ORIGINAL ARTICLE

Molecular Genetics & Genomic Medicine — Open Access — **WILEY**

A novel splicing pathogenic variant in *COL1A1* causing osteogenesis imperfecta (OI) type I in a Chinese family

Yaxin Han[1] iD | **Dongming Wang[2]** | **Jinli Guo[2]** | **Qiuhong Xiong[3]** iD | **Ping Li[3]** iD | **Yong-An Zhou[2]** iD | **Bin Zhao[2]**

[1]The Graduate School, Shanxi Medical University, Taiyuan, China

[2]The Second Hospital, Shanxi Medical University, Taiyuan, China

[3]Institutes of Biomedical Sciences, Shanxi University, Taiyuan, China

Correspondence
Ping Li, Institutes of Biomedical Sciences, Shanxi University, Taiyuan, China.
Email: pingli@sxu.edu.cn

Yong-An Zhou and Bin Zhao, The Second Hospital, Shanxi Medical University, Taiyuan, China.
Email: zya655903@163.com (Y. Z.) and zzbb3737@126.com (B. Z.)

Funding information
National Natural Science Foundation of China, Grant/Award Number: 31700731; Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi Province, China, Grant/Award Number: 2019L0007

Abstract

Background: Osteogenesis imperfecta (OI), a rare autosomal inheritable disorder characterized by bone fragility and skeletal deformity, is caused by pathogenic variants in genes impairing the synthesis and processing of extracellular matrix protein collagen type I. With the use of next-generation sequencing and panels approaches, an increasing number of OI patients can be confirmed and new pathogenic variants can be discovered. This study sought to identify pathogenic gene variants in a Chinese family with OI I.

Methods: Whole-exome sequencing was used to identify pathogenic variants in the proband, which is confirmed by Sanger sequencing and cosegregation analysis; MES, HSF, and Spliceman were used to analyze this splicing variant；qRT-PCR was performed to identify the mRNA expression level of *COL1A1* in patient peripheral blood samples; Minigene splicing assay was performed to mimic the splicing process of *COL1A1* variants in vitro; Analysis of evolutionary conservation of amino acid residues and structure prediction of the mutant protein.

Results: A novel splicing pathogenic variant (c.3814+1G>T) was identified in this OI family by using whole-exome sequencing, Sanger sequencing, and cosegregation analysis. Sequencing of RT-PCR products from the *COL1A1* minigene variant reveals a 132-nucleotide (nt) insertion exists at the junction between exons 48 and exon 49 of the *COL1A1* cDNA. Splicing assay indicates that the mutated minigene produces an alternatively spliced transcript which may cause a frameshift resulting in early termination of protein expression. The molecular analysis suggested that the altered amino acid is located at the C-terminus of type I procollagen.

Conclusion: Our study reveals the pathogenesis of a novel *COL1A1* splicing pathogenic variant c.3814+1G>T in a Chinese family with OI I.

KEYWORDS

c.3814+1G>T, *COL1A1*, novel splicing pathogenic variant, osteogenesis imperfecta, whole-exome sequencing

Yaxin Han and Dongming Wang are contributed equally as first author.

1 | INTRODUCTION

Osteogenesis imperfecta (OI), also known as brittle bone disease, is a rare genetic disorder characterized by increased bone fragility, skeletal deformities, and growth deficiency (Forlino & Marini, 2016). Extraskeletal manifestations are highly variable including the blue sclera, joint hyperlaxity, dentinogenesis imperfecta (DI), and hearing loss (Pollitt et al., 2006). Based on the severity of bone fragility, clinical presentation, and radiographic features, OI was classified into four types (types I–IV; OI I, OMIM#166200; OI II, OMIM #166210; OI III, OMIM #259420; OI IV, OMIM #166220) by Silence in the late 1970s (Sillence, Senn, & Danks, 1979).

According to the Sillence's classification system, all OI types from I to IV were caused by abnormal type I collagen or quantitative changes of the collagen. Type I collagen, the main protein building block in bone and other connective tissue, is a triple helix with nonhelical terminal peptides-telopeptides. The triple helix is folded by two pro-α1(Ⅰ)-collagen chains and one pro-α2(Ⅰ)-collagen chain which is encoded by *COL1A1* (OMIM #120150) and *COL1A2* (OMIM #120160), respectively (Lu, Zhang, Wang, Ren, & Han, 2019). The structure of triple helix is Gly-X-Y repeats with glycine in every third position, and proline and hydroxyproline are often in the X and Y position of collagen (Brodsky & Eikenberry, 1982; Kelly, Tanaka, Hardt, Eikenberry, & Brodsky, 1988). So far, hundreds of OI pathogenic variants in *COL1A1* and *COL1A2* have been recruited in the Osteogenesis Imperfecta Variant Database (https://oi.gene.le.ac.uk). These pathogenic variants include nonsense, missense, frameshift, and splicing. Nonsense and frameshift pathogenic variants generate truncating protein resulting in haploinsufficiency and mild phenotypes (Forlino & Marini, 2000). But the substitution of glycine within the Gly-X-Y triplet domain of the triple helix causes synthesize incorrectly of pro-α chains and structural abnormalities resulting in a dominant-negative effect and severe phenotypes (Bodian et al., 2009; Liu et al., 2007; Zhang et al., 2012).

