Clin. Lab. 2026;72:XXX-XXX

ORIGINAL ARTICLE

Mendelian Randomization and Transcriptomic Analysis Identify HE4 as a Potential Biomarker for Chronic Obstructive Pulmonary Disease

Mei-Xue Li [1, *], Zhang-Fang Zeng [2, *], Zhao-Wen Cao [1]

[*] *Mei-Xue Li and Zhang-Fang Zeng contributed equally to this study*
[1] *Department of Respiratory, Pingyang Hospital of Traditional Chinese Medicine, Pingyang, Zhejiang Province, P.R. China*
[2] *Department of Nursing, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang Province, P.R. China*

SUMMARY

Background: **Human epididymis protein 4 (HE4), a member of the WFDC protein family, has been shown to exhibit abnormal expression in patients with chronic obstructive pulmonary disease (COPD). However, the causal relationship between HE4 and COPD, as well as the involvement of other members of the WFDC family in COPD pathogenesis, remains unclear. This study aims to investigate the potential of WFDC proteins, particularly HE4, as biological markers for COPD through mendelian randomization (MR) and transcriptomic analysis.**

Methods: **A comprehensive search of the IEU OpenGWAS database was conducted, and single nucleotide polymorphisms (SNPs) associated with WFDC family members, including HE4 (WFDC2), WFDC3, WFDC5, WFDC10A, WFDC12, and WFDC13, were selected as instrumental variables for MR analysis. Two-sample MR was performed to examine the direct effects of WFDC proteins on COPD. Additionally, transcriptomic data related to COPD were obtained from the GEO database and analyzed using bioinformatics tools. Furthermore, the plasma level of HE4 was quantified using enzyme-linked immunosorbent assay (ELISA).**

Results: **MR analysis revealed a significant positive association between elevated *HE4* expression and increased COPD risk (OR = 1.10, 95% CI = 1.02 - 1.18, p = 0.01). However, no causal relationship was found between the expression of other WFDC family members and COPD (all p > 0.05). Transcriptomic analysis showed that *HE4* was overexpressed in the lung tissues of COPD patients compared to normal controls. In contrast, two other WFDC family members, *WFDC3* and *WFDC12*, exhibited significantly lower expression in COPD lung tissue (both p < 0.05). ROC curve analysis indicated that *HE4* (*WFDC2*) demonstrated a diagnostic potential for COPD with an area under the curve (AUC) of 0.78, while *WFDC3* and *WFDC12* showed lower AUCs of 0.50 and 0.57, respectively. ELISA detection demonstrated that the plasma level of secretory HE4 in COPD patients was significantly higher than that in healthy controls (p < 0.05).**

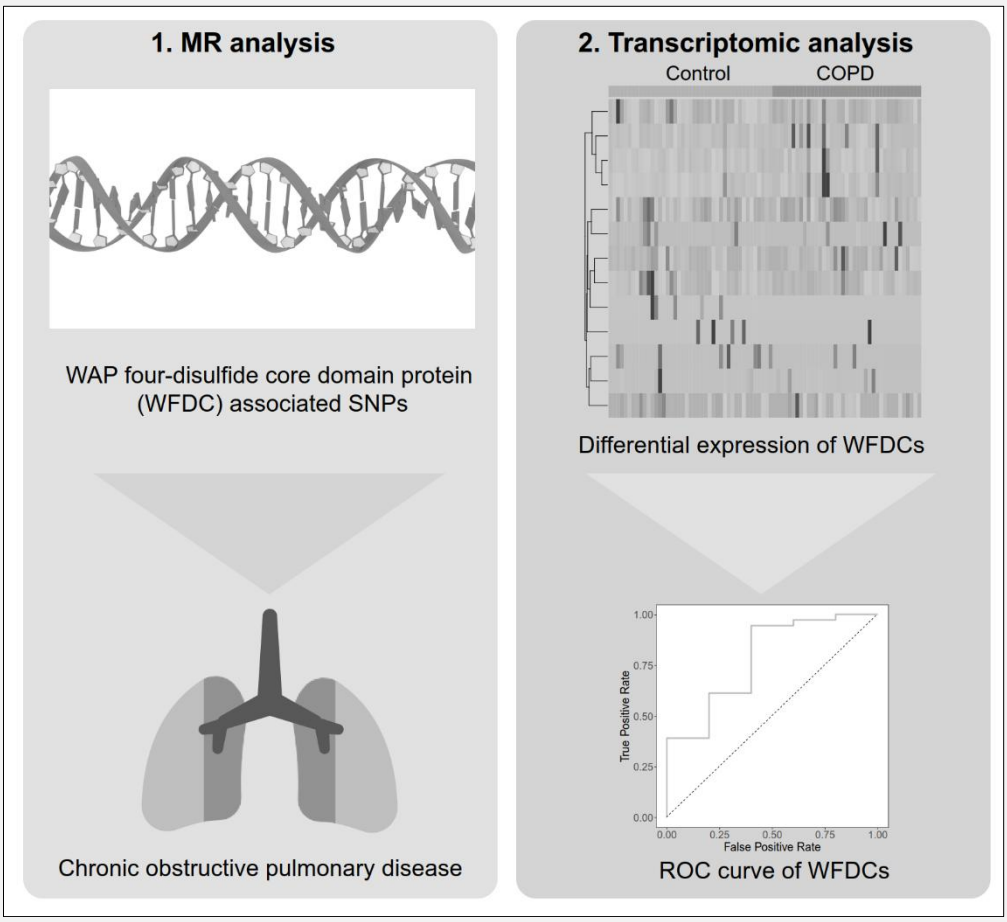
Conclusions: **Clinical data and transcriptomic analysis each confirmed that HE4 is highly expressed in plasma, and *HE4* is highly expressed in lung tissue. Increased expression of HE4 (WFDC2) is associated with an elevated risk of COPD, and its high expression in lung tissue holds promise as a diagnostic biomarker for COPD. This study provides valuable insights into the role of WFDC proteins in COPD and highlights the potential of HE4 as a biomarker for early diagnosis and risk stratification.**

