

Genome-Wide Association Study Reveals Host Genetic Factors for Liver Diseases

Nao Nishida^{1,2}, Katsushi Tokunaga² and Masashi Mizokami¹

¹The Research Center for Hepatitis & Immunology, National Center for Global Health and Medicine, Ichikawa, Chiba, Japan;

²Department of Human Genetics, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

Abstract

A number of disease-associated genetic markers for common liver diseases have been identified using genome-wide association studies (GWASs). The GWAS strategy is based on genome-wide single-nucleotide polymorphism typing technologies, which are now commercially available, accompanied by statistical methods to identify host genetic factors that are associated with target diseases or complex genetic traits. One of the most striking features of the GWAS strategy is the ability to identify unexpected disease-associated genetic markers across the entire human genome. Here, we describe the technological aspects of the GWAS strategy with examples from actual GWAS reports related to hepatitis research, including drug response for patients with chronic hepatitis C, susceptibility to primary biliary cirrhosis, and hepatitis-B-related hepatocellular carcinoma.

© 2013 The Second Affiliated Hospital of Chongqing Medical University. Published by XIA & HE Publishing Ltd. All right reserved.

Introduction

Numerous single-nucleotide polymorphisms (SNPs) have been discovered and deposited in public databases (e.g. the National Center for Biotechnology Information [http://www.ncbi.nlm.nih.gov], Ensembl [http://asia.ensembl.org/index.html], and the MEXT Integrated Database Project [http://dbcls.rois.ac.jp]) through international SNP discovery projects such as the Human Genome Project,¹ the International HapMap project (http://hapmap.ncbi.nlm.nih.gov/index.html.en), and the 1000 Genomes project (www.1000genomes.org). Together with the development of technologies for large-scale SNP genotyping, genome-wide association studies (GWASs) using hundreds of thousands of SNPs allow the identification of candidate genetic loci for multifactorial diseases. Disease-associated SNPs have also been deposited in public databases, such as the database of Genotypes and Phenotypes (www.ncbi.

nlm.nih.gov/gap). Moreover, a number of SNPs have been reported to be associated with complex genetic traits, such as body mass index,² height,³ and hair thickness.⁴ In the National Human Genome Research Institute (NHGRI) GWAS catalog (www.genome.gov), more than 8,800 trait- or disease-associated SNPs with genome-wide significance ($p < 5 \times 10^{-8}$) have been archived from a total of 1,551 published GWAS (through March, 2013).⁵

Here, we describe a GWAS strategy to identify disease-associated SNPs, including SNP genotyping technologies for both the GWAS stage and the following replication stage. Based on this GWAS strategy, we have identified associations of genetic variations with diseases related to hepatitis B and C viruses (HBV and HCV), including drug response in patients with chronic HCV infection,⁶ susceptibility to primary biliary cirrhosis (PBC),⁷ and HBV-related hepatocellular carcinoma (HCC).⁸

Technologies for GWAS and replication analysis

A number of SNP typing methods have been used to analyze a single SNP, or SNPs at multiple sites of a template or templates simultaneously. Most of the methods employ single or multiple site-specific amplifications and a genotyping step based on various types of chemical reactions, including Sanger direct sequencing, 5' exonuclease fluorescence-based assay (TaqMan),⁹ pyrosequencing,¹⁰ DigiTag2 assay,¹¹ single-base extension,¹² and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF).¹³

Together with technology developments in large-scale SNP genotyping, the most recent versions of commercially available genotyping platforms allow the simultaneous analysis of more than one million SNPs across the whole genome in a single experiment. Two platforms are commercially available and widely used for genome-wide SNP typing: Affymetrix SNP GeneChip arrays¹⁴ and Illumina BeadArray genotyping technology.¹⁵ The number of SNPs embedded in both platforms has been gradually increasing since 2003, when the first commercial genome-wide SNP genotyping platform was released by Affymetrix.¹⁶ The first platform of the Affymetrix GeneChip Mapping 10K Array included 14,548 SNPs, which enabled the performance of whole-genome linkage analyses and was indeed used to identify a disease-associated missense mutation in the *HOXD10* gene with Charcot-Marie-Tooth disease through a family-based linkage study.¹⁷ The current versions of the commercial platforms from Affymetrix and Illumina include more than 900,000 SNPs (Genome-Wide Human SNP Array 6.0) and 4.3 million SNPs (HumanOmni5-Quad BeadChip), respectively. A newly released genome-wide SNP typing platform, named the

Keywords: GWAS; Hepatitis B infection; Hepatitis C infection; Primary biliary cirrhosis; HLA-DP; Hepatocellular carcinoma; Host genetic factors.

