RESEARCH ARTICLE

Editorial Process: Submission:02/12/2025 Acceptance:19/08/2026 Published:21/08/2026

Mendelian Randomization Using a Japanese GWAS Identifies an HLA-Linked Causal Effect of Chronic Hepatitis B on Cholangiocarcinoma Risk

Khaled Elgeshy[1], Katsuya Nagaoka[1], Hajime Yamazaki[2], Munenori Honda[1], Maiko Nagaoka[3], Takahiro Mizuta[1], Toshinori Toyota[1], Daiki Maeda[1], Sotaro Kurano[1], Kentaro Tanaka[1], Satoshi Narahara[1], Hiroki Inada[1], Takayuki Tokunaga[1], Etsuko Iio[1], Takehisa Watanabe[1], Hiroko Setoyama[1], Haruki Uojima[1], Yasuhito Tanaka[1]*

Abstract

Background: Cholangiocarcinoma (CCA) is a highly malignant cancer that develops in the bile ducts. Its incidence is particularly high in East Asian populations, but the underlying genetic factors remain unclear. To investigate potential risk factors for CCA, we conducted a Mendelian randomization study to infer causality. **Methods:** Using large-scale genome-wide association study data from the BioBank Japan resource, we systematically investigated the causal effects of genetic predisposition to seven conditions, chronic hepatitis B (CHB), chronic hepatitis C, autoimmune hepatitis, type 1 diabetes, type 2 diabetes, chronic gastritis, and chronic pancreatitis, on CCA risk. **Results:** Our analysis reveals a significant association between genetic susceptibility to CHB with a 24% higher likelihood of developing CCA than non-susceptible individuals (Inverse-Variance Weighted Odds Ratio = 1.24, 95% Confidence Interval: 1.08–1.42; p = 0.002). This genetic association is significantly driven by instrumental variables enriched in the immune-regulatory HLA class II region (6p21), suggesting a plausible biological mechanism. For the primary outcome (CCA), statistical significance was assessed across seven exposures at a Bonferroni-corrected threshold (two-sided p<0.0071). Notably, the CHB–CCA association remains significant after correction. This primary finding is strongly supported by comprehensive sensitivity analyses that showed no evidence of confounding by horizontal pleiotropy or heterogeneity. Conversely, no significant causal effects on CCA were identified for the other six conditions. **Conclusions:** Our MR analysis supports a causal role of HBV infection in CCA development, highlighting the importance of targeted HBV screening and surveillance.

Keywords: Cholangiocarcinoma, Hepatitis B virus, Mendelian randomization, Genome-wide association study

Asian Pac J Cancer Prev, **27 (8)**, 2923-2930

Introduction

Cholangiocarcinoma (CCA) is a rare but highly aggressive malignancy of the biliary epithelium, accounting for approximately 15% of all primary liver cancers and having a poor prognosis because of late diagnosis and limited therapeutic options [1]. Although its incidence is relatively low, the global burden of CCA is increasing, particularly in East Asia, including Japan, where risk factors such as viral hepatitis remain common [2].

Viral hepatitis infections, metabolic disorders, and gastrointestinal inflammatory conditions are prevalent in the Japanese population [3, 4] and multiple epidemiological studies have reported associations between these conditions and CCA. Chronic hepatitis B and C (CHB and CHC) are among the most consistently reported risk factors [5–7]. However, some studies have found no significant association between hepatitis C virus infection and CCA development [8]. Autoimmune liver diseases such as autoimmune hepatitis (AIH), have also been linked to CCA in observational studies [9, 10]. Metabolic and inflammatory conditions, including type 1 diabetes (T1D), type 2 diabetes (T2D), chronic gastritis, and chronic pancreatitis, have also been implicated in CCA and other gastrointestinal cancers [11, 12], but whether these associations reflect causality or confounding effects remains unclear.

*[1]Department of Gastroenterology and Hepatology, Graduate School of Medical Sciences, Kumamoto University, Kumamoto, Japan. [2]Section of Clinical Epidemiology, Department of Community Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [3]Health Care Center, Kumamoto University, Kumamoto, Japan. *For Correspondence: ytanaka@kumamoto-u.ac.jp*

Several of these conditions are also known or suspected risk factors for other gastrointestinal malignancies. Hepatocellular carcinoma (HCC) is strongly linked to CHB and CHC, with well-established causal roles [13], and may also be influenced by diabetes and chronic pancreatitis [14, 15].

Mendelian Randomization (MR) is an approach that addresses this uncertainty, using genetic variants as instrumental variables to estimate the causal effects of exposures on disease risk [16]. Because genetic variants are randomly inherited at conception and remain fixed throughout life, MR can minimize confounding effects and eliminate reverse causation, approximating certain features of randomized controlled trials within an observational framework [17].

To date, few MR studies have examined the causal determinants of CCA. A recent MR analysis identified gallstones and liver fat accumulation as independent risk factors [18], but no study has systematically evaluated liver diseases, autoimmune conditions, or metabolic disorders as potential causal determinants of CCA, within a Japanese population using large-scale genetic data.

Here, we performed an MR analysis using GWAS summary statistics from the BioBank Japan (BBJ) project to assess whether genetic predisposition to CHB, CHC, AIH, T1D, T2D, chronic gastritis, and chronic pancreatitis causally influences CCA risk. In addition to CCA, we examined HCC as a secondary outcome.