(Clin. Lab. 2026;72:xx-xx. DOI: 10.7754/Clin.Lab.2025.250536)

Correspondence:
Zhao-Wen Cao
Department of Respiratory
Pingyang Hospital of Traditional Chinese Medicine
No. 516, West Road of Xingao, Aojiang Town
Pingyang 325401, Zhejiang Province
P.R. China
Email: czwen97@outlook.com

Manuscript accepted August 30, 2025

Graphical abstract



KEYWORDS

whey acidic protein four-disulfide core domain, human epididymis protein 4, Mendelian randomization, transcriptome analysis, chronic obstructive pulmonary disease

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a prevalent and preventable chronic lung disease characterized by persistent airflow limitation, ranking as the third leading cause of death globally and emerging as a major public health concern [1]. According to projections based on a health-augmented macroeconomic model, both China and the United States are expected to experience substantial macroeconomic losses due to COPD between 2020 and 2050 [2]. Furthermore, based on the Global Burden of Diseases, Injuries, and Risk Factors Study 2021 database, despite improvements in the age-standardized prevalence, mortality, and disability-adjusted life years of COPD since 1990, the overall burden continues to escalate, with an estimated global prevalence of 213.39 million cases in 2021 [3].

COPD is influenced by a wide range of risk factors and has a complex pathophysiology, exhibiting significant heterogeneity in clinical phenotypes, treatment response, disease progression, and prognosis. Identifying reliable biomarkers for early diagnosis, disease assessment, and prognosis prediction is therefore a key focus in current COPD research. The establishment of the COPD Biomarker Qualification Consortium in 2010 has accelerated biomarker development, aiming to evaluate potential biomarkers and support drug development by providing robust diagnostic and monitoring indicators [4]. For instance, the association between blood eosinophil counts and the therapeutic efficacy of inhaled corticosteroids has been validated in prospective clinical trials [5], and its utility is explicitly recommended in the Global Initiative for Chronic Obstructive Lung Disease 2020 Report [6]. However, additional biomarkers are

needed to improve early diagnosis, precision assessment, and outcome prediction in COPD.

Human epididymis protein 4 (HE4), also known as WFDC2, was first identified in the human epididymis by Kirchhoff et al. in 1991 and belongs to the whey acidic protein (WAP) four-disulfide core domain (WFDC) family. HE4 plays a key role in the development of inflammatory and fibrotic diseases and, together with other WFDC family members, may contribute to airway remodeling through the protease-antiprotease imbalance pathway [7,8]. However, due to the inherent complexity of COPD, the causal relationships between HE4 or other WFDC proteins and COPD remain poorly understood. Mendelian randomization (MR), which uses genetic variants as unconfounded proxies for modifiable exposures, provides a powerful approach to infer causality between risk factors and disease outcomes [9]. In this study, we employed MR analysis to investigate the potential causal effects of HE4 and other WFDC family members on the risk of COPD. Furthermore, we integrated transcriptomic analysis using publicly available datasets from the Gene Expression Omnibus (GEO) to explore the expression profiles of WFDC proteins in COPD and evaluate their potential as diagnostic biomarkers. These findings aim to provide novel insights into the pathogenesis of COPD and support the development of early screening strategies.

