



## Integrative RNA-sequencing analysis of COPD-related genes in association with individual PM<sub>2.5</sub> exposure

Jeeyoung Kim <sup>a</sup>, Ha Won Song <sup>a</sup>, Hyun Woo Lee <sup>b</sup>, Ye Jin Lee <sup>c</sup>, Sooim Sin <sup>d</sup>, Ji Yeon Lee <sup>d</sup>, Junghyun Kim <sup>e</sup>, Sun Mi Choi <sup>f</sup>, Kyoung-Nam Kim <sup>g</sup>, Chang-Hoon Lee <sup>f</sup>, Chang Hyun Lee <sup>h</sup>, Woo Jin Kim <sup>a,\*</sup>

<sup>a</sup> Department of Internal Medicine and Environmental Health Center, Kangwon National University Hospital, Kangwon National University, School of Medicine, Chuncheon, 24341, Republic of Korea

<sup>b</sup> Division of Respiratory and Critical Care, Department of Internal Medicine, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul National University, College of Medicine, Seoul, 07061, Republic of Korea

<sup>c</sup> Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Seoul National University Bundang Hospital, Seongnam, 13620, Republic of Korea

<sup>d</sup> Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, National Medical Center, Seoul, 04564, Republic of Korea

<sup>e</sup> Division of Pulmonary, Allergy and Critical Care Medicine, Department of Internal Medicine, Hallym University Dongtan Sacred Heart Hospital, Hallym University, College of Medicine, Hwaseong 18450, Republic of Korea

<sup>f</sup> Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Seoul National University Hospital, Seoul National University, College of Medicine, Seoul, 03080, Republic of Korea

<sup>g</sup> Department of Preventive Medicine, College of Medicine, Yonsei University, Seoul, 03722, Republic of Korea

<sup>h</sup> Department of Radiology and Institute of Radiation, Seoul National University Hospital, Seoul National University, College of Medicine, Seoul, 03080, Republic of Korea

### ARTICLE INFO

#### Keywords:

PM<sub>2.5</sub>  
COPD  
RNA sequencing  
Lung function  
Biomarker

### ABSTRACT

**Background:** Airborne fine particulate matter (PM<sub>2.5</sub>) is associated with chronic obstructive pulmonary disease (COPD); however, the precise mechanism remains unclear. Here, we examined distinct gene and pathway characteristics across varying personal and ambient PM<sub>2.5</sub> exposure durations within a prospective COPD cohort and the associations between differentially expressed genes (DEGs) and clinical phenotypes.

**Methods:** Blood samples for RNA-sequencing were collected from 50 patients with COPD who underwent spirometry and quantitative computed tomography. We estimated personal and ambient PM<sub>2.5</sub> exposure levels using hybrid and land use regression models. Associations between DEGs and PM<sub>2.5</sub> exposure were examined in relation to lung function indicators (FEV<sub>1</sub>, FVC, and FEV<sub>1</sub>/FVC ratio) using Pearson correlation analysis adjusted for factors such as hospitalization, age, sex, season, Charlson Comorbidity Index score, and smoking status.

**Results:** We analyzed DEGs across three cumulative PM<sub>2.5</sub> exposure periods using personal and ambient exposure assessments. Gene ontology annotation and biological pathway analysis of the identified DEGs using the individual air pollution exposure prediction model revealed significant associations with gas transport, cellular processes related to cell cycle, cell proliferations, and neuron projection morphogenesis. The ambient air pollution prediction model revealed significant biological responses related to purine metabolism and antigen processing and presentation. *EDAR*, *NKILA*, *HSD11B2*, *LOC100130027*, *LOC105378367*, *SENCR*, *CAMP*, *CEACAM6*, *CHIT1*, *EREG*, *HSD17B3*, *NPPA-AS1*, and *TRPV4* showed increased expression with higher PM<sub>2.5</sub>, correlating with reduced lung function.

**Conclusions:** Our findings offer insights into the role of gene expression in patients with COPD in response to personal and ambient PM<sub>2.5</sub> exposure, suggesting strategies to enhance respiratory conditions linked to air pollution.

\* Corresponding author. Department of Internal Medicine, Kangwon National University, School of Medicine, 1 Gangwondaehak-gil, Chuncheon, 24341, Republic of Korea

E-mail address: [pulmo2@kangwon.ac.kr](mailto:pulmo2@kangwon.ac.kr) (W.J. Kim).

## 1. Background

Recent estimates show a 66 % increase in the mortality rate linked to air pollution over the last two decades. This increase is attributed to various factors, such as increased industrialization and climate change (Fuller et al., 2022). Owing to its anatomical location and function, the respiratory system is continuously exposed to harmful air pollutants, making it the primary target of their toxic effects (Wallbanks et al., 2024). Although the connection between air pollutants and respiratory illnesses is not fully understood, researchers have observed that exposure to inhaled pollutants, such as particulate matter (PM) and nitrogen oxides (NOx), triggers a series of harmful biological responses (Losacco and Perillo, 2018; Newman et al., 2020). These responses involve inflammatory and oxidative stress pathways, which negatively impact the development and worsening of respiratory diseases.

The peripheral blood transcriptome offers insights into the physiological and pathological changes occurring in various tissues throughout the body (Pickrell et al., 2010). Analyzing changes in the transcriptome and associated signaling pathways provides a more effective way to understand the biological mechanisms underlying air pollution-related chronic obstructive pulmonary disease (COPD) (Sin et al., 2023). Several studies have connected air pollution exposure to alterations in blood transcriptomic profiles (Yao et al., 2021). The physiological mechanisms underlying COPD, including inflammatory imbalance, reactive oxygen species signaling, mitochondrial genome maintenance, and apoptotic processes, have been widely investigated (Chen et al., 2020; Yao et al., 2021). These mechanisms have been studied in relation to various COPD phenotypes, particularly lung function and injury. A more in-depth exploration of these biological mechanisms is crucial to clarify the respiratory health impacts of air pollution.

The role of air pollution in COPD is a subject of debate due to limitations in epidemiological design and overlapping risk factors impacting respiratory health. Inconsistencies in findings on transcriptomic changes associated with air pollution are attributed to methodological differences (Kakati et al., 2019), exposure terms (Merid et al., 2021; Vlaanderen et al., 2022), race (Kelchtermans et al., 2025), and geographical variations (Favé et al., 2018). The complex relationship between air pollution and human activities presents challenges in accurately evaluating individual exposure to air pollutants. To effectively measure PM<sub>2.5</sub> exposure, spatial and temporal factors must be carefully considered in personal exposure monitoring (Evangelopoulos et al., 2021; Steinle et al., 2015). Many studies investigating the link between air pollution and COPD have primarily used spatial modeling methods to estimate PM<sub>2.5</sub> exposure levels based on data from fixed outdoor sampling sites (Ma et al., 2024). While efforts have been made to understand the spatial and temporal distribution of PM<sub>2.5</sub>, enhancing tools and information for modeling PM<sub>2.5</sub> exposure across diverse microenvironments encountered in daily life is crucial for COPD research.

We aimed to investigate the relationship between PM<sub>2.5</sub> exposure levels and peripheral blood transcriptome responses through personal and ambient exposure to PM<sub>2.5</sub> assessment methods in a prospective COPD cohort. This builds on previous research by identifying differentially expressed genes (DEGs) and signaling pathways affected by PM<sub>2.5</sub> across short-, medium-, and long-term exposure cycles (Ji et al., 2024; Rundblad et al., 2025; Yao et al., 2021). Additionally, it explores the potential impact of these genetic changes on respiratory disease by examining specific genes linked to clinical phenotypic changes in COPD. This study emphasizes the intricate genetic implications of PM<sub>2.5</sub> exposure in patients with COPD and contributes to a deeper understanding of this multifaceted relationship.

## 2. Methods

### 2.1. Study populations

This prospective cohort study utilized data and biological specimens

collected between 2019 and 2020 from COPD participants in a prospective cohort study (NCT03813810). The study was conducted in five hospitals in South Korea: Seoul National University Hospital (SN), National Medical Center (NM), Seoul Metropolitan Government-Seoul National University Boramae Medical Center (BR), Seoul National University Bundang Hospital (SB), and Kangwon National University Hospital (KW). The primary objective was to investigate the effects of air pollution on the clinical outcomes of chronic respiratory diseases and identify novel genetic, chemical, and imaging biomarkers. The study included four participant categories: normal individuals, individuals with asthma, individuals with idiopathic pulmonary fibrosis, and individuals with COPD. Participants with conditions potentially affecting lung function decline, such as severe bronchiectasis, tuberculous lung destruction, pleural thickening, significant spinal abnormalities, pleural effusion, pneumothorax, or suspected lung cancer as indicated by chest X-ray, were excluded. COPD was defined as a post-bronchodilator ratio of forced expiratory volume in 1 s to forced vital capacity (FEV<sub>1</sub>/FVC) of less than 0.7. At baseline and the one-year follow-up, participants underwent inspiratory-expiratory computed tomography, blood tests, and pulmonary function tests. Air pollutant exposure was measured using a portable device for 24 h at baseline, followed by measurements twice at 3, 6, and 9 months, and again at 1 year. The study received approval from the Seoul National University Hospital Institutional Review Board (SNUH, 1810-036-977) in compliance with the Declaration of Helsinki. Written consent was obtained from all participants upon enrollment in the cohort.

### 2.2. Exposure assessment

This study evaluated individual exposure by measuring PM<sub>2.5</sub> levels in participants using personal monitors and predicting air pollution concentrations at each location through geocoding based on participant addresses.

#### 2.2.1. Measure of individual air pollution exposure using direct-reading device

We conducted a study to evaluate exposure to PM<sub>2.5</sub> utilizing a direct-reading device that they carried continuously for 24 h. During the sampling program, participants wore a personal monitor to measure PM<sub>2.5</sub> at each follow-up visit every 3 months from January 2019 to December 2020, using the Aerocet 831 Handheld direct-reading instrument from Met One Instruments. The instruments automatically recorded data at 1-min intervals for 22–24 h, and we subsequently downloaded the data. Measured PM<sub>2.5</sub> exposure concentrations for participants were consistently collected, with Direct exposure levels quantified at each subject's time to calculate the 24-h average. The final calculation for individual PM<sub>2.5</sub> concentration exposure during the study period is the average of overall PM<sub>2.5</sub> concentrations.

#### 2.2.2. Time-activity diaries

All participants were instructed to complete a time-activity diary as part of the questionnaire. This diary served to collect data during the monitoring period, with activity logs recorded on an hourly basis. Activities were categorized into four distinct classifications: indoor, other indoor spaces, outdoor activities, and specific detailed activities (such as walking, cleaning, and smoking indoors). The individual time-activity data collected in the hourly diaries were subsequently analyzed to discern patterns of indoor and outdoor occupancy for each participant. This analysis facilitated the determination of both hourly and daily exposure levels to PM<sub>2.5</sub> for each individual, differentiating between indoor and outdoor environments.

