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. EGA Institute of Women's Health, Population Health Sciences, UCL, 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.

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: National Institute of Health and Care Research (NIHR). 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

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 South Asian population discovered new genetic mutations.

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

Novel variants that contribute to ICP were identified.

Abstract (250 words)

Objectives: This study aimed to assess 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 high rates of parental relatedness and 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. Variants associated with a phenotype or predicted to be likely pathogenic (LP) underwent protein structure and modelling analysis.

Results: Out of 314 non-synonymous variants, 180 fulfilled the inclusion criteria and were heterozygous. 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 associated with gallstone disease (n=31), intrahepatic cholestasis of pregnancy (ICP, n=16), cholangiocarcinoma and cirrhosis (n=2). Fourteen heterozygous novel LoF variants were identified: 7 frameshift, 5 introduction of premature stop codon and 2 splice acceptor variants. A total of 73 variants underwent protein modelling. The study identified 10 variants that resulted in predicted significant structural damage at a protein level, of which 4 were unique to this cohort.

Conclusions: This study identified novel LP and pathogenic variants. We provide further insight into rare variants associated with ICP, cholangiocarcinoma, and gallstone disease. This study highlights the importance of studying 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* (adenosine triphosphate-binding cassette (ATP), subfamily B, member 4) gene coding for the multidrug resistance 3 (MDR3) protein (a canalicular phospholipid translocator), and the *ABCB11* (ATP-binding cassette, subfamily B, member 11) 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]. Several other canalicular transporters or their regulators are implicated in the pathogenesis of cholestatic liver disease e.g., *ATP8B1* (ATPase phospholipid transporting 8B1) [3], *NR1H4* (FXR, a principal bile acid sensor) gene[5], and *TJP2* (tight junction protein 2)[6]. Homozygous mutations of these genes are implicated in a small number of cases of severe familial cholestasis. However, the evidence base for a role of heterozygous mutations in milder forms of liver disease is very limited. *ABCB4* and *ABCB11* are involved in 20% of patients with severe intrahepatic cholestasis of pregnancy (ICP). However, also considerable numbers of heterozygous *ABCB4* variants have been reported in women with ICP[7-13]. ICP is the commonest gestational liver disease [14] and may be complicated by preterm birth, fetal asphyxia, meconium-stained amniotic fluid and stillbirth, in association with maternal serum bile acid concentrations ≥40µmol[15, 16]. In white 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].

Large bioresources with genome-wide genetic data and linked electronic healthcare records are powerful tools for exploring genetic disease. UK efforts such as the Genomics England initiative,

NIHR Bioresource[20], and UK Biobank[21] offer considerable resources for discovery science. However, the majority of participants in these initiatives are of White European ancestry.

In contrast, the Genes and Health study (Genes & Health) is a long-term population-based cohort that assesses the health and disease in British Bangladeshi and British Pakistani people. Genes & Health is distinguished by its genomics (whole exome sequencing (WES)) and its data linkage to electronic health records (EHR)[22]. This study aimed to address the paucity of non-White ethnic groups in genomic research[22] in the field of cholestatic liver disease. We sought to investigate rare variation in a unique British Bangladeshi and Pakistani cohort around 5 candidate loci (*ABCB4*, *ABCB11*, *ATP8B1*, *NR1H4*, and *TJP2*) previously described and implicated in cholestatic liver disease.

Methods

Study population

Genes & Health has previously been described by Finer et al.[22]. 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).

A total of 5236 exome sequencing samples from 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). In a secondary analysis, a phenotype to genotype analysis was used to validate these findings specifically using ICP as an example. 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 additionally to verify a diagnosis of ICP.

Exome sequencing & bioinformatic pipeline

Low/mid exome sequencing was performed using Agilent in solution capture and Illumina short read sequencing as previously described in detail by Narasimhan et al.[23]. 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%.

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. Non-exonic or synonymous variants were excluded from further analysis. Variants were filtered and annotated if they met 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[24], SIFT[25], CADD[26], Revel[27], MetaLR[28], MetaSVM[28] and M_CAP[29]. Open-source databases (Leiden Open Variation Database – LoVD, and ClinVar) including a commercial database (Mastermind) were used to 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[30].

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 used in this study (see Supplementary Figure 2). For the assessment of the variants on the protein sequence the web tool VarMap was used[31]. The server allows to map genomic coordinates to protein sequence and structure whilst also retrieving protein structural annotations. The 2D representations were

