



## OPEN ACCESS

## EDITED BY

Shilpa S. Dhar,  
University of Texas MD Anderson Cancer  
Center, United States

## REVIEWED BY

Jie Peng,  
Central South University, China  
Tianqi Lu,  
Chengdu Third People's Hospital, China

## \*CORRESPONDENCE

Tiejun Liu  
✉ liutiejun699@163.com

RECEIVED 03 January 2024

ACCEPTED 14 March 2024

PUBLISHED 25 March 2024

## CITATION

Huang X, Deng H, Zhang B, Wang K, Qu Y,  
Li T and Liu T (2024) The causal relationship  
between cathepsins and digestive system  
tumors: a Mendelian randomization study.  
*Front. Oncol.* 14:1365138.  
doi: 10.3389/fonc.2024.1365138

## COPYRIGHT

© 2024 Huang, Deng, Zhang, Wang, Qu, Li and  
Liu. 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.

# The causal relationship between cathepsins and digestive system tumors: a Mendelian randomization study

Xupeng Huang<sup>1</sup>, Houbo Deng<sup>2</sup>, Bo Zhang<sup>1</sup>, Kuisong Wang<sup>1</sup>,  
Yi Qu<sup>1</sup>, Ting Li<sup>2</sup> and Tiejun Liu<sup>2\*</sup>

<sup>1</sup>Graduate School, Changchun University of Traditional Chinese Medicine, Changchun, China,

<sup>2</sup>Department of Hepatology, First Affiliated Hospital to Changchun University of Chinese Medicine, Changchun, China

**Background:** Multiple studies have confirmed the significant role of cathepsins in the development and progression of digestive system tumors. However, further investigation is needed to determine the causal relationships.

**Methods:** We conducted a two-sample bidirectional Mendelian randomization (MR) study using pooled data from a genome-wide association study (GWAS) to assess the causal associations between nine cathepsins (cathepsin B, E, F, G, H, L2, O, S, and Z) and six types of digestive system tumors, including hepatocellular carcinoma (HCC), pancreatic cancer (PCa), biliary tract cancer (BTC), colorectal cancer (CRC), gastric carcinoma (GC), and esophageal cancer (EC). We employed the following methods including inverse variance weighting (IVW), MR-Egger, weighted median (WM), Cochran's Q, MR-PRESSO, MR-Egger intercept test and leave-one-out sensitivity analysis. The STROBE-MR checklist for the reporting of MR studies was used in this study.

**Results:** The risk of HCC increased with high levels of cathepsin G (IVW:  $p = 0.029$ , odds ratio (OR) = 1.369, 95% confidence interval (CI) = 1.033-1.814). Similarly, BTC was associated with elevated cathepsin B levels (IVW:  $p = 0.025$ , OR = 1.693, 95% CI = 1.070-2.681). Conversely, a reduction in PCa risk was associated with increased cathepsin H levels (IVW:  $p = 0.027$ , OR = 0.896, 95% CI = 0.812-0.988). Lastly, high levels of cathepsin L2 were found to lower the risk of CRC (IVW:  $p = 0.034$ , OR = 0.814, 95% CI = 0.674-0.985).

**Conclusion:** Our findings confirm the causal relationship between cathepsins and digestive system tumors, which can offer valuable insights for the diagnosis and treatment of digestive system tumors.

## KEYWORDS

Mendelian randomization, cathepsins, digestive system tumors, genome-wide association, causal relationship

## 1 Introduction

Digestive system tumors are a significant contributor to cancer-related deaths globally. Their incidence and mortality rates have been increasing consistently, presenting a significant challenge to human health (1). The primary types of digestive system tumors include hepatocellular carcinoma (HCC), pancreatic cancer (PCa), biliary tract cancer (BTC), colorectal cancer (CRC), gastric carcinoma (GC), and esophageal cancer (EC) (2). The occurrence and development of digestive system tumors are complex processes that involve multiple risk factors. One important factor in the formation of these tumors is the ability of carcinoma cells to sustain internal homeostasis (3). Proteins have a critical function in maintaining the metabolic equilibrium of cancer cells, with the activity of the proteolytic system being especially important for the proliferation of cancer cells (4). Cathepsins are lysosomal proteolytic enzymes that are responsible for maintaining cellular homeostasis. They primarily function as endopeptidases within the lysosomal vesicles of normal cells (5). Cathepsins are involved in various physiological processes such as protein turnover, differentiation, and apoptosis. They also play important roles in signaling cellular stress, breaking down the extracellular matrix, causing lysosome-mediated cell death, and have been associated with the progression of a diversity of diseases, including malignancies (6).

According to previous studies, cathepsins may contribute to the development and progression of diverse types of cancers, including breast cancer (7, 8), bladder cancer (9), thyroid cancer (10), and others. Notably, multiple cancer types have been linked to cathepsin B (11). A recent Mendelian randomization (MR) study revealed that elevated levels of cathepsin H may contribute to an increased risk of lung cancer (12). Previous observational studies have indicated a correlation between cathepsins and tumors in the digestive system, specifically cathepsin L (13) and cathepsin W (14), which have been linked to a poor prognosis for PCa. Cathepsin B has been linked to HCC (15), PCa (16, 17), BTC (18, 19), and CRC (20, 21). Furthermore, Cathepsin L (22) plays a significant role in angiogenesis of GC. Due to the lack of clear causal relationship between cathepsins and digestive system tumors in previous observational studies, we employed MR analysis. MR is a statistical method commonly used in genetic epidemiology studies, which combines data from genome-wide association studies (GWAS). One of the key advantages of MR is its ability to handle confounding factors and eradicate reverse causality by using genetic variants as instrumental variables (IVs) (23).

The objective of this research was to examine the correlation between cathepsins and digestive system tumors by analyzing genetic variants associated with nine cathepsins and six digestive system tumors obtained from a large GWAS. The analysis was conducted using a two-sample bidirectional MR approach. Our research findings significantly augment the comprehension of the causal relationship between cathepsins and these digestive system tumors.

