Multiomics links global surfactant dysregulation with airflow obstruction and emphysema in COPD

Ventzislava A. Hristova[1,15], Alastair Watson[2,3,15], Raghothama Chaerkady[1], Matthew S. Glover[1], Jodie Ackland[2], Bastian Angerman[4], Graham Belfield[5], Maria G. Belvisi[6,7], Hannah Burke[2,8], Doriana Cellura[2], Howard W. Clark[9,10], Damla Etal[5], Anna Freeman[2,8], Ashley I. Heinson[2], Sonja Hess[1], Michael Hühn[4], Emily Hall[2], Alex Mackay[4,7], Jens Madsen[2], Christopher McCrae[11], Daniel Muthas[4], Steven Novick[12], Kristoffer Ostridge[2,4], Lisa Öberg[4], Adam Platt[13], Anthony D. Postle[2], C. Mirella Spalluto[2], Outi Vaarala[14], Junmin Wang[1], Karl J. Staples[2,8] and Tom M.A. Wilkinson[2,8] on behalf of the MICA II Study group[16]

[1]Dynamic Omics, Centre for Genomics Research, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, MD, USA. [2]Faculty of Medicine, University of Southampton, Southampton, UK. [3]School of Clinical Medicine, University of Cambridge, Cambridge, UK. [4]Translational Science and Experimental Medicine, Research and Early Development, Respiratory and Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden. [5]Translational Genomics, Discovery Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden. [6]Research and Early Development, Respiratory and Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden. [7]NHLI, Imperial College, London, UK. [8]NIHR Southampton Biomedical Research Centre, University Hospital, Southampton, UK. [9]Neonatology, Faculty of Population Health Sciences, EGA Institute for Women's Health, University College London, London, UK. [10]NIHR University College London Hospital Biomedical Research, University College London Hospital, London, UK. [11]Translational Science and Experimental Medicine, Research and Early Development, Respiratory and Immunology, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, USA. [12]Data Sciences and Quantitative Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, MD, USA. [13]Translational Science and Experimental Medicine, Research and Early Development, Respiratory and Immunology, BioPharmaceuticals R&D, AstraZeneca, Cambridge, UK. [14]Faculty of Medicine, University of Helsinki, Helsinki, Finland. [15]These authors contributed equally. [16]For a list of the MICA II Study group members see the Acknowledgements.

Corresponding author: Tom M.A. Wilkinson (T.Wilkinson@soton.ac.uk)

Shareable abstract (@ERSpublications)
Multiomics demonstrates global surfactant dysregulation in COPD, associating with emphysema and airway obstruction severity. These findings frame the need for future studies to explore the potential for novel surfactant-targeting therapeutics. https://bit.ly/3TVUDB8

Cite this article as: Hristova VA, Watson A, Chaerkady R, *et al*. Multiomics links global surfactant dysregulation with airflow obstruction and emphysema in COPD. *ERJ Open Res* 2023; 9: 00378-2022 [DOI: 10.1183/23120541.00378-2022].

Received: 26 July 2022
Accepted: 16 Oct 2022

Abstract

Rationale Pulmonary surfactant is vital for lung homeostasis as it reduces surface tension to prevent alveolar collapse and provides essential immune-regulatory and antipathogenic functions. Previous studies demonstrated dysregulation of some individual surfactant components in COPD. We investigated relationships between COPD disease measures and dysregulation of surfactant components to gain new insights into potential disease mechanisms.

Methods Bronchoalveolar lavage proteome and lipidome were characterised in ex-smoking mild/moderate COPD subjects (n=26) and healthy ex-smoking (n=20) and never-smoking (n=16) controls using mass spectrometry. Serum surfactant protein analysis was performed.

Results Total phosphatidylcholine, phosphatidylglycerol, phosphatidylinositol, surfactant protein (SP)-B, SP-A and SP-D concentrations were lower in COPD *versus* controls (log_2 fold change (log_2FC) −2.0, −2.2, −1.5, −0.5, −0.7 and −0.5 (adjusted p<0.02), respectively) and correlated with lung function. Total phosphatidylcholine, phosphatidylglycerol, phosphatidylinositol, SP-A, SP-B, SP-D, napsin A and CD44 inversely correlated with computed tomography small airways disease measures (expiratory to inspiratory mean lung density) (r= −0.56, r= −0.58, r= −0.45, r= −0.36, r= −0.44, r= −0.37, r= −0.40 and r= −0.39 (adjusted p<0.05)). Total phosphatidylcholine, phosphatidylglycerol, phosphatidylinositol, SP-A, SP-B, SP-D and NAPSA inversely correlated with emphysema (% low-attenuation areas): r= −0.55, r= −0.61,

r= −0.48, r= −0.51, r= −0.41, r= −0.31 and r= −0.34, respectively (adjusted p<0.05). Neutrophil elastase, known to degrade SP-A and SP-D, was elevated in COPD *versus* controls (log_2FC 0.40, adjusted p=0.0390), and inversely correlated with SP-A and SP-D. Serum SP-D was increased in COPD *versus* healthy ex-smoking volunteers, and predicted COPD status (area under the curve 0.85).
Conclusions Using a multiomics approach, we demonstrate, for the first time, global surfactant dysregulation in COPD that was associated with emphysema, giving new insights into potential mechanisms underlying the cause or consequence of disease.

Introduction

COPD is a leading cause of morbidity and mortality worldwide. There is still much to be understood about the mechanistic processes underlying its pathology and it drives such an important unmet clinical need [1–4]. Pulmonary surfactant homeostasis is critical to healthy lung function as it coats the air–liquid interface which reduces surface tension and prevents alveolar collapse at end-expiration [5, 6]. Tightly controlled synthesis, secretion and subsequent recycling of surfactant are key to facilitate these essential functions. Emphysematous changes in COPD and loss of alveolar type II (AT2) cells, which produce surfactant, may lead to disrupted surfactant synthesis and homeostasis, and require study.

Pulmonary surfactant is comprised of ∼90% lipids and 10% proteins. Phosphatidylcholine (PC) accounts for >80% of surfactant lipids, with phosphatidylglycerol (PG) for ∼15%, and the remainder is phosphatidylethanolamine, phosphatidylinositol (PI), sphingomyelin and other lipids [7–10]. Surfactant proteins (SP)-B and SP-C are small hydrophobic proteins with essential biophysical roles in surfactant packaging, recycling and maintaining surfactant structure [5, 11]. SP-B is essential for reducing surface tension and its production is regulated by napsin A [11, 12]. SP-C regulation is not fully understood, but SP-B, cathepsin H (CTSH) and Nedd4 have been suggested to facilitate its production [13–15]. In contrast, SP-A, composed of SP-A1 and SP-A2, and SP-D are large, soluble, innate immune defence molecules with essential immunomodulatory and homeostatic lung functions [16–20]. These prevent infection and help clear bacterial, viral and fungal pathogens, whilst preventing aberrant inflammation and damage to the delicate epithelial–endothelial barrier [19–22].

Surfactant dysregulation may play a role in the pathological processes underlying COPD through changes in alveolar tension and development of emphysema [23]. In addition, SP-D, specifically, has long been known to be deficient in COPD, which may predispose to both exacerbations and inflammatory processes [24]. However, there are contradicting reports around pulmonary levels of SP-A in COPD, and levels of SP-B and SP-C remain to be fully elucidated [25–27]. A recent study reported surfactant lipids to be in lower abundance in a small cohort of COPD subjects as compared with nonsmoking controls [28]; however, the impact of disease on surfactant in the absence of current smoking remains to be elucidated. We used an unbiased comprehensive multiomics approach to characterise proteome and lipidome differences in bronchoalveolar lavage fluid (BAL) in well-characterised COPD subjects and healthy ex-smoking controls to better understand surfactant dysregulation in COPD and glean insights about potential mechanisms underlying the cause or consequence of disease.

Methods

Subjects

The MICA II study recruited subjects with mild or moderate COPD [29], alongside healthy ex-smoking volunteers (HV-ES); all had ⩾10 pack year history, but had stopped smoking ⩾6 months prior to enrolment [30–33]. Healthy volunteer never-smokers (HV-NS) were also recruited. All MICA II study subjects with recovered BAL supernatants suitable for proteomic and lipidomic analysis were included in this study (demographics given in table 1). Total subject numbers per group, therefore, differ slightly from previous publications on the MICA II study. Further details about this cohort have previously been reported [30–33]. Matched serum was also used for complementary proteomic analysis. This included both participants within the main cohort and some additional participants who were removed from the study prior to bronchoscopy due to numerous reasons, including subject request, not being suitable for bronchoscopy or not fitting the inclusion criteria as set out in the methodology (demographics given in table S1). Subjects were recruited from a combination of sources, including established research databases held within the University Hospital Southampton, contact by clinicians involved or aware of the study within the hospital and local healthcare facilities, and through subjects responding to study adverts/posters. All subjects gave written informed consent. The study was approved by National Research Ethics Service South Central – Hampshire A and Oxford C Committees (LREC no: 15/SC/0528).

