Journal of Hepatocellular Carcinoma

Dovepress

open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Fine Mapping of the MHC Region Identifies Novel Variants Associated with HBV-Related Hepatocellular Carcinoma in Han Chinese

Haoming Mai,[1,*] Jiaxuan Chen,[1,*] Haitao Chen,[1,2,*] Zhiwei Liu,[3,*] Guanlin Huang,[1] Jialin Wang,[1] Qianyi Xiao,[4] Weihua Ren,[5] Bin Zhou,[1] Jinlin Hou,[1] Deke Jiang[1]

[1]State Key Laboratory of Organ Failure Research, Guangdong Key Laboratory of Viral Hepatitis Research, Guangdong Institute of Liver Diseases, Department of Infectious Diseases and Hepatology Unit, Nanfang Hospital, Southern Medical University, Guangzhou, 510515, People's Republic of China; [2]School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, 518107, People's Republic of China; [3]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA; [4]School of Public Health, Fudan University, Shanghai, 200032, People's Republic of China; [5]Central Laboratory, First Affiliated Hospital, Henan University of Science and Technology, Luoyang, Henan Province, 471009, People's Republic of China

*These authors contributed equally to this work

Correspondence: Deke Jiang; Jinlin Hou
State Key Laboratory of Organ Failure Research, Guangdong Key Laboratory of Viral Hepatitis Research, Guangdong Institute of Liver Diseases, Department of Infectious Diseases and Hepatology Unit, Nanfang Hospital, Southern Medical University, Guangzhou, 510515, People's Republic of China
Tel +86-20-62786533; +86-20-61641941
Email dekejiang17@smu.edu.cn; jlhousmu@163.com

Introduction: Genome-wide association studies identified susceptibility loci in the major histocompatibility complex region for hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC). However, the causal variants underlying HBV-related HCC pathogenesis remain elusive.

Methods: With a total of 1,161 HBV-related HCC cases and 1,353 chronic HBV carriers without HCC, we imputed human leukocyte antigen (HLA) variants based on a Chinese HLA reference panel and evaluated the associations of these variants with the risk of HBV-related HCC. Conditional analyses were used to identify independent signals associated with the risk of HBV-related HCC (P false-discovery rate (FDR) <0.20). A total of 14,930 variants within the MHC region were genotyped or imputed.

Results: We identified two variants, rs114401688 ($P = 1.05 \times 10^{-6}$, $P_{FDR} = 2.43 \times 10^{-3}$) and rs115126566 ($P = 9.04 \times 10^{-5}$, $P_{FDR} = 1.77 \times 10^{-1}$), that are independently associated with the risk of HBV-related HCC. Single nucleotide polymorphism (SNP) rs114401688 is in linkage disequilibrium with a previously reported SNP rs9275319. In the current study, we found that its association with HCC could be explained by HLA-DQB1*04 and HLA-DRB1*04. SNP rs115126566 is a novel risk variant and may function by regulating transcriptions of HLA-DPA1/DPB1 through enhancer-mediated mechanisms. HLA zygosity analysis showed that homozygosity at HLA-DQB1 gene is suggestively associated with a higher risk of HCC (P = 0.10) and the risk was more pronounced in the older age group (age ≥50, P = 0.03).

Discussion: Our findings further the understanding of the genetic basis for HBV-related HCC predisposition in chronic HBV carriers.

Keywords: fine mapping, susceptibility, MHC, hepatocellular carcinoma, HBV

Introduction

Liver cancer is the sixth most common malignancy worldwide. Globally, it is estimated that approximately 841,000 new liver cancer cases were identified in 2018. Liver cancer is also the fourth leading cause of cancer mortality in 2018, resulting in approximately 781,000 deaths.[1] About 85–90% of primary liver malignancies are hepatocellular carcinoma (HCC).[2] Approximately 85% of HCC in China is hepatitis B virus (HBV)-related.[3] However, only a small fraction of chronic HBV carriers develop HCC. Despite the fact that pathogenesis of HBV-related HCC has not been fully understood,[4,5] host susceptibility is considered as one of the important risk factors for the progression of HBV-related HCC among chronic HBV carriers.[6–9]

Received: 26 May 2021
Accepted: 29 July 2021
Published: 16 August 2021

Genome-wide association studies (GWASs) have been widely used to identify different risk loci of various diseases.[10,11] Twelve susceptibility loci of HBV-related HCC have been identified over the past decade, of which six variants are located in the major histocompatibility complex (MHC) region.[6–9,12–14] Human leukocyte antigens (HLA), cell-surface proteins encoded from the MHC region, play a pivotal role in cancer immune surveillance and viral clearance in various virus-related cancers by presenting antigens to T-cells.[15–17] Alleles and polymorphisms of HLA may have diverse impacts on antigen presentations in different processes and therefore show disease-specific associations.[15] HLA gene variation is associated with risk of cancers, particularly for those with infectious etiology.[18] For instance, variants of *HLA-DPA1* and *HLA-DPB1* have been associated with cervical cancer, a cancer that is caused by human papillomavirus (HPV) infection.[19] As another example, *HLA-A*11:01* could be a protective marker for nasopharyngeal carcinoma caused by Epstein–Barr virus (EBV).[20] These studies suggest the role of HLA polymorphisms in virus-related cancer immune surveillance. Identification and characterization of HLA polymorphisms may help to deepen our understanding of cancer risk.

It is challenging to interpret MHC associations as it is one of the most polymorphic regions in the human genome with the strongest linkage disequilibrium (LD) and high gene density. Recently, Zhou et al have performed a deep sequencing of the entire MHC region in more than 10,000 Han Chinese healthy individuals.[21] This database provides us with a high-quality reference to further infer the HLA alleles in Han Chinese.[21] In the present study, we performed a fine-mapping analysis based on our previously published GWAS by imputing HLA variations using this Chinese HLA reference panel to dissect in detail the genetic causal factors underlying HBV-related HCC pathogenesis.

