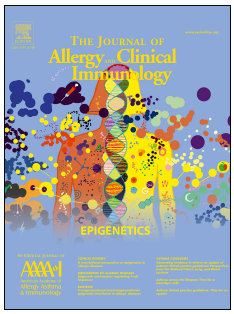
Accepted Manuscript



Stratification of asthma phenotypes by airway proteomic signatures

James P.R. Schofield, PhD, Dominic Burg, PhD, Ben Nicholas, PhD, Fabio Strazzeri, PhD, Joost Brandsma, PhD, Doroteya Staykova, PhD, Caterina Folisi, PhD, Aruna T. Bansal, PhD, Yang Xian, PhD, Yike Guo, PhD, Anthony Rowe, PhD, Julie Corfield, PhD, Susan Wilson, PhD, Jonathan Ward, Bsc, Rene Lutter, PhD, Dominick E. Shaw, MD, Per S. Bakke, MD, Massimo Caruso, MD, Sven-Erik Dahlen, MD, Stephen J. Fowler, MD, Ildikó Horváth, MD, Peter Howarth, MD, Norbert Krug, MD, Paolo Montuschi, MD, Marek Sanak, MD, Thomas Sandström, MD, Kai Sun, PhD, Ioannis Pandis, PhD, John Riley, PhD, Charles Auffray, PhD, Bertrand De Meulder, PhD, Diane Lefaudeux, PhD, Ana R. Sousa, PhD, Ian M. Adcock, PhD, Kian Fan Chung, MD, Peter J. Sterk, MD, Paul J. Skipp, PhD, Ratko Djukanović, MD, on behalf of the U-BIOPRED Study Group

PII: S0091-6749(19)30415-4

DOI: https://doi.org/10.1016/j.jaci.2019.03.013

Reference: YMAI 13943

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 12 November 2018

Revised Date: 14 February 2019

Accepted Date: 8 March 2019

Please cite this article as: Schofield JPR, Burg D, Nicholas B, Strazzeri F, Brandsma J, Staykova D, Folisi C, Bansal AT, Xian Y, Guo Y, Rowe A, Corfield J, Wilson S, Ward J, Lutter R, Shaw DE, Bakke PS, Caruso M, Dahlen S-E, Fowler SJ, Horváth I, Howarth P, Krug N, Montuschi P, Sanak M, Sandström T, Sun K, Pandis I, Riley J, Auffray C, De Meulder B, Lefaudeux D, Sousa AR, Adcock IM, Chung KF, Sterk PJ, Skipp PJ, Djukanović R, on behalf of the U-BIOPRED Study Group, Stratification of asthma phenotypes by airway proteomic signatures, *Journal of Allergy and Clinical Immunology* (2019), doi: https://doi.org/10.1016/j.jaci.2019.03.013.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please

note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Title: Stratification of asthma phenotypes by airway proteomic signatures

James P R Schofield, PhD[1, 2*], Dominic Burg, PhD[1, 2*], Ben Nicholas, PhD[2], Fabio Strazzeri, PhD[1, 3], Joost Brandsma, PhD[2], Doroteya Staykova, PhD[1], Caterina Folisi, PhD[1], Aruna T Bansal, PhD[4], Yang Xian, PhD[5], Yike Guo, PhD[5], Anthony Rowe, PhD[6], Julie Corfield, PhD[7], Susan Wilson, PhD[2], Jonathan Ward, Bsc[2], Rene Lutter, PhD[8, 9], Dominick E Shaw, MD[10], Per S Bakke, MD[11], Massimo Caruso, MD[12], Sven-Erik Dahlen, MD[13], Stephen J. Fowler, MD[14], Ildikó Horváth, MD[15], Peter Howarth, MD[2], Norbert Krug, MD[16], Paolo Montuschi, MD[17], Marek Sanak, MD[18], Thomas Sandström, MD[19], Kai Sun, PhD[5], Ioannis Pandis, PhD[5], John Riley, PhD[21], Charles Auffray, PhD[20], Bertrand De Meulder, PhD[20], Diane Lefaudeux, PhD[20], Ana R Sousa, PhD[21], Ian M Adcock, PhD[22], Kian Fan Chung, MD[22], Peter J Sterk, MD[9], Paul J Skipp, PhD[1, #], Ratko Djukanović, MD[2, #], on behalf of the U-BIOPRED Study Group

*Joint first authors

[#] Joint senior authors

[1]Centre for Proteomic Research, Biological Sciences, University of Southampton, UK

[2]NIHR Southampton Biomedical Research Centre, Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, UK

[3]Mathematical Sciences, University of Southampton, UK

[4]Acclarogen Ltd, Cambridge, UK

[5]Data Science Institute, Imperial College, London, UK

[6]Janssen Research & Development, Buckinghamshire, UK

[7]Areteva Ltd, Nottingham, UK

[8]AMC, Department of Experimental Immunology, University of Amsterdam, Amsterdam, The Netherlands

ACCEPTED MANUSCRIPT

[9]Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

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

[11]Institute of Medicine, University of Bergen, Bergen, Norway

[12]Dept. of Clinical and Experimental Medicine Hospital University, University of Catania, Catania, Italy.

[13]The Centre for Allergy Research, The Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.

[14]Respiratory and Allergy Research Group, University of Manchester, Manchester, UK.

[15]Dept. of Pulmonology, Semmelweis University, Budapest, Hungary

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

[17]Faculty of Medicine, Catholic University of the Sacred Heart, Rome, Italy.

[18]Laboratory of Molecular Biology and Clinical Genetics, Medical College, Jagiellonian University, Krakow, Poland

[19]Dept. of Medicine, Dept of Public Health and Clinical Medicine Respiratory Medicine Unit, Umeå University, Umeå, Sweden.

[20]European Institute for Systems Biology and Medicine, CNRS-ENS-UCBL-INSERM, Université de Lyon, France

[21]Respiratory Therapeutic Unit, GSK, Stockley Park, UK.

[22]Cell and Molecular Biology Group, Airways Disease Section, National Heart and Lung Institute, Imperial College London, Dovehouse Street, London, UK

To whom correspondence should be addressed: Ratko Djukanović at R.Djukanovic@soton.ac.uk, Mailpoint 810, Level F, Sir Henry Wellcome Laboratories,

3 South Block, Clinical and Experimental Sciences, University Hospital Southampton,
4 Tremona Road, Southampton SO16 6YD Phone: +44 2381204195

Sources of funding

The U-BIOPRED consortium receives funding from the European Union and from the European Federation of Pharmaceutical Industries and Associations as an IMI JU funded project (no. 115010).

