ORIGINAL RESEARCH

A Genetic Variant of PPP1CB Influences Risk of Hepatitis B Virus-Related Hepatocellular Carcinoma in Han Chinese: A Pathway Based Analysis

Haoming Mai*
Haisheng Xie*
Jia Hou (iD)*
Haitao Chen
Bin Zhou
Jinlin Hou
Deke Jiang

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, Guangdong, 510515, People's Republic of China

*These authors contributed equally to this work

Purpose: Activation of actin cytoskeleton remodeling is an important stage preceding cancer cell metastasis. Previous genome-wide association studies (GWAS) have identified multiple hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC)-associated risk loci. However, limited sample size or strict significance threshold of GWAS may cause HBV-related HCC risk-associated genetic loci to be undetected. We aimed to investigate the performance of the SNP rs13025377 in *PPP1CB* in HCC.

Patients and Methods: We performed a case–control study including 1161 cases and 1353 controls to evaluate associations between single nucleotide polymorphisms (SNPs) from 98 actin-cytoskeleton regulatory genes and risk of HBV-related HCC. The effects of SNPs on HBV-related HCC risk were assessed under logistic regression model and corrected by false discovery rate (FDR).

Results: We found that rs13025377 in *PPP1CB* was significantly associated with HBV-related HCC risk [odds ratio (OR) = 0.81, 95% confidence interval (CI) = 0.72~0.91, P = 4.88×10^{-4}]. The risk allele A of rs13025377 increased *PPP1CB* expression levels in normal liver tissue. SNP rs4665434 was tagged by rs13025377 (r^2 = 0.9) and its protective allele disrupted *CTCF* and *Cohesin* motifs. According to public datasets, *PPP1CB, CTCF* and *Cohesin* expression levels are increased in tumor tissues. Kaplan–Meier plots demonstrated that higher *PPP1CB* expression was significantly associated with shorter overall survival (OS). Moreover, we observed strong correlation between *CTCF, Cohesin*, and *PPP1CB* in various liver tissues. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis confirmed that *PPP1CB* plays a role in HCC through actin-cytoskeleton regulation.

Conclusion: Thus, these findings indicated that *PPP1CB* may be a key gene in actin-cytoskeleton regulation and rs13025377 contributes to the risk of HBV-related HCC by regulating *PPP1CB* expression.

Keywords: genome-wide association studies, PPP1CB, CTCF, cohesin, SNP, HBV-related hepatocellular carcinoma

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, Guangdong, 510515, People's Republic of China
Tel +86-20-62786533; +86-20-61641941
Email dekejiang17@smu.edu.cn; jlhousmu@163.com

Introduction

Liver cancer is the sixth most common malignancy worldwide. In 2020, there were about 906,000 new cases of liver cancer. Liver cancer is also the third leading cause of cancer death and resulted in approximately 830,000 deaths in 2020.[1] Hepatocellular carcinoma (HCC) comprises 75–85% of primary liver cancers and Hepatitis B virus (HBV) is the top determinant for HCC in China.[1,2] Due to the lack of diagnostic and therapeutic methods, many HBV-related HCC cases are diagnosed at an advanced stage with poor prognosis.[3] In the last decade, eight HBV-related HCC-associated

genetic loci have been identified by genome-wide association studies (GWAS).[4–8] However, genetic risk loci may have been missed because of limited sample size and strict significance threshold. Pathway-based deep mining of GWAS data is an effective strategy to detect missing disease-related genetic loci. This strategy has been successfully implemented in studies of various diseases including colorectal cancer, gastric cancer, and breast cancer.[9–11]

Cancer cell metastasis is a multi-step biological process that is driven by dynamic reorganization of the actin cytoskeleton.[12] Ample evidence supports the etiological role of dysregulation of actin-cytoskeleton regulation in tumorigenesis.[13] The actin cytoskeleton consists of the polymerization of actin monomers (G-actin) which form microfilaments (F-actin), and actin-cytoskeleton remodeling is crucial for epithelial-mesenchymal transition (EMT) in HCC.[14] Upon binding to cell surface receptors, signaling molecules stimulate intracellular signaling pathways to remodel the actin cytoskeleton.[13]

Previous studies have reported that actin-cytoskeleton regulatory genes might contribute to tumorigenesis in HCC. As an HCC invasion suppressor, *KLHL23* blocks the polymerization of F-actin and subsequent lamellipodia and filopodia formation by binding directly to F-actin.[14] Moreover, *CAPZA1* inhibits HCC cells' metastasis via regulating actin cytoskeleton.[15] Meanwhile, *Cofilin* is a major actin-binding protein overexpressed in multiple cancers whose activation is usually an early event in cell migration. *Cofilin* promotes EMT and cancer metastasis through involvement in cytoskeletal reorganization and lamellipodium formation.[16] The dephosphorylation of *Cofilin* on serine 3 is facilitated by phosphatase type 1 (*PP1*), which results in activation of *Cofilin* and actin binding.[16,17]

In this study, we analyzed associations between 529 single nucleotide polymorphisms (SNPs) in 98 genes involved in actin-cytoskeleton regulatory pathway and identified rs13025377 in PPP1CB gene which may contribute to the HBV-related HCC risk in Chinese Han population. Our findings shed light on the relationship between genetic variants in the actin-cytoskeleton regulatory pathway and the risk of HBV-related HCC.