Abbreviations: GWAS, genome-wide association study; SNP, single nucleotide polymorphism; HBV, hepatitis B virus; HCV, hepatitis C virus; PBC, primary biliary cirrhosis; HCC, hepatocellular carcinoma; PCR, polymerase chain reaction; HLA, human leukocyte antigen; CHB, chronic hepatitis B.

Received: 2 April 2013; Revised: 13 May 2013; Accepted: 22 May 2013

DOI of original article: http://dx.doi.org/10.14218/JCTH.2013.010XX.

Correspondence to: Masashi Mizokami, The Research Center for Hepatitis & Immunology, National Center for Global Health and Medicine, 1-7-1 Kohnodai, Ichikawa, Chiba 272-8516, Japan. Tel: +81-47-372-3501 Ext. 1250, Fax: +81-47-375-4766, E-mail: mmizokami@hospk.ncgm.go.jp

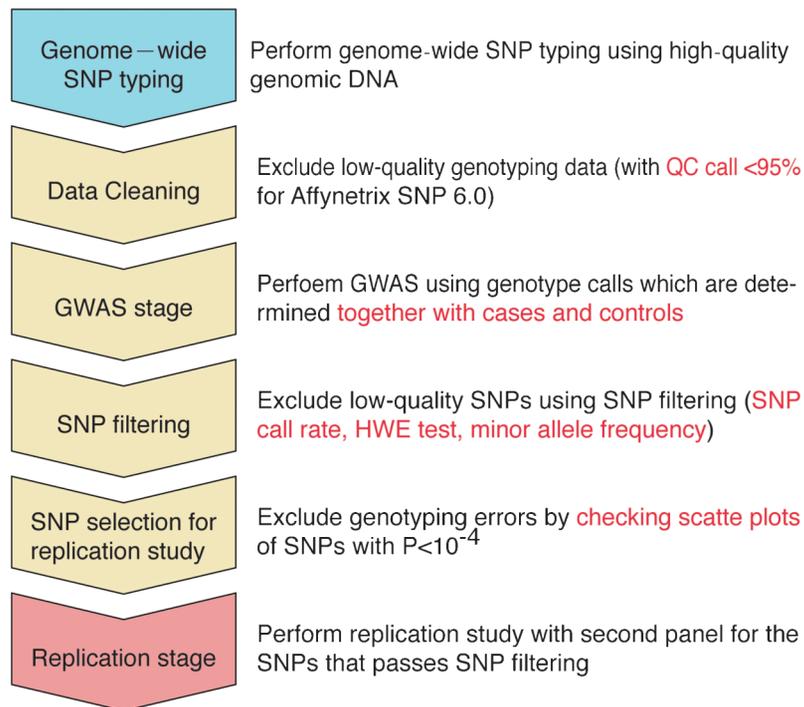


Fig. 1. GWAS strategy from genome-wide SNP typing to replication analysis.

Affymetrix Axiom Genome-Wide ASI 1 Array, has a probe set for SNPs (including rare and common variants) that are optimized for Asian populations. These platforms open a new approach for researchers to perform GWASs with hundreds of thousands of SNPs, allowing the identification of candidate genetic loci for multifactorial diseases.

In 2002, the first GWAS using 92,788 gene-based SNPs was reported by a Japanese group (RIKEN), which identified the lymphotoxin- α gene as being associated with susceptibility to myocardial infarction.¹⁸ The RIKEN group developed its own platform to perform a GWAS based on the Invader assay¹⁹ with multiplex polymerase chain reaction (PCR).²⁰ Since 2002, the number of published genome-wide associations with genome-wide significance ($p < 5 \times 10^{-8}$) has increased annually, reaching 1,551 publications in the NHGRI GWA catalog (through March, 2013).⁵

For a replication study following a GWAS stage, several candidate genetic regions that have been detected in the initial GWAS need to be analyzed. Suitable platforms for replication analyses have the ability to perform multiplex detections in a single reaction, such as the mini-sequencing (SNaPshot) technique,²¹ chip-based genotyping by mass spectrometry (Sequenom),²² and the DigiTag2 assay.¹¹ The DigiTag2 assay is our own technology for multiplex SNP typing, and represents a simple and cost-effective approach by combining multiplex PCR to enrich genetic regions including the target SNPs with an oligonucleotide ligation assay to determine the genotype of the target locus. For a single locus analysis the TaqMan assay would be more commonly used to determine the genotype of the target locus, as opposed to conventional Sanger sequencing, which is more commonly used when a large number of samples need to be analyzed.

