

ORIGINAL ARTICLE

Mutational analysis of *ATP7B* in north Chinese patients with Wilson disease

Kui Li¹, Wei-Min Zhang², Sheng Lin^{1,6}, Lu Wen^{1,7}, Zi-Feng Wang¹, Dan Xie³, Min Wei⁴, Zheng-Qing Qiu⁴, Yi Dai⁵, Marie C M Lin³, Hsiang-Fu Kung³ and Feng-Xia Yao²

Wilson disease (WD) is an autosomal recessive inherited disease caused by abnormalities of the copper-transporting protein encoding gene *ATP7B*. In this study, we examined *ATP7B* for mutations in 114 individuals of Chinese Han population living in north China who were diagnosed as WD. Totally, we identified 36 mutations and 11 single-nucleotide polymorphisms (SNPs), of which 14 mutations have never been reported previously and 5 were firstly described in Chinese. Among these, p.R778L (21.5%), p.A874V (7.5%) and p.P992L (6.1%) were the most frequent mutations. A genotype of p.L770L + p.R778L + p.P992L was the most frequent triple mutations and two pairs of mutations, p.L770L/p.R778L and p.A874V/p.I929V, were closely related. In addition, a database was established to summarize all *ATP7B* mutations, including those reported previously and those identified in this study. Popular algorithms were used to predict the functional effects of these mutations, and finally, by comparative genomics approaches, we predicted a group of mutation hot spots for *ATP7B*. Our study will broaden our knowledge about *ATP7B* mutations in WD patients in north China, and be helpful for clinical genetic testing.

Journal of Human Genetics (2013) 58, 67–72; doi:10.1038/jhg.2012.134; published online 13 December 2012

Keywords: ATP7B; Chinese; genotype; mutational analysis; susceptibility; Wilson disease

INTRODUCTION

Wilson disease (WD/WND, NCBI MIM No. 277900) is an autosomal recessive inherited disease characterized by excess accumulation of intracellular hepatic copper with subsequent hepatic and neurologic abnormalities. It was first reported by Dr Samuel Alexander Kinnier Wilson in 1912 and since then a growing number of cases were identified worldwide.¹ It is estimated that the disease frequency is about 1/5000 to 1/30 000 and the carrier frequency is even higher.² WD patients have a very poor prognosis if not diagnosed at an early stage and treated appropriately. A life-span treatment is required for WD patients with severe symptom or early onset age, and an early diagnosis and therapy will remarkably improve life quality with a hope of complete symptomatic recovery.^{3–5} Therefore, quick and accurate genetic testing for WD patients and efficient screening for carriers in the population are urgently required.

WD is one of the most prevalent inherited diseases caused by abnormalities of a single gene, the *ATP7B*, which encodes a copper-transporting ATPase 2 β protein.^{3,6} It is located at 13q14.3, spanning a genomic region of ~78 kb and including 21 exons and 20 introns.

ATP7B and its homolog ATP7A are the two main copper-transporting proteins.⁷ Defects of ATP7A impair the entry of copper from enterocytes into circulation and thereby cause deficiency of copper, known as the Menkes disease, an x-linked disorder.⁸ The role of ATP7B is quite different and it mainly functions as transforming apoceruloplasmin into ceruloplasmin and excreting copper into the biliary canaliculi. Defects of ATP7B reduce the blood ceruloplasmin and are harmful to hepatocytes. Many other organs, such as the brain and kidney, can also be affected.⁹

Up to now, more than 500 forms of genetic mutations in *ATP7B* have been reported and there are 1343 known single-nucleotide polymorphisms (SNPs), suggesting that *ATP7B* is susceptible to mutation.¹⁰ In this study, we identified 36 mutations from 114 individuals in north China Han populations who were diagnosed as WD, of which 14 mutations have never been reported previously and 5 are firstly described in the Chinese. We also performed a bioinformatic analysis to predict the functional effects of these mutations, and predicted a group of amino-acid residues of the *ATP7B* protein to be mutation hot spots.

¹Laboratory of Integrated Biosciences, Life Science School, Sun Yat-sen University, Guangzhou, Guangdong, China; ²Clinical Research Lab, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; ³State Key Laboratory of Oncology in South China, Cancer Center, Sun Yat-sen University, Guangzhou, China; ⁴Department of Pediatrics, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China and ⁵Department of Neurology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
⁶Current address: Laboratory of Medical Genetics, Research Institute of Population and Family Planning, Shenzhen, China.

⁷Current address: Biodynamic Optical Imaging Center, College of Life Sciences, Peking University, Beijing, China.

Correspondence: Dr F-X Yao, Clinical Research Lab, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical college, Beijing 100730, China.

