Abstract
Genome-wide association studies (GWAS) for autoimmune hepatitis (AIH) and GWAS/genome-wide meta-analyses (GWMA) for primary biliary cholangitis (PBC) and primary sclerosing cholangitis (PSC) have been successful over the past decade, identifying about 100 susceptibility loci in the human genome, with strong associations with the HLA locus and many susceptibility variants outside the HLA locus with relatively low risk. However, identifying causative variants and genes and determining their effects on liver cells and their immunological microenvironment is far from trivial. Polygenic risk scores (PRSs) based on current genome-wide data have limited potential to predict individual disease risk. Interestingly, results of mediated expression score regression analysis provide evidence that a substantial portion of gene expression at susceptibility loci is mediated by genetic risk variants, in contrast to many other complex diseases. Genome- and transcriptome-wide comparisons between AIH, PBC, and PSC could help to better delineate the shared inherited component of autoimmune liver diseases (AILDs), and statistical fine-mapping, chromosome X-wide association testing, and genome-wide in silico drug screening approaches recently applied to GWMA data from PBC could potentially be successfully applied to AIH and PSC. Initial successes through single-cell RNA sequencing (scRNA-seq) experiments in PBC and PSC now raise high hopes for understanding the impact of genetic risk variants in the context of liver-resident immune cells and liver cell subpopulations, and for bridging the gap between genetics and disease.

Keywords | Autoimmune hepatitis · Primary biliary cholangitis · Primary sclerosing cholangitis · Genome-wide association studies (GWAS) · Genome-wide meta-analyses (GWMA) · HLA · Non-HLA · Polygenic risk score · Mediated expression score regression

Introduction
The liver can be affected by three leading forms of complex autoimmune liver diseases (AILDs), in which the immune system attacks different sites in the liver: autoimmune hepatitis (AIH), primary biliary cholangitis (PBC), and primary sclerosing cholangitis (PSC). The exact pathogenesis of these diseases is poorly understood, and available therapeutic approaches are only partially effective. AIH is a rare, chronic progressive disease with a prevalence of approximately 4–42 per 100,000 depending on geographical location [1]. It is characterized by elevated levels of serum transaminases and immunoglobulin G (IgG), inflammatory liver histology, the presence of autoantibodies, and the absence of markers for viral hepatitis [2]. AIH occurs predominantly in middle-aged women but can affect all age groups of both sexes [3]. The exact cause of AIH is unknown, although loss of tolerance against liver antigens is thought to be the major pathophysiological mechanism caused by an unknown trigger in individuals with a genetic susceptibility [4]. Concordance of AIH and familial clustering of AIH were found in twins, with an estimated pair-wise concordance rate in monozygotic twins of 8.7%, suggesting genetic risk factors for AIH [5]. PBC is a rare disease in which a cycle of immune-mediated damage to biliary epithelial cell, cholestasis, and progressive fibrosis over time can lead to end-stage biliary cirrhosis. PBC more commonly affects women, with a prevalence of about 4–58
per 100,000 people [6], and is characterized by autoimmune destruction of small to medium size intrahepatic bile ducts and the presence of anti-mitochondrial antibodies (AMAs), which are present in > 90% of cases [7,8]. The onset of the disease is thought to be due to the interaction of environmental triggers and a genetic predisposition. Genetic risk is consistent with other complex autoimmune diseases, with an estimated sibling relative risk of 10.5 [9] and a pairwise concordance rate in identical twins of 63%, which is among the highest reported in autoimmune diseases [10]. PSC is a rare disease characterized by multifocal biliary strictures and progressive liver disease. In PSC, there is autoimmune damage to the medium to large bile ducts leading to concentric and obliteratorive fibrosis and structuring [11]. The prevalence of PSC is approximately 10 per 100,000 [12]. The most common positive autoantibodies are perinuclear antineutrophil autoantibodies (pANCA), which are detected in about 80% of patients but are not specific enough for diagnosis [13,14]. Unlike AIH and PBC, men are more commonly affected by PSC than women [15]. First-degree relatives of patients with PSC have an 11.5-fold increased risk of PSC [16].

AIH, PBC, and PSC are not yet curable. Progression of the disease, especially in PSC, often leads to liver transplantation or death. AIH is usually treated by regulating the immune system with steroid and thiopurine-based treatments [17], which is why therapy in AIH is often associated with significant side effects [4]. In many cases, the drugs can permanently suppress AIH and allow patients to live a normal life. If the disease progresses to cirrhosis despite consistent therapy, liver transplantation is the only option. Two drugs, the natural bile acid ursodeoxycholic acid (UDCA) and the semi-synthetic bile acid obeticholic acid (OCA), are approved and mainly used for the treatment of PBC. Long-term therapy with UDCA is often successful if started early [18], but many patients respond poorly to both agents, putting them at risk of progressive liver disease [19]. Therapy of PSC is also carried out with UDCA, but in contrast to PBC, the success of therapy in PSC is rather limited. In certain circumstances, liver transplantation is the only therapeutic option. Biologics, such as anti-tumor necrosis factor alpha (TNF-α) or B cell-depleting antibodies [20], and many new drugs are currently being investigated for the treatment of AILDs [21]. Unfortunately, there are still no promising drugs that target the (unknown) key pathogenic processes in the early phase of disease progression [22]. Of central importance is the improvement of risk stratification strategies, which requires in-depth, longitudinal phenotyping of patients using multi-omics data analysis. The different course of the disease and the different response of patients to treatment could also be related to the heterogeneous genetic background of individual patients, which translates into a heterogeneous clinical phenotype [23]. Elucidating the genetic architecture of AIH, PBC, and PSC is likely to contribute to a better understanding of these diseases by identifying causative genes and downstream signaling pathways that can be influenced pharmacologically. Genetic research should complement future work to identify the as yet unknown environmental risk factor(s) responsible for the development of autoimmune liver disease through interaction with genetic factors [24]. One of the greatest challenges in genetic epidemiological studies remains deriving a functional biological interpretation of the results from GWAS. In the following, I describe the current status from GWAS studies for AIH, PBC, and PSC and briefly outline what additional work I believe is promising to better understand the genetic component and its biological contribution.

**HLA-related genetic associations from GWAS**

AIH, PBC, and PSC show a strong association with classical human leukocyte antigen genes (class I and II HLA genes; region on chromosome 6p21), which is a common feature of autoimmune diseases, with HLA susceptibility variants usually having a much greater impact than any other risk variant in the genome [25]. Proteins encoded by HLA genes are expressed on cell surfaces and present processed antigens to immune cells, which then activate downstream immune processes. A GWAS is an association study involving several million single nucleotide polymorphisms (SNPs) to determine the contribution of genetic variants to disease susceptibility [26]. The first GWAS for AIH [27], PBC [28], and PSC [29] revealed genetic associations close to classical HLA susceptibility alleles discovered before the GWAS era [30]. Subsequent GWAS and genome-wide meta-analyses (GWMA) used HLA imputation methods to investigate a variety of HLA alleles as possible susceptibility alleles. HLA imputation is a method of deriving HLA types for patients and controls in GWAS studies by imputing (predicting) genotypes of HLA genes using regional SNPs and a SNP-HLA-allele reference panel for imputation [31]. In a first GWAS for AIH [27], involving 649 adult AIH patients and 13,436 controls, followed by replication in 451 patients and 4103 controls, the strongest genome-wide significant ($P < 5 \times 10^{-8}$) association signal for SNP rs2187668 was found at 6p21.32, and HLA imputation assigned the SNP signals to HLA-DRB1*03:01, which is considered the primary AIH susceptibility allele (Table 1); HLA-DRB1*04:01 was identified as another independent (i.e., secondary) AIH susceptibility allele by conditional forward stepwise logistic regression analysis [32]. Recently, a meta-analysis of two GWAS study populations from China (1622 Chinese AIH type 1 patients and 10,466 population controls) identified a SNP association signal near HLA-B [33]. Results of GWAS
cases and 8514 controls of European ancestry [34], with four HLA alleles with genome-wide significance were identified in a HLA Immunochip fine-mapping study of 2861 PBC patients and 1196 controls found a primary association signal in the Han Chinese population [39]. Subsequently, a GWAS and HLA fine mapping study of 1126 Han Chinese PBC patients and 1770 controls showed that DRB1 (with DRB1*08:03) and/or DQB1 (with DQB1*03:01) picked up most of the signals, with DPB1 (DPB1*17:01) being an independent locus [40]. Interestingly, the protective DQB1*03:01 allele for PBC has been identified as a secondary association signal in populations of European origin and in Japan, whereas it is considered a primary association signal in the Han Chinese population. For a complete list of HLA susceptibility alleles for PBC, including alleles from candidate studies, see Gerussi et al. [41]. For PSC, five independent HLA association clusters for PBC identified by conditional logistic regression (Table 1). The first three independent HLA susceptibility alleles (clusters representatives) DQA1*04:01, DQB1*06:02 and DQB1*03:01 confirmed findings from serological studies [35–37], with the fourth allele DRB1*04:04 not previously associated with PBC. In another Immunochip HLA fine-mapping study of 676 Italian PBC cases and 1440 controls, three DRB1 (DRB1*08, DRB1*11, DRB1*14) and one DPB1 (DPB1*03:01) susceptibility cluster were identified through conditional analysis [38], although DRB1*14 and DPB1*03:01 did not meet the genome-wide significance threshold and are therefore not listed in Table 1. An HLA genotyping study of 1200 Japanese PBC patients and 1196 controls found a primary contribution of DQB1*06:04 and DQB1*03:01 to PBC susceptibility [39]. Subsequently, a GWAS and HLA fine mapping study of 1126 Han Chinese PBC patients and 1770 controls showed that DRB1 (with DRB1*08:03) and/or DQB1 (with DQB1*03:01) picked up most of the signals, with DPB1 (DPB1*17:01) being an independent locus [40]. Interestingly, the protective DQB1*03:01 allele for PBC has been identified as a secondary association signal in populations of European origin and in Japan, whereas it is considered a primary association signal in the Han Chinese population.