designed using the open source tool protter[32], and TOPO2[33]. Variants that were LP or associated with a phenotype underwent further analysis. Software tools used to assess potential mutations effects were used: (1) Dynamut[34] (2) Missense 3D[35] (3) CUPSAT[36] (4) SNPMuSiC[37]. Dynamut analyses the effects amino acid substitutions on protein stability and flexibility as well as interaction amongst amino acid residues. The stability is calculated by structure-based methods including SDM, DUET, mCSM algorithms (ouput in kcal/mol) and flexibility is measured by entropy energy (output in kcal/mol/K)[34]. CUPSAT uses the environment of the mutation to predict changes in stability[38]. Missense3D was used in addition to assess a variety of structural features that may result in damage to the protein structure[35]. SNPMuSiC uses a knowledge based system that assess whether stabilising or destabilising mutations are disease-causing[37]. 3D structural representations were generated using PyMOL software. The recent release of *ABCB4* with phosphatidylcholine was used for modelling (PDB ID: 7NIV)[38], *ABCB11* open structure (PDB ID: 6LR0)[39], to model *ATP8B1* using a yeast - *Chaetomium thermophilum* (PDB ID: 6LCP)[40] after aligning to the human sequence, and *NR1H4* (PDB ID: 1OSH)[41]. There is no 3D protein structure available at present for *TJP2*.

Results

Genotype to phenotype analysis

Screening of the 5 candidate genes identified 300 variants. After meeting inclusion criteria 166 variants were included in the analysis (Figure 1). Most variants identified were heterozygous unless otherwise specified. Further variant interpretation including clinical information, literature search, scoring details of individual *in-silico* predictors, the impact of the coding substitution on disease propensity, and conservation information for all variants included in this study are presented in Supplementary Table 1.

Phenotype to genotype analysis

We were able to validate variants of interest in 15 volunteers with ICP, most of whom had documented raised BA concentrations (Table 2.). Further, variants discovered in 10 patients with raised BA but an uncertain or no diagnosis of ICP based on their EHR are described in Supplementary Table 2. This secondary analysis missed to identify 2 volunteers from the primary analysis with a diagnosis of ICP as there were no bile acid concentrations available for them.

ABCB4 variants

There were a total of 68 *ABCB4* variants identified (Figures 1 and 2). 41 variants fulfilled the inclusion criteria. Specifically, we identified variants with the following cholestatic liver disease phenotypes within Genes & Health: ICP (n=5), 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 loss of function (LoF) variants (n=4): three frameshift (one associated to an ICP phenotype) and one introduction of a premature stop codon (no associated phenotype) (Table 4).

ABCB11 variants

There were a total of 77 *ABCB11* variants identified of which 48 were included in the analysis (Figures 1 and 3). The cholestatic liver disease phenotypes identified in Genes & Health were: ICP (n=5), and gallstone disease (n=10) (Table 3). Some variants had a 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 were not associated with a known disease phenotype (Table 4).

ATP8B1 variants

We identified a total of 50 *ATP8B1* variants and 31 were included in the final analysis (Figures 1 and 4). There were a total of 7 variants associated with gallstone disease; three variants appeared to be in linkage disequilibrium (LD) noted in three volunteers: M674T, I577V, and H78Q (see LD calculation in Supplementary materials). The R384H variant was associated with a gallstone disease and with an ICP phenotype (in separate individuals). A further variant was associated with hepatitis-induced liver cirrhosis, and a final variant was associated with a liver cirrhosis, secondary malignant neoplasm of liver and bile duct, and gallstone disease phenotype (Table 3). In addition, previously reported cholestatic liver disease phenotype (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-gained variant. The latter variant was associated to 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 3). We only identified an ICP phenotype (n=2) (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 a total of 83 *TJP2* variants identified of which 37 were analysed (Figure 1). People with *TJP2* variants had ICP (n=3), gallstone disease (n=8), previously reported cholestatic liver disease phenotype (n=5), and 21 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

Results of the protein structure and molecular modelling software tools are presented in Supplementary Table 8. Specifically, the *ABCB4* variants R788L (gallstone phenotype), Q1106H, and D804N appeared to result in expansion of the cavity although the significance of this is not certain as ABC transporters are flexible proteins and very likely interact with their transport ligands by an induced fit mechanism. The variants G826R (gallstone phenotype), I1153K, G1113R, and G124C variant demonstrated in addition a significant change of a buried wild-type residue. Apart from G1113R the variants were novel and demonstrated a high disease propensity in highly conserved area of the protein. There were 6 *ABCB11* variants that resulted in structural changes on a protein level. R1235Q (not previously reported) and V284A (ICP phenotype) resulted in an expansion of the cavity volume, A1260P (gallstone phenotype) introduced a buried proline variant, T463I (previously reported) disrupted all side-chain H-bonds, A311T (gallstone phenotype) triggered a clash alert, D94N (associated to a gallstone phenotype) replaced a buried charge with an uncharged residue and causing a disruption of a salt bridge. Two other variants (G308S and G308S) in the *ATP8B1* gene resulted in a possible significant structural change as a buried glycine residue was replaced. Variants that resulted in possible significant structural damage and were unique to Genes & Health are illustrated in pymol (Supplementary Figure 4).

