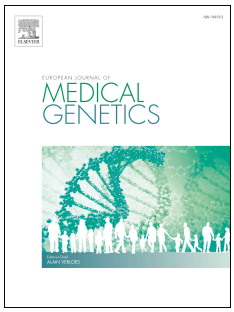
Journal Pre-proof



Biallelic variants in four genes underlying recessive osteogenesis imperfecta

Amir Hayat, Shabir Hussain, Muhammad Bilal, Mehran Kausar, Bader Almuzzaini, Safdar Abbas, Adeena Tanveer, Amjad Khan, Saima Siddiqi, Jia Nee Foo, Farooq Ahmad, Feroz Khan, Bushra Khan, Mariam Anees, Outi Mäkitie, Majid Alfadhel, Wasim Ahmad, Muhammad Umair

PII: S1769-7212(20)30064-1

DOI: https://doi.org/10.1016/j.ejmg.2020.103954

Reference: EJMG 103954

To appear in: *European Journal of Medical Genetics*

Received Date: 23 January 2020

Revised Date: 8 April 2020

Accepted Date: 9 May 2020

Please cite this article as: A. Hayat, S. Hussain, M. Bilal, M. Kausar, B. Almuzzaini, S. Abbas, A. Tanveer, A. Khan, S. Siddiqi, J.N. Foo, F. Ahmad, F. Khan, B. Khan, M. Anees, O. Mäkitie, M. Alfadhel, W. Ahmad, M. Umair, Biallelic variants in four genes underlying recessive osteogenesis imperfecta, *European Journal of Medical Genetics* (2020), doi: https://doi.org/10.1016/j.ejmg.2020.103954.

Author Statement or contributions

Hayat A, Hussain S, Bilal M, and Kausar M analyzed the patients, wrote the first draft of the manuscript and performed the molecular work. Almuzzaini1 B, Abbas S, and Tanveer A, performed protein modeling. Khan A, Foo NJ, Ahamd F, Khan F, analyzed the genetic data and clinical follow-up of the patients. Siddiqi S, Khan B, Anees M, Mäkitie O, Alfadhel M, Ahmad W, Umair M supervised the projects, students and edited the final manuscript.

Article Type: Original Article

Biallelic variants in four genes underlying recessive osteogenesis imperfecta

Amir Hayat[1*], Shabir Hussain[2*], Muhammad Bilal[2*], Mehran Kausar[2,3], Bader Almuzzaini[12], Safdar Abbas[2], Adeena Tanveer[4], Amjad Khan[5], Saima Siddiqi[6], Jia Nee Foo[7], Farooq Ahmad[8], Feroz Khan[9], Bushra Khan[1], Mariam Anees[2], Outi Mäkitie[10], Majid Alfadhel[11,12], Wasim Ahmad[2,#], Muhammad Umair[12,#]

[1]Department Biochemistry, Faculty of Life and Chemical Sciences, Abdul Wali khan University, Mardan, KPK, Pakistan.

[2]Department of Biochemistry, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan.

[3]Rehman College of Allied Health Sciences, RMI, Phase-5, Hayatabad, Peshawar, Pakistan.

[4]National Center for Bioinformatics, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan.

[5]Laboratoire d'ImmunoRhumatologie Moléculaire, Plateforme GENOMAX, INSERM UMR_S 1109, Faculté de Médecine, Fédération Hospitalo-Universitaire OMICARE, Fédération de Médecine Translationnelle de Strasbourg (FMTS), LabEx TRANSPLANTEX, Université de Strasbourg, 67085, Strasbourg, France. Service d'Immunologie Biologique, Plateau Technique de Biologie, Pôle de Biologie, Nouvel Hôpital Civil, Hôpitaux Universitaires de Strasbourg, 1 Place de l'Hôpital, 67091, Strasbourg, France.

[6]Institute of Biomedical & Genetic Engineering (IB&GE), Mauve area, G-9, Islamabad, Pakistan

[7]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore. Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore.

[8]Department of Chemistry, Women University Swabi, Swabi, Khyber Pakhtunkhwa (KPK), Pakistan

[9]Department of Zoology and Biology, Pir Mehr Ali Shah Arid Agriculture University, Rawalpindi, Pakistan.

[10]Folkhälsan Institute of Genetics and Research Program for Clinical and Molecular Metabolism, University of Helsinki, Helsinki, Finland.

[11]Division of Genetics, Department of Pediatrics, King Abdullah Specialized Children's Hospital, King Abdulaziz Medical City, Riyadh, Saudi Arabia

[12]Medical Genomics Research Department, King Abdullah International Medical Research Center (KAIMRC), King Saud bin Abdulaziz University for Health Sciences, Ministry of National Guard–Health Affairs P.O. Box 3660, Riyadh 11481 Saudi Arabia (KSA).

[*]These authors have contributed equally in the study presented. Therefore, they are considered as first author.

[#]**Corresponding Author**

Muhammad Umair, PhD, Medical Genomics Research Department, King Abdullah International Medical Research Center (KAIMRC), King Saud bin Abdulaziz University for Health Sciences, Ministry of National Guard–Health Affairs P.O. Box 3660, Riyadh 11481 Saudi Arabia (KSA).

Tel: +966-596020281, E-mail: umairmu@ngha.med.sa; khugoo4u@yahoo.com

Wasim Ahmad PhD, Department of Biochemistry, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad Pakistan

Tel:+92-51-90643003, Fax: +92-51-90643003, E-mail: wahmad@qau.edu.pk

Conflict of interest: Declared None

Abstract

Osteogenesis imperfecta (OI) is an inherited heterogeneous rare skeletal disorder characterized by increased bone fragility and low bone mass. The disorder mostly segregates in an autosomal dominant manner. However, several rare autosomal recessive and X-linked forms, caused by mutations in 18 different genes, have also been described in the literature.

Here, we present five consanguineous families segregating OI in an autosomal recessive pattern. Affected individuals in the five families presented severe forms of skeletal deformities. It included frequent bone fractures with abnormal healing, short stature, facial dysmorphism, osteopenia, joint laxity, and severe scoliosis. In order to search for the causative variants, DNA of at least one affected individual in three families (A-C) were subjected to whole exome sequencing (WES). In two other families (D-E), linkage analysis using highly polymorphic microsatellite markers was followed by Sanger sequencing. Sequence analysis revealed two novels and three previously reported disease-causing variants. The two novel homozygous variants including [c.824G>A; p.(Cys275Tyr)] in the *SP7* gene and [c.397C>T, p.(Gln133*)] in the *SERPINF1* gene were identified in families A and B, respectively. The three previously reported homozygous variants including [c.497G>A; p.(Arg166His)] in the *SPARC* gene, (c.359-3C>G; intron 2) and [c.677C>T; p.(Ser226Leu)] in the *WNT1* gene were identified in family C, D, and E.

In conclusion, our findings provided additional evidence of involvement of homozygous sequence variants in the *SP7*, *SERPINF1*, *SPARC* and *WNT1* genes causing severe OI. It also highlights the importance of extensive genetic investigations to search for the culprit gene in each case of skeletal deformity.

Keywords:

Osteogenesis Imperfecta, whole exome sequencing, Sanger sequencing, linkage analysis, *SP7*, *SERPINF1*, *SPARC*, *WNT1*, novel variants

1 Introduction

Osteogenesis imperfecta (OI) is genetically heterogeneous heritable skeletal dysplasia also known as “brittle bone disease” and prevails in about 1 in 10,000 to 20,000 live births (Monti et al., 2010). It is mostly characterized by progressive bone deformity, increased fracture susceptibility, low bone mass, and growth retardation. Osteogenesis imperfecta is mostly caused (85%) by dominant variants in the type I collagen genes (*COL1A1, COL1A2*) exhibiting mild to lethal phenotypic spectrum. Additional features associated with OI include bone deformities, short stature, mild osteopenia to severe osteoporosis, wormian bones, blue/grey sclera, dentinogenesis imperfecta (DI), increased vascular fragility and hearing loss. Recently, advances in rapid diagnostic technologies such as next generation sequencing have led to identification of multiple genes causing autosomal recessive and X-linked forms of OI, such as *CRTAP*, *P3H1/LEPRE1*, *PPIB*, *FKBP10*, *PLOD2*, *SERPINH1*, *BMP1*, *SERPINF1*, *IFITM5*, *SP7*, *TMEM38B*, *WNT1*, *SPARC*, *SEC24D*, *TENT5A*, *MESD*, *PLS3* and *MBTPS2* (Caparros-Martin et al., 2016; Diener et al., 2016; Umair et al., 2016, 2017; Moosa et al., 2019).

A functional metabolic classification has been **propsoed** for OI for genetic and clinical requirements. OI types were **classified** into five functional groups based on the **mechanism** that gene product having the same function and pathway might share the same disease mechanism (Forlino and Marini, 2016). These include group A, involving genes having defect in the collagen processing or structure (*COL1A1*, *COL1A2*, *BMP1*); group B, genes responsible for defect in collagen modification (*CRTAP*, *PPIB*, *TMEM38B*, and *LEPRE1*); group C, involved in cross-linking and collagen folding defects (*FKBP10*, *SERPINH1*, *PLOD2*, *MBTPS2*); group D, genes influencing bone mineralization (*SERPINF1*, *IFITM5*, *FAM46A*); and group E includes

gene involved in osteoblast development (*CREB3L1*, *WNT1*, *SP7*, *MESD*) (Forlino and Marini, 2016).

