Stratification of asthma by lipidomic profiling of induced sputum supernatant

Joost Brandsma, PhD,[a,b] James P.R. Schofield, PhD,[b,c] Xian Yang, PhD,[d] Fabio Strazzeri, PhD,[e] Clair Barber, PhD,[b] Victoria M. Goss, PhD,[a,b] Grielof Koster, PhD,[a,b] Per S. Bakke, MD,[f] Massimo Caruso, PhD,[g] Pascal Chanez, MD,[h] Sven-Erik Dahlén, MD,[i] Stephen J. Fowler, MD,[j,k] Ildikó Horváth, MD,[l] Norbert Krug, MD,[m] Paolo Montuschi, MD,[n,o] Marek Sanak, PhD,[p] Thomas Sandström, MD,[q] Dominick E. Shaw, MD,[r] Kian Fan Chung, MD,[o] Florian Singer, PhD,[s,t] Louise J. Fleming, MD,[o] Ian M. Adcock, PhD,[o] Ioannis Pandis, PhD,[d] Aruna T. Bansal, PhD,[u] Julie Corfield, MSc,[v] Ana R. Sousa, PhD,[w] Peter J. Sterk, MD,[x] Rubén J. Sánchez-García, PhD,[e] Paul J. Skipp, PhD,[c] Anthony D. Postle, PhD,[a]† Ratko Djukanović, PhD,[a,b]† on behalf of the U-BIOPRED Study Group

[a] Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, United Kingdom

[b] National Institute for Health Research Southampton Biomedical Research Centre, Southampton, United Kingdom

[c] Centre for Proteomic Research, Biological Sciences, University of Southampton, Southampton, United Kingdom

[d] Data Science Institute, Imperial College, London, United Kingdom

[e] Mathematical Sciences, University of Southampton, Southampton, United Kingdom

[f] Department of Clinical Science, University of Bergen, Bergen, Norway

[g] Department of Biomedical and Biotechnological Sciences, University of Catania, Catania, Italy

[h] Department of Respiratory Diseases, Aix-Marseille University, Marseille, France

[i] Institute of Environmental Medicine, Karolinska Institute, Stockholm, Sweden

[j] Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, The University of Manchester, Manchester, United Kingdom

[k] Manchester Academic Health Centre and NIHR Manchester Biomedical Research Centre, Manchester University Hospitals NHS Foundation Trust, Manchester, United Kingdom

[l] Department of Pulmonology, Semmelweis University, Budapest, Hungary

[m] Fraunhofer Institute for Toxicology and Experimental Medicine, Hannover, Germany

[n] Department of Pharmacology, Faculty of Medicine, Catholic University of the Sacred Heart, Rome, Italy

[o] National Heart and Lung Institute, Imperial College, London, United Kingdom

[p] Department of Medicine, Jagiellonian University, Krakow, Poland

[q] Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden

[r] National Institute for Health Research Biomedical Research Unit, University of Nottingham, Nottingham, United Kingdom

[s] Division of Paediatric Respiratory Medicine and Allergology, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland

[t] Division of Paediatric Pulmonology and Allergology, Department of Paediatrics and Adolescent Medicine, Medical University of Graz, Austria

[u] Acclarogen Ltd, St John's Innovation Centre, Cambridge, United Kingdom

[v] Areteva Ltd, Nottingham, United Kingdom

[w] Respiratory Therapy Unit, GlaxoSmithKline, London, United Kingdom

[x] Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

† These authors contributed equally to this work as joint senior authors.

Corresponding author: Joost Brandsma, PhD

The Henry M. Jackson Foundation for the Advancement of Military Medicine

Austere environments Consortium for Enhanced Sepsis Outcomes

6720B Rockledge Dr

Bethesda, MD 20817

United States

Telephone: 0044 74 01568130

Email: JBrandsma@aceso-sepsis.org

Funding: The U-BIOPRED consortium receives funding from the European Union and from the European Federation of Pharmaceutical Industries and Associations as an IMI JU funded project (no. 115010). Additional funding for the analytical equipment was obtained from a Wellcome Trust equipment grant (no. 093500/Z/10/Z).

U-BIOPRED Ethics Board and Study Group: The study was overseen and approved by the U-BIOPRED Ethics Board which was comprised of Pim de Boer (chair), Jan-Bas Prins, Martina Gahlemann, Luigi Visintin, Hazel Evans, Martine Puhl, Lina Buzermaniene, Val Hudson, Laura Bond, Guy Widdershoven and Ralf Sigmund (http://www.europeanlung.org/en/projects-and-research/projects/u-biopred/home). A comprehensive list of members of the U-BIOPRED Study Group has been provided. All members are acknowledged for their help and expertise, without which the study would not have been possible.

Conflict of Interest: Pascal Chanez reports receiving grants and personal fees from Almirall, Boehringer Ingelheim, ALK, GSK, AstraZeneca, Novartis, Teva, and Chiesi; Kian Fan Chung has received honoraria for participating in Advisory Board meetings of GSK, AstraZeneca, Roche, Novartis, Merck, Nocion, Shionogi, and Rickett-Beckinson regarding treatments for asthma, COPD and cough, and has also been renumerated for speaking engagements; Sven-Erik Dahlén reports receiving personal fees from AstraZeneca, Cayman Chemical, GSK, Novartis, Regeneron, Sanofi, and Teva; Ratko Djukanović reports receiving grants from the Innovative Medicines Initiative, UK Medical Research Council, and UK National Institute for Health and Care Research, a grant and personal fees from Novartis, personal fees from TEVA, Sanofi, Boehringer Ingelheim, Synairgen, and Kymab, as well as holding stocks of Synairgen; Louise Fleming reports receiving grants from the Innovative Medicines Initiative during the conduct of the study, as well as personal fees from Novartis, AstraZeneca, and Sanofi; Graham Roberts reports receiving grants from the Innovative Medicines Initiative during the conduct of the study; Florian Singer reports receiving personal fees from Novartis and Vertex Pharmaceuticals, and non-financial support from Chiesi;

Paul Skipp is a director and shareholder of TopMD Precision Medicine Ltd; Peter Sterk is a scientific advisor to, and has an officially non-substantial interest, in SME Breathomix BV; Fabio Strazzeri is a director and shareholder of TopMD Precision Medicine Ltd. None of the disclosed conflicts of interest relate to any of the work presented in this manuscript. The remaining authors declare no Conflicts of Interest.

Abstract

Background: Asthma is a chronic respiratory disease with significant heterogeneity in its clinical presentation and pathobiology. There is need for improved understanding of respiratory lipid metabolism in asthma patients and its relation to observable clinical features.

Objective: To perform a comprehensive, prospective, cross-sectional analysis of the lipid composition of induced sputum supernatant obtained from asthma patients with a range of disease severities, as well as healthy controls.

Methods: Induced sputum supernatant was collected from 211 asthmatic adults and 41 healthy individuals enrolled in the U-BIOPRED study. Sputum lipidomes were characterised by semi-quantitative shotgun mass spectrometry, and clustered using topological data analysis to identify lipid phenotypes.

Results: Shotgun lipidomics of induced sputum supernatant revealed a spectrum of nine molecular phenotypes, highlighting not just significant differences between the sputum lipidomes of asthmatics and healthy controls, but within the asthmatic population as well. Matching clinical, pathobiological, proteomic and transcriptomic data informed on the underlying disease processes. Sputum lipid phenotypes with higher levels of non-endogenous, cell-derived lipids were associated with significantly worse asthma severity, worse lung function, and elevated granulocyte counts.

Conclusion: We propose a novel mechanism of increased lipid loading in the epithelial lining fluid of asthmatics, resulting from the secretion of extracellular vesicles by granulocytic inflammatory cells, which could reduce the ability of pulmonary surfactant to lower surface tension in asthmatic small airways, as well as compromise its role as an immune regulator.

Clinical Implication: Immunomodulation of extracellular vesicle secretion in the lungs may provide a novel therapeutic target for severe asthma.

Capsule Summary: We used lipid phenotyping of induced sputum to stratify a heterogeneous asthma cohort, and propose a novel mechanism of pulmonary surfactant dysregulation by extracellular vesicles secreted in asthmatic airways.

Keywords: Asthma, induced sputum, epithelial lining fluid, pulmonary surfactant, lipidomics, molecular phenotyping, extracellular vesicles, granulocytic inflammation

Abbreviations:

ACQ	Asthma Control Questionnaire
ATII	Alveolar type II
Chol	Cholesterol
CE	Cholesterol ester
Cer	Ceramide
DG	Diglyceride
DPPC	Dipalmitoyl-phosphatidylcholine
ELF	Epithelial lining fluid
EV	Extracellular vesicle
HC	Healthy control

142	HexCer	Hexosyl-ceramide
143	IgE	Immunoglobulin E
144	IPA	Ingenuity Pathway Analysis
145	JT-test	Jonckheere-Terpstra test
146	LC-MS/MS	Liquid chromatography tandem mass spectrometry
147	LPC	Lyso-phosphatidylcholine
148	MDS	Multi-dimensional scaling
149	MMA	Mild-to-moderate asthmatic
150	PC	Phosphatidylcholine
151	PE	Phosphatidylethanolamine
152	PG	Phosphatidylglycerol
153	PI	Phosphatidylinositol
154	PS	Phosphatidylserine
155	SM	Sphingomyelin
156	QC	Quality control
157	SAc/ex	Current or ex-smoking severe asthmatic
158	SAn	Non-smoking severe asthmatic
159	TDA	Topological data analysis
160	TG	Triglyceride

U-BIOPRED	Unbiased Biomarkers for the Prediction of Respiratory Disease Outcomes

Introduction

Asthma is a chronic respiratory disease characterised by recurrent attacks of breathlessness and wheezing, variable airflow limitation and loss of lung function, with airways inflammation and remodelling as the underlying pathobiological processes. The most significant challenge in asthma treatment is its heterogeneity in clinical presentation and underlying pathobiology. A variety of asthma phenotypes have been described to date based on demographic, clinical or pathophysiological characteristics. Amongst these, blood and sputum eosinophilia have been of greatest value for understanding the risk of exacerbations and response to treatments with inhaled corticosteroids and biologics.[1] However, there is still a large unmet need for understanding the underlying disease mechanisms and for finding correlations with specific pathobiological processes or treatment responses in order to provide a clearer delineation of the various disease phenotypes and endotypes.[2-4] Improved stratification along mechanistic lines will open up new directions for targeted drug development and more personalised disease management strategies.

The U-BIOPRED (Unbiased Biomarkers for the Prediction of Respiratory Disease Outcomes) study[5] has employed an 'unbiased' multi-omics systems biology approach to stratify patients with asthma, elucidate biochemical pathways, and define new sets of diagnostic molecular biomarkers.[6-11] The aims of the current study were: to provide a comprehensive analysis of the lipid composition of induced sputum supernatant across the entire disease spectrum, from health to severe asthma; to stratify the heterogeneous U-BIOPRED cohort according to its sputum lipid molecular phenotypes; and to infer mechanisms of lipid biology that are either affected by, or contribute to, the observed phenotypes and their clinical features.

Sputum supernatant comprises a mixture of pulmonary surfactant and soluble material secreted by immune cells and the respiratory epithelium within the lungs, with small quantities of saliva. In healthy adult volunteers the sputum lipidome is dominated by a comparatively restricted number of molecular species, in particular di-saturated phosphatidylcholines, and small amounts of other glycerophospholipids, sphingolipids, glycerolipids and sterols.[12-14] Thus, its lipid composition matches its primary source: pulmonary surfactant secreted by ATII cells in the alveolar epithelium.[15-17] The tight and rapid regulation of lipids, from the cellular to systemic level, combined with their large molecular diversity and involvement in a wide range of intra and inter-cellular signalling pathways, makes them a rich source of molecular biomarkers of disease and a valuable component of systems-based disease phenotyping studies.[18,19] However, despite significant interest in the role of lipids in respiratory diseases, studies of the sputum lipidome remain scarce.[12-14,20-22] We have previously reported on the lipid composition of sputum supernatant in a cohort of 41 healthy adults[14] and now extend our analysis to an additional 211 U-BIOPRED study participants from across the asthma severity spectrum. Semi-quantitative shotgun lipidomic measurements were clustered using topological data analysis (TDA) and complemented with matched clinical, immunoassay and transcriptomic data to inform on the underlying disease processes. This multi-dimensional approach to patient characterisation stratified healthy and asthmatic participants into nine different groups based on their sputum lipid phenotypes, suggesting distinct biological mechanisms that could provide novel targets for asthma therapeutics.

Methods

U-BIOPRED study design and tranSMART data repository

All samples were obtained from the U-BIOPRED cohort recruited in 14 European clinical centres.[5] Processed biological samples from clinical sites were fully blinded and stored in a central biobank (CIGMR Biobank, University of Manchester). All clinical, laboratory and 'omics data collected for U-BIOPRED are hosted on the tranSMART knowledge management platform, and these were only released to study group members upon completion of all 'omics analyses.

