Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE CHIRURGICHE

Ciclo XXXII

Settore Concorsuale: 06/B1

Settore Scientifico Disciplinare: MED09

Molecular characterization of gene defects associated with Progressive Familial Intrahepatic Cholestasis by Next Generation Sequencing

Presentata da: Alessandro Mattiaccio

Coordinatore Dottorato **Supervisore**

Prof. Annalisa Patrizi **Prof. Pietro Andreone**

Esame finale anno 2020

ABSTRACT

Progressive familial intrahepatic cholestasis (PFIC) is a group of autosomal recessive cholestatic diseases that affects especially new-borns and children. These disorders unavoidably progressive to liver cirrhosis and portal hypertension, requiring liver transplantation. In most severe cases the progression to liver failure may occur in the first decades of life. PFIC is considered a rare disease with an estimated incidence of 1/50,000 to 1/100,000 births but it is hard to establish because of difficulties in diagnosis.

The disease has been classified into five types (PFIC1-5) based on the genetic defect involved in bile transport. The three first described PFIC-associated genes are *ATP8B1* (PFIC1), *ABCB11* (PFIC2) and *ABCB4* (PFIC3) encoding FIC1 (a P-type ATPase), BSEP and MDR3 (ABC transporters) proteins, respectively. These proteins are all located in the canalicular membrane of hepatocytes and mediate bile production. Homozygous or compound heterozygous mutations in those tree genes cause severe forms of cholestasis with an early onset, requiring liver transplantation in childhood. Heterozygous variants have also been linked to milder cholestatic phenotypes or non-progressive forms. After, mutations in *TJP2*, coding for tight junction protein 2, were linked to PFIC4. Recently, mutations in *NR1H4* gene, coding the Farnesoid X receptor (FXR), and mutations in *MYO5B*, coding Myosin VB protein, have been associated to PFIC5 and to PFIC-like form without microvillus inclusion disease, respectively. Other genes are suspected to be involved in the disease.

The aim of the project is to develop and validate a broad, reliable, rapid and cost-saving NGS genetic test for PFIC patients. 96 patients were sequenced with the four-gene panel and 80 patients were sequenced with the latest discovered genes (fifteen-gene panel).

A total of 184 different variations has been identified in all patients in the two panels: 18 pathogenic, 46 of uncertain significance, 44 likely benign and 76 benign. P/LP mutations were found in 12% of patients. Many patients had multiple variants in all genes. Patients in our cohort have from 7 to 35 variants each and we statistically demonstrate that also some

SNPs can be significantly involved in biochemical parameters and in phenotypic features that could accelerate the progression to liver failure. The high SNPs prevalence let us to hypothesize a synergistic haplotype effect in determining different cholestasis phenotypes and overlapping features.

After all, our detection rate is comparable to the literature and to other studies proposing multi-gene panels. This innovative test may be useful for the molecular diagnostics of PFIC and a better characterization and understanding of the linking between molecular defects and different subtypes of the disease. NGS method guarantees to reach quickly medical reports for all patients and their related, so should have great consequences on disease management, patient treatment and national medical system.

SECTION I: INTRODUCTION

1.1 Bile: synthesis, transport and regulation

Bile or gall is a dark green to yellowish brown fluid, produced by the liver of most vertebrates, that aids the digestion of lipids in the small intestine. In humans, bile is produced continuously by the liver, stored and concentrated in the gallbladder. After eating, stored bile is discharged into the duodenum. The composition of hepatic bile is 97% water, 0.7% bile salts, 0.2% bilirubin, 0.51% fats (cholesterol, fatty acids and lecithin) and inorganic salts. About 400 to 800 ml of bile is produced per day in adult human beings.

It is a physiological key function that leads the endo- and xenobiotics excretory way, enabling digestion and lipid absorption from the intestinal lumen. Bile secretion happens because of the osmotic filtration process: in this mechanism the secretion of osmotically active elements from hepatocytes to canalicular space establish water flow through intercellular tight junctions that are impermeable to bigger size molecules [1]. Among these solutes secreted in the duct, there are bile acids (BA), glutathione and various organic anions, in addition to bicarbonate, proteins, organic cations and lipids. Bile acids are the main solutes in bile and are regarded as the dominant osmotic driving force in its production [2]. Canalicular excretion of bile acids involves three phases: first, a highly efficient uptake from portal blood in the hepatic sinusoid; second, intracellular transport in the hepatocyte that usually can occur with the aid of chemical modifications (such as conjugation with glucuronic acid or aminoacids, hydroxylation, sulfation, etc.) and finally, excretion to biliary canaliculi [3]. After secretion into the intestinal lumen, bile acids are efficiently reabsorbed (95% of secreted BA) by the liver *via* the portal vein for reuptake at the sinusoidal pole of hepatocytes. In the last decades, the elucidation of the molecular features of proteins involved in hepatocyte transport as well as of the biology of biliary epithelia and the role of several nuclear receptors in regulating the expression of key transporters and enzymes, have provided a more detailed knowledge of the regulation of hepatobiliary transport in physiological and pathological conditions such as cholestasis [4]. Moreover, due to the importance of bile acids in determining cell injury and death in both hepatocytes and cholangiocytes and its possible role as trigger in the inflammatory response

in cholestatic disease, many studies are focusing on new BA-based therapeutic strategies with promising results [5].

The vectorial transport of bile acids by hepatocytes involves several transport proteins and enzymes (**figure 1**), including the sinusoidal transporters Sodium Taurocholate Co-transporting Polypeptide (NTCP/SLC10A1), members of the Organic Anion Transporting Polypeptides (OATPs/SLCO) family and conjugation enzymes. ATP-dependent efflux pump BSEP (Bile Salt Export Pump) is responsible for excretion of monovalent bile acids and conjugate export pump MRP2 (Multidrug Resistance-Associated Protein 2) for excretion of bilirubin and divalent bile acids (on the canalicular membrane). The multidrug export pump MDR1 assists in the excretion of cationic drugs [6]. The phospholipid export pump MDR3 "flops" phosphatidylcholine from the inner to the outer membrane leaflets, which forms mixed micelles together with bile acids and cholesterol. Other bile acids export pumps, MRP3, MRP4 and organic solute transporter (OSTa/b) are present at the basolateral membrane and function as back-up pumps for alternative sinusoidal bile acid export. The cystic fibrosis transmembrane conductance regulator (CFTR) drives bicarbonate excretion and, in the liver, it is expressed only in cholangiocytes. The biliary epithelium also reabsorbs bile acids *via* an apical Na^+-dependent bile-salt transporter ASBT and the basolateral counterpart OSTa/b [6].

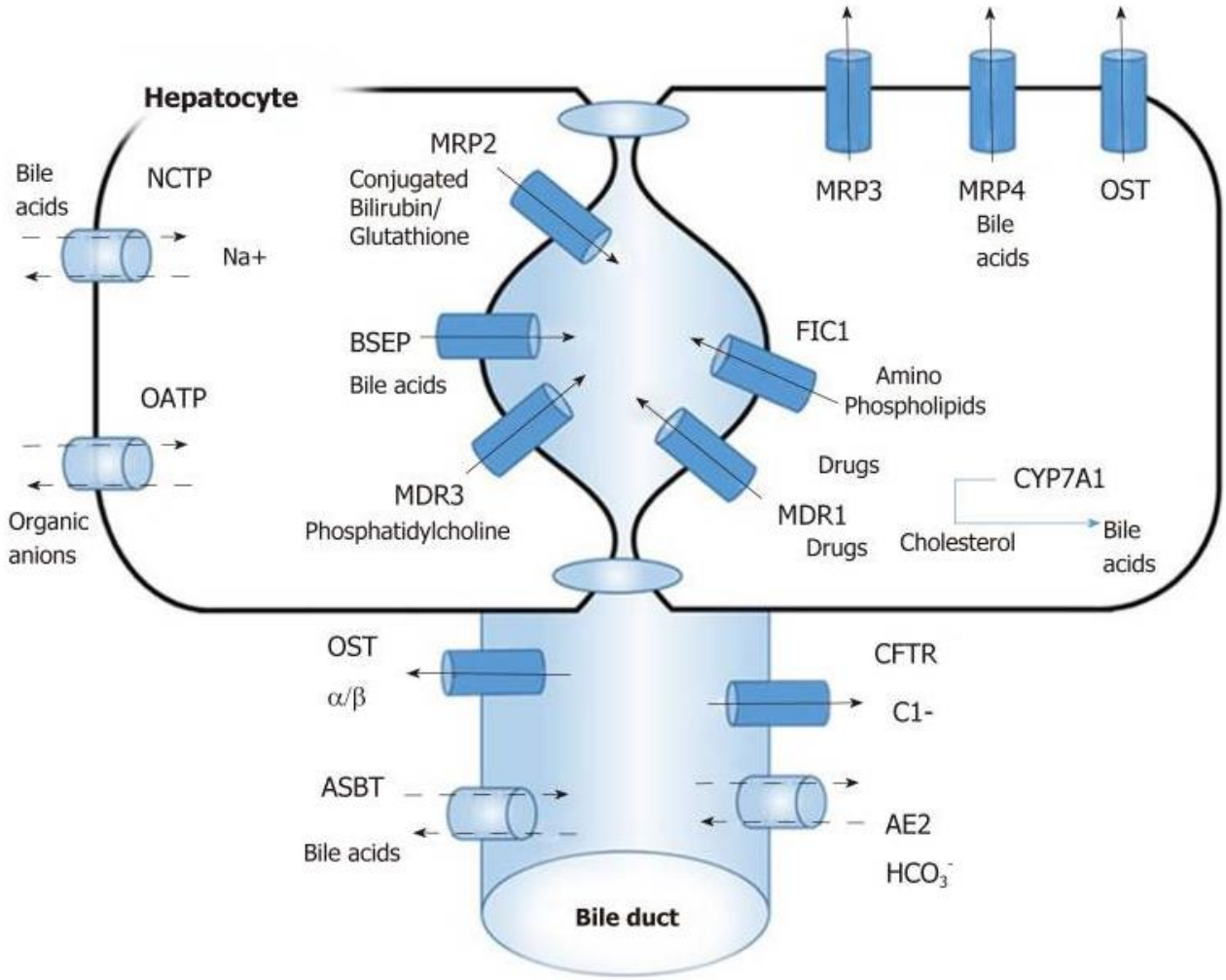


Figure 1. Overview of bile acid transport system. See details in the text. NTCP: sodium taurocholate co-transporting polypeptide; FIC1: Familial intrahepatic cholestasis 1; AE2: Solute Carrier Family 4 (Anion Exchanger), Member 2; OATPs: Organic anion transporters; BSEP: Bile-salt export pump; CFTR: Cystic fibrosis transmembrane conductance regulator; OST: Organic solute transporter; ASBT: Apical sodium-dependent bile-salt transporter. Samant H, Manatsathit W, Dies D, et al. *World J Clin Cases* 2019.

These proteins have a central role managing a rapid transition of bile acids from blood to bile and maintaining a low intracellular bile acids concentration. This is crucial to maintain hepatocyte integrity as bile acids are both signalling and detergent molecules that, at higher concentration (≥ 50 μM/L), may cause apoptosis, activate proinflammatory genes, and eventually induce cellular necrosis [7]. This cytotoxicity of bile acids leads to liver damage in cholestatic conditions: bile secretion is impaired, bile acids accumulate inside hepatocytes and, in the case of cholangiopathies, leak into the surrounding tissue due to injury of bile ducts. When this happens, the cholestatic setting alterations in the expression of

hepatobiliary transporters could represent a compensatory response aiming to limit the accumulation of potentially toxic biliary constituents. These adaptive responses are mediated by the activation of several nuclear receptors such as Farnesoid X Receptor (FXR), Pregnane X Receptor (PXR), Constitutive Active Response (CAR) and Small Heterodimer Partner (SHP) as well as by entero-hormones such as FGF-19 [8]. FXR is the major bile acids receptor and influences a myriad of pathways both in hepatocytes and in other resident cells such as Kupffer, endothelial and hepatic stellate cells. In hepatocytes, upon upregulation of SHP, FXR mediates a downregulation of NTCP and of a key enzyme in bile acid synthesis, CYP7A1 (Cytochrome P450 7A1). FXR also directly up-regulates BSEP, thus promoting bile acid excretion [9]. Lastly, alternative excretory transport proteins, located at the basolateral membrane of hepatocytes, like Organic Solute Transporter Subunit α, β (OSTα, β), Multidrug Resistance-Associated Protein 3 and 4 (MRP3 and MRP4) are expressed at low levels in physiological conditions and become up-regulated during cholestasis. Thus, if bile acid secretion is impaired, adaptive responses may limit their accumulation inside hepatocytes, preventing hepatocellular damage. If these responses are inadequate, cell damage and death may occur by either apoptosis or necrosis.

Of note, bile acids can trigger hepatocyte-specific inflammatory response in cholestatic environment that involves increased expression of cytokines such as C-C Motif Chemokine Ligand 2 (CCL2), C-X-C Motif Chemokine Ligand 2 (CXCL2), and Interleukin 8 (IL8), which in turn can contribute to neutrophil recruitment and augment local inflammation. This response is also dependent on the activation of the toll like receptor-9, maybe by BA-induced mitochondrial damage and consequent releasing of mitochondrial DNA. In addition to the local inflammation promoted by bile acids in other scenarios, such as in cholangiopathies or diseases of the bile duct, there are other types of related diseases. Mechanical obstruction is one of these and leads to an increase in biliary pressure, the occurrence of biliary infarcts and the loss of bile acids and other bile components in the surrounding tissue that can activate proliferative reactions and hepatic fibrogenesis leading to disease progression and finally to cirrhosis [10].

1.2 Different etiologies of cholestasis

Cholestasis is a clinical syndrome that can be either extra or intra-hepatic in etiology and springs from diminished bile production by hepatocytes or impaired bile secretion at the level of cholangiocyte to obstruction of bile flow by stones (cholelithiasis) or by tumor bulk. The common causes of extrahepatic cholestatic liver disease include choledocholithiasis, tumors and parasitic infections.

Several causes of intrahepatic cholestasis include immune-mediated conditions like primary biliary cholangitis (PBC), primary sclerosing cholangitis (PSC), exposure to several medications (steroids, nonsteroidal anti-inflammatory drugs, antibiotics, anti-diabetic agents) and inborn errors as genetic defects in cholesterol and bile acid biosynthesis or regarding their metabolism [11]. All these syndromes share the retention of products into blood or cholangiocytes, instead of being excreted into bile under normal activity. This condition of further retention of toxic hydrophobic bile acids cause persistent and extensive damage to the bile duct. In addition, depending on the triggering factors and its evolution, cholestatic disease is divided into progressive or non-progressive forms.

To adequately deal with cholestatic injury, it is necessary to first identify and appropriately target those defective secretory mechanisms or surgically remove or remediate lesions that obstruct or interfere with bile duct patency. The pathogenesis of many cholestatic liver disorders has been evolving with most research in bile production and secretion pathways, thus several molecular targets for diverse aspects of bile acids signalling and transport have been developed recently [5].

In the next paragraphs follows a description of the progressive forms, which are the subject of the PhD thesis, and non-progressive forms of cholestasis.

1.3 Progressive familial intrahepatic cholestasis

Progressive familial intrahepatic cholestasis (PFIC) is a group of autosomal recessive cholestatic diseases that affects especially new-borns and children. These diseases represent

a consolidated indication for liver transplantation [12]. These disorders unavoidably progressive to liver cirrhosis and portal hypertension. In most severe cases the progression to liver failure may occur in the first decades of life. PFIC is considered a rare disease with an estimated incidence of 1/50,000 to 1/100,000 births and all types of PFIC exist worldwide [13]; however, their incidence is hard to establish because of difficulties in diagnosis.

Clayton et al. first described this disease in 1965 as Byler disease in a population of Amish kindred [14]. The disease has been classified into three types (types 1, 2 and 3) based on the genetic defect involved in bile transport [15]. The three first described transport proteins associated to PFIC disease are familial intrahepatic cholestasis 1 (FIC1, gene symbol: *ATP8B1*), bile salt export pump (BSEP, PFIC2, *ABCB11*) and multidrug resistance P-glycoprotein 3 (MDR3, PFIC3, *ABCB4*). These proteins are all located in the canalicular membrane of hepatocytes and mediate bile formation. The FIC1 protein is encoded by the *ATP8B1* gene and represents a P-type ATPase, which flips aminophospholipids from the outer to the inner membrane leaflet, resulting in maintenance of membrane asymmetry, which is essential for the function of other canalicular transport proteins. BSEP (*ABCB11*) and MDR3 (*ABCB4*) are members of the subfamily B (MDR/TAP) of adenosine triphosphate (ATP)-binding cassette (ABC) transporters. As already mentioned above, BSEP excretes bile salts from the hepatocyte into the bile canaliculus, which represents the major driving force of bile salt-dependent bile flow [16]. MDR3 protein acts as a floppase and transports lipids of the phosphatidylcholine family from the inner to the outer leaflet of the canalicular membrane. A reduction or even an absence of FIC1, BSEP or MDR3 functionally active from the canalicular membrane of hepatocytes is the molecular mechanism underlying a wide spectrum of cholestatic disorders that includes both forms acquired, due to infections or drugs and hereditary cholestatic diseases. Mutations in the genes encoding FIC1 or BSEP may trigger a typical cholestatic disorder which is associated with normal levels of gamma-glutamyltransferase (low γ-GT cholestasis). In contrast, MDR3-related cholestasis is characterized by elevated γ-GT levels (high γ-GT cholestasis), which may be explained by impairment of mixed micelles due to reduced phospholipid concentrations in bile, resulting in higher relative concentrations of bile salts in bile with subsequent toxic effects on bile

ducts [17]. Progressive familial intrahepatic cholestasis (PFIC) type 1, 2 or 3 represent severe forms of cholestasis with an early onset phenotype linked to homozygous or compound heterozygous mutations in either the FIC1, BSEP or MDR3 genes, often requiring liver transplantation in early childhood. Heterozygous variants in these three genes have also been linked to milder cholestatic phenotypes or non-progressive forms like benign recurrent intrahepatic cholestasis (BRIC), intrahepatic cholestasis of pregnancy (ICP), contraceptive-induced cholestasis (CIC), drug-induced cholestasis (DIC), or low phospholipid-associated cholelithiasis (LPAC) for MDR3 variants [18]. Furthermore, late onset phenotypes have been described to have mutations in all three genes. It is also interesting that some genes responsible for PFIC are involved in the genesis of liver tumors such as hepatocellular carcinoma (HCC) and cholangiocarcinoma (CCA).

In last few years, the development of diagnostic methods such as next-generation sequencing (NGS) and whole-exome sequencing (WES) has allowed the detection of new genes responsible for PFIC4 and PFIC5. Mutations in *TJP2* (coding for tight junction protein-2) were linked to intrahepatic cholestasis with low GGT (PFIC4). The first association was described in 2014 [19] and the disease results clinically similar to PFIC2 and caused by homozygous or compound heterozygous mutations. Mutations in *NR1H4* gene, coding the farnesoid X receptor (FXR), a bile acid-activated nuclear hormone receptor that regulates bile acid metabolism, were associated for the first time to neonatal forms of cholestasis [20]. Finally, mutations in Myosin 5B gene (*MYO5B*), responsible for a PFIC-like form without microvillus inclusion disease, have been reported recently [21].

In **table 1** are represented various aspects of the different forms of PFIC, considering clinical characteristics, histology at diagnosis time and clinical outcomes.

	Progressive familial intrahepatic cholestasis					
	PFIC1	**PFIC2**	**PFIC3**	**PFIC4**	**PFIC5**	**PFIC***
locus	18q21-22	2q24	7q21	9q21.11	12q23.1	18q21.1
gene	ATP8B1	ABCB11	ABCB4	TJP2	NR1H4	MyosinVB
protein	FIC1	BSEP	MDR3	ZO-2	FXR	MYO5B
clinics	Early onset; severe jaundice/itching; growth retardation; diarrhoea, pancreatitis, deafness; leads to LT	Early onset, severe jaundice/itching; leads to LT; potential post-LT recurrence	Childhood/young adulthood onset; can be drug-triggered; hepatomegaly, growth retardation, HCC risk; leads to LT	Early severe cholestasis onset; Progression to liver failure in childhood; No post-LT recurrence; HCC risk	Neonatal onset, rapid progression to ESLD, vitK-independent; coagulopathy	Onset < 2 years: ± MVID, jaundice/itching; hepatomegaly
BA	High	Very high	High	High	High	High
GGT	Low or normal	Low or normal	High	Normal or mild elevation	Normal	Normal
AST/ALT	Mild elevation	Moderate elevation	Mild elevation	Elevation	Moderate elevation	Mild or moderate elevation
AFP	Normal	High	Normal	High	High	Normal
histology	Mild cholestasis, mild lobular fibrosis and inflammation with giant cells	Canalicular cholestasis, lobular/portal fibrosis and inflammation with giant cells	Loss of MDR3 expression, portal inflammation, portal fibrosis, cholestasis, ductular proliferation	Centrolobular cholestasis; mislocalization of claudin	Cholestasis, loss of BSEP expression	Cholestasis, Inflammation with giant cells, BSEP and MDR3 tissue expression, MYO5B and RAB11 A canalicular staining

Table 1. Summary of the typical features of progressive familial intrahepatic cholestasis associated with different genetic etiologies. Major deficiencies and relative genes involved underlying the different types of PFIC and their typical phenotypes. Abbreviations: BSEP: biliary salt export pump; MDR3: class III multidrug resistance P-glycoprotein; TJP2: tight junctions protein-2; ZO-2: zona occludens-2; FXR: fanesoid X receptor; MVID: microvillous Inclusion Disease; MYO5B: myosinVB protein; RAB11A: RAS-related protein RAB11 A; LT: liver transplantation; ESLD: end-stage liver disease; vitK: vitamin K; AST: aspartate aminotransferase; ALT: alanine aminotransferase; GGT: gamma-glutamyl transferase; AFP: alpha-1-fetoprotein; BA: bile acids. *MYO5B is classified by OMIM, online mendelian inheritance in man, as the gene responsible for microvillus inclusion disease but not yet for PFIC6. Vitale et al. *Digestive and Liver Disease* 2019.

1.3.1 Progressive familial intrahepatic cholestasis 1

Progressive Familial Intrahepatic Cholestasis Type 1, also known as PFIC1, *ATP8B1* disease, OMIM #211600 (OMIM® - Online Mendelian Inheritance in Man) and formerly Byler's disease [14], is the consequence of a severe defect in the gene encoding ATPase, BSEP/FIC1 (familial intrahepatic cholestasis 1), mapping on the long arm of the 18th chromosome (18q21.31) [22], reverse strand. This gene has a transcript length of 5935 base pairs (bp)

(ENST00000536015.1), divided into 28 exons. Its translation leads to a 1251 aminoacids long protein (NM_005603) (source: Ensembl genome browser).

It is a protein widely expressed, found in most tissues except brain and skeletal muscle and most abundant in pancreas, small intestine and liver. FIC1 protein acts as lipid flippase, carrying phosphatidylserine and phosphatidylethanolamine from ectoplasmic to the cytoplasmic leaflet of the hepatocyte canalicular membrane. FIC1 maintains the asymmetry between inner and outer leaflet of plasma membrane, having a protective function against high concentrations of bile salts. In healthy liver, elevated bile salts levels mediate activation of FXR. The FXR induces BSEP expression, thereby stimulating biliary bile salts output. In the intestine, FXR represses apical sodium-dependent bile acid transporter (ASBT) expression reducing intestinal bile salts reabsorption. Therefore, reduced FXR expression in PFIC1 represents a consequence, rather than a cause, of cholestatic phenotype in *ATP8B1* deficiency.

The main occurrences are the appearance of cholestasis, usually in the first few months of life, recurrent episodes of jaundice associated with uncontrollable itching with low gamma-glutamyl transferase. Progression to cirrhosis and end-stage liver disease (ESLD) occurs at a variable rate. In contrast to the other PFICs (PFIC2-5), clinical manifestation of PFIC1 also includes a wide range of extrahepatic symptoms, since the protein is also expressed in apical membrane of cholangiocytes, enterocytes and acinar cells of the pancreas. Extrahepatic manifestations of *ATP8B1* disease include elevated sweat chloride concentrations, delayed pubescence and growth, and watery diarrhoea as well as impaired hearing and pancreatitis [23]. These symptoms often persist, while diarrhoea even tends to worsen after liver transplantation. Liver allografts in PFIC1 patients may display diffuse steatosis with a variable necro inflammatory component, with subsequent fibrosis [24]. In fact, in a 2009 study, Miyahawa-Hayashino et al. showed that patients with post-transplant steatosis typically had more severe mutations in *ATP8B1* gene and were more likely to have systemic complications such as pancreatitis. Both diarrhoea and steatosis/steatohepatitis may improve after biliary diversion [25].

Laboratory findings show cholestasis with low serum levels of GGT, increased serum concentrations of primary bile salts and normal levels of cholesterol. Aminotransferases, starting within the reference range at the beginning of the disease, gradually increase by up to tenfold together with disease progression. From the histological point of view there are notable changes that mainly involve the canaliculi around the central veins with strong cholestatic stress. The interlobular ducts can be hypoplastic due to subnormal bile flow. Other features are absence of true ductular proliferation, portal and lobular fibrosis with inflammation and, more infrequent, giant cell transformation of hepatocytes (**figure 2**). Liver fibrosis progression corresponds with the relative stage of the disease and can lead to cirrhosis.

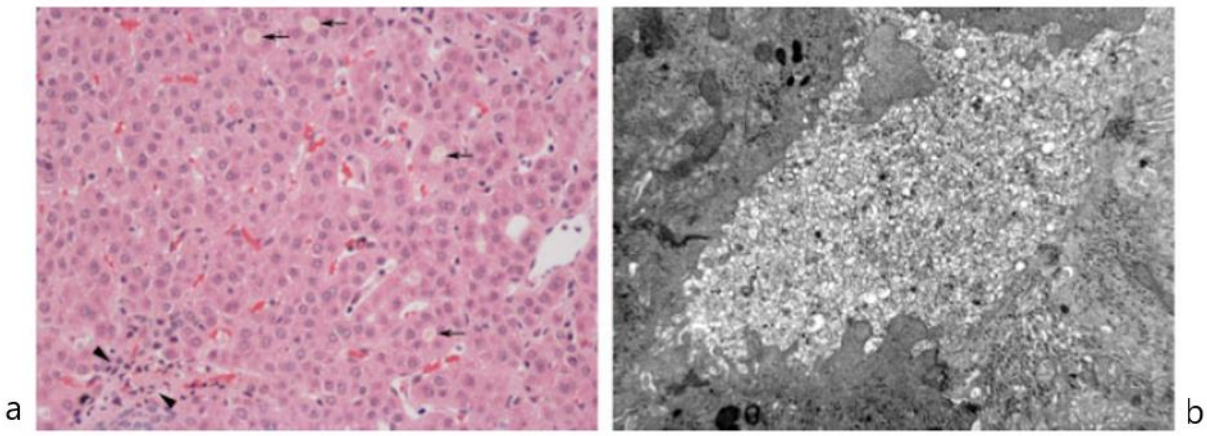


Figure 2. Familial intrahepatic cholestasis type 1 (FIC1) deficiency. (**a**) Needle biopsy of a 6-year-old girl with bland canalicular cholestasis with pseudorosette formation (arrows) and minimal inflammation. Portal tract (arrowheads) with mild fibrosis. (Hematoxylin and eosin stain, x200) (**b**) Electron microscopy of glutaraldehyde-fixed tissue: canaliculus distended by coarse, granular ''Byler bile,'' microvillous loss and swelling of residual microvilli. (Uranyl acetate/lead citrate, x8000) Morotti et al. *Semin Liver Dis*. 2011.

Ursodeoxycholic acid (UDCA) is the main treatment for early therapeutic management in children [26]. However, if there are also symptoms related to the intestine (including pancreatitis), pancreatic enzymes and fat-soluble vitamins are indicated. Cholestyramine and rifampicin have also been used for the treatment of the itching. When medical therapy fails, the partial biliary diversion (PBD) should be considered and the naso-biliary drainage may help to select potential responders to surgery. In approximately 80% of patients with

PFIC1 and 2, the partial biliary diversion results in improved growth, normalization or improvement of liver function and reduction of fibrosis [27]. If the latter are ineffective, liver transplantation is the only option but the development of the non-alcoholic fatty liver disease (NAFLD) and growth retardation can worsen the quality of life even after surgery [28].

1.3.2 Progressive familial intrahepatic cholestasis 2

Progressive Familial Intrahepatic Cholestasis Type 2, also known as PFIC2, *ABCB11* disease, (OMIM #601847) is caused by mutations in *ABCB11* gene, which is located on the long arm of the second chromosome (2q24) and encodes the canalicular transport protein ABCB11/BSEP (Bile Salt Export Pump). *ABCB11* is a reverse gene and it has a transcript length of 4775 bp (ENST00000263817.6), divided into 28 exons. Its translation leads to a 1321 aminoacids long protein (NM_003742).

BSEP protein is a liver-specific ATP -binding cassette transporter that mediates the excretion of monovalent bile salts from hepatocyte to canaliculi against a concentration gradient [29]. It is the main transporter of bile salts with a critical role in their physiologic maintenance in enterohepatic circulation. Expressed predominantly, if not exclusively in the liver, where it is further localized to the canalicular microvilli and to sub canalicular vesicles of the hepatocytes.

Long-term regulation of BSEP is complex and must fulfil many demands. It occurs in large parts on the level of transcription. Major regulators of BSEP expression are the farnesoid X receptor (FXR), the liver receptor homolog 1 (LRH1, *NR5A2* gene, Nuclear Receptor Subfamily 5 Group A Member 2) and the nuclear factor erythroid 2-related factor 2 (NRF2, *NFE2L2* gene). The farnesoid X receptor together with its obligate partner RXRα (Retinoid X Receptor Alpha) directly transactivates the BSEP promoter after bile acid binding to FXR. Chenodeoxycholate (CDCA), the predominant bile acid in humans, is the most powerful natural agonist for FXR. In mice, CDCA facilitates the recruitment of the coactivator complex ASCOM (Activating Signal Cointegrator-2 complex), including the coactivator NCOA6 and the H3K4 lysine methyltransferase MLL3, leading to methylation of histones

within the mouse Bsep promoter region, a process shown to be essential for Fxr-dependent Bsep expression [30]. The steroid receptor coactivator-2 (SRC-2; also known as nuclear receptor coactivator-2 *NCOA2* gene) was also shown to cooperatively interact with BSEP promoter stimulation by FXR. SRC-2 is activated *via* phosphorylation by the AMP activated protein kinase (AMPK), which is regarded as an energy depletion sensing kinase. SRC-2 possesses an intrinsic histone acetyltransferase activity and is recruited to promoter sites of genes. It was suggested that the AMPK-SRC2/FXR-BSEP cascade eventually would enhance fat absorption from the intestine *via* increased BSEP-dependent bile salt secretion. The importance of SRC-2 was highlighted by the observation that BSEP mRNA and protein expression was significantly reduced in hepatocytes-specific SRC-2 knockout mice, whereas deletion of SRC-1 or SRC-3 has no effect. The liver kinase B1 (LKB1), also known as serine/threonine kinase 11 (STK11), acts upstream of the AMPK and in addition activates other kinases from the AMPK-related kinase family. LKB1 regulates cell polarity and, *via* AMPK, cellular energy homeostasis. Interestingly, mice liver-specific knockout of LKB1 led to defective bile duct formation, impaired bile acid clearance and a marked retention of mBsep in intracellular pools along with reduced Ntcp mRNA (Sodium-taurocholate cotransporting polypeptide) and protein expression [31] (**figure 3**).

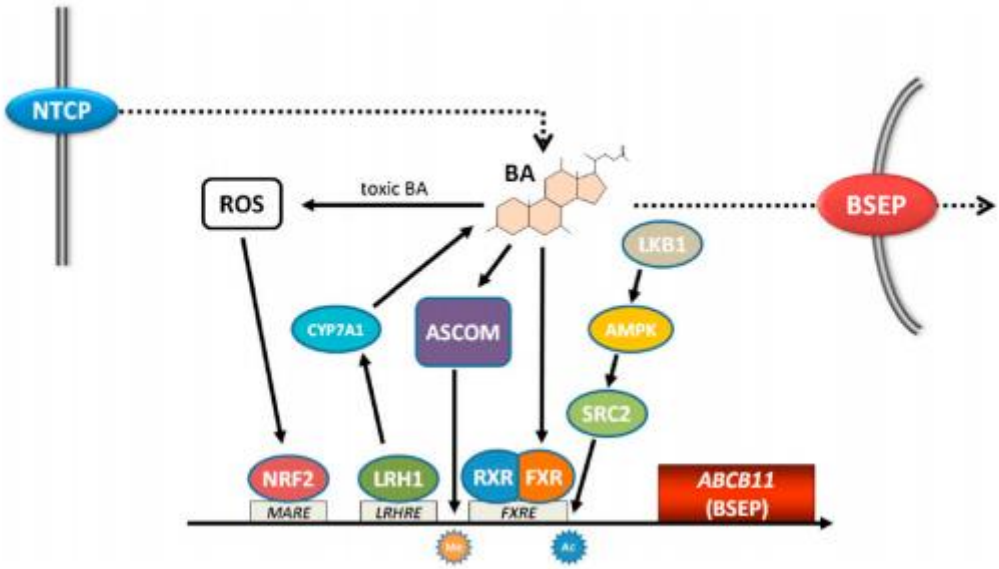


Figure 3. Regulation of the bile salt export pump (BSEP) promoter. Bile acids (BA) are major regulators of BSEP expression through activation of the farnesoid X receptor (FXR). The activating complex ASCOM (see text for details) is recruited by BA/FXR for the methylation (Me) of histones within the BSEP promoter. The nuclear factor erythroid 2-related factor 2 (NRF2) is activated by oxidative stress (e.g. by certain ''toxic'' BA), binds to a Maf recognition element (MARE) and transactivates the BSEP promoter. The steroid receptor coactivator-2 (SRC2) is activated by the liver kinase B1 (LKB1) and the ''energy sensor kinase'' AMP activated protein kinase (AMPK) and activates gene expression *via* acetylation (Ac) of histones. Liver receptor homolog 1 (LRH1) transactivates BSEP and CYP7A1, the rate-controlling enzyme of BA synthesis. Increased BSEP transcription eventually increases BSEP protein expression and removal of intracellular BA. (FXRE: FXR response element; LRHRE: LRH response element; NTCP: Na+-taurocholate cotransporting polypeptide; ROS: reactive oxygen species; RXR: Retinoid X receptor; the arrows indicate stimulation/activation; dotted lines represent BA flux.) Kubitz et al. *Clin Res Hepatol Gastroenterol.* 2012.