Materials and Methods

Data Sources and Study Design

This study used publicly available, de-identified GWAS summary statistics from the BBJ project. The original BBJ study was approved by the Institutional Review Board of the Institute of Medical Science, the University of Tokyo, and all participants provided informed consent. As only summary-level data were used, no additional ethical approval was required.

We conducted an MR analysis using GWAS summary statistics from the BBJ Project, with all exposure and outcome data obtained from the BBJ. The GWAS summary statistics were downloaded from the BBJ website between July 2025 and September 2025. The authors did not have access to any information that could identify individual participants during or after data collection. Partial sample overlap could not be excluded. During harmonization, effect alleles were aligned, strand inconsistencies were corrected, and ambiguous SNPs were removed. SNPs failing harmonization quality control were excluded from the analysis. The overall study design and analytic pipeline are summarized in Figure 1.

Exposures and Outcomes

We evaluated the causal effects of genetic predisposition to the following exposures: CHB, CHC, AIH, T1D, T2D, chronic gastritis, and chronic pancreatitis. The primary outcome was CCA. To explore broader potential associations, we also evaluated HCC.

Sample sizes for each trait were as follows: CHB comprised 2,234 cases and 169,588 controls, CHC comprised 7,110 cases and 169,588 controls, AIH comprised 85 cases and 166,529 controls, T1D comprised 1,219 cases and 132,032 controls, T2D comprised the largest sample size with 45,383 cases and 132,032 controls; chronic gastritis comprised 744 cases and 177,982 controls, and chronic pancreatitis comprised 457 cases and 177,471 controls.

For the primary and secondary outcomes, the sample sizes were: CCA with 418 cases and 159,201 controls, and HCC with 2,122 cases and 159,201 controls.

Instrument Selection

Genetic instruments were selected for each exposure using trait-specific p-value thresholds to balance instrument strength and the number of available SNPs. A conventional genome-wide significance threshold ($p < 5\times10^{-8}$) was applied for well-powered traits such as CHB, with a stricter threshold ($p < 1\times10^{-10}$) for T2D to minimize weak instrument bias. For traits with fewer genome-wide significant variants, including CHC, AIH, T1D, chronic gastritis, and chronic pancreatitis, more lenient thresholds ($p < 1\times10^{-6}$ to $p < 1\times10^{-5}$) were used to ensure sufficient instrument availability while maintaining validity. This approach is consistent with previous MR studies that adopted relaxed thresholds for underpowered GWAS to retain adequate statistical power [19].

To ensure independence among selected variants, we performed linkage disequilibrium clumping in PLINK (v1.9) using the 1000 Genomes Project Phase 3 East

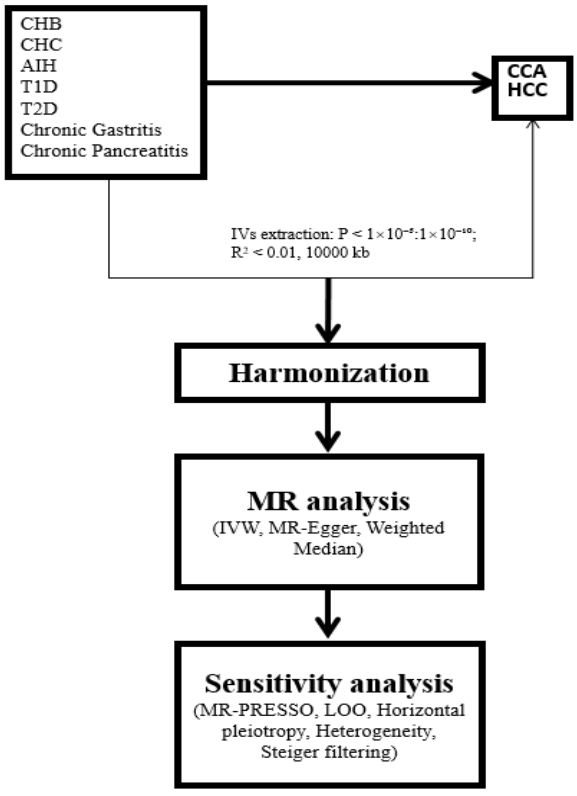

Figure 1. Study Workflow Showing Selection of Genetic Instruments, Data Harmonization, Mendelian Randomization (IVW, MR-Egger, weighted median), and Sensitivity Analyses (MR-PRESSO, Leave-one-out, Pleiotropy, Heterogeneity, Steiger filtering).

Asian reference panel, with $r^2 < 0.01$ and a 10,000 kb window to prune correlated SNPs. We calculated F-statistics to assess instrument strength, with values greater than 10 indicating sufficiently strong instruments. In addition, only autosomal SNPs (chromosomes 1–22) were considered to avoid potential complexities in the interpretation of X chromosome variants.

MR Analysis and Sensitivity Testing

Exposure and outcome datasets were harmonized to align effect alleles and correct strand ambiguities. MR analyses were performed using the TwoSampleMR package (v0.6.20) in R. For each exposure–outcome pair, inverse variance weighted (IVW)-MR was used as the primary method, with weighted median and MR-Egger serving as complementary sensitivity methods.