Osteogenesis imperfecta type I is the mildest form characterized by the blue sclera, near-normal stature, hearing loss, and dentinogenesis imperfecta, but phenotypes are variable between different OI I patients. Osteogenesis imperfecta I is caused by autosomal dominant pathogenic variants in *COL1A1* which were related to a quantitative deficiency of structurally normal type I collagen (Willing, Pruchno, & Byers, 1993). *COL1A1* (located at position17q21.33, OMIM #120150, HGNC 2197, NCBI reference sequence NM_000088.4) contains 51 exons encoding pro-α1 chain of type I collagen.

Splicing variants are less commonly reported than missense/nonsense variant types (Marini et al., 2007; Stenson et al., 2017). Although many *COL1A1* splicing pathogenic variants have been described in previous case reports (Peng et al., 2012; Xia et al., 2008; Zhao, Yang, Guo, & Li, 2014), the pathological mechanism was not well studied. Introns removing from pre-mRNA and exons bonding are significantly important in gene expression. These reactions were catalyzed by spliceosome depending on corrected splicing site recognition (Madhani & Guthrie, 1994; Nilsen, 1994). Most introns have a 5′ splice site beginning with GT and a 3′ splice site ending with AG, and some introns have distinct splice site consensus sequences and exhibit either AT-AC termini or AT-AG termini (Patel & Steitz, 2003).

In this study, we identified a novel *COL1A1* splicing pathogenic variant c.3814+1G>T with the first nucleotide of intron 48 is changed from G to T which alters the splice site sequence in an OI I patient. Therefore, we aimed to characterize the molecular effect and mechanism of the *COL1A1* GT-AG intron pathogenic variant that causes OI I.

2 | MATERIALS AND METHODS

2.1 | Ethical compliance

This study was approved by the ethics committee of the Second Hospital of Shanxi Medical University.

2.2 | Subjects

In this study, a Chinese family with OI I is recruited and all individuals have given the written informed consent. The proband accepted clinical and physical examinations and provided his medical records. Based on all medical records, clinical presentations, and gene diagnostics, the patient was diagnosed as OI I in the hospital.

2.3 | Whole-exome sequencing (WES) and variant confirmation

Whole-exome sequencing test was performed on the blood of the proband (III-1). Library preparation and sequencing were performed using Agilent SureSelect Human SureSelect All Exon V5 kit and Illumina HiSeq 2500. The read length at the end of the pair reaches 150 bp and the sequencing depth of the capture region is approximately 100×, sequencing covered also the intron-exon junction. The sequencing results were compared with the human genome reference sequence (UCSC and NCBI database). To identified single nucleotide variants (SNVs), multiple database annotations (such as dbSNP, OMIM, ESP, Clinvar, 1,000 Genomes, and HGMD) were performed simultaneously, and Sanger sequencing was used to confirm the variant.

2.4 | Sequencing analysis

The coding region and exon-intron splice junctions including the variant of *COL1A1* were analyzed using PCR amplification in combination with Sanger sequencing using the primers (Fw:5′-TCAAGAGAAGGCTCACGATGG-3′; Rv:5′-GGGGTTCTTGCTGATGTACCA-3′). The length of the product is 372bp included exon 48, intron 48, and exon 49. Genomic DNA was extracted from peripheral blood samples and PCR products were purified. These methods were similar to those we used previously (Li, Zhang, et al., 2019).

2.5 | qRT-PCR analysis

We used the qRT-PCR to evaluate the transcript variants of *COL1A1* in the blood cells from patients and healthy relatives were similar to those we used previously (Li, Zhang, et al., 2019). The primer pair for qRT-PCR of *COL1A1* were used: Fw: 5′-GAGAGTACTGGATTGAC-3′ and Rv: 5′-ACTGGAATCCATCGGTCATG-3′. The length of the product is 191 bp including exon 49 of the total (WT + mutated) *COL1A1* (NM_000088.4). For normalization, the expression of *GAPDH* (Fw: CGGAGTCAACGGATTTGGTCGTAT; Rv:AGCCTTCTCCATGGTGGTGAAGAC).

2.6 | Prediction of splicing

To assess the presumptive effect of this variant on splicing, three different in silico prediction tools: MaxEntScan (MES) using Maximum entropy score (http://genes.mit.edu/burgelab/maxent/Xmaxentscanscoreseq.html), Human Splicing Finder (HSF, http://www.umd.be/HSF/), and Spliceman (http://fairbrother.biomed.brown.edu/spliceman/) were used.

The output score of each tool indicates the splicing signal strength: the higher the value, the higher the possibility of being a splicing site in vivo. The following equation was used to calculate the percent score variation (%): [(WT score − Mut score)]/WT score × 100 (Miao et al., 2017). The magnitude of the variation corresponds to the increase in the likelihood of mis-splicing. Only the changes greater than 15% (yes/no) were considered to be deleterious significantly (Sharma et al., 2014).

2.7 | Cell culture

HEK293T cells were grown in DMEM supplemented with 10% (v/v) FBS, 100 U/ml penicillin, and 100 mg/ml streptomycin at 37°C and 5% CO_2. Transfection was performed using the Polyetherimide (PEI) (PolyScience, Cat.No. 23966-2)

according to the manufacturer's instructions. Transfected cells were incubated for 24–48 hr posttransfection.

