



OPEN ACCESS

EDITED BY

Kenta Nakai,
The University of Tokyo, Japan

REVIEWED BY

Marcos Edgar Herkenhoff,
University of São Paulo, Brazil
Yi Zheng,
The University of Tokyo, Japan

*CORRESPONDENCE

Xuena Wang,
✉ wangxuena_1214@163.com

[†]These authors have contributed equally to this work

RECEIVED 22 May 2024

ACCEPTED 28 August 2024

PUBLISHED 18 September 2024

CITATION

Wang N, Liu J, Chai B, Yao J, Du X, Mei Q and Wang X (2024) Bidirectional two-sample Mendelian randomization analysis investigates causal associations between cathepsins and inflammatory bowel disease.
Front. Genet. 15:1436407.
doi: 10.3389/fgene.2024.1436407

COPYRIGHT

© 2024 Wang, Liu, Chai, Yao, Du, Mei and Wang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Bidirectional two-sample Mendelian randomization analysis investigates causal associations between cathepsins and inflammatory bowel disease

Na Wang^{1†}, Jun Liu^{1†}, Bao Chai^{1†}, Jianhong Yao¹, Xufang Du¹, Qi Mei² and Xuena Wang^{1*}

¹Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Tongji Shanxi Hospital, Third Hospital of Shanxi Medical University, Taiyuan, China, ²Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

Background: Cathepsins, key regulators of the pathology of gastrointestinal disorders such as inflammatory bowel disease (IBD), are a target protease that has attracted much attention in recent years. IBD is a chronic and relapsing inflammatory disorder of the gut. Traditional studies have shown a correlation between cathepsin and the risk of IBD, while the causal relationship remains unclear.

Methods: This study utilized Mendelian randomization techniques to evaluate the causal relationships between eleven cathepsins and the subtypes of IBD, such as ulcerative colitis (UC) and Crohn's disease (CD). We also performed a series of sensitivity analyses to validate the primary Mendelian randomization (MR) results, including Cochran's Q test, the MR-PRESSO global test, and the MR pleiotropy test.

Results: The forward MR analyses showed no significant association between cathepsins and IBD. Reverse Mendelian randomization analyses suggested that UC might lead to elevated cathepsin G levels [inverse-variance weighted (IVW): $p = 0.038$, $b = 9.966$], and CD might cause a decrease in cathepsin B levels [IVW: $p = 0.002$, $b = -10.525$] and cathepsin L1 levels [IVW: $p = 0.045$, $b = -4.742$].

Conclusions: Our findings offer novel and comprehensive evidence on the impact of UC or CD on cathepsins, potentially providing valuable insights into the treatment and prognosis of IBD.

KEYWORDS

IBD, cathepsins, mendelian randomization, genetics, genome-wide association studies

1 Introduction

Inflammatory bowel disease (IBD) is a chronic inflammatory condition affecting the intestine, mainly defined as either ulcerative colitis (UC) or Crohn's disease (CD) (Guan, 2019; Rogler et al., 2021; Yang et al., 2021; Nóbrega et al., 2018; GBD, 2017 Inflammatory Bowel Disease Collaborators, 2020). In recent years, IBD has emerged as a significant global public health problem due to its elevated rates of morbidity and mortality, presenting a substantial threat to human health and well-being (Mak et al., 2020; Hasan et al., 2020;

Nielsen et al., 2022). Previous reports indicate that long-term IBD patients are at increased risk of gastrointestinal and extraintestinal malignancies. For instance, individuals with UC exhibit a heightened risk of developing pancreatic cancer (PC), with an odds ratio (OR) of 1.18 and a 95% confidence interval (CI) of 1.07–1.31⁹ (Huang et al., 2023). A Mendelian randomization study reported that IBD may be a risk factor in the development of oral cavity cancer (IBD: OR = 1.180, 95% CI: 1.059 to 1.316, $p = 0.003$; CD: OR = 1.112, 95% CI: 1.008 to 1.227, $p = 0.034$; UC: OR = 1.158, 95% CI: 1.041 to 1.288, $p = 0.007$) and breast cancer (IBD:OR = 1.059; 95% CI: 1.033 to 1.086; $p < 0.0001$; CD:OR = 1.029; 95% CI: 1.002 to 1.055; $p = 0.032$). Other studies have demonstrated a 17% greater overall cancer risk among IBD patients than those without the condition (Derikx et al., 2017; Singh et al., 2014; Faye et al., 2022).

Cathepsins are frequently used to catalyze the hydrolysis of unwanted proteins, play a key role in disease progression, and are important mediators of many inflammatory diseases (Reiser et al., 2010). Dysregulated cathepsins have been implicated in numerous pathologies, including arthritis, pancreatitis, and atherosclerosis (Yadati et al., 2020; Patel et al., 2018). Specific cathepsins have been identified as serological markers for certain diseases (Kyriakidi et al., 2016). In recent years, the role of cathepsins (such as cathepsin B, D, G, and L) has been widely studied in IBD (Menzel et al., 2006; Dong et al., 2022; Zhao et al., 2018). Studies reveal a significant increase in the cathepsin D protein expression in inflamed intestinal mucosa from IBD patients compared to non-inflamed mucosa (Hausmann et al., 2004). Inhibition of cathepsin D was followed by amelioration of mouse dextran sodium sulfate (DSS) colitis (Menzel et al., 2006); the level of anti-cathepsin G antibodies was significantly higher in patients with severe colitis than in those with mild or moderate colitis ($p < 0.05$) (Kuwana et al., 2000). Notably, Marta Dabek's study also verifies this point that cathepsin G has been found to be overexpressed in patients with UC compared to healthy controls (Dabek et al., 2009). Similar results have been found with other cathepsins, such as Zhao et al. (2018) and Zhao et al. (2016). However, our understanding of the association between other members of the cathepsin family and IBD remains limited, warranting further exploration and investigation. Meanwhile, due to the methodological constraints of observational studies, causal relationships between cathepsins and IBD cannot be conclusively established.

Genetics play a role in predisposition to IBD. The genetic risk ratio for IBD is in the range of 15–42 for CD (Monsén et al., 1991; Probert et al., 1993; Küster et al., 1989; Meucci et al., 1992; Orholm et al., 1991) and 7–17 for UC (Meucci et al., 1992; Orholm et al., 1991). Heredity as an etiological factor is stronger in Crohn's disease than in UC. Mendelian randomization (MR) analysis supplies a promising approach to exploring potential causal associations by leveraging instrumental variables (IVs) that are independent of common confounders, thus overcoming the limitations of traditional observational studies to some extent (Emdin et al., 2017; Smith and Ebrahim, 2004; Sekula et al., 2016). Recently, MR analysis has emerged as a tool for investigating relationships between IBD and various factors with the expanding size and scope of genome-wide association studies (GWAS) (Freuer et al., 2022; Zeng et al., 2023; Lund-Nielsen et al., 2018; Jones et al., 2020; Saadh et al., 2023). Long-term IBD patients have been reported to face an

up to three- to five-fold increased risk of developing colitis-associated colorectal cancer. To determine whether a potential causal relationship exists between cathepsins and IBD, we conducted a bidirectional two-sample MR analysis utilizing publicly available GWAS datasets.

2 Materials and methods

2.1 Study design

MR was employed to investigate the relationships between various cathepsins and IBD. The design of our study is shown in Figure 1. Initially, we extracted genetic variants that serve as IVs for 11 cathepsins. Subsequently, we collected the summary data, including all single-nucleotide polymorphisms (SNPs) derived from GWASs of UC and CD. Next, we performed bidirectional MR analyses with five distinct MR methods, including MR-Egger, weighted median, inverse-variance weighted (IVW), simple mode, and weighted mode. Furthermore, we conducted a series of sensitivity analyses, comprising Cochran's Q test, the MR-PRESSO global test, and the MR pleiotropy test, to evaluate the heterogeneity and horizontal pleiotropy of the MR results. Because the data employed in this study were based on published studies and public databases, no additional ethical approval from an institutional review board was necessary.

