 3
PLS3	A	300161	Xq23	PLASTIN 3
POC1A	A	614783	3p21.2	POC1 CENTRIOLAR PROTEIN, CHLAMYDOMONAS, HOMOLOG OF, A
POLR1C	A	610060	6p21.1	POLYMERASE (RNA) I POLYPEPTIDE C
POR	A	124015	7q11.23	CYTOCHROME P450 OXIDOREDUCTASE
POU1F1	A	173110	3p11.2	POU DOMAIN, CLASS 1, TRANSCRIPTION FACTOR 1
PPARG	B	601487	3p25.2	PEROXISOME PROLIFERATOR-ACTIVATED RECEPTOR-GAMMA
PPIB	A	123841	15q22.31	PEPTIDYL-PROLYL ISOMERASE B
PPM1A	D	606108	14q23.1	PROTEIN PHOSPHATASE, MAGNESIUM/MANGANESE-DEPENDENT, 1A
PPP2R3A	D	604944	3q22.2- q22.3	PROTEIN PHOSPHATASE 2, REGULATORY SUBUNIT B-DOUBLE PRIME, ALPHA
PRDM14	D	611781	8q13.3	PR DOMAIN-CONTAINING PROTEIN 14
PRDM4	D	605780	12q23.3	PR DOMAIN-CONTAINING PROTEIN 4
PRDM5	B	614161	4q27	PR DOMAIN-CONTAINING PROTEIN 5
PRKAR1A	A	188830	17q24.2	PROTEIN KINASE, CAMP-DEPENDENT, REGULATORY, TYPE I, ALPHA
PROK2	B	607002	3p13	PROKINETICIN 2
PROKR2	B	607123	20p12.3	PROKINETICIN RECEPTOR 2
PROP1	A	601538	5q35.3	PROP PAIRED-LIKE HOMEODOMAIN TRANSCRIPTION FACTOR 1
PRPSAP2	D	603762	17p11.2	PHOSPHORIBOSYLPYROPHOSPHATE SYNTHETASE-ASSOCIATED PROTEIN 2

PRSS50	C,D	607950	3p21.31	TESTIS-SPECIFIC PROTEASE 50
PSPH	D	172480	7p11.2	PHOSPHOSERINE PHOSPHATASE
PTDSS1	A	612792	8q22.1	PHOSPHATIDYLSERINE SYNTHASE 1
PTEN	B	601728	10q23.31	PHOSPHATASE AND TENSIN HOMOLOG
PTH	H	168450	11p15.3	PARATHYROID HORMONE
PTH1R	A, D	168468	3p22-21.1	PTH/PTHRP RECEPTOR 1
PTHLH	H	168470	12p11.22	PARATHYROID HORMONE-LIKE HORMONE (PARATHYROID HORMONE RELATED PEPTIDE, PTHRP)
PTPN11	A	176876	12q24	PROTEIN-TYROSINE PHOSPHATASE NONRECEPTOR-TYPE 11
PTPRQ	A	603317	12q21.31	PROTEIN-TYROSINE PHOSPHATASE, RECEPTOR-TYPE, Q
PXYLP1	D	H:26303	3q23	2-PHOSPHOXYLOSE PHOSPHATASE 1
PYCR1	A	179035	17q25.3	PYRROLINE-5-CARBOXYLATE REDUCTASE 1
RAB23	A	606144	6p11.2	RAS-ASSOCIATED PROTEIN RAB23
RAB24	D	612415	5q35.3	RAS-ASSOCIATED PROTEIN 24
RAB3IP	D	608686	12q13-q14	RAB3A-INTERACTING PROTEIN
RAF1	A	164760	3p25.2	V-RAF-1 MURINE LEUKEMIA VIRAL ONCOGENE HOMOLOG 1
RAPSN	B	601592	11p11.2	RECEPTOR-ASSOCIATED PROTEIN OF THE SYNAPSE, 43-KD
RASGRP2	A	605577	11q13	RAS GUANYL NUCLEOTIDE-RELEASING PROTEIN 2
RECQL4	A	603780	8q24.3	RECQ PROTEIN-LIKE 4
RGN	D	300212	Xp11.3	REGUCALCIN
RHBDD1	C	617515	2q36.3	RHOMBOID DOMAIN CONTAINING 1
RHBDF2	C	614404	17q25.1	RHOMBOID 5, DROSOPHILA, HOMOLOG OF, 2
RHOBTB3	D	607353	5q15	RHO-RELATED BTB DOMAIN-CONTAINING PROTEIN 3
RHOT2	D	613889	16p13.3	RAS HOMOLOG GENE FAMILY, MEMBER T2
RIC1	C	610354	9p24.1	CONNEXIN 43-INTERACTING PROTEIN
RIC8B	D	609147	12q23.3	RIC8, C. ELEGANS, HOMOLOG OF, B

RLN3	D	606855	19p13.12	RELAXIN 3
RNASE10	D	H:19275	14q11.1	RIBONUCLEASE A FAMILY MEMBER 10
RNF169	C	H:26961	11q13.4	RING FINGER PROTEIN 169
ROR2	A	602337	9q22	RECEPTOR TYROSINE KINASE-LIKE ORPHAN RECEPTOR 2
RPGRIP1L	A	610937	16q12.1	RPGRIP1-LIKE
RPS6KA3	A	300075	Xp22.12	RIBOSOMAL PROTEIN S6 KINASE, 90-KD, 3
RPS6KA5	H	603607	14q32.11	RIBOSOMAL PROTEIN S6 KINASE, 90-KD, 5
RSPO3	H	610574	6q22.33	R-SPONDIN 3
RTEL1	B	608833	20q13.33	REGULATOR OF TELOMERE ELONGATION HELICASE 1
RUNX2	A, D,H	600211	6p21	RUNT RELATED TRANSCRIPTION FACTOR 2
RXRG	C	180247	1q23.3	RETINOID X RECEPTOR, GAMMA
SALL1	A	602218	16q12.1	SAL-LIKE 1
SALL4	A	607343	20q13	SAL-LIKE 4
SBDS	A	607444	7q11	SBDS PROTEIN
SC5D	B	602286	11q23.3- q24.1	STEROL C5-DESATURASE-LIKE
SCARB2	B	602257	4q21.1	SCAVENGER RECEPTOR CLASS B, MEMBER 2
SCN10A	B	604427	3p22.2	SODIUM CHANNEL, VOLTAGE-GATED, TYPE X, ALPHA SUBUNIT
SCN11A	B	604385	3p22.2	SODIUM CHANNEL, VOLTAGE-GATED, TYPE XI, ALPHA SUBUNIT
SCN3B	D	608214	11q24.1	SODIUM CHANNEL, VOLTAGE-GATED, TYPE III, BETA SUBUNIT
SCN9A	B	603415	2q24.3	SODIUM CHANNEL, VOLTAGE-GATED, TYPE IX, ALPHA SUBUNIT
SDHC	D	602413	1q23.3	SUCCINATE DEHYDROGENASE COMPLEX, SUBUNIT C, INTEGRAL MEMBRANE PROTEIN, 15-KD
SEC24D	A	607186	4q24	SEC24-RELATED GENE FAMILY, MEMBER D
SELE	I	131210	1q24.2	SELECTIN E
SEMA3A	B	603961	7q21.11	SEMAPHORIN 3A

SEMA3F	D	601124	3p21.31	SEMAPHORIN 3F
SEMA5A	B	609297	5p15.31	SEMAPHORIN 5A
SERPINF1	A	172860	17p13.3	SERPIN PEPTIDASE INHIBITOR, CLADE F, MEMBER 1
SERPINH1	A	600943	11q13.5	SERPIN PEPTIDASE INHIBITOR, CLADE H, MEMBER 1
SETD2	D	612778	3p21.31	SET DOMAIN-CONTAINING PROTEIN 2
SETDB1	D	604396	1q21.3	SET DOMAIN PROTEIN, BIFURCATED, 1
SFRP1	C	604156	8p11.21	SECRETED FRIZZLED-RELATED PROTEIN 1
SFRP2	C	604157	4q31.3	SECRETED FRIZZLED-RELATED PROTEIN 2
SH3BP2	A	602104	4p16	SH3 DOMAIN-BINDING PROTEIN 2
SH3GL2	D	604465	9p22.2	SH3 DOMAIN, GRB2-LIKE, 2
SH3PXD2B	A	613293	5q35.1	TKS4
SHH	D	600725	7q36	SONIC HEDGEHOG
SHOX	A	312865	Xp22.33	SHORT STATURE—HOMEODOMAIN GENE
SHOX2	A	602504	3q25.32	SHORT STATURE HOMEODOMAIN 2
SLC20A2	D	158378	8p11.21	SOLUTE CARRIER FAMILY 20 (PHOSPHATE TRANSPORTER), MEMBER 2
SLC25A12	A	603667	2q31.1	SOLUTE CARRIER FAMILY 25 (MITOCHONDRIAL CARRIER, ARALAR), MEMBER 12
SLC25A13	H	603859	7q21.3	SOLUTE CARRIER FAMILY 25 (CITRIN), MEMBER 13
SLC25A21	D	607574	14q13.3	SOLUTE CARRIER FAMILY 25 (MITOCHONDRIAL OXODICARBOXYLATE CARRIER), MEMBER 21
SLC26A2	A	606718	5q32–q33	SOLUTE CARRIER FAMILY 26 (SULFATE TRANSPORTER), MEMBER 2
SLC29A1	D	602193	6p21.1	SOLUTE CARRIER FAMILY 29 (NUCLEOSIDE TRANSPORTER), MEMBER 1
SLC29A3	B	612373	10q22.1	SOLUTE CARRIER FAMILY 29 (NUCLEOSIDE TRANSPORTER), MEMBER 3
SLC34A1	B	182309	5q35.3	SOLUTE CARRIER FAMILY 34 (TYPE II SODIUM/PHOSPHATE COTRANSPORTER), MEMBER 1
SLC34A3	A	609826	9q34	SODIUM-PHOSPHATE COTRANSPORTER
SLC35D1	A	610804	1p31.3	SOLUTE CARRIER FAMILY 35 MEMBER D1; UDP-GLUCURONIC ACID/ UDP-N-

