UNIVERSITY OF TARTU

FACULTY OF SCIENCE AND TECHNOLOGY

INSTITUTE OF MOLECULAR AND CELL BIOLOGY

DEPARTMENT OF BIOTECHNOLOGY

Lidiia Zhytnik

Genetic Predisposition to *Osteogenesis Imperfecta*: Exome Analysis of Estonian OI Families

Master's Thesis

Supervisors Dr. Katre Maasalu, PhD

Tiit Nikopensius, PhD

TARTU 2014

TABLE OF CONTENTS

ABBREVIATIONS

5’UTR – 5’Untranslated Region

AD – Autosomal-Dominant

AR – Autosomal-Recessive

BMD – Bone Mineral Density

BMP – Bone Morphogenetic Protein

BRIL – Bone-Restricted Ifitm-Like Protein

BS – Bruck Syndrome

CCDS – Consensus Coding Sequence

cDNA – Complementary DNA

CNV – Copy Number Variation

COMP – Cartilage Oligomeric Matrix Protein

CRTAP – Cartilage-Associated Protein

CyPB – Cyclophilin B

DI – *Dentinogenesis Imperfecta*

ECM – Extracellular Matrix

EDS – Ehlers-Danlos Syndrome

ER – Endoplasmatic Reticulum

FKBP65 - 65 kDa FK506 Binding Protein 10

gDNA – Genomic DNA

HSP47 – Heat Shock Protein 47

IGV - Integrative Genomics Viewer

Indel – Insertion-Delition mutation

LH2 – Lysyl Hydroxylase 2

LRP5 – Lipoprotein Receptor-Related Protein 5

MAF – Minor Allele Frequency

M-CSF – Macrophage Colony Stimulation Factor

MLBR – Major Ligand Binding Region

MMP – Matrix Metalloproteinase

MSC – Mesenchymal Stem Cell

NCP - Non-Collagen Proteins

NGS – Next Generation Sequencing

NMD – mRNA Nonsense Mediated Decay

OI – *Osteogenesis Imperfecta*

OPG – Osteoprotegerin

OSX – Osterix

P3H1 – Prolyl 3-Hydroxylase 1

PEDF – Pigment Epithelium-Derived Factor

PG – Peptidoglycan

PPIase – Peptidyl-Prolyl Isomerase

PTC – Premature Termination

RANKL – Receptor Activator of NF-κβ Ligand

rER – Rough Endoplasmatic Reticulum

RUNX2 – Run-related Transcription Factor 2

SNP – Single Nucleotide Polymorphism

SPARC – Secreted Proteins Acidic Rich in Cysteine

TNFα – Tumor Necrosis Factor-α

TRIC – Monovalent Cation-Specific Transmembrane Channel

WGS – Whole Genome Sequencing

INTRODUCTION

Osteogenesis Imperfecta is a rare Mendelian disorder, characterized by fragility of the bone tissue. The expressivity of the disease is very wide, varying from mild forms, when patients may not even know about running cases of the disorder in their family, to severe lethal forms, which include extreme fragility and skeletal deformations. The current classification of *Osteogenesis Imperfecta* remains ambiguous, since a lot of cases are boundary and individual. Patients often develop different sets of phenotypic manifestations even within one family. Some cases of the disease overlap with other disorders of connective tissues.

Genetic background of *Osteogenesis Imperfecta* is also diverse. For now, 16 different genes are discovered to be involved into pathogenesis of the disease. Mostly genes are connected functionally to bone development and homeostasis. However, a few genes of unexpected function for bone tissue were described as *Osteogenesis Imperfecta* causing.

Generally, OI-causing mutations occur in collagen type I genes. Depending on the type of mutation and its position, patients with mutations in the same collagen type I gene may develop all ranges of phenotypes, starting from mild forms to lethal forms. As more information from *Osteogenesis Imperfecta* studies becomes available, genotype-phenotype correlations become more confusing since people with the same genotypes reveal diverse phenotypes and vice versa.

Osteogenesis Imperfecta patients have a poor quality of life since no effective treatment exists so far. Collecting information about genetic and phenotypic diversity may help understand genotype-phenotype correlations and molecular mechanisms related to pathogenicity of the disease and lead to the development of efficient treatment.

The main aim of this thesis is to give an overview of *Osteogenesis Imperfecta* phenotypic, genetic and molecular background. In the experimental part of the study exome analysis of three Estonian OI families is performed in order to identify the disease causing mutations.

1. LITERATURE REVIEW

1.1 *Osteogenesis Imperfecta* review

Osteogenesis Imperfecta (OI) is a group of heterogeneous genetic disorders, known also as a „brittle bone disease“, Vrolik syndrome or Lobstein syndrome. The prevalence of OI ranges about 1 per 20,000 to 30,000 births[1] according to OI type being considered (Byers *et al.,* 1992; Steiner *et al.,* 1993; Sillence *et al.,* 1979). The estimated number of OI families in Estonia is 32.

OI is a disorder of connective tissue matrix. The hallmark of the condition lays in bone fragility and easy fractures, caused by decreased bone mass (Kocher and Shapiro, 1998). No deviations of D_3 vitamin, osteocalcin and parathyroid hormone levels were identified in OI patients (Mörike *et al.,* 1993; Palomo *et al.,* 2014). Patients may develop secondary clinical features, such as a short stature, *Dentinogenesis Imperfecta* (DI), osteoporosis, scoliosis, joint laxity, blue or grayish eye sclera, progressive hearing loss, easy bruising, skeletal deformity, Wormian bones and skull sutures. The range of affection is spread from mild osteopenia to moderate and severe forms, which include limb deformity and lethal cases (Roughley *et al.,* 2003). Sillence classification differentiates four basic classical types of OI, by severity. Recently the fifth type was added to traditional classification. In addition to genetic predisposition, environment (climate, physical activity, diet, treatment) may influence the phenotype development of an OI patient (Roughley *et al.,* 2003).

Genetics of this disorder is extremely heterogeneous. The pattern of inheritance can be autosomal-dominant (AD), as well as autosomal-recessive (AR). More than 1,000 mutations in 16 different genes are already known to cause *Osteogenesis Imperfecta*.[2] On the basis of the affected genes, genetic OI classification differentiates 15 OI types.

Molecular pathology of OI is based on qualitative or quantitative defects of type I collagen. About 85-90% of OI cases are caused by dominant mutations in *COL1A1* and *COL1A2* genes, encoding for collagen I proα1 and proα2 polypeptide chains respectively (Byers and Pyott, 2012), and *IFITM5* gene (Lazarus *et al.,* 2014). The rest - 10-15% of OI cases represent

---

[1] http://www.orpha.net/

[2] https://oi.gene.le.ac.uk/home.php 25/05/2014

recessive forms of the disorder. Recessive OI is caused by the defects in genes, connected with collagen type I post-translational modifications (*CRTAP, LEPRE1, PPIB, PLOD2*), folding and transport (*SERPINH1, FKBP10),* osteoblasts gene transcription and proliferation (*OSX, CREB3L1),* and mineralization (*SERPINF1*) or cell signaling (*WNT1, TMEM38B*). However, some cases of AR-OI with mutations in *COL1A1* and *COL1A2* genes are also known (Forlino *et al.*, 2011; Glorieux and Moffatt, 2013).

1.2 Bone biology

Bones are endoskeleton organs of vertebrates, composed of dense connective osseous tissue. Adult human skeleton consists of 206 bones which accomplish a number of mechanic, synthetic and metabolic functions (Steele *et al.,* 1988). Bones act as an endocrine organ, store minerals, release growth factors and hormones, and provide hematopoiesis (Lee *et al.,* 2007). Bones provide movement of muscles, protect internal organs, support the body, take part in sound transduction, and are extremely important in functioning of the whole human body (Currey, 2002).

1.2.1 Cellular structure of bone tissue

Osseous tissue is a mineralized dynamic structure, composed of organic and non-organic parts. Bone mineral residues increase bone strength. Organic part of the bone supplies it with elasticity (Clarke *et al.,* 2008). Bone is constantly reorganized by osteoblasts, which produce matrix and mineral parts, and osteoclasts, which participate in bone resorption (Figure 1) (Clarke *et al.,* 2008).

Osteoblasts synthesize type I collagen and mineralize it by saturating intracellular matrix with calcium and phosphate ions in the form of hydroxyapatite $Ca_{10}(PO_4)_6(OH)_2$. Osteoblasts also secrete non-collagen proteins (NCP), hormones and enzymes which effect osteogenesis (Buckwalter *et al.,* 1996). Osteoblasts are differentiated from mesenchymal stem cell's (MSCs) under the regulation of pigment epithelium-derived factor (PEDF), Osterix (OSX), run-related transcription factor 2 (RUNX2) and SOX9 transcription factor (Long, 2012). Osteoblast differentiation is promoted by WNT and bone morphogenetic protein (BMP) signaling pathways (Lin and Hankenson, 2011).

Trapped in dense bone tissue, osteoblasts differentiate into osteocytes and gain regulatory function towards osteoblasts and osteoclasts, which are derived from monocyte-macrophages precursors (Buck and Dumanian, 2012). Osteocytes control oxygen and mineral levels in bones (Figure 1) (Currey 2012, Noble 2008). They are able to influence bone formation indirectly with parathyroid hormone or modulate bone mass by canonical WNT-β-catenin signaling pathway (Cui *et al.*, 2011; Clarke *et al.*, 2008).

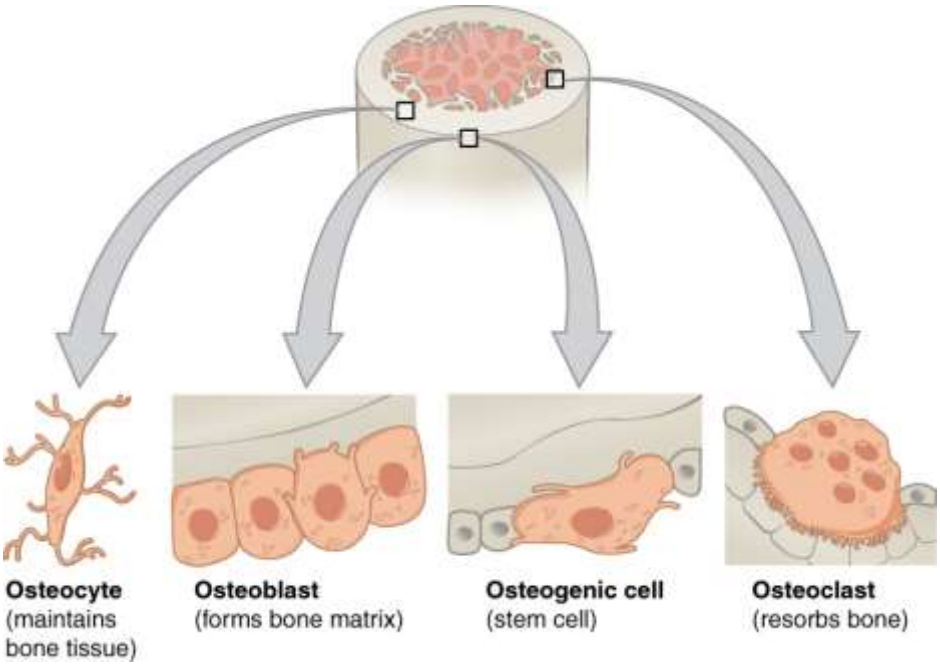

Figure 1 Bone cells: osteocyte, osteoblast, osteogenic stem cell, osteoclast (Betts *et al.*, 2013).

Osteoclasts are large multinuclear cells with phagocytic-like activity against mineral and organic components of the bone (Figure 1). Cells of osteoblastic lineage regulate osteoclastic proliferation and resorption processes via macrophage colony stimulation factor (M-CSF), receptor activator of NF-κβ ligand (RANKL), tumor necrosis factor-α (TNFα), and osteoprotegerin (OPG) (Clarke *et al.*, 2008).

1.2.2 Bone extracellular matrix and matrix mineralization

Bone extracellular matrix (ECM) is composed of collagen proteins on 85-90% (Miller 1984). There are 19 types of collagens which slightly differ by primary structure. Bone, skin, dentin and tendon collagen share the same structure and are recognized as type I collagen (Gelse *et al.*, 2003).

Only in vertebrates collagen undergoes transformation, becoming mineralized skeletal structure, which defines the main biomechanical properties of the bone. Mineral crystals are connected to collagen fibrils by ionic and Van der Waals forces (Currey 2012). Bone collagen fibrillar network is enriched by NCP – integrin-binding glycoproteins and proteoglycans – which regulate deposition of minerals into ECM (Horton *et al.*, 1995). For example, alkaline phosphatase binds to the surface of osteoblasts and increases matrix mineralization (Clarke *et al.*, 2008).

1.2.3 Collagen type I biosynthesis

Collagen type I is synthesized in the rough endoplasmatic reticulum (rER) as a procollagen heterotrimer. Precursor molecule consists of triple helix – two proα1 and one proα2 peptide chains, encoded by *COL1A1* and *COL1A2* genes respectively.

The amino acid sequence pattern of collagen is unusually rich with glycine (Gly) and proline (Pro). Glycine occupies almost every third position of the peptide chain, which is composed of triple amino acid motifs of Gly-X-Y. Preprocollagen is translated from mRNAs on ribosome of rER and has signal propeptides at N- and C-terminal ends (Boedtker *et al.*, 1983; Cundy 2012). Then, preprocollagen enters rER lumen, where N- and C-terminal signal peptide domains are removed by signal peptidase.

The formed procollagen undergoes hydroxylation of proline and lysine (Lys) residues, in specific positions relative to glycine locations. Pro is modified into hydroxyproline by prolyl-4 hydroxylase and prolyl-3 hydroxylase. Y-position Lys is hydroxylated with lysyl-hydroxylase-1. Hydroxylation is extremely important for forming a stable structure with intramolecular cross-links, responsible for the bone tissue strength (Gelse *et al.*, 2003; Cundy 2012; Currey 2012). Glucose and galactose are added by hydroxylysyl glycosyltransferase and hydroxylysyl galactosyltransferase to specific hydroxylysine residues. Preprocollagen is enriched with intra-chain disulfide bonds.

Heat shock protein 47 (HSP47) binds to N-terminus and acts as a chaperone in order to prevent premature triple helix formation (Ishida and Nagata, 2011). Afterwards prepropeptide at C-terminus is synthesized and molecule is folded into triple γ-helical structure from C-terminus to N-terminus. Folding is proceeded only with transpeptide bonds of Pro. Peptidyl-prolyl isomerase

(PPIase) changes peptide bond from cis position into trans position where needed. Then, triple collagen chain is formed (Gelse *et al.*, 2003). (Supplementary Figure 1).

Procollagen is transported into Golgi complex and packed into vesicles for extracellular transport. Outside the cell, registration peptides are removed from procollagen chains with procollagen peptidase. The processed molecule is called tropocollagen. Tropocollagen assembles into fibrils (Supplementary Figure 1). Collagen type I microfibril is composed of five overlapping monomers. Monomers overlap by 234 residues, forming 67 nm D-period, which is a basic repeat structure of a fibril. Each period consists of a "gap" zone and an "overlap" zone, composed by dark and light bands (Figure 2). The fibrils covalently cross-link by lysyl oxidase and form collagen type I fibers (Gelse *et al.*, 2003; Last and Reiser, 1984) (Supplementary Figure 1).

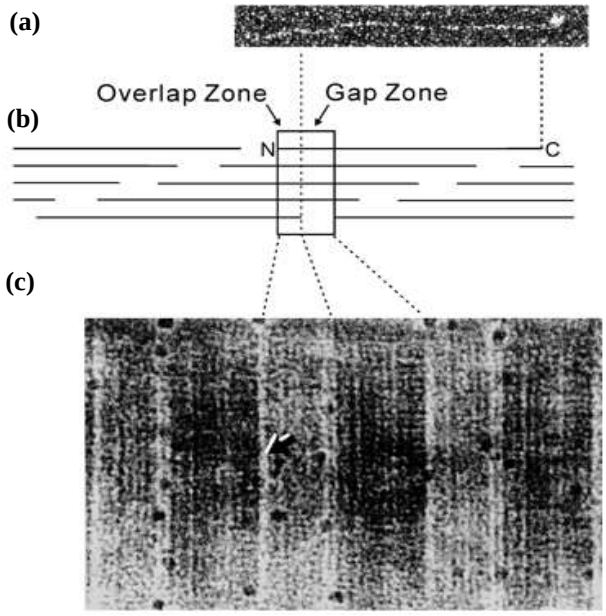

Figure 2 The structure of collagen type I fibril (a) Procollagen triple helix by electron microscopy with C-terminus on the right and N-terminus on the left (b) Collagen fibril model with a one D-period included into box region (c) Collagen fibril visualization in electron microscopy and positive-negative staining (Di Lullo *et al.*, 2002).

1.2.3.1 Structural domains of collagen type I α1 and α2 chains

The collagen fibril "overlap" zone consists of cell interaction and structural domains. Cell interaction domain is represented by integrin and peptidoglycan (PG) domain receptors. Binding sites for molecules, bridging cell surface and ECM, are also situated in the cell interaction domain. Usually ligands bind to a few sites of a fibril, which originate from different monomers, composing intramonomer multivalent binding pattern (Sweeney *et al.,* 2008).

C-terminal side of a collagen monomer is rich in ligand-binding sites. There are three "hotspots" of compact distribution of the ligand-binding sites, called major ligand binding regions (MLBRs). They are composed by residues 80-200 (MLBR1), 680-830 (MLBR2), and 920-C terminus (MLBR3) (Di Lullo *et al.,* 2002).

Integrin-binding sites are represented by 10 regions on the collagen monomer, one of which is a Gly-Phe-Hyp-Gly-Glu-Arg sequence, known as a GFPGER site. The integrin binding is important for tissue morphogenesis, ECM assembly, and cell signaling (Sweeney *et al.,* 2008). Fibrilogenesis also depends on PG, phosphophoryn, fibronectin, cartilage oligomeric matrix protein (COMP) binding sites etc. These proteins provide collagen assembly, arrangement and mineralization. Ligand-binding sites influence each other directly, by competition for binding proteins on overlapping regions, and indirectly, by interfering other binding sites (Di Lullo *et al.,* 2002).

Matrix interaction domain consists of binding sites for structural proteins and regions of fibril's cross-linking. In "gap" regions are situated binding sites for hydroxyapatite. Thus, the domain controls mineralization of the tissue (Sweeney *et al.,* 2008) (Supplementary Figure 2).

1.2.4 OI effect on bone structure

OI effect on bone structure depends on a particular OI causing defect. General tendencies include bone fragility and low bone mass phenotype. OI patients reveal the presence of immature woven bones with poor lamellar structure, decreased trabecular bone volume, and reduced osteoid quantity. If collagen processing enzymes are defective, collagen structure is overmodified, and the quantity of normal collagen, expression of NCP and transcription factors may be reduced (Sarafova *et al.,* 1998; Gioia *et al.,* 2012). Irregular NCP binding to collagen may cause some mineralization defects. When decreased, collagen matrix is combined with increased mineral content, and alkaline phosphatase activity in OI patients is increased as well as mineralization levels, compared to age-matched controls (Gioia *et al.,* 2012).

If extracellular signal transduction is defective, cell differentiation is disrupted by the absence of normal feedback signals from collagen protein. The number of osteoclasts and osteoblasts is increased (Baron *et al.*, 1983). However, the amount of mature differentiated osteoblasts may also be downshifted due to apoptosis, caused by stress responses to defective protein synthesis (Lisse *et al.*, 2008) (Figure 3).

