BMJ Open Respiratory Research

Shared genetic aetiology of respiratory diseases: a genome-wide multitraits association analysis

Zhe Chen ,[1] Ning Gao,[2] Xuanye Wang,[3] Xiangming Chen,[4] YaQi Zeng,[5] Cong Li,[6] Xiahong Yang,[7] Qidong Cai,[1] Xiang Wang[1]

To cite: Chen Z, Gao N, Wang X, *et al*. Shared genetic aetiology of respiratory diseases: a genome-wide multitraits association analysis. *BMJ Open Respir Res* 2024;**11**:e002148. doi:10.1136/bmjresp-2023-002148

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjresp-2023-002148).

ZC and NG contributed equally.

Received 22 October 2023
Accepted 10 May 2024

For numbered affiliations see end of article.

Correspondence to
Professor Xiang Wang; wangxiang@csu.edu.cn and

Dr Qidong Cai; 2204130810@csu.edu.cn

ABSTRACT

Objective This study aims to explore the common genetic basis between respiratory diseases and to identify shared molecular and biological mechanisms.

Methods This genome-wide pleiotropic association study uses multiple statistical methods to systematically analyse the shared genetic basis between five respiratory diseases (asthma, chronic obstructive pulmonary disease, idiopathic pulmonary fibrosis, lung cancer and snoring) using the largest publicly available genome wide association studies summary statistics. The missions of this study are to evaluate global and local genetic correlations, to identify pleiotropic loci, to elucidate biological pathways at the multiomics level and to explore causal relationships between respiratory diseases. Data were collected from 27 November 2022 to 30 March 2023 and analysed from 14 April 2023 to 13 July 2023.

Main outcomes and measures The primary outcomes are shared genetic loci, pleiotropic genes, biological pathways and estimates of genetic correlations and causal effects.

Results Significant genetic correlations were found for 10 paired traits in 5 respiratory diseases. Cross-Phenotype Association identified 12 400 significant potential pleiotropic single-nucleotide polymorphism at 156 independent pleiotropic loci. In addition, multitrait colocalisation analysis identified 15 colocalised loci and a subset of colocalised traits. Gene-based analyses identified 432 potential pleiotropic genes and were further validated at the transcriptome and protein levels. Both pathway enrichment and single-cell enrichment analyses supported the role of the immune system in respiratory diseases. Additionally, five pairs of respiratory diseases have a causal relationship.

Conclusions and relevance This study reveals the common genetic basis and pleiotropic genes among respiratory diseases. It provides strong evidence for further therapeutic strategies and risk prediction for the phenomenon of respiratory disease comorbidity.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Respiratory diseases are the most common cause of morbidity and mortality globally. Even more worryingly, the concurrence of multiple respiratory diseases is more frequent than expected.

WHAT THIS STUDY ADDS

⇒ It remains unclear whether and how common genetic components contribute to the comorbidity of respiratory diseases. This study reveals the mechanism of comorbidity in respiratory diseases and provides a theoretical basis for the diagnosis and prevention of multiple comorbidities in clinical practice.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our results have identified a number of old and new genetic targets for respiratory diseases that may help in the development of new drug targets or new utilisation of existing drugs.

INTRODUCTION

Respiratory diseases are one of the leading causes of morbidity and mortality worldwide, with chronic obstructive pulmonary disease (COPD) and asthma being the two largest contributors to the global respiratory disease burden.[1 2] Recently, numerous studies have reported a high level of coexistence between asthma, COPD, idiopathic pulmonary fibrosis (IPF), lung cancer (LC) and obstructive sleep apnoea (OSA).[3–7] However, the mechanism of such comorbidity remains unknown. Genome wide association studies (GWASs) have identified respiratory disease-associated susceptibility loci known as single-nucleotide polymorphism (SNP). Genes located near these SNPs, such as *IL1R1*, *CFTR* and *EPHX2*, have been observed in various diseases.[8–10] Considering that previous studies have revealed significant genetic association loci for these diseases, shared genetic mechanisms may provide important insights into the comorbidity of respiratory diseases.

A gene controlling two or more traits is commonly considered a pleiotropic locus. Identifying pleiotropic loci is an important strategy for resolving genetic mechanisms. It is a common way to identify pleiotropic loci by directly taking the intersection of significant associated loci of each trait. However, such a strategy has low power due to the

limited sample size and misses a large number of potentially shared loci. Indeed, enlarging the sample size entails substantial costs. Another effective approach is to use joint modelling[11] of genetic correlations between phenotypes for joint analysis, which will reveal new genetic loci and identify potential shared loci among diseases. Previous studies have also attempted to explore the genetic overlap[12 13] or causality[14] that exists between a number of respiratory diseases, generally confined to two-by-two and with limited sample sizes. The causes and structure of the existence of commonalities among respiratory diseases remain largely unknown. Therefore, it is important to further explore the common genetic factors that underlie the commonality between respiratory diseases.

This study uses a variety of statistical genetics methods to comprehensively explore the common genetic basis between five respiratory diseases, including asthma, COPD, IPF, LC and snoring. First, we evaluated global and regional genetic correlations between diseases. We then used the Cross-Phenotype Association (CPASSOC)[15] to identify pleiotropic genetic variants or loci between diseases. In addition, Hypothesis Prioritisation in multitrait Colocalisation (HyPrColoc)[16] analysis was conducted to identify colocalised loci and traits. We also performed gene-level analyses to identify candidate pleiotropic genes through various algorithms. Finally, Mendelian randomisation (MR) analysis was performed to probe different types of pleiotropy, namely vertical and horizontal pleiotropy.