Several sensitivity analyses were conducted to assess robustness and verify key MR assumptions: heterogeneity was evaluated using Cochran's Q statistic, horizontal pleiotropy was assessed via the MR-Egger intercept, leave-one-out analysis was performed to identify any single SNPs disproportionately influencing the causal estimates, MR-PRESSO was applied to detect and correct for outliers, and Steiger filtering confirmed causal direction. Statistical power was estimated using the non-centrality parameter approach.

Although TwoSampleMR is designed for two-sample MR, we applied its functions to harmonize data and run MR analyses. Possible partial sample overlap between exposure and outcome datasets may bias causal estimates toward observational associations.

The STROBE-MR reporting guideline was adhered to [20]. This study was not pre-registered. Analyses were conducted as part of an exploratory investigation, using publicly available summary statistics. For the primary outcome (CCA), statistical significance across the seven pre-specified exposures was assessed at a Bonferroni-corrected threshold of $\alpha = 0.05/7 \approx 0.0071$ (two-sided). The code used to generate the primary MR results is available from the corresponding author upon request.

Results

Cholangiocarcinoma

Among the tested exposures, only CHB showed a statistically significant causal association with CCA. Using the IVW method, genetic susceptibility to CHB was associated with a 24% increased risk of CCA (IVW OR = 1.24, 95% CI: 1.08–1.42, p = 0.002) and remained significant after Bonferroni correction for seven exposures (threshold p<0.0071) (Table 1 and Figure 2). There was no evidence of horizontal pleiotropy (MR-Egger intercept p = 0.995). Heterogeneity was low, with a Cochran's Q statistic of 9.17 (p = 0.76), supporting the robustness of this finding (Table 1 and Figure 3). Additionally, all instrumental variables used in the CHB-to-CCA MR analysis had F-statistics > 30, indicating strong instrument strength and minimizing the risk of weak instrument bias (Table S1).

In contrast, CHC showed no significant causal effect on CCA (IVW OR = 1.039, p = 0.912), consistent with earlier reports [8], nor did AIH, T1D, T2D, chronic gastritis, or chronic pancreatitis. A suggestive association between chronic gastritis and CCA was noted using MR-Egger (IVW OR = 0.544, p = 0.037), but this result should be interpreted with caution because of potential pleiotropy (p = 0.064) and lack of concordance across the methods (Table 1).

To assess the consistency and directionality of the genetic instruments, a scatter plot was generated showing the per-variant effect estimates on CHB (exposure) and CCA (outcome) (Figure 3). The plot illustrates that most

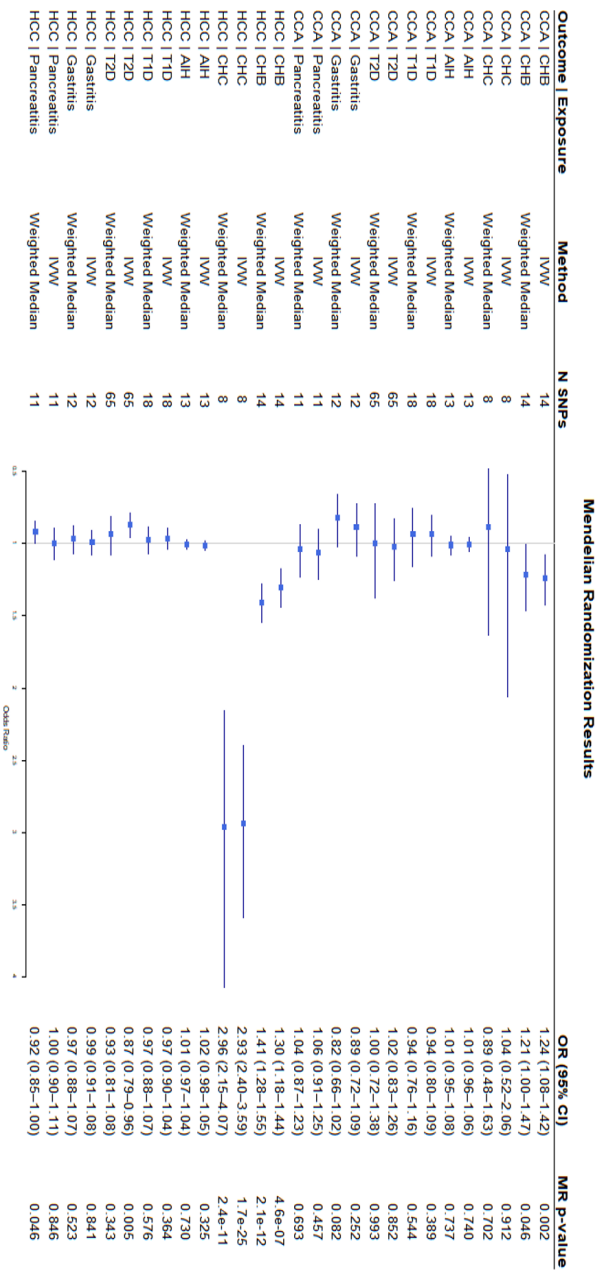

Figure 2. Forest Plot Showing the Causal Effects of Genetic Susceptibility to Liver, Autoimmune, and Metabolic Diseases on the Risk of CCA and HCC. Odds ratios (ORs) with 95% confidence intervals (CIs) are displayed for each exposure–outcome pair, estimated using inverse-variance weighted (IVW; primary method), weighted median, and MR-Egger approaches. The vertical dashed line at OR = 1 represents the null effect.