Lipid analysis and data processing

Shotgun lipidomics of induced sputum samples (n=252) followed the methodology previously published by us,[14] and a comprehensive description is given in this article's Online Repository at www.jacionline.org. Briefly, lipids were extracted using a modified Bligh-Dyer extraction protocol[23] and characterised by flow injection analysis on a Dionex 3000 ultra-high performance liquid chromatography system (Thermo Scientific Dionex, Sunnyvale, CA, USA), coupled to a MaXis 3G quadrupole time-of-flight mass spectrometer, equipped with an electrospray ionisation source (Bruker Daltonics, Billerica, MA, USA). Measurements were performed in full scan mode for both positive and negative ionisation at *m/z* 300-1000. Fragmentation analyses for lipid identification were performed on pooled QC samples, using the same instrumental setup, but in LC-MS/MS mode with a Waters Acquity C8 column (Waters, Milford, MA, USA). Data-independent product ion scans were acquired over the entire LC run via broadband collision induced dissociation. Precursor and fragment ions were matched by their chromatographic retention time and using well-established fragmentation rules for lipids[24] for identity confirmation.

After removal of ions with <60% detection rate, data were corrected for potential batch effects using the *R* script "SVA ComBat".[25] All ion counts were normalised using synthetic internal standards and the original sample volume to obtain semi-quantitative

results. Because sputum is subject to variable dilution of analytes during sampling and processing,[26,27] data were also normalised to the amount of total lipid.

Data analysis and statistics

In order to identify lipid phenotypes, topological data analysis (TDA)[9,28,29] was used to group participants with comparable sputum lipid profiles. TDA was performed on the AyasdiAI machine intelligence platform (Symphony AyasdiAI, Palo Alto, CA, USA), using a normalised correlation metric combined with MDS lenses. Groups of participants with similar sputum lipid profiles (i.e., sputum lipid phenotypes) were defined within the TDA structure using density mode clustering.[30,31] Statistical significance of trends across the TDA structure was assessed in SPSS Statistics 24 (IBM, Armonk, NY, USA): Jonckheere-Terpstra tests for ordered alternatives were performed for sputum lipids, as well as a range of demographic, clinical and pathobiological measurements, selected blood protein concentrations, and sputum cell pellet gene expression (see below). Trends were considered significant if $p<0.05$, and highly significant if $p<0.001$ (after Bonferroni correction).

Blood protein data

Concentrations of 32 protein markers of inflammation and tissue function were downloaded from the U-BIOPRED database hosted on the tranSMART knowledge management platform. Protein measurements were performed on plasma samples, using Mesoscale Discovery (MSD) electrochemiluminescence (11), or on serum samples, using either of the following immunoassay platforms: Luminex (16), Impact (2), Singulex (1), Elecsys (1) or Immulite (1).

Pathway analysis of sputum cell pellet transcriptomics data

Transcriptomic data from RNA extracts of 97 matching sputum cell pellets were downloaded from the U-BIOPRED dataset[32], and subjected to Ingenuity Pathway Analysis (QIAGEN Bioinformatics, Redwood City, CA, USA) to identify potential upstream regulators of differential gene expression in each of the lipid phenotypes. IPA core analysis was performed on the top 4000 differentially expressed genes, using the ordered TDA groups as described above. The results were subjected to a comparison analysis to identify trends of IPA-predicted upstream regulator activation/inhibition across the sputum lipidomics TDA structure.

Results

Study cohort

Of the 610 adult individuals recruited in U-BIOPRED, 252 successfully provided sputum samples that passed QC based on cell viability, resuspension volume and a squamous epithelial cell cut-off of ≤40% of total sputum inflammatory cells.[5] The study group thus comprised 137 females and 115 males of predominately white origin, clinically categorised as either non-smoking severe asthmatics (SAn; n=117), current or ex-smoking severe asthmatics (SAc/ex; n=51), mild-to-moderate asthmatics (MMA; n=43), or healthy controls (HC; n=41) (Table 1).

Topological data analysis of sputum lipid phenotypes

A total of 291 lipid molecular ions were quantified in the sputum samples (for methodology, QC and selection procedures see this article's Online Repository at www.jacionline.org). Of these, 92 lipid species were identified using LC-MS/MS of pooled QC samples. The remaining ions were classified as 'unknown lipids', but together these comprised only 5% of the total lipid signal.

Initial TDA of all samples produced a network comprising a tight, highly interconnected "core" group (~60% of participants), connected to a more diffuse "flare" (~40% of participants) via a small number of edges (Fig. 1A). This indicates that sputum lipid profiles were similar amongst members of the core group, but were markedly different in the flare group. To gain deeper insight, the "core" (C) and "flare" (F) sets were split and re-analysed in individual TDAs. This analysis yielded a ring-like network for the core set, consisting of four connected groups, which were labelled C1 to C4 (Fig. 1B). The flare set comprised a V-shaped string of five distinct groups, which were labelled F1 to F5 (Fig. 1B). Edges connecting the flare set to the core set in the original TDA network were restricted to C3, C4, F1 and F2 in the central part of the structure. The nine identified groups constitute a continuous spectrum of partly overlapping sputum lipid phenotypes, starting with the 'basal' phenotype of group C1, via 'intermediate' groups C2-C4, then F1-F2, to the 'terminal' groups F3-F5 (Fig. 1B).

To examine which components of the sputum lipidome were driving this structure, a trend analysis for ranked data was performed (JT-test). C1 and F5 were selected as the first and last group of the series, respectively, with the remaining intermediate groups ranked according to their distance from either end of the TDA structure (where groups were equidistant the order was assigned arbitrarily, e.g., C2 and C3). The results showed highly significant trends from C1 to F5 for 85% of the measured lipids (Fig. 2 and Table E1 in this article's Online Repository at www.jacionline.org). Relative amounts of DPPC and other palmitic acid-containing PC species progressively decreased from C1 onwards, being lowest in group F5. There was a reciprocal increase in the relative quantities of other lipids, including long-chain polyunsaturated fatty acid-containing PCs, mixed alkyl-acyl PCs, other glycerophospholipids such as PE and PS, sphingolipids, sterols (Chol and its esters), and triacylglycerols (TG).

Importantly, trends in relative abundance were not always matched by the actual lipid concentration data. For example, actual concentrations of the surfactant-specific lipid DPPC were comparable across all TDA groups (JT-test, p-value: 0.823, z-score: 0.223). Moreover, a number of differences between individual phenotypes did not conform to the general trend described. For example, relative to the other core groups, C2 was enriched in PC[16:0/18:0] and PC[16:0/18:1], whereas F3 was highly enriched in cholesterol and CE species, but not as enriched in PS[36:1] as the other flare groups (Table E1 in this article's Online Repository at www.jacionline.org).

Trends in matched clinical and pathobiological data

Trend tests were also performed for a variety of metadata available from the U-BIOPRED tranSMART repository, including demographic and clinical measurements, blood proteins and sputum cell pellet gene expression. There was a highly significant trend in asthma severity (JT-test, p-value <0.001, z-score: 6.336), as judged by the proportion of participants from each of the four clinically characterised U-BIOPRED categories (Table 1). The proportion of healthy controls was highest in basal group C1 (42%) and decreased progressively through the intermediate groups to 12% in C4, being only 6-7% in F1, F2 and F4. The two terminal flare groups (F3 and F5) contained only mild-to-moderate and severe asthmatics, but no healthy participants. Among the clinical and pathobiological variables, highly significant negative trends were observed for lung function measurements (spirometry and reversibility), whereas subject age, ACQ scores, serum IgE levels, blood inflammatory cells and blood platelets all increased significantly from C1 to F5 (Table E2 in this article's Online Repository at www.jacionline.org). The sputum differential cell counts showed reciprocal increases in eosinophils and neutrophils and significant decreases in macrophages and lymphocytes. Although median sputum eosinophil levels reached 3% of total

inflammatory cell counts in four of the TDA groups, a threshold viewed as clinically relevant,[33] they were significantly higher in group F3 (26.8%) than in any other group. Similarly, neutrophil levels gradually increased from 40% in C1 to 60% in groups F1-F3, and peaked in groups F4 and F5 (medians 83% and 91%, respectively).

Of the 32 blood protein biomarkers of inflammation and tissue function available for this analysis, more than half increased significantly from C1 to F4/F5 (Table E3 in this article's Online Repository at www.jacionline.org). Some of the between-group differences appeared to be independent of the overall trend. For example, high levels of serum Eotaxin-3 and IL-13 were associated with the high-eosinophil group F3, whereas levels of CCL17 and Galectin-3 in that group did not differ from the basal group C1. Finally, a number of upstream transcriptional regulators of inflammation, predicted by pathway analysis of the sputum cell pellet transcriptome, also showed consistent, significant trends across the TDA structure (Table E4 in this article's Online Repository at www.jacionline.org). These included increasing expressions (from C1 to F4/F5) of *RAB1B*, *PLA2R1*, *SYVN1*, *CD24*, *HSP90B1* and *DNMT3B*, and decreasing expressions of *miR-10*, *miR-122*, *WT1*, *SMARCA4*, *NANOG*, *KDM5B*, *ETS1* and *SMAD3*. However, for most of the regulators any differences in expression relative to C1 appeared to be specific to discrete, smaller parts of the structure.

Discussion

In this comprehensive cross-sectional assessment of sputum lipid biomarkers, we show the existence of a continuous spectrum of molecular phenotypes from health to severe neutrophilic and eosinophilic asthma (Figs. 1 and 3). TDA of the sputum lipidome showed a progressive reduction in relative quantities of DPPC and other di-

saturated PC species from the basal group C1 to the flare groups F1-F5 (Fig. 2). This trend was matched by a progressive increase in both absolute and relative abundances of alkyl-acyl PCs, longer-chain/polyunsaturated fatty acid-containing PCs, various PE and PS species, sphingolipids, and neutral lipids (Chol, CE, DAG and TG species), in particular in the eosinophilic and neutrophilic severely asthmatic flare groups. In contrast, absolute concentrations of surface-active di-saturated PC species, the main lipid component of pulmonary surfactant,[15,34,35] did not vary significantly between TDA groups. Lipid metabolism is highly dynamic, responding to developmental, nutritional and environmental challenges by up- or down-regulating lipid synthetic and catabolic pathways that maintain homeostasis. Due to its unique role in reducing surface tension at the air–liquid interface in the alveoli, and thereby preventing collapse of these structures at end expiration, the lipid composition of pulmonary surfactant is tightly regulated by the ATII cells.[35] The constant levels of surface-active di-saturated PC species observed in this study strongly suggest that surfactant production and secretion are not significantly altered in asthma. Rather, the progressive increase in lipids that are not secreted by ATII cells as part of the pulmonary surfactant points to the presence of another source for this material, particularly in the TDA flare groups. Given the nature of this sample type, the number of potential sources is limited. Saliva was ruled out as a major source since the numbers of squamous epithelial cells derived from the upper airways, an indicator of salivary contamination,[36,37] were on average only 10% and did not vary significantly between TDA groups. Moreover, concentrations of neutral lipids, which predominate in saliva,[38] were low. Plasma infiltration in the upper airways has been shown to disrupt the respiratory lipidome, in particular during asthma attacks and allergen challenges, leading to significantly increased levels in the ELF of typical plasma lipids, such as

linoleic acid-containing PCs.[39,40] This pattern did not match any of the phenotypes described here, and lipid species such as PC[34:1] and PC[34:2] either followed similar trends to DPPC, or did not show strong trends across the TDA network at all. This indicates that plasma infiltration is not a major factor in driving the sputum lipid phenotypes observed in this study.

In contrast, highly significant associations were seen between the lipidomic trends and inflammatory cell numbers. The flare groups all contained high numbers of granulocytic inflammatory cells, predominately eosinophils in F3, neutrophils in F4 and F5, and a combination of both in F1 and F2 (Fig. 3). We speculate, therefore, that the ELF lipidome in these phenotypes is enriched by material derived from airway granulocytes. The sampling protocol required removal (by centrifugation) of whole cells, and samples rich in dead or damaged cells were excluded from analysis as part of the QC. Thus, we postulate that the lipid material could be derived either from small membrane fragments or secreted extracellular vesicles (EVs). The latter were recently identified in both bronchoalveolar lavage fluid and induced sputum samples from mild allergic asthma patients.[41-43]

Knowledge of the lipid pathobiology of neutrophils and eosinophils is unfortunately limited. Neutrophils are rich in PC, PE, PS, PI, SM and cholesterol, with high levels of mixed alkyl-acyl.[44-47] In eosinophils, research has mainly focused on activation-induced formation of intracellular lipid droplets and their metabolism of arachidonic acid as the precursor for pro-inflammatory lipid mediators,[48-50] but there has been no systematic profiling of the eosinophil lipidome. Both neutrophils and eosinophils release EVs in response to inflammatory stimuli, either as endosome-derived exosomes or outer membrane-derived micro-vesicles.[51,52] Such EVs were shown to be rich in sphingolipids, PS and neutral lipids, and may also reflect the lipid

compositions of their progenitor cells.[53-55] The TDA flare groups were all significantly enriched in lipids that fit this inflammatory cell profile. For example, levels of arachidonic acid-containing lipids such as PC[16:0/20:4] and PC[18:0/20:4], as well as cholesterol and its esters, were highest in these groups (especially in the eosinophil-rich group F3), as were levels of SM species and other sphingolipids (especially in F4 and F5). Levels of mixed alkyl-acyl species such as PC[O-16:0/18:1], common in neutrophils, were only enriched in the two neutrophil-rich groups.

