Original Article

A multi-level investigation of the genetic relationship between gastroesophageal reflux disease and lung cancer

Dongsheng Wu[1#], Jian Zhou[1#], Lujia Song[2], Quan Zheng[1], Tengyong Wang[1], Zhizhen Ren[1], Yuchen Huang[3], Shuqiao Liu[3], Lunxu Liu[1]

[1]Department of Thoracic Surgery and Institute of Thoracic Oncology, West China Hospital, Sichuan University, Chengdu, China; [2]Department of Pulmonary and Critical Care Medicine, Frontiers Science Center for Disease-related Molecular Network, West China Hospital, and Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu, China; [3]West China School of Medicine, Sichuan University, Chengdu, China

Contributions: (I) Conception and design: D Wu, L Liu; (II) Administrative support: L Liu; (III) Provision of study materials or patients: J Zhou, L Liu; (IV) Collection and assembly of data: D Wu, J Zhou, L Song; (V) Data analysis and interpretation: D Wu, T Wang, Z Ren, Y Huang, S Liu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Lunxu Liu, MD, PhD, FRCS. Department of Thoracic Surgery and Institute of Thoracic Oncology, West China Hospital, Sichuan University, No. 37 Guoxue Alley, Chengdu 610041, China. Email: lunxu_liu@aliyun.com.

Background: Observational studies have revealed a potential association between gastroesophageal reflux disease (GERD) and lung cancer (LC), but the genetic role in their comorbidity have not been fully elucidated. This study aimed to comprehensively dissect the genetic link underlying GERD and LC.

Methods: Using large-scale genome-wide association study (GWAS) data, we investigated shared genetic architecture between GERD and LC. Our analyses encompassed genetic correlation, cross-trait meta-analysis, transcriptome-wide association studies (TWASs), and the evaluation of the causality though a bidirectional Mendelian randomization (MR) analysis with sufficient sensitivities.

Results: We identified a significant genome-wide genetic correlation between GERD and overall LC (r_g=0.33, P=1.58×10^{-14}), as well as across other subtype-specific LC (r_g ranging from 0.19 to 0.39). After separating the whole genome into approximately 2,353 independent regions, 5 specific regions demonstrated significant local genetic correlation, with most significant region located at 9q33.3. Cross-trait meta-analysis revealed 22 pleiotropic loci between GERD and LC, including 3 novel loci (rs537160, rs10156445, and rs17391694). TWASs discovered a total of 49 genes shared in multiple tissues, such as lung tissues, esophagus muscularis, esophagus mucosa, and esophagus gastroesophageal junction. MR analysis suggested a significantly causal relationship between GERD and overall LC [odds ratio (OR) =1.34, 95% confidence interval (CI): 1.19–1.51], as well as other subtype-specific LC (OR ranging from 1.25 to 1.76). No evidence supports a significant causal effect of LC on GERD.

Conclusions: Our findings suggest intrinsic genetic correlation underlying GERD and LC, which provides valuable insights for screening and management of LC in individuals with GERD.

Keywords: Gastroesophageal reflux disease (GERD); lung cancer (LC); genome-wide cross-trait analysis; Mendelian randomization (MR)

Submitted Apr 18, 2024. Accepted for publication Aug 12, 2024. Published online Sep 24, 2024.
doi: 10.21037/tlcr-24-345
View this article at: https://dx.doi.org/10.21037/tlcr-24-345

© AME Publishing Company.

Introduction

Gastroesophageal reflux disease (GERD), characterized by troublesome symptoms and complications caused by the reflux of duodenogastric contents, is a common and chronic condition affecting approximately 2.5% to 33.1% of the global population (1,2). Recurrent micro-aspiration from the refluxed contents is associated with higher risk of multiple lung diseases, including pneumonia, asthma, bronchiolitis obliterans syndrome, idiopathic pulmonary fibrosis, and chronic obstructive pulmonary disease (3-6). Additionally, underlying links between GERD and non-esophageal cancer have been also recognized (7). Lung cancer (LC), as one of the most prevalent malignancies, remains the leading cause in both incidence rate and mortality worldwide (8). Recent epidemiological studies have observed a significant phenotypic association between GERD and LC. Leveraging data from the National Health Insurance Research Database of Taiwan (15,412 cases and 60,957 controls), Hsu *et al.* found that patients with GERD were associated with significantly elevated risk of LC in comparison to those without (9). More recently, a multinational cohort study enrolled 812,617 patients with GERD to investigate the impact of anti-reflux surgery on the risk of distinct histological types of LC (10). Similarly, this study found that anti-reflux surgery significantly decreased the risk of small-cell lung cancer (SCLC) and lung squamous cell carcinoma (LUSC), and showed a protective trend for lung adenocarcinoma (LUAD). Despite this, phenotypic correlations revealed in conventional epidemiological studies were susceptible to potential biases, confounding factors, and reverse causality due to the observational nature (11).

Utilizing genetic data for phenotypic correlation analysis offers a distinct advantage over observational studies, as it can effectively circumvent the issue of reverse causality and can also minimize the potential confounding with meticulous design. With the increasing sample size of genome-wide association studies (GWAS), previous studies have identified a substantial number of genetic variants [single nucleotide polymorphisms (SNPs)] associated with GERD (88 SNPs) and LC (56 SNPs) (12,13). Furthermore, utilizing the design of the twin study, heritability of GERD and LC has been estimated as 30–31% (14,15) and 18–26% (16,17), respectively. This underscores a significant genetic component in disease susceptibility.

In this context, several Mendelian randomization (MR) studies have been conducted using genetic variants as instrumental variables (IVs), and consistently identified a causal association between GERD and LC, with odds ratio (OR) ranging from 1.25 to 1.37 (18-20). Nonetheless, multiple significant gaps in previous investigations remain to be filled. Firstly, prior MR studies used GWAS data with relatively small sample sizes (18,19), particularly for SCLC, which restricted the statistical power. Secondly, the insufficient sensitivity analyses did not guarantee the core model assumptions, thereby impeding the robustness of results (21). Finally, the adoption of limited confounders, such as smoking status and obesity, may not comprehensively account for potential pleiotropy in complex traits (20,22).

Therefore, a novel statistical genetic tool named genome-wide cross-trait analysis was utilized to dissect shared genetic components in complex traits, using summary data from the large-scale GWAS studies (11,23). Specifically, we measured the genetic correlation, identified the shared loci, and finally inferred a putative causal association through the bidirectional two-sample MR analysis. *Figure 1* illustrates the overall study design. We present this article in accordance with the STREGA reporting checklist (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-345/rc) (24).

Highlight box

Key findings

- This study represents the first comprehensive investigation into the shared genetic architecture between gastroesophageal reflux disease (GERD) and lung cancer (LC), providing valuable insights into this complex genetic interplay.

What is known and what is new?

- Observational studies have revealed a potential association between GERD and LC, but the genetic role in their comorbidity have not been fully elucidated.
- This study provides valuable evidence of genetic correlation, identifying pleiotropic loci, and suggesting a potential causal association between GERD and LC.

What is the implication, and what should change now?

- This study conveys a crucial public health message: managing individuals with GERD may potentially contribute to reducing the long-term burden of malignant diseases.

Methods

GWAS summary datasets

In the study, summary data from the hitherto largest GWAS of GERD and LC were leveraged for genetic analyses, both

© AME Publishing Company.

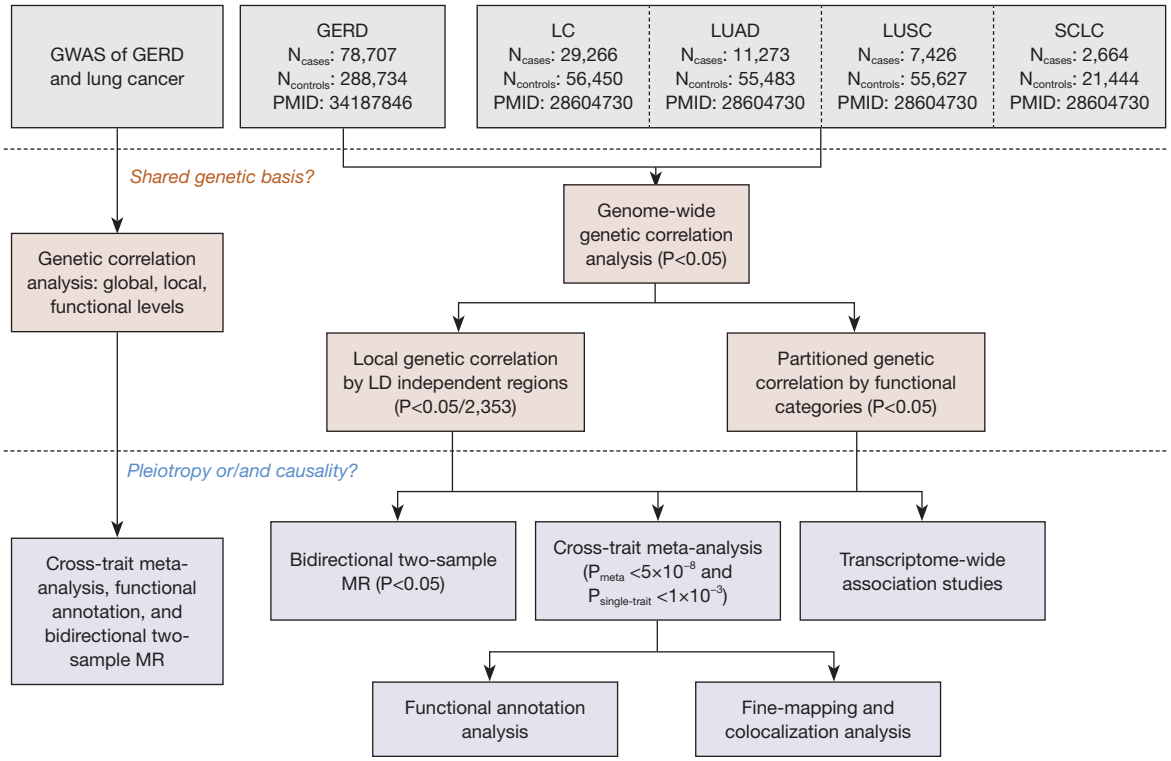

Figure 1 Overall study design of genome-wide cross-trait analysis. A global genetic correlation analysis between GERD and LC was performed. The global genetic correlation was further studied at LD independent regions and by functional categories. Cross-trait meta-analysis was used to identify pleiotropic loci, and a bidirectional two-sample Mendelian randomization analysis was applied to investigate potential causal association. GWAS, genome-wide association study; GERD, gastroesophageal reflux disease; LC, lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer; LD, linkage disequilibrium; MR, Mendelian randomization.

exclusively focused on individuals of European ancestry. The detailed information for the GWAS data of both traits is shown in Table S1.

The largest GWAS study on GERD derived from the meta-analyzing data by Ong *et al.* in 2022 (12), which combined up to 367,441 (78,707 cases) European individuals from the UK Biobank (UKBB) study (35,4285 individuals) and Queensland Sun and Health Study (QSKIN) study (13,156 individuals). GERD was defined based on a combination of self-reported GERD symptoms such as heartburn, the use of GERD medication, and hospital records [The International Classification of Diseases, Tenth Revision (ICD-10)]. The Haplotype Reference Consortium (HRC) reference panel was used to impute the genotype data. To combine the GWAS data from the UKB and QSKIN cohorts, a fixed-effect inverse variance-weighted (IVW) meta-analysis was performed.

For overall LC and subtype-specific LC, the largest GWAS data were a meta-analysis of data from McKay *et al.* in 2017 (25). The GWAS summary data from the International Lung Cancer Consortium (ILCCO) combined a total of 29,266 LC cases and 56,450 controls, which included 11,273 LUAD, 7,426 LUSC, and 2,664 SCLC. Imputation was performed on variants based on the 1000 Genomes Project (1KGP) Phase 3 panel. The fixed-effect IVW meta-analysis was carried out to combine the GWAS data. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Statistical analyses

Genome-wide genetic correlation analysis

We performed linkage disequilibrium (LD) score regression (LDSC) analysis to evaluate the genome-wide genetic correlations between two traits, utilizing GWAS summary statistics in the calculation (26). LDSC estimates genetic correlation (r_g) on a scale from −1 to +1. It leverages the fact that when estimating the effect size of a specific variant,

© AME Publishing Company.

the combined effects of all variants that are in LD with that variant were considered. Thus, the idea of substituting the χ^2 statistics with the product of z-scores from two traits and the LD scores allows us to calculate genetic correlations between traits: $E\left[z_{1j}z_{2j}\middle|l_j\right]=\frac{\sqrt{N_1N_2}\rho_g}{M}l_j+\frac{\rho N_s}{\sqrt{N_1N_2}}$, where N_i represents the sample size of each trait, M represents number of SNPs, ρ_g represents the genetic covariance, l_j represents the LD scores, ρ represents the phenotypic link within overlapping samples, N_s represents the overlapping sample size. By performing a regression of the product of z-scores from two GWASs based on the reference LD scores from 1KGP European ancestry with SNPs mapped in the Hapmap3 reference panel, the genetic covariance between two traits can be estimated. Then, the genetic correlation can be calculated as $r_g=\frac{\rho_g}{\sqrt{h_1^2h_2^2}}$, where h_i^2 represents the heritability for each trait. Given the potential overlap in population between the GWAS data of GERD and LC, we also conducted LDSC with a constrained intercept, which is more robust in handling sample overlap (26). For multiple testing, the false discovery rate (Benjamini-Hochberg correction) was employed.

Local genetic correlation analysis

The global genetic correlation offers an assessment of the collective impact of genome-wide variants. However, it is conceivable that, despite exhibiting minimal global genetic correlation, certain regions of the genome may still have an impact on both traits. Thus, we computed pairwise local genetic correlation using SUPERGNOVA (27). This algorithm separates the entire genome into approximately 2,353 independent regions, with each averaging about 1.6 centimorgans in length. It then measures the genetic correlation specific to each of these genomic regions. To account for multiple testing, Bonferroni correction (P<0.05/2,353) was applied.

Partitioned LDSC analysis

Using partitioned LDSC (28), we investigated the genetic correlation between GERD and LC in multiple functional categories. This study included 14 common functional categories, including coding region, conserved region, DNase I digital genomic foot-printing region (DGF), DNase I hypersensitive sites (DHS), fetal DHS, intron, promotor, super enhancer, transcription factor-binding site (TFBS), transcribed region, and histone marks H3K27ac, H3K4me1, H3K4me3, and H3K9ac (28,29). For SNPs classified within each specific category, recalculated LD scores were utilized to estimate the genetic correlation between GERD and LC within that functional category.

Cross-trait meta-analysis

A significant genetic correlation suggests the presence of either horizontal pleiotropy (pleiotropy) or vertical pleiotropy (causality). To further investigate the pleiotropic SNPs associated with both traits (GERD and LC), we performed a cross-phenotypic association (CPASSOC) analysis (30). Utilizing summary data from single SNP-trait associations in GWAS, CPASSOC provides two estimates, S_{Hom} and S_{Het}. Representing the maximum of the weighted sum of trait-specific genetic effects, S_{Hom} employs a fixed-effect meta-analysis approach, which was more powerful when genetic effect sizes cross traits were homogenous. S_{Het}, as an extension of S_{Hom}, assumes the presence of heterogeneity and computes corresponding P value via a sample size-weighted meta-analysis of GWAS summary data. For this analysis, we adopted the S_{Het} method to correct for potential heterogeneity and ensure more robust results.

After CPASSOC analysis, independent loci were obtained using software PLINK (v1.9) with parameters: --clump-p1 5E-8 --clump-p2 1E-5 --clump-r2 0.2 --clump-kb 500 (31). SNPs with the lowest P value within each independent locus were defined as index SNPs. Significant pleiotropic SNPs were defined as having $P_{CPASSOC}<5\times10^{-8}$ and $P_{single\text{-}trait}<1\times10^{-3}$ in both traits. These SNPs were further classified into four groups: (I) "known" shared SNPs, referring to SNPs that reach genome-wide significance in both traits ($P_{GERD}<5\times10^{-8}$ and $P_{LC}<5\times10^{-8}$); (II) "single-trait-driven" shared SNPs, referring to SNPs reaching genome-wide significance in one of the two traits, either $P_{GERD}<5\times10^{-8}$ or $P_{LC}<5\times10^{-8}$; (III) "LD-Tagged" shared SNPs, referring to SNPs not reaching genome-wide significance in both traits ($P_{GERD}>5\times10^{-8}$ and $P_{LC}>5\times10^{-8}$), but showing LD ($r^2\geq0.2$) with index SNPs previously identified by single-trait GWAS; and (IV) novel shared SNPs, referring to significant pleiotropic SNPs that did not reach genome-wide significance in both traits ($5\times10^{-8}< P_{single\text{-}trait}<1\times10^{-3}$) and were not in LD with previously identified SNPs in single-trait GWAS ($r^2<0.2$) (32). To gain further insights into the biological implications in the shared SNPs, the linear closest genes of pleiotropic loci were annotated using the Ensemble Variant Effect Predictor (VEP) (33).

Fine-mapping credible set analysis

Index SNPs may not always be causal variants due to

© AME Publishing Company.

the complex LD patterns across SNPs. To obtain a credible set of causal variants that have a 99% likelihood of encompassing causal variants for each shared loci, we employed the Bayesian fine-mapping algorithm—FM-summary (34). For each shared locus, variants located within 500 kb of the index SNP were extracted (35). The FM-summary prioritizes the primary signal and applies a flat prior along with a steepest descent approximation (36).

Colocalization analysis

To determine whether the association signals for GERD and LC co-occurred at identified shared loci, we performed the colocalization analysis using the R package Coloc (37). Coloc employs the Bayesian algorithm to obtain five posterior probabilities for five different hypotheses: (I) H0, no causal variant; (II) H1 or H2, causal variant only for one trait; (III) H3, two distinct variants associated with both traits; and (IV) H4, shared variant correlated with both traits. The posterior probability for H4 (PPH4) was calculated using summary data for variants near loci shared between GERD and LC that were within 500 kb of the index SNP. If PPH4 exceeded 0.5, a locus was labeled as a co-localized genetic variant.

Transcriptome-wide association studies (TWASs)

Many genetic variants have an effect on complex phenotypes by modulating gene expression. Therefore, determining overlapping genes underlying GERD and LC may shed light on the underlying causal mechanisms. Utilizing FUSION (38), the TWAS was performed to identify associations between GERD and LC regarding gene expression in multiple tissues. This involved integrating expression weights obtained from 49 tissues sourced from GTEx (version 8) with GWAS summary data (39). To obtain an independent set of gene-tissue pairs, a total of 49 TWASs for each trait were systematically conducted, focusing on one tissue-trait pairing at a time. Subsequently, by intersecting across traits, shared gene-tissue pairs were identified. The Benjamini-Hochberg correction was used to correct TWAS P values, and a false discovery rate <0.05 was deemed significant.

Bidirectional MR analysis

Next, we investigated the putative causal association between GERD and LC through the bidirectional two-sample MR analysis. For GERD, genome-wide significant SNPs ($P<5\times10^{-8}$) were selected and clumped for independent IVs (r^2=0.01 and window size =10 Mb). For LC, SNPs with P value $<5\times10^{-8}$ were obtained and clumped using parameters: r^2=0.01 and window size =10 Mb. *F*-statistic was calculated to assess strength of selected IVs, where a value less than 10 indicates a weak instrument (40). Additionally, the statistical power of MR was evaluated using an online calculator (https://shiny.cnsgenomics.com/mRnd/) (41).

We implemented the IVW method as the principal approach, assumes all IVs to be valid and offers the highest statistical power (42). Additionally, we performed several complementary sensitivity analyses to evaluate the robustness: (I) MR-Egger regression, identifying and mitigating bias resulting from directional pleiotropy (43); (II) weighted median, offering a consistent estimate of causality even with more than 50% invalid IVs (44); (III) Causal Analysis Using Summary Effect estimates (CAUSE) and MR-Pleiotropy Residual Sum and Outlier (MR-PRESSO), evaluating and adjusting for the potential correlated and uncorrelated horizontal pleiotropy (45,46); (IV) removing pleiotropic IVs associated with potential confounding factors based on the Phenoscanner (47); (V) removing palindromic IVs with strand ambiguity; and (VI) leave-one-out analyses, evaluating the potential impact of each SNP on the IVW estimate. We further utilized multivariable MR (MVMR) (48) to adjust for influence of significant confounding factors, including body mass index (BMI) (49), smoking status (50), alcohol consumption (50), physical activity (51), and sleep duration (52). These confounders were integrated individually as well as collectively with GERD to ensure a comprehensive analysis. Finally, a reverse-direction MR analysis was carried out to determine if genetic predisposition to LC has a causal impact on GERD.

All MR analyses were carried out utilizing the following R packages: "TwoSampleMR" (v0.5.6), "MRPRESSO" (v1.0), "CAUSE" (v1.2.0), and "MVMR" (v0.3), in R software (v4.2.3).

Results

Global genetic correlation

We observed a strongly significant global genetic correlation between GERD and overall LC (r_g=0.33, $P=1.58\times10^{-14}$) after adjusting for multiple testing (*Table 1*). The genetic correlation continued to be significant in subtype-specific LC (LUAD: r_g=0.19, $P=6.64\times10^{-6}$; LUSC: r_g=0.39, $P=2.22\times10^{-12}$; SCLC: r_g=0.39, $P=5.27\times10^{-12}$). Given the potential sample overlap in GWAS data, the intercepts of genetic covariance were constrained to zero,

© AME Publishing Company.

Table 1 Genome-wide genetic correlations between GERD and LC using constrained and unconstrained LDSC

Trait 1	Trait 2	Unconstrained LDSC			Constrained LDSC		
		r_g	r_g_se	P value	r_g	r_g_se	P value
GERD	Overall LC	0.33	0.04	1.58×10^{-14}	0.36	0.03	3.28×10^{-37}
GERD	LUAD	0.19	0.04	6.64×10^{-6}	0.20	0.03	5.27×10^{-12}
GERD	LUSC	0.39	0.05	2.22×10^{-12}	0.41	0.03	1.53×10^{-34}
GERD	SCLC	0.39	0.06	5.27×10^{-12}	0.39	0.03	7.31×10^{-25}

GERD, gastroesophageal reflux disease; LC, lung cancer; LDSC, linkage disequilibrium score regression; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer; r_g, genetic correlation; se, standard error.

which exhibited increased power, while also marginally reducing standard errors. Using constrained LDSC, the genetic correlation remained significant for LC (r_g=0.36, P=3.28×10^{-37}) as well as subtype-specific LC (LUAD: r_g=0.20, P=5.27×10^{-12}; LUSC: r_g =0.41, P=1.53×10^{-34}; SCLC: r_g=0.39, P=7.31×10^{-25}).

Local genetic correlation

After separating the genome into multiple LD-independent regions, 5 local regions with significant genetic correlation were detected, including 1 region (Chr2: 103,264,434–104,481,488) shared by overall LC and LUAD, and 1 region (Chr9: 126,927,204–128,926,989) shared by overall LC and LUSC (*Figure 2*). The most significant region was located at 9q33.3 (Chr9: 126,927,204–128,926,989, P=6.76×10^{-10}), which harbors *PBX3*, a factor interacts with the promoter of tumor suppressor *p53* associated with LC tumorigenesis (53,54).