MATERIALS AND METHODS

Data sources

The GWAS data for the WFDC protein family members HE4 (WFDC2), WFDC3, WFDC5, WFDC10A, WFDC 12, and WFDC13 were obtained from the plasma proteomic genetic map of 3,301 healthy individuals [10]. The COPD-related GWAS data were sourced from the FinnGen consortium (https://gwas.mrcieu.ac.uk/), comprising a summary of data from 6,915 COPD patients and 186,723 healthy controls [11]. Transcriptomic data associated with COPD (GSE239897) were retrieved from the GEO, including 39 lung tissue samples from COPD patients and 43 healthy controls [12]. Clinical data from COPD patients (n = 12) and healthy subjects (n = 12) were collected at Pingyang Hospital of Traditional Chinese Medicine, after informed consent was obtained from all participants.

Selection of instrumental variables

The selection of instrumental variables (IVs) for MR analysis must satisfy three key assumptions: I. relevance assumption: SNPs must be associated with the exposure factors, specifically the WFDC protein family members HE4 (WFDC2), WFDC3, WFDC5, WFDC10A, WFDC 12, and WFDC13; II. exclusivity assumption: SNPs should not be associated with COPD; III. independence assumption: SNPs must be independent of confounding factors. To meet these criteria, we set the genome-wide significance threshold for SNPs strongly associated with the exposure factors at $p < 1 \times 10^{-5}$ [13]. Additionally, to ensure the independence of IVs, SNPs in linkage disequilibrium (LD) were excluded using thresholds of $R^2 < 0.001$ and kb = 10,000. The F-statistic was then calculated, with an $F < 10$ indicating weak IVs. The formulas for calculating R^2 and F are as follows: $R^2 = 2 \times (1 - MAF) \times MAF \times \beta^2$; $F = (N - k - 1)/k \times R^2/(1 - R^2)$. Finally, we searched for all SNPs meeting these criteria in PhenoScanner (http://www.phenoscanner.medschl.cam.ac.uk/) and excluded those that were associated with confounding factors between the WFDC protein family members and COPD.

MR analysis

The study methods were conducted in accordance with the STROBE-MR checklist [14]. Statistical analyses in this study were performed using R software (version 4.3.2), with MR analysis conducted using the "Two SampleMR," "MRPRESSO," and "Mendelian Randomization" packages (https://github.com/czwen97/HE4-for-COPD). MR analysis was carried out employing five common MR methods, including inverse-variance weighted (IVW), MR-Egger, weighted median, MR-PRESSO, and simple mode, to explore the potential causal relationship between the WFDC protein family members and COPD. The results were visualized through forest plots and funnel plots. IVW, as the primary MR method, utilizes a meta-analysis approach that combines the Wald estimate for each SNP, providing an overall assessment of the effect of WFDC protein family members on COPD [15].

In comparison with IVW, MR-Egger incorporates an intercept term and uses the inverse of the outcome variance as weights to fit a linear regression model, with the significance of the intercept value used to assess the presence of horizontal pleiotropy [16]. Additionally, MR-PRESSO was used to detect and exclude outlier SNPs, followed by Cochran's Q test to evaluate heterogeneity, with a p-value greater than 0.05 indicating the absence of heterogeneity among the included SNPs [17]. Sensitivity analyses were also conducted using the leave-one-out method to visualize the impact of individual SNPs on the overall effect. This method involved sequentially removing one SNP at a time and observing whether the effect estimates changed substantially, allowing for an assessment of the influence of each SNP on the causal relationship [18].

Bioinformatics analysis

Statistical analyses were conducted using R software (version 4.3.2). The "stats" and "car" packages were employed to compare the gene expression levels of WFDC protein family members *HE4* (*WFDC2*), *WFDC3*, *WFDC5*, *WFDC10A*, *WFDC12*, and *WFDC13* in lung tissues from COPD patients and healthy controls. Heatmaps were generated using the "ComplexHeatmap" package. Additionally, receiver operating characteristic (ROC) analysis was performed using the "pROC" package (https://github.com/czwen97/HE4-for-COPD).

ELISA detection and analysis

Fasting venous blood samples (approximately 2 mL) were collected from COPD patients and healthy controls in the early morning, placed in non-anticoagulant tubes, and centrifuged at 3,000 r/minute for 8 minutes, after which the plasma samples were stored at -80℃ for subsequent analysis. Concentration of secretory HE4 was measured using an ELISA kit (Quanzhou Ruixin Biotechnology Co., Ltd, Quanzhou City, Fujian Province, China).

To ensure the reliability of the plasma HE4 measurements, each patient's plasma sample was tested in duplicate, and the average of the two results was used for analysis. Subsequently, statistical analysis was performed to compare HE4 expression between COPD patients and healthy controls using the Mann-Whitney U test.