2.2 Data sources

The summary-level GWAS data of 11 cathepsins ($\mu\text{g/L}$) (including cathepsin B, D, E, F, G, H, L1, L2, O, S, and Z) and IBD (including UC and CD) were sourced from the Integrative Epidemiology Unit (IEU) open GWAS project (<https://gwas.mrcieu.ac.uk/>). This project, supported by the MRC IEU at the University of Bristol, collated and analyzed GWAS data from different sources, such as the UK Biobank, published articles, and the FinnGen Biobank. The approval of the Ethical Review Authority was not required because the data used in our study were public, anonymized, and de-identified. Further details regarding the exposure and outcome datasets are supplied in Table 1.

2.3 Selection of IVs

To satisfy the basic assumption of MR that IVs must be closely related to the exposure factors, SNPs that were significantly associated with various cathepsins ($\mu\text{g/L}$), including cathepsin B, D, E, F, G, H, L1, L2, O, S, and Z, at the genome-wide level ($p < 5 \times 10^{-6}$, $r^2 < 0.001$, genetic distance = 10,000 kb) were screened. The same criteria were applied for the reverse Mendelian randomization analysis related to ulcerative colitis and Crohn's disease. Subsequently, the strength of each IV was calculated by the following formula: $F = R^2(N-2)/1-R^2$, in which R^2 represents the proportion of variability in the cathepsin explained by each IV, and N is the sample size of the GWAS for the SNP–cathepsin association. F -statistics greater than 10 were generally considered to have a strong association (Staiger and Stock, 1997; Bowden et al., 2016a).

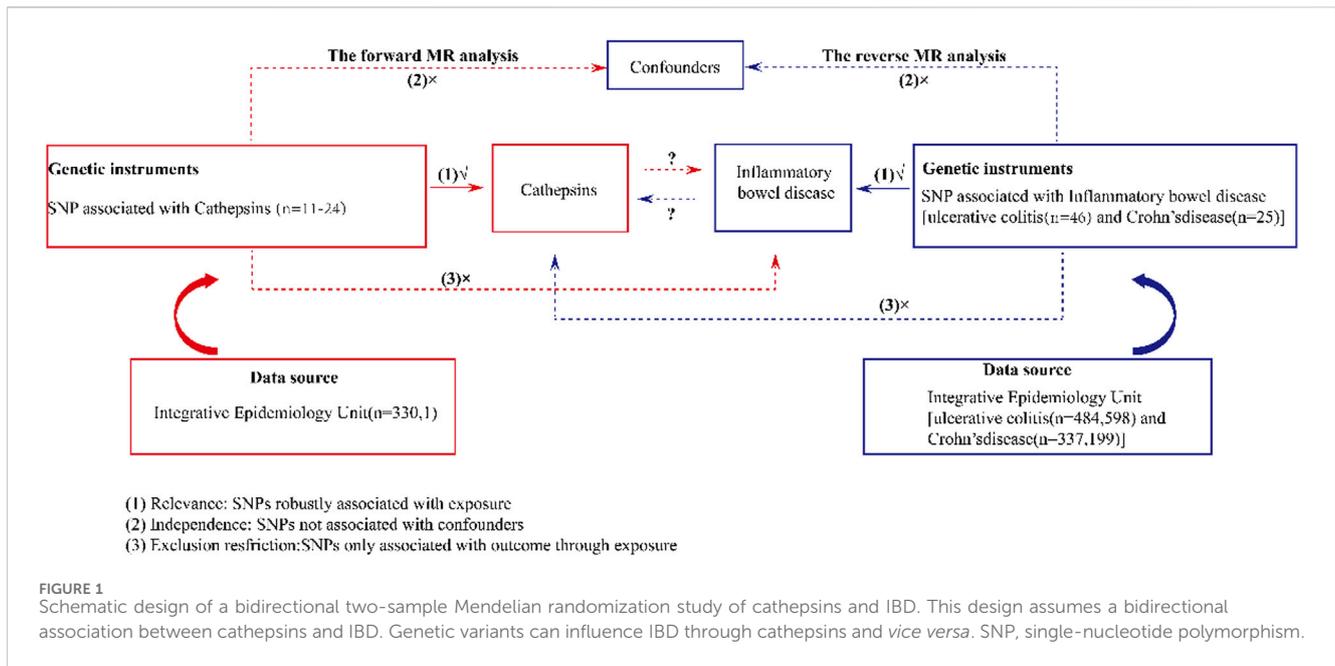


TABLE 1 Information about the exposures and outcome datasets.

IEU GWAS id	Exposure or outcome	Identified SNPs	Participants included in the analysis	F-statistic
prot-a-718	Cathepsin B	19	3,301 individuals of European descent	34.297
ebi-a-GCST90012053	Cathepsin D	20	21,758 individuals of European descent	92.98
prot-a-720	Cathepsin E	11	3,301 individuals of European descent	21.902
prot-a-722	Cathepsin F	13	3,301 individuals of European descent	35.678
prot-a-723	Cathepsin G	15	3,301 individuals of European descent	22.807
prot-a-725	Cathepsin H	11	3,301 individuals of European descent	126.409
ebi-a-GCST90012073	Cathepsin L1	32	21,758 individuals of European descent	43.212
prot-a-728	Cathepsin L2	11	3,301 individuals of European descent	24.669
prot-a-726	Cathepsin O	13	3,301 individuals of European descent	23.267
prot-a-727	Cathepsin S	24	3,301 individuals of European descent	39.691
prot-a-729	Cathepsin Z	13	3,301 individuals of European descent	53.426
ebi-a-GCST90038684	Ulcerative colitis	46	2,515 cases and 482,083 controls	32.190
ukb-a-552	Crohn's disease	25	732 cases and 336,467 controls	26.597

Abbreviation: SNPs, single-nucleotide polymorphisms; IEU, Integrative Epidemiology Unit; GWAS, Genome-wide association studies

SNPs in the MR analysis can be found in [Supplementary Figure S1](#) and [Supplementary Tables S1, S2](#).

2.4 Statistical analysis

First, the exposure and outcome data sets were harmonized, which ensured the reference allele (EA) remained consistently associated with the same allele across all analyses. Different MR methods (MR-Egger (Bowden et al., 2015), weighted median (Bowden et al., 2016b), IVW (Burgess et al., 2013), simple mode, and weighted mode (Hartwig et al., 2017)) were employed for the two-sample MR analysis to investigate the

causal relationship between cathepsins and IBD separately. The heterogeneity of the IVW model was evaluated by Cochran's Q test, and a Q test p -value <0.05 indicated significant heterogeneity. Outliers identified and addressed by the MR-PRESSO method were removed immediately once detected. Then, the MR analyses were performed again. The MR-Egger method, which allows for the existence of non-zero intercepts, was employed to detect directional pleiotropy (Verbanck et al., 2018). Additionally, a leave-one-out analysis was performed to ascertain whether the removal of individual SNPs significantly impacted the results. All statistical analyses were performed using R software (version 4.3.1) with the two-sample MR package (Yavorska and Burgess, 2017).

