Association between genetic prediction of 486 blood metabolites and the risk of idiopathic pulmonary fibrosis: A mendelian randomization study

FAN WU[1], BOYANG LI[1], JIAQING LI[1], WEISHAN YUAN[1], XUE ZHU[1] and XUE LIU[2]

[1]First Clinical Medical College, Shandong University of Traditional Chinese Medicine, Jinan, Shandong 250011, P.R. China; [2]Department of Pulmonary and Critical Care Medicine, Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan, Shandong 250011, P.R. China

Received September 17, 2024; Accepted January 16, 2025

DOI: 10.3892/br.2025.1930

Abstract. Metabolic disorders are a significant feature of fibrotic diseases. Nevertheless, the lack of sufficient proof regarding the cause-and-effect association between circulating metabolites and the promotion or prevention of idiopathic pulmonary fibrosis (IPF) persists. To assess the causal association between IPF and genetic proxies of 486 blood metabolites, a dual sample Mendelian randomization (MR) analysis was performed. Therefore, the two-sample MR technique and genome-wide association study data were employed to assess the association between 486 serum metabolites and IPF. To produce the primary outcomes, the inverse variance weighted (IVW) technique was applied, while to assess the stability and dependability of the outcomes, sensitivity analysis using MR-Egger analysis was performed. Additionally, weighted median, Cochran's Q-test, Egger intercept test and the leave-one-out method were used. The results of the present study revealed a total of 21 metabolites in blood circulation that could affect the risk of IPF (P_{IVW}<0.05). Among them, 10 compounds were already known, namely cotinine [odds ratio (OR)=1.206; 95% confidence interval (CI), 1.002-1.452; P=0.047], hypoxanthine (OR=0.225; 95% CI, 0.056-0.899; P=0.034), aspartyl phenylalanine (OR=4.309; 95% CI, 1.084-17.131; P=0.038), acetyl-carnitine (OR=5.767; 95% CI, 1.398-23.789; P=0.015), 2-aminobutyrate (OR=0.155; 95% CI, 0.033-0.713; P=0.016), Docosapentaenoic acid (PubChem ID: 5497182; OR=0.214; 95% CI, 0.055-0.833; P=0.026), octanoyl-carnitine (PubChem ID: 177508; OR=3.398; 95% CI, 1.179-9.794; P=0.023), alpha-hydroxy-isovalerate (PubChem ID: 857803-94-2; OR=0.324; 95% CI, 0.112-0.931; P=0.036), 1,7-dimethylurate (PubChem ID: 91611; OR=0.401; 95% CI, 0.172-0.931; P=0.033) and 1-linoleoyl-glycerophosphocholine (PubChem ID: 657272; OR=6.559; 95% CI, 1.060-40.557; P=0.043). Additionally, the study also identified 11 currently unknown chemical structures. The results of Cochran's Q-test indicated that there was no significant heterogeneity, while MR-Egger's intercept analysis verified the lack of horizontal pleiotropy. The retention of one method for plotting also supported the reliability of the MR analysis. Overall, the results of the current study supported the cause-and-effect association between IPF and 21 blood metabolites, including 10 with already known chemical composition and 11 which are still awaiting determination. These findings could provide novel insights for the further investigation of the mechanism underlying the development of IPF.

Correspondence to: Dr Xue Liu, Department of Pulmonary and Critical Care Medicine, Affiliated Hospital of Shandong University of Traditional Chinese Medicine, 42 Wenhua West Road, Lixia, Jinan, Shandong 250011, P.R. China
E-mail: liuxueszy@126.com

Key words: idiopathic pulmonary fibrosis, blood metabolites, Mendelian randomization, inverse variance weighting

Introduction

Idiopathic pulmonary fibrosis (IPF), a progressive fibrotic lung ailment (1), exhibits notably unfavorable prognosis, with a median survival rate of ~3-4 years subsequent to diagnosis (2). This condition frequently culminates in fatality due to respiratory failure, primarily caused by the extensive restructuring of the pulmonary tissue (3). Antifibrotic treatment has shown effectiveness in reducing the advancement of the disease and lessening the deterioration of lung function in individuals with IPF. Furthermore, other studies demonstrated that the survival rates of patients with IPF treated with antifibrotic medications were enhanced compared with untreated ones (4,5). However, several patients stop treatment due to the onset of adverse effects (6). Genetic studies suggested that the susceptibility to IPF could be significantly affected by genetics. Therefore, the investigation of sporadic and familial cases promoted the successful identification of multiple IPF-related variants (7). The establishment and verification of evidence-supported recommendations for gene-oriented screening for IPF could enable the redefinition and reclassification of this particular illness based on its molecular characteristics, thus promoting the adoption of precision medicine strategies (8). Therefore, the timely detection of possible genetic risk factors could

prevent the onset of IPF. Currently, there are limited methods for the early detection of IPF. However, the early detection and intervention are crucial for slowing down the progression of the disease. Due to their powerful monitoring capabilities, wide applicability and accessible analysis, serum metabolites have become promising biomarkers for IPF.

The understanding of disease progression has been improved with the introduction of transcriptomics, proteomics and metabolomics, thus providing novel insights into the cellular mechanisms involved. It has been reported that regulating metabolic balance could affect IPF (9). Pulmonary fibrosis can occur due to irregular collagen production and disrupted airway remodeling caused by metabolic abnormalities in alveolar epithelial cells, myofibroblasts and macrophages. Several metabolic pathways are involved in lipid, protein and carbohydrate metabolism, and more particularly in IPF (10). Metabolomics, via identifying altered metabolites or metabolic pathways, offer a fresh perspective for investigating the underlying mechanisms of different diseases, thus providing valuable insights into the biological processes associated with these diseases (11). Metabolic changes are not just a characteristic of fibrosis development, but they can also significantly affect the development of fibrosis, particularly in IPF (12). It has been also suggested that dysfunctions in the metabolic processes of glycine, glutamine and arginine, along with irregularities in glycolysis, can activate a pro-fibrotic feature via a TGF-β-dependent mechanism (10,13,14). Increased content of extended and moderate chain fatty acids has been documented in the lungs of patients with IPF. In addition, previous studies reported an enhanced reprogramming of macrophages, resulting in IPF improvement via fatty acid β-oxidation (15,16). Another study demonstrated that metabolites derived from arachidonic acid could hinder pulmonary fibrosis via altering the TGF-β1-induced pro-fibrotic signaling pathway (17). The aforementioned finding indicated that targeting these metabolites could be a promising approach for treating IPF (17). However, the effect and relevance of serum metabolomic profiles on IPF remains understudied.

