Published in final edited form as:

Nat Genet. 2013 January ; 45(1): 72–75. doi:10.1038/ng.2483.

Genetic variants in *STAT4* and HLA-DQ genes confer risk of hepatitis B virus–related hepatocellular carcinoma

A full list of authors and affiliations appears at the end of the article.

Abstract

To identify genetic susceptibility loci for hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC) in the Chinese population, we carried out a genome-wide association study (GWAS) in 2,514 chronic HBV carriers (1,161 HCC cases and 1,353 controls) followed by a 2-stage validation among 6 independent populations of chronic HBV carriers (4,319 cases and 4,966 controls). The joint analyses showed that HCC risk was significantly associated with two independent loci: rs7574865 at *STAT4*, $P_{\text{meta}} = 2.48 \times 10^{-10}$, odds ratio (OR) = 1.21; and rs9275319 at HLA-DQ, $P_{\text{meta}} = 2.72 \times 10^{-17}$, OR = 1.49. The risk allele G at rs7574865 was significantly associated with lower mRNA levels of *STAT4* in both the HCC tissues and nontumor tissues of 155 individuals with HBV-related HCC ($P_{\text{trend}} = 0.0008$ and 0.0002 , respectively). We also found significantly lower mRNA expression of *STAT4* in HCC tumor tissues compared with paired adjacent nontumor tissues ($P = 2.33 \times 10^{-14}$).

HCC is one of the most common cancers and the third most common cause of cancer mortality worldwide¹. Over 600,000 new HCC cases are diagnosed annually worldwide, among which ~55% are in China². Chronic HBV infection seems to be the most important risk factor for HCC^{1–3}. About 80% of individuals with HCC in China have a history of HBV infection⁴. Segregation analysis of familial HCC suggests interaction between HBV

© 2013 Nature America, Inc. All rights reserved.

Correspondence should be addressed to L.Y. (longyu@fudan.edu.cn) or J.X. (jxu@wfbmc.edu).

³²These authors contributed equally to this work.

³³These authors jointly directed this work.

AUTHOR CONTRIBUTIONS

L.Y. and J.X. designed and directed the study. D.-K.J. carried out overall project management. D.-K.J., X.-P.M., H.-X.H., D.W., L.-S.T., B.P., T.-T.L., P. Zhou, B.W. and H.-X.G.S.-Y. managed DNA samples. S.L.Z., D.-K.J., P. Zhang, X.-P.M. and H.-X.H. did genotyping. J.S., Z.W., D.-K.J., H.Y. and Y.Y.S. conducted data analysis. W.J., D.-K.J., X.-P.M. and H.C. did functional analyses. D.-K.J., Z.W., J.X. and L.Y. summarized results. D.-K.J., J.X. and J.S. wrote the manuscript. The following authors from the various collaborating groups carried out assembly of case-control series in their respective regions and collected data and samples: G.C., J.Y. and Hongwei Zhang in Shanghai; Y. Liu, H.S., L.L. and Z.H. in Nanjing; D.L., C.W., J.C. and K.Z. in Beijing; Y.-Z.G. in Suzhou; W.-H.R. and Y.P. in Luoyang; X.-D.L. and Y.-F.Z. in Baise; G.Z., Hongxing Zhang, Y.Z. and X.G. in Fusui; and T.-Y.C. and P.-X.L. in Qidong. The following authors collected partial control subjects in their respective regions: Z.M., Y.G. and A.T. in Guangxi Province. The following authors provided genotype data of typical Chinese population control subjects: X.-J.Z., L.-D.S., X.-B.Z., Y. Li, X.-O.S., W.Z., J.-R.L., Y.X., R.S., F.L., X.L., S.T., Q.D. and J.X. The following authors collected tissue samples of HCC cases: T.-Y.C. and N.-J.W. (major part); L.-X.Q., J.F. and Y.-K.L. (minor part). L.Y., J.X., J.L., Q.Y., D.-K.J. and H.S. obtained funding for the study.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests

Note: Supplementary information is available in the online version of the paper.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

infection and a major genetic locus^{5,6}. However, specific genetic factors predisposing to HCC are largely unknown.

Recently, a GWAS of HBV-related HCC in a Chinese population led to the identification of an HCC susceptibility locus at 1p36.22 (ref. 7). However, considering the small sample size in the initial discovery stage of that GWAS (348 cases and 359 controls), many HCC risk loci with modest effects could have been missed^{8,9}. To search for additional susceptibility genes for HCC, we conducted a multistage GWAS in 11,799 Chinese chronic HBV carriers (Supplementary Table 1).

In the initial GWAS stage, we genotyped 733,202 SNPs across the genome in 2,689 chronic HBV carriers, including 1,212 HCC cases and 1,477 non-HCC controls from Qidong, China. After a quality control analysis and filtering steps (Online Methods), 568,280 SNPs in 1,161 cases and 1,353 controls were retained for further analyses. Principal-component analysis showed that cases and controls had similar distributions of the top two eigenvectors, suggesting a common genetic background for these study subjects (Supplementary Fig. 1). Consistent with this result, we found no evidence of population stratification, with genomic control inflation factor (λ_{GC}) = 1.01 (Supplementary Fig. 2). To further increase genome coverage, we carried out an imputation analysis to infer the genotypes of additional common SNPs on the basis of HapMap Han Chinese in Beijing, China (CHB) and Japanese in Tokyo, Japan (JPT) samples (release 24; Online Methods). For all 2,462,060 imputed SNPs, we adopted the same quality control and filtering criteria as for the genotyped SNPs, and 1,672,517 imputed SNPs were retained for the GWAS analysis. Finally, we carried out an association test of HCC risk with each of these genotyped and imputed SNPs under an allelic model using PLINK (v1.07). This analysis suggested that HCC is associated with SNPs in multiple regions throughout the genome, although these associations did not reach the genome-wide significance level (we adopted a threshold of $P < 2.23 \times 10^{-8}$ for genome-wide significance, as we analyzed a total of 2.24 million genotyped and imputed SNPs in this study; Supplementary Fig. 3).

