RESEARCH   **Open Access**

Shared genetic factors and causal association between chronic hepatitis C infection and diffuse large B cell lymphoma

Leihua Fu[1*†], Jieni Yu[1†], Zhe Chen[1†], Feidan Gao[1], Zhijian Zhang[1], Jiaping Fu[1], Weiying Feng[1], Pan Hong[1] and Jing Jin[1]

Abstract

Background Epidemiological research and systematic meta-analyses indicate a higher risk of B-cell lymphomas in patients with chronic hepatitis C virus (HCV) compared to non-infected individuals. However, the genetic links between HCV and these lymphomas remain under-researched.

Methods Mendelian randomization analysis was employed to explore the association between chronic hepatitis C (CHC) and B-cell lymphomas as well as chronic lymphocytic leukemia (CLL). Approximate Bayes Factor (ABF) localization analysis was conducted to find shared genetic variants that might connect CHC with B-cell lymphomas and chronic lymphocytic leukemia (CLL). Furthermore, The Variant Effect Predictor (VEP) was utilized to annotate the functional effects of the identified genetic variants.

Results Mendelian randomization revealed a significant association between CHC and increased diffuse large B cell lymphoma (DLBCL) risk (OR: 1.34; 95% CI: 1.01–1.78; $P = 0.0397$). Subsequent colocalization analysis pinpointed two noteworthy variants, rs17208853 (chr6:32408583) and rs482759 (chr6:32227240) between these two traits. The annotation of these variants through the VEP revealed their respective associations with the butyrophilin-like protein 2 (BTNL2) and notch receptor 4 (NOTCH4) genes, along with the long non-coding RNA (lncRNA) TSBP1-AS1.

Conclusion This research provides a refined genetic understanding of the CHC-DLBCL connection, opening avenues for targeted therapeutic research and intervention.

Keywords Chronic Hepatitis C virus, Diffuse large B cell lymphoma, Mendelian randomization, Colocalization analysis

†Leihua Fu, Jieni Yu and Zhe Chen contributed equally to this work and share first authorship.

*Correspondence:
Leihua Fu
hm_henry@163.com
[1]Department of Hematology, Shaoxing People's Hospital, 312000 Shaoxing City, Zhejiang Province, China

Introduction

About 16% of human cancers are linked to infectious agents, predominantly due to viral and bacterial infections [1]. Chronic hepatitis C (CHC) virus infection, primarily affecting the liver, also poses systemic risks, significantly increasing the likelihood of developing B-cell lymphomas [2]. Extensive research, including comprehensive studies and meta-analyses, has underscored an elevated risk of B-cell non-hodgkin's lymphoma in patients with CHC virus compared to those without the infection [3, 4].

Pathophysiological processes from hepatitis C virus (HCV) infection to overt lymphoma involve mechanisms like sustained antigenic stimulation leading to monoclonal expansion and direct viral roles in cellular transformation, especially in large B-cell lymphoma [5, 6]. In vitro studies indicate that the CD81 receptor, crucial for HCV entry, plays a role in B cell infection [7]. HCV replication in B cells might initiate oncogenic events through intracellular viral proteins [8]. This process is linked to oxidative stress caused by viral proteins core and NS3, potentially leading to DNA mutations and repair abnormalities, resulting in cell transformation [9]. Additionally, genomic mutations are also implicated in this process, supported by identified mutations in Neurogenic Locus Notch Homolog Protein (NOTCH) 2, NOTCH1, and Phosphatase and Tensin Homolog (PTEN) in HCV-positive diffuse large B cell lymphoma (DLBCL) patients [10]. However, the molecular pathology linking HCV infection with lymphoma remains elusive. Advanced studies are essential for a comprehensive understanding of the genetic interplay in HCV-associated B-cell lymphoma [3].

Our research utilizes Genome-Wide Association Studies (GWAS) data for an advanced Mendelian randomization (MR) analysis [11–14], a robust method that identifies causal relationships between risk factors and health outcomes while effectively minimizing confounding influences, to establish a causal link between CHC virus infection and B-cell lymphomas as well as chronic lymphocytic leukemia (CLL). We further employ colocalization analysis [15], which plays a pivotal role in uncovering shared molecular mechanisms underlying various diseases and their associated intermediate phenotypes, to discern the causal variants that underlie this association, potentially elucidating the genetic interplay between these conditions. This approach is expected to yield advanced insights into the genetic underpinnings of the relationship between CHC virus infection and B-cell lymphoproliferative disorders.

Materials and methods

Study design

Figure 1 delineates the comprehensive study design. In the initial phase, the study leverages a MR framework, utilizing genetically derived instrumental variables (IVs) to robustly infer the causal nexus between CHC and B-cell lymphomas, inclusive of CLL. Subsequently, in the event of establishing a causal linkage between CHC and other traits, colocalization analysis is employed to discern shared causal genetic variants underpinning CHC and these traits. The final stage involves the application of advanced gene annotation methodologies to prognosticate the consequent effects on genomic and proteomic functionalities.