Discussion

This study demonstrates novel variants likely implicated in the aetiopathogenesis of cholestatic liver disease that appear to be associated to this unique British Bangladeshi and Pakistani cohort in comparison to other European genetic studies[42-44]. There have not been any other studies of this magnitude reported in the literature specifically analysing the burden of mutational variation in cholestatic liver disease in a large South Asian cohort. We discovered novel likely pathogenic variants that appear to be unique to this cohort and employed a genotype to phenotype analysis approach with the aim of validating these rare genetic variants using the ICP phenotype as an example. Thus,

this study will improve representation of this diverse population especially as prevalence of cholestatic liver disease is increased in the Genes & Health cohort for example 1.54% for ICP. This study demonstrates the importance of multi-ancestry genomic research und offers the potential of tailored treatment to this population.

In the Genes & Health cohort, we identified 50 variants that had a cholestatic liver disease phenotype reported in their linked EHR. Of those, 11 appeared to be unique to this British Bangladeshi and Pakistani population and a large number were predicted to be likely pathogenic based on *in-silico* prediction tools. For example, a novel *ABCB4* variant (G826R) associated with a gallstone phenotype was predicted to be LP and on prediction of the three-dimensional structure of the protein showing it could cause significant structural damage located in the intracellular domain 4 of the second transmembrane domain resulting in a change between buried and exposed state of the target variant residue. In addition, there were 35 variants that were previously reported in the literature with a cholestatic phenotype. However, a total of 81 variants had no phenotype previously reported in the literature. Out of those, 62 were novel (about 37% 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 pathological. Similarly, nearly half of the variants that likely caused significant structural impact at a protein level were novel (4 out of 10).

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. A multitude of diverse data sets is required including knowledge of the prevalence and mode of inheritance of the disease, awareness of the genetic heterogeneity, as well as the use of cutting-edge bioinformatics software tools[45]. The ACMG rare variant interpretation guideline[30] provides a standardised pathway for the analysis of variants. 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 mechanism are not clear yet [46, 47]. By employing computational protein modelling software tools, we were able to tease out a few variants that would likely have a significant impact on the conformation of the protein structure and were therefore likely 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. In brief, *in-silico* tools are not independent, and several use the same or similar sources of biological information[48]. Further, high probability of damage to the protein does not necessarily equate to high probability of pathogenicity[49]. Finally, exonic variants can have varied effects on different transcripts and the choice of which transcript to use has been widely discussed[50, 51]. 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 cancer, immune-mediated and cardiovascular diseases[52]. In addition, they have a higher gallstone-related morbidity as well as hepatitis C[53]. The risk of associated cholestatic liver disease is also increased in immediate family members[42].

Limitations

The use of electronic health records to determine phenotype is useful albeit dependent upon appropriate information having been coded and availability of linked data, which may be problematic particularly if historic. 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 for example the median age of volunteers in this study was around 45 years. In this study, we identified LP variants in volunteers with no phenotype of interest, or either no primary or secondary

care coding available or irrelevant disease coding for our phenotype of interest. It demonstrates the difficulty with interpreting variants when recalling the genotype first. Finally, empirical data on the effects of variants at the protein level would be helpful for further variant interpretation.

Conclusions

In this study we provide the first comprehensive evaluation of gene candidates associated to 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 who thus far remain understudied despite their distinct genetic background and increased risk to develop for example ICP in comparison to other populations. We were able to identify novel variants that have not been identified before and are likely implicated with disease. We demonstrated the ability to identify participants at risk both by a phenotype or genotype first approach. This demonstrates the importance to provide 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. Further, 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. 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.
16. 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.
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. Caulfield, M., J. Davies, and M. Dennys, *The 100,000 Genomes Project Protocol [Internet].* 2017.
21. Sudlow, C., J. Gallacher, N. Allen, et al., *UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age.* PLoS Med, 2015. **12**(3): p. e1001779.
22. 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.
23. 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.
24. 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.
25. 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.
26. 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.
27. 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.
28. Dong, C., P. Wei, X. Jian, et al., *Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies.* Human Molecular Genetics, 2014. **24**(8): p. 2125-2137.
29. 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.
30. 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.
31. 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.
32. Omasits, U., C.H. Ahrens, S. Müller, et al., *Protter: interactive protein feature visualization and integration with experimental proteomic data.* Bioinformatics, 2014. **30**(6): p. 884-6.
33. SJ, J. *TOPO2. Transmembrame protein display software*. Available from: http://www.sacs.ucsf.edu/TOPO2/.
34. 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.
35. Ittisoponpisan, S., S.A. Islam, T. Khanna, et al., *Can Predicted Protein 3D Structures Provide Reliable Insights into whether Missense Variants Are Disease Associated?* Journal of Molecular Biology, 2019. **431**(11): p. 2197-2212.
36. 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.
37. 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.
38. 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).

