Genetic analysis of ATP7B in 102 south Indian families with Wilson disease

Nivedita Singh¹, Pradeep Kallollimath², Mohd Hussain Shah¹, Saketh Kapoor¹, Vishwanath Kumble Bhat¹, Lakshminarayanapuram Gopal Viswanathan⁵, Madhu Nagappa², Parayil S. Bindu², Arun B. Taly², Sanjib Sinha², Arun Kumar ID¹*

¹ Department of Molecular Reproduction, Development and Genetics, Indian Institute of Science, Bangalore, India, ² Department of Neurology, National Institute of Mental Health and Neuro Sciences, Bangalore, India

* arunk@iisc.ac.in

Abstract

Wilson disease (WD) is an autosomal recessive disorder, characterized by excessive deposition of copper in various parts of the body, mainly in the liver and brain. It is caused by mutations in ATP7B. We report here the genetic analysis of 102 WD families from a south Indian population. Thirty-six different ATP7B mutations, including 13 novel ones [p.Ala58fs*19, p.Lys74fs*9, p.Gln281*, p.Pro350fs*12, p.Ser481*, p.Leu735Arg, p.Val752Gly, p.Asn812fs*2, p.Val845Ala, p.His889Pro, p.Ile1184fs*1, p.Val1307Glu and p.Ala1339Pro], were identified in 76/102 families. Interestingly, the mutation analysis of affected individuals in two families identified two different homozygous mutations in each ATP7B allele. Of 36 mutations, 28 were missense, thus making them the most prevalent mutations identified in the present study. Nonsense, insertion and deletion represented 3/36, 2/36 and 3/36 mutations, respectively. The haplotype analysis suggested founder effects for all the 14 recurrent mutations. Our study thus expands the mutational landscape of ATP7B with a total number of 758 mutations. The mutations identified during the present study will facilitate carrier and pre-symptomatic detection, and prenatal genetic diagnosis in affected families.

Introduction

Wilson disease (WD, MIM #277900) is an autosomal recessive disorder, characterized by the excessive deposition of copper in the body, mainly in the liver and brain. The worldwide disease incidence of WD is 1/5,000–1/30,000 live births. The disease presentation varies from as early as 2 years to as old as 72 years of age [1–3]. WD patients commonly manifest with hepatic and neuropsychiatric problems. The hepatic manifestations are acute hepatitis, chronic active hepatitis, cirrhosis and acute fulminant hepatic failure. Patients with neuropsychiatric manifestations have dystarthritis, dystonia, tremor, ataxia, parkinsonian features, behavioral problems and cognitive disturbances. ATP7B (MIM #606882; ATPase copper transporting beta), the causative gene of WD, is located on the chromosome 13q14.3-q21 [4–6]. It has 21 coding...
exons and codes for a 1,465 amino acid long protein of 165 kDa, which contains following
domains: six copper binding domains (CBD1-6), eight transmembrane domains (TMS1-8), A-
domain and ATP binding domain. It shows granular cytoplasmic expression in most tissues
(https://www.proteinatlas.org/ENSG00000123191-ATP7B/tissue), and resides mainly in the
trans-Golgi network (TGN). Under normal physiological conditions, ATP7B delivers copper
to apoceruloplasmin. As the copper level increases inside the cells, ATP7B traffics to the vesic-
ular compartments and lysosomes to remove excess of copper into the bile [7].

Currently, the diagnosis of WD is based mainly on a combination of different clinical fea-
tures (e.g., the presence of corneal Kayser-Fleischer ring, hepatic and neurological abnormali-
ties) and biochemical tests such as serum ceruloplasmin concentration, 24 hours urinary
copper excretion, hepatic copper determination, and serum copper levels. However, although
the first case of WD was reported in 1912, even after so many years of the disease identifica-
tion, its diagnosis often remains a challenge. Wrong and delayed diagnosis of WD patients is
not uncommon and might affect their outcome. The treatment of WD at present is performed
mainly by administration of chelating agents (eg., penicillamine, trientine and ammonium tet-
raithiomolybdate) or zinc salts (eg, zinc sulphate or zinc acetate) that prevent absorption of
copper in the body. The accurate diagnosis of WD could be enhanced by genetic testing.
Therefore, there is a need to perform a genetic study of WD families for carrier and pre-sym-
tomatic detection, and genotype-phenotype correlation. Several mutations in ATP7B have
been reported from different countries, including India (http://www.wilsondisease.med.
ualberta.ca/database.asp). However, only a few large cohort studies have been documented so
far from India [8–10]. In this study, we have screened the entire coding region of ATP7B in a
large cohort of 102 families from a south Indian population, mainly from the state of Karna-
taka, and identified 36 different mutations in the gene.

