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 BSc[b], Victoria M. Goss PhD[a,b], Grielof Koster PhD[a,b], Per S. Bakke MD[f], Massimo Caruso MD[g], Pascal Chanez MD[h], Sven-Erik Dahlén MD[i], Stephen J. Fowler MD[j,k], Ildikó Horváth MD[l], Peter H. Howarth MD[a,b], Norbert Krug MD[m], Paolo Montuschi MD[n], Marek Sanak PhD[o], Thomas Sandström MD[p], Dominick E. Shaw MD[q], Kian Fan Chung MD[r], Florian Singer MD[s], Louise J. Fleming MD[r], Ian M. Adcock PhD[r], Ioannis Pandis PhD[d], Aruna T. Bansal PhD[t], Julie Corfield MSc[u] Ana R. Sousa PhD[v], Peter J. Sterk MD[w], Graham Roberts[a,b], Ruben Sanchez Garcia PhD[e], Paul J. Skipp PhD[c], Anthony D. Postle PhD[a,b], Ratko Djukanović MD[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] Department of Medicine, Jagiellonian University, Krakow, Poland

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

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

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

[s] Department of Pediatrics, Bern University Hospital, Bern, Switzerland

[t] Acclarogen Ltd, St John's Innovation Centre, Cambridge, United Kingdom

[u] Areteva Ltd, Nottingham, United Kingdom

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

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

*** 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).

Disclosure of potential conflict of interest: PJ Sterk was the recipient of the U-BIOPRED grant from the Innovative Medicines Initiative (IMI), which was jointly funded by the European Union (EU) and the European Federation of Pharmaceutical Industries and Associations (EFPIA). P Chanez has provided consultancy services, served on advisory boards and/or received lecture fees from Boehringer Ingelheim,

GSK, Centocor, ALK, AstraZeneca, Novartis, TEVA, Chiesi, Sanofi, Boston Scientific and SNCF. R. Djukanović has provided consultancy services for AstraZeneca, TEVA and Novartis and has given talks at symposia organised by TEVA and Novartis. He is a co-founder of the University of Southampton spinout company Synairgen, where he is a consultant and has shares.

The remaining authors declare that they have no relevant conflicts of interest.

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.

Total word count: 3302

Abstract

Background: Asthma is a chronic respiratory disease with significant heterogeneity in its clinical presentation and pathobiology. There is need for improved stratification of asthma patients and a clearer understanding of the biomolecular processes underlying asthma pathogenesis.

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. Matching clinical, pathobiological, proteomic and transcriptomic data informed on the underlying disease processes.

Results: Shotgun lipidomics of induced sputum supernatant revealed nine distinct molecular phenotypes. These phenotypes aligned broadly with the asthma severity spectrum, from healthy controls and mild-to-moderate, well-controlled asthmatics to those with severe eosinophilic or neutrophilic disease. The lipidomic signature characteristic of pulmonary surfactant did not differ significantly between phenotypes. However, the sputum lipid phenotypes showing an additional, cell-derived lipid signal were associated with significantly worse asthma severity and lung function, and elevated granulocyte counts.

Conclusion: We propose a novel mechanism of increased lipid loading in the epithelial lining fluid of severe asthmatics, resulting from the secretion of extracellular

vesicles by granulocytic inflammatory cells. Such a mechanism would not only reduce the ability of pulmonary surfactant to lower surface tension in asthmatic small airways, but also compromise its role as an immune regulator.

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 severely asthmatic airways.

Key Messages:

- The sputum lipidome comprises nine distinct molecular phenotypes along an asthma severity spectrum
- Sputum lipid phenotypes in severe asthma contain elevated levels of cell-derived, non-surfactant lipids
- Secretion of lipid-rich extracellular vesicles by granulocytic inflammatory cells is proposed as a novel mechanism of pulmonary surfactant dysregulation in severe asthma and a potential target for new therapeutics

Key words: 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
HexCer	Hexosyl-ceramide
IgE	Immunoglobulin E
IPA	Ingenuity Pathway Analysis
JT-test	Jonckheere-Terpstra test
LC-MS/MS	Liquid chromatography tandem mass spectrometry
LPC	Lyso-phosphatidylcholine
MDS	Multi-dimensional scaling
MMA	Mild-to-moderate asthmatic
PC	Phosphatidylcholine

PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PS	Phosphatidylserine
SM	Sphingomyelin
QC	Quality control
SAc/ex	Current or ex-smoking severe asthmatic
SAn	Non-smoking severe asthmatic
TDA	Topological data analysis
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.[1-2] 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.[3] 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.[4-6] Better stratification will open up new directions for targeted drug development and more personalised disease management strategies.[2]

The U-BIOPRED (Unbiased Biomarkers for the Prediction of Respiratory Disease Outcomes) study[7,8] has employed an 'unbiased' multi-omics systems biology approach to stratify patients, elucidate biochemical pathways, and define new sets of diagnostic molecular biomarkers.[9-13] The current study is in keeping with this broad objective and provides the first ever comprehensive analysis in asthma of the lipid composition of induced sputum supernatant, a validated means of sampling epithelial lining fluid. Its aim is to stratify the heterogeneous U-BIOPRED cohort of asthma patients and healthy controls according to their sputum lipid molecular phenotypes.

Sputum supernatant comprises pulmonary surfactant and soluble material secreted by immune cells and the respiratory epithelium within the lungs, mixed 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.[14-16] Thus, its lipid composition matches its primary source: pulmonary surfactant secreted by ATII cells in the alveolar epithelium.[17-19] 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.[20,21] However, despite significant interest in the role of airway lipids, studies of the sputum lipidome remain scarce.[14-16,22-25] We have previously reported the lipid composition of sputum supernatant in healthy adults,[16] and now extend our analysis to U-BIOPRED participants with asthma of varying severity. Semi-quantitative shotgun lipidomic measurements were clustered using Topological Data Analysis (TDA) and complemented with matching 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 distinct sputum lipid phenotypes, highlighting biological mechanisms that may 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.[8] Processed biological samples from clinical sites were blinded and

stored in a central biobank. All clinical, laboratory and 'omics data collected for U-BIOPRED were stored on the tranSMART knowledge management platform for use by its study members.

Lipid analysis and data processing

Shotgun lipidomics of induced sputum samples was conducted as previously described[16] (for details, see Supplementary Materials). Briefly, lipids were extracted using the Bligh-Dyer extraction protocol,[26] 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 done in full scan mode for both positive and negative ionisation, with blank injections and pooled QC samples run for quality control. 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 (Waters, Milford, MA, USA) with a gradient of methanol and water. Data-independent product ion scans were acquired over the entire gradient using broadband collision induced dissociation. Precursor and fragment ions were matched by chromatographic retention time and using the well-established fragmentation rules for lipids for identity confirmation.[27]

After removal of ions with <60% detection rate, data were corrected for analytical batch variation using the *R* script "SVA ComBat".[28] 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, data were also normalised to the amount of total lipid.[29,30]

Data analysis and statistics

Participants with comparable sputum lipid profiles were grouped as previously described[11,31,32] in order to identify lipid phenotypes, using the AyasdiAI TDA Workbench (Symphony AyasdiAI, Palo Alto, CA, USA). Groups of participants with similar sputum lipid profiles were defined within the TDA structure using density mode clustering.[33] Statistical significance of trends and between-groups differences in lipidomic and supporting clinical and 'omics data were assessed in SPSS Statistics 24 (IBM, Armonk, NY, USA), using either the non-parametric Jonckheere-Terpstra test for ordered alternatives or the Mann-Whitney U test. For the limited number of nominal variables, a Pearson's Chi-squared test was used. Trends or differences were considered significant if $p<0.05$, and highly significant if $p<0.001$ (after Bonferroni correction).

Blood protein data

Plasma concentrations of 11 proteins were measured by electrochemiluminescence assay, and a further 21 proteins were measured in serum using a number of different immunoassays.

Pathway analysis of sputum cell pellet transcriptomics data

Transcriptomic data from RNA extracts of 97 matching sputum cell pellets[34] were subjected to Ingenuity Pathway Analysis (IPA; QIAGEN Bioinformatics, Redwood City, CA, USA) to identify upstream regulators of differential gene expression in each lipid phenotype. IPA core analysis was performed on the 4000 most differentially expressed genes, using the ordered TDA groups as described above. Results were subjected to comparison analysis to identify trends of IPA-predicted upstream regulator activation/inhibition across the sputum lipidomics TDA structure.

