

https://dspace.org/

Report of a missense TJP2 variant associated to PFIC4 with a pronounced phenotypic variability Focus on the structural effects on the protein level

Boudour Khabou; Houcemeddine Othman; Manel Guirat; Imen Chabchoub; Sana Kmiha ...

https://hdl.handle.net/10539/47357
WOS

Downloaded from a DSpace repository. https://dspace.org/

www.nature.com/jhg

ARTICLE

Report of a missense TJP2 variant associated to PFIC4 with a pronounced phenotypic variability: Focus on the structural effects on the protein level

Boudour Khabou[1] ✉, Houcemeddine Othman[2,3,4], Manel Guirat[5,6,7], Imen Chabchoub[8], Sana Kmiha[8], Bahri Mahjoub[9], Rania Abdelhadi[10,11,12], Afif Ben Mahmoud[13], Rim Kallel[14], Tahya Sellami Boudawara[14], Thouraya Kammoun[8], Faiza Fakhfakh[1], Hassen Hadj Kacem[15] and Hassen Kammoun[5]

© The Author(s), under exclusive licence to The Japan Society of Human Genetics 2025

PFIC4 is a chronic liver disease which cannot be diagnosed based on clinical and biochemical findings with an unpredictable evolution. Here, we reported three consanguineous families with 9 children suffering from intrahepatic cholestasis with low GGT-activity. Three probands were chosen to undergo genetic testing. In silico analyses were conducted to assess the functional impact of the identified variant, along with variants occurring at highly conserved positions within the protein. Additionally, close clinical monitoring was carried. Targeted-NGS sequencing ruled out the diagnosis of PFIC1 and PFIC2. Subsequently, WES allowed the establishment of PFIC4 diagnosis for the three families through the identification of a homozygous *TJP2* variant p. Gly532Arg classified as likely pathogenic with a structural damage predicted based on biomolecular modeling and simulation analysis. In-depth in silico analysis of 90 nsSNPs occurring in highly conserved residues in PDZ domains showed 14 ones seems to be relevant in the clinical practice. Clinically, a pronounced phenotypic variability is noted. In conclusion, our study described a homozygous missense PFIC4-related variant with a highlight on the pathogenic power of such types of variants. The clinical evaluation provided information about the importance of close monitoring to prevent liver failure and clarified the unexpected course of PFIC4.

Journal of Human Genetics (2025) 70:331–339; https://doi.org/10.1038/s10038-025-01338-w

INTRODUCTION

Progressive familial intrahepatic cholestasis (PFIC) is a heterogeneous group of autosomal recessive hepatobiliary diseases caused by defect in bile secretion or transport and lead, in several cases, to liver failure. Until 2013, three main subtypes of PFIC (PFIC1, PFIC2 and PFIC3) have been thoroughly characterized with a recognized clinical, histological and biochemical pattern for each form [1]. Lately, whole-exome sequencing (WES) have expanded the knowledge about the genetic background of this group and allowed the characterization of 9 others PFIC forms (From PFIC4 to PFIC12), all related to biallelic mutations in genes involved in the complex process of bile secretion (https://www.omim.org/). The clinical attributes of these forms seem to be overlapping, making the clinical diagnosis a challenging task [2].

PFIC4 has been firstly described in 2014 as a form of PFIC related to biallelic mutations in *TJP2* gene and presenting an early onset (< 6 months), low Gamma-Glutamyl Transferase (GGT) levels with a risk of extrahepatic manifestations [3]. Subsequently, others genetically-confirmed PFIC4 cases have been reported exhibiting a variable degree of severity of their liver disease and differential outcome [4, 5]. Besides, a biallelic mutation in *TJP2* gene has been also described, alone or in association with a *BAAT* mutation, in patients with familial hypercholanemia (FHC), a non-progressive liver disease characterized by pruritic and fat malabsorption [6]. Moreover, it has been documented that monoallelic TJP2 variant can be observed in Benign Recurrent Intrahepatic Cholestasis (BRIC) or Intrahepatic Cholestasis of Pregnancy (ICP) [7, 8]. PFIC4, FHC, ICP, and BRIC disorders all appear to fall within one spectrum, despite differences in severity. The amount of TJP2 alteration determines the developed phenotype. Nevertheless, the implication of other genetic or environmental factors must also be taken into consideration.

The *TJP2* gene encodes for TJP2 protein (1190 residues), also called Zonula Occludens-2 (ZO-2) protein. It is a multidomain

[1]Molecular and Functional Genetics Laboratory, Faculty of Sciences, University of Sfax, Sfax, Tunisia. [2]Department of Genetics, Farhat Hached University Hospital, Ibn El Jazzar St, Sousse 4000, Tunisia. [3]Laboratory of cytogenetics, molecular genetics and reproductive biology (LR03SP02), Farhat Hached University Hospital, Ibn El Jazzar St, Sousse 4000, Tunisia. [4]Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, 9 Jubilee Road, Parktown, Johannesburg 2193, South Africa. [5]Medical Genetic Department, Hédi Chaker University Hospital, Sfax, Tunisia. [6]Laboratory of Molecular and Cellular Screening Processes (LPCMC), Biotechnology Center of Sfax, Sfax, Tunisia. [7]Faculty of Medicine of Sfax, University of Sfax, Sfax, Tunisia. [8]Pediatric Department, Hedi Chaker University Hospital, Sfax, Tunisia. [9]Pediatric Department, Tahar Sfar Hospital, Mahdia, Tunisia. [10]Centre of Biotechnology of Sfax (Tunisia), Microorganisms and Biomolecule Laboratory, and Molecular and Cellular Screening Process Laboratory, Sfax, Tunisia. [11]Ksentini Private Laboratory of Cytogenetics of Sfax, Sfax, Tunisia. [12]Ribosite Biotech Company for Research and Development in Biotechnhology, Sfax, Tunisia. [13]Neurological Disorder Research Center, Qatar Biomedical Research Institute (QBRI), Qatar Foundation, Hamad Bin Khalifa University (HBKU), Doha, Qatar. [14]Department of Anatomy and Pathological Cytology, Habib Bourguiba Hospital, Sfax, Tunisia. [15]Department of Applied Biology, College of sciences, University of Sharjah, Sharjah, United Arab Emirates. ✉email: Boudour.khabou.bio@gmail.com

Received: 14 January 2025 Revised: 30 March 2025 Accepted: 30 March 2025
Published online: 18 April 2025

scaffold molecule with three N-terminal PDZ domains, SH3 module and GuK domain, in addition to an acidic and proline-rich C-terminal region. Through these domains, ZO-2 established a wide range of interactions [9, 10]. Particularly, in the liver, ZO-2 is involved in the maintenance of a correct hepatocytes' polarity through its interactions with cytoskeletal and integral membrane proteins. The disruption of this polarity affects the stability of canalicular transporters and their function, leading to an impaired bile secretion [11].

