Rare variant contribution to cholestatic liver disease in a South Asian population in the United Kingdom

Julia Zöllner[1], MBBS, PhD; Sarah Finer[2], MBBS, PhD; Kenneth J. Linton[3], PhD; Genes and Health Research Team[4]; David A. van Heel[4], BM BCh MA, DPhil; Catherine Williamson[1*], MBBS, PhD; Peter H. Dixon[1], PhD.

1. Department of Women and Children's Health, School of Life Course Sciences, King's College London, London, UK.

2. Institute for Population Health Sciences, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK.

3. Centre for Cell Biology and Cutaneous Research, Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK.

4. Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK.

*Correspondence to:

Catherine Williamson, catherine.williamson@kcl.ac.uk, 2.30W Hodgkin Building, Department of Women and Children's Health, School of Life Course Sciences, FOLSM, King's College London, Guy's Campus, London SE1 1UL, UK

Author Conflict of Interest/Study Support

Guarantor of the article: Catherine Williamson

Specific author contributions: CW and PD conceived the idea of the study and applied for data access. JZ accessed, extracted, analysed, and verified the data. SF and DvH helped with clinical data extraction. JZ, CW, PD and KL contributed to data interpretation. KL and JZ lead on protein modelling and interpretation. SF and DvH helped with clinical phenotype data interpretation and extractions. All authors gave final approval of the final draft to be published and have contributed to the manuscript.

Financial support: Disclose funding sources for the publication: CW is funded by a National Institute of Health and Care Research (NIHR) Senior Investigator grant (NIHR200254). The views expressed in this Article are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health. The funders had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Potential competing interests: SF receives research funding for Genes & Health from MRC, NIHR, Alnylam Pharmaceuticals, Takeda, Glaxo Smith Kline, Merck, Pfizer, NovoNordisk, Maze Pharmaceuticals, Bristol Myers Squibb. DvH receives research funding for Genes & Health from EU-H2020, NIH, MRC, HEFCE-Catalyst, Wellcome, HDR-UK, Alnylam Pharmaceuticals, Takeda, Glaxo Smith Kline, Merck, Pfizer, NovoNordisk, Maze Pharmaceuticals, Bristol Myers Squibb. CW consults for Mirum and GSK. JZ, PD, and KL have none to declare.

Key words: south Asian, cholestasis, exomes, mutations, disease, bioresource

Word count: 3010

WHAT IS KNOWN

Cholestatic liver diseases encompass a broad range of conditions.

Intrahepatic cholestasis of pregnancy (ICP) is the commonest gestational liver disease.

Genetic and environmental factors contribute to the aetiology of cholestatic disease.

South Asian populations are disproportionally affected.

WHAT IS NEW HERE

Exome sequencing analysis in a British Pakistani and Bangladeshi population discovered new genetic mutations.

Pathogenic variants were identified that increase risk of cholestatic liver disease.

Novel variants that contribute to ICP were identified.

Abstract

Objectives: This study assessed the contribution of five genes previously known to be involved in cholestatic liver disease in British Bangladeshi and Pakistani people in the United Kingdom as they are an understudied genetic ancestry group with disproportionate disease burden.

Methods: Five genes (*ABCB4*, *ABCB11*, *ATP8B1*, *NR1H4*, *TJP2*) were interrogated by low/mid whole exome sequencing data of 5236 volunteers. Included were non-synonymous or loss of function (LoF) variants with a minor allele frequency <5%. Variants were filtered and annotated. Rare variant burden analysis was conducted. Variants associated with a phenotype or predicted to be likely pathogenic (LP) underwent protein structure and modelling analysis in silico.

Results: Out of 314 non-synonymous variants, 180 fulfilled the inclusion criteria and were mostly heterozygous unless specified. 90 were novel and unique to this cohort and not previously reported in the GnomAD database. Of those novel variants, 22 were considered LP and 9 pathogenic. We identified variants in volunteers with gallstone disease (n=31), intrahepatic cholestasis of pregnancy (ICP, n=16), cholangiocarcinoma and cirrhosis (n=2). Fourteen novel LoF variants were identified: 7 frameshift, 5 introduction of premature stop codon and 2 splice acceptor variants. The rare variant burden was significantly increased in *ABCB11*. A total of 73 variants were assessed for impact at the protein level. Protein modelling demonstrated variants that appeared to likely cause significant structural damage.

Conclusions: This study highlights the significant genetic burden contributing to cholestatic liver disease. Novel likely pathogenic and pathogenic variants were identified addressing the underrepresentation of diverse ancestry groups in genomic research.

Introduction

Cholestatic liver disease encompasses a broad range of diagnoses that can present with fatigue, pruritus or jaundice[1]. Several genes are implicated, including the *ABCB4* gene coding for the canalicular phosphatidylcholine floppase ABCB4, and the *ABCB11* gene coding for the bile salt export pump (BSEP). Homozygous mutations in *ABCB4* and *ABCB11* cause a spectrum of disease from mild cholestasis to severe progressive familial intrahepatic cholestasis (PFIC), PFIC3 and PFIC2 respectively[2, 3]. *ABCB4* variants also increase the risk of developing drug-induced intrahepatic cholestasis, gallstone disease, gallbladder and bile duct carcinoma, liver cirrhosis and abnormal liver function tests[4]. Other canalicular transporters and their regulators are implicated in the pathogenesis of cholestatic liver disease e.g., *ATP8B1* (a P-type ATPase that flips phospholipids into the cytoplasmic leaflet of the membrane) [3], *NR1H4* (farnesoid X receptor (FXR)) gene[5], and *TJP2* (tight junction protein 2)[6]. While homozygous mutations of these genes are implicated in rare cases of severe familial cholestasis, the evidence base for a role of heterozygous mutations in milder forms of liver disease is limited. *ABCB4* and *ABCB11* are involved in 20% of patients with severe intrahepatic cholestasis of pregnancy (ICP). Heterozygous *ABCB4* variants have also been reported in ICP[7-13]. ICP is the commonest gestational liver disease[14] and may be complicated by preterm birth, meconium-stained amniotic fluid and stillbirth, in association with maternal serum bile acid concentrations ≥40µmol[15, 16]. In Europeans the incidence is 0.62% versus 1.24% in women of Indian and 1.46% of Pakistani origin[17]. Despite the increased prevalence of ICP and other liver conditions like non-alcoholic fatty liver disease in South Asian populations they often remain undiagnosed and under-investigated[18, 19].

Genes and Health is a long-term population-based cohort that assesses the health and disease in British Bangladeshi and British Pakistani people. Using the Genes & Health genomics (whole exome sequencing (WES)) and data linkage to electronic health records (EHR)[20], this study investigated

rare variation in a unique British Bangladeshi and Pakistani cohort around 5 candidate loci (*ABCB4*, *ABCB11*, *ATP8B1*, *NR1H4*, and *TJP2*) implicated in cholestatic liver disease.

Methods

Study population

Genes & Health has previously been described by Finer et al.[20]. Ethical approval for the study was provided by the South East London National Research Ethics Committee (14/LO/1240) including consent for publishing http://www.genesandhealth.org/volunteer-information)[22]. An individual application to support data access for this study was granted by Genes & Health (reference S00037).

Exome sequencing samples from 5236 Genes & Health volunteers reporting parental relatedness were available for analysis in variant call format files. For the initial analysis a genotype to phenotype approach was employed interrogating 5 gene candidate loci (Table 1). For the rare burden analysis female volunteers without ICP served as controls (n=3048). In a secondary analysis, a phenotype to genotype analysis was used to validate these findings, using ICP as the exemplar. For this approach, electronic health records allowed total serum bile acid concentrations ≥10 µmol to be retrieved from a network of acute hospitals that provide maternity care to (n=15,500) women per year living in east London to identify patients with a diagnosis of liver disease in pregnancy (ICD 10 diagnosis code O26.6), see Supplementary Figure 1. Maternal health records were screened by an experienced clinician to verify a diagnosis of ICP.

Exome sequencing & bioinformatic pipeline

Low/mid exome sequencing was performed as previously described[21]. The exome sequencing data is being held under a data access agreement at the European Genotype-phenome Archive (www.ebi.ac.uk/ega) under accession numbers EGAD00001005469. Minor allele frequency (MAF) was set at < 5% to include rare and low-frequency genetic variants to allow for a comprehensive evaluation.

Variant annotation

All protein-altering missense, non-sense, frameshift indels or splice site variants identified in the candidate gene set underwent the same processing as described below. Synonymous variants were excluded from further analysis. Variants were filtered and annotated if they met any of the following inclusion criteria (MAF < 5%): 1. associated with a phenotype; 2. known in the literature; 3. no recorded GnomAD allele frequency; 4. predicted to be likely pathogenic (LP) based on all 7 *in-silico* predictors. To assess the likelihood of functional impact of variants a variety of *in-silico* tools were employed, including Polyphen[22], SIFT[23], CADD[24], Revel[25], MetaLR[26], MetaSVM[26] and M_CAP[27]. Open-source databases (Leiden Open Variation Database – LoVD, and ClinVar) including a commercial database (Mastermind) were interrogated to assess whether variants were reported previously in the literature. The American College of Medical Genetics/Association for Molecular Pathology (ACMG/AMP) guidelines for rare variant interpretation was used to assess variant pathogenicity[28].

Rare variant burden analysis

To assess the significance of any rare variant burden of SNPs in all 5 gene candidates the exactCMC function in RVTESTS[29] was used. The burden was calculated as the proportion of all ICP cases versus control in the G&H cohort and who had at least one alternate allele. Variants with an allele frequency of 0.01 or less were included. ICP cases (n=18) from the Genomics England database (Project ID 747) – a predominantly European genetic cohort - were accessed to serve as a direct comparison to rare variant burden in G&H.

Protein structure and modelling analysis

Inclusion criteria for further protein structure and modelling analysis required variants to have an associated phenotype, and/or be predicted to be pathogenic based on all 7 *in-silico* tools. VarMap was used to assess protein sequence variants[30]. 2D representations were designed using the open

source tool TOPO2[31]. Variants that were LP or associated with a phenotype underwent further analysis using: (1) Dynamut[32] (2) CUPSAT[33] (3) SNPMuSiC[34]. 3D structural representations were generated using PyMOL software using *ABCB4* with phosphatidylcholine (PDB ID: 7NIV)[35], *ABCB11* open structure (PDB ID: 6LR0)[36], *ATP8B1* (PDB ID: 7PY4)[37], and *NR1H4* (PDB ID: 1OSH)[38]. There is no 3D protein structure available for *TJP2*.

Results

Genotype to phenotype analysis

Screening of the 5 candidate genes identified 300 variants; and 180 (166 non-synonymous and 14 loss of function (LoF) variants), were included in the analysis (Figure 1). Most variants identified were heterozygous unless otherwise specified. Further variant interpretation including scoring details of individual *in-silico* predictors, the impact of the coding substitution on disease propensity and conservation information for all variants are presented in Supplementary Table 1.

Phenotype to genotype analysis

We were able to validate variants of interest in 15 cases reporting ICP, most of whom had documented raised BA concentrations (Table 2). Further, variants discovered in 10 cases with raised BA but an uncertain or no diagnosis of ICP based on their EHR (Supplementary Table 2). This secondary analysis missed 2 individuals from the primary analysis with a diagnosis of ICP as there were no bile acid concentrations available for them. This analysis demonstrated a pragmatic approach to identifying disease causing variants and demonstrates the value of large genomic cohorts linked to electronic health data records.

Rare variant burden analysis

We observed in cases versus controls a significant enrichment of rare variants in *ABCB11* (p=0.00247). There was no enrichment in *ABCB4* (p=0.39138), *ATP8B1* (p=0.57957), *TJP2*

(p=0.17390), or *NR1H4* (p=0.70232). A further Fisher's exact analysis compared percentage of rare variants in ICP cases in G&H compared to Genomics England demonstrating that the rare genetic burden was significantly increased in tier 1 gene candidates in British Pakistani and Bangladeshi (*ABCB4*, p=0.0191; *ABCB11*, p=0.0191) but not in *ATP8B1* (p= 0.4857) *NR1H4* (p=0.2286) or *TJP2* (p=0.1039).

ABCB4 variants

There was a total of 68 *ABCB4* variants identified (Figures 1 and 2). 41 variants fulfilled the inclusion criteria. Variants were identified in people with cholestatic liver disease phenotypes: ICP (n=5 out of 88 women in the analysis), gallstone disease (n=7), and cholangiocarcinoma (n=1) (Table 3). For some identified variants a known cholestatic phenotype had previously been reported in the literature (n=9) whilst others had no phenotype reported (n=19) (Supplementary Table 3). We identified novel LoF variants (n=4): three frameshift (one associated with an ICP phenotype) and one introduction of a premature stop codon (no associated phenotype) (Table 4).