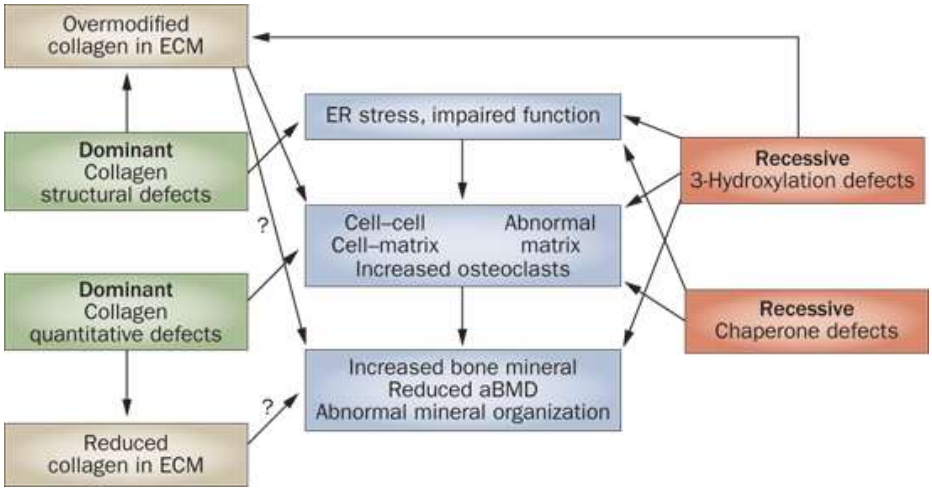

Figure 3 Features influenced by dominant (left column) and recessive (right column) forms of OI. In the middle are listed the features common for both types of OI (Forlino *et al.*, 2011).

1.3 OI classification

Classification of OI is complicated by heterogeneous phenotypes and variety of the genetic basis of molecular pathogenic mechanisms. The range of affection is spread from mild osteopenia to moderate and severe forms, which include limb deformity and lethal cases. In reality, the disorder represents the continuum of severity. Phenotypic traits can vary within a single type of OI interfamilially and even intrafamilially (Basel *et al.*, 2009). Early attempts of OI classification were based on the age of fracture occurrence. It included *osteogenesis imperfecta congenita*, with the presence of intrauterine fractures; and *osteogenesis imperfecta tarda* subclassified into *gravis* and *levis* forms, with fractures occurred respectively during the first year after birth and subsequently (Shapiro *et al.*, 2013).

1.3.1 Original Sillence Classification

In 1979, clinical classification based on skeletal features was proposed by Sillence (Sillence *et al.*, 1979). It characterizes four basic types of OI (I-IV) on the basis of clinical, radiological and inheritance patterns of the condition (Table 1). The numbers of types do not represent severity of OI form, but were given accordingly to a description order of OI group.

1.3.1.1 OI Type I (Classical non-deforming OI with blue sclera)

OI Type I is a mild, non-deforming form of OI, caused by autosomal dominant mutations in the *COL1A1* and *COL1A2* genes. Patients develop haploinsuffiency of collagen type I (Roughley *et al.*, 2003). The incidence of OI Type I is about 1 per 10,000 to 20,000 births.[1] The disorder is defined by osseous fragility, mild joint hypermobility, blue eye sclera, easy bruising, progressing hearing loss and progressing scoliosis. Blue sclera is developing due to thin collagen layers so that the choroid layer shines through. Patients usually develop near normal stature with slightly reduced height and a tendency to shorten with age. Height reduction is caused by insignificant lower limb and spinal deformity as well as the consequences of spinal osteoporosis – progressive platyspondyly and kyphosis (Sillence *et al.,* 1979) (Figure 4a). Some patients have specific – to OI – appearance with increased head circumference and triangular face. Skull radiographs can show islands of irregular calcification. Hearing loss usually develops during the third decade, although in the most severe hearing impairment cases, it is noticed within the second decade of life (Sillence *et al.,* 1979). *Dentinogenesis Imperfecta* and considerable bone deformity cases are rare. The first fractures occur at birth or during juvenile period. Patients are able to walk independently. The number of fractures varies individually and usually decreases after puberty. A new wave of susceptibility to fractures starts in menopausal females and males of early sixties, due to the development of osteoporosis (Sillence *et al.,* 1979; Basel *et al.*, 2009; Roughley *et al.*, 2003).

1.3.1.2 OI Type II (Perinatally lethal OI)

OI Type II is the most severe, perinatal lethal form of OI. Fetuses are spontaneously aborted or die during the first days after birth (Sillence *et al.,* 1979). The prevalence of OI Type II remains unknown. Individuals rarely succeed to survive more than few days. Death is a result of pulmonary insufficiency caused by small thorax, rib fractures, or flail chest. Intrauterine bone fractures, severe skeletal deformity, severe limb shortening, rhizomelia, diminution of calvarial

thickness, dark blue-black eye sclera, and small height and weight are present. The skull is soft, due to decreased mineralization, and large in proportion to the whole body. Femur tends to be crumpled (accordion-like) and angulated. The presence of beaded ribs, platyspondyly, fractures and deformity of shafts of bones of upper limbs is usually noticed (Sillence *et al.*, 1979). OI Type II arises from parental mosaicism for autosomal dominant mutations in the *COL1A1* and *COL1A2* genes (Edwards *et al.*, 1992). Autosomal recessive mutations in genes, connected with collagen type I post-translational modification, folding, transport or osteoblast proliferation and cell signaling, also may cause OI phenotype, overlapped with OI Type II (Basel *et al*., 2009; Roughley *et al.*, 2003; Marini *et al*., 2007).

1.3.1.3 OI Type III (Progressively deforming OI with normal sclera)

OI type III is the most severe survival form of OI with heterogeneous clinical symptoms with the incidence of 1 per 60,000 to 100,000.[1] The third type of OI is defined with extreme bone fragility and progressive severe skeletal deformity - bones are twisted and bowed. Skeletal deformities are caused by mechanical forces of muscles or past fractures. Severe spinal deformity in adults is also observed. Other symptoms include: severe scoliosis, multiple fractures at birth, extremely high number of fractures during whole life, Wormian bones, *Dentinogenesis Imperfecta*, and ligamentous laxity in childhood (Sillence *et al.*, 1979). Mobility of patients is available with assistance. With aging, individuals develop short stature - less than 1 m for adults (Figure 4c). Eye sclera is usually relatively normal – white, slightly grayish or pale blue, with the tendency to lose dense during aging. Triangular facial shape, and temporal and occipital bossing give patients typical OI appearance (Sillence *et al.*, 1979). Progressive hearing loss begins in puberty. OI Type III is caused by autosomal *de novo* dominant mutations in *COL1A1*, *COL1A2* and autosomal recessive mutations in genes connected with collagen modifications and transport, osteogenic cells proliferation or signaling pathways in bone tissue (Basel *et al.*, 2009; Roughley *et al.*, 2003; Marini *et al*., 2007).

1.3.1.4 OI Type IV (Common variable OI with normal sclera)

OI type IV is the most diverse group of phenotypes, varying from mild to severe forms of osseous fragility. Skeletal deformity is moderate (Figure 4b). Blue sclera tends to lighten after birth and results in normal scleral hue. The stature is shorter than in OI type I (Sillence *et al.*, 1979). DI and Basilar invagination, scoliosis and osteoporosis are also present. The pattern of

inheritance is autosomal dominant. OI IV associated mutations are present in *COL1A1* and *COL1A2* genes (Basel *et al.*, 2009; Roughley *et al.*, 2003).

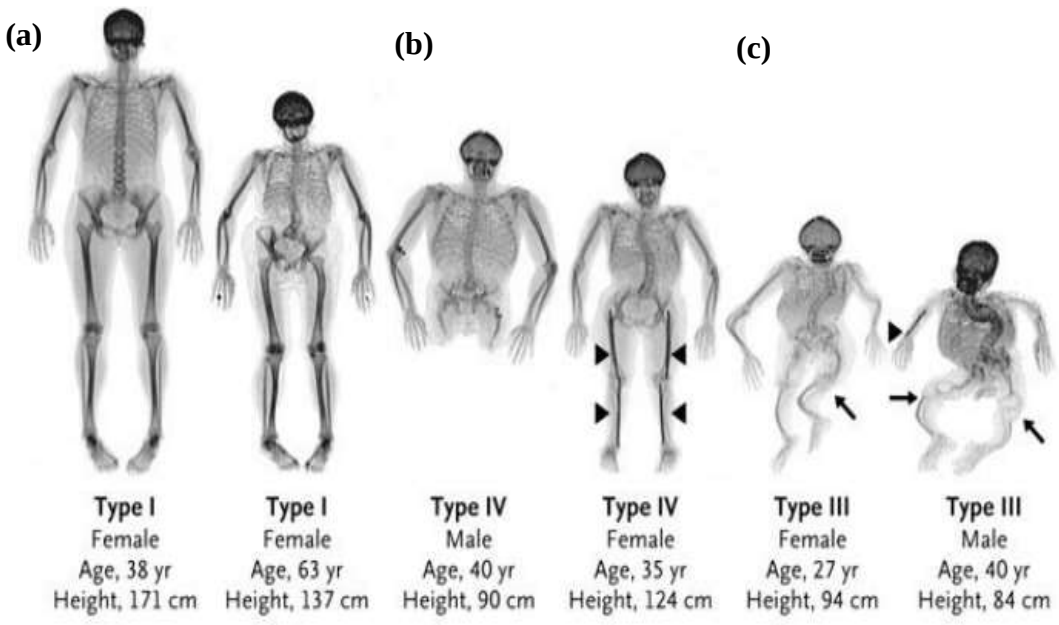

Figure 4 Radiographs of adults with OI types I (a), III (c), IV (b) (Reeder and Orwell, 2006).

1.3.2 Modern OI classification

After the discovery of non-collagen genes connected with OI, the original Sillence classification was expanded. In 2010, the International Nomenclature Group for Constitutional Disorders of the Skeleton suggested adding of the fifth OI type (OI Type V with calcification of interosseous membranes) to the classical OI types to avoid confusing molecular references in the clinical classification (Table 1). OI Type V was distinguished separately, due to special histological traits, which differed from changes in bone architecture of the previously described OI types. (Van Dijk *et al.*, 2010; Glorieux *et al.*, 2000).

1.3.2.1 OI Type V (OI with calcification of interosseous membranes)

Some phenotypes of OI Type IV differ during histological analysis of the bone tissue. These forms were distinguished in separate OI type. OI Type V is moderately deforming OI form with moderate to severe bone fragility (Glorieux *et al.*, 2000). Blue sclera, ligament laxity and DI are absent. Mostly, the height of patients is short but some are above average. Hearing loss and

Wormian bones are rare. Scoliosis was found in half of the patients. Skin, tendon, ligament, and fasciae have normal texture. Chest wall, elbow and forearm deformities may be present (Kim *et al.*, 2013). Special traits of OI Type V is hypertrophic callus in fracture sites, calcification of interosseous membranes, non-traumatic radial head dislocation, metaphyseal radiodende band adjacent to growth plates upon X-ray (Cheung *et al.*, 2007) (Figure 5a). Moreover, lamellar organization of the bone has an irregular mesh-like (butterfly-like) appearance (Figure 5b). OI Type V arises as a result of heterozygous mutations in the 5' untranslated region (5'UTR) of the *IFITM5* gene (Lazarus *et al.*, 2014). Recently, missense mutations causing OI type V in coding regions of *IFITM5* were also found (Farber *et al.*, 2014; Guillén-Navarro *et al.*, 2013).

(a) **(b)**

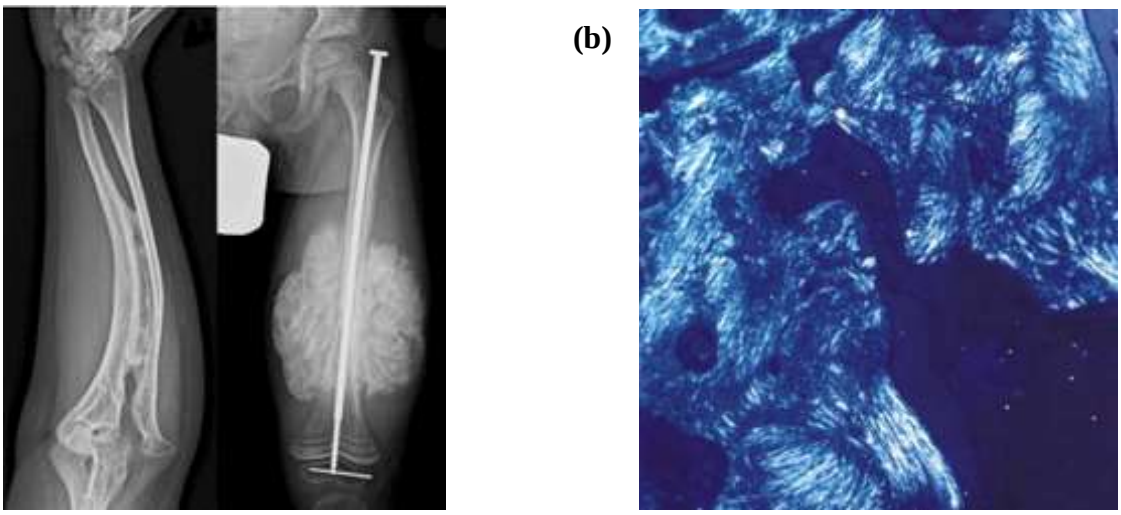

Figure 5 (a) Radial head dislocation and hypertrophic callus formation in a patient with OI type V (b) Irregular mesh-like pattern of bone tissue in a patient with OI type V (Lee *et al.*, 2006).

Table 1 OI Syndrome nomenclature according to International Nomenclature Group for Constitutional Disorders of the Skeleton (INCCDS) 2010* (Shapiro *et al.*, 2013).

Syndrome Name	OI Numerical Type	Subtypes
Classical non-deforming OI with blue sclera	I	2
Common variable OI with normal sclera	IV	2
OI with calcification of interosseous membranes	V	1
Progressively deforming OI with normal sclera	III	12
Perinatally lethal OI	II	6

*Syndromes are listed in order of increasing severity

1.3.3 Genetic OI classification

Genetic classification is based on the inheritance pattern and molecular mechanisms of an OI type. Each gene, affected by pathogenic OI mutation, generally forms a discrete OI type. Besides types I-V, genetic classification differs OI types VI-VII (Glorieux *et al.,* 2002; Homan *et al.,* 2011) and additional OI type VIII (Cabral *et al*., 2007). Also, the alternative classification suggests the presence of extra OI types IX-XV (Van Dijk *et al*., 2010; Barnes *et al.,* 2006, Christiansen *et al.,* 2010; Martinez-Glez *et al.,* 2012; Shaheen *et al.,* 2012; Pyott *et al.,* 2013; Barnes *et al.,* 2006) (Table 2). The difficulty stands in clinical impractice of such system, as many phenotypes tend to overlap. For example, types VII and VIII are sometimes excluded due to concordance of clinical and radiological features with types II-IV. Also, clinical OI classification usually considers genetic OI types VI-XV as subtypes of an OI type III.

Table 2 Genetic OI classification (Shapiro *et al.,* 2013)

OI Type	Phenotype severity	Inheritance	Gene	OMIM
I	Mild, non-deforming	AD	*COL1A1*	#166200
II	Perinatal lethal	AD/AR	*COL1A1, COL1A2, CRTAP, LEPRE1, PPIB, CREB3L1*	#166210
III	Severe, deforming	AD/AR	*COL1A1, COL1A2, CRTAP, LEPRE1, PPIB*	#259420
IV	Moderately deforming	AD	*COL1A1, COL1A2*	#166220
V	Mild to moderately deforming	AD	*IFITM5*	#610967
VI	Hyperosteoidosis	AR	*SERPINF1*	#613982
VII	Moderately deforming	AR	*CRTAP, LEPRE1*	#610682
VIII	Severe, lethal	AR	*CRTAP, LEPRE1*	#610915
IX	Moderate to severe phenotype	AR	*PPIB*	#259440
X	Moderate to severe phenotype	AR	*SERPINH1*	#613848
XI	Moderate to severe phenotype	AR	*FKBP10*	#610968
XII	Moderate to severe phenotype	AR	*OSX/SP7*	#613849
XIII	Moderate to severe phenotype	AR	*BMP1*	#614856
XIV	Moderate to severe phenotype	AR	*TMEM38B*	#615066
XV	Moderate to severe phenotype	AR	*WNT1*	#615220

1.3.4 Brittle bone overlap phenotypes

Cases of *Osteogenesis Imperfecta* that overlap with Ehlers-Danlos syndrome (EDS) and Bruck syndrome (BS) are not rare. EDS is an inherited connective tissue disorder with a variety of genotypes and phenotypes. EDS is characterized by hyperextensibility of skin, fragility of connective tissues and joint hypermobility. The classification of EDS suggests six major types. The arthroclasis type of EDS (types 7A, B) affects type I collagen. Mutations occur in N-propeptide cleavage region of *COL1A1* and *COL1A2* genes for EDS type 7A and 7B respectively. The pattern of inheritance is AD. Individuals reveal moderate fragility, skin hyperelasticity, congenital bilateral hip dislocation, muscular hypotonia etc. (Beighton *et al.,* 1998; Steinmann *et al.,* 1980). EDS and OI overlapping phenotypes arise if N-propeptide domains have not been removed and are integrated in the matrix, due to Gly mutations in first 90 residues of α1 chain (Makareeva *et al.,* 2008). Mutations in *PLOD3* gene also cause EDS/OI phenotype (Salo *et al.,* 2008).

Bruck syndrome is a genetic disorder which combines contractures and bone fragility. The condition also has genetic and phenotypic heterogeneity. BS and OI phenotype overlaps are caused by mutations in the *FKBP10* and *PLOD2* genes. OI and BS severity varies from mild to moderate (Puig-Hervás *et al.,* 2012; Kelley *et al.,* 2011).

Mutations in the *PLS3* (plastin 3) gene cause X-linked osteoporosis. The T-plastin protein takes part in forming the links between F-actin bundles. Through the assembly of cytoskeleton the protein influences mechanosensing of osteocytes and bone tissue homeostasis. The disease is connected with osteoporotic fractures and bone mineral density (BMD) decrease. Bone phenotypes overlap with mild OI (Van Dijk *et al.,* 2013; Fahiminiya *et al.,* 2014).

1.4 Genetic and molecular mechanisms of OI

Genetic background of the condition includes mutations in 16 different collagen and non-collagen genes responsible for osteogenesis processes. It explains the range of expressivity of OI phenotypes. About 60% (OI Type I, IV) – 100% (OI Type II, III) cases of OI arise as a result of *de novo* mutations (Shapiro *et al.,* 2013). Parental gonadal and somatic mosaicism occurs in 3-5% of OI cases. The frequencies of OI type occurrence vary depending on geographical location, and OI type and phenotype definition of a patient. The proportion of recessive cases may be higher in some populations due to founder effect, such as *LEPRE1* mutations in West Africans

and African Americans (Cabral *et al.,* 2012). Also, OI frequencies are higher in populations, where inbreeding and consanguineous unions are more common. To sum up, intrafamilial and interfamilial heterogeneity of the phenotype severity can be explained not only by affected the gene and a particular kind of mutation in it but also by the pattern of inheritance. It was also hypothesized that OI mutations tend to accumulate severity in subsequent generations (Moraes *et al.*, 2012). OI phenotype variability is also explained through incomplete penetrance. In some families, OI genotype may segregate phenotypic manifestations of OI. Such bias may occur due to differences in gene transcription, expression levels or the presence of modifying genes (Miko, 2008). Similar ectopic overexpressivity of a wild-type allele, due to variations in promoter regions, was reported for patients with Marfan syndrome (Li *et al.,* 2008).

1.4.1 *COL1A1* and *COL1A2* genes

OI Sillence classical types (I-IV) are caused by dominant or recessive mutations in collagen type I genes *COL1A1*and *COL1A2*. Mutations can lead to pathogenic biosynthesis of collagen and, as a result, collagen I structural dysfunction.

The *COL1A1* gene encodes for collagen type I α1 chain. It consists of 51 coding exons. OI mutations are identified in all exons, apart from exon 4. Exon 37 is supposed to be the richest of OI-associated mutations (Shapiro *et al.,* 2013). In addition to *Osteogenesis Imperfecta* types I-IV, mutations in *COL1A1* can lead to Ehlers-Danlos syndrome type VIIA, Ehlers-Danlos syndrome classical type, Caffey disease, idiopathic osteoporosis and *dermatofibrosarcoma protuberans* (Dalgliesh, 2008)[3].