E-mail: fxyao@hotmail.com

Received 7 April 2012; revised 19 September 2012; accepted 21 October 2012; published online 13 December 2012

MATERIALS AND METHODS

Subjects

A total of 114 individuals, who belonged to the Han population from north China and were diagnosed as WD, were recruited in this study, of which 76 were males and 38 were females, with onset ages from 7 to 44 years and a median age of 25. One hundred and eleven unrelated individuals and three siblings were set as the control, and all of them were fully informed and signed for agreement. All the patients were diagnosed as WD at the Peking Union Medical College Hospital generally based on the following criteria: (1) liver or brain failure symptoms; (2) presence of K-F ring in the cornea by slit-lamp examination; (3) reduced serum ceruloplasmin ($<0.20\text{ g l}^{-1}$) and/or elevated 24-hour urinary copper excretion ($>1.0\text{ }\mu\text{mol per day}$) and/or hepatic copper content $>250\text{ }\mu\text{g per g}$ of dry weight.¹¹

Data source

Reference *ATP7B* mRNA (NM_000053.3) and protein sequences (NP_000044.2) were retrieved from the NCBI Entrez database, and SNPs were retrieved from the NCBI database (db) SNP (<http://www.ncbi.nlm.nih.gov/snp/>), including all SNPs in the *ATP7B* gene region except the clinical submission items.¹²

Multiple approaches were applied to obtain the known genetic mutations of *ATP7B*, including (1) the Wilson Disease Mutation Database in the University of Alberta, maintained by the Cox Lab (<http://www.wilsonsdisease.med.ualberta.ca/database.asp>);¹⁰ (2) the UniProt (Universal Protein Resource) accession of *ATP7B* protein P35670 (<http://www.uniprot.org/uniprot/P35670>); (3) the Human Gene Mutation Database (HGMD) Professional at the Institute of Medical Genetics in Cardiff, a product of BIOBASE (<http://www.hgmd.org/>).¹³ HGMD is a commercial knowledge base service, being established and maintained manually, and the last update was at 30 September 2011. This database is thought to be the most precise and complete collection of the *ATP7B* gene variants. All the missense/nonsense mutations with an accession in the above databases were merged into a new set (Supplementary Table 1) and other forms of mutations, such as insert, deletion, indels, synonymous and noncoding mutations, were mainly based on the HGMD database annotation.

DNA extraction, PCR and sequencing

Genomic DNAs were extracted from the peripheral blood leukocytes of the patients by QIAamp blood kits (Qiagen, Hilden, Germany) according to the manufacturer's instruction. Eleven exons (2, 5, 8–12, 13, 16, 18 and 19) of the *ATP7B* gene and their flanking intron sequences, which are a total of 5687 bp in length, were amplified by PCR. These exons were selected because most *ATP7B* mutations in the Chinese WD patients are found to be located in them.^{14–18} Primer sets were designed by MutScreener or retrieved from published papers, and the sequences and the product length are described in Supplementary Table 1. Briefly, PCR was performed in a 25- μl reaction volume containing 20–100 ng genomic DNAs and 10 μM of each primer. Initial denaturation was at 95 °C for 5 min, followed by amplification for 30 cycles with denaturation at 95 °C for 30 s, annealing at 60–65 °C for 30 s and extension at 72 °C for 45 s. PCR products were subsequently purified using PCR purification kits (Guangzhou Dongsheng Biotech Ltd, Guangzhou, China) and sequenced bidirectionally using a dye-terminator cycle sequencing ready reaction kits (ABI PRISM 3130, Applied Biosystems, Life Technologies, Foster City, CA, USA).

Sequencing results were aligned to the *ATP7B* reference sequence (NM_000053.3) to find out mutations. To reduce false-positive discovery, (1) high-fidelity Taq enzyme was used for amplification and the reaction condition for PCR was optimized; (2) each sample was sequenced at least twice in order to validate the result.

Identification of novel mutations

A genetic variant is considered to be a novel mutation when it fits all of the following conditions: (1) it is not recorded in the dbSNP database as a polymorphism, (2) it is not included in the merged data set established by us and (3) it has not been reported in PubMed literatures.

Allele frequency and heterozygosity for each mutation were calculated by perl-script written by us, in order to find prevalent mutations and hotspot.

Meanwhile, allele frequency and heterozygosity of SNPs identified in this study were also analyzed and subsequently compared with those obtained from the dbSNP database.

A self-writing program was used for genotyping and paired relevance analysis of the identified mutations.

Computational prediction of deleterious variants of ATP7B

To predict the deleterious effects of a mutation to the structure and function of the protein, we applied three popular programs: PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/index.shtml>), PMut (<http://mmb2.pcb.ub.es:8080/PMut/>) and SNAP (<http://roslab.org/services/snap/>).^{19–22} To predict deleterious SNPs located in the introns, upstream and downstream of the coding region, the FastSNP program was used (http://fastsnp.ibms.sinica.edu.tw/pages/input_CandidateGeneSearch.jsp).¹⁷ All the programs were run by their default parameters with optimization.

Predication of mutation hot spots for ATP7B

ATP7B protein sequences of 13 non-human organisms, including monkey (XP_001103242.2), rat (NP_036643.2), mouse (NP_031537.2), dog (NP_001020438.1), cattle (XP_002691840.1), sheep (NP_001009732.1), horse (XP_001488500.1), duck (XP_001378265.2), chicken (XP_417073.3), panda (XP_002917723.1), platypus (XP_001513328.2), lizard (XP_003215344.1) and zebrafish (XP_684415.4), were retrieved from the NCBI database. Multiple sequences alignment of these sequences together with the human *ATP7B* protein sequence was performed by the Clustal x software (<http://www.clustal.org/>) with default parameters.¹⁸ Conservative scores of each residue were calculated and outputted, ranging from 0 to 100. An amino-acid residue with the conservative score >80 was considered to be a high-conservative site and more than five consecutive such sites were defined as a conservative segment.

