BMC Genomics

RESEARCH

Open Access

Identifying cross-tissue molecular targets of lung function by multi-omics integration analysis from DNA methylation and gene expression of diverse human tissues

Shisheng Peng[1], Jinlong Fang[1], Weiliang Mo[1], Guodong Hu[1*] and Senquan Wu[1*]

Abstract

Background Previous studies have reported several genetic loci associated with lung function. However, the mediating mechanism between these genetic loci and lung function phenotype is rarely explored. In this research, we used a cross-tissue multi-omics post-GWAS analysis to explain the associations between DNA methylation, gene expression, and lung function.

Methods We conducted integration analyses of lung function traits using genome-wide association study (GWAS) summary data alongside expression quantitative trait loci (eQTLs) and DNA methylation quantitative trait loci (mQTLs) derived from whole blood, utilizing multi-omics SMR and Bayesian colocalization analysis. Considering the genetic differences of tissues, we replicated the shared causal signals of eQTLs and lung function in 48 diverse tissues and the shared causal signals of mQTLs and lung function in 8 diverse tissues. Multi-trait colocalization analyses were utilized to identify the causal signals between gene expression in blood, blood cell traits, and lung function, as well as between cross-tissue gene expression in diverse tissues and lung function.

Results Eight genes from blood tissue were prioritized as FEV1 causal genes using multi-omics SMR analysis and COLOC colocalization analysis: *EML3*, *UBXN2A*, *ROM1*, *ZBTB38*, *RASGRP3*, *FAIM*, *PABPC4*, and *SNIP1*. Equally, five genes (*CD46*, *EML3*, *UBXN2A*, *ZBTB38*, and *LMCD1*) were prioritized as FVC causal genes and one gene (*LMCD1*) was prioritized as FEV1/FVC causal genes. The causal signals between 8 genes (*EML3*, *ROM1*, *UBXN2A*, *ZBTB38*, *RASGRP3*, *FAIM*, *PABPC4*, and *CD46*) and lung function were successfully replicated in diverse tissues. More importantly, MOLCO colocalization analysis showed that 3 genes (*CD46*, *LMCD1*, and *ZBTB38*) expression in blood, blood cell traits, and lung function traits shared the same causal signals. Finally, through cross-tissue colocalization analysis of multiple traits, we found that the heart–lung axis *EML3* expressions and lung function mediate the same causal signal.

Conclusion This study identified potential cross-tissue molecular targets associated with lung function traits from DNA methylation and gene expression of diverse tissues and explored the probable regulation mechanism of these molecular targets. This provides multi-omics and cross-tissue evidence for the molecular regulation mechanism of lung function and may provide new insight into the influence of crosstalk between organs and tissues on lung function.

Keywords Lung function, DNA methylation, Gene expression, Cross-tissue targets, Multi-omics analysis

*Correspondence:
Guodong Hu
huguodong@smu.edu.cn
Senquan Wu
wusenquan@stu2020.jnu.edu.cn

Background

Lung function is an important index for diagnosis and evaluation of respiratory diseases, including but not limited to chronic obstructive pulmonary disease (COPD), asthma, lung cancer, and COVID-19 [1]. Lung function, including forced expiratory volume in one second (FEV1), forced vital capacity (FVC), and their ratio (FEV1/FVC) is influenced by heredity and environment [2]. The variance in lung function phenotype remains incompletely explained by genetic variation, as it only accounts for heritability between 10 and 33%, implicating approximately 300 genetic loci [3–5]. The lung function trajectories of normal people change with the development and aging of lung in the whole life course [6]. Epigenetic modifications, such as DNA methylation, are established in early embryonic development and incorporate variation according to host genotype and changeable factors in the whole lifespan [7]. Previous large-scale studies have established significant associations between DNA methylation and lung function across both blood and lung tissue [8–10]. The impact of environmental exposure on respiratory health and lung function throughout the lifespan has been well recognized, notably in relation to factors such as cigarette smoke, environmental tobacco smoke, air pollution, and occupational exposures [11]. These environmental stressors have been shown to significantly affect DNA methylation patterns and gene expression in various tissues, including lung tissue. Owing to the complexity of epigenetics and genetics, clarifying the causal association between DNA methylation, gene expression, and lung function may provide new insights for explaining lung function and lung diseases.

Complex diseases and traits usually affect multiple tissues or organs, and interestingly damage and repair of tissues or organs will in turn affect diseases and traits. As the hub of human organs and tissues, blood is involved in the communications of RNA, protein, and cells in the whole body. The association between diverse tissues (including whole blood) genetic loci, blood components, and complex traits is still unclear. Therefore, the analysis based on multi-tissue genetic loci and blood components can provide new insights into the complex genetic mechanism of lung function traits.

Large-scale genome-wide association studies (GWASs) have been employed to identify genomic loci associated with lung function [3, 4]. However, due to the complicated linkage disequilibrium (LD) structure of the genome, the top associated variations may not be causal [12]. Furthermore, these genetic variations can potentially regulate chromatin accessibility, DNA splicing, DNA methylation, gene expression, and protein level to affect complex diseases and traits (including lung function traits). Consequently, there is a key challenge in dissecting the causal variants in trait-associated loci and interpreting their biological mechanisms. Multi-omics integration analysis of GWAS data is a method of post-GWAS analysis, which explains complex diseases and traits by identifying the genes and regulatory elements behind the associated loci of GWAS [13]. For instance, summary data-based Mendelian randomization (SMR) can integrate GWAS with various quantitative trait locus (QTLs) including gene expression quantitative trait loci (eQTL), methylation quantitative trait loci (mQTL), protein expression quantitative trait loci (pQTL), and chromatin accessibility quantitative trait loci (caQTL) to assess potential pleiotropic associations between the expression levels of QTLs and a complex trait of interest [13, 14]. In addition, Bayesian colocalization analysis can identify the same causal signals of QTLs and complex traits to detect the causal association between the expression level of QTLs and a complex trait of interest. The integration of the two methods serves as a test and compensation for each other.

In this research, we present a multi-omics-based post-GWAS study using *cis*-mQTLs and *cis*-eQTLs to identify the putative causal effects and molecular mechanisms of heredity in lung function. Utilizing SMR methods, we integrated lung function GWAS summary statistics with mQTLs and eQTLs in diverse human tissues to explain the effects of DNA methylation and gene expression on lung function. Furthermore, we used Bayesian colocalization analysis to verify the candidate genes of the same causal signals with lung function. For the key molecular targets obtained in blood, we replicated these results in the methylation of 8 human tissues and transcriptome of 48 human tissues. Finally, we identified the cross-tissue molecular targets with the same causal signal by multi-trait colocalization method in diverse tissues.

Methods

The design of this study is described in Fig. 1. In brief, we integrated cis-mQTLs in whole blood, cis-eQTLs in whole blood, and lung function through Multi-omics SMR and colocalization analysis to identify causal signals of DNA methylation, gene expression, and lung function traits. Then we replicated these causal signals in the other 13 datasets to verify their stability. For the stable results obtained in the discovery datasets and replication datasets, we replicated SMR or colocalization between cis-eQTLs/cis-mQTLs and lung function in diverse human tissues to verify the causal relationship between key genes and lung function in multiple tissues transcription and methylation levels. Finally, we conducted multi-trait colocalization analysis on multiple tissues cis-eQTLs and lung function to explain that cross-tissue gene expression affects lung function traits through cross-talk of diverse tissues.

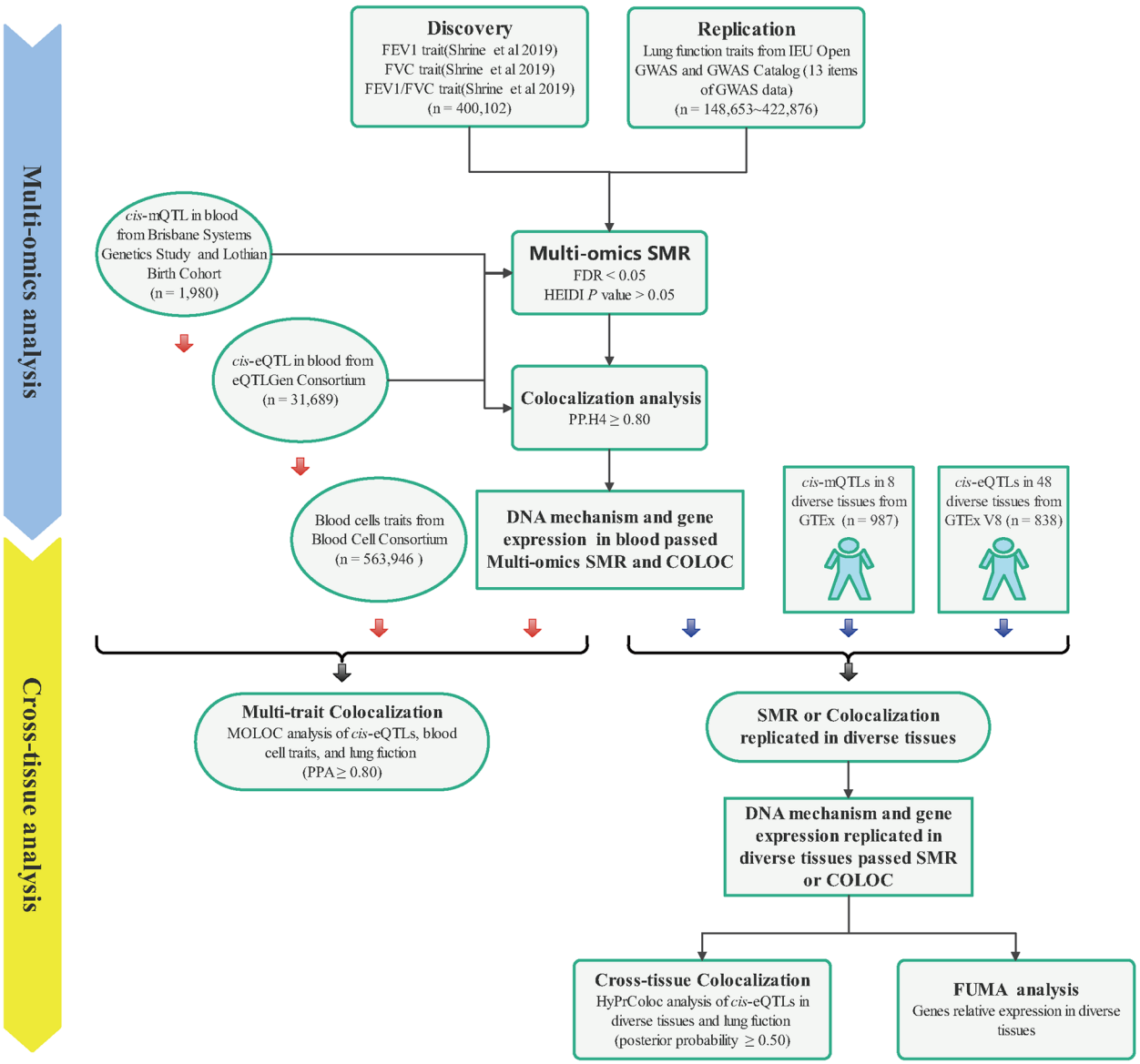


Fig. 1 Study design of this work. FEV1: Forced expiratory volume in one second; FVC: Forced vital capacity; FEV1/FVC: Forced expiratory volume in one second and forced vital capacity ratio; QTL: Quantitative trait loci; GWAS: Genome-wide association study; GTEx: Genotype-Tissue Expression; SMR: Summary data-based Mendelian randomization; COLOC: Bayesian colocalization; MOLOC: Multiple-trait-coloc; HyPrColoc: Hypothesis Prioritisation for multi-trait Colocalization; FUMA: Functional mapping and annotation of genetic associations; FDR: False discover rate; HEIDI: Heterogeneity in dependent instrument; PPH4: Posterior probability of hypothesis 4; PPA: An overall posterior probability; PPMT: Posterior probability of multiple traits