Referring to HGMD, the majority of *TJP2* mutations associated with progressive liver disease are predicted protein-truncating mutations (PPTMs) leading to a total absence of *TJP2* expression. The clinical interpretation of missense *TJP2* variants in the case of PFIC4 remains unclear [4]. To overcome this problem, a panoply of in silico tools could be used to shed light on the pathogenic effect of such type of variants. This approach is widely applied to identify pathogenic nsSNPs with putative relevance in the molecular genetic testing [12–15].

In this study, we reported three families including 9 affected children suffering from intrahepatic cholestasis with low GGT-activity. WES was conducted allowing the establishment of PFIC4 as the accurate diagnosis. Intensive in silico study and structural modeling were applied to understand the molecular mechanism behind the pathogenicity of the identified variant. In depth in silico analysis of the already described missense *TJP2* variants occurring in highly conserved residues was applied to assess the pathogenicity of such types of variants, poorly investigated in PFIC4. Comparative analysis of the symptoms and onset and the clinical course among the studied cases documented both inter and intrafamilial phenotypic variability.

RESULTS

Clinical features of the studied patients

Here, we described three consanguineous families with nine affected siblings presenting with cholestatic liver disease (Fig. 1). Since the first clinical examination, the UDCA and fat-soluble vitamins were administrated. A regular follow-up was maintained for seven patients. Clinical and biochemical data at the onset, as well as the current outcome are summarized in Table 1.

Family 1 : It is large, with six affected children; however, DNA samples were available for only four of them. Patient IV.3 was the proband selected for WES analysis. He underwent intensive clinical explorations including histological investigations at the moment of diagnosis revealed a mild portal fibrosis associated with canalicular and hepatocellular cholestasis, in addition to feathery degeneration and rosetting of the hepatocytes evolving towards a pronounced portal and septal fibrosis marking a pre-cirrhotic stage (Fig. 2, 3).

Families 2 and 3 : They included two twins and three patients, respectively with IV.2 from each family was chosen for genetic testing.

All patients were referred to the medical care at an age ranging from 2 to 10 months. Variable motifs at presentation as well variable course of the disease were noted. More details concerning these patients are provided in supplementary appendix.

Genetic testing

After ruling out diagnoses of PFIC1 and PFIC2, WES was performed. Following the implementation of the filtration process, a refined list of variants was obtained for further analysis. After applying the in silico panel of PanelApp (Neontal cholestasis: 96 genes), identical pathogenic homozygous variation was identified.

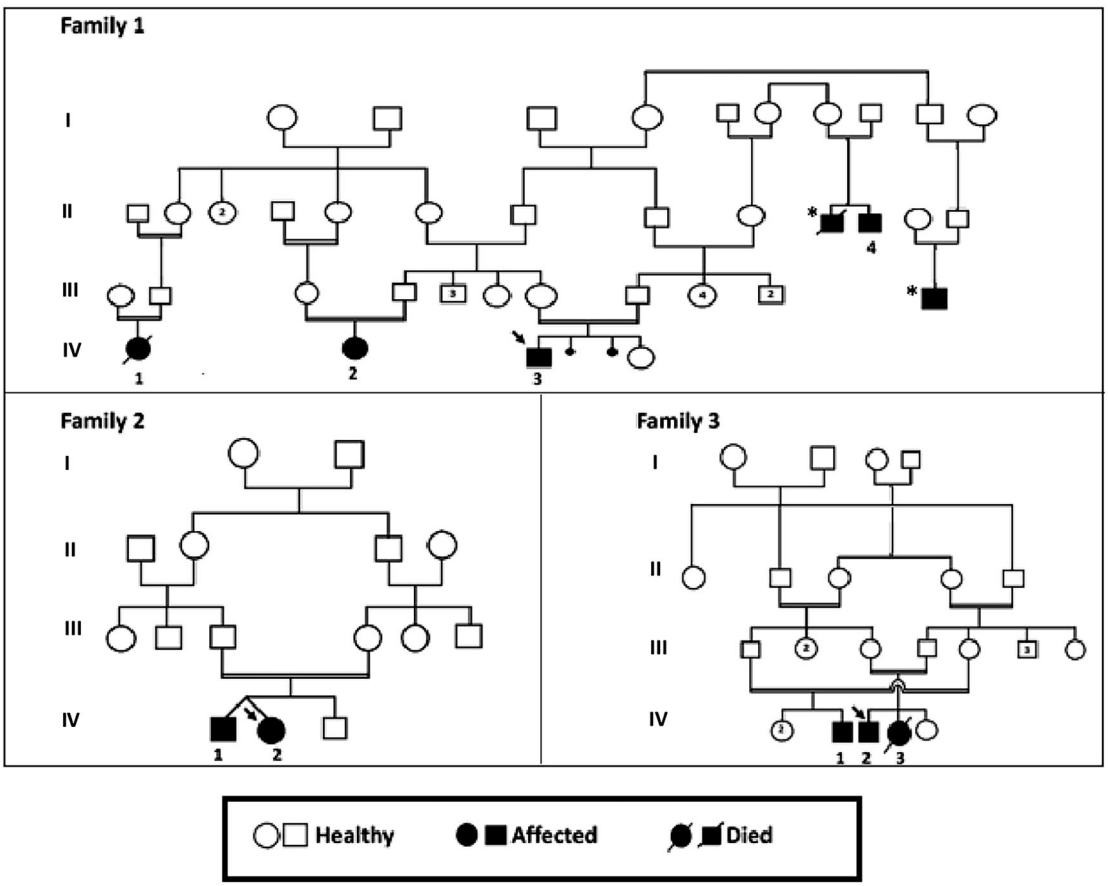


Fig. 1 Pedigrees of the studied families presenting with an intrahepatic cholestasis. Arrows indicates the proband choosen from each family for WES analysis. Asterisk indicates patients for whom neither DNA nor clinical data were available

Table 1. Clinical and biochemical findings at first examination and follow-up data for the studied patients