To identify mutation hot spots for *ATP7B*, all *ATP7B* missense/nonsense mutations, including those identified in this study and those reported previously, were mapped into the conservative track. We defined a mutation hot spot when (1) there are more than one form of missense/nonsense mutations for a high-conservative site, or (2) there is only one form of mutation, but it is within a conservative segment.

RESULTS

Identification of novel and/or prevalent ATP7B genetic variants of north China patients with WD

By PCR, sequencing and bioinformatic analysis, we examined 11 exons (exon 2, 5, 8, 9, 10, 11, 12, 13, 16, 18 and 19) of the *ATP7B* genes of 114 WD patients of the Chinese Han population living in north China. Totally, we identified 47 genetic variations, of which 36 were mutations and 11 were SNPs, which can be found in the dbSNP database.

Among the 36 mutations, there were 26 missense mutations, one nonsense mutation, three synonymous mutations, four indels and two nucleotide changes in the flanking introns that may disturb splicing (Table 1). Exon 8 was most frequent to suffer a mutation, followed by exon 13 and exon 18. These exons seem to be more susceptible to mutation in WD patients of the Chinese Han population living in north China, consistent with previous studies performed in patients living in other Chinese districts.^{23,24} Fourteen mutations were found to be located at the transmembrane segments-encoding region and four at the copper binding domain, and especially the mutation Mut-02 was exactly located at the copper binding residue. Some other mutations were located at the ATP-binding or catalyzing region. The mutant-allele frequencies were calculated; p.R778L and its linked p.L770L (Mut-13 and Mut-12) were the top prevalent mutations in this study and accounted for 38.6% cases, which is consistent with the previous studies in Asians including the Chinese.²⁴ p.A874V (Mut-17) was the third prevalent

Table 1 Information of *ATP7B* mutations

Symbol	Location	Functional region	NT change	AA change	Conservative score	Mutation type	Allele count	Allele freq.(%)	Patient count	Patient freq.(%)	Novelty ^a
Mut-01	Exon 2		c.130T>A	p.Y44N	32	Missense	1	0.4	1	0.9	W
Mut-02	Exon 2	MBD2 ^b	c.470G>T	p.C157F	100	Missense	1	0.4	1	0.9	W
Mut-03	Exon 2	MBD4	c.1162C>T	p.Q388X(Stop)	25	Nonsense	1	0.4	1	0.9	W
Mut-04	Exon 2	MBD4	c.1185C>T	p.A395A	73	Synonymous	1	0.4	1	0.9	W
Mut-05	Exon 5		c.1708-5T>G	—	N.A.	Splicing	1	0.4	1	0.9	C
Mut-06	Exon 5	MBD6	c.1817T>G	p.V606G	74	Missense	1	0.4	1	0.9	W
Mut-07	Exon 8	TMS2	c.2145C>T	p.Y715Y	100	Synonymous	2	0.9	2	1.8	C
Mut-08	Exon 8	TMS3	c.2195T>A	p.L732H	100	Missense	1	0.4	1	0.9	W
Mut-09	Exon 8	TMS3	c.2195T>C	p.L732P	100	Missense	1	0.4	1	0.9	W
Mut-10	Exon 8		c.2267C>G	p.A756G	100	Missense	1	0.4	1	0.9	
Mut-11	Exon 8		c.2305insC	—	N.A.	Insertion	1	0.4	1	0.9	C
Mut-12	Exon 8	TMS4	c.2310C>G	p.L770L	100	Synonymous	48	21.1	43	37.7	
Mut-13	Exon 8	TMS4	c.2333G>T	p.R778L	100	Missense	49	21.5	44	38.6	
Mut-14	Exon 8	TMS4	c.2333G>A	p.R778Q	100	Missense	1	0.4	1	0.9	
Mut-15	Exon 9	ATPase	c.2383C>T	p.L795F	100	Missense	1	0.4	1	0.9	C
Mut-16	Exon 11	ATPase	c.2620G>C	p.A874P	61	Missense	1	0.4	1	0.9	
Mut-17	Exon 11	ATPase	c.2621C>T	p.A874V	61	Missense	17	7.5	17	14.9	
Mut-18	Exon 11	ATPase	c.2668G>A	p.V890M	100	Missense	1	0.4	1	0.9	
Mut-19	Exon 12	ATPase	c.2755C>G	p.R919G	71	Missense	3	1.3	3	2.6	
Mut-20	Exon 12	TMS5	c.2763T>A	p.S921R	100	Missense	1	0.4	1	0.9	W
Mut-21	Exon 12	TMS5	c.2785A>G	p.I929V	81	Missense	2	0.9	2	1.8	
Mut-22	Exon 12		c.2828G>A	p.G943D	100	Missense	2	0.9	2	1.8	
Mut-23	Exon 13		c.2866-17G>C	—	N.A.	Splicing	1	0.4	1	0.9	W
Mut-24	Exon 13	TMS6	c.2924C>A	p.S975Y	60	Missense	4	1.8	4	3.5	
Mut-25	Exon 13	TMS6	c.2938T>C	p.C980L	29	Missense	2	0.9	1	0.9	W
Mut-26	Exon 13	TMS6	c.2939G>A	p.C980Y	29	Missense	1	0.4	1	0.9	
Mut-27	Exon 13	TMS6	c.2960T>C	p.L987P	61	Missense	1	0.4	1	0.9	W
Mut-28	Exon 13	TMS6	c.2975C>T	p.P992L	62	Missense	14	6.1	14	12.3	
Mut-29	Exon 16	ATP bind	c.3451C>T	p.R1151C	60	Missense	1	0.4	1	0.9	C
Mut-30	Exon 16	ATP bind	c.3532A>G	p.T1178A	60	Missense	1	0.4	1	0.9	
Mut-31	Exon 18		c.3700-1del G	—	N.A.	Deletion	1	0.4	1	0.9	
Mut-32	Exon 18		c.3765insAC	—	N.A.	Insertion	1	0.4	1	0.9	W
Mut-33	Exon 18	HAD	c.3797G>A	p.G1266E	63	Missense	1	0.4	1	0.9	W
Mut-34	Exon 18	HAD	c.3809A>G	p.N1270S	60	Missense	4	1.8	4	3.5	
Mut-35	Exon 18	HAD	c.3818C>T	p.P1273L	62	Missense	1	0.4	1	0.9	
Mut-36	Exon 19	TMS7	c.4006delA	—	N.A.	Deletion	3	1.3	3	2.6	W