Hepatitis research based on GWAS

In a GWAS, two groups of participants are compared to detect the “association(s)” of certain variants with a particular trait by examining differences in allele and/or genotype frequency of all SNPs, which exist across the entire genome. GWAS enables the effective detection of associated variations in strong linkage disequilibrium with the causal variants and genes, and the following replication analysis and high-density mapping identify the causal variants and genes using an independent set of participants with a larger number of samples. However, the association of SNPs with low minor allele frequency (below 1–5%; known as rare variants) would be difficult to detect in a SNP-based GWAS because of insufficient statistical power due to the limitation of sample number.²³ Fig. 1 outlines the GWAS strategy from whole-genome SNP typing to replication analysis.

The emerging strategy of GWAS has revealed disease-causing alleles, or variants that lead to susceptibility to complex polygenic diseases with small additive or multiplicative effects on the disease phenotype. For example, a recent GWAS and subsequent meta-analyses in populations of European descent identified human leukocyte antigen (HLA) and 21 non-HLA susceptibility loci, most of which are involved in interleukin (IL)-12/IL-12 receptor (IL-12R) signaling, tumor necrosis factor (TNF)/toll-like receptor (TLR)-nuclear factor (NF)- κ B signaling, and B-cell differentiation in the development of PBC.^{24–27} PBC is a chronic cholestatic liver disease characterized by chronic non-suppurative destructive cholangitis of the intrahepatic small bile ducts. A high concordance rate in monozygotic twins and familial clustering of patients with PBC indicates the involvement of strong genetic factors in the development of PBC.²⁸ To

Table 1. Replication analysis of Japanese samples for SNPs associated with PBC in previous studies, and two newly identified loci (*TNFSF15* and *POU2AF1*).

Gene name	SNP	OR	95% CI	P-value
Significant associations with PBC				
<i>TNFSF15</i>	rs4979462	1.57	1.76–1.40	1.85×10^{-14}
<i>POU2AF1</i>	rs4938534	1.38	1.55–1.23	3.27×10^{-8}
<i>IKZF3</i>	rs9303277	1.44	1.63–1.28	3.66×10^{-9}
<i>CD80</i>	rs2293370	1.48	1.68–1.29	3.04×10^{-9}
<i>IL7R</i>	rs6890853	1.47	1.69–1.28	3.66×10^{-8}
Suggestive associations with PBC				
<i>NFKB1</i>	rs7665090	1.35	1.52–1.21	1.42×10^{-7}
<i>STAT4</i>	rs7574865	1.35	1.52–1.19	1.11×10^{-6}
Marginal associations with PBC				
<i>CXCR5</i>	rs6421571	1.42	1.75–1.16	0.0004
<i>TNFAIP2</i>	rs8017161	1.22	1.38–1.08	0.0006
<i>MAP3K7IP1(TAB1)</i>	rs968451	1.29	1.52–1.10	0.0009
rs6974491	rs2717948	1.33	1.66–1.07	0.005
<i>DENND1B</i>	rs12134279	1.14	1.33–0.98	0.0405
No apparent associations with PBC				
rs11117432	rs8062669	1.21	1.52–0.96	0.0521
<i>IL12RB2/SCHIP1</i>	rs3790567	1.12	1.28–0.98	0.0540
<i>RPS6KA4</i>	rs538147	1.12	1.28–0.98	0.0554
<i>TNFRSF1A</i>	rs1800693	1.12	1.30–0.97	0.0607
<i>CLEC16A</i>	rs12924729	1.10	1.28–0.94	0.1197
<i>MMEL1</i>	rs3748816	1.07	1.20–0.95	0.1256
<i>PLCL2</i>	rs1372072	1.07	1.20–0.95	0.1396
<i>SPIB</i>	rs3745516	1.08	1.27–0.92	0.1803
<i>IRF5/TNPO3</i>	rs4728142	1.08	1.30–0.90	0.2027
<i>RAD51L1</i>	rs911263	1.07	1.30–0.89	0.2353
<i>IL12A</i>	rs6441286	1.02	1.15–0.91	0.3422

identify susceptibility loci for PBC in the Japanese population, we conducted a GWAS and subsequent replication study using a total of 1,327 PBC patients and 1,120 healthy controls.⁷ In addition to the most significant susceptibility region at *HLA*, two significant susceptibility loci (*TNFSF15* and *POU2AF1*) with *p*-values $< 5 \times 10^{-8}$ were identified (Table 1). Moreover, of the 21 non-*HLA* susceptibility loci that were identified in populations of European descent, three loci (*IKZF3*, *CD80*, and *IL7R*) showed significant associations and two loci (*NFKB1* and *STAT4*) showed suggestive associations with PBC in the Japanese population. Five other loci (*CXCR5*, *TNFAIP2*, *MAP3K7IP1*, rs6974491, and *DENND1B*) also showed marginal associations ($p < 0.05$) with PBC in the Japanese population (Table 1). These results indicate that additional important disease pathways (via *TNFSF15* and *POU2AF1*) – differentiation to T-helper 1 (Th1) cells (via *TNFSF15*, *CD80*, *IL12*, *IL12R*, and *STAT4*; Fig. 2), B-cell differentiation (via *POU2AF1*, *CXCR5*, *SPIB*, *IL7R*, and *IKZF3*), and NF- κ B signaling – in addition to the previously reported disease pathways have a role in the development of PBC in Japanese populations.