| Disease | HLA allele | Effect direction | Population | Lead/secondary signal | Reference |
|---------|------------|------------------|------------|-----------------------|-----------|
| AIH     | DRB1*03:01 | Risk             | Netherlands/Germany | Lead | De Boer et al. (2014) [27] |
| AIH     | DRB1*04:01 | Risk             | Netherlands/Germany | Secondary | De Boer et al. (2014) [27] |
| AIH     | HLA-B56   | Risk             | China       | Lead | Li et al. (2022) [33] |
| PBC     | DQA1*04:01 | Risk             | UK          | Lead | Liu et al. (2012) [34] |
| PBC     | DQB1*06:02 | Protective       | UK          | Secondary | Liu et al. (2012) [34] |
| PBC     | DQB1*03:01 | Protective       | UK          | Secondary | Liu et al. (2012) [34] |
| PBC     | DRB1*04:04 | Risk             | UK          | Secondary | Liu et al. (2012) [34] |
| PBC     | DRB1*08   | Risk             | Italy       | Lead | Invernizzi et al. (2012) [38] |
| PBC     | DRB1*11   | Protective       | Italy       | Secondary | Invernizzi et al. (2012) [38] |
| PBC     | DQB1*06:04 | Protective       | Japan       | Lead | Yasunami et al. (2017) [39] |
| PBC     | DQB1*03:01 | Protective       | Japan       | Secondary | Yasunami et al. (2017) [39] |
| PBC     | DQB1*03:01 | Protective       | China       | Lead | Wang et al. (2020) [40] |
| PBC     | DPB1*17:01 | Risk             | China       | Secondary | Wang et al. (2020) [40] |
| PBC     | DRB1*08:03 | Risk             | China       | Secondary | Wang et al. (2020) [40] |
| PSC     | B*08:01   | Risk             | Europe and North America | Lead | Liu et al. (2013) [42] |
| PSC     | DQA*01:03 | Risk             | Europe and North America | Secondary | Liu et al. (2013) [42] |
| PSC     | DQA*05:01 | Risk             | Europe and North America | Secondary | Liu et al. (2013) [42] |
| PSC     | DRB1*15:01 | Risk             | Europe and North America | Secondary | Liu et al. (2013) [42] |
| PSC     | DQA*01:01 | Risk             | Europe and North America | Secondary | Liu et al. (2013) [42] |

and HLA fine-mapping for other populations as well as transethnic studies (i.e., studies across globally different study populations with largely different haplotype structure) are not available for AIH but would be of great importance to identify additional HLA susceptibility alleles and determine susceptibility to AIH across populations. In PBC, a total of 14 HLA alleles with genome-wide significance were identified in a HLA Immunochip fine-mapping study of 2861 PBC cases and 8514 controls of European ancestry [34], with four independent HLA association clusters for PBC identified by conditional logistic regression (Table 1). The first three independent HLA susceptibility alleles (clusters representatives) DQA1*04:01, DQB1*06:02 and DQB1*03:01 confirmed findings from serological studies [35–37], with the fourth allele DRB1*04:04 not previously associated with PBC.
stratification, sex, and others. However, conditional forward stepwise regression analysis has several methodological disadvantages. First, as the number of conditioning steps increases, so does the number of statistical tests. If there are \( m \) alleles in the region of interest, about \( k^m \) statistical tests are performed after \( k \) consecutive steps, which significantly increases the probability of a false positive result. Second, if \( m \) is large, perhaps close to the number of individuals in the GWAS study, and if a liberal significance threshold is used to include alleles at each step, the forward selection procedure becomes unstable and is too optimistic about the disease variation explained by the selected allele. For this reason, I have listed in Table 1 only statistically independent and genome-wide significant HLA susceptibility alleles (cluster representatives) that are also genome-wide significant after conditioning regression analysis. The identification (fine-mapping) of a complete set of potentially “causal” HLA alleles in the overall context of all class I and I genes requires the use of high-quality multi-ethnic reference panels from different genetic backgrounds [46, 47], highly accurate HLA type imputation algorithms [31], the study of non-additive and interaction effects [48], inclusion of amino acid alleles composing HLA alleles [49], and functional fine-mapping approaches [50–52]. Future HLA fine-mapping studies for AIH, PBC, and PSC therefore have the potential to further refine these signals from previous GWAS and HLA imputation studies.

**Non-HLA-related genetic associations from GWAS**

As the number of GWAS studies for AIH, PBC, and PSC increased, it became clear that the effect sizes of non-HLA associations were much smaller compared with associations with HLA alleles. For AIH, a coding variant rs3184504 in \( SH2B3 \) was identified as a susceptibility variant with genome-wide significance in study populations from the Netherlands and Germany [27]; two additional non-coding variants (rs72929257 near \( CTLA4 \) and rs6809477 at \( SYNPR \)) were recently identified across two study populations from China [33]. For PBC, the largest GWMAs of European (Asian) case–control populations yielded 45 (12) loci with genome-wide significance, with a total of 55 genome-wide significant non-HLA susceptibility loci identified in one or the other GWMA [53]. GWMAs for PSC identified a total of 22 genome-wide significant non-HLA susceptibility loci in Europeans [42, 54, 55]; non-European as well as trans-ethnic GWMAs have not yet been performed for PSC. Table 2 summarizes all non-HLA susceptibility variants from GWMA studies for AIH, PBC, and PSC. For a review of non-HLA susceptibility variants for AIH, PBC, and PSC from monocentric studies, including candidate studies, see Engel et al. [56], Gerussi et al. [41], and Chung/Hirschfeld [57], respectively. For PBC, Bayesian fine-mapping was recently performed for the 55 non-HLA PBC susceptibility loci [53]. Bayesian methods are particularly well suited for fine-mapping of non-HLA loci to identify statistically “causal” sets of variants [58] and have been successfully used, for example, to fine-map inflammatory bowel disease (IBD) risk loci to single-variant resolution [59]. For 40 (9) of 55 non-HLA PBC susceptibility loci, the association signal was best explained by a single variant (posterior probability \( \geq 0.5 \)) across European and Asian (Asian only) populations (Table 2); for AIH and PSC, Bayesian fine-mapping for established non-HLA susceptibility regions remains to be performed. Chromosome X association analysis has unfortunately been neglected in most GWAS studies [60]. Recently, a chromosome X-wide association study for PBC identified a genome-wide significant locus at Xp11.23 (the locus includes the \( GRIPAP1 \) gene; see Table 2) in East Asian PBC case–control study populations, which also shows an association signal (not yet genome-wide significant) across European and Asian PBC case–control study sets [61]. X-linked inheritance models in GWAS/GWMA for PSC and AIH thus have the potential to reveal additional genetic associations. Figure 1 summarizes the polygenic landscape of genome-wide significant HLA and non-HLA susceptibility variants (each of which was named in association with a nearby candidate gene) for AIH, PBC, and PSC.

