UNIVERSITY OF VERONA

DEPARTMENT OF NEUROSCIENCES, BIOMEDICINE AND MOVEMENT SCIENCES

GRADUATE SCHOOL OF LIFE AND HEALTH SCIENCES

DOCTORAL PROGRAM IN APPLIED LIFE AND HEALTH SCIENCES

XXXI CYCLE

S.S.D. MED/03

A MULTI-OMIC APPROACH TO STUDY AN INTERESTING CASE OF TYPE VI OSTEOGENESIS IMPERFECTA

Coordinator: Prof. Giovanni Malerba

Tutor: Prof. Giovanni Malerba

Co-tutor: Dr. Maria Teresa Valenti

PhD student: Michela Deiana

Niente nella vita va temuto, dev'essere solamente compreso. Ora è tempo di comprendere di più, così possiamo temere di meno.
-Marie Curie-

INDEX

ABSTRACT

Background: Osteogenesis imperfecta (OI) also known as brittle bone disease, is a genetic pathology in which bones do not form properly and therefore are fragile and break easily. Is a heterogeneous congenital heritable disease that mainly affects connective tissues. Nowadays we number 18 types of OI, distinguished into autosomal dominant, recessive and X-linked inheritance. OI type VI is caused by loss-of-function mutations in SERPINF1, which shows recessive inheritance. Lack of the gene product, PEDF, causes an atypical bone mineralization defect determining a unique clinical phenotype.

Aim: the aim of the study is to identify genomic, epigenomic and metabolomic variations that are associated to the disease status in individuals that belong to a nuclear Pakistan family in which is supposed to be segregate type VI osteogenesis imperfecta.

Results: exome sequencing confirmed the consanguinity between the parents and shared regions of homozygosity between affected were observed in chr7, chr12 and chr22. In the hypothesis of Autosomal Recessive disease, any compatible mutation was found, and no clear pathogenic variant have been detected. Thus, we explore compound heterozygosis model, identifying and suggesting as potential candidate CERCAM gene, but it is role in bone homeostasis it is still unknown.

Epigenetic investigation highlights some interesting genes, known to be involved in bone metabolism, such as RXRA (Retinoid X Receptor Alpha), ELK3 (ETS Transcription Factor) and GLI2 (GLI Family Zinc Finger 2).

Metabolomic profiling found 4 modulated pathways: phenylalanine, tyrosine and tryptophan biosynthesis, tryptophan metabolism, pyrimidine metabolism, and Vitamin B6 metabolism.

Conclusions: the future investigation will try to enhance and integrate the results from the present omics (transcriptomic analysis is ongoing) into a context of system biology aimed to depict and clarify the defects and biological processes associated to the disease.

1. INTRODUCTION

1.1 Bone cells and tissue

Bone is a rigid organ that provides support and physical protection to various vital organs of the body. Throughout life, bone is in dynamic balance involving a complex coordination of multiple bone marrow cell types. It contains a relatively small number of cells entrenched in a matrix of collagen fibers that provide a surface for inorganic salt crystals to adhere. Although bone cells account for a small amount of the bone volume, they are crucial for bone functions. Four types of cells are found within bone tissue: osteoblasts, osteocytes, osteoclasts and mesenchymal stem cells. **Osteoblasts** are responsible for forming new bone and are found in the growing portions, including the periosteum and endosteum. Osteoblasts, which do not divide, synthesize and secrete the collagen (mainly type I) matrix and calcium salts. As the secreted matrix surrounding it calcifies, the osteoblast becomes trapped within it; as a result, it changes in structure and becomes an **osteocyte**, the primary cell of mature bone and the most common type of bone cell. Each osteocyte is located in a space called a lacuna and is surrounded by bone tissue. Osteocytes maintain the mineral concentration of the matrix via the secretion of enzymes. Like osteoblasts, osteocytes lack mitotic activity. They can communicate with each other and receive nutrients via long cytoplasmic processes that extend through canaliculi, channels within the bone matrix. The dynamic nature of bone means that new tissue is constantly formed, and old, injured, or unnecessary bone is dissolved for repair or for calcium release. The cell responsible for bone resorption, or breakdown, is the **osteoclast**. Osteoclasts are found on bone surfaces, they are multinucleated, and originate from monocytes and macrophages, two types of white blood cells, not from osteogenic cells. Osteoclasts are continually breaking down old bone while osteoblasts are continually forming new bone. The ongoing balance between osteoblasts and osteoclasts is responsible for the constant but subtle reshaping of bone.

If osteoblasts and osteocytes are incapable of mitosis, then how are they replaced when old ones die? The answer is a fourth category of cells: the **mesenchymal stem cells**. These cells are undifferentiated with high mitotic activity and they are the only bone forming cells able to divide.

1.2 Mesenchymal stem cells - osteoblast differentiation

Mesenchymal stem cells (MSCs), a non-hematopoietic stem cell population first discovered in bone marrow, are multipotent cells capable of differentiating into mature cells of several mesenchymal tissues, such as cartilage, bone, tendon, ligament, and adipose tissue. As common progenitor cells of adipocytes and osteoblasts, MSCs are delicately balanced for their differentiation commitment.[1] The differentiation of MSCs is a two-step process: lineage commitment (from MSCs to lineage-specific progenitors) and maturation (from progenitors to specific cell types). Recently it has been demonstrated that lineage commitment of MSCs involves the regulation of a large number of critical signalling pathways, including transforming growth factor-beta (TGFβ)/bone morphogenic protein (BMP) signalling, wingless-type MMTV integration site (Wnt) signalling, Hedgehogs (Hh), Notch, and fibroblast growth factors (FGFs). Wnt signalling is

involved in many critical biological processes, including development,[2]metabolism,[3] and stem cells maintenance.[4]Binding to the 7-transmembrane domain spanning Frizzled receptor (FZD) and LRP5/6 coreceptors, Wnt ligands stabilize β-catenin via preventing its phosphorylation and degradation. [5] The activation of Wnt signalling has been reported to promote osteogenic differentiation and inhibit adipogenic differentiation of MSCs.[6] Wnt3a has been specifically shown to stimulate osteogenic differentiation through activation of TAZ by PP1A-mediated dephosphorylation.[7] Most recently, it has been demonstrated that YAP/TAZ could mediate alternative Wnt signalling-induced osteogenesis.[8] Animal studies showed that activation of Wnt signalling by overexpression of Wnt10b or lithium supplementation could increase the thickness of trabecular bone.[9] Accordingly, deficiency of Wnt10b leads to a decrease in bone density. Aging-related increase in adipocytes is also thought to be related to the reduction of Wnt10b.[10,11] In addition, the loss of β-catenin in the mesenchyme of the developing mouse uterus was found to cause a switch to adipogenesis in the myometrium.[12] These studies provide strong evidence for the role of Wnt signalling in regulating the balance between adipogenic and osteogenic differentiation of MSCs. Committed MSCs are also known as **osteogenic cells**, they are found in the deep layers of the periosteum and the marrow. They differentiate and develop into osteoblasts. Runx2 and Osterix are considered as master transcription factors regulating osteogenic differentiation of MSCs. [13] During osteoblast differentiation, most signalling pathways investigated so far target Runx2, which in turn induces Osterix expression.[14] Upregulation of Runx2 in MSCs promotes their differentiation potential into immature osteoblasts, while inhibits their lineage commitment to adipocytes.[15] In addition, Runx2 has been shown to be required for the induction of major bone matrix genes in immature osteoblasts, while it is unnecessary for the maintenance of these genes in mature osteoblasts.

1.3 Bone development

Bone is a dynamic organ, able to replace old or disrupted tissue through a remodelling process.[16], Skeleton undergoes a lifelong continuous turnover depending on age, sex, ethnical group and on individual health condition. It has been estimated that the entire skeleton is remodelled every 10 years.[17]

Bone is a replacement tissue, that is, it uses a model tissue on which to lay down its mineral matrix. For skeletal development, the most common template is cartilage. During foetal development, a framework is laid down that determines where bones will form. This framework is a flexible, semi-solid matrix produced by chondroblasts and consists of hyaluronic acid, chondroitin sulphate, collagen (mainly type II) fibers, and water. Throughout foetal development and into childhood growth and development, bone forms on the cartilaginous matrix. By the time a foetus is born, most of the cartilage has been replaced with bone. Some additional cartilage will be replaced throughout childhood, and some cartilage remains in the adult skeleton. By the sixth or seventh week of embryonic life, the actual process of bone development, ossification (osteogenesis), begins.

1.4 Bone ossification in the foetus

Two osteogenic process characterize skeletal development: intramembranous and endochondral ossifications. Ossification is a complex process orchestrated by Mesenchymal Stem Cells (MSCs) able to differentiate either into osteoblasts or

chondrocytes. In intramembranous ossification MSCs direct differentiate into osteoblasts and usually occurs in the flat bones as skull, jaw and in the sub-periosteum areas of long bones. Early osteoblasts occur in cluster called ossification centre, although osteoblasts will be spreaded out by during formation of bone. Osteoblasts secrete osteoid, an uncalcified matrix, which within a few days calcifies trough a mineral salts deposition, thereby entrapping the osteoblasts in the matrix. Entrapped osteoblasts will become osteocytes. As osteoblasts-osteocytes transformation, also osteogenic cells present in the surrounding connective tissue may differentiate into new osteoblasts. Unmineralized bone matrix secreted around the capillaries became a trabecular matrix, while osteoblasts on the surface of the spongy bone form the periosteum. The periosteum represents the protective surface layer above trabecular bone. Intramembranous ossification that begins during foetal development prosecute until adolescence. At birth, the skull and clavicles bones are not fully ossified. Bones plasticity allow skull and shoulders to deform during passage through the birth canal. In endochondral ossification, process that occurs in long bones formation, mesenchymal stem cells differentiate into chondrocytes and secrete a cartilaginous matrix. Cartilage scaffolds are then replaced by bone, increasing the ability to counteract the compression.[19] For example, six-eight weeks after conception, mesenchymal cells differentiate into chondrocytes which form the cartilaginous skeletal precursor of a long bone. The canonical bone development model involves the differentiation of mesenchymal stem cells into a specific cell lineage fate, depending on the time and signalling. As more matrix is produced, more chondrocytes grow in size. Matrix calcification, inhibits nutrients reaching to the chondrocytes. During cartilage grows, capillaries penetrate it, starting the transformation of perichondrium into the bone-producing periosteum. Here, osteoblasts form a periosteal compact bone frame that surround the diaphysis cartilage. From the third month of foetal life, in the periosteal collar bone development create the primary ossification centre, where ossification begins. Chondrocytes and cartilage continue to grow at the ends of the bone, forming the future epiphyses. While length increases, bone is replacing cartilage in the diaphysis. When the foetal skeleton is fully formed, cartilage remains only at the joint surface and between the diaphysis and epiphysis. After birth, matrix mineralization, chondrocytes' death, invasion of blood vessels from the periosteum, and osteoblasts maturation occurs in the epiphyseal regions, and all these activities centres is referred to as a secondary ossification centre.

1.5 Childhood and adolescence

Childhood and adolescence are characterized by a progressive growing (fig. 1) of bone dimensions and by the enhanced bone mass. At birth the weight of the skeleton is around 75-90 g; it reaches 2400-3000 g in the young-adult.[19] The epiphyseal plate is the area of growth in a long bone. It is a layer of hyaline cartilage where ossification occurs in immature bones. The reserve zone is the region closest to the epiphyseal end of the plate and contains small chondrocytes within the matrix, these cells do not participate in bone growth but secure the epiphyseal plate to the osseous tissue of the epiphysis. The proliferative zone is the next layer toward the diaphysis and contains stacks of slightly larger chondrocytes. It makes new chondrocytes (via mitosis) to replace those that die at the diaphyseal end of the plate. Chondrocytes in the next layer, the zone of

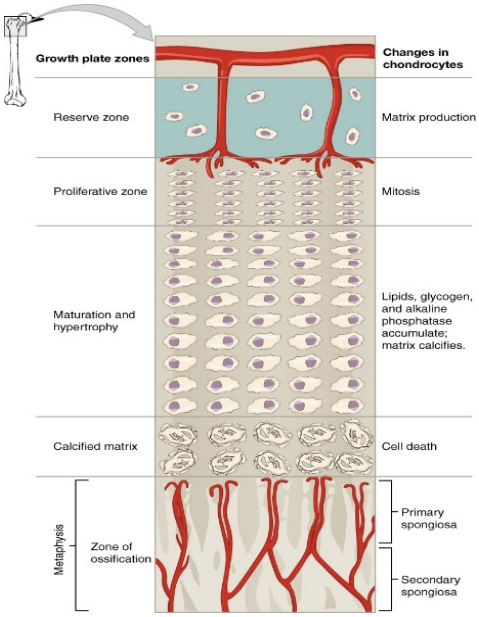

Figure 1: longitudinal Bone Growth

maturation and hypertrophy, are older and larger than those in the proliferative zone. The more mature cells are situated closer to the diaphyseal end of the plate. The longitudinal growth of bone is a result of cellular division in the proliferative zone and the maturation of cells in the zone of maturation and hypertrophy. Most chondrocytes in calcified matrix, the zone closest to the diaphysis, are dead because the matrix around them has calcified. Capillaries and osteoblasts from the diaphysis penetrate this zone; osteoblasts secrete bone tissue on the remaining calcified cartilage. Thus, the zone of calcified matrix connects the epiphyseal plate to the diaphysis. A bone grows in length when osseous tissue is added to the diaphysis. The growth rate is controlled by hormones, bones continue to grow in length until early adulthood. Bone growth includes also increasing in diameter, this can continue even after longitudinal growth ceases and it is called appositional growth. Osteoclasts resorb old bone that lines the medullary cavity, while osteoblasts, via intramembranous ossification, produce new bone tissue beneath the periosteum.

1.6 Bone remodelling

The process in which matrix is resorbed on one surface of a bone and deposited on another is known as bone modelling. Modelling primarily acts during a bone's growth. However, also during adult life, bone undergoes to remodelling, in which resorption of old or damaged bone takes place on the same surface where osteoblasts lay new bone to replace that which is resorbed. Injury, exercise, and other activities lead to remodelling. Normally about 5 to 10 percent of the skeleton is remodelled annually just by destroying old bone and renewing it with fresh bone. The remodelling process takes place under the control of various players, i.e., parathyroid hormone (PTH), calcitonin, vitamin D, growth hormone (GH), steroids, soluble cytokines and growth factors (i.e., Macrophage Colony-Stimulating Factor (M-CSF), Receptor Activator of Nuclear ϰB Ligand (RANKL), Vascular Endothelial Growth Factor (VEGF), Interleukin-6 (IL-6) family.[20] Different stimuli, such as microenvironmental factors, induce osteoblasts to produce RANKL which interacts with its receptor RANK (Receptor Activator of Nuclear _B), expressed by osteoclasts. This interaction determines osteoclast polarization and secretion of the enzymes required for bone resorption. Osteoblasts, which interact with bone matrix through integrins, synthesize type I collagen (that represents 90% of the proteins in the bone matrix). Unmineralized osteoid is composed of type I collagen together with other fibrillar collagens, bone proteins (osteopontin, bone sialoprotein and osteocalcin), proteoglycans, fibronectin, etc. PEDF (pigment epithelium derived factor) is necessary for osteoblast development; it stimulates the expression of osteogenic genes and mineral deposition[21] and it also stimulates the

production of osteoprotegerin, thus inhibiting osteoclast maturation. Mineralization then occurs thanks to the activity of osteoblast phosphatases, releasing phosphates that together with calcium form hydroxyapatite crystals.[22]

1.7 Osteogenesis imperfecta: history and phenotype

Perturbation or disruption of the molecular pathways controlling MSC proliferation and osteogenic commitment may be due to mutations in key genes in bone development. Osteogenesis Imperfecta (OI) also known as brittle bone disease, is a genetic pathology in which bones do not form properly and therefore are fragile and break easily[23].

Is a heterogeneous congenital heritable disease that mainly affects connective tissues. The estimated incidence is approximately 1 per 20,000 live births.[24] The various genetic mutations that cause OI in 85% of cases affect type I collagen- one of the critical components of bone matrix, either quantitatively or qualitatively,[25] The first studies on OI were done in 1788 by Olof Jakob Ekman; it can actually be considered a very ancient pathology. During archaeological studies some Egyptian mummies were found; the description of their skull, teeth and flat bones abnormalities suggests that they affected by OI. In 1979 Sillence et al described OI as a heterogeneous disease and, for the first time, they proposed a classification of at least four distinct types.[24] The hallmarks of the brittle bone dysplasia osteogenesis imperfecta are skeletal deformities and bone fragility that causes low bone mass and bone fractures.[87] The severity of the disease can range from mild to lethally severe. Frequent and multiple fractures typically lead to bone deformities and short stature. Since type I collagen is such an important structural protein in many parts of the body, people with OI may also experience fragile skin, weak muscles, loose joints, easy bruising, frequent nosebleeds, brittle teeth, blue sclerae, and hearing loss. The clinical features can be classified into skeletal and extra skeletal. Skeletal features include excess or atypical fractures, short stature, scoliosis, and basilar skull deformities. Extra skeletal manifestations include hearing loss which is a mixture of conductive and hearing loss and seen in 50% of adults by 50 years and in 5% of children with OI.[26] Dentinogenesis Imperfecta (DI), consist in small deformed teeth, which present opalescent and opaque dentin. Malocclusion and DI, are the main dental abnormalities which may occur. The phenotype is variable also within the same patient, with some teeth appearing normal and others being affected. The sclerae may be blue or gray.[27] Connective tissue abnormalities may result in dislocation of head of the radius and the joint. 36% of patients that manifest hypercalciuria, may result in renal calculi. Cardiovascular involvement leads to aortic root dilatation and mitral valve prolapse. Neurological manifestations include macrocephaly, hydrocephalus, basilar invagination and cervical spine kyphosis.[27]

1.8 Genetic classification and pathophysiology of OI

Mutations in genes coding for the α1 and α2 chains of type I collagen (COL1A1 and COL1A2) were associated to Osteogenesis imperfecta disease in 1980. Mutations in these genes are the most common cause and are transmitted as autosomal dominant traits. They cause a quantitative loss of α1(I) chain and, with few exceptions, they cause a mild form of OI (type I) because the other allele

produces collagen of normal quality and structure.[28] Other mutations in either COL1A1 or COL1A2, most often glycine substitutions in the Gly-X-Y repeat, cause structural defects of the collagen triple helix. They exert a dominant negative effect on the normal collagen chains upon trimers formation, and result in moderate, severe or lethal OI. Type I collagen is the most abundant protein of bone, skin, and tendon extracellular matrices; in bone and theeth it plays a fundamental role in the mineralization process, (fig. 2 type I collagen synthesis and processing).[29]

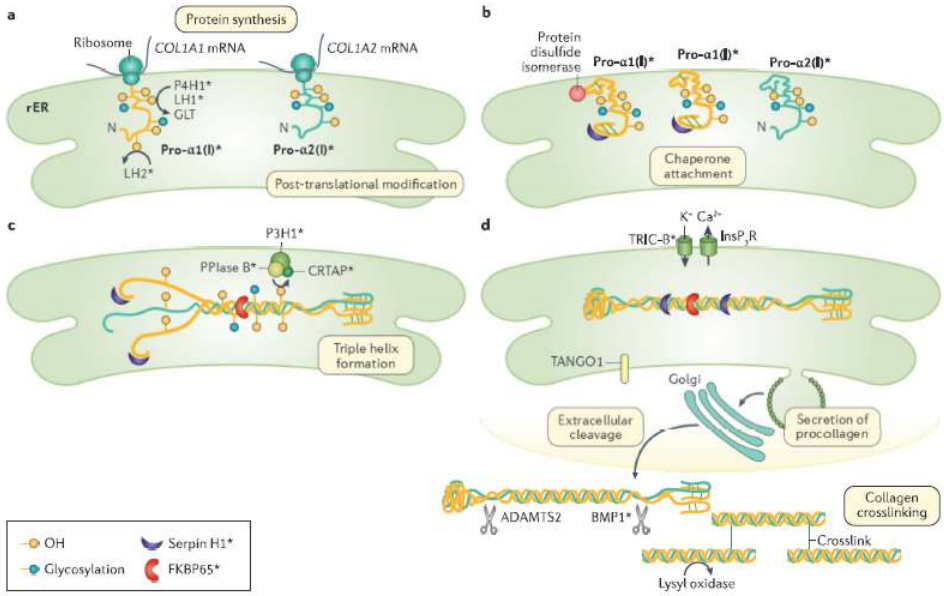

Figure 2: type I collagen synthesis and processing. A) translation and post-translational modifications of pro α1 and pro α2. B) Interaction with molecular chaperones to prevent premature triple helix formation. C) triple helix formation that comprises two pro α1 chains and one pro α2 chain. D) secretion of protocolagen, extracellular cleavage to collagen and crosslinking. ADAMTS2, a disintegrin and metalloproteinase with trombospondin motifs 2; BMP1, bone morphogenetic protein1; CTRAP, cartilage-associated protein; GLT, galactosyltransferase 1 ; $InsP_3R$, inositol-1,4,5-triphosphate receptor; LH, lysyl hydroxylase; P3H1, propyl 3-hydroxylase 1; P4H1, propyl 4-hydroxylase 1; PPIase B, peptidyl-prolyl cis-trans isomerase B; rER, rough endoplasmic reticulum; TANGO1, transport and Golgi organization protein 1; TRIC-B, trimeric intracellular cation channel type B (figure from Marini e al)[29].

In the past decade, a wide variety of genes encoding proteins involved in type I collagen synthesis, processing, secretion and post-translational modification, as well as in genes coding for proteins that regulate the differentiation and activity of bone-forming cells have been shown to cause osteogenesis imperfecta.[29] In 2006 The first gene causing recessive osteogenesis imperfect was identified, opening the way to a burst of exciting new information about the genetics and mechanism of this bone dysplasia.

The OI classification initially included four phenotypes (I-IV) involving COL1A1 and COL1A2 mutations.[30] The large number of causative genes discovered since 2006 has complicated the classic classification of the disease, and, although a new genetic classification system is widely used, it is still debated. Defects in proteins with very different functions, ranging from structural to enzymatic and from intracellular transport to chaperones, have been described in patients with osteogenesis imperfecta**.** Nowadays we number 18 types of OI, distinguished into autosomal dominant, recessive and X-linked inheritance. An appropriate

pathophysiology-based grouping of types, as proposed by JC Marini et al., appears helpful for researchers and clinicians (table 1, genetic classification of OI).

Mutaded gene	Encoded protein	Inheritance	Localization	Severity	OI type
Impairment of collagen suntheis and structure					
COL1A1	Collagen α1	AD	matrix structural component	Mild to lethal	I,II,III or IV
COL1A2	Collagen α2	AD	matrix structural component	Moderate to lethal	I,II,III or IV
Compromised bone mineralization					
IFITM5	BRIL	AD	bone-restriced interferon-induced transmembrane like protien	variable severity	V
SERPINF1	PEDF	AR	collagen-binding protein/pigment epithelium derived factor	Moderate to severe	VI
Abnormal collagen post-translational modification					
CTRP	CTRAP	AR	endoplasmic reticulum	Severe to lethal	VII
P3H1 (LEPRE1)	P3H1	AR	endoplasmic reticulum	Severe to lethal	VIII
PPIB	PPIase B	AR	endoplasmic reticulum	Moderate to severe	IX
Compromised collagen processing and crosslinking					
SERPINH1	HSP47	AR	endoplasmic reticulum-golgi	Severe to lethal	X
FKBP10	FKBP65	AR	endoplasmic reticulum	Moderate to severe	XI
PLOD2	LH2	AR	endoplasmic reticulum	Moderate to severe	no type
BMP1	BMP1	AR	endoplasmic reticulum	Moderate to severe	XII
Alterate osteoblast differentiation and function					
SP7	Transcription factor SP7/	AR	nucleus	Mild to moderate	XIII
TMEM38B	TRIC-B	AR	cation channel	Moderate to severe	XIV
WNT1	WNT1	AR/AD	secreted signal molecule	Moderate to severe	XV
CREB3L1	OASIS	AR	endoplasmic reticulum-golgi	Severe	XVI
SPARC	SPARC/osteone	AR	matrix	Moderate to severe	XVII
MBTPS2	S2P	XLR	endoplasmic reticulum-golgi	Moderate to severe	XVIII

Table 1: genetic classification of osteogenesis imperfecta

1.9 Defects in bone mineralization

Bone mineralization defects have been associated with mutations in two genes: IFITM5, which encodes for BRIL (Bone Restricted Ifitm-Like protein) and SERPINF1, which encodes for PEDF (pigment epithelium-derived factor) The causative mutations are directly responsible for OI type V and VI, respectively, (fig. 3 Defects in bone formation and mineralization in osteogenesis imperfecta).[29]

All cases of type V OI are caused by a dominantly inherited or "de novo" recurrent heterozygous mutation (c.-14C>T) in the 5'-UTR of IFITM5, a transmembrane protein enriched in osteoblasts during mineralization.[31] Patients with type V OI have a moderately severe bone dysplasia and fracture incidence, including vertebral compressions and, often, scoliosis. Features of type V include radial head dislocation, ossification of the forearm interosseous membrane, and hyperplastic callus. There is also a radiographically dense band prominent in the forearm metaphyses. [32,33]

Loss-of-function mutations in SERPINF1 are responsible for **OI type VI,** which shows recessive inheritance. Lack of the gene product, PEDF, causes an atypical bone mineralization defect determining a unique clinical phenotype.[34] Patients with type VI OI show bone dysplasia, with clinical characteristics and bone histology distinct from type V OI. Types V and VI osteogenesis imperfecta, although genetically different, share the distinctive feature of causing primary defects in endochondral bone ossification and/or mineralisation.

