University of Groningen

Novel HLA-DRB1 alleles contribute risk for disease susceptibility in primary biliary cholangitis

Li, Yanni; Liu, Xin; Wang, Yi; Zhou, Yi; Hu, Shixian; Yang, Hui; Zhong, Weilong; Zhao, Jingwen; Wang, Xiaoyi; Chu, Hongyu

Published in:
 Digestive and Liver Disease

DOI:
[10.1016/j.dld.2021.04.010](https://doi.org/10.1016/j.dld.2021.04.010)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
 Publisher's PDF, also known as Version of record

Publication date:
 2022

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Li, Y., Liu, X., Wang, Y., Zhou, Y., Hu, S., Yang, H., Zhong, W., Zhao, J., Wang, X., Chu, H., Zheng, Y., Zhang, J., Zhou, L., & Wang, B. (2022). Novel HLA-DRB1 alleles contribute risk for disease susceptibility in primary biliary cholangitis. *Digestive and Liver Disease*, 54(2), 228-236.
<https://doi.org/10.1016/j.dld.2021.04.010>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Liver, Pancreas and Biliary Tract

Novel HLA-DRB1 alleles contribute risk for disease susceptibility in primary biliary cholangitis

Yanni Li^{a,b,1}, Xin Liu^{a,1}, Yi Wang^{a,c,1}, Yi Zhou^a, Shixian Hu^b, Hui Yang^a, Weilong Zhong^a, Jingwen Zhao^a, Xiaoyi Wang^a, Hongyu Chu^a, Yanping Zheng^a, Jie Zhang^a, Lu Zhou^{a,d,1,*}, Bangmao Wang^{a,1,**}

^a Department of Gastroenterology and Hepatology, Tianjin Medical University General Hospital, Tianjin, China

^b Department of Gastroenterology and Hepatology, University of Groningen and University Medical Center Groningen, Groningen, the Netherlands

^c Department of General Surgery, YouAn Hospital, Capital Medical University, Beijing, China

^d Department of Gastroenterology and Hepatology, People's Hospital of Hetian District, Xinjiang Uygur Autonomous Region, China

ARTICLE INFO

Article history:

Received 7 January 2021

Revised 7 April 2021

Accepted 8 April 2021

Available online 17 May 2021

Keywords:

Primary biliary cholangitis

HLA-DRB1 alleles

Amino acid residues

Pedigree

ABSTRACT

Background: Primary biliary cholangitis (PBC) is a complex disease with high heritability. We investigated the association between human leukocyte antigen (HLA)-DRB1 alleles and PBC in families and sporadic cases to evaluate the genetic components of the disease.

Methods: We performed whole exome sequencing in three PBC families. We genotyped HLA-DRB1 and calculated the association between HLA-DRB1 alleles and the encoding amino acid sequences with the clinical features.

Results: Ten variants harboured the HLA-DRB1 gene associated with PBC. DRB1*07:01, 14:01 and 14:05 were highly increased in PBC. Ten coding region polymorphisms were associated with PBC that encode the amino acid variants of HLA-DR β 54, β 59 and β 66 located in the peptide-binding site of the MHC molecule. Glutamine at position 54 was confirmed as a risk amino acid, verifying the results of familial aggregation analysis of PBC families.

Discussion: Familial aggregation analysis indicated that HLA-DRB1 is a candidate gene for the risk of disease course. Considering that amino acid variations are critical to peptide-binding properties, they underlie the major component of MHC association with PBC.

© 2021 Editrice Gastroenterologica Italiana S.r.l. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Primary biliary cholangitis (PBC) is a chronic auto-immune cholestatic liver disease that progresses to liver dysfunction and potentially hepatic cirrhosis [1,2]. It is characterized by auto-immune-related destruction of small-to-medium intrahepatic bile ducts [3]. However, the exact pathogenesis of this disease has not been fully determined. Previous studies on familial clustering and monozygotic twins confirmed that genetic factors are critical in disease susceptibility to PBC [4–6]. Studies conducted in Europe and Japan found that the first-degree relatives (FDRs) of patients with PBC are at increased risk of this disease compared with con-

trols [7,8]. Moreover, the prevalence of anti-mitochondrial antibody increases among the FDRs of patients with PBC [9].

Whole exome sequencing (WES) refers to the targeted sequencing of a subset of the human genome that codes for proteins [10]. WES is adopted in exploring rare alleles that explain the heritability of complex diseases. Thousands of loci are associated with PBC [11–13]. Human leukocyte antigen (HLA) plays an important role in disease and immune defense and encodes the human major histocompatibility complex (MHC), which is expressed on the surface of antigen-presenting cells; its variants may play a key role in the process of antigen peptide preference [14]. HLA is closely related to extrahepatic autoimmune diseases (EHAIDs), including autoimmune thyroid diseases, Sjogren's syndrome and rheumatoid arthritis [15–19]. Although the molecular structure of HLA DR from MHC class II may contribute to susceptibility to PBC [20], HLA genotypes and the role of amino acid alignment have not been precisely analyzed.

Genetics play an important role in PBC, and HLA is the genetic system with a strong disease association. However, the HLA sys-

* Corresponding author at: Department of Gastroenterology and Hepatology, Tianjin Medical University General Hospital, Tianjin, China.

** Corresponding author.

E-mail addresses: lzhou01@tmu.edu.cn (L. Zhou), gi.tmu@yeah.net (B. Wang).

¹ These authors contributed equally to this study.

tem is complex and requires specialized testing and careful interpretation to support or exclude a diagnosis of PBC. In the present study, we identified the candidate mutations in three families via WES and tested the association between HLA-DRB1 and the clinical courses of PBC. We investigated the MHC-independent determinants of PBC by further checking the distribution of DRB1 amino acid residues in disease subjects.