ABCB11 variants

There were 77 *ABCB11* variants identified of which 48 were included in the analysis (Figures 1 and 3). The associated cholestatic liver disease phenotypes identified were: ICP (n=5 out of 83 women in the analysis), and gallstone disease (n=10) (Table 3). Some were linked to cholestatic phenotypes previously reported in the literature (n=14), whilst for 19, no phenotype had been previously reported (n=19) (Supplementary Table 4). Likely novel LoF variants were identified (n=3): a frameshift, a splice-acceptor and an introduction of premature stop codon variant. These variants

ATP8B1 variants

We identified a total of 50 *ATP8B1* variants and 31 were included in the final analysis (Figures 1 and 4). There were 7 variants associated with gallstone disease; three appeared to be in linkage disequilibrium (LD) noted in three volunteers: M674T, I577V, and H78Q. The R384H variant was

associated with gallstone disease and with an ICP phenotype (in separate individuals, n=2 out of 33 women in the analysis). A further variant was associated with hepatitis-induced liver cirrhosis (I513T), and a final variant (D70N) was associated with liver cirrhosis, secondary malignant neoplasm of liver and bile duct, and gallstone disease (Table 3). In addition, previously reported cholestatic liver disease phenotypes (n=7) and variants with no previous reported phenotype were seen (n=15) (Supplementary Table 5). Finally, we identified 4 novel LoF variants: a frameshift/splice region, frameshift, splice-acceptor/coding-sequence, and stop-gain variant. The latter variant was associated with a gallstone disease phenotype (Table 4).

NR1H4 variants

There were 22 *NR1H4* variants in the Genes & Health cohort and 9 variants in the final analysis (Figure 1 and Supplementary Figure 2). We only identified an ICP phenotype (n=2 out of 33 women in the analysis) (Table 3) and otherwise novel variants that had no previous phenotype reported (n=7) (Supplementary Table 6). Furthermore, one novel LoF variant was identified without demonstrating a phenotype (Table 4).

TJP2 variants

There were 83 *TJP2* variants identified of which 37 were analysed (Figure 1). People with *TJP2* variants had ICP (n=3 out of 103 women in the analysis), gallstone disease (n=8), previously reported cholestatic liver disease phenotype (n=4), and 22 did not have a previously reported phenotype (Table 3, Supplementary Table 7). There were two novel LoF variants without a clinical phenotype (Table 4).

Protein structure and molecular modelling

A flow chart illustrating the variants included in this analysis is shown in Supplementary Figure 3. Results of the protein structure and molecular modelling software tools are presented in Supplementary Table 8.

Some novel variants are in regions of transporters for which we can hypothesis a mechanistic impact. Of particular interest are Q1106H in ABCB4 and D191A in ABCB11. These ABC B-family transporters share 48% amino acid identity and are very likely have a common mechanism of action. The two amino acids are conserved in both proteins, and we propose that they are involved in energy transduction through the transporter in order to couple the substrate efflux cycle to the ATP binding and catalysis cycle.

In ABCB4 and ABCB11, two transmembrane domains (TMDs) bind the transport substrates (phosphatidylcholine (PC) and bile acids), respectively. The conformational changes required for substrate transport are driven by ATP hydrolysis at the interface between two nucleotide binding domains (NBDs). The TMDs and NBDs must therefore be intimately coupled, and this is achieved via four 'coupling helices' (CH) located at the base of the long intracellular loops extending from the transmembrane alpha helices of the TMDs (Supplementary figure 4A).

Q1106 (ABCB4) and D191 (ABCB11) are particularly interesting because they are located at this interface that is conserved in both ABCB4 and ABCB11. Q1106 is in a groove on the surface of NBD2 where it interacts with CH2 (Supplementary Figure 4B).

In the PC-bound conformation of ABCB4, Q1106 forms a weak electrostatic interaction with the peptide bond of G270. In the ATP-bound conformation (from which PC has most likely been released), Q1106 now interacts with Q272 which illustrates the movement of CH2 and its changing juxtaposition with the NBD during the transport cycle; essentially, a hinge region. The geometry of these interactions will not be preserved if Q1106 is replaced by histidine. In ABCB11, this triad is preserved in Q1150 and G295, with E297 providing a conservative change for the glutamine in CH2 (with respect to formation of an equivalent electrostatic bond with Q1150).

In the sole structural model that we have for ABCB11, D191 is in CH1 where it interacts with Y472 in a surface groove of NBD1 and also, intriguingly, with R946 which is in CH4, suggesting that CH1 and CH4 likely work together in energy transduction through the transporter (Supplementary Figure 4C).

These electrostatic bonds will not be possible if D191 is replaced with alanine. This triad and its bond architecture is perfectly conserved in ABCB4 in the ATP bound conformation through amino acids D166, Y446 and R902. However, in ABCB4 there is also an additional electrostatic interaction between the carbonyl of the D166 peptide bond and the side chain of Q1171. Q1171 (which is conserved in ABCB11) is adjacent to the ABC signature motif ^{1172}LSGGQ^{1176} which is involved in coordination of ATP and provides a direct mechanism for how CH1 influences, and responds to, the ATP catalytic cycle of these transporters.

Discussion

This study identified novel variants implicated in the aetiopathogenesis of cholestatic liver disease that occur uniquely in this British Bangladeshi and Pakistani cohort [39-42]. There have not been any other studies of this magnitude analysing the burden of mutational variation in cholestatic liver disease in a large South Asian cohort. Using a genotype to phenotype approach we discovered novel likely pathogenic variants that appear to be unique to this cohort. We then employed a phenotype to genotype analysis using the ICP phenotype as an exemplar, which offered a pragmatic interrogation of electronic health records to identify rare genetic variants that are likely pathogenic. Thus, this study improves representation of this distinct population especially as prevalence of cholestatic liver disease is increased in the Genes & Health cohort, e.g., 1.54% are affected by ICP compared to white Europeans (0.62%). This study demonstrates the importance of multi-ancestry genomic research and offers the potential of tailored treatment for this population.

In the Genes & Health cohort, out of 194 variants meeting inclusion criteria we identified 53 that had a cholestatic liver disease phenotype reported in their linked EHR. Of those, 16 are unique to this

British Bangladeshi and Pakistani population and a large number were predicted to be likely pathogenic or pathogenic based on *in-silico* prediction tools. In addition, there were 35 variants that were previously reported in the literature with a cholestatic phenotype. However, 87 variants had no previously reported phenotype; 67 were novel (34% of all variants analysed in this study) as they were also not previously reported in the GnomAD population database. Despite that, 9 were considered likely pathologic and 5 pathogenic. It is important to consider that heterozygosity as noted in most cases means that they are likely rescued by the wild-type allele but at higher risk of disease in later life or during times of liver stress, e.g. during pregnancy.

Our findings reflect the difficulty with interpretation of rare variants in clinically important genes when there is no previous evidence in the literature or functional data to interpret them further. The ACMG rare variant interpretation guideline[30] provides a standardised analysis pathway. However, it relies in part on the interpretation of the variant in the context of the literature and does not account for specific genes and diseases. It also may not be robust for flexible membrane proteins which do not work by lock and key mechanism. For example, the *ABCB11* variant V444A, considered as benign by the ACMG criteria, has been reported to increase the risk of ICP, hepatitis C disease progression, and drug-induced liver injury although the exact functional mechanisms are not clear yet[43, 44].

By employing computational protein modelling software tools, we were able to identify variants that likely have a significant impact on the conformation of the protein and could therefore be of clinical significance. It is important to bear in mind that all these tools have inherent flaws and are beyond the scope of this paper to discuss in detail. By taking ICP as a cholestatic liver disease example we were able to highlight further difficulties with rare variant interpretation in gestational syndromes as the inherent transient nature of the disease makes variant interpretation challenging. However, ICP is a clinically relevant example as the gestational disease consequences are not just relevant to their current pregnancy but also can result in later hepatobiliary disorders such as cancer, immune-

mediated and cardiovascular diseases[45]. In addition, they have a higher gallstone-related morbidity and a strong positive association between ICP and hepatitis C exists as well[46].

Limitations

The use of electronic health records to determine phenotype is extremely useful but dependent upon appropriate information having been coded. Participants with at-risk variants may not have presented yet with symptoms of disease but still be at high risk of developing complications at a later stage in their life, particularly given that the median age of volunteers in this study was around 45 years. It demonstrates the difficulty with interpreting variants when recalling the genotype first.

Conclusions

In this study we provide the first comprehensive evaluation of gene candidates associated with cholestatic liver diseases in a unique cohort of British Bangladeshi and Pakistani origin demonstrating the importance of characterising genetic disease in diverse ethnic groups. Our findings have demonstrated the increased mutational burden of cholestatic liver disease in British Bangladeshi and Pakistani people who thus far remain understudied despite their distinct genetic background and increased risk of developing ICP in comparison to other populations. We were able to identify novel variants that have not been previously identified and are likely implicated in disease. We demonstrated the ability to identify participants at risk both by a phenotype or genotype first approach. This demonstrates the importance of providing more personalised care in a clinical setting as identification of high-risk individuals and their family members enables early intervention to prevent adverse outcomes, for example hepato-protective drugs such as UDCA, in addition to hepatic surveillance. Furthermore, it provides the necessary foundation for improved therapy and drug development.

References

1. *EASL Clinical Practice Guidelines: management of cholestatic liver diseases.* J Hepatol, 2009. **51**(2): p. 237-67.
2. Sticova, E. and M. Jirsa, *ABCB4 disease: Many faces of one gene deficiency.* Annals of Hepatology, 2020. **19**(2): p. 126-133.
3. Huynh, M.T., T.T. Nguyen, S. Grison, et al., *Clinical characteristics and genetic profiles of young and adult patients with cholestatic liver disease.* Rev Esp Enferm Dig, 2019. **111**(10): p. 775-788.
4. Stättermayer, A.F., E. Halilbasic, F. Wrba, et al., *Variants in ABCB4 (MDR3) across the spectrum of cholestatic liver diseases in adults.* J Hepatol, 2020. **73**(3): p. 651-663.
5. Davit-Spraul, A., E. Gonzales, and E. Jacquemin, *NR1H4 analysis in patients with progressive familial intrahepatic cholestasis, drug-induced cholestasis or intrahepatic cholestasis of pregnancy unrelated to ATP8B1, ABCB11 and ABCB4 mutations.* Clin Res Hepatol Gastroenterol, 2012. **36**(6): p. 569-73.
6. Sambrotta, M., S. Strautnieks, E. Papouli, et al., *Mutations in TJP2 cause progressive cholestatic liver disease.* Nat Genet, 2014. **46**(4): p. 326-8.
7. de Vree, J.M., E. Jacquemin, E. Sturm, et al., *Mutations in the MDR3 gene cause progressive familial intrahepatic cholestasis.* Proc Natl Acad Sci U S A, 1998. **95**(1): p. 282-7.
8. Jacquemin, E., D. Cresteil, S. Manouvrier, et al., *Heterozygous non-sense mutation of the MDR3 gene in familial intrahepatic cholestasis of pregnancy.* The Lancet, 1999. **353**(9148): p. 210-211.
9. Dixon, P.H., N. Weerasekera, K.J. Linton, et al., *Heterozygous MDR3 missense mutation associated with intrahepatic cholestasis of pregnancy: evidence for a defect in protein trafficking.* Human Molecular Genetics, 2000. **9**(8): p. 1209-1217.
10. Lucena, J.-F., J.I. Herrero, J. Quiroga, et al., *A multidrug resistance 3 gene mutation causing cholelithiasis, cholestasis of pregnancy, and adulthood biliary cirrhosis.* Gastroenterology, 2003. **124**(4): p. 1037-1042.
11. Müllenbach, R., K.J. Linton, S. Wiltshire, et al., *ABCB4 gene sequence variation in women with intrahepatic cholestasis of pregnancy.* Journal of Medical Genetics, 2003. **40**(5): p. e70-e70.
12. Pauli-Magnus, C., T. Lang, Y. Meier, et al., *Sequence analysis of bile salt export pump (ABCB11) and multidrug resistance p-glycoprotein 3 (ABCB4, MDR3) in patients with intrahepatic cholestasis of pregnancy.* Pharmacogenetics and Genomics, 2004. **14**(2).
13. Floreani, A., I. Carderi, D. Paternoster, et al., *Intrahepatic cholestasis of pregnancy: three novel MDR3 gene mutations.* Alimentary Pharmacology & Therapeutics, 2006. **23**(11): p. 1649-1653.
14. Williamson, C. and V. Geenes, *Intrahepatic cholestasis of pregnancy.* Obstet Gynecol, 2014. **124**(1): p. 120-133.
15. Glantz, A., H.U. Marschall, and L.A. Mattsson, *Intrahepatic cholestasis of pregnancy: Relationships between bile acid levels and fetal complication rates.* Hepatology, 2004. **40**(2): p. 467-74.
16. Geenes, V., L.C. Chappell, P.T. Seed, et al., *Association of severe intrahepatic cholestasis of pregnancy with adverse pregnancy outcomes: a prospective population-based case-control study.* Hepatology, 2014. **59**(4): p. 1482-91.
17. Geenes, V. and C. Williamson, *Intrahepatic cholestasis of pregnancy.* World journal of gastroenterology, 2009. **15**(17): p. 2049-2066.
18. Alazawi, W., R. Mathur, K. Abeysekera, et al., *Ethnicity and the diagnosis gap in liver disease: a population-based study.* The British journal of general practice : the journal of the Royal College of General Practitioners, 2014. **64**(628): p. e694-e702.