Partitioned genetic correlation

We further partitioned genetic correlation across 14 distinct functional categories, considering the highly positive genetic correlations observed between GERD and LC (*Figure 3*, Table S2). In 13 of the 14 functional categories, GERD was significantly correlated with overall LC, of which r_g values ranged from 0.17 (super enhancer) to 0.38 (conserved regions). Extending to subtype-specific LC, we noted significant associations in 10/14, 12/14, and 12/14 functional categories for LUAD, LUSC, and SCLC, respectively. Notably, the conserved region (r_g=0.30), conserved region (r_g=0.38), and promotor (r_g=0.41) displayed strongest genetic correlation for LUAD, LUSC, and SCLC, respectively.

Cross-trait meta-analysis and pleiotropic loci

The strong genetic correlation inspired us to locate pleiotropic loci between GERD and LC by performing CPASSOC. Cross-trait meta-analysis included a total of 2,194,995, 2,197,591, 2,197,145, and 2,202,470 SNPs shared between GERD and overall LC, LUAD, LUSC, and SCLC, respectively. Finally, CPASSOC identified 22 independent loci with genome-wide significance ($P_{CPASSOC}$ <5×10^{-8} and $P_{single\text{-}trait}$ <1×10^{-3}), including 14 pleiotropic loci between GERD and overall LC, 4 pleiotropic loci between GERD and LUAD, 8 pleiotropic loci between GERD and LUSC, and 2 pleiotropic loci between GERD and SCLC (*Table 2*, Figure S1). Near these shared loci, some widely reported oncogenes, such as *PTPRF*, *PBX3*, *RAB5B*, and *TCF4* (related SNPs: rs2782641, rs10156445, rs773109, and rs4500831), were observed.

After removing loci identified in previously reported single-trait GWASs or loci in LD ($r^2\geq0.2$) with previously identified loci, 3 loci were categorized novel pleiotropic loci: 2 shared between GERD and overall LC, and 2 shared between GERD and LUAD, with 1 locus overlapped between overall LC and LUAD. The most significant novel locus was rs537160, which was mapped to complement factor B (CFB), a pivotal component of the alternative signaling pathway in complement activation (55). rs10156445, as the second most significant novel locus, was near *PBX3*, a member of the PBX family interacting with the promoter of tumor suppressor *p53* (54).

Identification of causal variants and colocalization

Using FM-summary algorithm, each of the identified pleiotropic variants established a 99% credible set of causal variants, which offers potential targets for subsequent

© AME Publishing Company.

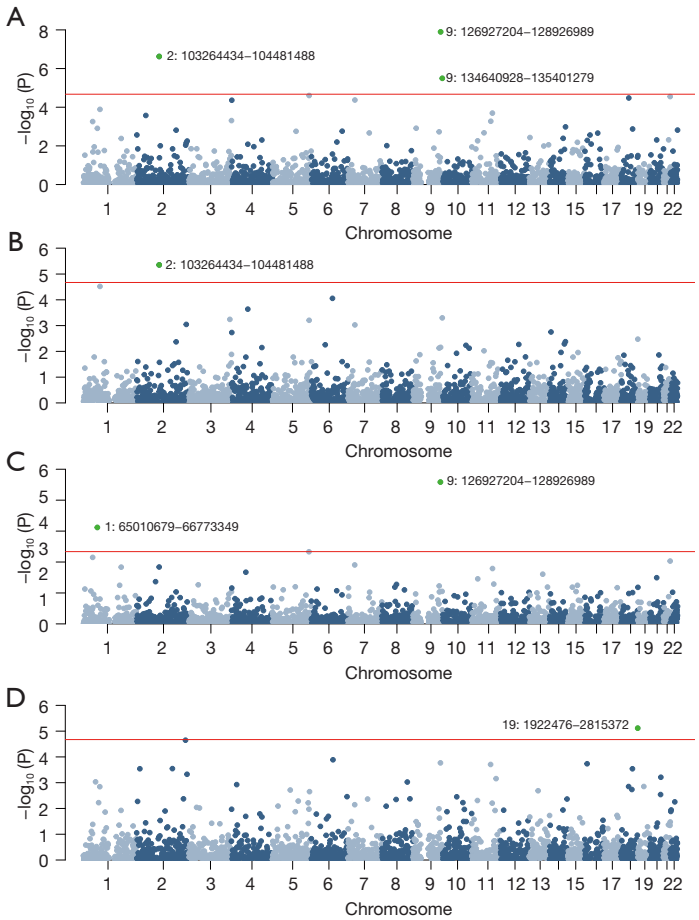

Figure 2 Manhattan plots for local genetic correlation between GERD and LC. (A) GERD and LC; (B) GERD and LUAD; (C) GERD and LUSC; (D) GERD and SCLC. The x-axis represents the chromosomal positions across human genome, while the y-axis represents the –log10 of P value. Each dot represents LD-independent genomic regions with the green dots representing significant regions (P<0.05/2,353). GERD, gastroesophageal reflux disease; LC, lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer.

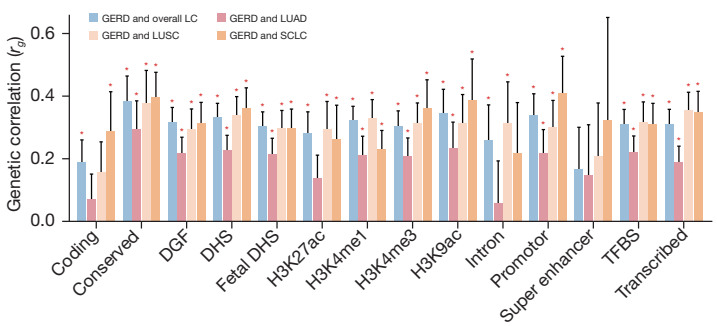

Figure 3 Partitioned genetic correlation between GERD and LC by genomic functional elements. Vertical axis represents genetic correlation. Horizontal axis represents 14 functional categories. Asterisks "*" represent significance (P<0.05), while error bars represent the standard error of genetic correlation. GERD, gastroesophageal reflux disease; LC, lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer; DGF, DNase I digital genomic foot-printing region; DHS, DNase I hypersensitive sites; TFBS, transcription factor-binding sites.

© AME Publishing Company.

Table 2 Significant pleiotropic SNPs identified by cross-trait meta-analysis ($P_{CPASSOC}$ <5×10^{-8} and $P_{single\text{-}trait}$ <1×10^{-3}, clumping r^2=0.2)

SNP	CHR: position	A1/A2	Beta		$P_{single\text{-}trait}$		$P_{CPASSOC}$	Gene[†]
			GERD	LC	GERD	LC		
GERD and over all LC								
rs17391694	Chr1: 78623626	C/T	−0.04	0.11	2.54×10^{-7}	2.62×10^{-8}	7.94×10^{-10}	–
rs2782641	Chr1: 44013355	G/A	−0.03	−0.04	4.33×10^{-8}	6.26×10^{-4}	5.67×10^{-10}	*PTPRF*
rs6711584	Chr2: 104421692	G/A	−0.03	0.04	2.66×10^{-11}	1.52×10^{-4}	2.13×10^{-13}	–
rs329122	Chr5: 133864599	G/A	0.03	−0.05	3.05×10^{-9}	1.69×10^{-5}	2.35×10^{-11}	*JADE2*
rs13207689	Chr6: 27369704	C/G	0.05	0.14	9.32×10^{-10}	9.29×10^{-11}	1.35×10^{-14}	*ZNF391*, *RP1-153G14.4*
rs13220495	Chr6: 26441640	C/T	0.04	0.13	1.96×10^{-8}	7.74×10^{-9}	2.91×10^{-12}	*BTN3A3*
rs17526722	Chr6: 25918855	G/A	0.03	0.13	5.59×10^{-5}	1.26×10^{-8}	1.58×10^{-8}	*SLC17A2*
rs2232423	Chr6: 28366151	A/G	0.05	0.15	1.37×10^{-11}	8.04×10^{-12}	2.54×10^{-17}	*ZSCAN12*
rs537160[‡]	Chr6: 31916400	A/G	−0.03	0.05	5.08×10^{-8}	3.98×10^{-5}	8.46×10^{-10}	*CFB*, *NELFE*, *C2*, *CYP21A2*
rs215614	Chr7: 32347335	G/A	0.03	0.04	4.08×10^{-11}	4.29×10^{-4}	1.31×10^{-13}	–
rs10156445[‡]	Chr9: 128617244	A/G	−0.02	−0.04	6.33×10^{-7}	7.81×10^{-4}	1.51×10^{-8}	*PBX3*
rs9328534	Chr9: 134874805	C/T	0.03	0.04	1.35×10^{-8}	4.67×10^{-4}	1.25×10^{-10}	*MED27*
rs773109	Chr12: 56374695	G/A	0.04	−0.04	8.71×10^{-14}	5.14×10^{-4}	5.40×10^{-16}	*RAB5B*, *RP11-603J24.7*
rs4500831	Chr18: 53097544	G/A	0.03	0.05	1.21×10^{-7}	3.42×10^{-4}	1.47×10^{-9}	TCF4
GERD and LUAD								
rs6695572	Chr1: 77945635	G/A	−0.02	0.12	4.09×10^{-4}	8.34×10^{-9}	2.14×10^{-8}	AK5
rs17391694[‡]	Chr1: 78623626	C/T	−0.04	0.14	2.54×10^{-7}	3.83×10^{-7}	8.34×10^{-9}	–
rs6711584	Chr2: 104421692	G/A	−0.03	0.07	2.66×10^{-11}	2.89×10^{-5}	8.68×10^{-13}	–
rs537160[‡]	Chr6: 31906797	A/G	−0.03	0.06	5.08×10^{-8}	7.89×10^{-4}	8.21×10^{-9}	*CFB*, *NELFE*, *C2*, *CYP21A2*
GERD and LUSC								
rs2782641	Chr1: 44013355	G/A	−0.03	−0.07	4.33×10^{-8}	2.76×10^{-4}	4.98×10^{-9}	*PTPRF*
rs329122	Chr5: 133864599	G/A	0.03	−0.08	3.05×10^{-9}	2.28×10^{-5}	5.12×10^{-10}	*JADE2*
rs13191445	Chr6: 26015489	G/A	0.03	0.25	5.35×10^{-5}	1.06×10^{-11}	5.56×10^{-11}	*HIST1H1A*, *HIST1H1PS2*, *U91328.22*
rs9379899	Chr6: 26603015	T/A	0.04	0.11	1.25×10^{-9}	2.17×10^{-4}	1.07×10^{-10}	*ABT1*
rs3922717	Chr6: 27030924	A/G	0.04	0.08	5.35×10^{-13}	3.75×10^{-4}	3.81×10^{-14}	*VN1R13P*
rs13219181	Chr6: 27136225	A/G	0.03	0.11	1.32×10^{-8}	2.56×10^{-5}	7.00×10^{-10}	–
rs200968	Chr6: 27859568	T/C	0.04	0.11	3.94×10^{-11}	4.25×10^{-5}	1.62×10^{-12}	*HIST1H2BO*, *HIST1H3J*, *HIST1H2AM*
rs2232426	Chr6: 28360659	G/C	0.05	0.22	1.39×10^{-11}	1.02×10^{-10}	1.63×10^{-14}	*ZSCAN12*
GERD and SCLC								
rs3172494	Chr3: 48731487	G/T	0.05	−0.15	6.71×10^{-9}	9.28×10^{-4}	5.12×10^{-9}	*IP6K2*
rs2232423	Chr6: 28366151	A/G	0.05	0.20	1.37×10^{-11}	2.14×10^{-4}	2.02×10^{-12}	*ZSCAN12*

[†], gene symbol mapped by VEP; [‡], novel SNPs, defined as shared SNPs that are neither driven by a single trait nor in LD with index SNPs identified in single-trait GWAS (LD r^2<0.2). SNPs, single nucleotide polymorphisms; CPASSOC, cross-phenotypic association; CHR, chromosome; A1, effect allele; A2, alternative allele; GERD, gastroesophageal reflux disease; LC, lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer; VEP, Variant Effect Predictor; LD, linkage disequilibrium; GWAS, genome-wide association study.

© AME Publishing Company.

experimental research (Tables S3-S6). As a result, we identified a set of 352, 42, 219, and 18 causal variants for overall LC, LUAD, LUSC, and SCLC. For the novel pleiotropic loci, we identified 1, 3, 11, 75 causal variants for rs17391694, rs537160, rs537160, and rs10156445, respectively.

To evaluate whether genetic variants influencing the association across traits were shared or distinct, the colocalization analysis was further performed. Approximately a half of pleiotropic loci showed colocalization at same candidate causal variants with PPH4 >0.5: 8/14 between GERD and overall LC, 2/4 between GERD and LUAD, 4/8 between GERD and LUSC, and 1/2 between GERD and SCLC (Table S7).

TWASs

After multiple testing (FDR <0.05) and intersecting the single-trait TWAS results across traits, multiple TWAS-significant gene-tissue pairs shared between GERD and LC were identified, including 30 genes shared between GERD and overall LC, 10 genes shared between GERD and LUAD, 11 genes shared between GERD and LUSC, and 12 genes shared between GERD and SCLC (Tables S8-S11, Figure S2). Among these gene-tissue pairs between GERD and overall LC, *ERAP1*, *FUBP1*, and *CEP57* were most frequently identified genes and simultaneously discovered in lung tissues and esophagus tissues (i.e., esophagus mucosa, esophagus muscularis, and esophagus gastroesophageal junction). As a member of the M1 family of aminopeptidases, *ERAP1* was previously implicated in autoimmunity and signals a role in susceptibility of LC (56). As a versatile DNA and RNA binding protein, *FUBP1* plays a role in multiple biological processes, and serves as an oncoprotein associated with multiple malignancies, including LC (57,58). CEP57, a centrosomal protein, is involved in the processes of microtubule nucleation and bundling associated with cell division error and thus potentially promote malignant progression (59,60). Additionally, *PBX3*, a factor interacting with the promoter of *p53* (54), was frequently identified in gene-tissue pairs between GERD and LUSC.

Bidirectional MR

Finally, we evaluated the causal association between GERD and LC by performing a two-sample MR. We identified a total of 91 GERD-associated SNPs as IVs, and *F*-statistics calculated >10 suggested strong IVs (Table S12). Utilizing the IVW method, GERD was found to be significantly associated with the risk of overall LC (OR =1.34, P=1.33×10^{-6}), which remained consistent in weight median (OR =1.28, P=2.03×10^{-4}), MR-PRESSO (OR =1.37, P=1.51×10^{-7}), and MR-CAUSE (OR =1.30, P=6.11×10^{-3}) (*Figure 4*, Table S13, Figures S3-S6). The estimates continued to be directionally consistent with MR-Egger regression, despite no significance (OR =1.08, P=0.82). Consistent results were also observed after excluding pleiotropic SNPs (OR =1.30, P=1.34×10^{-5}) or palindromic SNPs (OR =1.40, P=2.80×10^{-7}). No significant horizontal pleiotropy was observed ($P_{\text{MR-Egger intercept}}$ =0.54), and leave-one-out analyses detected no obvious outlying variants (Figure S7). Looking into the subtype-specific LC, significant causal associations also were identified in LUAD (IVW OR =1.25, P=2.71×10^{-3}), LUSC (IVW OR =1.52, P=9.59×10^{-9}), and SCLC (IVW OR =1.76, P=1.27×10^{-5}), which were further confirmed in sensitivity analyses except MR-Egger regression. Additionally, the power of all MR analyses was calculated to be 100% using estimates from IVW, suggesting a satisfactory statistical power (Table S14). Potential confounders were accounted for using MVMR, yielding estimates that exhibit a more pronounced magnitude and statistical significance, which suggests that the causal relationship between GER and LC remains independent of common confounding factors (Figure S8).

In the reverse-direction MR analysis, we identified a total of 14, 15, 13, and 4 SNPs for overall LC, LUAD, LUSC, and SCLC as IVs, with all *F*-statistics >10 suggesting strong IVs (Table S15). We observed no significant causal effect of LC on GERD: overall LC (IVW OR =1.02, P=0.24), LUAD (IVW OR =1.00, P=0.95), LUAD (IVW OR =0.99, P=0.46), and SCLC (IVW OR =0.98, P=0.27) (*Figure 5*).

Discussion

As far as we know, this genome-wide cross-trait analysis represents the first comprehensive investigation into the genetic correlation, pleiotropic loci, association between gene expression and trait, and causal relationship between GERD and LC, providing valuable insights into this complex genetic interplay. Our findings revealed a significantly genetic correlation underlying GERD and overall LC. After partitioning the whole genome, significant genetic correlations were identified within five genomic regions and multiple functional categories (e.g., conserved region, and promotor). The underlying genetic link was further divided into two categories: pleiotropy and causality, corresponding

© AME Publishing Company.

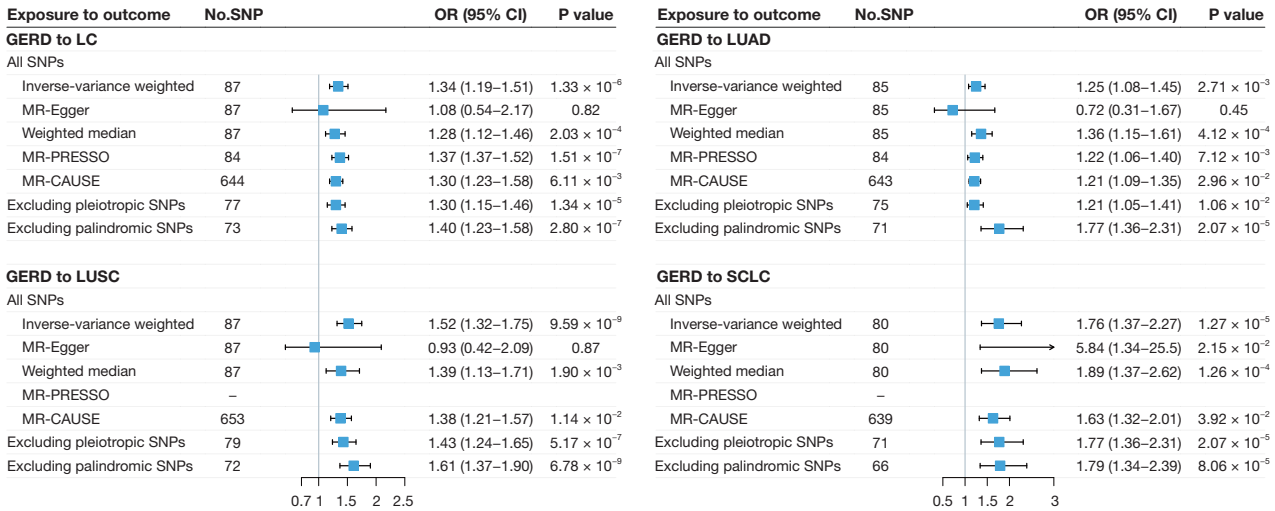

Figure 4 Estimates of causal effect sizes of GERD on LC using all GERD-associated SNPs, excluding pleiotropic SNPs or palindromic SNPs. Inverse variance-weighted approach was used as the primary outcome, while MR-Egger, weighted median, MR-PRESSO, and MR-CAUSE were applied as complementary analyses. GERD, gastroesophageal reflux disease; LC, lung cancer; MR, Mendelian randomization; PRESSO, Pleiotropy Residual Sum and Outlier; CAUSE, Causal Analysis Using Summary Effect estimates; SNPs, single nucleotide polymorphisms; LUSC, lung squamous cell carcinoma; LUAD, lung adenocarcinoma; SCLC, small-cell lung cancer; OR, odds ratio; CI, confidence interval.

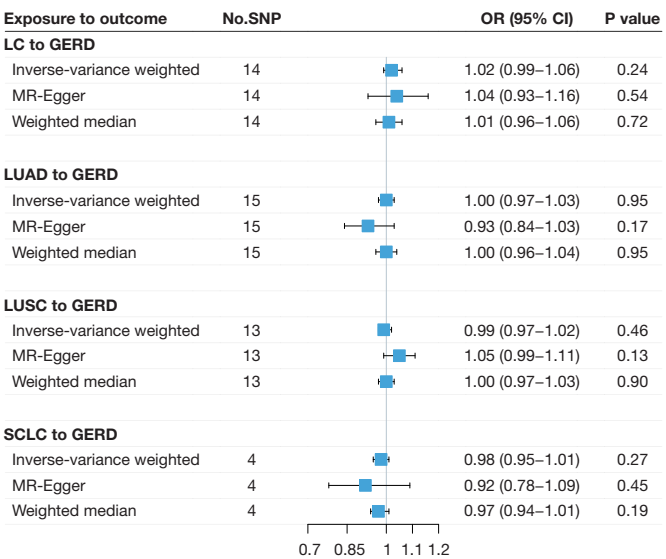

Figure 5 Estimates of causal effect sizes of LC on GERD using all LC-associated SNPs. Inverse variance-weighted approach was used as the primary outcome, while MR-Egger and weighted median were applied as complementary analyses. GERD, gastroesophageal reflux disease; LC, lung cancer; MR, Mendelian randomization; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer; SNPs, single nucleotide polymorphisms; OR, odds ratio; CI, confidence interval.

© AME Publishing Company.

with the identified pleiotropic loci from CPASSOC, the shared genes from TWAS, and the causal association through the bidirectional MR analysis. For subtype-specific LC, similar results were also observed for LUAD, LUSC, and SCLC. Taken together, these findings advance our understanding of the intricate link between a digestive disease and a respiratory malignancy, offering valuable implications for LC prevention in individuals with GERD.

Despite a significant global genetic correlation detected in LDSC, the estimated intercept of genetic covariance ranging from 0.001 to 0.008 suggested a presence of potential bias from sample overlap. Therefore, we employed LDSC with a constrained intercept to address this issue, and similarly, detected a significant global genetic correlation between GERD and LC (26). When separating the entire genome into approximately 2,353 distinct regions, a significant local genetic correlation between GERD and overall LC, as well as LUSC, was identified, specifically at 9q33.3. This genomic region contains *PBX3*, which has previously been reported to be associated with GERD and LC (12,53,54). Furthermore, we observed significant genetic correlations in multiple annotated regions of the genome using stratified LDSC. Notably, the conserved region exhibited the highest partitioned r_g, while other non-coding regions, including specific histone modification marks, such as H3K4me1 and H3K27me3, and histone acetylation marks, such as H3K27ac, H3K9ac, also showed significant partitioned r_g. These findings align with prior studies highlighting the crucial role of epigenetic modification in LC development (61,62).

Through the cross-trait meta-analysis, we revealed 22 pleiotropic loci between GERD and LC, among which 18 loci have been reported to be associated with either one or both traits. For instance, the shared SNP rs4500831 (18q21.2) showed LD with rs1942262 (r^2=0.21) previously identified in the GWAS study of GERD (12), which was mapped to *TCF4* implicated in the development of LC (63). Additionally, several pleiotropic loci were mapped to genes associated with risks of various carcinomas, such as *PTPRF*, *JADE2*, *SLC17A2*, *MED27*, and *RAB5B*. Multiple genes, including *PTPRF*, *JADE2*, *ZNF391*, *SLC17A2*, *MED27*, *RAB5B*, and *ZSCAN12*, exhibited significant evidence of colocalization (PPH4 >0.5), indicating etiological correlations. Cross-trait meta-analysis has the advantage of revealing signals that have not reached genome-wide significance in a single-trait analysis (64). Within these loci in our study, we identified four novel loci associated with both GERD and LC, among which we highlight two genes (*CFB* and *NELFE*), both mapped by the same locus (index SNP: rs537160).