2.8 | The minigene constructs and Western blotting

Minigene constructs encompassing exon 48, intron 48, and exon 49 were amplified using the primer pair of *COL1A1* E48-in49-E49 Fw: 5′-GCGCTCGAGGACACCACCCTC AAGAGCCTGA-3′ and E48-in49-E49 Rv: 5′-GCGCTCGA GTACACGCAGGTCTCACCAGTCTC-3′. The amplified minigene products were cloned into the pEGFP-C3 cloning vector at the Xho I and EcoR I sites. The complete sequences of the minigene constructs were validated by Sanger sequencing. Minigene constructs were transiently transfected in HEK293T cells with polyetherimide (PEI) (PolyScience, Cat.No. 23966-2) and RT-PCR was performed with the same primers. RT-PCR products were separated by electrophoresis analysis and isoforms were identified by Sanger sequencing. Besides, we used Western blotting to analyze the proteins of minigene. These methods were similar to those we used previously (Li, Zhang, et al., 2019).

2.9 | Analysis of evolutionary conservation of amino acid residues and structure prediction of the mutant protein

NCBI Protein blast was used to analyze the conservation of amino acid residue alteration by comparing across different species. The homology modeling programs Swiss-Model was used to develop an appropriate model to mimic the effects of the mutated region. The structures were displayed by PDB-Viewer software. These methods were similar to those we used previously (Li, Zhang, et al., 2019).

3 | RESULTS

3.1 | Family pedigree/patient information

This study included four affected individuals and six unaffected relatives from a Chinese family clinically characterized as type I OI. The pedigree of the family is shown in Figure . The proband (Patient 1, III-1), is a 30-year-old Chinese male, 175 cm tall, and 80 kg in weight. His clinical features include blue sclera, mildly excessive dorsiflexion of fingers, and five fractures. Hearing, teeth, and vision are normal. As he was concerned about his next generation and how this disease was inherited, a directed genetic study and genetic counseling were performed. All patients (II-2, II-5, I-2, and III-1) presented with blue sclerae, DI, and multiple
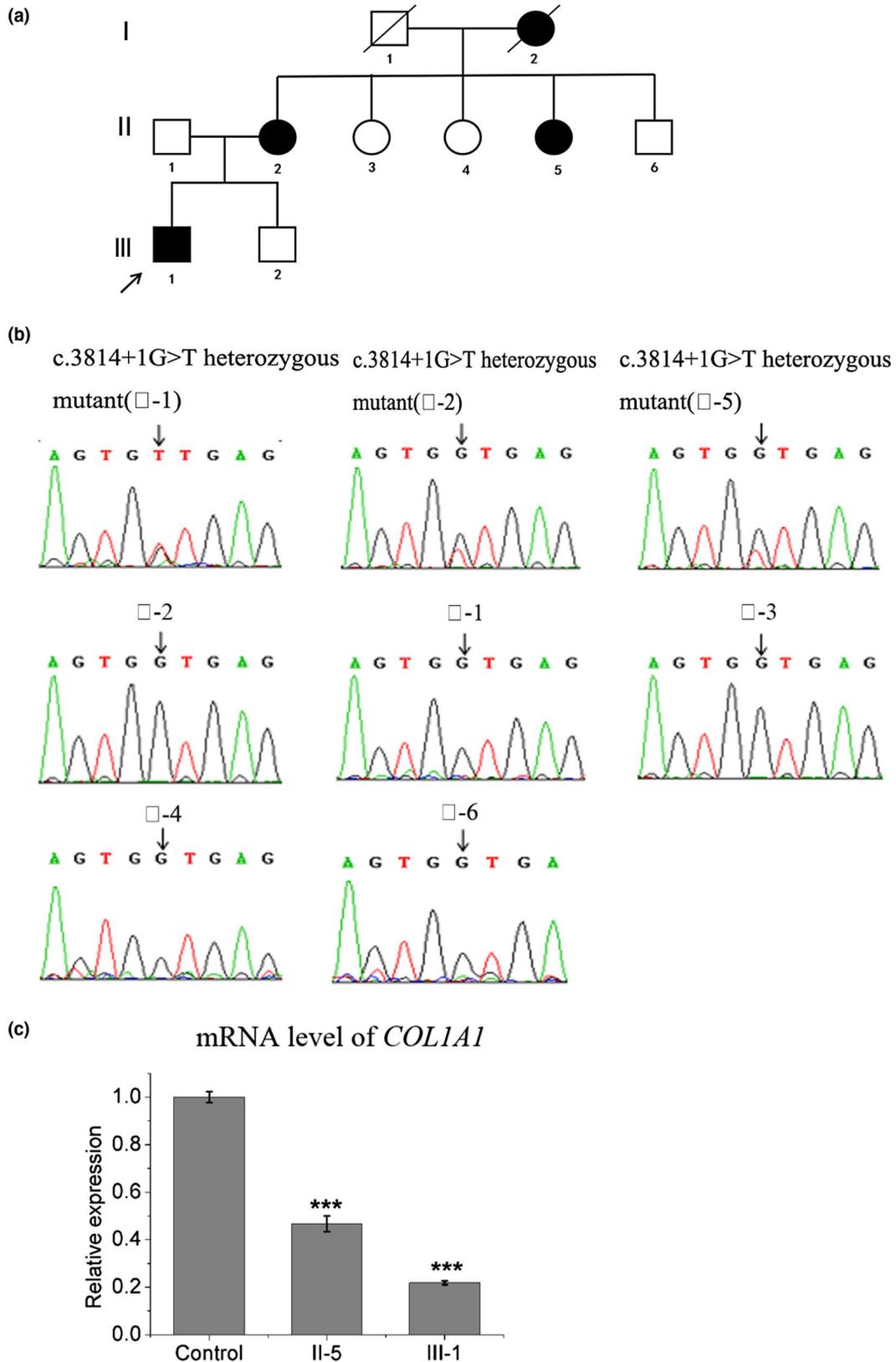
FIGURE 1 Pedigree and sequencing of this family (a) The pedigree of the family. Affected individuals are indicated by filled black, the proband is pointed by an arrow. (b) Sanger sequencing analysis performed on the genome DNA from this family. The gene variation is shown by black arrow. (c) qRT-PCR analysis performed on total (WT + mutated) RNA obtained from blood samples of patients (III-1 and II-5) and healthy relative (II-4). Level were normalized to the amount of GAPDH. Data represent the mean ± SE of independent measurements performed in triplicate(***$p < .001$)

