Mendelian randomization study of the association between cathepsins and melanoma

WENWEN WANG and JUN LI

Department of Dermatology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, P.R. China

Received April 23, 2024; Accepted June 21, 2024

DOI: 10.3892/wasj.2024.262

Abstract. Malignant melanoma is a skin tumor with a poor prognosis. Therefore, it is critical to explore the risk factors associated with the outcome of this tumor. In the present study, Mendelian randomization (MR) was used to investigate the causal association between cathepsins and malignant melanoma. Summary statistical data on five cathepsins from European participants were extracted as exposure data. Data on melanoma from a genome-wide association study of European ancestry were used as outcome data. Single nucleotide polymorphisms associated with cathepsins were used as instrumental variables (IVs). In a genome-wide association study of malignant melanoma including 3,751 melanoma cases and 372,016 European ancestry controls, MR analysis was conducted to examine the effects of these IVs on melanoma. The inverse variance-weighted method was used for MR analysis. In addition, MR-Egger, weighted median and MR pleiotropy residual sum were used for complementary analyses. Furthermore, a series of sensitivity analyses were performed to ensure the validity and robustness of the results. The gene-predicted results indicated no causal association between the five cathepsins and malignant melanoma (P>0.05). Cathepsin S (odds ratio (OR), 1.000; 95% confidence interval (CI), 0.999-1.001; P=0.943), cathepsin B (OR, 1.000; 95% CI, 0.999-1.001; P=0.763), cathepsin O (OR, 1.000; 95% CI, 0.999-1.001; P=0.646), cathepsin E (OR, 0.999; 95% CI, 0.998-1.001; P=0.375) and cathepsin L2 (OR, 1.101; 95% CI, 0.831-1.458; P=0.503) were not significantly associated with the risk of developing melanoma. Sensitivity analysis demonstrated no significant bias in the aforementioned results. On the whole, in the present study, MR analysis did not provide evidence that cathepsins (cathepsin S, cathepsin B, cathepsin O, cathepsin E and cathepsin L2) are causally related to malignant melanoma.

Introduction

Melanoma is an invasive malignant tumor originating from melanocytes (1). The incidence of melanoma exhibits a linear increase in the age group of 25 to 50 years (2). Although melanoma is less common than other types of skin cancer, it accounts for 73% of skin cancer-related mortality (3), with a 5-year survival rate of only ~18%. It is considered that the incidence and mortality rates of melanoma will continue to increase over the next 10 years (4,5). Regrettably, early surgical treatment and late-stage radiation therapy and chemotherapy have not yielded satisfactory therapeutic effects for patients with melanoma (6). Hence, the identification of the risk factors involved in the occurrence of melanoma is of utmost importance.

Studies have indicated a marked increase in the expression and secretion of cathepsins in cancer, which play recognized roles in cancer progression (7,8). Within tumor cells, cathepsins and their messenger RNAs are frequently upregulated, leading to the secretion of excess proenzyme (unprocessed form) (7). Extracellular cathepsins have been isolated from cancer cell culture media, constituting 40% of total secreted proteins, and have been observed in the serum of patients with cancer (9). Cathepsins play crucial roles in the occurrence, development and metastasis of melanoma (8,10). Secreted cathepsins promote melanoma invasion and metastasis by directly and indirectly facilitating extracellular matrix degradation (11,12). Saenger et al (13) utilized reverse transcription-polymerase chain reaction to measure preprocessed peripheral blood samples from 218 patients with melanoma, demonstrating that the expression levels of cathepsins in peripheral blood can predict the survival rates of patients with melanoma. Elevated levels of cathepsins were shown to be associated with the favorable prognosis of patients with melanoma (13). Kos et al (14) used quantitative enzyme-linked immunosorbent assay to measure cathepsin B levels in the serum of 43 patients with metastatic melanoma, 54 patients with in situ cutaneous melanoma (the absence of evidence of tumor metastasis) and 30 healthy blood donors. The results of their study revealed significantly higher levels of cathepsin B in the serum of patients with metastatic melanoma compared with healthy...
individuals, and patients with metastatic melanoma with high levels of cathepsin B exhibited markedly lower overall survival rates than those with low levels of cathepsin B (14). These associations suggest a link between cathepsins and malignant melanoma; however, causal associations cannot be established from observational studies due to confounding and reverse causality. These studies may be influenced by various confounding factors, such as the voluntary treatment choices of patients and incomplete data records, potentially resulting in the decreased internal validity of the research findings. Furthermore, causal associations cannot be established from these studies. Manipulating independent variables through data collection and outcome observation does not allow for the verification of causality.