Conflicts of interest: Ratko Djukanović has consulted and presented at symposia organised by TEVA, Novartis, GlaxoSmithKline and AstraZeneca, has shares in and consults for Synairgen; Charles Auffray reports grants from Innovative Medicine Initiative; Jeanette Bigler reports that she owns stock in Amgen Inc; Kian Fan Chung has received honoraria for participating in Advisory Board meetings of the pharmaceutical industry regarding treatments for asthma and chronic obstructive pulmonary disease and has also been renumerated for speaking engagements; Stephen Fowler reports personal fees from AstraZeneca, personal fees from Boehringer Ingelheim, personal fees from Novartis, personal fees from TEVA, outside the submitted work; Ildiko Horvath reports personal fees from TEVA, personal fees from Novartis, personal fees from Chiesi, personal fees from AstraZeneca, personal fees from GlaxoSmithKline, personal fees from Boehringer-Ingelheim, personal fees from Berlin-Chemie, personal fees from CSL Behring, personal fees from Roche, personal fees from Orion Pharma, personal fees from Sager Pharma, outside the submitted work; Peter Howarth reports personal fees from GSK, grants from Boehringer-Ingelheim, outside the submitted work; Norbert Krug reports grants from several pharmaceutical companies, outside the submitted work; Rene Lutter reports personal fees from Boehringer Ingelheim, grants from Medimmune, grants from GlaxoSmithKline, outside the submitted work. Dominick Shaw reports personal fees from Novartis, GSK and TEVA. Peter Sterk reports grants from IMI:

Innovative Medicines Initiative, during the conduct of the study; Thomas Sandstrom reports fee from Boehringer Ingelheim, outside the submitted work; Anthony Rowe is an Employee and Shareholder of Janssen Research and Development, a Johnson and Johnson company; The rest of the authors have nothing to disclose.

Running title Sputum proteomic signatures in asthma

Abstract

Background: Stratification by eosinophil and neutrophil counts increases our understanding of asthma and helps target therapy, but there is room for improvement in our accuracy to predict treatment responses and a need for better understanding of the underlying mechanisms.

Objective: Identify molecular sub-phenotypes of asthma defined by proteomic signatures for improved stratification.

Methods: Unbiased label-free quantitative mass spectrometry and topological data analysis were used to analyse the proteomes of sputum supernatants from 246 participants (206 asthmatics) as a novel means of asthma stratification. Microarray analysis of sputum cells provided transcriptomics data additionally to inform on underlying mechanisms.

Results: Analysis of the sputum proteome resulted in 10 clusters, proteotypes, based on similarity in proteomics features, representing discrete molecular sub-phenotypes of asthma. Overlaying granulocyte counts onto the 10 clusters as metadata further defined three of these as highly eosinophilic, three as highly neutrophilic, and two as highly atopic with relatively low granulocytic inflammation. For each of these three phenotypes, logistic regression analysis identified candidate protein biomarkers, and matched transcriptomic data pointed to differentially activated underlying mechanisms.

Conclusion: This study provides further stratification of asthma currently classified by quantifying granulocytic inflammation and gives additional insight into their underlying mechanisms which could become targets for novel therapies.

Clinical implications: By further stratifying asthma, the sputum proteomes and molecular pathways identified in this study could explain the variability in treatment response to asthma therapies, in particular the biologics.

Capsule summary: This study advances asthma stratification, shows that further stratification based on granulocyte counts is possible, points to novel mechanisms and identifies candidate biomarkers for use in research and clinical practice.

Key words: asthma, proteomics, biomarkers, eosinophils, neutrophils

Abbreviations:

U-BIOPRED: Unbiased BIOmarkers for the PREDiction of respiratory disease outcomes

FEV1: forced expiratory volume in one second

FeNO: fraction of exhaled nitric oxide

NGS: next generation sequencing

IPA: ingenuity pathway analysis

TDA: topological data analysis

PAM: Partitioning Around Medoids

LC/HDMS^{E}: Liquid Chromatography with High Definition Mass Spectrometry

ROC: receiver operating characteristic

ENLR: elastic-net-regularized logistic regression

IPA: Ingenuity Pathway Analysis

ACQ: Asthma Control Questionnaire

AQLQ: Asthma Quality of Life Questionnaire

Introduction

Asthma is a heterogenous disease, involving inflammatory and structural cells and a multitude of molecular mediators. Development of new drugs, including biologics(1), with stratified medicine will improve patient outcomes. Current, phenotypes of asthma are defined predominantly by clinical characteristics and a limited number of biomarkers(2). The most widely studied asthma biomarkers to date are sputum and blood eosinophils, exhaled breath nitric oxide (eNO) and serum periostin(3). Blood eosinophilia predicts risk of acute asthma exacerbations(4) and treatment strategies incorporating normalisation of sputum eosinophil counts have resulted in marked reductions in exacerbations(5). Eosinophilia has been used to enrich target populations in trials of the anti-IL-5 monoclonal. By comparison, defining asthma phenotypes by other inflammatory cell types has been less rewarding.

In the first project to apply multiple 'omics methodologies in an unbiased manner to stratify asthma, U-BIOPRED (Unbiased BIOmarkers for the Prediction of respiratory disease outcomes) has shed light on conceptually novel mechanisms and phenotypes. In the current study (Fig. S1), quantitative data-independent LC/HDMS^E (HDMS^E) was used to profile proteomes of the lining fluid of the bronchial epithelium sampled by sputum induction. We hypothesised that clustering of sputum proteomic data would identify molecular phenotypes at higher resolution than the existing asthma eosinophilic, neutrophilic and pauci-granulocytic phenotypes based on sputum granulocyte counts. By investigating molecular pathways, defined by whole-genome array from the sputum cells of the same participants, we have begun to understand the mechanisms that underlie these new phenotypes.

METHODS

Study design

U-BIOPRED is a multi-centre study where participants were phenotyped using standardised protocols, lung function testing and assessment of atopy(6) and applying several 'omics unbiased technologies (genomics, transcriptomics, proteomics, lipidomics, breathomics and metabolomics) to enable novel concepts of asthma mechanisms and phenotypes to be developed. In this study, we report the results of proteomic analysis of sputum supernatants as a means of studying the bronchial epithelial lining fluid sampled by standard sputum induction widely used in asthma research. Of the total 610 adult participants enrolled into U-BIOPRED, 246 provided sputum samples that passed stringent quality control for proteomics analysis: 118 non-smoking severe asthmatics, 48 current or ex-smoking severe asthmatics, 40 mild-moderate asthmatics, and 40 healthy participants (Table 1).

