Discover Oncology

Analysis

The relationship between cathepsins and nasopharyngeal carcinoma: a Mendelian randomization analysis

Jiaxiang Hu[1] · Keyun Lin[2] · Mei Jiang[3] · Guanying Qiao[1]

Received: 4 December 2024 / Accepted: 11 April 2025
Published online: 08 May 2025
© The Author(s) 2025 OPEN

Abstract

Background Observational studies have shown the potential for cathepsins (CTS) to have an effect on nasopharyngeal carcinoma (NPC), but their conclusions are susceptible to confounding factors.

Methods To investigate the causal relationship between CTS and NPC, Mendelian randomization (MR) was conducted. Genetic data for nine CTS (CTS B, E, F, G, H, L2, O, S and Z) was obtained from a genome-wide association study. As the data on outcome, genetic data of NPC was utilized from a FinnGen study. MR was performed using five analytical methods including Inverse Variance-Weighted (IVW) method, MR-Egger test, Weighted Median test, Simple Mode test and Weighted Mode test, with the IVW as the main analysis method. Cochran's *Q* test, MR-PRESSO global test and "leave-one-out" sensitivity test were used in sensitivity analysis. Reverse MR was performed to investigate whether there is reverse causality between NPC and CTS. MR Steiger test was used to determine the direction of the interaction between CTS and NPC.

Results Overall, the authors found favorable evidence to support the association between Cathepsin F (CTSF) and NPC. CTSF was associated to increase the risk of NPC (odds ratio [OR] = 1.845, 95% confidence intervals [CI] = 1.086 ~ 3.136, $P = 0.024$) according to IVW. The results proved to be stable and robust in the sensitivity analysis. In the Steiger test, the causal effect of CTS on NPC was shown to be unidirectional.

Conclusions These findings suggest that CTSF may plays an important role in NPC thus providing new research ideas for future basic research endeavors and clinical applications.

Keywords Nasopharyngeal carcinoma · Cathepsins · Mendelian randomization · Genetic variant

1 Introduction

Nasopharyngeal carcinoma (NPC) is one of the most common head and neck cancers that often occurs in South China and Southeast Asia. In 2022, 120,434 new cases ranked 23rd globally and 73,482 deaths ranked 21 st globally. NPC often causes masses in the nasopharynx and neck, even hindering the function of the auditory tube, leading to a serious reduction in quality of life [1, 2].To date, the development of modern medicine has made some progress in the treatment

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12672-025-02378-6.

✉ Mei Jiang, jiangmei@gzzyydxdyfsyy.wecom.work; ✉ Guanying Qiao, 397897924@qq.com | [1]Department of Oncology I, ShunDe Hospital Guangzhou University of Chinese Medicine, Foshan 528300, Guangdong Province, China. [2]Department of Oncology, First Clinical Medical College, Guangzhou University of Chinese Medicine, Guangzhou 510405, Guangdong Province, China. [3]Department of Oncology, First Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou 510405, Guangdong Province, China.

of NPC, but many patients still die due to distant metastasis or local recurrence [3]. Genetics, Epstein-Barr virus, smoking and pickled foods are risk factors for NPC. Compared with other tumors, NPC is more prone to copy number variation, a type of genetic mutation. Among them, genes that frequently undergo somatic mutations include TP53, KMT2 C, NOTCH2, BRCA1, BRCA2. Functional genes, such as TP53, RASSF1 and MAD1L1 significantly affect the growth process of NPC by influencing the NF-kB pathway [3]. Although genome-wide association studies have shown that genetic factors play an important role in the development of NPC [4], specific genomic abnormalities are largely unknown. The identification of genetic loci constitutes a pivotal catalyst for propelling advancements across diverse domains of biomedical research, encompassing both the formulation of innovative therapeutic strategies and the deciphering of molecular pathways underpinning drug resistance. Consequently, elucidating the mechanistic contributions of genetic determinants emerges as an indispensable research priority [5]. In addition, it is necessary to learn more about the process of the proteolytic system in the complex progress of the growth and survival of cancer cells [6]. As members closely related to its action, CTS have attracted much attention [7].

Cathepsins (CTS) are lysosomal proteolytic enzymes that are widely present in various animal tissues and cells [8]. Its specific substrates such as high efficiency and high selection promote its role in maintaining cell homeostasis by regulating proteolysis, foreign body clearance, immune response and apoptosis [9]. Among them, papain is also known as cysteine cathepsin which participates in a variety of common pathophysiological activities in the human body. It has been demonstrated that the process of proteolytic regulation is also implicated in the modulation of various hallmarks of cancer, including tumor cell proliferation, invasion, metastasis, and angiogenesis [10]. At the same time, the process of gene amplification and post-transcriptional modification of NPC promotes the high expression of CTS. Transformed cells and tumor cells secrete a variety of CTS, which in turn are associated with tumor cell surface binding partners and participate in tumor progression [11]. In head and neck cancers, including NPC, elevation of cathepsins constitutes a bypass loop of the RAS, thereby promoting cancer cell proliferation, survival, metastasis, angiogenesis and other tumor progression processes [12]. The high expression of cathepsin in the serum of patients with NPC makes it a potential serological marker of NPC. As a target, CTS has been used to study the immune system regulation of tumor-infiltrating lymphocytes and effector T cells to understand the mechanism of NPC suppression and evasion of immunity [13]. However, at present, there still a low number of studies on the causal relationship between CTS and NPC. Most of them use clinical observation and cell experiment research methods, leading their results are susceptible to confounding factors. These circumstances hinder our understanding of NPC from a genetic perspective and are detrimental to the prevention and treatment of NPC in clinical practice.

Previous genetic studies have identified potential associations between NPC susceptibility and multiple loci, including 6p21, TERT, MST1R, and HLA. However, the high polymorphism and linkage disequilibrium within these gene-dense regions have complicated the interpretation of genetic contributions to NPC risk, as small-scale studies are limited in resolving the intricate relationships among these variants [14].

Therefore, we conducted a Mendelian randomization (MR) analysis to explore the causal relationship between CTS and NPC, which is beneficial for learning more about the gene-related properties of NPC and provides new directions for clinical research on NPC. MR leverages the principle of random allocation of genetic variation during meiosis from parents to offspring to infer causal relationships between exposure and outcome while minimizing the influence of confounding factors and reverse causality. Single nucleotide polymorphisms (SNPs) are instrumental variables (IVs) used to overcome the influence of confounding factors on outcome data, leading to stable results compared with other traditional methods [15].To our knowledge, there has been no MR study investigating the potential causal relationship between CTS and NPC. Therefore, MR was conducted to confirm the association between these 9 CTS (CTS B, E, F, G, H, L2, O, S, and Z) and NPC in the hope of providing potential therapeutic targets for the prevention and treatment of NPC better [16].