## 2 Materials and methods

### 2.1 Data sources

The MR analysis of nine cathepsin levels in this investigation obtained the genetic tools from the INTERVAL study, which comprised 3301 Europeans (24). Every contributor was obligated to fill out a consent form, and the INTERVAL study received approval from The National Research Ethics Service (11/EE/0538). The relevant data can be accessed openly at <https://gwas.mrcieu.ac.uk>. Statistics on digestive system tumors were collected from various GWAS databases, the International Classification of Diseases Tenth Revision codes (ICD10), definitions, diagnostic criteria or methods involving six digestive disorders are listed in [Additional File 2](#). The genetic variation data for HCC and PCa were publicly accessible at <https://www.ebi.ac.uk/gwas>. HCC comprised of 475,638 samples (379 cases and 475,259 controls) and 24,194,938 single-nucleotide polymorphisms (SNPs). PCa comprised of 476,245 samples (1,196 cases and 475,049 controls) and 24,195,229 SNPs. The genetic variation data for BTC, CRC, and GC can be accessed openly at [https://www.finngen.fi/en/access\\_results](https://www.finngen.fi/en/access_results), BTC consists of 218,792 samples (109 cases and 218,683 controls) with 16,380,466 SNPs. CRC consists of 218,792 samples (3,022 cases and 215,770 controls) with the same number of SNPs. GC consists of 218,792 samples (633 cases and 218,159 controls) with the same number of SNPs. The genetic variation data for EC can be obtained openly from <https://gwas.mrcieu.ac.uk>, it includes 372,756 samples (740 cases and 372,016 controls) with 8,970,465 SNPs. The data presented in this study were derived exclusively from European population samples. These samples were obtained from independent GWAS databases, ensuring minimal overlap and bias. It is important to note that every participant granted informed written consent, and all studies underwent review and approval by the institutional ethical review boards of the relevant institutions. Therefore, no additional ethical approvals or licenses were necessary for this MR study.

### 2.2 Study design and selection of IVs

The study methods were compliant with the STROBE-MR checklist (25), further details can be found in [Additional File 1](#). These IVs needed to satisfy three core assumptions (26): the hypothesis of correlation, the hypothesis of exclusivity, and the assumption of independence, as illustrated in [Figure 1](#). The first assumption establishes a robust link between SNPs and the variable of exposure. To identify SNPs associated with exposure factors and establish the validity and accuracy of the causal relationship between cathepsins and digestive system tumors, the following steps were followed to select the most suitable SNPs. Firstly, due to the restricted pool of SNPs accessible for MR analysis, a significance threshold of P value was less than  $5 \times 10^{-6}$  was

established for the detection of SNPs that exhibit strong associations with the investigated exposures. Moreover, to eliminate any presence of linkage disequilibrium (27), an  $r^2$  threshold of 0.001 and a clump window size of 10,000 kb were implemented (27). Secondly, the selected SNPs were ensured to have no association with any confounding factors that could influence the relationship between exposure and outcome. Lastly, the SNPs were confirmed to only impact the outcome through exposure factors.

In addition, the selected IVs were assessed for the weak IV bias by calculating the F-statistic (28). The F-statistic for each SNP was calculated using the formula  $F = R^2(N - K - 1) / [K(1 - R^2)]$ , where  $R^2$ ,  $N$ , and  $K$  represent the estimated exposure variance explained by the IVs, sample size, and the number of IVs (29). If the F-statistic was less than 10, the SNP was considered a weak IV and excluded from the analysis to mitigate bias caused by weak IVs.

### 3 MR analysis

In this study, two samples were analyzed bidirectionally by MR. Initially, the researchers examined the causal relationship between nine cathepsins and six digestive system tumors, with cathepsins considered as the exposure and digestive system tumors as the outcome. Subsequently, a reverse analysis was conducted using the same settings and data sets as the forward MR analysis. All analyses were conducted using R version 4.2.2, with the software packages ‘Two-SampleMR’ and ‘MR-PRESSO’ (30). Three analysis methods were employed in this study: Inverse variance weighting (IVW), MR-Egger, and weighted median (WM). The IVW method, considered the primary method for assessing causality (31), yielded a nominally significantly correlated result when the P

value was less than 0.05. The IVW method is a classical method for MR analysis, where the weighted average is calculated by taking the reciprocal of the variance of each IV as the weight, ensuring the effectiveness of all IVs. To ensure the robustness of the MR results, both MR-Egger and WM methods were employed as complementary approaches. MR-Egger utilizes a weighted linear regression analysis, providing robust estimates that are independent of the validity of instrumental variables. Nevertheless, it is crucial to acknowledge that these estimates may have lower statistical precision and can be influenced by outlier genetic variation (32). On the other hand, The problem of estimation accuracy variability is tackled by the WM approach. In a manner reminiscent of the IVW approach, the WM method assigns inverse weights that are contingent upon the variance of individual genetic variants, demonstrating reliability even when causal effects are violated (33). The Cochran’s Q test was used to estimate the heterogeneity of SNPs. A p-value greater than 0.05 indicated the absence of heterogeneity. In the case of significant heterogeneity in SNPs, a random effects model was applied. Conversely, a fixed effects model was utilized instead. Additionally, to ensure the reliability of the results, a leave-one-out analysis was carried out. This analysis aimed to remove SNPs that could have potentially extreme effects (30). To identify horizontal pleiotropy, the MR-egger intercept was utilized. A p-value greater than 0.05 indicated no horizontal pleiotropy. If the p-value was less than 0.05 in situations, the outlier test was employed to eliminate horizontal pleiotropy by implementing the MR-PRESSO global test (34). Causality was evaluated using the odds ratio (OR) and 95% confidence interval (CI). If the OR value was less than 1, the exposure will be regarded as a protective factor for the outcome. Conversely, if the OR value was greater than 1, the exposure will be classified as a risk factor for the outcome.