Results

Study cohort

A total of 252 participants 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. 137 females and 115 males of predominately white origin were categorised[8] 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[16] (see Supplementary Materials for methodology, QC and selection procedures). Of these, 92 lipid species could be identified using LC-MS/MS analysis of pooled QC samples, comprising on average 95% 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 indicated 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" and "flare" sets were re-analysed individually. This produced a ring-like structure for the core set, consisting of four connected groups (labelled C1 to C4; Fig. 1B), whereas the flare set comprised a V-shaped string of five distinct groups (labelled F1 to F5; Fig. 1B). Edges connecting the flare set to the core set were restricted to C3, C4, F1 and F2 in the central part of the structure. The nine identified groups constituted 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 ordered alternatives was performed (J-T test). C1 and F5 were selected as respectively the first and last group of the series, 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). This analysis yielded highly significant trends from C1 to F5 for most of the measured lipids (Fig. 2 and Supplementary Table S1). 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 and triacylglycerols.

Trends in relative abundance were not always matched by the actual lipid concentration data. Importantly, levels of the surfactant-specific lipid DPPC and other di-saturated PC-species were comparable across all TDA groups (JT-test for DPCC: p=0.823, z-score: 0.223). Moreover, some differences between phenotypes did not conform to the general trend described. For example, relative to the other core groups, C2 was specifically 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 (Supplementary Table S1).

Trends in clinical and pathobiological metadata

Trend analyses were also performed for a range of supporting data, including demographic and clinical measurements, serum protein levels and sputum cell pellet gene expression. There was a highly significant trend in asthma severity (JT-test, $p<0.001$, z-score: 6.336), as judged by the proportion of participants from each of the four clinically characterised U-BIOPRED categories. The proportion of healthy controls was highest in basal group C1 (42%) and decreased progressively through the intermediate groups to 12% in C4, to only 6-7% in F1, F2 and F4. The two terminal flare groups F3 and F5 contained only mild-to-moderate or severe asthmatics, but no healthy participants. Among the clinical variables overlaid as metadata onto the TDA structure (Supplementary Table S2), highly significant negative associations 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. The sputum differential cell counts showed reciprocal increases in eosinophil and neutrophil counts and significant decreases in macrophage and lymphocyte counts. Although median sputum eosinophil counts reached 3% of total inflammatory cells in four of the TDA groups, a threshold viewed as clinically relevant,[35] they were significantly higher in F3 (26.8%) than in any other group. Neutrophil levels gradually increased from 40% in C1 to 60% in groups F1-F3 and peaked in F4 and F5 (medians 83% and 91%, respectively).

Of the 32 serum protein biomarkers of inflammation available for this analysis, more than half increased significantly from C1 to F4/F5 (Supplementary Table S3). Some between-group differences were independent of the overall trend: e.g., serum Eotaxin-3 and IL-13 were associated with the high-eosinophil group F3, whereas CCL17 and Galectin-3 in that group did not differ from the basal group C1. Finally, a number of upstream transcriptional regulators predicted by pathway analysis of the sputum cell

pellet transcriptome also showed consistent, significant trends across the TDA structure (Supplementary Table S4). 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, most of the upstream transcriptional regulators appeared specific to discrete, smaller parts of the structure.

Discussion

Comprehensive, cross-sectional assessment of sputum lipid biomarkers identified a continuous spectrum of molecular phenotypes from health and mild, well-controlled asthma to severe neutrophilic and eosinophilic asthma (Figs. 1 and 3). Our key finding is a progressive increase, from basal group C1 to flare groups F4 and F5, of non-surfactant lipids, including alkyl-acyl PCs, longer-chain/polyunsaturated fatty acid-containing PCs, various PE and PS species, sphingolipids, and neutral lipids (Chol, CE, DG and TG species) (Fig. 2). This increase was associated with significantly worse asthma severity and lung function, and sputum neutrophilia or eosinophilia in the flare groups. Given that absolute concentrations of di-saturated PC species, the main surface-active component of pulmonary surfactant, did not vary significantly between TDA groups, we conclude that surfactant production itself is not significantly altered in asthma.

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. Most lipids identified in this study are not secreted by ATII cells as part of pulmonary surfactant,[36] and are likely derived from

353 immune and epithelial cells. Sputum dilution by saliva was limited, with low levels of
354 squamous epithelial cells derived from the upper airways[37,38] (average 10%), and low
355 concentrations of neutral lipids, which predominate in saliva.[39] In contrast, highly
356 significant associations were seen between the lipidomic trends and inflammatory cell
357 numbers (Supplementary Table S2). The flare groups all contained high numbers of
358 granulocytic inflammatory cells, predominately eosinophils in F3, neutrophils in F4 and
359 F5, and a combination of both in F1 and F2 (Fig. 3). Whilst this suggests that the
360 sputum lipidome in the flare groups was enriched by material from airway
361 granulocytes, the sampling protocol required removal (by centrifugation) of whole
362 cells, and samples rich in dead or damaged cells were excluded from analysis as part
363 of the QC. Thus, the non-surfactant lipid material in these granulocytic phenotypes
364 was most likely derived from either small membrane fragments or secreted
365 extracellular vesicles (EVs).

366 Knowledge of the lipid pathobiology of neutrophils and eosinophils is limited.
367 Neutrophils are rich in PC, PE, PS, and PI, with high levels of mixed alkyl-acyl bonds,
368 as well as SM and Chol.[40-42] In eosinophils, there has been no systematic profiling of
369 the lipidome, with research mainly focusing on activation-induced formation of
370 intracellular lipid droplets and their metabolism of arachidonic acid as the precursor
371 for pro-inflammatory lipid mediators.[43-45] Both neutrophils and eosinophils release Evs
372 in response to inflammatory stimuli, either as endosome-derived exosomes or outer
373 membrane-derived micro-vesicles.[46,47] Such Evs are rich in sphingolipids, PS and
374 neutral lipids and may further reflect the lipid compositions of their progenitor cells.[48-
375 50] The TDA flare groups were all significantly enriched in lipids that fit this inflammatory
376 cell profile (Supplementary Table S1). For example, levels of arachidonic acid-
377 containing lipids such as PC[16:0/20:4] and PC[18:0/20:4], as well as Chol and CE,

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.

The chemokine eotaxin, which is overproduced in asthmatic airways, [51] recruits both eosinophils and stimulates the formation of lipid droplets in these cells. Such lipid droplets not only act as sites of lipid mediator synthesis,[52,53] but may selectively contribute lipid material for exosome formation, mediated by eotaxin and other inflammatory stimuli.[46] A similar mechanism was recently proposed for neutrophil-derived exosomes, which were shown to contribute to airway smooth muscle remodelling.[47] In the current study, circulating levels of several chemokine and cytokine markers of inflammation were significantly higher in the serum of participants in the TDA flare groups (Supplementary Table S3). The upstream transcriptional regulator with the strongest positive trend across the TDA structure was RAB1B (Supplementary Table S4). Members of the Rab GTPases protein family are key regulators of intracellular membrane trafficking,[54] and are present on the membranes of lipid droplets.[55] 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,[56] and was shown to be overexpressed in the bronchial epithelium of children with atopic asthma.[57] We did not observe a direct effect of potentially increased phospholipase A2 activity on the sputum lipidome of groups F3-F5 (e.g., a measurable increase in lysophospholipid levels)[58,59]. 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 phenotype differences observed in the TDA core groups were more subtle than those in the flare groups (Figs. 1 and 3; Supplementary Table S1). Group C1 contained a mixed population of healthy and mild, well-controlled asthmatic participants with a sputum lipid profile that could be defined as 'normal'. The other core groups (C2-C4) contained smaller numbers of healthy participants, and mostly comprised a mixture of mild-to-moderate and non-inflammatory severe asthmatics. C4 represented a paucigranulocytic phenotype, with mean eosinophil and neutrophil counts of 2% and 43% respectively. Its lipid composition was intermediate between the healthy and severe granulocytic phenotypes, containing relatively less DPPC, elevated levels of the other lipid classes mentioned above, and specific enrichment in PI species. The lipid composition of group C3 was intermediate between C1 and C4, with notably low levels of PIs. The main pathobiological features distinguishing these groups of asthmatics were atopy and serum IgE levels, both of which were high in C4 and low in C3 (Supplementary Table S2), suggesting the presence of a distinct lipid phenotype for 'non-atopic' asthmatics.[60] Finally, group C2 had a lipid phenotype similar to that of 'healthy' group C1, but with relatively less DPPC and concomitantly increased levels of C18 fatty acid-containing PC species (Supplementary Table S1). This phenotype was characterised by a significantly 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 demonstrated dysregulation of lung lipid metabolism in obese animal models,[61-63] and we have previously reported a distinct sputum lipid phenotype in overweight, but otherwise healthy adults.[16]