39. 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.
40. 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.
41. 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.
42. 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.
43. 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.
44. 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.
45. 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.
46. 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.
47. 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.
48. Popovic, D., A. Sifrim, J. Davis, et al., *Problems with the nested granularity of feature domains in bioinformatics: the eXtasy case.* BMC Bioinformatics, 2015. **16**(4): p. S2.
49. Eilbeck, K., A. Quinlan, and M. Yandell, *Settling the score: variant prioritization and Mendelian disease.* Nature reviews. Genetics, 2017. **18**(10): p. 599-612.
50. DiStefano, M.T., S.E. Hemphill, B.J. Cushman, et al., *Curating Clinically Relevant Transcripts for the Interpretation of Sequence Variants.* The Journal of Molecular Diagnostics, 2018. **20**(6): p. 789-801.
51. MacArthur, D.G., T.A. Manolio, D.P. Dimmock, et al., *Guidelines for investigating causality of sequence variants in human disease.* Nature, 2014. **508**(7497): p. 469-476.
52. Wikström Shemer, E.A., O. Stephansson, M. Thuresson, et al., *Intrahepatic cholestasis of pregnancy and cancer, immune-mediated and cardiovascular diseases: A population-based cohort study.* J Hepatol, 2015. **63**(2): p. 456-61.
53. 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.

Tables

Table 1.

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

Table 2.

All variants in patients with raised BA and diagnosis of ICP					
Patient ID	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	ATP8B1	A1152T	hom	Non-synonymous	
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

Table 3.

Gene	Clinical phenotype	Transcript	Protein change	dbSNP	gnomAD AF	ELGH AF	ACMG-AMP classification	ACMG-AMP criteria	Clinvar	References
ABCB4	Intrahepatic cholestasis of pregnancy	ENSP00000496956.1:p.Gly1254Ser	G1254S *	rs781315185	0.00003656	0.00028843	LP	PM1, PM2, PP2, PP3	.	Yes
		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	Yes
		ENSP00000496956.1:p.Thr175Ala	T175A	rs58238559	0.01155000	0.01315030	LB	PM1, PP2, PP3, BS1, BS2, BP6	benign/likely benign	Yes
	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	rs11568364	0.02364000	0.01005750	Benign	PP2, BA1,BS3, BP6, BP4	benign	Yes
		ENSP00000497931.1:p.Asn591Ser	N591S	rs11568367	0.01436000	0.12647200	Benign	PM1, PP2, BA1, BP6	benign	Yes

		ENSP00000497931.1:p.Val284Ala	V284A	rs200739891	0.00026040	0.00009558	LP	PM1, PM2, PM5, PP2, PP3, BP6	conflicting	Yes
	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.Asn539Asp	N539D *	.	.	0.00022604	VUS	PM1, PM2, PP2, BP4	.	.
		ENSP00000497931.1:p.Arg487Cys	R487C	rs770693935	0.00002043	0.00009549	LP	PM1, PM2, PP2, PP3	.	yes
		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	.	Yes
	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 +	rs35470719	0.00456300	0.00632063	Benign	PP2, BA1, BP4, BP6	Benign/Likely benign	yes
		ENSP00000497896.1:p.Ile577Val	I577V +	rs3745078	0.00467800	0.00628992	Benign	PP2, BA1, BP6	Benign	yes
		ENSP00000497896.1:p.His78Gln	H78Q +	rs3745079	0.00421800	0.00495751	Benign	PP2, BP4,BP6, BS1, BS2	Benign	yes
		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	rs34719006	0.00313900	0.00302678	VUS	PM2, PP2, PP3	Conflicting	yes
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	Yes
TJP2	Intrahepatic cholestasis of pregnancy	ENSP00000497787.1:p.Thr377Ala	T377A	rs766748789	0.00000406	0.00047765	VUS	PM2, BP4	.	.
		ENSP00000496791.1:p.Gln105Lys	Q105K	rs41305539	0.05150000	0.12031400	Benign	BA1, BP4, BP6,	benign	Yes
		ENSP00000497861.1:p.Arg21His	R21H	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	.

+	Linkage disequilibrium
*	South Asian specific variant
#	multiple phenotype
^	Structural impact
^^	Significant structural impact

