ATP7A Clinical Genetics Resource – A comprehensive clinically annotated database and resource for genetic variants in ATP7A gene

Aditi Mhaske [a], K.V. Dileep [c], Mukesh Kumar [a,b], Mukta Poojary [a,b], Kavita Pandhare [a,b], Kam Y.J. Zhang [c], Vinod Scaria [a,b,*], Binukumar B.K. [a,b,*]

[a] CSIR Institute of Genomics and Integrative Biology, Mathura Road, Delhi 110 025, India
[b] Academy of Scientific and Innovative Research, CSIR-IGIB South Campus, Mathura Road, Delhi, India
[c] Laboratory for Structural Bioinformatics, Center for Biosystems Dynamics Research, RIKEN, 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa 230-0045, Japan

ABSTRACT

ATP7A is a critical copper transporter involved in Menkes Disease, Occipital horn Syndrome and X-linked distal spinal muscular atrophy type 3 which are X linked genetic disorders. These are rare diseases and their genetic epidemiology of the diseases is unknown. A number of genetic variants in the genes have been reported in published literature as well as databases, however, understanding the pathogenicity of variants and genetic epidemiology requires the data to be compiled in a unified format. To this end, we systematically compiled genetic variants from published literature and datasets. Each of the variants were systematically evaluated for evidences with respect to their pathogenicity and classified as per the American College of Medical Genetics and the Association of Molecular Pathologists (ACMG-AMP) guidelines into Pathogenic, Likely Pathogenic, Benign, Likely Benign and Variants of Uncertain Significance. Additional integrative analysis of population genomic datasets provides insights into the genetic epidemiology of the disease through estimation of carrier frequencies in global populations. To deliver a mechanistic explanation for the pathogenicity of selected variants, we also performed molecular modeling studies. Our modeling studies concluded that the small structural distortions observed in the local structures of the protein may lead to the destabilization of the global structure. To the best of our knowledge, ATP7A Clinical Genetics Resource is one of the most comprehensive compendium of variants in the gene providing clinically relevant annotations in gene.

1. Introduction

Copper (Cu) is an essential trace element vital to the health of all living organisms and is crucial for the normal development of the nervous system. It is required for a variety of biological processes in the body including myelin formation and connective tissue synthesis.It is required by cuproenzymes like ceruloplasmin, a major circulating transport protein for copper, lysyl oxidase, an enzyme essential for initiating the crosslinking of connective tissues by oxidizing peptidyl lysine in elastin and collagen [1]. Human body requires copper for its cellular metabolism. The copper transporter 1, CTR1, plays a major role for cellular Cu uptake in the intestinal epithelium and in hepatic cells. Various chaperons such as Antioxidant 1, Copper Chaperone (ATOX1) and, Human Cytochrome *c* oxidase copper chaperone (hCOX17) distribute the Cu to specific proteins or organelles. Cu is transported from enterocytes into blood circulation by ATP7A whereas ATP7B incorporates Cu in liver ceruloplasmin and also mediates biliary excretion of Cu [2].

Various genes are involved in Cu metabolism, such as CTR1 which is a high affinity Cu transporter mediating copper uptake, DMT1 another mediator of Cu uptake. Another important gene involved in Cu transport is ATOX1 which transports Cu from the cytosol to transporters ATP7A and ATP7B. ATP7A and ATP7B are located in the trans golgi network, ATP7A pumps Cu into blood and ATP7B is responsible for maturation of ceruloplasmin [3]. ATP7A located at chromosome Xq21.1 spans about 150 kb of genomic DNA and contains 23 exons. The size of individual coding exons varies between 77 and 726 bp and introns vary in size between 196 bp and approximately 60 kb [4]. As described by [5] ATP7A is a member of the P1B-subfamily of the P-type ATPases.

The gene encodes a transmembrane copper-transporting P-type ATPase. Copper-transporting ATPases (Cu-ATPases) mediate the copper transport, these ATPases are required for the growth, development and also various physiological processes. Cu-ATPases transport copper from the cytosol across cellular membranes, this process reduces the intracellular copper concentration and also controls the copper homeostasis in the body.

Genetic anomalies in the Cu-ATPase encoding ATP7A are associated with a fatal neurological disorder such as Menkes disease (MD), Occipital Horn Syndrome (OHS) which is a mild form of MD and X-linked distal spinal muscular atrophy. ATP7A is ubiquitously expressed in extrahepatic cells and tissues, therefore, systemic defects are caused due to its absence or inactivation, in MD or OHS patients. MD is an X-linked recessive disorder which results in copper deficiency. Two forms of MD have been described: classic MD and mild MD, based on the symptoms mild MD is a less severe form. Classical MD is characterized by neurological defects and peculiar "kinky" hair. MD is caused by mutations in the ATP7A gene. One-third of the MD cases are due to de novo mutations in ATP7A [6]. Copper absorption is highly impaired in MD and intracellular copper trafficking is disrupted by mutations in the ATP7A gene [4]. Similar to MD, OHS is also a rare disorder occurring due to impaired copper absorption. OHS patients are characterized by "occipital horns" which are downward pointing exostosis arising from the occipital bone and certain connective tissue deformities comprising cutis laxa, hernias, joint laxity and bladder diverticula [7]. X-linked distal spinal muscular atrophy type 3 is a rare distal hereditary motor neuropathy. This disease is characterized by weakness of distal muscles of hands and feet, distal motor weakness particularly of the lower limbs [8].

In the present study, we have developed a comprehensive clinically relevant resource for genetic variants in the ATP7A gene by systematic curation and annotation of the genetic variants, with appropriate classification according to ACMG-AMP guidelines. We have also estimated the carrier frequencies of the variants in global populations thus explicating the genetic epidemiology of the disease. In this manuscript, we also present our molecular modelling studies for six variants, to elucidate the stability of these variants and the mutation-structure relationship. To the best of our knowledge, this resource is one of the most comprehensive compendium of variants in the ATP7A gene, with systematic annotation and classification according to the ACMG-AMP guidelines.

2. Materials and methods

2.1. Data curation and contents

The genetic variants in ATP7A associated with diseases were curated from publications indexed in PubMed and PubMed Central databases using formatted queries including the gene name and disease names. The data was systematically entered into a preformatted spreadsheet. Additional variants listed in ClinVar and Locus Specific Databases were also checked and linked to the data curation. The preformatted spreadsheet also collected a number of fields including the genome build, variant loci including the chromosome, Start position, End position, Reference base (Ref), Altered base (Alt), gene and the amino acid change was collected. In addition, the Technique used for identification, Ethnicity, Geographical origin and Population from where the variant was identified were collected. The nomenclature of the variants as per the Human Genome Variation Society guidelines were also generated. In addition, the variant types as defined by nonsense, frameshift, splicing, and deletion/insertion, missense, UTR3′, UTR5′ were also annotated.

2.2. Validation of curated data

All the variants which were curated from the various platforms were rechecked in order to ensure the correctness of the variant positions as well as the variant notations as per the Human Genome Variation Society (HGVS) nomenclature guidelines. Two web based utilities – LUMC Mutalyzer and Variant Validator were utilized.

Mutalyzer (https://mutalyzer.nl/) is a Web interface used for constructing, validating, and transforming sequence variant descriptions. Mutalyzer evaluates sequence variant nomenclature according to the guidelines of the Human Genome Variation Society (HGVS) nomenclature. The DNA tool Position converter was used for finding the missing pieces of data obtained originally. The Position Converter depends on mapping information from the NCBI. The Position Converter converts the variation description positions from the chromosomal position for a specific human genome build i.e GRCh37 / hg19 to a position relative to RefSeq transcript reference sequences present in the local mapping database or vice versa.

