Injury 50 (2019) 2215–2219

Contents lists available at ScienceDirect

Injury

journal homepage: www.elsevier.com/locate/injury

Hybrid minigene splicing assay verifies the pathogenicity of a novel splice site variant in the *COL1A1* gene of a chinese patient with osteogenesis imperfecta type I

Yuxiang Lin[1], Xiaoli Li[1], Xinghua Huang, Dezhu Zheng, Yichu Liu, Fenghua Lan, Zhihong Wang*

Research Center for Molecular Diagnosis of Genetic Diseases, Dongfang Hospital, Xiamen University Medical College, Fuzhou, China

ARTICLE INFO

Article history:
Accepted 14 October 2019

Keywords:
Bioinformatics
Osteogenesis imperfecta type I
Hybrid minigene splicing assay
Next-generation sequencing

ABSTRACT

Background: Osteogenesis imperfecta (OI) is a rare genetic bone disease associated with brittle bones and fractures. Among all known types, OI type I is the most common type and characterized by increased bone fragility, low bone mass, distinctly blue-gray sclera, and susceptibility to conductive hearing loss beginning in adolescence. Mutations in genes encoding type I collagen (*COL1A1* and *COL1A2*) contribute to the main pathogenic mechanism of OI.
Methods: Subtle mutation of the *COL1A1* gene in the proband was detected by targeted next-generation sequencing (NGS) and confirmed by Sanger sequencing. We then assessed the effect of the mutation on the splicing of the *COL1A1* gene by bioinformatics prediction and hybrid minigene splicing assay (HMSA).
Results: A novel splice site mutation c.1821+1 *G>C* was discovered in the proband by NGS and further confirmed by Sanger sequencing, which was also simultaneously identified from the proband's mother and elder sister. Bioinformatics predicted that this mutation would result in a disappearance of the 5′ donor splice site in intron 26, thereby leading to abnormal splicing and generation of premature stop codon. The follow-up experimental data generated by HMSA was consistent with this prediction.
Conclusion: Our study identified a novel splice site mutation that caused OI type I in the proband by abnormal splicing and demonstrated that combined applications of NGS, bioinformatics and HMSA are comprehensive and effective methods for diagnosis and aberrant splicing study of OI.

Introduction

Osteogenesis imperfecta (OI), also known as brittle bone disease, is a systemic hereditary connective tissue disease. The genealogical analysis demonstrates that OI is an autosomal dominant disorder with an incidence of approximately 1:10,000 and no significant difference is found between genders [1,2]. Previous studies have shown that mutations in genes encoding type I collagen (*COL1A1* and *COL1A2*) are the primary cause of OI. Besides, *FK506, CRTAP, P3H1, PPIB, SERP1NH1, SERPF1*, and *SP7/OS* are also associated with OI progression [3,4]. Commonly, four different types, designed mild (OI type I), lethal (OI type II), severely deforming (OI type III), and moderately deforming (OI type IV), are classified basis on clinical features and disease severity of OI, in which OI type I is the most common type [3,5]. Clinical manifestations of OI type I include bone fragility, distinctly blue-gray sclera, hearing loss, and dental caries [6]. Moreover, low bone mass and multiple fractures are main manifestations in X-ray [3,6].

Sanger sequencing is an earlier laborious and time-consuming tool that detects *COL1A1/2* subtle mutations. Nowadays, next-generation sequencing (NGS) provides a new, easy and quick way to detect mutations with high throughput and sensitivity [7,8]. Here, we found a novel mutation in intron 26 of *COL1A1* by NGS, which was predicted to be likely pathogenic via changing the splice donor site. Moreover, hybrid minigene splicing assay (HMSA) further proved the molecular pathogenic mechanism of the novel splicing mutation [9,10].