#### 2.2.3. LUR model for predicting ambient air pollution

The air quality prediction model was utilized to calculate the predicted value for the daily to annual average concentration of PM<sub>2.5</sub> based on the participant's residential address. These LUR models

utilized a Near function analysis tool of ArcGIS and incorporated geographic predictors and spatial correlations using regulatory monitoring data from 2015 to 2020 in South Korea. Geographic variables, including traffic, demographics, land use, physical geography, transportation infrastructure, emission levels, vegetation, and altitude, were utilized to derive predictors through partial least squares analysis. The model's accuracy was assessed through cross-validation, yielding a cross-validated  $R^2$  value of 0.734 for PM<sub>2.5</sub>. By integrating this predictive model with participants' address data at the district level, the average concentration for each daily to annual factor was estimated in a district representative of the population.

#### 2.2.4. Hybrid model for predicting individual air pollution exposure prediction

We are employed to estimate unmeasured individual PM<sub>2.5</sub> exposure using a LUR model, a widely used method for exposure estimation. The methods for estimating PM<sub>2.5</sub> forecasts have been previously described (Ji et al., 2024; Liu et al., 2019). To quantify microenvironmental PM<sub>2.5</sub> concentration distributions and patient exposures, we employed the Stochastic Human Exposure and Dose Simulation model, which incorporates time–activity diaries alongside real-time data on personal indoor and outdoor activities, including residence times. PM<sub>2.5</sub> exposure levels were calculated using a 12-month time-weighted exposure equation, where the temporal framework is articulated in days:

$$Y_{\text{air concentrations}} = (\text{Chin} \times \text{Thin} + \text{Chout} \times \text{Thout} + \text{Cout} \times \text{Tout} + \text{IFh} \\ \times (\text{Cout} \times \text{Tin}) + \text{Ttrans} \times \text{IFT} \times (\text{Chout} + \text{Cout}) / 2) / 24$$

Here, the variable  $Y_{\text{air}}$  corresponds to the time-weighted air pollution concentrations specific to each subject. Indoor (Chin) and outdoor (Chout) PM<sub>2.5</sub> concentrations were derived from direct measurements, while Cout denotes the outdoor PM<sub>2.5</sub> levels predicted by the LUR model for areas without direct monitoring. Thin and Thout indicate the time allocated by participants in indoor and outdoor environments, respectively. Tin and Tout reflect the durations spent inside and outside various locations. The term Ttrans accounts for the average daily time dedicated to transportation and mobility. The infiltration factors IFh and IFT represent the PM<sub>2.5</sub> penetration coefficients into indoor spaces and the variations linked to transportation modes, respectively. The expression (Chout + Cwout)/2 accounts for the mean PM<sub>2.5</sub> concentration in residential and other outdoor locales utilized in this investigation to represent fine dust levels resulting from transportation and mobility. This average was integrated into the LUR model to accurately estimate indoor concentrations. By considering the distribution of time spent both indoors and outdoors, we calculated individual fine dust exposure for the entire study period.

#### 2.3. RNA sequencing and data processing

Blood sampling and RNA extraction were conducted as previously outlined (ISO, 2019). Whole blood samples were obtained from each hospital following approved guidelines and relevant regulations. Prior to study participation, all individuals provided written informed consent. A 2.5 mL whole blood sample from each participant was placed in PAXgene Blood RNA tubes (BD, Cat. No. 762165) and stored at -80 °C. RNA isolation was performed according to the manufacturer's instructions provided with the PAXgene Blood RNA kit v2 (Qiagen, Cat. No. 762174). Total RNA concentration was determined using Quant-IT RiboGreen (Invitrogen, #R11490), and RNA integrity was evaluated using the TapeStation RNA ScreenTape (Agilent, #5067-5576). Only RNA samples with an RNA integrity number (RIN) > 7.0 were utilized for library preparation. For each sample, 0.5 µg of total RNA was processed with the Illumina TruSeq Stranded Total RNA Library Prep Globin Kit (Illumina, Inc., San Diego, CA, USA, #20020613). The rRNA was removed, and the remaining mRNA was fragmented via divalent cations at elevated temperatures. Subsequently, first-strand cDNA was

synthesized using SuperScript II reverse transcriptase (Invitrogen, #18064014) and random primers, followed by second-strand cDNA synthesis using DNA Polymerase I, RNase H, and dUTP. The cDNA underwent end repair, A-base addition, adapter ligation, purification, and PCR enrichment. Libraries were quantified using KAPA Library Quantification kits (KAPA BIOSYSTEMS, #KK4854) and assessed for quality using TapeStation D1000 ScreenTape (Agilent Technologies, #5067-5582). Indexed libraries were then processed for paired-end sequencing (2 × 100 bp) on an Illumina NovaSeq system (Macrogen Inc., San, USA). The quality of FASTQ files from RNA sequencing was initially assessed using FastQC (version 0.11.7) (S-Andrews, 2010). Adapter sequences were removed using cutadapt (version 3.3) (Martin, 2011), and reads with quality scores below 20 and short reads <80 bp were excluded. For read mapping and subsequent quantification of transcript abundance, we used the STAR-RSEM pipeline (STAR version 2.6.0 (Dobin et al., 2013) and RSEM version 1.3.1 (Li and Dewey, 2011)) with these parameters. The GRCh38.p13 hg38 human genome from GENCODE was used as the reference genome.

#### 2.4. Statistical methods

After quality control processing, the obtained RPKM values were utilized for analysis. For DEG analysis, we employed the DESeq2 R package (version 1.44.0). The samples were divided into high- and low-exposure groups based on the median of each criterion. The covariates included the hospital of sample collection, sex, age, BMI, Charlson comorbidity index (CCI), smoking status, season, temperature, and humidity. The regression formula used was: expression ~ hospital of sample collection + sex + age + BMI + CCI + smoking status + season + temperature + humidity. Genes with a raw  $p$ -value <0.01 and an absolute fold change >2 were considered DEGs with significant differences. To understand the distribution of expression patterns among patients, we calculated the principal components using the principal component analysis (PCA) function from the FactoMineR (version 2.9) (Lê et al., 2008) R package. Additionally, to quantitatively evaluate the distribution of patients, we conducted a PERMANOVA with the adonis2 command from the vegan R package (version 2.6.8), using Euclidean distance as the method. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed using the Metascape web tool (version 3.5) (Metascape, 2024) and the R package clusterProfiler (version 4.12.5). The selected DEGs were compared against the GO framework genes and curated using the KEGG database. The plotted terms represent the cases with  $p$ -values <0.05 for GO and KEGG. Finally, the correlation between clinical phenotype (FEV<sub>1</sub>, FVC, FEV<sub>1</sub>/FVC) and DEG expression was calculated using Pearson's correlation coefficient with the cor.test functions in the stats library (version 4.4.1) after adjusting with sex, age, BMI, CCI, smoking status, season, temperature, humidity, and hospital of sample collection data. Covariate adjustments were performed using the lm(.) function.

### 3. Results

#### 3.1. Participant characteristics

Only 50 of the 64 participants with bulk RNA sequencing data were included in the analysis. Participants were excluded due to missing information ( $n = 12$ ) or acute exacerbations ( $n = 2$ ) from the covariates in the main analysis. Table 1 presents the participants' characteristics. The participants had an average age of 68.0 years (SD = 7.0), and 86 % were male. Among them, 18.0 % were current smokers, 68.0 % were former smokers, and 14 % were non-smokers. The body mass index was 23.57 (SD = 2.61). The seasonal enrollment distribution of study participants was 16, 13, 14, and 7 for spring, summer, fall, and winter, respectively, which influenced PM<sub>2.5</sub> exposure and respiratory health outcomes. The CCI, excluding COPD, was reported for 22 participants. Among them, CCI scores >0–4 were distributed as follows: 12 (24 %) with a score of 1,

**Table 1**  
Participant and clinical data.

	* <sup>1</sup> Mean (SD) or n (%)
<b>Subject (N = 50<sup>1</sup>)</b>	
Age	68 (7)
Sex	
Female	7 (14 %)
Male	43 (86 %)
BMI	23.57 (2.61)
Smoke	
Non-smoker	7 (14 %)
Former smoker	34 (68 %)
Smoker	9 (18 %)
CCI	
0	12 (24 %)
1	22 (44 %)
2	10 (20 %)
3	2 (4.0 %)
4	4 (8.0 %)
Season	
Spring	16 (32 %)
Summer	13 (26 %)
Autumn	14 (28 %)
Winter	7 (14 %)
Hospital	
BR	6 (12 %)
KW	7 (14 %)
NM	9 (18 %)
SN	8 (16 %)
SN	24 (40 %)
Lung function	
FEV <sub>1</sub> (%)	61 (16)
FVC (%)	83 (22)
FEV <sub>1</sub> /FVC (%)	50 (12)

Note: \*Data are presented as n or mean ± SD.

**Abbreviations:** BMI, Body Mass Index; CCI, Charlson Comorbidity Index; **Hospital:** SN, Seoul National University Hospital; NM, National Medical Center; BR, Seoul Metropolitan Government-Seoul National University Boramae Medical Center; SB, Seoul National University Bundang Hospital; KW, Kangwon National University Hospital; **Lung function:** FEV<sub>1</sub>, Forced Expiratory Volume in 1 s; FVC, Forced Vital Capacity.

22 (44 %) with a score of 2, 10 (20 %) with a score of 3, 2 (4 %) with a score of 4, and 4 (8 %) with a score above 4. The predicted mean values of FEV<sub>1</sub> (%), FVC (%), and FEV<sub>1</sub>/FVC after bronchodilator administration were 61, 83, and 50 % (SD = 16, 22, 12), respectively.