Variant Validator (https://variantvalidator.org/) is a web-based variant validation tool which provides an interface which allows the validation of genomic variations published in scientific literature or databases. It is a web interface which is used to describe sequence variants. We validated all the curated variants through Variant Validator, in order to correct the erroneous variants published in the literature sources if any [9].

2.3. Functional annotation of the variants

All the curated variants post-validation were systematically annotated using the ANNOVAR package. The gene based annotations were retrieved for the variants. In addition, Computational annotations of the variants such as deleteriousness-prediction scores such as SIFT, PolyPhen, CADD etc for predicting the deleterious nature of the variants were obtained. In addition, the allele frequencies across ExAC, gnomAD, ESP6500, 1000 genome datasets were retrieved.

2.4. Interpretation of pathogenicity of sequence variants

The variants were systematically analysed and classified as per the guidelines for interpretation of sequence variants as put forward by the ACMG & AMP.

The following attributes were decided on the basis of population data i.e allele frequencies of the variant from the three databases 1000 Genomes Project, ESP6500 exome-sequencing project, ExAC 65,000 exomes data set. BA1 (Allele frequency is >5%); BS1 (Allele frequency in between 0.01 and 0.05); PM2 (Allele frequency absent or <0.0005). The following attributes were decided on the bases of computational data like SIFT, Polyphen and CADD score generated through annovar run. PP3 (If two of the three computational data shows deleterious effect); BP4 (If two of the three computational data shows tolerated/benign effect). The following attributes were decided on the basis of functional data; PS3 (If in vitro or in vivo functional assay support the damaging effect of the variant on the gene or gene product); BS3 (If in vitro or in vivo functional assay support the benign effect of the variant on the gene or gene product); PM1 (Variant must be present in functionally important protein domain) this was evaluated from the pfam protein families database). The following attributes were decided on the basis of Clinvar significance; PP5 (If Clinvar reports the variant as pathogenic or likely pathogenic); BP6 (If Clinvar reports the variant as Benign or likely benign); The following attributes were decided on the bases of segregation data; PP1 (Cosegregation with disease in multiple affected family members); BS4

(Lack of segregation in affected members of a family); PS2 (De novo confirmed); PM6 (De novo assumed); PP4 (Large family segregation); The following attributes were decided on the bases of variants; BP7 (Synonymous variants); PVS1 (Frameshift, Non-sense, Splice-site variants); PP2 (Missense variant in a gene that has a low rate of benign missense variation); BP1 (Missense in gene where only truncating cause disease); Other attributes; PS1 (Same amino acid change, previously described as pathogenic); PM5 (Missense change at a position where a different amino acid change described as pathogenic before [10]. These criteria were used to annotate each variant to the five classes as "pathogenic," "likely pathogenic," "uncertain significance," "likely benign," and "benign" as per the algorithm.

2.5. Comparison of allele frequencies in different global datasets

The variants that were classified as Pathogenic and Likely pathogenic according to the ACMG guidelines, were mapped to different global datasets of 1000 Genomes, ExAC and gnomAD version 2. The allele frequencies in the different subpopulations of the variants mapping to different datasets were compared to the global allele frequency in their respective dataset using Fisher's Exact test. The p values thus obtained were corrected using bonferroni correction method. The allele frequencies in the subpopulations having corrected p value less than 0.05 were defined to be significantly different as compared to the global allele frequency of that variant in the respective dataset.

2.6. Molecular modeling studies

The variants that were classified as "pathogenic and likely pathogenic" according to ACMG guidelines, were subjected to detailed molecular modeling studies to provide a mechanistic explanation for their pathogenicity. Although we have multiple entries with pathogenic and likely pathogenic labels, we restricted our studies only to those that have tertiary structures. To select the variants that have tertiary structures, we systematically matched the position of variants against the structure of Copper-transporting ATPase 1 (ATP7A). Though there are 6 isoforms of ATP7A reported, we chose isoform-4 (contains 1500 amino acids) for our studies since it was reported as a canonical sequence.

2.7. Assessment of structural stability upon mutations

After systematically matching the respective positions of mutations onto the structure, we selected several mutations for our modeling studies. All those mutations underwent structural stability calculations using multiple tools. We assumed a consensus-based ranking would provide a better understanding of the deleterious nature of the mutations via their structural distortions. Hence, we used 9 tools which consist of standalone molecular modeling packages such as FoldX [11], Schrödinger with two differ(ent force fields, OPLS2005 [12] and OPLS3 [13], MOE [14] and web servers such as CUPSAT [15], mCSM [16], SDM [17], iMutant 2.0 [18] and POPMUSIC [19]. The outputs of these tools were categorized as 'destabilizing', 'stabilizing' and 'neutral' based on the reference values of individual tools and a heat map was plotted.

2.8. Molecular dynamics simulations

In order to corroborate the results obtained from the previously used tools, we performed molecular dynamics simulations for 50 ns. We found that the mutation A629P lies in the solution structure of the apo form of the sixth soluble domain of ATP7A (PDB ID: 1YJV) [20]. Similarly, mutations such as R844C, R844H, P852L and G876R lie in the solution structure of the actuator domain of the

ATP7A (PDB ID: 2KIJ) [21]. In the same way, mutation V1180D lies in the solution structure of the nucleotide binding domain of the ATP7A (PDB ID: 2KMV) [22]. To perform MD simulations, we selected the first structure from the ensembles of solution structures and mutations were introduced using COOT [23]. Later, all these mutants along with the wild types (WT) were subjected to MD simulations with Desmond molecular dynamics package by preparing the proteins using the protein preparation wizard of Schrödinger Maestro, where the protein preparation was executed by adding hydrogen atoms, specifying the bond orders, incorporating protonation states for protein residues, optimizing hydrogen bond network and running a short energy minimization with a RMSD cutoff of 0.30 Å. The prepared proteins (both WT and mutants) were soaked into an orthorhombic water box that contained TIP3P water molecules. The size of the box was set in such a way to include the entire protein. Also, the system was neutralized by applying the respective number of Cl- atoms. All MD simulations were performed using NPT ensemble where the pressure and temperature was set to default i.e. 1.01325 bar and 300 K respectively. We recorded the trajectory and energy at every 50 ps and 1.2 ps respectively.

2.9. Database and web server

The database search tool was designed to allow users to easily explore variants in gene ATP7A from the database. The variant data is stored in MongoDB v3.4.10. The data can be accessed through a web interface running on Apache HTTP server using PHP 7.0. The user-friendly web interface for querying the database is coded in PHP 7.0, AngularJS, HTML, Bootstrap 4 and CSS. MongoDB v3.4.10 was used to keep track of data processing through the web interface.

3. Results

3.1. Compendium and classification of genetic variants

The compendium of curated variants encompasses a total of 602 variant entries in the ATP7A gene. These variants were derived from a total of 64 publications [6, 8 and 24–82] and encompassed variants reported from 17 countries. Of the total compendium of variants, a total of 404 variants were unique and a large majority of variants mapped to exons 18,18% (328/404), while a small number were intronic 12% (49/404), splicing 6.18% (25/404) and UTR 0.49% (2/404) variants (Fig. 1A). The variants were also classified by their potential functional implication. A majority of variants were classified as non-synonymous (226/404), while (14/404) were classified as synonymous. (52/404) variants caused a stop-gain, while (31/404) variants caused a frameshift, (3/404) were non-frameshift and (2/404) were insertions. The variant classes are summarised in Fig. 1B.

