J Korean Med Sci. 2023 Jul 24;38(29):e220
https://doi.org/10.3346/jkms.2023.38.e220
eISSN 1598-6357·pISSN 1011-8934

JKMS

Original Article
Respiratory Diseases

Machine Learning-Based Proteomics Reveals Ferroptosis in COPD Patient-Derived Airway Epithelial Cells Upon Smoking Exposure

Jung-Ki Yoon [1*], **Sungjoon Park** [2*], **Kyoung-Hee Lee** [3*], **Dabin Jeong** [4], **Jisu Woo** [3], **Jieun Park** [5], **Seung-Muk Yi** [6,7], **Dohyun Han** [8,9], **Chul-Gyu Yoo** [3], **Sun Kim** [2,4] and **Chang-Hoon Lee** [3]

[1]Division of Pulmonary, and Critical Care Medicine, Department of Medicine, Stanford University, Stanford, CA, USA
[2]Department of Computer Science and Engineering, Seoul National University, Seoul, Korea
[3]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Seoul National University Hospital, Seoul, Korea
[4]Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, Korea
[5]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[6]Graduate School of Public Health, Seoul National University, Seoul, Korea
[7]Institute of Health and Environment, Seoul National University, Seoul, Korea
[8]Transdisciplinary Department of Medicine & Advanced Technology, Seoul National University Hospital, Seoul, Korea
[9]Proteomics Core Facility, Biomedical Research Institute, Seoul National University Hospital, Seoul, Korea

Received: Nov 15, 2022
Accepted: Mar 27, 2023
Published online: Jun 14, 2023

Address for Correspondence:
Chang-Hoon Lee, MD, PhD
Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Seoul National University Hospital, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea.
Email: kauri670@empal.com

*Jung-Ki Yoon, Sungjoon Park, and Kyoung-Hee Lee contributed equally to this work.

ORCID iDs
Jung-Ki Yoon
https://orcid.org/0000-0003-4947-1412
Sungjoon Park
https://orcid.org/0000-0003-2023-9433
Kyoung-Hee Lee
https://orcid.org/0000-0001-5327-2171
Dabin Jeong
https://orcid.org/0000-0001-9910-8051
Jisu Woo
https://orcid.org/0009-0005-8740-2241
Jieun Park
https://orcid.org/0000-0002-0523-2121
Seung-Muk Yi
https://orcid.org/0000-0001-6333-4399

ABSTRACT

Background: Proteomics and genomics studies have contributed to understanding the pathogenesis of chronic obstructive pulmonary disease (COPD), but previous studies have limitations. Here, using a machine learning (ML) algorithm, we attempted to identify pathways in cultured bronchial epithelial cells of COPD patients that were significantly affected when the cells were exposed to a cigarette smoke extract (CSE).
Methods: Small airway epithelial cells were collected from patients with COPD and those without COPD who underwent bronchoscopy. After expansion through primary cell culture, the cells were treated with or without CSEs, and the proteomics of the cells were analyzed by mass spectrometry. ML-based feature selection was used to determine the most distinctive patterns in the proteomes of COPD and non-COPD cells after exposure to smoke extract. Publicly available single-cell RNA sequencing data from patients with COPD (GSE136831) were used to analyze and validate our findings.
Results: Five patients with COPD and five without COPD were enrolled, and 7,953 proteins were detected. Ferroptosis was enriched in both COPD and non-COPD epithelial cells after their exposure to smoke extract. However, the ML-based analysis identified ferroptosis as the most dramatically different response between COPD and non-COPD epithelial cells, adjusted P value = 4.172×10^{-6}, showing that epithelial cells from COPD patients are particularly vulnerable to the effects of smoke. Single-cell RNA sequencing data showed that in cells from COPD patients, ferroptosis is enriched in basal, goblet, and club cells in COPD but not in other cell types.
Conclusion: Our ML-based feature selection from proteomic data reveals ferroptosis to be the most distinctive feature of cultured COPD epithelial cells compared to non-COPD epithelial cells upon exposure to smoke extract.

Dohyun Han
https://orcid.org/0000-0002-0841-1598
Chul-Gyu Yoo
https://orcid.org/0000-0003-1321-7834
Sun Kim
https://orcid.org/0000-0001-5385-9546
Chang-Hoon Lee
https://orcid.org/0000-0001-9960-1524

Funding

This work was supported by the Korean Association of Internal Medicine Research Grant 2020. The study was supported by the National Research Foundation of Korea grant funded by the Ministry of Science and ICT (No. 2019R1A2C1008095) and by grant (0420190370) from Seoul National University Hospital.

Disclosure

The authors have no potential conflicts of interest to disclose.

Author Contributions

Conceptualization: Lee KH, Lee CH. Data curation: Yoon JK, Park S, Lee KH, Lee CH. Formal analysis: Park S, Jeong D, Kim S. Investigation: Yoon JK, Lee KH, Woo J, Park S. Methodology: Kim S, Lee KH. Park S, Yi SM, Han D. Software: Yoon JK, Park S, Jeong D. Validation: Yoon JK. Writing - original draft: Yoon JK, Lee CH, Lee KH, Park S. Writing - review & editing: Jeong D, Woo J, Park J, Yi SM, Han D, Yoo CG, Kim S, Yoon JK, Lee CH.