ACETYL GALACTOSAMINE DUAL TRANSPORTER				
SLC38A10	D	616525	17q25.3	SOLUTE CARRIER FAMILY 38 (AMINO ACID TRANSPORTER), MEMBER 10
SLC39A13	A	608735	11p11.2	SOLUTE CARRIER FAMILY 39 (ZINC TRANSPORTER), MEMBER 13
SLC40A1	D	604653	2q32.2	SOLUTE CARRIER FAMILY 40 (IRON-REGULATED TRANSPORTER), MEMBER 1
SLC4A1	B	109270	17q21.31	SOLUTE CARRIER FAMILY 4 (ANION EXCHANGER), MEMBER 1
SLC7A1	A	104615	13q12.3	SOLUTE CARRIER FAMILY 7 (CATIONIC AMINO ACID TRANSPORTER, Y+ SYSTEM), MEMBER 1
SLC9A3R1	B	604990	17q25.1	SOLUTE CARRIER FAMILY 9, MEMBER 3, REGULATOR 1
SLC9A4	D	600531	2q12.1	SOLUTE CARRIER FAMILY 9, MEMBER 4
SMAD1	F	601595	4q31.21	MOTHERS AGAINST DECAPENTAPLEGIC, DROSOPHILA, HOMOLOG OF, 1
SMAD4	F	600993	18q21.2	MOTHERS AGAINST DECAPENTAPLEGIC, DROSOPHILA, HOMOLOG OF, 4
SMAD7	F	602932	18q21.1	MOTHERS AGAINST DECAPENTAPLEGIC, DROSOPHILA, HOMOLOG OF, 7
SMARCAL1	A	606622	2q34–q36	SWI/SNF-RELATED REGULATOR OF CHROMATIN SUBFAMILY A-LIKE PROTEIN 1
SMC1A	A	300040	Xp11.22	STRUCTURAL MAINTENANCE OF CHROMOSOMES 1A
SMC3	A,D	606062	10q25.2	STRUCTURAL MAINTENANCE OF CHROMOSOMES 3
SMOC1	H	608488	14q24.2	SPARC-RELATED MODULAR CALCIUM-BINDING 1
SMS	B,D	607642	17p11.2	RETINOIC ACID-INDUCED GENE 1
SNX10	B	614780	7p15.2	SORTING NEXIN 10
SOS1	A	182530	2p22.1	SON OF SEVENLESS, DROSOPHILA, HOMOLOG 1
SOST	H	605740	17q21.31	SCLEROSTIN
SOX10	B	602229	22q13.1	SRY-BOX 10
SOX2	A	184429	3q26.33	SRY-BOX 2
SOX3	A	313430	Xq27.1	SRY-BOX 3
SOX6	H	607257	11p15.2	SRY-BOX 6
SOX9	A	608160	17q24.3	SRY-BOX 9
SP7	A	606633	12q13.13	TRANSCRIPTION FACTOR SP7

SPARC	D	182120	5q33.1	SECRETED PROTEIN, ACIDIC, CYSTEINE-RICH
SPNS2	D	612584	17p13.2	SPINSTER, DROSOPHILA, HOMOLOG OF, 2
SPP1 (OPN)	H	166490	4q22.1	SECRETED PHOSPHOPROTEIN 1
SPRY4	B	607984	5q31.3	SPROUTY, DROSOPHILA, HOMOLOG OF, 4
SPTBN1	H	182790	2p16.2	SPECTRIN, BETA, NONERYTHROCYTIC, 1
SQSTM1	B	601530	5q35.3	SEQUESTOSOME 1
SRCAP	A	611421	16p11.2	SNF2-RELATED CBP ACTIVATOR PROTEIN
SSBP2	D	607389	5q14.1	SINGLE-STRANDED DNA-BINDING PROTEIN 2
SST	D	182450	3q28	SOMATOSTATIN
STAG3	D	608489	7q22.1	STROMAL ANTIGEN 3
STARD3NL	H	611759	7p14.1	STARD3 N-TERMINAL-LIKE
STAT3	B	102582	17q21.2	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3
STAT5B	A	604260	17q21.2	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5B
SUCO	D	H:1240	1q24.3	SUN DOMAIN CONTAINING OSSIFICATION FACTOR
SULF1	A	610012	8q13.2- q13.3	SULFATASE 1
SYN3	D	602705	22q12.3	SYNAPSIN III
TACR3	B	162332	4q24	TACHYKININ RECEPTOR 3
TBCE	A	604934	1q42-q43	TUBULIN-SPECIFIC CHAPERONE E
TBX15	A	604127	1p13	T-BOX GENE 15
TBX3	A	601621	12q24.21	T-BOX GENE 3
TBX4	A	601719	17q21-q22	T-BOX GENE 4
TBX5	A	601620	12q24.1	T-BOX GENE 5
TBXAS1	A	274180	7q34	THROMBOXANE A SYNTHASE 1
TCIRG1	B	604592	11q13	SUBUNIT OF ATPASE PROTON PUMP
TCOF1	A	606847	5q32	TREACHER COLLINS-FRANCESCHETTI SYNDROME 1

TCTN3	A	613847	10q24.1	TECTONIC FAMILY, MEMBER 3
TDP2	A	605764	6p22.3	TYROSYL-DNA PHOSPHODIESTERASE 2
TERC	B	602322	3q26.2	TELOMERASE RNA COMPONENT
TERT	B	187270	5p15.33	TELOMERASE REVERSE TRANSCRIPTASE
TFEB	D	600744	6p21.1	T-CELL TRANSCRIPTION FACTOR EB
TGFB1	A	190180	19q13	TRANSFORMING GROWTH FACTOR BETA 1
TGFBR1	F	190181	9q22.33	TRANSFORMING GROWTH FACTOR-BETA RECEPTOR, TYPE I
TGFBR2	F	190182	3p24.1	TRANSFORMING GROWTH FACTOR-BETA RECEPTOR, TYPE II
THBS1	F	188060	15q14	THROMBOSPONDIN I
THPO	A	600044	3q27	THROMBOPOIETIN
THRB	A	190160	3p24.2	THYROID HORMONE RECEPTOR, BETA
TIMM50	D	607381	19q13.2	TRANSLOCASE OF INNER MITOCHONDRIAL MEMBRANE 50, YEAST, HOMOLOG OF
TIMM8A	B	300356	Xq22.1	TRANSLOCASE OF INNER MITOCHONDRIAL MEMBRANE 8, YEAST, HOMOLOG OF, A
TINF2	B	604319	14q12	TRF1-INTERACTING NUCLEAR FACTOR 2
TMCO1	A	614123	1q24.1	TRANSMEMBRANE AND COILED-COIL DOMAINS PROTEIN 1
TMEM132A	D	617363	11q12.2	TRANSMEMBRANE PROTEIN 132A
TMEM135	H	616360	11q14.2	TRANSMEMBRANE PROTEIN 135
TMEM136	D	H:28280	11q23.3	TRANSMEMBRANE PROTEIN 136
TMEM189	D	610994	20q13.13	TRANSMEMBRANE PROTEIN 189
TMEM216	A	613277	11q12.2	TRANSMEMBRANE PROTEIN 216
TMEM38B (TRIC-B)	A	611236	9q31.2	TRANSMEMBRANE PROTEIN 38B
TNFAIP1	D	191161	17q11.2	TUMOR NECROSIS FACTOR-ALPHA-INDUCED PROTEIN 1
TNFRSF11A	A,H	603499	18q21.33	RECEPTOR ACTIVATOR OF NF-KAPPA-B
TNFRSF11B	D,H	602643	8q24	OSTEOPROTEGERIN

TNFSF11 (RANK)	B,H	602642	13q14.11	RECEPTOR ACTIVATOR OF NF-KAPPA-B LIGAND (TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY, MEMBER 11)
TNIK	D	610005	3q26.2- q26.3	TRAF2- AND NCK-INTERACTING KINASE
TP63	A	603273	3q28	TUMOR PROTEIN P63
TPGS2	D	H:24561	18q12.2	TUBULIN POLYGLUTAMYLASE COMPLEX SUBUNIT 2
TPPP	D	608773	5p15.33	TUBULIN POLYMERIZATION-PROMOTING PROTEIN
TRAM2	D	608485	6p12.2	TRANSLOCATION-ASSOCIATED MEMBRANE PROTEIN 2
TRAPPC2	A	300202	Xp22	TRACKING PROTEIN PARTICLE COMPLEX, SUBUNIT 2
TREM2	A	605086	6p21.2	TRIGGERING RECEPTOR EXPRESSED ON MYELOID CELLS 2
TRIM37	A	605073	17q22	TRIPARTITE MOTIF-CONTAINING PROTEIN 37
TRIM45	D	609318	1p13.1	TRIPARTITE MOTIF-CONTAINING PROTEIN 45
TRIP11	A	604505	14q32.12	GOLGI-MICROTUBULE-ASSOCIATED PROTEIN, 210-KDA; MAP210
TRMT10A	A	616013	4q23	TRNA METHYLTRANSFERASE 10, S. CEREVISIAE, HOMOLOG OF, A
TRPC3	D	602345	4q27	TRANSIENT RECEPTOR POTENTIAL CATION CHANNEL, SUBFAMILY C, MEMBER 3
TRPS1	A	604386	8q24	ZINC FINGER TRANSCRIPTION FACTOR
TRPV4	A	605427	12q24.1	TRANSIENT RECEPTOR POTENTIAL CATION CHANNEL, SUBFAMILY V, MEMBER 4
TSPAN12	B	613138	7q31.31	TETRASPANIN 12
TTC21B	A	612014	2q24.3	TETRATRICOPEPTIDE REPEAT DOMAIN-CONTAINING PROTEIN 21B
TTC28	D	615098	22q12.1	TETRATRICOPEPTIDE REPEAT DOMAIN-CONTAINING PROTEIN 28
TWIST1	A	601622	7p21.1	TWIST, DROSOPHILA, HOMOLOG OF, 1
TWIST2	A	607556	2q37.3	TWIST, DROSOPHILA, HOMOLOG OF, 2
TXNIP	D	606599	1q21.1	THIOREDOXIN-INTERACTING PROTEIN
TYROBP	A	604142	19q13.1	TYRO PROTEIN TYROSINE KINASE-BINDING PROTEIN
UBE2J1	D	616175	6q15	UBIQUITIN-CONJUGATING ENZYME E2J 1
UBE3C	D	614454	7q36.3	UBIQUITIN PROTEIN LIGASE E3C

UMOD	D	191845	16p12.3	UROMODULIN
UROS	D	606938	10q26.2	UROPORPHYRINOGEN III SYNTHASE
USB1	B	613276	16q21	CHROMOSOME 16 OPEN READING FRAME 57; C16ORF57
USP8	B	603158	15q21.2	UBIQUITIN-SPECIFIC PROTEASE 8
VANGL2	C	600533	1q23.3	VANG-LIKE 2
VDR	H	601769	12q13.11	VITAMIN D RECEPTOR
VIPAS39	B	613401	14q24.3	VPS33B-INTERACTING PROTEIN, APICAL-BASOLATERAL POLARITY REGULATOR, SPE39 HOMOLOG
VPS13A	D	605978	9q21.2	VACUOLAR PROTEIN SORTING 13, YEAST, HOMOLOG OF, A
VPS33B	B	610034	12q24.31	VACUOLAR PROTEIN SORTING 33, YEAST, HOMOLOG OF, A
WBP2	D	606962	17q25.1	WW DOMAIN-BINDING PROTEIN 2
WDR11	B	606417	10q26.12	WD REPEAT-CONTAINING PROTEIN 11
WDR35	A	613602	2p24.1	WD REPEAT-CONTAINING PROTEIN 35
WDR60	A	615462	7q36.3	WD REPEAT-CONTAINING PROTEIN 60
WISP3	A	603400	6q22–23	WNT1-INDUCIBLE SIGNALING PATHWAY PROTEIN 3
WLS	H	611514	1p31.3	WNTLESS, DROSOPHILA, HOMOLOG OF
WNT1	A,B	164820	12q13.12	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 1
WNT16	H	606267	7q31.31	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 16
WNT3	H	165330	17q21.31	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 3
WNT3A	B	606359	1q42.13	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 3A
WNT4	H	603490	1p36.12	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 4
WNT5A	A	164975	3p14.3	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 5A
WNT5B	H	606361	12p13.33	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 5B
WNT7A	A	601570	3p25	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 7A
WNT9B	H	602864	17q21.32	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 9B
WRAP53	B	612661	17p13.1	WD REPEAT-CONTAINING PROTEIN ANTISENSE TO TP53