Sputum induction and proteomic analysis

Sputum collection and analysis were performed according to methods detailed by Burg et al.(7). Proteins were extracted from sputum supernatants by precipitation and analysed in duplicate via LC- MS^E using a Waters G2Si mass spectrometer coupled to a nanoAcquity UPLC (Waters). Patient-matched cell pellets processed into RNAlater were assessed for global gene expression using Affymetrix HT HGU133+ microarray analysis. For full set of microarray data see previous report(8).

Data analysis

Patients were randomised into a training and test set with a 3:7 ratio. All proteomics data were log2 transformed. Batch effects were corrected using ComBat(9). Statistical analyses were performed in R, custom Python scripts, and Microsoft Excel. Clustering of patients was based solely on proteomic data using topological data analysis (TDA) performed in Ayasdi Core software applying a norm correlation metric and two MDS lenses (resolution 32, gain

3.4 ×, equalized). Sub-phenotypes of asthma were assigned based on the persistence of TDA structure and conserved groupings of nodes when i) changing sampling settings, modifying input data, e.g. log vs. natural data, ii) varying the cohort input, e.g. all participants vs. asthmatics only, and iii) through use of metadata from consensus models and experimental composition (see online supplement). Cluster boundaries were created through an iterative process of varying resolution and gain settings and comparing the different analysis datasets. For comparison and validation of the TDA approach, the same data were also clustered using the consensus cluster plus R package, applying Partitioning Around Medoids (PAM) and Pearson's correlation with settings of Maximum k=20 and 1000 repetitions with item resampling setting of 0.9 (Fig. S2).

In order to identify individual predictive proteins that could be used in future studies as candidate biomarkers, we applied elastic net-regularised logistic regression (ENLR) to data in the training set and then tested the predictive value of the proteins by receiver operating characteristic (ROC) curve in the test set. Where it was not appropriate to break up data into training and test datasets because of sample size, exploratory analysis to identify predictive biomarkers was performed using ENLR alone applied to the combined training and test dataset, comparing the cluster of interest to the rest of the participants. The logistic regression and ROC curves were conducted using scikit-learn in Python.

Pathway analysis of patient-matched sputum transcriptomics data

RNA extracts from sputum cell pellets from the same samples from which sputum supernatants were analysed for protein content by HDMS^{E} were subjected to microarray analysis and Ingenuity Pathway Analysis (IPA).

RESULTS

The main demographic characteristics and medication of participants from the U-BIOPRED cohort whose data were used in this study (n=246) are shown in Table 1.

Stratification of asthmatics by protein signatures

Using our previously reported rationale for selecting sets of proteins for analysis(7), 270 proteins, identified and quantified in ≥40% of samples, constituted the core dataset for statistical analysis (Table S1). Within the network constructed by TDA, 10 clusters were identified, representing phenotypes of asthmatics with distinct proteomic signatures that we termed "proteotypes" (Fig. 1). When sputum granulocyte counts were overlaid as metadata onto the TDA network, a strong association with sputum eosinophils and neutrophils across the proteotypes was observed: Proteotypes 1, 2 and 3 were highly eosinophilic (mean counts 18.7%, 23.1% and 17.0%, respectively) and were, therefore, defined as highly eosinophilic "sub-phenotypes" 1, 2 and 3. Collectively, they represented a common, highly eosinophilic "phenotype", comprising 31% of the cohort, with a mean eosinophil count of 21.1%. Clusters (sub-phenotypes) 5 and 6, had atopy as their main shared feature. Clusters (sub-phenotypes) 8, 9 and 10 were characterised by raised neutrophil counts (71.8% for all three combined, i.e. the neutrophil phenotype, p=1.03E^{-9} when compared to the rest of the participants), with counts progressively increasing from sub-phenotype 8 to 10 (mean 63.5%, 82.0% and 72.6% for proteotypes 8, 9 and 10, respectively). Two smaller, less well-defined phenotypes (composed of single proteotypes) were also identified: a mildly eosinophilic proteotype 4 (mean eosinophil count 7.7%) and a mildly neutrophilic proteotype 7 (mean neutrophil count 46.2%).

Sub-phenotypes of eosinophilic asthma

Sputum eosinophil counts in proteotypes 1, 2 and 3 were significantly (p=5.49E^{-9}) higher than the rest of the TDA network (mean 5.4%) (Fig. 2 and Table 2), with 80 sputum proteins

differentially abundant comparted to health (p<0.05) (Table S2) and 49 proteins compared to the rest of asthmatics (p<0.05) (Table S3). Additionally, 14 proteins were differentially abundant at a significance of p<0.05 when comparing the individual eosinophilic sub-phenotypes (Table 3). Unlike the sub-phenotypes, the three major phenotypes were represented by sufficient patient numbers in the predefined, randomly assigned, training and test cohorts to allow testing of predictive success of associated proteins (training:test ratios were 40:22, 29:21, 38:21 for highly eosinophilic, highly atopic and highly neutrophilic phenotypes, respectively). Logistic regression/ROC analysis showed that ten proteins were strongly predictive of the entire eosinophilic phenotype (Fig. 3 and Fig. S3A): histone H4, vitronectin, histidine-rich glycoprotein, immunoglobulin heavy constant gamma 3, complement C3, transthyretin, serotransferrin and alpha-1 antitrypsin (all raised relative to the rest of the cohort) and galectin-3-binding protein and ezrin(both reduced).

Other biomarkers from the U-BIOPRED database were also elevated in this phenotype including blood periostin (Fig. 2D), a biomarker strongly associated with the Th2 cytokine phenotype, eosinophilia and airway remodelling[4] (Fig. 3A, Table 2) and sputum IL13 (Fig. 2B) (Fig. 2C).

Slightly fewer than a third of patients with this phenotype were on OCS and 55% were atopic, with more frequent respiratory infections (93.5% vs. 67.0%, p=8.7e^{-7}), use of long-acting β-agonists (83.9% vs. 58.5%, p=1.5E^{-4}) and higher concentrations of nitric oxide in exhaled breath (eNO) (44.4±35.8 ppb) (Table 2), implying more severe asthma than the rest of the cohort. Sub-phenotype 1 asthmatics had a higher prevalence of sinusitis (69.2% vs. 30.6% of participants in sub-phenotypes 2 and 3. Eosinophilic sub-phenotype 2 had the highest eNO levels (mean 48.77 ± SD 36.83 ppb) and highest sputum eosinophil count (23.08%), while sub-phenotype 3 had an increased frequency of atopy (81.8% compared to 37.3% in the other eosinophilic sub-phenotypes).