bone fractures. The clinical features of the pedigree are summarized in Table 1. All patients have not accepted bisphosphonates or any other therapies.

3.2 | Sequencing analysis

Peripheral blood samples were collected from this family including three patients (Ⅱ-2, Ⅱ-5, Ⅲ-1) and five healthy subjects (Ⅱ-1, Ⅱ-3, Ⅱ-4, Ⅱ-6, Ⅲ-2).

Given the proband's requirements, genetic counseling and WES was conducted. After bioinformatics analysis, a novel *COL1A1* splicing variant c.3814+1G>T was detected which was confirmed by Sanger sequencing.

To investigate whether this variant cosegregated in this family, Sanger sequencing of genomic DNA isolated from the peripheral blood samples in this family was performed for the *COL1A1*. The results showed that three patients harbor the heterozygous *COL1A1* variant c.3814+1G>T comparing to the five unaffected relatives (Figure 1b). The results revealed that the proband (Ⅲ-1) inherited the described *COL1A1* variant (c.3814+1G>T) from his mother (Ⅱ-2) and this variant separated with phenotypes in this family.

To identify the mRNA expression level of *COL1A1* in patient samples, qRT-PCR was performed. Total RNA was isolated from the blood samples of the patients (Ⅲ-1 and Ⅱ-5) and health subject (Ⅱ-3) in this family, qRT-PCR was normalized to that of *GAPDH*. The results showed that the level of *COL1A1* mRNA was decreased to 50% (patient: II-5) or 35% (patient: III-1) in the samples from patients compared to control (Figure 1c). These results indicate that the mRNA of mutant *COL1A1* is not stable in the cells thus lead to a lower expression of the protein, and implies that an insufficient normal *COL1A1* in the patient.

3.3 | Prediction of splicing and analysis of evolutionary conservation of amino acid residues

The effect of the c.3814+1G>T variant was computationally analyzed using three splice site prediction programs: MES,

HSF, and Spliceman. The outcomes were summarized in Table 2.

The analysis of amino acid residues showed that these lost amino acid residues were highly evolutionary conserved among *COL1A1* proteins from different species (Figure 2b), the mutant protein results in a truncated protein lacking the C-terminal (Figure 2c).

3.4 | Splicing analysis of *COL1A1* c.3814+1G>T in the minigene

The minigene analysis was performed to identify abnormal splicing. We constructed a minigene containing part of exon 48, exon 49, and entire intron 48 with G or T at c.3814+1 of *COL1A1* (Figure 3a). Electrophoresis analysis of RT-PCR products shows that no different band size between untransfected and wildtype minigene transfected HEK293T cells which were identified as *COL1A1* exon 48 (121 bp) and exon 49 (88 bp) in both, but a dissimilar band was detected in mutant minigene transfected cells. Sanger sequencing of RT-PCR products reveals the wild type construct is a normal size band of 209 bp which is consistent with *COL1A1* exon 48 (121 bp) and exon 49 (88 bp). The c.3814+1G>T substitution resulted in a band of 341 bp which contains 132 nucleotides from intron 48 (Figure 3b). Third, Western blot analysis was performed to study the protein expression of the minigenes. In pEGFP-C3 vector-transfected HEK293T cells, a strong band was observed with the molecular weight of about 27 kD which represents the GFP protein alone (Figure 3c). In pEGFP-*COL1A1*-minigene wild type transfected cells, a band with a molecular weight at ~34 kD was detected which was identified as the protein products including part of exon 48 and exon 49 fusion with GFP (Figure 3c). But, in pEGFP-*COL1A1* minigene mutant transfected cells, one extra band at ~31 kD molecular weight was observed compared to wild type *COL1A1* minigene transfected cells, suggesting that the c.3814+1G>T substitution caused a frameshift resulting in early termination of protein expression which is corresponding to the results of alternative transcript analysis (Figure 3c).

TABLE 1 Clinical characteristics of family members with OI I

Subject	III-1	II-2	II-5	I-1 (Died)
Sex	Male	Female	Female	Female
Age (y)	30	51	40	68
Height (cm)	175	155	153	Unknown
Weight (kg)	80	45	42	Unknown
Sclerae	Blue	Blue	Blue	Blue
Fractures	Five	Six	Multiple	Multiple
Dentinogenesis imperfecta	Yes	Yes	Yes	Yes
Hearing loss	No	Yes	No	Yes
Stature	Normal	Normal	Normal	Unknown

TABLE 2 Bioinformatics prediction of splice site variants by HSF, MaxEntScan, and spliceman

