

A bidirectional Mendelian randomization study integrating genome-wide association studies, expression quantitative trait locus, and methylation quantitative trait locus data revealed causal relationship between heavy cigarette dependence and Barrett's esophagus

SAGE Open Medicine

Volume 13: 1–10

© The Author(s) 2025

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/20503121251316595

journals.sagepub.com/home/smo

Zhou An¹, Meichun Zeng² and Xianhua Wang³ 

Abstract

Background: The association between smoking dependence and the risk of developing Barrett's esophagus remains unclear. This study aimed to investigate whether a causal relationship exists between smoking dependence and Barrett's esophagus, using Mendelian randomization analysis.

Methods: Two-sample Mendelian randomization analysis was conducted to evaluate the impact of smoking and Barrett's esophagus. Additionally, we applied summary data-based Mendelian randomization techniques to combine information from genome-wide association studies (GWAS) with expression quantitative trait locus and methylation quantitative trait locus.

Results: Multivariable Mendelian randomization showed an association between smoking per day (odds ratio = 1.2, 95% confidence interval: 1.038–1.38, $p=0.014$) or current smoking (odds ratio = 2.41, 95% confidence interval: 1.06–5.5, $p=0.037$) and Barrett's esophagus. Inverse variance-weighted methods of bidirectional Mendelian randomization also revealed that smoking per day was significantly associated with elevated risks of Barrett's esophagus (odds ratio = 1.34, 95% confidence interval: 1.092–1.649, $p=0.005$), while Barrett's esophagus was also a susceptibility factor for smoking per day (odds ratio = 1.05, 95% confidence interval: 1.017–1.087, $p=0.003$). By incorporating consistent summary data-based Mendelian randomization associations between DNA methylation and smoke/Barrett's esophagus, gene expression and smoke/Barrett's esophagus, and DNA methylation and gene expression, we identified that the genetic variant-cg00935895-RBM43 (ENSG00000184898)-smoke/Barrett's esophagus axis exerted an effect on smoke/Barrett's esophagus by altering the DNA methylation level, which regulated the expression level of RBM43.

Conclusions: Our study provides evidence of a bidirectional causal association between smoking and Barrett's esophagus from a genetic perspective, which sheds new light on the potential role of RBM43 as a mediator in facilitating the impact of smoking and Barrett's esophagus.

Keywords

Barrett's esophagus, smoke, cigarette dependence, Mendelian randomization, SMR

Date received: 15 October 2024; accepted: 8 January 2025

Background

Barrett's esophagus (BE) is an intestinal metaplasia caused by chronic damage to the esophageal lining and is a known precursor of esophageal adenocarcinoma (EAC).^{1,2} The current guidelines suggest screening for BE in populations with multiple risk factors. A system is needed to identify individuals within the general population who are at elevated risk of

¹Department of Thoracic Surgery, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

²Department of Pathology, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

³Department of Endoscopy, Hangzhou Hospital of Traditional Chinese Medicine, Hangzhou, Zhejiang, China

Corresponding author:

Xianhua Wang, Department of Endoscopy, Hangzhou Hospital of Traditional Chinese Medicine, 453 Stadium Road, Hangzhou, Zhejiang 310007, China.
Email: xhwang1212@163.com



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons

Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

BE, a precancerous condition associated with EAC, to facilitate appropriate endoscopic surveillance.³

Previous studies have summarized the key risk factors for BE progression and have highlighted promising new risk stratification tools.⁴ Studies have also concluded that with the increasing prevalence of cannabis smoking, endoscopic surveillance guidelines may need to be modified for individuals who chronically smoke cannabis.⁵ A multicenter study⁶ conducted in Europe demonstrated that smoking, even in the absence of gastroesophageal reflux disease (GERD), acts as an independent risk factor for BE. The exact internal mechanisms by which smoking contributes to the development of BE remain unclear, warranting further investigation of the pathophysiological processes involved. Nonetheless, these findings underscore the importance of smoking cessation in mitigating risk, although not entirely in eliminating it.

Building on these findings, our study conducted bidirectional validation using databases from different sources. Furthermore, we aim to provide additional insights into BE by incorporating summary data-based Mendelian Randomization (SMR) analysis. Specifically, we will use whole-blood expression quantitative trait locus (eQTL) and methylation quantitative trait locus (mQTL) data to explore smoking and BE. This approach allowed us to leverage large-scale genomic data to strengthen the evidence base and to address the limitations of previous studies. Recent advances in genetic epidemiology have emphasized the importance of integrating multiomics data, such as eQTL, to better understand the complex disease mechanisms.^{7,8} By incorporating BE eQTL data, studies have identified candidate genes whose expression is regulated by risk variants in BE-relevant tissues and have facilitated the elucidation of BE pathophysiology.⁹

BE is associated with chronic GERD¹⁰ and socioeconomic status, such as smoking,¹¹ alcohol consumption, and BMI.¹² However, smoking habits that increase the risk of BE warrant further investigation.

Previous findings hold profound significance for screening patients with BE who may be at risk of deterioration, as they can provide timely medical interventions to reverse the condition.¹³ Physicians may also benefit from additional BE education and better methods of risk stratification.¹⁴ Thus, regarding smoking behavior, it is feasible to identify individuals who are most at risk of developing BE. Moreover, acknowledging the risks and implementing suitable measures can contribute to the prevention of BE, thus mitigating the associated health burden.

Methods

Study design

Multivariable Mendelian randomization (MR) analysis was used to investigate the causal relationship between five smoking habits and BE, and univariable MR analysis was performed

to investigate the results of multivariable MR. Representative phenotypic single-nucleotide polymorphisms (SNPs) were selected as instrumental variables¹⁵ (Figure 1(a)). To serve as effective instrumental variables in phenotype studies, SNPs must satisfy three key criteria: first, they must exhibit a robust association with the phenotype in question; second, they must remain independent of any confounders that could influence the relationship between exposure and outcome; and third, there should be no direct causal pathways between the SNPs and the outcome.¹⁶

The information used in this study was sourced from publicly accessible databases that had previously received ethical clearance from their original studies. Consequently, no further ethical approval was required for this analysis. However, to ensure the validity of the MR approach, it is critical that the study population be accurately phenotyped. Therefore, we reviewed the diagnostic procedures and ensured that the diagnostic tests for BE were uniformly applied across both BE patients and non-BE controls to minimize the risk of false negatives and other bias. The details of the study design are given in Figure 1. The inclusion criteria for participants were strictly defined to ensure comparability between the groups. Additional sensitivity analyses were conducted to account for the potential bias caused by incomplete phenotyping. Our analysis was conducted in January 2024, encompassing data processing, statistical analysis, and interpretation.