Eotaxin production is elevated in the airways of asthmatic patients, and this chemokine can both recruit eosinophils[56] and stimulate the formation of lipid droplets in these cells, which are enriched in arachidonic acid and can act as sites of eicosanoid formation.[57,58] Such lipid droplets may selectively contribute lipid material for EVs, in particular exosomes, the formation and secretion of which is also induced by eotaxin and other inflammatory stimuli.[51] A similar mechanism was recently proposed for neutrophil-derived exosomes, which were shown to contribute to airway smooth muscle remodelling.[52] In the current study, circulating levels of a range of chemokine and cytokine markers of inflammation and tissue function were significantly higher in the serum of participants in the TDA flare groups. In addition, the upstream transcriptional regulator with the strongest positive trend across the TDA structure was *RAB1B*. Members of the Rab GTPases protein family are key regulators of intracellular membrane trafficking and are present on the membranes of lipid droplets.[59,60] Combined with the sputum lipidomics results, this suggests that pro-inflammatory mediators have additional, and potentially damaging, biological effects beyond cell recruitment. We also note a significant upregulation of the protein-coding gene *PLA2R1* in the granulocytic flare groups (F3-F5). This receptor is produced to counteract the biological effects of secreted phospholipase A2 enzymes,[61] and

PLA2R1 was previously shown to be overexpressed in the bronchial epithelium of children with atopic asthma.[62] We did not observe a direct effect of potentially increased phospholipase A2 activity on the sputum lipidome of groups F3-F5 (e.g., an increase in lysophospholipid levels[63,64]). Nevertheless, phospholipase activity is likely an important driver in the pathobiology of eosinophilic and neutrophilic asthma, and additional insight is required into their localisation, substrate specificity and kinetics.

The lipid phenotypic differences observed within the TDA core group were more subtle than in the flare groups. Group C1 contained a mixed population of healthy and mildly asthmatic participants with a sputum lipid profile that matches that of healthy adults.[12-14] The remaining core groups (C2-C4) contained smaller numbers of healthy participants, and mostly comprised a mixture of mild-to-moderate and non-inflammatory severe asthmatics. Phenotype C4 was paucigranulocytic, with mean eosinophil and neutrophil counts of 2% and 43% respectively. Its lipid composition was intermediate between the healthy and severe granulocytic flare phenotypes, containing relatively less DPPC, elevated levels of the other lipid classes mentioned above, and a specific enrichment in PI species. The lipid phenotype of group C3 was intermediate between C1 and C4, with notably low levels of PIs. The main pathobiological features distinguishing these two groups of asthmatics were their atopy status and serum IgE levels, both of which were high in C4 and low in C3, suggesting the presence of a distinct sputum lipid phenotype for 'non-atopic' asthmatics.[65] Finally, group C2 had a lipid phenotype similar to that of the 'healthy' group, but with relatively less DPPC and concomitantly increased levels of C18 fatty acid-containing PC species. This group was characterised by a higher body mass index and waist circumference, suggesting that the differences were related to body weight status, rather than a particular type of asthma. Several studies have

458 demonstrated dysregulation of lung lipid metabolism in obese animal models,[66-68] and
459 we have previously reported on a distinct ELF lipid phenotype in overweight, but
460 otherwise healthy, human adults.[14]

461 Finally, we examined associations between the lipidome and upstream regulators of
462 inflammation identified by IPA of sputum cell pellets. In addition to upregulated *RAB1B*
463 and *PLA2R1* in the TDA flare groups, significant associations were observed between
464 asthma severity and expression of *SYVN1, CD24, HSP90B1* and *DNMT3B*
465 (upregulation), and *miR-10*, *miR-122*, *WT1*, *SMARCA4*, *NANOG*, *KDM5B*, *ETS1* and
466 *SMAD3* (downregulation). Many of these have previously been indicated in asthma
467 and other inflammatory diseases. The upregulation of *CD24*, known to be expressed
468 on granulocytes and lymphocytes, matched the high-neutrophil TDA groups F2, F4
469 and F5.[69] In contrast, *ETS1* is a negative regulator of Th17 cells, proposed to drive
470 specific phenotypes of asthma.[70-73] Overexpression of *DNMT3B* promotes
471 macrophage polarization into a 'classically-activated' M1 phenotype and enhances
472 macrophage inflammation,[74] and this gene was upregulated in all of the asthmatic
473 groups apart from C2 and C4. Finally, Wilms Tumour 1 (*WT1*) is known to regulate the
474 expression of Matrix metalloproteinase-9 (MMP-9), an enzyme responsible for
475 extracellular matrix degradation and airway remodelling in asthma.[75] Our results
476 suggest that this mechanism may be activated in the most severe neutrophilic
477 asthmatics (F4, F5), with *WT1* downregulation leading to more MMP-9 mediated
478 tissue.[76]

479 In summary, we have shown that lipidomic profiling of induced sputum stratifies
480 asthma into a spectrum of distinct molecular phenotypes, and that the abundance and
481 proportion of lipids that are non-endogenous to the pulmonary surfactant increases
482 significantly with asthma severity. Based on matching trends in the clinical,

immunoassay and transcriptomic data, we propose a hypothesis for a novel mechanism of surfactant dysregulation in severe asthma, wherein granulocytes recruited into the airways are activated to produce both intracellular lipid droplets and EVs (exosomes and/or microvesicles) (Fig. 4). Upon release, these lipid and protein-rich EVs could disrupt the tightly regulated structure of the pulmonary surfactant component of the ELF,[34] in a similar way as has been proposed for granulocyte-derived proteins in asthma.[77] In turn, this would reduce the surfactant's ability to lower surface tension in the small airways, leading to collapsibility, and potentially also compromise its immunological function.[34,42]

We wish to highlight that our findings require external validation, and that the variable efficacy of widely used asthma medication such as inhaled corticosteroids may play an important role in refining the respiratory lipid phenotype model presented here. Further *in vivo* or *in vitro* studies are needed to verify the proposed mechanism, including detailed analyses of the lipidomes of granulocytes and their EVs. Nonetheless, immunomodulation of EV secretion by granulocytes in the lungs could provide a new and attractive therapeutic target for severe asthma. If the proposed mechanism is substantiated, then efforts should be directed at identifying drugs that could modulate the observed sputum lipid phenotypes, thereby further exploring the relevance of respiratory lipid metabolic changes in asthma and the potential for additional therapeutics.

Acknowledgements

505 The authors thank John Langley, Julie Herniman and Jon Paul Townsend for their
506 analytical support, and Dominic Burg, Ben Nicholas, Kamran Tariq and Jeanne-Marie
507 Perotin-Collard for additional discussion.

References

1. Yancey SW, Keene ON, Albers FC, Ortega H, Bates S, Bleecker ER, et al. Biomarkers for severe eosinophilic asthma. J Allergy Clin Immunol 2017; 140: 1509-1518. PMID: 29221581
2. Anderson GP. Endotyping asthma: new insights into key pathogenic mechanisms in a complex, heterogeneous disease. Lancet 2008; 372: 1107-1119. PMID: 18805339
3. Lötvall J, Akdis CA, Bacharier LB, Bjermer L, Casale TB, Custovic A, et al. Asthma endotypes: a new approach to classification of disease entities within the asthma syndrome. J Allergy Clin Immunol 2011; 127: 355-360. PMID: 21281866
4. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention. 2017: www.ginasthma.org
5. Shaw DE, Sousa AR, Fowler SJ, Fleming LJ, Roberts G, Corfield J, et al. Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. Eur Respir J 2015; 46: 1308-1321. PMID: 26357963
6. Auffray C, Adcock IA, Chung KF, Djukanović R, Pison C, Sterk PJ. An integrative systems biology approach to understanding pulmonary diseases. Chest 2010; 137: 1410-1416. PMID: 20525651
7. Bel EH, Sousa A, Fleming L, Bush A, Fan Chung K, Versnel J, et al. Diagnosis and definition of severe refractory asthma: an international consensus statement from the Innovative Medicine Initiative (IMI). Thorax 2011; 66: 910-917. PMID: 21106547
8. Wheelock CE, Goss VM, Balgoma D, Nicholas B, Brandsma J, Skipp PJ, et al. Application of 'omics technologies to biomarker discovery in inflammatory lung diseases. Eur Respir J 2013; 42: 802-825. PMID: 23397306

9. Bigler J, Boedigheimer M, Schofield JPR, Skipp PJ, Corfield J, Rowe A, et al. A severe asthma disease signature from gene expression profiling of peripheral blood from UBIOPRED cohorts. Am J Respir Crit Care Med 2017; 195: 1311-1320. PMID: 27925796
10. Schofield JPR, Burg D, Nicholas B, Strazzeri F, Brandsma J, Staykova DK, et al. Stratification of asthma phenotypes by airway proteomic signatures. J Allergy Clin Immunol 2019. PMID: 30928653
11. Reinke SN, Naz S, Chaleckis R, Gallart-Ayala H, Kolmert J, Kermani NZ, et al. Urinary metabotype of severe asthma evidences decreased carnitine metabolism independent of oral corticosteroid treatment in the U-BIOPRED study. Eur Respir J 2022; 59: 2101733. PMID: 34824054
12. Dushianthan A, Cusack R, Goss VM, Cusack R, Grocott MPW, Postle AD. Phospholipid composition and kinetics in different endobronchial fractions from healthy volunteers. BMC Pulm Med 2014; 14: 10. PMID: 24484629
13. t'Kindt R, Telenga ED, Jorge L, Van Oosterhout AJ, Sandra P, Ten Hacken NH, et al. Profiling over 1500 lipids in induced lung sputum and the implications in studying lung diseases. Anal Chem. 2015; 87(9): 4957-4964. PMID: 25884268
14. Brandsma J, Goss VM, Yang X, Bakke PS, Caruso M, Chanez P, et al. Lipid phenotyping of lung epithelial lining fluid in healthy human volunteers. Metabolomics 2018; 14: 123. PMID: 30830396
15. Goss VM, Hunt AN, Postle AD. Regulation of lung surfactant phospholipid synthesis and metabolism. Biochim Biophys Acta 2012; 1831: 448-458. PMID: 23200861

16. Bernhard W. Lung surfactant: Function and composition in the context of development and respiratory physiology. Ann Anat 2016; 208: 146-150. PMID: 27693601
17. Fessler MB, Summer RS. Surfactant lipids at the host-environment interface: Metabolic sensors, suppressors, and effectors of inflammatory lung disease. Am J Respir Cell Mol Biol 2016; 54: 624-635. PMID: 26859434
18. Orešič M, Vidal-Puig A, Hänninen V. Metabolomic approaches to phenotype characterization and applications to complex diseases. Expert Rev Mol Diagn 2006; 6: 575-585. PMID: 16824031
19. Shevchenko A, Simmons K. Lipidomics: coming to grips with lipid diversity. Nat Rev Mol Cell Biol 2010; 11: 593-598. PMID: 20606693
20. Sahu S, Lynn WS. Lipid composition of sputum from patients with asthma and patients with cystic fibrosis. Inflammation 1978; 3: 27-36. PMID: 581080
21. Telenga ED, Hoffmann RF, t'Kindt R, Hoonhorst SJ, Willemse BW, van Oosterhout AJ, et al. Untargeted lipidomic analysis in chronic obstructive pulmonary disease: uncovering sphingolipids. Am J Respir Crit Care Med 2014; 190: 155-164. PMID: 24871890
22. Quinn RA, Phelan VV, Whiteson KL, Garg N, Bailey BA, Lim YW, et al. Microbial, host and xenobiotic diversity in the cystic fibrosis sputum metabolome. ISME J 2016; 10: 1483-1498. PMID: 26623545
23. Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. Can J Biochem Physiol 1959; 37: 911-917. PMID: 13671378
24. Hsu FF, Turk J. Electrospray ionization/tandem quadrupole mass spectrometric studies on phosphatidylcholines: the fragmentation processes. J Am Soc Mass Spectrom 2003; 14: 352-363. PMID: 12686482

25. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007; 8: 118-127. PMID: 16632515
26. Simpson JL, Timmins NL, Fakes K, Talbot P, Gibson PG. Effect of saliva contamination on induced sputum cell counts, IL-8 and eosinophil cationic protein levels. Eur Respir J 2004; 23: 759-762. PMID: 15176693
27. Kirwan JA, Weber RJM, Broadhurst DI, Viant MR. Direct infusion mass spectrometry metabolomics dataset: a benchmark for data processing and quality control. Sci Data 2014; 1: 14002. PMID: 25977770
28. Hinks TS, Brown T, Lau LC, Rupani H, Barber C, Elliott S, et al. Multidimensional endotyping in patients with severe asthma reveals inflammatory heterogeneity in matrix metalloproteinases and chitinase 3–like protein 1. J Allergy Clin Immunol 2016; 138: 61-75. PMID: 26851968
29. Siddiqui S, Shikotra A, Richardson M, Doran E, Choy D, Bell A, et al. Airway pathological heterogeneity in asthma: visualization of disease microclusters using topological data analysis. J Allergy Clin Immunol 2018; 142: 1457-1468. PMID: 29550052
30. Wasserman L. Topological Data Analysis. Annu Rev Stat Appl 2018; 5: 501-523.
31. Schofield JPR, Strazzeri F, Bigler J, Boedigheimer M, Adcock IA, Fan Chung K, et al. Morse-clustering of a topological data analysis network identifies phenotypes of asthma based on blood gene expression profiles. bioRxiv 2020. DOI: 10.1101/516328
32. Kuo CS, Pavlidis S, Loza M, Baribaud F, Rowe A, Pandis I, et al. T-helper cell type 2 (Th2) and non-Th2 molecular phenotypes of asthma using sputum

transcriptomics in U-BIOPRED. Eur Respir J 2017; 49: 1602135. PMID: 28179442

33. Green RH, Brightling CE, McKenna S, Hargadon B, Parker D, Bradding P, et al. Asthma exacerbations and sputum eosinophil counts: a randomised controlled trial. Lancet 2002; 360: 1715-1721. PMID: 12480423