**SNP-based (co)-heritability**

The proportion of genetic variance in liability (i.e., heritability explained by individual genetic variants for binary outcomes; additive model based on disease prevalence, relative risks, and allele frequencies [62]) for PBC explained by four major HLA alleles [34] (Table 1) and 26 independent genome-wide significant non-HLA susceptibility variants (subset from Table 2) was estimated to be 4.9% and 1.4%, respectively, which account together for 16.2% of total PBC heritability [34]. More recent calculations using the much larger PBC GWAS sets available today would be desirable, as would estimates for AIH. SNP-based heritability for PSC explained by 16 independent genome-wide significant loci (including major HLA alleles) account for 7.3% of total PSC heritability [42]; again, more up-to-date estimates would be desirable. The discrepancy between the variance caused by common SNPs and the expected heritability of AILDs from twin studies is referred to as missing heritability [63]. This gap in heritability can have several possible causes; Heritability from twin studies could be overestimated because common environmental factors or non-additive effects were not taken into account. On the other hand, part of the heritability could be due to genetic
| Disease | Variant | Chromosome:position | Candidate gene | Population | Fine-mapped to single variant | Reference |
|---------|---------|---------------------|----------------|------------|-----------------------------|-----------|
| AIH     | rs184504 | 12:111,884,608      | SH2B3          | Netherlands/Germany | NA | De Boer et al. (2014) [27] |
| AIH     | rs72929257 | 2:204,982,643       | CTLA4          | China       | NA | Li et al. (2022) [33]       |
| AIH     | rs6809477 | 3:63,563,282        | SYNPR          | China       | NA | Li et al. (2022) [33]       |
| PBC     | rs867436  | 1:2,523,723         | MMEL1          | European ancestry | rs867436 | Cordell et al. (2021) [53] |
| PBC     | rs6679356 | 1:67,820,194        | IL12RB2        | European ancestry | rs6679356 | Cordell et al. (2021) [53] |
| PBC     | rs10802191 | 1:117,065,083       | CD58           | European ancestry | rs10802191 | Cordell et al. (2021) [53] |
| PBC     | rs945635  | 1:157,670,290       | FCRL3          | European ancestry | rs945635 | Cordell et al. (2021) [53] |
| PBC     | rs12123169 | 1:197,780,966       | DENND1B        | European ancestry | rs12123169 | Cordell et al. (2021) [53] |
| PBC     | rs5734382 | 1:201,019,059       | IRAV           | European ancestry | rs5734382 | Cordell et al. (2021) [53] |
| PBC     | rs34655300 | 2:25,514,333       | DNM73A         | European ancestry | rs891058 | Cordell et al. (2021) [53] |
| PBC     | rs859767  | 2:135,341,200       | TMEM163        | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs3771317 | 2:191,543,962       | STAT4          | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs9876137 | 3:16,961,265        | PLC12          | European ancestry | rs9876137 | Cordell et al. (2021) [53] |
| PBC     | rs6550965 | 3:25,383,587        | RARB           | European ancestry | rs6550965 | Cordell et al. (2021) [53] |
| PBC     | rs2293370 | 3:119,219,934       | CD80           | European ancestry | rs2293370 | Cordell et al. (2021) [53] |
| PBC     | rs589446  | 3:159,733,527       | IL12A          | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs7674640 | 4:103,540,780       | NFKB1          | European ancestry | rs7674640 | Cordell et al. (2021) [53] |
| PBC     | rs7663401 | 4:106,128,954       | TET2           | European ancestry | rs7663401 | Cordell et al. (2021) [53] |
| PBC     | rs35467801 | 5:35,881,130       | IL7R           | European ancestry | rs35467801 | Cordell et al. (2021) [53] |
| PBC     | rs2546890 | 5:158,759,900       | IL12B          | European ancestry | rs2546890 | Cordell et al. (2021) [53] |
| PBC     | rs2327832 | 6:137,973,068       | TNFAIP3        | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs7805218 | 7:20,378,801        | ITG8           | European ancestry | rs7805218 | Cordell et al. (2021) [53] |
| PBC     | rs60600003 | 7:37,382,465       | ELMO1          | European ancestry | rs60600003 | Cordell et al. (2021) [53] |
| PBC     | rs12531711 | 7:128,617,466      | ICRF           | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs370193557 | 7:138,729,543     | ZC3HAVL        | European ancestry | rs370193557 | Cordell et al. (2021) [53] |
| PBC     | rs11390003 | 9:100,741,912      | TRIM14         | European ancestry | rs11390003 | Cordell et al. (2021) [53] |
| PBC     | rs7097397 | 10:50,025,396       | WDF4           | European ancestry | rs7097397 | Cordell et al. (2021) [53] |
| PBC     | rs58523027 | 11:64,646,986      | IRF7           | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs11601860 | 11:64,110,422      | CCDC88B        | European ancestry | rs11601860 | Cordell et al. (2021) [53] |
| PBC     | rs12419634 | 11:111,239,365     | POU2AF1        | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs201150316 | 11:118,740,104     | CXCR5          | European ancestry | rs201150316 | Cordell et al. (2021) [53] |
| PBC     | rs1800693 | 12:6,440,009        | TNFRSF1A       | European ancestry | rs1800693 | Cordell et al. (2021) [53] |
| PBC     | rs35350651 | 12:111,907,431     | SH2B3          | European ancestry | rs35350651 | Cordell et al. (2021) [53] |
| PBC     | rs9533122 | 13:43,055,002       | TNFSF11        | European ancestry | rs9533122 | Cordell et al. (2021) [53] |
| PBC     | rs9591325 | 13:50,811,220       | DLEU1          | European ancestry | rs9591325 | Cordell et al. (2021) [53] |
| PBC     | rs3784099 | 14:68,749,927       | RAD51B         | European ancestry | rs3784099 | Cordell et al. (2021) [53] |
| PBC     | rs72699866 | 14:93,114,787      | RIN3           | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs59643720 | 14:103,564,807     | TNFAIP2        | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs9652601 | 16:111,174,365      | CLEC16A        | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs1119132 | 16:27,403,469       | IL21R          | European ancestry | rs1119132 | Cordell et al. (2021) [53] |
| PBC     | rs79577483 | 16:68,036,939       | DPEP3          | European ancestry | rs79577483 | Cordell et al. (2021) [53] |
| PBC     | rs11117432 | 16:86,019,271      | IRF8           | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs33938760 | 17:38,044,893       | IKZF3          | European ancestry | rs33938760 | Cordell et al. (2021) [53] |
| PBC     | rs1029464 | 17:44,149,348       | MAPT           | European ancestry | rs1029464 | Cordell et al. (2021) [53] |
| PBC     | rs1808094 | 18:67,526,026       | CD226          | European ancestry | rs1808094 | Cordell et al. (2021) [53] |
| Disease | Variant | Chromosome:position | Candidate gene | Population | Fine-mapped to single variant | Reference |
|---------|---------|---------------------|----------------|------------|------------------------------|-----------|
| PBC     | rs2304256 | 19:10,475,652       | TYK2           | European ancestry | rs2304256 | Cordell et al. (2021) [53] |
| PBC     | rs3745516 | 19:50,926,742       | SPIB           | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs137687 | 22:39,740,078       | SYNCR1         | European ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs10802190 | 1:117,061,384       | CD58           | China/Japan | rs10802190 | Cordell et al. (2021) [53] |
| PBC     | rs4824394 | 2:135,342,452       | TMEM163        | China/Japan | rs4824394 | Cordell et al. (2021) [53] |
| PBC     | rs11889341 | 2:191,943,742      | STAT4          | China/Japan | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs4675370 | 2:204,646,499       | CD28           | China/Japan | rs4675370 | Cordell et al. (2021) [53] |
| PBC     | rs12695386 | 3:119,209,027       | POGLUT1        | China/Japan | rs12695386 | Cordell et al. (2021) [53] |
| PBC     | rs230534 | 4:103,449,041       | NFkB1          | China/Japan | rs230534 | Cordell et al. (2021) [53] |
| PBC     | rs34463936 | 5:35,850,149      | LOC105374724   | China/Japan | rs34463936 | Cordell et al. (2021) [53] |
| PBC     | rs4709148 | 6:167,521,676       | CCR6           | China/Japan | rs4709148 | Cordell et al. (2021) [53] |
| PBC     | rs56211063 | 9:117,585,897      | TNFSF15        | China/Japan | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs4938534 | 11:111,275,133      | BTG4           | China/Japan | rs4938534 | Cordell et al. (2021) [53] |
| PBC     | rs480958 | 11:118,577,990      | LOC105369519   | China/Japan | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs12942330 | 17:37,939,839      | IKZF3          | China/Japan | rs12942330 | Cordell et al. (2021) [53] |
| PBC     | rs13416555 | 2:8,441,735        | ID2            | European + Asian ancestry | rs891058 | Cordell et al. (2021) [53] |
| PBC     | rs10581773 | 2:204,660,748      | CD28           | European + Asian ancestry | rs10581773 | Cordell et al. (2021) [53] |
| PBC     | rs60643069 | 5:100,238,073      | STSSIA4        | European + Asian ancestry | rs141002831 | Cordell et al. (2021) [53] |
| PBC     | rs6874308 | 5:141,506,911       | NDFIP1         | European + Asian ancestry | rs10062349 | Cordell et al. (2021) [53] |
| PBC     | rs742108 | 6:106,582,920       | PRDM1          | European + Asian ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs968334 | 6:167,526,096       | CCR6           | European + Asian ancestry | rs3093024 | Cordell et al. (2021) [53] |
| PBC     | rs4733851 | 8:129,264,420       | PVT1           | European + Asian ancestry | Several signals | Cordell et al. (2021) [53] |
| PBC     | rs1322057 | 9:117,578,374       | TNFSF15        | European + Asian ancestry | rs6478109 | Cordell et al. (2021) [53] |
| PBC     | rs10893872 | 11:128,325,553     | ETS1           | European + Asian ancestry | rs10893872 | Cordell et al. (2021) [53] |
| PBC     | rs799469 | 14:35,444,425       | FAM177A1       | European + Asian ancestry | rs712315 | Cordell et al. (2021) [53] |
| PBC     | rs7059064 | X:48,837,087        | GRIPAP1        | China/Japan | NA | Asselta et al. (2021) [61] |
| PSC     | rs7556897 | 2:228,660,112       | AC073065.3     | European ancestry | NA | Ellinghaus et al. (2016) [54] |
| PSC     | rs17032705 | 4:103,432,974      | NFkB1          | European ancestry | NA | Ellinghaus et al. (2016) [54] |
| PSC     | rs12369214 | 12:107,198,611     | RIC8B          | European ancestry | NA | Ellinghaus et al. (2016) [54] |
| PSC     | rs11649613 | 16:11,319,357     | RP11-396B14.2  | European ancestry | NA | Ellinghaus et al. (2016) [54] |
| PSC     | rs3748816 | 1:2,526,746         | MMEL1          | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs74237826 | 2:111,933,001      | BCL2L11        | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs7426056 | 2:204,612,058      | CD28           | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs3749171 | 2:241,569,692      | GPR35          | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs3197999 | 3:49,721,532       | SIK2           | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs3745516 | 3:49,721,532       | GPR35          | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs11168249 | 12:48,208,368   | HDAC7          | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs3184504 | 12:111,884,608     | SH2B3          | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs725613 | 16:11,169,683      | CLEC16A        | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs1788097 | 18:67,543,688      | CD226          | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs60652743 | 19:47,205,707     | PRKDC          | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs2836883 | 21:40,466,744      | PSMG1          | European ancestry | NA | Ji et al. (2016) [55] |
| PSC     | rs1893592 | 21:43,855,067      | UBASH3A        | European ancestry | NA | Ji et al. (2016) [55] |
variants that have remained undetected so far, such as rare genetic variants or sex chromosome variants. Non-additive genetic effects, such as dominance effects ($\delta_2^SNP$) or additive-by-additive interaction effects ($\eta_2^SNP$; epistasis), might describe part of the disease heritability in AILDs. In a recent study of 70 (continuous) complex traits from the UK Biobank (with more than 60,000 individuals for each trait), the average epistatic variance across all traits ($\hat{\eta}_2^SNP = 0.055$) was estimated to be significantly higher than the average variance for dominance effects ($\hat{\delta}_2^SNP = 0.001$), but still significantly lower than the average variance for additive effects ($\hat{h}_2^SNP = 0.208$) [64]. Genome-wide interaction studies (GWIS) are therefore another interesting approach, but GWIS for PBC, PSC, and AIH would require many times the current sample size (which is difficult to realize) due to the exponential increase in statistical tests and extremely longer calculation times [65], although alternative computer architectures such as GPUs could help here [66, 67].