2. Patients and methods

2.1. Study subjects

2.1.1. Hospital-based recruitment of subjects diagnosed with PBC

The patients recruited have met the PBC criteria recommended by the American Association for the study of liver diseases and the European Association for the Study of the Liver [21,22]. All patients were negative for hepatitis B surface antigen and antibodies to hepatitis B core antigen, hepatitis C virus and human immunodeficiency virus. All patients with autoimmune hepatitis (AIH) or primary sclerosing cholangitis (PSC) were excluded. All subjects provided written informed consent, which was approved by the institutional ethics committee.

2.2. Collection of clinical data

Data were collected from the subjects at the initial diagnosis, including baseline demographic characteristics, family history and serum laboratory examination. Family history and the incidence of EHAIDs, including autoimmune thyroid disease (AITD), Sjogren's syndrome (SS), rheumatoid arthritis (RA), connective tissue disease (CTD) and psoriasis, were collected through telephone follow-up [23,24]. The serological protocol of antibodies included testing for anti-nuclear antibodies, anti-mitochondrial and anti-mitochondrial M2 antibodies (AMA/AMA-M2), smooth muscle actin antibodies, anti-gp210 antibody and anti-sp100 antibody reactivity.

3. Family study with WES

3.1. Patients

Three families with at least one of FDRs diagnosed as PBC were recruited. The family history of two generations was collected, and the pedigrees are shown in Fig. 1A. Nine members, including seven patients with PBC (II-2 and II-3 of Family 1; II-2 and II-3 of Family 2; II-2, II-3 and III-1 of Family 3) and two healthy controls (II-1 and III-2 of Family 3) donated DNA samples. Clinical information and laboratory testing were obtained at the time of enrolment. Histological results were achieved in patients diagnosed with PBC.

3.2. WES and variant calling

Genomic DNA was extracted from the participants' peripheral blood by using the QIAamp DNA blood mini kit (QIAGEN Sciences, Maryland) after the receipt of informed consent. WES libraries were prepared using Agilent SureSelect Human All ExonV6 (Agilent Technologies, California) and sequenced on Illumina HiSeq platform (Illumina Inc. California) as paired-end reads to 100 bp/read. Variant calling and filtering based on GCC pipeline were carried out (https://molgenis.gitbooks.io/ngs_dna/ngs-run.html). Variants with a coverage of $< 5 \times$ for single nucleotide variants (SNVs) and $10 \times$ for insertion/deletion (Indels) were excluded from further analysis. Phred quality scores were specified as Q15 for SNVs and Q20 for Indels. Rare mutations that passed the quality score were filtered with gnomAD [25] and ExAC [26] and annotated using the 1000 Genomes Project [27].

3.3. Variants annotation

Functional annotation and effect prediction of the variants were performed using SnpEff and ANNOVAR to integrate the information from public databases, including UCSC hg19 and dbSNP [27]. Downstream/upstream gene variants, 3'/5' prime UTR variants, intron variants, intergenic variants, synonymous variants and non-coding variants were removed for further analysis. Possible deleterious variants were filtered with the deletion of variants with SIFT [28] (score > 0.05 , tolerated), Polyphen-2 [29] (score ≤ 0.446 , benign) and CADD [30] (score < 15 , benign). Considering that rare genetic variants often change the protein function and can be pathogenic, possible "pathogenic" variants were selected using GAVIN [31]. We checked the population frequency of the selected variants by using the 1000 Genomes Project and Novo-Zhonghua genome database (<https://en.novogene.com/>) and narrowed down candidate variants by excluding variants with allele frequency of more than 0.01 in either database.

4. PCR-SBT HLA genotyping

We recruited 52 patients with PBC for HLA-DRB1 genotyping. Complete, detailed phenotypic data, including baseline characteristics, family history and extrahepatic manifestations were obtained from all patients. The genomic DNA of patients was isolated from 5 ml of peripheral blood by using the DNA extraction kit (Qiagen). After amplifying the second and third exon of HLA-DRB1, the genotyping of HLA-DRB1 was conducted from polymerase chain reaction with sequence-based typing (PCR-SBT) by ABI 3500 genetic analyser. The HLA allelic genotypes of 126 healthy subjects with matched age and sex were included as controls (obtained from Chinese Marrow Donor Program). The peptide sequences of HLA-DRB1 alleles were aligned in the IPD-IMGT/HLA database release 3.41.2 [32]. According to the comparison of the detected genotype from IPD-IMGT/HLA, the amino acid sequence of the HLA-DR β chain was obtained. The frequency and amino acid type at each position of the DR β chain were extracted, and the difference between PBC and control cohort were calculated. The 3D structure and molecular surface electrostatic potentials of HLA-DR molecules were generated using PyMOL (version 2.3.2 Schrödinger, LLC).

5. Statistical analysis

The HLA-DRB1 genotyping frequencies of patients and controls were compared using the Pearson's chi-squared test or Fisher's exact test when appropriate. Clinical continuous data conforming to normal distribution are presented as mean and standard deviation, otherwise as median, quartile 25 and quartile 75%. Categorical variables are expressed in percentages. Depending on the distribution of data, parametric (t -test) and non-parametric tests (Mann-Whitney U test) were used to determine the statistical significance of the differences. Results with a p -value less than 0.05 were considered statistically significant. Significant associations between HLA alleles and identified autoimmune antibodies were tested with a logistic regression model. The association strength was estimated by calculating the odds ratio (OR) and 95% confidence interval (CI).

6. Results

6.1. Mutational profiling of families with PBC and HLA-DRB1 deficiency

6.1.1. Characterization of a family with PBC by WES

The descriptive characteristics of the sequenced samples from three pedigrees are shown in Supplementary Table 1. The median age of seven patients with PBC was 57 (IQR, 54–73)

Fig. 1. Genetic and molecular features of two-generation PBC families. (A) Pedigree structures of three PBC families. Affected subjects are shown in dark symbols. The individuals carrying the HLA-DRB1 variants have been marked in the text box of black flame. * indicates individuals selected for whole exome sequencing analysis, indicates the patient as the proband, indicates the patient died of liver disease. (B) Schematic representation of the structure of HLA, Class II HLA antigens includes HLA-DR, -DQ, and -DP loci, are on antigen-presenting cells. Whole exome sequence analysis identified 10 missense mutations and 2 stop-gain variants, of which cDNA mutations and amino acid change have been list in the text box of black flame. The color bars show the exome sequence and 3'-UTR on HLA-DR molecular from chromosome 6.