34. Lopez-Rodriguez E, Pérez-Gil J. Structure-function relationships in pulmonary surfactant membranes: from biophysics to therapy. Biochim Biophys Acta 2014; 1838: 1568-1585. PMID: 24525076

35. Brandsma J, Postle AD. Analysis of the regulation of surfactant phosphatidylcholine metabolism using stable isotopes. Ann Anat 2017; 211: 176-183. PMID: 28351529

36. Belda J, Leigh R, Parameswaran K, O'Byrne PM, Sears MR, Hargreave FR. Induced sputum cell counts in healthy adults. Am J Respir Crit Care Med 2000; 161: 475-478. PMID: 10673188

37. Spanevello A, Confalonieri M, Sulotto F, Romano F, Balzano G, Migliori GB, et al. Induced sputum cellularity: reference values and distribution in normal volunteers. Am J Respir Crit Care Med 2000; 162: 1172-1174. PMID: 10988149

38. Larsson B, Olivecrona B, Ericson T. Lipids in human saliva. Arch Oral Biol 1996; 41: 105-110. PMID: 8833598

39. Heeley EL, Hohlfeld J, Krug N, Postle AD. Phospholipid molecular species of bronchoalveolar lavage fluid after local allergen challenge in asthma. Am J Physiol Lung Cell Mol Physiol 2000; 278: L305-L311. PMID 10666114

40. Wright SM, Hockey PM, Enhorning G, Strong P, Reid KB, Holgate ST, et al. Altered airway surfactant phospholipid composition and reduced lung function in asthma. J Appl Physiol 2000; 89: 1283-1292. PMID: 11007560

41. Torregrosa Paredes P, Esser J, Admyre C, Nord M, Rahman QK , Lukic A, et al. Bronchoalveolar lavage fluid exosomes contribute to cytokine and leukotriene production in allergic asthma. Allergy 2012; 67: 911-919. PMID: 22620679
42. Sánchez-Vidaurre S, Eldh M, Larssen P, Daham K, Martinez-Bravo M-J, Dahlén S-E, et al. RNA-containing exosomes in induced sputum of asthmatic patients. J Allergy Clin Immunol 2017; 140: 1459-1461.e2. PMID: 28629752
43. Pastor L, Vera E, Marin JM, Sanz-Rubio D. Extracellular vesicles from airway secretions: New insights in lung diseases. Int J Mol Sci 2021; 22: 583. PMID: 33430153
44. Klock JC, Pieprzyk JK. Cholesterol, phospholipids, and fatty acids of normal immature neutrophils: comparison with acute myeloblastic leukemia cells and normal neutrophiIs. J Lipid Res 1979; 20: 908-911. PMID: 290722
45. Postle AD, Madden G, Clark GT, Wright SM. Electrospray ionisation mass spectrometry analysis of differential turnover of phosphatidylcholine by human blood leukocytes. Phys Chem Chem Phys 2004; 6: 1018-1021.
46. Leidl K, Liebisch G, Richter D, Schmitz G. Mass spectrometric analysis of lipid species of human circulating blood cells. Biochim Biophys Acta 2008; 1781: 655-664. PMID: 18723117
47. Alarcon-Barrera JC, Von Hegedus JH, Brouwers H, Steenvoorden E, Ioan-Facsinay A, Mayboroda OA, et al. Lipid metabolism of leukocytes in the unstimulated and activated states. Anal Bioanal Chem 2020; 412; 2353-2363. PMID: 32055910
48. Weller PF, Monahan-Earley RA, Dvorak HF, Dvorak AM. Cytoplasmic lipid bodies of human eosinophils. Am J Pathol 1991; 138: 141-148. PMID: 1846262

654 49. Isobe Y, Kato T, Arita M. Emerging roles of eosinophils and eosinophil-derived
655 lipid mediators in the resolution of inflammation. Front Immunol 2012; 3: 270.
656 PMID: 22973272
657 50. Melo RCN, Weller PF. Unravelling the complexity of lipid body organelles in
658 human eosinophils. J Leukoc Biol 2014; 96: 703-712. PMID: 25210147
659 51. Akuthota P, Carmo LAS, Bonjour K, Murphy RO, Silva TP, Gamalier JP, et al.
660 Extracellular microvesicle production by human eosinophils activated by
661 "inflammatory" stimuli. Front Cell Dev Biol 2016; 4: 117. PMID: 27833910
662 52. Vargas A, Roux-Dalvai F, Droit A, Lavoie J-P. Neutrophil-derived exosomes: a
663 new mechanism contributing to airway smooth muscle remodelling. Am J Respir
664 Cell Mol Biol 2016; 55: 450-461. PMID: 27105177
665 53. Skotland T, Sandvig K, Llorente A. Lipids in exosomes: current knowledge and
666 the way forward. Prog Lipid Res 2017; 66: 30-41. PMID: 28342835
667 54. Hough KP, Wilson LS, Trevor JL, Strenkowski JG, Maina N, Kim YI, et al. Unique
668 lipid signatures of extracellular vesicles from the airways of asthmatics. Sci Rep
669 2018; 8: 10340. PMID: 29985427
670 55. Van Niel G, D'Angelo G, Raposo G. Shedding light on the cell biology of
671 extracellular vesicles. Nat Rev Mol Cell Biol 2018; 19: 213-228. PMID: 29339798
672 56. Pease JE, Williams TJ. Eotaxin and asthma. Curr Opin Pharmacol 2001; 1: 248-
673 253. PMID: 11712747
674 57. Bandeira-Melo C, Phoofolo M, Weller PF. Extranuclear lipid bodies, elicited by
675 CCR3-mediated signaling pathways, are the sites of chemokine-enhanced
676 leukotriene C4 production in eosinophils and basophils. J Biol Chem 2001; 276:
677 22779-22787. PMID: 11274187

678 58. Dichlberger A, Kovanen PT, Schneider WJ. Mast cells: from lipid droplets to lipid
679 mediators. Clin Sci 2013; 125: 121-130. PMID: 23577635
680 59. Hutagalung AH, Novick PJ. Role of Rab GTPases in membrane traffic and cell
681 physiology. Physiol Rev 2011; 91: 119-149. PMID: 21248164
682 60. Yang L, Ding Y, Chen Y, Zhang S, Huo C, Wang Y, et al. The proteomics of lipid
683 droplets: structure, dynamics, and functions of the organelle conserved from
684 bacteria to humans. J Lipid Res 2012; 53: 1245-1253. PMID: 22534641
685 61. Murakami M, Sato H, Miki Y, Yamamoto K, Taketomi Y. A new era of secreted
686 Phospholipase A2 (sPLA2). J Lipid Res 2015; 56: 1248-1261. PMID: 25805806
687 62. Nolin JD, Ogden HL, Lai Y, Altemeier WA, Frevert CW, Bollinger JG, et al.
688 Identification of epithelial Phospholipase A2 Receptor 1 as a potential target in
689 asthma. Am J Respir Cell Mol Biol 2016; 55: 825-836. PMID: 27448109
690 63. Hite RD, Seeds MC, Jacinto RB, Grier BL, Waite BM, Bass DA. Lysophospholipid
691 and fatty acid inhibition of pulmonary surfactant: non-enzymatic models of
692 phospholipase A2 surfactant hydrolysis. Biochim Biophys Acta 2005; 1720: 14-
693 21. PMID: 16376294
694 64. Ackerman SJ, Park GY, Christman JW, Nyenhuis S, Berdyshev E, Natarajan V.
695 Polyunsaturated lysophosphatidic acid as a potential asthma biomarker. Biomark
696 Med 2016; 10: 123-135. PMID: 26808693
697 65. Pillai P, Fang C, Chan YC, Shamji MH, Harper C, Wu SY, et al. Allergen-specific
698 IgE is not detectable in the bronchial mucosa of nonatopic asthmatic patients. J
699 Allergy Clin Immunol 2014; 133: 1770-1772.e11. PMID: 11712747
700 66. Inselman LS, Chander A, Spitzer AR. Diminished lung compliance and elevated
701 surfactant lipids and proteins in nutritionally obese young rats. Lung 2004; 182
702 101-117. PMID: 15136884

67. Foster DJ, Ravikumar P, Bellotto DJ, Unger RH, Hsia CCW. Fatty diabetic lung: altered alveolar structure and surfactant protein expression. Am J Physiol Lung Cell Mol Physiol 2010; 298: L392-L403. PMID: 20061442

68. Showalter MR, Nonnecke EB, Linderholm AL, Cajka T, Sa MR, Lönnerdal B, et al. Obesogenic diets alter metabolism in mice. PLoS ONE 2018; 13: e0190632. PMID: 29324762

69. Elghetany MT, Patel J. Assessment of CD24 expression on bone marrow neutrophilic granulocytes: CD24 is a marker for the myelocytic stage of development. Am J Hematol 2002; 71: 348-349. PMID: 12447971

70. Garrett-Sinha LA. Review of Ets1 structure, function, and roles in immunity. Cell Mol Life Sci 2013; 70: 3375-3390. PMID: 23288305

71. Choy DF, Hart KM, Borthwick LA, Shikotra A, Nagarkar DR, Siddiqui S, et al. TH2 and TH17 inflammatory pathways are reciprocally regulated in asthma. Sci Transl Med 2015; 7: 301ra129. PMID: 26290411

72. Liu W, Liu S, Verma M, Zafar I, Good JT, Rollins D, Groshong S, et al. Mechanism of T_H2/T_H17-predominant and neutrophilic T_H2/T_H17-low subtypes of asthma. J Allergy Clin Immunol 2017; 139: 1548-1558. PMID: 27702673

73. Östling J, van Geest M, Schofield JPR, Jevnikar Z, Wilson S, Ward J, et al. IL-17-high asthma with features of a psoriasis immunophenotype. J Allergy Clin Immunol 2019; 144: 1198-1213. PMID: 30998987

74. Yang X, Wang X, Liu D, Yu L, Xue B, Shi H. Epigenetic regulation of macrophage polarization by DNA methyltransferase 3b. Mol Endocrinol 2014; 28: 565-574. PMID: 24597547

75. Ohbayashi H, Shimokata K. Matrix metalloproteinase-9 and airway remodeling in asthma. Curr Drug Targets Inflamm Allergy 2005; 4: 177-181. PMID: 15853739

76. Marcet-Palacios M, Ulanova M, Duta F, Puttagunta L, Munoz S, Gibbings D, et al. The transcription factor Wilms tumor 1 regulates matrix metalloproteinase-9 through a nitric oxide-mediated pathway. J Immunol 2007; 179: 256-265. PMID: 17579045

77. Hohlfeld JM, Schmiedl A, Erpenbeck VJ, Venge P, Krug N. Eosinophil cationic protein alters surfactant structure and function in asthma. J Allergy Clin Immunol 2004; 113: 496-502. PMID: 15007353