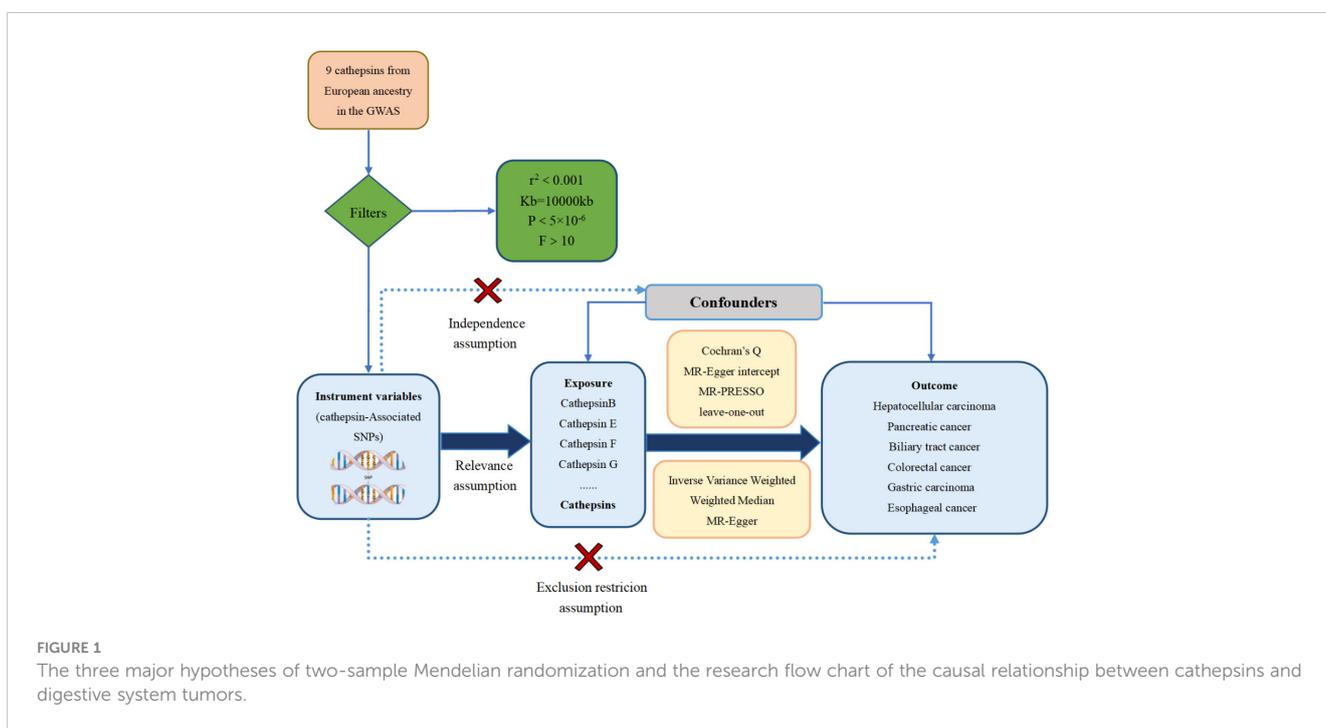


FIGURE 1 The three major hypotheses of two-sample Mendelian randomization and the research flow chart of the causal relationship between cathepsins and digestive system tumors.

To visualize the MR analysis, forest plots, scatter plots, and leave-one-out plots were generated using the data analysis function of the Rstudio platform.

## 4 Results

### 4.1 IV selection

The genetic variants utilized in our study to analyze 9 cathepsins were sourced from the INTERVAL study, while statistical data for 6 digestive system tumors were gathered from various GWAS databases. By employing genome-wide significance threshold screening ( $P < 5 \times 10^{-6}$ ), multiple SNPs were identified as IVs for each cathepsin and disease in the bidirectional MR analysis. Each selected IV had an F-statistic exceeding 10, indicating the absence of weak IV bias. Detailed information on the SNPs can be found in [Supplementary Table \(1, 3\)](#).

### 4.2 MR main analysis results

To evaluate the impact of various cathepsins on the risk of different types of digestive system tumors, a Two-SampleMR analysis was conducted. This analysis encompassed a total of nine cathepsins include cathepsin B, cathepsin E, cathepsin F, cathepsin G, cathepsin H, cathepsin L2, cathepsin O, cathepsin S, and cathepsin Z, and six digestive system tumors (HCC, PCa, BTC, CRC, GC, and EC). This study utilized the IVW method as the primary analysis technique. IVW is a method used for meta-analyzing the effects of multiple genetic variants, which can generate reliable causal estimates without the presence of directional pleiotropy. The findings from the univariate MR analysis indicated a causal relationship between four cathepsins and four digestive system tumors ([Table 1](#)). Specifically, The risk of HCC increased with high levels of cathepsin G (IVW:  $p = 0.029$ , odds ratio (OR) = 1.369, 95% CI = 1.033-1.814), the forest plot is shown in [Figure 2](#). Similarly, BTC was associated with elevated cathepsin B levels (IVW:  $p = 0.025$ , OR = 1.693, 95% CI = 1.070-2.681), the forest plot is shown in [Figure 3](#). Conversely, a reduction in PCa risk was associated with increased cathepsin H levels (IVW:  $p = 0.027$ , OR = 0.896, 95% CI = 0.812-0.988), the forest plot is shown in [Figure 4](#). Lastly, high levels of cathepsin L2 were found to lower the risk of CRC (IVW:  $p = 0.034$ , OR = 0.814, 95% CI = 0.674-0.985), the forest plot is shown in [Figure 5](#). The IVW results

indicated that cathepsin G was associated with a 36.9% increase in the risk of HCC (OR = 1.369), cathepsin B was linked to a 69.3% increase in the risk of BTC (OR = 1.693), while cathepsin H was found to reduce the risk of PCa by 10.4% (OR = 0.896), and cathepsin L2 was associated with an 18.6% reduction in the risk of CRC (OR = 0.814). The presence of a causal association between digestive system tumors and other forms of cathepsins was not found by employing the IVW method ([Supplementary Table 1](#)).

### 4.3 MR sensitivity analyses results

The IVW method may overlook potential pleiotropic effects, therefore we conducted sensitivity analyses such as MR-Egger and weighted median to evaluate the reliability and consistency of the findings. Additionally, The Cochran's Q test was used to estimate the heterogeneity of SNPs. A p-value greater than 0.05 indicated the absence of heterogeneity. To identify horizontal pleiotropy, the MR-egger intercept was utilized. A p-value greater than 0.05 indicated no horizontal pleiotropy. In [Supplementary Table 2](#), both Cochran's Q and MR-Egger intercepts did not indicate any evidence of heterogeneity or pleiotropy in these causal relationships. The scatter plot is shown in [Figure 6](#). The results of the leave-one-out analysis demonstrate the robustness of these findings, as illustrated in [Figure 7](#).

### 4.4 Reverse MR analysis

To address potential reverse causation, we conducted a reverse Two-SampleMR analysis using 6 digestive system tumors as exposures and 9 cathepsins as outcomes. The results of the MR analysis did not show any evidence of reverse causality ([Supplementary Table 3](#)).

## 5 Discussion

The occurrence and progression of tumors in the digestive system entail a remarkably intricate procedure, wherein proteolytic occurrences assume indispensable functions. A mounting quantity of investigations has been concentrating on scrutinizing the function of cathepsins in tumors of the digestive system (35). In this study, we employed genetic instruments to systematically examine the causal relationship between nine cathepsins and six

TABLE 1 Mendelian randomization analysis with forward causality.

