EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Original research article

Plasma proteins elevated in severe asthma despite oral steroid use and unrelated to Type-2 inflammation

Maria Sparreman Mikus, Johan Kolmert, Lars I. Andersson, Jörgen Östling, Richard G. Knowles, Cristina Gómez, Magnus Ericsson, John-Olof Thörngren, Payam Emami Khoonsari, Barbro Dahlén, Maciej Kupczyk, Bertrand De Meulder, Charles Auffray, Per S. Bakke, Bianca Beghe, Elisabeth H. Bel, Massimo Caruso, Pascal Chanez, Bo Chawes, Stephen J. Fowler, Mina Gaga, Thomas Geiser, Mark Gjomarkaj, Ildikó Horváth, Peter H. Howarth, Sebastian L. Johnston, Guy Joos, Norbert Krug, Paolo Montuschi, Jacek Musial, Ewa Niżankowska-Mogilnicka, Henric K. Olsson, Alberto Papi, Klaus F. Rabe, Thomas Sandström, Dominick E. Shaw, Nikolaos M. Siafakas, Mathias Uhlen, John H. Riley, Stewart Bates, Roelinde J.M. Middelveld, Craig E. Wheelock, Kian Fan Chung, Ian M. Adcock, Peter J. Sterk, Ratko Djukanovic, Peter Nilsson, Sven-Erik Dahlén, Anna James,on behalf of the U-BIOPRED (Unbiased Biomarkers for the Prediction of Respiratory Disease outcome) Study Group, the BIOAIR (Longitudinal Assessment of Clinical Course and Biomarkers in Severe Chronic Airway Disease) Consortium

Please cite this article as: Mikus MS, Kolmert J, Andersson LI, *et al*. Plasma proteins elevated in severe asthma despite oral steroid use and unrelated to Type-2 inflammation. *Eur Respir J* 2021; in press (https://doi.org/10.1183/13993003.00142-2021).

Plasma proteins elevated in severe asthma despite oral steroid use and unrelated to Type-2 inflammation

Maria Sparreman Mikus[1,2*], Johan Kolmert[2,3], Lars I. Andersson[2,3], Jörgen Östling[4], Richard G. Knowles[5], Cristina Gómez[2,3], Magnus Ericsson[6], John-Olof Thörngren[6], Payam Emami Khoonsari[7], Barbro Dahlén[2,3,8], Maciej Kupczyk[2,3,9], Bertrand De Meulder[10], Charles Auffray[10], Per S. Bakke[11], Bianca Beghe[12], Elisabeth H. Bel[13], Massimo Caruso[14], Pascal Chanez[15], Bo Chawes[16], Stephen J. Fowler[17], Mina Gaga[18], Thomas Geiser[19], Mark Gjomarkaj[20], Ildikó Horváth[21], Peter H. Howarth[22], Sebastian L. Johnston[23], Guy Joos[24], Norbert Krug[25], Paolo Montuschi[26], Jacek Musial[27], Ewa Niżankowska-Mogilnicka[27], Henric K. Olsson[28], Alberto Papi[29], Klaus F. Rabe[30], Thomas Sandström[31], Dominick E. Shaw[32], Nikolaos M. Siafakas[33], Mathias Uhlen[1,34], John H. Riley[35], Stewart Bates[35], Roelinde J.M. Middelveld[2,3], Craig E. Wheelock[3,36], Kian Fan Chung[23], Ian M. Adcock[23], Peter J. Sterk[13], Ratko Djukanovic[22], Peter Nilsson[1], Sven-Erik Dahlén[2,3], and Anna James[2,3]; on behalf of the U-BIOPRED (Unbiased Biomarkers for the Prediction of Respiratory Disease outcome) Study Group[37] and the BIOAIR (Longitudinal Assessment of Clinical Course and Biomarkers in Severe Chronic Airway Disease) Consortium[38]

Affiliations:

[1] Department of Protein Science, KTH Royal Institute of Technology, SciLifeLab, Stockholm, Sweden

[2] Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

[3] Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden

[4] PEXA AB, Gothenburg, Sweden

[5] Knowles Consulting, Stevenage Bioscience Catalyst, Stevenage, UK

[6] Department of Laboratory Medicine, Karolinska University Hospital Huddinge, Stockholm, Sweden

[7] Department of Biochemistry and Biophysics, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Solna, Sweden

[8] Department of Medicine, Karolinska University Hospital Huddinge, Stockholm, Sweden

[9] Department of Internal Medicine, Asthma and Allergy, Medical University of Lodz, University of Lodz, Lodz, Poland

[10] European Institute for Systems Biology and Medicine, Lyon, France

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

[12] Department of Medical and Surgical Sciences, University of Modena and Reggio Emilia, Modena, Italy

[13] Department of Respiratory Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

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

[15] Assistance Publique des Hôpitaux de Marseille, Clinique des Bronches, Allergies et Sommeil, Aix Marseille Université, Marseille, France

[16] COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark

[17] Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, The University of Manchester; Manchester Academic Health Science

Centre and NIHR Manchester Biomedical Research Centre, Manchester University Hospitals NHS Foundation Trust, Manchester, UK

[18] Respiratory Medicine Dept and Asthma Centre, Athens Chest Hospital “Sotiria”, University of Athens, Athens, Greece

[19] Department for Pulmonary Medicine, University Hospital and University of Bern, Bern, Switzerland

[20] Institute for Research and Biomedical Innovation, Italian National Research Council, Palermo, Italy

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

[22] NIHR Southampton Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust, and Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK

[23] National Heart and Lung Institute, Imperial College London, London, UK

[24] Department of Internal Medicine and Pediatrics, Ghent University, and Department of Respiratory Medicine, Ghent University Hospital, Ghent, Belgium

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

[26] Department of Pharmacology, Università Cattolica del Sacro Cuore, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy

[27] Department of Internal Medicine, Jagiellonian University Medical College, Krakow, Poland

[28] Translational Science and Experimental Medicine, Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

[29] Division of Internal and Cardiorespiratory Medicine, University of Ferrara, Ferrara, Italy

[30] Department of Internal Medicine, Christian Albrechts University Kiel, Kiel, Germany

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

[32] Respiratory Research Unit, University of Nottingham, Nottingham, UK

[33] Department of Thoracic Medicine, Medical School, University of Crete, Heraklion, Crete, Greece

[34] Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden

[35] Respiratory Therapeutic Unit, GlaxoSmithKline, London, UK

[36] Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden

[37,38] The collaborating members of the U-BIOPRED and the BIOAIR study groups are listed in the Acknowledgements section.

*** Corresponding author:**

Maria Sparreman Mikus, PhD

Institute of Environmental Medicine, Karolinska Institutet, Biomedicum 5B, SE-171 77

Stockholm, Sweden

E-mail: maria.mikus@ki.se

Take home message: Application of new proteomic panel in two established European asthma cohorts identifies plasma proteins associated with disease severity independently of Type-2 inflammation, suggesting potentially useful novel biomarkers and therapeutic targets.

Author contributions:

Conception, design or clinical conduct of the study: RK, BD, MK, CA, PSB, BB, EHB, MC, PC, BC, SJF, MG, TG, MG, IH, PHH, SLJ, GJ, NK, PM, JM, ENM, HKO, AP, KFR, TS, DES, NMS, MU, JHR, SB, RJMM, CEW, KFC, IMA, PJS, RD, PN, and SED.

Acquisition, analysis, or interpretation of data: MSM, JK, LIA, JÖ, CG, ME, JOT, PEK, BDM, CEW, PN, SED and AJ.

Drafting the work: MSM, JK, LIA, PN, SED and AJ.

Revising it critically for important intellectual content and final approval of the manuscript version to be published: all authors.

Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved: MSM, PN, SED and AJ.

Support statement:

This study was supported by grants from the ChAMP (Centre for Allergy Research Highlights Asthma Markers of Phenotype) consortium which is funded by the Swedish Foundation for Strategic Research, the Karolinska Institutet, AstraZeneca & Science for Life Laboratory Joint Research Collaboration, and the Vårdal Foundation, the Swedish Heart-Lung Foundation, the Swedish MRC, the Stockholm County Council Research Funds (ALF), the Swedish Asthma and Allergy Association's Research Foundation, and Karolinska Institutet. This project has received funding from an Innovative Medicines Initiative (IMI) Joint Undertaking (JU) under grant agreement No 115010 (U-BIOPRED [Unbiased Biomarkers for the Prediction of Respiratory Diseases Outcomes]), resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. Additional support from IMI 2 JU under grant agreement No 831434 (3TR [Taxonomy, Treatment, Targets and Remission]). This JU receives the support from the European Union's Horizon 2020 research and innovation programme and EFPIA. The BIOAIR consortium was funded by the Fifth and Sixth Framework Programmes of the European Union, contract numbers: QLG1-CT-2000-01185 (BIOAIR) and FOOD-CT-2004-506378 (GA2LEN), and several national funding bodies. AJ and MK were supported by the Bernard Osher Initiative for Severe Asthma. PEK was financially supported by the Knut and Alice Wallenberg Foundation as part of the National Bioinformatics Infrastructure Sweden at SciLifeLab.

This article has supplementary material available from erj.ersjournals.com

ABSTRACT

Rationale: Asthma phenotyping requires novel biomarker discovery**.**

Objectives: To identify plasma biomarkers associated with asthma phenotypes by application of a new proteomic panel to samples from two well-characterized cohorts of severe (SA) and mild-to-moderate (MMA) asthmatics, chronic obstructive pulmonary disease (COPD) subjects and healthy controls (HC).

Methods: An antibody-based array targeting 177 proteins predominantly involved in pathways relevant to inflammation, lipid metabolism, signal transduction and extracellular matrix was applied to plasma from 525 asthmatics and HC in the U-BIOPRED cohort, and 142 subjects with asthma and COPD from the validation cohort BIOAIR. Effects of oral corticosteroids (OCS) were determined by a two-week, placebo-controlled OCS trial in BIOAIR, and confirmed by relation to objective OCS measures in U-BIOPRED.

Results: In U-BIOPRED, 110 proteins were significantly different, mostly elevated, in SA compared to MMA and HC. Ten proteins were elevated in SA versus MMA in both U-BIOPRED and BIOAIR (alpha-1-antichymotrypsin, apolipoprotein-E, complement component 9, complement factor I, macrophage inflammatory protein-3, interleukin-6, sphingomyelin phosphodiesterase 3, RANK, TGF-β1, and glutathione S-transferase). OCS treatment decreased most proteins, yet differences between SA and MMA remained following correction for OCS use. Consensus clustering of U-BIOPRED protein data yielded six clusters associated with asthma control, quality of life, blood neutrophils, hsCRP, and BMI, but not Type-2 inflammatory biomarkers. The mast cell specific enzyme carboxypeptidase A3 was one major contributor to cluster differentiation.

Conclusions: The plasma proteomic panel revealed previously unexplored yet potentially useful Type-2-independent biomarkers, and validated several proteins with established involvement in the pathophysiology of severe asthma.

Word count: 250 out of 250

Keywords:

Biomarkers, consensus clustering, non-Type-2 asthma, phenotypes, protein profiling, suspension bead arrays

INTRODUCTION

Asthma is a prevalent chronic inflammatory disease with many different phenotypes sharing common clinical manifestations of episodic breathlessness, wheezing, cough, airflow obstruction, usually reversible, and airway hyperresponsiveness (1, 2). The underpinning pathobiology may however be widely different. As individuals respond differently to treatments targeting specific pathways, better predictive biomarkers are required to improve patient selection, guide treatment, and monitor response. Currently available Type-2 asthma biomarkers, such as blood eosinophils, total serum IgE or fraction of exhaled nitric oxide (FeNO), do not adequately enable endotyping which is required for personalized treatment (1). Another major clinical challenge is the identification of biomarkers reflecting non-Type-2 asthma, which is often more severe and lacking effective treatments (3, 4).

With the overall aim being biomarker discovery, the ChAMP (**C**entre for Allergy Research **h**ighlights **A**sthma **M**arkers of **P**henotype) consortium developed an affinity proteomics panel focused on proteins with potential involvement in airway or systemic inflammation. Protein selection was based on literature reviews, database searches, and recent research findings. The four main biological processes reflected by the proteins were 1) immune response, 2) lipid mediator pathways (predominantly sphingolipids), 3) signal transduction and 4) extracellular matrix (**figure 1**). The panel was enriched in non-Type-2-related proteins to address the unmet clinical need in this particular subgroup. Specifically, our objective was to examine plasma protein associations with asthma severity and oral corticosteroid (OCS) treatment. Furthermore, we aimed to test the hypothesis that specific plasma protein profiles can identify unique molecular subgroups of asthma patients.

Initially, subjects with mild-to-moderate or severe asthma, and healthy controls from the multicentre investigation U-BIOPRED (Unbiased Biomarkers for the Prediction of Respiratory Disease outcome) were screened (5). The discoveries in U-BIOPRED were then validated in a second cohort, BIOAIR (Longitudinal Assessment of Clinical Course and BIOmarkers in Severe Chronic AIRway Disease) (6), where samples were obtained from subjects with mild-to-moderate or severe asthma or COPD, before and after a placebo-controlled intervention with oral prednisolone. This permitted comparisons between the protein profiles of subjects with severe asthma and COPD, distinct respiratory disorders that may show certain overlapping characteristics, and importantly, an assessment of the pharmacological effect of oral glucocorticosteroids on the plasma proteins measured. The study design and main aims are shown in **figure 1**.

METHODS

Subjects

The discovery cohort included baseline EDTA plasma samples from U-BIOPRED, a prospective cohort study of severe asthma phenotypes (Clin.Trial.Gov NCT01976767) (5). Of 525 included subjects, 263 were non-smokers with severe asthma (SAn), 95 current or ex-smokers with severe asthma (SAs/ex), 76 non-smokers with mild-to-moderate asthma (MMA), and 91 healthy non-smoking controls (HC) (**figure 1**). Subject characteristics and data availability are summarized in **tables 1** and **S1**.

The validation cohort included 142 subjects from the BIOAIR study (Clin.Trial.Gov NCT00555607) (6). After a four-week treatment optimization period, subjects underwent a two-week, double-blind placebo-controlled OCS intervention (0.5 mg/kg/day prednisolone) added to regular treatment, followed by an identical open OCS treatment of the placebo group only (**figure 1**). Heparin plasma samples from 58 SAn, 48 MMA, and 36 patients with COPD, before and after the intervention, were analysed. Subject characteristics and data availability are shown in **tables 1** and **S1**.

Further information regarding the study design of each cohort and diagnostic criteria for each subject group are presented in the online supplement. Both studies were approved by the ethics committees of each participating clinical institution and participants provided written informed consent.

Panel design and suspension bead array protein profiling

Detailed information regarding the protein panel of 177 proteins (**table S2**) and the antibody bead array methodology is provided in the online data supplement. Briefly, this in-house developed array-based affinity proteomics method utilizes antibodies coupled to magnetic color-coded beads (MagPlex, Luminex Corp., Austin, TX, USA) to create a multiplex analysis platform (7, 8). Measurements were performed using FlexMAP3D instruments (Luminex corp.) and reported as relative fluorescent intensity values.

Statistical analysis

Statistical analysis and visualization were performed in R (9, 10). Nonparametric Kruskal-Wallis (multiple groups), Wilcoxon rank-sum (pairwise group), and Wilcoxon signed-rank test (paired) tests were used for comparisons of continuous variables. All reported *P* values for the proteins were adjusted for multiple testing using the

Benjamini and Hochberg (11) method and controlling the false discovery rate (FDR) at 5%.

Unsupervised consensus cluster analysis was used to identify potential subgroups of asthma patients, a process described in more detail in the online supplement. Briefly, a reduced set of variables (*n*=139) was selected for consensus cluster analysis of log2 transformed and z-scored intensity signals for each antibody, performed using the 'ConsensusClusterPlus' package in R (12, 13). The Euclidean distance measure was used to describe similarity between subjects and the partitioning around the medoids algorithm was used for clustering. For model validation, clustering was repeated 1000 times, randomly removing 10% of subjects at each iteration. The cluster stability of models with two to ten clusters was evaluated by the lowest proportion of ambiguously clustered subjects (14) and the lowest deviation of ideal stability (15). A six-cluster model had slightly lower stability compared to ten (**table S8**), but demonstrated greater consistency across all clusters (**figure S2e**) and is therefore presented.

To identify proteins important for classification of the six cluster groups, the Boruta algorithm (R package 'Boruta') (16), a wrapper built around the Random forest classification method, was used. For more details, see the online supplement.

RESULTS

Screening for asthma-associated proteins in U-BIOPRED

Plasma profiling revealed 110 proteins (measured by 139 antibodies) that showed significantly different levels between subject groups in the U-BIOPRED cohort (**table**

S4). The top 21 proteins, all with *P* values lower than 10^{-10}, are shown in **table 2**. The most significant differences were consistently identified between SAn or SAs/ex, and MMA or HC (**figure 2a**). No proteins were different between the non-smoking and smoking severe asthma groups (SAn and SAs/ex), nor between MMA and HC (**figure 2a**). The overlap of proteins showing significant differences in the pairwise group comparisons are shown in **figure S1a**.

Validation of asthma severity-associated proteins in BIOAIR

In the validation cohort BIOAIR, between-group comparisons revealed significantly different plasma levels in 23 proteins (25 antibodies), see **table S5**. Of these, 15 proteins were altered between MMA and COPD (all elevated in COPD) and 13 between SAn and MMA (all elevated in SAn). No differences in protein levels were found between SAn and COPD. As shown in **figure S1b**, the majority of the significantly different proteins were unique to the SAn vs MMA, or the MMA vs COPD pairwise group comparison.

Comparing the results of U-BIOPRED and BIOAIR, ten proteins were confirmed to be significantly different between severe and mild-to-moderate asthma in both cohorts (**table 3**). Alpha-1-antichymotrypsin (SERPINA3), apolipoprotein E (APOE), complement component 9 (C9), macrophage inflammatory protein-3 (CCL23 [MIP-3]), complement factor I (CFI), interleukin-6 (IL-6), sphingomyelin phosphodiesterase 3 (SMPD3), TNF receptor superfamily member 11a (TNFRSF11A [RANK]), transforming growth factor beta 1 (TGFB1), and glutathione S-transferase P (GSTP1), were all elevated in SAn compared to MMA (**figure 3**).

Influence of oral corticosteroids on plasma protein levels

The highly significant differences in plasma protein levels observed between severe and mild-to-moderate asthma could be due to differences in disease mechanisms,

pharmacological treatments, or both. We therefore used the BIOAIR cohort, where subjects underwent a two-week controlled OCS trial with prednisolone, to examine the effect of glucocorticoid treatment on plasma protein levels. The majority of proteins were found to be affected by prednisolone, most of which were decreased (**figure 4a-b**). Among the ten proteins that were elevated in severe asthma in both BIOAIR and U-BIOPRED (**figure 3**), apolipoprotein E, complement component 9, macrophage inflammatory protein-3, and alpha-1-antichymotrypsin all showed a significant change following oral steroid treatment. Apolipoprotein E and alpha-1-antichymotrypsin were increased after OCS intake (**figure 4c**), whereas complement component 9 and macrophage inflammatory protein-3 were decreased.

Oral steroid-induced plasma protein changes found in BIOAIR were confirmed by combining prescription data and the objective measurement of urinary prednisolone metabolites in U-BIOPRED. Comparing levels between SAn reportedly taking and not taking OCS, as confirmed by urinary analysis, 20 proteins (21 antibodies) in U-BIOPRED were associated with oral steroid use (**figure 2c** and **table S6**). In line with BIOAIR results, apolipoprotein E and alpha-1-antichymotrypsin showed profiles that were both severity and steroid-dependent, increasing from MMA to SAn not taking OCS, and further from SAn not taking OCS to SAn taking OCS (**figure 4c**). No proteins were significantly different when comparing the smaller groups of smoking severe asthmatics (SAs/ex) that used OCS and those who did not (**figure 2c**).

To take into account oral steroid effects, the multiple group comparison of SAn, SAs/ex, MMA and HC was repeated, limited to subjects reporting no current OCS use and in whom no urinary prednisolone metabolites were detected. Levels of 98 proteins (126 antibodies) remained significantly different between U-BIOPRED

subject groups (**figure 2b**). Pairwise comparisons of these groups revealed a similar pattern as for the whole dataset, irrespective of OCS use (compare **figure 2a** with **figure 2b**). Again, there was no difference in levels between SAn and SAs/ex (data not shown). More importantly, the ten replicated proteins remained elevated in severe asthma (significant after FDR-adjustment: apolipoprotein E, complement component 9, macrophage inflammatory protein-3, complement factor I, alpha-1-antichymotrypsin, and RANK). A summary of all proteins that were significantly changed using these two approaches, including or excluding OCS users, is shown in **table S7**.

Identification of protein-driven subgroups of asthma

To identify subgroups or phenotypes of asthma, a protein-driven consensus clustering algorithm was applied to U-BIOPRED asthmatics. We clustered subjects based on their plasma protein profiles alone, restricted to the asthma-associated proteins identified in the univariate analysis (**table S4**). Evaluation of cluster models identified six robust clusters with high stability (**table S8**, **figure 5a**) and that were defined only by within-group similarities of included protein profiles. By applying a classification algorithm, it was possible to identify 108 profiles (85 unique proteins) confirmed as important for classification of the six clusters (**figure S3**). Although all proteins were relevant for classification, the three most contributing proteins are shown in **figure 5b**.

A pattern was observed among the clusters when sorted by median plasma protein levels. Most commonly (63% of 139 antibody profiles), the highest levels were found in cluster 1, followed in decreasing order by clusters 6, 4, 3, 2 and 5. Decreasing protein plasma levels across the clusters were also associated with signs of decreasing disease severity (**table S9**). Although subjects with SA and MMA were

found in all clusters, those with SA were relatively more prevalent in clusters 1, 4 and 6 (more than 90% per cluster) and subjects with MMA were relatively more prevalent in clusters 2, 3, and 5 (21%-37% of subjects per cluster) which contained 78% of all MMA subjects.

The clusters aggregated subjects with similar clinical outcomes (**table S9**). In cluster 1, subjects experienced low lung function, the most uncontrolled asthma, the lowest quality of life, and the most frequent exacerbations. These subjects had the highest BMI (84% overweight, 60% obese) and evidence of systemic inflammation reflected by the most elevated high-sensitivity CRP (hsCRP) and highest blood neutrophils. They also showed signs of increased mast cell activation as reflected by elevated urinary tetranor-PGDM. Conversely, cluster 5 included the youngest subjects with the best lung function, best asthma control, earliest onset of asthma and lowest BMI. These subjects also exhibited the lowest blood neutrophil numbers, lowest hsCRP and near-normal urinary tetranor-PGDM concentration. A selection of clinical and biochemical variables associated with the clusters are shown in **figure 5b**.

Parameters that were not different among the clusters included gender distribution and objective evidence of OCS use. Indices of Type-2 inflammation, including blood and sputum eosinophils, FeNO, circulating periostin, total serum IgE and prevalence of atopy, were also not significantly different (**table S9**).

Proteins associated with Type-2-high asthma

To examine whether proteins in the array were altered in subjects with Type-2-high compared to Type-2-low inflammation, we examined the U-BIOPRED data in relation to two recently published strategies for defining Type-2 asthma. Dividing asthma subjects according to a composite Type-2 biomarker score based on blood

eosinophils, FeNO, and serum periostin as described by the Refractory Asthma Stratification Programme (RASP) (17), was not associated with any significant differences in protein levels. In U-BIOPRED, we have recently published that high urinary LTE_4 levels have a strong Type-2 association (18). In the current investigation, only 7 of the proteins in the panel were associated with LTE_4 levels (**table S10**). Taken together, this analysis confirms that the protein panel is suitable for its purpose to identify Type-2-independent signals.