The *COL1A2* gene encodes for collagen type I α2 chain. The gene is composed of 52 exons. No mutations are found in exons 1, 2, 5, and 10 but a lot of variants occur in exon 19 (Shapiro *et al.,* 2013). Besides OI types I-IV, mutations in *COL1A2* cause EDS VIIB, recessive EDS of the classical type, idiopathic osteoporosis, and atypical Marfan syndrome (Dalgliesh, 2008)[4]. Symptoms are less severe than in case of mutations in the *COL1A1* gene, due to the fact that only one α2 chain is incorporated into triple helix compared to two α1 chains (Dalgliesh, 2008)[4]. However, lethal mutations may appear in both genes.

[3] http://www.ncbi.nlm.nih.gov/gene/1277

[4] http://www.ncbi.nlm.nih.gov/gene/1278

1.4.1.1 Molecular mechanisms of OI mutations in collagen

Collagen gene mutations are divided into two types – quantitative and qualitative defects. In the case of quantitative defect, haploinsufficiency takes place. Collagen structure is normal but there is a lack of collagen in bone matrix. Only half of the normal amount of collagen is synthesized, because transcription of the mutated allele is interrupted with premature termination (PTC). Then, nonsense-mediated mRNA decay (NMD) is activated, and defective transcripts are degraded (Marini and Blissett, 2013). The rest of the matrix contains structurally normal collagen fibrils with proα1 and proα2 chains from a normal allele. A reduced number of collagen fibrils deposited in the ECM causes improper structure and integrity of ECM.

Dominant OI types II-IV result from mutations which alter collagen I quality. Collagen triple helix is extremely sensitive to alterations. The structure of collagen fibrils is changed due to single nucleotide substitutions in *COL1A1* and *COL1A2* genes. If glycine residues are substituted in proα1 or proα2 chains, the helical folding of collagen is postponed, and post-translational overmodification occurs. Phenotypes vary from mild forms of osteopenia to lethal forms.

1/5 and 1/3 of Gly substitutions in proα2 and proα1 chains, respectively, are lethal. Especially severe impact is caused by substitutions of Gly with branched nonpolar or charged amino acids (Glu, Arg, Asp, Val) (Marini *et al.*, 2007). The degree of severity of a disorder, based on substituted residues, is – the following – Ala ≤ Ser < Cys < Arg < Val < Glu ≤ Asp (Beck *et al.*, 2000). The more bulky amino acid incorporates into a collagen triple helix, the more unstructured collagen fibrils are. As a result, bones tend to be more fragile (Figure 6).

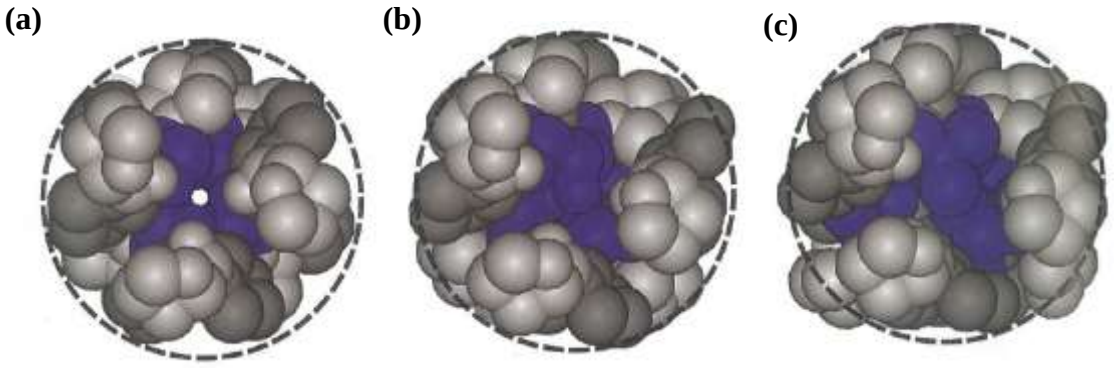

Figure 6 Collagen triple helix $(Pro\text{-}Pro\text{-}Gly)_{10}$ cross-section. Gly residues – dark blue, X residues – dark grey, Y residues – light grey. A circle illustrates normal diameter of the triple helix.
(a) Normal, (b) Gly substitution with Ser, (c) Gly substitution with Val. (Shapiro *et al.*, 2013).

Lethal Gly substitutions in proα1 chain are located in MLBR2 and MLBR3 (Makareeva *et al.,* 2008). Proα2 chain's lethal Gly substitutions are aligned with proteoglycan binding sites. Current data reveals the importance of collagen monomer interactions with NCPs – fibronectin, decorin, integrins, and matrix metalloproteinases (MMPs) (Marini *et al.,* 2007; Forlino *et al.,* 2011). Gly substitutions in N-proteinase cleavage sites can disrupt procollagen modification and cause mild OI and EDS.

Mutations in C-proteinase cleavage site do not interfere with posttranslational modification. Incorporated into fibrils uncleaved pC-collagen results in increased mineral density and high BMD OI (Lindahl *et al.,* 2011). Single nucleotide substitutions of X and Y positions in Gly-X-Y triplets cause significant frameshifts and prevent normal processing of propeptide, leading to combined OI/EDS condition. Substitutions of glycine at the same position may cause distinct forms of OI. Recently, p.Gly1088Glu and p.Gly1088Ala mutations were reported to cause OI type I and lethal OI respectively (Xia *et al.,* 2014).

Whole gene deletions of *COL1A1* result in haploinsufficiency and OI Type I phenotype. Multi-exon deletions of exons 23 to 25 (Barsh *et al.,* 1985; Chu *et al.,* 1985), 39 to 48 (Bodian *et al.,* 2009) cause OI Type II. Whole gene deletions of *COL1A2* are not identified. Deletions of exons 7 to 11, 39 to 40 cause OI Types I and II respectively (Mundlos *et al.,* 1996; Willing *et al.,* 1988). Collagen splice site mutations result in exon skipping, intronic retention, intron or exon cryptic splice site activation. Described frameshifts lead to PTC. Phenotypes develop mostly mild OI (Marini *et al.,* 2007).

1.4.2 Non-collagenous OI genes

Mutations in non-collagenous genes, apart from the *IFITM5* gene, respond to OI types with autosomal-recessive (AR) pattern of inheritance. OI is caused by defects of the mineralization (type V, VI OI; *IFITM5*, *SERPINF1* genes), collagen-3 hydroxylation defects (types VII-IX OI; *CRTAP, LEPRE1, PPIB* genes), collagen chaperones (types X, XI; *SERPINH1, FKBP10* genes), and C-propeptide cleavage enzyme (type XII, the *BMP1* gene). Defects of three more unclassified types of OI are connected with the *WNT1, TMEM38B, OSX*, and *CREB3L1* genes (Supplementary Figure 2).

1.4.2.1 The *IFITM5* gene

Interferon induced transmembrane protein 5 (*IFITM5*) gene encodes for a 14,8 kDa transmembrane protein bone-restricted Ifitm-like protein (BRIL) which is expressed specifically in skeleton. BRIL is important for collagen mineralization and is highly expressed in osteoblasts (Lazarus *et al*., 2014). Mineralization contribution can be direct through interactions with ECM or indirect via association with the membrane or intracellular mediators. BRIL connects directly to a *FKBP10* gene product - FK506 binding protein 10, which is known to take part in collagen folding via peptidyl-prolyl cis-trans isomerase activity. It is also possibly involved in the late stage of osteoblast maturation (Hanagata and Li X, 2011).

Previously all patients with OI type V were supposed to have the same heterozygous mutation in 5'UTR of the *IFITM5* gene. The mutation causes in-frame alternative initiation codon upstream of the annotated translation initiation codon. As a consequence, five amino acid residues (Met-Ala-Leu-Glu-Pro) are added to N-terminus of a protein (Semler *et al.,* 2012; Cho *et al.,* 2012). Recently, heterozygous missense mutation in a coding region of the *IFITM5* gene was also found (Hoyer-Kuhn *et al.,* 2014). Patients with the same mutations in the *IFITM5* gene develop distant phenotypes, even within one family. Generally, phenotypes are comparable to autosomal dominant non-collagen OI type V (Farber *et al.,* 2014).

1.4.2.2 The *SERPINF1* gene

The gene encodes for a serpin peptidase inhibitor, clade F (pigment epithelium-derived factor or PEDF). PEDF is a 50 kDa secreted glycoprotein, inhibitor of angiogenesis. Protein is highly expressed in osteoblasts, and is an important factor for bone homeostasis, matrix remodeling in particular (Becker *et al.,* 2011). Possible ways of PEDF control of bone homeostasis include – binding to ECM via collagen and glucoseaminoglycans binding motifs; regulation of intracellular signaling pathways, such as Notch and WNT; and modulation of nuclear signaling (Yasui *et al*, 2003; Hosomichi *et al.,* 2005). PEDF controls mineral deposition into ECM (Bogan *et al.,* 2013), and is supposed to inhibit osteoclast activity and bone resorption (Akiyama *et al.,* 2010).

Nonsense and frameshift mutations in this gene result in PTC and NMD of its transcripts (Becker *et al.,* 2011; Rauch *et al.,* 2012). Defects in the *SERPINF1* gene provide severe deforming OI type VI (Cho *et al.,* 2013; Rauch *et al.,* 2012). The exact molecular mechanism of the current OI type remains unknown. A patient with both recessive null-mutation in *SERPINF1* and non-

classical *IFITM5* mutation was reported. BRIL and PEDF expression differed from the expression levels in patients with *IFITM5* and *SERPINF1* mutations. As a result, mutual influence of PEDF and BRIL was predicted (Hoyer-Kuhn *et al.*, 2014).

1.4.2.3 *CRTAP*, *LEPRE1* and *PPIB* genes

Proline residues of collagen type I α1 and α2 chains (Pro986 of α1 and α2 chains and Pro707 of α2 chain) are modified by a heterotrimeric prolyl 3-hydroxylation complex, situated on endoplasmatic reticulum (ER).

The prolyl-lysyl-hydroxylase modification of procollagen provides intermolecular cross-links and stability of protein by attaching binding sites to carbohydrate units. This activity is crucial for collagen folding, protein-collagen interactions assembly, transport to extracellular matrix and, as a result, bone formation. Cofactors for the enzyme are iron and ascorbate. In addition to hydroxylase activity, the complex acts as a peptidyl-prolyl cis-trans isomerase (PPIase) and collagen chaperone (Forlino *et al.*, 2011). The activity of the complex is specific to Pro position and collagen substrate. The complex is composed of cartilage-associated protein, prolyl-3-hydroxylase and cyclophilin B proteins, coded by *CRTAP*, *LEPRE1*, and *PPIB* genes respectively. Mutations in this complex cause some of the most severe phenotypes of OI.

The *CRTAP* gene encodes for a cartilage-associated protein (CRTAP) – a helper protein of the complex. It is highly expressed in chondrocytes of growth plates (Morello *et al.*, 2006). The scaffolding protein controls the activity of proteins of the cytohesin/ARNO family in response to cell stimuli. CRTAP protein shows dependent costabilization with the product of the *LEPRE1* gene (leucine proline-enriched proteoglycan 1). Mutations in either of these genes lead to the absence or reduced amounts of CRTAP and prolyl 3-hydroxylase 1 (P3H1) (Chang *et al.*, 2010). The majority of mutations were identified in the first or fourth exon and surrounding introns. The defects of *CRTAP* cause NMD of the transcripts and deficiency of a CRTAP protein, where 3-hydroxylation complex activity is lost. Pro986 of procollagen α1 chain stays unhydroxylated. The collagen helix is overmodified by lysyl hydroxylase and prolyl-4-hydroxylase instead. The folding of collagen is delayed, as in the case of defects in C-termini of procollagen (Barnes *et al.*, 2006). Thus, phenotypes of *CRTAP* deficient OI type VII overlap with Sillence classical OI types II and III.

LEPRE1 gene product is a leprecan (P3H1). Leprecan provides prolyl 3-hydroxylase enzymatic function. The P3H1 protein has the same N-terminal domain as a CRTAP protein does. It is also

known as a tumor suppressor factor and signal molecule, providing intracellular communication (Cabral *et al.,* 2007, Van Dijk *et al.,* 2011). Homozygous and compound heterozygous mutations of the *LEPRE1* gene usually occur in exons 13, 15 and 5, and result in leprecan deficient cells. Splice-site mutations in the fifth exon bring alternative spliced transcripts with PTC (Cabral *et al.*, 2007). *LEPRE*-null fibroblasts produce up to 50% more collagen. Collagen is over-modified and lacks 3-prolyl-hydroxylation of Pro986 in proα chain, as in the cases of *CRTAP* deficiency. Null-alleles of *LEPRE1* result in connective tissue disorders, equivalent to *Osteogenesis Imperfecta* type VIII, phenotypically similar to OI type VII (II, III).

The *PPIB* (also known as a *CYPB*) gene codes for peptidyl-prolyl cis-trans isomerase B (cyclophilin B, CyPB). CyPB is the third component of the hydroxylation complex and contains cyclophilin-like PPIase domain. Cis-trans isomerase activity of CyPB changes the configuration of cis to trans-conformers, since only prolyl-containing trans-peptide bonds can be included into collagen triple helix (Van Dijk *et al.,* 2009; Pyott *et al.,* 2011). CyPB affects the folding of the Pro-rich regions of procollagen's C-termini, procollagen export and secretion with HSP47 protein. Deficient CyPB amount leads to inappropriate collagen helical folding, post-translational over modification and 3-hydroxyprolyl deficiency, forming OI type IX. Phenotype overlaps with phenotypes of OI types VII and VIII (II, III).

1.4.2.4 *FKBP10* and *SERPINH1* genes

Procollagen type I folding into triple helix is assisted by chaperones a 65 kDa FK506 binding protein 10 (FKBP65) and heat shock protein 47 (HSP47). These chaperons form ER resident complex.

HSP47 is coded by the *SERPINH1* gene (serpin peptidase inhibitor, clade H, member 1); it is a procollagen molecular chaperon. The protein prevents immature collagen folding. Due to defects of the *SERPINH1* gene, the distribution of procollagen to endoplasmic reticulum does not occur. Improper helical folding of collagen results in OI type X (severe progressive OI).

The *FKBP10* gene encodes for FKBP65. Protein can be secreted or binded to ER membrane. FKBP65 takes part in protein quality control and folding (Kelley *et al.,* 2011). FKBP65 takes part in lysyl hydroxylation by stabilizing the lysyl hydroxylase 2 (LH2) enzyme or promoting interaction between LH2 and substrate through PPIase activity (Barnes *et al.,* 2012; Schwarze *et al.,* 2013). Mutations in *FKBP10* gene result in the reduced amount of hydroxylated lysyl residues of procollagen in bone matrix as well as reduced collagen amount in ECM due to

reduced collagen cross-linking (Barnes *et al.,* 2012). Apart from OI type XI, phenotype defects of this gene can show contractures without OI (Yup'ik syndrome) and contractures combined with OI (Bruck syndrome).

1.4.2.5 The *BMP1* gene

The product of the *BMP1* gene – bone morphogenetic protein 1 (BMP1) – participates in bone and cartilage formation. BMP1 has peptidase activity and cleaves C-terminal propeptides of procollagen I, II and III. The enzyme also proteolytically activates lysyl oxydase. BMP1 increases the expression of collagen type 1 and osteocalcin in osteoblasts, promoting tissue mineralization and acting as a regulator of bone repair. BMP1-3 isoforms circulate in blood and can be used as bone biomarkers (Martinez-Glez *et al.,* 2012; Asharani *et al.,* 2012, Grgurevic *et al.,* 2011).

Mutations in the *BMP1* gene can alter catalytic peptidase domain or BMP1 signal peptide. Phenotypes are close to severe AR OI and are distinguished as an OI type XII with increased mineral density and recurrent fractures, as in the case of C-termini cleavage defects of procollagen (Asharani *et al.,* 2012). However, BMP1 protein has a range of substrates, and phenotype development of the condition may be connected with other molecular mechanisms.

1.4.2.6 The *OSX* gene

The *OSX/SP7* gene encodes osterix (OSX) – a zinc-finger transcription factor. Protein is bone-specific and connected to osteoblast differentiation and bone development. It contains N-terminal trans-activation motifs and three Cys2-His2 C-termini zinc finger DNA-binding domains. The described AR OI mutations cause deletion of one of the zinc finger domains. Detailed pathogenesis of an osterix-defficient OI type XIII remains unknown. *OSX*/*SP7* OI mutations are very rare, which are mainly deleterious mutations incompatible with life (Lapunzina *et al.,* 2010).

1.4.2.7 The *TMEM38B* gene

The *TMEM38B* gene (a transmembrane protein 38B) encodes a TRIC-B, ubiquitous component of the monovalent cation-specific transmembrane channel (TRIC), responsible for the release of calcium cations from intracellular stores, such as ER and sarcoplasmatic reticulum. Mutations in

this gene can lead to autosomal recessive OI type XIV (non-syndromic form of OI) (Volodarsky *et al.,* 2013; Shaheen *et al.,* 2012).

1.4.2.8 The *WNT1* gene

The *WNT1* (Wingless-type MMTV integration site family, member 1 (*INT1, BMND16*)) gene encodes for secreted signaling protein. The *WNT1* gene participates in body axis patterning, cell differentiation, migration, and proliferation. WNT1 activates canonical low density lipoprotein receptor-related protein 5 (LRP5) – mediated WNT/β-catenin signaling pathway which regulates gene expression in osteoblasts. As a result, the gene is supposed to influence bone development (Keupp *et al.,* 2013; Laine *et al.*, 2013).

Heterozygous and variant homozygous mutations in *WNT1* gene can induce bone fragility – age-related osteoporosis and severe recessive forms of OI indicated as OI type XV respectively (Pyott *et al.,* 2013; Fahiminiya *et al.*, 2013; Keupp *et al.,* 2013; Laine *et al.*, 2013). Patients with mutations in *WNT1* do not respond to bisphosphonate therapy. Bisphosphonates inhibit osteoclast activity and bone resorption, whereas *WNT1* mutations cause dysfunction of osteoblasts (Keupp *et al.,* 2013).

1.4.2.9 The *CREB3L1* gene

The *CREB3L1* (cAMP responsive element binding protein 3-like 1) gene encodes for ER-stress transducer protein OASIS (Old Astrocyte Specifically Induced Substance). During ER-stress protein activates the unfolded protein response target genes (Omori *et al.,* 2002). Mutations in this gene cause severe osteopenia, spontaneous fractures, as in severe recessive OI. Protein regulates tissue-specific gene transcription. Protein is also supposed to influence procollagen synthesis by osteoblasts (Symoens *et al.,* 2013).

1.4.3 Genetic diagnosis of OI

Clinical diagnosis of OI should be confirmed by genetic diagnosis, since phenotypes of different disorders tend to overlap. Firstly, the *COL1A1* and *COL1A2* genes are screened for mutations. As mentioned above, about 90% of OI cases are caused by mutations in collagen genes. If no causative mutations are detected, the sequencing of non-collagen genes, included into the recessive OI panel, occurs. All new causative variants are checked with functional analysis. For

this purpose, collagen expression analysis with mRNA/complementary DNA (cDNA) and proteins from cultured skin fibroblasts is done. It may reveal abnormalities in procollagen structure or expression. If no mutations in the known genes are found, the patient may have a mutation in an undiscovered OI gene (Van Dijk *et al.,* 2012) (Figure 7).

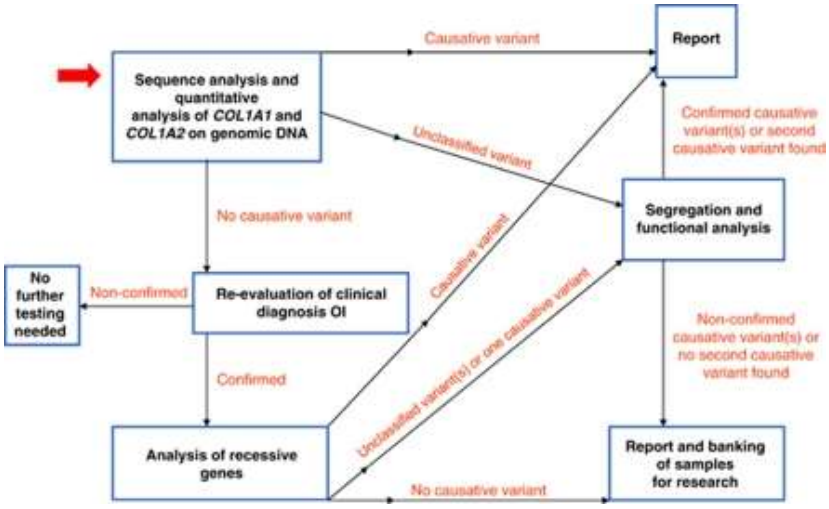

Figure 7 The workflow of genetic diagnostics for an OI patient (Van Dijk *et al.,* 2012).