Cathepsin	outcome	Inverse variance weighted		MR-Egger		Weighted median	
		OR (95%CI)	p_value	OR (95%CI)	p_value	OR (95%CI)	p_value
Cathepsin B	BTC	1.693(1.070-2.681)	0.025	1.893(0.651-5.505)	0.260	1.555(0.816-2.962)	0.180
Cathepsin G	HCC	1.369(1.033-1.814)	0.029	1.309(0.639-2.681)	0.479	1.559(1.049-2.318)	0.028
Cathepsin H	PCa	0.896(0.812-0.988)	0.027	0.948(0.837-1.075)	0.439	0.918(0.832-1.013)	0.089
Cathepsin L2	CRC	0.814(0.674-0.985)	0.034	0.785(0.464-1.325)	0.399	0.810(0.630-1.042)	0.101

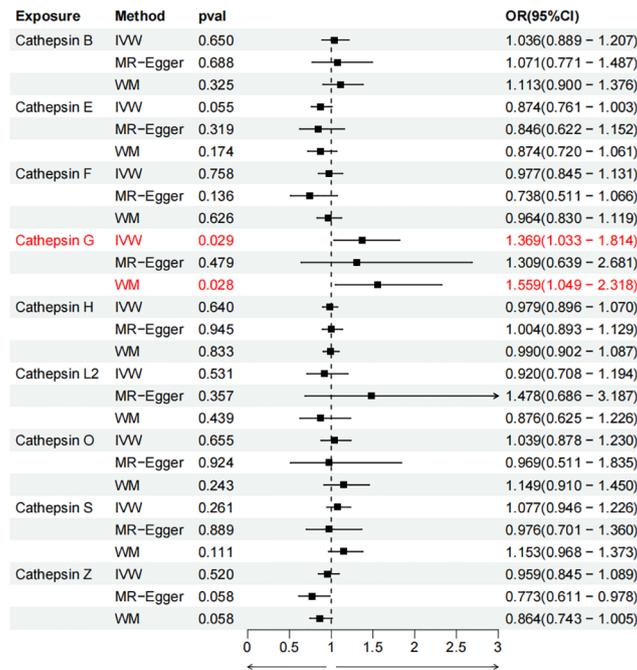


FIGURE 2 Forest plots of cathepsins on hepatocellular carcinoma. The forward causality were marked red in the figure. IVW, inverse variance weighting; WM, weighted median; OR, odds ratio; CI, confidence interval.

digestive system tumors. Through two-sample bidirectional MR analysis, we observed potential associations between four cathepsins and four digestive system tumors. Specifically, higher levels of cathepsin G were found to increase the risk of HCC, while

elevated levels of cathepsin B were linked to an increased risk of BTC. Conversely, elevated levels of cathepsin H were found to potentially decrease the risk of PCa, and Increased levels of cathepsin L2 were correlated with a potential decrease in the risk

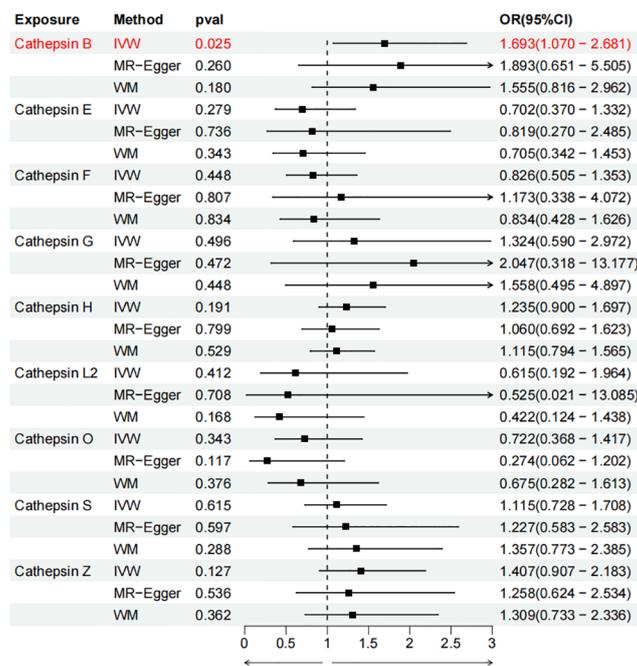


FIGURE 3 Forest plots of cathepsins on biliary tract cancer. The forward causality were marked red in the figure. IVW, inverse variance weighting; WM, weighted median; OR, odds ratio; CI, confidence interval.

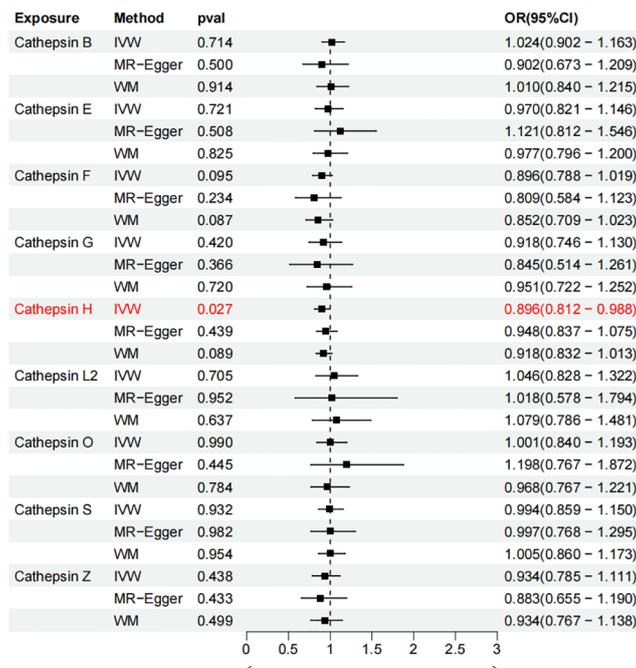


FIGURE 4 Forest plots of cathepsins on pancreatic cancer. The forward causality were marked red in the figure. IVW, inverse variance weighting; WM, weighted median; OR, odds ratio; CI, confidence interval.

of CRC. These findings suggest that different cathepsins may have varying effects on different types of digestive system tumors.