Sub-phenotypes of neutrophilic asthma

The neutrophilic phenotype, composed of sub-phenotypes 8, 9 and 10 (Fig. 2 and Table 2), represented 29.5% of the asthmatics and had significantly higher sputum neutrophil counts ($p = 1.03E^{-9}$). By comparison with other asthmatics, they had 134 differentially abundant proteins (Table S7). Fourteen proteins were differentially abundant between the neutrophilic sub-phenotypes (Table 3). Within this region of the TDA network, there was a concentration gradient of neutrophil-derived proteins, increasing from the left to the right side of the network (Fig. 2J and K): S100s, plastin2, leukocyte elastase inhibitor, MMP9 and leukotriene A4 hydrolase. This was associated with higher concentrations of sputum S100-A9 and matrix metalloproteinase-9 (MMP-9) proteins. The ten most predictive proteins identified by logistic regression/ROC analysis for this phenotype were histone H4 (of note also predictive of the eosinophilic phenotype), azurocidin, coronin-1A, chloride intracellular channel protein 1, annexin A1 and A3, neutrophil gelatinase-associated lipocalin (all raised in this phenotype) and transthyretin (reduced) (Fig. 3 and Fig S3A).

The number of ICU admissions in the last 12 months was highest in the neutrophilic phenotype (Fig. 2L). Their asthma symptoms were more likely to be triggered by fungus (54.2% vs. 23.6%, $p=1.4 \times 10^{-5}$) and dust (81.4% vs 51.3%, $p=2.3E^{-5}$). OCS consumption was highest in the neutrophilic sub-phenotypes 8-10, lowest in the highly atopic sub-phenotypes 5 and 6, and intermediate, but heterogeneous, in the eosinophilic sub-phenotypes 1-3 (Table 2 and Figure 2). However, statistical analysis showed that OCS use was significantly higher (compared to other asthmatics) only in sub-phenotype 9 which was also associated with decreased β-agonist reversibility. When compared to neutrophilic sub-phenotypes 8 and 9, sub-phenotype 10 (5.5% of asthmatics) was associated with higher frequency of ICU admissions (63.6% compared to 20.8% in sub-phenotypes 8 and 9; $p=8.9E^{-3}$) but less rescue inhaler use ($p=3.9E^{-3}$).

Sub-phenotypes of highly atopic asthma

This group constituted 25% of all asthmatics (Fig. 2E), characterised by mild asthma, with highest quality of life (ACQ lower and AQLQ higher), overall good lung function (mean FEV_1 78% of predicted), but with higher prevalence of atopy (78% vs. 53% in the other asthmatics, $p=9x10^{-4}$) (Table 2) and total serum IgE concentrations (Fig. 2F and H). They had high levels of sputum uteroglobin and clusterin proteins (Fig. 2E and G). By comparison with the other asthmatics and healthy participants, 134 and 20 proteins were, respectively, differentially abundant (Tables S4 and S5). The predictive biomarkers for this phenotype are also shown in Fig. 3 and Fig. S4A.

Comparisons of the two highly atopic sub-phenotypes showed significantly ($p=0.005$) higher abundance of lysozyme C in sub-phenotype 6 when compared to sub-phenotype 5. Conversely, eNO was significantly ($p=0.002$) lower in sub-phenotype 6, with a trend towards lower ACQ and higher AQLQ in sub-phenotype 5, suggesting that sub-phenotype 5 (9% of U-BIOPRED asthmatics) was less severe. Of note, sub-phenotype 5 had the highest incidence of atopy amongst all asthma sub-phenotypes and has a higher frequency of active hay fever, abut also had the highest FEV_1 and FVC of all sub-phenotypes. Lysozyme C was more highly abundant in sub-phenotype 6, compared to sub-phenotype 5 (Table 3), possibly due to higher neutrophil cell count (Table 2).

Pathway analysis of patient-matched sputum transcriptomics data

The sub-phenotypes were next compared for upstream regulators by applying IPA to the transcriptomics dataset. This revealed a general trend, i.e. a broad pattern, when moving from left to right across the TDA structure, with decreasing T2/atopy-associated and increasing T1 gene expression when placing the clusters approximately in the order shown in Table S9 and Fig. S6. The same analysis showed that gene expression in the neutrophilic sub-phenotypes was predicted to result from down-regulation of the T2 cytokine, IL13. Expression of T2

cytokine, IL5 mRNA, was highest in the eosinophilic sub-phenotypes (Fig. S5A). The pleiotropic cytokine Thymic Stromal Lymphopoietin (TSLP) was detected at low abundance and was not clearly distributed between the sub-phenotypes; however, some of the highest TSLP expression was measured in eosinophilic sub-phenotypes 1 and 2.

The TDA network was also characterised by increasing activation of IL2 from left to right. Neutrophilic sub-phenotypes had gene expression profiles predicted to result from down-regulation of COL18A1, a gene associated with atopy. Additionally, there was a left to-right trend of predicted activation of virally-induced transcription factors; interferon alpha, KDM5B and TNF. The top canonical pathways and upstream regulators of gene expression were predicted by IPA analysis, top upstream regulator for each sub-phenotype is shown in Table S8.

ACCEPTED MANUSCRIPT

DISCUSSION

Using unbiased proteomic profiling of induced sputum supernatants, we have achieved a greater degree of stratification than currently possible by granulocyte counts alone, sub-stratifying eosinophilic and neutrophilic phenotypes each into three sub-phenotypes. Application of TDA provided a new perspective on stratification by creating a network with patient clusters defined by shared airway proteomes. Asthma severity increased across the network, the most severe forms being at the extreme right end where neutrophilia was a striking feature. Analysis of the sputum cell transcriptome from the same patients pointed to mechanistic pathways that could inform further optimisation of asthma biologics and help development of new asthma drugs. Logistic regression and ROC analysis identified several candidate biomarkers to be explored further for application in clinical practice, possibly when selecting patients for novel drugs and existing biologic treatments.

As anticipated, the eosinophilic phenotype showed patterns of sputum cell gene expression normally associated with eosinophilia, i.e. IL-4 and IL-13 (Fig. S6), higher levels of serum IL-13 (Fig. 2B) and epithelium-derived IL-13-induced biomarkers (Fig. S6C), periostin (Fig. 2D) and Eno (Table 2). Of note, only 50% of these patients were atopic, less than the neutrophilic phenotype (68%) (Table 2), in keeping with the notion that T2 mechanisms are found in both atopic and non-atopic asthmatics. However, there were other, as yet unrecognised, associations. Differentially expressed genes were predicted to result from downregulated synoviolin 1 (Hrd1)-mediated signalling (Table S8), a pathway which may be partially responsible for eosinophilia. Hrd1 mediates clearance of misfolded proteins via Endoplasmic reticulum (ER)–associated degradation (ERAD)(10) and epithelial inflammation results from inhibition of Hrd1-associated ERAD(11). Under normal conditions, the Hrd1 ERAD complex degrades endoribonuclease IRE1α (a sensor of unfolded protein response); in response to ER stress, this degradation is inhibited and active IRE1α

accumulates, causing epithelial inflammation via inflammatory mediators such as JNK. Active IRE1α causes increased expression of Xbp1, which is highly activated during eosinophil commitment from granulocyte-monocyte progenitors (GMPs)(12). Additionally, the gene expression signatures of eosinophilic sub-phenotypes 1, 2 and 3 were predicted to result from activation of the nuclear retinoic acid receptor (RAR)-α.