Table 1
Identified HLA-DRB1 variants among three PBC families.

Chr	Pos	ID	Exon	Base mutation	Type	AA Change	Frequency
6	32,549,357	rs200809587	exon3	C629T	missense	T210I	0.1416
6	32,549,525	rs16822698	exon3	G461C	missense	G154A	0.2235
6	32,549,531	rs112796209	exon3	A455G	missense	Y152C	0.2238
6	32,549,582	rs200516145	exon3	C404A	missense	T135N	–
6	32,551,935	rs11554463	exon2	C321A	stopgain	Y107X	0.1952
6	32,552,039	rs150747106	exon2	G217A	missense	V73M	0.2728
6	32,552,048	rs56158521	exon2	G208A	missense	D70N	0.2694
6	32,552,095	rs17885382	exon2	G161A	missense	R54Q	–
6	32,548,632	rs1136881	exon4	A654T	missense	R218S	0.2578
6	32,548,026	rs9269744	exon5	C785G	missense	T262R	0.3868

Abbreviations: X, terminator; G, glycine; R, arginine; V, valine; M, methionine; L, leucine; T, threonine; I, isoleucine; Q, glutamine; A, alanine; Y, tyrosine; C, cysteine; N, asparagine; D, aspartic acid; S, serine.

years. The averages of alkaline phosphatase (ALKP), aspartate transaminase (AST), and average immunoglobulin G levels were 244.3 U/L, 47.3 U/L, and 17.9 g/L, respectively. The titers of ANA and AMA/AMA-M2 antibodies increased in both patients with PBC. Abdominal lymph node enlargement occurred in five patients with PBC and EHAIIDs, including ATD, SS, and RA, and were concurrent in two patients with PBC. The available pathological results of liver biopsy in the two patients with PBC suggested injury to small and medium bile ducts.

6.1.2. WES and selection of candidate variants

To identify potential disease-causing variants, we adopted a filtering and follow-up scheme, in which PBC cases in both family members were assumed to carry the same causal variant (Fig. 2). WES generated > 30 million sequence reads for each sample, and over 97% of the paired reads were uniquely mapped to the reference human genome UCSC GRCh37. Over 75% of the uniquely mapped reads were on the target regions, while over 96% of the target regions were covered with at least 10× depth. Approximately 10,000 nonsynonymous exonic SNVs and 200 indels were detected in each family member after filtering. The number of variants from the sequenced individuals is shown in Supplementary Table 2 and Fig. 1A. Rare mutations included variants in the affected patients but were absent in familial controls. The number of rare mutations of SNPs and Indels and the candidate genes selected within and among the three families are provided in Supplementary Table 3 and Fig. 2. Among the mapped genes identified in the three families, HLA-DRB1 was found in the reported GWAS of autoimmune liver disease [11,33,34].

From the three families with PBC, 10 variants harbored HLA-DRB1 (Table 1). Most of the identified variants encode amino acid mutations in exomes 2 and 3. These variants are involved in specifying peptide-binding amino acids, and this process may affect the antigen-presentation functions of MHC-II (Fig. 1B). The changes in the amino acids of each exonic variant are presented in Table 1. rs16822698 (G>C, p. G154A) and rs112796209 (A>G, p. Y152C) in exome 3 are located at the junction of the α/β chains of HLA-II molecules. rs11554463 (C>A, p. Y107X) is located in the antigenic peptide binding of MHC-II. rs17885382 (G>A, p. R54Q), rs1136881 (A>T, p. R218S) and rs9269744 (C>G, p. T262R) are encoded by DRB1×07:01, 08:01 and 03:01, respectively.

6.2. Risk genotyping of HLA-DRB1 in patients with PBC

To determine the effects of HLA risk alleles and clinical features of PBC, we analyzed the association of genotypes in an independent cohort of patients with PBC by PCR-SBT HLA genotyping. The characteristics of the selected patients at initial diagnosis are listed in Supplementary Table 4. The distribution of HLA-DRB1 allele carrier is shown in Fig. 3A. The frequency of HLA-DRB1 al-

Table 2

Association of HLA-DRB1 allele frequencies observed in patients with PBC and healthy controls.

DRB1	PBC 2n=104(%)	Control 2n=252(%)	OR	95%CI	P-value
01:01	5(4.81)	5(1.98)	2.495	0.707–8.807	0.164
03:01	5(4.81)	13(5.16)	0.929	0.322–2.674	0.891
04:01	1(0.96)	3(1.19)	0.806	0.083–7.837	1.000
04:05	3(2.88)	12(4.76)	0.594	0.164–2.150	0.567
04:06	1(0.96)	6(2.38)	0.398	0.047–3.348	0.647
04:10	1(0.96)	1(0.4)	2.437	0.151–39.331	0.500
07:01	19(18.27)	24(9.52)	2.124	1.107–4.073	0.021*
08:01	1(0.96)	0(0)	–	–	0.292
08:03	9(8.65)	16(6.35)	1.397	0.597–3.272	0.439
09:01	12(11.54)	37(14.68)	0.758	0.378–1.519	0.434
10:01	2(1.92)	4(1.59)	1.216	0.219–6.741	1.000
11:01	4(3.85)	14(5.56)	0.680	0.218–2.117	0.503
12:01	1(0.96)	7(2.78)	0.340	0.041–2.797	0.446
12:02	7(6.73)	22(8.73)	0.754	0.312–1.824	0.531
13:01	1(0.96)	4(1.59)	0.602	0.066–5.451	1.000
14:01	3(2.88)	0(0)	–	–	0.024*
14:05	8(7.69)	6(2.38)	3.417	1.155–10.106	0.032*
14:07	1(0.96)	1(0.4)	2.437	0.151–39.331	0.500
14:14	1(0.96)	0(0)	–	–	0.292
14:54	2(1.92)	6(2.38)	0.804	0.160–4.050	1.000
15:01	13(12.5)	29(11.51)	1.123	0.558–2.259	0.744
15:02	2(1.92)	8(3.17)	0.598	0.125–2.865	0.730
16:02	2(1.92)	8(3.17)	0.598	0.125–2.865	0.730

