# Using Blood Metabolites to Identify Genetic Factors in Respiratory Disease*

Dong Yiran, Li Huilong, Min Yiping, Li Yihan, Wang Jie, Zhao Jianping†
(School of Mathematics and System Science, Xinjiang University, Urumqi Xinjiang 830017, China)

**Abstract**: Genome-wide association study (GWAS) data are used to explore the associations between blood metabolites and 5 respiratory diseases: asthma, tuberculosis (TB), chronic obstructive pulmonary disease (COPD), cor pulmonale, and bronchitis. The main method of analysis used is the inverse-variance weighted (IVW) approach, complemented by several sensitivity analyses, including MR-Egger regression, the weighted median, the weighted mode, Cochran's Q test, and the pleiotropy test. Additional directional tests, Meta-analysis and metabolic pathway analyses are conducted for deeper insights. 3 metabolites showing significant causal relationships are identified. Catechol glucuronide levels as a protective factor have a positive causal relationship with asthma; the creatine to carnitine ratio has a negative causal relationship with COPD as a risk factor; and the adenosine 5'-diphosphate (ADP) to N-acetylglucosamine to N-acetylgalactosamine ratio as a protective factor has a positive causal relationship with bronchitis. Additionally, 13 metabolites demonstrate strong causal relationships. Furthermore, we delineate 14 metabolic pathways related to the outcomes, including 6 associated with asthma, 2 with TB, 1 with COPD, 4 with cor pulmonale, and 1 with bronchitis. A causal relationship between blood metabolites and 5 respiratory diseases has been established. The identified metabolites and pathways offer new insights into the underlying mechanisms of these diseases, necessitating further experimental validation.

**Key words**: Mendelian randomization; genome-wide association study; blood metabolites; respiratory disease

**DOI**: 10.13568/j.cnki.651094.651316.2025.09.16.0001

**CLC Number**: Q3-02 **Document Code**: A **Article ID**: 2096-7675(2026)02-0196-012

**引文格式**：董逸然,李慧龙,闵一萍,李逸晗,王杰,赵建平. 利用血液代谢物识别呼吸系统疾病的遗传因素[J]. 新疆大学学报(自然科学版中英文),2026,43(2):196-207.

**英文引文格式**：Dong Yiran, Li Huilong, Min Yiping, Li Yihan, Wang Jie, Zhao Jianping. Using blood metabolites to identify genetic factors in respiratory disease[J]. Journal of Xinjiang University(Natural Science Edition in Chinese and English),2026,43(2):196-207.

# 利用血液代谢物识别呼吸系统疾病的遗传因素

董逸然，李慧龙，闵一萍，李逸晗，王 杰，赵建平
（新疆大学 数学与系统科学学院，新疆 乌鲁木齐 830017）

**摘 要**：利用全基因组关联研究(GWAS)数据，本文系统探讨了血液代谢物与5种呼吸系统疾病(哮喘、结核病、慢性阻塞性肺疾病、肺源性心脏病及支气管炎)之间的关联性.采用逆方差加权法(IVW)作为核心分析方法，并辅以多项敏感

* **Received Date:** 2025-09-16; **Revised Date:** 2025-12-31; **Accepted Date:** 2026-01-03.

**Foundation Item:** The National Natural Science Foundation of China-Regional Science Foundation Project "Research on constructing cell communication networks based on single-cell and spatial transcriptomics data" (62362062); The Multimodal Major Chronic Disease Prevention and Control Science and Engineering Key Laboratory Project of MIIT "Identification of novel drug targets for lung cancer via Mendelian randomization analysis based on blood proteomics" (MCD-2023-1-15).

**Biography:** Dong Yiran (1999—), female, master student, research fields: applied statistics, E-mail: 15374717436@163.com.

† **Corresponding Author:** Zhao Jianping (1981—), male, professor, research fields: methods and applications in continuum fluid mechanics, numerical analysis, operations research and optimization, big data analysis, statistical computing, bioinformatics statistics, E-mail: zhaojianping@126.com.

性分析,包括MR-Egger回归、加权中位数法、加权众数法、Cochran's Q检验及多效性检验.通过方向性检验、Meta分析和代谢途径分析进一步深化研究结论.鉴定出3种具有显著因果关联的代谢物:作为保护性因素的儿茶酚葡萄糖醛酸苷水平与哮喘呈正向因果关系;作为风险因素的肌酸与肉碱比率与慢性阻塞性肺疾病呈负向因果关系;作为保护性因素的腺苷-5'-二磷酸(ADP)-N-乙酰葡糖胺-N-乙酰半乳糖胺比率与支气管炎呈正向因果关系.此外,还发现了存在强因果关联的13种代谢物.进一步解析出14条与目标疾病相关的代谢通路,其中6条通路与哮喘相关、2条与肺结核相关、1条与慢性阻塞性肺疾病相关、4条与肺源性心脏病相关、1条与支气管炎相关.本文证实血液代谢物与5种呼吸系统疾病之间存在因果关联,发现的代谢物及通路为揭示这些疾病的潜在机制提供了新视角,需通过后续实验进一步验证.

**关键词**:孟德尔随机化;全基因组关联研究;血液代谢物;呼吸系统疾病

# 0 Introduction

Respiratory diseases, such as asthma, tuberculosis (TB), chronic obstructive pulmonary disease (COPD), cor pulmonale, and bronchitis, are among the leading causes of global morbidity and mortality, affecting people of all ages and imposing a substantial health burden[1–2]. These conditions arise from complex interactions between genetic susceptibility and environmental exposures. Additionally, respiratory diseases increase the risk of other conditions, such as neurological problems[3] and muscle dysfunction[4]. Although observational studies have identified numerous risk factors[5], their results are often influenced by confounding biases and reverse causality, making it difficult to establish clear causal relationships.

Recent metabolomics advancements have significantly impacted disease diagnosis and treatment, attracting considerable attention. This high-throughput analytical technique provides a comprehensive profile of metabolites in the human body, thereby enhancing the possibilities for early detection and preventive strategies[6]. Metabolomics quantifies and tracks change in metabolite concentrations, thereby elucidating complex physiological and pathological processes. Unlike genomics and proteomics, metabolomics more accurately reflects the current physiological state, introducing novel biomarkers for early diagnosis, disease classification, and monitoring of disease progression. The evolving capabilities of metabolomics technologies have broadened their applications in respiratory disease research[7], aiding in the understanding of disease pathogenesis and identifying specific metabolic biomarkers[8–9]. Although observational studies in metabolomics have pinpointed certain biomarkers for conditions, like asthma[10], TB[11], and COPD[12], they typically involve small sample sizes and offer limited insight into underlying pathologies. Comprehensive metabolomics research promises to deepen our understanding of the metabolic dynamics in respiratory diseases.