3.2. Classification of variants as per the ACMG & AMP guidelines

The variants were further reclassified as per the ACMG & AMP guidelines for interpretation of the pathogenicity of variants. All variants were classified into one of the five categories - Pathogenic, Likely Pathogenic, Benign, Likely Benign and Variant of Uncertain Significance (VUS). The classification revealed that 87 variants could be classified as Pathogenic (21.53%), while 56/404 could be classified as Likely pathogenic (13.86%). 22/404 could be classified as likely benign 5.44% and 10/404 as benign (2.47%). A vast majority of variants, 229/404 could only be classified as a VUS (56.68%) for lack of evidence to classify them as Pathogenic or Benign (Fig. 1C). Of the Pathogenic / Likely Pathogenic variants, 136 are
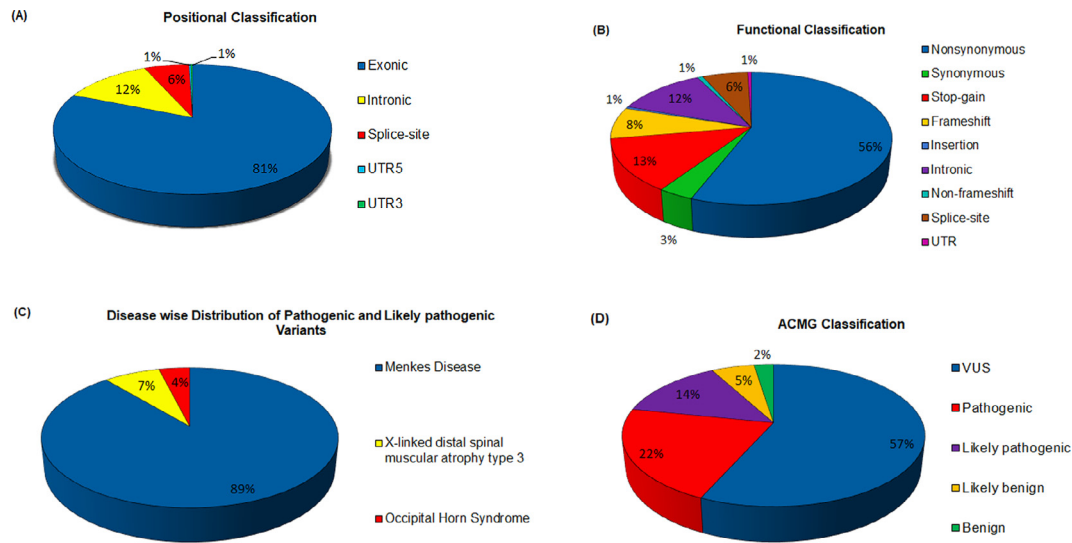
Fig. 1. The variants were annotated based on the context of the (A) genomic loci, (B) the functionality, (C) Distribution by Disease for the Pathogenic and Likely Pathogenic variants and (D) ACMG classification.

associated with Menkes disease, while 11 are associated with X-linked distal spinal muscular atrophy type 3 and 6 are associated with OHS (Fig. 1D).
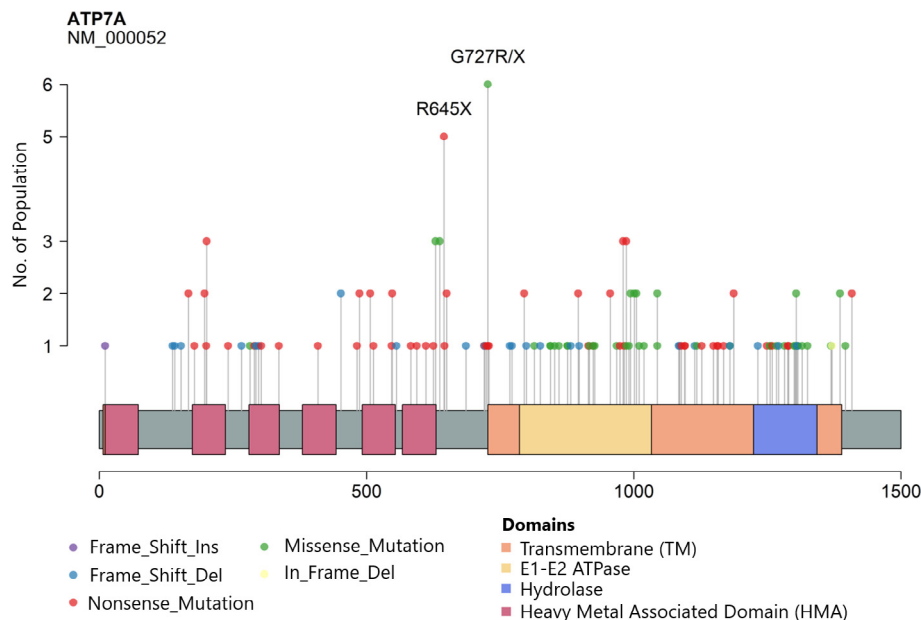
3.3. Distribution of likely pathogenic and pathogenic variants in the ATP7A gene

Our analysis revealed a total of 143 variants which could be classified as pathogenic / likely pathogenic as per the ACMG & AMP guidelines. The variants were systematically mapped to the ATP7A protein structure and domains. A total of four functional protein domains were annotated as per Pfam database using maftools package in r-programming. These included six heavy metal associated domains (HMA), E1-E2 ATPase, HAD like hydrolase and ATPase IB-1. With an exception of a few variants, all the pathogenic/likely pathogenic variants mapped to functional domains of the protein (Fig. 2).

3.4. Global allele frequencies

All the variants annotated as Pathogenic and Likely Pathogenic were further checked in global population genome datasets for their allele frequencies.

The datasets considered include the 1000 Genomes, ExAC and the gnomAD (version 2). Only a total of 5 likely pathogenic variants, of the 143 likely pathogenic and Pathogenic variants, mapped to any of the global datasets. The variant rs782237314 shows frequency of 0.00002 in the gnomAD dataset and ExAC dataset while it was absent in the 1000 Genomes dataset. This variant was only observed in the non-Finnish european subpopulation of both the datasets. The variant rs367775730 shows frequency of 0.0002 in the gnomAD dataset and ExAC dataset while it was absent in the 1000 Genomes dataset. This variant had a significantly high allele frequency in the Ashkenazi jewish subpopulation as compared to the global allele frequency of gnomAD dataset. The variant


Fig. 2. Distribution of pathogenic/likely pathogenic variants in ATP7A protein.

rs374162669 shows a frequency of 0.0003 in 1000 Genomes and 0.0001 in gnomAD and ExAC datasets. The variant rs138958687 shows frequency of 0.0006 in the 1000 Genomes and 0.0007 in the gnomAD and ExAC datasets. This variant has significantly dif-