Patients and Methods
Study Population and SNPs' Selection
GWAS data of our previous study have been applied in this study.[7] Briefly, in the present study we used GWAS data of 2514 chronic HBV carriers, including 1161 HBV-related HCC cases and 1353 chronic HBV carriers without HCC at recruitment as controls (Supplementary Table 1). The inclusion criteria of chronic HBV carriers and HCC diagnosis standard were described in a previous paper.[7] 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. These patients were enrolled from Qidong, Jiangsu Province, an area with high incidence of HCC in China. By using Illumina Human OmniExpress BeadChips, a total of 731,442 SNPs were inspected by Genome-wide scan. Imputation was performed by referencing HRC 1.1. Then 4,862,437 SNPs remained under the following post-imputation quality control standards: i) minor allele frequency (MAF) more than 5%; ii) Hardy–Weinberg equilibrium (HWE)$\geq 1\times 10^{-4}$; iii) call rate>95%. Finally, we performed association analysis with additive model and 240,150 SNPs ($P < 0.05$) were kept for further filtering process.

A total of 98 genes involved in actin cytoskeleton regulation were selected from the "actin-cytoskeleton pathway" of Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome Pathway Database and published literatures. 529 SNPs with suggestive significance ($P < 0.05$) were selected in those 98 candidate genes' regions. Then we used Rugulome DB (https://regulome.stanford.edu/index) to predict SNPs function and removed SNPs with ranking score $\geq$ 6. Next we conducted LD analysis to identify independent association and their tag SNPs ($r^2 \geq$ 0.8). Finally, 94 tag SNPs were selected for further exploration.

Bioinformatics Analysis and Expression Analysis
Normal liver RNA-seq transcriptome data were downloaded from Genotype-Tissue Expression (GTEx) (http://www.gtexportal.org/) and used to perform eQTL analysis for rs13025377 and rs4665434. Level 3 mRNA expression profiles were obtained in The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC). Other transcriptome data (GSE14520, GSE84044, GSE124535 and OED094052) were downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo) and National Omics Data Encyclopedia (NODE) database (https://www.biosino.org/node).

ChIP-seq, DNAse-seq and ChIA-PET Data Analysis

ChIP-seq data (ENCFF693EOH, ENCFF689CPO, ENCFF869INM), DNAse-seq data (ENCFF774LVT, ENCFF561QXW, ENCFF553BEF), and HepG2 ChIA-PET raw data (ENCLB951KEQ and ENCLB607ZTO) were downloaded from Encyclopedia of DNA Elements (ENCODE) database (https://www.encodeproject.org/). Two HepG2 ChIA-PET raw data were merged and we processed the data with ChIA-PET2 tool (https://github.com/GuipengLi/ChIA-PET2). ChIP-seq, DNAse-seq and ChIA-PET data were visualized with Integrative Genomics Viewer (IGV) (http://software.broadinstitute.org/software/igv/).

Motif Analysis

The effects of rs4665434 on *CTCF* and *Cohesin* (*RAD21* and *SMC3*) binding motifs were analyzed with flanking sequence of rs4665434 by HaploReg v4.1 (http://archive.broadinstitute.org/mammals/haploreg/haploreg.php).

Survival Analysis

Survival analyses of *PPP1CB, CTCF* and (*RAD21* and *SMC3*) were performed on TCGA-LIHC cohort using UCSC Xena (https://xenabrowser.net/heatmap/).

Pathway Enrichment Analysis

We used STRING (https://string-db.org/) to investigate *PPP1CB* protein interactome. Briefly, we applied the following settings: i) network type: full network; ii) meaning of network edges: evidence; iii) active interaction sources: experiments; iv) minimum required interaction score: low confidence (0.150); v) max number of interactors to show: 100.

We obtained the top 300 *PPP1CB*-correlated genes from TCGA-LIHC tumor RNA-seq by performing pairwise gene Pearson correlation analysis. These 300 genes were used to perform KEGG (Kyoto encyclopedia of genes and genomes) pathway enrichment analysis with the "clusterProfiler" R package.

Statistical Analysis

Chi-squared test was applied to examine the differences in demographic distributions between HBV-related HCC cases and chronic HBV carriers without HCC. For the ORs and their 95% CIs for HBV-related HCC risk, we used multivariate logistic regression analyses to adjust for age and gender. Chi-squared test was used to analyze HWE for genotypes. Benjamini–Hochberg false discovery rate (FDR)

correction was used to adjust P values for multiple testing. The significance of differences of gene expression between tumor tissues and normal tissues was tested by two-sided Mann–Whitney test. The significance of correlation between different genes' expression was tested by performing Pearson correlation analysis. With $P < 0.05$ as significance threshold value, all statistical analyses in this study were performed by R software (version 3.5.3) and PLINK-1.90.

Results
Associations of Selected SNPs with HBV-Related HCC Risk

The demographic features of studied subjects are shown in Supplementary Table 1. In this study, 1161 HBV-related HCC cases and 1353 CHB controls were Han Chinese enrolled from Qidong, China, which have been described in our previous GWAS.[7]

We selected 529 SNPs from 98 genes in actin-cytoskeleton regulatory pathway to test their associations with HBV-related HCC risk (Figure 1 and Supplementary Table 2). After SNPs' function prediction and pairwise LD analysis, 94 tag SNPs remained for FDR correction (Figure 1 and Supplementary Table 3). Finally, only rs13025377 in *PPP1CB* was found to be significantly associated with the risk of HBV-related HCC [odds ratio (OR) = 0.81, 95% confidence interval (CI) = 0.72~0.91, $P = 4.88×10^{-4}$, Supplementary Table 3 and Table 1].

Increased *PPP1CB* Expression May Affect HCC Progression

We found that rs13025377 was a significant eQTL for *PPP1CB* across various tissues in GTEx data and the risk allele A increased *PPP1CB* expression (Supplementary Figure 1 and Figure 2A). Kaplan–Meier curves showed that higher *PPP1CB* expression was correlated with shorter overall survival (OS) in HCC patients (Figure 2B). Meanwhile, TCGA and other public HCC transcriptome datasets showed elevated *PPP1CB* expression in HCC tumor tissue (Figure 2C–F). Besides, we observed correlation between *PPP1CB* expression and the pathological stages of HCC in TCGA data (Supplementary Figure 2). Collectively, overexpression of *PPP1CB* may affect HCC progression.