Table 1. Mendelian Randomization Estimates for the Causal Effect of Genetic Susceptibility to Liver, Autoimmune, and Metabolic Diseases on the Risk of Cholangiocarcinoma (CCA)

Outcomes	Exposures	Method	SNPs number	Odds Ratio	95% CI	MR p-value	Pleiotropy p-value	Heterogeneity p-value
CCA	CHB	IVW	14	1.24	1.08-1.42	0.002	0.995	0.76
CCA	CHB	Weighted Median	14	1.213	1.00-1.47	0.037		
CCA	CHB	Egger	14	1.238	0.76-2.01	0.402		0.689
CCA	CHC	IVW	8	1.039	0.52-2.06	0.912	0.435	0.021
CCA	CHC	Weighted Median	8	0.888	0.48–1.63	0.702		
CCA	CHC	Egger	8	0.551	0.11–2.85	0.504		0.022
CCA	AIH	IVW	13	1.008	0.96–1.06	0.74	0.127	0.62
CCA	AIH	Weighted Median	13	1.011	0.95–1.08	0.737		
CCA	AIH	Egger	13	0.915	0.81–1.0	0.19		0.78
CCA	T1D	IVW	18	0.935	0.80–1.09	0.389	0.627	0.587
CCA	T1D	Weighted Median	18	0.936	0.76–1.16	0.544		
CCA	T1D	Egger	18	1.008	0.72–1.41	0.963		0.534
CCA	T2D	IVW	65	1.02	0.83–1.26	0.852	0.349	0.054
CCA	T2D	Weighted Median	65	0.999	0.72–1.38	0.993		
CCA	T2D	Egger	65	0.833	0.52–1.33	0.449		0.054
CCA	Gastritis	IVW	12	0.888	0.72–1.09	0.252	0.064	0.082
CCA	Gastritis	Weighted Median	12	0.823	0.66–1.02	0.082		
CCA	Gastritis	Egger	12	0.544	0.33–0.89	0.037		0.249
CCA	Pancreatitis	IVW	11	1.063	0.91–1.25	0.457	0.297	0.086
CCA	Pancreatitis	Weighted Median	11	1.036	0.87–1.23	0.693		
CCA	Pancreatitis	Egger	11	1.287	0.89–1.87	0.219		0.105

For each exposure–outcome pair, the table shows the MR method, the number of SNPs used, causal estimate (log odds), odds ratio (OR), and corresponding p-values for the causal association (IVW), horizontal pleiotropy (MR-Egger intercept), and heterogeneity (Cochran's Q). Significant associations (two-sided p<0.05) are shown in bold. Pleiotropy p-values are derived from the MR-Egger intercept; and heterogeneity p-values are derived from Cochran's Q. IVW, inverse-variance weighted, CI, confidence interval. CCA, cholangiocarcinoma; CHB, chronic hepatitis B; CHC, chronic hepatitis C; AIH, autoimmune hepatitis; T1D, type 1 diabetes; T2D, type 2 diabetes.

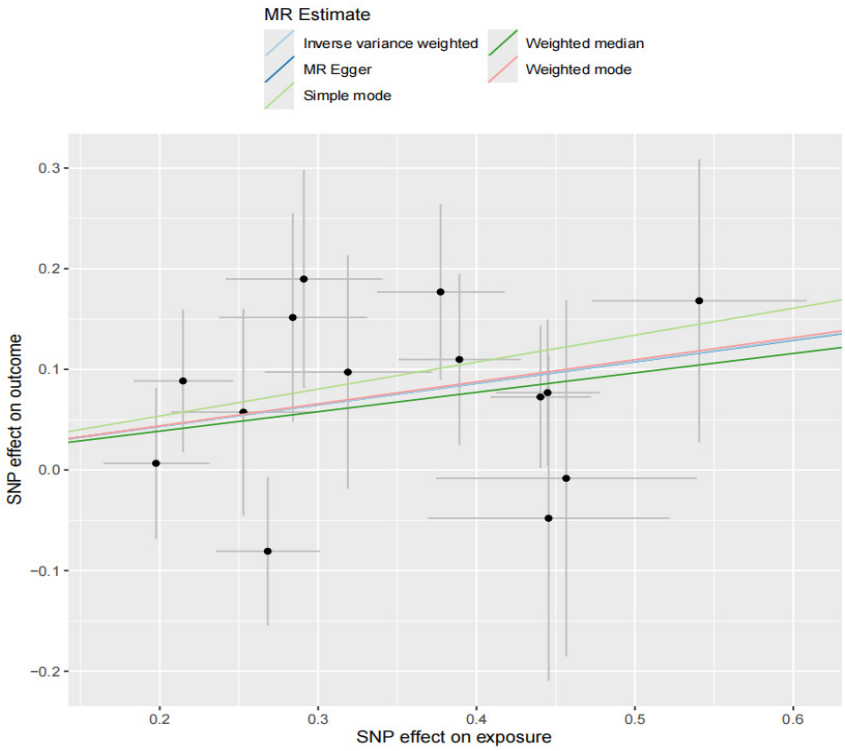

Figure 2. Scatter Plot of SNP-Specific Associations between CHB (x-axis) and CCA (y-axis). Axes are labeled with per-SD genetic effect estimates; fonts and tick intervals are harmonized across the figures. Lines indicate IVW (primary), weighted median, and MR-Egger estimates; shaded areas denote 95% confidence intervals.