ferent allele frequency in the South Asian, Non-Finnish european and Finnish subpopulations as compared to the global allele frequency in the gnomad dataset, while in the ExAC dataset, it showed a significant difference in the Finnish subpopulation as
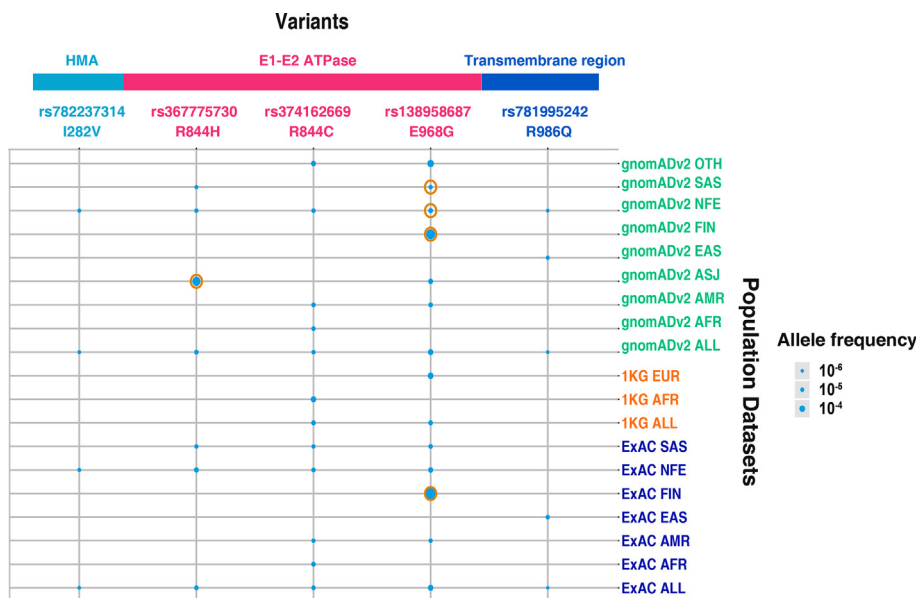
Fig. 3. Allele frequency distribution of pathogenic ATP7A variants in different population datasets. SAS, South Asian; NFE, Non-Finnish European; FIN, Finnish; EAS, Eastern Asian; AMR, American; AFR, African; OTH, Other; ALL, All Samples. 1KG_ALL, 1000 genomes. ESP6500si_ALL, ESP6500 All Samples.
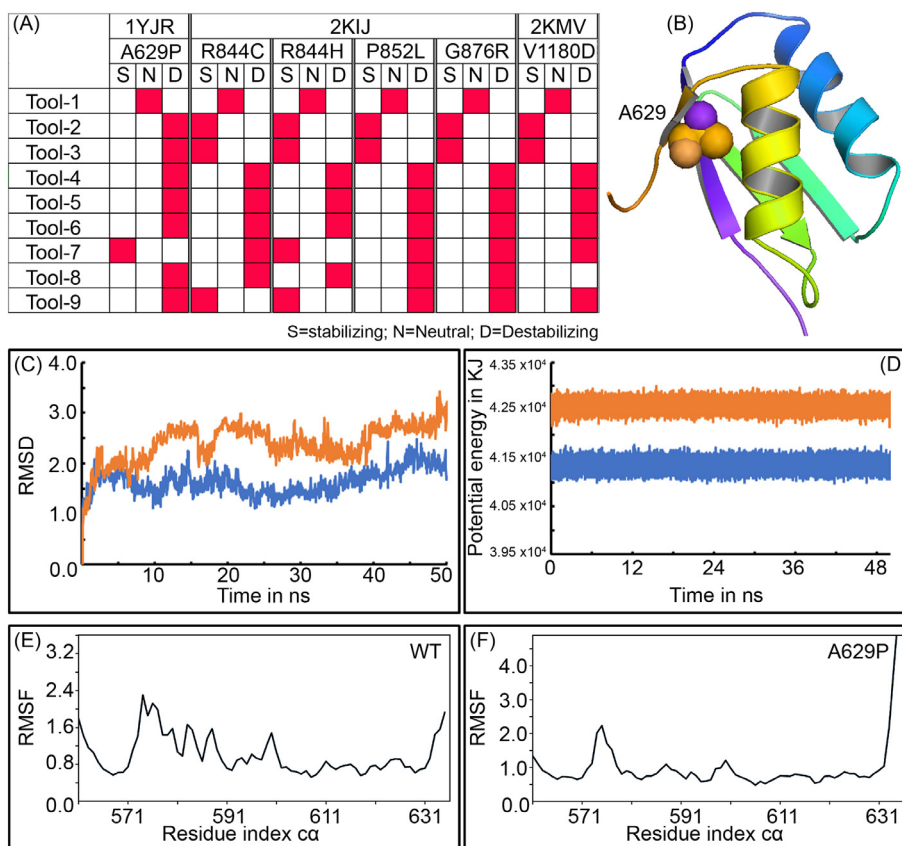
Fig. 4. A. Heatmap of the stability predictions using 9 different tools, where tool 1–9 represents FoldX, Schrödinger-OPLS2005, Schrödinger-OPLS3, MOE, CUPSAT, mCSM, SDM, iMutant 2.0 and POPMUSIC. B. Apo form of the sixth soluble domain of ATP7A, the position of the A629P mutation is labeled and marked with spheres. C. RMSD of the WT (blue) and A629P (orange). D. Potential energies of the WT (blue) and A629P (orange). E and F. RMSF of WT and A629P throughout 50 ns MD simulations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

compared to the global allele frequency. The variant rs781995242 shows frequency of 0.00001 in the gnomAD and ExAC datasets while it was absent in the 1000 Genome dataset. Fig. 3 shows the distribution of pathogenic variants in the population datasets.

3.5. Molecular modeling studies

After systematically matching the respective positions of mutations, we selected 6 variants for our studies; A629P, R844C, R844H, P852L, G876R and V1180D. These variants lie in three different domains of the ATP7A structure. In our stability predictions, we found FoldX predicted all of them as 'neutral'. Similarly, when applied to two different force fields, i.e., OPLS2005 and OPLS3, Schrödinger predicted as 'stabilizing' for variants R844C, R844H, P852L, G876R. In the same way, SDM predicted A629P and

R844H as 'stabilizing'. Also, POPMUSIC predicted R844C, R844H as 'stabilizing'. The iMutant 2.0 was unable to predict any effect for V1180D. However, most of the tools in our consensus-based stability prediction predicted the variants as 'destabilizing'. The stability prediction heatmap is shown in Fig. 4A.

Since the majority of the tools predicted these variants as destabilizing, we assumed that our MD simulations would provide more insights into the reasons for the destability. To elucidate the mutation-structure relationships we mainly computed and compared the root mean square deviation (RMSD) of the Cα and root mean square fluctuations (RMSF) of individual residues of mutant and WT protein over a 50 ns simulation time. The potential energy of mutant and WT was also calculated and compared between the mutants and WT.

