
Identification of a novel COL1A1 frameshift mutation, c.700delG, in a Chinese osteogenesis imperfecta family

Xiran Wang^{1,3,*}, Yu Pei^{1,*}, Jingtao Dou¹, Juming Lu¹, Jian Li², Zhaojun Lv¹

¹ Department of Endocrinology

² Department of Elderly Endocrinology

³ Department of Cadres Ward

* These authors contributed equally to this article.

10.1590/S1415-475738120130336- Editorial

Abstract

Osteogenesis imperfecta (OI) is a family of genetic disorders associated with bone loss and fragility. Mutations associated with OI have been found in genes encoding the type I collagen chains. People with OI type I often produce insufficient $\alpha 1$ -chain type I collagen because of frameshift, nonsense, or splice site mutations in *COL1A1* or *COL1A2*. This report is of a Chinese daughter and mother who had both experienced

two bone fractures. Because skeletal fragility is predominantly inherited, we focused on identifying mutations in *COL1A1* and *COL1A2* genes. A novel mutation in *COL1A1*, c.700delG, was detected by genomic DNA sequencing in the mother and daughter, but not in their relatives. The identification of this mutation led to the conclusion that they were affected by mild OI type I. Open reading frame analysis indicated that this frameshift mutation would truncate $\alpha 1$ -chain type I collagen at residue p263 (p.E234KfsX264), while the wild-type protein

would contain 1,464 residues. The clinical data were consistent with the patients' diagnosis of mild OI type I caused by haploinsufficiency of $\alpha 1$ -chain type I collagen. Combined with previous reports, identification of the novel mutation *COL1A1-c.700delG* in these patients suggests that additional genetic and environmental factors may influence the severity of OI.

Keywords: Osteogenesis imperfecta, Chinese OI type 1 family, type I collagen, sequence analysis, frameshift mutation

Received: 01/12/2013; Accepted: 09/08/2014.

Introduction

Osteogenesis imperfecta (OI), also known as brittle bone disease (Marini *et al.*, 2007), is a group of rare heritable connective tissue disorders characterized by skeletal fractures with mild trauma, secondary deformities, and extraskeletal manifestations (Ben Amor *et al.*, 2011; Marini *et al.*, 2007). Such manifestations may include blue sclerae, progressive hearing loss, dentinogenesis imperfecta, and hyperlaxity of ligaments and skin (Ben Amor *et al.*, 2011; Marini *et al.*, 2007). The prevalence of OI ranges from 1 in 10,000 to 1 in 25,000 (Byers and Cole, 2002).

OI was originally classified into four clinical types, I–IV, based on its clinical features and radiographs (Silence *et al.*, 1979; Table S1). Type V has been associated with hyperplastic callus formation (Glorieux *et al.*, 2000), type VI with a mineralization defect (Glorieux *et al.*, 2002) and type VII with a range of severity that may be lethal (Ward *et al.*, 2002). OI types VIII and IX are usually lethal (Barnes *et al.*, 2012; Moul *et al.*, 2013). OI severity ranges from intrauterine fractures and perinatal lethality to extremely mild forms without fractures (Ben Amor *et al.*, 2011). The severity of bone fragility of OI patients is highest in types VIII and IX, and declines progressively from type II to type VII. Type I exhibits the least severe fragility (Rauch and Glorieux, 2005). Most OI types (I–VI) show autosomal dominant inheritance, but two of the rarer types (VII and VIII) exhibit autosomal recessive disorders (Ward *et al.*, 2002; Ben Amor *et al.*, 2011; Table S1).

Type I collagen fibers contain three polypeptide chains. The vast majority of fibrils are composed of two $\alpha 1$ and one $\alpha 2$ chains, but a small number are composed of three $\alpha 1$ chains in a triple helix structure. The translated collagen α chains include propeptides at each end,

one at the amino terminus end and one at the carboxyl terminus end. These propeptides help to connect and arrange the three chains which subsequently intertwine in a zipper-like fashion into the triple helix configuration.

Most patients with types I–IV OI (90%) have a mutation in one of the two collagen type I genes, *COL1A1* or *COL1A2* (Ben Amor *et al.*, 2011; van Dijk *et al.*, 2012; Table S1). *COL1A1* extends 18 kb from 17q21.33 to 17q22 and has 52 exons. *COL1A2* extends 37 kb from 7q21.3 to 7q22.3 and has 52 exons. More than 200 mutations in collagen 1 genes have been associated with OI (Ben Amor *et al.*, 2011). These mutations occur in most exons and introns of *COL1A1* and *COL1A2* (<http://www.le.ac.uk/ge/collagen/>; Dalgleish, 1997, 1998).