Consistent with the knowledge of risk factors for asthma exacerbations(13), the eosinophilic phenotype was associated with a high prevalence of reported respiratory infections, use of long-acting β agonists and GORD. These patients had worse lung function and poor asthma symptom control, and 8% had been admitted to intensive care units because of exacerbations.

Ten proteins were strongly predictive of the eosinophilic phenotype. Alpha-1-antitrypsin neutralises neutrophil elastase that is known for its role in lung damage, bronchoconstriction and airway hyperreactivity, so its higher levels suggest negative feedback but also, possibly, interference with neutrophil-mediated anti-infection mechanisms. Serotransferrin is involved in transporting vitamin A (retinol) bound retinol-binding protein (RBP) and sequesters iron from invading bacteria, thereby inhibiting their ability to replicate and cause disease(14). It has been shown that amyloid fibrils from three different sources (α-synuclein, Sup35, and transthyretin) induce NADPH oxidase-dependent neutrophil extracellular traps (NETs) *in vitro* from human neutrophils(15). Transthyretin plasma levels have been reported to drop during systemic inflammation(16). We and others have previously observed reduced sputum transthyretin in COPD, a neutrophilic airways disease, and now show that in asthma the levels are increased in eosinophilic inflammation. Higher levels observed in asthmatics in the current study may also reflect corticosteroid use as suggested previously in an animal model(17).

The neutrophilic phenotype was observed in about a third of the asthmatics. Their symptoms were more likely triggered by dust or fungus, consistent with their high prevalence of atopy.

Lower levels of transthyretin observed in this phenotype, that were amongst the set of ten predictive proteins, is surprising considering the highest percentage of OCS use in the highly neutrophilic phenotype (Table 2). Other predictive proteins (Fig. 3 & S4E) included the neutrophil granular proteins, neutrophil gelatinase-associated lipocalin (NGAL), azurocidin, S100A9, and myeloperoxidase, all of which were increased. Annexin A1 and A3, coronin 1A and the chloride intracellular channel protein 1 were also increased. S100A9 is highly expressed in neutrophils, activated monocytes and differentiated macrophages. It has several functions in cellular inflammation, responding to intracellular Ca^{2+}. Extracellular S100 proteins act as damage-associated molecular pattern (DAMP) proteins, initiating pro-inflammatory immune responses. Annexin A1, a member of a large superfamily of glucocorticoid-regulated, calcium- and phospholipid-binding proteins, modulates neutrophil homeostasis and is an anti-inflammatory protein in innate immunity, modulating activation of several types of cells, including neutrophils. Annexin A1 inhibits NF-kB and blocks eicosanoid production by supressing phospholipase A2. Myeloperoxidase catalyses the formation of hypohalous acids that have antimicrobial properties. NGAL scavenges bacterial siderophores, thus depriving bacteria of iron. It is localised in azurophils and co-localises with myeloperoxidase(18). Azurocidin is a multifunctional inflammatory mediator with antimicrobial properties which binds endotoxin and is chemotactic for monocytes and macrophages(19). Sputum cell gene expression in the neutrophilic phenotype was predicted to result from higher activity of T1 cytokines, IL2 and IFNα (Fig. S6), and IL-1β, IL-17, IFN-γ and IL8. S100A9 protein drives neutrophilic inflammation in asthma, possibly by inducing IL-1β, IL-17 and IFN-γ. In response to inflammatory stimuli, recruited neutrophils release granular proteins, histones and chromatin DNA. These neutrophil extracellular structures (NETS) amplify the efficacy of antimicrobial substances from neutrophils by maintaining a high local concentration to degrade pathogens before engulfment. Amongst the eosinophilic

sub-phenotypes, Galectin-3-binding protein, LG3BP, was highest in sub-phenotype 3 (Table 3).

This study has also identified some important features outside the two main (eosinophilic and neutrophilic) phenotypes. IL-13 and MAPK1- mediated gene expression signatures were identified as increased relative to health in the highly atopic phenotype (Fig. S6C and E), suggestive of a T2 phenotype in this milder form of asthma characterised by lower sputum granulocyte counts. Clusterin, one of the biomarker proteins predictive of this phenotype (Fig. 3 & S4C), modulates NF-κB transcriptional activity(20), however, NF-κB-mediated gene expression was not upregulated in this study (Fig. S6L). The Interleukin-1 Receptor Antagonist Gene (IL1RN), a potent anti-inflammatory cytokine, was also activated in the highly atopic phenotype (Fig. S6B). The lack of eosinophilic inflammation in this phenotype may be also, in part, due to lower expression of the predictive biomarker protein, plastin-2 (Fig. 3B), which mediates priming of eosinophils(21). Similarly, lower expression of cofilin-1 and β-actin, proteins that are also involved in eosinophil priming, was predictive of classification of this highly atopic phenotype.

The use of TDA enabled an appreciation of the features of the spectrum of sub-phenotypes by examining the molecular and clinical characteristics across the TDA network. The order in which the clusters are shown in Table S9 (approximate left to right order) and Figure 5 (approximate circular order) should not be seen as absolute because there is some overlap between clusters and TDA networks do not have defined coordinates. Nevertheless, visualisation of the broad trends is helpful to explore the relative contribution of mechanisms across the whole network, something that is not possible with other methods like hierarchical clustering. For example, the type I interferon, Interferon-α and TNF, and KDM5B (JARID1B) demethylase were all identified as predicted upstream regulators of gene expression, seen as being increasingly activated from left to right in the TDA network, suggesting that the more