In the case of A629P (Fig. 4B), the RMSD of the mutant does not deviate much when compared to the WT (Fig. 4C). There was only
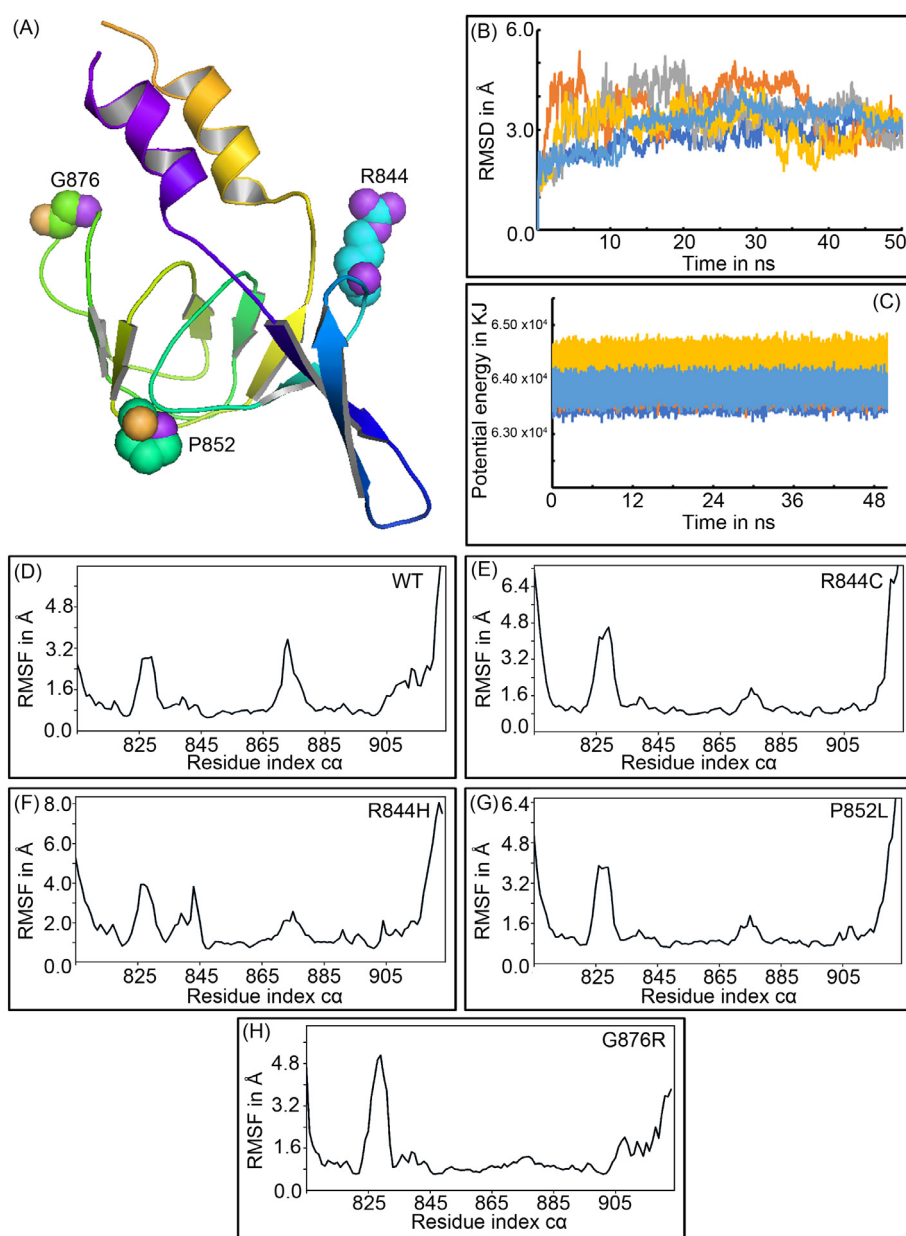
Fig. 5. A. Structure of the actuator domain of ATP7A, position of the mutations is labeled and marked as spheres. B. RMSD of the WT (dark blue), R844C (orange), R844H (grey), P852L (yellow) and G876R (pale blue). C. Potential energies of the WT (dark blue), R844C (orange), R844H (grey), P852L (yellow) and G876R (pale blue). 5D-H. RMSF of WT, R844C, R844H, P852L and G876R respectively throughout 50 ns MD simulations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

0.10 KJ difference between potential energies of mutant (4.25×10^4 KJ) and WT (4.15×10^4 KJ) throughout the 50 ns MD simulations (Fig. 4D). We found that the RMSF of mutant has deviated significantly at one of the flanking regions when compared to the WT (Fig. 4E and F). Our analysis revealed that the side chains of lysine and glutamic acid at the flanking regions were fluctuating drastically. It is known that RMSF calculations also include the side chain flexibility. When comparing the ensembles of the experimental structures, the RMSF of lysine and glutamic acid at the flanking region are also highly distorted as we observed in our MD simulations. From our simulation studies, we could conclude that the A629P mutation didn't impart any structural changes when compared to the WT. Our results are in good agreement with the previously reported experimental studies [20].

In the case of R844C, R844H, P852L and G876R (Fig. 5A), the RMSD of all the mutants agree with that of WT (Fig. 5B). RMSDs of all the proteins including WT were stabilized between 2 and 4 Å. Similarly, the potential energies (Fig. 5C) of the mutants and WT were between 6.35×10^4 KJ to 6.45×10^4 KJ which indicate the energies of the systems are well maintained throughout the MD simulations irrespective of the mutation. We then computed the RMSF of the mutants and compared with that of WT and found that the RMSF of both mutants and WT are stable (Fig. 5D–H). These results lead us to think that any of these mutations do not impose detectable structural changes when compared to the WT.

Finally, for V1180D (Fig. 6A), we found that the RMSD does not deviate much for the mutant when compared to the WT. For both mutant and WT, the RMSD was stabilized between 10 and 12 Å (Fig. 6B). To understand the drastic jump in RMSD, we critically analyzed the distortions in the structures throughout the MD simulations and found that most of the structural distortions happened in the region that connects β3 and β4. High flexibility for this region was previously reported in the experimental studies

[22]. The potential energy and RMSF of both mutant and WT were almost the same throughout the MD simulations (Fig. 6C and D). However, when we compared the overall secondary structural elements, we found that there was a 5% reduction in the beta strands for the mutant. The mutation V1180D was located at the β4 strand. When valine is replaced with aspartic acid, there is not sufficient space to accommodate the side chain of the aspartic acid and due to that there is structural rearrangement which is reflected in the reduction in the secondary structural elements.

3.6. Database interface and features

The ATP7A genetic variant resource features a user friendly web based interface which allows query and browsing the compendium of variants. The search interface allows query of the compendium using variant, ACMG, AAchange and dbSNP ID. Variants can be queried by formatted queries containing the chromosome, variant position, reference and alternate alleles. The query retrieves the list of all the matching entries from the database in a tabular form with an option to open the variant details. The variant details are compiled in 4 broad sections The first section includes basic information of the variant i.e. Gene name, Chromosome, Genome Build, HGVS Nomenclature, genomic locus, reference and alternate bases and the amino acid change. The second panel provides more information on the variant with respect to the functional effect of the mutation, and the classification as per the ACMG & AMP guidelines. This section also provides information on the disease along with the inheritance patterns, technique used for identification of the variant, ethnicity from which the variant has been reported and links to the variant in dbSNP database. The last sections provide information on the geographical / population information from where the variant was reported and the publication from which the variant information was retrieved.
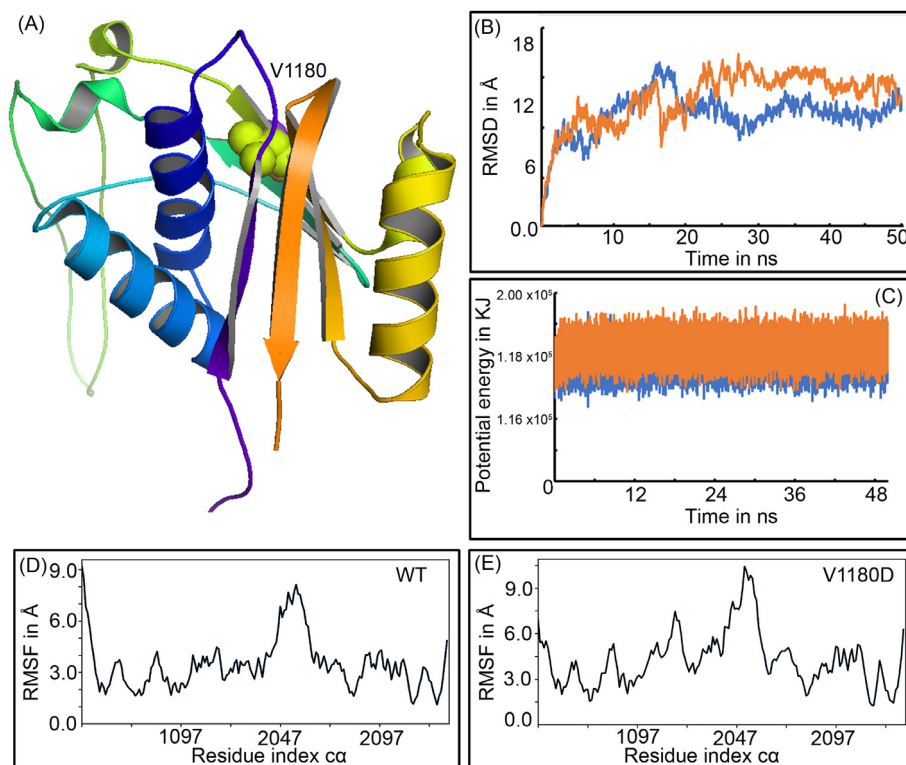
Fig. 6. Structure of the nucleotide binding domain of the ATP7A, position of the mutation is labeled and marked with spheres. B. RMSD of the WT (dark blue) and V1180D (orange). C. Potential energies of the WT (dark blue), V1180D (orange). D and E. RMSF of WT and V1180D throughout 50 ns MD simulations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Conclusions

The ATP7A Genetic Variant Resource is a comprehensive compendium of genetic variants in the ATP7A gene. The variant compendium indexes 404 variants in the ATP7A gene. The database also provides the annotations as per the ACMG & AMP guidelines putting together the disease associations and evidence to qualify the pathogenicity as derived from integrative analysis of datasets and published literature to aid clinicians and researchers in the clinical interpretation of variants.