In summary, through a series of multidimensional independent and combined analyses, we identified potential pleiotropic loci between five respiratory diseases that could be prioritised as potential targets for drug development and repurposing due to their potential to simultaneously prevent or treat these diseases.

MATERIAL AND METHODS

Data processing

We used five publicly available GWAS[17–19] summary statistics of respiratory diseases (online supplemental eTable 1). We attempted to use snoring as a substitute for OSA. On one hand, snoring is a primary symptom of OSA, and on the other hand, as an independent disease phenotype, snoring has a high genetic correlation with OSA. The total sample size for asthma is 137 6071, with 121 940 cases. For COPD, the total sample size is 995 917 with 58 559 cases. IPF has a total sample size of 953 873 with 6257 cases. LC has a total sample size of 85 716 with 29 266 cases. The total sample size for snoring is 408 317 with 152 302 cases. We performed rigorous quality control. The detailed characteristics of the GWAS summary statistics and quality control process are shown in online supplemental eMethods. The overall study design is shown in figure 1.

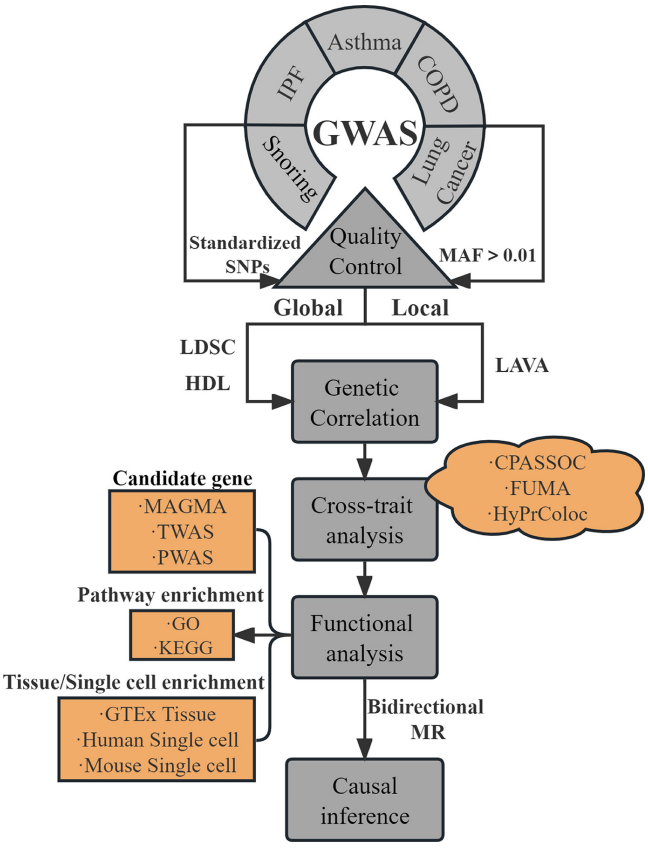

Figure 1 Overall study design. We conducted a comprehensive cross-trait analysis of five respiratory diseases from different perspectives. COPD, chronic obstructive pulmonary disease; CPASSOC, Cross Phenotype Association; FUMA, functional mapping and annotation; GWAS, genome-wide association studies; GO, Gene Ontology; HDL, high-definition likelihood; HyPrColoc, hypothesis prioritisation in multitrait colocalisation; IPF, idiopathic pulmonary fibrosis; KEGG, Kyoto Encyclopaedia of Genes and Genomes; LDSC, linkage disequilibrium score regression; LAVA, local analysis of variant association; MAGMA; multimarker analysis of GenoMic annotation; MR, Mendelian randomisation; PWAS, proteome-wide association study; SNP, single-nucleotide polymorphisms; TWAS, transcriptome-wide association study.

Global genetic and local correlation analysis

Both linkage disequilibrium score regression (LDSC)[20] and high-definition likelihood (HDL)[21] were applied to assess global genetic correlations between diseases. We did not restrict the intercept term for LDSC to assess population stratification within individual GWAS and sample overlap between pairs of GWAS.

We used the Local Analysis of Variant Association (LAVA)[22] to estimate genetic correlations in independent regions of the genome for each pair of traits. LAVA allows us to more effectively clarification of the specific effects of regional genetics on overall genetics. A false discovery rate (FDR) was used to correct for all of the above results. The significance threshold was set at p adjusted <0.05.

Multiple-trait meta-analysis

To identify potential pleiotropic SNPs of respiratory diseases, the SHet model provided by CPASSOC was executed to perform a multitrait meta-analysis.[15] The CPASSOC method enhances the sample size by incorporating information from multiple GWAS to identify novel significant SNPs. CPASSOC allows for heterogeneity and sample overlap. SNPs with $P_{CPASSOC}<5\times10^{-8}$ were considered significant pleiotropic SNP.

Genomic loci characterisation and functional annotation

We used functional mapping and annotation of genetic associations (FUMA)[23] to annotate significant genetic loci for the CPASSOC results, with the parameter settings shown in online supplemental eMethods. Combined Annotation Dependent Depletion (CADD) scores and RegulomeDB (RDB) scores were provided with FUMA, and SNP with CADD scores >12.37 were considered potentially deleterious variants. We also annotated the GWAS of five respiratory diseases for comparison with CAPSSOC results.

Multitrait colocalisation

We performed HyPrColoc[16] analyses to further identify common causal variants for each pleiotropic locus defined by FUMA. HyPrColoc divides traits into groups, with traits in each group sharing a casual SNP. Posterior prob >0.7 results in the final colocalised locus. Additionally, we conducted sensitivity analysis on the colocalisation analysis results using different prior probabilities (1×10^{-4}, 1×10^{-5}).