Data sources

Genetic instruments for smoking and BE were selected from summary statistics of genome-wide association studies (GWAS) meta-analysis performed on data from the UK Biobank (<https://www.ukbiobank.ac.uk/>)¹⁷ and the OpenGWAS database (<https://gwas.mrcieu.ac.uk>) in European samples, which is an open source, open database, through strict quality control.¹⁸ Details of the data sources are given in Figure 1(b). Additionally, only individuals of European ancestry were included to minimize population stratification effects.

Genetic instrumental variables selection

A threshold of $p < 5 \times 10^{-5}$ was used for genomewide significance in this study. Accordingly, we identified SNPs that showed a strong association with each dietary pattern using this criterion. To ensure the independence of these SNPs, linkage disequilibrium (LD) clumping was conducted using specific parameters ($r^2 < 0.001$ and a window size of 10,000 kb). Additionally, the F statistic was computed to evaluate the potential bias introduced by weak instrumental variables.¹⁹ An F -statistic greater than 10 signifies that the influence of such bias can be considered negligible.¹⁹ The exposure and outcome data were then integrated and standardized according to the effect alleles. SNPs linked to outcomes were excluded from the combined dataset.

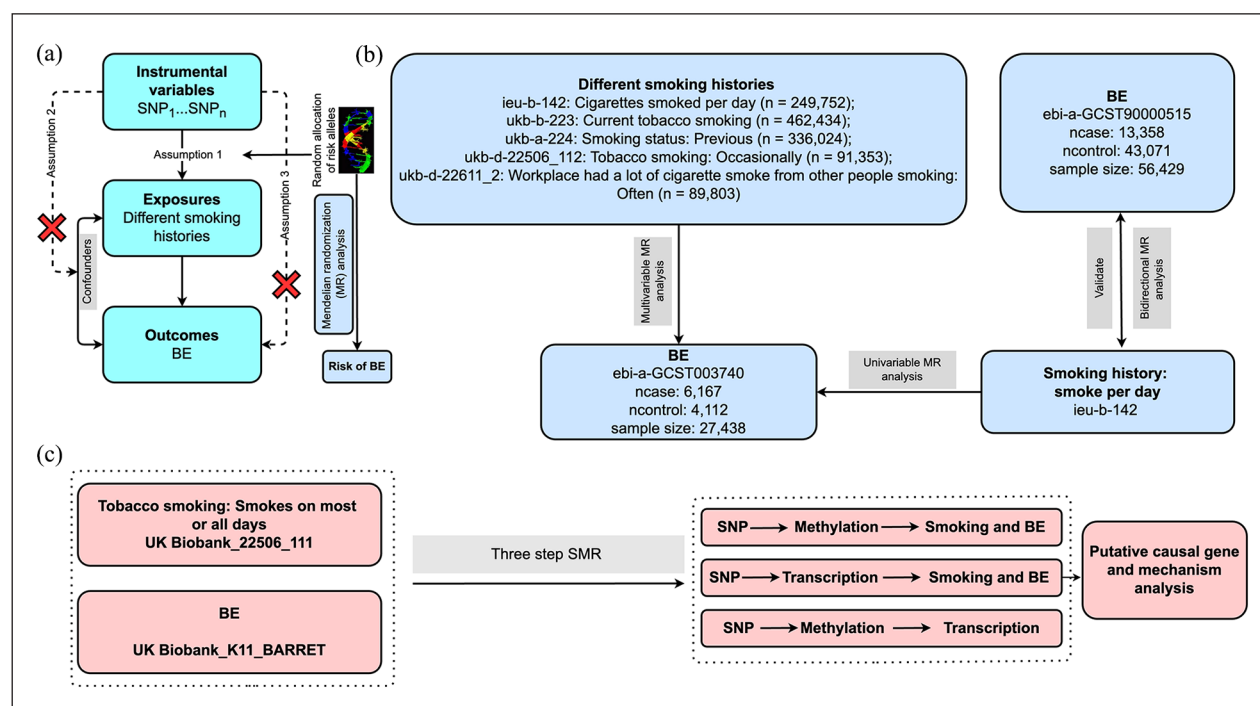


Figure 1. Flowchart of the study design. (a) MR analysis; (b) data sources utilized and the conduct of the MR analysis; and (c) summarizes the three-step SMR analysis.

BE, Barrett's esophagus; MR, Mendelian randomization; SMR, summary data-based Mendelian randomization.

Heterogeneity and pleiotropy analysis

Cochran's Q test was used to detect heterogeneity, with a significant p -value indicating the presence of variation among studies. To evaluate horizontal pleiotropy, MR-Egger intercept analysis was conducted. Additional sensitivity analyses were performed to confirm the robustness of the findings. A p -value for Cochran's Q statistic less than 0.05, it suggested the presence of heterogeneity. Forest plots and leave-one-out analyses were used to identify any single SNP that may have introduced bias into the results. After eliminating the identified outliers, the MR analysis was repeated to ensure accuracy.

MR analysis

The inverse variance-weighted (IVW) method served as the principal approach for evaluating the causal relationship between dietary habits and the risk of esophageal diseases. Complementary methods, including the MR-Egger and weighted median analyses, were also used. Multivariate MR analysis was conducted to control for potential confounders among various smoking patterns. MR findings were reported as odds ratios (ORs) and 95% confidence intervals (CIs). A two-sided p -value less than 0.05 was deemed indicative of statistical significance. All analyses in this study were performed using the "TwoSampleMR" package (version 0.6.7) in the R software (version 4.4.1), it is supported by the R Foundation for Statistical Computing, which is a non-profit

organization based in Vienna, Austria. For more detailed information, please visit the official R website at <https://www.r-project.org/>.

Gene expression and eQTL, and mQTL dataset

eQTL analysis is the most common approach to evaluate the effects of SNPs on gene expression.²⁰ The latest version of Genotype-Tissue Expression (GTEx) hosts data for 54 tissues obtained from 948 donors, totalling 17,382 samples.²¹ For this project, we obtained fully processed, filtered, and normalized gene expression matrices (in BED format) for each tissue using GTEx version 8 (v8). The gene expression and eQTL data used in this study were retrieved from the GTEx v8 database (<https://gtexportal.org/home/>). The mQTL data were a meta-analysis²² of blood mQTL from the Brisbane Systems Genetics Study ($n=614$) and the Lothian Birth Cohorts of 1921 and 1936 ($n=1366$).