As regards the clinical aspects of PFIC2, they are very similar to PFIC1, but no extrahepatic symptoms are present. BSEP deficiency can lead to cholestatic jaundice and itching in the neonatal period. Moreover, PFIC2 may leads to gallstone formation and the early increase in serum aminotransferases. Another difference with PFIC1, which may have serious complications in PFIC2, is that hepatobiliary neoplasms can develop, both hepatocellular carcinoma and cholangiocarcinoma [32]. Untreated PFIC2 usually leads to progressive liver fibrosis and cirrhosis, the development of which is more rapid than in PFIC1. For this reason, screening for liver tumours should be performed from the first year of life in patients with PFIC2. In liver biopsies of patients with PFIC2 the histology shows severe lobular injury, more pronounced lobular/portal fibrosis, inflammation and more severe hepatocellular necrosis respect to PFIC1. Another severe complication is giant cell

(syncytial) hepatitis with hepatocanalicular cholestasis, while a condition of extramedullary hemopoiesis is frequently detectable in the lobules. The interlobular bile ducts can be hypoplastic and ductular proliferation is commonly observed at the peripheries of the portal triads. A reduced expression or even the absence of BSEP protein at immunohistochemistry can help the diagnostic process [33] (**figure 4**).

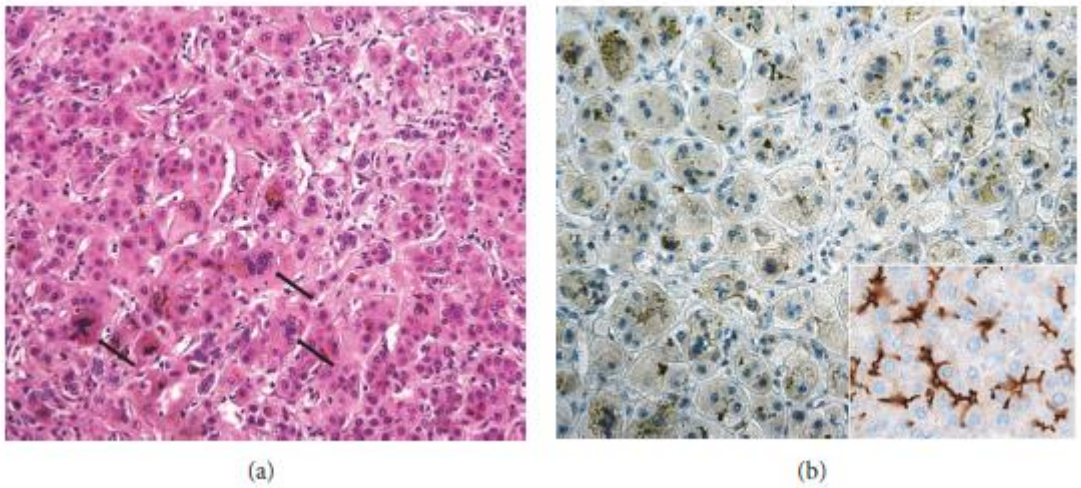


Figure 4. Histopathology of *ABCB11* disease. (a) Giant-cell hepatitis (arrows) with hepatocanalicular cholestasis and (b) complete absence of ABCB11/BSEP protein are typical findings in PFIC2 patients. Inset: immunohistochemical positivity of BSEP in the apical (canalicular) domain of hepatocytes in a healthy control. (a) Hematoxylin and eosin, original magnification x200. (b) Immunohistochemical staining with ABCB11 Rabbit Polyclonal Antibody (NBP1-89319, Novus Biologicals, USA), original magnification x200 (b), x400 (inset). Sticova et al. *Canadian Journal of Gastroenterology and Hepatology* 2018.

Ursodeoxycholic acid (UDCA) is also used as one medical therapy option. In a study of Gonzales et al. 2015, it has been observed a decrease of pruritus and serum bile acid concentration, as well as an improvement of serum liver tests by using 4-phenylbutyrate in patients carrying missense mutations. Pathological liver injuries improved and BSEP appeared at the canalicular membrane, where was not been detected before treatment, by inducing *de novo* canalicular BSEP expression [34]. This drug was previously used in patients with urea cycle defects. One curative method for patients with PFIC2 is liver transplantation and, unlike in PFIC1, steatosis and steatohepatitis are not present in liver allografts in the case of PFIC2. However, post-transplant PFIC2 phenotype recurrence has been observed,

with elevated serum levels of bile salts and bilirubin with almost normal GGT activity. Notably, in a patient's group with very high prevalence of severe BSEP mutations (e.g. splice site and premature stop codon mutations), *de novo* polyclonal inhibitory antibodies directed against the first extracellular loop of BSEP have been observed in the post-transplant serum, a condition known as Autoimmune BSEP Disease (AIBD) [35]. It is estimated that up to 8% of transplanted PFIC2 patients develop anti-BSEP antibodies and the exact mechanism by which bile salts transport is inhibited by anti-BSEP antibodies is still unknown. Interestingly, post-transplant development of inhibitory autoantibodies directed against FIC1 and MDR3 transporters with corresponding phenotypes has not been documented yet in patients with PFIC1 and PFIC3, respectively.

1.3.3 Progressive familial intrahepatic cholestasis 3

Progressive familial intrahepatic cholestasis 3, also known as PFIC3, *ABCB4* disease, (OMIM #602347) is caused by mutations in both alleles or just in one of the *ABCB4*/*MDR3* gene, on chromosome 7q21. *ABCB4* is a reverse gene and it has a transcript length of 4020 bp (ENST00000265723.4) divided into 28 exons too. Its translation leads to a 1286 aminoacids long protein. This gene encodes for the phospholipid transporter MDR3 (ATP Binding Cassette Subfamily B Member 4 or Multidrug Resistance Protein 3, NM_018849).

This protein is an energy-dependent phospholipid efflux translocator that acts as a positive regulator of biliary lipid secretion. It functions as a floppase that translocates specifically phosphatidylcholine from the inner to the outer leaflet of the canalicular membrane bilayer into the canalicular lumen of hepatocytes [36]. Phosphatidylcholine is a key component of bile salts micelles, as it reduces their detergent activity and thus protects the cholangiocytes from cellular damage. In fact, translocation of phosphatidylcholine makes the biliary phospholipids available for extraction into the canaliculi lumen by bile salt mixed micelles and therefore protects the biliary tree from the detergent activity of bile salts [37]. Cholestasis results from the toxicity of bile in which detergent bile salts are not inactivated by phospholipids, leading to bile canaliculi and biliary epithelium injuries. The expression of

ABCB4 leads to significant increases in the phosphatidylcholine, phosphatidylethanolamine and sphingomyelin contents in nonraft membranes and further enrichment of the latter and cholesterol in raft membranes (lipid rafts are specialized domains of cellular membranes that are enriched in saturated lipids) [38].

The age of PFIC3 onset ranges from 1 month to over 20 years. Patients may present less aggressive jaundice and itching is often triggered by certain drugs. Other clinical features are hepatomegaly, growth retardation and alcoholic stools. As occurs in other PFIC forms, there is evidence that either biallelic or monoallelic *ABCB4* defects may cause or predispose patients to a wide spectrum of human liver diseases: Low Phospholipid-Associated Cholelithiasis Syndrome (LPAC), Intrahepatic Cholestasis of Pregnancy (ICP), drug-induced liver injury (DILI), transient neonatal cholestasis, small duct sclerosing cholangitis, and adult biliary fibrosis or slowly evolution to cirrhosis [17]. Moreover, hepatocellular carcinoma and intrahepatic cholangiocarcinoma have been documented in patients with *ABCB4* mutations [39].

Clinical and laboratory findings associated with PFIC3 correspond to the other two forms of PFIC described above but are characterised by the absence of extrahepatic symptoms (except for cholelithiasis). Unlike PFIC1 and PFIC2, elevated serum GGT activity (often to over 10 times the normal value) and normal cholesterol levels are typically observed. The serum aminotransferases, conjugated bilirubin, and alkaline phosphatase are variably elevated and biliary phospholipids are markedly reduced.

Liver histology is characterized by non-specific portal inflammation, various stages of portal fibrosis, cholestasis with ductular proliferation and by the loss of MDR3 protein expression at immunohistochemistry (**figure 5**). Sometimes, lipid crystals within bile ducts and fibro obliterative bile duct lesions can be seen [40]. In children the disease may present in the first years of life, with early biopsies showing portal fibrosis and ductular proliferation without bile epithelial injury or periductal fibrosis. Cholestasis is present as diffuse hepatocellular cholestasis, but occasionally canalicular and ductular cholestasis can be present. Later, the liver biopsy shows biliary cirrhosis with preserved bile ducts, preserved bile duct epithelium, and no significant periductal fibrosis.

Like other PFIC forms, medical therapy with ursodeoxycholic acid (UDCA) should be considered in the initial therapeutic management of children with PFIC3. UDCA therapy may be effective in some patients, especially those with PFIC3 with missense mutations who have less severe disease (they respond to treatment with UDCA in 70% of cases) in comparison to patients with a mutation leading to a truncated protein. Liver transplantation is required in patients with PFIC3-associated liver failure. The mean age of surgical therapy is 9.6 years, ranging from 2 to 33 years [17]. There are no reports about the use of partial biliary diversion in patients with PFIC3. However, drugs that increase MDR3 expression *via* FXR activation could be good candidates for PFIC3 therapy [41].

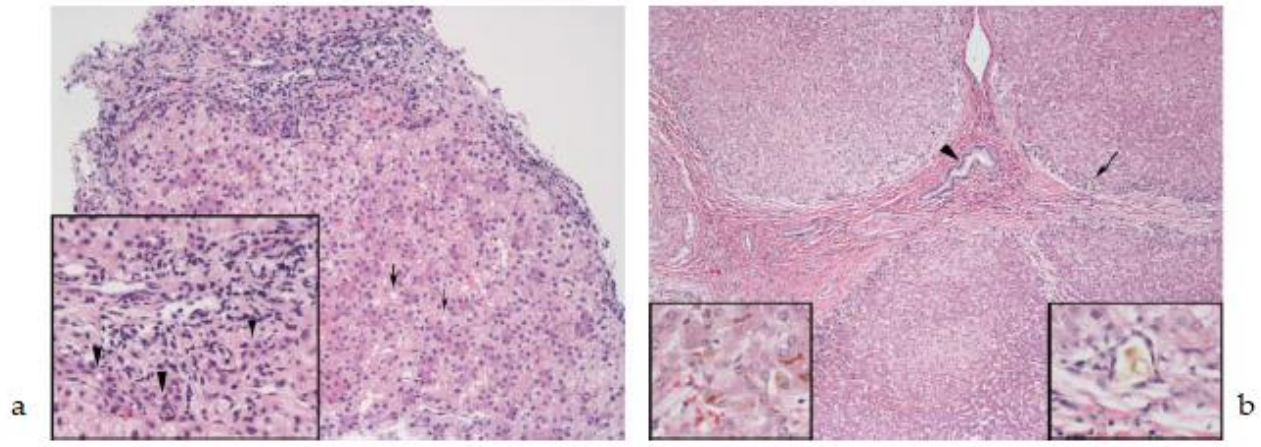


Figure 5. Multidrug resistance protein 3 (MDR3) deficiency. (**a**) Early: Needle biopsy from a 6-year-old boy with fibrous portal expansion and interportal bridging. The lobules contain scattered hepatocellular pseudorosettes (arrows). (Hematoxylin and eosin stain; 100) Inset: Portal area with bile ductular reaction (arrowheads) and mild chronic inflammation. (x200) (**b**) Late: Explanted liver from a 12-year-old girl with biliary-type cirrhosis, bile ductular reaction and occasional ductular bile plugs (arrow; also seen at higher magnification [x200] in the inset at the right lower corner); preserved bile duct (arrowhead); hepatocellular and canalicular cholestasis (better seen in the inset at the left lower corner, x200). (Hematoxylin and eosin stain; x40). Morotti et al. *Semin Liver Dis*. 2011.

1.3.4 Progressive familial intrahepatic cholestasis 4

The fourth form of PFIC (PFIC4, *TJP2* deficiency, OMIM #615878), first described in 2014 [19], is caused by homozygous or compound heterozygous mutations in the *TJP2* gene, located in the long arm of chromosome 9 (9q12). *TJP2* is a forward strand gene and it has a transcript

length of 4500 bp (ENST00000539225.1) divided into 23 exons. Its translation leads to a 1221 aminoacids long protein (NM_001170416).

Tight junction protein 2 (TJP2, also zona occludens-2, ZO-2) is a cytoplasmic component of cell-cell junctional complexes expressed in most epithelia and, through interaction with cytoskeletal proteins and integral membrane proteins, creates a link between transmembrane tight junction proteins and the actin cytoskeleton. Human *TJP2* gene generates 2 isoforms (A and C), both widely expressed. Isoform A is abundant in the heart and brain and detected in skeletal muscle. It is present almost exclusively in normal tissues. Isoform C is expressed at high level in the kidney, pancreas, liver, heart and placenta and does not detected in brain and skeletal muscle, found in normal as well as in most neoplastic tissues [42]. Complete TJP2 deficiency is associated with a significant reduction in an integral tight junction protein, Claudin-1 (*CLDN1* gene), predominantly in the canalicular membranes of liver cells. This deficiency subsequently leads to the disruption of intercellular connections and the leakage of bile through the paracellular space into the liver parenchyma [43].

Immunohistochemical studies [19] using liver tissue do not detected TJP2 protein (**figure 6, fig. 6a**). Claudin-1 and claudin-2 are known to be expressed in the liver and to manifest different subcellular localisations and they have been demonstrated to bind to the PDZ1 domain of TJPs in vitro [44]. Immunohistochemical analysis of CLDN1 expression in controls showed its presence in tight junction. Staining was substantially reduced in liver tissue from patients (**fig. 6b**). Normal expression of CLDN1 was, however, identified by western blotting analysis (**fig. 6d**). As expected, CLDN2 showed a pericanalicular staining pattern in controls. This distribution was preserved in patient material with an apparent increase in abundance (**fig. 6c**). However, western blotting analysis showed normal levels of CLDN2 protein (**fig. 6d**). These findings suggest that CLDN1, although expressed, fails to localise in the absence of TJP2, whilst CLDN2 localisation is maintained.

The recently described paediatric cases [19] presented a severe cholestatic disease with normal or, at most, slightly increased serum GGT activity and were free of mutations in *ATP8B1* and *ABCB11* genes. Those patients displayed severe progressive cholestatic liver disease in

early childhood, which puts them at increased risk of developing hepatocellular carcinoma. In addition to liver impairment, extrahepatic features have been identified in PFIC4 patients, including neurological and respiratory disorders. The PFIC4 onset appeared within 3 years in twelve pediatric cases. Nine of twelve patients needed liver transplantation within 10 years and no recurrences after surgery were observed. All homozygous mutations were predicted to abolish protein translation, consistent with a complete loss of function. Eighteen of the twenty-nine families examined in this study were consanguineous [19]. Until then, Zhou et al. in 2015 [45] reported 2 patients with PFIC4 who developed hepatocellular carcinoma. One was compound heterozygous for *TJP2* mutations and the other was homozygous for a frameshift mutation. Carlton et al. in 2003 have been previously described a single homozygous missense mutation in *TJP2* as causing benign familial hypercholanemia, a rare disorder of oligogenic inheritance, which usually manifests in elevated serum bile salts concentrations, pruritus, and fat malabsorption, but which does not lead to the development of liver disease [46].

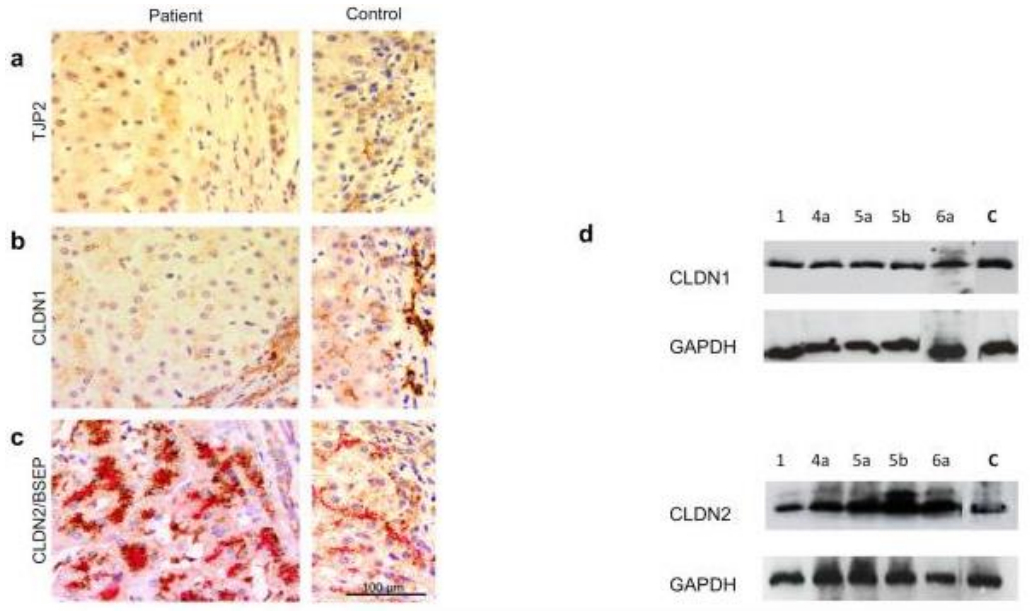


Figure 6. Protein consequences of *TJP2* mutation. Immunohistochemical staining of a PFIC4 patient (left panels), with control (right panels) for TJP2 (**a**), CLDN1 (**b**) and CLDN2/BSEP (**c**) proteins. TJP2 expression (brown) is absent from cholangiocytes and canalicular margins. CLDN1 expression (brown) is markedly reduced at both sites. CLDN2 expression (brown), clustered within cytoplasm adjoining bile canaliculi (highlighted by BSEP, red), as has been shown previously. Hematoxylin counterstain; scale bar= 100 μm. (**d**) Western blot for CLDN1 and CLDN2 proteins isolated from liver biopsy tissues (1 control and 5 patients). Normalization was assessed using GAPDH as a loading control. Sambrotta et al. *Nat Genet*. 2014.

1.3.5 Progressive familial intrahepatic cholestasis 5

PFIC5 (OMIM #617049) is a cholestatic disorder caused by mutations in *NR1H4*, a gene encoding farnesoid X receptor (FXR), the key regulator of bile salts metabolism. FXR is the receptor for chenodeoxycholic acid (CDCA), lithocholic acid, deoxycholic acid (DCA) and allocholic acid (ACA). *NR1H4* gene maps on the long arm of chromosome 12 in position 12q23.1 and has a transcript length of 1630 bp (ENST00000188403.7), divided in 9 exons, producing a 482 aminoacids long protein (NM_001206993). The receptor is predominantly expressed in the liver, intestine and to lesser extent in kidney, adipose tissue, muscle and adrenal glands.

The association between *NR1H4* mutations and cholestasis has been first described in 2016 and only 5 patients are reported in the literature. Gomez-Ospina et al. [20] performed whole-exome sequencing and single-nucleotide polymorphism (SNP) arrays of two unrelated probands with severe cholestasis. These analyses revealed homozygous loss of function mutations in *NR1H4*. Subsequently, further analysis confirmed the same *NR1H4* genotype for an additional affected individual in each family. Consanguinity in one family was confirmed by SNP array.

At the time of initial evaluation all patients had conjugated hyperbilirubinemia, elevated aminotransferases, low-to-normal GGT and elevated prothrombin time. Alpha-fetoprotein was strikingly increased early in the course of the disease and trended down with time. Two of the four patients developed liver failure in the first 2 years of life with worsening coagulopathy, hypoglycaemia and hyperammonaemia.

Histological examination of the liver in all four patients showed ductular reaction, diffuse giant cell transformation and ballooning of hepatocytes and intralobular cholestasis (**figure 7**). Liver immunohistochemistry showed no detectable expression of BSEP in bile canaliculi in all patients, but expression of the phospholipid transporter MDR3 was preserved.

In another study of 2019, Chen et al. detected the compound heterozygous mutations in the *NR1H4* gene, c.447_448insA (p.N150Kfs*6) and c.1099C>T (p.R367*) in a female baby born

at term with 3 healthy elder brothers with nonconsanguineous parents. The patient died of sepsis at age of 5 months, before being able to perform liver transplant [47].

Since FXR affects several metabolic pathways, some *NR1H4* variations or SNPs may be associated with susceptibility to various human diseases like gallstone formation, differences in glucose homeostasis, ICP, inflammatory bowel disease and several other disorders [48].

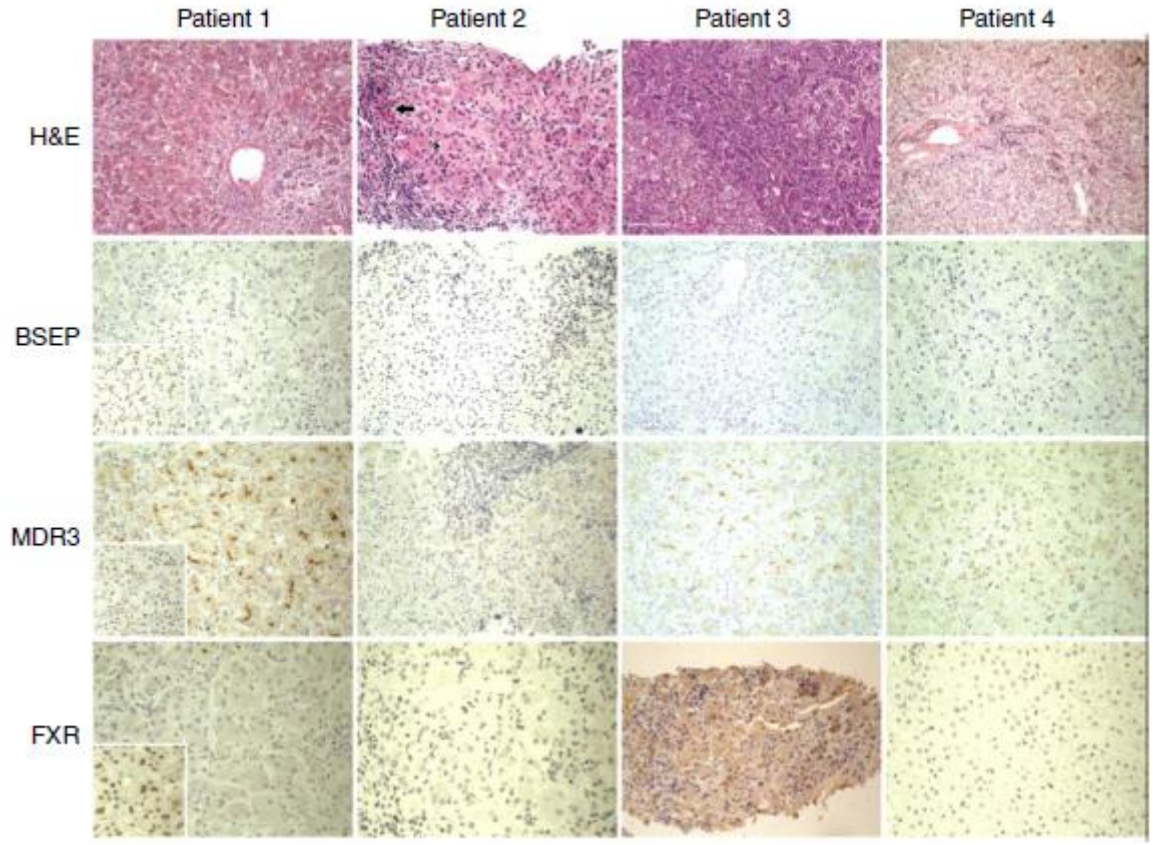


Figure 7. Histological and immunohistochemical findings in *NR1H4* mutant livers. Expression of BSEP, MDR3 and FXR in liver (diaminobenzidine chromogen and hematoxylin counterstain) in patients with *NR1H4*-related cholestasis; normal controls in insets. Hematoxylin/eosin views for orientation. Original magnification, all images, x200. BSEP and FXR are absent in all four patients when compared with positive control (inset). MDR3 is present in all four patients although reduced in patients 2 and 4. H&E, hematoxylin and eosin. Gomez-Ospina et al. *Nat Commun.* 2016.

1.3.6 PFIC linked to defects in MYO5B gene

MYO5B is a huge gene and maps on the long arm of chromosome 18 in position 18q21.1 and has a transcript length of 9505 bp (ENST00000285039.7), divided in 40 exons, producing a protein of 1848 aminoacids (NM_001080467).

MYO5B gene encodes for myosin-Vb, an actin-associated molecular motor. This protein interacts with recycling endosome-associated RAB family proteins, in particular with RAS-related protein RAB11A. This protein complex is essential for proper functioning of polarized epithelial cells, including hepatocytes, such as subcellular positioning of recycling endosomes and normal bile canaliculus formation [49]. This interaction could be responsible for different systemic manifestations. In fact, some patients also suffer short stature, though others have normal growth. Some patients with this disease can also have neurologic findings, though it is not clear if these are related to the gene defect [50]. *MYO5B* gene mutations cause microvillus inclusion disease (MVID, OMIM #251850), a congenital disorder of the enterocyte leading to intractable diarrhoea, requiring parenteral nutrition and often necessitating intestinal transplantation [51]. Cholestasis with normal-range GGT and intractable pruritus has been reported as an atypical symptom or as a side effect of parenteral nutrition in some MVID patients before or after intestinal transplantation [52].

Girard et al. in 2014 first described MVID patients developing a cholestatic liver disease akin to progressive familial intrahepatic cholestasis, characterized by intermittent jaundice, intractable pruritus, increased serum bile acid levels and normal GGT activity. Liver histology showed canalicular cholestasis, mild-to-moderate fibrosis, and ultrastructural abnormalities of bile canaliculi. No mutation in *ABCB11* or *ATP8B1* genes were identified but immunohistochemical studies demonstrated abnormal cytoplasmic distribution of MYO5B, RAB11A, and BSEP proteins in hepatocytes [53]. These first results suggested that cholestatic disease can arise by impairment of the MYO5B/RAB11A apical recycling endosome pathway in hepatocytes, resulting in an altered targeting of ABC transporter proteins, (BSEP is one of these) to the canalicular membrane.

In the study of Gonzales et al. 2017, they identified four compound heterozygous and one homozygous *MYO5B* mutation in five paediatric patients with normal-GGT PFIC phenotype and without clinical bowel disease, in whom other causative gene mutations had not been found [54]. The age of onset of symptoms was about 1 year and at a median age of 5 years none of the children have progressed to liver failure nor been listed for liver transplantation. Liver histology showed lobular and portal fibrosis, numerous multinucleate giant hepatocytes, hepatocellular and canalicular cholestasis and absence of bile duct proliferation (**figure 8**). Interestingly, Qiu et al. also showed a prevalence of *MYO5B* deficiency in 20% of previously undiagnosed low-GGT cholestasis cases in whom neither recurrent diarrhoea nor received parenteral alimentation have been reported [21].

In addition to supportive care for nutrition and diarrhoea, patients have been treated with antipruritic and anticholestatic agents, including UDCA, rifampin, cholestyramine, traditional Chinese medicine [55].

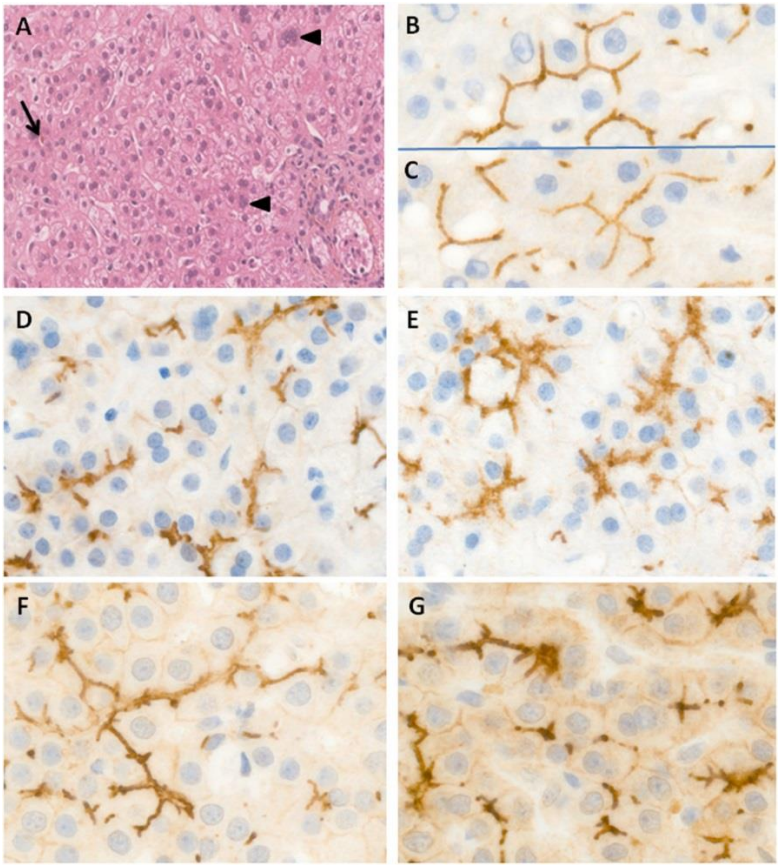


Figure 8. Liver histology and BSEP and MDR3 immunostaining studies in two children with *MYO5B* mutations and cholestasis with normal serum GGT activity. (A) Presence of giant hepatocytes (arrowhead) and of canalicular cholestasis (arrow). Hematoxylin and eosin, x200. (B) Normal pattern of canalicular expression. Normal human liver, BSEP immu-nostaining, x400. (C) Normal pattern of canalicular expression. Normal human liver, MDR3 immunostaining, x400. (D, F) Presence of an abnormal pattern of expression, with thickened canalicular staining and a granular and patchy pattern in the subcanalicular area. BSEP immunostaining, x400. (E, G) Presence of an abnormal pattern of expression, with thickened canalicular staining and a granular and patchy pattern in the subcanalicular area. MDR3 immunostaining, x400. Gonzales et al. *Hepatology* 2017.

1.4 Non-progressive familial intrahepatic cholestasis

1.4.1 Benign recurrent intrahepatic cholestasis

Benign recurrent intrahepatic cholestasis (BRIC) is a group of genetically heterogeneous autosomal recessive disorder, characterised by intermittent episodes of cholestasis. BRIC is caused by homozygous or compound heterozygous mutations in *ATP8B1* and *ABCB11* genes. For this reason, it is divided into two different phenotypes: BRIC1 (Summerskill-Walshe-Tygstrup syndrome, OMIM #243300) and BRIC2 (OMIM #605479), caused by partial deficiency in *ATP8B1* [56] and *ABCB11* [57].

The symptoms and physical aspects of BRIC are summarized in **table 2**. Typical clinical manifestations are pruritus and jaundice. One fourth of patients may never develop pruritus while in others the scleral icterus is barely noticeable. Other common symptoms include illness, irritability, nausea, vomiting, and anorexia. A small percentage of patients experience fever, arthralgias, headaches, and urticaria. Steatorrhea caused by fat malabsorption and anorexia can lead to significant weight loss during prolonged episodes of cholestasis. Other complications dependent on vitamin K malabsorption are coagulopathy and haemorrhagic complications such as bruising, epistaxis, vaginal and gingival bleeding [58].

The first attack of pruritus and jaundice typically occurs during the patient's teens, twenties or more, with itching that usually anticipates the appearance of jaundice (by 2 to 4 weeks). The duration and number of attacks vary widely in each patient and each attack can last from 2 weeks to 18 months. The mean duration is approximately 3 months. During the patient's life, there can be up to thirty attacks. Between each attack, asymptomatic periods range from as little as 1 month to as long as 33 years and the average frequency of attacks is slightly more than once every 2 years. The triggering factors can be the pregnancy, the infections or drugs such as oral contraceptives. Biochemical findings during cholestatic attacks reveal increase of bilirubin and bile salts levels in serum, while GGT activity and serum cholesterol levels tend to be normal. Enzyme levels of aminotransferase activity are usually in normal fields or slightly increased, while biochemical parameters fall within

normal ranges. Liver biopsies performed during cholestatic attacks have demonstrated hepatocanalicular cholestasis without fibrosis, while tissue expression of BSEP is reduced or mostly maintained at immunohistochemistry in BRIC2 [57]. During the asymptomatic period, the histological features are completely normal.

	Symptomatic period	Asymptomatic period
Symptom		
Pruritus	Severe	Absent
Nausea	Sometimes present	Absent
Steatorrhea	Frequently present	Absent
Weight loss	Sometimes present	Weight regained
Bleeding	Sometimes present	Resolved
Physical finding		
Jaundice	Present	Absent
Excoriations	Frequently present	Resolved
Hepatomegally	Sometimes present	Resolved
Splenomegally	Never	Never
Ecchymoses	Sometimes present	Resolved

Table 2. Symptoms and physical findings of BRIC during symptomatic and asymptomatic periods.

To perform a differential diagnosis for BRIC is necessary to exclude other diseases that could be overlapping or related to cholestatic manifestations, as the extrahepatic biliary obstructions, the acute viral hepatitis, the autoimmune hepatitis, and the immune-related cholestatic disorders [15]. After clinical exclusion of other diseases, molecular diagnosis of BRIC1 and BRIC2 should be done, demonstrating evidence of mutations in both alleles of the corresponding genes, *ATP8B1* and *ABCB11*.

While mutations causing PFIC are often located in the conserved regions of the genes that encode for important functional domains of the corresponding proteins, mutations in BRIC usually only partially impact protein function and expression. However, several cases of initially benign episodic cholestasis that have subsequently transitioned to a persistent progressive form of the disease have been reported, especially if the disease has an early onset. Van Ooteghem et al. in 2002 described four infant patients in whom liver biopsies performed in the early stages of the disease showed normal liver architecture, while late

stage biopsies revealed evident fibrosis with porto-portal septa formation [59]. Thus, BRIC and PFIC seem to represent two extremes of a continuous spectrum of phenotypes of the one disease.