1. Study population and data resources

No ethical approval was required for the present study, as all analyses utilized publicly available summary statistics and did not involve access to individual-level data. GWAS summary statistics of discovery sets for FEV1, FVC, and FEV1/FVC were derived from Shrine et al. 20194. The replication sets were 13 GWAS data of lung function of European ancestry from GWAS Catalog (https://www.ebi.ac.uk/gwas/) and IEU OpenGWAS project (https://gwas.mrcieu.ac.uk/). The basic information of lung function datasets is shown in Table S1.

2. Data resources of cis-mQTLs, cis-eQTLs, and blood cell traits

Summary statistics for blood cis-eQTLs were sourced from the eQTLGen Consortium, which includes genetic data on blood gene expression from a total of 31,684 individuals across 37 datasets [15]. Blood cis-mQTLs summary statistics were collected from a meta-analysis of Brisbane Systems Genetics Study (BSGS) and Lothian Birth Cohort (LBC) ($n = 1980$) [14, 16]. The data on tissue-specific cis-eQTLs came from the Genotype-Tissue Expression (GTEx) v8 project ($n = 838$), which included eQTLs data for 49 different human tissues [17, 18]. In addition, the tissue-specific cis-mQTLs data was sourced from a study by Oliva et al. 2023 [19], encompassing 987 human samples from the GTEx project, representing nine tissue types and 424 subjects. The GWAS data for blood cells was obtained from the Blood Cell Consortium ($n = 563{,}946$) [20]. Basic information about these datasets is presented in Table S2.

3. Summary data-based Mendelian randomization analysis

The SMR multi-tool (version 1.3.1) was utilized to assess whether the effects of single nucleotide polymorphisms SNPs on lung function phenotypes are mediated by molecular traits, including DNA methylation levels, gene expression, and protein expression [14]. This framework allows for the exploration of causal relationships between genetic variants and phenotypic traits by integrating multi-omics data. Multi-omics SMR analysis used the SMR multi-tool to determine the causal inference of genes and the 1000 Genomes European reference to calculate LD [21]. The steps of performing multi-omics SMR analysis include: first, SNPs were instruments, gene expressions were exposure, and lung function (FEV1, FVC, FEV1/FVC) was outcome; second, SNPs were instruments, DNA methylations were exposure, and lung function (FEV1, FVC, FEV1/FVC) was outcome; finally, for the important signals from step 1 and step 2, SNPs were instruments, DNA methylations were exposure, and gene expressions were outcome [22]. The top associated cis-QTLs were selected by considering a window centered around the corresponding gene (± 500 kb) and a P value threshold of 5.0×10^{-8}. Heterogeneity in the dependent instrument (HEIDI) test was used to distinguish pleiotropy of linkage, the results of PHEIDI < 0.05 were considered pleiotropy, so the results of PHEIDI > 0.05 were retained. The P values were adjusted to control the false discovery rate (FDR) < 0.05 by the Benjamini–Hochberg method. The final candidate signals were excluded if the SNPs were located within the Major Histocompatibility Complex (MHC) region (chr6:25.5–34.0 Mb) due to its complex structure of linkage disequilibrium.

4. Bayesian colocalization analysis

We conducted Bayesian colocalization (COLOC) to distinguish causality from confounding by LD using the "coloc" R package [23]. The colocalization analysis included five hypotheses: 1) there was no causal variant for either QTLs or lung function (H0); 2) there was one causal variant for QTLs only (H1); 3) there was one causal variant for lung function only (H2); 4) there were two distinct causal variants for QTLs and lung function, one for QTLs and one for lung function (H3); 5) there was a shared causal variant for QTLs and lung function (H4). For each QTL, we included SNPs within a ± 500 kb window around the gene position. We used the default COLOC priors of $p1 = 10^{-4}$, $p2 = 10^{-4}$, and $p12 = 10^{-5}$, where p1 is the probability that a given variant is associated with GWAS, p2 is the probability that a given variant is a significant QTL, and p12 is the probability that a given variant is significant in both GWAS and QTL [23]. When the phenotype was a continuous variable, the parameter "quant" was selected. A posterior colocalization probability (PPH4) ≥ 0.80 was used to denote a shared causal signal.

5. Multi-trait colocalization analysis by MOLOC

For the robust results of SMR (SMR PFDR < 0.05, HEIDI $P > 0.05$) and colocalization (PPH4 > 0.80), we explored whether these causal variations are shared in gene expression, blood cell traits, and lung function traits. We employed the multiple-trait-coloc (MOLOC) implemented in the "moloc" R package for our analysis24. In this approach, we utilized default prior probabilities set to 10^{-4} for any single layer of association, 10^{-6} for any two layers of associations, and 10^{-7} for colocalization involving all three layers of associations. These parameters were applied in the colocalization analysis to assess whether the association signals of the multiple traits were derived from the same genetic variants within the tested region [24]. To infer colocalization, we considered an overall colocalization probability of the three traits calculated as (PPA, Pa,bc + Pab,c + Pac,b + Pabc) ≥ 0.80. This threshold indicates that the evidence is strong enough to suggest that the three association signals are likely to colocalize within the specified genomic region, supporting the hypothesis that they may share a common underlying genetic basis. We use the "gassocplot" R package (https://github.com/jrs95/gassocplot) to visualize the results of multi-trait colocalization.

6. FUMA GENE2FUNC analysis

To evaluate the expression profiles of the identified genes across various human tissues, we utilized the

FUMA (Functional mapping and annotation of genetic associations) GENE2FUNC tool to generate a gene expression heatmap [25]. This heatmap visually represents the expression levels of the selected genes, highlighting their presence in different tissues. The heatmap displays zero mean normalized log2 transformed expression values; deeper shades of red signify higher expression levels, while deeper shades of blue indicate lower expression levels. The construction of the gene expression heatmap relied on data from the Genotype-Tissue Expression v8 project, following established methodologies to ensure accuracy and reliability.

7. Cross-tissue colocalization analysis by HyPrColoc

We conducted a colocalization analysis of eQTLs from diverse human tissues related to lung function to investigate whether the genetic association between gene expression in diverse human tissues and lung function was caused by the same causal variation. We estimated the posterior probability of multiple traits sharing the same causal SNP by employing the multi-trait colocalization approach known as Hypothesis Prioritisation for multi-trait Colocalization (HyPrColoc) [26]. This method enhances the established COLOC framework by providing a more refined estimation of the colocalization posterior probability. Specifically, HyPrColoc approximates the true posterior probability of colocalization based on the posterior probability associated with a single causal variable while considering multiple traits simultaneously. By integrating information from different traits, HyPrColoc allows for a more robust inference regarding the potential shared genetic basis of the traits under investigation. This approach not only improves the sensitivity of detecting colocalization among multiple traits but also helps to clarify the relationships between SNPs and the phenotypic outcomes, ultimately advancing our understanding of the complex genetics involved in these traits. Prior to colocalization analysis, all variants on each dataset were included within a ± 500 kb window around gene position and harmonized to the same effect allele. The posterior probability of multiple traits (PPMT) ≥ 0.50 would suggest that the multiple association signals were to share the same causal SNP simultaneously. The "HyPrColoc" R software package was used to perform the colocalization analysis.

Results

Multi-omics integration SMR and colocalization analysis identify the association among blood *cis*-mQTLs, blood *cis*-eQTLs, and lung function

We screened the possible association of blood *cis*-mQTLs, blood *cis*-eQTLs, and lung function (FEV1, FVC, and FEV1/FVC) by multi-omics integration SMR (Fig. 2A-F). Then the gene loci with the same causal signal were identified by colocalization analysis. In concrete, we integrated *cis*-eQTLs from the eQTLGen Consortium ($n = 31,684$) and FEV1 GWAS summary statistics to result in 701 genes by SMR analysis (SMR $P_{FDR} < 0.05$, HEIDI $P > 0.05$) (Table S3). Meanwhile, the SMR analysis of *cis*-mQTLs from a meta-analysis of the BSGS and LBC data ($n = 1,980$) and FEV1 GWAS summary statistics resulted in 3913 CpG sites (SMR $P_{FDR} < 0.05$, HEIDI $P > 0.05$) (Table S4). Further integration of putative FEV1 causal *cis*-mQTLs and *cis*-eQTLs prioritized 2368 CpG sites near 186 genes. With the above multi-omics SMR method, there are 186 genes and 2368 CpG sites coequally involved in the association between *cis*-eQLT and FEV1 GWAS summary statistics, as well as *cis*-mQTL and FEV1 GWAS summary statistics (Table S5). We performed Bayesian colocalization analysis on these 186 genes to report the probability that the eQTLs and FEV1 GWAS shared the same variant, referred to as hypothesis 4 (PPH4) (Table S6). This analysis found that 28 of the 186 genes provided evidence of genetic colocalization based on a PPH4 ≥ 0.80, indicating that play important roles in FEV1. Similarly, we conducted multi-omics integration SMR and colocalization analysis on FVC and FEV1/FVC. The multi-omics integration SMR analysis found 192 genes were related to FVC. After colocalization analysis, 33 gene shares causality with FVC (Table S7, Table S8, Table S9, Table S10). For FEV1/FVC, the multi-omics SMR analysis and colocalization analysis found 19 genes were related to this association (Table S11, Table S12, Table S13, Table S14). As one of these robust results, we mapped the common genetic locus of DNA methylation and gene expression of *ZBTB38* by SMR multi-omics tool and colocalization map (Fig. 3). In addition, the multi-omics SMR mappings of *LMCD1* and lung function (FVC and FEV1/FVC) were shown in Figure S1 and Figure S2.