Table 2. Instrumental Variables Used in the MR Analysis of Chronic Hepatitis B as Exposure and Cholangiocarcinoma as Outcome

SNP	Position	Mapped gene (Consequences)
rs9277562	chr6:33088985	*HLA-DPB1* (3 Prime UTR Variant)
rs9274664	chr6:32668751	-
rs3997848	chr6:32714617	-
rs3129877	chr6:32440820	*HLA-DRA* (Intron Variant)
rs28891489	chr6:32682585	-
rs6903053	chr6:32203502	*NOTCH4* (Intron Variant)
rs2507993	chr6:31352532	-
rs2894295	chr6:32811313	-
rs201417696	chr6:30345185	*RPP21* (Missense Variant)
		TRIM39-RPP21 (Intron Variant)
rs2116264	chr6:133410467	*EYA4* (Intron Variant)
rs12528890	chr6:33121826	*HLA-DPB2* (Intron Variant)
rs943465	chr6:33763886	*IP6K3* (2KB Upstream Variant)
rs12214820	chr6:27569325	-
rs9259998	chr6:29937583	*LOC124901298* (Intron Variant)

Genomic position, nearest gene, and predicted consequences were annotated using NCBI dbSNP (build 157).

instruments show a positive association, with the slope of the IVW line indicating a significant positive causal effect. The MR-Egger and weighted median estimates are generally consistent in direction with the IVW estimate, suggesting robustness of the observed causal relationship. No major outliers were visually apparent (Figure 3).

Leave-one-out analysis, in which each SNP is sequentially excluded to assess whether any single variant unduly drives the overall result, demonstrated that no single SNP disproportionately influenced the causal estimate between chronic hepatitis B and cholangiocarcinoma (Table S2 and Figure S1). The direction and magnitude of the association remained consistent across all leave-one-out iterations, supporting the robustness of the MR findings.

A funnel plot was generated to visually assess potential directional pleiotropy in the MR analysis of CHB on CCA (Figure S2). The SNP-specific causal estimates (Wald ratios) were symmetrically distributed around the IVW estimate, indicated by the red dashed vertical line. This symmetry suggests no strong evidence of directional pleiotropy affecting the MR results.

The MR-PRESSO (Mendelian Randomization-Pleiotropy RESidual Sum and Outlier) test, a method that detects and corrects for outlier variants indicative of horizontal pleiotropy, did not identify any outliers or evidence of horizontal pleiotropy ($p = 0.78$), supporting the robustness of the causal estimate of CHB on CCA risk (Table S3). All 14 SNPs used as instruments showed more variance in the exposure (CHB) than the outcome (CCA), as indicated by Steiger filtering. Steiger p-values ranged from 1.8×10^{-19} to 0.0044, with a median of 7.0×10^{-5}, confirming correct causal direction (see the right-hand column of Table S3).

Although only three of the 14 CHB-associated SNPs mapped directly to classical *HLA* class II genes (*HLA-DPB1, HLA-DRA, HLA-DPB2*), all but one (rs2116264) lies within the extended HLA region (chromosome 6: 25–34 Mb), a locus densely packed with immune-related genes and characterized by extensive linkage disequilibrium. In a sensitivity analysis excluding all instruments in this region, the CHB–CCA association could not be reliably estimated because of insufficient independent instruments ($N = 1$), suggesting that the observed causal signal is concentrated within this immunogenetic hotspot.

Colocalization analysis in the 6p21 region did not support a shared causal variant between CHB and CCA (posterior probability of $H4 = 0.039$), suggesting the MR association reflects a causal pathway (CHB → chronic inflammation → CCA) rather than pleiotropy via a single variant influencing both traits, consistent with MR-Egger results showing no evidence of directional pleiotropy (intercept $p = 0.995$).

Additionally, we screened all 14 CHB instruments for associations with known cholangiocarcinoma risk factors, including alcohol consumption, smoking, BMI, type 2 diabetes, gallstones, liver enzymes (ALT/AST), and chronic biliary diseases, using the GWAS Catalog (accessed November 2025). None showed genome-wide significant ($p < 5\times10^{-8}$) associations with these potential confounders. The only reported associations were with immune-related traits (e.g., complement C4, B2M, LY96) and HBV-related outcomes, consistent with the immunogenetic role of the HLA region in viral clearance.

Hepatocellular carcinoma

As expected, both CHC and CHB were shown to be strongly associated with HCC risk: CHC shows an OR = 2.933, $p = 1.75\times10^{-25}$ and CHB shows an OR = 1.302, $p = 5\times10^{-7}$ (Table S4). These results serve as positive controls, confirming the validity of the MR framework for detecting known causal relationships.

Discussion

This MR study revealed a significant causal relationship between genetic susceptibility to CHB and increased risk of CCA in a large Japanese cohort. Multiple sensitivity analyses confirmed the robustness of this association, with no evidence of pleiotropy or heterogeneity. The associations of CHB and CHC with HCC further validated the analytic framework. In contrast, no causal effects were identified for other liver, autoimmune, or metabolic diseases examined.

