Summary

Paralleling the first 30 years of the Journal of Hepatology we have witnessed huge advances in our understanding of liver disease and physiology. Genetic advances have played no small part in that. Initial studies in the 1970s and 1980s identified the strong major histocompatibility complex associations in autoimmune liver diseases. During the 1990s, developments in genomic technologies drove the identification of genes responsible for Mendelian liver diseases. Over the last decade, genome-wide association studies have allowed for the dissection of the genetic susceptibility to complex liver disorders, in which also environmental co-factors play important roles. Findings have allowed the identification and elaboration of pathophysiological processes, have indicated the need for reclassification of liver diseases and have already pointed to new disease treatments. In the immediate future genetics will allow further stratification of liver diseases and contribute to personalized medicine. Challenges exist with regard to clinical implementation of rapidly developing technologies and interpretation of the wealth of accumulating genetic data. The historical perspective of genetics in liver diseases illustrates the opportunities for future research and clinical care of our patients.

Introduction

Over the last 30 years, genetics has enhanced our understanding and management of liver disease (Table 1). The fundamental technologies available for genetic analysis encompass positional cloning of unknown disease genes, single gene tests for known single nucleotide variants (SNV), genome-wide association studies (GWAS) that compare genotype frequencies across the whole genome between cases and controls to identify unknown genetic risk factors, and next-generation sequencing (NGS) of selected genes, all exons, or the whole genome in individual patients.

To illustrate the developments and opportunities of these genetic methods and technologies, we here pose and address five key questions that should help students, researchers, and health care professionals to understand the strides in the genetics of liver disease. Our answers cover both historical aspects and indicate how the advances in our understanding have led to a rapidly developing framework for diagnostic and therapeutic strategies. Rather than a comprehensive representation of the complete genetics of liver disease (for which comprehensive reviews can be found [41]), we aim to present the reader with key examples of the current state of the field and our associated reflections. We focus on germline mutations and deliberately omit the utility of genomics in the management of hepatocellular carcinoma and cholangiocarcinoma, which is also covered elsewhere [42] and in this special issue.

Has genetics improved our understanding of the pathophysiology of liver diseases?

The first genes of monogenic diseases with predominant liver phenotypes that were mapped and cloned were the Wilson disease gene ATP7B (1993) and the haemochromatosis gene HFE (1996) (Table 1; Fig. 1). The Wilson disease gene was localized in the 1980s (Table 1; Fig. 1). The Wilson disease gene was localized...
by linkage disequilibrium and haplotype analysis in more than 100 families [8]. Functional studies demonstrated that it encodes a P-type ATPase gene with metal binding regions similar to those found in prokaryotic heavy metal transporters [9]. Building on this discovery, subsequent studies were able to determine vast aspects of previously unknown aspects of copper transport and the pathophysiology of Wilson disease [43]. Using a similar strategy three years later, the HFE gene in the extended major histocompatibility complex (MHC) was shown to be mutated in patients with autosomal-recessive haemochromatosis [10]. Furthermore, additional types of non-HFE hereditary haemochromatosis were subsequently described, linked to mutations in the ferroportin gene (SLC40A1) [17,18], the transferrin receptor 2 gene (TFR2) [19], the hepcidin gene (HAMP) [20] and the hemoglobin gene (HJV) [21], respectively. These genetic discoveries jointly paved the way for the full characterization of hepatic iron metabolism and its regulators [44], allowing for dissection of mechanisms responsible for the development of liver disease in the presence of detected mutations.

![Key Points](https://www.ncbi.nlm.nih.gov/gene/)

Different liver diseases are at different stages of the following sequence of steps:

- The first genetic discoveries identified the major loci for Mendelian diseases
- Numerous studies have identified susceptibility to complex diseases
- All diseases are actually complex with multiple genes contributing to the phenotype
- With the identification of different underlying genetic contributors to disease those diseases are being reclassified
- A better understanding of genetic aetiology will suggest new modes of treatment for disease
- A complete understanding of all the contributors to a individual’s disease will allow personally tailored treatment

The mutation profile of disease genes varies. Hereditary haemochromatosis is most often caused by a predominant founder mutation (p.C282Y) which accounts for approximately 95% of patients [45]. Wilson disease shows a different genetic profile, with extensive allelic heterogeneity with more than 500 mutations described in the ATP7B gene to date. For this reason, genetic testing in haemochromatosis is undertaken via targeted genotyping of the predominant variants (always starting with p.C282Y), whereas in Wilson disease there is a need for gene sequencing and variant analysis to obtain genetic support for the diagnosis.