Replicate the results of multi-omics integration SMR and colocalization in multiple lung function data sets

In our previous findings, we performed a verification of existing publicly available GWAS data on lung function in the European population using SMR analysis and colocalization analysis. Our focus was on lung function data that were suitable for SMR and Bayesian colocalization analyses, while excluding datasets that did not meet the required criteria for these analyses (Table S15, Table S16, and Table S17 for replication sets details). Our analysis successfully replicated causal signals for eight genes associated with FEV1: *EML3*, *FAIM*, *PABPC4*, *RASGRP3*, *ROM1*, *SNIP1*, *UBXN2A*, and *ZBTB38* (SMR $P_{FDR} < 0.05$, HEIDI $P > 0.05$, and PPH4 ≥ 0.80). Additionally, we confirmed findings for five genes related to FVC:

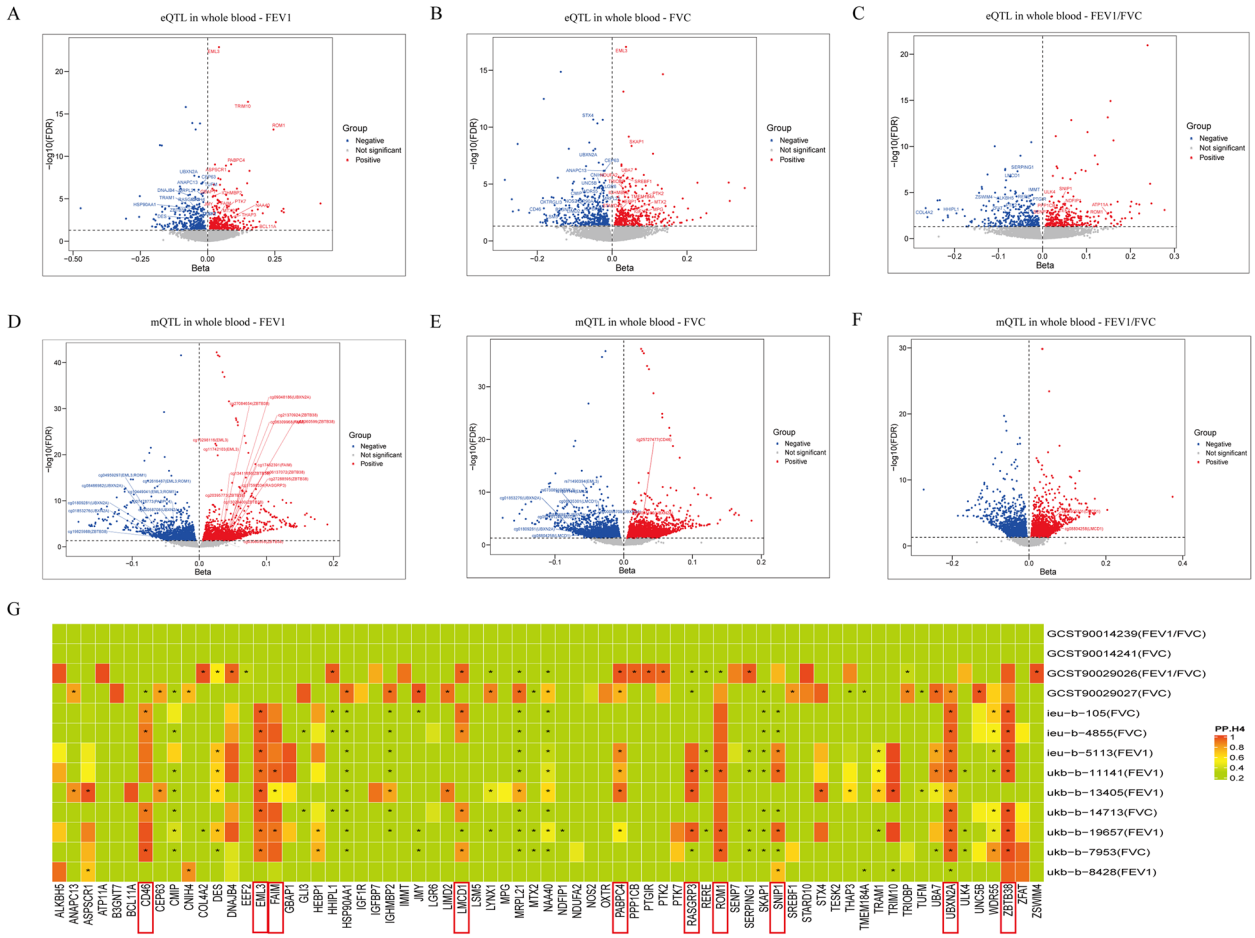


Fig. 2 SMR and colocalization analysis of discovery sets and replication sets. Volcano map shows the results of SMR analysis of mQTL/eQTL to lung function discovery sets (FEV1, FVC, and FEV1/FVC); FDR < 0.05 and beta < 0 are negatively correlated; FDR < 0.05 and beta > 0 are positively correlated; FDR > 0.05 is no significance. The thermogram shows the colocalization results of gene expression and GWAS data of lung function replication sets; The color gradient represents the value of PPH4; Robust results obtained in the replication sets are highlighted in red boxes. **A** SMR from eQTLs in blood to FEV1; **B** SMR from eQTLs in blood to FVC; **C** SMR from eQTLs in blood to FEV1/FVC; **D** SMR from mQTLs in blood to FEV1; **E** SMR from mQTLs in blood to FVC; **F** SMR from mQTLs in blood to FEV1/FVC; **G** colocalization of gene expression and 13 replication lung function data; *: pass all three steps of multi-omics SMR (all three steps FDR < 0.05); Beta: effect value of SMR analysis; FDR: false discover rate; PPH4: posterior probability of hypothesis 4

CD46, *EML3*, *LMCD1*, *UBXN2A*, and *ZBTB38* (SMR $P_{FDR} < 0.05$, HEIDI $P > 0.05$, and PPH4 ≥ 0.80). Moreover, we identified *LMCD1* associated with the FEV1/FVC (SMR $P_{FDR} < 0.05$, HEIDI $P > 0.05$, and PPH4 ≥ 0.80). The results of replication are shown in Fig. 2G.

The shared causal effect between gene expression, blood cell traits, and lung function

For the intricate nature of blood components, we hypothesize that eQTLs in blood may impact lung function by regulating blood cells. To explore this, we performed a multi-trait colocalization analysis involving eQTLs from the eQTLGen Consortium ($n = 31{,}684$), blood cell traits from Blood Cell Consortium ($n = 563{,}946$), and lung function traits (Table S18). Our findings revealed that a shared causal signal influenced the expression of the *CD46* in blood, monocyte count (PPA = 0.87), white blood cell count (PPA = 0.87), and lung function (Fig. 4A and B). Additionally, *LMCD1* expression in blood, white blood cell count, and the FEV1/FVC ratio were also associated with the same causal signal (PPA = 0.83) (Fig. 4C). Moreover, *ZBTB38* expression, lymphocyte count, and FEV1 were also mediated by a common causal signal (PPA = 0.91) (Fig. 4D).

Identify the association between methylation/transcription in diverse tissues and lung function by SMR analysis

To provide a broader explanation of how DNA methylation and gene expression regulate lung function in human tissues, we identified the associations between

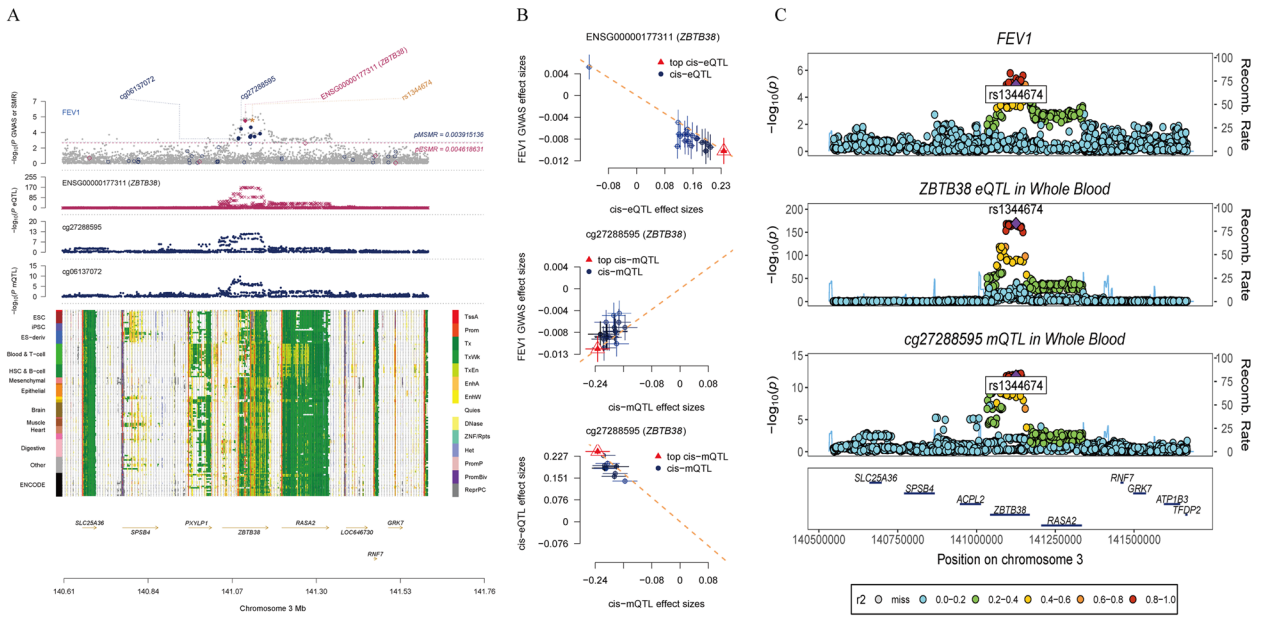


Fig. 3 Drawing SMR multi-omics integration map and multi-omics colocalization map, taking *ZBTB38* as an example. **A** Multi-omics integration map of SMR analysis; **B** Correlation of methylation, gene expression, and FEV1, respectively; **C** multi-omics colocalization map

mQTLs in 8 diverse tissues and eQTLs in 48 diverse tissues with lung function traits by SMR analysis (Figure S3, Figure S4, and Figure S5). The results showed that DNA methylation and gene expression in four tissues except for blood jointly affect lung functional traits: transverse colon, lung, skeletal muscle, and prostate (Fig. 5). Among them, the DNA methylation and expression of *UXNB2A* in transverse colon, lung, and skeletal muscle were related to FEV1 and FVC.