Data sources

We obtained CHC-related data for the European population ($N=352,013$) from the IEU Open GWAS project and selected SNPs associated with CHC as IVs from the GWAS summary database (https://gwas.mrcieu.ac.uk/). From the Finnish database (https://www.finngen.fi/en), we downloaded various types of B-cell lymphoma outcome data, including DLBCL ($N=315,243$), marginal
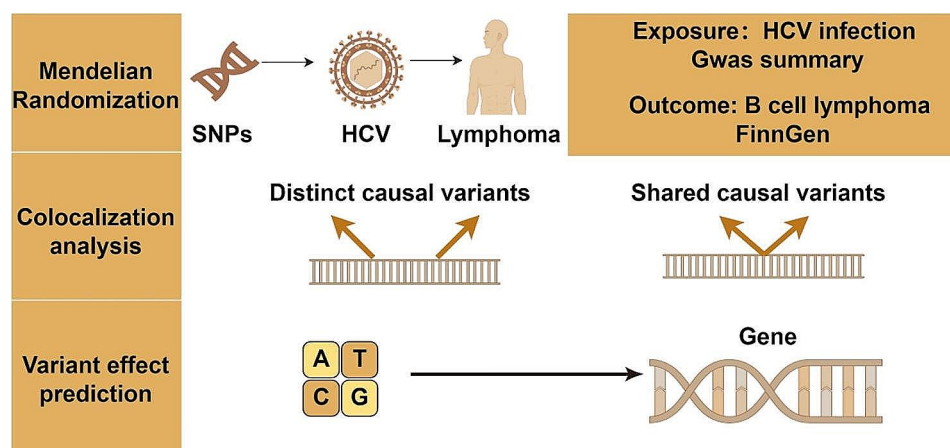
Fig. 1 A detailed flowchart showing the steps of the study

zone lymphoma (MZL) (N=314,395), mantle cell lymphoma (MCL) (N=314,403), Waldenström's macroglobulinemia (WM) (N=314,281), and CLL(N=314,857) as well. Detailed data information is provided in Supplementary Data (Tab. S1).

Selection of instrumental variables

In our approach, we set a genome-wide significance threshold at $P<5\times10^{-8}$ for identifying SNPs strongly correlated with CHC. To select independent SNPs and reduce the influence of linkage disequilibrium (LD), we applied a threshold of 0.01 for the LD parameter (r^2) and a genetic distance criterion of 1000 kb ($r^2<0.01$ and clump distance=1000 kb), based on the European 1000 Genomes Project reference panel. The association strength between the instrumental variables and exposure factors was evaluated using the *F* statistic. To avoid bias from weak instrumental variables, our criteria included only SNPs with an *F* statistic greater than 10 [16, 17]. Detailed data information is provided in Supplementary Data (Tab. S2).

Mendelian randomization

In our study, we conducted a two-sample MR analysis to explore the association between CHC and B-cell lymphomas, as well as CLL, utilizing the "TwoSampleMR" package (Version 0.5.8) in R software. The MR analysis incorporated several methods, notably the inverse variance-weighted model (IVW), MR-Egger, and the weighted median approach [18–21]. IVW was the primary analytical method, with MR-Egger and the weighted median serving as complementary techniques to enhance the reliability and robustness of our findings [21].

To validate our findings on the causal relationship between CHC and B-cell lymphomas, including CLL, we executed a series of sensitivity analyses. We assessed heterogeneity using Cochran's Q statistic in both the IVW and MR-Egger methods [22], considering a p-value over 0.05 as indicative of non-significant heterogeneity [20]. Additionally, to test for horizontal pleiotropy, we employed the Egger intercept and MR-PRESSO method [20], with a p-value above 0.05 suggesting no evidence of horizontal pleiotropy.

Colocalization analysis

We conducted an approximate Bayes factor (ABF) localization analysis to investigate potential shared genetic causal variants in gene regions that may link CHC, B-cell lymphomas, and CLL, utilizing the "COLOC" package (Version 5.2.3) in R software.

ABF localization analysis have five hypotheses [15]:

Hypothesis 0 (H0): Neither Trait 1 nor Trait 2 is associated with the genetic variant.

Hypothesis 1 (H1): Only Trait 1 is associated with the genetic variant.

Hypothesis 2 (H2): Only Trait 2 is associated with the genetic variant.

Hypothesis 3 (H3): Both traits are associated with the genetic variant, but each association is independent of the other.

Hypothesis 4 (H4): Both traits are associated with the genetic variant due to a shared causal mechanism.

To perform this colocalization analysis, we created windows of ±20 kb around the SNPs used for CHC trait IVs. We then executed ABF analysis using standard settings to calculate the posterior probability of the CHC, B-cell lymphomas and CLL having a shared causal signal (under Hypothesis 4: both traits are influenced by a common causal variant in the gene region). By standard practice, a posterior probability for H4 (PP.H4)>85% was regarded as indicative of colocalization [15]. This suggests the presence of a causal genetic variant linking CHC, B cell lymphoma and CLL, possibly pointing to a shared genetic basis in their association.

Variant effect prediction

Utilizing the Variant Effect Predictor (VEP) tool (https://asia.ensembl.org/Tools/VEP) for functional annotation of causal variants identified through colocalization analysis between CHC, B-cell lymphomas and CLL [23].

Results

Causal effect on B cell lymphomas and CLL

Upon screening, five independent SNPs linked to CHC were identified, with *F*-statistics ranging from 33.61 to 92.94. This range suggests a reduced probability of bias due to weak instrumental variables.

Subsequently, a Two Sample MR analysis was conducted to assess the causal relationship between CHC and various B cell lymphoma subtypes, including DLBCL, MCL, MZL, WM, as well as CLL. The primary outcome, derived via IVW method, indicated a causal association between CHC infection and DLBCL (OR: 1.34; 95% CI: 1.01–1.78; *P*=0.0397) (Fig. 2). However, our analysis did not reveal similar causal associations for CHC with other subtypes of B cell lymphoma and CLL.