To address these limitations, Mendelian randomization (MR) has been developed as an epidemiological method that uses genetic variants as instrumental variables (IVs) to infer causality[13–15]. The approach is based on Mendel's laws of inheritance, whereby genetic variants are randomly assigned at conception, largely independent of confounding environmental or lifestyle factors, and less susceptible to reverse causation. With the expansion of genome-wide association studies (GWAS), large-scale genetic data have become widely accessible, facilitating the application of two-sample MR (TSMR) frameworks. This method uses summary-level data from existing GWAS to estimate causal effects without requiring individual-level data, significantly reducing research costs.

Foundational studies have been critical in enabling MR analyses of metabolites. For example, large-scale GWAS of blood metabolites[16] have provided comprehensive maps of genetic variants influencing metabolite levels, creating essential resources for subsequent causal inference studies. In the field of respiratory diseases, MR has already demonstrated its value. For instance, studies of COPD have systematically evaluated causal relationships between hundreds of blood metabolites and disease risk, identifying specific metabolites involved in disease pathogenesis[17–18]. Similar approaches have been applied to other respiratory diseases[19–20]. Furthermore, multi-omics integration approaches have begun to combine genomic and metabolomics data to build more complete models of disease etiology[21–23].

In this context, the present study aims to systematically investigate the causal relationships between blood metabolites and 5 major respiratory diseases: asthma, TB, COPD, cor pulmonale, and bronchitis. Using a TSMR

framework and large-scale GWAS summary data, we use genetic variants as instrumental variables to infer the causal effects of blood metabolites on these diseases. This study not only seeks to identify specific metabolite biomarkers with causal roles, but also intends to conduct sensitivity analyses and metabolic pathway analyses to explore underlying biological mechanisms. By integrating genetic and metabolomics evidence, this research aims to advance our understanding of the etiology of respiratory diseases, and provide a reliable causal foundation for future early screening, risk prediction, and targeted intervention strategies.

# 1 Data Sources and Analysis Methods

## 1.1 Data Sources

We selected 1 400 blood metabolites as exposures and asthma, TB, COPD, cor pulmonale, and bronchitis as outcomes for MR analysis. Data on these metabolites were sourced from the July 2023 Nature Genetics article, "Genomic atlas of the plasma metabolome prioritises metabolites implicated in human diseases"[24]. This study involved 8 299 older Canadians (average age 62 ± 10 years), and a GWAS was conducted on 1 091 blood metabolites and 309 metabolite ratios. Outcome data were obtained from the Finnish database[25]. In addition, we selected another set of data from different years and sources for replication analysis and Meta-analysis.

TSMR analysis requires two independent samples from different populations with similar gender, age, and other characteristics. Therefore, different databases were used for the analysis and the populations in this study are all European. To ensure that there are no ethical issues in this study, the data used in this paper are all publicly available.

## 1.2 Study Design

To ensure accurate results, MR relies on 3 essential assumptions. Firstly, the genetic variations used as IVs need to be significantly associated with the risk factors under investigation (relevance assumption). Secondly, these IVs must be independent of any confounders, whether known or unknown (independence assumption). Lastly, the IVs should influence the outcome exclusively through the risk factors, without any other direct causal pathways (exclusivity assumption). Simply put, the chosen IVs should impact the occurrence of respiratory diseases solely through blood metabolites and not via confounders or direct effects on respiratory diseases. Genetic data for blood metabolites and 5 respiratory diseases were sourced from independent GWAS datasets to prevent sample overlap. Figure 1 shows a summary of the study.

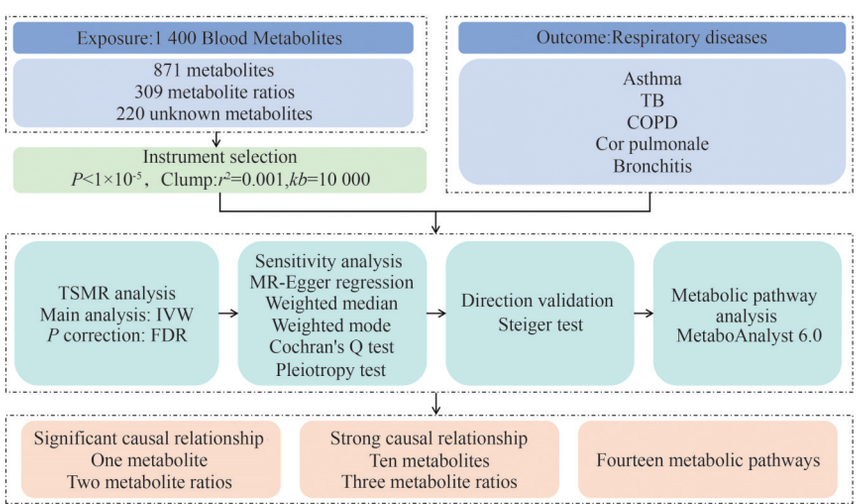


**Figure 1 Research design plan**

## 1.3 Selection of IVs

This study used the R package (TwoSampleMR) for clump processing, specifying single nucleotide polymor-

phisms (SNPs) with $P<1\times10^{-5}$, linkage disequilibrium (LD) $r^2=0.001$, and an LD screening region width of 10 000 kb. Missing SNPs were substituted with highly correlated SNPs, and palindromic sites were excluded. In total, 1 352 metabolites were analyzed.

Weak IVs may introduce bias, making their selection critical. We screened for weak IVs by calculating the *F*-statistic ($F=(R^2/(1-R^2))\times(N-K-1/K)$, where $R^2=\beta^2\times2\times MAF\times(1-MAF)$)[26]. In these equations, $R^2$ is the proportion of variance explained by SNPs associated with exposure, $N$ represents the sample size, $K$ represents the number of SNPs included, $\beta$ denotes the effect size of SNPs associated with blood metabolites, and $MAF$ signifies the minor allele frequency. Generally, an *F*-statistic less than 10 indicates a weak association, potentially affecting MR results. Consequently, we excluded IVs with an *F*-statistic below 10.

All data used for MR analysis in this study underwent rigorous screening. The aforementioned data processing helps to reduce the influence of potential confounding factors on the study results. The data quality in this study is satisfactory, providing a robust foundation for interpreting the findings.

