The Journal of Molecular Diagnostics, Vol. 25, No. 1, January 2023

the Journal of Molecular Diagnostics
jmdjournal.org

Pathogenicity of Intronic and Synonymous Variants of *ATP7B* in Wilson Disease

Wan-Qing Xu, Rou-Min Wang, Yi Dong, and Zhi-Ying Wu

From the Departments of Neurology and Medical Genetics, Second Affiliated Hospital, and Key Laboratory of Medical Neurobiology of Zhejiang Province, Zhejiang University School of Medicine, Hangzhou, China

Accepted for publication October 18, 2022.

Address correspondence to Zhi-Ying Wu, M.D., Ph.D., Departments of Neurology and Medical Genetics, Second Affiliated Hospital, and Key Laboratory of Medical Neurobiology of Zhejiang Province, Zhejiang University School of Medicine, 88 Jiefang Rd., Hangzhou 310009, China. E-mail: zhiyingwu@zju.edu.cn.

Wilson disease (WD) is a hereditary disorder of copper metabolism, resulting from mutations within *ATP7B*. Early diagnosis is essential for affected individuals. However, there are still patients with clinically suspected WD who do not have detectable pathogenic variants, which makes diagnosis difficult and delays treatment. This study included such patients from the authors' center and screened for the full-length sequence of *ATP7B* by next-generation sequencing. Newly identified synonymous and intronic variants were then analyzed with *in silico* tools. A minigene system was constructed to determine the pathogenicity of these variants in terms of splicing and blood RNA extraction, and RT-PCR experiments were performed on several patients to verify the splicing alterations. The phenotypes of the patients were also analyzed. Fourteen suspected pathogenic variants, including nine synonymous and five intronic variants, were detected in 12 patients with clinically suspected WD. Among them, four synonymous variants (c.1050G>A, c.1122C>G, c.3243G>A, and c.4014T>A) and four intronic variants (c.1543 +40G>A, c.1707+6_1707+16del, c.1870-49A>G, and c.2731-67A>G) resulted in splicing changes in *ATP7B*. After the above analysis, the diagnosis of WD could be confirmed in eight clinically suspected patients with WD who showed a late age of onset. *(J Mol Diagn 2023, 25: 57—67; https://doi.org/10.1016/j.jmoldx.2022.10.002)*

Hepatolenticular degeneration, also known as Wilson disease (WD), is an autosomal recessive disorder of copper metabolism, in which the only known causative gene is *ATP7B* on chromosome 13. The ATPase copper transporting beta (ATP7B) protein is a P-type ATPase that transports copper to ceruloplasmin and excretes excess copper through cytosol.[1] Defects in its function will lead to copper accumulation in organs and tissues, such as the liver, brain, cornea, and kidney, triggering oxidative damage and mitochondrial dysfunction and progressively causing different combinations of symptoms, including liver damage, neuropsychiatric symptoms, corneal pigment ring (Kayser-Fleischer ring), and nephritis.[2] Current estimates of the worldwide incidence are about 1 in 2600 to 30,000 individuals and about 1 in 90 carriers of the causative gene, but some statistical studies in East Asia in recent decades suggest that the incidence should be higher in China.[3]

On the basis of the aforementioned pathogenic mechanisms, the age at onset (AAO) and clinical symptoms of patients with WD are highly heterogeneous, which also makes their diagnosis extra difficult. According to the 2021 Chinese guidelines for the diagnosis and treatment of WD,[4] in addition to the corresponding clinical symptoms and manifestations of low serum ceruloplasmin, the diagnosis of WD must be determined by pathogenicity analysis and lineage cosegregation to identify patients carrying pathogenic variants of *ATP7B* gene on both chromosomes, which underlines the importance of genetic diagnosis. With current high-throughput next-generation sequencing (NGS) technologies enabling a rapid increase in the number of variants identified, >1000 *ATP7B* variants have been reported, including common variants with clear pathogenesis: 571 missense mutations, 222 shift mutations, 78 nonsense

Supported by National Natural Science Foundation of China grants 81125009 and 81701126; and the research foundation (188020-193810101/089) for distinguished scholar of Zhejiang University (Z.-Y.W.).

Disclosures: None declared.

Copyright © 2023 Association for Molecular Pathology and American Society for Investigative Pathology. Published by Elsevier Inc. All rights reserved.
https://doi.org/10.1016/j.jmoldx.2022.10.002

mutations, 76 traditional splice site mutations, and 27 large fragment deletions; and less common variants with pathogenesis yet to be investigated: 10 noncoding region mutations and 27 splicing variants (source: Human Gene Mutation Database). For the pathogenesis of unconventional splicing site variants, previous studies in China have found that the synonymous mutations cause exon skipping by affecting mRNA splicing.[5,6] In 2020, a study in Canada also demonstrated that the missense mutation (c.1934T>G, p.M645R) can also cause exon skipping by alternation in splicing elements.[6]

Accurate pre-mRNA splicing is essential for accurate protein translation, and approximately 15% to 50% of human disease gene mutations alter the function of basic splicing elements or auxiliary regulatory sequences, resulting in pathogenicity.[7] Moreover, existing *in silico* tools cannot fully predict whether synonymous variants as well as noncoding region variants lead to aberrant splicing of mRNAs,[8] thus necessitating functional studies of such variants and determination of their pathogenicity. At present, the genetic diagnosis rate of WD in the current center has reached >95%,[9,10] but there are still some patients presenting clinical manifestations suspected of WD with only one or none *ATP7B* pathogenic variants detected through Sanger sequencing and multiplex ligation-dependent probe amplification. Therefore, we further analyzed the NGS screening results of the full-length *ATP7B* gene in these patients and found synonymous and intronic variants of possible pathogenic significance for study with the aim of further refining the diagnosis.