To confirm these suggestive associations, we selected 35 candidate SNPs for replication (Online Methods) in an independent case-control population of chronic HBV carriers (replication 1, 582 HCC cases and 672 controls from Suzhou, China; Supplementary Table 1). Suggestive confirmation was found for two SNPs, rs7574865 (genotyped) at 2q32.2-2q32.3 and rs9275319 (imputed) at 6p21.3, at $P < 0.05$, with the same direction of association as observed in the initial GWAS stage (Supplementary Table 2).

We then evaluated these 2 SNPs in 5 additional independent case-control populations of chronic HBV carriers, including 3,737 HCC cases and 4,294 controls (replication 2a, 1,016 cases and 924 controls from Shanghai; replication 2b, 967 cases and 1,031 controls from Nanjing, China; replication 2c, 790 cases and 616 controls from Beijing; replication 2d, 502 cases and 934 controls from Luoyang, China; and replication 2e, 462 cases and 789 controls from Fusui and Baise, China; Supplementary Table 1). Both SNPs were significantly associated with HCC risk in these five populations at $P_{\text{meta}} < 1.0 \times 10^{-7}$ (Supplementary Table 3).

When we performed a joint analysis of subjects from the GWAS stage and all subsequent replication cohorts, the statistical evidence for association between HCC risk and these two SNPs reached genome-wide significance at $P_{\text{meta}} = 1.66 \times 10^{-11}$ for rs7574865 and $P_{\text{meta}} = 8.65 \times 10^{-19}$ for rs9275319. The OR for HCC was 1.22 (95% confidence interval (CI) = 1.15–1.29) for allele G at rs7574865 and was 1.51 (95% CI = 1.38–1.66) for allele A at rs9275319. We found similar results when we adjusted for age and gender under an additive genetic model: $P_{\text{meta}} = 2.48 \times 10^{-10}$ and OR = 1.21 (95% CI = 1.14–1.28) for rs7574865; $P_{\text{meta}} = 2.72 \times 10^{-17}$ and OR = 1.49 (95% CI = 1.36–1.63) for rs9275319. We found no evidence for heterogeneity of OR for these two SNPs among these seven populations, with $P_{\text{het}} = 0.74$ and 0.12, respectively (Table 1 and Supplementary Table 4).

The association of HCC risk with the rs7574865 SNP at 2q32.2-32.3 is previously unknown. This SNP is located in the third intron of *STAT4*, a transcription factor and a critical regulator of immune responses (Supplementary Fig. 4). *STAT4* transmits signals from interleukin-12 and type I interferon (IFN- α or IFN- β) to induce IFN- γ production^{10–12}. IFN- γ is a pleiotropic cytokine that has a crucial role in host defense¹³. Impaired activation of IFN- γ via *STAT4* may reduce its antiviral and antitumor activities^{14,15}, whereas overexpression of IFN- γ may lead to the development of autoimmune diseases¹⁶. To examine the effect of genotypes at rs7574865 on *STAT4* expression, we measured the mRNA levels of *STAT4* in 155 individuals with HCC who are carriers of chronic HBV using quantitative RT-PCR. We found that subjects with the risk genotype GG at rs7574865 had the lowest mRNA level of *STAT4*, followed by subjects with the TG and TT genotypes, in both HCC tissues and nontumor tissues ($P_{\text{trend}} = 0.0008$ and 0.0002, respectively) (Fig. 1). We also observed significantly lower mRNA expression of *STAT4* in HCC tissues compared with paired adjacent nontumor tissues ($P = 2.33 \times 10^{-14}$; Fig. 1). The latter finding is also supported by data from the OncoDB.HCC database (Supplementary Fig. 5).

Recently, multiple GWAS and candidate gene studies have reported associations of *STAT4* variants, particularly the rs7574865 SNP, with various autoimmune diseases^{17–26}. Notably, in contrast to our finding that the G allele at rs7574865 is associated with greater risk for HCC, this allele was associated with lower risk for autoimmune diseases (Supplementary Figs. 6 and 7 and Supplementary Table 5). These results support the dual roles of *STAT4* in autoimmune diseases and HBV-related HCC.

The association of HCC risk with rs9275319 at 6p21.3 was not reported in earlier GWASs of HCC. In addition to the strong statistical evidence, this association was supported by two other observations. First, this SNP is located between *HLA-DQB1* and *HLA-DQA2*, which are biologically relevant in HCC (Supplementary Fig. 8). HLA-DQ proteins make up the human leukocyte antigen (HLA) class II complex, an α - β heterodimeric membrane glycoprotein that is expressed on the surface of antigen-presenting cells, such as macrophages, dendritic cells and B lymphocytes²⁷. HLA-DQ genes have a central role in immune-mediated diseases, including liver diseases and cancer^{28,29}. SNPs in these genes have been implicated in risk for primary biliary cirrhosis^{30,31}, HBV³² and hepatitis C virus (HCV)-related HCC³³ from several GWASs (Supplementary Table 6). SNPs in *HLA-DQB1* were also associated with HBV and HCC risk in a candidate gene approach^{34,35}. Second, the observed association is likely not confounded by population stratification, a common

concern for SNPs in the HLA region owing to highly variable allele frequencies for some SNPs among and within populations. The allele frequency of the rs9275319 SNP in controls was similar across all the populations in this study (0.87–0.90 for allele A) and was consistently lower than that in cases within each population (0.91–0.93 for allele A), except for the Guangxi population (Table 1).