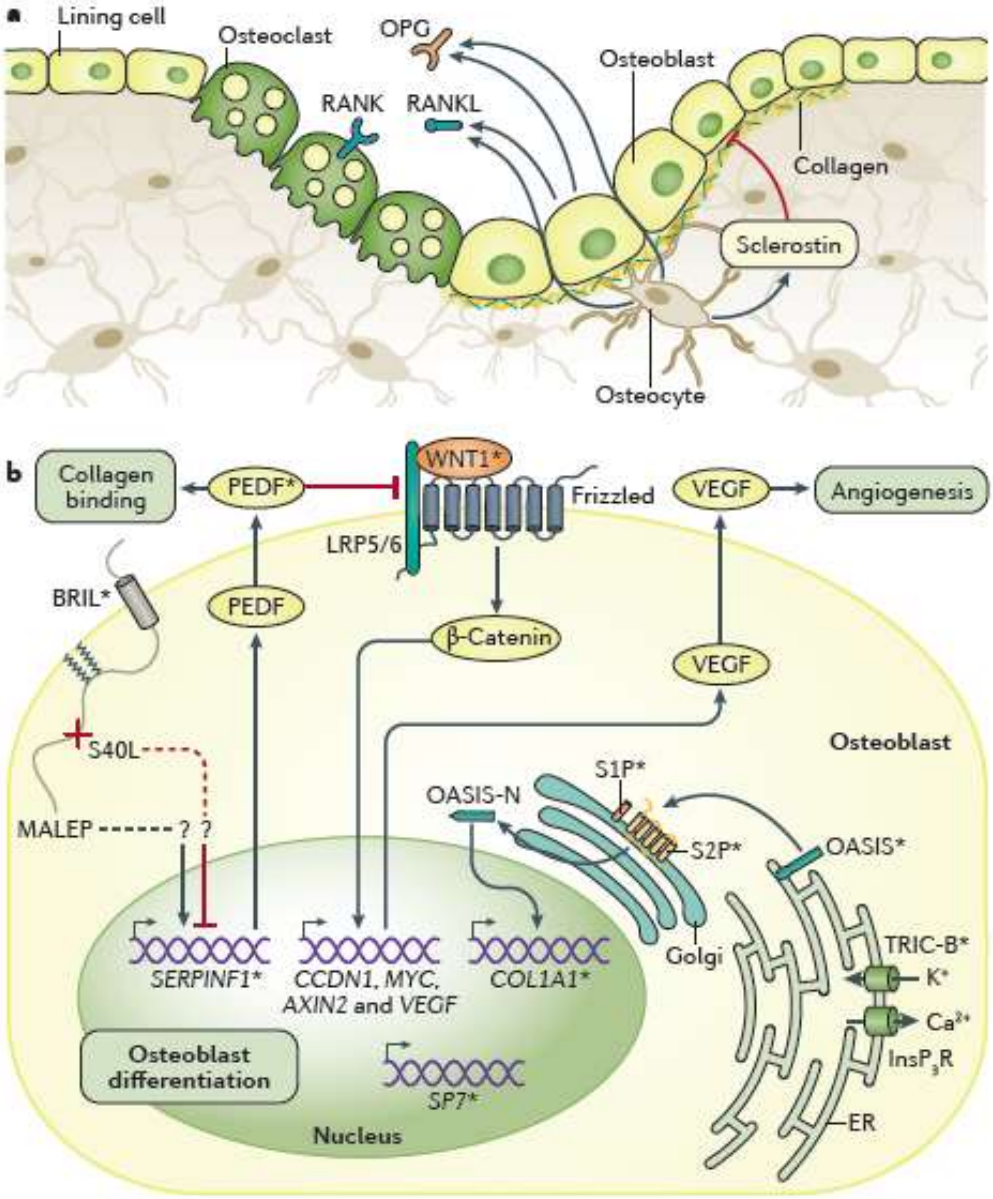

Figure 3: defects in bone formation and mineralization in osteogenesis imperfecta. a | Bone remodelling. Bone formation consists of the secretion of bone extracellular matrix components (mainly type I collagen) by osteoblasts. The unmineralized bone matrix (osteoid) subsequently becomes mineralized. In addition, osteoblasts and osteocytes release many cytokines, including receptor activator of nuclear factor-κB ligand (RANKL; also known as TNFSF11) and osteoprotegerin (OPG), which regulate bone resorption by osteoclasts. RANKL acts on osteoclast precursor cells by binding to receptor activator of nuclear factor-κB (RANK; also known as TNFRSF11A) on their surface, thereby favouring their differentiation to osteoclasts. OPG, by interacting with RANKL, prevents the binding of RANKL to RANK. Some osteoblasts become embedded in the mineralized bone matrix and differentiate to osteocytes, which produce, among other factors, sclerostin, an inhibitor of the WNT pathway that is known to stimulate bone formation by stimulating osteoblast activity. b | WNT1 is a secreted ligand that stimulates the transcription of genes involved in osteoblast differentiation by interacting with the receptors low-density lipoprotein receptor-related protein 5/6 (LRP5/6) and Frizzled on osteoblast precursor cells. Endopeptidase S2P is a protease in the Golgi membrane that is involved in regulated

intramembrane proteolysis (RIP) of transcription factors, such as old astrocyte specifically induced substance (OASIS), which is transported from the endoplasmic reticulum (ER) membrane for processing in times of ER stress. Pigment epithelium-derived factor (PEDF) is a multifunctional secreted protein with anti-angiogenic activity; bone-restricted interferon-induced transmembrane protein-like protein (BRIL) is a transmembrane protein involved in mineralization. Crosstalk between the two proteins had been elucidated following mutations identified in patients with osteogenesis imperfecta. Dashed lines indicate unknown pathways. InsP3R, inositol-1,4,5-triphosphate receptor; OASIS-N, N-terminal part of OASIS; MALEP, methionine, alanine, leucine, glutamate and proline pentapeptide; S40L, p.Ser40Leu substitution; TRIC-B, trimeric intracellular cation channel type B; VEGF, vascular endothelial growth factor (figure from Marini e al)[29].

1.10 Phenotypical aspects OI type VI

Bone histology is remarkable for broad bands of unmineralized osteoid and a fish-scale pattern under polarized light[25] Patients affected by OI type VI do not have fractures at birth, but starting from ≥6 mo. age, they suffer from frequent fractures, progressive bone deformity, vertebral compressions and scoliosis. Growth deficiency in type VI is moderately severe, sclerae are white and teeth are normal.[35] A variety of recessive null mutations in SERPINF1 has been reported. [36,37,38,39] Serum PEDF is virtually absent in type VI OI patients, while normal PEDF values are reported in type V OI, as well as in OI caused by collagen defects (types I, III, and IV).[39] Children with type VI OI have elevated serum alkaline phosphatase levels; with a mean of 409 U/L (range 200-650 U/L).[29] Other biochemical parameters of bone and mineral metabolism are within the reference range.[35]

1.11 Molecular and cellular features of OI type VI

PEDF is a ubiquitously expressed secreted protein best known as a potent anti-angiogenic factor that inhibits tumour growth and metastasis. [40,31] This protein is able to bind two distinct sites on type I collagen[41], and this binding is critical to its anti-angiogenic function.[42,43] One binding site overlaps the heparin and heparan sulphate proteoglycan binding site with the α1(I) C-terminal major ligand binding region. The second collagen-binding site, in the α1(I) N-terminal, overlaps integrin collagen-binding sites. The variety of recessive null mutations in *SERPINF1* determine the absence of PEDF in the serum.[44] In bone, PEDF acts at many levels to maintain bone homoeostasis and regulate osteoid mineralisation. PEDF positively affects osteoblast development by favouring the expression of osteogenic genes and mineral deposition. It inhibits osteoclast maturation by stimulating osteoprotegerin expression.[45] The absence of PEDF increases osteoclast number and bone resorption by favouring RANKL binding to the osteoclast RANK receptor. Moreover, it impairs the proper transition of osteoblast to osteocyte, consequently causing a defective mineralization that leads to brittle bone.[46] Importantly, an intact collagen-binding site on PEDF is required for its anti-angiogenic activity. PEDF, appears also to be involved in a poorly characterized signalling pathway that includes the plasma membrane receptor BRIL, encoded by *IFITM5*.[34]

1.12 Multi-omics analysis

The advent of genotyping arrays and next generation sequencing (NGS) platform enabled large-scale genome-wide studies opening the arrival of genomics, transcriptomics and metabolomics.

Osteogenesis imperfecta had been known since the early 1980s as a disease caused by mutations in either of the genes encoding type I collagen (COL1A1 and COL1A2). Since 2006 the NGS revolution applied to OI, combined with the traditional approaches, has unravelled new mutations in collagen-related genes with different inheritance patterns.[47] Fifteen novel disease OI associated loci have been discovered in ten years' time.[48]

The appearance of "omics" to a molecular term implies a comprehensive, or global, assessment of a set of molecules.[53]

Genomics is an interdisciplinary field of science focusing on the structure, function, evolution, mapping, and editing of genomes. A genome is a complete set of DNA, including all genes of an individual. In contrast to genetics, which refers to the study of individual genes and their roles in inheritance, genomics aims at the collective characterization and quantification of genes. Likewise, **transcriptomics** refers to all RNA molecules in one cell or a population of cells. Despite the DNA, that remains unchanged in every cell of one individual, RNA molecules reflect the expression profile a set of transcripts in of a specific cell type. **Proteomics** is used to quantify peptide abundance, modification, and interaction. **Metabolomics** simultaneously quantifies multiple small molecule types, such as amino acids, fatty acids, carbohydrates, or other products of cellular metabolic functions. Metabolite levels and relative ratios reflect metabolic function, and out of normal range perturbations are often indicative of disease. Quantitative measures of metabolite levels have made possible the discovery of novel genetic loci regulating small molecules, or their relative ratios, in plasma and other tissues.[49,50]Additionally, metabolomics in combination with modelling has been used extensively to study metabolite flux. Associated technologies include Mass Spectrometry (MS)-based approaches to quantify both relative and targeted small molecule abundances. [51,52]

The omics field has been driven largely by technological advances that have made cost-efficient, high-throughput analysis of biologic molecules possible.[53] Taken together, these understanding provides a rationale for the development of system biology technologies that involve the integration of different omics data to identify molecular patterns associated with disease.[53] The advantage of multi-omics approach is to offers the opportunity to understand the flow of information that underlies disease.

2. AIMS OF THE STUDY

The aim of the study is to investigate the possible defects that might be detected at genomic, epigenomic and metabolomic level, related to the disease status in individuals that belong to a nuclear Pakistan family in which is supposed to segregate a form of type VI osteogenesis imperfecta (see fig. 4).
We performed a multi-omics approach in order to:

- Discover possible genetic defects by exome sequencing
- Assess epigenetic modifications associated with the disease status of individuals, in particular we investigated DNA methylation using Next Generation Sequencing technology (NGS)
- Detect the metabolic pathways associated with the disease performing an untargeted metabolomic scan.

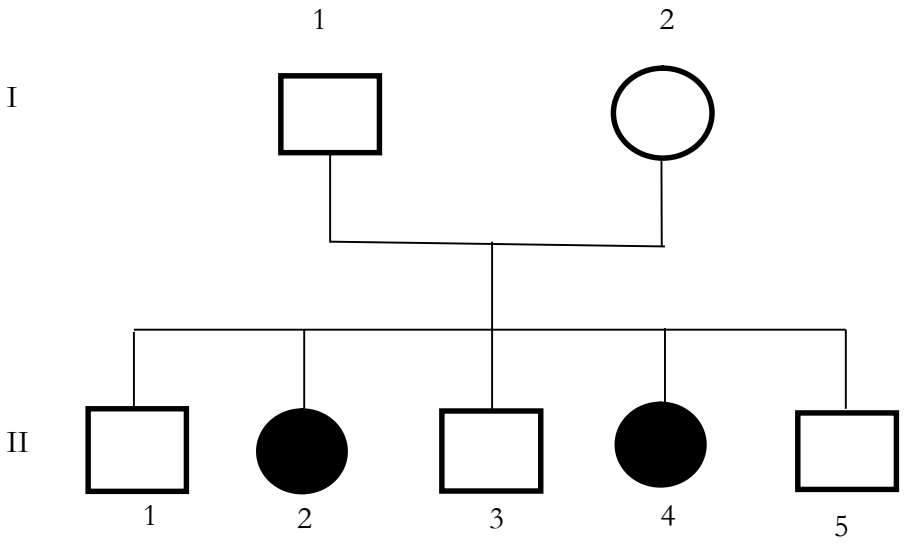

Figure 4: pedigree of the family that has been investigated at genetic, epigenetic and metabolomic levels.

3. MATERIALS AND METHODS

3.1 Subjects

The family settled in Italy in 2009, it is composed by 2 parents (individuals I-1, I-2) and 5 children, 3 males (individuals II-1, II-3, II-5) and 2 females (individuals II-2, II-4) (in fig.4 pedigree of the family). Since 2010 the family, and in particular the affected sisters (individuals II-2, II-4), have been monitored and supervised by an interdisciplinary team of Italian paediatricians and bone experts at Verona University Hospital (AOUI). The two daughters present all phenotypical traits characteristic of Osteogenesis Imperfecta type VI. At birth both babies (individual II-2 and II-4) were fractureless, but at the age of 6 months they manifested the first fractures and bone mineralization defects.

In the absence of any causative mutation in known candidate genes that could explain that pathology we decided to proceed with a multi-omics approach.

3.2 Bone biopsies - hystomorfometric analysis

Iliac crest bone biopsy was collected from the affected children. The samples were fixed in 70% and 100% ethanol and embedded undecalcified in methyl-methacrylate resin (Merck 800590). Bone sections were cut by using a microtome (Polycut S, Leica Microsystems) equipped with a carbide-tungsten blade, stained with Goldner's stain and mounted on microscope slides for histomorphometric measurements. Measurements were performed by means of an image analysis system consisting of an epifluorescent microscope (Leica DM2500) connected to a digital camera (Leica DFC420 C) and a computer equipped with a specific software for histomorphometric analyses (Bone 3.5, Explora Nova, France).

3.3 Peripheral blood sample collection

Peripheral blood mononuclear cells (PBMCs) were obtained by venepuncture from 10 mL of whole blood from parents, two healthy brothers, subject 3 and 4 (subject 7 was excluded from the study because on this time he was still a baby) and the affected sisters. Blood were diluted 1:1 with Phosphate Saline Buffer (PBS) 1X and then stratified in 10 mL of Ficoll solution to allow the separation of PBMCs from whole blood. White blood cells ring was collected and washed two times whit PBS 1x to eliminate traces of contaminant solutions; PBMCs pellet was obtained of 10 minutes centrifugation at 12000 g.

3.4 Sera collection

Sera samples were collected from all member of the family. Each serum sample was obtained from 5 mL of peripheral blood by centrifugation at 400x for 15 minutes at 4°C and after it was harvested and frozen in aliquots at − 80 °C until use.

3.5 Mesenchymal stem cell isolation

We collect 50 mL of whole blood from the mother, two healthy brothers (subject 4 and 5) and the affected sisters. MSCs were isolated from 15 mL of heparinized blood by two gradient centrifugation (Ficoll GE). The enriched cell pellet (PBMCs) was mixed with 5 ml of additional peripheral blood (from the same patient) as red cells are necessary to apply 200µl antibodies cocktail (RosetteSep

Mesenchymal Enrichment Cocktail; StemCells); this mix contains bi-specific antibody complexes against red blood cells (glycophorin A) and CD3, CD14, CD19, CD38 and CD66b positive cells (Stem Cell Technologies Inc). These antibody complexes crosslinked the unwanted cells with red blood cells in the sample, causing them to pellet together when centrifuged during the second Ficoll. The pellet enriched from MSCs was washed in PBS and then cultured in 24 well plates with Mesencult Basal Medium (StemCell) containing 10% of the Stimulatory Supplements for hMSCs (StemCell).[54]

3.6 Phenotype analysis of MSCs

We investigated the cell phenotype analysing some mRNA markers. Since positive MSCs are expected express CD90 and CD105 transcripts and not to express CD3, CD14, CD19, CD45, CD34, we studied the expression of the above cited mRNAs.[54] This method allows the phenotypic analysis of small numbers of cells obtained with stringent stem cell purification techniques.

3.7 DNA and RNA extraction

We performed DNA extraction from PBMCs or MSCs pellet QIAamp® Blood Mini Kit according to datasheet (see supplementary, point A). Part of MSCs dry pellet was used to RNA extraction. Procedure was performed using RNeasy Mini Kit according to manufactural protocol (see supplementary, point B).

Quantification of nucleic acids was performed in two steps: in the first step we used ThermoScientific™ NanoDrop™ spectrophotometers to evaluate purity and concentration, while in second step we used Qubit 4 Fluorometer (Invitrogen) also to check the integrity. Nucleics acid will be used for whole exome and sequencing and transcriptomic analysis.

3.9 Whole Exome sequencing

Samples were sent to a facility centre for genomic at University of Verona specialized that prepared library fragments of whole exome. Exome capture was performed starting from 500 ng of library as input material with Nextera rapid Expanded Exome Capture Kit (Illumina, San Diego, CA, USA), able to target exons, UTRs and miRNA DNA fragments. It captures 201,121 exons, the 96.0% of CCDS, for a total target region of 62 Mb size (for each sample ware sequenced 60M of fragments). Whole exome library was sequenced with an Illumina NextSeq 500 sequencer (Illumina Inc., San Diego, CA, USA) and 150-bp paired-end sequences were acquired.

3.10 Methylation sequencing

Methylation analysis by sequencing was performed by the following steps: 1) DNA extraction from and enzymatic fragmentation (450-500 bp), 2) library preparation, 3) conversion of unmethylated 5-C into T by bisulphite treatment and 4) sequencing. 1) 500 μg of DNA extracted from MSCs were subjected to library preparation using the SeqCap® Epi CpGiant Probes (Roche) kit. We performed the sequencing using the as reported in the datasheet. For each sample has been picked up 500 ng of DNA to perform the enzymatic fragmentation. 2) After fragmentation we proceed immediately to adapter ligation and library amplification. 3) Only fragments with a size of ~450-500 bp were selected for bisulphite conversion step (allow the conversion of unmethylated cytosine in

thymine, while methylated one remain cytosine). To evaluate conversion rate, we added also a control in 1% concentration. Fagoλ DNA present a unmethylated genome and we used his genome as positive conversion control. At this point, samples marked with different adapters were polled together for the hybridization step. It has been used a probe that allows to interrogate >5.5 million methylation sites per sample at single-nucleotide resolution and covered a target region of 80 Mb. Captured library was amplified using LM-PCR and quantified to evaluate concentration, size distribution and quality. 4) Library, prepared as described before, was sequenced using an Illumina NextSeq 500sequencer (Illumina Inc., San Diego, CA, USA) and 150 bp paired-end sequences were generated.

3.11 Metabolomic profile

We investigate an untargeted profile, that describes the global profiling of small molecules in a biological system without bias. The aim of untargeted metabolomics is to determine which of these features is dysregulated between two or more samples groups, and in our case between healthy and affected subjects. For each subject we used 100 μL of plasma, 200 μL of acetonitrile was added and then vortexed for 1 minute. Shake the mixture for 20 minutes, centrifugate at 10000 rpm for 10 minutes at 4 ° C. Take 150 μL of supernatant and transfer it in a glass vial. Plasma metabolites were detected using liquid chromatography combined with electrospray ionization tandem mass spectrometry (HPLC–ESI-MS/MS). Detail above methods are reported by Syslová et al. 2011.[55]

The injection volume was 5 μl. Metabolomic profile was measured by HPLC-Orbitrap in positive and negative polarity. Data was acquired by ThermoSieve v2.2 software (Thermo Scientific™) and the analyses were performed using XCMS software (v. 3.7.1).

3.12 Written informed consent

Written informed consent was obtained from all subjects of the family.

3.13 Bioinformatic analyses

3.13.1 Exome sequencing

The first level of QC was at read level filtering the reads at single base quality score, per base sequence content, duplication rate and over-represented sequences. Sequenced data (fastq files) were processed with 2 tools: 1) Scythe for adapters removal and 2) Sickle for low base quality trimming in 3' and 5' extremities. Then aligned with the Burrows Wheeler Aligner software BWA (version 0.5.7)[56] using hg38 human genome assembly as a reference. The family members were subject to **identity by descent** (IBD) estimation and **run of homozygosity (ROH)** test. With these tests we investigated allele frequencies, the degree of genetic similarity between individuals, and the presence of long chromosomal regions presenting homozygosity within the same individual and pair of affected and unaffected individuals.

Variant sites were called using the GATK Haplotypecaller module, following the indications (best practices) of GATK. Single nucleotide variant calls were filtered by using the variant quality score recalibration method.[57] Each variant was annotated reporting, when possible the associated gene, location, predicted functional effects, amino acid changes. Multiple pathogenicity prediction programs such as SIFT[58], Polyphen2[59], LRP, MutationTaster, FATHMM, PROVEAN, MetaSVM, MetaLR, M-cap, CADD prediction scores and gnomAD

(exome genomes Minor Allele Frequencies) were used to examine the impact of identified variants. All the predictors used are implemented in the Annovar software (v. 2018Apr16).

3.13.2 Methylation sequencing

Quality control was performed as described in 3.13.1 section. Bisulphite sequencing (BS-Seq) data were aligned with Bismark Bisulphite Mapper (v.0.19.0). The pipeline of the analyses can be summarized in three main steps: I) preparation of the converted version of reference genome step (all Cs to Ts, and Gs in As), II) reads alignment step and III) extraction of methylation level of each cytosine. Given the step I, all the alignments in step II used as reference two genomes (converted and unconverted). Only reads with best alignment score were kept. Multiple reads were discarded. A summary report along with a file containing single base methylation information (CG, CHG or CHH contexts) were prepared. In order to capture bisulphite converted DNA, probes are designed to hybridize both strands of fully methylated, partially methylated and fully unmethylated derivatives of the genomic target and then pool them together. The identification of differentially methylated regions (DMR) was carried out using DMRfinder software pipeline (v 0.3).[60]

DMR results were processed by Annovar (2018Apr16) annotation tool performing a **gene-based** annotation to identify SNPs or CNVs, **region-based** annotation to identify variants in specific genomic regions (for example predicted transcription factor binding sites, etc.) and a **filter-based annotation to** identify variants that are documented in specific databases such as variant reported in dbSNP.

Differential methylated genes were also submitted to the STRING portal (https://string-db.org/) for independent inspection of their predicted connections.

3.13.3. Metabolomic profile

Metabolomic untarget analysis was performed using XCMS software (v. 3.7.1), running under the R package (v. 3.5). XCMS identifies features whose relative intensity varies between sample groups (i.e. affected and controls) and calculates p-values as well as fold changes regarding differences in each of features (that are likely to correspond to metabolites in most cases). Dysregulated features were reported as plots in a so-called "mirror plot", reporting intensities of signals among groups according to statistical thresholds (p-value ≤ 0.05, fold change $\geq$ 1,5 The identification of features was based on METLIN database (https://metlin.scripps.edu/). Statistical analysis of metabolite-associated features was performed with MetaboAnalyst (v.4.0)

4. RESULTS

4.1 Previous results of genetic, hystomorfometry and biochemical markers

As after mentioned, the family have been monitored from 2010, here we reported few values of OI clinical markers harvest over the time.

Since type VI OI is characterized by the absence of PEDF protein, increased alkaline phosphatise levels, while the other biochemical parameters of bone and mineral metabolism remain unchanged. It has been reported that PEDF is present in the plasma of normal individuals at a concentration of approximately 2,76-28,57μg/ml.[61] PEDF serum levels have been assessed in five of six family members, while ALP was dosed only in the affected subjects (II-2 and II-4) (table 2).

SUBJECT #	PEDF conc. (2.76-28.57μg/ml μg/mL)	ALP conc. (2-16 μg/L)	STATUS
Father (I-1)	9.68	-	Healthy
Mother (I-2)	3.98	-	Healthy
Male Child (II-1)	7.51	-	Healthy
Female Child (II-2)	2.31	21	Affected
Male Child (II-3)	4.67	-	Healthy
Female Child (II-4)	3.08	68	Affected
Male Child (II-5)	-	-	Healthy

Table 2: concentration of PEDF and ALP in the family members (see fig.4)

PEDF concentration levels within the normal range were found in II-4 affected girls, while II-2 showed a lower expression compared to normal range. However, considering that in all other type VI patients described so far, included those supervised at Verona AOUI PEDF levels were close to 0 μg/mL, PEDF concentration of individual II-2 has been considered within the normal range.

A complete panel of known OI related genes has been previously sequenced by dr. Alberto Gandini in all family individuals (table 3, list of investigated genes).

Gene	Chomosome	Length of target region (bp)	OI type	Inheritance
LEPRE1	Chr 1	4791	VIII	AR
PLOD2	Chr 3	5079	-*	AR
COL1A2	Chr 7	8063	I,II,III,IV	AD
BMP1	Chr 8	5263	XII	AR
TMEM38B	Chr 9	2382	XIV	AR
LRP5	Chr 11	6301	-*	AR
IFITM5	Chr 11	835	V	AR
SERPINH1	Chr 11	2626	X	AR
LPR6	Chr 12	11256	-*	AR
SP7	Chr 12	3320	XIII	AR
WNT1	Chr 12	2488	XV	AR,AD
PPIB	Chr 15	1283	IX	AR
SMPD3	Chr 16	5728	-*	AR
COL1A1	Chr 17	8528	I,II,III,IV	AD

Gene	Chomosome	Length of target region (bp)	OI type	Inheritance
FKBP10	Chr 17	3393	XI	AR
SERPINF1	Chr 17	1941	VI	AR
LEPREL4	Chr 17	3010	-*	AR
SERPINB5	Chr 18	2990	-*	AR

Table 3: candidates investigated genes. Table reports the list of candidate genes that have been sequenced in family members. Genes were sequenced in their CDS and UTR regions. * Genes not directly associated to OI spectrum, but related to bone diseases such as Bruck syndrome, primary osteoporosis, bone mass decrease, etc.