The most commonly used functional experimental methods for mRNA splicing studies are RT-PCR and *in vitro* construction of minigene plasmid systems.[11] Minigene detects mRNA splicing changes caused by gene variants in exons or introns by encapsulating all related splicing elements into a special plasmid. Compared with RT-PCR, the minigene system allows for parallel analysis of wild type and mutant in the same cellular context to better determine the effect of single allele variant. Because most ATP7B protein is expressed in the liver and different transcripts may be present in the blood, this study chose to validate the pathogenicity of these suspected variants *in vitro* using liver-derived cells and constructing minigene systems.

Materials and Methods

Patients

This study included 12 patients admitted to the authors' center between November 2012 and February 2021, whose clinical presentation was suspected of WD but genetic diagnosis could not be obtained by Sanger sequencing and multiplex ligation-dependent probe amplification.[5,9] However, these patients had synonymous or intronic variants of unknown significance in the full-length *ATP7B* gene sequenced by NGS. The clinical diagnosis was made according to the Leipzig score. Each patient was evaluated for clinical history and physical examination by at least two senior neurologists. Clinical data were collected, including AAO, age at diagnosis, symptom presentations, Unified Wilson's Disease Rating Scale scores,[12] laboratory test results, such as serum ceruloplasmin and 24-hour urine copper, and imaging findings in liver ultrasound and cranial magnetic resonance imaging. The study was approved by the ethics committee of Second Affiliated Hospital, Zhejiang University School of Medicine (Hangzhou, China). The informed consents were obtained by the participants or their guardians.

Genetic Analysis

Genomic DNA was extracted from peripheral blood using a DNA isolation kit (Qiagen Inc., Valencia, CA). Full-length *ATP7B* gene was sequenced using NGS technology. When no suspected pathogenic variants were found in exon and conventional upstream and downstream splice site, intronic and synonymous variants were included. Variants with frequencies of >0.05 in the Genome Aggregation Database (gnomAD), 1000Genomes, and Exome Aggregation Consortium (ExAc) databases were removed. Variants classified as benign were excluded in accordance with the American College of Medical Genetics and Genomics guidelines, in conjunction with evidence of lineage cosegregation. The screened variants were used for the next step of splicing-related *in silico* analysis and functional experiments.

In Silico Splicing Analysis

The following databases and *in silico* analysis software were used for changes in *ATP7B* splicing process caused by mutations, according to American College of Medical Genetics and Genomics guidelines[13] and recent literature. MaxEntScan was used for predicting canonical splice sites. Exonic splicing enhancer (ESE) Finder and RESCUE-ESE were used for predicting possible splicing changes caused by synonymous variants in exons. CADD version 1.6, SpliceAI,[14] and Human Splicing Finder version 3.1 were used for compound analysis.

Minigene System

The pSPL3 vector is widely utilized in exon-trapping procedure. It contains a coding sequence and a functional pair of splice donor and splice acceptor. The effect of variants on gene splicing can be verified by inserting the corresponding target gene fragment into its intron sequence. Wild-type and variant-containing minigene systems were constructed using pSPL3. The primers used are listed in Table 1. The pSPL3 vector was linearized by double digestion using XhoI and BamHI (Takara, Shiga, Japan) endonucleases in a 37°C water bath for 4 to 6 hours. The target gene fragment and the linearized vector were then ligated using a homologous

Table 1 Primers Used for Minigene Design

Variant	Primers
c.1050G>A p.P350P	F: 5′-TCACCAGAATTCTGGAGCTCGAGCGTGCTGCATTCCATATGT-3′
	R: 5′-GGATCACCAGATATCTGGGATCCAATTCAGCCATTCTGTTAT-3′
c.1122C>G p.V374V	F: 5′-TCACCAGAATTCTGGAGCTCGAGCGTGCTGCATTCCATATGT-3′
	R: 5′-GGATCACCAGATATCTGGGATCCAATTCAGCCATTCTGTTAT-3′
c.1543+40G>A	F: 5′-TCACCAGAATTCTGGAGCTCGAGTATAGAGAACTATTTACAA-3′
	R: 5′-GGATCACCAGATATCTGGGATCCCCTAACTGAGAAGTCTATC-3′
c.1707+6_1707+16del	F: 5′-TCACCAGAATTCTGGAGCTCGAGCGAGATCATGCTACTGTAC-3′
	R: 5′-GGATCACCAGATATCTGGGATCCTTCTCTTAGAATGGCACTT-3′
c.1870-49A>G	F: 5′-TCACCAGAATTCTGGAGCTCGAGTCCTCATACATGGAATGAA-3′
	R: 5′-GGATCACCAGATATCTGGGATCCCAGTCACACTAGCTGCATT-3′
c.2145C>T p.Y715Y	F: 5′-TCACCAGAATTCTGGAGCTCGAGGGTCTCACATGCTCTTGGTC-3′
	R: 5′-GGATCACCAGATATCTGGGATCCAGACAATGTAGGCTCTGCCC-3′
c.2268G>A p.A756A	F: 5′-TCACCAGAATTCTGGAGCTCGAGTGTGCTGTCTACAGTACAA-3′
	R: 5′-GGATCACCAGATATCTGGGATCCTGCACCAAGAGACAATGTA-3′
c.2292C > T p.F764F	F: 5′-TCACCAGAATTCTGGAGCTCGAGGCCTTGTGCTGTCTACAGT-3′
	R: 5′-GGATCACCAGATATCTGGGATCCATACACCTGAATGATGGTT-3′
c.2583C>T p.A861A	F: 5′-TCACCAGAATTCTGGAGCTCGAGCCATGGAGCTGGTGCAGCG-3′
	R: 5′-GGATCACCAGATATCTGGGATCCAGTTGACATGATGATGATA-3′
c.2731-67A>G	F: 5′-TCACCAGAATTCTGGAGCTCGAGGTATTCAGCAGCTGCACGA-3′
	R: 5′-GGATCACCAGATATCTGGGATCCTCCTACTCTGGCTTAGATT-3′
c.2866-13G>C	F: 5′-TCACCAGAATTCTGGAGCTCGAGTCCTATGCCAGGTGTTATG-3′
	R: 5′-GGATCACCAGATATCTGGGATCCTGTTGCTACTGTTGTTATT-3′
c.3243G>A p.E1081E	F: 5′-TCACCAGAATTCTGGAGCTCGAGTTGACTGGACACATAGGGA-3′
	R: 5′-GGATCACCAGATATCTGGGATCCCACTCCTGTTGTTCTCCTG-3′
c.4014T>A p.I1338I	F: 5′-TCACCAGAATTCTGGAGCTCGAGTAAAGGGAAGAAAGTCGCC-3′
	R: 5′-GGATCACCAGATATCTGGGATCCGGGATCAGAAAATACAGCC-3′
c.4251A>G p.T1417T	F: 5′-TCACCAGAATTCTGGAGCTCGAGGTGCAAGGCATTTGCTTAG-3′
	R: 5′-GGATCACCAGATATCTGGGATCCGCAGGCCTGCATTAGAGAA-3′