The first GWAS in hepatobiliary diseases was performed for gallstone disease and confirmed the candidacy of the hepatobiliary cholesterol transporter ABCG5/G8 as a major susceptibility gene for gallstones worldwide, with p.D19H representing the likely causal single nucleotide polymorphism (SNP) [22,46]. Other rare loss-of-function SNVs in this transporter had been identified previously, in individual patients, to underlie the monogenic disease sitosterolemia, which is characterized by unrestricted intestinal absorption of both cholesterol and phytosterols such as sitosterol [47]. The adjacent, oppositely oriented ABCG5/G8 genes encode two ATP-binding cassette (ABC) hemitransporters that are localized in the apical membranes of enterocytes and hepatocytes. The case also illustrates a

<table>
<thead>
<tr>
<th>Mendelian diseases</th>
<th>Complex diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Degree of heritability</strong></td>
<td>Relative sibling risk increased &gt;200-1000 times compared with general population risk</td>
</tr>
<tr>
<td><strong>Technology preferences</strong></td>
<td>Whole-exome and whole-genome sequencing</td>
</tr>
<tr>
<td><strong>Study design</strong></td>
<td>Linkage analysis, <em>in silico</em> candidate variant filtering and prioritization, functional assessments of disease variants</td>
</tr>
<tr>
<td><strong>Key challenges</strong></td>
<td>Large number of candidate variants makes conclusive identification difficult, mechanistic support for causality needed and may require extensive studies</td>
</tr>
</tbody>
</table>

Table 1. Present day technologies involved in the detection of disease relevant genetic variation. These genomic technologies are principally deriving from a variety of sequencing techniques (targeting genes, the coding regions of the genome [exome] or the whole genome) and microarrays for genome-wide single nucleotide or copy number variant detection. Cost (price) and the availability of relevant bioinformatic tools also need considerations in making the ideal selection of method for investigating any observed trait.
fundamental difference between risk factors in monogenic vs. polygenic/complex traits (Table 2). For sitosterolemia, the SNVs appear to be sufficient to cause the observed phenotype. For risk alleles of SNPs detected by GWAS (that have an odds ratio of approximately 2), interacting genetic and environmental factors are required for gallstone development [46,48,49].

GWAS for autoimmune liver diseases have led to the detection of several dozen risk genes in primary biliary cirrhosis (PBC) [25], primary sclerosing cholangitis (PSC) [32,50] and recently, autoimmune hepatitis (AIH) [37]. With few exceptions, the detected risk loci overlap with those of other autoimmune and immune-mediated diseases (Fig. 2) [51–53]. This means that the majority of the genetic risk factors for these phenotypes involves an increasing susceptibility to autoimmunity rather than liver affections per se. A large fraction of the risk loci and the strong MHC associations found in both PBC, PSC, and AIH point to the pathophysiological importance of adaptive immune responses [46]. The overall genetic contribution to disease susceptibility in PBC, PSC, and AIH is relatively low, with the current risk gene pool representing less than 10% of the overall disease liability [51,52]. Larger studies will increase this fraction, but more than 50% of the susceptibility to these diseases is still likely to be of an environmental origin.