Materials and Methods

Study Participants and Genotyping

A total of 2,514 chronic HBV carriers were included in the present study, including 1,161 HBV-related HCC cases and 1,353 chronic HBV carriers without HCC at recruitment as controls (Supplementary Table 1). Detailed information on these subjects was described elsewhere.[6] Briefly, chronic HBV carriers were defined as individuals who were positive for both HBV surface antigen (HBsAg) and immunoglobulin G antibody to HBV core antigen for at least 6 months. Diagnosis with HCC was based on (i) positive images on angiogram, ultrasonography, computed tomography and/or magnetic resonance imaging, combined with an α-fetoprotein concentration of ≥400 ng/mL and/or (ii) positive findings on cytological or pathological examination. All patients did not have HIV or HCV antibodies and other types of liver disease, such as primary biliary cirrhosis, toxic hepatitis and autoimmune hepatitis. Before participating in this study, informed consent was obtained from all participants. The ethical committees of all institutions involved in this study have approved the study. This study was conducted in accordance with Declaration of Helsinki principles.

Genome-wide scan was performed using Illumina Human OmniExpress BeadChips at Genergy Biotech covering 733,202 SNPs, from which 6,271 SNPs located within the MHC regions (29 to 34 Mb on chromosome 6, NCBI Build 37) were extracted.

Imputation of HLA Variants and Quality Control

We performed HLA imputation using the software SNP2HLA[22] based on the reference panel of the Han Chinese population (n = 10,689).[21] A total of 29,948 HLA variants (ie, SNPs, HLA amino acid polymorphisms and HLA alleles) were included in this reference panel. We applied post-imputation quality control to filter imputed HLA variants with minor allele frequency (MAF) less than 5% and imputation quality (INFO) less than 0.7 (Figure 1). 'Best-guess' genotypes were generated by SNP2HLA for imputed variants. These variants are major and minor alleles for biallelic variants or presence and absence of each allele for multiallelic variants.[22]

Consistency Between Imputation Method and Sequence-Based Typing (SBT) Method

To validate the imputation performance, we compared the imputed HLA alleles to the 4-digit HLA sequencing data (obtained from sequence-based typing [SBT] approach) on three HLA class II loci (ie, *HLA-DQA1, -DQB1*, and *-DRB1*) for 936 subjects (478 HCC cases and 458 controls) included in this study. Sensitivity, specificity, and accuracy for each HLA allele obtained from SBT were investigated.

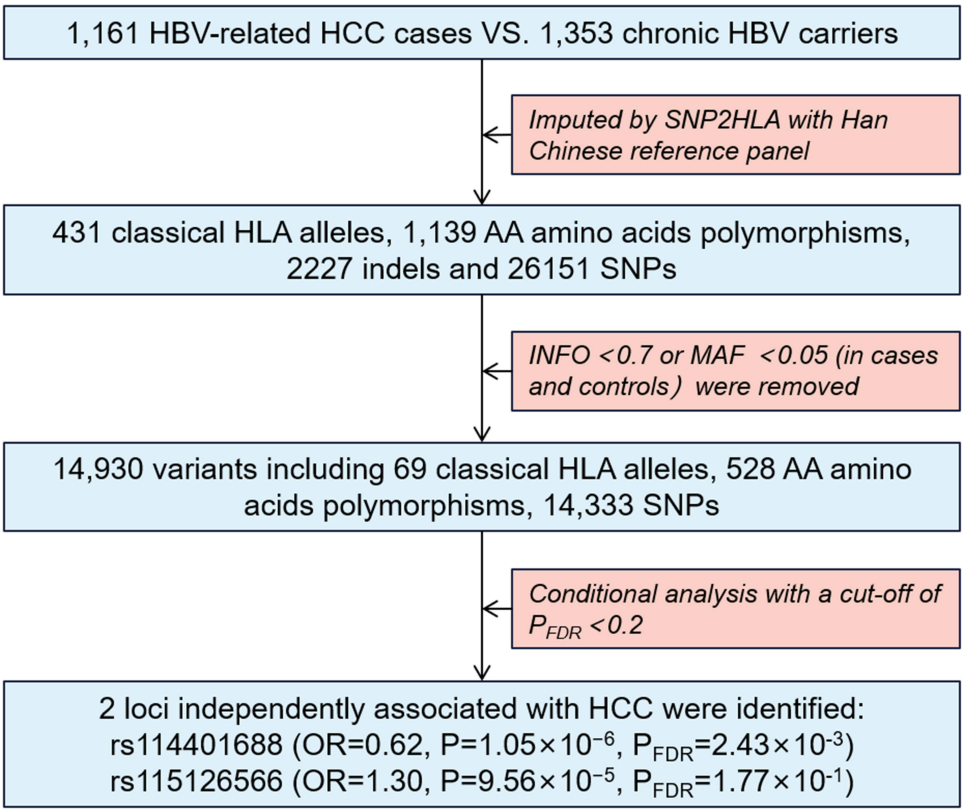

Figure 1 Schematic of the study design and workflow.

Association Analysis of HLA Variants with HBV-Related HCC Risk

All the HLA variants were defined as biallelic variable (including biallelic SNPs, two- and four-digit biallelic classic HLA alleles, and biallelic HLA amino acid polymorphisms for respective residues). By assuming an additive model, logistic regression was used to evaluate the association of HLA variants with HCC risk adjusting for age and sex. Principal components were not included in the analysis because no population substructure is observed. Conditional analyses were used to identify the independent signals, that is, the most significant variant was included in the next model as covariates and repeated the same step until no variants reached the study-wide significance threshold, which was defined as $P_{FDR} < 0.20$ in our study.

Association of HLA Zygosity with HBV-Related HCC Risk

It is hypothesized that diversity in the HLA genes is associated with disease outcomes – the heterozygote advantage hypothesis. We also tested this hypothesis by assessing the association between HLA zygosity and HBV-related HCC using logistic regression models. Individuals were classified as homozygotes at a given locus when imputed to carry the same 4-digit allele for the two HLA alleles at that locus.

Functional Annotation of Independent HBV-Related HCC Association

NESDA NTR Conditional eQTL Catalog (https://eqtl.onderzoek.io/) was used to perform the expression quantitative trait loci (eQTL) analysis in peripheral blood for SNPs independently associated with HCC.

To investigate their cis-regulatory roles, we mapped SNPs to H3K27AC ChIP-seq, DNase peaks, and transcription factors binding sites using WASHU epigenome browser (http://epigenomegateway.wustl.edu/browser/).

Results

General Description of the Study Population and Imputation Results

All participants were Han Chinese and recruited from Qidong, an area with the highest incidence of HCC in China (Supplementary Table 1). Based on the HLA

reference panel of Han Chinese population (n = 10,689),[21] we imputed 14,930 HLA variants from 1,161 HBV-related HCC cases and 1,353 CHB controls using SNP2HLA software.[22] Those HLA variants included 14,333 SNPs, 69 two-/four-digit HLA alleles and 528 amino acid polymorphisms (Figure 1).