Because up to 9% and 7% of AIH patients have clinical overlap with PBC and PSC [27], respectively, it is reasonable to assume that there are shared genetic factors for AIH, PBC, and PSC [68], and also for other immune-mediated diseases, as shown by Immunochip studies for PBC and PSC [34, 38, 42]. In AIH, genetic risk is shared with type 1 diabetes for $DRB1^*04:01$ [69] and systemic lupus erythematosus for $DRB1^*03:01$ [70]. The association of AIH with the $SH2B3$ locus has also been identified as a genetic risk factor for PSC and PBC; more specifically, even the same risk variant rs3184504 in SH2B3 has been identified for AIH and PSC (see Table 1). Genetic relationships between disease pairs on a genome-wide level can be investigated by genome-wide genetic correlation analyses that quantify genome-wide SNP-based heritability ($h_g^2SNP$) in a bivariate model and provide information on potential coheritability between diseases [71–73]. Unfortunately, the shared genetics of AIH, PBC, and PSC have been studied only to a limited extent genome-wide and
only partially with other immune-mediated comorbidities. When SNPs from the extended major histocompatibility complex (MHC) region (chromosome 6 region of 25–34 Mb including HLA genes) were excluded, a significant genetic correlation was observed between PSC and inflammatory bowel disease (ulcerative colitis, \( r_g = 0.29 \); Crohn’s disease, \( r_g = 0.04 \)). For 196 fine-mapped regions of the Immunochip, the genetic correlation between PSC and ulcerative colitis \( (r_g = 0.64) \), PSC and Crohn’s disease \( (r_g = 0.35) \), and PSC and ankylosing spondylitis \( (r_g = 0.33) \) was highest compared with non-immune diseases [54]. To accurately identify possible shared causal variants in AIH-, PBC-, and PSC-associated regions, Bayesian tests of colocalization may be useful [74] and could provide an indication of whether there are common or independent causal variants for the same genomic regions. For example, six of 14 loci associated with both PSC and IBD showed strong evidence of a shared causal variant with UC, CD, or both [55]; colocalization analyses for AIH, PBC, and PSC could provide further insight into shared genetic structure. Future genome-wide comparisons between (worldwide) study populations with AIH, PBC, and PSC would provide the opportunity to identify the potentially shared landscape of AILDs.

**Polygenic risk scores**

Genome-wide SNP-based (co-)heritability estimation provides information on the proportion of (co-)heritability explained for (pairs of) diseases and measures pleiotropy (vertical and horizontal [75]) between diseases, but does not provide an estimate of individual patient risk based on genetic markers. Given the polygenic nature of AILDs (Fig. 1) and the fact that individual risk variants from GWAS/GWMA describe only a fraction of the heritability, a combined genetic burden across all genetic variants can be calculated to identify individuals at significant increased risk. A polygenic risk score (PRS) is an estimate of an individual’s genetic susceptibility to a disease calculated based on that individual’s genotype profile and relevant data from GWAS. A study by Khera and colleagues [76] revived the topic of PRS for common complex diseases and showed for coronary artery disease (CAD), that a PRS identifies 20-times more individuals at comparable or greater risk than did previous studies for monogenic mutations. Therefore, identifying individuals with high (low) PRS in a population-based sample may provide an opportunity to identify those with the highest (lowest) genetic risk. However, the utility of PRS-based risk estimates for AIH, PBC, and PSC is limited by the small effect sizes of the identified susceptibility variants (see Fig. 1). Using genome-wide data from UK Biobank, Khera and colleagues showed that individuals in the top 5% of a PRS for CAD had a 3.34-fold risk [OR(CI95%) = 3.34(3.12–3.58)]; \( P_{\text{logistic\_regression}} = 6.5 \times 10^{-264} \) compared with the remaining 95% of the general population. To provide an estimate for PSC here (although our GWAS study data here is not a population-based sample), I calculated a PRS from the summary statistics of the most recent GWMA for PSC [55] and determined the distribution of the PRS for 628 GWMA-independent German PSC cases and 4,272 healthy controls (Methods). A PRS for PSC runs the risk of creating a mixed PRS for PSC and IBD, as patients with PSC have a highly increased incidence of IBD (called PSC with concomitant IBD or PSC-IBD); however, PSC-IBD has clinical differences from classical IBD [77] and appears to be genetically distinct from classical IBD phenotypes [54] (see below). Individuals in the top 5% of the PRS for PSC had a 5.99-fold risk \( \text{OR(CI95%) = 5.99(4.52–7.92);} \) \( P_{\text{logistic\_regression}} = 1.06 \times 10^{-46} \) and adjusted for sex and genetic ancestry] compared with the remaining 95% of the general population (Fig. 2a). The OR should be interpreted as the factor by which the chance of developing the disease increases if a person has a positive PRS test result (here, being in the top 5%). However, with an OR = 5.99 and an underlying false positive rate (1-specificity) of 5%, the detection rate (sensitivity) is only 24% (Fig. 2b), resulting in an OAPR (odds of being affected given a positive result) of 1:2087, compared with the overall prevalence of 1:10,000 for PSC in the general population. On this basis, diagnostic genetic testing would be inappropriate. The fact that even a very high OR is associated with low predictive power of a diagnostic test may seem counterintuitive. It is largely explained by the fact that the genetic risk variants are widespread in the general population, so almost everyone can be affected by these causes, even if not everyone is or becomes ill because of the genetic burden. Therefore, the PRSs for AILDs in their current form are not suitable for diagnostic testing. On the other hand, for example, a possibly increased PRS for PSC in patients with AIH (compared with healthy controls) could indicate shared genetic risk factors (pleiotropy) between PSC and AIH, but an increased cross-disease PRS could be due to multiple causes such as diagnostic misclassification, molecular subtypes, or excessive comorbidity (collectively referred to as heterogeneity). Cross-locus correlation analyses of loci associated with disease B in cases of disease A (and vice versa) can help distinguish pleiotropy from heterogeneity. For PSC and IBD, for example, we have shown that PSC-IBD is likely to be a distinct disease at the genetic level, sharing some genetic factors with IBD, but genetically distinct from classical IBD phenotypes such as CD and UC [54].
Genetically regulated expression and single cell analyses