Abbreviations: OR, odds ratio; CI, confidence interval; * $p<0.05$.

lele in patients with PBC was compared with that of healthy subjects through logistic analysis (Table 2). DRB1×07:01 increased the risk to PBC compared with healthy controls (OR 2.124, 95%CI 1.107–4.073, $p=0.021$). DRB1×14:05 increased the risk of PBC (OR 3.417, 95%CI 1.155–10.106, $p=0.032$). Despite the low frequency in controls, DRB1×14:01 still poses substantially risk to patients with PBC ($p=0.024$).

7. Association between HLA-DRB1 and clinical features of PBC

The clinical characteristics of patients with PBC with or without DRB1×07:01/DRB1×14:01/ DRB1×14:05 are summarized in Supplementary Table 5. The median immunoglobulin G and M levels were significantly high in patients with PBC with DRB1×14:01 (OR 1.004, 95%CI 1.000–1.007, $p=0.006$; and OR 1.003, 95%CI 1.000–1.006, $p=0.007$, respectively). Similar to patients with DRB1×14:05, a high level of complement C4 was associated with patients with PBC with DRB1×07:01, DRB1×14:01 or DRB1×14:05 alleles (OR 1.150, 95%CI 1.014–1.304, $p=0.018$). Notably, AMA/AMA-M2 positivity increased in patients with PBC with identified DRB1 alleles (OR 9.947, 95%CI 1.150–86.012, $p=0.029$), whereas the autoimmune antibody positivity values for anti-glycoprotein-210 and anti-sp100 antibodies was equally distributed.

Family	Variants	Number	Genes	Reported in previous GWAS
F1& F2	SNP	57	<i>PRAMEF1, FAM131C, SRGAP2B, PPIAL4G, PDE4DIP, SEC22B, NOTCH2NL, NBPFL14</i>	<i>HLA-DRB1, MST1</i>
	Indel	26	<i>FAM171B, KIF1A, TRAK1, MYH15, HTT</i>	-
F1& F3	SNP	18	<i>OBSCN, HLA-DRB1, OTOA</i>	<i>HLA-DRB1</i>
F2& F3	SNP	18	<i>OBSCN, HLA-DRB1, OTOA</i>	<i>HLA-DRB1</i>
F1&F2&F3	SNP	18	<i>OBSCN, HLA-DRB1, OTOA</i>	<i>HLA-DRB1</i>

Fig. 2. Filtering and follow-up scheme used in family analysis. After variant filtering, 1285 candidate variants were selected for further analysis. 18 variants were co-segregated in affected patients and were absent in familial controls among three PBC family, which were mapped in three genes. Out of them, HLA-DRB1 has been identified in genome-wide analysis study of autoimmune liver disease.

The association between HLA-DRB1 genotypes and autoimmune antibodies is shown in Table 3. HLA-DRB1*01 provides protective effects against the prevalence of anti-mitochondrial antibody (OR 0.06, 95%CI 0.01–0.60, $p=0.005$), whereas HLA-DRB1*03 and HLA-DRB1*16 are associated with the incidence of anti-sp100 antibodies (OR 11.14, 95%CI 1.58–78.75, $p=0.024$; and OR 6.53, 95%CI 1.37–31.23, $p=0.027$, respectively).

8. Distribution of DRB1 amino acid residues

The amino acid sequences of HLA-DRB1 were determined for each disease subject in accordance with codon rules. Ten risk amino acids associated with PBC were detected (Fig. 3B; Supplementary Table 6). In comparison with healthy controls, the prevalence of glycine at $\beta 40$ ($p=0.021$), tyrosine at $\beta 42$, $\beta 76$, and $\beta 141$

Fig. 3. HLA-DRB1 genotypes in PBC patients and healthy subjects. (A) The frequency of unique HLA-DRB1 alleles in PBC patients. (B) Associated polymorphic amino acid positions of HLA-DRB1 alleles within PBC cohort. Amino acid position 66 had the strongest association with PBC patients compared to healthy subjects, and positions 40, 42, 43, 45, 54, 59, 76, 103, and 141 were associated with PBC to a lesser extent. The y-axis shows the P value of each association tested and the x-axis represents the amino acid positions of the HLA-DRB1 protein. The dashed horizontal red line identifies a significance threshold of $p < 0.05$. (C) 3D structure for HLA-DR1 with PBC associated amino acid residues shown in yellow. The green chain shows a HLA-DR α chain, the blue chain shows HLA-DR β chain (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Table 3
Association between HLA-DRB1 genotyping and autoantibodies in patients with PBC