Keywords: Ferroptosis; Pulmonary Disease, Chronic Obstructive; Proteomics; Machine Learning; Epithelial Cells

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a respiratory disease with 10.3% global prevalence among people aged 30–79 years in 2019.[1,2] It was the third most common cause of death in 2019 worldwide, and as one of the leading causes of global disease burden, it is a major public health concern.[3,4] COPD is characterized by persistent respiratory symptoms and airflow obstruction caused by significant exposure to noxious gases or particles, including smoking.[5] Smoking generates aberrant inflammation that injures the airways and alveoli, contributing to the occurrence of COPD, and is regarded as the most critical risk factor for COPD.[6] However, the detailed molecular mechanisms in the pathogenesis of COPD have not been fully elucidated,[7] leading to setbacks in development of better therapies for COPD patients.

Recent advances in genomics and proteomics have enabled screens for the components that play essential roles in COPD pathogenesis.[8-10] However, previous studies have limitations. Firstly, most studies have shown the functional roles of cytokines and immune cells in COPD, and few studies have explored the direct changes in epithelial cells.[10] Considering that inhaled cigarette smoke has primary contact with the airway epithelial cells and that key histological findings in COPD, such as squamous cell metaplasia, and goblet cell hyperplasia, occur in epithelial cells, it is important to understand the pathological changes of epithelial cells in COPD. Secondly, smoking is one of the strongest factors causing COPD; however, most studies have focused on naïve COPD tissues and have yet to explore the differential response between COPD and non-COPD lungs to smoking exposure. To understand the link between smoking and COPD epithelial cells, we must understand the responses of the disease model to smoking exposure. Finally, high-throughput omics studies are vulnerable to false positive results since they produce numerous signals, especially in comparison to the number of samples. Therefore, feature selection algorithms need to be more resistant to noisy data in order to identify the key components in the pathogenesis of COPD.

In this study, we performed mass spectrometry (MS) with machine learning (ML)-based pathway analysis to identify the most distinctive pathways in COPD epithelial cells exposed to cigarette smoke extract (CSE), compared to cells from patients without non-COPD that were also exposed. The pathways were also explored using publicly available COPD single-cell RNA sequencing data relating to cells from patients with COPD and healthy controls.

METHODS

Sample collection and preparation

COPD patients and non-COPD participants who underwent bronchoscopy evaluation for their clinical needs were enrolled with informed consent. During the examination, bronchial epithelial cells were collected by bronchoscopic brushing using the protected sheath on the distal small airways. The COPD group was defined as smokers with emphysema on lung computed tomography (CT) with post-bronchodilator (BD) forced expiratory volume in 1 second (FEV_1) over post-BD forced vital capacity (FVC) < 70%. The non-COPD group was

defined as non-smokers with no emphysema on CT, whose post-BD FEV_1/FVC was > 80%. Specimens were obtained from the lobe with emphysema in patients with COPD, and from the right middle lobe in patients without COPD. The brushed cells were cultured in defined keratinocyte-SFM (GIBCO, Life Technologies, Grand Island, NY, USA) with 1% penicillin/streptomycin. Cells were grown at 37°C with 5% CO_2, and the media were changed every other day. CSE was prepared from 20 commercial cigarettes (THIS; KT&G Corp., Daejeon, Korea) continuously bubbled in 60 mL of phosphate buffered saline using a bottle system connected to a vacuum machine (Gast Manufacturing Inc., Benton Harbor, MI, USA). To determine the optimal concentration of CSE, cultured bronchial epithelial cells were treated with 0, 0.5, 1, 2, or 4% CSE for 24 hours and then the cell viability was measured with an MTT assay. Based on these results, cells exposed to 2% CSE-containing culture medium were compared to those treated with control medium for 24 hours before harvest were used.

Mass spectrometry

Cultured primary cells, either treated or untreated with 2% CSE, were detached with TrypLE (Thermo Fisher Scientific, Waltham, MA, USA) and centrifuged at 300 g for 5 minutes. From each sample, cells were lysed using 100 µL of buffer (4% sodium dodecyl sulfate [SDS], 2 mM tris (2-carboxyethyl) phosphine in 100 mM Tris; pH 8.5) and incubated at 95°C for 20 minutes. Protein concentrations were measured using a reducing agent-compatible BCA assay (Thermo Fisher Scientific). After acetone precipitation of 100 µg of proteins, precipitated protein samples were digested using the filter-aided sample preparation procedure as previously described.[11] Briefly, samples were dissolved with 2% SDS and 2 mM tris (2-carboxyethyl) phosphine in 100 mM Tris/HCl, pH 7.5, mixed with 0.3 mL of 8 M urea in 0.1 M Tris/HCl, pH 8.5, and loaded onto a 30 K Amicon filter (EMD Millipore, Billerica, MA, USA). The buffer was exchanged with urea solution by centrifugation. The reduced cysteines in the proteins were alkylated with iodoacetamide solution in the dark at room temperature for 30 minutes. An additional 40 mM of ammonium bicarbonate was added to exchange the urea solution. Finally, proteins were digested at 37°C overnight with trypsin at an enzyme-to-protein ratio of 1:100. After overnight incubation, the filtration unit was transferred to new collection tubes and centrifuged for 20 minutes. Peptides that were retained in the filtration units were eluted with 50 µL of 0.5 M NaCl to enhance the yield of digested protein. The resulting supernatants were acidified with 1% trifluoroacetic acid. Peptides were desalted and fractionated on homemade styrene divinylbenzene reversed-phase sulfonate-StageTips[12] by basic reverse-phase using a stepwise gradient of acetonitrile (40%, 60%, and 80%) in 1% ammonium hydroxide.