Finally, we examined associations between the lipidome and upstream regulators of inflammation identified by IPA of sputum cell pellet RNA expression. In addition to upregulated *RAB1B* and *PLA2R1* in the TDA flare groups, significant associations were observed between asthma severity and expression of *SYVN1, CD24, HSP90B1* and *DNMT3B* (upregulation), and *miR-10*, *miR-122*, *WT1*, *SMARCA4*, *NANOG*, *KDM5B*, *ETS1* and *SMAD3* (downregulation) (Supplementary Table S4). Many of these have been implicated previously in asthma and other inflammatory diseases. The upregulation of *CD24*, expressed on granulocytes and lymphocytes, matched the high-neutrophil TDA groups F2, F4 and F5.[64] In contrast, *ETS1* is a negative regulator of Th17 cells, proposed to drive specific phenotypes of asthma.[65-68] Overexpression of *DNMT3B* promotes macrophage polarization into a 'classically-activated' M1 phenotype and enhances macrophage inflammation,[69] and this gene was upregulated in each of the asthmatic groups apart from C2 and C4. Wilms Tumour 1 (WT1) is known to regulate the expression of Matrix metalloproteinase-9 (MMP-9), an enzyme involved in extracellular matrix degradation and airway remodelling in asthma.[70] Our results show that this mechanism may be activated in the most severe neutrophilic asthmatics (F4, F5), with WT1 downregulation leading to increased MMP-9 mediated tissue remodelling.[71]

Summary and outlook

We have shown that lipidomic profiling of sputum supernatant can stratify respiratory asthma into molecular phenotypes and that the abundance of lipids that are non-endogenous to the pulmonary surfactant increases significantly with asthma severity. Pulmonary surfactant has a highly specific structure-function relationship, which depends crucially on its lipid and protein composition. Therefore, the observed mixing

of endogenous and cell-derived lipids in the TDA flare groups has implications for functional impairment of surfactant in asthmatic airways.[72] We propose a novel mechanism of surfactant dysregulation in severe asthmatics (Fig. 4), wherein granulocytes recruited into the airways are activated to produce both intracellular lipid droplets and EVs (exosomes and/or microvesicles). Upon release, these lipid and protein-rich EVs disrupt the tightly regulated structure of the pulmonary surfactant component of the epithelial lining fluid. This reduces the surfactant's ability to lower surface tension in the small airways, leading to collapsibility, and may additionally compromise its immunological function. Although our findings require external validation and verification of the proposed mechanism, immunomodulation of extracellular vesicle secretion by granulocytes in the lungs could provide a new and attractive therapeutic target for severe asthma.

Acknowledgements

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

References

1. McDowell JP, Heaney LG. Different endotypes and phenotypes drive the heterogeneity in severe asthma. Allergy 2020; 75: 302-310. PMID: 31267562
2. Ray A, Camiolo M, Fitzpatrick A, Gauthier M, Wenzel SE. Are We Meeting the Promise of Endotypes and Precision Medicine in Asthma? Physiol Rev 2021; 100: 983-1017. PMID: 31917651
3. 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
4. Anderson GP. Endotyping asthma: new insights into key pathogenic mechanisms in a complex, heterogeneous disease. Lancet 2008; 372: 1107-1119. PMID: 18805339
5. 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
6. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention. 2017: https://www.ginasthma.org
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. 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

9. 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
10. 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
11. 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
12. 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; 144: 70-82. PMID: 30928653
13. 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 2021; in press. PMID: 34824054
14. Dushianthan A, Goss VM, Cusack R, Grocott MPW, Postle AD. Bronchoalveolar lavage, tracheal wash and induced sputum surfactant phospholipid kinetics from healthy volunteers. BMC Pulm Med 2014; 14: 10. PMID: 24484629
15. 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: 4957-4964. PMID: 25884268

16. 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
17. Goss VM, Hunt AN, Postle AD. Regulation of lung surfactant phospholipid synthesis and metabolism. Biochim Biophys Acta 2012; 1831: 448-458. PMID: 23200861
18. Bernhard W. Lung surfactant: Function and composition in the context of development and respiratory physiology. Ann Anat 2016; 208: 146-150. PMID: 27693601
19. 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
20. 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
21. Shevchenko A, Simmons K. Lipidomics: coming to grips with lipid diversity. Nat Rev Mol Cell Biol 2010; 11: 593-598. PMID: 20606693
22. Sahu S, Lynn WS. Lipid composition of sputum from patients with asthma and patients with cystic fibrosis. Inflammation 1978; 3: 27-36. PMID: 581080
23. Wright SM, Hockey PM, Enhorning G, Strong P, Reid KBM, Holgate ST et al. Altered surfactant phospholipid composition and reduced lung function in asthma. J Appl Physiol (1985) 2000; 89: 1283-1292. PMID: 11007560
24. 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

25. 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

26. Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. Can J Biochem Physiol 1959; 37: 911-917. PMID: 13671378

27. 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

28. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007; 8: 118-127. PMID: 16632515

29. 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

30. 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

31. 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

32. 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

33. Strazzeri F. A Morse-theoretical clustering algorithm for annotated networks and spectral bounds for fuzzy clustering. 2018: PhD Thesis, University of Southampton, pp. 133.

34. 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

35. 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

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

37. 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

38. 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

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

589 40. Klock JC, Pieprzyk JK. Cholesterol, phospholipids, and fatty acids of normal
590 immature neutrophils: comparison with acute myeloblastic leukemia cells and
591 normal neutrophils. J Lipid Res 1979; 20: 908-911. PMID: 290722
592 41. Postle AD, Madden G, Clark GT, Wright SM. Electrospray ionisation mass
593 spectrometry analysis of differential turnover of phosphatidylcholine by human
594 blood leukocytes. Phys Chem Chem Phys 2004; 6: 1018-1021.
595 42. Leidl K, Liebisch G, Richter D, Schmitz G. Mass spectrometric analysis of lipid
596 species of human circulating blood cells. Biochim Biophys Acta 2008; 1781: 655-
597 664. PMID: 18723117
598 43. Weller PF, Monahan-Earley RA, Dvorak HF, Dvorak AM. Cytoplasmic lipid
599 bodies of human eosinophils. Am J Pathol 1991; 138: 141-148. PMID: 1846262
600 44. Isobe Y, Kato T, Arita M. Emerging roles of eosinophils and eosinophil-derived
601 lipid mediators in the resolution of inflammation. Front Immunol 2012; 3: 270.
602 PMID: 22973272
603 45. Melo RCN, Weller PF. Unravelling the complexity of lipid body organelles in
604 human eosinophils. J Leukoc Biol 2014; 96: 703-712. PMID: 25210147
605 46. Akuthota P, Carmo, LAS, Bonjour K, Murphy RO, Silva TP, Gamalier JP, et al.
606 Extracellular Microvesicle Production by Human Eosinophils Activated by
607 "Inflammatory" Stimuli. Front Cell Dev Biol 2016; 4: 117. PMID: 27833910
608 47. Vargas A, Roux-Dalvai F, Droit A, Lavoie J-P. Neutrophil-derived exosomes: a
609 new mechanism contributing to airway smooth muscle remodelling. Am J Respir
610 Cell Mol Biol 2016; 55: 450-461. PMID: 27105177
611 48. Skotland T, Sandvig K, Llorente A. Lipids in exosomes: current knowledge and
612 the way forward. Prog Lipid Res 2017; 66: 30-41. PMID: 28342835