The conclusion of the study further clarified the previously observed correlation between cathepsin B and BTC (18, 19),

indicating a potential causal relationship. Cathepsin B, a cysteine protease, acts mainly as an endopeptidase in the normal cellular endolysosomal compartment. However, during tumor development, the regulation of cathepsin B may suffer disturbances at different

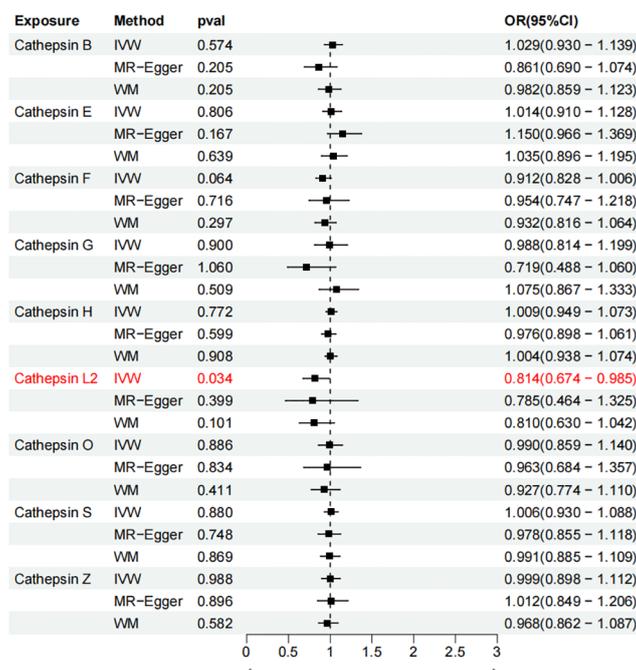
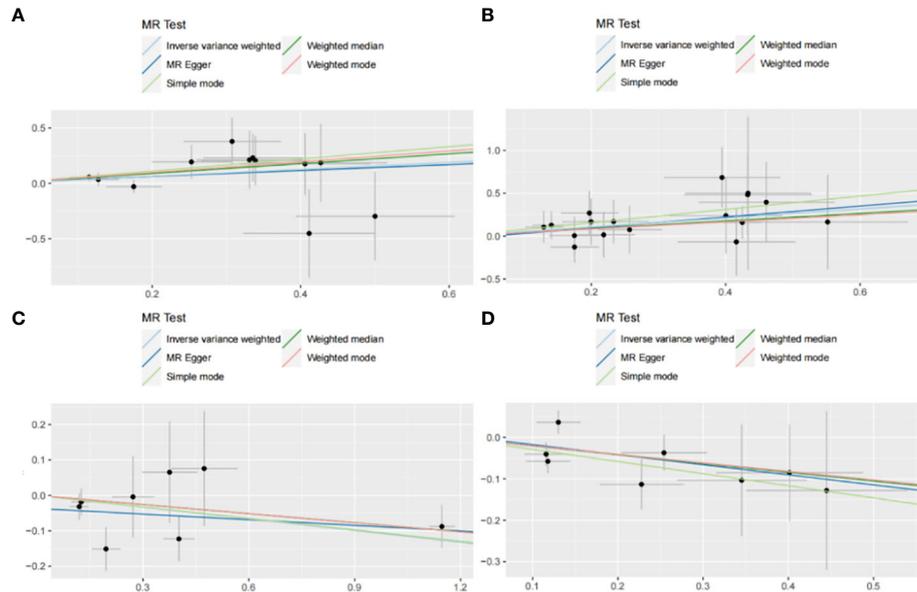


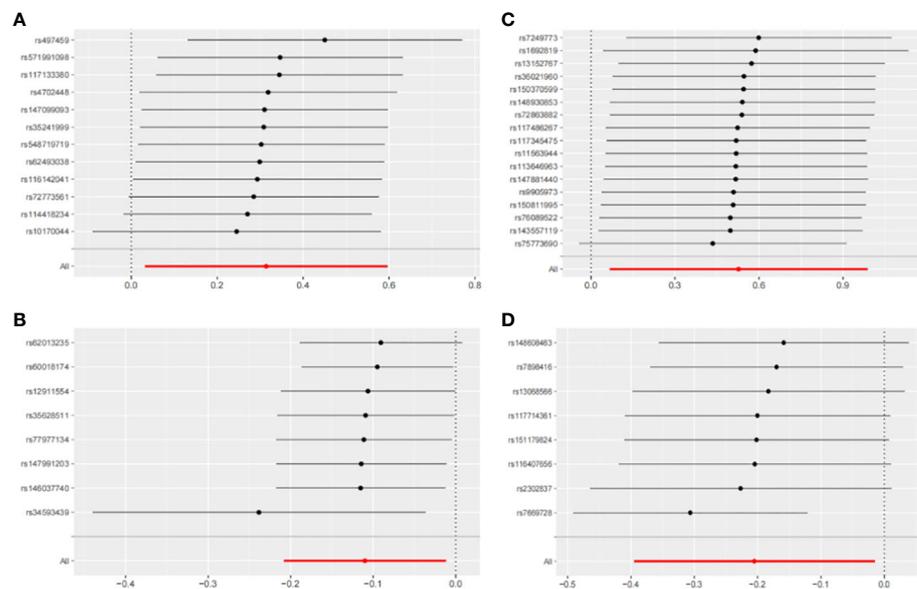
FIGURE 5 Forest plots of cathepsins on colorectal cancer. The forward causality were marked red in the figure. IVW, inverse variance weighting; WM, weighted median; OR, odds ratio; CI, confidence interval.



**FIGURE 6** Scatter plot with forward causality in Mendelian randomization. (A) Cathepsin G on hepatocellular carcinoma. (B) Cathepsin B on biliary tract cancer. (C) Cathepsin H on pancreatic cancer. (D) Cathepsin L2 on colorectal cancer.

stages, leading to its excessive expression and subsequent release into the extracellular space. Cathepsin B has an important function in controlling tumor growth, migration, invasion, angiogenesis, and metastasis, particularly in digestive system tumors (36). Moreover, previous studies have demonstrated the substantial impact of cathepsin B on the growth of HCC (15). However, the present MR analysis does not establish a causal relationship between the two variables, suggesting the involvement of intricate mechanisms that require further investigation.