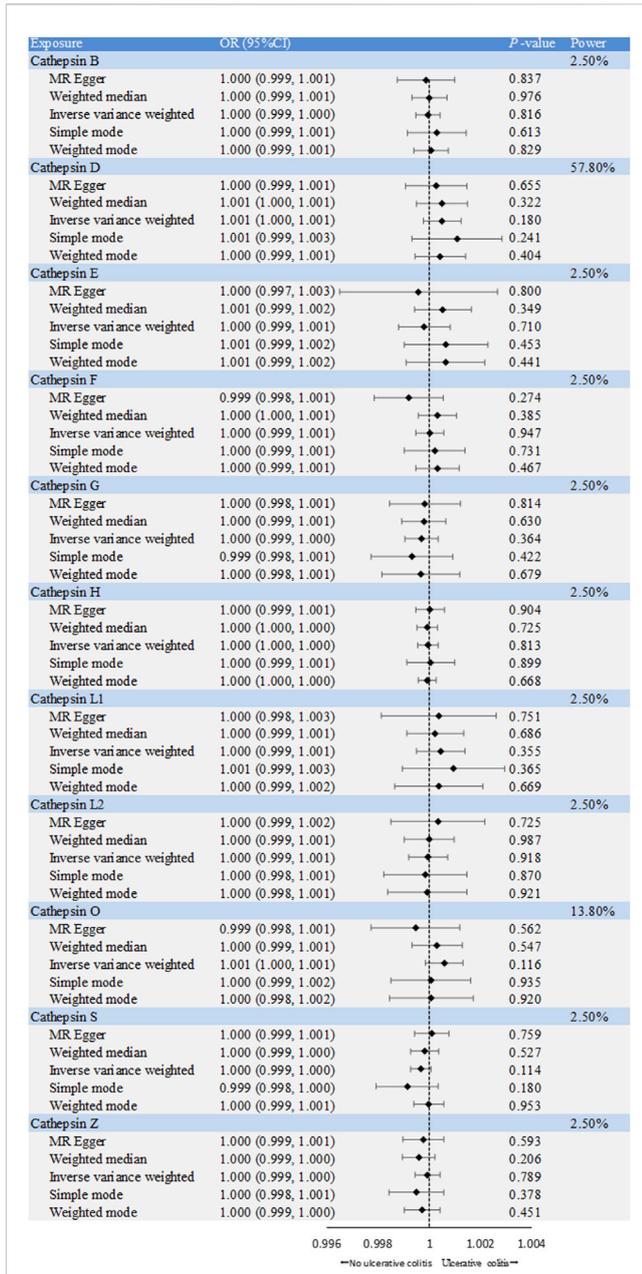


FIGURE 2 Forest plot of causal effects of cathepsin single-nucleotide polymorphisms (SNPs) on ulcerative colitis. We conducted inverse-variance weighted analyses, among others, to evaluate the causal relationship between cathepsins and ulcerative colitis. (error bars represent 95% confidence intervals).

Outcome	b (95%CI)	se	P-value	Power
Cathepsin B				
MR Egger	2.609 (-17.419,22.637)	10.218	0.800	21%
Weighted median	-0.915 (-15.021,13.191)	7.197	0.899	
Inverse variance weighted	0.363 (-8.378,9.104)	4.460	0.935	
Simple mode	-8.744 (-34.866,17.377)	13.327	0.515	
Weighted mode	-3.593 (-20.917,13.731)	8.839	0.686	
Cathepsin D				
MR Egger	3.743 (-7.508,14.994)	5.740	0.518	60.20%
Weighted median	2.035 (-3.430,7.500)	2.788	0.466	
Inverse variance weighted	0.272 (-3.725,4.270)	2.040	0.894	
Simple mode	4.336 (-6.326,14.999)	5.440	0.430	
Weighted mode	3.743 (-5.056,12.542)	4.489	0.409	
Cathepsin E				
MR Egger	13.822 (-6.591,34.235)	10.415	0.191	100%
Weighted median	6.366 (-8.738,21.470)	7.706	0.409	
Inverse variance weighted	6.804 (-2.157,15.765)	4.572	0.137	
Simple mode	15.547 (-14.488,45.582)	15.324	0.316	
Weighted mode	10.836 (-8.594,30.267)	9.913	0.280	
Cathepsin F				
MR Egger	-4.564 (-24.017,14.889)	9.925	0.648	100%
Weighted median	-4.403 (-17.738,8.932)	6.803	0.517	
Inverse variance weighted	-8.100 (-16.681,0.482)	4.378	0.064	
Simple mode	3.896 (-23.530,31.321)	13.993	0.782	
Weighted mode	-0.986 (-19.057,17.084)	9.220	0.915	
Cathepsin G				
MR Egger	0.590 (-20.766,21.946)	10.896	0.957	100%
Weighted median	4.301 (-9.304,17.907)	6.941	0.535	
Inverse variance weighted	9.966 (0.553,19.378)	4.802	0.038	
Simple mode	13.681 (-11.879,39.241)	13.041	0.300	
Weighted mode	2.939 (-15.238,21.116)	9.274	0.753	
Cathepsin H				
MR Egger	-1.094 (-20.547,18.359)	9.925	0.913	100%
Weighted median	7.733 (-6.422,21.888)	7.222	0.284	
Inverse variance weighted	2.931 (-5.651,11.513)	4.379	0.503	
Simple mode	8.496 (-16.261,33.252)	12.631	0.505	
Weighted mode	8.943 (-8.838,26.724)	9.072	0.330	
Cathepsin L1				
MR Egger	2.036 (-10.113,14.185)	6.198	0.744	100%
Weighted median	-0.358 (-6.081,5.365)	2.920	0.902	
Inverse variance weighted	2.671 (-1.894,7.237)	2.329	0.251	
Simple mode	-2.350 (-14.339,9.638)	6.116	0.703	
Weighted mode	-3.670 (-14.308,6.969)	5.428	0.503	
Cathepsin L2				
MR Egger	-1.847 (-21.301,17.607)	9.926	0.853	36.60%
Weighted median	2.242 (-11.656,16.140)	7.091	0.752	
Inverse variance weighted	-0.509 (-9.091,8.073)	4.379	0.907	
Simple mode	12.179 (-14.123,38.480)	13.419	0.369	
Weighted mode	3.193 (-15.766,22.151)	9.673	0.743	
Cathepsin O				
MR Egger	1.919 (-20.131,23.970)	11.250	0.865	100%
Weighted median	12.174 (-1.346,25.694)	6.898	0.078	
Inverse variance weighted	3.638 (-5.982,13.257)	4.908	0.459	
Simple mode	10.145 (-19.539,39.829)	15.145	0.506	
Weighted mode	12.054 (-5.948,30.056)	9.185	0.196	
Cathepsin S				
MR Egger	23.952 (4.497,43.407)	9.926	0.020	100%
Weighted median	5.438 (-7.223,18.099)	6.460	0.400	
Inverse variance weighted	3.560 (-5.362,12.482)	4.552	0.434	
Simple mode	-18.194 (-46.680,10.291)	14.534	0.217	
Weighted mode	12.414 (-7.481,32.308)	10.150	0.228	
Cathepsin Z				
MR Egger	-16.552 (-35.993,2.889)	9.919	0.102	100%
Weighted median	-4.134 (-16.598,8.330)	6.359	0.516	
Inverse variance weighted	-6.788 (-15.370,1.794)	4.379	0.121	
Simple mode	-1.619 (-30.882,27.643)	14.930	0.914	
Weighted mode	-9.031 (-28.645,10.584)	10.007	0.372	

FIGURE 3 Two-sample Mendelian randomization study based on the MR method estimates the causal effects of ulcerative colitis on cathepsins.

3 Results

3.1 Defining the causal relationship between various cathepsins and UC

The causal associations between eleven types of cathepsins (cathepsin B, E, F, G, H, L2, O, S, and Z) and IBD were analyzed. The forward MR analysis did not reveal any causal associations between eleven types of cathepsins and the risk of IBD (Figure 2).