WRN/RECQL2	A	604611	8p12	RECQ PROTEIN-LIKE 2
XBP1	C	194355	22q12.1	X BOX-BINDING PROTEIN 1
XYLT1	A	608124	16p12.3	XYLOSYLTRANSFERASE 1
XYLT2	A	608125	17q21.33	XYLOSYLTRANSFERASE 2
ZBTB40	H	612106	1p36.12	ZINC FINGER- AND BTB DOMAIN-CONTAINING PROTEIN 40
ZBTB45	D	H:23715	19q13.43	ZINC FINGER AND BTB DOMAIN-CONTAINING PROTEIN 45
ZFYVE28	D	614176	4p16.3	ZINC FINGER FYVE DOMAIN-CONTAINING PROTEIN 28
ZMPSTE24	A,D	606480	1p34	ZINC METALLOPROTEINASE
ZNF408	B	616454	11p11.2	ZINC FINGER PROTEIN 408
ZNF469	B	612078	16q24.2	ZINC FINGER PROTEIN 469

I. Patient information sheet for young people

Sheffield Children's **NHS**
NHS Foundation Trust



PARTICIPANT INFORMATION SHEET

FOR YOUNG PEOPLE

Study title

Novel molecular mechanisms leading to Osteogenesis Imperfecta

(DNA studies to find new causes of Brittle Bone Disease)

Part 1 – to give you first thoughts about the project

1. Invitation paragraph

We would like you to help us with our research study. Please read this information carefully and talk to your mum, dad or carer about the study. Ask us if there is anything that is not clear or if you want to know more. Take time to decide if you want to take part. It is up to you if you want to do this. If you don't then that's fine, you'll be looked after at the hospital just the same.

2. Why are we doing this research?

We want to try and find out if mistakes in inherited factors called genes can cause fractures and a condition called Brittle Bone Disease. Genes are carried on genetic material called DNA. We inherit genes from our parents. Each gene is made up of thousands of individual genetic letters known as the gene sequence. A mistake in the gene sequence (a mutation) can disrupt the genetic message and in some instances cause bones to not receive the correct instructions for growth and development.



If you have brittle bones, they're not as strong as other peoples and you're more likely to fracture them. Some people have an inherited condition called brittle bone disease (also called osteogenesis imperfecta) where their bones are much more likely to fracture and in some severe cases may not form properly.

Although some genes have already been studied in people with fractures and brittle bone disease, it is likely that more genes exist. We hope that by studying some new genes in families who have fractures that this study will help us to identify currently unknown causes of fractures and brittle bone disease. If we can discover more about the genes that cause fractures and brittle bone disease it may lead to important advances in the management of children and their families with fractures and we may be able to develop new treatments to combat the condition.

The study is being conducted as part of a research degree.

3. Why have I been asked to take part?

You have been chosen because you and some of your family have had fractures and a previous gene test did not find the cause of these fractures.

4. Do I have to take part?

No! It is entirely up to you. If you do decide to take part:

- you will be asked to sign a form to say that you agree to take part (an assent form)

- you will be given this information sheet and a copy of your signed assent form to keep.

You are free to stop taking part at any time during the research without giving a reason. If you decide to stop, this will not affect the care you receive whilst in hospital.



5. What will happen to me if I take part?

We would like to do some tests for some new genes by using a sample of your DNA. which we hope may explain the cause of the fractures in your family. For most people we have a DNA sample stored in our laboratory. This DNA is left over from when you had your previous gene test.

If there is not enough DNA leftover for us to finish our tests we would like to take a small blood sample from you, so we can extract more DNA for analysis. We would need to take about 1 teaspoon of blood (5mls), which can be taken at your next Genetics clinic, if you have one, or at your local surgery by your GP or Practice Nurse. Taking blood samples may take up to 10 minutes, and will be very similar to having blood samples taken at a routine clinic visit. Local anaesthetic cream can be used to help numb any pain and your parent/carer can stay with you while you give your blood sample.

We may ask you questions about:

- How many brothers & sisters you have
- Do your Mum and Dad have brothers and sisters, uncles & aunts and what other family members are there?
- If any of them have fractures or other problems

Our testing of your DNA is difficult and may take some years. We will only use your DNA to look at genes which may be involved in brittle bones.

6. What will I be asked to do?

For most people, taking part in the study will only mean consenting to use of their stored DNA sample.

Some people may be asked of give a small sample of blood.

You may have to give your Genetics clinician/nurse more information on your family tree. This should take about 10 minutes.

7. Is there anything else to be worried about if I take part?

Your normal medical care will continue unchanged.

You would not need to attend any extra clinic appointments or have any extra tests to participate in this study. The study does not involve any treatment, or follow up clinic visits. You will not have to miss any meals to take part.

If we find out something that we think is important about the cause of your fractures we will talk to your mum, dad or carer and ask them if they want to come back and have you checked again at the hospital.

8. Will the study help me?

No, but the information we get might help treat young people with fractures and brittle bones better in the future.

9. What happens when the research study stops?

We will collect all the information together and decide if it is useful in telling us if we can manage fractures and brittle bone disease better in the future.

10. Contact for further information

If you would like any further information about this study you could contact:

Dr M Balasubramanian
Consultant Clinical Geneticist
Sheffield Clinical Genetics Service
Sheffield Children's NHS Foundation Trust
Western Bank
Sheffield S10 2TH
Phone: 0114 2717025



Thank you for reading so far - if you are still interested, please go to Part 2:

Part 2 - more detail – information you need to know if you still want to take part.

11. What if new information comes along?

Sometimes during research, new things are found out about your condition. If this happens, someone from the research team will tell you all about it and discuss whether you want to continue in the study. If you change your mind this will not affect any care you will receive whilst in hospital. If you decide to continue in the study you will be asked to sign an updated assent form.

12. What if I don't want to do the research anymore?

Just tell your mum, dad, carer, doctor or nurse at any time. They will not be cross with you. You will still have the same care whilst you are at hospital.

13. What if there is a problem or something goes wrong?

Tell us if there is a problem and we will try and sort it out straight away. You and your mum, dad or carer can either contact the project co-ordinator:



Dr M Balasubramanian
Consultant Clinical Geneticist
Sheffield Clinical Genetics Service
Sheffield Children's NHS Foundation Trust
Western Bank
Sheffield S10 2TH
Phone: 0114 2717025

or the hospital complaints co-ordinator:

Mrs Linda Towers

Patient Advice & Liaison Co-ordinator

Sheffield Children's NHS Foundation Trust

Tel: 0114 271 7594

14. Will anyone else know I'm doing this?

The people in our research team will know you are taking part. The Genetics doctor looking after you will also know.

Your medical notes may also be looked at by other people who work at the hospital to check that the study is being carried out correctly.

All information that is collected about you during the research will be kept strictly confidential. You will be given a number which will be used instead.

Once the study is complete all information will be kept by the research team for 30 years. This will include results of the work that we undertake whilst looking for a mistake in genes which may cause brittle bones.

15. What will happen to any samples I give?

Your DNA sample will be held in special fridges/freezers at Sheffield Diagnostic Genetics Service for 30 years. At the end of this period, your sample will be destroyed.

Your DNA sample will be used only for studying genes which may be a cause of brittle bones. This may involve looking for possible causes of brittle bones that cannot yet be specified. Only researchers involved in this work will have access to your sample.

Any significant findings during/after the study will be fed back to you by your Genetics doctor.

16. What will happen to the results of the research study?

When the study has finished we will present our findings to other doctors, and we will put the results in medical magazines and websites that doctors read. We would also like to put a brief summary on the hospital research website so that you will be able to read about our results too. This will be available at the end of the study, in <insert date>, on www.sheffieldchildrenscrf.nhs.uk. The results will also be included as part of the chief investigator's educational

qualification. They will be anonymous, which means that you will not be able to be identified from them.

17. Who is organising and funding the research?

The research is being organised by Sheffield Children's NHS Foundation Trust and paid for by the Childrens Hospital Charity Research Fund

18. Who has checked the study?

Before any research goes ahead it has to be checked by a Research Ethics Committee. This is a group of people who make sure that the research is OK to do.

It has also been checked by the Research Department at this hospital.

19. How can I find out more about research?

The Clinical Research Facility at this hospital has an **Information for families** section on its website www.sheffieldchildrenscrf.nhs.uk or you could contact the hospital Clinical Research Facility:

Mrs Wendy Swann

R&D Manager

Clinical Research Facility

Sheffield Children's NHS Foundation Trust

Tel: 0114 2717417

Thank you for taking the time to read this – please ask any questions if you need to.

J. Approved funding application for Targeted Exome Sequencing Project

FULL STUDY TITLE

Investigation of the *SERPINH1*, *SERPINF1*, *FKBP10* and *SP7* genes to determine novel mutational mechanism responsible for Osteogenesis Imperfecta in families lacking a mutation in the type I collagen genes.

SHORT STUDY TITLE

Novel molecular mechanisms leading to Osteogenesis Imperfecta.

STUDY NUMBER

DATE AND VERSION NUMBER

09th March 2011 Version 1

LAY SUMMARY

Osteogenesis imperfecta (OI) is a group of inherited disorders characterized by increased bone fragility and is often referred to as brittle bone disease. The severity of OI can range from very mild, with few fractures, to lethal. At present it can be difficult to exclude OI in children with recurrent fractures, particularly where there is a family history of bone fragility. A genetic test can help to establish the diagnosis of OI in some children.

The majority, but not all, of OI is caused by a genetic mistake (mutation) in one of the two genes which code for a protein called type I collagen, which is a major component of bone. If a mutation is found in the type I collagen genes, it will help to confirm a diagnosis of OI. However ~10% of families with OI do not have a mutation and therefore a negative test does not completely exclude OI.

Research to identify the cause of OI in individuals without mutations in the type I collagen genes has focused on individuals with lethal forms of the disease. Recently it has been found that some severely affected individuals may have mutations in new genes called *FKBP1*, *SERPINF1*, *SERPINH1* and *SP7*.