Activated neutrophils secrete Cathepsin G, which is a serine protease and is closely associated with tumor diseases (37). The interaction of cathepsin G, found on neutrophils, with the receptor for advanced glycation end products (RAGE), found on the surface of tumors, plays a critical role in the cancer-fighting abilities of neutrophils (38). Furthermore, cathepsin G can enhance the adhesion and migration of cancer cells (39), as well as facilitate tumor formation by regulating tumor-matrix interactions (40). Increasing numbers of studies have shown that neutrophil



**FIGURE 7** Leave-one-out sensitivity analysis with forward causality in Mendelian randomization. (A) Cathepsin G on hepatocellular carcinoma. (B) Cathepsin B on biliary tract cancer. (C) Cathepsin H on pancreatic cancer. (D) Cathepsin L2 on colorectal cancer.

extracellular traps (NET) are associated with tumor progression. In the process, cathepsin G, in conjunction with matrix metalloproteinase 9 (MMP-9) and neutrophil elastase (NE), actively participates (41). Previous experimental studies have confirmed that the interplay of tumor cells and neutrophils promotes the metastasis of liver cancer through the cathepsin G component related to NET (42), indicating a correlation between cathepsin G and HCC. This study strengthens the evidence of a causal link between cathepsin G and HCC, implying the potential of cathepsin G as a target for therapeutic intervention in HCC. These findings provide valuable insights for the treatment of HCC.

The present study unveiled a previously unmentioned finding that cathepsin H reduces the risk of PCa. Cathepsin H belongs to the papain superfamily of lysosomal cysteine proteases and is the sole known aminopeptidase in this family (43). It is synthesized as an inactive precursor and becomes activated through protein hydrolysis to remove its prepeptide (44). Cathepsin H has been reported to be associated with cancer and other major diseases. However, there is an insufficiency of existing investigations concerning the correlation between cathepsin H and PCa. Additional research is imperative to acquire a comprehensive understanding of the intricate linkage between the two entities. The results of the present MR analysis serve as a valuable reference for future studies on cathepsin H and PCa. Previous studies have indicated that several cathepsins are linked to the occurrence and progression of PCa (45). Specifically, cathepsin E has been found to have high expression and activity in pancreatic ductal adenocarcinoma (PDAC) (46), while cathepsin B has been shown to play a significant role in the autonomous growth of PDAC, potentially promoting tumor cell proliferation through the regulation of cathepsin L (16). Nevertheless, the MR analysis conducted in this study failed to establish a conclusive causal link between cathepsin E, cathepsin B, and the development of PCa.

According to MR analysis, the presence of cathepsin L2 has been found to reduce the risk of CRC. Cathepsin L2, also known as cathepsin V, is a human lysosomal cysteine peptidase that plays specific roles in pathological mechanisms and is crucial in the breakdown of the extracellular matrix (47). Cathepsin L2 is involved in the release of antigenic peptides, the maturation of major histocompatibility complex (MHC) class II molecules, the turnover of elastinogen fibers, as well as the cleavage of intracellular and extracellular substrates (48). It has been discovered that cathepsin L2, which is widely expressed in human tumors, plays an essential part in promoting the proliferation of bladder cancer by increasing NF- $\kappa$ B activity (9). Furthermore, it has been found that cathepsin L2 can also drive the progression of lung cancer by influencing the immunosuppressive environment and the cleavage of adhesion molecules (47). Furthermore, the activity of cathepsin L2 has been demonstrated to govern the advancement of the cell cycle and maintain the stability of histones within the nucleus of malignant cells found in breast cancer (49). It is also closely associated with the prognosis of HCC and the tumor microenvironment (50). Interestingly, this study reveals that cathepsin L2 reduces the risk of CRC, which has not been reported in previous studies. This suggests a potentially more complex relationship between cathepsin L2 and CRC, and further

research is required to comprehend the underlying mechanism of action.

Previous research have highlighted the importance of cathepsin B in the development, invasion, and metastasis of CRC (21). Furthermore, the expression of cathepsin S in CRC has been reported to be substantial (51). However, the current MR analysis did not establish a causal relationship between cathepsin B, S, and CRC. Apart from BTC, HCC, PCa, and CRC as mentioned earlier, there are also relevant reports suggesting that cathepsin B serves as a sensitive indicator for gastrointestinal malignant tumors (52). This indicates the potential importance of cathepsin B in digestive system tumors, which warrants further investigation.

Tumor screening is gaining more attention in the prevention of tumor diseases. The screening of serum biomarkers is also becoming more convenient and efficient. The objective of this investigation is to examine the causal relationship between various cathepsins and different digestive system tumors through MR analysis. The study employs genetic variations as a means to minimize the impact of confounding factors and reverse causal connections. Using genetics as a foundation for analysis helps to establish a stronger correlation between variables and reduce the likelihood of unrelated factors affecting the results. This approach allows for a more reliable and accurate understanding of the data, as it circumvents the potential biases that may be introduced by external factors. By focusing on genetic variations, the study aims to provide a clearer picture of the relationships under investigation and offers a valuable contribution to the field. The findings of this study can help in the search for effective tumor markers and provide potential value for further research on digestive system tumors. Nevertheless, it is critical to recognize the constraints of our investigation. Firstly, the databases used in the study only included individuals of European ancestry. To obtain stronger evidence, it is necessary to expand the databases to include other ethnic groups such as those from Asia and Africa. Secondly, the threshold of P value was less than  $5 \times 10^{-8}$  is generally considered to indicate genome-wide significance when screening for instrumental variables. In this study, we aimed to increase the number of SNPs to mitigate bias from a limited number of IVs, therefore, we consulted relevant literature on Mendelian randomization of cathepsin and established a significance threshold of P value less than  $5 \times 10^{-6}$ . Nevertheless, caution is advised when interpreting the results. Third, the MR analysis method is a theoretical causal analysis method that requires further validation through animal experiments to establish the causal relationship. This will help in understanding the intricate mechanism linking cathepsins and digestive system tumors.

## 6 Conclusion

In conclusion, the results demonstrate a potential causal relationship between specific cathepsins and digestive system tumors. These findings may offer potential targets and new biomarkers for the diagnosis and treatment of digestive system tumors. However, further interventional trials are needed to clarify the underlying mechanisms.

## 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

XH: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. HD: Methodology, Writing – review & editing. BZ: Data curation, Writing – review & editing. KW: Data curation, Writing – review & editing. YQ: Visualization, Writing – review & editing. TL: Project administration, Writing – review & editing. TJL: Supervision, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by National Key R & D Program of China (2018YFC1704100, 2018YFC1704105) and the Natural Science Foundation of Jilin Province (20200201590JC).