19. Szanto, K.B., J. Li, P. Cordero, et al., *Ethnic differences and heterogeneity in genetic and metabolic makeup contributing to nonalcoholic fatty liver disease.* Diabetes, metabolic syndrome and obesity : targets and therapy, 2019. **12**: p. 357-367.
20. Finer, S., H.C. Martin, A. Khan, et al., *Cohort Profile: East London Genes & Health (ELGH), a community-based population genomics and health study in British Bangladeshi and British Pakistani people.* International Journal of Epidemiology, 2019. **49**(1): p. 20-21i.
21. Narasimhan, V.M., K.A. Hunt, D. Mason, et al., *Health and population effects of rare gene knockouts in adult humans with related parents.* Science, 2016. **352**(6284): p. 474-477.
22. Adzhubei, I.A., S. Schmidt, L. Peshkin, et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.
23. Sim, N.-L., P. Kumar, J. Hu, et al., *SIFT web server: predicting effects of amino acid substitutions on proteins.* Nucleic Acids Research, 2012. **40**(W1): p. W452-W457.
24. Kircher, M., D.M. Witten, P. Jain, et al., *A general framework for estimating the relative pathogenicity of human genetic variants.* Nature Genetics, 2014. **46**(3): p. 310-315.
25. Ioannidis, N.M., J.H. Rothstein, V. Pejaver, et al., *REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants.* The American Journal of Human Genetics, 2016. **99**(4): p. 877-885.
26. Liu, X., C. Li, C. Mou, et al., *dbNSFP v4: a comprehensive database of transcript-specific functional predictions and annotations for human nonsynonymous and splice-site SNVs.* Genome Medicine, 2020. **12**(1): p. 103.
27. Jagadeesh, K.A., A.M. Wenger, M.J. Berger, et al., *M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity.* Nature Genetics, 2016. **48**(12): p. 1581-1586.
28. Richards, S., N. Aziz, S. Bale, 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. **17**(5): p. 405-24.
29. Zhan, X., Y. Hu, B. Li, et al., *RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data.* Bioinformatics, 2016. **32**(9): p. 1423-1426.
30. Stephenson, J.D., R.A. Laskowski, A. Nightingale, et al., *VarMap: a web tool for mapping genomic coordinates to protein sequence and structure and retrieving protein structural annotations.* Bioinformatics, 2019. **35**(22): p. 4854-4856.
31. Johns, S.J. *TOPO2. Transmembrame protein display software*. Available from: http://www.sacs.ucsf.edu/TOPO2/.
32. Rodrigues, C.H., D.E. Pires, and D.B. Ascher, *DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability.* Nucleic Acids Research, 2018. **46**(W1): p. W350-W355.
33. Parthiban, V., M.M. Gromiha, and D. Schomburg, *CUPSAT: prediction of protein stability upon point mutations.* Nucleic Acids Research, 2006. **34**(suppl_2): p. W239-W242.
34. Ancien, F., F. Pucci, M. Godfroid, et al., *Prediction and interpretation of deleterious coding variants in terms of protein structural stability.* Scientific Reports, 2018. **8**(1): p. 4480.
35. Nosol, K., R. Bang-Sørensen, R.N. Irobalieva, et al., *Structures of ABCB4 provide insight into phosphatidylcholine translocation.* Proc Natl Acad Sci U S A, 2021. **118**(33).
36. Wang, L., W.T. Hou, L. Chen, et al., *Cryo-EM structure of human bile salts exporter ABCB11.* Cell Res, 2020. **30**(7): p. 623-625.
37. He, Y., J. Xu, X. Wu, et al., *Structures of a P4-ATPase lipid flippase in lipid bilayers.* Protein Cell, 2020. **11**(6): p. 458-463.
38. Downes, M., M.A. Verdecia, A.J. Roecker, et al., *A chemical, genetic, and structural analysis of the nuclear bile acid receptor FXR.* Mol Cell, 2003. **11**(4): p. 1079-92.

39. Dixon, P.H., M. Sambrotta, J. Chambers, et al., *An expanded role for heterozygous mutations of ABCB4, ABCB11, ATP8B1, ABCC2 and TJP2 in intrahepatic cholestasis of pregnancy.* Scientific Reports, 2017. **7**(1): p. 11823.
40. Liu, X., H. Lai, S. Xin, et al., *Whole-exome sequencing identifies novel mutations in ABC transporter genes associated with intrahepatic cholestasis of pregnancy disease: a case-control study.* BMC pregnancy and childbirth, 2021. **21**(1): p. 110-110.
41. Avena, A., S. Puggelli, M. Morris, et al., *ABCB4 variants in adult patients with cholestatic disease are frequent and underdiagnosed.* Digestive and Liver Disease, 2021. **53**(3): p. 329-344.
42. Wright, C.F., B. West, M. Tuke, et al., *Assessing the Pathogenicity, Penetrance, and Expressivity of Putative Disease-Causing Variants in a Population Setting.* Am J Hum Genet, 2019. **104**(2): p. 275-286.
43. Dixon, P.H., S.W. van Mil, J. Chambers, et al., *Contribution of variant alleles of ABCB11 to susceptibility to intrahepatic cholestasis of pregnancy.* Gut, 2009. **58**(4): p. 537-44.
44. Ali, I., S. Khalid, B. Stieger, et al., *Effect of a Common Genetic Variant (p.V444A) in the Bile Salt Export Pump on the Inhibition of Bile Acid Transport by Cholestatic Medications.* Molecular pharmaceutics, 2019. **16**(3): p. 1406-1411.
45. Wikström Shemer, E., H.U. Marschall, J.F. Ludvigsson, et al., *Intrahepatic cholestasis of pregnancy and associated adverse pregnancy and fetal outcomes: a 12-year population-based cohort study.* Bjog, 2013. **120**(6): p. 717-23.
46. Marschall, H.-U., E. Wikström Shemer, J.F. Ludvigsson, et al., *Intrahepatic cholestasis of pregnancy and associated hepatobiliary disease: A population-based cohort study.* Hepatology, 2013. **58**(4): p. 1385-1391.

Legends

Figure legends

Figure 1.

Overall summary of mutational burden discovered in the Genes and Health cohort for all five gene candidates. a. Overall summary, describes all non-synonymous variants identified; b. Pathogenicity, describes breakdown of pathogenicity; c. Phenotypes, describes phenotypes identified. Inclusion criteria (<5% MAF), non-synonymous and at least any of the following: 1. Include all variants with a phenotype 2. Include all variants known in the literature 3. Include all variants with no GnomAD allele frequency but ELGH allele frequency 4. Include all variants with *in-silico* prediction of 7

Figure 2.

ABCB4 variant summary in a 2-dimensional illustration. Variants are divided into their phenotypic presentation: No phenotype previously reported, cholestatic phenotype reported in the literature, gallstone disease, cholangiocarcinoma, and intrahepatic cholestasis of pregnancy. Bold border represents variants that are unique to the Genes & Health cohort.

Figure 3.

ABCB11 variant summary in a 2-dimensional illustration. Variants are divided into their phenotypic presentation: No phenotype previously reported, cholestatic phenotype reported in the literature, gallstone disease, and intrahepatic cholestasis of pregnancy. Bold border represents variants that are unique to the Genes & Health cohort.

Figure 4.

ATP8B1 variant summary in a 2-dimensional illustration. Variants are divided into their phenotypic presentation: No phenotype previously reported, cholestatic phenotype reported in the literature, gallstone disease, liver cirrhosis, Liver cirrhosis, multiple cholestatic phenotype, and intrahepatic cholestasis of pregnancy. Bold border represents variants that are unique to the Genes & Health cohort.

Supplementary Figure legends

Supplementary Figure 1.

Flow diagram illustrating the patient inclusion for the phenotype to genotype analysis. This diagram illustrates the number of volunteers included with a phenotypic presentation of intrahepatic cholestasis of pregnancy selected from the 5236 exome sequencing data available in the Genes and Health database.

Supplementary Figure 2.

NR1H4 variant summary in a 2-dimensional illustration. NR1H4 also known as FXR is a nuclear receptor within hepatocytes. Variants are divided into their phenotypic presentation: No phenotype previously reported, and intrahepatic cholestasis of pregnancy. Bold border represents variants that are unique to the Genes & Health cohort.

Supplementary Figure 3.

Overall summary of all variants included for the protein modelling in the Genes and Health cohort for all five gene candidates. All variants included for the analysis were modelled and stratified by their predicted structural effect on the 3D structure. *TJP2 structure not included as no 3D structure available.

Supplementary Figure 4.

Rare variants at the energy transduction interface between the ligand binding and ATP catalytic domains of ABCB4 and ABCB11

A. Topological cartoon showing the organisation of the transmembrane domains (TMDs) and the interaction with the nucleotide binding domains (NBDs) of ABCB4 and ABCB11. Each TMD comprises 6 transmembrane alpha helices (TMH 1-6 of TMD1 in rainbow colouring, TMH 7-12 of TMD2 in grey). The intracellular loops between TMH 2 and 3, 4 and 5, 8 and 9, and 10 and 11 extend into the cytosol to interact with the NBDs (blue and pink squares represent NBD1 and NBD2 respectively). At the base of each of these intracellular loops is a coupling helix (labelled CH1-4) which interdigitate with grooves on the top surface of the NBDs. The coupling helices interact with the surface of the NBDs and are assumed to convey conformational changes in a bidirectional manner: to the NBDs from the TMDs to signal substrate binding to the latter and, vice versa to the TMDs from the NBDs to drive the conformational changes associated with ATP binding (resulting in substrate efflux) and the conformational changes associated with ATP hydrolysis to reset the transporter for binding of the next transport substrate. The domain colour coding is preserved in 4B and 4C.

B. ABCB4 variant Q1106H is in a functionally sensitive location for energy transduction, disrupting the interaction between CH2 and NBD2. Left panel; cartoon representation of ABCB4 with bound PC. PC is shown in sphere format and coloured elementally (pdb: 7NIV). Q1106 also shown as spheres with elemental colouring is located in a groove on the top surface of NBD2 (pink). The groove is occupied by CH2 formed between TMH4 (yellow) and TMH5 (orange). The distance between the side chain nitrogen (blue) of Q1106 and the nearest atom in the coupling helix, the oxygen (red) of the peptide bond formed by G270 is indicated in the zoomed image where the relevant amino acids are shown in stick format with elemental colouring (nitrogen in blue, oxygen in red, carbon in green). Right panel; cartoon representation of ABCB4 with 2 x Mg^{2+} and 2 x ATP (pdb: 6S7P). Q1106 is now in close proximity to Q272 in CH2. The distance between the side chain nitrogen (blue) of Q1106 and the oxygen (red) of the Q272 is indicated in the zoomed image. Interatom distances were measured in Pymol.

C. ABCB11 variant D191A is in a functionally sensitive location for energy transduction, disrupting the interaction between CH1 and NBD1. Left panel; cartoon representation of ABCB11 (pdb: 6LR0). D191 which is shown as spheres with elemental colouring is located in CH1 where it interacts with both Y472 on the top surface of NBD1 (blue) and R946 from CH4. Minimum distances between the oxygens of the carboxylic acid side chain of D191 and the oxygen of Y472 and the nitrogen of the guanidinium group of R946 are shown in the zoomed image where the relevant amino acids are shown in stick format with elemental colouring as above.

Table 1. ***Description of the five gene candidates.***

Genes	Chr	Gene product	Omim ID	Exons	Length	Associated disease
ABCB4	7 (7q21.1)	.	171060	27	81kb	PFIC-3
ABCB11	2 (2q31.1)	BSEP	603201	28	115kb	PFIC-2
ATP8B1	18 (18q21.31)	FIC1	602397	28	157kb	PFIC-1
TJP2	9 (9q21.11)	.	607709	25	153kb	PFIC-4
NR1H4	12 (12q23.1)	FXR	603826	11	90kb	PFIC-5

Footnote: Abbreviations: Chr, chromosomes.