severe, granulocytic phenotypes are associated with viral and/or bacterial infections. Gene expression on the left side of the TDA network (sub-phenotypes 5, 6 and 4) was predicted by collagen type XVIII alpha 1 (COL18A1), shown to be associated with atopy(22). Furthermore, the shape of the TDA network revealed a broad gap between eosinophilic and neutrophilic asthma proteotypes (Fig. 1). However, as indicated by the links between the eosinophilic sub-phenotype 1 and the neutrophilic sub-phenotype 10 (see lines joining these in Fig. 1), the proteomes of the eosinophilic and neutrophilic sub-phenotypes can have some associations. In contrast, the absence of links between sub-phenotypes 4 and 10 indicated large proteomic differences. The highly atopic phenotypes (5 and 6) had the lowest asthma symptoms (ACQ) and the best lung function (FEV_1 78% of predicted), lower symptom severity, hence best quality of life, compared to the highly granulocytic phenotypes. At the opposite, right end of the network were sub-phenotypes 1 and 10), where asthma symptoms (ACQ) were most severe, suggesting that the asthma spectrum progresses in severity from the left to right ends of the network, worsening either through the neutrophilic or eosinophilic pathways. There was increasing IL-2-mediated and decreasing IL-13 gene expression from left to right across the TDA network, suggesting a shift from T2 to T1 mechanisms with increasing severity (Fig. 5). Additionally, there was a trend across the TDA structure (Table S9 & Fig. S6I) of gene expression associated with the upstream regulator, Brahma-related gene-1 (Brg1, also known as SMARCA4), a chromatin remodelling factor known to inhibit expression of CD44(23) and E-cadherin(24), drivers of the T2 phenotype(25).

Not surprising for a study of severe asthma, the use of OCS was high, with mean percentages of patients requiring OCS for disease control ranging from 27% in the eosinophilic sub-phenotype 3 to 58% in the neutrophilic sub-phenotype 9. Not surprisingly, OCS use was highest in patients with neutrophilic sub-phenotypes possibly due to the pro-neutrophilic effects of corticosteroids on neutrophil numbers, although the higher rates of intensive care

unit admission in these patients (Figure 2, panel L) suggest particularly severe pathogenetic mechanisms that result in the most severe forms of exacerbation. We speculate that in the atopic, predominantly pauci-leukocytic sub-phenotypes corticosteroids were effective at reducing eosinophilic inflammation, while the persistence of eosinophilic inflammation in sub-phenotypes 1-3 points to at least partial insensitivity to corticosteroids. Full elucidation of the mechanisms to explain the levels of granulocytic inflammation and responses to OCS requires appropriately designed mechanistic studies.

This study has limitations. The individual biomarkers and the sub-phenotypes need to be validated in a separate (validation) cohort and assessed for intra-subject reproducibility in a longitudinal study. Stratification of the highly eosinophilic, neutrophilic and atopic phenotypes into sub-phenotypes resulted in insufficient numbers of participants in the sub-phenotypes prohibiting the extensive analysis we were able to conduct on the phenotypes. Furthermore, like other studies involving large cohorts, we have not undertaken a stringent analysis of adherence to treatment, using methods like FENO suppression testing which have been recently validated (26) Such additional analyses, and application of biomarkers in mechanistic studies with new asthma drugs, especially the range of biologics available and others in development, could point to further associations with exacerbations and provide why some but not all patients benefit from individual biologics(27,28). As in all studies of severe asthma, treatments varied significantly between participants and these variations are likely to impact on the biomarker profiles. Perhaps of greatest relevance is the variable use of oral corticosteroids even though there is hope that the majority of patients will no longer be dependent on them as maintenance treatment.

ACKNOWLEDGEMENTS

This paper is presented on behalf of the U-BIOPRED Study Group with input from the U-BIOPRED Patient Input Platform, Ethics Board and Safety Management Board. We thank all the members of each recruiting centre for their dedicated effort, devotion, promptness and care in the recruitment and assessment of the participants in this study. U-BIOPRED is supported through an Innovative Medicines Initiative Joint Undertaking under grant agreement no. 115010, 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 (www.imi.europa.eu). We would also like to acknowledge help from the IMI funded eTRIKS project (EU Grant Code No.115446).

The members of the U-BIOPRED Study Group are as follows: H. Ahmed, European Institute for Systems Biology and Medicine, University of Lyon, France; D. Allen, North West Severe Asthma Network; Pennine Acute Hospital NHS Trust; P. Badorrek, Fraunhofer ITEM; S. Ballereau, European Institute for Systems Biology and Medicine, University of Lyon, France; F. Baribaud, Janssen R&D, USA; M.K. Batuwitage, Imperial College, London, UK; A. Bedding, Roche Diagnostics GmbH, Mannheim, Germany; A.F. Behndig, Umeå University; A. Berglind, Karolinska University Hospital and Karolinska Institutet; A. Berton, Boehringer Ingelheim Pharma GmbH & Co. KG; J. Bigler, Amgen Inc; M.J. Boedigheimer, Amgen Inc; K. Bønnelykke, University of Copenhagen and Danish Pediatric Asthma Center, Gentofte Hospital, University of Copenhagen, Denmark; P. Brinkman, Academic Medical Centre, University of Amsterdam; A. Bush, Department of Paediatrics and National Heart and Lung Institute, Imperial College, London; Department of Respiratory Paediatrics, Royal Brompton Hospital, London, UK; D. Campagna, University of Catania; C. Casaulta, University Children's Hospital Bern, Switzerland; A. Chaiboonchoe, European Institute for Systems Biology and Medicine, University of Lyon, France; T. Davison, Janssen R&D, USA; B. De

Meulder, European Institute for Systems Biology and Medicine, University of Lyon, France; I. Delin, Institute of Environmental Medicine, Karolinska Instituet, Stockholm, Sweden; P. Dennison, NIHR Southampton Respiratory Biomedical Research Unit and University of Southampton; P. Dodson, AstraZeneca, Mölndal, Sweden; L. El Hadjam, European Institute for Systems Biology and Medicine, University of Lyon, France; D. Erzen, Boehringer Ingelheim Pharma GmbH & Co. KG; C. Faulenbach, Fraunhofer ITEM; K. Fichtner, Boehringer Ingelheim Pharma GmbH & Co. KG; N. Fitch, BioSci Consulting, Belgium; E. Formaggio, PhD, Project manager, Verona Italy; M. Gahlemann, Boehringer Ingelheim (Schweiz) GmbH; G. Galffy, Semmelweis University, Budapest, Hungary; D. Garissi, Global Head Clinical Research Division, CROMSOURCE, Italy; T. Garret, BioSci Consulting, Belgium; J. Gent, Royal Brompton and Harefield NHS Foundation Trust; E. Guillmant-Farry, Royal Brompton Hospital, London, UK; E. Henriksson, Karolinska Institutet; U. Hoda, Imperial College; J.M. Hohlfeld, Fraunhofer ITEM; X. Hu, Amgen Inc; A. James, Karolinska Institutet; K. Johnson, Centre for respiratory medicine and allergy, Institute of Inflammation and repair, University Hospital of South Manchester, NHS Foundation Trust, Manchester, UK; N. Jullian, European Institute for Systems Biology and Medicine, University of Lyon, France; G. Kerry, Centre for respiratory medicine and allergy, Institute of Inflammation and repair, University Hospital of South Manchester, NHS Foundation Trust, Manchester, UK; M. Klüglich, Boehringer Ingelheim Pharma GmbH & Co. KG; R. Knowles, Arachos Pharma, Stevenge, UK; J.R. Konradsen, Karolinska University Hospital and Karolinska Institutet; K. Kretsos, UCB, Slough, UK; L. Krueger, University Children's Hospital Bern, Switzerland; A-S. Lantz, Karolinska University Hospital and Karolinska Institutet; C. Larminie, GSK, London, UK; P. Latzin, University Children's Hospital Bern, 3010 Bern, Switzerland; D. Lefaudeux, European Institute for Systems Biology and Medicine, University of Lyon, France; N. Lemonnier, European Institute for Systems Biology and Medicine, University of