Susceptibility variants identified by GWAS are often located at genomic positions with methylation, expression, and protein-quantitative trait loci (mQTLs, eQTLs, pQTLs), but it remains unclear whether this overlap is due to methylation, expression, and protein levels "mediating" genetic effects on disease. Cordell and colleagues applied Bayesian tests for colocalization between GWMA summary statistics of PBC and genome-wide mQTLs, eQTLs, and pQTLs data from large-scale consortium projects ALSPAC [78], GTEx [79], and INTERVAL [80] and suggested that the genetic architecture of PBC influences susceptibility to the disease primarily by affecting the regulation of expression of potentially causal genes [53]. To assess whether this might also be expected for PSC, I calculated the correlation between GWMA summary statistics of PSC [55] and summary statistics of large tissue-specific eQTL studies from GTEx using an approach developed by Yao et al. [81] called mediated expression score regression (MESC). I estimated that the heritability for PSC mediated by the cis genetic component of gene expression levels ($h^2_{med}/h^2_g$) averaged 38.4% for 48 tissues used in the GTEx project (Methods). This value of 38.4% for PSC is among the top of all disease-specific MESC values published in the work of Yao et al. who studied 42 diseases and human traits in the same 48 tissues, including ulcerative colitis with a similarly high published value of 38.2%. Therefore, it is hypothesized that in PBC and PSC (perhaps also in AIH), the gap between genetic approaches and the resulting disease phenotype can be reduced by the transcriptome. Because eQTL data from bulk tissues are thought to be a poor surrogate for eQTL data in causal cell types/contexts and little is known about the composition of intrahepatic immune cells and their contribution to disease pathogenesis, measurement of context-specific expression [82] and expression in single cells may allow the identification of genetic variants that impact key regulatory networks in AILDs [83]. Using single-cell RNA sequencing (scRNA-seq) techniques, Poch and colleagues generated the first atlas of intrahepatic T cells in PSC and identified a previously uncharacterized population of liver-resident CD4+ T cells that likely contribute to the pathogenesis of PSC [84]. Xiang and colleagues [85] developed a computational framework to integrate GWAS summary statistics with scRNA-seq data and revealed genetically modulated liver cell subpopulations for PBC. They found that cholangiocytes show significant enrichment with PBC-related genetic association signals, with the ORMDL3 gene showing the highest expression level in cholangiocytes compared with other liver cells. Such combined genetics/single cell omics studies have the potential to identify the causative genes for AIH, PBC, and PSC in a disease-specific context.

![Fig. 2](image-url)
Genome-wide screenings for drug reuse

More than 25% of drugs entering clinical development fail because of lack of efficacy, but drugs with supportive genetic evidence are twice as likely to succeed in clinical development as drugs without supportive genetic evidence [86]. Thus, one potential approach is to test drugs with genetic support that have been successfully used in practice for other immune-mediated diseases for their transferability to AILDs. To this end, Cordell and colleagues have developed an elegant in-silico method for identifying drugs that can improve (or exacerbate) PBC in this prediction [53], highlighting the potential of genomic screening approaches for drug discovery and prediction of opposing drug effects in complex diseases. Briefly, they adapted a network-based approach to drug proximity screening from Guney et al. [87] for PBC candidate genes from GWAS risk loci by calculating a measure of proximity ($z$-score) between candidate genes and known drug targets (from agents stored in Drugbank), where a low $z$-score indicates recommended use of these agents (because an agent’s gene targets are closer to susceptibility genes than expected by chance) and a high $z$-score represents non-recommended use (because an agent’s targets are not closer to susceptibility risk genes than expected by chance). Major drugs predicted to improve PBC included ustekinumab, a monoclonal antibody against IL-12/23 used to treat psoriasis and Crohn’s disease [88]; however, a proof-of-concept study has not shown benefit of ustekinumab for patients with PBC [89]. Major drugs that could exacerbate PBC included the pharmacologic interferons interferon alpha-2a and interferon beta-1b. The drugs already approved for PBC, fenofibrate, bezafibrate, and OCA were confirmed; interestingly, UDCA did not achieve a significant result, suggesting that genetics does not play a role in this case. A similar analysis would be desirable for AIH and PSC.

Conclusion

A GWAS for AIH and several GWMAs for PBC and PSC have been successfully conducted and have identified a variety of genetic factors associated with AIH, PBC, and PSC. Some of these studies have already identified disease-causing variants by statistical fine-mapping and provided important biological insights into pathogenesis. Some statistical epidemiological approaches, such as statistical fine-mapping, chromosome X-wide association testing, and genome-wide screens for drug reuse, that have already been successfully performed in PBC, could also be applied to AIH and PSC. Large-scale cross-disease GWMAs to explore the shared genetic landscape of AIH, PBC, and PSC are still lacking. Merging genetic and statistical results with single-cell transcriptomic data from relevant cell types and liver tissue is likely to provide more accurate insights into the effects of genetic factors on liver cells and their immunologic microenvironment.

Software

PRS derivation: The LDpred2 software [90] (https://github.com/private/bigsnpr) was used to generate a PRS. The MESC software [81] (https://github.com/douglassroy/mesc) was used for estimating heritability mediated by assayed gene expression levels.

Acknowledgements I thank the IPSCSG for permission to using GWMA summary statistics of PSC [55] for PRS and MESC analysis. Written, informed consent was obtained from all study participants, and the institutional ethical review committees of the participating centers approved all protocols. I would like to thank Elke Matthias Wacker for his support in preparing the figures. The study received infrastructure support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Cluster of Excellence 2167 “Precision Medicine in Chronic Inflammation (PMI)” (EXC 2167-390884018).

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The author declares no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Boberg KM (2002) Prevalence and epidemiology of autoimmune hepatitis. Clin Liver Dis 6(3):635–647
2. Hennes EM, Zeniya M, Czaja AJ, Pares A, Dalekos GN, Krawitt EL, Bittencourt PL, Porta G, Boberg KM, Hofer H, Bianchi FB, Shibata M, Schramm C, Eisenmann de Torres B, Galle PR, McFarlane I, Dienes HP, Lohse AWG (2008) International autoimmune hepatitis, simplified criteria for the diagnosis of autoimmune hepatitis. Hepatology 48(1):169–76
20. Than NN, Hodson J, Schmidt-Martin D, Taubert R, Wawman RE, Botton M, Gautam N, Bock J, Jones R, Appanna GD, Godkin A, Montano-Loza AJ, Lammert F, Schramm C, Manns MP, Swain M, Burak KW, Adams DH, Hirschfeld GM, Oo YH (2019) Efficacy of rituximab in difficult-to-manage autoimmune hepatitis: Results from the International Autoimmune Hepatitis Group. JHEP Rep 1(6):437–445

21. Gerussi A, Luca M, Cristofieri L, Ronca V, Mancuso C, Milani C, D’Amato D, O’Donnell SE, Carbone M, Invernizzi P (2020) New therapeutic targets in autoimmune cholangiopathies. Front Med (Lausanne) 7:117

22. Bossi L, Gerussi A, Lygoura V, Mells GF, Carbone M, Invernizzi P (2018) Support of precision medicine through risk-stratification in autoimmune liver diseases - histology, scoring systems, and non-invasive markers. Autoimmun Rev 17(9):854–865