Abbreviations: AA, amino acid; ATPase, copper-(or silver)-translocating P-type ATPase segment; ATP bind, ATP binding segment; HAD, haloacid dehalogenase-like hydrolases segment; MBD, metal-binding domain; NT, nucleotide; TMS, transmembrane segment.

^aNovelty 'W' indicated a novel mutation that has not been reported previously and 'C' indicated a mutation firstly identified in Chinese.

^bRight at the heavy metal binding site.

mutation and p.P992L (Mut-28) the fourth. Others were rare mutations with no more than three cases.

We applied an approach to determine whether a mutation was novel or not (see the Materials and methods). We found that 14 mutations identified in this study have never been reported previously, and 5 mutations were firstly identified in the Chinese (Table 1, and Supplementary Figure 1 for electropherogram of the mutations). A user database containing all the missense/nonsense mutations reported previously and those identified in this study are listed in Supplementary Table 2.

The minor allele frequency and heterozygosity of the SNPs identified in this study were compared with those in the NCBI dbSNP annotations (Supplementary Table 3). We found that most of the SNPs identified in this study had lower minor allele frequency and

heterozygosity in individuals recruited in this study than in the whole population. This may be due to the differences of allele frequency among ethnic populations.

Genotyping the mutations of each WD patient and identification of the linked mutations

Genotypes of each WD patient are summarized in Supplementary Table 4. On average, one patient harbored 0 to 3 mutations in the genomic region we examined. In detail, 24 patients harbored three mutations, and in these patients, p.L770L, p.R778L and p.P992L were the most prevalent genotype. In all, 35 patients harbored 2 mutations, 21 harbored 1 mutation and 34 harbored no mutations. Homozygous mutations were only found in six patients, and five of them were p.L770L and p.R778L homozygote and one was p.C980L homozygote.

In order to investigate the potential linkage between mutations, the paired emerging of the mutations with an allele count of no less than two was analyzed. We found that two pairs of mutations, p.L770L/p.R778L and p.A874V/p.I929V, were closely related (Figure 1). A full

p.L770L	43							
p.R778L	43	44						
p.A874V	6	6	17					
p.R919G	2	2	0	3				
p.I929V	0	0	2	0	2			
p.P992L	7	7	0	0	0	14		
p.N1270S	2	2	0	0	0	0	4	
c.4006delA	3	3	0	0	0	0	0	3
	p.L770L	p.R778L	p.A874V	p.R919G	p.I929V	p.P992L	p.N1270S	c.4006delA

Figure 1 Paired relevance of mutations. Grids in gray: patient count of specified linked mutations; grids in blue: patient count of a single prevalent mutation; numbers in red: triple linked mutations. Notice that all the observed p.L770L mutations are linked with p.R778L and p.I929V are linked with p.A874V.

map showing the paired relevance of all the mutations found in this study is shown in Supplementary Figure 1.

Predicting the functional effects of mutations and SNPs

We applied three popular programs to predict the functional effects of these mutations and SNPs. PolyPhen-2 predicts the possible impact of an amino-acid substitution on the structure and function of a human protein by using straightforward physical and comparative considerations. SNAP is a neural-network-based method that uses *in silico*-derived protein information (for example, . secondary structure, conservation, solvent accessibility, etc.) in order to make predictions regarding the functionality of mutated proteins. It returns a score for each substitution. These scores can then be translated into binary predictions of effect (present/absent) and reliability indices . PMut is based on the use of different kinds of sequence information to label mutations, and neural networks to process this information to estimate whether a mutation is pathological (that is, it can lead to disease for the carrier) or non-pathological/neutral (no effect on the carrier's health). The cross-validated performance of PMut is 84% overall success rate and 67% improvement over random. The combined prediction results are shown in Table 2. Notice that the nonsense mutation that truncated the protein sequence significantly was commonly considered to be most deleterious, as well as the one identified in our study, p.Q388X.