Candidate gene analysis

We searched for genes that overlapped with the pleiotropic loci and then subjected them to gene-based association analysis with multimarker analysis of GenoMic annotation (MAGMA).[24] FDR-corrected p<0 .05 was considered a significant result. Based on lung and whole blood tissues provided by GTEx V.8,[25] we used Functional Summary-based Imputation to perform transcriptome-wide association[26] (TWAS) analyses on GWAS for single traits. TWAS results for single traits were combined to clarify whether there is gene sharing between multiple traits at the transcriptome level. We also conducted the proteome-wide association study[27] (PWAS) study. PWAS assesses associations between plasma proteins and respiratory traits with an analytical strategy consistent with TWAS. Significant results were defined as p adjusted <0.05 after correction using the FDR method.

Pathway and GTEx tissue enrichment analysis

We conducted a MAGMA[24] gene-set enrichment analysis based on the Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway databases. We performed tissue/single cell-specific enrichment with a phenotype-cell-gene association (PCGA) analysis platform.[28–30] Significant results were defined as p adjusted <0.05 after correction using the FDR method. In PCGA, tissue results are based on GTEx V.8 and single cell results are derived from human(based on PanglaoDB, Human Cell Landscape and Allen Brain Atlas) and mouse (based on PanglaoDB) datasets. FDR-corrected p<0.05 was considered a significant result.

Mendelian randomisation analysis

We used bidirectional two-sample MR analysis to explore the causal relationships between the five phenotypes. The primary method is Multiplicative Random Effects Inverse variance weighted (IVW-MRE). IVW-MRE provides more accurate outcome estimates even in the presence of heterogeneity. Complementarily, we applied MR-Egger regression[31] and the weighted median[32] method as additional analytical strategies alongside IVW. To verify the accuracy of the results, we used sensitivity analyses such as MR-Egger intercept, leave-one-out analysis, MR-Steiger directionality test and F-statistic. The detailed methods are in online supplemental eMethods. For more information on SNPs, please refer to online supplemental eTables 2 and 3. Significant results were defined as p adjusted <0.05 after correction using FDR method.

Patient and public involvement

This study employed GWAS for data analysis and did not directly involve patients or the public.

RESULTS

Global and local genetic correlation

The LDSC results indicated significant heritability for all diseases (online supplemental eTable 4), with estimates ranging from 0.3% to 8.4%. Significant genetic associations (p adjusted <0.05) were found for all 10 pairs of traits in 5 respiratory diseases. The strongest genetic association was found between asthma and COPD (rg=0.7052), indicating a highly shared genetic component between the two. COPD and LC were also strongly correlated (rg=0.5944), suggesting a high degree of concordance between the two in terms of genetic components, but not identical. HDL described nearly identical results to LDSC except that IPF and LC suggested no genetic association in HDL (figure 2A,C and online supplemental eTables 5 and 6). Additionally, the results suggest a potential sample overlap between GWAS with selected respiratory diseases (online supplemental eTable 5). There may be some common genetic components among respiratory diseases, but there is no evidence of the extent to which these genetic components are shared and the specific mechanisms involved in these five diseases.

The results of LAVA showed a total of 499 regions (figure 2B,D and online supplemental eTable 7) where at least one pair of trait pairs existed with significant localised genetic correlations (p adjusted <0.05). Of the 499 localised regions, 87.58% were positively correlated

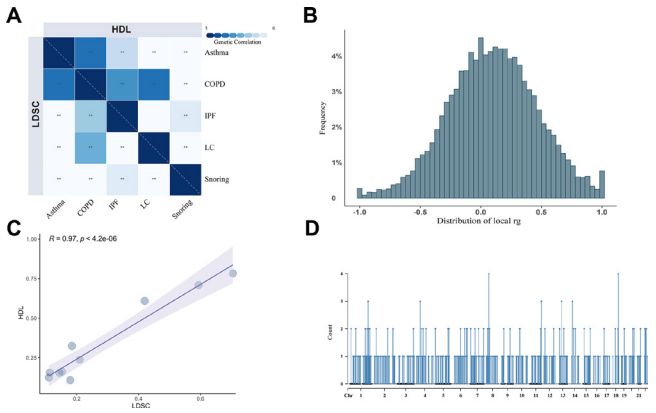

Figure 2 Genetic correlation among five respiratory diseases. (A) Global genetic correlations among five respiratory diseases were explored with LDSC and HDL methods. **p adjusted <0.05. (B) Frequency distribution of localised genetic correlations for five respiratory diseases determined by the LAVA method. (C) High consistency of the LDSC and HDL methods for investigating global genetic correlations. (D) Counts of trait pairs with local genetic associations in specific regions of the chromosome. COPD, chronic obstructive pulmonary disease; HDL, high-definition likelihood; IPF, idiopathic pulmonary fibrosis; LC, lung cancer; LDSC, linkage disequilibrium score regression.

and 12.42% were negatively correlated. The region 6:32629240–32682213 had the most six pairs of related traits and all traits occurred at least once. The asthma and COPD trait pairs had the highest significance in LDSC. LAVA indicated that this trait pair possessed the most regional genetic associations, all of which were positive.