23. Karlsen TH, Chung BK (2015) Genetic risk and the development of autoimmune liver disease. Dig Dis 33(Suppl 2):13–24

24. Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodier YK, Lush MJ, Povey S, Talbot CC Jr, Wright MW, Wain HM, Trowsdale J, Ziegler A, Beck S (2004) Gene map of the extended human MHC. Nat Rev Genet 5(12):889–899

25. Tam V, Patel N, Turcotte M, Bosse Y, Pare G, Meyre D (2019) Benefits and limitations of genome-wide association studies. Nat Rev Genet 20(8):467–484

26. de Boer YS, van Gerven NM, Zwiers A, Verwer BJ, van Hoek B, van Erpecum KJ, Beuers U, van Buuren HR, Drenth JP, den Ouden JW, Verdonck RC, Koek GH, Brouwer JT, Guichelaar MM, Vrolijk JM, Kraai G, Mulder CJ, van Nieuwkerk CM, Fischer J, Berg T, Stiekel F, Sarrazin C, Schramm C, Lohse AW, Weiler-Normann C, Lerch MM, Naukkarinen A, Wirjatmi C, Franke L, Bouma G, Dutch Autoimmune Hepatitis Study G, LifeLines Cohort S, P. (2014) Study of health in, Genome-wide association study identifies variants associated with autoimmune hepatitis type 1. Gastroenterology 147(2):443–52:e5

27. Hirschfeld GM, Liu X, Xu C, Lu Y, Xie G, Lu Y, Gu X, Walker EL, Jing K, Juran BD, Mason AL, Myers RP, Pellekian KM, Ghent CN, Coltscu C, Atkinson EJ, Heathcote EL, Lazaridis KN, Amos CI, Siminovich KA (2009) Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. N Engl J Med 360(24):2544–2555

28. Karlsen TH, Franke A, Melum E, Kaser A, Hoy JR, Balschun T, Lie BA, Bergquist A, Schramm C, Weissmuller TI, Gotthardt R, Rust C, Philipp EE, Fritz T, Henckaerts L, Weersma RK, Stokkers P, Ponsioen CY, Wijmenga C, Sterneck M, Nothnagel M, Hampe J, Teufel A, Runz H, Rosenstiel P, Stielh A, Vermeire S, Beuers U, Manns MP, Schumpf E, Boberg KM, Schreiber S (2010) Genome-wide association analysis in primary sclerosing cholangitis. Gastroenterology 138(3):1102–1111

29. Czaja AJ, Strettell MD, Thomson LJ, Santrach PJ, Moore SB, Donaldson PT, Williams R (1997) Associations between alleles of the major histocompatibility complex and type 1 autoimmune hepatitis. Hepatology 25(2):317–323

30. Naito T, Okada Y (2022) HLA imputation and its application to genetic and molecular fine-mapping of the MHC region in autoimmune diseases. Semin Immunopathol 44(1):15–28

31. Morris AP (2014) Fine mapping of type 2 diabetes susceptibility loci. Curr Diab Rep 14(11):549

32. Li Y, Sun Y, Liu Y, Wang B, Li J, Wang H, Zhang H, Wang X, Han X, Lin Q, Zhou Y, Hu L, Song Y, Bao J, Gong L, Sun M, Yuan X, Zhang X, Lian M, Xiao X, Mao Q, Wang Q, Li KK, Du S, Ma A, Li Y, Xu J, Tang S, Shi J, Xu Y, Yang L, Zhang J, Huang Z, Zhou L, Cui Y, Seldin MF, Gershwin ME, Yan H, Zou Z, Zuo X, Tang R, Ma X, Chinese AIHC (2022) Genome-wide

33. Bahar R, Wong KA, Liu CH, Bowls CL (2018) Update on new drugs and those in development for the treatment of primary biliary cholangitis. Gastroenterol Hepatol (N Y) 14(3):154–163

34. Weiler-Normann C, Schramm C, Quasas A, Weigard C, Glaubke C, Pannicke N, Moller S, Lohse AW (2013) Infliximab as a rescue treatment in difficult-to-treat autoimmune hepatitis. J Hepatol 58(3):529–534

35. Krawitt EL (2006) Autoimmune hepatitis. N Engl J Med 354(1):54–66

36. Pape S, Schramm C, Gevers TJ (2019) Clinical management of autoimmune hepatitis. United European Gastroenterol J 7(9):1156–1163

37. Gronbaek L, Vilstrup H, Pedersen L, Christensen K, Jepsen P (2018) Family occurrence of autoimmune hepatitis: a Danish nationwide registry-based cohort study. J Hepatol 69(4):873–877

38. Tanaka A, Leung PSC, Gershwin ME (2019) The genetics of primary biliary cholangitis. Curr Opin Gastroenterol 35(2):93–98

39. European Association for the Study of the Liver (2017) Electronic address, L. European Association for the Study of the, EASL Clinical Practice Guidelines: the diagnosis and management of patients with primary biliary cholangitis. J Hepatol 67(1):145–172

40. Lleo A, Leung PSC, Hirschfeld GM, Gershwin EM (2020) The pathogenesis of primary biliary cholangitis: a comprehensive review. Semin Liver Dis 40(1):34–48

41. Jones DE, Watt FE, Metcalf IV, Bassendine MF, James OF (1999) Familial primary biliary cirrhosis reassessed: a geographically-based population study. J Hepatol 30(3):402–407

42. Selmi C, Mayo MJ, Bach N, Ishibashi H, Invernizzi P, Gish RG, Gordon SC, Wright HI, Zweiban A, Podda M, Gershwin ME (2004) Primary biliary cirrhosis in monzygotic and dizygotic twins: genetics, epigenetics, and environment. Gastroenterology 127(2):485–492

43. Karlsen TH, Folseraas T, Thorburn D, Vesterhus M (2017) Primary sclerosing cholangitis - a comprehensive review. J Hepatol 67(6):1298–1323

44. Folseraas T, Melum E, Franke A, Karlsen TH (2011) Genetics in primary sclerosing cholangitis. Best Pract Res Clin Gastroenterol 25(6):713–726

45. Bansu D, Chapman R, Fleming K (1996) Antineutrophil cytoplasmic antibodies in chronic liver diseases: prevalence, titre, specificity, and IgG subclass. J Hepatol 24(5):581–586

46. Chapman RW (1995) The enigma of anti-neutrophil antibodies in ulcerative colitis primary sclerosing cholangitis: important genetic marker or epiphenomenon? Hepatology 21(5):1473–1474

47. Delle Monache M, Salvio A, Fiocca F, Basoli A, Ricci GL (1992) Primary sclerosing cholangitis: an analysis of 37 retrospective cases. J Hepatol 24(9):485–492

48. Bergquist A, Montgomery SM, Bahmanyar S, Olsson R, Danilsson A, Lindgren S, Prytz H, Hultcrantz R, Loof LA, Sandberg-Gertzen H, Almer S, Asking J, Ehnin A, Ekborn A (2008) Increased risk of primary sclerosing cholangitis and ulcerative colitis in first-degree relatives of patients with primary sclerosing cholangitis. Clin Gastroenterol Hepatol 6(8):939–943

49. Mieli-Vergani G, Vergani D, Czaja AJ, Manns MP, Krawitt EL, Vierling JM, Lohse AW, Montano-Loza AJ (2018) Autoimmune hepatitis. Nat Rev Dis Primers 4:18017

50. Carbone M, Mells GF, Pells G, Dawmas WF, Newton JL, Heneghan MA, Neuberger JM, Day DB, Ducker SJ, Consortium UP, Sandford RN, Alexander GJ, Jones DE (2013) Sex and age are determinants of the clinical phenotype of primary biliary cirrhosis and response to ursodeoxycholic acid. Gastroenterology 144(3):560–569 e7; quiz e13–4

51. Bahar R, Wong KA, Liu CH, Bowls CL (2018) Update on new drugs and those in development for the treatment of primary biliary cholangitis. Gastroenterol Hepatol (N Y) 14(3):154–163

52. Weiler-Normann C, Schramm C, Quasas A, Weigard C, Glaubke C, Pannicke N, Moller S, Lohse AW (2013) Infliximab as a rescue treatment in difficult-to-treat autoimmune hepatitis. J Hepatol 58(3):529–534
meta-analysis identifies novel susceptibility loci for autoimmune hepatitis Type 1. Hepatology

34. Liu JZ, Almarri MA, Gaffney DJ, Mells GF, Jostins L, Cordell HJ, Ducker SJ, Day DB, Heneghan MA, Neuberger JM, Donaldson PT, Bathgate AJ, Burroughs A, Davies MH, Jones DE, Alexander GJ, Barrett JC, Sandford RN, Anderson CA, UKPBC Consortium, C. (2012) Wellcome Trust case control, dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. Nat Genet 44(10):1137–41