No mutation showing a segregation of mendelian pattern associated with the disease was detected for any gene.

In 2012 our laboratory performed a hystomorphometric analysis (fig.5) which confirmed the presence of the "fish-scale pattern" originally described by F. Gloriuex[62], a peculiar feature of OI type VI. Fig. 6 show radiographs of subject II-2 (according to pedigree figure) in which we can appreciate bone thickness al limbs and spinal level

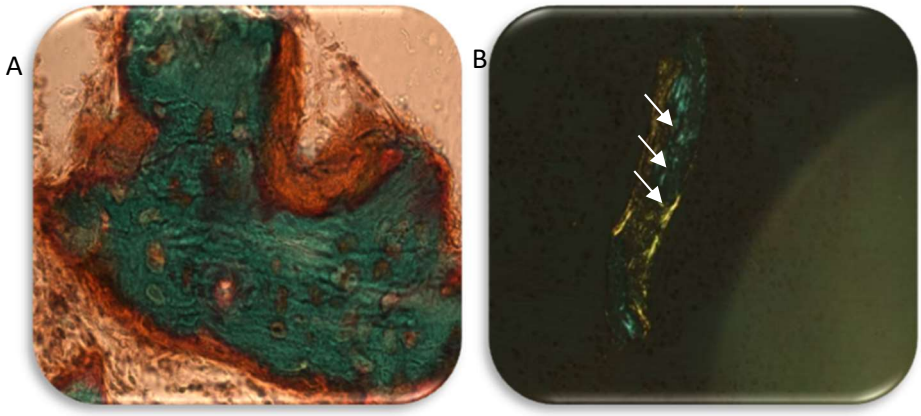

Figure 5: hystomorphometric analysis of affected children performed by Goldener staining. Figure A correspond to subject II-2, while figure B correspond to subject 6. Green part in the staining represent mineralized bone, while pink staining represents unmineralized osteoid. White arrows in figure B indicates fish scales.

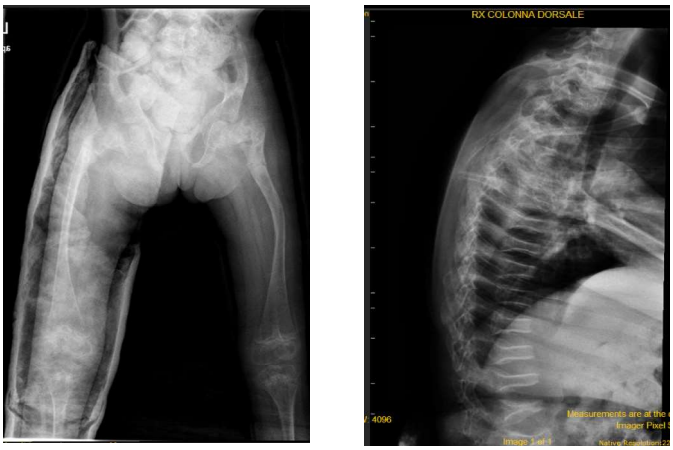

Figure 6: radiographs of lower limbs and spine of subject II-2

4.2 Whole exome sequencing

Whole exome sequencing (WES) on DNA samples was performed on all the family members with the exception of II-5 because too young at time of blood collection. Table 4 reports overall mean and median depth value of sequencing of the target region for each sample.

Genetic similarity tested by identity by descend (IBD), was estimate in a sample of unrelated Italian individuals showing an average score 0,005 (proportion of IBD, pi_IBD). IBD estimation revelled that the similarity between the two parents was higher among the population (0,1167). Thus, suggests the parents present some degree of consanguinity close to the class of first cousins. In fact, we presumed that they could be cousins, considering also their ethnic origin, in which it is allowed marriages between relatives.

Run of homozygosity (ROH) test, aims to find the presence of long chromosomal regions presenting homozygosity in each family member. Consanguineous shared long ROH, while outbreed individuals should present short ROH regions. In the hypothesis of an autosomal recessive disease, considering that they are consanguineous, we identified in affected subjects (II-2, II-4) three ROH chromosomic regions: Chr7 position from 8 to 12 millions bp, Chr12 position from 12 to 27 millions bp and Chr22 position from 24 to 28 million bp. Therefore chr7, chr12 and chr22 presented a homozygosity region (in the affected and not in the unaffected individuals) whose length was 4 Mb, 15 Mb and 4 Mb respectively (see fig. 7).

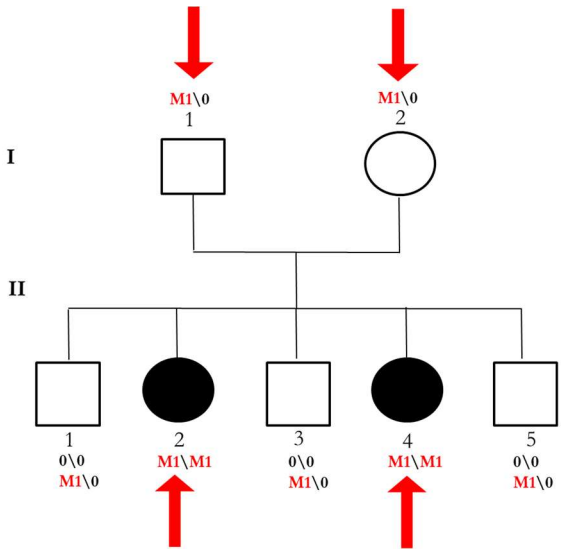

Figure 7: segregation of disease allele (single disease allele) under an AR and consanguinity model (ARC model). Figure depicts the family structure and the expected genotypes in the case of an AR disease. The parents are supposed to inherited disease allele from a common ancestor. Allele 0 refers to wild type allele, while M1 to mutated disease allele. The search for pathogenic variants were conducted by studying the segregation of disease allele in all the sibship, (full model) or only the affected children (Affected based analysis) (only I-1, I-2, II-2 and II-4) indicated by ⬇ red arrows.

As additional quality control check, we evaluated sex chromosomes to assess samples identity and avoid errors during library preparation. In table 5 are reported the number of loci on chrX in which each individual present homozygosity (i.e 0\0 or 1\1\) and heterozygosity. Males, due to their hemizygosity status (one

copy of chrX) are expected to show homozygosity to the most of loci (exception made for pseudo autosomal region).
Exome analysis called 22374 high- confidence individual sequence variants.
When searching for the best pathogenetic candidates we filtered the variants according to mapping variant quality (phred_scale > 200) and their depth of reading (200 independent reads for the six samples). Then, variants predicted to be synonymous or not having a location on either a coding exon, UTR, splice site junction or flanking intron and not showing a pattern of AR disease were filtered out. Thus, variants have been filtered according to the reported frequency in the gnomAD MAF (< 5%), to the prediction deleterious for SIFT and pathogenic Polyphen, according to their bioinformatically-estimated severity score of the CADD score (>0,85). The analysis did not identify any likely pathogenic variant. However, when the same model was tested ignoring the unaffected children (II-1 and II3) and relating the filtering parameters, 7 variants showed compatible segregation. In table 6a,b are reported the variant and their depth reading for each individual. It is considerably interesting that SERPINF1 is one of the seven genes that come out of the analyses. In the compound heterozygosis model (see fig.8) performed on a whole family, a final set of 163 (from 55 genes) candidate variants were selected. In thus search, candidates have been filtered in a first step only for MAF < 5%, in which we identified 87 variants from 42 genes (genes are listed in supplementary at point D). Looking at predictions (SIFT, Polyphen and Cadd scores), in addition to the variant frequency we identified 4 candidates genes: CERCAM, CHAT, SDHA and WHRN, in table 7a,b.

		DEPTH OF READINGS			
Sample_id	**Total bp**	**Mean**	**Median**	**%bases above 4**	**%bases above 10**
I-1	4116809001	66.11	41	90.1	82.3
I-2	5134113040	82.44	51	91.2	85.1
II-1	4494447654	72.17	43	90.2	82.6
II-2	4329328387	69.52	42	89.8	82.0
II-3	4842587640	77.76	49	91.6	85.5
II-4	5074565589	81.49	52	91.8	86.2

Table 4: depth of coverage per sample, based on intervals of the target region. In the last two columns depth in at least 4 and 10 calls for region. Mean= average reading depth, median= median reading depth, %bases above 4 and above 10= % of bases called more than 4 or 10 times.

Calls/Individual	#1	#2	#3	#4	#5	#6
Genotype 0\0	178	96	152	108	155	69
Genotype 0\1	51	213	52	199	61	217
Genotype 1\1	167	39	187	91	172	81
Sex	XY	XX	XY	XX	XY	XX

Table 5: comparison of phenotypic and genetic sex. Loci on chrX showing homozygosity and heterozygosity (number of loci presenting the genotype 0\1, over the total number of

loci). Heterozygosity was low in individuals I-1, II-1 and II-3, confirming their phenotype sex (males). Genotype 0\0: allele wild type, genotype 1\1: alternative allele.

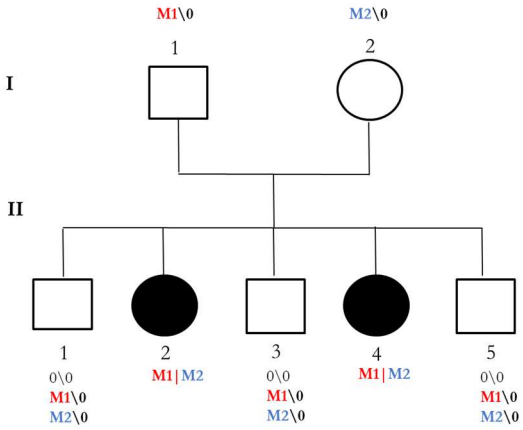

Figure 8: segregation of disease allele (two disease alleles M1 and M2) under an AR allele heterogeneity model (ARAH model). Figure represents the family structure in which 2 disease alleles segregates into affected children, one from father (M1) and one for mother (M2). Allele 0 indicate the wild type allele, M1|M2 indicate the presence of both mutation in the two chromosomes (one from I-1 and one from I-2). The three listed genotypes under individuals II-1, II-3 and II-5 represent all the possible expected genotypes.

Chr	Pos	Wt allele	Alt allele	Gene	Other info	RefSNP	Freq
2	21008652	G	A	APOB	exon26 c.C8216T p.P2739L	rs676210	0.292
6	656143	C	A	HUS1B	exon1 c.G802T p.D268Y	rs1211554	0.848
6	36479198	G	C	KCTD20	exon3 c.G77C p.S26T	rs2239808	0.237
10	95336798	C	T	SORBS1	exon28 c.G3362A p.R1121H	rs114088633	0.014
17	1776631	G	A	SERPINF1	exon7 c.G886A p.E296K	rs776811200	0.0002
17	11640293	A	G	DNAH9	exon10 c.A1810G p.M604V	rs61740059	0.0236
19	32881908	G	A	CEP89	exon18 c.C2071T p.R691W	rs34626245	0.095

Table 6a: variants detected under the AR model analysis, including affected children only. This investigation was conducted under the hypothesis that healthy subjects might have inherited the disease genotype, but they do not show the disease because a possible low penetrance or a very mild phenotype. APOB, HUSB1 and CEP89 showed a compatible segregation with AR model but they have been excluded because of the high frequency.

Gene	Gen I-1	Gen I-2	Gen II-1	Gen II-3	Gen II-2
APOB	0/1 30,14	0/1 22,37	0/0 34,0	0/1 27,28	1/1 0,39
HUS1B	0/1 27,15	0/1 25,29	0/1 16,25	0/1 17,37	1/1 0,32
KCTD20	0/1 12,19	0/1 23,11	1/1 0,23	0/1 43,17	1/1 0,36
SORBS1	0/1 22,14	0/1 28,13	1/1 0,32	0/1 18,16	1/1 0,31
SERPINF1	0/1 46,20	0/1 48,26	0/1 36,14	1/1 0,39	1/1 0,44
DNAH9	0/1 130,81	0/1 141,118	1/1 0,201	0/1 145,123	1/1 0,179
CEP89	0/1 172,120	0/1 168,147	0/1 206,11	0/1 154,131	1/1 0,264

Table 6b: genotype 0\0 allele wild type, genotype 1\1 alternative allele. Behind the genotype is reported depth of reading for each allele. Behind the genotype is reported depth of reading for each allele.

Chr	Position	Wt allele	Alt. allele	Gene	Other info	RefSNP	Frequency
chr5	236502	G	T	SDHA	exon9:c.G1191T:p.S397S	rs200223188	0.0001
chr5	236534	C	T	SDHA	exon9:c.C1223T:p.S408L	rs76896145	0.0207
chr5	236504	T	C	SDHA	exon9:c.T1193C:p.V398A	rs201741295	0.0009889
chr5	236534	C	T	SDHA	exon9:c.C1223T:p.S408L	rs76896145	0.0207
chr5	236513	C	T	SDHA	exon9:c.C1202T:p.A401V	rs201139275	0.000989
chr5	236534	C	T	SDHA	exon9:c.C1223T:p.S408L	rs76896145	0.0207
chr5	236534	C	T	SDHA	exon9:c.C1223T:p.S408L	rs76896145	0.0207
chr5	236538	C	A	SDHA	exon9:c.C1227A:p.L409L	rs75091805	0.0209
chr9	128423190	G	A	CERCAM	exon3:c.G119A:p.R40K	rs556586561	0.0003
chr9	128434147	G	A	CERCAM	exon10:c.G1015A:p.V339M	rs766133803	0.0004954
chr9	114403926	G	T	WHRN	exon6:c.C1335A:p.N445K	rs2274158	0.2476
chr9	114478722	C	T	WHRN	exon2:c.G668A:p.R223H	rs146273185	0.0067
chr9	114403966	A	G	WHRN	exon6:c.T1295C:p.V432A	rs2274159	0.04786
chr9	114478722	C	T	WHRN	exon2:c.G668A:p.R223H	rs146273185	0.0067
chr9	114424432	C	T	WHRN	exon2:c.G265A:p.A89T	rs4978584	0.2530
chr9	114478722	C	T	WHRN	exon2:c.G668A:p.R223H	rs146273185	0.0067
chr10	49616071	G	A	CHAT	exon2:c.G19A:p.D7N	rs1880676	0.2112
chr10	49622125	C	T	CHAT	exon5:c.C373T:p.L125F	rs8178990	0.0467
chr10	49616573	G	A	CHAT	exon2:c.G4A:p.A2T	rs3810950	0.2269
chr10	49622125	C	T	CHAT	exon5:c.C373T:p.L125F	rs8178990	0.0467

Table 7a: variant detected in compound heterozygosis model analysis.

Gene	Genotype I-1	Genotype I-2	Genotype II-1	Genotype II-3	GenotypeII-2	Genotype II-4
SDHA	0/0 34,0	0/0 33,0	0/0 35,0	0/0 43,0	0/1 90,16	0/1 119,18
SDHA	0/0 34,0	0/0 33,0	0/0 35,0	0/0 43,0	0/1 129,20	0/1 155,22
SDHA	0/0 34,0	0/0 33,0	0/0 35,0	0/0 43,0	0/1 99,16	0/1 127,18
SDHA	0/0 34,0	0/0 33,0	0/0 35,0	0/0 43,0	0/1 129,20	0/1 155,22
SDHA	0/0 34,0	0/0 33,0	0/0 35,0	0/0 43,0	0/1 122,20	0/1 147,20
SDHA	0/0 34,0	0/0 33,0	0/0 35,0	0/0 43,0	0/1 129,20	0/1 155,22
SDHA	0/0 34,0	0/0 33,0	0/0 35,0	0/0 43,0	0/1 129,20	0/1 155,22
SDHA	0/0 34,0	0/0 33,0	0/0 35,0	0/0 43,0	0/1 132,20	0/1 160,23
CERCAM	0/1 76,16	0/0 36,0	0/1 87,11	0/0 47,0	0/1 39,34	0/1 64,26
CERCAM	0/0 118,0	0/1 189,128	0/0 136,0	0/0 140,0	0/1 183,101	0/1 174,139
WHRN	0/1 72,45	0/0 36,0	0/1 75,63	0/0 36,0	0/1 69,50	0/1 60,72
WHRN	0/0 40,0	0/1 35,29	0/0 45,0	0/0 37,0	0/1 34,18	0/1 49,39
WHRN	0/1 144,86	0/0 36,0	0/1 123,111	0/0 36,0	0/1 132,101	0/1 106,138
WHRN	0/0 40,0	0/1 35,29	0/0 45,0	0/0 37,0	0/1 34,18	0/1 49,39
WHRN	0/1 134,136	0/0 46,0	0/1 158,139	0/0 40,0	0/1 156,135	0/1 190,138
WHRN	0/0 40,0	0/1 35,29	0/0 45,0	0/0 37,0	0/1 34,18	0/1 49,39
CHAT	0/1 24,13	0/0 51,0	0/0 41,0	0/0 35,0	0/1 26,11	0/1 22,11
CHAT	0/0 33,0	0/1 61,47	0/1 57,49	0/0 36,0	0/1 61,28	0/1 56,31
CHAT	0/1 40,34	0/0 36,0	0/0 44,0	0/0 34,0	0/1 33,23	0/1 67,35
CHAT	0/0 33,0	0/1 61,47	0/1 57,49	0/0 36,0	0/1 61,28	0/1 56,31

Table 7b: allele 0: allele wild type, allele 1: alternative allele. Behind the genotype is reported depth of reading for each allele.

4.2 Methylation sequencing

After alignments (Bismark aligner) we evaluated quality of sequencing and mapping efficiency by estimating several parameters as reported in table 8.

OI	Total Pairs (M)	Unique best hit (M)	Mapping efficiency %	Leftover (M)	Cov TR	Depth Cov
I-2	42.3	37.4	88.60	27.5	99.4 %	37.6
II-2	40.3	35.9	89.10	30.0	99.4 %	51.2
II-3	40.3	36.8	91.20	21.7	99 %	24.7
II-4	33.1	29.4	88.80	25.3	99.4 %	39.5
II-5	40.1	35.5	88.70	27.2	99.4 %	39.1

Table 8: mapping quality control (Bismark alignment). Total pairs = total number of PE fragments that have been sequenced; unique best hit = PE with the single match; mapping efficiency = number of PE sequenced mapping properly (i.e. both ends mapped in the same region and in the expected DNA strand; leftover = number of alignments remained after quality control filtering; Cov TR = number of PE mapping in the target region; Depth Cov = average of reading depth in the target region of leftover PE. There are the PE used for the following steps to estimate methylation status.

Subject	MetC (Million)	UnMetC (Million)	% MetC
I-2	37.3	195	16.10%
II-1	40.5	211	16.10%
II-2	34.1	216	13.60%
II-3	27.4	153	15.20%
II-4	30.3	202	13.10%

Table 9: CpG methylation status. It is reported the number of methylated and unmethylated Cs (MetC, UnmetC) and the overall methylation status (%MetC) of the genome of each individual.

CpG methylation in the genome often occurs in clusters and CpGs in the same region are correlated and perform a similar function. Identification of DRM was achieved using DMRfinder tool, that extracts methylation counts and performs a modified single-linkage clustering of methylation sites into genomic regions. It then compares methylation levels using beta-binomial hierarchical modelling and Wald tests. The program determines the valid CpG sites that meet minimum coverage criteria for one or more samples. Then it groups sites that are within a specified distance of each other into regions. This single-linkage clustering can lead to the chaining effect, in which distant CpG sites are placed into the same region due an extended series of intermediate sites. After clustering, each region is additionally required to meet a total methylation count minimum (we choose twenty as cut of) in each sample before being tested for differential methylation. In the final step, DMRfinder tool conducts pairwise tests of sample groups to identify genomic regions that are differentially methylated. The underlying statistics are based on Bayesian beta-binomial hierarchical modelling, which accounts for both the biological variation of replicates and the binomial nature of the methylation data. Regions that meet the minimum methylation difference (10%) and the maximum p- and q-values ($q < 0.05$) are reported as DMRs. The analysis highlights 124 differentially methylated region between affected children and the control group, composed by two healthy brothers and the mother. To better understand

the genome localization of DMRs, we annotated the regions and we saw that 55 of 121 were methylations that occurs into genes (gene names in supplements point C), 7 were recognized as promoter, 1 as Transcription factor binding site (TF binding site) and 1 as enhancer. We identified also 10 regions that mapped as promoter flanking region, 12 as CTCF binding site and 35 open chromatic regions. Transcriptional repressor CTCF also known as 11-zinc finger protein is a transcription factor encoded by the CTCF gene. CTCF is involved in many cellular processes, including transcriptional regulation, insulator activity and regulation of chromatin architecture. Methylation status of 124 DMR identified can be divided into 94 that showed a lower methylation in the affected group, while 30 exhibited a higher methylation. In table 10 the schematic summary of DMR types and status of methylation.

DMR type	# DMR	# of hypo methylated in Affected group	# of hyper methylated in Affected group
Genes	55	-	55
Open chromatic regions (OCR)	35	28	7
CTCF binding site	12	8	4
Promoter flanking region	10	2	8
Promoter			
TF binding site	7	3	4
Enhancer	1	1	-
Promoter flanking region	1	1	-
Total	121	43	78

Table 10: annotation of 121 Differentially Methylated Region (q< 0,05) identified. Genes show a high methylation, despite OCR that present a lower methylation in the disease group.

Overall, affected group showed a higher methylation profile in comparison with healthy subjects. The majority of DMR regions are represented by genes and all of the seem to be more methylated in the affected group. In order to discover which genes could play a role in the ethiopathogenesis of this strange disease, we annotate our sequences using Annovar tool. We perform a String analysis to asses if there was any possible correlation between identified genes based on type of interaction, confidence and molecular action. We exclude 15 entries because they were identified as RNA genes. In figure 9a it is showed PPI (protein-protein interaction) based on confidence, line thickness indicates the strength of data support, colours type of interaction. Exploring the molecular action in the interaction showed by STRING, we see the four molecular action clusters (showed in figure 9b), that see as protagonist 1) RXRA, RAC3 and PSD3; 2) MUC4 and GALTN8; 3) SHANK2, BRD1, SETD1B, DOT1L and IGHMB2; 4) SNTG2 and SNTA1. Details and complete gene names are reported in supplementary materials, point D.

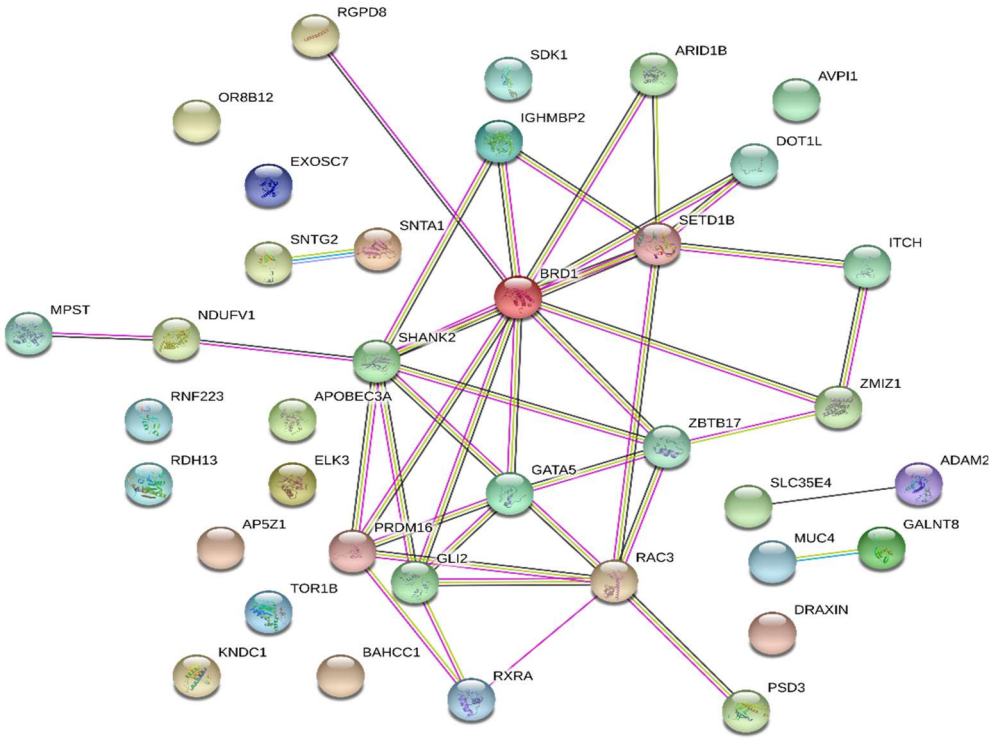

Figure 9a: STRING network of physical and functional interactions for DMR genes. analysis based on confidence (0.150). Known interaction is showed by cerulean and fuchsia lines (from curated database and experimentally determinated); red, green and blue lines show the predicted interaction (gene neighbour, gene fusion, gene co-occurrence); light green, black and light blue display other info as textmining, co-expression and protein homology.