736 **Table 1**

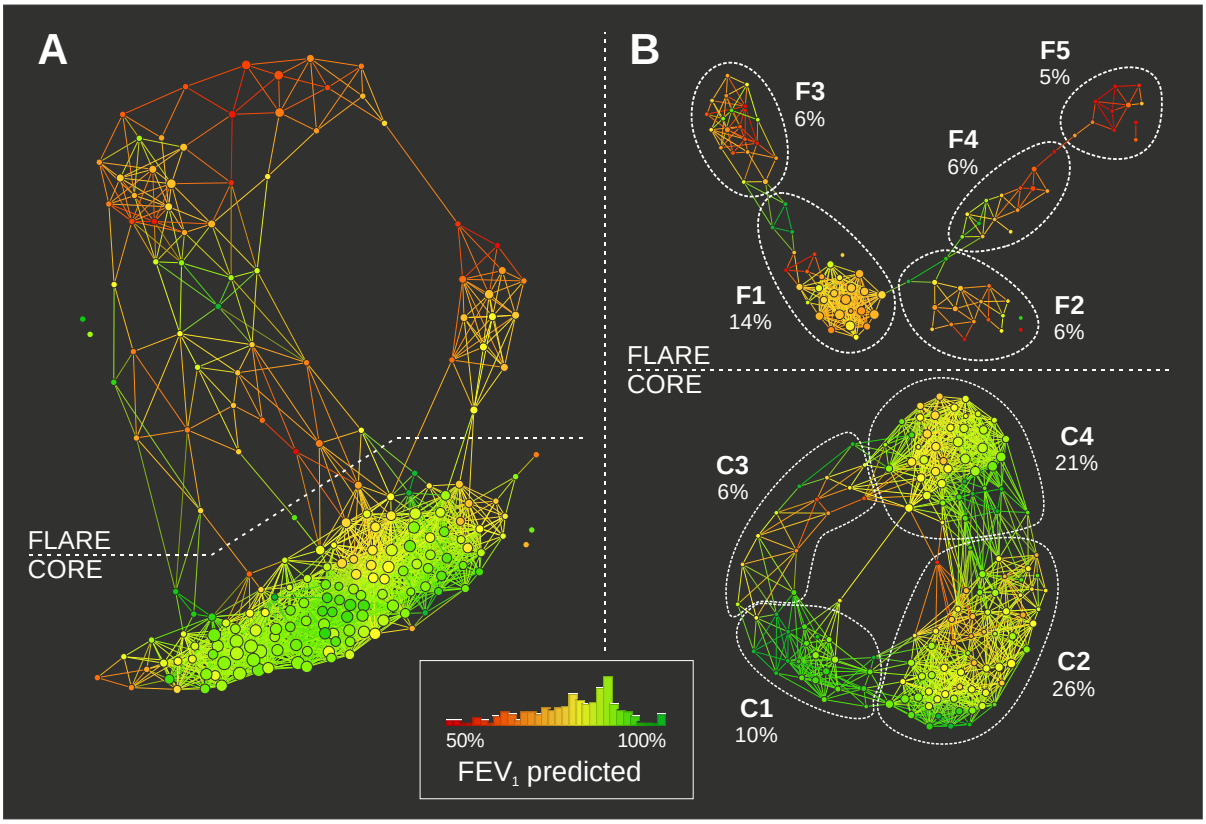
737

		Severe asthmatic Active or ex-smoker n=51	Severe ast Non-sm n=11
Age	mean [range]	55 [29-74]	53 [21-
Sex (m/f)	ratio	19/32	43/74
Race (Caucasian/non-Caucasian)	ratio	49/2	109/8
Age at first diagnosis	mean [range]	35 [1-67]	24 [0-6
BMI	mean [range]	30.1 [20.6-48.4]	29.1 [17.8
Serum IgE (mL-1)	mean [range]	274 [5-2690]	398 [0-68
FEV1 (% predicted)	mean [range]	66.1 [24.3-113.0]	65.4 [18.4-
FVC (% predicted)	mean [range]	90.6 [54.6-129.1]	87.8 [40.2-
FEV1/FVC ratio	mean [range]	62.6 [35.2-90.0]	62.1 [31.0
Exacerbations (past 12 months)	mean [range]	2.5 [0-10]	2.1 [0-
Smoking pack-years	mean [range]	24.2 [5-70]	0.4 [0-
Intubation (ever)	count [pct]	1 [2%]	13 [11
ICU admission (ever)	count [pct]	7 [14%]	28 [24
Positive atopy test	count [pct]	27 [53%]	73 [62
ACQ1-5 score	mean [range]	2.2 [0.2-4.4]	2.2 [0-5
ACQ7 score	mean [range]	4.2 [0-7.0]	3.9 [0-7
AQLQ score	mean [range]	4.4 [2.3-6.8]	4.6 [1.9-
Oral corticosteroid use (current)	count [pct]	25 [49%]	49 [42
Inhaled corticosteroid use (current)	count [pct]	50 [98%]	113 [97
Injectable corticosteroid use (current)	count [pct]	0 [0%]	8 [7%
Long-acting ß-agonist use (current)	count [pct]	48 [94%]	112 [96
Short-acting ß-agonist use (current)	count [pct]	37 [73%]	91 [78
Corticosteroid dose (mg day-1)	mean [range]	13.7 [2.5-40.0]	12.4 [5.0-

738 Demographics of study participants according to the U-BIOPRED cohorts (see main
739 text for definitions). Ex-smokers with a pack-year smoking history of ≤5 were
740 considered to have a 'negative' smoking status, whereas ex-smokers with pack-year
741 ≥5 were only included in the study if also diagnosed with severe asthma.
742 Abbreviations: BMI = body mass index; FEV_1 = forced expiratory volume in 1 second;

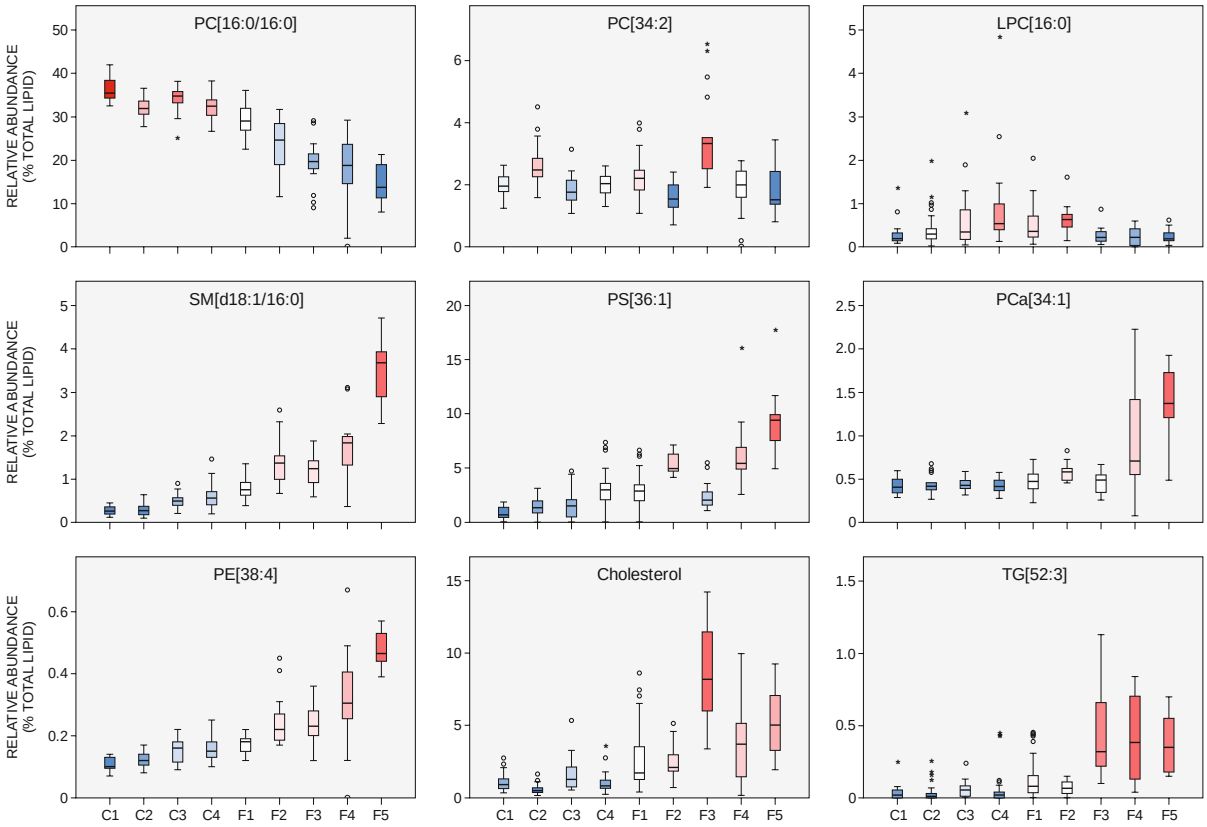
743 FVC = forced vital capacity; IgE = Immunoglobulin E; ICU = intensive care unit; ACQ
744 = Asthma Control Questionnaire; AQLQ = Asthma Quality of Life Questionnaire; NA =
745 not applicable/not assessed. Systemic dosage of corticosteroids for severe asthmatic
746 participants is expressed in prednisolone-equivalent doses.

Figure 1



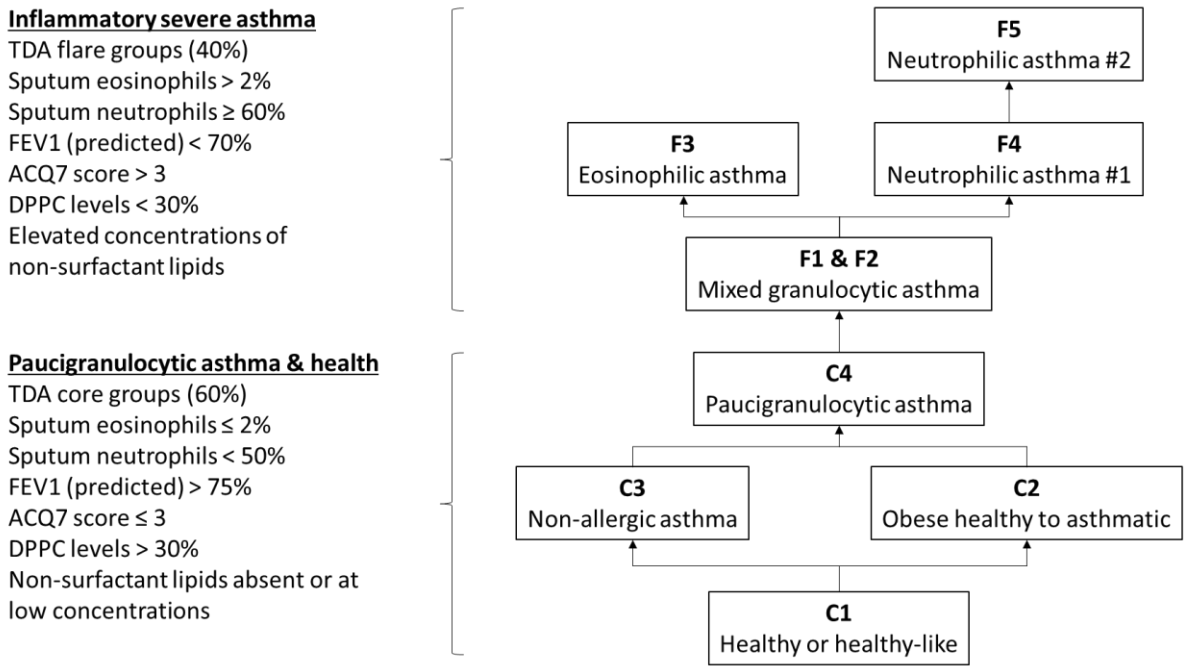
TDA structures of (A) the complete study cohort (n=252) and (B) the 'core' and 'flare' subgroups side by side (n=164 and n=107 respectively), coloured by FEV_1 (forced expiratory volume in 1 second) from 50% (red) to 100% (green; see histogram in inset). TDA was performed on 291 sputum lipid ions, using a normalised correlation metric and two MDS lenses. The TDA groups, as delineated by density mode clustering, along with the proportion of participants present in each group are shown in B. The original figures were obtained with the Symphony AyasdiAI machine intelligence platform (www.ayasdi.com).

Figure 2



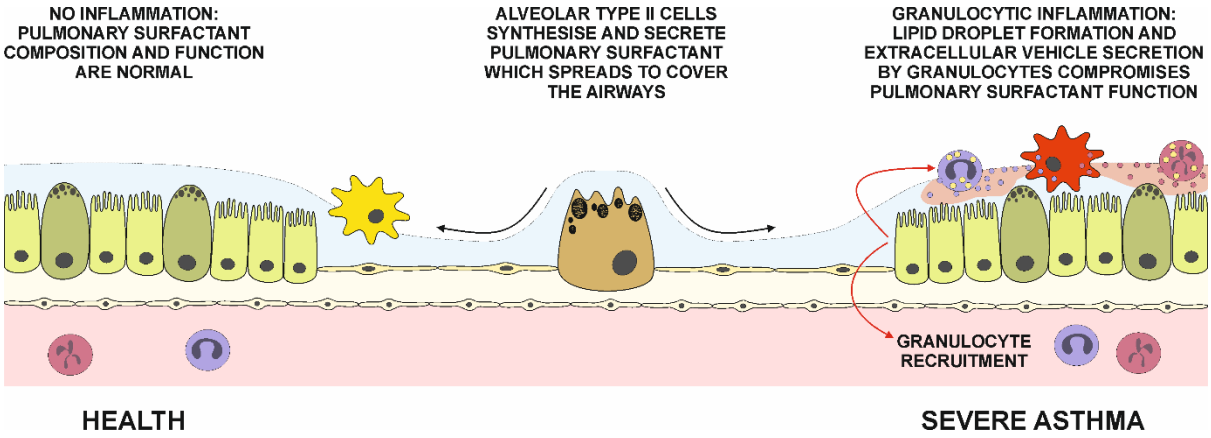
Box plots of representative lipid species demonstrating the trends across the TDA structure. Relative abundances are given as a percentage of the total lipid, and boxplots were coloured from blue (low) to red (high) to highlight trends. The original plots were created in SPSS Statistics 24 which defines outliers as 'near' (open circles: more than 1.5 times the interquartile range) and 'far' (stars: more than 3 times the interquartile range). The abbreviations of the lipid species, e.g. PC, are explained in the text.

Figure 3



Summary of the sputum lipid phenotypes found in this study and their assignments based on associations with the demographic, clinical and pathobiological data (predominately sputum differential cell counts). As shown in Fig. 1 and discussed in the main text, the nine phenotypes represent a spectrum of asthma severity from C1 (low) to F3 and F5 (high). Key characteristics of the main 'core' and 'flare' groups are listed on the left and can be found in Tables E1-E4 in this article's Online Repository (www.jacionline.org).

Figure 4



Conceptual representation of the potential role of granulocytic inflammation in producing EVs (exosomes and/or micro-vesicles) that may alter the lipid composition of the ELF (shown as light blue layer) in asthma. Pro-inflammatory chemokines and cytokines (red arrows) recruit eosinophils and neutrophils into the airways and stimulate intracellular lipid droplet formation and the secretion of EVs. The latter are rich in cellular lipids and proteins, which could impair the function of the pulmonary surfactant component of ELF, thereby reducing its ability to lower surface tension in the small airways and potentially compromising its role as an immunological barrier.