As previously described, all subjects underwent volumetric computed tomography (CT) chest scans in full inspiration and maximum expiration using a Siemens Sensation 64 scanner [32]. Low-attenuation areas

TABLE 1 Demographics of healthy volunteer ex-smoker (HV-ES) controls compared with COPD

	Control			COPD	p-value (HV-ES *versus* COPD)
	HV-NS	HV-ES	p-value (HV-NS *versus* HV-ES)		
Subjects[#]	16	20		26	
Males/females	9/7	11/9	0.9402	20/6	0.1159
Age, years	63.5 (9.5)	67.5 (6.75)	0.0871	70.0 (9.75)	>0.9999
Smoking exposure, pack-years	0.1 (1.6)	25.0 (18.6)	**<0.0001**	40.5 (37.8)	0.3009
BMI, kg·m^{-2}	27.6 (4.6)	27.7 (3.6)	>0.9999	28.5 (5.4)	>0.9999
FEV_1, % of predicted	104.5 (13.5)	100.5 (11.75)	>0.9999	74.0 (14.75)	**<0.0001**
FEV_1/FVC ratio	79.5 (5.0)	77.5 (4.5)	0.8400	58.5 (15.0)	**<0.0001**
T_{LCO}, % of predicted	95.5 (15.5)	89.5 (9.25)	0.9127	73.0 (23.0)	**0.0196**
HRCT %LAA	5.32 (4.165)	5.86 (4.98)	0.8877	11.96 (8.68)	**0.0062**
HRCT E/I MLD	0.800 (0.048)	0.800 (0.060)	>0.9999	0.870 (0.080)	**0.0034**
ICS use, n (%)	0 (0)	0 (0)		14 (53.65)	**<0.0001**
Bronchodilator use, n (%)	0 (0)	1 (5.00)	0.3681	20 (76.92)	**<0.0001**

Data are presented as median (interquartile range) unless otherwise indicated. Statistical testing performed using Chi-squared test for categorical variables (sex: male or female) and Kruskal–Wallis with Dunn's *post hoc* test for continuous variables (all other variables). This table is similar to other research previously reported in the MICA II population [30–33]. HV-NS: healthy volunteer ex-smoker; BMI: body mass index; FEV_1: forced expiratory volume in 1 s; FVC: forced vital capacity; T_{LCO}: transfer factor of the lung for carbon monoxide; HRCT: high-resolution computed tomography; %LAA: % low-attenuation areas; E/I MLD: expiratory to inspiratory mean lung density; ICS: inhaled corticosteroids. [#]: N=62.

(LAAs) below −950 Hounsfield units (%LAA) was calculated as a measure of emphysema and prebronchodilator, single-breath diffusion was performed, as per guidelines, with percent predicted transfer coefficient of the lung for carbon monoxide (T_{LCO}) calculated. A surrogate marker for small airways disease was measured using the ratio of expiratory to inspiratory mean lung density (E/I MLD) scans.

Sample collection

Sampling was undertaken using fibreoptic bronchoscopy and BAL was recovered and processed as previously described [30, 31]. Macrophages were sorted by flow cytometry using forward scatter width and forward scatter area, and subsequently CD45, CD163 and human leukocyte antigen-DR expression. Serum was isolated from blood as previously described [34].

Experimental design of analysis

The three groups and other statistically modelled covariates (*e.g.* age and gender) were balanced through statistical D-optimal block design *via* the optBlock() function in the AlgDesign library in R. Each tandem mass tag (TMT) 11-plex contained all three groups and balance was achieved across the 10 plexes.

Proteomics and lipidomics

BAL supernatants were processed using an S-Trap-based method (protifi.com). Proteins were digested with trypsin/lysC (Promega). Resulting peptides were desalted and subjected to TMT (Thermo Fisher Scientific) labelling for 11-plex TMT analysis, according to the manufacturer's instructions. Liquid chromatography–tandem mass spectrometry (LC-MS/MS) analysis was carried out on a Q Exactive HF-X (Thermo Fisher Scientific) mass spectrometer interfaced with a Dionex 3000 RSLCnano (acquisition parameters outlined in supplementary methods). Data analysis was undertaken with Proteome Discoverer 2.3 (Thermo Fisher Scientific) and Mascot (version 2.6.0) using the latest Uniprot human protein database (search parameters outlined in supplementary methods). Total protein levels were consistent across BAL samples and no outliers were observed before normalisation. Protein quantitation was analysed using Perseus software version 1.6.15.0 and protein abundances were normalised to total protein levels [35, 36].

Serum was depleted with HighSelect Top14 Abundant Protein Depletion Resin and digested using the EasyPep 96 MS Prep Kit (Thermo Fisher Scientific), according to manufacturers' instructions. Serum analysis was undertaken by data-independent acquisition (DIA) on an Exploris 480 mass spectrometer interfaced with a Dionex 3000 RSLCnano (Thermo Fisher Scientific) and DIA analysis was performed with Spectronaut v15 (Biognosys) (sample processing, acquisition and analysis parameters outlined in supplementary methods).

Lipid extraction from BAL supernatants were performed using a modified Maytash method [37]. LC-MS/MS analysis was performed on a Vanquish UHPLC–Orbitrap ID-X Tribrid MS (Thermo Scientific).

Lipidomic data was analysed using MS-DIAL version 4 [38]. Detailed parameters of lipidomic experiments performed by LC-MS/MS and data analysis are outlined in supplementary methods.

SP-D ELISA was performed to confirm mass spectrometry results using a rabbit polyclonal anti-recombinant fragment of human SP-D capture antibody and biotinylated mouse anti-human SP-D (Hyb246-04) detection antibody, with streptavidin–horseradish peroxidase [39]. Quantification was through comparison with a recombinant full length human SP-D standard [40].

Bioinformatic and statistical analysis

We fitted a linear model separately for each lipid and lipid class, and a linear mixed model for each protein, to understand differences between COPD and control cohorts, while accounting for effects of age, sex and experimental design. For participants with two BAL samples, the average of the two samples per subject was taken prior to constructing and fitting the models. The model for any lipid was specified as:

$$y_{jk} = \mu_j + \beta_1 \times \text{age}_k + \beta_2 \times \text{sex}_k + e_{jk}$$

where $y_{jk}$was the $\log_2$-transformed abundance or composition of the lipid, the jth group (COPD, HV-ES or HV-NS) and the kth subject (of a certain age and sex). $e_{jk} \sim N(0, \sigma_e^2)$ denoted subject-to-subject variability. The model for a lipid class had the same formula, except that $y_{jk}$was the $\log_2$-transformed summed abundance of the lipids belonging to that class.

For BAL proteomics, TMT batches were added as a random effect. Two samples of the same subject were placed in the same TMT batch except for one subject, which was excluded from model fitting. The model for any protein was specified as:

$$y_{jkl} = \mu_j + \beta_1 \times \text{age}_k + \beta_2 \times \text{sex}_k + T_l + e_{jkl}$$

where all terms were defined the same as above, but with $T_l \sim N(0, \sigma_T^2)$ and $e_{jkl} \sim N(0, \sigma_e^2)$ respectively denoting TMT batch and subject-to-subject variability.

The model coefficients were estimated using the lmer() function in the lme4 R package [41] (R version 3.6.0) [13, 14]. Pairwise comparisons of estimated marginal means were conducted using the emmeans R package. The abundance of a protein/lipid, lipid composition or summed abundance of a lipid class was considered significantly different between two groups if its adjusted p-value (p-value adjusted *via* Benjamini–Hochberg false discovery rate (FDR) method) was <0.05. Box plots were made after effects of age, sex and random effects, if present, were subtracted from the model fits. The Spearman's rank correlation test was conducted to determine the association between two variables (p-value adjusted *via* FDR method). For participants with two BAL samples, the average of the two samples was taken prior to the correlation analysis.

A separate analysis was undertaken to evaluate the predictive value of serum SPs for use as a potential COPD biomarker. This was done through development of a logistic regression model that classifies COPD status based on serum SP-D. The model was trained using data from 29 donor-matched serum samples in which serum SP-D was detected.

Transcriptomics

For subjects where sufficient alveolar macrophages were able to be purified, total RNA was extracted from BAL fluid purified macrophages using the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen), as previously described [33]. Gene expression profiles of isolated macrophages from 47 subjects (15 HV-NS, 18 HV-ES and 14 COPD; demographics given in table S3) were assessed using total RNA sequencing on an Illumina NovaSeq 6000 platform as outlined in the supplementary methods and as previously described [33]. To explore surfactant transcriptomic differences in macrophages, differential gene expression analysis was performed with DESeq2 (version 1.26.0) and weighted gene correlation network analysis [42] was also implemented (full details in supplementary methods).

Results

Subject demographics

BAL analysis was undertaken in 26 COPD subjects, 20 HV-ES and 16 HV-NS, and clinical characteristics are summarised in table 1. There were no significant differences between HV-NS and HV-ES in age, sex,

TABLE 2 Bronchoalveolar lavage cell counts

	Control			COPD	p-value (HV-ES controls *versus* COPD)
	HV-NS	HV-ES	p-Value (HV-NS *versus* HV-ES)		
Macrophages, %	32.93 (22.10)	36.42 (14.84)	0.4538	32.48 (19.63)	0.3003
Neutrophils, %	2.25 (3.45)	1.20 (2.51)	>0.9999	0.88 (3.89)	>0.9999
Eosinophils, %	0.48 (0.44)	0.60 (1.15)	>0.9999	0.22 (0.25)	0.3492
Lymphocytes, %	0.08 (1.20)	1.02 (2.03)	0.0975	0.03 (0.85)	**0.0340**
Epithelial cells, %	63.95 (16.33)	57.4 (10.25)	0.3202	62.18 (15.31)	0.2307
Squamous cells, %	0.16 (0.51)	0.33 (0.60)	>0.9999	0.23 (0.69)	>0.9999

Data are presented as median (interquartile range) unless otherwise indicated. Statistical testing performed using Kruskal–Wallis test with Dunn's *post hoc* test. HV-NS: healthy volunteer never-smokers; HV-ES: healthy volunteer ever-smokers.

body mass index (BMI) or lung function (forced expiratory volume in 1 s (FEV_1) and FEV_1/forced vital capacity (FVC), but expected differences were seen in pack-year history. There were no significant differences between HV-ES and COPD volunteers in age, sex or BMI. However, there were expected significant differences in lung function, CT and physiological measures of emphysema and small airway disease. There were no significant differences in proportion of eosinophils, macrophages or neutrophils between HV-NS and HV-ES or HV-ES and COPD BAL. There was a significantly lower median proportion of lymphocytes in BAL in COPD *versus* HV-ES (1.02% *versus* 0.03%, p=0.034) (table 2).

Surfactant lipid dysregulation in COPD

There were no significant differences in BAL lipid composition between males and females (figure S1A). No differences were seen in total PC, PG and PI lipids or surfactant-specific PC and PG species between HV-ES and HV-NS (figure 1a). However, a lower total concentration of PC, PG and PI were observed in COPD compared to HV-ES, log_2FC of −2.0, −2.2 and −1.5 (all adjusted p<0.0001), respectively (figure 1b). Concentrations of PC, PG and PI were further decreased when compared with HV-NS (figure 1b).