Liquid chromatography-MS/MS analysis was performed using a Q Exactive HF-X Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled to an Ultimate 3000 RSLC system (Dionex, Sunnyvale, CA, USA) via a nano-electrospray source as previously described with modifications.[12,13] Peptide samples were separated on a two-column system, consisting of a trap column and an analytic column (75 µm × 50 cm) with a 90 minutes gradient from 7% to 32% acetonitrile at 300 nL/min and analyzed via MS. The column temperature was maintained at 60°C using a column heater. Survey scans (300–1,650 m/z) were acquired with a resolution of 70,000 at m/z 200. The top-15 method was used to select precursor ions with an isolation window of 1.2 m/z. MS/MS spectra were acquired at an higher-energy collisional dissociation-normalized collision energy of 30 with a resolution of 17,500 at m/z 200. The maximum ion injection times for the full scan and MS/MS scans were 20 ms and 100 ms, respectively.

https://doi.org/10.3346/jkms.2023.38.e220

Mass spectra were processed using MaxQuant version 1.6.1.10.[14] MS/MS spectra were searched against the Human Uniprot protein sequence database (December 2014, 88,657 entries) using the Andromeda search engine.[15] Primary searches were performed using a 6-ppm precursor ion tolerance for total protein level analysis. MS/MS ion tolerance was set to 20 ppm. Cysteine carbamidomethylation was used as a fixed modification. N-acetylation of proteins and oxidation of methionine were set as the variable modifications. Enzyme specificity was set to full tryptic digestion. Peptides with a minimum length of six amino acids and up to two missed cleavages were considered. The required false discovery rate was set to 1% at the peptide, protein, and modification levels. We also enabled the "Match between Runs" option on the MaxQuant platform to maximize the number of quantification events across samples. For label-free quantification (LFQ), the intensity-based MaxLFQ algorithms[16] were used as a part of the MaxQuant platform.

Differentially expressed peptide (DEP) analysis

MaxQuant data were processed and analyzed via the R package 'DEP' version 1.12.0, using default parameters. Briefly, LFQ intensities were normalized using variance-stabilizing transformation and missing data were imputed using random draws from a Gaussian distribution centered around a minimal value with a q-value of 0.05. Differential enrichment analysis was performed using linear and empirical Bayesian statistics. Significant proteins were denoted based on the following criteria: *P* value < 0.05 and log fold change > 1.5.

Feature selection by machine learning method

Next, we used an ML-based approach to select features from the raw LFQ intensity data to investigate the effects of various environmental factors on cells from COPD patients. To differentiate proteins that are enriched when exposed to CSE, we used Extra Trees Classifier, a variation of the random forest.

First, we trained Extra Trees Classifiers with raw LFQ data to distinguish the features of CSE-treated samples in the non-COPD group from the matched non-COPD samples. We interpreted the feature importance of the trained model, denoted as set A (**Fig. 1**), as the featured proteomic response to CSE. Next, the proteomic response of the COPD group was computed using Extra Trees Classifier trained on CSE-treated and non-CSE treated samples in the COPD group, denoted as set B (**Fig. 1**). Then, we subtracted the proteomic response of the non-COPD group from the COPD group and selected the top 100 proteins with the highest differential response. Pathway analysis was performed using Enrichr[17] with the Gene Ontology gene set. All pathways were sorted on the adjusted *P* value. Z-score and combined score, log(p)·z, are adopted from Enrichr.

Validation with single cell RNA sequencing data

Data from the Gene Expression Omnibus accession number GSE136831 representing single-cell RNA sequencing data, including data obtained from 18 lungs from patients with COPD and from 28 control donor lungs, were re-analyzed.[18] Only cells annotated in metadata were included, and cell types with at least 100 cells in both COPD and control groups were used for further analysis.