Materials and methods

Subjects

The proband was a 31-years-old woman, whose height was 1.49 m. She was suspected to have OI type I based on clinical symptoms

* Corresponding author. Zhihong Wang, Research Center for Molecular Diagnosis of Genetic Diseases, Fuzhou General Hospital, 156 Xi'erhuanbei Road, Fuzhou City, Fujian Province, 350025 China
E-mail address: xmzhwang@xmu.edu.cn (Z. Wang).
[1] These authors contributed equally to this work

https://doi.org/10.1016/j.injury.2019.10.033
0020-1383/

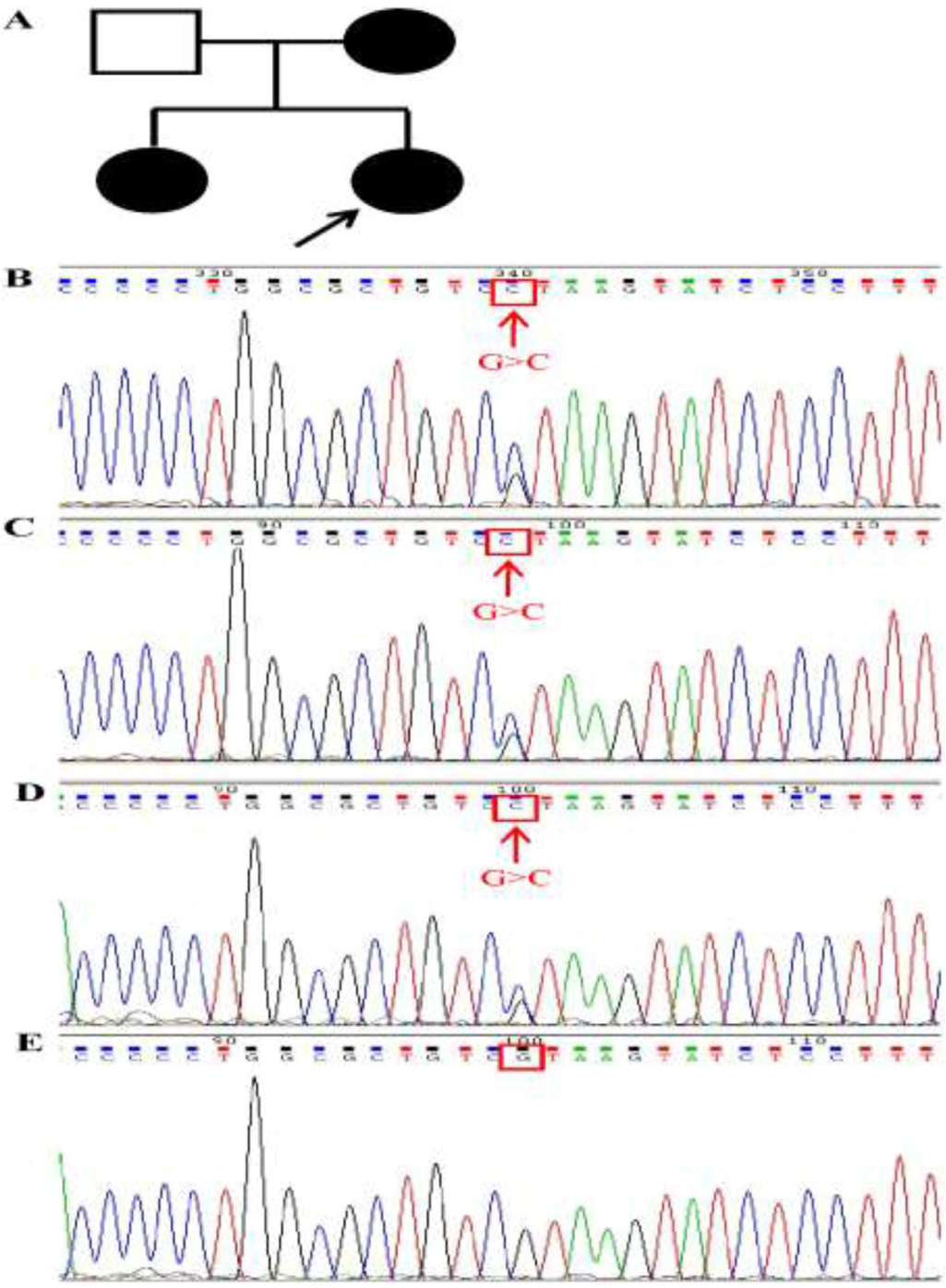

Fig. 1. Confirmation of the mutation by Sanger sequencing. (A) The pedigree chart of this family. (B)The proband had a heterozygous mutation of base G/C; (C) and (D) The proband's mother and elder sister had the same mutation of base G/C; (E) The proband's father had a homozygous base G;.

of frequent fractures, low bone mass, and dental caries. Her mother aged 60 years was suffered from splintered fracture of right femur once and severe osteoporosis after menostasis, while her elder sister aged 34 years only had low bone mass and no other typical clinical symptoms about OI, their height was almost the same as the proband (Fig. 1A). The protocols for this study were evaluated and approved by the Ethics Committee of Fuzhou General Hospital. Written informed consent was obtained from the proband.