Investigating metabolites associated with the pathogenesis and progression of IPF is not just significant for early detection and prevention, but also serve a key role in comprehending the biological mechanisms of IPF therapy and uncovering possible therapeutic objectives. Several metabolites associated with IPF have been identified, such as Krebs von den Lungen 6, matrix metalloproteinases and periostin (18). Nevertheless, the uncertain causal association between metabolites and IPF persists, since there is a dearth of prospective studies investigating metabolites and IPF. Randomized controlled trials are widely considered as the benchmark for establishing causal effects. However, the complexities and obstacles in this particular scenario pose challenges in reaching conclusive findings regarding the causal association between metabolites and IPF. Mendelian randomization (MR) is an extensively employed and novel epidemiological approach. In the absence of randomized controlled trials, MR stands out as a highly effective method for investigating causal associations between exposures of interest and outcomes. This can be achieved via utilizing genetic variants as instrumental variables (IVs) to estimate the causal associations between exposures and outcomes. Furthermore, MR can help in reducing confounding factors and the bias of reverse causality, which is inherent in observational studies. Due to the limited comprehension of the cause-and-effect association between blood metabolites and IPF, more studies are urgently needed. Therefore, the current study employed MR analysis to thoroughly investigate the potential causality of 486 blood metabolites in the progression of IPF via using pooled data from genome-wide association studies (GWAS). Exploring these particular metabolites and their association with IPF could promote the improved understanding of the pathogenesis of IPF and lay a solid foundation for exploring the potential mechanisms and therapeutic targets of IPF in the future.

Materials and methods

Data of 486 blood metabolites and IPF from GWAS. Data for 486 metabolites were obtained from the Metabolomics GWAS summary database (https://metabolomics.helmholtz-muenchen.de/gwas/), which includes a total of 2,163,597 single nucleotide polymorphisms (SNPs) associated with them (19). The names of the 486 metabolites are listed in detail in Table SI. X indicates an unidentified chemical composition. The FinnGen biospecimen repository (https://www.finngen.fi/en) was utilized as an outcome variable for conducting GWAS analysis on IPF. In the current study the analysis included genotypic data from 1,028 patients with IPF and 196,986 controls (20).

Selection of IVs. To evaluate the possible cause-and-effect association between circulating metabolites and IPF, a two-sample MR analysis was conducted (21). The present study assessed blood metabolites and IPF as factors of exposure and outcome. The three essential assumptions (22) that should be met for IVs to be valid are depicted in Fig. 1. Therefore, it is crucial that the genetic variants employed as IVs exert a robust association with blood metabolites. Furthermore, genetic mutations should not be associated with any potential confounding variables. Thirdly, the genetic variations should exclusively affect IPF via blood metabolites rather than any alternative pathways. Therefore, to meet the aforementioned three assumptions, a significance threshold of $P<1x10^{-5}$ was set for autonomy (23). A lower P-value threshold, such as that of $P<1x10^{-5}$, could effectively control the false positive rate. However, it could also increase the risk of false negative results. Choosing this value could balance the two to a particular extent, ensuring that the results would have a certain level of reliability when discovering potential related SNPs. In large sample studies, lower P-value thresholds, such as $P<5x10^{-8}$, are commonly used in GWAS, since they can effectively reduce false positive results caused by multiple comparisons. However, if the sample size is small, using overly strict thresholds can miss out several truly relevant SNPs. Therefore, choosing a $P<1x10^{-5}$ could be more suitable for the current research scenario. Additionally, to guarantee the autonomy of SNPs and eliminate linkage disequilibrium (LD) (24), a LD threshold of $r^2<0.001$ and a distance of 10,000 kb were implemented. r^2 is a statistical measure of LD between two loci. Therefore, the closer the value is to 1, the stronger the LD between these two loci. $r^2<0.001$ indicates that there is almost no linkage imbalance between these two loci and they can

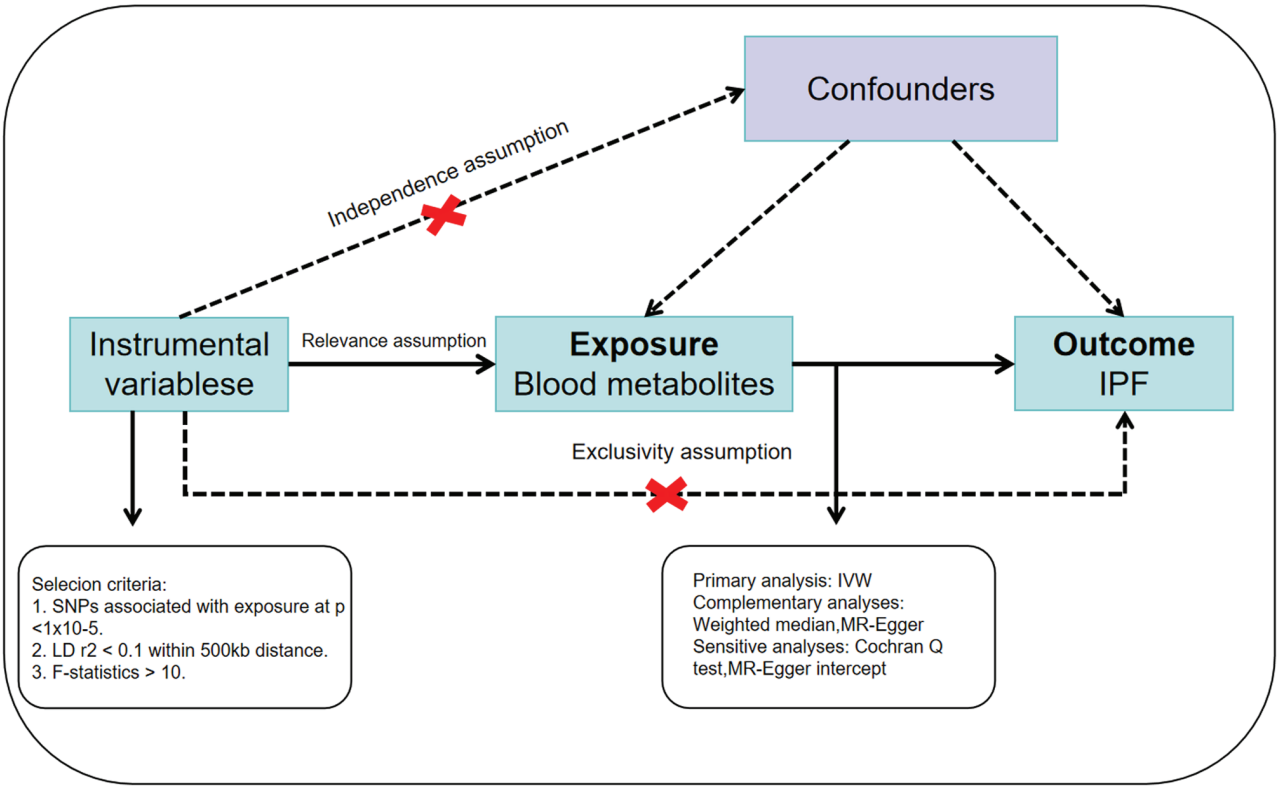

Figure 1. Overview of the analysis. MR, Mendelian randomization; IPF, idiopathic pulmonary fibrosis.

be therefore considered independent from each other. Setting r^2<0.001 could help identify markers that are independent of the phenotype and avoid LD-caused bias. This strict selection could improve the accuracy of the results, ensuring that the discovered effects are independent and not caused by the effects of other loci. Furthermore, the PhenoScanner (25) (http://www.phenoscanner.medschl.cam.ac.uk/) website was utilized to evaluate if these SNPs were associated with confounding variables in IPF. Currently, several risk factors for IPF have been identified, such as smoking, exposure to dust and reflux esophagitis. Therefore, in the present study, when any associations were identified, individuals with the aforementioned variations were excluded.