Differentially expressed genes (DEGs) between COPD and non-COPD samples were analyzed for each cell type using the Wilcoxon rank sum test. Gene set enrichment analysis was performed using the Gene Ontology Biological Process and Kyoto Encyclopedia of Gene and Genomes human databases under the following criteria: adjusted *P* value < 0.05, log

https://doi.org/10.3346/jkms.2023.38.e220

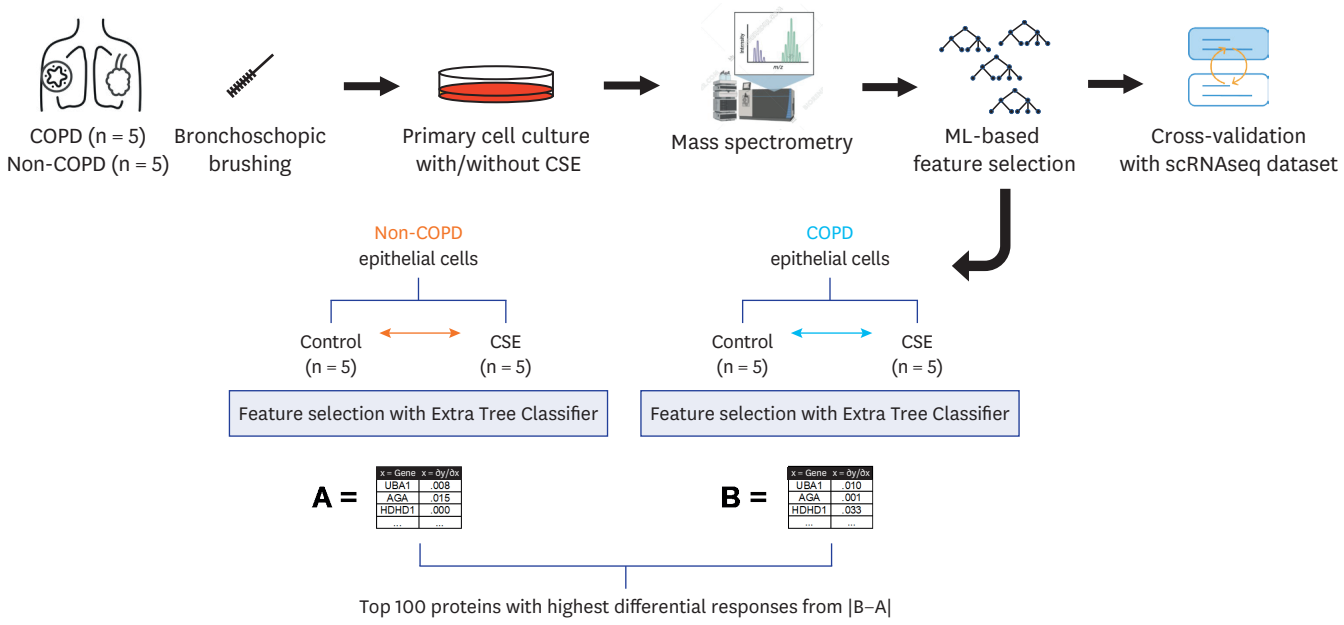

Fig. 1. Schematic of the study protocol with detailed ML-based feature selection.
CSE = cigarette smoking extract, ML = machine learning, COPD = chronic obstructive pulmonary disease.

fold-change > 1.5. An adjusted *P* value is a Bonferroni-corrected *P* value to avoid multiple comparison problem.

Ethics statement

The collection and use of human specimens was reviewed and approved by the Institutional Review Board (IRB) of Seoul National University Hospital, IRB number: H-2105-032-1216. Informed consent was submitted by all subjects when they were enrolled.

RESULTS

Five patients with COPD and five without COPD were included in the analysis. All COPD patients were men and had lower pulmonary function than the non-COPD participants (**Table 1**). Epithelial cells were successfully cultured from bronchoscopic brushing specimens obtained from all participants without complications. The MTT assay showed that 2% or less CSE did not affect cell viability but 4% induced cell death (**Supplementary Fig. 1**). Cultured

Table 1. Basic characteristics of participants

Characteristics	COPD patients (n = 5)	Non-COPD participants (n = 5)
Demographic		
Age, yr	65 (57–84)	65 (48–71)
Male	5 (100)	1 (20)
Smoking, pack-years	60 (1.5–80)	0 (0–0)
Post-bronchodilator spirometry		
FVC, %	78 (74–115)	90 (105–111)
FEV_1, %	56 (33–109)	113 (99–124)
FEV_1/FVC,	46.5 (29–68)	78 (74–83)

Continuous variables are expressed as medians (ranges), and categorical variables are expressed as number (%).
COPD = chronic obstructive pulmonary disease, FVC = forced vital capacity, FEV_1 = forced expiratory volume in one second.

https://doi.org/10.3346/jkms.2023.38.e220

cells were treated as either the control (medium only) or with 2% CSE for 24 hours and employed for proteomic analysis (**Fig. 1**).