The treatment of BRIC is symptomatic and aims to reduce the frequency of the cholestatic attacks and to prevent the relapses. The administration of statins, corticosteroids, cholestyramine, or ursodeoxycholic acid (UDCA) is usually not very effective in BRIC patients. In this regard, rifampicin, antihistamines, phenobarbital and carbamazepine have been proposed in recent years. Nevertheless, severe hepatotoxicity after long-term administration of rifampicin has been reported in patients with cholestatic disorders [60, 61]. Non-medical treatments include short-term naso-biliary drainage, which can improve the pruritus and reduce the duration of cholestatic attacks [62].

1.4.2 Intrahepatic Cholestasis of Pregnancy

Intrahepatic cholestasis of pregnancy (ICP, OMIM #147480), also called obstetric cholestasis, is the most common liver disease of the pregnancy. It is a liver condition of pregnancy characterized by pruritus and the biochemical finding of elevated serum bile acids, often in the presence of other signs of liver dysfunction. By definition, ICP is confined to pregnancy and the peripartum period and usually resolves soon after delivery. It is diagnosed after exclusion of other causes of cholestasis. In addition to the maternal symptomatology, ICP can be associated with adverse fetal outcomes: spontaneous preterm birth and meconium staining of the amniotic fluid can be registered especially in ICP patients with high levels of bile salts (>40 mol/L), while the risk of stillbirth is increased when serum bile salts concentrations are ≥100 mol/L [63]; hence, appropriate recognition and management are extremely important.

Several studies confirmed ICP to be characterized by glucose intolerance and dyslipidaemia, consistent with the changes observed in the metabolic syndrome. This adverse effect was secondary to maternal aberrant bile acid homeostasis. From the 12th week of pregnancy, lipid parameters, including total cholesterol, triglycerides, low-density lipoprotein-

cholesterol (LDL) and high-density lipoprotein-cholesterol (HDL) are increased, especially in the second and third trimesters [64].

The incidence of ICP ranges from 0.1% to 15.6% worldwide and it is dependent on geographic location and ethnicity. For example, the incidence is 4% in Chile; in the United Kingdom it is 0.7% but is higher in women of Indian or Pakistani origin, whereas the incidence in China has recently been estimated at 1.2%, based on more than 100,000 hospital births. Within the Californian population, genetic ancestry mapping has indicated higher levels of ICP in women of Native American ethnicity [65].

The heterozygous mutations in *ATP8B1*, *ABCB11* and *ABCB4* genes in ICP have been reported [66]. Mutations in *ABCB4* were the first to be described in ICP and are implicated in up to 20% of ICP cases. In *ABCB11* gene, polymorphism c.1331T>C (p.V444A) in exon 13, has been consistently observed in pregnant women [67]. It was reported in 83% of Caucasians with ICP and the susceptibility to this disorder was increased if combined with other mutations in the same gene or in the other two for PFIC1 and PFIC2.

The heterozygous mutations have also been identified in *ATP8B1* gene during ICP, but just few cases are described [68]. Another gene involved in PFIC5, *NR1H4*, has been previously associated with the ICP. Van Mil et al. identified 4 *novel* heterozygous FXR variants (c.-1G>T, p.M1V, p.W80R, p.M173T) in ICP. They demonstrate functional defects in either translation efficiency or activity for 3 of the 4 variants (c.-1G>T, p.M1V, p.M173T), suggesting a predisposition of those alleles for ICP [69]. In an ICP cohort of 147 women Dixon et al. have identified some evidence for involvement of *TJP2* gene, with three occurrences of variants where bioinformatics prediction tools did not agree on the clinical relevance of these changes but that may predispose to the disease [70].

Differential diagnosis should include other causes of jaundice such as HELLP syndrome (a pregnancy-associated disease characterized by Haemolysis, Elevated Liver enzymes, and Low Platelets in the mother), acute fatty liver of pregnancy, viral hepatitis, Budd–Chiari syndrome and Primary Biliary Cholangitis [15]. Cholestyramine, dexamethasone, S-Adenosyl-l-methionine, rifampicin and UDCA have been used as medical therapy in ICP. The UDCA represents the first-line therapy being able to reduce maternal and fetal complications of

ICP, improving maternal symptoms and biochemistry values. Pregnant women with previous ICP and fetal mortality should start UDCA therapy early in subsequent pregnancies [71]. The rifampicin, a potent pregnane-X receptor agonist, has reported to be useful as a second line therapy in one third of pregnant women not responsive to UDCA alone. S-Adenosyl-l-methionine improves itching in the mother while dexamethasone promotes only the fetal lung maturity. Oral contraceptives represent a risk factor for further liver impairments in women having past ICP history. Usually, in presence of ICP phenotype and when the clinical picture in well established, physicians induce the birth between 37 and 38 weeks of gestation to reduce the risk of intrauterine death. Although the blood tests and liver parameters are usually normalized within 2–8 weeks after delivery, a follow-up is necessary for ICP patients because these women may continue to have liver disfunctions, with the possibility to develop cirrhosis, non-alcoholic pancreatitis and cholelithiasis overtime [72].

1.4.3 Drug induced cholestasis

Drug-induced liver injury (DILI) is caused by a different number of triggering factors, including drug-induced cholestasis (DIC) and it is another form of acquired liver disease, accounting for approximately 2 to 5% of hospitalizations for jaundice, 10% of cases of hepatitis in all adults, and more than 40% of hepatitis cases in adults older than 50 [73]. Approximately 30% of DILI are from cholestatic origin.

The liver pathology of drug-induced liver injury covers a wide spectrum of lesions from bland cholestasis to hepatitis and mixed forms. It occurs with many drugs through a variety of mechanisms, which might differ in their clinical presentations, ranging from asymptomatic mild biochemical abnormalities to an acute illness with jaundice that resembles viral hepatitis. Although good epidemiologic data exist on the entire spectrum of drug-induced liver injury (DILI), data on the incidence of cholestatic forms of DILI are scarce.

Cholestatic liver injury is typically characterized by a predominant elevation in alkaline phosphatase and bilirubin levels. Bland canalicular cholestasis is typically associated with minimal hepatocellular inflammation and normal or only slightly elevated aminotransferase levels and is often seen with anabolic steroids or oral contraceptives. In contrast, portal inflammation is seen in hepatocanalicular cholestasis, often associated with an elevation of aminotransferases. Acute or chronic hepatitis can be clinical manifestations of DIC. Sometimes the itching, the xanthomas and/or the melanoderma are present.

The main limit during the systemic clearance of lipophilic drugs and their metabolites by the liver is their excretion into the bile. Several ATP-dependent canalicular transporters regulate this process: BSEP protein (*ABCB11*), the multidrug resistance protein-2 (MRP2, *ABCC2* gene) able to cause bile salts independent flow by excretion of glutathione, the multidrug resistance-1 protein (MDR1, *ABCB1*) which transports organic cations and the MDR3 protein (*ABCB4*).

The list of the drugs that can induce cholestasis includes the non-steroidal anti-inflammatory drugs, antihypertensives, antidiabetic, anticonvulsants, lipid-lowering agents, and psychotropic drugs [74]. Many drugs cause cholestasis through the interaction with hepatic transporters. To date, a relatively strong association between DIC and attenuated BSEP activity has been proposed. Inhibition of BSEP due to a drug activity, should lead to reduced bile salts secretion and their retention within hepatocytes, leading to cholestasis. The drugs such as rifampicin, cyclosporine A, rifamycin, bosentan, troglitazone, erythromycin, glibenclamide inhibit the BSEP protein with a dose-dependent action. Chlorpromazine, imipramine, itraconazole, haloperidol, ketoconazole, sequinavir, clotrimazole, ritonavir and troglitazone restrain the MDR3 activity *in vitro*, and resulting in an exposure of the biliary epithelium to the toxic detergent effects of bile salts, thus increasing susceptibility to DIC or to the vanishing bile duct syndrome.

The diagnosis of DIC can be difficult, due to the presence of the concomitant medications or of other chronic liver diseases [75]. The diagnosis of DIC is dependent on the accurate clinical history: the use of drugs, herbal remedies and naturopathic substances should be always investigated. Temporal relationships between taking the drug and development of liver

injury can be different, ranging from a few days to several weeks or longer (1–12 months). Furthermore, the use of mixed pharmacotherapies increases the risk of drug-drug interactions. The differential diagnosis should be made versus primary biliary cholangitis, sepsis, autoimmune hepatitis and obstructive biliary diseases. Liver biopsy may be helpful when the diagnosis is difficult; a loss of BSEP/MDR3 expression at immunohistochemistry increases the chance for DIC diagnosis. In DIC, the canalicular cholestasis and minimal hepatocellular inflammation are present. Given the different presentation patterns of DIC, its management should not be limited to the discontinuation of a specific drug but should include genetic analysis to determine which bile transporters variants might be involved, in order to avoid relapses of DIC. The most associated disease-causing mutations and polymorphisms are in *ABCB11* and *ABCB4* genes, but it cannot be excluded that other genes responsible for PFIC are involved in DIC.

1.4.4 Low phospholipid-associated cholelithiasis

Low phospholipid-associated cholestasis and cholelithiasis (LPAC or Gallbladder Disease 1, GBD1, OMIM #600803) is a syndrome first described in 2001 by Rosmorduc et al. The disorder is caused by mutations in FIC genes, in particular in the *ABCB4* gene, which encodes the bile canalicular protein MDR3. In their initial observations, Rosmorduc et al. [76] described six patients (two men and four women, aged 26-55 years) with a peculiar form of intrahepatic cholelithiasis and a family history of gallstones. All six had a mutation in *ABCB4*, and all exhibited remarkable improvement on treatment with ursodeoxycholic acid. The exact prevalence of LPAC remains unknown. The disease is more common in young adults, the usual age at the onset of the symptoms is typically lower than 40 years. The male to female ratio is estimated at approximately 1:3. Up to 10% of the European and American population carry gallstones, approximately 25% of cases have symptoms and less than 2% present with severe complications (cholangitis or pancreatitis). Based on previous epidemiological data, it is considered that LPAC syndrome is infrequent and corresponds to a peculiar subgroup of patients with symptomatic gallstone disease. The LPAC is usually

associated with the intrahepatic biliary lithiasis and is often symptomatic after cholecystectomy and presents a good response to ursodeoxycholic acid. A reduced secretion of phospholipids in the bile decreases the solubility of cholesterol promoting its crystallization and the gall-stone formation.

At least two of the following criteria are necessary for the diagnosis of LPAC: (1) age at onset of biliary symptoms ≤40 years; (2) intrahepatic echogenic *foci* or microlithiasis; (3) recurrence of biliary symptoms after cholecystectomy. Severe biliary complications such as acute pancreatitis, recurrent cholangitis, segmental spindle-shape dilatation of the biliary tree filled with gallstones, and ICP may be observed in some patients [77]. The diagnosis should also be suspected when previous episodes of intrahepatic cholestasis of pregnancy and a family history of gallstones in first-degree relatives are present. The onset of the symptoms has been reported at the end of or soon after pregnancy. In fact, 56% of symptomatic women who are carriers of *ABCB4* variants also presented a history of ICP. The cholestasis results in an increase of GGT while abdominal ultrasonography, computing tomography scan and magnetic resonance cholangiopancreatography, help to evidence the intrahepatic gallstones (**table 3**). In a large cohort of 156 patients, Poupon et al. [78] discovered the *ABCB4* variants in 79 cases (63 heterozygous variants): similar features between subjects with and those without mutations were found and the authors hypothesized the possible involvement of unexplored genes. For example, the FXR regulates ABCB4 and ABCB11 expression and a down-regulation of this gene, due to the mutations or to the high levels of the progesterone metabolites, could promote defective bile salts and phospholipid secretion, leading to LPAC and ICP. While the homozygous variants cause PFIC3, the heterozygous variants lead to LPAC/ICP disorders. The differential diagnosis should be made versus inflammatory cholestatic disorders which may cause biliary symptoms and/or intrahepatic gallstones such as Caroli disease and primary sclerosing cholangitis. A definitive diagnosis is obtained by *ABCB4* gene analysis. A careful management is recommended and an indication to cholecystectomy should consider high rates of recurrence after surgery. Moreover, a long-term therapy with UDCA could be useful to prevent complications.

Non-progressive familial intrahepatic cholestasis				
	BRIC	ICP	DIC	LPAC
Locus/gene/protein	18q21-22/ATP8B1/FIC1 2q24/ABCB11/BSEP 18q21.1/MyosinVB/MYO5B	18q21-22/ATP8B1/FIC1 2q24/ABCB11/BSEP 7q21/ABCB4/MDR3 9q21.11/TJP2/Z0-2 12q23.1/NR1H4/FXR	2q24/ABCB11/BSEP 7q21/ABCB4/MDR3	7q21/ABCB4/MDR3
Clinics	Intermittent severe cholestasis (intervals weeks to years); hearing loss, pancreatitis, diarrhea	Transient cholestasis + itching during pregnancy; post-natal resolution; potentially serious fetal complications	Chronic liver injury; acute hepatitis; fulminant hepatic failure Use of herbal remedies and naturopathic substances should be investigated; Onset < 1-12 months by drug adiministration	<40y cholelithiasis; intrahepatic microlithiasis; recurrence of biliary symptoms after cholecystectomy; previous episodes of ICP; familial history of gallstones
Laboratory				
BA	High during attack	High during pregnancy	Normal or mild elevation	High during obstruction
GGT	Low or normal	Normal or mild elevation	Variable	High
ALP	High	Normal or mild elevation	High	Normal to high
AST/ALT	Normal or mild elevation	Normal or mild elevation	Moderate or severe elevation	Normal or mild elevation
AFP	No data	No data	No data	No data
Histology	Centrolobular cholestasis, no alterations of liver structure, no BSEP tissue expression	Not performed	Loss of BSEP/MDR3 expression, canalicular cholestasis, hepatocellular inflammation	Not required; imaging based diagnosis

Table 3. Laboratory, clinical, genetic and histological characteristics of different non-progressive familial intrahepatic cholestasis. BRIC: benign recurrent intrahepatic cholestasis; ICP: intrahepatic cholestasis of pregnancy; DIC: drug induced cholestasis; LPAC: low-phospholipid-associated cholelithiasis; BSEP: biliary salt export pump; MDR-3: class III multidrug resistance P-glycoprotein; TJP-2: tight junctions protein-2; zo-2: zona occludens-2; FXR: fanesoid X receptor; MYO5B: myosinVB protein; AST: aspartate aminotransferase; ALT: alanine aminotransferase; GGT: gamma-glutammyl transferase; ALP: alkaline phosphatase; AFP: alpha- 1-fetoprotein BA: bile acids. Vitale et al. *Digestive and Liver Disease* 2019.

1.5 Current therapeutic options

1.5.1 Medical therapy

In addition to the intake of fat-soluble vitamins to prevent deficiencies, two drugs can be used: ursodeoxycholic acid (UDCA) and rifampicin.

UDCA is safe and without major side effects and stimulates hepatobiliary secretion of bile salts and it can improve bile flow. Through a complex network of signalling pathways, it stimulates the impaired targeting of transport proteins as BSEP or the MRP2 conjugate export pump (protein 2 associated with multidrug resistance, coded from *ABCC2*) to the canalicular membrane by activation of a complex signalling network [61]. Moreover, UDCA conjugates can have antiapoptotic effect, contributing to the protection of hepatocytes. UDCA therapy is the first used drug for PFIC. It has been shown that UDCA is effective in more than half of the patients with PFIC3 (high GGT). In patients with low-GGT PFIC, the

response to UDCA therapy was less promising, but still resulted in improvement of serum transaminase levels and pruritus in some patients [79].

Rifampicin might work, at least in part, by upregulating detoxification enzymes and export pumps by mechanisms dependent of pregnane X receptor. In low-GGT PFIC patients, rifampicin treatment does not result in improvement of serum transaminases and bilirubin and reduces pruritus in only a few patients [61].

Recently, serotonin reuptake inhibitors were suggested for the management of refractory cholestatic pruritus in patients with PFIC or Alagille syndrome, resulting in an improvement of this symptom [80]. Since no severe adverse events were reported in this group, serotonin reuptake inhibitors like sertraline could be considered for uncontrolled pruritus.

Another non-surgical intervention is nasobiliary drainage, a temporary diversion of bile established by endoscopically introducing a nasobiliary drain [81]. Due to the temporary character, this therapy is especially useful to abort long-term cholestatic episodes in BRIC patients. The risk of procedure-related pancreatitis needs to be considered.

1.5.2 Surgical therapy

The purpose of partial biliary diversion (PBD) is to reduce the enterohepatic circulation of bile salts, thereby diminishing the accumulation of these compounds and preventing hepatic injury. Currently, PBD is therapy of choice in all non-cirrhotic children with low-GGT PFIC, in whom permanent cholestasis and intractable pruritus is present. This intervention was introduced in 1988, as partial external biliary diversion (PEBD) by cholecystojejunocutaneostomy [82]. This technique has the aim to reduce the reabsorption surface, directly secreting bile acids. In up to 85% of the non-cirrhotic patients, PEBD is successful in improving pruritus as well as biochemical parameters of cholestasis such as serum bile acids, liver enzymes and bilirubin and long-term results indicate that PEBD can delay or even reverse hepatic injury. However, the more severe disadvantage of this surgical intervention PEBD is that patients must live with a permanent stoma [80].

Liver transplantation is indicated for patients with advanced liver disease or persistent uncontrolled pruritus, in whom PBD have had no effect. Living-donor liver transplantation is also a safe and effective option. PFIC1 patients with ATP8B1 deficiency usually have extrahepatic manifestations, due to the widely expression of FIC1 protein. This situation can cause significant morbidity even after transplantation. Features such as pancreatitis, short stature and hearing loss will persist, and diarrhoea can be induced or exacerbate. ATP8B1 is highly expressed in the small intestine where it might have a role in the regulation of reabsorption of bile salts. After transplantation, normal amounts of bile salts are excreted by the hepatic graft while intestinal ATP8B1 is still impaired, resulting in high concentrations of bile salts in the colon and subsequent diarrhoea. Furthermore, significant hepatic steatosis often occurs after transplantation which can progress to cirrhosis and may require re-transplantation [80]. Liver transplantation usually gives complete correction of phenotype in patients with ABCB4 and ABCB11 deficiency. However, in some patients with severe ABCB11 deficiency, recurrence has been observed post liver transplantation as a result of the formation of autoantibodies against BSEP.

1.6 Future therapeutic options

1.6.1 Total biliary diversion

In some patients, PBD is not effective in reducing clinical symptoms and improving biochemical parameters. For these patients, total biliary diversion (TBD) might be useful. In four patients with ATP8B1 deficiency, of which three did not benefit from PEBD, total external biliary diversion resulted in a significant reduction or complete disappearance of pruritus and cholestasis [83]. No clinical signs of fat malabsorption were encountered, although fat soluble vitamin levels turned out to be more difficult to manage than in PBD. Currently, TBD is not recommended as primary therapy since PBD is already effective in most of the patients and the experience with TBD is limited to only few patients. Moreover, the long-term effects of total absence of intestinal bile salts are unknown.

Recently, biliary diversion has also been used in combination with liver transplantation to improve diarrhoea and nutritional status in two patients with severe ATP8B1 deficiency, which proved to be effective in terms of clinical, histological and biochemical outcomes [84].

1.6.2 Hepatocyte transplantation

Given the shortage of donor organs, hepatocyte transplantation might become an alternative to liver transplantation. Advantages of hepatocyte transplantation are that the procedure is less invasive, can be repeated several times and leaves the native liver *in situ*. The procedure may delay or even eliminate the liver transplantation. However, scarcity of donor hepatocytes again is the limiting factor. In addition, immunosuppression with its associated morbidity is still necessary. Therefore, alternative cell sources are being investigated, such as induced pluripotent stem cells (iPSCs) derived hepatocyte-like cells or hepatocytes generated by trans-differentiation. Takebe et al. succeeded in generating vascularized and functional human liver from iPSCs by transplantation of three-dimensional liver buds created *in vitro* [85]. However, the hepatocyte-like cells were not fully differentiated (absence of cholangiocytes), as evidenced by lower albumin secretion and lower expression of hepatocyte specific CYP450 enzymes compared to primary human hepatocytes.

A promising alternative approach is the clonal expansion of single Lgr 5+ bipotent liver progenitor cells into transplantable liver organoids [86]. A three-dimensional culture system allows long-term clonal expansion of single Lgr5(+) stem cells into transplantable organoids (budding cysts) that retain many characteristics of the original epithelial architecture. In contrast to iPSC derived hepatocytes or liver buds, the organoid cells are obtained directly from the liver without the need for genetic modification or introduction of reprogramming factors, diminishing the risk of malignancies. Although several issues need to be addressed, these techniques might enable personalized autologous hepatocyte transplantation in the future.

1.6.3 Gene therapy

Gene therapy corrects the defective gene responsible for disease development and can be applied to different tissues. However, multiple barriers are still present, including the vector genome and transcriptional persistence as well as the immune response that can limit the viability of transduced cells [80]. In addition, genes like ATP8B1 and ABCB11 are large and therefore difficult to insert into conventional viral vectors. Alternatively, autologous transplant with *ex vivo* genetically corrected stem cells or hepatocyte-like cells generated from patient derived iPSCs, may hold promise of success. Because liver organoids can be based on the *in vitro* expansion of a single progenitor cell, these might also be suitable for genetic modifications. Also *in vivo* gene editing using the bacterial CRISPR/Cas system might become an option in the future [87].

1.7 Sequencing analysis

1.7.1 Next-generation sequencing

The development of Next-Generation Sequencing (NGS) technology has enormously increased the possibility of genome analysis, having the ability to perform numerous measurements simultaneously. In fact, all the NGS platforms available today present a common technological feature, the massive parallel sequencing of cloned amplified DNA molecules. Sequencing in NGS technologies is performed by repeated cycles of nucleotide extensions or by iterative oligonucleotide ligation cycles. Considering the ability of these platforms to work in parallel, these tools have the possibility of sequencing from hundreds of Mb (millions of base pairs) up to Gb (billions of base pairs) of DNA in a single analytical session, according to type of NGS technology used [88].

NGS technologies involve a series of subsequent steps (§ **2.4**) and the unique combination of specific protocols distinguishes one technology from the others and determines the type of data produced by each platform. In this PhD project Ion Torrent platform was used.

1.7.2 Ion Torrent technology

The Ion Torrent platform uses semiconductor technology: the heart of the platform is represented by a semiconductor chip containing an array of millions of ISFET ion sensor positioned under as many micro-wells (an ISFET is an ion-sensitive field-effect transistor, that is a field-effect transistor used for measuring ion concentrations in solution). Each micro-well can accommodate only a single sphere (bead) containing multiple copies of the same single-strand DNA (ssDNA). The principle behind platform is the sequencing-by-synthesis. During the synthesis, the incorporation of a nucleotide complementary to the target sequence by the polymerase produces the release of two waste products: the pyrophosphate (PP_i) and hydrogen ions (H^+). A microwell containing a template DNA strand to be sequenced is sequentially flooded with a single species of deoxyribonucleotide triphosphate (dATP, dTTP, dCTP, or dGTP). A dNTP will only be incorporated if it is complementary to the leading unpaired template nucleotide. Whenever a single dNTP is added to the growing chain of the ssDNA template, pH-sensitive semiconductors below each well detect the increase in protons (**figure 9**). The pH lowering, caused by the increase in H^+ ions generates a current increase in the sensitive surface that is converted into a potential variation [89]. After the chip is flooded with a single type of dNTP, the signal is recorded and follows a wash that restores the initial pH conditions. This step is necessary to allow a new detection when adding the next dNTP.

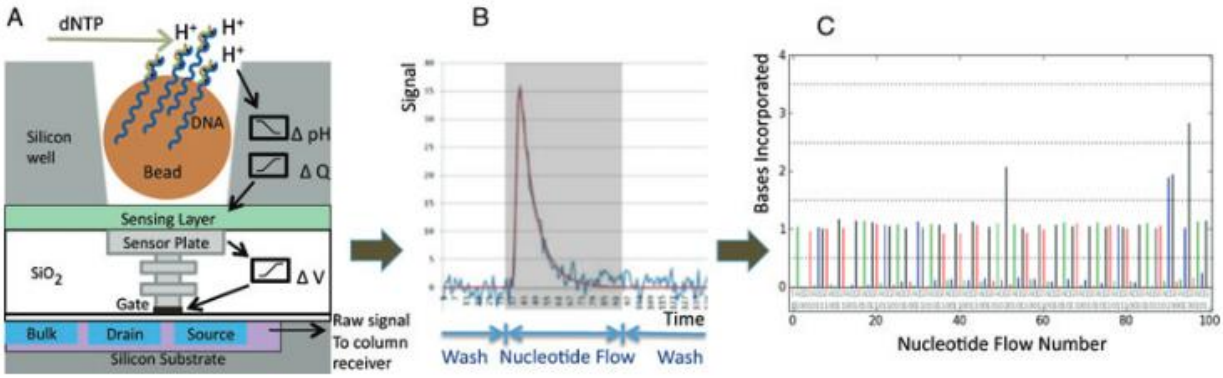


Figure 9. Overview of sequencing using a semiconductor pH sensor device. (A) The sequencing chemistry configuration, with templated DNA bead (brown sphere) deposited in a well above the ISFET sensor. The clonally templated bead is single-stranded, primed for extension and bound with polymerase. After dNTP incorporation, H^+ ions are released. The resulting H^+ release (Delta pH) is converted into a +charge buildup (ΔQ) on the sensing layer, which the conductive "via" metal pillar transmits as a voltage change (ΔV) at the transistor gate. This in turn gates the transistor current, and the time-accumulated current signal is sent off-chip, in the form of a voltage signal, via a chip-wide polling process. (B) These pH signals are sampled from the well at high frequency (blue curve), e.g. 100 Hz, and then fit to a model that extracts the net signal (red curve). (C) The net incorporation signals for each flow become the primary input, reflecting no incorporation, 1-mer incorporation, 2-mer incorporation, etc. for each trial flow, and these raw data are the input for the base calling algorithm that converts the full series of net incorporation signals for that well, over hundreds of trial nucleotide flows, into one corresponding sequence read (Merriman et al. *Electrophoresis* 2012).

In this way, the reconstructed sequence will be the result of positive or null signals each addition of dNTP. The received signals are shown graphically (**figure 9c**), with the type of dNTP added to the chip in a sequential manner in the abscissa and in ordinate the quantity of signal recorded at each incorporation (base calling). In case of homopolymers, in fact, there will be more incorporations of the same dNTP and therefore a greater pH variation will be recorded (height of each bar).

The number of sensors in the chips determines its sequencing capability and, therefore, the number of reads reproducible at each analytical session. Increasing sequencing capabilities can be achieved by increasing the chip area and decreasing the distance between various sensor arrays.

In this study we used two different instruments: Ion PGM™ and Ion GeneStudio™ S5 System (Thermo Fisher Scientific, Waltham, MA, USA, US). Each instrument has its specific chips. For Ion PGM system there are three different chips: 314, 316 and 318. For our purpose

we adopted the last one, a chip characterized by an output of 1 Gb (Ion 318), implementing several times the sequencing capability of the first chip.

The Ion GeneStudio S5 systems is the evolution of the Ion PGM and is designed to enable a broad range of targeted next-generation sequencing (NGS) applications with speed and scalability. Five Ion S5 chips enable a sequencing throughput range of 2M to 130M reads per run. The new Ion 550 Chip generates 100–130 million sequencing reads on the Ion GeneStudio S5 Plus and Prime Systems and it is not compatible with Ion GeneStudio S5 System. For the second part of the PhD project, when we updated our NGS panel with 15 genes **(§ 2.10)**, we had to use the Ion 520™ chip, with up to 8 million reads per run.

1.8 Project aim

The goal of this project is to develop and validate a broad, reliable, rapid and cost-saving genetic test for PFIC patients, based on NGS technology, with a sensitivity comparable with Sanger sequencing.

Up to date, Sanger sequencing represent the most used method in molecular diagnostics. However, considering the new large panel size, the probability to have also other genes involved in PFIC and the molecular complexity of the disease, like high allele heterogeneity and changes in normal autosomal recessive hereditary manner, Sanger method has become not suitable for these complex diseases, technically challenging, labour intensive and costly. This innovative test may be useful for the molecular diagnostics of PFIC and a better characterization and understanding the linking between molecular defects and different subtypes of the disease. NGS method guarantees to reach quickly medical reports for all patients and their related, so should have great consequences on disease management, patient treatment and national medical system.

Italian adult population data lack clinical significance of mutational profiles of the four genes associated to PFIC so far. With this purpose we wanted to perform a screening for these genes and assess the role of genetic contributions in an adult cohort, to show that

heterozygous mutations are mainly linked to a later onset of the disease or to patients with cryptogenic cholestasis.

In addition, another aim was to find potential alterations on other eleven genes (fifteen-gene panel) probably involved in cholestasis and to evaluate their impact on the disease.

Lastly, the integration of standard variant prediction and classification tools with more advanced bioinformatic analyses, to get a deeper understanding on the role of genetic variation from large NGS datasets.

SECTION II: MATERIALS AND METHODS

2.1 Patients recruitment

Patients to be included in the project were selected by the S. Orsola-Malpighi Policlinic clinicians who mostly deal with cholestatic disease. The groups are mainly: Prof. Fabio Piscaglia, dr. Laura Gramantieri, dr. Giovanni Vitale and Prof. Maria Cristina Morelli from Internal Medicine Operational Unit, Prof. Pietro Andreone from Internal Metabolic Nutritional Medicine of the Policlinic of Modena.

In the enrolled patients, other causes of cholestasis that cannot be ascribed to the PFIC phenotype were excluded, such as primary biliary cholangitis, primitive sclerosing cholangitis, viral hepatitis, alcohol abuse, hemochromatosis, Wilson disease and alpha-1-antitrypsin deficiency. Instead, patients presenting the classic forms of PFIC were included, but also BRIC, LPAC, ICP and DIC phenotypes. During the three-year PhD project, from December 2016 to August 2019, 176 patients were recruited. All patients signed a specific informed consent for genetic analysis for diagnostic purposes.

2.2 DNA extraction and normalization

The genomic DNA of the patients to be analysed was extracted from peripheral blood lymphocytes in EDTA (Ethylenediaminetetraacetic Acid) using the Maxwell® 16 Instrument automatic extractor and the related Maxwell® 16 DNA Purification kit (Promega). The automatic extractor is prepared for the simultaneous extraction of genomic DNA of 16 samples in just over 30 minutes, using a system of magnetic beads (MagneSil® Paramagnetic Particles) to capture DNA. The specific purification kit consists of disposable cartridges, divided into various sections containing the necessary reagents, from disposable plungers to cuvettes in which the extracted samples are collected (**figure 10**).

The extraction process develops through successive phases, each taking place at a specific section of the cartridge. The first step is represented by the cell lysis, which takes place partly chemically and mechanically. In the second section of the cartridge, magnetic spheres can bind nucleic acids. At this point the plunger is loaded with marbles and then reinserted into

the first compartment of the cartridge, where it collects the DNA. After, DNA undergoes to purification steps. Finally, the DNA is eluted in a cuvette containing 300 µl of Milli-Q® water (deionized and sterilized).

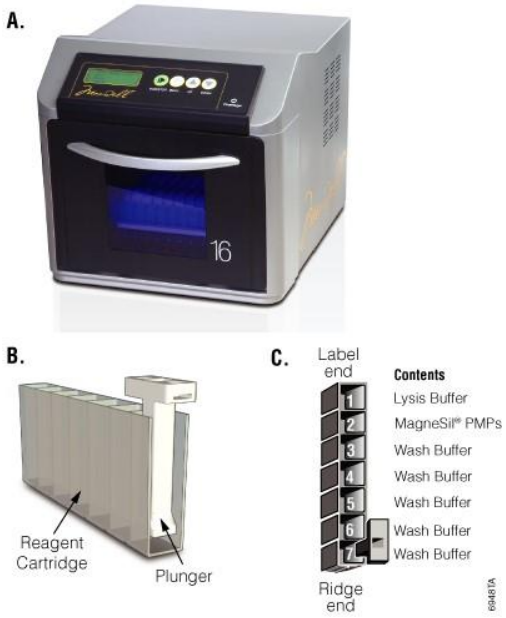


Figure 10. Maxwell® 16 Extractor. A. Instrument. **B.** Disposable cartridge with plunger facing the operator. C. Seen from the top of the cartridge, in position 1, where the lysis buffer is located, the blood sample is loaded, in the well number 7 the eluted DNA will be collected after about 30 minutes.

The extracted DNA samples were quantified using the Nanodrop 3.0.0 spectrophotometer (Celbio S.P.A. Milan, Italy). This instrument, using 1 µl of DNA, allows to determine the concentration (reading at 260 nm wavelength) and the purity (reading at 280 nm wavelength) of the DNA. Once the DNA concentration was measured, each sample was normalized to a concentration of 10 ng/µl.

2.3 Experimental design

The Ion AmpliSeq™ Designer is a free and available online tool created for Ion Torrent users. By entering the chromosomal coordinates and the name of the target genes, the software provides a primer list divided in two primer tubes for multiplex-PCR amplification (for panel size up to 50 genes). All theoretical data related to the quality are also provided by the design. At the beginning of PhD project, the first design has been created including

the four genes involved in PFIC: *ATP8B1*, *ABCB11*, *ABCB4* and *TJP2*. The design was conceived to cover the entire coding region of the four genes and 50 base pairs flanking the beginning and end of each exon, to be able to identify also intronic variants potentially implicated in the splicing process. A real coverage of 99% was obtained with 194 amplicons for a total panel size of 42 kb. With this four-genes panel we analysed 96 patients.