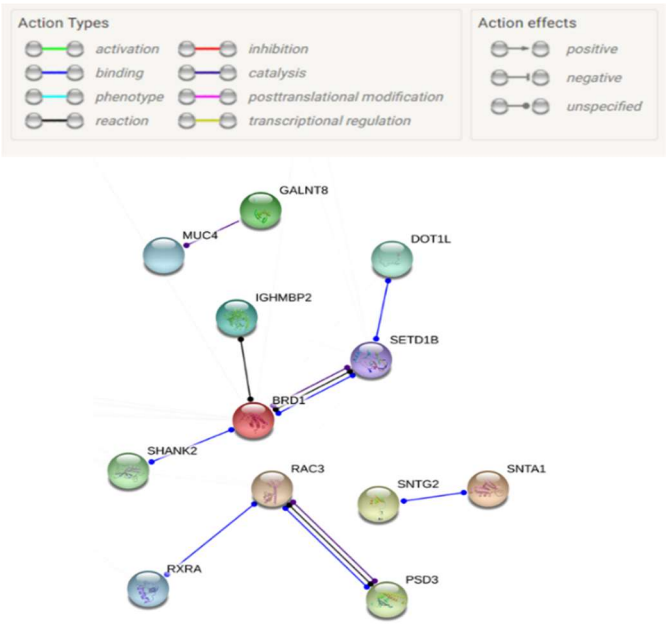

Figure 9b: Network of interaction focusing on PPI molecular action. Action type are explained by different colour lines (as reported in the legend), while the effect types are represented by arrow, circles or straight line. Only binding and catalysis action types appears in the network.

4.3 Metabolomic profile

To investigate the modulation of metabolomic profile we prefer to analyse only the children, excluding the parents, to avoid false positive modulation due to the age. Modulated and statistical significative features are plotted in the mirror plot (or cloud plot) in fig.10. Green circles represent the upregulated features in affected children, otherwise red circles represent the downregulates ones. The size of each circle corresponds to the (log) fold change of the feature (the average difference in relative intensity of the peak be-tween sample groups). Larger circles correspond to peaks with greater fold changes. The shade of colour is used to represent p-value, with brighter circles having lower p-values. The retention time corrected (TIC) total ion chromatograms are also overlaid in gray in the background of the figure. Metabolomic profile show a general down regulation of some metabolites in comparison with the healthy group.

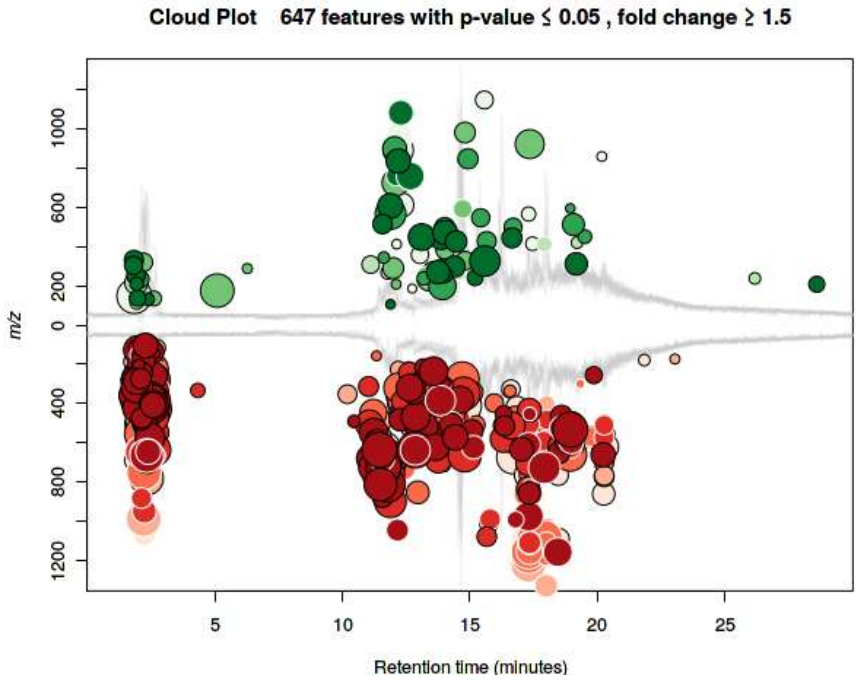

Figure 10: cloud plot of the modulated features (fold change ≥1.5, p-value≤ 0.05). Feature signals belonging to molecules defined by m/z and retention time. Differential analysis is based on the intensity of the signals in affected versus unaffected individuals. Green up modulation, red down modulation) between affected and healthy group.

Cloud plot recognized 647 modulated features, the known metabolites (in METLIN database) are represented by the circles whit black borders, in total 445.

PCA is an unsupervised method aiming to best explain the variance in a data without referring to class labels. The data are summarized into much fewer variables called scores which are weighted average of the original variables. The weighting profiles are called loadings. Fig. 11 shows the 3-D scores plot between selected PCs.

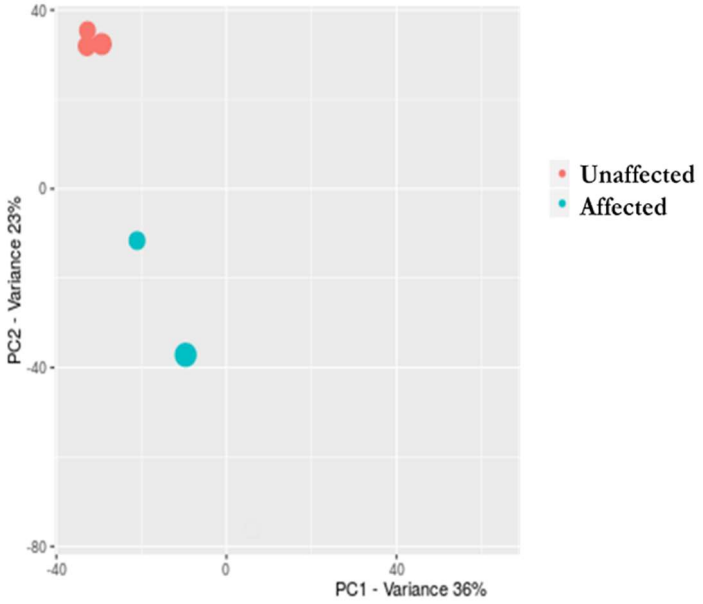

Figure 11: PCA plot. Individuals are plotted according to their values of PC1 and PC2 from PCA that included all the detected features. PC1 and PC2 count for the 36 and 23% of the total variance, respectively.

Information (m/z, rt and intensity) of all identified metabolites are processed by MetaboAnalyst tool (v.3.0) to perform statistical analyses. Data were processed using a univariate analyses method such as Fold-change (FC), T-test and Volcano plot analysis. The purpose of fold change is to compare absolute value changes between two group means; important features selected by fold-change analysis with threshold 2. A T-test is a type of statistical test that is used to compare the means of two groups. has been considered as significative only features with a p-value adjusted ≤ 0.05. Data are reported in table 11.

m/z	Matabolites Name	Adjusted p-value	FDR
266.9896	dopamine 3-sulfate	1.1074e-12	5.4374e-10
146.9842	2,3-diphospho-D-glycerate	2.7473e-07	6.7447e-05
129.0204	1,5-anhydro-D-glucitol	1.1254e-05	0.0011052
179.0575	dopaquinone	1.9603e-05	0.0016042
225.0632	2-keto-6-aminocaproate	3.1295e-05	0.0021951
145.0881	N-acetyl-serotonin glucuronide	4.9289e-05	0.0030251
425.0129	3-hydroxy-L-kynurenine	0.00019605	0.010695
240.0974	2-imino-3-(indol-3-yl)propanoate	0.00025665	0.011677
198.0762	N-acetyl-&beta-D-glucosamine(anhydrous)-N-acetylmuramate	0.0002616	0.011677
269.1412	&gamma-L-glutamyl 5-phosphate	0.00029578	0.012103
272.0751	S-acetyldihydrolipoamide	0.00081008	0.027177
185.0445	salicyl-6-hydroxy-2-cyclohexene-on-oyl	0.00087264	0.027177
209.0682	3-hydroxyquinine	0.0008856	0.027177
325.1676	11-deoxycorticosterone	0.0011099	0.032057
353.1771	3-&beta-hydroxyandrost-5-en-17-one 3-sulfate	0.0016607	0.045301
165.0763	dTDP	0.001755	0.045353
280.1175	8-oxo-GTP	0.0020966	0.049958
522.9683	triiodothyronine sulfate	0.0021367	0.049958

Table 11: modulated metabolites between affected and unaffected individuals (t-test < 0.05). Metabolites have been named according to their known m/z values.

Volcano plot (fig. 12) is a type of scatter-plot that is used to quickly identify changes in a data set, it plots significance versus fold-change on the y and x axes, respectively. Were considered as important (red circles) only metabolites above the thresholds (-log10(p-val) 0,05 and FC ≥ 2).

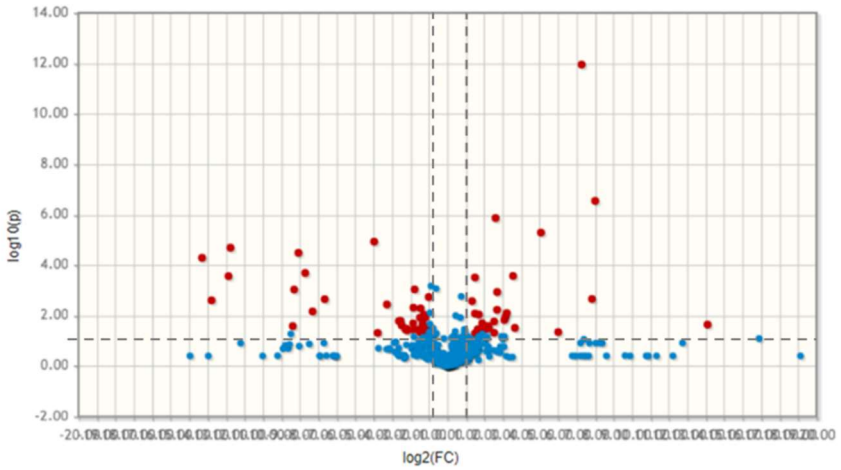

Figure 12: volcano plot of modulated metabolites between affected and unaffected individuals. Pink circles represent features above the threshold. Both fold changes and p-values are log transformed. The further position from the 0 of y axis the more significant association.

In hierarchical cluster analysis, each sample begins as a separate cluster and the algorithm proceeds to combine them until all samples belong to one cluster. Differences in the metabolomic profile can of each cluster can be visualized though a plot called heatmap (fig. 13), that summarized the high and low expression of a metabolite by colours.

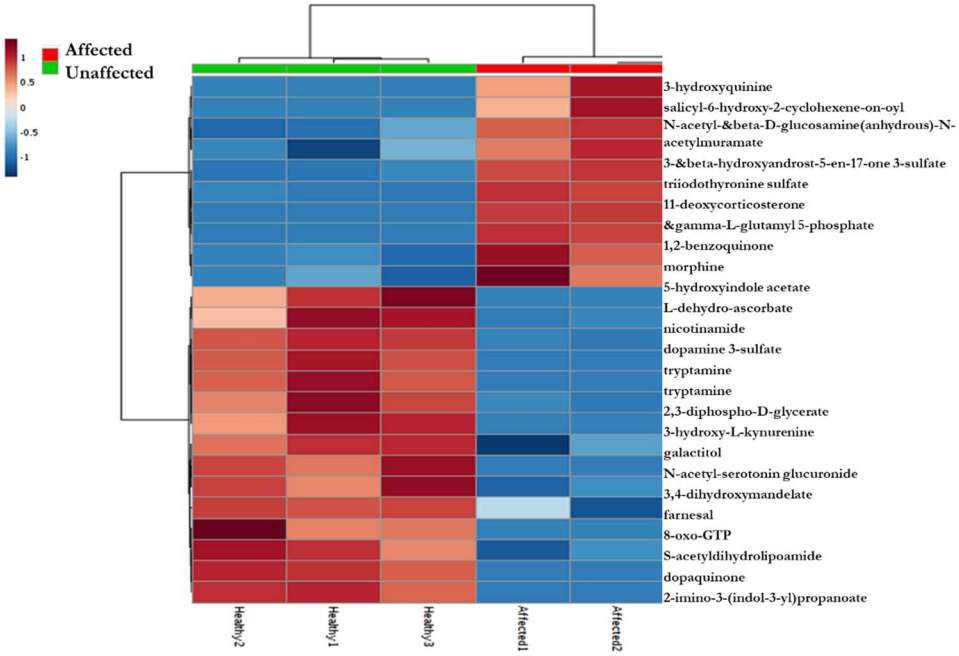

Figure 13: heatmap of the 25 most significant modulated metabolites between affected and unaffected individuals. Metabolites levels define profiles able to group affected and

unaffected separately. Blue represent lower abundance respect the average abundance, as red present a higher abundance (distance measure using Euclidean distance, and clustering algorithm using ward.D).

Significative metabolites identified by univariate statistical analysis were used to explore which pathway were involved. This analysis refers to KEGG database to build the network between the query metabolites. The significative metabolites (query) are represented in red, the neighbours coloured in blue represent the metabolites directly connected with the queries. Statistical analyses of identified pathway have been reported in table 12.

Pathway	Pathway size	Overlap size	p-value
Phenylalanine, tyrosine and tryptophan biosynthesis	27	4	0.001
Tryptophan metabolism	79	6	0.02
Pyrimidine metabolism	60	6	0.02
Vitamin B6 metabolism	32	4	0.03

Table 12: most significative modulated pathways suggested by metabolomic analysis between the two group.

5. DISCUSSION

Whole Exome sequencing

Osteogenesis imperfecta is a rare genetic heterogenous and pathology that can generate a wide spectrum of different and peculiar phenotypes. Since 2006, NGS identified novel disease associated loci. In the last twelve years, fifteen OI-related genes have been discovered, clarifying the extremely wide genetic heterogeneity of OI. Despite the great achievements during this years, there are OI types that still lack of a genetic explanation.

Here we described a study performed at multi-omics levels to elucidate potential changes that lead or are associated to an unexplained OI type segregating in a nuclear Pakistan family. The study has been undertaken because we estimated that the exome analysis was powered enough (power > 80%) to identify the locus (i.e. variants or restricted chromosomes regions) linked to the disease, under the hypothesis of an Autosomal Recessive disease and Consanguinity (ACR) of the parents. In this scenario we firstly searched for chromosomal regions presenting homozygosity among the affected individuals (single disease allele in double dosage in each affected individual is expected) and not in the unaffected children.

Exome sequencing, performed in the six family subjects (median depth of reading was 40x for each sample), confirmed the consanguinity between the parents (probably first cousins). Then, we searched for shared regions of homozygosity between affected, since they represent the best candidate regions to contain a disease variant. Hence, we found 3 long chromosomal regions presenting homozygosity (homozygosity mapping) on chr7, chr12 and chr22. However, we cannot exclude that other smaller shared chromosomes regions are present.

The searches for the pathogenic variant under the ARC hypothesis in the 3 chromosomal regions did not revealed any strong candidate. These results can have different explanations 1) the pathogenic variant maps on a detected homozygosity region but is not located inside the protein coding region (thus not sequenced in the present study and therefore not detected); 2) the pathogenic variant is in a protein coding region that has been sequenced at a very low (or null) depth (we estimated that the 18% of the target region presented a depth of reading low than 10 calls per base); 3) the disease genotype present an incomplete penetrance (there might be healthy subject harbouring the disease genotype, see fig,7) 4) parents harbour two different pathogenic variants that can be detected searching for disease allele gene in the hypothesis that affected individuals carry two different mutations (one per chromosome, see fig. 8).

Although it was hard impossible for the present study to explore point 1 (it would require a genome sequencing) and point 2 (it would require additional sequencing and the usage of a new system to capture exons), but it was possible to investigate point 3 and 4.

When excluding the two unaffected children (affected based analysis, see fig 7, red arrows) from the analysis we found some intriguing results. The analysis showed 7 candidates (reported in table 6), one of them was SERPINF1 gene, the gene of Osteogenesis imperfecta type VI. The identified missense variant, E296K, it is quite rare (0,024%) and leads an amino acid change in position 296, a glutamic acid became a lysine. Prediction analysis suggest that E296K is a high pathogenic and damaging mutation.

To our knowledge, this rare variant has not been reported as disease allele for any bone disorder or others. According to its rare frequency in the general population

and the possible pathogenetic impact, suggested by bioinformatic predictors, we could speculate that this variant might manifest a relative severe effect when in combination with other possible and still unknown genetic factor (i.e. other relative severe mutation in other OI-related genes risk of disease). The disease will appear in subjects presenting both variant (in different genes) that are not associated to the disease when present alone. This highly speculative hypothesis would imply that the second gene, other than the strong candidate SERPINF1 map in the region oh homozygosity (chr7, chr12, chr22) in according with the high likely ARC model. This will be investigated in the nearest future.
Compound heterozygosity model (see fig. 8), identified 163 variants from 55 genes. When applying the filter described in results, point 4.1, 7 genes remained (see table 7). CERCAM gene captured our attention. The analysis showed that the parents were carrier for two different mutation (G119A in the mother, G1015A in the father).
It is known that collagen is modified by hydroxylation and glycosylation of hydroxylysine residues. This glycosylation is initiated by the β1, O galactosyltransferases COLGALT1 and COLGALT2.[65] CERCAM protein has been described as structurally similar to his paralogs COLGALT1 and COLGALT2.[65] However, it did not show any galactosyltransferase activity. Nevertheless, it is intriguing that CERCAM has been associated to idiopathic scoliosis[66], suggesting that it may play a role on bone homeostasis.
Currently we are not able to determine if one of these mutations could be associated to the OI-like phenotype, and the results suggest a widely complicated scenario, in which were involved two different genes (SERPINF1 and CERCAM). Certainly we will perform further investigation to evaluate the SERPINF1 mutation, evaluating also the penetrance. It is possible that SERPINF1 mutation may produce a mild phenotype, explaining that healthy state of individual II-3 whom presented E296K variant, while the co presence of other variants (such as CERCAM), also in different genes may lead to this complex phenotype.

Methylation sequencing

Epigenetic modulation has been reported to play a crucial role in several disease conditions.[67] In 2014 it has been described a type V OI associated to methylation, in which a de novo C>T transition in the 5' UTR of the IFITM5 gene was reported. This dinucleotide shown to be highly methylated in several tissues and particularly in the sperm DNA.[68]
Comprehensive analysis of human DNA methylomes revealed distinct DNA methylation patterns among different cell types, tissue types and individuals, potentially underscoring divergent epigenetic regulation at different scales of phenotypic diversity.[89]
In our study, methylation profiling (in individuals I-2, II-2, II-3, II-4 and II-5) was carried out in MSCs. We choose to study MSCs because are known to be multipotent cells and can differentiate into different cells types, including osteoblast and chondrocytes, who play a crucial role in Osteogenesis Imperfecta. Met-Seq, is a new and challenging analysis, able to evaluate the methylation status at single base resolution. Over the entire genome we observed an overall lower cytosines methylation in the affected sisters (M-value ~13%), in comparison with the unaffected individuals (~16%), as reported in table 8.

When studying regions with multiple methylation sites it could be more fruitful to investigate the overall methylation at the region level rather than to single sites independently. This may be more biologically plausible and would reduce the number of statistical tests when searching for differentially methylated regions. Therefore, we studied regions of CpG sites rather than the single CpG.
Differential analysis, identified 121 differential methylated regions (DMRs) in the affected individuals, most of them were annotated to map into genes and open chromatin regions (OCR), while 31 were mapped as genome regulatory features. We saw a higher methylation in gene regions rather than OCR. According with the theories which assumed that the transcriptional activity in gene and OCR is under DNA methylation control, these foundings suggest that the activity of DMRs were different in affected sisters than the unaffected family members.
It is known that DNA methylation has different functions in different genomic contexts. DNA methylation in the protein coding sequence seems to be involved in transcription elongation and alternative splicing.[69] While, there is evidence that methylated CpG islands at transcription start sites (TSS) controls genes transcription.[70]
In this study we looked at the methylation status as possible marker to distinguish affected and unaffected individuals. To explore how methylation affects biological process we would need to perform additional studies that focus on gene expression and conformational status of chromatin. This would give a broad picture on several steps that play a role in the regulation of genes whose products might impact on bone homeostasis. The study of methylation performed at high resolution (single base) permits to study methylation under many different hypotheses. For instance, methylation could be studied at single base level, or as average level of target regions (i.e. genes, promoters, transcription factor binding sites, enhancers, chromatin regions etc). or investigate if there are nucleotide variants that inserts or remove potential methylation target sites. This is still a challenging open issue (many statistics tests when methylation is studied at single base level) and the still present lack of knowledge on the impact of methylation on the different regions of the genome.
Among the 55 annotated genes, we identified some interesting genes, that showed molecular interaction during STRING analysis. Moreover, they are known to be involved in bone metabolism, such as RXRA (Retinoid X Receptor Alpha), ELK3 (ETS Transcription Factor) and GLI2 (GLI Family Zinc Finger 2).
It is well known that terminal differentiation of multipotent stem cells is achieved through a coordinated cascade of activated transcription factors and epigenetic modifications that drive gene transcription responsible for unique cell fate.[74] In a recent study, has been assessed that RUNX2, C/EBP, retinoid X receptor (RXRA), and vitamin D receptor (VDR) binding sites, are genes vital for osteogenic identity.[74] RXRs is known to be partners for a huge number of other nuclear receptors, playing a coordinating role in several signalling pathways, but also RXRs homodimers has been described. Recently, it has been showed that RXRs is involved in osteoclastogenesis and bone remodelling[75]. It is able to control transcription of bone-forming osteoblasts and bone-resorbing osteoclasts, both through heterodimerization with other nuclear receptors and through RXR homodimerization.[75]
ETS Transcription Factor, or ELK3, it is known to play a crucial role as negative regulator of transcription, but it is able to activate transcription when it is co-

expressed with RAS, SRC or MOS. Several studies described ELK3, SOX10 and SOX18 as critical for neural crest development,[76] but, surprisingly, they were described also as up-regulated gene in distraction osteogenesis, also if they are not involved in postnatal bone repair.[77]
In 2010, Kesper et al, reported that GLI2 activity is capable to induce osteoblast differentiation, stimulating the expression of Ihh target genes in skeletal tissues.[78] In fact, it is showed that a repressor function of Gli2 is likely dispensable for endochondral ossification.
Considering this foundings, the crucial role of retinoid X receptor in bone homeostasis, the activity of GLI2 during ossification and the up regulation of ELK3 in distraction osteogenesis, we want to deepen if the methylation status of this genes reflects the expression level, (trough transcriptomics data analysis). The comparison of methylation and expression data may allow to evaluate the role of these genes in the pathology. For these reasons we are performing transcriptomic investigation, aimed to compare expression levels to epigenetic data, to truly inhabits methylation signals.

Metabolomic analysis

Metabolomics represent the unique biochemical fingerprint of all cellular processes, it allows to deeper into the actual phenotype of any biological system.[79] Here we describe the results obtained by an untarget analysis performed on the serum of the family members. Most of the "metabolites signals" that have been detected cannot be easily recognize in a known metabolites because the database that associates a metabolite to a given signal is still incomplete.
The metabolomic analysis on children, (in which we investigated only II-1, II-2, II-3, II4 and II-5 individuals), identified 5340 signals, 647 of them resulted identified as significative modulated ($p \leq 0,05$) metabolites between affected and unaffected individuals. It is intriguing that the principal component analysis (PCA) that included all the detected signals showed affected and unaffected individuals in distinct regions of the PC1 and PC2 plot (fig. 11). This analysis presents a strong bias due to the low sample size and to the fact that the affected individuals are female and unaffected individuals are males and therefore we cannot exclude that PCA is distinguishing gender effects from disease effects. In general, we observed a lower expression of most of identified metabolites in affected than in unaffected individuals.
Enrichment analysis revealed the following 4 significant "phenylalanine, tyrosine and tryptophan biosynthesis", "tryptophan metabolism, pyrimidine metabolism", and "Vitamin B6 metabolism".
Tryptophan metabolism (see table 12) and serotonin N-acetyl glucuronide (in tab. 11) could be associated with some signalling molecules in the nervous system, including neuropeptides and neurotransmitters that have been identified in bone.[80,81] Recently, it has been suggested that peripheral serotonin produced in the gut was a major negative regulator of osteoblast proliferation.[82] In fact, osteoclasts are able to synthetize serotonin that acts locally to induce osteoclast precursors differentiation.[83] The current hypothesis about serotonin role in bone asserts that a low amount of serotonin in bone microenvironment would trigger the proliferation of both osteoclast and osteoblast precursors.[83] It is noteworthy that serotonin biosynthesis starts from tryptophan and 5-HTP (5-hydroxy-tryptophan), and it needs Vitamin B6.[84]

Another modulated pathway, found in our analysis was the phenylalanine, tyrosine and tryptophan biosynthesis. It has been demonstrated that 3,4-dihydroxy-L-phenylalanine, also known as DOPA, enhances the osteogenic differentiation in vitro of human bone marrow-derived mesenchymal stem cells (h BM-MSCs).[85] Moreover many studies showed that damaged or missing sympathetic nerves leads to the abnormal bone formation and bone mass[86] suggesting that osteoblasts may respond to neurotransmitters such as dopamine.[83] Dopamine biosynthesis starts from tyrosine amino acid, that in turn derived from the conversion of phenylalanine to tyrosine (table 12). Some studies reported that dopamine receptors are expressed in osteoblast and were functional to promote bone mineralization.

All these reports together with our results would suggest that tryptophan, phenylalanine, tyrosine and L-dopa alteration may reflect an impaired proliferation and mineralization of bone cells, affecting and being markers of bone remodelling process. Moreover, it should be noted that Vitamin B6 it always necessary to catalyse the synthesis of serotonin and L-DOPA metabolites.

6. CONCLUSIONS

The here reported study describes the results from exome, methylome and metabolome studies performed on a family in which is likely to segregate a form of type VI OI. The exome analysis was not able to identify a clear pathogenic variant. However, several chromosomes regions and possible candidate genes have been suggested when testing different segregation models. The intriguing results is that among the associated variants, many of them belong to gene associated to bone homeostasis raising the still highly speculative hypothesis that more than one gene might be concurrently associated to the affected status within the family. Methylation analysis showed an overall different methylation level between affected and unaffected as well as several chromosomal regions linked to RXRA, ELK3 and GLI2 bone homeostasis genes. These regions may serve as markers for the disease status and might help in shedding light on the possible pathogenic steps associated to the disease. Metabolome studies offered similar results, adding further information of the defective mechanism involved in the disease.