CFB is a factor that binds C3 to form C3B in the alternative pathway, playing a pivotal role in labeling target particles and thereby contributing to effective target clearance (65). Through the integration of proteomic analysis, CFB has been identified as a potential biomarker for pancreatic cancer (66). Also, a recent study found that elevated CFB expression serves as an independent predictor of long-term survival of LUAD (65). Furthermore, complementary pathway may play a critical role in the development of GERD. Previous studies reported that the transcription factor NF-κB is associated with the development of GERD, and the activation of NF-κB is mediated through the alternative pathway (67,68). These observations underscore the potential etiology of CFB underlying GERD and LC.

Additionally, NELFE is RNA-binding protein that plays a role in tumor biology and progression (69,70). Prior study has revealed that NELFE has the potential to induce hepatocellular carcinoma by regulating the MYC signaling pathway (71). Furthermore, NELFE may promote the tumorigenesis and metastasis of pancreatic cancer via the Wnt/β-catenin signaling pathway (72). Through the whole-exome sequencing of early non-smokers with LUAD, NELFE was also identified as a candidate driver marker (73). Nevertheless, further study is warranted to validate and explore the biological mechanism of NELFE in the tumorigenesis of LC.

The TWAS analysis evaluated pleiotropy at the level of gene expression by combining GTEx tissue-specific expression data and GWAS summary data. Specifically, both CPASSOC and TWAS identified *PTPRF* and *PBX3* as relevant genes. Furthermore, *PBX3*, located at 9q33.3, was also identified in the local genetic correlation analysis. Two shared genes, *CEP57* and *FUBP1*, were also identified by TWAS, and have been reported to have a direct or indirect association with GERD and LC (12,56,58). In summary, these shared biological targets between GERD and LC suggest potential therapeutic strategies for the coexisting groups in clinical practice. Further studies are warranted to elucidate the underlying mechanisms.

Utilizing a comprehensive bidirectional MR analysis, our results revealed a significant causal association between GERD and LC, further extending to subtype-specific LC. Of note, the strength of the causal estimates between GERD and LC largely aligns with the genetic correlation; specifically, the correlation is strongest between GERD

© AME Publishing Company.

and LUSC/SCLC, whereas the correlation with LUAD is the weakest. Compared with prior MR studies, our research significantly advances previous findings in several crucial aspects (18-20). We leveraged the GWAS of LC with an expanded sample size, substantially enhancing the statistical ability to discover causal relationships. For example, the causal association between GERD and SCLC was not discovered from the MR study by Liu *et al.*, who utilized limited sample size GWAS data from the FinnGen database (only 461 cases) (18). Additionally, the sensitivity analyses were performed to scrutinize the assumptions of MR, thereby offering further support for the reliability of our main findings. To ensure causal estimates independent of potential confounding factors, comprehensive MVMR analyses were carried out. Through a reverse directional MR design, we found no significant causal association of genetically predicted LC on the risk of GERD, which had not been previously explored in prior MR studies (18-20). Taken together, the estimated causal effects were consistently affirmed among multiple sensitivity analyses and statistical approaches, indicating its robustness. In line with our findings, previous population-based epidemiologic studies also reported positive associations of GERD and LC (9,10,74). Interestingly, a large-scale cohort study reported that anti-reflux surgery led to a significant reduction in the risk of LUSC [standardized incidence ratio (SIR) =0.75, 95% CI: 0.60–0.92] and SCLC (SIR =0.63, 95% CI: 0.44–0.90), with a protective trend in LUAD (SIR =0.80, 95% CI: 0.62–1.03) (10). These observations closely resemble the findings of our study, finding a significantly positive causal effect between GERD and LUSC/SCLC, while the correlation with LUAD is marginally significant.

Several limitations in the current study should be acknowledged. Firstly, to mitigate potential bias from population stratification, we focused exclusively on individuals of European ancestry for our genetic data. However, it is important to note that the incidence of GERD may exhibit racial disparities (75), suggesting the need for further research to validate the generalizability of our findings in other ethnic populations. Secondly, our study was limited to data from autosomes due to existing limitations in the analytical software, which does not support the analysis of sex chromosomes. Thirdly, while we mapped pleiotropic SNPs to relevant genes, further investigations are warranted to pinpoint the causal genes responsible for the observed signals. Finally, our study relied on summary-level data rather than individual-level data, determined by data limitations. While summary-level data provide a larger sample size, leading to increased statistical power in causal estimates (42), it is important to acknowledge its drawbacks. Compared with individual-level data, summary-level data do not account for some important confounders for each individual, such as local socioeconomic, medical situations, and other factors.

Conclusions

In summary, using a novel statistical genetic approach based on the hitherto largest GWAS summary data, the study sheds light on the observational association between GERD and LC. These findings provide valuable evidence of genetic correlation, identifying pleiotropic loci, and suggesting a potential causal association between GERD and LC. This study conveys a crucial public health message: managing individuals with GERD may potentially contribute to reducing the long-term burden of malignant diseases.

Acknowledgments

Funding: This study was supported by 1.3.5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (ZYJC21002 to Dr. L.L.), and National Natural Science Foundation of China (82102968 to Dr. J.Z.).

Footnote

Reporting Checklist: The authors have completed the STREGA reporting checklist. Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-345/rc

Peer Review File: Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-345/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-345/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The ethical approval for each summary-level data can be found from the corresponding studies. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

© AME Publishing Company.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Vakil N, van Zanten SV, Kahrilas P, et al. The Montreal definition and classification of gastroesophageal reflux disease: a global evidence-based consensus. Am J Gastroenterol 2006;101:1900-20; quiz 1943.
2. The global, regional, and national burden of gastro-oesophageal reflux disease in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Gastroenterol Hepatol 2020;5:561-81.
3. Broers C, Tack J, Pauwels A. Review article: gastro-oesophageal reflux disease in asthma and chronic obstructive pulmonary disease. Aliment Pharmacol Ther 2018;47:176-91.
4. Hayes D Jr. A review of bronchiolitis obliterans syndrome and therapeutic strategies. J Cardiothorac Surg 2011;6:92.
5. Gaillet G, Favelle O, Guilleminault L, et al. Gastroesophageal reflux disease is a risk factor for severity of organizing pneumonia. Respiration 2015;89:119-26.
6. Baqir M, Vasirreddy A, Vu AN, et al. Idiopathic pulmonary fibrosis and gastroesophageal reflux disease: A population-based, case-control study. Respir Med 2021;178:106309.
7. Herbella FA, Neto SP, Santoro IL, et al. Gastroesophageal reflux disease and non-esophageal cancer. World J Gastroenterol 2015;21:815-9.
8. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69:7-34.
9. Hsu CK, Lai CC, Wang K, et al. Risk of lung cancer in patients with gastro-esophageal reflux disease: a population-based cohort study. PeerJ 2016;4:e2753.
10. Yanes M, Santoni G, Maret-Ouda J, et al. Antireflux surgery and risk of lung cancer by histological type in a multinational cohort study. Eur J Cancer 2020;138:80-8.
11. Zhu Z, Hasegawa K, Camargo CA Jr, et al. Investigating asthma heterogeneity through shared and distinct genetics: Insights from genome-wide cross-trait analysis. J Allergy Clin Immunol 2021;147:796-807.
12. Ong JS, An J, Han X, et al. Multitrait genetic association analysis identifies 50 new risk loci for gastro-oesophageal reflux, seven new loci for Barrett's oesophagus and provides insights into clinical heterogeneity in reflux diagnosis. Gut 2022;71:1053-61.
13. Wang C, Dai J, Qin N, et al. Analyses of rare predisposing variants of lung cancer in 6,004 whole genomes in Chinese. Cancer Cell 2022;40:1223-1239.e6.
14. Cameron AJ, Lagergren J, Henriksson C, et al. Gastroesophageal reflux disease in monozygotic and dizygotic twins. Gastroenterology 2002;122:55-9.
15. Mohammed I, Cherkas LF, Riley SA, et al. Genetic influences in gastro-oesophageal reflux disease: a twin study. Gut 2003;52:1085-9.
16. Mucci LA, Hjelmborg JB, Harris JR, et al. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. JAMA 2016;315:68-76.
17. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med 2000;343:78-85.
18. Liu Y, Lai H, Zhang R, et al. Causal relationship between gastro-esophageal reflux disease and risk of lung cancer: insights from multivariable Mendelian randomization and mediation analysis. Int J Epidemiol 2023;52:1435-47.
19. Li L, Ren Q, Zheng Q, et al. Causal associations between gastroesophageal reflux disease and lung cancer risk: A Mendelian randomization study. Cancer Med 2023;12:7552-9.
20. Yang J, Nie D, Chen Y, et al. The role of smoking and alcohol in mediating the effect of gastroesophageal reflux disease on lung cancer: A Mendelian randomization study. Front Genet 2022;13:1054132.
21. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 2003;32:1-22.
22. Solovieff N, Cotsapas C, Lee PH, et al. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet 2013;14:483-95.
23. Zhu Z, Lee PH, Chaffin MD, et al. A genome-wide cross-trait analysis from UK Biobank highlights the shared genetic architecture of asthma and allergic diseases. Nat Genet 2018;50:857-64.
24. Little J, Higgins JP, Ioannidis JP, et al. STrengthening the REporting of Genetic Association studies (STREGA): an extension of the STROBE Statement. Ann Intern Med 2009;150:206-15.

25. McKay JD, Hung RJ, Han Y, et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. Nat Genet 2017;49:1126-32.
26. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet 2015;47:1236-41.
27. Zhang Y, Lu Q, Ye Y, et al. SUPERGNOVA: local genetic correlation analysis reveals heterogeneous etiologic sharing of complex traits. Genome Biol 2021;22:262.
28. Finucane HK, Bulik-Sullivan B, Gusev A, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet 2015;47:1228-35.
29. Bernstein BE, Stamatoyannopoulos JA, Costello JF, et al. The NIH Roadmap Epigenomics Mapping Consortium. Nat Biotechnol 2010;28:1045-8.
30. Li X, Zhu X. Cross-Phenotype Association Analysis Using Summary Statistics from GWAS. Methods Mol Biol 2017;1666:455-67.
31. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-75.
32. Qu Y, Xiao C, Wu X, et al. Genetic Correlation, Shared Loci, and Causal Association Between Sex Hormone-Binding Globulin and Bone Mineral Density: Insights From a Large-Scale Genomewide Cross-Trait Analysis. J Bone Miner Res 2023;38:1635-44.
33. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. Genome Biol 2016;17:122.
34. Schaid DJ, Chen W, Larson NB. From genome-wide associations to candidate causal variants by statistical fine-mapping. Nat Rev Genet 2018;19:491-504.
35. Wu X, Xiao C, Han Z, et al. Investigating the shared genetic architecture of uterine leiomyoma and breast cancer: A genome-wide cross-trait analysis. Am J Hum Genet 2022;109:1272-85.
36. Farh KK, Marson A, Zhu J, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature 2015;518:337-43.
37. Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet 2014;10:e1004383.
38. Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet 2016;48:245-52.
39. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580-5.
40. Burgess S, Thompson SG; . Avoiding bias from weak instruments in Mendelian randomization studies. Int J Epidemiol 2011;40:755-64.
41. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. Int J Epidemiol 2013;42:1497-501.
42. Burgess S, Scott RA, Timpson NJ, et al. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. Eur J Epidemiol 2015;30:543-52.
43. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44:512-25.
44. Bowden J, Davey Smith G, Haycock PC, et al. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 2016;40:304-14.
45. Morrison J, Knoblauch N, Marcus JH, et al. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. Nat Genet 2020;52:740-7.
46. 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;50:693-8.
47. Kamat MA, Blackshaw JA, Young R, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics 2019;35:4851-3.
48. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. Am J Epidemiol 2015;181:251-60.
49. Pulit SL, Stoneman C, Morris AP, et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet 2019;28:166-74.
50. Liu M, Jiang Y, Wedow R, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet 2019;51:237-44.
51. Klimentidis YC, Raichlen DA, Bea J, et al. Genome-wide association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple variants including CADM2 and APOE. Int J Obes (Lond) 2018;42:1161-76.
52. Dashti HS, Jones SE, Wood AR, et al. Genome-wide association study identifies genetic loci for self-reported

© AME Publishing Company.

habitual sleep duration supported by accelerometer-derived estimates. Nat Commun 2019;10:1100.
53. Plowright L, Harrington KJ, Pandha HS, et al. HOX transcription factors are potential therapeutic targets in non-small-cell lung cancer (targeting HOX genes in lung cancer). Br J Cancer 2009;100:470-5.
54. Li WF, Herkilini A, Tang Y, et al. The transcription factor PBX3 promotes tumor cell growth through transcriptional suppression of the tumor suppressor p53. Acta Pharmacol Sin 2021;42:1888-99.
55. Ricklin D, Hajishengallis G, Yang K, et al. Complement: a key system for immune surveillance and homeostasis. Nat Immunol 2010;11:785-97.
56. Babaie F, Hosseinzadeh R, Ebrazeh M, et al. The roles of ERAP1 and ERAP2 in autoimmunity and cancer immunity: New insights and perspective. Mol Immunol 2020;121:7-19.
57. Debaize L, Troadec MB. The master regulator FUBP1: its emerging role in normal cell function and malignant development. Cell Mol Life Sci 2019;76:259-81.
58. Byun J, Han Y, Li Y, et al. Cross-ancestry genome-wide meta-analysis of 61,047 cases and 947,237 controls identifies new susceptibility loci contributing to lung cancer. Nat Genet 2022;54:1167-77.
59. Mang J, Korzeniewski N, Dietrich D, et al. Prognostic Significance and Functional Role of CEP57 in Prostate Cancer. Transl Oncol 2015;8:487-96.
60. Snape K, Hanks S, Ruark E, et al. Mutations in CEP57 cause mosaic variegated aneuploidy syndrome. Nat Genet 2011;43:527-9.
61. Li W, Tian W, Yuan G, et al. Molecular basis of nucleosomal H3K36 methylation by NSD methyltransferases. Nature 2021;590:498-503.
62. Sengupta D, Zeng L, Li Y, et al. NSD2 dimethylation at H3K36 promotes lung adenocarcinoma pathogenesis. Mol Cell 2021;81:4481-4492.e9.
63. Zhang HJ, Chang WJ, Jia CY, et al. Destrin Contributes to Lung Adenocarcinoma Progression by Activating Wnt/β-Catenin Signaling Pathway. Mol Cancer Res 2020;18:1789-802.
64. Zhu Z, Anttila V, Smoller JW, et al. Statistical power and utility of meta-analysis methods for cross-phenotype genome-wide association studies. PLoS One 2018;13:e0193256.
65. He C, Li Y, Zhang R, et al. Low CFB expression is independently associated with poor overall and disease-free survival in patients with lung adenocarcinoma. Oncol Lett 2021;21:478.
66. Lee MJ, Na K, Jeong SK, et al. Identification of human complement factor B as a novel biomarker candidate for pancreatic ductal adenocarcinoma. J Proteome Res 2014;13:4878-88.
67. Zhang ML, Ran LQ, Wu MJ, et al. NF-κB: A novel therapeutic pathway for gastroesophageal reflux disease? World J Clin Cases 2022;10:8436-42.
68. Liu T, Zhang L, Joo D, et al. NF-κB signaling in inflammation. Signal Transduct Target Ther 2017;2:17023.
69. Roesch A, Becker B, Meyer S, et al. Retinoblastoma-binding protein 2-homolog 1: a retinoblastoma-binding protein downregulated in malignant melanomas. Mod Pathol 2005;18:1249-57.
70. Hodson DJ, Screen M, Turner M. RNA-binding proteins in hematopoiesis and hematological malignancy. Blood 2019;133:2365-73.
71. Dang H, Pomyen Y, Martin SP, et al. NELFE-Dependent MYC Signature Identifies a Unique Cancer Subtype in Hepatocellular Carcinoma. Sci Rep 2019;9:3369.
72. Han L, Zan Y, Huang C, et al. NELFE promoted pancreatic cancer metastasis and the epithelial-to-mesenchymal transition by decreasing the stabilization of NDRG2 mRNA. Int J Oncol 2019;55:1313-23.
73. Renieri A, Mencarelli MA, Cetta F, et al. Oligogenic germline mutations identified in early non-smokers lung adenocarcinoma patients. Lung Cancer 2014;85:168-74.
74. Amarnath S, Starr A, Chukkalore D, et al. The Association Between Gastroesophageal Reflux Disease and Non-Small Cell Lung Cancer: A Retrospective Case-Control Study. Gastroenterology Res 2022;15:173-9.
75. Richter JE, Rubenstein JH. Presentation and Epidemiology of Gastroesophageal Reflux Disease. Gastroenterology 2018;154:267-76.

Cite this article as: Wu D, Zhou J, Song L, Zheng Q, Wang T, Ren Z, Huang Y, Liu S, Liu L. A multi-level investigation of the genetic relationship between gastroesophageal reflux disease and lung cancer. Transl Lung Cancer Res 2024;13(9):2373-2387. doi: 10.21037/tlcr-24-345

© AME Publishing Company.

Supplementary

Table S1 Details of GWAS summary data

Phenotype	N_{cases}	$N_{controls}$	No. SNPs	Consortium	Population	Year	PubMed ID
Gastroesophageal reflux disease	78,707	288,734	2,324,711	UKBB and QSKIN	European	2022	34187846
Lung cancer	29,266	56,450	7,884,164	ILCCO	European	2017	28604730
Lung adenocarcinoma	11,273	55,483	7,876,477	ILCCO	European	2017	28604730
Lung squamous cell carcinoma	7,426	55,627	7,865,405	ILCCO	European	2017	28604730
Small-cell lung cancer	2,664	21,444	7,644,095	ILCCO	European	2017	28604730

GWAS, genome-wide association study; UKBB, UK Biobank; QSKIN, Queensland Sun and Health Study; ILCCO, International Lung Cancer Consortium.

Table S2 Partitioned genetic correlation between GERD and LC

Functional category	Phenotype1	Phenotype2	r_g	r_g_se	P value
Coding regions	GERD	LC	0.1894	0.0707	7.40E−03
		LUAD	0.0705	0.0804	3.81E−01
		LUSC	0.1565	0.0974	1.08E−01
		SCLC	0.2891	0.1249	2.06E−02
Conserved regions	GERD	LC	0.3838	0.0804	1.81E−06
		LUAD	0.2950	0.0899	1.00E−03
		LUSC	0.3790	0.1031	2.00E−04
		SCLC	0.3974	0.0787	4.36E−07
DNaseI digital genomic foot-printing region	GERD	LC	0.3156	0.0482	5.78E−11
		LUAD	0.2166	0.0521	3.23E−05
		LUSC	0.2957	0.0633	3.05E−06
		SCLC	0.3121	0.0679	4.35E−06
DNase I hypersensitive site	GERD	LC	0.3329	0.0441	4.48E−14
		LUAD	0.2266	0.0482	2.62E−06
		LUSC	0.3382	0.0603	2.00E−08
		SCLC	0.3616	0.0649	2.51E−08
Fetal DNase I hypersensitivve site	GERD	LC	0.3043	0.0454	2.08E−11
		LUAD	0.2138	0.0516	3.47E−05
		LUSC	0.2988	0.0554	6.82E−08
		SCLC	0.2987	0.0600	6.52E−07
H3K4me1	GERD	LC	0.3236	0.0439	1.74E−13
		LUAD	0.2103	0.0616	6.00E−04
		LUSC	0.3301	0.0588	1.94E−08
		SCLC	0.2309	0.0594	1.00E−04
H3K4me3	GERD	LC	0.3039	0.0490	5.47E−10
		LUAD	0.2087	0.0576	3.00E−04
		LUSC	0.3147	0.0635	7.26E−07
		SCLC	0.3612	0.0910	7.22E−05
H3K9ac	GERD	LC	0.3458	0.0760	5.30E−06
		LUAD	0.2325	0.0844	5.90E−03
		LUSC	0.3123	0.0928	8.00E−04
		SCLC	0.3873	0.1314	3.20E−03
H3K27ac	GERD	LC	0.2823	0.0675	2.89E−05
		LUAD	0.1364	0.0750	6.91E−02
		LUSC	0.2944	0.0885	9.00E−04
		SCLC	0.2639	0.1071	1.38E−02
Intron	GERD	LC	0.2579	0.1141	2.37E−02
		LUAD	0.0567	0.1363	6.77E−01
		LUSC	0.3125	0.1331	1.89E−02
		SCLC	0.2177	0.1617	1.78E−01

Table S2 *(continued)*

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S2 *(continued)*

Functional category	Phenotype1	Phenotype2	r_g	$r_{g_}$se	P value
Promotor	GERD	LC	0.3382	0.0693	1.05E−06
		LUAD	0.2164	0.0767	4.80E−03
		LUSC	0.3014	0.0849	4.00E−04
		SCLC	0.4093	0.1180	5.00E−04
Super enhancer	GERD	LC	0.1664	0.1340	2.14E−01
		LUAD	0.1466	0.1619	3.65E−01
		LUSC	0.2087	0.1694	2.18E−01
		SCLC	0.3248	0.3265	3.20E−01
Transcription factor-binding sites	GERD	LC	0.3097	0.0478	8.78E−11
		LUAD	0.2209	0.0518	2.02E−05
		LUSC	0.3184	0.0631	4.45E−07
		SCLC	0.3117	0.0652	1.77E−06
Transcribed regions	GERD	LC	0.3106	0.0472	4.55E−11
		LUAD	0.1881	0.0520	3.00E−04
		LUSC	0.3539	0.0585	1.46E−09
		SCLC	0.3474	0.0679	3.08E−07

r_g, genetic correlation; se, standard error; GERD, gastroesophageal reflux disease; LC, lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer.