In another study that aimed to identify host genetic factors related to drug response to pegylated interferon- α plus ribavirin treatment for HCV infected patients, comparatively small number of samples were analyzed in a GWAS, including

samples from 154 Japanese HCV patients undergoing pegylated interferon- α /ribavirin treatment, 78 null virologic responders, and 64 virologic responders.⁶ Despite the small number of samples in the GWAS in comparison with other studies in European descendants (European American,²⁹ Australian,³⁰ and Swiss³¹), the same disease-causing locus of *IL28B* was identified with the strongest association in the Japanese population. In general, the number of samples affects the statistical power of detection in a GWAS. Moreover, false-positive associations can increase when low-quality genotype data are incorporated in the analysis, presumably caused by accidental errors in genotyping steps or low-quality genomic DNA. The Japanese GWAS was able to successfully identify the risk factors in a small number of samples because: (1) *IL28B* is a strong host risk factor for drug response in Asian and white populations; and (2) quality controls were used in sample collection in terms of clinical characteristics, and the genotype data were checked for quality.¹⁴

As for HBV-related HCC, a GWAS using chronic HBV carriers with and without HCC in five independent Chinese samples found that one SNP (rs17401966) in *KIF1B* was associated with susceptibility to HBV-related HCC.³² Moreover, in the most recent report on this topic, genetic variants in the *STAT4* and *HLA-DQ* genes were identified as