To understand specific BAL surfactant phospholipid differences in COPD, we performed a detailed examination of distinct PC and PG molecular species abundances as a percentage of each phospholipid type. Lower concentrations of desaturated PC 32:0, corresponding to dipalmitoylphosphatidylcholine, were observed in COPD *versus* HV-ES (log_2FC −2.5, adjusted p<0.0001). As a percentage of total PC, PC 32:0 was significantly lower in COPD *versus* HV-ES BAL (log_2FC −0.47, adjusted p=0.001) (figure 1c). As a percentage of total PC abundance, PC 30:0 was not significantly lower in COPD BAL (log_2FC −0.16, adjusted p=0.4080), while PC 34:1 was a higher percentage of total PC in COPD subjects compared to HV-ES (log_2FC 0.28, adjusted p=0.0244) (figure 1c). PG 34:1 and 36:2, as a percentage of total PG, were not significantly different between COPD and HV-ES subject BAL, with log_2FC −0.06 (adjusted p=0.6433) and−0.03 (adjusted p=0.8709), respectively (figure 1d). The percentage of PG 36:1 out of total PG was significantly lower in COPD *versus* HV-ES BAL, with log_2FC −0.29 (adjusted p=0.0322) (figure 1d).

BAL ceramides CER 42:1, CER 40:1, CER 34:1, CER 42:2, CER 38:1 and CER 36:1 were significantly higher in COPD compared to HV-NS, with log_2FC of 0.6, 0.9, 1.2, 0.8, 1.2 and 1.1, respectively (all adjusted p<0.05). Other sphingolipids were not significantly different. No differences in sphingolipids were seen between COPD and HV-ES.

Correlation of lipid surfactant with lung disease measures

We next explored if BAL surfactant dysregulation associated with markers of lung function, CT and physiological measures of emphysema (high %LAA and low T_{LCO} % predicted, respectively) and a surrogate CT marker of small airway disease (E/I MLD). Low BAL surfactant levels were seen in subjects with worse lung function. Total PC, PG and PI BAL levels correlated with both FEV_1 (r=0.52, r=0.56 and r=0.47 respectively; all adjusted p<0.01) and FEV_1/FVC (r=0.68, r=0.73 and r=0.60 respectively; all adjusted p<0.01) (figure 2). Low BAL surfactant concentrations were seen in subjects with the most emphysema. Total BAL PC, PG and PI correlated with T_{LCO} % predicted (r=0.41, r=0.42 and r=0.39, respectively; all adjusted p<0.01), and inversely correlated with %LAA (r= −0.55, r= −0.61 and r= −0.48, respectively; all adjusted p<0.01). E/I MLD negatively correlated with total PC (r= −0.56), total PG (r= −0.58) and total PI (r= −0.45) (all adjusted p<0.01).

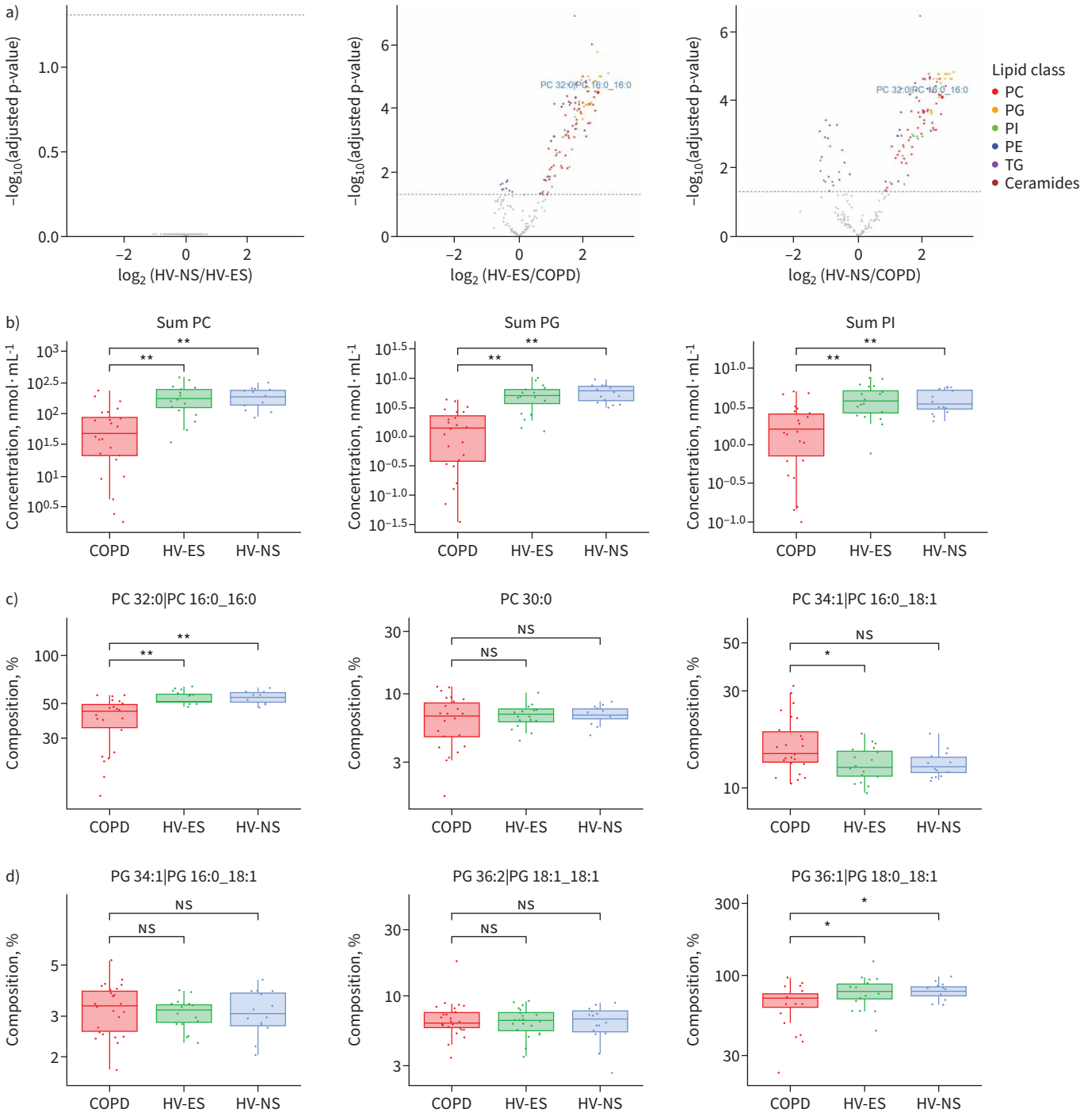


FIGURE 1 Bronchoalveolar lavage lipidomic analysis showed reduced levels of phosphatidylcholine (PC) phospholipids, specifically PC 32:0, dipalmitoyl phosphatidylcholine (DPPC), as well as phosphatidylglycerol (PG) in COPD. a) Volcano plots of lipid abundance in healthy volunteer never-smokers (HV-NS) *versus* healthy volunteer ex-smokers (HV-ES) (left), HV-ES *versus* COPD subjects (middle) and HV-NS *versus* COPD subjects (right). The *x*-axis displays $\log_2$(fold change) and the *y*-axis displays $-\log_{10}$(adjusted p-value). The dashed horizontal line represents an adjusted p-value threshold of 0.05. DPPC is labelled. Lipid classes, including PC, PG, phosphatidylinositols (PI), phosphatidylethanolamines (PE) and triglycerides (TG) are coloured. b) Covariate-adjusted box plots showing the summed abundance of PC, PG and PI compared across COPD and HV-ES/HV-NS cohorts. c) Covariate-adjusted box plot showing the composition of the top three most abundant PC lipids. d) Covariate-adjusted box plots showing the composition of the top three most abundant PG lipids. For details regarding covariate adjustment see the supplementary methods. *: adjusted p<0.05; **: adjusted p<0.01; NS: not significant.