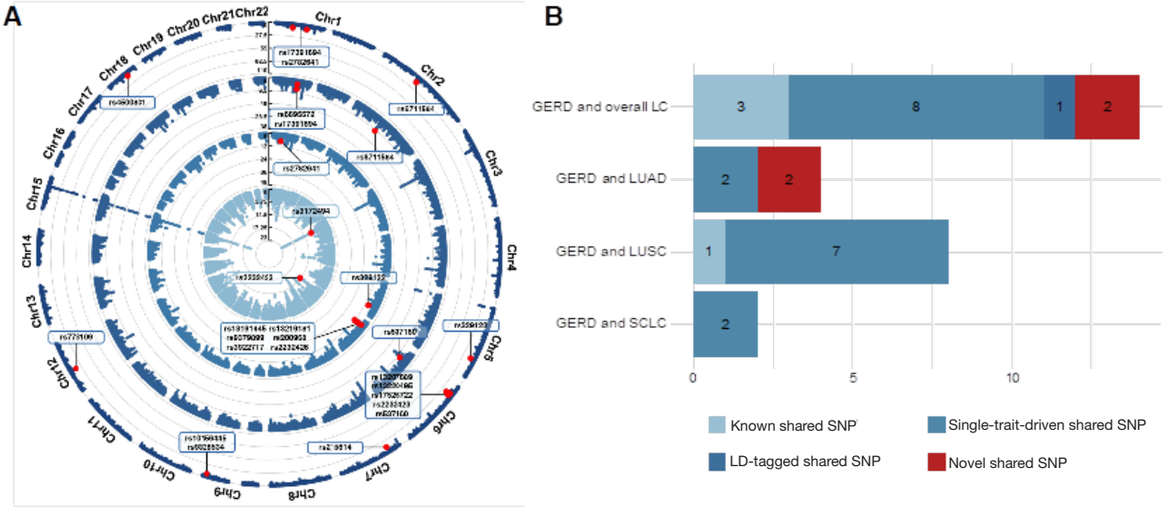

Figure S1 Cross-trait meta-analysis between GERD and LC. (A) Pleiotropic loci shared between GERD and LC. From periphery to center, each circular Manhattan plot represents results of the cross-trait meta-analysis between GERD and LC, as well as subtype-specific LC (LUAD, LUSC, and SCLC). The red dots represent significant pleiotropic loci in the cross-trait meta-analysis ($P_{CPASSOC}$ <5×10^{-8} and $P_{single-trait}$ <1×10^{-3} in both traits). (B) Bar plot of significant pleiotropic loci between GERD and LC. SNPs are categorized into four groups based on their single-trait and cross-trait characteristics: (I) "known" shared SNPs, referring to SNPs that reached genome-wide significance in both traits (P_{GERD} <5×10^{-8} and PLC <5×10^{-8}); (II) "single-trait-driven" shared SNPs, referring to SNPs reaching genome-wide significance in one of the two traits, either P_{GERD} <5×10^{-8} or P_{LC} <5×10^{-8}; (III) "LD-tagged" shared SNPs, referring to SNPs not reaching genome-wide significance in both traits (P_{GERD} >5×10^{-8} and P_{LC} >5×10^{-8}), but showing LD (r^2=0.2) with index SNPs previously identified by single-trait GWAS; and (IV) novel shared SNPs, referring to significant pleiotropic SNPs that did not reach genome-wide significance in both traits (1×10^{-3}< $P_{single-trait}$ <5×10^{-8}) and were not in LD with previously identified SNPs in single-trait GWAS (r2<0.2). GERD, gastroesophageal reflux disease; LC, lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer; SNP, single nucleotide polymorphism; CPASSOC, cross-phenotypic association; GWAS, genome-wide association study; LD, linkage disequilibrium.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S3 List of SNPs in the 99% credible set identified from fine-mapping analysis for each CPASSOC-identified locus shared between GERD and overall LC

Index SNP	99% credible-set SNPs	Chromosome	Position	CumSum
rs17391694	rs17391694	1	78623626	0.9978
rs2782641	rs1013793	1	44207553	0.9874
rs2782641	rs10890281	1	44245939	0.9863
rs2782641	rs11210860	1	43982527	0.9387
rs2782641	rs11210869	1	44026040	0.8560
rs2782641	rs11577403	1	43989773	0.9286
rs2782641	rs11805774	1	44152715	0.8890
rs2782641	rs12410155	1	44188465	0.9467
rs2782641	rs17371903	1	44070691	0.8352
rs2782641	rs17401357	1	44153619	0.9040
rs2782641	rs2152113	1	43983569	0.7905
rs2782641	rs2270972	1	44158129	0.8735
rs2782641	rs2274465	1	44121557	0.9533
rs2782641	rs2782641	1	44013355	0.2861
rs2782641	rs2819333	1	44014573	0.7594
rs2782641	rs2819334	1	44014735	0.6716
rs2782641	rs2842185	1	44019731	0.3525
rs2782641	rs2842187	1	44014949	0.4727
rs2782641	rs2842188	1	44014280	0.5761
rs2782641	rs3791034	1	44145130	0.9837
rs2782641	rs3791035	1	44154479	0.9183
rs2782641	rs3791040	1	44202733	0.9820
rs2782641	rs3791043	1	44219546	0.9905
rs2782641	rs3791134	1	44049156	0.4183
rs2782641	rs3791136	1	44049947	0.5259
rs2782641	rs4660257	1	44148168	0.8136
rs2782641	rs4660740	1	44148617	0.9851
rs2782641	rs489319	1	44131794	0.9642
rs2782641	rs501299	1	44051834	0.6241
rs2782641	rs605709	1	44058467	0.7168
rs2782641	rs618678	1	44133299	0.9690
rs2782641	rs6429637	1	44198531	0.9802
rs2782641	rs660899	1	44117006	0.9588
rs2782641	rs6664485	1	44225532	0.9895
rs2782641	rs6673970	1	44199469	0.9784
rs2782641	rs673253	1	44062154	0.9729
rs2782641	rs7516647	1	44237465	0.9884
rs2782641	rs9787076	1	44141149	0.9764
rs6711584	rs17343925	2	104426377	0.8771
rs6711584	rs2678670	2	104469564	0.7454
rs6711584	rs2945452	2	104469422	0.5700
rs6711584	rs4347858	2	104423042	0.3914
rs6711584	rs6707445	2	104420858	0.9993
rs6711584	rs6711584	2	104421692	0.2127
rs329122	rs11242219	5	133848760	0.9941
rs329122	rs1476096	5	133837021	0.9594
rs329122	rs1981627	5	133838180	0.9463
rs329122	rs329117	5	133860101	0.9292
rs329122	rs329120	5	133861756	0.6415
rs329122	rs329122	5	133864599	0.3310
rs329122	rs329124	5	133865452	0.9065
rs329122	rs4958244	5	133845380	0.9887
rs329122	rs6891328	5	133851526	0.9722
rs329122	rs7701346	5	133828356	0.9812
rs13207689	rs10484399	6	27534528	0.9741
rs13207689	rs12190473	6	27024687	0.8650
rs13207689	rs12215241	6	27023081	0.9638
rs13207689	rs12215773	6	27039233	0.9704
rs13207689	rs13191227	6	27390115	0.5158
rs13207689	rs13193542	6	27702425	0.9879
rs13207689	rs13194781	6	27815639	0.9723
rs13207689	rs13195040	6	27413924	0.7365
rs13207689	rs13195636	6	27509493	0.9552
rs13207689	rs13196692	6	27379119	0.6278
rs13207689	rs13199649	6	27868792	0.4014
rs13207689	rs13199906	6	27834139	0.9792
rs13207689	rs13202295	6	27698837	0.9203
rs13207689	rs13207082	6	27251379	0.9905
rs13207689	rs13207689	6	27369704	0.2241
rs13207689	rs13209332	6	27520752	0.9584
rs13207689	rs13210634	6	27646492	0.9683
rs13207689	rs13212318	6	27688841	0.9808
rs13207689	rs13215275	6	27647509	0.9661
rs13207689	rs13217620	6	27653120	0.8913
rs13207689	rs17695758	6	27837183	0.9776
rs13207689	rs17749927	6	27669976	0.9615
rs13207689	rs17750424	6	27701122	0.9865
rs13207689	rs17750747	6	27730334	0.9851
rs13207689	rs17751184	6	27775028	0.9759
rs13207689	rs200965	6	27866384	0.9301
rs13207689	rs200968	6	27859568	0.9418
rs13207689	rs200974	6	27855845	0.9469
rs13207689	rs200975	6	27855625	0.9519
rs13207689	rs200977	6	27854301	0.9822
rs13207689	rs200979	6	27852357	0.9360
rs13207689	rs3922717	6	27030924	0.8154
rs13207689	rs483143	6	27846744	0.9837
rs13207689	rs6904596	6	27491299	0.9060
rs13207689	rs6940638	6	27046250	0.9892
rs13220495	rs10484442	6	26555879	0.9680
rs13220495	rs1056667	6	26510564	0.9448
rs13220495	rs10946834	6	26533664	0.9366
rs13220495	rs12526680	6	26550954	0.9421
rs13220495	rs1321479	6	26501897	0.9631
rs13220495	rs1321481	6	26538210	0.9832
rs13220495	rs1407045	6	26476155	0.6332
rs13220495	rs1570059	6	26573325	0.8899
rs13220495	rs1570060	6	26573562	0.8804
rs13220495	rs1884946	6	26545308	0.9338
rs13220495	rs2024970	6	26497520	0.7693
rs13220495	rs2145318	6	26496603	0.3722
rs13220495	rs2179152	6	26325888	0.2660
rs13220495	rs2255070	6	26501777	0.9813
rs13220495	rs3736781	6	26505362	0.9879
rs13220495	rs3736782	6	26505403	0.9310
rs13220495	rs4145910	6	26313305	0.1380
rs13220495	rs4573	6	26546808	0.9892
rs13220495	rs4713006	6	26519872	0.9474
rs13220495	rs4713008	6	26538268	0.9133
rs13220495	rs4871	6	26545632	0.9751
rs13220495	rs6456733	6	26566804	0.9030
rs13220495	rs6456735	6	26574149	0.8635
rs13220495	rs6903973	6	26499942	0.9793
rs13220495	rs6918360	6	26577867	0.8751
rs13220495	rs6918506	6	26577857	0.8514
rs13220495	rs6918854	6	26577924	0.8945
rs13220495	rs6922824	6	26553815	0.9253
rs13220495	rs6926629	6	26499903	0.9655
rs13220495	rs6930120	6	26555484	0.9164

Table S3 *(continued)*

© AME Publishing Company.

https://dx.doi.org/10.21037/tlcr-24-345

Table S3 *(continued)*

Index SNP	99% credible-set SNPs	Chromosome	Position	CumSum
rs13220495	rs6933176	6	26540178	0.9580
rs13220495	rs6939048	6	26327953	0.5561
rs13220495	rs6940188	6	26562029	0.9194
rs13220495	rs6941022	6	26553531	0.9065
rs13220495	rs742090	6	26415637	0.9905
rs13220495	rs766406	6	26319588	0.8851
rs13220495	rs766407	6	26319534	0.4694
rs13220495	rs767471	6	26557854	0.9225
rs13220495	rs7753565	6	26560012	0.9865
rs13220495	rs7763910	6	26472655	0.8285
rs13220495	rs9295694	6	26512994	0.9606
rs13220495	rs9295695	6	26528250	0.9528
rs13220495	rs9357010	6	26527945	0.9704
rs13220495	rs9461267	6	26525455	0.9282
rs13220495	rs9461270	6	26544110	0.9772
rs13220495	rs9461271	6	26554968	0.9099
rs13220495	rs9461272	6	26579648	0.8693
rs13220495	rs9467703	6	26318903	0.7086
rs13220495	rs9467774	6	26505036	0.9394
rs13220495	rs9467779	6	26536687	0.9728
rs13220495	rs9467782	6	26542773	0.9849
rs13220495	rs9467791	6	26562486	0.9554
rs13220495	rs9467798	6	26575697	0.8989
rs13220495	rs9467800	6	26578525	0.8575
rs13220495	rs9467804	6	26583129	0.8453
rs13220495	rs9986382	6	26550619	0.9501
rs17526722	rs2179152	6	26325888	0.4994
rs17526722	rs4145910	6	26313305	0.2592
rs17526722	rs6939048	6	26327953	0.8447
rs17526722	rs766406	6	26319588	0.9952
rs17526722	rs766407	6	26319534	0.6820
rs17526722	rs9467703	6	26318903	0.9863
rs2232423	rs13204012	6	28201531	0.9815
rs2232423	rs13205211	6	28203056	0.9889
rs2232423	rs13208096	6	28225311	0.9736
rs2232423	rs13213152	6	28349698	0.9634
rs2232423	rs13213986	6	28358009	0.7324
rs2232423	rs13214023	6	28332141	0.8535
rs2232423	rs13217619	6	28306671	0.9915
rs2232423	rs2232423	6	28366151	0.2047
rs2232423	rs2232426	6	28360659	0.4072
rs2232423	rs2232429	6	28359632	0.5948
rs537160	rs1150754	6	32050758	1.0000
rs537160	rs1150755	6	32038550	0.4307
rs537160	rs1150758	6	32028149	0.7368
rs215614	rs1014242	7	32272305	0.9693
rs215614	rs10226228	7	32315613	0.9865
rs215614	rs10233045	7	32264492	0.9895
rs215614	rs10236197	7	32291761	0.9846
rs215614	rs10237329	7	32265725	0.9315
rs215614	rs10259431	7	32281397	0.9596
rs215614	rs10264177	7	32370862	0.9265
rs215614	rs1450869	7	32311672	0.9824
rs215614	rs215600	7	32333642	0.8983
rs215614	rs215605	7	32336965	0.8264
rs215614	rs215611	7	32341438	0.9407
rs215614	rs215614	7	32347335	0.5231
rs215614	rs215622	7	32357659	0.8705
rs215614	rs215625	7	32358313	0.8512
rs215614	rs215629	7	32360464	0.9630

Table S3 *(continued)*

Table S3 *(continued)*

Index SNP	99% credible-set SNPs	Chromosome	Position	CumSum
rs215614	rs215632	7	32368524	0.7112
rs215614	rs215634	7	32369148	0.9449
rs215614	rs215639	7	32373639	0.8889
rs215614	rs215669	7	32378979	0.9489
rs215614	rs215670	7	32379218	0.9069
rs215614	rs215695	7	32397908	0.9880
rs215614	rs215696	7	32398028	0.9364
rs215614	rs215697	7	32398041	0.9202
rs215614	rs4368879	7	32307925	0.9775
rs215614	rs4723147	7	32398156	0.9662
rs215614	rs6955346	7	32369553	0.9138
rs215614	rs7780515	7	32305274	0.9561
rs215614	rs7798739	7	32292961	0.9801
rs215614	rs7806224	7	32273107	0.9909
rs215614	rs7806397	7	32303339	0.9748
rs215614	rs929456	7	32293644	0.9721
rs215614	rs9771228	7	32322496	0.9526
rs10156445	rs10115938	9	128628571	0.2152
rs10156445	rs10118570	9	128476464	0.4374
rs10156445	rs10121853	9	128550854	0.6120
rs10156445	rs10156445	9	128617244	0.0607
rs10156445	rs10441758	9	128496337	0.7878
rs10156445	rs10513454	9	128673226	0.9182
rs10156445	rs10513456	9	128719849	0.9378
rs10156445	rs10739664	9	128684577	0.8977
rs10156445	rs10739665	9	128729977	0.9799
rs10156445	rs10760400	9	128524269	0.7591
rs10156445	rs10760401	9	128538841	0.7250
rs10156445	rs10760403	9	128608495	0.3092
rs10156445	rs10760404	9	128653015	0.1146
rs10156445	rs10760405	9	128711455	0.9871
rs10156445	rs10819079	9	128583899	0.7425
rs10156445	rs10819081	9	128629174	0.8119
rs10156445	rs10819082	9	128645617	0.4763
rs10156445	rs10819087	9	128695180	0.9518
rs10156445	rs10819089	9	128720289	0.9451
rs10156445	rs10986778	9	128263417	0.9291
rs10156445	rs10986780	9	128265664	0.9404
rs10156445	rs10986936	9	128569777	0.7999
rs10156445	rs10986965	9	128595268	0.5835
rs10156445	rs10986983	9	128620009	0.6621
rs10156445	rs10987017	9	128668715	0.9048
rs10156445	rs10987043	9	128689740	0.9322
rs10156445	rs10987054	9	128705227	0.9814
rs10156445	rs10987055	9	128707568	0.9351
rs10156445	rs1105727	9	128725543	0.9625
rs10156445	rs11789020	9	128616376	0.8550
rs10156445	rs11789188	9	128705012	0.9828
rs10156445	rs11791242	9	128705189	0.9768
rs10156445	rs11999260	9	128521473	0.8236
rs10156445	rs12336219	9	128732275	0.9641
rs10156445	rs12341976	9	128546359	0.8646
rs10156445	rs12345427	9	128732475	0.9557
rs10156445	rs12353435	9	128709253	0.9474
rs10156445	rs12552782	9	128564325	0.9101
rs10156445	rs12553980	9	128646863	0.8899
rs10156445	rs12686660	9	128335467	0.9913
rs10156445	rs13293667	9	128601761	0.8452
rs10156445	rs13299979	9	128625923	0.7063
rs10156445	rs1411350	9	128670179	0.7744

Table S3 *(continued)*

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S3 *(continued)*

Index SNP	99% credible-set SNPs	Chromosome	Position	CumSum
rs10156445	rs1411352	9	128655115	0.1664
rs10156445	rs1477147	9	128736511	0.9538
rs10156445	rs1952669	9	128660676	0.6372
rs10156445	rs2026134	9	128680148	0.9428
rs10156445	rs2041943	9	128717131	0.9885
rs10156445	rs2058850	9	128731818	0.9843
rs10156445	rs2111358	9	128724683	0.9722
rs10156445	rs2149992	9	128738591	0.9591
rs10156445	rs2149993	9	128702846	0.9574
rs10156445	rs2416983	9	128730959	0.9658
rs10156445	rs3793622	9	128727170	0.9737
rs10156445	rs3829098	9	128727035	0.9753
rs10156445	rs4258094	9	128586487	0.3978
rs10156445	rs4515655	9	128616073	0.8347
rs10156445	rs4837022	9	128475379	0.5146
rs10156445	rs4838296	9	128591530	0.8733
rs10156445	rs4838298	9	128629488	0.6857
rs10156445	rs6478712	9	128514608	0.8817
rs10156445	rs7024493	9	128712242	0.9857
rs10156445	rs7026855	9	128724831	0.9784
rs10156445	rs7849076	9	128640702	0.2625
rs10156445	rs7849781	9	128736180	0.9220
rs10156445	rs7859110	9	128709600	0.9259
rs10156445	rs7862061	9	128731930	0.9496
rs10156445	rs7864769	9	128736041	0.9142
rs10156445	rs7868182	9	128721311	0.9899
rs10156445	rs7868811	9	128729746	0.9674
rs10156445	rs7869867	9	128675819	0.3543
rs10156445	rs7871599	9	128721828	0.9608
rs10156445	rs7875710	9	128597392	0.5509
rs10156445	rs872524	9	128723819	0.9706
rs10156445	rs888230	9	128724304	0.9690
rs9328534	rs10901120	9	134905524	0.7423
rs9328534	rs4145638	9	134880190	0.8844
rs9328534	rs4246173	9	134877278	0.4441
rs9328534	rs4246175	9	134930808	0.6184
rs9328534	rs4266768	9	134892125	0.6833
rs9328534	rs4287057	9	134898653	0.8353
rs9328534	rs4292819	9	134929122	0.5960
rs9328534	rs4363310	9	134890430	0.9603
rs9328534	rs4382592	9	134870755	0.2118
rs9328534	rs4402000	9	134891688	0.8519
rs9328534	rs4564007	9	134854280	0.9309
rs9328534	rs4592148	9	134917623	0.5725
rs9328534	rs4617289	9	134926958	0.7615
rs9328534	rs4962181	9	134880575	0.9001
rs9328534	rs6597540	9	134856576	0.4713
rs9328534	rs7019796	9	134863453	0.5239
rs9328534	rs7025089	9	134881443	0.3511
rs9328534	rs7025683	9	134865815	0.4142
rs9328534	rs7026534	9	134907263	0.2716
rs9328534	rs7032884	9	134858482	0.3836
rs9328534	rs7039772	9	134865891	0.4981
rs9328534	rs7040224	9	134886837	0.7032
rs9328534	rs7043386	9	134866354	0.3150
rs9328534	rs7048083	9	134858097	0.7806
rs9328534	rs7467596	9	134912678	0.5486
rs9328534	rs7858113	9	134853684	0.9156
rs9328534	rs9328534	9	134874805	0.1280
rs9328534	rs9328536	9	134914385	0.7232
rs9328534	rs9411331	9	134883419	0.8183
rs9328534	rs9411334	9	134896887	0.8684
rs9328534	rs9411335	9	134901224	0.9746
rs9328534	rs9411336	9	134901901	0.7995
rs9328534	rs9411337	9	134902172	0.9849
rs9328534	rs9411340	9	134943989	0.9929
rs9328534	rs9411424	9	134864382	0.6626
rs9328534	rs9411430	9	134907819	0.6407
rs9328534	rs9411433	9	134921485	0.9460
rs773109	rs1873914	12	56379427	0.9812
rs773109	rs705698	12	56384687	0.7284
rs773109	rs705699	12	56384804	0.9857
rs773109	rs705702	12	56390636	0.9653
rs773109	rs773107	12	56369506	0.8518
rs773109	rs773108	12	56369911	0.5324
rs773109	rs773109	12	56374695	0.2923
rs773109	rs773114	12	56379060	0.9741
rs773109	rs877636	12	56480583	0.9902
rs4500831	rs1020169	18	53066079	0.7585
rs4500831	rs12605773	18	53064423	0.9225
rs4500831	rs12607679	18	53059748	0.9417
rs4500831	rs1261097	18	52868977	0.5526
rs4500831	rs1348047	18	53050058	0.8670
rs4500831	rs1440477	18	53086455	0.8405
rs4500831	rs1452789	18	53115879	0.9888
rs4500831	rs1623427	18	53046319	0.8537
rs4500831	rs1788019	18	53050523	0.8905
rs4500831	rs1788025	18	53048678	0.9670
rs4500831	rs1942262	18	52873317	0.4951
rs4500831	rs2872041	18	53064491	0.7150
rs4500831	rs2919450	18	53084545	0.7757
rs4500831	rs2919451	18	53084300	0.9815
rs4500831	rs2924321	18	53125435	0.4291
rs4500831	rs2958163	18	53085412	0.6440
rs4500831	rs2958166	18	53087632	0.9120
rs4500831	rs2958169	18	53089138	0.6807
rs4500831	rs2958171	18	53072832	0.9594
rs4500831	rs2958175	18	53074958	0.7927
rs4500831	rs2958178	18	53069972	0.9322
rs4500831	rs2958186	18	53077795	0.6038
rs4500831	rs3794889	18	53061937	0.9510
rs4500831	rs3794891	18	53055212	0.9012
rs4500831	rs4374254	18	53092547	0.2546
rs4500831	rs4468713	18	53104019	0.9908
rs4500831	rs4500831	18	53097544	0.1278
rs4500831	rs4524013	18	53096708	0.8094
rs4500831	rs4801149	18	53063229	0.8257
rs4500831	rs4801150	18	53065110	0.7398
rs4500831	rs624244	18	53183396	0.9899
rs4500831	rs7228159	18	53104253	0.3461
rs4500831	rs9320010	18	53053897	0.9744
rs4500831	rs9320016	18	53095471	0.8791
rs4500831	rs9950000	18	53052169	0.9871

SNPs, single nucleotide polymorphisms; CPASSOC, cross-phenotypic association; GERD, gastroesophageal reflux disease; LC, lung cancer.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S4 List of SNPs in the 99% credible set identified from fine-mapping analysis for each CPASSOC-identified locus shared between GERD and LUAD

Index SNP	99% credible-set SNPs	Chromosome	Position	CumSum
rs6695572	rs10782651	1	77934335	0.8769
rs6695572	rs10873947	1	77986516	0.9825
rs6695572	rs11162350	1	77929479	0.8468
rs6695572	rs11162351	1	77944732	0.8905
rs6695572	rs11806197	1	77964855	0.9178
rs6695572	rs11811611	1	77982017	0.9619
rs6695572	rs12040471	1	77968401	0.9000
rs6695572	rs12042177	1	77986638	0.9554
rs6695572	rs12042881	1	77968443	0.9089
rs6695572	rs12047928	1	77932750	0.9675
rs6695572	rs12049202	1	77967523	0.9780
rs6695572	rs12093263	1	77964339	0.9899
rs6695572	rs12729914	1	77980235	0.9341
rs6695572	rs17384946	1	77978205	0.9260
rs6695572	rs1874819	1	77912969	0.9915
rs6695572	rs2088518	1	77951330	0.8132
rs6695572	rs2647506	1	77914159	0.9864
rs6695572	rs3113637	1	77927144	0.9845
rs6695572	rs6603950	1	77938087	0.9420
rs6695572	rs6695572	1	77945635	0.3824
rs6695572	rs6698295	1	77945965	0.7380
rs6695572	rs6704141	1	77932095	0.9803
rs6695572	rs7514937	1	77981099	0.9882
rs6695572	rs7522356	1	77951460	0.9730
rs6695572	rs7542588	1	77938165	0.9488
rs6695572	rs9324162	1	77954804	0.8632
rs17391694	rs17391694	1	78623626	0.9940
rs6711584	rs17343925	2	104426377	0.6405
rs6711584	rs2678670	2	104469564	0.9991
rs6711584	rs2945452	2	104469422	0.8904
rs6711584	rs4347858	2	104423042	0.4785
rs6711584	rs6707445	2	104420858	0.7798
rs6711584	rs6711584	2	104421692	0.2622
rs537160	rs1052248	6	31556581	0.9846
rs537160	rs1799964	6	31542308	0.7980
rs537160	rs2072633	6	31919578	0.6528
rs537160	rs2534671	6	31465661	0.9877
rs537160	rs2844480	6	31564821	0.9196
rs537160	rs3093661	6	31543758	0.9729
rs537160	rs3093668	6	31546495	0.8591
rs537160	rs3828917	6	31465917	0.7326
rs537160	rs389883	6	31947460	0.9799
rs537160	rs537160	6	31916400	0.5421
rs537160	rs589428	6	31848220	0.9903

SNPs, single nucleotide polymorphisms; CPASSOC, cross-phenotypic association; GERD, gastroesophageal reflux disease; LUAD, lung adenocarcinoma.