DISCUSSION

Using a high-throughput, array-based protein profiling technique to simultaneously analyse 177 selected plasma proteins in 667 subjects from two established European asthma cohorts, we identified both known and previously unknown candidate proteins associated with asthma severity. Despite observed effects of oral corticosteroids on multiple proteins, associations with asthma severity remained for most proteins following adjustment for OCS use, a confounding factor often overlooked in biomarker discovery efforts. Furthermore, the patterns of protein profiles grouped asthmatics into clusters with clinically meaningful differences that were independent of Type-2 inflammation.

Among the 110 proteins significantly altered in severe compared to milder disease, many were known to be involved in regulation of immunological pathways. For example, the chitinases (CHI3L1 [YKL-40] and CHIT1 [chitotriosidase]) were elevated in both severe asthma groups irrespective of smoking status compared with HC, in line with previous data (19). Multiple inflammatory cytokines associated with the pathobiology of severe asthma were also differentially abundant, including IL-4,

IL-6, IL-10, IL-13, IL-17A, IL-26, and TNFα (20, 21). Severe asthmatics also showed increased levels of surfactant proteins, such as surfactant protein D, a pattern-recognition molecule involved in innate immune responses with antimicrobial activity (22). Increased plasma levels are in line with previous findings (23), presumably reflecting increased epithelial damage and permeability within the peripheral airways.

Of the lipid mediator pathways, our findings suggest involvement of sphingolipids in asthma pathogenesis as multiple enzymes of sphingolipid metabolism were elevated in SA including ceramide synthases, sphingosine kinases, and sphingomyelin phosphodiesterases. Components of the sphingolipid pathways, that have been shown to be involved in asthma and linked to disease severity (24, 25), represent potential therapeutic targets (26). Several genome-wide association studies have also linked polymorphisms in the 17q21 locus where the ORMDL3 gene resides to both childhood and adult asthma (27, 28). ORMDL3 is a regulator of sphingolipid synthesis and genetic variants lead to increased expression as well as decreased sphingolipid *de novo* synthesis in children with asthma (29). Certain eicosanoid pathway enzymes were also elevated in severe asthma such as hematopoietic prostaglandin D synthase (HPGDS) and cyclooxygenase-1 (PTGS1 [COX-1]) which are involved in many processes, and in particular mast cell activation (30).

Further proteins found to be increased in severe compared to mild asthma included members of the complement and coagulation cascades, confirming recent reports (31-33), as well as metabolic factors such as insulin and leptin. Proteins related to oxidative stress, including superoxide dismutase (SOD2) and glutathione S-transferase (GSTP1), were also elevated. These proteins have previously been

associated with air pollution and mild asthma (34), but their pathobiological role in severe asthma is largely unknown.

The follow-up investigation in the BIOAIR cohort replicated some of the findings with ten proteins being significantly increased in SA compared to MMA in both cohorts. Many more proteins also followed the same trends as observed in U-BIOPRED, but did not reach statistical significance, possibly due to the smaller size of BIOAIR. It is also possible that this number may have been greater if plasma had been collected in exactly the same way in both studies (i.e. both heparin, or both EDTA plasma samples). However, the studies were planned and conducted at different times by different consortia, and therefore procedures were not identical. This study nevertheless demonstrates the complementary power of BIOAIR and U-BIOPRED, despite minor differences in methodology.

Among replicated findings were proteins with known involvement in airway fibrosis and remodelling such as TGF-β1 (35) and alpha-1-antichymotrypsin. The latter might be of particular relevance in COPD, where plasma levels have been found to be elevated (36). IL-6, which may play a role in the pathobiology of a specific, exacerbation-prone, asthma sub-phenotype (37), was also increased in SA patients. Interestingly, plasma macrophage inflammatory protein-3 was elevated, consistent with a strong association of the gene encoding for macrophage inflammatory protein-3 with suboptimal asthma control (38). Indeed, in the cluster analy-sis, macrophage inflammatory protein-3 showed the highest levels in the cluster with the worst ACQ-7 and AQLQ scores and was among the most significant proteins differentiating between clusters. In the context of asthma, apolipoprotein E has received attention due to its ability to suppress airway inflammation, and one may speculate that the increased levels observed in SA represent a protective

function (39). However, apolipoproteins may also be affected by statin therapy and for this reason we investigated whether APOE differed between MMA and SA following removal of patients with cardiovascular comorbidities. In both U-BIOPRED and BIOAIR, APOE remained significantly greater in SA than MMA, data not shown.

One particularly novel aspect of this study was the exploration of plasma protein profiles for the clustering of asthmatics into consensus groups beyond their clinical diagnosis. This yielded six clusters of subjects with differences in clinical and biochemical parameters. Clusters sorted by decreasing median plasma levels of the studied proteins were associated with phenotypes of decreasing severity. In summary, we observed associations between the cluster groups and age, age of asthma onset, BMI, FEV_1, exacerbations, blood neutrophils, serum hsCRP, asthma control and quality of life. However, no significant differences across clusters were observed in measures associated with Type-2 inflammation, including blood and sputum eosinophils, FeNO, total serum IgE, serum periostin, RASP Type-2 score, or atopy. Along with the relative lack of differences in protein levels observed when dividing asthmatics into Type-2-high and Type-2-low subgroups, these findings confirm that the protein panel was fit for the purpose of identifying Type-2-independent signals. Furthermore, the proportion of subjects with detectable urinary levels of prednisolone metabolites was similar among the clusters, suggesting that OCS use was not a factor driving the observed cluster differences.

Plasma proteins that were particularly important for cluster classification included carboxypeptidase A3 (CPA3, a mast cell specific carboxypeptidase stored in secretory granules), tripartite motif containing protein 33 (TRIM33), and TRAF3 interacting protein 2 (TRAF3IP2). In fact, the cluster-driven increase in carboxypeptidase A3 was mimicked by elevated urinary tetranor-PGDM, the major

urinary metabolite of the main mast cell prostanoid PGD_2. Collectively, this evidence suggests a pivotal role for mast cell activation in Type-2-independent severe asthma. Interestingly, and in contrast to these findings, we previously showed that PGD_2 is associated with Type-2 inflammation (18). Thus, taken together our findings indicate that mast cells are diametric immune cells with a wide-spread involvement in both Type-2 and non-Type-2 inflammation.

Several new biological treatments have been approved for insufficiently controlled Type-2 asthma, although what currently represents a major unmet medical need is improved therapy for the subgroup of asthmatics with little or no Type-2 inflammation (1, 2, 40). Interestingly, the most striking differences in clinical characteristics between clusters included the high BMI observed in cluster 1, a group which also had the highest blood neutrophils, highest serum CRP, worst quality of life, and most exacerbations, possibly reflecting a non-Type-2 asthma group. Serum CRP has been suggested as a biomarker for neutrophilic asthma (41), and accordingly, a neutrophilic inflammatory phenotype of asthma was associated with increased systemic inflammation, as reflected by elevated serum CRP levels compared with non-neutrophilic asthma (42). Our findings may therefore provide a set of new, potential Type-2-independent biomarkers.

A major strength of the current investigation is the use of two cohorts of well-characterised asthmatic subjects, enabling validation of findings in an independent population and providing a comparison with COPD. Incidentally, severe asthma and COPD could not be differentiated based on levels of the measured proteins in plasma, providing support for “The Dutch Hypothesis” (43) and common molecular features in severe obstructive inflammatory airway diseases. One may speculate that the lack of difference between SA and COPD in BIOAIR could be due to the

presence of patients with asthma/COPD overlap in either of these groups. This is however unlikely as the inclusion/exclusion criteria of the BIOAIR study were designed to reduce possible overlap between diagnoses. An exclusion factor for severe asthma was current smoking (>10 cigarettes per day) and a history of greater than 5 pack years. In fact, the median number of pack years in the BIOAIR SA group was 0, and reversibility was also significantly higher, as expected, in SA compared to COPD (44). An exclusion factor for COPD was diagnosed asthma or allergy, and furthermore, COPD patients were required to be current or ex-smokers with a history of >15 pack years.

The two study designs were also complementary with U-BIOPRED providing data from a larger, observational, cross-sectional study. BIOAIR, on the other hand, although smaller, included two interventions: first four weeks of standardized treatment, which reduced biological variability due to differences in therapy, and then a placebo-controlled OCS trial. A further strength was the design of the array by the ChAMP consortium, based on multiple hypotheses concerning signalling pathways of relevance to asthma, airway inflammation, and immune-mediated diseases. The outcome was successful as the majority of proteins indeed displayed differential levels between disease severity and treatment groups. Finally, a methodological advantage of the assay itself was that for several analytes, multiple antibodies targeting the same protein were used, often showing supporting profiles (**figure S4**).

A further advantage of the current investigation is that the analysis was based on the readily available plasma matrix, supporting future development of less invasive tools for molecular asthma phenotyping at the point-of-care. The detection of elevated airway proteins, such as surfactant proteins, in plasma from subjects with severe asthma and COPD, suggesting increased leakage in severe disease, also

confirms that disease processes in the lung tissue may be reflected in plasma. Generally, the findings suggest a greater systemic impact of severe compared to mild-to-moderate asthma as protein levels were similar in mild asthma and healthy controls.

Possible effects of therapy are often overlooked in biomarker discovery efforts but the BIOAIR oral prednisolone intervention enabled a thorough assessment of the effect of oral corticosteroid therapy on plasma protein profiles. Results showed that levels of most proteins were reduced by OCS treatment. Interestingly, when U-BIOPRED subjects taking OCS were removed from analyses (**figure 2b**), the differences between mild-to-moderate and severe asthma were maintained for most proteins, although not all, suggesting that plasma levels were affected both by disease severity and OCS treatment. Moreover, proteins that were altered by OCS in the non-smoking severe asthmatics did not change with OCS use in the smoking severe asthma group where, in fact, no proteins changed (**figure 2c**). This may reflect the known decreased responsiveness to corticosteroids and differences in inflammatory biomarkers among smoking asthmatics (45, 46) and/or be a power issue. The severe asthma groups in both U-BIOPRED and BIOAIR were taking higher doses of inhaled corticosteroids than the mild-to-moderate groups, as per inclusion criteria, and it is important to note that ICS could potentially also affect protein expression. Unfortunately, a limitation of the current investigation is that we were not able to examine this possibility as rigorously as for OCS.

In conclusion, the antibody array we developed identified multiple plasma proteins associated with asthma severity, both confirming the involvement of known proteins such as inflammatory cytokines, growth factors, and chitinases, as well as suggesting novel targets for further investigation. For example, our findings advocate

components of the sphingolipid pathway and complement cascades as well as mast cell related proteins as being worthy of more detailed future examinations. These discoveries were found to be independent of effects potentially related to oral corticosteroid therapy. Furthermore, we showed that protein profiles, driven by amongst others a mast cell-specific carboxypeptidase, could be used to group asthmatic subjects into six clinically distinct clusters. Taken together, we present a platform that is able to suggest novel biomarker candidates for molecular phenotyping, particularly relevant to non-Type-2 asthma. The panel may also aid discovery of future pharmacotherapeutic targets by exposing previously unexplored pathways.

ACKNOWLEDGEMENTS

Support by NBIS (National Bioinformatics Infrastructure Sweden) is gratefully acknowledged. We also thank the entire staff of the Human Protein Atlas and the group of Affinity proteomics at SciLifeLab Stockholm for their great efforts. All the members of U-BIOPRED Study Group and the BIOAIR consortium are gratefully acknowledged:

U-BIOPRED Study Group members: Hassan Ahmed, David Balgoma, Aruna T Bansal, Frédéric Baribaud, Jeanette Bigler, Bo Billing, Hans Bisgaard, Michael J Boedigheimer, Klaus Bønnelykke, Joost Brandsma, Paul Brinkman, Enrica Bucchioni, Dominic Burg, Andrew Bush, Amphun Chaiboonchoe, Toni Checa, Chris H Compton, Julie Corfield, Danen Cunoosamy, Arnaldo D'Amico, Rosalia Emma, Veit J Erpenbeck, Damijan Erzen, Klaus Fichtner, Neil Fitch, Louise J Fleming, Elena Formaggio, Urs Frey, Martina Gahlemann, Victoria Goss, Yi-Ke Guo, Simone

Hashimoto, John Haughney, Gunilla Hedlin, Pieter-Paul W Hekking, Tim Higenbottam, Jans M Hohlfeld, Cecile TJ Holweg, Alan J Knox, Jon Konradsen, Nikolaos Lazarinis, Diane Lefaudeux, Tracy Li, Matthew J Loza, Rene Lutter, Alexander Manta, Sarah Masefield, John G Matthews, Alexander Mazein, Andrea Meiser, Montse Miralpeix, Nadia Mores, Clare S Murray, David Myles, Shama Naz, Björn Nordlund, Laurie Pahus, Ioannis Pandis, Stelios Pavlidis, Anthony Postle, Pippa Powel, Navin Rao, Stacey Reinke, Amanda Roberts, Graham Roberts, Anthony Rowe, James PR Schofield, Wolfgang Seibold, Anna Selby, Ralf Sigmund R, Florian Singer, Marcus Sjödin, Paul J Skipp, Ana R Sousa, Kai Sun, Bob Thornton, Mohib Uddin, Wim M van Aalderen, Marleen van Geest, Jorgen Vestbo, Nadja H Vissing, Arianne H Wagener, Scott S Wagers, Zsoka Weiszhart, Craig E Wheelock, Åsa Wheelock, Susan J Wilson, and Valentyna Yasinska.

BIOAIR consortium members: Guy G Brusselle, Deborah A Campbell, Marco Contoli, Katarina Damm, Isabelle de Rudder, Ingrid Delin, Catherine Devautour, Mariusz Duplaga, Marianne Eduards, Alexandra Ek, Tommy Ekström, Ewa Figiel, Flora Gaber, Stefanie Gauw, Agnieszka Gawlewicz-Mroczka, Daisy Gerding, Shushila Haque, Lorraine Hewitt, Pieter S Hiemstra, Stephen T Holgate, John Holloway, Aleksander Kania, Frank Kanniess, Östen Karlsson, Johan C Kips, Maria Kumlin, Ann-Sofie Lantz, Nikos Lazarinis, Helgo Magnussen, Patrick Mallia, Ingrid Martling, Lahouari Meziane, Erasmia Oikonomidou, Marianne Olsson, Elisabetta Pace, Eva Papadopouli, Nikos Papadopoulos, Maria Plataki, Mirella Profita, Lovisa E Reinius, Kai Richter, Douglas S Robinson, Micaela Romagnoli, Katerina Samara, Vanessa Schelfhout, Maria Skedinger, Evangelia Stamataki, Anneke ten Brinke, Isabelle

Vachier, Eva Wallén-Nielsen, Ilonka van Veen, Els Weersink, Susan J Wilson, Valentyna Yasinska, Eleftherios Zervas, and Bozena Ziolkowska-Graca.

ABBREVIATIONS

BIOAIR = Longitudinal Assessment of Clinical Course and BIOmarkers in Severe Chronic AIRway Disease, ChAMP = Centre for Allergy Research highlights Asthma Markers of Phenotype, COPD = chronic obstructive pulmonary disease, FDR = false discovery rate, FeNO = fraction of exhaled nitric oxide, FEV_1 = forced expiratory volume in 1 s, HC = healthy non-smoking controls, HPA = Human Protein Atlas, MMA = non-smoking subjects with mild-to-moderate asthma, OCS = oral corticosteroid, SAn = non-smoking subjects with severe asthma, SAs/ex = current smokers or ex-smokers with severe asthma, U-BIOPRED = Unbiased Biomarkers for the Prediction of Respiratory Disease outcome

REFERENCES

1. Pavord ID, Beasley R, Agusti A, Anderson GP, Bel E, Brusselle G, Cullinan P, Custovic A, Ducharme FM, Fahy JV, Frey U, Gibson P, Heaney LG, Holt PG, Humbert M, Lloyd CM, Marks G, Martinez FD, Sly PD, von Mutius E, Wenzel S, Zar HJ, Bush A. After asthma: redefining airways diseases. *Lancet* 2018; 391: 350-400.
2. Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat Med* 2012; 18: 716-725.
3. Papi A, Brightling C, Pedersen SE, Reddel HK. Asthma. *Lancet* 2018; 391: 783-800.
4. Sze E, Bhalla A, Nair P. Mechanisms and therapeutic strategies for non-T2 asthma. *Allergy* 2020; 75: 311-325.
5. Shaw DE, Sousa AR, Fowler SJ, Fleming LJ, Roberts G, Corfield J, Pandis I, Bansal AT, Bel EH, Auffray C, Compton CH, Bisgaard H, Bucchioni E, Caruso M, Chanez P, Dahlen B, Dahlen SE, Dyson K, Frey U, Geiser T, Gerhardsson de Verdier M, Gibeon D, Guo YK, Hashimoto S, Hedlin G, Jeyasingham E, Hekking PP, Higenbottam T, Horvath I, Knox AJ, Krug N, Erpenbeck VJ, Larsson LX, Lazarinis N, Matthews JG, Middelveld R, Montuschi P, Musial J, Myles D, Pahus L, Sandstrom T, Seibold W, Singer F, Strandberg K, Vestbo J, Vissing N, von Garnier C, Adcock IM, Wagers S, Rowe A, Howarth P, Wagener AH, Djukanovic R, Sterk PJ, Chung KF, U-BIOPRED Study Group. Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. *Eur Respir J* 2015; 46: 1308-1321.

6. Kupczyk M, Haque S, Sterk PJ, Nizankowska-Mogilnicka E, Papi A, Bel EH, Chanez P, Dahlen B, Gaga M, Gjomarkaj M, Howarth PH, Johnston SL, Joos GF, Kanniess F, Tzortzaki E, James A, Middelveld RJ, Dahlen SE, BIOAIR investigators. Detection of exacerbations in asthma based on electronic diary data: results from the 1-year prospective BIOAIR study. *Thorax* 2013; 68: 611-618.

7. Mikus M, Drobin K, Gry M, Bachmann J, Lindberg J, Yimer G, Aklillu E, Makonnen E, Aderaye G, Roach J, Fier I, Kampf C, Gopfert J, Perazzo H, Poynard T, Stephens C, Andrade RJ, Lucena MI, Arber N, Uhlen M, Watkins PB, Schwenk JM, Nilsson P, Schuppe-Koistinen I. Elevated levels of circulating CDH5 and FABP1 in association with human drug-induced liver injury. *Liver Int* 2017; 37: 132-140.

8. Schwenk JM, Gry M, Rimini R, Uhlen M, Nilsson P. Antibody suspension bead arrays within serum proteomics. *J Proteome Res* 2008; 7: 3168-3179.

9. Ihaka R, Gentleman R. R: a language for data analysis and graphics. *Journal of computational and Graphical Statistics* 1996: 299-314.

10. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

11. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological* 1995; 57: 289-300.

12. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. *Machine Learning* 2003; 52: 91-118.

13. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 2010; 26: 1572-1573.

14. Senbabaoglu Y, Michailidis G, Li JZ. Critical limitations of consensus clustering in class discovery. *Sci Rep* 2014; 4: 6207.

15. Lefaudeux D, De Meulder B, Loza MJ, Peffer N, Rowe A, Baribaud F, Bansal AT, Lutter R, Sousa AR, Corfield J, Pandis I, Bakke PS, Caruso M, Chanez P, Dahlen SE, Fleming LJ, Fowler SJ, Horvath I, Krug N, Montuschi P, Sanak M, Sandstrom T, Shaw DE, Singer F, Sterk PJ, Roberts G, Adcock IM, Djukanovic R, Auffray C, Chung KF, U-BIOPRED Study Group. U-BIOPRED clinical adult asthma clusters linked to a subset of sputum omics. *J Allergy Clin Immunol* 2017; 139: 1797-1807.

16. Kursa MB, Rudnicki WR. Feature Selection with the Boruta Package. *Journal of Statistical Software* 2010; 36: 1-13.

17. Heaney LG, Busby J, Hanratty CE, Djukanovic R, Woodcock A, Walker SM, Hardman TC, Arron JR, Choy DF, Bradding P, Brightling CE, Chaudhuri R, Cowan DC, Mansur AH, Fowler SJ, Niven RM, Howarth PH, Lordan JL, Menzies-Gow A, Harrison TW, Robinson DS, Holweg CTJ, Matthews JG, Pavord ID, on behalf of the investigators for the MRC Refractory Asthma Stratification Programme. Composite type-2 biomarker strategy versus a symptom-risk-based algorithm to adjust corticosteroid dose in patients with severe asthma: a multicentre, single-blind, parallel group, randomised controlled trial. *Lancet Respir Med* 2020.

18. Kolmert J, Gomez C, Balgoma D, Sjodin M, Bood J, Konradsen JR, Ericsson M, Thorngren JO, James A, Mikus M, Sousa AR, Riley JH, Bates S, Bakke PS,

Pandis I, Caruso M, Chanez P, Fowler SJ, Geiser T, Howarth P, Horvath I, Krug N, Montuschi P, Sanak M, Behndig A, Shaw DE, Knowles RG, Holweg CTJ, Wheelock AM, Dahlen B, Nordlund B, Alving K, Hedlin G, Chung KF, Adcock IM, Sterk PJ, Djukanovic R, Dahlen SE, Wheelock CE, U-BIOPRED Study Group. Urinary Leukotriene E4 and Prostaglandin D2 Metabolites Increase in Adult and Childhood Severe Asthma Characterized by Type-2 Inflammation. *Am J Respir Crit Care Med* 2020.

19. James AJ, Reinius LE, Verhoek M, Gomes A, Kupczyk M, Hammar U, Ono J, Ohta S, Izuhara K, Bel E, Kere J, Soderhall C, Dahlen B, Boot RG, Dahlen SE, BIOAIR Consortium. Increased YKL-40 and Chitotriosidase in Asthma and Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* 2016; 193: 131-142.
20. Konradsen JR, Nordlund B, Levanen B, Hedlin G, Linden A. The cytokine interleukin-26 as a biomarker in pediatric asthma. *Respir Res* 2016; 17: 32.
21. Linden A, Dahlen B. Interleukin-17 cytokine signalling in patients with asthma. *Eur Respir J* 2014; 44: 1319-1331.
22. Sorensen GL. Surfactant Protein D in Respiratory and Non-Respiratory Diseases. *Front Med (Lausanne)* 2018; 5: 18.
23. Mackay RM, Grainge CL, Lau LC, Barber C, Clark HW, Howarth PH. Airway Surfactant Protein D Deficiency in Adults With Severe Asthma. *Chest* 2016; 149: 1165-1172.
24. Jiang Y, Gruzieva O, Wang T, Forno E, Boutaoui N, Sun T, Merid SK, Acosta-Pérez E, Kull I, Canino G, Antó JM, Bousquet J, Melén E, Chen W, Celedón JC. Transcriptomics of atopy and atopic asthma in white blood cells from children and adolescents. *European Respiratory Journal* 2019: 1900102.

25. Reinke SN, Gallart-Ayala H, Gómez C, Checa A, Fauland A, Naz S, Kamleh MA, Djukanović R, Hinks TS, Wheelock CE. Metabolomics analysis identifies different metabotypes of asthma severity. *Eur Respir J* 2017; 49.

26. Heras AF, Veerappan A, Silver RB, Emala CW, Worgall TS, Perez-Zoghbi J, Worgall S. Increasing Sphingolipid Synthesis Alleviates Airway Hyperreactivity. *Am J Respir Cell Mol Biol* 2020; 63: 690-698.

27. Ono JG, Worgall TS, Worgall S. 17q21 locus and ORMDL3: an increased risk for childhood asthma. *Pediatr Res* 2014; 75: 165-170.