### 1.4 TSMR

A primary method commonly used in MR studies is inverse variance weighted (IVW)[27]. It assigns weights to effect estimates based on each IV effect estimate and its variance, giving greater weight to IVs with smaller variance. IVW is the most efficient method for estimating MR effects, and thus this study uses IVW as the primary method to explore the causal associations between blood metabolites ($P_{IVW}<0.05$) and respiratory diseases. The FDR method[28] was used to correct *P*-values for multiple comparisons.

### 1.5 Sensitivity Analysis

To assess the robustness of our results, we used MR-Egger regression[29], weighted median[30] and weighted mode[31] as additional methods. The weighted median can provide an accurate estimate of causal effects even when less than 50% of the genetic variants violate the core assumptions of the MR. The MR-Egger method fits a regression model of gene-outcome and gene-exposure associations to test and correct for biases caused by IV pleiotropy. Cochran's Q test[32] and the pleiotropy test were used to detect the presence of heterogeneity and pleiotropy, respectively. Statistically significant results ($P<0.05$) from Cochran's Q test and the pleiotropy test indicate substantial heterogeneity and horizontal pleiotropy in the analysis. The criteria for sensitivity analysis are: (1) the direction and magnitude of 4 MR methods are similar; (2) no heterogeneity detected ($P_{\text{Cochran's Q test}}>0.05$); and (3) no pleiotropy detected ($P_{\text{the pleiotropy test}}>0.05$).

### 1.6 Identification of Candidate Biomarkers

The following criteria were used to screen candidate biomarkers: (1) if $P_{IVW\text{-}FDR}<0.05$, the candidate biomarker is considered to be significant in causally relating the outcome; (2) if $P_{IVW\text{-}FDR}>0.05$ but $P_{IVW}<0.05$, and the *P*-values for MR-Egger regression, weighted median, and weighted mode are all less than 0.05, the candidate biomarker is considered to have a strong causal relationship with the outcome; (3) the candidate biomarker passes the sensitivity analysis, indicating no heterogeneity or pleiotropy.

### 1.7 Direction Validation

The Steiger test[33] is a statistical method used to examine reverse causality. It operates under the assumption of causal direction and determines the causal relationship between variables by comparing the goodness of fit of 2 models. When the Steiger_Pval is less than 0.05 and the correct_causal_direction result is TRUE, it indicates the absence of reverse causality between exposure and outcome.

### 1.8 Meta-Analysis

To strengthen the evidence supporting the candidate biomarkers, we performed replication analyses. A Meta-analysis of the IVW results from both the primary and replication analyses was conducted to derive a comprehensive causal relationship. The Q test and $I^2$ statistic were applied to evaluate heterogeneity within the Meta-analysis. Het-

erogeneity was considered significant when $I^2$>50% and $P$<0.05. In the absence of heterogeneity, a fixed-effect model was used, while a random-effect model was used when significant heterogeneity was present. The statistical threshold for the Meta-analysis was set at 0.05.

### 1.9 Metabolic Pathway Analysis

Based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) IDs provided by the author, pathway analysis of known metabolites ($P_{IVW}$<0.05) was performed using MetaboAnalyst 6.0 (https://www.metaboanalyst.ca/). The Small Molecule Pathway Database (SMPDB) and KEGG databases were used to extract and analyse potentially positive metabolites. The objective was to identify metabolic pathways that may be associated with the biological processes of 5 respiratory diseases ($P$<0.05).

### 1.10 Reported Results and Software

Given that the outcomes in question are binary variables, the results are presented as odds ratios (*OR*) and 95% confidence intervals (*CI*) per standard deviation. These figures are expressed as estimates (Estimate) for the sake of simplicity. A two-tailed *P*-value of 0.05 was deemed to represent a statistically significant threshold. All statistical analyses were conducted using the TwoSampleMR package (version 0.5.7) in R (version 4.5.1), code is available at https://github.com/DAIOKD/Using-blood-metabolites-to-identify-genetic-factors-in-respiratory-disease.

## 2 Statistical Analysis

### 2.1 Results of TSMR

After conducting TSMR analysis, we identified 91 metabolites with potential causal relationships to asthma, including 74 known and 17 unknown. For TB, 68 metabolites were identified, comprising 58 known and 10 unknown. The analysis for COPD revealed 77 metabolites, with 68 known and 9 unknown. For cor pulmonale, 88 metabolites were identified, including 74 known and 14 unknown. Lastly, 75 metabolites with potential causal relationships to bronchitis were found, consisting of 61 known and 14 unknown. Notably, catechol glucuronide levels ($P_{FDR}=1.01\times10^{-2}$), adenosine 5'-diphosphate (ADP) to N-acetylglucosamine to N-acetylgalactosamine ratio ($P_{FDR}=2.00\times10^{-2}$), and creatine to carnitine ratio ($P_{FDR}=3.72\times10^{-3}$) all passed Bonferroni correction. This study does not examine unknown metabolites. The volcano map (Figure 2) highlights the risk association between blood metabolites and outcomes.

### 2.2 Sensitivity Analysis

Among the metabolites that displayed comparable direction and magnitude across 4 methods, IVW, MR-Egger regression, weighted median and weighted mode, 70 metabolites showed potential causal relationships with asthma. Similarly, 52 metabolites were potentially causally related to TB, 59 metabolites to COPD, 61 metabolites to cor pulmonale, and 50 metabolites to bronchitis.

Additionally, among the 70 metabolites with potential causal relationships with asthma, 12 did not pass the heterogeneity test, 1 did not pass the pleiotropy test, and 3 did not pass both tests. Among the 52 metabolites with potential causal relationships with TB, 3 did not pass the pleiotropy test. For COPD, among the 59 metabolites, 4 did not pass the heterogeneity test and 1 did not pass the pleiotropy test. For cor pulmonale, among the 61 metabolites, 3 did not pass the heterogeneity test and 2 did not pass the pleiotropy test. For bronchitis, all 50 metabolites passed both the heterogeneity and pleiotropy tests.

### 2.3 Candidate Biomarker

An *OR*>1 indicates the exposure increases disease risk; *OR*<1 indicates it decreases risk; *OR*=1 indicates no association. According to 3 criteria for identifying candidate biomarkers, we discovered 3 metabolites with significant causal relationships. Catechol glucuronide levels ($P$: $7.48\times10^{-6}$, *OR*: 0.91, 95%*CI*: 0.87-0.95) as a protective factor has a positive causal relationship with asthma; creatine to carnitine ratio ($P$: $2.75\times10^{-6}$, *OR*: 0.76, 95%*CI*: 0.67-0.85) as a protective factor has a positive causal relationship with COPD; and adenosine 5'-diphosphate (ADP) to N-acetylglucosamine to N-acetylgalactosamine ratio ($P$: $1.48\times10^{-5}$, *OR*: 1.13, 95%*CI*: 1.07-1.19) as a

risk factor has a negative causal relationship with bronchitis. Additionally, we identified 13 metabolites with strong causal relationships.