Materials and methods

Families

A total of 102 WD families (Figure A in S1 File) were recruited in the Wilson’s disease clinic of
the Department of Neurology, National Institute of Mental Health and Neuro Sciences (NIM-
HANS), Bangalore, Karnataka. The diagnosis was based on the presence of typical clinical fea-
tures, Kayser-Fleischer ring by slit lamp and biochemical tests (viz., low serum copper, low
serum ceruloplasmin and elevated 24 hours urinary copper), which were accompanied by
brain MRI and other tests wherever required (Table A in S1 File). The informed written con-
sent for research was obtained from individuals following the approval of the Institutional Eth-
ics Committee of the National Institute of Mental Health and Neuro Sciences (NIMHANS),
Bangalore. All methods were performed in accordance with the relevant guidelines and
regulations.

Genetic analysis

For genetic analysis, 3–5 ml of peripheral blood was drawn from each individual in a Vacutai-
ner™ EDTA tube (Becton Dickinson, Franklin Lakes, NJ) for genomic DNA isolation, using a
Wizard™ genomic DNA extraction kit (Promega, Madison, WI). Genomic DNA was isolated
from a total of 314 individuals, including 113 affected individuals. To determine if the cause of
WD in these families is due to ATP7B gene (GenBank accession # NM_000053.3) mutation(s),
its entire coding region, including the intron-exon boundaries, was amplified using specific
primers (Table B in S1 File). The 5’and 3’UTRs of the ATP7B gene were screened in patients
without mutations in the coding region using specific primers (Table C in S1 File). For the
mutation identification, PCR products were Sanger sequenced from one affected individual
from each family on an ABIprism A310-automated sequencer (Life Technologies, Carlsbad, CA). Once the mutation was identified, the remaining family members were also sequenced for the presence or absence of the same mutation. To rule out the possibility that the novel variants are also present in the general population, single-strand conformation polymorphism (SSCP) and allele-specific PCR (ASP) (Table D in S1 File) were performed. The ClustalW2 program (http://www.ebi.ac.uk/Tools/msa/clustalw2/) was used for multiple sequence alignment to see the conservation of an amino acid residue affected due to a missense mutation. To determine if novel missense variants identified in the study are mutations, different bioinformatics tools (e.g., SIFT, PolyPhen-2 and Mutation Taster) were used. Other databases such as the 1000 Genomes (http://www.1000genomes.org/) and the ExAC (Exome Aggregation Consortium; http://exac.broadinstitute.org/) were used to confirm if the variants identified in this study are novel. The allele frequency of a mutation was calculated as the number of mutant alleles present in affected individuals/total number of alleles in the affected individuals.

Genotype-phenotype correlation
For the genotype-phenotype correlation, the mutations were divided into two groups: missense and nonsense/insertion/deletion. The mutations present in only homozygous and compound heterozygous states were included in the correlation study. The phenotypes that were included for the study were as follows: age of onset, serum copper, serum ceruloplasmin, 24 hr urinary copper, Kayser-Fleischer ring, dysphagia, dysarthria, tremor, dystonia, writing difficulty, chorea, athetosis, parkinsonism, rigidity, bradykinesia, cerebral atrophy, cerebellar atrophy, brain stem atrophy, giant panda sign, hepatomegaly, splenomegaly, and jaundice. For the phenotypes which are based on the presence or absence of the symptoms, the analysis was performed by assigning ‘+’ for the presence of symptom and ‘-’ for the absence of symptoms. The statistical significance was assessed by two-tailed unpaired t-test with Welch’s correction using the GraphPad PRISM5 software (GraphPad Software Inc., San Diego, CA, USA).

Haplotype analysis
To establish founder effects for recurrent mutations, the haplotype analysis was performed, using microsatellite markers flanking the ATP7B gene (Table E in S1 File) and intragenic SNPs. The marker genotyping was carried out as described by Kumar et al. [11].

