



OPEN Investigating the potential of oxidative stress-related gene as predictive markers in idiopathic pulmonary fibrosis

Yuhao Wang^{1,2,8}, Zhao Zhang^{3,8}, Hongnan Zhang⁴, Luyao Xu⁵, Shuling Huang⁶, Ying Wen⁷, Zhiming Zhuang³, Xiaoxin Li³, Jinyi Lu³ & Xudong Li¹✉

This study investigates genes linking oxidative stress to idiopathic pulmonary fibrosis (IPF) through multi-omics data integration. We collected oxidative stress-related genes from GeneCards and integrated data for gene expression (eQTLs), DNA methylation (mQTLs), and protein expression (pQTLs). Genome-wide association study (GWAS) data on IPF from Allen et al. served as the discovery set, with FinnGen R10 for validation. Summary data-based Mendelian randomization (SMR) and colocalization analyses assessed interactions and shared causal variants, followed by multi-omics integration with tissue-specific validation. SMR and colocalization screening identified 90 mQTLs, 15 eQTLs, and 2 pQTLs (*KRT18* and *FOXO1*) linked to IPF in the discovery cohort. Twelve mQTLs were validated in the FinnGen cohort, with *MUC1* showing strong SMR and colocalization evidence (eQTL). Multi-omics integration validated *NDUFA9* (mQTL-eQTL) level and *FOXO1* (mQTL-eQTL-pQTL). Our study identified key oxidative stress-related genes (i.e., *FOXO1* and *NDUFA9*) in IPF pathogenesis, highlighting the need for further research to inform prevention and treatment.

Keywords Oxidative stress, Idiopathic pulmonary fibrosis, Multi-omics, Mendelian randomization analysis, Colocalization analysis

Idiopathic pulmonary fibrosis (IPF) is a progressive lung disease that presents significant clinical challenges, with approximately 50,000 to 100,000 new cases diagnosed annually worldwide. The disease primarily affects older adults, and its prognosis is poor, with 3- and 5-year mortality rates estimated at 50% and 80%, respectively¹. Patients typically exhibit gradually worsening dyspnea and dry cough, which severely impact their quality of life². Current treatment strategies focus on slowing disease progression and improving quality of life, such as using antifibrotic agents like pirfenidone and nintedanib to reduce the rate of lung function decline³. Despite these advancements, the underlying mechanisms of the disease remain poorly understood, which has limited the development of targeted drugs and treatment strategies.

The progression of IPF is a complex process influenced by multiple factors, including chronic inflammation, epithelial cell injury, fibroblast activation, and extracellular matrix remodeling. Recent studies have emphasized the importance of understanding the mechanisms by which oxidative stress promotes IPF progression, with OS playing a central role in driving these processes⁴. Oxidative stress is characterized by an imbalance between the production of reactive oxygen species (ROS) and the body's antioxidant defenses, which has been shown to damage lung epithelial cells, further activate fibroblasts, and exacerbate extracellular matrix deposition, thereby promoting fibrosis⁵. Moreover, the accumulation of ROS can lead to mitochondrial dysfunction, endoplasmic reticulum stress, induction of apoptosis, cellular senescence, and activation of inflammatory responses, all of which exacerbate the pathological processes of IPF^{6,7}. These abnormal pathways can result in further

¹Guangdong Research Center of Occupational Hygiene, Guangdong Province Hospital for Occupational Disease Prevention and Treatment, Guangzhou 510399, Guangdong, China. ²Guangzhou Center for Disease Control and Prevention (Guangzhou Health Supervision institute), Guangzhou 510440, Guangdong, China. ³School of Public Health, Southern Medical University, Guangzhou 510515, Guangdong, China. ⁴Longgui Street Community Health Service Center, Baiyun District, Guangzhou 510445, Guangdong, China. ⁵School of Public Health, Sun Yat-sen University, Guangzhou 510275, Guangdong, China. ⁶School of Public Health, Guangzhou Medical University, Guangzhou 511436, Guangdong, China. ⁷Guangzhou Women and Children's Medical Centre, Institute of Pediatrics, Guangzhou Medical University, Guangzhou 510623, Guangdong, China. ⁸Yuhao Wang and Zhao Zhang contributed equally to this work. ✉email: lixd116@126.com

accumulation of oxidative stress, creating a feedback loop that contributes to a vicious cycle of oxidative stress in the pathogenesis of IPF, thereby accelerating the irreversibility of the disease⁸. Given the critical role of oxidative stress in IPF progression, it has become a potential therapeutic target. Inhibiting ROS production or enhancing antioxidant defenses may slow or halt the progression of IPF⁵. However, despite the recognition of oxidative stress as a key factor in IPF progression, the specific roles and mechanisms by which oxidative stress-related genes impact IPF remain unclear.

Summary data-based Mendelian randomization (SMR) is a powerful tool for exploring the genetic basis of complex traits, including oxidative stress in IPF⁹. By leveraging genome-wide association study (GWAS) summary data, SMR enhances statistical power and uncovers subtle genetic associations that are critical for advancing IPF research¹⁰. Recent studies utilizing SMR have begun to explore the role of oxidative stress-related genes in IPF by analyzing GWAS in conjunction with expression quantitative trait loci (eQTLs), DNA methylation QTLs (mQTLs), and protein QTLs (pQTLs) to identify potential causal genes involved in IPF pathogenesis^{11,12}. However, despite laying an important foundation, these studies, due to insufficient sample sizes and a primary focus on associations, still leave significant gaps in understanding the specific genetic changes related to oxidative stress in IPF^{11,12}. Therefore, to address these shortcomings and ensure the robustness and clinical relevance of the findings, there is an urgent need for comprehensive multi-omics integration and extensive validation across multiple cohorts.

In this study, we constructed a genetic and molecular interaction landscape of IPF by integrating multi-level data related to oxidative stress, including gene expression (eQTLs), DNA methylation (mQTLs), and protein expression (pQTLs). This approach provides an in-depth analysis of the complex relationships between genetics and disease susceptibility¹⁴. By integrating diverse QTL data, we sought to preliminarily explore the causal genetic determinants and molecular pathways that mediate the oxidative stress-driven pathogenesis of IPF, thereby deepening our understanding of IPF genetics and providing a theoretical foundation for identifying new therapeutic targets.

Materials and methods

Study design

In this study, we selected genes related to oxidative stress as instrumental variables across three biological layers: DNA methylation, gene expression, and protein abundance. Independent MR analyses were conducted at each biological level to investigate their associations with IPF. The GWAS data on IPF from Allen et al.¹⁵ was used as the primary discovery set, while the FinnGen R10 cohort dataset was employed for validation. Notably, there were no sample overlaps between the exposure and outcome groups. To strengthen causal inference, we also performed colocalization analyses^{16–18}. By integrating the results from these three MR analyses, we identified candidate causal genes, which were further subjected to tissue-specific validation. Figure 1 summarizes the study design and the workflow for the selection of genetic variants and analysis methods.

Data sources

Oxidative stress-related genes were extracted from the GeneCards database. By restricting the Category to “Protein Coding” and filtering based on relevance scores, we retained 991 protein-coding genes within the top 10% of relevance scores for subsequent analysis. The summary-level data for blood mQTLs were derived from a meta-analysis of two cohort studies: the Brisbane Systems Genetics Study ($n=614$) and the Lothian Birth Cohort ($n=1366$) (https://yanglab.westlake.edu.cn/data/SMR/LBC_BSGS_meta.tar.gz)¹⁹. Blood eQTL data were obtained from the eQTLGen consortium, which encompasses blood gene expression data from 31,684 individuals (<https://molgenis26.gcc.rug.nl/downloads/eqtgen/cis-eqt/2019-12-11-cis-eQTLsFDR-ProbeLevel-CohortInfoRemoved-BonferroniAdded.txt.gz>)²⁰. The blood pQTL summary data were sourced from the UK Biobank Proteomics-Genomics Project (UKB-PPP) by Benjamin et al., which investigates the proteome-genome intersection in human diseases and includes data from 54,219 donors (<https://www.synapse.org/Synapse:syn51365303>)²¹.