In the second year of the project we decided to expand our gene panel because of new genes probably associated to the disease by the literature. In scientific literature we found a considerable quantitative and qualitative diversity of genetic panels analysed between the different groups and the data were quite conflicting. Therefore, considering also the clinical and phenotypic data of our patient pool we decided to include other eleven genes with a biological rationale:

the *ABCC2* gene, coding for a protein member of the superfamily of ATP-binding cassette (ABC) transporters and expressed in the canalicular (apical) part of the hepatocyte and functions in biliary transport. Several different mutations in this gene have been observed in patients with Dubin-Johnson syndrome (DJS), an autosomal recessive disorder characterized by conjugated hyperbilirubinemia. The *CLDN1* gene, which encodes the Claudin-1 protein, is part of the tight junction complex and causes sclerosing-neonatal ichthyosis (NISCH syndrome); *GPBAR1*, which codes for the G Protein-Coupled Bile Acid Receptor 1, a cell membrane receptor for bile acids associated with sclerosing cholangitis; *NR1H4*, associated with progressive familial intrahepatic cholestasis 5 (PFIC5, ¶ **1.3.5**); *JAG1*, (involved in Notch signalling) and *NOTCH2* (Jagged-1 protein receptor) whose mutations cause Alagille syndrome; *MYO5B*, associated with Myo5b-Related PFIC (¶ **1.3.6**); *SLC25A13*, associated with citrullinemia; *SLCO1B1* and *SLCO1B3* (Solute Carrier Organic Anion Transporter Family, associated with hyperbilirubinemia and ICP; *VPS33B* associated with PFIC1 phenotype.

The new panel included (and still includes today) 15 genes: *ATP8B1*, *ABCB11*, *ABCB4*, *TJP2*, *NR1H4*, *MYO5B*, *JAG1*, *ABCC2*, *NOTCH2*, *CLDN1*, *GPBAR1*, *SLC25A13*, *SLCO1B1*, *SLCO1B3* and *VPS33B*, for a total of 436 amplicons and 85 kb, with a total coverage of 99.9% of the target. Considering the new panel size we had to upgrade the equipment used, in

this regard we started to use Ion S5 GeneStudio (Thermo Fisher Scientific, Waltham, MA, USA), available for sharing with the departments of Pathological Anatomy and Medical Genetics of the Sant'Orsola-Malpighi Hospital (¶ **2.10**).

2.4 Workflow

2.4.1 Library construction

Libraries were prepared using the recommended Ion AmpliSeq™ Library Kit 2.0 (Thermo Fisher Scientific, Waltham, MA, USA) which includes a five-step structured workflow:

2.4.1.1 Amplification

The amplification step was performed by setting up a multiplex-PCR. The reaction mix and amplification conditions are shown in the tables below and are intended for single sample, amplified for each pool.

Reagents	**Volume (µl)**
5X Ion AmpliSeq™ HiFi Master Mix	2
2X Ion AmpliSeq™ Primer Pool	5
gDNA (10 ng)	1
Nuclease-free H_2O	2
Final volume	**10**

Temperature	**Time**	**Cycles**
99 °C	2′	1
99 °C	15″	19
60 °C	4′	
10 °C	Hold	End

2.4.1.2 Primer digestion

Following amplification, 1 µl of FuPa Reagent was added to each sample, allowing partial digestion of primers and phosphorylation of amplicons. Samples were then incubated in thermal cycler under conditions shown in the table.

Temperature	Time
50 °C	10′
55 °C	10′
60 °C	20′
10 °C	Hold*

* for up to 1 hour

2.4.1.3 Adapters and barcode binding and purification

In this step the reaction mix shown in the table below was added to each sample, amplified with both pools, and samples were incubated in thermal cycler, according to the described conditions.

Reagents	Volume (µl)
Switch solution	2
Ion P1 adapter	1
Ion Xpress™ barcode	1
DNA ligase	1
Nuclease-free H_2O	1
Final volume for each sample	**6**

Temperature	Time
22 °C	30′
72 °C	10′
10 °C	Hold*

* for up to 1 hour

For purification phase, special magnetic spheres that select fragments greater than 100 bp have been used. The processed samples were transferred into a 1.5 ml Eppendorf, resuspended with Agencourt AMPure XP Reagent 1.8X and incubated at room temperature for 5 minutes. After incubation, the Eppendorf were placed in the special DynaMag™ 2 magnetic rack (**figure 11**) for about 3 minutes. In this way, the magnetic beads that have bound the 175-275 bp fragments are arranged along the wall of the Eppendorf, allowing the removal of the supernatant. Subsequently, three washes were carried out with 70% ethanol (freshly prepared). The Eppendorf were removed from magnetic rack, pellet was resuspended in 25 µl of Low TE and incubated at room temperature for about two minutes, to allow the fragments to detach from beads. Finally, the Eppendorf were again placed in

the magnetic rack and the supernatant, this time containing only the fragments of interest, was transferred into a 0.2 ml micro-tube.

Figure 11. Magnetic rack. On the left DynaMag-2 Magnet compatible with 1.5 ml Eppendorf. On the right DynaMag -96 Side Magnet, compatible with 0.2 ml micro-tube, in strip or singles (Thermo Fisher Scientific, Waltham, MA, USA).

2.4.1.4 Library quantitation

Libraries obtained from previous steps were quantified by Real Time-PCR. The quantitation was carried out using the Ion Library TaqMan™ Quantitation Kit (Applied Biosystems by Thermo Fisher Scientific, Waltham, MA, USA). The assay allows to obtain a quantity measurement of fragments to which adapters have been bound. This is possible because of marked probes are complementary to their sequences. For this purpose, serial dilutions of a known stock concentration (68pM) of control library, *E. coli* DH10B Ion Control Library (6.8 pM, 0.68 pM, 0.068 pM and 0.0068 pM), have been prepared. This step is fundamental in order to construct a calibration curve on which the concentrations of libraries are determined (previously diluted 1:500). For qRT-PCR assay, the reaction mix reported in the following table was used and samples were incubated in the StepOne™ Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA, USA) under reported conditions.

Reagents	Volume (µl)
2X Ion TaqMan™ Master Mix	5
20X Ion TaqMan™ Assay	0.5
Samples*	**Volume (µl)**
E. coli serial dilutions	4.5
Diluted library (1:500)	4.5
Final volume for each sample	**10**

* 8 wells for 4 *E. coli* different concentrations and *x**2 wells for *x* samples (each in double)

Temperature	Time	Cycles
50 °C	2′	Hold
95 °C	20″	Hold
95 °C	1″	40**
60 °C	20″	

** at last cycle real time will automatically end

2.4.2 Library pooling

After quantitation, libraries of both pools were combined into a single 0.2 ml reaction tube at equimolar concentration (9 pM).

2.4.3 Template preparation

2.4.3.1 Emulsion PCR

For this step the Ion PGM™ Hi-Q™ View OT2 kit (Thermo Fisher Scientific, Waltham, MA, USA) was used. The preparation of template was performed using emulsion-PCR technique. Pool fragments of all libraries have been mixed with an excess of spheres (Ion Sphere™ Particles), so that each amplicon binds to the surface of a sphere. The library dilution therefore represents the key step to avoid, or at least minimize, the production of polyclonal spheres.

The emulsion was created by properly mixing the PCR reagents (shown in the table below) with mineral oil, in a special reaction filter. The filter was immediately loaded on the Ion OneTouch™ 2 System instrument (Thermo Fisher Scientific, Waltham, MA, USA) (**figure 12**), which contains a heating block with the amplification plate inserted in and produces a temperature gradient from 95 ° C to 64 ° C. The plate contains a channel that allows the reaction mix to move and allows the temperature changes needed for the reaction.

To a 2-mL tube containing 800 µL of Ion PGM™ Hi-Q™ View Reagent Mix, add the following components in the designated order:

Reagents	**Volume (µl)**
Nuclease-free H_2O	25
Ion PGM™ Hi-Q™ View Enzyme Mix	50
Diluted library	25
Ion PGM™ Hi-Q™ View ISPs	100
Final volume	**1000**

Figura 12. Ion OneTouch™ 2 System. (Thermo Fisher Scientific, Waltham, MA, USA).

2.4.3.2 Quality control

The quality control of emu-PCR is carried out with the Ion Sphere™ Quality Control Kit, which involves two fluorochromes: one binds the spheres (Alexa Fluor 488 dye) and the other binds the external adapter (Alexa Fluor 647 dye), therefore binding together the template. Samples are prepared in special 0.2 ml tubes and incubated in the thermocycler under the reported conditions.

Reagents	**Volume (µl)**
Annealing Buffer	19
Ion Probe™	1
Unenriched sample	2
Final volume for each sample	**22**

Temperature	**Time**
95 °C	2′
37 °C	2′

2.4.3.3 Enrichment

The enrichment of the emulsion-PCR product allows the selection of the spheres that have bound the template and the removal of the empty spheres but does not allow to exclude the polyclonal ones. This phase exploits the action of streptavidin, which recognizes and binds the amplicon bound to the sphere. In this step, the Ion PGM™ Enrichment Beads kit and the Ion OneTouch™ ES instrument were used to automate the process. The instrument uses a mechanical arm equipped with a tip that moves along a special 8-well-strip filled as follows (**figure 13**).

Well number	Reagent to dispense in well
Well 1[1]	*Entire* template-positive ISP sample [100 µL; prepared in step 1 of this procedure]
Well 2	130 µL of Dynabeads™ MyOne™ Streptavidin C1 Beads resuspended in MyOne™ Beads Wash Solution] prepared in "Wash and resuspend the Dynabeads™ MyOne™ Streptavidin C1 Beads" on page 5]
Well 3	300 µL of Ion OneTouch™ Wash Solution
Well 4	300 µL of Ion OneTouch™ Wash Solution
Well 5	300 µL of Ion OneTouch™ Wash Solution
Well 6	Empty
Well 7	300 µL of freshly-prepared Melt-Off Solution [prepared in "Prepare Melt-Off Solution" on page 5]
Well 8	Empty

[1] Well closest to the square-shaped tab

Figure 13. Ion OneTouch™ System. On the left strip preparation (numbers in the list correspond to the number of the strip well). **B.** strip into the Ion OneTouch ES (Ion PGM™ Hi-Q™ OT2 Kit *Quick Reference*).

2.5 Chip loading

Final sample is released into a 0.2 ml tube, previously placed in a special housing in the OneTouch™ ES. Five microliters of Control ISPs from Ion PGM™ Hi-Q™ View Sequencing Kit were added. The sample was centrifuged at 15500 rpm for 2 minutes and the supernatant was removed. Subsequently, 12 µl of Sequencing Primer are added and the tube is incubated in the thermocycler, according to the conditions reported in the table.

Temperature	Time	Cycles
95 °C	2′	1
37 °C	2′	1

The thermal-cycler incubation allows the annealing of sequencing primers to the template attached to the spheres. After incubation, 3 μL of Ion PGM™ Hi-Q™ View Sequencing Polymerase was added to the sample.

Several sequencing sessions were performed for this project. At the beginning of the project we used Chip 318 v.2 and Ion Torrent PGM sequencer; later (see below) we used Ion GeneStudio S5 and Ion Chef. By the Pipet-Lite XLS+ pipette, enriched pool was inserted inside the chip (through the loading port) carefully, avoiding the introduction of bubbles. The chip was later inserted in a special mini centrifuge for 30 seconds. After, the chip is tilted by 45° and the liquid inside was resuspended. The subsequent centrifuges allow the spheres that have bound the template and the polymerase to settle inside the chip's micro-wells. These last steps were repeated 3 times. After the third centrifugation, the liquid is removed entirely from the chip and was loaded into the Ion Torrent PGM™ sequencer (**figure 14**).

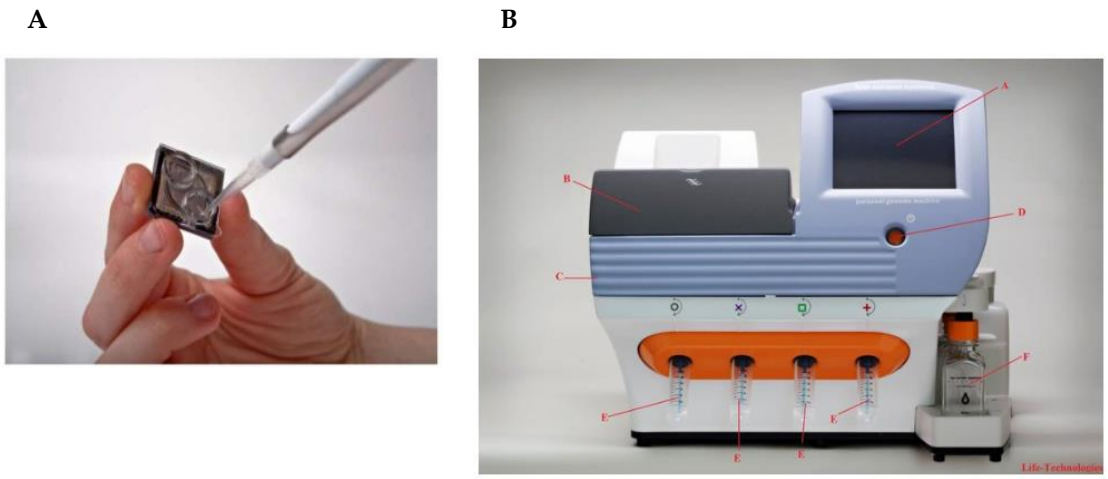


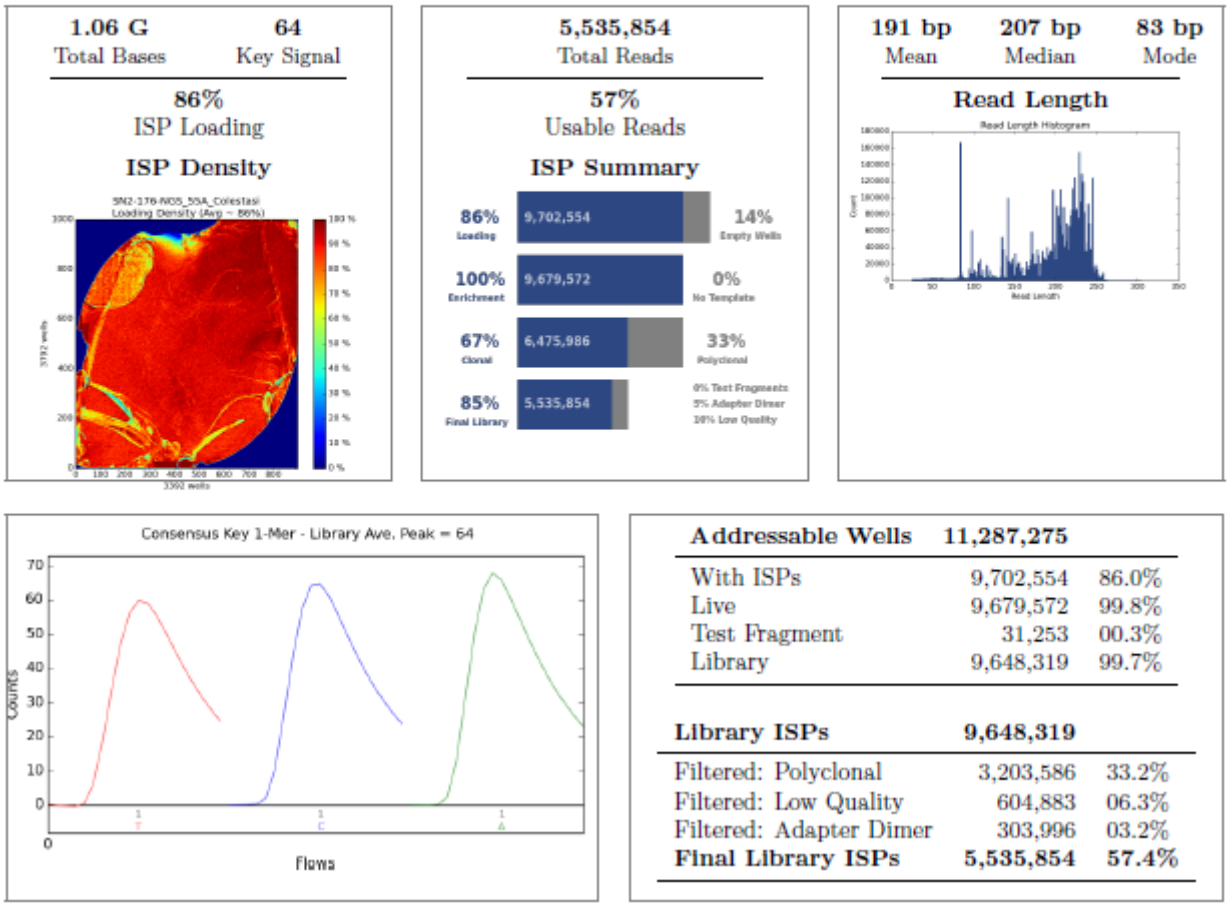
Figure 14. A. Chip loading. B. Ion-Torrent PGM sequencer. A) touch screen control, B) Ion-chip loading deck clamping mechanism, C) special material grounding plate, D) power button, E) Reagent bottles, F) Wash bottles (https://www.researchgate.net/publication/316668112).

After about half an hour from the beginning of the sequencing, a preliminary Run Report is available (**figure 15**). The report shows some details about the chip loading, the ISPs and reads quality. In the first box, at the top left, there is a photo of the chip loading, with the total number of bases produced (1.06 G in this case) and the percentage of micro-wells containing a live ISP with template attached (**figure 15a**). In the central box, at the top, the total number of reads, the percentage of empty wells and the percentage of wells that contain an ISP with bounded template (enrichment = 100%) are described. The 67% of clonal spheres indicates the spheres that have tied a clone coming from a single template, so they possess sequences that has the same origin of one template; the remaining 33% indicate the polyclonal spheres. In the box at the top right mean, median, mode and read length graph are shown. **figure 15b** shows the list of patients with respective barcodes and data of the read's quality.

A

Run Report for Auto_user_SN2-176-NGS_55A_Colestasi_154

Run Summary



Addressable Wells	11,287,275	
With ISPs	9,702,554	86.0%
Live	9,679,572	99.8%
Test Fragment	31,253	00.3%
Library	9,648,319	99.7%

Library ISPs	9,648,319	
Filtered: Polyclonal	3,203,586	33.2%
Filtered: Low Quality	604,883	06.3%
Filtered: Adapter Dimer	303,996	03.2%
Final Library ISPs	**5,535,854**	**57.4%**

B

Barcode Name	Sample	Bases	$\geq Q20$	Reads	Mean Read Length	Read Length Histogram
IonXpress_005	17692	58,108,598	54,450,394	311,037	186 bp	
IonXpress_006	17700	69,107,827	65,133,046	360,227	191 bp	
IonXpress_007	17718	40,468,501	38,170,429	211,284	191 bp	

Figure 15. Run report from one of our sequencing runs. A. Summary report in PDF, downloaded from Ion Torrent server (see text for details). **B. Samples details.** Output file showing barcode name, sample name, bases sequenced, reads exceeding quality Q20 cut-off, number and mean of the reads and a length graph.

2.6 NGS data analysis

Ion PGM™ Sequencer transfers row sequencing data to the Torrent Server, which converts raw data to base calls and the Torrent Suite™ generates the summary sequencing report and Fastq files. Reads are aligned to human genome (hg19 version) and Binary Alignment Map (BAM) in conjunction with Binary Alignment Index (BAI) and Variant Call Format (VCF) files are produced.

It is possible to visualize reads aligned of BAM/BAI files using Integrative Genome Viewer (IGV) Software. IGV is currently used to assess the coverage depth of the sequencing reads, zygosity, quality of the sequencing reads and mapping quality.

All the VCF files are uploaded into the Ion Reporter software (https://ionreporter.thermofisher.com/ir/), selecting the Annotation Variant workflow in order to associate each variant identified to the nucleotide change in cDNA and mRNA transcript, the aminoacidic change, the exons or IVS and its function. Moreover, the Ion Reporter can readily access to several prediction programs used to evaluating the pathogenic potential of missense change, providing important information about the pathogenicity of the variants analysed (buying license).

2.6.1 Variant filtering strategy

The read files obtained from sequencing were mapped to the GRCh37/hg19 assembly and the sequence variants were identified by Variant Caller and Ion Reporter software.

The filtering step of sequence variants was carried out according to default criteria for germline mutations: insertions or deletions, non-sense variants, splicing-site variants, and missense with minor allele frequency (MAF) ≤ 0.05 were considered. Three bioinformatics tools were used to predict the role of missense variants. Sorting Intolerant From Tolerant (SIFT) (http://www.jcvi.org/cms/home/), PolyPhen-2 and Mutation Taster software can evaluate whether an amino-acid substitution influences protein structure and function, according to physical modifications and the degree of conservation of protein sequence

among species. Variants that meet at least one of the following criteria were also considered: ever described before, reported in Human Gene Mutation Database (HGMD®), predicted not benign by at least one of the bioinformatic tools.

All the variant found were classified following The American College of Medical Genetics and Genomics 2015 (Richards et al. [90]) in five principal categories:

1. **Pathogenic** (**P**): all the Loss of Function (LOF) mutations as nonsense, frameshift, indel non-frameshift, canonical splicing variants, and the SNP occurring into the ATG starting codon;
2. **Likely Pathogenic** (**LP**): all the missense described as pathogenic variants listed in HGMD® (Human Gene Mutation Database). Novel missense variants with a Minor Allele Frequency (MAF) < 0,01 on free online databases like gnomAD (Genome Aggregation Database, https://gnomad.broadinstitute.org/), predicted as damaging or not tolerated by main bioinformatic tools as PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT sequence (Sort Intolerant From Tolerant) (https://sift.bii.a-star.edu.sg/www/SIFT_seq_submit2.html);
3. **Variant of Uncertain Significance** (**VUS**): all novel missense rare variants with discordant prediction results or where poor data are available.
4. **Likely Benign** (**LB**): all frequent missense variants and rare synonymous and intronic (near the canonical splicing site) variants predicted to not affect function by Human Splicing finder (http://www.umd.be/HSF3/);
5. **Benign** (**B**): all the deep intronic variants and frequent synonyms variant in non-conserved sequence domains.

Variants passing the above filtering process were mapped to Uniprot canonical sequences and subjected to Mechismo analysis to predict their functional consequences at biomolecular interaction interfaces. The approach matches protein sequence aminoacids to positions within structures and identifies sites affecting interactions with other proteins, DNA/RNA or small molecules. We considered low confidence predictions including known

structures or close (C 30% sequence identity) homologs and only very confident, physical protein–protein interactions (as defined by Mechismo based on a benchmark for the accuracy of perturbed interfaces).

We annotated mutations and PTMs (phosphorylations and acetylations) from Phosphosite (https://www.phosphosite.org/homeAction.action) of the considered genes on corresponding Uniprot canonical sequences using lollipop diagrams.

2.7 Statistical analysis

Categorical and quantitative variables were used to perform statistical analysis. Quantitative variables were Body mass index (BMI), level of liver fibrosis (based on Fibroscan® scores), laboratory features like alkaline phosphatase (AP), γ-glutamyltransferase (GGT), aspartate and alanine aminotransferases (AST and ALT), total and direct bilirubin and albumin levels. Presence of P/LP mutations and number of variants per patient were also considered variables.

Continuous variables were also categorized according to established pathological cut-offs.

For each continuous parameter the normality of the distribution was checked by the Kolmogorov-Smirnov and Shapiro-Wilk tests, then parametric or non-parametric test were used accordingly to assess differences between genotypes.

Contingency tables were developed by matrices of frequency distribution of the variables and used to compare categorical variables. while for quantitative variables the one-way analysis of variance (ANOVA) was used.

Binary logistic regression was performed for univariate analysis to identify alleles with potential risk factors. A $p < 0.05$ was considered significant for all tests.

About non-parametric tests, independent samples Mann-Whitney U (2 samples) test and Kruskal-Wallis one-way ANOVA (k samples) test were performed.

The IBM SPSS® statistical software, version 21.0 (SPSS® Inc., Chicago, IL, USA) was used for statistical analyses.

Allele frequency (AF) of common single-nucleotide polymorphisms (SNPs) was matched with data reported in the 1000 Genomes Browser and compared with TSI population (Toscani in Italy) (https://www.internationalgenome.org/1000-genomes-browsers/) by using de Finetti test. This analysis is part of the case-control studies and tests the deviation from the Hardy-Weinberg principle (HWP) performing Pearson's Chi-squared test by using the observed genotype frequencies obtained from the data and the expected genotype frequencies obtained using the HWP. The output of the de Finetti test shows also the allele and genotype associations with the disease, comparing alleles and genotypes frequency between controls and patients [91], as shown in the **table 4**. Still, a $p < 0.05$ was considered statistically significant in this analysis.

Table 4. The de Finetti output. HW equilibrium and association for *SLCO1B3* c.699G>A (p.M233I), rs7311358. In this association test significant *p* and Odds ratio are shown in bold.

		Risk allele 2				
n11=82 (83.46) n12=25 (22.08) n22=0 (1.46) f_a1=0.88 +/-0.020 F=-0.13228 p=0.171229 (Pearson) p=0.068584 (Llr) p=0.352367 (Exact)	n11=55 (52.00) n12=19 (24.99) n22=6 (3.00) f_a1=0.81 +/-0.035 F=0.23981 **p=0.031959** **(Pearson)** p=0.044187 (Llr) p=0.063375 (Exact)	**[1]<->[2]**	**[11]<->[12]**	**[11+]<->[22]**	**[11]<->[12+22]**	**common odds ratio**
		Odds_ratio=1.817 C.I.=[1.025-3.220] chi2=4.26 **p=0.03912 (P)**	Odds_ratio=1.133 C.I.=[0.570-2.253] chi2=0.13 p=0.72157	Odds_ratio=19.324 C.I.=[1.067-349.975] chi2=8.42 **p=0.00371**	Odds_ratio=1.491 C.I.=[0.777-2.859] chi2=1.45 p=0.22804	Odds_ratio=3.340 chi2=3.96 **p=0.04673**
		Risk allele 1				
		[2]<->[1]	**[22]<->[12]**	**[22]<->[11]**	**[11+12]<->[22]**	**common odds ratio**
		Odds_ratio=0.550 C.I.=[0.311-0.976] chi2=4.26 **p=0.03912 (P)**	Odds_ratio=0.059 C.I.=[0.003-1.109] chi2=6.82 **p=0.00902**	Odds_ratio=0.052 C.I.=[0.003-0.937] chi2=8.42 **p=0.00371**	Odds_ratio=0.053 C.I.=[0.003-0.961] chi2=8.29 **p=0.00398**	Odds_ratio=0.566 chi2=3.96 **p=0.04673**

2.8 Validation of NGS data by Sanger sequencing

Sanger sequencing was performed in order to validate all the variants classified as pathogenic, likely pathogenic and VUS. Primers have been designed for each of these variants. After amplification in thermal cycler, sequencing of purified amplicons was performed by ABI Prism 3730 DNA Analyzer (Applied Biosystem, Foster City, CA, USA), following the instruction reported above.

Reaction mix	Volume (μl)
Nuclease-free H_2O	4
Big-Dye™ Buffer (5X)	2
Big-Dye™ Terminator 3.1	1
Primers (3,2 pmol/μl)	1
Purified template	2
Final volume	**10**

Temperature	Time	Cycles
96 °C	30″	1
96 °C	10′	25
60 °C	3′	
10 °C	4	Hold

N.B. Mix and thermal cycler conditions for Sanger sequencing reaction. Previous amplification conditions are not shown, due to different and specific condition of each genomic segment in which variants has been identified.

2.9 MLPA analysis

MLPA® (Multiplex Ligation-dependent Probe Amplification, MRC-Holland, Amsterdam, NL) is a multiplex PCR method detecting large genomic rearrangement and copy number variations (CNVs). Commercial kit containing specific probe for one gene analysis are available by this company. This technique is not suitable for all genes. Thus, we have kit for the *ABCB4* analysis. Patients resulting wild type to NGS test and carrying typical PFIC3 phenotypes were choose by physicians for MLPA test. MLPA for *ABCB4* gene was performed by Salsa® MLPA® P109 ABCB4 probe mix.

The protocols consist of three different steps (not shown).

2.10 Ion GeneStudio S5 system

As soon as we started to perform new NGS-panel analysis with 15 genes (and a target region of 85 kb), we decided to upgrade it. So, when new equipment was available, we started to use the Ion S5 GeneStudio™ System (Thermo Fisher Scientific, Waltham, MA, USA), available for sharing with the departments of Pathological Anatomy and Medical Genetics of the Sant'Orsola-Malpighi Hospital.

The Ion GeneStudio™ S5 System is a semiconductor-based next-generation sequencing (NGS) system that enables simple targeted sequencing workflows exploiting the same

theory of its older relative PGM, but it is more efficient (**figure 16a**). The differences with Ion PGM are: high throughput to achieve from 2 to 80 million reads per run; reduce setup time and complexity with cartridge-based plug-and-play reagents; reduced time to complete run and analysis. A new chip series has been developed by the company to better fit with this instrument and to increment the number of reads produced per run. This evolution let us to analyse more samples in less time.

In the last six months we started to use the Ion Chef™ Instrument. The Ion Chef System is the next generation of workflow simplification products for the Ion GeneStudio™ S5 Systems. The Ion Chef System provides automated library preparation, template preparation, and chip loading. In less than 1 hour of up-front hands-on time and with the use of pre-packaged library preparation reagent kits, the Ion Chef System provides a convenient walk-away workflow resulting in equalized, pooled libraries ready for templating and producing one or two chips ready for sequencing (**figure 16b**). We realize that the better way to use the Ion Chef was from template preparation to the end, continuing to manually prepare libraries in order to have more cost savings.

A

B

Figure 16. A. Ion GeneStudioTM S5 Sequencer. B. Ion Chef System

SECTION III: RESULTS AND DISCUSSION

3.1 Assay analysis

The Next-Generation Sequencing technology, characterized by high throughput and convenient benchtop workflows, was used to develop a rapid and cost-effective molecular diagnostic method for PFIC.

The NGS platform used in this work was the Ion Torrent Personal Genome Machine™ (PGM) for the first two years of the project and Ion GeneStudio™ S5 for the last one (Thermo Fisher Scientific, Waltham, MA, USA), two highly suitable instruments for targeted-resequencing assay for DNA, RNA and small genomes.

For the purpose of this study Ion Ampliseq™ technology for the target-resequencing was developed. Through the Ion Ampliseq Designer™ software we designed two different NGS panels, as mentioned above (¶ **2.3**): the four-gene panel with *ATP8B1*, *ABCB11*, *ABCB4* and *TJP2*, for 194 overlapping amplicons and a target regions of about 48 kb (IAD61312_182); the fifteen-gene panel, including *ATP8B1*, *ABCB11*, *ABCB4*, *TJP2*, *NR1H4*, *MYO5B*, *JAG1*, *ABCC2*, *NOTCH2*, *CLDN1*, *GPBAR1*, *SLC25A13*, *SLCO1B1*, *SLCO1B3* and *VPS33B*, for a total of 436 amplicons and 85 kb, with a coverage of 99.9% of the target (IAD142012_1000) (**table 5**, **figure 17**). Those two panels characterized by overlapping amplicons divided in two pools each were used to allow us to achieve ultra-high multiplexing-PCR for the complete coverage of the entire coding sequences, i.e. exons and their flanking regions.

As regard the fifteen-gene panel, all the missed regions (74 bp in exon 2 of *NOTCH2*, 25 bp in exon 18 of *SLCO1B1* and 1 bp in exon 20 of *VPS33B*) were investigated by Sanger sequencing.

The library barcoding allowed the simultaneous sequencing of more patients per run. Despite the remarkable dimension of the fifteen-gene panel target region, an average coverage of about 1000X was obtained for the simultaneous sequencing of 15 patients on a 520 chip by Ion S5. For the four-genes panel an average coverage of over 800X was obtained submitting 12/14 patients in a single run on Ion PGM 318 chip.

The two NGS panels resulted both reliable in detecting variants, let us capable of obtain the same detection rate of other research groups. It is also true that Ion GeneStudio S5 was more

efficient than Ion PGM, resulting in an easier workflow with less probability to induce the operator into errors.

Gene	Gene Name	Score	Synonyms	Map Location	Exons	In-Silico Coverage	Gene Uniformity (Wet Lab)	Amplicons	Research Areas	HGNC ID	Ensembl ID	OMIM ID
ABCB11	ATP binding cassette subfamily B member 11	0.412	ABC16, ABCB11, BSEP, PFIC-2, PFIC2, PGY4, SPGP	2q24	27	100%	98.42%	43	Neoplasms, Nutritional and Metabolic Diseases, Hemic and Lymphatic Diseases,...	HGNC:42	ENSG00000073734	MIM:603201
ABCB4	ATP binding cassette subfamily B member 4	0.848	ABCB4, GBD1, MDR2, MDR3, PFIC-3, PGY3	7q21	27	100%	98.3%	33	Female Urogenital Diseases and Pregnancy Complications, Neoplasms, Nervous...	HGNC:45	ENSG00000005471	MIM:171060
ABCC2	ATP binding cassette subfamily C member 2	0.648	ABCC2, CMOAT, cMRP, DJS	10q24	32	100%	98.82%	38	Skin and Connective Tissue Diseases, Digestive System Diseases, Immune System...	HGNC:53	ENSG00000023839	MIM:601107
ATP8B1	ATPase phospholipid transporting 8B1	0.495	ATP8B1, ATPIC, BRIC, FIC1, PFIC1, PFIC	18q21	27	100%	99.8%	34	Female Urogenital Diseases and Pregnancy Complications, Digestive System...	HGNC:3706	ENSG00000081923	MIM:602397
CLDN1	claudin 1	0.615	CLDN1, ILVASC, SEMP1	3q28-q29	4	100%	100%	6	Female Urogenital Diseases and Pregnancy Complications, Neoplasms, Digestive...	HGNC:2032	ENSG00000163347	MIM:603718
GPBAR1	G protein-coupled bile acid receptor 1	0.123	BG37, GPBAR1, GPCR19, GPR131, M-BAR, MGC40597, TGR5	2q35	1	100%	100%	6	Digestive System Diseases	HGNC:19680	ENSG00000179921	MIM:610147
JAG1	jagged 1	0.799	AGS, AHD, AWS, CD339, HJ1, JAG1, JAGL1	20p12.1-p11.23	26	100%	98.47%	35	Eye Diseases, Otorhinolaryngologic Diseases, Cardiovascular Diseases,...	HGNC:6188	ENSG00000101384	MIM:601920
MYO5B	myosin VB	0.764	KIAA1119, MYO5B	18q	40	100%	96.61%	50	Congenital, Hereditary, and Neonatal Diseases and Abnormalities, Nervous...	HGNC:7603	ENSG00000167306	MIM:606540
NOTCH2	notch 2	0.543	NOTCH2	1p13-p11	34	99%	99.06%	54	Endocrine System Diseases, Congenital, Hereditary, and Neonatal Diseases and...	HGNC:7882	ENSG00000134250	MIM:600275
NR1H4	nuclear receptor subfamily 1 group H member 4	0.423	FXR, HRR-1, HRR1, NR1H4, RIP14	12q23.1	10	100%	99.74%	13	Digestive System Diseases, Female Urogenital Diseases and Pregnancy...	HGNC:7967	ENSG00000012504	MIM:603826
SLC25A13	solute carrier family 25 member 13	71.000	ARALAR2, CITRIN, CTLN2, SLC25A13	7q21.3	18	100%	97.42%	22	Congenital, Hereditary, and Neonatal Diseases and Abnormalities, Nervous...	HGNC:10983	ENSG00000004864	MIM:603859
SLCO1B1	solute carrier organic anion transporter family member 1B1	0.249	OATP1B1, OATP-C, SLC21A6, SLCO1B1	12p12	14	98%	93.36%	20	Congenital, Hereditary, and Neonatal Diseases and Abnormalities,...	HGNC:10959	ENSG00000134538	MIM:604843
SLCO1B3	solute carrier organic anion transporter family member 1B3	0.241	OATP1B3, OATP8, SLC21A8, SLCO1B3	12p12	14	100%	94.67%	25	Respiratory Tract Diseases, Congenital, Hereditary, and Neonatal Diseases and...	HGNC:10961	ENSG00000111700	MIM:605495
TJP2	tight junction protein 2	0.482	DFNA51, TJP2, X104, ZO-2, ZO2	9q13-q21	24	100%	99.79%	34	Eye Diseases, Digestive System Diseases, Nutritional and Metabolic Diseases,...	HGNC:11828	ENSG00000119139	MIM:607709
VPS33B	VPS33B, late endosome and lysosome associated	0.485	FLJ14848, VPS33B	15q26.1	23	99%	91.95%	23	Nutritional and Metabolic Diseases, Digestive System Diseases, Musculoskeletal...	HGNC:12712	ENSG00000184056	MIM:608552

Table 5. Fifteen-gene panel by Ion Ampliseq™ Designer tool. Table view of the genes with several parameters (www.ampliseq.com).