We want to find out if children and their families with moderate or mild OI also have mutations in these new genes. We will collect a group of 20 patients with moderate/mild OI without mutations in the type I collagen genes and use them to study the new genes. If a mutation can be found in any of these patients then this would influence how we investigate unexplained and recurrent fractures in infancy and childhood, improve the diagnosis and possible treatment of future OI patients and allow family members to find out if they carry the OI mutation.

GENERAL INFORMATION

Sponsor: Sheffield Children's NHS Foundation Trust	Sponsor's Representative/Contact: Dr Jim Bonham
Western Bank	Director of Research & Development
Sheffield	
S10 2TH	Phone: 0114 2717404
United Kingdom	Fax: 0114 2717417
	Email: jim.bonham@sch.nhs.uk
Chief Investigator: Name – Rebecca Pollitt	Principal Investigator Name –
Address –	Address –
Sheffield Diagnostic Genetics Department	Telephone –
Sheffield Children's NHS Foundation Trust	Fax –
Western Bank	Email -
Sheffield	
S10 2TH	
United Kingdom	
Telephone – +441142717016	
Fax +441142750629	
Email – rebecca.pollitt@sch.nhs.uk	
Co-Investigator: Name – Dr A Dalton	Co-Investigator: Name – Prof N J Bishop
Address –	Address –
Sheffield Diagnostic Genetics Department	Academic Unit of Child Health, University of Sheffield
Sheffield Children's NHS Foundation Trust	Stephenson Wing

Western Bank	Sheffield Children's Hospital
Sheffield	Western Bank
S10 2TH	Sheffield S10 2TH
United Kingdom	
Telephone – +441142717004	Telephone – Tel +441142717303
Fax +441142750629	Fax +441142755364
Email – ann.dalton@sch.nhs.uk	
	Email – N.J.Bishop@scheffield.nhs.uk
Statistician: Name –	Other: Name –
Address –	Address –
Telephone –	Telephone –
Fax –	Fax –
Email -	Email -

GLOSSARY

A list of abbreviations and definitions

Osteogenesis Imperfecta (OI) – inherited bone disease resulting in bone fragility and fractures.

Mutation – a genetic mistake responsible for an inherited disease such as OI

Collagen – the fibrous protein constituent of bone, cartilage, tendon and other connective tissue.

COL1A1 and *COL1A2* – human genes that encode for the type I collagen protein.

FKBP10 – the gene that encodes the FK506-binding protein that acts as a molecular chaperone. Chaperones are proteins that assist in the folding, unfolding, assembly and disassembly of other larger molecules.

SERPINH1 – the gene that encodes for the HSP47 protein that is a collagen specific chaperone

SERPINF1 - this gene encodes pigment epithelium-derived factor (PEDF), a multifunctional protein and one of the strongest inhibitors of angiogenesis currently known in humans

Osterix/SP7 – a putative master regulator of bone cell differentiation

Osteoblast – a cell that contributes to the formation of bone

CRAP, *LEPRE1* and *PPIB* – genes which encode for the cartilage associated protein, leprecan and cyclophilin B proteins. These proteins form a complex that helps process certain forms of collagen

UKNEQAS – United Kingdom National External Quality Assessment Service

EMQN – European Molecular Quality Network

1.0 BACKGROUND

Osteogenesis imperfecta (OI), also known as brittle bone disease, is an inherited connective tissue disorder characterized by low bone mass and increased bone fragility estimated to affect 1:15,000 to 1:20,000 individuals¹⁻³. Clinical severity can be highly variable ranging from mild, with few or no fractures, to lethal. Recurrent fractures in children, affects approximately 1 in 5 of those with a previous fracture, and is often accompanied by a reported family history of bone fragility. The scientific basis for excluding Osteogenesis Imperfecta (OI) in these children is poorly defined and relies heavily on clinical expertise.

In the majority of OI patients the disorder is caused by an autosomal dominant mutation in one of the two genes that encode the alpha chains of type I collagen, *COL1A1* and *COL1A2*. Research to identify the cause of OI in families without a mutation in these two genes (~ 10% of families) has focused on individuals with lethal forms of the disease. The discovery of an intracellular collagen-modifying complex, which is encoded by the *CRTAP*, *LEPRE1* and *PPIB* genes, led to the identification of severe autosomal recessive OI, with a proposed frequency of ~ 5% in lethal OI⁴. However, these genes have not been reported to be associated with mild or moderate OI phenotypes.

Recent evidence has shown that mutations in the osteoblast-specific transcription factor Osterix (Osx), encoded by the *SP7* gene⁸⁻⁹, can also be associated with severe OI, with homozygous (Knock-out) mutations in this gene being reported⁷. Mutations in genes involved in the regulation of bone formation are known to be associated with a variety of bone diseases, for example homozygous mutations in *LRP5* are associated with the severe osteoporotic disorder Osteoporosis Pseudoglioma (OPPG) whereas heterozygous mutations have a variable phenotype including Idiopathic Juvenile Osteoporosis (IJO) and reduced bone mass¹⁰. Heterozygous loss of function mutations in the *CBFA1* gene (Runx2), another transcription factor which acts upstream of Osx, are associated with Cleidocranial Dysplasia¹¹. Autosomal recessive mutations in the type I collagen chaperone genes, *FKBP10*¹²⁻¹³ and *SERPINH1*¹⁴ have also recently been reported to be associated with severe OI, as have mutations in *SERPINF1*¹⁷.

The prevalence and range of phenotypes associated with mutations in *FKBP10*, *SERPINH1*, *SERPINF1* and *SP7* is unknown and the phenotype of carriers of these mutations is not well documented. However these findings identify a previously unrecognized molecular mechanism in the pathogenesis of OI.

Hypothesis: Mutations in *FKBP10*, *SERPINH1*, *SERPINF1* and *SP7* are associated with mild/moderate OI phenotypes.

This is a pilot study to identify genes associated with mild/moderate OI by investigating families with multiple affected individuals, previously shown to lack a mutation in the type I collagen genes, for mutations in *FKBP10*, *SERPINH1*, *SERPINF1* and *SP7*. Identification of mutations in any of these genes in this cohort would have implications for the strategies by which individuals thought to have OI are evaluated at a genetic level as well as influence the investigation of recurrent fractures in infancy and childhood. The identification of causative mutations provides information to patients and

their families about their condition and its inheritance as well as improving the diagnosis and possible treatment of future OI patients.

Sheffield Diagnostic Genetics Service (SDGS) has been providing a national and international service for OI in collaboration with Professor Bishop, Professor of Paediatric Metabolic Bone disease, since 2002 based on analysis of the type I collagen genes with results published^{5,15-16}. A service for severe autosomal recessive OI by analysis of *CRTAP*, *LEPRE1* and *PP1B* was developed and introduced in 2009. The Laboratory has recently developed a comprehensive NCG funded service for Ehlers Danlos Syndrome as part of a portfolio of connective tissue disorders.

2.0 STUDY OBJECTIVES AND PURPOSE

The primary objective is to identify and recruit 20 individuals and their family members from Sheffield and around the UK who have been referred to the SDGS for investigation of a clinical diagnosis of mild/moderate OI and who are known from previous genetic analysis to lack a mutation in the type I collagen genes. This patient cohort will form the basis for analysis of the *SERPINH1*, *SERPINF1*, *FKBP10* and *SP7* genes.

By analysis of these genes we aim to identify novel mutations causing mild/moderate OI and improve the strategies by which children with recurrent fractures should be evaluated. A better understanding of the pathways and molecular mechanisms of OI has implications for treatment, genetic counselling and family studies. At present these families cannot receive the same standard of genetics counselling or make informed reproductive decisions without knowledge of the mutation responsible for their disease.

If the research is productive we plan to undertake further studies to ascertain the incidence of mutations in these genes more generally in the OI population and a possible role for these genes in influencing the phenotype of individuals with an identified type I collagen mutation would be investigated. Children with low bone mass and recurrent fractures could also be studied for mutations and polymorphisms in these genes.

It is not envisaged that a mutation will be identified in these new genes in all patients recruited, but a well characterised group of families will have been compiled which would be suitable for further studies to identify novel genetic causes of OI.

3.0 STUDY DESIGN

The study design is represented in figure 1. Clinical data supplied with the initial referral for type I collagen gene analysis will be retrospectively reviewed by the chief investigator for all mutation negative individuals. Any available pedigree information detailing further affected family members will also be reviewed. Those with a clear phenotype and strong family history will be identified and invited to participate in the study through their local referring Clinical Genetics Service.

Consent for genetic investigation for OI and storage of a DNA sample will have been obtained prior to previous analysis of the type I collagen genes. Recruitment will involve obtaining informed consent for further genetic analysis of the stored DNA sample for the *SP7*, *FKBP10*, *SERPINF1* and *SERPINH1* genes

in this study. If there is insufficient DNA or the DNA is of poor quality it may be necessary to obtain a blood sample (2-5mls EDTA) that would be obtained through the referring Clinical Genetics Service at a routine clinic appointment.

An extensive pedigree identifying index cases and affected plus unaffected family members should already be available in the patient's records but may require further clarification by the referring clinician, and will be used to provide information about possible modes of inheritance of the clinical phenotype in the family.

Analysis of DNA by sequencing the coding region and intron/exon boundaries of the candidate genes will be performed at the SDGS for each index case. Changes identified will be sought on locus specific database and dbSNP to identify previously reported variants. *In silico* analysis using amino acid conservation and splice prediction tools to indicate possible pathogenicity and to aid in prioritising further analysis will be performed. Family members will be investigated for any putative mutations identified to determine inheritance/penetrance.

At the end of the study findings will be returned to the local Clinical Genetics Services for dissemination to the participants.

There are no therapeutic interventions in this study and it is not envisaged that participants will withdraw from the study following initial recruitment. However, if for any reason patients decide to withdraw a discussion will be held regarding retaining/destroying samples and records already collected.

Primary and Secondary Endpoints

The primary end point will be recruitment of 20 index cases with mild/moderate OI without a mutation in the type I collagen genes.

The secondary endpoint will be completion of analysis of *SP7*, *SERPINH1*, *SERPINF1* and *FKBP10* for index patients and their families.

General Information

The risk to patients and their families is limited to one standard venepuncture, if required

The referring Clinical Geneticist, or their nominated genetic counsellor, will discuss the study with patients and their family and take consent. Participation in this study will not influence patient treatment.

4.0 SELECTION OF PARTICIPANTS

Individuals will be identified for the study if they have previously been referred to the Sheffield Diagnostic Genetics Service for diagnostic analysis of the type I collagen genes.

Inclusion Criteria

Individuals with a clinical diagnosis of mild/moderate OI and a strong family history but without a previously identified type one collagen gene mutation will be invited to participate.

Exclusion Criteria

Individuals without a clear clinical phenotype will be excluded as will those where analysis has been performed to exclude a diagnosis of OI in cases of suspected non-accidental injury (medico-legal testing).