1.5 Exome sequencing as a tool for Mendelian disorder gene discovery

Genetic causes of more than 3,000 single-gene disorders with known phenotypes (1/5 of OMIM database queries) remain unidentified. However, after next-generation sequencing (NGS) approaches became available, more than 100 genes causing Mendelian disorders were revealed (Rabbani *et al.,* 2012). The majority of discovered mutations alter the coding regions of genes, which makes sequencing of the exon regions a constructive approach for the discovery of new genes and mutations connected with Mendelian disorders (Bamshad *et al.,* 2011; Ku *et al.,* 2011). Such selective sequencing of protein-coding regions of the genome is provided by exome sequencing studies. It is a rapid, high-throughput, sensitive and a cost-effective strategy of gene discovery for Mendelian disorders, compared to Whole Genome Sequencing (WGS).

The estimated number of exons in the human genome is about 180,000, with the length equal to 27.9 Mb, composing about 1% of the whole human genome (Ng *et al.,* 2009). In addition to exons, exome sequencing kits also include flanking intron regions, UTR regions, promoters, miRNA genes and non-coding RNAs, which may affect gene regulation and functioning

(Bamshad *et al.,* 2011). Annotated and unknown regions remain uncovered because exome sequencing methods are mostly based on sequence information from consensus coding sequence (CCDS) and RefSeq databases (Ku *et al.,* 2011). The analysis of copy number variations (CNV) and other structural variations is limited as well (Gilissen *et al.,* 2011). Some regulatory regions and evolutionary conserved motifs miss covering as well. It does not affect the method's sensitivity as non-synonymous variants have weak or absent effect in non-coding regions. At the same time, pathogenic missense or nonsense substitutions and indels in exons are more important as they cause fitness decrease and are deleterious. SureSelect Human All Exon kit V5 (Agilent Technologies Inc.) includes 70 Mb of human genome. The reduction of sequence volumes allows to deepen sequencing coverage up to 60-65x and increase the method's sensitivity (Ng *et al.,* 2010).

Previously, OI was supposed to arise due to mutations in collagen type I (*COL1A1* and *COL1A2*) genes. 14 other OI genes were recently revealed with Sanger sequencing, real-time quantitative polymerase chain reaction (RT-qPCR), linkage analysis, homozygosity analysis, and NGS. In some OI patients, mutations in known OI genes are absent, which remains a great interest to investigators. Traditional methods, such as linkage analysis, lose effectiveness if they have to deal with the limited number of affected individuals, *de novo* mutations, unrelated patients and phenotypic heterogeneity (Ku *et al.,* 2011), which are common features for OI disorders (Roughly *et al.,* 2003). In this case, exome sequencing represents the most effective method for the discovery of new OI genes and mutations. A few novel OI genes have been discovered with the assistance of exome sequencing. Among these OI genes are *SERPINF1* (Becker *et al.,* 2011), *WNT1* (Pyott *et al.,* 2013), *IFITM5* (Cho *et al.,* 2012), and *PLOD2* with BS/OI overlap phenotype (Puig-Hervás *et al.,* 2012).

1.5.1 Strategy of exome sequencing and detecting disease-causing variants

The algorithm of exome sequencing is as follows:

Firstly, the DNA is randomly cleaved, and a shotgun *in vitro* library is formed. Then, DNA segments are flanked with adaptors. Probes are hybridized with biotinylated DNAs or RNAs. Biotin-streptavidin connections allow the capturing of probes to array for exome enrichment. Then, amplification and massively parallel sequencing take place. Finally, potential causal variants are detected and mapped (Bamshad *et al.,* 2011). Bar-coding, which is done during

library construction or amplification, allows sequencing of a multiple number of samples at the same time (up to 16) (Ku *et al.*, 2011).

Data is processed in order to find rare variants, shared between unrelated or closely related patients. There are four basic strategies of gene discovery. The first one is based on the sequencing and filtering of unrelated affected individuals. Exome analysis identifies the variants common for all affected patients. Polymorphisms that are found in healthy controls or public variant databases (dbSNP, 1000 Genomes Project) are excluded (Figure 8a). Another way is to sequence and filter affected individuals from a single family. If mapping data is not available, sequencing of the two most distantly related affected individuals will be best. If mapping data is accessible, sequenced individuals have to share the smallest single haplotype (Figure 8b). Parents-child trio analysis is successful for detecting *de novo* mutations. Affected offspring will have unique variant, which was not inherited from parents but arose as a result of spontaneous mutation (Figure 8c).

Extreme phenotype approach is based on the principle of fitness. Therefore, affected individuals will concentrate in one extreme of phenotype distribution (Figure 8d). In case of rare recessive disorders, filtering mostly uses minor allele frequency (MAF) below 1%. Moreover, to promote variant selection, alleles can be classified by functional consequence and biological function (Bamshad *et al.*, 2011; Ku *et al.*, 2011).

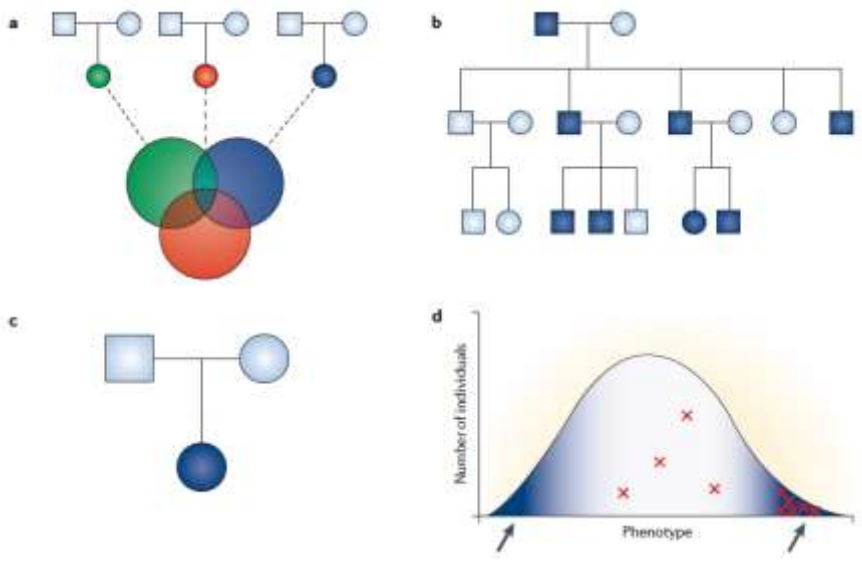

Figure 8 Basic gene discovery strategies in exome sequencing analysis. (a) Sequencing of unrelated affected individuals; (b) Sequencing of multiple affected individuals from a single pedigree; (c) Sequencing of a parent-child trio; (d) Comparing phenotype distribution extremes (pointed at with an arrow red crosses represent affected phenotypes) (Bamshad *et al.*, 2011).

1.6 OI in Estonia

Since 1995 the center for clinical follow-up, treatment and research of OI in Estonia is the Clinic of Traumatology and Orthopaedics of Tartu University Hospital. During this period, patients from 32 families from different counties of Estonia are treated and investigated in Tartu University Hospital. Taking into account that prevalence of OI is 1 per 20,000 to 30,000 it is close to 100% of patients in Estonia.

The aim is to help people with *Osteogenesis Imperfecta* with medical and surgical treatment, to improve quality of life of OI patients, and provide support for the families. At the same time, with clinical follow-up the genealogical information, clinical and medical history and phenotype description of OI patients and family members are collected to OI database for a future research. Additionally, blood samples for bone markers, DNA and RNA analysis are collected from all available family members. (Maasalu *et al.*, 2003).

2. EXPERIMENTAL PART

2.1 Aim of the study

The main aim of this study is to perform exome sequencing analysis in three selected Estonian families with OI history in order to identify causative mutations of the disorder.

The study includes:

1. Description of genealogical information of the selected families with OI history.
2. Overview of phenotype manifestations of the affected OI family members.
3. Identification of OI causing mutations with exome sequencing.

2.2 Methods and materials

2.2.1 Material collection and selection of families

The OI patient database of the Department of Traumatology and Orthopaedics, Tartu University Hospital (Estonia) was used for selection of families for current study[5] The study was approved by the Ethics Review Committee on Human Research of the University of Tartu. The participants included in the study or their legal representatives gave prior consent to take part in the study.

OI families with unusual clinical features in phenotypes, inheritance patterns or clinical history were selected for priority exome analysis. On the basis of the selection criteria three families were selected (2, 7 and 10) with cases of severe OI (Figures 9-11). Families 2 and 10 were outstanding with extreme intrafamilial variety in phenotype severity. Family 7 represented the case of *de novo* mutation with phenotype of OI type II "survivor". Every family was provided with genealogical information, clinical and medical history and phenotype description of the family members. Blood samples were collected from all available family members.

---

[5] http://www.kliinikum.ee/ortopeedia/

2.2.1.1 Registration of genealogical information and OI phenotype description

The genealogical information requested included OI history in the family and family consanguinity data. All healthy and affected individuals of the family, miscarriages and early death cases were reported. Pedigree trees were created with a Progeny Free Online Pedigree Tool (by Progeny Biosoftware, LLC; USA).[6] Phenotypes and clinical history of the families were described previously by Dr. Katre Maasalu.

Phenotype description was based on the following information:

- Presence of intrauterine fractures
- Deviations of height and weight, skeletal deformities, presence of fractures, eye sclera color after birth
- Current height, weight and skeletal deformities (including scoliosis)
- Age of the first fracture and number of fractures for current time, description of fractures
- Walking and standing ability
- Head and face shape
- Eye sclera color
- Presence of hearing loss and age when it started
- Presence of *Dentinogenesis Imperfecta*
- Additional specific features of the patient (presence of other hereditary disorders and health problems etc)
- Medical history of both parents

2.2.1.2 Sample collection and DNA extraction

Blood samples were collected by employees of Tartu University Hospital. Genomic DNA (gDNA) was extracted from peripheral white blood cells with standard high-salt extraction methods and stored at −80°C until analyzing.

2.2.1 Exome sequencing

Whole-exome sequencing was performed from gDNA by the Sequencing Core Laboratory of the Estonian Genome Center, University of Tartu (Estonia). Exome capture was performed using the TruSeq Exome Enrichment kit (Illumina) following the manufacturer's protocol.

[6] http://www.progenygenetics.com/online-pedigree/

DNA fragments were sequenced with Illumina HiSeq2000 with 100-bp paired-end reads. Sequence reads were aligned to the human reference genome (hg19, GRCh37) with the Burrows-Wheeler Aligner (BWA, version 0.6.1) (Li *et al.,* 2009). Exome analysis was performed in probands 219, 221, 234, 228 of family 2; individuals 710, 713, 716 of family 7; and probands 1009, 1013, 1019 of family 10 (Figures 9-11).

2.2.1.1 Bioinformatic analysis and variant prioritization

All bioinformatics analysis procedures, including variant prioritization, were conducted by researchers at the Estonian Genome Center (University of Tartu) as previously described (Nikopensius *et al.,* 2013). Single-nucleotide polymorphisms (SNPs) and small insertion-deletion (indels) variants were called with SAMtools (version 0.1.18), Picard tools (version 1.60), and a Genome Analysis Toolkit (GATK, version 1.5.21) (McKenna *et al.,* 2010; Li *et al.,* 2009). The main points of interest were non-synonymous mutations, insertions-deletions and variants in canonical splicesites, absent from public databases (dbSNP135 and the 1000 Genomes Project). PolyPhen-2, SIFT, Alamut and Condel software tools were used to predict the functional effects of mutations (Kumar *et al.,* 2009; Adzhubei *et al.,* 2010; Gonzalez-Perez *et al.,* 2011).

2.2.1.2 Mutation analysis with Sanger sequencing

The validation of a novel pathogenic variant, discovered in family 7 with exome sequencing, was performed with Sanger sequencing. Primers for Sanger sequencing were designed with online tool.[7] Primer sequence quality was checked with UCSC browser (UCSC Genome Bioinformatics Group).[8] Designed primer sequences (T_m 60°C, length 22 nt, GC % 40-60%) for *COL1A1* variant were as follows:

- forward primer: 5'GAAACCCAGACACAAGCAGAAC3'
- reverse primer: 5'TAGTAGATGACCCCAGGAGAGC3'

[7] http://bioinfo.ut.ee/primer3-0.4.0/primer3/

[8] http://genome.ucsc.edu/

PCR reaction was performed with a PTC-200 Peltier Thermal Cycler (MJ Research, USA) PCR machine. The PCR *touchdown* program, used for the reaction of amplification, was as follows:

1=95,0°; 15:00 min
2=95,0°; 0:25 min
3=64,0°; 0:30 min
4=72,0°; 0:40 min
5=Go to 2, 4 times
6=95,0°; 0:25 min
7=62,0°; 0:30 min
8=72,0°; 0:40 min
9=Go to 6, 30 times
10=72,0°; 5:00 min
11=6,0° forever

PCR reaction mix was performed in a 19 µl volume which contained 10× reaction buffer B (Solis BioDyne, Estonia), 0,2 mM of dNTPs (Fermentas, Lethuania), 25 mM $MgCl_2$ (Solis BioDyne, Estonia), 10 pmol of *forward* and 10 pmol of *reverse* praimers, 1 U Taq HOT FIREPol® DNA polymerase (Solis BioDyne, Estonia), 50 ng of gDNA and Milli-Q water to the final volume of a PCR mix. Products were separated on 1,5% agarose gel electrophoresis with 0,5× TBE buffer.

PCR products were purified, precipitated and sequenced in the Sequencing Core Laboratory of the Estonian Biocentre (Tartu, Estonia). Sanger sequencing was performed with Applied Biosystems 3730xl DNA Analyzer.[9] Sequence reads were analyzed visually with BioEdit program, version 7.0.9.0.[10]

[9] http://www.lifetechnologies.com/ee/en/home.html

[10] http://www.mbio.ncsu.edu/bioedit/bioedit.html

2.3 Results

2.3.1 Genealogical information and phenotype description of the selected OI families

All families included in the study are non-consanguineous families of Estonian origin. Families 2 and 10 have history of OI segregating with AD inheritance pattern, whereas family 7 was analyzed as a case-parents trio to search for *de novo* mutation (Figures 9-11).

2.3.1.1 Family 2

Family was choosed for study due to an intrafamilial variety in phenotype severity. Probands 219, 221, 234, and 228 were selected for exome sequencing using the strategy of multiple affected individuals in a single pedigree (Figure 9). The majority of family members developed mild bone fragility, which started in the childhood. Secondary features of OI phenotype included hearing loss and blue eye sclera, normal dentition. Height is normal. The number of fractures varied from <10 (individual 228) up to 20 (individuals 221, 234). The clinical signs are typical to OI type I. Proband 219 revealed distant phenotype from the rest of the affected individuals. The OI phenotype differs in its severity. The number of fractures is considerably higher, major skeletal deformations, reduced height, severe osteoporosis and triangular face are present. These findings are typical to type III of OI.

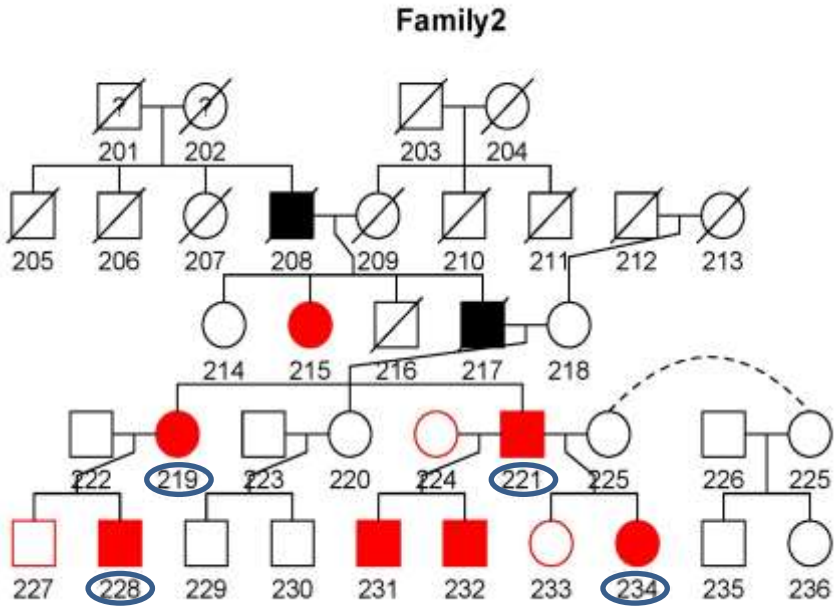

Figure 9 The pedigree tree of family 2 with AD inheritance pattern of OI.

2.3.1.2 Family 10

The case represents the diversity of OI forms within a single family (Figure 10). All affected family members are females, they have blue sclera, but DI and hearing loss are not present. The individual 1009 developed severe OI with plenty of fractures, long bones deformations, scoliosis, body disproportions, and significantly reduced height. The individual was clinically diagnosed with OI type III. The proband 1013 has mild OI with normal height. The number of fractures is up to 10. Individuals 1017 and 1019 have moderate OI with reduced height, and reoccurred fractures. The phenotype is typical to OI type IV. Exome sequencing was performed in probands 1009, 1013 and 1019 using the strategy of multiple affected individuals in a single pedigree.

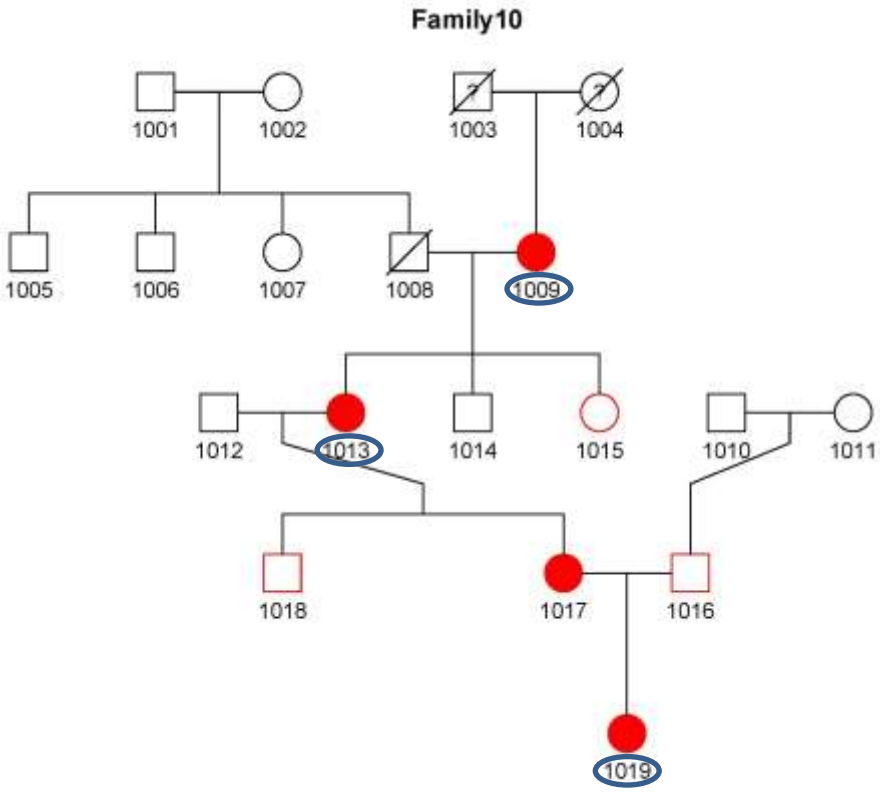

Figure 10 AD inheritance pattern of OI in family 10.