Our MD studies revealed no large changes in the local structures of the protein. However, it may be possible that even the small structural distortions that we observed in our studies may affect the stability of the global structure. Also changes in the microenvironment of the mutations that we could not detect could play an important role in protein destabilization. We also admit that longer-timescale simulation may reveal more significant structural changes which could lead to the destabilization of the protein. We agree that our MD simulation studies have limitations in terms of time duration. Due to the incomplete structure of ATP7A, the predictions may not be fully indicative of the global stability of ATP7A. Additionally, the lack of experimental support to corroborate the simulation results (or vice versa) means further studies are necessary.

4.1. Genotype to phenotype correlation

Molecular diagnosis of MD can play a vital role in better prognosis and thus resulting in better disease management. In order to overcome the problem of clinical heterogeneity in MD patients, the defects should be properly characterized. Mutations like nonsense, frameshift-insertion, deletions which are truncating can cause complete disruption of the protein function, such mutations can lead to severe clinical phenotype. A young patient harboring ATP7A Q1168X mutation showed severe manifestation of MD [83]. Patients showing classical MD phenotype have shown to harbor ATP7A R201X, Q303X, R980X, R986X but exceptionally K1408X was found in a patient with milder phenotype of MD. ATP7A N1304S variant was observed in patient with mild form of MD i.e. OHS. It is observed that MD with milder form are characterized by mutations showing some residual activity while mutations with no residual activity manifested classical form of MD [41].

To the best of our knowledge, this is the most comprehensive collection of genetic variants in ATP7A gene and central point of information for the genetic epidemiology, carrier screening and functional genomics in MD and OHS.

Author contribution

VS and BK conceived and designed the study. AM, and MK, collected the data and ACMG classification. DKV and KYJZ performed molecular dynamics simulations. AM, MP and MK designed the figures, KP contributed to the search engine. All authors reviewed and approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Solomons NW. Biochemical, metabolic, and clinical role of copper in human nutrition. J Am Coll Nutr 1985. https://doi.org/10.1080/07315724.1985.10720069.

[2] Kaler SG. ATP7A-related copper transport diseasesg-emerging concepts and future trends. Nat Rev Neurol 2011. https://doi.org/10.1038/nrneurol.2010.180.

[3] González M, Reyes-Jara A, Suazo M, Jo WJ, Vulpe C. Expression of copper-related genes in response to copper load. Am. J. Clin. Nutr. 2008. https://doi.org/10.1093/ajcn/88.3.830s.

[4] DiDonato M, Sarkar B. Copper transport and its alterations in Menkes and Wilson diseases. Biochim Biophys Acta – Mol Basis Dis 1997. https://doi.org/10.1016/S0925-4439(96)00064-6.

[5] León-García G, Santana A, Villegas-Sepúlveda N, Pérez-González C, Henrríquez-Esquíroz JM, de León-García C, et al. The T1048I mutation in ATP7A gene causes an unusual Menkes disease presentation. BMC Pediatr 2012. https://doi.org/10.1186/1471-2431-12-150.

[6] De Bie P, Muller P, Wijmenga C, Klomp LWJ. Molecular pathogenesis of Wilson and Menkes disease: Correlation of mutations with molecular defects and disease phenotypes. J Med Genet 2007. https://doi.org/10.1136/jmg.2007.052746.

[7] Beyens A, Van Meensel K, Pottie L, De Rycke R, De Bruyne M, Baeke F, et al. Defining the clinical, molecular and ultrastructural characteristics in occipital horn syndrome: two new cases and review of the literature. Genes (Basel) 2019. https://doi.org/10.3390/genes10070528.

[8] Kennerson ML, Nicholson GA, Kaler SG, Kowalski B, Mercer JFB, Tang J, et al. Missense mutations in the copper transporter gene ATP7A cause X-linked distal hereditary motor neuropathy. Am J Hum Genet 2010. https://doi.org/10.1016/j.ajhg.2010.01.027.

[9] Freeman PJ, Hart RK, Gretton LJ, Brookes AJ, Dalgleish R. VariantValidator: accurate validation, mapping, and formatting of sequence variation descriptions. Hum Mutat 2018. https://doi.org/10.1002/humu.23348.

[10] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the american college of medical genetics and genomics and the association for molecular pathology. Genet Med 2015. https://doi.org/10.1038/gim.2015.30.

[11] Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: an online force field. Nucleic Acids Res 2005. https://doi.org/10.1093/nar/gki387.

[12] Banks JL, Beard HS, Cao Y, Cho AE, Damm W, Farid R, et al. Integrated modeling program, applied chemical theory (IMPACT). J Comput Chem 2005. https://doi.org/10.1002/jcc.20292.

[13] Harder E, Damm W, Maple J, Wu C, Reboul M, Xiang JY, et al. OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. J Chem Theory Comput 2016. https://doi.org/10.1021/acs.jctc.5b00864.

[14] Vilar S, Cozza G, Moro S. Medicinal chemistry and the molecular operating environment (MOE): application of QSAR and molecular docking to drug discovery. Curr Top Med Chem 2008. https://doi.org/10.2174/156802608786786624.

[15] Parthiban V, Gromiha MM, Schomburg D. CUPSAT: prediction of protein stability upon point mutations. Nucleic Acids Res 2006. https://doi.org/10.1093/nar/gkl190.

[16] Pires DEV, Ascher DB, Blundell TL. MCSM: predicting the effects of mutations in proteins using graph-based signatures. Bioinformatics 2014. https://doi.org/10.1093/bioinformatics/btt691.

[17] Worth CL, Preissner R, Blundell TL. SDM – A server for predicting effects of mutations on protein stability and malfunction. Nucleic Acids Res 2011. https://doi.org/10.1093/nar/gkr363.

[18] Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res 2005. https://doi.org/10.1093/nar/gki375.

[19] Dehouck Y, Kwasigroch JM, Gilis D, Rooman M. PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. BMC Bioinf 2011. https://doi.org/10.1186/1471-2105-12-151.

[20] Banci L, Bertini I, Cantini F, Migliardi M, Rosato A, Wang S. An atomic-level investigation of the disease-causing A629P mutant of the menkes protein, ATP7A. J Mol Biol 2005. https://doi.org/10.1016/j.jmb.2005.07.034.

[21] Banci L, Bertini I, Cantini F, Migliardi M, Natile G, Nushi F, et al. Solution structures of the actuator domain of ATP7A and ATP7B, the menkes and wilson disease proteins. Biochemistry 2009. https://doi.org/10.1021/bi901003k.

[22] Banci L, Bertini I, Cantini F, Inagaki S, Migliardi M, Rosato A. The binding mode of ATP revealed by the solution structure of the N-domain of human ATP7A. J Biol Chem 2010. https://doi.org/10.1074/jbc.M109.054262.

[23] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr Sect D Biol Crystallogr 2010. https://doi.org/10.1107/S0907444910007493.

[24] Tang J, Donsante A, Desai V, Patronas N, Kaler SG. Clinical outcomes in Menkes disease patients with a copper-responsive ATP7A mutation, G727R. Mol Genet Metab 2008. https://doi.org/10.1016/j.ymgme.2008.06.015.