Stratification of asthma by lipidomic profiling of induced sputum supernatant

Online Repository

Methods

U-BIOPRED study design and tranSMART data repository

All samples used in this study were obtained from the U-BIOPRED cohort recruited in14 clinical centres across Europe (Shaw *et al*. 2015). The study protocols were approved by local Ethics Review Boards and participants gave their written informed consent for in-depth characterisation using U-BIOPRED standardised protocols for clinical assessment and biological sample collection, as well as molecular analysis by a variety of 'omics platforms. Processed biological samples from all clinical sites were blinded and stored in a central biobank (CIGMR Biobank, University of Manchester) and after completion of recruitment analysed in the Mass Spectrometry Unit of the NIHR Southampton Biomedical Research Centre. The study IDs and clinical metadata of the participants providing the samples were un-blinded only after completing all the analyses, data processing and quality control.

All clinical, laboratory and 'omics data collected as part of the U-BIOPRED study is hosted on the tranSMART knowledge management platform and available to study group members. To gain insight in the pathobiology underlying the sputum supernatant lipid phenotypes described in this work, we acquired from this database a variety of demographic, clinical and laboratory data, as well as blood protein and sputum cell pellet gene expression data (tranSMART query date: 28 September 2017).

Lipid analysis and data processing

A detailed description of the experimental procedures and data analysis methods can be found in Brandsma *et al*. (2018). Briefly, lipids were extracted from 100 µl of sputum using semi-automated Bligh-Dyer extraction protocol (Bligh and Dyer 1959) on a TECAN Freedom EVO100 robotic liquid handling platform (Tecan, Männedorf, Switzerland). Untargeted 'shotgun' mass spectra were acquired by flow injection analysis on a Dionex 3000 ultra-high performance liquid chromatography system (Thermo Scientific Dionex, Sunnyvale, CA, USA), coupled to a MaXis 3G quadrupole time-of flight mass spectrometer equipped with an electrospray ionisation source (Bruker Daltonics, Billerica, MA, USA). Measurements were done in full scan mode over an *m/z* range of 350-1200 with separate injections for positive and negative ionisation. Blank injections were performed after every four samples (no significant carry-over was detected) and a pooled QC sample was run after every four samples to check for changes in instrument performance. Fragmentation analysis for lipid identification was performed on the pooled QC using the same instrumental setup, but in LC-MS/MS mode using a Waters Acquity C8 column (1.7µm, 2.1mm x 100mm; Waters, Milford, MA, USA) and a 50 min gradient of methanol and water (both with 50 mM NH4HCO2 and 0.2% formic acid). Data-independent product ion scans were acquired over the entire gradient using broadband collision induced dissociation. Precursor and fragment ions were matched retrospectively by their LC retention time and using the well-established fragmentation rules for lipids (Hsu & Turk 2003) to provide confirmation of identities. Lipid nomenclature followed the framework set out by Liebisch *et al*. (2013) where sufficient structural information was available. The following abbreviations for lipid classes were used in this study: phosphatidylcholine (PC); phosphatidylglycerol (PG); phosphatidylserine (PS); phosphatidylinositol (PI); phosphatidylethanolamine (PE); lyso-phosphatidylcholine (LPC); ceramide (Cer); hexosyl-ceramide (HexCer); sphingomyelin (SM); cholesterol (Chol); cholesterol ester (CE); diglyceride (DG); and triglyceride (TG).

All raw screening mass spectra were smoothed, lock mass calibrated and aligned using a hierarchical clustering-based algorithm (adapted from Yang 2016). After background subtraction and removal of ions with <60% detection rate, the data were corrected for potential batch effects due to instrument performance or differences in sample work-up date using the R script “SVA ComBat” (Johnson *et al*. 2007). All ion counts were normalised using internal standards and the original sample volume to obtain semi-quantitative results. However, biofluids and particularly induced sputum are subject to variable dilution of analytes during sampling and subsequent workup (Simpson *et al*. 2004; Kirwan *et al*. 2014), hence the data were also normalised to the amount of total lipid.

Data analysis and statistics

Topological data analysis (TDA) was used to group participants with comparable sputum lipid profiles in an unbiased manner (Hinks *et al*. 2016; Bigler *et al*. 2017; Siddiqui *et al*. 2018), and identify trends and lipid phenotypes within the study cohort. TDA was performed using the AyasdiAI machine intelligence platform (Symphony AyasdiAI, Palo Alto, CA, USA) on the sputum lipid data set, employing a normalised correlation metric combined with two multidimensional scaling (MDS) lenses. Groups of participants with similar sputum lipid profiles were defined within the TDA structure using density mode clustering (Wasserman 2018, Schofield *et al*. 2020). By its very essence, TDA captures the continuous nature of data (Lum *et al*. 2013), and it therefore allows for a degree of overlap between connected groups of cases (*i.e.*, a study subject can be a member of more than one group at the same time). As this negatively affects the efficacy of statistical tests for between-group comparisons, the resolution settings of the TDA were adjusted to limit the degree of overlap between groups (less than 10% of participants were allowed to be shared), whilst at the same time maintaining the integrity of the TDA structure.

Note that the TDA group designations in this paper reflect the placement of each individual group within the semi-continuous TDA network. Groups C1 through to C4 constitute the ring-like structure of the “core group”, with C1 forming the outer edge of the overall TDA network

and C4 residing near its centre. F1 through to F5 constitute the V-shaped “flare”, a string of TDA groups connected to the “core” groups at F1 and F2, and with F3 and F4/F5 as its respective terminal branches. Thus, sputum lipid profiles are most different between patients in groups C1 and F3 or F4/F5, but each of the intermediate groups represents a step along the gradient between these extremes. Consequently, ranking the TDA groups along this gradient enables the use of ranks-based statistical tests, such as the Jonckheere-Terpstra test for ranked alternatives.

The statistical significance of trends across the TDA structure and differences between individual groups, were examined in SPSS Statistics 24 (IBM, Armonk, NY, USA). Analyses were performed for the sputum lipids, a variety of demographic, clinical and pathobiological measurements, as well as blood protein biomarker concentrations. Distribution analysis showed that, with few exceptions, the variables were not normally distributed within groups. Therefore, a non-parametric Jonckheere-Terpstra test for ordered alternatives was selected to identify trends in the ordinal and continuous variables, assuming an ascending hypothesis order. Trends were considered significant if $p<0.05$, and highly significant if $p<0.001$ (after Bonferroni correction).The significance of individual between-group differences was assessed pairwise by either Mann-Whitney U test or Pearson's Chi-squared test (with Bonferroni correction). However, the small group sizes likely led to this analysis being underpowered and prone to returning false positives, and consequently the results are not reported here.

Blood protein data

Protein expression data from peripheral blood were acquired as a routine analysis for all U-BIOPRED participants. Levels of 11 proteins were measured in plasma using a Mesoscale Discovery (MSD) electrochemiluminescence assay, whereas a further 21 proteins were measured in serum using either of the following immunoassay platforms: 16 by Luminex, 2 by Impact, 1 by Singulex, 1 by Elecsys, and 1 by Immulite. For this study, curated protein levels were downloaded directly from the U-BIOPRED data repository.

Pathway analysis of sputum cell pellet transcriptomics data

Transcriptomic data from RNA extracts of 97 matching sputum cell pellets (Kuo *et al*. 2017) were acquired from the U-BIOPRED data repository. These were subjected to Ingenuity Pathway Analysis (IPA; QIAGEN Bioinformatics, Redwood City, CA, USA) in order to identify potential upstream regulators of differential gene expression in each of the lipid phenotypes. IPA core analysis was performed on the top 4000 differentially expressed genes, using the ordered TDA groups as described above. The results were subjected to a comparison analysis to identify trends of IPA-predicted upstream regulator activation/inhibition across the sputum lipidomics TDA structure.

References

1. Bigler J, Boedigheimer M, Schofield JPR, Skipp PJ, Corfield J, Rowe A, *et al*. A severe asthma disease signature from gene expression profiling of peripheral blood from UBIOPRED cohorts. Am J Respir Crit Care Med. 2017; 195(10): 1311-1320. PMID: 27925796 847
2. Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. Can J Biochem Physiol. 1959; 37(8): 911-917. PMID: 13671378
3. Brandsma J, Goss VM, Yang X, Bakke PS, Caruso M, Chanez P, *et al*. Lipid phenotyping of lung epithelial lining fluid in healthy human volunteers. Metabolomics. 2018; 14(10): 123. PMID: 30830396
4. Hinks TS, Brown T, Lau LC, Rupani H, Barber C, Elliott S, *et al*. Multidimensional endotyping in patients with severe asthma reveals inflammatory heterogeneity in matrix metalloproteinases and chitinase 3–like protein 1. J Allergy Clin Immunol. 2016; 138(1): 61-75. PMID: 26851968

5. Hsu FF, Turk J. Electrospray ionization/tandem quadrupole mass spectrometric studies on phosphatidylcholines: the fragmentation processes. J Am Soc Mass Spectrom. 2003; 14(4): 352-363. PMID: 12686482
6. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007; 8(1): 118-127. PMID: 16632515
7. Kirwan JA, Weber RJM, Broadhurst DI, Viant MR. Direct infusion mass spectrometry metabolomics dataset: a benchmark for data processing and quality control. Sci Data. 2014; 1: 14002. PMID: 25977770
8. Kuo CS, Pavlidis S, Loza M, Baribaud F, Rowe A, Pandis I, *et al*. T-helper cell type 2 (Th2) and non-Th2 molecular phenotypes of asthma using sputum transcriptomics in U-BIOPRED. Eur Respir J. 2017; 49(2): 1602135. PMID: 28179442
9. Liebisch G, Vizcaíno JA, Köfeler H, Trötzmüller M, Griffiths WJ, Schmitz G, *et al*. Shorthand notation for lipid structures derived from mass spectrometry. J Lipid Res. 2013; 54(6): 1523-1530. PMID: 23549332
10. Schofield JPR, Strazzeri F, Bigler J, Boedigheimer M, Adcock IA, Fan Chung K, *et al*. Morse-clustering of a topological data analysis network identifies phenotypes of asthma based on blood gene expression profiles. bioRxiv. 2020. DOI: 10.1101/516328
11. Shaw DE, Sousa AR, Fowler SJ, Fleming LJ, Roberts G, Corfield J, *et al*. Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. Eur Respir J. 2015; 46(5): 1308-1321. PMID: 26357963
12. Siddiqui S, Shikotra A, Richardson M, Doran E, Choy D, Bell A, *et al*. Airway pathological heterogeneity in asthma: visualization of disease microclusters using topological data analysis. J Allergy Clin Immunol. 2018; 142(5): 1457-1468. PMID: 29550052
13. Simpson JL, Timmins NL, Fakes K, Talbot P, Gibson PG. Effect of saliva contamination on induced sputum cell counts, IL-8 and eosinophil cationic protein levels. Eur Respir J. 2004; 23(5): 759-762. PMID: 15176693
14. Wasserman L. Topological Data Analysis. Annu Rev Stat Appl. 2018; 5: 501-523.