2.3.1.3 Family 7

The case represents arising of a new mutation in the otherwise healthy family. Proband 716 was born to healthy parents (Figure 11). Ultrasound discovered delayed intrauterine development of the skeleton. The child was born with multiple fractures, deformations of upper and lower limbs, soft strained skull, and severe respiratory failure. The presence of extreme osteopenia, and accordion-like ribs was confirmed. The proband was clinically diagnosed with OI type II. Due to extreme care, the proband survived, and at the age of 14, the height of the proband was 80 cm, and the weight 15 kilos. Mental development was normal. The patient died from respiratory insufficiency at the age of 14.5 years. Exome sequencing was performed in individuals 710, 713, and 716 using the strategy of a parents-child trio.

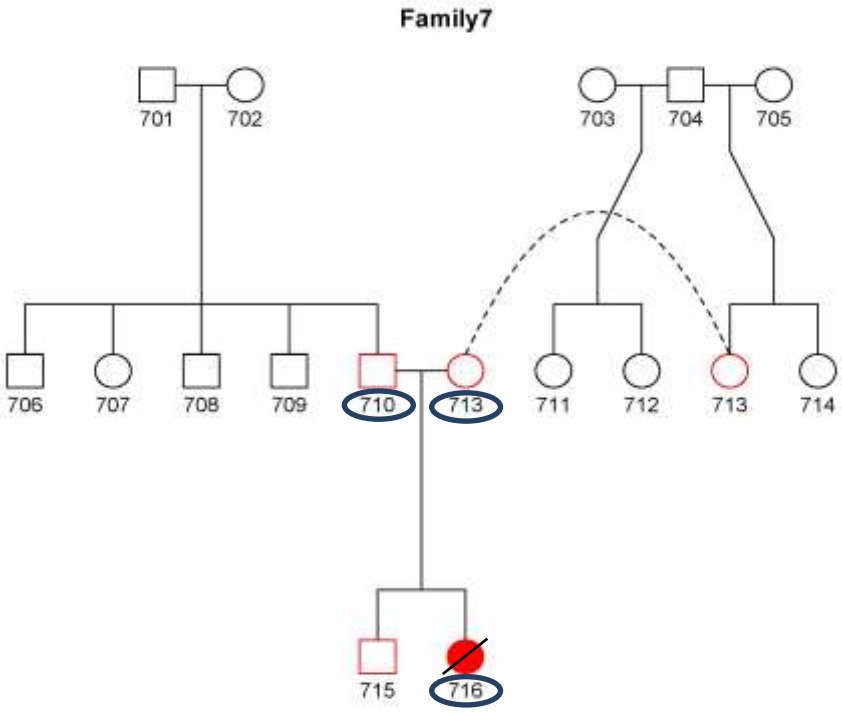

Figure 11 The pedigree tree of OI family 7.

2.3.3 Genetic studies

2.3.3.1 Exome sequencing in family 2

During exome sequencing, heterozygous mutation c.1821+1G>A in intron-exon junction 26 of the *COL1A1* gene (rs66555264 in dbSNP) was discovered in probands 219, 221, 234, 228 (Figure 12). The mutation causes the loss of canonical splicing donor site. The c.1821 position is highly conserved (phyloP score 5,165) (Table 4).

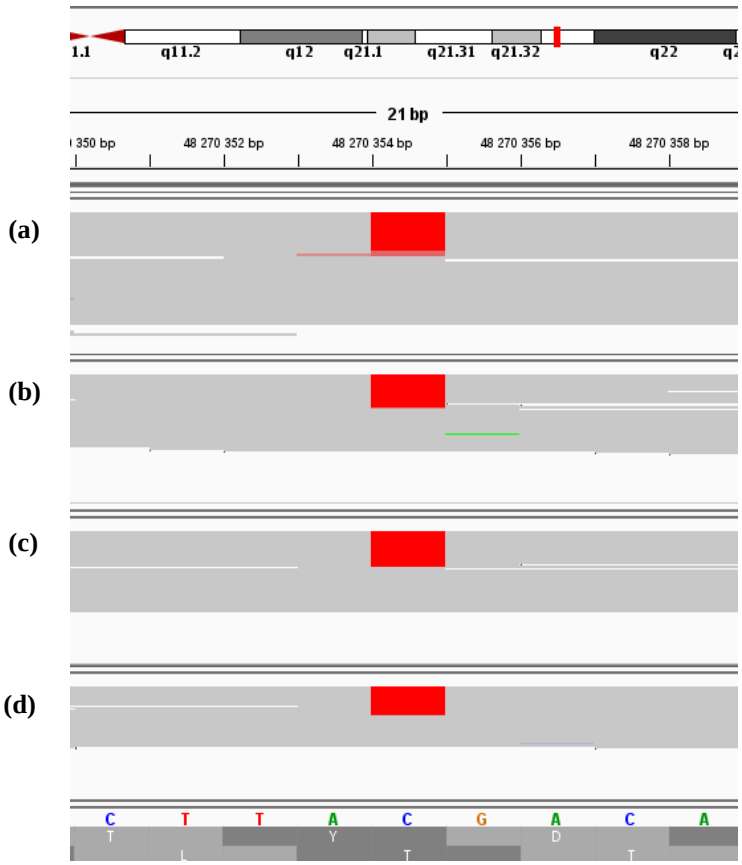

Figure 12 Integrative Genomics Viewer (IGV) image of chromosome 17 region: c.1821+1G>A transition in the *COL1A1* gene in sequenced probands of family 2 (a) 219, (b) 221, (c) 234, (d) 228. Figure represents the sequence of the sense DNA strand. Red squares show C>T transition (complementary to G>A on the antisense strand). Approximately half of the reads contain mutated allele (red), indicating the heterozygous state of mutation.

2.3.3.2 Exome sequencing in family 10

Heterozygous frameshift mutation c.1128delT (referred to as p.Gly377AlafsX164 in OI database) was found in exon 17 of the *COL1A1* gene of probands 1009, 1013, 1019 from family 10 (rs72645370 in dbSNP) (Figure 13). The frameshift caused by the deletion of T in position 1128 results in the synonymous substitution of Pro376 residue. However, the appeared frameshift alters the following sequence, starting from Gly377, substituted with Ala residue. The shift frame length composes 164 bp, including stop codon (Table 4). Gly377 residue and the following seven amino acids are highly conserved from human up to zebrafish (Table 3).

Table 3 The conservation of Gly377 residue in different species

	Gly377 ↓						
Human	G	P	A	G	A	A	G	P
Chimp	G	P	A	G	A	A	G	P
Macaque	G	P	A	G	A	A	G	P
Rat	G	P	A	G	A	A	G	P
Mouse	G	P	A	G	A	A	G	P
Dog	G	P	A	G	A	A	G	P
Cat	G	P	A	G	A	A	G	P
Opossum	G	P	A	G	A	A	G	P
Frog	G	Q	A	G	A	A	G	P
Zebrafish	G	P	A	G	A	A	G	P

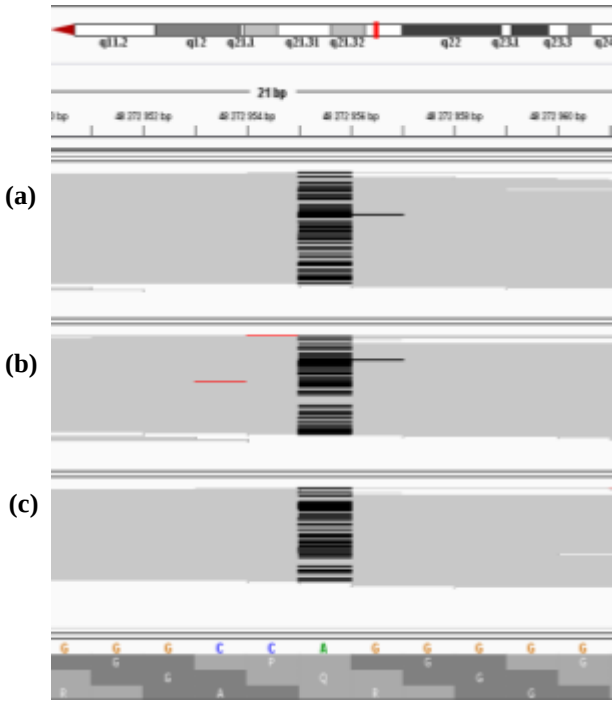

Figure 13 The IGV image of chromosome 17: c.1128delT (p.Gly377AlafsX164) frameshift mutation in the *COL1A1* gene in sequenced probands from family 10 (a) 1009, (b) 1013, (c) 1019. The sequence of the sense DNA strand is presented above. Black stripes represent deletion of an A nucleotide at position 1128 (complementary to T nucleotide for the antisense strand). The mutation is in heterozygous state.

2.3.3.3 Exome sequencing in family 7

Heterozygous missense c.2317G>T (p.Gly773Cys) mutation was found in exon 33/34 of the *COL1A1* gene in the affected proband 716 from family 7. The mutation was absent in healthy parents and a healthy brother of the affected proband. The presence of mutation was confirmed with Sanger sequencing (Figure 14). Codon 773 is highly conserved between species (phyloP score 5,418) (Table 4). The deleterious effect of the mutation was confirmed *in silico* (Condel score 0,945, SIFT score 0, PolyPhen-2 score 1).

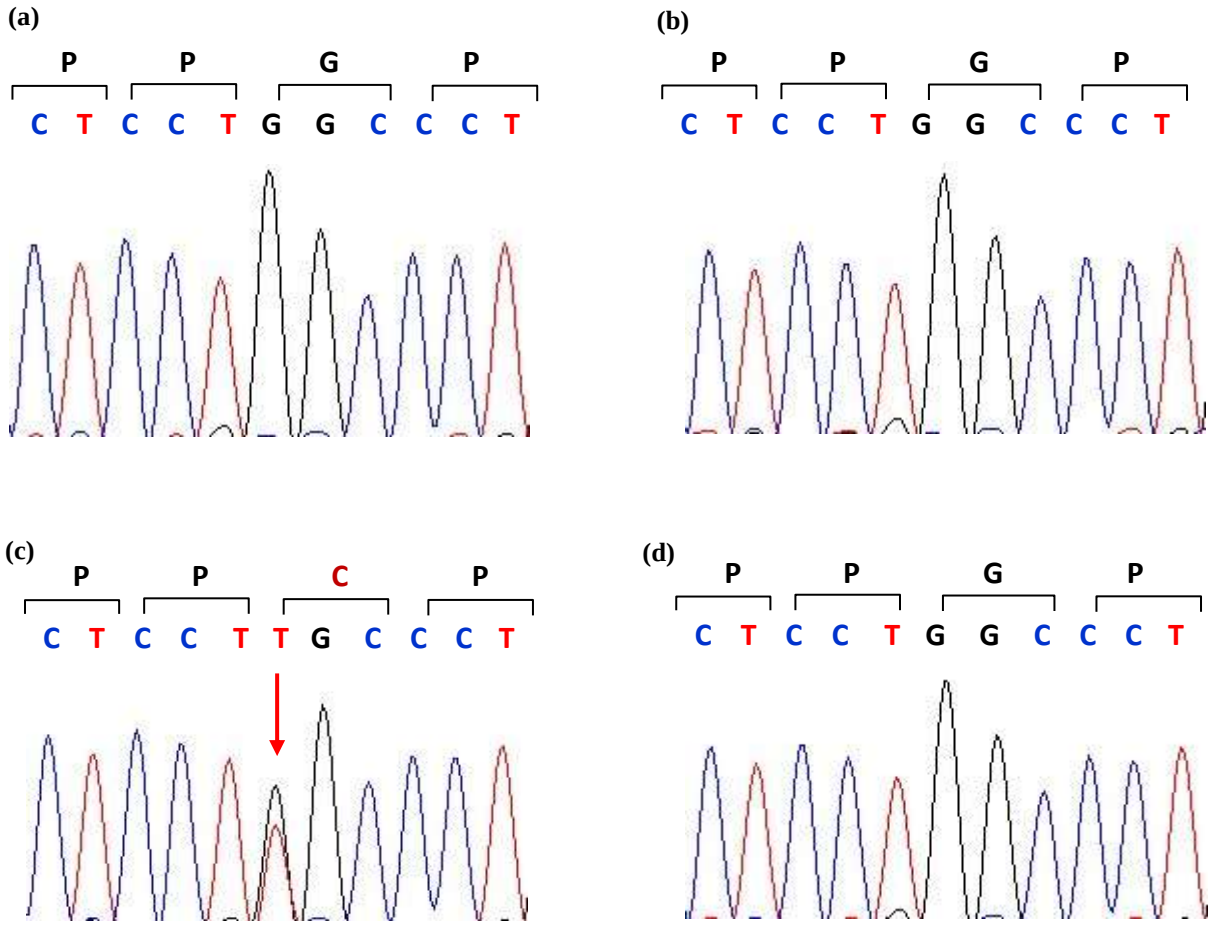

Figure 14 Validation of the presence of a c.2317G>T mutation in the *COL1A1* gene by Sanger sequencing.

Reference allele c.2317G in healthy family members 710 (a), 713 (b), 715 (d).

Heterozygous c.2317G >T transversion in proband 716 (c) leads to p.Gly773Cys substitution.

Table 4. Mutations, discovered with exome sequencing in three Estonian OI families

	Family 2	**Family 10**	**Family 7**
Gene	*COL1A1*	*COL1A1*	*COL1A1*
Mutation	c.1821+1G>A	c.1128delT	c.2317G>T
Allelic state	heterozygous	heterozygous	heterozygous
Function	splice site loss	frameshift	missense
Amino acid substitution	-	p.Gly377AlafsX164	p.Gly773Cys
dbSNP rs #	rs66555264	rs72645370	-
phyloP score	5,165	5,753	5,418
Pattern of inheritance	AD	AD	*de novo*

2.3.3.4 Epidemiological studies

All the mutations detected were absent from the Exome Variant Server of the NHLBI-ESP database and the 1000 Genomes Project database. Also, 89 Estonian control exomes and 87 Estonian full genomes were free from the presence of the pathogenic variants discovered.

2.4 Discussion

2.4.1 *COL1A1* splice site mutation c.1821+1G>A causes mild OI

Pathogenic c.1821+1G>A variant was discovered with exome sequencing in the *COL1A1* gene. The mutation alters consensus sequence GT at the 5' canonic splicing donor site of intron 26. The variant has been previously described 12 times in OI database (OI types I, IV) (Marini *et al.*, 2007; Qin *et al.*, 2005; Johnson *et al.*, 2000; etc).[2] Two of the reported cases had OI type IV. The rest of the individuals revealed mild OI phenotype matched with OI type I. Our study reported c.1821+1G>A mutation in patients from a single family with OI type I and IV.

The mutation causes retention of intron 26, the frame-shift and origin of the down-stream stop codons. Due to splice site mutations, intron retention and exon skipping depend on the length of

an intron. If an intron is short and has a mutated splicing site with stop codons downstream not further than 300 bp, as in the case of intron 26, U1RNA will consider the intronic region as a possible exon sequence and include it into the transcript. With mutations in splicing site of longer introns, exon skipping will occur (Robberson *et al.,* 1990; Stamm *et al.,* 2000).

Previous studies showed the absence of transcripts of mutated allele in the cytoplasm of skin fibroblasts of patients with OI type I caused by c.1821+1G>A mutation (Stover *et al.,* 1993; Johnson *et al.,* 2000). The amount of mutated collagen I α1 mRNA in the nucleus was normal with longer half-life of altered mRNAs. Longer half-life could also be connected to the change in stability of mutated transcripts through the presence of additional Cis-elements in a retained intron (Galante *et al.,* 2004).

The absence of mutated mRNA in the cytoplasm was not caused by NMD but was a consequence of defective mRNA export (Johnson *et al.,* 2000). Pre-mRNA undergoes a checkpoint of packaging and screening in SC-35 domain of the nucleus before transporting to the cytoplasm. The retention of introns into mRNA after splicing might prevent mRNA from leaving the SC-35 domain (Shopland *et al.,* 2002; Hall *et al.,* 2006). As a result, only half of collagen I α1 is translated into protein, which causes collagen quantitative deficiency and mild OI.

The absence of mutated collagen type I α1 transcripts in the cell cytoplasm is the key difference between mild and severe OI. As described above, the variety in phenotype forms between carriers of the same mutation might come from variants in the *COL1A1* promoters, which could change the regulation of wild-type allele expression and influence the amount of mRNA and translated protein in the cell. Mechanisms of mRNA export from the nucleus and NMD could also input into phenotype diversity.

Apparently, phenotype differences between proband 219 and the rest of the affected family 2 members might be caused by deviations in the *COL1A1* gene expression or mRNA transport mechanisms. To clearify the effect of the mutation in the current family and the cause of an intrafamiliar variety, further collagen expression and mRNA splicing, metabolizing and transporting analysis are required.

2.4.2 Frameshift mutation c.1128delT (p.Gly377AlafsX164) in *COL1A1* may cause OI of different phenotypes

Exome sequencing in three probands of family 10 discovered c.1128delT (p.Gly377AlafsX164) frameshift mutation in exon 17 (triple helical domain of collagen I α1 chain) of the *COL1A1* gene. The mutation has been reported 7 times in OI database before, as a variant causing mild OI (Zhang *et al.,* 2012; Roschger *et al.,* 2008; Willing *et al.,* 1996; etc).[2] Probands of family 10 were clinically diagnosed with OI types I, IV and III.

The discovered frameshift mutation results in a PTC, followed by NMD, which is the most common cause of OI type I (Willing *et al.,* 1996). A signal to NMD is the presence of special mRNA-binding exon-exon junction protein complexes, which are attached to mRNA after splicing. If, after the translation, the protein complex stays attached to the mRNA downstream of the stop codon, the ribosome was released before the stop codon preterminally, and the NMD mechanism is launched (Chang *et al.,* 2007). NMD results in the *COL1A1* null allele. The number of mutated transcripts of collagen I α1 in the nucleus and the cytoplasm is significantly decreased. It is likely that there are other mechanisms to provide nonsense mRNA decay for *COL1A1* transcripts. The stretch of conserved amino acids indicates an important functional meaning of the mutation position for protein.

As mentioned above, one of the common features of all OI type I cases, is elevated mineralization of ECM, regardless of the kind of mutation. ECM mineralization analysis might help in distinguishing of OI types of the family 10 members. The analysis of collagen expression levels might reveal the reasons of intrafamilial diversity of family 10.

2.4.3 Novel c.2317G>T *COL1A1* mutation causes severe life-threatening OI

Parent-child trio sequencing revealed novel missense c.2317G>T mutation in proband 716 of family 10. The proband appeared to be an OI type II "survivor". The mutation leads to p.Gly773Cys substitution in collagen I α1 chain, and is absent from the OI database.

The Gly773 residue of the collagen I α1 chain is located at MLBR2 (amino acid residues 680-830 of the α1 chain). The region is rich in α1β1/α2β1 integrin binding sites. Right next to 773 site, there are MMP 1, 2 and 13 interaction domains, HSP47, COMP, discoidin domain receptor 2 and fibronectin binding sites. The region also acts as a fibrillogenesis inhibition site and

binding site for secreted, acidic, rich in Cys proteins (SPARC) (Sweeney *et al.,* 2008) (Supplementary Figure 2).

The inability to bind HSP47 chaperone may result in procollagen early folding and severe OI, similar to OI caused by mutations in the *SERPINH1* gene. The defect in binding of MMP1 could influence interstitial collagen catabolism, altering proteolysis of the defective collagen fibrils (Fields *et al.,* 2013). The lack of MMP13 binding might result in defective collagen mineralization as enzyme restructures fibrils for proper mineralization (Johansson *et al.,* 2000). Defective SPARC binding may cause problems in Ca^{2+} binding with collagen (Yiu *et al.,* 2001). The absence of integrin binding to collagen may cause the loss of interactions between cells and collagen so that ECM and cell binding, and osteoblast homeostasis would be interrupted, and the bone might gain fragility. The loss of fibronectin binding sites might affect fibril cross-linking (Kadler *et al.,* 2008). As mentioned above, the region is crucially important for collagen interaction with NCP, fibril self-assembly, tissue mineralization, and ECM formation. In this way, the mutation in the MLBR2 region may simultaneously affect the tissue morphogenesis in diverse ways. However, the main load of the mutation effect must be provided on the fibril assembly since the function of the most binding NCPs of the MLBR2 region consists of fibrillogenesis.