Statistical analysis. Various strong statistical methods were employed to guarantee the dependability and authenticity of the results. However, owing to its exceptional resilience, the inverse variance weighted (IVW) method was used as the main analytical approach (26). The IVW method is commonly applied in fixed- and random-effect models, which help mitigate the heterogeneity-related bias (27). Furthermore, the robustness of the findings was assessed via employing the MR-Egger regression and weighted median techniques. Several studieshave argued that the MR-Egger regression technique lacks statistical validity for estimating causality (27,28). Therefore, it was suggested that this method should only serve as a sensitivity analysis to assess the violation of core assumptions of IVs, rather than being used as a substitute for the IVW method (27). It is generally recognized that SNPs with F-statistics <10 are weakly instrumented variables, while those of >10 indicate no significant weakly instrumented variable bias. In the present study, SNPs with F-statistics of <10 in the MR analysis were excluded. To calculate the F value, the following formula was used: $F=R^2 \times (n-k-1)/(1-R^2) \times k$. The sample size, denoted by n, indicates the genetic variance denoted by R^2, while k indicates the number of SNPs present in the sample (29).

Sensitivity analysis. The heterogeneity of IVW and MR Egger regression methods was evaluated using Cochran's Q test and funnel plots. The results of Cochran's Q test (30) with P>0.05 indicated the absence of significant heterogeneity. The intercept in the MR-Egger regression represents the magnitude of horizontal pleiotropy. Therefore, the smaller the horizontal pleiotropy, the more the intercept tends to 0. In addition, to evaluate whether the findings were significantly affected by a solitary SNP, a leave-one-out examination was performed (31). This involved excluding each SNP at a time and subsequently conducting an MR analysis.

Results

Analysis of the associations between 486 blood metabolites and risk of IPF. After rigorous quality evaluation of the IVs, a total of 486 blood metabolites were obtained for MR analysis. LD analysis and removal of palindromic sequences to enhance accuracy and reliability of the selected IVs were also performed. Among the 486 blood metabolites identified, five appeared in two forms, thus resulting in a total of 491 metabolites. All SNPs demonstrated adequate validity as indicated by the F-statistics, which exceeded the empirical threshold of 10. The results of the analysis of the 486 metabolites are displayed in Table SII. Additionally, using the IVW methodology, a total

of 21 circulating metabolites were identified, that could exert a causal role in IPF risk. Among the aforementioned circulating metabolites, 10 were known compounds and 11 were unknown (Fig. 2). Namely, the circulating metabolites identified were the following: Cotinine (OR=1.206; 95% CI, 1.002-1.452; P=0.047), hypoxanthine (OR=0.225; 95% CI, 0.056-0.899; P=0.034), aspartyl phenylalanine (OR=4.309; 95% CI, 1.084-17.131; P=0.038), acetyl-carnitine (OR=5.767; 95% CI, 1.398-23.789; P=0.015), 2-aminobutyrate (OR=0.155; 95% CI, 0.033-0.713; P=0.016), docosapentaenoate (OR=0.214; 95% CI, 0.055-0.833; P=0.026), octanoyl-carnitine (OR=3.398; 95% CI, 1.179-9.794; P=0.023), alpha-hydroxy-isovalerate (OR=0.324; 95% CI, 0.112-0.931; P=0.036), 1,7-dimethylurate (OR=0.401; 95% CI, 0.172-0.931; P=0.033), 1-linoleoylglycerophosphocholine (OR=6.559; 95% CI, 1.060-40.557; P=0.043), X-10395 (OR=3.592; 95% CI, 1.263-10.210; P=0.016), X-10810 (OR=3.42; 95% CI, 1.263-10.210; P=0.016), X-11261 (OR=2.321; 95% CI, 1.009-5.335; P=0.047), X-11795 (OR=5.13; 95% CI, 1.150-22.885; P=0.032), X-12189 (OR=0.803; 95% CI; 0.669-0.964; P=0.019), X-12428 (OR=0.503; 95% CI, 0.259-0.973; P=0.041), X-12443 (OR=1.806; 95% CI, 1.051-3.104; P=0.032), X-12786 (OR=3.371; 95% CI, 1.206-9.423, P=0.020), X-13671 (OR=23.162; 95% CI; 3.091-173.553; P=0.002), X-14374 (OR=0.102; 95% CI; 0.011-0.936; P=0.043) and X-14745 (OR=4.947; 95% CI; 1.024-23.895; P=0.046). The analysis also revealed that the high levels of cotinine, aspartyl-phenylalanine, acetyl-carnitine, octanoyl-carnitine and 1-lipoyllglycophorophosphacholine were potential risk metabolites for IPF. By contrast, the high levels of hypoxanthine, 2-aminobutyrate, docosapentaenoate alpha hydroxy-isovalerate and 1,7-dimethylurate could reduce the risk of IPF. These findings could be helpful for detecting the incidence and predicting the prognosis of IPF in clinical practice. The results of the MR analysis are visualized in Table SIII, while those of sensitivity analysis are listed in Table I. Cochran's Q-test yielded insignificant heterogeneity, while MR-Egger's intercept analysis verified the lack of horizontal pleiotropy. The reliability of the MR analysis was also verified by the leave-one-out graphing (Fig. 3). The scatter plot is displayed in Fig. 4.

Discussion

The present study aimed to investigate the possible causal effect of 486 blood metabolites on IPF progression. Therefore, via combining two extensive GWAS datasets and implementing a rigorous MR design, a total of 13 genetic metabolites, namely cotinine, aspartyl phenylalanine, acetyl-carnitine, octanoyl-carnitine, 1-linoleoylglycerophosphocholine, X-10395, X-10810, X-11261, X-11795, X-12443, X-12786, X-13671 and X-14745, were identified, which were associated with enhanced susceptibility to IPF. By contrast, the analysis predicted eight genetic metabolites, namely hypoxanthine, 2-aminobutyrate, docosapentaenoate, alpha-hydroxy-isovalerate, 1,7-dimethylurate, X-12189, X-12428 and X-14374, that could reduce the risk of IPF.