Case	Age of onset/ age at first presentation	Clinical findings at first examination	Liver function tests at first examination						Duration of follow-up, years	Present condition
			GGT (UI/L)	AST/ ALT (UI/L)	ALP	BiliT/biliC (umol/L)	TBA (umol/L)	PT (%)		
Family 1										
IV.1	2 mo/ 4 ye	Jaundice, differentiated HCC	53	247/ 178	NA	41/81	NA	44%	Unregular follow-up	Died at the age of years because of terminal LF
IV.2	1 mo /6 mo	pruritus	28	128/95	888	35.9/6.9	183.7	100	3 years	Present age 3.6 years : persistent pruritus, stable liver disease with normalized liver enzymes except PAL
IV.3	4 mo/ 5 mo	Prolonged jaundice, acholic stools, dark urine, pruritus	18	41/26		188.8/ 113.9	256	76.9%	7 years	Present age 7,5 years : decompensated cirrhosis (Stade 4) with evidence of PH, LT waiting list
II.4	1 mo/ 3 mo	Jaundice, acholic stools, dark urine,	NA	205/ 200*	570*	62/30	NA	100	Unregular follow-up	Present age 28 years, stable liver disease with normalized liver enzymes except PAL
Family 2										
IV.1	9mo/10mo	Jaundice, pruritus	19	39.7/ 32.4	591	21.9/15.8	142.2	100	5 years	Present age 6 years, persistence of biochemical cholestasis with pruritus with a correct liver function
IV.2	4 mo/10mo	Jaundice, pruritus, hepatomegaly	8	42/49	955	43/27	NA	99	5 years	Present age 6 years, persistence of biochemical and clinical cholestasis without progression of the disease
Family 3										
IV.1	6 mo	Pruritus	17	52/80	848	NA	99	100	11 years	Present age 12 years : asymptomatic with an evidence of biochemical cholestasis
IV.2	6 mo / 14 mo	Pruritus, hepatomegaly, splenomegaly	39	60/41	223	17/4	NA	100	15 years	Present age 16 years : compensated cirrhosis with evidence of PH
IV.3	6 mo	Prolonged jaundice	26	253/ 173	738	28/19	NA	93	9 years	Died at the age of 10 years because of terminal LF

AST aspartate aminotransferase, *ALT* alanine aminotransferase, *GGT* γ-glutamyl transpeptidase, *ALP* alkaline phosphatase, *TBA* total biliary acids, *LT* liver transplantation, *LF* Liver failure, *PH* portal hypertension, *HCC* hepatocellular carcinoma

(NA : Not available), *at the age of 2 years

Normal values: AST < 35 U/L; ALAT < 40 U/L; GGT < 30 U/L; ALP < 250 U/L; TBA <14umol/l

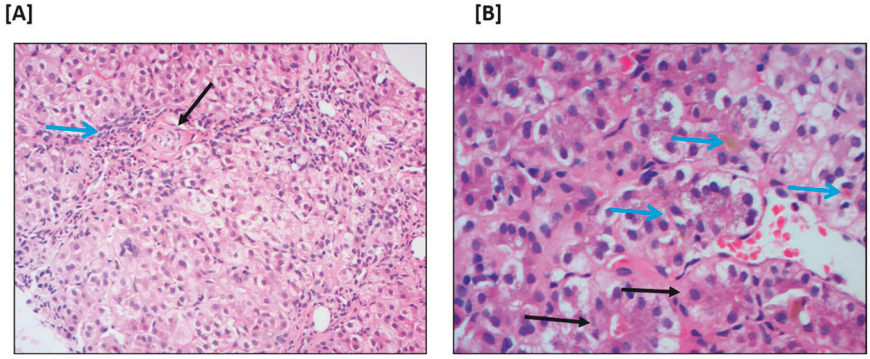


Fig. 2 H&E-stained sections of patient 3 from family 1 at the moment of diagnosis (5 months). **A** Inflammatory infiltrates (blue Arrow) associated with mild portal fibrosis (black arrow) (original magnification ×100). **B** Canalicular and hepatocellular cholestasis (blue arrow) with feathery degeneration and rosetting of the hepatocytes (black arrow) (original magnification ×200)

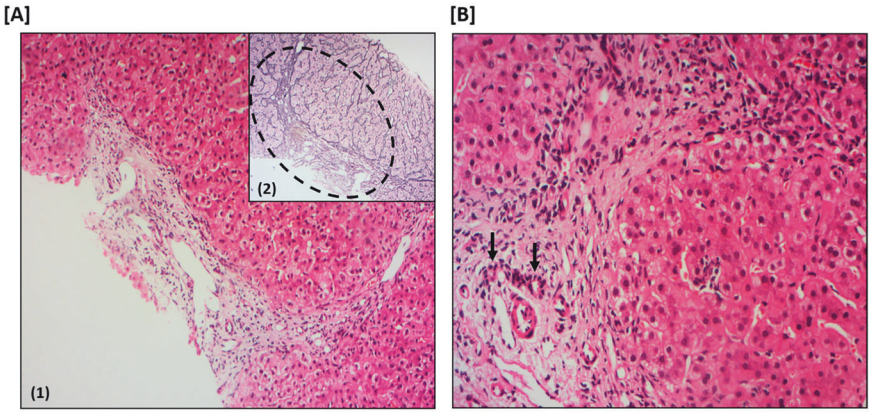


Fig. 3 Histological findings of patient 3 from family 1 at the age of 4 years showing the progression of the liver disease. **A** Pronounced portal and septal fibrosis marking a pre-cirrhotic stage revealed by H&E staining (1) and confirmed by reticulum coloration (2) (highlighted by dotted circle) (original magnification ×100). **B** Original magnifcation ×200 confirming the advanced fibrosis and showing a neoductulare proliferation (black arrow)

The mutation, localized at the exon 11 of *TJP2* gene (c.1594 G > A), leads to the substitution of the Glycine at the position 532 by an arginine (p.Gly532Arg). This variant occurs in PDZ3 domain, considered as a functional domain within TJP2 protein

According to GnomAD v4, this variant possessed a minor allele frequency (MAF) equal 2.052e-06 and has been never described at homozygous state. Furthermore, Mobidetails' pathogenicity prediction tools, as well as Mutpred results, all produced predictions supporting its causality (CADD:31, Revel:0.72, Poly-Phen:1, Alpha-Missense:0.993, ClinPred:0.999, Mutpred : 0.945)

Targeted sequencing of exon 11 confirmed parental heterozygosity and the presence of the variant in a homozygous state in all affected individuals in line with an autosomal recessive inheritance pattern.

Taken together, these data allowed the classification of the p.Gly532Arg as likely pathogenic according to the ACMG classification (PM1, PM2, PP1, PP3). Consequently, PFIC4 was retained as the accurate diagnosis.

In depth in silico analysis of the identified variant

Predictive results of Gly532Arg. Once the variant was retained, additional analysis were performed. This analysis denied a putative impact of the change c.1594 G > A on the splicing process. As on the protein level, results showed that the substituted residue belongs to the PDZ3 domain of TJP2 protein (509-590). It is a highly conserved residue (ConSurf score equal to 9) and exhibits a slight intolerance (0.57) to genetic variants. A thorough examination of PDZ3 via ConSurf showed that 70% of their residues are conserved (ConSurf score 8/9) in favor of its importance. Besides, the intolerance to genetic variants is a trait shared by most PDZ3 residues, whose MetaDome scores range from 0.68 to 0.41 (Fig. S1). These findings highlighted the importance of the PDZ3 domain harboring the identified variant.