In the present study, Mendelian randomization (MR) was used to validate the causal associations in assessing the role of cathepsins in the development of melanoma. With MR, the impact of exposure (cathepsins) on the risk of developing melanoma is assessed using genetic instrumental variables (IVs) (15). As IVs are randomly distributed in conceptualization, they are not influenced by confounding factors. In genome-wide association studies (GWASs), the common genetic variants associated with cathepsins serve as IVs for measuring cathepsins. The present study also discusses the causal risk of melanoma development from a genetic perspective, providing a basis for further prevention and diagnosis.

Materials and methods

Study design. A two-sample MR was conducted utilizing GWAS data. The MR method was based on the following three assumptions (Fig. 1) (16,17): i) The genetic variants selected as IVs are associated with cathepsins; ii) the genetic variants are independent of any unmeasured confounding factors; iii) the genetic variants are associated with malignant melanoma solely through cathepsins and not through other pathways.

Data sources. The GWAS involved in malignant melanoma and cathepsins utilized data from GWAS databases. Specifically, the GWAS data for malignant melanoma were sourced from the UK Biobank (https://www.ukbiobank.ac.uk), and the data for cathepsins were obtained from the GWAS Catalog (https://gwas.mrcieu.ac.uk/datasets). The summarized levels of cathepsin S, cathepsin B, cathepsin O, cathepsin E and cathepsin L2 in the European population were extracted from the comprehensive summary data of GWAS protein-level summaries described by Sun et al (18). As regards the outcome data details, the melanoma data from the UK Biobank included 3,751 cases and 372,016 controls, covering a total of 11,396,019 single nucleotide polymorphisms (SNPs). All melanoma cases met the diagnostic criteria for melanoma: immunohistochemical staining revealed positivity for characteristic markers of melanocytes, including S100, Melan-A, and HMB45 (19). The study design, quality control procedures, phenotype definition and inference methods for GWAS were described in the previous publication where sample collection was reported (20). Participants in these data sources were of European ancestry. The intentional selection of European participants helps ensure sample population homogeneity, reduce confounding factors and enhance result reliability. As the data included in the present study have been previously published, ethical approval or informed consent were not required (20). Detailed information regarding the data sources is provided in Table I, along with corresponding links to the specific data sources (Table S1).

Genetic variants. As regards the exposure data, genetic variants exceeding the genome-wide association thresholds ($P<5.0\times10^{-8}$) were selected as IVs, and a clumping algorithm ($\pi^2$ threshold=0.001; kb=10,000 mB) was used to eliminate linkage disequilibrium. Additionally, the strength of the instrument was assessed using the F-statistic (F-value calculation formula below), which is approximated by the squared SNP-phenotype association divided by its variance. A potential weak instrument is indicated by an F-statistic <10 (21). Therefore, strongly correlated IVs, based on an F-statistic >10, were selected to meet the assumption of correlation in MR analysis as follows:

$$F = \frac{\beta^2}{SE^2}$$

where beta represents the effect size of SNP exposure, and SE is the standard error of SNP exposure. A total of 23 SNPs were found to be associated with cathepsin S, 17 SNPs with cathepsin B, 11 SNPs with cathepsin O, 11 SNPs with cathepsin E and 11 SNPs with cathepsin L2. The effect sizes of changes in cathepsins for each additional effect allele are presented in Tables SII-SVI. No linkage disequilibrium was observed among the SNPs.