28. Sturgill JL. Sphingolipids and their enigmatic role in asthma. *Adv Biol Regul* 2018; 70: 74-81.

29. Ono JG, Kim BI, Zhao Y, Christos PJ, Tesfaigzi Y, Worgall TS, Worgall S. Decreased sphingolipid synthesis in children with 17q21 asthma–risk genotypes. *The Journal of Clinical Investigation* 2020; 130: 921-926.

30. Daham K, James A, Balgoma D, Kupczyk M, Billing B, Lindeberg A, Henriksson E, FitzGerald GA, Wheelock CE, Dahlen SE, Dahlen B. Effects of selective COX-2 inhibition on allergen-induced bronchoconstriction and airway inflammation in asthma. *J Allergy Clin Immunol* 2014; 134: 306-313.

31. de Boer JD, Majoor CJ, van 't Veer C, Bel EH, van der Poll T. Asthma and coagulation. *Blood* 2012; 119: 3236-3244.

32. Pandya PH, Wilkes DS. Complement system in lung disease. *Am J Respir Cell Mol Biol* 2014; 51: 467-473.

33. Zhang X, Kohl J. A complex role for complement in allergic asthma. *Expert Rev Clin Immunol* 2010; 6: 269-277.

34. Kelly FJ, Sandstrom T. Air pollution, oxidative stress, and allergic response. *Lancet* 2004; 363: 95-96.

35. Aschner Y, Downey GP. Transforming Growth Factor-beta: Master Regulator of the Respiratory System in Health and Disease. *Am J Respir Cell Mol Biol* 2016; 54: 647-655.

36. Hollander C, Westin U, Wallmark A, Piitulainen E, Sveger T, Janciauskiene SM. Plasma levels of alpha1-antichymotrypsin and secretory leukocyte proteinase inhibitor in healthy and chronic obstructive pulmonary disease (COPD) subjects with and without severe alpha1-antitrypsin deficiency. *BMC Pulm Med* 2007; 7: 1.

37. Jevnikar Z, Ostling J, Ax E, Calven J, Thorn K, Israelsson E, Oberg L, Singhania A, Lau LCK, Wilson SJ, Ward JA, Chauhan A, Sousa AR, De Meulder B, Loza MJ, Baribaud F, Sterk PJ, Chung KF, Sun K, Guo Y, Adcock IM, Payne D, Dahlen B, Chanez P, Shaw DE, Krug N, Hohlfeld JM, Sandstrom T, Djukanovic R, James A, Hinks TSC, Howarth PH, Vaarala O, van Geest M, Olsson H, Unbiased Biomarkers in Prediction of Respiratory Disease Outcomes study group. Epithelial IL-6 trans-signaling defines a new asthma phenotype with increased airway inflammation. *J Allergy Clin Immunol* 2019; 143: 577-590.

38. Croteau-Chonka DC, Qiu W, Martinez FD, Strunk RC, Lemanske RF, Jr., Liu AH, Gilliland FD, Millstein J, Gauderman WJ, Ober C, Krishnan JA, White SR, Naureckas ET, Nicolae DL, Barnes KC, London SJ, Barraza-Villarreal A, Carey VJ, Weiss ST, Raby BA, Asthma BioRepository for Integrative Genomic Exploration Consortium. Gene Expression Profiling in Blood Provides Reproducible Molecular Insights into Asthma Control. *Am J Respir Crit Care Med* 2017; 195: 179-188.

39. Yao X, Gordon EM, Figueroa DM, Barochia AV, Levine SJ. Emerging Roles of Apolipoprotein E and Apolipoprotein A-I in the Pathogenesis and Treatment of Lung Disease. *Am J Respir Cell Mol Biol* 2016; 55: 159-169.

40. Fahy JV. Type 2 inflammation in asthma--present in most, absent in many. *Nat Rev Immunol* 2015; 15: 57-65.

41. Schleich F, Demarche S, Louis R. Biomarkers in the Management of Difficult Asthma. *Curr Top Med Chem* 2016; 16: 1561-1573.

42. Wood LG, Baines KJ, Fu J, Scott HA, Gibson PG. The neutrophilic inflammatory phenotype is associated with systemic inflammation in asthma. *Chest* 2012; 142: 86-93.

43. Postma DS, Weiss ST, van den Berge M, Kerstjens HA, Koppelman GH. Revisiting the Dutch hypothesis. *J Allergy Clin Immunol* 2015; 136: 521-529.

44. Delgado-Eckert E, James A, Meier-Girard D, Kupczyk M, Andersson LI, Bossios A, Mikus M, Ono J, Izuhara K, Middelveld R, Dahlén B, Gaga M, Siafakas NM, Papi A, Beghe B, Joos G, Rabe KF, Sterk PJ, Bel EH, Johnston SL, Chanez P, Gjomarkaj M, Howarth PH, Niżankowska-Mogilnicka E, Dahlén SE, Frey U. Lung function fluctuation patterns unveil asthma and COPD phenotypes unrelated to type 2 inflammation. *J Allergy Clin Immunol* 2021; Feb 3:S0091-6749(21)00162-7.

45. Thomson NC, Chaudhuri R, Heaney LG, Bucknall C, Niven RM, Brightling CE, Menzies-Gow AN, Mansur AH, McSharry C. Clinical outcomes and inflammatory biomarkers in current smokers and exsmokers with severe asthma. *J Allergy Clin Immunol* 2013; 131: 1008-1016.

46. Thomson NC, Shepherd M, Spears M, Chaudhuri R. Corticosteroid insensitivity in smokers with asthma : clinical evidence, mechanisms, and management. *Treat Respir Med* 2006; 5: 467-481.

FIGURE LEGENDS

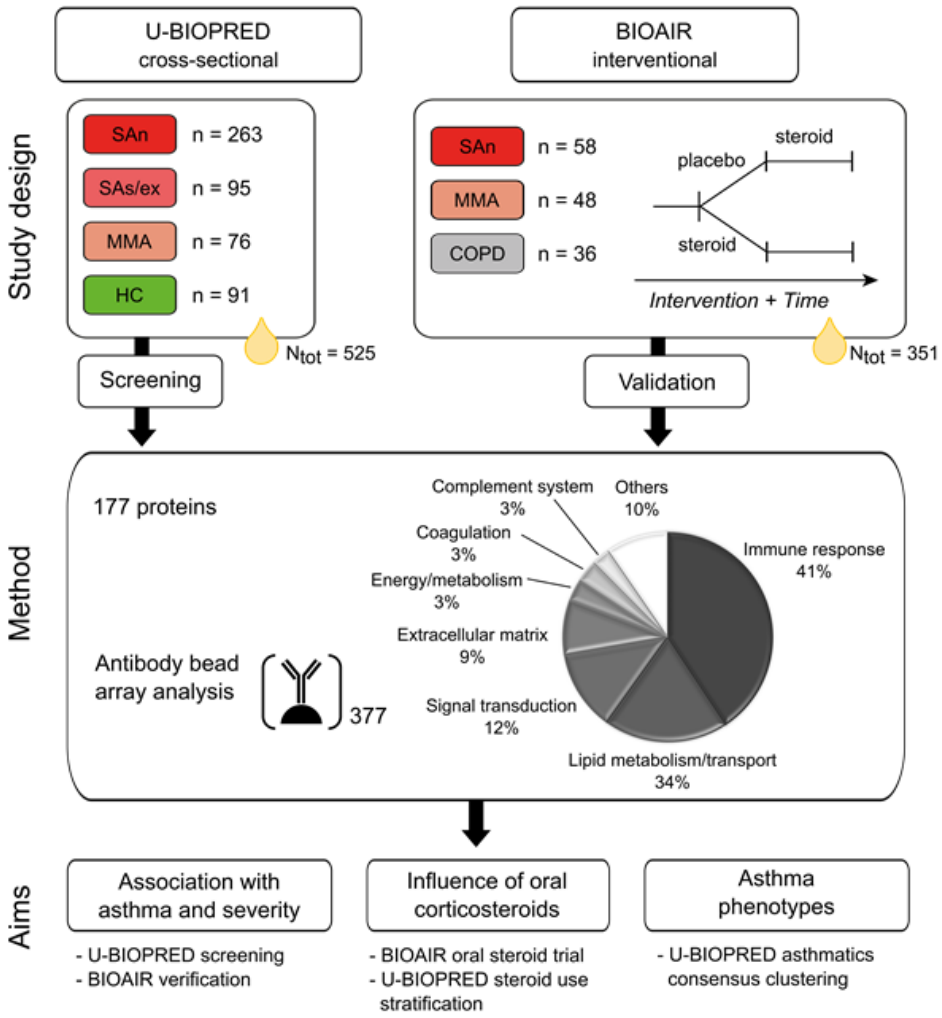

Figure 1. Study overview. Two independent cohorts, U-BIOPRED and BIOAIR, were investigated in this study. In a first screening, the U-BIOPRED cohort including 525 baseline plasma samples from 525 subjects was profiled using antibody bead arrays detecting 177 proteins with 377 antibodies. In the validation stage, the same array was used to profile the BIOAIR cohort comprising 351 plasma samples from 142 subjects. The BIOAIR cohort included a double-blind placebo-controlled oral corticosteroid intervention trial where the placebo group received additional open steroid treatment. These samples were used to study the influence of steroids on plasma protein levels. Asthmatic subjects from the U-BIOPRED cohort were used to identify potential phenotypes using consensus clustering of protein profiles. COPD = chronic obstructive pulmonary disease; HC = healthy non-smoking controls; MMA = non-smokers with mild-to-moderate asthma SAn = non-smokers with severe asthma; SAs/ex = smokers or ex-smokers with severe asthma.

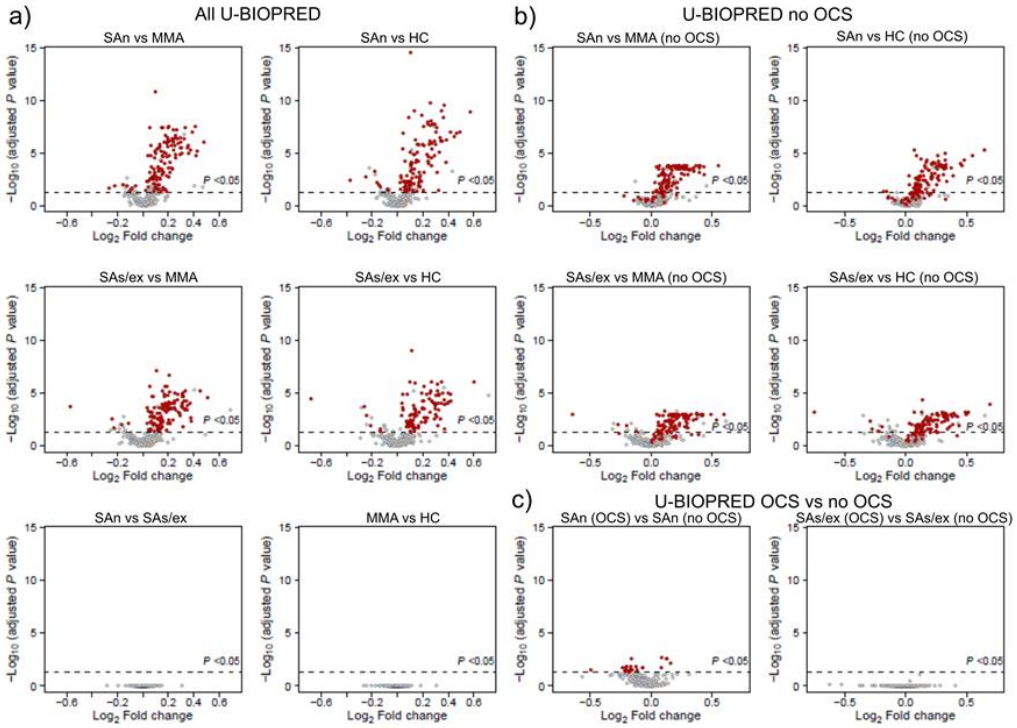

Figure 2. (a) Volcano plots of pairwise group comparisons in U-BIOPRED highlight multiple proteins elevated in subjects with severe asthma, but not in mild-to-moderate. Each dot represents a protein measured by the antibody array. Dashed lines represent adjusted *P* values <0.05. Highlighted in red are proteins significantly different in the respective pairwise group comparison as well as in the multiple group comparison (i.e. any of the 110 proteins). If multiple antibodies for the same target protein were significant, they all needed to show the same sign of log2 fold change. (b) Group comparisons in U-BIOPRED when limited to subjects where OCS use was not reported and confirmed negative by urinary analysis (SAn *n*=103, SAs/ex *n*=42, MMA *n*=63, HC *n*=90). Highlighted in red are the proteins found to be significantly different in the respective pairwise group comparison based on all subjects (i.e. proteins that were significantly different in the respective comparisons in figure 2a).

The majority of the proteins were still significantly different, seen by enrichment of red circles above the *P* value threshold. (c) Analysis of steroid effect in U-BIOPRED. SAn and SAs/ex stratified by reported OCS use and OCS metabolite detection. OCS use defined as prescribed and confirmed positive in urine (SAn *n*=51, SAs/ex *n*=21) and no OCS defined as not prescribed and confirmed negative in urine (SAn *n*=103, SAs/ex *n*=42). HC = healthy non-smoking controls; MMA = non-smokers with mild-to-moderate asthma; OCS = oral corticosteroids; SAn = non-smokers with severe asthma; SAs/ex = smokers or ex-smokers with severe asthma.

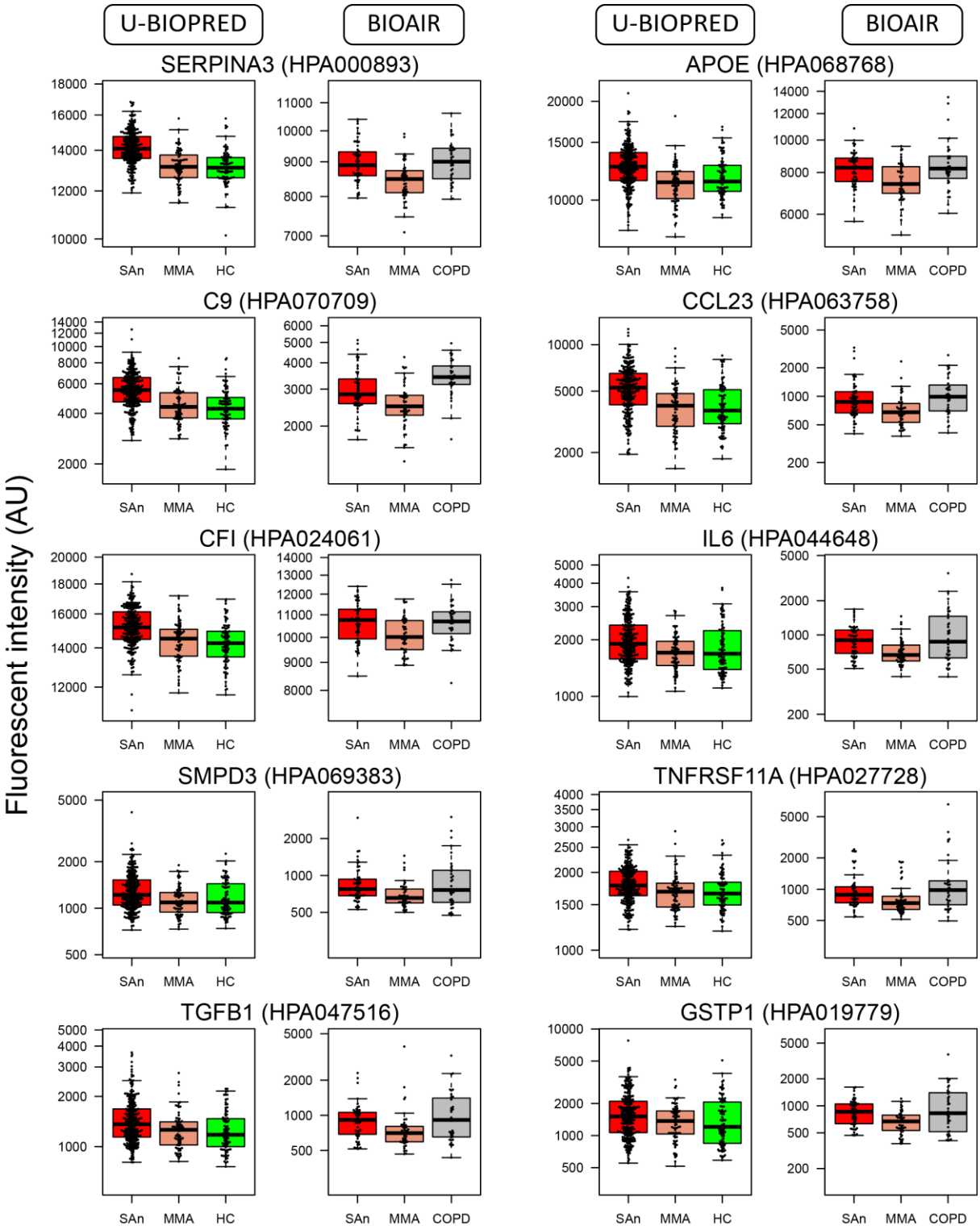

Figure 3. Proteins validated in two independent cohorts. Boxplots of the 10 proteins showing significantly different plasma levels between SAn and MMA in both studied cohorts (adjusted p<0.05). The screening cohort U-BIOPRED and the validation

cohort BIOAIR are shown side by side for each protein. COPD = chronic obstructive pulmonary disease; HC = healthy non-smoking controls; MMA = non-smokers with mild-to-moderate asthma; SAn = non-smokers with severe asthma.

SERPINA3 = Alpha-1-antichymotrypsin (ACT), APOE = Apolipoprotein E (Apo-E), C9 = Complement component 9 (C9), CCL23 = Macrophage inflammatory protein 3 (MIP-3), CFI = Complement factor I (C1), IL6 = Interleukin-6 (IL-6), SMPD3 = Sphingomyelin phosphodiesterase 3, TNFRSF11A = Receptor activator of nuclear factor kappa B (RANK), TGFB1 = Transforming growth factor beta-1 (TGF-β1), GSTP1 = Glutathione S-transferase P (GSTP1-1).

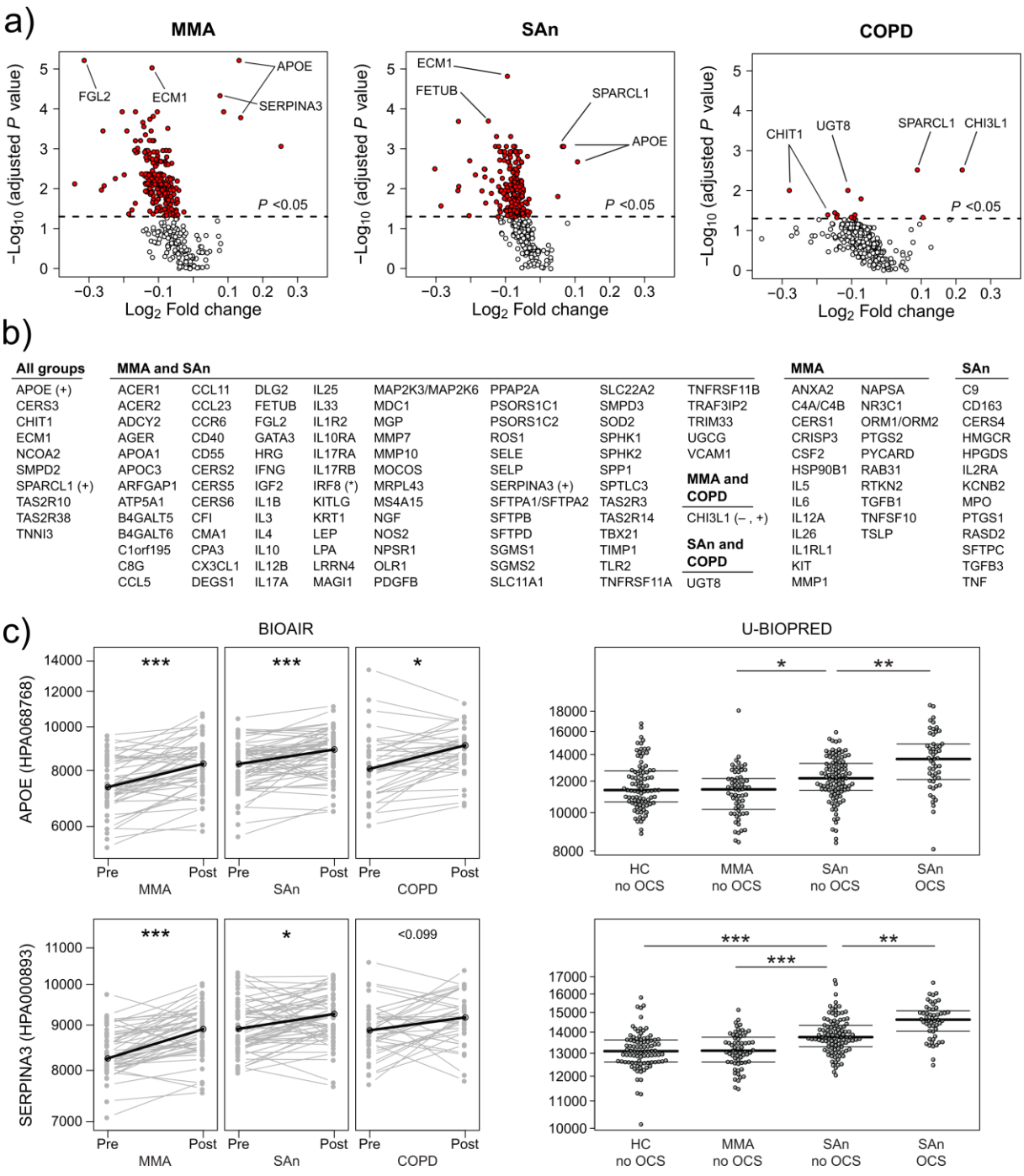

Figure 4. Influence of oral corticosteroids. (a) Volcano plots of comparisons in BIOAIR show that multiple proteins were affected by the steroid treatment, with the majority being decreased. Each dot represents a protein measured by the antibody array, with significantly changing proteins highlighted in red (above the dashed lines representing adjusted *P* values <0.05). The two most significant proteins on the decreasing and increasing side are labelled. Fold change calculated as log2 of the

median of individual ratios (post/pre). (b) List of proteins where the signal of at least one of the multiple antibodies targeting that protein in the array was affected by OCS. Decreasing levels indicated by no mark and increasing levels marked with (+). Mixed directions provided by multiple antibodies are marked with (*) and mixed directions of effect in different subject groups are marked with (-,+). (c) The plasma levels of apolipoprotein E (APOE) and alpha 1-antichymotrypsin (SERPINA3) increased after oral corticosteroids (OCS) in BIOAIR subject groups. In U-BIOPRED, APOE and SERPINA3 were associated with the severity of asthma as well as with the use of OCS among the severe asthmatics. Adjusted Wilcoxon signed-rank (BIOAIR) and rank-sum test (U-BIOPRED) *P* values *<0.05, ** <0.01, *** <0.001. COPD = chronic obstructive pulmonary disease; HC = healthy non-smoking controls; MMA = non-smokers with mild-to-moderate asthma SAn = non-smokers with severe asthma.