In the other part, a Grantham distance of 125 was attributed to the Gly to Arg substitution, reflecting the high difference of the biochemical properties between both residues. In fact, Glycine's simple side chain allows for more flexibility of the backbone. On the other hand, the Arg is a very hydrophilic residue that prefers to reside on the surface of proteins and has a high potential for interactions with other hydrophilic residues, water, or organic compounds. Considering these observations, the deleterious effect of such change, especially at the structural level, seems to be accentuated. To delve deeper in this hypothesis, we performed a structural phenotyping of the variant through SNPeffect. Results showed that the p.Gly532Arg increased the aggregation tendency of the mutated protein with a dTANGO = 97.8 > 50. Besides, the Gly532Arg leads to a less soluble and stable TJP2 protein compared to the WT one according to the predictions supplied by SODA (-6.7) and FoldX (ΔΔG = 4.99 kcal/mol), respectively.

Structural modeling and simulation. The homology model of the PDZ3 domain exhibits a typical fold, characterized by six β strands and three α helices arranged in a β barrel fold (Fig. 4A). The p. Gly532Arg variant resides at the C-terminal edge of the loop connecting the β2 and β3 strands within PDZ3. By performing a structural alignment with the PSD-95 PDZ domain [16], we inferred the putative binding site of PDZ3. This interaction groove involves residues from the β strand and α2 helix (Fig. 4A).

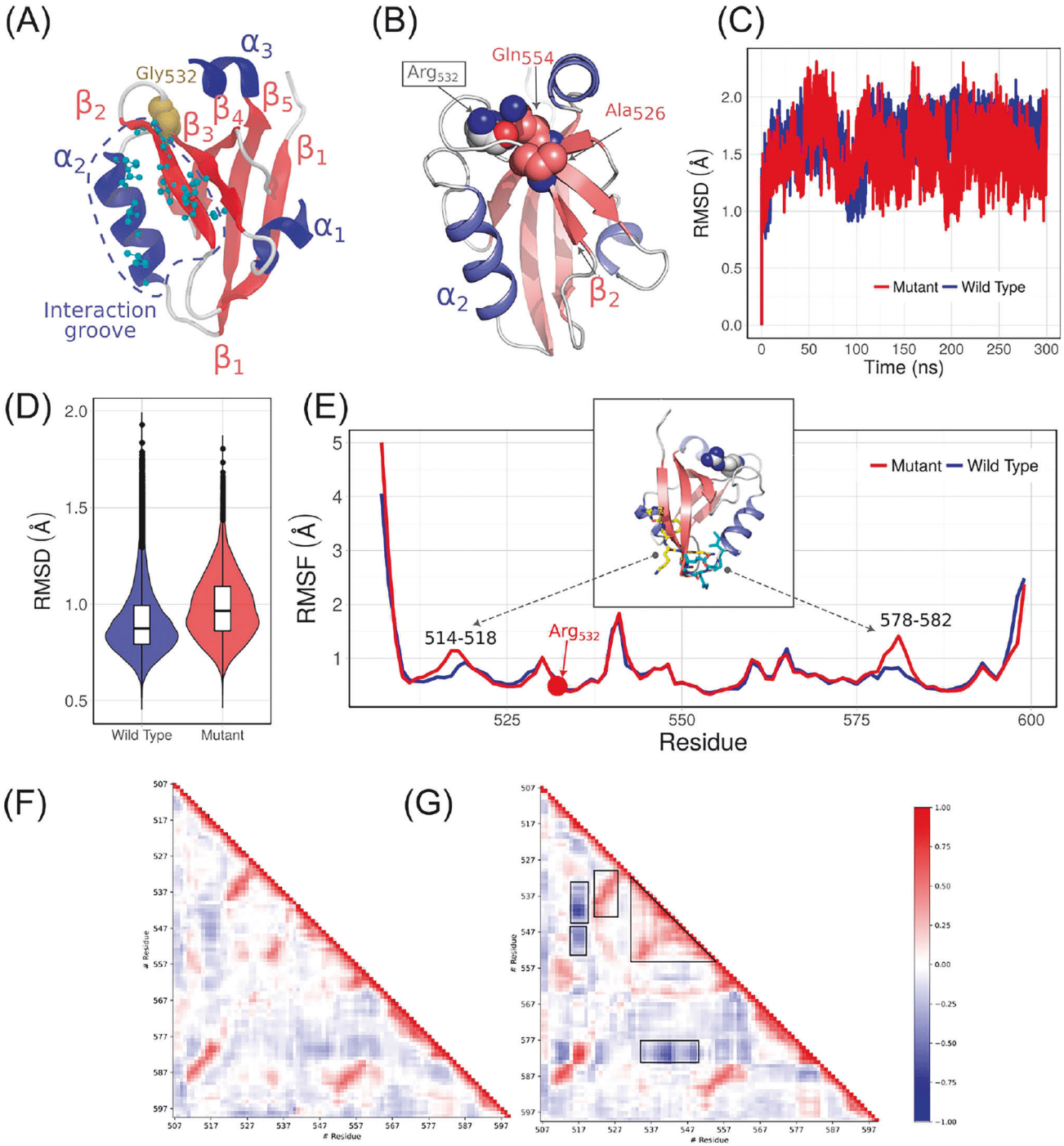


Fig. 4 Molecular modeling and dynamics simulation of the TJP2 PDZ3 domain. **A** Homology model of the TJP2 PDZ3 domain highlighting the location of residue Gly532. The blue sticks indicate residues of the interaction groove located on the α2 helix and β2 strand. **B** Interaction of the residue Arg532 in the variant structure with residues Q554 and A526, as mapped on the homology model. **C** RMSD time evolution of Cα atoms during the molecular dynamics simulation, fitted and calculated relative to the first frame of the trajectory. **D** Violin plot showing the RMSD distribution, fitted and calculated relative to the average structure of each trajectory. **E** Plot of per-residue RMSF. **F**, **G** Heatmap of the cross-correlation matrix for the reference and variant structures, respectively. Black rectangles highlight segments where the dynamic profile has changed the most

Although residue 532 is not part of the groove, it is situated in close proximity to residue A526 (Fig. 4B). Upon mutation, we observed that residue R532 can establish several polar contacts, utilizing its elongated side chains to interact with residues Q554 and A526 (Fig. 4B).