We conducted reverse MR analyses to explore the possibility of reverse causality. The reverse MR analysis provided evidence that UC elevated cathepsin G levels (IVW: $p = 0.038$, $b = 9.966$) (Figures 3, 4), and the p -values of the Cochran's Q test, the MR-PRESSO global test, and the MR-Egger intercept showing no signs of heterogeneity and directional pleiotropy (0.167, 0.168, and 0.343, respectively). The statistical power of this analysis reached 100%. The effect estimate from MR-Egger is 0.590, while

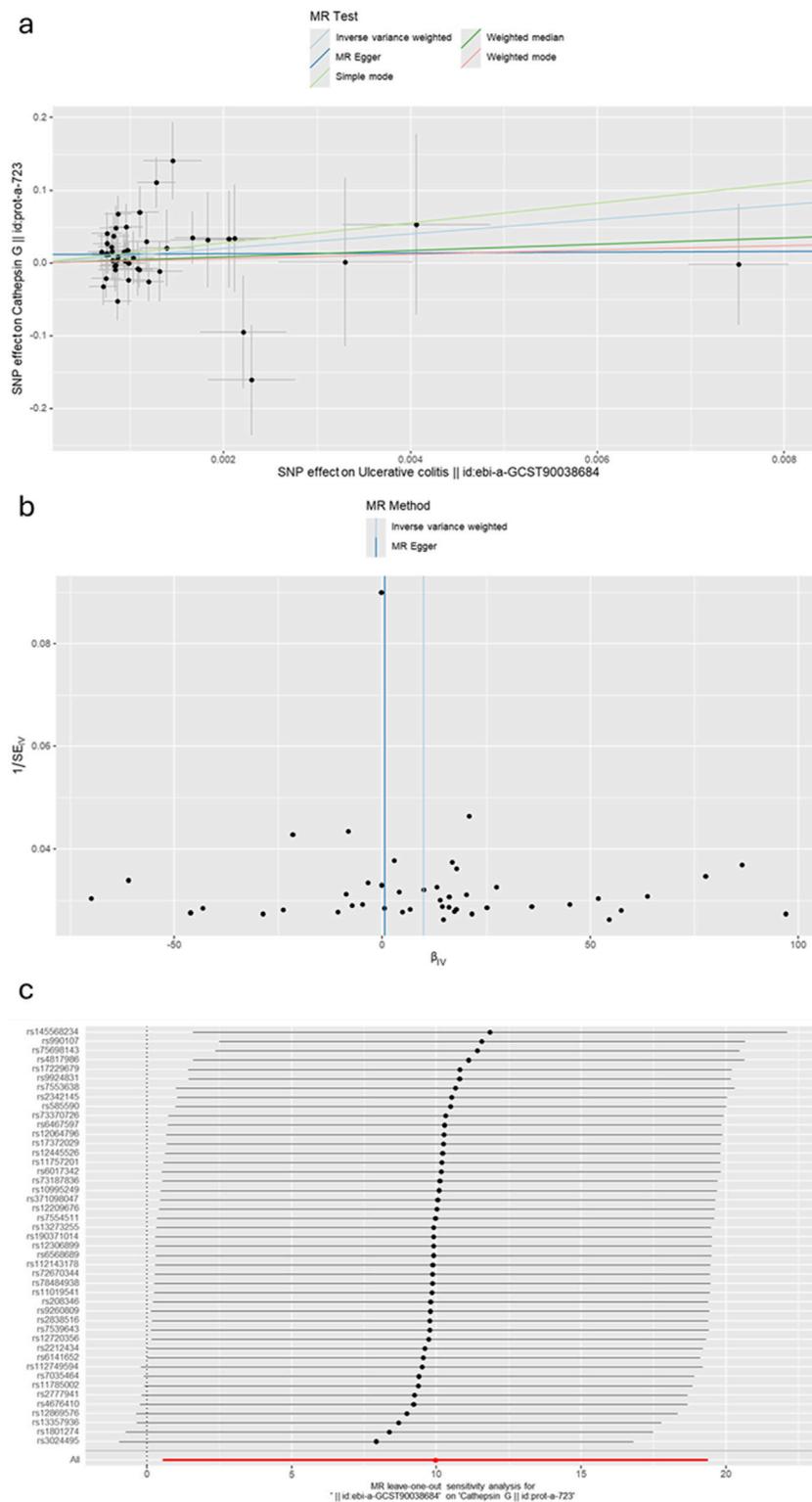


FIGURE 4
(A) A scatter plot represents genetic IV association between ulcerative colitis and cathepsin G. **(B)** A funnel plot of SNPs associated with ulcerative colitis and cathepsin G. **(C)** A leave-one-out analysis of the estimations for ulcerative colitis and cathepsin G.

the IVW estimate is 9.966. Although the *p*-value for the MR-Egger intercept test is 0.343, we note that the b-value from MR-Egger (0.590) is considerably smaller than its standard error

(9.966), which may suggest potential pleiotropy. No evidence supported a causal association between UC and other types of cathepsins (Figure 3).

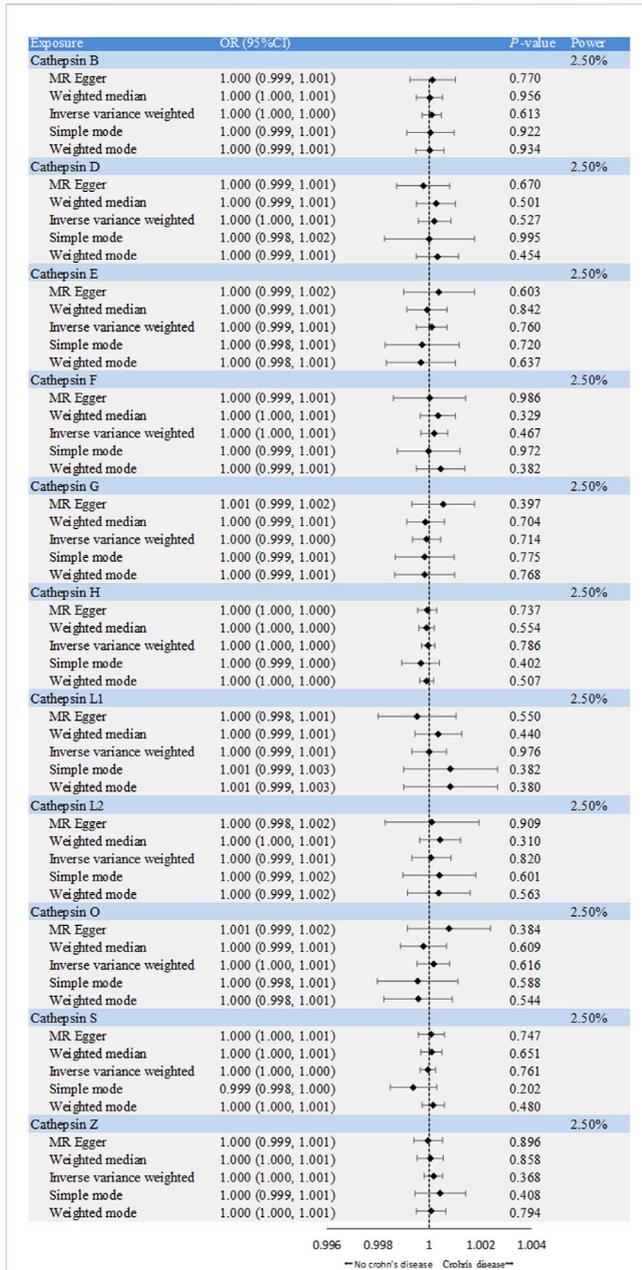


FIGURE 5 Forest plot of causal effects of cathepsin single-nucleotide polymorphisms (SNPs) on Crohn's disease. We conducted inverse-variance weighted analyses to evaluate the causal relationship between cathepsins and Crohn's disease. (error bars represent 95% confidence intervals).