Two-sample MR analysis. MR analysis, a prominent tool for determining the causal effect of exposure variables on outcomes via genetic variation, was used to confirm the causal association between cathepsins and melanoma (22,23). Various MR techniques, including inverse variance weighted (IVW), MR-Egger, weighted median (WM) and weighted mode, were utilized. Notably, the IVW and MR-Egger techniques, commonly utilized as fundamental methodologies in MR analysis worldwide, play essential roles. An IVW analysis was conducted, utilizing multiplicative random effects assuming balanced pleiotropy, with SNP-specific Wald estimates, i.e., SNP-outcome association divided by SNP-exposure association, as the main analysis. Additionally, sensitivity analyses were performed using WM and MR-Egger (24). The WM assumes that 50% of the weight comes from valid SNPs. Furthermore, MR-Egger exhibited robustness to genetically invalid instruments due to instrument strength independent of direct effect, meaning the instruments do not confound the exposure of the outcomes. A zero MR-Egger intercept indicated no evidence for this genetic pleiotropy. IVW, serving as the main tool, determines causal links between exposure variables and outcomes in MR analysis, with a result deemed significant when the P-value of IVW is <0.05. Additionally, under the condition of IVW, the direction of MR-Egger and the WM method must align with that of IVW. To eliminate bias, MR-PRESSO (https://github.com/roundlab/MR-PRESSO) was used to detect and correct potentially pleiotropic outliers (SNPs) for all reported results. Heterogeneity, quantified using Cochran Q statistics and I² statistics, increased with larger I² values. Further analysis included leave-one-out analysis...
conducted by removing each SNP to assess the stability and reliability of the MR results. The final results, depicted using forest plots, scatter plots and funnel plots, showcased the effectiveness of the MR analysis methods. These methods were implemented using the ‘Two-SampleMR’ (https://mrcieu.github.io/TwoSampleMR) and ‘MR-PRESSO’ (https://github.com/rondolab/MR-PRESSO) R packages in R (version 4.0.3).

Statistical analysis. IVW: IVW is a method used in MR for the meta-analysis of the effects of multiple SNPs across different locus. The prerequisite for applying IVW is that all SNPs serve as valid IVs and are completely independent of each other. WM: WM represents the median of the distribution function obtained by sorting the individual SNP effect values according to weights. When at least 50% of the information comes from valid IVs, WM can provide robust estimates. Weighted mode: The weighted mode method evaluates the combined effects of different genotypes on the phenotype by calculating the weighted average of each genotype. This approach can better control the impact of genotype frequency differences on the analysis results, thereby improving the robustness and accuracy of the analysis. MR-Egger method: MR-Egger does not enforce the regression line to pass through the origin, allowing for the inclusion of IVs with directional pleiotropy. The presence of genetic pleiotropy is indicated when the regression intercept is non-zero and P<0.05. MR-PRESSO method: MR-PRESSO can be used to obtain more accurate estimates by identifying and excluding specific SNPs as outliers, thus reducing the influence of potential confounding factors in MR analysis.

Results

Selection of IVs. Following an extensive quality control review, 72 SNPs associated with five cathepsins were identified as IVs for melanoma. Notably, the cathepsin S levels were linked to 23 SNPs, cathepsin B to 17 SNPs, cathepsin O to 11 SNPs, cathepsin E to 11 SNPs and cathepsin L2 to 11 SNPs (Tables I and SII-SV1).

Causal role of cathepsins in melanoma. In the analysis, IVW was used to provide estimates of the causal association between cathepsins and melanoma, indicating that cathepsin S [odds ratio (OR), 1.000; 95% confidence interval (CI)]