In summary, the SMR methodology leveraged the concept of MR by integrating GWAS, eQTL, and mQTL summary statistics. This approach aimed to assess the pleiotropic associations between gene expression, DNA methylation, and smoke/BE, stemming from a potentially causal variant shared within a locus. The Heterogeneity in Dependent Instruments (HEIDI) test was used to evaluate the homogeneity of these associations. Rejection of the null hypothesis (HEIDI $p < 0.05$) suggests that the observed association could be attributed to two genetically distinct variants with significant LD.

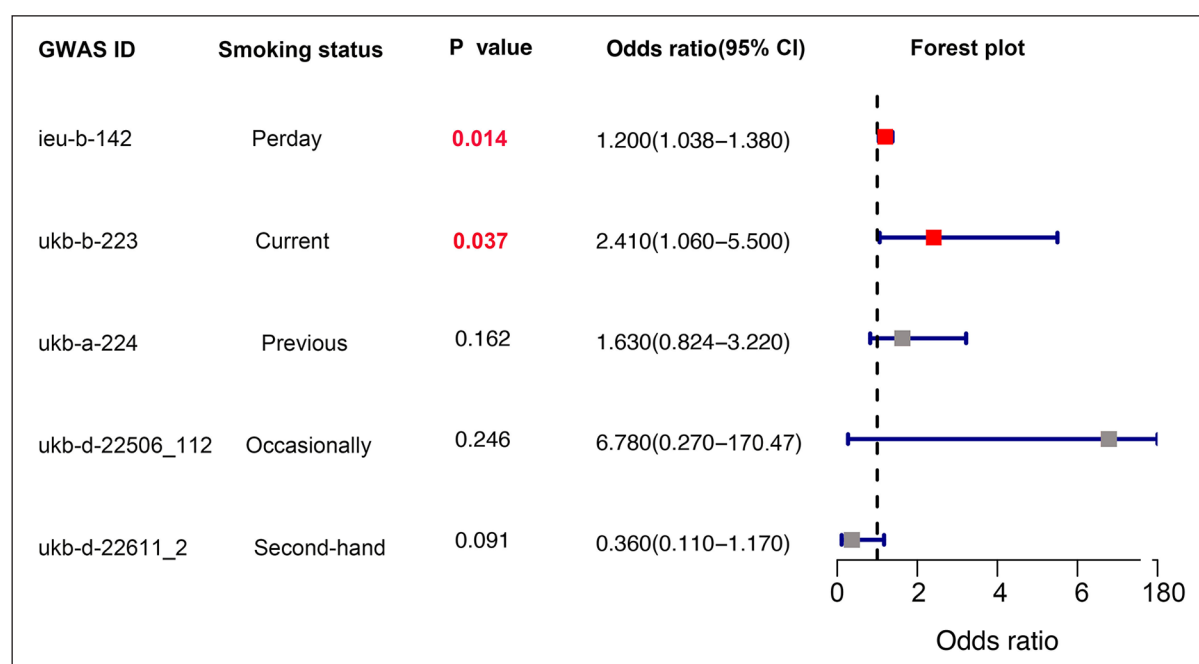


Figure 2. IVW-methods of multivariable MR analysis of smoking status on BE.
BE, Barrett's esophagus; CI, confidence interval; IVW, inverse variance-weighted; MR, Mendelian randomization.

We employed standard configurations within the SMP framework, such as a minor allele frequency exceeding 0.01, while omitting SNPs in extremely strong LD ($r^2 > 0.9$) with the primary QTL. Additionally, SNPs with weak or negligible LD ($r^2 < 0.05$) with the leading QTL were excluded. Multiple comparisons were adjusted using the False Discovery Rate (FDR) method. The study design and data sources for the three-step SMR are given in Figure 1(c). Data cleaning and statistical analysis were performed using R version 4.4.1 (<https://www.rproject.org/>).

Investigating overlapping genetic variations potentially linked to multiple traits is the focus of colocalization analysis. By combining GWAS results with molecular QTL datasets using techniques such as SMR or co-localization, researchers can enhance the identification of possible causal SNPs that operate through specific biological pathways.²³ This approach evaluates whether a common causal variant influences the two traits within the same genomic region. To assess shared genetic effects between traits, the “coloc” R package was employed, setting a threshold of PPH4 greater than 0.5 to determine significant associations.²⁴

Results

Instrumental variables for smoking factors and BE outcomes

Overall, six smoking-related GWAS datasets and three BE-related GWAS datasets were included in our analysis

(Figure 1). Detailed information regarding these data is shown in Figure 1.

Multivariable MR analysis

We lack information on which smoking behaviors, including exposure to passive smoking environments, are linked to the outcome of BE. Therefore, we performed a multivariable MR analysis based on five causal associations using the IVW method. The association between cigarette smoking per day (OR = 1.2, 95% CI: 1.038–1.38, $p = 0.014$) or current tobacco smoking (OR = 2.410, 95% CI: 1.06–5.5, $p = 0.037$) and BE remained significant in the multivariate MR analysis. However, the remaining three smoking statuses were no longer significant ($p > 0.05$; Figure 2).

Bidirectional MR analysis

In the forward MR analysis, 87 SNPs were extracted, with cigarettes smoked per day as exposure and BE as outcome (Table S1). The results showed IVW (OR = 1.15, 95% CI: 1.013–1.303, $p = 0.031$), and weighted median (OR = 1.233, 95% CI: 1.013–1.501, $p = 0.033$; Figure 3 and Table 1). MR results using other validated BE GWAS datasets (ebi-a-GCST90000515) showed IVW (OR = 1.342, 95% CI: 1.092–1.649, $p = 0.005$; Figure 3 and Table S2). Moreover, reverse MR analysis also showed the BE effect on cigarettes smoked per day using the IVW (OR = 1.051, 95% CI: 1.017–1.087, $p = 0.003$) and weighted median (OR = 1.08,