Results and discussion
Sanger Sequencing of the entire coding region of ATP7B in 102 WD families identified a total of 36 different mutations in 76/102 (74.51%) families, with a frequency ranging from 0.4% to 8.4% (Table 1). These mutations were missense, nonsense, insertions and deletions (Table 1). Of these, 13 were novel mutations [p.Ala58fs*19, p.Lys74fs*9, p.Gln281*, p.Pro350fs*12, p.Ser481*, p.Leu735Arg, p.Val752Gly, p.Asn812fs*2, p.Val845Ala, p.His889Pro, p.Ile1184fs*1, p.Val1307Glu and p.Ala1339Pro] (Table 1; Figs 1–7). Further, 28/36 (77.8%) mutations were missense, thus making them the most prevalent mutations identified in the present study. Based on this analysis, a mutational landscape was constructed, which depicts the spread of mutations across different exons (Fig 8A) and domains (Fig 8B) of ATP7B. Of 76 families with mutations, 20 families had a single mutation in a heterozygous condition, and thus the second mutation was not identified (Table 1). We used the following criteria to name a novel variant as a mutation. 1) The novel variant was segregating in the family. 2) The novel variant was not observed in 50 normal controls. 3) The novel variant was absent in the 1000 Genomes and ExAC databases. 4) The affected amino acid residue was conserved across species (Fig 9). 5) At least two of the following three mutation prediction programs, PolyPhen-2, Mutation Taster
Table 1. Mutations detected in the ATP7B gene in Wilson disease patients during present study.