<table>
<thead>
<tr>
<th>Trait</th>
<th>SNPs</th>
<th>OR (95% CI)</th>
<th>P-value</th>
<th>IVW MR-Egger Weighted mode</th>
<th>OR (95% CI)</th>
<th>P-value</th>
<th>OR (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cathepsin S</td>
<td>23</td>
<td>1.000 (0.999-1.001)</td>
<td>0.943</td>
<td>0.943</td>
<td>1.000 (0.999-1.001)</td>
<td>0.24</td>
<td>1.000 (0.999-1.001)</td>
<td>0.24</td>
</tr>
<tr>
<td>Cathepsin B</td>
<td>17</td>
<td>1.000 (0.999-1.001)</td>
<td>0.646</td>
<td>0.646</td>
<td>1.000 (0.999-1.001)</td>
<td>0.29</td>
<td>1.000 (0.999-1.001)</td>
<td>0.29</td>
</tr>
<tr>
<td>Cathepsin O</td>
<td>10</td>
<td>1.000 (0.999-1.001)</td>
<td>0.375</td>
<td>0.375</td>
<td>1.000 (0.999-1.001)</td>
<td>0.37</td>
<td>1.000 (0.999-1.001)</td>
<td>0.37</td>
</tr>
<tr>
<td>Cathepsin E</td>
<td>1</td>
<td>1.000 (0.999-1.001)</td>
<td>0.872</td>
<td>0.872</td>
<td>1.000 (0.999-1.001)</td>
<td>0.87</td>
<td>1.000 (0.999-1.001)</td>
<td>0.87</td>
</tr>
<tr>
<td>Cathepsin L2</td>
<td>11</td>
<td>1.000 (0.999-1.001)</td>
<td>0.441</td>
<td>0.441</td>
<td>1.000 (0.999-1.001)</td>
<td>0.44</td>
<td>1.000 (0.999-1.001)</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Estimates of inverse-variance weighted, MR-Egger and weighted mode of matrix metalloproteinases on sepsis in the Mendelian randomization analysis. SNPs, single nucleotide polymorphisms; IVW, inverse variance weighted.
0.999-1.001; P=0.943], cathepsin B (OR, 1.000; 95% CI, 0.999-1.001; P=0.763), cathepsin O (OR, 1.000; 95% CI, 0.999-1.001; P=0.646), cathepsin E (OR, 0.999; 95% CI, 0.998-1.001; P=0.375) and cathepsin L2 (OR, 1.101; 95% CI, 0.831-1.458; P=0.503) were not significantly associated with the risk of developing melanoma. The overall estimates of the MR Egger and WM methods were consistent with the IVW analysis (Table I).

Sensitivity analyses. For cathepsin S, neither the IVW test (Q=27.421, P=0.196) nor the MR-Egger test (Q=27.415, P=0.158) revealed significant heterogeneity (P>0.05). Similarly, for cathepsin B (IVW test: Q=19.013, P=0.268; MR-Egger test: Q=18.782, P=0.224), cathepsin O (IVW test: Q=5.165, P=0.820; MR-Egger test: Q=4.114, P=0.847), cathepsin E (IVW test: Q=4.086, p = 0.943; MR-Egger test: Q=3.566, P=0.938) and cathepsin L2 (IVW test: Q=12.683, P=0.242; MR-Egger test: Q=11.644, P=0.234), no apparent heterogeneity was observed (Table SVII). Additionally, the MR-Egger intercept test indicated the absence of horizontal pleiotropy (P>0.05) (Table SVIII). Scatter plots demonstrated the genetic association between cathepsins and melanoma (Fig. S1). Furthermore, funnel plots revealed no significant asymmetry, suggesting negligible publication bias and directional horizontal pleiotropy (Fig. S2). Following a systematic elimination of SNPs through leave-one-out analysis, the outcomes showed minimal alteration, indicating that no single variant influenced the association between cathepsins (cathepsin S, cathepsin B, cathepsin O, cathepsin E and cathepsin L2) and melanoma (Fig. 2).

Discussion

In the present study, genetic variants were utilized to assess whether a risk factor has a causal effect on the results of MR studies. To the best of our knowledge, this is the first MR study aimed at evaluating the impact of cathepsins on the occurrence of melanoma. The present study found no causal association between several cathepsins (cathepsin S, cathepsin B, cathepsin O, cathepsin E and cathepsin L2) and melanoma among Europeans. In other words, alterations in cathepsins levels are not associated with the occurrence of melanoma.