SNP rs4665434 May Be the Functional Variant which Regulates *PPP1CB* Expression

In our GWAS data, four SNPs were in strong linkage disequilibrium (LD) with rs13025377 ($r^2>0.8$, Table 2),
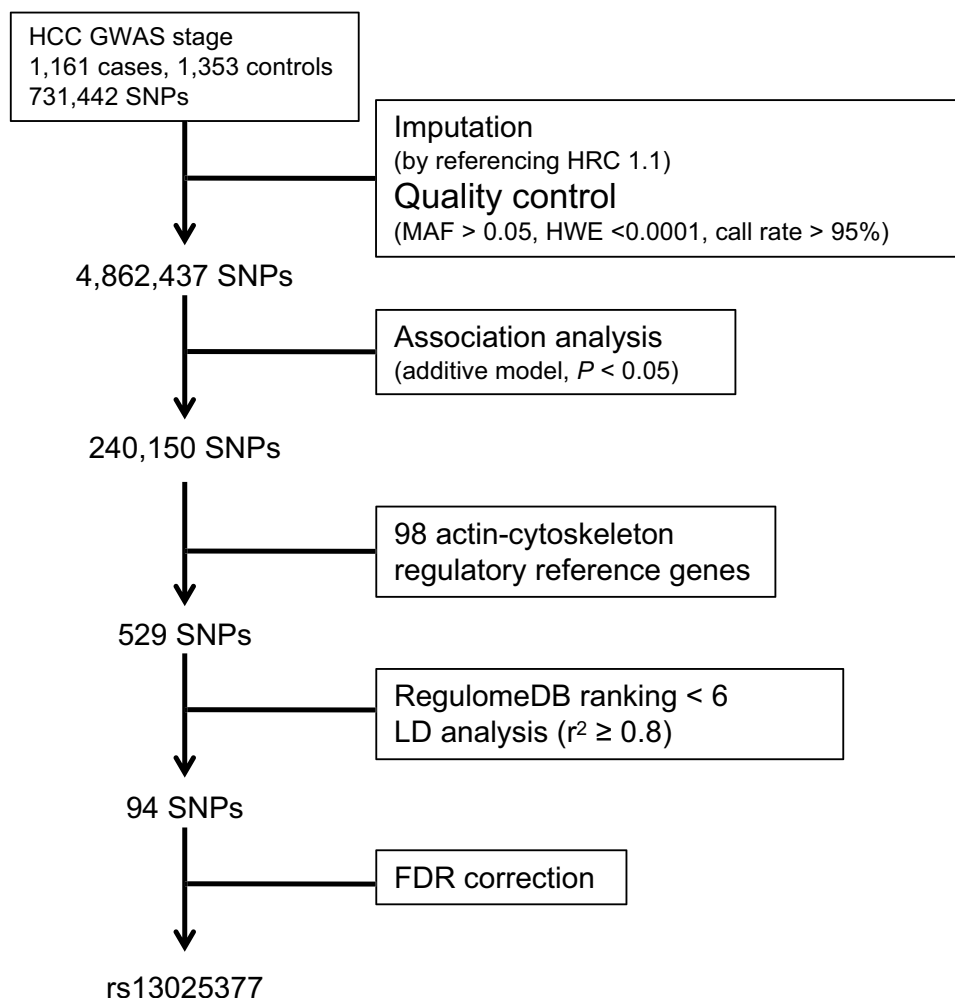
Figure 1 Schematic of the study design and work flow.

and the association at rs13025377 may be driven by some of these SNPs. We then investigated whether SNPs tagged by rs13025377 acted as cis-regulatory elements by HCC. DNase I hypersensitive sites were markers for cis-regulatory elements and enriched in topologically associated domain (TAD) boundaries.[18] Among SNPs that were in strong LD with rs13025377, ENCODE DNase-seq data suggested that rs4665434, an SNP located in the first intron of *PPP1CB*, may be an active regulatory element in hepatocytes and liver cancer cell lines (Figure 3A). In addition, rs4665434 was co-occupied by *CTCF* and *Cohesin* (*RAD21, SMC3*) in HepG2 cell line,

Table 1 Association Between rs13025377 in *PPP1CB* and HBV-Related HCC Risk

Genotypes	Cases		Controls		OR (95% CI)[a]	P[a]
	Number	**%**	**Number**	**%**		
GG	196	16.88	296	21.88	–	–
GA	576	49.61	666	49.22	1.34 (1.08~1.67)	8.49E-03
AA	372	32.04	379	28.01	1.54 (1.22~1.95)	3.35E-04
Additive model					1.23 (1.10~1.38)	4.88E-04
Dominant model (GA+AA vs GG)					1.41 (1.15~1.74)	1.05E-03
Recessive model (AA vs GA+GG)					1.25 (1.05~1.49)	1.35E-02