Many OI type I patients have *COL1A1* and *COL1A2* frameshift mutations, splice site mutations, and insertion mutations resulting in haploinsufficiency, or the expression of approximately half the amount of collagen type I mRNA and protein, leading to weaker bones (Redford-Badwal *et al.*, 1996; Slayton *et al.*, 2000; Peng *et al.*, 2012). Other deletions, insertions, and point mutations in *COL1A1* or *COL1A2* induce conformational changes in the collagen $\alpha 1$ and $\alpha 2$ chains and lead to knot microfibril structure disorders, altered deposition of hydroxyapatite, brittleness, reduced resistance to external force, altered stability of the collagen triple helix, and fragility (Pallos *et al.*, 2001; Witecka *et al.*, 2008; Ben Amor *et al.*, 2011).

Clinical diagnosis of osteogenesis imperfecta is straightforward in individuals with a positive family history and several typical OI features (van Dijk *et al.*, 2012). However, despite the presence of several physical manifestations, some patients are difficult to diagnose in the absence of an affected family member and without apparent association between bone fragility and obvious extraskeletal abnormalities. In these cases, analyses of type I collagen genes for mutations can provide helpful information since 90% of OI-associated mutations arise in collagen I (van Dijk *et al.*, 2012).

Here, we present the results of *COL1A1* and *COL1A2* sequencing analysis from DNA samples collected from two related Chinese patients with several typical features of OI. A novel mutation not previously associated with *COL1A1* was detected. These results will help to provide a better diagnosis of mild OI type I when physical symptoms are mild.

Materials and Methods

Patients

Two related Han Chinese patients, mother and daughter, presented with blue sclerae, bone pain, and dentinogenesis imperfecta. Each had experienced two bone fractures. We reviewed their clinical features, medical history, physical examination, radiologic images, and biochemical parameters, and diagnosed them with type I osteogenesis imperfecta. Both affected family members, unaffected family members (maternal grandmother, father, and two sisters), and controls agreed to a genetic analysis and provided informed consent for the study approved by the Endocrinology Department of General Hospital of PLA. Their pedigree is shown in Figure 1.

Genetic analysis

Blood samples were obtained from the two patients, their family members, and 50 ethnically matched, unrelated controls. Genomic DNA was isolated from peripheral leukocytes by using a whole blood DNA extraction kit (Qiagen Co., Valencia, CA) according to the manufacturer's instructions. PCR was performed to detect sequences in all exons and exon-intron boundaries of *COL1A1* and *COL1A2*. Fifty-seven pairs of primers, 17 for *COL1A1* and 40 for *COL1A2* (sequences available upon request), were designed and verified with Primer 5 software (Softonic, Shanghai, China) using the genomic sequences of *COL1A1* (GenBank accession NG_007400.1) and *COL1A2* (Genbank accession NG_007405.1).

DNA sequence analysis was performed by automated sequencing in an ABI genetic analyzer (Model 377; ABI, USA) by the Ying Jun Gene Company (Shanghai, China). Sequencing results were compared with the GenBank database and confirmed through sequencing of clones. DNA was cloned into a PGEM-T Easy Vector following the manufacturer's instructions (Promega, Madison, WI). The osteogenesis imperfecta and Ehlers-Danlos syndrome variant databases compiled of previously reported OI mutations in *COL1A1* and *COL1A2*, the 1000 Genomes database, and the UniProKB database were also used for comparison. PCR and sequencing of *COL1A1* and *COL1A2* genes in samples of other family members were performed with the same methods.

Open reading frame analysis of the mutated sequence was performed by the ORF Finder, part of the BLAST software suite. The parental protein sequence and the protein sequence corresponding to the c.700delG mutation were aligned using the align function of BLAST software (NCBI). The 1000 Genomes database was also used to search for the presence of this mutation by using the alignment function of the

BLAST software and the following search sequence: 5'-TTTTCTCTCCCTCTCAGGGGA AGCTGGAAAAC-CTGGT CGTCCTGGTGAGCG-3'.