The future investigation will try to enhance and integrate the results from the present omics (transcriptomic analysis is ongoing) into a context of system biology aimed to depict and clarify the defects and biological processes associated to the disease.

7. SUPPLEMENTARY MATERIALS
A. DNA EXTRACTION PROTOCOL

QIAamp® Blood Mini Kit
The QIAamp Blood Mini Kit (cat. nos. 51104 and 51106) can be stored at room temperature (15–25°C) for up to 12 months. Reconstituted QIAGEN Protease is stable for 2 months when stored at 2–8°C.

- Further information
- *QIAamp DNA Mini and Blood Mini Handbook* **www.qiagen.com/HB-0329**
- Safety Data Sheets **www.qiagen.com/safety**
- Technical assistance **support.qiagen.com**

Notes before starting

- Perform all centrifugation steps at room temperature (15–25°C).
- Dissolve any precipitates in Buffer AL by warming at 56°C until the precipitate has dissolved.
- Add ethanol to Buffer AW1 and Buffer AW2 concentrates, as indicated on the bottle.
- Add Protease Solvent to lyophilized QIAGEN Protease, as indicated on the label.
- Equilibrate samples to room temperature (15–25°C).
- Preheat a water bath or heating block to 56°C.

1. Pipet 20 µl QIAGEN Protease into a 1.5 ml microcentrifuge tube. Add 200 µl sample.

If the sample volume is less than 200 µl, add the appropriate volume of PBS.
2. Add 200 µl Buffer AL. Mix thoroughly by vortexing.

3. Incubate at 56°C for 10 min. Briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from the lid.
4. Add 200 µl ethanol (96–100%). Mix thoroughly by vortexing. Briefly centrifuge the tube to remove drops from the lid.

5. Pipet the mixture onto the QIAamp Mini spin column (in a 2 ml collection tube) and centrifuge at 6000 x *g* (8000 rpm) for 1 min. Discard the flow-through and collection tube.

Note When preparing DNA from buffy coat or lymphocytes, centrifugation at full speed is recommended to avoid clogging.
6. Place the QIAamp Mini spin column in a new 2 ml collection tube and add 500 µl Buffer AW1. Centrifuge at 6000 x *g* (8000 rpm) for 1 min. Discard the flow-through and collection tube.

7. Place the QIAamp Mini spin column in a new 2 ml collection tube and add 500 µl Buffer AW2. Centrifuge at full speed (20,000 x *g*; 14,000 rpm) for 3 min. Discard the flow-through and collection tube.

8. **Recommended** Place the QIAamp Mini spin column in a new 2 ml collection tube (not provided) and centrifuge at full speed for 1 min. This eliminates the chance of possible Buffer AW2 carryover.

9. Place the QIAamp Mini spin column in a new 1.5 ml microcentrifuge tube (not provided), add 200 µl Buffer AE or distilled water and incubate at room temperature (15–25°C) for 1 min. Centrifuge at 6000 x *g* (8000 rpm) for 1 min to elute the DNA.

B. RNA EXTRACTION

RNeasy® Mini Kit,
Notes before starting
If purifying RNA from cell lines rich in RNases, or tissue, add either 10 µl β-mercaptoethanol (β-ME), or 20 µl 2 M dithiothreitol (DTT),* to 1 ml Buffer RLT. Buffer RLT with β-ME or DTT can be stored at room temperature for up to 1 month.

- Add 4 volumes of ethanol (96–100%) to Buffer RPE for a working solution.
- Remove RNA*later*® stabilized tissue from the reagent using forceps.
- For RNeasy Protect Mini Kit (cat. nos. 74124 and 74126), please start with the *Quick-Start Protocol RNA*later *RNA Stabilization Reagent, RNAlater TissueProtect Tubes, and RNeasy Protect Kits*.

* This option not included for cells in handbook; handbook to be updated.
1. **Cells** Harvest a maximum of 1 x 107 cells, as a cell pellet or by direct lysis in the vessel. Add the appropriate volume of Buffer RLT (see Table 1).
Tissues Do not use more than 30 mg tissue. Disrupt the tissue and homogenize the lysate in the appropriate volume of Buffer RLT (see Table 1). Centrifuge the lysate for 3 min at maximum speed. Carefully remove the supernatant by pipetting, and use it in step 2.
2. Add 1 volume of 70% ethanol to the lysate, and mix well by pipetting. Do not centrifuge. Proceed immediately to step 3.
Transfer up to 700 µl of the sample, including any precipitate, to an RNeasy Mini spin column placed in a 2 ml collection tube (supplied). Close the lid, and centrifuge for 15 s at ≥8000 x *g*. Discard the flow-through.

Optional For DNase digestion, follow steps 1–4 of "On column DNase digestion" in *Quick-Start Protocol RNeasy Mini Kit, Part 2*.
4. Add 700 µl Buffer RW1 to the RNeasy spin column. Close the lid, and centrifuge for 15 s at ≥8000 x *g*. Discard the flow-through.
5. Add 500 µl Buffer RPE to the RNeasy spin column. Close the lid, and centrifuge for 15 s at ≥8000 x *g*. Discard the flow-through.
6. Add 500 µl Buffer RPE to the RNeasy spin column. Close the lid, and centrifuge for 2 min at ≥8000 x *g*.

Optional Place the RNeasy spin column in a new 2 ml collection tube (supplied). Centrifuge at full speed for 1 min to dry the membrane.
7. Place the RNeasy spin column in a new 1.5 ml collection tube (supplied). Add 30–50 µl RNase-free water directly to the spin column membrane. Close the lid, and centrifuge for 1 min at ≥8000 x *g* to elute the RNA.
8. If the expected RNA yield is >30 µg, repeat step 7 using another 30–50 µl of RNase-free water, or using the eluate from step 7 (if high RNA concentration is required). Reuse the collection tube from step 7.

C. GENE NAMES OF SIGNIFICANT DMR

AC007881.2	AC073869.4	ADAM2
AC009093.10	AC087855.1	AL139317.1
AC009093.2	AC090993.1	AP5Z1
AC009093.8	AC127526.3	APOBEC3A
AC011476.2	AC127526.4	ARID1B
AC026412.1	AC243964.2	AVPI1
AC026412.3	AL008719.1	BAHCC1

BAK1P1	KNDC1	RXRA
BRD1	LINC00887	SDK1
DOT1L	MPST	SETD1B
DRAXIN	MUC4	SHANK2
ELK3	NDUFV1	SLC35E4
EXOSC7	OR8B12	SNTA1
GALNT8	PRDM16	SNTG2
GATA5	PSD3	TOR1B
GLI2	RAC3	ZBTB17
IGHE	RDH13	ZMIZ1
IGHMBP2	RGPD8	
ITCH	RNF223	

D. CANDIDATES GENE INDENTIFIED IN THE COMPOUND HETEROZYGOSIS MODEL

ADPRHL1	HRNR	SLC35E2
ARHGEF7	HSPG2	TNFRSF25
ASMTL	JAG1	PLEKHG5
CERCAM	KAZN	VCX3A
CHAT	KLHDC7A	WASHC2A
COL2A1	LAMC3	WDFY4
COLEC11	MAN1B1	WHRN
DDX11	MYO16	
DNM1	NELFB	
ENTPD2	NQO2	
FAAP20	PEPD	
FANCA	PLEKHG5	
FBN2	PLOD1	
FBXL21	PNPLA7	
FBXO18	PRAMEF1	
FCGBP	PRDM10	
FPGS	RAP1GAP	
HIP1R	SDHA	

Chr	Position	Wt allele	Alt. allele	Gene	Other info	RefSNP	Frequency	Genotype I-1	Genotype I-2	Genotype II-1	Genotype II-3	GenotypeII-2	Genotype II-4
chr1	26190776	A	G	CATSPER4	CATSPER4:NM_198137:exon1:c.A149G:p.H50R	rs41284333	0.1620	0/0:37,0	0/1:52,43	0/0:34,0	0/1:52,45	0/1:37,34	0/1:46,42
chr1	26201460	G	A	CATSPER4	CATSPER4:NM_198137:exon9:c.G1306A:p.D436	rs6657616	0.2235	0/1:65,59	0/0:38,0	0/1:81,56	0/0:37,0	0/1:93,57	0/1:70,56
chr1	26191282	T	C	CATSPER4	.	rs11247865	0.1596	0/0:31,0	0/1:21,34	0/0:35,0	0/1:25,19	0/1:21,15	0/1:18,29
chr1	26201460	G	A	CATSPER4	CATSPER4:NM_198137:exon9:c.G1306A:p.D436	rs6657616	0.2235	0/1:65,59	0/0:38,0	0/1:81,56	0/0:37,0	0/1:93,57	0/1:70,56
chr1	26191303	A	G	CATSPER4	CATSPER4:NM_198137:exon2:c.A230G:p.Q77R	rs11247866	0.1547	0/0:38,0	0/1:42,49	0/0:35,0	0/1:36,36	0/1:31,28	0/1:36,52
chr1	26201460	G	A	CATSPER4	CATSPER4:NM_198137:exon9:c.G1306A:p.D436	rs6657616	0.2235	0/1:65,59	0/0:38,0	0/1:81,56	0/0:37,0	0/1:93,57	0/1:70,56
chr1	26191400	C	T	CATSPER4	CATSPER4:NM_198137:exon2:c.C327T:p.I109I	rs11247867	0.1534	0/0:38,0	0/1:106,9	0/0:35,0	0/1:75,74	0/1:71,60	0/1:87,78
chr1	26201460	G	A	CATSPER4	CATSPER4:NM_198137:exon9:c.G1306A:p.D436	rs6657616	0.2235	0/1:65,59	0/0:38,0	0/1:81,56	0/0:37,0	0/1:93,57	0/1:70,56
chr1	26199948	A	G	CATSPER4	CATSPER4:NM_198137:exon7:c.A877G:p.I293V	rs17257155	0.1621	0/0:33,0	0/1:55,60	0/0:35,0	0/1:53,49	0/1:38,44	0/1:60,56
chr1	26201460	G	A	CATSPER4	CATSPER4:NM_198137:exon9:c.G1306A:p.D436	rs6657616	0.2235	0/1:65,59	0/0:38,0	0/1:81,56	0/0:37,0	0/1:93,57	0/1:70,56
chr1	2193733	G	A	FAAP20	FAAP20:NM_001256946:exon3:c.C376T:p.P126S,I	rs1058411	0.1225	0/1:35,34	0/0:36,0	0/1:37,24	0/0:34,0	0/1:32,20	0/1:42,25
chr1	2193733	G	A	FAAP20	FAAP20:NM_001256946:exon3:c.C376T:p.P126S,I	rs1058411	0.1225	0/1:35,34	0/0:36,0	0/1:37,24	0/0:34,0	0/1:32,20	0/1:42,25
chr1	152214138	C	G	HRNR	HRNR:NM_001009931:exon3:c.G7491C:p.S2497S	rs879592321	.	0/1:48,21	0/0:40,0	0/1:51,10	0/1:67,9	0/1:29,12	0/1:59,22
chr1	152215152	A	G	HRNR	HRNR:NM_001009931:exon3:c.T6477C:p.Y2159Y	rs75639578	0.1729	0/0:540,0	0/1:424,8	0/0:55,0	0/0:517,64	0/1:491,63	0/1:557,80
chr1	152214138	C	G	HRNR	HRNR:NM_001009931:exon3:c.G7491C:p.S2497S	rs879592321	.	0/1:48,21	0/0:40,0	0/1:51,10	0/1:67,9	0/1:29,12	0/1:59,22
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152214290	G	A	HRNR	HRNR:NM_001009931:exon3:c.C7339T:p.R2447C	rs4845747	0.4386	0/1:241,19	0/1:282,5	0/1:243,1	0/1:292,14	0/1:260,12	0/1:369,137
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152214707	C	A	HRNR	HRNR:NM_001009931:exon3:c.G6922T:p.G2308(	rs12126691	0.4739	0/1:222,83	0/1:204,28	0/1:228,6	0/1:216,49	0/1:165,45	0/1:217,49
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152215078	C	T	HRNR	HRNR:NM_001009931:exon3:c.G6551A:p.S2184N	rs61814935	0.4830	0/1:173,12	0/1:177,72	0/1:117,1	0/1:200,76	0/1:186,46	0/1:253,78
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152215152	A	G	HRNR	HRNR:NM_001009931:exon3:c.T6477C:p.Y2159Y	rs75639578	0.1729	0/0:540,0	0/1:424,8	0/0:55,0	0/0:517,64	0/1:491,63	0/1:557,80
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152215459	C	T	HRNR	HRNR:NM_001009931:exon3:c.G6170A:p.R2057C	rs61814938	0.0093	0/1:135,36	0/1:188,25	0/1:148,2	0/1:205,22	0/1:124,35	0/1:206,26
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152215469	C	T	HRNR	HRNR:NM_001009931:exon3:c.G6160A:p.G2054	rs61814939	0.0230	0/1:143,53	0/1:180,3	0/1:149,5	0/1:201,47	0/1:123,50	0/1:197,42
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152215633	T	A	HRNR	HRNR:NM_001009931:exon3:c.A5996T:p.Q1999I	rs61814940	0.4897	0/1:167,47	0/1:108,2(	0/1:169,4	0/1:145,33	0/1:88,19	0/1:142,43
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152216371	A	G	HRNR	HRNR:NM_001009931:exon3:c.T5258C:p.V1753A	rs145667921	0.2283	0/1:152,29	0/1:96,15	0/1:143,2	0/0:144,12	0/1:80,12	0/1:127,16
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152216444	C	T	HRNR	HRNR:NM_001009931:exon3:c.G5185A:p.G1729	rs61814942	0.4012	0/1:234,93	0/1:127,54	0/1:220,9	0/1:206,62	0/1:109,51	0/1:159,79
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152216522	C	T	HRNR	HRNR:NM_001009931:exon3:c.G5107A:p.G1703	rs80018286	0.3654	0/1:168,59	0/1:124,1	0/1:197,4	0/1:162,43	0/1:118,20	0/1:144,43
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152218359	G	A	HRNR	HRNR:NM_001009931:exon3:c.C3270T:p.G1090(	rs76046733	0.2308	0/0:34,0	0/1:43,33	0/0:34,0	0/1:52,35	0/1:31,35	0/1:41,36
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152218469	C	T	HRNR	HRNR:NM_001009931:exon3:c.G3160A:p.E1054I	rs77376932	0.2286	0/0:33,0	0/1:20,19	0/0:34,0	0/1:33,22	0/1:20,22	0/1:31,34
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152219638	C	T	HRNR	HRNR:NM_001009931:exon3:c.G1991A:p.R664Q	rs7520249	0.7143	0/1:26,84	0/1:83,39	0/1:20,75	0/1:86,39	0/1:64,30	0/1:79,45
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152219687	G	T	HRNR	HRNR:NM_001009931:exon3:c.C1942A:p.R648S	rs61814943	9.061e-05	0/1:110,56	0/1:157,26	0/1:135,3	0/1:151,26	0/1:107,17	0/1:181,27
chr1	152215700	G	A	HRNR	HRNR:NM_001009931:exon3:c.C5929T:p.R1977C	rs77464974	0.2932	0/0:448,50	0/0:142,0	0/0:177,0	0/0:374,47	0/1:243,45	0/1:360,48
chr1	152219743	G	C	HRNR	HRNR:NM_001009931:exon3:c.C1886G:p.T629S	rs79513582	0.0011	0/1:100,56	0/1:128,23	0/1:129,3	0/1:143,27	0/1:107,17	0/1:180,27

Chr	Position	Wt allele	Alt. allele	Gene	Other info	RefSNP	Frequency	Genotype I-1	Genotype I-2	Genotype II-1	Genotype II-3	GenotypeII-2	Genotype II-4
chr1	21823625	C	T	HSPG2	HSPG2:NM_001291860:exon96:c.G12997A:p.V43	rs1138469	0.0585	0/1:17,34	0/0:40,0	0/1:21,14	0/0:30,0	0/1:16,14	0/1:24,15
chr1	21831627	G	T	HSPG2	.	rs2270701	0.1950	0/0:34,0	0/1:43,52	0/0:49,0	0/1:48,36	0/1:42,28	0/1:39,52
chr1	21823625	C	T	HSPG2	HSPG2:NM_001291860:exon96:c.G12997A:p.V43	rs1138469	0.0585	0/1:17,34	0/0:40,0	0/1:21,14	0/0:30,0	0/1:16,14	0/1:24,15
chr1	21832515	C	A	HSPG2	HSPG2:NM_001291860:exon81:c.G11190T:p.V37	rs2229487	0.1731	0/0:34,0	0/1:101,1	0/0:39,0	0/1:116,13	0/1:121,10	0/1:129,141
chr1	21823625	C	T	HSPG2	HSPG2:NM_001291860:exon96:c.G12997A:p.V43	rs1138469	0.0585	0/1:17,34	0/0:40,0	0/1:21,14	0/0:30,0	0/1:16,14	0/1:24,15
chr1	21833618	C	T	HSPG2	.	rs17459139	0.1757	0/0:48,0	0/1:26,42	0/0:52,0	0/1:25,39	0/1:23,52	0/1:37,47
chr1	21823625	C	T	HSPG2	HSPG2:NM_001291860:exon96:c.G12997A:p.V43	rs1138469	0.0585	0/1:17,34	0/0:40,0	0/1:21,14	0/0:30,0	0/1:16,14	0/1:24,15
chr1	21834782	G	A	HSPG2	HSPG2:NM_001291860:exon77:c.C10620T:p.V354	rs2229492	0.1748	0/0:33,0	0/1:124,9	0/0:34,0	0/1:117,93	0/1:99,67	0/1:121,88
chr1	21823625	C	T	HSPG2	HSPG2:NM_001291860:exon96:c.G12997A:p.V43	rs1138469	0.0585	0/1:17,34	0/0:40,0	0/1:21,14	0/0:30,0	0/1:16,14	0/1:24,15
chr1	21839494	G	A	HSPG2	HSPG2:NM_001291860:exon73:c.C9769T:p.H325	rs2291827	0.1728	0/0:44,0	0/1:87,55	0/0:46,0	0/1:68,57	0/1:49,50	0/1:82,56
chr1	21823625	C	T	HSPG2	HSPG2:NM_001291860:exon96:c.G12997A:p.V43	rs1138469	0.0585	0/1:17,34	0/0:40,0	0/1:21,14	0/0:30,0	0/1:16,14	0/1:24,15
chr1	21872278	G	A	HSPG2	HSPG2:NM_001291860:exon33:c.C4132T:p.P1378	rs759523072	0.0002	0/0:38,0	0/1:210,1	0/0:44,0	0/1:178,17	0/1:181,15	0/1:184,131
chr1	21823625	C	T	HSPG2	HSPG2:NM_001291860:exon96:c.G12997A:p.V43	rs1138469	0.0585	0/1:17,34	0/0:40,0	0/1:21,14	0/0:30,0	0/1:16,14	0/1:24,15
chr1	21872752	C	T	HSPG2	HSPG2:NM_001291860:exon32:c.G3900A:p.V130	rs2229482	0.6066	0/0:36,0	0/1:70,64	0/0:46,0	0/1:63,46	0/1:48,52	0/1:55,61
chr1	21831627	G	T	HSPG2	.	rs2270701	0.1950	0/0:34,0	0/1:43,52	0/0:49,0	0/1:48,36	0/1:42,28	0/1:39,52
chr1	21836909	G	A	HSPG2	HSPG2:NM_001291860:exon75:c.C10251T:p.S341	rs41310390	0.0976	0/1:73,59	0/0:49,0	0/1:55,81	0/0:44,0	0/1:85,60	0/1:121,76
chr1	21832515	C	A	HSPG2	HSPG2:NM_001291860:exon81:c.G11190T:p.V37	rs2229487	0.1731	0/0:34,0	0/1:101,1	0/0:39,0	0/1:116,13	0/1:121,10	0/1:129,141
chr1	21836909	G	A	HSPG2	HSPG2:NM_001291860:exon75:c.C10251T:p.S341	rs41310390	0.0976	0/1:73,59	0/0:49,0	0/1:55,81	0/0:44,0	0/1:85,60	0/1:121,76
chr1	21833618	C	T	HSPG2	.	rs17459139	0.1757	0/0:48,0	0/1:26,42	0/0:52,0	0/1:25,39	0/1:23,52	0/1:37,47
chr1	21836909	G	A	HSPG2	HSPG2:NM_001291860:exon75:c.C10251T:p.S341	rs41310390	0.0976	0/1:73,59	0/0:49,0	0/1:55,81	0/0:44,0	0/1:85,60	0/1:121,76
chr1	21834782	G	A	HSPG2	HSPG2:NM_001291860:exon77:c.C10620T:p.V354	rs2229492	0.1748	0/0:33,0	0/1:124,9	0/0:34,0	0/1:117,93	0/1:99,67	0/1:121,88
chr1	21836909	G	A	HSPG2	HSPG2:NM_001291860:exon75:c.C10251T:p.S341	rs41310390	0.0976	0/1:73,59	0/0:49,0	0/1:55,81	0/0:44,0	0/1:85,60	0/1:121,76
chr1	21836909	G	A	HSPG2	HSPG2:NM_001291860:exon75:c.C10251T:p.S341	rs41310390	0.0976	0/1:73,59	0/0:49,0	0/1:55,81	0/0:44,0	0/1:85,60	0/1:121,76
chr1	21839494	G	A	HSPG2	HSPG2:NM_001291860:exon73:c.C9769T:p.H325	rs2291827	0.1728	0/0:44,0	0/1:87,55	0/0:46,0	0/1:68,57	0/1:49,50	0/1:82,56
chr1	21836909	G	A	HSPG2	HSPG2:NM_001291860:exon75:c.C10251T:p.S341	rs41310390	0.0976	0/1:73,59	0/0:49,0	0/1:55,81	0/0:44,0	0/1:85,60	0/1:121,76
chr1	21872278	G	A	HSPG2	HSPG2:NM_001291860:exon33:c.C4132T:p.P1378	rs759523072	0.0002	0/0:38,0	0/1:210,1	0/0:44,0	0/1:178,17	0/1:181,15	0/1:184,131
chr1	21836909	G	A	HSPG2	HSPG2:NM_001291860:exon75:c.C10251T:p.S341	rs41310390	0.0976	0/1:73,59	0/0:49,0	0/1:55,81	0/0:44,0	0/1:85,60	0/1:121,76
chr1	21872752	C	T	HSPG2	HSPG2:NM_001291860:exon32:c.G3900A:p.V130	rs2229482	0.6066	0/0:36,0	0/1:70,64	0/0:46,0	0/1:63,46	0/1:48,52	0/1:55,61
chr1	14960672	C	T	KAZN	.	rs41301991	0.0530	0/1:67,73	0/0:41,0	0/1:49,66	0/0:46,0	0/1:71,47	0/1:78,79
chr1	15056111	G	A	KAZN	KAZN:NM_001017999:exon5:c.G465A:p.T155T,I	rs79632856	0.0478	0/0:34,0	0/1:36,41	0/0:39,0	0/0:45,0	0/1:24,32	0/1:56,43
chr1	18481057	G	A	KLHDC7	KLHDC7A:NM_152375:exon1:c.G76A:p.A26T	rs201286096	0.0029	0/1:151,86	0/0:55,0	0/1:161,1	0/0:41,0	0/1:147,84	0/1:196,99
chr1	18481798	C	A	KLHDC7	KLHDC7A:NM_152375:exon1:c.C817A:p.H273N	rs2992753	0.6539	0/0:35,0	0/1:109,6	0/0:35,0	0/1:101,36	0/1:76,44	0/1:107,69
chr1	18481057	G	A	KLHDC7	KLHDC7A:NM_152375:exon1:c.G76A:p.A26T	rs201286096	0.0029	0/1:151,86	0/0:55,0	0/1:161,1	0/0:41,0	0/1:147,84	0/1:196,99
chr1	18482032	A	C	KLHDC7	KLHDC7A:NM_152375:exon1:c.A1051C:p.T351I	rs2992752	0.6552	0/0:34,0	0/1:24,26	0/0:30,0	0/1:25,24	0/1:27,16	0/1:26,32
chr1	3067399	G	A	LINC009	.	rs2297828	0.3098	0/0:48,0	0/1:107,7	0/0:55,0	0/0:61,0	0/1:100,76	0/1:113,75
chr1	3067523	C	A	LINC009	.	rs2297829	0.3914	0/1:41,49	0/0:50,0	0/1:36,42	0/0:61,0	0/1:35,32	0/1:27,37
chr1	11121270	C	T	MTOR	MTOR:NM_004958:exon49:c.G6909A:p.L2303L	rs11121691	0.2184	0/0:47,0	0/1:29,47	0/0:45,0	0/0:36,0	0/1:27,30	0/1:39,39
chr1	11145001	C	T	MTOR	MTOR:NM_004958:exon33:c.G4731A:p.A1577A	rs1057079	0.6779	0/1:42,20	0/0:45,0	0/1:55,9	0/0:63,0	0/1:46,28	0/1:55,36
chr1	11121270	C	T	MTOR	MTOR:NM_004958:exon49:c.G6909A:p.L2303L	rs11121691	0.2184	0/0:47,0	0/1:29,47	0/0:45,0	0/0:36,0	0/1:27,30	0/1:39,39
chr1	11241657	A	G	MTOR	MTOR:NM_004958:exon10:c.T1437C:p.D479D	rs1135172	0.7170	0/1:119,70	0/0:76,0	0/1:97,54	0/0:69,0	0/1:85,47	0/1:88,62
chr1	6466123	G	A	TNFRSF2	.	rs3007417	0.1694	0/0:35,0	0/1:65,72	0/0:41,0	0/0:35,0	0/1:74,69	0/1:69,77
chr1	6468505	G	A	PLEKHG	PLEKHG5:NM_001042664:exon20:c.C2331T:p.S7	rs61749272	0.0097	0/1:74,87	0/0:33,0	0/1:99,82	0/0:45,0	0/1:107,72	0/1:107,98
chr1	11949898	C	T	PLOD1	PLOD1:NM_000302:exon3:c.C294T:p.F98F,PLOI	rs7529452	0.3281	0/0:36,0	0/1:112,7	0/0:43,0	0/0:48,0	0/1:74,52	0/1:84,84
chr1	11958608	C	T	PLOD1	PLOD1:NM_000302:exon9:c.C936T:p.H312H,PL	rs374787907	0.0006	0/1:146,44	0/0:33,0	0/1:135,1	0/0:45,0	0/1:134,87	0/1:157,83
chr1	11958608	C	T	PLOD1	PLOD1:NM_000302:exon9:c.C936T:p.H312H,PL	rs374787907	0.0006	0/1:146,44	0/0:33,0	0/1:135,1	0/0:45,0	0/1:134,87	0/1:157,83
chr1	11964178	C	T	PLOD1	PLOD1:NM_000302:exon12:c.C1206T:p.N402N,I	rs1130529	0.3631	0/0:33,0	0/1:92,63	0/0:33,0	0/0:48,0	0/1:44,67	0/1:95,87
chr1	19325569	G	A	PQLC2	PQLC2:NM_001040125:exon4:c.G369A:p.T123T,	rs11558390	0.2477	0/0:35,0	0/1:17,17	0/0:39,0	0/1:16,18	0/1:28,16	0/1:21,12
chr1	19326380	G	A	PQLC2	PQLC2:NM_001040125:exon5:c.G518A:p.G173D	rs41264079	0.0893	0/1:52,49	0/0:38,0	0/1:50,59	0/0:43,0	0/1:36,50	0/1:43,70
chr1	12794279	C	T	PRAMEF	PRAMEF1:NM_023013:exon3:c.C652T:p.P218S	rs1769774	0.3045	0/1:8,36	0/0:57,0	0/1:27,27	0/1:36,28	0/1:39,22	0/1:28,29
chr1	12794381	G	C	PRAMEF	PRAMEF1:NM_023013:exon3:c.G754C:p.E252Q	rs1063776	.	0/0:45,0	0/1:35,33	0/0:45,0	0/0:52,0	0/1:24,32	0/1:29,29
chr1	21619037	C	T	RAP1GA	RAP1GAP:NM_001145658:exon5:c.G246A:p.P82I	rs77539762	0.0551	0/1:147,78	0/0:130,0	0/1:129,9	0/0:128,0	0/1:132,68	0/1:135,87
chr1	21649771	G	A	RAP1GA	RAP1GAP:NM_001145658:exon2:c.C70T:p.R24W	rs372126308	0.0026	0/0:44,0	0/1:54,21	0/0:34,0	0/1:47,24	0/1:46,20	0/1:48,24