3.2 Defining the causal relationship between various cathepsins and CD

The causal associations between eleven types of cathepsins (cathepsin B, D, E, F, G, H, L1, L2, O, S, and Z) and CD were analyzed. The forward MR analysis did not reveal any causal associations between eleven types of cathepsins and CD risk (Figure 5). We conducted reverse MR analyses to explore the possibility of reverse causality. The reverse MR analysis provided evidence that CD decreased the level of cathepsin B (cathepsin B:

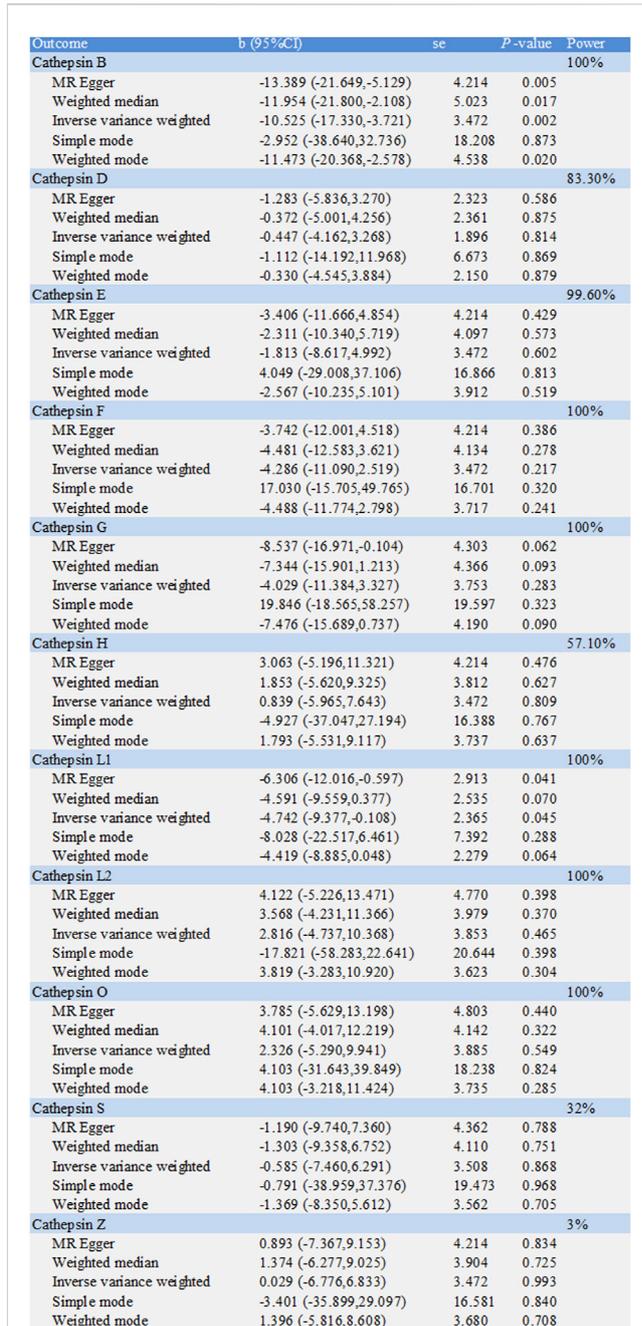


FIGURE 6 Two-sample Mendelian randomization study based on the MR method estimates the causal effects of Crohn's disease on cathepsins.

IVW: $p = 0.002$, $b = -10.525$) (Figures 6, 7). The p -values of Cochran's Q test, the MR-PRESSO global test, and the MR-Egger intercept showed no signs of heterogeneity, outliers, or directional pleiotropy (0.828, 0.709, and 0.245, respectively). The reverse MR analysis provided evidence that CD decreased the level of cathepsin L1 (IVW: $p = 0.045$, $b = -4.742$) (Figure 8), and the p -value of the Cochran's Q test was 0.047, suggesting the presence of slight but statistically significant heterogeneity among the genetic instruments deployed. The p -values of the MR-PRESSO global test and the MR-Egger intercept showed no signs of outliers or directional pleiotropy (0.147 and 0.365,

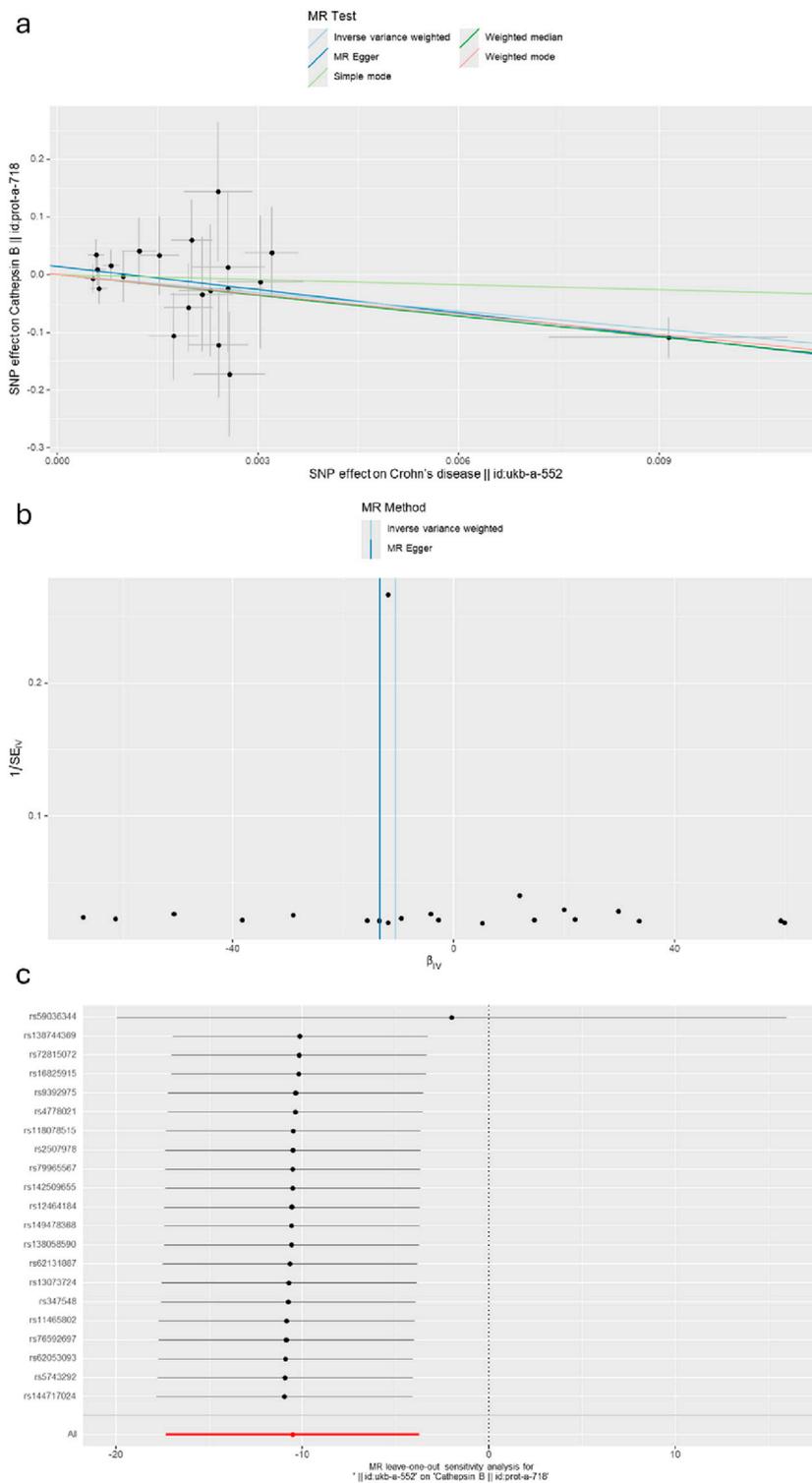


FIGURE 7 (A) A scatter plot represents genetic IV association between Crohn's disease and cathepsin B. (B) A funnel plot of SNPs associated with Crohn's disease and cathepsin B. (C) A leave-one-out analysis of the estimations for Crohn's disease and cathepsin B.

respectively). The statistical power of these analyses reached 100%. No evidence supported a causal association between CD and other types of cathepsins (Figure 6).