Multiple-trait meta-analysis and pleiotropic loci

A total of 7 169 288 SNPs were included for CPASSOC analysis (figure 3 and online supplemental eTable 8). The 12 400 significant SNPs ($p<5\times10^{-8}$) in the CPASSOC results were annotated by FUMA, yielding 156

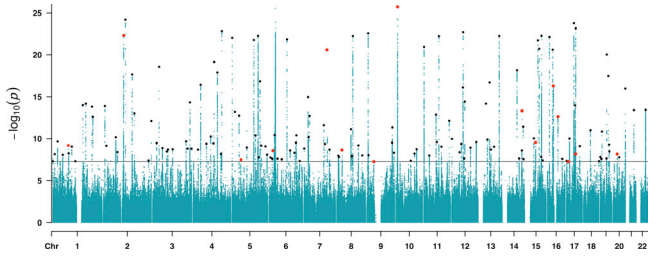

Figure 3 Manhattan plot of pleiotropic loci Manhattan plot of pleiotropic loci analysed by the CPASSOC method, with the x-axis denoting chromosomal location and the y-axis denoting the –log10 p value. The horizontal line indicates the genome-wide significance threshold of $p=5\times10^{-8}$. 156 pleiotropic loci were identified at the genome-wide significance level, of which, 15 were colocalised loci (black dots represent index SNPs of pleiotropic loci and red dots represent index SNPs of colocalised loci). CPASSOC, Cross-Phenotype Association; SNP, single-nucleotide polymorphism.

independent loci (online supplemental eTable 9). 112 loci (72%) did not overlap with any of the 5 single-trait GWAS significant loci and were considered as newly identified pleiotropic loci (online supplemental eTable 10). SNP category annotation indicated that 83 (53%) index SNPs were intronic variants and 61 (39%) index SNPs were intergenic variants. There were 12 (8%) exonic variants, including 10 messenger RNA (mRNA) exonic variants and 2 non-coding RNA exonic variants.

16 indexed SNPs had CADD scores >12.37 and 6 were mRNA exon variants (online supplemental eTable 9). For example, index SNP rs11571833 is in the exonic region of the gene BRCA2 with a CADD score=36. BRCA2 is a common tumour susceptibility gene, and studies[33] have found that BRCA2 mutations directly double the risk of developing LC. The index SNP rs34712979 with a CADD score of 22.8 is in the intronic region of the gene *NPNT*. *NPNT* was found to be significantly expressed in alveolar cells and lung fibroblasts,[34] and to regulate[35 36] the pathogenic risk of respiratory diseases such as COPD and LC.

Identification of colocalised loci

HyPrColoc obtained 15 (10%) loci (PP>0.7) (figure 3 and online supplemental eTable 11). Detailed information on all SNPs within the colocalised locis can be found in online supplemental eTable 12. The trait set with the highest posterior probability was asthma and COPD, colocalised at locus 6:19837774–19844117 (PP=0.9687, casual SNP: rs9350191). Nine of the 15 colocalised loci contained Asthma and COPD. Of all casual SNPs, five (33%) were intronic variants and eight (53%) were intergenic variants. Two (13%) were exonic variants and both were mRNA exonic variants: rs28929474 (*SERPINA1*), rs1641512 (*ATP1B2*). Notably, 7 of the 15 colocalised loci were located within the newly discovered pleiotropic loci. It is worth mentioning that in loci5, the SNP with the highest posterior probability is rs34517439. The nearby gene, *DNAJB4*, is believed to be involved in regulating the growth of non-small cell LC.[37] Recent research reports that the deletion of this gene leads to a novel type of myopathy primarily characterised by early-onset respiratory failure. This suggests that *DNAJB4* is closely associated with the development and progression of respiratory system diseases.[38] The sensitivity analysis results include heat-maps and similarity matrices (online supplemental eTable 11, marked with an asterisk**). We conducted quantitative statistics (similarity matrices) on the sensitivity analysis for different prior probabilities and found that the variation in seven locis (5, 57, 73, 86, 119, 135 and 152) did not exceed 20%. The remaining locis seem sensitive to changes in prior probabilities but still form a good cluster of colocalised traits. Even with different prior probabilities provided, six locis (17, 45, 123, 131, 132 and 139) with larger fluctuations also have a probability greater than 50% of forming stable colocalisation clusters. We also examined the LD relationships between the causal SNPs identified in the Hyprcoloc analysis and

surrounding SNPs, finding that most regions do not have variants in extremely high LD with the causal variants (online supplemental eTable 13). In five colocalised locis (90, 123, 132, 135 and 139), causal variants exhibit an LD relationship with $r^2 \geq 0.9$, accounting for 33.3% of all locis.

Candidate gene identification

We identified 739 protein-coding genes that overlapped with the 156 pleiotropic loci through gene position mapping. MAGMA analysis further identified 678 significant pleiotropic genes (online supplemental eTable 14). Among them, the most significant gene was *IL1R1* ($p=2.30\times10^{-16}$). The significant genes obtained based on MAGMA were analysed by TWAS and PWAS. A total of 3600 tests were performed for the TWAS analysis (online supplemental eTable 15). 593 tissue-gene-trait pairs were significantly associated (p adjusted <0.05): asthma (196), COPD (125), IPF (75), LC (93), snoring (104). 199 (46%) genes reached significant levels in at least one tissue, suggesting that the effects of pleiotropic genes on phenotype are influenced by the amount of mRNA expression. Of these, 108 genes were shared among different traits and were mostly tissue-specific. PWAS results (online supplemental eTable 16) indicated that 19 plasma proteins were expressed at significant levels (p adjusted <0.05). Nine of these plasma proteins share expression in two or more respiratory diseases. TWAS was compared with PWAS, and two significant protein-coding genes (*IL1R1*, *PRSS8*) were identified in the colocalised loci. *IL1R1* and *PRSS8* were associated with two or more traits at all levels.