One of the most notable findings of our study is that the genetic markers for CHB are concentrated in immune-regulatory regions, particularly within the *HLA class II* region on chromosome 6p21. Three of the 14 SNPs mapped to *HLA-DPB1, HLA-DRA*, and *HLA-DPB2,* which present viral antigens to helper T cells and influence clearance of hepatitis B virus (HBV) [21, 22]. Variants in *HLA-DPB1, HLA-DQB1*, and *HLA-DRB1* have been associated with HBV persistence and hepatocarcinogenesis [23, 24], while *HLA-B* and *NOTCH4* contribute to immune regulation and cancer pathways (Table 2) [25]. These data provide a plausible molecular mechanism supporting our MR findings: individuals with certain genotypes may mount less effective immune responses to HBV, leading to persistent infection [22, 23]. Persistent HBV infection causes chronic hepatic and biliary inflammation, with cytokine release and reactive oxygen species, inducing DNA damage and genomic instability in cholangiocytes, promoting malignant transformation [26]. Power analysis suggested that the study had adequate statistical power (>80%) to detect the observed effect size (OR = 1.24) at the Bonferroni-corrected threshold ($\alpha = 0.0071$), supported by strong instrument strength (NCP = 46.9).

Interestingly, our analyses suggested an inverse association between genetic predisposition to T2D and the risk of HCC. This finding contrasts with numerous observational studies reporting T2D as a risk factor for these malignancies [14]. Several explanations could account for this discrepancy. First, the confounding effect of medication may play a role: antidiabetic drugs, particularly metformin, have been reported to reduce cancer risk through multiple mechanisms, including AMPK activation and inhibition of tumor proliferation [27, 28]. Second, survival or selection bias may also contribute. For example, individuals with severe diabetes may die from cardiovascular or metabolic complications before cancer develops, or may be underrepresented in biobank cohorts. Such forms of bias can persist, even in a large GWAS, despite the substantial case (n = 45,383) and control (n =132,032) numbers [29]. Third, MR estimates reflect lifelong genetic risk for T2D, which is not identical to clinically diagnosed diabetes that develops later in life and is influenced by treatment. Finally, sample overlap and residual pleiotropy cannot be entirely excluded. Taken together, these results should be regarded as exploratory and hypothesis-generating rather than definitive evidence of a protective effect.

The strengths of this work include the use of a large, ethnically homogeneous dataset that minimizes population stratification, and the application of multiple MR methods and sensitivity tests (MR-Egger). These collectively support the robustness of our findings.

Our MR analyses did not detect a causal effect of CHC and AIH on CCA; however, a null result does not definitively indicate the absence of causality. The CHC analysis relied on only eight SNPs (mean F-statistic = 38.2, $R^2 \approx 0.048$), limiting the power to detect even large causal effects. Given the limited instrument strength for CHC, our study had low power (< 10%) to detect an odds ratio of 2.0 for CCA at the Bonferroni-corrected threshold ($\alpha = 0.0071$). Thus, the null finding for CHC should be interpreted as inconclusive rather than evidence of no effect. Additional limitations include partial sample overlap between the exposure and outcome datasets, which could bias estimates towards observational associations, residual pleiotropy, and the need for replication in independent and ethnically diverse cohorts. While similar GWAS have been carried on European populations, validation in independent ancestries is beyond the scope of this Japan-focused study and is recommended for future work. Despite strong instrument strength (F-statistics > 10) and sensitivity analyses suggesting robustness, these factors indicate that null findings, particularly for CHC, should be interpreted with caution.

Our analyses used summary statistics from a single large Japanese biobank, which may feature partial sample overlap and limit generalizability beyond East Asian populations. While such overlap can bias two-sample MR towards observational associations, our findings are consistent across multiple sensitivity analyses (MR-Egger intercept, MR-PRESSO, Cochran's Q, leave-one-out), and the instrument strength exceeded conventional thresholds, mitigating weak-instrument concerns. Replication in independent East Asian and non-East Asian cohorts will be an important next step. However, recent methodological evaluations have shown that two-sample MR methods can be reliably applied within single large biobank datasets, with the exception of MR-Egger which remains sensitive to sample overlap [30]. This provides reassurance that our analytic framework is robust despite reliance on a single cohort.

European datasets (e.g., UK Biobank, FinnGen) could theoretically enable cross-ancestry validation, but they face major limitations for this specific analysis: extremely low CHB prevalence in European populations, resulting in underpowered GWAS, and potential differences in viral subtypes, host immune response, and environmental co-factors that may limit biological comparability. As such, while cross-population replication is a valuable future direction, it falls beyond the scope of this Japan-focused etiological study. We recommend this approach as a priority for international consortia with sufficient sample sizes and harmonized phenotypes.

Given the rising incidence of intrahepatic cholangiocarcinoma in East Asia, our Japanese population–based MR findings strengthen the rationale for integrating HBV vaccination, antiviral treatment access, and long-term surveillance into national cancer prevention strategies. Prioritizing high-risk groups identified by HBV status may yield measurable reductions in CCA burden.