### 3.2. Assessment of PM<sub>2.5</sub> exposure levels through personal evaluations

This study evaluated participants' exposure to PM<sub>2.5</sub>, utilizing both direct and indirect assessment methodologies. Direct exposure levels were quantified using a portable device, with subjects measuring their PM<sub>2.5</sub> levels during quarterly hospital visits over the course of a year. The mean individual PM<sub>2.5</sub> exposure recorded via this direct method was 27.22 µg/m<sup>3</sup>. Baseline exposure levels for the 3-, 6-, 9-, and 12-month intervals were 23.50, 34.34, 35.37, 22.19 and 25.13 µg/m<sup>3</sup>, respectively (Table S1). Subsequently, we evaluated PM<sub>2.5</sub> exposure concentrations using personal and ambient methods to confirm exposure during specific periods (Fig. 1A). Changes in PM<sub>2.5</sub> concentration were categorized into short, medium, and long terms to assess the effects of exposure over different periods. The average concentrations at 7 d, 90 d, and 1 year before blood sampling were verified. The average PM<sub>2.5</sub> concentration over one year represented the yearly average for the year of blood collection from the patient (e.g., May 2019 = average for 2019, March 2020 = average for 2020). Table 2 presents the median (IQR) distributions of PM<sub>2.5</sub> concentrations by exposure period for each model. When the hybrid model was used, the estimated PM<sub>2.5</sub> exposure levels were 22.11 µg/m<sup>3</sup> for short-term, 24.01 µg/m<sup>3</sup> for medium-term, and 25.49 µg/m<sup>3</sup> for long-term exposure; when the LUR model was used, the ambient PM<sub>2.5</sub> exposure levels were predicted to be 20.42, 21.54, and

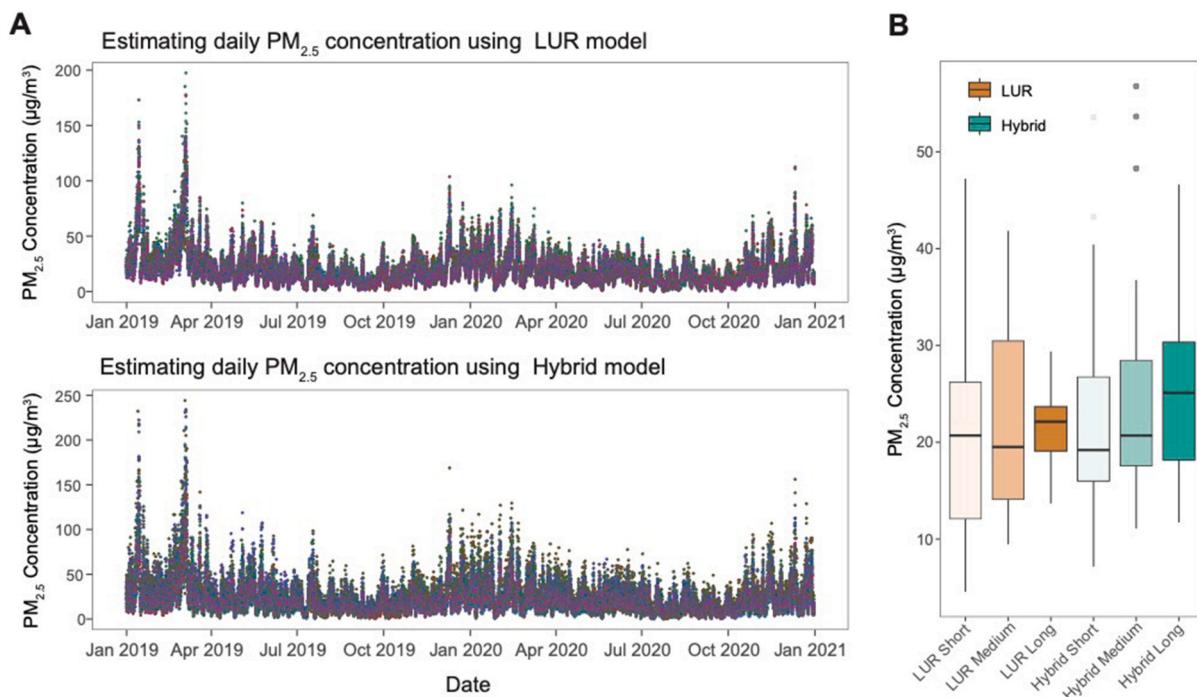
21.59 µg/m<sup>3</sup>, respectively (Fig. 1B). The hybrid model yielded slightly higher daily averages compared to the LUR model. A significant correlation was observed for short- and medium-term exposure, but not for long-term exposure. Both models recorded daily average PM<sub>2.5</sub> concentrations exceeding the WHO AQG level of 5 µg/m<sup>3</sup>.

### 3.3. Association between DEGs and PM<sub>2.5</sub> exposure

We conducted bulk RNA sequencing on the collected blood samples to analyze gene expression patterns. Subsequently, we divided the patients into high- and low-exposure groups based on their PM<sub>2.5</sub> exposure levels. We identified DEGs between these groups (Fig. 2A). The median value for each criterion was utilized to categorize patients while adjusting for potential confounding variables, such as the sampling hospital, sex, age, BMI, CCI score, smoking status, temperature, humidity and sampling season. Using the LUR model with a *p* < 0.01 threshold, we identified 39, 90, and 40 DEGs for short-term, mid-term, and long-term PM<sub>2.5</sub> exposure, respectively. For short-term exposure, there were 17 up-regulated and 22 down-regulated genes; mid-term exposure showed 14 up-regulated and 76 down-regulated genes; and long-term exposure identified 21 up-regulated and 24 down-regulated genes. Applying the hybrid model also revealed 39 DEGs across all exposure periods under the same *p* < 0.01 criterion. Using this model, short-term exposure exhibited 14 up-regulated and 25 down-regulated genes, mid-term showed 21 up-regulated and 18 down-regulated genes, and long-term exposure revealed 14 up-regulated and 25 down-regulated genes. Importantly, using an FDR criterion of <0.05, no genes were found to be differentially expressed based on individual and ambient PM<sub>2.5</sub> exposure levels. The up- and down-regulated genes corresponding to the identified DEGs by PM<sub>2.5</sub> exposure period in each model can be found in Table S4. The PCA results, derived from the DEG-based exposure assessment, revealed significant differences between the high- and low-exposure groups across the three exposure windows, except for mid-term exposure in the LUR model (Fig. 2B). Interestingly, for the hybrid-based personal exposure estimates, the analysis of gene expression across different exposure windows, seven genes were found to overlap after short-to mid-term exposure: *FAM83A*, *TEK*, *C4BPA*, *COL6A4P2*, *CSMD1*, *CYP3A7-CYP3A51P*, and *PCLO*. Furthermore, two genes, *GRIK1-AS2* and *EGR3*, were found to overlap after med-to long-term exposure. On examining the LUR-based ambient exposure estimates, five genes exhibited a significant overlap for short-to medium-term exposure: *CRYGN*, *HKDC1*, *CCDC144A*, *TM4SF1*, and *TMEM51-AS1*. Additionally, two genes, *GLYATL1B* and *LINC01106* were significantly differentially expressed after medium-to long-term exposure. Notably, none of the genes overlapped across the three temporal categories of exposure. However, two genes, *MOCOS* (short-term) and *CCR10* (mid-term), were significantly affected by both personal and ambient PM<sub>2.5</sub>, indicating a potential difference in the biological responses induced by short-to medium-term versus long-term exposures. This divergence suggests that variable exposure durations may lead to the activation of distinct molecular pathways. Identified DEGs that secured statistical significance as a cutoff criterion based on *p*-value (*p* < 0.01) and log fold change ( $|log_2| > 1$ ) by the PM<sub>2.5</sub> exposure method and period were selected for further analysis.

### 3.4. GO terms and biological pathway associated with PM<sub>2.5</sub> exposure

Furthermore, we conducted GO analysis to evaluate the functional changes in genes linked to PM<sub>2.5</sub>, as shown in Fig. 3A and detailed in Table 3. For the hybrid model, short-term exposure to PM<sub>2.5</sub> was associated with changes in gas transport, including carbon dioxide transport and nitric oxide transport. In contrast, medium-term exposure mainly affected blood vessel development and regulated exocytosis. Long-term exposure was associated with cell proliferation and neuron projection morphogenesis. The LUR model identified significant gene expression changes in response to short- and mid-term PM<sub>2.5</sub> exposure, primarily



**Fig. 1.** The distribution of estimated PM<sub>2.5</sub> exposure concentrations in all participants during follow-up (Jan 2019 to Jan 2021). **A.** PM<sub>2.5</sub> exposure levels based on individual and ambient assessment methods. **B.** Boxplot showing the distribution of the PM<sub>2.5</sub> concentration ( $\mu\text{g}/\text{m}^3$ ) to which participants were exposed on each day preceding sample collection.

**Table 2**

Predicted individual and ambient concentrations ( $\mu\text{g}/\text{m}^3$ ) over short-, mid-, and long-term PM<sub>2.5</sub> exposure for 50 patients with COPD (2019–2020) across exposure models.

Model	Time window	Mean(SD)	Min	Median	Max
Hybrid	Short	22.11 (10)	7.15	19.20	53.58
	Medium	24.01 (10)	11.09	20.69	56.77
	Long	25.49 (8)	11.72	25.12	46.63
LUR	Short	20.42 (10)	4.56	20.70	47.23
	Medium	21.54 (9)	9.47	19.52	41.85
	Long	21.59 (4)	13.68	22.13	29.37
Temperature	Short	16.76 (8)	-1.30	17.79	28.46
	Medium	15.31 (8)	1.70	16.67	25.87
	Long	13.34 (1)	9.76	13.41	14.60
Humidity	Short	63.82 (13)	43.00	64.00	90.71
	Medium	63.25 (9)	48.83	63.98	89.58
	Long	63.10 (5)	56.21	61.97	76.22

**Abbreviations:** PM<sub>2.5</sub>, Particulate Matter less than 2.5  $\mu\text{m}$  in diameter; COPD, Chronic Obstructive Pulmonary Disease. Hybrid model, Individual PM<sub>2.5</sub> exposure concentrations by combining data from a direct-reading device, 24-h time-activity diary along with hourly and daily through LUR modeling; LUR, Land Used Regression model; SD, Standard Deviation; Range: Min, Minimum; Max, Maximum.

involving mitotic cell cycle, nucleoside phosphate metabolic process and purine ribonucleotide catabolic processes. However, the DEGs associated with long-term exposure did not exhibit any significant GO terms.