RESULTS

MR analysis

Based on the screening criteria set according to the three core assumptions of MR, a total of 23, 24, 24, 25, 23, and 23 SNPs were selected as instrumental variables for analyzing the causal relationship between HE4 (WFDC2), WFDC3, WFDC5, WFDC10A, WFDC12, and WFDC13 and COPD, respectively, with each SNP having an F-value greater than 10. IVW method analysis showed a significant association between elevated HE4 levels and an increased risk of COPD (OR = 1.10, 95% CI = 1.02 - 1.18, p = 0.01). MR-PRESSO analysis further confirmed this significant association (OR = 1.10, 95% CI = 1.03 - 1.18, p = 0.01) (Table 1). Cochran's Q test confirmed that there was no heterogeneity in the results (Q = 24.61, p = 0.32), and the MR-PRESSO intercept indicated the absence of horizontal pleiotropy (intercept = 0.03, p = 0.12). Sensitivity analysis scatter plot (Figure 1A) and symmetrical funnel plot (Figure 1B) showed that the effect direction and magnitude of multiple MR methods were consistent with IVW and demonstrated the absence of horizontal pleiotropy, further confirming the validity of the causal relationship. The leave-one-out method (Figure 2) demonstrated that the removal of individual SNPs did not significantly affect the outcome, indicating the robustness of the results.

Transcriptomic analysis

The results of the Wilcoxon rank-sum test revealed that *HE4* (*WFDC2*) was significantly upregulated in the lung tissue of COPD patients compared to normal lung tissue, whereas *WFDC3* and *WFDC12* were significantly downregulated in COPD patients' lung tissue (all $p < 0.05$) (Figure 3). ROC curve analysis demonstrated that among the three genes of WFDC protein family members, only *HE4* exhibited an area under the curve (AUC) greater than 0.70, indicating excellent diagnostic performance, with an optimal cutoff value of 0.71 (Figure 4).

ELISA analysis

To investigate the expression of HE4 in COPD patients, we performed ELISA to measure its levels in peripheral blood plasma. The ELISA results revealed that the plasma level of secretory HE4 in COPD patients (97.69 pmol/L) was significantly higher than that in healthy controls ($p < 0.05$) (Figure 5).

DISCUSSION

COPD is a multifactorial condition with a complex pathophysiological process, and its risk factors remain incompletely understood. This study employs MR analysis to explore the association between six members of the WFDC protein family and COPD. Our findings confirm, for the first time, that elevated levels of HE4 (WFDC2) increase the risk of COPD, with HE4 showing high expression in the lung tissues of COPD patients and demonstrating good diagnostic efficacy.

The *WFDC2* gene is located on the human chromosome 20q12-q13.1 region. HE4, a cysteine-rich protein encoded by the *WFDC2* gene, is a small acidic polypeptide approximately 11.78 kb in length, known to function as a protease inhibitor. HE4 has been found to be abnormally expressed in various malignant tumors, and it has been approved by the U.S. Food and Drug Administration as a new tumor marker for monitoring the therapeutic response and recurrence of ovarian cancer [19]. Furthermore, HE4 is widely distributed across multiple tissues and organs. Bingle et al. [20] used immunohistochemistry to demonstrate that HE4 is predominantly expressed in the epithelial cells of the upper respiratory tract and in the glands and ducts of submucosal glands. They also noted that HE4 expression is significantly elevated in the airways of cystic fibrosis patients. Additionally, several studies have found a significant negative correlation between HE4 expression levels and pulmonary function in cystic fibrosis patients, suggesting that HE4 could serve as a potential inflammatory marker for evaluating treatment efficacy in these patients [21]. This may be related to HE4's protease inhibitory activity. Under normal conditions, HE4, in coordination with other protease inhibitors, regulates the dynamic balance of lung development, extracellular matrix degradation and synthesis, inflammation, immune modulation, and tissue repair [22]. Elevated HE4 levels may disrupt the protease/antiprotease balance, leading to excessive degradation of the extracellular matrix and epithelial junctions, thereby compromising the integrity of the airway epithelium. Moreover, the protease/antiprotease imbalance induced by abnormal HE4 expression may exacerbate airway inflammation in COPD by promoting the accumulation of inflammatory cells and the release of inflammatory mediators, contributing to disease progression.

COPD has an insidious onset, and it is often diagnosed in the presence of multiple comorbidities. Therefore, identifying potential biomarkers for early screening is

Table 1. MR Analysis of six WFDC protein family members and COPD.