Note: [a]Adjusted for age and gender in logistic regression model.
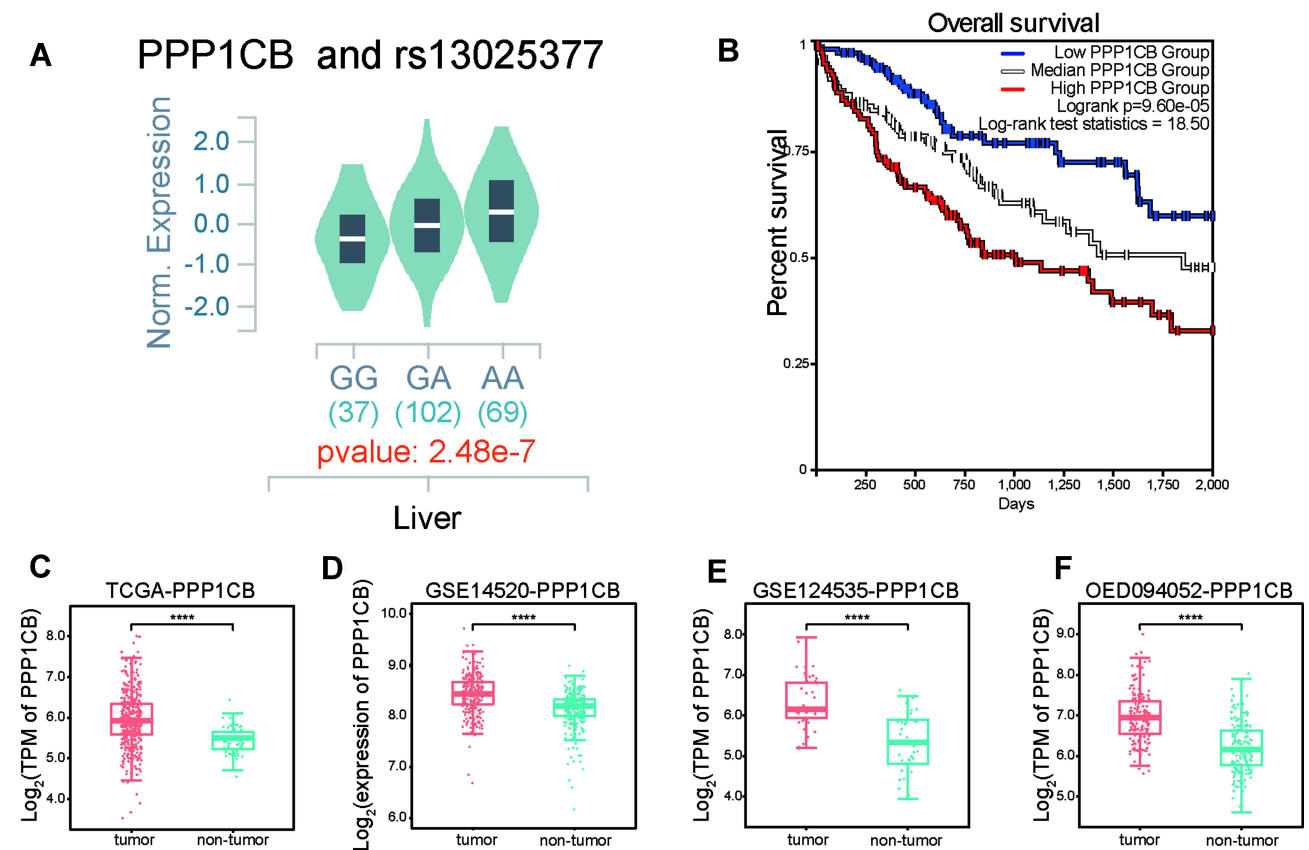
Figure 2 PPP1CB expression may affect HCC progression. (**A**) Expression of quantitative trait loci (eQTL) analysis for rs13025377 based on GTEx liver RNA-seq data. (**B**) Kaplan–Meier plot of OS based on PPP1CB expression in TCGA-LIHC data. (**C–F**) In public datasets, HCC tumor samples exhibited upregulated PPP1CB when compared with non-tumor tissues.

respectively (Figure 3A). And HepG2 *RAD21* ChIA-PET confirmed that rs4665434 was located at a *CTCF/Cohesin*-mediated interaction anchor of TAD in various tissues (Figure 3A). Meanwhile, we found that the protective allele A of rs4665434 could disrupt *CTCF, RAD21* and *SMC3* motifs at this TAD anchor (Figure 3B). GTEx eQTL data showed that the rs4665434 risk allele G increased the *PPP1CB* expression in liver tissue (Supplementary Figure 3). Thus, we speculate that *CTCF* and Cohesin might regulate *PPP1CB* transcription via rs4665434.

Then we analyzed the expression correlation between *PPP1CB* and *CTCF, RAD21, SMC3* in liver tissue. Strong positive correlations between *PPP1CB* and *CTCF, RAD21, SMC3* were observed in TCGA and GTEx liver data (Figure 4A–F), which were also confirmed by other transcriptome data from HBV-related HCC or liver cirrhosis (Supplementary Figure 4). Further, gene expression microarray analysis of a previous study has shown that *PPP1CB* experienced a significant decrease after knockdown of *CTCF* in HCC cell line PLC5.[19] This evidence suggested that rs4665434 may regulate

Table 2 SNPs Tagged by rs13025377 in Previous GWAS Data ($r^2 > 0.8$)

CHR[a]	SNP	OR (95% CI)[b]	EA	P	SNP Location	Gene
2	rs13025377	0.81 (0.72~0.91)	G	4.88E-04	Intergenic	*PPP1CB*
2	rs4665434	0.81 (0.73~0.92)	A	7.72E-04	Intron	*PPP1CB*
2	rs7586405	0.82 (0.73~0.92)	A	9.17E-04	Intergenic	*PPP1CB*
2	rs7594363	0.82 (0.73~0.93)	C	1.46E-03	Intergenic	*PPP1CB*
2	rs6547870	0.83 (0.74~0.93)	G	1.99E-03	Intergenic	*PPP1CB*