Clinically, these findings emphasize the value of

targeted HBV screening and surveillance in high-risk populations to improve CCA prevention and early detection. Future work should investigate hepatitis B viral variants and subtypes, and explore gene–environment interactions that may influence bile duct carcinogenesis. Overall, our MR analysis provides strong genetic evidence for a causal role of HBV in CCA development, while the metabolic and autoimmune conditions examined here do not seem to contribute substantially.

Author Contribution Statement

Khaled Elgeshy: Conceptualization, Data Curation, Formal Analysis, Methodology, Software, Visualization, Writing – Original Draft Preparation. Katsuya Nagaoka: Conceptualization, Data Curation, Methodology, Project Administration, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing. Hajime Yamazaki: Conceptualization, Formal Analysis, Software, Methodology, Supervision, Validation, Visualization, Writing – Review & Editing, Maiko Nagaoka: Conceptualization, Validation. Takahiro Mizuta: Conceptualization. Toshinori Toyota: Conceptualization. Daiki Maeda: Investigation. Sotaro Kurano: Investigation. Kentaro Tanaka: Conceptualization. Satoshi Narahara: Conceptualization, Investigation, Software. Hiroki Inada: Conceptualization, Investigation. Takayuki Tokunaga: Conceptualization, Investigation. Etsuko Iio: Investigation. Takehisa Watanabe: Conceptualization, Investigation, Methodology. Hiroko Setoyama: Investigation. Haruki Uojima: Conceptualization. Yasuhito Tanaka: Conceptualization, Resources, Funding Acquisition, Supervision, Validation, Writing – Review & Editing.

Acknowledgements

We thank the BioBank Japan project for making GWAS summary statistics publicly available. This research received no specific grant from any funding agency.

Funding

This work was supported by a grant-in-aid from the Research Program on Hepatitis from the Japan Agency for Medical Research and Development (AMED JP25fk0210172)

Conflict of Interest

Yasuhito Tanaka received honoraria from AbbVie GK, Gilead Sciences, Inc, Chugai Pharmaceutical Co., Ltd., ASKA Pharmaceutical Holdings Co., Ltd., GlaxoSmithKline PLC, AstraZeneca, Eisai and HU frontier. He received research funds from AbbVie GK, FUJIREBIO Inc, Sysmex Corp, GlaxoSmithKline PLC., Gilead Sciences, Inc., AstraZeneca and OTSUKA Pharmaceutical Co., Ltd. Katsuya Nagaoka received a research fund from Eli Lilly Japan K.K. Hiroki Inada received a research fund from the Takeda Science Foundation.

References

1. Banales JM, Marin JJG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, et al. Cholangiocarcinoma 2020: The next horizon in mechanisms and management. Nat Rev Gastroenterol Hepatol. 2020;17(9):557-88. https://doi.org/10.1038/s41575-020-0310-z.
2. Khan SA, Tavolari S, Brandi G. Cholangiocarcinoma: Epidemiology and risk factors. Liver Int. 2019;39 Suppl 1:19-31. https://doi.org/10.1111/liv.14095.
3. Kubo S, Shinkawa H, Asaoka Y, Ioka T, Igaki H, Izumi N, et al. Liver cancer study group of japan clinical practice guidelines for intrahepatic cholangiocarcinoma. Liver Cancer. 2022;11(4):290-314. https://doi.org/10.1159/000522403.
4. Enomoto H, Ueno Y, Hiasa Y, Nishikawa H, Hige S, Takikawa Y, et al. The transition in the etiologies of hepatocellular carcinoma-complicated liver cirrhosis in a nationwide survey of japan. J Gastroenterol. 2021;56(2):158-67. https://doi.org/10.1007/s00535-020-01748-x.
5. El-Serag HB, Engels EA, Landgren O, Chiao E, Henderson L, Amaratunge HC, et al. Risk of hepatobiliary and pancreatic cancers after hepatitis c virus infection: A population-based study of u.S. Veterans. Hepatology. 2009;49(1):116-23. https://doi.org/10.1002/hep.22606.
6. Zhang H, Zhu B, Zhang H, Liang J, Zeng W. Hbv infection status and the risk of cholangiocarcinoma in asia: A meta-analysis. Biomed Res Int. 2016;2016:3417976. https://doi.org/10.1155/2016/3417976.
7. Lin B, He Q, Lu Y, Zhang W, Jin J, Pan H. Viral hepatitis increases the risk of cholangiocarcinoma: A systematic review and meta-analysis. Transl Cancer Res. 2023;12(6):1602-16. https://doi.org/10.21037/tcr-23-892.
8. Cho IR, Yi SW, Choi JS, Yi JJ. Comparison of risk factors for cholangiocarcinoma and hepatocellular carcinoma: A prospective cohort study in korean adults. Cancers (Basel). 2022;14(7). https://doi.org/10.3390/cancers14071709.
9. Garg R, Khan U, AlRajjal A, Kafri Z. Autoimmune hepatitis: A risk factor for cholangiocarcinoma. Case Rep Gastroenterol. 2017;11(3):672-7. https://doi.org/10.1159/000484131.
10. Saengboonmee C, Seubwai W, Lert-Itthiporn W, Sanlung T, Wongkham S. Association of diabetes mellitus and cholangiocarcinoma: Update of evidence and the effects of antidiabetic medication. Can J Diabetes. 2021;45(3):282-90. https://doi.org/10.1016/j.jcjd.2020.09.008.
11. Jing W, Jin G, Zhou X, Zhou Y, Zhang Y, Shao C, et al. Diabetes mellitus and increased risk of cholangiocarcinoma: A meta-analysis. Eur J Cancer Prev. 2012;21(1):24-31. https://doi.org/10.1097/CEJ.0b013e3283481d89.
12. Chang JS, Tsai CR, Chen LT. Medical risk factors associated with cholangiocarcinoma in taiwan: A population-based case-control study. PLoS One. 2013;8(7):e69981. https://doi.org/10.1371/journal.pone.0069981.
13. El-Serag HB. Epidemiology of viral hepatitis and hepatocellular carcinoma. Gastroenterology. 2012;142(6):1264-73.e1. https://doi.org/10.1053/j.gastro.2011.12.061.
14. Wang P, Kang D, Cao W, Wang Y, Liu Z. Diabetes mellitus and risk of hepatocellular carcinoma: A systematic review and meta-analysis. Diabetes Metab Res Rev. 2012;28(2):109-22. https://doi.org/10.1002/dmrr.1291.
15. Quammie S, Crooks CJ, Aliyu A, Aithal GP, Aravinthan AD. Chronic pancreatitis and extra pancreatic cancers- a systematic review and meta analysis. Pancreatology. 2025;25(4):516-27. https://doi.org/10.1016/j.pan.2025.05.003.
16. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133-63. https://doi.org/10.1002/