Comprehensive sensitivity analyses, including heterogeneity assessment using Cochran's Q test, revealed significant heterogeneity from CHC to MZL (Q-value=0.0124, I2=68.72%). Consequently, a random effects IVW model was employed to estimate the MR
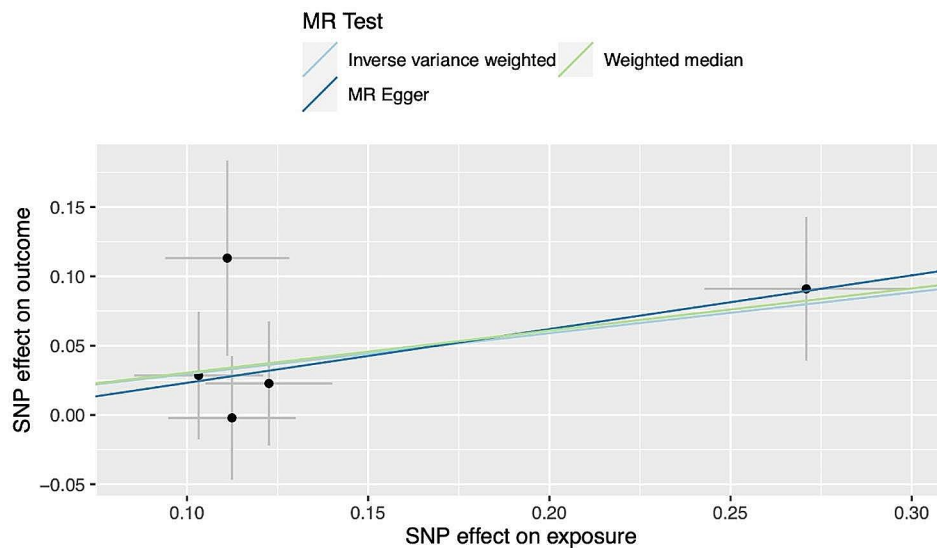
Fig. 2 Scatter plots for causal association between CHC and DLBCL

effect size for this relationship. In contrast, analyses of other lymphoma subtypes did not exhibit such heterogeneity. Further, horizontal pleiotropy was evaluated using MR-Egger regression, and supplementary MR-PRESSO testing did not reveal any data outliers (Fig. 3, Supplementary Data Tab. S3).

Colocalization analyses of CHC with DLBCL

To further investigate the potential causal variant linking CHC and DLBCL, a colocalization analysis was conducted in gene regions for these traits (Fig. 4). The analysis indicated a high posterior probability (94.00% PP.H4) of a shared causal variant within the ±20 kb gene region of rs9501400 (chr6: 32394184), which leads to the causal variant rs17208853. Additionally, a similar high posterior probability of 93.40% PP.H4 was ascertained for the sharing of a causal variant within the gene region (±20 kb) of rs412657 (chr6: 32211085), culminating in the causal variant rs482759. And notably, In the chr6:32642425–32682425 region, our analysis revealed that CHC and DLBCL share independent variants with a high posterior probability of 99.90% PP.H3, indicating distinct genetic influences in this specific genomic area (Fig. 5; Table 1).
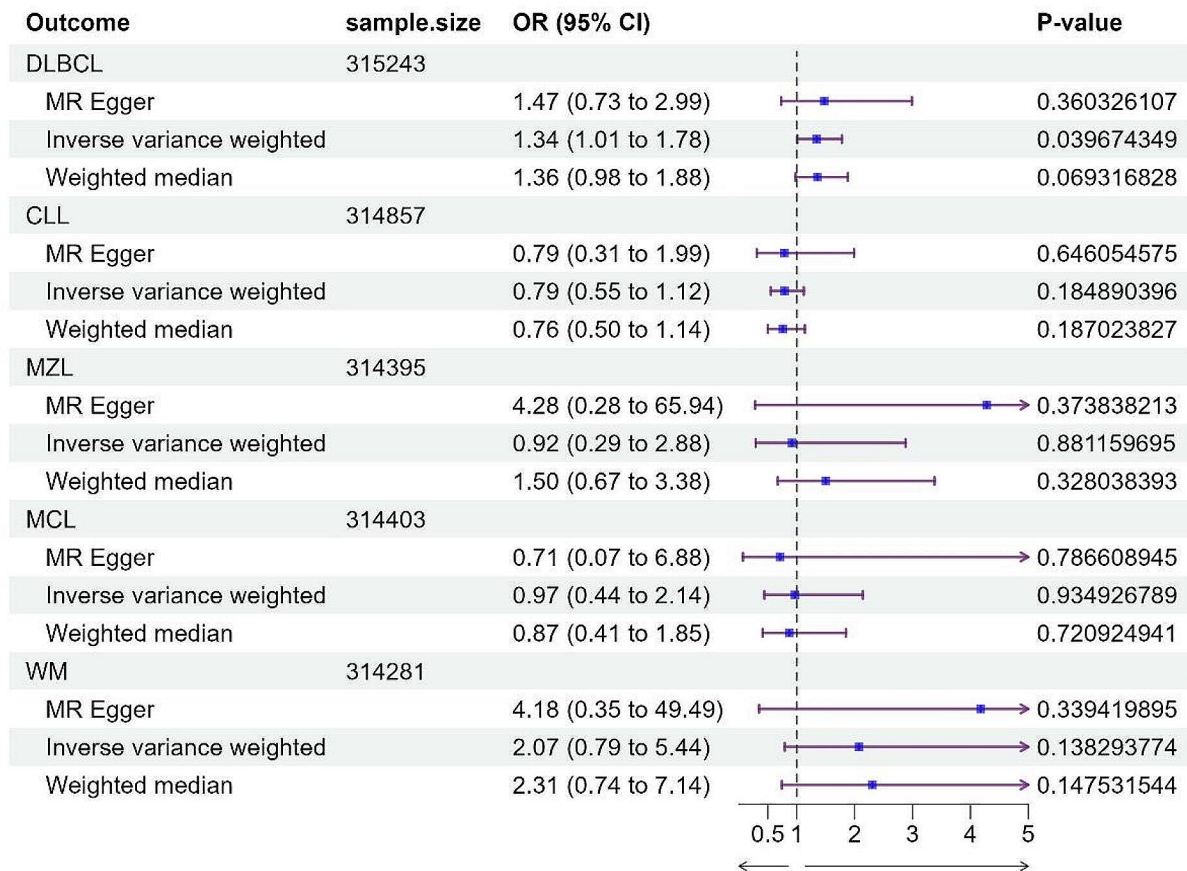
Variant effect prediction