Lyon, France; L.A. Lowe, Centre for respiratory medicine and allergy, Institute of Inflammation and repair, University Hospital of South Manchester, NHS Foundation Trust, Manchester, UK; R. Lutter, Academic Medical Centre, University of Amsterdam; A. Manta, Roche Diagnostics GmbH, Mannheim, Germany; A. Mazein, European Institute for Systems Biology and Medicine, University of Lyon, France; L. McEvoy, University Hospital, Department of Pulmonary Medicine, Bern, Switzerland; A. Menzies-Gow, Royal Brompton and Harefield NHS Foundation Trust; N. Mores, Università Cattolica del Sacro Cuore; C.S. Murray, Centre for Respiratory Medicine and Allergy, The University of Manchester, Manchester Academic Health Science Centre, University Hospital of South Manchester NHS Foundation Trust, Manchester, UK; K. Nething, Boehringer Ingelheim Pharma GmbH & Co. KG; U. Nihlén, Department of Respiratory Medicine and Allergology, Skåne University Hospital, Lund, Sweden; AstraZeneca R&D, Mölndal, Sweden; R. Niven, North West Severe Asthma Network, University Hospital South Manchester NHS Trust; B. Nordlund, Astrid Lindgren Children's Hospital, Karolinska University Hospital, Stockholm, Sweden; Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden; S. Nsubuga, Royal Brompton Hospital, London, UK; J. Pellet, European Institute for Systems Biology and Medicine, University of Lyon, France; C. Pison, European Institute for Systems Biology and Medicine, University of Lyon, France; G. Praticò, CROMSOURCE, Verona, Italy; M. Puig Valls, CROMSOURCE, Barcelona, Spain; K. Riemann, Boehringer Ingelheim Pharma GmbH & Co. KG; J.P. Rocha, Royal Brompton and Harefield NHS Foundation Trust; C. Rossios, Imperial College; G. Santini, Università Cattolica del Sacro Cuore; M. Saqi, European Institute for Systems Biology and Medicine, University of Lyon, France; S. Scott, North West Severe Asthma Network; Countess of Chester NHS Trust; N. Sehgal, North West Severe Asthma Network; Pennine Acute Hospital NHS Trust; A. Selby, NIHR Southampton Respiratory Biomedical Research Unit, Clinical and Experimental Sciences and Human

Development and Health, Southampton, UK; P. Söderman, Astrid Lindgren Children's Hospital, Karolinska University Hospital, Stockholm, Sweden; Department of Women's and Children's Health, Stockholm, Sweden; A. Sogbesan, Royal Brompton and Harefield NHS Foundation Trust; F. Spycher, University Hospital, Department of Pulmonary Medicine, Bern, Switzerland; S. Stephan, Centre for respiratory medicine and allergy, Institute of Inflammation and repair, University Hospital of South Manchester, NHS Foundation Trust, Manchester, UK; J. Stokholm, University of Copenhagen and Danish Pediatric Asthma Center, Gentofte Hospital, University of Copenhagen, Denmark; M. Sunther, Centre for respiratory medicine and allergy, Institute of Inflammation and repair, University Hospital of South Manchester, NHS Foundation Trust, Manchester, UK; M. Szentkereszty, Semmelweis University, Budapest, Hungary; L. Tamasi, Semmelweis University, Budapest, Hungary; K. Tariq, NIHR Southampton Respiratory Biomedical Research Unit and University of Southampton; S. Valente, Università Cattolica del Sacro Cuore; W.M. van Aalderen, Academic Medical Centre, University of Amsterdam; C.M. van Drunen, Academic Medical Centre, University of Amsterdam; J. Van Eyll, UCB, Slough, UK; A. Vyas, North West Severe Asthma Network; Lancashire Teaching Hospitals NHS Trust; W. Yu, Amgen Inc; W. Zetterquist, Department of Woman and Child Health, Karolinska Institutet, Department of Woman and Child Health, Karolinska Institutet, Stockholm, Sweden; Z. Zolkipli, NIHR Southampton Respiratory Biomedical Research Unit, University Hospital Southampton NHS Foundation Trust, Southampton, UK; Clinical and Experimental Sciences and Human Development in Health Academic Unit, University of Southampton Faculty of Medicine, Southampton, UK; The David Hide Asthma and Allergy Research Centre, St Mary’s Hospital, Isle of Wight, UK; A.H. Zwinderman, Academic Medical Centre, University of Amsterdam.

The U-BIOPRED consortium wishes to acknowledge the help and expertise of the following individuals and groups without whom, the study would not have been possible.