[25] Møller LB, Horn N. Mutation detection in the menkes gene ATP7A using the protein truncation test. Clin Med Pathol 2008. https://doi.org/10.4137/cpath.s565.

[26] Vonk WIM, De Bie P, Wichers CGK, Van Den Berghe PVE, Van Der Plaats R, Berger R, et al. The copper-transporting capacity of ATP7A mutants associated with Menkes disease is ameliorated by COMMD1 as a result of improved protein expression. Cell Mol Life Sci 2012. https://doi.org/10.1007/s00018-011-0743-1.

[27] Hicks JD, Donsante A, Pierson TM, Gillespie MJ, Chou DE, Kaler SG. Increased frequency of congenital heart defects in Menkes disease. Clin Dysmorphol 2012. https://doi.org/10.1097/MCD.0b013e32834ea52b.

[28] J.Y. K, H.S. C, T.-J. P, H.J. S, D.W. S, H.S. N, et al. Screening for 392 polymorphisms in 141 pharmacogenes. Biomed Reports 2014. https://doi.org/10.3892/br.2014.

[29] Rizk T, Mahmoud A, Jamali T, Al-Mubarak S. Menkes disease presenting with epilepsia partialis continua. Case Rep Neurol Med 2014. https://doi.org/10.1155/2014/525784.

[30] Bansagi B, Lewis-Smith D, Pal E, Duff J, Griffin H, Pyle A, et al. Phenotypic convergence of menkes and wilson disease. Neurol Genet 2016. https://doi.org/10.1212/NXG.0000000000000119.

[31] Mercer SW, Wang J, Burke R. In vivo modeling of the pathogenic effect of copper transporter mutations that cause Menkes and Wilson diseases, motor neuropathy, and susceptibility to Alzheimer's disease. J Biol Chem 2017. https://doi.org/10.1074/jbc.M116.756163.

[32] Bansagi B, Griffin H, Whittaker RG, Antoniadi T, Evangelista T, Miller J, et al. Genetic heterogeneity of motor neuropathies. Neurology 2017. https://doi.org/10.1212/WNL.0000000000003772.

[33] Caicedo-Herrera G, Candelo E, Pinilla J, Vidal A, Cruz S, Pachajoa HM. Novel ATP7A gene mutation in a patient with Menkes disease. Appl Clin Genet 2018. https://doi.org/10.2147/TACG.S180087.

[34] Cao Y, Tokita MJ, Chen ES, Ghosh R, Chen T, Feng Y, et al. A clinical survey of mosaic single nucleotide variants in disease-causing genes detected by exome sequencing. Genome Med 2019. https://doi.org/10.1186/s13073-019-0658-2.

[35] Seidel J, Caca K, Schwab SG, Berr F, Wildenauer DB, Mentzel HJ, et al. Disturbed copper transport in humans. Part 2: mutations of the ATP7B gene lead to Wilson disease (WD). Cell Mol Biol (Noisy-Le-Grand) 2001.

[36] Watanabe A, Shimizu N. Identification of three novel mutations in Japanese patients with Menkes disease and mutation screening by denaturing high performance liquid chromatography. Pediatr Int 2005. https://doi.org/10.1111/j.1442-200x.2004.02012.x.

[37] Gourdon P, Liu XY, Skjørringe T, Morth JP, Møller LB, Pedersen BP, et al. Crystal structure of a copper-transporting PIB-type ATPase. Nature 2011. https://doi.org/10.1038/nature10191.

[38] Møller LB, Rea G, Yasmeen S, Skjørringe T, Thorborg SS, Morrison PJ, et al. A silent nucleotide substitution in the ATP7A gene in a child with Menkes disease. Mol Genet Metab 2013. https://doi.org/10.1016/j.ymgme.2013.09.012.

[39] Wang Q, Ding Y, Wang JM, Huang QH, Zhao CF, Liu YP, et al. Clinical and ATP7A gene analysis of three infants with Menkes disease and prenatal diagnosis for a fetus at risk. Chin J Contemp Pediatr 2014. https://doi.org/10.7499/j.issn.1008-8830.2014.06.013.

[40] Cao B, Yang X, Chen Y, Huang Q, Wu Y, Gu Q, et al. Identification of novel ATP7A mutations and prenatal diagnosis in Chinese patients with Menkes disease. Metab Brain Dis 2017. https://doi.org/10.1007/s11011-017-9985-4.

[41] de Gemmis P, Enzo MV, Lorenzetto E, Cattelan P, Segat D, Hladnik U. 13 novel putative mutations in ATP7A found in a cohort of 25 Italian families. Metab Brain Dis 2017. https://doi.org/10.1007/s11011-017-0010-8.

[42] Králík L, Flachsová E, Hansíková H, Saudek V, Zeman J, Martásek P. Molecular diagnostics of copper-transporting protein mutations allows early onset individual therapy of menkes disease. Folia Biol (Czech Republic) 2017.

[43] Ogawa A, Yamamoto S, Takayanagi M, Kogo T, Kanazawa M, Kohno Y. Identification of three novel mutations in the MNK gene in three unrelated Japanese patients with classical Menkes disease. J Hum Genet 1999. https://doi.org/10.1007/s100380050144.

[44] Ogawa A, Yamamoto S, Kanazawa M, Ogawa E, Takayanagi M, Hasegawa S, et al. Novel mutation of L718X in the ATP7A gene in a Japanese patient with classical Menkes disease, and four novel polymorphisms in the Japanese population. J Hum Genet 2000. https://doi.org/10.1007/s100380070024.

[45] Gu YH, Kodama H, Murata Y, Mochizuki D, Yanagawa Y, Ushijima H, et al. ATP7A gene mutations in 16 patients with Menkes disease and a patient with occipital horn syndrome. Am J Med Genet 2001. https://doi.org/10.1002/1096-8628(2001)9999:9999<::AID-AJMG1167>3.0.CO;2-R.

[46] Hahn S, Cho K, Ryu K, Kim J, Pai K, Kim M, et al. Identification of four novel mutations in classical menkes disease and successful prenatal DNA diagnosis. Mol Genet Metab 2001. https://doi.org/10.1006/mgme.2001.3169.

[47] Ozawa H, Kodama H, Murata Y, Takashima S, Noma S. Transient temporal lobe changes and a novel mutation in a patient with Menkes disease. Pediatr Int 2001. https://doi.org/10.1046/j.1442-200X.2001.01402.x.

[48] Kim BE, Smith K, Meagher CK, Petris MJ. A conditional mutation affecting localization of the Menkes disease copper ATPase: Suppression by copper supplementation. J Biol Chem 2002. https://doi.org/10.1074/jbc.M208737200.

[49] Gu YH, Kodama H, Sato E, Mochizuki D, Yanagawa Y, Takayanagi M, et al. Prenatal diagnosis of Menkes disease by genetic analysis and copper measurement. Brain Dev 2002. https://doi.org/10.1016/S0387-7604(02)00093-1.

[50] Poulsen L, Horn N, Møller LB. X-linked recessive Menkes disease: carrier detection in the case of a partial gene deletion. Clin Genet 2002. https://doi.org/10.1034/j.1399-0004.2002.620604.x.

[51] Tümer Z, Møller LB, Horn N. Screening of 383 unrelated patients affected with menkes disease and Finding of 57 gross deletions in ATP7A. Hum Mutat 2003. https://doi.org/10.1002/humu.10287.

[52] Møller LB, Bukrinsky JT, Mølgaard A, Paulsen M, Lund C, Tümer Z, et al. Identification and analysis of 21 novel disease-causing amino acid substitutions in the conserved part of ATP7A. Hum Mutat 2005. https://doi.org/10.1002/humu.20190.