We conducted molecular dynamics simulations for the reference and p.Gly532Arg variant structures, totaling 300 ns. During the simulation, the Root Mean Square Deviation (RMSD) exhibited equilibration around 15 ns, followed by two plateau phases (90 ns to 110 ns and 110 ns to 300 ns) (Fig. 4C). Overall, the simulation converged, with RMSD remaining below 2 Å. Notably, the p.Gly532Arg variant displayed more fluctuation throughout the 300 ns, occasionally exceeding 2 Å. When calculating RMSD using an average reference structure, the variant's structure showed

greater variability (Fig. 4D). Specifically, the Mean RMSD values were 0.91 Å for the reference and 0.98 Å for the variant. The variances of RMSD between the two simulations significantly differed based on an F-test (p value = 1.11E–16).

Subsequently, we investigated the per-residue fluctuation by constructing the Root Mean Square Fluctuation (RMSF) profile (Fig. 4E). In both the reference and variant trajectories, we observed increased flexibilities in specific segments. Specifically, segment 514-518 encompasses the β1 strand and part of the β1-β2 connecting loop, while segment 578-582 covers the α2-β5 connecting loop and the two N-terminal residues of the β5 strand. Notably, the RMSF peaks differ by 0.47 Å and 0.59 Å for the first and second segments, respectively.

After the equilibration phase of each trajectory, we computed the dynamic cross-correlation matrix (Fig. 4F, G). The resulting heatmap revealed distinct patterns between the reference and variant structures. Notably, the p.Gly532Arg variant exhibited stronger correlated movements near the variant site, spanning segment 532-555. In contrast, the same segment showed reduced movement correlation in the reference sequence. Additionally, certain other segments displayed anti-correlated movements in the variant.

We then analyzed the hydrogen bond network which revealed that Gly532 in the reference structure formed a polar contact with Gly528, persisting for at least 10% of the simulation time. This contact occurred via a main chain-to-main chain interaction. In contrast, Arg532 in the variant's structure not only maintained the polar contact with Gly528 but also established new contacts with Asn529 and Arg564. Overall, Arg532 in the variant's structure participated in a total of 40 hydrogen bonds with varying occupancies. In comparison, the reference structure exhibited only three hydrogen bonds involving Gly532.

In silico analysis of nsSNPs in highly conserved positions of PDZ domains

Referring to GnomAD v4.1.1, we retrieved 90 nsSNPs altering highly conserved residues in PDZ domains and possessing a very low MAF. CADD scoring showed 66 nsSNPs present a CADD score between 20 and 29.9, thus possessing a putative pathogenic effect while 21 ones present a CADD score>30, so selected as the most pathogenic variants (Table S1).

The 21 nsSNPs were subject of additional analysis. Firstly, a comparative analysis of Grantham distance showed that the changes between WT and mutated residues were radical or moderately radical for only 8 variants. In the other part, FoldX results revealed that TJP2 is destabilized to varying degrees by 14 nsSNPs, with p.Asp76Asn and p.Gly528Ser being the most destabilizing. According to SODA study, 14 and 7 nsSNPs result in less soluble and more soluble TJP2 protein, respectively (Table S2).

DISCUSSION

In this study, we reported three consanguineous families including nine individuals suffering from intrahepatic cholestasis with low GGT activity. One proband from each family underwent WES after a negative customized sequencing panel followed by and in-depth in silico research to gain a deeper understanding of the variant's impact. Clinically, a regular follow-up was maintained for 5 patients for a period ranging from 3 to 15 years.

Singleton WES analysis revealed the presence of the same substitution c.1594 G > A (p.Gly532Arg), in the three probands, in homozygous state in the exon 11 of *TJP2* gene. This variant has been reported once in two double-diagnosed patients harboring two causative mutations in *TJP2* and *DNAI1* behind PFIC4 and Ciliary dyskinesia primary 1, respectively without conducting additional research into the potential effects of these variants on the mutated proteins [17]. In our study, the p.Gly532Arg variant was confirmed to be homozygous in all affected individuals. Its pathogenicity was predicted using in silico tools. While such analyses cannot replace functional characterization, the damaging effect appears highly likely, as p.Gly532Arg received high pathogenicity scores from MutPred, CADD, and REVEL—three tools that have demonstrated excellent predictive performance in previous studies [18–20]. Supporting this assertion, we observed a correlation between functional impact and predictive scores in other TJP2 variants known to disrupt protein expression, including p.Leu79Arg, p.Leu525Ser, and p.Gly737Ser. Similar to p.Gly532Arg, these variants exhibited high MutPred scores (> 0.9), CADD scores (> 28), and REVEL scores (> 0.7). Furthermore, additional analyses aimed at understanding the molecular disease mechanism of p.Gly532Arg revealed a potential effect on both protein solubility and stability, likely due to structural disruption.

To give deeper insights into the structural changes caused by the Gly to Arg replacement, molecular modeling and dynamics simulation of the TJP2 PDZ3 domain were performed. Significant differences in stability between the reference structure and the mutated structure were revealed. In the reference structure, Gly532 is located at the beginning of the β3 strand, where it likely contributes to the adaptability and stability of the β strand by reducing conformational strain [21]. Glycine's small size and flexibility are advantageous in maintaining the integrity of the β strand and the overall structure. In contrast, the introduction of the larger and more rigid Arginine residue introduces steric hindrance and potential for unfavorable interactions, leading to structural instability. Such effects could impair the protein function [22, 23]. This instability is not confined to the local environment around the residue but extends to key segments, particularly those affecting the rigid secondary structure elements. The increase in flexibility at the α2-β2 connecting loop suggests that the interaction groove interface, critical for the domain's function, may also be destabilized. Our results also showed that the introduction of Arg532 disrupts the dynamic pattern of the PDZ3 domain, potentially propagating its effects to distant segments of the protein. The ability of Arginine to form strong polar interactions could further explain these observations. Unlike Glycine, Arg532 engaged in many hydrogen bond interactions, which may alter the local environment and protein dynamics significantly. This propensity for establishing polar interactions could lead to an overall disruption in the PDZ3 domain's structural and dynamic integrity. Our findings align with previous studies, which have demonstrated that the dynamic pattern of the PDZ3 domain is highly sensitive to changes in the protein structure [24].

Since TJP2 has been characterized as a scaffold protein, which associated several protein, it seems that the TJP2 protein harboring the p.Gly532Arg will lost its binding affinity to ZASP, the binding partner of PDZ3 [10]. Further, we sought that similar consequences could be observed with missense variants in PDZ1, PDZ2 or PDZ3 domains, three crucial protein-binding domains of TJP2 [25]. Indeed, our study showed that 21 ns SNPs occurring in highly conserved residues of these domains, are highly pathogenic (CADD > 30) with an impact on structural stability and/or solubility of 14 ones. Despite their pathogenicity, the biochemical difference between WT and Mutated residues for 13 nsSNPs is low. This finding elucidated that the deleterious effect of missense substitution is mostly related to the functional importance of the WT residue, rather than to changes in the physic-chemical properties between WT and mutated residues [26, 27]. Consequently, our in silico analysis shed light on the deleteriousness of missense variants in *TJP2*, especially those altering highly conserved residues occurring in the functional domains. We suggested that this type of variant must be considered in the context of TJP2-related diseases.