The rising morbidity and significant mortality of IPF in recent years have imposed a substantial burden on individuals globally, emphasizing the need for prompt screening and prevention as a critical strategy. The emergence of metabolomics technologies has sparked growing curiosity in investigating the significance of metabolites associated with IPF. Due to the chemical nature of exposures originating from both internal and external sources, blood samples can serve as a means to identify exposure groups (32). A previous study supported the reliability of S100 calcium-binding protein A12 as a serum biomarker for assessing the severity and prognosis of IPF (33). 5-methoxytryptophan (5-MTP) has been recently identified as a tryptophan metabolite. Therefore, a study indicated that 5-MTP could combat inflammation, inhibit tumor growth, protect blood vessels and prevent fibrosis in kidney disorders. Additionally, the aforementioned study showed that 5-MTP could be also significantly involved in IPF and it was therefore anticipated to serve as a treatment strategy for pulmonary fibrosis (34). Glutaminolysis, a crucial metabolic pathway in IPF, involves the conversion of glutamine to glutamate by glutaminase and subsequently to α-ketoglutarate, a circulating metabolite of the tricarboxylic acid cycle (35). Emerging evidence has suggested that the immune system also exerts a significant role in the development of IPF. Therefore, it has been reported that IL-1β, IL-6, IL-23 and IL-17A, involved in type 17 immunity, play a significant role in both pulmonary fibrosis and acute exacerbation of pulmonary fibrosis. Targeting type 17 immunity could emerge as a novel therapeutic approach to hinder the advancement or onset of pulmonary fibrosis (36). Although the existing literature has provided conflicting results for particular metabolites, significant differences in the expression of serum metabolites have been identified between patients with IPF and healthy controls. However, further studies are needed to establish a definitive cause-and-effect association and gain a deeper comprehension of the underlying mechanisms. In the present study, the c MR method was used to elucidate the cause-and-effect association between blood metabolites and IPF, along with the metabolic pathways implicated. The present study aimed to offer guidance for the screening and management of IPF.

The study revealed that elevated levels of cotinine, aspartyl phenylalanine, acetyl-carnitine, octanoyl-carnitine and 1-linoleoylglycerophosphocholine were associated with a heightened susceptibility to IPF. A previous study demonstrated that cotinine, which serves as a biomarker for smoking exposure, was associated with an increased risk of kidney stones, coronary heart disease and chronic obstructive pulmonary disease (37). In addition, high levels of serum cotinine could also reduce muscle mass, hepatic steatosis and liver fibrosis (38-40). The present study verified that cotinine could be involved in elevated risk of IPF, thus suggesting that the smoking cessation management should be strengthened in clinical practice. Aspartyl phenylalanine, which is formed when cholecystokinin-8 breaks down, but not when angiotensin breaks down by angiotensin-converting enzyme, has been associated with an increased risk of acute coronary syndromes (41). Additionally, another study revealed a cause-and-effect association between aspartyl phenylalanine and hypertension and hypertriglyceridemia (41). The deficiency of alpha aspartyl phenylalanine hydrolase activity is considered to be the reason for the possible adverse reactions of aspartame intake (42,43). Therefore, reducing the intake of aspartame could be beneficial for preventing IPF. However, further research is needed to confirm this finding. Fatty acid metabolites, such

exposure.x	method	nsnp	pval	OR(95%CI)	or
cotinine	Inverse variance weighted	15	0.047562392	1.21(1.00 to 1.45)	1.2064783
	MR Egger	15	0.130712965	1.29(0.95 to 1.77)	1.2944580
	Weighted median	15	0.291452926	1.15(0.88 to 1.51)	1.1542492
hypoxanthine	Inverse variance weighted	21	0.034839368	0.23(0.06 to 0.90)	0.2250932
	MR Egger	21	0.722362228	0.50(0.01 to 21.49)	0.5007223
	Weighted median	21	0.250301067	0.30(0.04 to 2.33)	0.3006352
aspartylphenylalanine	Inverse variance weighted	5	0.038016710	4.31(1.08 to 17.13)	4.3095591
	MR Egger	5	0.455244431	5.67(0.11 to 302.50)	5.6711128
	Weighted median	5	0.080568033	4.71(0.83 to 26.84)	4.7142918
X-10395	Inverse variance weighted	29	0.016416881	3.59(1.26 to 10.21)	3.5924052
	MR Egger	29	0.163299246	2.93(0.67 to 12.74)	2.9298066
	Weighted median	29	0.048489124	4.48(1.01 to 19.87)	4.4801752
X-10810	Inverse variance weighted	16	0.016605714	3.42(1.25 to 9.34)	3.4176025
	MR Egger	16	0.272516890	2.33(0.55 to 9.91)	2.3268153
	Weighted median	16	0.100544005	3.16(0.80 to 12.51)	3.1634376
acetylcarnitine	Inverse variance weighted	18	0.015359067	5.77(1.40 to 23.79)	5.7676367
	MR Egger	18	0.477734138	3.96(0.10 to 162.45)	3.9639695
	Weighted median	18	0.382917199	2.52(0.32 to 19.98)	2.5154518
2-aminobutyrate	Inverse variance weighted	40	0.016673592	0.16(0.03 to 0.71)	0.1550673
	MR Egger	40	0.560806544	0.25(0.00 to 26.74)	0.2455331
	Weighted median	40	0.187864735	0.23(0.03 to 2.06)	0.2275049
docosapentaenoate (n3 DPA; 22:5n3)	Inverse variance weighted	10	0.026193605	0.21(0.06 to 0.83)	0.2141320
	MR Egger	10	0.650514433	0.41(0.01 to 17.38)	0.4057479
	Weighted median	10	0.033431888	0.12(0.02 to 0.85)	0.1249143
X-11261	Inverse variance weighted	16	0.047360276	2.32(1.01 to 5.34)	2.3212830
	MR Egger	16	0.124830637	5.38(0.71 to 40.48)	5.3752496
	Weighted median	16	0.021224382	3.73(1.22 to 11.40)	3.7252695
X-11795	Inverse variance weighted	12	0.032085026	5.13(1.15 to 22.89)	5.1304852
	MR Egger	12	0.207018727	18.79(0.27 to 1332.81)	18.7940363
	Weighted median	12	0.179797747	4.03(0.53 to 30.94)	4.0327712
X-12189	Inverse variance weighted	27	0.019036519	0.80(0.67 to 0.96)	0.8034431
	MR Egger	27	0.417886719	0.89(0.68 to 1.17)	0.8909256
	Weighted median	27	0.087715325	0.80(0.62 to 1.03)	0.8002393
X-12428	Inverse variance weighted	9	0.041332397	0.50(0.26 to 0.97)	0.5030240
	MR Egger	9	0.427908793	0.58(0.16 to 2.07)	0.5790571
	Weighted median	9	0.149447108	0.49(0.19 to 1.29)	0.4902578
X-12443	Inverse variance weighted	12	0.032292450	1.81(1.05 to 3.10)	1.8064532
	MR Egger	12	0.387084349	1.71(0.53 to 5.49)	1.7123766
	Weighted median	12	0.336725176	1.46(0.68 to 3.14)	1.4575584
octanoylcarnitine	Inverse variance weighted	15	0.023466553	3.40(1.18 to 9.79)	3.3989603
	MR Egger	15	0.273158298	3.05(0.45 to 20.63)	3.0514126
	Weighted median	15	0.268202064	2.06(0.57 to 7.41)	2.0609784
alpha-hydroxyisovalerate	Inverse variance weighted	15	0.036434844	0.32(0.11 to 0.93)	0.3240534
	MR Egger	15	0.641244244	2.08(0.10 to 42.07)	2.0791505
	Weighted median	15	0.587794304	0.66(0.14 to 2.99)	0.6577501
1,7-dimethylurate	Inverse variance weighted	11	0.033479711	0.40(0.17 to 0.93)	0.4012497
	MR Egger	11	0.102601161	0.24(0.05 to 1.12)	0.2375157
	Weighted median	11	0.080696629	0.36(0.11 to 1.13)	0.3569612
1-linoleoylglycerophosphocholine	Inverse variance weighted	15	0.043011471	6.56(1.06 to 40.56)	6.5594952
	MR Egger	15	0.139562572	58.69(0.37 to 9359.93)	58.6933841
	Weighted median	15	0.479506823	2.54(0.19 to 33.75)	2.5420146
X-12786	Inverse variance weighted	11	0.020439783	3.37(1.21 to 9.42)	3.3719032
	MR Egger	11	0.325352140	2.67(0.42 to 16.91)	2.6656631
	Weighted median	11	0.153375819	2.77(0.68 to 11.19)	2.7672317
X-13671	Inverse variance weighted	15	0.002225249	23.16(3.09 to 173.55)	23.1628982
	MR Egger	15	0.924161043	1.49(0.00 to 4619.10)	1.4890313
	Weighted median	15	0.163260705	7.70(0.44 to 135.77)	7.7003996
X-14374	Inverse variance weighted	8	0.043599864	0.10(0.01 to 0.94)	0.1025390
	MR Egger	8	0.839056886	0.33(0.00 to 8594.67)	0.3331344
	Weighted median	8	0.171921537	0.15(0.01 to 2.26)	0.1528362
X-14745	Inverse variance weighted	12	0.046612994	4.95(1.02 to 23.90)	4.9472083
	MR Egger	12	0.415854665	65.99(0.00 to 1049343.24)	65.9877372
	Weighted median	12	0.266212445	3.49(0.39 to 31.50)	3.4858081