To investigate whether the association of HCC risk with rs9275319 indirectly reflects association of established alleles in HLA-DQ genes, we analyzed *HLA-DQB1* and *HLA-DQA1* genotypes in 936 HBV carriers, including 478 HCC cases and 458 controls (a subset of the Qidong cohort samples used in the GWAS stage) using the sequence-based typing (SBT) method. Of the 33 alleles of the 2 HLA genes, 2 alleles (*HLA-DQB1**0401 and *HLA-DQA1**0303) were in moderate linkage disequilibrium (LD) with rs9275319 ($r^2 = 0.48$ and 0.47 , respectively) and were significantly associated with HBV-related HCC risk ($P = 4.51 \times 10^{-4}$ and 4.44×10^{-4} , respectively). However, the strength of association was weaker than that of rs9275319 in this subset ($P = 5.82 \times 10^{-6}$). Furthermore, when the rs9275319 SNP was included in the regression model, *HLA-DQB1**0401 and *HLA-DQA1**0303 were no longer significantly associated with HCC risk, suggesting that these two existing HLA alleles are probably not causal variants for HCC risk (Supplementary Table 7). Additional studies are needed to dissect the causal association at this region.

Considering that HBV genotype (the major HBV genotypes are B and C in the Chinese population) and HBV viral load may affect the associations of HCC risk with the two SNPs, we measured the HBV genotype and HBV viral load in a subset of the Shanghai cohort and carried out association analyses of HCC risk with rs7574865 and rs9275319 by stratifying HBV genotype status or viral load. The ORs for HCC at each SNP were not significantly different between HBV genotype groups ($P_{\text{het}} = 0.77$ for rs7574865 and 0.48 for rs9275319; Supplementary Table 8). Similarly, the ORs for HCC at each SNP were not significantly different between the two HBV viral groups (HBV DNA < 2,000 and ≥ 2,000 copies/ml; $P_{\text{het}} = 0.54$ for rs7574865 and 0.67 for rs9275319; Supplementary Table 8).

To examine whether the two HCC risk-associated SNPs (rs7574865 at *STAT4* and rs9275319 at HLA-DQ) are also associated with risk for chronic HBV infection, we compared 3,935 chronic HBV carriers without HCC (used in the replication 1 and 2 stages of our HCC study, except for the Nanjing population) with 4,585 healthy subjects who did not carry HBV (recruited in geographical areas corresponding to those of the HBV carriers; Supplementary Table 9). We found that the A allele at rs9275319 in HLA-DQ had consistently higher frequency in chronic HBV carriers in each population, and the overall evidence among all populations was significant (OR = 1.29, $P_{\text{meta}} = 3.54 \times 10^{-6}$; Supplementary Table 10). This result suggested that this SNP is also associated with susceptibility to chronic HBV infection. For rs7574865 in *STAT4*, we found similar results. There was a trend of higher frequency for the G allele in chronic HBV carriers in each population. However, the overall statistical evidence was weak (OR = 1.12, $P_{\text{meta}} = 0.003$; Supplementary Table 10).

A recent GWAS on HBV-related HCC in Chinese individuals reported a strong association with rs17401966 at 1p36.22 (ref. 7). In our study, we found no evidence of association

between rs17401966 and risk for HCC in the first-stage samples of the GWAS (OR = 0.98, $P = 0.79$; Supplementary Table 11). Similarly, a lack of association at this locus has been reported in a recent study of the eastern Chinese population (OR = 1.01, $P = 0.92$)³⁵. Additional studies are needed to assess the association between this SNP and HCC risk^{8,9}.

Finally, we examined associations of HCC risk with SNPs that have been implicated in other HBV-related diseases and HCV-induced HCC. We found no evidence of association for any of these SNPs in our GWAS stage (Supplementary Table 11). Our study had >90% statistical power to detect associations for SNPs with risk allele frequency of >0.15 and with OR of >1.4 (Supplementary Table 12).

In summary, we conducted a three-stage GWAS in seven case-control populations for HBV-related HCC among the Chinese population. We identified two previously unknown HCC susceptibility loci at *STAT4* and HLA-DQ. Additional studies are warranted to further confirm the findings.

URLs

PLINK, <http://pngu.mgh.harvard.edu/~purcell/plink/>; EIGENSTRAT, <http://genepath.med.harvard.edu/~reich/Software.htm>; IMPUTE, <http://mathgen.stats.ox.ac.uk/impute/impute.html>; the International HapMap Project, <http://hapmap.ncbi.nlm.nih.gov/>; R statistical environment, <http://www.cran.r-project.org/>; LocusZoom, <http://csg.sph.umich.edu/locuszoom/>; OncoDB.HCC database, <http://oncodb.hcc.ibms.sinica.edu.tw/index.htm>.