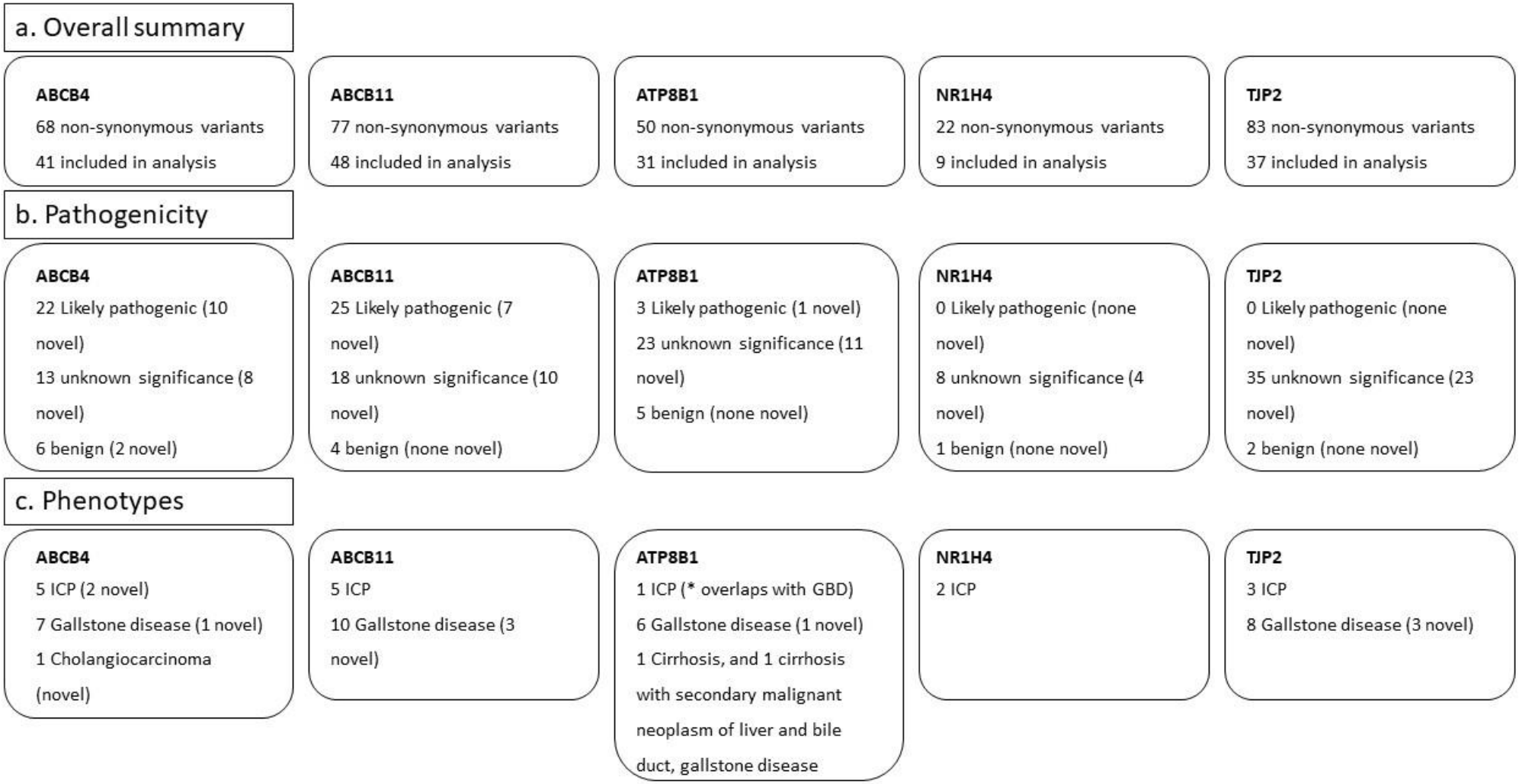
Table 4.

Info	Clinical phenotype	Transcript	Protein change	Info	gnomAD AF	ELGH AF	ACMG-AMP classification	ACMG-AMP criteria	Clinvar	References
ICP	ABCB4	ENSP00000395716.1:p.Ser99LeufsTer11	S99x	frameshift	0.00039970	0.00336538	P	PVS1, PM2, PP3	.	.
.	ABCB4	ENSP00000392983.1:p.Leu759TyrfsTer38	F758x	frameshift	.	0.00009610	LP	PVS1, PM2	.	.
.	ABCB4	ENSP00000392983.1:p.Lys30GlyfsTer7	Lys30GlyfsTer7	frameshift	0.00000408	0.00020243	LP	PVS1, PM2	.	.
.	ABCB4	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	.	.
.	ABCB11	ENST00000263817.7:c.2611-2A>T	c.2611-2A>T	splice-acceptor-variant	0.00000407	0.00057870	P	PVS1, PM2, PP3	.	.

.	ABCB11	ENSP00000497931.1:p.Trp239Ter	W239x	stop-gained	.	0.00009566	P	PVS1, PM2, PP3	.	.
. . Gallstone disease	ATP8B1 ATP8B1 ATP8B1 ATP8B1	ENSP00000283684.4:p.Gln1179GlufsTer56 ENSP00000283684.4:p.Pro792HisfsTer8 ENST00000283684.9:c.182-4_183del ENSP00000283684.4:p.Glu20Ter	IQ1178-1179IX F791X ?-61 E20*	frameshift_variant & splice_region_variant frameshift_variant splice_acceptor_variant&coding_sequence_variant&intron_variant stop_gained	. . . .	0.00193798 0.00019069 0.00048956 0.00009566	P P	PVS1, PM2, PP3 PVS1, PM2, PP3	. . . .	. . . .
.	NR1H4	ENSP00000446760.1:p.Lys4Ter	K4*	stop_gained	.	0.00014188	P	PVS1, PM2, PP3	.	.
. .	TJP2 TJP2	ENSP00000438262.1:p.Glu44Ter ENSP00000345893.4:p.Gly5ArgfsTer26	E44* M1MPVX	stop_gained frameshift_variant&start_lost	. 0.00001343	0.00009604 0.00029768	LP P	PVS1, PM2 PVS1, PM2, PP3	. .	. .