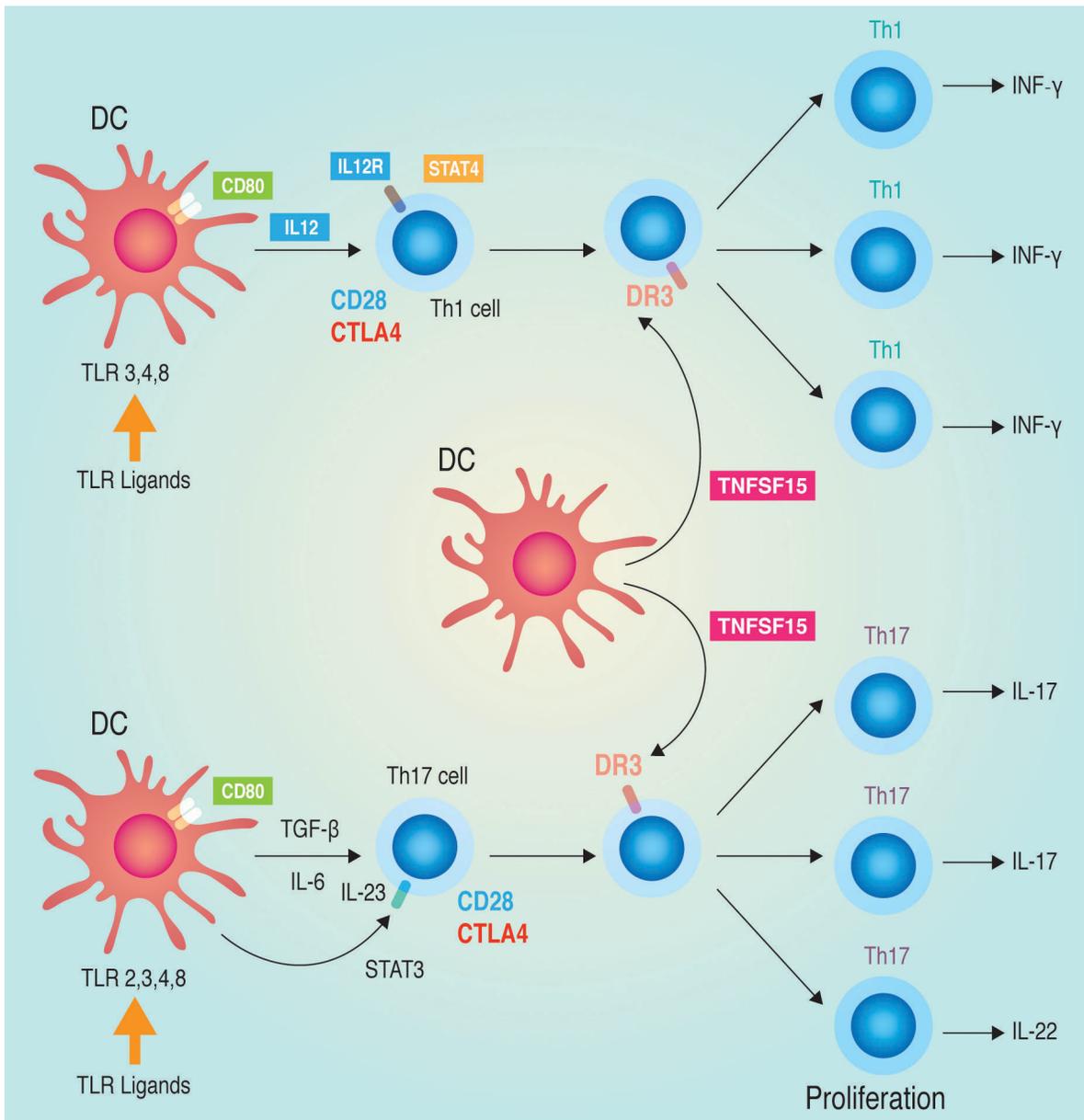


Fig. 2. T-cell proliferation via *TNFSF15* and other related genes. A proportion of susceptibility genes associated with PBC (*CD80*, *IL12A*, *IL12RB2*, *STAT4*, and *TNFSF15*) are related to T-cell proliferation via both Th1 and Th17 cells.

genetic susceptibility loci for HBV-related HCC in the Chinese population.³³ We performed SNP genotyping of rs17401966 on the *KIF1B* gene in Japanese, Korean, and Hong Kong populations for the purpose of replication analysis of a previous GWAS report.⁸ We first examined two independent Japanese HBV-related HCC populations and healthy controls, including 179 patients and 769 controls from Biobank Japan, and 142 patients and 251 controls from various hospitals. We did not detect any associations between rs17401966 and HCC in the Japanese population. We also detected no association of the SNP with HBV-related HCC in Korean and Hong Kong populations using 164 patients and 144 controls, and 94 patients and 187 controls, respectively. In a recent report from another group, no significant association of the

KIF1B gene was observed in HBV-related HCC patients of Saudi Arabian ethnicity.³⁴ These results may be explained by genetic diversity among the Chinese, Japanese, Korean, Hong Kong, and Saudi Arabian populations. The complexity of multivariate interactions in the pathogenesis of HCC may lead to difficulties in identifying the gene(s) associated with HBV-related HCC.

In a previous report that studied 179 Japanese individuals with chronic HBV infection (CHB) and 934 control participants, a GWAS identified significant associations of CHB with a region including *HLA-DPA1* and *HLA-DPB1*.³⁵ The same group was also conducted a second GWAS with a total of 2,667 Japanese patients with persistent HBV and 6,496 controls, which confirmed significant associations between