A total of 125,724 peptides, including 98,831 unique peptides, were characterized using MS, and subsequently, 7,953 proteins were identified. Initially, conventional analysis of DEPs was performed. A total of 156 DEPs were identified between CSE treated vs. non-CSE treated from non-COPD samples, and 213 DEPs from the corresponding samples derived from patients with COPD (**Fig. 2**, **Supplementary Table 1**). Only 70 DEPs were shared between these two DEPs above. Ferroptosis-related proteins, including ferritin heavy chain 1 (FTH1) and ferritin light chain (FTL), were identified in this set of overlapping DEPs, however, there were no significantly enriched pathways from these 70 DEPs. In addition, 106 DEPs were found between non-CSE treated COPD and non-COPD samples whereas 82 DEPs were found between CSE-treated COPD and non-COPD samples. There were no ferroptosis-related proteins identified from these DEPs (**Fig. 2**).

To find the proteins that are most affected upon CSE exposure and to compare them between COPD and non-COPD directly, first we used an Extra Trees Classifier with the raw LFQ intensity data from COPD and non-COPD, separately. This ML-based algorithm calculated the importance of each protein in the process of classification between CSE-treated vs. non-CSE treated group. Set A was the feature importance of each protein from the non-COPD group, and set B was from COPD group (**Fig. 1**). Next, the top 100 proteins showing the highest differences between the COPD and non-COPD groups were selected (**Supplementary Table 2**), and a pathway analysis was performed. Interestingly, ferroptosis-related proteins, including FTH1, FTL, SLC7A11, and TP53, were shown to be the proteins that changed most in expression in response to CSE (adjusted *P* value = 4.172×10^{-6}; **Table 2**); additionally, glutathione (GSH) metabolism was selected as a featured pathway in response to CSE (adjusted *P* value = 4.527×10^{-3}).

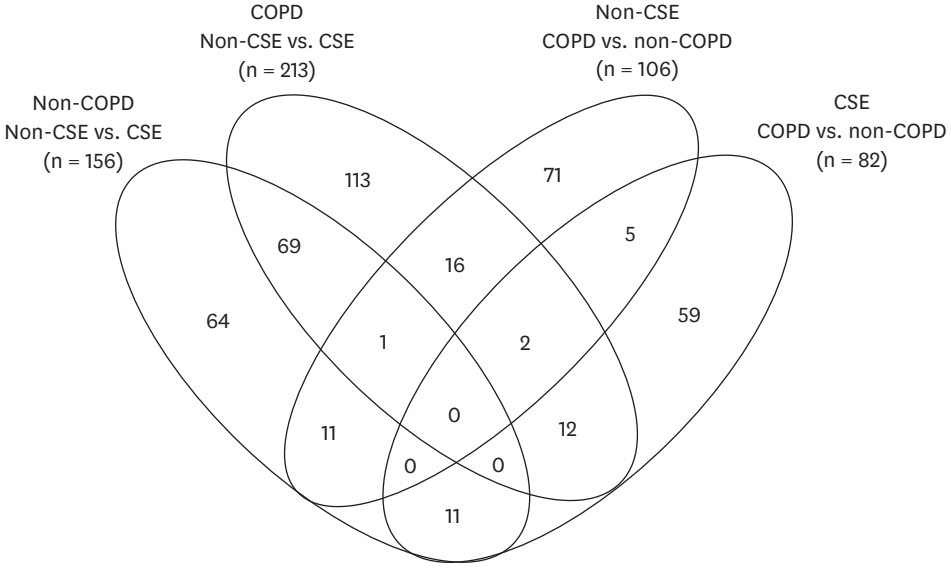

Fig. 2. Venn diagram and numbers of DEPs.
DEP = differentially expressed protein, CSE = cigarette smoking extract, COPD = chronic obstructive pulmonary disease.

https://doi.org/10.3346/jkms.2023.38.e220

Table 2. Featured pathways responsive to smoking exposure in the COPD group compared to the non-COPD group

Pathway	Adjusted *P* value	Combined score
Ferroptosis	8.906×10^{-6}	607
FTH1, FTL, GCLC, GCLM, HMOX1, SLC7A11		
Protein processing in endoplasmic reticulum	7.466×10^{-3}	64
DNAJA1, HSP90AA1, HSPH1, HSP90AB1, SAR1B, HSPA1A		
Fluid shear stress and atherosclerosis	5.870×10^{-3}	91
HMOX1, IL1A, HSP90AA1, HSP90AB1, SUMO1, ARHGEF2		

Adjusted *P* value was from Bonferroni-correction method, combined score, log(p)·z, was adopted from Enrichr.
COPD = chronic obstructive pulmonary disease.

As an external validation of the featured pathways identified in this study, we reanalyzed publicly available single-cell RNA sequencing data from lung parenchyma cells from healthy controls and patients with COPD. Based on DEGs for each cell type, ferroptosis was selected as a featured pathway in basal cells, club cells, and goblet cells but not in ciliated cells (**Table 3**). We hypothesized that the ferroptosis-related signals from these epithelial cells were diluted in the bulk sequencing or proteomics data. Congruently, when we performed the same analysis of single-cell RNA sequencing data with broader cell categories, such as epithelial, myeloid, lymphoid, endothelial, and stromal cells, ferroptosis was not a distinct pathway in any of the other cell types (**Supplementary Table 3**).