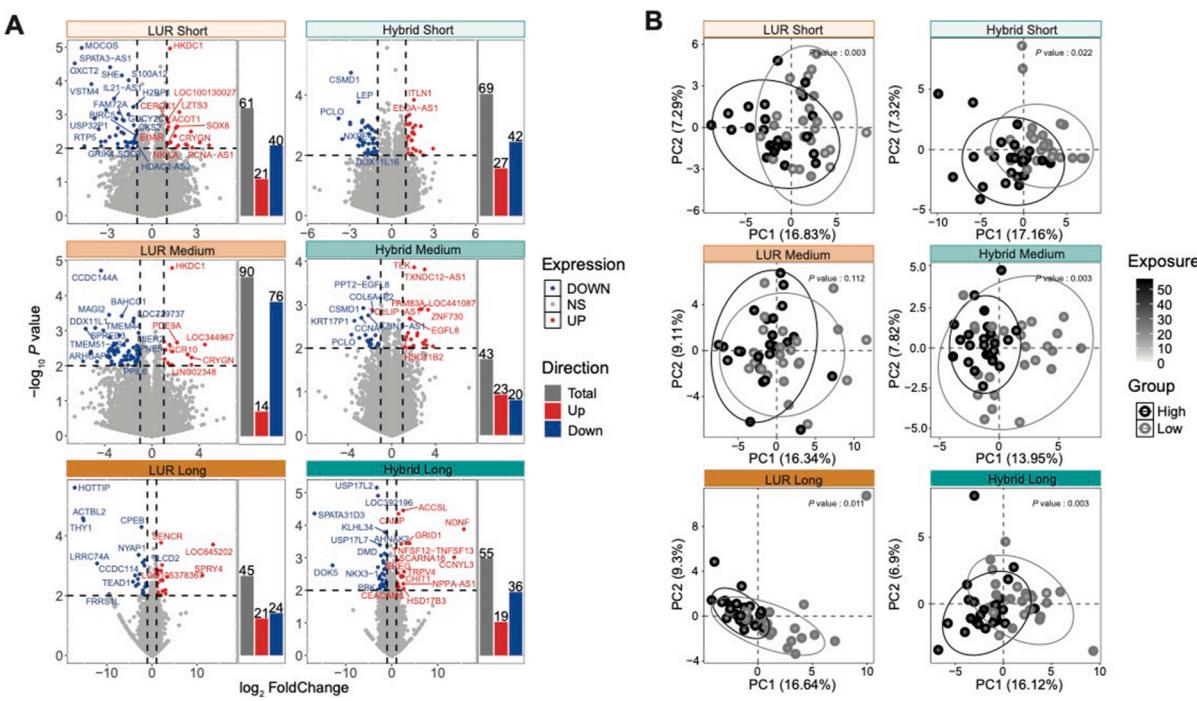
Furthermore, KEGG pathway analysis was conducted to identify the biological pathways linked to these genes in the context of air pollution exposure, as shown in Fig. 3B. In the hybrid model, the pathways were associated with three cumulative exposure periods of PM<sub>2.5</sub>, including the viral infection, cell cycle, and c-type lectin receptor pathways related to immune response. In contrast, the LUR model emphasized pathways related to purine metabolism, antigen processing and presentation. The full list of the GO terms and KEGG pathways associated with PM<sub>2.5</sub> exposure is available in Tables S5–S6.

### 3.5. Altered gene expression associated with COPD phenotype

Furthermore, we investigated the correlation between DEGs and lung function indices, focusing on the exposure methods and time points. Our analysis revealed a significant correlation between long-term PM<sub>2.5</sub> exposure in the hybrid model and FVC(%), indicating potential areas for further investigation (Fig. 4A). However, it is important to acknowledge that most other instances did not show significant correlations (Fig. 4B and C). By examining the DEGs associated with PM<sub>2.5</sub> exposure concentrations, we identified several genes that demonstrated significant correlations with lung function indicators (Fig. 4D and Table 4). Table S7 provides a complete list of identified genes relevant to lung function indices and PM<sub>2.5</sub> exposure, offering further opportunities for research and intervention strategies. Specifically, 15 genes, including EDAR, NKILA, HSD11B2, LOC100130027, LOC105378367, SENCR, CAMP, CEACAM6, CHIT1, EREG, HSD17B3, NPPA-AS1, and TRPV4, were found to have increased expression levels in response to higher PM<sub>2.5</sub> concentrations, correlating with a decline in lung function. Conversely, we observed that the expression of 27 genes, including BIRC5, CKS2, GUCY1A2, IL21-AS1, S100A12, SDC3, ARHGAP8, ACTBL2, FRRS1L, LRRC74A, TREAD1, THY1, C1orf54, DOK5, PBK, GRIK4, SPRED3, FAM72A, HDAC-AS2, LGALS2, RTP5, CPNE5, DDX11L1, ACTRT3, DDX11L16, NKX3-1, and USP17L2 decreased upon exposure to PM<sub>2.5</sub>, which was associated with improved lung function outcomes. This analysis provides valuable insights into the relationship between gene expression and lung function in the context of PM<sub>2.5</sub> exposure.

## 4. Discussion

In this study, we conducted a thorough transcriptome analysis to evaluate the effects of short-, medium-, and long-term PM<sub>2.5</sub> exposures in a prospective cohort of patients with COPD. We identified significant genes in peripheral blood associated with PM<sub>2.5</sub> exposure across different durations using two distinct exposure assessment methodologies. These genes are involved in critical biological pathways related to



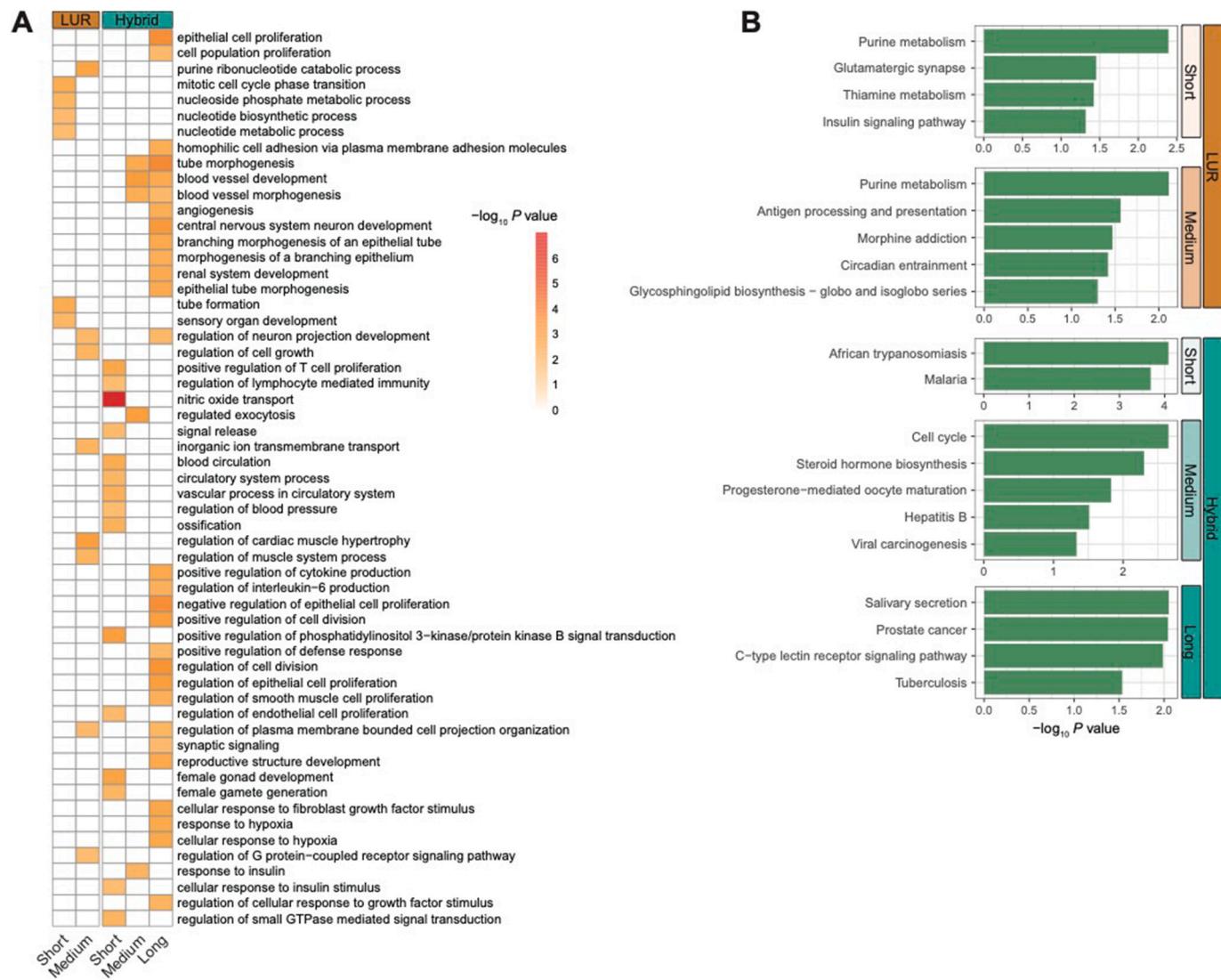
**Fig. 2.** The differentially expressed genes are based on personal and ambient PM<sub>2.5</sub> exposure during specific exposure periods. **A.** Volcano/bar-plot showing the number of DEGs identified based on the selected cutoff for each criterion ( $|\log_2| > 1$ ,  $p\text{-value} < 0.01$ ). Directions represent the expression of up- or down-regulated DEGs. The gray bar represents the total number of DEGs, including up-and down-regulated genes. **B.** PCA plot of identified DEGs. The figure illustrates PM<sub>2.5</sub> exposure levels across different exposure windows for each model, with ovals representing the 95 % confidence intervals for high- and low-exposure groups.

cellular processes dysfunction, inflammation, and immune response, thereby enhancing our understanding of the impact of PM<sub>2.5</sub> on the progression of COPD. Additionally, we analyzed the correlations between PM<sub>2.5</sub> exposure and various lung function indices across the three exposure periods. The genes significantly correlated with PM<sub>2.5</sub> exposure during each exposure period may serve as potential biomarkers, providing insights into the biological mechanisms underlying the development and progression of respiratory diseases. This study highlights the necessity to further investigate these associations to enhance disease management and prevention strategies.

Personal exposure to pollutants, such as particulate matter, can be assessed using direct and indirect methods. Direct methods involve physically measuring pollutant levels at an individual's location or through biomonitoring. Indirect methods use models to estimate exposure based on factors such as air quality and time-activity patterns. We analyzed real-time PM<sub>2.5</sub> exposure direct-reading devices based on RNA sequencing (Fig. S1), which identified significant genes with altered expression levels at cumulative exposure levels. Notably the expression level of *PDE10A* (FDR < 0.05), which encodes an enzyme that modulates cAMP and cGMP signaling, was found to be altered. GO analyses revealed that PM<sub>2.5</sub> exposure impacted genes involved in purine ribonucleotide and nucleoside phosphate metabolism (Table S2). We found an inverse correlation between the expression levels purine-metabolism-related genes, such as *ENPP3*, and lung function, particularly the FEV<sub>1</sub>/FVC ratio (Table S3). This suggests a potential role in PM<sub>2.5</sub>-related COPD pathogenesis and suggests avenues for therapeutic intervention. Real-time data collected from portable devices are the most accurate means of measuring PM<sub>2.5</sub> exposure levels. The biological responses elicited by real-time exposure concentrations can provide valuable scientific insights. However, inherent limitations remain, including missing data for continuous exposure metrics and the potential for inaccuracies, because individual characteristics may not have been fully captured. Continuous improvements in personal exposure measurement techniques are crucial to enhance the accuracy of exposure assessments. Long-term assessments of individual exposure to air pollution are

infrequent in epidemiological research. Nevertheless, advances in technology have enhanced the feasibility of these measurements, leading to their gradual integration into epidemiological investigations. Accordingly, future studies should incorporate long-term individual exposure assessments to improve our understanding of the health impacts associated with air pollution.