15. Yang X. Analysing datafied life. 2016: PhD Thesis, Imperial College London, pp. 288.

Supplementary Table E1

Lipid	p-value	z-score	C1 10%	C2 26%	C3 6%	C4 21%	F1 14%	F2 6%	F3 6%	F4 6%	F5 5%
SM[d18:1/24:0]	<0.001	16.639	0.04	0.05	0.09	0.10	0.14	0.25	0.24	0.29	0.54
SM[d18:1/24:1]	<0.001	16.341	0.09	0.08	0.17	0.16	0.28	0.44	0.51	0.66	1.15
PC[30:1]	<0.001	16.303	0.13	0.14	0.19	0.23	0.31	0.54	0.50	0.72	1.44
SM[d18:1/16:0]	<0.001	16.248	0.25	0.26	0.49	0.56	0.74	1.36	1.23	1.83	3.67
PE[40:4]	<0.001	14.941	0.05	0.05	0.08	0.07	0.09	0.12	0.16	0.21	0.36
SM[d18:1/18:0]	<0.001	14.519	0.00	0.00	0.06	0.10	0.11	0.28	0.18	0.22	0.33
PE[38:4]	<0.001	13.239	0.10	0.12	0.16	0.15	0.18	0.22	0.23	0.31	0.47
PC[36:0]	<0.001	13.211	0.18	0.17	0.28	0.41	0.40	0.83	0.33	0.83	1.49
CE[18:2]	<0.001	12.377	0.06	0.04	0.16	0.07	0.30	0.28	2.02	0.85	0.66
PCa[40:2]	<0.001	12.269	0.02	0.02	0.05	0.04	0.06	0.14	0.07	0.12	0.24
CE[18:1]	<0.001	12.218	0.04	0.02	0.10	0.04	0.16	0.28	0.68	0.35	0.36
Cholesterol	<0.001	11.943	0.91	0.47	1.30	0.83	1.70	2.09	8.17	3.70	5.01
PC[38:4]	<0.001	11.501	0.13	0.15	0.15	0.16	0.19	0.22	0.51	0.36	0.39
PCa[36:0]	<0.001	11.281	0.09	0.09	0.11	0.13	0.14	0.26	0.13	0.20	0.27
PS[36:1]	<0.001	11.144	0.69	1.34	1.52	3.00	2.88	4.95	2.04	5.42	9.40
PC[38:0]	<0.001	10.951	0.01	0.02	0.02	0.03	0.03	0.06	0.05	0.07	0.09
PCa[40:1]	<0.001	10.838	0.00	0.01	0.00	0.02	0.02	0.07	0.03	0.07	0.15
PE[32:1]	<0.001	10.825	0.00	0.00	0.00	0.01	0.05	0.11	0.33	0.12	0.14
PE[40:6]	<0.001	10.346	0.10	0.13	0.12	0.16	0.16	0.25	0.15	0.24	0.33
TG[52:3]	<0.001	10.339	0.02	0.01	0.06	0.02	0.08	0.07	0.32	0.39	0.35
TG[52:2]	<0.001	10.041	0.03	0.02	0.08	0.03	0.10	0.09	0.40	0.66	0.44
Cer[d18:0/16:0]	<0.001	9.951	0.01	0.00	0.02	0.02	0.02	0.11	0.03	0.03	0.07
Cer[d18:1/16:0]	<0.001	9.949	0.03	0.01	0.05	0.04	0.05	0.17	0.06	0.08	0.15
PCa[32:2]	<0.001	9.891	0.04	0.03	0.04	0.04	0.05	0.09	0.05	0.08	0.10
PE[40:5]	<0.001	9.828	0.09	0.11	0.09	0.13	0.14	0.14	0.19	0.22	0.35
PCa[36:1]	<0.001	9.726	0.10	0.10	0.12	0.11	0.14	0.18	0.16	0.23	0.55
HexCer[d18:1/16:0]	<0.001	9.622	0.02	0.02	0.03	0.03	0.03	0.06	0.04	0.06	0.07
TG[50:0]	<0.001	9.488	0.00	0.00	0.02	0.01	0.03	0.03	0.04	0.08	0.07
PCa[42:2]	<0.001	9.380	0.00	0.00	0.00	0.01	0.03	0.04	0.03	0.11	0.18
PC[38:5]	<0.001	9.379	0.13	0.18	0.15	0.16	0.19	0.21	0.38	0.29	0.29
TG[54:3]	<0.001	9.260	0.00	0.01	0.05	0.02	0.05	0.06	0.15	0.21	0.22
TG[54:5]	<0.001	9.152	0.02	0.01	0.03	0.01	0.04	0.04	0.09	0.12	0.15
TG[54:4]	<0.001	9.101	0.02	0.01	0.04	0.01	0.04	0.05	0.14	0.16	0.20
TG[52:4]	<0.001	9.060	0.00	0.00	0.01	0.01	0.04	0.02	0.14	0.16	0.13
TG[52:1]	<0.001	9.015	0.01	0.01	0.03	0.01	0.04	0.04	0.10	0.23	0.15
CE[16:1]	<0.001	8.949	0.00	0.00	0.00	0.02	0.05	0.08	0.08	0.06	0.09
SM[d18:1/14:0]	<0.001	8.552	0.00	0.01	0.00	0.02	0.02	0.00	0.07	0.08	0.11
TG[54:2]	<0.001	8.441	0.02	0.02	0.03	0.03	0.03	0.05	0.04	0.07	0.13
TG[50:1]	<0.001	8.310	0.01	0.01	0.04	0.01	0.05	0.06	0.13	0.35	0.20
PC[38:1]	<0.001	8.196	0.03	0.02	0.05	0.06	0.05	0.11	0.05	0.07	0.13
TG[50:2]	<0.001	7.917	0.00	0.00	0.03	0.00	0.04	0.02	0.16	0.22	0.17
PCa[32:1]	<0.001	7.891	0.16	0.17	0.17	0.19	0.20	0.28	0.17	0.29	0.41
PCa[34:1]	<0.001	7.244	0.40	0.41	0.42	0.41	0.47	0.58	0.48	0.70	1.37
PCa[36:2]	<0.001	7.020	0.07	0.08	0.04	0.06	0.10	0.11	0.14	0.17	0.72
PC[36:1]	<0.001	7.014	0.44	0.49	0.45	0.44	0.57	0.46	0.67	0.75	1.37
PE[36:2]	<0.001	6.994	0.12	0.13	0.09	0.12	0.15	0.22	0.17	0.22	0.60
PC[36:5]	<0.001	6.598	0.16	0.22	0.18	0.22	0.24	0.23	0.38	0.25	0.31
LPC[18:0]	<0.001	6.269	0.04	0.04	0.06	0.09	0.08	0.08	0.10	0.14	0.10
PC[38:2]	<0.001	5.492	0.02	0.04	0.04	0.11	0.08	0.16	0.00	0.04	0.19

Continued on next page

Supplementary Table E1 (continued)

Lipid	p-value	z-score	C1 10%	C2 26%	C3 6%	C4 21%	F1 14%	F2 6%	F3 6%	F4 6%	F5 5%
Cer[d18:1/18:0]	<0.001	4.880	0.01	0.00	0.02	0.01	0.02	0.09	0.01	0.00	0.06
PC[40:6]	<0.001	4.875	0.04	0.03	0.06	0.00	0.05	0.00	0.14	0.09	0.12
PE[38:6]	<0.001	4.863	0.02	0.12	0.05	0.43	0.18	0.31	0.16	0.11	0.42
PC[36:2]	<0.001	4.814	0.74	0.99	0.72	0.75	1.04	0.71	1.72	1.44	1.63
PE[38:5]	<0.001	4.163	0.11	0.15	0.10	0.16	0.15	0.18	0.16	0.20	0.22
PC[44:11]	<0.001	3.505	0.02	0.04	0.03	0.06	0.04	0.06	0.03	0.05	0.06
PE[36:4]	0.001	3.327	0.04	0.05	0.04	0.06	0.05	0.06	0.04	0.06	0.08
Cer[d18:1/18:1]	0.001	3.243	0.01	0.01	0.03	0.02	0.03	0.06	0.01	0.03	0.03
LPC[18:2]	0.005	2.795	0.03	0.02	0.06	0.05	0.06	0.12	0.03	0.03	0.02
LPC[18:1]	0.008	2.656	0.06	0.06	0.08	0.10	0.09	0.18	0.05	0.12	0.07
PG[34:1]	0.014	2.451	2.51	3.02	2.45	2.89	2.83	3.44	2.82	3.77	4.85
PC[38:6]	0.037	2.085	0.10	0.12	0.07	0.09	0.12	0.07	0.33	0.16	0.12
PC[32:2]	0.057	1.907	0.12	0.16	0.13	0.13	0.15	0.25	0.14	0.21	0.41
PC[36:4]	0.157	1.415	0.54	0.81	0.57	0.67	0.67	0.53	1.21	0.81	0.78
LPC[18:3]	0.163	1.395	0.03	0.05	0.09	0.14	0.08	0.13	0.04	0.04	0.03
PCa[32:0]	0.173	1.364	0.89	0.81	0.86	0.84	0.73	0.70	0.60	0.79	1.11
PE[36:3]	0.174	1.361	0.11	0.13	0.11	0.13	0.11	0.12	0.11	0.14	0.19
PCa[34:0]	0.222	1.221	0.72	0.65	0.76	0.70	0.58	0.54	0.47	0.53	0.55
LPC[24:0]	0.328	0.979	0.01	0.01	0.02	0.01	0.02	0.04	0.00	0.01	0.01
LPC[16:0]	0.581	0.553	0.20	0.30	0.35	0.54	0.36	0.63	0.22	0.23	0.19
PE[32:0]	0.611	-0.509	0.06	0.05	0.07	0.06	0.05	0.00	0.00	0.07	0.10
PI[36:1]	0.214	-1.242	0.63	1.00	0.01	1.74	0.99	0.43	0.01	0.79	0.01
PE[36:5]	0.022	-2.299	0.07	0.07	0.08	0.07	0.07	0.08	0.05	0.06	0.06
PC[36:3]	0.017	-2.397	0.47	0.60	0.34	0.30	0.48	0.28	0.87	0.45	0.28
PC[34:2]	0.012	-2.524	1.96	2.48	1.76	2.04	2.21	1.54	3.33	2.01	1.52
LPC[16:1]	0.009	-2.631	0.02	0.01	0.04	0.01	0.01	0.05	0.00	0.00	0.00
PI[36:2]	0.001	-3.255	0.56	0.95	0.00	1.35	0.71	0.01	0.01	0.97	0.01
PC[34:4]	0.001	-3.404	0.24	0.35	0.28	0.33	0.30	0.31	0.20	0.26	0.29
PE[34:0]	0.001	-3.421	0.89	0.81	0.86	0.84	0.73	0.70	0.60	0.79	1.11
PG[36:1]	<0.001	-5.898	1.90	1.94	1.34	1.81	1.70	1.65	1.29	1.26	0.01
PC[34:1]	<0.001	-6.205	5.43	6.64	5.19	5.45	5.62	4.49	5.13	5.36	4.99
PG[36:2]	<0.001	-7.237	1.44	1.82	1.09	1.40	1.35	0.66	0.01	0.87	0.01
PC[34:3]	<0.001	-8.571	1.99	2.52	2.06	2.68	2.05	1.73	1.38	1.18	0.88
PC[34:0]	<0.001	-8.822	1.48	1.55	1.40	1.28	1.36	1.08	0.97	1.22	1.05
PC[32:1]	<0.001	-11.653	4.17	4.91	3.33	3.54	3.13	2.50	2.09	2.10	1.46
PC[30:0]	<0.001	-12.117	4.26	4.08	3.68	3.44	2.96	2.54	2.21	2.13	1.32
PC[32:0]	<0.001	-12.463	35.52	31.87	34.78	32.43	29.04	24.71	19.68	18.76	13.72

Heat map of lipidomics trends across the TDA structure, sorted by p-value and z-score of a Jonckheere-Terpstra test for ranked alternatives. Values in the table show the median relative abundance of a lipid for each TDA group. Note that unidentified ions were not included in this table, and see main text for lipid nomenclature and the abbreviations used.

Supplementary Table E2

Clinical and pathobiological data	p-value	z-score	C1 10%	C2 26%	C3 6%	C4 21%	F1 14%	F2 6%	F3 6%	F4 6%	F5 5%
Astmha severity	<0.001	6.366	0.58	0.76	0.71	0.88	0.93	0.94	1.00	0.94	1.00
Inhaled corticosteroid dose (mg day-1)	<0.001	5.917	0.4	0.5	0.6	0.6	0.8	0.8	0.8	0.8	1.0
Sputum eosinophils (%)	<0.001	5.857	0.2%	0.7%	0.2%	2.0%	4.4%	1.4%	26.8%	3.5%	2.9%
ACQ7 score	<0.001	5.026	1.5	2.5	2.5	3.0	3.2	3.8	4.2	3.5	4.5
Sputum neutrophils (%)	<0.001	4.907	40.5%	45.7%	43.9%	43.2%	59.5%	60.8%	59.6%	83.0%	91.4%
Exacerbations (past 12 months)	<0.001	4.445	0.8	1.1	1.2	1.7	2.1	1.7	1.9	2.2	2.9
Blood leukocytes (µL-1)	<0.001	4.306	6000	6900	6050	6900	7800	9450	7900	9350	9100
Oral corticosteroid dose (mg day-1)	<0.001	4.247	2.6	2.1	1.9	4.1	5.4	6.7	6.0	5.9	3.4
Age	<0.001	3.994	45	48	48	49	55	54	58	57	56
Blood neutrophils (µL-1)	<0.001	3.849	3200	4000	3500	3900	4750	5700	4700	5300	5550
Exhaled nitric oxide (ppb)	0.003	2.970	23	20	27	27	32	24	54	23	19
Blood eosinophils (µL-1)	0.006	2.756	100	200	200	200	200	200	400	250	250
ICU admission (ever)	0.006	2.748	0.1	0.1	0.0	0.2	0.1	0.1	0.1	0.3	0.6
SNOT score	0.007	2.685	22	24	23	26	30	29	26	33	29
Serum IgE (mL-1)	0.160	1.405	86	89	68	109	93	91	120	91	129
Age at first diagnosis	0.194	1.299	24	17	24	22	21	16	43	35	14
HADS score	0.850	0.189	8.8	10.6	12.1	11.3	13.8	10.8	9.6	9.5	9.6
Smoking status (% current or ex-smokers)	0.752	-0.316	0%	11%	4%	14%	14%	7%	0%	6%	0%
Positive atopy test (%)	0.512	-0.656	37%	38%	26%	44%	43%	33%	53%	37%	36%
Sex (% male)	0.488	-0.734	0.4	0.3	0.7	0.6	0.5	0.6	0.6	0.4	0.2
BMI	0.111	-1.592	25	30	27	25	27	27	26	28	24
Sputum squamous epithelial cells (%)	0.170	-2.376	9.5%	5.3%	13.3%	6.7%	9.9%	6.9%	3.6%	7.8%	0.9%
FEV1/FVC ratio (predicted)	<0.001	-3.510	79.3	78.9	79.4	79.3	77.6	77.9	77.1	77.9	77.9
FVC (% predicted)	<0.001	-3.643	107	98	98	96	97	89	93	90	75
Sputum lymphocytes (%)	<0.001	-3.707	1.4%	1.1%	1.3%	1.2%	0.9%	0.8%	0.7%	0.6%	0.4%
sGAW (1/kPA x sec)	<0.001	-4.162	1.18	1.29	1.06	0.86	0.77	0.84	0.72	0.72	0.25
FEV1 (% predicted)	<0.001	-5.495	91	80	82	79	69	66	62	60	42
FEF25-75 (predicted)	<0.001	-5.699	3.71	3.56	3.64	3.76	3.34	3.35	3.29	3.36	3.29
Sputum macrophages (%)	<0.001	-8.764	57.6%	49.0%	54.8%	41.8%	25.7%	20.3%	12.1%	9.1%	4.8%

Heat map of trends in the demographic, clinical and pathobiological data across the sputum lipidomics TDA structure, with p-values and z-scores of a Jonckheere-Terpstra test for ranked alternatives. Values in the table show the median for each TDA group. See Shaw *et al*. (2015) for a description of the variables and methods used in the U-BIOPRED study. Abbreviations: BMI = Body Mass Index; IgE = Immunoglobulin E; ACQ = Asthma Control Questionnaire; ICU = intensive care unit; SNOT = SinoNasal Outcomes Test; HADS = Hospital Anxiety and Depression Scale; FEV1 = forced expiratory volume in 1 second; FVC = forced vital capacity; FEF 25-75 = forced expiratory flow at 25-75% of the pulmonary volume; sGAW = specific airway conductance. Asthma severity is expressed as the ratio of any asthmatic (MMA, SAc/ex and SAn) versus healthy (HC) participants; systemic dosage of corticosteroids for severe asthmatic participants is expressed in prednisolone-equivalent doses.