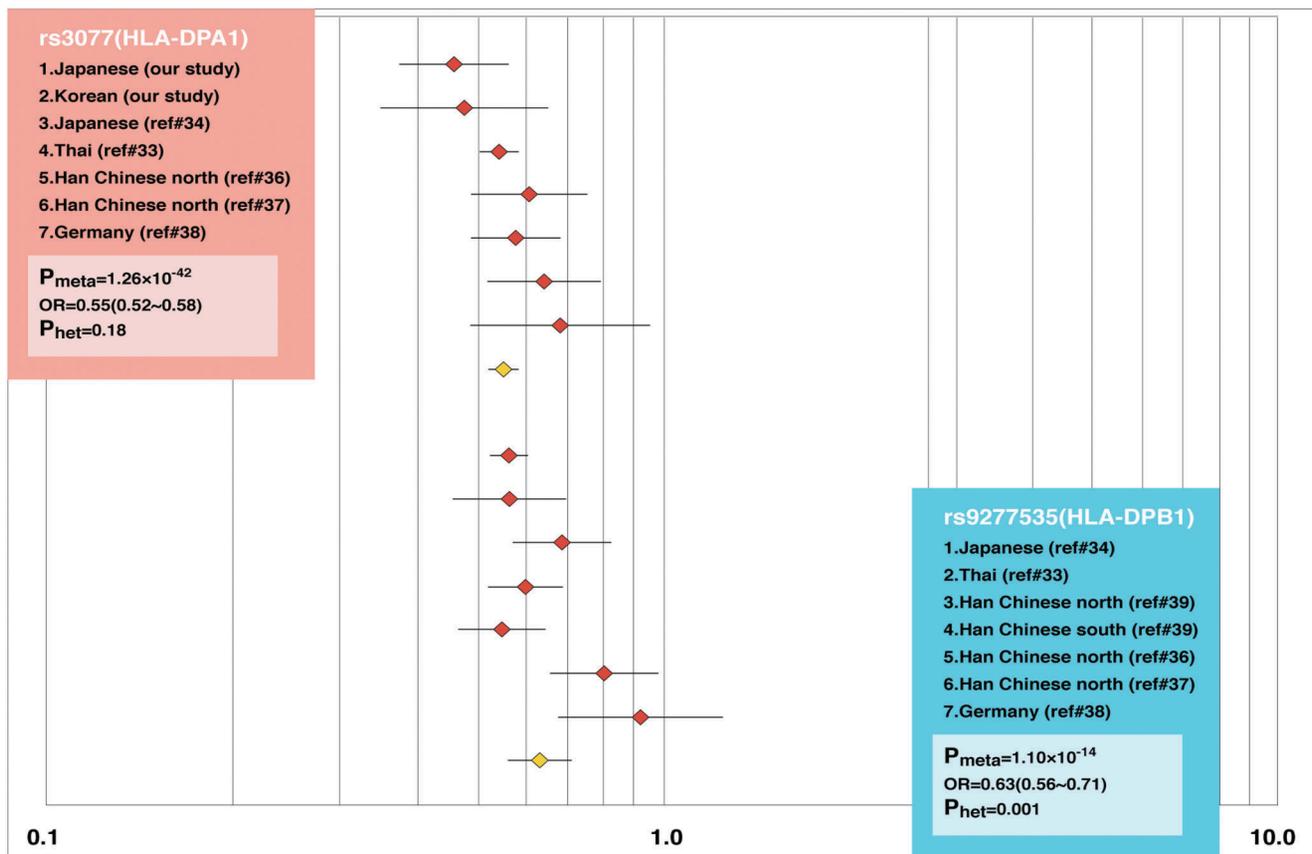


Fig. 3. Associations of HLA-DP with CHB and HBV clearance in Asian populations. Meta-analysis using the random-effects model across seven independent studies, including six additional published data sets, showed $P_{meta}=1.26 \times 10^{-42}$, odds ratio (OR) 0.55 for rs3077, and $P_{meta}=1.10 \times 10^{-14}$, OR 0.63 for rs9277535. Heterogeneity was tested using a general variance-based method.

the *HLA-DP* locus and CHB, in addition to associations with another two SNPs located in a genetic region including the *HLA-DQ* gene.³⁶ We performed a GWAS using samples from Japanese HBV carriers, healthy controls, and individuals who spontaneously resolved HBV infections (hepatitis B surface antigen [HBsAg] negative and hepatitis B core antibody [anti-HBc] positive), in order to confirm or identify the host genetic factors related to CHB and viral clearance.³⁷ In the subsequent replication analysis, we validated the associated SNPs in the GWAS using two independent sets of Japanese and Korean individuals. In our study, healthy controls with no clinical evidence of HBV exposure were randomly selected; therefore, HBV-resolved individuals were prepared to clearly identify the host genetic factors related to CHB or HBV clearance. An association analysis conducted with the Japanese and Korean data identified the *HLA-DPA1* and *HLA-DPB1* genes with $P_{meta}=1.89 \times 10^{-12}$ for rs3077 and $P_{meta}=9.69 \times 10^{-10}$ for rs9277542. We also found that the *HLA-DPA1* and *HLA-DPB1* genes were significantly associated with protective effects against CHB in Asian populations including Japanese, Korean, Chinese, and Thai individuals ($P_{meta}=1.26 \times 10^{-42}$ for rs3077 and $P_{meta}=1.10 \times 10^{-14}$ for rs9277535) (Fig. 3).³⁵⁻⁴¹ The SNP rs9277535 was located about 4 kb upstream from rs9277542 and showed strong linkage disequilibrium of r^2 0.955 in the HapMap JPT (Japanese in Tokyo, Japan) population. These results suggest that the associations between the *HLA-DP* locus and the

protective effects against persistent HBV infection and with clearance of HBV are widely replicated in East Asian populations; however, there are few reports of GWASs in Caucasian or African populations. Further studies are necessary to clarify the pathogenesis of CHB and the mechanisms of HBV clearance, including functional analyses of the *HLA-DP* molecule.

The GWASs described above have successfully identified disease-associated genes or SNPs using different types of genome-wide SNP typing platforms. The embedded SNPs are varied among platforms by selecting the tagging SNPs and the suitable SNPs for their own genotyping strategy; however, the genome coverage among platforms revealed no differences over 60% between the HapMap CEU samples and the HapMap JPT+CHB samples.⁴² Moreover, the necessary to clarify the pathogenesis of CHB and the mechanisms of HBV clearance, including functional analyses of the *HLA-DP* molecule. The genome coverage of the current version of the Affymetrix Genome-Wide Human SNP Array 6.0 platform has been estimated to reach 75% in the Japanese population.¹⁴

Conclusions

Together with technology developments, GWASs are a promising strategy with which to identify host genetic factors for multactorial diseases, including common liver diseases, and various host genetic traits. The GWAS strategy may allow researchers to identify unexpected significant associations. Recently, new strategies and emerging technologies for

massive parallel sequencing (also termed next-generation sequencing) have allowed whole-genome analysis to identify single-nucleotide variations and structural variations (including insertion, deletion, duplication, translocation, and transposition events). The costs of using these emerging technologies are currently high; therefore, common SNP-based GWASs using the genome-wide SNP analysis technologies introduced in this paper still have an important potential role in the fields of clinical and basic research.