When evaluating the relationship between air pollution exposure and respiratory diseases, it is essential to consider the spatial and temporal variations in air pollution resulting from human movement (Ma et al., 2024). When evaluating the relationship between air pollution exposure and respiratory diseases, it is essential to consider the spatial and temporal variations in air pollution resulting from human movement (Ma et al., 2024). Previous studies on COPD associated with air pollution and PM<sub>2.5</sub> exposure have mainly focused on residential areas, overlooking other locations, such as workplaces and sites of leisure facilities. Neglecting the exposure to different microenvironments can introduce biases and uncertainties into exposure assessments. This study addressed the importance of human mobility by measuring PM<sub>2.5</sub> levels among participants using a direct-reading device and incorporating time-activity diary data to calculate estimated PM<sub>2.5</sub> concentrations. Based on RNA-sequencing related to COPD, we analyzed DEGs following short-, medium-, and long-term exposure to PM<sub>2.5</sub> using two air prediction models employing indirect methodologies. Although previous studies have predominantly focused on single-term exposure to PM<sub>2.5</sub> (Chen et al., 2020; Yao et al., 2021), studies evaluating the gene expression patterns and biological pathways enriched in relation to short-, medium-, and long-term exposure remain limited. We employed individualized and ambient PM<sub>2.5</sub> exposure estimates obtained from air quality prediction models to identify DEGs corresponding to each exposure period. Notably, our analyses showed overlapping genes in the short- and medium-term exposure groups using both the hybrid and LUR models. However, a distinct set of overlapping genes that were not differentially expressed after long-term exposure was identified. Furthermore, the functional annotations and biological pathways associated with the DEGs varied with the exposure period. The absence of



**Fig. 3.** Functional and biological pathways related to the identified genes based on individual and ambient PM<sub>2.5</sub> exposure during specific exposure periods. **A.** Heatmap of significant GO terms identified from DEGs based on each criterion. **B.** Boxplots showing KEGG pathways significantly associated with DEGs based on individual and ambient PM<sub>2.5</sub> exposure across three exposure windows. GO terms and KEGG pathway with a *p*-value of less than 0.05 are shown; color indicates log<sub>10</sub> (*p*-value). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

common DEGs across all exposure durations suggests that the biological responses elicited by short- and medium-term exposure to PM<sub>2.5</sub> are likely to be divergent, potentially influencing different disease pathways.

We conducted GO and KEGG pathway analyses to enhance our understanding of the associated signaling pathways. In the hybrid model, our findings indicate that short-term PM<sub>2.5</sub> exposure impacts genes related to hemoglobin, which is essential for transporting gases such as carbon dioxide and oxygen. Exposure to CO can induce tissue hypoxia by forming carboxyhemoglobin (COHb), which reduces the oxygen-carrying capacity and may increase sensitivity among patients with COPD (Yasuda et al., 2005). Medium-term exposure revealed genes related to vascular development and exocytosis regulation (Table S5). It is well documented that PM<sub>2.5</sub> contributes to mitochondrial dysfunction and oxidative stress, both of which can disrupt calcium signaling, an essential process for exocytosis, and affect proteins involved in vesicle fusion, such as SNARE proteins (Itakura et al., 2012). This regulated exocytosis is vital for cellular function; however, PM<sub>2.5</sub> exposure can compromise cellular function by inducing cellular damage and altering the signaling pathways involved in exocytosis (Liu et al., 2023), leading to inhibited secretion by lung cells. This inhibition is notable in the

context of COPD, as it can impair mucus clearance, hinder immune responses, and disrupt lung architecture (Chen et al., 2024; Gu et al., 2017; Zhao et al., 2019). Moreover, long-term exposure to PM<sub>2.5</sub> has been associated with morphogenetic changes in neuronal projections. This aligns with the existing literature indicating a strong correlation between sustained exposure to PM<sub>2.5</sub>, and neurodevelopmental disorders (Ahadullah et al., 2021). Notably, PM<sub>2.5</sub> disrupts Hoxa5-mediated axon, dendritic, and synaptic morphogenesis, which may lead to cognitive impairment, particularly in spatial learning and memory (Liu et al., 2022). These mechanisms are crucial for understanding the relationship between PM<sub>2.5</sub> exposure and neurodevelopmental deficits. Although COPD is characterized as a pulmonary condition, emerging research suggests that it can affect neurological health by affecting neuronal development and function (Yu et al., 2025). Investigating the molecular pathways through which chronic inflammation, hypoxia, and oxidative stress from PM<sub>2.5</sub> exposure disrupt neuronal projections in COPD is vital for developing targeted strategies to counteract these adverse respiratory health effects. Additionally, KEGG pathway analysis of long-term PM<sub>2.5</sub> exposure highlighted the activation of C-type lectin receptors (CLRs). Components of PM<sub>2.5</sub>, including endotoxins and allergens, are recognized by immune cell pattern recognition receptors, which

**Table 3**

GO analysis of DEGs in relation to both personal and ambient PM<sub>2.5</sub> exposure levels.

GO	Description	Hits	-Log <sub>10</sub> (p-value)	Z-score
GO:0030185	nitric oxide transport	<i>HBA1, HBA2, HBB</i>	6.986442	28.45187
GO:0035239	tube morphogenesis	<i>EGR3, ENPEP, EREG, FGFR2, FOXD1, LEP, BCAM, NKX3-1, PML, TEK, THY1, NOG, MICAL2, EDAR, HEY1, SOX8, CSMD1, NDNF, C1orf54, EGFL8, VSTM4</i>	4.795477	5.182673
GO:0001568	blood vessel development	<i>EGR3, ENPEP, EREG, FGFR2, LEP, BCAM, NKX3-1, TEK, THY1, NOG, SCN1A, HEY1, NDNF, EGFL8, VSTM4</i>	3.246654	4.036933
GO:0048514	blood vessel morphogenesis	<i>EGR3, ENPEP, EREG, FGFR2, LEP, BCAM, TEK, THY1, NOG, HEY1, NDNF, EGFL8, VSTM4</i>	3.074567	3.951498
GO:0001525	angiogenesis	<i>EGR3, ENPEP, EREG, FGFR2, LEP, BCAM, TEK, THY1, HEY1, NDNF, VSTM4</i>	2.931967	3.908657
GO:0050680	negative regulation of epithelial cell proliferation	<i>EREG, FGFR2, IL12A, NKX3-1</i>	3.619667	6.751854
GO:0050673	epithelial cell proliferation	<i>EREG, FGFR2, NKX3-1, C1orf54</i>	3.610289	6.729739
GO:0051302	regulation of cell division	<i>EREG, FGFR2, NKX3-1, INSC</i>	3.493276	6.457598
GO:0021954	central nervous system neuron development	<i>FGFR2, MAP2, NDNF</i>	3.241679	7.082482
GO:0051781	positive regulation of cell division	<i>EREG, FGFR2, NKX3-1</i>	3.176067	6.881011
GO:0050678	regulation of epithelial cell proliferation	<i>EGR3, EREG, FGFR2, IL12A, NKX3-1</i>	3.169782	5.167727

**Table 3 (continued)**

GO	Description	Hits	-Log <sub>10</sub> (p-value)	Z-score
GO:0044344	cellular response to fibroblast growth factor stimulus	<i>EGR3, FGFR2, NDNF</i>	2.796511	5.792252
GO:0048754	branching morphogenesis of an epithelial tube	<i>FGFR2, FOXD1, NKX3-1</i>	2.751044	5.669963
GO:0048608	reproductive structure development	<i>EREG, FGFR2, HSD17B3, NKX3-1</i>	2.734001	4.842343
GO:0001666	response to hypoxia	<i>FGFR2, NKX3-1, TRPV4, NDNF</i>	2.728815	4.832092
GO:0071456	cellular response to hypoxia	<i>FGFR2, NKX3-1, NDNF</i>	2.724593	5.599563
GO:0061138	morphogenesis of a branching epithelium	<i>FGFR2, FOXD1, NKX3-1, PML, NOG, SOX8, CSMD1</i>	2.692785	3.984367
GO:0072001	renal system development	<i>ENPEP, FGFR2, FOXD1, NKX3-1, TEK, NOG, MAGI2, SOX8, C1orf54, PCSK9</i>	2.678196	3.683733
GO:0060562	epithelial tube morphogenesis	<i>FGFR2, FOXD1, NKX3-1, C1orf54</i>	2.672975	4.722318
GO:0048660	regulation of smooth muscle cell proliferation	<i>EREG, FGFR2, IL12A</i>	2.474833	4.960311
GO:0008283	cell population proliferation	<i>EREG, FGFR2, NKX3-1, PTGES, C1orf54</i>	2.115634	3.436613
GO:0009154	purine ribonucleotide catabolic process	<i>PDE4A, PDE9A, HKDC1</i>	2.884144	6.038833
GO:0008277	regulation of G protein-coupled receptor signaling pathway	<i>GSK3A, PDE4A, PDE6H</i>	2.031951	3.931337
GO:0044772	mitotic cell cycle phase transition	<i>BIRC5, CDC25C, CKS2</i>	2.554433	5.161838
GO:0008015	blood circulation	<i>DMD, ENPEP, GSK3A, HBB, HSD11B2, LEP, SCN1B, TEK, SCN1A, TRPM4, TRPV4, PLVAP, VSTM4</i>	3.18437	4.065591
GO:0003013	circulatory system process	<i>DMD, ENPEP, GSK3A, HBB, HSD11B2, LEP, SCN1B, TEK, SLC28A2, SCN1A, TRPM4, TRPV4, PLVAP, VSTM4</i>	2.899821	3.727093

(continued on next page)