Utilizing the Variant effect prediction (VEP) online tool, we executed a rigorous annotation of SNPs rs482759 and rs17208853. The results designate rs482759 as an upstream regulatory variant for NOTCH4, involved in protein coding modulation and intronic sequence maintenance. Simultaneously, rs17208853, mapped upstream of butyrophilin-like protein 2(BTNL2) gene, emerges as a multifunctional regulatory agent, influencing both

protein-coding and non-coding RNA transcripts, including those marked for nonsense-mediated decay. This highlights its intricate regulatory capacity in BTNL2 gene expression. Additionally, The SNP rs17208853, located downstream of the TSBP1 and BTNL2 antisense RNA 1(TSBP1-AS1), is classified as a modifier and may impact the expression and regulatory mechanisms of this lncRNA (Supplementary Data Tab. S4).

Discussion

In our own study, we employed MR to investigate the association between CHC and B-cell lymphoma subtypes, along with CLL. Our findings indicated a statistically significant elevation in the risk of DLBCL associated with CHC infection (OR: 1.34; 95% CI: 1.01–1.78; $P=0.0397$). This evidence corroborates the causal inference drawn from prior observational correlations [24–27]. Additionally, previous systematic meta-analysis found that, besides DLBCL, CHC is also associated with other types of B-cell lymphoma such as MZL, lymphoplasmacytic lymphoma (LPL), and follicular lymphoma (FL) [25, 28–30]. However, our study did not establish a causal relationship between CHC and these B-cell lymphoma subtypes as well as CLL. Several factors could account for this observed discrepancy. Firstly, MR utilizes genetic variants as IVs to investigate the causal relationship between exposure and outcome, with a critical assumption being the exclusion of confounders [11, 12]. For instance, during the selection of the exposure-related IVs, factors such as chronic inflammatory stimuli are intentionally excluded as confounders, given that chronic inflammation is a recognized pathogenic mechanism for lymphoma [31]. In other words, a known contributing factor to lymphoma pathogenesis is deliberately

Fu *et al. Infectious Agents and Cancer*   (2024) 19:15

Page 5 of 9

Outcome	sample.size	OR (95% CI)		P-value
DLBCL	315243			
MR Egger		1.47 (0.73 to 2.99)		0.360326107
Inverse variance weighted		1.34 (1.01 to 1.78)		0.039674349
Weighted median		1.36 (0.98 to 1.88)		0.069316828
CLL	314857			
MR Egger		0.79 (0.31 to 1.99)		0.646054575
Inverse variance weighted		0.79 (0.55 to 1.12)		0.184890396
Weighted median		0.76 (0.50 to 1.14)		0.187023827
MZL	314395			
MR Egger		4.28 (0.28 to 65.94)		0.373838213
Inverse variance weighted		0.92 (0.29 to 2.88)		0.881159695
Weighted median		1.50 (0.67 to 3.38)		0.328038393
MCL	314403			
MR Egger		0.71 (0.07 to 6.88)		0.786608945
Inverse variance weighted		0.97 (0.44 to 2.14)		0.934926789
Weighted median		0.87 (0.41 to 1.85)		0.720924941
WM	314281			
MR Egger		4.18 (0.35 to 49.49)		0.339419895
Inverse variance weighted		2.07 (0.79 to 5.44)		0.138293774
Weighted median		2.31 (0.74 to 7.14)		0.147531544

Fig. 3 The forest plot, utilizing IVW, MR-Egger, and weighted median methods, demonstrated causal associations between CHC and Bcell lymphomas, including CLL. A significant causal relationship was found between CHC infection and DLBCL (OR: 1.34; 95% CI: 1.01–1.78; $P = 0.0397$)

eliminated. Secondly, the pathogenesis of CHC-related lymphoma involves a complex cascade of molecular and cellular events. The development of HCV-related lymphoma involves ongoing stimulation of lymphocytes by viral antigens, leading to cell proliferation; oncogenic effects from HCV replication in B cells, mediated by viral proteins; and lasting B-cell damage from transient viral presence, a phenomenon known as the "hit and run" theory [32]. Nevertheless, these critical pathogenic factors are not assessed in MR studies. In a systematic meta-analysis [33, 34], however, the calculated incidence rate of lymphoma accounts for the possibility of all occurrences of the disease. The exclusion or infeasibility of evaluating these factors may result in MR being less capable to provide as comprehensive and conclusive a result as a systematic meta-analysis might, potentially contributing to the discrepancy between their findings. The last potential reason for inconsistencies might be inadequate statistical power in MR studies, influenced by the prevalence of genetic variants used, their effect size on the risk factor, and the study sample size [12, 14]. These factors may lead

to discrepancies between the conclusions of epidemiological studies and those derived from MR, yet the two sets of findings are not necessarily contradictory.