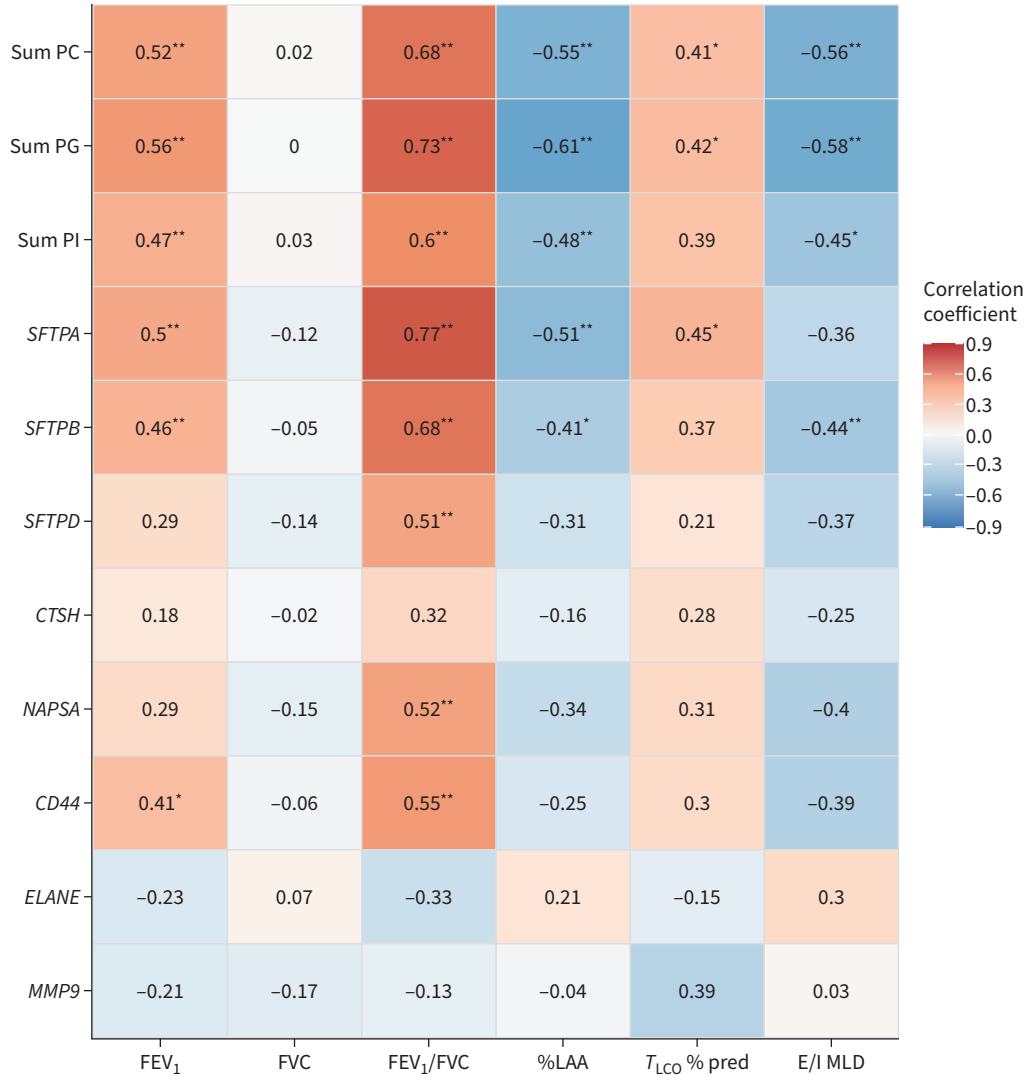


FIGURE 2 Correlation analysis showed correlation between bronchoalveolar lavage fluid phosphatidylcholine (PC), phosphatidylglycerol (PG), phosphatidylinositols (PI), surfactant protein A (SFTPA), surfactant protein B (SFTPB), surfactant protein D (SFTPD) and napsin A (NAPSA), and forced expiratory value in 1 s (FEV_1) to forced vital capacity (FVC) ratio in COPD. The colour of each voxel of the heatmap represents the calculated Spearman's correlation coefficient between a COPD lung function parameter and the abundance of a surfactant protein, the abundance of a surfactant-associated protein or the summed abundance of a lipid category. The *y*-axis displays protein symbols or lipid abbreviations and the *x*-axis displays lung function parameters. CTSH: cathepsin H; ELANE: neutrophil elastase; MMP: matrix metalloproteinase; %LAA: % low-attenuation areas; T_{LCO}: transfer coefficient of the lung for carbon monoxide; E/I MLD: expiratory to inspiratory mean lung density. *: adjusted p<0.05; **: adjusted p<0.01.

SP dysregulation in COPD

There were no significant differences in male *versus* female BAL proteomes (figure S1B). No significant proteome differences, including in pulmonary SP levels, were observed between HV-ES and HV-NS (figure 3a). In contrast, SP-B, SP-A and SP-D BAL levels were significantly lower in COPD *versus* HV-ES (log_2FC −0.5, −0.7 and −0.5; all adjusted p<0.02), respectively (figure 3b). Napsin A and pro-cathepsin H, proteins important in SP-B synthesis [11, 12], were lower in COPD *versus* HV-ES BAL, with log_2FC of −0.6 (adjusted p=0.0017) and −0.4 (adjusted p=0.0466), respectively. BAL SP-B correlated with napsin A abundance (r=0.72, $p<2.2\times10^{-16}$) and cathepsin H levels (r=0.5, $p=4.8\times10^{-5}$) (figure S2A and S2B). CD44 antigen, which has been reported to play a role in surfactant homeostasis [43, 44], was lower in COPD *versus* HV-ES BAL with a log_2FC of −0.6 (adjusted p=0.0025) (figure 3b).

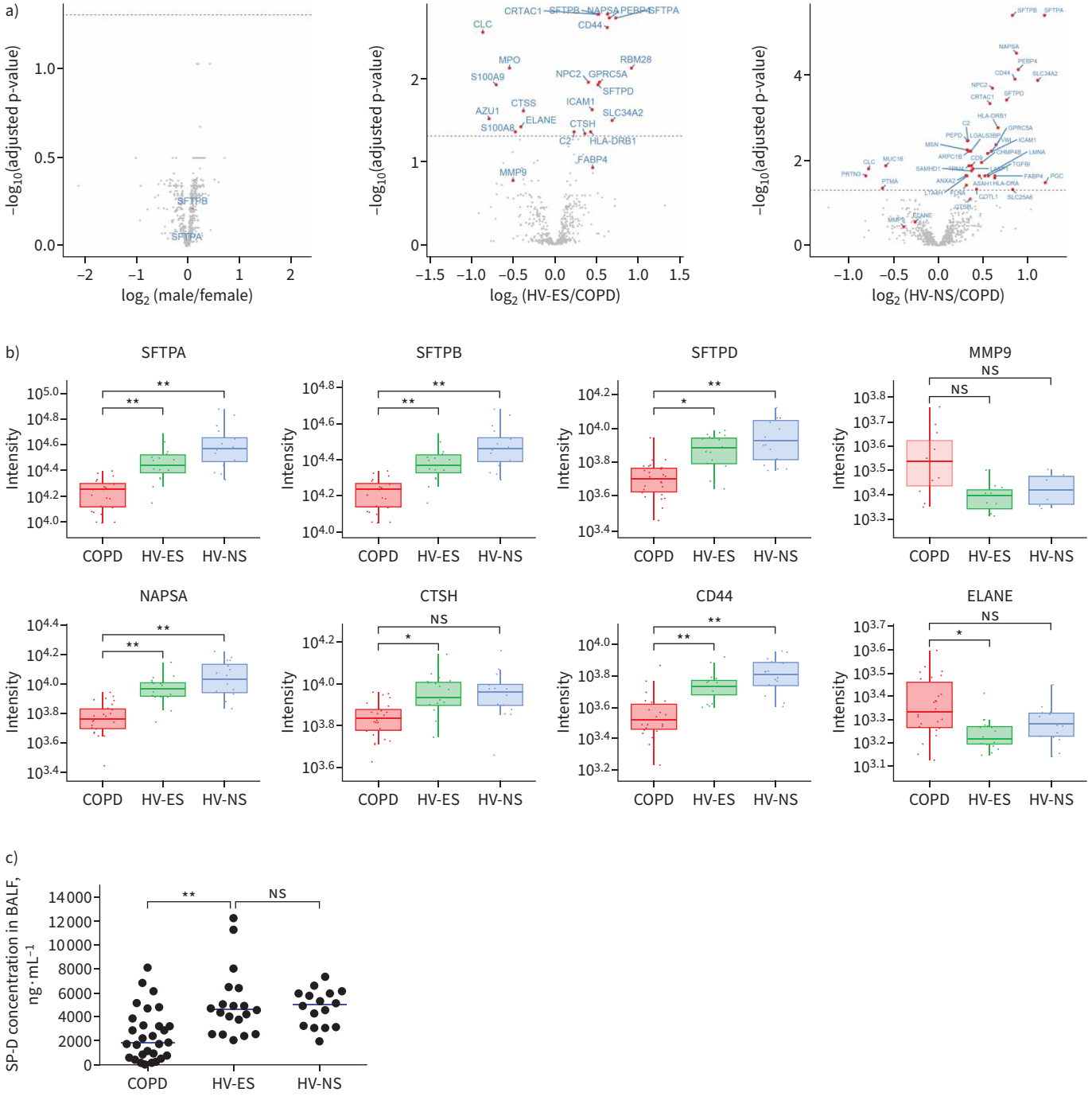


FIGURE 3 Bronchoalveolar lavage fluid (BALF) proteomic analysis showed lower surfactant proteins and proteins involved in surfactant synthesis and secretion in COPD. a) Volcano plots of protein abundance in healthy volunteer non-smoking subjects (HV-NS) *versus* healthy volunteer ex-smoking subjects (HV-ES) (left), HV-ES *versus* COPD subjects (middle) and HV-NS *versus* COPD subjects (right). The *x*-axis displays $\log_2$(fold change) and the *y*-axis displays $-\log_{10}$(adjusted p-value). The dashed horizontal line represents an adjusted p-value threshold of 0.05. Proteins whose abundance is significantly altered in COPD compared to HV-ES and HV-NS donors ($-\log_{10}$(adjusted p)>1.3) are labelled on the volcano plots. These include surfactant and surfactant-associated proteins surfactant protein A (SFTPA), surfactant protein B (SFTPB), surfactant protein D (SFTPD), cathepsin H (CTSH), napsin A (NAPSA), CD44, neutrophil elastase (ELANE) and matrix metalloproteinase (MMP) 9. b) Covariate-adjusted box plots showing the abundance of surfactant proteins SFTPA, SFTPB, SFTPD and NAPSA, CTSH, CD44, ELANE and MMP9 across COPD and HV-ES/HV-NS cohorts. For details regarding covariate adjustment see the supplementary methods. c) ELISA showing concentrations of BALF surfactant protein (SP)-D. ELISA was performed using a rabbit polyclonal anti-recombinant fragment of human SP-D capture antibody and biotinylated mouse anti-human SP-D detection antibody with streptavidin–horseradish peroxidase. Quantification was through comparison with a recombinant full-length human SP-D standard. *: adjusted p<0.05; **: adjusted p<0.01.

Neutrophil elastase, an enzyme known to degrade SP-A and SP-D [45–47], was elevated in COPD *versus* HV-ES BAL (log_2FC 0.40, adjusted p=0.0390) (figure 3b). Furthermore, concentrations of neutrophil elastase inversely correlated with concentrations of SP-A (r= −0.36, p=0.0046) and SP-D (r= −0.33, p=0.0083) (figure S2C and D). No other proteases were significantly different in concentration between COPD and HV-ES BAL.

To validate our findings, we subsequently confirmed our mass spectrometry data of lower BAL SP-D in COPD using ELISA. As expected, SP-D concentrations were significantly lower in COPD *versus* HV-ES, median concentration of 1878.90 *versus* 4684.92 ng·mL^{-1} (p<0.01) (figure 3c). There was no difference in SP-D concentration in BAL between HV-NS and HV-ES.