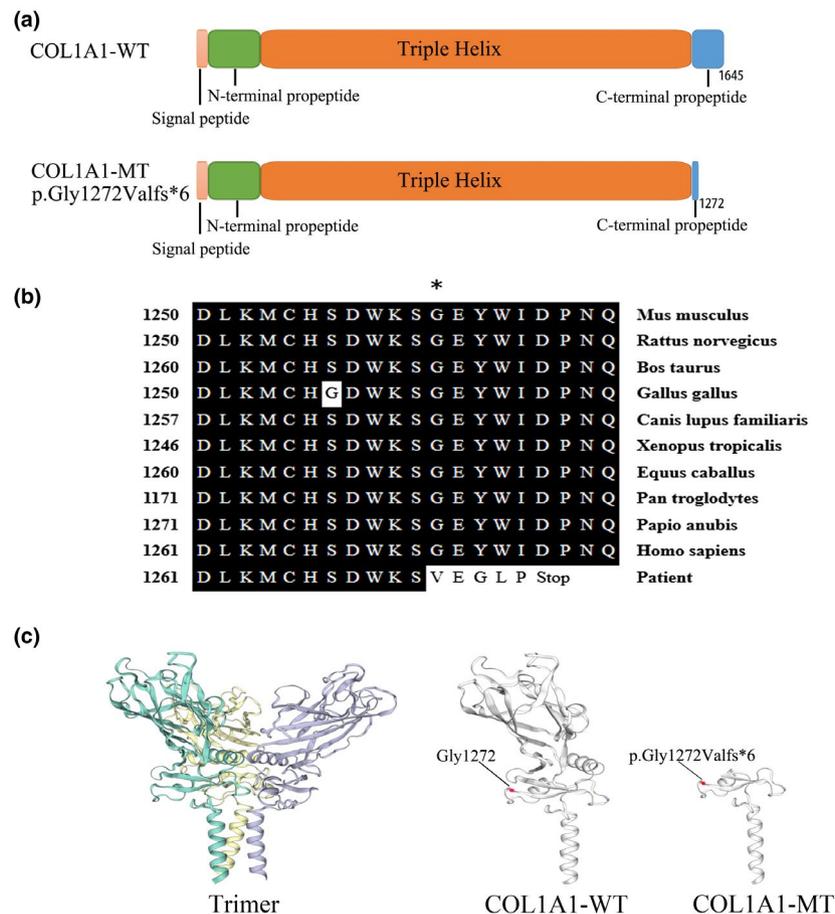| Splicing prediction tool	Prediction score values		
	HSF(%)	MaxEnt scan(%)	Spliceman
COL1A1 c.3814+1G>T	0.90/0.64 (30)	6.13/−2.38 (138)	YES:66

FIGURE 2 Analysis of COL1A1 mutation. (a) Schematics of the COL1A1 gene and the secondary structure of the proα1 protein. (b) Evolutionary cservation of amino acid residues altered by c.3814+1G>T (p.Gly1272Valfs*6) across different species. NCBI accesssion numbers are: Mus musculus: NP_031768.2; Rattus norvegicus: NP_445756.1; Bos taurus: NP_001029211.1; Gallus gallus: XP_024999899.1; Canis lupus familiaris: NP_001003090.1; Xenopus tropicalis: NP_001011005.1; Equus caballus: XP_023508478.1; Pan troglodytes: XP_009431085.1; Papio anubis: XP_021784688.1; Homo sapiens: NP_000079.2. (c) Left side: The structure of the type Ι procollagen C-propeptide trimer; Right side: The structure of proα1(I)-C-propeptide (WT and MT). The mutant protein p.Gly1272Valfs*6 was predicted by Swiss-Modle online, which alters the structure of preα1 C-propeptide compared to the wild type. The WT and MT monomer are shown in gray. The altered amino acid is shown as red point

According to the American College of Medical Genetics and Genomics (ACMG) guidelines for interpretation of sequence variants (Richards et al., 2015), we consider this variant to be pathogenic (PM2 + PVS1+PP1 + PP4).

4 | DISCUSSION

Pathogenic variants in type I collagen-encoded genes (*COL1A1* and *COL1A2*) or other genes (*CRTAP*: MIM#605497, *WNT1*: MIM#164820, and *BMP1*: MIM#112264 et, al) responsible