Recent advancements have decrypted various pathogenic pathways linking CHC to non-Hodgkin lymphoma, emphasizing the significant role of genetic determinants in HCV-associated lymphoproliferative disorders [9, 10, 35, 36]. In a comprehensive French study with 87 HCV-associated lymphoma patients, the TNF alpha induced protein 3 (TNFAIP3)/A20 gene's rs2230926G allele was found more frequently in patients exhibiting rheumatoid factor (RF) activity (20%). This correlation suggests that even minimal A20 dysfunction, potentially resulting in heightened NF-kB activity, might be enough to trigger the lymphomatous transformation in autoimmune B cells, particularly under conditions of chronic RF+B cell stimulation [37]. In our study, we utilized colocalization analysis [15], which plays a pivotal role in uncovering shared molecular mechanisms underlying various diseases and their associated intermediate phenotypes, we pinpointed two causal variants - rs17208853
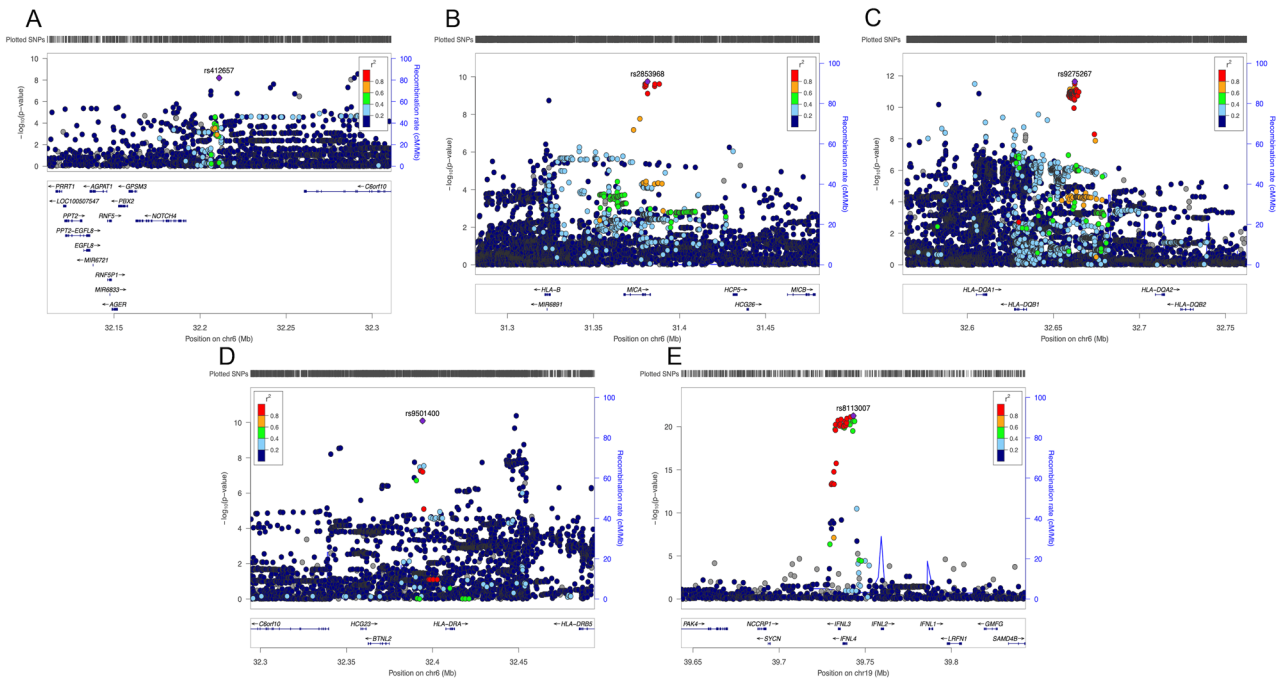
Fig. 4 The LocusZoom plot in the study displays the SNPs used for analyzing the CHC trait, featuring five key variants located on chromosomes 6 and 19. The plot displays -log10 *P*-value on the y-axis and physical position on the x-axis. Points identify individual variants whose colour indicates their LD r2 with key variants. A: rs412657(chr6: 32211085), B: rs2853968(chr6:31381351), C: rs9275267 (chr6:32662425). D: rs9501400(chr6:32394184). E: rs8113007(chr19: 39743103)
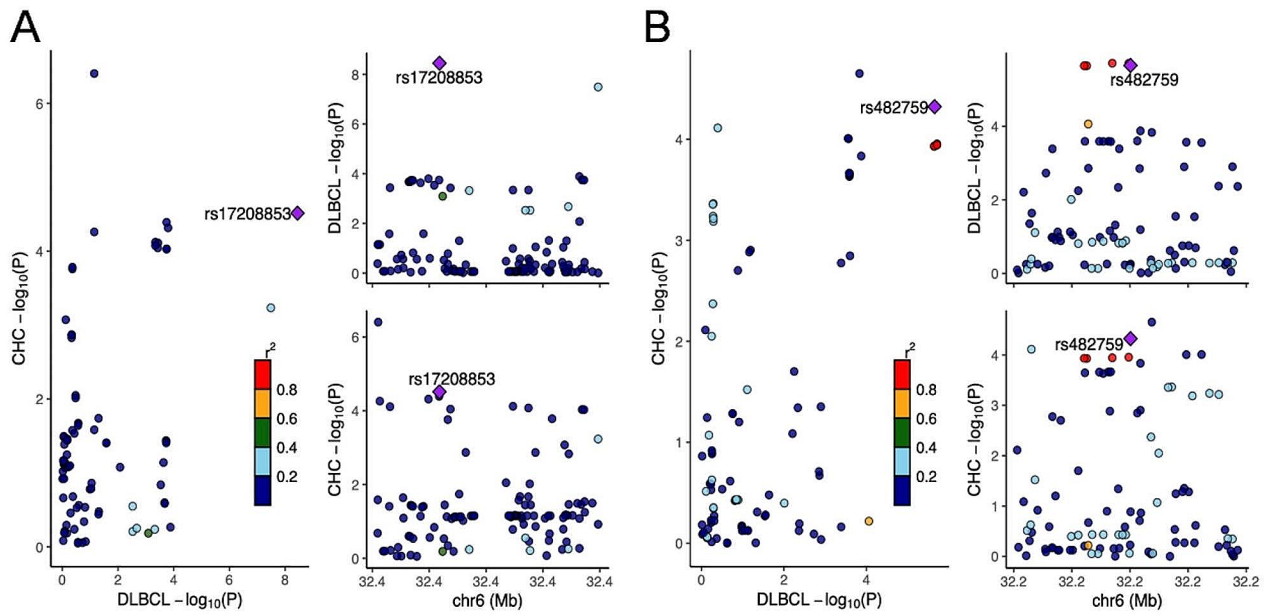
Fig. 5 Colocalization analysis results for the association between CHC and DLBCL. A: A shared causal variant within the ± 20 kb gene region of rs9501400, which leads to the causal variant rs17208853. B: A shared causal variant within the ± 20 kb gene region of rs412657, which leads to the causal variant rs482759

and rs482759 - that mediate the correlation between CHC infection and DLBCL, implying a potential shared genetic underpinning for these two traits.