2 Methods

2.1 Source of genetic data of CTS and NPC

The CTS GWAS data were grouped into nine clusters: CTSB, CTSE, CTSF, CTSG, CTSH, CTSL2, CTSS, CTSO, and CTSZ. Also, after referring to other studies that employed MR as a research method, we involved these 9 CTS in our study [8, 17]. Genetic data of these 9 CTS was utilized from a genome-wide association study (IEU Open GWAS) (Table 1). The study included 2994 plasma proteins from 3301 European populations. The 10.5 million estimated autosomal variants were examined genome-wide, resulting in 1927 relationships between genomic regions and 1478 proteins through short

strands of DNA bound to specific molecular targets (aptamers) [18]. In relevant studies, the genetic data of these 9 CTS were proven to be sufficiently justified for sample size [19]. In order to reduce the sample overlap of exposure data and outcome data, so as to reduce the interference of confounding factors on the results and improve the stability of the results, the latest Genetic data for NPC was from FinnGen (Table 1). Referred to other studies which include genetic data for NPC into MR analysis, the data were obtained from the European population including a total of 89 cases and 345,118 healthy controls [20]. All GWAS data referenced in this study were sourced from authoritative databases (e.g., IEU Open GWAS and FinnGen) and underwent rigorous quality control (QC) procedures, encompassing both sample-wise QC and variant-wise QC [19, 21]. These QC workflows minimize technical noise and population stratification, thereby enhancing the authenticity of association signals and reducing false-positive genetic loci. This ensures robust exposure-instrument variable associations, mitigating weak IVs bias in MR. A detailed overview of the GWAS data in this study are presented in supplementary materials (Supplementary file Table S1).

2.2 Study design

We conducted a two-sample MR study to explore the potential association between CTS and NPC. The SNPs related to CTS (exposure data) were used as IVs for analysis with NPC (outcome data). The three assumptions of MR are the conditions for the design of this study: (1) correlation assumption: IVs and exposure data are closely related; (2) independence assumption: the confounding factors of exposure data and outcome data have no significant effect on IVs; (3) exclusivity hypothesis: IVs are not directly related to the outcome data and have nothing to do with confounding factors, but only affect the outcome data through exposure data. The data for this study were obtained from public databases and had informed consents from the participants. Thus, no additional ethical application was required for this study. Schematic design of MR study of cathepsins and NPC is as seen in Fig. 1.

This IVs need to meet the following three hypotheses: (1) correlation assumption: IVs and exposure data are closely related; (2) independence assumption: the confounding factors of exposure data and outcome data have no significant effect on IVs; (3) exclusivity hypothesis: IVs are not directly related to the outcome data and have nothing to do with confounding factors, but only affect the outcome data through exposure data.

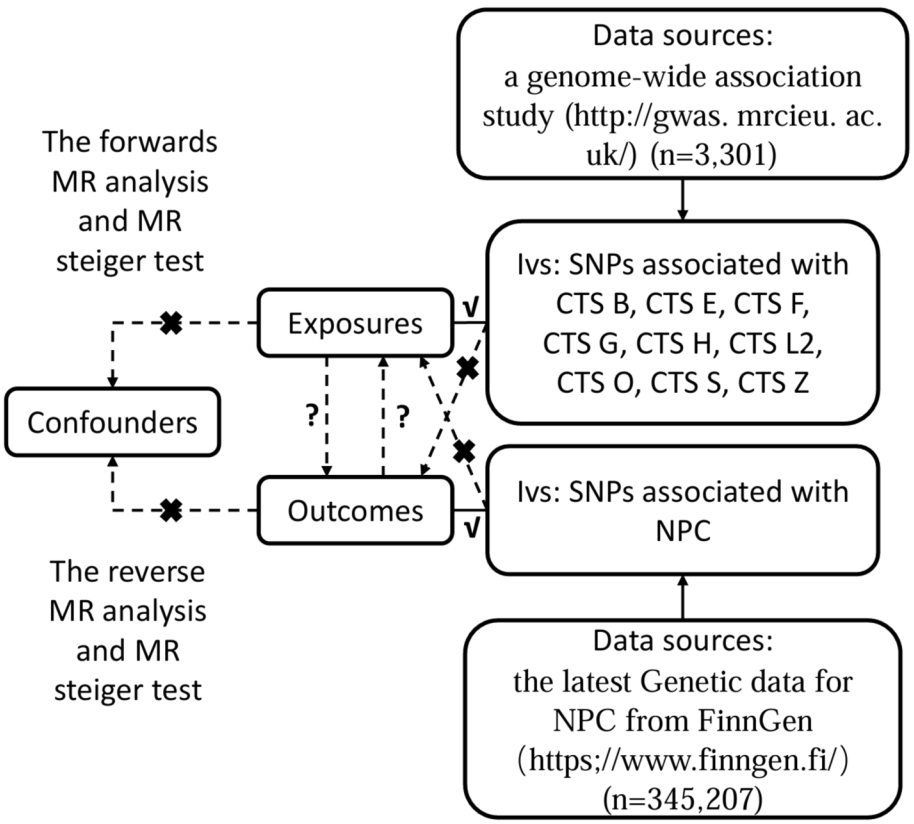

Fig. 1 Schematic design of MR study of CTS and NPC

Table 1 Information on the datasets for CTS and NPC

Trait	Population	Year	Sample size	Number of SNPS
Cathepsin B	European	2018	3301	10,534,735
Cathepsin E	European	2018	3301	10,534,735
Cathepsin F	European	2018	3301	10,534,735
Cathepsin G	European	2018	3301	10,534,735
Cathepsin H	European	2018	3301	10,534,735
Cathepsin O	European	2018	3301	10,534,735
Cathepsin S	European	2018	3301	10,534,735
Cathepsin L2	European	2018	3301	10,534,735
Cathepsin Z	European	2018	3301	10,534,735
Nasopharyngeal carcinoma	European	2023	345,207	20,442,898