Table 2. ***Variants identified and TSBA concentration in volunteers with a diagnosis of ICP based on electronic health records.***

All variants in volunteers with raised BA and diagnosis of ICP					
Volunteers	**Gene**	**Variants**	**Zygosity**	**Type**	**Highest BA (umol/L)**
3	ABCB4	G1254S / G1261S	het	Non-synonymous	80.1
3	ABCB11	N591S	het	Non-synonymous	
4	ABCB4	P1050S	het	Non-synonymous	53
5	TJP2	Q105K	hom	Non-synonymous	15
6	ABCB11	N591S	het	Non-synonymous	17
13	TJP2	Q105K	het	Non-synonymous	25
15	ATP8B1	R384H	het	Non-synonymous	31.6
15	TJP2	Q105K	hom	Non-synonymous	
18	ABCB11	N591S	het	Non-synonymous	115
18	TJP2	R21H	het	Non-synonymous	
20	ABCB11	N591S	het	Non-synonymous	18
20	NH1R4	M173T	het	Non-synonymous	
21	ABCB11	M677V	het	Non-synonymous	14
21	NH1R4	N358H	het	Non-synonymous	
22	ABCB11	N591S	het	Non-synonymous	14
25	ABCB4	S99x	het	frameshift	55
26	ABCB11	V284A	het	Non-synonymous	35
26	ABCB4	N510S	het	Non-synonymous	
31	ABCB4	T175A	het	Non-synonymous	32
32	ABCB11	D1284N	het	Non-synonymous	.
32	ABCB11	N591S	het	Non-synonymous	
38	ABCB11	N591S	het	Non-synonymous	45.2

Foot note: Abbreviations: TSBA, total serum bile acid concentrations; ICP, intrahepatic cholestasis of pregnancy.

Table 3. ***Non-synonymous variants identified in the five gene candidates associated with a cholestatic phenotype in the Genes and Health cohort.***

Gene	Clinical phenotype	Transcript	Protein change	dbSNP	gnomAD AF	G&H AF ^	ACMG-AMP	ACMG-AMP criteria	Clinvar	Ref
ABCB4	Intrahepatic cholestasis of pregnancy	ENSP00000496956.1:p.Gly1254Ser	G1254S *	rs781315185	0.00003656	0.00028843	LP	PM1, PM2, PP2, PP3	.	[47]
		ENSP00000497931.1:p.Asp1284Asn	P1050S *	.	.	0.00019146	LP	PM1, PM2, PP2, PP3	.	.
		ENSP00000496956.1:p.Ala833Thr	A833T *	.	.	0.00009638	LP	PM1, PM2, PP2, PP3	.	.
		ENSP00000496956.1:p.Asn510Ser	N510S	rs375315619	0.00019110	0.00057394	LP	PM1, PM2, PP2, PP3, PP5	LP	[48-55]
		ENSP00000496956.1:p.Thr175Ala	T175A [1.]	rs58238559	0.01155000	0.01315030	LB	PM1, PP2, PP3, BS1, BS2, BP6	benign/likely benign	[3, 12, 39, 48, 50, 51, 53-67]
	Gallstone disease	ENSP00000496956.1:p.Arg1137Gln	R1137Q	rs780738927	0.00003250	0.00028846	LP	PM1, PM2, PP2, PP3	.	.
		ENSP00000496956.1:p.Gly826Arg	G826R *	.	.	0.00028681	LP	PM1, PM2, PP2, PP3	.	.
		ENSP00000496956.1:p.Arg788Leu	R788L	rs8187801	0.00000813	0.00009579	LP	PM1, PM2, PP2, PP3	benign	.
		ENSP00000496956.1:p.Asp686Asn	D686N	rs78653500	.	0.00009586	VUS	PM2, PP2, BP4	.	.
		ENSP00000496956.1:p.Met676Ile	M676I	rs376702091	0.00002033	0.00038278	VUS	PM2, PP2, BP4	.	.

		ENSP00000496956.1:p.Thr651Asn	T651N	rs45476795	0.0005776	0.0006719	LB	PM2, PP2, BP4, BP6	conflicting	·
		ENSP00000496956.1:p.Lys391Glu	K391E	rs781347049	0.00002440	0.00009582	LB	PM2, PP2, BP4, BP6	.	·
	Cholangiocarcinoma	ENSP00000496956.1:p.Gln668His	Q668H *	.	.	0.00009586	LB	PM2, PP2, BP4	.	·
ABCB11	Intrahepatic cholestasis of pregnancy	ENSP00000497931.1:p.Asp1284Asn	D1284N	rs766784155	0.00001228	0.00028780	LP	PM1, PM2, PP2, PP3	.	·
		ENSP00000497931.1:p.Arg1050His	R1050H	rs72549398	0.00000421	0.00019135	VUS	PM2, PP2, BP4	.	·
		ENSP00000497931.1:p.Met677Val	M677V [2.]	rs11568364	0.02364000	0.01005750	Benign	PP2, BA1, BS3, BP6, BP4	benign	[12, 67-76]
		ENSP00000497931.1:p.Asn591Ser	N591S [3.]	rs11568367	0.01436000	0.12647200	Benign	PM1, PP2, BA1, BP6	benign	[3, 12, 39, 43, 53, 70, 76-84]
		ENSP00000497931.1:p.Val284Ala	V284A	rs200739891	0.00026040	0.00009558	LP	PM1, PM2, PM5, PP2, PP3, BP6	conflicting	[56, 72, 74]
	Gallstone disease	ENSP00000497931.1:p.Ala1260Pro	A1260P	rs772097949	0.00001641	0.00028153	LP	PM1, PM2, PP2, PP3	VUS	·
		ENSP00000497931.1:p.Gln976Arg	Q976R	rs199940188	0.00054840	0.00066883	VUS	PM1, PM2, PP2, BP4	conflicting	·
		ENSP00000497931.1:p.Ala926Ser	A926S *	.	.	0.00040667	LP	PM1, PM2, PM5, PP2,PP3	.	·
		ENSP00000497931.1:p.Ala679Val	A679V	rs200912109	0.00045560	0.00143761	VUS	PM2, PP2, BP4	conflicting	·
		ENSP00000497931.1:p.Asn539As	N539D *	.	.	0.0002260	VUS	PM1, PM2, PP2, BP4	.	·

		p				4				
		ENSP00000497931.1:p.Arg487Cys	R487C	rs770693935	0.00002043	0.00009549	LP	PM1, PM2, PP2, PP3	.	[85]
		ENSP00000497931.1:p.Ala311Thr	A311T	rs200509511	0.00004073	0.00028969	LP	PM1, PM2, PP2, PP3	.	.
		ENSP00000497931.1:p.Val95Ile	V95I	rs201735739	0.00009766	0.00028708	Benign	PM1, PM2, PP2, BP4	.	.
		ENSP00000497931.1:p.Asp94Asn	D94N	rs760920706	0.00010170	0.00095621	LP	PM1, PM2, PP2	conflicting	.
		ENSP00000497931.1:p.Lys12Arg	K12R *	.	.	0.00010378	LP	PM1, PM2, PP2, PP3	.	.
ATP8B1	Intrahepatic cholestasis of pregnancy Gallstone disease	ENSP00000497896.1:p.Arg384His	R384H #	rs2271260	0.00026400	0.00048040	VUS	PM2, PP2, BP6	.	[3]
	Gallstone disease	ENSP00000497896.1:p.Val1161Ala	V1161A	rs1255793857	0.00000406	0.00009549	VUS	PM2, PP2	.	.
		ENSP00000497896.1:p.Thr1092Ile	T1092I	rs780425796	0.00001220	0.00030581	VUS	PM2, PP2, PP3	.	.
		ENSP00000497896.1:p.Met674Thr	M674T + [4.]	rs35470719	0.00456300	0.00632063	Benign	PP2, BA1, BP4, BP6	Benign/Likely benign	[76, 86-91]
		ENSP00000497896.1:p.Ile577Val	I577V + [5.]	rs3745078	0.00467800	0.00628992	Benign	PP2, BA1, BP6	Benign	[76, 86-89, 91]
		ENSP00000497896.1:p.His78Gln	H78Q + [6.]	rs3745079	0.00421800	0.00495751	Benign	PP2, BP4,BP6, BS1, BS2	Benign	[76, 86-

										89, 91]
		ENSP00000497896.1:p.Asp14Tyr	D14Y *	.	.	0.00009560	VUS	PM1, PM2, PP2, BP4	.	.
	Cirrhosis	ENSP00000497896.1:p.Ile513Thr	I513T	rs772028343	0.00008531	0.00066973	VUS	PM2, PP2	.	.
	Cirrhosis, secondary malignant neoplasm of liver and bile duct, gallstone disease	ENSP00000497896.1:p.Asp70Asn	D70N [7.]	rs34719006	0.00313900	0.00302678	VUS	PM2, PP2, PP3	Conflicting	[49, 87, 92-98]
NR1H4	Intrahepatic cholestasis of pregnancy	ENSP00000496908.1:p.Asn358His	N358H	rs149287629	0.00041020	0.00038307	VUS	PM2	VUS	.
		ENSP00000496908.1:p.Met173Thr	M173T	rs61755050	0.00374800	0.00267482	Likely benign	PM1, PP3, BS1, BS2, BP6	likely benign	[6, 99]
TJP2	Intrahepatic cholestasis of pregnancy	ENSP00000497787.1:p.Thr377Ala	T377A	rs766748789	0.00000406	0.00047765	VUS	PM2, BP4	.	.
		ENSP00000496791.1:p.Gln105Lys	Q105K [8.]	rs41305539	0.05150000	0.12031400	Benign	BA1, BP4, BP6,	benign	[68, 100, 101]
		ENSP00000497861.1:p.Arg21His	R21H [9.]	rs4493966	0.07416000	0.04555170	Benign	BA1, BP4, BP6,	benign	.
	Gallstone disease	ENSP00000496791.1:p.Gln8Arg	Q8R *	.	.	0.00009553	VUS	PM2, PP3	.	.
		ENSP00000496791.1:p.Thr68Asn	T68N *	.	.	0.00019150	VUS	PM1, PM2	.	.

		ENSP00000496791.1:p.Pro152Leu	P152L	rs754300892	0.00007876	0.00046685	VUS	BP4	.	.
		ENSP00000497787.1:p.Arg178Cys	R178C	rs199761505	0.00043060	0.00223305	VUS	PP3	.	.
		ENSP00000497787.1:p.Arg255His	R255H	rs532438219	0.00012990	0.00066947	VUS	.	.	.
		ENSP00000497787.1:p.Arg461Pro	R461P	rs748523814	0.00009746	0.00078125	VUS	PP3	.	.
		ENSP00000496791.1:p.Thr902Met	T902M	rs774198938	0.00010970	0.00010449	VUS	PP3	.	.
		ENSP00000496791.1:p.Arg1070Lys	R1070K *	.	.	0.00009564	VUS	PM2, BP4	VUS	.

Foot note: Abbreviations: AF, allele frequency; ACMG-AMP, American College of Medical Genetics and Genomics and the Association for Molecular Pathology (BP, benign supporting; PM, pathogenic moderate; PP, pathogenic supporting); G&H, Genes & Health; LP, likely pathogenic; Ref, references; VUS, variant of unknown significance. Symbols: * South Asian specific variant; # multiple phenotype; +, Linkage disequilibrium. ^Allele frequency specific to East London Genes & Health cohort. [1.] T175A, Hom (n) 6, Het (n) 126; [2.] M677V, Hom (n) 3, Het (n) 99; [3.] N591S, Hom (n) 99, Het (n) 760; [4.] M674T, Hom (n) 2, Het (n) 67; [5.] I577V, Hom (n) 2, Het (n) 61; [6.] H78Q, Hom (n) 1, Het (n) 47; [7.] D70N, Hom (n) 1, Het (n) 24; [8.] Q105K, Hom (n) 76, Het (n) 711; [9.] R21H, Hom (n) 16, Het (n) 443.

Table 4. ***Loss of function variants identified in the five gene candidates.***

Phenotype	Gene	Transcript	Protein change	Info	gnomAD AF	G&H AF ^	ACMG-AMP	ACMG-AMP criteria	Clinvar
ICP	ABCB4	ENSP00000395716.1:p.Ser99LeufsTer11	S99x	frameshift	0.00039970	0.00336538	P	PVS1, PM2, PP3	.
.		ENSP00000392983.1:p.Leu759TyrfsTer38	F758x	frameshift	.	0.00009610	LP	PVS1, PM2	.
.		ENSP00000392983.1:p.Lys30GlyfsTer7	Lys30Glyfster	frameshift	0.0000040	0.0002024	LP	PVS1, PM2	.

			7		8	3			
.		ENSP00000392983.1:p.Arg595Ter	R595*	stop-gained	0.00001627	0.00009593	P	PVS1, PM2, PP3, PP5	pathogenic
Gallstone disease	ABCB11	ENSP00000497931.1:p.Ala1044LeufsTer53	A1044x	frameshift	.	0.00009562	P	PVS1, PM2, PP3	.
.		ENST00000263817.7:c.2611-2A>T	c.2611-2A>T	splice-acceptor-variant	0.00000407	0.00057870	P	PVS1, PM2, PP3	.
.		ENSP00000497931.1:p.Trp239Ter	W239x	stop-gained	.	0.00009566	P	PVS1, PM2, PP3	.
.	ATP8B1	ENSP00000283684.4:p.Gln1179GlufsTer56	IQ1178-1179IX	frameshift_variant & splice_region_variant	.	0.00193798			.
.		ENSP00000283684.4:p.Pro792HisfsTer8	F791X	frameshift_variant	.	0.00019069	P	PVS1, PM2, PP3	.
Gallstone disease		ENST00000283684.9:c.182-4_183del	?-61	splice_acceptor_variant & coding_sequence_variant & intron_variant	.	0.00048956			.
		ENSP00000283684.4:p.Glu20Ter	E20*	stop_gained	.	0.00009566	P	PVS1, PM2, PP3	.
.	NR1H4	ENSP00000446760.1:p.Lys4Ter	K4*	stop_gained	.	0.00014188	P	PVS1, PM2, PP3	.
.	T	ENSP00000438262.1:p.Glu44Ter	E44*	stop_gained	.	0.0000960	LP	PVS1, PM2	.