Further VEP analysis suggested that rs17208853 and rs482759 are regulatory variants affecting BTNL2 and NOTCH4genes, with rs17208853 also influencing the lncRNA profile of TSBP1-AS1. The BTNL2

Table 1 The result from colocalization analysis of chronic hepatitis C infection with DLBCL

SNPs	CHR	Position	Region	PP.H3	PP.H4	Best causal	SNP.PP.H4
rs2853968	6	31381351	Chr6:31361351–31401351	0.707%	2.200%	NA	NA
rs9275267	6	32662425	Chr6:32642425–32682425	99.900%	0.021%	NA	NA
rs9501400	6	32394184	Chr6:32374184–32414184	4.570%	94.000%	rs17208853	0.9918
rs412657	6	32211085	Chr6:32191085–32231085	1.280%	93.400%	rs482759	0.3506
rs8113007	19	39743103	Chr19:39723103–39763103	0.000%	0.001%	NA	NA

gene functions predominantly as an immune regulator, with high expression in lymphoid tissues and structural homology to B7 co-stimulatory molecules [38, 39]. It is implicated in dampening T-cell proliferation and cytokine production, thereby modulating T-cell-mediated responses, especially in the gastrointestinal tract [40]. Research by Waller RG et al. recognizes BTNL2 among genes that elevate the risk of multiple myeloma, indicating potential genetic overlap with other lymphoid malignancies [41]. Additionally, Vijai J et al. identify loci near BTNL2 linked to MZL, emphasizing its role in lymphoma pathogenesis [42]. Although no direct genetic link between CHC and MZL was found in our study, it is important to note that MZL carries a risk of transformation into DLBCL [43]. The NOTCH pathway is recurrently mutated in DLBCL associated with hepatitis C virus infection [10]. NOTCH2, NOTCH1 and PTEN mutations have been identified in respectively, 20%, 4% and 2% of HCV-positive DLBCL patients [10]. NOTCH4, a key component of the NOTCH pathway, has been shown to have significant associations with both benign and malignant lymphoproliferative diseases related to HCV [44, 45]. Our colocalization analysis has identified rs482759 as a regulatory variant located upstream of the NOTCH4 gene. This variant is involved in the encoding of NOTCH4 proteins and may play a role in the pathogenesis of DLBCL related to CHC. The TSBP1-AS1 gene, spanning 139 kb and characterized as a lncRNA, is highly expressed in immune system cells [46]. It is notable for its partial overlap with the protein-coding genes TSBP1 (also known as C6orf10) and BTNL2, which are transcribed from the opposite strand [47, 48], and have pleiotropic effects on autoimmune disorders [49, 50]. The precise function of TSBP1-AS1 in lymphomas remains to be elucidated, necessitating further focused research to unravel its potential roles in lymphomagenesis.

Our research constitutes a pioneering application of MR to investigate the risk association between CHC and B-cell proliferative diseases, establishing CHC as a risk factor for DLBCL and identifying two causal genetic variants. This contributes significantly to understanding the genetic interplay between these pathologies. However, our results also have some limitations. Firstly, the study's concentration on a European population limits the extension of our findings to other ethnic groups. Additionally, despite the identification of a colocalized signal, the complex underlying genetic mechanisms still require more comprehensive exploration. Furthermore, our analysis, while comprehensive for various B-cell lymphoma subtypes, faced limitations due to data availability, which restricted the inclusion and examination of certain lymphomas, such as FL.

Conclusion

In summary, our study utilizes Mendelian randomization to explore the genetic relationship between CHC and B-cell lymphoproliferative disorders, particularly focusing on DLBCL. We successfully identified CHC as a contributing risk factor for DLBCL and uncovers specific genetic variants that provide a causal link. These findings offer significant insights into the genetic interplay between these conditions. Further clinical and experimental studies are required to validate the relationship between CHC and B-cell lymphoproliferative disorders.

Abbreviations

CHC	Chronic hepatitis C
HCV	Hepatitis C virus
NOTCH	Neurogenic Locus Notch Homolog Protein
PTEN	Phosphatase and Tensin Homolog
DLBCL	Diffuse large B cell lymphoma
GWAS	Genome-Wide Association Studies
MR	Mendelian randomization
CLL	Chronic lymphocytic leukemia
IVs	Instrumental variables
MZL	Marginal zone lymphoma
MCL	Mantle cell lymphoma
WM	Waldenström's macroglobulinemia
LD	Linkage disequilibrium
IVW	Inverse variance-weighted model
ABF	Approximate Bayes factor
VEP	Variant Effect Predictor
BTNL2	Butyrophilin-like protein 2
TSBP1-AS1	TSBP1 and BTNL2 antisense RNA 1
LPL	Lymphoplasmacytic lymphoma
FL	Follicular lymphoma
TNFAIP3	TNF alpha induced protein 3
RF	Rheumatoid factor

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13027-024-00577-4.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Declarations