Overall, in primary MR analyses, IBD was statistically associated with cathepsin B and cathepsin G, suggesting the development of IBD might affect the level of these cathepsins.

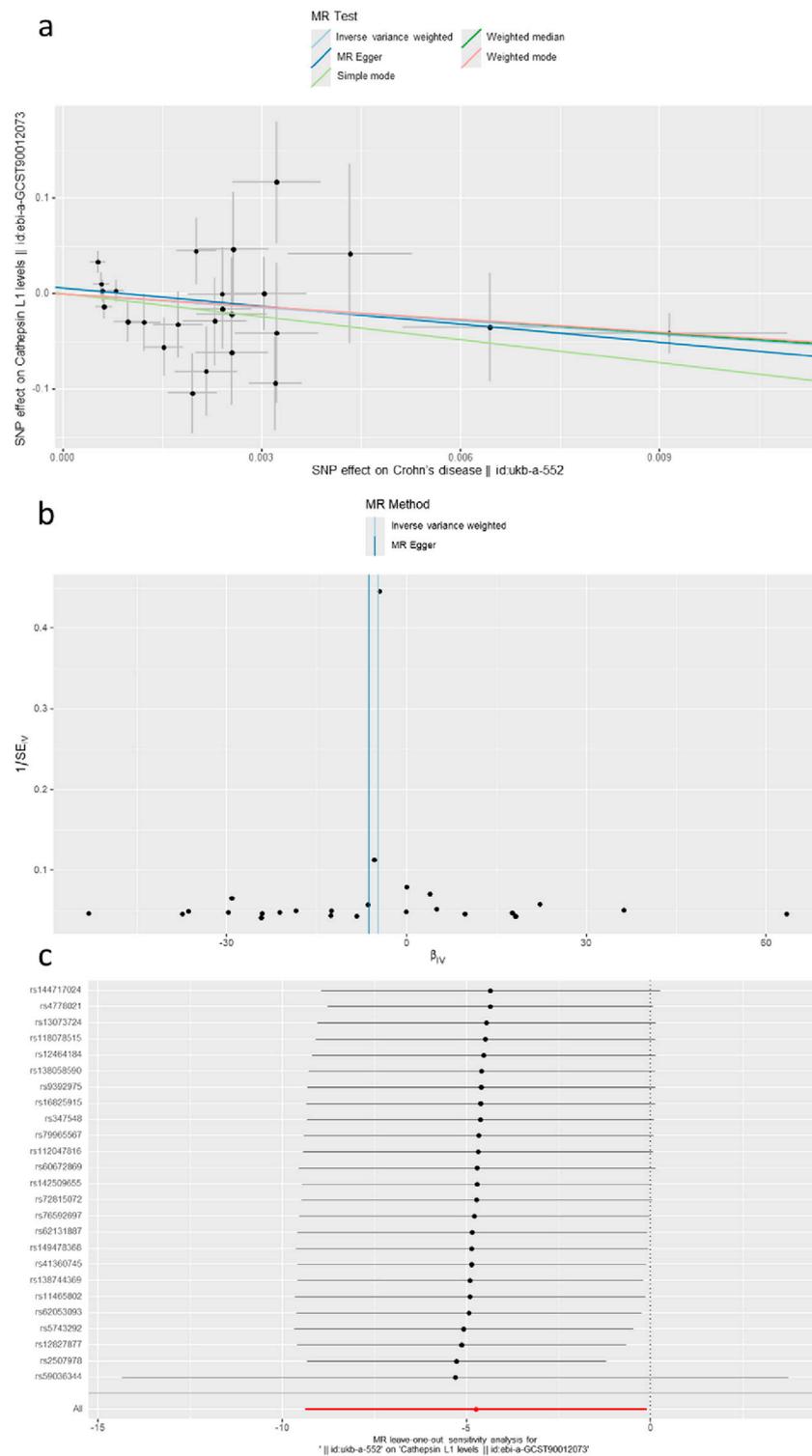


FIGURE 8 (A) A scatter plot represents the genetic IV association between Crohn’s disease and cathepsin L1. (B) A funnel plot of SNPs associated with Crohn’s disease and cathepsin L1. (C) A leave-one-out analysis of the estimations for Crohn’s disease and cathepsin L1.

4 Discussion

IBD is a prevalent chronic immune-mediated inflammatory disorder of the digestive tract that, if left untreated, can lead to

prolonged bowel damage. Recent observational analyses have suggested that intestinal pathophysiological conditions can disrupt proteolytic homeostasis, resulting in inappropriate substrate degradation and abnormal accumulation (Sekula et al.,

2016). Among the key players in these processes, cathepsins have garnered significant attention from researchers. In this study, we conducted a systematic analysis of the causal relationship between eleven distinct cathepsins and IBD using genetic instruments. Our findings revealed that cathepsin levels do not significantly influence the onset or progression of IBD. Reverse Mendelian randomization analyses suggested that ulcerative colitis (UC) might lead to elevated cathepsin G levels, while Crohn's disease (CD) may cause a decrease in cathepsin B and L1 levels.

Previous observational research partially corroborates our findings regarding the association between UC and cathepsin G. Marta Dabek et al. demonstrated through RT-PCR and Western blotting that Cat-G is overexpressed in UC patients, with higher mRNA expression in biopsies than healthy subjects (0.44 ± 0.08 vs 0.24 ± 0.03 , $p < 0.05$) (Dabek et al., 2009). Additionally, Kuwana's study showed that anti-neutrophil cytoplasmic antibodies targeting Cat-G are present in the sera of more than half of UC patients (Kuwana et al., 2000). These findings suggest that Cat-G may play a crucial role in UC pathogenesis. However, due to the larger uncertainty in the MR-Egger estimate, we cannot entirely rule out the presence of pleiotropy. Thus, further investigation is necessary to confirm the causal relationship between UC and cathepsin G.

Interestingly, our study revealed that CD occurrence correlates with lower levels of cathepsin B and cathepsin L1, which contradicts previous observational research and clinical studies. Menzel et al. confirmed the upregulation of cathepsin B and cathepsin L in areas of tissue damage and mucosal ulceration in IBD patients using immunohistochemistry and gene expression analysis (Menzel et al., 2006). Moreover, elevated cathepsin B levels have been observed in various pathological conditions, including inflammation, infection, and cancer. Benjamin Bian et al. reported increased mRNA and activated levels of cathepsin B in human adenomas and colorectal cancers (CRCs) at all stages (Bian et al., 2016). Zhenhu Zhang's study demonstrated significantly higher expression levels of CTSL in ESCC tissues than adjacent non-cancerous tissues (Zhang et al., 2024); Shengnan Zhao's research revealed increased cathepsin B levels in colonic tissue and exacerbated dextran sodium sulfate (DSS)-induced colitis (Zhao et al., 2016). The discrepancy between our findings and previous observational studies could be attributed to bias from reverse causality or residual confounding in the latter. These divergent results underscore the need for further research to elucidate the complex relationship between cathepsins and IBD.