We evaluated the sensitivity, specificity and accuracy of HLA alleles imputed by SNP2HLA, against 12 HLA alleles obtained based on SBT (Supplementary Table 2). Overall, we found high concordance rates for all 12 HLA alleles obtained from SBT, with at least 95% sensitivity, 97% specificity, and 97% accuracy.

Association and Conditioning Analyses Identified 2 Independent Signals Driving HBV-Related HCC Risk

We conducted logistic regression analysis of the SNPs, HLA alleles, and amino acid polymorphisms within the MHC region assuming additive effects of the allele dosages on the log-odds scale. We observed 566 significant variants, including 522 SNPs, 7 HLA alleles (3 two-digit HLA alleles and 4 four-digit HLA alleles) and 37 amino acid polymorphisms associated with HBV-related HCC after the FDR correction ($P_{FDR} < 0.2$, Supplementary Table 3). The regional association results are shown in Figure 2A. The most significant association was observed for SNP rs114401688 located in the MHC class II region (odds ratio [OR] = 0.62, $P = 1.05 \times 10^{-6}$, $P_{FDR} = 2.43 \times 10^{-3}$, Table 1), which was in perfect LD with a previously reported SNP rs9275319 ($r^2 = 1$).

We then conducted analysis conditioning on rs114401688 and identified another SNP (rs115126566) that satisfied the statistical significance after FDR correction (OR = 1.3, $P = 9.56 \times 10^{-5}$, $P_{FDR} = 1.77 \times 10^{-1}$, Figure 2B and Table 1). In our data, this SNP was not in strong LD with any imputed HLA-alleles and has not been reported previously, suggesting that this variant could be novel in its association with HBV-related HCC risk.

After conditioning on rs114401688 and rs115126566, no additional significant association was observed in the MHC region (Figure 2C).

Functional Exploration of rs114401688 Risk

The SNP rs114401688 represented the same signal rs9275319 ($r^2 = 1$) previously identified by our GWAS.[6] However, to our knowledge, the causal variants of this association have not been dissected in detail.

Although no statistical significance ($P > 0.05$) was observed in conditioning analysis on the top SNP (rs114401688, Supplementary Table 3), in analyses without conditioning, the most significant association was observed at HLA-DQB*04 (OR = 0.56, P = 4.43×10^{-6}, $P_{FDR} = 2.43 \times 10^{-3}$, Table 2), followed by HLA-DRB*04 (OR = 0.66, P = 1.96×10^{-5}, $P_{FDR} = 5.98 \times 10^{-3}$, Table 2). In our results, HLA-DQB*04 could be tagged by residues Leu56, Glu70 and Asp71, which are all located in α-helix wall of HLA-DQ peptide-binding groove (Supplementary Table 4, Figure 3A). And *HLA-DRB*04* was in perfect LD with amino acid polymorphisms at 5 positions, among which His13 and His33 are located within the β-sheet floor of HLA-DR peptide-binding groove (Supplementary Table 4, Figure 3B). We found SNP rs114401688 was in strong LD with HLA-DQB1*04 ($r^2 = 0.59$) and DRB1*04 ($r^2 = 0.78$). When conditioning on HLA-DQB1*04 or DRB1*04, the significance of rs114401688 experienced a dramatic reduction (Table 3). Further, when jointly conditioning on HLA-DQB1*04 and DRB1*04, the significance of rs114401688 was eliminated (Table 3).

Among the 4-digit HLA alleles, the most significant association was observed at HLA-DQB1*04:01 (OR = 0.61, $P = 1.70 \times 10^{-4}$, $P_{FDR} = 2.87 \times 10^{-2}$, Table 2), followed by HLA-DQA1*03:03, HLA-DRB1*04:05 and HLA-B*46:01. SNP rs114401688 was in moderate LD with HLA-DQB1*04:01 ($r^2 = 0.50$), HLA-DQA1*03:03 ($r^2 = 0.41$) and HLA-DRB1*04:05 ($r^2 = 0.49$). Similarly, the significance of all these HLA variants was eliminated when conditioning on rs114401688 (Supplementary Table 3).

Collectively, our results demonstrate that the association between rs114401688 and HBV-related HCC may be driven jointly by HLA-DQB*04 and HLA-DRB*04. And both of these HLA alleles could be protective markers for HBV-related HCC.

Functional Annotation of rs115126566 Risk

Meanwhile, as mentioned above, we did not identify any HLA alleles or amino acid polymorphisms were in LD ($r^2 > 0.2$) with rs115126566, which indicated that polymorphisms within HLA peptide-binding groove could not account for the signal at rs115126566.