						4			
.		ENSP00000345893.4:p.Gly5ArgfsTer26	M1MPVX	frameshift_variant & start_lost	0.00001343	0.00029768	P	PVS1, PM2, PP3	.

Foot note: Abbreviations: AF, allele frequency; ACMG-AMP, American College of Medical Genetics and Genomics and the Association for Molecular Pathology (BP, benign supporting; PM, pathogenic moderate; PP, pathogenic supporting); G&H, Genes & Health; LP, likely pathogenic; P, pathogenic. ^Allele frequency specific to East London Genes & Health cohort.

Table references

1. *EASL Clinical Practice Guidelines: management of cholestatic liver diseases.* J Hepatol, 2009. **51**(2): p. 237-67.
2. Sticova, E. and M. Jirsa, *ABCB4 disease: Many faces of one gene deficiency.* Annals of Hepatology, 2020. **19**(2): p. 126-133.
3. Huynh, M.T., T.T. Nguyen, S. Grison, et al., *Clinical characteristics and genetic profiles of young and adult patients with cholestatic liver disease.* Rev Esp Enferm Dig, 2019. **111**(10): p. 775-788.
4. Stättermayer, A.F., E. Halilbasic, F. Wrba, et al., *Variants in ABCB4 (MDR3) across the spectrum of cholestatic liver diseases in adults.* J Hepatol, 2020. **73**(3): p. 651-663.
5. Davit-Spraul, A., E. Gonzales, and E. Jacquemin, *NR1H4 analysis in patients with progressive familial intrahepatic cholestasis, drug-induced cholestasis or intrahepatic cholestasis of pregnancy unrelated to ATP8B1, ABCB11 and ABCB4 mutations.* Clin Res Hepatol Gastroenterol, 2012. **36**(6): p. 569-73.
6. Sambrotta, M., S. Strautnieks, E. Papouli, et al., *Mutations in TJP2 cause progressive cholestatic liver disease.* Nat Genet, 2014. **46**(4): p. 326-8.
7. de Vree, J.M., E. Jacquemin, E. Sturm, et al., *Mutations in the MDR3 gene cause progressive familial intrahepatic cholestasis.* Proc Natl Acad Sci U S A, 1998. **95**(1): p. 282-7.
8. Jacquemin, E., D. Cresteil, S. Manouvrier, et al., *Heterozygous non-sense mutation of the MDR3 gene in familial intrahepatic cholestasis of pregnancy.* The Lancet, 1999. **353**(9148): p. 210-211.
9. Dixon, P.H., N. Weerasekera, K.J. Linton, et al., *Heterozygous MDR3 missense mutation associated with intrahepatic cholestasis of pregnancy: evidence for a defect in protein trafficking.* Human Molecular Genetics, 2000. **9**(8): p. 1209-1217.
10. Lucena, J.-F., J.I. Herrero, J. Quiroga, et al., *A multidrug resistance 3 gene mutation causing cholelithiasis, cholestasis of pregnancy, and adulthood biliary cirrhosis.* Gastroenterology, 2003. **124**(4): p. 1037-1042.
11. Müllenbach, R., K.J. Linton, S. Wiltshire, et al., *ABCB4 gene sequence variation in women with intrahepatic cholestasis of pregnancy.* Journal of Medical Genetics, 2003. **40**(5): p. e70-e70.
12. Pauli-Magnus, C., T. Lang, Y. Meier, et al., *Sequence analysis of bile salt export pump (ABCB11) and multidrug resistance p-glycoprotein 3 (ABCB4, MDR3) in patients with intrahepatic cholestasis of pregnancy.* Pharmacogenetics and Genomics, 2004. **14**(2).
13. Floreani, A., I. Carderi, D. Paternoster, et al., *Intrahepatic cholestasis of pregnancy: three novel MDR3 gene mutations.* Alimentary Pharmacology & Therapeutics, 2006. **23**(11): p. 1649-1653.
14. Williamson, C. and V. Geenes, *Intrahepatic cholestasis of pregnancy.* Obstet Gynecol, 2014. **124**(1): p. 120-133.
15. Glantz, A., H.U. Marschall, and L.A. Mattsson, *Intrahepatic cholestasis of pregnancy: Relationships between bile acid levels and fetal complication rates.* Hepatology, 2004. **40**(2): p. 467-74.
16. Geenes, V., L.C. Chappell, P.T. Seed, et al., *Association of severe intrahepatic cholestasis of pregnancy with adverse pregnancy outcomes: a prospective population-based case-control study.* Hepatology, 2014. **59**(4): p. 1482-91.
17. Geenes, V. and C. Williamson, *Intrahepatic cholestasis of pregnancy.* World journal of gastroenterology, 2009. **15**(17): p. 2049-2066.
18. Alazawi, W., R. Mathur, K. Abeysekera, et al., *Ethnicity and the diagnosis gap in liver disease: a population-based study.* The British journal of general practice : the journal of the Royal College of General Practitioners, 2014. **64**(628): p. e694-e702.

19. Szanto, K.B., J. Li, P. Cordero, et al., *Ethnic differences and heterogeneity in genetic and metabolic makeup contributing to nonalcoholic fatty liver disease.* Diabetes, metabolic syndrome and obesity : targets and therapy, 2019. **12**: p. 357-367.
20. Finer, S., H.C. Martin, A. Khan, et al., *Cohort Profile: East London Genes & Health (ELGH), a community-based population genomics and health study in British Bangladeshi and British Pakistani people.* International Journal of Epidemiology, 2019. **49**(1): p. 20-21i.
21. Narasimhan, V.M., K.A. Hunt, D. Mason, et al., *Health and population effects of rare gene knockouts in adult humans with related parents.* Science, 2016. **352**(6284): p. 474-477.
22. Adzhubei, I.A., S. Schmidt, L. Peshkin, et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.
23. Sim, N.-L., P. Kumar, J. Hu, et al., *SIFT web server: predicting effects of amino acid substitutions on proteins.* Nucleic Acids Research, 2012. **40**(W1): p. W452-W457.
24. Kircher, M., D.M. Witten, P. Jain, et al., *A general framework for estimating the relative pathogenicity of human genetic variants.* Nature Genetics, 2014. **46**(3): p. 310-315.
25. Ioannidis, N.M., J.H. Rothstein, V. Pejaver, et al., *REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants.* The American Journal of Human Genetics, 2016. **99**(4): p. 877-885.
26. Liu, X., C. Li, C. Mou, et al., *dbNSFP v4: a comprehensive database of transcript-specific functional predictions and annotations for human nonsynonymous and splice-site SNVs.* Genome Medicine, 2020. **12**(1): p. 103.
27. Jagadeesh, K.A., A.M. Wenger, M.J. Berger, et al., *M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity.* Nature Genetics, 2016. **48**(12): p. 1581-1586.
28. Richards, S., N. Aziz, S. Bale, 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. **17**(5): p. 405-24.
29. Zhan, X., Y. Hu, B. Li, et al., *RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data.* Bioinformatics, 2016. **32**(9): p. 1423-1426.
30. Stephenson, J.D., R.A. Laskowski, A. Nightingale, et al., *VarMap: a web tool for mapping genomic coordinates to protein sequence and structure and retrieving protein structural annotations.* Bioinformatics, 2019. **35**(22): p. 4854-4856.
31. Johns, S.J. *TOPO2. Transmembrame protein display software*. Available from: http://www.sacs.ucsf.edu/TOPO2/.
32. Rodrigues, C.H., D.E. Pires, and D.B. Ascher, *DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability.* Nucleic Acids Research, 2018. **46**(W1): p. W350-W355.
33. Parthiban, V., M.M. Gromiha, and D. Schomburg, *CUPSAT: prediction of protein stability upon point mutations.* Nucleic Acids Research, 2006. **34**(suppl_2): p. W239-W242.
34. Ancien, F., F. Pucci, M. Godfroid, et al., *Prediction and interpretation of deleterious coding variants in terms of protein structural stability.* Scientific Reports, 2018. **8**(1): p. 4480.
35. Nosol, K., R. Bang-Sørensen, R.N. Irobalieva, et al., *Structures of ABCB4 provide insight into phosphatidylcholine translocation.* Proc Natl Acad Sci U S A, 2021. **118**(33).
36. Wang, L., W.T. Hou, L. Chen, et al., *Cryo-EM structure of human bile salts exporter ABCB11.* Cell Res, 2020. **30**(7): p. 623-625.
37. He, Y., J. Xu, X. Wu, et al., *Structures of a P4-ATPase lipid flippase in lipid bilayers.* Protein Cell, 2020. **11**(6): p. 458-463.
38. Downes, M., M.A. Verdecia, A.J. Roecker, et al., *A chemical, genetic, and structural analysis of the nuclear bile acid receptor FXR.* Mol Cell, 2003. **11**(4): p. 1079-92.

39. Dixon, P.H., M. Sambrotta, J. Chambers, et al., *An expanded role for heterozygous mutations of ABCB4, ABCB11, ATP8B1, ABCC2 and TJP2 in intrahepatic cholestasis of pregnancy.* Scientific Reports, 2017. **7**(1): p. 11823.
40. Liu, X., H. Lai, S. Xin, et al., *Whole-exome sequencing identifies novel mutations in ABC transporter genes associated with intrahepatic cholestasis of pregnancy disease: a case-control study.* BMC pregnancy and childbirth, 2021. **21**(1): p. 110-110.
41. Avena, A., S. Puggelli, M. Morris, et al., *ABCB4 variants in adult patients with cholestatic disease are frequent and underdiagnosed.* Digestive and Liver Disease, 2021. **53**(3): p. 329-344.
42. Wright, C.F., B. West, M. Tuke, et al., *Assessing the Pathogenicity, Penetrance, and Expressivity of Putative Disease-Causing Variants in a Population Setting.* Am J Hum Genet, 2019. **104**(2): p. 275-286.
43. Dixon, P.H., S.W. van Mil, J. Chambers, et al., *Contribution of variant alleles of ABCB11 to susceptibility to intrahepatic cholestasis of pregnancy.* Gut, 2009. **58**(4): p. 537-44.
44. Ali, I., S. Khalid, B. Stieger, et al., *Effect of a Common Genetic Variant (p.V444A) in the Bile Salt Export Pump on the Inhibition of Bile Acid Transport by Cholestatic Medications.* Molecular pharmaceutics, 2019. **16**(3): p. 1406-1411.
45. Wikström Shemer, E., H.U. Marschall, J.F. Ludvigsson, et al., *Intrahepatic cholestasis of pregnancy and associated adverse pregnancy and fetal outcomes: a 12-year population-based cohort study.* Bjog, 2013. **120**(6): p. 717-23.
46. Marschall, H.-U., E. Wikström Shemer, J.F. Ludvigsson, et al., *Intrahepatic cholestasis of pregnancy and associated hepatobiliary disease: A population-based cohort study.* Hepatology, 2013. **58**(4): p. 1385-1391.
47. Yeap, S.P., H. Harley, R. Thompson, et al., *Biliary transporter gene mutations in severe intrahepatic cholestasis of pregnancy: Diagnostic and management implications.* J Gastroenterol Hepatol, 2019. **34**(2): p. 425-435.
48. Delaunay, J.-L., A.-M. Durand-Schneider, C. Dossier, et al., *A functional classification of ABCB4 variations causing progressive familial intrahepatic cholestasis type 3.* Hepatology, 2016. **63**(5): p. 1620-1631.
49. Aydın, G.A., G. Özgen, and O. Görükmez, *The role of genetic mutations in intrahepatic cholestasis of pregnancy.* Taiwanese Journal of Obstetrics and Gynecology, 2020. **59**(5): p. 706-710.
50. Poupon, R., O. Rosmorduc, P.Y. Boëlle, et al., *Genotype-phenotype relationships in the low-phospholipid-associated cholelithiasis syndrome: A study of 156 consecutive patients.* Hepatology, 2013. **58**(3): p. 1105-1110.
51. Anzivino, C., M.R. Odoardi, E. Meschiari, et al., *ABCB4 and ABCB11 mutations in intrahepatic cholestasis of pregnancy in an Italian population.* Dig Liver Dis, 2013. **45**(3): p. 226-32.
52. Condat, B., D. Zanditenas, V. Barbu, et al., *Prevalence of low phospholipid-associated cholelithiasis in young female patients.* Digestive and Liver Disease, 2013. **45**(11): p. 915-919.
53. Stalke, A., B. Skawran, B. Auber, et al., *Diagnosis of monogenic liver diseases in childhood by next-generation sequencing.* Clinical Genetics, 2018. **93**(3): p. 665-670.
54. Gudbjartsson, D.F., H. Helgason, S.A. Gudjonsson, et al., *Large-scale whole-genome sequencing of the Icelandic population.* Nature Genetics, 2015. **47**(5): p. 435-444.
55. Degiorgio, D., A. Crosignani, C. Colombo, et al., *ABCB4 mutations in adult patients with cholestatic liver disease: impact and phenotypic expression.* Journal of Gastroenterology, 2016. **51**(3): p. 271-280.
56. Jirsa, M., J. Bronský, L. Dvořáková, et al., *ABCB4 mutations underlie hormonal cholestasis but not pediatric idiopathic gallstones.* World journal of gastroenterology, 2014. **20**(19): p. 5867-5874.