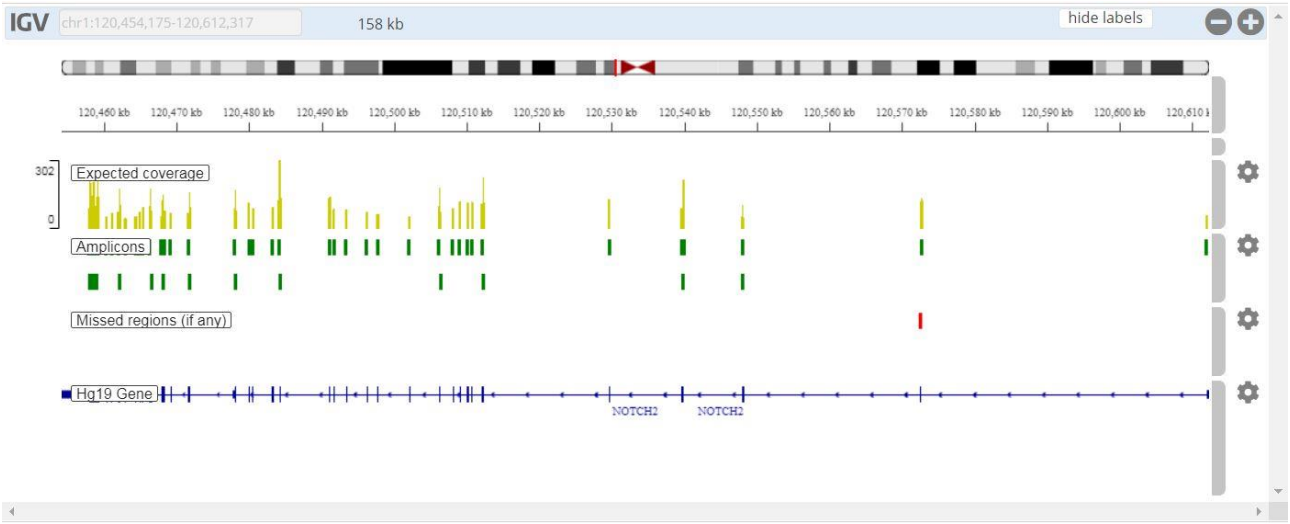


Figure 17. Integrative Genomics Viewer (IGV) of *NOTCH2* gene. Genomic region (158 kb, on the top) of *NOTCH2* gene visualized on IGV through Ion Ampliseq™ Designer tool. From the top to the bottom: yellow segments show the theoretical coverage under standard conditions; green segments show amplicons produced by the tool sometimes overlapping; in this gene there are no missed region; blue line indicates genomic space of the gene and vertical segments represent exons (www.ampliseq.com).

3.2 Results from four-gene panel analysis

Of the 96 patients sequenced with the four-gene panel, we had clinical and laboratory information for 48 of them. Chronic idiopathic cholestasis was defined as GGT e/o alkaline phosphatase persistently ≥ 1.5-fold the upper normal values in at least two tests or as history of itching combined with elevated serum bile acids concentration (10 μmol/l) for more than 6 months. PFIC patients underwent laboratory analysis, including bile acids serum concentration, liver fibrosis evaluation by transient elastography (Fibroscan®) and, if indicated, liver biopsy within 6 months from the execution of the genetic tests. Comprehensive molecular analysis of *ABCB11*, *ATP8B1*, *ABCB4* and *TJP2* genes was performed by Ion Torrent PGM (Thermo Fisher Scientific, Waltham, MA, USA). Molecular analysis was simultaneously carried out in 12 barcoded samples in an Ion 318 v2 Chip for each run. We used our variation filtering strategy according to ACMG standards (¶ **2.6.1**), patients' phenotypes and their familial history.

In **table 6** are shown the main features of our cohort. A history of familial cholestatic diseases and DIC was present in 35%, itching in 27%, neonatal jaundice in 21%, juvenile

cholelithiasis in 17%, personal or family history of ICP in 13%. Histologic features of cholestasis were present in 65% of cases (overall 44 patients agreed to a liver biopsy).

N	48
Male *N* (%)	28 (58.3)
Adult population ≥ 18 years *N* (%)	45 (93.8)
Age (years, mean ± SD)	
At time of genetic test	42 ± 15
At presentation	32 ± 15
Risk factor for cholestasis *N* (%)	
DIC history	17 (35.4)
Neonatal jaundice	10 (20.8)
Itching history	13 (27.1)
ICP history	6 (12.5)
Juvenile cholelithiasis	8 (16.7)
Familiarity	17 (35.4)
Laboratory (median, range)	
GGT (UI/l)	139 (5–597)
FA (UI/l)	283 (122–1012)
ALT (UI/l)	45 (9–387)
Bilirubin (mg/dl)	0.7 (0.4–11.2)
Bile acids (μmol/l)	11 (2.3–403)
Cholesterol (mg/dl)	211 (87–328)
Albumin (g/dl) 4.3 (2.9–5.2)	4.3 (2.9–5.2)
Platelets ($n \times 10^3$/μl)	229 (61–417)
FibroScan (kPa) (median, range)	5.3 (3.1–35.8)
Histologic features of cholestasis *N* (%)	31 (64.6)

Table 6. Schematic representation of age, time, risk factors and laboratory variables of our cohort. N: number, SD: standard deviation, DIC: drug-induced cholestasis, ICP: intrahepatic cholestasis of pregnancy, GGT: gamma-glutamyltransferase, AP: alkaline phosphatase, ALT: alanine transaminases, kPa: kilopascal (Vitale et al. *J Gastroenterol*. 2018).

Gene analysis revealed the presence of P/LP mutations in about one-fourth of subjects; clinical features and type of mutations found in the 12 patients (eleven unrelated probands and one affected sister) were reported in **table 7.**

Table 7

ID	Sex	*ATP8B1*	*ABCB11*	*ABCB4*	*TJP2*	Age	GGT (U/L)	AP (U/L)	AST/ ALT (U/L)	Bilirubin (mg/dL)	Fibro scan (kPa)	Notes
1	F	Ht p.R952Q LB **Ht p.D1219H LP**	Ht p.V444A LB Ht p.M677V LB	WT	**Ht p.R353W LP** Ht p.M699I LB	71	121	394	23 19	1.02 0.17	5,5	-
2	M	**Ht p.P23L LP**	Ht p.V444A LB Ht p.M677V LB	WT	WT	31	28	138	19 29	3.08 0.77	7.4	Familiarity; Juvenile Cholelithiasis; Neonatal Jaundice
3	F	WT	**Ht p.Y93S LP** Ht p.V444A LB **Ht p.V597L LP** **Ht p.R1128C LP**	Ht p.I237= LB	WT	29	15	183	81 67	18.31 9.21	12	DIC; Neonatal Jaundice; Itching; Juvenile Cholelithiasis; ICP
4	F	Ht p.R952Q LB	**Hm p.A523G LP** Hm p.V444A LB	Ht p.I237= LB	WT	20	9	358	20 9	0.80 0.30	20.9	Familiarity; DIC; Neonatal Jaundice; Itching
4a	F	WT	**Hm p.A523G LP** Hm p.V444A LB	WT	WT	7	13	1012	69 55	3.45 2.12	10.1	Familiarity; DIC; Neonatal Jaundice; Itching
5	M	WT	**Ht p.E135K LP** HM p.V444A B Ht **p.S1100Qfs*38 P**	WT	WT	16	26	457	70 41	9.16 3.98	11.5	Neonatal Jaundice; Itching
6	F	Ht R952Q (LB)	Ht G278E VUS Hm V444A LB **Ht R832C LP**	Ht T175A VUS	WT	40	50	191	30 41	1.23 0.37	3.5	Liver transplantation for Byler disease at 16y
7	F	WT	WT	**Ht Y403C LP** Ht R652G LB	WT	42	54	106	24 22	0.8 n.a.	n.a.	ICP, LPAC familiarity for juvenile cholelithiasis
8	F	WT	Hm p.V444A LB	Ht p.I237= LB **Ht p.K672* P**	Ht p.T1155= LB	39	363	562	64 77	0.59 0.12	7.7	Familiarity; DIC; Itching; Cholelithiasis
9	M	WT	WT	Ht p.I237= LB **Ht p.A364V LP**	Ht p.Q159K LB Ht p.M699I LB	57	379	754	23 31	0.66 0.21	35.8	DIC; HCC
10	F	WT	Ht p.V444A LB Ht p.M677V LB	Hm p.I237= LB	**Ht p.T93M LP**	51	84	31 41	205	0.63 0.13	4,4	DIC; ICP
11	M	WT	Hm p.V444A LB	**Ht p.T775M LP**	Hm p.R24H LB **Ht p.I906T LP**	37	235	30 74	399	0.6	11	Familiarity

Table 7. Patients with P/LP mutations identified in the four-gene panel. ID 4 and 4a are affected sibling. Sequence variants are reported according to HGVS recommendations (http://varnomen.hgvs.org).

3.2.1 *ATP8B1* variants

Five variants were identified in the coding region of *ATP8B1*. According to ACMG standards, two variants were labelled as LP: c.3655G>C (p.D1219H) in exon 28 and c.68C>T (p.P23L) in exon 2; the first was combined with heterozygous *TJP2* LP mutation c.1057C>T in a 71-year-old woman (case 1 in **table 7**) with a long history of cholestasis, 5.5 kPa at Fibroscan (F1). The second is a novel mutation found in a 31-year-old male (case 2 in **table 7**) who experienced neonatal jaundice and juvenile cholelithiasis, with low GGT, 7.4 kPa at Fibroscan (F2) and having also LB variants in *ABCB11*.

We classified the missense c.134A>C (p.N45T) in exon 2 and c.607A>G (p.K203E) in exon 7 as VUS and each was previously described as risk allele in ICP (CM051009 and CM051010, Painter et al. 2005) [68]. We found these variants in compound heterozygous in a 36-year-old female who showed a cholestatic disease characterized by normal GGT and high bile acids levels, neonatal jaundice, juvenile cholelithiasis complicated by LPAC, and a history of ICP. We found the N45T three times (two with this panel and one with 15-gene panel) and the K203E two times. The second patient having the N45T is a 60-year-old woman having also three LB variants in *ABCB11*, *ABCB4* and *TJP2*.

The second patient carrying the K203E is a 71-year-old woman with a late onset.

We found one B/LB variant: the c.2855G>A (p.R952Q) in exon 23. The SNP p.R952Q was found in five cases and in two patients was combined with LP mutations, one on *ABCB11* and one on *TJP2*, respectively. Allele frequency of p.R952Q was significantly higher in comparison with the one of East Asian population (10.4 vs. 0.03%: $p < 0.0001$) and comparable with European and worldwide data (**table 8**).

Variants' analysis and structural annotations on available crystalized structures or homology models from PDB (Protein Data Bank, www.rcsb.org) revealed that two variants, i.e., p.K203E and p.F305I are found in spatial proximity within the ATPase catalytic domain (**figure 18**), suggesting potential converging effect on the enzymatic activity, while the p.R952Q is found at the transmembrane (i.e., phospholipid translocating) domain. The LP

variants P23L and D1219H, reside at the N- and C-terminals of the protein, outside the structured portion (**figure 21** page 77).

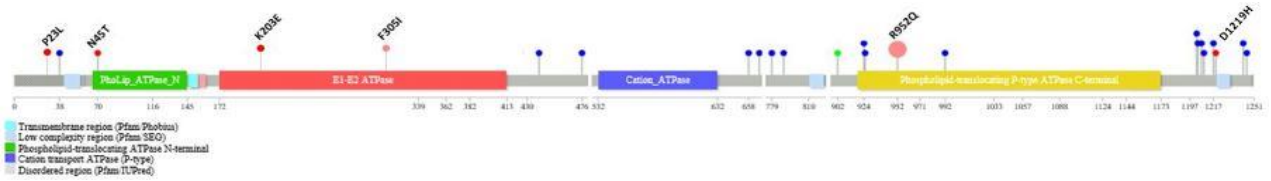


Figure 18. ***ATP8B1*** **lollipops.** Protein sequence annotations of PFIC1 non-synonymous mutations displayed as pink and red lollipops to indicate respectively B/LB or VUS and P/LP predicted consequences. Circle diameters are proportional to the number of patients affected. Blue and green lollipops indicate phosphorylation and acetylation sites. Protein sequence regions corresponding to conserved domains are highlighted by colored boxes and are indicated by their respective Pfam names.

3.2.2 *ABCB11* variants

Eleven variants were identified in the coding region of *ABCB11* and seven resulted as P/LP mutations. Three causative mutations, c.278A>C (p.Y93S) in exon 5, c.1789G>C (p.V597L) in exon 15 and c.3382C>T (p.R1128C) in exon 25 were combined in a young female (29 years) presenting with a history of neonatal jaundice, DIC, ICP, juvenile cholelithiasis, itching, normal GGT and high bilirubin levels (case 3 in **table 7**). Fibroscan analysis revealed a score of 12 kPa (F3).

Two sisters (20 and 7 years) were carriers of the LP homozygous mutation c.1568C>G (p.A523G) [92] in exon 14: both had history of neonatal jaundice, itching and DIC (cases 4 - 4a in **Table 7**), but there were differences at Fibroscan analysis (20.9, F4 vs 10.1 kPa, F3). Last two P/LP mutations, c.3297delC (p.S1100Qfs*38) in exon 25 and c.403G>A (p.E135K) in exon 6, were responsible for a double heterozygosis in a boy with history of itching and neonatal jaundice, 11.5 kPa (F3) at Fibroscan and listed for liver transplantation (case 5 in **table 7**).

The disease-causing mutation c.2494C>T (p.R832C) in exon 21 was previously reported in HGMD (CM067617) as PFIC2-causing mutation by Walkowiak et al. 2006. We found it in a 40-year-old female who had normal GGT and undergone liver transplantation at age of sixteen due to Byler's disease, and no information about familial lineage were present. The

patient had also the *novel* VUS (private variant) c.833G>A (p.G278E) in exon 9 of *ABCB11*, another VUS in *ABCB4* and the LB variant R952Q in *ATP8B1* (case 6 in **table 7**).

Multiple lines of computational evidence support a deleterious effect on the gene or gene product by the G278E (13 pathogenic predictions from DANN, DEOGEN2, FATHMM, FATHMM-MKL, FATHMM-XF, LRT, M-CAP, MutationAssessor, MutationTaster, PROVEAN, PrimateAI, SIFT and SIFT4G vs no benign predictions), it is not present in ClinVar and for VarSome it is considered as VUS, due to poor information.

Another VUS was the novel variant c.209T>C (p.F70S) in exon 5.

The c.1268A>G (p.H423R) in exon 12 is a rare variant (MAF: 0.0002) predicted to be tolerated (0.53) by SIFT and probably damaging (1.000) by PolyPhen (discordant predictions by the two tools) and it was identified in a 48-years-old female with elevated GGT, carrying also 2 LB variants in *ABCB11* and *ATP8B1*. MLPA analysis for large genomic rearrangements of *ABCB4* gene was also performed on patient's DNA, resulting negative.

Benign mutations detected were among those previously described [74]: c.2029A>G (p.M677V) in exon 17, detected in nine patients, and c.1331T>C (p.V444A) in exon 13, found in 40 cases (83.3% of population studied). Of note, SNP c.1331T>C was reported associated to DIC and ICP in previous reports [74] and recently to cholestatic phenotype in patients without disease-causing mutations in the respective gene. Allele frequency of p.V444A was significantly more frequent in our cohort of patients without P/LP mutations in comparison with European and worldwide populations (81 vs. 60%: $p = 0.008$; 81 vs. 57%: $p = 0.002$). We also observed a higher allele frequency of p.M677V in our cohort with and without P/LP mutations respect to all other populations (13.5 vs. 1.76% European, 0% East Asian, 2.7% worldwide: all $p < 0.0001$) (**table 8**).

All nine reported mutations for *ABCB11* were mapped to homolog available structures from PDB through Mechismo (**figure 19**). P/LP variants were found to be located either at ATP-binding domain, like p.A523G, p.V597L, p.S1100Qfs*38 and p.R1128C, or at transmembrane domain, like p.E135K or p.Y93S, which are also found in spatial proximity (**figure 20** page 77). Moreover, the variant p.H423R was found at the interface with the ATP-binding pocket.

This overall suggests that deleterious mutations on *ABCB11* might similarly affect the ATP-binding cassette (ABC) transporter's functionality by impairing ABC catalytic activity, or by perturbing transmembrane region.

Gene	Mutation	Protein	dbSNPs	AF full cohort (a) (%)	AF Pts w/o P/LP mutations (b) (%)	AF EU (1) (%)	AF East Asia (2) (%)	AF world (3) (%)	*p* values *f*	
ATP8B1	c.2855G>A	p.R952Q	rs12968116	10.4	10.8	12	0.03	8.3	a-1 = 0.9151	b-1 = 0.8297
									a-2 < 0.0001	**b-2 < 0.0001**
									a-3 = 0.5888	b-3 = 0.5744
ABCB11	c.1331T>C	p.V444A	rs2287622	83	81	60	73	57	**a-1 < 0.0001**	**b-1 = 0.008**
									a-2 = 0.1329	b-2 = 0.3318
									a-3 = <0.001	**b-3 = 0.003**
	c.2029A>G	p.M677V	rs11568364	18.8	13.5	1.76	0	2.73	**a-1 < 0.0001**	**b-1 < 0.0001**
									a-2 < 0.0001	**b-2 < 0.0001**
									a-3 < 0.0001	**b-3 < 0.0001**
ABCB4	c.711A>T	p.I237=	rs2109505	33.3	16.2	17.7	26.1	20.9	**a-1 = 0.0083**	b-1 = 0.0903
									a-2 = 0.3319	b-2 = 0.7558
									a-3 = 0.0338	b-3 = 0.1854
	c.1954A>G	p.R652G	rs2230028	8.3	10.8	7.5	27	10.5	a-1 = 0.8320	b-1 = 0.6558
									a-2 = 0.0476	**b-2 = 0.0106**
									a-3 = 0.6184	b-3 = 0.9574
TJP2	c.71G>A	p.R12H	rs4493966	4.2	2.7	3.8	4.9	6.7	a-1 = 0.8817	b-1 = 0.7357
									a-2 = 0.7924	b-2 = 0.7937
									a-3 = 0.4862	b-3 = 0.333
	c.475C>A	p.Q128K	rs41305539	8.3	8.1	1.8	5.7	2.9	**a-1 = 0.0005**	**b-1 = 0.0033**
									a-2 = 0.0777	b-2 = 0.1742
									a-3 = 0.4304	b-3 = 0.5266
	c.2097G>A	p.M668I	rs34774441	14.6	12.1	6.3	0.03	5.3	**a-1 = 0.0385**	b-1 = 0.4267
									a-2 < 0.0001	**b-2 < 0.0001**
									a-3 = 0.0043	b-3 = 0.1372

Table 8. Allele frequencies of detected SNPs. The allele frequencies (AF) of common single-nucleotide polymorphisms (SNPs) in *ATP8B1*, *ABCB11*, *ABCB4* and *TJP2*, in the full cohort (a) of patients with and without (w/o) pathogenic/likely pathogenic (P/LP) mutations (b) were compared to AF available from the Genome Aggregation Database (gnomAD) either of the non-Finnish European population (1), East Asian population (2), and worldwide population (3) (gnomAD: http://gnomad.broadinstitute.org/). *p* values were calculated using the Chi-square test by GraphPad web software (https://www.graphpad.com/quickcalcs/contingency1.cfm). $p < 0.05$ are in bold (Vitale et al. *J Gastroenterol*. 2018).

3.2.3 *ABCB4* Variants

We identified seven variants in the coding region of *ABCB4* and four were P/LP mutations. The mutation c.1208A>G p.Y403C was a *novel* variant. It is located in the same position of a mutation previously described by Degiorgio [93] Y403H (CM075960 on HGMD) as associated to PFIC phenotype. For this reason, we can assess that this aminoacid residue is a mutational

hotspot, so the probability that the variant is pathogenic is very high. We found this mutation in a 42-year-old female (case 7 in **table 7**) with history of ICP, LPAC and juvenile cholelithiasis. The same variant was found in his affected brother. Only the man had also a LB variant in *ABCB4*.

The c.2014A>T (p.K672*) in exon 16, previously undescribed, was found in a 39-year-old woman (case 8 in **table 7**) with familiarity for cholestatic diseases and cholelithiasis (mutation was inherited from the affected father), history of DIC and itching, with high GGT and a Fibroscan score of 7.7 kPa (F2).

The c.1091C>T (p.A364V) in exon 10, described by Degiorgio et al. [93], was present in a 57-year-old cirrhotic patient with high AP and GGT, 35.8 kPa at Fibroscan (F4), DIC and hepatocellular carcinoma (case 9 in **table 7**).

The c.2324C>T (p.T775M) seemed to be a VUS and found in a 41-year-old male with a mutation in TJP2 (I906T). Sanger sequencing confirmed these two mutations to segregates in the affected brother (31 years). For this reason, we reclassified the T775M as LP mutation.

Three VUS, p.L73V (c.217C>G) in exon 4, c.523A>G (p.T175A) in exon 5 and c.2324C>T (p.T775M) in exon 19 were identified, previously associated to different PFIC3 phenotypes. The c.523A>G (p.T175A) was seen two times in our cohort (4- and 15-gene panel).

The missense variant p.T775M was detected together with LP c.2717T>C (p.I906T) in exon 18 of *TJP2* in a 37-year-old male presenting a familiarity for cholestasis (case 11 in **table 7**), while the variant p.L73V was detected in a 23-year-old female with high GGT levels, without evidence of significant liver fibrosis at Fibroscan.

The known LB variants c.711A>T (p.I237=) in exon 8 and c.1954A>G (p.R652G) in exon 16 were found in 16 and four cases, respectively. The SNP p.I237= was associated in previous reports with elevated GGT-cholestasis, ICP, and gallstones [56].

Allele frequency of p.I237= was significantly more common in our cohort of patients with disease-causing mutations than in European and worldwide population (33 vs. 18%: p = 0.004; 33 vs. 21%: p = 0.033; **table 8**). Frequency of p.R652G was similar between our cohort and both European and worldwide population.

ABCB4 non-synonymous mutations, with exception of the stop-gain p.L672*, were mapped to the same available homolog structure used for *ABCB11* but revealed no defined patterns (**figure 19**). Mutations were essentially located in intrinsically disordered regions (**figure 21** page 77).

Among our cohort, physicians selected 21 patients with a phenotype attributable to *ABCB4* gene. Thus, *ABCB4* MLPA analysis was performed and only one patient had a heterozygous deletion from the exon 16 to 28. Unfortunately, we had no clinical information.

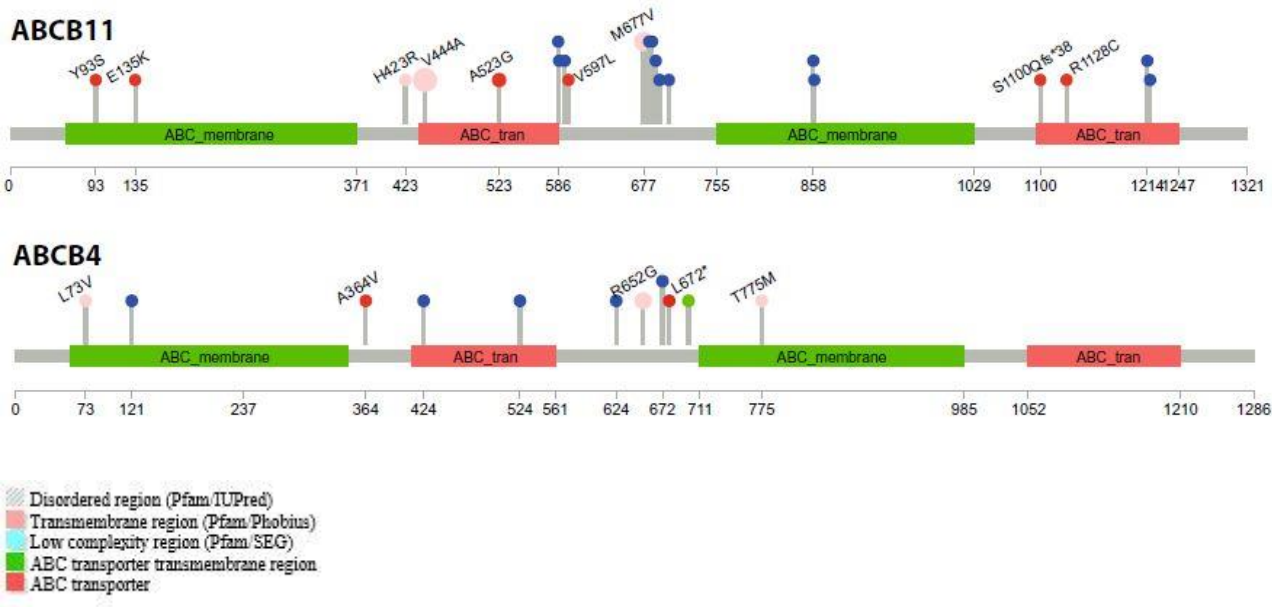


Figure 19. *ABCB11* and *ABCB4* lollipops. Legend is the same of the Figure 18 (ATP8B1). ABCB11 and ABCB4 have the same Pfam domains, as they are both ATP-binding cassette sub-family B members.

3.2.4 *TJP2* variants

Variants identified in the coding region of *TJP2* were thirteen: three LP mutations, one VUS, and nine B/LB variants.

Among LP mutations, the c.278C>T (p.T93M) in exon 3 was found in two unrelated cases (case 10 in **table 7**), one who experienced DIC and ICP.

The c.1057C>T (p.R353W) in exon 6, combined to LP heterozygous p.D1219H on *ATP8B1* in a 71-year-old man presenting only high GGT (case 1), and c.2717T>C (p.I906T) in exon 18 associated with LP p.T775M in *ABCB4* in one patient with familiarity for cholestatic diseases

(case 11 in **table 7**). Differently from what has been reported in a previous report [19], all three patients presented high GGT.

The only VUS identified was the c.46A>G (p.K16E) in exon 1. B/LB mutations were c.43G>C (p.V15L) and c.71G>A (p.R24H) in exon 1, c.475C>A (p.Q159K) and c.860C>T (p.A287V) in exon 5 (in two unrelated probands), c.1571G>A (p.R524K) in 3 unrelated probands in exon 10, c.1686C>T (p.Val562=) in exon 11, c.2097G>A (p.M699I) and c.2224T>C (p.S742P) in exon 14 and c.3465G>A (p.T1155=) in exon 22.

VUS p.K16E was present in a 31-year-old female associated to benign c.1954A>G and c.711A>T mutations in *ABCB4*: the patient had a history of ICP, juvenile cholelithiasis and recurrent DIC to different drugs used for multiple sclerosis. Finally, a 31-year-old man with high GGT and bile acids levels presented the LB variant V15L.

Most deleterious *TJP2* mutations were found at the level of PDZ, SH3, and guanylate kinase domains (**figure 20**). This suggests a potential perturbation of TJP2's function, i.e., organizing tight and adherent junctions by binding to the cytoplasmic C termini of junctional transmembrane proteins and linking them to the actin cytoskeleton.

Indeed, one variant (p.R353W) is predicted to affect the homo- and hetero-dimerization interface with *TJP1* (**figure 21** page 77).

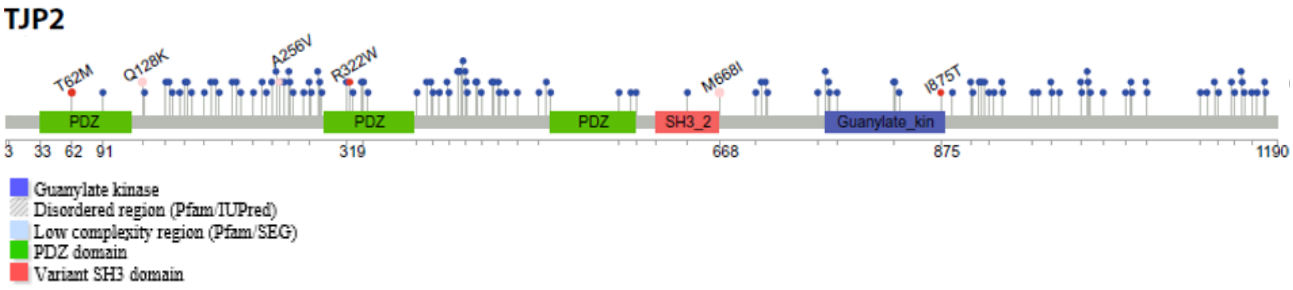


Figure 20. *TJP2* lollipops. Legend is the same of other lollipops figures. V15L, K16E and R24H are not present in figure. T93M in the text is at position 62 in the figure, Q159K is 128, A287V is 256, R353W is 322, M699I is 668 and I906T is 875. This is due to the following reason: our VCF were generated from ENST00000539225.1, NM_001170416 as reference transcript, but we were forced to use NM_004817.3 for annotation of variants on protein sequences through the lollipop approach (https://github.com/pbnjay/lollipops) and the Mechismo analysis, both requiring Uniprot canonical isoforms.

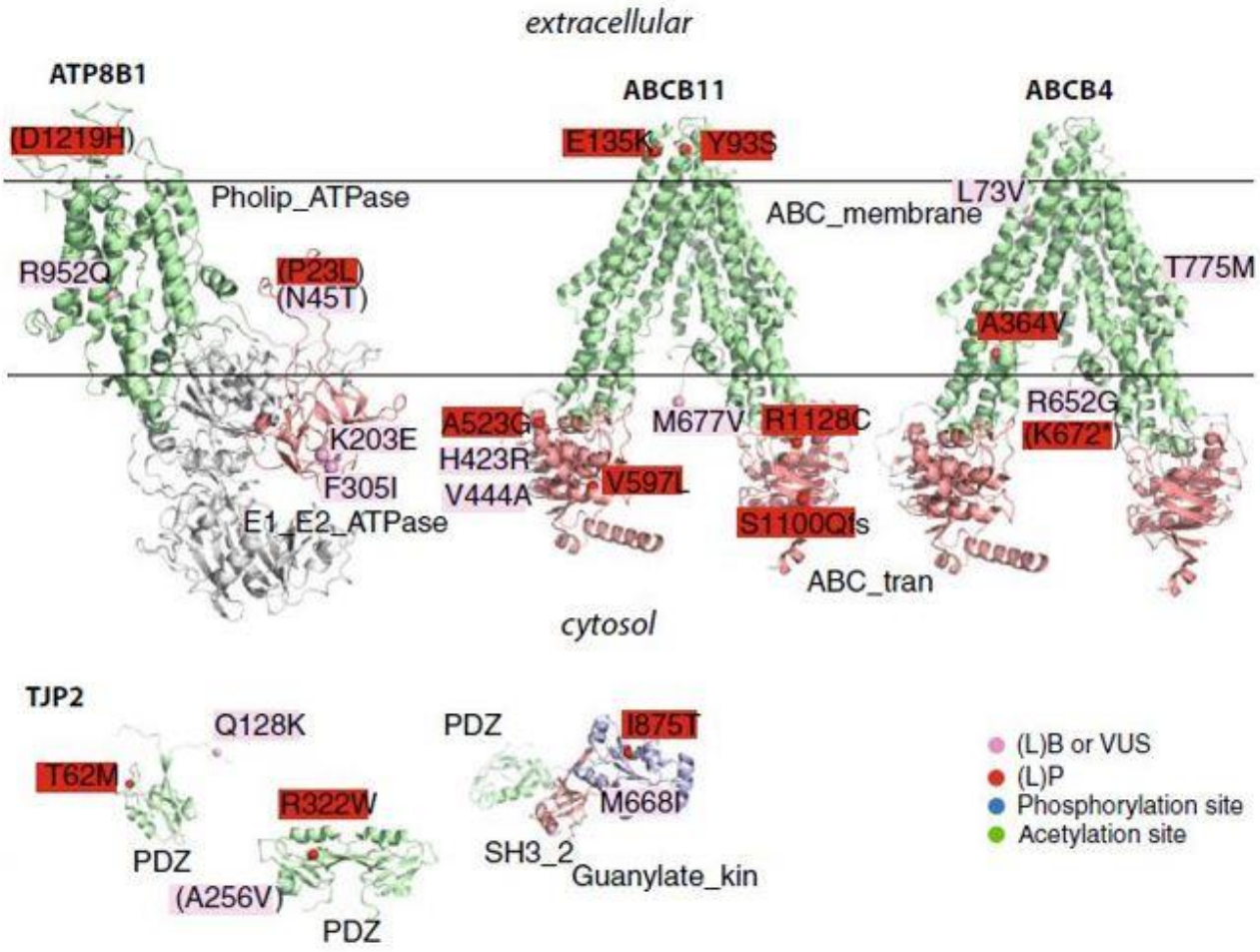


Figure 21. PFIC protein structures. Non-synonymous mutation annotations on available 3D structures as assessed through Mechismo, are displayed following the same colouring scheme as in **figures 20**, **21**, **22**. Labels in parenthesis indicate mutations with no structural information. PDB numbers for crystalized protein structures as follow. ATP8B1: 3TLM, Crystal Structure of Endoplasmic Reticulum Ca2+-ATPase (SERCA) from Bovine Muscle of *Bos taurus.* ABCB11 and ABCB4: 4F4C, Crystal Structure of the Multi-Drug Transporter, from *Caenorhabditis elegans.* TJP2: 2OSG, PDZ and SH3 domains from *Homo sapiens*.