613 49. Hough KP, Wilson LS, Trevor JL, Strenkowski JG, Maina N, Kim YI, et al. Unique
614 lipid signatures of extracellular vesicles from the airways of asthmatics. Sci Rep
615 2018; 8: 10340. PMID: 29985427
616 50. Van Niel G, D'Angelo G, Raposo G. Shedding light on the cell biology of
617 extracellular vesicles. Nat Rev Mol Cell Biol 2018; 19: 213-228. PMID: 29339798
618 51. Pease JE, Williams TJ. Eotaxin and asthma. Curr Opin Pharmacol 2001; 1: 248-
619 253. PMID: 11712747
620 52. Bandeira-Melo C, Phoofolo M, Weller PF. Extranuclear lipid bodies, elicited by
621 CCR3-mediated signaling pathways, are the sites of chemokine-enhanced
622 leukotriene C4 production in eosinophils and basophils. J Biol Chem 2001; 276:
623 22779-22787. PMID: 11274187
624 53. Dichlberger A, Kovanen PT, Schneider WJ. Mast cells: from lipid droplets to lipid
625 mediators. Clin Sci 2013; 125: 121-130. PMID: 23577635
626 54. Hutagalung AH, Novick PJ. Role of Rab GTPases in membrane traffic and cell
627 physiology. Physiol Rev 2011; 91: 119-149. PMID: 21248164
628 55. Yang L, Ding Y, Chen Y, Zhang S, Huo C, Wang Y, et al. The proteomics of lipid
629 droplets: structure, dynamics, and functions of the organelle conserved from
630 bacteria to humans. J Lipid Res 2012; 53: 1245-1253. PMID: 22534641
631 56. Murakami M, Sato H, Miki Y, Yamamoto K, Taketomi Y. A new era of secreted
632 Phospholipase A2 (sPLA2). J Lipid Res 2015; 56: 1248-1261. PMID: 25805806
633 57. Nolin JD, Ogden HL, Lai Y, Altemeier WA, Frevert CW, Bollinger JG, et al.
634 Identification of epithelial Phospholipase A2 Receptor 1 as a potential target in
635 asthma. Am J Respir Cell Mol Biol 2016; 55: 825-836. PMID: 27448109

636 58. Ackerman SJ, Park GY, Christman JW, Nyenhuis S, Berdyshev E, Natarajan V.
637 Polyunsaturated lysophosphatidic acid as a potential asthma biomarker. Biomark
638 Med 2016; 10: 123-135. PMID: 26808693
639 59. Hite RD, Seeds MC, Jacinto RB, Grier BL, Waite BM, Bass DA. Lysophospholipid
640 and fatty acid inhibition of pulmonary surfactant: non-enzymatic models of
641 phospholipase A2 surfactant hydrolysis. Biochim Biophys Acta 2005; 1720: 14-
642 21. PMID: 16376294
643 60. Pillai P, Fang C, Chan YC, Shamji MH, Harper C, Wu SY, et al. Allergen-specific
644 IgE is not detectable in the bronchial mucosa of nonatopic asthmatic patients. J
645 Allergy Clin Immunol 2014; 133: 1770-1772.e11. PMID: 11712747
646 61. Inselman LS, Chander A, Spitzer AR. Diminished lung compliance and elevated
647 surfactant lipids and proteins in nutritionally obese young rats. Lung 2004; 182:
648 101-117. PMID: 15136884
649 62. Foster DJ, Ravikumar P, Bellotto DJ, Unger RH, Hsia CCW. Fatty diabetic lung:
650 altered alveolar structure and surfactant protein expression. Am J Physiol Lung
651 Cell Mol Physiol 2010; 298: L392-L403. PMID: 20061442
652 63. Showalter MR, Nonnecke EB, Linderholm AL, Cajka T, Sa MR, Lönnerdal B, et
653 al. Obesogenic diets alter metabolism in mice. PLoS ONE 2018; 13: e0190632.
654 PMID: 29324762
655 64. Elghetany MT, Patel J. Assessment of CD24 expression on bone marrow
656 neutrophilic granulocytes: CD24 is a marker for the myelocytic stage of
657 development. Am J Hematol 2002; 71: 348-349. PMID: 12447971
658 65. Garrett-Sinha LA. Review of Ets1 structure, function, and roles in immunity. Cell
659 Mol Life Sci 2013; 70: 3375-3390. PMID: 23288305

66. 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

67. 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

68. 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

69. 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

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

71. 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

72. 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

682 **Table 1**

	Asthma severity	Severe		Mild-to-moderate	Healthy
	Smoking status	Negative	Positive	Negative	Negative
	U-BIOPRED cohort	**SAn** n=117	**SAc/ex** n=51	**MMA** n=43	**HC** n=41
Age (yr)	mean [range]	53 [21-78]	55 [29-74]	42 [18-72]	37 [18-65]
Gender (m/f)	ratio	43/74	19/32	24/19	29/12
Race (Caucasian/other)	ratio	109/8	49/2	40/3	37/4
Age at first diagnosis (yr)	mean [range]	24 [0-68]	35 [1-67]	20 [1-63]	NA
BMI (kg m-2)	mean [range]	29.1 [17.8-51.1]	30.1 [20.6-48.4]	25.5 [17.9-36.6]	25.5 [18.9-32.0]
Serum IgE (IU mL-1)	mean [range]	274 [5-2690]	398 [0-6811]	328 [7-3520]	89 [0-574]
FEV1 (% predicted)	mean [range]	65.4 [18.4-120.6]	66.1 [24.3-113.0]	90.4 [40.9-132.3]	102.1 [66.9-123.6]
FVC (% predicted)	mean [range]	87.8 [40.2-131.3]	90.6 [54.6-129.1]	107.4 [68.6-151.6]	108.5 [72.2-136.4]
FEV1/FVC ratio (%)	mean [range]	62.1 [31.0-92.2]	62.6 [35.2-90.0]	75.0 [52.4-93.9]	NA
Exacerbations (previous year)	mean [range]	2.1 [0-8]	2.5 [0-10]	0.4 [0-4]	NA
Pack-years (yr)	mean [range]	0.4 [0-5]	24.2 [5-70]	0.7 [0-5]	0.3 [0-5]
Intubation (ever)	count [pct]	13 [11%]	1 [2%]	0 [0%]	NA
ICU admission (ever)	count [pct]	28 [24%]	7 [14%]	0 [0%]	NA
Positive atopy test	count [pct]	73 [62%]	27 [53%]	36 [84%]	14 [34%]
ACQ1-5 score	mean [range]	2.2 [0.0-5.0]	2.2 [0.2-4.4]	0.8 [0.0-2.4]	NA
ACQ7 score	mean [range]	3.9 [0.0-7.0]	4.2 [0.0-7.0]	1.8 [0.0-7.0]	NA
AQLQ score	mean [range]	4.6 [1.9-6.8]	4.4 [2.3-6.8]	5.9 [3.2-6.9]	NA
Oral corticosteroid (current use)	count [pct]	49 [42%]	25 [49%]	0 [0%]	NA
Inhaled corticosteroid (current use)	count [pct]	113 [97%]	50 [98%]	43 [100%]	NA
Injectable corticosteroid (current use)	count [pct]	8 [7%]	0 [0%]	0 [0%]	NA
Long-acting ß-agonist (current use)	count [pct]	112 [96%]	48 [94%]	1 [2%]	NA
Short-acting ß-agonist (current use)	count [pct]	91 [78%]	37 [73%]	32 [74%]	NA
Corticosteroid dose (mg day-1)†	mean [range]	12.4 [5.0-37.5]	13.7 [2.5-40.0]	NA	NA

683

684 **Table 1**: Demographics of study participants according to the U-BIOPRED cohorts.[8]
685 Ex-smokers with a pack-year smoking history of ≤5 were considered to have a
686 'negative' smoking status, whereas ex-smokers with pack-year ≥5 were included in
687 the SAc/ex cohort if diagnosed with severe asthma, but excluded from the study
688 otherwise. Abbreviations: BMI = body mass index; FEV_1 = forced expiratory volume in
689 1 second; FVC = forced vital capacity; IgE = Immunoglobulin E; ICU = intensive care
690 unit; ACQ = Asthma Control Questionnaire; AQLQ = Asthma Quality of Life
691 Questionnaire; NA = not applicable/not assessed. † Systemic dosage of
692 corticosteroids for severe asthmatic participants is expressed in prednisolone-
693 equivalent doses.