ONLINE METHODS

Study subjects

The 2,689 chronic HBV carriers, including 1,212 HCC cases and 1,477 controls (Supplementary Table 1), used in the initial GWAS were recruited by the Qidong Liver Cancer Institute (Supplementary Fig. 9). The surrounding area has the highest incidence of HCC in China³⁶. Chronic HBV carriers were defined as individuals that 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 findings on cytological or pathological examination and/or (ii) positive images on angiogram, ultra-sonography, computed tomography and/or magnetic resonance imaging, combined with an α -fetoprotein concentration of ≥ 400 ng/ml. We confirmed that none of the individuals with HCC had other cancers through an initial screening. All controls had, by self-report, no history of HCC or other cancers.

In the validation study, 6 independent sample sets totaling 9,285 chronic HBV carriers, including 4,319 individuals with HCC and 4,966 controls (Supplementary Table 1), were enrolled. These samples were recruited from eastern China (Suzhou, replication 1; Nanjing, replication 2; Shanghai, replication 2), northern China (Beijing, replication 2), central China (Luoyang, replication 2) and southern China (Fusui and Baise, replication 2) (Supplementary Fig. 9). To examine whether the HCC risk-associated SNPs identified in our study were also associated with risk for chronic hepatitis B, 4,585 healthy controls

without a history of HBV infection were also recruited from eastern China (Suzhou and Shanghai), northern China (Beijing), central China (Luoyang) and southern China (Fusui and Baise) (Supplementary Fig. 9 and Supplementary Table 9). All cases and controls were recruited using the same criteria as was used for the samples in the GWAS stage and have been described in detail^{7,35,37–39}. All subjects did not have antibodies to HCV or HIV and did not have other types of liver disease, such as autoimmune hepatitis, toxic hepatitis and primary biliary cirrhosis. Informed consent was obtained from all subjects before their participation in the study. The study was approved by the ethical committees of all institutions involved in the study and conducted in accordance with Declaration of Helsinki principles.

Genotyping and quality control in GWAS

Genome-wide genotyping analysis was conducted using Illumina Human OmniExpress BeadChips at Genergy Biotech. Genotyping was performed according to the Infinium HD protocol from Illumina. SNPs were excluded if they were not mapped to autosomal chromosomes or showed either a call rate of <95% in cases or controls, a minor allele frequency of <1% in the population or significant deviation from Hardy-Weinberg equilibrium (HWE) both in the cases and controls ($P = 0.001$; Supplementary Fig. 10). The most common cause of SNP exclusion was low allele frequency (minor allele frequency of <0.01), followed by presence on the X chromosome. Some 1,212 HBV-related HCC cases and 1,477 controls were genotyped initially. Seven samples with genotyping call rate of <95% were removed (two cases and five controls). We then examined potential genetic relatedness on the basis of pairwise identity by state for all of the successfully genotyped samples, using PLINK (v1.07) software. In total, 157 samples were removed owing to sample duplications and genetic relatedness (44 cases and 113 controls). In addition, 11 population outliers (5 cases and 6 controls) were excluded in principal-component analysis using the software package EIGENSTRAT. After quality control and filtering processes, genotype data for 568,280 SNPs in 1,161 cases and 1,353 controls were used for the GWAS analysis.

Imputation

We imputed all known SNPs across the genome on the basis of the genotyped SNPs in the 1,161 cases and 1,353 controls used in the initial GWAS and haplotype data from HapMap Phase 2 CHB and JPT subjects (release 24) using the IMPUTE program (v2.0). A posterior probability of 0.90 was used as a threshold to call genotypes. The quality control and filtering criteria for all 2,462,060 imputed SNPs were the same as for the genotyped SNPs, and 1,672,517 imputed SNPs were retained for the GWAS analysis (Supplementary Fig. 10). Many of the imputed SNPs were removed from further analyses because of low call rate and deviation from HWE, probably owing to low LD between genotyped and imputed SNPs.

SNP selection for replication stages

First, we chose SNPs on the basis of a CLUMP analysis⁴⁰ requiring $P < 5.0 \times 10^{-3}$ in the association test in the initial GWAS and LD r^2 of 0.2 between markers. SNPs were further

filtered by comparing the allele frequency of an additional 6,450 Chinese samples. These samples were recruited to represent the typical Chinese population and have been genotyped in a series of GWAS, which have been described^{41–43}. SNPs were removed from further analysis if the allele frequency of the controls used in the current study differed from the allele frequency of the additional 6,450 Chinese samples by 0.02. However, these 6,450 samples were not included in the association analysis of the validation study or in the final joint association analysis of the combined samples. In total, 35 SNPs were selected for the validation analysis (Supplementary Table 2).

Genotyping and quality control in replication studies

Genotyping analyses for replications 1 and 2 were conducted using the Sequenom MassArray system at the Fudan-VARI laboratory and TaqMan assays (Applied Biosystems) at the State Key Laboratory of Genetic Engineering, Fudan University, according to the manufacturers' instructions. Duplicate test samples and two water samples (PCR negative controls) were included in each 96-well plate, with the technician blinded to the identity of the samples. The rate of concordant results in 200 duplicate samples was >99%. All 35 SNPs in replication 1 and 2 SNPs in replication 2 satisfied the quality control criteria, with a call rate of >95% in both cases and controls and with no evidence of departure from HWE in the controls ($P = 0.001$).

Quantitative RT-PCR

Tumor tissue and paired adjacent nontumor liver tissue samples were collected for gene expression analysis from 155 subjects with HCC, who were randomly selected from the HCC cases used in the initial GWAS. The genotypes at rs7574865 in these 155 individuals were determined in the initial GWAS as described (78 GG, 63 GT and 14 TT genotypes). RNA templates were extracted from tissue samples using the RNeasy miniprep kit (Qiagen). cDNA was synthesized from 1 µg of RNA templates using reverse transcriptase and oligo(dT) primer (Promega). The sequences of the primers used for amplification are given in Supplementary Table 13. Quantitative RT-PCR for *STAT4* was carried out in triplicate using SYBR Green on a Roche LightCycler 480 II Real-Time PCR System. Expression of all samples was normalized to the relative expression level of $\beta 2$ microglobulin using the C_T method.