Individual whole-exome sequencing (WES) has been successfully used to identify unknown genetic variants for inherited diseases or conditions of indeterminate aetiology such as

<table>
<thead>
<tr>
<th>Year</th>
<th>Landmark genetic publication</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972</td>
<td>MHC associations in autoimmune hepatitis</td>
<td>[1]</td>
</tr>
<tr>
<td>1979</td>
<td>MHC associations in primary biliary cirrhosis</td>
<td>[2]</td>
</tr>
<tr>
<td>1982</td>
<td>MHC associations in primary sclerosing cholangitis</td>
<td>[3]</td>
</tr>
<tr>
<td>1989</td>
<td>CFTR in cystic fibrosis</td>
<td>[4]</td>
</tr>
<tr>
<td>1992/95</td>
<td>UGT1A1 in Gilbert and Crigler-Najjar syndromes</td>
<td>[5-7]</td>
</tr>
<tr>
<td>1993</td>
<td>ATP7B in Wilson disease</td>
<td>[8, 9]</td>
</tr>
<tr>
<td>1996</td>
<td>HFE in haemochromatosis (type 1)</td>
<td>[10]</td>
</tr>
<tr>
<td>1997</td>
<td>JAG1 in Alagille syndrome</td>
<td>[12,13]</td>
</tr>
<tr>
<td>1998</td>
<td>ATP8B1 in PFIC (type 1)</td>
<td>[14]</td>
</tr>
<tr>
<td>1998</td>
<td>ABCB11 in PFIC (type 2)</td>
<td>[15]</td>
</tr>
<tr>
<td>1998</td>
<td>ABCB4 in PFIC (type 3)</td>
<td>[16]</td>
</tr>
<tr>
<td>2001</td>
<td>SLC40A1 (ferroportin) in haemochromatosis (type 4)</td>
<td>[17,18]</td>
</tr>
<tr>
<td>2002</td>
<td>TFR2 in haemochromatosis (type 3)</td>
<td>[19]</td>
</tr>
<tr>
<td>2003</td>
<td>HAMP in haemochromatosis (type 2B)</td>
<td>[20]</td>
</tr>
<tr>
<td>2004</td>
<td>HJV in haemochromatosis (type 2A)</td>
<td>[21]</td>
</tr>
<tr>
<td>2006</td>
<td>NOTCH2 in Alagille syndrome</td>
<td>[12]</td>
</tr>
<tr>
<td>2007</td>
<td>ABCG8 in gallstone disease (GWAS)</td>
<td>[22]</td>
</tr>
<tr>
<td>2008</td>
<td>PNPLA3 in non-alcoholic fatty liver disease (GWAS)</td>
<td>[23]</td>
</tr>
<tr>
<td>2008</td>
<td>First GWAS on genetic factors associated with plasma liver enzyme activities</td>
<td>[24]</td>
</tr>
<tr>
<td>2009</td>
<td>First GWAS in primary biliary cirrhosis</td>
<td>[25]</td>
</tr>
<tr>
<td>2009</td>
<td>MHC associations in flucloxacinil DILI (GWAS)</td>
<td>[26]</td>
</tr>
<tr>
<td>2009</td>
<td>MHC associations in HBV clearance (GWAS)</td>
<td>[27]</td>
</tr>
<tr>
<td>2009</td>
<td>IL28B in HCV treatment response (GWAS)</td>
<td>[28]</td>
</tr>
<tr>
<td>2009</td>
<td>IL28B in spontaneous HCV clearance (GWAS)</td>
<td>[29]</td>
</tr>
<tr>
<td>2009</td>
<td>Genetic modifiers for CFTR-associated liver disease</td>
<td>[30]</td>
</tr>
<tr>
<td>2010</td>
<td>PNPLA3 in alcoholic liver disease</td>
<td>[31]</td>
</tr>
<tr>
<td>2010</td>
<td>First GWAS in primary sclerosing cholangitis</td>
<td>[32]</td>
</tr>
<tr>
<td>2010</td>
<td>ITPA in ribavirin-induced anaemia (GWAS)</td>
<td>[33]</td>
</tr>
<tr>
<td>2010</td>
<td>First GWAS in hepatitis B-related hepatocellular carcinoma (GWAS)</td>
<td>[34]</td>
</tr>
<tr>
<td>2011</td>
<td>MHC associations in amoxicillin-clavulanate DILI (GWAS)</td>
<td>[35]</td>
</tr>
<tr>
<td>2012</td>
<td>OATP1B1 and OATP1B3 in Rotor syndrome</td>
<td>[36]</td>
</tr>
<tr>
<td>2012</td>
<td>First GWAS on liver fibrosis in chronic HCV infection</td>
<td>[37]</td>
</tr>
<tr>
<td>2014</td>
<td>First GWAS in autoimmune hepatitis</td>
<td>[38]</td>
</tr>
<tr>
<td>2014</td>
<td>TJR2 in PFIC (type 4)</td>
<td>[39]</td>
</tr>
<tr>
<td>2014</td>
<td>MHC associations in hepatitis B vaccine response (GWAS)</td>
<td>[40]</td>
</tr>
</tbody>
</table>