**Table 3 (continued)**

GO	Description	Hits	-Log <sub>10</sub> (p-value)	Z-score
GO:0003018	vascular process in circulatory system	<i>HBB, LEP, TEK, PLVAP</i>	2.470511	4.340864
GO:0008217	regulation of blood pressure	<i>ENPEP, HBB, LEP</i>	2.037115	3.94099
GO:0051897	positive regulation of phosphatidylinositol 3-kinase/protein kinase B signal transduction	<i>IGF2, LEP, TEK, TPBG</i>	3.073068	5.544635
GO:0042102	positive regulation of T cell proliferation	<i>IGF2, LEP, TNFSF9</i>	2.782143	5.75866
GO:0032868	response to insulin	<i>FGFR2, GSK3A, HSD11B2, IGF2, LEP, TEK, TRPV4, CPEB1, PCSK9</i>	2.66048	3.734541
GO:0001503	ossification	<i>IGF2, LEP, TEK, MRC2</i>	2.38421	4.178065
GO:0032869	cellular response to insulin stimulus	<i>FGFR2, GSK3A, IGF2, LEP, TEK, CPEB1, PCSK9</i>	2.299271	3.451964
GO:0001936	regulation of endothelial cell proliferation	<i>IGF2, LEP, TEK</i>	2.125733	4.139712
GO:0045055	regulated exocytosis	<i>DOC2B, SCN11A, PCLO</i>	3.051701	6.500997
GO:0010611	regulation of cardiac muscle hypertrophy	<i>GSK3A, PDE9A, PI16</i>	3.049912	6.51018
GO:0001558	regulation of cell growth	<i>GSK3A, MAP1B, PML, CPNE5, PI16</i>	2.332544	3.794386
GO:0090257	regulation of muscle system process	<i>GSK3A, PDE9A, TRPM4, PI16</i>	2.321248	4.064234
GO:0008585	female gonad development	<i>LEP, DMC1, CSMD1</i>	2.880968	6.028941
GO:0007292	female gamete generation	<i>LEP, DMC1, PAQR5</i>	2.215141	4.344391
GO:0023061	signal release	<i>LEP, SYN1, PCLO</i>	2.06169	3.995701
GO:0002706	regulation of lymphocyte mediated immunity	<i>C4BPA, LEP, C17orf99</i>	2.001293	3.861768
GO:0001819	positive regulation of cytokine production	<i>CAMP, EREG, IL12A, ORM1, TRPV4</i>	2.730102	4.424171
GO:0032675	regulation of interleukin-6 production	<i>EREG, ORM1, TRPV4</i>	2.453999	4.90896
GO:0031349	positive regulation of defense response	<i>EREG, IL12A, TRPV4, USP17L2</i>	2.132761	3.708205
GO:0006753	nucleoside phosphate metabolic process	<i>GUCY1A2, GUCY2C, HKDC1, AK7, ACOT1</i>	2.304255	3.739454
GO:0009165	nucleotide biosynthetic process	<i>GUCY1A2, GUCY2C, AK7</i>	2.093161	4.064051
GO:0009117	nucleotide metabolic process	<i>GUCY1A2, GUCY2C, HKDC1, AK7</i>	2.028021	3.522553
GO:0098660	inorganic ion transmembrane transport	<i>GRINA, KCNJ9, SCN1B, ANO7, TRPM4,</i>	2.339878	3.49085

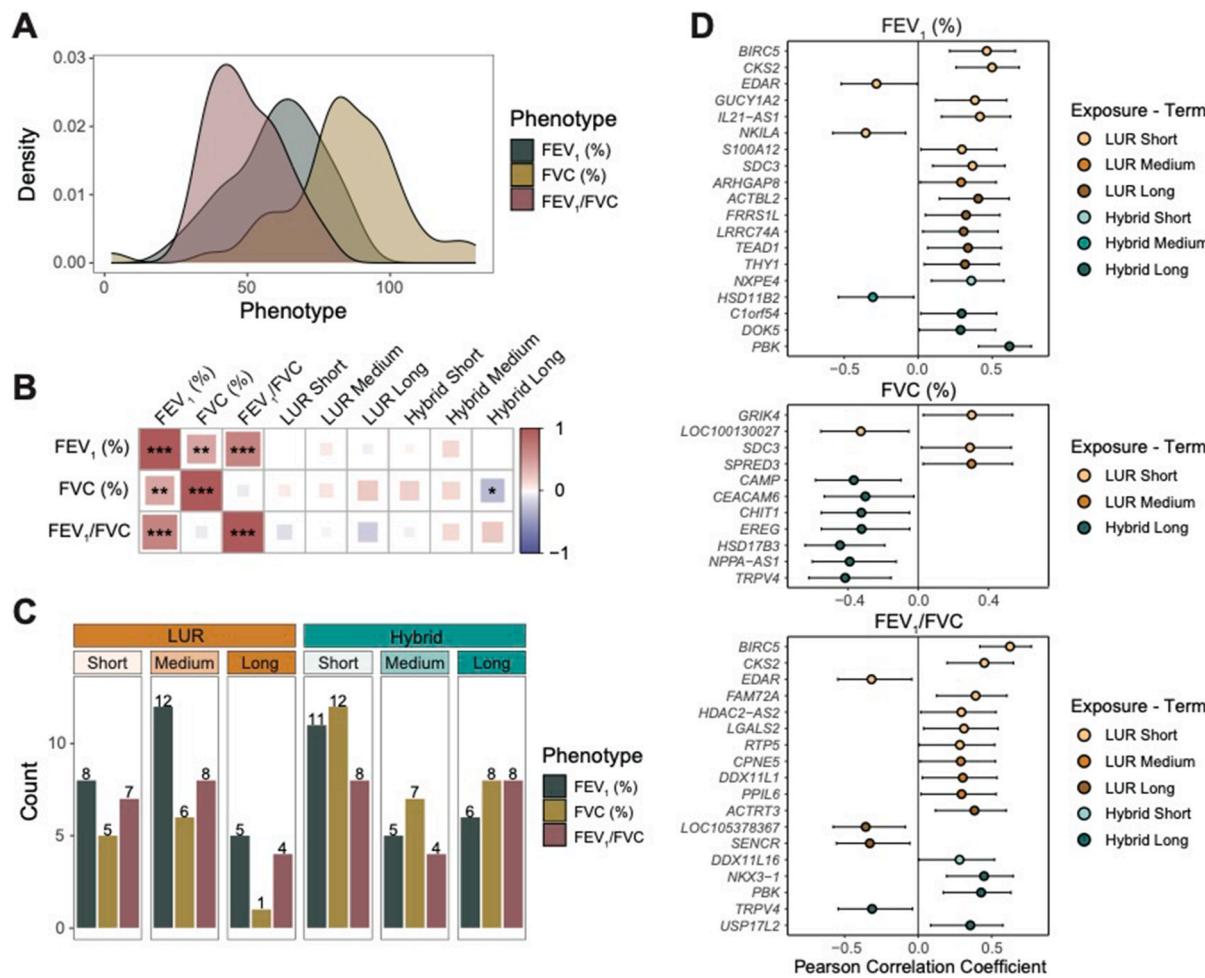
**Table 3 (continued)**

GO	Description	Hits	-Log <sub>10</sub> (p-value)	Z-score
		<i>SLC48A1, BEST4</i>		
GO:0035148	tube formation	<i>NOG, EDAR, SOX8</i>	2.579914	5.226458
GO:0007423	sensory organ development	<i>NOG, SOX8, CRYGN, VSTM4, TMIE</i>	2.209859	3.58935
GO:0010975	regulation of neuron projection development	<i>DMD, GSK3A, MAP1B, MAP2, NEFL, SCN1B, THY1, LZTS3, MAGI2, TRPV4, NDNF, TRIM67</i>	2.536675	3.419013
GO:0120035	regulation of plasma membrane bounded cell projection organization	<i>DMD, GSK3A, MAP1B, MAP2, NEFL, RAP1GAP, SCN1B, THY1, LZTS3, MAGI2, CDHR2, TRPV4, NDNF, USP17L2, TRIM67</i>	2.417992	3.193123
GO:0007156	homophilic cell adhesion via plasma membrane adhesion molecules	<i>CEACAM6, PCDHGB1, PCDHGA7</i>	2.517689	5.066859
GO:0090287	regulation of cellular response to growth factor stimulus	<i>DMD, FOXD1, IL12A, DOK5</i>	2.370027	4.143872
GO:0051056	regulation of small GTPase mediated signal transduction	<i>TEK, ARHGEF10, ARHGEF25, TRIM67</i>	2.337984	4.091695
GO:0099536	synaptic signaling	<i>DMD, EGR3, GRID1, MYCBPAP</i>	2.067677	3.590893

**Abbreviations:** GO, Gene Ontology; DEG, Differentially Expressed Gene.

subsequently activate inflammatory pathways (Wang and Liu, 2023). The activation of CLEC5A among CLRs has been implicated in enhanced inflammation through the engagement of downstream signaling pathways, such as those involving Syk kinase (Sung et al., 2020). This cascade may lead to enhanced inflammation and the exacerbation of COPD (Heyl Kerstin et al., 2014; Wortham et al., 2016). These insights enhance our understanding of the mechanisms underlying PM<sub>2.5</sub>-induced inflammatory responses that contribute to the exacerbation of pulmonary diseases.

In the LUR model, genes associated with PM<sub>2.5</sub> exposure were analyzed across three cumulative windows: short-, medium-, and long-term. After short-term exposure, genes involved in mitotic cell cycle phase transitions, tube formation, and nucleoside phosphate metabolic processes were significantly enriched. Conversely, medium- and long-term exposure were associated with genes pertinent to neuronal projection morphogenesis and general cell morphogenesis. KEGG pathway analysis revealed that PM<sub>2.5</sub> exposure was associated with biological pathways involved in purine metabolism. This was further corroborated by real-time PM<sub>2.5</sub> exposure studies through GO-term and KEGG pathway analyses. Purine, a crucial metabolic precursor for DNA and RNA synthesis in all biological systems, is notably affected by PM<sub>2.5</sub>.



**Fig. 4.** Association between PM<sub>2.5</sub> exposure-related DEGs and COPD phenotypes. **A.** Distribution of clinical phenotypes in patients with COPD, including FEV<sub>1</sub> (%), FVC (%), and FEV<sub>1</sub>/FVC. **B.** Correlation between clinical phenotypes in patients with COPD and PM<sub>2.5</sub> exposure concentration. The color and size of each cell represent the Pearson correlation coefficient, with the asterisks indicating statistical significance (\*\*p < 0.01; \*\*p < 0.01; \*p < 0.05). **C.** Bar-plot showing the number of DEGs, identified based on each criterion, that correlate with COPD clinical phenotypes. **D.** Point-range plot showing Pearson correlation coefficient between gene expression levels and clinical phenotypes. Error bars denote 95 % confidence intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

exposure, which alters the levels of purine metabolites (Huang et al., 2018). These alterations may disrupt purine synthesis and degradation mechanisms, thereby influencing COPD severity and exacerbation. Elevated concentrations of purines, particularly uric acid, have been documented in the airways of patients with COPD and are correlated with clinical indicators of disease severity (Huang et al., 2018). In addition, the uric acid-to-creatinine ratio serves as a significant biomarker for assessing the severity and risk of exacerbation of COPD (Rumora et al., 2020). Future research is essential to systematically delineate the relationships between PM<sub>2.5</sub>, purine metabolic alterations, and COPD pathophysiology. Long-term exposure may result in the cumulative effects of shorter exposure periods. The effect of cycle-specific biological effects on COPD progression remains unclear, highlighting the need for additional research to elucidate the mechanisms governing air pollution exposure cycles.