Figures

Figure 1.



Inclusion criteria (<5% MAF), non-synonymous: 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.

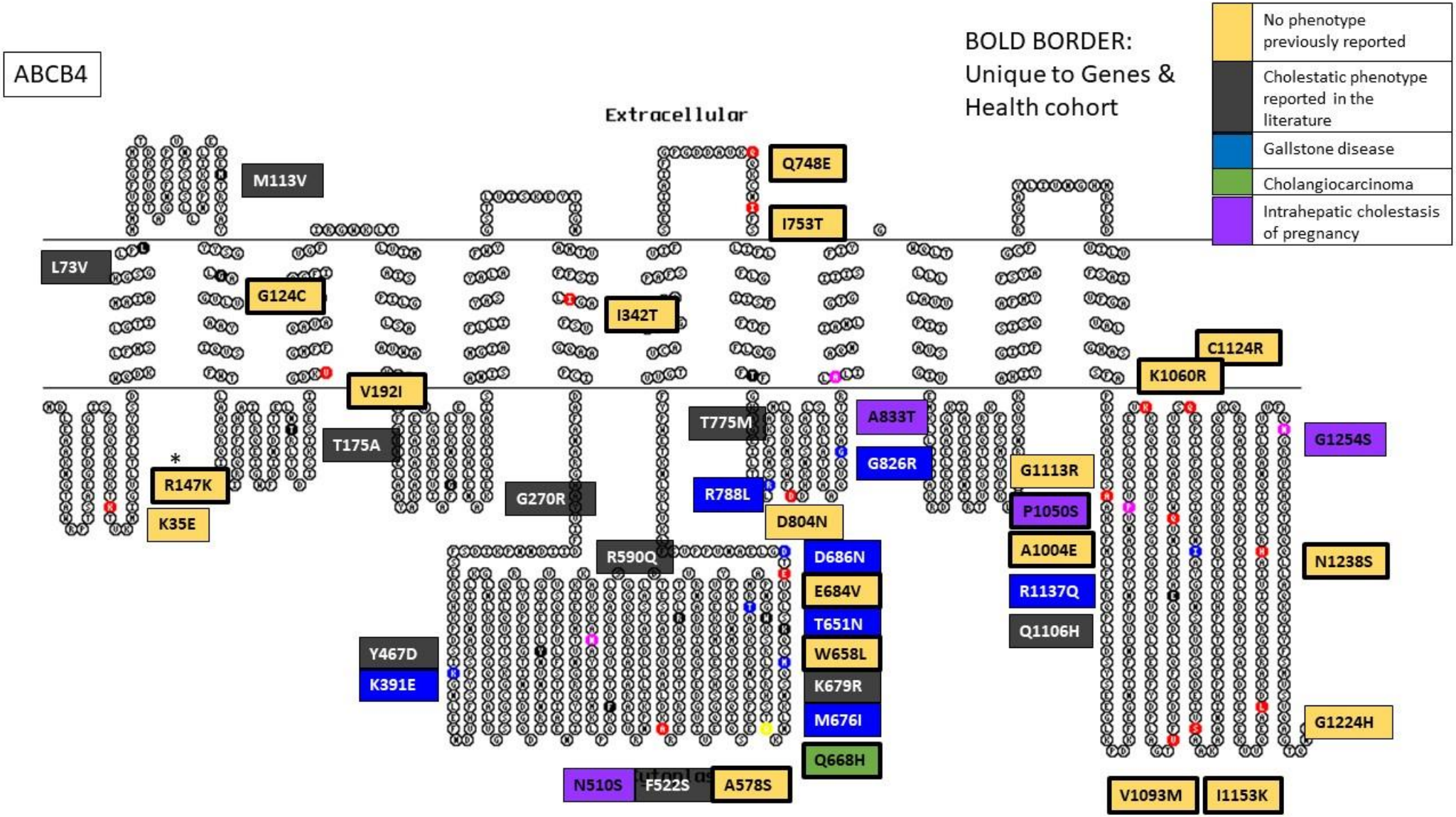

ABCB4
BOLD BORDER:
Unique to Genes &
Health cohort
No phenotype previously reported
Cholestatic phenotype reported in the literature
Gallstone disease
Cholangiocarcinoma
Intrahepatic cholestasis of pregnancy
Extracellular
M113V
Q748E
I753T
L73V
G124C
I342T
C1124R
K1060R
V192I
T175A
T775M
A833T
G826R
G1254S
R147K
G270R
R788L
G1113R
P1050S
K35E
D804N
R590Q
D686N
A1004E
N1238S
E684V
R1137Q
T651N
Q1106H
Y467D
W658L
K391E
K679R
M676I
G1224H
Q668H
N510S
F522S
A578S
V1093M
I1153K

Figure 3.