Biological pathway, GTEx tissue and SNP-heritability enrichment

MAGMA gene-set analysis identified 130 significantly enriched biological pathways (p adjusted <0.05), including 115 GO pathways and 15 KEGG pathways (figure 4A and online supplemental eTable 17). The enriched pathways are mainly focused on the immune system. An example is interleukin-21 (GO: 0032625), which has been found in a large number of studies to play a huge effect on immune system diseases and cancer.

GTEx tissue enrichment analysis showed that five respiratory diseases were significantly enriched in 33 tissues (figure 4B). The most significant tissue was lung. A total of 167 human monocytes were significantly enriched, mainly in lung, kidney and tracheal tissue single cells. A total of 493 mouse monocytes were significantly enriched, mainly in lung, spleen, arterial and tracheal tissue single cells.

Mendelian randomisation

Bidirectional MR analysis of 20 exposure–outcome trait pairs in 5 respiratory diseases. IVW results showed a total of 10 pairs of traits with significant causal associations (p adjusted <0.05), all of which were positively correlated (figure 5 and online supplemental eTable 18). Evidence of horizontal pleiotropy existed for two trait pairs: asthma-COPD and asthma-LC. Asthma-COPD, asthma-LC obtained significant estimates consistent with IVW using MR-Egger. This suggests that they remain causally related after accounting for horizontal pleiotropy.[31] All other sensitivity analyses supported significant results (online supplemental eTable 19). Causality among the five respiratory diseases in the MR analysis was unidirectional, and the associations of exposure–outcome trait pairs were not driven by a single SNP.

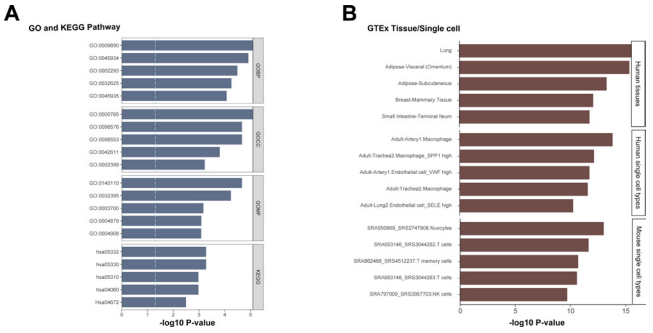

Figure 4 Biological functional and tissue and single cell-specific enrichment of candidate pleiotropic genes. (A) Top five pathways most significantly enriched for GO and KEGG gene sets. (B) Tissue-specific enrichment analysis using the PCGA (based on GTEx and PanglaoDB) identified top five significantly enriched tissues and single cell (p adjusted <0.05). BP, biological process; CC, cellular component; KEGG, Kyoto Encyclopaedia of Genes and Genomes; GO, Gene Ontology; GTEx, Genotype-Tissue Expression; KEGG, Kyoto Encyclopaedia of Genes and Genomes; MF, molecular function.

DISCUSSION

We used multiple trait association analysis to explore shared genetic factors among five respiratory diseases. This study presents a comprehensive analysis revealing the potential genetic basis of pleiotropic association loci, colocalised trait subsets, biological pathways and tissue specificity of pleiotropic genes. Different respiratory diseases often exhibit highly specific clinical features, but respiratory comorbidity is common. The results of this study support the idea that comorbidity between respiratory diseases may be driven by a common genetic basis.

We identified 61 related candidate genes in the 15 colocalised regions identified. The most significant gene, *IL1R1*, is located in the region 2: 102681836–102 801 334. Asthma and COPD trait pairs colocalise at this locus, which has a shared causal SNP of rs11679146. TWAS results from asthma showed that IL1R1 was significantly expressed in lung tissues but negative in whole blood tissues. Similarly, TWAS results in COPD had positive results for *IL1R1* in lung tissue but not in whole blood tissue. This suggests the presence of tissue specificity in the *IL1R1* gene. *IL1R1* is bound by *IL-1α* and *IL-1β* inflammatory factors. *IL-1α* and *IL-β* have been identified

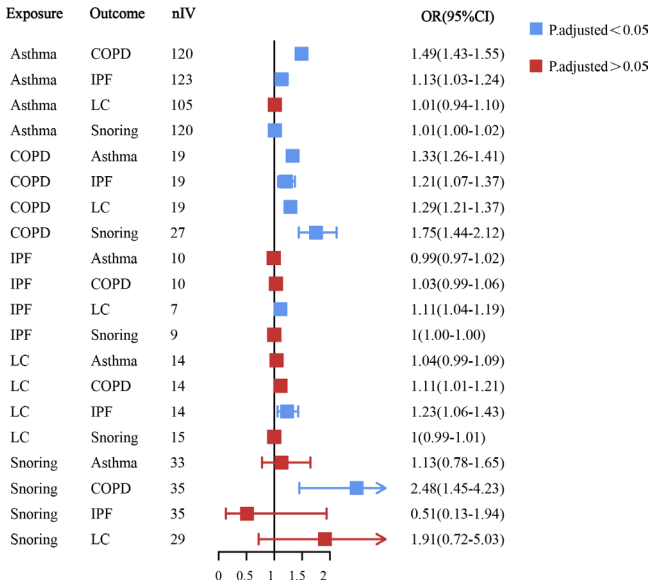

Figure 5 A bidirectional causal effect estimated with random effects IVW method. Error bars represent the 95% CI of the corresponding MR estimates. P adjusted, p value after corrected using false discovery rate; P.pleiotropy, the resultant pleiotropy remained significant after sensitivity analysis. COPD, chronic obstructive pulmonary disease; IPF, idiopathic pulmonary fibrosis; IVW, inverse variance weighted; LC, lung cancer.

in many preclinical models as mediators that play a huge role in the respiratory inflammatory response. They regulate the secretion of neutrophils, macrophages[39 40] and have an active role[41] in the development of emphysema. Clinical cohort studies have found that upregulation of *IL-1* pathway mediators is associated with frequent worsening of obstructive airway disease.[42] *IL1R1* appears to be a marker of neutrophil inflammation and airflow obstruction and is a potential therapeutic target for Asthma. *IL1R1* appears to be a marker of neutrophil inflammation and airflow obstruction and a potential therapeutic target for asthma.[8] Thus, IL1R1 gene expression in lung tissue may affect both Asthma and COPD. It may be feasible to develop targeted drugs against *IL1R1*, which needs to be justified by more research data. A feasible therapeutic target for asthma, *IL1R1* appears to be a marker of neutrophil inflammation and airflow obstruction.