57. Bronský, J., M. Jirsa, J. Nevoral, et al., *Role of common canalicular transporter gene variations in aetiology of idiopathic gallstones in childhood.* Folia Biol (Praha), 2010. **56**(1): p. 9-13.
58. Schatz, S.B., C. Jüngst, V. Keitel-Anselmo, et al., *Phenotypic spectrum and diagnostic pitfalls of ABCB4 deficiency depending on age of onset.* Hepatology Communications, 2018. **2**(5): p. 504-514.
59. Gordo-Gilart, R., S. Andueza, L. Hierro, et al., *Functional analysis of ABCB4 mutations relates clinical outcomes of progressive familial intrahepatic cholestasis type 3 to the degree of MDR3 floppase activity.* Gut, 2015. **64**(1): p. 147-155.
60. Krawczyk, M., M. Rau, F. Grünhage, et al., *The ABCB4 p.T175A variant as potential modulator of hepatic fibrosis in patients with chronic liver diseases: Looking beyond the cholestatic realm.* Hepatology, 2017. **66**(2): p. 666-667.
61. Nakken, K.E., K.J. Labori, O.K. Rødningen, et al., *ABCB4 sequence variations in young adults with cholesterol gallstone disease.* Liver International, 2009. **29**(5): p. 743-747.
62. Denk, G.U., H. Bikker, R.H. Lekanne dit Deprez, et al., *ABCB4 deficiency: A family saga of early onset cholelithiasis, sclerosing cholangitis and cirrhosis and a novel mutation in the ABCB4 gene.* Hepatology Research, 2010. **40**(9): p. 937-941.
63. Wendum, D., V. Barbu, O. Rosmorduc, et al., *Aspects of liver pathology in adult patients with MDR3/ABCB4 gene mutations.* Virchows Archiv, 2012. **460**(3): p. 291-298.
64. Tomaiuolo, R., D. Degiorgio, D.A. Coviello, et al., *An MBL2 haplotype and ABCB4 variants modulate the risk of liver disease in cystic fibrosis patients: A multicentre study.* Digestive and Liver Disease, 2009. **41**(11): p. 817-822.
65. Degiorgio, D., C. Colombo, M. Seia, et al., *Molecular characterization and structural implications of 25 new ABCB4 mutations in progressive familial intrahepatic cholestasis type 3 (PFIC3).* European Journal of Human Genetics, 2007. **15**(12): p. 1230-1238.
66. Ziol, M., V. Barbu, O. Rosmorduc, et al., *ABCB4 Heterozygous Gene Mutations Associated With Fibrosing Cholestatic Liver Disease in Adults.* Gastroenterology, 2008. **135**(1): p. 131-141.
67. Lang, C., Y. Meier, B. Stieger, et al., *Mutations and polymorphisms in the bile salt export pump and the multidrug resistance protein 3 associated with drug-induced liver injury.* Pharmacogenetics and Genomics, 2007. **17**(1).
68. Vitale, G., S. Gitto, F. Raimondi, et al., *Cryptogenic cholestasis in young and adults: ATP8B1, ABCB11, ABCB4, and TJP2 gene variants analysis by high-throughput sequencing.* Journal of Gastroenterology, 2018. **53**(8): p. 945-958.
69. Davit-Spraul, A., C. Oliveira, E. Gonzales, et al., *Liver transcript analysis reveals aberrant splicing due to silent and intronic variations in the ABCB11 gene.* Molecular Genetics and Metabolism, 2014. **113**(3): p. 225-229.
70. Ho, R.H., B.F. Leake, D.M. Kilkenny, et al., *Polymorphic variants in the human bile salt export pump (BSEP; ABCB11): functional characterization and interindividual variability.* Pharmacogenet Genomics, 2010. **20**(1): p. 45-57.
71. Meier, Y., C. Pauli-Magnus, U.M. Zanger, et al., *Interindividual variability of canalicular ATP-binding-cassette (ABC)-transporter expression in human liver.* Hepatology, 2006. **44**(1): p. 62-74.
72. Lang, T., M. Haberl, D. Jung, et al., *Genetic Variability, Haplotype Structures, and Ethnic Diversity of Hepatic Transporters MDR3 (ABCB4) and Bile Salt Export Pump (ABCB11).* Drug Metabolism and Disposition, 2006. **34**(9): p. 1582-1599.
73. Davit-Spraul, A., M. Fabre, S. Branchereau, et al., *ATP8B1 and ABCB11 analysis in 62 children with normal gamma-glutamyl transferase progressive familial intrahepatic cholestasis (PFIC): Phenotypic differences between PFIC1 and PFIC2 and natural history.* Hepatology, 2010. **51**(5): p. 1645-1655.

74. Pauli-Magnus, C., R. Kerb, K. Fattinger, et al., *BSEP and MDR3 haplotype structure in healthy Caucasians, primary biliary cirrhosis and primary sclerosing cholangitis.* Hepatology, 2004. **39**(3): p. 779-791.
75. Kim, S.R., Y. Saito, M. Itoda, et al., *Genetic variations of the ABC transporter gene ABCB11 encoding the human bile salt export pump (BSEP) in a Japanese population.* Drug Metab Pharmacokinet, 2009. **24**(3): p. 277-81.
76. Dröge, C., M. Bonus, U. Baumann, et al., *Sequencing of FIC1, BSEP and MDR3 in a large cohort of patients with cholestasis revealed a high number of different genetic variants.* Journal of Hepatology, 2017. **67**(6): p. 1253-1264.
77. Mitra, S., A. Das, B. Thapa, et al., *Phenotype-Genotype Correlation of North Indian Progressive Familial Intrahepatic Cholestasis type2 Children Shows p.Val444Ala and p.Asn591Ser Variants and Retained BSEP Expression.* Fetal Pediatr Pathol, 2020. **39**(2): p. 107-123.
78. Lam, P., C.L. Pearson, C.J. Soroka, et al., *Levels of plasma membrane expression in progressive and benign mutations of the bile salt export pump (Bsep/Abcb11) correlate with severity of cholestatic diseases.* American Journal of Physiology-Cell Physiology, 2007. **293**(5): p. C1709-C1716.
79. Byrne, J.A., S.S. Strautnieks, G. Ihrke, et al., *Missense mutations and single nucleotide polymorphisms in ABCB11 impair bile salt export pump processing and function or disrupt pre-messenger RNA splicing.* Hepatology, 2009. **49**(2): p. 553-567.
80. Chong, C.P., P.B. Mills, P. McClean, et al., *Bile acid-CoA ligase deficiency--a new inborn error of bile acid metabolism.* J Inherit Metab Dis, 2012. **35**(3): p. 521-30.
81. Sharma, A., U. Poddar, S. Agnihotry, et al., *Spectrum of genomic variations in Indian patients with progressive familial intrahepatic cholestasis.* BMC Gastroenterology, 2018. **18**(1): p. 107.
82. Chen, X.-Q., L.-L. Wang, Q.-W. Shan, et al., *Multidrug resistance protein 3 R652G may reduce susceptibility to idiopathic infant cholestasis.* World journal of gastroenterology, 2009. **15**(46): p. 5855-5858.
83. Varma, S., N. Revencu, X. Stephenne, et al., *Retargeting of bile salt export pump and favorable outcome in children with progressive familial intrahepatic cholestasis type 2.* Hepatology, 2015. **62**(1): p. 198-206.
84. Roustit, M., X. Fonrose, D. Montani, et al., *CYP2C9, SLCO1B1, SLCO1B3, and ABCB11 Polymorphisms in Patients With Bosentan-Induced Liver Toxicity.* Clinical Pharmacology & Therapeutics, 2014. **95**(6): p. 583-585.
85. Liu, T., R.-X. Wang, J. Han, et al., *Changes in plasma bile acid profiles after partial internal biliary diversion in PFIC2 patients.* Annals of Translational Medicine, 2020. **8**(5): p. 185.
86. Abuduxikuer, K., J.Y. Feng, Y. Lu, et al., *Novel methionyl-tRNA synthetase gene variants/phenotypes in interstitial lung and liver disease: A case report and review of literature.* World J Gastroenterol, 2018. **24**(36): p. 4208-4216.
87. van der Woerd, W.L., D.Y. van Haaften-Visser, S.F.J. van de Graaf, et al., *Mutational Analysis of ATP8B1 in Patients with Chronic Pancreatitis.* PLOS ONE, 2013. **8**(11): p. e80553.
88. Hasegawa, Y., H. Hayashi, S. Naoi, et al., *Intractable itch relieved by 4-phenylbutyrate therapy in patients with progressive familial intrahepatic cholestasis type 1.* Orphanet Journal of Rare Diseases, 2014. **9**(1): p. 89.
89. Hayashi, H., S. Naoi, T. Togawa, et al., *Assessment of ATP8B1 Deficiency in Pediatric Patients With Cholestasis Using Peripheral Blood Monocyte-Derived Macrophages.* EBioMedicine, 2018. **27**: p. 187-199.
90. Nagasaka, H., T. Yorifuji, K. Hirano, et al., *Effects of bezafibrate on dyslipidemia with cholestasis in children with familial intrahepatic cholestasis-1 deficiency manifesting progressive familial intrahepatic cholestasis.* Metabolism, 2009. **58**(1): p. 48-54.

91. Qiu, Y.-L., J.-Y. Gong, J.-Y. Feng, et al., *Defects in myosin VB are associated with a spectrum of previously undiagnosed low γ-glutamyltransferase cholestasis.* Hepatology, 2017. **65**(5): p. 1655-1669.
92. Klomp, L.W.J., J.C. Vargas, S.W.C. van Mil, et al., *Characterization of mutations in ATP8B1 associated with hereditary cholestasis.* Hepatology, 2004. **40**(1): p. 27-38.
93. Müllenbach, R., A. Bennett, N. Tetlow, et al., *ATP8B1 mutations in British cases with intrahepatic cholestasis of pregnancy.* Gut, 2005. **54**(6): p. 829.
94. Folmer, D.E., V.A. van der Mark, K.S. Ho-Mok, et al., *Differential effects of progressive familial intrahepatic cholestasis type 1 and benign recurrent intrahepatic cholestasis type 1 mutations on canalicular localization of ATP8B1.* Hepatology, 2009. **50**(5): p. 1597-605.
95. Stolz, A., V. Navarro, P.H. Hayashi, et al., *Severe and protracted cholestasis in 44 young men taking bodybuilding supplements: assessment of genetic, clinical and chemical risk factors.* Alimentary Pharmacology & Therapeutics, 2019. **49**(9): p. 1195-1204.
96. Takatsu, H., G. Tanaka, K. Segawa, et al., *Phospholipid Flippase Activities and Substrate Specificities of Human Type IV P-type ATPases Localized to the Plasma Membrane *.* Journal of Biological Chemistry, 2014. **289**(48): p. 33543-33556.
97. Giovannoni, I., F. Callea, E. Bellacchio, et al., *Genetics and Molecular Modeling of New Mutations of Familial Intrahepatic Cholestasis in a Single Italian Center.* PLoS One, 2015. **10**(12): p. e0145021.
98. McKay, K., C. Bruce, J. Hartley, et al., *Mutation detection in cholestatic patients using microarray resequencing of ATP8B1 and ABCB11 [version 2; peer review: 2 approved, 1 approved with reservations].* F1000Research, 2013. **2**(32).
99. Van Mil, S.W., A. Milona, P.H. Dixon, et al., *Functional variants of the central bile acid sensor FXR identified in intrahepatic cholestasis of pregnancy.* Gastroenterology, 2007. **133**(2): p. 507-16.
100. Kim, M.-A., Y.-R. Kim, B. Sagong, et al., *Genetic Analysis of Genes Related to Tight Junction Function in the Korean Population with Non-Syndromic Hearing Loss.* PLOS ONE, 2014. **9**(4): p. e95646.
101. Li, J., D. Jiang, X. Xiao, et al., *Evaluation of 12 Myopia-Associated Genes in Chinese Patients With High Myopia.* Investigative Ophthalmology & Visual Science, 2015. **56**(2): p. 722-729.