This study exhibits several notable strengths and limitations. First, MR leverages genetic variants to estimate the causal effects of circulating proteins on IBD, effectively overcoming the bias caused by reverse causality and confounding. Second, we performed sensitivity and pleiotropic analyses to ensure the accuracy of MR analysis. We used exposures and outcomes in European populations from different countries to minimize the potential for population stratification bias. Despite these insights, several limitations warrant consideration, such as (1) the potential overlap of participants between the exposure and outcome GWAS in the two-sample MR analyses could not be ascertained in this study; (2) the presence of pleiotropy, where instrumental SNPs may influence multiple traits, could not be entirely ruled out; however, no evidence of pleiotropy was observed in the MR analyses conducted in any of the above-mentioned MR approaches; (3) given that our study

cohort comprised exclusively of individuals of European ancestry, caution should be exercised when generalizing these findings to broader populations.

In conclusion, this study's primary genetic evidence reveals a nuanced relationship between cathepsins and IBD. Cathepsins do not appear to have a direct impact on IBD. While the occurrence of UC may be associated with elevated levels of cathepsin G, CD was linked to lower levels of cathepsins B and L1. Future research should aim to definitively establish and validate these causal relationships. These insights may prove valuable in identifying biochemical markers for the prediction, screening, early diagnosis, and prognosis of IBD.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

Author contributions

NW: writing—original draft, methodology, writing—review and editing, conceptualization, formal analysis, funding acquisition, investigation, resources, software, validation, and visualization. JL: conceptualization, investigation, validation, formal analysis, methodology, software, resources, and writing—original draft. BC: supervision, data curation, methodology, conceptualization, software, writing—original draft, investigation, writing—review and editing, and resources. JY: visualization, methodology, validation, writing—original draft, and investigation. XD: visualization, validation, writing—original draft, methodology, and investigation. QM: methodology, investigation, supervision, software, resources and writing—review and editing. XW: investigation, visualization, writing—original draft, formal analysis, resources, data curation, project administration, funding acquisition, conceptualization, supervision, validation, and writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. Grant information is as follows: NW is supported by the Talent Introduction and Research Initiation Project from Shanxi Province of China (2023RC48), XW is supported by the Talent Introduction and Research Initiation Project from Shanxi, and QM is supported by the 2023 COVID-19 Emergency Project of Shanxi Health Commission (2023XG02) and the 2023 annual translational open project of the National Emergency Medical Rescue Base at Shanxi Bethune Hospital (JYJD 2023-03).

Acknowledgments

The authors gratefully acknowledge all the people who have made this study possible.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2024.1436407/full#supplementary-material>

References

- Bian, B., Mongrain, S., Cagnol, S., Langlois, M. J., Boulanger, J., Bernatchez, G., et al. (2016). Cathepsin B promotes colorectal tumorigenesis, cell invasion, and metastasis. *Mol. Carcinog.* 55 (5), 671–687. doi:10.1002/mc.22312
- Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* 44 (2), 512–525. doi:10.1093/ije/dyv080
- Bowden, J., Davey Smith, G., Haycock, P. C., and Burgess, S. (2016b). Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* 40 (4), 304–314. doi:10.1002/gepi.21965
- Bowden, J., Del Greco, M. F., Minelli, C., Davey Smith, G., Sheehan, N. A., and Thompson, J. R. (2016a). Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I² statistic. *Int. J. Epidemiol.* 45 (6), 1961–1974. doi:10.1093/ije/dyw220
- Burgess, S., Butterworth, A., and Thompson, S. G. (2013). Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* 37 (7), 658–665. doi:10.1002/gepi.21758
- Dabek, M., Ferrier, L., Roka, R., Gecse, K., Annahazi, A., Moreau, J., et al. (2009). Luminal cathepsin g and protease-activated receptor 4: a duet involved in alterations of the colonic epithelial barrier in ulcerative colitis. *Am. J. Pathol.* 175 (1), 207–214. doi:10.2353/ajpath.2009.080986
- Derikx, L. A., Smits, L. J., van Vliet, S., Dekker, E., Aalfs, C. M., van Kouwen, M. C., et al. (2017). Colorectal cancer risk in patients with lynch syndrome and inflammatory bowel disease. *Clin. Gastroenterol. Hepatol.* 15 (3), 454–458. doi:10.1016/j.cgh.2016.08.005
- Dong, L., Xie, J., Wang, Y., Jiang, H., Chen, K., Li, D., et al. (2022). Mannose ameliorates experimental colitis by protecting intestinal barrier integrity. *Nat. Commun.* 13 (1), 4804. doi:10.1038/s41467-022-32505-8
- Emdin, C. A., Khera, A. V., and Kathiresan, S. (2017). Mendelian randomization. *Jama* 318 (19), 1925–1926. doi:10.1001/jama.2017.17219
- Faye, A. S., Holmer, A. K., and Axelrad, J. E. (2022). Cancer in inflammatory bowel disease. *Gastroenterol. Clin. North Am.* 51 (3), 649–666. doi:10.1016/j.gtc.2022.05.003
- Freuer, D., Linseisen, J., and Meisinger, C. (2022). Asthma and the risk of gastrointestinal disorders: a Mendelian randomization study. *BMC Med.* 20 (1), 82. doi:10.1186/s12916-022-02283-7
- GBD 2017 Inflammatory Bowel Disease Collaborators (2020). The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol. Hepatol.* 5 (1), 17–30. doi:10.1016/S2468-1253(19)30333-4
- Guan, Q. (2019). A comprehensive review and update on the pathogenesis of inflammatory bowel disease. *J. Immunol. Res.* 2019, 7247238. doi:10.1155/2019/7247238
- Hartwig, F. P., Davey Smith, G., and Bowden, J. (2017). Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.* 46 (6), 1985–1998. doi:10.1093/ije/dyx102
- Hasan, B., Senada, P. A., Wallace, M. B., Iacucci, M., and Shen, B. (2020). “Chapter 19 - chromoendoscopy in inflammatory bowel disease,” in *Atlas of endoscopy imaging in inflammatory bowel disease*. Editor B. Shen (Academic Press), 265–309.
- Hausmann, M., Obermeier, F., Schreiter, K., Spottl, T., Falk, W., Schölmerich, J., et al. (2004). Cathepsin D is up-regulated in inflammatory bowel disease macrophages. *Clin. Exp. Immunol.* 136 (1), 157–167. doi:10.1111/j.1365-2249.2004.02420.x
- Huang, J., Li, X., Hong, J., Huang, L., Jiang, Q., Guo, S., et al. (2023). Inflammatory bowel disease increases the risk of hepatobiliary pancreatic cancer: a two-sample Mendelian randomization analysis of European and East Asian populations. *Cancer Med.* 12 (12), 13599–13609. doi:10.1002/cam4.6057
- Jones, D. P., Richardson, T. G., Davey Smith, G., Gunnell, D., Munafò, M. R., and Wootton, R. E. (2020). Exploring the effects of cigarette smoking on inflammatory bowel disease using mendelian randomization. *Crohn's Colitis* 360 2 (1), otaa018. doi:10.1093/crocol/otaa018
- Küster, W., Pascoe, L., Purrmann, J., Funk, S., and Majewski, F. (1989). The genetics of Crohn disease: complex segregation analysis of a family study with 265 patients with Crohn disease and 5,387 relatives. *Am. J. Med. Genet.* 32 (1), 105–108. doi:10.1002/ajmg.1320320122
- Kuwana, T., Sato, Y., Saka, M., Kondo, Y., Miyata, M., Obara, K., et al. (2000). Anti-cathepsin G antibodies in the sera of patients with ulcerative colitis. *J. Gastroenterol.* 35 (9), 682–689. doi:10.1007/s005350070047
- Kyriakidi, K. S., Tsianos, V. E., Karvounis, E., Christodoulou, D. K., Katsanos, K. H., and Tsianos, E. V. (2016). Neutrophil anti-neutrophil cytoplasmic autoantibody proteins: bactericidal increasing protein, lactoferrin, cathepsin, and elastase as serological markers of inflammatory bowel and other diseases. *Ann. Gastroenterol.* 29 (3), 258–267. doi:10.20524/aog.2016.0028
- Lund-Nielsen, J., Vedel-Krogh, S., Kobylecki, C. J., Brynskov, J., Afzal, S., and Nordestgaard, B. G. (2018). Vitamin D and inflammatory bowel disease: mendelian randomization analyses in the Copenhagen studies and UK Biobank. *J. Clin. Endocrinol. Metab.* 103 (9), 3267–3277. doi:10.1210/je.2018-00250
- Mak, W. Y., Zhao, M., Ng, S. C., and Burisch, J. (2020). The epidemiology of inflammatory bowel disease: east meets west. *J. Gastroenterol. Hepatol.* 35 (3), 380–389. doi:10.1111/jgh.14872
- Menzel, K., Hausmann, M., Obermeier, F., Schreiter, K., Dunger, N., Bataille, F., et al. (2006). Cathepsins B, L and D in inflammatory bowel disease macrophages and potential therapeutic effects of cathepsin inhibition *in vivo*. *Clin. Exp. Immunol.* 146 (1), 169–180. doi:10.1111/j.1365-2249.2006.03188.x
- Meucci, G., Vecchi, M., Torgano, G., Arrigoni, M., Prada, A., Rocca, F., et al. (1992). Familial aggregation of inflammatory bowel disease in northern Italy: a multicenter study. The Gruppo di Studio per le Malattie Infiammatorie Intestinali (IBD Study Group). *Gastroenterology* 103 (2), 514–519. doi:10.1016/0016-5085(92)90841-1
- Monsén, U., Bernell, O., Johansson, C., and Hellers, G. (1991). Prevalence of inflammatory bowel disease among relatives of patients with Crohn's disease. *Scand. J. gastroenterology* 26 (3), 302–306. doi:10.3109/00365529109025046
- Nielsen, D. L., Juhl, C. B., Chen, I. M., Kellermann, L., and Nielsen, O. H. (2022). Immune checkpoint inhibitor-induced diarrhea and colitis: incidence and management. A systematic review and meta-analysis. *Cancer Treat. Rev.* 109, 102440. doi:10.1016/j.ctrv.2022.102440
- Nóbrega, V. G., Silva, I. N. N., Brito, B. S., Silva, J., Silva, M., and Santana, G. O. (2018). THE ONSET OF CLINICAL MANIFESTATIONS IN INFLAMMATORY BOWEL DISEASE PATIENTS. *Arq. Gastroenterol.* 55 (3), 290–295. doi:10.1590/S0004-2803.201800000-73
- Orholm, M., Munkholm, P., Langholz, E., Nielsen, O. H., Sørensen, T. I., and Binder, V. (1991). Familial occurrence of inflammatory bowel disease. *N. Engl. J. Med.* 324 (2), 84–88. doi:10.1056/NEJM199101103240203
- Patel, S., Homaei, A., El-Seedi, H. R., and Akhtar, N. (2018). Cathepsins: proteases that are vital for survival but can also be fatal. *Biomed. Pharmacother.* 105, 526–532. doi:10.1016/j.biopha.2018.05.148
- Probert, C. S., Jayanthi, V., Hughes, A. O., Thompson, J. R., Wicks, A. C., and Mayberry, J. F. (1993). Prevalence and family risk of ulcerative colitis and Crohn's disease: an epidemiological study among Europeans and south Asians in Leicestershire. *Gut* 34 (11), 1547–1551. doi:10.1136/gut.34.11.1547
- Reiser, J., Adair, B., and Reinheckel, T. (2010). Specialized roles for cysteine cathepsins in health and disease. *J. Clin. Invest.* 120 (10), 3421–3431. doi:10.1172/JCI42918
- Rogler, G., Singh, A., Kavanaugh, A., and Rubin, D. T. (2021). Extraintestinal manifestations of inflammatory bowel disease: current concepts, treatment, and implications for disease management. *Gastroenterology* 161 (4), 1118–1132. doi:10.1053/j.gastro.2021.07.042