Figure 1

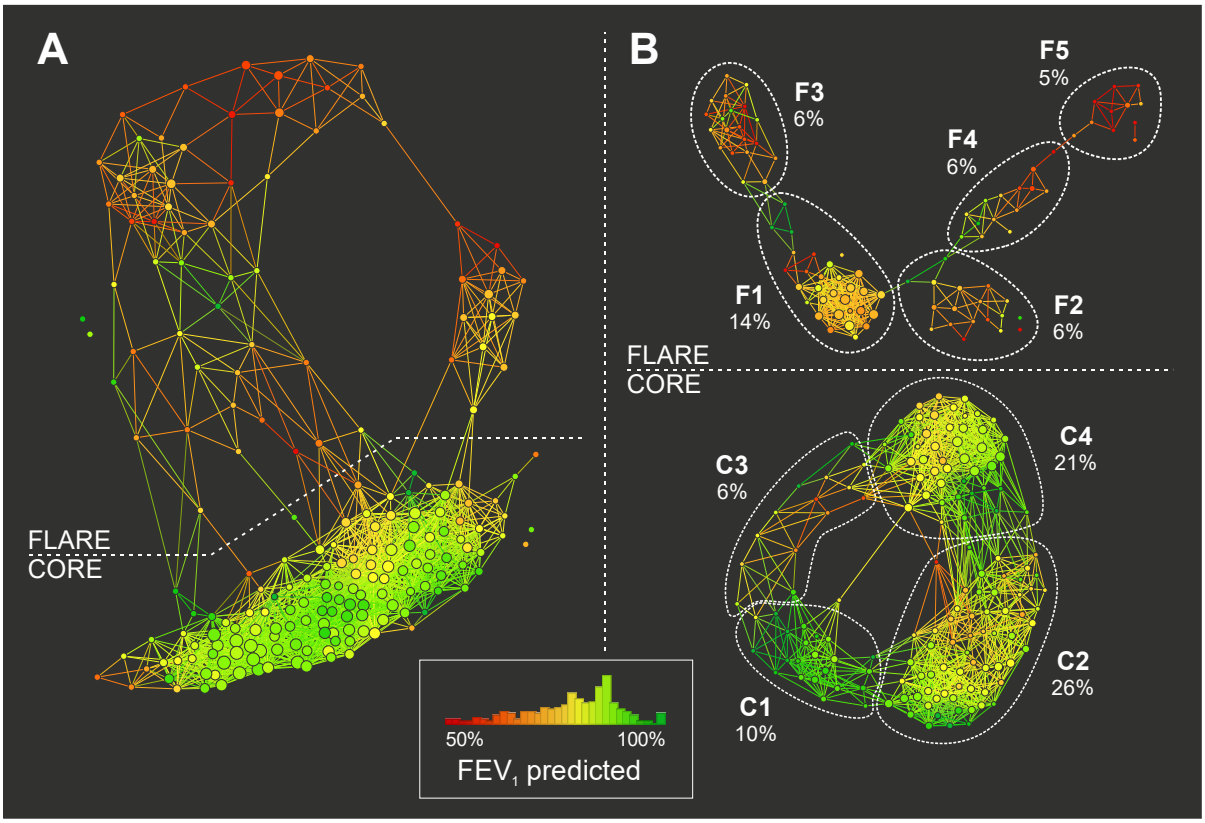

Figure 1: TDA structures of (A) the complete study cohort (n=252) and (B) the 'core' and 'flare' subgroups side by side (n=164 and n=107 respectively), coloured by FEV_1 (forced expiratory volume in 1 second). TDA was performed on 291 sputum lipid ions, using a normalised correlation metric and two MDS 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.

Figure 2

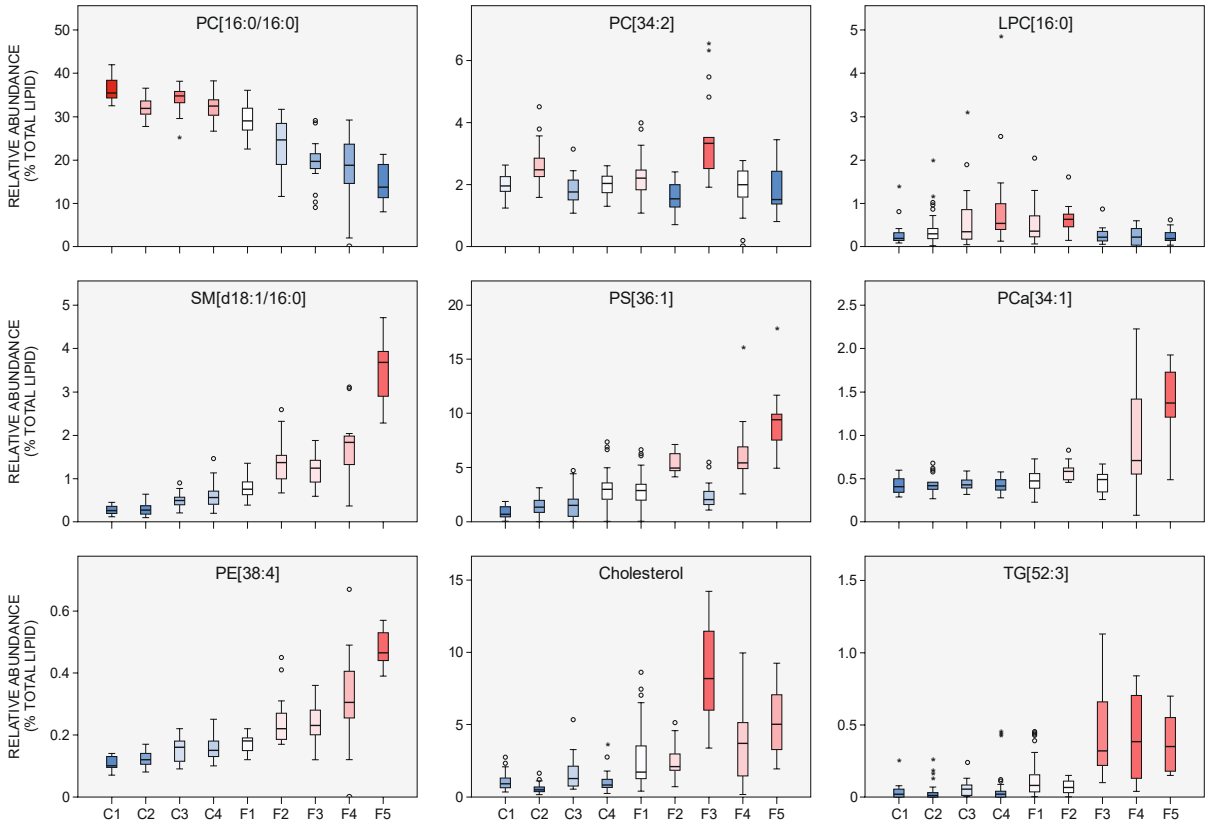

Figure 2: Box plots of representative lipid species demonstrating the trends across the TDA structure. Relative abundances are given as a percentage of the total lipid. The original boxplots were created in SPSS Statistics 24 which defines outliers as 'near' (open circles: more than 1.5 times the interquartile range) and 'far' (stars: more than 3 times the interquartile range).