NGS

To find a disease-causing mutation, targeted next-generation sequencing was carried out from the proband's DNA sample. Illumina Genome Analyzer Technology platform was used to test genes related to skeletal dysplasia (*COL1A1, COL1A2, SLC26A2, COMP, COL9A1, FGFR1,* etc.). Library preparation and sequencing were performed using SureSelect Human All Exome enrichment kit (Agilent, USA) and Illumina HiSeq sequencer (Illumina, USA) [11]. The sequencing data were processed using the SAMtools 0.1.16 alignment tool, and human genome NCBI database (build 38) was used as the reference [12]. The identified single nucleotide variants (SNVs)

were compared with the data from the 1000 Genomes Project and dbSNP database (build 142).

Sanger sequencing

The potential mutated base and flanking sequences of the *COL1A1* gene were amplified and sequenced by Sanger sequencing (forward primer: GAGTAAATGAGAGGCCCCAG, and reverse primer: GACCCTGCACAGAGAGAACAC). PCR conditions were as follows: 94 °C 5 min, 35 × (94 °C 30 s, 58 °C 45 s, 72 °C 30 s), 72 °C 10 min, and the PCR products are 432 bp. PCR products of normal controls (100 alleles) were analyzed by direct sequencing simultaneously to exclude the possibility of polymorphism.

Bioinformatics analysis

Bioinformatics analysis was performed to investigate the effect of the novel intronic mutation on the splicing of the *COL1A1* gene by using splicing prediction tool Human Splice Finder (http://www.umd.be/HSF3) [13].

HSMA

Construction of the minigenes

A "minigene" spanning exon 26-intron 26-exon 27 of the *COL1A1* gene was amplified from the proband's and a normal person's genome DNA using a forward primer (F14-H): CCCAAGCTTGGGGAGTAAATGAGAGGCCCCAG with the restriction site HindIII, and a reverse primer (R14-B): CGCGGATCCGCGGACCCTGCACAGAGAGAACAC with the restriction site *Bam*HI. This minigene was then cloned into pEGFP-C1 vector (Promega, USA). All these recombinant vectors were sequenced, and a mutated and a wild-type plasmid were chosen for further work.

Transfection of HeLa cells

The wild-type or mutated plasmids were transiently transfected into HeLa cells using Lipofectamine 2000 (Invitrogen, USA) according to the instructions. After transfecting and then culturing for 48 h, all cells were collected to extract total RNA using Trizol reagent (Invitrogen, USA).

Reverse transcription-polymerase chain reaction (RT-PCR) analysis

Total RNA was reverse transcribed with reverse transcriptase (Invitrogen, USA) according to the instructions. The cDNA was then amplified with the primers F14-H and R14-B. PCR fragments were analyzed by agarose gel electrophoresis and individual bands were excised and sequenced.

Results

Subtle mutation detection

The average depth for the targeted NGS region was approximately 212 ×, and the average depth for each base pair was around 118 ×. On average 98.4% of base pairs with > 30 × coverage were successfully detected, indicating high capability for variants identification. A novel mutation at the 5′ donor splice site of intron 26 (c. 1821+1 *G*>*C*) was detected in the *COL1A1* gene of the proband by NGS, which was confirmed by Sanger sequencing. DNA sequencing indicated that the proband's mother was a carrier, who coincidently transferred the mutation to her daughters, including the proband and the elder sister of the proband (Fig. 1).

Bioinformatics prediction

Human Splice Finder predicted that the 5′donor splice site of intron 26 would disappear due to the mutation, which might cause two types of aberrant splicing including intron retention and alternative donor site using (Fig. 2).