35. Wassmuth R, Depner F, Danielsen A, Hultcrantz R, Loof L, Olson R, Frytz H, Sandberg-Gertzen H, Wallerstedt S, Lindgren S (2002) HLA class II markers and clinical heterogeneity in Swedish patients with primary biliary cirrhosis. Tissue Antigens 59(5):381–387

36. Donaldson PT, Baragiotta A, Heneghan MA, Floreani A, Venturi C, Underhill JA, Jones DE, James OF, Bassendine MF (2006) HLA class II alleles, genotypes, haplotypes, and amino acids in primary biliary cirrhosis: a large-scale study. Hepatology 44(3):667–674

37. Mullarkey ME, Stevens AM, McDonnell WM, Loubiere LS, Brackensiek JA, Pang JM, Porter AJ, Galloway DA, Nelson JL (2005) Human leukocyte antigen class II alleles in Caucasian women with primary biliary cirrhosis. Tissue Antigens 65(2):199–205

38. Invernizzi P, Ransom M, Raychaudhuri S, Kosoy R, Lleo A, Shugeta R, Franke A, Bossa F, Amos CI, Gregersen PK, Siminovich KA, Cussi D, de Bakker PI, Podda M, Gershwin ME, Seldin MF, Italian PBCSGSG (2012) Classical HLA-DRB1 and DPB1 alleles account for HLA associations with primary biliary cirrhosis. Genes Immun 13(6):461–8

39. Yasunami M, Nakamura H, Tokunaga K, Kawashima M, Nishida N, Hitomi Y, Nakamura M (2017) Principal contribution of HLA-DQ alleles, DQB1*06:04 and DQB1*03:01, to disease resistance against primary biliary cholangitis in a Japanese population. Sci Rep 7(1):11093

40. Wang C, Zheng X, Tang R, Han C, Jiang Y, Wu J, Shao Y, Gao Y, Yu J, Hu Z, Zang Z, Zhao Y, Dai N, Liu G, Wu X, Nie J, Jiang B, Lin M, Li L, Wei Y, Li Y, Gong Y, Dai Y, Wang L, Ding N, Xu P, Chen S, Jiang P, Wang L, Qiu P, Wu Q, Zhang M, Jawed R, Chen R, Zhang Y, Shi X, Zhu Z, Pei H, Huang L, Tian Y, Zhang K, Qiu H, Zhao W, Gershwin ME, Chen W, Seldin MF, Liu X, Ma X, Sun L. (2020) Fine mapping of the MHC region identifies major independent variants associated with Han Chinese primary biliary cholangitis. J Autoimmun 107:102372

41. Gerussi A, Carbone M, Corpechot C, Schramm C, Asselta R, Invernizzi P (2021) The genetic architecture of primary biliary cirrhosis. Eur J Med Genet 64(9):104292

42. Liu JZ, Hof JR, Folseeras T, Ellinghaus E, Rushbrook SM, Doncheva NT, Andreassen OA, Weersma RK, Weismuller TJ, Ekstien B, Invernizzi P, Hirschfeld GM, Gotthardt DN, Pares A, Ellinghaus D, Shah T, Juran BD, Worthington J, Wouters MM, Klareskog L, Elder JT, Forer L, Schonherr S, Fuchsberger C, Smith AV, Fellay J, Carrington M, Haas DW, Guo X, Palmer ND, Shen Y, Rotter JI, Taylor KD, Rich SS, Correa A, Wilson JG, Kathiresan S, Cho MH, Metspalu A, Esko T, Okada Y, Han B, NT-OF PM Consortium, McLaren PJ, Raychaudhuri S (2021) A high-resolution HLA reference panel capturing global population diversity enables multi-ancestry fine-mapping in HIV host response. Nat Genet 53(10):1504–1516

43. Lenz TL, Deutsch AJ, Han B, Hu X, Okada Y, Eyre S, Knapp M, Zhernakova A, Huizinga TW, Abecasis G, Becker J, Boeckxstaens GE, Chen WM, Franke A, Gladman DD, Gockel I, Gutierrez-Achury J, Martin J, Nair RP, Nothen MM, Onengut-Gumuscu S, Rahman P, Rantapaa-Dahlqvist S, Stuart PE, Tsai LC, van Heel DA, Worthington J, Wouters MM, Klareskog L, Elder JT, Gregersen PK, Schumacher J, Rich SS, Wijmenga C, Sunyaev SR, de Bakker PI, Raychaudhuri S (2015) Widespread non-additive and interaction effects within HLA loci modulate the risk of autoimmune diseases. Nat Genet 47(9):1085–1090

44. Hu X, Deutsch AJ, Lenz TL, Onengut-Gumuscu S, Han B, Chen MW, Howson JM, Todd JA, de Bakker PI, Rich SS, Raychaudhuri S (2015) Additive and interaction effects at three amino acid positions in HLA-DQA and HLA-DR molecules drive type 1 diabetes risk. Nat Genet 47(8):898–905

45. Aguilar VRC, Cesar J, Delanoue O, Dermitzakis ET, Meyer D (2019) Expression estimation and eQTL mapping for HLA genes with a personalized pipeline. PLoS Genet 15(4):e1008091

46. Gutierrez-Arcelus M, Baglaenko Y, Arora J, Hannes S, Luo Y, Amariuta T, Teslovich N, Rao DA, Ermann J, Jonsson AH, NT-OPM Consortium, Navarrete C, Rich SS, Taylor KD, Rotter JI, Gregersen PK, Esko T, Brenner MB, Raychaudhuri S (2020) Allele-specific expression changes dynamically during T cell activation in HLA class II and other autoimmune loci. Nat Genet 52(3):247–253

47. Yamamoto F, Suzuki S, Mizutani A, Shigenari A, Ito S, Kametani Y, Kato S, Fernandez-Vina M, Murata M, Morishima Y, Morishima Y, Tanaka M, Kulsik JK, Bahram S, Shina T (2020) Capturing differential allele-level expression and genotypes of all classical HLA loci and haplotypes by a new capture RNA-Seq method. Front Immunol 11:941

48. Cordell HJ, Fryett JJ, Ueno K, Darlay R, Aiba Y, Hitomi Y, Kawashima M, Nishida N, Khor SS, Gervais O, Kawai Y, Nagasaki M, Tokunaga K, Tang R, Shi Y, Li Z, Juran BD, Atkinson EI, Gerussi A, Carbone M, Asselta R, Cheung A, de Andrade M, Baras A, Horowitz J, Ferreira MAR, Sun D, Jones DE, Flack...
S. Spicer A. Mulcahy VL, Byan J, Han Y, Sandford RN, Lazaaridis KN, Amos CI, Hirschfeld GM, Seldin MF, Invernizzi P, Siminovich KA, Ma X, Nakamura M, Mells GF, Consortia PBC, Canadian PBCCC, Chinese PBCCC, Italian PBCSG, Japan PBCGC, Consortium U-P (2021) An international genome-wide meta-analysis of primary biliary cholangitis: novel risk loci and candidate drugs. J Hepatol 75(3):572–581

54. Ellinghaus D, Jostins L, Spain SL, Cortes A, Bethune J, Han B, Park YR, Raychaudhuri S, Pojung JG, Hubenthal M, Folseraa T, Wang Y, Esko T, Metspalu A, Westra HJ, Franke L, Pers TH, Weersma RK, Collil Y, D’Amato M, Halfvarson J, Jensen AB, Lieb W, Degenhardt F, Forstner AJ, Hofmann A, IBDCG International, C. International Genetics of Ankylosing Spondyli-

55. P.S.C.G. International, C. Genetic Analysis of Psoriasis, E. Psoriasis Association Genetics, S. Schreiber, U. Mrowietz, B.D. Juran, K.N. Lazaaridis, S. Brunak, A.M. Dale, R.C. Trembath, S. Weidinger, M. Weichenthal, E. Ellinghaus, J.T. Elder, J.N. Barker, O.A. Andreassen, D.P. McGovern, T.H. Karlsen, J.C. Barrett, M. Parkes, M.A. Brown, Franke A (2016) Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. Nat Genet 48(5):510–518