## References

1. Ashktorab H, Kupfer SS, Brim H, Carethers JM. Racial disparity in gastrointestinal cancer risk. *Gastroenterology*. (2017) 153:910–23. doi: 10.1053/j.gastro.2017.08.018
2. Chirnomas D, Hornberger KR, Crews CM. Protein degraders enter the clinic - a new approach to cancer therapy. *Nat Rev Clin Oncol*. (2023) 20:265–78. doi: 10.1038/s41571-023-00736-3
3. Khaket TP, Kwon TK, Kang SC. Cathepsins: Potent regulators in carcinogenesis. *Pharmacol Ther*. (2019) 198:1–19. doi: 10.1016/j.pharmthera.2019.02.003
4. Rose TC, Pennington A, Kyridemos C, Chen T, Subhani M, Hanefeld J, et al. Analysis of the burden and economic impact of digestive diseases and investigation of research gaps and priorities in the field of digestive health in the European Region-White Book 2: Executive summary. *United Eur Gastroenterol J*. (2022) 10:657–62. doi: 10.1002/ueg2.12298
5. Ruiz-Blázquez P, Pistorio V, Fernández-Fernández M, Moles A. The multifaceted role of cathepsins in liver disease. *J Hepatol*. (2021) 75:1192–202. doi: 10.1016/j.jhep.2021.06.031
6. Vermeulen L, Snippert HJ. Stem cell dynamics in homeostasis and cancer of the intestine. *Nat Rev Cancer*. (2014) 14:468–80. doi: 10.1038/nrc3744
7. Kang J, Yu Y, Jeong S, Lee H, Heo HJ, Park JJ, et al. Prognostic role of high cathepsin D expression in breast cancer: a systematic review and meta-analysis. *Ther Adv Med Oncol*. (2020) 12:1758835920927838. doi: 10.1177/1758835920927838
8. Xiao Y, Cong M, Li J, He D, Wu Q, Tian P, et al. Cathepsin C promotes breast cancer lung metastasis by modulating neutrophil infiltration and neutrophil extracellular trap formation. *Cancer Cell*. (2021) 39:423–37.e7. doi: 10.1016/j.ccell.2020.12.012
9. Xia Y, Ge M, Xia L, Shan G, Qian H. CTSV (cathepsin V) promotes bladder cancer progression by increasing NF- $\kappa$ B activity. *Bioengineered*. (2022) 13:10180–90. doi: 10.1080/21655979.2022.2061278
10. Wang Y, Mei J, Zhang Y, He X, Zheng X, Tan J, et al. Cathepsin F genetic mutation is associated with familial papillary thyroid cancer. *Am J Med Sci*. (2022) 364:414–24. doi: 10.1016/j.amjms.2022.03.017

## Acknowledgments

The authors thank the studies or consortiums cited and included in this analysis for providing public datasets.

## 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/fonc.2024.1365138/full#supplementary-material>

11. Mijanović O, Branković A, Panin AN, Savchuk S, Timashev P, Ulasov I, et al. Cathepsin B: A sellsword of cancer progression. *Cancer letters*. (2019) 449:207–14. doi: 10.1016/j.canlet.2019.02.035
12. Li J, Tang M, Gao X, Tian S, Liu W. Mendelian randomization analyses explore the relationship between cathepsins and lung cancer. *Commun Biol*. (2023) 6:1019. doi: 10.1038/s42003-023-05408-7
13. Singh N, Das P, Datta Gupta S, Sahni P, Pandey RM, Gupta S, et al. Prognostic significance of extracellular matrix degrading enzymes-cathepsin L and matrix metalloproteinases-2 [MMP-2] in human pancreatic cancer. *Cancer Invest*. (2013) 31:461–71. doi: 10.3109/07357907.2013.820318
14. Khojasteh-Leylakooi F, Mohit R, Khalili-Tanha N, Asadnia A, Naderi H, Pourali G, et al. Down regulation of Cathepsin W is associated with poor prognosis in pancreatic cancer. *Sci Rep*. (2023) 13:16678. doi: 10.1038/s41598-023-42928-y
15. Nguyen T, Kumar RP, Park PH. Cathepsin B maturation plays a critical role in leptin-induced hepatic cancer cell growth through activation of NLRP3 inflammasomes. *Arch Pharmacol Res*. (2023) 46:160–76. doi: 10.1007/s12272-023-01437-2
16. Gopinathan A, Denicola GM, Frese KK, Cook N, Karreth FA, Mayerle J, et al. Cathepsin B promotes the progression of pancreatic ductal adenocarcinoma in mice. *Gut*. (2012) 61:877–84. doi: 10.1136/gutjnl-2011-300850
17. Rinderknecht H, Renner IG. Increased cathepsin B activity in pancreatic juice from a patient with pancreatic cancer. *New Engl J Med*. (1980) 303:462–3. doi: 10.1056/nejm198008213030819
18. Monsouvanh A, Proungvitaya T, Limpaboon T, Wongkham C, Wongkham S, Luvira V, et al. Serum cathepsin B to cystatin C ratio as a potential marker for the diagnosis of cholangiocarcinoma. *Asian Pacific J Cancer Prev APJCP*. (2014) 15:9511–5. doi: 10.7314/apjcp.2014.15.21.9511
19. Terada T, Ohta T, Minato H, Nakanuma Y. Expression of pancreatic trypsinogen/trypsin and cathepsin B in human cholangiocarcinomas and hepatocellular carcinomas. *Hum pathol*. (1995) 26:746–52. doi: 10.1016/0046-8177(95)90222-8
20. Herszényi L, Farinati F, Cardin R, István G, Molnár LD, Hritz I, et al. Tumor marker utility and prognostic relevance of cathepsin B, cathepsin L, urokinase-type