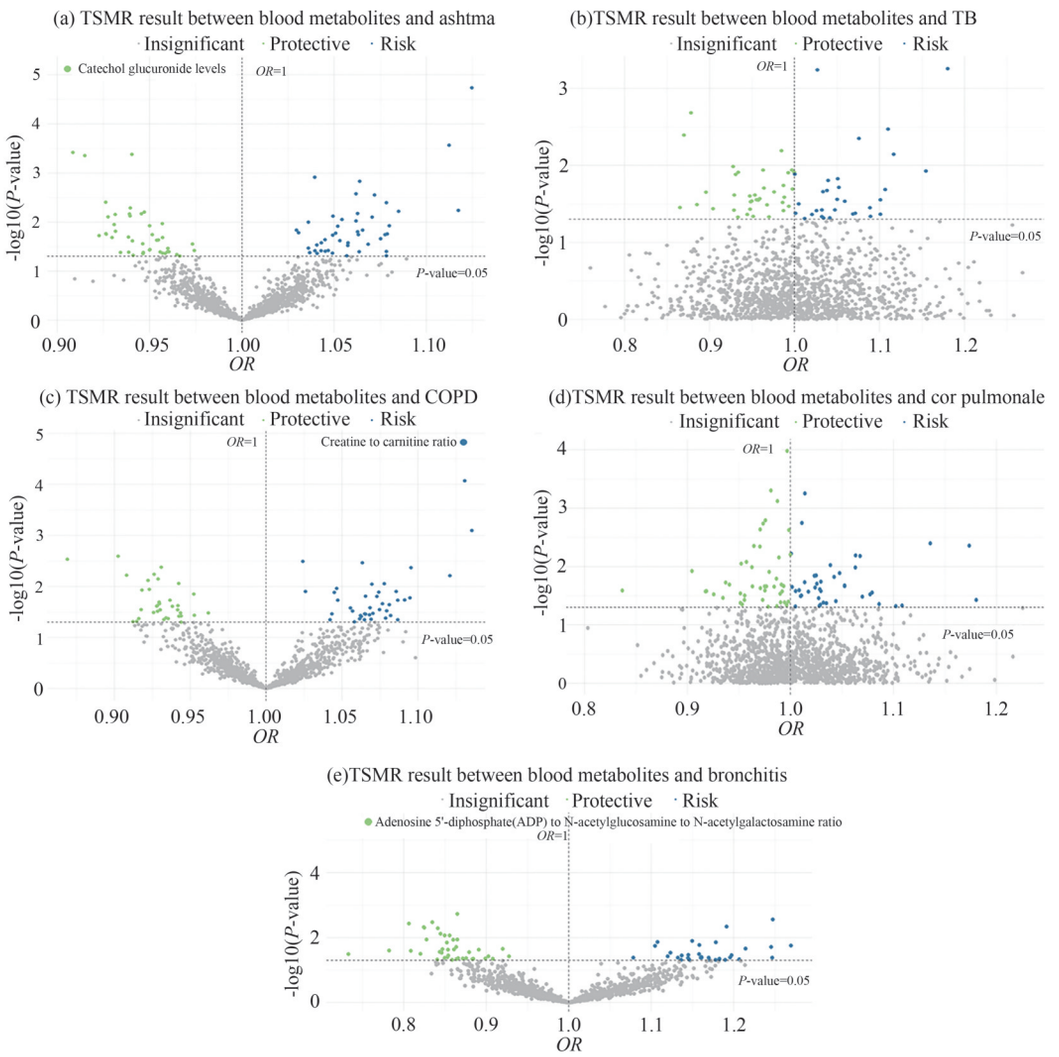


**Figure 2 Volcano map of TSMR results**

## 2.4 Direction Validation

Direction validation confirms a one-way causal link from exposure to outcome, and the Steiger test verifies that SNPs affect the outcome via the exposure rather than the reverse. We conducted the Steiger test on the 16 aforementioned metabolites, and the results showed that all Steiger_Pval values were less than 0.05, and the correct causal direction results were all TRUE, indicating that there were no reverse causal relationships between blood metabolites and outcomes.

## 2.5 Meta-Analysis

Replication analysis enhances the credibility of causal inference by repeatedly validating MR results across different datasets or cohorts. Meta-analysis can synthesize the results of multiple independent MR studies to provide more precise effect estimates. $I^2$ reflects the proportion of total variation that is due to non-sampling error. When $I^2$> 50%, a random-effect model is used; when $I^2$<50%, a fixed-effect model is applied. If the *P*-value is below 0.05, the result is regarded as statistically significant. An $OR_{combined}$<1 indicates the metabolite is protective for the disease, whereas $OR_{combined}$>1 indicates it is a risk factor. We performed a Meta-analysis of the MR estimates derived from both the primary and replication TSMR analyses. The results indicated that 3 candidate metabolic biomarkers were statistically significant for asthma, 3 for COPD, and 4 for cor pulmonale. In asthma, the following metabolites

were protective: catechol glucuronide levels ($OR_{combined}$<1, $P$=0.02), histidine levels ($OR_{combined}$<1, $P$=0.02), and the cholesterol to linoleoyl-arachidonoyl-glycerol (18∶2 to 20∶4)[2] ratio ($OR_{combined}$<1, $P$=0.02). In COPD, the following metabolites were associated with significant risk: creatine to carnitine ratio ($OR_{combined}$>1, $P$=0.02), orotidine levels ($OR_{combined}$>1, $P$<0.01), and glycerol to sulfate ratio ($OR_{combined}$>1, $P$<0.01). In cor pulmonale, the following metabolites were protective: 1-palmitoyl-2-dihomo-linolenoyl-GPC (16∶0/20∶3n3 or 6) levels ($OR_{combined}$<1, $P$<0.01) and 1,2-dilinoleoyl-GPE (18∶2/18∶2) levels ($OR_{combined}$<1, $P$=0.02). However, gamma-glutamylglycine levels ($OR_{combined}$>1, $P$<0.01) and glycine levels ($OR_{combined}$>1, $P$<0.01) were linked to significant risk. Meta-analysis summary graph is shown in Figure 3.