DILI, drug-induced liver injury; HBV, hepatitis B virus; HCV, hepatitis C virus; PFIC, progressive intrahepatic cholestasis. For gene name abbreviations, see http://www.ncbi.nlm.nih.gov/gene/.
as acute liver failure in children [54]. Recently this has been illustrated in two families with unsuspected mitochondrial hepatopathies, where WES of germline DNA indicated homozygous mutations in \textit{MPV17} and \textit{SERAC1} as causative in paediatric cases of acute liver failure [55]. In fact, liver disease-targeted re-sequencing of genes has been implemented in several liver centres and we believe that it ultimately will be fully integrated into care and get a firm place in the diagnostic evaluation of patients with unexplained liver disease, in particular in the setting of a strong family history of similar afflictions. Due to the massive amounts of data being generated, establishing the pathophysiological connection for detected variants within a clinically relevant time frame remains a major challenge.

\textbf{Is genetics helpful for the definition of liver diseases and disease subtypes?}

Fifty years ago, Byler disease was first described in an Amish kindred [56]. The same disorder in non-Amish populations was termed Byler syndrome, which represented a heterogeneous group of autosomal-recessive liver diseases characterized by early onset of cholestasis that progresses to hepatic fibrosis, cirrhosis, and end-stage liver disease often before adulthood. In the past two decades basic research teams identified and cloned hepatobiliary transport proteins that are involved in the apical uptake and canalicular secretion of bile acids as well as other endo- and xenobiotics (Fig. 3) [57,58]. The landmark study by Smit \textit{et al.} [59] demonstrated that homozygous disruption of the murine \textit{Mdr2} P-glycoprotein gene, now known as \textit{Abcb4}, leads to a complete absence of phosphatidylcholine from bile and chronic cholangiopathy. Massive gene amplification and rearrangements also cause hepatocellular cancer in this model [60]. Studies in humans demonstrated that the phenotypic consequences of mutations of the orthologous human gene \textit{MDR3}/\textit{ABCB4} range from neonatal cholestasis to progressive familial intrahepatic cholestasis (PFIC) and cirrhosis in adults [16,61]; they can also result in intrahepatic cholestasis of pregnancy (ICP) [62] as well as intrahepatic and gallbladder cholesterol cholelithiasis, termed low phospholipid-associated cholelithiasis (LPAC) syndrome [63,64].

The related gene \textit{ABCB11} encodes the bile salt export pump (BSEP) and was shown to be mutated in consanguineous families with PFIC but generally low serum \(\gamma\)-glutamyl transferase activities [15,65]. In contrast, patients with Byler disease and others were shown to carry mutations in the \textit{ATP8B1} (\textit{FIC1}) gene [14,66] which encodes a P-type ATPase that flips phosphatidylserine from the outer to the inner leaflet of the hepatocanalicular membrane, thereby maintaining membrane integrity [67]. Studies comparing the phenotypic presentation and progression between BSEP and FIC1 deficiencies [68,69] reported that the former patients were more likely to develop neonatal cholestasis and manifested more severe hepatobiliary disease (as indicated by higher serum ALT activities and bile salt...
In 1997, Oda et al. [76] and Li et al. [77] demonstrated that another hereditary cholestatic liver disease, Alagille syndrome, is caused by mutations in the human homolog of Jagged-1 (JAG1), which encodes a ligand in the Notch signalling pathway. Alagille syndrome is a rare autosomal-dominant disease with patients showing not only paucity of intrahepatic bile ducts, but characteristic dysmorphic facies and cardiac anomalies, notably peripheral pulmonary stenosis, along with posterior embryotoxon and notched or butterfly vertebrae as well as failure to thrive. In less than 1% of patients, the disease is caused by NOTCH2 mutations [12,13], and the phenotypes are now known to be modified by other genes such as thrombospondin 2 [78].