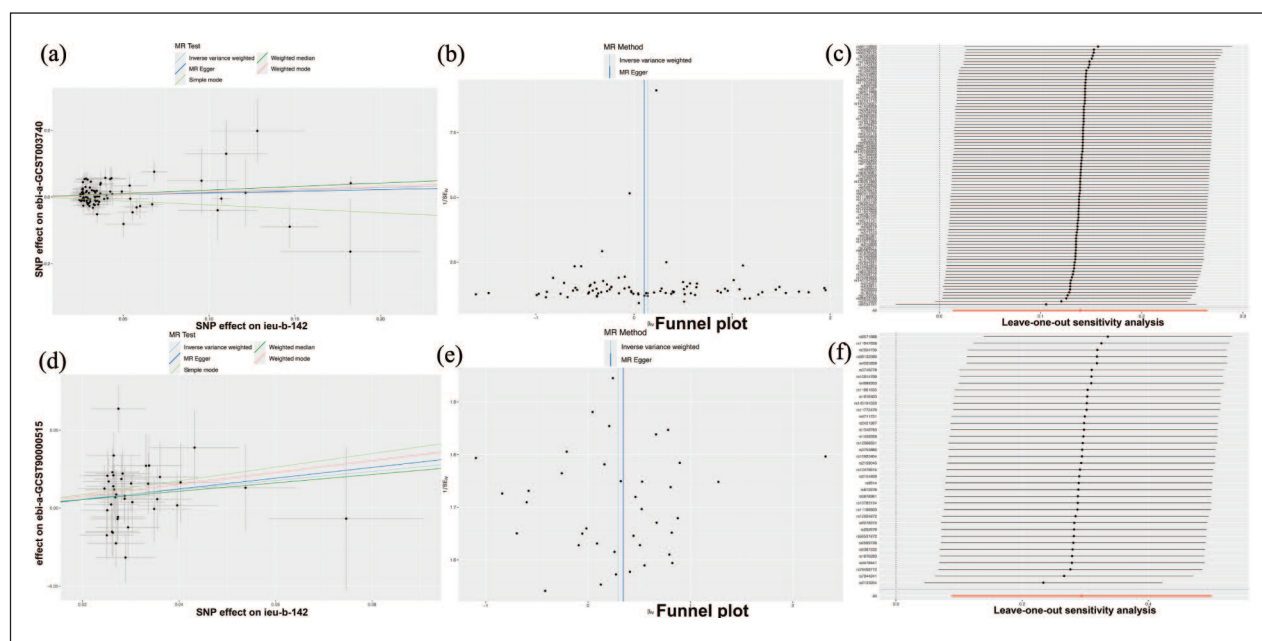


Figure 3. Forward MR results between cigarettes smoked per day and BE. (a) Scatter plot depicting the effects of SNPs on BE (ebi-a-GCST003740); (b) funnel plots comparing IVW and MR-Egger methods for assessing bias; (c) leave-one-out sensitivity analysis; (d) scatter plot depicting the effects of SNPs on BE (ebi-a-GCST90000515); (e) funnel plots comparing IVW and MR-Egger methods for assessing bias; (f) leave-one-out sensitivity analysis. BE, Barrett's esophagus; MR, Mendelian randomization; SNPs, single-nucleotide polymorphisms.

95% CI: 1.032–1.12, $p=0.001$) methods (Figure 4, Table 1 and Table S3).

Integration of GWAS and smoking-related cis-QTL data from whole-blood

We identified 81 genes showing smoking associations from the GWAS datasets (dataset: UK Biobank_22506_111; SMR FDR < 0.05 and HEIDI $p > 0.05$) (Tables S4 and S5). Figure 5(a) shows the eQTL results for the probe ENSG00000184898 tagged with RBM43 on chromosome 2 (Figure 5(a)). Figure 5(b)–(d) shows that the SNP signals associated with RBM43 were significant across the data from smoking GWAS, eQTL, and mQTL studies. The RBM43 expression level was negatively associated with smoking (beta SMR = −0.034). DNA methylation levels by probe cg00935895 showed a positive effect on smoking (beta SMR = 0.464) and a negative effect on RBM43 expression (beta SMR = −0.303). Taken together, our results suggest a putative mechanism wherein a lower DNA methylation level at the enhancer region of RBM43 upregulates the expression of RBM43 and subsequently decreases smoking risk (Figure 5).

Integration of GWAS and BE-related cis-QTL data from whole-blood

After SMR analysis of BE associations from the UK Biobank_K11_BARRET dataset (Table S6), we obtained six intersection genes with the UK Biobank_22506_111: ITIH4, MUSTN1,

TMEM110, HEMK1, RBM43, and RP11-263K19.4 (Figure 6(a)). As a result, only RBM43 (ENSG00000184898) passed the integrated GWAS summary statistics with putative BE-causal cis-eQTLs by colocalization analysis with the threshold of PPH4 > 0.5 (Figure 6(b)). Figure 6(c) shows the SNP and SMR associations across the mQTL, eQTL, and BE GWAS at the RBM43 locus (Figure 6(c)). Figure 6(d)–(f) shows that the SNP signals associated with RBM43 were significant across the data from smoking GWAS, eQTL, and mQTL studies. The RBM43 expression level was negatively associated with BE outcome (beta SMR = −0.212). DNA methylation levels by probe cg00935895 showed a positive effect on BE (beta SMR = 0.267) and a negative effect on RBM43 expression (beta SMR = −0.303). Our results also suggest a putative mechanism wherein a lower DNA methylation level in the enhancer region of RBM43 upregulates the expression of RBM43 and subsequently decreases the BE risk (Figure 6(d)–(f)).

Discussion

In this study, we first conducted MR analysis to explore the potential association between smoking and BE. Subsequently, we integrated the GWAS and cis-QTL data to identify putative genes that showed a pleiotropic/potentially causal association with smoking dependence and BE.

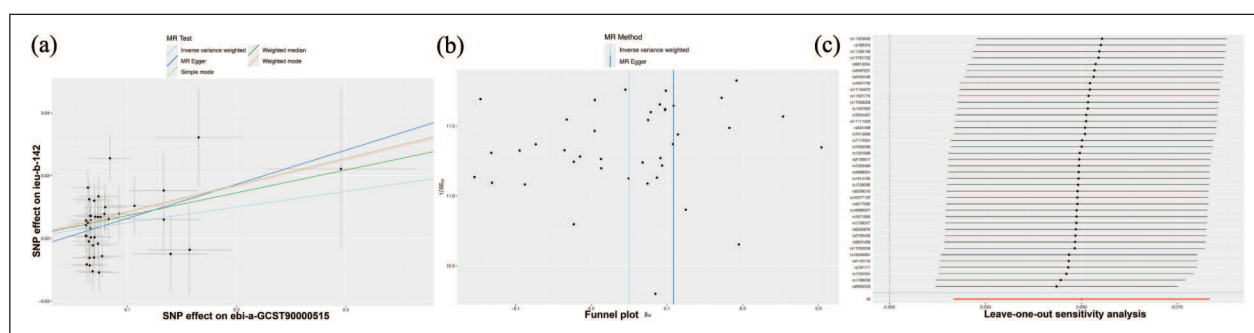
We confirmed that smoking dependence significantly increased the risk of BE and further identified significant genes that may be involved in the pathogenesis of both smoking dependence and BE. Our findings provide important insight

Table 1. The bidirectional MR estimates for the relationship between genetically instrumented smoking dependence and BE.