Table 2 Prediction of the functional effects of the mutations

Mutation	Novelty	PMut ^a			SNAP ^b			PolyPhen-2 ^c	
		Prediction	Reliability	Score	Prediction	Reliability	Accuracy (%)	Prediction	Score
p.Y44N	W	—	4	0.26	+	0	58	++	0.99
p.C157F	W	+	4	0.71	+	6	93	++	1.00
p.Q388X	W	—	—	—	—	—	—	—	—
p.V606G	W	—	5	0.23	+	5	87	++	0.99
p.L732H	W	—	8	0.10	+	5	87	++	1.00
p.L732P	W	—	6	0.20	+	6	93	++	1.00
p.A756G	—	—	7	0.11	+	1	63	++	1.00
p.R778L	—	—	3	0.32	+	3	78	++	1.00
p.R778Q	—	+	1	0.55	+	3	78	++	1.00
p.L795F	C	—	6	0.17	+	4	82	++	1.00
p.A874P	—	+	7	0.86	+	4	82	++	1.00
p.A874V	—	+	0	0.52	+	4	82	++	1.00
p.V890M	—	—	2	0.39	+	3	78	++	1.00
p.R919G	—	—	3	0.31	+	2	70	+	0.83
p.S921R	W	+	3	0.65	+	5	87	++	1.00
p.I929V	—	—	9	0.04	—	3	78	—	0.03
p.G943D	—	+	3	0.65	+	6	93	++	1.00
p.S975Y	—	+	4	0.72	—	0	53	++	1.00
p.C980L	W	+	0	0.54	+	1	63	+	0.95
p.C980Y	—	+	8	0.91	+	2	70	++	1.00
p.L987P	W	+	0	0.55	+	6	93	++	1.00
p.P992L	—	+	0	0.55	+	4	82	++	1.00
p.R1151C	C	+	8	0.91	+	4	82	++	1.00
p.T1178A	—	—	3	0.32	+	3	78	++	0.91
p.G1266E	W	+	5	0.78	+	7	96	++	1.00
p.N1270S	—	—	1	0.43	+	5	87	++	1.00
p.P1273L	—	+	7	0.87	+	3	78	++	1.00

^aPMut: predictions +: pathological and -: neutral; reliability ranged from 0 to 10, higher is better; score ranged from 0 to 1, higher is better.

^bSNAP: prediction +: non-neutral and -: neutral; reliability ranged from 0 to 10, higher is better; accuracy ranged from 0 to 100%, higher is better.

^cPolyPhen-2: ++: probably damage and +: possibly damage and -: benign; score ranged from 0 to 1, higher is better.

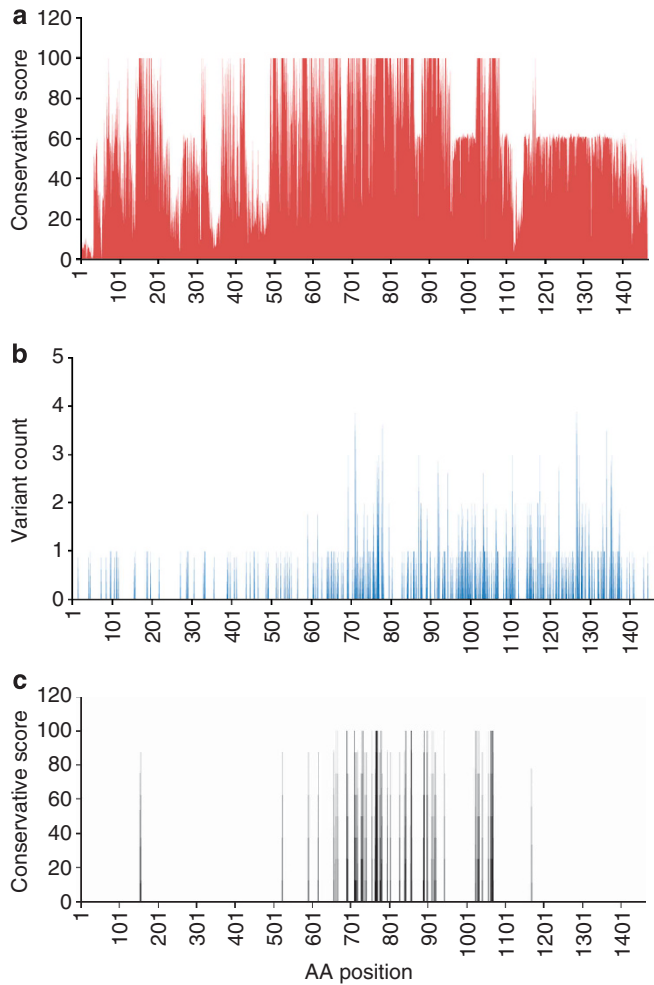


Figure 2 Tracks presenting the conservation and mutation hot spots for *ATP7B*. (a) Conservative score of *ATP7B* full length calculated by multiple sequences alignments among 14 species (ranged from 0 to 100). (b) All currently known missense/nonsense mutations of *ATP7B* accompanied with our findings are mapped into the protein sequence. Vertical coordinates indicate the count of variant forms of one AA residue site. (c) Merged tracks indicate the mutation hot spots for *ATP7B* in WD patients.