Correlation of SPs with lung disease measures

SP-A, SP-B, SP-D and CD44 BAL levels correlated with FEV_1 (r=0.50, r=0.46, r=0.29 and r=0.41; all adjusted p<0.05). SP-A, SP-B, SP-D and napsin A BAL levels also correlated with FEV_1/FVC (r=0.77, r=0.68, r=0.51 and r=0.52, respectively; all adjusted p<0.01) (figure 2). BAL SP-A and SP-B correlated with T_{LCO} % predicted (r=0.45 and r=0.37, respectively; both adjusted, p<0.01). Furthermore, BAL levels of SP-A, SP-B, SP-D and napsin A inversely correlated with %LAA (r= −0.51, r= −0.41, r= −0.31 and r= −0.34, respectively; all adjusted p<0.05). Finally, low SP abundance in BAL was associated with a CT measure of small airways disease; E/I MLD negatively correlated with SP-A (r= −0.36), SP-B (r= −0.44), SP-D (r= −0.37), napsin A (r= −0.40) and CD44 (r= −0.39) (all p<0.01).

SP differences in serum: potential of SP-D as a biomarker

To understand if dysregulation of surfactant could be detected in blood, we undertook proteomic analysis of serum samples. Serum analysis were performed in 35 COPD subjects, 22 HV-ES and 19 HV-NS from the MICA II cohort (demographics for included participants given in table S1 and serum protein identifications in table S2). Serum SP-D was significantly higher in COPD *versus* HV-ES (mean SP-D abundance intensities of ∼5000 and ∼3000, respectively; p=0.0095). Significantly more serum SP-D was present in COPD *versus* HV-NS (mean abundance intensities of ∼5000 and ∼2450, respectively; p=0.0030) (figure 4a). SP-D levels in circulation negatively correlated with BAL SP-D abundance in donor-matched samples (r= −0.37, p=0.05) (figure 4b). Serum SP-B was detected in six COPD subjects and two HV-ES but not in HV-NS (figure 4c). Serum cathepsin H was detected in 30 donors but did not show differential abundance in COPD *versus* controls. Serum cathepsin H levels did not correlate with BAL cathepsin H abundance (r= −0.15, p=0.49) (figure 4d).

Notably, serum SP-D levels were able to predict COPD status by the logistic regression model with an area under the curve of 0.85 (figure 4e).

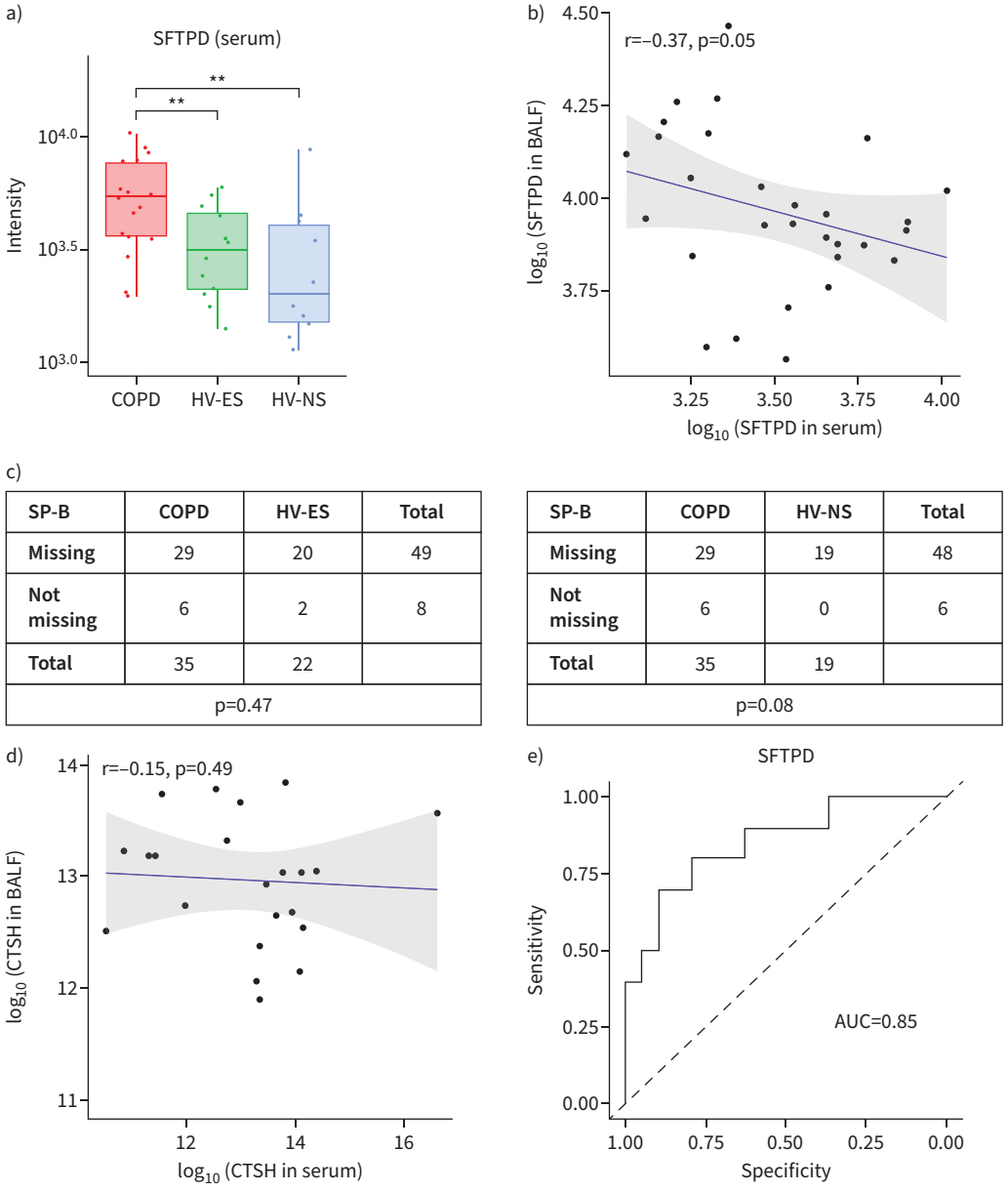
Alveolar macrophage gene expression

A recent study in *Cd44*$^{-/-}$ mice demonstrated that deficiency of CD44 on alveolar macrophages disrupted surfactant lipid homeostasis [44]. We analysed the transcriptome of purified BAL macrophages (demographics for included participants given in table S3) and found no significant activation of transcriptomic pathways involved in macrophage lipid turnover or potentially involved in surfactant metabolism using either differentially expressed gene analysis or WGCNA (figure S3) [48].

Discussion

This study delineates the BAL proteome and lipidome of a well-characterised cohort of mild-to-moderate COPD subjects and ex-smoking and never-smoking controls. Using multiomics, we gained a comprehensive understanding of surfactant dysregulation in COPD, independent of current smoking effects, to glean insights about potential explanatory mechanisms [49].

We report lower concentrations of surfactant lipids, SPs and proteins involved in surfactant synthesis in BAL from COPD subjects *versus* controls, which correlated with airflow obstruction. Furthermore, we demonstrate an association with emphysema, highlighting that decreased surfactant concentrations could be driving mechanisms underlying this pathology or be a consequence of emphysematous changes, or both. Network modelling previously suggested lung surface tension to be important in emphysema pathophysiology through its influence on lung recoil [23]. Our study adds to these findings by demonstrating that surfactant, the key lung surface tension regulator, is decreased in COPD *versus* HV-ES. This expands on prior small studies reporting lower surfactant lipids in BAL from COPD *versus* healthy nonsmokers' and smokers' induced sputum [10, 28, 50].



c)

SP-B	COPD	HV-ES	Total
Missing	29	20	49
Not missing	6	2	8
Total	35	22	
p=0.47			

SP-B	COPD	HV-NS	Total
Missing	29	19	48
Not missing	6	0	6
Total	35	19	
p=0.08			

FIGURE 4 Serum proteomic analysis detected increased surfactant protein D (SFTPD) and surfactant protein B (SFTPB) in COPD patients. a) Box plot of serum SFTPD abundance in healthy non-smoking volunteer subjects (HV-NS), healthy ex-smoking volunteer subjects (HV-ES) and COPD subjects. The *y*-axis is the intensity corresponding to SFTPD for each donor where it was detected. b) Spearman's rank correlation of serum and donor-matched bronchoalveolar lavage fluid (BALF) SFTPD abundance across the cohort. c) Two-by-two contingency table for missing values in SFTPB in relation to disease status (*i.e.* COPD, HV-ES and HV-NS). p-values are obtained from Fisher's exact test, which compares the proportion of missing values in COPD and HV-ES/HV-NS cohorts. The left table compares COPD and HV-ES cohorts, whereas the right table compares COPD and HV-NS cohorts. d) Spearman's rank correlation of serum and donor-matched BAL cathepsin H (CTSH) abundance across cohort. e) receiver operator characteristics-curve of the logistic regression model trained on all donor-matched serum samples. AUC: area under the curve. *: p<0.05; **: p<0.01.

SP-A, SP-B and SP-D single-nucleotide polymorphisms have been shown to associate with COPD, highlighting a potential causative role in COPD pathogenesis [51–54]. We, for the first time, report lower SP-B concentrations in BAL from ex-smoking COPD subjects *versus* matched healthy controls, adding to a prior small study of decreased BAL SP-B in a mixture of COPD current and ex-smokers *versus* healthy

controls [55]. SP-B has key roles in surfactant packaging, recycling and maintaining surfactant structure function, and SP-B deficiency is lethal [12]. Lower SP-B concentrations could therefore have important physiological consequences.

We demonstrate lower BAL SP-A concentrations in COPD, building on two prior contradictory studies in lung tissue from subjects with moderate COPD which report higher or lower SP-A expression *versus* healthy controls [25, 27]. We also confirm lower BAL SP-D concentrations in COPD [24]. SP-A and SP-D have key roles in lung homeostasis through neutralising, opsonising, agglutinating and clearing pathogens, particles and apoptotic cells, whilst preventing aberrant inflammatory pathways and damage to the delicate lung epithelium [19]. Increased utility of these roles in COPD could potentially deplete levels of these proteins [11, 19, 51, 56]. SP-D knock-out mice develop an emphysematous phenotype with impaired surfactant regulation, influx of inflammatory cells and increased apoptotic cells, metalloproteinases and cytokines, highlighting that low SP-D levels could contribute to emphysema pathogenesis. SP-A and SP-D knock-out models demonstrate increased susceptibility to an array of respiratory viruses and bacteria with associated host-mediated inflammation following infection or allergen challenge [19], highlighting the potential importance of our finding of deficient SP-A and SP-D in COPD on risk of infectious exacerbations and inflammation. Delivery of a recombinant SP-D fragment largely resolves the emphysematous phenotype in the SP-D knock-out mice, as well as the susceptibility to respiratory pathogens, raising the potential of this as a novel therapeutic [19, 57, 58].