The *TNFSF12* gene can be found in the region of 17:7447375–7466207. Asthma and snoring share this locus and the most likely causal SNP is rs1641512. *TNFSF12-encoded TWEAK* and its receptor factor-inducible 14 (*Fn14*) have been shown to induce the production of *IL-8* and *GM-CSF* by the human bronchial epithelium[43] and to contribute to airway inflammation by activating the *NF-B/STAT3* pathway to produce a variety of inflammatory mediators.[44] A more intriguing finding[45] was the direct detection of significantly higher *TWEAK* in the sputum of asthmatic patients, and the quantity was positively connected with the degree of the disease. Obesity is the main OSA risk factor.[46] Certain clinical studies found that obese patients' adipose tissue contained increased concentrations of *TWEAK/Fn14*.[47] Snoring, the most noticeable OSA symptom and a simple source of intermittent hypoxia. It has been reported that the mechanism of Hypoxia-inducible factor (*HIF)-1α* in OSA-related complications[48 49] while *TWEAK* has been discovered to increase *HIF-1* expression.[50] Thus, *TWEAK/Fn14* may be a good path to explore for intervention treatment of patients with asthma and OSA.

In addition to the genes already mentioned, additional genes, including *CFTR*, *EPHX2*, *FTO* and others, were found to be often associated with respiratory illnesses. Additionally, it was found that several genes, including *BCKDK*, *SETD1A*, *VKORC1*, *PRSS8* and others, temporarily lack associations with respiratory illnesses. Significantly, the *PRSS8* gene, positioned at 16:31137754–31152151, shows colocalisation in both asthma and snoring. Both TWAS lung tissue and whole blood and PWAS contained this common gene. According to reports, the *PRSS8* pathway plays a crucial role in controlling sodium transport in alveolar epithelial cells and pulmonary fluid balance,[51] but further research is needed to determine how it relates to respiratory illnesses.

The majority of the shared pathways among the five respiratory illnesses emphasised immunology, which nearly matched single-cell enrichment findings like the high expression of macrophages in the trachea and T cells in lung tissue. The immune system's undeniable implication in COPD has dominated basic research for years.[52] Much earlier, asthma has been connected to the immunological response[53] and it has also been demonstrated that immune dysregulation triggers IPF.[54] Moreover, an important factor in controlling OSA's inflammatory response is the activation of the *NLRP3* inflammasome. LC has also received a lot of attention lately in the immunological microenvironment[55] and immunotherapy[56] of LC. This study provides additional favourable support that the immune system may be an important pathway for respiratory diseases to be codriven.

Bidirectional MR analysis further explores possible genetically related causal relationships between multiple diseases.[57] Asthma affects COPD, IPF and snoring while COPD and IPF threaten LC. The MR results demonstrated that the relationship between the five respiratory diseases was interactive and intricate. This provides evidence of the effectiveness of clinical prevention of respiratory disease complications.

This study reveals the mechanism of co-morbidity in respiratory diseases and provides a theoretical basis for the diagnosis and prevention of multiple comorbidities in clinical practice. Our results have identified a number of old and new genetic targets for respiratory diseases that may help in the development of new drug targets or new utilisation of existing drugs. The limitation that has to be recognised is that the GWAS included in this study were all from a single European population so the results may not necessarily match other ancestries. More subsequent

GWASs of other ancestries are needed to validate this result or to mine new applicable loci. Proof of this result or exploration of new applicable loci will require subsequent GWAS of other ancestries to validate. Furthermore, the presence of excessively high LD relationships may obscure the accurate identification of causal SNPs in colocalisation analysis (specifically at loci 90, 123, 132, 135 and 139), necessitating further validation.

CONCLUSION

In conclusion, this study reveals the presence of comorbid genetic correlations between five respiratory diseases, identifies pleiotropic loci, defines common biological mechanisms dominated by immune responses and infers potential causal relationships between the diseases. We suggest that respiratory diseases are not entirely independent but are closely linked and share specific genetic loci. These findings provide strong evidence for further therapeutic strategies and risk prediction.