F, forward; R, reverse.

recombination kit (ClonExpressII; Vazyme, Nanjing, China). The above ligation products were transformed into DH5α-receptive *Escherichia coli*, and inoculated onto solid medium plates containing ampicillin antibiotics overnight, and then the constructed plasmids were extracted using DP118 kit (Tiangen, Beijing, China). Minigene constructs containing wild or mutant alleles were transfected into HepG2 cells using Lip3000 (Invitrogen, Waltham, MA). After 48 hours, total RNA was extracted from the cells with Trizol reagent (Takara). The cDNA strands were obtained by reverse transcription using PrimeScript II 1st Strand cDNA Synthesis Kit (Takara) and then amplified with TaKaRa Ex Taq (Takara) by primers located in the vector. The sequences of the primers are as follows: forward splice donor, 5′-TCTGAGTCACCTGGACAACC-3′; and reverse splice acceptor, 5′-TCAGTGGTATTTGTGAGCCA-3′. PCR products were analyzed by 1.5% agarose gel electrophoresis and proved by sequencing of extracted DNA.

Transcript Analysis *in Vivo*

Total RNA was extracted from peripheral leukocytes of three patients carrying c.4014T>A (Patient 4), c.1543+40G>A (Patient 6), and exon8-c.2292C>T (Patient 9). Fibroblasts from Patient 5 carrying the c.1870-49A>G variant were cultured, and total RNA was extracted. Blood samples and fibroblasts from healthy volunteers were used as normal controls for mRNA. Reverse transcription to cDNA was performed using Takara PrimeScript RT Master Mix. PCR was then performed (primers shown in Table 2), and the products were separated on 1.5% agarose.

Table 2 Primers Used for PCR

Variant	Primers
c.1543+40G>A	F: 5′-ATAGGCAATGGGAAGAAC-3′
	R: 5′-TGCCAGTTATACAAGGAC-3′
c.1870-49A>G	F: 5′-CTCTTCTGAGAAGGGTGG-3′
	R: 5′-ACACTAGCTGCATTTGAAG-3′
c.2292C > T p.F764F	F: 5′-GCCTTGTGCTGTCTACAG-3′
	R: 5′-CTCTGCCCTGAAGGCCA-3′
c.4014T>A p.I1338I	F: 5′-AGGTCCAGGAGCTCCAGA-3′
	R: 5′-TGAGCTGCAGGGATGAGAG-3′

F, forward; R, reverse.