Notes: [a]Chromosome number; [b]adjusted for age and gender in logistic regression model.
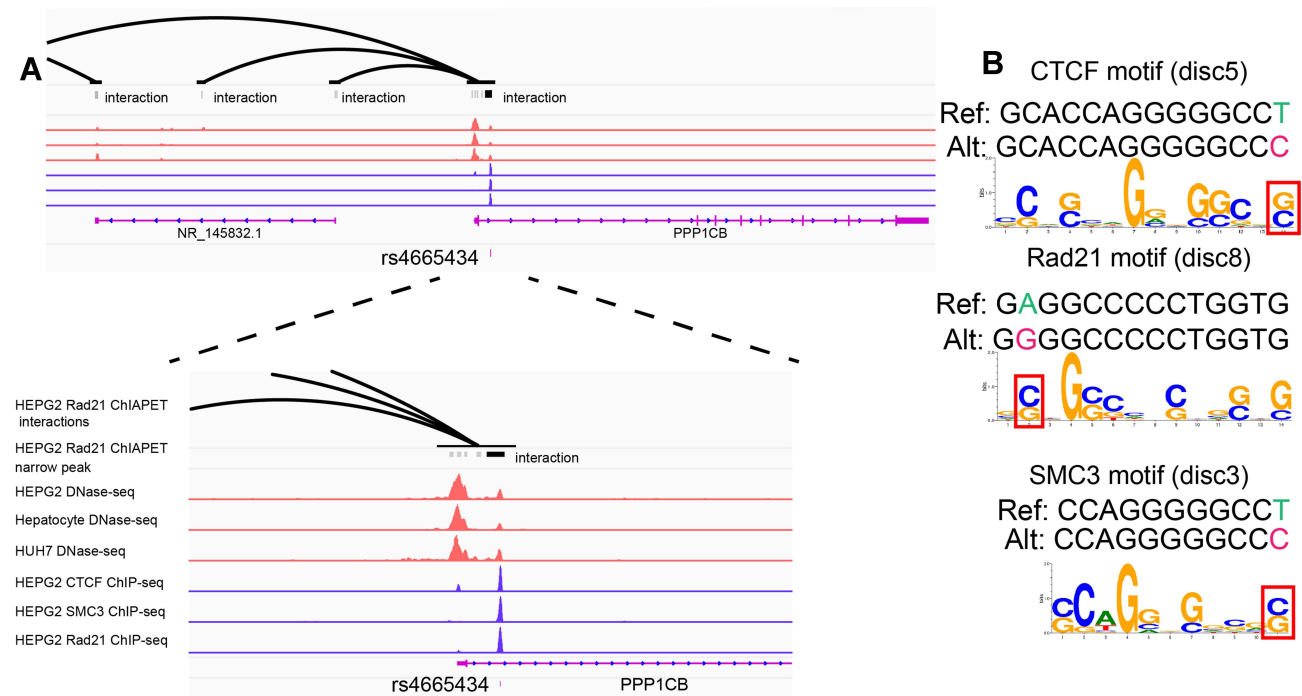**Abbreviation:** EA, effect allele.

Figure 3 SNP rs4665434 with TAD markers resided in motif region of *CTCF* and *Cohesin* (*RAD21* and *SMC3*). (**A** and **B**) SNP rs4665434 located in the first intron of *PPP1CB*. HepG2 *RAD21* ChIA-PET peak 70841 and HepG2 ChIP-seq (*CTCF*, *RAD21* and *SMC3*) peaks indicated that rs4665434 resided in a TAD boundary. HepG2, Huh7 and hepatocyte DNAse-seq peaks indicated rs4665434 resided in a cis-regulatory element. (**B**) Predicted *CTCF*, *RAD21* and *SMC3* motifs for the rs4665434 flanking sequences containing the reference (protective) and alternate alleles.

PPP1CB expression by disrupting binding motif of both *CTCF* and *Cohesin*.

Enrichment Analysis of *PPP1CB*-Related Genes

As one of the three catalytic subunits of *PP1*, *PPP1CB* is involved in a series of biological processes by interacting with a large proteomic interactome.[20] To further investigate the role of *PPP1CB* in HCC, we extracted the top 300 genes that were co-expressed with *PPP1CB* from TCGA data to perform KEGG pathway enrichment analysis. As shown in Figure 5A, the result of enrichment analysis validated that *PPP1CB* participated in "actin-cytoskeleton regulation" (Figure 5A and Supplementary Table 4). Besides, both "Hepatitis B" and "Viral carcinogenesis" were significant pathways identified from *PPP1CB*-coexpressed genes. Then we screened out 100 proteins interacting with *PPP1CB* based on experimental evidence by STRING tool. Similarly, the interaction network also indicated that *PPP1CB* was significantly involved in KEGG pathways such as "actin-cytoskeleton regulation", "Hepatitis B", and "Viral carcinogenesis" (Figure 5B).

Discussion

In the last decade, GWAS has identified 8 genetic variants associated with HBV-related HCC risk (P $\leq$ 5×10^{-8}). However, a large portion of risk loci may remain undetected due to limited sample size and the strict significance threshold of GWAS. By analyzing SNPs within actin-cytoskeleton regulatory pathway, we identified rs13025377 as a novel variant associated with HBV-related HCC risk in this study. SNP rs13025377 was an eQTL for *PPP1CB* expression and we found that rs4665434 may be the causal variant for this risk locus by modulating *CTCF* and *Cohesin* binding affinity.