for OI. Osteogenesis imperfecta patients can be confirmed and new pathogenic mutations can be discovered by the genetic tests which are significantly important for us to analyze the relationship between the genotype and phenotype of OI comprehensively. Whole-exome sequencing is a currently available effective approach for screening pathogenic variants of OI (Mackenroth et al., 2016). We use WES to detect a pathogenic gene variant which confirmed by Sanger sequencing and perform a series of molecular biology methods to study the pathogenic mechanism of the variant. A novel splicing variant (NM_ 000088.4: c.3814+1G>T;
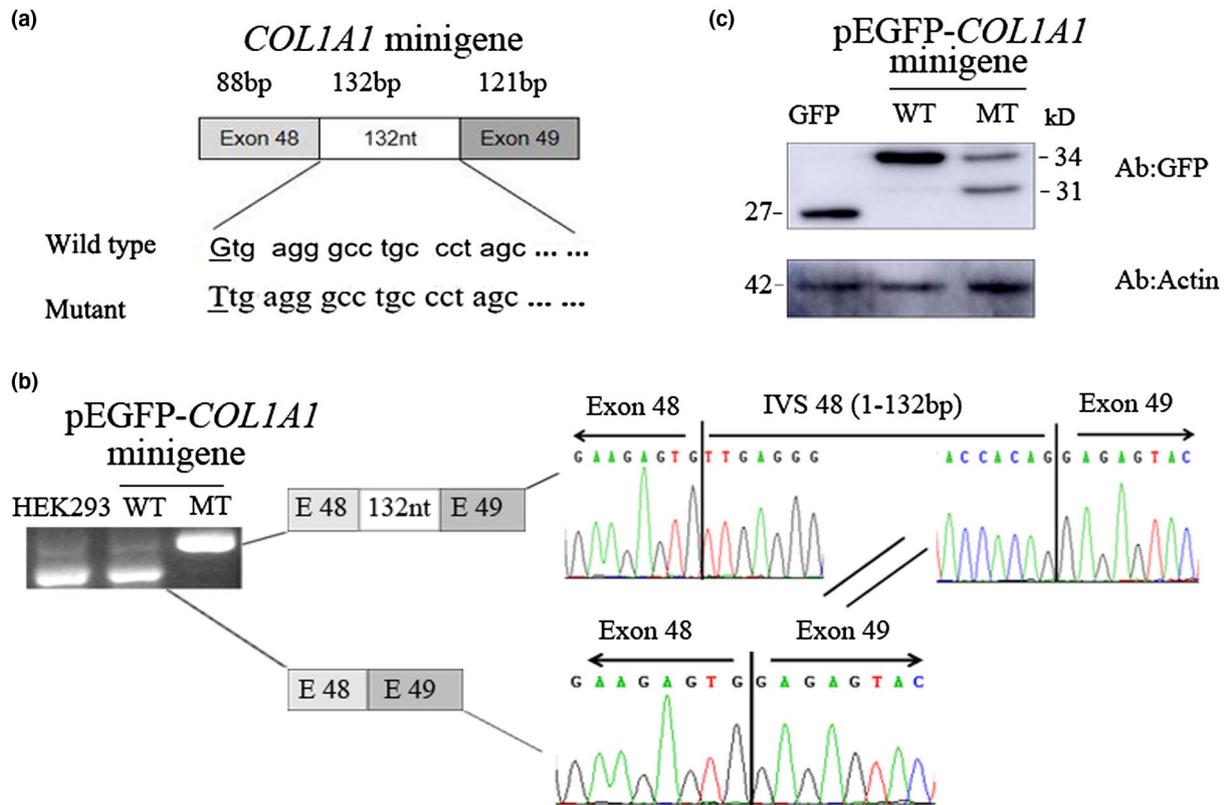
(a)

COL1A1 minigene

88bp   132bp   121bp

| Exon 48 | 132nt | Exon 49 |

Wild type   Gtg agg gcc tgc cct agc … …

Mutant   Ttg agg gcc tgc cct agc … …

(c)

pEGFP-*COL1A1* minigene

(b)

pEGFP-*COL1A1* minigene

FIGURE 3 Minigene constructs for splicing pattern investigation. (a) Schematic representation of the COL1A1 minigene construct, extending from exon 48 to exon 49. The nucleotide sequences of the wild type (WT) and mutant (MT) minigenes are shown below. The changed nucleotide in intro 48 of both wild type and mutant COL1A1 minigene are labeled by underline. (b) Minigene splicing assay. COL1A1 minigenes harboring the wild type or mutant COL1A1 pseudoexon were transiently transfected into HEK293T cells. After RNA isolation the splicing products were analyzed by RT-PCR. The lower bands represent correctly spliced exons, whereas the higher band represent the COL1A1 pseudoexon inserted 132 nt from intro 48 between minigene exons 48 and 49. (c) Western blot analysis for the expression of the wild type and mutant pEGFP-COL1A1-minigenes. Whole cell lysates were separated by SDS—PAGE (12% acrylamide). GFP monoclonal antibodies were used

NP_000079.2: p.Gly1272Valfs*6) in *COL1A1* was identified in this Chinese family with OI I. Three patients (Ⅱ-2, Ⅱ-5, and Ⅲ-1) carried the novel splicing variant c.3814+1G>T. Five healthy family members (Ⅱ-1, Ⅱ-3, Ⅱ-4, Ⅱ-6, and Ⅲ-2) were free of this variant. These results indicate that the novel splicing variant c.3814+1G>T is cosegregated with the disease in the pedigree.

Given the variant is located at the boundary of exon 48 and intron 48, and alternative splicing depended on the flanking intronic regions recognized correctly (Patel & Steitz, 2003). Owing to pathogenic variants of splice sites could have led to splicing defects (Scotti & Swanson, 2016). To further investigate whether this mutation affects the alternative splicing or not, minigene was performed to mimic the process of splicing in vitro. The minigene results suggest that the novel splicing variant c.3814+1G>T contributed to the disappearance of the 5′ donor splice site of intron 48, which leads to intron 48 retention. This in-frame insertion created a premature termination stop codon (PTC). Mutations that created premature termination codons (PTCs) can cause haploinsufficiency via nonsense-medicated decay (NMD) (Maquat, 1995).