Results

Patient description

The proband was a 13-year-old girl of Han descent from the Shanxi province of China. Her spontaneous full-term delivery was uneventful. The proband began walking at 15 months. She had suffered two bone fractures in the long bone and vertebra. The first fracture occurred in the right tibia/fibula after a slight injury at 16 months old and its healing was achieved with fixation of a plate. The second fracture was caused by a slight hit to the abdomen. The proband did not grow in height during the last two years. Mental development was normal. Physical examination revealed short stature, blue sclerae, discrete signs of dentinogenesis imperfecta, pigeon breast, deformities of slight scoliosis and joint hypermobility. X-rays showed thoracic-lumbar vertebra compression fracture and osteoporosis, acetabulum, and femoral head protrusion acetabulum into the pelvis.

The mother is a 43-year-old of Han Chinese descent. She had two fractures in childhood: one in the right tibia/fibula that occurred at age 5 and the other in the right radius at age 8. She presented with hearing loss that had onset near the age of 20. Physical examination revealed blue sclerae, discrete signs of dentinogenesis imperfecta, deformity of scoliosis, and joint hypermobility. Blood samples were analyzed from the two patients and blood serum calcium, phosphorus, and parathyroid hormone levels were all within the normal range (data not shown).

Mutation analysis

Sequence analysis revealed no mutations in the *COL1A2* gene from the proband or her mother. However, sequence analysis of *COL1A1* identified a heterozygous frameshift mutation in exon 10 from both subjects at nucleotide 700 of the mRNA, named as c.700delG, with nucleotide 1 defined as A of the initiating ATG (Figure 2). This mutation was not identified in other family members (two sisters, father, and maternal grandmother of the proband) or ethnically matched controls. Importantly, a comparison with the compiled list of *COL1A1* mutations associated with OI in the osteogenesis imperfecta and Ehlers-Danlos syndrome variant databases (Dalglish, 1997, 1998 and <http://www.uniprot.org/uniprot/P02452a>) revealed that

the mutation is novel and had not been previously reported.

This novel, heterozygous c.700delG frameshift mutation in *COL1A1* was inherited in an autosomal dominant manner and was correlated with OI symptoms (Figure 2). An open reading frame (ORF) analysis indicated that translation of the *COL1A1*-c.700delG mRNA would initiate at the standard methionine and would be 100% homologous until amino acid 233. However, the mutation would produce distinct amino acids from p234 to p263 when compared to wild-type collagen α_1 and would terminate at the UGA stop codon at nucleotide 792 in exon 10 (Figure 3). This *COL1A1* mutation would be named p.E234KfsX264. According to the ORF analysis, the putative 263 amino acid collagen α_1 700delG mutation would contain the signal peptide (p1–22), propeptide (p23–161) and the first 102 residues of the main collagen α_1 chain (p162–1218 and the triple helical region from p179–1192), but no C-terminal propeptide (p1219–1464). Thus, the mutation produces a truncated form of *COL1A1*. The second allele was normal in the proband and mother. The mutation was not described in genomes submitted to the 1000 Genomes database.

Discussion

Mutations in *COL1A1* and *COL1A2* account for approximately 90% of OI cases and usually occur in patients with OI types I–IV. OI symptoms can also occur in patients with mutations in genes that code proteins that interact with type I collagen, including cartilage associated protein (CRTAP; Starman *et al.*, 1989; Barnes *et al.*, 2006; Morello *et al.*, 2006; Valli *et al.*, 2012), prolyl-3-hydroxylase 1 (P3H1/LEPRE1), cyclophilin B (PPIB; Morello *et al.*, 2006; Fratzl-Zelman *et al.*, 2010; Marini *et al.*, 2010; Valli *et al.*, 2012; Zhang *et al.*, 2012), FKBP10 (Barnes *et al.*, 2012), WNT-1 (Pyott *et al.*, 2013), and serpin peptidase inhibitor clade F member 1 (SERPINF1; Homan *et al.*, 2011). In our report, the clinical presentation of OI in the proband and her mother was mild, and the associated finding of a frameshift mutation in both patients allowed diagnosis of type I osteogenesis imperfecta.