This study investigated the gene expression differences to elucidate the relationship between clinical indicators, air pollution, and respiratory diseases. The main goal of this study was to identify the genes that influence lung function changes in patients with COPD during PM<sub>2.5</sub> exposure periods. Our analysis evaluated the association between PM<sub>2.5</sub> exposure period estimates and gene expression levels. The findings revealed that the up-regulated expression levels of *CAMP*, *CEACAM6*, *CHIT1*, *EREG*, *HSD17B3*, *NPPA-AS1*, and *TRPV4* were significantly

associated with decreased lung function as air pollutant concentrations increased. The *TRPV4* gene plays a pivotal role in the cellular response to prolonged exposure to PM<sub>2.5</sub>. As an ion channel, *TRPV4* facilitates the influx of calcium ions across the cell membrane, which is essential for respiratory homeostasis (Li et al., 2011). Its expression in airway epithelial cells is crucial for the regulation of pulmonary blood flow, the maintenance of fluid homeostasis, and the modulation of immune responses (Morgan et al., 2018). Furthermore, research specifically investigating the expression of *TRPV4* in patients can elicit itch transmission response to histamine or chloroquine through TRPA1 receptor signaling (Kim et al., 2016). Moreover, reactive oxygen species generated during the inflammatory response have been implicated in TRPA1 activation (Mukhopadhyay et al., 2011), which further amplifies the immune response in COPD. Genotyping studies have linked the T allele of rs7819749 in TRPA1 to a decline in lung function (Naumov et al., 2021). This finding underscores the potential significance of these cation channels in COPD pathogenesis and supports our research approach aimed at elucidating the role of *TRPV4* in lung function decline due to PM<sub>2.5</sub>. The *CAMP* gene, which encodes cathelicidin antimicrobial peptide, is located on chromosome 3 at position p21.31. Exposure to PM<sub>2.5</sub>, has been shown to decrease lung cathelicidin levels (Burkes et al., 2020). There is emerging evidence linking low plasma cathelicidin levels to an increased risk of exacerbation in patients with COPD. Furthermore,

**Table 4**

Significantly up- and down-regulated DEGs associated with COPD phenotypes across exposure time windows in each model.

	LUR			Hybrid		
	Short	Medium	Long	Short	Medium	Long
FEV <sub>1</sub> (%)	<i>BIRC5</i> <i>CKS2</i> <i>EDAR</i> ↑ <i>GUCY1A2</i> <i>IL21-AS1</i> <i>NKILA</i> ↑ <i>S100A12</i> <i>SDC3</i> <i>GRIK4</i> <i>LOC100130027†</i> <i>SDC3</i>	<i>ARHGAP8</i>	<i>ACTBL2</i> <i>FRRS1L</i> <i>LRRC74A</i> <i>TEAD1</i> <i>THY1</i>	<i>NXPE4</i>	<i>HSD11B2†</i> <i>C1orf54</i> <i>DOK5</i> <i>PBK</i>	
FVC(%)		<i>SPRED3</i>				<i>CAMP</i> ↑ <i>CEACAM6</i> ↑ <i>CHIT1</i> ↑ <i>EREG</i> ↑ <i>HSD17B3†</i> <i>NPPA-AS1</i> ↑ <i>TRPV4</i> ↑
FEV <sub>1</sub> /FVC	<i>BIRC5</i> <i>CKS2</i> <i>EDAR</i> ↑ <i>FAM72A</i> <i>HDAC2-AS2</i> <i>LGALS2</i> <i>RTP5</i>	<i>CPNE5</i> <i>DDX11L1</i> <i>PPIL6</i>	<i>ACTRT3</i> <i>LOC105378367†</i> <i>SENCR</i> ↑	<i>DDX11L16</i>		<i>NKX3-1</i> <i>PBK</i> <i>TRPV4</i> ↑ <i>USP17L2</i>

Note: \*†, represented by the up-regulated genes.

Abbreviations: FEV<sub>1</sub>, Forced Expiratory Volume in 1 s; FVC, Forced Vital Capacity; FEV<sub>1</sub>/FVC ratio, Ratio of forced expiratory volume in 1 s to forced vital capacity.

cathelicidin has been recognized for its roles in both antimicrobial defense and immune modulation in COPD (Singanayagam et al., 2019). Li et al. documented the differential expression of *CAMP*, along with other critical genes such as *DEFA4*, *ELANE*, *LCN2*, and *LTF*, in COPD (Li et al., 2022). The *CHIT1* gene, which encodes the chitotriosidase enzyme, is upregulated in macrophages in COPD and its expression levels are correlated with elevated chitinase activity in smokers (Agapov et al., 2009), likely exacerbating the inflammatory response and tissue damage that are characteristic of the disease. Previous studies have indicated that *CHIT1* levels in the blood or bronchoalveolar lavage fluid can serve as biomarkers of lung disease severity and progression (Chang et al., 2020). Additionally, *EREG* encodes a protein that binds to the epidermal growth factor receptor (EGFR), thereby activating downstream pathways that promote cell proliferation, survival, and differentiation during the progression of lung cancer in the context of PM<sub>2.5</sub> exposure (Chen et al., 2025). In COPD, pathways involving *EREG* are associated with airway remodeling in smokers, featuring basal cell and mucus cell hyperplasia, squamous metaplasia, and alterations in ciliated cell differentiation (Zuo et al., 2017). *EREG* expression is particularly sensitive to PM<sub>2.5</sub>, and its upregulation has been implicated in tumorigenesis and tumor invasion. However, the study design had some limitations, including repeated measurements of individual exposure levels and lung function at each time point. Therefore, better study designs are required to achieve greater precision through repeated measurements categorized according to exposure and clinical indicators in patients with COPD. While the specific functions of NPPA-AS1, SENCR in human exposure studies related to COPD and PM<sub>2.5</sub> remain unclear, previous research has highlighted the significant role of long non-coding RNAs (lncRNAs) in biological responses to air pollution and respiratory diseases (Xie et al., 2023). For example, one study revealed that lncRNA LOC101927514 regulates inflammation induced by PM<sub>2.5</sub> exposure by interacting with the p-STAT3 protein in human bronchial epithelial cells (Tan et al., 2020). Another study suggested that long-term PM<sub>2.5</sub> exposure may enhance the malignancy of human bronchial epithelial cells by activating upstream EGFR via the lncRNA SOX2-OT (Fu et al., 2021). Additionally, the lncRNA PAET, associated with PM<sub>2.5</sub> exposure, was identified as a potential biomarker due to its role in promoting DNA damage through OXPHOS-dependent regulation, particularly in childhood asthma (Zheng et al., 2024). In this study, we identified a

significant correlation between MAST4-AS1, located at chromosomal locus 5q12.3, and a decline in FVC, marking its first recognition in this context. Further investigations into the roles and functions of these genes as COPD-specific markers of air pollution are warranted.

This study has several limitations that should be acknowledged. First, the limited sample size and absence of a control group restrict the generalizability of the findings, which are primarily applicable to individuals with COPD. Moreover, these results have not been validated in other populations, making it challenging to broaden the applicability of these findings. The accuracy and reliability of the air pollution exposure assessments are crucial to the validity of this study. It is well established that patients with COPD generally exhibit lower levels of physical activity than control groups. In this study, participants maintained physical activity diaries to integrate their time and movement patterns into the exposure measurement data more accurately. However, the potential for inaccuracies remains as individual characteristics may not have been fully captured. This study introduced a methodology to examine the effects of individual air pollution exposure on COPD by employing both a land regression model and an individual-exposure modeling approach. Future studies should prioritize the development and implementation of survey methods that can effectively delineate air-polluted microenvironments, thereby validating and generalizing the findings on a larger scale. Additionally, a limitation of this study is its exclusive focus on PM<sub>2.5</sub> as the sole air pollutant measured. Given the complexity of air pollution, unmeasured pollutants or correlations with the assessed factors may contribute to its effects. Furthermore, the association analysis conducted with the identified genes was based on a single pulmonary function test measurement, and variations in PM<sub>2.5</sub> levels at the time of testing may have had a limited effect on lung function. To enhance individual exposure assessments, future research should incorporate repeated measurements and follow-up tests.

Despite some limitations, our method for assessing personal and ambient exposure to PM<sub>2.5</sub> has significantly enhanced our understanding of the molecular mechanisms linking air pollution and COPD. We have identified transcriptomic reactions and biological pathways over three successive exposure periods, as well as new genes associated with the clinical phenotypes of COPD. This discovery may enhance strategies for preventing and managing the disease. However, further investigation is essential to elucidate the causal relationships between the

differential gene expression and their possible pathobiological functions in the adverse effects of air pollution on COPD.

#### CRediT authorship contribution statement

**Jeeyoung Kim:** Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Ha Won Song:** Visualization, Methodology, Formal analysis, Data curation. **Hyun Woo Lee:** Resources, Investigation, Data curation. **Ye Jin Lee:** Resources, Investigation, Data curation. **Sooim Sin:** Resources, Investigation, Data curation. **Ji Yeon Lee:** Resources, Investigation, Data curation. **Junghyun Kim:** Resources, Investigation, Data curation. **Sun Mi Choi:** Resources, Investigation, Data curation. **Kyoung-Nam Kim:** Resources, Investigation, Data curation. **Chang-Hoon Lee:** Resources, Investigation, Data curation. **Chang Hyun Lee:** Resources, Investigation, Data curation. **Woo Jin Kim:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Investigation, Funding acquisition, Data curation, Conceptualization.

#### Ethics approval and informed consent

This study was approved by the Seoul National University Hospital IRB (SNUH, 1810-036-977), and all study protocols were conducted in accordance with the Institutional Review Board of Seoul National University Hospital. Written informed consent was obtained from all participants.

#### Consent for publication

All authors agree to this publication.

#### Funding

This study was financially supported by grants from the Korea Environment Industry & Technology Institute (KEITI), through the Core Technology Development Project for Environmental Disease Prevention and Management funded by the Korea Ministry of Environment (Grant No. RS-2022-KE002052), and the National Research Foundation of South Korea, also funded by the Ministry of Education (NRF-2021 R1I1A1A01047585). Furthermore, this study received partial support from the Particulate Matter Management Specialized Graduate Program through the KEITI, funded by the Ministry of Environment (MOE), South Korea.

#### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

#### Acknowledgments

We would like to express our gratitude to the Genome Informatics Laboratory of Professor Sanghyuk Lee at the Department of Life Science, Ewha Research Center for Systems Biology, Ewha Womans University in South Korea, for their invaluable support and access to the quality control of the RNA-seq raw data.