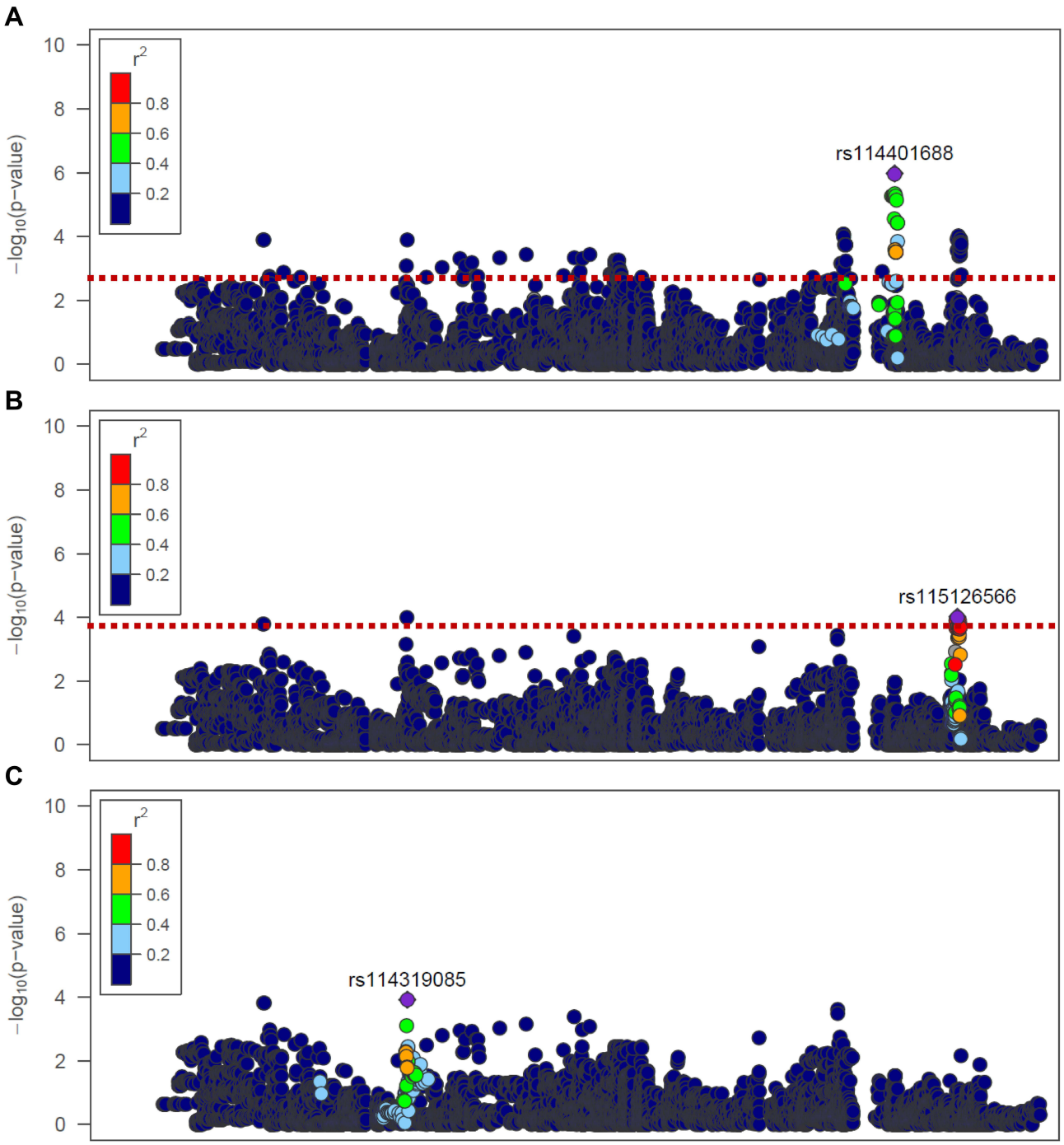

Figure 2 Regional association plots of HLA loci independently associated with HBV-related HCC risk. Each dot represents the $-\log_{10} P$ of HLA variants, including SNPs, classical alleles and amino acid polymorphisms. The red horizontal dashed line represents $P_{FDR} = 0.20$. (**A**) The top signal was SNP rs114401688. (**B**) After conditioning on rs114401688, the most significant independent association was rs115126566. (**C**) After conditioning on rs114401688 and rs115126566, no additional significant association was observed in the MHC region.

Increasing evidence suggests that SNPs associated with complex traits are more likely to be eQTLs.[23,24] To explore the potential mechanisms underlying the associations between SNPs within the MHC region and HCC risk, we conducted eQTL analysis for rs115126566. In the public data from NESDA NTR Conditional eQTL Catalog, we found that rs115126566 has been identified as a strong eQTL for HLA-DPB1 ($P = 3.75\times10^{-34}$) and HLA-DPA1 ($P = 1.05\times10^{-23}$) in peripheral blood.[25] Then, we used Haploreg4.1 to explore the potential function of rs115126566 (rs9277053). We found that SNP rs116477415 (rs9277027), an SNP in a strong LD with

Table 1 Associations of Two Independent Variants of HBV-Related HCC Genetic Susceptibility Among Chronic HBV Carriers

Variants	Nearby Gene	Effect Allele	EAF[a]		OR[b] (95% CI)	P^b	$P_{FDR}{}^b$
			Cases	Controls			
rs114401688	*HLA-DQB1*	G	0.081	0.120	0.61 (0.51–0.75)	1.05×10^{-6}	2.43×10^{-3}
rs115126566	*HLA-DPA1*	A	0.280	0.120	1.29 (1.19–1.47)	9.04×10^{-5}	2.37×10^{-2}

Note: [b]Adjusted for age and gender.
Abbreviation: [a]EAF, effect allele frequency.

Table 2 Associations of Top HLA Alleles of HBV-Related HCC Genetic Susceptibility Among Chronic HBV Carriers

Variants	EAF[a]		OR[b] (95% CI)	P^b	$P_{FDR}{}^b$
	Cases	Controls			
*HLA-DQB1*04*	0.046	0.077	0.57 (0.44–0.72)	5.38×10^{-6}	2.43×10^{-3}
*HLA-DRB1*04*	0.082	0.12	0.66 (0.54–0.80)	1.96×10^{-5}	5.98×10^{-3}
*HLA-DQB1*04:01*	0.041	0.064	0.61 (0.47–0.79)	1.70×10^{-4}	2.87×10^{-2}
*HLA-DQA1*03:03*	0.049	0.074	0.63 (0.50–0.81)	2.04×10^{-4}	3.25×10^{-2}
*HLA-DRB1*04:05*	0.041	0.062	0.63 (0.48–0.81)	4.56×10^{-4}	5.12×10^{-2}
*HLA-B*46:01*	0.170	0.140	1.27 (1.08–1.49)	3.08×10^{-3}	1.48×10^{-1}

Note: [b]Adjusted for age and gender.
Abbreviation: [a]EAF, effect allele frequency.

rs115126566 ($r^2 = 0.98$), was overlapped with the active enhancer marker H3K27ac in immune cells, especially B cells (Figure 4), which indicated that rs116477415 may regulate HLA gene expression in a tissue-specific manner. Moreover, according to ChIP-seq data from the database Encyclopedia of DNA Elements (ENCODE), a peak could be observed in this region by Early B Cell Factor 1 (EBF1) ChIP-seq (Figure 4), a key regulator of B lineage specification and differentiation.[26] In fact, rs116477415 has been demonstrated as a B cell-specific eQTL for *HLA-DPA1* in a previous study ($P = 1.99\times10^{-9}$).[27] These evidences indicated that