<table>
<thead>
<tr>
<th>Sl. #</th>
<th>Mutation (Exon/Intron)</th>
<th>Type of mutation</th>
<th>Novel/reported</th>
<th>Family (zygosity)</th>
<th>Region of protein</th>
<th>Ethnic origin</th>
<th>Frequency of mutation in affected individuals (%)</th>
<th>Frequency of mutation in other population (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>c.172_173insC (p.Ala58fs*19)</td>
<td>E2</td>
<td>Insertion</td>
<td>Novel</td>
<td>Family-B (compound heterozygous with c.3741C&gt;G &amp; Family-90 (homozygous))</td>
<td>CBD1</td>
<td>Indian (South)</td>
<td>1.3</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>c.220_222delA (p.Lys74fs*9)</td>
<td>E2</td>
<td>Deletion</td>
<td>Novel</td>
<td>Family-128 (homozygous)</td>
<td>CBD1</td>
<td>Indian (South)</td>
<td>0.8</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>c.813C&gt;A (p.Cys271*)</td>
<td>E2</td>
<td>Nonsense</td>
<td>Reported</td>
<td>Family-N, Family-Q, Family-62, Family-75, Family-D, Family-69, Family-93 &amp; Family-64 (homozygous); Family-A, Family-34 &amp; Family-W (heterozygous)</td>
<td>CBD3</td>
<td>German, Turkish, Indian (East), Indian (South), Thai, Indian (West) &amp; Indian (East)</td>
<td>8.4</td>
<td>1.2 (German), 1.9 (Turkish), 16 Indian (East), 10 Indian (South), 2.6 (Thai), 20.2 Indian (West) &amp; 24.2 Indian (East)</td>
</tr>
<tr>
<td>4</td>
<td>c.841C&gt;T (p.Gln281*)</td>
<td>E2</td>
<td>Nonsense</td>
<td>Novel</td>
<td>Family-31 (homozygous)</td>
<td>CBD3</td>
<td>Indian (South)</td>
<td>0.8</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>c.1048delC (p.Pro350fs*12)</td>
<td>E2</td>
<td>Deletion</td>
<td>Novel</td>
<td>Between CBD3 &amp; CBD4</td>
<td>Indian (South)</td>
<td>0.8</td>
<td>-</td>
<td>This study</td>
</tr>
<tr>
<td>6</td>
<td>c.1442C&gt;G (p.Ser481*)</td>
<td>E3</td>
<td>Nonsense</td>
<td>Novel</td>
<td>Family-506 (heterozygous)</td>
<td>Between CBD4 &amp; CBD5</td>
<td>Indian (South)</td>
<td>0.4</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>c.1847G&gt;A (p.Arg616Gln)</td>
<td>E5</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-67 (heterozygous)</td>
<td>CBD6</td>
<td>British, Mediterranean, Indian (South) &amp; Chinese–Han</td>
<td>0.4</td>
<td>1 (British), 6 (Mediterranean), 4 Indian (South) &amp; 1 Chinese–Han</td>
</tr>
<tr>
<td>8</td>
<td>c.2128G&gt;A (p.Gly710Ser)</td>
<td>E8</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-56 (heterozygous)</td>
<td>TMS2</td>
<td>European, Middle Eastern, Hunagarian &amp; Indian (East)</td>
<td>0.4</td>
<td>1.78 (European), &lt;1 (Middle Eastern), 1.2 (Hunagarian) &amp; 0.8% Indian (East)</td>
</tr>
<tr>
<td>10</td>
<td>c.2204T&gt;G (p.Leu735Arg)</td>
<td>E8</td>
<td>Missense</td>
<td>Novel</td>
<td>Family-M (homozygous)</td>
<td>TMS3</td>
<td>Indian (South)</td>
<td>0.8</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>c.2255T&gt;G (p.Val752Gly)</td>
<td>E8</td>
<td>Missense</td>
<td>Novel</td>
<td>Family-41B (compound heterozygous with c.3182G&gt;A)</td>
<td>TMS3</td>
<td>Indian (South)</td>
<td>0.4</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>c.2267C&gt;T (p.Ala756Val)</td>
<td>E8</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-72 (homozygous)</td>
<td>Between TMS3 &amp; TMS4</td>
<td>Italian</td>
<td>0.8</td>
<td>16.6 (Italian)</td>
</tr>
<tr>
<td>SL. #</td>
<td>Mutation</td>
<td>Exon (E)/Intron (I)</td>
<td>Type of mutation</td>
<td>Novel/Reported</td>
<td>Family (zygosity)</td>
<td>Region of protein</td>
<td>Ethnic origin</td>
<td>Frequency of mutation in affected individuals (%)</td>
<td>Frequency of mutation in other population (%)</td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>---------------------</td>
<td>------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>------------------</td>
<td>--------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>13</td>