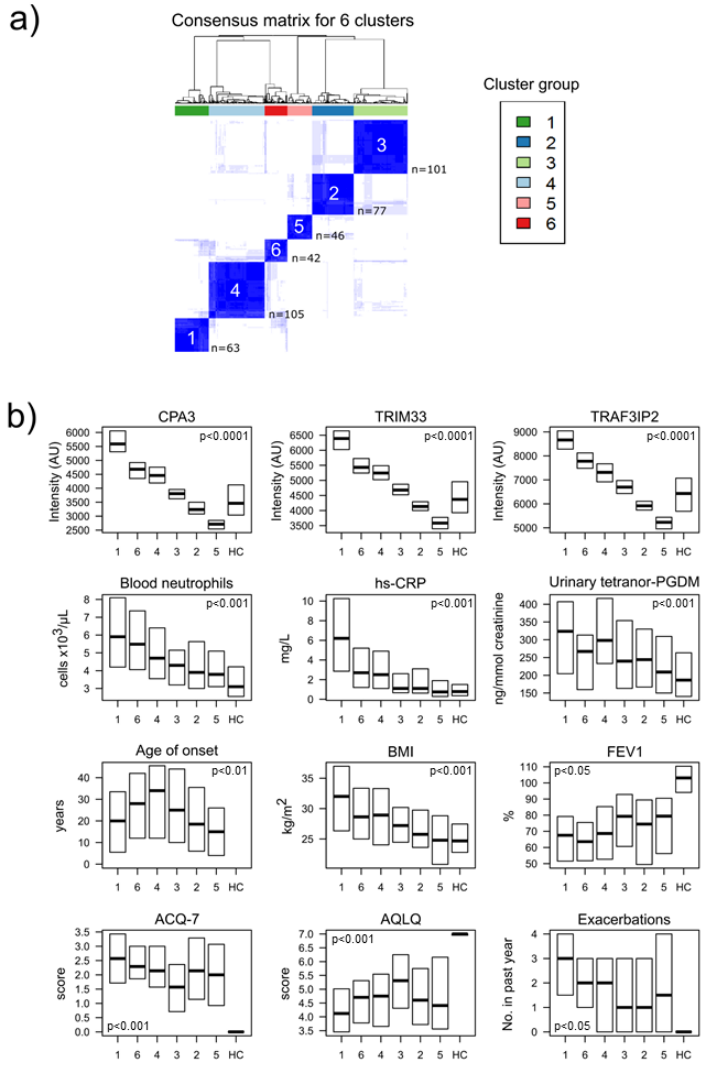

Figure 5. Phenotypes of asthma. (a) Six clusters were identified in a consensus clustering of asthmatic subjects (*n*=434, mild-to-moderate or severe asthma) from the U-BIOPRED cohort. The clustering was driven by the profiles of 110 proteins (139 antibodies). (b) The three proteins CPA3 (carboxypeptidase A3), TRIM33 (tripartite motif containing protein 33), and TRAF3IP2 (TRAF3 interacting protein 2) with the

greatest importance for cluster classification (for ranked importance of all proteins, see Figure S3) and a selection of variables with a significant association with clusters. Boxes show the first quartile, median and third quartile. P values are from the Kruskal Wallis test. ACQ-7 = asthma control questionnaire-average of 7; AQLQ = asthma quality of life questionnaire-average; BMI = body mass index; FEV1 = forced expiratory volume in 1 s; HC = non-smoking healthy controls; hsCRP = high-sensitivity C-reactive protein.

Table 1. Demographic and clinical characteristics of study subjects.

U-BIOPRED cohort

	HC (*n* = 91)	**MMA (*n* = 76)**	**San (*n* = 263)**	**SAs/ex (*n* = 95)**	***P* value† (incl. HC)**	***P* value† (excl. HC)**
Age (years)	36 (27-49)	43 (29-51)	53 (44-62)	55 (48-61)	<0.0001	<0.0001
Gender (% females)	37.4	48.7	66.2	51.6	<0.0001	0.004
BMI (kg/m^2)	24.7 (22.8-27.5)	24.8 (22.9-28.7)	27.5 (24.5-33.4)	29.0 (25.6-32.6)	<0.0001	<0.0001
Age of onset or diagnosis (years)	NA	15 (6-34)	20 (7-38)	38 (17-47)	-	<0.0001
FEV_1 (% predicted)	103.1 (94.1-110.3)	91.7 (78.5-101.0)	65.8 (50.4-84.5)	68.6 (55.1-78.2)	<0.0001	<0.0001
FEV_1/FVC	79.9 (75.3-82.9)	73.3 (67.1-78.6)	64.4 (52.9-74.4)	62.8 (54.5-70.2)	<0.0001	<0.0001
FeNO (ppb)	20.0 (13.4-29.2)	27.5 (19.0-56.0)	27.0 (16.0-47.0)	23.0 (12.0-44.0)	0.0006	0.1501
FEV_1 (% predicted)*	NA	91.4 (76.6-101.4)	64.2 (50.2-79.6)	67.8 (53.4-75.6)	-	<0.0001
FEV_1/FVC*	NA	73.0 (66.0-77.5)	63.1 (53.2-72.7)	62.1 (53.4-69.5)	-	<0.0001
Reversibility (%)	NA	9.6 (5.6-16.1)	13.3 (5.8-22.1)	14.6 (7.7-21.1)	-	0.090
OCS dose (normalized to mg Prednisolone)	NA	NA	10.0 (7.5-19.0)	13.8 (10.0-20.0)	-	0.065
OCS users (*n*)	NA	NA	137	46	-	
LTRA users (*n*)	NA	NA	114	41	-	
Blood eosinophils (cells/µL)	100 (90-200)	200 (100-300)	220 (100-465)	225 (110-403)	<0.0001	0.260
Blood neutrophils (cells/µL)	3100 (2550-4215)	3155 (2700-4425)	4720 (3600-6600)	4845 (3993-6728)	<0.0001	<0.0001
Blood WBC (cells/µL)	5700 (4725-6905)	5700 (5100-7300)	7980 (6400-9963)	8450 (6735-10300)	<0.0001	<0.0001
Serum periostin (ng/mL)	49.9 (43.9-57.8)	48.8 (41.6-54.7)	50.0 (42.0-60.5)	43.8 (36.2-59.6)	0.023	0.014
Serum total IgE (IU/mL)	23.4 (8.1-69.7)	94.7 (52.5-255.2)	120.0 (44.4-356.5)	119.4 (49.2-352.5)	<0.0001	0.666
hsCRP (mg/L)	0.8 (0.4-1.5)	0.7 (0.4-1.8)	2.0 (0.9-4.9)	2.4 (1.1-4.8)	<0.0001	<0.0001
Sputum eosinophils (%)	0.0 (0.0-0.2)	0.8 (0.2-3.4)	2.9 (0.6-21.5)	3.3 (0.7-13.3)	<0.0001	0.017
Sputum neutrophils (%)	41 (21-62)	42 (25-63)	55 (35-79)	55 (36-65)	0.015	0.065
ACQ-7	0.0 (0.0-0.0)	0.7 (0.4-1.3)	2.3 (1.7-3.1)	2.3 (1.6-3.0)	<0.0001	<0.0001
AQLQ	7.0 (7.0-7.0)	6.2 (5.4-6.5)	4.5 (3.6-5.4)	4.4 (3.5-5.2)	<0.0001	<0.0001

BIOAIR cohort

	MMA (*n* = 48)	**SAn (*n* = 58)**	**COPD (*n* = 36)**	***P* value†**
Age (years)	42 (32-53)	51 (42-58)	66 (55-70)	<0.0001
Gender (% females)	66.7	60.3	36.1	0.015
BMI (kg/m^2)	24.6 (22.4-27.1)	27.4 (25.6-30.7)	25.9 (22.4-29.6)	0.0007
Age of diagnosis (years)	18 (4-33)	31 (20-42)	60 (49-66)	<0.0001
FEV_1 (% predicted)*	89.8 (79.3-100.9)	72.8 (56.2-88.4)	44.3 (36.9-57.4)	<0.0001
FEV_1/FVC (%)*	71.1 (61.9-76.2)	67.0 (57.3-76.2)	44.9 (40.8-54.0)	<0.0001
FeNO (ppb)	34.0 (20.7-59.1)	35.1 (14.4-77.0)	10.8 (7.5-18.3)	0.0007
Reversibility (%)	9.5 (7.3-14.1)	8.9 (3.5-13.7)	4.4 (1.9-5.7)	<0.0001
OCS users (*n*)	0	7	0	-
Blood eosinophils (cells/µL)	250 (160-380)	260 (100-480)	200 (100-280)	0.088
Blood neutrophils (cells/µL)	3700 (2610-4360)	4880 (3680-6880)	4920 (3760-5920)	0.0001
Blood WBC (cells/µL)	6330 (5400-7220)	7770 (6700-10300)	7500 (6560-9250)	<0.0001
Serum periostin (ng/L)	86 (71-104)	78 (69.5-101)	74 (60-91)	0.089
Serum total IgE (IU/mL)	172 (43.1-320)	152 (50.1-326)	54.5 (29.3-146)	0.013
hsCRP (mg/L)	1.0 (0.38-2.5)	2.2 (0.72-5.3)	3.1 (1.8-6.6)	0.0004
Sputum eosinophils (%)	1.2 (0.4-8.8)	11 (2.4-30)	0.11 (0-0.63)	<0.0001
Sputum neutrophils (%)	45 (17-69)	35 (25-63)	68 (50-77)	0.033
ACQ-7	0.86 (0.32-1.4)	1.7 (1.0-2.9)	2.0 (1.3-2.8)	<0.0001
SGRQ	17 (11-39)	42 (31-61)	41 (32-59)	<0.0001

Data are presented as median (interquartile range), unless otherwise stated.
*Pre-bronchodilator.
† Kruskal-Wallis, Wilcoxon rank-sum or Chi-squared test, unadjusted.

Definition of abbreviations: ACQ-7 = asthma control questionnaire-average of 7; AQLQ = asthma quality of life questionnaire-average; BMI = body mass index; COPD = chronic obstructive pulmonary disease; FeNO = fraction of exhaled nitric oxide; FEV_1 = forced expiratory volume in 1 s; FVC = forced vital capacity; HC = healthy non-smoking controls; hsCRP = high-sensitivity C-reactive protein; MMA = non-smokers with mild-to-moderate asthma; NA = Not applicable; OCS = oral corticosteroids; SAn = non-smokers with severe asthma; SAs/ex = smokers or ex-smokers with severe asthma; SGRQ = St. George's Respiratory Questionnaire score; WBC = white blood cells.

Table 2. Plasma proteins significantly different between subject groups in the U-BIOPRED cohort. There were 21 proteins targeted by 23 antibodies with an adjusted Kruskal-Wallis *P* value $<10^{-10}$. All plasma proteins were found elevated in SAn and SAs/ex compared to MMA and HC. Comparisons of SAn *vs.* SAs/ex and MMA *vs.* HC showed nonsignificant differences for all proteins (data not shown). Table was sorted by the adjusted Kruskal-Wallis *P* value, lowest on top. A list of all the 110 proteins that were significantly different is shown in Table E4.

Protein	Gene	Antibody	*-log P* value*	*-log P* value†			
				SAn *vs*. MMA	SAn *vs*. HC	SAs/ex *vs*. MMA	SAs/ex *vs*. HC
Alpha-1-antichymotrypsin (ACT)	SERPINA3	HPA000893	22.01	10.82	14.52	7.08	9.05
Complement component 9 (C9)	C9	HPA029577	14.55	7.51	9.77	5.11	6.10
Complement component 9 (C9)	C9	HPA070709	14.03	7.36	9.51	5.00	6.10
Lung surfactant protein D (SPD)	SFTPD	HPA044582	13.72	7.36	9.00	5.21	5.92
Tumor necrosis factor alpha (TNFα)	TNF	HPA055037	13.04	7.51	8.39	5.21	5.03
SPARC-like protein 1	SPARCL1	HPA067641	12.70	7.36	8.60	4.72	5.01
Fractalkine	CX3CL1	HPA056729	12.37	6.01	8.85	4.59	6.10
Interleukin-2 receptor subunit alpha (IL-2Rα)	IL2RA	HPA046738	12.19	7.21	7.74	4.96	5.24
Complement factor I (C1)	CFI	HPA024061	11.96	5.72	8.38	4.66	6.10
Complement component 8 gamma chain	C8G	HPA046269	11.96	6.62	8.05	4.66	5.19
Fibrinogen alpha chain	FGA	HPA064755	11.96	7.36	6.85	5.60	4.92
Complement factor I (C1)	CFI	HPA001143	11.77	6.01	7.77	4.85	5.85
Membrane-associated guanylate kinase inverted 1 (MAGUK)	MAGI1	HPA031852	11.62	7.16	7.77	4.08	4.43
Macrophage inflammatory protein-3 (MIP-3)	CCL23	HPA063758	11.30	6.92	6.92	4.85	4.85
Fibroleukin	FGL2	HPA026682	11.29	6.17	7.96	4.08	5.01
Interleukin-1 receptor beta (IL-1Rβ)	IL1R2	HPA027598	11.11	6.48	7.44	4.32	4.85
Cyclooxygenase-1 (COX-1)	PTGS1	HPA002834	11.10	6.34	7.64	4.15	4.85
Ceramide synthase 4 (CerS4)	CERS4	HPA023621	11.10	6.25	7.74	4.21	4.74
E-selectin	SELE	HPA057891	11.03	6.62	7.54	4.08	4.26
Uteroglobin	SCGB1A1	HPA031828	10.96	6.48	7.00	4.66	4.74
Apolipoprotein E (Apo-E)	APOE	HPA068768	10.82	7.36	4.32	6.72	4.26
Leptin	LEP	HPA068565	10.80	5.46	8.96	1.77	3.57
Granulocyte-macrophage colony-stimulating factor (GM-CSF)	CSF2	HPA071579	10.74	6.32	6.11	5.24	4.96

*Kruskal-Wallis test, adjusted.
†Wilcoxon rank-sum test, adjusted.
Definition of abbreviations: HC = healthy non-smoking controls; MMA = non-smokers with mild-to-moderate asthma; SAn = non-smokers with severe asthma; SAs/ex = smokers or ex-smokers with severe asthma.

Table 3. Proteins successfully validated in both the U-BIOPRED and BIOAIR cohort. All proteins showed higher plasma levels in SAn compared to MMA.

Protein	Gene	Antibody	*-log P* value* SAn *vs.* MMA	
			U-BIOPRED	BIOAIR
Alpha-1-antichymotrypsin (ACT)	SERPINA3	HPA000893	10.82	2.22
Apolipoprotein E (Apo-E)	APOE	HPA068768	7.36	1.66
Complement component 9 (C9)	C9	HPA070709	7.36	1.57
Macrophage inflammatory protein 3 (MIP-3)	CCL23	HPA063758	6.92	1.55
Complement factor I (C1)	CFI	HPA024061	5.72	1.44
Interleukin-6 (IL-6)	IL6	HPA044648	3.28	1.66
Sphingomyelin phosphodiesterase 3	SMPD3	HPA069383	3.14	1.72
Receptor activator of nuclear factor kappa B (RANK)	TNFRSF11A	HPA027728	2.72	1.66
Transforming growth factor beta-1 (TGF-β1)	TGFB1	HPA047516	2.41	1.55
Glutathione S-transferase P (GSTP1-1)	GSTP1	HPA019779	1.33	1.72

*Wilcoxon rank-sum test, adjusted

Definition of abbreviations: MMA = non-smokers with mild-to-moderate asthma; SAn = non-smokers with severe asthma

ONLINE DATA SUPPLEMENT

Plasma proteins elevated in severe asthma despite oral steroid use and unrelated to Type-2 inflammation

Maria Sparreman Mikus, Johan Kolmert, Lars I. Andersson, Jörgen Östling, Richard G. Knowles, Cristina Gómez, Magnus Ericsson, John-Olof Thörngren, Payam Emami Khoonsari, Barbro Dahlén, Maciej Kupczyk, Bertrand De Meulder, Charles Auffray, Per S. Bakke, Bianca Beghe, Elisabeth H. Bel, Massimo Caruso, Pascal Chanez, Bo Chawes, Stephen J. Fowler, Mina Gaga, Thomas Geiser, Mark Gjomarkaj, Ildikó Horváth, Peter H. Howarth, Sebastian L. Johnston, Guy Joos, Norbert Krug, Paolo Montuschi, Jacek Musial, Ewa Niżankowska-Mogilnicka, Henric K. Olsson, Alberto Papi, Klaus F. Rabe, Thomas Sandström, Dominick E. Shaw, Nikolaos M. Siafakas, Mathias Uhlen, John H. Riley, Stewart Bates, Roelinde J.M. Middelveld, Craig E. Wheelock, Kian Fan Chung, Ian M. Adcock, Peter J. Sterk, Ratko Djukanovic, Peter Nilsson, Sven-Erik Dahlén, and Anna James; on behalf of the U-BIOPRED (Unbiased Biomarkers for the Prediction of Respiratory Disease outcome) Study Group and the BIOAIR (Longitudinal Assessment of Clinical Course and Biomarkers in Severe Chronic Airway Disease) Consortium

SUPPLEMENTARY METHODS

Subjects, inclusion and exclusion criteria and study design

U-BIOPRED

Asthma diagnosis was based on fulfilling one of following four criteria:
1) Airflow reversibility (an in increase in forced expiratory volume in 1 second (FEV1) ≥12% of predicted or improvement of 200 ml after inhalation of 400 µg salbutamol);

2) Airway hyperresponsiveness (methacholine provocative concentration causing 20% fall in FEV1 <8 mg/ml);

3) Diurnal peak expiratory flow (PEF) amplitude >8% of mean;

4) A decrease in FEV1 >12% of predicted and >200 ml within 4 weeks after tapering maintenance treatment (one or more of: inhaled corticosteroids (ICS), oral corticosteroids (OCS), long-acting β-agonists (LABA), short-acting β-agonists (SABA))

Plus a history of wheeze, either spontaneous or on exertion.

Subjects with severe asthma had been under follow-up by a respiratory specialist for at least 6 months prior to enrolment in the study. During this time, investigations to exclude other diagnoses were performed and steps taken to optimise asthma control (treatment of co-morbidities, assessment of adherence, and reduction in allergen exposure). Non-smokers with severe asthma (SAn) were defined according to GINA guidelines [1] and/or frequent exacerbations (more than 2 per year) despite high-dose ICS (ICS ≥1000 µg fluticasone propionate/day or equivalent dose, or a doubling of the dose of maintenance ICS for at least three days or requiring hospitalisation), with or without OCS, plus at least one other controller medication. They had been non-smokers for at least the past 12 months with a pack-year history of ≤5. Smoker and ex-smokers with severe asthma (SAs/ex) were defined as SAn except being either current smokers or former smokers with a pack-year history of >5.

Non-smokers with mild-to-moderate asthma (MMA) were defined by GINA guidelines [1] and had controlled or partially controlled asthma symptoms whilst receiving a dose of less than 500 µg fluticasone propionate/day or equivalent. They had been non-smokers for at least the past 12 months with a pack-year history of ≤5.

Healthy non-smoking controls (HC) had no history of asthma, wheeze or other chronic respiratory disease. They had a pre-bronchodilator FEV1 of ≥80% predicted and were non-smokers for at least the past 12 months with pack-year history of ≤5.

Exclusion criteria included a history of recreational drug use, allergy which contraindicated participation, or a recent history of incapacitating psychiatric disorders. Female subjects who were pregnant or breast feeding (or up to 6 weeks post-partum or cessation of breast feeding) were also excluded. No participation within 3 months after the first dose using a new molecular entity, drug or invasive procedure in another study. In addition, subjects were to have no history of or current evidence of upper or lower respiratory infection or symptoms within 2 weeks of baseline assessment (in which case assessments were deferred).

Study design

In this prospective cohort study, all subjects were identified and recruited locally and attended a screening visit when eligibility for the study was confirmed. All subjects underwent a baseline visit (up to 28 days later) at which study data and biological samples were collected.

BIOAIR

Asthma and COPD were diagnosed by pulmonary specialists according to published criteria [2-5]. The diagnosis of asthma had to be confirmed by at least one of the four criteria for reversible airway obstruction as documented during the last 5 years before the study or at screening visit, namely:

1) an increase in FEV1 ≥9% of predicted (or improvement of 200 ml) after administration of four puffs of 100 µg salbutamol dose-aerosol inhaled via a spacer, or after additional inhalation of four puffs of 20 µg ipratropium bromide administered through a large volume spacer;

2) mean diurnal variation in PEF of ≥15% on ≥4 days/week for at least 2 weeks, as calculated by the following equation: (highest PEF–lowest PEF)/mean PEF;

3) an increase in FEV1 of at least 400 ml after a course of prednisolone 0.5 mg/kg/day for 14 days;

4) in patients with a FEV1 ≥70% predicted, demonstrated bronchial hyperresponsiveness to histamine, methacholine, isocapnic hyperventilation, exercise or other indirect challenges (according to established local methods).

Subjects in the severe asthma group had been under specialist treatment for at least one year and had experienced at least one exacerbation in the past year. Exacerbations were defined either by the initiation of a course of OCS therapy in those individuals on regular ICS treatment, or for those on regular OCS therapy, a significant temporary increase in their dose of oral steroids for an acute deterioration in their disease control. Severe asthmatics had to require continuous treatment with high doses of ICS (at least 1600 µg/day budesonide or beclomethasone, 800 µg/day fluticasone or equivalent). For those taking oral steroids, the inhaled dose of steroids had to be at least 800 µg/day budesonide or beclomethasone, 400 µg/day fluticasone or equivalent. In addition, the subjects had to require continuous treatment with LABA or oral theophylline, as documented for at least one year.

Non-smokers with mild-to-moderate asthma (MMA) had stable disease and received daily treatment with a maximum of 800 µg/day budesonide or beclomethasone, 500 µg/day fluticasone or equivalent. MMA subjects used SABA as needed but did not require treatment with LABA and had had no exacerbations, nor hospitalisations in the past year.

Subjects recruited to the COPD group had a FEV1/FVC (Forced Vital Capacity) ratio of less than 0.7, and a post-bronchodilator FEV1 value between 30 and 80% of predicted, that increased by less than 9% (or 200 ml) after inhalation of bronchodilator. They were required to have been treated with ICS in the range of

800-1600 µg/day (beclomethasone, budesonide or equivalent) for at least three months prior to the study. A smoking history of more than 15 pack years, either as a current or ex-smoker, was required in the COPD group, whereas asthma patients who had smoked for more than 5 pack years were excluded.

Subjects were excluded if they were pregnant, had a history of alcohol or illicit drug abuse, had other acute or chronic pulmonary disorders or had clinically relevant psychiatric disease. They were not allowed to be receiving immunosuppressants other than corticosteroids, or undergoing immunotherapy. Subjects receiving chronic oxygen therapy were also excluded.

Study design

Subjects were screened during a half-day visit in one of the participating clinics following which all subjects underwent a 4-week treatment optimisation period. Medication details are shown below:

MMA 1. Fixed moderate daily dose of inhaled steroids (800 µg budesonide or beclomethasone, or 500 µg fluticasone). 2. Fixed daily dose of LABA (formoterol 9-24 µg b.i.d. or salmeterol 50 µg b.i.d.). 3. SABA prescribed as needed. 4. No other drugs should be taken except for topical nasal drugs or treatment of comorbidity.

SAn 1. Fixed high daily dose of inhaled steroids (1600 µg budesonide or 1500 µg beclomethasone, or 1000 µg fluticasone). 2. Fixed daily dose of LABA (formoterol 9-24 µg b.i.d. or salmeterol 50 µg b.i.d.). 3. SABA prescribed as needed. 4. If applicable, dose of oral prednisolone unchanged. 5. No other drugs should be taken except for topical nasal drugs or treatment of comorbidity. (1 and 2 could be administered as a fixed combination, provided that the dosage is accurate)

COPD 1. Fixed daily dose of LABA (formoterol 9-24 µg b.i.d. or salmeterol 50 µg b.i.d.). 2. Ipratropium 40 µg q.i.d. 3. Fixed daily dose of inhaled steroid (800 µg budesonide or beclomethasone, or 500 µg fluticasone). 4. SABA prescribed as needed. 5. No other drugs should be taken except for treatment of co-morbidity.