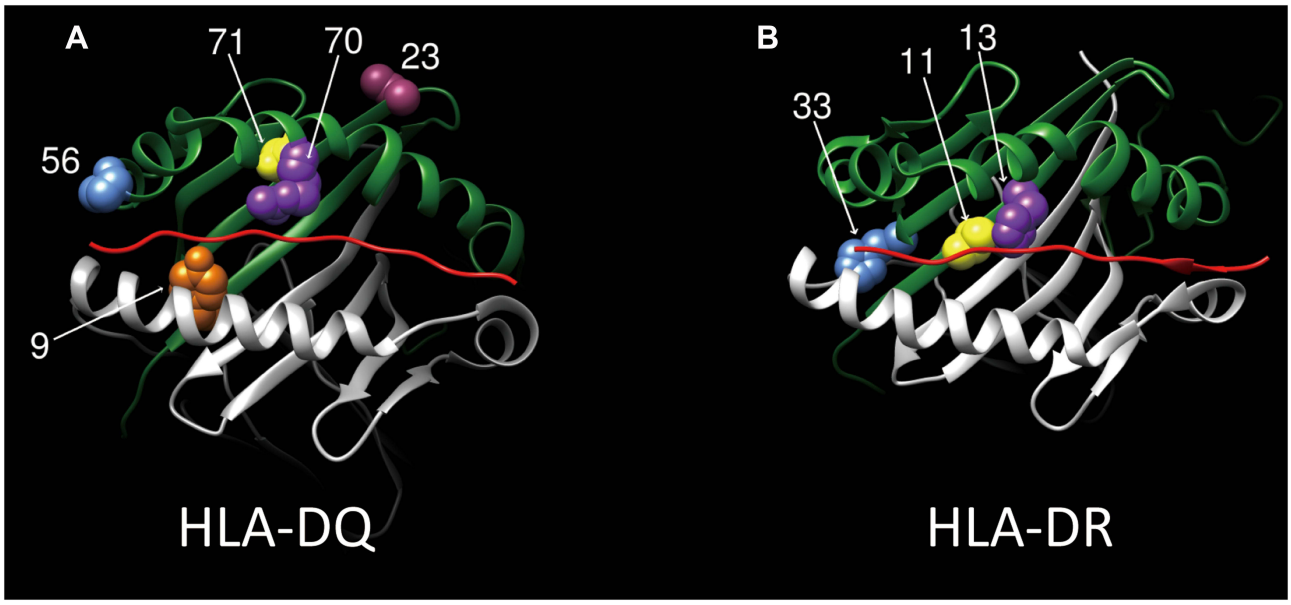

Figure 3 Three-dimensional ribbon models of HLA amino acid polymorphisms tagged by HLA-DQB1*04 and HLA*DRB1*04 (peptide binding grooves) associated with HBV-related HCC risk. The protein structures of HLA-DQ (**A**) and HLA-DR (**B**) are based on Protein Data Bank entries 4Z7W and 5LAX, respectively. This figure was prepared by using UCSF Chimera.

Table 3 The Association of HBV-Related Risk at rs114401688 After Conditioning on HLA-DQB1*04 and HLA-DRB1*04

CHR	Variants	Reference Allele	Effect Allele	OR[a]	P[a]
6	rs114401688	A	G	0.62	1.05×10^{-6}*
Conditional on *HLA-DQB1*04*					
6	rs114401688	A	G	0.72	3.06×10^{-2}*
Conditional on *HLA-DRB1*04*					
6	rs114401688	A	G	0.61	1.54×10^{-2}*
Conditional on *HLA-DQB1*04* and *HLA-DRB1*04*					
6	rs114401688	A	G	0.72	2.02×10^{-1}

Notes: [a]Adjusted for age and gender; *Statistically significant results.

rs115126566 may influence HBV-related HCC risk, through SNP rs116477415, by regulating HLA class II gene expression in B cells.

Association Analysis for Previously Reported HLA Variants

Of the previously reported six SNPs within the MHC region, associations with HBV-related HCC were confirmed for two SNPs, rs9272105 ($P = 1.22 \times 10^{-3}$, Supplementary Table 5),[7] and rs9275319 (in perfect LD with rs114401688). Of note, the association with rs9272105 was attenuated after conditioning on SNP rs114401688 (ie, rs9275319, P>0.05).

Previous studies have reported 8 HLA-alleles (DQA1*03:01, DQA1*06:01, DQB1*04:01, DR B1*04:05, DRB1*09:01, DQB1*03:02/3, DQB1*04:02 and A*33:03) that were associated with HBV-related HCC.[7,9,28] Among those, we confirmed the associations with HLA-DQB1*04:01 and HLA-DRB1*04:05 but not others. This could be partly due to the modest sample size and different distributions of HLA alleles across ethnic groups.