Although such substitutions in *TJP2* could have drastic impact, our study constitutes the second report of homozygous missense mutation related to PFIC4. Indeed, the majority of PFIC4 patients

lack the TJP2 expression due to PPTMs. Only five missense variants (p.Leu79Arg, p.Leu525Ser, p.Gly737Ser, p.Glu401Gly, p.Val69Met) have been described in compound heterozygous status with truncating ones [3, 4, 28, 29]. Referring to the literature, three homozygous *TJP2* variants have been described as associated to ICP (p.Gln128Lys & p.Arg20His) or FHC (p.Val48Ala), two mild liver diseases [6, 30]. An in silico analysis including all missense variants was performed to compare their pathogenicity (Table 2). Results showed that both Gln128Lys and Arg20His are predicted as neutral, while the remaining ones present equivalent predictive scores with the exception of p.Glu401Gly, whose impact appears to be less pronounced. We suggest that a threshold for residual *TJP2* expression/function should exist in order to ascertain if homozygous missense variants cause PFIC4 or less common hepatic disorders like ICP or FHC.

Clinically, p.Gly532Arg has been already reported in two patients presenting a variable clinical severity at diagnosis ranging from hypertransaminasemia and hepatosplenomegaly to liver cirrhosis requiring a liver transplant [17]. In our study, we reported likewise a phenotypic heterogeneity, both in terms of symptoms and clinical outcomes. Indeed, four patients present a stable liver disease. In the other part, three ones manifested an unfavorable evolution and waiting for liver transplant or died at an ages of 11- and 10-years, with one of them developing HCC. A similar issue has already been documented in PFIC4 patients [31]. In the other part, this phenotypic variability has been reported among PFIC4 patients lacking the *TJP2* expression. Besides, an intrafamilial variability in the expression of PFIC4 has been described in a Syrian family with six children, harboring p.Gln1112* and present either a cirrhosis with signs of portal hypertension or only elevated enzymes levels. Authors suggested an age-dependent penetrance as a putative explanation of this variability [5]. In a similar context, an incomplete penetrance of a *TJP2* mutation responsible for Congenital Hearing loss has been shown in an Iranian family [32]. Additionally, a recent study reported, for the first time, the coexistence of PFIC4 and hearing impairment in Chinese child carrier of biallelic *TJP2* variants [33]. All these data provide cumulative evidence about the variable expressivity of *TJP2* variants and deny the presence of a genotype-phenotype correlation in PFIC4. Consequently, the evolution of PFIC4 disease cannot be predicted based solely on the genotype and an ongoing clinical monitoring remains crucial to avoid rapid progression to end-stage liver disease.

In conclusion, our study constitutes the second report of the homozygous missense variant (p.Gly532Arg) in PFIC4 with an elucidation of its impact on the structural level. The pathogenic effect of such substitutions on *TJP2* function and structure has been also highlighted. Furthermore, the clinical follow-up of the cases under study revealed PFIC4's intrafamilial and interfamilial phenotypic variability, demonstrating that the disease's unpredictable progression calls for constant clinical surveillance. However, the absence of a functional characterization of the found variant is a major limitation in our study. Such analysis requires labor-intensive methods and technical expertise that are unavailable.

PATIENTS AND METHODS

Patients

Our study focused on three consanguineous families comprising nine individuals who experienced a liver disease with varying degrees of severity. Interviews with guardians and elders in the family helped to establish the medical history of each family as well as the relationships among its members (Fig. 1). Subjects IV.3 (Family 1), IV.2 (family 2) IV.2 (Family 3) were referred to the pediatric department of Hedi Chaker (Sfax) and Taher Sfar (Mahdia) hospitals, respectively following a prolonged jaundice during the neonatal period. Intensive explorations allowed the retention of the diagnosis of intrahepatic cholestasis with normal GGT range after the exclusion of the other causes (infections, biliary atresia, primary sclerosing cholangitis.). These patients were chosen as probands for the genetic screening.

Due to the family history of the cholestasis, the same diagnosis was retained for the remaining affected individuals presenting all a persistent jaundice except IV.2 from family 1 which suffered only from intense pruritus at the onset.

The three probands were known to be free of mutations in ABCB11 (NM_003742.4) and ATP8B1 (NM_001374385.1) following a genetic analysis elaborated at common laboratory of Biology and Molecular Genetics of Saint-Antoine Hospital in Paris. Subsequently, PFIC1 and PFIC2 are excluded and a WES remains the only way to establish the diagnosis.

Methods

DNA extraction. The study was carried out in accordance with the local ethics committee's ethical guidelines. For the genomic study, informed consent was acquired from each participant or his or her legal representative before the collection of peripheral blood samples. Subsequently, a genomic DNA extraction was conducted using the *MagicPure*® Blood Genomic DNA Kit from Transgen biotech [34].

Whole exome sequencing. Firstly, the three probands were the subject of a genetic testing through a customized panel (10 genes) designed in 2014 for the diagnosis of hereditary cholestasis and cholelithiasis. The framework of this analysis was detailed previously [35]. This analysis revealed the absence of disease-causative variants.

Subsequently, a WES was conducted for a proband from each family to determine the causative mutation. Full details concerning the WES analysis and data analysis are included in the supplementary appendix.

Sanger sequencing. To confirm the segregation pattern of the identified variant in *TJP2* gene (NM_004817.4), we applied a Sanger sequencing using the DNA samples of the remaining patients and the parents of the probands. The sequence of interest was amplified by PCR using the following couple of primers TJP2_F 5′ TCTTCACCAGTAACCAGGGACCT3′ and TJP2_R 5′ GAAGACAGGACGCATTTGAGGA3′. The PCR reaction has been conducted according to the following conditions (30 cycles of denaturation at 95 °C for 30 s, hybridization at 60 °C for 30 s and extension at 72 °C for 40 s). Purified amplicons have been sequenced using ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit (ABI PRISM/Biosystems).

In silico analysis

Pathogenicity prediction: To establish an accurate interpretation of the TJP2 variant in the context of molecular diagnosis, we referred to five in silico tools including ConSurf, Mobidetails, SNPeffect, FoldX et SODA (Table S1).

Exhaustive in silico analysis of missense *TJP2* variants in PDZ domains: In order to give deeper insights into the pathogenicity of missense *TJP2* variants occurring in highly conserved residues (ConSurf score =9) of PDZ domains, we referred to GnomAD v4.1.1 to retrieve nsSNPs affecting these positions and possessing MAF < 0.01%.