Colocalization of tissue-specific *cis*-eQTLs and lung function

Considering the intricate nature of lung function traits, we propose that diverse tissue-specific eQTLs may collectively influence lung function. To examine the impact of gene expression levels across various tissues and organs on lung function, we conducted a colocalization analysis of homologous *cis*-eQTLs and lung function traits in 48 different tissues (Fig. 6A). Our findings revealed that the causal relationships between the expression of seven genes (*EML3*, *UBXN2A*, *ROM1*, *ZBTB38*, *RASGRP3*, *FAIM*, and *PABPC4*) and FEV1 were consistently replicated across distinct tissues (PPH4 ≥ 0.80). Additionally, we observed that the causal signals linking the expression of three genes (*CD46*, *EML3*, and *UBXN2A*) to FVC were also replicated across various tissues (PPH4 ≥ 0.80). In short, *EML3 cis*-eQTLs were identified in 28 different human tissues, *FAIM cis*-eQTLs in 6 tissues, *PABPC4 cis*-eQTLs in 1 tissue, *RASGRP3 cis*-eQTLs in 1 tissue, *ROM1 cis*-eQTLs in 16 tissues, *UBXN2A cis*-eQTLs in 11 tissues, and *ZBTB38 cis*-eQTLs in 13 tissues, all demonstrating the same causal signals with FEV1. For FVC, *CD46 cis*-eQTLs were observed in 14 tissues, *EML3 cis*-eQTLs in 27 tissues, and *UBXN2A cis*-eQTLs in 13 tissues, all sharing similar causal signals. Details of the colocalization of gene expression in 48 diverse tissues and lung function are shown in Table S21.

In addition, we used FUMA GENE2FUNC to generate a gene expression heat map to test if the identified genes that shared causal signals with lung function are expressed in diverse human tissues based on the GTEx v8 project (Fig. 6B). All these genes are relatively highly expressed in the lung. *CD46*, *EML3*, and *PABPC4* are relatively highly expressed in almost 49 tissues. *UBXN2A* is relatively moderately expressed in transverse colon, lung, skeletal muscle, and prostate. However, *LMCD1* and *ZBTB38* are relatively low expressed in whole blood.

Cross-tissue eQTLs colocalization identified the causal association between the cardiopulmonary axis and lung function

Lung function is influenced by the condition of various tissues and organs throughout the body. To identify shared causal signals between multi-tissue eQTLs and lung function, we utilized multi-traits colocalization analysis. The SMR and colocalization analysis indicated that increased levels of FEV1 and FVC were associated with higher expression levels of *EML3* in aorta (FEV1: $Beta_{SMR} = 0.12$, 95%CI 0.08–0.16, PPH4 = 0.95; FVC: $Beta_{SMR} = 0.11$, 95%CI 0.07–0.15, PPH4 = 0.96), atrial

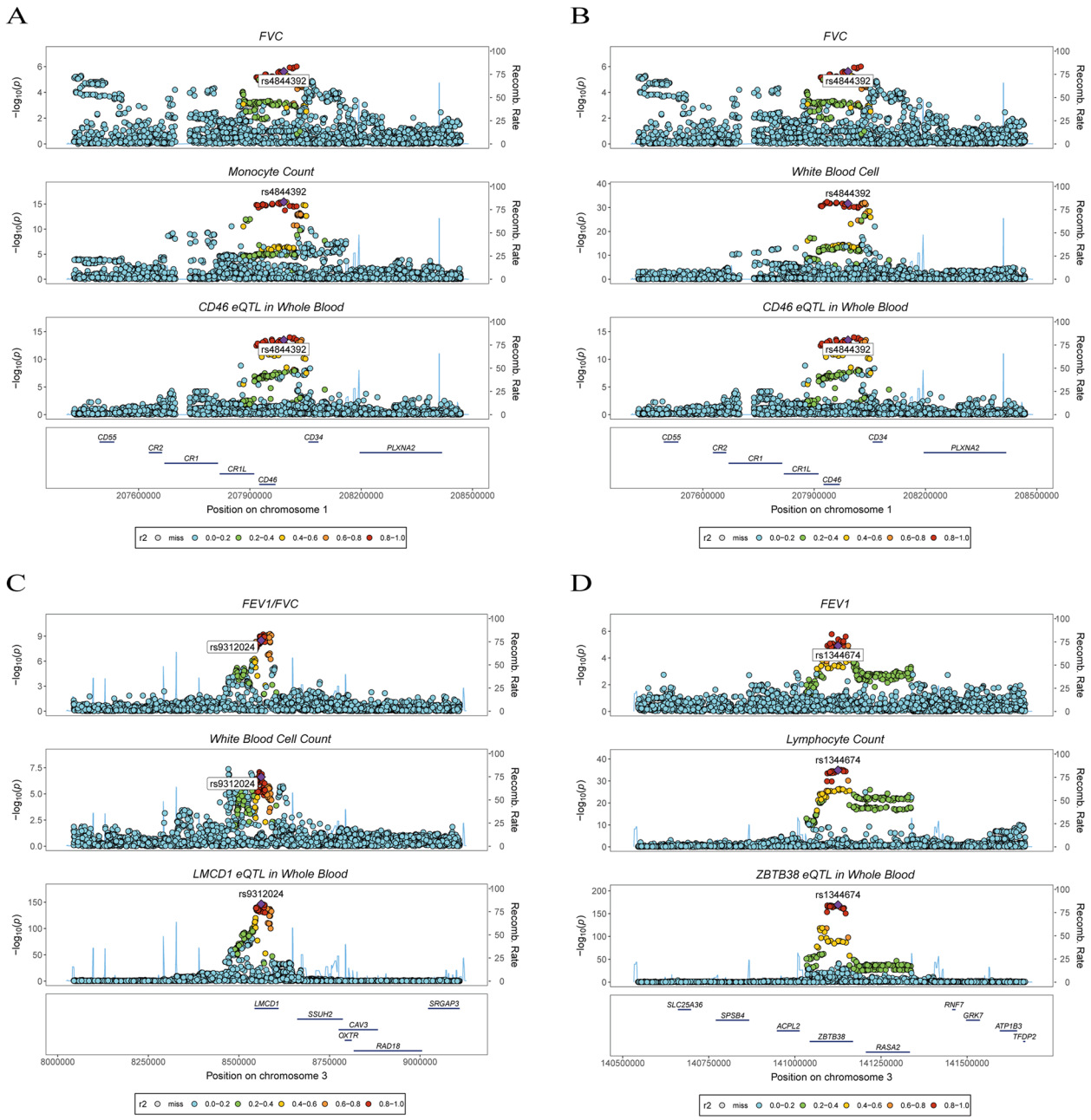


Fig. 4 MOLOC colocalization analysis identified the association signals of gene expression, blood cell traits, and lung function. **A** *CD46* expression in blood, monocyte count, and FVC shared the same causal signal (PPA = 0.87); **B** *CD46* expression in blood, white blood cell count, and FVC shared the same causal signal (PPA = 0.87); **C** *LMCD1* expression in blood, white blood cell count, and FEV1/FVC shared the same causal signal (PPA = 0.83); **D** *ZBTB38* expression in blood, lymphocyte count, and FEV1 shared the same causal signal (PPA = 0.91). eQTL: expression quantitative trait loci; FEV1: Forced expiratory volume in one second; FVC: Forced vital capacity; FEV1/FVC: Forced expiratory volume in one second and forced vital capacity ratio; PPA: An overall posterior probability of the three traits (PPA ≥ 0.80 was considered to share causal signals)

appendage (FEV1: $Beta_{SMR}$ = 0.14, 95%CI 0.10–0.18, PPH4 = 0.96; FVC: $Beta_{SMR}$ = 0.12, 95%CI 0.08–0.16, PPH4 = 0.98), left ventricle (FEV1: $Beta_{SMR}$ = 0.20, 95%CI 0.13–0.27, PPH4 = 0.94; FVC: $Beta_{SMR}$ = 0.18, 95%CI 0.11–0.24, PPH4 = 0.98), and lung (FEV1: $Beta_{SMR}$ = 0.12, 95%CI 0.08–0.16, PPH4 = 0.95; FVC: $Beta_{SMR}$ = 0.10, 95%CI 0.07–0.13, PPH4 = 0.94) (Fig. 7B). We discovered that *EML3* eQTLs in aorta, atrial appendage, left ventricle, and lung jointly shared the same causal signal with lung function, which were mediated by SNPs such as

A

SMR analysis between cis-eQTLs in four diverse tissues and lung function traits

Tisues	Traits	Genes	Probe.ID	Chr	Top.SNP	Beta(95%CI)	p_SMR	FDR
Colon_Transverse	FEV1	EML3	ENSG00000149499	11	rs2849030	0.11(0.07 to 0.14)	8.67e−11	2.05e−08
		PABPC4	ENSG00000090621	1	rs755249	0.09(0.05 to 0.12)	1.11e−06	7.57e−05
		UBXN2A	ENSG00000173960	2	rs4665238	−0.08(−0.11 to −0.04)	7.56e−06	3.51e−04
	FVC	EML3	ENSG00000149499	11	rs2849030	0.09(0.06 to 0.12)	6.88e−10	2.08e−07
		UBXN2A	ENSG00000173960	2	rs4665238	−0.08(−0.11 to −0.04)	1.40e−05	7.57e−04
Lung	FEV1	ROM1	ENSG00000149489	11	rs1801144	−0.07(−0.09 to −0.05)	2.46e−13	9.06e−11
		UBXN2A	ENSG00000173960	2	rs4665244	−0.08(−0.12 to −0.05)	7.76e−06	4.08e−04
	FVC	UBXN2A	ENSG00000173960	2	rs4665244	−0.08(−0.12 to −0.05)	5.92e−06	3.83e−04
Muscle_Skeletal	FEV1	EML3	ENSG00000149499	11	rs2428551	0.16(0.11 to 0.21)	5.48e−12	1.59e−09
		FAIM	ENSG00000158234	3	rs654528	0.03(0.01 to 0.04)	2.95e−05	1.06e−03
		PABPC4	ENSG00000090621	1	rs61232586	0.09(0.06 to 0.12)	2.97e−07	2.55e−05
		UBXN2A	ENSG00000173960	2	rs6751857	−0.08(−0.11 to −0.05)	2.35e−07	2.05e−05
	FVC	UBXN2A	ENSG00000173960	2	rs6751857	−0.08(−0.11 to −0.05)	1.46e−07	1.60e−05
Prostate	FEV1	EML3	ENSG00000149499	11	rs11553576	0.12(0.07 to 0.16)	5.22e−08	4.57e−06
	FVC	EML3	ENSG00000149499	11	rs11553576	0.10(0.07 to 0.14)	1.38e−07	1.76e−05