Splicing study by HMSA

The minigene vectors were constructed and the correct insertions were verified by Sanger sequencing. After transient transfection of the wild-type (WT), mutated and empty plasmid into HeLa cells for 48 h, mRNA was analyzed by RT-PCR using primers F14-H and R14-B. Normally, splicing after transcription in WT minigene containing exon 26, intron 26 and exon 27 resulted in the removal of intron 26, thereby producing a fragment of 289 bp with exon 26 and exon 27. However, aberrant splicing after transcription in mutated minigene harboring a novel mutation at the 5′ donor splice site of intron 26 (c. 1821+1 *G*>*C*) generated a product of 432 bp (Fig. 3), which was consistent with the intron 26 retention predicted by Human Splice Finder (Fig. 2). The DNA sequences of the fragments were further confirmed by Sanger sequencing (data not shown).

Discussion

Among all the gene mutational patterns that result in the occurrence of inherited diseases, pre-mRNA splicing mutation has recently attracted more and more attention. Pre-mRNA splicing is an essential step for producing the functional protein product [14,15]. The process of U1/U2-based mRNA splicing involves a variety of snRNP and correct identification of donor and acceptor splice sites [15]. By estimation, pre-mRNA splicing occurs in over 60% of human genes [16]. Any mutations that lead to exon skipping, intron inclusion, leaky splicing, or cryptic splicing, introduction of pseudo-exons into the processed mRNA will generate aberrant proteins that destroy biochemical pathways and/or interfere cell growth regulations and eventually result in clinically abnormal phenotypes [17].

Although high-throughput DNA sequencing accelerates the pace of discovery of splice-site mutations, there remain many challenges in reliably illuminating the mechanisms of how these mutations change splicing patterns, which is important for exploiting the mutation pathogenicity [18,19]. In some cases, RNA samples from affected individuals are unavailable, which certainly impedes subsequent transcriptional analysis. However, HMSA provides an effective, available and relatively simple approach to study the effect of splice-site mutations on the splicing process [20,21].

Osteogenesis imperfecta (OI) is an inherited systemic connective tissue disease with highly varied clinical consequences. Mutations in genes encoding type I collagen (*COL1A1* and *COL1A2*) are the main pathogenic mechanism, in which mutations in *COL1A1* account for 60%–70% of OI [4]. To date, over a hundred kinds of mutations have been recorded into the osteogenesis imperfecta & Ehlers-Danlos syndrome variant database of the University of Leicester, including 54% of glycine substitutions, 24% of frame-shift mutation, 15% of splicing mutation, 7% of polymorphism [3, 4]. Given the highly varied clinical consequences of OI, genetic diagnosis that examines mutations in genes associated with OI is essential for asymptomatic and atypical patients. Next-generation sequencing (NGS) is a powerful tool to screen mutations with ultra-high sensitivity, fidelity, throughput, and speed. In the present study, we used targeted NGS panel for skeletal dysplasia to detect subtle variants of OI-related genes and identified a splicing variant. Subsequent RNA-based studies were performed to assess

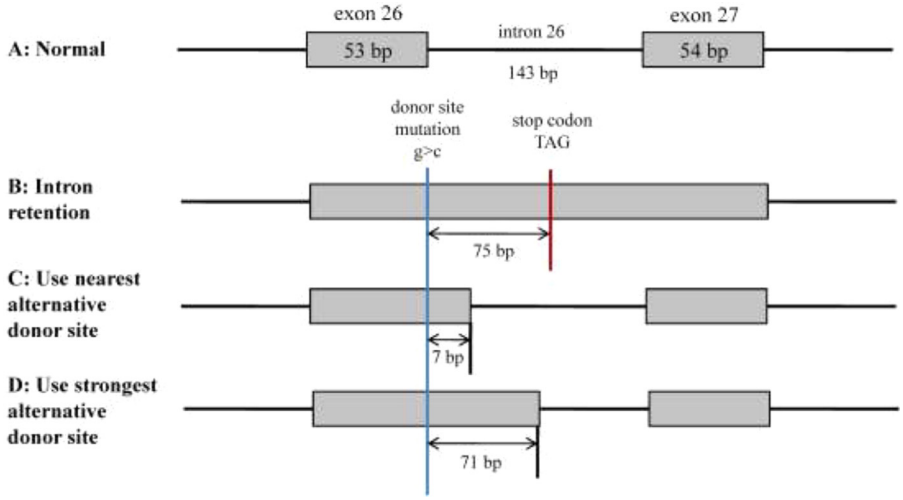

Fig. 2. Bioinformatics prediction of the effect of the mutation on the splicing of the *COL1A1* gene. (A) Normal splicing; (B) Aberrant splicing with intron retention; (C) Aberrant splicing using the nearest alternative donor site; (D) Aberrant splicing using the strongest alternative donor site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