GWAS ID (outcome)	GWAS ID (exposure)	Number of SNPs	Method	p Value	OR (95% CI)	Cochran Q (p value)	MR-Egger intercept (p value)
ebi-a-GCST003740 (BE)	ieu-b-142 (Smoke per day)	86	MR Egger	0.405	1.107 (0.873–1.404)	0.052	0.715
		86	Weighted median	0.033	1.233 (1.013–1.501)		
		86	IVW	0.031	1.15 (1.013–1.303)	0.059	
		86	Simple mode	0.328	0.79 (0.501–1.246)		
		86	Weighted mode	0.134	1.167 (0.979–1.39)		
ebi-a-GCST90000515 (BE)		35	MR Egger	0.559	1.411 (0.449–4.432)	0.136	0.931
		35	Weighted median	0.056	1.31 (1.004–1.706)		
		35	IVW	0.005	1.342 (1.092–1.649)	0.162	
		35	Simple mode	0.132	1.55 (0.879–2.724)		
		35	Weighted mode	0.197	1.46 (0.857–2.493)		
ieu-b-142 (Smoke per day)	ebi-a-GCST90000515 (BE)	40	MR Egger	0.137	1.11 (0.969–1.283)	0.025	0.404
		40	Weighted median	0.001	1.08 (1.032–1.12)		
		40	IVW	0.003	1.051 (1.017–1.087)	0.026	
		40	Simple mode	0.078	1.09 (0.996–1.191)		
		40	Weighted mode	0.085	1.09 (0.999–1.184)		

BE, Barrett's esophagus; CI, confidence interval; IVW, inverse variance-weighted; MR, Mendelian randomization; OR, odds ratio; SNP, single-nucleotide polymorphism.

Bold fonts indicate statistically significant differences.

**Figure 4.** Reverse MR analysis of BE (ebi-a-GCST90000515) on cigarettes smoked per day (ieu-b-142). (a) Scatter plot displaying the effects of SNPs associated with BE; (b) funnel plots comparing results from the IVW and MR-Egger methods; (c) leave-one-out sensitivity analysis summarizing the causal effects of BE-associated SNPs on smoking.

BE, Barrett's esophagus; IVW, inverse variance-weighted; MR, Mendelian randomization; SNP, single-nucleotide polymorphism.

into the mechanisms underlying the association between smoking dependence and BE. This study also identified potential therapeutic targets for effective prevention and treatment of BE.

BE is the precursor of EAC²⁵ and guidelines suggest screening for BE in populations with multiple risk factors.²⁶ Understanding the potential association between smoking and BE, as revealed by our MR study, may facilitate early intervention strategies, thereby reducing the likelihood of recurrence in populations with dependence on smoking. However, failure to screen high-risk individuals represents a missed opportunity for EAC prevention and early detection.²⁷

Our MR study provides strong evidence supporting a causal link between heavy cigarette dependence and increased

risk of BE. Furthermore, multivariable MR analysis demonstrated a significant association between the number of cigarettes smoked per day and current tobacco smoking status and the risk of BE in comparison with other forms of smoking exposure, such as secondhand smoke. These findings are consistent with those of previous observational studies, which have indicated a potential relationship between smoking and BE.^{5,28} However, unlike observational studies, MR designs reduce the risk of confounding and reverse causation, thereby strengthening causal inference.²⁹ This robust methodology suggests that the observed associations are likely causal rather than merely correlational.³⁰ The consistency of these results with those obtained from multivariable MR analysis reinforces the reliability of our findings and underscores the