2.3 The selection of IVs

Given that there were few SNPs with CTS satisfying the $P < 5 \times 10^{-8}$ condition, we selected $P < 5 \times 10^{-6}$ as the threshold of significance level after comprehensive reference to the relevant literature, ensuring that IVs and CTS are closely related [22]. The exclusion of confounding factors is important for the independence assumption; thus, SNPs were assessed for confounding factors by manual search in the LDlink database. SNPs that were related to NPC were thought to be confounding factors in our study and were excluded according to the independence assumption. To obtain independent IVs, the linkage disequilibrium coefficient was set to $r^2 = 0.001$ and the region width was set to kb > 10,000. The *F*-value is closely related to the stability of the MR Analysis results; therefore, weak instrumental variables need to be excluded. This study used the formula for calculating *F*-scores to exclude weak instrumental variables: $F = R^2$ (n-k-1)/k (1-R^2) (*n*: the sample size, *k*: the number of IVs, R^2: the extent to which exposure is explained by the selected SNPs), $R^2 = 2 \times$ (1-MAF) $\times$ MAF $\times \beta^2$ (MAF is the minimum allele frequency, β is the effect size of the allele), and only IVs with $F > 10$ were retained [23]. Meanwhile, the palindromic SNPs with intermediate allele frequencies were removed by coordinating the effect alleles between the exposure and outcome data. In the reverse MR analysis, the SNP screening criteria we applied was the same as the forwards MR analysis, which is $P < 5 \times 10^{-6}$.

2.4 The forwards MR analysis

The"TwoSampleMR"package in R software (version 4.4.1) was used for data analysis. The results of Inverse-Variance Weighted (IVW), MR-Egger, Weighted Median, Simple Mode and Weighted Mode were used. Because the IVW method can fully refer to the Wald estimate of each IV in the absence of horizontal pleiotropy, the results are more stable and accurate. Therefore, in the absence of horizontal pleiotropy, the IVW was used as the primary analysis method. MR-Egger, Weighted Median, Simple Mode and Weighted Mode were used to validate the results from IVW.

2.5 Sensitivity analysis

Sensitivity analyses included Cochran's *Q* test, MR-PRESSO global test and leave-one-out test. Cochran's *Q* test as used to detect heterogeneity among IVs in the heterogeneity test [24].It was considered to have no significant heterogeneity when $P > 0.05$. MR-PRESSO global test was used to detect horizontal pleiotropy [25]. It can be considered as no horizontal pleiotropy when $P > 0.05$, which corresponds to MR-Egger's intercept approximates to 0. When horizontal pleiotropy ($P < 0.05$) was detected, the analysis would be repeated after outlier variants were removed. MR-PRESSO global test can also be used to validate MR Analysis results and detect outliers simultaneously. To verify that the causal relationship between CTS and NPC is not affected by a single SNP,"leave-one-out"sensitivity test was used [26]. In"leave-one-out"sensitivity test, forest plot can be used to show the MR estimate and 95% confidence interval values of every SNP [27]. The Bonferroni correction formula is 0.05/(number of exposures included in the study * number of outcomes included in the study) [28].

2.6 The reverse MR analysis and MR Steiger test

In the reverse MR analysis, genetic data of NPC were used as exposure, while genetic data of 9 CTS were used as outcome to determine whether there was a reverse causality between NPC and CTS. The screening criteria for SNPs were as serious as those in the forwards MR analysis.

To confirm the direction of causal action of CTS on NPC, MR Steiger test was used to examine the direction of the interaction between CTS and NPC [29]. Results were considered statistically significant at $P < 0.05$.

3 Results

3.1 The forwards MR analysis

We analyzed the causal relationship between nine common CTS and NPC. To account for any linkage disequilibrium, we excluded IVs with an r^2 value below 0.01 within a range of 1000 kb. The calculation formula of the *F* value was used to value the *F* values of the IVs selected in this study. The effect of them were found to be strong, because the *F* values were all greater than 10. No confounding factors were association with NPC. After Bonferroni correction, the *P*-values were below the Bonferroni threshold. According to the results of the forwards MR analysis, CTSF increased the risk of NPC (odds ratio [OR] = 1.845, 95% confidence intervals [CI] = 1.086 ~ 3.136, $P = 0.024$). In addition, the results of IVW and MR-Egger test, Weighted Median test, Simple Mode test and Weighted Mode test were consistent in the direction, which further proves the promoting effect of CTSF on the risk of NPC, as seen in Fig. 2 and Fig. 3.