In the present study, we have investigated five families exhibiting severe features of OI segregating in an autosomal recessive pattern. Using linkage analysis and whole exome sequencing we have identified two novel and three previously reported mutations in the four genes *SP7*, *SERPINF1*, *WNT1* and *SPARC* causing severe recessive form of OI. Our findings indicate an overlap in the phenotypic presentation and underscore the need for extensive genetic **investigations** to identify the causative gene defect in similar cases of severe OI from Pakistani population.

2 Materials and Methods

2.1 Ethics Statement

Five families, exhibiting severe form of skeletal fragility, were recruited from different remote areas of Pakistan. The study was approved by the Institutional Review Board (IRB) of Quaid-i-Azam University, Islamabad and Abdul Wali Khan University Mardan, Pakistan. Written informed consent to conduct the the study and presenting the data including radiographs and photographs in research journals and conferences was obtained from the proband, patients and other family members in compliance with the Helsinki declaration. **Pedigrees (Fig. 1-2a, 3-4a, b) were constructed after careful interview with elders in the families.** Clinical and radiological assessments were performed at local government hospitals. Further, the authors reviewed the data and compared it with previously reported clinical spectrum of the patients with similar disorder. Genomic DNA was extracted from peripheral-blood lymphocytes and quantified using standard methods.

2.2 Whole exome sequencing (WES)

DNA of at least one affected member in three families (A-C) were directly subjected to WES on different platforms including Illumina HiSeq2500 (Illumina, Inc., San Diego, CA, USA) and Agilent 2100 Bioanalyser/Illumina HiSeq platform (BGI, Hong Kong). Exome enrichment was performed using standared methods and the obtained reads were aligned (BWA-MEM). Further, duplicates were removed (Picard), inDel were realigned and base quality was recalled. For variant calling different tools including PINDEL, SAM tools and Exome Depth were used. SAM tools varFilter script was used for variant filtering. Single nucleotide variants (SNVs) and InDels were detected using GATK HaplotypeCaller and annotated using the SnpEff tool. All the obtained reads were aligned against the human assembly hg19 (GRCh37) using Burrows-Wheeler Aligner. The WES mechanisim was followed as described previously (Umair, Shah et al., 2017; Umair, Eckstein et al., 2018). Considering the family pedigrees clearly depicted recessive pattern of inheritance of the disorder, on priority basis the filtering process screened the data for previously thirteen known causative genes (*SERPINF1, CRTAP*, *P3H1*, *PPIB, SERPINH1, FKBP10, SP7*, *BMP1*, *TMEM38B*, *WNT1*, *CREB3L1*, *MBTPS2*, *SPARC*) (Table 1).

2.3 Linkage analysis

In two families (D-E), all the available affected and unaffected members were subjected to genotyping using highly polymorphic microsatellite markers mapped in the region encompassing previously reported genes involved in causing autosomal recessive form of OI. This included *SERPINF1, CRTAP*, *P3H1, PPIB, SERPINH1, FKBP10, SP7, BMP1, TMEM38B, WNT1, CREB3L1, MBTPS2* and *SPARC*. The procedure followed for genotyping markers was as describe earlier (Umair, Rafique et al., 2017). Both the families showed convincing linkage to the markers (D12S1701, D12S1661, D12S2196, D12S1290, D12S339, D12S1627, D12S1620) linked to the *WNT1* gene on chromosome 12q13.12.

2.4 Segregation of the identified variants

Primer sequences for PCR amplification of the *WNT1* exons and identified variants in the *SP7, SERPINF1* and *SPARC* genes were designed using online Primer-3 software (http://bioinfo.ut.ee/primer3-0.4.0/). Sanger sequencing was performed using standard methods as described previously (Umair, Hassan et al., 2016). Sequencig data was analyzed via BIOEDIT sequence alignment editor version 6.0.7.

2.5 *In silico* analysis

Pathogenicity index for the detected variants was calculated using MutationTaster (http:// www.mutationtaster.org/), SIFT (http://sift.bii.a-star.edu.sg/), VarSome (https://varsome.com/), PolyPhen-2 (http://genetics. bwh.harvard.edu/pph2/), splic site variant tools (NNSplice; [https://omictools.com/nnsplice-tool], MutPred Splice [http://www.mutdb.org/mutpredsplice], SKIPPY [https://research.nhgri.nih.gov/skippy] and Human Splice Finder (v2.4.1) [PMC2685110]. Frequency of the variant in the general population was determined using ExAC (http://exac.broadinstitute.org/), genomAD (http://gnomad.broadinstitute.org/), 1000 Genomes and 135 Pakistani exomes (Inhouse).

2.6 Sequence retrieval and 3D structure prediction

UniProt database was used to retrive the SP7 (431 aa) protein sequence in the FASTA format with accession number Q8TDD2-1 (Venter et al., 2001). Comparative modelling was imployed in the absence of expremently known structures. Thus, I-TASSER was used to predict the SP7 protein structure. The I-TASSER model was evaluated using the provided scores (Yang et al., 2015). Using the AMBERff14 S B force field, the selected structure was optimized through 1000 steps of steepest- decent and 1000 steps of conjugate-gradient minimization by UCSF Chimera

version 1.11 (Meng et al., 2006). Finally, the Ramachandran outliers and poor rotamers were corrected through WinCoot and their optimized and reliable structure was constructed (Emsley et al., 2010). 3D structure of mutated protein was generated by MODELLER 9.19 (Webb and Sali, 2014). The MODELLER assists in 3D structure prediction of proteins by satisfaction of spatial restraints (Eswar et al., 2008) and the model was selected based on MODELLER evaluation score.

2.7 Model Evaluation

In structural bioinformatics the theoretical and expremental models of protein structures is a major concern (Melo et al., 1997). Thus, different evaluation tools were used for the assessment of protein structure such as RAMPAGE (Lovell et al., 2003), ERRAT (Colovos et al., 1993) and ProSA (Protein Structure Analysis). RAMPAGE generates Ramachandran plot for the assessment of models along with distribution of residues in favoured, allowed and outlier regions. ERRAT generated a plot indicating the confidence and overall quality of model and ProSA was used to calculate an overall quality score of the predicted structure (Wiederstein et al., 2007).

3. Results

3.1 Clinical Assessment

Family A: The proband was 32 years old male (IV-2). Parents and siblings of the affected member were healthy (Fig. 1a-g). The affected individual (IV-2) suffered fractures 5-15 days after birth. Since then, numerous fractures occurred. His height and weight were 95 cm and 30 kg. Physical examination of the affected member revealed bilateral bowing of femurs, tibiae and fibulae. Due to occurrance of numerous fractures and self-healing, the humeri, radii and ulnae

had zig-zag appearance. He did not have blue sclera, hearing impairment, dentinogenesis imperfecta, intellectual disability or skin manifestations (Fig. 1b-g, Table 2). Serum levels of alkaline phosphatase, phosphate, calcium and parathyroid hormone were within age-appropriate reference ranges.

Radiological assessment of the proband (IV-2) demonstrated multiple fractures and deformities in clavicle, acromion, glenoid cavity, scapula, sternum and ribs resulting in compression of intercostal ribs spaces, which led to decreased total lung capacity (Table 2). The coxal bone (hip bone) and the femur were severely deformed. Other bones including llium, pubic, ischium, sacrum and acetabulum were fused and unidentifiable. Femur had severe complex bending/fractures along with femoral head and neck. Lower limbs had fixed flexion deformity (FFD), having bilateral flexion contractures of knee as a result the affected individual was unable to ambulate freely. Severe kyphoscoliosis associated with platyspondyly, osteopenia and fractures in humerus, radii and ulna were also observed (Fig. 1b-g).

Family B: All three affected individuals (IV-1, IV-2, IV-3) in family B showed classical osteogenesis imperfecta phenotypes. At the time of the medical examination, the patients IV-1, IV-2 and IV-3 were 10 years, 12 years, and 14 years old, respectively. Patients were born at full term without any complications. Affected individuals showed fragile and brooked bones, and multiple fractures in arms, hands, legs, feet and spine (Table 2). Legs were bowed and the long bones were fragile and thin. The patients also had white sclerae, kyphosis, and sarcopenia. The long bones in the extremities were the most common sites of fractures (Fig. 2b-d).

Family C: This family had two affected members (IV-2, IV-4) (Fig. 3a). The affected member IV-4 died at the age of 22 months. They suffered from moderate to severe upper and lower limbs

fractures (Fig. 3b, c). Afected individual IV-2 was four years old girl, born to healthy first cousin parents. Pregnancy was normal however she was under-weight with noticeable poor mineralization of the skeleton (Table 2).. Multiple recurrent fractures of long bones especially in the left femur occurred during the first month of life. She had extremely fragile skeleton with moderate to severe osteoporotic bones (Fig. 3d, e). Left femur was more prone to fracture in contrast to other long bones. She was short and underweight having progressive kyphoscoliosis. Hearing loss, dentinogenesis imperfecta and blue sclerea were not observed (Fig. 3b). Serum biochemistry showed normal calcium, phosphate, alkaline phosphatase and vitamin D levels. Radiographic examination indicated severe osteopenia and multiple fractures (Fig. 3c).

Family D: In the family there were two affected individuals IV-2 and IV-3 aged 26 and 30 years, respectively. They suffered from severe long bone fractures at an early age and were confined to wheel chairs. Both had multiple fractures of the femur, tibia, fibula, humerus, radius and ulna, bowed extremities, short stature, joint laxity and severe scoliosis (Fig. 4c-f). Hearing impairment, facial dysmorphism and intellectual disability were not observed (Table 2)..