Table S5 List of SNPs in the 99% credible set identified from fine-mapping analysis for each CPASSOC-identified locus shared between GERD and LUSC

Index SNP	99% credible-set SNPs	Chromosome	Position	CumSum
rs2782641	rs1013793	1	44207553	0.9865
rs2782641	rs10789442	1	44140075	0.9807
rs2782641	rs10890281	1	44245939	0.9823
rs2782641	rs11210860	1	43982527	0.9116
rs2782641	rs11210869	1	44026040	0.8064
rs2782641	rs11577403	1	43989773	0.9005
rs2782641	rs11805774	1	44152715	0.8892
rs2782641	rs12410155	1	44188465	0.9211
rs2782641	rs1472661	1	44209075	0.9878
rs2782641	rs17371903	1	44070691	0.8279
rs2782641	rs17401357	1	44153619	0.8602
rs2782641	rs1887402	1	44036085	0.7272
rs2782641	rs2152113	1	43983569	0.7607
rs2782641	rs2270972	1	44158129	0.8449
rs2782641	rs2274465	1	44121557	0.9300
rs2782641	rs2782640	1	44009033	0.2575
rs2782641	rs2782641	1	44013355	0.1587
rs2782641	rs2819333	1	44014573	0.6930
rs2782641	rs2819334	1	44014735	0.6575
rs2782641	rs2842185	1	44019731	0.3231
rs2782641	rs2842187	1	44014949	0.4322
rs2782641	rs2842188	1	44014280	0.4825
rs2782641	rs3791034	1	44145130	0.9702
rs2782641	rs3791035	1	44154479	0.8748
rs2782641	rs3791040	1	44202733	0.9770
rs2782641	rs3791043	1	44219546	0.9851
rs2782641	rs3791101	1	44366250	0.9905
rs2782641	rs3791134	1	44049156	0.3796
rs2782641	rs3791136	1	44049947	0.5317
rs2782641	rs4660257	1	44148168	0.7838
rs2782641	rs4660740	1	44148617	0.9791
rs2782641	rs489319	1	44131794	0.9381
rs2782641	rs501299	1	44051834	0.5753
rs2782641	rs605709	1	44058467	0.6168
rs2782641	rs618678	1	44133299	0.9460
rs2782641	rs6429637	1	44198531	0.9749
rs2782641	rs660899	1	44117006	0.9531
rs2782641	rs6673970	1	44199469	0.9726
rs2782641	rs673253	1	44062154	0.9664
rs2782641	rs7516647	1	44237465	0.9838
rs2782641	rs7520053	1	44226657	0.9892
rs2782641	rs9787076	1	44141149	0.9597
rs329122	rs10056247	5	133898136	0.9910
rs329122	rs11242219	5	133848760	0.9688
rs329122	rs1476096	5	133837021	0.8845
rs329122	rs1981627	5	133838180	0.8515
rs329122	rs2241699	5	133899872	0.9809
rs329122	rs329117	5	133860101	0.8101
rs329122	rs329120	5	133861756	0.5201
rs329122	rs329122	5	133864599	0.2687
rs329122	rs329124	5	133865452	0.7620
rs329122	rs4958241	5	133829990	0.9749
rs329122	rs4958244	5	133845380	0.9564

Table S5 *(continued)*

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S5 *(continued)*

Index SNP	99% credible-set SNPs	Chromosome	Position	CumSum
rs329122	rs6871635	5	133830395	0.9862
rs329122	rs6891328	5	133851526	0.9129
rs329122	rs7701346	5	133828356	0.9390
rs13191445	rs10456045	6	26404958	0.9900
rs13191445	rs1056667	6	26510564	0.9267
rs13191445	rs13191445	6	26015489	0.9723
rs13191445	rs13202688	6	25993469	0.9855
rs13191445	rs13212534	6	25983010	0.9882
rs13191445	rs1321479	6	26501897	0.9491
rs13191445	rs13220495	6	26441640	0.9893
rs13191445	rs1407045	6	26476155	0.3564
rs13191445	rs1624440	6	26433329	0.9209
rs13191445	rs17526722	6	25918855	0.9869
rs13191445	rs1796520	6	26410800	0.9825
rs13191445	rs1796521	6	26421392	0.9760
rs13191445	rs2024970	6	26497520	0.4776
rs13191445	rs2145318	6	26496603	0.1869
rs13191445	rs2179152	6	26325888	0.6592
rs13191445	rs2255070	6	26501777	0.9681
rs13191445	rs3736781	6	26505362	0.9794
rs13191445	rs3736782	6	26505403	0.9381
rs13191445	rs4145910	6	26313305	0.5763
rs13191445	rs6903973	6	26499942	0.9637
rs13191445	rs6926629	6	26499903	0.9544
rs13191445	rs6939048	6	26327953	0.8525
rs13191445	rs742090	6	26415637	0.9150
rs13191445	rs766406	6	26319588	0.9592
rs13191445	rs766407	6	26319534	0.7289
rs13191445	rs7763910	6	26472655	0.7931
rs13191445	rs806973	6	26148326	0.9906
rs13191445	rs9295694	6	26512994	0.9324
rs13191445	rs9467703	6	26318903	0.9084
rs13191445	rs9467774	6	26505036	0.9438
rs9379899	rs10484442	6	26555879	0.9829
rs9379899	rs1056667	6	26510564	0.9864
rs9379899	rs10946834	6	26533664	0.9835
rs9379899	rs12190473	6	27024687	0.8085
rs9379899	rs12215241	6	27023081	0.8554
rs9379899	rs12215773	6	27039233	0.8324
rs9379899	rs12526680	6	26550954	0.9811
rs9379899	rs1321479	6	26501897	0.9901
rs9379899	rs1407045	6	26476155	0.8895
rs9379899	rs1570059	6	26573325	0.9712
rs9379899	rs1570060	6	26573562	0.9704
rs9379899	rs1624440	6	26433329	0.9847
rs9379899	rs1884946	6	26545308	0.9779
rs9379899	rs2024970	6	26497520	0.9171
rs9379899	rs2145318	6	26496603	0.8733
rs9379899	rs2179152	6	26325888	0.9345
rs9379899	rs3736782	6	26505403	0.9880
rs9379899	rs3922717	6	27030924	0.5104
rs9379899	rs4145910	6	26313305	0.9265
rs9379899	rs4713006	6	26519872	0.9858
rs9379899	rs4713008	6	26538268	0.9758
rs9379899	rs4871	6	26545632	0.9852
rs9379899	rs6456733	6	26566804	0.9694

Table S5 *(continued)*

Table S5 *(continued)*

Index SNP	99% credible-set SNPs	Chromosome	Position	CumSum
rs9379899	rs6456735	6	26574149	0.9685
rs9379899	rs6918360	6	26577867	0.9655
rs9379899	rs6918506	6	26577857	0.9645
rs9379899	rs6918854	6	26577924	0.9721
rs9379899	rs6922824	6	26553815	0.9751
rs9379899	rs6930120	6	26555484	0.9773
rs9379899	rs6933176	6	26540178	0.9817
rs9379899	rs6939048	6	26327953	0.9529
rs9379899	rs6940188	6	26562029	0.9786
rs9379899	rs6940638	6	27046250	0.9055
rs9379899	rs6941022	6	26553531	0.9729
rs9379899	rs742090	6	26415637	0.9799
rs9379899	rs766407	6	26319534	0.9411
rs9379899	rs767471	6	26557854	0.9765
rs9379899	rs7763910	6	26472655	0.9473
rs9379899	rs9295694	6	26512994	0.9874
rs9379899	rs9295695	6	26528250	0.9841
rs9379899	rs9357010	6	26527945	0.9896
rs9379899	rs9461267	6	26525455	0.9792
rs9379899	rs9461270	6	26544110	0.9891
rs9379899	rs9461271	6	26554968	0.9737
rs9379899	rs9461272	6	26579648	0.9675
rs9379899	rs9467703	6	26318903	0.9583
rs9379899	rs9467774	6	26505036	0.9885
rs9379899	rs9467779	6	26536687	0.9869
rs9379899	rs9467791	6	26562486	0.9805
rs9379899	rs9467798	6	26575697	0.9744
rs9379899	rs9467800	6	26578525	0.9665
rs9379899	rs9467804	6	26583129	0.9634
rs9379899	rs9986382	6	26550619	0.9823
rs3922717	rs10484442	6	26555879	0.9884
rs3922717	rs10946834	6	26533664	0.9891
rs3922717	rs12190473	6	27024687	0.8678
rs3922717	rs12215241	6	27023081	0.9181
rs3922717	rs12215773	6	27039233	0.8935
rs3922717	rs12526680	6	26550954	0.9865
rs3922717	rs13191227	6	27390115	0.9520
rs3922717	rs13195040	6	27413924	0.9628
rs3922717	rs13196692	6	27379119	0.9680
rs3922717	rs13207689	6	27369704	0.9464
rs3922717	rs1570059	6	26573325	0.9764
rs3922717	rs1570060	6	26573562	0.9755
rs3922717	rs1884946	6	26545308	0.9844
rs3922717	rs3922717	6	27030924	0.5479
rs3922717	rs4713008	6	26538268	0.9822
rs3922717	rs4871	6	26545632	0.9897
rs3922717	rs6456733	6	26566804	0.9745
rs3922717	rs6456735	6	26574149	0.9734
rs3922717	rs6904596	6	27491299	0.9790
rs3922717	rs6918360	6	26577867	0.9702
rs3922717	rs6918506	6	26577857	0.9691
rs3922717	rs6918854	6	26577924	0.9773
rs3922717	rs6922824	6	26553815	0.9814
rs3922717	rs6930120	6	26555484	0.9837
rs3922717	rs6933176	6	26540178	0.9872
rs3922717	rs6940188	6	26562029	0.9851
rs3922717	rs6940638	6	27046250	0.9354
rs3922717	rs6941022	6	26553531	0.9782
rs3922717	rs767471	6	26557854	0.9829

Table S5 *(continued)*

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S5 *(continued)*

Index SNP	99% credible-set SNPs	Chromosome	Position	CumSum
rs3922717	rs9461271	6	26554968	0.9798
rs3922717	rs9461272	6	26579648	0.9724
rs3922717	rs9467779	6	26536687	0.9903
rs3922717	rs9467791	6	26562486	0.9858
rs3922717	rs9467798	6	26575697	0.9807
rs3922717	rs9467800	6	26578525	0.9713
rs3922717	rs9467804	6	26583129	0.9574
rs3922717	rs9986382	6	26550619	0.9878
rs13219181	rs12190473	6	27024687	0.8972
rs13219181	rs12215241	6	27023081	0.9492
rs13219181	rs12215773	6	27039233	0.9237
rs13219181	rs13191227	6	27390115	0.9842
rs13219181	rs13195040	6	27413924	0.9897
rs13219181	rs13196692	6	27379119	0.9951
rs13219181	rs13207689	6	27369704	0.9785
rs13219181	rs3922717	6	27030924	0.5664
rs13219181	rs6940638	6	27046250	0.9670
rs200968	rs13195291	6	28169241	0.9352
rs200968	rs13197574	6	28060239	0.9444
rs200968	rs13197633	6	28174757	0.9249
rs200968	rs13199649	6	27868792	0.9910
rs200968	rs13200214	6	28017250	0.9881
rs200968	rs13201308	6	28130089	0.9535
rs200968	rs13204012	6	28201531	0.8717
rs200968	rs13205211	6	28203056	0.9029
rs200968	rs13205911	6	28124114	0.9849
rs200968	rs13207689	6	27369704	0.9814
rs200968	rs13208096	6	28225311	0.8393
rs200968	rs13213152	6	28349698	0.5590
rs200968	rs13213986	6	28358009	0.3121
rs200968	rs13214023	6	28332141	0.8019
rs200968	rs13217619	6	28306671	0.9140
rs200968	rs200965	6	27866384	0.9607
rs200968	rs200968	6	27859568	0.9653
rs200968	rs200974	6	27855845	0.9737
rs200968	rs200975	6	27855625	0.9776
rs200968	rs200979	6	27852357	0.9698
rs2232426	rs13195291	6	28169241	0.9838
rs2232426	rs13197574	6	28060239	0.9879
rs2232426	rs13197633	6	28174757	0.9791
rs2232426	rs13201308	6	28130089	0.9921
rs2232426	rs13204012	6	28201531	0.9550
rs2232426	rs13205211	6	28203056	0.9691
rs2232426	rs13208096	6	28225311	0.9403
rs2232426	rs13213152	6	28349698	0.8131
rs2232426	rs13213986	6	28358009	0.7011
rs2232426	rs13214023	6	28332141	0.9233
rs2232426	rs13217619	6	28306671	0.9742
rs2232426	rs2232423	6	28366151	0.3781
rs2232426	rs2232426	6	28360659	0.1932
rs2232426	rs2232429	6	28359632	0.5595

SNPs, single nucleotide polymorphisms; CPASSOC, cross-phenotypic association; GERD, gastroesophageal reflux disease; LUSC, lung squamous cell carcinoma.

Table S6 List of SNPs in the 99% credible set identified from fine-mapping analysis for each CPASSOC-identified locus shared between GERD and SCLC

Index SNP	99% credible-set SNPs	Chromosome	Position	CumSum
rs3172494	rs13084037	3	49214066	0.9005
rs3172494	rs7617480	3	49210732	0.5461
rs3172494	rs9586	3	49213637	0.9997
rs2232423	rs13195291	6	28169241	0.9747
rs2232423	rs13197574	6	28060239	0.9875
rs2232423	rs13197633	6	28174757	0.9813
rs2232423	rs13201308	6	28130089	0.9935
rs2232423	rs13204012	6	28201531	0.9462
rs2232423	rs13205211	6	28203056	0.9596
rs2232423	rs13208096	6	28225311	0.9151
rs2232423	rs13213152	6	28349698	0.8969
rs2232423	rs13213986	6	28358009	0.6773
rs2232423	rs13214023	6	28332141	0.7902
rs2232423	rs13217619	6	28306671	0.9681
rs2232423	rs200965	6	27866384	0.9324
rs2232423	rs2232423	6	28366151	0.1908
rs2232423	rs2232426	6	28360659	0.3752
rs2232423	rs2232429	6	28359632	0.5484

SNPs, single nucleotide polymorphisms; CPASSOC, cross-phenotypic association; GERD, gastroesophageal reflux disease; SCLC, small-cell lung cancer.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S7 Results from colocalization analysis for each pleiotropic locus identified from CPASSOC

Index SNP	Chromosome	Position	Gastroesophageal reflux disease		Lung cancer		$P_{CPASSOC}$	PPH3.abf	PPH4.abf
			Beta	P value	Beta	P value			
Gastroesophageal reflux disease and lung cancer									
rs17391694	1	78623626	−0.035685568	2.54E−07	−0.106292	2.62E−08	7.94E−10	0.000	0.999
rs2782641	1	44013355	−0.027088204	4.33E−08	−0.041668	6.26E−04	5.67E−10	0.040	0.592
rs6711584	2	104421692	−0.032254466	2.66E−11	−0.044136	1.52E−04	2.13E−13	0.028	0.848
rs329122	5	133864599	0.028952886	3.05E−09	0.050511	1.69E−05	2.35E−11	0.004	0.970
rs13207689	6	27369704	0.046620595	9.32E−10	−0.139724	9.29E−11	1.35E−14	0.836	0.164
rs13220495	6	26441640	0.043286473	1.96E−08	−0.126045	7.74E−09	2.91E−12	1.000	0.000
rs17526722	6	25918855	0.032422152	5.59E−05	−0.131428	1.26E−08	1.58E−08	1.000	0.000
rs2232423	6	28366151	0.050652994	1.37E−11	−0.14533	8.04E−12	2.54E−17	0.009	0.991
rs537160	6	31916400	−0.027209424	5.08E−08	0.052519	3.98E−05	8.46E−10	0.989	0.001
rs215614	7	32347335	0.032854105	4.08E−11	0.042536	4.29E−04	1.31E−13	0.050	0.634
rs10156445	9	128617244	−0.024850116	6.33E−07	−0.040924	7.81E−04	1.51E−08	0.350	0.353
rs9328534	9	134874805	0.029341563	1.35E−08	0.043651	4.67E−04	1.25E−10	0.034	0.539
rs773109	12	56374695	0.038057215	8.71E−14	0.04369	5.14E−04	5.40E−16	0.008	0.736
rs4500831	18	53097544	0.028003873	1.21E−07	0.045982	3.42E−04	1.47E−09	0.049	0.191
Gastroesophageal reflux disease and lung adenocarcinoma									
rs6695572	1	77945635	−0.021700617	4.09E−04	−0.12281	8.34E−09	2.14E−08	0.218	0.472
rs17391694	1	78623626	−0.035685568	2.54E−07	−0.135426	3.83E−07	8.34E−09	0.000	0.999
rs6711584	2	104421692	−0.032254466	2.66E−11	−0.067323	2.89E−05	8.68E−13	0.007	0.969
rs537160	6	31916400	−0.027209424	5.08E−08	0.059983	7.89E−04	8.21E−09	0.049	0.160
Gastroesophageal reflux disease and lung squamous cell carcinoma									
rs2782641	1	44013355	−0.027088204	4.33E−08	−0.070961	2.76E−04	4.98E−09	0.077	0.879
rs329122	5	133864599	0.028952886	3.05E−09	0.079537	2.28E−05	5.12E−10	0.006	0.973
rs13191445	6	26015489	0.032472655	5.35E−05	−0.251115	1.06E−11	5.56E−11	1.000	0.000
rs9379899	6	26603015	0.042970721	1.25E−09	−0.112059	2.17E−04	1.07E−10	1.000	0.000
rs3922717	6	27030924	0.040691795	5.35E−13	−0.083107	3.75E−04	3.81E−14	0.916	0.084
rs13219181	6	27136225	0.034914578	1.32E−08	−0.106462	2.56E−05	7.00E−10	0.803	0.197
rs200968	6	27859568	0.0401529	3.94E−11	−0.105147	4.25E−05	1.62E−12	0.022	0.978
rs2232426	6	28360659	0.050647495	1.39E−11	−0.220279	1.02E−10	1.63E−14	0.013	0.987
Gastroesophageal reflux disease and small-cell lung cancer									
rs3172494	3	48731487	0.045352619	6.71E−09	0.153362	9.28E−04	5.12E−09	0.016	0.158
rs2232423	6	28366151	0.050652994	1.37E−11	−0.200395	2.14E−04	2.02E−12	0.016	0.920

CPASSOC, cross-phenotypic association; SNP, single nucleotide polymorphism.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S8 TWAS-identified shared gene-tissue pairs between GERD and LC

Gene	Tissue type	CHR	SNP. Num	Gastroesophageal reflux disease				Lung cancer			
				BEST.GWAS. ID	BEST. GWAS.Z	TWAS.P	FDR	BEST.GWAS. ID	BEST. GWAS.Z	TWAS.P	FDR
FUBP1	Adipose subcutaneous	1	332	rs17391694	5.16	2.54E−07	4.50E−05	rs17391694	5.57	2.62E−08	2.48E−05
ERAP1	Adipose subcutaneous	5	494	rs28129	−4.45	1.08E−04	1.04E−02	rs27039	3.68	5.43E−04	4.47E−02
MAPKAP1	Adipose visceral omentum	9	355	rs12352375	−4.95	2.20E−03	2.96E−02	rs13291556	4.24	2.27E−04	3.45E−02
FUBP1	Adipose visceral omentum	1	332	rs17391694	5.16	2.91E−05	2.52E−03	rs17391694	5.57	3.59E−07	2.63E−04
CEP57	Adipose visceral omentum	11	617	rs568668	−4.58	8.39E−06	1.03E−03	rs10831447	−3.95	1.72E−04	4.01E−02
PBX3	Adipose visceral omentum	9	377	rs12352375	−4.93	4.24E−07	1.16E−04	rs13291556	4.24	4.73E−04	3.45E−02
MTMR2	Artery aorta	11	602	rs568668	−4.58	9.51E−05	8.52E−03	rs10831447	−3.95	3.94E−04	4.39E−02
YIPF4	Artery aorta	2	314	rs176403	3.61	8.05E−04	2.91E−02	rs458628	4.16	2.69E−05	6.16E−03
ZDHHC5	Artery aorta	11	393	rs12790660	4.83	4.72E−06	2.11E−03	rs11229042	−4.88	4.78E−04	4.39E−02
MTMR2	Artery tibial	11	604	rs568668	−4.58	1.45E−04	7.24E−03	rs10831447	−3.95	4.79E−04	3.78E−02
MEGF9	Artery tibial	9	341	rs12377227	−3.40	2.69E−03	3.66E−02	rs1060817	−3.78	3.24E−04	3.95E−02
PTPRF	Artery tibial	1	444	rs2782641	−5.48	7.05E−07	1.52E−04	rs4660740	−3.69	1.22E−04	2.78E−02
FUBP1	Artery tibial	1	332	rs17391694	5.16	2.54E−07	7.28E−05	rs17391694	5.57	2.62E−08	2.39E−05
CEP57	Artery tibial	11	613	rs568668	−4.58	3.89E−04	1.19E−02	rs10831447	−3.95	1.04E−03	4.78E−02
ERAP1	Brain amygdala	5	498	rs28129	−4.45	1.16E−03	4.52E−02	rs27039	3.68	3.53E−04	2.11E−02
MAP2K5	Brain caudate basal ganglia	15	408	rs8025889	4.74	1.47E−03	4.04E−02	rs1001870	−3.65	2.66E−04	5.73E−03
ERAP1	Brain caudate basal ganglia	5	497	rs28129	−4.45	6.90E−04	4.17E−02	rs27039	3.68	2.84E−04	4.17E−02
ENSG00000276517	Brain cerebellar hemisphere	2	378	rs176403	3.61	2.03E−03	4.38E−02	rs458628	4.16	1.08E−04	2.15E−02
FUBP1	Brain cerebellum	1	332	rs17391694	5.16	2.54E−07	8.50E−05	rs17391694	5.57	2.62E−08	1.87E−05
ERAP1	Brain cortex	5	496	rs28129	−4.45	1.25E−03	3.91E−02	rs27039	3.68	3.62E−04	3.18E−02
ENSG00000276334	Brain frontal cortex BA9	2	375	rs176403	3.61	1.10E−03	2.60E−02	rs458628	4.16	6.81E−05	2.13E−02
ERAP1	Brain nucleus accumbens basal ganglia	5	499	rs28129	−4.45	1.18E−03	3.42E−02	rs27039	3.68	3.18E−04	3.14E−02
VPS33B-DT	Brain nucleus accumbens basal ganglia	15	487	rs11073964	3.53	1.01E−03	2.86E−02	rs7601	−3.50	1.76E−04	4.97E−03
MAP2K5	Brain putamen basal ganglia	15	408	rs8025889	4.74	3.66E−07	3.31E−05	rs1001870	−3.65	5.85E−03	4.87E−02
ERAP1	Brain putamen basal ganglia	5	499	rs28129	−4.45	8.01E−04	4.03E−02	rs27039	3.68	2.74E−04	3.64E−02
DISP2	Brain spinal cord cervical c-1	15	396	rs11070293	−4.18	1.92E−03	4.02E−02	rs1992272	−3.99	7.38E−05	2.34E−03
ERAP1	Brain substantia nigra	5	500	rs28129	−4.45	1.17E−03	2.66E−02	rs27039	3.68	1.47E−04	1.41E−02
FUBP1	Breast mammary tissue	1	332	rs17391694	5.16	2.54E−07	1.46E−04	rs17391694	5.57	2.62E−08	1.61E−05
CEP57	Breast mammary tissue	11	613	rs568668	−4.58	1.57E−05	1.50E−03	rs10831447	−3.95	1.02E−04	1.96E−02
TIE1	Cells cultured fibroblasts	1	460	rs2782641	−5.48	3.92E−05	7.10E−03	rs4660740	−3.69	1.35E−04	3.28E−02
MEPCE	Cells cultured fibroblasts	7	311	rs12532238	3.52	7.58E−04	3.63E−02	rs314370	3.32	1.56E−05	8.38E−03
PBX3	Cells cultured fibroblasts	9	377	rs12352375	−4.93	2.02E−06	7.27E−04	rs13291556	4.24	2.51E−05	9.71E−03
ERAP1	Cells EBV-transformed lymphocytes	5	490	rs28129	−4.45	2.70E−04	2.50E−02	rs27039	3.68	1.70E−04	1.10E−02
CEP57	Cells EBV-transformed lymphocytes	11	614	rs568668	−4.58	8.25E−06	1.25E−03	rs10831447	−3.95	2.06E−04	1.95E−02