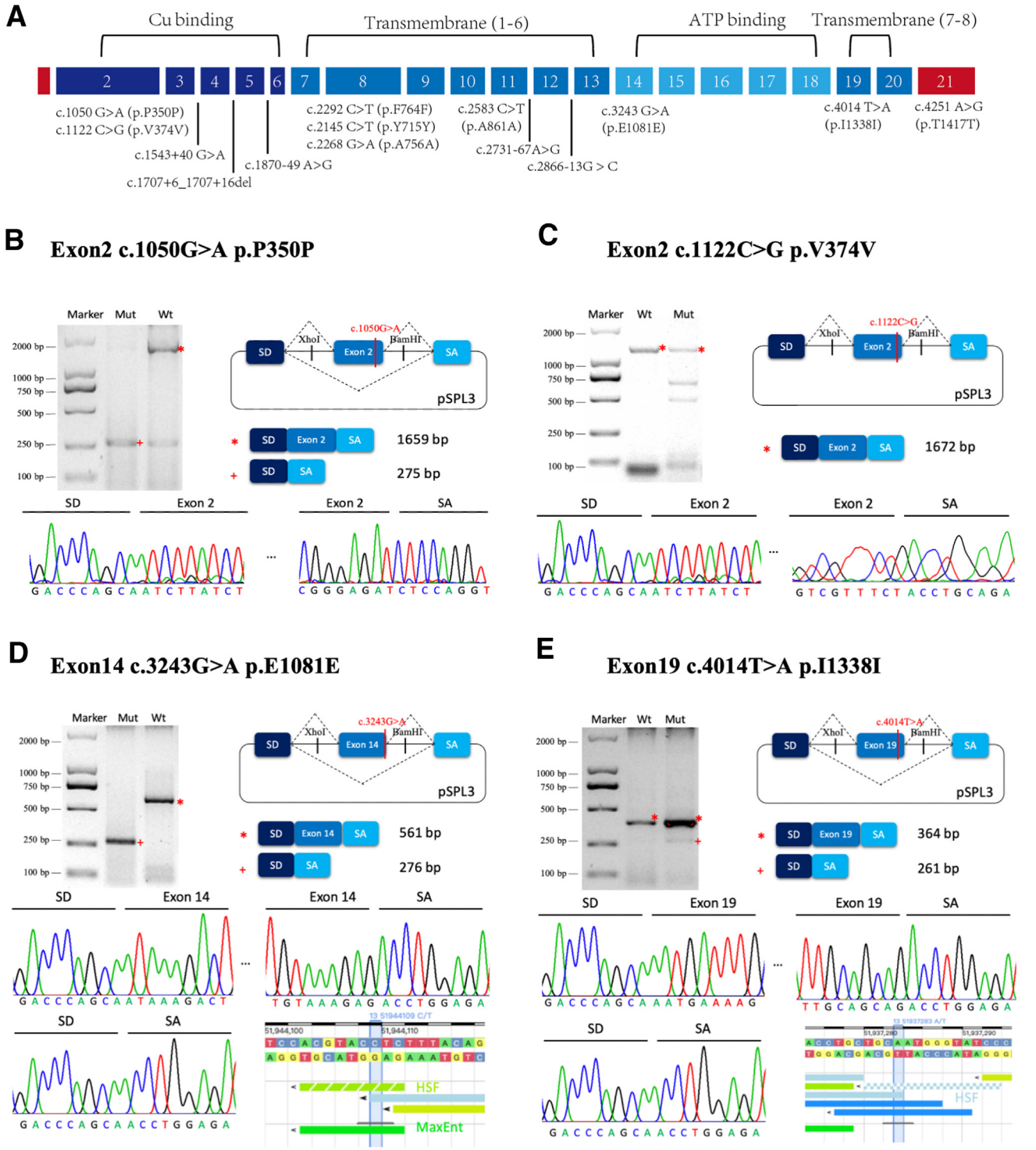

Figure 1 Minigene construction and altered mRNA splicing results in synonymous variants. The + and **asterisk** symbols represent different bands. **A:** Structural diagram of *ATP7B* gene and the variants involved in this study. **B:** Variant c.1050G>A, p.P350P, caused exon 2 skipping. **C:** Variant c.1122C>G, p.V374V, caused partial deletion of exon 2. **D:** Variant c.3243G>A, p.E1081E, caused exon 14 skipping. **E:** Variant c.4014T>A, p.I1338I, caused partial exon 19 skipping. Mut, mutated; SA, splice acceptor; SD, splice donor; Wt, wild type.

Statistics and Analysis

AAO data from the team's previous study cohort[15] ($n = 715$) were used for comparison with data from the present study ($n = 8$) using the Wilcoxon signed rank test. The threshold for statistical significance was $P < 0.05$, and the software used for statistics and plotting was GraphPad Prism 9.0 (GraphPad Inc., La Jolla, CA) and Microsoft Excel 2021 (Microsoft, Redmond, WA).

Table 3 Description of the Patient's Genetic Variants and Clinical Symptoms

Patient no.	*ATP7B* variants			Sex	Age, years	AAO, years	CP, mg/L	Onset	K-F rings	Liver cirrhosis	Cranial MRI	Splenomegaly
1	c.2390C>T	c.2731-67A>G*		M	43	33	20	H	−	−	−	−
2	c.1449_1456del	c.2268G>A*	c.1122C>G*	F	45	36	47	H	−	+	−	−
3	c.2333G>T	c.1707+6_1707+16del*		M	25	11	26	H	+	−	−	−
4	c.2333G>T	c.4014T>A*		F	34	29	90	N	−	+	+	+
5	c.2333G>T	c.1870-49A>G*		M	21	14	24	N	+	−	+	+
6	c.2447+5G>T	c.1543+40G>A*		M	7	NA	47	A	−	−	−	−
7	c.2620G>C	c.1050G>A*		M	21	17	26	N	+	+	−	+
8	c.2333G>T	c.3243G>A*		F	41	38	32	N	+	+	+	+
9	c.2333G>T	c.2292C>T*	c.2866-13G>C*	M	8	NA	30	A	−	−	−	−
10	c.2975C>T	c.2145C>T*		M	47	39	20	H	+	+	−	+
11	c.3889G>A	c.4251A>G*		M	22	NA	20	A	−	−	−	−
12	c.2583C>T*			M	23	3	117	N	−	−	−	−

*The variants constructed in the minigene system in the current study.

F, female; M, male; +, the patient has this manifestation; −, the patient does not have this manifestation; A, asymptomatic; AAO, age of onset; CP, ceruloplasmin, with a normal range of 200 to 600 mg/L; H, hepatic; K-F, Kayser-Fleischer; MRI, magnetic resonance imaging; N, neurologic; NA, not available.

Results

General Information and *in Silico* Splicing Analysis

Fourteen variants in *ATP7B*, including nine synonymous variants and five intronic variants, were identified in the 12 patients with suspected WD (Figure 1). The age, sex, AAO, symptoms, variants, and clinical test results of the patients are shown in Table 3. After *in silico* splicing analysis, c.1050G>A, c.1543+40G>A, c.1870-49A>G, c.2583C>T, and c.2731-67A>G do not show any obvious splicing alteration. The c.1122C>G may generate or activate a new splice donor, whereas c.1707+6_1707+16del may generate or activate three splice acceptors and disrupt the original two splice donors, thus affecting the splicing process. Both synonymous variants in exon 8, c.2145C>T, and c.2292C>T may affect auxiliary sequences, resulting in significant changes in the ratio of exon splicing enhancer to exon splicing silencer motifs, whereas c.2268G>A may generate or activate a new splicing acceptor. The variant c.3243G>A in exon 14 can lead to the disruption of the original splice donor site, c.4014T>A in exon 19 may cause the generation of a new splice acceptor, and c.4251A>G in exon 21 may affect the auxiliary sequence and lead to a significant change in the ratio of exon splicing enhancer and exon splicing silencer motifs (Table 4).

Splicing Alterations Identified in Minigene System

After validation of the minigene system, it was found that the intronic variants, including c.1543+40G>A, c.1707+6_1707+16del, c.1870-49A>G, and c.2731-67A>G, and synonymous ones, including c.1050G>A (p.P350P), c.1122C>G (p.V374V), c.4014T>A (p.I1338I), and c.3243G>A (p.E1081E), may lead to splicing changes in the corresponding exon regions, as shown in Figure 1. However, the intronic variant (c.2866-13G>C) and the synonymous ones, including c.2145C>T (p.Y715Y), c.2268G>A (p.A756A), c.2292C>T (p.F764F), c.2583C>T (p. A861A), and c.4251A>G (p.T1417T), may not affect the splicing process of the corresponding exon regions. The remaining negative results are shown in Supplemental Figure S1.