<td>c.2297C&gt;T (p.Thr766Met)</td>
<td>E8 Missense</td>
<td>Reported</td>
<td>Family-72 (homozygous)</td>
<td>TMS4</td>
<td>British</td>
<td>0.8</td>
<td>0.2 (British)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>14</td>
<td>c.2303C&gt;T (p.Pro768Leu)</td>
<td>E8 Missense</td>
<td>Reported</td>
<td>Family-55 (compound heterozygous with c.3446G&gt;C)</td>
<td>TMS4</td>
<td>Indian (South) &amp; Spanish</td>
<td>0.4</td>
<td>7.4 Indian (South) &amp; 3.3 Spanish</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>15</td>
<td>c.2333G&gt;A (p.Arg778Gln)</td>
<td>E8 Missense</td>
<td>Reported</td>
<td>Family-P &amp; Family-85 (homozygous)</td>
<td>TMS4</td>
<td>Taiwanese, Chinese, Indian (West) &amp; Indian (East)</td>
<td>1.7</td>
<td>4.8 (Taiwanese), 1.5 Chinese, 1 Indian (West) &amp; 1 Indian (East)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>16</td>
<td>c.2435delA (p.Asn812fs2)</td>
<td>E9 Deletion</td>
<td>Novel</td>
<td>Family-77 (heterozygous)</td>
<td>A domain</td>
<td>Indian (South)</td>
<td>0.4</td>
<td>-</td>
<td>This study</td>
</tr>
<tr>
<td>17</td>
<td>c.2534T&gt;C (p.Val845Ala)</td>
<td>E10 Missense</td>
<td>Novel</td>
<td>Family-80 (homozygous)</td>
<td>A domain</td>
<td>Indian (South)</td>
<td>0.8</td>
<td>-</td>
<td>This study</td>
</tr>
<tr>
<td>18</td>
<td>c.2666A&gt;C (p.His889Pro)</td>
<td>E11 Missense</td>
<td>Novel</td>
<td>Family-123 (compound heterozygous with c.3155G&gt;T)</td>
<td>A domain</td>
<td>Indian (South)</td>
<td>0.4</td>
<td>-</td>
<td>This study</td>
</tr>
<tr>
<td>19</td>
<td>c.2690C&gt;T (p.Thr977Met)</td>
<td>E13 Missense</td>
<td>Reported</td>
<td>Family-76 (homozygous)</td>
<td>TMS6</td>
<td>European, British &amp; Indian (East)</td>
<td>0.8</td>
<td>5.3 European, 2.8 British, 0.5 Indian (East) &amp; 1 Indian (East)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>20</td>
<td>c.2998G&gt;A (p.Gly1001Arg)</td>
<td>E13 Missense</td>
<td>Reported</td>
<td>Family-46 &amp; Family-118 (homozygous)</td>
<td>Between TMS6 &amp; ATP binding domain</td>
<td>British &amp; Sardinian</td>
<td>1.7</td>
<td>1.6 (British) &amp; 0.4 (Sardinian)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>21</td>
<td>c.3008C&gt;T (p.Ala1003Val)</td>
<td>E13 Missense</td>
<td>Reported</td>
<td>Family-36, Family-54, Family-83 &amp; Family-100 (homozygous)</td>
<td>Between TMS6 &amp; ATP binding domain</td>
<td>Turkish, Indian (South) &amp; Indian (West)</td>
<td>3.5</td>
<td>1 (Turkish), 9.2 Indian (South) &amp; 11 Indian (West)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>22</td>
<td>c.3053C&gt;T (p.Ala1018Val)</td>
<td>E13 Missense</td>
<td>Reported</td>
<td>Family-63 (compound heterozygous with c.3809A&gt;G); Family-121 (Heterozygous)</td>
<td>Between TMS6 &amp; ATP binding domain</td>
<td>Czech, Chinese–Han &amp; Sardinian</td>
<td>0.8</td>
<td>0.2 (Czech), 0.7 (Chinese–Han) &amp; 1.3 (Sardinian)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>23</td>
<td>c.3155C&gt;T (p.Pro1052Leu)</td>
<td>E14 Missense</td>
<td>Reported</td>
<td>Family-123 (compound heterozygous with c.2666A&gt;C)</td>
<td>ATP binding domain</td>
<td>British</td>
<td>0.4</td>
<td>0.9 (British)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>24</td>
<td>c.3182G&gt;A (p.Gly1061Glu)</td>
<td>E14 Missense</td>
<td>Reported</td>
<td>Family-O, Family-V, Family-58 &amp; Family-101 (homozygous); Family-38 &amp; Family-78 (heterozygous); Family-41B (compound heterozygous with c.2255T&gt;C)</td>
<td>ATP binding domain</td>
<td>Pakistani, Turkish, Indian (South), Indian (East), Indian (North) &amp; Indian (West)</td>
<td>4.8</td>
<td>1.9 (Pakistani), 3 (Turkish), 3.7 Indian (South), 11 Indian (East), 3.3 Indian (North) &amp; 3 Indian (West)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
</tbody>
</table>