These nsSNPs were the subject of an assessment of Grantham distance and CADD scoring. Because they are expected to be among the 0.1% most harmful changes in the human genome, nsSNPs *TJP2* variations with CADD scores, greater than 30, are chosen as the most pathogenic [36]. These ones were the subject of additional in silico analysis for a better characterization of their effect.

Molecular modeling and simulation analysis: The reference sequence for the TJP2 protein was obtained from the UniProt entry Q9UDY2. This protein consists of 1190 amino acids. The variant p.Gly532Arg is situated within the PDZ3 domain, spanning residues 506 to 590. A homology search followed by a construction of TJP2 PDZ3 domain model (WT and mutated) was performed. More details were provided in the supplementary appendix. We conducted molecular dynamics simulations using the AMBER 20 suite [37] for both the reference and p.Gly532Arg variant structures. Details concerning the simulations were included in the supplementary appendix. Finally, we analyzed the molecular dynamics data using Pytraj [38] and MDtraj [39].

Table 2. Comparative in silico analysis between pathogenic missense variants described in TJP2-related diseases

Statuts	Missense TJP2 mutation	Associated truncating TJP2 mutation	Conservation score (ConSurf)	In silico analysis							Associated phentoype	Reference
				MobiDetails*					Structural changes			
				Protein domain (UniProt)	MetaDome	Mutpred	CADD/ Revel	Grantham distance [5-215]	Protein solublity (SODA)	Protein stability (FoldX, ΔΔG)		
Homozygous	p. Gln128Lys (c.382 C > A)	—	2	ND	1.02 (slightly Tolerant)	0.108	6.9/ 0.066	53	More soluble (0.215)	No effect	ICP	30
	p. Arg 20 His (c.59 G > A)	—	1	ND	1.27 (Tolerant)	0.18	16.7/ 0.028	25	Less soluble (-1.41)	No effect		
	p.Val48Ala (c.143 T > C)	—	9	PDZ1	0.82 (Neutral)	0.91	27.3/ 0.637	64		Destablization effect (1.94 kcal/mol)	FHC	6
Homozygous	p. Gly 532 Arg (c. 1594 G > A)	—	9	PDZ3	0.57 (slightly intolerant)	0.945	31/ 0.72	125	Less soluble (-6.7)	Destablization effect (4.99 kcal/mol)	PFIC4	Our study
Coumpound Heterozygous	p.Leu79Arg (c.236 T > G)	p.Arg334*	9	PDZ1	0.71 (Neutral)	0.94	28.2/ 0.71	102	More soluble (3.34)	Destablization effect (8.27 kcal/mol)		4
	p.Leu525Ser (c.1574 T > C)	p.Val558*	9	PDZ3	0.53 (slightly intolerant)	0.937	29.7/ 0.77	145	More soluble (6.14)	Destablization effect (4.47 kcal/mol)		
	p.Gly737Ser (c.2209 G > A)	p.Arg970*	9	Guanylate kinase-like	0.58 (slightly intolerant)	0.925	31/ 0.794	56	Less soluble (-2.89)	Destablization effect (12.37 kcal/mol)		
	p.Glu401Gly (c.1202 A > G)	p.Ala273Profs*38	7	ND	0.79 (Neutral)	0.299	25.8/ 0.208	98	More soluble (4.72)	No effect (0.01 kcal/mol)		29
	p.Val69Met (c.205 G > A)	p.Arg367*	9	PDZ1	0.83 (Neutral)	0.685	26.7/ 0.55	21	More soluble (15.31)	Destablization effect (1.62 kcal/mol)		28
	p.Leu 791 Pro (c.2372 T > C)	Deletion 46 kb ex1-ex16	9	Guanylate kinase-like	0.26 (Intolerant)	0.969	28.6/ 0.876	98	More soluble (10.33)	Destablization effect (9.96 kcal/mol)		40

*Referring to Mobidetails, all tested missense variants have no effect on the splicing process
ND : Not in a defined domain