It has been reported that variants at splicing sites can cause OI (Schleit et al., 2015; Schwarze, Starman, & Byers, 1999). These variants included the typical splicing variants which were located at splicing donor/acceptor sites in introns (Schleit et al., 2015), and the atypical splicing sites beyond the splicing sites (GT-AG) (Li, Cao, et al., 2019). A previous study identified a *COL1A1* C-propeptide pathogenic variant (one out-of-frame splicing error), generating a PTC and causing haploinsufficiency via nonsense-medicated decay (NMD), was associated with OI I (Symoens et al., 2014). Synthesizing of single *COL1A1* allele results in haploinsufficiency which involves nonsense-mediated mRNA decay or frameshift/splicing mutation-induced pretermination codon, most of them being mild OI type (Rauch, Lalic, Roughley, & Glorieux, 2010). In our case, the *COL1A1* novel splicing variant c.3814+1G>T (p.Gly1272Valfs*6) is located at the C-terminal propeptide and generated a premature termination codon (PTC). The level of *COL1A1* mRNA was decreased in patients (III-1, II-5), so this PTC leads to nonsense-mediated mRNA decay (NMD), by which mRNAs are degraded before they produce large quantities of truncated proteins. The

NMD process contributes to the *COL1A1* haploinsufficiency which may be the major mechanism underlying the OI I in this family.

In conclusion, we report a novel *COL1A1* intronic pathogenic variant c.3814+1G>T alters the alternative splice of *COL1A1* and consequently caused Type OI I. Our study validates that c.3814+1G>T is a type OI I-causative mutation and data in this study enrich mutation OI database and provide an OI causative mutation for molecular diagnosis as well as scientific information.

CONFLICT OF INTEREST
Yaxin Han, Dongming Wang, Jinli Guo, Qiuhong Xiong, Ping Li, Yong-An Zhou and Bin Zhao declare that they have no conflict of interest.

AUTHOR CONTRIBUTIONS
YH wrote the manuscript and performed the practical work. DW collected patient's data. QX and GJ analyzed the patient's data and were assisted by DW. PL designed the study. BZ and YA-Z conceived the study and edited the manuscript.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID
Yaxin Han https://orcid.org/0000-0001-6340-6225
Qiuhong Xiong https://orcid.org/0000-0003-1531-7353
Ping Li https://orcid.org/0000-0002-6137-6154
Yong-An Zhou https://orcid.org/0000-0001-5015-7688

REFERENCES
Bodian, D. L., Chan, T.-F., Poon, A., Schwarze, U., Yang, K., Byers, P. H., … Klein, T. E. (2009). Mutation and polymorphism spectrum in osteogenesis imperfecta type II: Implications for genotype-phenotype relationships. *Human Molecular Genetics*, *18*(3), 463–471. https://doi.org/10.1093/hmg/ddn374

Brodsky, B. & Eikenberry, E. F. (1982). Characterization of fibrous forms of collagen. *Methods in Enzymology*, *82*, 127–174. https://doi.org/10.1016/0076-6879(82)82062-4

Forlino, A., & Marini, J. C. (2000). Osteogenesis imperfecta: Prospects for molecular therapeutics. *Molecular Genetics and Metabolism*, *71*(1–2), 225–232. https://doi.org/10.1006/mgme.2000.3039

Forlino, A. & Marini, J. C. (2016). Osteogenesis imperfecta. *Lancet (London, England)*, *387*(10028), 1657–1671. https://doi.org/10.1016/S0140-6736(15)00728-X

Kelly, J., Tanaka, S., Hardt, T., Eikenberry, E. F., & Brodsky, B.(1988). Fibril-forming collagens in lamprey. *The Journal of Biological Chemistry*, *263*(2), 980–987.

Li, L., Cao, Y., Zhao, F., Mao, B., Ren, X., Wang, Y., … Zhao, X. (2019). Validation and classification of atypical splicing variants associated with osteogenesis imperfecta. *Frontiers in Genetics*, *10*, 979. https://doi.org/10.3389/fgene.2019.00979

Li, P., Zhang, L., Zhao, N., Xiong, Q., Zhou, Y. A., Wu, C., & Xiao, H. (2019). A novel alpha-galactosidase a splicing mutation predisposes to fabry disease. *Frontiers in Genetics*, *10*, 60. https://doi.org/10.3389/fgene.2019.00060

Liu, W., Gu, F., Ji, J., Lu, D., Li, X., & Ma, X. (2007). A novel COL1A1 nonsense mutation causing osteogenesis imperfecta in a Chinese family. *Molecular Vision*, *13*, 360–365.

Lu, Y., Zhang, S., Wang, Y., Ren, X., & Han, J. (2019). Molecular mechanisms and clinical manifestations of rare genetic disorders associated with type I collagen. *Intractable & Rare Diseases Research*, *8*(2), 98–107. https://doi.org/10.5582/irdr.2019.01064

Mackenroth, L., Fischer-Zirnsak, B., Egerer, J., Hecht, J., Kallinich, T., Stenzel, W., … Gerhold, K.(2016). An overlapping phenotype of Osteogenesis imperfecta and Ehlers-Danlos syndrome due to a heterozygous mutation in COL1A1 and biallelic missense variants in TNXB identified by whole exome sequencing. *American Journal of Medical Genetics. Part A*, *170*, 1080–1085. https://doi.org/10.1002/ajmg.a.37547

Madhani, H. D., & Guthrie, C. (1994). Dynamic RNA-RNA interactions in the spliceosome. *Annual Review of Genetics*, *28*, 1–26. https://doi.org/10.1146/annurev.ge.28.120194.000245

Maquat, L. E. (1995). When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. *RNA*, *1*(5), 453–465.