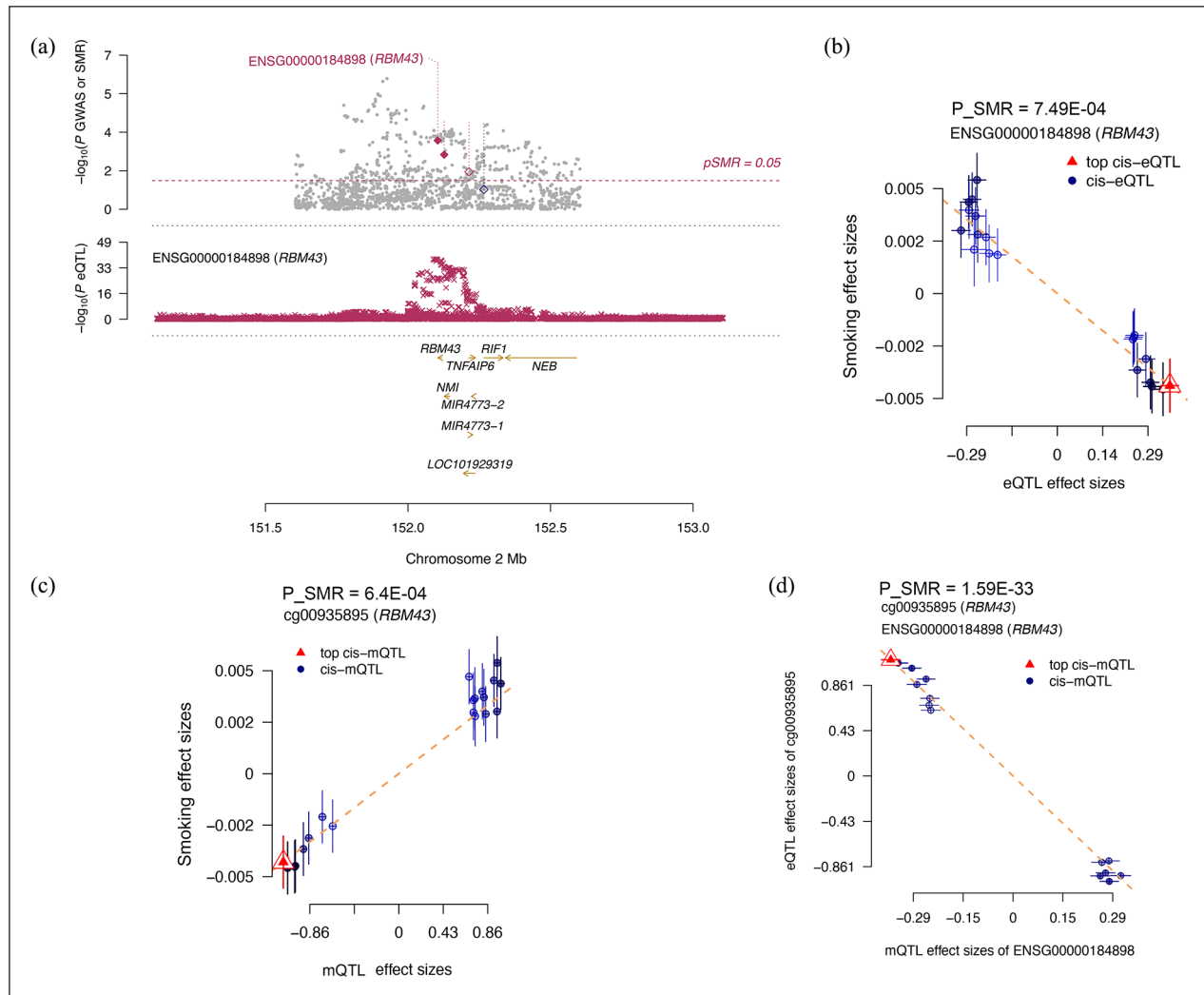


Figure 5. Locus plots showing the significant gene RBM43 and the SMR effect plots. (a) Locus plots display eQTL results for the probe ENSG00000184898 tagging RBM43 within chromosome 2; (b) SMR effect on gene expression and smoking; (c) SMR effect on methylation and smoking; and (d) SMR effect between methylation and gene expression. The x-axis represents cis-QTL effect sizes, while the y-axis represents GWAS/eQTL effect sizes.

eQTL, expression quantitative trait locus; SMR, summary data-based Mendelian randomization.

importance of genetic factors in modulating the risk of BE among individuals with a propensity for heavy smoking dependence. Together, these analyses suggest that genetic variants that influence smoking behavior also play a role in the pathogenesis of BE.

Additionally, our bidirectional MR analysis was designed to assess the possibility of reciprocal causation between BE and cigarette smoking. These results support the hypothesis that cigarette smoking increases the risk of BE, with evidence in the reverse direction. The results were consistent with both previous MR studies³¹ and observational studies³²; notably, our study also revealed a significant association in either direction. This could imply one of several scenarios: (1) both smoking and BE might influence each other over time. In this case, smoking could increase the risk of BE; however, the presence of BE might also influence smoking behavior, perhaps due to symptom exacerbation leading to

changes in smoking habits. (2) There may be common genetic variants that influence both smoking behaviors and the risk of BE independently. This could mean that individuals who are genetically predisposed to smoking might also be predisposed to developing BE, but not necessarily because one directly causes the other; (3) there could be unmeasured confounders that affect both smoking and BE, such as socioeconomic status, dietary habits, or other lifestyle factors that were not accounted for in the analysis; and (4) it is also possible that the bidirectional analysis picked up indirect effects or pleiotropy, where genetic instruments affecting one phenotype also affect another phenotype through unrelated pathways. In our bidirectional MR analysis, the presence of heterogeneity indicates the influence of potential confounders. However, the analysis passed the test for horizontal pleiotropy, further confirming the potential genetic association between smoking and BE.