Table 3. Featured pathways of chronic obstructive pulmonary disease single cell transcriptomics of epithelial cells

Pathways	Adjusted *P* value	Combined score
Basal cell		
Ferroptosis[a]	1.017×10^{-2}	1,309
Cholesterol metabolism	1.017×10^{-2}	996
Mineral absorption	1.017×10^{-2}	973
Necroptosis	2.410×10^{-2}	229
Club cell		
Ferroptosis[a]	7.903×10^{-2}	223
Ciliated cell		
Ribosome	2.486×10^{-2}	31
Goblet cell		
Prion diseases	2.633×10^{-2}	1,541
Ferroptosis[a]	2.633×10^{-2}	1,310
Cholesterol metabolism	2.633×10^{-2}	996
Mineral absorption	2.633×10^{-2}	973
Legionellosis	2.633×10^{-2}	886
Antigen processing and presentation	2.907×10^{-2}	585
Longevity regulating pathway	2.907×10^{-2}	412
Toxoplasmosis	2.907×10^{-2}	362
Spliceosome	2.907×10^{-2}	292
Estrogen signaling pathway	2.907×10^{-2}	284
Measles	2.907×10^{-2}	281
Necroptosis	2.907×10^{-2}	229
Protein processing in endoplasmic reticulum[a]	2.907×10^{-2}	224
Influenza A	2.907×10^{-2}	214
Endocytosis	3.857×10^{-2}	135
MAPK signaling pathway	4.360×10^{-2}	105
Alveolar type I cell	n.s.	n.s.
Alveolar type II cell	n.s.	n.s.

Adjusted *P* value was from Bonferroni-correction method, combined score, log(p)·z, was adopted from Enrichr.
n.s. = not significantly enriched pathways.
[a]Pathways were also enriched in this study.

https://doi.org/10.3346/jkms.2023.38.e220

DISCUSSION

Airway epithelial cells play vital roles in the maintenance of respiratory homeostasis, including host defense, inflammation control, and tissue repair, all of which are important in the pathogenesis of COPD.[19] COPD is associated with different types of regulatory cell death, including apoptosis[20] and necroptosis[21] in airway epithelial cells. Recent studies have reported that ferroptosis, a newly discovered form of regulated cell death, is also linked to COPD.[22,23] Ferroptosis has been identified in cancer cells and is widely observed in other diseases, including COPD.

In this study, we identified ferroptosis as a distinct signature of smoking exposure in cultured epithelial cells from COPD patients compared to those from non-COPD patients. The ferroptosis-related proteins, including FTH1, FTL, and SLC7A11, were also present in the results of conventional analysis; the DEP lists of CSE vs. non-CSE of COPD and those of non-COPD (**Supplementary Table 1**), which are consistent with the previous report.[24] However, when we compared the effects of CSE treatment between the COPD group and the non-COPD group directly, the ferroptosis-related genes were not shown in the DEP list derived by conventional analysis (**Supplementary Table 1**). However, ML-based analysis provided the feature importance scores during the training process of Extra Trees Classifier, enabling us to directly compare the response to CSE. With our ML-based analysis, we revealed that the ferroptosis-related proteins responded more dramatically in COPD than in non-COPD.

The genes and proteins identified in this study have been reported to be significant signals of COPD in previous studies. A gene expression profile study using microarray analysis showed that FTL is one of 48 genes associated with COPD progression.[25] A MS proteomic study of lung tissues also showed that FTL and FTH1 were two of the 25 proteins significantly associated with COPD.[26] However, since only a few of the ferroptosis markers were present amongst dozens of markers detected in the previous studies, they were understandably dismissed.[25,26] This is mainly because lung tissue has many corresponding immune cells,[18,27] and ciliated cells, another dominant cell type in the airway,[28] are not involved in ferroptosis, according to our analysis. In contrast to other studies, only epithelial cells were enriched in our study; therefore, ferroptosis could be selected as the most important featured pathway of COPD epithelial cells. Indeed, the single-cell RNA sequencing data of COPD and healthy controls showed that only basal cells, club cells, and goblet cells were enriched in ferroptosis, supporting our hypothesis.

In addition, the activity of post-translational modifications that regulate the degradation rate of proteins cannot be detected at the transcriptional level, and this might be another reason why ferroptosis was not distinguished in previous studies. For example, SLC7A11 was detected in this analysis and sustained the production of GSH,[29] a key compound in the ferroptosis pathway. SLC7A11 is known to be regulated by ubiquitination in ferroptosis,[29] however, previous studies did not establish SLC7A11 as a gene differentially expressed in COPD.