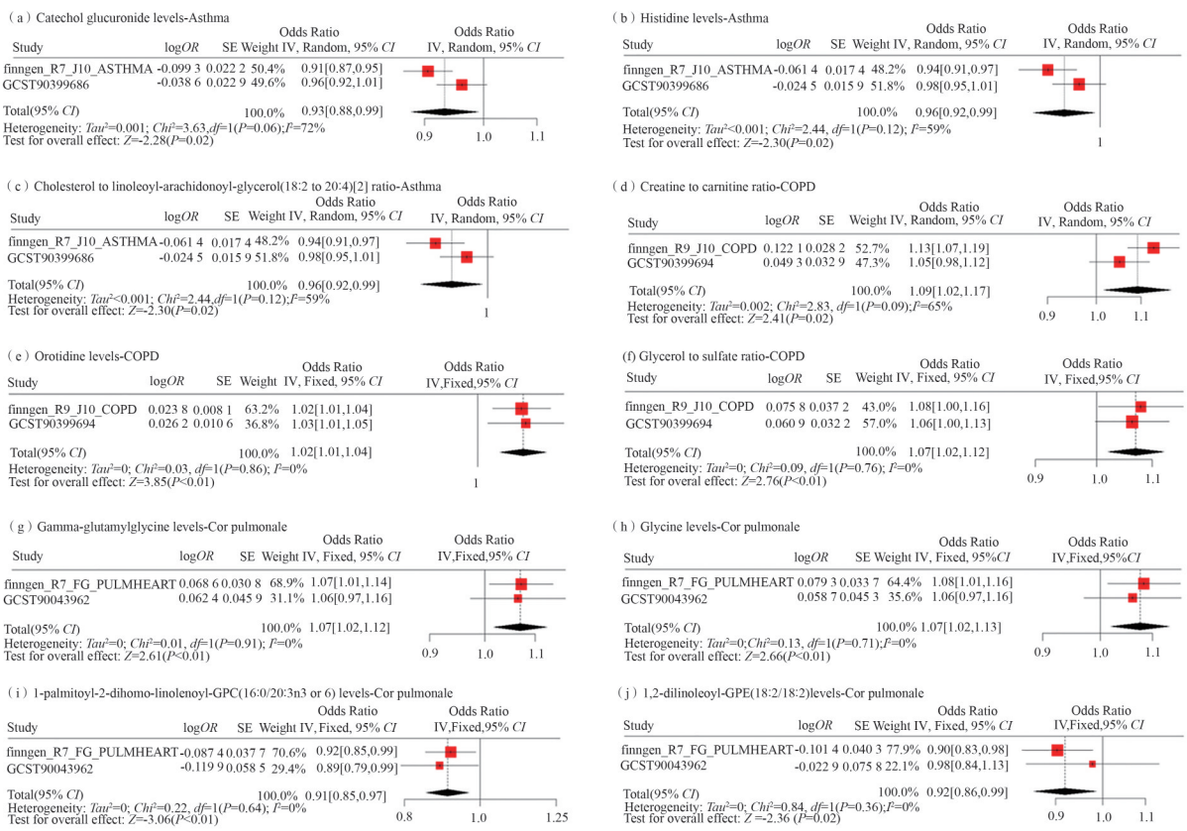


**Figure 3 Summary plot of Meta-analysis of 10 metabolites and 3 diseases**

## 2.6 Metabolic Pathway Analysis

We analyzed the metabolic pathways of known metabolites with $P_{IVW}$<0.05, identifying 14 important disease-related metabolic pathways (Figure 4). Specifically, we found 6 pathways related to asthma, 2 pathways related to TB, 1 pathway related to COPD, 4 pathways related to cor pulmonale, and 1 pathway related to bronchitis. The results indicate that glycine, serine and threonine metabolism ($P$=0.000 2), glycerophospholipid metabolism ($P$=0.004 1), arginine biosynthesis ($P$=0.007 2), butanoate metabolism ($P$=0.008 3), alanine, aspartate and glutamate metabolism ($P$=0.027 7), and glyoxylate and dicarboxylate metabolism ($P$=0.035 6) might be associated with asthma; tryptophan metabolism ($P$=0.021 2) and taurine and hypotaurine metabolism ($P$=0.044 9) might be associated with TB; caffeine metabolism ($P$=0.049 8) might be associated with COPD; glycine, serine and threonine metabolism ($P$=0.014 0), biosynthesis of unsaturated fatty acids ($P$=0.016 5), primary bile acid biosynthesis ($P$=0.026 4), and linoleic acid metabolism ($P$=0.028 3) might be associated with cor pulmonale; pentose and glucuronate interconversions ($P$=0.007 1) might be associated with bronchitis. Notably, asthma and cor pulmonale share a common metabolic pathway: glycine, serine and threonine metabolism.