HLA-DQB1 and HLA-DQA1 genotyping

We analyzed HLA-DQ genotypes in 936 HBV carriers, including 478 HCC cases and 458 controls (a subset of the Qidong cohort samples used in the GWAS stage). The SBT method⁴⁴ (BGI *HLA-DQB1* and *DQA1* SBT kit) was used to assay the genotypes of the *HLA-DQB1* and *HLA-DQA1* genes at BGI, Shenzhen, China. *HLA-DQB1* and *HLA-DQA1* alleles were determined using the alignment database of dbMHC.

HBV genotyping and serum viral load detection

We carried out HBV genotyping and serum viral load detection in the Shanghai cohort samples. HBV genotype determination has been described⁴⁵. Briefly, HBV DNA was extracted from 200 µl of HBsAg-positive serum using High Pure Viral Nucleic Acid kits

(Roche) according to the manufacturer's instructions. HBV genotype was determined using a multiplex PCR assay. HBV genotypes of samples with low amounts of HBV DNA were identified by nested multiplex PCR. Viral load was measured with the LightCycler 480 (Roche) using quantitative HBV PCR fluorescence diagnostic kits (Fosun Diagnostics).

Statistical analysis

The association of SNPs with HCC risk in the GWAS and replication stages was carried out with an allelic model using PLINK (v1.07). In the GWAS, population structure was evaluated by principal-component analysis in the EIGENSTRAT software package. The quantile-quantile plot and Manhattan plot were generated using the R software package (v2.10.1). The regional plot was plotted using an online tool, LocusZoom 1.1 (see URLs). For the two identified SNPs and HLA alleles, the ORs and 95% CIs were calculated using a logistic regression model, assuming an additive model and adjusting for gender and age. Heterogeneity across studies was examined using the χ^2 -based Cochran's Q statistic. Meta-analysis was performed using the fixed-effects model (Mantel-Haenszel model) if the result of the heterogeneity test showed $P_{\text{het}} \geq 0.05$ or using the random-effects model (DerSimonian and Laird model) if $P_{\text{het}} < 0.05$. Differences in expression between tumors and paired adjacent nontumor samples were assessed by a Wilcoxon signed rank test, and differences between the rs7574865 GG, GT and TT genotypes were assessed by a nonparametric trend test. $P < 0.05$ was used as the criterion for significance, and all statistical tests were two sided.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

De-Ke Jiang^{1,32}, Jielin Sun^{2,32}, Guangwen Cao^{3,32}, Yao Liu^{4,5,6,32}, Dongxin Lin^{7,32}, Yu-Zhen Gao⁸, Wei-Hua Ren^{1,9}, Xi-Dai Long¹⁰, Hongxing Zhang¹¹, Xiao-Pin Ma¹, Zhong Wang², Wei Jiang¹, Tao-Yang Chen¹², Yong Gao^{13,14,15}, Liang-Dan Sun^{16,17,18}, Ji-Rong Long¹⁹, Hui-Xing Huang¹, Dan Wang¹, Hongjie Yu¹⁵, Pengyin Zhang¹⁵, Li-Sha Tang¹, Bo Peng¹, Hao Cai¹, Ting-Ting Liu¹, Ping Zhou¹, Fang Liu¹⁵, Xiaoling Lin¹⁵, Sha Tao², Bo Wan¹, He-Xi Ge Sai-Yin¹, Lun-Xiu Qin^{20,21}, Jianhua Yin³, Li Liu²², Chen Wu⁷, Yan Pei⁹, Yuan-Feng Zhou¹⁰, Yun Zhai¹¹, Pei-Xin Lu¹², Aihua Tan^{13,14}, Xian-Bo Zuo^{16,17,18}, Jia Fan^{20,21}, Jiang Chang⁷, Xiaoli Gu¹¹, Neng-Jin Wang¹², Yang Li^{16,17,18}, Yin-Kun Liu^{20,21}, Kan Zhai⁷, Hongwei Zhang³, Zhibin Hu^{4,5,6}, Jun Liu^{1,23}, Qing Yi^{1,24,25,26}, Yongbing Xiang²⁷, Rong Shi²⁸, Qiang Ding²⁹, Wei Zheng¹⁹, Xiao-Ou Shu¹⁹, Zengnan Mo^{13,14}, Yin Yao Shugart³⁰, Xue-Jun Zhang^{16,17,18}, Gangqiao Zhou¹¹, Hongbing Shen^{4,5,6}, S Lilly Zheng², Jianfeng Xu^{2,15,29,33}, and Long Yu^{1,31,33}

Affiliations

¹State Key Laboratory of Genetic Engineering, Institute of Genetics, School of Life Sciences, Fudan University, Shanghai, China