The MLBR2 region was reported as exclusively lethal (Marini *et al.,* 2007), although the nearby regions contain residues where mutations may cause unlethal OI types I, IV, III (Sweeney *et al.,* 2008). Severity of the mutation effect depends - in addition to mutation position - on amino acid substitution. Two cases of c.2317G>A missense mutations causing p.Gly773Ser substitution have been described before (Marini *et al.,* 2007; Chamberlain *et al.,* 2004). Patients were diagnosed with OI type III. Gly substitutions with Ser and Cys are quite close to each other in degree of severity, as described previously. The key difference is bigger volume of Cys, in contrast to Ser and its ability to form disulfide bridges. Collagen fibril structure was altered more in the case of p.Gly773Cys mutation and might result in lethal phenotype. However, the patient survived until the age of 14 years, which indicates a transitional OI form between types II and III. OI type II survivor phenotype might have been determined not only by extreme care of the parents but also deviations in collagen expressivity or tissue mineralization that facilitated the effect of lethal mutation. Other reported cases of OI caused by missense mutations in the p.Gly773 position could shed light on lethality potential of the substitutions of the current amino acid residue.

CONCLUSION

Osteogenesis Imperfecta is a group of hereditary disorders connected to bone brittleness and skeletal deformities. Phenotypes of the disorder differ with high variety ranging from mild OI type I, to moderate OI type IV, interosseous calcification OI type V, severe OI type III, and lethal OI type II.

OI is caused by mutations in 16 different genes. About 90% of OI cases are represented by autosomal-dominant mutations in collagen I, which is the main structural part of the bone. Mild OI is usually caused by quantitative collagen defects, whereas severe OI is caused by qualitative collagen defects.

In the current study exome analysis of three Estonian OI families was performed. All causative mutations were found in the *COL1A1* gene, what supports the primary strategy of OI diagnosis, including priority screening of the *COL1A1* and *COL1A2* genes.

Family 2 with an autosomal-dominant inheritance pattern of the disease showed segregation of the autosomal-dominant 5'canonical splice site mutation c.1821+1G>A. The mutation causes retention of intron 26 and defective mRNA transport to the cytoplasm. The deficient amount of collagen type I α1 chain mRNAs results in null-allele, collagen quantitative defect and mild OI type I, as the main part of the affected probands of family 2 did. However, one of the patients had a more severe phenotype, which correlated with OI type III.

Exome analysis of family 10 discovered segregation of the autosomal dominant frameshift mutation c.1128delT (p.Gly377AlafsX164). The mutation results in mRNA decay and as a result, *COL1A1* null-allele, collagen quantitative defect and mild OI phenotype. However, the probands revealed phenotypes of OI type I, IV, III.

In family 7, novel autosomal-dominant missense mutation c.2317G>T (p.Gly773Cys) was found. The mutation appeared in the MLBR2 region of the collagen I α1 chain, which is important for collagen interactions with NCPs and fibril self-assembly. The importance of the region is noted with the high number of reported lethal mutations. Current proband appeared to be an OI type II survivor. Obviously, the case studied is an in-between form of lethal and unlethal forms of OI, which indicated deficient OI classification and poor genotype-phenotype correlation concepts.

The results of the study give additional information about diversity of the disease genotypes and raise questions on intrafamilial diversity. Further studies of collagen expressivity have to be provided in families 2 and 10 in order to produce information about the causes of intrafamilial diversity.

The results of exome analysis of family 7 give a new insight into lethality prediction and classification of the disease. Future reports about mutations at the same p.Gly773 residue will provide a decision on pathogenicity and genotype-phenotype correlations of the current mutation. The results of the study also confirm, that exome analysis is an effective method for OI causing mutation discovery.

Osteogenesis Imperfecta geneetiline eelsoodumus: eksoomi analüüs Eesti perekondades

Lidiia Zhytnik

RESÜMEE

Osteogenesis Imperfecta (OI) ehk "habraste luude haigus" on geneetiline haigus, mille tüüpilisteks tunnusteks on sagedased luumurrud, skeleti deformatsioonid ning sinised silmad. OI teisteks enamlevinud sümptomiteks, mis võivad esineda on skolioos, vähenenud luude mass (osteoporoos või osteopeenia), *Dentinogenesis Imperfecta* (DI), kasvupeetus, liigeste lõtvus, kuulmislangus ja kolmnurkse kujuga nägu. Fenotüübid on individuaalsed ja sageli erinevad mitte ainult sama mutatsiooni kandjate vahel, vaid isegi ühe perekonna sees. OI on väga heterogeenne haigus, mis varieerub kergest osteopeeniast kuni perinataalselt letaalsuse vormini. Haiguse üldine esinemissagedus on umbes 1/20 000-30 000 sünnituse kohta ning esinemissagedus sõltub OI tüübist. Eestis on hetkel teada 32 OI perekonda, kellest enamikul esineb haigus läbi mitme põlvkonna. OI-d põhjustavad mutatsioonid 16 erinevates geenides. 85-90% juhtumitest tekkivad kollageeni I α1 ja α2 ahelate (*COL1A1, COL1A2* geenid) mutatsioonide tõttu.

Sillence'i klassifikatsiooni järgi eristatakse nelja klassikalist OI tüüpi.. I OI tüüp on tavaline kerge, mittedeformeeruv OI, mille peamiseks kliiniliseks tunnuseks on sinised skleerad. II OI tüüp on perinataalselt letaalne. Kolmas OI tüüp on raske progressiivselt deformeeruv OI. IV OI tüüp on varieeruv, milleraskusaste on mõõdukas. Viimasel ajal eristatakse ka V OI tüüpi, mille puhul esinevad teistest tüüpidest erinevad histoloogilised ja röntgenoloogilised muutused Tegemist on interosseuste membraanide kaltsifitseerimisega OI ja hüpertroofilise kallusega.

Käesolevas töö on eesmärgiks oli teostada eksoomi sekveneerimisanalüüs kolmes eesti OI perekondades ja määrata OI-tekitavaid mutatsioone.

Perekonnas 2 leiti AD päranduv 5'splaissaiti mutatsioon c.1821+1G>A, mis põhjustas kerget OI. Mutatsiooni tagajärjel lülitati sisse 26. intron kollageeni I α1 ahela mRNAsse. Defektne mRNA ei läbi tuuma SC-35 domääni kontrolli ja mRNA eksporti tsütoplasmasse ei toimu. Tulemuseks on haplopuudulikkus ja kollageeni kvantitatiivne defekt. Sekveneeritud patsientidel esines fenotüübi järgi I ja IV tüüpi OI.

Perekonnas 10 määrati AD päranduv c.1128delT (p.Gly377AlafsX164) raaminihke mutatsioon, mis põhjustas patsientidel erineva raskusastmega OI (tüübid I, IV, III). Antud mutatsioon põhjustab mRNAde lagundamist NMD kaudu ja haplopuudulikkust. Kvantitatiivne kollageeni defekt peaks avalduma kui kerge OI, aga fenotüübi varieeruvus perekonnas oli suurem.

Perekonnas 7 esines *de novo* tekkinud OI mutatsiooni juhtum. Patsiendil esines II tüüpi OI ning ta suri 14,5-aastasena. Tal leiti uus, OI andmebaasist puuduv *missense* c.2317G>T (p.Gly773Cys) mutatsiooni, mis põhjustas eriti raske ja eluohtliku OI. Mutatsioon esines kollageeni I α1 ahela piirkonnas, mis vastutab ligandide seondumise eest. Tulemuseks on defektid, mis ilmnevad kollageeni ja fibrillide voltimises, mineralisatsiooni puudulikkuses, ekstratsellulaarmaatriksi kujunemise häires ja osteoblastide homeostaasis.

Uuringu tulemused annavad lisainformatsiooni OI genotüüpide varieeruvusest, aga ei anna vastust perekonnasisese haiguse varieeruvuse kohta. Selleks on vajalikud täiendavad kollageeni ekspressiooni uuringud, luude mineralisatsiooni analüüs ning detailne mRNA analüüs perekondades 2 ja 10, mis aitaks selgitada perekonnasisese varieeruvuse põhjusi. Perekonna 7 eksoomi analüüsi tulemused annavad uut informatsiooni haiguse letaalsuse ennustamiseks. Järgnevad raportid c.2317G>T (p.Gly773Cys) mutatsiooni kandlusega OI juhtumitest annaksid uusi võimalusi hinnata OI letaalsuse ja mitte-letaalsuse suhteid antud positsioonis. Kokkuvõtteks võiks öelda, et eksoomi analüüs on efektiivne meetod uute OI variantide leidmiseks.

ACKNOWLEDGEMENTS

I would like to express gratitude to my supervisor Katre Maasalu for her time, patience, sharing of unique experience and providing a chance to work on current theme.

Also I would like to thank my co-supervisor Tiit Nikopensius for supportive attitude and teaching of basics of exome analysis.

REFERENCES

A) Magazines

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat Methods. 7(4):248-249.

Akiyama, T., Dass, C.R., Shinoda, Y., Kawano, H., Tanaka, S., Choong, P.F. (2010). PEDF regulates osteoclasts via osteoprotegerin and RANKL. Biochem Biophys Res Commun. 391(1):789-794.

Asharani, P.V.1, Keupp, K., Semler, O., Wang, W., Li, Y., Thiele, H., Yigit, G., Pohl, E., Becker, J., Frommolt, P., Sonntag, C., Altmüller, J., Zimmermann, K., Greenspan, D.S., Akarsu, N.A., Netzer, C., Schönau, E., Wirth, R., Hammerschmidt, M., Nürnberg, P., Wollnik, B., Carney, T.J. (2012). Attenuated BMP1 function compromises osteogenesis, leading to bone fragility in humans and zebrafish. Am J Hum Genet. 90(4):661-674.

Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A., Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet. 12(11):745-755.

Barnes, A.M., Cabral, W.A., Weis, M., Makareeva, E., Mertz, E.L., Leikin, S., Eyre, D., Trujillo, C., Marini, J.C. (2012). Absence of FKBP10 in recessive type XI osteogenesis imperfecta leads to diminished collagen cross-linking and reduced collagen deposition in extracellular matrix. Hum Mutat. 33(11):1589-1598.

Barnes, A.M., Chang, W., Morello, R., Cabral, W.A., Weis, M., Eyre, D.R., Leikin, S., Makareeva, E., Kuznetsova, N., Uveges, T.E., Ashok, A., Flor, A.W., Mulvihill, J.J., Wilson, P.L., Sundaram, U.T., Lee, B., Marini, J.C. (2006). Deficiency of cartilage-associated protein in recessive lethal osteogenesis imperfecta. N Engl J Med. 355(26):2757-2764.

Baron, R., Gertner, J.M., Lang, R., Vignery, A. (1983). Increased bone turnover with decreased bone formation by osteoblasts in children with osteogenesis imperfecta tarda. Pediatr Res. 17(3):204-207.

Barsh, G.S., Roush, C.L., Bonadio, J., Byers, P.H., Gelinas, R.E. (1985). Intron-mediated recombination may cause a deletion in an alpha 1 type I collagen chain in a lethal form of osteogenesis imperfecta. Proc Natl Acad Sci U S A. 82(9):2870-2874.

Basel, D., Steiner, R.D. (2009). Osteogenesis imperfecta: recent findings shed new light on this once well-understood condition. Genet Med. 11(6):375-385

Beck, K., Chan, V.C., Shenoy, N., Kirkpatrick, A., Ramshaw, J.A., Brodsky, B. (2000). Destabilization of osteogenesis imperfecta collagen-like model peptides correlates with the identity of the residue replacing glycine. Proc Natl Acad Sci U S A. 97(8):4273-4278

Becker, J., Semler, O., Gilissen, C., Li, Y., Bolz, H.J., Giunta, C., Bergmann, C., Rohrbach, M., Koerber, F., Zimmermann, K., de Vries, P., Wirth, B., Schoenau, E., Wollnik, B., Veltman, J.A., Hoischen, A., Netzer, C. (2011). Exome sequencing identifies truncating mutations in human SERPINF1 in autosomal-recessive osteogenesis imperfecta. Am J Hum Genet. 88(3):362-371.

Beighton, P., De Paepe, A., Steinmann, B., Tsipouras, P., Wenstrup, R.J. (1998). Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). Am J Med Genet. 77(1):31-37.

Bodian, D.L., Chan, T.F., Poon, A., Schwarze, U., Yang, K., Byers, P.H., Kwok, P.Y., Klein, T.E. (2009). Mutation and polymorphism spectrum in osteogenesis imperfecta type II: implications for genotype-phenotype relationships. Hum Mol Genet. 18(3):463-471.

Boedtker, H., Fuller, F., Tate, V. (1983). The structure of collagen genes. Int Rev Connect Tissue Res. 10:1-63.

Bogan, R.1, Riddle, R.C., Li, Z., Kumar, S., Nandal, A., Faugere, M.C., Boskey, A., Crawford, S.E., Clemens, T.L. (2013). A mouse model for human osteogenesis imperfecta type VI. J Bone Miner Res. 28(7):1531-1536.

Buck, D.W. 2nd, Dumanian, G.A. (2012) Bone biology and physiology: Part I. The fundamentals. Plast Reconstr Surg. 129(6):950e-956e.

Buckwalter, J.A, Glimcher, M.J., Cooper, R.R., Recker, R. (1996). Bone biology. I: Structure, blood supply, cells, matrix, and mineralization. Instr Course Lect. 45:371-386.

Byers, P.H and Steiner, R.D. (1992). Osteogenesis Imperfecta. Annu Rev Med. 43:269-282.

Byers, P.H., Pyott, S.M. (2012). Recessively inherited forms of osteogenesis imperfecta. Annu Rev Genet. 46:475-497.

Cabral, W.A., Barnes, A.M., Adeyemo, A., Cushing, K., Chitayat, D., Porter, F.D., Panny, S.R., Gulamali-Majid, F., Tishkoff, S.A., Rebbeck, T.R., Gueye, S.M., Bailey-Wilson, J.E., Brody, L.C., Rotimi, C.N., Marini, J.C. (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. 543-551.

Cabral, W.A., Chang, W., Barnes, A.M., Weis, M., Scott, M.A., Leikin, S., Makareeva, E., Kuznetsova, N.V., Rosenbaum, K.N., Tifft, C.J., Bulas, D.I., Kozma, C., Smith, P.A., Eyre, D.R., Marini, J.C. (2008). Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. Nat Genet. 39(3):359-365.

Cabral, W.A., Chang, W., Barnes, A.M., Weis, M., Scott, M.A., Leikin, S., Makareeva, E., Kuznetsova, N.V., Rosenbaum, K.N., Tifft, C.J., Bulas, D.I., Kozma, C., Smith, P.A., Eyre, D.R., Marini, J.C. (2007). Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. Nat Genet. 39(3):359-365

Chamberlain, J.R.1, Schwarze, U., Wang, P.R., Hirata, R.K., Hankenson, K.D., Pace, J.M., Underwood, R.A., Song, K.M., Sussman, M., Byers, P.H., Russell, D.W. (2004). Gene targeting in stem cells from individuals with osteogenesis imperfecta. Science. 303(5661):1198-1201.

Chang, W., Barnes, A.M., Cabral, W.A., Bodurtha, J.N., Marini, J.C. (2010). Prolyl 3-hydroxylase 1 and CRTAP are mutually stabilizing in the endoplasmic reticulum collagen prolyl 3-hydroxylation complex. Hum Mol Genet. 19(2):223-234.

Chang, Y.F., Imam, J.S., Wilkinson, M.F. (2007). The nonsense-mediated decay RNA surveillance pathway. Annu Rev Biochem. 76:51-74.

Cheung, M.S., Glorieux, F.H., Rauch, F.(2008). Natural history of hyperplastic callus formation in osteogenesis imperfecta type V. J Bone Miner Res. 22(8):1181-1186.

Cho, S.Y.1, Ki, C.S., Sohn, Y.B., Kim, S.J., Maeng, S.H., Jin, D.K. (2013). Osteogenesis imperfecta Type VI with severe bony deformities caused by novel compound heterozygous mutations in SERPINF1. J Korean Med Sci. 28(7):1107-1110.

Cho, T.J., Lee, K.E., Lee, S.K., Song, S.J., Kim, K.J., Jeon, D., Lee, G., Kim, H.N., Lee, H.R., Eom, H.H., Lee, Z.H., Kim, O.H., Park, W.Y., Park, S.S., Ikegawa, S., Yoo, W.J., Choi, I.H., Kim, J.W. (2012). A single recurrent mutation in the 5'-UTR of IFITM5 causes osteogenesis imperfecta type V. Am J Hum Genet. 91(2):343-348.

Christiansen, H.E., Schwarze, U., Pyott, S.M., AlSwaid, A., Al Balwi, M., Alrasheed, S., Pepin, M.G., Weis, M.A., Eyre, D.R., Byers, P.H. (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(3):389-398.

Chu, M.L., Gargiulo, V., Williams, C.J., Ramirez, F. (1985). Multiexon deletion in an osteogenesis imperfecta variant with increased type III collagen mRNA. J Biol Chem. 260(2):691-694.

Clarke, B. (2008). Normal bone anatomy and physiology. Clin J Am Soc Nephrol. 3:S131-139.

Cui, Y., Niziolek, P.J., MacDonald, B.T., Zylstra, C.R., Alenina, N., Robinson, D.R., Zhong, Z., Matthes, S., Jacobsen, C.M., Conlon, R.A., Brommage, R., Liu, Q., Mseeh, F., Powell, D.R., Yang, Q.M., Zambrowicz, B., Gerrits, H., Gossen, J.A., He, X., Bader, M., Williams, B.O., Warman, M.L., Robling, A.G. (2011). Lrp5 functions in bone to regulate bone mass. Nat Med. 17(6):684-691.

Cundy, T. (2012). Recent advances in osteogenesis imperfecta. Calcif Tissue Int. 90(6):439-449.

Currey, J.D. (2012). The structure and mechanics of bone. J Mater Sci. 47:41-54.

Di Lullo, G.A., Sweeney, S.M., Korkko, J., Ala-Kokko, L., San Antonio, J.D. (2002). Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen. J Biol Chem. 277(6):4223-4231.

Edwards, M.J., Wenstrup, R.J., Byers, P.H., Cohn, D.H. (1992). Recurrence of lethal osteogenesis imperfecta due to parental mosaicism for a mutation in the COL1A2 gene of type I collagen. The mosaic parent exhibits phenotypic features of a mild form of the disease. Hum Mutat. 1(1):47-54.

Fahiminiya, S., Majewski, J., Al-Jallad, H., Moffatt, P., Mort, J., Glorieux, F.H., Roschger, P., Klaushofer, K., Rauch, F. (2014). Osteoporosis Caused by Mutations in PLS3 - Clinical and Bone Tissue Characteristics. J Bone Miner Res. [Epub ahead of print]

Fahiminiya, S., Majewski, J., Mort, J., Moffatt, P., Glorieux, F.H., Rauch, F. (2013). Mutations in WNT1 are a cause of osteogenesis imperfecta. J Med Genet. 50(5):345-348.

Farber, C.R, Reich, A., Barnes, A.M., Becerra, P., Rauch, F., Cabral, W.A., Bae, A., Quinlan, A., Glorieux, F.H., Clemens, T.L., Marini, J.C. (2014). A Novel IFITM5 Mutation in Severe Atypical Osteogenesis Imperfecta Type VI Impairs Osteoblast Production of Pigment Epithelium-Derived Factor. J Bone Miner Res. [Epub ahead of print]

Fields, G.B. (2013). Interstitial collagen catabolism. J Biol Chem. 288(13):8785-8793.

Forlino, A., Cabral, W.A., Barnes, A.M., Marini, J.C. (2011). New perspectives on osteogenesis imperfecta. Nat Rev Endocrinol. 7(9):540-557.

Galante, P.A., Sakabe, N.J., Kirschbaum-Slager, N., de Souza, S.J. (2004). Detection and evaluation of intron retention events in the human transcriptome. RNA. 10(5):757-765.

Gelse, K., Pöschl, E., Aigner, T. (2003). Collagens--structure, function, and biosynthesis. Adv Drug Deliv Rev. 55(12):1531-1546.

Gilissen, C., Hoischen, A., Brunner, H.G., Veltman, J.A. (2011). Unlocking Mendelian disease using exome sequencing. Genome Biol. 12(9):228.