Surfactant has been proposed to be dysregulated in respiratory diseases through various mechanisms [19]. Damage to AT2 cells through noxious stimuli and AT2 cell loss through emphysematous changes and alteration to lung parenchymal architecture could lead to decreased surfactant production [23]. Our data support this by showing an association between decreased surfactant and emphysema. Due to the nature of bronchoscopy sampling within our study, it was not possible to directly sample the distal airways to look at AT2 gene expression or to correlate transcriptomics of surfactant genes with our findings. Future studies with paired samples taken from resected lung tissue will be important and could add further clarity to the relative contribution of gene expression *versus* other mechanisms involved in our finding of surfactant regulation in COPD.

Inflammation-related damage to the delicate epithelial–endothelial barrier could also lead to surfactant loss through leakage into the blood [24, 59–61]. This aligns with our findings of higher serum SP-D in COPD, as well as the negative correlation between serum and BAL SP-D levels. Serum SP-B was observed in some donors, predominantly COPD subjects and two HV-ES with hiatus hernia, and was absent in HV-NS. Despite its more hydrophilic nature, SP-A was not detected in serum, potentially due to its larger size. Increased serum SP-B and SP-D in COPD, complemented by corresponding BAL findings of reduced SP and lipids, suggest SPs could have utility as lung-specific peripheral biomarkers for COPD. We investigated this through our logistic regression analysis on serum SP-D levels, which had good predictive value for COPD status and demonstrated the potential of SP-D in particular as a COPD biomarker.

We found neutrophil elastase to be increased in COPD *versus* control BAL, and inversely correlated with SP-A and SP-D concentrations. Neutrophil elastase, alongside other host- and pathogen-associated enzymes, degrades SP-A and SP-D, and imbalances could contribute to decreased SP-A and SP-D levels in the COPD lung [19, 46, 47, 62–64]. Potential altered lipid metabolism and surfactant catabolism by alveolar macrophages in COPD could also lead to altered surfactant turnover and dysregulation [11, 19, 51, 56]. In contrast to a recent study of the COPD alveolar macrophage transcriptome [48], we did not see signs of altered lipid metabolism in alveolar macrophage expression signatures. However, that previous study included more severe and predominantly smoking COPD patients, which may explain differences with our observations. We demonstrated higher levels of ceramides in COPD *versus* HV-NS. Ceramides have been reported to influence surfactant production and activity, and could therefore lead to surfactant dysregulation [65]. However, we did not see differences in other sphingolipids, as have previously been reported [66, 67].

We recognise that associations may not indicate causation and that it is impossible to fully rule out other potentially confounding clinical parameters or pathological mechanisms. We normalised our BAL analyses for protein content and performed statistical testing to address potential confounding effects. However, there are additional factors, which may be difficult to completely address. Inhaled therapeutics have been reported to influence surfactant regulation [68], and SP-A, SP-B and SP-D expression have been reported to be increased by corticosteroids [69]. In this study, we saw lower levels of these proteins in COPD subjects, a large proportion of whom were on inhaled therapies. Furthermore, although samples were frozen and stored at −80°C, and were analysed immediately after thawing, we cannot rule out the potential

for endogenous enzymes to have degraded components and influenced our results. Due to intensive study sampling, our well characterised cohort was mild and relatively small, making it impossible to rule out false-negative findings in macrophage lipid metabolism. SP-C was below the limit of detection in both BAL and serum. However, SP-C has previously been reported to be the human lung SP of lowest abundance by weight [70, 71]. Due to cohort heterogeneity and the distinct nature of the various omics datasets, we used an adjusted p-value for the BAL multiomic analysis and trends in surfactant-associated protein. However, lipid downregulation was significantly different between COPD and HV-ES and HV-NS. Notably, the proteome coverage reported here may vary from previous reports due to differences in methodology, analysis stringency, cohort composition, volume instilled into the lungs and sample volume [72, 73]. In our study, we performed TMT mass spectrometry analysis on peptides originating from 25 µL BAL and detected >900 proteins per donor sample. In contrast, Tu *et al*. [72] performed BAL proteome profiling across 20 donors, using 10 mL BAL each and quantified 423 proteins, less than half the proteome coverage we achieved from a 400 times larger sample input. Cohort size is reflective of donor heterogeneity and while studies with small populations may identify more differentially abundant proteins, they have limited statistical power and disease representation. Furthermore, here we report proteomic findings from TMT and label-free mass spectrometry analysis of matched BAL and serum, respectively, that provided a global, unbiased view of proteome dynamics in COPD that may be overlooked by targeted, antibody-based applications, despite potentially higher protein identification rates. Our comprehensive multiomic study used a deeply phenotyped and well-characterised mild–moderate COPD cohort to demonstrate global dysregulation of surfactant in the COPD lung, which was associated with emphysematous changes and airway obstruction. Longitudinal studies in early disease and different COPD endotypes will add further clarity to the causes of surfactant dysregulation, the impact on disease progression and importantly the potential for novel surfactant-replacement and surfactant-targeting therapeutics for the future.

Provenance: Submitted article, peer reviewed.

Acknowledgements: The authors thank all the study volunteers for their contribution towards furthering knowledge about COPD. They also thank the nursing staff in the Southampton Centre for Biomedical Research. The authors thank VIDA for the image analysis which formed part of an academic collaboration. The authors thank Grith L. Sørensen and Uffe Holmskov (University of Southern Denmark) for provision of the biotinylated mouse anti-human surfactant protein-D (Hyb246-04) antibody.

MICA II Study group: Ventzislava A. Hristova (Dynamic Omics, Centre for Genomics Research (CGR), Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, MD, USA), Alastair Watson (Faculty of Medicine, University of Southampton, Southampton, and School of Clinical Medicine, University of Cambridge, Cambridge, UK), Raghothama Chaerkady, Matthew S. Glover (both Dynamic Omics, CGR, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, MD, USA), Jodie Ackland (Faculty of Medicine, University of Southampton, Southampton, UK), Bastian Angerman (Translational Science and Experimental Medicine, Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Graham Belfield (Translational Genomics, Discovery Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Maria G. Belvisi (Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden, and NHLI, Imperial College, London, UK), Hannah Burke (Faculty of Medicine, University of Southampton, and NIHR Southampton Biomedical Research Centre, University Hospital, Southampton, UK), D. Cellura (Faculty of Medicine, University of Southampton, Southampton, UK), Damla Etal (Translational Genomics, Discovery Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Anna Freeman (Faculty of Medicine, University of Southampton, and NIHR Southampton Biomedical Research Centre, University Hospital, Southampton, UK), Ashley I. Heinson (Faculty of Medicine, University of Southampton, Southampton, UK), Sonja Hess (Dynamic Omics, CGR, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, MD, USA), Michael Hühn (Translational Science and Experimental Medicine, Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Alex Mackay (Translational Science and Experimental Medicine, Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, and NHLI, Imperial College, London, UK), Christopher McCrae, Daniel Muthas (both Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Steven Novick (Translational Genomics, Discovery Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Kristoffer Ostridge (Faculty of Medicine, University of Southampton, Southampton, UK, and Translational Science and Experimental Medicine, Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Lisa Öberg (Translational Science and Experimental Medicine, Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca,

Gothenburg, Sweden), Adam Platt (Translational Science and Experimental Medicine, Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Cambridge, UK), Anthony D. Postle, C. Mirella Spalluto (both Faculty of Medicine, University of Southampton, Southampton, UK), Outi Vaarala (Faculty of Medicine, University of Helsinki, Helsinki, Finland), Junmin Wang (Dynamic Omics, CGR, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, USA), Karl J. Staples, Tom M.A Wilkinson (nominated consortia representative) (both Faculty of Medicine, University of Southampton, and NIHR Southampton Biomedical Research Centre, University Hospital, Southampton, UK), Stephanie Ashenden (Translational Genomics, Discovery Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Sarah Bawden (NIHR Southampton Biomedical Research Centre, University Hospital, Southampton, UK), Aurelie Bornot (Translational Genomics, Discovery Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Jerome Bouquet, Carolina Caceres (both Microbial Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Chia-Chien Chiang (Data Science and Artificial Intelligence, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Kerry Day (Faculty of Medicine, University of Southampton, and NIHR Southampton Biomedical Research Centre, University Hospital, Southampton, UK), Antonio DiGiandomenico (Microbial Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Hanna Duàn (Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Vancheswaran Gopalakrishnan (Microbial Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Alex Hicks (Faculty of Medicine, University of Southampton, and NIHR Southampton Biomedical Research Centre, University Hospital, Southampton, UK), Fredrik Karlsson (Translational Genomics, Discovery Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Shameer Khader (Data Science and Artificial Intelligence, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Glenda Lassi (Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Christopher Morehouse (Microbial Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Karl Nordström (Translational Genomics, Discovery Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Esther Nyimbili, Laura Presland, Nicola Rayner, Pedro Rodrigues (all NIHR Southampton Biomedical Research Centre, University Hospital, Southampton, UK), Bret Sellman (Microbial Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Gary Sims (Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Andria Staniford (NIHR Southampton Biomedical Research Centre, University Hospital, Southampton, UK), Paul Warrener (Microbial Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Nicholas P. Williams (Faculty of Medicine, University of Southampton, and NIHR Southampton Biomedical Research Centre, University Hospital, Southampton, UK), Wen Yu, Xiaotao Qu (both Data Science and Artificial Intelligence, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Bairu Zhang, Tianhui Zhang, Natalie van Zuydam (all Translational Genomics, Discovery Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden), Bruce Thompson (Swinburne University of Technology Melbourne, Australia), Ulrika Edvardsson (Business Development and Licensing, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden) and Stephen Harden (University Hospital Southampton NHS Foundation Trust, Southampton, UK).