Table S8 *(continued)*

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S8 *(continued)*

Gene	Tissue type	CHR	SNP. Num	Gastroesophageal reflux disease				Lung cancer			
				BEST.GWAS. ID	BEST. GWAS.Z	TWAS.P	FDR	BEST.GWAS. ID	BEST. GWAS.Z	TWAS.P	FDR
PBX3	Cells EBV-transformed lymphocytes	9	378	rs12352375	−4.93	3.47E−04	6.71E−03	rs13291556	4.24	1.83E−05	1.67E−03
FUBP1	Colon sigmoid	1	331	rs17391694	5.16	1.58E−08	9.07E−06	rs17391694	5.57	1.36E−08	8.24E−06
ERAP1	Colon sigmoid	5	493	rs28129	−4.45	3.96E−04	2.27E−02	rs27039	3.68	2.69E−04	2.04E−02
CEP57	Colon sigmoid	11	614	rs568668	−4.58	9.41E−06	8.42E−04	rs10831447	−3.95	1.32E−04	2.37E−02
PBX3	Colon sigmoid	9	377	rs12352375	−4.93	9.37E−07	1.06E−04	rs13291556	4.24	1.79E−05	4.31E−03
ENSG00000272109	Colon sigmoid	5	493	rs28129	−4.45	7.56E−04	2.81E−02	rs27039	3.68	9.91E−05	1.00E−02
ERAP1	Colon transverse	5	496	rs28129	−4.45	4.08E−05	5.76E−03	rs27039	3.68	2.77E−05	8.37E−03
PBX3	Colon transverse	9	375	rs12352375	−4.93	1.58E−07	3.73E−05	rs13291556	4.24	1.13E−04	1.50E−02
FUBP1	Esophagus gastroesophageal junction	1	331	rs17391694	5.16	2.54E−07	5.02E−05	rs17391694	5.57	2.62E−08	1.65E−05
FUBP1	Esophagus mucosa	1	332	rs17391694	5.16	2.54E−07	5.55E−05	rs17391694	5.57	2.62E−08	2.41E−05
ERAP1	Esophagus mucosa	5	493	rs28129	−4.45	1.50E−04	9.75E−03	rs27039	3.68	7.76E−05	7.95E−03
ENSG00000272109	Esophagus mucosa	5	493	rs28129	−4.45	1.48E−03	4.96E−02	rs27039	3.68	3.94E−04	3.23E−02
FUBP1	Esophagus muscularis	1	332	rs17391694	5.16	5.47E−07	1.54E−04	rs17391694	5.57	8.49E−10	7.52E−07
CEP57	Esophagus muscularis	11	613	rs568668	−4.58	1.30E−05	1.68E−03	rs10831447	−3.95	1.92E−04	3.99E−02
ZDHHC5	Heart atrial appendage	11	393	rs12790660	4.83	2.08E−06	3.92E−04	rs11229042	−4.88	1.82E−04	3.46E−02
MED19	Heart atrial appendage	11	386	rs12790660	4.83	6.09E−06	5.74E−04	rs11229042	−4.88	6.28E−04	4.77E−02
FUBP1	Heart atrial appendage	1	332	rs17391694	5.16	2.84E−08	8.89E−06	rs17391694	5.57	1.15E−08	7.59E−06
CEP57	Heart atrial appendage	11	613	rs568668	−4.58	3.22E−04	1.01E−02	rs10831447	−3.95	4.64E−04	4.41E−02
PBX3	Heart atrial appendage	9	375	rs12352375	−4.93	1.11E−06	2.79E−04	rs13291556	4.24	2.79E−04	3.79E−02
SKOR1	Heart atrial appendage	15	350	rs8025889	4.98	5.39E−04	2.07E−02	rs1001870	−3.65	9.09E−04	2.08E−02
MTMR2	Heart left ventricle	11	603	rs568668	−4.58	1.13E−03	2.84E−02	rs10831447	−3.95	8.81E−04	3.22E−02
ZDHHC5	Heart left ventricle	11	393	rs12790660	4.83	4.72E−06	5.13E−04	rs11229042	−4.88	4.78E−04	2.62E−02
MED19	Heart left ventricle	11	386	rs12790660	4.83	4.02E−06	5.13E−04	rs11229042	−4.88	7.02E−04	3.15E−02
ERAP1	Heart left ventricle	5	492	rs28129	−4.45	7.55E−04	2.82E−02	rs27039	3.68	6.56E−05	1.66E−02
CEP57	Heart left ventricle	11	612	rs568668	−4.58	8.02E−06	5.25E−04	rs10831447	−3.95	1.69E−04	1.39E−02
FAM76B	Liver	11	614	rs568668	−4.58	8.26E−06	9.09E−04	rs10831447	−3.95	1.64E−04	1.21E−02
ERAP1	Liver	5	495	rs28129	−4.45	6.56E−04	2.38E−02	rs27039	3.68	3.51E−04	2.71E−02
MTMR2	Lung	11	602	rs568668	−4.58	1.89E−05	1.55E−03	rs10831447	−3.95	1.47E−04	2.04E−02
MAPKAP1	Lung	9	355	rs12352375	−4.95	1.43E−03	2.48E−02	rs13291556	4.24	1.15E−04	3.78E−02
FUBP1	Lung	1	332	rs17391694	5.16	2.54E−07	6.68E−05	rs17391694	5.57	2.62E−08	2.20E−05
ERAP1	Lung	5	495	rs28129	−4.45	2.71E−04	2.70E−02	rs27039	3.68	1.35E−04	1.88E−02
CEP57	Lung	11	613	rs568668	−4.58	9.16E−06	1.20E−03	rs10831447	−3.95	1.65E−04	2.04E−02
DISP2	Minor salivary gland	15	392	rs11070293	−4.18	1.10E−03	2.57E−02	rs1992272	−3.99	1.30E−04	3.80E−03
ERAP1	Minor salivary gland	5	491	rs28129	−4.45	2.15E−04	1.15E−02	rs27039	3.68	8.89E−04	3.44E−02
TRIM38	Muscle skeletal	6	666	rs1614887	7.05	1.24E−05	6.57E−04	rs2393592	6.48	4.63E−05	5.73E−03
FUBP1	Muscle skeletal	1	332	rs17391694	5.16	2.54E−07	6.71E−05	rs17391694	5.57	2.62E−08	2.19E−05
HYKK	Muscle skeletal	15	422	rs34016249	3.31	4.58E−04	1.54E−02	rs12914385	21.45	3.12E−26	5.19E−24
FUBP1	Nerve tibial	1	332	rs17391694	5.16	5.26E−07	1.78E−04	rs17391694	5.57	1.58E−09	1.71E−06
CEP57	Nerve tibial	11	616	rs568668	−4.58	3.40E−05	3.62E−03	rs10831447	−3.95	1.97E−04	2.53E−02
TMX2	Nerve tibial	11	386	rs12790660	4.83	1.09E−05	1.39E−03	rs11229042	−4.88	1.49E−04	2.53E−02
MTMR2	Ovary	11	604	rs568668	−4.58	1.22E−05	2.07E−03	rs10831447	−3.95	1.44E−04	2.46E−02
TRIM38	Pancreas	6	661	rs1614887	7.04	5.71E−04	1.99E−02	rs2393592	6.48	9.69E−04	3.97E−02
FUBP1	Pancreas	1	332	rs17391694	5.16	3.11E−06	3.09E−04	rs17391694	5.57	3.88E−10	2.06E−07
ERAP1	Pancreas	5	492	rs28129	−4.45	2.24E−04	8.19E−03	rs27039	3.68	1.23E−04	1.64E−02
LINC02609	Pituitary	1	415	rs12089815	4.60	9.98E−04	4.30E−02	rs1342780	−4.28	1.84E−05	1.03E−02

Table S8 *(continued)*

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S8 *(continued)*

Gene	Tissue type	CHR	SNP. Num	Gastroesophageal reflux disease				Lung cancer			
				BEST.GWAS. ID	BEST. GWAS.Z	TWAS.P	FDR	BEST.GWAS. ID	BEST. GWAS.Z	TWAS.P	FDR
ERAP1	Prostate	5	495	rs28129	−4.45	1.30E−03	3.75E−02	rs27039	3.68	1.67E−05	3.11E−03
CEP57	Prostate	11	616	rs568668	−4.58	4.41E−05	5.36E−03	rs10831447	−3.95	7.85E−05	1.92E−02
ENSG00000272109	Prostate	5	495	rs28129	−4.45	2.31E−04	1.78E−02	rs27039	3.68	5.41E−04	2.81E−02
ENSG00000276334	Prostate	2	375	rs176403	3.61	7.00E−04	2.94E−02	rs458628	4.16	5.81E−05	1.91E−02
FAM76B	Skin not sun exposed suprapubic	11	618	rs568668	−4.58	6.64E−06	1.74E−03	rs10831447	−3.95	2.78E−04	2.75E−02
PRPF6	Skin not sun exposed suprapubic	20	371	rs6011118	−4.50	1.53E−04	9.75E−03	rs7264220	−4.71	6.25E−05	1.61E−02
FUBP1	Skin not sun exposed suprapubic	1	332	rs17391694	5.16	2.54E−07	7.50E−05	rs17391694	5.57	2.62E−08	2.45E−05
CEP57	Skin not sun exposed suprapubic	11	617	rs568668	−4.58	2.54E−04	1.33E−02	rs10831447	−3.95	3.83E−04	2.88E−02
FUBP1	Skin sun exposed lower leg	1	332	rs17391694	5.16	7.73E−11	7.58E−08	rs17391694	5.57	5.28E−10	5.46E−07
CEP57	Skin sun exposed lower leg	11	617	rs568668	−4.58	1.00E−04	6.01E−03	rs10831447	−3.95	1.88E−04	2.85E−02
VPS33B-DT	Skin sun exposed lower leg	15	485	rs11073964	3.53	1.13E−03	2.78E−02	rs7601	−3.50	1.55E−03	2.83E−02
ERAP1	Small intestine terminal ileum	5	490	rs28129	−4.45	1.41E−04	2.06E−02	rs27039	3.68	2.20E−04	8.31E−03
ZDHHC5	Spleen	11	393	rs12790660	4.83	1.86E−06	6.21E−04	rs11229042	−4.88	6.48E−04	4.62E−02
CEP57	Spleen	11	613	rs568668	−4.58	6.92E−06	6.81E−04	rs10831447	−3.95	2.72E−04	4.58E−02
ENSG00000272109	Spleen	5	497	rs28129	−4.45	1.06E−03	3.91E−02	rs27039	3.68	1.68E−04	1.09E−02
DISP2	Stomach	15	398	rs11070293	−4.18	1.10E−03	1.67E−02	rs1992272	−3.99	1.30E−04	4.78E−03
ERAP1	Stomach	5	492	rs28129	−4.45	1.05E−04	8.12E−03	rs27039	3.68	3.94E−05	3.15E−03
CEP57	Stomach	11	612	rs568668	−4.58	8.50E−06	1.21E−03	rs10831447	−3.95	1.65E−04	1.58E−02
SEMA6D	Testis	15	426	rs8034783	4.88	7.16E−05	7.93E−03	rs7165678	5.43	9.18E−05	3.16E−03
DISP2	Testis	15	398	rs11070293	−4.18	1.80E−03	4.98E−02	rs1992272	−3.99	1.23E−04	3.93E−03
ERAP1	Testis	5	494	rs28129	−4.45	1.37E−04	8.74E−03	rs27039	3.68	3.33E−04	3.25E−02
FAM76B	Thyroid	11	616	rs568668	−4.58	9.56E−06	1.18E−03	rs10831447	−3.95	2.98E−04	3.08E−02
SERPING1	Thyroid	11	386	rs12790660	4.83	8.08E−05	4.96E−03	rs11229042	−4.88	1.59E−04	2.47E−02
FUBP1	Thyroid	1	332	rs17391694	5.16	2.54E−07	8.56E−05	rs17391694	5.57	2.62E−08	2.82E−05
CEP57	Thyroid	11	615	rs568668	−4.58	5.50E−06	1.18E−03	rs10831447	−3.95	1.50E−04	2.47E−02
ENSG00000254602	Thyroid	11	397	rs12790660	4.83	1.09E−05	1.18E−03	rs11229042	−4.88	1.24E−04	2.47E−02
N4BP2L2-IT2	Uterus	13	469	rs7988462	−4.49	7.45E−04	1.19E−02	rs9943888	−4.77	6.90E−04	3.31E−02
ENSG00000272109	Vagina	5	492	rs28129	−4.45	2.33E−03	3.65E−02	rs27039	3.68	1.77E−03	3.51E−02
FAM76B	Whole blood	11	618	rs568668	−4.58	6.64E−06	6.10E−04	rs10831447	−3.95	2.78E−04	2.57E−02
SERPING1	Whole blood	11	386	rs12790660	4.83	1.30E−06	2.07E−04	rs11229042	−4.88	1.87E−04	2.16E−02

TWAS, transcriptome-wide association study; GERD, gastroesophageal reflux disease; LC, lung cancer; CHR, chromosome; SNP, single nucleotide polymorphism; GWAS, genome-wide association study; FDR, false discovery rate.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S9 TWAS-identified shared gene-tissue pairs between GERD and LUAD

Gene	Tissue type	CHR	SNP. Num	Gastroesophageal reflux disease				Lung adenocarcinoma			
				BEST.GWAS. ID	BEST. GWAS.Z	TWAS.P	FDR	BEST.GWAS. ID	BEST. GWAS.Z	TWAS.P	FDR
FUBP1	Adipose subcutaneous	1	332	rs17391694	5.16	2.54E−07	4.50E−05	rs6695572	5.76	3.83E−07	3.63E−04
FUBP1	Adipose visceral omentum	1	332	rs17391694	5.16	2.91E−05	2.52E−03	rs6695572	5.76	1.55E−08	1.14E−05
FUBP1	Artery tibial	1	332	rs17391694	5.16	2.54E−07	7.28E−05	rs6695572	5.76	3.83E−07	3.50E−04
CEP192	Brain amygdala	18	398	rs1786263	-2.92	5.81E−03	4.18E−02	rs1981354	3.45	1.90E−03	3.41E−02
FUBP1	Brain cerebellum	1	332	rs17391694	5.16	2.54E−07	8.50E−05	rs6695572	5.76	3.83E−07	2.73E−04
FUBP1	Breast mammary tissue	1	332	rs17391694	5.16	2.54E−07	1.46E−04	rs6695572	5.76	3.83E−07	2.36E−04
MEPCE	Cells cultured fibroblasts	7	311	rs12532238	3.52	7.58E−04	3.63E−02	rs314311	3.41	4.44E−05	2.38E−02
FUBP1	Colon sigmoid	1	331	rs17391694	5.16	1.58E−08	9.07E−06	rs6695572	5.76	5.48E−07	3.32E−04
FUBP1	Esophagus gastroesophageal junction	1	331	rs17391694	5.16	2.54E−07	5.02E−05	rs6695572	5.76	3.83E−07	2.43E−04
FAM227B	Esophagus gastroesophageal junction	15	364	rs769136	2.74	7.30E−04	1.86E−02	rs10519226	−6.20	2.12E−04	1.10E−02
FUBP1	Esophagus mucosa	1	332	rs17391694	5.16	2.54E−07	5.55E−05	rs6695572	5.76	3.83E−07	3.54E−04
ERAP1	Esophagus mucosa	5	493	rs28129	-4.45	1.50E−04	9.75E−03	rs3853202	2.81	4.78E−04	4.90E−02
FUBP1	Esophagus muscularis	1	332	rs17391694	5.16	5.47E−07	1.54E−04	rs6695572	5.76	1.49E−10	1.32E−07
FUBP1	Heart atrial appendage	1	332	rs17391694	5.16	2.84E−08	8.89E−06	rs6695572	5.76	4.37E−08	2.89E−05
FUBP1	Lung	1	332	rs17391694	5.16	2.54E−07	6.68E−05	rs6695572	5.76	3.83E−07	3.23E−04
FUBP1	Muscle skeletal	1	332	rs17391694	5.16	2.54E−07	6.71E−05	rs6695572	5.76	3.83E−07	3.21E−04
HYKK	Muscle skeletal	15	422	rs34016249	3.31	4.58E−04	1.54E−02	rs1051730	14.48	5.88E−12	9.79E−10
FUBP1	Nerve tibial	1	332	rs17391694	5.16	5.26E−07	1.78E−04	rs6695572	5.76	5.59E−09	6.04E−06
FUBP1	Pancreas	1	332	rs17391694	5.16	3.11E−06	3.09E−04	rs6695572	5.76	1.73E−07	9.17E−05
ERAP1	Prostate	5	495	rs28129	-4.45	1.30E−03	3.75E−02	rs3853202	2.81	1.25E−04	1.31E−02
PRPF6	Skin not sun exposed suprapubic	20	371	rs6011118	-4.50	1.53E−04	9.75E−03	rs3761121	5.78	1.16E−04	2.98E−02
FUBP1	Skin not sun exposed suprapubic	1	332	rs17391694	5.16	2.54E−07	7.50E−05	rs6695572	5.76	3.83E−07	3.58E−04
FUBP1	Skin sun exposed lower leg	1	332	rs17391694	5.16	7.73E−11	7.58E−08	rs6695572	5.76	1.67E−07	1.73E−04
TNFRSF6B	Skin sun exposed lower leg	20	429	rs6011118	-4.50	2.85E−05	3.16E−03	rs3761121	5.78	1.00E−05	1.40E−03
SLC6A3	Stomach	5	527	rs2550948	-3.22	1.30E−03	3.51E−02	rs2853677	11.75	2.10E−04	1.23E−02
ERAP1	Stomach	5	492	rs28129	-4.45	1.05E−04	8.12E−03	rs3853202	2.82	9.38E−04	3.77E−02
DISP2	Testis	15	398	rs11070293	-4.18	1.80E−03	4.98E−02	rs679882	3.37	1.23E−03	3.93E−02
FUBP1	Thyroid	1	332	rs17391694	5.16	2.54E−07	8.56E−05	rs6695572	5.76	3.83E−07	4.12E−04

TWAS, transcriptome-wide association study; GERD, gastroesophageal reflux disease; LUAD, lung adenocarcinoma; CHR, chromosome; SNP, single nucleotide polymorphism; GWAS, genome-wide association study; FDR, false discovery rate.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S10 TWAS-identified shared gene-tissue pairs between GERD and LUSC

Gene	Tissue type	CHR	SNP.Num	Gastroesophageal reflux disease				Lung squamous cell carcinoma			
				BEST.GWAS.ID	BEST.GWAS.Z	TWAS.P	FDR	BEST.GWAS.ID	BEST.GWAS.Z	TWAS.P	FDR
PBX3	Adipose subcutaneous	9	377	rs12352375	−4.93	1.86E−06	6.05E−04	rs6478712	−4.38	4.72E−05	1.64E−02
PBX3	Adipose visceral omentum	9	377	rs12352375	−4.93	4.24E−07	1.16E−04	rs6478712	−4.38	2.61E−05	7.70E−03
PBX3	Artery aorta	9	377	rs12352375	−4.93	9.87E−07	2.87E−04	rs6478712	−4.38	2.97E−05	9.33E−03
PTPRF	Artery tibial	1	444	rs2782641	−5.48	7.05E−07	1.52E−04	rs4660740	−4.57	9.98E−06	7.78E−03
PBX3	Artery tibial	9	377	rs12352375	−4.93	1.07E−06	1.82E−04	rs6478712	−4.38	2.97E−05	1.09E−02
MAP2K5	Brain caudate basal ganglia	15	408	rs8025889	4.74	1.47E−03	4.04E−02	rs12905397	4.03	1.41E−03	2.74E−02
SZT2	Brain hypothalamus	1	455	rs2782641	−5.48	1.08E−07	3.43E−05	rs4660740	−4.57	1.25E−04	1.42E−02
TMEM125	Brain putamen basal ganglia	1	452	rs2782641	−5.48	1.53E−05	3.64E−03	rs4660740	−4.57	1.83E−05	7.89E−03
PBX3	Breast mammary tissue	9	377	rs12352375	−4.93	2.26E−06	5.18E−04	rs6478712	−4.38	1.39E−04	1.70E−02
TIE1	Cells cultured fibroblasts	1	460	rs2782641	−5.48	3.92E−05	7.10E−03	rs4660740	−4.57	4.75E−05	4.16E−02
PBX3	Cells cultured fibroblasts	9	377	rs12352375	−4.93	2.02E−06	7.27E−04	rs6478712	−4.38	1.90E−05	7.39E−03
PBX3	Cells EBV-transformed lymphocytes	9	378	rs12352375	−4.93	3.47E−04	6.71E−03	rs6478712	−4.38	1.73E−04	1.59E−02
NA	Cells EBV-transformed lymphocytes	1	420	rs596522	−5.48	1.63E−05	1.24E−03	rs4660740	−4.57	1.07E−04	2.59E−02
PBX3	Colon sigmoid	9	377	rs12352375	−4.93	9.37E−07	1.06E−04	rs6478712	−4.38	5.58E−06	1.36E−03
PBX3	Colon transverse	9	375	rs12352375	−4.93	1.58E−07	3.73E−05	rs6478712	−4.38	2.42E−05	6.41E−03
PBX3	Esophagus gastroesophageal junction	9	377	rs12352375	−4.93	1.84E−06	3.64E−04	rs6478712	−4.38	6.57E−05	1.70E−02
SKOR1	Esophagus mucosa	15	341	rs8025889	4.98	3.47E−04	1.47E−02	rs12905397	4.03	3.31E−04	1.26E−02
PBX3	Esophagus muscularis	9	377	rs12352375	−4.93	2.21E−06	6.76E−04	rs6478712	−4.38	9.63E−06	3.20E−03
PBX3	Heart atrial appendage	9	375	rs12352375	−4.93	1.11E−06	2.79E−04	rs6478712	−4.38	7.00E−06	1.90E−03
SKOR1	Heart atrial appendage	15	350	rs8025889	4.98	5.39E−04	2.07E−02	rs12905397	4.03	2.57E−04	1.77E−02
PBX3	Heart left ventricle	9	375	rs12352375	−4.93	2.27E−06	5.02E−04	rs6478712	−4.38	2.18E−05	5.23E−03
PBX3	Lung	9	377	rs12352375	−4.93	2.05E−06	6.40E−04	rs6478712	−4.38	3.28E−05	1.09E−02
HYKK	Muscle skeletal	15	422	rs34016249	3.31	4.58E−04	1.54E−02	rs8040868	13.47	1.46E−12	2.43E−10
TRIM38	Pancreas	6	661	rs1614887	7.04	5.71E−04	1.99E−02	rs13212534	6.62	2.77E−05	6.81E−03
FUBP1	Pancreas	1	332	rs17391694	5.16	3.11E−06	3.09E−04	rs17101224	−3.37	2.74E−05	1.46E−02
PBX3	Pancreas	9	378	rs12352375	−4.93	3.26E−06	6.55E−04	rs6478712	−4.38	1.19E−05	2.75E−03
SKOR1	Pancreas	15	351	rs8025889	4.98	9.68E−04	4.63E−02	rs12905397	4.03	2.82E−04	1.09E−02
SORCS3	Pituitary	10	488	rs1021362	6.17	4.94E−06	1.20E−03	rs4532962	3.68	3.21E−04	4.11E−02
PBX3	Skin not sun exposed suprapubic	9	376	rs12352375	−4.93	2.27E−06	2.40E−04	rs6478712	−4.38	2.18E−05	7.63E−03
SKOR1	Skin not sun exposed suprapubic	15	341	rs8025889	4.98	7.45E−04	2.18E−02	rs12905397	4.03	3.89E−04	2.33E−02
PBX3	Testis	9	377	rs12352375	−4.93	9.63E−07	1.88E−04	rs6478712	−4.38	9.19E−05	3.83E−02