- Saadh, M. J., Pal, R. S., Arias-González, J. L., Orosco Gavilán, J. C., Jc, D., Mohany, M., et al. (2023). A mendelian randomization analysis investigates causal associations between inflammatory bowel diseases and variable risk factors. *Nutrients* 15 (5), 1202. doi:10.3390/nu15051202
- Sekula, P., Del Greco, M. F., Pattaro, C., and Köttgen, A. (2016). Mendelian randomization as an approach to assess causality using observational data. *J. Am. Soc. Nephrol.* 27 (11), 3253–3265. doi:10.1681/ASN.2016010098
- Singh, S., Naggal, S. J., Murad, M. H., Yadav, S., Kane, S. V., Pardi, D. S., et al. (2014). Inflammatory bowel disease is associated with an increased risk of melanoma: a systematic review and meta-analysis. *Clin. Gastroenterol. Hepatol.* 12 (2), 210–218. doi:10.1016/j.cgh.2013.04.033
- Smith, G. D., and Ebrahim, S. (2004). Mendelian randomization: prospects, potentials, and limitations. *Int. J. Epidemiol.* 33 (1), 30–42. doi:10.1093/ije/dyh132
- Staiger, D., and Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica* 65 (3), 557–586. doi:10.2307/2171753
- Verbanck, M., Chen, C. Y., Neale, B., and Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* 50 (5), 693–698. doi:10.1038/s41588-018-0099-7
- Yadati, T., Houben, T., Bitorina, A., and Shiri-Sverdlov, R. (2020). The ins and outs of cathepsins: physiological function and role in disease management. *Cells* 9 (7), 1679. doi:10.3390/cells9071679
- Yang, M., Gu, Y., Li, L., Liu, T., Song, X., Sun, Y., et al. (2021). Bile acid-gut microbiota Axis in inflammatory bowel disease: from bench to bedside. *Nutrients* 13 (9), 3143. doi:10.3390/nu13093143
- Yavorska, O. O., and Burgess, S. (2017). MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int. J. Epidemiol.* 46 (6), 1734–1739. doi:10.1093/ije/dyx034
- Zeng, R., Wang, J., Jiang, R., Yang, J., Zheng, C., Wu, H., et al. (2023). Investigating causality and shared genetic architecture between neurodegenerative disorders and inflammatory bowel disease. *Aging Dis.* 14 (4), 1349–1359. doi:10.14336/AD.2022.12209
- Zhang, Z., Wang, J., Shi, Y., Wang, B., and Wang, D. (2024). Cathepsin L promotes oesophageal squamous cell carcinoma development and may be associated with tumour-associated macrophages. *Heliyon* 10 (7), e29273. doi:10.1016/j.heliyon.2024.e29273
- Zhao, S., Gong, Z., Du, X., Tian, C., Wang, L., Zhou, J., et al. (2018). Deoxycholic acid-mediated sphingosine-1-phosphate receptor 2 signaling Exacerbates DSS-induced colitis through promoting cathepsin B Release. *J. Immunol. Res.* 2018, 2481418. doi:10.1155/2018/2481418
- Zhao, S., Gong, Z., Zhou, J., Tian, C., Gao, Y., Xu, C., et al. (2016). Deoxycholic Acid triggers NLRP3 inflammasome activation and aggravates DSS-induced colitis in mice. *Front. Immunol.* 7, 536. doi:10.3389/fimmu.2016.00536