a. Overall summary

ABCB4
68 non-synonymous variants
41 included in analysis

ABCB11
77 non-synonymous variants
48 included in analysis

ATP8B1
50 non-synonymous variants
31 included in analysis

NR1H4
22 non-synonymous variants
9 included in analysis

TJP2
83 non-synonymous variants
37 included in analysis

b. Pathogenicity

ABCB4
22 Likely pathogenic (10 novel)
13 unknown significance (8 novel)
6 benign (2 novel)

ABCB11
25 Likely pathogenic (7 novel)
18 unknown significance (10 novel)
4 benign (none novel)

ATP8B1
3 Likely pathogenic (1 novel)
23 unknown significance (11 novel)
5 benign (none novel)

NR1H4
0 Likely pathogenic (none novel)
8 unknown significance (4 novel)
1 benign (none novel)

TJP2
0 Likely pathogenic (none novel)
35 unknown significance (23 novel)
2 benign (none novel)

c. Phenotypes

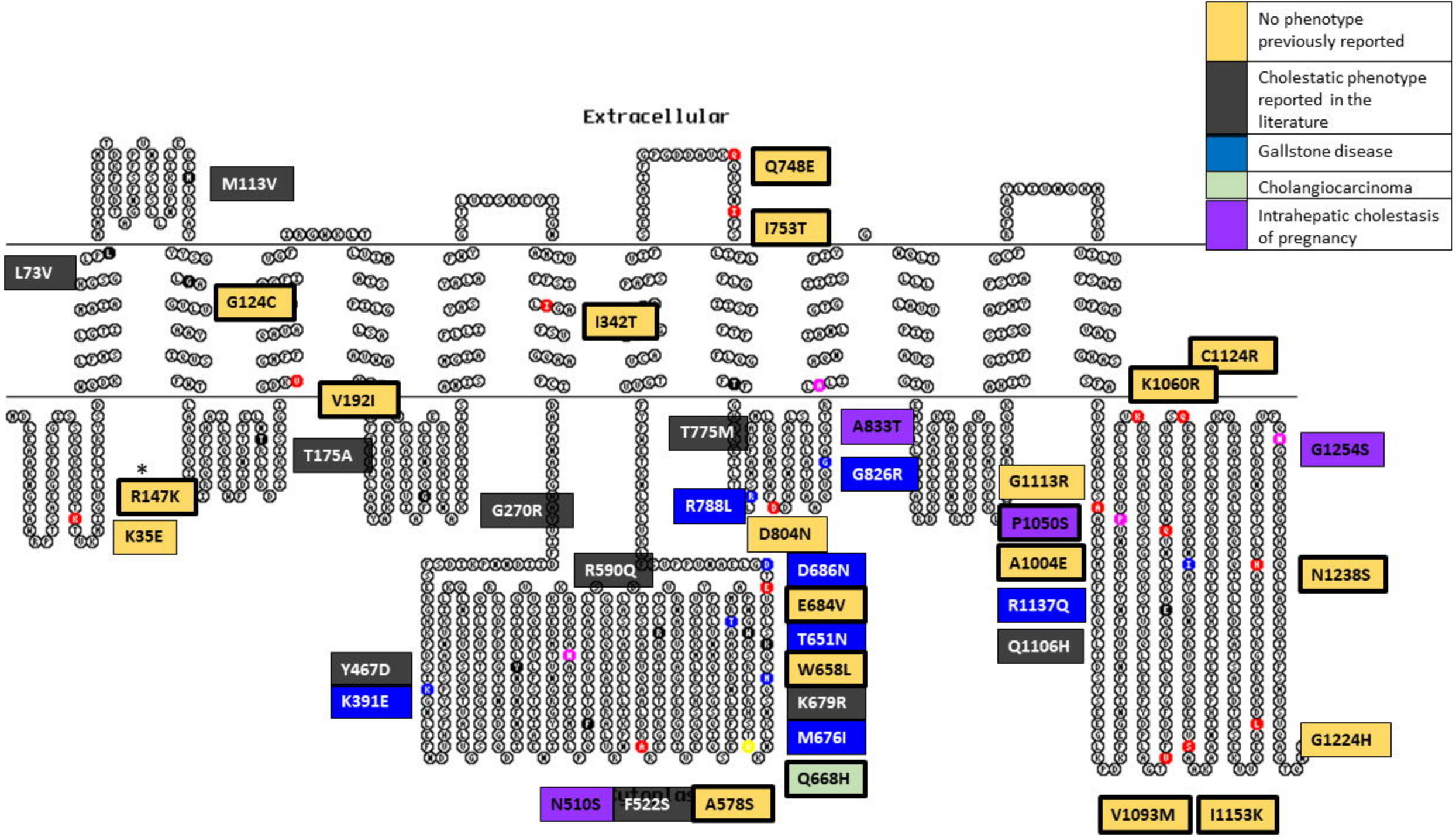
ABCB4
5 ICP (2 novel)
7 Gallstone disease (1 novel)
1 Cholangiocarcinoma (novel)

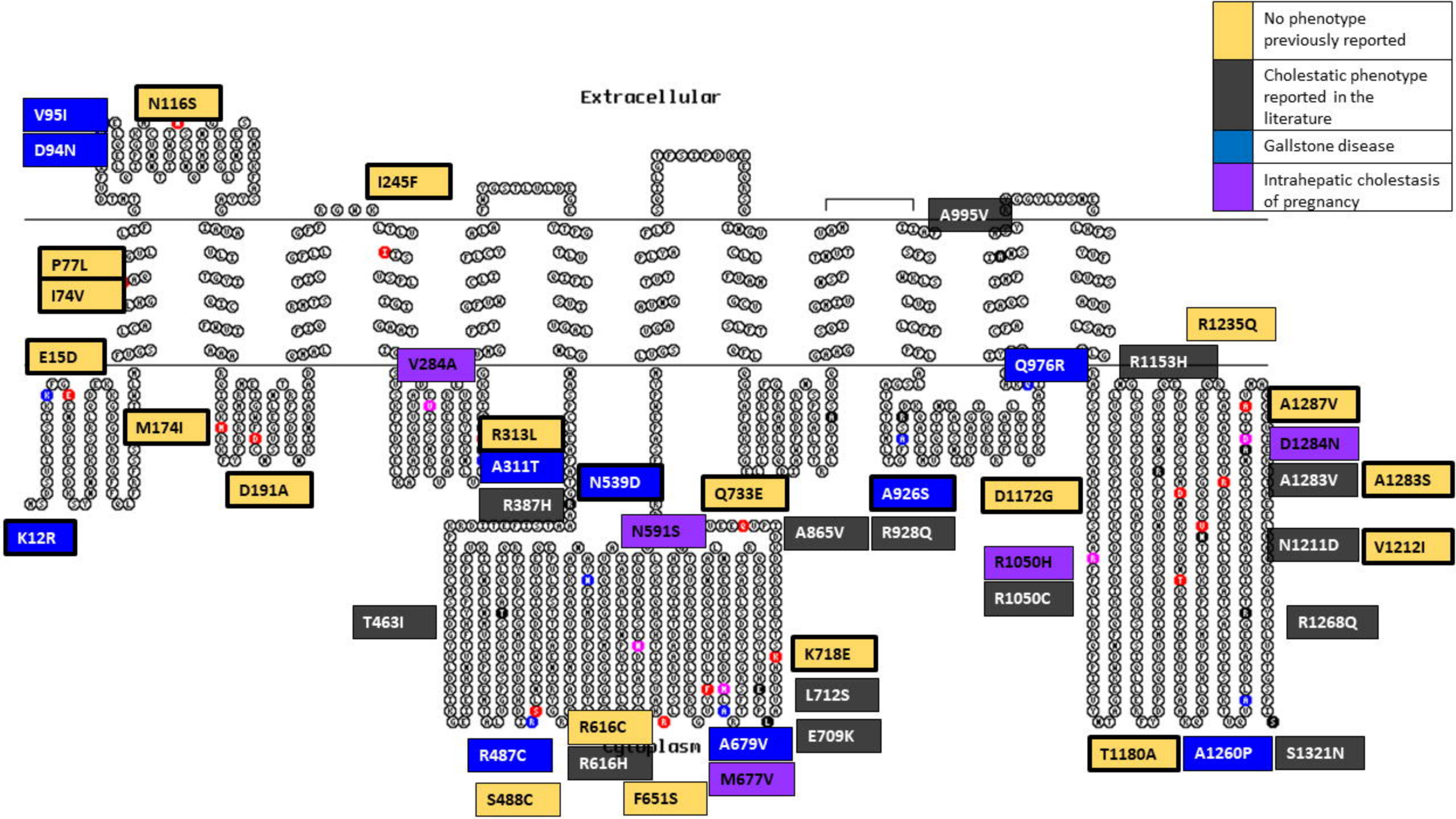
ABCB11
5 ICP
10 Gallstone disease (3 novel)

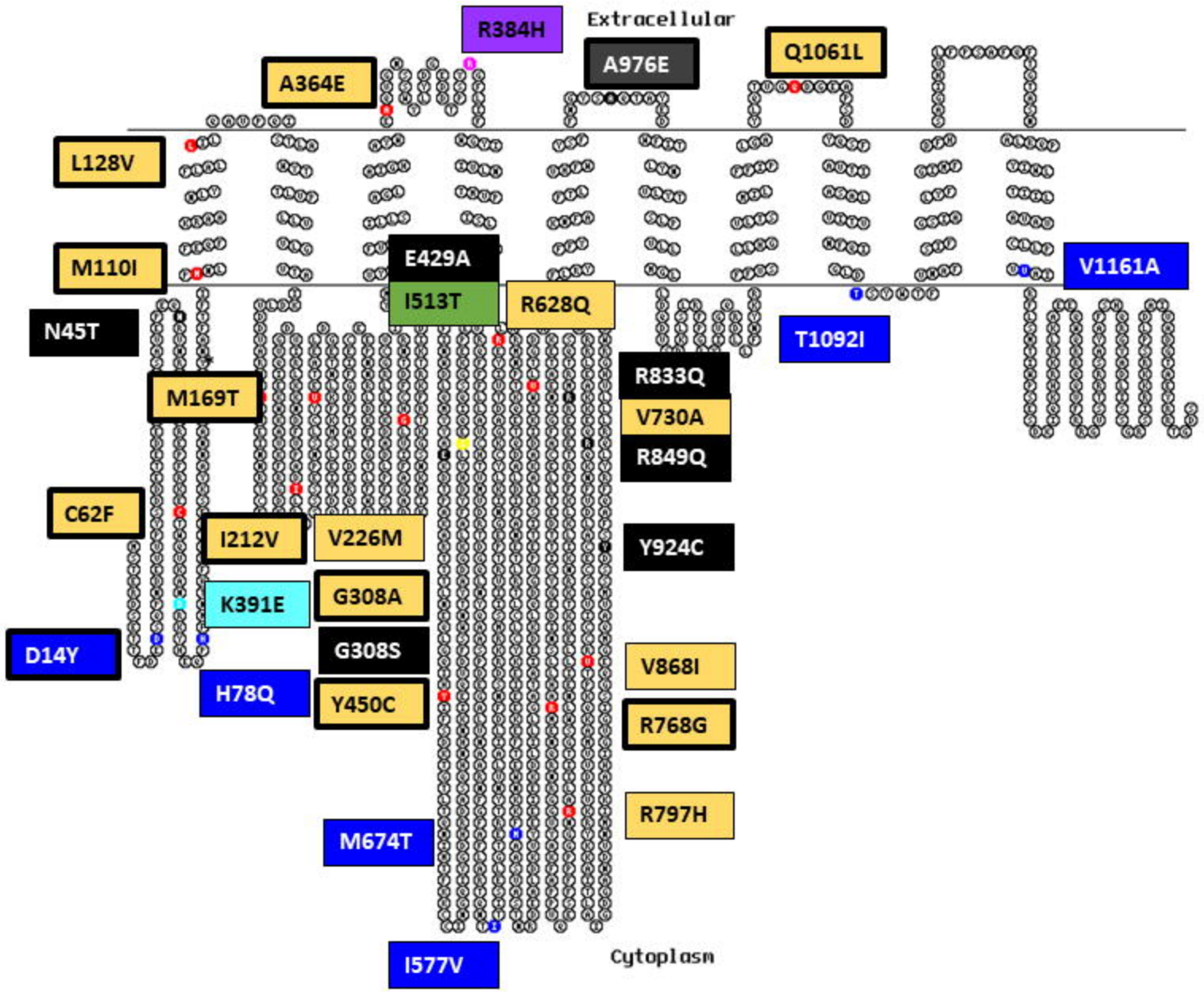
ATP8B1
1 ICP (* overlaps with GBD)
6 Gallstone disease (1 novel)
1 Cirrhosis, and 1 cirrhosis with secondary malignant neoplasm of liver and bile duct, gallstone disease