Gioia, R., Panaroni, C., Besio, R., Palladini, G., Merlini, G., Giansanti, V., Scovassi, I.A., Villani, S., Villa, I., Villa, A., Vezzoni, P., Tenni, R., Rossi, A., Marini, J.C., Forlino, A. (2012). Impaired osteoblastogenesis in a murine model of dominantosteogenesis imperfecta: a new target for osteogenesis imperfecta pharmacological therapy. Stem Cells. (7):1465-1476.

Glorieux, F.H., Moffatt, P. (2013). Osteogenesis imperfecta, an ever-expanding conundrum. J Bone Miner Res. 28(7):1519-1522.

Glorieux, F.H., Rauch, F., Plotkin, H., Ward, L., Travers, R., Roughley, P., Lalic, L., Glorieux, D.F., Fassier, F., Bishop, N.J. (2000). Type V osteogenesis imperfecta: a new form of brittle bone disease. J Bone Miner Res. 15(9):1650-1658.

Glorieux, F.H.1, Ward, L.M., Rauch, F., Lalic, L., Roughley, P.J., Travers, R. (2002). Osteogenesis imperfecta type VI: a form of brittle bone disease with a mineralization defect. J Bone Miner Res. 17(1):30-38.

González-Pérez, A.1, López-Bigas, N. (2011). Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. Am J Hum Genet. 88(4):440-449.

Grgurevic, L.1, Macek, B., Mercep, M., Jelic, M., Smoljanovic, T., Erjavec, I., Dumic-Cule, I., Prgomet, S., Durdevic, D., Vnuk, D., Lipar, M., Stejskal, M., Kufner, V., Brkljacic, J., Maticic, D., Vukicevic, S. (2011). Bone morphogenetic protein (BMP)1-3 enhances bone repair. Biochem Biophys Res Commun. 408(1):25-31.

Guillén-Navarro, E., Ballesta-Martínez, M.J., Valencia, M., Bueno, A.M., Martinez-Glez, V., López-González, V., Burnyte, B., Utkus, A., Lapunzina, P., Ruiz-Perez, V.L. (2014). Two mutations in IFITM5 causing distinct forms of osteogenesis imperfecta. Am J Med Genet A. 164(5):1136-1142.

Hall, L.L.1, Smith, K.P., Byron, M., Lawrence, J.B. (2006). Molecular anatomy of a speckle. Anat Rec A Discov Mol Cell Evol Biol. 288(7):664-675.

Hanagata, N., Li, X., Morita, H., Takemura, T., Li, J., Minowa, T. (2011). Characterization of the osteoblast-specific transmembrane protein IFITM5 and analysis of IFITM5-deficient mice. J Bone Miner Metab. 29(3):279-290.

Homan, E.P.1, Rauch, F., Grafe, I., Lietman, C., Doll, J.A., Dawson, B., Bertin, T., Napierala, D., Morello, R., Gibbs, R., White, L., Miki, R., Cohn, D.H., Crawford, S., Travers, R., Glorieux, F.H., Lee, B. (2011). Mutations in SERPINF1 cause osteogenesis imperfecta type VI. J Bone Miner Res. 26(12):2798-2803.

Horton, M.A., Nesbit, M.A., Helfrich, M.H. (1995). Interaction of osteopontin with osteoclast integrins. Ann N Y Acad Sci. 760:190-200.

Hosomichi, J., Yasui, N., Koide, T., Soma, K., Morita, I. (2005). Involvement of the collagen I-binding motif in the anti-angiogenic activity of pigment epithelium-derived factor. Biochem Biophys Res Commun. 335(3):756-761.

Hoyer-Kuhn, H.1, Semler, O., Garbes, L., Zimmermann, K., Becker, J., Wollnik, B., Schoenau, E., Netzer, C. (2014). A Nonclassical IFITM5 Mutation Located in the Coding Region

Causes Severe Osteogenesis Imperfecta With Prenatal Onset. J Bone Miner Res. 29(6):1387-1391.

Ishida, Y., Nagata, K. (2011). Hsp47 as a collagen-specific molecular chaperone. Methods Enzymol. 499:167-182.

Johansson, N., Ala-aho, R., Uitto, V., Grénman, R., Fusenig, N.E., López-Otín, C., Kähäri, V.M. (2000). Expression of collagenase-3 (MMP-13) and collagenase-1 (MMP-1) by transformed keratinocytes is dependent on the activity of p38 mitogen-activated protein kinase. J Cell Sci. 113 Pt 2:227-235.

Johnson, C., Primorac, D., McKinstry, M., McNeil, J., Rowe, D., Lawrence, J.B. (2000). Tracking COL1A1 RNA in osteogenesis imperfecta. splice-defective transcripts initiate transport from the gene but are retained within the SC35 domain. J Cell Biol. 150(3):417-432.

Kadler, K.E.1, Hill, A., Canty-Laird, E.G. (2008). Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. Curr Opin Cell Biol. 20(5):495-501

Kelley, B.P.1, Malfait, F., Bonafe, L., Baldridge, D., Homan, E., Symoens, S., Willaert, A., Elcioglu, N., Van Maldergem, L., Verellen-Dumoulin, C., Gillerot, Y., Napierala, D., Krakow, D., Beighton, P., Superti-Furga, A., De Paepe, A., Lee, B. (2011). Mutations in FKBP10 cause recessive osteogenesis imperfecta and Bruck syndrome. J Bone Miner Res. 26(3):666-672.

Keupp, K.1, Beleggia, F., Kayserili, H., Barnes, A.M., Steiner, M., Semler, O., Fischer, B., Yigit, G., Janda, C.Y., Becker, J., Breer, S., Altunoglu, U., Grünhagen, J., Krawitz, P., Hecht, J., Schinke, T., Makareeva, E., Lausch, E., Cankaya, T., Caparrós-Martín, J.A., Lapunzina, P., Temtamy, S., Aglan, M., Zabel, B., Eysel, P., Koerber, F., Leikin, S., Garcia, K.C., Netzer, C., Schönau, E., Ruiz-Perez, V.L., Mundlos, S., Amling, M., Kornak, U., Marini, J., Wollnik, B. (2013). Mutations in WNT1 cause different forms of bone fragility. Am J Hum Genet. 92(4):565-574.

Kim, O.H., Jin, D.K., Kosaki, K., Kim, J.W., Cho, S.Y., Yoo, W.J., Choi, I.H., Nishimura, G., Ikegawa, S., Cho, T.J. (2013). Osteogenesis imperfecta type V: clinical and radiographic manifestations in mutation confirmed patients. Am J Med Genet A. 161A(8):1972-1979.

Kocher, M.S., Shapiro, F. (1998). Osteogenesis Imperfecta. J Am Acad Orthop Surg. 6(4):225-236

Ku, C.S., Naidoo, N., Pawitan, Y. (2011). Revisiting Mendelian disorders through exome sequencing. Hum Genet. 129(4):351-370.

Kumar, P., Henikoff, S., Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 4(7):1073-1081.

Laine, C.M., Joeng, K.S., Campeau, P.M., Kiviranta, R., Tarkkonen, K., Grover, M., Lu, J.T., Pekkinen, M., Wessman, M., Heino, T.J., Nieminen-Pihala, V., Aronen, M., Laine, T., Kröger, H., Cole, W.G., Lehesjoki, A.E., Nevarez, L., Krakow, D., Curry, C.J., Cohn, D.H., Gibbs, R.A., Lee, B.H., Mäkitie, O. (2013). WNT1 mutations in early-onset osteoporosis and osteogenesis imperfecta. N Engl J Med. 368(19):1809-1816.

Lapunzina, P., Aglan, M., Temtamy, S., Caparrós-Martín, J.A., Valencia, M., Letón, R., Martínez-Glez, V., Elhossini, R., Amr, K., Vilaboa, N., Ruiz-Perez, V.L. (2010). Identification of a frameshift mutation in Osterix in a patient with recessive osteogenesis imperfecta. Am J Hum Genet. 87(1):110-4.

Last, J.A., Reiser, K.M. (1984).Collagen biosynthesis. Environ Health Perspect. 55:169-177.

Lazarus, S., Moffatt, P., Duncan, E.L., Thomas, G.P. (2014). A brilliant breakthrough in OI type V. Osteoporos Int. 25(2):399-405.

Lee, N.K., Sowa, H., Hinoi, E., Ferron, M., Ahn, J.D., Confavreux, C., Dacquin, R., Mee, P.J., McKee, M.D., Jung, D.Y., Zhang, Z., Kim, J.K., Mauvais-Jarvis, F., Ducy, P., Karsenty, G. (2007). "Endocrine Regulation of Energy Metabolism by the Skeleton". Cell. 130(3): 456–469.

Lee, D.Y., Cho, T.J., Choi, I.H., Chung, C.Y., Yoo, W.J., Kim, J.H., Park, Y.K. (2006). Clinical and radiological manifestations of osteogenesis imperfecta type V. J Korean Med Sci. 21(4):709-714.

Li, D., Yu, J., Gu, F., Pang, X., Ma, X., Li, R., Liu, N., Ma, X. (2008). The roles of two novel FBN1 gene mutations in the genotype-phenotype correlations of Marfan syndrome and ectopia lentis patients with marfanoid habitus. Genet Test. 12(2):325-330.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.; 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25(16):2078-2079.

Li, H.1, Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25(14):1754-1760.

Lin, G.L., Hankenson, K.D. (2011). Integration of BMP, Wnt, and notch signaling pathways in osteoblast differentiation. J Cell Biochem. 112(12):3491-3501.

Lindahl, K., Barnes, A.M., Fratzl-Zelman, N., Whyte, M.P., Hefferan, T.E., Makareeva, E., Brusel, M., Yaszemski, M.J., Rubin, C.J., Kindmark, A., Roschger, P., Klaushofer, K., McAlister, W.H., Mumm, S., Leikin, S., Kessler, E., Boskey, A.L., Ljunggren, O., Marini, J.C. (2011). COL1 C-propeptide cleavage site mutations cause high bone mass osteogenesis imperfecta. Hum Mutat. 32(6):598-609.

Lisse, T.S., Thiele, F., Fuchs, H., Hans, W., Przemeck, G.K., Abe, K., Rathkolb, B., Quintanilla-Martinez, L., Hoelzlwimmer, G., Helfrich, M., Wolf, E., Ralston, S.H., Hrabé de Angelis, M. (2008). ER stress-mediated apoptosis in a new mouse model of osteogenesis imperfecta. PLoS Genet. 4(2):e7.

Long, F. (2012). Building strong bones molecular regulation of the osteoblast lineage. Nat Rev Mol Cell Biol. 13(1):27-38.

Maasalu, K., Haviko, T., Märtson, A. (2003). Treatment of children with Osteogenesis imperfecta in Estonia. Acta Paediatrica, 92: 452–455

Makareeva, E.1, Mertz, E.L., Kuznetsova, N.V., Sutter, M.B., DeRidder, A.M., Cabral, W.A., Barnes, A.M., McBride, D.J., Marini, J.C., Leikin, S.(2008). Structural heterogeneity of type I collagen triple helix and its role in osteogenesis imperfecta. J Biol Chem. 283(8):4787-4798.

Marini, J.C., Forlino, A., Cabral, W.A., Barnes, A.M., San Antonio, J.D., Milgrom, S., Hyland, J.C., Körkkö, J., Prockop, D.J., De Paepe, A., Coucke, P., Symoens, S., Glorieux, F.H., Roughley, P.J., Lund, A.M., Kuurila-Svahn, K., Hartikka, H., Cohn, D.H., Krakow, D., Mottes, M., Schwarze, U., Chen, D., Yang, K., Kuslich, C., Troendle, J., Dalgleish, R., Byers, P.H. (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.

Marini, J.C.1, Blissett, A.R. (2013). New genes in bone development: what's new in osteogenesis imperfecta. J Clin Endocrinol Metab. 98(8):3095-3103.

Martínez-Glez, V., Valencia, M., Caparrós-Martín, J.A., Aglan, M., Temtamy, S., Tenorio, J., Pulido, V., Lindert, U., Rohrbach, M., Eyre, D., Giunta, C., Lapunzina, P., Ruiz-Perez, V.L. (2012). Identification of a mutation causing deficient BMP1/mTLD proteolytic activity in autosomal recessive osteogenesis imperfecta. Hum Mutat. 33(2):343-350.

McKenna, A.1, Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20(9):1297-1303.

Miko, I. (2008). Phenotype variability: penetrance and expressivity. Nature Education 1(1):137

Miller, A. (1984). Collagen: the organic matrix of bone. Philos Trans R Soc Lond B Biol Sci. 304(1121):455-477.

Moraes, M.V., Milanez, M., Almada, B.V., Sipolatti, V., Rebouças, M.R., Nunes, V.R., Akel, A.N. Jr., Zatz, M., Errera, F.I., Louro, I.D., Paula, F. (2012). Variable expressivity of osteogenesis imperfecta in a Brazilian family due to p.G1079S mutation in the COL1A1 gene. Genet Mol Res. 11(3):3246-3255.

Morello, R., Bertin, T.K., Chen, Y., Hicks, J., Tonachini, L., Monticone, M., Castagnola, P., Rauch, F., Glorieux, F.H., Vranka, J., Bächinger, H.P., Pace, J.M., Schwarze, U., Byers, P.H., Weis, M., Fernandes, R.J., Eyre, D.R., Yao, Z., Boyce, B.F., Lee, B. (2006). CRTAP is required for prolyl 3- hydroxylation and mutations cause recessive osteogenesis imperfecta. Cell. 127(2):291-304.

Mörike, M., Schulz, M., Brenner, R.E., Bushart, G.B., Teller, W.M., Vetter, U. (1993). In vitro expression of osteoblastic markers in cells isolated from normal fetal and postnatal human bone and from bone of patients with osteogenesis imperfecta. J Cell Physiol. 157(3):439-444.

Mundlos, S., Chan, D., Weng, Y.M., Sillence, D.O., Cole, W.G., Bateman, J.F. (1996). Multiexon deletions in the type I collagen COL1A2 gene in osteogenesis imperfecta type

IB. Molecules containing the shortened alpha2(I) chains show differential incorporation into the bone and skin extracellular matrix. J Biol Chem. 271(35):21068-21074.

Ng, S.B., Buckingham, K.J., Lee, C., Bigham, A.W., Tabor, H.K., Dent, K.M., Huff, C.D., Shannon, P.T., Jabs, E.W., Nickerson, D.A., Shendure, J., Bamshad, M.J. (2010). Exome sequencing identifies the cause of a mendelian disorder. Nat Genet. 42(1):30-35.

Nikopensius, T., Saag, M., Jagomägi, T., Annilo, T., Kals, M., Kivistik, P.A., Milani, L., Metspalu, A. (2013). A missense mutation in DUSP6 is associated with Class III malocclusion. J Dent Res. 92(10):893-898.

Noble, B.S. (2008). The osteocyte lineage. Arch Biochem Biophys. 473(2):106-111Pyott, S.M., Tran, T.T., Leistritz, D.F., Pepin, M.G., Mendelsohn, N.J., Temme, R.T., Fernandez, B.A., Elsayed, S.M., Elsobky, E., Verma, I., Nair, S., Turner, E.H., Smith, J.D., Jarvik, G.P., Byers, P.H. (2013). WNT1 mutations in families affected by moderately severe and progressive recessiveosteogenesis imperfecta. Am J Hum Genet. 92(4):590-597.

Omori, Y.1, Imai, J., Suzuki, Y., Watanabe, S., Tanigami, A., Sugano, S. (2002). OASIS is a transcriptional activator of CREB/ATF family with a transmembrane domain. Biochem Biophys Res Commun. 293(1):470-477.

Palomo, T., Glorieux, F.H., Rauch, F. (2014). Circulating sclerostin in children and young adults with heritable bone disorders. J Clin Endocrinol Metab. 99(5):E920-925.

Puig-Hervás, M.T.1, Temtamy, S., Aglan, M., Valencia, M., Martínez-Glez, V., Ballesta-Martínez, M.J., López-González, V., Ashour, A.M., Amr, K., Pulido, V., Guillén-Navarro, E., Lapunzina, P., Caparrós-Martín, J.A., Ruiz-Perez, V.L. (2012). Mutations in PLOD2 cause autosomal-recessive connective tissue disorders within the Brucksyndrome--osteogenesis imperfecta phenotypic spectrum. Hum Mutat. 33(10):1444-1449.

Pyott, S.M.1, Schwarze, U., Christiansen, H.E., Pepin, M.G., Leistritz, D.F., Dineen, R., Harris, C., Burton, B.K., Angle, B., Kim, K., Sussman, M.D., Weis, M., Eyre, D.R., Russell, D.W., McCarthy, K.J., Steiner, R.D., Byers, P.H. (2011). Mutations in PPIB (cyclophilin B) delay type I procollagen chain association and result in perinatal lethal to moderate osteogenesis imperfecta phenotypes. Hum Mol Genet. 20(8):1595-1609.

Qin, W., He, J.X., Shi, J., Xing, Q.H., Gao, J.J., He, L., Qian, X.Q., Liu, Z.J., Shu, A.L., He, L. (2005). [Mutation detection of COL1A1 gene in a pedigree with osteogenesis imperfecta]. Yi Chuan Xue Bao. 32(3):248-252. (Chinese).

Rabbani, B.1, Mahdieh, N., Hosomichi, K., Nakaoka, H., Inoue, I. (2012). Next-generation sequencing: impact of exome sequencing in characterizing Mendelian disorders. J Hum Genet. 57(10):621-632.

Rauch, F., Husseini, A., Roughley, P., Glorieux, F.H., Moffatt, P. (2012). Lack of circulating pigment epithelium-derived factor is a marker of osteogenesis imperfecta type VI. J Clin Endocrinol Metab. 97(8):E1550-1556.

Reeder, J., Orwoll, E. (2006). Images in clinical medicine. Adults with osteogenesis imperfecta. N Engl J Med. 355(26):e28.

Robberson, B.L.1, Cote, G.J., Berget, S.M. (1990). Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol Cell Biol. 10(1):84-94.

Roschger, P., Fratzl-Zelman, N., Misof, B.M., Glorieux, F.H., Klaushofer, K., Rauch, F. (2008). Evidence that abnormal high bone mineralization in growing children with osteogenesisimperfecta is not associated with specific collagen mutations. Calcif Tissue Int. 82(4):263-70.

Roughley, P.J., Rauch, F., Glorieux, F.H. (2003). Osteogenesis imperfecta--clinical and molecular diversity. Eur Cell Mater. 5:41-47.

Salo, A.M.1, Cox, H., Farndon, P., Moss, C., Grindulis, H., Risteli, M., Robins, S.P., Myllylä, R. (2008). A connective tissue disorder caused by mutations of the lysyl hydroxylase 3 gene. Am J Hum Genet. 83(4):495-503.

Sarafova, A.P., Choi, H., Forlino, A., Gajko, A., Cabral, W.A., Tosi, L., Reing, C.M., Marini, J.C. (1998). Three novel type I collagen mutations in osteogenesis imperfecta type IV probands are associated with discrepancies between electrophoretic migration of osteoblast and fibroblast collagen. Hum Mutat. 11(5):395-403.

Schwarze, U.1, Cundy, T., Pyott, S.M., Christiansen, H.E., Hegde, M.R., Bank, R.A., Pals, G., Ankala, A., Conneely, K., Seaver, L., Yandow, S.M., Raney, E., Babovic-Vuksanovic, D., Stoler, J., Ben-Neriah, Z., Segel, R., Lieberman, S., Siderius, L., Al-Aqeel, A., Hannibal, M., Hudgins, L., McPherson, E., Clemens, M., Sussman, M.D., Steiner,

R.D., Mahan, J., Smith, R., Anyane-Yeboa, K., Wynn, J., Chong, K., Uster, T., Aftimos, S., Sutton, V.R., Davis, E.C., Kim, L.S., Weis, M.A., Eyre, D., Byers, P.H. (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.

Semler, O., Garbes, L., Keupp, K., Swan, D., Zimmermann, K., Becker, J., Iden, S., Wirth, B., Eysel, P., Koerber, F., Schoenau, E., Bohlander, S.K., Wollnik, B., Netzer, C. (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(2):349-357.