3.2.5 FIC patients' profile

Excluding the common SNP p.V444A on *ABCB11*, (present in 83.3% of cases) and p.I237= on *ABCB4* (33%), 26/48 patients (58.3%), have at least one mutation in one or more genes: 8 patients had only one mutated allele while 18 patients had two or more mutated alleles. *ABCB11* and *TJP2* resulted the most affected genes and they are mutual exclusive.

It is possible that patients with multiple alleles mutated have more severe phenotypes by a synergistic effect at different bile transporters' sites as predicted by structural analysis.

However, patients having ≥ 2 non-synonymous mutations (with at least one predicted P/LP), tend to have higher liver stiffness comparing to patients with < 2 non-synonymous mutations (kPa at Fibroscan 7.9 vs. 4.9, $p = 0.028$), indicating a potential role of these variants as disease modifiers.

Significant differences between subjects with and without P/LP mutations were not observed in sex, age at symptoms presentation, and the following risk factors for PFIC: DIC history, family or personal history of ICP, juvenile cholelithiasis and family history for cholestatic diseases (**table 9**). However, at least one of these risk factors was present in 9/11 patients with P/LP mutations. Patients with P/LP mutations had more frequently neonatal jaundice (45.5 vs. 13.5%, $p = 0.036$) and itching (54.5 vs. 18.9%, $p = 0.029$). Regarding laboratory tests, significant differences were observed between the two subgroups only in terms of bile acids concentration (23.8 [4.4–403] vs. 8.8 [2.3–114], $p = 0.003$).

Liver stiffness was greater in subjects with P/LP mutations in comparison with the others (8.1 [3.3–35.8] vs. 4.8 [3.1–29.9], $p = 0.009$). All subjects of the first group showed histological features of intrahepatic cholestasis while 72% of patients without P/LP mutations presented it ($p = 0.018$).

	Pts with P/LP mutations $N = 11$	Pts without P/LP mutations $N = 37$	p value
Male *N*/Tot *N* (%)	5/11 (45.5)	23/37 (62.2)	0.260
Age: (years, mean ± SD)			
At time of genetic test	37 ± 19	44 ± 14	0.244
At symptoms presentation	27 ± 12	33 ± 15	0.308
Risk factor of cholestasis *N*/Tot *N* (%)			
DIC history	6/11 (54.4)	11/37 (29.7)	0.126
Neonatal jaundice	5/11 (45.5)	5/37 (13.5)	0.036
Itching history	6/11 (54.5)	7/37 (18.9)	0.029
ICP history	2/11 (18.2)	4/37 (10.8)	0.420
Juvenile cholelithiasis	3/11 (27.3)	5/37 (13.5)	0.259
Family history	5/11 (45.5)	12/37 (32.4)	0.327
Laboratory, median (range)			
GammaGT (UI/l)	121 (5–379)	148 (18–597)	0.425
AP (UI/l)	467 (138–1012)	280 (122–610)	0.061
ALT (UI/l)	41 (9–387)	46 (9–358)	0.608
Bilirubin (mg/dl)	1.2 (0.4–11.2)	0.7 (0.4–6.2)	0.484
Bile acids (μmol/l)	23.8 (4.4–403)	8.8 (2.3–114)	0.034
Cholesterol (mg/dl)	182 (132–303)	215 (87–328)	0.484
Albumin (g/dl) 4.3 (2.9–5.2)	4 (3.6–4.7)	4.3 (2.9–5.2)	0.126
Platelets (10^3/μL)	223 (128–412)	233 (61–417)	0.873
FibroScan (kPa), median (range)	8.1 (3.3–35.8)	4.8 (3.1–29.9)	0.009
Histologic features of cholestasis *N*/Tot *N* (%)	10/10 (100)	21/34 (61.8)	0.018

Table 9. Main features of patients according to presence of P/LP mutations. *N*: number, *Tot N*: total number, *DIC*: drug-induced cholestasis, *SD*: standard deviation, *P/LP*: mutations pathogenic and likely mutations, *ICP*: intrahepatic cholestasis of pregnancy, *AP*: alkaline phosphatase, *ALT*: alanine transaminases, *kPa*: kilopascal (Vitale et al. *J Gastroenterol.* 2018).

3.3 Results from fifteen-gene panel analysis

During the three-years PhD-project, 179 patients were enrolled at the S.Orsola-Malpighi Policlinic in Bologna. In about two years of the project we performed the four-gene analysis on 96 patients, while in about the last year of the project we used the new fifteen-gene panel for the molecular analysis of PFIC on 80 patients. In addition to the NGS variants validation, Sanger sequencing were executed on three patients to confirm the variant segregation of their relatives.

A total of 184 different variations was identified in all patients in the two panels. To assess the variant's category, we have adopted more stringent ACMG standards criteria than ones mentioned before. Other two tools helped us to understand the role of the variants. The first, ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) is a freely accessible, public archive of reports of the relationships among human variations and phenotypes, with supporting evidence, contributing to define the pathogenicity of a certain allele. The second, VarSome [94], is a search engine for human genomic variation, which enables users to look up variants in their genomic context, collects data from multiple databases in a central location and most importantly, aims to enable the community to freely and easily share knowledge on human variation. In the next paragraphs variants certainly benign will not described.

3.3.1 *ATP8B1* variants

More than 100 different mutations have been detected in *ATP8B1* gene so far in patients with PFIC1 (HGMD®), but the number of identified variants with no clear definition of pathogenicity is still increasing.

We identified 8 VUS, 1 likely benign and 2 benign. Among VUS, the third patient carrying the N45T in our cohort (¶ **3.2.1**) is a 65-year-old male subjected to fifteen-gene panel analysis who has also LB variants in *ABCB11*, *ABCB4*, *MYO5B*, *SLCO1B1* and *SLCO1B3*.

An 18-year-old man, who presented histological signs of biliary ductopathy, increased levels of bile acids and an history of metabolic liver disease in the father, carries two VUS.

The c.208G>A (p.D70N) in exon 3 is a rare variant (MAF: 0.004) associated to BRIC (CM043812, Klomp et al. 2004) and the c.1288A>G (p.K430E) in exon 13 is a *novel* variant predicted to be pathogenic by *in-silico* predictions. The patient had also LB variants in *ABCB11*, *ABCC2*, *SLCO1B1* and *SLCO1B3*.

The c.913T>A (p.F305I) in exon 10 was previously associated to PFIC [56] and present in a 58-year-old woman who had also LB variants in *ABCB11* gene.

The c.1177A>G (p.I393V) in exon 12 is present in HGMD (CM1716476, Meng et al. 2017), it is a rare variant (MAF: 0.006) and was identified in 2 patients.

The c.1498T>C (p.Y500H) in exon 15 was previously associated to PFIC (CM043818, Klomp et al. 2004), but all the other tools referred to it as VUS.

The c.2771A>G (p.Y924C) in exon 23 is a rare VUS (MAF: 0.00004) with discordant predictions of pathogenicity. It was found in a 68-year-old man with other two VUS in *MYO5B* and *NOTCH2*. Other LB variants were found in *ABCB11*, *ABCB4*, *ABCC2*, *JAG1*, *SLCO1B1* and *TJP2*.

The c.3589G>T (p.V1197L) in exon 28 is a rare VUS ever described before, but all bioinformatic tools agree for no functional impact for it. This variant was found in a patient with a frameshift in *JAG1* (¶ **3.3.7**).

The c.3707C>G (p.A1236G) in exon 28, is a novel variant predicted not to affect the protein and, even for the poor knowledge about it, it is classified as VUS.

The only LB variant found in *ATP8B1* is once again the c.2855G>A (p.R952Q) in exon 23. This variant is a polymorphism because of the MAF (0.12). It has also the same frequency between the two major population databases and our patients' cohort: 0.12 for European population by gnomAD (genome aggregation database, https://gnomad.broadinstitute.org/), 0.11 for TSI (Tuscan from Italy) in 1000Genome and 0.15 in our cohort (27/179). All identified *ATP8B1* variants during the project are in **table 10**.

ATP8B1

ex	DNA (c.)	Protein (p.)	HGMD/dbSNP/*novel*	MAF	VarSome	ClinVar	our	*N*
2	68C>T	P23L	rs1345769999	-	VUS	-	LP	1
2	134A>C	N45T	CM051009[1]/rs146599962	0.007	VUS	-	VUS	3
3	208G>A	D70N	CM043812[2]/rs34719006	0.00449	VUS	VUS	VUS	1
7	607A>G	K203E	CM051010[1] DM?/rs56355310	0.001	VUS	VUS	VUS	2
10	913T>A	F305I	CM1722276[3] DM?/rs150860808	0.001	LB	LB/VUS	VUS	1
12	1177A>G	I393V	CM1716476[4]/rs34315917	0.00615	LB	B/LB/VUS	VUS	2
13	1288A>G	K430E	*novel*	-	VUS	-	VUS	1
15	1498T>C	Y500H	CM043818[2]/rs147642236	0.000589	LB	VUS	VUS	1
23	2771A>G	Y924C	rs145287364	0.0000439	VUS	VUS	VUS	1
23	2855G>A	R952Q	rs12968116	0.121	B	B/LB	LB	27
28	3589G>T	V1197L	rs1376578283	0.0000122	LB	VUS	VUS	1
28	3655C>G	D1219H	rs753339861	0.0000183	VUS	VUS	LP	1
28	3707C>G	A1236G	*novel*	-	VUS	-	VUS	1

Table 10. All identified *ATP8B1* variants in all patients during the project. 1. Painter (2005) *Eur J Hum Genet*; 2. Klomp (2004) *Hepatology*; 3. Dröge (2017) *J Hepatol*; 4. Meng (2017) *JAMA*; DM?: likely disease-causing mutation; *N*: number of cases. benign variants are not shown.

3.3.2 *ABCB11* variants

So far, more than 400 missense variants have been reported in *ABCB11* gene on HGMD. We identified 11 variants: 3 VUS, 4 LB variants and 4 benign variants (not reported here). Among VUS, we identified the *novel* variant c.209T>C (p.F70S) in exon 5 combined to a frameshift in *ABCB4* (§ **3.3.3**).

Always in the exon five, the c.383G>A (p.R128H) is a rare variant (MAF: 0.00003) never associated to the disease and with discordant bioinformatic predictions. It was found in a 50-years-old man carrying another VUS in *NOTCH2* (F1209V) and LB variants in *MYO5B*, *SLCO1B1* and *SLCO1B3*.

The synonymous variant c.2739A>G (p.L913=) was classified as VUS by Varsome because is absent from controls in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium (MAF: $9*10^{-6}$) and it is located in a mutational hot spot and/or critical and well-established functional domain (e.g., active site of an enzyme). HSF (Human Splicing Finder) analysis revealed a potential alteration of splicing by activation of an exonic cryptic donor site (algorithm: HSF Matrices) or by creation of an exonic ESS site. ESSs inhibit

or silence splicing of the pre-mRNA and contribute to constitutive and alternate splicing. However, this *in-silico* analysis should be confirmed with other *in-vitro* studies.

Among LB variants, we identified the c.957A>G (p.G319=) in exon 10. This variant was previously described by Byrne et al. 2009 as splicing modificatory by *in vitro* functional studies (CM092797). Nevertheless, due to the high and different frequencies in worldwide population (spanning from 0.327 in African to 0.000613 in Est Asian) it is considered LB in our cohort, formed mainly by European population.

The following two LB variants are described in the previous section (¶ **3.2.2**): the c.1331T>C (p.V444A) in exon 13 and c.2029A>G (p.M677V) in exon 17. At the end of the project the polymorphism c.1331T>C was present in 153/176 patients (86%), while population data from 1000 Genome in TSI is 63% and allele frequency from gnomAD is 59%.

The polymorphism c.2029A>G (p.M677V) was present in 20/176 (11%), 0.017 gnomAD 0.0327 1000 Genome.

The c.3084A>G (p.A1028=) in exon 24 is a polymorphism present in HGMD (CM092757, Byrne et al. 2009) with additional functional evidence and associated with primary biliary cirrhosis. It is also present in dbSNP (rs497692) and gnomAD with a G allele frequency of 55%. In our cohort, instead, frequency of G allele is 148/176 (84%).

All identified *ABCB11* variants during the project are in **table 11**.

ABCB11

ex	DNA (c.)	Protein (p.)	HGMD/dbSNP/*novel*	MAF	VarSome	ClinVar	our	*N*
5	209T>C	F70S	*novel*	-	VUS	-	VUS	1
5	278A>C	Y93S	rs1185733805	0	VUS	-	VUS	1
5	383G>A	R128H	rs181533618	0.0000268	LB	VUS	VUS	1
6	403G>A	E135K	CM092737[1]/rs752992432	0.0000579	VUS	LP	LP	1
9	833G>A	G278E	*novel*	-	VUS	-	VUS	1
10	957A>G	G319=	CM092797[1] FP/rs7563233	0.02	LB	LB	LB	20
12	1268A>G	H423R	rs147522210	0.000231	VUS	-	VUS	1
13	1331T>C	V444A	CM071525[2] DFP/rs2287622	0.5972	B	LB	LB	153
14	1568C>G	A523G	CM1611166[3]/rs769652427	0.0000267	VUS	-	LP	2
15	1789G>C	V597L	*novel*	-	VUS	-	VUS	1
17	2029A>G	M677V	CM1723787[4] DM?/rs11568364	0.0178	B	LB	LB	20
21	2494C>T	R832C	CM067617[5]/rs772294884	0.00000888	LP	LP	LP	1
22	2739A>G	L913=	rs775098358	0.000008	VUS	-	VUS	1
24	3084A>G	A1028=	CM092757[1] DFP/rs497692	0.5	B	B	LB	148
25	3297delC	S1100QfsX38	CD1513805[6]	0	P	-	P	1
25	3382C>T	R1128C	CM081482[7]	0	LP	-	LP	1

Table 11. All identified *ABCB11* variants in all patients during the project. 1. Byrne (2009) *Hepatology*; 2. Lang (2007) *Pharmacogenet Genomics*; 3. Vitale (2016) *Ann Hepatol*; 4. Dröge (2017) *J Hepatol*; 5. Walkowiak (2006) *J Pediatr Gastroenterol Nutr*; 6. Giovannoni (2015) *PLoS One*; 7. Strautnieks (2008) *Gastroenterology*; FP: in vitro or in vivo functional polymorphism; DFP: disease-associated polymorphism with additional functional evidence; DM?: likely disease-causing mutation; *N*: number of cases; benign variants are not shown.

3.3.3 *ABCB4* variants

More than 300 missense/non sense variants have been reported in *ABCB4* gene on HGMD. We identified 15 variants divided as follows: 3 P/LP mutations, 2 VUS, 3 LB variants and 7 benign (not reported here).

As regard P/LP mutations, the c.959C>T (p.S320F) in exon 9 was previously associated to PFIC (CM013506) [75] as disease-causing mutation. This mutation was found in a 37-year-old female with a familial history of biliary lithiasis. The patient had also LB variants in *ABCB11* (V444A), *SLCO1B1*, *SLCO1B3* and *VPS33B* (case 1 in **table 17**).

Among disease-causing variants, the small insertion c.1135_1136insAA (exon 11) causes an aminoacid change (Lysine) and a translation's shift (p.S379Kfs*35) that induce a premature stop codon after 35 aminoacids. The novel mutation produces a 416 aminoacids long

protein, 1/3 of the wild-type protein (1286aa). This variant was found in a 58-year-old female with familial history of ICP and cholelithiasis and a F1/F2 Fibroscan score of 7 kPa (case 2 in **table 17**).

The c.1442T>G (p.L481R) in exon 13 was previously described by Degiorgio et al. 2014 and it is present in HGMD (CM144425). We found this mutation in a 33-year-old female with high GGT and alkaline phosphatase with a familial history of hepatic disease with high GGT (case 3 in **table 17**).

Two VUS were identified: c.523A>G (p.T175A) in exon 6 and c.1769G>A (p.R590Q) in exon 15. The latter was seen three times in our cohort, and both were previously described in HGMD (CM013505 and CM075942). They were reclassified by ClinVar and Varsome as VUS.

LB variants were three: c.579A>G (p.G193=) in exon 7, c.711A>T (p.I237=) in exon 8 and c.1954A>G (p.R652G) in exon 16.

The G193= is a *novel* variant found in heterozygous in a 74-year-old male without LP mutations. Varsome classifies this variant as VUS, because it is located in a mutational hot-spot without benign variations and it is absent from controls in the main population databases, like Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium. The variant is also a synonymous for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site. The patient had also heterozygosity for I237=. This variant is present on HGMD as previously associated polymorphism to ICP (CM031116). We found it in 54/179 patients (30%), while in gnomAD MAF is 0.177 and 0.20 in 1000 Genomes.

The c.1954A>G (p.R652G) in exon 16 was previously associated to PFIC in HGMD. After revision by some authors, now it is considered a disease-associated polymorphism. We found it in 33/179 patients (18%), while in gnomAD MAF is 0.07 in European population and in 1000 Genomes is 0.10 in TSI population.

All identified *ABCB4* variants during the project are in **table 12**.

ABCB4

ex	DNA (c.)	Protein (p.)	HGMD/dbSNP/*novel*	MAF	VarSome	ClinVar	our	*N*
4	217C>G	L73V	CM110312[1] DM?/rs8187788	0.0013	VUS	VUS	VUS	1
6	523A>G	T175A	CM013505[2] DM?/rs58238559	0.01	VUS	B/LB/P	VUS	2
7	579A>G	G193=	*novel*	-	VUS	-	LB	1
8	711A>T	I237=	CM031116[3]/rs2109505	0.17	B	LB	LB	54
9	959C>T	S320F	CM013506[3]/rs72552778	0.000291	LP	VUS/LP	LP	1
10	1091C>T	A364V	CM075946[4]	-	VUS	-	LP	1
11	c.1135_1136insAA	S379KfsX35	*novel*	-	P	-	P	1
11	1028A>G	Y403C	*novel*	-	LP	-	LP	1
13	1442T>G	L481R	CM144425[5]	-	LP	-	LP	1
15	1769G>A	R590Q	CM075942[4] DM?/rs45575636	0.007	VUS	VUS	VUS	3
16	1954A>G	R652G	CM072814[6] DP/rs2230028	0.0737	B	B	LB	33
16	2014A>T	K672*	*novel*	-	P	-	P	1
19	2324C>T	T775M	CM075938[4]/rs148052192	0.001	VUS	-	LP	1

Table 12. All identified *ABCB4* variants in all patients during the project. 1. Colombo (2011) *J Pediatr Gastroenterol Nutr*; 2. Rosmorduc (2001) *Gastroenterology*; 3. Mullenbach (2003) *J Med Genet*; 4. Degiorgio (2007) *Eur J Hum Genet*; 5. Degiorgio (2014) *Eur J Hum Genet*; 6. Liu (2007) *Gastroenterology*; DM?: likely disease-causing mutation; DP: disease-associated polymorphism; *N*: number of cases; benign variants are not shown.

3.3.4 *TJP2* variants

Besides many variants identified with the four-gene panel in *TJP2*, we identified also several variants in the new genetic panel, confirming the high variability of this gene. It is also true that this is a relatively new gene involved in the disease and we are finding always missense variants in a gene for which primarily truncating variants are known to cause the disease.

We identified the VUS c.1411A>G (p.S471G) in exon 8, c.1669G>T (p.A557S) exon 11 and c.2530T>G (p.F844V) in exon 17.

We found the S471G in a 48-year-old male born with Tetralogy of Fallot and with a familial history of hepatic cirrhosis. The patient had also variants in *ABCC2*, *JAG1* and *SLCO1B1*.

The A557S is a very rare variant (MAF: 0.0000791) found in a 44-year-old female with no clinical information.

The F844V is a *novel* variant and was found in a 31-year-old female with no clinical information, who carries one VUS in *MYO5B*.

LB variants identified were V15L in two unrelated patients and the c.1571G>A p.R524K. All identified *TJP2* variants during the project are in **table 13**. Eleven benign variants were also found more times in several patients.

TJP2

ex	DNA (c.)	Protein (p.)	HGMD/dbSNP/*novel*	MAF	VarSome	ClinVar	our	*N*
1	43G>C	V15L	rs73450853	0.0000715	LB	B	LB	2
1	46A>G	K16E	rs151117327	0.0006	VUS	VUS	VUS	1
3	278C>T	T93M	CM1714190[1] DM?/rs138241615	0.00239	VUS	LB/VUS	LP	3
5	860C>T	A287V	rs1266640794	-	LB	-	LB	2
6	1057C>T	R353W	rs759484247	0.00004	VUS	-	LP	2
8	1411A>G	S471G	rs762001932	0.0000352	VUS	-	VUS	1
10	1571G>A	R524K	rs41277901	0.00179	LB	B/LB	LB	3
11	1669G>T	A557S	rs746223152	0.0000791	VUS	VUS	VUS	1
11	1686C>T	V562=	rs781308072	0.00000879	VUS	VUS	LB	1
14	2097G>A	M699I	rs34774441	0.0629	B	B/LB	LB	28
14	2224T>C	S742P	rs35797487	0.000387	B	B/LB	LB	1
17	2530T>G	F844V	novel	-	VUS	-	VUS	1
18	2717T>C	I906T	rs756322608	0.0000176	VUS	-	LP	1
21	3122C>T	S1041F	rs41277907	0.08	B	B/LB	LB	31
22	3465G>A	T1155=	rs745427593	0.0000439	VUS	LB/VUS	LB	2

Table 13. All identified *TJP2* variants in all patients during the project. 1. Dixon (2017) *Sci Rep*; DM?: likely disease-causing mutation; *N*: number of cases; benign variants are not shown.

3.3.5 *ABCC2* variants

Just forty *ABCC2* mutations (missense and nonsense) have been described in HGMD so far. We found 17 variants in *ABCC2* gene. One is LP mutation, five are VUS, five are LB variants and six are benign variants.

The only LP mutation identified is the c.2303G>A (p.R768Q) in exon 18. This variant is present in dbSNP (rs536840524), it is very rare (MAF: 0.0000352) and it is found at an amino acid residue where a different missense change determined to be pathogenic and is described in HGMD (R768W). We found this mutation in a 48-year-old male with high GGT and alkaline phosphatase and having also tree LB variants in *ABCB11*, *TJP2* and *MYO5B* (case 4 in **table 17**).

Among VUS, the c.1109C>T (p.A370V) in exon 9 is also rare (MAF: 0.0000264) and was considered LB/VUS respectively by Varsome and ClinVar. We found it in a 55-year-old male having LB variants in *ABCB4*, *SLCO1B1*, *SLCO1B3* and *TJP2*. No clinical and familial information were provided.

The c.1249_1250inv (p.V417T) in exon 10 is a private variant does not present in population databases. *In-silico* predictions are discordant, and we found it in a 35-year-old female with history of itching and cholestatic disease. The brother of the proband, sequenced in Sanger, has the same variant of the sister. They have also LB variants in *ABCC2* (P871R) and *MYO5B*.

The c.3236G>A (p.R1079Q) in exon 23 was present in a 50-year-old female with other LB variants in *ATP8B1*, *ABCB11*, *ABCB4* and *MYO5B*.

The c.3500T>C (p.V1167A) in exon 25 is a rare variant (MAF: 0.000466) with poor information and predicted to be deleterious/probably damaging by SIFT and PolyPhen. We found it in a 21-year-old female with juvenile lithiasis and cholestatic disease. The patient carried also LB variants in *ABCB11*, *ABCB4*, *SLCO1B1* and *SLCO1B3*.

The c.4430C>T (p.T1477M) in exon 31 is a VUS with discordant predictions of pathogenicity. We found it in a 56-year-old male with another VUS in *GPBAR1* and many LB variants in *ATP8B1*, *ABCB11*, *ABCB4*, *SLCO1B1*, *SLCO1B3* and *TJP2*.

Among LB variants, the c.842G>A (p.S281N) in exon 7, the c.2546T>G (p.L849R) in exon 19, the c.3107T>C (p.I1036T) in exon 23, the c.3563T>A (p.V1188E) in exon 25 and the c.4544G>A (p.C1515Y) in exon 32. Only the last one seems to be involved in the disease because it is reported in HGMD as reducing the pump efflux activity (CM119536).

All identified *ABCC2* variants are in **table 14** and shown as lollipops in **figure 22**.

ABCC2

ex	DNA (c.)	Protein (p.)	HGMD/dbSNP/*novel*	MAF	VarSome	ClinVar	our	*N*
7	842G>A	S281N	rs56131651	0.00794	LB	VUS/LB	LB	2
9	1109C>T	A370V	rs144390464	0.0000264	LB	VUS/LB	VUS	1
10	1249_1250inv	V417T	*novel*	-	LB	VUS	VUS	1
18	2303G>A	R768Q	rs536840524	0.0000352	LP	VUS	LP	1
19	2546T>G	L849R	rs17222617	0.0179	LB	LB	LB	3
23	3107T>C	I1036T	rs45441199	0.00821	LB	B	LB	1
23	3236G>A	R1079Q	rs145672804	0.00821	LB	VUS	VUS	1
25	3500T>C	V1167A	rs140680467	0.000466	VUS	VUS	VUS	2
25	3563T>A	V1188E	rs17222723	0.056	B	LB	LB	16
31	4430C>T	T1477M	rs142573385	0.00173	LB	VUS	VUS	1
32	4544G>A	C1515Y	CM119536[1] FP/rs8187710	0.057	B	LB	LB	17

Table 14. All identified *ABCC2* variants. 1. Elens (2011) *Pharmacogenet Genomics*; FP: in vitro or in vivo functional polymorphism; *N*: number of cases; benign variants are not shown.

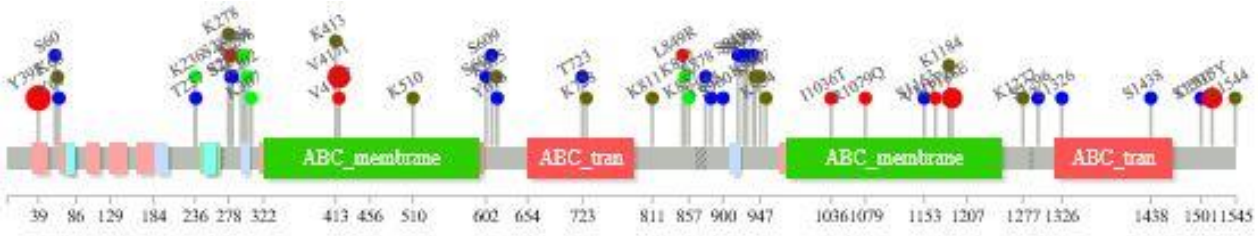


Figure 22. *ABCC2* lollipops. Protein sequence annotations of *ABCC2* variants displayed as red lollipops. Diameters are proportional to the number of patients affected. Blue, green and khaki lollipops indicate respectively phosphorylation, acetylation and ubiquitylation sites. The Y93F is a false positive.

3.3.6 *GPBAR1* variants

GPBAR1 gene encodes for the G Protein-Coupled Bile Acid Receptor 1. This enzyme functions as a cell surface receptor for bile acids. Treatment of cells expressing this GPCR with bile acids induces the production of intracellular cAMP, activation of a MAP kinase signalling pathway, and internalization of the receptor. The receptor is implicated in the suppression of macrophage functions and regulation of energy homeostasis by bile acids. Eight mutations have been described in HGMD so far. Associated phenotypes are primary sclerosing cholangitis, heterotaxy and severe obesity with an early onset.

The gene has only two exons and the coding region starts in exon 2. We found three VUS and one benign variant.

The c.648_649inv (p.A217P) is a VUS for VarSome and does not have any information about frequency in the population. It has the same aminoacid change of one of the eight variants reported in HGMD (CM105602, Hov et al. 2010) as a functional polymorphism with *in vitro* or *in vivo* evidences. However, we found a different nucleotide change, the c.648_649inv that is a deletion/insertion: GG are substituted by CC (delGGinsCC) and this is called "inversion" and our variant is a *novel*.

We found this variant in a 56-year-old male with no clinical information and who carries also another VUS in *ABCC2* and many LB variants in *ATP8B1*, *ABCB11*, *ABCB4*, *SLCO1B1*, *SLCO1B3* and *TJP2*.

The c.856C>G (p.R286G) is another *novel* VUS predicted tolerated and possibly damaging respectively by SIFT and PolyPhen. We found it in a 53-year-old female carries also LB variants in *ABCB11*, *ABCB4*, *JAG1*, *MYO5B* and *SLCO1B1*. The patient has a familial history for biliary calculi and breast cancer and has low GGT but high AP and transaminases and a 8.7 kPa (F2) at Fibroscan (score from F0 to F4, F>2 significant fibrosis).

The c.923G>C (p.R308P) is present in dbSNP (rs200024655) with no frequency information and it has discordant bioinformatics predictions of pathogenicity (SIFT: not tolerated, PolyPhen: benign). We found it in a 36-year-old female with no clinical information. The patient has also LB variants in *ABCB11*, *ABCB4*, *ABCC2* and a VUS in *MYO5B*.

All identified *GPBAR1* variants during the project are in **table 15**.

GPBAR1

ex	DNA (c.)	Protein (p.)	HGMD/dbSNP/*novel*	MAF	VarSome	ClinVar	our	N
2	648_649inv	A217P	*novel*	-	VUS	-	VUS	1
2	856C>G	R286G	*novel*	-	VUS	-	VUS	1
2	923G>C	R308P	rs200024655	-	VUS	-	VUS	1

Table 15. All identified *GPBAR1* variants. *N*: number of cases; benign variants are not shown.

3.3.7 *JAG1* variants

JAG1 gene has 26 exons and produces a protein of 1218 aminoacids. It is associated with Alagille syndrome 1 and Tetralogy of Fallot.

We identified 13 variants: one was pathogenic, one was VUS, two LB and nine benign variants.

The causative mutation c.2029dup (p.H677Pfs*12) in exon 16 is an insertion of one C nucleotide that cause the frameshift mutation H677Pfs*12 with a premature truncated protein. We found it in a 47-year-old male presenting biliary calculi in adolescence with subsequent cholecystectomy and developing a chronic renal failure and HCC (case 5 in **table 17**). No information about his family were provided.

The VUS c.3638G>A p.R1213Q in exon 26 is present in HGMD (CM023743, Kohsaka et al. 2002). It is a rare variant (MAF: 0.0000527) and it is considered VUS by VarSome and ClinVar. It was found in a 54-year-old woman with no clinical information.

The two LB variants were the c.436G>A (p.V146I) in exon 3 and the c.2612C>G (p.P871R) in exon 22, found in thirteen patients. All identified *JAG1* variants during the project are in **table 16**.

JAG1

ex	DNA (c.)	Protein (p.)	HGMD/dbSNP/*novel*	MAF	VarSome	ClinVar	our	*N*
3	436G>A	V146I	rs6040067	0.000193	B	B/LB	LB	1
16	c.2029dup	H677Pfs*12	*novel*	-	P	-	P	1
22	2612C>G	P871R	rs35761929	0.0669	B	B/LB	LB	13
26	3638G>A	R1213Q	CM023743[1]/rs138007561	0.0000527	VUS	VUS	VUS	1

Table 16. All identified *JAG1* variants. 1. Kohsaka (2002) *Hepatology*; *N*: number of cases; Benign variants are not shown.

Table 17

ID	Sex	*ATP8B1*	*ABCB11*	*ABCB4*	*TJP2*	*ABCC2*	*JAG1*	*NOTCH2*	Other genes	Age	GGT[1] (U/L)	AP (U/L)	AST/ ALT	Tot. Bilir. Dir. Bilir. (mg/dL)	Fibroscan (kPa)	Notes
1	F	WT	Ht V444A LB	**Ht S320F LP**	WT	WT	WT	WT	**SLCO1B1**: Hm N130D (LB) **SLCO1B3**: Hm V619= (LB) **VPS33B**: Ht N54= (LB)	30	47	56	14 21	1.7 -	-	LPAC, familial history of biliary calculi
2	F	Ht R952Q LB	Ht V444A LB	**Ht S379KfsX35 LP** Ht R652G LB	Ht M699I LB	WT	Ht P871R LB		**MYO5B:** Ht I612V (LB) **SLCO1B3**: Ht G256A (LB)	58	98	393	31 38	0.77 0.27		ICP, juvenile cholelithiasis, history of familial hepatic disease
3	F	WT	Ht V444A LB	**Ht L481R LP**	WT	WT	WT	WT	Benign variants	33	278	300	33 51	0.58 0.13	-	ICP, juvenile cholelithiasis, history of familial hepatic disease
4	M	WT	Ht V444A LB	WT	Ht S1041F LB	**Ht R768Q LP**	WT	WT	**MYO5B**: Ht R918H (LB)	25	690	382	141 228	2.78 0.57	-	Gilbert syndrome, jaundice, cholecystectomy
5	M	Ht R952Q LB Ht V1197L VUS	Hm V444A LB			Ht V1188E LB Ht C1515Y LB	**Ht H677Pfs*12 P**	WT	**CLDN1**: Ht K189E (VUS) **MYO5B**: Ht G1321E (LB) **SLCO1B1**: Ht N130D (LB) Ht V174A (LB) Ht L643F (LB) **SLCO1B3**: Ht G256A (LB)	47	104	180	39 37	1.99	F4	HCC, chronic renal failure, biliary calculi, cholecystectomy, no information about family
6	M	WT	Ht V444A LB	WT	WT	Ht V1188E LB Ht C1515Y LB	WT	**Ht R2003* P**	**SLCO1B1**: Ht N130D (LB) **SLCO1B3**: Ht G256A (LB)	26	14	200	66 208	0.75 0.32	F1	ileal volvulus syndrome, no familiarity

Table 17. P/LP mutation identified by 15-gene analysis. 1: within one year from genetic test**.**

3.3.8 *MYO5B* variants

We identified 22 variants: 6 VUS, 7 likely benign and 9 benign variants.