The optimisation period was followed by a 2-week double-blind placebo-controlled steroid intervention, consisting of a standard course of oral prednisolone of 0.5 mg/kg of body weight/day or placebo added to regular treatment. This was followed by an identical open OCS treatment of the placebo group only.

Designing a protein panel for plasma profiling

We designed a panel for protein profiling based on knowledge within the consortium, searches in databases (Ensembl, Genotator, the Human Protein Atlas, Ingenuity Pathway Analysis, Intomics, MetaCore) and the literature. The selection covered proteins reported or proposed to be dysregulated in asthma or chronic obstructive pulmonary disease (COPD) and/or being expressed in blood, lung or during inflammation, and resulted in 177 proteins (Table S2). Biological processes reflected by these proteins were the immune response, lipid mediator pathways, signal

transduction, extracellular matrix, energy/metabolic, coagulation and complement system. These proteins were targeted by 377 antibodies, the majority affinity-purified polyclonal antibodies from the Human Protein Atlas project (HPA, www.proteinatlas.org) [6] in addition to anti-CHI3L1 (AF2599 and MAB25991, R&D Systems, Minneapolis, MN, USA) and anti-HPGDS (MAB6487, R&D Systems). The antibody set included at least two and up to five antibodies per protein target for 125 of the proteins (71% of the full protein panel), see Table S3. These so-called paired antibodies were raised against the same or different amino acid sequences of the protein (ranging from partial to full protein sequence coverage and an average length of 80 amino acids).

Protein profiling using suspension bead arrays

Protein profiling in plasma was performed using antibody suspension bead arrays. The in-house developed array-based affinity proteomics method enables, in combination with the HPA antibodies, a dual dimensional multiplexing capacity for the analysis of 384 samples in parallel on 384 proteins. To create the bead array, antibodies were separately and covalently coupled to magnetic color-coded beads (MagPlex, Luminex Corp., Austin, TX, USA) before being combined into a multiplex suspension as previously described [7, 8].

Layouts of 96-well microtiter plates were designed for each of the two cohorts, U-BIOPRED and BIOAIR, separately. This design controlled for a balanced distribution of disease groups and controls (U-BIOPRED cohort: mild-to-moderate asthma [MMA], non-smokers with severe asthma [SAn], smokers or ex-smokers with severe asthma [SAs/ex], healthy non-smoking controls [HC]; BIOAIR cohort: MMA, SAn, COPD); intervention group (applicable for BIOAIR); gender and age across the multiple plates. In BIOAIR, where sampling was done at multiple time points, all samples from each subject were analysed on the same plate. The samples were then randomized within each plate and distributed using a liquid handling device (EVO150, Tecan Group, Männedorf, Switzerland). In addition, sample-free PBS buffer in duplicates and technical replicates constituting a pool of all samples within a cohort were included on each plate as negative controls and for quality assessment, respectively.

Plasma samples were diluted in PBS and the protein content directly labelled with biotin. Labelled samples were diluted in assay buffer, heat treated at 56°C before being combined with antibody-coupled beads in 384-well microtiter plates. Incubation took place overnight. Measurements were performed using FlexMAP3D instruments (Luminex corp.), in parallel where applicable, reporting relative fluorescent intensity values.

Data pre-processing

Outliers were identified and removed based on robust principal component analysis (PCA) and the R package “rrcov” [9]. This resulted in exclusion of samples from 6 subjects (4 SAn, 1 MMA, 1 HC) from the U-BIOPRED dataset and samples from all timepoints of 2 COPD subjects in the BIOAIR dataset. These samples are excluded from the numbers presented and not included in Table 1 and Table S1 or in the

statistical analysis. The raw intensity values were processed to minimize batch effects using the R package “MDimNormn” and function “normn_MA” [10], such that for each antibody the geometric mean of each plate was adjusted to match the geometric mean of the geometric mean values of multiple plates, thereby resulting in equal means for each plate.

Statistical analysis

Statistical analysis and visualization were performed in R [11, 12].

Univariate statistics
Nonparametric Kruskal-Wallis and Wilcoxon rank-sum tests were used for multiple group and pairwise group comparisons of continuous variables, respectively. Wilcoxon signed-rank tests were used in paired comparisons. For all proteins, *P* values were adjusted for multiple testing using the Benjamini and Hochberg [13] method and controlling the false discovery rate (FDR) at 5%.

Consensus cluster analysis
Consensus clustering is a method for class discovery where the clustering results are validated by a resampling scheme that simulates perturbations of the data, thereby allowing for assessment of the robustness of the clusters [14]. From the U-BIOPRED data, a reduced set of variables (n=139 antibodies) from 434 subjects with asthma (263 SAn, 95 SAs/ex, and 76 MMA) was selected for consensus cluster analysis on log2 transformed and z-scored intensity signals of each antibody. Variables were selected if significantly different (adjusted P <0.05) in the multiple group comparison in U-BIOPRED (Table S4). Consensus cluster analysis was performed using the ‘ConsensusClusterPlus’ package in R [15]. The Euclidean distance measure was used to describe similarity between subjects and partitioning around the medoids algorithm used for clustering. The clustering was repeated 1000 times, removing randomly 10% of the subjects at each iteration. Models of two to ten clusters were generated and cluster stability evaluated by two methods: 1) localizing the cluster model having the lowest proportion of ambiguously clustered subjects [16] which is demonstrated by the cluster model number having the longest flat line in the cumulative distribution function plot; and 2) the lowest deviation from ideal stability also based on the cumulative distribution function [17]. Evaluation of cluster models identified ten clusters to have slightly better numerical stability than six (Table S8, Figure S2). However, all clusters in the six-cluster model demonstrated greater consistency compared to all clusters in the ten-cluster model (Figure S2E) and was therefore selected for clinical and biochemical characterization in more detail. In an exploratory analysis, Kruskal-Wallis and Chi-squared tests were used to identify clinical and haematological variables that were significantly different between clusters.

Classification analysis with variable selection and random forest
To identify which of the 139 variables used in the consensus clustering were relevant for classification of the six clusters, machine learning was applied with feature (variable) selection based on a random forest classification algorithm.

In the pre-processing steps, there were no variables that needed to be removed due to zero or near-zero variance in their values and no variables with missing data.

Feature selection was performed using the Boruta algorithm (R package ‘Boruta’, all parameters set to default except number of trees ntree=1000, maximal number of iterations maxRuns=1000) [18]. Boruta is a so-called all-relevant variable method, meaning that it identifies all variables that are relevant for the classification. Boruta assesses the relevance by comparing the importance of the original variables with the importance of random permutations of the original variables. The importance of each variable is provided by a specified machine learning method. Here, the default method Random forest (R package ‘ranger’) was applied. To alleviate the problem of class imbalance, we adjusted the weights for the classes in the splitting rule of the random-forest algorithm according to the frequency of the classes. In short, Boruta works by first making a copy of each variable and randomly mixes the values of these to create “shadow variables”. This is followed by training a Random forest model on the extended dataset and extracting the importance measures (ranger normalized permutation importance). Original variables with a higher importance than the maximum importance seen for any of the shadow variables are recorded as a hit. The process is then repeated for a set number of times and a statistical test is used to test if a variable has significantly higher importance score than the shadow variables. *P* values were adjusted with the Bonferroni method and *P* values <0.01 considered significant. Variables are defined as confirmed important, unimportant or tentatively important (i.e. unresolved).

SUPPLEMENTARY TABLES AND FIGURES

Table S1. Percent (%) available values for each variable reported in Table 1 from the U-BIOPRED and BIOAIR cohorts.

U-BIOPRED

	HC (*n* = 91)	**MMA (*n* = 76)**	**SAn (*n* = 263)**	**SAs/ex (*n* = 95)**
Age (years)	100%	100%	100%	100%
BMI (kg/m^2)	100%	100%	100%	100%
Age of onset or diagnosis (years)	NA	96%	98%	99%
FEV_1 (% predicted)	100%	99%	99%	100%
FEV_1/FVC	100%	99%	99%	100%
FeNO (ppb)	97%	99%	95%	94%
FEV_1 (% predicted)*	NA	99%	99%	100%
FEV_1/FVC*	NA	99%	99%	100%
Reversibility (%)	NA	97%	98%	97%
Blood eosinophils (cells/µL)	100%	100%	98%	97%
Blood neutrophils (cells/µL)	100%	100%	98%	97%
Blood WBC (cells/µL)	100%	100%	99%	97%
Serum periostin (ng/mL)	90%	91%	90%	86%
Serum total IgE (IU/mL)	97%	97%	99%	97%
hsCRP (mg/L)	98%	99%	99%	100%
Sputum eosinophils (%)	41%	54%	44%	54%
Sputum neutrophils (%)	41%	54%	44%	54%
ACQ-7	32%	97%	97%	98%
AQLQ	31%	97%	99%	96%

BIOAIR

	MMA (*n* = 48)	**SAn (*n* = 58)**	**COPD (*n* = 36)**
Age (years)	100%	100%	100%
BMI (kg/m^2)	100%	100%	100%
Age of diagnosis (years)	90%	88%	97%
FEV_1 (% predicted)*	92%	98%	97%
FEV_1/FVC (%)*	92%	98%	97%
FeNO (ppb)	52%	48%	45%
Reversibility (%)	100%	97%	97%
Blood eosinophils (cells/µL)	94%	97%	100%
Blood neutrophils (cells/µL)	90%	95%	95%
Bloood WBC (cells/µL)	98%	100%	100%
Serum periostin (ng/L)	90%	98%	95%
Serum total IgE (IU/mL)	96%	98%	97%
hsCRP (mg/L)	92%	95%	87%
Sputum eosinophils (%)	60%	48%	37%
Sputum neutrophils (%)	60%	48%	37%
ACQ-7	92%	88%	66%
SGRQ	81%	91%	84%

*Pre-bronchodilator.

Definition of abbreviations: ACQ-7 = asthma control questionnaire-average of 7; AQLQ = asthma quality of life questionnaire-average; BMI = body mass index; COPD = chronic obstructive pulmonary disease; FeNO = fraction of exhaled nitric oxide; FEV_1 = forced expiratory volume in 1 s; FVC = forced

vital capacity; HC = healthy non-smoking controls; hsCRP = high-sensitivity C-reactive protein; MMA = non-smokers with mild-to-moderate asthma; OCS = oral corticosteroids; SAn = non-smokers with severe asthma; SAs/ex = smokers or ex-smokers with severe asthma; SGRQ = St. George's Respiratory Questionnaire score; WBC = white blood cells.

Table S2. List of the 177 proteins analysed with the antibody bead array, grouped by biological process.

Gene*	Gene description	Ensembl ID	UniProt ID	Number of antibodies
Immune response				
AGER	advanced glycosylation end-product specific receptor	ENSG00000204305	Q15109	2
ARFGAP1	ADP ribosylation factor GTPase activating protein 1	ENSG00000101199	Q8N6T3	3
CCL5	C-C motif chemokine ligand 5	ENSG00000271503	P13501	3
CCL11	C-C motif chemokine ligand 11	ENSG00000172156	P51671	1
CCL23	C-C motif chemokine ligand 23	ENSG00000274736	P55773	2
CCR6	C-C motif chemokine receptor 6	ENSG00000112486	P51684	2
CD40	CD40 molecule	ENSG00000101017	P25942	2
CD55	CD55 molecule (Cromer blood group)	ENSG00000196352	P08174	1
CD163	CD163 molecule	ENSG00000177575	Q86VB7	2
CHI3L1	chitinase 3 like 1	ENSG00000133048	P36222	5
CHIT1	chitinase 1	ENSG00000133063	Q13231	3
CMA1	chymase 1	ENSG00000092009	P23946	1
CPA3	carboxypeptidase A3	ENSG00000163751	P15088	2
CSF2	colony stimulating factor 2	ENSG00000164400	P04141	3
CX3CL1	C-X3-C motif chemokine ligand 1	ENSG00000006210	P78423	3
ELANE	elastase, neutrophil expressed	ENSG00000197561	P08246	2
EPX	eosinophil peroxidase	ENSG00000121053	P11678	1
FKBP5	FK506 binding protein 5	ENSG00000096060	Q13451	3
GATA3	GATA binding protein 3	ENSG00000107485	P23771	2
GZMB	granzyme B	ENSG00000100453	P10144	1
HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	ENSG00000196735	P01909	1
HSP90B1	heat shock protein 90 beta family member 1	ENSG00000166598	P14625	4
IFNG	interferon gamma	ENSG00000111537	P01579	4
IGF2	insulin like growth factor 2	ENSG00000167244	P01344	2
IL1B	interleukin 1 beta	ENSG00000125538	P01584	3
IL1R2	interleukin 1 receptor type 2	ENSG00000115590	P27930	2
IL1RL1	interleukin 1 receptor like 1	ENSG00000115602	Q01638	2
IL2RA	interleukin 2 receptor subunit alpha	ENSG00000134460	P01589	2
IL3	interleukin 3	ENSG00000164399	P08700	3
IL4	interleukin 4	ENSG00000113520	P05112	3
IL5	interleukin 5	ENSG00000113525	P05113	1
IL6	interleukin 6	ENSG00000136244	P05231	3
IL10	interleukin 10	ENSG00000136634	P22301	3
IL10RA	interleukin 10 receptor subunit alpha	ENSG00000110324	Q13651	3
IL12A	interleukin 12A	ENSG00000168811	P29459	1
IL12B	interleukin 12B	ENSG00000113302	P29460	2

IL13	interleukin 13	ENSG00000169194	P35225	2
IL17A	interleukin 17A	ENSG00000112115	Q16552	3
IL17RA	interleukin 17 receptor A	ENSG00000177663	Q96F46	3
IL17RB	interleukin 17 receptor B	ENSG00000056736	Q9NRM6	3
IL25	interleukin 25	ENSG00000166090	Q9H293	2
IL26	interleukin 26	ENSG00000111536	Q9NPH9	3
IL33	interleukin 33	ENSG00000137033	O95760	3
IRF5	interferon regulatory factor 5	ENSG00000128604	Q13568	1
IRF8	interferon regulatory factor 8	ENSG00000140968	Q02556	2
ISG15	ISG15 ubiquitin-like modifier	ENSG00000187608	P05161	1
KIT	KIT proto-oncogene receptor tyrosine kinase	ENSG00000157404	P10721	2
KITLG	KIT ligand	ENSG00000049130	P21583	2
LRRN4	leucine rich repeat neuronal 4	ENSG00000125872	Q8WUT4	2
MPO	myeloperoxidase	ENSG00000005381	P05164	2
NFKB2	nuclear factor kappa B subunit 2	ENSG00000077150	Q00653	1
NGF	nerve growth factor	ENSG00000134259	P01138	2
OLR1	oxidized low density lipoprotein receptor 1	ENSG00000173391	P78380	4
PDGFB	platelet derived growth factor subunit B	ENSG00000100311	P01127	2
PYCARD	PYD and CARD domain containing	ENSG00000103490	Q9ULZ3	1
SELE	selectin E	ENSG00000007908	P16581	4
SELP	selectin P	ENSG00000174175	P16109	2
SFTPA1/SFTPA2	surfactant protein A1/surfactant protein A2	ENSG00000122852/ ENSG00000185303	Q8IWL2/ Q8IWL1	3
SFTPB	surfactant protein B	ENSG00000168878	P07988	2
SFTPC	surfactant protein C	ENSG00000168484	P11686	1
SFTPD	surfactant protein D	ENSG00000133661	P35247	2
TGFB1	transforming growth factor beta 1	ENSG00000105329	P01137	2
TGFB3	transforming growth factor beta 3	ENSG00000119699	P10600	1
TLR2	toll like receptor 2	ENSG00000137462	O60603	4
TNF	tumor necrosis factor	ENSG00000232810	P01375	3
TNFRSF11A	TNF receptor superfamily member 11a	ENSG00000141655	Q9Y6Q6	2
TNFRSF11B	TNF receptor superfamily member 11b	ENSG00000164761	O00300	1
TNFSF10	tumor necrosis factor superfamily member 10	ENSG00000121858	P50591	4
TRAF3IP2	TRAF3 interacting protein 2	ENSG00000056972	O43734	3
TRIM33	tripartite motif containing 33	ENSG00000197323	Q9UPN9	1
TSLP	thymic stromal lymphopoietin	ENSG00000145777	Q969D9	2
VCAM1	vascular cell adhesion molecule 1	ENSG00000162692	P19320	3
Lipid metabolism/transport (phospholipid/eicosanoid/apolipo/sphingolipid)				
ACER1	alkaline ceramidase 1	ENSG00000167769	Q8TDN7	1
ACER2	alkaline ceramidase 2	ENSG00000177076	Q5QJU3	2
APOA1	apolipoprotein A1	ENSG00000118137	P02647	1
APOA4	apolipoprotein A4	ENSG00000110244	P06727	1

APOC3	apolipoprotein C3	ENSG00000110245	P02656	2
APOE	apolipoprotein E	ENSG00000130203	P02649	2
APOH	apolipoprotein H	ENSG00000091583	P02749	2
CERS1	ceramide synthase 1	ENSG00000223802	P27544	1
CERS2	ceramide synthase 2	ENSG00000143418	Q96G23	1
CERS3	ceramide synthase 3	ENSG00000154227	Q8IU89	3
CERS4	ceramide synthase 4	ENSG00000090661	Q9HA82	3
CERS5	ceramide synthase 5	ENSG00000139624	Q8N5B7	2
CERS6	ceramide synthase 6	ENSG00000172292	Q6ZMG9	2
DEGS1	delta 4-desaturase, sphingolipid 1	ENSG00000143753	O15121	3
GBA	glucosylceramidase beta	ENSG00000177628	P04062	1
GSTP1	glutathione S-transferase pi 1	ENSG00000084207	P09211	2
HMGCR	3-hydroxy-3-methylglutaryl-CoA reductase	ENSG00000113161	P04035	1
HPGDS	hematopoietic prostaglandin D synthase	ENSG00000163106	O60760	2
NAPSA	napsin A aspartic peptidase	ENSG00000131400	O96009	3
ORM1/ORM2	orosomucoid 1/orosomucoid 2	ENSG00000229314/ ENSG00000228278	P02763/ P19652	2
PLPP1	phospholipid phosphatase 1	ENSG00000067113	O14494	1
PTGS1	prostaglandin-endoperoxide synthase 1	ENSG00000095303	P23219	1
PTGS2	prostaglandin-endoperoxide synthase 2	ENSG00000073756	P35354	1
SGMS1	sphingomyelin synthase 1	ENSG00000198964		2
SGMS2	sphingomyelin synthase 2	ENSG00000164023	Q8NHU3	2
SGPL1	sphingosine-1-phosphate lyase 1	ENSG00000166224	O95470	1
SMPD1	sphingomyelin phosphodiesterase 1	ENSG00000166311	P17405	1
SMPD2	sphingomyelin phosphodiesterase 2	ENSG00000135587	O60906	1
SMPD3	sphingomyelin phosphodiesterase 3	ENSG00000103056	Q9NY59	4
SPHK1	sphingosine kinase 1	ENSG00000176170	Q9NYA1	4
SPHK2	sphingosine kinase 2	ENSG00000063176	Q9NRA0	3
SPTLC3	serine palmitoyltransferase long chain base subunit 3	ENSG00000172296	Q9NUV7	3
UGCG	UDP-glucose ceramide glucosyltransferase	ENSG00000148154	Q16739	2
UGT8	UDP glycosyltransferase 8	ENSG00000174607	Q16880	2
Signal transduction				
ADCY2	adenylate cyclase 2	ENSG00000078295	Q08462	2
DLG2	discs large MAGUK scaffold protein 2	ENSG00000150672	Q15700	3
KCNB2	potassium voltage-gated channel subfamily B member 2	ENSG00000182674	Q92953	1
MAGI1	membrane associated guanylate kinase, WW and PDZ domain containing 1	ENSG00000151276	Q96QZ7	3

MAP2K3	mitogen-activated protein kinase kinase 3	ENSG00000034152	P46734	1
MAP2K3/MAP2K6	mitogen-activated protein kinase kinase 3/mitogen-activated protein kinase kinase 6	ENSG00000034152/ ENSG00000108984	P46734/ P52564	1
MS4A15	membrane spanning 4-domains A15	ENSG00000166961	Q8N5U1	2
NCOA2	nuclear receptor coactivator 2	ENSG00000140396	Q15596	2
NOS2	nitric oxide synthase 2	ENSG00000007171	P35228	3
NPSR1	neuropeptide S receptor 1	ENSG00000187258	Q6W5P4	4
NR3C1	nuclear receptor subfamily 3 group C member 1	ENSG00000113580	P04150	1
RAB31	RAB31, member RAS oncogene family	ENSG00000168461	Q13636	1
RASD2	RASD family member 2	ENSG00000100302	Q96D21	1
RGS18	regulator of G-protein signaling 18	ENSG00000150681	Q9NS28	4
ROS1	ROS proto-oncogene 1, receptor tyrosine kinase	ENSG00000047936	P08922	2
RTKN2	rhotekin 2	ENSG00000182010	Q8IZC4	3
S100A12	S100 calcium binding protein A12	ENSG00000163221	P80511	2
TAS2R3	taste 2 receptor member 3	ENSG00000127362	Q9NYW6	4
TAS2R10	taste 2 receptor member 10	ENSG00000121318	Q9NYW0	3
TAS2R14	taste 2 receptor member 14	ENSG00000212127	Q9NYV8	4
TAS2R38	taste 2 receptor member 38	ENSG00000257138	P59533	2
ZNF688	zinc finger protein 688	ENSG00000229809	P0C7X2	1
Extracellular matrix				
ANXA2	annexin A2	ENSG00000182718	P07355	2
ECM1	extracellular matrix protein 1	ENSG00000143369	Q16610	1
KRT1	keratin 1	ENSG00000167768	P04264	1
MGP	matrix Gla protein	ENSG00000111341	P08493	1
MMP1	matrix metallopeptidase 1	ENSG00000196611	P03956	4
MMP7	matrix metallopeptidase 7	ENSG00000137673	P09237	3
MMP9	matrix metallopeptidase 9	ENSG00000100985	P14780	2
MMP10	matrix metallopeptidase 10	ENSG00000166670	P09238	2
POSTN	periostin	ENSG00000133110	Q15063	1
SERPINA1	serpin family A member 1	ENSG00000197249	P01009	3
SERPINA3	serpin family A member 3	ENSG00000196136	P01011	1
SPARCL1	SPARC like 1	ENSG00000152583	Q14515	2
SPP1	secreted phosphoprotein 1	ENSG00000118785	P10451	2
SPRR3	small proline rich protein 3	ENSG00000163209	Q9UBC9	2
TIMP1	TIMP metallopeptidase inhibitor 1	ENSG00000102265	P01033	1
Energy/metabolic				
ATP5A1	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit 1, cardiac muscle	ENSG00000152234	P25705	1
B4GALT5	beta-1,4-galactosyltransferase 5	ENSG00000158470	O43286	3
B4GALT6	beta-1,4-galactosyltransferase 6	ENSG00000118276	Q9UBX8	2
FETUB	fetuin B	ENSG00000090512	Q9UGM5	3

INS	insulin	ENSG00000254647	P01308	1
LEP	leptin	ENSG00000174697	P41159	4
Coagulation				
A2M	alpha-2-macroglobulin	ENSG00000175899	P01023	1
F7	coagulation factor VII	ENSG00000057593	P08709	2
FGA	fibrinogen alpha chain	ENSG00000171560	P02671	2
FGL2	fibrinogen like 2	ENSG00000127951	Q14314	2
HRG	histidine rich glycoprotein	ENSG00000113905	P04196	2
LPA	lipoprotein(a)	ENSG00000198670	P08519	3
Complement system				
C1QB	complement C1q B chain	ENSG00000173369	P02746	1
C4A/C4B	complement C4A (Rodgers blood group)/complement C4B (Chido blood group)	ENSG00000244731/ ENSG00000224389	P0C0L4/ P0C0L5	1
C8G	complement C8 gamma chain	ENSG00000176919	P07360	2
C9	complement C9	ENSG00000113600	P02748	2
CFI	complement factor I	ENSG00000205403	P05156	2
Others				
C1orf195	chromosome 1 open reading frame 195	ENSG00000204464		2
CRISP3	cysteine rich secretory protein 3	ENSG00000096006	P54108	2
GAP43	growth associated protein 43	ENSG00000172020	P17677	2
MDC1	mediator of DNA damage checkpoint 1	ENSG00000137337	Q14676	3
MOCOS	molybdenum cofactor sulfurase	ENSG00000075643	Q96EN8	4
MRPL43	mitochondrial ribosomal protein L43	ENSG00000055950	Q8N983	4
PSORS1C1	psoriasis susceptibility 1 candidate 1	ENSG00000204540	Q9UIG5	2
PSORS1C2	psoriasis susceptibility 1 candidate 2	ENSG00000204538	Q9UIG4	3
RETNLB	resistin like beta	ENSG00000163515	Q9BQ08	1
RNASE3	ribonuclease A family member 3	ENSG00000169397	P12724	1
SCGB1A1	secretoglobin family 1A member 1	ENSG00000149021	P11684	1
SLC11A1	solute carrier family 11 member 1	ENSG00000018280	P49279	2
SLC22A2	solute carrier family 22 member 2	ENSG00000112499	O15244	2
SOD2	superoxide dismutase 2	ENSG00000112096	P04179	1
TBX21	T-box 21	ENSG00000073861	Q9UL17	3
TNNI3	troponin I3, cardiac type	ENSG00000129991	P19429	2
TTR	transthyretin	ENSG00000118271	P02766	1

* Corresponding protein name can be found at https://www.uniprot.org/

Table S3. The protein panel included 177 proteins. In the antibody bead array, two or more antibodies for each protein target were included for the majority of the proteins, resulting in 377 antibodies.