How might genetics contribute to the classification of liver diseases?

Usually fatty liver disease is attributed to alcohol or metabolic stress, resulting in the terms alcoholic steatohepatitis (ASH) or non-alcoholic steatohepatitis (NASH) [79]. However, ASH may present as acute alcoholic hepatitis or chronic steatohepatitis, and both alcohol and environmental factors (over-nutrition, physical inactivity) can contribute to liver disease. In addition, chemotherapy and other drugs might also cause steatohepatitis. Interestingly, only a subgroup of patients (10–20%) who drink alcohol and/or are overweight develop severe steatohepatitis, advanced fibrosis, and cirrhosis. Since 2008, GWAS-based studies have led to the identification of a major genetic determinant of fatty liver and steatohepatitis as well as progression to fibrosis, cirrhosis, and hepatocellular cancer. Romeo et al. [23] identified the amino acid substitution p.I148M of the lipid droplet-associated triglyceride lipase PNPLA3 (also termed adiponutrin) to confer susceptibility to non-alcoholic fatty liver disease. It appears also to be the major genetic risk factor aggravating alcoholic liver disease [31,80]. This SNP is highly prevalent in European populations (risk allele frequency 21–28%) and confers a 3-fold risk for liver disease and NASH as well as a 12-fold risk for hepatocellular cancer in comparison to healthy controls [81]. The effect is modified by other genetic risk factors such as TM6SF2 p.E167K [82] but the population-attributable risk for fatty liver disease is highest for PNPLA3 p.I148M [81]. Therefore, we have proposed to use this novel genetic information to define a group of patients with fatty liver disease in whom the susceptibility gene PNPLA3 appears to be a major driver of disease progression, often in combination with ethanol and Western diet but also without, hence the suggested name PNPLA3-associated steatohepatitis (“PASH”) [83]. Although the PNPLA3 risk genotypes confer a high risk of NASH, the patients appear to be sensitive to the beneficial effects of lifestyle modification and dietary intervention should be encouraged to do so [84,85].

Common, conserved, aetiopathogenic pathways have been proposed across diseases, often related to failed resolution of responses to infections or stress and damage caused by the patient’s immune system [86]. Recently a new cytokine-based taxonomy has been proposed for chronic inflammatory diseases, comprising tumour necrosis factor (TNF)-α, interleukin (IL)-1, IL-6, IL-17, and IL-23 associated diseases, which are characterized by inflammation in different organs such as gut, skin, joints and brain [87]. Preclinical studies and randomised controlled trials with targeted cytokine intervention in humans indicate the existence of a cytokine hierarchy that was used to define these novel
Drug-induced liver injury (DILI) represents a clinically diverse spectrum. Hyman J. Zimmerman discriminated direct and idiosyncratic hepatotoxicity [99]. Whereas direct hepatotoxic drugs lead to dose-related liver injury that is reproducible in animal models, idiosyncratic liver injury shows an unpredictable outcome and is rare. The clinical manifestation of idiosyncratic DILI is acute hepatitis, and to date the diagnosis has been made by exclusion [100]. Idiosyncratic DILI cases are most often caused by antibiotics and generic drugs that have been approved decades ago. In terms of aetiology, DILI is not a single, uncommon disease but rather many, rare diseases converging on a similar presentation of immune-mediated liver injury. A GWAS for a common type of DILI caused by amoxicillin-clavulanate exposure showed the strongest finding in the MHC region, with signals localizing to human leukocyte antigen (HLA) class I and II genes. Although the respective HLA genotypes (HLA-A*02:01 and DQB1*06:02) have limited utility as predictive biomarkers because of low positive-predictive values, they can post hoc serve to assess causality in individual severe cases. Furthermore, the MHC associations point to the involvement of adaptive immunity, likely representing immune reactions to the drug itself, to drug metabolites, or to drug-generated neoantigens (protein adducts), any of which can trigger immunity in otherwise susceptible individuals. Such mechanisms might even be amendable to early treatment with steroids or other types of immunosuppression [101,102]. To date, additional GWAS have indicated only a limited contribution of common genetic variants outside of the MHC to risk of DILI [103].