Exposures	SNP (n)	Methods	β	*SE*	OR (95% CI)	p
HE4 (WFDC2)	23	IVW	0.10	0.04	1.10 (1.02 - 1.18)	0.01
		MR-Egger	-0.05	0.09	0.95 (0.79 - 1.15)	0.63
		Weighted median	0.04	0.05	1.04 (0.95 - 1.15)	0.46
		MR-PRESSO	0.10	0.03	1.10 (1.03 - 1.18)	0.01
		Simple mode	0.01	0.10	1.00 (0.83 ~ 1.21)	0.96
WFDC3	24	IVW	0.01	0.03	1.00 (0.94 - 1.07)	0.90
		MR-Egger	-0.02	0.07	0.98 (0.86 - 1.12)	0.80
		Weighted median	-0.03	0.04	0.97 (0.90 - 1.05)	0.50
		MR-PRESSO	0.01	0.03	1.00 (0.94 - 1.07)	0.90
		Simple mode	-0.01	0.07	0.99 (0.87 - 1.13)	0.91
WFDC5	24	IVW	0.01	0.03	1.01 (0.95 - 1.07)	0.78
		MR-Egger	-0.03	0.06	0.97 (0.87 - 1.08)	0.58
		Weighted median	0.02	0.04	1.02 (0.93 - 1.11)	0.68
		MR-PRESSO	0.01	0.02	1.01 (0.96 - 1.06)	0.73
		Simple mode	0.04	0.08	1.04 (0.90 - 1.21)	0.58
WFDC10A	25	IVW	0.01	0.03	1.00 (0.94 - 1.07)	0.98
		MR-Egger	0.11	0.08	1.11 (0.96 - 1.29)	0.17
		Weighted median	-0.07	0.05	0.94 (0.85 - 1.03)	0.16
		MR-PRESSO	0.01	0.03	1.00 (0.94 - 1.07)	0.98
		Simple mode	-0.07	0.09	0.93 (0.79 - 1.10)	0.4311
WFDC12	23	IVW	-0.04	0.03	0.96 (0.90 - 1.03)	0.24
		MR-Egger	-0.11	0.07	0.90 (0.78 - 1.04)	0.16
		Weighted median	-0.04	0.04	0.96 (0.88 - 1.04)	0.31
		MR-PRESSO	-0.04	0.02	0.96 (0.92 - 1.00)	0.09
		Simple mode	-0.05	0.07	0.95 (0.83 - 1.10)	0.52
WFDC13	23	IVW	0.06	0.03	1.06 (0.99 - 1.12)	0.08
		MR-Egger	-0.01	0.07	0.99 (0.87 - 1.13)	0.91
		Weighted median	0.03	0.06	1.03 (0.93 - 1.16)	0.55
		MR-PRESSO	0.05	0.03	1.05 (0.99 - 1.12)	0.08
		Simple mode	-0.01	0.12	1.00 (0.79 - 1.25)	0.98

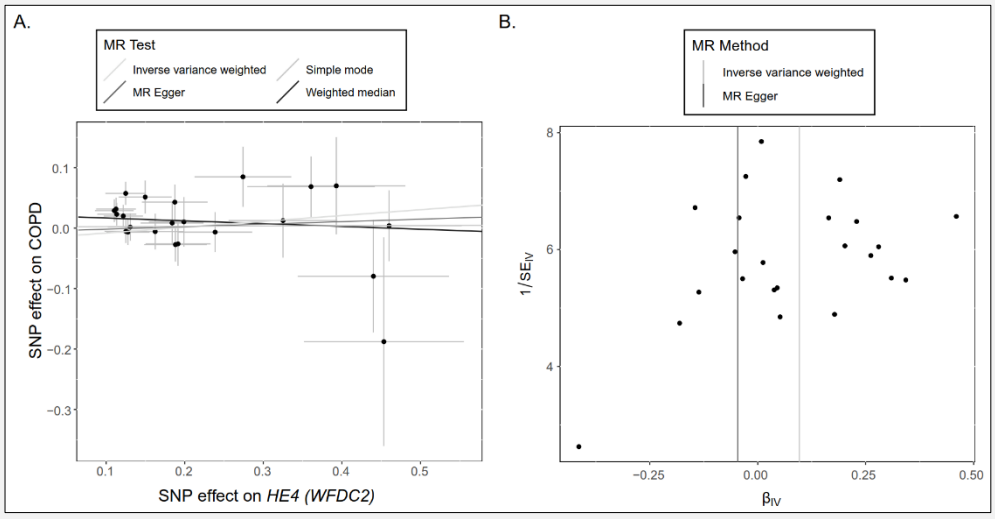


Figure 1. MR analysis of the causal relationship between WFDC protein family members and COPD.

A - Sensitivity analysis scatter plot, B - Funnel plot.