- ²Center for Cancer Genomics, Wake Forest University School of Medicine, Winston-Salem, North Carolina, USA
- ³Department of Epidemiology, Second Military Medical University, Shanghai, China
- ⁴Ministry of Education Key Laboratory of Modern Toxicology, School of Public Health, Nanjing Medical University, Nanjing, China
- ⁵Jiangsu Key Laboratory of Cancer Biomarkers, Prevention and Treatment, School of Public Health, Nanjing Medical University, Nanjing, China
- ⁶State Key Laboratory of Reproductive Medicine, School of Public Health, Nanjing Medical University, Nanjing, China
- ⁷Department of Etiology and Carcinogenesis, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
- ⁸Department of Forensic Medicine, Medical College of Soochow University, Suzhou, China
- ⁹Department of Surgery, The First Affiliated Hospital, Henan University of Science and Technology, Luoyang, China
- ¹⁰Department of Pathology, Youjiang Medical College for Nationalities, Baise, China
- ¹¹State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Radiation Medicine, Beijing, China
- ¹²Qidong Liver Cancer Institute, Qidong, China
- ¹³Institute of Urology and Nephrology, The First Affiliated Hospital of Guangxi Medical University, Guangxi Medical University, Nanning, China
- ¹⁴Center for Genomic and Personalized Medicine, Guangxi Medical University, Nanning, China
- ¹⁵Fudan-VARI Center for Genetic Epidemiology, Fudan University, Shanghai, China
- ¹⁶Institute of Dermatology, Anhui Medical University, Hefei, China
- ¹⁷Department of Dermatology at No. 1 Hospital, Anhui Medical University, Hefei, China
- ¹⁸State Key Laboratory Incubation Base of Dermatology, Ministry of National Science and Technology, Hefei, China
- ¹⁹Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee, USA
- ²⁰Liver Cancer Institute, Fudan University, Shanghai, China
- ²¹Department of Liver Surgery, Zhongshan Hospital, Fudan University, Shanghai, China

²²Department of Hepatobiliary Surgery, Nantong Tumor Hospital, Nantong, China

²³Department of Pharmacology and Molecular Sciences, The Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

²⁴Department of Lymphoma/Myeloma, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

²⁵Division of Cancer Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

²⁶Center for Cancer Immunology Research, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

²⁷Shanghai Cancer Institute, Shanghai, China

²⁸School of Public Health, Shanghai Jiaotong University, Shanghai, China

²⁹Fudan Institute of Urology, Huashan Hospital, Fudan University, Shanghai, China

³⁰Unit of Statistical Genomics, Division of Intramural Research Program, National Institute of Mental Health, US National Institutes of Health, Bethesda, Maryland, USA

³¹Institute of Biomedical Science, Fudan University, Shanghai, China

Acknowledgments

We thank all the subjects who participated in this study. We also thank Y. Shi and L. He from Bio-X Institutes of Shanghai Jiao Tong University for helping with data analysis of typical Chinese population control subjects, and L. Jin, Y. Zhong, H.-H. Wang and W.-Y. Yan from the School of Life Sciences of Fudan University and H.-B. Shu from College of Life Sciences of Wuhan University for suggestions on this study. The study is supported by the National Natural Science Foundation of China for Creative Research Groups (30024001 to L.Y.), the National Key Sci-Tech Special Project of China (2013ZX10002010 and 2008ZX10002-020 to L.Y.), the Project of the Shanghai Municipal Science and Technology Commission (to L.Y.), the National Natural Science Foundation of China (31071193 to L.Y. and 31100895 to D.-K.J.), Director Foundation of the State Key Laboratory of Genetic Engineering (to L.Y.), the Fund from X.-J. Qu and Chang-An Capital of Beijing (to L.Y.), the Fund from S.-J. Yan and Tianying S.T. Group of Jiangsu (to L.Y.), MOE Foundation of Cheung Kong Scholars Program (to J.L. and Q.Y.), an intramural research grant from Huashan Hospital, Fudan University (to J.X.) and an intramural research grant from Fudan-VARI Center for Genetic Epidemiology, Fudan University (to J.X.). This work is partly funded by the National Natural Science Foundation of China (30800946 to H.S.), Foundation for the Author of National Excellent Doctoral Dissertation (201081 to H.S.), the Program for New Century Excellent Talents in University (NCET-10-0178 to H.S.), and the Priority Academic Program Development of Jiangsu Higher Education Institutions (to H.S.).

References

1. Yang JD, Roberts LR. Hepatocellular carcinoma: a global view. *Nat Rev Gastroenterol Hepatol*. 2010; 7:448–458. [PubMed: 20628345]
2. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin*. 2005; 55:74–108. [PubMed: 15761078]
3. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology*. 2007; 132:2557–2576. [PubMed: 17570226]
4. Tanaka M, et al. Hepatitis B and C virus infection and hepatocellular carcinoma in china: a review of epidemiology and control measures. *J Epidemiol*. 2011; 21:401–416. [PubMed: 22041528]
5. Shen FM, Lee MK, Gong HM, Cai XQ, King MC. Complex segregation analysis of primary hepatocellular carcinoma in Chinese families: interaction of inherited susceptibility and hepatitis B viral infection. *Am J Hum Genet*. 1991; 49:88–93. [PubMed: 1648308]