DRB1	AMA/AMA-M2				Gp210				Sp100			
	Pos 2n=80 (%)	Neg 2n=20 (%)	OR (95%CI)	P	Pos 2n=26 (%)	Neg 2n=38 (%)	OR (95%CI)	P	Pos 2n=10 (%)	Neg 2n=54 (%)	OR (95%CI)	P
1	1(0.01)	4(0.20)	0.06(0.01–0.60)	0.005*	3(0.12)	2(0.05)	2.35(0.36–15.14)	0.389	1(0.10)	4(0.07)	1.39(0.14–13.90)	1.000
3	3(0.04)	2(0.10)	0.35(0.06–2.26)	0.261	3(0.12)	2(0.05)	2.35(0.36–15.14)	0.389	3(0.30)	2(0.04)	11.14(1.58–78.75)	0.024*
4	4(0.05)	2(0.10)	0.47(0.08–2.79)	0.597	2(0.08)	3(0.08)	0.97(0.15–6.27)	1.000	0	5(0.09)	-	1.000
7	16(0.20)	2(0.10)	2.25(0.47–10.71)	0.515	4(0.15)	5(0.13)	1.20(0.29–4.97)	1.000	0	9(0.17)	-	0.333
8	8(0.10)	1(0.05)	2.11(0.25–17.93)	0.682	1(0.04)	4(0.11)	0.34(0.04–3.23)	0.640	0	5(0.09)	-	1.000
9	10(0.13)	2(0.10)	1.29(0.26–6.39)	0.101	3(0.12)	5(0.13)	0.86(0.19–3.96)	1.000	2(0.20)	6(0.11)	2.00(0.34–11.70)	0.405
11	2(0.03)	0	-	1.000	0	0	-	-	0	0	-	-
12	2(0.03)	2(0.10)	0.23(0.03–1.75)	0.178	2(0.08)	0	-	0.161	0	2(0.04)	-	1.000
13	4(0.05)	3(0.15)	0.30(0.06–1.46)	0.140	1(0.04)	2(0.05)	0.72(0.06–8.38)	1.000	0	4(0.07)	-	1.000
14	0	1(0.05)	-	0.200	1(0.04)	0	-	0.406	0	1(0.02)	-	1.000
15	15(0.19)	0	-	0.037*	3(0.12)	7(0.18)	0.58(0.14–2.48)	0.510	0	9(0.17)	-	0.333
16	11(0.14)	3(0.15)	0.90(0.23–3.60)	1.000	3(0.12)	6(0.16)	0.70(0.16–3.07)	0.728	4(0.40)	5(0.09)	6.53(1.37–31.23)	0.027*

Abbreviations: AMA/AMA-M2, antimitochondrial antibody and antimitochondrial M2 antibody; gp210, anti-glycoprotein-210 antibodies; sp100, anti-sp100 antibodies; OR, odds ratio; CI, confidence interval; *with p-value<0.05.

($p=0.021$, 0.037 , 0.024 , respectively), lysine at $\beta 43$ ($p=0.021$), glutamine at $\beta 45$, $\beta 54$, and $\beta 103$ ($p=0.032$, 0.021 , 0.021 respectively), leucine at $\beta 59$ ($p=0.021$) and phenylalanine at $\beta 66$ ($p=0.0025$) significantly increased in patients with PBC. Glutamine at position 54 was confirmed as a risk amino acid, thus verifying the results of WES in the families with PBC.

The potential functional effects of critical PBC-susceptible amino acid residues were explored by calculating the electrostatic potential in 3D space surrounding the HLA-DRB1 structure (Fig. 3C). We focused on the identified amino acid residues in the regression analysis. Three amino acid variants of HLA-DR $\beta 54$, $\beta 59$ and $\beta 66$ are located in the peptide-binding site of the HLA molecule. HLA-DR $\beta 66$ forms part of the interface of the $\alpha\beta$ dimer and could interfere with TCR contact, implying that the antigen-binding affinities of the protein have different potential affects.

9. Discussion

Previous genetic analysis demonstrated that HLA alleles had a strong link to PBC susceptibility [35,36]. We identified that HLA-DRB1 was a candidate gene for the risk of disease course via familial aggregation analysis in our study. DRB1 $\times 07:01$, DRB1 $\times 14:01$ and DRB1 $\times 14:05$ were indicated an increased risk of PBC through HLA-DRB1 genotyping within sporadic PBC cohort. Considering that amino acid variations were critical to peptide-binding properties, we identified coding region polymorphisms that encode the amino acid variants of HLA-DR $\beta 54$, $\beta 59$ and $\beta 66$, located in the peptide-binding site of the MHC molecule, were associated with PBC.

PBC has a unique genetic background compared with other autoimmune diseases because of its strong familial clustering. Masatoshi et al. reported the familial clustering phenomenon in 5.8% cases of patients with PBC [8]. Studies conducted in Europe and North America reported that the prevalence of PBC among family members ranges from 3.8% to 9.0% [37]. The analysis of the genetic causes of highly similar familial structures can help in identifying co-segregations in affected patients and understanding the pathogenic mechanisms of PBC. In the present study, we conducted whole exome sequencing of peripheral blood genomic DNA of families with PBC and performed downstream bioinformatics filtering to identify biologically plausible genetic variants. We compared the genes reportedly associated with PBC and identified HLA-DRB1 as the best candidate in our familial analysis. Ten associated variants in the HLA class II gene DRB1 had been previously implicated in PBC risk [38]. (Table 1; Fig. 1B) Most of the identified variants encode amino acid mutations involved in specifying peptide-binding amino acids.

HLA genes encode cell-surface molecules through peptide presentation and could mediate key immunological events [39]. Genome-wide association studies (GWAS) of antinuclear autoantibody sub-phenotypes in PBC identified HLA-DRB1 and -DPB1 were the major determinants for speckled 100 kDa (sp100) production [40]. We summarized the studies on HLA-DRB1 associations in PBC courses in Supplementary Table 7. HLA-DRB1 $\times 08:01$ confers an increased risk to PBC among Caucasian populations, whereas DRB1 $\times 11$ and DRB1 $\times 13$ have a protective role in PBC [36,41]. Nakamura et al. reported a strong association between HLA-DRB1 $\times 04:05$ and *08:03 with PBC only in a subset of patients positive for anti-sp100 and anti-centromere antibodies [42]. By comparison, Genaro et al. found no significant association between HLA-DRB1 and PBC among Mexican patients [43]. Hence, ethnic differences have a substantial contribution to the determination of HLA-DRB1 alleles.