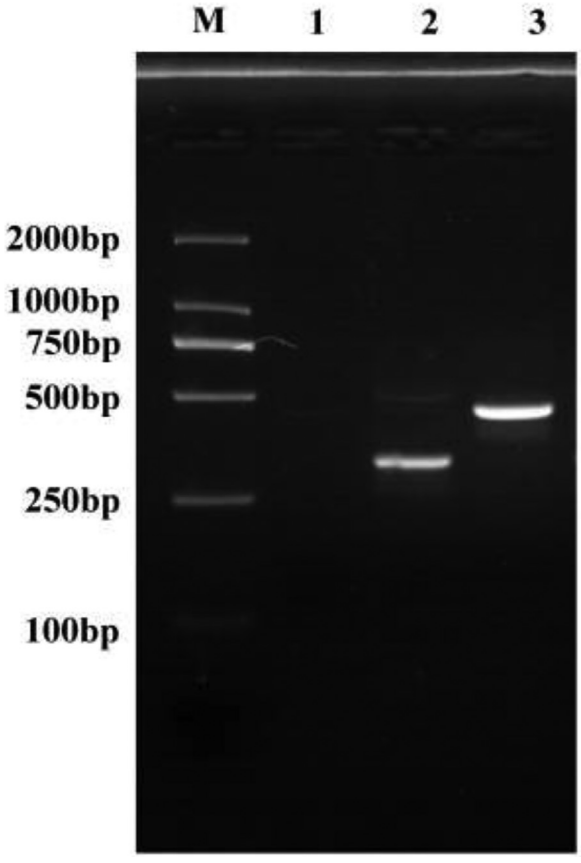

Fig. 3. Electrophoretic pattern of the RT-PCR products. M: DNA marker (from up to bottom: 2000, 1000, 750, 500, 250 and 100 bp); (1) The empty plasmid; (2) The wild-type plasmid with a normal splicing fragment of 289 bp; (3) The mutant plasmid with a splicing fragment of 432 bp.

whether the variant represents a disease-causing mutation or a benign polymorphism and study the mechanism involved in normal and aberrant splicing.

In this work, we identify a novel splice donor site mutation c. 1821+1 *G>C* in the *COL1A1* gene. Three lines of evidence supported our current hypothesis that this reported variant might destroy the biological function of *COL1A1*. Firstly, splice-site mutations in the *COL1A1* gene reported thus far are generally linked to the loss of function and have implications for pathogenicity. Secondly, Human Splice Finder predicted that the mutation would generate new alternative splice sites, probably resulting in intron retention or alternative donor site using. Thirdly, HMSA results proved that the mutation contributed to the disappearance of the 5′ donor splice site of intron 26, which led to intron 26 retention, introduction of the premature stop codon, and dysfunction of protein ultimately. According to the American College of Medical Genetics and Genomics (ACMG) guidelines for interpretation of sequence variants [22], we consider this mutation to be pathogenic (PVS1+PM2+PM3+PP1+PP4). It was worth noting that this pedigree reflected a feature of OI, while highly varied clinical consequences. Three carriers of this pedigree had the same mutation site but difference in severity of symptoms. The possible explanations include modification of transcription factors, editing of mRNA, unknown compensatory mechanism, interaction of other gene products and environmental factors [3,4,23].

In conclusion, we identified a novel splice donor site mutation in the proband by applying NGS. Moreover, both of bioinformatics and in vitro experimentation (HMSA) results demonstrated that the present mutation caused OI progression via aberrant splicing. Therefore, we strongly recommend that the combination of NGS, bioinformatics and HMSA be applied to the genetic diagnosis and abnormal splicing study of OI. Furthermore, since the proband and her sister carried the same mutation, prenatal diagnosis or preimplantation genetic diagnosis can be one of the choices for their future reproduction.

Declaration of Competing Interest

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Acknowledgements

This work was supported by the Natural Science Foundation of Fujian Province, China (No. 2018J01348).