Author affiliations
[1]Department of Thoracic Surgery, The Second Xiangya Hospital of Central South University Department of Thoracic Surgery, Changsha, Hunan, China
[2]Department of Cardiovascular Surgery, The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China
[3]Department of Oncology, Xi'an Jiaotong University Second Affiliated Hospital Department of Oncology, Xi'an, Shaanxi, China
[4]Department of Orthopaedics, Xiangya Hospital Central South University, Changsha, Hunan, China
[5]Department of Psychiatry, Brain Hospital of Hunan Province, Changsha, Hunan, China
[6]Department of Radiology, The Second Xiangya Hospital of Central South University Department of Radiology, Changsha, Hunan, China
[7]Department of Anesthesiology, The Second Xiangya Hospital of Central South University Department of Anesthesiology, Changsha, Hunan, China

X Xiangming Chen @No twitter

Contributors The study's conception, data analysis and interpretation, and manuscript writing were all done by ZC and NG. XuW, XC, YZ, CL, and XY analysed and interpreted the data. QC provides methodological support and revises the manuscript. XiW concepted and designed the study and revised the manuscript. XiW is responsible for the overall content as the guarantor. The final manuscript was read and approved by the authors. GWAS Catalog (https://www.ebi.ac.uk/gwas/), Global Biobank Meta-analysis Initiative (http://results.globalbiobankmeta.org/) provided access to genome-wide association study data. We thank all research participants who provided DNA samples for these studies

Funding This work was supported by the Central South University Research Programme of Advanced Interdisciplinary Studies (2023QYJC030, XiW).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository.

ORCID iD
Zhe Chen http://orcid.org/0009-0008-6526-1013

REFERENCES

1 Christenson SA, Smith BM, Bafadhel M, *et al*. Chronic obstructive pulmonary disease. *The Lancet* 2022;399:2227–42.
2 The global asthma report 2022. *Int J Tuberc Lung Dis* 2022;26:1–104.
3 Gibson PG, McDonald VM. Asthma-COPD overlap 2015: now we are six. *Thorax* 2015;70:683–91.
4 Teodorescu M, Barnet JH, Hagen EW, *et al*. Association between asthma and risk of developing obstructive sleep apnea. *JAMA* 2015;313:156–64.
5 Soler X, Gaio E, Powell FL, *et al*. High prevalence of obstructive sleep apnea in patients with moderate to severe COPD. *Annals ATS* 2015;150414075541005.
6 Houghton AM. Mechanistic links between COPD and lung cancer. *Nat Rev Cancer* 2013;13:233–45.
7 Brown S-AW, Dobelle M, Padilla M, *et al*. Idiopathic pulmonary fibrosis and lung cancer. A systematic review and meta-analysis. *Annals ATS* 2019;16:1041–51.
8 Evans MD, Esnault S, Denlinger LC, *et al*. Sputum cell IL-1 receptor expression level is a marker of airway neutrophilia and airflow obstruction in asthmatic patients. *J Allergy Clin Immunol* 2018;142:415–23.
9 Patel SD, Bono TR, Rowe SM, *et al*. CFTR targeted therapies: recent advances in cystic fibrosis and possibilities in other diseases of the airways. *Eur Respir Rev* 2020;29:190068.
10 Yang J, Bratt J, Franzi L, *et al*. Soluble epoxide hydrolase inhibitor attenuates inflammation and airway hyperresponsiveness in mice. *Am J Respir Cell Mol Biol* 2015;52:46–55.
11 Lincoln MR, Connally N, Axisa P-P, *et al*. Joint analysis reveals shared autoimmune disease associations and identifies common mechanisms. *Genetic and Genomic Medicine* [Preprint].
12 John C, Guyatt AL, Shrine N, *et al*. Genetic associations and architecture of asthma-COPD overlap. *Chest* 2022;161:1155–66.
13 Parris BA, O'Farrell HE, Fong KM, *et al*. Chronic obstructive pulmonary disease (COPD) and lung cancer: common pathways for pathogenesis. *J Thorac Dis* 2019;11:S2155–72.
14 Zhu J, Zhou D, Wang J, *et al*. A causal atlas on comorbidities in idiopathic pulmonary fibrosis. *Chest* 2023;164:429–40.
15 Zhu X, Feng T, Tayo BO, *et al*. Meta-analysis of correlated traits via summary Statistics from GWASs with an application in hypertension. *Am J Hum Genet* 2015;96:21–36.
16 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;12:764.
17 Zhou W, Kanai M, Wu K-HH, *et al*. Global Biobank meta-analysis initiative: powering genetic discovery across human disease. *Cell Genom* 2022;2:100192.
18 Campos AI, García-Marín LM, Byrne EM, *et al*. Insights into the aetiology of snoring from observational and genetic investigations in the UK Biobank. *Nat Commun* 2020;11:817.
19 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.
20 Bulik-Sullivan BK, Loh P-R, Finucane HK, *et al*. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015;47:291–5.
21 Ning Z, Pawitan Y, Shen X. High-definition likelihood inference of genetic correlations across human complex traits. *Nat Genet* 2020;52:859–64.
22 Werme J, van der Sluis S, Posthuma D, *et al*. An integrated framework for local genetic correlation analysis. *Nat Genet* 2022;54:274–82.
23 Watanabe K, Taskesen E, van Bochoven A, *et al*. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017;8:1826.
24 de Leeuw CA, Mooij JM, Heskes T, *et al*. MAGMA: generalized gene-set analysis of GWAS data. *PLOS Comput Biol* 2015;11:e1004219.