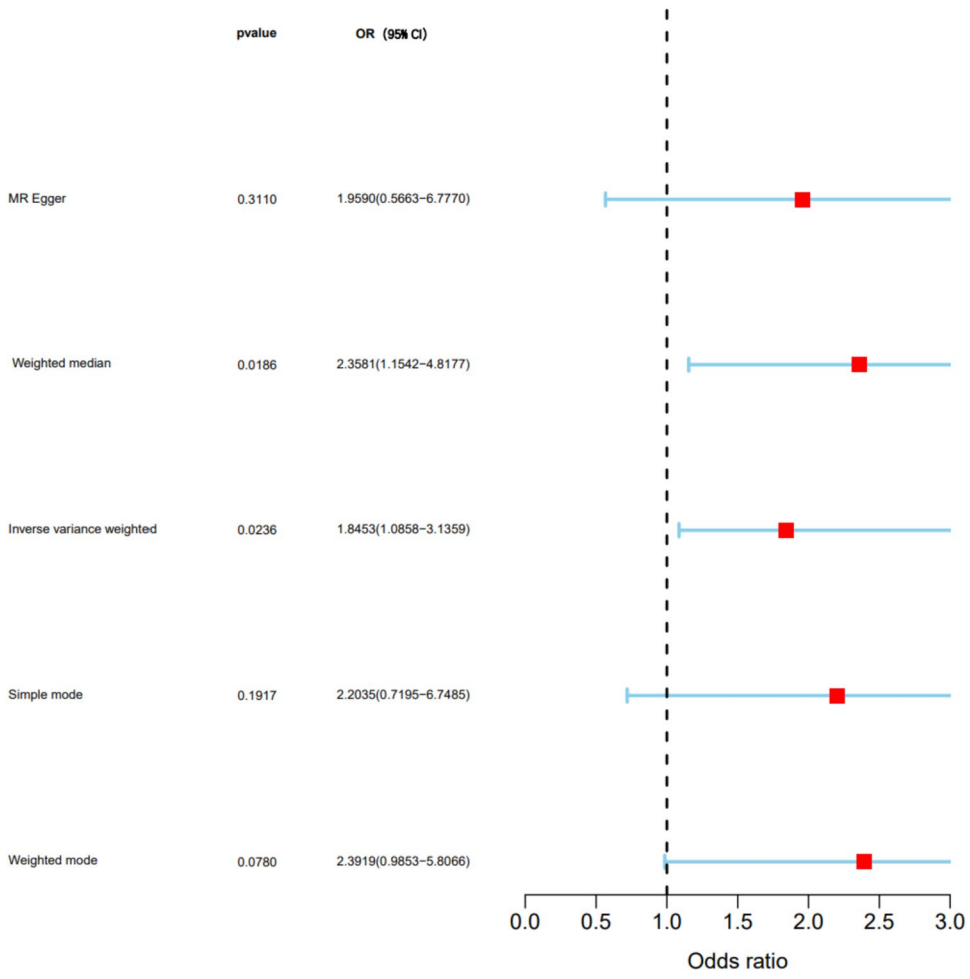

Fig. 2 Forest plot of OR values between CTSF and NPC. (*OR* odds ratio, 95% *CI* 95% confidence intervals)

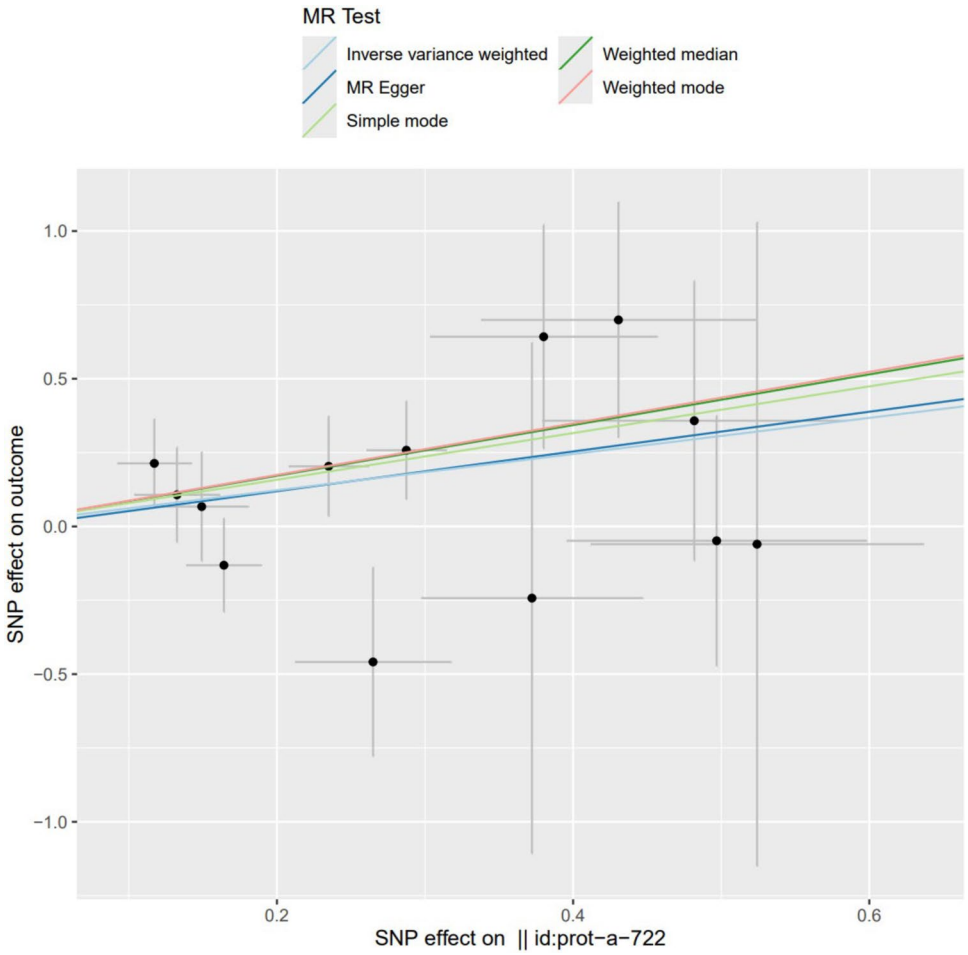

Fig. 3 Scatter plot of association between CTSF and NPC. The causality of each analysis is indicated by the slope of the respective line. (*OR* odds ratio, 95% *CI*, 95% confidence intervals)

However, in addition to CTSF, the other eight CTS did not show an obvious causal relationship with NPC risk in the MR Analysis. The results of the forwards MR analysis are as seen in Table 2.

3.2 Sensitivity analysis

Cochran's *Q* test indicated that there was no significant heterogeneity (IVW: $P = 0.467$, MR Egger: $P = 0.552$). The results for MR-PRESSO ($P = 0.579$) demonstrated the absence of horizontal pleiotropy and MR-Egger's intercept approximates to 0 (intercept = – 0.015, $P = 0.919$). The results of the Cochran's *Q* test and MR-PRESSO indicated that there was no heterogeneity among the included SNPs and no potential bias in the association between CTS and NPC, such that CTS only affects NPC through IVs. According to the"leave-one-out"sensitivity test, a single SNP had no significant effect on causality, confirming the robustness of the results. The forest plot of it showed that CTS played a positive role in regulating the risk of NPC. In the absence of SNPs exhibiting substantial heterogeneity, the majority of SNPs demonstrated an obvious impact on NPC. The 95% confidence intervals indicate that the effects were unidirectional. It confirmed the results of the forwards MR analysis, as seen in Fig. 4.

3.3 The reverse MR analysis and MR Steiger test

In the reverse MR analysis, although the SNP screening criteria we applied were the same as those in forwards MR analysis, there were no suitable SNPs. This suggests that genetic susceptibility to NPC does not significantly affect genetic