P<0.05 was considered statistically significant

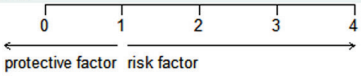

Figure 2. Forest plot of the MR-analyzed blood metabolites on idiopathic pulmonary causality. MR, Mendelian randomization; CI, confidence interval; OR, odds ratio; SNPs, single nucleotide polymorphisms.

Table I. Sensitivity analysis.

Metabolites	Q_pval (IVW)	MRegger_intercept	MRegger_interpreter_pval
Cotinine	0.634204819	-0.016253799	0.594349888
Hypoxanthine	0.64539929	-0.014255373	0.658951597
Aspartylphenylalanine	0.507548038	-0.0100362	0.8937339
X-10395	0.421558055	0.005325463	0.697704685
X-10810	0.942302112	0.016006743	0.482089942
Acetylcarnitine	0.854090843	0.007536244	0.833125497
2-aminobutyrate	0.94823233	-0.004991164	0.84015343
Docosapentaenoate (n3 DPA; 22:5n3)	0.857218151	-0.020539902	0.729874938
X-11261	0.752776285	-0.029700565	0.386080027
X-11795	0.812512936	-0.032772067	0.537999752
X-12189	0.19702192	-0.02414237	0.332548791
X-12428	0.49662327	-0.009779736	0.805574022
X-12443	0.748525852	0.004634389	0.921132003
Octanoylcarnitine	0.198556909	0.003739538	0.894565868
Alpha-hydroxyisovalerate	0.480656564	-0.057310933	0.218243317
1,7-dimethylurate	0.86047744	0.026635303	0.450327514
1-linoleoylglycerophosphocholine	0.806629731	-0.040877387	0.380697155
X-12786	0.824668697	0.009070153	0.770926683
X-13671	0.667042516	0.037205512	0.501665959
X-14374	0.124767596	-0.026570989	0.822873085
X-14745	0.384835525	-0.057805481	0.606030326

MR, Mendelian randomization; IVW, inverse variance weighted.

as acetyl-carnitine and octanoyl-carnitine, serve significant roles in several cellular energy metabolic pathways and are commonly utilized as nutritional supplements. The identification of acyl-carnitines has become increasingly crucial for the study of metabolism in various diseases, such as metabolic disorders, heart disease, diabetes, depression, neurological disorders and particular types of cancer (44). Therefore, it was hypothesized that its effect on IPF could be associated to its involvement in histone acetylation. However, further studies are needed to verify this finding. 1-linoleoylglycerophosphocholine is a phospholipid, which is formed from linoleic acid and glycerol. It is involved in several processes in the human body, such as building and upkeeping cell membranes and acting as a signaling molecule. It is also used as an ingredient in nutritional supplements and medications. A previous study suggested that 1-oleoylglycerophosphate choline could notably reduce the risk of type 2 diabetes (45). In the present study, elevated 1-lipoloylglycophorophosphacholine levels were associated with increased susceptibility to IPF. During the pathological process of IPF, the cell membrane structure can be damaged, while the cellular function can be disrupted. When the levels of 1-lipoloylglycophorophosphacholine are abnormally enhanced, it can disrupt the lipid balance of the cell membrane, and affect membrane protein function and the exchange and signal transmission of substances inside and outside the cell. For example, it can interfere with the binding of cell surface receptors and ligands, thus affecting the response of cells to signaling molecules, such as growth factors and cytokines, thereby promoting abnormal activation and proliferation of fibroblasts, excessive deposition of extracellular matrix, and advancing the process of pulmonary fibrosis (46,47).

In the present study, the following metabolites with protective effects on IPF were identified: Hypoxanthine, 2-aminobutyrate, docosapentaenoate, alpha-hydroxy-isovalerate and 1,7-dimethylurate. Therefore, the particular mechanism of their protective effects on IPF requires further in-depth study. Hypoxanthine is a purine base, which plays a significant role in living organisms. It is a component of DNA and RNA, and is also involved in the process of nucleotide synthesis. In addition, hypoxanthine can be converted to uric acid via metabolic pathways, thus affecting purine metabolism and acting as an indicator of hypoxia. Purine plays a key role in regulating immune response during wound healing and fibrosis and is a key signal hub for scar tissue formation in the fibrotic region. The current study strengthened the study on the improvement of IPF by hypoxanthine, thus providing novel insights into the identification of potential targets for IPF. 2-aminobutyrate, an amino acid, plays multiple functions within the human body. Protein synthesis, metabolism, as well as neurotransmitter synthesis are its areas of involvement. It has been reported that 2-aminobutyrate can also have antioxidant and anti-inflammatory effects, which could account for its protective effects against IPF. Docosapentaenoic acid, a type of omega-3 fatty acid, is specifically classified as a long-chain polyunsaturated fatty acid. It is also known as DPA and is particularly found in fish and marine oils. Docosapentaenoic acid was found to exert similar health benefits with other