Family E: The proband (IV-1) in the family was 5 years old at time of the study (Fig. 4g). He suffered from bone fragility and frequent fractures at an early age (Table 2). He was confined to bed. Radiographs of the affected individual revealed deformed irregular humerus, radius, and ulna, anterior bowing of bones and decreased bone mineral density. In addition, he had improperly healed fractures in both lower limbs resulting in to severe lower limb bowing (Fig. 4h-i).

3.2 Sequencing candidate genes

As described above, search for the disease causing variants was carried out using WES in three families (A-C) and by genotyping microsatellite markers followed by Sanger sequencing in family D and E. **In family A-C step-by-step filtering and validation by Sanger sequencing of different homozygous and compound heterozygous variants revealed two novel and one previously reported variant.** The novel homozygous variants included a missense [c.824G>A, p.(Cys275Tyr)] in exon 2 of the *SP7* gene (NM_152860.1; NG_023391.1; **LOVD:00295528**) in family A and a nonsense [c.397C>T, p.(Gln133*)] in exon 4 of the *SERPINF1* gene **(NM_002615.7; NP_002606.3; LOVD:00295529)** in family B. A previously reported homozygous missense variant [c.497G>A, p.(Arg166His)] in exon 7 of the *SPARC* gene (NM_152860.1; NG_042174.1; **LOVD:00295530**) was identified in family C. **Filteration steps for WES data are presented in Supplementary Table 1.**

In family D and E analysis of the haplotypes established linkage to the *WNT1* gene (NM_005430.3; NG_033141.1) located on human chromosome 12q13.12. Subsequently, all exons and 5ÚTR and 3ÚTR of the *WNT1* gene were Sanger sequenced in both the families. Sequence analysis revealed two previously reported homozygous variants including a splice acceptor site (c.359-3C>G; intron 2; **LOVD:00295531**) in family D (Fig. 4l-m) and a missense [c.677C>T, p.(Ser226Leu)] in family E **[LOVD:00295532]** (Fig. 4o-q). All the four identified homozygous sequence variants were present in the heterozygous state in the obligate carriers in the families. The identified variants were not present in 135 Pakistani exomes and 230 ethnically matched control individuals.

The pathogenicity index of the detected variants was calculated using different online analysis tools including MutationTaster, SIFT, VarSome, PolyPhen-2, splic site variant tools, MutPred Splice, SKIPPY, Human Splice Finder and were predicted as disease causing. The variants were

also not reported in online public databases (ExAC, genomAD, 1000 Genomes) in homozygous form.

3.3 3D-modeling of SP7

Homology/3D modeling of a mutated SP7 protein [p.(Cys275Tyr)] revealed substantial structural changes. It is highly likely that such changes altered secondary structure and also affect binding ability of the SP7 protein. A heavy tyrosine (Molar mass: 181.19 gm/mol) replaced cysteine residue (Molar mass: 121.16 gm/mol) in a beta sheet which distorted protein conformation in the local structure (Fig. 1n-q). Relative thermal stability indicated that the mutant protein was less stable (stability=–11.05 kcal/mol) as compared to the wild-type (stability=–13.75 kcal/mol).

4. Discussion

The study, presented here, describes clinical and genetic characterization of five consanguineous families segregating osteogenesis imperfecta (OI) in autosoml recessive manner. Previously reported OI-related clinical features including frequent fractures and poor healing, osteopenia, and kyphoscoliosis with platyspondyly (Umair et al., 2016, 2017; Mendoza-Londono et al., 2015; Kausar et al., 2018) were found in affected individuals of all five families. Severity of the features however varied among the families. Affected members in four families (A, B, D, E) had severe forms of various skeletal deformities. An affected member in family A showed decreased bone density, bowing of the tibia and fibula, fracture of tibia, improper healing of the bones and multiple craniofacial features. Similar features were reported previously by Lapunzina et al. (2010). In both the cases mutations in the same *SP7* gene caused such severe form of bone deformities. Features such as white sclera, multiple fracures, kyphosis and muscle sarcopenia recorded in affected members in family B were reported previously in families of French, UAE,

and Turkish origin (Homan et al., 2011; Becker et al., 2011). Hearing impairment reported previously in several studies (Fiscaletti et sal., 2018; Lindahl et al., 2018; Carré et al., 2019) was not observed in the affected individuals in our families.

Using whole exome sequencing in three families and genotyping followed by Sanger sequencing in two other families revealed two novel and two previously reported disease causing variants in four culprit genes. The novel homozygous variants including a missense p.(Cys275Tyr) and a nonsense p.(Gln133*) were identified in the *SP7* and *SERPINF1*, respectively. The other two previously known homozygous missense variants p.(Arg166His) and p.(Ser226Leu) were detected in the *SPARC* and *WNT1*, respectively.

The p.(Cys275Tyr) is the third homozygous variant identified in the *SP7* gene located on human chromosome 12q13.13. The other two homozygous variants were reported in families of Egyptian and Australian origin (Lapunzina et al., 2010; Fiscaletti et al., 2018). In our study, structural modeling of the mutated SP7 revealed the variant p.(Cys275Tyr) affect secondary structure and ability of the SP7 to interact with other proteins. All these three variants caused osteogenesis imperfecta type XII. The SP7 (specific protein), also called as Osterix (OSX), belongs to a subgroup of the Kruppel-like family of transcription factor (NP_690599.1) containing zinc finger in its DNA-binding domain. It plays a major role, along with Runx2 and *Dlx5*, in driving the differentiation of mesenchymal precursor cells into osteoblasts and eventually osteocytes (Sinha and Zhou, 2013; Fukuda et al., 2018). During development, a mouse embryo model with Sp7 expression knocked out had no formation of bone tissue (Nakashima et al., 2002). In human, Timpson et al. (2009) have shown association of *SP7* locus with bone mass density.

Prior to the present study, several disease causing variants have been reported in the *SERPINF1* gene, located on human chromosome 17p13.3. The variant p.(Gln133*), identified in our family, is one of very few nonsense variants reported in this gene. All the variants in this gene cause OI type XVI (Caparrós-Martin et al., 2013; Homan et al., 2011; Wang et al., 2017). The gene encodes 50kDa Serpin Family F Member 1 (SERPINF1) protein, which is also called as pigment epithelium-derived factor (PEDF). This protein is shown to inhibit the vascular endothelial growth factor (VEGF) downstream actions, which is expressed during endochondral bone by the chondrocytes (Quan et al., 2005). It also binds to collagen type I neighboring the a1b1 and a1b2 integrin binding site, signifying its important role in integrin–collagen interactions, that has been reported in angiogenesis and cell adhesion (Meyer et al., 2002).

In three of our families we have found previously known missense variants in two causative genes. One of the variant p.(Arg166His) was identified in the *SPARC* gene (NM_003118.4), located on chromosome 5q33.1, which is responsible for causing OI type XVII (MIM 616507). To date, only two variants have been reported in the *SPARC* gene (Mendoza-Londono et al., 2015). The variant p.(Arg166His) was reported previously in a family of North African and Indian origin (Mendoza-Londono et al., 2015). The *SPARC* gene encodes a 40 kDa Secreted Protein Acidic and Cysteine Rich (SPARC) which is also called as osteonectin (ON). This protein has been implicated in several biological functions including mineralization of bone and cartilage, and modulation of cell proliferation (Guweidhi et al., 2005). The second known variant p.(Ser226Leu) was detected in the *WNT1* gene (NM_005430.3). To date, only 22 variants have been reported in the *WNT1* gene and all caused OI type XV (Umair et al., 2017). The *WNT1* gene encodes a secreted signaling protein which has been implicated in several developmental

processes, including regulation of cell fate and patterning during embryogenesis (Laine et al., 2013).

In conclusion, we have identified five disease causing variants in four different genes causing various types of osteogenesis imperfecta. Severity of the phenotypes varied from mild to severe form in the five families presented here. **Previously, only five OI families have been reported from Pakistan (Umair et al., 2016, 17; Kausar et al., 2018). The variant identified in the *SP7*, *SERPINF1* and *SPARC* genes are for the first time reported from Pakistani population (Umair et al., 2019).** Considering high rate of consanguinity, the variants identified here will support genetic conseling of Paksitani families segregating various types of skeletal deformities. In addition, the variants will further expand spectrum of mutations in the genes causing osteogenesis imperfecta.

Acknowledgments

We highly appreciate members of all five families for participating in the study. Amir Hayat and Shabir Hussain were supported by HEC sponsored IRSIP. M Bilal was supported by Indigenous PhD fellowship. This work was sponsored by King Abdullah International Medical Research Centre (KAIMRC), Riyadh Saudi Arabia.

Conflict of interest: Declared None

Funding

This work was funded by King Abdullah International Medical Research Centre (KAIMRC), project number: RC19/352/R.

References

Becker, J., Semler, O., Gilissen, C., Li, Y., Bolz, HJ., Giunta, C., Bergmann, C., Rohrbach, M., Koerber, F., Zimmermann, K., de Vries, P., Wirth, B., Schoenau, E., Wollnik, B., Veltman, J.A., Hoischen, A., Netzer, C., 2011. Exome sequencing identifies truncating mutations in human SERPINF1 in autosomal-recessive osteogenesis imperfecta. Am. J. Hum. Genet. 88(3), 362-71. doi: 10.1016/j.ajhg.2011.01.015.