REFERENCES

1. Amer S, Hajira A. A comprehensive review of progressive familial intrahepatic cholestasis (PFIC): genetic disorders of hepatocanalicular transporters. Gastroenterol Res. 2014;7:39.
2. Vitale G, Mattiaccio A, Conti A, Turco L, Seri M, Piscaglia F, et al. Genetics in familial intrahepatic cholestasis: clinical patterns and development of liver and biliary cancers: a review of the literature. Cancers. 2022;14:3421.
3. Sambrotta M, Strautnieks S, Papouli E, Rushton P, Clark BE, Parry DA, et al. Mutations in TJP2 cause progressive cholestatic liver disease. Nat Genet. 2014;46:326–8.
4. Zhang J, Liu LL, Gong JY, Hao CZ, Qiu YL, Lu Y, et al. TJP2 hepatobiliary disorders: novel variants and clinical diversity. Hum Mutat. 2020;41:502–11.
5. Wei CS, Becher N, Friis JB, Ott P, Vogel I, Grønbæk H. New tight junction protein 2 variant causing progressive familial intrahepatic cholestasis type 4 in adults: A case report. World J Gastroenterol. 2020;26:550.
6. Carlton VE, Harris BZ, Puffenberger EG, Batta AK, Knisely AS, Robinson DL, et al. Complex inheritance of familial hypercholanemia with associated mutations in TJP2 and BAAT. Nat Genet. 2003;34:91–96.
7. Dixon PH, Sambrotta M, Chambers J, Taylor-Harris P, Syngelaki A, Nicolaides K, et al. An expanded role for heterozygous mutations of ABCB4, ABCB11, ATP8B1, ABCC2 and TJP2 in intrahepatic cholestasis of pregnancy. Sci Rep. 2017;7. 11823.
8. Kornitzer GA, Alvarez F. Case report: a novel single variant TJP2 mutation in a case of benign recurrent intrahepatic cholestasis. JPGN Rep. 2021;2:e087.
9. Beatch M, Jesaitis LA, Gallin WJ, Goodenough DA, Stevenson BR. The tight junction protein ZO-2 contains three PDZ (PSD-95Discs-LargeZO-1) domains and an alternatively spliced region. J Biol Chem. 1996;271:25723–6.
10. Gonzalez-Mariscal L, Bautista P, Lechuga S, Quiros M. ZO-2, a tight junction scaffold protein involved in the regulation of cell proliferation and apoptosis. Ann N Y Acad Sci. 2012;1257:133–41.
11. Itoh M, Terada M, Sugimoto H. The zonula occludens protein family regulates the hepatic barrier system in the murine liver. Biochimica et Biophysica Acta (BBA)-Mol Basis Dis. 2021;1867:165994.
12. Shinwari K, Guojun L, Deryabina SS, Bolkov MA, Tuzankina IA, Chereshnev VA. Predicting the most deleterious missense nonsynonymous single-nucleotide polymorphisms of hennekam syndrome-causing CCBE1 gene, in silico analysis. Sci World J. 2021;2021:1–19.
13. Behairy MY, Abdelrahman AA, Abdallah HY, Ibrahim EEDA, Sayed AA, Azab MM. In silico analysis of missense variants of the C1qA gene related to infection and autoimmune diseases. J Taibah Univ Med Sci. 2022;17:1074–82.
14. Poon KS. In silico analysis of BRCA1 and BRCA2 missense variants and the relevance in molecular genetic testing. Sci Rep. 2021;11. 11114.
15. Mansouri M, El Haddoumi G, Bendani H, Boumajdi N, Hakmi M, Abbou H, et al. In silico analyses of all STAT3 missense variants leading to explore divergent AD-HIES clinical phenotypes. Evolut Bioinforma. 2023;19. 11769343231169374.
16. Doyle DA, Lee A, Lewis J, Kim E, Sheng M, MacKinnon R. Crystal structures of a complexed and peptide-free membrane protein–binding domain: molecular basis of peptide recognition by PDZ. Cell. 1996;85:1067–76.
17. Rosina E, Pezzani L, Pezzoli L, Marchetti D, Bellini M, Pilotta A, et al. Atypical, composite, or blended phenotypes: how different molecular mechanisms could associate in double-diagnosed patients. Genes. 2022;13:1275.
18. Walters-Sen LC, Hashimoto S, Thrush DL, Reshmi S, Gastier-Foster JM, Astbury C, et al. Variability in pathogenicity prediction programs: impact on clinical diagnostics. Mol Genet Genom Med. 2015;3:99–110.
19. Thusberg J, Olatubosun A, Vihinen M. Performance of mutation pathogenicity prediction methods on missense variants. Hum Mutat. 2011;32:358–68.
20. Wang D, Li J, Wang Y, Wang E. A comparison on predicting functional impact of genomic variants. NAR genomics Bioinforma. 2022;4. lqab122.
21. Dou J, Vorobieva AA, Sheffler W, Doyle LA, Park H, Bick MJ, et al. De novo design of a fluorescence-activating β-barrel. Nature. 2018;561:485–91.
22. Tokuriki N, Tawfik DS. Stability effects of mutations and protein evolvability. Curr Opin Struct Biol. 2009;19:596–604.
23. Idicula-Thomas S, Balaji PV. Correlation between the structural stability and aggregation propensity of proteins. silico Biol. 2007;7:225–37.
24. Dudola D, Hinsenkamp A, Gáspári Z. Ensemble-Based analysis of the dynamic allostery in the PSD-95 PDZ3 domain in relation to the general variability of PDZ structures. Int J Mol Sci. 2020;21:8348.
25. Heinemann U, Schuetz A. Structural features of tight-junction proteins. Int J Mol Sci. 2019;20:6020.
26. MacGowan SA, Madeira F, Britto-Borges T, Schmittner MS, Cole C, Barton GJ. Human missense variation is constrained by domain structure and highlights functional and pathogenic residues. BioRxiv. 2017;127050.
27. Iqbal S, Pérez-Palma E, Jespersen JB, May P, Hoksza D, Heyne HO, et al. Comprehensive characterization of amino acid positions in protein structures reveals molecular effect of missense variants. Proc Natl Acad Sci. 2020;117:28201–11.
28. Lipiński P, Ciara E, Jurkiewicz D, Pollak A, Wypchło M, Płoski R, et al. Targeted next-generation sequencing in diagnostic approach to monogenic cholestatic liver disorders—single-center experience. Front Pediatr. 2020;8:414.
29. Tang J, Tan M, Deng Y, Tang H, Shi H, Li M, et al. Two novel pathogenic variants of TJP2 gene and the underlying molecular mechanisms in progressive familial intrahepatic cholestasis type 4 patients. Front Cell Dev Biol. 2021;9. 661599.
30. Zöllner J, Finer S, Linton KJ, van Heel DA, Williamson C, Dixon PH. Rare variant contribution to cholestatic liver disease in a South Asian population in the United Kingdom. Sci Rep. 2023;13. 8120.
31. Zhou S, Hertel PM, Finegold MJ, Wang L, Kerkar N, Wang J, et al. Hepatocellular carcinoma associated with tight-junction protein 2 deficiency. Hepatology. 2015;62:1914–6.
32. Rajabi S, Dastmalchi R, Dehghan MH, Eftekharian A, Aghazadeh E, Ghaderian SMH. TJP2 Gene Mutation c. G1012A May Responsible for Congenital Hearing Loss with Incomplete Penetrance in An Iranian Pedigree. J Genet Resour. 2019;5:143–8.
33. Zhang J, Guo S, Mei TL, Zhou J, Guan DX, Wang GL. Novel mutation of the TJP2 gene in a Chinese child with progressive cholestatic liver disease coexistent with hearing impairment. Hepatobiliary Pancreat Dis Int. 2021;20:198–200.
34. Gong R, Li S. Extraction of human genomic DNA from whole blood using a magnetic microsphere method. Int J Nanomed. 2014;9:3781–9.
35. Khabou B, Mahjoub B, Barbu V, Balhoudi N, Wardani A, Sfar MT, et al. Phenotypic variability in Tunisian PFIC3 patients harboring a complex genotype with a differential clinical outcome of UDCA treatment. Clin Chim Acta. 2018;486:122–8.
36. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic acids Res. 2019;47:D886–D894.
37. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, et al. The Amber biomolecular simulation programs. J Comput Chem. 2005;26:1668–88.
38. Roe DR, Cheatham TE III. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J Chem theory Comput. 2013;9:3084–95.
39. McGibbon RT, Beauchamp KA, Harrigan MP, Klein C, Swails JM, Hernández CX, et al. MDTraj: a modern open library for the analysis of molecular dynamics trajectories. Biophys J. 2015;109:1528–32.
40. Mirza N, Bharadwaj R, Malhotra S, Sibal A. Progressive familial intrahepatic cholestasis type 4 in an Indian child: presentation, initial course and novel compound heterozygous mutation. BMJ Case Rep CP. 2020;13:e234193.

ACKNOWLEDGEMENTS

We acknowledge the Center for High Performance Computing (CHPC), South Africa, for providing computational resources to conduct the molecular simulation analysis. And Ribosite Biotechnology company for advanced molecular biology technical support and customized services. This project is carried out under the MOBIDOC scheme, funded by the EU through the SWAFY proJECT and managed by the ANPR.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s10038-025-01338-w.

Correspondence and requests for materials should be addressed to Boudour Khabou.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.