Shaheen, R.1, Alazami, A.M., Alshammari, M.J., Faqeih, E., Alhashmi, N., Mousa, N., Alsinani, A., Ansari, S., Alzahrani, F., Al-Owain, M., Alzayed, Z.S., Alkuraya, F.S. (2012). Study of autosomal recessive osteogenesis imperfecta in Arabia reveals a novel locus defined by TMEM38B mutation. J Med Genet. 49(10):630-635.

Shopland, L.S., Johnson, C.V., Lawrence, J.B. (2002). Evidence that all SC-35 domains contain mRNAs and that transcripts can be structurally constrained within these domains. J Struct Biol. 140(1-3):131-139.

Sillence, D.O., Senn, A., Danks, D.M. (1979). Genetic heterogeneity in osteogenesis imperfecta. J Med Genet. 16(2):101-116.

Stamm, S., Zhu, J., Nakai, K., Stoilov, P., Stoss, O., Zhang, M.Q. (2000). An alternative-exon database and its statistical analysis. DNA Cell Biol. 19(12):739-756.

Steiner, R.D., Pepin, M.G., Byers, P.H. (1993). Osteogenesis Imperfecta, in Pagon RA, Bird TD, Dolan CR, Stephens K (eds): GeneReviews

Steinmann, B., Tuderman, L., Peltonen, L., Martin, G. R., McKusick, V. A., Prockop, D. J. (1980). Evidence for a structural mutation of procollagen type I in a patient with the Ehlers-Danlos syndrome type VII. J. Biol. Chem. 255: 8887-8893.

Stover, M.L., Primorac, D., Liu, S.C., McKinstry, M.B., Rowe, D.W. (1993). Defective splicing of mRNA from one COL1A1 allele of type I collagen in nondeforming (type I) osteogenesis imperfecta. J Clin Invest. 92(4):1994-2002.

Sweeney, S.M., Orgel, J.P., Fertala, A., McAuliffe, J.D., Turner, K.R., Di Lullo, G.A., Chen, S., Antipova, O., Perumal, S., Ala-Kokko, L., Forlino, A., Cabral, W.A., Barnes, A.M., Marini, J.C., San Antonio, J.D. (2008). Candidate cell and matrix interaction domains on the collagen fibril, the predominant protein of vertebrates. J Biol Chem. 283(30):21187-21197

Symoens, S., Malfait, F., D'hondt, S., Callewaert, B., Dheedene, A., Steyaert, W., Bächinger, H.P., De Paepe, A., Kayserili, H., Coucke, P.J. (2013). Deficiency for the ER-stress transducer OASIS causes severe recessive osteogenesis imperfecta in humans. Orphanet J Rare Dis. 8:154.

Van Dijk, F.S., Zillikens, M.C., Micha, D., Riessland, M., Marcelis, C.L., de Die-Smulders, C.E., Milbradt, J., Franken, A.A., Harsevoort, A.J., Lichtenbelt, K.D., Pruijs, H.E., Rubio-Gozalbo, M.E., Zwertbroek, R., Moutaouakil, Y., Egthuijsen, J., Hammerschmidt, M., Bijman, R., Semeins, C.M., Bakker, A.D., Everts, V., Klein-Nulend, J., Campos-Obando, N., Hofman, A., te Meerman, G.J., Verkerk, A.J., Uitterlinden, A.G., Maugeri, A., Sistermans, E.A., Waisfisz, Q., Meijers-Heijboer, H., Wirth, B., Simon, M.E., Pals, G. (2013). PLS3 mutations in X-linked osteoporosis with fractures. N Engl J Med. 369(16):1529-1536.

Van Dijk, F.S., Nikkels, P.G., den Hollander, N.S., Nesbitt, I.M., van Rijn, R.R., Cobben, J.M., Pals, G. (2011). Lethal/severe osteogenesis imperfecta in a large family: a novel homozygous LEPRE1 mutation and bone histological findings. Pediatr Dev Pathol. 14(3):228-234.

Van Dijk, F.S., Pals, G., Van Rijn, R.R., Nikkels, P.G., Cobben, J.M. (2010). Classification of Osteogenesis Imperfecta revisited. Eur J Med Genet. 53(1):1-5.

Van Dijk, F.S., Nesbitt, I.M., Zwikstra, E.H., Nikkels, P.G., Piersma, S.R., Fratantoni, S.A., Jimenez, C.R., Huizer, M., Morsman, A.C., Cobben, J.M., van Roij, M.H., Elting, M.W., Verbeke, J.I., Wijnaendts L.C., Shaw, N.J., Högler, W., McKeown, C., Sistermans, E.A., Dalton, A., Meijers-Heijboer, H., Pals, G. (2009). PPIB mutations cause severe osteogenesis imperfecta. Am J Hum Genet. 85(4):521-527.

Van Dijk, F.S.1, Byers, P.H., Dalgleish, R., Malfait, F., Maugeri, A., Rohrbach, M., Symoens, S., Sistermans, E.A., Pals, G. (2012). EMQN best practice guidelines for the laboratory diagnosis of osteogenesis imperfecta. Eur J Hum Genet. 20(1):11-19.

Volodarsky, M.1, Markus, B., Cohen, I., Staretz-Chacham, O., Flusser, H., Landau, D., Shelef, I., Langer, Y., Birk, O.S. (2013). A deletion mutation in TMEM38B associated with autosomal recessive osteogenesis imperfecta. Hum Mutat. 34(4):582-586.

Willing, M.C., Cohn, D.H., Starman, B., Holbrook, K.A., Greenberg, C.R., Byers, P.H. (1988). Heterozygosity for a large deletion in the alpha 2(I) collagen gene has a dramatic effect on type I collagen secretion and produces perinatal lethal osteogenesis imperfecta. J Biol Chem. 263(17):8398-8404.

Willing, M.C.1, Deschenes, S.P., Slayton, R.L., Roberts, E.J. (1996). Premature chain termination is a unifying mechanism for COL1A1 null alleles in osteogenesis imperfecta type I cell strains. Am J Hum Genet. 59(4):799-809.

Xia, X.Y., Li, W.W., Li, N., Wu, Q.Y., Cui, Y.X., Li, X.J. (2014). A novel mild variant of osteogenesis imperfecta type I caused by a Gly1088Glu mutation in COL1A1. Mol Med Rep. 9(6):2187-2190.

Yasui, N., Mori, T., Morito, D., Matsushita, O., Kourai, H., Nagata, K., Koide, T. (2003). Dual-site recognition of different extracellular matrix components by anti-angiogenic/neurotrophic serpin, PEDF. Biochemistry. 42(11):3160-3167.

Yiu, G.K.1, Chan, W.Y., Ng, S.W., Chan, P.S., Cheung, K.K., Berkowitz, R.S., Mok, S.C. (2001). SPARC (secreted protein acidic and rich in cysteine) induces apoptosis in ovarian cancer cells. Am J Pathol. 159(2):609-622.

Zhang, Z.L.1, Zhang, H., Ke, Y.H., Yue, H., Xiao, W.J., Yu, J.B., Gu, J.M., Hu, W.W., Wang, C., He, J.W., Fu, W.Z. (2012). The identification of novel mutations in COL1A1, COL1A2, and LEPRE1 genes in Chinese patients with osteogenesis imperfecta. J Bone Miner Metab. 30(1):69-77.

B) Books

Betts, J.G., DeSaix, P., Johnson, E., Johnson, J. E. Korol, O., Kruse, D.H., Poe, B., Wise, J. A., Young, K. A. (2013). Anatomy and physiology textbook. OpenStax College

Shapiro, J.R., Byers, P.H., Glorieux, F.H., Sponseller, P.D. (2013). Osteogenesis Imperfecta: A Translational Approach to Brittle Bone Disease. Elsevier. Academic Press

Steele, D.G., Claud A.B. (1988). The Anatomy and Biology of the Human Skeleton. Texas A&M University Press. p. 4.

C) Used web addresses

[1]http://www.orpha.net

[2]https://oi.gene.le.ac.uk/home.php

[3] http://www.ncbi.nlm.nih.gov/gene/1277

[4] http://www.ncbi.nlm.nih.gov/gene/1278

[5]http://www.kliinikum.ee/ortopeedia

[6]http://www.progenygenetics.com/online-pedigree

[7]http://bioinfo.ut.ee/primer3-0.4.0/primer3

[8]http://genome.ucsc.edu

[9]http://www.lifetechnologies.com/ee/en/home.html

[10]http://www.mbio.ncsu.edu/bioedit/bioedit.html

SUPPLEMENTARY DATA

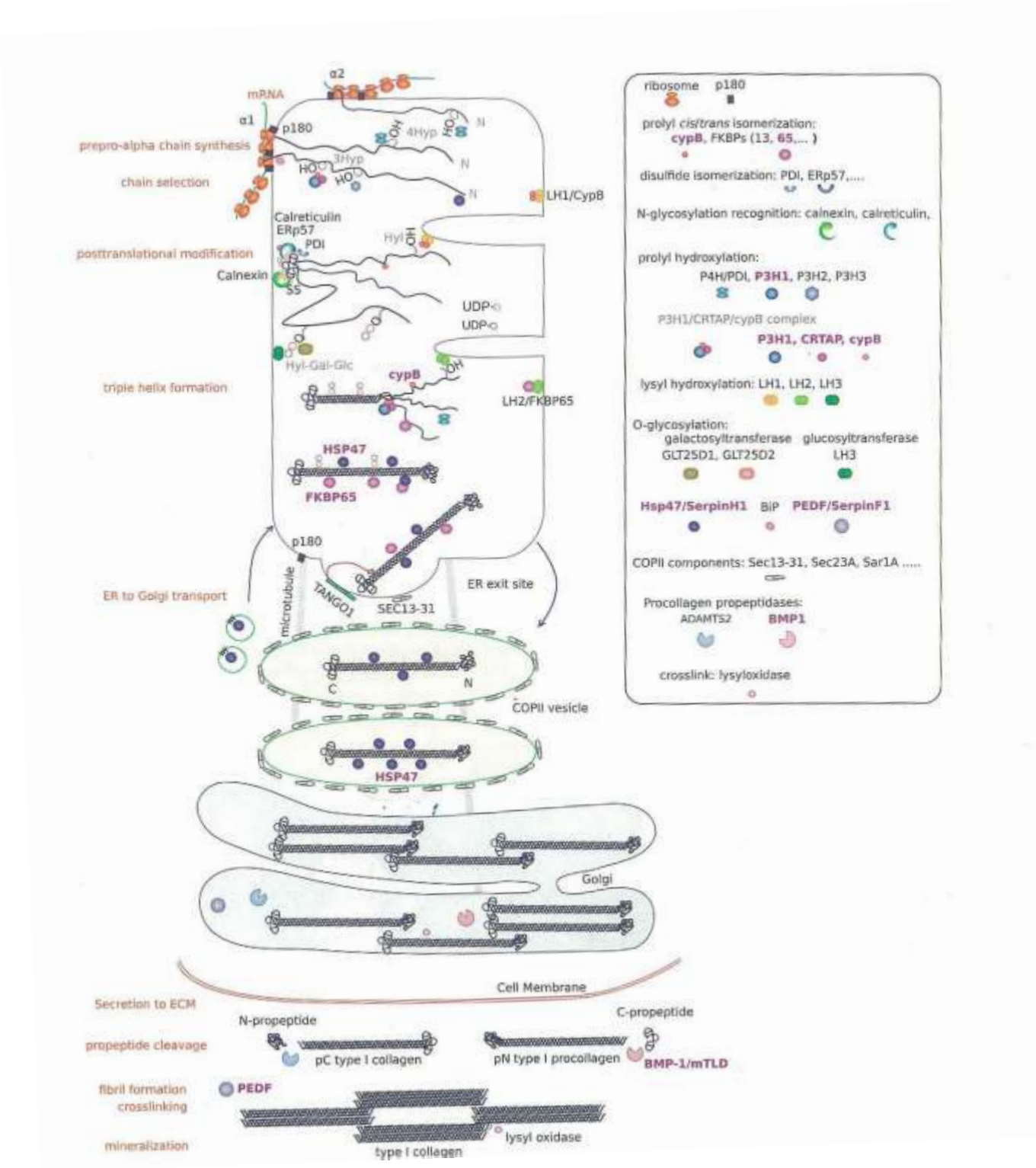

Figure 1 Collagen type I biosynthesis. Proteins involved into recessive OI are indicated in purple color (Shapiro *et al.*, 2013).

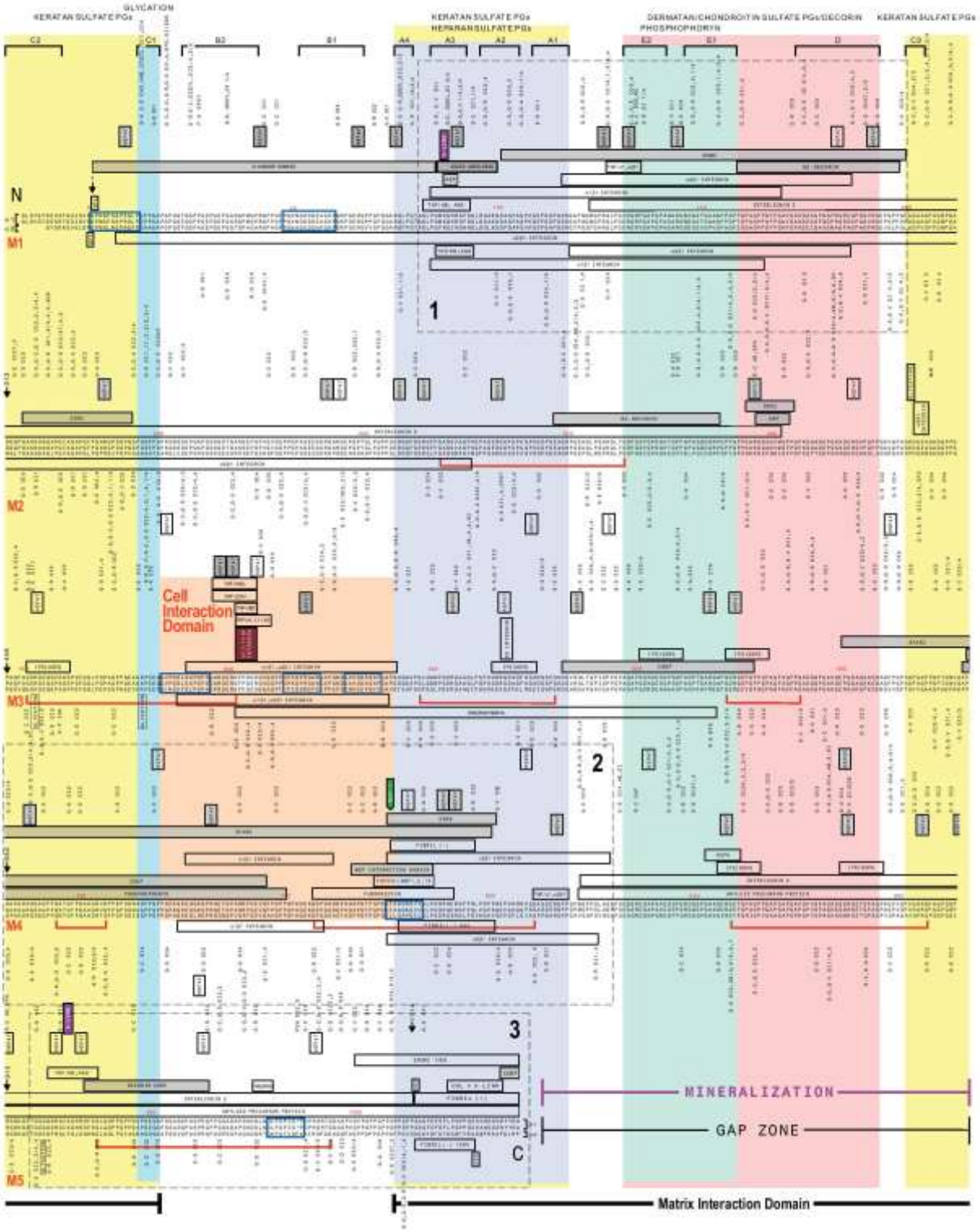

Figure 2 Map of ligand binding site domains of collagen type I fibrils. The legend is represented below (Sweeney *et al.*, 2008).

α1β1 integrin - α1β1 integrin binding sites
α2β1 integrin - α2β1 integrin binding sites
α11β1 integrin - α11β1 integrin binding sites
AD - Mutation associated with arterial dissections
ANE - Mutation associated with aneurysm
CAD - Mutation associated with cervical artery dissections
CAF - Mutation associated with Caffey Disease (Infantile cortical hyperstole)
Col V X-link - Collagen V cross-link site
COMP - Cartilage oligomeric matrix protein binding sites
CPC - C-Proteinase cleavage site
CTW - Mutation associated with connective tissue weakness
DC integrin - Integrin binding site on denatured collagen
DDR2 - Proposed site of discoidin domain receptor 2 binding
Decorin - core Decorin core protein binding site
DI - Mutations associated with Dentinogenesis Imperfecta
[PS] DS PG - Proposed site of dermatan sulfate proteoglycan binding
EDS - Mutations associated with Ehlers-Danlos Syndrome
EDS7L - Mutations associated with Ehlers-Danlos VII-Like Syndrome
ENDO180 - Collagen binding factor
Fibril (+) - C-telopeptide fibrillogenesis nucleation domain
Fibril (-) - Fibrillogenesis inhibition site
Fibril (-) - 50%, 100% - 50 or 100% inhibition of fibrillogenesis by peptides
G X - Substitution mutation
GE-decorin - Guanidine-extracted decorin binding sites
Glycation - Glycation sites
HEP - Heparin, heparan sulfate proteoglycan binding site
HSP47 - Potential heat shock protein 47(chaperone) binding sites
ICA - Mutation associated with intracranial aneurysms
[PS] KS PG - Proposed sites of keratan sulfate proteoglycan binding
MFS - Mutation associated with Marfan Syndrome
MMP-1 - Matrix Metalloproteinase-1 cleavage site
MMP 1, 2, 13 - 1, 2, 13
MMP Interaction Domain - Matrix Metalloproteinase Interaction Domain
N-ANCHOR DOMAIN - Proposed N-anchor Domain
NPC - N-Proteinase cleavage site
OI - Mutations associated with Osteogenesis imperfecta
types 1, 2, 3, 4; ?/US - denotes difficult diagnosis/unusual phenotype
OPA - Mutation associated with Osteopaenia; ? denotes unusual phenotype
OPO - Mutations associated with Osteoporosis; ? denotes unusual phenotype
P-3-H - Site of Prolyl-3-hydroxylation
PHOSPH - Phosphophoryn binding site
SPARC - Secreted protein acidic and rich in cystein binding site
THP/α1, 2, 11-β1 integrin - Triple helical peptide: binds integrin receptors
THP/ANG- - Triple helical peptide: inhibits angiogenesis
THP/ECA+ - Triple helical peptide: promotes endothelial cell activation
THP/FA+; - Triple helical peptide: promotes fibroblast adhesion, substrate for MMPs
THP/HB; ANG- - Triple helical peptide: binds heparin, inhibits angiogenesis
THP/OBD - Triple helical peptide: supports osteoblastic differentiation
μ - Unfolding Proposed Microunfolding domain
VWF - Proposed Site of von Willebrand Factor binding
X-LINK - Intermolecular cross-link site
Orange - Cell interaction domain
Yellow - Keratan sulfate proteoglycans binding region on the fibril

Blue - Glycation region on the fibril
Green - Phosphophoryn binding region on the fibril
Pink - Dermatan/chondroitin sulfate proteoglycans/decorin binding region on the fibril
Black - Brackets Matrix interaction domain
Red Brackets - Delineate clusters of lethal OI mutations on the α2(I) chain
Blue Boxes - Regions containing 3 or more first position glycines on both chains where no human mutations have been reported.

Non-exclusive licence to reproduce thesis and make thesis public

I, Lidiia Zhytnik

(date of birth: 25.06.1989),

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

1.1. reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

1.2. make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

Genetic Predisposition to Osteogenesis Imperfecta: Exome Analysis of Estonian OI Families,

supervised by Dr. Katre Maasalu and Tiit Nikopensius,

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, 26.05.2014