B

SMR analysis between cis-mQTLs in four diverse tissues and lung function traits

Tisues	Traits	Genes	Probe.ID	Chr	Top.SNP	Beta(95%CI)	p_SMR	FDR
Colon Transverse	FEV1	RASGRP3	cg01109219	2	rs62147110	0.01(0.01 to 0.02)	3.17e−05	2.45e−03
		RASGRP3	cg17598334	2	rs4077711	0.01(0.01 to 0.02)	1.42e−05	1.37e−03
		UBXN2A	cg04619854	2	rs72788206	−0.01(−0.02 to −0.01)	1.84e−04	9.17e−03
	FVC	LMCD1	cg06266100	3	rs13075586	−0.02(−0.04 to −0.01)	1.21e−03	3.34e−02
		UBXN2A	cg04619854	2	rs72788206	−0.01(−0.02 to −0.00)	1.68e−03	4.15e−02
	FEV1FVC	LMCD1	cg06266100	3	rs13075586	0.04(0.02 to 0.05)	1.90e−05	1.56e−03
Lung	FEV1	RASGRP3	cg17598334	2	rs4077711	0.01(0.01 to 0.02)	1.58e−05	1.45e−03
		RASGRP3	cg18200741	2	rs10182608	−0.03(−0.05 to −0.01)	2.60e−03	5.21e−02
		UBXN2A	cg04619854	2	rs72788206	−0.01(−0.02 to −0.01)	1.22e−04	6.46e−03
		UBXN2A	cg08212172	2	rs4665651	0.02(0.01 to 0.03)	1.10e−08	4.66e−06
	FVC	CD46	cg25727477	1	rs7144	0.01(0.00 to 0.02)	1.39e−03	3.49e−02
		LMCD1	cg06266100	3	rs6806939	−0.02(−0.03 to −0.01)	6.67e−04	2.16e−02
		UBXN2A	cg04619854	2	rs72788206	−0.01(−0.02 to −0.00)	1.36e−03	3.45e−02
		UBXN2A	cg08212172	2	rs4665651	0.02(0.01 to 0.03)	1.41e−08	7.57e−06
	FEV1FVC	LMCD1	cg06266100	3	rs6806939	0.04(0.02 to 0.05)	2.78e−06	3.23e−04
Muscle Skeletal	FEV1	UBXN2A	cg08212172	2	rs61463699	0.02(0.01 to 0.02)	8.06e−06	1.11e−03
	FVC	UBXN2A	cg08212172	2	rs61463699	0.02(0.01 to 0.02)	8.64e−06	1.24e−03
Prostate	FEV1	RASGRP3	cg01109219	2	rs62147110	0.01(0.01 to 0.02)	5.00e−05	3.98e−03
		RASGRP3	cg17598334	2	rs4077711	0.01(0.01 to 0.02)	4.86e−05	3.91e−03
	FVC	CD46	cg25727477	1	rs7144	0.01(0.00 to 0.02)	2.04e−03	4.60e−02

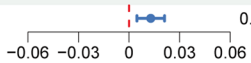


Fig. 5 SMR analysis showed that gene expression and DNA methylation in four tissues were involved in the regulation of lung function. **A** SMR analysis between eQTLs in four diverse tissues and lung function traits; **B** SMR analysis between mQTLs in four diverse tissues and lung function traits. Chr: chromosome; Beta: effect value of SMR analysis; FDR: false discover rate

rs2849031 (PPMT = 0.54) (Fig. 7C). In short, SNPs such as rs2849031 promote the expression of *EML3* in aorta, atrial appendage, left ventricle, and lung to jointly contribute to FEV1 and FVC (Fig. 7A).

Discussion

Tissue-specific DNA methylation and gene expression quantitative trait locus could provide new insights for exploring the genetic mechanism of complex diseases and traits in post-GWAS analysis. In this research, we performed multi-omics SMR analysis and Bayesian colocalization analysis to integrate lung function GWAS with *cis*-eQTLs and *cis*-mQTLs from diverse human tissues. We identified eight genes in whole blood that were associated with FEV1: *EML3*, *FAIM*, *PABPC4*, *RASGRP3*, *ROM1*, *SNIP1*, *UBXN2A*, and *ZBTB38*. Additionally, five genes in whole blood were associated with FVC: *CD46*, *EML3*, *LMCD1*, *UBXN2A*, and *ZBTB38*. Furthermore,

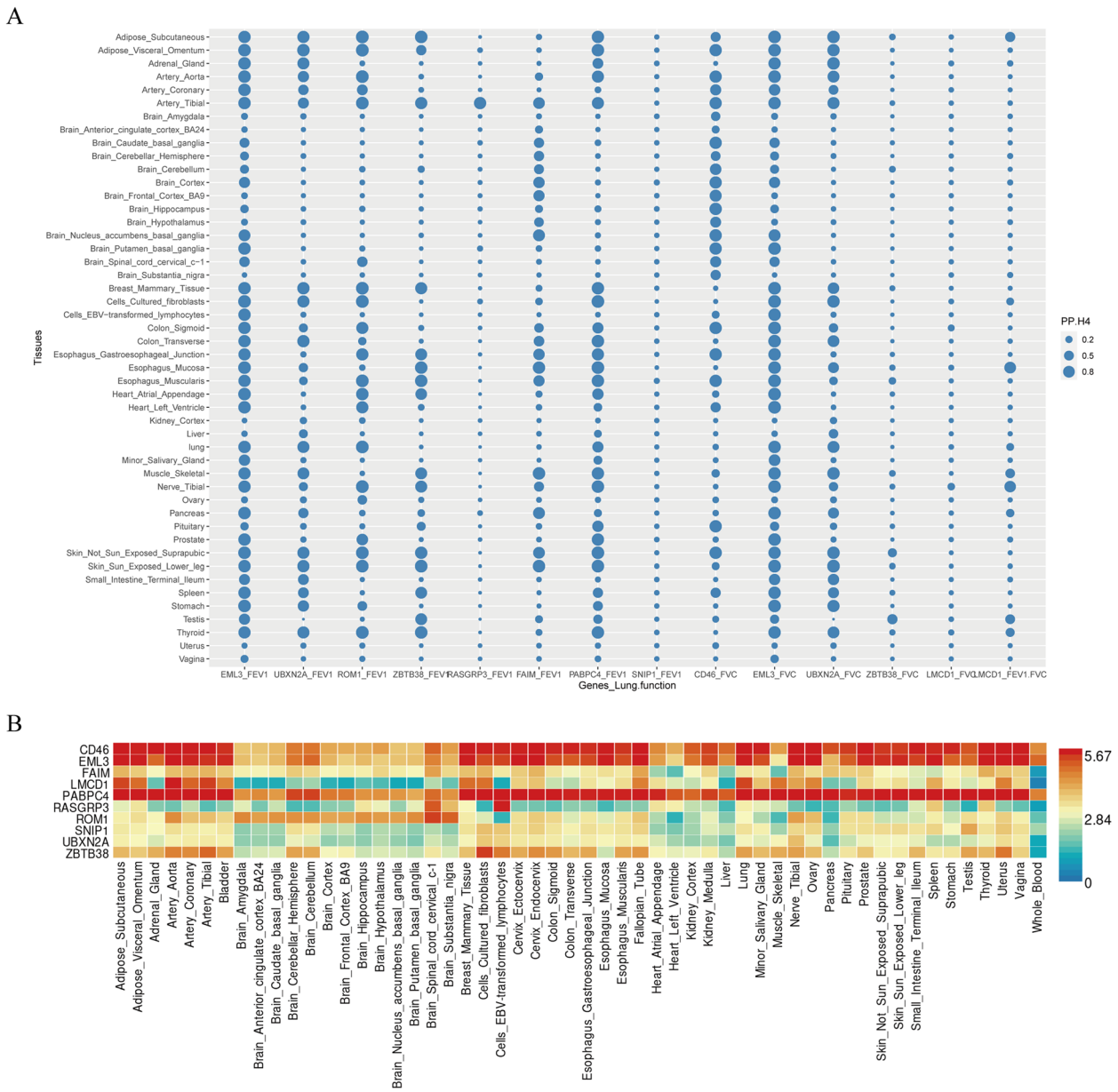
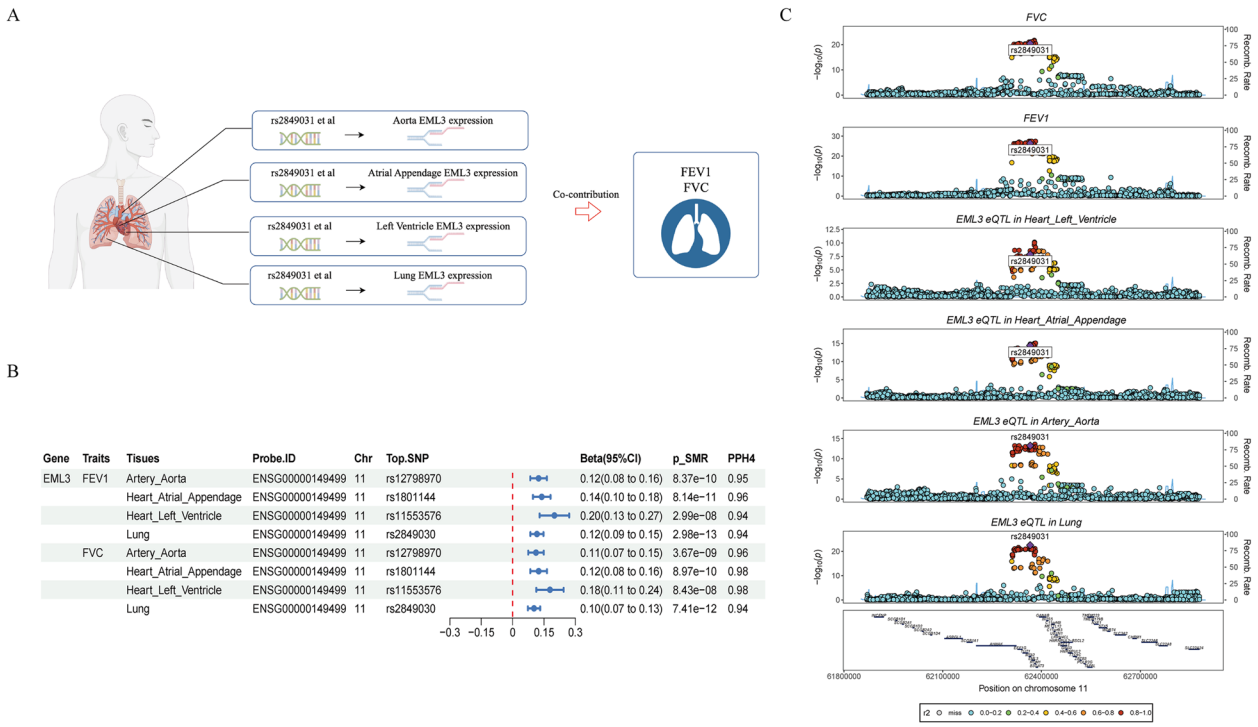