Figure 3

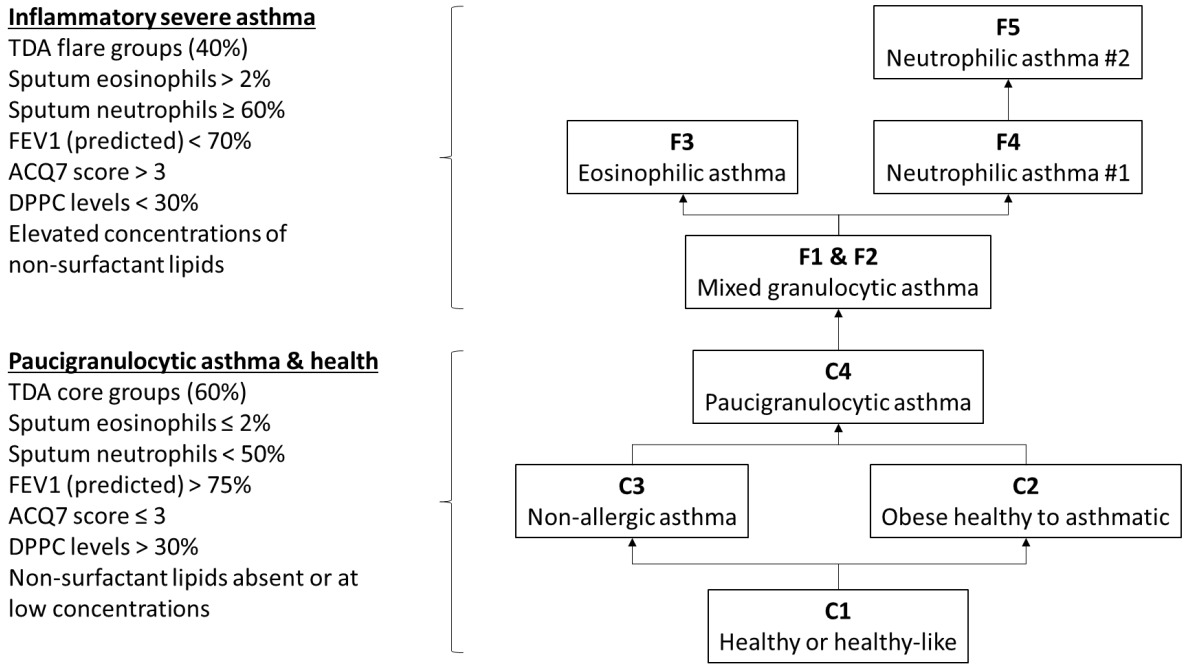

Figure 3: Summary of the sputum lipid phenotypes found in this study and their assignments based on associations with the demographic, clinical and pathobiological data (predominately sputum differential cell counts). As shown in Fig. 1 and discussed in the main text, the nine phenotypes represent a spectrum of asthma severity from C1 (healthy) to F3 and F5 (severe eosinophilic or neutrophilic). Key characteristics of the main 'core' and 'flare' groups are listed on the left.

Figure 4

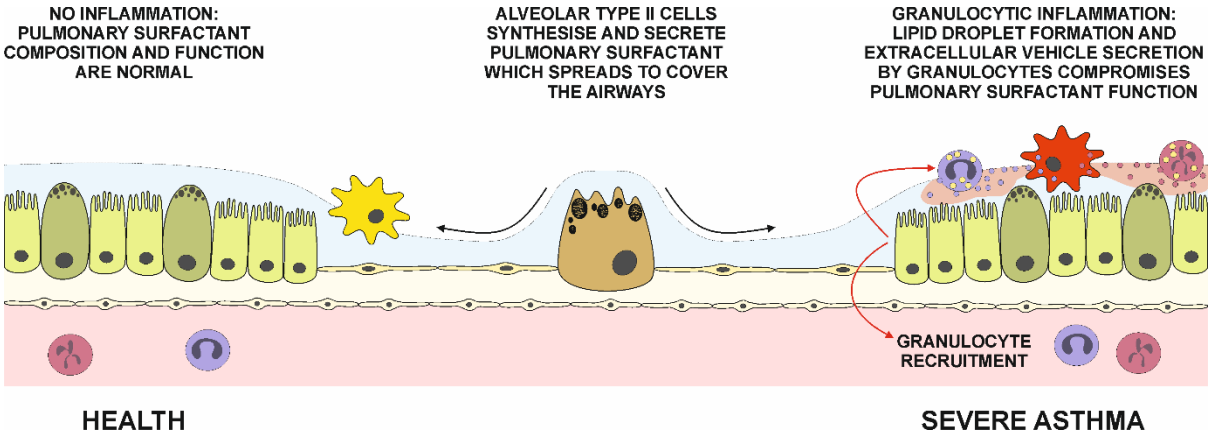

Figure 4: Conceptual representation of the role of granulocytic inflammation in producing EVs (exosomes and/or micro-vesicles) that may alter the lipid composition of the epithelial lining fluid (ELF; shown in light blue) in asthma. Pro-inflammatory chemokines and cytokines (red arrows) recruit eosinophils and neutrophils into the airways and stimulate intracellular lipid droplet formation, as well as the secretion of EVs. The latter are rich in cellular lipids and proteins which 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.

Online supplement

Methods

U-BIOPRED study design and tranSMART data repository

All samples used in this study were obtained from the U-BIOPRED cohort recruited in 14 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

753 TECAN Freedom EVO100 robotic liquid handling platform (Tecan, Männedorf,
754 Switzerland). Untargeted 'shotgun' mass spectra were acquired by flow injection
755 analysis on a Dionex 3000 ultra-high performance liquid chromatography system
756 (Thermo Scientific Dionex, Sunnyvale, CA, USA), coupled to a MaXis 3G quadrupole
757 time-of flight mass spectrometer equipped with an electrospray ionisation source
758 (Bruker Daltonics, Billerica, MA, USA). Measurements were done in full scan mode
759 over an *m/z* range of 350-1200 with separate injections for positive and negative
760 ionisation. Blank injections were performed after every four samples (no significant
761 carry-over was detected) and a pooled QC sample was run after every four samples
762 to check for changes in instrument performance. Fragmentation analysis for lipid
763 identification was performed on the pooled QC using the same instrumental setup, but
764 in LC-MS/MS mode using a Waters Acquity C8 column (1.7µm, 2.1mm x 100mm;
765 Waters, Milford, MA, USA) and a 50 min gradient of methanol and water (both with 50
766 mM NH_4HCO_2 and 0.2% formic acid). Data-independent product ion scans were
767 acquired over the entire gradient using broadband collision induced dissociation.
768 Precursor and fragment ions were matched retrospectively by their LC retention time
769 and using the well-established fragmentation rules for lipids (Hsu and Turk 2003) to
770 provide confirmation of identities. Lipid nomenclature followed the framework set out
771 by Liebisch et al. (2013) where sufficient structural information was available. The
772 following abbreviations for lipid classes were used in this study: phosphatidylcholine
773 (PC); phosphatidylglycerol (PG); phosphatidylserine (PS); phosphatidylinositol (PI);
774 phosphatidylethanolamine (PE); lyso-phosphatidylcholine (LPC); ceramide (Cer);
775 hexosyl-ceramide (HexCer); sphingomyelin (SM); cholesterol (Chol); cholesterol ester
776 (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 (Strazzeri 2018). By its very essence, TDA captures the continuous nature of data (Chazal and Michel 2021), 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.

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 (JT-test) for ordered alternatives was selected to identify trends in the ordinal and continuous variables, assuming an ascending hypothesis order. The significance of individual between-group differences was assessed pairwise by Mann-Whitney U tests. For the limited number of nominal variables, a Pearson's Chi-squared test was used instead. Trends or differences were considered significant if $p<0.05$, and highly significant if $p<0.001$ (after Bonferroni correction).

Blood protein data

Protein expression data from peripheral blood were acquired as a routine analysis for all U-BIOPRED participants, and obtained for this study from the U-BIOPRED data repository. 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: 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 (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

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

Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. Can J Biochem Physiol 1959; 37: 911-917. PMID: 13671378

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

Chazal F, Michel B. An Introduction to Topological Data Analysis: Fundamental and Practical Aspects for Data Scientists. Front Artif Intell; 4: 667963. PMID: 34661095

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

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

Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007; 8: 118-127. PMID: 16632515

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

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

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

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

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

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

872 Strazzeri F. A Morse-theoretical clustering algorithm for annotated networks and
873 spectral bounds for fuzzy clustering. 2018: PhD Thesis, University of Southampton,
874 pp. 133.