Caparros-Martin, J.A., Aglan, M.S., Temtamy, S., Otaify, G.A., Valencia, M., Nevado, J., Vallespin, E., Del Pozo, A., Prior de Castro, C., Calatrava-Ferreras, L., Gutierrez, P., Bueno, A.M., Sagastizabal, B., Guillen-Navarro, E., Ballesta-Martinez, M., Gonzalez, V., Basaran, S.Y., Buyukoglan, R., Sarikepe, B., Espinoza-Valdez, C., Cammarata-Scalisi, F., Martinez-Glez, V., Heath, K.E., Lapunzina, P., Ruiz-Perez, V.L., 2016. Molecular spectrum and differential diagnosis in patients referred with sporadic or autosomal recessive osteogenesis imperfecta. Mol. Genet. Genomic. Med. 5(1),28-39. doi: 10.1002/mgg3.257.

Caparrós-Martin, J.A., Valencia, M., Pulido, V., Martínez-Glez, V., Rueda-Arenas, I., Amr, K., Farra, C., Lapunzina, P., Ruiz-Perez, V.L., Temtamy, S., Aglan, M., 2013. Clinical and molecular analysis in families with autosomal recessive osteogenesis imperfecta identifies mutations in five genes and suggests genotype-phenotype correlations. Am. J. Med. Genet A. 161A(6),1354-69. doi: 10.1002/ajmg.a.35938.

Carré, F., Achard, S., Rouillon, I., Parodi, M., Loundon, N., 2019. Hearing impairment and osteogenesis imperfecta: Literature review. Eur. Ann. Otorhinolaryngol. Head. Neck. Dis. 136(5),379-383. doi: 10.1016/j.anorl.2019.05.004.

Colovos, C., Yeates, T.O., 1993. Verification of protein structures: patterns of nonbonded atomic interactions. Protein. Sci. 2, 1511-1519. doi: 10.1002/pro.5560020916.

Diener, Sm, Bayer, S., Sabrautzki, S., Wieland, T., Mentrup, B., Przemeck, G.K., Rathkolb, B., Graf, E., Hans, W., Fuchs, H., Horsch, M., Schwarzmayr, T., Wolf, E., Klopocki, E., Jakob, F., Strom, T.M., Hrabě de Angelis, M., Lorenz-Depiereux, B., 2016. Exome sequencing identifies a nonsense mutation in Fam46a associated with bone abnormalities in a new mouse model for skeletal dysplasia. Mamm. Genome. 27(3-4),111-21.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta. Crystallogr. D. Biol. Crystallogr. 66(Pt 4),486-501.

Eswar, N., Eramian, D., Webb, B., Shen, M.Y., Sali, A., 2008. Protein structuremodeling with MODELLER. Methods. Mol. Biol. 426, 145-159.

Fiscaletti, M., Biggin, A., Bennetts, B., Wong, K., Briody, J., Pacey, V., Birman, C., Munns, C.F., 2018. Novel variant in Sp7/Osx associated with recessive osteogenesis imperfecta with bone fragility and hearing impairment. Bone. 110, 66-75.

Forlino, A., Marini, J.C., 2016. Osteogenesis imperfecta. Lancet. 387(10028), 1657-71. doi: 10.1016/S0140-6736(15)00728-X.

Fukuda, M., Yoshizawa, T., Karim, M.F., Sobuz, S.U., Korogi, W., Kobayasi, D., Okanishi, H., Tasaki, M., Ono, K., Sawa, T., Sato, Y., Chirifu, M., Masuda, T., Nakamura, T., Tanoue, H., Nakashima, K., Kobashigawa, Y., Morioka, H., Bober, E., Ohtsuki, S., Yamagata, Y., Ando, Y., Oike, Y., Araki, N., Takeda, S., Mizuta, H., Yamagata, K., 2018. SIRT7 has a critical role in bone formation by regulating lysine acylation of SP7/Osterix. Nat. Commun. 9(1), 2833. doi: 10.1038/s41467-018-05187-4.

Guweidhi, A., Kleeff, J., Adwan, H., Giese, N.A., Wente, M.N., Giese, T., Büchler, M.W., Berger, M.R., Friess, H., 2005. Osteonectin influences growth and invasion of pancreatic cancer cells. Ann. Surg. 242(2), 224-34. doi: 10.1097/01.sla.0000171866.45848.68.

Homan, E.P., Rauch, F., Grafe, I., Lietman, C., Doll, J.A., Dawson, B., Bertin, T., Napieral,a D., Morello, R., Gibbs, R., White, L., Miki, R., Cohn, D.H., Crawford, S., Travers, R., Glorieux, F.H., Lee, B., 2017. Mutations in SERPINF1 cause osteogenesis imperfecta type VI. J. Clin. Invest. 127(7), 2678-2688. doi: 10.1002/jbmr.487.

Kausar, M., Siddiqi, S., Yaqoob, M., Mansoor, S., Makitie, O., Mir, A., Khor, C.C., Foo, J.N., Anees, M., 2018. Novel mutation G324C in WNT1 mapped in a large Pakistani family with severe recessively inherited Osteogenesis Imperfecta. J. Biomed. Sci. 25(1),82. doi: 10.1186/s12929-019-0525-x.

Laine, C.M., Joeng, K.S., Campeau, P.M., Kiviranta, R., Tarkkonen, K., Grover, M., Lu, J.T., Pekkinen, M., Wessman, M., Heino, T.J., Nieminen-Pihala, V., Aronen, M., Laine, T., Kröger, H., Cole, W.G., Lehesjoki, A.E., Nevarez, L,, Krakow, D., Curry, C.J., Cohn, D.H,. Gibbs, R.A., Lee., B.H., Mäkitie, O., 2013. WNT1 mutations in early-onset osteoporosis and osteogenesis imperfecta. N. Engl. J. Med. 368(19):1809-16. doi: 10.1056/NEJMoa1215458.

Lapunzina, P., Aglan, M., Temtamy, S., Caparrós-Martín, J.A., Valencia, M., Letón, R., Martínez-Glez, V., Elhossini, R., Amr, K., Vilaboa, N., Ruiz-Perez, V.L., 2010. Identification of a frameshift mutation in Osterix in a patient with recessive osteogenesis imperfecta. Am. J. Hum. Genet. 87(1), 110-4. doi: 10.1016/j.ajhg.2010.05.016.

Lindahl, K., Aström, E., Dragomir, A., Symoens, S., Coucke, P., Larsson, S., Paschalis, E., Roschger, P., Gamsjaeger, S., Klaushofer, K., Fratzl-Zelman, N., Kindmark, A., 2018. Homozygosity for CREB3L1 premature stop codon in first case of recessive osteogenesis imperfecta associated with OASIS-deficiency to survive infancy. Bone. 114, 268-277. doi: 10.1016/j.bone.2018.06.019.

Lovell, S.C., Davis, I.W., Arendall, W.B. 3rd., de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., 2003. Richardson DC. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins. 50(3), 437-50. doi: 10.1002/prot.10286.

Melo, F., Devos, D., Depiereux, E., Feytmans, E., 1997. ANOLEA: a www server to assess protein structures. Proc. Int. Conf. Intell. Syst. Mol. Biol. 5, 187-90.

Mendoza-Londono, R., Fahiminiya, S., Majewski, J., Care4Rare, Canada. Consortium., Tétreault, M., Nadaf, J., Kannu, P., Sochett, E., Howard, A., Stimec, J., Dupuis, L., Roschger, P., Klaushofer, K., Palomo, T., Ouellet, J., Al-Jallad, H., Mort, J.S., Moffatt,

P., Boudko, S., Bächinger, H.P., Rauch, F., 2015. Recessive osteogenesis imperfecta caused by missense mutations in SPARC. Am. J. Hum. Genet. 96(6), 979-85. doi: 10.1016/j.ajhg.2015.04.021.

Meng, E.C., Pettersen, E.F., Couch, G.S., Huang, C.C., Ferrin, T.E., 2006. Tools for integrated sequence-structure analysis with UCSF Chimera. BMC. Bioinf. 7; 339. doi: 10.1186/1471-2105-7-339.

Meyer, C., Notari, L., Becerra, S.P., 2002. Mapping the type I collagenbinding site on pigment epithelium-derived factor. Implications for its antiangiogenic activity. J. Biol. Chem. 277(47): 45400–7. doi: 10.1074/jbc.M208339200.

Monti, E., Mottes, M., Fraschini, P., Brunelli, P., Forlino, A., Venturi, G., Doro, F., Perlini, S., Cavarzere, P., Antoniazzi F., 2010. Current and emerging treatments for the management of osteogenesis imperfecta. Ther. Clin. Risk. Manag. 6, 367–81. doi: 10.2147/tcrm.s5932.

Moosa, S., Yamamoto, G.L., Garbes, L., Keupp, K., Beleza-Meireles, A., Moreno, C.A., Valadares, E.R., de Sousa, S.B., Maia, S., Saraiva, J., Honjo, R.S., Kim, C.A., Cabral de Menezes, H., Lausch, E., Lorini, P.V., Lamounier ,A. Jr., Carniero, T.C.B., Giunta, C., Rohrbach, M., Janner, M., Semler, O., Beleggia, F., Li, Y., Yigit, G., Reintjes, N., Altmüller, J., Nürnberg, P., Cavalcanti, D.P., Zabel, B., Warman, M.L., Bertola, D.R., Wollnik, B., Netzer, C., 2019. Autosomal-Recessive Mutations in MESD Cause Osteogenesis Imperfecta. Am. J. Hum. Genet. 105(4),836-843. doi: 10.1016/j.ajhg.2019.08.008.