PP1 catalyzes diversified protein dephosphorylation in eukaryotic cells with a large interactome and *PPP1CB* is one of the three catalytic subunits of it.[20] One of the *PPP1CB*-interacting proteins is *Cofilin*, which is an evolutionarily conserved protein which binds both G-actin (monomeric) and F-actin (filamentous actin). Recent studies reported that the expression levels of *Cofilin* were upregulated in various tumor tissues compared with normal tissues. And *Cofilin* knockdown could disrupt lamellipodium formation and arrest cell cycle in G1 phase.[16] In our study, increased *PPP1CB* expression was observed in HCC tumor tissues, and higher *PPP1CB* expression was
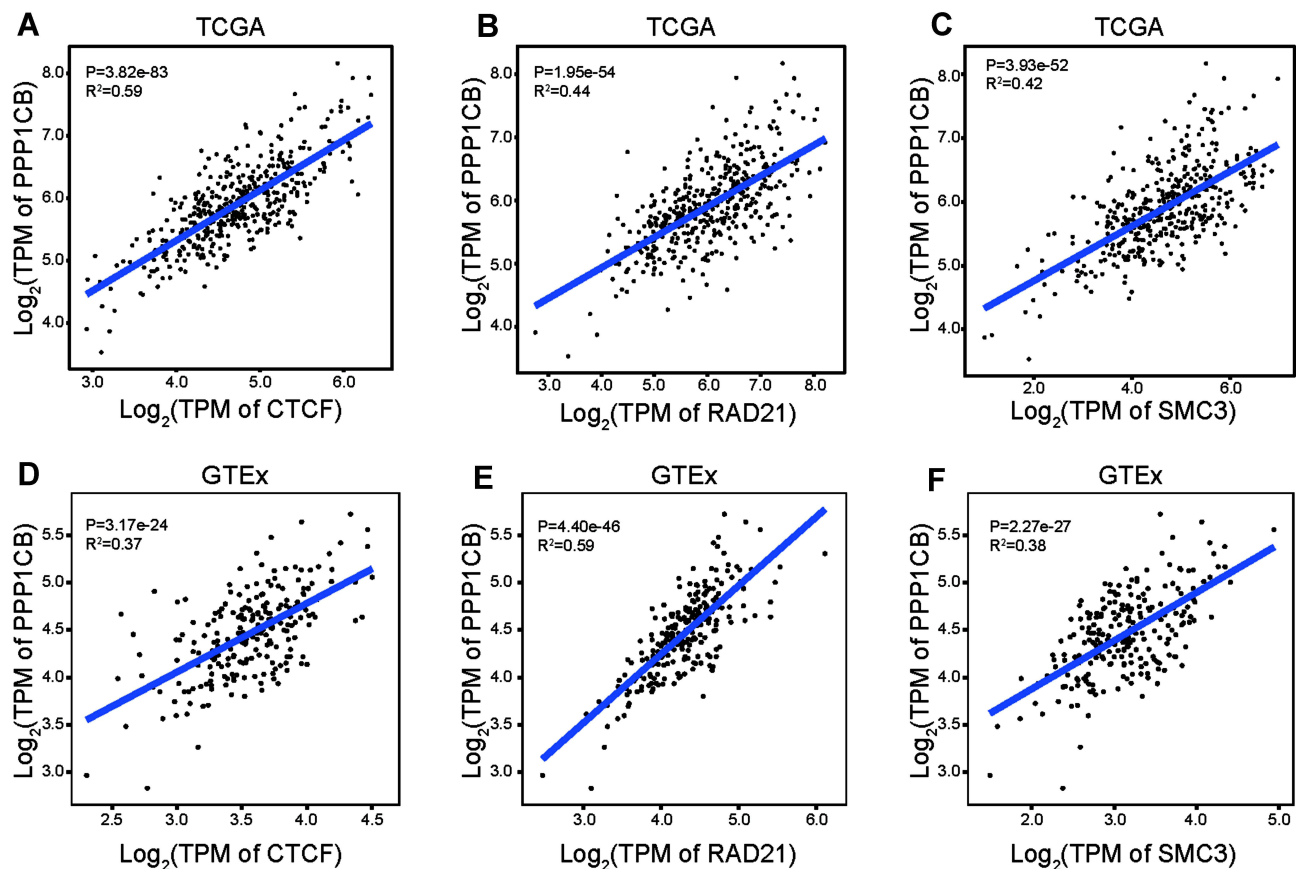
Figure 4 The expression correlation between *PPP1CB* and *CTCF, RAD21, SMC3*. (**A–C**) PPP1CB expression showed positive correlation with *CTCF, RAD21* and *SMC3* in HCC tumor and non-tumor tissues from TCGA datasets. (**D–F**) *PPP1CB* expression showed positive correlation with *CTCF, RAD21* and *SMC3* in normal liver tissues from GTEx datasets.

significantly correlated with shorter OS of HCC patients. A previous study reported that *PP1/PP2A* phosphatase could promote dephosphorylation of *Cofilin* on serine 3. In addition, dephosphorylated *Cofilin* was activated to dynamically reorganize cytoskeleton by facilitating polymerization and depolymerization of actin filaments.[16] Based on our results, we speculate that rs13025377 may be involved in *Cofilin*-mediated actin-cytoskeleton regulation by increasing *PPP1CB* expression. However, biological function studies are needed to validate the speculation.

We found that rs4665434 is located in *CTCF* and *Cohesin* (*RAD21* and *SMC3*) motif at TAD boundary (Figure 3). Both *CTCF* and *Cohesin* are chromatin loops genome-wide organizers and broadly regulate gene expression.[21] SNPs within *CTCF* and *Cohesin* binding motifs could alter chromatin loops topology and link to disease.[22] Previous studies have shown that *CTCF* and *RAD21* may contribute to tumorigenesis in HCC.[19,23,24] We also found that both *CTCF* and *Cohesin* were increased in HCC tumor tissues (Supplementary

Figure 5) and correlated with shorter OS (Supplementary Figure 6). Moreover, expression of *CTCF* and *Cohesin* have shown strong correlation with *PPP1CB* expression and public dataset GSE100533 validated that *CTCF* knockdown could reduce *PPP1CB* expression.[19] A previous study demonstrated *CTCF/Cohesin* binding site mutations accumulated with preferential A•T>C•G and A•T>G•C substitutions in gastrointestinal tumors.[25] Our results also found similar substitutions at rs4665434, thus germline mutation within *CTCF/Cohesin* binding sites also showed promising research prospects for HCC.