TWAS, transcriptome-wide association study; GERD, gastroesophageal reflux disease; LUSC, lung squamous cell carcinoma; CHR, chromosome; SNP, single nucleotide polymorphism; GWAS, genome-wide association study; FDR, false discovery rate.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S11 TWAS-identified shared gene-tissue pairs between GERD and SCLC

Gene	Tissue type	CHR	SNP. Num	Gastroesophageal reflux				Small-cell lung cancer			
				BEST. GWAS.ID	BEST. GWAS.Z	TWAS.P	FDR	BEST. GWAS.ID	BEST. GWAS.Z	TWAS.P	FDR
ASCC3	Artery aorta	6	297	rs12524934	−4.56	1.62E−05	1.51E−03	rs12524934	3.88	3.71E−04	3.89E−02
ASCC3	Artery coronary	6	299	rs12524934	−4.56	7.10E−05	2.07E−03	rs12524934	3.88	2.26E−04	3.96E−02
ENSG00000265055	Artery tibial	17	320	rs6504573	4.80	9.50E−05	5.00E−03	rs11867618	4.28	7.83E−05	4.49E−02
C6orf163	Brain frontal cortex BA9	6	409	rs2268992	4.39	2.07E−03	2.99E−02	rs12663587	3.50	2.57E−04	4.86E−02
ENSG00000237854	Brain frontal cortex BA9	17	318	rs6504573	4.80	1.18E−04	3.91E−03	rs11867618	4.28	8.59E−05	1.98E−02
NBN	Brain putamen basal ganglia	8	375	rs40457	3.33	2.34E−03	3.98E−02	rs16902897	4.28	3.93E−05	6.01E−03
ENSG00000278730	Colon sigmoid	17	321	rs6504573	4.71	7.11E−04	1.72E−02	rs11867618	4.34	2.54E−04	3.51E−02
BPTF	Heart atrial appendage	17	318	rs6504573	4.51	7.11E−05	3.73E−03	rs7216064	4.10	6.74E−05	1.91E−02
ENSG00000265055	Heart atrial appendage	17	320	rs6504573	4.80	6.63E−05	3.73E−03	rs11867618	4.28	9.19E−05	1.91E−02
GMPPB	Minor salivary gland	3	313	rs2526743	−9.82	3.70E−14	2.78E−12	rs11130208	3.69	2.22E−04	3.31E−02
HYKK	Muscle skeletal	15	422	rs34016249	3.31	4.58E−04	1.54E−02	rs2036527	9.35	2.48E−07	4.13E−05
C17orf58	Pituitary	17	329	rs6504573	4.66	1.02E−04	5.20E−03	rs7216064	4.10	1.00E−04	2.81E−02
ENSG00000265055	Prostate	17	319	rs6504573	4.80	3.41E−04	7.65E−03	rs11867618	4.31	6.72E−05	1.88E−02
ENSG00000267708	Testis	17	320	rs6504573	4.35	2.42E−04	1.23E−02	rs11867618	4.40	6.64E−05	2.20E−02
ENSG00000274712	Uterus	17	343	rs6504573	4.32	5.66E−04	1.43E−02	rs11867618	4.46	2.16E−04	3.24E−02

GERD, gastroesophageal reflux disease; SCLC, small-cell lung cancer; CHR, chromosome; SNP, single nucleotide polymorphism; GWAS, genome-wide association study; FDR, false discovery rate.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

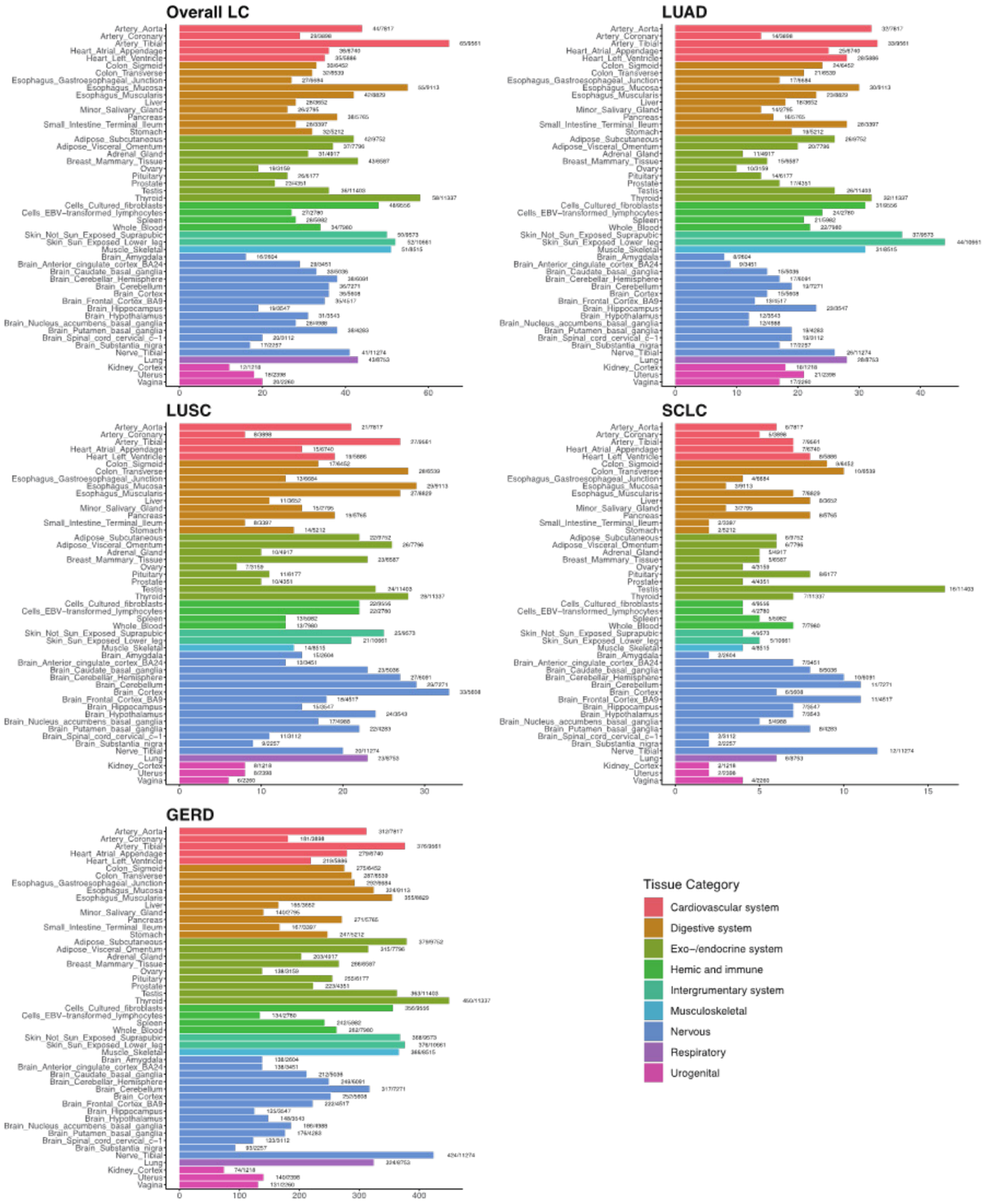

Figure S2 Number of TWAS significant genes for GERD and LC across 49 GTEx tissues (version 8). The X axis showcases the count of genes from GTEx tissues that meet the significance thresholds for multiple testing for each trait. The Y axis lists the respective GTEx tissues. Different tissue categories are indicated by distinct colors. The null hypothesis of TWAS assumes no expression-trait association, implying no genetic correlation between expression and a trait, under the condition of the observed GWAS statistics at the corresponding locus. In total, there are approximately 290,000 TWAS gene-tissue pairs undergoing testing across these 49 GTEx tissues. TWAS, transcriptome-wide association study; GERD, gastroesophageal reflux disease; LC, lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S12 Details of instrumental variables selected for GERD

SNP	Chromosome	Position	Effect allele	Other allele	Effect allele frequency	Beta	Standard error	Sample size	P value	R^2	F-statistic	Pleiotropic traits
rs12357321	10	21790476	G	A	0.6889	–0.0317	0.0052	367411	1.33E–09	4.31E–04	158.48	
rs1021363	10	106610839	A	G	0.3580	0.0312	0.0050	367411	5.10E–10	4.48E–04	164.66	
rs761777	10	134938075	A	G	0.7460	–0.0345	0.0055	367411	4.71E–10	4.52E–04	166.14	
rs10837002	11	38565727	C	G	0.6488	–0.0276	0.0050	367411	4.03E–08	3.48E–04	128.05	
rs7942368	11	76465362	C	T	0.7853	0.0340	0.0059	367411	9.54E–09	3.89E–04	142.99	
rs7104724	11	112837559	T	A	0.8833	–0.0416	0.0075	367411	3.01E–08	3.57E–04	131.39	Past tobacco smoking
rs2734839	11	113286490	C	T	0.3933	0.0283	0.0049	367411	8.79E–09	3.84E–04	140.96	
rs1479405	12	15387519	C	T	0.6783	–0.0315	0.0052	367411	9.85E–10	4.33E–04	159.01	
rs773109	12	56374695	G	A	0.6647	0.0381	0.0051	367411	8.71E–14	6.46E–04	237.34	
rs324769	12	83969240	C	T	0.5508	0.0268	0.0048	367411	3.05E–08	3.55E–04	130.33	
rs1716171	12	123716376	C	T	0.2100	–0.0384	0.0059	367411	7.82E–11	4.89E–04	179.81	
rs9542729	13	31833578	C	G	0.7976	0.0363	0.0060	367411	1.41E–09	4.26E–04	156.57	
rs1334297	13	58335375	G	A	0.2658	0.0388	0.0055	367411	1.14E–12	5.87E–04	215.96	
rs9540720	13	66922705	A	G	0.5211	–0.0267	0.0048	367411	3.01E–08	3.56E–04	130.77	
rs9517313	13	99105892	G	C	0.6168	–0.0331	0.0049	367411	2.05E–11	5.18E–04	190.55	
rs957345	14	75276079	C	G	0.4598	–0.0291	0.0048	367411	1.72E–09	4.19E–04	154.12	
rs17701934	14	89394474	T	C	0.5629	0.0265	0.0048	367411	4.60E–08	3.45E–04	126.91	
rs942065	14	94032065	G	A	0.3660	–0.0307	0.0050	367411	8.45E–10	4.38E–04	161.17	
rs10133111	14	103377321	G	A	0.8370	–0.0418	0.0065	367411	1.35E–10	4.76E–04	175.14	
rs11645288	16	51172677	G	A	0.8087	–0.0339	0.0061	367411	2.78E–08	3.56E–04	130.93	
rs9940128	16	53800754	G	A	0.5782	–0.0333	0.0049	367411	8.06E–12	5.39E–04	198.25	Type II diabetes
rs12598916	16	60658751	C	G	0.7252	0.0333	0.0054	367411	6.87E–10	4.41E–04	162.08	
rs7206608	16	82872628	C	G	0.6771	–0.0292	0.0051	367411	1.46E–08	3.72E–04	136.61	
rs12453010	17	50316131	C	T	0.6052	–0.0297	0.0049	367411	1.75E–09	4.21E–04	154.90	
rs12967855	18	35138245	A	G	0.3296	0.0365	0.0051	367411	1.09E–12	5.90E–04	216.97	
rs1431196	18	50832102	A	G	0.5716	–0.0324	0.0049	367411	2.65E–11	5.15E–04	189.22	
rs1942262	18	52873317	G	A	0.7086	–0.0315	0.0053	367411	2.60E–09	4.10E–04	150.64	
rs7241572	18	77580712	G	A	0.7909	–0.0366	0.0060	367411	9.49E–10	4.42E–04	162.42	
rs9636202	19	18449238	G	A	0.7334	0.0350	0.0055	367411	1.51E–10	4.80E–04	176.54	
rs2023878	19	18834124	C	T	0.8076	0.0363	0.0061	367411	3.04E–09	4.09E–04	150.37	
rs569356	1	29136686	A	G	0.8592	0.0379	0.0069	367411	4.07E–08	3.48E–04	127.89	
rs3766823	1	32197257	G	A	0.8285	–0.0394	0.0064	367411	7.09E–10	4.40E–04	161.80	
rs2782641	1	44013355	G	A	0.3873	–0.0271	0.0049	367411	4.33E–08	3.48E–04	128.00	
rs1937450	1	66478840	T	G	0.4623	–0.0316	0.0048	367411	7.07E–11	4.96E–04	182.31	
rs2815749	1	72814783	A	G	0.1990	–0.0389	0.0060	367411	1.07E–10	4.82E–04	177.13	Whole body fat mass
rs17379561	1	98340139	A	T	0.8556	–0.0531	0.0069	367411	1.08E–14	6.96E–04	255.87	
rs861575	1	184725099	T	C	0.5759	–0.0276	0.0049	367411	1.63E–08	3.73E–04	136.94	
rs7527682	1	189172684	A	G	0.4628	0.0267	0.0048	367411	3.13E–08	3.54E–04	130.12	
rs7541875	1	190957589	A	G	0.5739	–0.0274	0.0048	367411	1.61E–08	3.67E–04	134.92	
rs903678	1	201809918	G	A	0.6606	–0.0277	0.0051	367411	4.89E–08	3.45E–04	126.80	Whole body fat mass
rs1883842	20	41223062	T	G	0.7207	–0.0308	0.0054	367411	9.27E–09	3.83E–04	140.66	
rs2834005	21	34291708	T	C	0.6850	–0.0297	0.0052	367411	9.42E–09	3.81E–04	139.91	
rs2183588	21	42626882	A	G	0.3506	–0.0288	0.0051	367411	1.22E–08	3.79E–04	139.28	
rs2838771	21	46501576	G	C	0.3533	0.0281	0.0051	367411	2.91E–08	3.61E–04	132.60	
rs9615905	22	48875699	C	T	0.5418	–0.0276	0.0048	367411	1.21E–08	3.77E–04	138.67	
rs4300861	2	22549441	C	T	0.6179	–0.0307	0.0049	367411	5.43E–10	4.45E–04	163.72	
rs12997558	2	41704580	G	A	0.6412	–0.0278	0.0050	367411	3.04E–08	3.56E–04	130.86	
rs1011407	2	60665768	A	G	0.8784	0.0421	0.0074	367411	1.09E–08	3.78E–04	138.94	
rs4851239	2	100489966	C	T	0.6204	0.0329	0.0050	367411	3.24E–11	5.08E–04	186.88	
rs6722661	2	100806588	G	A	0.6353	0.0323	0.0050	367411	1.15E–10	4.82E–04	177.20	
rs6711584	2	104421692	G	A	0.5480	–0.0323	0.0048	367411	2.66E–11	5.15E–04	189.45	Whole body fat mass

Table S12 *(continued)*

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S12 *(continued)*

SNP	Chromosome	Position	Effect allele	Other allele	Effect allele frequency	Beta	Standard error	Sample size	P value	R^2	F-statistic	Pleiotropic traits
rs13409451	2	144257639	A	G	0.6076	0.0277	0.0049	367411	1.93E–08	3.66E–04	134.56	
rs2358016	2	162007430	C	G	0.5019	–0.0283	0.0048	367411	4.17E–09	4.00E–04	146.88	
rs1596747	2	193802478	A	G	0.5059	–0.0311	0.0048	367411	1.00E–10	4.83E–04	177.59	
rs7600261	2	212622818	C	T	0.6936	–0.0338	0.0052	367411	9.47E–11	4.86E–04	178.53	
rs7612999	3	35678337	G	A	0.7547	–0.0305	0.0056	367411	4.90E–08	3.45E–04	126.80	
rs6441814	3	44049114	G	A	0.5294	0.0284	0.0048	367411	3.86E–09	4.03E–04	147.97	
rs3172494	3	48731487	G	T	0.8930	0.0454	0.0078	367411	6.71E–09	3.93E–04	144.51	
rs2240326	3	50128386	G	A	0.5262	0.0472	0.0048	367411	1.13E–22	1.11E–03	408.04	
rs2016933	3	65653157	C	G	0.2699	0.0310	0.0054	367411	1.04E–08	3.79E–04	139.45	
rs6780459	3	104624105	A	T	0.2534	–0.0306	0.0055	367411	3.14E–08	3.53E–04	129.79	
rs7685686	4	3207142	A	G	0.5776	0.0279	0.0049	367411	1.14E–08	3.80E–04	139.82	
rs2164300	4	67813017	C	T	0.4767	0.0265	0.0048	367411	4.13E–08	3.50E–04	128.53	
rs7675588	4	80734978	C	A	0.2054	0.0335	0.0060	367411	1.80E–08	3.67E–04	134.81	
rs13107325	4	103188709	C	T	0.9256	–0.0701	0.0092	367411	2.20E–14	6.78E–04	249.29	Body mass index
rs1510719	4	140938116	T	C	0.6166	0.0389	0.0049	367411	3.84E–15	7.15E–04	262.85	
rs10010963	4	159839313	C	T	0.3836	0.0270	0.0049	367411	4.92E–08	3.44E–04	126.52	
rs1592757	5	103889998	G	C	0.6442	–0.0311	0.0050	367411	6.00E–10	4.44E–04	163.02	
rs11953061	5	120144025	C	T	0.6611	–0.0282	0.0051	367411	3.10E–08	3.55E–04	130.60	
rs329122	5	133864599	G	A	0.5804	0.0290	0.0049	367411	3.05E–09	4.08E–04	150.08	Type 2 diabetes
rs4461735	6	16946758	A	G	0.7620	0.0346	0.0056	367411	8.78E–10	4.35E–04	159.82	
rs9358901	6	26024436	G	T	0.3067	0.0327	0.0052	367411	3.68E–10	4.54E–04	166.77	
rs2145318	6	26496603	T	A	0.5134	–0.0353	0.0048	367411	2.03E–13	6.24E–04	229.45	
rs3828917	6	31465917	G	T	0.9582	–0.0671	0.0120	367411	2.27E–08	3.61E–04	132.68	
rs4713692	6	33807638	C	T	0.6322	0.0276	0.0050	367411	3.07E–08	3.55E–04	130.32	
rs205262	6	34563164	A	G	0.7328	–0.0348	0.0054	367411	1.38E–10	4.75E–04	174.47	Trunk predicted mass
rs9372625	6	98344031	G	A	0.6170	0.0377	0.0050	367411	2.62E–14	6.73E–04	247.33	
rs9373363	6	143150043	A	G	0.7464	0.0327	0.0056	367411	4.13E–09	4.04E–04	148.65	
rs12204714	6	152235339	C	T	0.3678	0.0288	0.0050	367411	7.92E–09	3.86E–04	141.94	
rs11762636	7	2061111	C	A	0.8197	0.0515	0.0063	367411	1.88E–16	7.83E–04	288.04	
rs10242223	7	3521573	A	G	0.3302	0.0282	0.0051	367411	3.79E–08	3.52E–04	129.48	Past tobacco smoking
rs2043539	7	12253880	G	A	0.5813	–0.0272	0.0049	367411	2.24E–08	3.60E–04	132.42	
rs215614	7	32347335	G	A	0.3703	0.0329	0.0050	367411	4.08E–11	5.03E–04	185.03	Whole body fat mass
rs2396133	7	109197067	A	G	0.5247	–0.0294	0.0048	367411	1.11E–09	4.30E–04	157.98	
rs2396766	7	114318071	G	A	0.5269	–0.0322	0.0048	367411	2.33E–11	5.17E–04	190.09	
rs2106353	7	126506598	G	T	0.7685	–0.0367	0.0057	367411	1.37E–10	4.80E–04	176.61	
rs3863241	8	73890335	C	T	0.4730	–0.0325	0.0048	367411	1.49E–11	5.27E–04	193.55	
rs903959	8	142630782	T	A	0.6007	–0.0292	0.0049	367411	2.99E–09	4.08E–04	149.96	
rs3793577	9	23737627	A	G	0.4617	–0.0270	0.0048	367411	2.49E–08	3.63E–04	133.49	
rs7032155	9	122672771	C	A	0.4082	–0.0278	0.0049	367411	1.63E–08	3.72E–04	136.74	
rs4382592	9	134870755	T	G	0.3005	0.0303	0.0053	367411	8.20E–09	3.85E–04	141.55	

F-statistic was calculated using the following formulas: $R^2=2\times\beta^2\times EAF\times(1 - EAF)$ and $F=R^2 (n - 2)/(1 - R^2)$, where R^2 represents the phenotypic variance explained by a genetic instrument, *F* represents *F* statistic, n is the sample size, β is the estimated genetic association of SNP with the exposure, EAF is the effect allele frequency. SNP, single nucleotide polymorphism; GERD, gastroesophageal reflux disease.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S13 Causal Analysis Using Summary Effect estimates (CAUSE) results for GERD on LC

Exposure	Outcome	Model 1	Model 2	Delta_ ELPD	se_delta_ ELPD	z score	γ (95% CI)	η (95% CI)	q (95% CI)	P value
GERD	LC	Null	Sharing	−19.0	5.2	−3.6	NA	0.49 (0.32, 0.75)	0.41 (0.22, 0.63)	1.30E−04
GERD	LC	Null	Causal	−24.0	6.6	−3.6	0.26 (0.17, 0.35)	0.02 (−1.11, 1.21)	0.04 (0, 0.26)	1.80E−04
GERD	LC	Sharing	Causal	−4.5	1.8	−2.5	NA	NA	NA	6.10E−03
GERD	LUAD	Null	Sharing	−4.6	2.4	−1.9	NA	0.53 (0.21, 1.19)	0.21 (0.03, 0.47)	2.60E−02
GERD	LUAD	Null	Causal	−8.0	4.0	−2.0	0.19 (0.09, 0.30)	0.04 (−1.52, 1.96)	0.03 (0, 0.25)	2.30E−02
GERD	LUAD	Sharing	Causal	−3.4	1.8	−1.9	NA	NA	NA	3.00E−02
GERD	LUSC	Null	Sharing	−13.0	4.4	−3.0	NA	0.73 (0.42, 1.25)	0.33 (0.14, 0.56)	1.40E−03
GERD	LUSC	Null	Causal	−17.0	5.6	−3.0	0.32 (0.19, 0.45)	0.07 (−1.77, 2.08)	0.04 (0, 0.25)	1.40E−03
GERD	LUSC	Sharing	Causal	−3.8	1.7	−2.3	NA	NA	NA	1.10E−02
GERD	SCLC	Null	Sharing	−13.0	4.4	−2.9	NA	1.14 (0.66, 1.87)	0.32 (0.14, 0.55)	1.60E−03
GERD	SCLC	Null	Causal	−16.0	5.7	−2.8	0.49 (0.28, 0.70)	−0.01 (−2.92, 2.73)	0.04 (0, 0.26)	2.50E−03
GERD	SCLC	Sharing	Causal	−3.1	1.7	−1.8	NA	NA	NA	3.90E−02

Model 1 and model 2 represent the models being compared (null, sharing, or causal). The fit of the models is assessed by examining the change in Expected Log Pointwise Posterior Density (ELPD), denoted as Delta_ELPD; negative values of z indicate that model 2 fits the data better. Delta_ELPD, estimated difference in ELPD; se_delta_ELPD, estimated standard error of delta_ELPD; z, delta_ELPD/se_delta_ELPD; GERD, gastroesophageal reflux disease; LC, lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer; γ (gamma), estimate of causal effect if causal model is correct; η (eta), estimate of correlated pleiotropy; q, proportion of effect due to correlated pleiotropy; CI, confidence interval; NA, not applicable.