Table 2 MR results of 5 methods between CTS and NPC

Exposure	Outcome	No. of SNPs	Methods	OR (95%CI)	P
CTS B	NPC	19	MR Egger	1.186(0.373 ~ 3.778)	0.776
		19	Weighted median	0.806(0.396 ~ 1.641)	0.552
		19	Inverse variance weighted	0.916(0.562 ~ 1.495)	0.727
		19	Simple mode	0.900(0.273 ~ 2.969)	0.864
		19	Weighted mode	0.795(0.375 ~ 1.684)	0.556
CTS E	NPC	9	MR Egger	0.870(0.331 ~ 0.283)	0.785
		9	Weighted median	1.198(0.561 ~ 2.558)	0.641
		9	Inverse variance weighted	0.990(0.554 ~ 1.767)	0.970
		9	Simple mode	1.360(0.386 ~ 4.792)	0.646
		9	Weighted mode	1.247(0.448 ~ 3.471)	0.684
CTS F	NPC	13	MR Egger	1.959(0.566 ~ 6.778)	0.311
		13	Weighted median	2.358(1.141 ~ 4.875)	0.021
		13	Inverse variance weighted	1.845(1.086 ~ 3.136)	0.024
		13	Simple mode	2.203(0.732 ~ 6.630)	0.185
		13	Weighted mode	2.392(0.923 ~ 6.201)	0.098
CTS G	NPC	13	MR Egger	1.847(0.571 ~ 5.980)	0.328
		13	Weighted median	1.063(0.444 ~ 2.548)	0.891
		13	Inverse variance weighted	0.783(0.428 ~ 1.433)	0.427
		13	Simple mode	1.568(0.389 ~ 6.317)	0.639
		13	Weighted mode	1.153(0.477 ~ 2.787)	0.757
CTS H	NPC	16	MR Egger	1.178(0.288 ~ 4.820)	0.823
		16	Weighted median	1.042(0.453 ~ 2.398)	0.923
		16	Inverse variance weighted	1.120(0.557 ~ 2.252)	0.751
		16	Simple mode	1.380(0.365 ~ 5.520)	0.642
		16	Weighted mode	1.164(0.467 ~ 2.901)	0.748
CTS L2	NPC	11	MR Egger	0.729(0.094 ~ 5.623)	0.768
		11	Weighted median	0.921(0.341 ~ 2.487)	0.872
		11	Inverse variance weighted	0.986(0.460 ~ 2.116)	0.972
		11	Simple mode	0.751(1.177 ~ 3.193)	0.707
		11	Weighted mode	0.796(0.237 ~ 2.672)	0.720
CTS O	NPC	12	MR Egger	0.280(0.056 ~ 1.391)	0.151
		12	Weighted median	0.979(0.396 ~ 2.416)	0.963
		12	Inverse variance weighted	0.950(0.477 ~ 1.889)	0.882
		12	Simple mode	0.927(0.250 ~ 3.430)	0.912
		12	Weighted mode	0.917(0.259 ~ 3.251)	0.895
CTS S	NPC	23	MR Egger	0.811(0.386 ~ 1.703)	0.586
		23	Weighted median	0.843(0.454 ~ 1.566)	0.588
		23	Inverse variance weighted	1.070(0.694 ~ 1.651)	0.758
		23	Simple mode	1.241(0.444 ~ 3.469)	0.684
		23	Weighted mode	0.890(0.462 ~ 1.716)	0.732
CTS Z	NPC	13	MR Egger	1.369(0.530 ~ 3.535)	0.530
		13	Weighted median	1.224(0.624 ~ 1.403)	0.557
		13	Inverse variance weighted	1.455(0.830 ~ 2.552)	0.191
		13	Simple mode	1.273(0.460 ~ 3.520)	0.650
		13	Weighted mode	1.432(0.721 ~ 2.843)	0.325

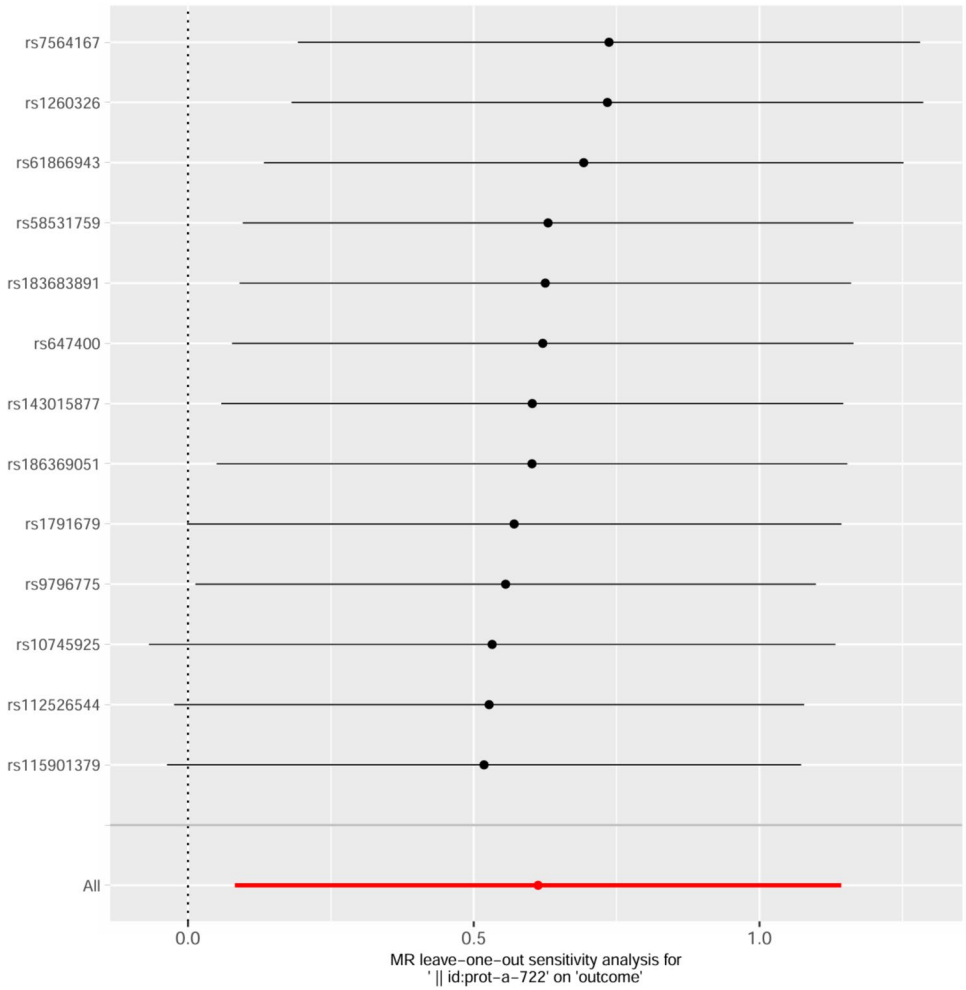

Fig. 4 The results of leave-one-out analysis of the causal effect between CTSF and NPC

susceptibility to CTS. The causal relationship between NPC and CTS is influenced by complex confounding factors, leading to the need to discuss the potential association between CTS and NPC.