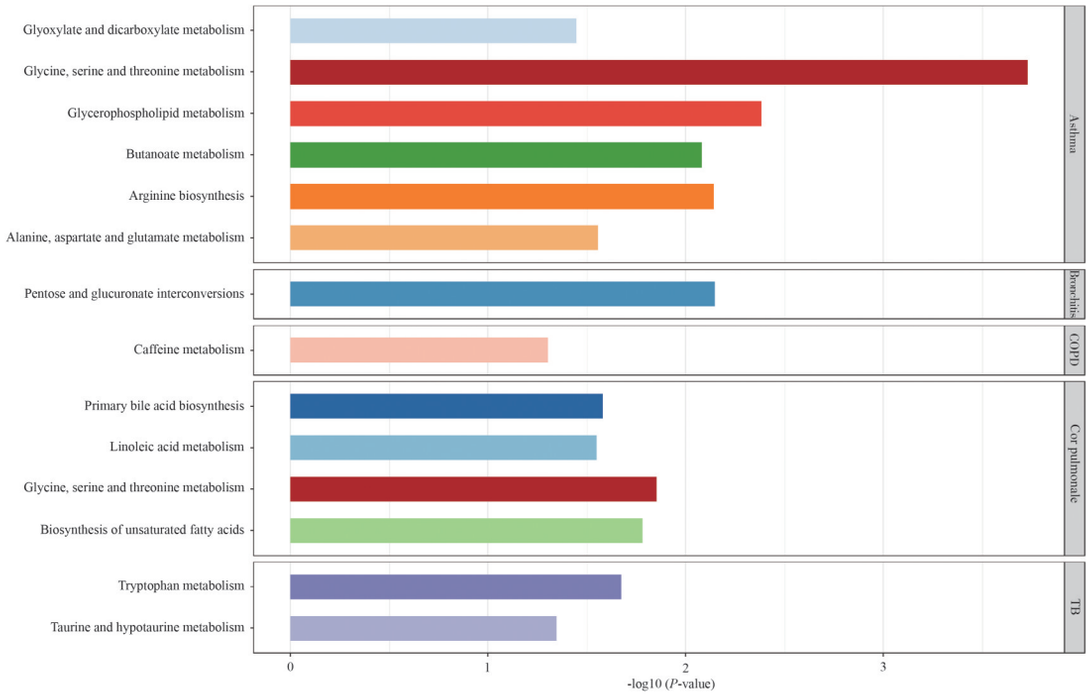


**Figure 4 Significant metabolic pathways of 5 respiratory diseases**

# 3 Statistical Analysis

## 3.1 Asthma

Consistent with previous studies, our results indicate that asthma is related to glycine, serine and threonine metabolism. Glycine, the simplest amino acid and a non-essential amino acid, often serves as a basic structural unit in protein synthesis. In the body, glycine participates in the urea cycle, helping to eliminate nitrogen metabolism products. Additionally, glycine plays a crucial role in collagen synthesis, maintaining the health of skin, muscles, and bones. Serine and threonine are essential amino acids that help maintain specific protein structures and regulate protein activity through modifications such as phosphorylation. Glycine and serine are also important in carbohydrate metabolism, particularly in glycogen synthesis and degradation. Moreover, glycine and serine participate in antioxidant reactions, helping remove free radicals and protect cells from oxidative damage. Recent studies have shown significant changes in plasma metabolites in children with worsening asthma, with glycine, serine and threonine metabolism being crucial pathways distinguishing between exacerbated and stable asthma in children[34-37]. Liao et al.'s experiments found that the adverse effects of asthma caused by exposure to air pollutants are related to decreased levels of glycine, serine and threonine metabolism[38], consistent with Wang et al.'s findings in mouse experiments[39]. The pathogenesis of asthma is complex, with oxidative stress considered an important factor, as asthma patients have significantly higher oxidative stress levels than the general population. Given that serine is a precursor for the synthesis of glycine and cysteine, and that glycine and cysteine are essential for the synthesis of antioxidant glutathione, glycine, serine and threonine metabolism might prevent asthma by influencing oxidative stress pathways.

## 3.2 TB

The relationship between tryptophan metabolism and TB has garnered significant interest. Consistent with previous studies, we posit that TB is linked to tryptophan metabolism. Tryptophan, an essential amino acid, is a component of proteins and a precursor to neurotransmitters such as melatonin. Research indicates that in organisms infected with TB, tryptophan metabolism often alters, leading to decreased tryptophan levels and subsequently affecting the immune system's function[40]. This metabolic disorder may impair the function of monocytes and T cells, increasing susceptibility to bacterial invasion and exacerbating TB progression. Tryptophan metabolism is inti-

mately associated with the survival and growth of *Mycobacterium tuberculosis*. Studies have shown that *M. tuberculosis* manipulates the host's tryptophan metabolism pathway to ensure its survival after infection[41]. Additionally, abnormalities in tryptophan metabolism have been observed in patients with intracranial tuberculosis[42], suggesting that tryptophan metabolism may be involved in the localization and proliferation of *M. tuberculosis* in the nervous system. Moreover, some studies have found that intervening in the tryptophan metabolism pathway can reduce the incidence and progression of TB. For instance, some inhibitors of the tryptophan metabolism pathway have demonstrated therapeutic effects on TB[43]. Exploring the potential applications of tryptophan metabolism in TB treatment, understanding and regulating this pathway, may provide new approaches for TB prevention and treatment.

### 3.3 COPD

Caffeine is a chemical commonly found in foods such as coffee, chocolate, and tea. It plays a significant role in the stimulation of the respiratory system, making it a common ingredient in bronchodilators. Our research indicates that COPD is related to caffeine metabolism. Caffeine has both positive and negative effects on the human body. On one hand, it can prevent oxidative stress associated with Alzheimer's disease; on the other hand, it can negatively impact individuals with hypertension. The effect of caffeine on COPD is similarly dual-faceted. Some studies have found that caffeine has anti-inflammatory and antioxidant properties, which help alleviate symptoms and inflammatory responses in COPD patients, drinking coffee significantly benefits the respiratory system[44]. Other studies also suggest that caffeine can dilate the bronchi and improve respiratory function, thereby alleviating dyspnea symptoms. However, there are also findings indicating that excessive caffeine intake may adversely affect COPD patients. Within a certain range, an increased intake of coffee and caffeine has been demonstrated to exert a more pronounced pro-inflammatory effect on COPD[45]. In summary, the relationship between caffeine and COPD is complex and multifaceted, necessitating further research and exploration.

### 3.4 Cor Pulmonale

Similar to asthma, our study indicates that the occurrence of cor pulmonale is influenced by glycine, serine and threonine metabolism. Currently, there is no direct evidence linking glycine, serine and threonine metabolism with cor pulmonale, but evidence suggests an association with pulmonary arterial hypertension, a major cause of cor pulmonale[46]. Xu et al. confirmed that glycine, serine and threonine metabolism can promote one-carbon metabolism[47], while Xu et al. observed an increased one-carbon metabolism in pulmonary artery endothelial cells from patients with pulmonary hypertension[48]. At present, many mysteries remain regarding the relationship between glycine, serine and threonine metabolism and cor pulmonale. For example, it is still unclear whether the abnormal metabolism of these amino acids is a mechanism of cor pulmonale pathogenesis or merely a byproduct of its pathophysiological process. Furthermore, detailed mechanisms of these amino acids' metabolism need to be further studied to provide more targeted methods for the treatment and prevention of cor pulmonale.

### 3.5 Bronchitis

Bronchitis is a prevalent respiratory condition characterised by inflammatory processes within the airways and an increased production of mucus. Our research suggests that bronchitis is associated with pentose and glucuronate interconversions. Pentose and glucuronate are 2 different metabolic products that can be interconverted. Pentose is a sugar containing 5 carbon atoms, while glucuronate is a compound derived from glucose through a series of reactions. The interconversions of pentose and glucuronate include D-ribulose 5-phosphate, xylitol, and glucose-1-phosphate. Glucuronate can be converted into xylulose 5-phosphate via the uronic acid pathway, which is part of the pentose phosphate pathway[49]. Oxidative stress is well known to be a key factor in the pathophysiology of bronchitis. The respiratory system is frequently exposed to environmental pollutants and toxins, leading to the generation of reactive oxygen species and subsequent oxidative damage. In order to counteract this oxidative stress, the antioxidant system within the respiratory tract, including the pentose phosphate pathway, is activated to maintain redox balance and protect cells from damage. Additionally, the pentose phosphate pathway participates in producing ribose-5-phosphate, which

serves as an essential precursor for the production of nucleotides and nucleic acids. These are important components for cellular growth and repair processes, including the regeneration of bronchial epithelial cells. There may be an association between pentose and glucuronate interconversions and bronchitis, but research on their specific relationship is currently limited. Further experiments and studies are needed to confirm this association.