- plasminogen activator, plasminogen activator inhibitor type-1, CEA and CA 19-9 in colorectal cancer. *BMC cancer*. (2008) 8:194. doi: 10.1186/1471-2407-8-194
21. Bian B, Mongrain S, Cagnol S, Langlois MJ, Boulanger J, Bernatchez G, et al. Cathepsin B promotes colorectal tumorigenesis, cell invasion, and metastasis. *Mol carcinogene*. (2016) 55:671–87. doi: 10.1002/mc.22312
22. Pan T, Jin Z, Yu Z, Wu X, Chang X, Fan Z, et al. Cathepsin L promotes angiogenesis by regulating the CDP/Cux/VEGF-D pathway in human gastric cancer. *Gastric Cancer*. (2020) 23:974–87. doi: 10.1007/s10120-020-01080-6
23. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations: update for summer 2023. *Wellcome Open Res*. (2019) 4:186. doi: 10.12688/wellcomeopenres.15555.3
24. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. *Nature*. (2018) 558:73–9. doi: 10.1038/s41586-018-0175-2
25. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomization: the STROBE-MR statement. *Jama*. (2021) 326:1614–21. doi: 10.1001/jama.2021.18236
26. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int J Epidemiol*. (2016) 45:1717–26. doi: 10.1093/ije/dyx028
27. Arnold M, Raffler J, Pfeufer A, Suhre K, Kastenmüller G. SNIIPA: an interactive, genetic variant-centered annotation browser. *Bioinf (Oxford England)*. (2015) 31:1334–6. doi: 10.1093/bioinformatics/btu779
28. Burgess S, Thompson SG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol*. (2011) 40:755–64. doi: 10.1093/ije/dyr036
29. Xiang K, Zhang JJ, Xu YY, Zhong X, Ni J, Pan HF. Genetically predicted causality of 28 gut microbiome families and type 2 diabetes mellitus risk. *Front endocrinol*. (2022) 13:780133. doi: 10.3389/fendo.2022.780133
30. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet*. (2017) 13:e1007081. doi: 10.1371/journal.pgen.1007081
31. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. (2013) 37:658–65. doi: 10.1002/gepi.21758
32. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. (2015) 44:512–25. doi: 10.1093/ije/dyv080
33. 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) 40:304–14. doi: 10.1002/gepi.21965
34. 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) 50:693–8. doi: 10.1038/s41588-018-0099-7
35. Herszényi L, Plebani M, Carraro P, De Paoli M, Roveroni G, Cardin R, et al. Proteases in gastrointestinal neoplastic diseases. *Clinica chimica Acta*. (2000) 291:171–87. doi: 10.1016/s0009-8981(99)00227-2
36. Zamyatnin AA Jr, Gregory LC, Townsend PA, Soond SM. Beyond basic research: the contribution of cathepsin B to cancer development, diagnosis and therapy. *Expert Opin Ther targets*. (2022) 26:963–77. doi: 10.1080/14728222.2022.2161888
37. Morimoto-Kamata R, Yui S. Insulin-like growth factor-1 signaling is responsible for cathepsin G-induced aggregation of breast cancer MCF-7 cells. *Cancer sci*. (2017) 108:1574–83. doi: 10.1111/cas.13286
38. Sionov RV. Leveling up the controversial role of neutrophils in cancer: when the complexity becomes entangled. *Cells*. (2021) 10:2486. doi: 10.3390/cells10092486
39. Woo SM, Min KJ, Seo SU, Kim S, Park JW, Song DK, et al. Up-regulation of 5-lipoxygenase by inhibition of cathepsin G enhances TRAIL-induced apoptosis through down-regulation of survivin. *Oncotarget*. (2017) 8:106672–84. doi: 10.18632/oncotarget.22508
40. Wilson TJ, Nannuru KC, Singh RK. Cathepsin G recruits osteoclast precursors via proteolytic activation of protease-activated receptor-1. *Cancer Res*. (2009) 69:3188–95. doi: 10.1158/0008-5472.Can-08-1956
41. Liu Y, Liu L. The pro-tumor effect and the anti-tumor effect of neutrophils extracellular traps. *Biosci trends*. (2020) 13:469–75. doi: 10.5582/bst.2019.01326
42. Guan X, Lu Y, Zhu H, Yu S, Zhao W, Chi X, et al. The crosstalk between cancer cells and neutrophils enhances hepatocellular carcinoma metastasis via neutrophil extracellular traps-associated cathepsin G component: A potential therapeutic target. *J hepatocell carcinoma*. (2021) 8:451–65. doi: 10.2147/jhc.S303588
43. Hao Y, Purtha WE, Cortesio C, Rui H, Gu Y, Chen H, et al. Correction: Crystal structures of human procathepsin H. *PLoS One*. (2019) 14:e0218063. doi: 10.1371/journal.pone.0218063
44. Rojnik M, Jevnikar ZR, Doljak B, Turk S, Zidar N, Kos J. The influence of differential processing of procathepsin H on its aminopeptidase activity, secretion and subcellular localization in human cell lines. *Eur J Cell Biol*. (2012) 91:757–64. doi: 10.1016/j.ejcb.2012.05.004
45. Sulpizio S, Franceschini N, Piattelli A, Di Sebastiano P, Innocenti P, Selvaggi F. Cathepsins and pancreatic cancer: the 2012 update. *Pancreatol*. (2012) 12:395–401. doi: 10.1016/j.pan.2012.07.011
46. Pontious C, Kaul S, Hong M, Hart PA, Krishna SG, Lara LF, et al. Cathepsin E expression and activity: Role in the detection and treatment of pancreatic cancer. *Pancreatol*. (2019) 19:951–6. doi: 10.1016/j.pan.2019.09.009
47. Zhu L, Zeng Q, Wang J, Deng F, Jin S. Cathepsin V drives lung cancer progression by shaping the immunosuppressive environment and adhesion molecules cleavage. *Aging*. (2023) 15:13961–79. doi: 10.18632/aging.205278
48. Lecaillon F, Chazeirat T, Saidi A, Lalmanach G. Cathepsin V: Molecular characteristics and significance in health and disease. *Mol aspects Med*. (2022) 88:101086. doi: 10.1016/j.mam.2022.101086
49. Sereesongsang N, Burrows JF, Scott CJ, Brix K, Burden RE. Cathepsin V regulates cell cycle progression and histone stability in the nucleus of breast cancer cells. *Front Pharmacol*. (2023) 14:1271435. doi: 10.3389/fphar.2023.1271435
50. Liu J, Zhang W, Wang Z, Wang Y, Li T, Wang Y, et al. Cathepsin V is correlated with the prognosis and tumor microenvironment in liver cancer. *Mol carcinogene*. (2023) 63:400–16. doi: 10.1002/mc.23660
51. Ghanadi K, Ashorzadeh S, Aliyepoor A, Anbari K. Evaluation of serum levels of cathepsin S among colorectal cancer patients. *Ann Med Surg*. (2012) 78:103831. doi: 10.1016/j.amsu.2022.103831
52. Herszényi L, István G, Cardin R, De Paoli M, Plebani M, Tulassay Z, et al. Serum cathepsin B and plasma urokinase-type plasminogen activator levels in gastrointestinal tract cancers. *Eur J Cancer Prev*. (2008) 17:438–45. doi: 10.1097/CEJ.0b013e328305a130