Supplementary Table E3

Protein biomarker	Sample	Assay	p-value	z-score	C1 10%	C2 26%	C3 6%	C4 21%	F1 14%	F2 6%	F3 6%	F4 6%	F5 5%
IL-8	Plasma	MSD	<0.001	4.031	2.79	3.13	2.73	3.01	3.03	3.73	3.24	3.87	3.97
Serpin-E1	Serum	Luminex	<0.001	3.875	77469	85894	86691	90934	91605	101345	96137	111359	91426
MCP-4	Serum	Luminex	<0.001	3.740	98.3	127.2	111.9	128.2	131.7	145.1	146.0	129.9	159.6
Periostin	Serum	Elecsys	0.001	3.328	48.3	45.0	44.3	44.5	47.9	53.7	61.3	50.4	53.9
Eotaxin	Plasma	MSD	0.001	3.326	87.8	93.9	92.7	106.0	106.5	138.0	104.5	113.0	135.0
CCL18	Serum	Impact	0.001	3.185	134.9	161.3	164.1	191.4	192.1	186.0	221.3	179.8	276.6
Galectin-3	Serum	Luminex	0.003	3.011	5288	5486	5182	5298	5722	6256	5112	6232	6740
IL-6	Plasma	MSD	0.004	2.906	0.52	0.91	0.61	0.73	0.60	0.89	1.07	1.37	1.23
IL-17AA	Serum	Singulex	0.005	2.806	0.29	0.30	0.36	0.37	0.35	0.38	0.46	0.53	0.36
CCL17	Plasma	MSD	0.006	2.754	55.7	69.3	51.6	63.2	65.6	129.0	54.7	117.0	73.8
IL-13	Serum	Impact	0.006	2.726	0.49	0.49	0.54	0.62	0.60	0.47	1.32	0.53	0.48
MCP-1	Plasma	MSD	0.008	2.641	93.6	98.7	102.0	98.2	95.7	125.0	110.5	109.0	112.5
hs-CRP	Serum	Immulite	0.010	2.577	1.2	1.4	1.2	1.6	1.2	1.4	1.3	6.2	2.8
TNF-alpha	Plasma	MSD	0.016	2.400	1.56	1.86	1.68	1.89	1.71	1.83	1.96	1.85	2.12
C5a	Serum	Luminex	0.020	2.335	33.7	38.2	48.9	36.6	39.8	39.0	35.6	54.4	49.5
MMP-3	Serum	Luminex	0.025	2.245	13817	13569	13730	14477	17158	23242	20967	12788	15431
MIP-1b	Plasma	MSD	0.035	2.114	46.2	49.4	44.1	45.7	45.2	67.8	52.1	61.8	54.6
Eotaxin-3	Plasma	MSD	0.058	1.894	17.8	15.4	15.5	19.1	18.4	18.0	26.9	14.2	15.4
INF-gamma	Plasma	MSD	0.067	1.834	4.38	5.77	5.40	4.67	4.80	5.73	7.41	7.43	11.02
CD40L	Serum	Luminex	0.104	1.624	4717	4287	4750	4698	4910	5275	5546	4725	4253
IL-6R-alpha	Serum	Luminex	0.127	1.525	10398	10927	11221	11227	11696	12783	11108	10787	10884
IL-1-alpha	Serum	Luminex	0.231	1.199	34.1	33.6	32.5	35.9	34.1	36.6	37.5	31.9	32.3
Alpha-1-microglobulin	Serum	Luminex	0.257	1.134	6355	6551	7592	7056	6935	7610	6889	6971	7392
IL-18	Serum	Luminex	0.267	1.110	192.4	219.0	214.4	222.0	210.7	209.1	218.1	244.7	221.9
IP-10	Plasma	MSD	0.347	0.940	224	326	297	278	250	234	305	353	473
CD30	Serum	Luminex	0.377	0.883	31.7	38.0	38.1	38.5	35.7	37.1	36.3	40.7	40.5
LBP	Serum	Luminex	0.466	0.729	2056309	2166450	2160387	2019250	1999580	2623438	1722458	2770901	2508327
Lumican	Serum	Luminex	0.502	0.672	130969	133245	132268	133135	128644	141913	148609	125866	141143
RAGE	Serum	Luminex	0.809	0.242	1358	1298	1382	1275	1250	1265	1151	1372	1585
SHBG	Serum	Luminex	0.512	-0.655	4637586	2777497	1693569	2969730	2709649	1989644	2396730	2470851	3401339
CCL22	Plasma	MSD	0.502	-0.672	869	890	810	794	700	858	829	880	893
DPPIV	Serum	Luminex	0.216	-1.238	90742	101947	100085	101725	95053	92193	93394	86834	91931

Heat map of trends in blood protein biomarker levels across the lipidomics TDA structure, sorted by p-value and z-score of a Jonckheere-Terpstra test for ranked alternatives. Values in the table show the median concentration of a given protein for each TDA group.

Supplementary Table E4

Gene	p-value	z-score	C1 10%	C2 26%	C3 6%	C4 21%	F1 14%	F2 6%	F3 6%	F4 6%	F5 5%
RAB1B	0.011	2.535	0.00	0.00	0.00	0.00	0.00	2.00	1.71	2.98	2.35
PLA2R1	0.017	2.381	0.00	0.00	0.00	0.00	0.00	0.00	2.22	2.14	3.12
SYVN1	0.017	2.381	0.00	0.00	0.00	0.00	0.00	0.00	2.26	2.25	3.28
CD24	0.021	2.304	0.00	-0.65	0.00	0.00	0.00	1.89	0.00	2.84	2.31
HSP90B1	0.025	2.234	0.00	0.71	0.79	0.00	0.00	2.14	2.51	2.52	2.21
DNMT3B	0.037	2.085	0.00	-0.51	1.34	-0.49	1.74	2.33	1.90	1.46	2.14
DAP3	0.168	1.379	0.00	0.00	-2.00	-2.45	0.00	0.00	0.00	0.00	2.00
ERG	0.242	1.170	0.00	-1.09	0.94	-0.69	-0.54	0.00	0.00	-0.28	1.13
IL13	0.242	1.170	0.00	-1.63	-2.19	-3.62	-2.32	0.00	-0.92	0.00	2.76
CXCL8	0.259	1.128	0.00	1.42	0.00	0.00	0.00	0.00	0.00	2.36	1.91
MYC	0.259	1.128	0.00	-1.13	0.00	-1.28	0.00	0.00	0.67	0.00	0.00
NONO	0.357	0.922	0.00	0.00	0.00	0.38	1.63	1.41	0.00	0.00	0.82
EGLN	0.380	0.877	0.00	1.19	0.00	0.00	0.00	0.00	1.86	2.44	0.00
HDAC1	0.380	0.877	0.00	-1.00	0.00	-0.56	1.02	0.00	0.00	0.00	0.00
HEXIM1	0.380	0.877	0.00	0.00	0.00	0.20	0.00	2.00	0.00	0.45	0.00
TCF7L2	0.381	0.876	0.00	0.00	-1.13	-1.65	-2.08	-2.61	0.00	-1.80	0.00
EOMES	0.531	0.627	0.00	-2.45	0.00	0.00	0.28	-0.94	0.00	0.00	0.00
PRL	0.707	0.376	0.00	0.00	0.00	3.48	0.00	4.97	2.53	0.00	0.00
RXRA	0.707	0.376	0.00	-0.15	0.00	0.00	0.00	-1.67	0.00	-0.24	0.00
P38 MAPK	0.900	0.125	0.00	0.00	0.00	-3.11	-2.86	0.00	-2.18	0.00	0.00
MGEA5	0.915	0.106	0.00	1.81	2.79	2.59	2.27	2.34	0.00	2.89	0.00
GATA6	1.000	0.000	0.00	0.00	0.00	0.00	-0.92	-0.59	0.00	-0.08	1.14
HIF1A	1.000	0.000	0.00	-0.89	0.00	-2.00	-1.96	0.00	-2.05	0.00	0.00
PGR	1.000	0.000	0.00	0.00	0.80	1.77	1.11	0.00	0.00	0.00	0.41
USP7	0.908	-0.116	0.00	-2.00	0.00	0.00	-2.83	-2.24	0.00	-2.00	0.00
SMARCD3	0.707	-0.376	0.00	0.00	-1.34	-1.13	0.00	0.00	0.00	0.00	-1.98
TP73	0.531	-0.627	0.00	0.68	0.00	1.66	2.00	0.00	0.00	0.00	0.00
INHBA	0.489	-0.691	0.00	0.82	0.00	0.66	0.00	0.00	-1.82	0.73	0.00
RARA	0.489	-0.691	0.00	1.22	0.00	-0.78	0.08	-2.16	0.00	0.00	0.00
POU2F2	0.311	-1.014	0.00	0.00	0.90	0.32	0.00	0.00	0.00	0.00	0.00
GATA1	0.300	-1.037	0.00	0.00	0.00	0.00	-0.20	-2.21	0.00	0.00	-1.40
IFNA	0.273	-1.095	0.00	2.74	1.95	3.73	0.00	2.77	0.00	-1.89	0.00
IGFBP2	0.273	-1.095	0.00	0.00	0.00	0.03	0.18	-2.77	0.00	-0.98	-1.35
GLI1	0.249	-1.152	0.00	0.10	0.00	-0.04	0.00	-2.31	0.00	-0.92	0.00
IL1RN	0.249	-1.152	0.00	0.00	0.00	-2.52	0.00	-2.21	-2.63	-2.43	0.00
Cg	0.167	-1.383	0.00	-0.53	0.00	0.00	0.00	0.00	-3.24	-2.84	-1.87
ITGB1	0.167	-1.383	0.00	-0.15	1.93	0.00	0.00	0.00	-1.86	0.00	-2.11
JUN	0.107	-1.613	0.00	-1.59	0.00	0.00	0.00	0.00	-1.95	-2.14	-1.77
EPAS1	0.103	-1.629	0.00	0.00	0.00	0.00	0.00	-2.57	-1.93	0.00	-1.95
HGF	0.103	-1.629	0.00	2.00	0.52	2.00	0.00	0.00	0.00	0.00	0.00
CST5	0.080	-1.753	0.00	-0.39	0.00	0.00	-0.25	0.00	-1.94	-3.17	-1.46
miR-10	0.049	-1.972	0.00	1.05	0.24	0.00	0.06	-0.73	0.00	0.00	-1.86
miR-122	0.021	-2.304	0.00	0.00	-1.83	0.00	0.00	0.00	-2.48	-3.00	-3.69
WT1	0.021	-2.304	0.00	1.92	0.00	0.06	0.00	0.00	0.00	-0.38	-0.62
SMARCA4	0.016	-2.410	0.00	0.00	0.00	-1.20	0.00	-3.17	-1.76	-4.09	-2.95
NANOG	0.016	-2.411	0.00	-1.46	0.08	0.00	-1.29	-1.79	-3.21	-2.77	-2.87
KDM5B	0.011	-2.535	0.00	1.90	0.00	0.00	0.00	0.00	-2.39	-2.24	-3.09
ETS1	0.008	-2.668	0.00	0.49	0.00	0.00	0.00	0.00	-1.84	-2.07	-2.07
SMAD3	0.004	-2.848	0.00	0.00	0.00	0.00	-1.64	-2.19	-2.14	-2.37	-2.32

Upstream transcriptional regulators in matched sputum cell pellets from the U-BIOPRED cohort (n=97), as determined by Ingenuity Pathway Analysis for the sputum lipidomics TDA groups. All fold changes in gene expression are relative to the basal TDA group C1 (‘healthy’). The heat map is sorted by p-values and z-score of a Jonckheere-Terpstra test for ranked alternatives.