In addition to the actin-cytoskeleton regulatory pathway, our pathway enrichment analysis indicated that *PPP1CB* may be involved in HBV-related viral carcinogenesis (Figure 5). A previous study reported that *PPP1CB* was required for HBV life cycle by playing an essential role in Core protein dephosphorylation and pregenomic RNA encapsidation.[26] This evidence suggests that *PPP1CB* and rs13025377 may influence the
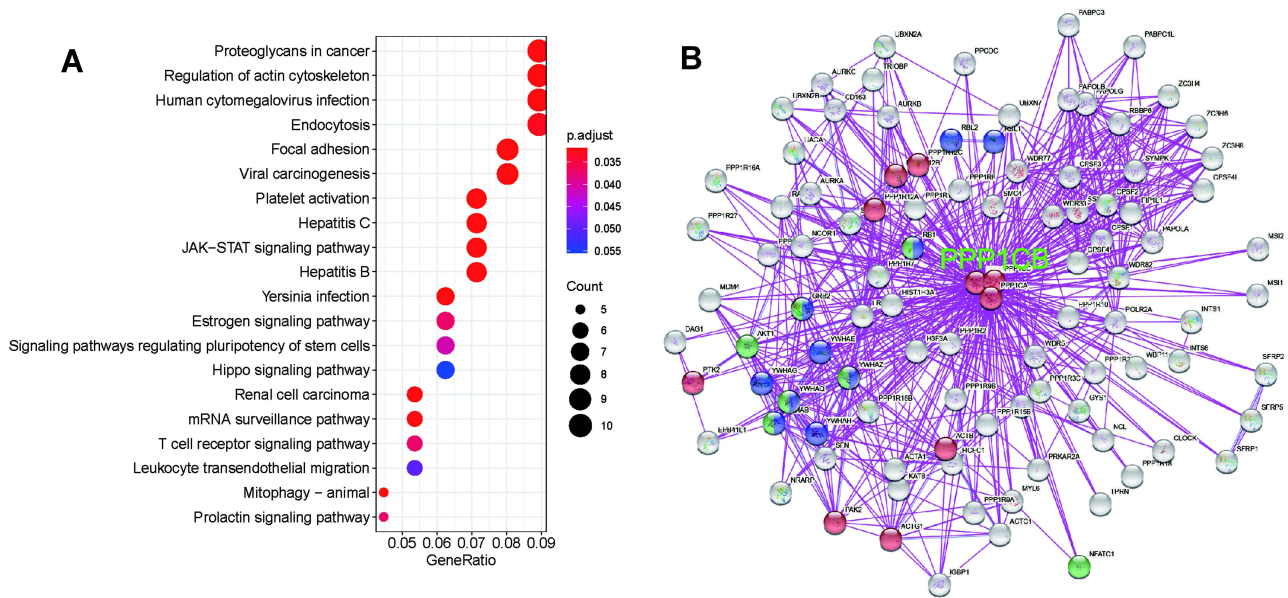
Figure 5 *PPP1CB*-related gene enrichment analysis. (**A**) KEGG pathway analysis based on *PPP1CB* expression-related genes in TCGA-LIHC datasets. (**B**) Experimentally determined *PPP1CB*-binding protein generated by String tool. Genes involved in the actin-cytoskeleton regulatory pathway (KEGG) are shown in red. Genes involved in the KEGG "Hepatitis B" are shown in green. Genes involved in the KEGG "Viral Carcinogenesis" are shown in blue.

HCC tumorigenesis in other ways except the actin-cytoskeleton regulatory pathway.

To our knowledge, this is the first study that comprehensively assessed the associations of SNPs within genes of actin-cytoskeleton regulatory pathway with HBV-related HCC risk. However, this study still has some limitations. First, due to the modest sample size of previous GWAS data, we could not identify rs13025377 under genome-wide significance. The significant findings may need further replication set to be validated. Second, although we integrated multiple bioinformatics data to decipher *CTCF/RAD21*-rs4665434-*PPP1CB* axis in actin-cytoskeleton regulatory pathway, functional experimental evidence is needed to validate biological mechanisms in this study.

Conclusion

In conclusion, we identified rs13025377 in *PPP1CB* gene which may contribute to the HBV-related HCC risk through actin-cytoskeleton regulatory pathway in Chinese Han population. The risk allele of rs13025377 is related to increased *PPP1CB* expression. SNP rs4665434 at *CTCF/Cohesin* mediating TAD boundaries may be the causal variant for the rs13025377 signal. Our results provide deeper insight into genetic variants in actin-cytoskeleton regulatory pathway with the risk of HBV-related HCC and future functional studies of biological mechanisms are warranted.

Abbreviations

GWAS, genome-wide association studies; HBV, Hepatitis B virus; HCC, hepatocellular carcinoma; SNPs, single nucleotide polymorphisms; FDR, false discovery rate; OR, odds ratio; CI, confidence interval; OS, overall survival; KEGG, Kyoto Encyclopedia of Genes and Genomes; EMT, epithelial-mesenchymal transition; PP1, Phosphatase type 1; eQTL, expression quantitative trait loci; LD, linkage disequilibrium; TAD, topologically associated domain; MAF, minor allele frequency; HWE, Hardy–Weinberg equilibrium; GTEx, Genotype-Tissue Expression; TCGA-LIHC, The Cancer Genome Atlas Liver Hepatocellular Carcinoma; GEO, Gene Expression Omnibus; NODE, National Omics Data Encyclopedia; ENCODE, Encyclopedia of DNA Elements; IGV, Integrative Genomics Viewer.