56. Ji SG, Juran BD, Mucha S, Folseraa T, Jostins L, Melum E, Kumakata N, Atkinson EJ, Schlicht EM, Liu JZ, Shah T, Gutierrez-Achury J, Bober KM, Bergquist A, Vermeiren S, Eksteine B, Durie PR, Forkkila M, Muller T, Schramm C, Sternek M, Weismuller TJ, Gotthardt DN, Ellinghaus D, Braun F, Teufel A, Laudes M, Lieb W, Jacobs G, Beuers U, Weersma RK, Wijmenge C, Marschall HU, Milkiewicz P, Pares A, Kontula K, Chazouilleres O, Invernizzi P, Goede E, Spiess K, Moore C, Sambrook J, Ouwehand WH, Roberts DJ, Danesh J, Floreni A, Gulamhusein AF, Eaton JE, Schreiber S, Coltescu C, Bowls CL, Luuktic VA, Oin JA, Chopra KB, Kowdley KV, Chalasani N, Manns MP, Srivastava B, Mells G, Sandford RN, Alexander CL, Luketic VA, Odin JA, Chopra KB, Kowdley KV, Anderson CA (2017) Genome-wide association study of primary sclerosing cholangitis identifies new risk loci and quantifies the genetic relationship with inflammatory bowel disease. Nat Genet 49(2):269–273

57. Engel B, Laschoschwitz A, Janik MK, Junge N, Baumann U, Milkiewicz P, Taubert R, Sebode M (2021) Genetic aspects of adult and pediatric autoimmune hepatitis: A concise review. Eur J Med Genet 64(6):104214

58. Chung BK, Hirschfeld GM (2017) Immunogenetics in primary sclerosing cholangitis. Curr Opin Gastroenterol 33(2):93–98

59. Schaid DJ, Chen W, Larson NB (2018) From genome-wide associations to candidate causal variants by statistical fine-mapping. Nat Rev Genet 19(8):491–504

60. Huang H, Fang M, Jostins L, Umicevic Mirkov M, Boucher G, Anderson CA, Andersen V, Cleyne I, Cortes A, Crins F, D’Amato M, Deffontaine V, Dmitrijev J, Docampe E, Elansary M, Farh KK, Franke A, Gori AS, Goyette P, Halfvarson J, Harritunians T, Knight J, Lawrence IC, Lees CW, Louis E, Mariman M, Meuwissen T, Mni M, Momozawa Y, Parkes M, Spain SL, Theatre E, Trynka G, Satsangi J, van Someren S, Vermeiren S, Xavier RJ, C International inflammatory bowel disease genetics, Weersma RK, Duerr RH, Mathew CG, Rioux JD, McGovern DPB, Cho JH, Georges M, Daly MI, Barrett JC (2017) Fine-mapping inflammatory bowel disease loci to single-variant resolution. Nature 547(7662):173–178

61. Konig IR, Loley C, Erdmann J, Ziegler A (2014) How to include chromosome X in your genome-wide association study. Genet Epidemiol 38(2):97–103

62. Asselta R, Paraboschi EM, Gerussi A, Cordell HJ, Mells GF, Sandford RN, Jones DE, Nakamura M, Ueno K, Hitomi Y, Kawashima M, Nishida N, Tokunaga K, Nagasaki M, Tanaka A, Tang R, Li Z, Shi Y, Liu X, Xiong M, Hirschfeld G, Siminovich KA, USPBCB Canadian A, International PBCGSG, Consortium U-P, Japan PBCGC, Cardone M, Cardamone G, Duga S, Gershwin ME, Seldin MF, Invernizzi P X (2021) Chromosome contribution to the genetic architecture of primary biliary cholangitis. Gastroenterology 160(7):2483–2495 e26

63. So HC, Gui AH, Cherny SS, Sham PC (2011) Evaluating the heritability explained by known susceptibility variants: a survey of ten complex diseases. Genet Epidemiol 35(5):310–317

64. Genin E (2020) Missing heritability of complex diseases: case solved? Hum Genet 139(1):103–113

65. Hivert V, Sideronenko J, Rohart F, Goddard ME, Yang J, Wray NR, Yengo L, Visscher PM (2021) Estimation of non- additive genetic variance in human complex traits from a large sample of unrelated individuals. Am J Hum Genet 108(5):786–798

66. Wei WH, Hamani G, Haley CS (2014) Detecting epistasis in human complex traits. Nat Rev Genet 15(11):722–733

67. Hamani G, Theocaris D, Wei W, Haley C (2011) EpiGPU: exhaustive pairwise epistasis scans parallelized on consumer level graphics cards. Bioinformatics 27(11):1462–1465

68. Wierbrandt L, Kassens JC, Ellinghaus D (2021) SNPInt-GPU: Tool for Epistasis Testing with Multiple Methods and GPU Acceleration. Methods Mol Biol 221:17–35

69. Boberg KM, Chapman RW, Hirschfeld GM, Lobho AW, Manns MP, Schrumpf E, G (2011) International autoimmune hepatitis, overlap syndromes: the International Autoimmune Hepatitis Group (IAIHG) position statement on a controversial issue. J Hepatol 54(2):374–85

70. Erlich H, Valdes AM, Noble J, Carlson JA, Varney M, Con
cannon P, Mychaleckyj JC, Todd JA, Bonella P, Fear AL, Lavant E, Louey A, Moonsamy P, Type 1 Diabetes Genetics C (2008) HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1 diabetes genetics consortium families. Diabetes 57(4):1084–92

71. Morris DL, Taylor KE, Fernando MM, Nithiam J, Alarcon-Riquelme ME, Barcellos LF, Behrens TW, Cotsapas C, Gaffney PM, Graham RR, Pons-Estel BA, Gregersen PK, Harley JB, Hauser SL, Hom G, International MHC, Autoimmunity Genetics N, Langefeld CD, Noble JA, Rioux JD, Seldin MF, Systemic Lupus Erythematosus Genetics C, Criswell LA, Vys T (2012) Unraveling multiple MHC gene associations with systemic lupus erythematosus: model choice indicates a role for HLA alleles and non-HLA genes in Europeans. Am J Hum Genet 91(5):778–93

72. Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR (2012) Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics 28(19):2540–2542

73. Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88(1):76–82

74. BK Bulik-Sullivan, Loh PR, Finucane HK, Ripke S, Yang J, Psychizophrenia Working Group of the Psychiatric Genomics C, Patterson N, Daly MJ, Price AL, Neale BM (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47(3):291–5

75. Giambartolomei C, Vuckevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V (2014) Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet 10(5):e1004383
van Rheenen W, Peyrot WJ, Schork AJ, Lee SH, Wray NR (2019) Genetic correlations of polygenic disease traits: from theory to practice. Nat Genet 20(10):567–581

Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Chi SH, Natrajan P, Lander ES, Lubitz SA, Ellinor PT, Kathiresan S (2018) Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 50(9):1219–1224

de Vries AB, Janse M, Blokzijl H, Weersma RK (2015) Distinctive inflammatory bowel disease phenotype in primary sclerosing cholangitis. World J Gastroenterol 21(6):1956–1971

Relton CL, Gaunt T, Mcardle W, Ho K, Duggirala A, Shihab H, Woodward G, Lyttleton O, Evans DM, Reik W, Paul YL, Ficz G, Wipat A, Plenge RM, Danesh J, Runz H, Butterworth AS (2021) Single-cell atlas of hepatic T cells reveals expansion of liver-resident naïve-like CD4(+) T cells in primary sclerosing cholangitis. J Hepatol 75(2):414–423

Xiang B, Deng C, Qiu F, Li J, Li S, Zhang H, Lin X, Huang Y, Zhou Y, Su J, Lu M, Ma Y (2021) Single cell sequencing analysis identifies genetics-modulated ORMDL3(+) cholangiocytes having higher metabolic effects on primary biliary cholangitis. J Nanobiotechnol 19(1):406

Nelson MR, Tipney H, Painter J, Shen J, Nicoletti P, Shen Y, Floratos A, Sham PC, Li MJ, Wang J, Cardon LR, Whittaker JC, Sanseau P (2015) The support of human genetic evidence for approved drug indications. Nat Genet 47(8):856–860

Gunev E, Menche J, Vidal M, Barabasi AL (2016) Network-based in silico drug efficacy screening. Nat Commun 7:10331

Hedin CRH, Sonkoly E, Eberhardson M, Stahle M (2021) Inflammatory bowel disease and psoriasis: modernizing the multidisciplinary approach. J Intern Med 290(2):257–278

Hirschfield GM, Gershwin ME, Strauss R, Mayo MJ, Levy C, Zou B, Johanss N, Nane P, Dasgupta B, Li K, Selmi C, Marshall HU, Jones D, Lindor K, Group PPS (2016) Ustekinumab for patients with primary biliary cholangitis who have an inadequate response to ursodeoxycholic acid: a proof-of-concept study. Hepatology 64(1):189–199

Prive F, Arbél J, Vilhhjalmsson BJ (2020) LDpred2: better, faster, stronger, Bioinformatics

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