Many patients diagnosed with mild OI type I have null variants in either *COL1A1* or *COL1A2* genes and a second allele with the wild-type sequence. The phenotype that results from these null variants is due to approximately 50% decrease of the wild-type production of the type 1 collagen α_1 or α_2 chains, known as haploinsufficiency (Redford-Badwal *et al.*, 1996; Slayton *et al.*, 2000; Ben Amor *et al.*, 2013; Peng *et al.*, 2012; Zhang *et al.*, 2012). The phenotypes of patients with haploin-

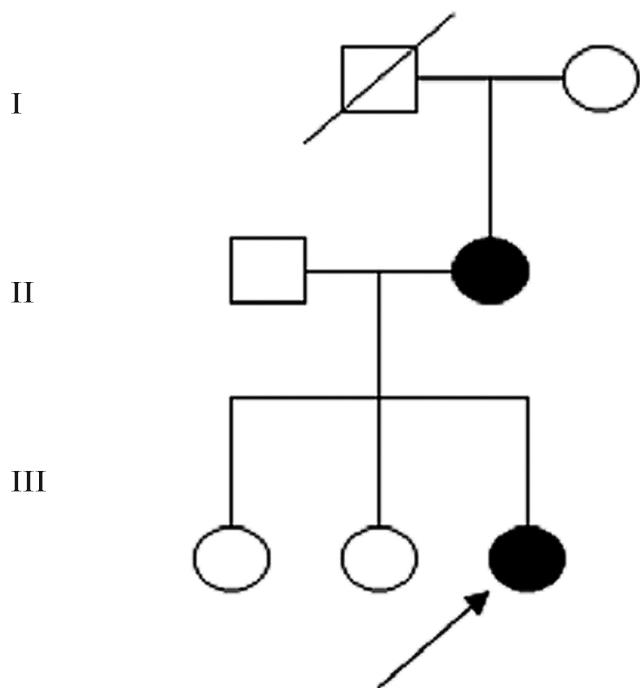


Figura 1: The osteogenesis imperfecta pedigree of the nuclear family of the proband.

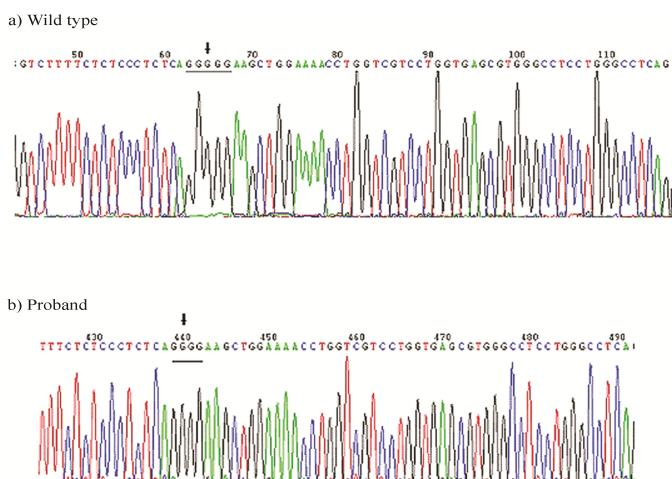


Figura 2: Sequence analysis of *COL1A1* from a wild-type control and the same DNA region from the proband containing the mutation. DNA sequencing was initiated in the ninth intron and extended into the tenth exon. (A) Genomic DNA sequence of the control wild-type allele. (B) DNA sequence of a *COL1A1* clone from the proband. The DNA sequence of the same region from the mother was identical to the proband (data not shown).

sufficiency may be caused by increased degradation of the abnormal type 1 collagen α chain (Garnero *et al.*, 2009), alterations of the mutant mRNA, such as non-sense mediated decay of the mRNA from the mutant allele, or inefficient transport of mRNA to the cytoplasm (Redford-Badwal *et al.*, 1996; Slayton *et al.*, 2000). Non-sense mutations, frameshift mutations, and splice site mutations can induce a premature termination codon in the ORF. For example, 23 of 33 Chinese patients with OI type I had mutations leading to haploinsufficiency in *COL1A1*; the mutations included eight frameshift mutations, eight splice site mutations, and seven nonsense mutations (Zhang *et al.*, 2012). Mutations associated with haploinsufficiency occurred in exons 2, 3, 5, 7, 9, 11, 17, 18, 19, 24, 28, 31, 36, 42, 43, 46, 47, 49, and 52 (Slayton *et al.*, 2000; Zhang *et al.*, 2012) and introns 44i (Peng *et al.*, 2012). Several of the previously described null mutations of *COL1A1* appear to be transcribed into (hnRNA) at equivalent levels to wild-type *COL1A1*, but the mRNA levels of the null mutations are very low or undetectable in the cytoplasm (Slayton *et al.*, 2000).

Similar mutations have been observed less frequently in *COL1A2* (Zhang *et al.*, 2012). In a recent study, eighty six Canadians with OI type I due to haploinsufficiency that resulted from nonsense and frameshift mutations showed a risk of 0.62 bone fractures per year which was approximately 95 times greater than that of healthy person of British decent (Ben Amor *et al.*, 2013). The proband and mother described here had suffered only two fractures each in their lifetimes.