Meanwhile, a comprehensive analysis to predict the functional effects of the SNPs were run by the FastSNP program, and 28 SNPs were predicted to be functional or deleterious (Supplementary Table 5).

Prediction of mutation hot spots for *ATP7B*

Multiple sequences alignment of *ATP7B* protein sequences of human and other 13 non-human species were performed, and conservation of each amino-acid residue was quantitatively scored from 0 to 100 (Figure 2a and Supplementary Table 6). All *ATP7B* missense/nonsense mutations, including those reported previously and those identified in this study, were mapped into this conservative track to show the mutant forms counts (Figure 2b).

Totally, we identified 34 mutation hot spots, which are schematically shown in Figure 2c and listed in Supplementary Table 7. Mutations of these hot spots should also be more deleterious to protein structure and functions than those of other sites, and therefore should be preferentially selected for genetic testing of WD.

The reported mutation hot spot p.R778L of East Asians and p.H1069Q were both included in this group, verifying our strategy.²⁵

DISCUSSION

WD is known to be caused by abnormalities of the copper-transporting protein encoding gene *ATP7B*. Nonsense mutations of *ATP7B* can result in truncated proteins with severe defects of copper transporting, and finally lead to serious syndrome and an early onset age.²⁶ Missense mutations influence protein structure and can be found in several functional domains, including those functioning as an interaction between *ATP7B* and its chaperone proteins.²⁷ Splicing and transcription regulation can also be disturbed by the mutations located in the introns and the promoter region.^{28,29} All these kinds of mutations are identified in this study. It is interesting that some synonymous mutations are also linked to WD, such as c.2310C>G, p.L770L, which is almost completely linked with the most prevalent mutation p.R778L identified in this study and reported in WD patients of the Hong Kong Chinese.²⁴ The p.L770L mutation was rare in the normal population but appeared frequently in WD patients, being found in more than one-third of the patients we examined. However, this silent mutation seems to be a polymorphism in European population. A similar mutation is p.I1148T, which is found to be a polymorphism in our study, but is only found in WD patients in France.³⁰ Further expression analysis will be helpful to elucidate the functions of these mutations.

A growing number of *ATP7B* mutations have been identified in WD patients and some mutation hot spots were identified among races and areas. For example, p.R778L was the most prevalent mutation identified in East Asians with WD, with a frequency of up to 60%.^{31–33} In this study, we found that p.R778L (21.5%) and p.P992L (6.1%) were the most frequent mutations, which is consistent with several previous reports in the Chinese WD patients.^{14–18} The p.A874V (7.5%) and p.N1270S (1.8%) were also among the most frequent mutations in this study, which was reported in two groups of patients from Korea³³ and Shanghai, China,¹⁴ suggesting similar *ATP7B* mutation spectrum among these districts and ours.^{14,34} Also consistent with the previous reports, we did not found the H1069Q mutation in this study, which is a common WD mutation in Europeans and North Americans,^{35–37} and has not been reported in Chinese WD patients.^{14–18} Only six homozygous mutations were found in our study, five of which were p.L770L and p.R778L homozygote, which are considered to be less deleterious but will lead to early onset of the disease. The other homozygote was p.C980L, which was a novel mutation identified in this study. It was predicted to be deleterious by all the three programs and it will be interesting to investigate the functional effects of this mutation in the future. It should be noted that in this study, the rate of patients with no mutations is relatively high (34/114, 29.8%), and this may be due to the fact that we only sequenced 11 out of 21 exons of the *ATP7B* gene. More mutations are expected to be found when more exons, as well as the regulatory regions, are sequenced in the future.

A group of 28 frequently observed mutations were suggested to be used for rapid diagnosis of WD.³⁸ In this study, we applied a different strategy to find out a group of 34 mutation hot spots. It should be noted that two groups overlap by 5 sites and 29 sites in our group are not included in the previous panel. Thus, our predication should be a useful supplement for clinical genetic testing of WD. Recently, the *ATP7B* gene was found to be associated with sporadic Alzheimer's disease and some cancers, and thus the genetic testing of *ATP7B* should also benefit studies of these diseases.^{39–43}

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (No. 81000252).