Ethics approval and consent to participate
Due to the focus on collecting previous data for analysis, no additional ethical approval was required.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. de Martel C, Ferlay J, Franceschi S, Vignat J, Bray F, Forman D, Plummer M. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. Lancet Oncol. 2012;13(6):607–15.
2. Gill K, Ghazinian H, Manch R, Gish R. Hepatitis C virus as a systemic disease: reaching beyond the liver. Hepatol Int. 2016;10(3):415–23.
3. Couronne L, Bachy E, Roulland S, Nadel B, Davi F, Armand M, Canioni D, Michot JM, Visco C, Arcaini L, et al. From hepatitis C virus infection to B-cell lymphoma. Ann Oncol. 2018;29(1):92–100.
4. Ponzetto A, Carloni G. Hepatitis C virus and lymphoma. Hepatology. 2016;64(5):1813.
5. Carbonari M, Caprini E, Tedesco T, Mazzetta F, Tocco V, Casato M, Russo G, Fiorilli M. Hepatitis C virus drives the unconstrained monoclonal expansion of VH1-69-expressing memory B cells in type II cryoglobulinemia: a model of infection-driven lymphomagenesis. J Immunol. 2005;174(10):6532–9.
6. Charles ED, Green RM, Marukian S, Talal AH, Lake-Bakaar GV, Jacobson IM, Rice CM, Dustin LB. Clonal expansion of immunoglobulin M + CD27 + B cells in HCV-associated mixed cryoglobulinemia. Blood. 2008;111(3):1344–56.
7. Robbiani DF, Deroubaix S, Feldhahn N, Oliveira TY, Callen E, Wang Q, Jankovic M, Silva IT, Rommel PC, Bosque D, et al. Plasmodium infection promotes genomic instability and AID-Dependent B cell lymphoma. Cell. 2015;162(4):727–37.
8. Dai B, Chen AY, Corkum CP, Peroutka RJ, Landon A, Houng S, Muniandy PA, Zhang Y, Lehrmann E, Mazan-Mamczarz K, et al. Hepatitis C virus upregulates B-cell receptor signaling: a novel mechanism for HCV-associated B-cell lymphoproliferative disorders. Oncogene. 2016;35(23):2979–90.
9. Pawelczyk A, Kubisa N, Jablonska J, Bukowska-Osko I, Caraballo Cortes K, Fic M, Laskus T, Radkowski M. Detection of hepatitis C virus (HCV) negative strand RNA and NS3 protein in peripheral blood mononuclear cells (PBMC): CD3+, CD14 + and CD19+. Virol J. 2013;10:346.
10. Arcaini L, Rossi D, Lucioni M, Nicola M, Bruscaggin A, Fiaccadori V, Riboni R, Ramponi A, Ferretti VV, Cresta S, et al. The NOTCH pathway is recurrently mutated in diffuse large B-cell lymphoma associated with hepatitis C virus infection. Haematologica. 2015;100(2):246–52.
11. Bowden J, Holmes MV. Meta-analysis and mendelian randomization: a review. Res Synth Methods. 2019;10(4):486–96.
12. Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. JAMA. 2017;318(19):1925–6.
13. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, VanderWeele TJ, Higgins JPT, Timpson NJ, Dimou N, et al. Strengthening the reporting of Observational studies in Epidemiology using mendelian randomization: the STROBE-MR Statement. JAMA. 2021;326(16):1614–21.
14. Sekula P, Del Greco MF, Pattaro C, Kottgen A. Mendelian randomization as an Approach to assess causality using Observational Data. J Am Soc Nephrol. 2016;27(11):3253–65.
15. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014;10(5):e1004383.
16. Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in mendelian randomization studies. Int J Epidemiol. 2011;40(3):755–64.
17. Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, Davey Smith G, Sterne JA. Using multiple genetic variants as instrumental variables for modifiable risk factors. Stat Methods Med Res. 2012;21(3):223–42.
18. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in mendelian randomization with some Invalid instruments using a weighted median estimator. Genet Epidemiol. 2016;40(4):304–14.
19. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44(2):512–25.
20. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from mendelian randomization between complex traits and diseases. Nat Genet. 2018;50(5):693–8.
21. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for mendelian randomization. Stat Methods Med Res. 2017;26(5):2333–55.
22. Burgess S, Thompson SG. Interpreting findings from mendelian randomization using the MR-Egger method. Eur J Epidemiol. 2017;32(5):377–89.
23. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F. The Ensembl variant effect predictor. Genome Biol. 2016;17(1):122.
24. Arcari A, Tabanelli V, Merli F, Marcheselli L, Merli M, Balzarotti M, Zilioli VR, Fabbri A, Cavallo F, Casaluci GM, et al. Biological features and outcome of diffuse large B-cell lymphoma associated with hepatitis C virus in elderly patients: results of the prospective 'Elderly Project' by the Fondazione Italiana Linfomi. Br J Haematol. 2023;201(4):653–62.
25. Iqbal T, Mahale P, Turturro F, Kyvernitakis A, Torres HA. Prevalence and association of hepatitis C virus infection with different types of lymphoma. Int J Cancer. 2016;138(4):1035–7.
26. Besson C, Canioni D, Lepage E, Pol S, Morel P, Lederlin P, Van Hoof A, Tilly H, Gaulard P, Coiffier B, et al. Characteristics and outcome of diffuse large B-cell lymphoma in hepatitis C virus-positive patients in LNH 93 and LNH 98 Groupe d'Etude Des Lymphomes De l'Adulte programs. J Clin Oncol. 2006;24(6):953–60.
27. Hosry J, Miranda RN, Samaniego F, Economides MP, Torres HA. Clinicopathologic characteristics and outcomes of transformed diffuse large B-cell lymphoma in hepatitis C virus-infected patients. Int J Cancer. 2018;142(5):940–8.
28. Anderson LA, Pfeiffer R, Warren JL, Landgren O, Gadalla S, Berndt SI, Ricker W, Parsons R, Wheeler W, Engels EA. Hematopoietic malignancies associated with viral and alcoholic hepatitis. Cancer Epidemiol Biomarkers Prev. 2008;17(11):3069–75.
29. Maciocia N, O'Brien A, Ardeshna K. Remission of Follicular Lymphoma after Treatment for Hepatitis C virus infection. N Engl J Med. 2016;375(17):1699–701.
30. Turner NC, Dusheiko G, Jones A. Hepatitis C and B-cell lymphoma. Ann Oncol. 2003;14(9):1341–5.
31. Feldmann G, Nischalke HD, Nattermann J, Banas B, Berg T, Teschendorf C, Schmiegel W, Duhrsen U, Halangk J, Iwan A, et al. Induction of interleukin-6 by hepatitis C virus core protein in hepatitis C-associated mixed cryoglobulinemia and B-cell non-hodgkin's lymphoma. Clin Cancer Res. 2006;12(15):4491–8.
32. Peveling-Oberhag J, Arcaini L, Hansmann ML, Zeuzem S. Hepatitis C-associated B-cell non-hodgkin lymphomas. Epidemiology, molecular signature and clinical management. J Hepatol. 2013;59(1):169–77.
33. Muka T, Glisic M, Milic J, Verhoog S, Bohlius J, Bramer W, Chowdhury R, Franco OH. A 24-step guide on how to design, conduct, and successfully publish a