The summary statistics for the IPF GWAS were obtained from the dataset published by Allen et al.¹⁵ (3-way meta-GWAS of IPF susceptibility), which includes 2668 cases and 8591 controls (<https://github.com/genomicslTER/Pfgenetics/tree/master>). For the validation phase, we used the FinnGen R10 cohort dataset, which consists of 2189 cases and 407,609 controls (https://storage.googleapis.com/finngen-public-data-r10/summary_stats/finngen_R10_IPF.gz).

All summary statistics used for MR analysis were derived from previously published studies (Supplementary Table S1), all of which received ethical approval.

Summary-based Mendelian randomization analysis

We used SMR analysis to estimate the relationships between the methylation, expression, and protein abundance of oxidative stress-related genes and IPF. SMR offers greater statistical power than traditional MR analyses based on the most significant cis-QTLs, particularly when the exposure and outcome are derived from two independent cohorts with large sample sizes²². In this study, we selected the most significant cis-QTLs within a window centered on the corresponding gene (± 1000 kb) and a p-value threshold of 5.0×10^{-823} . Single nucleotide polymorphisms (SNPs) with an allele frequency difference exceeding a specified threshold (0.2 in this study) between any pairwise datasets, including the LD reference sample, QTL summary data, and outcome summary data, were excluded. For eQTLs, mQTLs, and pQTLs, the allowable allele frequency difference was set to 0.05 by default.

Building on the SMR analysis, we employed a multi-SNP SMR approach that integrates mQTL, eQTL, and pQTL data to investigate the causal relationships between DNA methylation and gene expression, as well as between gene expression and protein abundance²². In SMR analysis, mQTLs were used as exposures and eQTLs

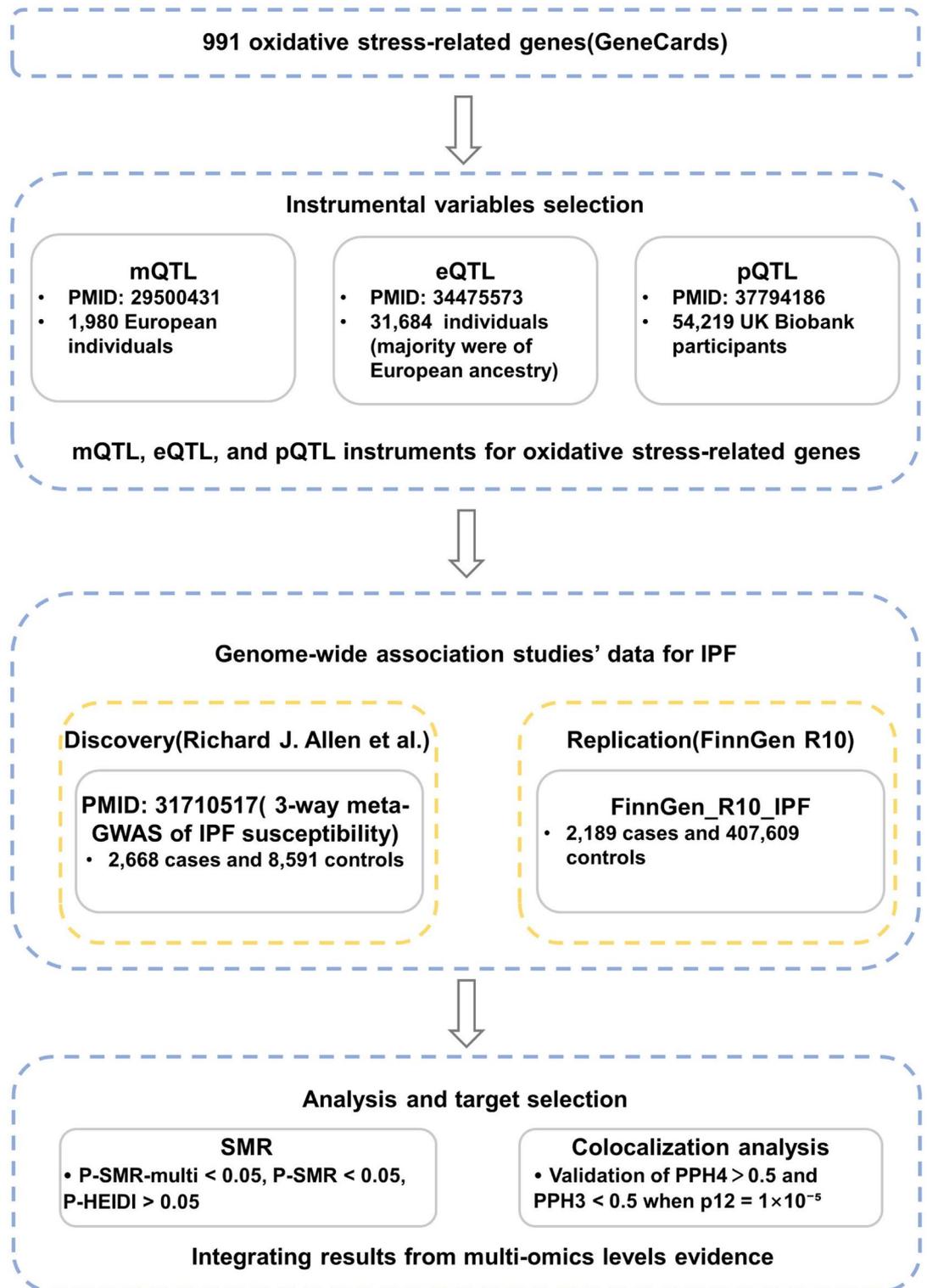


Fig. 1. Flowchart of the analyses performed.

as outcomes, or eQTLs were used as exposures and pQTLs as outcomes. These analyses aimed to determine whether the expression of target genes is regulated by the methylation of specific CpG sites within their functional regions or to verify whether the expression of target genes regulates the abundance of their encoded proteins. This study focuses on the results obtained using this method.

This approach considered all SNPs within the QTL probe window region (default of 500 kb) with p-values below the default threshold of 5×10^{-8} and an LD r^2 value below the default threshold of 0.9 for the SNP most strongly associated with the SMR analysis²⁴. In this study, we thoroughly evaluated the significance of the results

obtained using this method. Subsequently, results without pleiotropy were screened using the HEIDI test, with a p-value threshold of 0.05²⁵. Therefore, under the condition that P-SMR is less than 0.05, results satisfying P-SMR-multi < 0.05 and P-HEIDI > 0.05 were used for subsequent eQTL, mQTL, and pQTL colocalization and integration analyses. SMR and HEIDI tests were implemented using the SMR software tool (version 1.3.1).

Colocalization analysis

We performed colocalization analysis using the “coloc” R package to detect shared causal variants between identified oxidative stress-related mQTLs, eQTLs, or pQTLs and IPF. Specifically, we hypothesized that GWAS loci might influence the phenotype by altering gene-related biological processes when colocalization between GWAS signals and QTLs is observed. The colocalization analysis reports five different posterior probabilities corresponding to five independent hypotheses: (1) no genetic association with either trait (H0); (2) genetic association with gene expression only (H1); (3) genetic association with disease risk only (H2); (4) both traits are associated with SNPs, but with different causal variants (H3); (5) both traits share the same causal variant (H4).

According to the published literature, the colocalization region windows for mQTL-GWAS, eQTL-GWAS, and pQTL-GWAS colocalization analysis were set at ± 500 kb, ± 1000 kb, and ± 1000 kb, respectively^{26–28}. The default prior probability of an SNP being associated with both exposure and outcome was set at $p_{12} = 1 \times 10^{-5}$. Although a PPH4 (posterior probability of colocalization) > 0.8 has been shown to indicate strong Bayesian evidence for colocalization, Breen et al. observed that many loci with PPH4 > 0.5 qualitatively exhibit colocalization-like patterns^{29,30}. Therefore, to enhance the sensitivity of colocalization analysis and ensure the capture of more potential colocalization signals, particularly during the exploratory phase of the study, we considered QTL and GWAS signals to have strong evidence of colocalization if they met the conditions (1) PPH4 > 0.5 when $p_{12} = 5 \times 10^{-5}$ and (2) PPH3 < 0.5 when $p_{12} = 1 \times 10^{-5}$. This approach provides a more comprehensive preliminary indication of colocalization³¹.