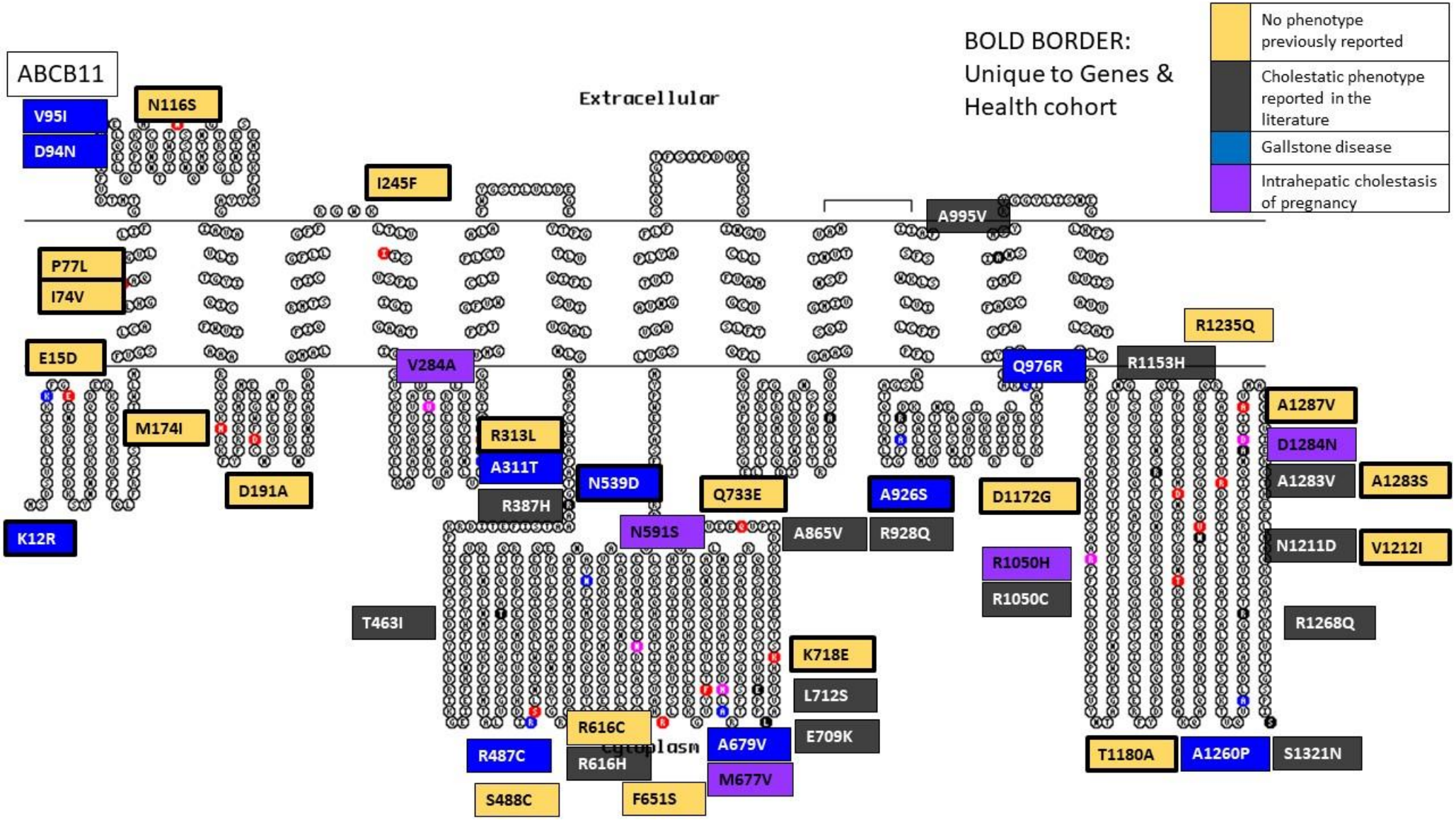
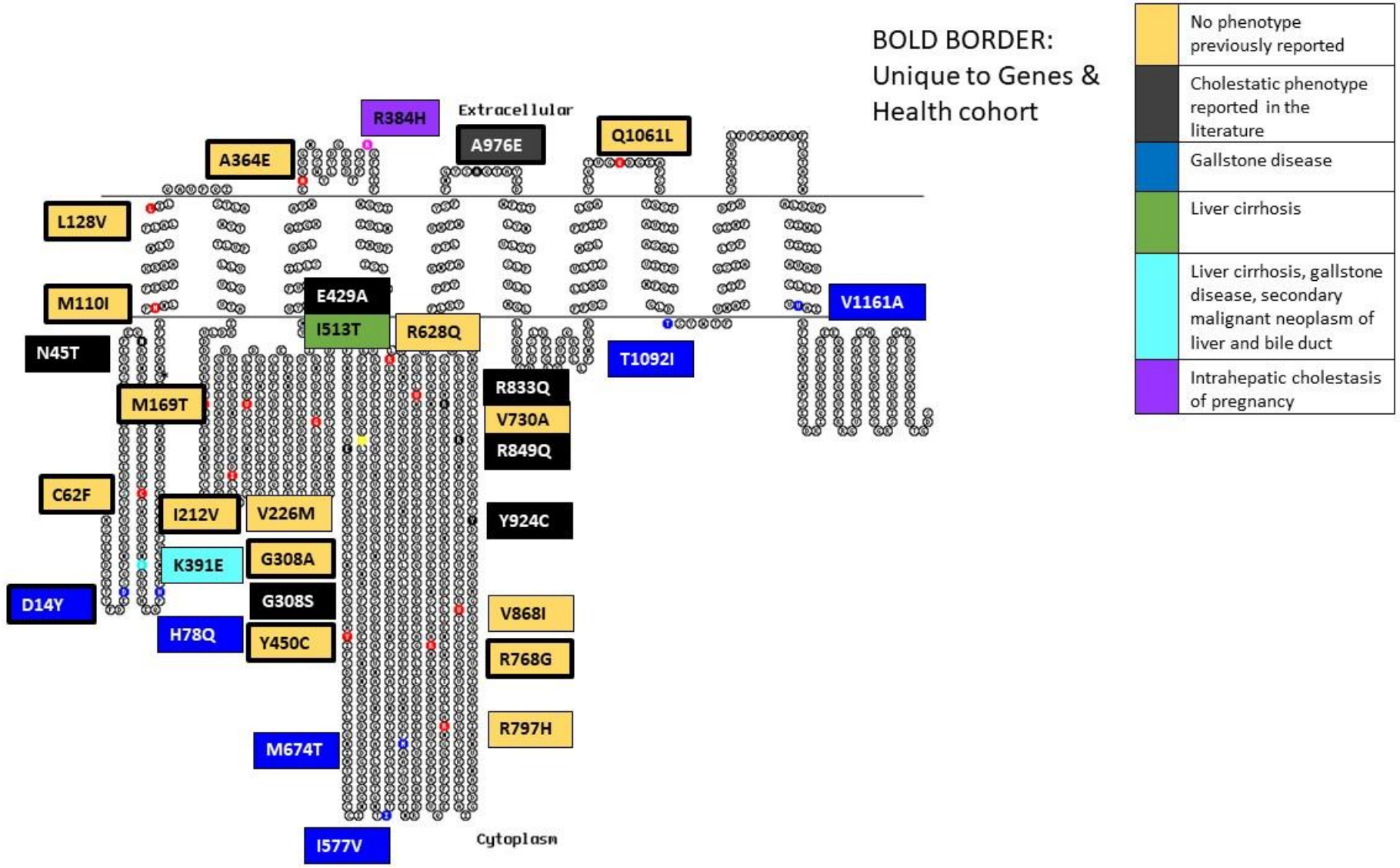