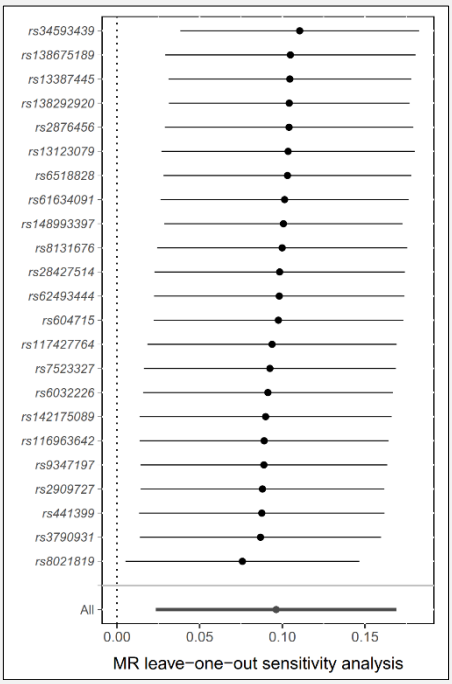


Figure 2. Leave-one-out analysis of the causal relationship between WFDC protein family members and COPD.

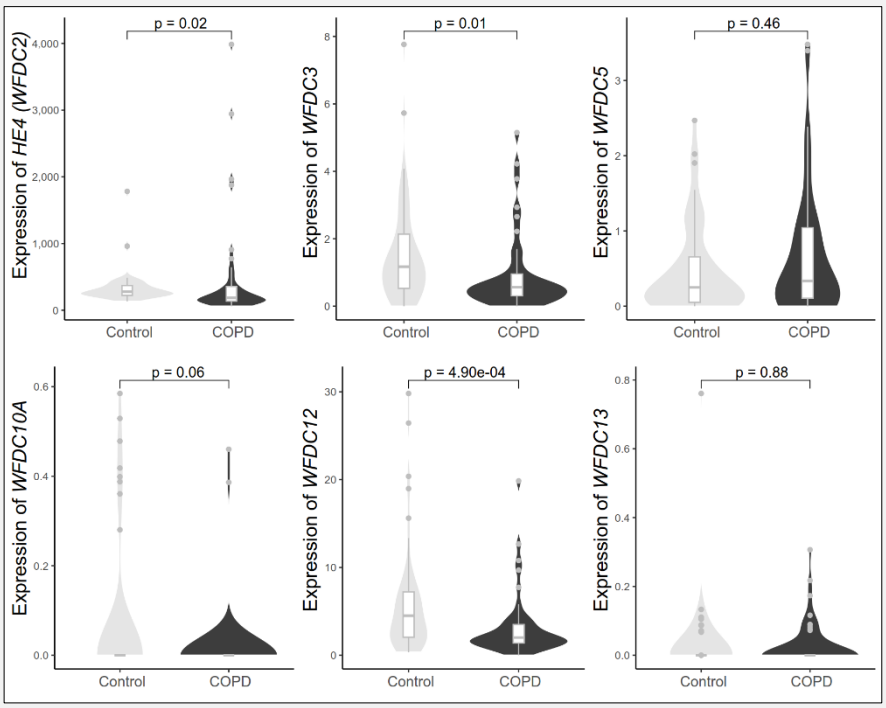


Figure 3. Gene expression profiles of six WFDC protein family members in lung tissue samples from COPD patients and normal controls.

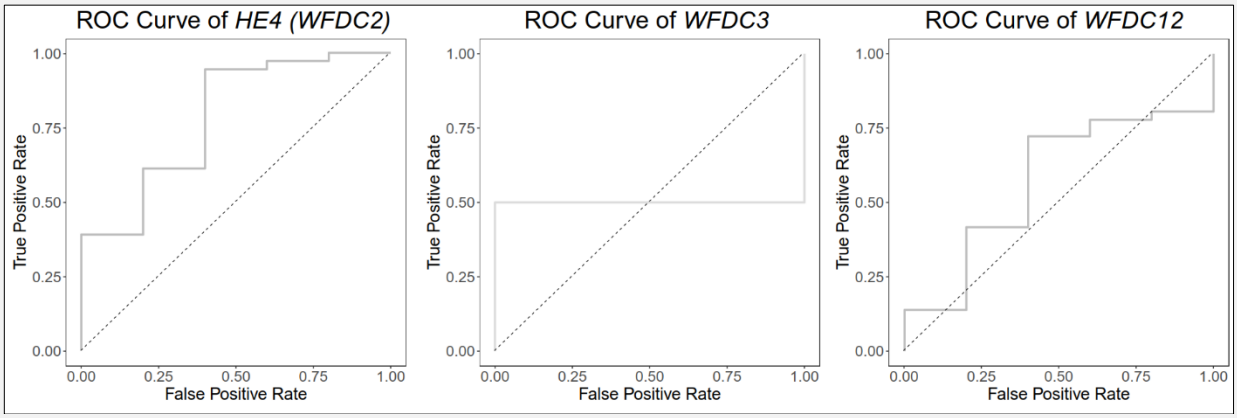


Figure 4. ROC curve analysis of the diagnostic efficacy of *HE4* (*WFDC2*), *WFDC3*, and *WFDC12* for COPD.