Ferroptosis has an iron-dependent mechanism and biological characteristics that are distinct from other regulated cell death processes, such as apoptosis and autophagy. Intracellular GSH depletion and iron accumulation led to an increased level of lipid peroxides through the Fenton reaction, resulting in a large increase in the number of radical oxygen species (ROS), which in turn promotes cell death.[30] An in vitro and in vivo study showed a correlation between ferroptosis in airway epithelial cells and COPD pathogenesis. Cigarette smoke

https://doi.org/10.3346/jkms.2023.38.e220

exposure initiates ferritin-selective autophagy (ferritinophagy), which leads to labile iron deposits, enhanced lipid peroxidation, and concomitant regulated cell death (ferroptosis) that are associated with the COPD phenotype, including emphysema. This CSE-induced ferroptosis-related emphysema was augmented in glutathione peroxidase 4 $(Gpx4)^{+/-}$ mice.[24] GSH and Gpx4 depletion results in ferroptosis, suggesting that ferroptosis is linked to oxidative stress.[31] GSH is depleted in acute smoking, leading to the prevalence of ROS and lung injury.[32] Necrotic cell death via ferroptosis also provides danger-associated molecular patterns that induce and enhance the proinflammatory state.[33] Hence, ferroptosis may contribute to COPD pathogenesis through cell death, increased oxidative stress, and proinflammatory responses.

Our study has limitations; 1) All COPD patients were male whereas there was only one male in the non-COPD participants in this study. This sex difference may cause sex-specific biomarkers to be mis-classified as COPD-specific biomarkers. The prevalence of COPD is higher in males and one of the challenging problems in mining biomarkers in COPD.[34] Although ferroptosis is not thought to be a sex-specific biologic process, the confounding effects on ferroptosis from sex difference cannot be excluded in this study and further study is required. 2) The single cell data are based on tissue from patients with COPD and healthy controls but lacks information on smoking exposure prior to donation of the tissues. A better validation would be possible with larger set of single cell transcriptomics data with a detailed smoking history. However, we consider it is still valid for supporting the importance of ferroptosis in the airway epithelial cells, but not other cell types.

In conclusion, ML-based feature selection revealed that ferroptosis is the most distinctive difference between epithelial cells from patients with COPD vs. non-COPD patients that were exposed to CSE. Our study contributes to the understanding of the pathogenesis of COPD and the development of improved therapeutic strategies for COPD.

ACKNOWLEDGMENTS

We appreciate the participants of this study.

SUPPLEMENTARY MATERIALS

Supplementary Table 1

List of differentially expressed proteins

Click here to view

Supplementary Table 2

The top 100 differentially responded proteins to smoke extract exposure between COPD and non-COPD epithelial cells

Click here to view

Supplementary Table 3

Featured pathways of chronic obstructive pulmonary disease single-cell transcriptomics based on broad cell type definitions

Click here to view

Supplementary Fig. 1

CSE concentration titration for primary human bronchial epithelial cells. Primary HBECs were treated with CSE (0, 0.5, 1, 2, 4%) for 24 hours. MTT cell viability assay was performed. Data represent the mean ± standard deviation.

Click here to view

REFERENCES

1. Adeloye D, Song P, Zhu Y, Campbell H, Sheikh A, Rudan I, et al. 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.
PUBMED | CROSSREF
2. Rhee CK. High prevalence of chronic obstructive pulmonary disease in Korea. *Korean J Intern Med* 2016;31(4):651-2.
PUBMED | CROSSREF
3. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020;396(10258):1204-22.
PUBMED | CROSSREF
4. Park SC, Kim DW, Park EC, Shin CS, Rhee CK, Kang YA, et al. Mortality of patients with chronic obstructive pulmonary disease: a nationwide populationbased cohort study. *Korean J Intern Med* 2019;34(6):1272-8.
PUBMED | CROSSREF
5. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for prevention, diagnosis and management of chronic obstructive pulmonary disease: 2022 report. https://goldcopd.org/. Updated 2022. Accessed September 29, 2022.
6. Buist AS, Vollmer WM, McBurnie MA. Worldwide burden of COPD in high- and low-income countries. Part I. The burden of obstructive lung disease (BOLD) initiative. *Int J Tuberc Lung Dis* 2008;12(7):703-8.
PUBMED
7. Agustí A, Hogg JC. Update on the pathogenesis of chronic obstructive pulmonary disease. *N Engl J Med* 2019;381(13):1248-56.
PUBMED | CROSSREF
8. Liu Y, Liu H, Li C, Ma C, Ge W. Proteome profiling of lung tissues in chronic obstructive pulmonary disease (COPD): platelet and macrophage dysfunction contribute to the pathogenesis of COPD. *Int J Chron Obstruct Pulmon Dis* 2020;15:973-80.
PUBMED | CROSSREF
9. Zhu Y, Zhou A, Li Q. Whole transcriptome analyis of human lung tissue to identify COPD-associated genes. *Genomics* 2020;112(5):3135-41.
PUBMED | CROSSREF
10. Morrow JD, Chase RP, Parker MM, Glass K, Seo M, Divo M, et al. RNA-sequencing across three matched tissues reveals shared and tissue-specific gene expression and pathway signatures of COPD. *Respir Res* 2019;20(1):65.
PUBMED | CROSSREF
11. Han D, Jin J, Woo J, Min H, Kim Y. Proteomic analysis of mouse astrocytes and their secretome by a combination of FASP and StageTip-based, high pH, reversed-phase fractionation. *Proteomics* 2014;14(13-14):1604-9.
PUBMED | CROSSREF