### 3.6 Discussion

It is noteworthy that the present study examined 5 respiratory diseases and identified significant shared molecular mechanisms involving glycine, serine and threonine metabolism in asthma and cor pulmonale.

Firstly, defects in glutathione synthesis and oxidative stress imbalance represent core pathological pathways. Reduced levels of glycine and serine, which are precursors for glutathione synthesis, lead to glutathione deficiency, exacerbating airway inflammation in asthma patients and pulmonary artery endothelial cell apoptosis in patients with cor pulmonale. Secondly, abnormalities in one-carbon metabolism have been demonstrated to drive pathological proliferation. SHMT2 has been shown to promote eosinophil activation in asthma and pulmonary artery vascular remodelling in cor pulmonale by generating glycine and one-carbon units. The metabolites of the former provide substrates for nucleic acid synthesis, thereby fueling immune cell proliferation and vascular smooth muscle cell hyperplasia. In order to address these mechanisms, therapeutic strategies are required that regulate metabolism and block pathological pathways concurrently.

### 3.7 Strengths and Limitations

This study has several notable strengths. Firstly, the study uses the most extensive and systematic blood metabolite dataset to date as exposure, incorporating 1 352 blood metabolites into the MR analysis. Secondly, the study comprehensively explores the causal relationships between blood metabolites and 5 respiratory diseases, representing potentially the most in-depth and systematic examination of metabolic processes in respiratory diseases to date. Thirdly, the study integrates a range of methodologies and designs a series of experiments to ensure the reliability of MR estimation results. Fourthly, the study discloses shared molecular mechanisms between asthma and cor pulmonale, thus paving the way for common therapeutic strategies.

However, there are some limitations to our study. Firstly, limitations in data quality and genetic IV may compromise the reliability of causal inference. Secondly, insufficient sample size and diversity restrict the generalizability of findings. Furthermore, existing studies predominantly rely on cross-sectional data, making it difficult to capture dynamic changes in metabolites and inadequately accounting for the complexity of environmental factors.

## 4 Conclusion

This study systematically analyzed the causal effects of blood metabolites on 5 respiratory diseases using an MR framework based on GWAS data, identified molecular targets with potential intervention value, and uncovered shared metabolic pathways across diseases.

## Reference:

[1] Levine S M, Marciniuk D D. Global impact of respiratory disease: What can we do, together, to make a difference?[J]. Chest, 2022, 161(5): 1153-1154.

[2] Ferkol T, Schraufnagel D. The global burden of respiratory disease[J]. Annals of the American Thoracic Society, 2014, 11(3): 404-406.

[3] Kirsch D B, Józefowicz R F. Neurologic complications of respiratory disease[J]. Neurologic Clinics, 2002, 20(1): 247-264.

[4] Gea J, Casadevall C, Pascual S, et al. Respiratory diseases and muscle dysfunction[J]. Expert Review of Respiratory Medicine, 2012, 6(1): 75-90.

[5] Soriano J B, Kendrick P J, Paulson K R, et al. Prevalence and attributable health burden of chronic respiratory diseases, 1990—2017: A systematic analysis for the Global Burden of Disease Study 2017[J]. The Lancet Respiratory Medicine, 2020, 8(6):

585-596.

[6] Mamas M, Dunn W B, Neyses L, et al. The role of metabolites and metabolomics in clinically applicable biomarkers of disease[J]. Archives of Toxicology, 2011, 85(1): 5-17.

[7] Snowden S, Dahlén S E, Wheelock C E. Application of metabolomics approaches to the study of respiratory diseases[J]. Bioanalysis, 2012, 4(18): 2265-2290.

[8] Devillier P, Salvator H, Naline E, et al. Metabolomics in the diagnosis and pharmacotherapy of lung diseases[J]. Current Pharmaceutical Design, 2017, 23(14): 2050-2059.

[9] Shi M Y, Han S Y, Klier K, et al. Identification of candidate metabolite biomarkers for metabolic syndrome and its five components in population-based human cohorts[J]. Cardiovascular Diabetology, 2023, 22(1): 141.

[10] Comhair S A A, McDunn J, Bennett C, et al. Metabolomic endotype of asthma[J]. The Journal of Immunology, 2015, 195(2): 643-650.

[11] Collins J M, Walker D I, Jones D P, et al. High-resolution plasma metabolomics analysis to detect *Mycobacterium tuberculosis*-associated metabolites that distinguish active pulmonary tuberculosis in humans[J]. PloS One, 2018, 13(10): e0205398.

[12] Cruickshank-Quinn C I, Jacobson S, Hughes G, et al. Metabolomics and transcriptomics pathway approach reveals outcome-specific perturbations in COPD[J]. Scientific Reports, 2018, 8: 17132.

[13] Emdin C A, Khera A V, Kathiresan S. Mendelian randomization[J]. JAMA, 2017, 318(19): 1925-1926.

[14] Smith G D, Ebrahim S. Mendelian randomization: Prospects, potentials, and limitations[J]. International Journal of Epidemiology, 2004, 33(1): 30-42.

[15] Lawlor D A, Harbord R M, Sterne J A C, et al. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology[J]. Statistics in Medicine, 2008, 27(8): 1133-1163.

[16] Draisma H H M, Pool R, Kobl M, et al. Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels[J]. Nature Communications, 2015, 6: 7208.

[17] Cao S R, Gu Y Q, Lu G Y, et al. Causal correlations between plasma metabolites, inflammatory proteins, and chronic obstructive pulmonary disease: A Mendelian randomization and bioinformatics-based investigation[J]. Journal of Inflammation Research, 2025, 18: 4057-4073.

[18] Zhang Y, Ma X L, Wu F, et al. The relationship between genetic prediction of 486 blood metabolites and the risk of COPD: Mendelian randomization study[J]. Scientific Reports, 2025, 15(1): 7349.

[19] Zhao C Y, Song X F, Yang D, et al. A causal relationship between circulating immune cells, plasma metabolites, and pulmonary diseases: A mediated Mendelian randomization study[J]. Journal of Genetic Engineering and Biotechnology, 2025, 23(3): 100537.