In our case, we used high-throughput technique with HLA-DRB1 genotyping in deep sequence and found that HLA-DRB1 $\times 07:01$ was strongly associated with PBC manifestations,

which distinguishes patients with overlap syndromes from those with autoimmune hepatitis [44]. Associations between PBC and DRB1×14 have not been demonstrated. Notably, patients concurrent with DRB1×07:01, 14:01 or 14:05 are associated with anti-mitochondrial antibody and M2 subtype positivity, which are favorable serum markers with high specificities and relative moderate sensitivities for PBC diagnosis [45].

A previous study determined the amino acid sequences encoded by the HLA-DRB1 allele in relation to disease susceptibility in a Japanese population [46]. In patients with PBC, the incidence of glycine 13, tyrosine 16 and leucine 74 encoded by DRB1×08:03 is high, whereas that of serine 13, histidine 16 and phenylalanine 47 encoded by DRB1×11 and DRB1×13 is low. The strongest amino acid association at position 74 in the antigen-binding pocket of DRB1 was observed in Caucasian patients with PBC [47]. In the present study, we found 10 amino acid residues associated with the risk of PBC, and these residues were partly gathered in a region of conserved amino acids from positions 40 to 70 (Fig. 3B). Through the three-dimensional modeling of protein structures, we also demonstrated that the HLA-DR molecules located in the antigenic peptide binding of MHC-II and the HLA-DRB1 amino acid substitutions conferring PBC susceptibility were correlated.

The ability to calculate DRB1 genotype among patients with PBC and their clinical characteristics is limited, which could lead to some correlations might be neglected. Although other sequence variations that affect HLA have remarkable contributions, our data indicated that a limited set of classical DRB1 alleles are sufficient to explain HLA associations with PBC. Furthermore, it's relatively hard to collect enough power East Asian PBC patients considering the prevalence of PBC is in 39–40.2 per 100,000 people. The current data provided cogent information for understanding HLA associations in patients with PBC in East Asian populations, the analysis with an enlarged dataset from the same ethnic group still need to be conducted for further investigation.

In conclusion, HLA-DRB1 alleles contribute to the risk of disease susceptibility and autoimmune antibody productions in patients with PBC. The results of our analyses in the familial clustering PBC by identifying HLA-DRB1 genotypes and amino acid residues showed functional relevance to the MHC region, highlighting new avenues for future functional investigation.

CRedit authorship contribution statement

Yanni Li: Formal analysis, Data curation, Writing – original draft. **Xin Liu:** Formal analysis, Data curation, Writing – original draft. **Yi Wang:** Formal analysis, Data curation, Writing – original draft. **Yi Zhou:** Formal analysis, Writing – original draft. **Shixian Hu:** Formal analysis, Writing – original draft. **Hui Yang:** Formal analysis, Writing – original draft. **Weilong Zhong:** Formal analysis, Writing – original draft. **Jingwen Zhao:** Formal analysis, Writing – original draft. **Xiaoyi Wang:** Formal analysis, Writing – original draft. **Hongyu Chu:** Formal analysis, Writing – original draft. **Yanping Zheng:** Formal analysis, Writing – original draft. **Jie Zhang:** Formal analysis, Writing – original draft. **Lu Zhou:** Writing – review & editing, Formal analysis, Writing – original draft. **Bangmao Wang:** Writing – review & editing, Formal analysis, Writing – original draft.

Acknowledgments

We would like to thank the study nurses, technicians, laboratory staff and clinicians for the collaboration. We would like to acknowledge Prof. Weidong Li for the scientific guidance. We would like to acknowledge Ying Ran, Liping Guo, Hongxia Zhang and Simin Zhou for their statistical guidance. Yanni Li was supported by the [China Scholarship Council](#) (No. 201706940045). The work

was supported by grants from the [National Natural Science Foundation of China](#) (No. 81860109, No. 81470834) and the Project of Science and Technology Supporting Xinjiang (2018E02100).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.dld.2021.04.010](https://doi.org/10.1016/j.dld.2021.04.010).