What are the future hurdles and perspectives for translation of genetic research into clinics?

The best strategy for the transformation of the large number of genetic findings into novel therapeutic opportunities for complex liver diseases is not known. Almost 40 years after the discovery of the gene defect in cystic fibrosis [4], genetic information provides information on the future disease course of individual patients, and variant-specific treatment is being deployed [104]. In PBC, an attempt was made to target the IL-12-related genetic findings by ustekinumab, a monoclonal antibody which targets the p40 subunit of IL-12 and IL-23. The study was terminated due to lack of efficacy (https://clinicaltrials.gov/ct2/show/NCT01389973) and serves to exemplify the point that genetic findings do not have clinically useful predictive power. Modified from [109].

**Fig. 4. Distinguishing Mendelian from complex liver diseases.** For complex phenotypes, the contribution from genetics to overall disease liability is limited (typically less than 50%). In addition, only a fraction (typically 10–20%) of the genetic susceptibility is known. In both Mendelian and complex disease manifestations the gene findings serve clues as to the underlying pathophysiology. Only for the case of Mendelian liver affections do genetic findings have clinically useful predictive power. Modified from [109].
necessarily identify therapeutic targets. It is also clear that for the complex phenotypes, genetic findings are also not useful for disease prediction (genetic testing) (Fig. 4) [105,106]. Both these aspects pertain to the lack of knowledge on interacting environmental and other genetic) and the Genuine function of the relevant gene products. In our opinion, the main purpose of genetics in complex phenotypes is to specify experimental research into disease pathophysiology (“pathway detection”). A challenge in forming relevant hypotheses for experimental and translational research following the application of GWAS is that most of the research into the relevant biology has been performed without the prior knowledge of the liver disease associations. Surprising and novel insights are likely to arise from correctly designed experiments on relevant models and human material, exemplified by the unravelling of IL-29 in HCV clearance [89]. These efforts are now urgently needed and will be the source of genuine clinical relevance and therapeutic applications – which so far are lacking for the gene findings per se.

Evident from the above discussions, for monogenic and even oligogenic liver diseases, genetic testing is already relevant. Here, the obstacles to clinical implementation are that of access to technology and associated expertise. WES has now been used extensively, and successfully, for research. Whole-genome sequencing (WGS) has advantages and disadvantages in comparison to WES. There is complete genomic coverage, sample preparation is easier, and there is less bias in the generated sequence [107]. Yields, in terms of a conclusive diagnosis is therefore higher [108]. The costs of WGS are decreasing rapidly, but major hurdles to widespread use of either technique for routine diagnostics remains in the interpretation of data. A number of commercial and mainly research-oriented software packages are now available, but none are perfect and there is no integration with the regular patient information systems. The most critical issue remains however the huge number of variants found in each individual. Within these variants lie not only the “cause” of Mendelian diseases, but, as discussed above, disease-modifiers and predictors of treatment response. Since disease severity varies even for “classic” monogenic liver diseases (Table 1), it is difficult to tell the difference between a susceptibility allele and a disease modifier, and ultimately every disease is more or less complex (Fig. 4). We simply do not yet have the knowledge and tools to implement most of the new data in clinical practice. The only way we will be able to approach this information, and thereby get closer to a vision of “personalized medicine”, is through large cohort studies incorporating detailed clinical and molecular phenotyping alongside the WGS data. Such efforts are needed now more than ever.

Conflict of interest
The authors declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

References


**Review**


[88] Zimmerman HJ. Hepatotoxicity: the adverse effects of drugs and other chemicals on the liver. Lippincott Williams & Wilkins; 1999, 789.


[99] Zimmerman HJ. Hepatotoxicity: the adverse effects of drugs and other chemicals on the liver. Lippincott Williams & Wilkins; 1999, 789.