25 Lonsdale J, Thomas J, Salvatore M, *et al*. The genotype-tissue expression (GTEx) project. *Nat Genet* 2013;45:580–5.
26 Gusev A, Ko A, Shi H, *et al*. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet* 2016;48:245–52.
27 Zhang J, Dutta D, Köttgen A, *et al*. Plasma Proteome analyses in individuals of European and African ancestry identify Cis-PQTLs and models for Proteome-wide Association studies. *Nat Genet* 2022;54:593–602.
28 Xue C, Jiang L, Zhou M, *et al*. PCGA: a comprehensive web server for phenotype-cell-gene association analysis. *Nucleic Acids Res* 2022;50:W568–76.
29 Jiang L, Xue C, Dai S, *et al*. DESE: estimating driver tissues by selective expression of genes associated with complex diseases or traits. *Genome Biol* 2019;20:233.
30 Li M, Jiang L, Mak TSH, *et al*. A powerful conditional gene-based association approach implicated functionally important genes for schizophrenia. *Bioinformatics* 2019;35:628–35.
31 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.
32 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.
33 Wang Y, McKay JD, Rafnar T, *et al*. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. *Nat Genet* 2014;46:736–41.
34 Yoshiji S, Butler-Laporte G, Lu T, *et al*. Proteome-wide Mendelian randomization Implicates nephronectin as an actionable mediator of the effect of obesity on COVID-19 severity. *Nat Metab* 2023;5:248–64.
35 Saferali A, Xu Z, Sheynkman GM, *et al*. Characterization of a COPD-associated NPNT functional splicing genetic variant in human lung tissue via long-read sequencing. *Genetic and Genomic Medicine* [Preprint] 2020.
36 Wang P, Zhang Y, Lv X, *et al*. Lncrna Adamts9-As1 inhibits the Stemness of lung adenocarcinoma cells by regulating Mir-5009-3p/NPNT axis. *Genomics* 2023;115:110596.
37 She K, Yu S, He S, *et al*. CircRNA 0009043 suppresses non-small-cell lung cancer development via targeting the miR-148A-3P/DNAJB4 axis. *Biomark Res* 2022;10:61.
38 Weihl CC, Töpf A, Bengoechea R, *et al*. Loss of function variants in DNAJB4 cause a myopathy with early respiratory failure. *Acta Neuropathol* 2023;145:127–43.
39 Pauwels NS, Bracke KR, Dupont LL, *et al*. Role of IL-1 and the Nlrp3/Caspase-1/IL-1 axis in cigarette smoke-induced pulmonary inflammation and COPD. *Eur Respir J* 2011;38:1019–28.
40 Bucher H, Mang S, Keck M, *et al*. Neutralization of both IL-1A/IL-1B plays a major role in suppressing combined cigarette smoke/virus-induced pulmonary inflammation in mice. *Pulm Pharmacol Ther* 2017;44:96–105.
41 Couillin I, Vasseur V, Charron S, *et al*. IL-1R1/Myd88 signaling is critical for elastase-induced lung inflammation and emphysema. *J Immunol* 2009;183:8195–202.
42 Baines KJ, Fu J-J, McDonald VM, *et al*. Airway gene expression of IL-1 pathway mediators predicts exacerbation risk in obstructive airway disease. *Int J Chron Obstruct Pulmon Dis* 2017;12:541–50.
43 Xu H, Okamoto A, Ichikawa J, *et al*. TWEAK/Fn14 interaction stimulates human bronchial epithelial cells to produce IL-8 and GM-CSF. *Biochem Biophys Res Commun* 2004;318:422–7.
44 Wang A, Zhang F, Xu H, *et al*. TWEAK/Fn14 promotes pro-inflammatory cytokine secretion in hepatic stellate cells via NF-κB/Stat3 pathways. *Mol Immunol* 2017;87:67–75.
45 Kim SY, Kim JD, Sol IS, *et al*. Sputum TWEAK expression correlates with severity and degree of control in non-eosinophilic childhood asthma. *Pediatr Allergy Immunol* 2018;29:42–9.
46 Li M, Li X, Lu Y. Obstructive sleep apnea syndrome and metabolic diseases. *Endocrinology* 2018;159:2670–5.
47 Maymó-Masip E, Fernández-Veledo S, Garcia España A, *et al*. The rise of soluble TWEAK levels in severely obese subjects after Bariatric surgery may affect Adipocyte-cytokine production induced by TNFα. *J Clin Endocrinol Metab* 2013;98:E1323–33.
48 Kasivisvanathan V, Shalhoub J, Lim CS, *et al*. Hypoxia-inducible Factor-1 in arterial disease: a putative therapeutic target. *Curr Vasc Pharmacol* 2011;9:333–49.
49 Belaidi E, Joyeux-Faure M, Ribuot C, *et al*. Major role for hypoxia inducible Factor-1 and the Endothelin system in promoting myocardial infarction and hypertension in an animal model of obstructive sleep apnea. *J Am Coll Cardiol* 2009;53:1309–17.
50 Wang M, Xie Z, Xu J, *et al*. TWEAK/Fn14 axis in respiratory diseases. *Clin Chim Acta* 2020;509:139–48.
51 Planès C, Randrianarison NH, Charles R-P, *et al*. Enac-mediated alveolar fluid clearance and lung fluid balance depend on the channel-activating protease 1. *EMBO Mol Med* 2010;2:26–37.
52 Kheradmand F, Zhang Y, Corry DB. Contribution of adaptive immunity to human COPD and experimental models of emphysema. *Physiol Rev* 2023;103:1059–93.
53 Holgate ST. Innate and adaptive immune responses in asthma. *Nat Med* 2012;18:673–83.
54 Shenderov K, Collins SL, Powell JD, *et al*. Immune dysregulation as a driver of idiopathic pulmonary fibrosis. *J Clin Invest* 2021;131:e143226.
55 Sorin M, Rezanejad M, Karimi E, *et al*. Single-cell spatial landscapes of the lung tumour immune microenvironment. *Nature* 2023;614:548–54.
56 Lahiri A, Maji A, Potdar PD, *et al*. Lung cancer immunotherapy: progress, pitfalls, and promises. *Mol Cancer* 2023;22:40.
57 van Rheenen W, Peyrot WJ, Schork AJ, *et al*. Genetic correlations of polygenic disease traits: from theory to practice. *Nat Rev Genet* 2019;20:567–81.