Conflict of interest

None

Author contributions

Genotyping and statistical analyses for hepatitis studies (NN), acquisition of genotyping data on hepatitis researches (KT), manuscript writing (NN, KT), critical review (MM).

References

- [1] Collins FS, Mansoura MK. The Human Genome Project. Revealing the shared inheritance of all humankind. *Cancer* 2001;91:221–225.
- [2] Yang J, Loos RJ, Powell JE, Medland SE, Speliotes EK, Chasman DI, *et al*. FTO genotype is associated with phenotypic variability of body mass index. *Nature* 2012;490:267–272.
- [3] Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, *et al*. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010;467:832–838.
- [4] Fujimoto A, Ohashi J, Nishida N, Miyagawa T, Morishita Y, Tsunoda T, *et al*. A replication study confirmed the *EDAR* gene to be a major contributor to population differentiation regarding head hair thickness in Asia. *Hum Genet* 2008;124:179–185.
- [5] Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, *et al*. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 2009;106:9362–9367.
- [6] Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, *et al*. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet* 2009;41:1105–1109.
- [7] Nakamura M, Nishida N, Kawashima M, Aiba Y, Tanaka A, Yasunami 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;91:721–728.
- [8] Sawai H, Nishida N, Mbarek H, Matsuda K, Mawatari Y, Yamaoka M, *et al*. No association for Chinese HBV-related hepatocellular carcinoma susceptibility SNP in other East Asian populations. *BMC Med Genet* 2012;13:47.
- [9] Holland PM, Abramson RD, Watson R, Gelfand DH. Detection of specific polymerase chain reaction product by utilizing the 5'-----3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc Natl Acad Sci USA* 1991;88:7276–7280.
- [10] Pourmand N, Elahi E, Davis RW, Ronaghi M. Multiplex pyrosequencing. *Nucleic Acids Res* 2002;30:e31.
- [11] Nishida N, Mawatari Y, Sageshima M, Tokunaga K. Highly parallel and short-acting amplification with locus-specific primers to detect single nucleotide polymorphisms by the DigiTag2 assay. *PLoS One* 2012;7:e29967.
- [12] Lindroos K, Liljedahl U, Raitio M, Syvanen AC. Minisequencing on oligonucleotide microarrays: comparison of immobilisation chemistries. *Nucleic Acids Res* 2001;29:E69–E69.
- [13] Bray MS, Boerwinkle E, Doris PA. High-throughput multiplex SNP genotyping with MALDI-TOF mass spectrometry: practice, problems and promise. *Hum Mutat* 2001;17:296–304.
- [14] Nishida N, Koike A, Tajima A, Ogasawara Y, Ishibashi Y, Uehara Y, *et al*. Evaluating the performance of Affymetrix SNP Array 6.0 platform with 400 Japanese individuals. *BMC Genomics* 2008;9:431.
- [15] Shen R, Fan JB, Campbell D, Chang W, Chen J, Doucet D, *et al*. High-throughput SNP genotyping on universal bead arrays. *Mutat Res* 2005;573:70–82.
- [16] Kennedy GC, Matsuzaki H, Dong S, Liu WM, Huang J, Liu G, *et al*. Large-scale genotyping of complex DNA. *Nat Biotechnol* 2003;21:1233–1237.