(Continued)
Table 1. (Continued)

<table>
<thead>
<tr>
<th>SL #</th>
<th>Mutation</th>
<th>Exon (E)/Intron (I)</th>
<th>Type of mutation</th>
<th>Novel/Reported</th>
<th>Family (zygosity)</th>
<th>Region of protein</th>
<th>Ethnic origin</th>
<th>Frequency of mutation in affected individuals (%)</th>
<th>Frequency of mutation in other population (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>c.3236G&gt;T (p. Cys1079Phe)</td>
<td>E14</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-73 (homozygous)</td>
<td>ATP binding domain</td>
<td>Chinese</td>
<td>0.8</td>
<td>0.6 (Chinese)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>26</td>
<td>c.3301G&gt;A (p. Gly1101Arg)</td>
<td>E15</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-32, Family-43, Family-52, Family-122 &amp; Family-40 (homozygous)</td>
<td>ATP binding domain</td>
<td>Indian (West) &amp; Indian (East)</td>
<td>4.4</td>
<td>3 Indian (West) &amp; 23 Indian (East)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>27</td>
<td>c.3446G&gt;C (p. Gly1149Ala)</td>
<td>E16</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-‘T, Family-35 &amp; Family-47 (homozygous); Family-55 (compound heterozygous with c.2303C&gt;T)</td>
<td>ATP binding domain</td>
<td>Filipino</td>
<td>3.0</td>
<td>0.1 (Filipino)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>28</td>
<td>c.3532A&gt;G (p.Thr1178Ala)</td>
<td>E16</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-59 &amp; Family-91 (homozygous); Family-E, Family-G &amp; Family-44 (heterozygous)</td>
<td>ATP binding domain</td>
<td>Taiwanese, Indian (East), Chinese &amp; Indian (East)</td>
<td>3.0</td>
<td>3.4 (Taiwanese), 0.57 Indian (East), 3.5 (Chinese) &amp; 0.5 Indian (East)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>29</td>
<td>c.3551InsT (p.Ile1184fs)</td>
<td>E16</td>
<td>Insertion</td>
<td>Novel</td>
<td>Family-86 (homozygous)</td>
<td>ATP binding domain</td>
<td>Indian (South)</td>
<td>0.8</td>
<td>-</td>
<td>This study</td>
</tr>
<tr>
<td>30</td>
<td>c.3722C&gt;T (p.Ala1241Val)</td>
<td>E18</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-98 (heterozygous)</td>
<td>ATP binding domain</td>
<td>Indian (South) &amp; Indian (East)</td>
<td>0.4</td>
<td>1.8 Indian (South) &amp; 5 Indian (East)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>31</td>
<td>c.3741C&gt;G (p.His1247Gln)</td>
<td>E18</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-B (compound heterozygous with c.172_173insC); Family-90 (homozygous)</td>
<td>ATP binding domain</td>
<td>Indian (East)</td>
<td>1.3</td>
<td>0.2 Indian (East)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>32</td>
<td>c.3809A&gt;G (p.Asp1270Ser)</td>
<td>E18</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-89 (homozygous); Family-L (compound heterozygous with c.2131G&gt;T) &amp; Family-63 (compound heterozygous with c.3053C&gt;T)</td>
<td>ATP binding domain</td>
<td>Indian (South), Indian (East), Chinese, Indian (North) &amp; Indian (West)</td>
<td>1.7</td>
<td>5.5 Indian (South), 1.7 Indian (East), 0.8 (Chinese), 2.2 Indian (North) &amp; 3 Indian (West)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>33</td>
<td>c.3895C&gt;T (p.Leu1299Phe)</td>
<td>E18</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-U &amp; Family-70 (heterozygous)</td>
<td>Between ATP binding domain &amp; TMS7</td>
<td>Indian (South) &amp; Indian (West)</td>
<td>0.8</td>
<td>1.8 Indian (South) &amp; 3 Indian (West)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
<tr>
<td>34</td>
<td>c.3920T&gt;A (p.Val1307Glu)</td>
<td>E19</td>
<td>Missense</td>
<td>Novel</td>
<td>Family-49 (homozygous)</td>
<td>Between ATP binding domain &amp; TMS7</td>
<td>Indian (South)</td>
<td>0.8</td>
<td>-</td>
<td>This study</td>
</tr>
<tr>
<td>35</td>
<td>c.4015G&gt;C (p.Ala1339Pro)</td>
<td>E19</td>
<td>Missense</td>
<td>Novel</td>
<td>Family-C (heterozygous)</td>
<td>TMS7</td>
<td>Indian (South)</td>
<td>0.4</td>
<td>-</td>
<td>This study</td>
</tr>
<tr>
<td>36</td>
<td>c.4021G&gt;A (p.Gly1341Ser)</td>
<td>E19</td>
<td>Missense</td>
<td>Reported</td>
<td>Family-K (homozygous); Family-Y &amp; Family-57 (heterozygous)</td>
<td>TMS7</td>
<td>Indian (South) &amp; Indian (West)</td>
<td>1.7</td>
<td>3.7 Indian (South) &amp; 1 Indian (West)</td>
<td><a href="http://www.wilsondisease.med.ualberta.ca/">http://www.wilsondisease.med.ualberta.ca/</a></td>
</tr>
</tbody>
</table>

https://doi.org/10.1371/journal.pone.0215779.t001
Fig 1. DNA sequence analysis of individuals from family-90. (Upper panel) Sequencing chromatograms from the affected individual IV-1 and the parent III-1 showing c.172_173insC mutation. Arrows mark the insertion of C in a homozygous state in the affected individual IV-1 and in a heterozygous state in the parent III-1. (Lower panel) Sequencing chromatograms from the affected individual IV-1 and the parent III-1 showing c.3741C>G mutation. Arrows mark the C>G change in a homozygous state in the affected individual IV-1 and in a heterozygous state in the parent III-1. + denotes the wild-type allele. m1 and m2 denote different mutations.

https://doi.org/10.1371/journal.pone.0215779.g001

Fig 2. DNA sequence analysis of individuals from family-128 and family-31. (Upper panel) Sequencing chromatogram of the affected individual IV-5 from family-128. Arrow marks the deletion of A residue in a homozygous state. (Lower panel) Sequencing chromatograms of the affected individual II-3 and the parent I-1 from family-31. Arrows mark the C>T change in a homozygous state in the affected individual II-3 and in a heterozygous state in the parent I-1. + and m denote the wild-type and mutant alleles, respectively.

https://doi.org/10.1371/journal.pone.0215779.g002
and SIFT, predicted a novel missense variant to be disease causing (Table F in S1 File). Interestingly, the mutation analysis of affected individuals in family-90 (Fig 1) and family-72 (Fig 10) identified two different homozygous mutations in each family, and thus each affected individual from these families harbored 2 mutations in each ATP7B allele.