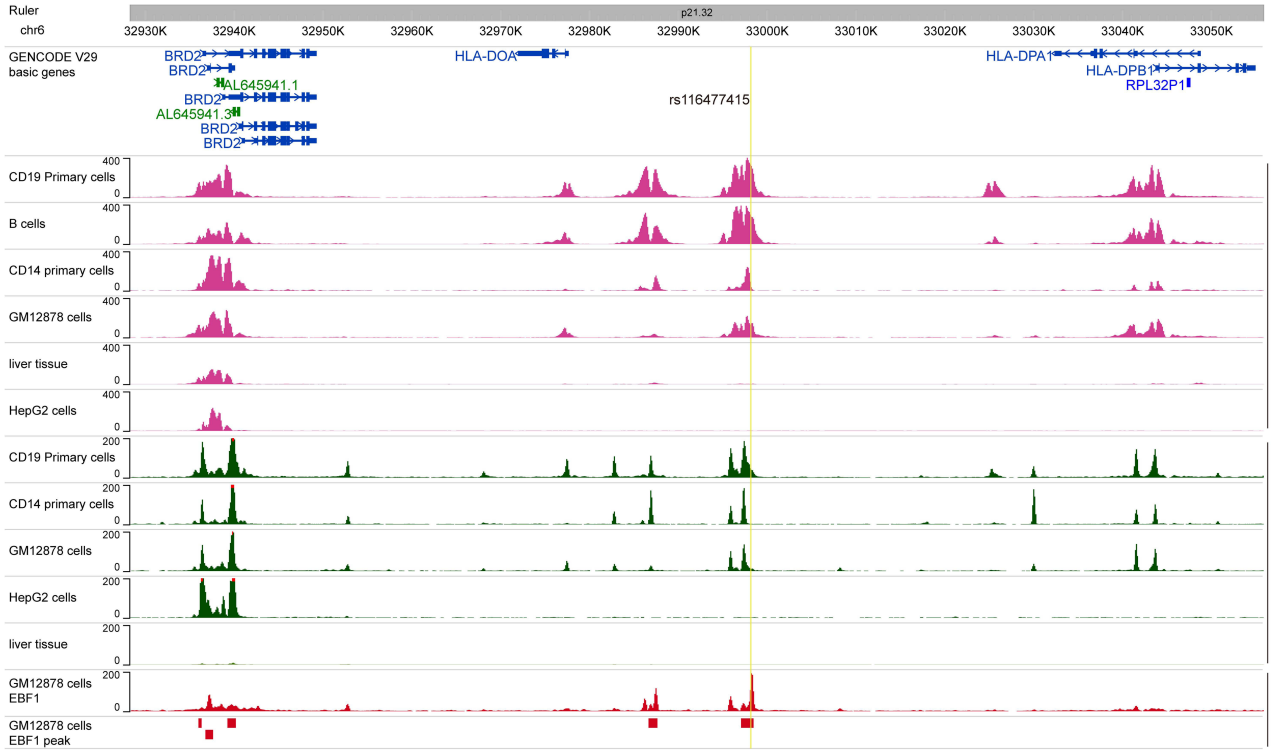

Figure 4 WASHU epigenome browser present in the MHC region around rs116477415.

Association of HLA Zygosity with HBV-Related HCC Risk

To test the heterozygosity advantage hypothesis,[29] we evaluated the association of the HLA zygosity at class II genes with the risk of HCC (Supplementary Table 6). We found that the zygosity at *HLA-DQB1* is likely to be associated with the risk of HBV-related HCC, with an adjusted OR of 1.19 (0.97–1.46, $P = 0.10$). This effect seems to be more pronounced in the older age group (≥ 50 vs < 50, P for heterogeneity = 0.14), with an adjusted OR of 1.38 (95% CI: 1.04–1.84, $P = 0.03$).

Discussion

MHC is an extremely polymorphic region and is considered as the most gene-dense region in the human genome.[15] Recent studies have successfully identified several susceptibility variants within the MHC region in many diseases, and fine-mapping has been proved to be a powerful analysis in these studies.[30–40] In the present study, we imputed the HLA variants based on the newly developed Chinese Han reference panel,[21] and identified 2 variants that independently confer risk for HBV-related HCC. SNP rs114401688 was in perfect LD with the top signal in our previous GWAS, rs9275319.[6] Our findings also suggest that SNP rs115126566, an intergenic variant, may explain the HBV-related HCC risk associated within the MHC region among the Chinese Han population.

Previous GWAS have reported rs114401688 was a validated risk locus for HBV-related HCC and chronic HBV infection.[6,41] However, the causal variants for this locus have not been deciphered so far. The SNP rs114401688 is highly correlated with several HLA alleles and amino acid polymorphisms. In the current study, conditional analysis suggested that this previously reported HBV-related HCC susceptibility locus could be explained by HLA-DQB1*04 and HLA-DRB1*04. Therefore, both of these HLA alleles could be protective markers for HBV-related HCC. We also found that *HLA-DQB1*04* (OR = 0.78, $P = 1.56\times10^{-2}$) and *HLA-DRB1*04* (OR = 0.71, $P = 1.85\times10^{-5}$) were also inversely associated with chronic hepatitis B (CHB) risk (unpublished data),[42] which was consistent with previous studies, which identify rs9275319 as a protective marker for chronic HBV infection.[43,44] In addition, the associations between *HLA-DQB1*04:01* and *HLA-DRB1*04:05* and HBV-related HCC risk have been observed by ours and one previous GWAS in Chinese Han.[7] Such associations need to be replicated to validate their independent effect on HBV-related HCC.

In contrast, the association between rs115126566 and HCC risk is independent of all HLA alleles of interest. In addition, we did not observe any LD between rs115126566 and other reported HBV-related HCC risk-associated variants. Several studies reported that HLA genes expression could also affect immune response, in which the SNP may play a role as cis-regulatory element.[23,45] These studies prompted us to decipher the association of these SNPs beyond polymorphism in HLA peptide-binding groove. Our eQTL analysis showed that the rs115126566 might cis-regulate the expression of HLA-DPA1 and HLA-DPB1 in peripheral blood. If replicated, further studies are needed to understand its role in HCC development. These results may pave the way to understand the underlying mechanisms of HBV-related HCC mediated by HLA variants.[15]

Our findings also support the heterozygote advantage hypothesis in the case of HBV-related HCC, particularly in the older age group. It is expected that heterozygotes are able to recognize and bind a more diverse set of peptides than homozygotes, leading to more T cell clonal expansion and more efficient-specific cytotoxic T lymphocyte responses against infections.[46–48] The observed association with HBV-related HCC for *HLA-DQB1* gene is in line with results from a Taiwan study (unpublished work) and further suggests an important role of *HLA-DQB1* gene on the progression of HCC among CHB.

Strengths of this study include its large sample size and setting in a region that is ethnically homogeneous. Based on the newly developed reference panel, to the best of our knowledge, this is the first study among the Chinese population that comprehensively assessed the risk of both HLA class I and class II genes in association with HBV-related HCC among chronic HBV carriers. However, there are some limitations. First, we lacked replication set to validate the significant findings. Second, information on HLA alleles and amino acid polymorphisms were obtained by imputation, thereby misclassification cannot be ruled out. Third, although this is the largest study so far studying the HLA association with HCC, our sample size is still modest to identify rare variants and HLA alleles in 4-digit resolution. Fourth, the functional annotation of rs114401688 and rs115126566 in this study need further experimental characterization.

Conclusions

In summary, our fine-mapping analysis using HLA imputation with the large-scale population-specific reference panel identified one novel locus that is independently associated with HBV-related HCC risk among chronic HBV carriers in China. Our results provide deeper insight into the disease biology of HBV-related HCC. Future studies focusing on biological mechanisms underlying the HLA-HCC association are warranted.

Abbreviations

HCC, hepatocellular carcinoma; HBV, hepatitis B virus; GWAS, genome-wide association study; MHC, major histocompatibility complex; HLA, human leukocyte antigen; HPV, human papillomavirus; EBV, Epstein–Barr virus; LD, linkage disequilibrium; HBsAg, HBV surface antigen; MAF, minor allele frequency; eQTL, expression quantitative trait loci; ENCODE, Encyclopedia of DNA Elements; EBF1, Early B Cell Factor 1; CHB, chronic hepatitis B.