Fig. 6 The causal association between cross-tissue gene expression and lung function. **A** Colocalization map of tissue-specific eQTLs and lung function; The size of the point represents the value of PPH4. **B** The FUMA GENE2FUNC tool generated a gene expression heatmap in diverse human tissues of 10 genes associated with lung function; The heatmap displays zero mean normalized log2 transformed expression values; deeper shades of red signify higher expression levels, while deeper shades of blue indicate lower expression levels. PPH4: Posterior probability of hypothesis 4

one gene, *LMCD1*, was associated with the FEV1/FVC, supported by multi-omics evidence. In the SMR analysis of tissue-specific eQTLs/mQTLs and lung function, the multi-omics evidence for *UBXN2A* was replicated in tissues such as the lung, colon, and skeletal muscle. Moreover, the causal relationships between eight genes (*CD46*, *EML3*, *FAIM*, *PABPC4*, *RASGRP3*, *ROM1*, *UBXN2A*, and *ZBTB38*) and lung function were confirmed through colocalization analyses of tissue-specific cis-eQTLs, showing replication in at least one human tissue other than whole blood.

Experimental studies have indicated that *CD46* (Complement regulatory protein), as a Treg-activated costimulatory molecule, may help to inhibit asthma inflammation together with the production of IL-10/granzyme B [27]. The higher expression of *CD46* protects former smokers from emphysema and chronic obstructive pulmonary disease by eliminating inflammation that hinders the



Gene	Traits	Tisues	Probe.ID	Chr	Top.SNP	Beta(95%CI)	p_SMR	PPH4
EML3	FEV1	Artery_Aorta	ENSG00000149499	11	rs12798970	0.12(0.08 to 0.16)	8.37e−10	0.95
		Heart_Atrial_Appendage	ENSG00000149499	11	rs1801144	0.14(0.10 to 0.18)	8.14e−11	0.96
		Heart_Left_Ventricle	ENSG00000149499	11	rs11553576	0.20(0.13 to 0.27)	2.99e−08	0.94
		Lung	ENSG00000149499	11	rs2849030	0.12(0.09 to 0.15)	2.98e−13	0.94
	FVC	Artery_Aorta	ENSG00000149499	11	rs12798970	0.11(0.07 to 0.15)	3.67e−09	0.96
		Heart_Atrial_Appendage	ENSG00000149499	11	rs1801144	0.12(0.08 to 0.16)	8.97e−10	0.98
		Heart_Left_Ventricle	ENSG00000149499	11	rs11553576	0.18(0.11 to 0.24)	8.43e−08	0.98
		Lung	ENSG00000149499	11	rs2849030	0.10(0.07 to 0.13)	7.41e−12	0.94

Fig. 7 Cross-tissue *EML3* expression jointly affects lung function through the cardiopulmonary axis. **A** Schematic diagram of common genetic loci promoting the expression of *EML3* in the cardiopulmonary axis and then affecting the expression of lung function (Drawing by Figdraw). **B** SMR and colocalization analysis showed that the expression of *EML3* in aorta, atrial appendage, left ventricle, and lung respectively had a causal association with FEV1 and FVC. **C** HyPrColoc colocalization analysis showed that *EML3* in aorta, atrial appendage, left ventricle, and lung jointly had a causal association with FEV1 and FVC. Chr: chromosome; Beta: effect value of SMR analysis; p_SMR: *P* value of SMR analysis; PPH4: posterior probability of hypothesis 4

proliferation and necrosis of CD8(+) T cells [28]. Our findings further substantiated the importance of *CD46* in respiratory health, demonstrating that genetically predicted levels of *CD46* methylation and expression in whole blood, lung, and prostate tissues were significantly associated with FVC. Notably, *CD46* expression in whole blood, monocyte count, and FVC shared the same causal signal (PPA = 0.87, Fig. 4A). These results underscore the potential of *CD46* as a novel immune target for respiratory diseases and lung function. However, the underlying mechanisms through which *CD46* exerts its effects remain to be thoroughly investigated. Future research should focus on elucidating these mechanisms in practical settings to validate the therapeutic potential of *CD46* in the context of respiratory disease management and prevention.

EML3 (Echinoderm Microtubule Associated Protein Like 3) is a nuclear microtubule-binding protein, which is necessary for the correct arrangement of chromosomes in metaphase [29]. As reported, *EML3* is related to biological processes such as mitosis, embryonic development, and methylation regulation [29–31]. Additionally, a Mendelian randomized analysis on smoking-related DNA methylation and lung function showed that cg12616487 (*EML3*) had a causal relationship with FEV1 [32]. A latest genome-wide association study of preserved ratio impaired spirometry (PRISm), defined as FEV1 < 80% predicted and FEV1/FVC ratio ≥ 0.70, showed that *EML3* is a heritable component for the development of PRISm [33]. In our research, the DNA methylation and gene expression of *EML3* are causally related to FEV1 and FVC. Moreover, *EML3 cis*-eQTLs in 28 diverse human tissues shared the same causal signals with FEV1, and *EML3 cis*-eQTLs in 27 diverse human tissues shared the same causal signals with FVC. An observational study has reported that decreased FEV1 was associated with the risk of coronary artery disease (CAD) [34]. However, a Mendelian randomized analysis found that FVC, not FEV1, was independently related to the risk of CAD [35]. Notably, we identified the causal association between the cardiopulmonary axis of *EML3 cis*-eQTLs and FEV1 or FVC (PPMT = 0.54, Fig. 7), which seemingly explains the causal relationship between lung function and cardiopulmonary diseases from a cross-tissue perspective.

LMCD1 (LIM and cysteine-rich domains 1, also named Dyxin) is a member of the LIM domain family of zinc finger proteins that act as transcription co-regulators. *LMCD1* plays a key regulatory role in a variety of cell functions, such as gene expression, cell growth, cell adhesion, cell differentiation, and cytoskeleton remodeling [36, 37]. *LMCD1* has been reported to be involved in pulmonary fibrosis in systemic sclerosis-related interstitial lung disease [38]. Previously, genome-wide association studies found that *LMCD1* was the key genetic locus of mitral valve prolapse [39, 40]. Findings from our multi-omics investigation suggested that DNA methylation of cg08935301 and cg08804258 in whole blood may decrease the gene expression of *LMCD1* in whole blood and thus decrease the level of FVC, while DNA methylation of cg24259363 in whole blood may increase the gene expression of *LMCD1* in whole blood and thus increase the level of FVC (Figure S1). Meanwhile, the hypermethylation levels of cg08935301 and cg08804258 and the underexpression of *LMCD1* in whole blood are protective factors for FEV1/FVC, while the hypermethylation level of cg24259363 and the overexpression of *LMCD1* in whole blood are damaging factors for FEV1/FVC (Figure S2). Moreover, we found that *LMCD1* gene expression in whole blood, white blood cell count, and FEV1/FVC shared the same causal signal (PPA = 0.83, Fig. 4C). In brief, *LMCD1* is a potential target of FVC and FEV1/FVC, and further regulation mechanisms need to be verified in practice.

UBXN2A (Ubiquitin-like-domain-containing protein 2A) serves as a ubiquitin-like protein-coding gene to be involved in several processes, including autophagosome assembly, nuclear membrane reassembly, and proteasome-mediated ubiquitin-dependent protein catabolic process. In colon cancer, *UBXN2A* was reported to be able to competitively bind mot-2 with *p53*, thus saving *p53* tumor suppressor function and reconstructing the inactivated *p53*-dependent apoptosis pathway [41]. According to our research, DNA methylations of cg04619854 in transverse colon, cg04619854 in lung, cg08212172 in lung, cg08212172 in skeletal muscle, cg01809281 in whole blood, cg01853276 in whole blood, cg00058708 in whole blood, cg08466982 in whole blood, and cg09048186 in whole blood were associated with FEV1 and FVC. Meanwhile, *UBXN2A* expressions in transverse colon, lung, skeletal muscle, and whole blood were associated with FEV1 and FVC. It seems to suggest the importance of *UBXN2A* in the lung-colon axis for the respiratory system.

ZBTB38 (Zinc finger and BTB domain containing 38) is a methyl-CpG binding protein, binding to and repressing methylated DNA. In a multi-omics analysis, *ZBTB38* mRNA expression and DNA methylation were found to be related to cisplatin resistance in non-small cell lung cancer [42]. Based on our findings, DNA methylation of cg27288595 may decrease the gene expression of *ZBTB38* and thus decrease the level of FEV1 (Fig. 3B). Furthermore, we found that *ZBTB38* gene expression in whole blood, lymphocyte count, and FEV1 shared the same causal signal (PPA = 0.91, Fig. 4D). For colocalization of tissue-specific *cis*-eQTLs and lung function, *ZBTB38 cis*-eQTLs shared the same causal signals with FEV1 in 13 diverse human tissues including lung, heart, and colon. Further studies are needed to better elucidate the role of *ZBTB38* in FEV1 and respiratory diseases.