In the intronic variants, c.1707+6_1707+16del caused complete skipping of exon 4, which is consistent with the prediction of *in silico* tools, whereas c.1543+40G>A and c.1870-49A>G caused complete skipping of exons 3 and 5, respectively, with *in silico* predictions indicating no possible splicing alterations. And c.2731-67A>G resulted in the retention of introns 10 and 11 in the minigene system (Figure 2).

The synonymous variant c.1050G>A (p.P350P) lead to exon 2 skipping in contrast to the negative *in silico* predictions, whereas also in exon 2, c.1122C>G (p.V374V) caused the generation of multiple splicing results. The synonymous variant c.3243G>A (p.E1081E) is located at the last one position of exon 14, and minigene system showed that exon 14 is completely skipped in splicing, which is consistent with the prediction of *in silico* tools that this variant could disrupt the original splicing donor site. And the synonymous variant c.4014T>A (p.I1338I) caused partial skipping of exon 19 in minigene construct, which is consistent with the prediction that the generation or activation of a new splice acceptor by this variant could lead to exon skipping.

RT-PCR Analysis *in Vivo*

In vivo validation experiments were performed by RT-PCR using mRNA extracted from blood leukocytes of Patients 4, 6, and 9 and from fibroblast cultures of Patient 5. The results are shown in Figure 2. No corresponding splicing changes were verified in Patients 4 and 9, whereas splicing changes

Table 4 Pathogenicity Interpretation of 14 *ATP7B* Intronic and Synonymous Variants

Site	NA alterations	AA alterations	1000g (China)	ExAC	gnomAD	SpliceAI
Exon 2	c.1050G>A	p.P350P		0.0003	0.0026	
Exon 2	c.1122C>G	p.V374V	0	0.0008	0.0026	
IVS3	c.1543+40G>A			0.0021	0.0026	
IVS4	c.1707+6_1707+16del				0.0026	
IVS5	c.1870-49A>G				0.0026	Damaging
Exon 8	c.2145C>T	p.A756A		0.0008		
Exon 8	c.2268G>A	p.F764F		0.0001		
Exon 8	c.2292C>T	p.Y715Y				
Exon 11	c.2583C>T	p.A861A	0	0.0039		
IVS11	c.2731-67A>G					
IVS12	c.2866-13G>C		0.0505	0.2357		
Exon 14	c.3243G>A	p.E1081E		0	0	
Exon 19	c.4014T>A	p.I1338I			0	
Exon 21	c.4251A>G	p.T1417T	0	0.002	0	

(*table continues*)

AA, amino acid; AS, acceptor site; CADD, Combined Annotation Dependent Depletion; DS, donor site; ESE, exon splicing enhancer; ExAC, Exome Aggregation Consortium; gnomAD, Genome Aggregation Database; HSF, Human Splicing Finder; IVS, intervening sequence; LB, likely benign; LP, likely pathogenic; NA, nucleic acid; WT, wild type.

consistent with minigene results were observed in Patients 5 and 6.

Pathogenicity Interpretation and Phenotypic Analysis

On the basis of the results of the *in silico* analysis and minigene experiments described above, the current 14 variants were reinterpreted for American College of Medical Genetics and Genomics pathogenicity, which led to genetic confirmation in 8 of the 12 patients with clinically suspected WD included in the study. The pathogenicity of c.1050G>A and c.1122C>G, which were originally likely benign, can be classified as likely pathogenic (Table 4). The pathogenicity of c.1543+40G>A, c.1707+6_1707+16del, c.1870-49A>G, c.2731-67A>G, c.3243G>A, and c.4014T>A, which were originally of unknown significance, can be classified as likely pathogenic. The pathogenicity of c.4251A>G, which was originally of unknown significance, can be classified as likely benign. The pathogenicity of c.2145C>T, c.2268G>A, c.2292C>T, c.2583C>T, and c.2866-13G>C remained as likely benign.

Table 4 (*continued*)

CADD	HSF matrices	MaxEnt	ESE finder	ESE motifs	Minigene	Pathogenicity interpretation
					Exon 2 skipping	LP
PHRED > 10	New DS		New site in SRp55 (-13.15%)		Partial deletion of exon 2	LP
			New site in SRp55 (-4.17%)/site broke		Exon 3 skipping	LP
	New AS Broken WT DS	New AS Broken WT DS	New site in SRp55 (+3.11%)/ site broke		Exon 4 skipping	LP
PHRED > 10				New site (+6.33%)	Exon 5 skipping	LP
	Alteration of auxiliary sequences		New site in SRp40 (+13.78%)			LB
	New AS		New site in SRp40 (+0.42%)	New site (+7.68%)		LB
	Alteration of auxiliary sequences		New site in SRp40 (+2.37%)	Site broke		LB
PHRED > 10				Site broke (-16.24%)		LB
			New site in SC35 (+1.04%)		Retention of intron 11	LP
			New site in SRp40 (+7.19%)			LB
			New site in SC35 (+7.06%)			
PHRED > 10	Broken WT DS	Broken WT DS			Exon 14 skipping	LP
	New AS		New site in SRp40 (+0.42%); site broke in SRp55 (-13.84%)		Exon 19 skipping	LP
	Alteration of auxiliary sequences		Site broke	New site (+1.43%)		LB

Among the eight patients with genetically confirmed WD (Patients 1 through 8) (Table 3), five were men and three were women, of whom Patient 6 was asymptomatic and the rest had a mean AAO of 25.2 ± 11.2 years. And the mean level of serum ceruloplasmin was 39.0 ± 23.0 mg/L. Four patients had neurologic onset, four patients had hepatic onset, and four patients had significant corneal Kayser-Fleischer rings. Three cases of cranial magnetic resonance imaging suggested metal deposits, two of which were symmetrical lesions in the basal ganglia region, thalamus, and midbrain, and one case was symmetrical abnormal signal in the dorsal aspect of the cerebral peduncle and cerebral bridge. Patients 1 and 2 were lost to follow-up, and scale data could not be obtained. Patients 3 and 6 with hepatic onset had no significant change in Unified Wilson's Disease Rating Scale scores before and after treatment, whereas four patients with neurologic onset had significantly lower Unified Wilson's Disease Rating Scale scores over a treatment period of >1 year. For comparison, the mean AAO was 17.0 ± 27.0 years in the previously genetically confirmed WD control group ($n = 715$) in the center ($P < 0.01$).[15]