References

- [1] Yokoda RT, Carey EJ. Primary biliary cholangitis and primary sclerosing cholangitis. *Am J Gastroenterol* 2019 Published online. doi:[10.14309/ajg.0000000000000268](https://doi.org/10.14309/ajg.0000000000000268).
- [2] Rodrigues PM, Perugorria MJ, Santos-Laso A, Bujanda L, Beuers U, Banalles JM. Primary biliary cholangitis: a tale of epigenetically-induced secretory failure? *J Hepatol* 2018 Published online. doi:[10.1016/j.jhep.2018.08.020](https://doi.org/10.1016/j.jhep.2018.08.020).
- [3] Poupon R, Chazouillères O, Balkau B, Poupon RE. Clinical and biochemical expression of the histopathological lesions of primary biliary cirrhosis. *J Hepatol* 1999 Published online. doi:[10.1016/S0168-8278\(99\)80098-1](https://doi.org/10.1016/S0168-8278(99)80098-1).
- [4] Selmi C, Mayo MJ, Bach N, et al. Primary biliary cirrhosis in monozygotic and dizygotic twins: genetics, epigenetics, and environment. *Gastroenterology* 2004 Published online. doi:[10.1053/j.gastro.2004.05.005](https://doi.org/10.1053/j.gastro.2004.05.005).
- [5] Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. *J Hepatol* 2012 Published online. doi:[10.1016/j.jhep.2011.10.025](https://doi.org/10.1016/j.jhep.2011.10.025).
- [6] Örnolfsson KT, Olafsson S, Bergmann OM, Gershwin ME, Björnsson ES. Using the Icelandic genealogical database to define the familial risk of primary biliary cholangitis. *Hepatology* 2018 Published online. doi:[10.1002/hep.29675](https://doi.org/10.1002/hep.29675).
- [7] Floreani A, Naccarato R, Chiaramonte M. Prevalence of familial disease in primary biliary cirrhosis in Italy. *J Hepatol* 1997;26(3):737–8. doi:[10.1016/S0168-8278\(97\)80444-8](https://doi.org/10.1016/S0168-8278(97)80444-8).
- [8] Yanagisawa M, Takagi H, Takahashi H, et al. Familial clustering and genetic background of primary biliary cirrhosis in Japan. *Dig Dis Sci* 2010 Published online. doi:[10.1007/s10620-009-1057-0](https://doi.org/10.1007/s10620-009-1057-0).
- [9] Lazaridis KN, Juran BD, Boe GM, et al. Increased prevalence of anti-mitochondrial antibodies in first-degree relatives of patients with primary biliary cirrhosis. *Hepatology* 2007 Published online. doi:[10.1002/hep.21749](https://doi.org/10.1002/hep.21749).
- [10] Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for mendelian disease gene discovery. *Nat Rev Genet* 2011 Published online. doi:[10.1038/nrg3031](https://doi.org/10.1038/nrg3031).
- [11] Wang C, Zheng X, Tang R, et al. Fine mapping of the MHC region identifies major independent variants associated with han chinese primary biliary cholangitis. *J Autoimmun* 2020 Published online. doi:[10.1016/j.jaut.2019.102372](https://doi.org/10.1016/j.jaut.2019.102372).
- [12] Nakamura M, Nishida N, Kawashima M, et al. Genome-wide association study identifies TNFSF15 and POU2AF1 as susceptibility loci for primary biliary cirrhosis in the Japanese population. *Am J Hum Genet* 2012 Published online. doi:[10.1016/j.ajhg.2012.08.010](https://doi.org/10.1016/j.ajhg.2012.08.010).
- [13] Mells GF, Kaser A, Karlsen TH. Novel insights into autoimmune liver diseases provided by genome-wide association studies. *J Autoimmun* 2013 Published online. doi:[10.1016/j.jaut.2013.07.004](https://doi.org/10.1016/j.jaut.2013.07.004).
- [14] Dendrou CA, Petersen J, Rossjohn J, Fugger L. HLA variation and disease. *Nat Rev Immunol* 2018 Published online. doi:[10.1038/nri.2017.143](https://doi.org/10.1038/nri.2017.143).
- [15] Valdés-Corona LF, Hernández-Doño S, Rodríguez-Reyna TS, et al. Aspartic acid 70 in the HLA-DRB1 chain and shared epitope alleles partially explain the high prevalence of autoimmunity in mexicans. *J Transl Autoimmun* 2020 Published online. doi:[10.1016/j.jtauto.2020.100057](https://doi.org/10.1016/j.jtauto.2020.100057).
- [16] Ramgopal S, Rathika C, Padma MR, et al. Interaction of HLA-DRB1* alleles and CTLA4 (+ 49 AG) gene polymorphism in autoimmune thyroid disease. *Gene* 2018 Published online. doi:[10.1016/j.gene.2017.11.057](https://doi.org/10.1016/j.gene.2017.11.057).
- [17] Carapito R, Gottenberg JE, Kotova I, et al. A new MHC-linked susceptibility locus for primary Sjögren's syndrome: MICA. *Hum Mol Genet* 2017 Published online. doi:[10.1093/hmg/ddx135](https://doi.org/10.1093/hmg/ddx135).
- [18] Chun S, Bang S, Ha E, et al. Allele-specific quantification of HLA-DRB1 transcripts reveals imbalanced allelic expression that modifies the amino acid effects in HLA-DRB1. *Arthr Rheumatol* 2020 Published online. doi:[10.1002/art.41535](https://doi.org/10.1002/art.41535).
- [19] Fernandez CA, Smith C, Yang W, et al. HLA-DRB1×07:01 is associated with a higher risk of asparaginase allergies. *Blood* 2014 Published online. doi:[10.1182/blood-2014-03-563742](https://doi.org/10.1182/blood-2014-03-563742).
- [20] Chow IT, James EA, Gates TJ, et al. Differential binding of pyruvate dehydrogenase complex-E2 epitopes by DRB1×08:01 and DRB1×11:01 is predicted by their structural motifs and correlates with disease risk. *J Immunol* 2013 Published online. doi:[10.4049/jimmunol.1202445](https://doi.org/10.4049/jimmunol.1202445).
- [21] Lindor KD, Gershwin ME, Poupon R, Kaplan M, Bergasa NV, Heathcote EJ. Primary biliary cirrhosis. *Hepatology* 2009 Published online. doi:[10.1002/hep.22906](https://doi.org/10.1002/hep.22906).
- [22] European association for the study of the liver. EASL clinical practice guidelines: management of cholestatic liver diseases. *J Hepatol* 2009 Published online. doi:[10.1016/j.jhep.2009.04.009](https://doi.org/10.1016/j.jhep.2009.04.009).
- [23] Guo L, Zhou L, Zhang N, Deng B, Wang B. Extrahepatic autoimmune diseases in patients with autoimmune liver diseases: a phenomenon neglected by gastroenterologists. *Gastroenterol Res Pract* 2017 Published online. doi:[10.1155/2017/2376231](https://doi.org/10.1155/2017/2376231).