Marini, J. C., Forlino, A., Cabral, W. A., Barnes, A. M., San Antonio, J. D., Milgrom, S., … Byers, P. H. (2007). Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: Regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans. *Human Mutation*, *28*(3), 209–221. https://doi.org/10.1002/humu.20429

Miao, Q., Ping, X., Tang, X., Zhang, L., Zhang, X., Cheng, Y., & Shentu, X. (2017). Experimental assessment of novel PAX6 splicing mutations in two Chinese families with aniridia. *Gene*, *630*, 44–48. https://doi.org/10.1016/j.gene.2017.07.073

Nilsen, T. W. (1994). RNA-RNA interactions in the spliceosome: Unraveling the ties that bind. *Cell*, *78*(1), 1–4. https://doi.org/10.1016/0092-8674(94)90563-0

Patel, A. A., & Steitz, J. A. (2003). Splicing double: Insights from the second spliceosome. *Nature Reviews Molecular Cell Biology*, *4*(12), 960–970. https://doi.org/10.1038/nrm1259

Peng, H., Zhang, Y., Long, Z., Zhao, D., Guo, Z., Xue, J., … Hu, Z. (2012). A novel splicing mutation in COL1A1 gene caused type I osteogenesis imperfecta in a Chinese family. *Gene*, *502*(2), 168–171. https://doi.org/10.1016/j.gene.2012.04.023

Pollitt, R., McMahon, R., Nunn, J., Bamford, R., Afifi, A., Bishop, N., & Dalton, A. (2006). Mutation analysis of COL1A1 and COL1A2 in patients diagnosed with osteogenesis imperfecta type I-IV. *Human Mutation*, *27*(7), 716.

Rauch, F., Lalic, L., Roughley, P., & Glorieux, F. H. (2010). Relationship between genotype and skeletal phenotype in children and adolescents with osteogenesis imperfecta. *Journal of Bone and Mineral Research*, *25*(6), 1367–1374. https://doi.org/10.1359/jbmr.091109

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., … Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the

Association for Molecular Pathology. *Genetics in Medicine*, *17*(5), 405–424. https://doi.org/10.1038/gim.2015.30

Schleit, J., Bailey, S. S., Tran, T., Chen, D., Stowers, S., Schwarze, U., & Byers, P. H. (2015). Molecular outcome, prediction, and clinical consequences of splice variants in COL1A1, which encodes the proalpha1(I) chains of type I procollagen. *Human Mutation*, *36*(7), 728–739. https://doi.org/10.1002/humu.22812

Schwarze, U., Starman, B. J., & Byers, P. H. (1999). Redefinition of exon 7 in the COL1A1 gene of type I collagen by an intron 8 splice-donor-site mutation in a form of osteogenesis imperfecta: Influence of intron splice order on outcome of splice-site mutation. *American Journal of Human Genetics*, *65*(2), 336–344. https://doi.org/10.1086/302512

Scotti, M. M., & Swanson, M. S. (2016). RNA mis-splicing in disease. *Nature Reviews Genetics*, *17*(1), 19–32. https://doi.org/10.1038/nrg.2015.3

Sharma, N., Sosnay, P. R., Ramalho, A. S., Douville, C., Franca, A., Gottschalk, L. B., … Cutting, G. R. (2014). Experimental assessment of splicing variants using expression minigenes and comparison with in silico predictions. *Human Mutation*, *35*(10), 1249–1259. https://doi.org/10.1002/humu.22624

Sillence, D. O., Senn, A., & Danks, D. M. (1979). Genetic heterogeneity in osteogenesis imperfecta. *Journal of Medical Genetics*, *16*(2), 101–116. https://doi.org/10.1136/jmg.16.2.101

Stenson, P. D., Mort, M., Ball, E. V., Evans, K., Hayden, M., Heywood, S., … Cooper, D. N. (2017). The Human Gene Mutation Database: Towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Human Genetics*, *136*(6), 665–677. https://doi.org/10.1007/s00439-017-1779-6

Symoens, S., Hulmes, D. J., Bourhis, J. M., Coucke, P. J., De Paepe, A., & Malfait, F. (2014). Type I procollagen C-propeptide defects: Study of genotype-phenotype correlation and predictive role of crystal structure. *Human Mutation*, *35*(11), 1330–1341. https://doi.org/10.1002/humu.22677

Willing, M. C., Pruchno, C. J., & Byers, P. H. (1993). Molecular heterogeneity in osteogenesis imperfecta type I. *American Journal of Medical Genetics*, *45*(2), 223–227. https://doi.org/10.1002/ajmg.1320450214

Xia, X.-Y., Cui, Y.-X., Huang, Y.-F., Pan, L.-J., Yang, B., Wang, H.-Y., … Zhou, Y.-C. (2008). A novel RNA-splicing mutation in COL1A1 gene causing osteogenesis imperfecta type I in a Chinese family. *Clinica Chimica Acta*, *398*(1–2), 148–151. https://doi.org/10.1016/j.cca.2008.07.030

Zhang, Z.-L., Zhang, H., Ke, Y.-H., Yue, H., Xiao, W.-J., Yu, J.-B., … Fu, W.-Z. (2012). The identification of novel mutations in COL1A1, COL1A2, and LEPRE1 genes in Chinese patients with osteogenesis imperfecta. *Journal of Bone and Mineral Metabolism*, *30*(1), 69–77. https://doi.org/10.1007/s00774-011-0284-6

Zhao, D., Yang, J., Guo, Z., & Li, R. (2014). A novel splicing mutation in COL1A1 gene caused type I osteogenesis imperfecta in a Chinese family. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi*, *31*(2), 189–191. https://doi.org/10.3760/cma.j.issn.1003-9406.2014.02.013