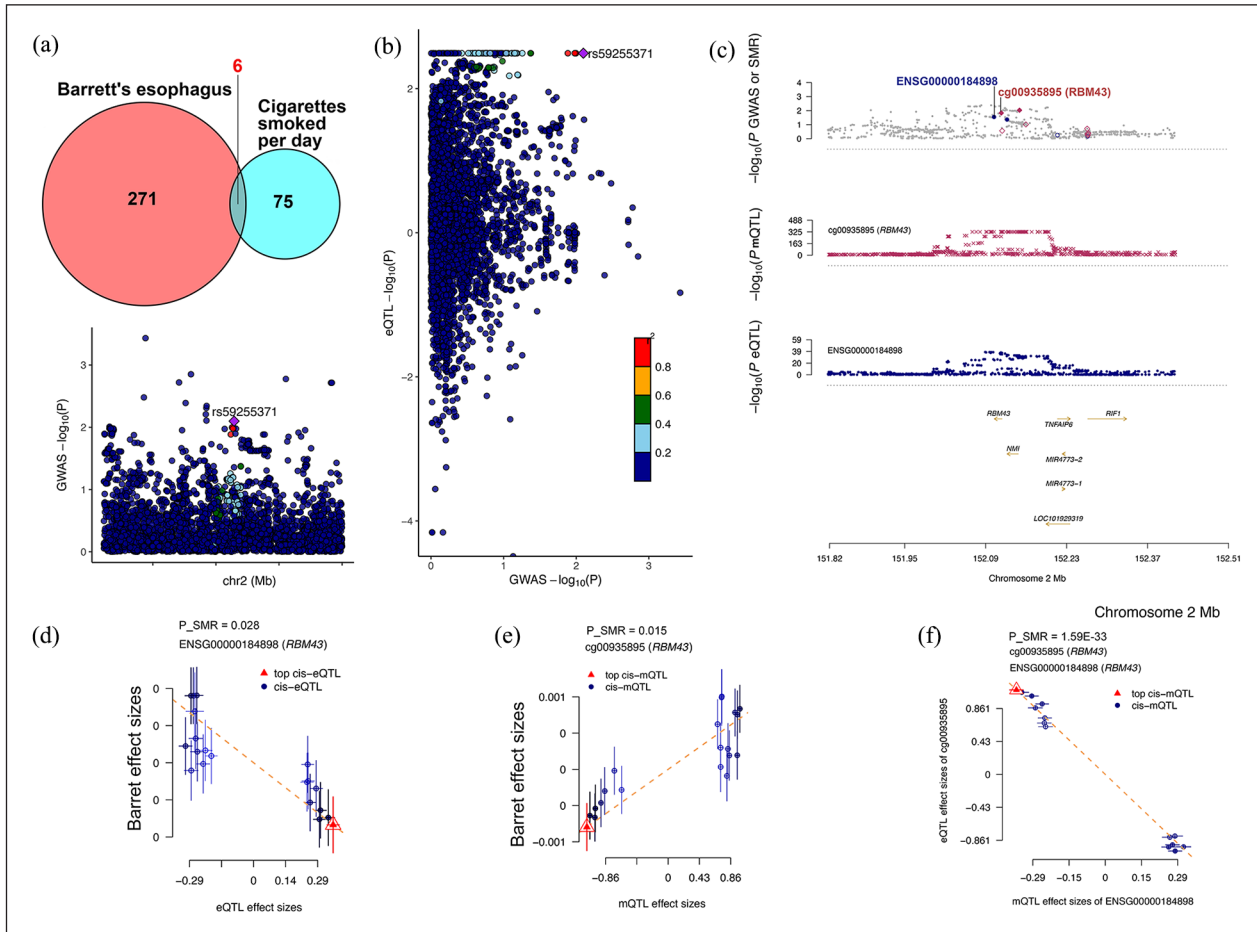


Figure 6. RBM43 gene screening, colocalization analysis, and three-step SMR analysis prioritized putative causal RBM43 gene with BE. (a) Intersection gene screening from SMR results for smoking and BE; (b) locus comparisons between cis-eQTLs and BE GWAS by colocalization analysis (PPH4 > 0.5); (c) locus plots illustrating consistent genetic effects from BE GWAS, cis-mQTL, and cis-eQTL near RBM43 (all minimum $p < 1.0E-05$). (d–f) Three-step SMR indicates significant causal relationships between RBM43 gene expressions and BE onset mediated by methylation (all HEIDI test $p > 0.05$). The r^2 value indicates the LD between the variants and the top SNPs. BE, Barrett's esophagus; eQTL, expression quantitative trait locus; HEIDI, Heterogeneity in Dependent Instruments; mQTL, methylation quantitative trait locus; SMR, summary data-based Mendelian randomization; SNP, single-nucleotide polymorphism.

Our study employed a SMR approach to integrate a comprehensive set of meta-analytic cis-QTL information derived from the GTEx project.^{33,34} Integrative analysis identified the specific gene RBM43 as the top hit, exhibiting a negative association with both smoking dependence and BE. This protein functions as a tumor suppressor by regulating CCNB1 expression, thereby controlling the cell cycle.³⁵ Recent studies have revealed that RBM43 inhibits the expression of mitochondrial and thermogenic genes in a PGC1 α -dependent manner, and the loss of RBM43 protects cells from cytokine-induced mitochondrial dysfunction.³⁶ Therefore, the regulatory mechanisms of RBM43 in smoking dependence and BE require further investigation. This discovery also highlights the potential role of the genetic regulation of gene expression in mediating the effects of smoking dependence on BE risk.

The observed bidirectional association between smoking and BE, where both forward and reverse MR analyses suggest

a promoting effect of one on the other, presents intriguing findings but is not without limitations. One key limitation is the potential for pleiotropy or horizontal pleiotropy, where genetic variants influence multiple traits independently of the BE smoking pathway, which can lead to biased estimates in this MR study. Although the MR-Egger intercept analysis did not reach statistical significance ($p > 0.05$), suggesting no substantial directional pleiotropy, the significant Cochran's Q test in the reverse analysis indicated heterogeneity among the instrumental variables, raising concerns regarding the validity of the genetic instruments used. This heterogeneity may reflect the presence of unaccounted confounders or nonlinear relationships that were not captured by the linear models typically employed in this analysis. Additionally, the bidirectional causality suggested by the analyses might be an artifact of shared genetic predispositions or lifestyle factors that influence both smoking behavior and the risk of developing BE, rather than a true causal relationship.

Additionally, the generalizability of our findings may be limited by the predominant European ancestry of the GTEx cohort, which requires replication in diverse populations. It is important to note that our study predominantly utilized data from European cohorts, limiting the applicability of our findings to diverse populations. Genetic architecture can vary significantly across ethnicities, potentially leading to distinct causal pathways and effect sizes. Although our results provide robust insights into the genetic mechanisms underlying smoking and BE in European populations, they should be interpreted with caution when applied to other ethnic groups. Future studies should aim to replicate our findings in diverse populations, including individuals of African, Asian, and Latin American descent, to ensure broader applicability and enhance the generalizability of our conclusions. Incorporating multiethnic datasets will help identify population-specific genetic effects and improve the precision of the analyses. Therefore, future studies should focus on functional validation of the identified genetic variants and their interactions with environmental exposure to better understand the complex etiology of BE.

The discovery of RBM43 opens new avenues for translational research aimed at developing novel therapeutic strategies for BE. Understanding how genetic variations influence susceptibility to BE can guide the development of personalized medical approaches that target underlying molecular mechanisms.³⁷ Such efforts could lead to the identification of new biomarkers for the early detection and more effective treatment of BE, potentially improving patient outcomes and reducing the burden of the disease on the healthcare system.

Conclusions

Our study provides compelling evidence of a causal relationship between heavy cigarette dependence and the risk of developing BE. Integration of cis-QTL data from GTEx has identified RBM43 as a key gene in this relationship, offering new targets for future research and potential therapeutic interventions. Public health strategies aimed at reducing smoking rates may be pivotal in lowering the incidence of BE and its subsequent complications. Further research is needed to confirm these findings and to explore the underlying mechanisms responsible for the observed associations.

Acknowledgements

We express our gratitude to OpenGWAS (<https://gwas.mrcieu.ac.uk>) and the UK Biobank (<https://www.ukbiobank.ac.uk/>) for providing the publicly available data for our study.

Author contributions

All authors have made substantial contributions to the development and completion of this study. Z.A. was involved in the conception and design of the study; the acquisition, analysis, and interpretation of data; and drafted the article. M.Z. contributed to data analysis and revised the article. X.W. contributed to study design, data

analysis, and article drafting. All authors provided final approval and agreed to be accountable for all aspects of the study.

Availability of data and materials

All data generated or analyzed during this study were included in the datasets of the OpenGWAS database (<https://gwas.mrcieu.ac.uk>), the UK Biobank (<https://www.ukbiobank.ac.uk/>), and the study's Supplemental information file.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethics approval

For the current research, it was not necessary to seek ethical approval since the initial investigation had already secured appropriate permission.

Informed consent

Informed consent was not required for this study because the original research participants provided written informed consent.

Consent for publication

Not applicable.

Trial registration

Not applicable.

ORCID iD

Xianhua Wang  <https://orcid.org/0009-0006-7205-7876>

Supplemental material

Supplemental material for this article is available online.

References

1. Kota L and Tyberg A. Risk of Barrett's esophagus progression to esophageal adenocarcinoma. *Clin Endosc* 2024; 57(6): 763–764.
2. Hvid-Jensen F, Pedersen L, Drewes AM, et al. Incidence of adenocarcinoma among patients with Barrett's esophagus. *N Engl J Med* 2011; 365(15): 1375–1383.
3. Kodama S, Watanabe K, Shimodaira Y, et al. Development of a prediction score for Barrett's esophagus in Japanese health checkup settings. *Esophagus* 2024; 21(4): 552–562.
4. Honing J and Fitzgerald RC. Categorizing risks within Barrett's esophagus to guide surveillance and interception; suggesting a new framework. *Cancer Prev Res (Phila)* 2023; 16(6): 313–320.
5. Levy J, Buhl K, Fernandez C, et al. Does smoking cannabis increase the risk of Barrett's esophagus? *Cureus* 2020; 12(2): e6913.