5.0 PARTICIPANT RECRUITMENT

Consultant Clinical Geneticists who have referred individuals eligible for the study will be contacted with information on the study. Consultants will be asked to discuss the study with relevant patients and carers who are interested in participation either directly or through their Genetic counselling team. Patients who are interested in participation will be sent an invitation letter and age appropriate patient information sheet describing the study and its objective. After the patient (and/or carer) has had the opportunity to discuss the study with the referring clinician they will be given up to 6 weeks to decide if they wish to participate.

Consent for diagnostic genetic testing for OI and storage of DNA samples will have previously been obtained by the referring Clinical Genetics department prior to analysis of the type I collagen genes. Recruitment will involve obtaining consent for genetic analysis for the purpose of research and will be obtained by the clinician responsible for their care.

Where non-English speakers are invited to participate, standard arrangements used by Clinical Genetics to communicate with individuals and their families will be followed.

There will be no payment made for participating in this study.

Withdrawal of participants

It is not envisaged that withdrawal from the study will happen frequently. However, if a patient decides to withdraw from the study, all analysis will be halted. Clarification will be sought from the patient about information and samples to be retained/destroyed. A further eligible patient would be invited to participate in the study. Clinical care of any participant would not be affected by withdrawal from the study.

Data collection

A list of data to be collected and its source is shown in table 1.

The Chief Investigator will collate data into the research specific StarLims database.

6.0 DATA HANDLING AND RECORD KEEPING

The chief investigator will be responsible for the data collection, recording and quality.

Data will primarily consist of paper copies of referral forms, clinical and pedigree data. These will be stored in swipe card access controlled offices in the SDGS. All data will be scanned into the research specific Starlims Genetics database in accordance with standard operating procedures. Paper copies will then be destroyed in accordance with standard operating procedures. The Starlims database is secure and password protected and will be accessed by the research team only. The database is stored on a network computer drive which is password protected.

Laboratory worksheets, which record details of genetic analysis undertaken for the study, will be retained and stored as per standard operation procedures. Electronic records of DNA sequence will be generated during the course of the study..

STANDARD STATEMENT ON DATA PROTECTION

Data will be collected and retained in accordance with the Data Protection Act 1998.

STANDARD STATEMENT ON STORAGE OF RECORDS

Study documents (paper and electronic) will be retained in a secure location during and after the study has finished. All source documents will be retained for a period of 5 years following the end of the study. Where study related information is documented in the medical records – those records will be identified by a “Do not destroy before dd/mm/yyyy” label where date is 5 years after the last patient last visit.

7.0 ACCESS TO SOURCE DATA

The chief investigator will be the custodian of the data. Access to all data will be restricted to members of the research team.

The sponsor will permit monitoring and audits by the relevant authorities, including the Research Ethics Committee and the Medicines and Healthcare products Regulatory Agency (MHRA). The investigator will also allow monitoring and audits by these bodies and the sponsor, and they will provide direct access to source data and documents.

8.0 STATISTICAL ANALYSIS

This study is focused on finding changes in the *SP7*, *SERPINH1*, *SERPINF1* and *FKBP10* genes. No statistical analysis is required.

9.0 SAFETY ASSESSMENTS

A few participants may have a venous blood sample drawn. This may result in pain/discomfort at the site of venepuncture. Some participants may have subsequent bruising. These events will not require reporting.

Monitoring

The study will be monitored and audited in accordance with the Monitoring Standard Operating Procedures of the Clinical Research Support Unit. All study related documents will be made available on request for monitoring and audits by the Sponsor, the relevant Research Ethics Committee and for inspection by the MHRA or other licensing bodies.

The SDGS participates in UKNEQAS and EMQN quality assessment schemes.

10.0 ETHICAL CONSIDERATIONS

The study will be conducted in compliance with a Research Ethics Committee favourable opinion, including any provisions for Site Specific Assessment, and local Research and Development approval. The study will also be conducted in accordance with the International Conference for Harmonisation of Good Clinical Practice (ICH GCP), and the Research Governance Framework for Health and Social Care (2nd Edition).

Index cases recruited will already have had genetic analysis undertaken, but without the identification of a type I collagen gene mutation, so may be disappointed. Participants may be pleased to have the opportunity to possibly find out the cause of the OI in their family.

Each participant will be bled, the venepuncture may cause transient discomfort and bruising in a small number of cases. Where appropriate paediatric blood sampling equipment will be used to minimise this.

The patient and their family members will be asked about family relationships to establish or confirm a previously ascertained pedigree. Where relationships have broken down, this may represent unwelcome intrusion into the family.

DNA will be prepared and stored for several years. Samples will be securely stored and used only to seek mutations responsible for OI in the patients and their families. This study is expected to produce a cohort of well characterised patients and their family members, and novel mutations responsible for their OI will be identified in only a proportion, therefore a further application for funding is expected to follow. Participants will be asked for permission to store and use samples over several years and not just for this initial study,

however if patients do not consent to have their samples stored, then samples will be destroyed after the initial testing in this study.

Information on the patients and their family members will be held on computer. To ensure confidentiality, access to this information will be restricted to the study team.

Results of the study will be presented at relevant scientific meetings and in peer-reviewed scientific journals. Participants will be consented for publication at recruitment. The data will be anonymised prior to any publication.

11.0 FINANCE AND INDEMNITY

This is an NHS sponsored study. For NHS sponsored research HSG (96) 48 reference no. 2 refers. If there is negligent harm during the study when the NHS body owes a duty of care to the person harmed, NHS Indemnity will cover NHS staff, medical academic staff with honorary contracts and those conducting the study. NHS Indemnity does not offer no-fault compensation and is unable to agree in advance to pay compensation for non-negligent harm. Ex-gratia payments may be considered in the case of a claim.

12.0 REPORTING AND DISSEMINATION

Throughout the project, where significant findings are made, they will be presented at relevant national and international meetings (eg British Society for Human Genetics, International Conference on Osteogenesis Imperfecta). Manuscripts on the study findings will be submitted to relevant international journals. Mutations identified will be submitted to the international OI mutation database¹⁶, the most relevant means of communicating genetic analysis results with other laboratories undertaking OI genetic analysis.

For any mutations identified, laboratory analysis will be established and validated to NHS diagnostic standards (if not already) and the mutation reported to the patient's Clinician, thus any novel mutations will be added to the laboratories' diagnostic repertoire shortly after having been identified.

The results of the study will also be disseminated on the Trust R&D website and in the R&D newsletter.

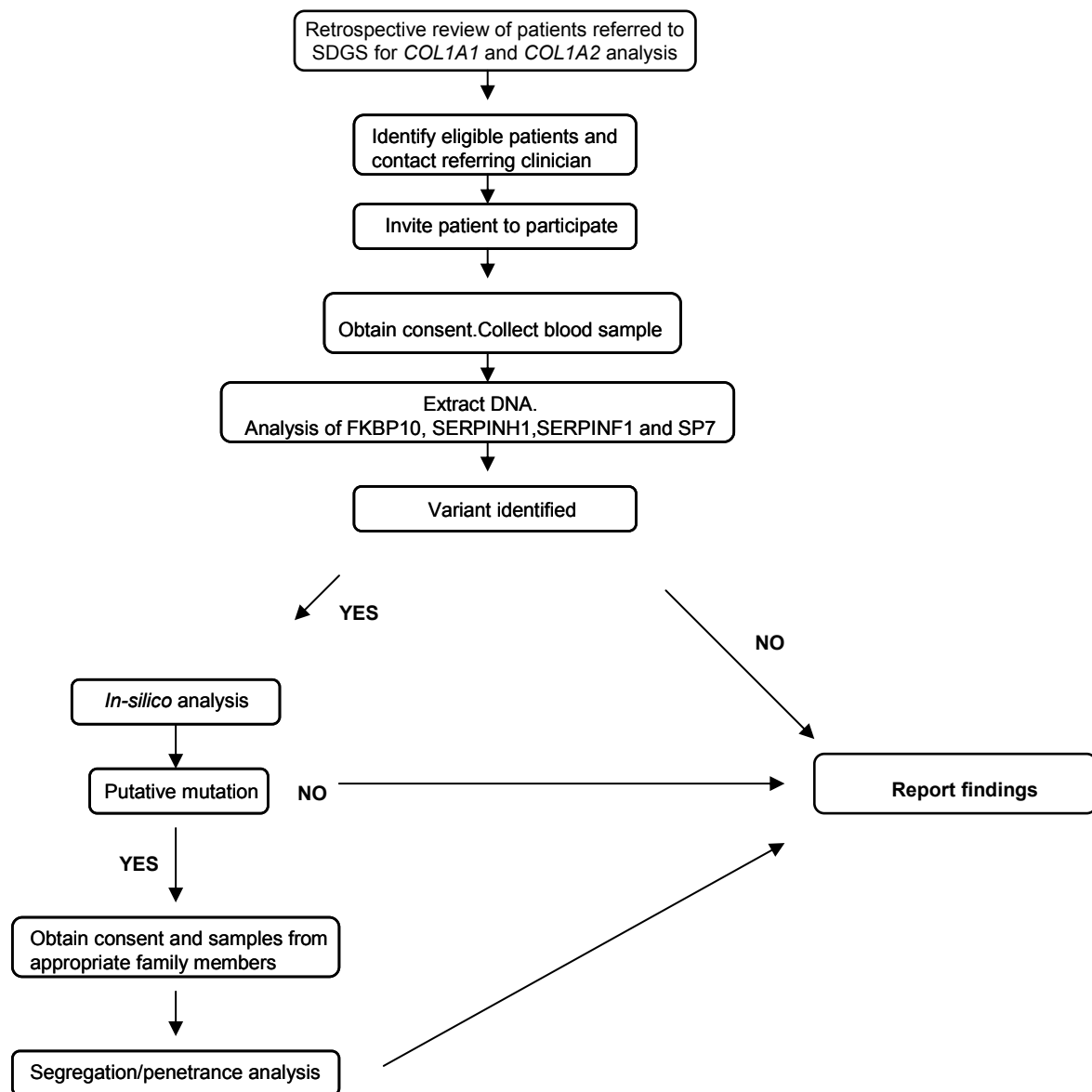
TABLES, FIGURES AND REFERENCES

Table 1_Data items to be collected (section 5.0).

Data item	Comment	Source of data
Full Name	Allows results of previous genetic analysis for OI to be traced and confirmed	Recruiting centre
Date of birth	Part of unique patient identifier	Recruiting centre
Genetic Counselling Family code	Numeric code, used to identify index case and family members	Recruiting centre
Clinical examination findings	Presence or absence of extra-skeletal manifestation of OI	Obtained from patient records by treating clinician
Fracture history	Number, age range for fractures and type of fractures sustained.	Obtained from patient records by treating clinician
Extent of previous OI analysis	Extent of previous OI analysis including; gene tested and methodology eg MLPA, DNA, RNA, collagen species analysis.	Copies of reports sent to referring clinician
Pedigree	Hand drawn or transferred to computer using Cyrillic software during previous interview prior to investigation of type I collagen genes.	Obtained from patient records by treating clinician.
Results of <i>FKBP10</i> , <i>SERPINH1</i> , <i>SERPINF1</i> , and <i>SP7</i> gene analysis	Results for sequence analysis of each gene.	SDGS data generated during the study.
Result of <i>in-silico</i> analysis	Extent and results of all <i>in-silico</i> analysis performed for any unclassified variant identified.	Data generated during the study.