All statistical analyses were performed using R (v4.3.0). The R package “ggplot2” was used for Manhattan plot generation, “ggrepel” for Manhattan plot annotations, and “forestplot” for forest plot generation. SMRLocusPlot and SMREffectPlot generation codes were sourced from Zhu et al.²³.

Results

Gene expression of oxidative stress-related genes in IPF

Applying the same criteria, i.e., P-SMR-multi < 0.05, P-SMR < 0.05, and P-HEIDI > 0.05, we found that the expression of 21 genes was significantly associated with IPF (Supplementary Table S2). However, it is noteworthy that while these 21 genes showed significant associations in our primary analysis, their expression levels did not consistently replicate in the FinnGen cohort (Supplementary Table S3). Additionally, further analysis in the discovery cohort revealed that 15 of these genes showed strong evidence of colocalization (PPH4 > 0.5, PPH3 < 0.5) (Fig. 2).

Methylation of oxidative stress-related genes in IPF

Using the criteria of P-SMR-multi < 0.05, P-SMR < 0.05, and P-HEIDI > 0.05, we identified methylation sites for 156 oxidative stress-related genes corresponding to 93 genes in the discovery set (Supplementary Table S4). Among the results from the SMR analysis, 12 sites corresponding to 10 genes were validated in the FinnGen cohort (Supplementary Table S5), providing strong evidence of the stability and consistency of these sites across different populations.

In the discovery cohort, we also identified 90 sites corresponding to 53 genes that showed strong evidence of colocalization (PPH4 > 0.5, PPH3 < 0.5). Figure 3 illustrates key CpG sites whose associated gene expression is significantly related to IPF risk. Among these, 12 sites were validated in the FinnGen cohort. Notably, *MUC1* (cg15699386) (OR = 0.6, 95% CI = 0.43–0.83) was also validated in the gene expression analysis in the discovery set and showed strong evidence of colocalization (Fig. 3). The validation results of these genes further support their important role in IPF and suggest that they may influence lung health through oxidative stress pathways.

Protein abundance of oxidative stress-related genes in IPF

At the protein abundance level, using the same criteria (P-SMR-multi < 0.05, P-SMR < 0.05, and P-HEIDI > 0.05), we found that the protein abundance of four oxidative stress-related genes was associated with IPF risk (Supplementary Table S6). Among these, *KRT18* (OR = 10.83, 95% CI = 2.1–55.94) and *FOXO1* (OR = 2.15, 95% CI = 1.02–4.53) were supported by strong colocalization evidence in the colocalization region window (PPH4 > 0.5, PPH3 < 0.5) (Fig. 4). This suggests that the expression levels of these gene-related proteins may play key roles in the pathological process of IPF, particularly in mechanisms related to oxidative stress. However, despite these significant findings in the discovery cohort, these findings were not validated in the FinnGen cohort (Supplementary Table S7).

Integration of blood mQTL and eQTL level data

We subsequently integrated key results from the SMR analysis to further explore the regulation of key oxidative stress genes in IPF by methylation in blood. Figure 5 shows the distribution of key loci, genes, and their encoded proteins on chromosomes. Based on single-level SMR analyses of IPF and oxidative stress-related mQTLs and eQTLs, we identified that the genes *MUC1*, *MAP3K7*, *HSF1*, *NDUFA9*, *HMGB1*, *FOXO1*, *GPX2*, and *SMAD2* may have causal associations with idiopathic pulmonary fibrosis. Furthermore, SMR analysis was conducted with blood mQTLs as the exposure and eQTLs as the outcome to explore whether methylation of CpG sites in these cross-results significantly regulates the expression of the associated genes. We found that *NDUFA9* (cg18779092, cg03680150) and *FOXO1* (cg23413567, cg11244402) were significantly validated in the above results (P-SMR-multi < 0.05, P-SMR < 0.05, and P-HEIDI > 0.05) (Table 1, Supplementary Table S8).

probeID	Symbol		OR (95%CI)	Pvalue	Pvalue (Multi)	PPH4
ENSG00000105327	BBC3		4.66 (1.21 - 17.89)	0.03	0.03	0.46
ENSG00000078401	EDN1		2.78 (1.58 - 4.91)	4.05E-04	0.03	0.97 *
ENSG00000185122	HSF1		2.62 (1.21 - 5.7)	0.02	0.02	0.63 *
ENSG00000139180	NDUFA9		1.89 (1.22 - 2.92)	4.42E-03	0.01	0.77 *
ENSG00000169032	MAP2K1		1.88 (1.27 - 2.77)	1.51E-03	6.82E-03	0.9 *
ENSG00000185499	MUC1		1.8 (1.16 - 2.78)	8.53E-03	1.78E-03	0.71 *
ENSG00000140945	CDH13		1.67 (1.03 - 2.7)	0.04	6.76E-03	0.24
ENSG00000115419	GLS		1.52 (1.13 - 2.04)	5.97E-03	0.04	0.53 *
ENSG00000132170	PPARG		1.44 (1.1 - 1.87)	7.31E-03	3.58E-03	0.63 *
ENSG00000165672	PRDX3		1.34 (1.01 - 1.8)	0.05	5.42E-03	0.21
ENSG00000131389	SLC6A6		1.27 (1.04 - 1.55)	0.02	0.04	0.03
ENSG00000157017	GHRL		0.61 (0.43 - 0.86)	4.91E-03	4.00E-03	0.7 *
ENSG00000135341	MAP3K7		0.55 (0.36 - 0.86)	7.84E-03	5.91E-03	0.57 *
ENSG00000006638	TBXA2R		0.5 (0.26 - 0.97)	0.04	0.04	0.27
ENSG00000175387	SMAD2		0.45 (0.21 - 0.98)	0.05	0.05	0.57 *
ENSG00000176153	GPX2		0.39 (0.21 - 0.75)	4.73E-03	1.99E-03	0.82 *
ENSG00000156273	BACH1		0.32 (0.13 - 0.8)	0.02	0.01	0.52 *
ENSG00000189403	HMGB1		0.31 (0.11 - 0.85)	0.02	0.02	0.53 *
ENSG00000101182	PSMA7		0.3 (0.12 - 0.77)	0.01	0.01	0.62 *
ENSG00000183765	CHEK2		0.29 (0.1 - 0.88)	0.03	0.03	0.44
ENSG00000150907	FOXO1		0.22 (0.06 - 0.85)	0.03	5.03E-03	0.78 *

Fig. 2. Associations of genetically predicted oxidative stress-related gene expression with idiopathic pulmonary fibrosis (IPF) in Mendelian randomization analysis. Odds ratios (ORs) with 95% confidence intervals (CIs) were derived from the SMR analysis. The horizontal bars denote the ORs (x-axis), where values greater than 1 indicate a positive association with IPF risk, and values less than 1 indicate a protective effect. P-value indicates the nominal significance level from SMR. P-value (Multi) refers to the p-value adjusted for multiple testing. PPH4 is the posterior probability that the same causal variant influences both gene expression and IPF risk; values above 0.5 are considered supportive of colocalization (marked with an asterisk “*”).