6. Filiberti R, Fontana V, De Ceglie A, et al. Smoking as an independent determinant of Barrett's esophagus and, to a lesser degree, of reflux esophagitis. *Cancer Causes Control* 2015; 26(3): 419–429.
7. Lin YL, Yao T, Wang YW, et al. Potential drug targets for gastroesophageal reflux disease and Barrett's esophagus identified through Mendelian randomization analysis. *J Hum Genet* 2024; 69(6): 245–253.
8. Ali MW, Chen J, Yan L, et al. A risk variant for Barrett's esophagus and esophageal adenocarcinoma at chr8p23.1 affects enhancer activity and implicates multiple gene targets. *Hum Mol Genet* 2022; 31(23): 3975–3986.
9. Schroder J, Schuller V, May A, et al. Identification of loci of functional relevance to Barrett's esophagus and esophageal adenocarcinoma: cross-referencing of expression quantitative trait loci data from disease-relevant tissues with genetic association data. *PLoS One* 2019; 14(12): e0227072.
10. Gharahkhani P, Tung J, Hinds D, et al. Chronic gastroesophageal reflux disease shares genetic background with esophageal adenocarcinoma and Barrett's esophagus. *Hum Mol Genet* 2016; 25(4): 828–835.
11. Saha B, Verma A, Dierkhising R, et al. Barrett's esophagus risk factors in patients without gastroesophageal reflux: a large population-based study. *Clin Gastroenterol Hepatol* 2024; 22(10): 2150.e2–2152.e2.
12. Huang Y, Cai Y, Chen Y, et al. Cholelithiasis and cholecystectomy increase the risk of gastroesophageal reflux disease and Barrett's esophagus. *Front Med (Lausanne)* 2024; 11: 1420462.
13. Li C, Shu P, Shi T, et al. Predicting the potential deterioration of Barrett's esophagus based on gut microbiota: a Mendelian randomization analysis. *Mamm Genome* 2024; 35(3): 399–413.
14. Cruz JD, Paculdo D, Ganesan D, et al. Clinical variation in surveillance and management of Barrett's esophagus: a cross-sectional study of gastroenterologists and gastrointestinal surgeons. *Medicine (Baltimore)* 2022; 101(51): e32187.
15. Verbanck M, Chen CY, Neale B, et al. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet* 2018; 50(5): 693–698.
16. Davies NM, Holmes MV and Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018; 362: k601.
17. Kuo CL, Pilling LC, Liu Z, et al. Genetic associations for two biological age measures point to distinct aging phenotypes. *Aging Cell* 2021; 20(6): e13376.
18. Ramdas S, Judd J, Graham SE, et al. A multi-layer functional genomic analysis to understand noncoding genetic variation in lipids. *Am J Hum Genet* 2022; 109(8): 1366–1387.
19. Palmer TM, Lawlor DA, Harbord RM, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res* 2012; 21(3): 223–242.
20. Vosa U, Claringbould A, Westra HJ, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet* 2021; 53(9): 1300–1310.
21. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* 2020; 369(6509): 1318–1330.
22. Wu Y, Zeng J, Zhang F, et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat Commun* 2018; 9(1): 918.
23. Xu S, Li X, Zhang S, et al. Oxidative stress gene expression, DNA methylation, and gut microbiota interaction trigger Crohn's disease: a multi-omics Mendelian randomization study. *BMC Med* 2023; 21(1): 179.
24. Huang Y, Shan Y, Zhang W, et al. Deciphering genetic causes for sex differences in human health through drug metabolism and transporter genes. *Nat Commun* 2023; 14(1): 175.
25. Shaheen NJ, Othman MO, Taunk J, et al. Use of a two-gene methylated DNA biomarker assay and non-endoscopic balloon for detection of Barrett's esophagus among high-risk individuals in a screening population. *Am J Gastroenterol* 2024. (Online ahead of print).
26. Hoefnagel SJM, Mostafavi N, Timmer MR, et al. A genomic biomarker-based model for cancer risk stratification of non-dysplastic Barrett's esophagus patients after extended follow up; results from Dutch surveillance cohorts. *PLoS One* 2020; 15(4): e0231419.
27. Crowe BR, Krigel A, Li T, et al. Veterans with multiple risk factors for Barrett's esophagus are infrequently evaluated with upper endoscopy. *Dis Esophagus* 2023; 36(9): doad007.
28. Westra WM, Lutzke LS, Mostafavi NS, et al. Smokeless tobacco and cigar and/or pipe are risk factors for Barrett esophagus in male patients with gastroesophageal reflux disease. *Mayo Clin Proc* 2018; 93(9): 1282–1289.
29. Birney E. Mendelian randomization. *Cold Spring Harb Perspect Med* 2022; 12(4): a041302.
30. Elhage KG, Kranyak A, Jin JQ, et al. Mendelian randomization studies in atopic dermatitis: a systematic review. *J Invest Dermatol* 2024; 144(5): 1022–1037.
31. Liu YX, Bin CL, Zhang L, et al. Socioeconomic traits and the risk of Barrett's esophagus and gastroesophageal reflux disease: a Mendelian randomization study. *World J Gastrointest Oncol* 2024; 16(6): 2631–2645.
32. Byrne CJ, Brennan P, Carberry J, et al. Long-term risk factors for developing Barrett's oesophagus in patients with gastroesophageal reflux disease: a longitudinal cohort study. *BMJ Open Gastroenterol* 2024; 11(1): e001307.
33. Yan KK, Zhao H, Wu JT, et al. An enhanced machine learning tool for cis-eQTL mapping with regularization and confounder adjustments. *Genet Epidemiol* 2020; 44(8): 798–810.
34. Sobin L, Barcus M, Branton PA, et al. Histologic and quality assessment of Genotype-Tissue Expression (GTEx) research samples: a large postmortem tissue collection. *Arch Pathol Lab Med* 2024. (Online ahead of print).
35. Liu Y, Feng H, Zhao Q, et al. RNA binding motif protein 43 (RBM43) suppresses hepatocellular carcinoma metastasis by regulating Slug mRNA stability. *Genes Dis* 2024; 11(6): 101192.
36. Dumesic PA, Wilensky SE, Bose S, et al. RBM43 links adipose inflammation and energy expenditure through translational regulation of PGC1alpha. *bioRxiv*. 2023. DOI: 10.1101/2023.01.06.522985
37. Lartigue S, Yalaoui M, Belliard J, et al. Consistent variations in personality traits and their potential for genetic improvement in biocontrol agents: *Trichogramma evanescens* as a case study. *Evol Appl* 2022; 15(10): 1565–1579.