APPENDICES

Figure 1 Schematic diagram of study design



References

1. Kuurila *et al* . Hearing loss in Finnish adults with osteogenesis imperfecta: a nationwide survey. *Ann Otol Rhinol Laryngol*. 2002;111:939–46
2. Wynne-Davies *et al* .Clinical and genetic patterns in osteogenesis imperfecta. *Clin Orthop*. 1981;159:26–35.
3. Forlino *et al* . Osteogenesis Imperfecta:Prospects for Molecular Therapies. *Mol Genet Meta.b* 2000; 71(1-2):225-32.
4. Bodian *et al* . Mutation and polymorphism spectrum in osteogenesis imperfecta type II:implications for genotype-phenotype relationships.*Hum Mol Gen*. 2009;18(3):463-471.
5. Van Dijk *et al* . PPIB mutations cause severe osteogenesis imperfecta. *Am J Hum Genet*. 2009;85(4):521-7.
6. Baldrige *et al* . CRTAP and LEPRE1 mutations in recessive osteogenesis imperfecta. *Hum Mut*. 2009;29(12):1435-42.
7. Lapunzina *et al* . Identification of a frameshift mutation in Osterix in a patient with Recessive Osteogenesis Imperfecta. *Am J Hum Genet*. 2010;87 110-112.
8. Yun-Feng *et al* .Sp7/Osterix is involved in the up-regulation of the mouse pro- α 1(V) collagne gene (Col5a1) in osteoblastic cells. *Matrix Biol*. 2010;29 701-706.
9. Yun-Feng *et al* .Sp7/Osterix up-regulates the mouse pro- α 3(V) collagne gene (Col5a β) during the osteoblast differentiation. *Biochem Res*. 2010;394 503-508.
10. http://chromium.liacs.nl/lovd2/home.php?select_db=LRP5
11. Otto *et al* Mutations in the RUNX2 gene in patients with cleidocranial dysplasia *Hum Mutat* 2002 19(3) 209-216
12. Alanay *et al* . Mutations in the gene encoding the RER protein FKBP65 cause autosomal recessive Osteogenesis Imperfecta. *Am J Hum Genet*. 2010;86(4) 551-559
13. Kelley *et al* . Mutations in FKBP10 cause recessive osteogenesis imperfecta and bruck syndrome. *J Bone Miner Res*. 2011;26(3)666-672.
14. Chritansen *et al* . Homozygosity for a missense mutation in SERPINH1 which encodes the collagen chaperone protein HSP47, results in Severe Recessive osteogenesis imperfecta. *Am J Hum Genet*.;2010 86 389-398.
15. Pollitt *et al* . Mutation Analysis of COL1A1 and COL1A2 in Patients diagnosed with osteogenesis imperfecta type 1-IV. *Hum Mut*. 2006; MIB #901
16. <http://www.le.ac.uk/ge/collagen>
17. Becker *et al*, Exome sequencing identifies truncating mutations in human SERPINF1 in autosomal recessive osteogenesis imperfecta. *Am J Hum Genet* 2011 88(3) 362-71.

K. DDD complementary research proposal submission form

DDD COMPLEMENTARY RESEARCH PROPOSAL SUBMISSION FORM

Proposal #	[to be filled in by central office]
Proposed by	Rebecca Pollitt/ Meena Balasubramanian
Research Group	Sheffield Diagnostic Genetics Service/ Sheffield Clinical Genetics Service Sheffield Childrens NHS Foundation Trust Western Bank Sheffield S10 2TH
Contact details	Rebecca.pollitt@sch.nhs.uk / Meena.balasubramanian@sch.nhs.uk
Date	30.1.13
Title of project	Novel molecular mechanisms leading to Osteogenesis Imperfecta
Brief description of project	<p>We aim to identify novel candidate genes for mild/moderate OI in patients without mutations in the type I collagen genes.</p> <p>How required subset of samples will be identified (e.g. specific HPO terms or codes) We will search the DDD database for phenotypes using the following HPO terms:</p> <ul style="list-style-type: none"> • Abnormal susceptibility to fractures • Fractures of the long bones • Fractures of vertebral bodies • Generalised osteoporosis with pathologic fractures • Increased fracture rate • Increased susceptibility to fractures • Recurrent fractures • Pathologic fractures • Vertebral compression fractures • Generalized osteopenia • Generalized osteopenia with pathologic fractures • Moderate generalised osteoporosis • Severe osteoporosis • Osteopenia • Spontaneous fractures <p>The type of genetic data required (e.g CNVs, exome variants)</p> <p>We will assess the phenotype in each case and review the copy number</p>

	<p>and exome data. Any variants identified will be assessed against the list of genes known to be involved in susceptibility fractures and those in pathways involved in skeletal development and bone homeostasis.</p> <p>The level of data required (e.g. variant files, raw data files) Variant files with array CGH and exome sequence results</p>
<p>Will the project require additional local ethical approval? (e.g. will additional investigations be required that are outside standard clinical care)</p>	No
<p>Will data be shared with investigators outside DDD? (i.e. investigators outside of the group of a DDD researcher)</p>	Yes, possibly
<p>If yes, what type of data and with whom?</p>	Nationally-commissioned OI service via Dr Balasubramanian
<p>Other DDD researchers who might be interested in this proposal</p>	Not at this moment
<p>Additional comments</p>	[any additional information you feel is relevant]

L. Confirmation of Ethical Approval for Target Exomes



Health Research Authority

NRES Committee Yorkshire & The Humber - Humber Bridge

Yorkshire and the Humber Research Ethics Office
First Floor
Millside
Mill Pond Lane
Leeds
LS6 4RA

Telephone: 0113 3050127
Facsimile: 0113 8556191

08 February 2012

Ms Rebecca C Pollitt
Lead Scientist, Connective Tissue Service
Sheffield Children's NHS Foundation Trust
Sheffield Diagnostic Genetics Service
Western Bank, Sheffield
S10 2TH

Dear Ms Pollitt

Study title: Investigation of the SERPINH1, SERPINF1, FKBP10 and SP7 genes to determine novel mutational mechanism responsible for Osteogenesis Imperfecta in families lacking a mutation in the type I collagen genes.

REC reference: 12/YH/0021
Protocol number: SCH/11/062

The Research Ethics Committee reviewed the above application at the meeting held on 25 January 2012. Thank you for attending to discuss the study.

Ethical opinion

You explained to the Committee that you considered the main ethical issues with your study was genetic testing involving children and the potential need for extra blood samples.

The Committee asked you if you intend to inform the participant's GP about their involvement in the study. You explained that you will not be informing the participant's GP and that this should be removed from study documentation. You confirmed that results will be passed on to the participant's Clinical Geneticist.

Members asked you how findings will be discussed with the participant's family. You explained that the tests that will be carried out as part of the research project are not currently available as diagnostic tests. You stated these patients are used to receiving results from research tests and understand the distinction between diagnostic and research findings. You explained that the test results will be passed on to the Clinical Geneticist who will confirm the result through diagnostic testing.

The Committee informed you that the consent form should include a clause stating that the participant understands that the tests are for research purposes, but that any findings will be fed back to the clinical team.

Members asked you what will happen to the tissue should a participant reach 16 years of age during the time the sample is stored. You confirmed that consent is taken from participants before their stored tissue is used in research. When a child reaches 16 years

old, consent will be taken from them and not the parent.

The Committee asked you how you will obtain extra blood samples for participants. You stated that blood will be taken at routine clinic appointments or by their local phlebotomists at the GP practice.

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Ethical review of research sites

NHS Sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at <http://www.rdforum.nhs.uk>.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations

1. The participant information sheet should explain that the study is part of a research degree.
2. The consent form should include a clause stating that the participant understands that although the test results are for research purposes they will still be fed back to them and their clinical geneticist.
3. The clause allowing participant's to consent to their GP being informed of their involvement in the study should be removed from the consent form.

It is responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

You should notify the REC in writing once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised

documentation with updated version numbers. Confirmation should also be provided to host organisations together with relevant documentation

Approved documents

The documents reviewed and approved at the meeting were:

<i>Document</i>	<i>Version</i>	<i>Date</i>
Covering Letter		20 December 2011
Investigator CV		20 December 2011
Letter from Sponsor		27 October 2011
Other: CV - N Bishop (Supervisor)		20 December 2011
Participant Consent Form: Assent Form	1.0	08 December 2011
Participant Consent Form	1.0	08 December 2011
Participant Information Sheet: Aged 0-5 years	1.0	08 December 2011
Participant Information Sheet: Aged 6-12 years	1.0	08 December 2011
Participant Information Sheet: Aged 13-15 years	1.0	08 December 2011
Participant Information Sheet: Aged 16+ years	1.0	08 December 2011
Participant Information Sheet: Parent/legal guardian	1.0	08 December 2011
Protocol	1.0	09 March 2011
REC application	1.0	20 December 2011

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National

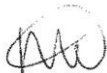
Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

Further information is available at National Research Ethics Service website > After Review

12/YH/0021	Please quote this number on all correspondence
------------	--

With the Committee's best wishes for the success of this project

Yours sincerely



nl **Dr Lynn Cawkwell**
Chair

Email: nicola.mallender-ward@nhs.net

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*

"After ethical review – guidance for researchers"

Copy to: *Ms Wendy Swann, Sheffield Children's Hospital NHS Foundation Trust*

NRES Committee Yorkshire & The Humber - Humber Bridge

Attendance at Committee meeting on 25 January 2012

Committee Members:

<i>Name</i>	<i>Profession</i>	<i>Present</i>	<i>Notes</i>
Reverend Annabel Barber	Chaplain	Yes	
Dr Stephen Beer	Consultant Physician	Yes	
Mrs Kate Bollington	Customer Services Manager	Yes	
Dr Lynn Cawkwell	Chair	Yes	
Dr Fiona Cowdell	Senior Research Fellow	Yes	
Mr Michael Davidson	Retired Senior Personnel Manager	Yes	
Mrs Dorothy Fagge	Specialist in Health Education	No	
Mr Michael Hockey	Consultant in Accident & Emergency	No	
Dr Sandeep Kapoor	Consultant Paediatrician	Yes	
Mrs Wendy Witter	Farm Administrator	Yes	

Also in attendance:

<i>Name</i>	<i>Position (or reason for attending)</i>
Nicola Mallender-Ward	REC Co-ordinator

M. NBAS Gene Dossier Approved by UKGTN



UK Genetic Testing Network

Proposal form for the evaluation of a genetic test for NHS Service Gene Dossier/Additional Provider