Number of antibodies per protein	Number of proteins	Number of antibodies
1	52	52
2	68	136
3	40	120
4	16	64
5	1	5
Total:	**177**	**377**

Table S4. Plasma protein group comparisons in the U-BIOPRED cohort. 110 proteins showed significantly different plasma levels in multiple group comparisons (adjusted Kruskal-Wallis test*). Pairwise group comparisons (adjusted Wilcoxon rank-sum test†) as well as directionality, indicated with (+) or (-), meaning higher or lower in the first group mentioned, are also reported. Table sorted by the adjusted Kruskal-Wallis *P* values in increasing order.

			***P* value†**											
Gene‡	**Antibody**	***P* value***	**SAn *vs.* SAs/ex**		**SAn *vs.* MMA**		**SAn *vs.* HC**		**SAs/ex *vs.* MMA**		**SAs/ex *vs.* HC**		**MMA *vs.* HC**	
SERPINA3	HPA000893	9.8E-23	1.0E+00	(-)	1.5E-11	(+)	3.0E-15	(+)	8.4E-08	(+)	8.9E-10	(+)	9.8E-01	(+)
C9	HPA029577	2.8E-15	1.0E+00	(-)	3.1E-08	(+)	1.7E-10	(+)	7.7E-06	(+)	8.0E-07	(+)	9.8E-01	(+)
C9	HPA070709	9.4E-15	1.0E+00	(+)	4.4E-08	(+)	3.1E-10	(+)	9.9E-06	(+)	8.0E-07	(+)	9.8E-01	(+)
SFTPD	HPA044582	1.9E-14	1.0E+00	(-)	4.4E-08	(+)	1.0E-09	(+)	6.1E-06	(+)	1.2E-06	(+)	9.9E-01	(-)
TNF	HPA055037	9.1E-14	1.0E+00	(-)	3.1E-08	(+)	4.1E-09	(+)	6.1E-06	(+)	9.3E-06	(+)	9.8E-01	(-)
SPARCL1	HPA067641	2.0E-13	1.0E+00	(+)	4.4E-08	(+)	2.5E-09	(+)	1.9E-05	(+)	9.8E-06	(+)	9.8E-01	(-)
CX3CL1	HPA056729	4.3E-13	1.0E+00	(-)	9.7E-07	(+)	1.4E-09	(+)	2.6E-05	(+)	8.0E-07	(+)	9.8E-01	(+)
IL2RA	HPA046738	6.5E-13	1.0E+00	(+)	6.2E-08	(+)	1.8E-08	(+)	1.1E-05	(+)	5.7E-06	(+)	9.8E-01	(+)
CFI	HPA024061	1.1E-12	1.0E+00	(-)	1.9E-06	(+)	4.2E-09	(+)	2.2E-05	(+)	8.0E-07	(+)	9.8E-01	(+)
C8G	HPA046269	1.1E-12	1.0E+00	(-)	2.4E-07	(+)	8.9E-09	(+)	2.2E-05	(+)	6.5E-06	(+)	9.8E-01	(+)
FGA	HPA064755	1.1E-12	1.0E+00	(+)	4.4E-08	(+)	1.4E-07	(+)	2.5E-06	(+)	1.2E-05	(+)	9.8E-01	(-)
CFI	HPA001143	1.7E-12	1.0E+00	(-)	9.7E-07	(+)	1.7E-08	(+)	1.4E-05	(+)	1.4E-06	(+)	9.8E-01	(+)
MAGI1	HPA031852	2.4E-12	1.0E+00	(-)	6.9E-08	(+)	1.7E-08	(+)	8.4E-05	(+)	3.7E-05	(+)	9.8E-01	(+)
CCL23	HPA063758	5.0E-12	1.0E+00	(+)	1.2E-07	(+)	1.2E-07	(+)	1.4E-05	(+)	1.4E-05	(+)	9.8E-01	(+)
FGL2	HPA026682	5.1E-12	1.0E+00	(-)	6.8E-07	(+)	1.1E-08	(+)	8.4E-05	(+)	9.8E-06	(+)	9.8E-01	(+)
IL1R2	HPA027598	7.8E-12	1.0E+00	(-)	3.3E-07	(+)	3.6E-08	(+)	4.8E-05	(+)	1.4E-05	(+)	9.8E-01	(+)
PTGS1	HPA002834	7.9E-12	1.0E+00	(+)	4.6E-07	(+)	2.3E-08	(+)	7.1E-05	(+)	1.4E-05	(+)	9.9E-01	(+)
CERS4	HPA023621	8.0E-12	1.0E+00	(-)	5.6E-07	(+)	1.8E-08	(+)	6.2E-05	(+)	1.8E-05	(+)	9.8E-01	(+)
SELE	HPA057891	9.4E-12	1.0E+00	(+)	2.4E-07	(+)	2.9E-08	(+)	8.4E-05	(+)	5.5E-05	(+)	9.8E-01	(+)
SCGB1A1	HPA031828	1.1E-11	1.0E+00	(-)	3.3E-07	(+)	9.9E-08	(+)	2.2E-05	(+)	1.8E-05	(+)	9.8E-01	(+)
APOE	HPA068768	1.5E-11	1.0E+00	(-)	4.4E-08	(+)	4.8E-05	(+)	1.9E-07	(+)	5.5E-05	(+)	9.8E-01	(-)
LEP	HPA068565	1.6E-11	1.0E+00	(+)	3.5E-06	(+)	1.1E-09	(+)	1.7E-02	(+)	2.7E-04	(+)	9.8E-01	(+)
CSF2	HPA071579	1.8E-11	1.0E+00	(-)	4.8E-07	(+)	7.8E-07	(+)	5.7E-06	(+)	1.1E-05	(+)	9.8E-01	(+)
CD163	HPA046404	1.8E-10	1.0E+00	(-)	5.0E-07	(+)	7.5E-07	(+)	5.8E-05	(+)	9.5E-05	(+)	9.8E-01	(+)
PDGFB	HPA011972	5.2E-10	1.0E+00	(+)	3.3E-07	(+)	1.4E-06	(+)	1.3E-04	(+)	4.4E-04	(+)	9.8E-01	(-)
MRPL43	HPA047247	6.1E-10	1.0E+00	(-)	1.9E-06	(+)	8.3E-07	(+)	9.5E-05	(+)	7.0E-05	(+)	9.8E-01	(+)
C8G	HPA073328	6.1E-10	1.0E+00	(+)	1.8E-06	(+)	3.3E-07	(+)	2.6E-04	(+)	1.2E-04	(+)	9.9E-01	(+)
RASD2	HPA005839	6.3E-10	1.0E+00	(-)	1.0E-06	(+)	1.7E-06	(+)	5.8E-05	(+)	1.0E-04	(+)	9.8E-01	(-)
ARFGAP1	HPA047382	7.7E-10	1.0E+00	(-)	3.6E-06	(+)	7.8E-07	(+)	9.5E-05	(+)	5.0E-05	(+)	1.0E+00	(-)
CPA3	HPA006664	8.9E-10	1.0E+00	(-)	2.8E-06	(+)	1.3E-06	(+)	8.4E-05	(+)	5.5E-05	(+)	1.0E+00	(+)
TBX21	HPA046626	9.9E-10	1.0E+00	(+)	5.9E-06	(+)	1.4E-07	(+)	7.3E-04	(+)	1.2E-04	(+)	9.8E-01	(+)
MS4A15	HPA054563	9.9E-10	1.0E+00	(+)	1.1E-05	(+)	9.8E-08	(+)	1.4E-03	(+)	7.4E-05	(+)	9.8E-01	(+)
ARFGAP1	HPA056273	1.1E-09	1.0E+00	(+)	9.7E-07	(+)	1.7E-06	(+)	1.2E-04	(+)	3.1E-04	(+)	9.8E-01	(-)
TGFB3	HPA063582	1.1E-09	1.0E+00	(-)	3.4E-06	(+)	6.3E-07	(+)	2.6E-04	(+)	9.5E-05	(+)	9.8E-01	(+)
HPGDS	MAB6487	1.3E-09	1.0E+00	(+)	1.1E-05	(+)	3.1E-07	(+)	3.6E-04	(+)	5.2E-05	(+)	9.8E-01	(+)
TRIM33	HPA004345	1.9E-09	1.0E+00	(+)	1.3E-06	(+)	1.9E-06	(+)	1.9E-04	(+)	3.3E-04	(+)	9.8E-01	(+)
PPAP2A	HPA047815	2.0E-09	1.0E+00	(+)	1.0E-06	(+)	1.4E-06	(+)	3.7E-04	(+)	6.2E-04	(+)	9.8E-01	(+)
IL10	HPA027476	2.2E-09	1.0E+00	(-)	1.1E-06	(+)	3.8E-06	(+)	1.5E-04	(+)	3.4E-04	(+)	9.8E-01	(-)
C4A/C4B	HPA046356	2.8E-09	1.0E+00	(-)	2.0E-03	(+)	6.4E-07	(+)	1.9E-04	(+)	2.5E-06	(+)	9.8E-01	(+)
CERS2	HPA027262	4.6E-09	1.0E+00	(+)	1.3E-05	(+)	1.6E-06	(+)	2.3E-04	(+)	9.5E-05	(+)	9.8E-01	(+)
SPRR3	HPA044467	6.3E-09	1.0E+00	(+)	2.0E-06	(+)	8.6E-06	(+)	1.5E-04	(+)	4.2E-04	(+)	9.8E-01	(-)
CHI3L1	AF2599	9.2E-09	1.0E+00	(+)	7.1E-06	(+)	7.8E-07	(+)	2.2E-03	(+)	1.0E-03	(+)	9.8E-01	(-)
APOE	HPA065539	9.2E-09	1.0E+00	(-)	3.2E-06	(+)	1.4E-03	(+)	2.5E-06	(+)	1.5E-04	(+)	9.8E-01	(-)
FGA	HPA051370	1.3E-08	1.0E+00	(+)	1.7E-05	(+)	3.8E-06	(+)	2.7E-04	(+)	1.0E-04	(+)	9.8E-01	(+)
IL4	HPA007714	2.0E-08	1.0E+00	(+)	2.7E-06	(+)	1.5E-05	(+)	2.3E-04	(+)	1.1E-03	(+)	9.8E-01	(+)
APOA4	HPA001352	2.5E-08	1.0E+00	(-)	2.1E-04	(+)	6.5E-05	(+)	2.2E-05	(+)	1.4E-05	(+)	1.0E+00	(-)
FGL2	HPA019229	3.5E-08	1.0E+00	(-)	2.8E-05	(+)	3.2E-05	(+)	9.5E-05	(+)	1.0E-04	(+)	9.8E-01	(+)
TAS2R14	HPA042371	3.7E-08	1.0E+00	(+)	1.7E-06	(+)	1.5E-05	(+)	1.8E-03	(+)	5.8E-03	(+)	9.8E-01	(-)
SMPD2	HPA018125	1.1E-07	1.0E+00	(+)	7.7E-07	(+)	2.1E-03	(+)	2.2E-05	(+)	9.5E-03	(+)	9.8E-01	(-)
SPP1	HPA005562	1.2E-07	1.0E+00	(+)	8.5E-06	(+)	5.4E-05	(+)	3.3E-04	(+)	1.4E-03	(+)	9.8E-01	(+)
ISG15	HPA004627	1.2E-07	1.0E+00	(-)	1.6E-05	(+)	6.2E-05	(+)	1.9E-04	(+)	6.6E-04	(+)	9.8E-01	(-)
VCAM1	HPA001618	2.1E-07	1.0E+00	(+)	6.2E-06	(+)	1.1E-04	(+)	2.8E-04	(+)	3.7E-03	(+)	9.8E-01	(+)
SGMS2	HPA015076	2.1E-07	1.0E+00	(+)	4.0E-06	(+)	9.3E-05	(+)	6.8E-04	(+)	4.7E-03	(+)	9.8E-01	(+)
IL17A	HPA052258	2.3E-07	1.0E+00	(-)	2.5E-04	(+)	1.4E-05	(+)	8.8E-04	(+)	1.5E-04	(+)	9.8E-01	(+)
TRAF3IP2	HPA036352	2.6E-07	1.0E+00	(+)	7.1E-06	(+)	7.4E-05	(+)	7.6E-04	(+)	4.6E-03	(+)	9.8E-01	(+)
SPRR3	HPA024330	2.6E-07	1.0E+00	(+)	1.7E-05	(+)	7.5E-05	(+)	4.0E-04	(+)	1.6E-03	(+)	9.8E-01	(-)
IL2RA	HPA054622	3.7E-07	1.0E+00	(-)	1.1E-05	(+)	9.0E-05	(+)	6.8E-04	(+)	4.9E-03	(+)	9.8E-01	(-)
IL17RB	HPA002837	6.5E-07	1.0E+00	(+)	8.1E-05	(+)	1.8E-05	(+)	3.9E-03	(+)	2.1E-03	(+)	9.8E-01	(+)

			P value†											
Gene‡	Antibody	*P* value*	SAn *vs.* SAs/ex		SAn *vs.* MMA		SAn *vs.* HC		SAs/ex *vs.* MMA		SAs/ex *vs.* HC		MMA *vs.* HC	
SFTPA1/SFTPA2	HPA045752	6.6E-07	1.0E+00	(-)	5.1E-06	(+)	6.4E-04	(+)	3.6E-04	(+)	6.5E-03	(+)	9.8E-01	(-)
APOH	HPA003732	7.4E-07	8.7E-01	(-)	2.5E-04	(+)	4.5E-02	(+)	2.5E-06	(+)	2.9E-04	(+)	9.8E-01	(-)
ATP5A1	HPA040622	1.2E-06	1.0E+00	(+)	3.3E-05	(+)	7.8E-05	(+)	2.4E-03	(+)	5.4E-03	(+)	9.8E-01	(+)
MGP	HPA013949	1.4E-06	9.6E-01	(-)	2.2E-04	(+)	1.3E-02	(+)	1.4E-05	(+)	6.0E-04	(+)	9.8E-01	(-)
LEP	HPA030721	1.5E-06	1.0E+00	(+)	2.1E-05	(+)	1.6E-04	(+)	2.4E-03	(+)	8.0E-03	(+)	9.8E-01	(-)
UGCG	HPA050554	1.5E-06	1.0E+00	(-)	6.8E-04	(+)	9.7E-06	(+)	9.8E-03	(+)	9.8E-04	(+)	9.8E-01	(+)
TBX21	HPA028935	1.5E-06	1.0E+00	(+)	1.2E-05	(+)	3.4E-04	(+)	1.4E-03	(+)	1.2E-02	(+)	9.8E-01	(+)
MMP1	HPA031456	2.6E-06	1.0E+00	(+)	1.5E-03	(+)	2.5E-05	(+)	6.2E-03	(+)	3.4E-04	(+)	9.8E-01	(+)
TGFB1	HPA047516	3.9E-06	1.0E+00	(-)	3.9E-03	(+)	6.4E-04	(+)	2.4E-04	(+)	9.5E-05	(+)	9.8E-01	(+)
INS	HPA004932	4.1E-06	1.0E+00	(-)	9.6E-05	(+)	1.3E-03	(+)	3.0E-04	(+)	2.1E-03	(+)	9.8E-01	(+)
LEP	HPA057322	4.3E-06	1.0E+00	(+)	1.9E-03	(+)	8.6E-06	(+)	4.3E-02	(+)	1.9E-03	(+)	9.8E-01	(+)
TNNI3	HPA046428	4.5E-06	8.7E-01	(+)	2.1E-02	(-)	3.9E-03	(-)	1.9E-04	(-)	3.7E-05	(-)	9.8E-01	(-)
F7	HPA004826	4.7E-06	1.0E+00	(+)	2.0E-05	(+)	2.7E-03	(+)	4.5E-04	(+)	1.2E-02	(+)	9.8E-01	(-)
SPHK1	HPA022829	5.7E-06	1.0E+00	(+)	6.6E-05	(+)	3.4E-04	(+)	2.8E-03	(+)	1.1E-02	(+)	9.8E-01	(-)
IL10RA	HPA065647	6.8E-06	1.0E+00	(-)	2.2E-05	(+)	7.9E-03	(+)	1.7E-04	(+)	1.4E-02	(+)	9.8E-01	(-)
MDC1	HPA006795	6.9E-06	1.0E+00	(+)	1.9E-05	(+)	8.5E-04	(+)	4.0E-03	(+)	4.4E-02	(+)	9.8E-01	(-)
SLC22A2	HPA008549	1.7E-05	1.0E+00	(-)	3.6E-04	(+)	5.5E-03	(+)	3.1E-04	(+)	2.8E-03	(+)	9.8E-01	(-)
SPARCL1	HPA067587	1.9E-05	1.0E+00	(+)	1.8E-05	(+)	5.5E-03	(+)	1.8E-03	(+)	8.1E-02	(+)	9.8E-01	(-)
SERPINA1	HPA001291	2.5E-05	1.0E+00	(-)	5.0E-03	(+)	1.6E-03	(+)	6.3E-04	(+)	2.7E-04	(+)	9.9E-01	(-)
SERPINA1	HPA000927	3.2E-05	1.0E+00	(-)	5.1E-03	(+)	2.6E-03	(+)	5.4E-04	(+)	4.4E-04	(+)	1.0E+00	(+)
SMPD3	HPA069383	3.2E-05	1.0E+00	(-)	7.3E-04	(+)	6.0E-03	(+)	3.1E-04	(+)	2.3E-03	(+)	9.8E-01	(+)
IL6	HPA044648	3.5E-05	1.0E+00	(-)	5.2E-04	(+)	1.3E-02	(+)	1.2E-04	(+)	4.6E-03	(+)	9.8E-01	(+)
SELP	HPA005990	6.5E-05	1.0E+00	(+)	3.7E-04	(+)	2.4E-03	(+)	3.1E-03	(+)	1.3E-02	(+)	9.8E-01	(-)
C1QB	HPA052116	1.9E-04	1.0E+00	(+)	3.3E-03	(+)	5.0E-04	(+)	2.9E-02	(+)	1.1E-02	(+)	9.8E-01	(+)
HSP90B1	HPA008424	2.0E-04	1.0E+00	(-)	7.7E-03	(+)	1.5E-02	(+)	8.0E-04	(+)	1.3E-03	(+)	9.8E-01	(-)
B4GALT5	HPA067597	2.0E-04	1.0E+00	(+)	5.9E-02	(-)	6.7E-03	(-)	3.1E-03	(-)	1.9E-04	(-)	9.8E-01	(-)
C1orf195	HPA045811	2.4E-04	1.0E+00	(+)	2.7E-02	(+)	1.0E-04	(+)	1.4E-01	(+)	5.9E-03	(+)	9.8E-01	(+)
CHIT1	HPA010115	2.7E-04	8.7E-01	(-)	1.5E-01	(+)	1.5E-02	(+)	4.8E-03	(+)	1.2E-04	(+)	9.8E-01	(-)
GAP43	HPA013603	3.5E-04	1.0E+00	(+)	3.7E-04	(+)	7.4E-03	(+)	1.4E-02	(+)	1.2E-01	(+)	9.8E-01	(-)
TNFRSF11A	HPA027728	3.5E-04	1.0E+00	(+)	1.9E-03	(+)	4.9E-03	(+)	4.7E-03	(+)	1.7E-02	(+)	9.8E-01	(+)
IL4	HPA070010	3.6E-04	1.0E+00	(+)	5.3E-04	(+)	4.9E-03	(+)	1.6E-02	(+)	9.3E-02	(+)	9.8E-01	(-)
SELE	HPA006225	4.3E-04	1.0E+00	(-)	2.5E-04	(+)	2.9E-02	(+)	3.3E-03	(+)	1.1E-01	(+)	9.8E-01	(-)
SMPD3	HPA044442	4.3E-04	1.0E+00	(+)	1.7E-03	(+)	6.5E-03	(+)	6.2E-03	(+)	1.7E-02	(+)	9.8E-01	(+)
HMGCR	HPA008338	5.8E-04	1.0E+00	(-)	1.9E-03	(+)	7.7E-03	(+)	7.8E-03	(+)	2.0E-02	(+)	9.8E-01	(-)
HPGDS	HPA024035	6.2E-04	1.0E+00	(-)	4.1E-02	(+)	1.6E-04	(+)	2.1E-01	(+)	1.4E-02	(+)	9.8E-01	(+)
CD40	HPA031566	7.4E-04	1.0E+00	(+)	8.0E-04	(+)	5.5E-03	(+)	3.7E-02	(+)	1.4E-01	(+)	9.8E-01	(+)
GATA3	HPA029731	7.7E-04	1.0E+00	(+)	1.1E-03	(+)	1.2E-02	(+)	9.4E-03	(+)	6.2E-02	(+)	9.8E-01	(-)
SERPINA1	HPA001292	7.7E-04	1.0E+00	(-)	3.5E-02	(+)	6.7E-03	(+)	7.5E-03	(+)	1.3E-03	(+)	9.8E-01	(+)
HSP90B1	HPA003901	1.2E-03	1.0E+00	(-)	2.4E-02	(+)	3.8E-02	(+)	1.6E-03	(+)	3.9E-03	(+)	9.8E-01	(-)
TLR2	HPA051188	1.2E-03	1.0E+00	(+)	3.7E-04	(+)	3.8E-02	(+)	2.7E-02	(+)	3.4E-01	(+)	9.8E-01	(-)
S100A12	HPA002881	1.5E-03	1.0E+00	(+)	2.8E-01	(-)	5.0E-04	(-)	2.8E-01	(-)	1.5E-03	(-)	9.8E-01	(-)
KCNB2	HPA061960	1.7E-03	1.0E+00	(-)	2.9E-03	(+)	7.4E-03	(+)	3.2E-02	(+)	7.2E-02	(+)	9.8E-01	(-)
MRPL43	HPA065191	1.7E-03	1.0E+00	(-)	8.0E-04	(+)	8.8E-02	(+)	3.6E-03	(+)	1.3E-01	(+)	9.8E-01	(-)
CCR6	HPA014510	1.7E-03	1.0E+00	(-)	1.7E-03	(+)	2.6E-02	(+)	9.8E-03	(+)	7.5E-02	(+)	9.8E-01	(-)
CERS6	HPA063527	2.9E-03	1.0E+00	(+)	1.3E-03	(+)	3.6E-02	(+)	2.8E-02	(+)	2.5E-01	(+)	9.8E-01	(-)
SOD2	HPA001814	3.0E-03	1.0E+00	(-)	1.1E-02	(+)	5.7E-02	(+)	4.2E-03	(+)	1.8E-02	(+)	9.8E-01	(-)
SGPL1	HPA021125	4.1E-03	1.0E+00	(+)	2.2E-02	(+)	3.7E-03	(+)	1.2E-01	(+)	3.3E-02	(+)	9.8E-01	(+)
CHIT1	HPA010575	6.1E-03	1.0E+00	(-)	1.7E-01	(+)	3.4E-02	(+)	1.7E-02	(+)	3.1E-03	(+)	9.8E-01	(+)
CERS1	HPA045724	6.1E-03	1.0E+00	(-)	5.1E-02	(-)	1.9E-03	(-)	4.3E-01	(-)	8.3E-02	(-)	9.8E-01	(-)
CERS4	HPA049826	6.2E-03	1.0E+00	(+)	2.9E-03	(+)	4.2E-02	(+)	3.6E-02	(+)	1.7E-01	(+)	9.8E-01	(+)
CD55	HPA002190	6.3E-03	1.0E+00	(-)	4.4E-02	(+)	2.8E-02	(+)	1.5E-02	(+)	1.0E-02	(+)	9.8E-01	(-)
CCL5	HPA010552	7.6E-03	1.0E+00	(+)	3.6E-01	(+)	6.4E-04	(+)	9.4E-01	(+)	1.1E-01	(+)	9.8E-01	(+)
APOC3	HPA056395	7.9E-03	1.0E+00	(+)	1.4E-02	(-)	6.0E-02	(-)	8.1E-03	(-)	4.5E-02	(-)	9.8E-01	(-)
MAP2K3/MAP2K6	HPA067854	9.1E-03	1.0E+00	(-)	3.0E-02	(+)	6.0E-02	(+)	8.3E-03	(+)	2.2E-02	(+)	9.8E-01	(-)
NPSR1	HPA007048	9.3E-03	1.0E+00	(-)	2.4E-03	(+)	2.1E-01	(+)	7.8E-03	(+)	2.8E-01	(+)	9.8E-01	(-)
SLC22A2	HPA008567	1.1E-02	1.0E+00	(+)	2.9E-03	(+)	7.3E-02	(+)	8.0E-02	(+)	4.6E-01	(+)	9.8E-01	(-)
CERS5	HPA006780	1.2E-02	1.0E+00	(-)	6.5E-02	(+)	3.1E-02	(+)	2.9E-02	(+)	1.5E-02	(+)	9.8E-01	(+)
IL33	HPA022899	1.2E-02	1.0E+00	(+)	1.2E-02	(-)	5.2E-02	(-)	2.0E-02	(-)	9.3E-02	(-)	9.8E-01	(+)
GSTP1	HPA019779	1.6E-02	1.0E+00	(-)	4.7E-02	(+)	3.8E-02	(+)	3.2E-02	(+)	2.5E-02	(+)	9.8E-01	(+)
IL17A	HPA045886	1.6E-02	1.0E+00	(-)	2.9E-03	(+)	2.7E-01	(+)	2.0E-02	(+)	3.6E-01	(+)	9.8E-01	(-)
NCOA2	HPA069172	1.6E-02	1.0E+00	(-)	6.7E-02	(+)	6.8E-03	(+)	1.7E-01	(+)	4.9E-02	(+)	9.8E-01	(+)
SMPD1	HPA001823	1.8E-02	1.0E+00	(-)	1.4E-02	(-)	2.7E-02	(-)	2.5E-01	(-)	4.0E-01	(-)	9.8E-01	(+)
CX3CL1	HPA040361	1.9E-02	1.0E+00	(-)	8.4E-03	(+)	2.4E-01	(+)	8.1E-03	(+)	2.4E-01	(+)	9.8E-01	(-)
OLR1	HPA050798	2.0E-02	8.7E-01	(+)	8.7E-02	(+)	6.7E-02	(+)	4.8E-01	(-)	7.4E-01	(+)	9.8E-01	(+)