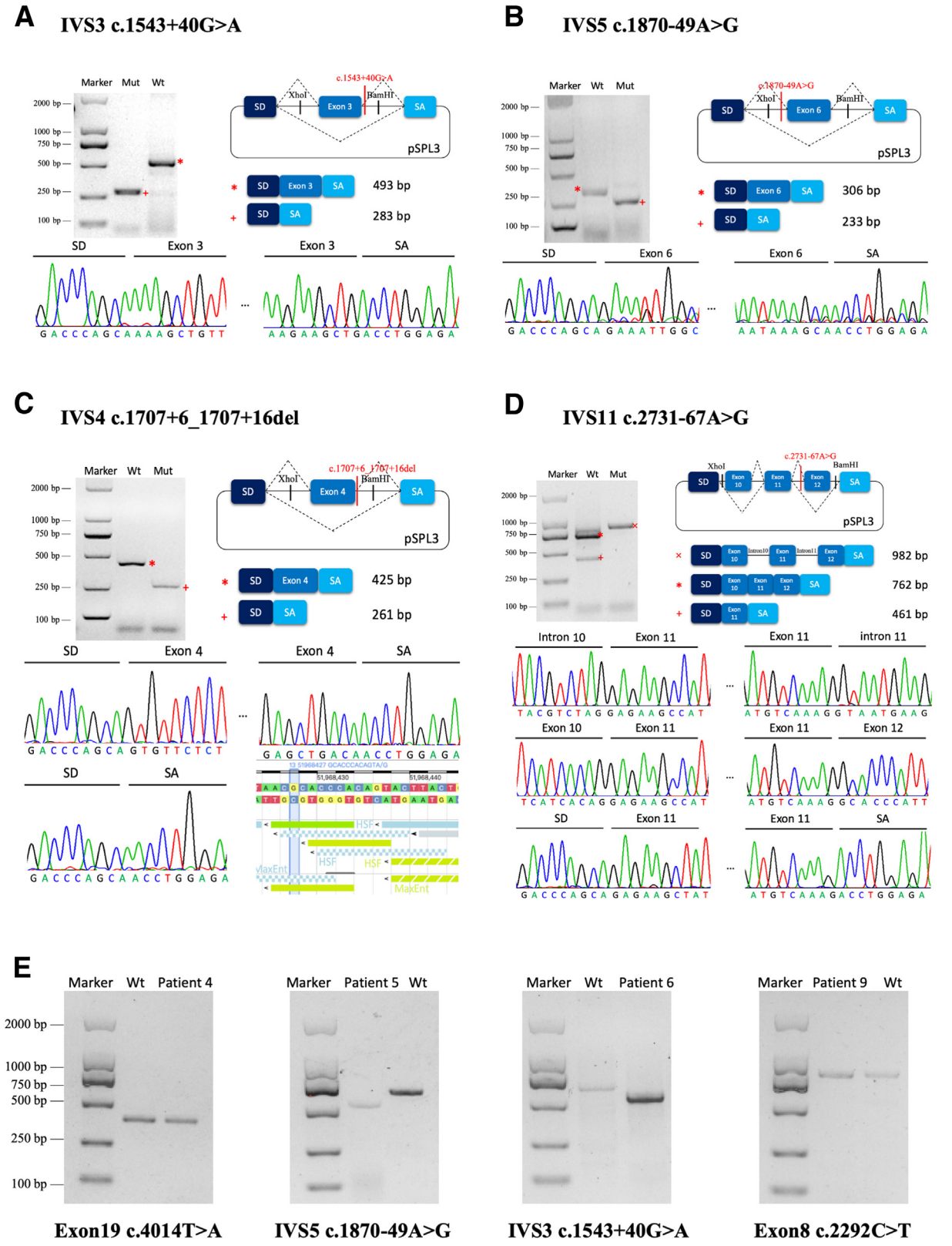

Figure 2 Minigene construction and altered mRNA splicing results in intronic variants. The +, **asterisk**, and **x** symbols represent different bands. **A:** Variant c.1543+40G>A caused exon 3 skipping. **B:** Variant c.1870-49A>G caused exon 5 skipping. **C:** Variant c.1707+6_1707+16del caused exon 4 skipping. **D:** Variant c.2731-67A>G caused the retention of intron. **E:** RT-PCR analysis *in vivo*. Mut, mutated; SA, splice acceptor; SD, splice donor; Wt, wild type.

Discussion

The genetic diagnosis rate of WD is yet to be improved, and there are still some patients who clinically exhibit suspected symptom combinations, such as neuropsychiatric symptoms, low serum ceruloplasmin, or high urinary copper, abnormal liver function, and Kayser-Fleischer rings, but cannot get a definite genetic diagnosis by screening *ATP7B* mutations. Therefore, for this monogenic genetic disorder, full-length NGS of *ATP7B* should be performed, and functional studies of variants, such as synonymous and intronic variants, should be enhanced.

More than 95% of human genes undergo the splicing process of pre-mRNAs to produce multiple protein forms from the same transcript,[16] and any change in this process involving splicing regulatory sequences can lead to abnormal splicing results and affect protein functions.[17,18] With this perspective, synonymous and intronic variants previously considered as benign alterations may also be pathogenic through mechanisms such as causing disruption of existing splicing sites, generation of new sites, or activation of cryptic sites.[19,20] Therefore, in this study, we performed NGS and reanalyzed the full lengths of *ATP7B* in patients who failed to obtain a genetic diagnosis in a previous WD cohort, and screened nine suspected synonymous variants and five intronic variants for *in silico* prediction and minigene function experiments, which demonstrated that eight variants had an effect on splicing function and clarified the genetic diagnosis of WD for 8 of 12 patients.