No mutations were detected in the coding or untranslated regions of the ATP7B gene in 26/102 (25.49%) families. Our inability to detect mutations in these families suggests that the mutations could be present in deep intronic regions or there could be large insertions, deletions and duplications which are not amenable to Sanger sequencing. Several non-disease-causing variants were also observed in different families (Table G in S1 File).

Several studies on the mutation analysis of ATP7B in WD patients have been reported from India [8–10, 12–15]. However, only a few large cohort studies have been reported from India. Three such studies by Gupta et al. [8], Mukherjee et al. [10], and Gupta et al. [12], which included 109, 119, and 62 WD families respectively, were reported from the eastern Indian population. Similarly, a study of 52 WD families was reported from the western Indian population [9]. This is the first largest cohort study of 102 WD families from a south Indian population, which has identified 36 different mutations, including 13 novel ones. Moreover, 15/36 known mutations were reported previously from the north, south, east and west Indian populations [9–10, 12, 14–15], whereas 8/36 mutations identified in the present study have never been reported previously from the Indian population (Table 1). The difference in mutations identified from different regions of India might be due to geographic, ethnic, and genetic differences in the Indian population. Therefore, a genetic study performed in a region of India cannot be extrapolated to the other parts of the country.

Fig 3. DNA sequence analysis of individuals from family-60 and family-506. (Upper panel) Sequencing chromatograms from the affected individual V-1 and parent IV-1 from family-60. Arrows mark the deletion of the C residue in a homozygous state in the affected individual V-1 and in a heterozygous state in the parent IV-1. (Lower panel) Sequencing chromatograms from the affected individual II-1 and parent I-1 from family-506. Arrows mark the C>G change in a heterozygous state in the affected individual II-1 and parent I-1. + and m denote the wild type and mutant alleles, respectively.

https://doi.org/10.1371/journal.pone.0215779.g003
Genotype-phenotype correlation studies help in understanding the molecular mechanism of a disease, predict its course of progression and provide better monitoring of the individuals at risk. Our genotype-phenotype analysis suggested that clinical features such as dystonia, rigidity, bradykinesia, hepatomegaly and splenomegaly are significantly more associated with patients having truncating than those with missense mutations (Table H in S1 File). For example, 20/45 (44.44%) patients with missense mutations have dystonia, whereas 11/12 (91.66%) patients with truncating mutation have it (Table H in S1 File). Further, tremor is significantly more associated with patients having missense than those with truncating mutations (Table H in S1 File). No significant difference is observed between patients with missense or truncating mutations and other clinical features (Table H in S1 File).

There are a very few reports on the genotype-phenotype correlations in WD patients [16–18]. One of the best-known examples of the genotype-phenotype correlation in WD patients is the high frequency mutation p.H1069Q, which is associated with late-onset neurological symptoms in the European population [17]. Similar to the present study, no significant correlations were seen between the phenotypes like KF ring and ceruloplasmin levels and genotypes in another study [18]. Interestingly, differences in clinical features were seen even in patients...
Fig 5. DNA sequence analysis of individuals from family-77 and family-80. (Upper panel) Sequencing chromatograms from the affected individual II-2 and parent I-1 from family-77. Arrows mark the deletion of the A residue in the affected individual II-2 and parent I-1 in a heterozygous state. (Lower panel) Sequencing chromatograms from the affected individual IV-1 and the parent II-2 from family-80. Arrows mark the T>C change in a homozygous state in the affected individual IV-1 and in a heterozygous state in the parent II-2. + and m denote the wild type and mutant alleles, respectively.

https://doi.org/10.1371/journal.pone.0215779.g005

Fig 6. DNA sequence analysis of individuals from family-123. (Upper panel) Sequencing chromatogram from the affected individual II-2 from family-123. Arrow marks the A>C change in a heterozygous state in the affected individual II-2. (Lower panel) Sequencing chromatogram from the affected individual II-2 and the parent I-2 from family-123. Arrows mark the C>T change in a heterozygous state in the affected individual II-2 and parent I-2. + denotes the wild type allele. m1 and m2 denote different mutations.