sim.3034.
17. Davey Smith G, Hemani G. Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014;23(R1):R89-98. https://doi.org/10.1093/hmg/ddu328.
18. Chen L, Fan Z, Sun X, Qiu W, Mu W, Chai K, et al. Examination on the risk factors of cholangiocarcinoma: A mendelian randomization study. Front Pharmacol. 2022;13:900424. https://doi.org/10.3389/fphar.2022.900424.
19. Ruth KS, Day FR, Tyrrell J, Thompson DJ, Wood AR, Mahajan A, et al. Using human genetics to understand the disease impacts of testosterone in men and women. Nat Med. 2020;26(2):252-8. https://doi.org/10.1038/s41591-020-0751-5.
20. Skrivankova VW, Richmond RC, Woolf BAR, Davies NM, Swanson SA, VanderWeele TJ, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (strobe-mr): Explanation and elaboration. BMJ. 2021;375:n2233. https://doi.org/10.1136/bmj.n2233.
21. Zhao Y, Chen K, Yang H, Zhang F, Ding L, Liu Y, et al. Hla-dr genetic polymorphisms and hepatitis b virus mutations affect the risk of hepatocellular carcinoma in han chinese population. Virol J. 2023;20(1):283. https://doi.org/10.1186/s12985-023-02253-2.
22. Karra VK, Chowdhury SJ, Ruttala R, Gumma PK, Polipalli SK, Chakravarti A, et al. Hla-dqa1 & dqb1 variants associated with hepatitis b virus-related chronic hepatitis, cirrhosis & hepatocellular carcinoma. Indian J Med Res. 2018;147(6):573-80. https://doi.org/10.4103/ijmr.IJMR_1644_15.
23. Ou G, Liu X, Xu H, Ji X, Liu X, Wang J. Variation and expression of hla-dpb1 gene in hbv infection. Immunogenetics. 2021;73(3):253-61. https://doi.org/10.1007/s00251-021-01213-w.
24. Huang J, Xiong L, Wang J, Liu Y, Zhu Q, Lei J, et al. Association between the hla-dqb1 polymorphisms and the susceptibility of chronic hepatitis b: A comprehensive meta-analysis. Biomed Rep. 2016;4(5):557-66. https://doi.org/10.3892/br.2016.632.
25. Chu YJ, Yang HI, Hu HH, Liu J, Lin YL, Chang CL, et al. Hbv genotype-dependent association of hla variants with the serodecline of hbsag in chronic hepatitis b patients. Sci Rep. 2023;13(1):359. https://doi.org/10.1038/s41598-023-27570-y.
26. Ehling J, Tacke F. Role of chemokine pathways in hepatobiliary cancer. Cancer Lett. 2016;379(2):173-83. https://doi.org/10.1016/j.canlet.2015.06.017.
27. Evans JM, Donnelly LA, Emslie-Smith AM, Alessi DR, Morris AD. Metformin and reduced risk of cancer in diabetic patients. BMJ. 2005;330(7503):1304-5. https://doi.org/10.1136/bmj.38415.708634.F7.
28. Decensi A, Puntoni M, Goodwin P, Cazzaniga M, Gennari A, Bonanni B, et al. Metformin and cancer risk in diabetic patients: A systematic review and meta-analysis. Cancer Prev Res (Phila). 2010;3(11):1451-61. https://doi.org/10.1158/1940-6207.Capr-10-0157.
29. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008 Sep 20.
30. Minelli C, Del Greco MF, van der Plaat DA, Bowden J, Sheehan NA, Thompson J. The use of two-sample methods for mendelian randomization analyses on single large datasets. Int J Epidemiol. 2021;50(5):1651-9. https://doi.org/10.1093/ije/dyab084