In previous research, the genes *SNIP1*, *PABPC4*, *FAIM*, and *RASGRP3* have emerged as significant players in the context of lung cancer. Notably, *SNIP1* (Smad Nuclear Interacting Protein 1) has been shown to enhance the migration and invasion of lung cancer cells through its interaction with actin filament-associated protein 1 antisense RNA 1 (AFAP1 − AS1), which is a long noncoding RNA known for its role in cancer progression [43]. Additionally, *PABPC4* (Poly(A) Binding Protein Cytoplasmic 4) has been identified through quantitative proteomic analysis as a promising biomarker for lung adenocarcinoma, highlighting its potential as a target for therapeutic strategies [44]. A study showed that knocking out *FAIM* (Fas apoptosis inhibitor molecule) can induce autophagy in lung adenocarcinoma cells by inhibiting the mTOR pathway, suggesting a mechanism by which cancer cells evade apoptosis [45]. Furthermore, *RASGRP3* (RAS Guanyl Releasing Protein 3) has been implicated in promoting the risk of lung cancer, as evidenced by findings from SNP microarray analyses that evaluated genomic instability in airway epithelial cells [46]. This highlights *RASGRP3*'s potential role as a biomarker in lung cancer risk assessment. Taken together, these studies not only underscored the importance of these molecular targets in lung cancer development but also paved the way for further investigations into their mechanisms and potential therapeutic implications. As we delve deeper into the impact of these findings, we can better understand the relationships between these genes and lung function, ultimately leading to more effective interventions in lung cancer treatment and prevention.

One of the key strengths of our research lies in the integration of methylation and transcriptomic evidence through multi-omics SMR. This approach enhances the understanding of the relationships between individual omics components and lung function. By merging data from different omics layers, we were able to provide a more comprehensive perspective on the biological

mechanisms influencing lung health. Furthermore, we employed colocalization analysis to evaluate the causal effects of gene expression on lung function, which helps to mitigate potential confounding factors arising from linkage disequilibrium. This methodological rigor solidifies our findings, ensuring that the observed associations are more likely to reflect true biological relationships rather than artifacts of genetic correlation. To enhance the robustness of our results, we replicated the findings from both the multi-omics SMR and colocalization analysis across several large datasets of lung function data. This replication in diverse datasets added credibility to our conclusions and underscored the reliability of our findings. Additionally, we examined tissue-specific *cis*-eQTLs and *cis*-mQTLs to further elucidate the communication between different organs and tissues in the context of omics regulation. This exploration not only strengthens our understanding of how various tissues interact to influence lung function but also emphasizes the importance of a systems biology approach in studying complex traits like lung function. Overall, our research contributes to a more nuanced understanding of the regulatory networks that underpin lung function, providing new insight for future studies aimed at exploring the intricate interplay between genetic and epigenetic factors of lung function.

Our research does have several limitations that should be acknowledged. First, the impact of racial and ethnic differences on the genetic determinants of lung function cannot be overlooked. The majority of large genome-wide association studies on lung function have primarily focused on populations of European descent. Consequently, the generalizability of our findings to other populations may require further investigation and discussion to ensure broader applicability. Second, our analysis was limited to *cis*-eQTLs. While these provide valuable insight into local regulatory variations, *trans*-eQTLs have the potential to exert widespread influence on gene regulatory networks that could be equally significant but were not addressed in this study. Future research should consider examining *trans*-eQTLs to gain a more comprehensive understanding of gene regulation. Third, Bayesian colocalization analysis is a powerful tool for identifying shared genetic signals between traits; however, it is important to note that this method does not confirm causality. While colocalization can suggest that traits may be influenced by the same genetic variant, it does not establish a direct causal relationship. Further experimental validation and complementary approaches are necessary to elucidate the causal mechanisms underlying these associations. Fourth, the exploration of the effects of methylation on gene expression was limited to just nine diverse tissues, constrained by the availability of mQTLs data from other tissues. This limitation restricts our ability to fully elucidate the relationships between methylation and gene expression across a wider range of biological contexts. Expanding the dataset to include additional tissues could provide a more thorough understanding of the implications of methylation in gene regulation and lung function. Finally, despite the advancements in understanding the genetic and epigenetic factors influencing lung function, there remains a significant gap in identifying mQTLs with single-cell resolution. Current studies predominantly focus on bulk tissue analyses, which can obscure the heterogeneity present within cellular populations. The lack of single-cell resolution in mQTL studies limits our ability to associate specific methylation changes with cellular phenotypes and functions accurately.

Conclusion

In this study, the causal effect among tissue-specific DNA methylation, tissue-specific gene expression, and lung function were identified through multi-omics integration analysis. Leveraging methylation and transcriptome genetic databases, we identified several genes associated with lung function and explained them in terms of DNA methylation and gene expression from diverse human tissues. This not only provides multi-omics evidence for the regulation mechanism of lung function but also provides a new insight into the influence of crosstalk between organs and tissues on lung function. Furthermore, our study identifies potential molecular targets that could facilitate cross-tissue therapeutic strategies aimed at improving lung function and addressing lung diseases in the future.

Abbreviations

FEV1	Forced expiratory volume in one second
FVC	Forced vital capacity
FEV1/FVC	Forced expiratory volume in one second and forced vital capacity ratio
COPD	Chronic obstructive pulmonary disease
GWAS	Genome-wide association study
eQTL	Expression quantitative trait loci
mQTL	DNA methylation quantitative trait loci
GTEx	Genotype-Tissue Expression
LD	Linkage disequilibrium
SNP	Single nucleotide polymorphism
SMR	Summary data-based Mendelian randomization
HEIDI	Heterogeneity in dependent instruments
FDR	False discover rate
MHC	Major Histocompatibility Complex
PPH4	Posterior probability of hypothesis 4
PPA	An overall posterior probability
PPMT	Posterior probability of multiple traits
COLOC	Bayesian colocalization
MOLOC	Multiple-trait-coloc
FUMA	Functional mapping and annotation of genetic associations
HyPrColoc	Hypothesis Prioritisation for multi-trait Colocalization

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11476-2.

Supplementary Material 1: Table S1. The basic information of lung function replication datasets. Table S2. Data resources of cis-mQTLs, cis-eQTLs, and blood cell traits. Table S3. Summary-based Mendelian randomization (SMR) analysis from blood gene expression to FEV1 (P HEIDI > 0.05, FDR < 0.05). Table S4. SMR analysis from blood DNA methylation to FEV1 (P HEIDI > 0.05, FDR < 0.05). Table S5. SMR analysis from blood DNA methylation to blood gene expression (associated with FEV1) (P HEIDI > 0.05, FDR < 0.05). Table S6. Colocalization analysis between gene expression and FEV1. Table S7. SMR analysis from blood gene expression to FVC (P HEIDI > 0.05, FDR < 0.05). Table S8. SMR analysis from blood DNA methylation to FVC (P HEIDI > 0.05, FDR < 0.05). Table S9. SMR analysis from blood DNA methylation to blood gene expression (associated with FVC) (P HEIDI > 0.05, FDR < 0.05). Table S10. Colocalization analysis between gene expression and FVC. Table S11. SMR analysis from blood gene expression to FEV1/FVC (P HEIDI > 0.05, FDR < 0.05). Table S12. SMR analysis from blood DNA methylation to FEV1/FVC (P HEIDI > 0.05, FDR < 0.05). Table S13. SMR analysis from blood DNA methylation to blood gene expression (associated with FEV1/FVC) (P HEIDI > 0.05, FDR < 0.05). Table S14. Colocalization analysis between gene expression and FEV1/FVC. Table S15. SMR analysis from blood DNA methylation to 13 replication sets of lung function (P HEIDI > 0.05, FDR < 0.05). Table S16. SMR analysis from blood gene expression to 13 replication sets of lung function. Table S17. Colocalization analysis between gene expression and 13 replication sets of lung function. Table S18. Colocalization analysis between gene expression and blood cell traits. Table S19. SMR analysis from tissue-specific DNA methylation to lung function (FDR < 0.05). Table S20. SMR analysis from tissue-specific gene expression to lung function (FDR < 0.05). Table S21. Colocalization analysis between tissue-specific gene expression and lung function (PPH4 ≥ 0.80)

Supplementary Material 2: Figure S1. Drawing SMR multi-omics integration map of LMCD1 and FVC. A: Multi-omics integration map of SMR analysis; B: Correlation of methylation of cg08935301, gene expression of LMCD1, and FVC, respectively; C: Correlation of methylation of cg08804258, gene expression of LMCD1, and FVC, respectively; D: Correlation of methylation of cg24259363, gene expression of LMCD1, and FVC, respectively. Figure S2. Drawing SMR multi-omics integration map of LMCD1 and FEV1/FVC. A: Multi-omics integration map of SMR analysis; B: Correlation of methylation of cg08935301, gene expression of LMCD1, and FEV1/FVC, respectively; C: Correlation of methylation of cg08804258, gene expression of LMCD1, and FEV1/FVC, respectively; D: Correlation of methylation of cg24259363, gene expression of LMCD1, and FEV1/FVC, respectively. Figure S3. SMR analysis of tissue-specific eQTLs and FEV1. Figure S4. SMR analysis of tissue-specific eQTLs and FVC. Figure S5. SMR analysis of tissue-specific mQTLs and lung function (FEV1, FVC, and FEV1/FVC)

Acknowledgements
Not applicable.

Authors' contributions
SP: study design, conceptualization, drafting of the manuscript, and statistical analysis. JF and WM: collating and statistical data. GH and SW: conceptualization, drafting of the manuscript, and statistical analysis. All authors read and approved the final manuscript.

Funding
Not applicable.

Data availability
Discovery sets for FEV1, FVC, and FEV1/FVC were derived from Shrine et al. 2019 (PMID: 30,804,560). The 13 GWAS data replication sets (Table S1 for details) of lung function of European ancestry were from GWAS Catalog (https://www.ebi.ac.uk/gwas/downloads) and IEU OpenGWAS project (https://gwas.mrcieu.ac.uk/). The standard tools used in this study are available at Yang Lab (https://yanglab.westlake.edu.cn/software/smr/#SMR). Publicly available summary statistics of blood eQTLs and mQTLs were obtained from eQTLGen Consortium (http://www.eqtlgen.org/) and Yang Lab, respectively. Tissue-specific eQTLs data and tissue-specific mQTLs data come from Genotype-Tissue Expression (GTEx) (https://www.gtexportal.org/). The GWAS data of blood cell traits are from the Blood Cell Consortium (http://www.mhi-humangenetics.org/).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]The Tenth Affiliated Hospital, Southern Medical University, Dongguan People's Hospital, Dongguan, Guangdong 523059, China.