Chr	Position	Wt allele	Alt. allele	Gene	Other info	RefSNP	Frequency	Genotype I-1	Genotype I-2	Genotype II-1	Genotype II-3	GenotypeII-2	Genotype II-4
chr1	1732392	C	T	SLC35E2	.	rs3817856	.	0/0:97,0	0/1:55,35	0/0:121,0	0/0:105,0	0/1:75,28	0/1:82,43
chr1	1734812	G	A	SLC35E2	SLC35E2:NM_182838:exon5:c.C610T:p.R204W,SI	rs77655487	0.3081	0/1:53,15	0/0:46,0	0/0:52,0	0/0:64,0	1/1:0,5	1/1:0,12
chr1	1734736	C	T	SLC35E2	SLC35E2:NM_182838:exon5:c.G686A:p.R229H,SI	rs61777509	0.4048	0/0:36,0	0/1:83,46	0/1:73,48	0/1:141,52	0/0:8,0	0/0:11,0
chr1	1734812	G	A	SLC35E2	SLC35E2:NM_182838:exon5:c.C610T:p.R204W,SI	rs77655487	0.3081	0/1:53,15	0/0:46,0	0/0:52,0	0/0:64,0	1/1:0,5	1/1:0,12
chr1	18839800	G	A	TAS1R2	TAS1R2:NM_152232:exon6:c.C2319T:p.S773S	rs12033832	0.3122	0/1:36,31	0/0:45,0	0/1:30,27	0/0:46,0	0/1:33,26	0/1:55,51
chr1	18854521	G	C	TAS1R2	TAS1R2:NM_152232:exon3:c.C949G:p.R317G	rs34447754	0.2886	0/0:40,0	0/1:46,30	0/0:41,0	0/1:46,27	0/1:43,22	0/1:53,32
chr1	18839800	G	A	TAS1R2	TAS1R2:NM_152232:exon6:c.C2319T:p.S773S	rs12033832	0.3122	0/1:36,31	0/0:45,0	0/1:30,27	0/0:46,0	0/1:33,26	0/1:55,51
chr1	18854588	A	C	TAS1R2	TAS1R2:NM_152232:exon3:c.T882G:p.T294T	rs28470550	0.2893	0/0:40,0	0/1:156,85	0/0:41,0	0/1:162,62	0/1:148,70	0/1:208,58
chr10	49616071	G	A	CHAT	CHAT:NM_001142933:exon2:c.G19A:p.D7N	rs1880676	0.2112	0/1:24,13	0/0:51,0	0/0:41,0	0/0:35,0	0/1:26,11	0/1:22,11
chr10	49622125	C	T	CHAT	CHAT:NM_001142929:exon5:c.C373T:p.L125F,C	rs8178990	0.0467	0/0:33,0	0/1:61,47	0/1:57,49	0/0:36,0	0/1:61,28	0/1:56,31
chr10	49616573	G	A	CHAT	CHAT:NM_001142929:exon2:c.G4A:p.A2T,CHA	rs3810950	0.2269	0/1:40,34	0/0:36,0	0/0:44,0	0/0:34,0	0/1:33,23	0/1:67,35
chr10	49622125	C	T	CHAT	CHAT:NM_001142929:exon5:c.C373T:p.L125F,C	rs8178990	0.0467	0/0:33,0	0/1:61,47	0/1:57,49	0/0:36,0	0/1:61,28	0/1:56,31
chr10	5914241	C	T	FBXO18	FBXO18:NM_001258452:exon8:c.C1146T:p.L382I	rs78214360	0.0078	0/1:21,31	0/0:44,0	0/0:39,0	0/0:35,0	0/1:21,18	0/1:38,27
chr10	5917639	C	T	FBXO18	FBXO18:NM_001258452:exon12:c.C1704T:p.D56	rs11553630	0.0266	0/0:58,0	0/1:71,48	0/1:61,54	0/0:65,0	0/1:53,40	0/1:66,36
chr10	27155399	C	A	MASTL	.	rs7919803	0.6445	0/0:120,0	0/0:75,5	0/0:132,0	0/0:78,9	0/1:61,20	0/0:183,0
chr10	27170817	C	G	MASTL	MASTL:NM_001172303:exon8:c.C1858G:p.P620A	rs3802526	0.0842	0/1:19,21	0/1:36,24	0/1:31,17	0/0:36,0	0/1:21,16	1/1:0,30
chr10	27170817	C	G	MASTL	MASTL:NM_001172303:exon8:c.C1858G:p.P620A	rs3802526	0.0842	0/1:19,21	0/1:36,24	0/1:31,17	0/0:36,0	0/1:21,16	1/1:0,30
chr10	27186515	T	C	MASTL	MASTL:NM_001172304:exon11:c.T2502C:p.T834	rs10741130	0.5971	0/0:36,0	0/0:34,0	0/0:35,0	0/1:30,12	0/1:18,14	0/0:31,0
chr10	50068003	A	G	WASHC2	.	rs4935425	0.2883	0/1:88,99	0/0:136,0	0/0:87,0	0/0:102,0	0/1:100,95	0/1:94,103
chr10	50068174	G	A	WASHC2	WASHC2A:NM_001005751:exon2:c.G73A:p.V25M	rs55875820	0.4623	0/0:54,0	0/1:72,62	0/1:60,66	0/0:52,0	0/1:84,74	0/1:83,74
chr10	50068174	G	A	WASHC2	WASHC2A:NM_001005751:exon2:c.G73A:p.V25M	rs55875820	0.4623	0/0:54,0	0/1:72,62	0/1:60,66	0/0:52,0	0/1:84,74	0/1:83,74
chr10	50078599	T	C	WASHC2	.	rs2305300	.	0/1:25,29	0/0:33,0	0/0:43,0	0/0:44,0	0/1:28,18	0/1:51,28
chr10	50068174	G	A	WASHC2	WASHC2A:NM_001005751:exon2:c.G73A:p.V25M	rs55875820	0.4623	0/0:54,0	0/1:72,62	0/1:60,66	0/0:52,0	0/1:84,74	0/1:83,74
chr10	50095550	C	T	WASHC2	.	rs781790179	0.4174	0/1:31,12	0/0:42,0	0/0:35,0	0/0:45,0	0/1:17,15	0/1:22,21
chr10	50068174	G	A	WASHC2	WASHC2A:NM_001005751:exon2:c.G73A:p.V25M	rs55875820	0.4623	0/0:54,0	0/1:72,62	0/1:60,66	0/0:52,0	0/1:84,74	0/1:83,74
chr10	50095689	G	A	WASHC2	WASHC2A:NM_001005751:exon15:c.G1331A:p.C	rs2305302	0.4964	0/1:31,7	0/0:39,0	0/0:52,0	0/0:43,0	0/1:30,19	0/1:40,8
chr10	50068174	G	A	WASHC2	WASHC2A:NM_001005751:exon2:c.G73A:p.V25M	rs55875820	0.4623	0/0:54,0	0/1:72,62	0/1:60,66	0/0:52,0	0/1:84,74	0/1:83,74
chr10	50095690	T	G	WASHC2	WASHC2A:NM_001005751:exon15:c.T1332G:p.C	rs2305303	0.4760	0/1:30,7	0/0:39,0	0/0:52,0	0/0:43,0	0/1:26,19	0/1:35,8
chr10	48721286	G	A	WDFY4	WDFY4:NM_020945:exon4:c.G375A:p.Q125Q	rs11818741	0.0683	0/0:73,0	0/1:136,55	0/1:101,6	0/0:73,0	0/1:111,45	0/1:118,73
chr10	48817351	G	A	WDFY4	WDFY4:NM_020945:exon32:c.G5447A:p.R1816Q	rs7097397	0.3754	0/1:95,53	0/0:36,0	0/0:39,0	0/0:34,0	0/1:82,55	0/1:100,80
chr11	129905704	C	T	PRDM10	PRDM10:NM_199439:exon15:c.G2802A:p.Q934Q	rs3734076	0.0764	0/0:54,0	0/0:37,0	0/1:27,13	0/0:36,0	0/1:18,10	0/1:26,8
chr11	129915699	G	A	PRDM10	PRDM10:NM_199439:exon12:c.C2229T:p.P743P,I	rs3734073	0.2889	0/0:36,0	0/1:26,37	0/0:35,0	0/1:41,48	0/1:31,23	0/1:30,34
chr11	129914869	C	T	PRDM10	PRDM10:NM_199439:exon13:c.G2418A:p.T806T	rs117750366	0.0191	0/1:45,32	0/0:34,0	0/1:69,17	0/0:41,0	0/1:54,28	0/1:66,26
chr11	129915699	G	A	PRDM10	PRDM10:NM_199439:exon12:c.C2229T:p.P743P,I	rs3734073	0.2889	0/0:36,0	0/1:26,37	0/0:35,0	0/1:41,48	0/1:31,23	0/1:30,34
chr12	47980015	G	C	COL2A1	COL2A1:NM_033150:exon39:c.C2466G:p.G822C	rs41272029	0.0147	0/1:97,33	0/0:52,0	0/1:109,4	0/0:41,0	0/1:97,48	0/1:132,59
chr12	47982816	C	T	COL2A1	.	rs3829735	0.0269	0/0:40,0	0/1:46,37	0/0:46,0	0/0:33,0	0/1:25,29	0/1:36,40

Chr	Position	Wt allele	Alt. allele	Gene	Other info	RefSNP	Frequency	Genotype I-1	Genotype I-2	Genotype II-1	Genotype II-3	GenotypeII-2	Genotype II-4
chr12	31091769	C	T	DDX11	DDX11:NM_001257144:exon10:c.C1140T:p.A380	rs74439745	0.0131	0/1:276,50	0/0:436,50	0/0:452,5	0/1:414,65	0/1:291,58	0/1:447,78
chr12	31096774	C	T	DDX11	.	rs71455622	0.0706	0/0:92,0	0/1:99,85	0/0:110,0	0/0:99,0	0/1:80,62	0/1:98,92
chr12	31091769	C	T	DDX11	DDX11:NM_001257144:exon10:c.C1140T:p.A380	rs74439745	0.0131	0/1:276,50	0/0:436,50	0/0:452,5	0/1:414,65	0/1:291,58	0/1:447,78
chr12	31103583	C	T	DDX11	DDX11:NM_004399:exon25:c.C2393T:p.A798V,I	rs140335402	0.0725	0/0:33,0	0/1:36,53	0/0:68,0	0/0:34,0	0/1:32,36	0/1:34,32
chr12	31096774	C	T	DDX11	.	rs71455622	0.0706	0/0:92,0	0/1:99,85	0/0:110,0	0/0:99,0	0/1:80,62	0/1:98,92
chr12	31103919	C	T	DDX11	DDX11:NM_001257144:exon27:c.C2799T:p.C933	rs368711403	5.353e-05	0/1:33,35	0/0:57,0	0/1:52,25	0/0:44,0	0/1:56,25	0/1:35,27
chr12	31103583	C	T	DDX11	DDX11:NM_004399:exon25:c.C2393T:p.A798V,I	rs140335402	0.0725	0/0:33,0	0/1:36,53	0/0:68,0	0/0:34,0	0/1:32,36	0/1:34,32
chr12	31103919	C	T	DDX11	DDX11:NM_001257144:exon27:c.C2799T:p.C933	rs368711403	5.353e-05	0/1:33,35	0/0:57,0	0/1:52,25	0/0:44,0	0/1:56,25	0/1:35,27
chr12	122861209	G	GC	HIP1R	.	rs3215148	0.8341	0/0:118,0	0/1:83,77	0/1:54,38	0/0:145,0	0/1:66,64	0/1:42,63
chr12	122861209	G	GC	HIP1R	.	rs3215148	0.8341	0/0:118,0	0/1:83,77	0/1:54,38	0/0:145,0	0/1:66,64	0/1:42,63
chr13	113428957	G	A	ADPRHL	ADPRHL1:NM_138430:exon4:c.C641T:p.T214M,	rs563174718	3.372e-05	0/0:88,0	0/1:85,75	0/0:103,0	0/0:117,0	0/1:39,62	0/1:35,87
chr13	113429001	C	T	ADPRHL	ADPRHL1:NM_138430:exon4:c.G597A:p.A199A,	rs9635114	0.0356	0/1:137,82	0/0:136,0	0/1:179,9	0/0:117,0	0/1:150,78	0/1:191,91
chr13	111153898	A	G	ARHGEF	.	rs61966721	0.4176	0/1:64,35	0/0:49,0	0/1:80,40	0/0:37,0	0/1:73,29	0/1:88,39
chr13	111286164	T	C	ARHGEF	ARHGEF7:NM_001320854:exon12:c.T1263C:p.P4	rs72653559	0.0127	0/0:97,0	0/1:118,8	0/0:102,0	0/0:107,0	0/1:80,60	0/1:125,68
chr13	110166251	T	G	COL4A1	COL4A1:NM_001845:exon45:c.A4002C:p.Q1334I	rs3742207	0.3098	0/0:37,0	0/1:34,9	0/0:33,0	0/0:40,0	0/1:30,20	0/1:40,25
chr13	110175233	C	T	COL4A1	COL4A1:NM_001845:exon37:c.G3183A:p.G1061	rs874204	0.3515	0/1:18,18	0/0:43,0	0/1:26,17	0/0:42,0	1/1:0,31	1/1:0,45
chr13	110166251	T	G	COL4A1	COL4A1:NM_001845:exon45:c.A4002C:p.Q1334I	rs3742207	0.3098	0/0:37,0	0/1:34,9	0/0:33,0	0/0:40,0	0/1:30,20	0/1:40,25
chr13	110183113	C	T	COL4A1	.	rs72654112	0.1152	0/1:36,27	0/0:42,0	0/1:22,26	0/0:48,0	0/1:26,20	0/1:13,22
chr13	110458826	G	A	COL4A2	COL4A2:NM_001846:exon22:c.G1488A:p.P496P	rs7990214	0.5288	0/0:46,0	0/1:202,89	0/0:51,0	0/0:52,0	0/1:165,11	0/1:176,106
chr13	110493286	C	T	COL4A2	.	rs2274544	0.0916	0/1:124,70	0/0:34,0	0/1:126,1	0/0:37,0	0/1:106,10	0/1:86,79
chr13	110458888	G	A	COL4A2	COL4A2:NM_001846:exon22:c.G1550A:p.R517K	rs7990383	0.5390	0/0:46,0	0/1:71,40	0/0:51,0	0/0:52,0	0/1:65,61	0/1:83,62
chr13	110493286	C	T	COL4A2	.	rs2274544	0.0916	0/1:124,70	0/0:34,0	0/1:126,1	0/0:37,0	0/1:106,10	0/1:86,79
chr13	110469273	C	T	COL4A2	COL4A2:NM_001846:exon28:c.C2152T:p.P718S	rs9583500	0.2554	0/0:117,0	0/1:215,10	0/0:136,0	0/0:161,0	0/1:163,10	0/1:183,121
chr13	110493286	C	T	COL4A2	.	rs2274544	0.0916	0/1:124,70	0/0:34,0	0/1:126,1	0/0:37,0	0/1:106,10	0/1:86,79
chr13	109055109	A	G	MYO16	MYO16:NM_001198950:exon26:c.A3112G:p.I103	rs146716803	0.0017	0/0:33,0	0/1:30,20	0/0:34,0	0/0:40,0	0/1:17,9	0/1:37,21
chr13	109140829	G	A	MYO16	MYO16:NM_001198950:exon32:c.G4617A:p.K15	rs150813880	0.0309	0/1:27,38	0/0:44,0	0/1:54,39	0/0:35,0	0/1:51,28	0/1:27,26
chr13	109140829	G	A	MYO16	MYO16:NM_001198950:exon32:c.G4617A:p.K15	rs150813880	0.0309	0/1:27,38	0/0:44,0	0/1:54,39	0/0:35,0	0/1:51,28	0/1:27,26
chr13	109141076	G	T	MYO16	MYO16:NM_001198950:exon32:c.G4864T:p.A162	rs113648411	0.2225	0/0:41,0	0/1:11,25	0/0:39,0	0/0:42,0	0/1:15,6	0/1:18,19
chr14	70457520	C	A	ADAM21	ADAM21:NM_003813:exon2:c.C21A:p.L7L	rs3751520	0.2560	0/0:58,0	0/1:50,25	0/0:57,0	0/1:72,24	0/1:62,10	0/1:73,14
chr14	70457733	C	T	ADAM21	ADAM21:NM_003813:exon2:c.C234T:p.V78V	rs58247196	0.3772	0/1:19,10	0/0:46,3	0/1:32,6	0/0:66,6	0/1:20,3	0/1:44,11
chr14	70457733	C	T	ADAM21	ADAM21:NM_003813:exon2:c.C234T:p.V78V	rs58247196	0.3772	0/1:19,10	0/0:46,3	0/1:32,6	0/0:66,6	0/1:20,3	0/1:44,11
chr14	70458135	A	C	ADAM21	ADAM21:NM_003813:exon2:c.A636C:p.L212L	rs3829455	0.4255	0/0:43,0	0/1:69,25	0/0:50,0	0/1:65,22	0/1:60,14	0/1:82,21
chr16	89767168	G	C	FANCA	FANCA:NM_000135:exon27:c.C2574G:p.S858R,F	rs17233141	0.0103	0/0:36,0	0/1:49,9	0/0:33,0	0/1:40,17	0/1:31,5	0/1:58,10
chr16	89816599	A	T	FANCA	FANCA:NM_000135:exon1:c.T17A:p.V6D,FANC	rs1800282	0.1252	0/0:50,0	0/1:112,12	0/0:91,0	0/1:107,10	0/1:94,66	0/1:137,71
chr19	39875734	G	C	FCGBP	FCGBP:NM_003890:exon23:c.C10257G:p.P3419F	rs2053089	0.2047	0/1:86,85	0/0:68,0	0/1:102,7	0/0:54,0	0/1:99,65	0/1:107,92
chr19	39885638	A	G	FCGBP	.	rs756374614	6.99e-05	0/0:52,5	0/0:65,5	0/0:44,0	0/1:64,8	0/1:41,23	0/1:56,8
chr19	39885638	A	G	FCGBP	.	rs756374614	6.99e-05	0/0:52,5	0/0:65,5	0/0:44,0	0/1:64,8	0/1:41,23	0/1:56,8
chr19	39893486	G	A	FCGBP	FCGBP:NM_003890:exon14:c.C5881T:p.P1961S	.	.	0/1:106,21	0/0:35,0	0/1:122,2	0/0:33,0	0/1:70,16	0/1:122,24