A second class of *COL1A1* and *COL1A2* mutations are translated into procollagen α chains that exhibit altered structure. These mutated procollagen α chains become en-twined with normal collagen α chains and the resulting triple helix has a modified configuration, reducing bone strength. The most common mutations result in replacement of one single glycine in the many essential Gly-X-Y triplets in the triple helical region (p.179–1192). The substituted amino acid may include serine, arginine, aspartic acid, valine, or alanine at a single position (Ben Amor *et al.*, 2011; Xu *et al.*, 2011). These point mutations that replace glycine can occur throughout most of the length of *COL1A1* and affect collagen stability (Witecka *et al.*, 2008; Xu *et al.*, 2008; Xiao *et al.*, 2011). These expressed *COL1A1* and *COL1A2* peptides usually cause more severe OI types (II, III, IV) and usually result from point mutations, insertions, deletions, duplications, and frameshift mutations in the last 200 nucleotides of the reading frame (Xiao *et al.*, 2011).

Table 1 is a summary of current known mutations of *COL1A1* that have been documented. Seven mutations were described in the ninth intron that may reduce the efficiency of the 3' splice site. Interestingly,

three patients with the same mutation presented with OI of varying severity (OI type I or type IV). Six mutations were described in exon 10 in two *COL1A1* databases. The OI patients with mutations c.716G > A, c.725G > A, c.740G > T, c.742G > A, and c.743G > A were diagnosed with type IV, unclassified OI, OI type I, OI type III, and OI type III, respectively using the osteogenesis imperfecta and Ehlers-Danlos syndrome variant databases (Dagleish, 1997, 1998) and the UniProKB database. Taken together, these data, combined with the description of the novel mutation c.700delG described here, suggest that additional genes and possibly environmental factors such as nutritional status during development may influence the severity of OI.

Tabela 1: Comparison of OI types with a mutation in the ninth intron or the tenth exon of *COL1A1*.

Location	Mutation	Type
09i	c.696+2T > G	Substitution
09i	c.697-2A > G	Substitution
09i	c.697-2A > T	Substitution
09i	c.697-2delA	Deletion
09i	c.697_2_697-1del	Deletion
09i	c.697-1G > C	Substitution
09i	c.697-1G > T (3)	Substitution
10	c.700_G-1del	Deletion
10	c.716G > A	Substitution
10	c.725G > A	Substitution
10	c.740C > T	Substitution
10	c.742G > A	Substitution
10	c.743G > A	Substitution

Adapted from the osteogenesis imperfecta and Ehlers-Danlos syndrome variant databases (Dagleish, 1997; Dagleish, 1998).

* Location of clinician who submitted the mutation and characteristics of the carrier of the database.

If translated, the c.700delG allele would result in a mutant collagen $\alpha 1$ that would be significantly shorter than the wild-type collagen $\alpha 1$ (1464 residues) and would not contain two regions, helix positions p.691–893 and p.910–964, in which single point mutations are usually lethal ((Ben Amor *et al.*, 2011). Because of the mild OI type I symptoms and the dominant inheritance in the pedigree, the *COL1A1-c.700delG* variant likely leads to haploinsufficiency. The *COL1A1-c.700delG* was not detected in samples from the proband's maternal grandmother, father, two sisters and was not present in the 50 control samples from unrelated individuals who did not express the *COL1A1* mutation.

Limitations of this study include the fact that we only sequenced *COL1A1* and *COL1A2* genes. While the

dominant inheritance pattern suggests involvement of *COL1A1* or *COL1A2*, other known OI candidate genes which induce a recessive OI were not examined, including CRTAP (Barnes *et al.*, 2006; Morello *et al.*, 2006; Valli *et al.*, 2012), P3H1/LEPRE1, PPIB (Morello *et al.*, 2006; Fratzl-Zelman *et al.*, 2010; Marini *et al.*, 2010; Zhang *et al.*, 2012), FKBP10 (Barnes *et al.*, 2012), WNT-1 (Pyott *et al.*, 2013), and SERPINF1 (Homan *et al.*, 2011). It is theoretically feasible that variants in the aforementioned proteins that interact with collagen may compensate for a collagen α_1 or α_2 variant and thereby contribute to the heterogeneity of the OI symptoms.