6. Yu MW, et al. Familial risk of hepatocellular carcinoma among chronic hepatitis B carriers and their relatives. *J Natl Cancer Inst.* 2000; 92:1159–1164. [PubMed: 10904089]
7. Zhang H, 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. [PubMed: 20676096]
8. Casper M, Grunhage F, Lammert F. Cancer risk in chronic hepatitis B: do genome-wide association studies hit the mark? *Hepatology.* 2011; 53:1390–1392. [PubMed: 21480342]
9. Budhu A, Wang XW. Power play: scoring our goals for liver cancer with better GWAS study design. *J Hepatol.* 2011; 54:823–824. [PubMed: 21167853]
10. Thierfelder WE, et al. Requirement for Stat4 in interleukin-12-mediated responses of natural killer and T cells. *Nature.* 1996; 382:171–174. [PubMed: 8700208]
11. Nguyen KB, et al. Critical role for STAT4 activation by type 1 interferons in the interferon- γ response to viral infection. *Science.* 2002; 297:2063–2066. [PubMed: 12242445]
12. Yu H, Pardoll D, Jove R. STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat Rev Cancer.* 2009; 9:798–809. [PubMed: 19851315]
13. Dunn GP, Koebel CM, Schreiber RD. Interferons, immunity and cancer immunoediting. *Nat Rev Immunol.* 2006; 6:836–848. [PubMed: 17063185]
14. Saha B, Jyothi Prasanna S, Chandrasekar B, Nandi D. Gene modulation and immunoregulatory roles of interferon γ . *Cytokine.* 2010; 50:1–14. [PubMed: 20036577]
15. Horras CJ, Lamb CL, Mitchell KA. Regulation of hepatocyte fate by interferon- γ . *Cytokine Growth Factor Rev.* 2011; 22:35–43. [PubMed: 21334249]
16. Schwarting A, et al. IL-12 drives IFN- γ -dependent autoimmune kidney disease in MRL-*Fas*^{lpr} mice. *J Immunol.* 1999; 163:6884–6891. [PubMed: 10586090]
17. Remmers EF, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N Engl J Med.* 2007; 357:977–986. [PubMed: 17804842]
18. Hom G, et al. Association of systemic lupus erythematosus with *C8orf13-BLK* and *ITGAM-ITGAX*. *N Engl J Med.* 2008; 358:900–909.
19. Harley JB, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in *ITGAM*, *PXK*, *KIAA1542* and other loci. *Nat Genet.* 2008; 40:204–210. [PubMed: 18204446]
20. Han JW, et al. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat Genet.* 2009; 41:1234–1237. [PubMed: 19838193]
21. Yang W, et al. Genome-wide association study in Asian populations identifies variants in *ETS1* and *WDFY4* associated with systemic lupus erythematosus. *PLoS Genet.* 2010; 6:e1000841. [PubMed: 20169177]
22. Chung SA, et al. Differential genetic associations for systemic lupus erythematosus based on anti-dsDNA autoantibody production. *PLoS Genet.* 2011; 7:e1001323. [PubMed: 21408207]
23. Kochi Y, et al. A regulatory variant in *CCR6* is associated with rheumatoid arthritis susceptibility. *Nat Genet.* 2010; 42:515–519. [PubMed: 20453841]
24. Stahl EA, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet.* 2010; 42:508–514. [PubMed: 20453842]
25. Allanore Y, et al. Genome-wide scan identifies *TNIP1*, *PSORS1C1*, and *RHOB* as novel risk loci for systemic sclerosis. *PLoS Genet.* 2011; 7:e1002091. [PubMed: 21750679]
26. Korman BD, Kastner DL, Gregersen PK, Remmers EF. STAT4: genetics, mechanisms, and implications for autoimmunity. *Curr Allergy Asthma Rep.* 2008; 8:398–403. [PubMed: 18682104]
27. Choi NM, Majumder P, Boss JM. Regulation of major histocompatibility complex class II genes. *Curr Opin Immunol.* 2011; 23:81–87. [PubMed: 20970972]
28. Handunnethi L, Ramagopalan SV, Ebers GC, Knight JC. Regulation of major histocompatibility complex class II gene expression, genetic variation and disease. *Genes Immun.* 2010; 11:99–112. [PubMed: 19890353]

29. Liu C, Cheng B. Association of polymorphisms of human leucocyte antigen-DQA1 and DQB1 alleles with chronic hepatitis B virus infection, liver cirrhosis and hepatocellular carcinoma in Chinese. *Int J Immunogenet.* 2007; 34:373–378. [PubMed: 17845309]
30. Hirschfield GM, et al. Primary biliary cirrhosis associated with *HLA*, *IL12A*, and *IL12RB2* variants. *N Engl J Med.* 2009; 360:2544–2555. [PubMed: 19458352]
31. Liu X, et al. Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. *Nat Genet.* 2010; 42:658–660. [PubMed: 20639880]
32. Mbarek H, 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. [PubMed: 21750111]
33. Kumar V, et al. Genome-wide association study identifies a susceptibility locus for HCV-induced hepatocellular carcinoma. *Nat Genet.* 2011; 43:455–458. [PubMed: 21499248]
34. Xin YN, et al. Specific *HLA-DQB1* alleles associated with risk for development of hepatocellular carcinoma: a meta-analysis. *World J Gastroenterol.* 2011; 17:2248–2254. [PubMed: 21633537]
35. Hu L, et al. Genetic variants in human leukocyte antigen/DP-DQ influence both hepatitis B virus clearance and hepatocellular carcinoma development. *Hepatology.* 2012; 55:1426–1431. [PubMed: 22105689]
36. Chen JG, Zhang SW. Liver cancer epidemic in China: past, present and future. *Semin Cancer Biol.* 2011; 21:59–69. [PubMed: 21144900]
37. He Y, et al. IkB α gene promoter polymorphisms are associated with hepatocarcinogenesis in patients infected with hepatitis B virus genotype C. *Carcinogenesis.* 2009; 30:1916–1922. [PubMed: 19797428]
38. Long XD, Ma Y, Zhou YF, Ma AM, Fu GH. Polymorphism in xeroderma pigmentosum complementation group C codon 939 and aflatoxin B1-related hepatocellular carcinoma in the Guangxi population. *Hepatology.* 2010; 52:1301–1309. [PubMed: 20658464]
39. Gao Y, et al. An insertion/deletion polymorphism at miRNA-122-binding site in the interleukin-1 α 3' untranslated region confers risk for hepatocellular carcinoma. *Carcinogenesis.* 2009; 30:2064–2069. [PubMed: 19917630]
40. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007; 81:559–575. [PubMed: 17701901]
41. Quan C, et al. Genome-wide association study for vitiligo identifies susceptibility loci at 6q27 and the MHC. *Nat Genet.* 2010; 42:614–618. [PubMed: 20526339]
42. Zheng W, et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat Genet.* 2009; 41:324–328. [PubMed: 19219042]
43. Long J, et al. Identification of a functional genetic variant at 16q12.1 for breast cancer risk: results from the Asia Breast Cancer Consortium. *PLoS Genet.* 2010; 6:e1001002. [PubMed: 20585626]
44. Li J, et al. Characterization of 236 novel alleles at the HLA-A, -B, -C, -DRB1, -DQB1 and -DPB1 loci from China Marrow Donor Program. *Tissue Antigens.* 2011; 78:267–270. [PubMed: 21732917]
45. Chen J, et al. Improved multiplex-PCR to identify hepatitis B virus genotypes A–F and subgenotypes B1, B2, C1 and C2. *J Clin Virol.* 2007; 38:238–243. [PubMed: 17224304]