Integration of blood eQTL and pQTL level data

Based on our key findings from the comprehensive analysis of IPF and oxidative stress-related mQTL-eQTLs, we further explored the relationship between oxidative stress-related blood pQTLs and IPF GWAS. This series of analyses aimed to elucidate how genetic variants influence the development of IPF by affecting protein and gene expression levels. Notably, we observed significant validation of *FOXO1* protein in the pQTL-GWAS SMR analysis (Fig. 4). However, when investigating the regulatory relationship between the eQTL and pQTL signals of *FOXO1*, we found that the SMR P-value (P-SMR) was 0.054, which narrowly missed the conventional threshold for statistical significance (Table 1, Supplementary Table S9). Meanwhile, the P-value from the HEIDI test was 0.027, which did not meet the strict P-HEIDI > 0.05 threshold (Table 1, Supplementary Table S9). Given that single-omic analyses revealed a significant association between *FOXO1* and IPF, and considering the proximity of the P-value to the significance threshold in the integrative analysis, this finding may still deserve further investigation.

Integration of multi-omics evidence

We used locus zoom plots to display the distribution of *FOXO1* at various levels (Supplementary Figure S1) and SMR effect plots to show the impact of *FOXO1* at various levels on IPF risk (Supplementary Figure S2). By assessing the odds ratio (OR) values to determine risk relevance and direction of regulation, it was observed that

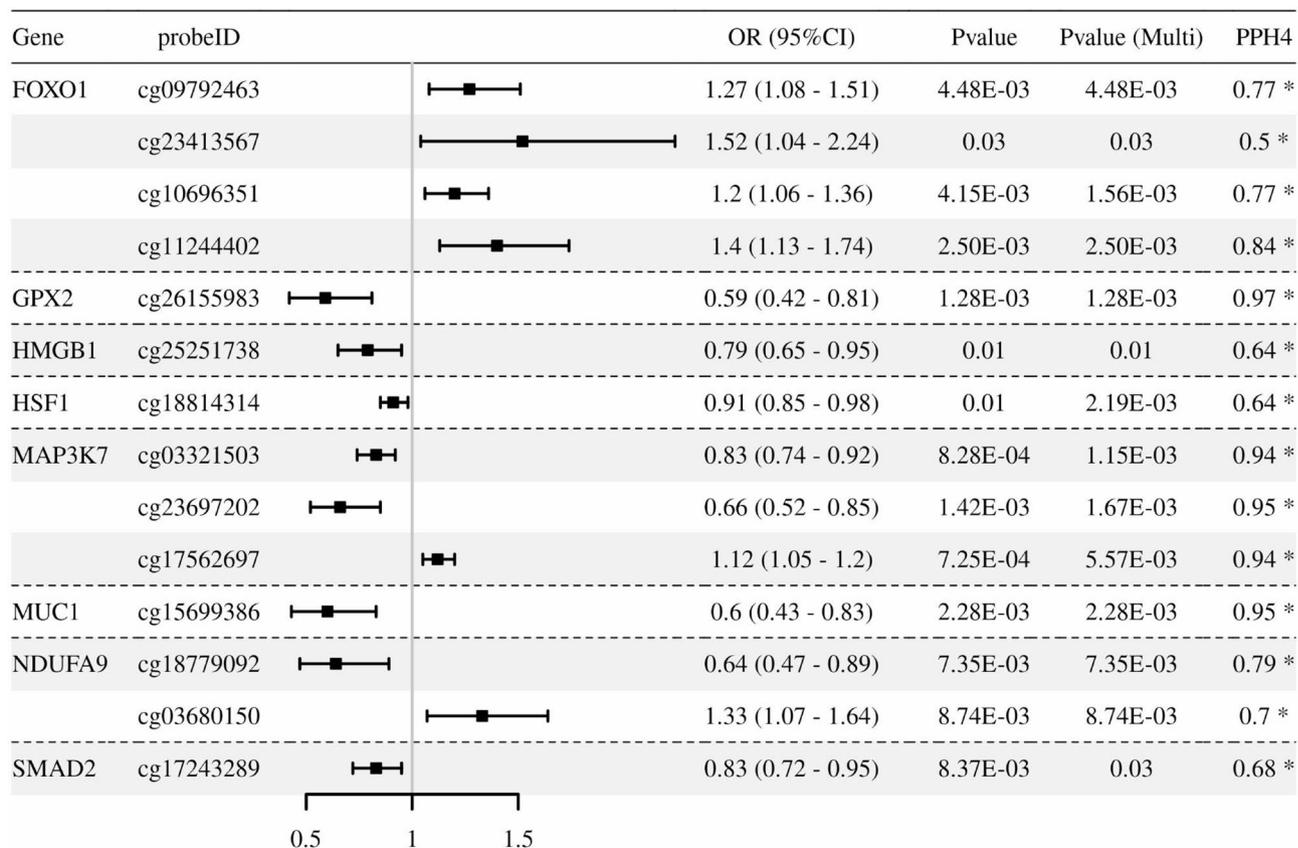


Fig. 3. Associations of genetically predicted oxidative stress-related gene methylation with idiopathic pulmonary fibrosis (IPF) in Mendelian randomization analysis.

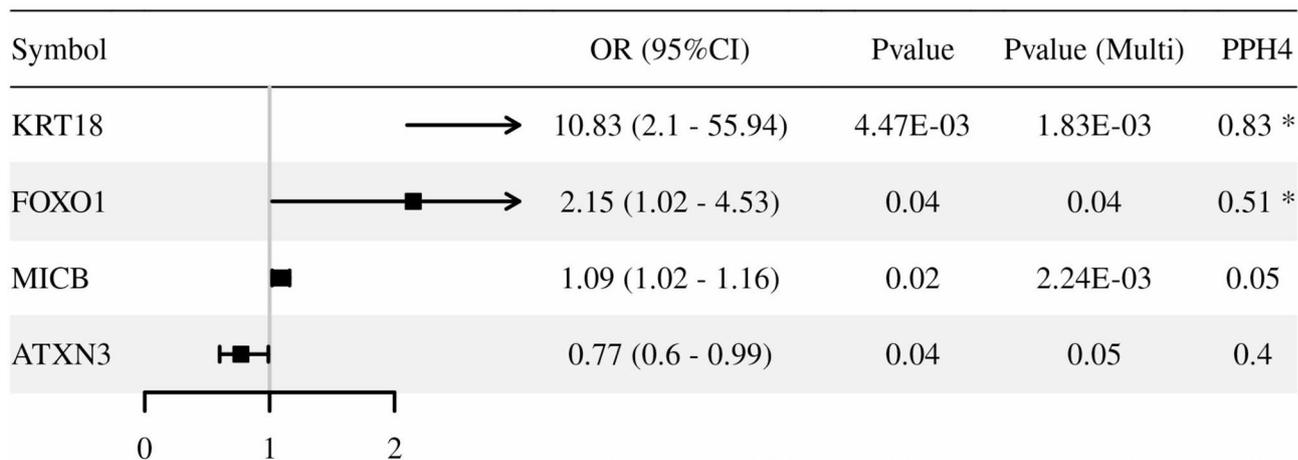


Fig. 4. Associations of genetically predicted oxidative stress-related protein abundance with idiopathic pulmonary fibrosis (IPF) in Mendelian randomization analysis.

FOXO1 gene expression levels were negatively correlated with IPF risk (OR=0.22, 95% CI=0.06–0.85) (Fig. 2), while methylation levels of CpG sites cg23413567 (OR=1.52, 95% CI=1.04–2.24) and cg11244402 (OR=1.4, 95% CI=1.13–1.74) were positively correlated with IPF risk (Fig. 3). Forkhead box protein O1 (FOXO1) protein abundance was positively correlated with IPF risk (OR=2.15, 95% CI=1.02–4.53) (Fig. 4); the methylation of these sites negatively regulated the expression levels of the associated gene, while gene expression levels were inversely correlated with the abundance of the encoded protein (Table 1). Based on single-level evidence, we hypothesized a potential model where higher methylation levels of cg23413567 and cg11244402 downregulate *FOXO1* expression, leading to an increase in FOXO1 protein abundance, which in turn elevates the risk of IPF.