Nakashima, K., Zhou, X., Kunkel, G., Zhang, Z., Deng, J.M., Behringer, R.R., de Crombrugghe, B., 2002. The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation. Cell. 108, 17–29. doi: 10.1016/s0092-8674(01)00622-5.

Quan, G.M., Ojaimi, J., Li, Y., Kartsogiannis, V., Zhou, H., Choong, P.F., 2005. Localization of pigment epithelium-derived factor in growing mouse bone. Calcif. Tissue. Int. 76(2),146–53. doi: 10.1007/s00223-004-0068-2.

Sinha KM, Zhou X. Genetic and molecular control of osterix in skeletal formation. J Cell Biochem. 2013 May;114(5):975-84. doi: 10.1002/jcb.24439.

Timpson, N.J., Tobias, J.H., Richards, J.B., Soranzo, N., Duncan, E.L., Sims, A.M., Whittaker, P., Kumanduri, V., Zhai, G., Glaser, B., Eisman, J., Jones, G., Nicholson, G., Prince, R., Seeman, E., Spector, T.D., Brown, M.A., Peltonen, L., Smith, G.D., Deloukas, P., Evans, D.M., 2009. Common variants in the region around Osterix are associated with bone mineral density and growth in childhood. Hum. Mol. Genet.18(8),1510-7. doi: 10.1093/hmg/ddp052.

Umair, M., Alhaddad, B., Rafique, A., Jan, A., Haack, T.B., Graf, E., Ullah, A., Ahmad, F., Strom, T.M., Meitinger, T., Ahmad, W., 2017. Exome Sequencing Revealed a Novel Homozygous Splice Site Variant in WNT1 Gene Underlying Osteogenesis Imperfecta Type 3. Pediatr. Res. 82(5),753-758. doi: 10.1038/pr.2017.149.

Umair, M., Eckstein, G., Rudolph, G., Strom, T., Graf, E., Hendig, D., Hoover, J., Alanay, J., Meitinger, T., Schmidt, H., Ahmad, W., 2018. Homozygous XYLT2 variants as a cause of spondyloocular syndrome. Clin. Genet. 93(4),913-918. doi: 10.1111/cge.13179.

Umair, M., Hassan, A., Jan, A., Ahmad, F., Imran, M., Samman, M,I., Basit, S., Ahmad, W., 2016. Homozygous sequence variants in the FKBP10 gene underlie osteogenesis imperfecta in consanguineous families. J Hum Genet. 2016; 61(3):207-213. doi: 10.1038/jhg.2015.129.

Umair, M., Shah, K., Alhaddad, B., Haack, T.B., Graf, E., Strom, T.M., Meitinger, T., Ahmad, W., 2017. Exome sequencing revealed a splice site variant in the IQCE gene underlying post-axial polydactyly type A restricted to lower limb. Eur. J. Hum. Genet. 25(8), 960-965. doi: 10.1038/ejhg.2017.83.

Umair, M., Ahmad, F., Bilal, M., Asiri, A.A., Younus, K., Khan, A., 2019. Genetic skeletal disorders in Pakistan: A brief commentary. Meta. Gene. 20, 100559. https://doi.org/10.1016/j.mgene.2019.100559.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., et al., 2001. The sequence of the human genome. Science. 291,1304-1351. doi: 10.1126/science.1058040.

Wang, J.Y., Liu, Y., Song, L.J., Lv, F., Xu, X.J., San, A., Wang, J., Yang, H.M., Yang, Z.Y., Jiang, Y., Wang, O., Xia, W.B., Xing, X.P., Li, M., 2017. Novel Mutations in SERPINF1 Result in Rare Osteogenesis Imperfecta Type VI. Calcif. Tissue. Int. 100(1), 55-66. doi: 10.1007/s00223-016-0201-z.

Webb, B., Sali, A., 2014. Protein structure modeling with MODELLER. Methods. Mol. Biol. 1137,1-15. doi: 10.1007/978-1-4939-0366-5_1.

Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic. Acids. Res. 35, W407-410. doi: 10.1093/nar/gkm290.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y., 2015. The I-TASSER Suite: protein structure and function prediction. Nat. Methods.12(1),7-8. doi: 10.1038/nmeth.3213.

Figure Legends

Figure 1: (a) Pedigree of family A segregating OI in an autosomal recessive manner. (**b**) Radiographs of the proband (IV-2) showing skeletal features with severe fractures in the upper and lower limbs. **(c)** The lower limbs showing severe bending and fixed flexion deformity (FFD). (**d**) Severe kyphoscolisis along with both dorsal kyphosis and lateral scoliosis. **(e)** Shrieked rib cage. **(f-g)** Proband showing enlarged skull. **(h)** Schematic representation of *SP7* gene and red arrow indicating the missense mutation identified in the present study. Intronic and exonic regions are not drawn to scale. **(i-k)** Sanger sequence electrograms of the *SP7* gene shows a substitution of nucleotide G with A at position 824 [c.824G>A, p.(Cys275Tyr)] in homozygous normal, heterozygous carrier and homozygous affected individual. **(l)** Conservation of cysteine275 across different species. **(m)** Schematic representation of SP7 protein with red arrow indicating position of the mutation p.(Cys275Tyr). Strutural modeling of wild type **(n, o)**, mutated **(p)** and comparision of the both the proteins **(q).**

Figure 2: (a) Pedigree of family B segregating OI in an autosomal recessive manner. **(b-d)** The proband (IV-1, IV-2) showed typical features of OI phenotypes such as severe fractures and healing of upper and lower limbs exhibiting zig zag limbs. **(e)** Schematic representation of SERPINF1 (PEDF) protein and blue arrow indicating position of the nonsense mutation p.(Gln133*). **(f-h)** Sanger sequence electrograms of the *SERPINF1* gene shows a substitution of nucleotide C with T at position 397 [c.397C>T, p.(Gln133*)] in homozygous normal, heterozygous carrier and homozygous affected individual. **(i)** Conservation of Gln133 across different species.

Figure 3: (a) Pedigree of family C segregating OI in an autosomal recessive manner. **(b-e)** Image and radiographs of the affected individual (IV-2) showing mild to severe fractures in the

lower limbs. **(f, g)** Schematic representation of the *SPARC* gene and protein showing position of the muataion [c.497G>A, p.(Arg166His)]. Sanger sequence electrograms of the *SPARC* gene showing a substitution of nucleotide G with A at position 497 (c.497G>A) in homozygous normal, heterozygous carrier and homozygous affected individual. **(i)** Conservation of Arg166 across different species.

Figure 4: (a, b) Pedigrees of family D and E segregating OI in an autosomal recessive manner. **(c-f)** Photograph and radiographs of affected individual (IV-2) showing multiple fracures in the upper and lower limb, resulting in confinement of the patient to wheelchair. **(g-i)** Photograph and radiographs of affected individual (IV-1) showing severe fractures in the lowe limbs. **(j, k)** Schematic representation of the *WNT1* gene and protein representing all the mutations reported to date including the one, in red, identified in the present study. **(l-n)** Sanger electrograms of family D in which a missense mutation [c.677C>T, p.(Ser226Leu)] was identified in the *WNT1* gene. **(o-q)** Sanger electrograms of family E carrying a splice site mutation (c.359-3C>G) in the *WNT1* gene.

Table 1: Latest OI classification.

Table 2: Clinical phenotypes comparison of all the five families investigated here.

Table 3: To-date *SP7* genes mutations.

Table 4: *SERPINF1* mutation update.

Table 5: *SPARC* mutation update.

Table 6: *WNT1* mutation update.

Table 1. Latest OI classification.

Type	Gene	Locus	OMIM	Protein	Inheritance
OI I	***COL1A1***	**17q21.33**	**166200**	**COL1A1**	**AD**
OI II	***COL1A1***	**17q21.33**	**166210**	**COL1A1**	**AD**
OI III	***COL1A1***	**17q21.33**	**259420**	**COL1A1**	**AD**
OI IV	***COL1A1***	**17q21.33**	**166220**	**COL1A1**	**AD**
OI V	***IFITM5***	**11p15.5**	**610967**	**IFITM5**	**AD**
OI VI	***SERPINF1***	**17p13.3**	**613982**	**PEDF**	**AR**
OI VII	***CRTAP***	**3p22.3**	**610682**	**CRTAP**	**AR**
OI VIII	***LEPRE1***	**1p34.2**	**610915**	**P3H1**	**AR**
OI IX	***PPIB***	**15q22.31**	**259440**	**CYPB**	**AR**
OI X	***SERPINH1***	**11q13.5**	**613848**	**HSP47**	**AR**
OI XI	***FKBP10***	**17q21.2**	**610968**	**FKBP65**	**AR**
OI XII	***SP7***	**12q13.13**	**613848**	**OSX (Osterix)**	**AR**
OI XIII	***BMP1***	**8p21.3**	**614856**	**BMP1**	**AR**
OI XIV	***TMEM38B***	**9q31.2**	**615066**	**TMEM38B**	**AR**
OI XV	***WNT1***	**12q13.12**	**615220**	**WNT1**	**AR**
OI XVI	***CREB3L1***	**11p11.2**	**616229**	**CREB3L1**	**AR**
OI XVII	***SPARC***	**5q33.1**	**616507**	**SPARC**	**AR**
OI XVIII	***TENT5A***	**6q14.1**	**617952**	**TENT5A**	**AR**
OI XIX	***MBTPS2***	**Xp22.12**	**301014**	**MBTPS2**	**XLR**
OIXX	***MESD***	**15q25.1**	**607783**	**MESD**	**AR**

Table 2: Clinical phenotypes comparison of all the five families investigated here.