Fu *et al. Infectious Agents and Cancer*   (2024) 19:15

Page 9 of 9

systematic review and meta-analysis in medical research. Eur J Epidemiol. 2020;35(1):49–60.
34. Siddaway AP, Wood AM, Hedges LV. How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, Meta-analyses, and Meta-syntheses. Annu Rev Psychol. 2019;70:747–70.
35. Ishikawa T, Shibuya K, Yasui K, Mitamura K, Ueda S. Expression of hepatitis C virus core protein associated with malignant lymphoma in transgenic mice. Comp Immunol Microbiol Infect Dis. 2003;26(2):115–24.
36. Kasama Y, Sekiguchi S, Saito M, Tanaka K, Satoh M, Kuwahara K, Sakaguchi N, Takeya M, Hiasa Y, Kohara M, et al. Persistent expression of the full genome of hepatitis C virus in B cells induces spontaneous development of B-cell lymphomas in vivo. Blood. 2010;116(23):4926–33.
37. Nocturne G, Boudaoud S, Besson C, Davi F, Canioni D, Cacoub P, Hermine O, Mariette X. Genetic variants of TNFAIP3 in patients with HCV related Lymphoma are Associated with the Presence of Rheumatoid factor (RF). Blood. 2014;124(21):1641.
38. Nguyen T, Liu XK, Zhang Y, Dong C. BTNL2, a butyrophilin-like molecule that functions to inhibit T cell activation. J Immunol. 2006;176(12):7354–60.
39. Rhodes DA, Reith W, Trowsdale J. Regulation of immunity by Butyrophilins. Annu Rev Immunol. 2016;34:151–72.
40. Du Y, Peng Q, Cheng D, Pan T, Sun W, Wang H, Ma X, He R, Zhang H, Cui Z, et al. Cancer cell-expressed BTNL2 facilitates tumour immune escape via engagement with IL-17A-producing gammadelta T cells. Nat Commun. 2022;13(1):231.
41. Waller RG, Klein RJ, Vijai J, McKay JD, Clay-Gilmour A, Wei X, Madsen MJ, Sborov DW, Curtin K, Slager SL, et al. Sequencing at lymphoid neoplasm susceptibility loci maps six myeloma risk genes. Hum Mol Genet. 2021;30(12):1142–53.
42. Vijai J, Wang Z, Berndt SI, Skibola CF, Slager SL, de Sanjose S, Melbye M, Glimelius B, Bracci PM, Conde L, et al. A genome-wide association study of marginal zone lymphoma shows association to the HLA region. Nat Commun. 2015;6:5751.
43. Rossi D, Bertoni F, Zucca E. Marginal-Zone Lymphomas. N Engl J Med. 2022;386(6):568–81.
44. Gragnani L, Lorini S, Marri S, Basile U, Santarlasci V, Monti M, Madia F, Petraccia L, Stasi C, Marello N, et al. Hematological and genetic markers in the Rational Approach to patients with HCV sustained Virological Response with or without persisting Cryoglobulinemic Vasculitis. Hepatology. 2021;74(3):1164–73.
45. Zignego AL, Wojcik GL, Cacoub P, Visentini M, Casato M, Mangia A, Latanich R, Charles ED, Gragnani L, Terrier B, et al. Genome-wide association study of hepatitis C virus- and cryoglobulin-related vasculitis. Genes Immun. 2014;15(7):500–5.
46. Papatheodorou I, Moreno P, Manning J, Fuentes AM, George N, Fexova S, Fonseca NA, Fullgrabe A, Green M, Huang N, et al. Expression Atlas update: from tissues to single cells. Nucleic Acids Res. 2020;48(D1):D77–83.
47. Volders PJ, Anckaert J, Verheggen K, Nuytens J, Martens L, Mestdagh P, Vandesompele J. LNCipedia 5: towards a reference set of human long non-coding RNAs. Nucleic Acids Res. 2019;47(D1):D135–9.
48. The RC. RNAcentral: a hub of information for non-coding RNA sequences. Nucleic Acids Res. 2019;47(D1):D221–9.
49. Anaparti V, Agarwal P, Smolik I, Mookherjee N, El-Gabalawy H. Whole blood targeted Bisulfite sequencing and Differential methylation in the C6ORF10 gene of patients with rheumatoid arthritis. J Rheumatol. 2020;47(11):1614–23.
50. Jin Y, Birlea SA, Fain PR, Gowan K, Riccardi SL, Holland PJ, Bennett DC, Herbstman DM, Wallace MR, McCormack WT, et al. Genome-wide analysis identifies a quantitative trait locus in the MHC class II region associated with generalized vitiligo age of onset. J Invest Dermatol. 2011;131(6):1308–12.

Publisher's Note