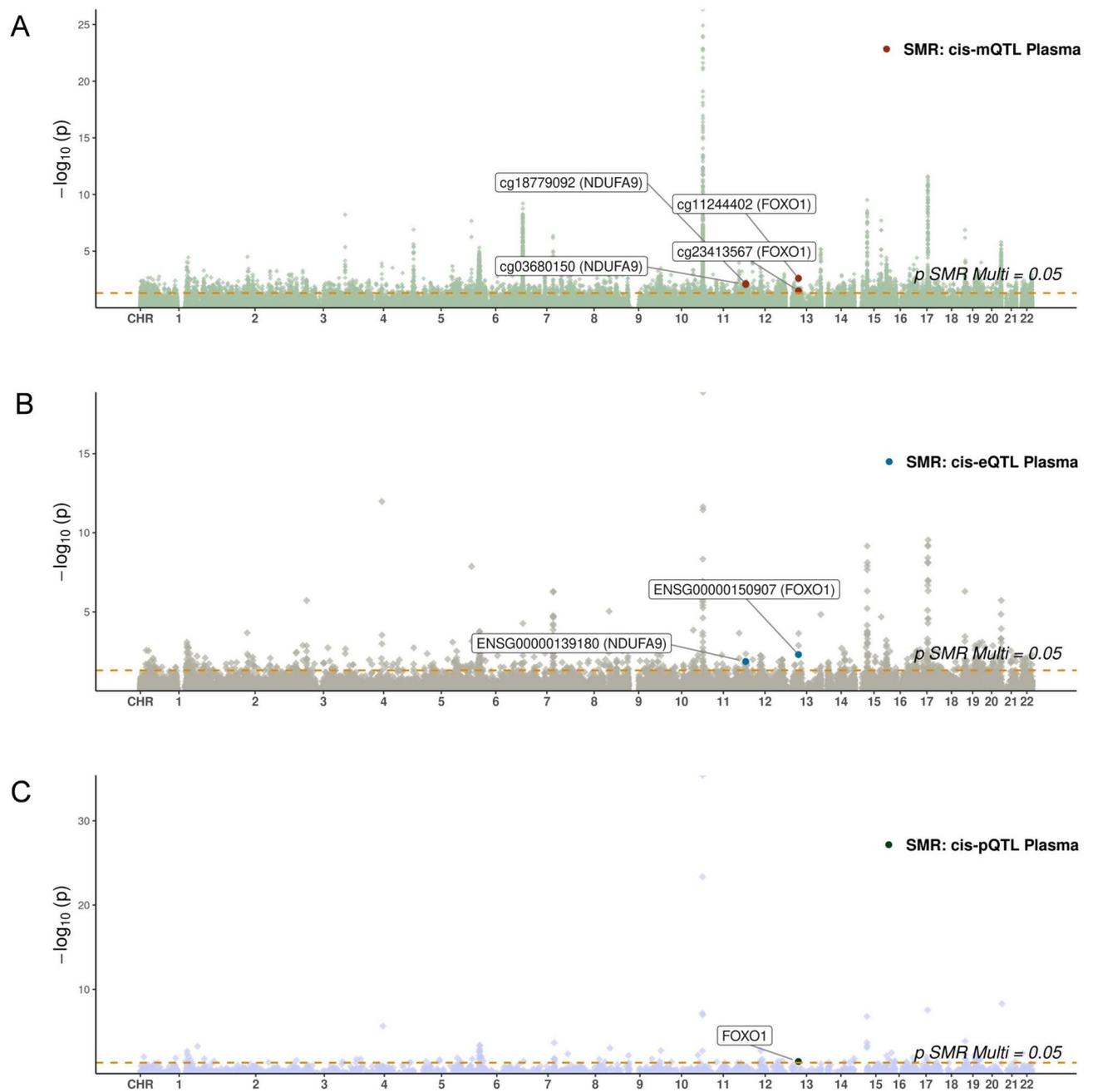


Fig. 5. Manhattan plot for associations between oxidative stress-related gene molecular features and idiopathic pulmonary fibrosis (IPF). Manhattan plot for oxidative stress-related gene methylation (A), expression (B) and protein abundance (C). Each point represents a genetic variant associated with a molecular trait (methylation, expression, or protein level), plotted by chromosomal position (x-axis) and $-\log_{10}(\text{p-value})$ from the SMR test (y-axis). The dashed orange line indicates the SMR multi-test significance threshold ($p = 0.05$). Significant loci, including FOXO1 and NDUFA9, are highlighted with annotations showing probe IDs or Ensembl gene IDs. These consistent associations across omics layers suggest potential causal relationships with idiopathic pulmonary fibrosis (IPF).

Tissue-specific validation

We further explored the causal relationship between the expression of identified genes in tissues and IPF. Specifically, we used eQTL data from lung tissue in the GTEx V8 database for analysis. However, the results from the integrated analysis mentioned above did not replicate in the tissue eQTLs' SMR analysis using the screening criteria of $p_{\text{SMR_multi}} < 0.05$, $p_{\text{SMR}} < 0.05$, and $P\text{-HEIDI} > 0.05$ (Supplementary Table S10).

Expo ID	Outco gene	p SMR	p SMR multi	p HEIDI	OR SMR (95% CI)
cg15699386(mQTL)	MUC1(eQTL)	2.87E-07	2.87E-07	2.20E-05	0.76 (0.68-0.84)
cg03321503(mQTL)	MAP3K7(eQTL)	8.27E-36	6.84E-33	6.02E-07	1.24 (1.2-1.28)
cg23697202(mQTL)	MAP3K7(eQTL)	1.27E-14	3.61E-15	0.021594	1.58 (1.41-1.78)
cg17562697(mQTL)	MAP3K7(eQTL)	1.22E-50	1.37E-42	2.74E-11	0.88 (0.86-0.89)
cg18814314(mQTL)	HSF1(eQTL)	4.40E-11	2.66E-10	0.045205	0.92 (0.89-0.94)
cg18779092(mQTL)	NDUFA9(eQTL)	1.56E-09	1.56E-09	0.132647	0.5 (0.4-0.63)
cg03680150(mQTL)	NDUFA9(eQTL)	1.09E-15	1.09E-15	0.247113	1.59 (1.42-1.78)
cg25251738(mQTL)	HMGB1(eQTL)	5.05E-10	5.05E-10	0.019692	1.19 (1.13-1.26)
cg09792463(mQTL)	FOXO1(eQTL)	4.84E-06	4.84E-06	0.009601	0.9 (0.86-0.94)
cg23413567(mQTL)	FOXO1(eQTL)	1.34E-05	1.34E-05	0.655891	0.75 (0.66-0.86)
cg10696351(mQTL)	FOXO1(eQTL)	9.47E-08	3.12E-07	0.004635	0.92 (0.89-0.95)
cg11244402(mQTL)	FOXO1(eQTL)	3.66E-07	3.66E-07	0.079158	0.86 (0.82-0.91)
cg26155983(mQTL)	GPX2(eQTL)	2.84E-06	2.84E-06	0.001973	1.27 (1.15-1.4)
cg17243289(mQTL)	SMAD2(eQTL)	3.50E-14	2.84E-12	0.001039	1.18 (1.13-1.23)
FOXO1(eQTL)	Forkhead box protein O1(pQTL)	0.054	0.011	0.027	0.83 (0.68-1)

Table 1. Results of integrating multi-omics evidence.

Discussion

This study systematically investigated the causal relationships between the methylation, expression, and protein abundance of oxidative stress-related genes and IPF through multi-omics approaches and SMR analysis. By integrating multi-omics evidence, we found that *FOXO1* may play a significant role in IPF. Additionally, combining single-level validation data, we also identified genes such as *NDUFA9*, *MUC1*, and *KRT18* as noteworthy candidates.