Cathepsins are a group of enzymes involved in determining the metastatic potential of cancer cells (10), including cysteine cathepsins (cathepsin B and cathepsin L) and aspartyl protease (cathepsin D), which are typically present in the form of inactivezymogens in lysosomes. Once released into the extracellular space, cathepsins contribute to the enhancement of tumor metastatic potential by promoting cell migration and invasion capabilities (25,26). However, the current understanding of the association between cathepsins and melanoma is still at a laboratory experiment phase. Only a limited number of clinical observational studies are currently available; however, due to limitations in sample size, the results of these observational studies are not yet reliable (13,14,27). In the present study, not all capabilities were selected as exposure factors for MR analysis, as other capabilities are rarely reported to be closely associated with cutaneous melanoma. Previous studies have indicated that various cathepsins are overexpressed in melanoma. Furthermore, cathepsin B co-localizes with HMB-45-positive cells in malignant melanoma (28). It is noteworthy that the levels of cathepsin B and H are significantly elevated in metastatic melanoma patient groups compared to healthy control groups (14). The transition of human melanoma cells from a non-metastatic to a highly metastatic phenotype is directly associated with the secretion of cathepsin L (29). The precursor of cathepsin-L is a protease capable of cleaving human C3 (the third component of complement), imparting high tumorigenicity and metastatic potential to human melanoma cells (29). Cathepsin S is a single-chain non-glycosylated enzyme, the only member among 11 cathepsins (Cat B, C, F, H, L, K, O, S, V, W, X/Z) possessing protein hydrolysis activity at neutral pH (30). Research has also revealed increased expression of cathepsin S in certain types of cancer, such as colorectal, gastric and breast cancer (31-33), with its crucial role demonstrated in tumor invasion and metastasis through inducing tumor angiogenesis and degradation of the extracellular matrix (34,35). The inhibition of cathepsin S expression has been shown to be associated with the inhibition of malignant phenotypes in cancer cells and improved clinical outcomes in patients with breast cancer (33,36). Additionally, specific cathepsin S activity in primary choroidal melanoma is the strongest predictor of tumor metastatic behavior (37). Notably, unlike other cathepsins, cathepsin E has been found to be associated with antitumor activity in vitro, as research using mice carrying human and murine tumor xenografts has demonstrated the antitumor activity of cathepsin E (38). In addition, the injection of purified cathepsin E into nude mice carrying human tumor xenografts has been shown to induce tumor cell apoptosis in a dose-dependent manner and to inhibit tumor growth (38).

Although some of the aforementioned studies strongly suggest that cathepsins in serum can serve as serum markers for the early diagnosis and prognostic prediction of melanoma, the present study did not provide sufficient evidence to support this conclusion. Serum markers possess advantages, such as easy detection, minimal patient trauma and a low cost, rendering them markedly advantageous for the early screening and follow-up care of patients with cancer, and they are also a current focus and priority of research (39,40). Tumor markers for melanoma lack specificity; hence, broad-spectrum tumor markers, such as CA50, CA199, etc., can serve as melanoma tumor markers. Since these tumor markers are also highly expressed in patients with other types of cancer, there is a need to discover sensitive and specific biomarkers for cutaneous malignant melanoma (41). The cathepsins hold value for the early diagnosis and prognosis assessment of patients with melanoma; however, the results are typically derived from clinical studies with small sample sizes, lacking authenticity and reliability (14,27). In conclusion, diagnosing melanoma using cathepsins remains challenging; therefore, further evaluation using large sample sizes, multicenter, and rigorously designed studies are required to validate their authenticity. Further prospective studies are warranted to verify their effectiveness, and concrete steps should be taken to establish standards for guiding the effective use of diagnostic and prognostic assessment tools.

In addition, the present study aimed to increase the number of SNPs to alleviate bias caused by limited IVs; therefore, relevant literature on MR of cathepsins was examined and a
Figure 2. Leave-one-out analysis of cathepsins on melanoma skin cancer in the Mendelian randomization study. MR, Mendelian randomization. (A) Associations between cathepsin S and melanoma. (B) Associations between cathepsin B and melanoma. (C) Associations between cathepsin O and melanoma. (D) Associations between cathepsin E and melanoma. (E) Associations between cathepsin L2 and melanoma.
significance threshold with a P-value <5x10^{-6} was established. However, caution is required in interpreting the results. Third, the MR analysis method is a theoretical causal analysis method that requires further validation through animal experiments to establish causality. On the whole, the findings in the present study may aid in the understanding of the complex mechanisms linking cathepsins and melanoma.

In conclusion, the present study found no evidence of a causal association between cathepsins and melanoma, which contradicts the results of the majority of previous observational studies, as aforementioned. Prior research has suggested a potential association between melanoma and cathepsins (42); however, confounding factors or reverse causality may also be at play. Further investigations are required to fully elucidate the association between melanoma and cathepsins.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author's contributions

WW contributed to the conceptualization and methodology of the study. JL and WW were involved in the data analysis and visualization of the results. WW participated in the drafting and reviewing of the main manuscript. JL and WW have reviewed and approved the final manuscript. JL and WW confirm the authenticity of all the raw data.

Ethics approval and consent to participate

Two-sample MR was conducted using GWAS data. The study did not require ethical approval since all data were derived from summary statistics from previously published GWASs.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References