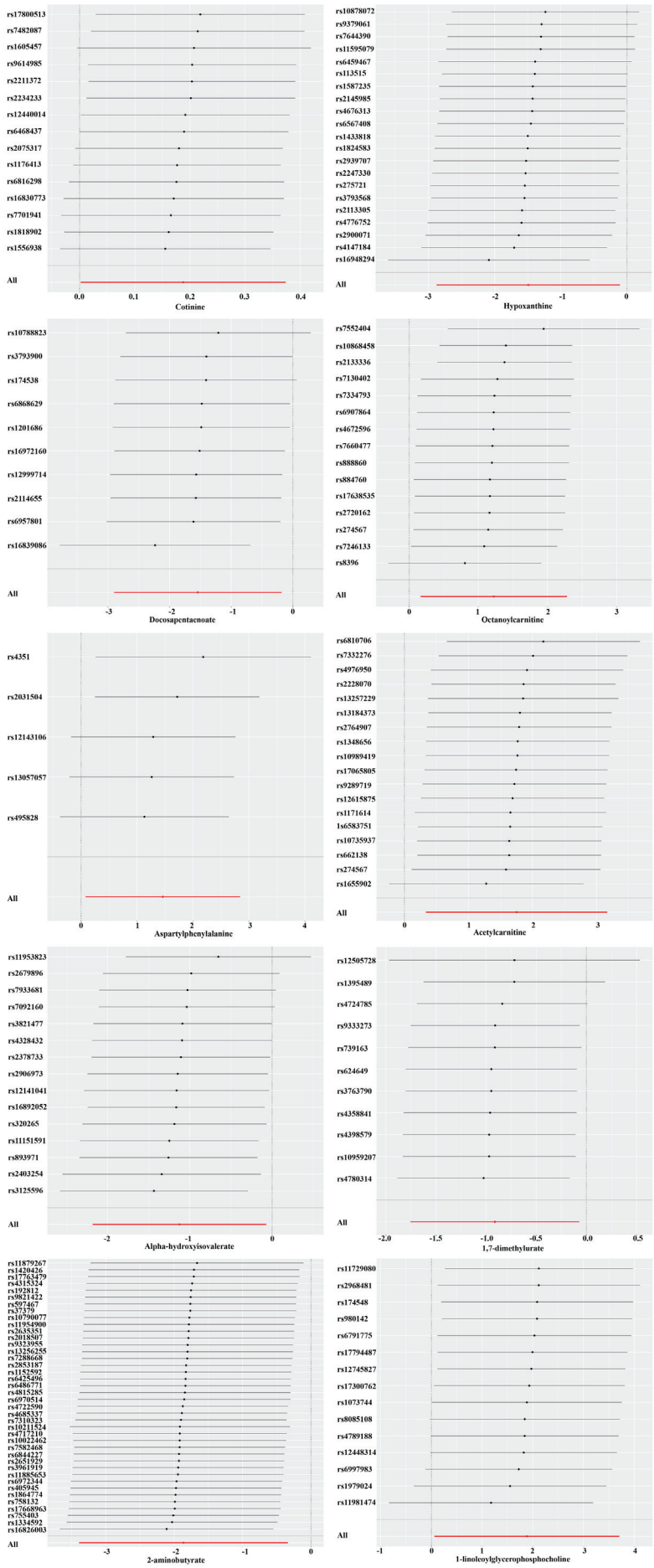

Figure 3. Leave-one-out results are shown.

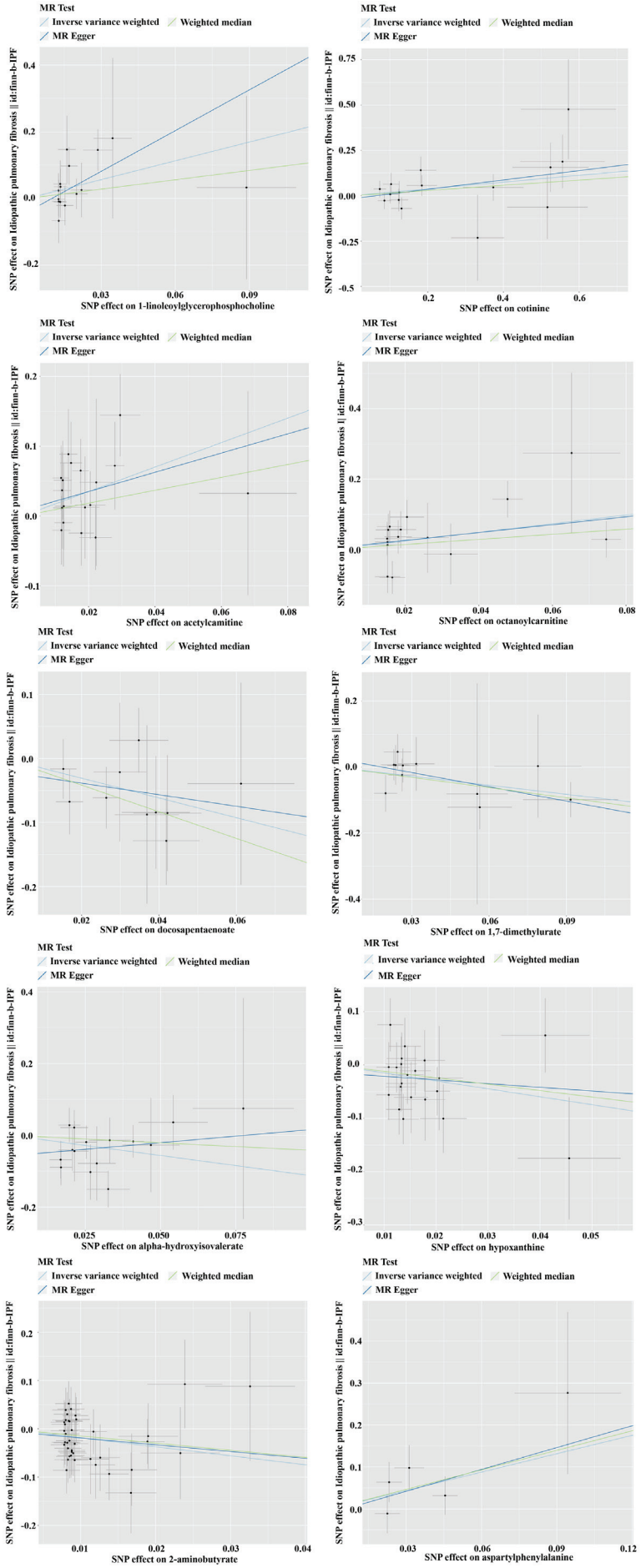

Figure 4. Scatter plots of the Mendelian randomization analysis are presented.

omega-3 fatty acids, such as reducing inflammation, supporting heart health and potentially improving cognitive function (48). During the degradation of leucine, alpha-hydroxy-isovalerate is produced as a metabolite. Alpha-hydroxy-isovalerate is an organic compound involved in several metabolic pathways in the body. Researchers are currently examining its possible involvement in the identification and treatment of different ailments, such as maple syrup urine disease and other conditions associated with the metabolism of branched-chain amino acids (49). 1,7-Dimethylurate is a metabolite of caffeine, also known as dimethylxanthine. It is produced through the metabolism of caffeine in the body and has antioxidant and anti-inflammatory properties (50). For the unknown serum metabolites found to be associated with IPF risk, further attention should be paid in future studies to explore their possible mechanisms underlying their effect on IPF, thus providing assistance and insights for the clinical prevention and treatment of IPF.