- [24] Efe C, Torgutalp M, Henriksson I, et al. Extrahepatic autoimmune diseases in primary biliary cholangitis: prevalence and significance for clinical presentation and disease outcome. *J Gastroenterol Hepatol* 2020 Published online. doi:[10.1111/jgh.15214](https://doi.org/10.1111/jgh.15214).
- [25] Koch L. Exploring human genomic diversity with gnomAD. *Nat Rev Genet* 2020 Published online. doi:[10.1038/s41576-020-0255-7](https://doi.org/10.1038/s41576-020-0255-7).
- [26] Karczewski KJ, Weisburd B, Thomas B, et al. The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Res* 2017 Published online. doi:[10.1093/nar/gkw971](https://doi.org/10.1093/nar/gkw971).
- [27] Consortium T 1000 GP. A global reference for human genetic variation the 1000 genomes project consortium. *Nature* 2015 Published online. doi:[10.1038/nature15393](https://doi.org/10.1038/nature15393).
- [28] Ng PC, Henikoff SSIFT. Predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003 Published online. doi:[10.1093/nar/gkg509](https://doi.org/10.1093/nar/gkg509).
- [29] Adzhubei I, Schmidt S, Peshkin L, et al. PolyPhen-2: prediction of functional effects of human nsSNPs. *Nat Methods* 2010 Published online.
- [30] Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* 2019 Published online. doi:[10.1093/nar/gky1016](https://doi.org/10.1093/nar/gky1016).
- [31] van der Velde KJ, de Boer EN, van Diemen CC, et al. GAVIN: gene-aware variant interpretation for medical sequencing. *Genome Biol* 2017 Published online. doi:[10.1186/s13059-016-1141-7](https://doi.org/10.1186/s13059-016-1141-7).
- [32] Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SGE. IPD-IMGT/HLA database. *Nucleic Acids Res* 2020 Published online. doi:[10.1093/nar/gkz950](https://doi.org/10.1093/nar/gkz950).
- [33] Hirschfield GM, Liu X, Han Y, et al. Variants at IRF5-TNPO3, 17q12-21 and MMEL1 are associated with primary biliary cirrhosis. *Nat Genet* 2010 Published online. doi:[10.1038/ng.631](https://doi.org/10.1038/ng.631).
- [34] Liu JZ, Almarri MA, Gaffney DJ, et al. Dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. *Nat Genet* 2012 Published online. doi:[10.1038/ng.2395](https://doi.org/10.1038/ng.2395).
- [35] Hirschfield GM, Liu X, Xu C, et al. Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. *N Engl J Med* 2009 Published online. doi:[10.1056/nejmoa0810440](https://doi.org/10.1056/nejmoa0810440).
- [36] Invernizzi P, Selmi C, Poli F, et al. Human leukocyte antigen polymorphisms in Italian primary biliary cirrhosis: a multicenter study of 664 patients and 1992 healthy controls. *Hepatology* 2008 Published online. doi:[10.1002/hep.22567](https://doi.org/10.1002/hep.22567).
- [37] Smyk D, Cholongitas E, Kriese S, Rigopoulou EI, Bogdanos DP. Primary biliary cirrhosis: family stories. *Autoimmun Dis* 2011 Published online. doi:[10.4061/2011/189585](https://doi.org/10.4061/2011/189585).
- [38] Invernizzi P, Selmi C, Mackay IR, Podda M, Gershwin ME. From bases to basis: linking genetics to causation in primary biliary cirrhosis. *Clin Gastroenterol Hepatol* 2005 Published online. doi:[10.1016/S1542-3565\(04\)00678-0](https://doi.org/10.1016/S1542-3565(04)00678-0).
- [39] Klein J, Sato A. The HLA system. Second of two parts. *N Engl J Med* 2000;343(11):782–6. doi:[10.1056/NEJM200009143431106](https://doi.org/10.1056/NEJM200009143431106).
- [40] Wang C, Zheng X, Jiang P, et al. Genome-wide association studies of specific antinuclear autoantibody subphenotypes in primary biliary cholangitis. *Hepatology* 2019;70(1):294–307. doi:[10.1002/hep.30604](https://doi.org/10.1002/hep.30604).
- [41] Donaldson PT, Baragiotta A, Heneghan MA, et al. HLA class II alleles, genotypes, haplotypes, and amino acids in primary biliary cirrhosis: a large-scale study. *Hepatology* 2006 Published online. doi:[10.1002/hep.21316](https://doi.org/10.1002/hep.21316).
- [42] Nakamura M, Yasunami M, Kondo H, et al. Analysis of HLA-DRB1 polymorphisms in Japanese patients with primary biliary cirrhosis (PBC): the HLA-DRB1 polymorphism determines the relative risk of antinuclear antibodies for disease progression in PBC. *Hepatol Res* 2010 Published online. doi:[10.1111/j.1872-034X.2010.00631.x](https://doi.org/10.1111/j.1872-034X.2010.00631.x).
- [43] Vázquez-Elizondo G, Ponciano-Rodríguez G, Uribe M, Méndez-Sánchez N. Human leukocyte antigens among primary biliary cirrhosis patients born in Mexico. *Ann Hepatol* 2009 Published online. doi:[10.1016/s1665-2681\(19\)31808-3](https://doi.org/10.1016/s1665-2681(19)31808-3).
- [44] Adame EC, Granados J, Uribe M, Torre A. Does HLA-DR7 differentiate the overlap syndrome of auto-immune hepatitis-primary biliary cirrhosis (AIH-PBC) from those with auto-immune hepatitis type 1? *Ann Hepatol* 2011 Published online. doi:[10.1016/s1665-2681\(19\)31583-2](https://doi.org/10.1016/s1665-2681(19)31583-2).
- [45] Hu S, Zhao F, Wang Q, Chen WX. The accuracy of the anti-mitochondrial antibody and the M₂ subtype test for diagnosis of primary biliary cirrhosis: a meta-analysis. *Clin Chem Lab Med* 2014 Published online. doi:[10.1515/cclm-2013-0926](https://doi.org/10.1515/cclm-2013-0926).
- [46] Umemura T, Joshita S, Ichijo T, et al. Human leukocyte antigen class II molecules confer both susceptibility and progression in Japanese patients with primary biliary cirrhosis. *Hepatology* 2012 Published online. doi:[10.1002/hep.24705](https://doi.org/10.1002/hep.24705).
- [47] Invernizzi P, Ransom M, Raychaudhuri S, et al. Classical HLA-DRB1 and DPB1 alleles account for HLA associations with primary biliary cirrhosis. *Genes Immun* 2012 Published online. doi:[10.1038/gene.2012.17](https://doi.org/10.1038/gene.2012.17).