Canonical splicing variants are defined as intronic splicing donor/acceptor sites and typically involve residues +1 and +2 at the donor site at the 5′ end of the exon and residues +1 and +2 at the acceptor site at the 3′ end. These sites are easily identified in sequencing for diagnostic or research purposes.[7] Therefore, this study does not include these relatively obvious pathogenic splicing variants, but focuses on deeper positions in the intron, including c.1543+40G>A, c.1707+6_1707+16del, c.1870-49A>G, c.2731-67A>G, and c.2866-13G>C. Among them, except for c.2866-13G>C, the remaining four variants showed positive results in minigene assays, indicating that noncanonical splicing site variants in noncoding regions are also important in disease. It has been suggested that intronic variants may also generate new acceptor or donor sites that bind to splicing complexes and pair with already existing intronic cryptic splice sites, or generate new regulatory elements that result in the recognition of specific intron sequences as exon sequences.[21,22] Thus, such variants may result in the incorporation of intron fragments (ie, pseudoexons) into mature transcripts.[23] In the present study, the c.2731-67A>G variant of intron 11 was not successfully spliced in the minigene system to remove this intron. It can be concluded that, contrary to most previous views, the deep intron region also plays a role in physiological and pathologic mechanisms related to mRNA processing. Therefore, when the standard screening of coding regions and exon/intron boundaries does not lead to conclusions, the analysis of the possible pathogenicity of variants in deeper noncoding regions should be considered.

Recent studies have also revealed that synonymous variants can affect gene function through a variety of mechanisms.[24] However, the functional prediction of synonymous variants is also as difficult as the intronic variants. Existing methods for splicing prediction are limited, partially because of the limited experimental data available to assess variant effects and the lack of data for training or testing.[25] In the present study, nine synonymous mutations of *ATP7B* were included, of which c.1050G>A, c.1122C>G, c.3243G>A, and c.4014T>A were reclassified as likely pathogenic by functional experiments. However, synonymous variants analyzed by several available *in silico* software programs as potentially altering splicing process, such as c.2145 C>T, c.2268G>A, c.2292C>T, and c.4251A>G, were not found to have significant splicing changes by minigene validation, whereas c.1050G>A, which was not predicted to have significant changes by *in silico* analysis, reflected suspicious splicing alternation in minigene systems. It is clear that *in silico* analysis software is not a substitute for functional experiments, and functional studies at the RNA or protein level should be performed for suspected synonymous variants in the process of exploring the pathogenicity.

However, the experimental results of minigene cannot be used as a final judgment of pathogenicity, and there is a possibility of imperfection in this research method. For example, while the current study was ongoing, a UK study found that the synonymous variant c.2292C>T in *ATP7B* resulted in a lack of exon 8 expression in the reverse transcription products from fibroblasts in both heterozygous and homozygous individuals, but this change was not seen in RNA extracted from leukocytes from the same patients.[26] In the minigene system of the present study, there was also no manifestation of exon 8 skipping in HepG2 cells. This suggests the possible tissue specificity of such splicing changes. However, minigene still has greater advantages compared with the direct extraction of blood leukocyte mRNA for RT-PCR validation. As in the present study, the splicing changes of c.4014T>A and c.2292C>T could not be verified by RT-PCR.

WD is a rare disease with a large number of low-frequency heterozygous variants, and therefore genotype-phenotype studies with large cohorts are relatively lacking. The relevant studies available today have focused on a few high-frequency variants, and it is difficult to obtain statistically significant results for other relatively rare variants, and even less for intronic and synonymous variants that were previously considered benign. Therefore, this study only made a preliminary attempt to characterize the phenotype of these variants. After analyzing the symptoms, AAO, and clinical tests, it is hypothesized that synonymous and intronic variants causing abnormal splicing process may

retain some normal proteins due to a certain percentage of mRNA damage (as c.4014T>A), and thus exhibit a milder clinical phenotype, as reflected by a later AAO, milder symptoms, and better response to copper chelation therapy. However, variants that cause complete skipping of exons may also completely impair or terminate the subsequent expression process. This suggests a complex variation in the splicing process and its relevance to the phenotype. Thus, the phenotypes of synonymous and intron variants causing splicing abnormalities may be related to the specific splicing changes and cannot be compared uniformly.

WD is one of the few rare diseases that can be controlled by early diagnosis and treatment. In the era of NGS, the ultimate goal of genomic medicine is to use patients' sequencing results for genomic diagnosis and personalized therapy, and therapeutic manipulation methods for RNA splicing are under progressive investigations.[27] Thus, the *in vitro* study of synonymous variants and intronic variants causing splicing abnormalities in this study could also pave the way for future gene therapy for the disease.

Conclusion

Research on inherited rare diseases needs to revisit the significance of synonymous or intronic variants, and elucidating the pathogenicity of these variants through splicing functional experiments could help further improve genetic diagnosis, avoid delayed treatment, and possibly provide future directions for gene therapy.

Supplemental Data

Supplemental material for this article can be found at http://doi.org/10.1016/j.jmoldx.2022.10.002.