Availability of data and materials: The datasets generated and analysed during the current study are not publicly available in order to protect the privacy of all individuals whose data we have collected, stored and analysed. However, data may be made available upon reasonable request by applying through the established Data Request Portal through which researchers can request access to de-identified clinical data (https://vivli.org), after which, clinical data may be made available upon review of the patient consent forms and scientific merit of the proposal, and signature of a data sharing/collaboration agreement. This mechanism allows controlled, risk-managed accessibility of the data and at the same time safeguards subjects' confidentiality. Proteomic and lipidomic data will be made available as required by the journal.

Author contributions: V.A. Hristova analysed proteomics, lipidomics and subject data; V.A. Hristova wrote the manuscript with input from all authors and approved the final version of the manuscript. A. Watson conceptualised parts of the project; contributed to methodology, sample collection and processing and project administration; wrote the original draft with input from all authors and approved the final version of the manuscript. R. Chaerkady designed, conducted and analysed the proteomics experiments; R. Chaerkady contributed to writing the manuscript and approved the final version of the manuscript. M.S. Glover designed, conducted and analysed the lipidomics experiments; M.S. Glover contributed to writing of the manuscript and approved the final version of the manuscript. J. Ackland undertook analysis, helped write the original draft, reviewed and edited the manuscript, and approved the final version of the manuscript. B. Angerman analysed data, contributed to writing of the manuscript and approved the final version of the manuscript. G. Belfield contributed to the design, optimisation and analysis of the transcriptomic experiments, contributed to writing of the manuscript, and approved the final version of the manuscript. M.G. Belvisi conceptualised parts of the project, provided resources, contributed to writing of the manuscript and approved the final version of the

manuscript. H. Burke contributed to sample collection methodology, sample collection and the original draft, and approved the final version of the manuscript. D. Cellura contributed sample collection and processing, and approved the final version of the manuscript. H.W. Clark conceptualised parts of the project, and edited and approved the final version of the manuscript. D. Etal contributed sample post-processing methodology and optimisation, conducted the transcriptomic experiments, contributed to writing of the manuscript, and approved the final version of the manuscript. A. Freeman contributed sample collection methodology and sample collection, contributed to the original draft, and approved the final version of the manuscript. E. Hall contributed ELISA analysis, and edited and approved the final version of the manuscript. A.I. Heinson contributed to writing of the manuscript and approved the final version of the manuscript. S. Hess designed the proteomics and lipidomics experiments, and analysed the data; curated subject and medication metadata for subsequent data and statistical analysis; oversaw the omics studies and provided resources; and contributed to writing of the manuscript and approved the final version of the manuscript. M. Hühn conceptualised parts of the project, contributed to analysis and writing the manuscript, and approved the final version of the manuscript. A. Mackay conceptualised parts of the project, contributed to analysis, supervised the project, contributed to writing the manuscript, and approved the final version of the manuscript. J. Madsen conceptualised parts of the project, and edited and approved the final version of the manuscript. C. McCrae conceptualised parts of the project, contributed to analysis and writing of manuscript, and approved the final version of the manuscript. D. Muthas conceptualised parts of the project, contributed to analysis and writing the manuscript, and approved the final version of the manuscript. K. Ostridge conceptualised the project and designed the study, provided resources, supervised the project, contributed to sample collection methodology, sample collection and writing the manuscript, and approved the final version of the manuscript. L. Öberg analysed data, contributed to writing the manuscript and approved the final version of the manuscript. A. Platt conceptualised parts of the project, provided resources, contributed to writing of the manuscript and approved the final version of the manuscript. A.D. Postle contributed to the original draft and approved the final version of the manuscript. C. Mirella Spalluto contributed sample collection methodology, sample collection and processing, and project administration; supervised the project; contributed to writing the manuscript; and approved the final version of the manuscript. O. Vaarala conceptualised parts of the project, contributed to writing the manuscript and approved the final version of the manuscript. J. Wang and S. Novick analysed proteomics and lipidomics data. J. Wang contributed to writing of the manuscript and approved the final version of the manuscript. K.J. Staples conceptualised the project; performed sample collection methodology, sample collection and processing, and project administration; supervised the project and contributed to the original draft; and approved the final version of the manuscript. T.M.A. Wilkinson conceptualised the project and designed the study, provided resources, acquired funding, supervised the project, contributed to writing of the manuscript, and approved the final version of the manuscript.

Conflicts of interest: This project was funded by AstraZeneca. V.A. Hristova, R. Chaerkady, M.S. Glover, B. Angermann, G. Belfield, M.G. Belvisi, D. Etal, S. Hess, M. Hühn, C. McCrae, D. Muthas, S. Novick, K. Ostridge, L. Öberg, A. Platt and J. Wang are employees of AstraZeneca, and hold AstraZeneca employee stocks and/or stock options. A. Mackay was an employee of AstraZeneca during the conduct of the study and an employee of Novartis upon submission of this article; Novartis played no role and made no contribution, financial or otherwise, to the work in this manuscript. O. Vaarala was employee of AstraZeneca from 2014 to 2019, and an employee of OrionPharma from 2019 and during the conduct of this study, and owns AstraZeneca stock. K.J. Staples reports receiving grants from AstraZeneca within the submitted work. T.M.A Wilkinson reports grants and personal fees from AstraZeneca during the conduct of the study; and personal fees and other support from MMH, grants and personal fees from GSK, personal fees from BI, and grants and personal fees from Synairgen, outside the submitted work. A. Watson, J. Ackland, H. Burke, D. Cellura, H.W. Clark, A. Freeman, E. Hall, A.I. Heinson, J. Madsen, A.D. Postle and C. Mirella Spalluto report no conflicts of interest.

Support statement: The study was funded by AstraZeneca. AstraZeneca reviewed the publication, without influencing the opinions of the authors, to ensure medical and scientific accuracy, and the protection of intellectual property. The clinical study was supported by the NIHR Southampton BRC. The corresponding author had access to all data in the study and had the final responsibility for the decision to submit the manuscript for publication. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 WHO. The global burden of disease: 2004 update. https://apps.who.int/iris/handle/10665/43942 Date last accessed: 17 May 2021.
2 Burke H, Wilkinson TMA. Unravelling the mechanisms driving multimorbidity in COPD to develop holistic approaches to patient-centred care. *Eur Respir Rev* 2021; 30: 210041.
3 Kong CW, Wilkinson TMA. Predicting and preventing hospital readmission for exacerbations of COPD. *ERJ Open Res* 2020; 6: 00325-2019.

4 Watson A, Wilkinson TMA. Digital healthcare in COPD management: a narrative review on the advantages, pitfalls, and need for further research. *Ther Adv Respir Dis* 2022; 16: 17534666221075493.
5 Cañadas O, Olmeda B, Alonso A, *et al.* Lipid-protein and protein-protein interactions in the pulmonary surfactant system and their role in lung homeostasis. *Int J Mol Sci* 2020; 21: 3708.
6 Agassandian M, Mallampalli RK. Surfactant phospholipid metabolism. *Biochim Biophys Acta* 2013; 1831: 612–625.
7 Batenburg JJ. Surfactant phospholipids: synthesis and storage. *Am J Physiol* 1992; 262: L367–L385.
8 Holm BA, Wang Z, Egan EA, *et al.* Content of dipalmitoyl phosphatidylcholine in lung surfactant: ramifications for surface activity. *Pediatr Res* 1996; 39: 805–811.
9 Haagsman HP, van Golde LM. Synthesis and assembly of lung surfactant. *Annu Rev Physiol* 1991; 53: 441–464.
10 Agudelo CW, Samaha G, Garcia-Arcos I. Alveolar lipids in pulmonary disease. A review. *Lipids Health Dis* 2020; 19: 122.
11 Wright JR. Immunoregulatory functions of surfactant proteins. *Nat Rev Immunol* 2005; 5: 58–68.
12 Dunbar AE, Wert SE, Ikegami M, *et al.* Prolonged survival in hereditary surfactant protein B (SP-B) deficiency associated with a novel splicing mutation. *Pediatr Res* 2000; 48: 275–282.
13 Possmayer F, Nag K, Rodriguez K, *et al.* Surface activity *in vitro*: role of surfactant proteins. *Comp Biochem Physiol A Mol Integr Physiol* 2001; 129: 209–220.
14 Perez-Gil J. Structure of pulmonary surfactant membranes and films: the role of proteins and lipid-protein interactions. *Biochim Biophys Acta* 2008; 1778: 1676–1695.
15 Conkright JJ, Apsley KS, Martin EP, *et al.* Nedd4-2-mediated ubiquitination facilitates processing of surfactant protein-C. *Am J Respir Cell Mol Biol* 2010; 42: 181–189.
16 Fakih D, Pilecki B, Schlosser A, *et al.* Protective effects of surfactant protein D treatment in 1, 3–β-glucan-modulated allergic inflammation. *Am J Physiol Lung Cell Mol Physiol* 2015; 309: L1333–L1343.
17 Littlejohn JR, da Silva RF, Neale WA, *et al.* Structural definition of hSP-D recognition of Salmonella enterica LPS inner core oligosaccharides reveals alternative binding modes for the same LPS. *PLoS One* 2018; 13: e0199175.
18 Ujma S, Carse S, Chetty A, *et al.* Surfactant protein a impairs genital HPV16 pseudovirus infection by innate immune cell activation in a murine model. *Pathogens* 2019; 8: 288.
19 Watson A, Madsen J, Clark HW. SP-A and SP-D: dual functioning immune molecules with antiviral and 1 immunomodulatory properties. *Front Immunol* 2020; 11: 622598.
20 Watson A, Phipps MJ, Clark HW, *et al.* Surfactant proteins A and D: trimerized innate immunity proteins with an affinity for viral fusion proteins. *J Innate Immun* 2019; 11: 13–28.
21 Nayak A, Dodagatta-Marri E, Tsolaki A, *et al.* An insight into the diverse roles of surfactant proteins, SP-A and SP-D in innate and adaptive immunity. *Front Immunol* 2012; 3: 131.
22 Watson A, Kronqvist N, Spalluto CM, *et al.* Novel expression of a functional trimeric fragment of human SP-A with efficacy in neutralisation of RSV. *Immunobiology* 2017; 222: 111–118.
23 Ingenito EP, Tsai LW, Majumdar A, *et al.* On the role of surface tension in the pathophysiology of emphysema. *Am J Respir Crit Care Med* 2005; 171: 300–304.
24 Winkler C, Atochina-Vasserman EN, Holz O, *et al.* Comprehensive characterisation of pulmonary and serum surfactant protein D in COPD. *Respir Res* 2011; 12: 29.
25 Ohlmeier S, Vuolanto M, Toljamo T, *et al.* Proteomics of human lung tissue identifies surfactant protein A as a marker of chronic obstructive pulmonary disease. *J Proteome Res* 2008; 7: 5125–5132.
26 Bækvad-Hansen M, Nordestgaard BG, Dahl M. Surfactant protein B polymorphisms, pulmonary function and COPD in 10,231 individuals. *Eur Respir J* 2011; 37: 791–799.
27 Liu Z, Chen S, Xu Y, *et al.* Surfactant protein A expression and distribution in human lung samples from smokers with or without chronic obstructive pulmonary disease in China. *Medicine (Baltimore)* 2020; 99: e19118.
28 Agudelo CW, Kumley BK, Area-Gomez E, *et al.* Decreased surfactant lipids correlate with lung function in chronic obstructive pulmonary disease (COPD). *PLoS One* 2020; 15: e0228279.
29 Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for Prevention, Diagnosis and Management of COPD: 2022 report. Available from: www.goldcopd.org.
30 Day K, Ostridge K, Conway J, *et al.* Interrelationships among small airways dysfunction, neutrophilic inflammation, and exacerbation frequency in COPD. *Chest* 2021; 159: 1391–1399.
31 Watson A, Spalluto CM, McCrae C, *et al.* Dynamics of IFN-β responses during respiratory viral infection. Insights for therapeutic strategies. *Am J Respir Crit Care Med* 2020; 201: 83–94.
32 Ostridge K, Gove K, Paas KHW, *et al.* Using novel computed tomography analysis to describe the contribution and distribution of emphysema and small airways disease in chronic obstructive pulmonary disease. *Ann Am Thorac Soc* 2019; 16: 990–997.
33 Watson A, Öberg L, Angermann B, *et al.* Dysregulation of COVID-19 related gene expression in the COPD lung. *Respir Res* 2021; 22: 164.