Figure 1.

STAT4 relative expression levels according to rs7574865 genotypes in paired tumor tissues and adjacent nontumor tissues from 155 subjects with HCC who are chronic HBV carriers by quantitative RT-PCR. Lines indicate the median with interquartile range (25% and 75%). *P* values were calculated by Wilcoxon rank test.

Table 1

Results of GWAs, replication and joint analysis

Chr.	Gene	SNP	Allele ^a	Study	Population ^b	RAF		OR _{add} (95% CI) ^{c,e}	P _{add} ^e	P _{het} ^f
						Cases	Controls			
2	STAT4	rs7574865	T/G	GWAS	Qidong	0.72	0.67	1.25 (1.11–1.41)	3.21 × 10 ⁻⁴	1.53 × 10 ⁻⁴
				Replication 1	Suzhou	0.71	0.67	1.19 (1.00–1.42)	4.69 × 10 ⁻²	5.19 × 10 ⁻²
				Replication 2a	Shanghai	0.72	0.68	1.23 (1.07–1.41)	3.93 × 10 ⁻³	1.06 × 10 ⁻²
				Replication 2b	Nanjing	0.71	0.68	1.17 (1.02–1.34)	2.45 × 10 ⁻²	1.96 × 10 ⁻²
				Replication 2c	Beijing	0.71	0.68	1.14 (0.97–1.34)	1.23 × 10 ⁻¹	1.50 × 10 ⁻¹
				Replication 2d	Henan	0.72	0.67	1.23 (1.03–1.46)	2.07 × 10 ⁻²	2.74 × 10 ⁻¹
				Replication 2e	Guangxi	0.74	0.68	1.33 (1.11–1.60)	1.88 × 10 ⁻³	1.64 × 10 ⁻³
6	HLA-DQ	rs9275319	G/A	Meta-analysis ^g				1.22 (1.15–1.29)	1.66 × 10 ⁻¹¹	2.48 × 10 ⁻¹⁰
				GWAS	Qidong	0.92	0.88	1.58 (1.31–1.92)	2.27 × 10 ⁻⁶	1.11 × 10 ⁻⁶
				Replication 1	Suzhou	0.92	0.89	1.53 (1.16–2.03)	2.84 × 10 ⁻³	4.73 × 10 ⁻³
				Replication 2a	Shanghai	0.92	0.88	1.45 (1.17–1.80)	6.89 × 10 ⁻⁴	2.05 × 10 ⁻³
				Replication 2b	Nanjing	0.93	0.87	1.84 (1.49–2.28)	1.61 × 10 ⁻⁸	2.42 × 10 ⁻⁸
				Replication 2c	Beijing	0.93	0.90	1.49 (1.14–1.95)	3.32 × 10 ⁻³	1.35 × 10 ⁻²
				Replication 2d	Henan	0.91	0.89	1.33 (1.01–1.73)	4.28 × 10 ⁻²	1.06 × 10 ⁻¹
				Replication 2e	Guangxi	0.94	0.93	1.10 (0.79–1.53)	5.74 × 10 ⁻¹	6.59 × 10 ⁻¹
				Meta-analysis ^g				1.51 (1.38–1.66)	8.65 × 10 ⁻¹⁹	2.72 × 10 ⁻¹⁷

Chr., chromosome; RAF, risk allele frequency (G allele for rs7574865; A allele for rs9275319); all, allelic model; add, additive model; het, heterogeneity.

^aMinor allele/major allele.

^bSamples used in our study were recruited in eastern China (Qidong, Suzhou, Shanghai and Nanjing), northern China (Beijing), central China (Henan) and southern China (Guangxi).

^cORs and 95% CIs were calculated by considering the nonrisk allele (allele frequency of the cases is lower than that of the controls in the initial GWAS stage) as a reference.

^dData were analyzed under an allelic model.

^eData were analyzed under an additive model and adjusted for gender and age.

^fResults of heterogeneity tests.

^gMeta-analysis was carried out using the fixed-effects model, as no significant heterogeneity across the study was observed for both SNPs.