FOXO1 is a crucial transcription factor widely expressed in various tissues and organs. It plays a significant role in regulating key biological processes such as cell cycle, metabolism, oxidative stress response, apoptosis, and aging³². The findings of this study unveil the intricate relationship between *FOXO1* expression, the methylation levels of specific CpG sites, and the risk of IPF. The negative correlation between *FOXO1* gene expression and IPF risk suggests a protective role of *FOXO1* at the transcriptional level. Previous studies have demonstrated that *FOXO1* mediates oxidative stress responses and apoptosis, both essential for maintaining lung homeostasis and preventing fibrosis progression³². In contrast, the methylation of CpG sites cg23413567 and cg11244402 correlates positively with IPF risk, indicating an epigenetic mechanism where increased methylation suppresses *FOXO1* expression, thus promoting the development of IPF. This finding aligns with studies showing that hypermethylation in gene promoter regions typically leads to decreased gene expression, especially in genes involved in cellular stress responses³³. Kim et al. suggested that epigenetic therapies aimed at reversing hypermethylation might restore *FOXO1* expression levels, offering a novel approach to modulate IPF fibrotic responses³⁴. Interestingly, our results also reveal a positive correlation between *FOXO1* protein abundance and IPF risk. This could indicate a post-transcriptional regulation of *FOXO1*, where factors beyond transcriptional control affect protein stability or translation efficiency³⁵. It is supported by findings that oxidative stress can influence the post-translational modification of proteins like FOXO1, thereby affecting their stability and activity³⁶. Given these findings, the therapeutic potential of targeting the *FOXO1* pathway and the epigenetic modifications of these CpG sites warrants further exploration.

FOXO1 plays a complex role in fibrosis, acting as a potential suppressor of fibrogenic processes in various organs, including the liver, kidney, lung, and heart, by inhibiting fibroblast activation and extracellular matrix production³⁷⁻⁴⁰. Indeed, *FOXO1* can inhibit the activation of fibrogenic effector cells (like myofibroblasts), which are key players in fibrosis^{38,40}. In addition, *FOXO1* has been shown to reduce the production of ECM, a hallmark of fibrosis, in various organs^{38,40}. The transforming growth factor beta (TGF- β) pathway is a major driver of fibrosis, and *FOXO1* interacts with this pathway. TGF- β can increase *FOXO1* expression in cardiac fibroblasts and activate fibrogenic effector cells while simultaneously upregulating *FOXO1/3*, potentially limiting the effects of TGF- β ⁴¹. Platelet-derived growth factor (PDGF) phosphorylates *FOXO1* via the PI3K/Akt pathway, leading to *FOXO1* translocation from the nucleus to the cytosol, which can contribute to PDGF-induced proliferation of fibrogenic effector cells⁴⁰. *FOXO1* regulates macrophage functionality, which is crucial in tissue homeostasis and fibrosis⁴². *FOXO1* plays a role in regulating oxidative stress responses, which are implicated in fibrosis⁴². Furthermore, *FOXO1* activity is regulated by various post-translational modifications, which can either activate or inactivate it³⁸. Notch1 signaling is also essential for the *FOXO1*-mediated regulation of inflammation and fibrosis in macrophages⁴³. In IPF, additional potential mechanisms of *FOXO1* include the regulation of mesenchymal progenitor cells (MPCs)⁴⁴, autophagy⁴⁵, and the PGE-2, Akt, and JNK pathways^{46,47}. In the present study, decreased *FOXO1* expression levels were negatively correlated with IPF risk (OR = 0.22), while increased methylation of the *FOXO1* gene (which should decrease transcription) was positively associated with IPF (OR = 1.40), indicating that decreased *FOXO1* expression was associated with IPF, as described above. On the other hand, increased *FOXO1* protein expression in IPF (OR = 2.15) could also be a compensatory mechanism for protecting against IPF⁴⁴. Nevertheless, in vitro and in vivo experiments are necessary to determine the exact mechanisms. Our study highlights the critical role of *NDUFA9* in the pathogenesis of IPF.

The *NDUFA9* gene encodes NADH dehydrogenase (ubiquinone) Fe-S protein 9 (*NDUFA9*), a key subunit of mitochondrial complex I, which is essential for maintaining mitochondrial function and generating ROS. Specifically, our findings revealed a negative correlation between the methylation level at CpG site cg18779092 and IPF risk, suggesting that higher methylation at this site may suppress *NDUFA9* gene expression. As *NDUFA9* is crucial for mitochondrial function and ROS production, increased methylation at cg18779092 could lead to reduced *NDUFA9* expression, thereby alleviating mitochondrial dysfunction and oxidative stress and subsequently lowering the risk of IPF⁴⁸. In contrast, the positive correlation between methylation at cg03680150 and IPF risk indicates that higher methylation at this site may enhance *NDUFA9* gene expression. Elevated *NDUFA9* expression could exacerbate mitochondrial dysfunction and ROS production, thereby promoting the fibrotic processes characteristic of IPF⁴⁹. These results align with the significant role of oxidative stress in the pathogenesis of IPF^{50,51}. The differential effects of cg18779092 and cg03680150 methylation on *NDUFA9* expression underscore the complexity of epigenetic regulation in the disease context and highlight the potential of these methylation patterns as biomarkers for IPF risk assessment and therapeutic targets. We also observed a positive correlation between *NDUFA9* gene expression levels and IPF. Given that *NDUFA9* is involved in the function of mitochondrial complex I, we hypothesized that there may be a mechanistic link between *NDUFA9* and *FOXO1*, particularly in the context of oxidative stress and cellular metabolism. Specifically, increased *NDUFA9* expression may indicate enhanced complex I activity, leading to the generation of more ROS, which in turn could activate *FOXO1*, exacerbating oxidative stress and promoting the fibrotic process. However, further studies are needed to confirm the direct association between *NDUFA9* and *FOXO1* and to elucidate their specific mechanistic roles in IPF.

In addition, although evidence based on single-level data is not yet sufficient, the potential relevance of *MUC1* and *KRT18* in the pathology of IPF warrants further investigation. The *MUC1* gene encodes the Mucin 1 (*MUC1*) protein, which serves as a protective barrier on the cell surface and can mitigate oxidative stress-induced cellular damage. Through its surface glycosylation modifications, *MUC1* can neutralize ROS, reducing oxidative damage to cell membranes and internal organs⁵². Our study revealed a complex role for *MUC1* in IPF. Specifically, we found that DNA methylation at *MUC1* (cg15699386) is negatively correlated with IPF risk, and the gene expression level of *MUC1* is also negatively correlated with IPF risk. It suggests that *MUC1* may have a dual role in IPF, potentially influencing disease risk both through epigenetic regulation and gene expression levels. Higher levels of *MUC1* DNA methylation are generally associated with reduced expression of *MUC1*. This finding implies that increased methylation of *MUC1* might contribute to lower expression, potentially reducing the risk of IPF. This aligns with previous studies indicating that *MUC1* overexpression might exacerbate IPF through mechanisms involving oxidative stress and fibrosis⁵³. Despite the fact that high *MUC1* expression is typically associated with increased IPF risk, suggesting a potential promotive role in the disease process⁵⁴, the discrepancy between methylation and gene expression data highlights the need for further research to clarify the specific mechanisms by which *MUC1* contributes to IPF.

The *KRT18* gene encodes Keratin 18 (*KRT18*), which plays a crucial role in epithelial cell structure and stress response and is associated with oxidative stress and fibrotic diseases such as non-alcoholic fatty liver disease. Our study found a positive correlation between *KRT18* protein levels and the risk of IPF, aligning with its role in fibrotic processes⁵⁵. *KRT18* undergoes cleavage during apoptosis, with its fragment (CK18) serving as a biomarker for liver diseases^{56,57}. Our results indicated that increased *KRT18* protein abundance is associated with a higher risk of IPF, potentially reflecting its role in chronic injury and fibrosis. Additionally, *KRT18* mutations are linked to fibrosis in other organs, including the lungs⁵⁸, further supporting its potential as a biomarker for IPF.

The multi-omics analysis in this study has potential clinical significance. Key CpG sites and associated genes (such as *FOXO1*, *NDUFA9*, *MUC1*, and *KRT18*) with their methylation status and expression levels could serve as biomarkers for IPF, aiding in early diagnosis and risk assessment. Additionally, targeting the epigenetic regulation of these genes may offer new therapeutic strategies.