34 Burke H, Freeman A, Cellura DC, *et al.* Inflammatory phenotyping predicts clinical outcome in COVID-19. *Respir Res* 2020; 21: 245.
35 Tyanova S, Temu T, Sinitcyn P, *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 2016; 13: 731–740.
36 Mao S, Chaerkady R, Yu W, *et al.* Resistance to pyrrolobenzodiazepine dimers is associated with SLFN11 downregulation and can be reversed through inhibition of ATR. *Mol Cancer Ther* 2021; 20: 541–552.
37 Matyash V, Liebisch G, Kurzchalia TV, *et al.* Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *J Lipid Res* 2008; 49: 1137–1146.
38 Tsugawa H, Ikeda K, Takahashi M, *et al.* A lipidome atlas in MS-DIAL 4. *Nat Biotechnol* 2020; 38: 1159–1163.
39 Leth-Larsen R, Nordenbaek C, Tornoe I, *et al.* Surfactant protein D (SP-D) serum levels in patients with community-acquired pneumonia. *Clin Immunol* 2003; 108: 29–37.
40 Leth-Larsen R, Garred P, Jensenius H, *et al.* A common polymorphism in the *SFTPD* gene influences assembly, function, and concentration of surfactant protein D. *J Immunol* 2005; 174: 1532–1538.
41 Clements JA. Surface tension of lung extracts. *Proc Soc Exp Biol Med* 1957; 95: 170–172.
42 Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008; 9: 559.
43 Xiang X, Yuan F, Zhao J, *et al.* Deficiency in pulmonary surfactant proteins in mice with fatty acid binding protein 4-Cre-mediated knockout of the tuberous sclerosis complex 1 gene. *Exp Physiol* 2013; 98: 830–841.
44 Dong Y, Arif AA, Guo J, *et al.* CD44 loss disrupts lung lipid surfactant homeostasis and exacerbates oxidized lipid-induced lung inflammation. *Front Immunol* 2020; 11: 29.
45 Bratcher PE, Weathington NM, Nick HJ, *et al.* MMP-9 cleaves SP-D and abrogates its innate immune functions in vitro. *PLoS One* 2012; 7: e41881.
46 Rubio F, Cooley J, Accurso FJ, *et al.* Linkage of neutrophil serine proteases and decreased surfactant protein-A (SP-A) levels in inflammatory lung disease. *Thorax* 2004; 59: 318–323.
47 Hirche TO, Crouch EC, Espinola M, *et al.* Neutrophil serine proteinases inactivate surfactant protein D by cleaving within a conserved subregion of the carbohydrate recognition domain. *J Biol Chem* 2004; 279: 27688–27698.
48 Fujii W, Kapellos TS, Bassler K, *et al.* Alveolar macrophage transcriptomic profiling in COPD shows major lipid metabolism changes. *ERJ Open Res* 2021; 7: 00915-2020.
49 Otto-Verberne C, Ten Have-Opbroek A, Willems L, *et al.* Lack of type II cells and emphysema in human lungs. *Eur Respir J* 1991; 4: 316–323.
50 Telenga ED, Hoffmann RF, Ruben TK, *et al.* Untargeted lipidomic analysis in chronic obstructive pulmonary disease. Uncovering sphingolipids. *Am J Respir Crit Care Med* 2014; 190: 155–164.
51 Sorensen GL. Surfactant protein D in respiratory and non-respiratory diseases. *Front Med (Lausanne)* 2018; 5: 18.
52 Silveyra P, Floros J. Genetic variant associations of human SP-A and SP-D with acute and chronic lung injury. *Front Biosci (Landmark Ed)* 2012; 17: 407–429.
53 Guo X, Lin H-M, Lin Z, *et al.* Surfactant protein gene A, B, and D marker alleles in chronic obstructive pulmonary disease of a Mexican population. *Eur Respir J* 2001; 18: 482–490.
54 Guan J, Liu X, Xie J, *et al.* Surfactant protein a polymorphism is associated with susceptibility to chronic obstructive pulmonary disease in Chinese Uighur population. *J Huazhong Univ Sci Technolog Med Sci* 2012; 32: 186–189.
55 Um SJ, Lam S, Coxson H, *et al.* Budesonide/formoterol enhances the expression of pro Surfactant Protein-B in lungs of COPD patients. *PLoS One* 2013; 8: e83881.
56 Dong Q, Wright JR. Degradation of surfactant protein D by alveolar macrophages. *Am J Physiol* 1998; 274: L97–L105.
57 Knudsen L, Ochs M, MacKay R, *et al.* Truncated recombinant human SP-D attenuates emphysema and type II cell changes in SP-D deficient mice. *Respir Res* 2007; 8: 70.
58 Watson A, Sørensen GL, Holmskov U, *et al.* Generation of novel trimeric fragments of human SP-A and SP-D after recombinant soluble expression in E. coli. *Immunobiology* 2020; 225: 151953.
59 Robin M, Dong P, Hermans C, *et al.* Serum levels of CC16, SP-A and SP-B reflect tobacco-smoke exposure in asymptomatic subjects. *Eur Respir J* 2002; 20: 1152–1161.
60 Leung JM, Mayo J, Tan W, *et al.* Plasma pro-surfactant protein B and lung function decline in smokers. *Eur Respir J* 2015; 45: 1037–1045.
61 Papaioannou AI, Konstantelou E, Papaporfyriou A, *et al.* Serum surfactant protein levels in patients admitted to the hospital with acute COPD exacerbation. *Lung* 2018; 196: 201–205.
62 Ostridge K, Williams N, Kim V, *et al.* Relationship between pulmonary matrix metalloproteinases and quantitative CT markers of small airways disease and emphysema in COPD. *Thorax* 2016; 71: 126–132.
63 Ostridge K, Williams N, Kim V, *et al.* Distinct emphysema subtypes defined by quantitative CT analysis are associated with specific pulmonary matrix metalloproteinases. *Respir Res* 2016; 17: 92.
64 University Hospital Southampton. ACCORD Trial Platform. https://www.accord-trial.org/. Date last accessed: 29 June 2021.

65 Tibboel J, Reiss I, de Jongste JC, *et al.* Ceramides: a potential therapeutic target in pulmonary emphysema. *Respir Res* 2013; 14: 96.

66 Koike K, Berdyshev EV, Mikosz AM, *et al.* Role of glucosylceramide in lung endothelial cell fate and emphysema. *Am J Respir Crit Care Med* 2019; 200: 1113–1125.

67 Berdyshev EV, Serban KA, Schweitzer KS, *et al.* Ceramide and sphingosine-1 phosphate in COPD lungs. *Thorax* 2021; 76: 821–825.

68 Sims MW, Tal-Singer RM, Kierstein S, *et al.* Chronic obstructive pulmonary disease and inhaled steroids alter surfactant protein D (SP-D) levels: a cross-sectional study. *Respir Res* 2008; 9: 13.

69 Tillis CC, Huang HW, Bi W, *et al.* Glucocorticoid regulation of human pulmonary surfactant protein-B (SP-B) mRNA stability is independent of activated glucocorticoid receptor. *Am J Physiol* 2011; 300: L940–L950.

70 Weaver TE, Whitsett JA. Function and regulation of expression of pulmonary surfactant-associated proteins. *Biochem J* 1991; 273: 249.

71 Hickling TP, Clark H, Malhotra R, *et al.* Collectins and their role in lung immunity. *J Leukoc Biol* 2004; 75: 27–33.

72 Tu C, Mammen MJ, Li J, *et al.* Large-scale, ion-current-based proteomics investigation of bronchoalveolar lavage fluid in chronic obstructive pulmonary disease patients. *J Proteome Res* 2014; 13: 627–639.

73 Pastor MD, Nogal A, Molina-Pinelo S, *et al.* Identification of oxidative stress related proteins as biomarkers for lung cancer and chronic obstructive pulmonary disease in bronchoalveolar lavage. *Int J Mol Sci* 2013; 14: 3440–3455.