- [17] Shrimpton AE, Levinsohn EM, Yozawitz JM, Packard DS Jr, Cady RB, Middleton FA, *et al*. A *HOX* gene mutation in a family with isolated congenital vertical talus and Charcot-Marie-Tooth disease. *Am J Hum Genet* 2004;75:92–96.
- [18] Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, *et al*. Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nat Genet* 2002;32:650–654.
- [19] Kwiatkowski RW, Lyamichev V, de Arruda M, Neri B. Clinical, genetic, and pharmacogenetic applications of the Invader assay. *Mol Diagn* 1999;4:353–364.
- [20] Ohnishi Y, Tanaka T, Ozaki K, Yamada R, Suzuki H, Nakamura Y. A high-throughput SNP typing system for genome-wide association studies. *J Hum Genet* 2001;46:471–477.
- [21] Makridakis NM, Reichardt JK. Multiplex automated primer extension analysis: simultaneous genotyping of several polymorphisms. *Biotechniques* 2001;31:1374–1380.
- [22] Tang K, Fu DJ, Julien D, Braun A, Cantor CR, Köster H. Chip-based genotyping by mass spectrometry. *Proc Natl Acad Sci USA* 1999;96:10016–10020.
- [23] Ohashi J, Tokunaga K. The power of genome-wide association studies of complex disease genes: statistical limitations of indirect approaches using SNP markers. *J Hum Genet* 2001;46:478–482.
- [24] Hirschfield GM, Liu X, Xu C, Lu Y, Xie G, Lu Y, *et al*. Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. *N Engl J Med* 2009;360:2544–2555.
- [25] Hirschfield GM, Liu X, Han Y, Gorlov IP, Lu Y, Xu C, *et al*. Variants at IRF5-TNPO3, 17q12-21 and MMEL1 are associated with primary biliary cirrhosis. *Nat Genet* 2010;42:655–657.
- [26] Liu X, Invernizzi P, Lu Y, Kosoy R, Lu Y, Bianchi I, *et al*. Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. *Nat Genet* 2010;42:658–660.
- [27] Mells GF, Floyd JA, Morley KI, Cordell HJ, Franklin CS, Shin SY, *et al*. Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. *Nat Genet* 2011;43:329–332.
- [28] Selmi C, Mayo MJ, Bach N, Ishibashi H, Invernizzi P, Gish RG, *et al*. Primary biliary cirrhosis in monozygotic and dizygotic twins: genetics, epigenetics, and environment. *Gastroenterology* 2004;127:485–492.
- [29] Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, *et al*. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 2009;461:399–401.
- [30] Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, *et al*. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet* 2009;41:1100–1104.
- [31] Rauch A, Kutalik Z, Descombes P, Cai T, Di Iulio J, Mueller T, *et al*. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology* 2010;138:1338–1345, 1345.e1–e7.
- [32] Zhang H, Zhai Y, Hu Z, Wu C, Qian J, Jia W, *et al*. Genome-wide association study identifies 1p36.22 as a new susceptibility locus for hepatocellular carcinoma in chronic hepatitis B virus carriers. *Nat Genet* 2010;42:755–758.
- [33] Jiang DK, Sun J, Cao G, Liu Y, Lin D, Gao YZ, *et al*. Genetic variants in *STAT4* and *HLA-DQ* genes confer risk of hepatitis B virus-related hepatocellular carcinoma. *Nat Genet* 2013;45:72–75.
- [34] Al-Qahtani A, Al-Anazi M, Viswan NA, Khalaf N, Abdo AA, Sanai FM, *et al*. Role of single nucleotide polymorphisms of *KIF1B* gene in HBV-associated viral hepatitis. *PLoS One* 2012;7:e45128.
- [35] Kamatani Y, Wattanapokayakit S, Ochi H, Kawaguchi T, Takahashi A, Hosono N, *et al*. A genome-wide association study identifies variants in the *HLA-DP* locus associated with chronic hepatitis B in Asians. *Nat Genet* 2009;41:591–595.
- [36] Mbarek H, Ochi H, Urabe Y, Kumar V, Kubo M, Hosono N, *et al*. A genome-wide association study of chronic hepatitis B identified novel risk locus in a Japanese population. *Hum Mol Genet* 2011;20:3884–3892.
- [37] Nishida N, Sawai H, Matsuura K, Sugiyama M, Ahn SH, Park JY, *et al*. Genome-wide association study confirming association of HLA-DP with protection against chronic hepatitis B and viral clearance in Japanese and Korean. *PLoS One* 2012;7:e39175.
- [38] Guo X, Zhang Y, Li J, Ma J, Wei Z, Tan W, *et al*. Strong influence of human leukocyte antigen (HLA)-DP gene variants on development of persistent chronic hepatitis B virus carriers in the Han Chinese population. *Hepatology* 2011;53:422–428.
- [39] An P, Winkler C, Guan L, O'Brien SJ, Zeng Z. A common HLA-DPA1 variant is a major determinant of hepatitis B virus clearance in Han Chinese. *J Infect Dis* 2011;203:943–947.
- [40] Vermehren J, Lötsch J, Susser S, Wicker S, Berger A, Zeuzem S, *et al*. A common HLA-DPA1 variant is associated with hepatitis B virus infection but fails to distinguish active from inactive Caucasian carriers. *PLoS One* 2012;7:e32605.
- [41] Li J, Yang D, He Y, Wang M, Wen Z, Liu L, *et al*. Associations of HLA-DP variants with hepatitis B virus infection in southern and northern Han Chinese populations: a multicenter case-control study. *PLoS One* 2011;6:e24221.
- [42] Barrett JC, Cardon LR. Evaluating coverage of genome-wide association studies. *Nat Genet* 2006;38:659–662.