In this study, key genes such as *FOXO1* and *NDUFA9* were identified through multi-omics analysis, with promising associations to IPF. However, the lack of multi-level validation for some of these findings, along with the absence of tissue-specific data for these genes in available databases, may help explain the observed discrepancies. Firstly, while peripheral blood is a useful proxy for systemic changes, it may not fully capture lung-specific regulatory signals, especially those from cells involved in fibrogenesis. This could explain why tissue-level validation using GTEx lung eQTLs did not replicate some blood-derived findings. Secondly, GTEx lung data are from healthy individuals, while IPF involves distinct molecular alterations. Disease-specific regulatory mechanisms may not be adequately represented, potentially explaining the lack of replication. Thirdly, differences in sample sizes and population structure may also contribute. GTEx has relatively fewer lung tissue samples compared to blood QTL datasets, limiting statistical power. Additionally, inter-individual variability in cell composition across GTEx lung samples could obscure signals from relevant cell populations, such as alveolar epithelial cells or fibroblasts.

There are limitations to this study. The present study was exclusively performed using publicly available databases. Unfortunately, no such databases of lung biopsies, bronchial brushes, or bronchoalveolar lavage fluid were available for the purpose of the present study, and blood data of pQTL had to be used. In the present study, the datasets from Allen et al.¹⁵ and the FinnGen study yielded inconsistent results. Different databases are often from different populations with different genetic backgrounds and environmental exposures, and IPF is a complex trait influenced by genetics and the environment⁸. The FinnGen database exclusively includes Finnish individuals of Finnish ancestry, while the dataset by Allen et al.¹⁵ (used as the discovery cohort) is, in reality, from five cohorts from the United States of America, the United Kingdom, and Spain. In addition, the diagnoses in the FinnGen study are solely based on the ICD10 classification, while the cohorts used by Allen et al.¹⁵ used the ATS/ERS criteria for IPF diagnosis. Results obtained from smaller GWAS datasets could

also lack reproducibility and confidence. Finally, different GWASs can examine different numbers of SNPs and not necessarily the exact same ones. All those factors could contribute to different genes being identified between different datasets. It is why the present study focused on two genes (*FOXO1* and *NDUFA9*) that could be validated through a multi-omics approach involving multiple databases. In addition, MR and SMR analyses examine the associations between genetically predicted exposures and outcomes and cannot take into account the post-transcriptional and post-translational regulation mechanisms that can ultimately affect the protein expression of the identified genes. Additional studies will be necessary to explore the potential involvement of the other genes that were not consistently validated using multiple databases. Although *FOXO1* and *NDUFA9* are both known to be involved in oxidative stress and fibrotic diseases^{2–6}, experimental evidence is necessary for IPF to be able to better understand their role in the pathogenesis of IPF and their potential as therapeutic targets. Further experimental validation is needed to support these findings. Future research should integrate multi-omics data and functional experiments to better confirm the association of these genes with IPF, enhancing the reliability and clinical applicability of the study.

In conclusion, this study preliminarily explored the potential role of oxidative stress-related genes in IPF through SMR analysis and multi-omics approaches. It also examined the potential mechanisms by which *FOXO1* might be involved in oxidative stress and IPF, as well as the roles of *NDUFA9*, *MUC1*, and *KRT18* in this process, based on existing literature. These findings not only reflect the nuanced role of epigenetic regulation in disease progression but also suggest that future research should focus on unraveling the complex interactions among these genes. This is crucial for fully elucidating the intricate pathophysiological mechanisms linking oxidative stress and IPF, thereby providing more precise targets for disease prevention and treatment.

Data availability

All data generated or analyzed during this study are included in this article and supplementary information files.

Received: 13 November 2024; Accepted: 14 May 2025

Published online: 01 July 2025

References

- Klingsberg, R. C., Mutsaers, S. E. & Lasky, J. A. Current clinical trials for the treatment of idiopathic pulmonary fibrosis. *Respirology* **15**, 19–31 (2010).
- Richeldi, L. et al. Efficacy of a tyrosine kinase inhibitor in idiopathic pulmonary fibrosis. *N. Engl. J. Med.* **365**, 1079–1087 (2011).
- Teoh, A. K. Y. & Corte, T. J. Nonspecific interstitial pneumonia. *Semin. Respir. Crit. Care Med.* **41**, 184–201 (2020).
- Jelic, M. D., Mandic, A. D., Maricic, S. M. & Srdjenovic, B. U. Oxidative stress and its role in cancer. *J. Cancer Res. Ther.* **17**, 22–28 (2021).
- Ornatowski, W. et al. Complex interplay between autophagy and oxidative stress in the development of pulmonary disease. *Redox Biol.* **36**, 101679 (2020).
- Zhang, Y. et al. SIRT1 prevents cigarette smoking-induced lung fibroblasts activation by regulating mitochondrial oxidative stress and lipid metabolism. *J. Transl. Med.* **20**, 222 (2022).
- Matsunaga, T. et al. Supersulphides provide airway protection in viral and chronic lung diseases. *Nat. Commun.* **14**, 4476 (2023).
- Saha, P. & Talwar, P. Idiopathic pulmonary fibrosis (IPF): disease pathophysiology, targets, and potential therapeutic interventions. *Mol. Cell. Biochem.* **479**, 2181–2194 (2024).
- Bowden, J. & Holmes, M. V. Meta-analysis and Mendelian randomization: A review. *Res. Synth. Methods.* **10**, 486–496 (2019).
- Yengo, L. et al. Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum. Mol. Genet.* **27**, 3641–3649 (2018).
- Liu, X. et al. Deciphering the role of oxidative stress genes in idiopathic pulmonary fibrosis: a multi-omics Mendelian randomization approach. *Genes Immun.* **25**, 389–396 (2024).
- Zhao, W. et al. The role of oxidative stress-related genes in idiopathic pulmonary fibrosis. *Sci. Rep.* **15**, 5954 (2025).
- Li, Y. et al. Machine learning-based radiomics to distinguish pulmonary nodules between lung adenocarcinoma and tuberculosis. *Thorac. Cancer.* **15**, 466–476 (2024).
- Badia-Bringué, G. et al. Summary-data based Mendelian randomization identifies gene expression regulatory polymorphisms associated with bovine paratuberculosis by modulation of the nuclear factor kappa B (NF- κ B)-mediated inflammatory response. *BMC Genom.* **24**, 605 (2023).
- Allen, R. J. et al. Genome-wide association study across five cohorts identifies five novel loci associated with idiopathic pulmonary fibrosis. *Thorax* **77**, 829–833 (2022).
- Giambartolomei, C. et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
- Guo, Z. et al. Assessing the causal relationships between human blood metabolites and the risk of NAFLD: A comprehensive Mendelian randomization study. *Front. Genet.* **14**, 1108086 (2023).
- Sun, J., Wu, Y., Burgess, S., Weng, Y. & Wang, Z. Mitochondrial-related genome-wide Mendelian randomization identifies putatively causal genes in the pathogenesis of sepsis. *Surgery* **181**, 109150 (2025).
- Wu, F., Huang, Y., Hu, J. & Shao, Z. Mendelian randomization study of inflammatory bowel disease and bone mineral density. *BMC Med.* **18**, 312 (2020).
- Wu, Y. et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat. Commun.* **9**, 918 (2018).
- Vösa, U. et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310 (2021).
- Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK biobank. *Nature* **622**, 329–338 (2023).
- Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
- Chen, J. et al. Multi-omic insight into the molecular networks of mitochondrial dysfunction in the pathogenesis of inflammatory bowel disease. *EBioMedicine* **99**, 104934 (2024).
- Xu, Q. et al. Causal relationship between gut microbiota and autoimmune diseases: A Two-Sample Mendelian randomization study. *Front. Immunol.* **12**, 746998 (2021).
- Maimaiti, A. et al. DNA methylation regulator-mediated modification patterns and risk of intracranial aneurysm: a multi-omics and epigenome-wide association study integrating machine learning, Mendelian randomization, eQTL and mQTL data. *J. Transl. Med.* **21**, 660 (2023).