Data Sharing Statement

Normal liver RNA-seq transcriptome data were downloaded from Genotype-Tissue Expression (GTEx). Level 3 mRNA expression profiles were obtained from The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC). Other transcriptome data (GSE14520, GSE84044, GSE124535 and OED094052) were downloaded from the Gene Expression Omnibus (GEO) database and National Omics Data Encyclopedia (NODE) database.

Ethics Approval and Informed Consent

Consent for Publication

Written informed consent for publication was obtained from all participants.

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

Disclosure

Dr Jinlin Hou reports personal fees from AbbVie, Arbutus, Gilead Sciences, and Roche; grants, personal fees from Bristol Myers Squibb and Johnson&Johnson, outside the submitted work. The authors declare that they have no other competing interests.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249. doi:10.3322/caac.21660
2. Society, Health Communication, and Chinese Preventive Medicine Association. [Strategies of primary prevention of liver cancer in China: expert consensus (2018)]. *Zhonghua zhong liu za zhi [Chin J Oncol]*. 2018;40(7):550–557. Chinese.
3. Kulik L, El-Serag HB. Epidemiology and management of hepatocellular carcinoma. *Gastroenterology*. 2019;156(2):477–491. doi:10.1053/j.gastro.2018.08.065
4. Al-Qahtani A, Khalak HG, Alkuraya FS, et al. 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.
5. 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
6. 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
7. 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
8. 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.
9. Li S, Xie L, Du M, et al. Association study of genetic variants in estrogen metabolic pathway genes and colorectal cancer risk and survival. *Arch Toxicol*. 2018;92(6):1991–1999. doi:10.1007/s00204-018-2195-y
10. Wang X, Wu X, Xin J, et al. Genetic variants in Ras/Raf/MEK/ERK pathway are associated with gastric cancer risk in Chinese Han population. *Arch Toxicol*. 2020;94(8):2683–2690. doi:10.1007/s00204-020-02771-w
11. Purrington KS, Slettedahl S, Bolla MK, et al. Genetic variation in mitotic regulatory pathway genes is associated with breast tumor grade. *Hum Mol Genet*. 2014;23(22):6034–6046. doi:10.1093/hmg/ddu300
12. Olson MF, Sahai E. The actin cytoskeleton in cancer cell motility. *Clin Exp Metastasis*. 2009;26(4):273–287. doi:10.1007/s10585-008-9174-2
13. Yamaguchi H, Condeelis J. Regulation of the actin cytoskeleton in cancer cell migration and invasion. *Biochim Biophys Acta*. 2007;1773(5):642–652. doi:10.1016/j.bbamcr.2006.07.001
14. Peng J, Bera R, Chiou C, et al. Actin cytoskeleton remodeling drives epithelial-mesenchymal transition for hepatoma invasion and metastasis in mice. *Hepatology*. 2018;67(6):2226–2243. doi:10.1002/hep.29678
15. Huang D, Cao L, Zheng S. CAPZA1 modulates EMT by regulating actin cytoskeleton remodelling in hepatocellular carcinoma. *J Exp Clin Cancer Res*. 2017;36(1):13. doi:10.1186/s13046-016-0474-0
16. Xu J, Huang Y, Zhao J, et al. Cofilin: a promising protein implicated in cancer metastasis and apoptosis. *Front Cell Dev Biol*. 2021;9:599065. doi:10.3389/fcell.2021.599065
17. Ambach A, Saunus J, Konstandin M, Wesselborg S, Meuer S, Samstag YJ. The serine phosphatases PP1 and PP2A associate with and activate the actin-binding protein cofilin in human T lymphocytes. *Eur J Immunol*. 2000;30(12):3422–3431. doi:10.1002/1521-4141(2000012)30:12<3422::AID-IMMU3422>3.0.CO;2-J

18. Hong S, Kim D. Computational characterization of chromatin domain boundary-associated genomic elements. *Nucleic Acids Res*. 2017;45 (18):10403–10414. doi:10.1093/nar/gkx738

19. Zhang B, Zhang Y, Zou X, et al. The CCCTC-binding factor (CTCF)-forkhead box protein M1 axis regulates tumour growth and metastasis in hepatocellular carcinoma. *J Pathol*. 2017;243(4): 418–430. doi:10.1002/path.4976

20. Bollen M, Peti W, Ragusa M, Beullens M. The extended PP1 toolkit: designed to create specificity. *Trends Biochem Sci*. 2010;35(8):450–458. doi:10.1016/j.tibs.2010.03.002

21. Grubert F, Srivas R, Spacek DV, et al. Landscape of cohesin-mediated chromatin loops in the human genome. *Nature*. 2020;583(7818):737–743. doi:10.1038/s41586-020-2151-x

22. Tang Z, Luo OJ, Li X, et al. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell*. 2015;163(7):1611–1627. doi:10.1016/j.cell.2015.11.024

23. Perez S, Gevor M, Davidovich A, et al. Dysregulation of the cohesin subunit RAD21 by Hepatitis C virus mediates host-virus interactions. *Nucleic Acids Res*. 2019;47(5):2455–2471. doi:10.1093/nar/gkz052

24. Wang J, Zhao H, Yu J, et al. MiR-320b/RAD21 axis affects hepatocellular carcinoma radiosensitivity to ionizing radiation treatment through DNA damage repair signaling. *Cancer Sci*. 2021;112 (2):575–588. doi:10.1111/cas.14751

25. Katainen R, Dave K, Pitkanen E, et al. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nat Genet*. 2015;47(7):818–821. doi:10.1038/ng.3335

26. Hu Z, Ban H, Zheng H, Liu M, Chang J, Guo JT. Protein phosphatase 1 catalyzes HBV core protein dephosphorylation and is co-packaged with viral pregenomic RNA into nucleocapsids. *PLoS Pathog*. 2020;16(7):e1008669. doi:10.1371/journal.ppat.1008669