ABCB11
BOLD BORDER:
Unique to Genes &
Health cohort
No phenotype previously reported
Cholestatic phenotype reported in the literature
Gallstone disease
Intrahepatic cholestasis of pregnancy
Extracellular
Cytoplasm
V95I
D94N
N116S
I245F
A995V
P77L
I74V
E15D
V284A
Q976R
R1235Q
R1153H
M174I
R313L
A311T
D191A
N539D
Q733E
A926S
D1172G
A1287V
D1284N
A1283V
A1283S
R387H
N591S
A865V
R928Q
K12R
R1050H
R1050C
N1211D
V1212I
T463I
R1268Q
K718E
L712S
E709K
R616C
R616H
A679V
M677V
R487C
S488C
F651S
T1180A
A1260P
S1321N

Figure 4.


ATP8B1
BOLD BORDER:
Unique to Genes &
Health cohort
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
Extracellular
Cytoplasm
R384H
A976E
Q1061L
A364E
L128V
M110I
E429A
I513T
R628Q
V1161A
N45T
T1092I
M169T
R833Q
V730A
R849Q
C62F
I212V
V226M
Y924C
K391E
G308A
D14Y
G308S
V868I
H78Q
Y450C
R768G
R797H
M674T
I577V

Legends

Table legends

Table W.

Description of the five gene candidates.

Table O.

Variants identified and TSBA concentration in volunteers with a diagnosis of ICP based on electronic health records. If serial TSBA concentrations are taken in pregnancy, diagnostic threshold of 19-20µmol/L for ICP, using non-fasting samples should be used. Abbreviations: TSBA, total serum bile acid concentrations; ICP, intrahepatic cholestasis of pregnancy.

Table X.

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

Table Y.

Loss of function variants identified in the five gene candidates.

Figure legends

Figure 1.

Overall summary of mutational burden discovered in the Genes and Health cohort for all five gene candidates. a. Describes all non-synonymous variants identified; b. describes breakdown of pathogenicity; c. describes phenotypes identified.

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.