- 1 Compston, A. Progressive lenticular degeneration: a familial nervous disease associated with cirrhosis of the liver, by S. A. Kinnier Wilson, (From the National Hospital, and the Laboratory of the National Hospital, Queen Square, London) *Brain* 1912: 34; 295–509. *Brain* **132**, 1997–2001 (2009).
- 2 Figus, A., Angius, A., Loudianos, G., Bertini, C., Dessi, V., Loi, A. *et al.* Molecular pathology and haplotype analysis of Wilson disease in Mediterranean populations. *Am. J. Hum. Genet.* **57**, 1318–1324 (1995).
- 3 Rosencrantz, R. & Schilsky, M. Wilson disease: pathogenesis and clinical considerations in diagnosis and treatment. *Semin. Liver Dis.* **31**, 245–259 (2011).
- 4 Huster, D. Wilson disease. *Best Pract. Res. Clin. Gastroenterol.* **24**, 531–539 (2010).
- 5 Schilsky, M. L. Diagnosis and long-term management of Wilson disease. *Gastroenterol. Hepatol. (NY)* **3**, 27–29 (2007).
- 6 Bull, P. C., Thomas, G. R., Rommens, J. M., Forbes, J. R. & Cox, D. W. The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. *Nat. Genet.* **5**, 327–337 (1993).
- 7 Harris, E. D. Cellular copper transport and metabolism. *Annu. Rev. Nutr.* **20**, 291–310 (2000).
- 8 Menkes, J. H. Menkes disease and Wilson disease: two sides of the same copper coin. Part I: Menkes disease. *Eur. J. Paediatr. Neurol.* **3**, 147–158 (1999).
- 9 Menkes, J. H. Menkes disease and Wilson disease: two sides of the same copper coin. Part II: Wilson disease. *Eur. J. Paediatr. Neurol.* **3**, 245–253 (1999).
- 10 Kenney, S. M. & Cox, D. W. Sequence variation database for the Wilson disease copper transporter, ATP7B. *Hum. Mutat.* **28**, 1171–1177 (2007).
- 11 Roberts, E. A. & Schilsky, M. L. Diagnosis and treatment of Wilson disease: an update. *Hepatology* **47**, 2089–2111 (2008).
- 12 Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
- 13 Cooper, D. N., Stenson, P. D. & Chuzhanova, N. A. The Human Gene Mutation Database (HGMD) and its exploitation in the study of mutational mechanisms. *Curr. Protoc. Bioinform.* **Ch. 1**, Unit 1.13 (2006).
- 14 Li, X. H., Lu, Y., Ling, Y., Fu, Q. C., Xu, J., Zang, G. Q. *et al.* Clinical and molecular characterization of Wilson's disease in China: identification of 14 novel mutations. *BMC Med. Genet.* **12**, 6 (2011).
- 15 Wang, L. H., Huang, Y. Q., Shang, X., Su, Q. X., Xiong, F., Yu, Q. Y. *et al.* Mutation analysis of 73 southern Chinese Wilson's disease patients: identification of 10 novel mutations and its clinical correlation. *J. Hum. Genet.* **56**, 660–665 (2011).
- 16 Wan, L., Tsai, C. H., Hsu, C. M., Huang, C. C., Yang, C. C., Liao, C. C. *et al.* Mutation analysis and characterization of alternative splice variants of the Wilson disease gene ATP7B. *Hepatology* **52**, 1662–1670 (2010).
- 17 Yuan, H. Y., Chiou, J. J., Tseng, W. H., Liu, C. H., Liu, C. K., Lin, Y. J. *et al.* FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Res.* **34**, W635–W641 (2006).
- 18 Thompson, J. D., Gibson, T. J. & Higgins, D. G. Multiple sequence alignment using ClustalW and ClustalX. *Curr. Protoc. Bioinform.* **Ch. 2**, Unit 2.3 (2002).
- 19 Flanagan, S. E., Patch, A. M. & Ellard, S. Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. *Genet. Test Mol. Biomarkers* **14**, 533–537 (2010).
- 20 Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
- 21 Ferrer-Costa, C., Gelpi, J. L., Zamakola, L., Parraga, I., de la Cruz, X. & Orozco, M. PMUT: a web-based tool for the annotation of pathological mutations on proteins. *Bioinformatics* **21**, 3176–3178 (2005).
- 22 Bromberg, Y., Yachdav, G. & Rost, B. SNAP predicts effect of mutations on protein function. *Bioinformatics* **24**, 2397–2398 (2008).
- 23 Gu, Y. H., Kodama, H., Du, S. L., Gu, Q. J., Sun, H. J. & Ushijima, H. Mutation spectrum and polymorphisms in ATP7B identified on direct sequencing of all exons in Chinese Han and Hui ethnic patients with Wilson's disease. *Clin. Genet.* **64**, 479–484 (2003).
- 24 Mak, C. M., Lam, C. W., Tam, S., Lai, C. L., Chan, L. Y., Fan, S. T. *et al.* Mutational analysis of 65 Wilson disease patients in Hong Kong Chinese: identification of 17 novel mutations and its genetic heterogeneity. *J. Hum. Genet.* **53**, 55–63 (2008).
- 25 Ferenci, P. Regional distribution of mutations of the ATP7B gene in patients with Wilson disease: impact on genetic testing. *Hum. Genet.* **120**, 151–159 (2006).