NR1H4
2 ICP

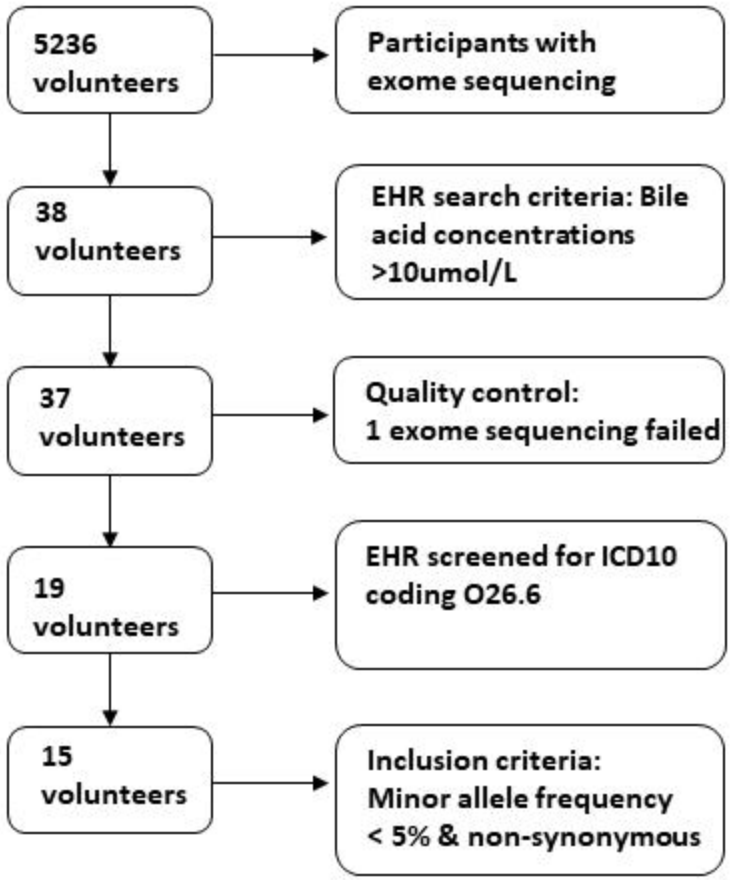
TJP2
3 ICP
8 Gallstone disease (3 novel)


No phenotype previously reported
Cholestatic phenotype reported in the literature
Gallstone disease
Cholangiocarcinoma
Intrahepatic cholestasis of pregnancy
Extracellular
M113V
Q748E
I753T
L73V
G124C
I342T
C1124R
K1060R
V192I
T175A
T775M
A833T
G826R
G1254S
R147K
K35E
G270R
R788L
D804N
G1113R
P1050S
A1004E
R1137Q
Q1106H
R590Q
D686N
E684V
T651N
W658L
K679R
M676I
Q668H
N1238S
Y467D
K391E
G1224H
N510S
F522S
A578S
V1093M
I1153K

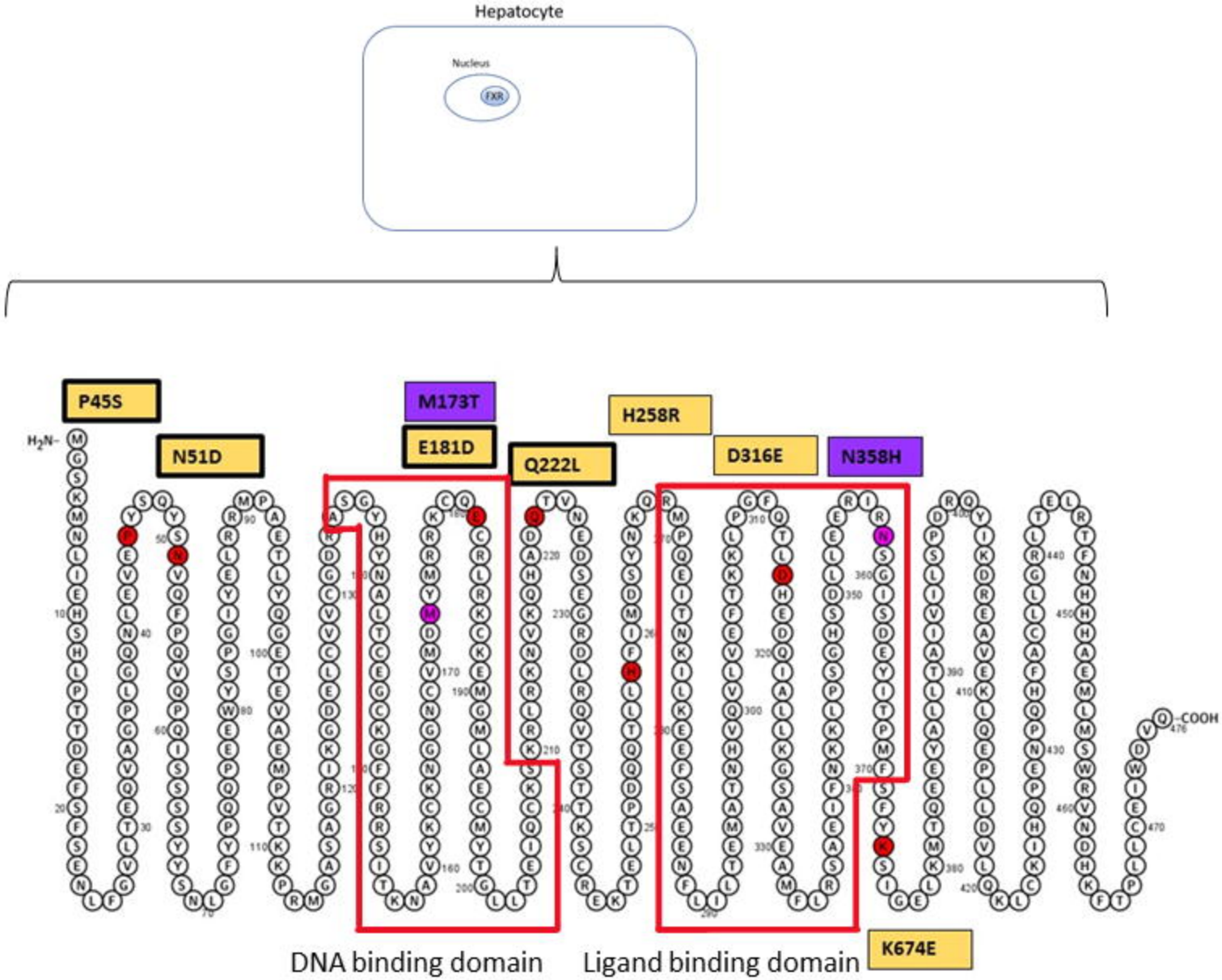

No phenotype previously reported
Cholestatic phenotype reported in the literature
Gallstone disease
Intrahepatic cholestasis of pregnancy
Extracellular
Cytoplasm
V95I
D94N
N116S
I245F
A995V
P77L
I74V
E15D
V284A
Q976R
R1235Q
R1153H
M174I
R313L
A311T
D191A
N539D
Q733E
A926S
D1172G
A1287V
D1284N
A1283V
A1283S
K12R
R387H
N591S
A865V
R928Q
N1211D
V1212I
R1050H
R1050C
T463I
R1268Q
K718E
L712S
E709K
R616C
R487C
R616H
A679V
M677V
S488C
F651S
T1180A
A1260P
S1321N



	No phenotype previously reported
	Cholestatic phenotype reported in the literature
	Gallstone disease
	Liver cirrhosis
	Liver cirrhosis, gallstone disease, secondary malignant neoplasm of liver and bile duct
	Intrahepatic cholestasis of pregnancy


5236 volunteers
Participants with exome sequencing
38 volunteers
EHR search criteria: Bile acid concentrations >10umol/L
37 volunteers
Quality control: 1 exome sequencing failed
19 volunteers
EHR screened for ICD10 coding O26.6
15 volunteers
Inclusion criteria: Minor allele frequency < 5% & non-synonymous

NR1H4 (FXR)



BOLD BORDER:
Unique to Genes & Health cohort

	Not previously reported
	Intrahepatic cholestasis of pregnancy

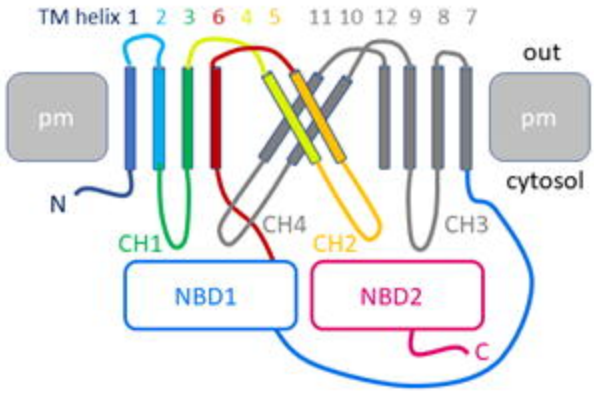
Inclusion criteria: 1. associated phenotype

2. and/or be predicted to be pathogenic based on all 7 *in-silico* tools

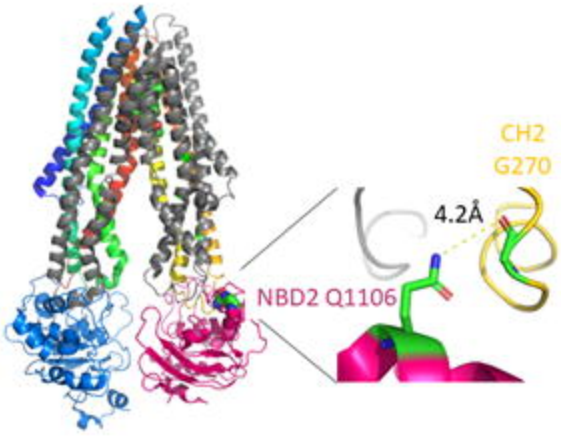
ABCB4	*ABCB11*	*ATP8B1*	*NR1H4*
41 variants	48 variants	31 variants	22 variants
↓ include variants with phenotype, &/or likely pathogenic	↓ include variants with phenotype, &/or likely pathogenic	↓ include variants with phenotype, &/or likely pathogenic	↓ include variants with phenotype, &/or likely pathogenic
27 variants	30 variants	8 variants	2 variants
↓ Unresolved 3D structure	↓ Unresolved 3D structure	↓ not conserved in PDB 6LCP	↓ Unresolved 3D structure
22 variants	27 variants	6 variants	1 variant
↓ structural damage	↓ structural damage	↓ significant change / structural damage	↓ significant change / structural damage
7 variants	6 variants	2 variants	No variant
↓ significant change / structural damage	↓ significant change / structural damage		
4 variants	4 variants		
PDB 7NIV	PDB 6LR0	PDB 6LCP	PDB 1OSH

* *TJP2* structure not resolved

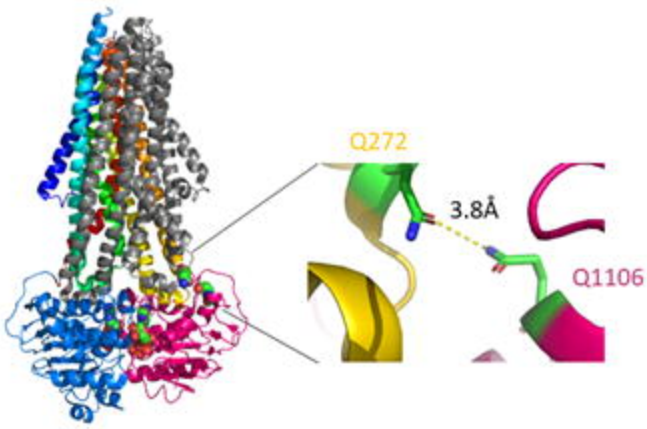
a



b

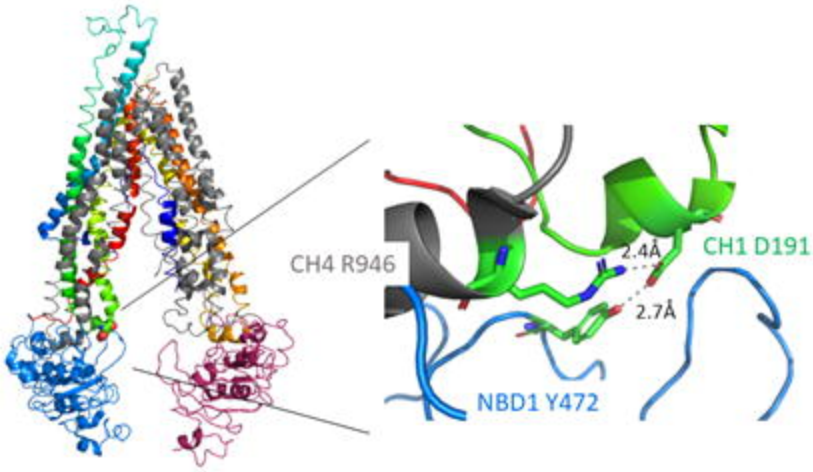


ABCB4 with PC bound
Pdb: 7NIV



ABCB4 with ATP bound
Pdb: 6S7P

c



ABCB11
Pdb: 6LR0