Among VUS, the c.1069T>C (p.Y357H) in exon 10 is a very rare because it was seen just in one European female (MAF: $9*10^{-6}$) before us. It is located in a mutational hot spot in Myosin head functional domain without benign variation. We found it in a 36-year-old female with no clinical information. LB variants were found in *ABCB11*, *ABCB4*, *ABCC2* and *MYO5B*.

The c.1834A>G (p.I612V) in exon 15 is also a rare variant and located in the functional domain of the Myosin head and we found it in a 67-year-old female with low GGT, a familial history of hepatic disease, ICP and juvenile cholelithiasis. The patient had also LB variants in *ABCB11*, *ABCB4*, *ABCC2*, one LB and the following VUS in *MYO5B*.

The c.2123G>A (p.R708Q) in exon 18 has an allele frequency of 0.0009, bioinformatic predictions are deleterious but still poor information are available to assess the role of this variant. We found this variant in two unrelated patients. The second is a 21-year-old male with high GGT and AP with a Fibroscan score of 10.7 kPa (F3) and neonatal jaundice. Multiple LB variants were found: *ABCB11*, *ABCB4*, *ABCC2*, *MYO5B*, *SLCO1B1* and *SLCO1B3*.

The c.2343G>C (p.K781N) in exon 19 was found in a 71-year-old female with high ALT, GGT and AP and a Fibroscan score of 6 kPa (F1) and no familiarity. The patient had also LB variants in *ABCB11* and *ABCB4*.

The c.5144A>G (p.N1715S) in exon 38 is a *novel* variant predicted to be pathogenic by the main bioinformatic tools. We found it in a 68-year-old male (BMI = 29) with high bilirubin, AP and slightly GGT. The patient had also the higher Fibroscan score in the cohort: 75 kPa (F4 starts from 13 kPa). The patient's haplotype is very interesting, with another VUS in *ATP8B1* and different LB variants in *ABCB11*, *ABCB4*, *ABCC2*, *JAG1*, *NOTCH2*, *SLCO1B1* and *TJP2*.

The last VUS, c.5411G>C (p.R1804P) in exon 40 is also rare (MAF: 0.0000443) and we found it in a 31-year-old female with no clinical information, but carrying also LB variants in *ABCB11*, *ABCB4*, *ABCC2* and the novel VUS in *TJP2* **(¶ 3.3.4).**

Likely benign variants identified were: the c.145G>C (p.E49Q) and c.176T>C (p.L59P) in exon 3, the c.921G>T (p.K307N) in exon 8, the c.2753G>A (p.R918H) in exon 21, the c.3645C>T (p.D1215=) in exon 28, the c.4145C>T (p.T1382M) in exon 31 and c.5395-6C>T in IVS39. All identified *MYO5B* variants during the project are in **table 18**.

MYO5B

ex	DNA (c.)	Protein (p.)	HGMD/dbSNP/*novel*	MAF	VarSome	ClinVar	our	N
3	145G>C	E49Q	rs141998504	0.0137	LB	LB	LB	1
3	176T>C	L59P	rs78626055	0.0258	B	B/LB	LB	4
8	921G>T	K307N	rs17659179	0.0633	B	B/LB	LB	4
10	1069T>C	Y357H	rs772554828	0.000008932	VUS	-	VUS	1
15	1834A>G	I612V	rs1267348047	0.00000883	VUS	-	VUS	1
18	2123G>A	R708Q	rs201670299	0.0009	VUS	-	VUS	2
19	2343G>C	K781N	rs61737448	0.00204	LB	-	VUS	1
21	2753G>A	R918H	rs2298624	0.139	B	LB	LB	12
28	3645C>T	D1215=	rs200991250	0.000339	VUS	-	LB	1
31	4145C>T	T1382M	rs145598498	0.00365	B	LB	LB	1
38	5144A>G	N1715S	*novel*	-	VUS	-	VUS	1
IVS39	5395-6C>T	-	rs140275825	0.00459	LB	VUS	LB	1
40	5411G>C	R1804P	rs201080553	0.0000443	VUS	VUS	VUS	1

Table 18. All identified *JAG1* variants in all patients during the project. *N*: number of cases; benign variants are not shown.

3.3.9 *NOTCH2* variants

NOTCH2 gene has been associated to Alagille syndrome 2 and Hajdu-Cheney syndrome so far. We identified 11 variants: 1 P/LP mutation, 1 VUS, 4 LB and 5 benign variants.

The c.6007C>T (p.R2003*) in exon 33 is a nonsense mutation described in HGMD (CM1110718, Kamath et al. 2012) as a disease-causing mutation associated to Alagille syndrome. We identified this mutation in a 26-year-old male born with ileal volvulus syndrome and now having high GPT and AP, low GTT F1 at Fibroscan and carrying other LB variants in *ABCB11*, *ABCC2*, *SLCO1B1* and *SLCO1B3* (case 6 in **table 17**).

The c.2765A>G (p.N922S) in exon 18 is a VUS not found in dbSNP. This missense variant was seen in two European individuals (as reported by gnomAD), so the MAF is very low

(0.00001553). The variant is classified as VUS because it is in a gene for which primarily truncating variants are known to cause disease and multiple lines of computational evidence support a deleterious effect on the gene or gene product. We found it in a patient with high GTP, AP, low GGT and 6.7 kPa at Fibroscan (F1). The grandmother of the proband had NAFLD. The patient had also LB variants in *ATP8B1*, *ABCB11*, *JAG1*, *SLCO1B1* and *TJP2*.

The LB variants were c.1918G>T (p.V640F) in exon 12, c.3625T>G (p.F1209V) in exon 22, c.3980A>G p.D1327G in exon 24 and c.6223G>A p.V2075M in exon 34. They are all rare variants. All identified *NOTCH2* variants during the project are in **table 19**.

NOTCH2

ex	DNA (c.)	Protein (p.)	HGMD/dbSNP/*novel*	MAF	VarSome	ClinVar	our	N
12	1918G>T	V640F	rs782463006	0.0000353	LB	-	LB	1
18	2765A>G	N922S	*novel*	0.00000882	VUS	-	VUS	1
22	3625T>G	F1209V	CM156998[1] DM?/rs147223770	0.00382	B	LB	LB	3
24	3980A>G	D1327G	CM165062[2] DM?/rs61752484	0.00763	B	B	LB	3
33	6007C>T	R2003*	CM1110718[3]/rs312262801	-	P	P	P	1
34	6223G>A	V2075M	rs150516342	0.00198	B	LB	LB	1

Table 19. All identified *NOTCH2* variants in all patients. 1. Aloraifi (2015) *FEBS J*; 2. Priest (2016) *PLoS Genet* 3. Kamath (2012) *J Med Genet*; DM?: likely disease-causing mutation; *N*: number of cases; benign variants are not shown.

3.3.10 Other genes

The *NR1H4* gene has 3 mutation recorded in HGMD, it is a recently PFIC-associated gene and for this reason its knowledge is still poor. Our results are in accordance with literature. We found just one variant: the c.548T>C (p.M183T) in exon 3. It is a rare variant (MAF: 0.00282), computational analysis supports a deleterious effect on the gene but ClinVar refers to it as likely benign variant. For this reason, it is very difficult to assess the role of this variant. We found it in a 67-year-old male with PFIC and a score of 28 kPa at Fibroscan (F4). The patient had also LB variants in *ABCB11*, *ABCB4*, *SLCO1B1* and *TJP2*.

CLDN1 is a little gene with a transcript length of 3481 bp divided into 4 exons. Its translation produces a 211 aminoacid long protein. The knowledge about this gene is still poor, in fact,

there are only two nonsense mutations described in HGMD so far. The phenotypes associated with these two mutations are ichthyosis, alopecia and sclerosing cholangitis.

We identified just one VUS and three benign variants.

The only variant identified in *CLDN1* gene is the c.565A>G p.K189E in exon 4. The K189E is very rare, with a MAF in European population of 0.000149 and it is predicted to be deleterious and possibly damaging by SIFT and PolyPhen. We found it in a 47-year-old male having a severe mutation in *JAG1* (§ **3.3.7**) and other variants.

The *SLC25A13* gene causes intrahepatic cholestasis by citrin deficiency (NICCD). In this gene we found just 4 variants: 1 VUS, 1 LB and 2 benign variants.

The c.190G>A (p.V64M) in exon 3 is a *novel* variant does not present in HGMD. The aminoacid change from Valine to Methionine is predicted to affect the protein by SIFT and PolyPhen, however in this gene primarily truncating variants are known to cause disease. We found it in a 47-year-old female with ICP and normal GGT. Fibroscan analysis revealed a score of 7.9 kPa (F2).

The synonymous c.609A>G (p.L203=) is defined VUS by VarSome. Because it is a synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site, we refer to it as LB variant.

In the *SLCO1B1* gene we found 4 LB variants and 4 benign variants.

The c.317T>C (p.I106T) in exon 4, the c.388A>G (p.N130D) in exon 5, c.521T>C (p.V174A) in exon 6 and c.1929A>C (p.L643F) in exon 15.

The frequency of the SNP c.388A>G (p.N130D) in our cohort is higher than the population one (MAF: 0.612 vs 0.404). This variant is described in HGMD as disease-associated polymorphism with additional functional evidence and an altered transport activity (CM043776). In the present study we demonstrated that this SNP is significantly associated to normal AST levels ($p = 0.047$) and low Fibroscan scores ($p = 0.044$) by contingency tables (¶ **3.4**).

In the *SLCO1B3* gene we found 4 LB and 6 benign variants. SNPs were c.767G>C (p.G256A) in exon 9, c.1074C>T (p.Y358=) in exon 10, c.1347A>G (p.A449=) in exon 12 and c.1857A>T

(p.V619=) in exon 15. For SNP c.767G>C, p.G256A (rs60140950), the frequency in our cohort is higher than the population one (MAF: 0.275 vs 0.159). By the de Finetti association analysis, we shown that homozygous patients carrying Alanine in both allele have an increased risk for the disease: OR = 5.053 (p = 0.02877) for CC genotype (¶ **3.4**).

Gissen et al. (2004) identified mutations in the *VPS33B* gene as the cause of the arthrogryposis-renal dysfunction-cholestasis syndrome. We found nine variants in this gene: three VUS, two LB and four benign variants. Among VUS, the c.163G>A (p.V55I) in exon 2, c.944G>A (p.R315Q) in exon 13 and the c.1170+5G>A in IVS15.

The V55I is very rare (MAF: 0.0000176) and present in dbSNP. It was found in a 54-year-old woman with no clinical information. The patient had also a VUS in *JAG1* and LB variants in *ABCB11*, *ABCC2*, *SLCO1B1* and *SLCO1B3.*

The R315Q was found in a 56-year-old male with low GGT, HCC and a score of 12.8 kPa at Fibroscan (F3). The patient had also SNPs in seven genes: four LB variants in *ABCB11*, variants in *ABCC2*, *MYO5B*, *NOTCH2*, *SLCO1B1*, *SLCO1B3* and *TJP2.*

The c.1170+5G>A is predicted to alter the WT splicing donor site by HSF. We found this variant in a 44-year-old male with cryptogenetic cholestasis, high GGT and AP and a score of 4.9 kPa at Fibroscan (F0/F1). He showed a reduced expression of MDR3 protein at immunohistochemistry. *ABCB11*, *ABCB4*, and *NOTCH2* LB variants were also identified in this patient. All variants identified in other genes during the project are in **table 20**.

Other genes

ex	DNA (c.)	Protein (p.)	HGMD/dbSNP/novel	MAF	VarSome	ClinVar	our	*N*
NR1H4								
3	548T>C	M183T	rs61755050	0.00629	VUS	LB	VUS	1
CLDN1								
4	565A>G	K189E	rs199998460	0.000149	VUS	-	VUS	1
SLC25A13								
3	190G>A	V64M	*novel*	-	VUS	-	VUS	1
6	609A>G	L203=	rs751566503	0.0000176	VUS	-	LB	1
SLCO1B1								
4	317T>C	I106T	rs200227560	0.00074	LB	-	LB	1
5	388A>G	N130D	CM043776[1] DFP/rs2306283	0.404	B	B/VUS	LB	51
6	521T>C	V174A	CM043777[2] DFP/rs4149056	0.157	B	B/drug response	LB	30
15	1929A>C	L643F	CM120190[3] FP/rs34671512	0.05	B	LB	LB	8
SLCO1B3								
9	767G>C	G256A	rs60140950	0.159	B	B/LB	LB	23
10	1074C>T	Y358=	rs145036538	0.000159	VUS	-	LB	1
12	1347A>G	A449=	rs79382866	0.00659	VUS	VUS	LB	1
15	1857A>T	V619=	rs143827641	0.000961	VUS	VUS	LB	2
VPS33B								
2	162T>C	N54=	rs769227036	0.0000439	VUS	-	LB	1
2	163G>A	V55I	rs763310786	0.0000176	VUS	-	VUS	1
13	944G>A	R315Q	rs145303578	0.00135	VUS	VUS	VUS	1
15	1166G>A	R389Q	rs145070485	0.00238	LB	-	LB	1
15	1170+5G>A	IVS15	rs201431055	0.000114	VUS	VUS	VUS	1

Table 20. All identified variants in *NR1H4, CLDN1, SLC25A13, SLCO1B1, SLCO1B3* and *VPS33B* in all patients during the project. 1. Iwai (2004) *Pharmacogenetics*; 2. Pasanen (2008) *Pharmacogenet Genomics*; 3. Ramsey (2012) *Genome Res*; DFP: disease-associated polymorphism with additional functional evidence; FP: in vitro or in vivo functional polymorphism; *N*: number of cases; Benign variants are not shown.

3.4 Statistical analysis

Baseline features of the population analysed are shown in **table 21.** Laboratory findings were available for 56/80 patients, Fibroscan for 43 patients and BMI for 55. Familiarity for the disease, DIC and itching history, ICP history, neonatal jaundice and juvenile cholelithiasis episodes were not fully available for all patients and were not used for statistical analysis.

Table 21. Baseline features of the 80 patients in fifteen-gene analysis

Sex	41 male (51,3%)
	39 female (48,8%)
Age	47.7 ± 15 (2-78)
Laboratory	
AST (U/l)	39.0 (14-141)
ALT (U/l)	63.2 (11-378)
GGT (U/l)	146.2 (10-687)
AP (U/l)	251.2 (56-719)
Tot. Bilirubin (mg/dl)	1.1 (0.30-6.86)
Dir. Bilirubin (mg/dl)	0.24 (0.08-0.65)
Albumin (g/dl)	4.3 ± 0.5 (3.1-5.2)
Fibroscan (kPa)	10.2 (2.7-75)
BMI	25.3 (19-38)

AST: aspartate transaminase; ALT: alanine transaminases; GGT: gamma-glutamyl transferase; AP: alkaline phosphatase; kPa: kilopascal; BMI: body mass index. In the right column mean, standard deviation and range are shown.

For laboratory data disease markers cut-offs were used to test the association between SNPs and biochemical parameters as follows: AST over 35 U/l, ALT over 45 U/l, GGT over 50 U/l and AP over 220 U/l. A BMI over the normal (> 25) was used as cut-off for this variable. Fibroscan scores were divided into two classes: F3 and F4 fibrosis (high fibrosis) against F1 and F2 fibrosis (low scores). These two classes were used to test if SNPs were associated with high rates of hepatic fibrosis. With these analyses, some SNPs resulted as a protective or risk factors against the disease. Only statistically significant Pearson's Chi squares are reported in **tables 22** and **23**. **Table 22** shows protective SNPs.

Patients carrying one allele of the *ABCB11* variant c.270T>C (rs4148777) or the *NOTCH2* variant c.7341T>A (rs6685892) (there were not homozygous patients for these variants) seem to have a normal BMI. The *ABCB11* variant c.957A>G (rs7563233) is significantly associated to normal GGT levels. Homozygosis for the benign variant c.388A>G (N130D, rs2306283) in

SLCO1B1 gene was present in patients who did not have AST elevated, while just one allele of this variant could protect from liver fibrosis.

Table 22. SNPs with a potential protective effect by contingency tables

Gene	DNA/Protein	dbSNP	Variable tested	Pearson's Chi-square
ABCB11	c.270T>C/p.F90=	rs4148777	High BMI	$p = 0.043$
ABCB11	c.957A>G/p.G319=	rs7563233	High GGT	$p = 0.041$
NOTCH2	c.7341T>A/p.G2447=	rs6685892	High BMI	$p = 0.024$
SLCO1B1	c.388A>G/p.N130D	rs2306283	High AST	$p = 0.047$
			High Fibroscan	$p = 0.044$
TJP2	c.1230A>G/p.L410=	rs17062695	High ALT	$p = 0.001$

Table 23 shows SNPs that could act as potential risk factors.

High ALT levels are associated with patients having at least one mutated C allele of the c.571T>C in *SLCO1B1* (rs4149057): in fact, hetero and homozygous patients for this variant were doubly represented in ALT > 35 U/l category. The *TJP2* variant Q159K (rs41305539) was also associated to high ALT. The presence of at least one allele of the IVS32 variant of the *MYO5B* gene c.4315+5G>C (rs488890) was associated to high GGT levels with a $p = 0.025$. The M1688V variant (rs112417235), still in *MYO5B* gene, was associated to high AP levels ($p = 0.034$). Homozygosis for the c.1197A>G (rs2301629) of the *SLC25A13* gene and the presence of one allele of the c.2808C>T (rs2282336) and c.2820G>A (rs2282336) in *TJP2* are also associated to high AP levels with $p = 0.004$ and $p = 0.028$ (for both in *TJP2*) respectively. The c.2808C>T and c.2820G>A appear always together in patients. One allele of the c.1249G>A, p.V417I in *ABCC2* (rs2273697), c.1230A>G (rs17062695) and c.2820G>A in *TJP2* are associated to a high BMI. F3 and F4 fibrosis stages are significantly associated to the presence of one allele of c.1230A>G in *TJP2* and homozygous c.411G>A variants (rs11045818) in *SLCO1B1* with $p = 0.003$ and $p = 0.043$ respectively.

Table 23. SNPs with a potential risk factor by contingency tables

Gene	DNA/Protein	dbSNP	Tested variable	Pearson's Chi-square
SLCO1B1	c.571T>C/p.L191=	rs4149057	High ALT	$p = 0.017$
TJP2	c.475C>A/p.Q159K	rs41305539	High ALT	$p = 0.022$
MYO5B	c.4315+5G>C	rs488890	High GGT	$p = 0.025$
MYO5B	c.5062A>G/p.M1688V	rs112417235	High AP	$p = 0.034$
SLC25A13	c.1197A>G/p.L399=	rs2301629	High AP	$p = 0.004$
TJP2	c.2808C>T/p.T936=	rs2282336	High AP	$p = 0.028$
TJP2	c.2820G>A/p.A940=	rs2095876	High AP	$p = 0.028$
			High BMI	$p = 0.017$
ABCC2	c.1249G>A/p.V417I	rs2273697	High BMI	$p = 0.034$
TJP2	c.1230A>G/p.L410=	rs17062695	High BMI	$p = 0.013$
			High Fibroscan	$p = 0.003$
SLCO1B1	c.411G>A/p.S137=	rs11045818	High Fibroscan	$p = 0.043$

Due to the absence of normality of the distribution for several continuous parameters checked by the Kolmogorov-Smirnov test, non-parametric test was used accordingly to assess differences between genotypes: independent samples Mann-Whitney U test for testing differences with the mutated allele and Kruskal-Wallis one-way ANOVA test for analysing genotypes when wild-type, heterozygotes and homozygotes were comparable. Statistically significant SNPs associated to continuous parameters are reported below. SNPs with a potential protective effect are shown in **table 24**.

Table 24. SNPs with a potential protective effect by Nonparametric Tests

Gene	DNA/Protein	dbSNP	Association	*p* value	Previously
ABCB11	c.270T>C/p.Phe90=	rs4148777	≤BMI	0.036	*
SLCO1B3	c.699G>A/M233I	rs7311358	≤BMI	0.010	
	c.1557A>G/p.A519=	rs2053098	≤BMI	0.014	
ABCB4	c.1954A>G/p.R652G	rs2230028	≤tot. Bil.	0.047	
MYO5B	c.4315+5G>C	rs488890	≤tot. Bil.	0.033	
	c.4315+5G>C	rs488890	≤dir. Bil.	0.008	
ABCC2	c.1249G>A/p.V417I	rs2273697	≤Albumin	0.037	
TJP2	c.2097G>A/p.M699I	rs34774441	≤Albumin	0.014	
	c.1230A>G/p.L410=	rs17062695	≤ALT	0.012	*

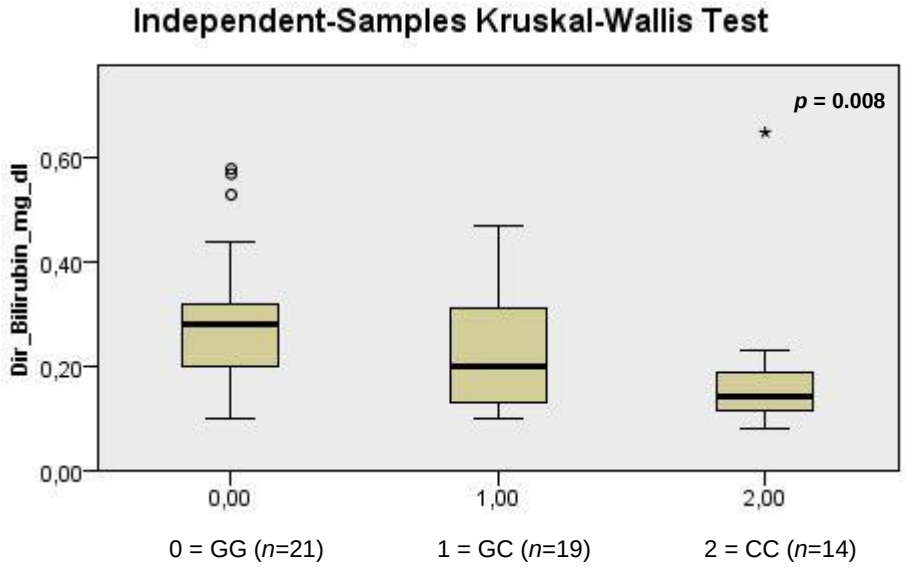
Table 24 continued. SNPs with a potential protective effect by Nonparametric Tests.

Gene	DNA/Protein	dbSNP	Association	*p* value	Previously
ABCC2	c.4290G>T/p.V1430=	rs1137968	≤GGT	0.013	
	c.4488C>T/p.H1496=	rs8187707	≤GGT	0.019	
TJP2	c.1230A>G/p.L410=	rs17062695	≤GGT	0.028	

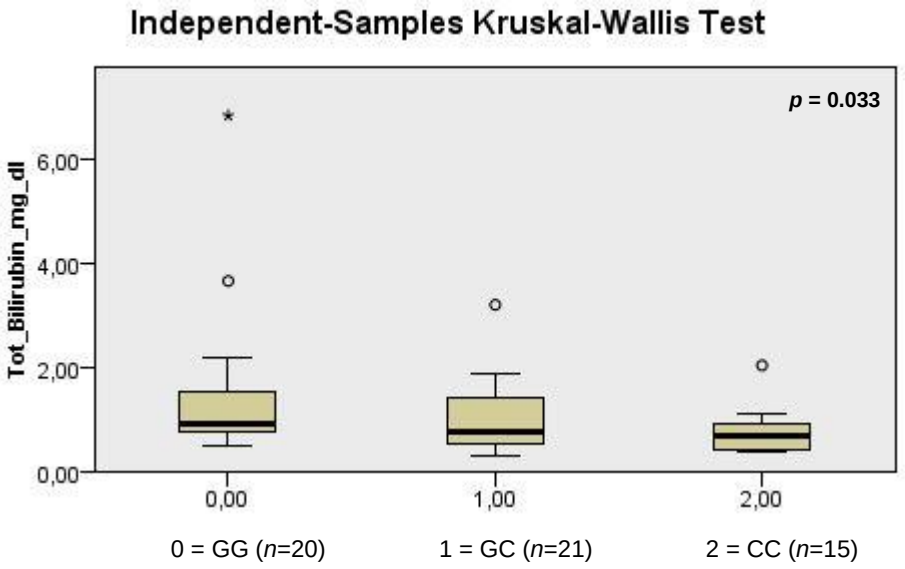
*** Previously associated by contingency tables' Chi-square**

Two SNPs, c.270T>C (p.Phe90=) in *ABCB11* and c.1230A>G (p.L410=) in *TJP2* were previously tested for the same category by contingency tables (see above) and were confirmed by the latter analysis as associated to a low BMI and low ALT levels, respectively. Among potential protective SNPs, the C allele of the c.4315+5G>C in *MYO5B* gene (rs488890) was found to act as a protective factor for both direct and total bilirubin levels, showing a trend from wild-type, hetero and homozygous population in the whisker plots below.

***MYO5B* c.4315+5G>C IVS32 rs488890**



***MYO5B* c.4315+5G>C IVS32 rs488890**



SNPs with a potential risk factor are shown in **table 25**.

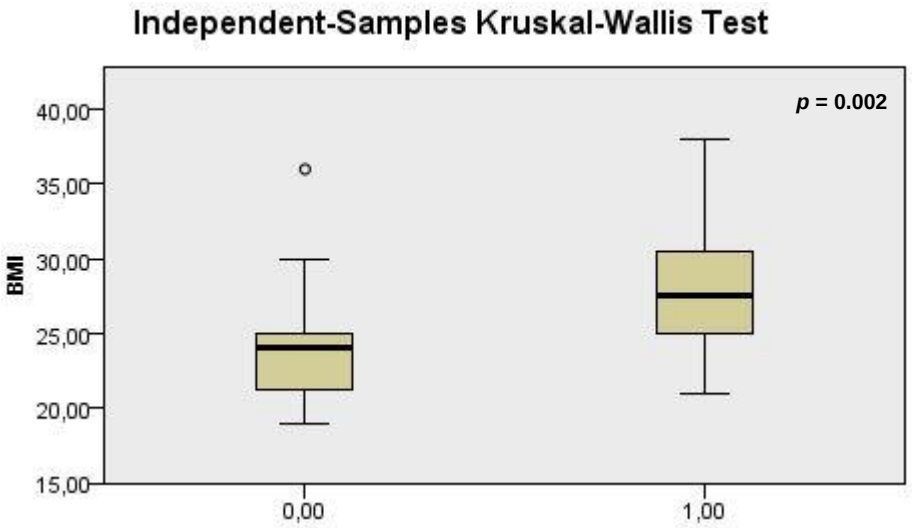
Table 25. SNPs with a potential risk by Nonparametric Tests

Gene	DNA/Protein	dbSNP	Association	*p* value	Previously
ABCC2	c.1249G>A/p.V417I	rs2273697	>BMI	0.002	*
NOTCH2	c.3625T>G/p.F1209V	rs147223770	>BMI	0.046	
VPS33B	c.151C>A/p.R51=	rs11542638	>BMI	0.033	
SLCO1B3	c.1557A>G/p.A519=	rs2053098	>ALT	0.018	
SLCO1B3	c.1833G>A/p.G611=	rs3764006	>ALT	0.018	
TJP2	c.1230A>G/p.L410=	rs17062695	> Fibroscan	0.025	*
TJP2	c.2808C>T/T936=	rs2282336	>AP	0.003	*
TJP2	c.2820G>A/p.A940=	rs2095876	>AP	0.003	*

*** Previously associated by contingency tables' Chi-square**

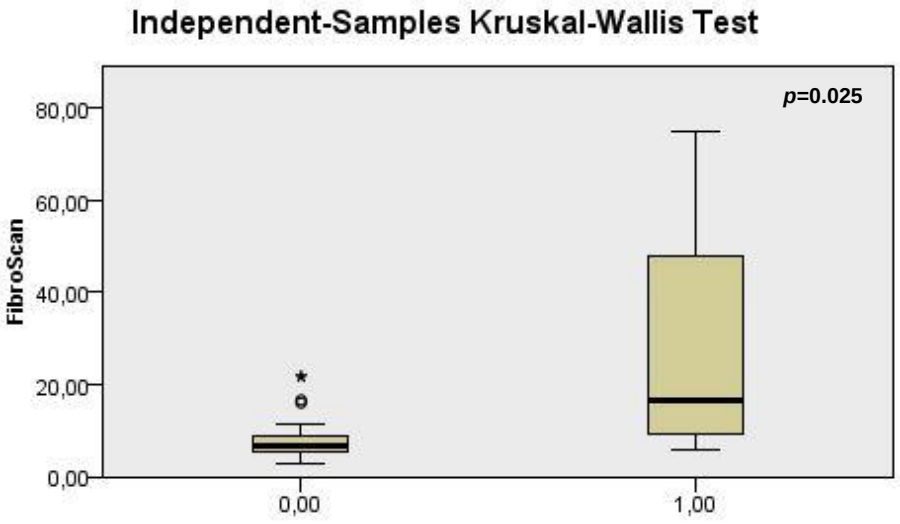
Among those SNPs, rs2273697 in *ABCC2* and rs17062695, rs2282336 and rs2095876 in *TJP2* were previously tested for the same category by contingency tables (see above) and were confirmed by nonparametric test as associated to high BMI, high fibrosis (F3/F4) and high AP levels, respectively, as shown in the whisker plots below. The rs2282336 and rs2095876 have the same *p* (0.003) at the Kruskal-Wallis ANOVA analysis because they are always both mutated in patients (rs2095876 not shown).

***ABCC2* c.1249G>A p.V417I rs2273697**



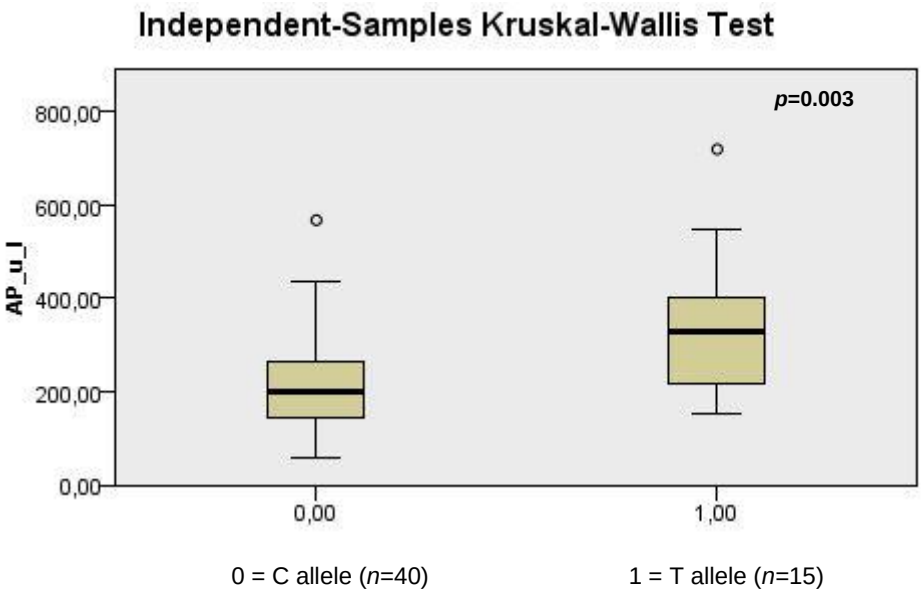
0 = T allele (*n*=36) 1 = C allele (*n*=18)

***TJP2* c.1230A>G p.Leu410= rs17062695**



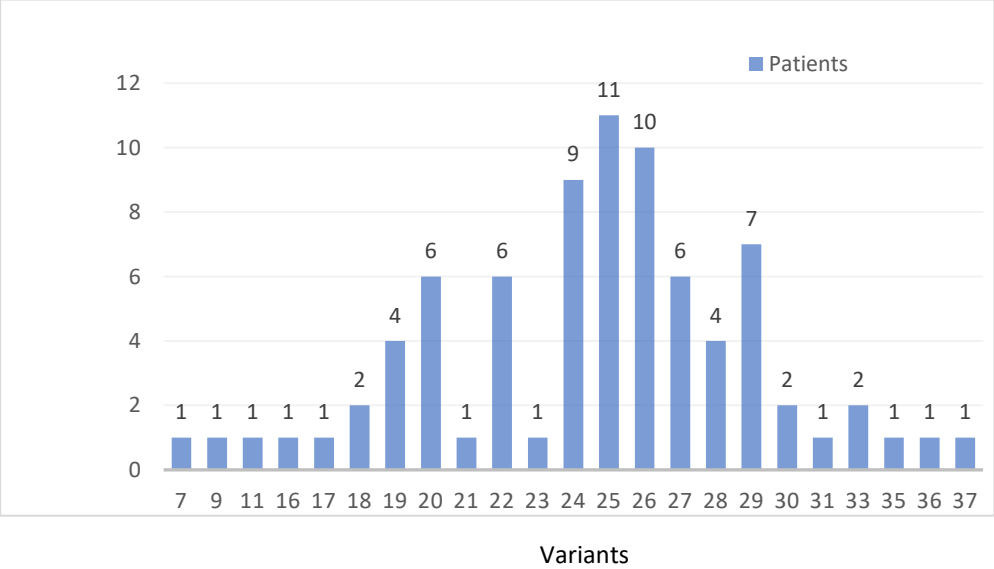
0 = AA (*n*=46) 1 = AG (*n*=9)

***TJP2* c.2808C>T p.T936= rs2282336**



Six patients have P/LP mutations, 30/80 (37,5 %) patients have at least one VUS, but all patients have several SNPs, ranging from one patient having 7 SNPs to another one having 37 SNPs. The distribution of the variants' number per patient is shown in **figure 23**.

Figure 23



The graph shows a Gaussian distribution. Interestingly, the patient carrying 37 SNPs had also the highest Fibroscan score in the population studied. The ANOVA analysis between the variant's number per patient and Fibroscan scores revealed a significant $p = 0.000$. The more variants are present, the more probability to have high liver fibrosis.