Features observed	**Family A**	**Family B**			**Family C**		**Family D**		**Family E**
Consanguineous pedigree	Yes	Yes			Yes		Yes		Yes
Mutation type	Missense	Nonsense			Missense		Splice site		Missense
Variant	c.824G>A	c.397C>T			c.497G>A		c.359-3C>G		c.677C>T
Protein change	p.Cys275Tyr	p.Gln133*			p.Arg166His		-----		p.Ser226Leu
Race	Pakistani	Pakistani			Pakistani		Pakistani		Pakistani
Sex (male: female)	1 Male	Male	Male	Male	Female	Female	Male	Male	1 Male
Patient #	IV-2	IV-1	IV-2	IV-3	IV-2	IV-4	IV-2	IV-3	IV-1
Age (years)	**32**	**10**	**12**	**14**	**4**	**22 months (deceased)**	**26**	**30**	**5**
Height [cm] (Z-score)	**116.1 (-1.95)**	**118 (-0.976)**	**114 (-1.12)**	**121 (-0.869)**	**122 (-1.192)**		**111 (-1.23)**	**112 (-1.192)**	**110 (-1.939)**
Weight [kg] (Z-score)	**30 (-1.649)**	**33 (-0.735)**	**34 (-0.679)**	**30 (-0.902)**	**22 (-1.658)**		**31 (-1.101)**	**28 (-1.229)**	**21 (-1.88)**
Head Circumference [cm] (Z-score)	**61 (-1.633)**	**56 (0.304)**	**55 (-0.456)**	**54 (-1.066)**	**54 (-1.3719)**		**58 (0.697)**	**59 (1.27)**	**53 (-1.377)**
Fractures of extremities	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Dentinogenesis imperfecta	No	No	No	No	No		No	No	No
Bowing of extremities	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Hypermobility of joints	Yes	Yes	Yes	Yes	No		Yes	Yes	Yes
Craniofacial features affected	Yes	Yes	Yes	Yes	Normal		Yes	Yes	Normal
Hearing impairment	Normal	Normal	Normal	Normal	Normal		Normal	Normal	Normal

Growth retardation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intellectual development	Normal	Normal	Normal	Normal	Normal	Normal	Normal	Normal
flexion deformity	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cardiac impairment	No	No	No	No	No	No	No	No
Kidney stones	Yes	No	No	No	No	No	No	Yes
Treatment	Genetic counseling suggested	Genetic counseling suggested			Genetic counseling suggested	Genetic counseling suggested		Genetic counseling suggested

Table 3: To-date *SP7* genes mutations.

S.No	Amino acid change	Nucleotide change	Reported phenotype
1	p.Arg316Cys	c.946C>T	Autosomal recessive osteogenesis imperfecta, with bone fragility & hearing impairment
2	p.Glu351Glyfs*19	c.1052delA	Autosomal recessive osteogenesis imperfecta

Table 4: *SERPINF1* mutation update.

S.No	Amino acid change	Nucleotide change	Reported phenotype
1	p.Met1Val	c.1A>G	Osteogenesis imperfecta IV
2	p.Ala56Gly	c.167C>G	Otosclerosis
3	p.Gly62Ser	c.184G>A	Osteogenesis imperfecta III
4	p.Ser81Cys	c.242C>G	Osteogenesis imperfecta
5	p.Arg99*	c.295C>T	Osteogenesis imperfecta
6	p.Gln133*	c.397C>T	Osteogenesis imperfecta III
7	p.Gln178*	c.532C>T p	Osteogenesis imperfecta
8	p.Trp217*	c.651G>A	Osteogenesis imperfecta
9	p.Tyr232*	c.696C>G	Osteogenesis imperfecta
10	p.Arg303*	c.907C>T	Osteogenesis imperfecta III
11	p.Val356Glu	c.1067T>A	Osteogenesis imperfecta
12	p.Trp364*	c.1091G>A	Osteogenesis imperfecta
13	p.Gln378*	c.1132C>T	Osteogenesis imperfecta
14	--------	c.283+1G>T	Osteogenesis imperfecta III
15	--------	c.439+34C>T	Osteogenesis imperfecta VI
16	--------	c.787-10C>G	Osteogenesis imperfecta VI
17	--------	c.787-617G>A	Osteogenesis imperfecta IV
18	p.Leu83Glnfs*28	c.248_249delTC	Osteogenesis imperfecta III
19	p.Ile142Serfs*9	c.423delG	Osteogenesis imperfecta
20	p.Arg167Serfs*35	c.498_499delCA	Osteogenesis imperfecta III
21	p.Val218Glufs*22	c.653delT	Autosomal recessive osteogenesis imperfecta
22	p.Trp217*	c.650_653delGGGT	Osteogenesis imperfecta
23	p.Phe277del	c.829_831delTTC	Osteogenesis imperfecta VI
24	p.Leu280Glufs*20	c.838_839delCT	Osteogenesis imperfecta
25	p.Thr294Profs*8	c.879delC	Osteogenesis imperfecta III
26	p.Pro373Glnfs*18	c.1118_1119delCC	Osteogenesis imperfecta
27	p.Phe384Leufs*9	c.1152_1170del19	Osteogenesis imperfecta VI
28	p.Thr401Argfs*28	c.1202_1203delCA	Osteogenesis imperfecta III
29	--------	c.-9+2dupT	Osteogenesis imperfecta
30	p.Glu27Glyfs*38	c.77dupC	Osteogenesis imperfecta
31	p.Ser84Lysfs*28	c.250dupA	Osteogenesis imperfecta
32	p.Leu89Argfs*26	c.261_265dupGGCCC	Osteogenesis imperfecta III
33	p.Ala91_Ser93dup	c.271_279dupGCCCTCTCG	Osteogenesis imperfecta VI
34	p.yr109Serfs*5	c.324_325dupCT	Osteogenesis imperfecta
35	p.Arg141Profs*5	c.421dupC	Osteogenesis imperfecta
36	p.His389Glnfs*4	c.1163_1166dupATCA	Osteogenesis imperfecta
37	p.Lys147_Gly215delinsArg	c.439+127_643+545del1310	Osteogenesis imperfecta VI
38	--------	c.752_753insKC847088.1:g.51_393	

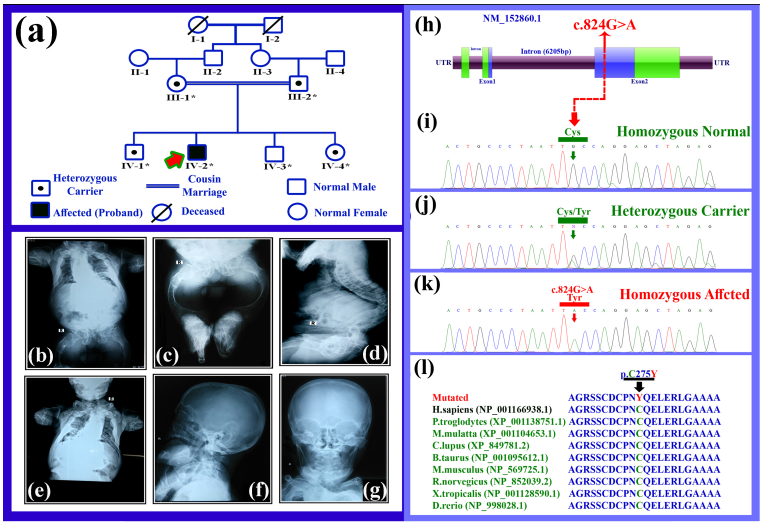
Table 5: *SPARC* mutation update.

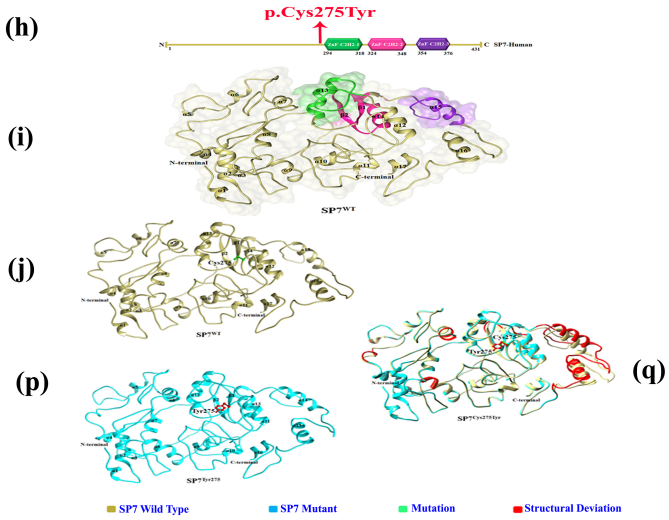
S.No	Amino acid change	Nucleotide change	Reported phenotype
1	p.Arg166His	c.497G>A	Osteogenesis imperfecta
2	p.Glu263Lys	c.787G>A	Osteogenesis imperfecta
3	--------	c.*582C>G	Osteogenesis imperfecta