References

[1] Gatti D, Rossini M, Viapiana O, Povino MR, Liuzza S, Fracassi E, et al. Teriparatide treatment in adult patients with osteogenesis imperfecta type I. Calcif. Tissue Int. 2013;93:448–52.
[2] Harrington J, Sochett E, Howard A. Update on the evaluation and treatment of osteogenesis imperfecta. Pediatr. Clin. North Am. 2014;61:1243–57.
[3] Van Dijk FS, Sillence DO. Osteogenesis imperfecta: clinical diagnosis, nomenclature and severity assessment. Am. J. Med. Genet. A 2014;164A:1470–81.
[4] Forlino A, Cabral WA, Barnes AM, Marini JC. New perspectives on osteogenesis imperfecta. Nat. Rev. Endocrinol. 2011;7:540–57.
[5] Makareeva E, Sun G, Mirigian LS, Mertz EL, Vera JC, Espinoza NA, et al. Substitutions for arginine at position 780 in triple helical domain of the alpha1(I) chain alter folding of the type I procollagen molecule and cause osteogenesis imperfecta. PLoS ONE 2018;13:e0200264.
[6] Scheres LJJ, van Dijk FS, Harsevoort AJ, van Dijk ATH, Dommisse AM, Janus GJM, et al. Adults with osteogenesis imperfecta: clinical characteristics of 151 patients with a focus on bisphosphonate use and bone density measurements. Bone Rep. 2018;8:168–72.
[7] Hui P. Next generation sequencing: chemistry, technology and applications. Top Curr Chem. 2014;336:1–18.
[8] Bahassi el M, Stambrook PJ. Next-generation sequencing technologies: breaking the sound barrier of human genetics. Mutagenesis 2014;29:303–10.
[9] Kallel-Bouattour R, Belguith-Maalej S, Zouari-Bradai E, Mnif M, Abid M, Hadj Kacem H. Intronic variants of SLC26A4 gene enhance splicing efficiency in hybrid minigene assay. Gene 2017;620:10–14.
[10] Wang Z, Lin Y, Qiu L, Zheng D, Yan A, Zeng J, et al. Hybrid minigene splicing assay verified the pathogenicity of a novel splice site variant in the dystrophin gene of a chinese patient with typical duchenne muscular dystrophy phenotype. Clin. Chem. Lab. Med. 2016;54:1435–40.
[11] Xie S, Lan Z, Qu N, Wei X, Yu P, Zhu Q, et al. Detection of truncated dystrophin lacking the C-terminal domain in a chinese pedigree by next-generation sequencing. Gene 2012;499:139–42.
[12] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25:2078–9.
[13] Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human splicing finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009;37:e67.
[14] De Conti L, Baralle M, Buratti E. Exon and intron definition in pre-mRNA splicing. Wiley interdisciplinary reviews RNA 2013;4:49–60.
[15] Lee Y, Rio DC. Mechanisms and regulation of alternative pre-mrna splicing. Annu. Rev. Biochem. 2015;84:291–323.
[16] Liu X, Wang Y, Chang G, Wang F, Wang F, Geng X. Alternative splicing of hTERT pre-mrna: a potential strategy for the regulation of telomerase activity. Int. J. Mol. Sci. 2017:18.
[17] MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, et al. Guidelines for investigating causality of sequence variants in human disease. Nature 2014;508:469–76.
[18] Singh RK, Cooper TA. Pre-mRNA splicing in disease and therapeutics. Trends Mol. Med. 2012;18:472–82.
[19] Black DL. Mechanisms of alternative pre-messenger RNA splicing. Annu. Rev. Biochem. 2003;72:291–336.
[20] Baralle D, Baralle M. Splicing in action: assessing disease causing sequence changes. J. Med. Genet. 2005;42:737–48.
[21] Desviat LR, Perez B, Ugarte M. Minigenes to confirm exon skipping mutations. Methods in molecular biology (Clifton, NJ) 2012;867:37–47.
[22] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. 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: official journal of the American College of Medical Genetics 2015;17:405–24.
[23] Patel RM, Nagamani SC, Cuthbertson D, Campeau PM, Krischer JP, Shapiro JR, et al. A cross-sectional multicenter study of osteogenesis imperfecta in north america - results from the linked clinical research centers. Clin. Genet. 2015;87:133–40.