The MR Steiger test proved that CTS did not have a reverse effect on NPC. $P < 0.05$ confirmed that the result was considered statistically significant (Supplementary file Table S2).

4 Discussion

The study investigated the causal relationship between 9 common CTS (CTS B, E, F, G, H, L2, O, S, and Z) and NPC. To exclude the interference of confounding factors, we selected the genetic data of 9 common CTS and NPC as the exposure data and outcome data for the forwards MR analysis under the premise of meeting the three principles of correlation assumption, independence assumption, and exclusion assumption. The results showed that CTSF might increase the risk of NPC (OR = 1.845, 95% CI = 1.086 ~ 3.136, $P = 0.024$), whereas the other 8 types of CTS had no causal effect on NPC. In contrast, NPC had no causal effect on CTS. This result was confirmed to be not affected by a single SNP with no heterogeneity and horizontal pleiotropy in heterogeneity and horizontal pleiotropy test.

NPC is a common head and neck cancer, which has caused significant harm in South China and Southeast Asia. The pathogenesis of NPC is related to genetics, diet, smoking, oral hygiene, Epstein-Barr virus infection, and other factors [1]. Among these, genetic factors play a role in the pathogenesis of NPC by affecting the immune response of genes, exogenous metabolism, and the cell cycle. Radiation therapy represents the primary treatment modality for NPC, and some advanced patients may need to be treated in combination with chemotherapy. Radiation therapy is a therapeutic

modality that can elicit anti-tumor biological effects by inducing damage or death to tumor cells through the application of ionizing radiation. Despite the proven efficacy of radiation therapy in treating NPC, the treatment inevitably causes damage to patients'normal head and neck tissues, resulting in a decline in patients'quality of life. The majority of NPC cells exhibit poor differentiation and rapid proliferation, resulting in a high probability of tumor recurrence or metastasis following treatment, which can ultimately lead to mortality [3, 30].

CTS is widely found in cellular tissues of various animals. Abnormalities in its active proteolytic activity or signal transduction pathway often leads to the occurrence of a variety of diseases, including cancers. Research has revealed that CTS has the capacity to degrade the extracellular matrix of surrounding cells, promote neovascularization and alter the tumor microenvironment by affecting cytokines and chemokines. As a consequence of these effects, tumor cell invasion and metastasis are promoted [31, 32]. To date, 15 CTS have been identified in the human body [33]. CTSB, expressed on the surface of tumor cells, plays a critical role in tumor progression and metastasis by degrading components of the extracellular matrix [34]. CTSG, a potential modulator of protein stability, has been associated with alterations in tumor-stroma interactions and an enhancement in cancer cell adhesion and migration [35]. CTSS, which can be upregulated in endothelial cells, is involved in neovascularization, providing oxygen and nutrients for cancer cell proliferation [36]. CTSF, also known as CATSF/CLN13, is widely involved in various protein catabolic processes in the human body [37, 38]. It is a potential target for cancers when its genetic variation is incorrect [39]. The elevated expression levels of CTSF in diverse human cancers have led to the proposition that its measurement in serum could serve as a diagnostic tool for cancer detection, with the notable benefit of minimal trauma [40].

Research has revealed that CTS has the potential to serve as a tumor marker for NPC, and it has been associated with metastasis and drug resistance in this malignancy [31]. CTS plays a pivotal role in the motility and invasion of NPC cells, serving as a crucial reference indicator for tumor metastasis [41]. CTS showed high expression in NPC, and this expression was found to be closely associated with lymph node metastasis and distant metastasis [42]. Patients diagnosed with NPC who exhibit elevated cathepsin expression levels frequently demonstrate a more negative prognosis.

CTSF expression increases in a multitude of cancer cells; however, the role of CTSF in tumorigenesis and progression requires further elucidation [43]. There remains a paucity of relevant studies examining the potential association between CTSF and the risk of NPC, and the stability of results is compromised by the influence of confounding factors. Although CTSF is widely expressed in cancers, such as NPC, its role in cancer progression remains unclear.

The results of our study proved whether CTS, including CTSF serve as potential candidate biomarkers of NPC, which might help us to discover NPC as soon as possible. To date, extant studies have not examined the causal relationship between CTS and NPC at the genetic level. This study was the first to address this research gap using MR. Controlling for confounding factors and reverse causality, this study found that CTSF in CTS might participate in the development of NPC. This finding indicates that CTSF has the capacity to function as a tumor marker or risk stratification factor, with the purpose of alerting patients to their potential risk of developing NPC and evaluating treatment efficacy. This purpose can be accomplished by measuring CTSF levels in the blood, which prompts clinicians to offer patients further tumor screening or a modification in the appropriate anti-tumor treatment regimen. Conversely, CTSF can be utilized as a therapeutic target, reducing the promotion of CTSF in NPC cells through the interference with CTSF expression.

5 Limitations

Although this study contributes to the understanding of the causal relationship between CTS and NPC, there are limitations to this study. The present study is based on data analyzed from a European population, so it is uncertain whether the results of this study are appropriate for interpretation in other regions and populations. Although this study went through several assumptions to avoid the horizontal pleiotropy and heterogeneity that exist in IVs, there may still be unknown potential biases that could affect our findings. This study only explored the relationship between CTS and NPC from the perspective of statistical analysis of data, and could not provide evidence of causality at the level of specific mechanisms. Therefore, the results of this study must be interpreted and applied with caution in actual clinical situations. Comprehensive evidence based on studies such as cellular experiments and clinical trials is needed to demonstrate that the conclusions of this study are reliable.

6 Conclusion

The study investigated the causal relationship between 9 common CTS (CTS B, E, F, G, H, L2, O, S, and Z) and NPC. To exclude the interference of confounding factors, we selected the genetic data of 9 common CTS and NPC as the exposure data and outcome data for the forwards MR analysis under the premise of meeting the three principles of correlation assumption, independence assumption, and exclusion assumption. The results showed that CTSF might increase the risk of NPC (OR = 1.845, 95% CI = 1.086 ~ 3.136, $P = 0.024$), whereas the other 8 types of CTS had no causal effect on NPC. NPC had no causal effect on CTS. This result was confirmed to be not affected by a single SNP with no heterogeneity and horizontal pleiotropy in heterogeneity and horizontal pleiotropy.

However, this study only found a causal relationship between CTSF and NPC from the aspect of MR and did not involve other research, including clinical and mechanistic research. In the future, the limitations of the conclusions will lead us to adopt more analytical methods to explore the role of CTSF in NPC risk.