Gene‡	Antibody	*P* value*	*P* value†											
			SAn *vs.* SAs/ex		SAn *vs.* MMA		SAn *vs.* HC		SAs/ex *vs.* MMA		SAs/ex *vs.* HC		MMA *vs.* HC	
KRT1	HPA019797	2.1E-02	1.0E+00	(+)	2.2E-01	(+)	1.1E-02	(+)	2.0E-01	(+)	1.1E-02	(+)	9.8E-01	(+)
FETUB	HPA035132	2.3E-02	1.0E+00	(+)	2.3E-01	(+)	4.5E-03	(+)	8.7E-01	(+)	1.9E-01	(+)	9.8E-01	(+)
NOS2	HPA038086	2.4E-02	1.0E+00	(-)	1.1E-02	(-)	6.4E-02	(-)	7.6E-02	(-)	2.2E-01	(-)	9.8E-01	(+)
TAS2R38	HPA054366	2.6E-02	1.0E+00	(+)	6.8E-02	(+)	1.3E-02	(+)	2.0E-01	(+)	7.5E-02	(+)	9.8E-01	(+)
IL26	HPA055164	2.7E-02	1.0E+00	(-)	3.9E-02	(-)	1.5E-02	(-)	4.4E-01	(-)	2.9E-01	(-)	9.8E-01	(-)
SMPD3	HPA065535	2.8E-02	1.0E+00	(-)	1.1E-02	(+)	1.0E-01	(+)	5.9E-02	(+)	3.3E-01	(+)	9.8E-01	(-)
KIT	HPA004471	3.2E-02	1.0E+00	(-)	8.3E-02	(-)	1.0E-02	(-)	3.1E-01	(-)	1.1E-01	(-)	9.8E-01	(-)
APOA1	HPA046715	3.2E-02	1.0E+00	(+)	3.8E-01	(+)	4.5E-03	(+)	8.3E-01	(+)	8.1E-02	(+)	9.8E-01	(+)
NAPSA	HPA047744	3.2E-02	1.0E+00	(-)	2.0E-02	(-)	3.9E-02	(-)	1.8E-01	(-)	3.2E-01	(-)	9.8E-01	(+)
SPHK2	HPA049062	3.3E-02	1.0E+00	(+)	2.6E-01	(-)	5.7E-02	(-)	5.8E-02	(-)	1.4E-02	(-)	9.8E-01	(-)
MMP7	HPA051358	3.4E-02	8.7E-01	(+)	4.7E-01	(-)	6.6E-01	(-)	9.8E-03	(-)	2.8E-02	(-)	9.8E-01	(+)
AGER	HPA064436	3.6E-02	1.0E+00	(-)	3.0E-02	(+)	7.5E-01	(-)	7.2E-03	(+)	6.8E-01	(+)	9.8E-01	(-)
IL13	HPA042421	3.9E-02	1.0E+00	(-)	2.4E-02	(+)	1.0E-01	(+)	3.7E-02	(+)	1.5E-01	(+)	9.8E-01	(+)
TRAF3IP2	HPA049742	4.0E-02	1.0E+00	(+)	2.8E-02	(+)	4.8E-02	(+)	1.4E-01	(+)	2.1E-01	(+)	9.8E-01	(+)
SGMS1	HPA045191	4.4E-02	1.0E+00	(-)	1.1E-02	(+)	4.6E-01	(+)	1.7E-02	(+)	4.2E-01	(+)	9.8E-01	(-)
CERS4	HPA070214	4.4E-02	1.0E+00	(+)	1.2E-02	(+)	1.9E-01	(+)	5.9E-02	(+)	4.2E-01	(+)	9.8E-01	(-)
MPO	HPA061464	4.7E-02	1.0E+00	(+)	1.6E-02	(+)	2.5E-01	(+)	4.4E-01	(+)	8.2E-01	(+)	9.8E-01	(+)

Definition of abbreviations: HC = healthy non-smoking controls; MMA = non-smokers with mild-to-moderate asthma; SAn = non-smokers with severe asthma; SAs/ex = smokers or ex-smokers with severe asthma.

*Kruskal-Wallis test, adjusted
†Wilcoxon rank-sum test, adjusted
‡ Corresponding protein name can be found at https://www.uniprot.org/

Table S5. Group comparisons in the BIOAIR cohort. Proteins (*n*=23) and the corresponding antibodies (*n*=25) identified to show significantly different plasma levels in multiple group comparisons in the BIOAIR cohort (adjusted Kruskal-Wallis test*). Pairwise group comparisons (adjusted Wilcoxon rank-sum test†) as well as directionality, indicated with (+) or (-), meaning higher or lower in the first group mentioned, are also reported. Table sorted by the adjusted Kruskal-Wallis *P* values in increasing order.

Gene‡	Antibody	*P* value*	*P* value†		
			SAn *vs*. MMA	COPD *vs*. MMA	SAn *vs*. COPD
C9	HPA070709	2.0E-05	2.7E-02 (+)	3.5E-06 (+)	8.2E-02 (-)
SERPINA3	HPA000893	1.8E-03	6.0E-03 (+)	5.4E-03 (+)	9.8E-01 (-)
C9	HPA029577	1.0E-02	9.9E-02 (+)	2.9E-03 (+)	6.5E-01 (-)
CCL23	HPA063758	1.6E-02	2.8E-02 (+)	1.5E-02 (+)	7.4E-01 (-)
CHI3L1	AF2599	2.5E-02	3.3E-01 (+)	4.7E-03 (+)	6.5E-01 (-)
CHIT1	HPA010115	2.5E-02	9.9E-01 (+)	3.0E-02 (+)	7.8E-02 (-)
GSTP1	HPA019779	2.6E-02	1.9E-02 (+)	7.8E-02 (+)	9.7E-01 (+)
SFTPA1/SFTPA2	HPA049368	2.6E-02	1.9E-02 (+)	5.7E-02 (+)	9.7E-01 (-)
IFNG	HPA063125	2.6E-02	2.2E-02 (+)	6.5E-02 (+)	9.9E-01 (-)
IL6	HPA044648	2.6E-02	2.2E-02 (+)	6.5E-02 (+)	9.7E-01 (+)
APOE	HPA068768	2.6E-02	2.2E-02 (+)	4.8E-02 (+)	9.9E-01 (+)
TNFRSF11A	HPA027728	2.6E-02	2.2E-02 (+)	5.6E-02 (+)	8.4E-01 (-)
SFTPB	HPA034820	2.6E-02	1.7E-01 (+)	1.9E-02 (+)	6.8E-01 (-)
IGF2	HPA007556	2.6E-02	1.8E-01 (+)	1.5E-02 (+)	7.1E-01 (-)
GZMB	HPA003418	2.6E-02	8.3E-01 (+)	2.2E-02 (+)	2.0E-01 (-)
CHIT1	HPA010575	2.6E-02	9.3E-01 (-)	4.3E-02 (+)	8.2E-02 (-)
CFI	HPA024061	3.3E-02	3.6E-02 (+)	4.8E-02 (+)	1.0E+00 (+)
MMP1	HPA031456	3.3E-02	2.9E-01 (+)	2.2E-02 (+)	6.8E-01 (-)
APOA1	HPA046715	3.7E-02	2.8E-02 (+)	6.5E-02 (+)	9.9E-01 (-)
SMPD3	HPA069383	4.0E-02	1.9E-02 (+)	2.0E-01 (+)	9.9E-01 (+)
TGFB1	HPA047516	4.0E-02	2.8E-02 (+)	7.9E-02 (+)	9.2E-01 (+)
TSLP	HPA022816	4.0E-02	1.2E-01 (+)	3.8E-02 (+)	7.4E-01 (-)
SCGB1A1	HPA031828	4.0E-02	9.8E-01 (-)	3.4E-02 (+)	2.0E-01 (-)
S100A12	HPA003620	4.0E-02	6.5E-02 (+)	5.0E-02 (+)	1.0E+00 (+)
MGP	HPA013949	4.0E-02	9.9E-02 (+)	4.3E-02 (+)	7.9E-01 (-)

Definition of abbreviations: COPD = chronic obstructive pulmonary disease; MMA = non-smokers with mild-to-moderate asthma; SAn = non-smokers with severe asthma.

*Kruskal-Wallis test, adjusted
†Wilcoxon rank-sum test, adjusted
‡ Corresponding protein name can be found at https://www.uniprot.org/

Table S6. Proteins associated with OCS use in non-smoking severe asthmatics in the U-BIOPRED cohort. Proteins (*n*=20) and the corresponding antibodies (*n*=21) identified to show significantly different plasma levels in SAn subjects that did take OCS *vs*. those that did not take OCS and where urinary OCS markers were also not detected (adjusted Wilcoxon rank-sum test*).

Gene†	Antibody	*P* value*	Significant in BIOAIR
SERPINA3	HPA000893	2.2E-03	yes (MMA, SAn)
GBA	HPA006667	2.6E-03	-
APOE	HPA065539	2.7E-03	yes (MMA, SAn)
APOA1	HPA046715	2.7E-03	yes (MMA, SAn)
APOE	HPA068768	7.0E-03	yes (MMA, SAn, COPD)
AGER	HPA069474	1.4E-02	yes (MMA, SAn)
MOCOS	HPA039412	1.6E-02	yes (MMA, SAn)
SPARCL1	HPA067587	1.7E-02	yes (MMA, SAn, COPD)
CRISP3	HPA043282	1.7E-02	yes (MMA)
OLR1	HPA050798	1.9E-02	No
RGS18	HPA028081	2.0E-02	No
SPTLC3	HPA062197	2.9E-02	yes (SAn)
IL25	HPA053829	2.9E-02	yes (MMA, SAn)
KITLG	HPA061862	3.0E-02	yes (MMA, SAn)
CHI3L1	MAB25991	3.1E-02	yes (MMA)
CCR6	HPA066394	3.1E-02	yes (MMA, SAn)
TAS2R3	HPA061025	3.1E-02	yes (MMA, SAn)
RTKN2	HPA071940	3.6E-02	yes (MMA)
IL3	HPA030770	3.6E-02	yes (MMA, SAn)
MMP1	HPA008130	4.3E-02	yes (MMA)
CCL11	HPA011652	4.8E-02	yes (MMA, SAn)

Definition of abbreviations: COPD = chronic obstructive pulmonary disease; MMA = non-smokers with mild-to-moderate asthma; SAn = non-smokers with severe asthma.

*Wilcoxon rank-sum test, adjusted

†Corresponding protein name can be found at https://www.uniprot.org/

Table S7. Group comparisons in the U-BIOPRED cohort. The majority of antibodies were still significantly different in pairwise comparisons when the analysis was limited to include only subjects that did not take OCS and where urinary OCS metabolites were not detected.

	SAn *vs*. MMA	**SAn *vs*. HC**	**SAs/ex *vs*. MMA**	**SAs/ex *vs*. HC**
Number of significant antibodies in the whole dataset (all subjects)	123	120	116	104
Number of overlapping significant antibodies when excluding OCS users	102	84	96	84
% overlap	83	70	83	81

Definition of abbreviations: HC = healthy non-smoking controls; MMA = non-smokers with mild-to-moderate asthma; SAn = non-smokers with severe asthma; SAs/ex = smokers or ex-smokers with severe asthma.

Table S8. Identification of optimal number of clusters in consensus clustering of the U-BIOPRED cohort. To find the optimal number of clusters, cluster models with two and up to ten clusters were evaluated based on cluster stability using two different methods, “Proportion of ambiguous clustering” and “Deviation from ideal stability”. Marked in grey are the two lowest values within the respective methods, indicating that a model of six or ten clusters has the best stability.

Number of clusters	Proportion of ambiguous clustering*	Deviation from ideal stability†
2	0.190	0.0741
3	0.355	0.0971
4	0.258	0.0707
5	0.249	0.0573
6	0.176	0.0406
7	0.199	0.0506
8	0.217	0.0582
9	0.196	0.0562
10	0.150	0.0394

* from [16]
† from [17]

Table S9. Clinical characteristics of the six protein-driven clusters of asthmatics in the U-BIOPRED cohort. The order of clusters is presented in the order observed for most proteins when clusters were sorted by median protein levels in decreasing order.

Cluster	**1 (*n* = 63)**	**6 (*n* = 42)**	**4 (*n* = 105)**	**3 (*n* = 101)**	**2 (*n* = 77)**	**5 (*n* = 46)**	***P* value**
MMA per cluster	3 (5%)	3 (7%)	11 (10%)	26 (26%)	16 (21%)	17 (37%)	
SA per cluster	60 (95%)	39 (93%)	94 (90%)	75 (74%)	61 (79%)	29 (63%)	
SAn	42 (67%)	32 (76%)	67 (64%)	56 (55%)	47 (61%)	19 (41%)	
SAs/ex	18 (29%)	7 (17%)	27 (26%)	19 (19%)	14 (18%)	10 (22%)	
Age (years)	55 (47.5-61.5)	51.5 (43.2-58.0)	57.0 (46.0-63.0)	53.0 (40.0-62.0)	48.0 (37.0-58.0)	45.0 (26.0-54.8)	<0.001*
Females	37 (59%)	29 (69%)	65 (62%)	61 (60%)	41 (53%)	27 (59%)	n.s.†
BMI (kg/m^2)	32 (26.3-37)	28.6 (25.1-33.3)	28.9 (24-33.3)	27.2 (24.5-30.2)	25.8 (23.6-29.8)	24.8 (20.9-28.7)	<0.001*
BMI ≥25	53 (84%)	32 (76%)	73 (70%)	71 (70%)	46 (60%)	21 (46%)	<0.001†
BMI ≥30	38 (60%)	17 (40%)	42 (40%)	26 (26%)	19 (25%)	11 (24%)	<0.001†
Age of onset *or* first diagnosis (years)	20 (5.5-33.5)	28 (12-42)	34 (12.5-45.2)	25 (10-44)	18.5 (6-35.2)	15 (4-26)	<0.01*
OCS urinary metabolites detected	23 (37%)	7 (17%)	29 (28%)	25 (25%)	17 (22%)	15 (33%)	n.s. †
OCS prescribed (all categories)	33 (52%)	15 (36%)	53 (50%)	36 (36%)	31 (40%)	15 (33%)	n.s. †
OCS prescribed (at least daily)	23 (37%)	10 (24%)	43 (41%)	27 (27%)	25 (32%)	12 (26%)	n.s. †
FEV_1 (% predicted)	67.6 (51.5-79.2)	63.6 (51.9-75.3)	68.7 (52.7-85.3)	79.3 (60.6-92.8)	74.5 (49.6-89.4)	79.4 (56.2-90.5)	<0.05*
F_ENO (ppb)	22 (14.5-34.5)	26 (14-54.8)	29.8 (16-54.2)	24.2 (17.6-39)	27.5 (19-57)	24.5 (14.6-40.8)	n.s.*
Serum periostin (ng/mL)	46.5 (38.9 - 54.8)	51.1 (41.2 - 58)	51.4 (40.1 - 61)	50.5 (43 - 60.4)	48.1 (39.6 - 55.5)	47.2 (40.8 - 52.3)	n.s.*
Serum total IgE (IU/mL)	106 (33.5 - 313)	92 (39 - 259.7)	136 (71.5 - 361.5)	145 (61.3 - 415.5)	102 (41.3 - 318.1)	125 (52 - 233.9)	n.s.*
Atopy (positive skin prick test)	36 (57%)	21 (50%)	61 (58%)	60 (59%)	43 (56%)	32 (70%)	n.s. †
RASP Type 2 composite score‡							n.s. †
Low	11 (17%)	7 (17%)	15 (14%)	16 (16%)	14 (18%)	16 (35%)	
Intermediate	30 (48%)	12 (29%)	40 (38%)	43 (43%)	33 (43%)	19 (41%)	
High	13 (21%)	13 (31%)	29 (28%)	30 (30%)	19 (25%)	7 (15%)	
Exacerbation number in past year	3 (1.5-4)	2 (1-2.8)	2 (0-3)	1 (0-3)	1 (0-3)	1.5 (0-4)	<0.05*
Blood eosinophils (cells/µL)	200 (100-400)	205 (108-425)	300 (110-500)	200 (100-450)	220 (100-395)	200 (100-300)	n.s.*
Blood neutrophils (cells/µL)	5905 (4275-8050)	5485 (4095-7340)	4700 (3575-6350)	4300 (3200-5150)	3900 (3000-5630)	3790 (3100-5100)	<0.001*

Cluster	**1 (*n* = 63)**	**6 (*n* = 42)**	**4 (*n* = 105)**	**3 (*n* = 101)**	**2 (*n* = 77)**	**5 (*n* = 46)**	***P* value**
Blood hsCRP (mg/L)	6.2 (2.8-10.2)	2.7 (1.2-5.2)	2.5 (1.1-4.9)	1.1 (0.7-2.6)	1.1 (0.62-3.1)	0.7 (0.3-1.9)	<0.001*
Sputum eosinophils (%)	1.4 (0.2 - 6.9)	6.2 (0.8 - 22.2)	4.4 (1.2 - 17.4)	1.6 (0.8 - 5.6)	2.9 (0.2 - 8.2)	1.1 (0.1 - 5.0)	n.s.*
Sputum neutrophils (%)	59 (36.2 - 88.6)	51 (37.2 - 63.9)	55.6 (37.5 - 71.7)	49.4 (25.9 - 71)	55.1 (36.4 - 78.6)	48.9 (26.8 - 63.5)	n.s.*
Urinary tetranor-PGDM	324 (204 – 407)	267 (161 – 312)	298 (234 – 416)	240 (163 – 354)	244 (169 – 328)	209 (151 – 308)	<0.001*
ACQ-7	2.6 (1.7 - 3.4)	2.3 (1.9 - 3.0)	2.1 (1.6- 3.0)	1.6 (0.7 - 2.4)	2.1 (1.1 - 3.3)	2.0 (0.9 - 3.1)	<0.001*
ACQ-7 ≥1.5 (not well controlled)	51 (81%)	33 (79%)	79 (75%)	52 (51%)	48 (62%)	27 (59%)	<0.01†
AQLQ	4.1 (3.5-5.0)	4.7 (3.8 - 5.3)	4.8 (3.7-5.5)	5.3 (4.3-6.2)	4.6 (3.7-5.7)	4.4 (3.6-6.2)	<0.001*

Definition of abbreviations: ACQ-7 = asthma control questionnaire-average of 7; AQLQ = asthma quality of life questionnaire-average; BMI = body mass index; FeNO = fraction of exhaled nitric oxide; FEV_1 = forced expiratory volume in 1 s; FVC = forced vital capacity; hsCRP = high-sensitivity C-reactive protein; MMA = non-smoker with mild-to-moderate asthma; SA = severe asthma; SAn = non-smokers with severe asthma; SAs/ex = smokers or ex-smokers with severe asthma.
Data presented as *n* (%) or median (interquartile range).