References

1. Petrukhin K, Lutsenko S, Chernov I, Ross BM, Kaplan JH, Gilliam TC: Characterization of the Wilson disease gene encoding a P-type copper transporting ATPase: genomic organization, alternative splicing, and structure/function predictions. Hum Mol Genet 1994, 3: 1647−1656
2. Członkowska A, Litwin T, Dusek P, Ferenci P, Lutsenko S, Medici V, Rybakowski JK, Weiss KH, Schilsky ML: Wilson disease. Nat Rev Dis Primers 2018, 4:21
3. Xie J-J, Wu Z-Y: Wilson's disease in China. Neurosci Bull 2017, 33: 323−330
4. Chinese Society of Neurogenetics: Chinese clinical practice guidelines for Wilson′s disease 2021. Chin J Neurol 2021, 54:310−319
5. Wang C, Zhou W, Huang Y, Yin H, Jin Y, Jia Z, Zhang A, Liu Z, Zheng B: Presumed missense and synonymous mutations in ATP7B gene cause exon skipping in Wilson disease. Liver Int 2018, 38: 1504−1513
6. Merico D, Spickett C, O'Hara M, Kakaradov B, Deshwar AG, Fradkin P, Gandhi S, Gao J, Grant S, Kron K, Schmitges FW, Shalev Z, Sun M, Verby M, Cahill M, Dowling JJ, Fransson J, Wienholds E, Frey BJ: ATP7B variant c.1934T > G p.Met645Arg causes Wilson disease by promoting exon 6 skipping. NPJ Genom Med 2020, 5:16
7. Ule J, Blencowe BJ: Alternative splicing regulatory networks: functions, mechanisms, and evolution. Mol Cell 2019, 76:329−345
8. Squitti R, Siotto M, Bucossi S, Polimanti R: In silico investigation of the ATP7B gene: insights from functional prediction of non-synonymous substitution to protein structure. Biometals 2014, 27:53−64
9. Dong Y, Ni W, Chen W-J, Wan B, Zhao G-X, Shi Z-Q, Zhang Y, Wang N, Yu L, Xu J-F, Wu Z-Y: Spectrum and classification of ATP7B variants in a large cohort of Chinese patients with Wilson's disease guides genetic diagnosis. Theranostics 2016, 6:638−649
10. Chen Y-C, Yu H, Wang R-M, Xie J-J, Ni W, Zhang Y, Dong Y, Wu Z-Y: Contribution of intragenic deletions to mutation spectrum in Chinese patients with Wilson's disease and possible mechanism underlying ATP7B gross deletions. Parkinsonism Relat Disord 2019, 62:128−133
11. Desviat LR, Pérez B, Ugarte M: Minigenes to confirm exon skipping mutations. Methods Mol Biol 2012, 867:37−47
12. Członkowska A, Tarnacka B, Möller JC, Leinweber B, Bandmann O, Woimant F, Oertel WH: Unified Wilson's Disease Rating Scale - a proposal for the neurological scoring of Wilson's disease patients. Neurol Neurochir Pol 2007, 41:1−12
13. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL: 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. Genet Med 2015, 17:405−424
14. Zhang Z, Pan Z, Ying Y, Xie Z, Adhikari S, Phillips J, Carstens RP, Black DL, Wu Y, Xing Y: Deep-learning augmented RNA-seq analysis of transcript splicing. Nat Methods 2019, 16:307−310
15. Dong Y, Wang R-M, Yang G-M, Yu H, Xu W-Q, Xie J-J, Zhang Y, Chen Y-C, Ni W, Wu Z-Y: Role for biochemical assays and Kayser-Fleischer rings in diagnosis of Wilson's disease. Clin Gastroenterol Hepatol 2021, 19:590−596
16. Riolo G, Cantara S, Ricci C: What's wrong in a jump? prediction and validation of splice site variants. Methods Protoc 2021, 4:62
17. Shaikh SS, Nahorski MS, Rai H, Woods CG: Before progressing from "exomes" to "genomes"… don't forget splicing variants. Eur J Hum Genet 2018, 26:1559−1562
18. Lord J, Gallone G, Short PJ, McRae JF, Ironfield H, Wynn EH, Gerety SS, He L, Kerr B, Johnson DS, McCann E, Kinning E, Flinter F, Temple IK, Clayton-Smith J, McEntagart M, Lynch SA, Joss S, Douzgou S, Dabir T, Clowes V, McConnell VPM, Lam W, Wright CF, FitzPatrick DR, Firth HV, Barrett JC, Hurles ME: Deciphering Developmental Disorders study. Pathogenicity and selective constraint on variation near splice sites. Genome Res 2019, 29:159−170
19. Chen HI, Jagadeesh KA, Birgmeier J, Wenger AM, Guturu H, Schelley S, Bernstein JA, Bejerano G: An MTF1 binding site disrupted by a homozygous variant in the promoter of ATP7B likely causes Wilson disease. Eur J Hum Genet 2018, 26:1810−1818
20. Mameli E, Lepori MB, Chiappe F, Ranucci G, Di Dato F, Iorio R, Loudianos G: Wilson's disease caused by alternative splicing and Alu exonization due to a homozygous 3039-bp deletion spanning from intron 1 to exon 2 of the ATP7B gene. Gene 2015, 569:276−279
21. Jung H, Lee KS, Choi JK: Comprehensive characterisation of intronic mis-splicing mutations in human cancers. Oncogene 2021, 40: 1347−1361
22. Wang X, Zhang Y, Ding J, Wang F: mRNA analysis identifies deep intronic variants causing Alport syndrome and overcomes the problem of negative results of exome sequencing. Sci Rep 2021, 11:18097
23. Anna A, Monika G: Splicing mutations in human genetic disorders: examples, detection, and confirmation. J Appl Genet 2018, 59:253−268
24. Zeng Z, Aptekmann AA, Bromberg Y: Decoding the effects of synonymous variants. Nucleic Acids Res 2021, 49:12673−12691
25. Frisso G, Detta N, Coppola P, Mazzaccara C, Pricolo MR, D'Onofrio A, Limongelli G, Calabrò R, Salvatore F: Functional studies

and in silico analyses to evaluate non-coding variants in inherited cardiomyopathies. Int J Mol Sci 2016, 17:E1883

26. Panzer M, Viveiros A, Schaefer B, Baumgartner N, Seppi K, Djamshidian A, Todorov T, Griffiths WJH, Schott E, Schuelke M, Eurich D, Stättermayer AF, Bomford A, Foskett P, Vodopiutz J, Stauber R, Pertler E, Morell B, Tilg H, Müller T, Kiechl S, Jimenez-Heredia R, Weiss KH, Hahn SH, Janecke A, Ferenci P, Zoller H: Synonymous mutation in adenosine triphosphatase copper-transporting beta causes enhanced exon skipping in Wilson disease. Hepatol Commun 2022, 6:1611−1619
27. Shimo T, Maruyama R, Yokota T: Designing effective antisense oligonucleotides for exon skipping. Methods Mol Biol 2018, 1687:143−155