Chr	Position	Wt allele	Alt. allele	Gene	Other info	RefSNP	Frequency	Genotype I-1	Genotype I-2	Genotype II-1	Genotype II-3	GenotypeII-2	Genotype II-4
chr19	33391316	G	A	PEPD	PEPD:NM_001166056:exon11:c.C1008T:p.H336H	rs17569	0.1484	0/1:36,17	0/0:41,0	0/1:17,25	0/0:33,0	0/1:33,19	0/1:21,32
chr19	33463006	A	G	PEPD	PEPD:NM_001166057:exon7:c.T468C:p.Y156Y,PI	rs3745969	0.0378	0/0:44,0	0/1:43,16	0/0:42,0	0/1:27,12	0/1:25,29	0/1:43,10
chr2	3613266	G	A	COLEC1	.	rs7567724	0.1328	0/0:89,0	0/1:87,70	0/1:55,64	0/1:62,57	0/1:68,52	0/1:91,72
chr2	3613279	G	A	COLEC1	.	rs17017752	0.0540	0/1:74,72	0/0:40,0	0/0:42,0	0/0:43,0	0/1:57,75	0/1:78,109
chr2	3613266	G	A	COLEC1	.	rs7567724	0.1328	0/0:89,0	0/1:87,70	0/1:55,64	0/1:62,57	0/1:68,52	0/1:91,72
chr2	3613412	G	A	COLEC1	.	rs67826307	0.0554	0/1:15,27	0/0:40,0	0/0:42,0	0/0:31,0	0/1:13,27	0/1:22,28
chr2	3613279	G	A	COLEC1	.	rs17017752	0.0540	0/1:74,72	0/0:40,0	0/0:42,0	0/0:43,0	0/1:57,75	0/1:78,109
chr2	3637573	T	C	COLEC1	COLEC11:NM_001255987:exon2:c.T93C:p.R31R,	rs34347618	0.0444	0/0:36,0	0/1:41,29	0/1:33,15	0/1:38,26	0/1:28,10	0/1:34,20
chr2	3613412	G	A	COLEC1	.	rs67826307	0.0554	0/1:15,27	0/0:40,0	0/0:42,0	0/0:31,0	0/1:13,27	0/1:22,28
chr2	3637573	T	C	COLEC1	COLEC11:NM_001255987:exon2:c.T93C:p.R31R,	rs34347618	0.0444	0/0:36,0	0/1:41,29	0/1:33,15	0/1:38,26	0/1:28,10	0/1:34,20
chr2	11566567	T	C	GREB1	GREB1:NM_014668:exon4:c.T365C:p.V122A,GR	rs4669751	0.1939	0/1:153,12	0/0:57,0	0/0:45,0	0/1:177,17	0/1:138,13	0/1:165,165
chr2	11634198	G	A	GREB1	GREB1:NM_014668:exon29:c.G5059A:p.D1687N	rs2304402	0.4485	0/0:33,0	0/1:37,56	0/0:34,0	0/0:36,0	0/1:25,32	0/1:38,56
chr20	10652213	G	A	JAG1	JAG1:NM_000214:exon7:c.C924T:p.N308N	rs45575136	0.0337	0/1:50,18	0/0:39,0	0/1:30,17	0/0:36,0	0/1:28,11	0/1:42,20
chr20	10658574	G	A	JAG1	JAG1:NM_000214:exon4:c.C588T:p.C196C	rs1801138	0.1088	0/0:37,0	0/1:31,13	0/0:27,0	0/1:28,8	0/1:28,6	0/1:25,8
chr20	10656409	T	C	JAG1	JAG1:NM_000214:exon5:c.A744G:p.P248P	rs10485741	0.0716	0/1:33,8	0/0:47,0	0/1:21,18	0/0:37,0	0/1:24,31	0/1:41,21
chr20	10658574	G	A	JAG1	JAG1:NM_000214:exon4:c.C588T:p.C196C	rs1801138	0.1088	0/0:37,0	0/1:31,13	0/0:27,0	0/1:28,8	0/1:28,6	0/1:25,8
chr5	128261768	T	G	FBN2	FBN2:NM_001999:exon64:c.A8332C:p.K2778Q	rs371715068	0.0015	0/1:38,21	0/0:33,0	0/1:14,22	0/1:24,20	0/1:20,23	0/1:25,31
chr5	128395085	G	C	FBN2	.	rs28763953	0.0327	0/0:33,0	0/1:33,28	0/0:42,0	0/0:36,0	0/1:17,24	0/1:33,30
chr5	128278780	A	G	FBN2	FBN2:NM_001999:exon57:c.T7200C:p.S2400S	rs190450	0.7281	0/1:50,30	0/0:39,0	0/1:34,16	0/1:45,26	0/1:38,23	0/1:40,37
chr5	128395085	G	C	FBN2	.	rs28763953	0.0327	0/0:33,0	0/1:33,28	0/0:42,0	0/0:36,0	0/1:17,24	0/1:33,30
chr5	128312808	A	G	FBN2	.	rs10044959	0.0858	0/1:60,35	0/0:34,0	0/1:46,54	0/1:71,45	0/1:50,27	0/1:53,53
chr5	128395085	G	C	FBN2	.	rs28763953	0.0327	0/0:33,0	0/1:33,28	0/0:42,0	0/0:36,0	0/1:17,24	0/1:33,30
chr5	128395085	G	C	FBN2	.	rs28763953	0.0327	0/0:33,0	0/1:33,28	0/0:42,0	0/0:36,0	0/1:17,24	0/1:33,30
chr5	128537421	G	C	FBN2	FBN2:NM_001999:exon1:c.C183G:p.P61P	rs73348287	0.0920	0/1:76,47	0/0:33,0	0/1:120,7	0/1:113,70	0/1:90,67	0/1:86,93
chr5	128395085	G	C	FBN2	.	rs28763953	0.0327	0/0:33,0	0/1:33,28	0/0:42,0	0/0:36,0	0/1:17,24	0/1:33,30
chr5	128537493	C	T	FBN2	FBN2:NM_001999:exon1:c.G111A:p.P37P	rs55715053	0.1174	0/1:62,43	0/0:33,0	0/1:87,57	0/1:93,54	0/1:74,57	0/1:63,78
chr5	135936655	C	G	FBXL21	UNKNOWN	rs760974883	0.0001	0/0:33,0	0/1:46,26	0/0:37,0	0/0:36,0	0/1:49,16	0/1:56,22
chr5	135941158	T	C	FBXL21	UNKNOWN	rs31547	0.7350	0/1:29,17	0/0:20,0	0/0:41,0	0/0:33,0	0/1:22,4	0/1:25,12
chr5	236502	G	T	SDHA	SDHA:NM_001294332:exon9:c.G1191T:p.S397S,S	rs200223188	0.0001	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:90,16	0/1:119,18
chr5	236504	T	C	SDHA	SDHA:NM_001294332:exon9:c.T1193C:p.V398A,S	rs201741295	9.889e-05	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:99,16	0/1:127,18
chr5	236502	G	T	SDHA	SDHA:NM_001294332:exon9:c.G1191T:p.S397S,S	rs200223188	0.0001	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:90,16	0/1:119,18
chr5	236513	C	T	SDHA	SDHA:NM_001294332:exon9:c.C1202T:p.A401V,S	rs201139275	9.89e-05	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:122,20	0/1:147,20
chr5	236502	G	T	SDHA	SDHA:NM_001294332:exon9:c.G1191T:p.S397S,S	rs200223188	0.0001	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:90,16	0/1:119,18
chr5	236534	C	T	SDHA	SDHA:NM_001294332:exon9:c.C1223T:p.S408L,S	rs76896145	0.0207	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:129,20	0/1:155,22
chr5	236502	G	T	SDHA	SDHA:NM_001294332:exon9:c.G1191T:p.S397S,S	rs200223188	0.0001	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:90,16	0/1:119,18
chr5	236538	C	A	SDHA	SDHA:NM_001294332:exon9:c.C1227A:p.L409L,S	rs75091805	0.0209	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:132,20	0/1:160,23
chr5	236504	T	C	SDHA	SDHA:NM_001294332:exon9:c.T1193C:p.V398A,S	rs201741295	9.889e-05	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:99,16	0/1:127,18
chr5	236513	C	T	SDHA	SDHA:NM_001294332:exon9:c.C1202T:p.A401V,S	rs201139275	9.89e-05	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:122,20	0/1:147,20
chr5	236504	T	C	SDHA	SDHA:NM_001294332:exon9:c.T1193C:p.V398A,S	rs201741295	9.889e-05	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:99,16	0/1:127,18
chr5	236534	C	T	SDHA	SDHA:NM_001294332:exon9:c.C1223T:p.S408L,S	rs76896145	0.0207	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:129,20	0/1:155,22
chr5	236504	T	C	SDHA	SDHA:NM_001294332:exon9:c.T1193C:p.V398A,S	rs201741295	9.889e-05	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:99,16	0/1:127,18
chr5	236538	C	A	SDHA	SDHA:NM_001294332:exon9:c.C1227A:p.L409L,S	rs75091805	0.0209	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:132,20	0/1:160,23
chr5	236513	C	T	SDHA	SDHA:NM_001294332:exon9:c.C1202T:p.A401V,S	rs201139275	9.89e-05	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:122,20	0/1:147,20
chr5	236534	C	T	SDHA	SDHA:NM_001294332:exon9:c.C1223T:p.S408L,S	rs76896145	0.0207	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:129,20	0/1:155,22
chr5	236513	C	T	SDHA	SDHA:NM_001294332:exon9:c.C1202T:p.A401V,S	rs201139275	9.89e-05	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:122,20	0/1:147,20
chr5	236538	C	A	SDHA	SDHA:NM_001294332:exon9:c.C1227A:p.L409L,S	rs75091805	0.0209	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:132,20	0/1:160,23

Chr	Position	Wt allele	Alt. allele	Gene	Other info	RefSNP	Frequency	Genotype I-1	Genotype I-2	Genotype II-1	Genotype II-3	GenotypeII-2	Genotype II-4
chr5	236534	C	T	SDHA	SDHA:NM_001294332:exon9:c.C1223T:p.S408L,S	rs76896145	0.0207	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:129,20	0/1:155,22
chr5	236538	C	A	SDHA	SDHA:NM_001294332:exon9:c.C1227A:p.L409L,S	rs75091805	0.0209	0/0:34,0	0/0:33,0	0/0:35,0	0/0:43,0	0/1:132,20	0/1:160,23
chr6	3010103	A	G	NQO2	NQO2:NM_000904:exon3:c.A86G:p.E29G,NQO	rs17136117	0.0228	0/0:41,0	0/1:61,44	0/0:45,0	0/0:35,0	0/1:49,17	0/1:68,33
chr6	3012544	G	A	NQO2	NQO2:NM_000904:exon4:c.G173A:p.G58D,NQ(	rs17300141	0.0294	0/1:33,24	0/0:42,0	0/1:22,23	0/1:30,22	0/1:33,16	0/1:24,25
chr6	3012544	G	A	NQO2	NQO2:NM_000904:exon4:c.G173A:p.G58D,NQ(	rs17300141	0.0294	0/1:33,24	0/0:42,0	0/1:22,23	0/1:30,22	0/1:33,16	0/1:24,25
chr6	3015631	C	T	NQO2	NQO2:NM_000904:exon5:c.C405T:p.S135S,NQC	rs4149367	0.1148	0/0:75,0	0/1:55,64	0/0:86,0	0/0:86,0	0/1:50,46	0/1:70,36
chr7	77256452	A	G	CCDC14(	CCDC146:NM_020879:exon6:c.A627G:p.Q209Q	rs3108428	0.6255	0/1:44,20	0/0:36,0	0/0:35,0	0/0:33,0	0/1:53,15	0/1:39,24
chr7	77262168	A	G	CCDC14(	CCDC146:NM_020879:exon9:c.A1034G:p.N345S	rs1109968	0.1181	0/0:33,0	0/1:40,38	0/0:45,0	0/1:46,47	0/1:40,44	0/1:48,52
chr7	74518230	A	G	GTF2IRI	GTF2IRD1:NM_001199207:exon5:c.A609G:p.E2(	rs17851629	0.2026	0/0:37,0	0/1:111,7	0/0:34,0	0/1:99,78	0/1:75,67	0/1:116,70
chr7	74535081	T	C	GTF2IRI	.	rs59656369	0.0745	0/1:15,15	0/0:44,0	0/0:46,0	0/0:35,0	0/1:21,14	0/1:28,20
chr9	137986862	A	C	CACNA1	.	rs7034745	0.2662	0/0:115,0	0/1:68,96	0/0:148,0	0/0:128,0	0/1:106,71	0/1:78,69
chr9	138114424	C	T	CACNA1	CACNA1B:NM_000718:exon40:c.C5583T:p.F186	rs2229948	0.7159	0/1:24,8	0/0:37,0	0/0:33,0	0/1:22,9	0/1:30,10	0/1:31,18
chr9	138023327	G	T	CACNA1	CACNA1B:NM_000718:exon19:c.G2584T:p.A862	rs7873074	0.2824	0/0:28,0	0/1:17,36	0/0:35,0	0/0:26,0	0/1:29,16	0/1:20,26
chr9	138114424	C	T	CACNA1	CACNA1B:NM_000718:exon40:c.C5583T:p.F186	rs2229948	0.7159	0/1:24,8	0/0:37,0	0/0:33,0	0/1:22,9	0/1:30,10	0/1:31,18
chr9	128423190	G	A	CERCAM	CERCAM:NM_001286760:exon3:c.G119A:p.R40H	rs556586561	0.0003	0/1:76,16	0/0:36,0	0/1:87,11	0/0:47,0	0/1:39,34	0/1:64,26
chr9	128434147	G	A	CERCAM	CERCAM:NM_001286760:exon10:c.G1015A:p.V3	rs766133803	4.954e-05	0/0:118,0	0/1:189,12	0/0:136,0	0/0:140,0	0/1:183,10	0/1:174,139
chr9	128203516	C	A	DNM1	UNKNOWN	rs61757224	0.0335	0/1:103,88	0/0:52,0	0/1:143,8	0/0:38,0	0/1:109,78	0/1:107,70
chr9	128222476	C	T	DNM1	UNKNOWN	rs3003609	0.5325	0/0:45,0	0/1:56,66	0/0:53,0	0/0:56,0	0/1:38,50	0/1:76,74
chr9	137048977	G	A	ENTPD2	ENTPD2:NM_001246:exon8:c.C1179T:p.D393D,	rs7873815	0.3176	0/1:78,61	0/0:93,0	0/0:95,0	0/1:101,61	0/1:79,87	0/1:86,93
chr9	137051179	C	T	ENTPD2	.	rs78069653	0.0150	0/0:55,0	0/1:48,34	0/0:60,0	0/0:59,0	0/1:31,18	0/1:54,32
chr9	127803047	G	A	FPGS	FPGS:NM_001288803:exon1:c.G123A:p.P41P,FPC	rs11554716	0.0404	0/0:93,0	0/1:82,50	0/0:100,0	0/0:44,0	0/1:63,59	0/1:86,66
chr9	127813354	A	C	FPGS	FPGS:NM_001288803:exon14:c.A1436C:p.H479P,	.	.	0/1:55,25	0/0:98,32	0/1:89,31	0/1:82,21	0/1:85,32	0/1:98,33
chr9	127803047	G	A	FPGS	FPGS:NM_001288803:exon1:c.G123A:p.P41P,FPC	rs11554716	0.0404	0/0:93,0	0/1:82,50	0/0:100,0	0/0:44,0	0/1:63,59	0/1:86,66
chr9	127813358	A	C	FPGS	FPGS:NM_001288803:exon14:c.A1440C:p.P480P,I	.	.	0/1:57,13	0/0:44,0	0/1:85,21	0/1:81,10	0/1:85,20	0/1:93,23
chr9	133205977	GC	G	.	.	.	.	0/1:30,3	0/0:45,0	0/0:27,0	0/1:45,6	0/1:26,10	0/1:50,12
chr9	137109077	G	GC	MAN1B1	.	rs71803786	0.1356	0/0:67,0	0/1:37,28	0/0:68,0	0/0:81,0	0/1:33,27	0/1:40,57
chr9	133205977	GC	G	.	.	.	.	0/1:30,3	0/0:45,0	0/0:27,0	0/1:45,6	0/1:26,10	0/1:50,12
chr9	135549839	G	GC	OBP2A	OBP2A:NM_001293189:exon7:c.582delinsGC:p.L	rs112851559	0.2659	0/1:23,15	0/0:61,0	0/0:46,0	0/1:35,19	0/1:32,8	0/1:33,23
chr9	137109077	G	GC	MAN1B1	.	rs71803786	0.1356	0/0:67,0	0/1:37,28	0/0:68,0	0/0:81,0	0/1:33,27	0/1:40,57
chr9	135549839	G	GC	OBP2A	OBP2A:NM_001293189:exon7:c.582delinsGC:p.L	rs112851559	0.2659	0/1:23,15	0/0:61,0	0/0:46,0	0/1:35,19	0/1:32,8	0/1:33,23
chr9	131026432	C	A	LAMC3	LAMC3:NM_006059:exon2:c.C521A:p.P174H	rs2275137	0.2232	0/1:96,36	0/0:34,0	0/1:141,5	0/0:39,0	0/1:111,57	0/1:117,50
chr9	131068182	C	T	LAMC3	LAMC3:NM_006059:exon15:c.C2698T:p.R900C	rs757253079	1.665e-05	0/0:43,0	0/1:106,5	0/0:33,0	0/0:35,0	0/1:98,43	0/1:121,59
chr9	131026432	C	A	LAMC3	LAMC3:NM_006059:exon2:c.C521A:p.P174H	rs2275137	0.2232	0/1:96,36	0/0:34,0	0/1:141,5	0/0:39,0	0/1:111,57	0/1:117,50
chr9	131091853	T	C	LAMC3	.	rs818056	0.3152	0/0:19,0	0/1:14,41	0/0:20,0	0/0:21,0	0/1:16,36	0/1:6,40
chr9	131026433	C	G	LAMC3	LAMC3:NM_006059:exon2:c.C522G:p.P174P	rs2275136	0.1203	0/1:96,37	0/0:34,0	0/1:141,5	0/0:39,0	0/1:111,57	0/1:117,52
chr9	131068182	C	T	LAMC3	LAMC3:NM_006059:exon15:c.C2698T:p.R900C	rs757253079	1.665e-05	0/0:43,0	0/1:106,5	0/0:33,0	0/0:35,0	0/1:98,43	0/1:121,59
chr9	131026433	C	G	LAMC3	LAMC3:NM_006059:exon2:c.C522G:p.P174P	rs2275136	0.1203	0/1:96,37	0/0:34,0	0/1:141,5	0/0:39,0	0/1:111,57	0/1:117,52
chr9	131091853	T	C	LAMC3	.	rs818056	0.3152	0/0:19,0	0/1:14,41	0/0:20,0	0/0:21,0	0/1:16,36	0/1:6,40
chr9	131049014	G	A	LAMC3	.	rs192031352	0.0009	0/1:10,25	0/0:34,0	0/1:38,18	0/0:39,0	0/1:22,20	0/1:43,25

Chr	Position	Wt allele	Alt. allele	Gene	Other info	RefSNP	Frequency	Genotype I-1	Genotype I-2	Genotype II-1	Genotype II-3	GenotypeII-2	Genotype II-4
chr9	131068182	C	T	LAMC3	LAMC3:NM_006059:exon15:c.C2698T:p.R900C	rs757253079	1.665e-05	0/0:43,0	0/1:106,57	0/0:33,0	0/0:35,0	0/1:98,43	0/1:121,59
chr9	131049014	G	A	LAMC3	.	rs192031352	0.0009	0/1:10,25	0/0:34,0	0/1:38,18	0/0:39,0	0/1:22,20	0/1:43,25
chr9	131091853	T	C	LAMC3	.	rs818056	0.3152	0/0:19,0	0/1:14,41	0/0:20,0	0/0:21,0	0/1:16,36	0/1:6,40
chr9	131068182	C	T	LAMC3	LAMC3:NM_006059:exon15:c.C2698T:p.R900C	rs757253079	1.665e-05	0/0:43,0	0/1:106,57	0/0:33,0	0/0:35,0	0/1:98,43	0/1:121,59
chr9	131071462	C	CT	LAMC3	.	rs34910694	0.3710	0/1:116,15	0/0:40,0	0/1:90,14	0/0:42,0	0/1:89,8	0/1:108,12
chr9	131068182	C	T	LAMC3	LAMC3:NM_006059:exon15:c.C2698T:p.R900C	rs757253079	1.665e-05	0/0:43,0	0/1:106,57	0/0:33,0	0/0:35,0	0/1:98,43	0/1:121,59
chr9	131071522	G	A	LAMC3	LAMC3:NM_006059:exon18:c.G3108A:p.G1036C	rs10901344	0.3693	0/1:175,35	0/0:40,0	0/1:148,2	0/0:42,0	0/1:164,19	0/1:188,17
chr9	131068182	C	T	LAMC3	LAMC3:NM_006059:exon15:c.C2698T:p.R900C	rs757253079	1.665e-05	0/0:43,0	0/1:106,57	0/0:33,0	0/0:35,0	0/1:98,43	0/1:121,59
chr9	131071528	C	G	LAMC3	LAMC3:NM_006059:exon18:c.C3114G:p.L1038L	rs10901345	0.3694	0/1:179,37	0/0:40,0	0/1:149,2	0/0:42,0	0/1:164,21	0/1:189,19
chr9	131068182	C	T	LAMC3	LAMC3:NM_006059:exon15:c.C2698T:p.R900C	rs757253079	1.665e-05	0/0:43,0	0/1:106,57	0/0:33,0	0/0:35,0	0/1:98,43	0/1:121,59
chr9	131087466	C	T	LAMC3	.	rs4740411	0.2359	0/1:53,50	0/0:36,0	0/1:22,38	0/0:33,0	0/1:31,34	0/1:33,58
chr9	131068182	C	T	LAMC3	LAMC3:NM_006059:exon15:c.C2698T:p.R900C	rs757253079	1.665e-05	0/0:43,0	0/1:106,57	0/0:33,0	0/0:35,0	0/1:98,43	0/1:121,59
chr9	131087621	G	A	LAMC3	LAMC3:NM_006059:exon26:c.G4376A:p.R1459Q	rs4740412	0.2309	0/1:79,75	0/0:99,0	0/1:92,84	0/0:88,0	0/1:73,72	0/1:83,103
chr9	131068182	C	T	LAMC3	LAMC3:NM_006059:exon15:c.C2698T:p.R900C	rs757253079	1.665e-05	0/0:43,0	0/1:106,57	0/0:33,0	0/0:35,0	0/1:98,43	0/1:121,59
chr9	131087714	A	G	LAMC3	.	rs4740413	0.4080	0/1:66,70	0/0:79,0	0/1:58,58	0/0:36,0	0/1:63,55	0/1:76,80
chr9	131071462	C	CT	LAMC3	.	rs34910694	0.3710	0/1:116,15	0/0:40,0	0/1:90,14	0/0:42,0	0/1:89,8	0/1:108,12
chr9	131091853	T	C	LAMC3	.	rs818056	0.3152	0/0:19,0	0/1:14,41	0/0:20,0	0/0:21,0	0/1:16,36	0/1:6,40
chr9	131071522	G	A	LAMC3	LAMC3:NM_006059:exon18:c.G3108A:p.G1036C	rs10901344	0.3693	0/1:175,35	0/0:40,0	0/1:148,2	0/0:42,0	0/1:164,19	0/1:188,17
chr9	131091853	T	C	LAMC3	.	rs818056	0.3152	0/0:19,0	0/1:14,41	0/0:20,0	0/0:21,0	0/1:16,36	0/1:6,40
chr9	131071528	C	G	LAMC3	LAMC3:NM_006059:exon18:c.C3114G:p.L1038L	rs10901345	0.3694	0/1:179,37	0/0:40,0	0/1:149,2	0/0:42,0	0/1:164,21	0/1:189,19
chr9	131091853	T	C	LAMC3	.	rs818056	0.3152	0/0:19,0	0/1:14,41	0/0:20,0	0/0:21,0	0/1:16,36	0/1:6,40
chr9	131087466	C	T	LAMC3	.	rs4740411	0.2359	0/1:53,50	0/0:36,0	0/1:22,38	0/0:33,0	0/1:31,34	0/1:33,58
chr9	131091853	T	C	LAMC3	.	rs818056	0.3152	0/0:19,0	0/1:14,41	0/0:20,0	0/0:21,0	0/1:16,36	0/1:6,40
chr9	131087621	G	A	LAMC3	LAMC3:NM_006059:exon26:c.G4376A:p.R1459Q	rs4740412	0.2309	0/1:79,75	0/0:99,0	0/1:92,84	0/0:88,0	0/1:73,72	0/1:83,103
chr9	131091853	T	C	LAMC3	.	rs818056	0.3152	0/0:19,0	0/1:14,41	0/0:20,0	0/0:21,0	0/1:16,36	0/1:6,40
chr9	131087714	A	G	LAMC3	.	rs4740413	0.4080	0/1:66,70	0/0:79,0	0/1:58,58	0/0:36,0	0/1:63,55	0/1:76,80

Chr	Position	Wt allele	Alt. allele	Gene	Other info	RefSNP	Frequency	Genotype I-1	Genotype I-2	Genotype II-1	Genotype II-3	GenotypeII-2	Genotype II-4
chr9	131091853	T	C	LAMC3	.	rs818056	0.3152	0/0:19,0	0/1:14,41	0/0:20,0	0/0:21,0	0/1:16,36	0/1:6,40
chr9	137256024	C	T	NELFB	NELFB:NM_015456:exon2:c.C364T:p.L122L	rs139208158	0.0044	0/1:96,85	0/0:66,0	0/0:67,0	0/1:94,126	0/1:112,11	0/1:124,103
chr9	137272871	G	A	NELFB	NELFB:NM_015456:exon13:c.G1830A:p.A610A	rs148005191	0.0148	0/0:33,0	0/1:50,33	0/0:47,0	0/0:34,0	0/1:20,29	0/1:23,26
chr9	137467395	G	A	PNPLA7	PNPLA7:NM_152286:exon25:c.C2886T:p.I962I,PI	rs61738892	0.0497	0/0:44,0	0/1:94,68	0/0:36,0	0/0:47,0	0/1:88,59	0/1:142,61
chr9	137515425	C	A	PNPLA7	PNPLA7:NM_152286:exon11:c.G1104T:p.E368D	rs3750379	0.0397	0/0:48,0	0/0:54,0	0/0:51,0	0/0:46,0	0/1:88,30	0/1:136,33
chr9	137467395	G	A	PNPLA7	PNPLA7:NM_152286:exon25:c.C2886T:p.I962I,PI	rs61738892	0.0497	0/0:44,0	0/1:94,68	0/0:36,0	0/0:47,0	0/1:88,59	0/1:142,61
chr9	137515439	G	C	PNPLA7	PNPLA7:NM_152286:exon11:c.C1090G:p.Q364E	rs3750378	0.0424	0/0:48,0	0/0:54,0	0/0:51,0	0/0:46,0	0/1:86,30	0/1:137,33
chr9	137498164	G	A	PNPLA7	PNPLA7:NM_152286:exon16:c.C1764T:p.I588I,PI	rs61754920	0.1225	0/0:54,0	0/1:139,1	0/0:70,0	0/0:87,0	0/1:153,75	0/1:163,61
chr9	137515425	C	A	PNPLA7	PNPLA7:NM_152286:exon11:c.G1104T:p.E368D	rs3750379	0.0397	0/0:48,0	0/0:54,0	0/0:51,0	0/0:46,0	0/1:88,30	0/1:136,33
chr9	137498164	G	A	PNPLA7	PNPLA7:NM_152286:exon16:c.C1764T:p.I588I,PI	rs61754920	0.1225	0/0:54,0	0/1:139,1	0/0:70,0	0/0:87,0	0/1:153,75	0/1:163,61
chr9	137515439	G	C	PNPLA7	PNPLA7:NM_152286:exon11:c.C1090G:p.Q364E	rs3750378	0.0424	0/0:48,0	0/0:54,0	0/0:51,0	0/0:46,0	0/1:86,30	0/1:137,33
chr9	137515425	C	A	PNPLA7	PNPLA7:NM_152286:exon11:c.G1104T:p.E368D	rs3750379	0.0397	0/0:48,0	0/0:54,0	0/0:51,0	0/0:46,0	0/1:88,30	0/1:136,33
chr9	137515439	G	C	PNPLA7	PNPLA7:NM_152286:exon11:c.C1090G:p.Q364E	rs3750378	0.0424	0/0:48,0	0/0:54,0	0/0:51,0	0/0:46,0	0/1:86,30	0/1:137,33
chr9	129874641	C	T	USP20	USP20:NM_001008563:exon18:c.C1806T:p.H602H	rs35899714	0.0670	0/0:61,0	0/1:139,1	0/0:52,0	0/0:83,0	0/1:128,10	0/1:165,101
chr9	129875433	G	A	USP20	USP20:NM_001008563:exon20:c.G2172A:p.A724A	rs735115	0.3127	0/1:92,93	0/0:49,0	0/1:117,8	0/0:122,0	0/1:115,77	0/1:150,108
chr9	114403926	G	T	WHRN	WHRN:NM_001346890:exon6:c.C1335A:p.N445K	rs2274158	0.2476	0/1:72,45	0/0:36,0	0/1:75,63	0/0:36,0	0/1:69,50	0/1:60,72
chr9	114478722	C	T	WHRN	WHRN:NM_001173425:exon2:c.G668A:p.R223H,	rs146273185	0.0067	0/0:40,0	0/1:35,29	0/0:45,0	0/0:37,0	0/1:34,18	0/1:49,39
chr9	114403966	A	G	WHRN	WHRN:NM_001346890:exon6:c.T1295C:p.V432A	rs2274159	0.4786	0/1:144,86	0/0:36,0	0/1:123,1	0/0:36,0	0/1:132,10	0/1:106,138
chr9	114478722	C	T	WHRN	WHRN:NM_001173425:exon2:c.G668A:p.R223H,	rs146273185	0.0067	0/0:40,0	0/1:35,29	0/0:45,0	0/0:37,0	0/1:34,18	0/1:49,39
chr9	114424432	C	T	WHRN	WHRN:NM_001346890:exon2:c.G265A:p.A89T,W	rs4978584	0.2530	0/1:134,13	0/0:46,0	0/1:158,1	0/0:40,0	0/1:156,13	0/1:190,138
chr9	114478722	C	T	WHRN	WHRN:NM_001173425:exon2:c.G668A:p.R223H,	rs146273185	0.0067	0/0:40,0	0/1:35,29	0/0:45,0	0/0:37,0	0/1:34,18	0/1:49,39
chrX	1412803	C	T	ASMTL	ASMTL:NM_001173474:exon11:c.G1526A:p.R509	rs201860146	0.0010	0/0:98,0	0/1:42,49	0/0:52,0	0/0:93,0	0/1:31,36	0/1:54,43
chrX	1435014	G	A	ASMTL	.	rs28587213	0.1311	0/1:23,21	0/0:40,0	0/0:32,0	0/0:29,0	0/1:18,20	0/1:19,20
chrX	1421813	A	G	ASMTL	ASMTL:NM_001173474:exon8:c.T1042C:p.Y348H	rs35466875	0.1207	0/0:39,0	0/1:26,32	0/0:39,0	0/0:58,0	0/1:26,20	0/1:38,16
chrX	1435014	G	A	ASMTL	.	rs28587213	0.1311	0/1:23,21	0/0:40,0	0/0:32,0	0/0:29,0	0/1:18,20	0/1:19,20
chrX	1427949	A	G	ASMTL	ASMTL:NM_001173474:exon6:c.T634C:p.S212P,A	rs11553051	0.1266	0/0:196,0	0/1:131,14	0/0:201,0	0/0:257,0	0/1:151,79	0/1:184,142
chrX	1435014	G	A	ASMTL	.	rs28587213	0.1311	0/1:23,21	0/0:40,0	0/0:32,0	0/0:29,0	0/1:18,20	0/1:19,20
chrX	6533743	G	A	VCX3A	.	rs149365061	0.0146	0/0:10,0	0/0:304,0	0/0:70,0	0/0:78,0	0/1:4,48	1/1:4,101
chrX	6533823	C	T	VCX3A	VCX3A:NM_016379:exon3:c.G483A:p.E161E	rs143093189	0.2057	0/0:4,0	0/0:64,0	0/1:22,12	0/1:21,17	0/1:9,5	0/1:13,10
chrX	6533743	G	A	VCX3A	.	rs149365061	0.0146	0/0:10,0	0/0:304,0	0/0:70,0	0/0:78,0	0/1:4,48	1/1:4,101
chrX	6533828	C	T	VCX3A	VCX3A:NM_016379:exon3:c.G478A:p.V160M	rs35874450	0.1310	0/0:4,0	0/0:56,0	0/1:21,6	0/1:69,11	0/1:9,4	0/1:13,5
chrX	152914462	G	C	ZNF185	.	rs6526140	0.2657	./.:2,0:2:.:.	0/0:53,0	0/0:40,0	0/0:37,0	0/1:56,13	0/1:50,5
chrX	152914468	G	A	ZNF185	.	rs6526141	0.2736	0/0:2,0	0/0:53,0	0/0:40,0	0/0:37,0	0/1:56,13	0/1:51,6
chrX	152914462	G	C	ZNF185	.	rs6526140	0.2657	./.:2,0:2:.:.	0/0:53,0	0/0:40,0	0/0:37,0	0/1:56,13	0/1:50,5
chrX	152914515	G	A	ZNF185	ZNF185:NM_001178106:exon1:c.G26A:p.R9H,ZN	rs6526142	0.2975	0/1:1,3	0/0:53,0	0/0:31,0	0/0:37,0	0/1:34,17	0/1:35,8
chrX	152914468	G	A	ZNF185	.	rs6526141	0.2736	0/0:2,0	0/0:53,0	0/0:40,0	0/0:37,0	0/1:56,13	0/1:51,6
chrX	152914515	G	A	ZNF185	ZNF185:NM_001178106:exon1:c.G26A:p.R9H,ZN	rs6526142	0.2975	0/1:1,3	0/0:53,0	0/0:31,0	0/0:37,0	0/1:34,17	0/1:35,8