Acknowledgements This research has been conducted using IEU Open GWAS project and FinnGen Consortium. The authors are grateful to IEU Open GWAS and FinnGen studies.

Author contributions JXH designed the study. JXH and KYL collected the data, analyzed the data and wrote the original draft. JXH, KYL, MJ and GYQ revised the original draft. MJ and GYQ are co-corresponding authors of this paper and GYQ is a financial supporter of this paper.

Funding This work was supported by the Foshan "14th Five-Year Plan" key speciality construction project of traditional Chinese medicine (Foshan Health Commission Letter [2020] No. 15).

Data availability The data that support the findings of this study are available from the IEU Open GWAS and FinnGen at https://gwas.mrcieu.ac.uk/ and https://r11.finngen.fi/.

Declarations

Ethics approval and consent to participate The data used in this study were obtained from publicly aggregated IEU Open GWAS and FinnGen. Therefore, no additional ethical approval was required for the study.

Consent for publication All authors have read and approved the submission of the manuscript.

Competing interests The authors declare that the research was conducted in the absence of any non-financial competing interests or financial competing interests that could be construed as a potential conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Su ZY, Siak PY, Lwin YY, Cheah SC. Epidemiology of nasopharyngeal carcinoma: current insights and future outlook. Cancer Metastasis Rev. 2024;43(3):919–39. https://doi.org/10.1007/s10555-024-10176-9.
2. Wang Z, Fang M, Zhang J, et al. Radiomics and deep learning in nasopharyngeal carcinoma: a review. IEEE Rev Biomed Eng. 2024;17:118–35. https://doi.org/10.1109/RBME.2023.3269776.
3. Juarez-Vignon Whaley JJ, Afkhami M, Onyshchenko M, et al. Recurrent/metastatic nasopharyngeal carcinoma treatment from present to future: where are we and where are we heading? Curr Treat Options Oncol. 2023;24(9):1138–66. https://doi.org/10.1007/s11864-023-01101-3.
4. Bei JX, Li Y, Jia WH, Feng BJ, Zhou G, Chen LZ, Feng QS, Low HQ, Zhang H, He F, Tai ES, Kang T, Liu ET, Liu J, Zeng YX. A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. Nat Genet. 2010;42(7):599–603. https://doi.org/10.1038/ng.601. (**Epub 2010 May 30 PMID: 20512145**).
5. Mat Lazim N, Che Lah CI, Wan Juhari WK, Sulong S, Zilfalil BA, Abdullah B. The role of genetic pathways in the development of chemoradiation resistance in nasopharyngeal carcinoma (NPC) patients. Genes (Basel). 2021;12(11):1835. https://doi.org/10.3390/genes12111835.

6. Xu YF, Li YQ, Liu N, He QM, Tang XR, Wen X, Yang XJ, Sun Y, Ma J, Tang LL. Differential genome-wide profiling of alternative polyadenylation sites in nasopharyngeal carcinoma by high-throughput sequencing. J Biomed Sci. 2018;25(1):74. https://doi.org/10.1186/s12929-018-0477-6.
7. Moschovi M, Critselis E, Cen O, Adamaki M, Lambrou GI, Chrousos GP, Vlahopoulos S. Drugs acting on homeostasis: challenging cancer cell adaptation. Expert Rev Anticancer Ther. 2015;15(12):1405–17. https://doi.org/10.1586/14737140.2015.1095095. (**Epub 2015 Nov 2 PMID: 26523494**).
8. 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(1):1019. https://doi.org/10.1038/s42003-023-05408-7.
9. Deng T, Lu X, Jia X, et al. Cathepsins and cancer risk: a Mendelian randomization study. Front Endocrinol (Lausanne). 2024;15:1428433. https://doi.org/10.3389/fendo.2024.1428433.
10. Khaket TP, Kwon TK, Kang SC. Cathepsins: Potent regulators in carcinogenesis. Pharmacol Ther. 2019;198:1–19. https://doi.org/10.1016/j.pharmthera.2019.02.003. (**Epub 2019 Feb 11 PMID: 30763594**).
11. Mohamed MM, Sloane BF. Cysteine cathepsins: multifunctional enzymes in cancer. Nat Rev Cancer. 2006;6(10):764–75. https://doi.org/10.1038/nrc1949. (**PMID: 16990854**).
12. Featherston T, Brasch HD, Siljee SD, van Schaijik B, Patel J, de Jongh J, Marsh RW, Itinteang T, Tan ST. Cancer stem cells in head and neck cutaneous squamous cell carcinoma express cathepsins. Plast Reconstr Surg Glob Open. 2020;8(8): e3042. https://doi.org/10.1097/GOX.0000000000003042.
13. Wu CC, Hsu CW, Chen CD, Yu CJ, Chang KP, Tai DI, Liu HP, Su WH, Chang YS, Yu JS. Candidate serological biomarkers for cancer identified from the secretomes of 23 cancer cell lines and the human protein atlas. Mol Cell Proteomics. 2010;9(6):1100–17. https://doi.org/10.1074/mcp.M900398-MCP200.
14. Lung ML, Cheung AK, Ko JM, Lung HL, Cheng Y, Dai W. The interplay of host genetic factors and Epstein-Barr virus in the development of nasopharyngeal carcinoma. Chin J Cancer. 2014;33(11):556–68. https://doi.org/10.5732/cjc.014.10170.
15. Ference BA, Holmes MV, Smith GD. Using mendelian randomization to improve the design of randomized trials. Cold Spring Harb Perspect Med. 2021;11(7):a040980. https://doi.org/10.1101/cshperspect.a040980.
16. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014;23(R1):R89-98. https://doi.org/10.1093/hmg/ddu328.
17. Deng T, Lu X, Jia X, Du J, Wang L, Cao B, Yang M, Yin Y, Liu F. Cathepsins and cancer risk: a Mendelian randomization study. Front Endocrinol (Lausanne). 2024;31(15):1428433. https://doi.org/10.3389/fendo.2024.1428433.
18. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P, Oliver-Williams C, Kamat MA, Prins BP, Wilcox SK, Zimmerman ES, Chi A, Bansal N, Spain SL, Wood AM, Morrell NW, Bradley JR, Janjic N, Roberts DJ, Ouwehand WH, Todd JA, Soranzo N, Suhre K, Paul DS, Fox CS, Plenge RM, Danesh J, Runz H, Butterworth AS. Genomic atlas of the human plasma proteome. Nature. 2018;558(7708):73–9. https://doi.org/10.1038/s41586-018-0175-2.
19. Bu F, Yu K, Ye C, Huang G, Yang T, Chen K, Lu J, Rong L. Can alterations in cathepsin levels restrain the development of skin cancer?: a bidirectional multivariate Mendelian-randomization study. Medicine (Baltimore). 2024;103(38):e39628. https://doi.org/10.1097/MD.0000000000039628.
20. Yi T, Lin S. The protective role of vitamin d in nasopharyngeal carcinoma: insights from Mendelian randomization and meta-analysis. Discov Oncol. 2024;15(1):637. https://doi.org/10.1007/s12672-024-01511-1.
21. Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature. 2023. https://doi.org/10.1038/s41586-023-05837-8.
22. Fang P, Liu X, Qiu Y, et al. Exploring causal correlations between inflammatory cytokines and ankylosing spondylitis: a bidirectional mendelian-randomization study. Front Immunol. 2023;14:1285106. https://doi.org/10.3389/fimmu.2023.1285106.
23. Pierce BL, Ahsan H, Vanderweele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants[J]. Int J Epidemiol. 2011;40(3):740–52. https://doi.org/10.1093/ije/dyq151.
24. Zhang F, Yu Z. Mendelian randomization study on insulin resistance and risk of hypertension and cardiovascular disease. Sci Rep. 2024;14(1):6191. https://doi.org/10.1038/s41598-023-46983-3.
25. Wu F, Huang Y, Hu J, Shao Z. Mendelian randomization study of inflammatory bowel disease and bone mineral density. BMC Med. 2020;18(1):312. https://doi.org/10.1186/s12916-020-01778-5.
26. Ma Z, He W. Fatty acids and pregnancy-induced hypertension: a Mendelian randomization study. Lipids Health Dis. 2023;22(1):131. https://doi.org/10.1186/s12944-023-01889-x.
27. Yang X, Wei J, Sun L, Zhong Q, Zhai X, Chen Y, Luo S, Tang C, Wang L. Causal relationship between iron status and preeclampsia-eclampsia: a Mendelian randomization analysis. Clin Exp Hypertens. 2024;46(1):2321148. https://doi.org/10.1080/10641963.2024.2321148.
28. Curtin F, Schulz P. Multiple correlations and Bonferroni's correction. Biol Psychiatry. 1998;44(8):775–7. https://doi.org/10.1016/s0006-3223(98)00043-2. (**PMID: 9798082**).
29. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. Int J Epidemiol. 2013;42(5):1497–501. https://doi.org/10.1093/ije/dyt179.
30. Wang BR, Han JB, Jiang Y, et al. CENPN suppresses autophagy and increases paclitaxel resistance in nasopharyngeal carcinoma cells by inhibiting the CREB-VAMP8 signaling axis. Autophagy. 2024;20(2):329–48. https://doi.org/10.1080/15548627.2023.2258052.
31. Chen ZT, Liang ZG, Zhu XD. A review: proteomics in nasopharyngeal carcinoma. Int J Mol Sci. 2015;16(7):15497–530. https://doi.org/10.3390/ijms160715497.
32. Mijanović O, Branković A, Panin AN, et al. Cathepsin B: a sellsword of cancer progression. Cancer Lett. 2019;449:207–14. https://doi.org/10.1016/j.canlet.2019.02.035.
33. Patel S, Homaei A, El-Seedi HR, Akhtar N. Cathepsins: proteases that are vital for survival but can also be fatal. Biomed Pharmacother. 2018;105:526–32. https://doi.org/10.1016/j.biopha.2018.05.148.
34. Aggarwal N, Sloane BF. Cathepsin B: multiple roles in cancer. Proteomics Clin Appl. 2014;8(5–6):427–37. https://doi.org/10.1002/prca.201300105.