Received: 27 January 2025 Accepted: 12 March 2025
Published online: 24 March 2025

References

1. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for Prevention, Diagnosis and Management of COPD 2023 Report. https://goldcopd.org/2023-gold-report-2/.
2. Cosentino J, Behsaz B, Alipanahi B, et al. Inference of chronic obstructive pulmonary disease with deep learning on raw spirograms identifies new genetic loci and improves risk models. Nat Genet. 2023;55(5):787–95. https://doi.org/10.1038/s41588-023-01372-4.
3. Shrine N, Izquierdo AG, Chen J, et al. Multi-ancestry genome-wide association analyses improve resolution of genes and pathways influencing lung function and chronic obstructive pulmonary disease risk. Nat Genet. 2023;55(3):410–22. https://doi.org/10.1038/s41588-023-01314-0.
4. Shrine N, Guyatt AL, Erzurumluoglu AM, et al. New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. Nat Genet. 2019;51(3):481–93. https://doi.org/10.1038/s41588-018-0321-7.
5. Wyss AB, Sofer T, Lee MK, et al. Multiethnic meta-analysis identifies ancestry-specific and cross-ancestry loci for pulmonary function. Nat Commun. 2018;9(1):2976. https://doi.org/10.1038/s41467-018-05369-0.
6. Agusti A, Faner R. Lung function trajectories in health and disease. Lancet Respir Med. 2019;7(4):358–64. https://doi.org/10.1016/S2213-2600(18)30529-0.
7. Martino DJ, Bui DS, Li S, et al. Genetic and Epigenetic Associations with Pre-Chronic Obstructive Pulmonary Disease Lung Function Trajectories. Am J Respir Crit Care Med. 2023;208(10):1135–7. https://doi.org/10.1164/rccm.202306-1025LE.
8. Qiu W, Baccarelli A, Carey VJ, et al. Variable DNA methylation is associated with chronic obstructive pulmonary disease and lung function. Am J Respir Crit Care Med. 2012;185(4):373–81. https://doi.org/10.1164/rccm.201108-1382OC.
9. Lee M, Huan T, McCartney DL, et al. Pulmonary Function and Blood DNA Methylation: A Multiancestry Epigenome-Wide Association Meta-analysis. Am J Respir Crit Care Med. 2022;206(3):321–36. https://doi.org/10.1164/rccm.202108-1907OC.
10. Imboden M, Wielscher M, Rezwan FI, et al. Epigenome-wide association study of lung function level and its change. Eur Respir J. 2019;54(1):1900457. https://doi.org/10.1183/13993003.00457-2019.
11. Kitaba NT, Knudsen GTM, Johannessen A, et al. Fathers' preconception smoking and offspring DNA methylation. Clin Epigenetics. 2023;15(1):131. https://doi.org/10.1186/s13148-023-01540-7.
12. Farh KKH, Marson A, Zhu J, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature. 2015;518(7539):337–43. https://doi.org/10.1038/nature13835.

13. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48(5):481–7. https://doi.org/10.1038/ng.3538.
14. Wu Y, Zeng J, Zhang F, et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. Nat Commun. 2018;9(1):918. https://doi.org/10.1038/s41467-018-03371-0.
15. Võsa U, Claringbould A, Westra HJ, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat Genet. 2021;53(9):1300–10. https://doi.org/10.1038/s41588-021-00913-z.
16. McRae AF, Marioni RE, Shah S, et al. Identification of 55,000 Replicated DNA Methylation QTL. Sci Rep. 2018;8(1):17605. https://doi.org/10.1038/s41598-018-35871-w.
17. GTEx Consortium, Laboratory, Data Analysis &Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, et al. Genetic effects on gene expression across human tissues. *Nature*. 2017;550(7675):204–213. https://doi.org/10.1038/nature24277.
18. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*. 2020;369(6509):1318–1330. https://doi.org/10.1126/science.aaz1776.
19. Oliva M, Demanelis K, Lu Y, et al. DNA methylation QTL mapping across diverse human tissues provides molecular links between genetic variation and complex traits. Nat Genet. 2023;55(1):112–22. https://doi.org/10.1038/s41588-022-01248-z.
20. Chen MH, Raffield LM, Mousas A, et al. Trans-ethnic and Ancestry-Specific Blood-Cell Genetics in 746,667 Individuals from 5 Global Populations. Cell. 2020;182(5):1198-1213.e14. https://doi.org/10.1016/j.cell.2020.06.045.
21. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68–74. https://doi.org/10.1038/nature15393.
22. Xu S, Li X, Zhang S, et al. Oxidative stress gene expression, DNA methylation, and gut microbiota interaction trigger Crohn's disease: a multi-omics Mendelian randomization study. BMC Med. 2023;21(1):179. https://doi.org/10.1186/s12916-023-02878-8.
23. 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(5): e1004383. https://doi.org/10.1371/journal.pgen.1004383.
24. Giambartolomei C, Zhenli Liu J, Zhang W, et al. A Bayesian framework for multiple trait colocalization from summary association statistics. Bioinformatics. 2018;34(15):2538–45. https://doi.org/10.1093/bioinformatics/bty147.
25. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8(1):1826. https://doi.org/10.1038/s41467-017-01261-5.
26. Foley CN, Staley JR, Breen PG, et al. A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. Nat Commun. 2021;12(1):764. https://doi.org/10.1038/s41467-020-20885-8.
27. Tsai YG, Niu DM, Yang KD, et al. Functional defects of CD46-induced regulatory T cells to suppress airway inflammation in mite allergic asthma. Lab Invest. 2012;92(9):1260–9. https://doi.org/10.1038/labinvest.2012.86.
28. Grumelli S, Lu B, Peterson L, Maeno T, Gerard C. CD46 protects against chronic obstructive pulmonary disease. PLoS ONE. 2011;6(5): e18785. https://doi.org/10.1371/journal.pone.0018785.
29. Tegha-Dunghu J, Neumann B, Reber S, et al. EML3 is a nuclear microtubule-binding protein required for the correct alignment of chromosomes in metaphase. J Cell Sci. 2008;121(Pt 10):1718–26. https://doi.org/10.1242/jcs.019174.
30. Wang HD, Hou QF, Guo QN, et al. DNA methylation study of fetus genome through a genome-wide analysis. BMC Med Genomics. 2014;7:18. https://doi.org/10.1186/1755-8794-7-18.
31. Elbere I, Silamikelis I, Ustinova M, et al. Significantly altered peripheral blood cell DNA methylation profile as a result of immediate effect of metformin use in healthy individuals. Clin Epigenetics. 2018;10(1):156. https://doi.org/10.1186/s13148-018-0593-x.
32. Jamieson E, Korologou-Linden R, Wootton RE, et al. Smoking, DNA Methylation, and Lung Function: a Mendelian Randomization Analysis to Investigate Causal Pathways. Am J Hum Genet. 2020;106(3):315–26. https://doi.org/10.1016/j.ajhg.2020.01.015.
33. Higbee DH, Lirio A, Hamilton F, et al. Genome-wide association study of preserved ratio impaired spirometry (PRISm). Eur Respir J. 2024;63(1):2300337. https://doi.org/10.1183/13993003.00337-2023.
34. Schroeder EB, Welch VL, Couper D, et al. Lung function and incident coronary heart disease: the Atherosclerosis Risk in Communities Study. Am J Epidemiol. 2003;158(12):1171–81. https://doi.org/10.1093/aje/kwg276.
35. Higbee DH, Granell R, Sanderson E, Davey Smith G, Dodd JW. Lung function and cardiovascular disease: a two-sample Mendelian randomisation study. Eur Respir J. 2021;58(3):2003196. https://doi.org/10.1183/13993003.03196-2020.
36. Bespalova IN, Burmeister M. Identification of a novel LIM domain gene, LMCD1, and chromosomal localization in human and mouse. Genomics. 2000;63(1):69–74. https://doi.org/10.1006/geno.1999.6049.
37. Kadrmas JL, Beckerle MC. The LIM domain: from the cytoskeleton to the nucleus. Nat Rev Mol Cell Biol. 2004;5(11):920–31. https://doi.org/10.1038/nrm1499.
38. Bogatkevich GS, Atanelishvili I, Bogatkevich AM, Silver RM. Critical Role of LMCD1 in Promoting Profibrotic Characteristics of Lung Myofibroblasts in Experimental and Scleroderma-Associated Lung Fibrosis. Arthritis Rheumatol. 2023;75(3):438–48. https://doi.org/10.1002/art.42344.
39. Dina C, Bouatia-Naji N, Tucker N, et al. Genetic association analyses highlight biological pathways underlying mitral valve prolapse. Nat Genet. 2015;47(10):1206–11. https://doi.org/10.1038/ng.3383.
40. Roselli C, Yu M, Nauffal V, et al. Genome-wide association study reveals novel genetic loci: a new polygenic risk score for mitral valve prolapse. Eur Heart J. 2022;43(17):1668–80. https://doi.org/10.1093/eurheartj/ehac049.
41. Sane S, Abdullah A, Boudreau DA, et al. Ubiquitin-like (UBX)-domain-containing protein, UBXN2A, promotes cell death by interfering with the p53-Mortalin interactions in colon cancer cells. Cell Death Dis. 2014;5(3): e1118. https://doi.org/10.1038/cddis.2014.100.
42. Sui Q, Chen Z, Hu Z, et al. Cisplatin resistance-related multi-omics differences and the establishment of machine learning models. J Transl Med. 2022;20(1):171. https://doi.org/10.1186/s12967-022-03372-0.
43. Zhong Y, Yang L, Xiong F, et al. Long non-coding RNA AFAP1-AS1 accelerates lung cancer cells migration and invasion by interacting with SNIP1 to upregulate c-Myc. Signal Transduct Target Ther. 2021;6(1):240. https://doi.org/10.1038/s41392-021-00562-y.
44. Hsu CH, Hsu CW, Hsueh C, et al. Identification and Characterization of Potential Biomarkers by Quantitative Tissue Proteomics of Primary Lung Adenocarcinoma. Mol Cell Proteomics. 2016;15(7):2396–410. https://doi.org/10.1074/mcp.M115.057026.
45. Han T, Wang P, Wang Y, et al. FAIM regulates autophagy through glutaminolysis in lung adenocarcinoma. Autophagy. 2022;18(6):1416–32. https://doi.org/10.1080/15548627.2021.1987672.
46. Nakachi I, Rice JL, Coldren CD, et al. Application of SNP microarrays to the genome-wide analysis of chromosomal instability in premalignant airway lesions. Cancer Prev Res (Phila). 2014;7(2):255–65. https://doi.org/10.1158/1940-6207.CAPR-12-0485.

Publisher's Note