[53] Bhadhprasit W, Kodama H, Fujisawa C, Hiroki T, Ogawa E. Effect of copper and disulfiram combination therapy on the macular mouse, a model of Menkes disease. J Trace Elem Med Biol 2012. https://doi.org/10.1016/j.jtemb.2012.05.002.

[54] Kim BE, Petris MJ. Phenotypic diversity of Menkes disease in mottled mice is associated with defects in localisation and trafficking of the ATP7A protein. J Med Genet 2007. https://doi.org/10.1136/jmg.2007.049627.

[55] Donsante A, Tang J, Godwin SC, Holmes CS, Goldstein DS, Bassuk A, et al. Differences in ATP7A gene expression underlie intrafamilial variability in Menkes disease/occipital horn syndrome. J Med Genet 2007. https://doi.org/10.1136/jmg.2007.050013.

[56] Kaler SG, Holmes CS, Goldstein DS, Tang J, Godwin SC, Donsante A, et al. Neonatal diagnosis and treatment of Menkes disease. N Engl J Med 2008. https://doi.org/10.1056/NEJMoa070613.

[57] Kaler SG, Liew CJ, Donsante A, Hicks JD, Sato S, Greenfield JC. Molecular correlates of epilepsy in early diagnosed and treated Menkes disease. J Inherit Metab Dis 2010. https://doi.org/10.1007/s10545-010-9118-2.

[58] Donsante A, Johnson P, Jansen LA, Kaler SG. Somatic mosaicism in Menkes disease suggests choroid plexus-mediated copper transport to the developing brain. Am J Med Genet Part A 2010. https://doi.org/10.1002/ajmg.a.33632.

[59] Moizard MP, Ronce N, Blesson S, Bieth E, Burglen L, Mignot C, et al. Twenty-five novel mutations including duplications in the ATP7A gene. Clin Genet 2011. https://doi.org/10.1111/j.1399-0004.2010.01461.x.

[60] Bell CJ, Dinwiddie DL, Miller NA, Hateley SL, Ganusova EE, Mudge J, et al. Carrier testing for severe childhood recessive diseases by next-generation sequencing. Sci Transl Med 2011. https://doi.org/10.1126/scitranslmed.3001756.

[61] Skjørringe T, Tümer Z, Møller LB. Splice site mutations in the ATP7A gene. PLoS ONE 2011. https://doi.org/10.1371/journal.pone.0018599.

[62] Y.H. K, R. L, H.W. Y, M.-S. Y, S.H. B, S.C. C, et al. Identification of a novel mutation in the ATP7A gene in a Korean patient with menkes disease. J Korean Med Sci 2011. https://doi.org/10.3346/jkms.2011.26.7.951.

[63] Prasad AN, Levin S, Rupar CA, Prasad C. Menkes disease and infantile epilepsy. Brain Dev 2011. https://doi.org/10.1016/j.braindev.2011.08.002.

[64] Yi L, Donsante A, Kennerson ML, Mercer JFB, Garbern JY, Kaler SG. Altered intracellular localization and valosin-containing protein (p97 VCP) interaction underlie ATP7A-related distal motor neuropathy. Hum Mol Genet 2012. https://doi.org/10.1093/hmg/ddr612.

[65] Møller L, Lenartowicz M, Zabot MT, Josiane A, Burglen L, Bennett C, et al. Clinical expression of Menkes disease in females with normal karyotype. Orphanet J Rare Dis 2012. https://doi.org/10.1186/1750-1172-7-6.

[66] Gu YH, Kodama H, Kato T. Congenital abnormalities in Japanese patients with Menkes disease. Brain Dev 2012. https://doi.org/10.1016/j.braindev.2012.01.012.

[67] Tümer Z. An Overview and update of ATP7A mutations leading to menkes disease and occipital horn syndrome. Hum Mutat 2013. https://doi.org/10.1002/humu.22266.

[68] Lee JS, Lim BC, Kim KJ, Hwang YS, Cheon JE, Kim IO, et al. Menkes disease in Korea: ATP7A mutation and epilepsy phenotype. Brain Dev 2015. https://doi.org/10.1016/j.braindev.2014.04.012.

[69] Smpokou P, Samanta M, Berry GT, Hecht L, Engle EC, Lichter-Konecki U. Menkes disease in affected females: the clinical disease spectrum. Am J Med Genet Part A 2015. https://doi.org/10.1002/ajmg.a.36853.

[70] Tümer Z, Petris M, Zhu S, Mercer J, Bukrinski J, Bilz S, et al. A 37-year-old Menkes disease patient—Residual ATP7A activity and early copper administration as key factors in beneficial treatment. Clin Genet 2017. https://doi.org/10.1111/cge.13083.

[71] G. C-H, E. C, J. P, A. V, S. C, H.M. P. Novel ATP7 gene mutation in a patient with Menkes disease. Appl Clin Genet 2018. https://doi.org/10.2147/TACG.S180087.

[72] Das S, Levinson B, Whitney S, Vulpe C, Packman S, Gitschier J. Diverse mutations in patients with Menkes disease often lead to exon skipping. Am J Hum Genet 1994.

[73] Tümer Z, Lund C, Tolshave J, Vural B, Tønnesen T, Horn N. Identification of point mutations in 41 unrelated patients affected with Menkes disease. Am J Hum Genet 1997.

[74] Mori M, Nishimura M. A serine-to-proline mutation in the copper-transporting P-type ATPase gene of the macular mouse. Mamm Genome 1997. https://doi.org/10.1007/s003359900457.

[75] Murata Y, Kodama H, Abe T, Ishida N, Nishimura M, Levinson B, et al. Mutation analysis and expression of the mottled gene in the macular mouse model of Menkes disease. Pediatr Res 1997. https://doi.org/10.1203/00006450-199710000-00003.

[76] Ohta Y, Shiraishi N, Nishikimi M. Occurrence of two missense mutations in Cu-ATPase of the macular mouse, a Menkes disease model. Biochem Mol Biol Int 1997. https://doi.org/10.1080/15216549700204721.

[77] Jankov RP, Boerkoel CF, Hellmann J, Sirkin WL, Tümer Z, Horn N, et al. Lethal neonatal Menkes' disease with severe vasculopathy and fractures. Acta Paediatr Int J Paediatr 1998. https://doi.org/10.1080/080352598750031013.

[78] Skjørringe T, Pedersen PA, Thorborg SS, Nissen P, Gourdon P, Møller LB. Characterization of ATP7A missense mutants suggests a correlation between intracellular trafficking and severity of Menkes disease. Sci Rep 2017. https://doi.org/10.1038/s41598-017-00618-6.

[79] Kim JH, Lee BH, Kim YM, Choi JH, Kim GH, Cheon CK, et al. Novel mutations and clinical outcomes of copper-histidine therapy in Menkes disease patients. Metab Brain Dis 2014. https://doi.org/10.1007/s11011-014-9569-5.

[80] Tzschach A, Grasshoff U, Beck-Woedl S, Dufke C, Bauer C, Kehrer M, et al. Next-generation sequencing in X-linked intellectual disability. Eur J Hum Genet 2015. https://doi.org/10.1038/ejhg.2015.5.

[81] Park S, Park JY, Kim GH, Choi JH, Kim KM, Kim JB, et al. Identification of novel ATP7B gene mutations and their functional roles in Korean patients with Wilson disease. Hum Mutat 2007. https://doi.org/10.1002/humu.20574.

[82] 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 2009. https://doi.org/10.2133/dmpk.24.565.

[83] Lin Y-J, Ho C-S, Hsu C-H, Lin J-L, Chuang C-K, Tsai J-D, et al. A truncating de novo point mutation in a young infant with severe menkes disease. Pediatr Neonatol 2017;58:89–92. https://doi.org/10.1007/s11011-017-0010-8.