Submitting laboratory: Sheffield RGC
1. Disorder/condition – approved name (please provide UK spelling if different from US) and symbol as published on the OMIM database (alternative names will be listed on the UKGTN website). If NGS panel test, please provide a test name & the number of unique conditions across the whole of the panel test. If this submission is for a panel test please complete the Excel spread sheet, Appendix 1, available for download from the UKGTN website, and list all of the conditions grouped by sub panels if applicable.
NBAS related Disorder
2. OMIM number for disorder/condition If a panel test – see 1 above. If a number of subpanels exist with different clinical entry points e.g. cancer panel test but different subpanels for different types of cancer (breast cancer, colon, pheochromocytoma), then please list the sub panels here – providing name of each sub panel.
Short Stature, Optic Nerve Atrophy and Pelger-Huet Anomaly; SOPH (614800) Infantile Liver Failure Syndrome 2; ILFS2 (616483)
3a. Disorder/condition – to help commissioners to understand the impact of this condition please provide, in laymen's terms (e.g tubes in the kidney (renal tubule) or low sugar in the blood (hypoglycaemia), a brief (2-5 sentences/no more than 50 words) description of how the disorder(s) affect individuals and prognosis.
Individuals with this condition can have a range of symptoms including crises of recurrent acute liver failure in childhood that are life threatening and are triggered by infections that cause a high temperature (febrile infections) and / or by vaccinations. Other symptoms include short stature, fragile bones, problems with the immune system (immunological abnormalities) and vision (optic atrophy), and often a particular pattern of the white cells in the blood.
3b. Disorder/condition – if required please expand on the description of the disorder provided in answer to Q3a.
NBAS-related disorder results in a multi-system phenotype comprising bone fragility, growth failure, short stature, optic atrophy, recurrent liver failure (clinical or sub-clinical) and immune deficiency. Pelger-Huet anomaly is an inherited blood disorder where nuclei of white blood cells have an unusual bilobed or dumbbell shape, instead of the usual trilobed shape with a granulocyte shift to the left. It is usually a benign disorder although, in some instances, this can be indicative of a myelodysplastic syndrome necessitating a bone marrow examination. In NBAS-related conditions, PHA is usually a diagnostic (relatively inexpensive) clue to the diagnosis as this can be identified on blood film. However, it is neither a very specific or sensitive finding as can be seen in normal individuals as a benign finding or can be indicative of another condition (as mentioned above), hence genetic testing would be the gold standard for confirmatory diagnosis.
4. Disorder/condition – mode of inheritance r If this submission is for a panel test, please complete the mode of inheritance for each condition in the Excel spread sheet appendix 1 and if there is only one mode of inheritance across all conditions, please state it here or if it varies please provide proportion split here.
Autosomal recessive

Approval Date: Sept 2017

Copyright UKGTN © 2017

Submitting Laboratory: Sheffield RGC

<p>5. Gene – approved name(s) and symbol as published on HGNC database (alternative names will be listed on the UKGTN website) If this submission is for a panel test please complete the Excel spread sheet, Appendix 1, available for download from the UKGTN website, and list all of the genes grouped by sub panels if applicable.</p>
neuroblastoma amplified sequence; NBAS
<p>6a. OMIM number(s) for gene(s) If a panel test – see 5. above</p>
608025
<p>6b. HGNC number(s) for gene(s) If a panel test – see 5. above</p>
15625
<p>7a. Gene – description(s) If this submission is for a panel test, please provide total number of genes and if there are subpanels, please also list the number genes per sub panel.</p>
The NBAS gene encodes a protein, originally identified in neuroblastoma cells, that is thought to function as a component of an endoplasmic reticulum (ER) soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) complex. SNAREs mediate the docking and fusion of transport vesicles.
<p>7b. Number of amplicons to provide this test (molecular) or type of test (cytogenetic) (n/a for panel tests)</p>
52
<p>7c. GenU band (based on 2016 version) that this test is assigned to for index case testing. For NGS panel tests if there are sub panels, please provide GenU per subpanel.</p>
G
<p>8. Mutational spectrum for which you test including details of known common mutations (n/a for panel tests)</p>
Detection of all missense, nonsense, splice site mutations and small deletions and insertions. The mutations are scattered throughout the gene.
<p>9a. Technical method(s) – please describe the test.</p>
Currently Sanger sequencing. This will transfer to NGS at next panel revision.
<p>9b. For panel tests, please specify the strategy for dealing with gaps in coverage.</p>
<p>9c. Does the test include MLPA? (For panel tests, please provide this information in appendix 1)</p>
No
<p>9d. If NGS is used, does the lab adhere to the Association of Clinical Genetic Science Best Practice Guidelines for NGS?</p>
N/A
<p>10. Is the assay to be provided by the lab or is it to be outsourced to another provider? If to be outsourced, please provide the name of the laboratory and a copy of their ISO certificate or their CPA number.</p>
Provided by the laboratory

<p>11. Validation process</p> <p>Please explain how this test has been validated for use in your laboratory, including calculations of the sensitivity and specificity for the types of mutations reported to cause the clinical phenotype. Note that the preferred threshold for validation and verification is $\geq 95\%$ sensitivity (with 95% Confidence Intervals). Your internal validation documentation can be submitted as an appendix (and will be included in the published Gene Dossier available on the website). The validation information should include data on establishing minimum read depth and horizontal coverage for the regions of interest, reproducibility of the pipeline, accuracy of variant calling, filtering of common variants and artefacts.</p> <p>If this submission is for a panel test, please provide a summary of evidence of instrument and pipeline validation and complete the tables below. If the performance of the sub panels is expected to vary significantly to the data provided, please provide further details.</p>		
<p>Sanger sequencing has been extensively validated within the laboratory against scanning techniques (SSCP, CSGE). Sensitivity of DNA sequencing is over 95%. We participate in appropriate EQA</p>		
<p>12a. Are you providing this test already?</p>		
<p>Yes</p>		
<p>12b. If yes, how many reports have you produced?</p>		
	Sanger Based Tests	NGS Based Tests
	2	
<p>12c. Number of reports with a pathogenic (or likely pathogenic) mutation identified?</p>		
	Sanger Based Tests	NGS Based Tests
	1	
<p>12d. Please provide the time period in which these reports have been produced and whether in a research or a full clinical diagnostic setting.</p>		
<p>Clinical diagnostic testing offered 2016</p>		
<p>13a. Is there specialised local clinical/research expertise for this disorder?</p>		
<p>Yes</p>		
<p>13b. If yes, please provide details</p>		
<p>Dr Meena Balasubramanian, Consultant Clinical Geneticist, Sheffield Clinical Genetics Service Lead Consultant, OI-Genetics Service, Highly Specialised Severe, Complex and Atypical Osteogenesis Imperfecta Service Dr Fiona Shackley, Consultant Paediatric Immunology/Allergy/Infectious Diseases, Sheffield Children's Hospital Dr Sally Connolly, Consultant Paediatric Hepatologist, Sheffield Children's Hospital</p>		
<p>14. If using this form as an Additional Provider application, please explain why you wish to provide this test as it is already available from another provider.</p>		
<p></p>		

EPIDEMIOLOGY

15. Estimated prevalence and/or incidence of conditions in the general UK population

For panel tests, please provide estimates for the conditions grouped by phenotypes being tested.

Prevalence is total number of persons with the condition(s) in a defined population at a specific time (i.e. new and existing cases).

e.g. CF prevalence approx. 12 per 100,000 with UK population of approx. 63 million the prevalence of affected individuals in the UK is 7560

Incidence is total number of newly identified cases in a year in a defined population. e.g. CF incidence 1/2650 live births in a UK population with 724,000 live births in a year = 273 new cases a year

Please identify the information on which this is based.

No data available

16. Estimated gene frequency (Carrier frequency or allele frequency)

Please identify the information on which this is based.

n/a for panel tests.

No data available

17. Estimated penetrance of the condition. Please identify the information on which this is based

n/a for panel tests

No data available although Gene Review states 'penetrance is believed to be 100%'.

18. Estimated prevalence of conditions in the population of people that will be tested.

n/a for panel tests.

There are no figures for prevalence of this condition. It is a frequent cause of recurrent acute liver failure (RALF), and it may also be a frequent cause for recurrent, milder liver crises not fulfilling criteria of an ALF or isolated ALF in children (Staufner et al J Inherit Metab Dis (2016) 39:3-16).

INTENDED USE

19. Please tick either yes or no for each clinical purpose listed.

Panel Tests: a panel test would not be used for pre symptomatic testing, carrier testing and pre natal testing as the familial mutation would already be known in this case and the full panel would not be required.

Diagnosis	<input checked="" type="checkbox"/> Yes	<input type="checkbox"/> No
Treatment	<input checked="" type="checkbox"/> Yes	<input type="checkbox"/> No
Prognosis & management	<input checked="" type="checkbox"/> Yes	<input type="checkbox"/> No
Presymptomatic testing (n/a for Panel Tests)	<input checked="" type="checkbox"/> Yes	<input type="checkbox"/> No
Carrier testing for family members (n/a for Panel Tests)	<input checked="" type="checkbox"/> Yes	<input type="checkbox"/> No
Prenatal testing (n/a for Panel Tests)	<input checked="" type="checkbox"/> Yes	<input type="checkbox"/> No

TEST CHARACTERISTICS
<p>20. Analytical sensitivity and specificity</p> <p>The <i>analytical sensitivity</i> of a test is the proportion of positive results correctly identified by the test (true positive/true positive + false negative). The <i>analytical specificity</i> of a test is the proportion of negative results correctly identified by the test (true negative/true negative + false positive).</p> <p>This should be based on your own laboratory data for (a) the specific test being applied for or (b) the analytical sensitivity and specificity of the method/technique to be used in the case of a test yet to be set up. Please specify any types of mutations reported to cause the clinical phenotype that cannot be detected by the test.</p> <p>Note that the preferred threshold is $\geq 95\%$ sensitivity (with 95% Confidence Intervals).</p> <p>Sensitivity of DNA sanger sequencing is over 95%. Since all mutations are checked in two separate amplicons, if possible by two independent methods, specificity is 100% where the mutation or type of mutation (eg missense substitution, nonsense) has been previously reported.</p>
<p>21. Clinical sensitivity and specificity of test in target population</p> <p>The <i>clinical sensitivity</i> of a test is the probability of a positive test result when condition is known to be present; the <i>clinical specificity</i> is the probability of a negative test result when disorder is known to be absent. The denominator in this case is the number with the disorder (for sensitivity) or the number without condition (for specificity).</p> <p>Please provide the best estimate. UKGTN will request actual data after one year service.</p> <p>For a panel test, the expected percentage diagnostic yield for the test in the target population can be presented as an alternative to clinical sensitivity and specificity?</p> <p>There is no data available for either clinical sensitivity or specificity although they are both expected to be high.</p>
<p>22. Clinical validity (positive and negative predictive value in the target population)</p> <p>The <i>clinical validity</i> of a genetic test is a measure of how well the test predicts the presence or absence of the phenotype, clinical condition or predisposition. It is measured by its <i>positive predictive value</i> (the probability of getting the condition given a positive test) and <i>negative predictive value</i> (the probability of not getting the condition given a negative test).</p> <p>Not currently requested for panel tests</p> <p>There are no reports on clinical validity of testing for NBAS deficiency. However, literature reports and our data suggest that positive predictive value and penetrance is high.</p>
<p>23. Testing pathway for tests where more than one gene is to be tested sequentially</p> <p>Please include your testing strategy if more than one gene will be tested and data on the expected proportions of positive results for each part of the process. Please illustrate this with a flow diagram. This will be added to the published Testing Criteria.</p> <p>n/a for panel tests</p> <p>N/A</p>