Ethics Approval

Ethical Committee of the School of Life Sciences Fudan University approved this study (ID: 4007)

Author Contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

Funding

This study was supported by the Local Innovative and Research Teams Project of Guangdong Pearl River Talents Program (No. 2017BT01S131), the National Science and Technology Major Project (No. 2017ZX10202202 and 2018ZX10301202), the General Programs from the National Natural Science Foundation of China (No. 81472618, 81670535 and 81802828), the National Natural Science Foundation of China, China (No. 82060458 and 81802833) the General Program from the Natural Science Foundation of Guangdong Province (No. 2019A1515011423), the Key-Area Research and Development Program of Guangdong Province (No. 2019B020227004), the Innovative Research Team Project of Guangxi Province (No. 2017GXNSF GA198002), the Dean Fund of Nanfang Hospital, Southern Medical University (No. 2018Z005), the Grant for Recruited Talents to Start Scientific Research from Nanfang Hospital, and the Outstanding Youth Development Scheme of Nanfang Hospital, Southern Medical University (No. 2017J001), Natural Science Foundation of Tibet Autonomous Region (ZRKX2018000153).

Disclosure

Professor Jinlin Hou reports personal fees from AbbVie, personal fees from Arbutus, grants, personal fees from Bristol Myers Squibb, personal fees from Gilead Sciences, grants, personal fees from Johnson & Johnson, personal fees from Roche, outside the submitted work. The authors declare that they have no other conflicts of interest in this work.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424. doi:10.3322/caac.21492
2. Sartorius K, Sartorius B, Aldous C, Govender PS, Madiba TE. Global and country underestimation of hepatocellular carcinoma (HCC) in 2012 and its implications. *Cancer Epidemiol*. 2015;39(3):284–290. doi:10.1016/j.canep.2015.04.006
3. Society, Health Communication, and Chinese Preventive Medicine Association. [Strategies of primary prevention of liver cancer in China: expert consensus (2018)]. *Zhonghua Yu Fang Yi Xue Za Zhi [Chin J Prev Med]*. 2019;53(1):36–44. Chinese
4. Levrero M, Zucman-Rossi J. Mechanisms of HBV-induced hepatocellular carcinoma. *J Hepatol*. 2016;64:S84–S101. doi:10.1016/j.jhep.2016.02.021
5. Wang S, Yeh S, Chen P. Unique features of hepatitis B virus-related hepatocellular carcinoma in pathogenesis and clinical significance. *Cancers*. 2021;13(10):2454. doi:10.3390/cancers13102454
6. Jiang DK, Sun J, Cao G, et al. Genetic variants in STAT4 and HLA-DQ genes confer risk of hepatitis B virus–related hepatocellular carcinoma. *Nat Genet*. 2013;45(1):72–75. doi:10.1038/ng.2483
7. Li S, Qian J, Yang Y. GWAS identifies novel susceptibility loci on 6p21.32 and 21q21.3 for hepatocellular carcinoma in chronic hepatitis B virus carriers. *PLoS Genet*. 2012;8(7):e1002791. doi:10.1371/journal.pgen.1002791
8. Li Y, Zhai Y, Song Q. Genome-wide association study identifies a new locus at 7q21.13 associated with hepatitis B virus-related hepatocellular carcinoma. *Clin Cancer Res*. 2018;24(4):906–915. doi:10.1158/1078-0432.CCR-17-2537
9. Sawai H, Nishida N, Khor SS, et al. Genome-wide association study identified new susceptible genetic variants in HLA class I region for hepatitis B virus-related hepatocellular carcinoma. *Sci Rep*. 2018;8(1):7958. doi:10.1038/s41598-018-26217-7
10. Spracklen CN, Horikoshi M, Kim YJ. Identification of type 2 diabetes loci in 433,540 East Asian individuals. *Nature*. 2020;582(7811):240–245. doi:10.1038/s41586-020-2263-3