12. Kim JE, Han D, Jeong JS, Moon JJ, Moon HK, Lee S, et al. Multisample mass spectrometry-based approach for discovering injury markers in chronic kidney disease. *Mol Cell Proteomics* 2021;20:100037.
PUBMED | CROSSREF

13. Park J, Kim H, Kim SY, Kim Y, Lee JS, Dan K, et al. In-depth blood proteome profiling analysis revealed distinct functional characteristics of plasma proteins between severe and non-severe COVID-19 patients. *Sci Rep* 2020;10(1):22418.
PUBMED | CROSSREF

14. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 2016;11(12):2301-19.
PUBMED | CROSSREF

15. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 2011;10(4):1794-805.
PUBMED | CROSSREF

16. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* 2014;13(9):2513-26.
PUBMED | CROSSREF

17. Xie Z, Bailey A, Kuleshov MV, Clarke DJ, Evangelista JE, Jenkins SL, et al. Gene set knowledge discovery with Enrichr. *Curr Protoc* 2021;1(3):e90.
PUBMED | CROSSREF

18. Adams TS, Schupp JC, Poli S, Ayaub EA, Neumark N, Ahangari F, et al. Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci Adv* 2020;6(28):eaba1983.
PUBMED | CROSSREF

19. Croasdell Lucchini A, Gachanja NN, Rossi AG, Dorward DA, Lucas CD. Epithelial cells and inflammation in pulmonary wound repair. *Cells* 2021;10(2):339.
PUBMED | CROSSREF

20. Demedts IK, Demoor T, Bracke KR, Joos GF, Brusselle GG. Role of apoptosis in the pathogenesis of COPD and pulmonary emphysema. *Respir Res* 2006;7(1):53.
PUBMED | CROSSREF

21. Pierotti CL, Silke J. Necroptosis in chronic obstructive pulmonary disease, a smoking gun? *Immunol Cell Biol* 2022;100(2):79-82.
PUBMED | CROSSREF

22. Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, et al. Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell* 2012;149(5):1060-72.
PUBMED | CROSSREF

23. Jiang X, Stockwell BR, Conrad M. Ferroptosis: mechanisms, biology and role in disease. *Nat Rev Mol Cell Biol* 2021;22(4):266-82.
PUBMED | CROSSREF

24. Yoshida M, Minagawa S, Araya J, Sakamoto T, Hara H, Tsubouchi K, et al. Involvement of cigarette smoke-induced epithelial cell ferroptosis in COPD pathogenesis. *Nat Commun* 2019;10(1):3145.
PUBMED | CROSSREF

25. Mostafaei S, Kazemnejad A, Azimzadeh Jamalkandi S, Amirhashchi S, Donnelly SC, Armstrong ME, et al. Identification of novel genes in human airway epithelial cells associated with chronic obstructive pulmonary disease (COPD) using machine-based learning algorithms. *Sci Rep* 2018;8(1):15775.
PUBMED | CROSSREF

26. Zhang YH, Hoopmann MR, Castaldi PJ, Simonsen KA, Midha MK, Cho MH, et al. Lung proteomic biomarkers associated with chronic obstructive pulmonary disease. *Am J Physiol Lung Cell Mol Physiol* 2021;321(6):L1119-30.
PUBMED | CROSSREF

27. Travaglini KJ, Nabhan AN, Penland L, Sinha R, Gillich A, Sit RV, et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* 2020;587(7835):619-25.
PUBMED | CROSSREF

28. Yaghi A, Dolovich MB. Airway epithelial cell cilia and obstructive lung disease. *Cells* 2016;5(4):40.
PUBMED | CROSSREF

29. Tang D, Chen X, Kang R, Kroemer G. Ferroptosis: molecular mechanisms and health implications. *Cell Res* 2021;31(2):107-25.
PUBMED | CROSSREF

30. Li J, Cao F, Yin HL, Huang ZJ, Lin ZT, Mao N, et al. Ferroptosis: past, present and future. *Cell Death Dis* 2020;11(2):88.
PUBMED | CROSSREF

31. Ursini F, Maiorino M. Lipid peroxidation and ferroptosis: the role of GSH and GPx4. *Free Radic Biol Med* 2020;152:175-85.
PUBMED | CROSSREF

32. Engelen MP, Orozco-Levi M, Deutz NE, Barreiro E, Hernández N, Wouters EF, et al. Glutathione and glutamate levels in the diaphragm of patients with chronic obstructive pulmonary disease. *Eur Respir J* 2004;23(4):545-51.
PUBMED | CROSSREF

33. Li JY, Yao YM, Tian YP. Ferroptosis: a trigger of proinflammatory state progression to immunogenicity in necroinflammatory disease. *Front Immunol* 2021;12:701163.
PUBMED | CROSSREF

34. DeMeo DL. Sex and gender omic biomarkers in men and women with COPD: considerations for precision medicine. *Chest* 2021;160(1):104-13.
PUBMED | CROSSREF