27. Morrow, J. D. et al. Human lung DNA methylation quantitative trait loci colocalize with chronic obstructive pulmonary disease Genome-Wide association loci. *Am. J. Respir. Crit. Care Med.* **197**, 1275–1284 (2018).
28. Battle, A., Brown, C. D., Engelhardt, B. E. & Montgomery, S. B. Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).
29. Yoshiji, S. et al. Proteome-wide Mendelian randomization implicates nephronectin as an actionable mediator of the effect of obesity on COVID-19 severity. *Nat. Metab.* **5**, 248–264 (2023).
30. Dobbyn, A. et al. Landscape of conditional eQTL in dorsolateral prefrontal cortex and Co-localization with schizophrenia GWAS. *Am. J. Hum. Genet.* **102**, 1169–1184 (2018).
31. Breen, M. S. et al. Global landscape and genetic regulation of RNA editing in cortical samples from individuals with schizophrenia. *Nat. Neurosci.* **22**, 1402–1412 (2019).
32. Pairo-Castineira, E. et al. GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19. *Nature* **617**, 764–768 (2023).
33. Zhang, Y. e. Role of FOXO1 in cellular apoptosis and oxidative stress responses in pulmonary cells. *J. Respir. Cell. Mol. Biol.* **59**, 487–495 (2018).
34. Kim, Y. H. et al. Targeting epigenetic modifications to modulate FOXO1 expression in lung fibrosis. *Clin. Epigeneti.* **15**, 87 (2023).
35. Liu, M. al. e. Post-transcriptional regulation of FOXO1 protein in myofibroblasts and its implications in IPF. *Am. J. Physiol. Lung Cell Mol. Physiol.* **319**, L1091–L1103 (2020).
36. Walker, R. L. & Turner, J. P. Epigenetic therapy approaches in pulmonary fibrosis. *Curr. Opin. Pulm. Med.* **27**, 421–429 (2021).
37. Huang, F. et al. FoxO1-mediated inhibition of STAT1 alleviates tubulointerstitial fibrosis and tubule apoptosis in diabetic kidney disease. *eBioMedicine.* **48**, 491–504 (2019).
38. Sergi, C., Shen, F. & Liu, S. M. Insulin/IGF-1R, SIRT1, and FOXOs Pathways—An intriguing interaction platform for bone and osteosarcoma. *Front. Endocrinol. (Lausanne).* **10**, 93 (2019).
39. Yu, W., Chen, C. & Cheng, J. The role and molecular mechanism of FoxO1 in mediating cardiac hypertrophy. *ESC Heart Fail.* **7**, 3497–3504 (2020).
40. Xin, Z. et al. FOXO1/3: potential suppressors of fibrosis. *Ageing Res. Rev.* **41**, 42–52 (2018).
41. Vivar, R. et al. FoxO1 mediates TGF-beta1-dependent cardiac myofibroblast differentiation. *Biochim. Et Biophys. Acta (BBA) Mol. Cell. Res.* **1863**, 128–138 (2016).
42. Rong, S. J. et al. The essential role of FoxO1 in the regulation of macrophage function. *Biomed. Res. Int.* **2022**, 1068962 (2022).
43. Xu, D. et al. The Foxo1-YAP-Notch1 axis reprograms STING-mediated innate immunity in NASH progression. *Exp. Mol. Med.* **56**, 1843–1855 (2024).
44. Jbeli, A. H. et al. Brg1/PRMT5 nuclear complex epigenetically regulates FOXO1 in IPF mesenchymal progenitor cells. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **326**, L344–L352 (2024).
45. Fan, G., Liu, J., Wu, Z., Li, C. & Zhang, Y. Development and validation of the prognostic model based on autophagy-associated genes in idiopathic pulmonary fibrosis. *Front. Immunol.* **13**, 1049361 (2022).
46. Meng, Z. X. et al. Prostaglandin E2 regulates Foxo activity via the Akt pathway: implications for pancreatic islet beta cell dysfunction. *Diabetologia.* **49**, 2959–2968 (2006).
47. Kawamori, D. et al. The forkhead transcription factor Foxo1 bridges the JNK pathway and the transcription factor PDX-1 through its intracellular translocation**. *J. Biol. Chem.* **281**, 1091–1098 (2006).
48. Miller, J. T. & Davis, P. E. Molecular mechanisms underlying the stability of FOXO1 protein in oxidative stress conditions. *Free Radic. Biol. Med.* **180**, 236–244 (2022).
49. Lin, C. S. et al. Role of mitochondrial function in the invasiveness of human colon cancer cells. *Oncol. Rep.* **39**, 316–330 (2018).
50. Dlasková, A., Clarke, K. J., Rooney, M. F. & Porter, R. K. The use of reactive oxygen species production by Succinate-Driven reverse Electron flow as an index of complex I activity in isolated brown adipose tissue mitochondria. *Methods Mol. Biol.* **2310**, 247–258 (2021).
51. Gadicherla, A. K., Stowe, D. F., Antholine, W. E., Yang, M. & Camara, A. K. Damage to mitochondrial complex I during cardiac ischemia reperfusion injury is reduced indirectly by anti-anginal drug Ranolazine. *Biochim. Biophys. Acta.* **1817**, 419–429 (2012).
52. Pillai, K., Pourgholami, M. H., Chua, T. C. & Morris, D. L. MUC1 as a potential target in anticancer therapies. *Am. J. Clin. Oncol.* **38**, 108–118 (2015).
53. Kasprzak, A. & Adamek, A. Mucins: the old, the new and the promising factors in hepatobiliary carcinogenesis. *Int J. Mol. Sci.* **20**, (2019).
54. Stroopinsky, D., Kufe, D. & Avigan, D. MUC1 in hematological malignancies. *Leuk. Lymphoma.* **57**, 2489–2498 (2016).
55. Milbank, E. et al. Liver lipopolysaccharide binding protein prevents hepatic inflammation in physiological and pathological non-obesogenic conditions. *Pharmacol. Res.* **187**, 106562 (2023).
56. Chen, L. et al. Proteomic response of the rat liver in differential swimming modes. *Clin. Exp. Pharmacol. Physiol.* **45**, 581–590 (2018).
57. Boonanuntasarn, S., Nakharuthai, C., Schrama, D., Duangkaew, R. & Rodrigues, P. M. Effects of dietary lipid sources on hepatic nutritive contents, fatty acid composition and proteome of Nile tilapia (*Oreochromis niloticus*). *J. Proteom.* **192**, 208–222 (2019).
58. Ku, N. O. et al. Keratins as susceptibility genes for end-stage liver disease. *Gastroenterology.* **129**, 885–893 (2005).

Acknowledgements

None.

Author contributions

(I) Conception and design: Y.W. (II) Administrative support: H.Z. (III) Provision of study materials or patients: L.X., S.H. (IV) Collection and assembly of data: Y.W., Z.Z., X.L. and Z.Z. (V) Data analysis and interpretation: Y.W., J.L. (VI) Manuscript writing: All authors. (VII) Final approval of manuscript: All authors.

Funding

This study was partially supported by grants from the following sources: (1) Natural Science Foundation of Guangdong Province (Grants No. 2023A1515011329 and 2025A1515012984); (2) Research Funds of the Joint Research Center for Occupational Medicine and Health of IHM (Grant No. OMH-2023-19); and (3) Open Project of the Anhui Province Key Laboratory of Occupational Health (Grant No. 2024ZYJKC001).

Declarations

Ethics approval and consent to participate

This article is a Mendelian randomization study. The data for this study were obtained from publicly available

databases and published literature data and do not require ethical approval and written informed consent.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-02579-7>.

Correspondence and requests for materials should be addressed to X.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025