- 26 Thomas, G. R., Forbes, J. R., Roberts, E. A., Walshe, J. M. & Cox, D. W. The Wilson disease gene: spectrum of mutations and their consequences. *Nat. Genet.* **9**, 210–217 (1995).
- 27 Wernimont, A. K., Yatsunyk, L. A. & Rosenzweig, A. C. Binding of copper(I) by the Wilson disease protein and its copper chaperone. *J. Biol. Chem.* **279**, 12269–12276 (2004).
- 28 Loudianos, G., Lovicu, M., Dessi, V., Tzetis, M., Kanavakis, E., Zancan, L. *et al.* Abnormal mRNA splicing resulting from consensus sequence splicing mutations of ATP7B. *Hum. Mutat.* **20**, 260–266 (2002).
- 29 Cullen, L. M., Prat, L. & Cox, D. W. Genetic variation in the promoter and 5' UTR of the copper transporter, ATP7B, in patients with Wilson disease. *Clin. Genet.* **64**, 429–432 (2003).
- 30 Bost, M., Piguet-Lacroix, G., Parant, F. & Wilson, C. M. Molecular analysis of Wilson patients: direct sequencing and MLPA analysis in the ATP7B gene and Atox1 and COMMD1 gene analysis. *J. Trace Elem. Med. Biol.* **26**, 97–101 (2012).
- 31 Ye, S., Gong, L., Shui, Q. X. & Zhou, L. F. Wilson disease: identification of two novel mutations and clinical correlation in Eastern Chinese patients. *World J. Gastroenterol.* **13**, 5147–5150 (2007).
- 32 Park, S., Park, J. Y., Kim, G. H., Choi, J. H., Kim, K. M., Kim, J. B. *et al.* Identification of novel ATP7B gene mutations and their functional roles in Korean patients with Wilson disease. *Hum. Mutat.* **28**, 1108–1113 (2007).
- 33 Okada, T., Shiono, Y., Hayashi, H., Satoh, H., Sawada, T., Suzuki, A. *et al.* Mutational analysis of ATP7B and genotype-phenotype correlation in Japanese with Wilson's disease. *Hum. Mutat.* **15**, 454–462 (2000).
- 34 Yoo, H. W. Identification of novel mutations and the three most common mutations in the human ATP7B gene of Korean patients with Wilson disease. *Genet. Med.* **4**, 43S–48S (2002).
- 35 Caca, S., Ferenci, P., Kuhn, H. J., Polli, C., Willgerodt, H., Kunath, B. *et al.* High prevalence of the H1069Q mutation in East German patients with Wilson disease: rapid detection of mutations by limited sequencing and phenotype-genotype analysis. *J. Hepatol.* **35**, 575–581 (2001).
- 36 Cauza, E., Ulrich-Pur, H., Polli, C., Gangl, A. & Ferenci, P. Distribution of patients with Wilson disease carrying the H1069Q mutation in Austria. *Wien Klin Wochenschr* **112**, 576–579 (2000).
- 37 Kumar, S., Thapa, B., Kaur, G. & Prasad, R. Analysis of most common mutations R778G, R778L, R778W, I1102T and H1069Q in Indian Wilson disease patients: correlation between genotype/phenotype/copper ATPase activity. *Mol. Cell. Biochem.* **294**, 1–10 (2007).
- 38 Mak, C. M., Lam, C. W., Lai, S. T., Hui, Y. & Tam, S. Rapid diagnosis of Wilson disease by a 28-mutation panel: real-time amplification refractory mutation system in diagnosing acute Wilsonian liver failure. *Clin. Chim. Acta.* **398**, 39–42 (2008).
- 39 Bucossi, S., Mariani, S., Ventriglia, M., Polimanti, R., Gennarelli, M., Bonvicini, C. *et al.* Association between the c. 2495 A>G ATP7B Polymorphism and Sporadic Alzheimer's Disease. *Int. J. Alzheimers Dis.* **2011**, 973692 (2011).
- 40 Martinez-Balibrea, E., Martinez-Cardus, A., Musulen, E., Gines, A., Manzano, J. L., Aranda, E. *et al.* Increased levels of copper efflux transporter ATP7B are associated with poor outcome in colorectal cancer patients receiving oxaliplatin-based chemotherapy. *Int. J. Cancer* **124**, 2905–2910 (2009).
- 41 Fukushima-Uesaka, H., Saito, Y., Maekawa, K., Kurose, K., Sugiyama, E., Katori, N. *et al.* Genetic polymorphisms of copper- and platinum drug-efflux transporters ATP7A and ATP7B in Japanese cancer patients. *Drug Metab. Pharmacokinet.* **24**, 565–574 (2009).
- 42 Xu, W., Cai, B., Chen, J. L., Li, L. X., Zhang, J. R., Sun, Y. Y. *et al.* ATP7B antisense oligodeoxynucleotides increase the cisplatin sensitivity of human ovarian cancer cell line SKOV3ipl. *Int. J. Gynecol. Cancer* **18**, 718–722 (2008).
- 43 Nakagawa, T., Inoue, Y., Kodama, H., Yamazaki, H., Kawai, K., Suemizu, H. *et al.* Expression of copper-transporting P-type adenosine triphosphatase (ATP7B) correlates with cisplatin resistance in human non-small cell lung cancer xenografts. *Oncol. Rep.* **20**, 265–270 (2008).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)