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

Supplementary Table S1

Lipid†	P-value	Z-score	C1	C2	C3	C4	F1	F2	F3	F4	F5
			10%	26%	6%	21%	14%	6%	6%	6%	5%
SM[d18:1/24:0]	0.000	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]	0.000	16.341	0.09	0.08	0.17	0.16	0.28	0.44	0.51	0.66	1.15
PC[30:1]	0.000	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]	0.000	16.248	0.25	0.26	0.49	0.56	0.74	1.36	1.23	1.83	3.67
PE[40:4]	0.000	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]	0.000	14.519	0.00	0.00	0.06	0.10	0.11	0.28	0.18	0.22	0.33
PE[38:4]	0.000	13.239	0.10	0.12	0.16	0.15	0.18	0.22	0.23	0.31	0.47
PC[36:0]	0.000	13.211	0.18	0.17	0.28	0.41	0.40	0.83	0.33	0.83	1.49
PE[40:7]	0.000	13.211	0.18	0.17	0.28	0.41	0.40	0.83	0.33	0.83	1.49
CE[18:2]	0.000	12.377	0.06	0.04	0.16	0.07	0.30	0.28	2.02	0.85	0.66
PCa[40:2]	0.000	12.269	0.02	0.02	0.05	0.04	0.06	0.14	0.07	0.12	0.24
CE[18:1]	0.000	12.218	0.04	0.02	0.10	0.04	0.16	0.28	0.68	0.35	0.36
Cholesterol	0.000	11.943	0.91	0.47	1.30	0.83	1.70	2.09	8.17	3.70	5.01
PC[38:4]	0.000	11.501	0.13	0.15	0.15	0.16	0.19	0.22	0.51	0.36	0.39
PCa[36:0]	0.000	11.281	0.09	0.09	0.11	0.13	0.14	0.26	0.13	0.20	0.27
PS[36:1]	0.000	11.144	0.69	1.34	1.52	3.00	2.88	4.95	2.04	5.42	9.40
PC[38:0]	0.000	10.951	0.01	0.02	0.02	0.03	0.03	0.06	0.05	0.07	0.09
PCa[40:1]	0.000	10.838	0.00	0.01	0.00	0.02	0.02	0.07	0.03	0.07	0.15
PE[32:1]	0.000	10.825	0.00	0.00	0.00	0.01	0.05	0.11	0.33	0.12	0.14
PE[40:6]	0.000	10.346	0.10	0.13	0.12	0.16	0.16	0.25	0.15	0.24	0.33
TG[52:3]	0.000	10.339	0.02	0.01	0.06	0.02	0.08	0.07	0.32	0.39	0.35
TG[52:2]	0.000	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]	0.000	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]	0.000	9.949	0.03	0.01	0.05	0.04	0.05	0.17	0.06	0.08	0.15
PCa[32:2]	0.000	9.891	0.04	0.03	0.04	0.04	0.05	0.09	0.05	0.08	0.10
PE[34:2]	0.000	9.891	0.04	0.03	0.04	0.04	0.05	0.09	0.05	0.08	0.10
PE[40:5]	0.000	9.828	0.09	0.11	0.09	0.13	0.14	0.14	0.19	0.22	0.35
PCa[36:1]	0.000	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]	0.000	9.622	0.02	0.02	0.03	0.03	0.03	0.06	0.04	0.06	0.07
TG[50:0]	0.000	9.488	0.00	0.00	0.02	0.01	0.03	0.03	0.04	0.08	0.07
PCa[42:2]	0.000	9.380	0.00	0.00	0.00	0.01	0.03	0.04	0.03	0.11	0.18
PC[38:5]	0.000	9.379	0.13	0.18	0.15	0.16	0.19	0.21	0.38	0.29	0.29
TG[54:3]	0.000	9.260	0.00	0.01	0.05	0.02	0.05	0.06	0.15	0.21	0.22
TG[54:5]	0.000	9.152	0.02	0.01	0.03	0.01	0.04	0.04	0.09	0.12	0.15
TG[54:4]	0.000	9.101	0.02	0.01	0.04	0.01	0.04	0.05	0.14	0.16	0.20
TG[52:4]	0.000	9.060	0.00	0.00	0.01	0.01	0.04	0.02	0.14	0.16	0.13
TG[52:1]	0.000	9.015	0.01	0.01	0.03	0.01	0.04	0.04	0.10	0.23	0.15
CE[16:1]	0.000	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]	0.000	8.552	0.00	0.01	0.00	0.02	0.02	0.00	0.07	0.08	0.11
TG[54:2]	0.000	8.441	0.02	0.02	0.03	0.03	0.03	0.05	0.04	0.07	0.13
TG[50:1]	0.000	8.310	0.01	0.01	0.04	0.01	0.05	0.06	0.13	0.35	0.20
PC[38:1]	0.000	8.196	0.03	0.02	0.05	0.06	0.05	0.11	0.05	0.07	0.13
TG[50:2]	0.000	7.917	0.00	0.00	0.03	0.00	0.04	0.02	0.16	0.22	0.17
PCa[32:1]	0.000	7.891	0.16	0.17	0.17	0.19	0.20	0.28	0.17	0.29	0.41
PE[34:1]	0.000	7.891	0.16	0.17	0.17	0.19	0.20	0.28	0.17	0.29	0.41
PCa[34:1]	0.000	7.244	0.40	0.41	0.42	0.41	0.47	0.58	0.48	0.70	1.37
PE[36:1]	0.000	7.244	0.40	0.41	0.42	0.41	0.47	0.58	0.48	0.70	1.37
PCa[36:2]	0.000	7.022	0.07	0.08	0.04	0.06	0.10	0.11	0.14	0.17	0.72
PC[36:1]	0.000	7.014	0.44	0.49	0.45	0.44	0.57	0.46	0.67	0.75	1.37
PCa[34:2]	0.000	6.994	0.12	0.13	0.09	0.12	0.15	0.22	0.17	0.22	0.60
PE[36:2]	0.000	6.994	0.12	0.13	0.09	0.12	0.15	0.22	0.17	0.22	0.60
PC[36:5]	0.000	6.598	0.16	0.22	0.18	0.22	0.24	0.23	0.38	0.25	0.31
LPC[18:0]	0.000	6.269	0.04	0.04	0.06	0.09	0.08	0.08	0.10	0.14	0.10
PC[38:2]	0.000	5.492	0.02	0.04	0.04	0.11	0.08	0.16	0.00	0.04	0.19
Cer[d18:1/18:0]	0.000	4.880	0.01	0.00	0.02	0.01	0.02	0.09	0.01	0.00	0.06
PC[40:6]	0.000	4.875	0.04	0.03	0.06	0.00	0.05	0.00	0.14	0.09	0.12
PE[38:6]	0.000	4.863	0.02	0.12	0.05	0.43	0.18	0.31	0.16	0.11	0.42
PC[36:2]	0.000	4.814	0.74	0.99	0.72	0.75	1.04	0.71	1.72	1.44	1.63
PE[38:5]	0.000	4.163	0.11	0.15	0.10	0.16	0.15	0.18	0.16	0.20	0.22
PC[44:11]	0.000	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
PCa[30:0]	0.611	-0.509	0.06	0.05	0.07	0.06	0.05	0.00	0.00	0.07	0.10
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	0.01	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]	0.000	-5.898	1.90	1.94	1.34	1.81	1.70	1.65	1.29	1.26	0.01
PC[34:1]	0.000	-6.205	5.43	6.64	5.19	5.45	5.62	4.49	5.13	5.36	4.99
PG[36:2]	0.000	-7.237	1.44	1.82	1.09	1.40	1.35	0.66	0.01	0.87	0.01
PC[34:3]	0.000	-8.571	1.99	2.52	2.06	2.68	2.05	1.73	1.38	1.18	0.88
PC[34:0]	0.000	-8.822	1.48	1.55	1.40	1.28	1.36	1.08	0.97	1.22	1.05
PC[32:1]	0.000	-11.653	4.17	4.91	3.33	3.54	3.13	2.50	2.09	2.10	1.46
PC[30:0]	0.000	-12.117	4.26	4.08	3.68	3.44	2.96	2.54	2.21	2.13	1.32
PC[32:0]	0.000	-12.463	35.52	31.87	34.78	32.43	29.04	24.71	19.68	18.76	13.72

878 Heat map of lipidomics trends across the TDA structure, sorted by z-score of a JT-test
879 for ranked alternatives. Values in the table show the median relative abundance of a
880 lipid for each TDA group. Note that unidentified ions were not included in this table. †
881 See main text for lipid nomenclature and the abbreviations used.