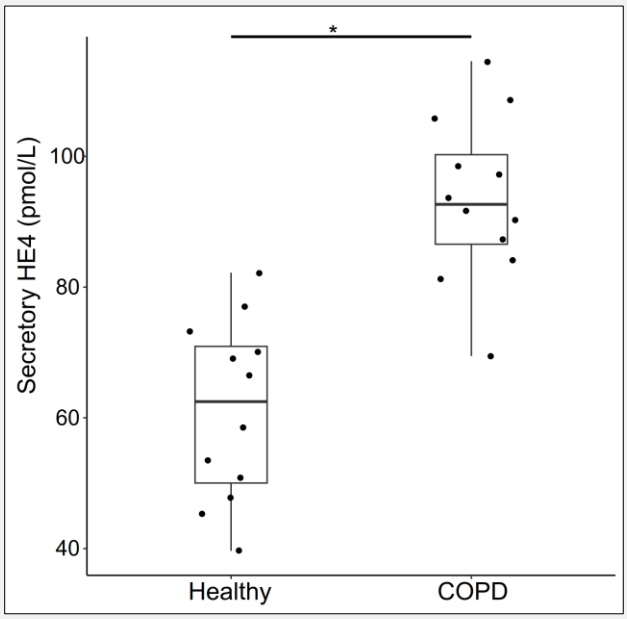


Figure 5. Secretory HE4 expression in the plasma of COPD patients and healthy subjects.

of significant scientific importance. In this study, we did not observe a causal relationship between the other five WFDC protein family members and the development of COPD, aside from HE4. This may be attributed to HE4's involvement in mechanisms beyond its protease inhibitory activity. It has been shown that cigarette smoke, a major exogenous source of reactive oxygen species, induces oxidative stress that increases HE4 secretion in bronchial epithelium, resulting in elevated HE4 expression. It is hypothesized that HE4 may exacerbate airway inflammation and contribute to COPD pathogenesis by promoting the release of inflammatory

cytokines through oxidative stress and inflammatory pathways [23].

Our findings demonstrated that HE4 is highly expressed in plasma, and *HE4* is significantly upregulated in the lung tissues of COPD patients, consistent with the results of Zhan et al. [23]. However, these findings contradict those of Zhu et al. [24], which may be attributed to differences in the methodologies used for detection. In our study, human lung tissue samples were used, which aligns with the study by Zhan et al. [23]. Additionally, two other genes of WFDC protein family members, *WFDC3* and *WFDC12*, showed differential expression in COPD patients' lung tissues compared to normal lung tissues. However, to date, no studies have explored the association between *WFDC3* or *WFDC12* and COPD. ROC curve analysis revealed that *HE4*, among the three WFDC proteins, had a significant diagnostic efficacy for COPD, with an AUC of 0.78, while *WFDC3* and *WFDC12* showed AUC values of 0.50 and 0.57, respectively.

This study has several limitations. First, the data used in this study were derived from European populations, and due to inherent population stratification in MR analysis, the generalizability of our findings to other populations may be limited. Second, while the GWAS dataset employed in this study includes a large sample size, it remains insufficiently comprehensive, and the reliance on protein quantitative trait loci data as proxies for the expression levels of WFDC protein family members may not fully capture tissue-specific expression patterns. Third, this study consists solely of bioinformatics analysis of public databases which are not accompanied by validation. Lastly, with ongoing research, HE4's diagnostic value in other respiratory diseases has been identified [25], raising concerns about its specificity as a biomarker. Furthermore, other studies have indicated that age is a significant factor influencing HE4 levels [26], suggesting that setting population-specific reference ranges will enhance the clinical application of HE4.

In conclusion, our study demonstrates that elevated HE4 expression may increase the risk of COPD and provides further insight into the complex relationship between HE4 and COPD. Additionally, we found that *HE4* is highly expressed in the lung tissues of COPD patients and exhibits excellent diagnostic efficacy, suggesting that HE4 may serve as a potential biomarker for early COPD screening.

Declaration of Interest:

None.