However, the present study has certain limitations. Firstly, a limited number of SNPs, that could be used for genome-wide level exposure, were detected. In order to tackle this problem, slightly more lenient criteria were established for MR analysis, a widely adopted approach. Nevertheless, the F-test statistics for all chosen SNPs surpassed 10, thus suggesting that, in the present study, the instrumental variable was adequately strong. Furthermore, the research sample was presently restricted to individuals of European descent. Therefore, the applicability to other populations requires further investigation and confirmation. Further exploration is required to uncover the precise mechanisms underlying the effects of particular metabolites on the development of IPF. Due to limited data sources, in the current study, only 486 metabolites were covered. Therefore, more studies on additional blood metabolites associated with IPF should be performed. A purely bioinformatic analysis of existing data from one source indeed presents several limitations. Firstly, the data may not be representative of the entire population. For example, if the source database has a bias in terms of geographical location or ethnicity, it could lead to inaccurate or incomplete conclusions about the relationship between metabolites and the studied condition. Secondly, the lack of experimental verification means that the bioinformatic predictions may not accurately reflect the *in vivo* or *in vitro* biological processes. There could be confounding factors or interactions that are not accounted for in the data. Thirdly, the static nature of the existing data may not capture the dynamic changes in metabolite levels over time or in response to different stimuli. To address these limitations, future research could involve collecting data from multiple diverse sources to improve representativeness. Experimental studies such as cell culture and animal models could be conducted to validate the bioinformatic findings and explore the underlying mechanisms. Longitudinal studies could also be designed to monitor metabolite fluctuations and their associations with disease progression.

In conclusion, the present magnetic resonance study indicated that numerous blood metabolites could play a key role in IPF, thus offering valuable perspectives on possible approaches for the early detection, avoidance and management of IPF. By combining genomics and metabolomics, this analysis could provide a point of reference for investigating the cause and development of IPF.

Acknowledgements

Not applicable.

Funding

The present study was supported by the Shandong Taishan Scholar Project (grant no. tsqn202306392).

Availability of data and materials

The data generated in the present study are included in the figures and/or tables of this article.

Authors' contributions

FW and XL designed the study and drafted the manuscript. FW, JL, BL, WY, XZ and XL performed the data collection and analysis. All authors read and approved the final version of the manuscript. FW and XL confirm the authenticity of all the raw data.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Podolanczuk AJ, Thomson CC, Remy-Jardin M, Richeldi L, Martinez FJ, Kolb M and Raghu G: Idiopathic pulmonary fibrosis: State of the art for 2023. Eur Respir J 61: 2200957, 2023.
2. Moss BJ, Ryter SW and Rosas IO: Pathogenic mechanisms underlying idiopathic pulmonary fibrosis. Annu Rev Pathol 17: 515-546, 2022.
3. Thiam F, Phogat S, Abokor FA and Osei ET: In vitro co-culture studies and the crucial role of fibroblast-immune cell crosstalk in IPF pathogenesis. Respir Res 24: 298, 2023.
4. Platenburg MGJP, van Moorsel CHM, Wiertz IA, van der Vis JJ, Vorselaars ADM, Veltkamp M and Grutters JC: Improved survival of IPF patients treated with antifibrotic drugs compared with untreated patients. Lung 201: 335-343, 2023.
5. Tomassetti S, Ravaglia C, Piciucchi S, Ryu J, Wells A, Donati L, Dubini A, Klersy C, Luzzi V, Gori L, *et al*: Historical eye on IPF: A cohort study redefining the mortality scenario. Front Med (Lausanne) 10: 1151922, 2023.
6. Ogura T, Inoue Y, Azuma A, Homma S, Kondoh Y, Tanaka K, Ochiai K, Sugiyama Y and Nukiwa T: Real-world safety and tolerability of nintedanib in patients with idiopathic pulmonary fibrosis: Interim report of a post-marketing surveillance in Japan. Adv Ther 40: 1474-1493, 2023.
7. Spagnolo P and Lee JS: Recent advances in the genetics of idiopathic pulmonary fibrosis. Curr Opin Pulm Med 29: 399-405, 2023.
8. Alonso-Gonzalez A, Tosco-Herrera E, Molina-Molina M and Flores C: Idiopathic pulmonary fibrosis and the role of genetics in the era of precision medicine. Front Med (Lausanne) 10: 1152211, 2023.
9. Rajesh R, Atallah R and Bärnthaler T: Dysregulation of metabolic pathways in pulmonary fibrosis. Pharmacol Ther 246: 108436, 2023.