Although mutations of *COL1A1* and *COL1A2* have been associated with OI for several decades, the relationship between the clinical symptoms and location of the variants need further elucidation. Additional genetic or epigenetic factors must influence the variety and severity of symptoms since clinical symptoms of patients with the same mutations can vary among family members (Rauch *et al.*, 2013). The effect of collagen α_1 and α_2 variants on proteins that interact with collagen, such as CRTAP, LEPRE1, PPIB, FKBP10, SERPINH1, and SP7 (Garnero *et al.*, 2009; Homan *et al.*, 2011; Ben Amor *et al.*, 2013), would further clarify the mechanisms for the pathology of OI. The novel frameshift variant in exon 10 described here is located only 16 nucleotides upstream of another previously described exon 10 point mutation, c.716G > A, which was detected in a patient with OI type IV (Sarafova *et al.*, 1998). Comparison of the interacting factors for hnRNA, mRNA, and mRNA transport of these two *COL1A1* variants may help elucidate the phenotypic differences.

In summary, the *COL1A1*-c.700delG variant described here in a proband and her mother, if translated, would produce a truncated form of collagen α_1 chain without the two essential helical regions (691–893 and 910–964) found in the wild-type form. Because of the mild OI type I phenotype, we propose that the novel *COL1A1*-c.700delG variant likely leads to haploinsufficiency of the collagen α_1 chain and that this frameshift variant is the molecular basis of OI type I in this Han Chinese family. Its identification adds further insight to *COL1A1* mutations in the Han Chinese population with OI type I.

Acknowledgments

The authors thank Dr. Weibo Xia (Department of Endocrinology, Key Laboratory of Endocrinology, Ministry of Health, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing, China) who provided helpful discussion. The work was supported

by the special funds from the Beijing Natural Science Foundation of China (No. 7072073) and the Nursery Fund of General Hospital of Chinese PLA (No. 06MP01).

a)

CCU GGU CCC CAA GGC UUC CAA GGU CCC CCU GGU GAG CCU GGC GAG
CCU GGA GCU UCA GGU CCC AUG GGU CCC CGA GGU CCC CCA GGU CCC
CCU GGA AAG AAU GGA GAU GAU ²³⁴ **GGG** **GAA** GCU GGA AAA CCU GGU CGU
CCU GGU GAG CGU GGG CCU CCU GGG CCU CAG GGU GCU CGA GGA UUG
CCC GGA ACA GCU GGC CUC CCU GGA ²⁶⁴ **AUG** AAG GGA CAC AGA GGU UUC

b)

CCU GGU CCC CAA GGC UUC CAA GGU CCC CCU GGU GAG CCU GGC GAG
CCU GGA GCU UCA GGU CCC AUG GGU CCC CGA GGU CCC CCA GGU CCC
CCU GGA AAG AAU GGA GAU GAU ²³⁴ **GGG** **AAG** CUG GAA AAC CUG GUC GUC
CUG GUG AGC GUG GGC CUC CUG GGC CUC AGG GUG CUC GAG GAU UGC
CCG GAA CAG CUG GCC UCC CUG GAA ²⁶⁴ **UGA**

Figura 3: Comparison of mRNA sequences from the wild-type control and proband. The mRNA sequence is shown from nucleotide 586 (codon 196) in exon 9 to nucleotide 810 (codon 270) in exon 10 with nucleotide 1 defined as the A of the AUG/ATG translational start codon. The splicing of introns between exon 9 and 10 removed the first G of a five-G sequence in the parental cDNA, providing four Gs from nucleotide c.697–700, whereas the proband mutant COL1A1 mRNA only contained a 3-G sequence. This c.700delG genotype induced a frameshift mutation in the proband and her mother. (A) Wild-type COL1A1 mRNA. (B) Proband COL1A1 mRNA with the 700delG variant.

References

- Barnes AM, Chang W, Morello R, Cabral WA, Weis M, Eyre DR, Leikin S, Makareeva E, Kuznetsova N, Uveges TE, *et al.* (2006) Deficiency of cartilage-associated protein in recessive lethal osteogenesis imperfecta. *N Engl J Med* 355:2757–2764.
- Barnes AM, Cabral WA, Weis M, Makareeva E, Mertz EL, Leikin S, Eyre D, Trujillo C and Marini JC (2012) Absence of FKBP10 in recessive type XI osteogenesis imperfecta leads to diminished collagen cross-linking and reduced collagen deposition in extracellular matrix. *Hum Mutat* 33:1589–1598.
- _____. (2011) Genotypephenotype correlations in autosomal dominant osteogenesis imperfecta. *J Osteoporos* 2011:540178.