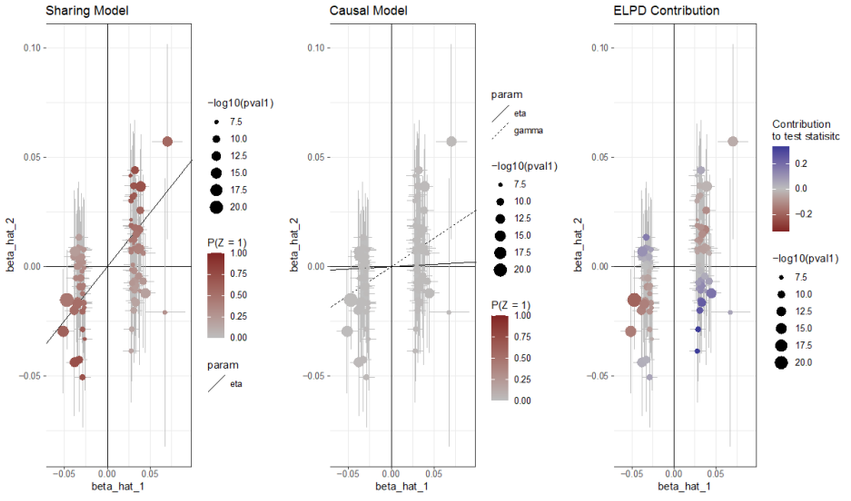

Figure S3 Causal Analysis Using Summary Effect estimates (CAUSE) results for GERD on overall LC. Plots illustrate the sharing, causal and expected log pointwise posterior density (ELPD) models for GERD on LC. CAUSE shows that there is significant evidence supporting both the sharing (correlated pleiotropy) and causal models compared to the null (no effect) model. Com-paring both shared and causal models, there is significant evidence indicating that the causal model fits the data better than the sharing model, indicating that correlated pleiotropy could be discounted. Gamma represents the estimate of causal effect if causal model is correct, while Eta represents the estimate of correlated pleiotropy. GERD, gastroesophageal reflux disease; LC, lung cancer.

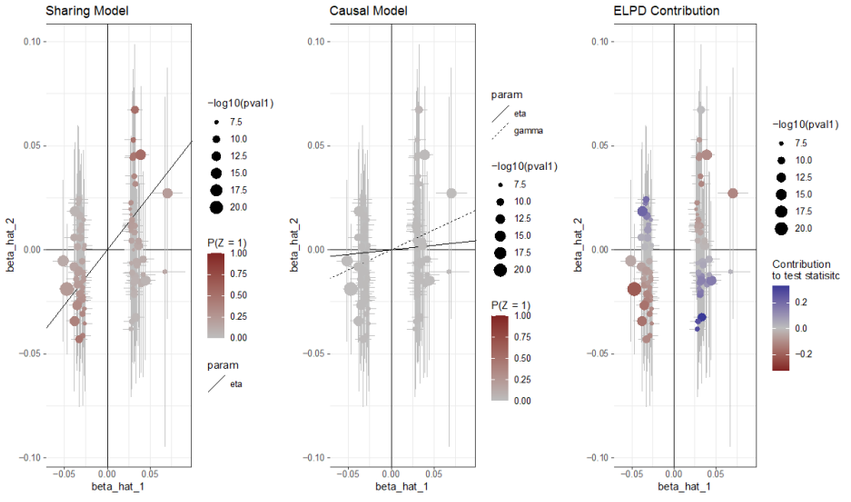

Figure S4 Causal Analysis Using Summary Effect estimates (CAUSE) results for GERD on LUAD. Plots illustrate the sharing, causal and expected log pointwise posterior density (ELPD) models for GERD on LUAD. CAUSE shows that there is significant evidence supporting both the sharing (correlated pleiotropy) and causal models compared to the null (no effect) model. Comparing both shared and causal models, there is significant evidence indicating that the causal model fits the data better than the sharing model, indicating that correlated pleiotropy could be discounted. Gamma represents the estimate of causal effect if causal model is correct, while Eta represents the estimate of correlated pleiotropy. GERD, gastroesophageal reflux disease; LUAD, lung adenocarcinoma.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

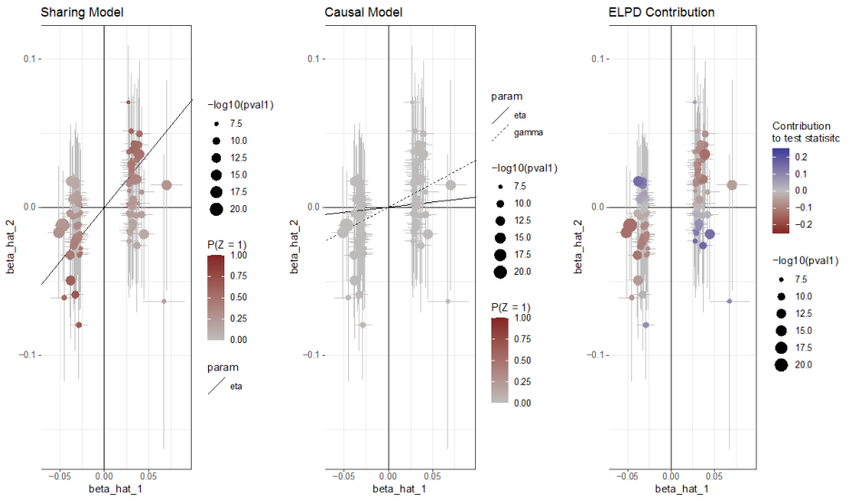

Figure S5 Causal Analysis Using Summary Effect estimates (CAUSE) results for GERD on LUSC. Plots illustrate the sharing, causal and expected log pointwise posterior density (ELPD) models for GERD on LUSC. CAUSE shows that there is significant evidence supporting both the sharing (correlated pleiotropy) and causal models compared to the null (no effect) model. Comparing both shared and causal models, there is significant evidence indicating that the causal model fits the data better than the sharing model, indicating that correlated pleiotropy could be discounted. Gamma represents the estimate of causal effect if causal model is correct, while eta represents the estimate of correlated pleiotropy. GERD, gastroesophageal reflux disease; LUSC, lung squamous cell carcinoma.

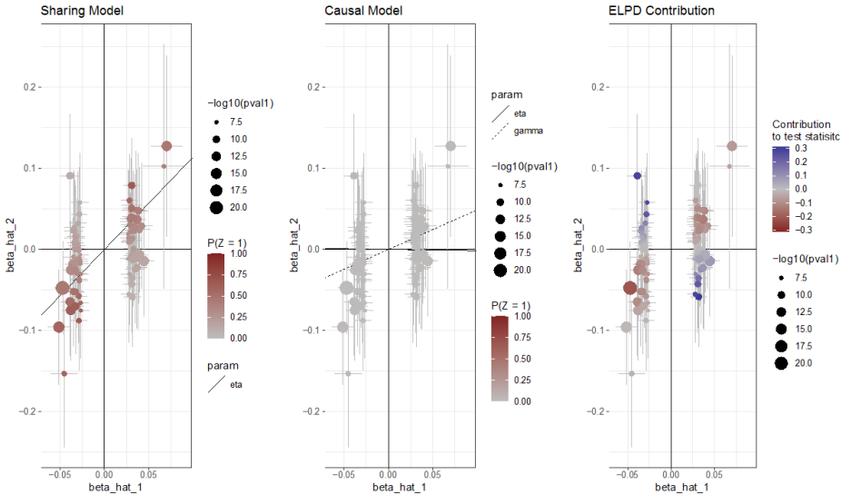

Figure S6 Causal Analysis Using Summary Effect estimates (CAUSE) results for GERD on SCLC. Plots illustrate the sharing, causal and expected log pointwise posterior density (ELPD) models for GERD on SCLC. CAUSE shows that there is significant evidence supporting both the sharing (correlated pleiotropy) and causal models compared to the null (no effect) model. Comparing both shared and causal models, there is significant evidence indicating that the causal model fits the data better than the sharing model, indicating that correlated pleiotropy could be discounted. Gamma represents the estimate of causal effect if causal model is correct, while eta represents the estimate of correlated pleiotropy. GERD, gastroesophageal reflux disease; SCLC, small-cell lung cancer.

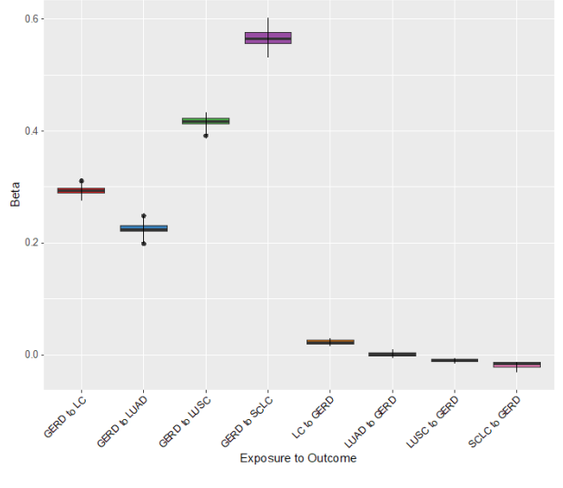

Figure S7 Box plot of the leave-one-out analysis. Each SNP was systematically removed at a time, and inverse-variance weighted analysis was performed using the remaining SNPs. Beta, effect allele beta coefficient; GERD, gastroesophageal reflux disease; LC, lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S14 Power of mendelian randomization analysis of GERD on LC

Exposure	Outcome	Power
Gastroesophageal reflux disease	Lung cancer	100%
	Lung adenocarcinoma	100%
	Lung squamous cell carcinoma	100%
	Small-cell lung cancer	100%

The power of MR analyses was calculated based on an online calculator (https://shiny.cnsgenomics.com/mRnd/). The calculation incorporated the type I error of 0.05, sample size, proportion of cases (Table S1), explained genetic variation (Table S12), and odds ratios from IVW analyses (*Figure 4*). GERD, gastroesophageal reflux disease; LC, lung cancer; MR, Mendelian randomization; IVW, inverse variance-weighted.

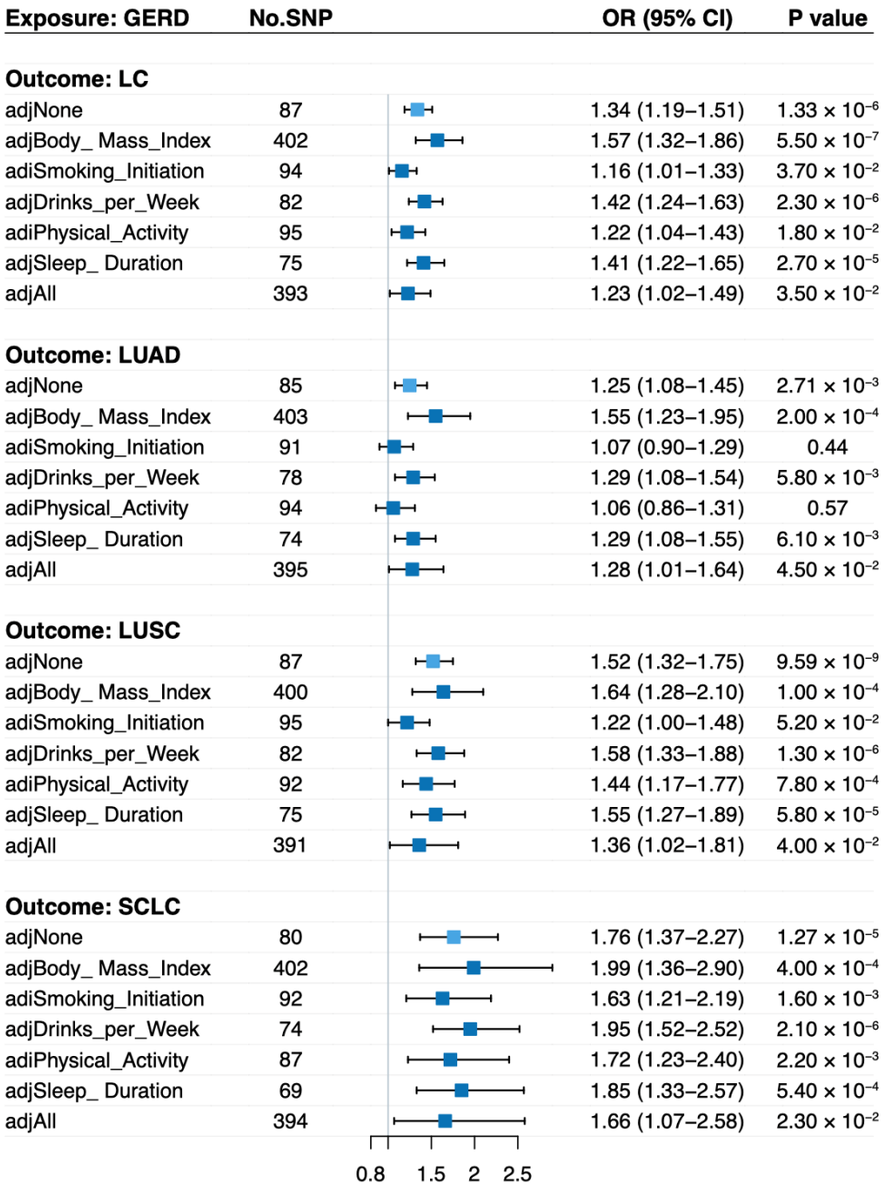

Figure S8 Multivariable Mendelian randomization of genetically predicted GERD on LC. The estimated effect sizes were adjusted for each potential confounder separately and combined. The y-axis indicates the genetically predicted confounder for which adjustment was made. The boxes represent the point estimates of the causal effects of genetically predicted GERD on LC, with error bars representing the 95% confidence intervals. "adjNone" refers to the point estimates of the causal effects of genetically predicted GERD on LC using univariable Mendelian randomization. GERD, gastroesophageal reflux disease; LC, lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small-cell lung cancer; SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345

Table S15 Details of instrumental variables selected for lung cancer

Phenotype	SNP	Chromosome	Position	Effect allele	Other allele	Effect allele frequency	Beta	Standard error	Sample size	P value	R^2	F-statistic
Lung cancer	rs10265693	7	130720805	G	A	0.0930	0.0861	0.0197	85716	1.87E−06	1.25E−03	107.45
Lung cancer	rs1039766	2	65520145	T	C	0.1544	0.1541	0.0163	85716	2.05E−06	6.20E−03	534.80
Lung cancer	rs10904377	10	4978419	T	G	0.0380	0.0491	0.0284	85716	2.24E−06	1.76E−04	15.11
Lung cancer	rs11780471	8	27344719	A	G	0.0638	0.0596	0.0250	85716	1.69E−08	4.25E−04	36.40
Lung cancer	rs12081674	1	90337025	G	A	0.1810	0.1841	0.0154	85716	4.14E−06	1.00E−02	869.67
Lung cancer	rs150211	21	19415773	C	T	0.0411	0.0383	0.0326	85716	3.59E−06	1.16E−04	10.00
Lung cancer	rs17181550	17	70299958	G	T	0.4170	0.4265	0.0119	85716	1.98E−07	8.84E−02	8316.61
Lung cancer	rs17775239	8	128897079	A	T	0.2609	0.2615	0.0135	85716	2.78E−06	2.64E−02	2322.19
Lung cancer	rs1991625	2	153831482	G	A	0.3056	0.3221	0.0125	85716	3.99E−06	4.40E−02	3948.86
Lung cancer	rs239935	6	167411788	G	A	0.4816	0.4785	0.0118	85716	1.29E−08	1.14E−01	11062.39
Lung cancer	rs380286	5	1320247	A	G	0.4371	0.4226	0.0119	85716	1.51E−32	8.79E−02	8260.00
Lung cancer	rs631644	18	2280388	G	A	0.3878	0.4040	0.0139	85716	3.06E−06	7.75E−02	7200.82
Lung cancer	rs7805022	7	9825541	A	T	0.3398	0.3256	0.0125	85716	1.92E−06	4.76E−02	4281.52
Lung cancer	rs9869622	3	16671821	T	C	0.1863	0.1808	0.0152	85716	1.73E−06	9.91E−03	857.92
Lung adenocarcinoma	rs1039766	2	65520145	T	C	0.1539	0.1054	0.0224	66756	2.60E−06	2.89E−03	193.78
Lung adenocarcinoma	rs10445262	17	4943176	G	A	0.2437	0.0852	0.0186	66756	4.76E−06	2.68E−03	179.20
Lung adenocarcinoma	rs1056562	11	118125625	T	C	0.4731	0.1021	0.0162	66756	2.76E−10	5.20E−03	349.01
Lung adenocarcinoma	rs11591710	10	105687632	C	A	0.1368	0.1506	0.0230	66756	6.30E−11	5.35E−03	359.20
Lung adenocarcinoma	rs11848063	14	47235445	G	A	0.4083	−0.0777	0.0167	66756	3.31E−06	2.91E−03	195.05
Lung adenocarcinoma	rs11855650	15	70431773	T	G	0.3825	0.0829	0.0166	66756	5.60E−07	3.25E−03	217.65
Lung adenocarcinoma	rs12369136	12	20576874	G	A	0.0400	0.2042	0.0438	66756	3.21E−06	3.20E−03	214.32
Lung adenocarcinoma	rs1512829	11	9951257	G	A	0.2384	0.0915	0.0191	66756	1.66E−06	3.04E−03	203.69
Lung adenocarcinoma	rs17181550	17	70299958	G	T	0.4263	−0.0767	0.0165	66756	3.21E−06	2.88E−03	192.64
Lung adenocarcinoma	rs2320614	4	164070122	C	T	0.4002	0.0846	0.0164	66756	2.48E−07	3.44E−03	230.38
Lung adenocarcinoma	rs2608029	8	129170126	G	C	0.3510	0.0810	0.0168	66756	1.44E−06	2.99E−03	199.96
Lung adenocarcinoma	rs421629	5	1320136	A	G	0.4267	−0.1566	0.0163	66756	9.75E−22	1.20E−02	810.54
Lung adenocarcinoma	rs4236709	8	32410110	G	A	0.2178	0.1243	0.0193	66756	1.28E−10	5.27E−03	353.34
Lung adenocarcinoma	rs8108034	19	39813853	G	T	0.0988	0.1302	0.0281	66756	3.50E−06	3.02E−03	201.96
Lung adenocarcinoma	rs885518	9	21830157	G	A	0.1011	0.1545	0.0253	66756	9.96E−10	4.34E−03	291.05
Lung squamous cell carcinoma	rs1108581	9	136505241	G	A	0.2034	−0.1154	0.0235	63053	9.09E−07	4.31E−03	273.09
Lung squamous cell carcinoma	rs13031455	2	17784157	C	T	0.4048	−0.0899	0.0190	63053	2.23E−06	3.89E−03	246.29
Lung squamous cell carcinoma	rs1333040	9	22083404	C	T	0.4579	0.0936	0.0189	63053	7.02E−07	4.35E−03	275.62
Lung squamous cell carcinoma	rs1534979	20	2331513	T	C	0.3242	0.0908	0.0199	63053	4.90E−06	3.61E−03	228.56
Lung squamous cell carcinoma	rs2674946	17	13066819	T	A	0.2791	0.0978	0.0211	63053	3.74E−06	3.85E−03	243.67
Lung squamous cell carcinoma	rs3754287	1	41952597	T	C	0.1426	0.1299	0.0271	63053	1.66E−06	4.13E−03	261.18
Lung squamous cell carcinoma	rs4453114	10	4961021	C	T	0.0508	0.2072	0.0436	63053	2.04E−06	4.14E−03	262.27
Lung squamous cell carcinoma	rs467095	5	1336221	C	T	0.4269	−0.1788	0.0191	63053	6.73E−21	1.56E−02	1002.37
Lung squamous cell carcinoma	rs6957511	7	130668618	C	T	0.3978	0.0951	0.0194	63053	9.78E−07	4.33E−03	274.23
Lung squamous cell carcinoma	rs7591446	2	45834076	T	C	0.1903	−0.1075	0.0235	63053	4.73E−06	3.56E−03	225.23
Lung squamous cell carcinoma	rs7658584	4	89096641	A	G	0.1546	0.1194	0.0253	63053	2.41E−06	3.73E−03	235.95
Lung squamous cell carcinoma	rs8040868	15	78911181	C	T	0.4135	0.2550	0.0189	63053	2.50E−41	3.15E−02	2053.12
Lung squamous cell carcinoma	rs9602270	13	84281063	T	A	0.0486	0.2385	0.0467	63053	3.28E−07	5.26E−03	333.56
Small-cell lung cancer	rs1703426	5	133185535	T	C	0.1749	−0.1861	0.0388	24108	1.59E−06	1.00E−02	243.55
Small-cell lung cancer	rs17185553	9	17934120	C	G	0.0810	0.2575	0.0524	24108	8.94E−07	9.88E−03	240.45
Small-cell lung cancer	rs3134425	11	122709178	T	C	0.3810	−0.1497	0.0311	24108	1.47E−06	1.06E−02	257.62
Small-cell lung cancer	rs6463739	7	7906724	C	G	0.4033	−0.1445	0.0304	24108	1.94E−06	1.01E−02	244.84

F-statistic was calculated using the following formulas: $R^2=2\times\beta^2\times EAF\times(1-EAF)$ and $F=R^2(n-2)/(1-R^2)$, where R^2 represents the phenotypic variance explained by a genetic instrument, *F* represents *F* statistic, n is the sample size, β is the estimated genetic association of SNP with the exposure, EAF is the effect allele frequency. SNP, single nucleotide polymorphism.

© AME Publishing Company.
https://dx.doi.org/10.21037/tlcr-24-345