References:

1. Adeloye D, Song P, Zhu Y, Campbell H, Sheikh A, Rudan I. Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis. Lancet Respir Med 2022; 10(5):447-58. (PMID: 35279265)

2. Chen S, Kuhn M, Prettner K, et al. The global economic burden of chronic obstructive pulmonary disease for 204 countries and territories in 2020 - 50: a health-augmented macroeconomic modelling study. Lancet Glob Health 2023;11(8):e1183-e1193. (PMID: 37474226)

3. Wang Z, Lin J, Liang L, et al. Global, regional, and national burden of chronic obstructive pulmonary disease and its attributable risk factors from 1990 to 2021: an analysis for the Global Burden of Disease Study 2021. Respir Res 2025;26(1):2. (PMID: 39748260)

4. Casaburi R, Celli B, Crapo J, et al. The COPD biomarker qualification consortium (CBQC). COPD 2013;10(3):367-77. (PMID: 23713597)

5. Lipson DA, Barnhart F, Brealey N, et al. Once-daily single-inhaler triple versus dual therapy in patients with COPD. N Engl J Med 2018;378(18):1671-180. (PMID: 29668352)

6. Halpin DMG, Criner GJ, Papi A, et al. Global initiative for the diagnosis, management, and prevention of chronic obstructive lung disease. The 2020 GOLD science committee report on COVID-19 and chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2021;203(1):24-36. (PMID: 33146552)

7. Kirchhoff C, Habben I, Ivell R, Krull N. A major human epididymis-specific cDNA encodes a protein with sequence homology to extracellular proteinase inhibitors. Biol Reprod 1991;45(2): 350-7. (PMID: 1686187)

8. Wen Y, Jiang N, Wang Z, Xiao Y. Versatile whey acidic protein four-disulfide core domain proteins: biology and role in diseases. Front Cell Dev Biol 2024;12:1459129. (PMID: 39296934)

9. Sekula P, Del Greco MF, Pattaro C, Kottgen A. Mendelian randomization as an approach to assess causality using observational data. J Am Soc Nephrol 2016;27(11):3253-65. (PMID: 27486138)

10. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. Nature 2018;558(7708):73-9. (PMID: 29875488)

11. Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature 2023;613(7944):508-18. (PMID: 36653562)

12. de Fays C, Geudens V, Gyselinck I, et al. Mucosal immune alterations at the early onset of tissue destruction in chronic obstructive pulmonary disease. Front Immuno 2023;14:1275845. (PMID: 37915582)

13. Sanna S, van Zuydam NR, Mahajan A, et al. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. Nat Genet 2019;51(4):600-5. (PMID: 30778224)

14. Au Yeung SL, Gill D. Standardizing the reporting of Mendelian randomization studies. BMC Med 2023;21(1):187. (PMID: 37198682)

15. Mounier N, Kutalik Z. Bias correction for inverse variance weighting Mendelian randomization. Genet Epidemiol 2023;47 (4):314-31. (PMID: 37036286)

16. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44(2):512-25. (PMID: 26050253)

17. Greco MF, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. Stat Med 2015;34(21):2926-40. (PMID: 25950993)

18. Bowden J, Spiller W, Del Greco MF, et al. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. Int J Epidemiol 2018;47(4):1264-78. (PMID: 29961852)

19. Yang WL, Lu Z, Guo J, et al. Human epididymis protein 4 antigen-autoantibody complexes complement cancer antigen 125 for detecting early-stage ovarian cancer. Cancer 2020;126(4):725-36. (PMID: 31714597)

20. Bingle L, Cross SS, High AS, et al. WFDC2 (HE4): a potential role in the innate immunity of the oral cavity and respiratory tract and the development of adenocarcinomas of the lung. Respir Res 2006;7(1):61. (PMID: 16600032)

21. Nagy B Jr, Nagy B, Fila L, et al. Human Epididymis Protein 4: A Novel Serum Inflammatory Biomarker in Cystic Fibrosis. Chest 2016;150(3):661-72. (PMID: 27105680)

22. Wang W, Mei A, Qian H, et al. The role of glucagon-like peptide-1 receptor agonists in chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis 2023;18:129-37. (PMID: 36815056)

23. Zhan Y, Chen J, Wu J, et al. Human epididymis protein 4 aggravates airway inflammation and remodeling in chronic obstructive pulmonary disease. Respir Res 2022;23(1):120. (PMID: 35550579)

24. Zhu XH, Jia M, Xu JY, Meng YQ, Qiao YY, Yao X. The expression and meaning of HE4 in chronic obstructive pulmonary disease. Int J Respir. 2019;39(18):1391-1396. https://rs.yiigle.com/cmaid/1164393

25. Luo J, Liang J, Wang S, et al. Serum human epididymis secretory protein 4 correlates with sepsis-associated acute respiratory distress syndrome and 28-day mortality in critically ill patients. Ann Clin Biochem 2022;59(5):338-46. (PMID: 35549539)

26. Bolstad N, Oijordsbakken M, Nustad K, Bjerner J. Human epididymis protein 4 reference limits and natural variation in a Nordic reference population. Tumour Biol 2012;33(1):141-8. (PMID: 22105734)