* Kruskal-Wallis test, unadjusted

† Chi-squared test, unadjusted

‡ defined in [19]

Table S10. High urinary LTE_4 levels were associated with increased levels in 7 proteins in the U-BIOPRED cohort.

Protein	Gene	Antibody	LTE_4 high	LTE_4 low	*P* value*
Complement component 8 gamma chain	C8G	HPA046269	3637 (3110-4075)	3085 (2732-3631)	0.019
Fibroleukin	FGL2	HPA026682	14502 (12800-16233)	13207 (11631-14615)	0.034
Interferon regulatory factor 8	IRF8	HPA002531	4322 (3311-6041)	3352 (2379-4970)	0.038
Psoriasis susceptibility 1 candidate gene 2 protein	PSORS1C2	HPA051817	11281 (9778-13060)	9667 (8351-11366)	0.019
Alfa-1-antitrypsin (AAT)	SERPINA1	HPA000927	11036 (10616-11550)	10729 (10451-11064)	0.034
Alpha-1-antichymotrypsin (ACT)	SERPINA3	HPA000893	14278 (13689-14958)	13755 (13041-14467)	0.019
Lung surfactant protein D (SPD)	SFTPD	HPA044582	34920 (32968-36656)	33327 (31470-35722)	0.044

Data presented as median (interquartile range).

*Wilcoxon rank-sum test, adjusted

SUPPLEMENTARY FIGURES

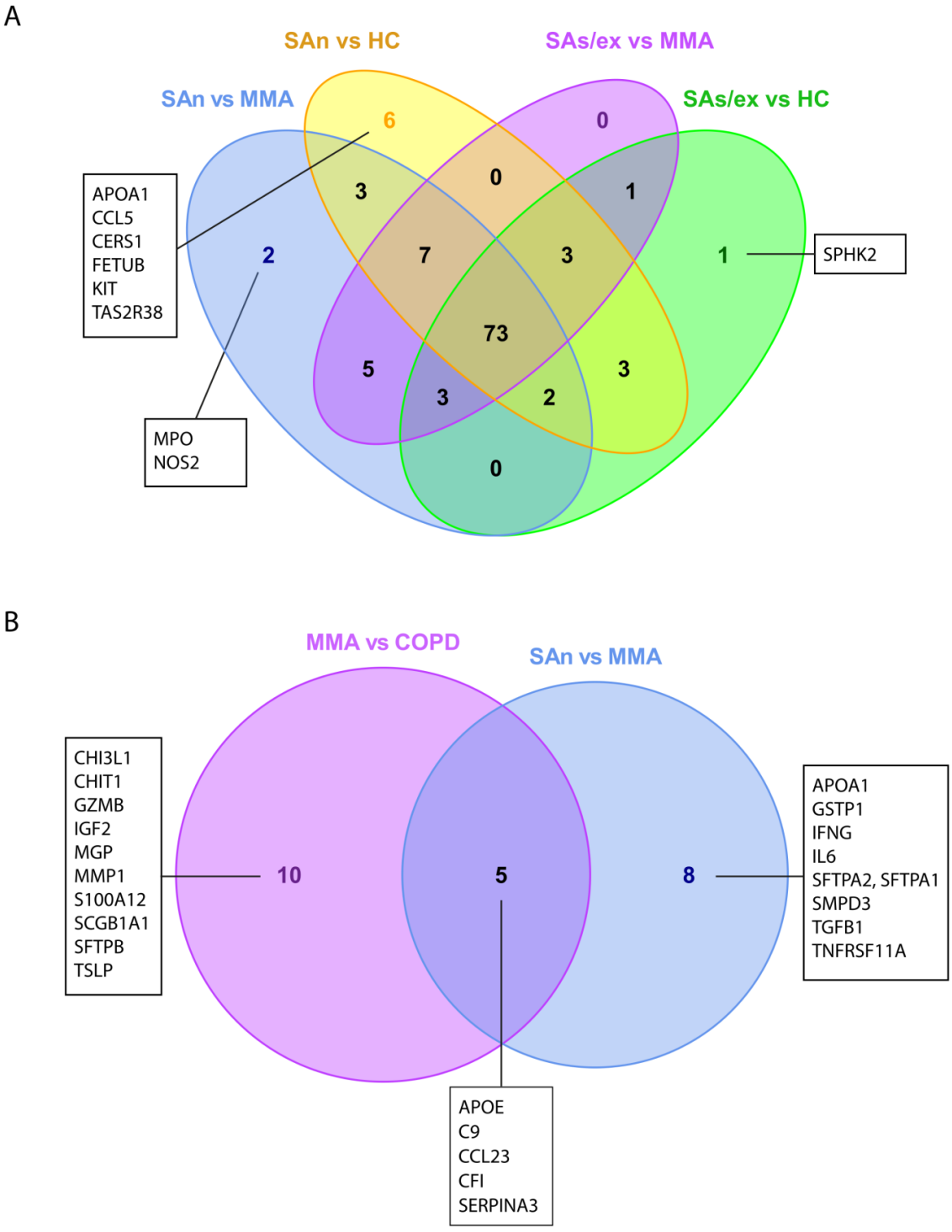

Figure S1. Venn diagram illustrating the number of plasma proteins that were significantly different (adjusted P <0.05) in pairwise group comparisons of **(A)** SAn vs. MMA, SAn vs. HC, SAs/ex vs. MMA, and SAs/ex vs. HC in U BIOPRED, or **(B)** in comparisons of SAn vs. MMA and MMA vs. COPD in BIOAIR (B). Note that no proteins showed differential levels between SAn vs. COPD. COPD = chronic obstructive pulmonary disease; HC = healthy non-smoking controls; MMA = non-smokers with mild-to-moderate asthma; SAn = non-smokers with severe asthma; SAs/ex = smokers or ex-smokers with severe asthma.

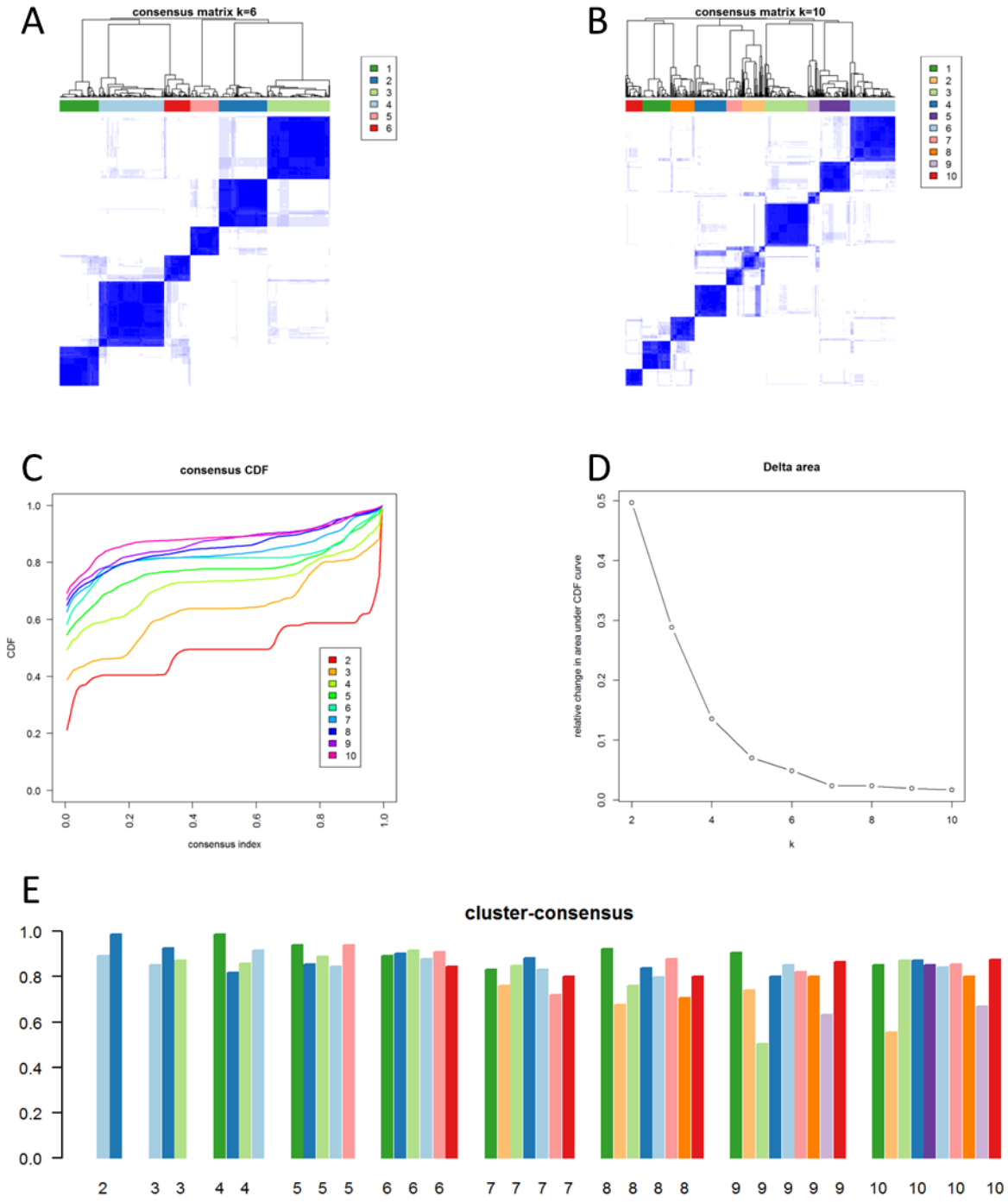

Figure S2. Consensus clustering in U-BIOPRED. Clustering was applied to U-BIOPRED subjects with asthma (*n*=434) using the input of protein profiles from 139 antibodies (identified to show a significant difference between asthma severity groups or healthy controls). **(A-B)** Consensus matrices for models with optimal number of clusters. Six clusters (K=6; A) and ten clusters (K=10; B) visualized in heatmaps. Values are between 0 (shown in white; meaning two samples are never clustered together across the 1000 iterations) and 1 (shown in dark blue; meaning two samples are always clustered together across the 1000 iterations). **(C)** Consensus CDF plot for cluster models of two (K=2) and up to ten (K=10) clusters. **(D)** Delta area plot with the relative change in the area under the CDF curve when comparing models with "K" clusters and "K_{-1}" clusters. For K=2, the total area under the curve is shown. **(E)** Cluster-consensus plot where, for a given cluster model K, the mean of all pairwise consensus values is higher if subjects are more often clustered together.

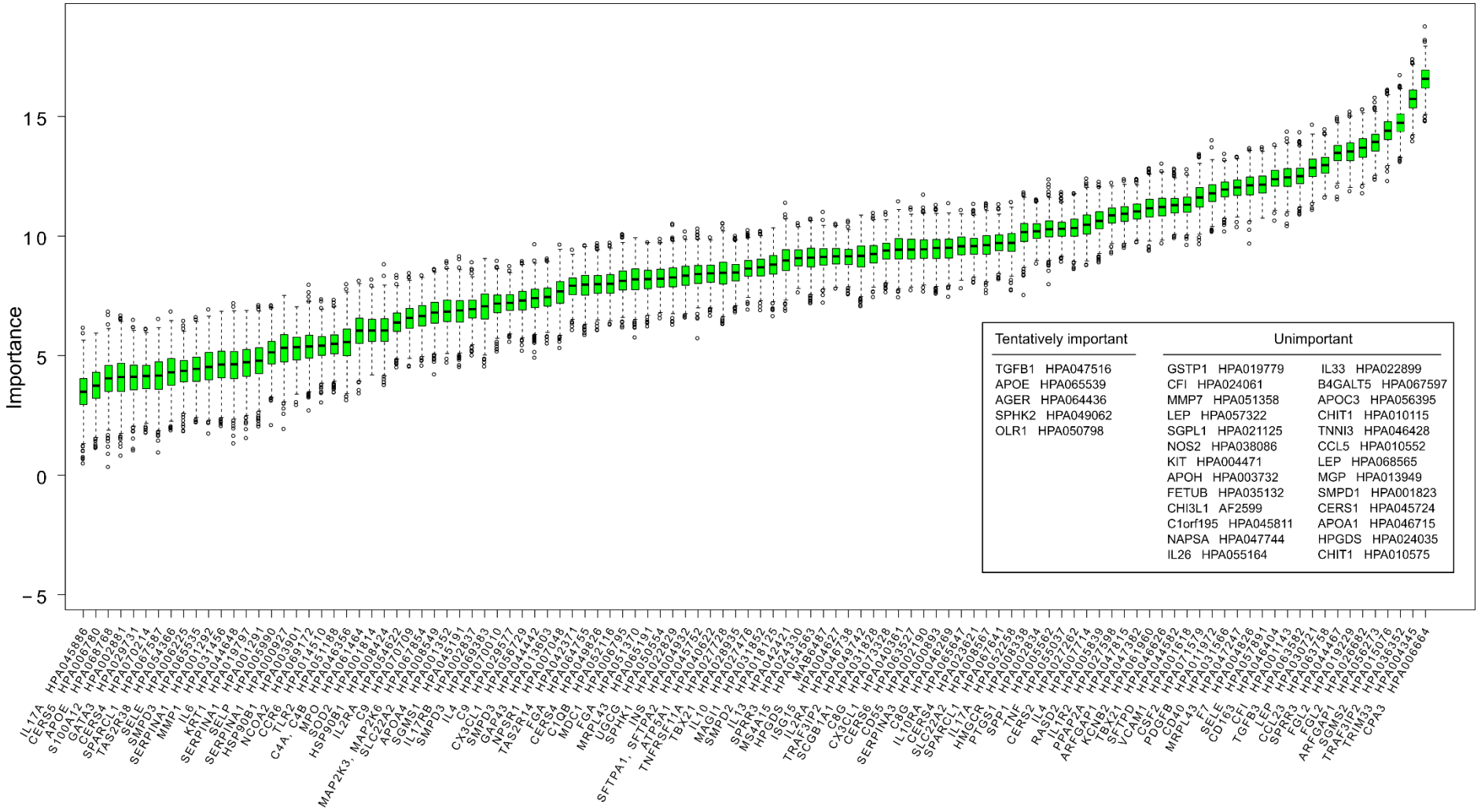

Figure S3. Importance of proteins for classification of clusters. The Boruta algorithm identifies all variables with significant importance for the classification of six clusters and furthermore defines each variable as confirmed (*n*=108, with increasing importance from left to right), tentatively important (*n*=5) or unimportant (*n*=26).

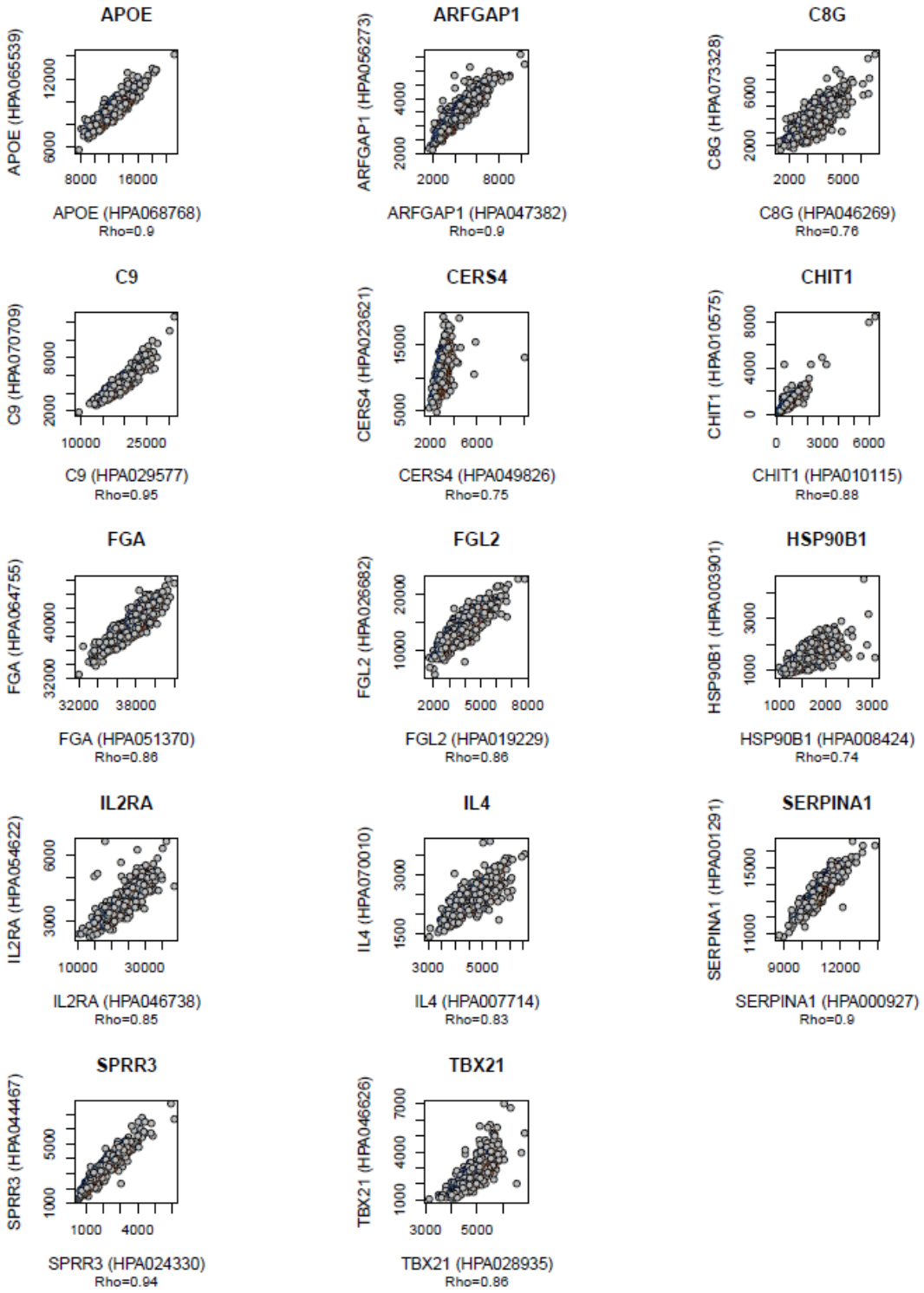

Figure S4. Supporting antibody profiles. Examples of proteins in the U-BIOPRED dataset where antibodies targeting different or same regions of a protein showed concordant profiles (Spearman's Rho >0.7).

REFERENCES

1. Bousquet J. Global initiative for asthma (GINA) and its objectives. *Clin Exp Allergy* 2000; 30 Suppl 1: 2-5.
2. The ENFUMOSA cross-sectional European multicentre study of the clinical phenotype of chronic severe asthma. *European Respiratory Journal* 2003; 22: 470.
3. Bel EH, Sousa A, Fleming L, Bush A, Chung KF, Versnel J, Wagener AH, Wagers SS, Sterk PJ, Compton CH. Diagnosis and definition of severe refractory asthma: an international consensus statement from the Innovative Medicine Initiative (IMI). *Thorax* 2011; 66: 910-917.
4. Bousquet J, Mantzouranis E, Cruz AA, Aït-Khaled N, Baena-Cagnani CE, Bleecker ER, Brightling CE, Burney P, Bush A, Busse WW, Casale TB, Chan-Yeung M, Chen R, Chowdhury B, Chung KF, Dahl R, Drazen JM, Fabbri LM, Holgate ST, Kauffmann F, Haahtela T, Khaltaev N, Kiley JP, Masjedi MR, Mohammad Y, O'Byrne P, Partridge MR, Rabe KF, Togias A, van Weel C, Wenzel S, Zhong N, Zuberbier T. Uniform definition of asthma severity, control, and exacerbations: document presented for the World Health Organization Consultation on Severe Asthma. *J Allergy Clin Immunol* 2010; 126: 926-938.
5. Moore WC, Bleecker ER, Curran-Everett D, Erzurum SC, Ameredes BT, Bacharier L, Calhoun WJ, Castro M, Chung KF, Clark MP, Dweik RA, Fitzpatrick AM, Gaston B, Hew M, Hussain I, Jarjour NN, Israel E, Levy BD, Murphy JR, Peters SP, Teague WG, Meyers DA, Busse WW, Wenzel SE, , National Heart, Lung Blood Institute's Severe Asthma Research Program. Characterization of the severe asthma phenotype by the National Heart, Lung, and Blood

Institute's Severe Asthma Research Program. *The Journal of allergy and clinical immunology* 2007; 119: 405-413.

6. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Ponten F. Proteomics. Tissue-based map of the human proteome. *Science* 2015; 347: 1260419.
7. Mikus M, Drobin K, Gry M, Bachmann J, Lindberg J, Yimer G, Aklillu E, Makonnen E, Aderaye G, Roach J, Fier I, Kampf C, Gopfert J, Perazzo H, Poynard T, Stephens C, Andrade RJ, Lucena MI, Arber N, Uhlen M, Watkins PB, Schwenk JM, Nilsson P, Schuppe-Koistinen I. Elevated levels of circulating CDH5 and FABP1 in association with human drug-induced liver injury. *Liver Int* 2017; 37: 132-140.
8. Schwenk JM, Gry M, Rimini R, Uhlen M, Nilsson P. Antibody suspension bead arrays within serum proteomics. *J Proteome Res* 2008; 7: 3168-3179.
9. Hubert M, Rousseeuw PJ. ROBPCA: A New Approach to Robust Principal Component Analysis. *Technometrics* 2005; 47: 64-79.
10. Hong MG, Lee W, Nilsson P, Pawitan Y, Schwenk JM. Multidimensional Normalization to Minimize Plate Effects of Suspension Bead Array Data. *J Proteome Res* 2016; 15: 3473-3480.
11. Ihaka R, Gentleman R. R: a language for data analysis and graphics. *Journal of computational and Graphical Statistics* 1996: 299-314.

12. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
13. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological* 1995; 57: 289-300.
14. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. *Machine Learning* 2003; 52: 91-118.
15. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 2010; 26: 1572-1573.
16. Senbabaoglu Y, Michailidis G, Li JZ. Critical limitations of consensus clustering in class discovery. *Sci Rep* 2014; 4: 6207.
17. Lefaudeux D, De Meulder B, Loza MJ, Peffer N, Rowe A, Baribaud F, Bansal AT, Lutter R, Sousa AR, Corfield J, Pandis I, Bakke PS, Caruso M, Chanez P, Dahlen SE, Fleming LJ, Fowler SJ, Horvath I, Krug N, Montuschi P, Sanak M, Sandstrom T, Shaw DE, Singer F, Sterk PJ, Roberts G, Adcock IM, Djukanovic R, Auffray C, Chung KF, U-BIOPRED Study Group. U-BIOPRED clinical adult asthma clusters linked to a subset of sputum omics. *J Allergy Clin Immunol* 2017; 139: 1797-1807.
18. Kursa MB, Rudnicki WR. Feature Selection with the Boruta Package. *Journal of Statistical Software* 2010; 36: 1-13.
19. Heaney LG, Busby J, Hanratty CE, Djukanovic R, Woodcock A, Walker SM, Hardman TC, Arron JR, Choy DF, Bradding P, Brightling CE, Chaudhuri R, Cowan DC, Mansur AH, Fowler SJ, Niven RM, Howarth PH, Lordan JL,

Menzies-Gow A, Harrison TW, Robinson DS, Holweg CTJ, Matthews JG, Pavord ID, on behalf of the investigators for the MRC Refractory Asthma Stratification Programme. Composite type-2 biomarker strategy versus a symptom-risk-based algorithm to adjust corticosteroid dose in patients with severe asthma: a multicentre, single-blind, parallel group, randomised controlled trial. *Lancet Respir Med* 2020.