CLINICAL UTILITY
24. How will the test change the management of the patient and/or alter clinical outcome? Please summarise in 2-3 sentences – no more than 50 words.
Impact on diagnosis based on the multi-system phenotype, avoids further unnecessary invasive investigations such as liver biopsy, bone marrow examination (as has been implemented locally in our patient cohort) and alters clinical outcome by institution of appropriate treatment (Pamidronate for bone fragility; immunoglobulin replacement therapy) and medications to prevent acute onset liver failure which is pyrexia-dependent in these patients.
25. Please provide full description on likely impact on management of patient and describe associated benefits for family members. If there are any cost savings AFTER the diagnosis, please detail them here.
Clarify recurrence risks, carrier testing for unaffected family members. As this is a recessively inherited condition, it is important to clarify the exact nature of mutations to confirm carrier status and provide the option of prenatal genetic testing for parents of children with this disorder. Cost savings would be avoidance of unnecessary investigations, sometimes invasive, saving substantially on not just the investigations but hospital stay etc.
26a. If this test was not available, what would be the consequences for patients and family members? Please describe in not more than 50 of words.
Prolonged patient journey with search for a diagnosis, several investigations some of which are invasive and unpleasant such as bone marrow examination, bone biopsy, liver biopsy which can be avoided if this testing was available and confirmed the diagnosis early on. Recurrence risk is up to 25%, hence a higher chance of recurrence in families without a proven diagnosis
26b. The consequences for patients and family members if this test was not available – if required please expand on the response provided in question 26a.
Patients would need to undergo invasive and sometimes unpleasant testing. The recurrence risk would be unknown. There would be unnecessary hospital admissions.
27. Is there an alternative means of diagnosis or prediction that does not involve molecular diagnosis? If so (and in particular if there is a biochemical test), please state the added advantage of the molecular test.
No as there are no hallmarks specifically linked to the condition. Abnormal liver function tests, immune deficiency are all very non-specific diagnostic criteria and the gold standard would be genetic testing.
28. Please list any genes where the main phenotype associated with that gene is unrelated to the phenotype being tested by the panel. For example, lung cancer susceptibility when testing for congenital cataract because ERCC6 gene (primarily associated with lung cancer) is included in a panel test for congenital cataract.
N/A
29. If testing highlights a condition that is very different from that being tested for, please outline your strategy for dealing with this situation.
N/A
30. If a panel test, is this replacing an existing panel/multi gene test and/or other tests currently carried out by your lab e.g. Noonan Spectrum Disorders 12 Gene Panel replaced multigene Sanger test for KRAS, RAF1, PTPN11 and SOS1? If so, please provide details below.
Not applicable
31. Please describe any specific ethical, legal or social issues with this particular test.
Nothing that we are aware of

Approval Date: Sept 2017

Copyright UKGTN © 2017

Submitting Laboratory: Sheffield RGC

32. REAL LIFE CASE STUDY**Please provide a case study that illustrates the benefits of this test**

H is a little boy now 10-years of age that we have been following up since birth. He has had several investigations- genetic and otherwise to identify a unifying cause for his difficulties. He presented at birth with significant failure-to-thrive, short stature, fragile bones resulting in multiple vertebral crush fractures needing treatment with Pamidronate.

He also went on to have persistently abnormal liver function tests for which no cause was identified necessitating a liver biopsy. He has immune deficiency needing regular immunoglobulin replacement therapy. He had bone marrow examinations to exclude a myelodysplasia.

Trio exome sequencing identified compound heterozygous mutations in NBAS which was confirmed through Sanger sequencing and parents were confirmed to be carriers.

This provided a unifying diagnosis for his problems and ensured that we avoided subjecting him to a planned bone marrow and liver biopsy. It also provided an explanation to the family 10 years on. Now that we know and understand the phenotypic spectrum, we have identified several more patients with this phenotype that are waiting testing for NBAS.

UKGTN Testing Criteria

Test name: NBAS Related Disorder	
Approved name and symbol of disorder/condition(s): Short Stature, Optic Nerve Atrophy and Pelger-Huet Anomaly; SOPH Infantile Liver Failure Syndrome 2, ILFS2	OMIM number(s): 614800 616483
Approved name and symbol of gene(s): neuroblastoma amplified sequence; NBAS	OMIM number(s): 608025

Referrals will only be accepted from one of the following:	
Referrer	Tick if this refers to you.
Consultant Clinical Geneticist	
Consultant Paediatric Immunologist	
Consultant Adult Immunologist	
Consultant Paediatric Hepatologist	
Consultant Adult Hepatologist	

Minimum criteria required for testing to be appropriate as stated in the Gene Dossier:	
Criteria	Tick if this patient meets criteria
• Bone fragility or slender bones AND	
• Short stature AND	
• Compatible with autosomal recessive inheritance	
AND at least one of the following:	
• Optic Atrophy (with Achromotopsia)	
• Liver disease (recurrent acute, childhood onset)	
• Immunodeficiency	
• Pelger-Huet anomaly (hypossegmented granulocytes)	
OR At risk family members where familial mutation is known.	

Additional Information:

If the sample does not fulfil the clinical criteria or you are not one of the specified types of referrer and you still feel that testing should be performed please contact the laboratory to discuss testing of the sample.

36. Based on experience what will be the national (UK wide) expected activity for requesting this test, per annum, for:

Index cases 20

Family members where mutation is known 40
If a NGS panel test, it is recognised that the full panel will not be used to test family members where the familial mutation is known. Please provide expected number of tests to inform completion of Q40

37. If your laboratory does not have capacity to provide the full national need please suggest how the national requirement may be met.

For example, are you aware of any other labs (UKGTN members or otherwise) offering this test to NHS patients on a local area basis only? This question has been included in order to gauge if there could be any issues in equity of access for NHS patients. If you are unable to answer this question please write "unknown".

We have the capacity to provide the full national need.

38. In order to establish the potential costs/savings that could be realised in the diagnostic care pathway, please list the tests/procedures that are no longer required to make a diagnosis for index cases where index cases would have the molecular genetic test proposed in this gene dossier at an earlier stage in the pathway. It is the tests/procedures that would be stopped for patients that are eligible for the gene test.

This information will be used to calculate the overall investment / savings required in Q39

Example:
 The introduction of a 95 gene panel for syndromic and non syndromic hearing loss would allow those patients who are recognised early enough in their pathway to diagnosis to be offered the genetic test instead of having sequential gene tests for individual genes already available and repeated ECGs, ERGs & renal ultrasounds as part of the diagnostic pathway although these may still be required as part of management after diagnosis.

	Type of test	Cost (£)
Imaging procedures		
Laboratory pathology tests (other than molecular/cyto genetic test proposed in this Gene Dossier)		
Physiological tests (e.g. ECG)		
Other investigations/procedures (e.g. biopsy)	bone marrow and liver biopsies	£440 £1710
Associated inpatient stays in the diagnostic pathway		£1005 £1150
Total cost of tests/procedures to be stopped (please write n/a if the genetic test does not replace any other tests procedures in the diagnostic care pathway)		£4305
If any of the tests/procedures listed above would be carried out on individuals after having the genetic test because the genetic test did not pick up a pathogenic mutation (i.e. negatives), please indicate the costs for these tests to continue to diagnosis.		£4305
<i>For example a panel test replaces single gene tests that have been included above, but after the panel test an individual that tests negative would not need to have these single gene tests, because the genes were on the NGS panel.</i>		

Approval Date: Sept 2017

Copyright UKGTN © 2017

Submitting Laboratory: Sheffield RGC

39. Please complete the Excel spread sheet available to download from the UKGTN website to calculate the estimated investment or savings, based on the expected annual activity of index & family cases (Q36 above) and using the information provided in Q38.

Number of index cases expected annually	20
Number of family member tests expected annually	40
Cost to provide index case test	£750
Cost to provide family member test	£155
Costs associated with tests/procedures for index cases if the genetic test in this Gene Dossier was not available	£4,305
Costs associated with tests/procedures for index cases that test negative for the genetic test in this Gene Dossier	£4,305
Total annual costs for diagnostic tests prior to introduction of the genetic test submitted for evaluation in this Gene Dossier	£86,100
Total annual costs to provide genetic test	£15,000
Additional savings or investment for 100% pick up rate for index cases	£71,100
Percentage of index cases expected not to find a pathogenic mutation (negatives)	50%
Number of index cases estimated to not find a pathogenic mutation (negatives)	10
Costs or savings to provide additional tests for index cases that test negative	£43,050
Total savings / investment prior to application of marginal reduction if applicable	£28,050
If a panel test and there are genes on the panel test that are already available on either other panel tests or single gene tests please estimate/suggest a marginal percentage reduction of the investment/savings. If you feel this is NOT applicable please leave this as 0%.	0%
Marginal percentage reduction if applicable applied to the savings/investment	£0
TOTAL SAVINGS / INVESTMENT for tests for INDEX CASES	£28,050
Total costs for family members	£6,200
If family testing is already available for any of the genes on this panel across the Network, please estimate the associated funding for these tests.	£0
TOTAL SAVINGS / INVESTMENT for tests for FAMILY MEMBERS	£6,200
ADDITIONAL INVESTMENT / SAVINGS FOR ALL ACTIVITY EXPECTED PER ANNUM	£21,850

40. Please indicate the healthcare outcomes that apply to this test after diagnosis. It is recognised that all tests recommended by the UKGTN for NHS service improve clinical management and, if a familial mutation is found, allows for prenatal testing and therefore these are not included in the list below. This information provides a useful guide to commissioners on the utility of the test.

Healthcare outcomes	Does this apply to this test?
1. Alerts significant clinical co-morbidities	Yes
2. Reduces mortality/saves lives	Yes
3. Avoids irreversible harm	No
4. Avoids diagnostic procedures/tests (some of which may be invasive) and/or multiple hospital appointments	Yes
5. Avoids incorrect management (e.g. medication or treatment) that could be harmful	Yes
6. Confirms targeted therapy/management	Yes
7. Earlier diagnosis allowing commencement of treatment earlier with associated improved prognosis	Yes
8. Enables access to educational and social support	Yes
9. At risk family members that test negative for a familial mutation can be discharged from follow up	Yes
10. At risk family members that test positive for a familial mutation have appropriate follow up	Yes