3.4.1 Case-control studies

Case-control studies were performed by the de Finetti analysis.
This analysis is part of the tests for deviation from Hardy-Weinberg equilibrium and tests for association. Several genotypes resulted in an increase or diminished risk for developing the disease, as shown in **table 26**.

Table 26. SNPs resulted from de Finetti analysis

Gene	DNA/Protein	dbSNP	*p* value	OR	notes
ABCB11	c.3084A>G/p.A1028A	rs497692	0.01993	1.603	G allele risk
JAG1	c.588C>T/p.C196=	rs1801138	0. 01293	2.808	T allele risk
JAG1	c.765C>T/p.Y255=	rs1131695	0.00698	1.696	T allele risk
SLCO1B3	c.699G>A/p.M233I	rs7311358	0.00371	3.340	AA genotype risk
SLCO1B3	c.767G>C/p.G256A	rs60140950	0.02877	5.053	CC genotype risk
ABCC2	c.3972C>T/p.I1324=	rs3740066	0.01960	0.698	T allele protection
CLDN1	c.369T>C/p.G123=	rs9869263	0.00356	0.464	C allele protection
SLCO1B3	c.334T>G/p.S112A=	rs4149117	0.00326	0.698	G allele protection
VPS33B	c.1540G>A/p.G514S	rs11073964	0.00001	0.383	A allele protection

OR: Odds Ratio, from 0 to 1 potential protection from the disease, over 1 potential risk for the disease.

In *ABCB11*, the c.3084A>G, rs497692, showed a $p = 0.01993$ and a common Odds ratio (OR) of 1.603, indicating a slightly increased risk for the presence of the G allele.
In *JAG1*, the c.588C>T, rs1801138 represents a risk factor for patients who carry the T allele (Odds ratio = 2.808 with a $p = 0. 01293$); another SNP, the c.765C>T, rs1131695, cause a potential risk factor for the T allele with an Odds ratio of 1.696 with a $p = 0.00698$.

In *ABCC2*, the c.3972C>T, rs3740066 showed a protective effect for the presence of the G allele (OR = 0.698).

In *CLDN1*, the c.369T>C, rs9869263, resulted in a protective effect. In fact, patients carrying the C allele have a halved risk to develop the disease (Odds ratio = 0.464 with a $p = 0.00356$).

In *VPS33B* gene, with an Odds ratio = 0.383 and a $p = 0.00001$, the A allele resulted slightly protective in the c.1540G>A (p.G514S), rs11073964.

In the *SLCO1B3* gene, three SNPs have significant deviation from Hardy-Weinberg equilibrium and an association with the disease: the c.334T>G, rs4149117 had a slight protective effect for the presence of the G allele (Odds ratio = 0.698, $p = 0.00326$); the c.699G>A, rs7311358 is a risk factor for population carrying the AA genotype with a $p = 0.00371$ and an OR = 3.340; in the SNP c.767G>C (p.G256A) rs60140950, only homozygosis (CC genotype) resulted in a very increased risk (OR = 5.053, $p = 0.02877$).

3.5 Conclusions

PFIC are usually considered pediatric diseases related to liver failure. Mutations in *ATP8B1*, *ABCB11*, *ABCB4* and *TJP2* have been associated to several cholestatic disorders. Hepatocellular carcinoma and cholangiocarcinoma to *ABCB11* and *TJP2*, ICP is linked to *ATP8B1*, *ABCB11* and *ABCB4*, LPAC to *ABCB4*, DIC to *ABCB11* and *ABCB4* and BRIC to *ATP8B1* and *ABCB11.* Those mutations, genes and phenotypes are related and may coexist in the same patient. Furthermore, all four genes may be responsible for progressive forms of cholestasis or some benign forms could evolve to them.

Of interest, we found a mean age of 36 years in subjects with disease-causing variants to confirm that P/LP mutations are not present only in children. Few studies linked mutations in PFIC genes with non-progressive diseases, especially in heterozygous subjects. In our cohort of patients with at least one disease-causing mutation, only four patients were homozygous or compound heterozygous for disease-causing variants in the same gene.

Colombo et al. 2011, described pathogenic mutations in *ABCB4* in about a quarter of asymptomatic children where cholestatic disease was incidentally discovered via liver enzyme abnormalities and, not surprisingly, some of these patients carried a single heterozygous mutation. Gordo-Gilart et al. 2016, reported similar results in another cohort of pediatric patients in whom defects in a single allele of *ABCB4* were identified in 9/67 subjects. Some authors suggested that cryptogenic cholestasis in adults should be added to the spectrum of conditions associated with *ABCB4* mutations.

In our cohort, 12% of patients had at least one P/LP mutation and 20% of them presented a liver disorder in adulthood, suggesting once again that variants in PFIC genes are not exclusive to cholestatic diseases in childhood, but can also start with benign forms and lead to progressive forms and ESLD after years.

Patients with P/LP mutations had higher Fibroscan scores, higher liver stiffness and a more aggressive phenotypes, but no significant correlation was observed.

Many patients had multiple variants in the four classic PFIC genes. The same trend has been observed with the fifteen-gene panel. Patients in our cohort ranging from 7 to 35 variants and we demonstrate that also some SNPs can be significantly involved in biochemical parameters (such as high AST, ALT and GGT levels) and in phenotypic features (high liver fibrosis and tendency to have higher weights) that could accelerate the progression to liver failure. It is thus believable to hypothesize a synergistic haplotype effect in determining different cholestasis phenotypes and overlapping features, or specific interactions with environmental factors, in particular certain drugs. A reduction of bile flow represents a pathophysiological state and can be critical for metabolism of many drugs. This is supported by MDR3 and BSEP deficiencies being known to be associated to different levels of DIC. Examples are SNPs V444A and M677V in *ABCB11* and I237= in *ABCB4*. We found V444A in 83.3% of patients (38% in homozygous state): this benign variant was reported with AF in the general population of 56.9 and of 65.9% in cholestatic patients with no BSEP disease-causing mutations. It is associated with DIC, ICP and reduced expression of BSEP levels in previous studies.

All the fifteen genes had at least one SNP more frequent in our cohort than in TSI population from 1000 Genome. Some of them had a great deviation from HW equilibrium and a statistically significant association with the disease. Among genes with a high number of SNP in relation to their total variants' number we found *CLDN1*, *SLCA25A13*, *SLCO1B1* and *SLCO1B3*. These genes are all involved in cell adhesion and organic anion transport and it is possible that the simultaneous presence of more variants at the same time could also changes the normal activity of this protein chain.

TJP2 resulted the most affected gene with 25 variants identified. *ABCB11*, *MYO5B* and *ABCB4* also had a high number of variants (22 the first two genes and 20 the latter) but only *ABCB11* and *ABCB4* had also different P/LP mutations. The *ATP8B1* gene had the higher VUS rate and the lowest rate of SNPs.

We did not find mutations or SNPs in the recently PFIC5 associated gene *NR1H4*: we found just one VUS. One explanation could be that this gene is involved mainly in pediatric or childhood phenotypes, while here we have a young-adult and adult population cohort.

After all, our detection rate is comparable to the literature and to other studies proposing multi-gene panels. Some research groups are able to facilitate genetic diagnosis in children with overt intrahepatic cholestasis phenotypes, but it remains elusive in adult populations. The failure to find pathogenic mutations in the major part of our cohort may be explained by the following, not mutually exclusive, reasons. First, the deep intronic regions, promoters and untranslated regions (5'-3'UTR) of all genes have never been analysed by NGS-method. Apart for MLPA analysis of the *ABCB4* gene, large genomic deletions of the other eleven genes have not been analysed. In fact, NGS sequencing does not allow to detect those genomic rearrangements. Second, coding variants currently classified as non-pathogenic, including silent mutations and some missense mutations, may be of pathogenic relevance through modulation of mRNA splicing and/or stability; this possibility has not been systematically tested to date. Third, some "PFIC" cases may have been wrongly clinically diagnosed. Finally, the high proportion of unsolved cases suggests novel genetic etiologies that remain to be elucidated. With this purpose, exome analysis on some interesting families

could highlight new mutations on other genes potentially involved in biliary synthesis, transport and reuptake.

The NGS-based developed protocols turned out to be both faster and cost-saving compared to the traditional Sanger sequencing, allowing us to obtain rapid medical reports for PFIC patients in just few weeks.

Analysing reagent costs and their respective labour time to reach a complete molecular diagnosis between the conventional Sanger sequencing usually adopted in many public hospital laboratories and the NGS method, we can assess that the last one is absolutely more efficient, time and cost saving. Moreover, NGS technique resulted very suitable for our purpose of sequencing a multi-gene panel.

Patients can be genotyped with NGS protocols at reagent costs per patient five-times less than Sanger sequencing method. Regarding the benchwork, the hands-on time is required just for preparing libraries. With the introduction of Ion Chef System in our workflow, in fact, all the other steps are automatized. The subsequent data analysis is considerably faster and can be completed in just few hours.

In conclusion, the NGS-based genetic analysis for PFIC disease is a highly accurate and reliable approach for mutation analysis achieving high sensitivity with a faster turnaround time and lower cost.

NGS would be an appropriate new gold-standard for clinical genetic testing of Progressive Familial Intrahepatic Cholestatic disease.

Bibliography

1. Boyer JL. Bile formation and secretion. *Compr Physiol* 2013:3: 1035-78.
2. Arrese M, Trauner M. Molecular aspects of bile formation and cholestasis. *Trends Mol Med* 2003: 9: 558-64.
3. Trauner M, Boyer JL. Bile salt transporters: molecular characterization, function, and regulation. *Physiol Rev* 2003:83:633-71.
4. Dawson PA, Karpen SJ. Intestinal transport and metabolism of bile acids. *J Lipid Res* 2015: 56: 1085-99.
5. Hegade VS, Speight RA, Etherington RE, Jones DE. Novel bile acid therapeutics for the treatment of chronic liver diseases. *Therap Adv Gastroenterol* 2016: 9: 376-91.
6. Samant, H., Manatsathit, W., Dies, D., Shokouh-Amiri, H., Zibari, G., Boktor, M., & Alexander, J. S. Cholestatic liver diseases: An era of emerging therapies. *World Journal of Clinical Cases*, 2019. 7(13), 1571–1581.
7. Jansen PL, Ghallab A, Vartak N, et al. The ascending pathophysiology of cholestatic liver disease. *Hepatology* 2017:65: 722-38.
8. Wagner M, Zollner G, Trauner M. Nuclear receptor regulation of the adaptive response of bile acid transporters in cholestasis. *Semin Liver Dis* 2010: 30: 160-77.
9. Matsubara T, Li F, Gonzalez FJ. FXR signaling in the enterohepatic system. *Mol Cell Endocrinol* 2013: 368: 17-29.
10. Cai SY, Ouyang X, Chen Y, Soroka CJ, Wang J, Mennone A, Wang Y, et al. Bile acids initiate cholestatic liver injury by triggering a hepatocyte-specific inflammatory response. *JCI Insight* 2017: 2: e90780.
11. Trauner M, Meier PJ, Boyer JL. Molecular pathogenesis of cholestasis. *N Engl J Med* 1998; 339: 1217-1227.
12. Hori T, Nguyen JH, Uemoto S. Progressive familial intrahepatic cholestasis. *Hepatobiliary Pancreat Dis Int*. 2010; 9:570–8.

13. Davit-Spraul A, Gonzales E, Baussan C, Jacquemin E. The spectrum of liver diseases related to ABCB4 gene mutations: pathophysiology and clinical aspects. *Semin Liver Dis* 2010; 30:134–46.

14. Clayton RJ, Iber FL, Ruebner BH, McKusick VA. Byler disease. Fatal familial intrahepatic cholestasis in Amish kindred. *Am J Dis Child* 1969; 117:112–24.

15. Srivastava A. Progressive familial intrahepatic cholestasis. *J Clin Exp Hepatol*. 2014; 4(1):25–36.

16. Strautnieks SS, Kagalwalla AF, Tanner MS, Knisely AS, Bull L, Freimer N, et al. Identification of a locus for progressive familial intrahepatic cholestasis PFIC2 on chromosome 2q24. *Am J Hum Genet* 1997; 61:630–633.

17. Jacquemin E, De Vree JM, Cresteil D, Sokal EM, Sturm E, Dumont M, et al. The wide spectrum of multidrug resistance 3 deficiency: from neonatal cholestasis to cirrhosis of adulthood. *Gastroenterology* 2001;120:1448–1458.

18. Pauli-Magnus C, Lang T, Meier Y, Zodan-Marin T, Jung D, Breymann C, 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* 2004;14:91–102

19. Sambrotta M, Strautnieks S, Papouli E, Rushton P, Clark BE, Parry DA, Logan CV, Newbury LJ, Kamath BM, Ling S, Grammatikopoulos T, Wagner BE, Magee JC, Sokol RJ, Mieli-Vergani G; University of Washington Center for Mendelian Genomics, Smith JD, Johnson CA, McClean P, Simpson MA, Knisely AS, Bull LN, Thompson RJ. Mutations in TJP2 cause progressive cholestatic liver disease. *Nat Genet*. 2014 Apr;46(4):326-8.

20. Gomez-Ospina N, Potter CJ, Xiao R, Manickam K, Kim M-S, Kim KH. Mutations in the nuclear bile acid receptor FXR cause progressive familial intrahepatic cholestasis. *Nat Commun.* 2016; 7:10713.

21. Qiu YL, Gong JY, Feng JY, Wang RX, Han J, Liu T, Lu Y, Li LT, Zhang MH, Sheps JA, Wang NL, Yan YY, Li JQ, Chen L, Borchers CH, Sipos B, Knisely AS, Ling V, Xing QH, Wang JS. Defects in myosin VB are associated with a spectrum of previously

undiagnosed low γ-glutamyltransferase cholestasis. *Hepatology* 2017 May;65(5):1655-1669.

22. Bull LN, van Eijk MJ, Pawlikowska L, DeYoung JA, Juijn JA, Liao M, Klomp LW, Lomri N, Berger R, Scharschmidt BF, Knisely AS, Houwen RH, Freimer NB. A gene encoding a P-type ATPase mutated in two forms of hereditary cholestasis. *Nature Genetics* vol. 18, no. 3, pp. 219–224, 1998.

23. S.W. C. VanMil, L.W. J. Klomp, L. N. Bull, andR.H. J. Houwen. FIC1 disease: A spectrum of intrahepatic cholestatic disorders. *Seminars in Liver Disease* vol. 21, no. 4, pp. 535–544, 2001.

24. A. Miyahawa-Hayashino, H. Egawa, T. Yorifuji, M. Hasegawa, H. Haga, T. Tsuruyama, M. C. Wen, R. Sumazaki, T. Manabe, S. Uemoto. Allograft steatohepatitis in progressive familial intrahepatic cholestasis type 1 after living donor liver transplantation. *Liver Transplantation* vol. 15, no. 6, pp. 610–618, 2009.

25. Nicastro E, Stephenne X, Smets F, Fusaro F, de Magnee C, Reding R, Sokal E. M. Recovery of graft steatosis and protein-losing enteropathy after biliary diversion in a PFIC 1 liver transplanted child. *Pediatr Transplant*. 2012;16(5):E177–82.

26. Jacquemin E, Hermans D, Myara A, Habes D, Debray D, Hadchouel M, Sokal EM, Bernard O. Ursodeoxycholic acid therapy in pediatric patients with progressive familial intrahepatic cholestasis. *Hepatology*. 1997.

27. Davis AR, Rosenthal P, Newman TB. Nontransplant surgical interventions in progressive familial intrahepatic cholestasis. *J Pediatr Surg*. 2009;44:821–7.

28. Lykavieris P, van Mil S, Cresteil D. Progressive familial intrahepatic cholestasis type 1 and extrahepatic features: no catch-up of stature growth, exacerbation of diarrhoea, and appearance of liver steatosis after liver transplantation. *J Hepatol*. 2003;39:447–52.

29. Kubitz R, Dröge C, Stindt J, Weissenberger K, Häussinger D. The bile salt export pump (BSEP) in health and disease. *Clin Res Hepatol Gastroenterol*. 2012; 36:536–53.

30. Ananthanarayanan M, Li Y, Surapureddi S, Balasubramaniyan N, Ahn J, Goldstein JA, et al. Histone H3K4 trimethylation by MLL3 as part of ASCOM complex is critical

for NR activation of bile acid transporter genes and is downregulated in cholestasis. *Am J Physiol Gastrointest Liver Physiol* 2011;300:G771—81

31. Woods A, Heslegrave AJ, Muckett PJ, Levene AP, Clements M, Mobberley M, Ryder TA, Abu-Hayyeh S, Williamson C, Goldin RD, Ashworth A, Withers DJ, Carling D. LKB1 is required for hepatic bile acid transport and canalicular membrane integrity in mice. *Biochem J*. 2011; 434:49—60.

32. Scheimann AO, Strautnieks SS, Knisely AS, Byrne JA, Thompson RJ, Finegold MJ. Mutations in Bile Salt Export Pump (ABCB11) in Two Children with Progressive Familial Intrahepatic Cholestasis and Cholangiocarcinoma. *Journal of Pediatrics*. vol. 150, no. 5, pp. 556–559, 2007.

33. Sticova E, Jirsa M, Pawlowska J. New Insights in Genetic Cholestasis: From Molecular Mechanisms to Clinical Implications. *Canadian Journal of Gastroenterology and Hepatology.* 2018. doi.org/10.1155/2018/2313675

34. Gonzales E, Grosse B, Schuller B, Davit-Spraul A, Conti F, Guettier C, Cassio D, Jacquemin E. Targeted pharmacotherapy in progressive familial intrahepatic cholestasis type 2: Evidence for improvement of cholestasis with 4-phenylbutyrate. *Hepatology* 2015. doi.org/10.1002/hep.27767.

35. Kubitz R, Dröge C, Kluge S, Stross C, Walter N, Keitel V, Häussinger D, Stindt J. Autoimmune BSEP disease: disease recurrence after liver transplantation for progressive familial intrahepatic cholestasis. *Clin Rev Allergy Immunol*. 2015 Jun;48(2-3):273-84. doi: 10.1007/s12016-014-8457-4.

36. de Vree JM, Jacquemin E, Sturm E, Cresteil D, Bosma PJ, Aten J, Deleuze JF, Desrochers M, Burdelski M, Bernard O, Oude Elferink RP, Hadchouel M. Mutations in the MDR3 gene cause progressive familial intrahepatic cholestasis. *Proc Natl Acad Sci U S A*. 1998 Jan 6;95(1):282-7. DOI: 10.1073/pnas.95.1.282.

37. Groen A., Romero M.R., Kunne C., Hoosdally S.J., Dixon P.H., Wooding C., Williamson C., Seppen J., Van den Oever K., Mok K.S., Paulusma C.C., Linton K.J., Oude Elferink R.P. Complementary functions of the flippase ATP8B1 and the

floppase ABCB4 in maintaining canalicular membrane integrity. *Gastroenterology* 141:1927-1937(2011).

38. Morita S.Y., Tsuda T., Horikami M., Teraoka R., Kitagawa S., Terada T. Bile salt-stimulated phospholipid efflux mediated by ABCB4 localized in nonraft membranes. *J. Lipid Res*. 2013 May;54(5):1221-30. doi: 10.1194/jlr.M032425.

39. Wendum D. Barbu V. Rosmorduc O. Arriv´e L. Fl´ejou J.-F, and Poupon R. Aspects of liver pathology in adult patients with MDR3/ABCB4 gene mutations. *Virchows Archiv*, 2012, vol. 460, no. 3, pp. 291–298.

40. Morotti RA, Suchy FJ, Magid MS. Progressive familial intrahepatic cholestasis (PFIC) type 1, 2, and 3: a review of the liver pathology findings. *Semin Liver Dis*. 2011 Feb;31(1):3-10.

41. Fiorucci S, Clerici C, Antonelli E, Orlandi S, Goodwin B, Sadeghpour BM, Sabatino G, Russo G, Castellani D, Willson TM, Pruzanski M, Pellicciari R, Morelli A. Protective effects of 6-ethyl chenodeoxycholic acid, a farnesoid X receptor ligand, in estrogen-induced cholestasis. *J Pharmacol Exp Ther*. 2005; 313:604–12.

42. Sambrotta M. and Thompson R. J. Mutations in TJP2, encoding zona occludens 2, and liver disease. *Tissue Barriers*, 2015 vol. 3, no. 3, Article ID e1026537.

43. Holczbauer Á, Gyöngyösi B, Lotz G, Szijártó A, Kupcsulik P, Schaff Z, Kiss A. Distinct claudin expression profiles of hepatocellular carcinoma and metastatic colorectal and pancreatic carcinomas. *J Histochem Cytochem*. 2013 Apr; 61(4):294-305.

44. Itoh M, Furuse M, Morita K, Kubota K, Saitou M, Tsukita S. Direct binding of three tight junction-associated MAGUKs, ZO-1, ZO-2, and ZO-3, with the COOH termini of claudins. *J Cell Biol*. 1999 Dec 13; 147(6):1351-63.

45. Zhou, S., Hertel, P. M., Finegold, M. J., Wang, L., Kerkar, N., Wang, J., Wong, L.-J. C., Plon, S. E., Sambrotta, M., Foskett, P., Niu, Z., Thompson, R. J., Knisely, A. S. Hepatocellular carcinoma associated with tight-junction protein 2 deficiency. *Hepatology*. 62: 1914-1916, 2015.

46. Carlton VE, Harris BZ, Puffenberger EG, Batta AK, Knisely AS, Robinson DL, Strauss KA, Shneider BL, Lim WA, Salen G, Morton DH, Bull LN. Complex inheritance of

familial hypercholanemia with associated mutations in TJP2 and BAAT. *Nat Genet*. 2003 May;34(1):91-6. DOI: 10.1038/ng1147.

47. Chen HL, Li HY, Wu JF, Wu SH, Chen HL, Yang YH, Hsu YH, Liou BY, Chang MH, Ni YH. Panel-Based Next-Generation Sequencing for the Diagnosis of Cholestatic Genetic Liver Diseases: Clinical Utility and Challenges. *J Pediatr* 2019; 205: 153-159.e6 DOI: 10.1016/j.jpeds.2018.09.028.

48. Koutsounas I, Theocharis S, Delladetsima I, Patsouris E, and Giaginis C. Farnesoid x receptor in human metabolism and disease: the interplay between gene polymorphisms, clinical phenotypes and disease susceptibility. *Expert Opinion on Drug Metabolism & Toxicology*, 2015, vol. 11, no. 4, pp. 523–532.

49. Gissen P, Arias IM. Structural and functional hepatocyte polarity and liver disease. *J Hepatol*. 2015;63:1023–37.

50. Bull LN, Thompson RJ. Progressive Familial Intrahepatic Cholestasis. *Clin Liver Dis* 2018. 22: 657–669.

51. Müller T, Hess MW, Schiefermeier N, Pfaller K, Ebner HL, Heinz-Erian P, Ponstingl H, Partsch J, Röllinghoff B, Köhler H, Berger T, Lenhartz H, Schlenck B, Houwen RJ, Taylor CJ, Zoller H, Lechner S, Goulet O, Utermann G, Ruemmele FM, Huber LA, Janecke AR. MYO5B mutations cause microvillus inclusion disease and disrupt epithelial cell polarity. *Nat Genet*. 2008 Oct;40(10):1163-5. doi: 10.1038/ng.225.

52. Siahanidou T, Koutsounaki E, Skiathitou A, Stefanaki K, Marinos E, Panajiotou I, Chouliaras G. Extraintestinal manifestations in an infant with microvillus inclusion disease: complications or features of the disease? *Eur J Pediatr.* 2013;172:1271-1275.

53. Girard M, Lacaille F, Verkarre V, Mategot R, Feldmann G, Grodet A, Sauvat F, Irtan S, Davit-Spraul A, Jacquemin E, Ruemmele F, Rainteau D, Goulet O, Colomb V, Chardot C, Henrion-Caude A, Debray D. MYO5B and bile salt export pump contribute to cholestatic liver disorder in microvillous inclusion disease. *Hepatology* 2014; 60:301-310.

54. Gonzales E, Taylor SA, Davit-Spraul A, Thébaut A, Thomassin N, Guettier C, Whitington FP, Jacquemin E. MYO5B mutations cause cholestasis with normal serum

gamma-glutamyl transferase activity in children without microvillous inclusion disease. *Hepatology* 2017; 65:164–73.

55. Van Mil S.W. C., Houwen R. H. J., and Klomp L.W. J. Genetics of familial intrahepatic cholestasis syndromes. *Journal of Medical Genetics*, 2005. vol. 42, no. 6, pp. 449–463.

56. Dröge C. Bonus M. Baumann U. Klindt C. Lainka E. Kathemann S. Brinkert F. Grabhorn E. Pfister ED. Wenning D. Fichtner A. Gotthardt DN. Weiss KH. McKiernan P. Puri RD. Verma IC. Kluge S. Gohlke H. Schmitt L. Kubitz R. Häussinger D. Keitel V. Sequencing of FIC1, BSEP and MDR3 in a large cohort of patients with cholestasis revealed a high number of different genetic variants. *J Hepatol*. 2017 Dec;67(6):1253-1264.

57. Brenard R, Geubel AP, Benhamou JP. Benign recurrent intrahepatic cholestasis. *J Clin Gastroenterol.* 1989;11:546–51.

58. Vitale G, Gitto S, Vukotic R, Raimondi F, Andreone P. Familial intrahepatic cholestasis: New and wide perspectives. *Dig Liver Dis*. 2019 Jul;51(7):922-933.

59. Van Ooteghem NA, Klomp LW, van Berge-Henegouwen GP, Houwen RH. Benign recurrent intrahepatic cholestasis progressing to progressive familial intrahepatic cholestasis: low GGT cholestasis is a clinical continuum. *J Hepatol*. 2002 Mar;36(3):439-43.

60. Folvik G, Hilde O, Helge GO. Benign recurrent intrahepatic cholestasis: review and long-term follow-up of five cases. *Scand J Gastroenterol.* 2012; 47:482–8.

61. Beuers U, Trauner M, Jansen P, Poupon R. New paradigms in the treatment of hepatic cholestasis: From UDCA to FXR, PXR and beyond. *Journal of Hepatology*, vol. 62, no. 1, pp. S25–S37, 2015.

62. Stapelbroek JM, van Erpecum KJ, Klomp LW, Venneman NG, Schwartz TP, van Berge Henegouwen GP, Devlin J, van Nieuwkerk CM, Knisely AS, Houwen RH. Nasobiliary drainage induces long-lasting remission in benign recurrent intrahepatic cholestasis. *Hepatology* 2006; 43:51–3.

63. Ovadia C, Seed PT, Sklavounos A, Geenes V, Di Ilio C, Chambers J, Kohari K, Bacq Y, Bozkurt N, Brun-Furrer R, Bull L, Estiú MC, Grymowicz M, Gunaydin B, Hague WM, Haslinger C, Hu Y, Kawakita T, Williamson C. Association of adverse perinatal outcomes of intrahepatic cholestasis of pregnancy with biochemical markers: results of aggregate and individual patient data meta-analyses. *The Lancet* 2019;393:899–909.

64. Jin WY, Lin SL, Hou RL, Chen XY, Han T, Jin Y, Tang L, Zhu ZW, Zhao ZY. Associations between maternal lipid profile and pregnancy complications and perinatal outcomes: a population-based study from China. *BMC Pregnancy Childbirth*. 2016 Mar 21;16:60.

65. Williamson C, Geenes V. Intrahepatic cholestasis of pregnancy. *Obstet Gynecol*. 2014; 124:120–33.

66. Van der Woerd WL, van Mil SW, Stapelbroek JM, Klomp LW, van de Graaf SF,Houwen RH. Familial cholestasis: progressive familial intrahepatic cholestasis, benign recurrent intrahepatic cholestasis and intrahepatic cholestasis of pregnancy. *Best Pract Res Clin Gastroenterol.* 2010; 24:541–53.

67. Pauli-Magnus C, Meier PJ, Stieger B. Genetic Determinants of Drug-induced Cholestasis and Intrahepatic Cholestasis of Pregnancy. *Semin Liver Dis* 2010;30:147–159.

68. Painter JN, Savander M, Ropponen A, Nupponen N, Riikonen S, Ylikorkala O, Lehesjoki AE, Aittomäki K. Sequence variation in the *ATP8B1* gene and intrahepatic cholestasis of pregnancy. *European Journal of Human Genetics* (2005) 13, 435–439.

69. Van Mil SW, Milona A, Dixon PH, Mullenbach R, Geenes VL, Chambers J, Shevchuk V, Moore GE, Lammert F, Glantz AG, Mattsson LA, Whittaker J, Parker MG, White R, Williamson C. Functional variants of the central bile acid sensor FXR identified in intrahepatic cholestasis of pregnancy. *Gastroenterology* 2007;133:507–16.

70. Dixon PH, Sambrotta M, Chambers J, Taylor-Harris P, Syngelaki A, Nicolaides K, Knisely AS, Thompson RJ, Williamson C. An expanded role for heterozygous mutations of *ABCB4*, *ABCB11*, *ATP8B1*, *ABCC2* and *TJP2* in intrahepatic cholestasis of pregnancy. *Sci Rep* 2017; 7:11823.

71. Westbrook RH, Dusheiko G, Williamson C. Pregnancy and liver disease. *J Hepatol*. 2016; 64:933–45.

72. Ropponen A, Sund R, Riikonen S, Ylikorkala O, Aittomäki K. Intrahepatic cholestasis of pregnancy as an indicator of liver and biliary diseases: a population-based study. *Hepatology* 2006;43:723–8.

73. Shapiro MA, Lewis JH. Causality assessment of drug induced hepatotoxicity: promises and pitfalls. *Clin Liver Dis.* 2007;11(3):477–505.

74. Padda MS, Sanchez M, Akhtar AJ, Boyer JL. Drug-induced cholestasis. *Hepatology* 2011;53:1377–87.

75. Rosmorduc O, Hermelin B, Poupon R. MDR3 gene defect in adults with symptomatic intrahepatic and gallbladder cholesterol cholelithiasis. *Gastroenterology* 2001; 120:1459- 67.

76. Rosmorduc O, Poupon R. Low phospholipid associated cholelithiasis: association with mutation in the MDR3/ABCB4 gene. *Orphanet J Rare Dis* 2007; 11(2):29.

77. Erlinger S. Low phospholipid-associated cholestasis and cholelithiasis. *Clin Res Hepatol Gastroenterol.* 2012; 36(Suppl. 1):S36–40.

78. Poupon R, Rosmorduc O, Boëlle PY, Chrétien Y, Corpechot C, Chazouillères O, Housset C, Barbu V. Genotype-phenotype relationships in the low-phospholipid-associated cholelithiasis syndrome: a study of 156 consecutive patients. *Hepatology* 2013; 58:1105–10.

79. Stapelbroek JM, van Erpecum KJ, Klomp LW, Houwen RH. Liver disease associated with canalicular transport defects: current and future therapies. *J Hepatol* 2010; 52: 258-271.

80. Van der Woerd WL, Houwen RHJ, van de Graaf SFJ. Current and future therapies for inherited cholestatic liver diseases. *World J Gastroenterol*. 2017 Feb 7; 23(5): 763-775.

81. Stapelbroek JM, van Erpecum KJ, Klomp LW, Venneman NG, Schwartz TP, van Berge Henegouwen GP, Devlin J, van Nieuwkerk CM, Knisely AS, Houwen RH.

Nasobiliary drainage induces long lasting remission in benign recurrent intrahepatic cholestasis. *Hepatology* 2006; 43: 51-53.

82. Whitington PF, Whitington GL. Partial external diversion of bile for the treatment of intractable pruritus associated with intrahepatic cholestasis. *Gastroenterology* 1988; 95: 130-136.

83. Van der Woerd WL, Kokke FT, van der Zee DC, Houwen RH. Total biliary diversion as a treatment option for patients with progressive familial intrahepatic cholestasis and Alagille syndrome. *J Pediatr Surg.* 2015; 50: 1846-1849.

84. Mali VP, Fukuda A, Shigeta T, Uchida H, Hirata Y, Rahayatri TH, Kanazawa H, Sasaki K, de Ville de Goyet J, Kasahara M. Total internal biliary diversion during liver transplantation for type 1 progressive familial intrahepatic cholestasis: a novel approach. *Pediatr Transplant*. 2016; 20: 981-986.

85. Takebe T, Sekine K, Enomura M, Koike H, Kimura M, Ogaeri T, Zhang RR, Ueno Y, Zheng YW, Koike N, Aoyama S, Adachi Y, Taniguchi H. Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature* 2013; 499: 481-484.

86. Huch M, Dorrell C, Boj SF, van Es JH, Li VS, van de Wetering M, Sato T, Hamer K, Sasaki N, Finegold MJ, Haft A, Vries RG, Grompe M, Clevers H. In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. *Nature* 2013; 494: 247-250.

87. Wang L, Wu J, Fang W, Liu GH, Izpisua Belmonte JC. Regenerative medicine: targeted genome editing in vivo. *Cell Res* 2015; 25: 271-272.

88. Voelkerding KV, Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. *Clin Chem.* 2009 Apr;55(4):641-58.

89. Merriman B, Ion Torrent R&D Team, Rothberg JM. Progress in ion torrent semiconductor chip-based sequencing. *Electrophoresis*. 2012 Dec;33(23):3397-41.

90. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG 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 May;17(5):405-24.

91. Mendell NR, Simon GA. A general expression for the variance-covariance matrix of estimates of gene frequency: the effects of departures from Hardy-Weinberg equilibrium. *Ann Hum Genet* 1984, 48, 283-286

92. Vitale G, Pirillo M, Mantovani V, Marasco E, Aquilano A, Gamal N, Francalanci P, Conti F, Andreone P. Bile salt export pump deficiency disease: two novel, late onset, ABCB11 mutations identified by next generation sequencing. *Ann Hepatol*. 2016 Sep - Oct;15(5):795-800.

93. Degiorgio D, Colombo C, Seia M, Porcaro L, Costantino L, Zazzeron L, Bordo D, Coviello DA. Molecular characterization and structural implications of 25 new ABCB4 mutations in progressive familial intrahepatic cholestasis type 3 (PFIC3). *Eur J Hum Genet*. 2007 Dec;15(12):1230-8.

94. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Albarca Aguilera M, Meyer R, and Massouras A. VarSome: The Human Genomic Variant Search Engine. *Oxford Bioinformatics*, bty897, 30 October 2018.