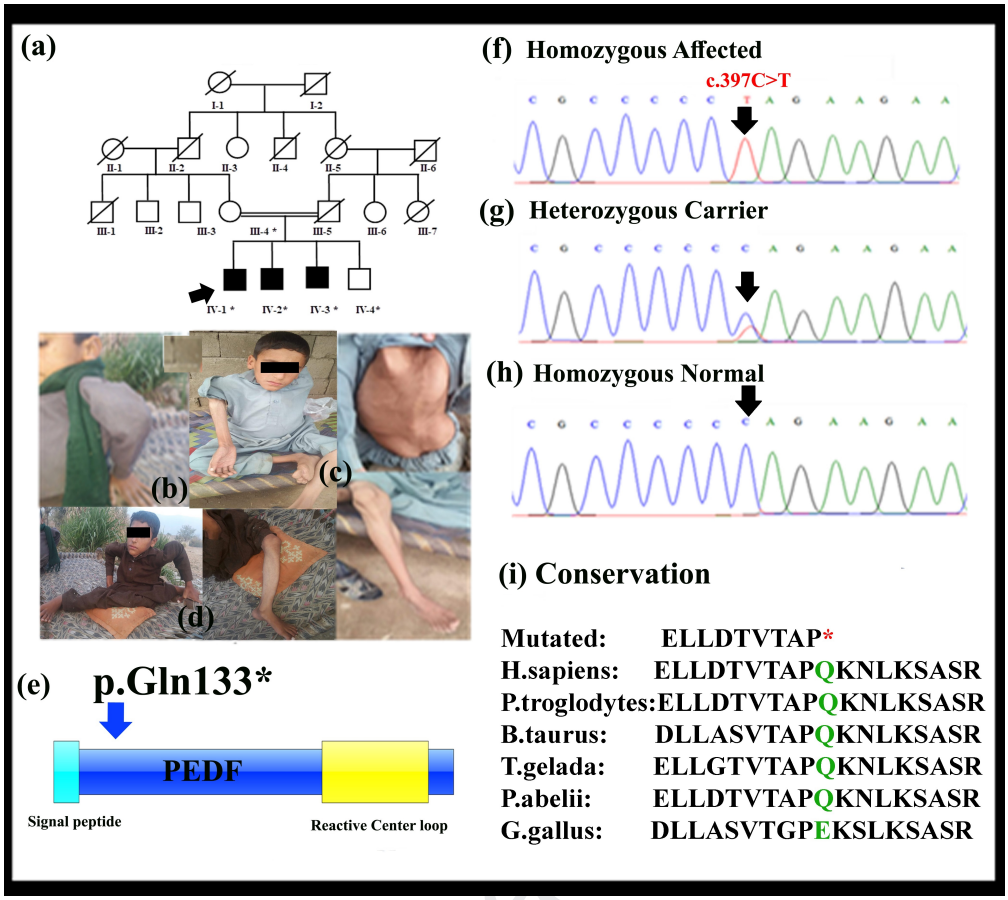
Table 6: *WNT1* mutation update.

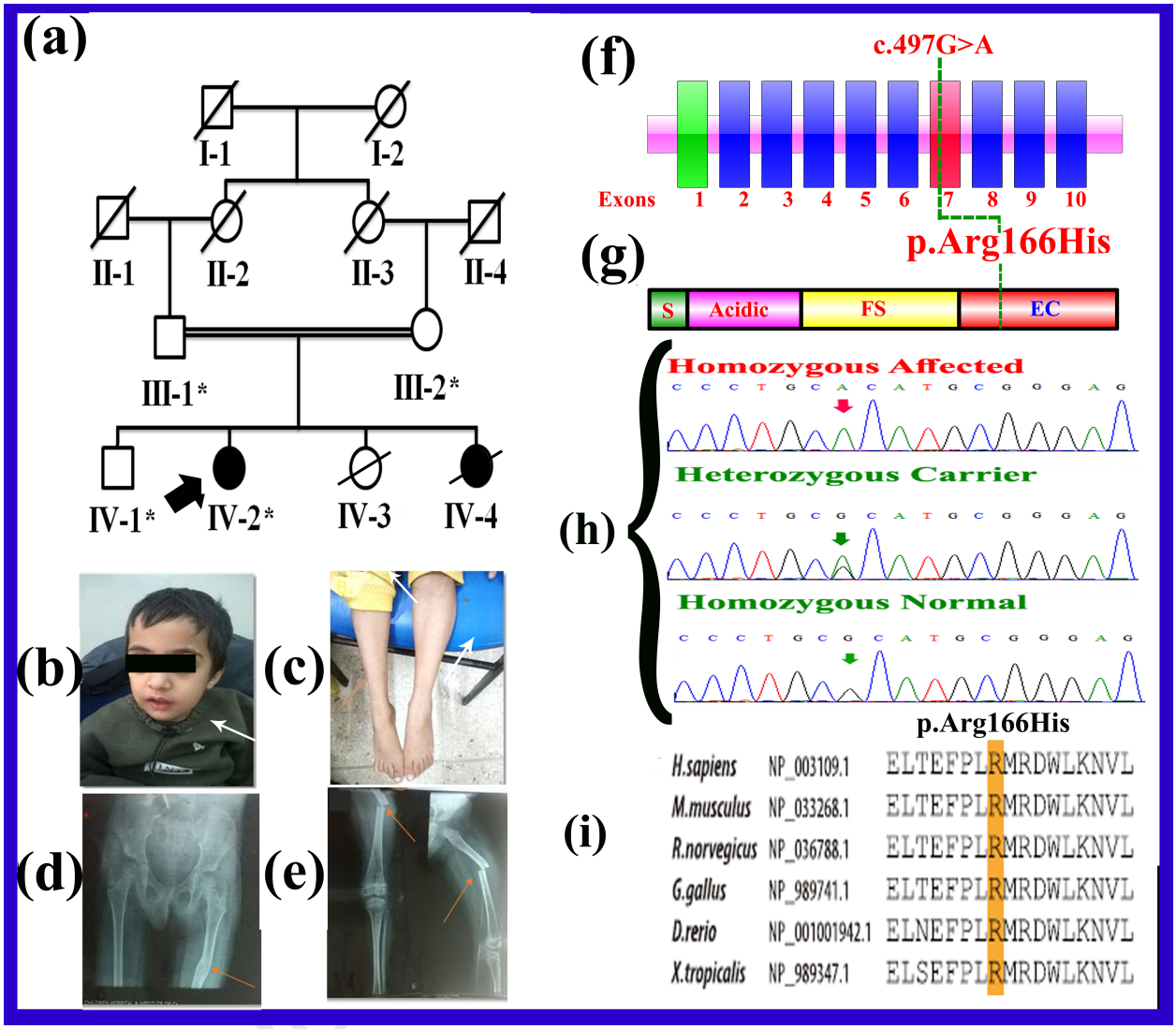
S.No	Amino acid change	Nucleotide change	Reported phenotype
1	p.Met1Thr	c.2T>C	Osteogenesis imperfecta
2	p.Gly36Asp	c.107G>A	Osteoporosis
3	p.Ile37Thr	c.110T>C	Osteogenesis imperfecta III
4	p.Gln62*	c.184C>T	Osteogenesis imperfecta
5	p.Gln87*	c.259C>T	Osteogenesis imperfecta
6	p.Arg101Cys	c.301C>T	Osteogenesis imperfecta III
7	p.Trp102Cys	c.306G>T	Osteogenesis imperfecta
8	p.Glu123Asp	c.369A>C	Osteogenesis imperfecta
9	p.Thr124Met	c.371C>T	Osteogenesis imperfect IV
10	p.Phe128Val	c.382T>G	Osteogenesis imperfecta III
11	p.Ala129Thr	c.385G>A	Osteogenesis imperfect IV
12	p.Ala133Thr	c.397G>A	Osteogenesis imperfecta
13	p.Gly134Val	c.401G>T	Osteoporosis
14	p.Arg141Cys	c.421C>T	Osteogenesis imperfect IV
15	p.Cys143Phe	c.428G>T	Osteogenesis imperfecta
16	p.Gly146Arg	c.436G>C	Osteoporosis
17	p.Cys153Gly	c.457T>G	Osteogenesis imperfecta
18	p.Trp167Cys	c.501G>C	Osteogenesis imperfect IV
19	p.Gly169Asp	c.506G>A	Osteogenesis imperfect IV
20	p.Gly169Cys	c.505G>T	Osteogenesis imperfecta III
21	p.Gly177Cys	c.529G>T	Osteogenesis imperfecta
22	p.Arg207His	c.620G>A	Osteogenesis imperfect IV
23	p.Cys218Gly	c.652T>G	Autosomal dominant, Osteogenesis imperfecta
24	p.Gly225Val	c.674G>T	Osteogenesis imperfecta
25	p.Ser226Leu	c.677C>T	Osteogenesis imperfecta
26	p.Cys227Term	c.681C>A	Osteogenesis imperfecta III
27	p.Arg235Trp	c.703C>T	Osteogenesis imperfecta
28	p.Leu257Pro	c.770T>C	Osteogenesis imperfecta I
29	p.Tyr258*	c.774C>A	Osteogenesis imperfecta III
30	p.Ser295*	c.884C>A	Osteogenesis imperfecta
31	p.Phe298Cys	c.893T>G	Osteogenesis imperfecta
32	p.Arg313Cys	c.937C>T	Osteogenesis imperfecta III
33	p.Gly324Cys	c.970G>T	Osteogenesis imperfecta
34	p.Cys330*	c.990C>A	Osteogenesis imperfecta
35	p.Arg337Pro	c.1010G>C	Osteogenesis imperfecta III
36	p.Trp351Arg	c.1051T>C	Osteogenesis imperfecta
37	p.Val355Phe	c.1063G>T	Osteogenesis imperfecta
38	--------	c.105-2A>G	Osteogenesis imperfecta
39	--------	c.104+1G>A	Osteogenesis imperfecta IV
40	--------	c.359-3C>G	Osteogenesis imperfecta
41	--------	c.624+4A>G	Osteogenesis imperfecta