35. Woo SM, Min KJ, Seo SU, et al. Up-regulation of 5-lipoxygenase by inhibition of cathepsin G enhances TRAIL-induced apoptosis through down-regulation of survivin. Oncotarget. 2017;8(63):106672–84. https://doi.org/10.18632/oncotarget.22508.
36. Smyth P, Sasiwachirangkul J, Williams R, Scott CJ. Cathepsin S (CTSS) activity in health and disease - A treasure trove of untapped clinical potential. Mol Aspects Med. 2022;88:101106. https://doi.org/10.1016/j.mam.2022.101106.
37. Ji C, Zhao Y, Kou YW, et al. Cathepsin F knockdown induces proliferation and inhibits apoptosis in gastric cancer cells. Oncol Res. 2018;26(1):83–93. https://doi.org/10.3727/096504017X14928634401204.
38. Song L, Wang X, Cheng W, Wu Y, Liu M, Liu R, Zhang S, Xia H, Liu H, Tai X, Zhao H, Li X, Ji F. Expression signature, prognosis value and immune characteristics of cathepsin F in non-small cell lung cancer identified by bioinformatics assessment. BMC Pulm Med. 2021;21(1):420. https://doi.org/10.1186/s12890-021-01796-w.
39. Santamaría I, Velasco G, Pendás AM, Paz A, López-Otín C. Molecular cloning and structural and functional characterization of human cathepsin F, a new cysteine proteinase of the papain family with a long propeptide domain. J Biol Chem. 1999;274(20):13800–9. https://doi.org/10.1074/jbc.274.20.13800. (**PMID: 10318784**).
40. Wei S, Liu W, Xu M, Qin H, Liu C, Zhang R, Zhou S, Li E, Liu Z, Wang Q. Cathepsin F and Fibulin-1 as novel diagnostic biomarkers for brain metastasis of non-small cell lung cancer. Br J Cancer. 2022;126(12):1795–805. https://doi.org/10.1038/s41416-022-01744-3.
41. Tan G, Liu Q, Tang X, et al. Diagnostic values of serum cathepsin B and D in patients with nasopharyngeal carcinoma. BMC Cancer. 2016;16:241. https://doi.org/10.1186/s12885-016-2283-4.
42. Wu CC, Hsu CW, Chen CD, et al. Candidate serological biomarkers for cancer identified from the secretomes of 23 cancer cell lines and the human protein atlas. Mol Cell Proteomics. 2010;9(6):1100–17. https://doi.org/10.1074/mcp.M900398-MCP200.
43. Zdravkova K, Mijanovic O, Brankovic A, et al. Unveiling the roles of cysteine proteinases F and W: from structure to pathological implications and therapeutic targets. Cells. 2024;13(11):917. https://doi.org/10.3390/cells13110917.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.