11. Ishigaki K, Akiyama M, Kanai Met al. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nat Genet*. 2020;52:669–679.
12. Al-Qahtani A, Khalak HG, Alkuraya FS. Genome-wide association study of chronic hepatitis B virus infection reveals a novel candidate risk allele on 11q22.3. *J Med Genet*. 2013;50(11):725–732. doi:10.1136/jmedgenet-2013-101724
13. Zhang H, Zhai Y, Hu Z. Genome-wide association study identifies 1p36.22 as a new susceptibility locus for hepatocellular carcinoma in chronic hepatitis B virus carriers. *Nat Genet*. 2010;42(9):755–758. doi:10.1038/ng.638
14. Chan KYK, Wong CM, Kwan JSH. Genome-wide association study of hepatocellular carcinoma in Southern Chinese patients with chronic hepatitis B virus infection. *PLoS One*. 2011;6(12):e28798. doi:10.1371/journal.pone.0028798
15. Dendrou CA, Petersen J, Rossjohn J, Fugger L. HLA variation and disease. *Nat Rev Immunol*. 2018;18(5):325–339. doi:10.1038/nri.2017.143
16. Liu XS, Mardis ER. Applications of immunogenomics to cancer. *Cell*. 2017;168(4):600–612. doi:10.1016/j.cell.2017.01.014
17. Anczurowski M, Hirano N. Mechanisms of HLA-DP antigen processing and presentation revisited. *Trends Immunol*. 2018;39(12):960–964. doi:10.1016/j.it.2018.10.008
18. Liu Z, Derkach A, Yu KJ, et al. Patterns of human leukocyte antigen class I and class II associations and cancer. *Cancer Res*. 2021;81:1148–1152.
19. Lin GW, Xu C, Chen K, et al. Genetic risk of extranodal natural killer T-cell lymphoma: a genome-wide association study in multiple populations. *Lancet Oncol*. 2020;21:306–316.
20. Bei JX, Li Y, Jia WH, et al. A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. *Nat Genet*. 2010;42(7):599–603. doi:10.1038/ng.601
21. Zhou F, Cao H, Zuo X. Deep sequencing of the MHC region in the Chinese population contributes to studies of complex disease. *Nat Genet*. 2016;48(7):740–746. doi:10.1038/ng.3576
22. Jia X, Han B, Onengut-Gumuscu S. Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One*. 2013;8(6):e64683. doi:10.1371/journal.pone.0064683
23. Mika KM, Li X, DeMayo FJ, Lynch VJ. An ancient fecundability-associated polymorphism creates a GATA2 binding site in a distal enhancer of HLA-F. *Am J Hum Genet*. 2018;103(4):509–521. doi:10.1016/j.ajhg.2018.08.009
24. Mika KM, Lynch VJ. An ancient fecundability-associated polymorphism switches a repressor into an enhancer of endometrial TAP2 expression. *Am J Hum Genet*. 2016;99(5):1059–1071. doi:10.1016/j.ajhg.2016.09.002
25. Jansen R, Hottenga JJ, Nivard MG, et al. Conditional eQTL analysis reveals allelic heterogeneity of gene expression. *Hum Mol Genet*. 2017;26(8):1444–1451. doi:10.1093/hmg/ddx043
26. Treiber T, Mandel EM, Pott S. Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription- independent poising of chromatin. *Immunity*. 2010;32(5):714–725. doi:10.1016/j.immuni.2010.04.013
27. Fairfax BP, Makino S, Radhakrishnan J. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet*. 2012;44(5):502–510. doi:10.1038/ng.2205
28. Karra VK, Chowdhury SJ, Ruttala R, et al. HLA-DQA1 & DQB1 variants associated with hepatitis B virus-related chronic hepatitis, cirrhosis & hepatocellular carcinoma. *Indian J Med Res*. 2018;147(6):573–580. doi:10.4103/ijmr.IJMR_1644_15
29. Thursz M, Thomas H, Greenwood B, Hill AV. Heterozygote advantage for HLA class-II type in hepatitis B virus infection. *Nat Genet*. 1997;17(1):11–12. doi:10.1038/ng0997-11
30. Huang M, Zhu M, Jiang T, et al. Fine mapping the MHC region identified rs4997052 as a new variant associated with nonobstructive azoospermia in Han Chinese males. *Fertil Steril*. 2019;111(1):61–68. doi:10.1016/j.fertnstert.2018.08.052
31. Zhu M, Dai J, Wang C, et al. Fine mapping the MHC region identified four independent variants modifying susceptibility to chronic hepatitis B in Han Chinese. *Hum Mol Genet*. 2016;25(6):1225–1232. doi:10.1093/hmg/ddw003
32. Chin YM, Mushiroda T, Takahashi A, et al. HLA-A SNPs and amino acid variants are associated with nasopharyngeal carcinoma in Malaysian Chinese. *Int J Cancer*. 2015;136(3):678–687.
33. 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;107:102372.
34. Chu X, Yang M, Song ZJ. Fine mapping MHC associations in Graves' disease and its clinical subtypes in Han Chinese. *J Med Genet*. 2018;55(10):685–692. doi:10.1136/jmedgenet-2017-105146
35. Qin N, Wang C, Zhu M, et al. Fine-mapping the MHC region in Asian populations identified novel variants modifying susceptibility to lung cancer. *Lung Cancer*. 2017;112:169–175. doi:10.1016/j.lungcan.2017.08.016
36. Ferreiro-Iglesias A, Lesseur C, McKay J, et al. Fine mapping of MHC region in lung cancer highlights independent susceptibility loci by ethnicity. *Nat Commun*. 2018;9(1):3927. doi:10.1038/s41467-018-05890-2
37. Kim K, Bang SY, Lee HS, et al. The HLA-DRbeta1 amino acid positions 11-13-26 explain the majority of SLE-MHC associations. *Nat Commun*. 2014;5:5902. doi:10.1038/ncomms6902
38. Tian C, Hromatka BS, Kiefer AK, et al. Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. *Nat Commun*. 2017;8(1):599. doi:10.1038/s41467-017-00257-5
39. Tziotzios C, Petridis C, Dand N, et al. Genome-wide association study in frontal fibrosing alopecia identifies four susceptibility loci including HLA-B*07:02. *Nat Commun*. 2019;10(1):1150. doi:10.1038/s41467-019-09117-w
40. Gutierrez-Achury J, Zhernakova A, Pulit SL, et al. Fine mapping in the MHC region accounts for 18% additional genetic risk for celiac disease. *Nat Genet*. 2015;47(6):577–578. doi:10.1038/ng.3268
41. Kim L, Cheong H, Namgoong S, et al. Replication of genome wide association studies on hepatocellular carcinoma susceptibility loci of STAT4 and HLA-DQ in a Korean population. *Infect Genet Evol*. 2015;33:72–76. doi:10.1016/j.meegid.2015.04.013
42. Jiang DK, Ma XP, Yu H, et al. Genetic variants in five novel loci including CFB and CD40 predispose to chronic hepatitis B. *Hepatology*. 2015;62(1):118–128. doi:10.1002/hep.27794
43. Ji X, Zhang Q, Li B, et al. Impacts of human leukocyte antigen DQ genetic polymorphisms and their interactions with hepatitis B virus mutations on the risks of viral persistence, liver cirrhosis, and hepatocellular carcinoma. *Infect Genet Evol*. 2014;28:201–209. doi:10.1016/j.meegid.2014.09.032
44. Fan J, Hou S, Qing-Ling L, et al. Association of HLA-DQ and IFNL4 polymorphisms with susceptibility to hepatitis B virus infection and clearance. *Ann Hepatol*. 2016;15(4):532–539.
45. Kulkarni S, Savan R, Qi Y, et al. Differential microRNA regulation of HLA-C expression and its association with HIV control. *Nature*. 2011;472(7344):495–498. doi:10.1038/nature09914
46. Godkin A, Davenport M, Hill AV. Molecular analysis of HLA class II associations with hepatitis B virus clearance and vaccine nonresponsiveness. *Hepatology*. 2005;41(6):1383–1390. doi:10.1002/hep.20716
47. Chowell D, Krishna C, Pierini F, et al. Evolutionary divergence of HLA class I genotype impacts efficacy of cancer immunotherapy. *Nat Med*. 2019;25(11):1715–1720. doi:10.1038/s41591-019-0639-4
48. Pierini F, Lenz TL. Divergent allele advantage at human MHC genes: signatures of past and ongoing selection. *Mol Biol Evol*. 2018;35(9):2145–2158. doi:10.1093/molbev/msy116