#### Abbreviation

Boramae Medical Center (BR)  
Bundang Hospital (SB)  
Seoul National University Hospital (SN)  
Kangwon National University Hospital (KW)  
Charlson Comorbidity Index (CCI)  
Chronic Obstructive Pulmonary Disease (COPD)

Differential Gene Expression (DEG)  
Forced Vital Capacity (FVC)  
Forced Expiratory Volume in 1 s (FEV<sub>1</sub>)  
Particulate Matter 2.5 (PM<sub>2.5</sub>)

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envres.2025.122377>.

#### Data availability

Data will be made available on request.

#### References

- Agapov, E., et al., 2009. Macrophage chitinase 1 stratifies chronic obstructive lung disease. *Am. J. Respir. Cell Mol. Biol.* 41, 379–384.
- Ahadullah, et al., 2021. PM2.5 as a potential risk factor for autism spectrum disorder: its possible link to neuroinflammation, oxidative stress and changes in gene expression. *Neurosci. Biobehav. Rev.* 128, 534–548.
- Andrews, S., 2010. FastQC: a Quality Control Tool for High Throughput Sequence Data.
- Burkes, R.M., et al., 2020. Plasma cathelicidin is independently associated with reduced lung function in COPD: analysis of the subpopulations and intermediate outcome measures in COPD study cohort. *Chronic. Obstr. Pulm. Dis.* 7, 370–381.
- Chang, D., et al., 2020. Chitotriosidase: a marker and modulator of lung disease. *Eur. Respir. Rev.* 29.
- Chen, X., et al., 2020. Susceptibility of individuals with chronic obstructive pulmonary disease to air pollution exposure in Beijing, China: a case-control panel study (COPDB). *Sci. Total Environ.* 717, 137285.
- Chen, L., et al., 2024. Effects of PM2.5 on mucus hypersecretion in airway through miR-133b-5p/EGFR/Claudin1/MUC5AC axis. *Aging* 16, 8472–8483.
- Chen, C.-Y., et al., 2025. The role of PM2.5 exposure in lung cancer: mechanisms, genetic factors, and clinical implications. *EMBO Mol. Med.* 17, 31–40, 40.
- Dobin, A., et al., 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Evangelopoulos, D., et al., 2021. Personal exposure to air pollution and respiratory health of COPD patients in London. *Eur. Respir. J.* 58.
- Favé, M.-J., et al., 2018. Gene-by-environment interactions in urban populations modulate risk phenotypes. *Nat. Commun.* 9, 827.
- Fu, Y., et al., 2021. lncRNA SOX2-OT ceRNA network enhances the malignancy of long-term PM2.5-exposed human bronchial epithelia. *Ecotoxicol. Environ. Saf.* 217, 112242.
- Fuller, R., et al., 2022. Pollution and health: a progress update. *Lancet Planet. Health* 6, e535–e547.
- Gu, X.-y., et al., 2017. Effects of PM2.5 exposure on the Notch signaling pathway and immune imbalance in chronic obstructive pulmonary disease. *Environ. Pollut.* 226, 163–173.
- Heyl Kersting, A., et al., 2014. Dectin-1 is expressed in human lung and mediates the proinflammatory immune response to nontypeable *Haemophilus influenzae*. *mBio* 5. <https://doi.org/10.1128/mBio.01492-14>.
- Huang, Q., et al., 2018. The modification of indoor PM2.5 exposure to chronic obstructive pulmonary disease in Chinese elderly people: a meet-in-metabolite analysis. *Environ. Int.* 121, 1243–1252.
- ISO, 2019. Molecular in Vitro Diagnostic Examinations — Specifications for Pre-examination Processes for Venous Whole Blood — Part 1: Isolated Cellular RNA ISO Standard No. 20186-1(E).
- Itakura, E., et al., 2012. The hairpin-type tail-anchored SNARE syntaxin 17 targets to autophagosomes for fusion with endosomes/lysosomes. *Cell* 151, 1256–1269.
- Ji, H.W., et al., 2024. The association between cumulative exposure to PM(2.5) and DNA methylation measured using methyl-capture sequencing among COPD patients. *Respir. Res.* 25, 335.
- Kakati, T., et al., 2019. Comparison of methods for differential Co-expression analysis for disease biomarker prediction. *Comput. Biol. Med.*, 103380.
- Kelchtermans, J., et al., 2025. Genetic modifiers of asthma response to air pollution in children: an African ancestry GWAS and PM2.5 polygenic risk score study. *Environ. Res.* 267, 120666.
- Kim, S., et al., 2016. Facilitation of TRPV4 by TRPV1 is required for itch transmission in some sensory neuron populations. *Sci. Signal.* 9, ra71–ra71.
- Lê, S., Josse, J., Husson, F., 2008. FactoMineR: an R package for multivariate analysis. *J. Stat. Software* 25 (1), 1–18.
- Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* 12, 323.
- Li, J., et al., 2011. TRPV4-Mediated calcium influx into human bronchial epithelia upon exposure to diesel exhaust particles. *Environ. Health Perspect.* 119, 784–793.
- Li, R., et al., 2022. Differential expression of serum proteins in chronic obstructive pulmonary disease assessed using label-free proteomics and bioinformatics analyses. *Int. J. Chronic Obstr. Pulm. Dis.* 17, 2871–2891.
- Liu, X., et al., 2019. A weekly time-weighted method of outdoor and indoor individual exposure to particulate air pollution. *MethodsX* 6, 2439–2442.

- Liu, C., et al., 2022. Role of necroptosis in airflow limitation in chronic obstructive pulmonary disease: focus on small-airway disease and emphysema. *Cell Death Discov.* 8, 363.
- Liu, Q., et al., 2023. Attenuation of PM2.5-induced alveolar epithelial cells and lung injury through regulation of mitochondrial fission and fusion. *Part. Fibre Toxicol.* 20, 28.
- Losacco, C., Perillo, A., 2018. Particulate matter air pollution and respiratory impact on humans and animals. *Environ. Sci. Pollut. Res. Int.* 25, 33901–33910.
- Ma, X., et al., 2024. A comprehensive review of the development of land use regression approaches for modeling spatiotemporal variations of ambient air pollution: a perspective from 2011 to 2023. *Environ. Int.* 183, 108430.
- Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12.
- Merid, S.K., et al., 2021. Integration of gene expression and DNA methylation identifies epigenetically controlled modules related to PM2.5 exposure. *Environ. Int.* 146, 106248.
- Metascape, 2024. A Gene Annotation & Analysis Resource.
- Morgan, J.T., et al., 2018. The mechanosensitive ion channel TRPV4 is a regulator of lung development and pulmonary vasculature stabilization. *Cell. Mol. Bioeng.* 11, 309–320.
- Mukhopadhyay, I., et al., 2011. Expression of functional TRPA1 receptor on human lung fibroblast and epithelial cells. *J. Recept. Signal Transduction* 31, 350–358.
- Naumov, D.E., et al., 2021. Effect of TRPM8 and TRPA1 polymorphisms on COPD predisposition and lung function in COPD patients. *J. Personalized Med.* 11.
- Newman, J.D., et al., 2020. Cardiopulmonary impact of particulate air pollution in high-risk populations: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* 76, 2878–2894.
- Pickrell, J.K., et al., 2010. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature* 464, 768–772.
- Rumora, L., et al., 2020. Uric acid and uric acid to creatinine ratio in the assessment of chronic obstructive pulmonary disease: potential biomarkers in multicomponent models comprising IL-1beta. *PLoS One* 15, e0234363.
- Rundblad, A., et al., 2025. Exposure to fine particulate matter in adults is associated with immune cell gene expression related to inflammation, the electron transport chain, and cell cycle regulation. *Environ. Epigenet.* 11, dvaf008.
- Sin, D.D., et al., 2023. Air pollution and COPD: GOLD 2023 committee report. *Eur. Respir. J.* 61.
- Singanayagam, A., et al., 2019. Inhaled corticosteroid suppression of cathelicidin drives dysbiosis and bacterial infection in chronic obstructive pulmonary disease. *Sci. Transl. Med.* 11.
- Steinle, S., et al., 2015. Personal exposure monitoring of PM2.5 in indoor and outdoor microenvironments. *Sci. Total Environ.* 508, 383–394.
- Sung, P.-S., et al., 2020. CLEC5A: a promiscuous pattern recognition receptor to microbes and beyond. In: Hsieh, S.-L. (Ed.), *Lectin in Host Defense against Microbial Infections*. Springer Singapore, Singapore, pp. 57–73.
- Tan, Y., et al., 2020. LncRNA LOC101927514 Regulates PM(2.5)-driven Inflammation in Human Bronchial Epithelial Cells through Binding P-STAT3 Protein.
- Vlaanderen, J., et al., 2022. Impact of long-term exposure to PM2.5 on peripheral blood gene expression pathways involved in cell signaling and immune response. *Environ. Int.* 168, 107491.
- Wallbanks, S., et al., 2024. Impact of environmental air pollution on respiratory health and function. *Phys. Rep.* 12, e70006.
- Wang, Q., Liu, S., 2023. The effects and pathogenesis of PM2.5 and its components on chronic obstructive pulmonary disease. *Int. J. Chronic Obstr. Pulm. Dis.* 18, 493–506.
- Wortham, B.W., et al., 2016. Cutting edge: CLEC5A mediates macrophage function and chronic obstructive pulmonary disease pathologies. *J. Immunol.* 196, 3227–3231.
- Xie, J., et al., 2023. The role of lncRNA in the pathogenesis of chronic obstructive pulmonary disease. *Heliyon* 9, e22460.
- Yao, Y., et al., 2021. Differences in transcriptome response to air pollution exposure between adult residents with and without chronic obstructive pulmonary disease in Beijing: A panel study. *J. Hazard Mater.* 416, 125790.
- Yasuda, H., et al., 2005. Increased arterial carboxyhemoglobin concentrations in chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* 171, 1246–1251.
- Yu, X., et al., 2025. The lung-brain Axis in chronic obstructive pulmonary disease-associated neurocognitive dysfunction: mechanistic insights and potential therapeutic options. *Int. J. Biol. Sci.* 21, 3461–3477.
- Zhao, J., et al., 2019. Role of PM(2.5) in the development and progression of COPD and its mechanisms. *Respir. Res.* 20, 120.
- Zheng, R., et al., 2024. PM(2.5)-derived exosomal long noncoding RNA PAET participates in childhood asthma by enhancing DNA damage via m(6)A-dependent OXPHOS regulation. *Environ. Int.* 183, 108386.
- Zuo, W.L., et al., 2017. EGF-amphiregulin interplay in airway stem/progenitor cells links the pathogenesis of smoking-induced lesions in the human airway epithelium. *Stem Cell.* 35, 824–837.