Supplementary Table S2

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%
Asthma severity*	0.000	6.366	0.58	0.76	0.71	0.88	0.93	0.94	1.00	0.94	1.00
Age (yr)	0.000	3.994	45	48	48	49	55	54	58	57	56
Gender ratio (m/f)	0.488	-0.734	0.7	0.6	1.3	1.2	0.9	1.1	1.1	0.8	0.4
BMI (kg m-2)	0.111	-1.592	25	30	27	25	27	27	26	28	24
Smoking status (% current/ex)	0.752	-0.316	0	11	4	14	14	7	0	6	0
Age of onset or first diagnosis (yr)	0.194	1.299	24	17	24	22	21	16	43	35	14
Positive atopy test (%)	0.512	-0.656	37	38	26	44	43	33	53	37	36
Serum IgE (IU mL-1)	0.160	1.405	86	89	68	109	93	91	120	91	129
ACQ7 score	0.000	5.026	1.5	2.5	2.5	3.0	3.2	3.8	4.2	3.5	4.5
ICU admissions (previous year)	0.006	2.748	0.1	0.1	0.0	0.2	0.1	0.1	0.1	0.3	0.6
Exacerbations (previous year)	0.000	4.445	0.8	1.1	1.2	1.7	2.1	1.7	1.9	2.2	2.9
Oral corticosteroids dose (mg day-1)†	0.000	4.247	2.6	2.1	1.9	4.1	5.4	6.7	6.0	5.9	3.4
Inhaled corticosteroids dose (mg day-1)†	0.000	5.917	0.37	0.45	0.58	0.62	0.80	0.81	0.82	0.81	1.00
SNOT score	0.007	2.685	21.8	23.6	22.7	25.8	29.9	28.9	26.2	33.4	29.3
HADS score	0.850	0.189	8.8	10.6	12.1	11.3	13.8	10.8	9.6	9.5	9.6
Exhaled nitric oxide (ppb)	0.003	2.970	22.5	20.0	26.5	27.0	32.0	23.8	54.0	23.0	19.0
FEV1 (% predicted)	0.000	-5.495	90.8	79.9	82.4	79.0	69.3	65.5	61.9	59.6	42.2
FVC (% predicted)	0.000	-3.643	107.4	98.4	98.3	95.8	96.9	88.8	92.6	89.7	74.7
FEV1/FVC ratio (predicted)	0.000	-3.510	79.3	78.9	79.4	79.3	77.6	77.9	77.1	77.9	77.9
FEF 25-75 (predicted)	0.000	-5.699	3.71	3.56	3.64	3.76	3.34	3.35	3.29	3.36	3.29
sGaw (1/kPA x sec)	0.000	-4.162	1.18	1.29	1.06	0.86	0.77	0.84	0.72	0.72	0.25
Sputum eosinophils (%)	0.000	5.857	0.20	0.71	0.20	2.00	4.39	1.37	26.79	3.49	2.90
Sputum neutrophils (%)	0.000	4.907	40.47	45.65	43.90	43.20	59.54	60.78	59.58	83.01	91.40
Sputum macrophages (%)	0.000	-8.764	57.64	48.95	54.76	41.84	25.74	20.30	12.14	9.08	4.84
Sputum lymphocytes (%)	0.000	-3.707	1.38	1.14	1.27	1.24	0.94	0.76	0.68	0.59	0.39
Sputum squamous epithelial cells (%)	0.017	-2.376	9.50	5.25	13.25	6.70	9.90	6.90	3.55	7.80	0.90
Blood neutrophils (µL-1)	0.000	3.849	3200	4000	3500	3900	4750	5700	4700	5300	5550
Blood eosinophils (µL-1)	0.006	2.756	100	200	200	200	200	200	400	250	250
Blood leukocytes (µL-1)	0.000	4.306	6000	6900	6050	6900	7800	9450	7900	9350	9100

Heat map of trends in the demographic, clinical and pathobiological data across the sputum lipidomics TDA structure, with z-scores of a JT-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; FEV_1 = 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 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 S3

Protein biomarker	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	f	0.000	4.031	2.79	3.13	2.73	3.01	3.03	3.73	3.24	3.87	3.97
Serpin-E1	a	0.000	3.875	77469	85894	86691	90934	91605	101345	96137	111359	91426
MCP-4	a	0.000	3.740	98.3	127.2	111.9	128.2	131.7	145.1	146.0	129.9	159.6
Periostin	d	0.001	3.328	48.3	45.0	44.3	44.5	47.9	53.7	61.3	50.4	53.9
Eotaxin	f	0.001	3.326	87.8	93.9	92.7	106.0	106.5	138.0	104.5	113.0	135.0
CCL18	b	0.001	3.185	134.9	161.3	164.1	191.4	192.1	186.0	221.3	179.8	276.6
Galectin-3	a	0.003	3.011	5288	5486	5182	5298	5722	6256	5112	6232	6740
IL-6	f	0.004	2.906	0.52	0.91	0.61	0.73	0.60	0.89	1.07	1.37	1.23
IL-17AA	c	0.005	2.806	0.29	0.30	0.36	0.37	0.35	0.38	0.46	0.53	0.36
CCL17	f	0.006	2.754	55.7	69.3	51.6	63.2	65.6	129.0	54.7	117.0	73.8
IL-13	b	0.006	2.726	0.49	0.49	0.54	0.62	0.60	0.47	1.32	0.53	0.48
MCP-1	f	0.008	2.641	93.6	98.7	102.0	98.2	95.7	125.0	110.5	109.0	112.5
hs-CRP	e	0.010	2.577	1.20	1.40	1.20	1.60	1.15	1.40	1.30	6.20	2.80
TNF-alpha	f	0.016	2.400	1.56	1.86	1.68	1.89	1.71	1.83	1.96	1.85	2.12
C5a	a	0.020	2.335	33.7	38.2	48.9	36.6	39.8	39.0	35.6	54.4	49.5
MMP-3	a	0.025	2.245	13817	13569	13730	14477	17158	23242	20967	12788	15431
MIP-1b	f	0.035	2.114	46.2	49.4	44.1	45.7	45.2	67.8	52.1	61.8	54.6
Eotaxin-3	f	0.058	1.894	17.8	15.4	15.5	19.1	18.4	18.0	26.9	14.2	15.4
IFN-gamma	f	0.067	1.834	4.38	5.77	5.40	4.67	4.80	5.73	7.41	7.43	11.02
CD40L	a	0.104	1.624	4717	4287	4750	4698	4910	5275	5546	4725	4253
IL-6R-alpha	a	0.127	1.525	10398	10927	11221	11227	11696	12783	11108	10787	10884
IL-1-alpha	a	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	a	0.257	1.134	6355	6551	7592	7056	6935	7610	6889	6971	7392
IL-18	a	0.267	1.110	192.4	219.0	214.4	222.0	210.7	209.1	218.1	244.7	221.9
IP-10	f	0.347	0.940	223.5	326.0	297.0	278.0	250.0	234.0	304.5	353.0	473.0
CD30	a	0.377	0.883	31.7	38.0	38.1	38.5	35.7	37.1	36.3	40.7	40.5
LBP	a	0.466	0.729	2056309	2166450	2160387	2019250	1999580	2623438	1722485	2770901	2508327
Lumican	a	0.502	0.672	130969	133245	132268	133135	128644	141913	148609	125866	141143
RAGE	a	0.809	0.242	1358	1298	1382	1275	1250	1265	1151	1372	1585
SHBG	a	0.512	-0.655	4637586	2777497	1693569	2969730	2709649	1989644	2396730	2470851	3401339
CCL22	f	0.502	-0.672	868.5	889.5	810.0	794.0	699.5	858.0	828.5	880.0	893.0
DPPIV	a	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 z-score of a JT-test for ranked alternatives. Protein concentrations were measured in serum by either [a] Luminex, [b] Impact, [c] Singulex, [d] Elecsys or [e] Immulite immunoassays, and in plasma by [f] MSD assay. Values in the table show the median concentration of a given protein for each TDA group.

Supplementary Table S4

Upstream regulator	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
T	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	-1.98	-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 z-score of a JT-test for ranked alternatives.