8. BIBLIOGRAPHY

1. Q Chen, P Shou, C Zheng , M Jiang , G Cao , Q Yang , J Cao, N Xie, T Velletri, X Zhang, C Xu, L Zhang, H Yang, J Hou, Y Wang and Y Shi. Fate decision of mesenchymal stem cells adipocytes or osteoblasts? Cell Death and Differentiation (2016) 23, 1128–1139 & 2016 Macmillan Publishers Limited All rights reserved 1350-9047/16.
2. Clevers H, Loh KM, Nusse R. Stem cell signaling. An integral program for tissue renewal and regeneration Wnt signaling and stem cell control. Science 2014; 346 1248012. 39.
3. Sherwood V. WNT signaling an emerging mediator of cancer cell metabolism? Mol Cell Biol 2015; 35 2–10. 40.
4. Holland JD, Klaus A, Garratt AN, Birchmeier W. Wnt signaling in stem and cancer stem cells. Curr Opin Cell Biol 2013; 25 254–264
5. Muruganandan S, Roman AA, Sinal CJ. Adipocyte differentiation of bone marrow-derived Mesenchymal stem cells cross talk with the osteoblastogenic program. Cell Mol Life Sci 2009; 66 236–253
6. Lv FJ, Tuan RS, Cheung KM, Leung VY. Concise review the surface markers and identity of human mesenchymal stem cells. Stem Cells 2014; 32 1408–1419.
7. Park HW, Kim YC, Yu B, Moroishi T, Mo JS, Plouffe SW et al. Alternative Wnt signaling activates YAP/TAZ. Cell 2015; 162 780–794.
8. Byun MR, Hwang JH, Kim AR, Kim KM, Hwang ES, Yaffe MB et al. Canonical Wnt signalling activates TAZ through PP1A during osteogenic differentiation. Cell Death Differ 2014; 21 854–863
9. Bennett CN, Ouyang H, Ma YL, Zeng Q, Gerin I, Sousa KM et al. Wnt10b increases postnatal bone formation by enhancing osteoblast differentiation. J Bone Miner Res 2007; 22 1924–1932.
10. Stevens JR, Miranda-Carboni GA, Singer MA, Brugger SM, Lyons KM, Lane TF. Wnt10b deficiency results in age-dependent loss of bone mass and progressive reduction of mesenchymal progenitor cells. J Bone Miner Res 2010; 25 2138–2147.
11. Arango NA, Szotek PP, Manganaro TF, Oliva E, Donahoe PK, Teixeira J. Conditional deletion of beta-catenin in the mesenchyme of the developing mouse uterus results in a switch to adipogenesis in the myometrium. Dev Biol 2005; 288 276–283.
12. Komori T. Regulation of osteoblast differentiation by transcription factors. J Cell Biochem 2006; 99 1233–1239. 89.
13. Nakashima K, Zhou X, Kunkel G, Zhang Z, Deng JM, Behringer RR et al. The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation. Cell 2002; 108 17–29. 90.
14. Komori T. Regulation of osteoblast differentiation by Runx2. Adv Exp Med Biol 2010; 658 43–49. 91.
15. Komori T. Regulation of bone development and extracellular matrix protein genes by RUNX2. Cell Tissue Res 2010; 339 189–195.
16. Maria Teresa Valenti, Luca Dalle Carbonare and Monica Mottes. Osteogenic Differentiation in Healthy and Pathological Conditions. Int. J. Mol. Sci. 2017, 18, 41; doi 10.3390/ijms18010041
17. Arantza Infante and Clara I. Rodríguez. Osteogenesis and aging lessons from mesenchymal stem cells. z Stem Cell Research & Therapy (2018) 9 244 doi.org/10.1186/s13287-018-0995-x
18. Xu, J.F.; Yang, G.H.; Pan, X.H.; Zhang, S.J.; Zhao, C.; Qiu, B.S.; Gu, H.F.; Hong, J.F.; Cao, L.; Chen, Y.; et al. Altered microrna expression profile in exosomes during osteogenic differentiation of human bone marrow-derived mesenchymal stem cells. PLoS ONE 2014, 9, e114627.
19. Bailey DA, McKay HA, Mirwald RL, Crocker PR, Faulkner RA. A six-year longitudinal study of the relationship of physical activity to bone mineral accrual in growing children the university of Saskatchewan bone mineral accrual study. J Bone Miner Res. 1999 Oct;14(10) 1672-9.
20. Maria Teresa Valenti , Luca Dalle Carbonare and Monica Mottes. Osteogenic Differentiation in Healthy and Pathological Conditions. Int. J. Mol. Sci. 2017, 18, 41; doi 10.3390/ijms18010041
21. Valenti, M.T.; Carbonare, L.D.; Mottes, M. Hypophosphatasia and Mesenchymal. Int. J. Stem. Cell Res. Ther. 2016, 3, 20.
22. Cao, X.; Chen, D. The BMP signaling and in vivo bone formation. Gene 2005, 357, 1–8. [CrossRef] [PubMed]
23. Lowenstein EJ. Osteogenesis imperfecta in a 3,000 year old mummy. Childs Nerv Syst 2009;25 515 6.
24. Kuurila K, Kaitila I, Johansson R, Grénman R. Hearing loss in Finnish adults with osteogenesis imperfecta A nationwide survey. Ann Otol Rhinol Laryngol 2002;111 939 46.
25. Sam JE, harmalingam M. Osteogenesis imperfecta. Indian J Endocr Metab 2017;21 903-8
26. Thakker RV, Whyte MP, Eisman J, Igarashi T, editors. Genetics of Bone Biology and Skeletal Disease. San Diego Academic Press; 2013.
27. Robichon J, Germain JP. Pathogenesis of osteogenesis imperfecta. Can Med Assoc J 1968;99 975-9.
28. Osteogenesis imperfecta and therapeutics. Roy Morello. © 2018 Elsevier B.V. All rights reserved.

29. Joan C. Marini, Antonella Forlino, Hans Peter Bächinger, Nick J. Bishop, Peter H. Byers, Anne De Paepe, Francois Fassier, Nadja Fratzl-Zelman, Kenneth M. Kozloff, Deborah Krakow, Kathleen Montpetit & Oliver Semler. Osteogenesis imperfecta. Nat Rev Dis Primers. 2017 Aug 18;3 17052. doi 10.1038/nrdp.2017.52.

30. Martin E1, Shapiro JR. Osteogenesis imperfecta epidemiology and pathophysiology. Curr Osteoporos Rep. 2007 Sep;5(3) 91-7.

31. Moffatt P, Gaumond MH, Salois P, Sellin K, Bessette MC, Godin E, et al. Bril a novel bonespecific modulator of mineralization. J Bone Miner Res. Sep; 2008 23(9) 1497–508. [PubMed 18442316]

32. Glorieux FH, Rauch F, Plotkin H, Ward L, Travers R, Roughley P, et al. Type V osteogenesis imperfecta a new form of brittle bone disease. J Bone Miner Res. Sep; 2000 15(9) 1650–8. [PubMed 10976985]

33. Cheung MS, Glorieux FH, Rauch F. Natural history of hyperplastic callus formation in osteogenesis imperfecta type V. J Bone Miner Res. Aug; 2007 22(8) 1181–6. [PubMed 17451374]

34. Antonella Forlino, Joan C Marini. Osteogenesis imperfecta. Lancet 2016; 387 1657–71

35. Dragan Primorac, David W. Rowe1 ,Monica Mottes2 , Ingeborg Barišiæ3 ,Darko Antièeviæ4 , Stefania Mirandola2 , Macarena Gomez Lira2 , Ivo Kalajziæ1 , Vesna Kušec3 , Francis H. Glorieux5. Osteogenesis Imperfecta at the Beginning of Bone and Joint Decade. Croat Med J 2001;42 393-415

36. Sekiya A, Okano-Kosugi H, Yamazaki CM, Koide T. Pigment epithelium-derived factor (PEDF) shares binding sites in collagen with heparin/heparan sulfate proteoglycans. J Biol Chem. Jul 29; 2011 286(30) 26364–74. [PubMed 21652703]

37. Rauch F, Husseini A, Roughley P, Glorieux FH, Moffatt P. Lack of circulating pigment epithelium-derived factor is a marker of osteogenesis imperfecta type VI. J Clin Endocrinol Metab. Aug; 2012 97(8) E1550–6. [PubMed 22669302]

38. Venturi G, Gandini A, Monti E, Dalle Carbonare L, Corradi M, Vincenzi M, et al. Lack of expression of SERPINF1, the gene coding for pigment epithelium-derived factor, causes progressively deforming osteogenesis imperfecta with normal type I collagen. J Bone Miner Res. Mar; 2012 27(3) 723–8. [PubMed 22113968]

39. Farber CR, Reich A, Barnes AM, Becerra P, Rauch F, Cabral WA, et al. A Novel IFITM5 Mutation in Severe Atypical Osteogenesis Imperfecta Type VI Impairs Osteoblast Production of Pigment Epithelium-Derived Factor. J Bone Miner Res. Jan 13.2014 [First to describe novel mutation (S40L) in IFITM5 and suggest a relationship between BRIL and PEDF.]

40. Becerra SP, Notario V. The effects of PEDF on cancer biology mechanisms of action and therapeutic potential. Nat Rev Cancer. 2013 Apr;13(4) 258–71. [PMC free article] [PubMed]

41. Sekiya A, Okano-Kosugi H, Yamazaki CM, Koide T. Pigment epithelium-derived factor (PEDF) shares binding sites in collagen with heparin/heparan sulfate proteoglycans. J Biol Chem 2011; 286 26364–74.

42. Sekiya A, Okano-Kosugi H, Yamazaki CM, Koide T. Pigment epithelium-derived factor (PEDF) shares binding sites in collagen with heparin/heparan sulfate proteoglycans. J Biol Chem. 2011 Jul 29;286(30) 26364–74. [PMC free article] [PubMed]

43. Becerra SP, Notario V. The effects of PEDF on cancer biology mechanisms of action and therapeutic potential. Nat Rev Cancer. Apr; 2013 13(4) 258–71. [PubMed 23486238]

44. Rauch F, Husseini A, Roughley P, Glorieux FH, Moffatt P. Lack of circulating pigment epithelium-derived factor is a marker of osteogenesis imperfecta type VI. J Clin Endocrinol Metab. 2012 Aug;97(8) E1550–6. [PubMed]

45. Li F, Song N, Tombran-Tink J, Niyibizi C. Pigment epithelium derived factor enhances diff erentiation and mineral deposition of human mesenchymal stem cells. Stem Cells 2013; 31 2714–23.

46. 85 Akiyama T, Dass CR, Shinoda Y, Kawano H, Tanaka S, Choong PF. PEDF regulates osteoclasts via osteoprotegerin and RANKL. Biochem Biophys Res Commun 2010; 391 789–94.

47. Joan C. Marini, Antonella Forlino, Hans Peter Bächinger, Nick J. Bishop, Peter H. Byers, Anne De Paepe, Francois Fassier, Nadja Fratzl-Zelman, Kenneth M. Kozloff, Deborah Krakow, Kathleen Montpetit and Oliver Semler. Osteogenesis imperfecta. NATURE REVIEWS DISEASE PRIMERS Article number 17052 doi 10.1038/nrdp.2017.52

48. Mottes M, Malerba G (2017). Hunting Novel Disease Genes in the Next Generation Sequencing Era Lessons from Osteogenesis Imperfecta. J of Genetics and genomes 2018 1 e102.

49. Ghazalpour A, Bennett BJ, Shih D, Che N, Orozco L, Pan C, et al. Genetic regulation of mouse liver metabolite levels. Mol Syst Biol. 2014;10 730.

50. Gieger C, Geistlinger L, Altmaier E, Hrabe de Angelis M, Kronenberg F, Meitinger T, et al. Genetics meets metabolomics a genome-wide association study of metabolite profiles in human serum. PLoS Genet. 2008;4 e1000282.

51. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. Mass Spectrom Rev. 2007;26 51–78

52. Dunn WB, Broadhurst DI, Atherton HJ, Goodacre R, Griffin JL. Systems level studies of mammalian metabolomes the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. Chem Soc Rev. 2011;40 387–426

53. Yehudit Hasin, Marcus Seldin1 and Aldons Lusis. Multi-omics approaches to disease. DOI 10.1186/s13059-017-1215-1

54. Maria Teresa Valenti, Luca Dalle Carbonare, Luca Donatelli, Francesco Bertoldo, Mirko Zanatta, Vincenzo Lo Cascio. Gene expression analysis in osteoblastic differentiation from peripheral blood mesenchymal stem cells. Bone 43 (2008) 1084–1092

55. KamilaSyslová, Lukáš Rambousekabc, Marek Kuzmad, Věra Najmanová, Věra Bubeníková-Valešováb, Romana Šlamberová, Petr Kačer. Monitoring of dopamine and its metabolites in brain microdialysates Method combining freeze-drying with liquid chromatography–tandem mass spectrometry. Journal of Chromatography A, Volume 1218, Issue 21, 27 May 2011, Pages 3382-3391

56. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009, 25, 1754–1760.

57. DePristo, M.A.; Banks, E.; Poplin, R.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; del Angel, G.; Rivas, M.A.; Hanna, M.; et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 2011, 43, 491–498.

58. Sim, NL.; Kumar, P.; Hu, J.; Henikoff, S.; Schneider, G.; Ng, P.C. SIFT web server predicting effects of amino acid substitutions on proteins. Nucleic Acid Res. 2012, 40, 452–457.

59. Sim, NL.; Kumar, P.; Hu, J.; Henikoff, S.; Schneider, G.; Ng, P.C. SIFT web server predicting effects of amino acid substitutions on proteins. Nucleic Acid Res. 2012, 40, 452–457. [Google Scholar] [CrossRef] [PubMed]

60. John M. Gasparcorresponding ,and Ronald P. Hart DMRfinder efficiently identifying differentially methylated regions from MethylC-seq data. BMC Bioinformatics. 2017; 18 528.

61. Petersen SV, Valnickova Z, Enghild JJ 2003 Pigment epithelium-derived factor (PEDF) occurs at a physiologically relevant concentration in human blood purification and characterization. Biochem J 374 199–206

62. Ref Glorieux F.H et al. 2002. Osteogenesis Imperfecta type VI a form of brittle bone disease with a mineralization defect. J BMR17; 30-38

63. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv 1303.3997v2.

64. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A et al. The Genome Analysis Toolkit a MapReduce framework for analysing next-generation DNA sequencing data. Genome Res 2010; 20 1297–1303

65. Claire Perrin-Tricaud, Christoph Rutschmann, Thierry Hennet. Identification of Domains and Amino Acids Essential to the Collagen Galactosyltransferase Activity of GLT25D1. PLoS ONE 6(12) e29390. doi 10.1371/journal.pone.0029390

66. Nancy H. Miller, Cristina M. Justice, Beth Marosy, Kandice Swindle, Yoonhee Kim, Marie-Hélène Roy-Gagnon, Heejong Sung, Dana Behneman, Kimberly F. Doheny, Elizabeth Pugh, and Alexander F. Wilson. Intra-familial tests of association between Familial Idiopathic Scoliosis and markers on 9q31.3-q34.3 and 16p12.3-q22.2. Hum Hered. 2012 ; 74(1) 36–44. doi 10.1159/000343751.

67. David I.K. Martin, Jennifer E. Cropley & Catherine M. Suter (2011) Epigenetics in disease Leader or follower?, Epigenetics, 6 7, 843-848, DOI 10.4161/epi.6.7.16498

68. Corradi Massimiliano, Monti Elena, Venturi Giacomo, Gandini Alberto, Mottes Monica, Antoniazzi Franco. The recurrent causal mutation for osteogenesis imperfecta type V occurs at a highly methylated CpG dinucleotide within the IFITM5 gene. Journal of Pediatric Genetics, vol. 3, no. 1, pp. 35-39, 2014 DOI 10.3233/PGE-14079

69. Peter A. Jones. Functions of DNA methylation islands, start sites, gene bodies and beyond. Nature Reviews Genetics volume 13, pages 484–492 (2012)

70. Hashimshony, T., Zhang, J., Keshet, I., Bustin, M. & Cedar, H. The role of DNA methylation in setting up chromatin structure during development. Nature Genet. 34, 187–192 (2003)

71. Ooi, S. K. & Bestor, T. H. The colorful history of active DNA demethylation. Cell 133, 1145–1148 (2008).

72. Wu, H. & Zhang, Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. Genes Dev. 25, 2436–2452 (2011).

73. Branco, M. R., Ficz, G. & Reik, W. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. Nature Rev. Genet. 13, 7–13 (2012).

74. Mark B. Meyer, Nancy A. Benkusky, Buer Sen, Janet Rubin, and J. Wesley Pike. Epigenetic Plasticity Drives Adipogenic and Osteogenic Differentiation of Marrow-Derived Mesenchymal Stem Cells. THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 291, NO. 34, pp. 17829 –17847, August 19, 2016©

75. María P. Menéndez-GutiérrezMercedes Ricote. The multi-faceted role of retinoid X receptor in bone remodelling. Cell. Mol. Life Sci. (2017) 74 2135. https //doi.org/10.1007/s00018-017-2458-4

76. Rogers, C. D., Phillips, J. L. & Bronner, M. E. Elk3 is essential for the progression from progenitor to definitive neural crest cell. Dev. Biol. 374, 255–263 (2013).

77. Ryan C. Ransom, Ava C. Carter, Ankit Salhotra, Tripp Leavitt, Owen Marecic, Matthew P. Murphy, Michael L. Lopez, Yuning Wei, Clement D. Marshall, Ethan Z. Shen, Ruth Ellen Jones, Amnon Sharir, Ophir D. Klein, Charles K. F. Chan, Derrick C. Wan, Howard Y. Chang, & Michael T. Longaker. Mechanoresponsive

stem cells acquire neural crest fate in jaw regeneration. 2018 Nature Research. https //doi.org/10.1038/s41586-018-0650-9

78. Do ¨rthe Andrea Kesper, Lydia Didt-Koziel, and Andrea Vortkamp. Gli2 Activator Function in Preosteoblasts IsSufficient to Mediate Ihh-Dependent OsteoblastDifferentiation, Whereas the RepressorFunction of Gli2 Is Dispensable forEndochondral Ossification. DEVELOPMENTAL DYNAMICS 239 1818–1826, 2010

79. Alyssa K. Kosmides, Kubra Kamisoglu, Steve E. Calvano, Siobhan A. Corbett, and Ioannis P. Androulakis. Metabolomic Fingerprinting Challenges and Opportunities. Crit Rev Biomed Eng. 2013; 41(3) 205–221

80. K. Irie, F. Hara-Irie, H. Ozawa, T. Yajima, Calcitonin gene-related peptide (CGRP)-containing nerve fibers in bone tissue and their involvement in bone remodeling, Microsc. Res. Tech. 58 (2002) 85e90.

81. E.L. Hill, R. Elde, Distribution of CGRP-, VIP-, D beta H-, SP-, and NPYimmunoreactive nerves in the periosteum of the rat, Cell Tissue Res. 264 (1991) 469e480.

82. Yadav VK, Ryu JH, Suda N, Tanaka KF, Gingrich JA, Schütz G et al. . Lrp5 controls bone formation by inhibiting serotonin synthesis in the duodenum. Cell 2008;135 825–837. [PMC free article] [PubMed]

83. Marie-Christine de Vernejoul, Corinne Collet, and Yasmine Chabbi-Achengli. Serotonin good or bad for bone. Bonekey Rep. 2012; 1 120.doi [10.1038/bonekey.2012.120]

84. Walther DJ, Peter JU, Bashammakh S, Hortnagl H, Voits M, Fink H et al. . Synthesis of serotonin by a second tryptophan hydroxylase isoform. Science 2003;299 76. [PubMed]

85. Wan Geun La, Jung Youn Shin, Suk Ho Bhang, Min Jin, Hee Hun Yoon, Seong Seo Noh, Gun Il Im, Chang Sung Kim, Byung Soo Kim. Culture on a 3,4-dihydroxy-L-phenylalanine-coated surface promotes the osteogenic differentiation of human mesenchymal stem cells. Tissue Engineering - Part A, 19(9-10), 1255-1263. https //doi.org/10.1089/ten.tea.2012.0165

86. ChantalChenu,MassimoMarenzana. Sympathetic nervous system and bone remodelling. Joint Bone Spine Volume 72, Issue 6, December 2005, Pages 481-483

87. 87. D.0. Sillence, Alison Senn, D.M. Danks. Genetic heterogeneity in osteogenesis imperfecta. Journal of Medical Genetics, 1979, 16, 101-116

88. Kovacs CS 2003. Foetal mineral homeostasis. In Gloriuex FH, Perrifor JM, Juppner H. Pediatrics Bone Biology and diseases. Acc

89. Bo Zhang, Yan Zhou, Nan Lin, Rebecca F. Lowdon, Chibo Hong, Raman P. Nagarajan, Jeffrey B. Cheng, Daofeng Li, Michael Stevens, Hyung Joo Lee, Xiaoyun Xing, Jia Zhou, Vasavi Sundaram, GiNell Elliott,1 Junchen Gu,1 Taoping Shi, Philippe Gascard, Mahvash Sigaroudinia, Thea D. Tlsty, Theresa Kadlecek, Arthur Weiss, Henriette O'Geen, Peggy J. Farnham, Ce´cile L. Maire, Keith L. Ligon, Pamela A.F. Madden, Angela Tam, Richard Moore, Martin Hirst, Marco A. Marra, Baoxue Zhang, Joseph F. Costello, and Ting Wang. Functional DNA methylation differences between tissues, cell types, and across individuals discovered using the M&M algorithm. Genome Research 23 1522–1540