[20] Lin L Y, Wu Z L, Zhong A Q, et al. Causal effects of genetically determined blood metabolites on asthma: A bidirectional Mendelian randomization study[J]. Journal of Asthma, 2024, 61(12): 1727-1737.

[21] Cao Z H, Wu T, Fang Y K, et al. Dissecting causal relationships between immune cells, plasma metabolites, and COPD: A mediating Mendelian randomization study[J]. Frontiers in Immunology, 2024, 15: 1406234.

[22] Johnson R K, Brunetti T, Quinn K, et al. Discovering metabolite quantitative trait loci in asthma using an isolated population[J]. Journal of Allergy and Clinical Immunology, 2022, 149(5): 1807-1811.

[23] Wang L Z, Lyu Z. Causal associations among gut microbiota, 1 400 plasma metabolites, and asthma: A two-sample Mendelian randomization study[J]. Frontiers in Molecular Biosciences, 2024, 11: 1370919.

[24] Chen Y H, Lu T Y, Pettersson-Kymmer U, et al. Genomic atlas of the plasma metabolome prioritizes metabolites implicated in human diseases[J]. Nature Genetics, 2023, 55(1): 44-53.

[25] Kurki M I, Karjalainen J, Palta P, et al. FinnGen: Unique genetic insights from combining isolated population and national health register data[PP/OL]. medRxiv(2022-03-06)[2025-09-01]. https://doi.org/10.1101/2022.03.03.22271360.

[26] Burgess S, Thompson S G, CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies[J]. International Journal of Epidemiology, 2011, 40(3): 755-764.

[27] Slob E A W, Burgess S. A comparison of robust Mendelian randomization methods using summary data[J]. Genetic Epidemiology, 2020, 44(4): 313-329.

[28] Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing[J]. Journal of the Royal Statistical Society: Series B(Statistical Methodology), 1995, 57(1): 289-300.

[29] Bowden J, Smith G D, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression[J]. International Journal of Epidemiology, 2015, 44(2): 512-525.
[30] Bowden J, Smith G D, Haycock P C, et al. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator[J]. Genetic Epidemiology, 2016, 40(4): 304-314.
[31] Hartwig F P, Smith G D, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption[J]. International Journal of Epidemiology, 2017, 46(6): 1985-1998.
[32] del Greco M F, Minelli C, Sheehan N A, et al. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome[J]. Statistics in Medicine, 2015, 34(21): 2926-2940.
[33] Hemani G, Tilling K, Smith G D. Orienting the causal relationship between imprecisely measured traits using GWAS summary data[J]. PLoS Genetics, 2017, 13(11): e1007081.
[34] Lukkarinen H, Söderlund-Venermo M, Vuorinen T, et al. Children with severe asthma have unique oxidative stress-associated metabolomic profiles[J]. Journal of Allergy and Clinical Immunology, 2014, 133(1): 258-261.
[35] Tian M, Chen M, Bao Y L, et al. Sputum metabolomic profiling of bronchial asthma based on quadruple time-of-flight mass spectrometry[J]. International Journal of Clinical and Experimental Pathology, 2017, 10(10): 10363-10373.
[36] Cottrill K A, Stephenson S T, Mohammad A F, et al. Exacerbation-prone pediatric asthma is associated with arginine, lysine, and methionine pathway alterations[J]. Journal of Allergy and Clinical Immunology, 2023, 151(1): 118-127.
[37] Cottrill K A, Chandler J D, Kobara S, et al. Metabolomics identifies disturbances in arginine, phenylalanine, and glycine metabolism as differentiating features of exacerbating atopic asthma in children[J]. Journal of Allergy and Clinical Immunology: Global, 2023, 2(3): 100115.
[38] Liao J W, Gheissari R, Thomas D C, et al. Transcriptomic and metabolomic associations with exposures to air pollutants among young adults with childhood asthma history[J]. Environmental Pollution, 2022, 299: 118903.
[39] Wang Z T, Gao S L, Xie J F, et al. Identification of multiple dysregulated metabolic pathways by GC-MS-based profiling of liver tissue in mice with OVA-induced asthma exposed to $PM_{2.5}$[J]. Chemosphere, 2019, 234: 277-286.
[40] Collins J M, Siddiqa A, Jones D P, et al. Tryptophan catabolism reflects disease activity in human tuberculosis[J]. JCI Insight, 2020, 5(10): e137131.
[41] Lott J S. The tryptophan biosynthetic pathway is essential for *Mycobacterium tuberculosis* to cause disease[J]. Biochemical Society Transactions, 2020, 48(5): 2029-2037.
[42] Ardiansyah E, Pacheco J A, Nhat L T H, et al. Tryptophan metabolism determines outcome in tuberculous meningitis: A targeted metabolomic analysis[PP/OL]. MedRxiv(2023-01-09)[2025-09-01]. https://doi.org/10.1101/2023.01.08.23284316.
[43] Consalvi S, Scarpecci C, Biava M, et al. Mycobacterial tryptophan biosynthesis: A promising target for tuberculosis drug development?[J]. Bioorganic & Medicinal Chemistry Letters, 2019, 29(23): 126731.
[44] Moitra S, Bandyopadhyay A, Lacy P. Metabolomics of respiratory diseases[M]//Ghini V, Stringer K A, Luchinat C. Metabolomics and its impact on health and diseases. Cham: Springer, 2022: 339-365.
[45] Liao W Z, Li J X, Feng W Y, et al. Association between coffee and caffeine intake and risk of COPD: Findings based on NHANES 2007—2012[J]. Heart & Lung, 2024, 67: 53-61.
[46] Xu W L, Janocha A J, Erzurum S C. Metabolism in pulmonary hypertension[J]. Annual Review of Physiology, 2021, 83: 551-576.
[47] Xu W L, Comhair S A A, Chen R Y, et al. Integrative proteomics and phosphoproteomics in pulmonary arterial hypertension[J]. Scientific Reports, 2019, 9: 18623.
[48] Xu B Z, Huang C H, Zhang C J, et al. NMR-based metabolomic analysis of plasma in patients with adult congenital heart disease and associated pulmonary arterial hypertension: A pilot study[J]. Metabolites, 2022, 12(9): 845.
[49] Alfaro T M, Monteiro R A, Cunha R A, et al. Chronic coffee consumption and respiratory disease: A systematic review[J]. The Clinical Respiratory Journal, 2018, 12(3): 1283-1294.

**责任编辑：** 刘 敏