10. Roque W and Romero F: Cellular metabolomics of pulmonary fibrosis, from amino acids to lipids. Am J Physiol Cell Physiol 320: C689-C695, 2021.
11. Johnson CH, Ivanisevic J and Siuzdak G: Metabolomics: Beyond biomarkers and towards mechanisms. Nat Rev Mol Cell Biol 17: 451-459, 2016.
12. Selvarajah B, Azuelos I, Anastasiou D and Chambers RC: Fibrometabolism-An emerging therapeutic frontier in pulmonary fibrosis. Sci Signal 14: eaay1027, 2021.
13. Zhao YD, Yin L, Archer S, Lu C, Zhao G, Yao Y, Wu L, Hsin M, Waddell TK, Keshavjee S, *et al*: Metabolic heterogeneity of idiopathic pulmonary fibrosis: A metabolomic study. BMJ Open Respir. Res 4: e000183, 2017.
14. Gaugg MT, Engler A, Bregy L, Nussbaumer-Ochsner Y, Eiffert L, Bruderer T, Zenobi R, Sinues P and Kohler M: Molecular breath analysis supports altered amino acid metabolism in idiopathic pulmonary fibrosis. Respirology 24: 437-444, 2019.
15. Mamazhakypov A, Schermuly RT, Schaefer L and Wygrecka M: Lipids - two sides of the same coin in lung fibrosis. Cell Signal 60: 65-80, 2019.
16. Tedesco S, Scattolini V, Albiero M, Bortolozzi M, Avogaro A, Cignarella A and Fadini GP: Mitochondrial calcium uptake is instrumental to alternative macrophage polarization and phagocytic activity. Int J Mol Sci 20: 4966, 2019.
17. Kim HS, Moon SJ, Lee SE, Hwang GW, Yoo HJ and Song JW: The arachidonic acid metabolite 11,12-epoxyeicosatrienoic acid alleviates pulmonary fibrosis. Exp Mol Med 53: 864-874, 2021.
18. Wang Q, Xie Z, Wan N, Yang L, Jin Z, Jin F, Huang Z, Chen M, Wang H and Feng J: Potential biomarkers for diagnosis and disease evaluation of idiopathic pulmonary fibrosis. Chin Med J (Engl) 136: 1278-1290, 2023.
19. Kanehisa M, Goto S, Sato Y, Furumichi M and Tanabe M: KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res 40(Database issue): D109-D114, 2012.
20. Dhindsa RS, Mattsson J, Nag A, Wang Q, Wain LV, Allen R, Wigmore EM, Ibanez K, Vitsios D, Deevi SVV, *et al*: Identification of a missense variant in SPDL1 associated with idiopathic pulmonary fibrosis. Commun Biol 4: 392, 2021.
21. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, VanderWeele TJ, Higgins JPT, Timpson NJ, Dimou N, *et al*: Strengthening the reporting of observational studies in epidemiology using mendelian randomization: The STROBE-MR statement. JAMA 326: 1614-1621, 2021.
22. Slob EAW and Burgess S: A comparison of robust Mendelian randomization methods using summary data. Genet Epidemiol 44: 313-329, 2020.
23. Yang X, Zhu Q, Zhang L, Pei Y, Xu X, Liu X, Lu G, Pan J and Wang Y: Causal relationship between gut microbiota and serum vitamin D: Evidence from genetic correlation and Mendelian randomization study. Eur J Clin Nutr 76: 1017-1023, 2022.
24. Roze D: Causes and consequences of linkage disequilibrium among transposable elements within eukaryotic genomes. Genetics 224: iyad058, 2023.
25. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, Butterworth AS and Staley JR: PhenoScanner V2: An expanded tool for searching human genotype-phenotype associations. Bioinformatics 35: 4851-4853, 2019.
26. Sekula P, Del Greco MF, Pattaro C and Köttgen A: Mendelian randomization as an approach to assess causality using observational data. J Am Soc Nephrol 27: 3253-3265, 2016.
27. Zhao L, Wu R, Wu Z, Liu X, Li J, Zhang L and Zhang S: Genetically predicted 486 blood metabolites concerning risk of systemic lupus erythematosus: A Mendelian randomization study. Sci Rep 13: 22543, 2023.
28. Burgess S and Thompson SG: Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol 32: 377-389, 2017.
29. Pierce BL, Ahsan H and VanderWeele TJ: Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. Int J Epidemiol 40: 740-752, 2011.
30. Hemani G, Bowden J and Davey Smith G: Evaluating the potential role of pleiotropy in mendelian randomization studies. Hum Mol Genet 27: R195-R208, 2018.
31. Burgess S, Bowden J, Fall T, Ingelsson E and Thompson SG: Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. Epidemiology 28: 30-42, 2017.
32. Rappaport SM, Barupal DK, Wishart D, Vineis P and Scalbert A: The blood exposome and its role in discovering causes of disease. Environ Health Perspect 122: 769-774, 2014.
33. Li Y, He Y, Chen S, Wang Q, Yang Y, Shen D, Ma J, Wen Z, Ning S and Chen H: S100A12 as biomarker of disease severity and prognosis in patients with idiopathic pulmonary fibrosis. Front Immunol 13: 810338, 2022.
34. Fang L, Chen H, Kong R and Que J: Endogenous tryptophan metabolite 5-Methoxytryptophan inhibits pulmonary fibrosis by downregulating the TGF-β/SMAD3 and PI3K/AKT signaling pathway. Life Sci 260: 118399, 2020.
35. Bai L, Bernard K, Tang X, Hu M, Horowitz JC, Thannickal VJ and Sanders YY: Glutaminolysis epigenetically regulates antiapoptotic gene expression in idiopathic pulmonary fibrosis fibroblasts. Am J Respir Cell Mol Biol 60: 49-57, 2019.
36. Senoo S, Higo H, Taniguchi A, Kiura K, Maeda Y and Miyahara N: Pulmonary fibrosis and type-17 immunity. Respir Investig 61: 553-562, 2023.
37. Lei T, Li M, Zhu Z, Yang J, Hu Y and Hua L: Comprehensive evaluation of serum cotinine on human health: Novel evidence for the systemic toxicity of tobacco smoke in the US general population. Sci Total Environ 892: 164443, 2023.
38. Álvarez-Hernández J, Matía-Martín P, Cáncer-Minchot E and Cuerda C; NUTRICOVID study group of SENDIMAD: Long-term outcomes in critically ill patients who survived COVID-19: The NUTRICOVID observational cohort study. Clin Nutr 42: 2029-2035, 2023.
39. Chen Z, Li H, Song C, Sun J and Liu W: Association between serum cotinine and muscle mass: Results from NHANES 2011-2018. BMC Public Health 24: 2093, 2024.
40. She D, Jiang S and Yuan S: Association between serum cotinine and hepatic steatosis and liver fibrosis in adolescent: A population-based study in the United States. Sci Rep 14: 11424, 2024.
41. Qiu G, Lin Y, Ouyang Y, You M, Zhao X, Wang H, Niu R, Li W, Xu X, Yan Q, *et al*: Nontargeted metabolomics revealed novel association between serum metabolites and incident acute coronary syndrome: A mendelian randomization study. J Am Heart Assoc 12: e028540, 2023.
42. Stegink LD, Lindgren SD, Brummel MC, Stumbo PJ and Wolraich ML: Erythrocyte L-aspartyl-L-phenylalanine hydrolase activity and plasma phenylalanine and aspartate concentrations in children consuming diets high in aspartame. Am J Clin Nutr 62: 1206-1211, 1995.
43. Meng J, Liu W, Wu Y, Xiao Y, Tang H and Gao S: Is it necessary to wear compression stockings and how long should they be worn for preventing post thrombotic syndrome? A meta-analysis of randomized controlled trials. Thromb Res 225: 79-86, 2023.
44. Dambrova M, Makrecka-Kuka M, Kuka J, Vilskersts R, Nordberg D, Attwood MM, Smesny S, Sen ZD, Guo AC, Oler E, *et al*: Acylcarnitines: Nomenclature, biomarkers, therapeutic potential, drug targets, and clinical trials. Pharmacol Rev 74: 506-551, 2022.
45. Vangipurapu J, Fernandes Silva L, Kuulasmaa T, Smith U and Laakso M: Microbiota-related metabolites and the risk of type 2 diabetes. Diabetes Care 43: 1319-1325, 2020.
46. Zou X, Wang L, Wang S, Zhang Y, Ma J, Chen L, Li Y, Yao TX, Zhou H, Wu L, *et al*: Promising therapeutic targets for ischemic stroke identified from plasma and cerebrospinal fluid proteomes: A multicenter Mendelian randomization study. Int J Surg 110: 766-776, 2024.
47. Meng J, Tang H, Xiao Y, Liu W, Wu Y, Xiong Y and Gao S: Appropriate thromboprophylaxis strategy for COVID-19 patients on dosage, antiplatelet therapy, outpatient, and postdischarge prophylaxis: A meta-analysis of randomized controlled trials. Int J Surg 110: 3910-3922, 2024.
48. Lu Y, Li D, Wang L, Zhang H, Jiang F, Zhang R, Xu L, Yang N, Dai S, Xu X, *et al*: Comprehensive investigation on associations between dietary intake and blood levels of fatty acids and colorectal cancer risk. Nutrients 15: 730, 2023.
49. Kaplan RC, Williams-Nguyen JS, Huang Y, Mossavar-Rahmani Y, Yu B, Boerwinkle E, Gellman MD, Daviglus M, Chilcoat A, Van Horn L, *et al*: Identification of dietary supplements associated with blood metabolites in the hispanic community health study/study of latinos cohort study. J Nutr 153: 1483-1492, 2023.
50. Purpura M, Jäger R and Falk M: An assessment of mutagenicity, genotoxicity, acute-, subacute and subchronic oral toxicity of paraxanthine (1,7-dimethylxanthine). Food Chem Toxicol 158: 112579, 2021.