https://doi.org/10.1371/journal.pone.0215779.g006
Fig 7. DNA sequence analysis of individuals from family-86, family-49 and family-C. (Upper panel) Sequencing chromatograms from the affected individual II-2 and the parent I-1 from family-86. Arrows mark the insertion of the A residue in a homozygous state in the affected individual II-2 and in a heterozygous state in the parent I-1. (Middle panel) Sequencing chromatograms of the affected individual IV-1 and the parent II-1 from family-49. Arrows mark the T>A change in a homozygous state in the affected individual IV-1 and in a heterozygous state in the parent II-1. (Lower panel) Sequencing chromatogram from the affected individual II-3 from family-C. Arrow marks the G>C change in a heterozygous state in the affected individual II-3. + and m denote the wild type and mutant alleles, respectively.

https://doi.org/10.1371/journal.pone.0215779.g007

Fig 8. Mutation landscape of the ATP7B gene and protein. (A) The intron–exon structure of the gene. The novel mutations are shown in bold. (B) Different domains of the protein. Abbreviations: aa; amino acid; CBDs, copper binding domains; TMS, transmembrane segment and A-domain (actuator domain). The numbers refer to amino acid positions.

https://doi.org/10.1371/journal.pone.0215779.g008
with the same mutation, suggesting that other factors like environment, genetic, and epigenetic changes might affect the disease outcome [19]. It is interestingly to note that affected individuals from family-35, family-54 and family-73 have a normal ceruloplasmin level of 16, 18 and 28 mg/dL, respectively (Table I in S1 File). However, this finding is not surprising as around 5% of affected individuals homozygous for \textit{ATP7B} mutations have normal ceruloplasmin levels [20].

Founder mutations help in tracing the evolutionary aspect of a disease, evolution of human population, its migration and growth [21]. Interestingly, in a country like India, it is seen that founder mutations contribute more to the disease burden of recessive disorders than consanguinity [22]. This is mainly because of the traditional practice of same-caste marriages in Indian population which contribute to the presence of founder mutations [23]. In the present study, out of the 36 mutations identified, 14 (38.8%) were recurrent (Table 1). This prompted us to look into the possible founder effect for these 14 recurrent mutations. The haplotype analysis, using microsatellite and SNP (single nucleotide polymorphism) markers, suggested that all the 14 mutations have founder effects (Figures B-O in S1 File). For example, family-B and family-90 with p.Ala58fs share a common disease haplotype of 632.8 Kb (Figure B in Fig 9. Conservation of the amino acid residues across different species in ATP7B. Arrows mark the conservation of amino acid residues Leu735, Val752, His889, Val1307, and Ala1339 across different species. GenBank accession numbers of ATP7B are also given.

https://doi.org/10.1371/journal.pone.0215779.g009
Length of the disease haplotype ranged from 45.8 kb for p.Asp1270Ser (Figure M in S1 File) to 698.2 kb for p.Cys271\textsuperscript{\textbullet}, p.Gly711Trp, p.Ala1003Val, p.Gly1061Glu, p.Gly1101Arg, p.Gly1149Ala, p.Thr1178Ala, p.His1247Gln and p.Leu1299Phe mutations (Figures C-D, G-L and N in S1 File). Further, the mutation analysis identified c.813\textsuperscript{C}>A (p.Cys271\textsuperscript{\textbullet}) as the most common mutation with an allele frequency of 8.4% (Table 1). A common haplotype was identified in all the 11 families carrying this mutation (Figure C in S1 File). Our data thus add up to the previous studies from India where c.813\textsuperscript{C}>A (p.Cys271\textsuperscript{\textbullet}) mutation was identified as the most prevalent mutation with an allele frequency of 16%, 10%, and 20.2% in eastern, southern, and western Indian populations respectively [9,12,14].

In summary, this is the first study on the genetic analysis of a large cohort of 102 WD families from a south Indian population, predominantly from the state of Karnataka. Of 36 mutations identified in 76/102 families, 13 (36.1%) were novel, and thus the total number of \textit{ATP7B} mutations increases to 758 (Table J in S1 File). We hope that the mutations identified during the present study will facilitate carrier and pre-symptomatic detection, and prenatal genetic diagnosis in affected families.

**Supporting information**

S1 File. Supporting multiple Figures and Tables.

(PDF)

**Acknowledgments**

We thank the patients and their family members for participation in this study. We are grateful to anonymous reviewers for their valuable suggestions to improve the manuscript.

**Author Contributions**

Conceptualization: Sanjib Sinha, Arun Kumar.
Formal analysis: Nivedita Singh.
Funding acquisition: Sanjib Sinha, Arun Kumar.
Methodology: Nivedita Singh.
Project administration: Arun Kumar.
Supervision: Arun Kumar.
Writing – original draft: Nivedita Singh.
Writing – review & editing: Arun Kumar.

References