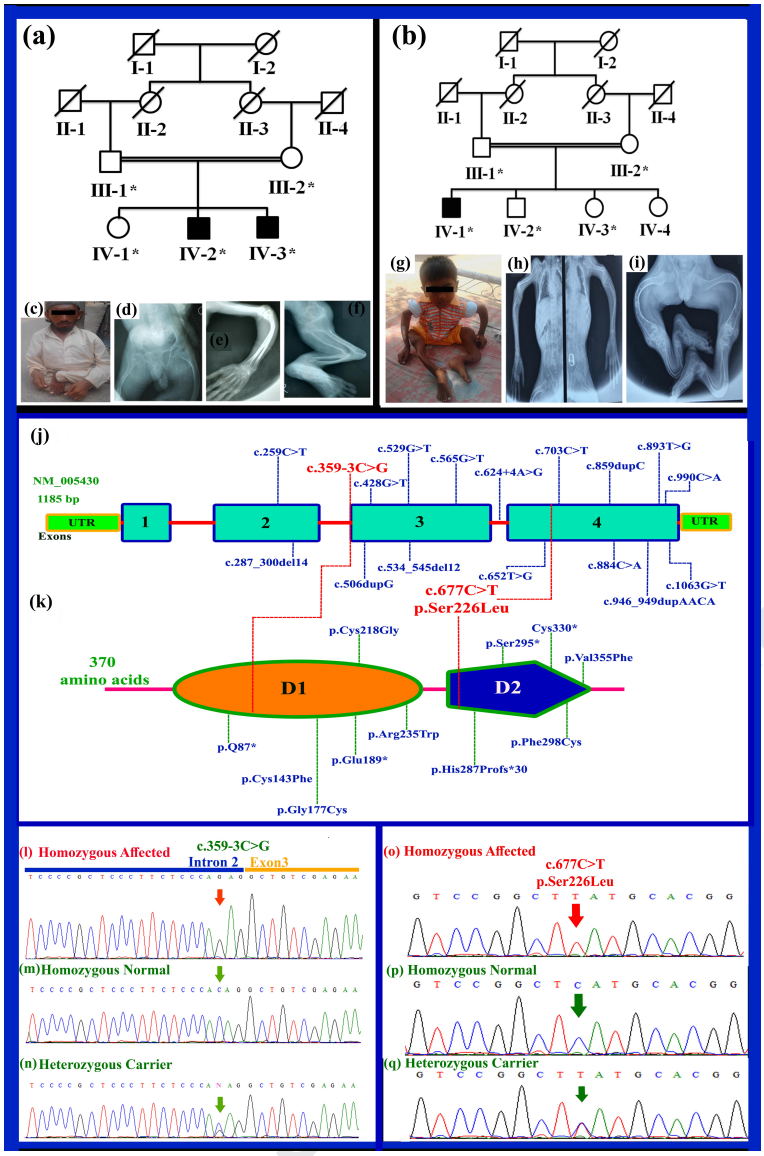
42	p.Leu3Serfs*36	c.6delG	Osteogenesis imperfecta
43	p.Leu64*	c.189delG	Osteogenesis imperfecta IV
44	p.Leu86Cysfs*113	c.255delG	Osteogenesis imperfecta
45	p.Gln96Profs*54	c.287_300del14	Osteogenesis imperfecta
46	p.Arg156Glyfs*43	c.466delC	Osteogenesis imperfecta III
47	p.Leu179_Arg182del	c.534_545del12	Osteogenesis imperfecta
48	p.Val229Hisfs*86	c.685_689delGTGCG	Osteogenesis imperfecta
49	p.His287Thrfs*106	c.859delC	Osteogenesis imperfecta
50	p.Glu343Serfs*50	c.1026delC	Osteogenesis imperfecta
51	p.Arg73Thrfs*82	c.216dupA	Osteogenesis imperfecta
52	p.Cys170Leufs*6	c.506dupG	Osteogenesis imperfecta
53	p.His287Profs*30	c.859dupC	Osteogenesis imperfecta
54	p.Ser317Lysfs*153	c.946_949dupAACA	Osteogenesis imperfecta
55	--------	c.104+4_104+44del41	Osteogenesis imperfecta


(a)
I-1
I-2
II-1
II-2
II-3
II-4
III-1^
III-2^
IV-1^
IV-2^
IV-3^
IV-4^
Heterozygous Carrier
Cousin Marriage
Normal Male
Affected (Proband)
Deceased
Normal Female
(b)
(c)
(d)
(e)
(f)
(g)
(h)
NM_152860.1
c.824G>A
UTR
Intron (6205bp)
UTR
Exon1
Exon2
(i)
Cys
Homozygous Normal
(j)
Cys/Tyr
Heterozygous Carrier
(k)
c.824G>A
Tyr
Homozygous Affcted
(l)
p.C275Y
Mutated AGRSSCDCPNYQELERLGAAAA
H.sapiens (NP_001166938.1) AGRSSCDCPNCQELERLGAAAA
P.troglodytes (XP_001113875.1) AGRSSCDCPNCQELERLGAAAA
M.mulatta (XP_001104653.1) AGRSSCDCPNCQELERLGAAAA
C.lupus (XP_849781.2) AGRSSCDCPNCQELERLGAAAA
B.taurus (NP_001095612.1) AGRSSCDCPNCQELERLGAAAA
M.musculus (NP_569725.1) AGRSSCDCPNCQELERLGAAAA
R.norvegicus (NP_852039.2) AGRSSCDCPNCQELERLGAAAA
X.tropicalis (NP_001128590.1) AGRSSCDCPNCQELERLGAAAA
D.rerio (NP_998028.1) AGRSSCDCPNCQELERLGAAAA


(h)
p.Cys275Tyr
SP7-Human
(i)
N-terminal
C-terminal
SP7WT
(j)
SP7WT
(p)
SP7Tyr275
(q)
SP7Cys275Tyr
SP7 Wild Type
SP7 Mutant
Mutation
Structural Deviation


(a)
I-1
I-2
II-1
II-2
II-3
II-4
II-5
II-6
III-1
III-2
III-3
III-4 *
III-5
III-6
III-7
IV-1 *
IV-2*
IV-3 *
IV-4*
(b)
(c)
(d)
(e)
p.Gln133*
PEDF
Signal peptide
Reactive Center loop
(f) Homozygous Affected
c.397C>T
(g) Heterozygous Carrier
(h) Homozygous Normal
(i) Conservation
Mutated: ELLDTVTAP*
H.sapiens: ELLDTVTAPQKNLKSASR
P.troglodytes: ELLDTVTAPQKNLKSASR
B.taurus: DLLASVTAPQKNLKSASR
T.gelada: ELLGTVTAPQKNLKSASR
P.abelii: ELLDTVTAPQKNLKSASR
G.gallus: DLLASVTGPEKSLKSASR


(a)
I-1
I-2
II-1
II-2
II-3
II-4
III-1*
III-2*
IV-1*
IV-2*
IV-3
IV-4
(b)
(c)
(d)
(e)
(f)
c.497G>A
Exons 1 2 3 4 5 6 7 8 9 10
p.Arg166His
(g)
S
Acidic
FS
EC
(h)
Homozygous Affected
Heterozygous Carrier
Homozygous Normal
(i)
p.Arg166His
H.sapiens NP_003109.1 ELTEFPLRMRDWLKNVL
M.musculus NP_033268.1 ELTEFPLRMRDWLKNVL
R.norvegicus NP_036788.1 ELTEFPLRMRDWLKNVL
G.gallus NP_989741.1 ELTEFPLRMRDWLKNVL
D.rerio NP_001001942.1 ELNEFPLRMRDWLKNVL
X.tropicalis NP_989347.1 ELSEFPLRMRDWLKNVL


(a)
I-1
I-2
II-1
II-2
II-3
II-4
III-1*
III-2*
IV-1*
IV-2*
IV-3*
(b)
I-1
I-2
II-1
II-2
II-3
II-4
III-1*
III-2*
IV-1*
IV-2*
IV-3*
IV-4
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
NM_005430
1185 bp
UTR
Exons
1
2
3
4
UTR
c.259C>T
c.359-3C>G
c.428G>T
c.529G>T
c.565G>T
c.624+4A>G
c.703C>T
c.859dupC
c.893T>G
c.990C>A
c.287_300del14
c.506dupG
c.534_545del12
c.652T>G
c.677C>T
p.Ser226Leu
c.884C>A
c.946_949dupAACA
c.1063G>T
(k)
370
amino acids
D1
D2
p.Cys218Gly
p.Ser295*
Cys330*
p.Val355Phe
p.Q87*
p.Arg235Trp
p.Glu189*
p.Cys143Phe
p.Gly177Cys
p.Phe298Cys
p.His287Profs*30
(l) Homozygous Affected
c.359-3C>G
Intron 2
Exon3
(m) Homozygous Normal
(n) Heterozygous Carrier
(o) Homozygous Affected
c.677C>T
p.Ser226Leu
(p) Homozygous Normal
(q) Heterozygous Carrier