Investigators and contributors: Nora Adriaens, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands; Antonios Aliprantis, Merck Research Laboratories, Boston, USA; Kjell Alving, Dept Women's and Children's Health, Uppsala University, Uppsala, Sweden; Per Bakke, Department of Clinical Science, University of Bergen, Bergen, Norway; David Balgoma, Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden; Clair Barber, NIHR Southampton Respiratory Biomedical Research Unit and Clinical and Experimental Sciences, Southampton, UK; Frédéric Baribaud, Janssen R&D, USA; Stewart Bates, Respiratory Therapeutic Unit, GSK, London, UK; An Bautmans, MSD, Brussels, Belgium; Jorge Beleta, Almirall S.A., Barcelona, Spain; Grazyna Bochenek, II Department of Internal Medicine, Jagiellonian University Medical College, Krakow, Poland; Joost Brandsma, University of Southampton, Southampton, UK; Armin Braun, Fraunhofer Institute for Toxicology and Experimental Medicine, Hannover, Germany; Dominic Burg, Centre for Proteomic Research, Institute for Life Sciences, University of Southampton, Southampton, UK; Leon Carayannopoulos, previously at: MSD, USA; João Pedro Carvalho da Purificação Rocha, Royal Brompton and Harefield NHS Foundation Trust, London, UK; Romanas Chaleckis, Centre of Allergy Research, Karolinska Institutet, Stockholm, Sweden; Arnaldo D'Amico, University of Rome 'Tor Vergata', Rome Italy; Jorge De Alba, Almirall S.A., Barcelona, Spain; Inge De Lepeleire, MSD, Brussels, Belgium; Tamara Dekker, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands; Annemiek Dijkhuis, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands; Aleksandra Draper, BioSci Consulting, Maasmechelen, Belgium; Jessica Edwards, Asthma UK, London, UK; Rosalia Emma, Department of Clinical and Experimental Medicine, University of Catania, Catania, Italy; Magnus Ericsson, Karolinska University Hospital, Stockholm, Sweden; Breda Flood, European Federation of Allergy and Airways Diseases Patient's Associations, Brussels, Belgium; Hector Gallart,

Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden; Cristina Gomez, Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden; Kerry Gove, NIHR Southampton Respiratory Biomedical Research Unit and Clinical and Experimental Sciences, Southampton, UK; Neil Gozzard, UCB, Slough, UK; John Haughney, International Primary Care Respiratory Group, Aberdeen, Scotland; Lorraine Hewitt, NIHR Southampton Respiratory Biomedical Research Unit, Southampton, UK; Jens Hohlfeld, Fraunhofer Institute for Toxicology and Experimental Medicine, Hannover, Germany; Cecile Holweg, Respiratory and Allergy Diseases, Genentech, San Francisco, USA; Richard Hu, Amgen Inc. Thousand Oaks, USA; Sile Hu, National Heart and Lung Institute, Imperial College, London, UK; Juliette Kamphuis, Longfonds, Amersfoort, The Netherlands; Erika J. Kennington, Asthma UK, London, UK; Dyson Kerry, CromSource, Stirling, UK; Hugo Knobel, Philips Research Laboratories, Eindhoven, The Netherlands; Johan Kolmert, Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden; Maxim Kots, Chiesi Pharmaceuticals, SPA, Parma, Italy; Scott Kuo, National Heart and Lung Institute, Imperial College, London, UK; Maciej Kupczyk, Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden; Bart Lambrecht, University of Gent, Gent, Belgium; Saeeda Lone-Latif, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands; Matthew J. Loza, Janssen R&D, USA; Lisa Marouzet, NIHR Southampton Respiratory Biomedical Research Unit, Southampton, UK; Jane Martin, NIHR Southampton Respiratory Biomedical Research Unit, Southampton, UK; Sarah Masefield, European Lung Foundation, Sheffield, UK; Caroline Mathon, Centre of Allergy Research, Karolinska Institutet, Stockholm, Sweden; Sally Meah, National Heart and Lung Institute, Imperial College, London, UK; Andrea Meiser, Data Science Institute, Imperial College, London, UK; Leanne Metcalf, previously at: Asthma UK, London, UK; Maria Mikus, Science for Life Laboratory and The Royal Institute of Technology, Stockholm, Sweden; Montse Miralpeix, Almirall, Barcelona, Spain; Philip

Monk, Synairgen Research Ltd, Southampton, UK; Shama Naz, Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden; Ben Nicholas, University of Southampton, Southampton, UK; Peter Nilsson, Science for Life Laboratory and The Royal Institute of Technology, Stockholm, Sweden; Jörgen Östling, AstraZeneca, Mölndal, Sweden; Antonio Pacino, Lega Italiano Anti Fumo, Catania, Italy; Susanna Palkonen, European Federation of Allergy and Airways Diseases Patient's Associations, Brussels, Belgium; Stelios Pavlidis, National Heart and Lung Institute, Imperial College, London, UK; Giorgio Pennazza, University of Rome 'Tor Vergata', Rome Italy; Anne Petrén, Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden; Sandy Pink, NIHR Southampton Respiratory Biomedical Research Unit, Southampton, UK; Anthony Postle, University of Southampton, UK; Pippa Powell, European Lung Foundation, Sheffield, UK; Malayka Rahman-Amin, Previously at: Asthma UK, London, UK; Navin Rao, Janssen R&D, USA; Lara Ravanetti, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands; Emma Ray, NIHR Southampton Respiratory Biomedical Research Unit, Southampton, UK; Stacey Reinke, Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden; Leanne Reynolds, previously at: Asthma UK, London, UK; John Riley, Respiratory Therapeutic Unit, GSK, London, UK; Martine Robberechts, MSD, Brussels, Belgium; Amanda Roberts, Asthma UK, London, UK; Kirsty Russell, National Heart and Lung Institute, Imperial College, London, UK; Michael Rutgers, Longfonds, Amersfoort, The Netherlands; Marco Santoninco, University of Rome 'Tor Vergata', Rome Italy; Corinna Schoelch, Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany; James P.R. Schofield, Centre for Proteomic Research, Institute for Life Sciences, University of Southampton, Southampton, UK; Marcus Sjödin, Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden; Paul J. Skipp, Centre for Proteomic Research, Institute for Life Sciences, University of Southampton, Southampton, UK; Barbara Smids, Academic Medical Centre, University of

Amsterdam, Amsterdam, The Netherlands; Caroline Smith, NIHR Southampton Respiratory Biomedical Research Unit, Southampton, UK; Jessica Smith, Asthma UK, London, UK; Katherine M. Smith, University of Nottingham, UK; Doroteya Staykova, University of Southampton, Southampton, UK; Kai Sun, Data Science Institute, Imperial College, London, UK; John-Olof Thörngren, Karolinska University Hospital, Stockholm, Sweden; Bob Thornton, MSD, USA; Jonathan Thorsen, COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark; Marianne van de Pol, Academic Medical Centre, University of Amsterdam, Amsterdam ,The Netherlands; Marleen van Geest, AstraZeneca, Mölndal, Sweden; Jenny Versnel, previously at: Asthma UK, London, UK; Anton Vink, Philips Research Laboratories, Eindhoven, The Netherlands; Frans Wald, Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany; Samantha Walker, Asthma UK, London, UK; Jonathan Ward, Histochemistry Research Unit, Faculty of Medicine, University of Southampton, Southampton, UK; Zsoka Weiszhart, Semmelweis University, Budapest, Hungary; Kristiane Wetzel, Boehringer Ingelheim Pharma GmbH, Biberach, Germany; Craig E. Wheelock, Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden; Coen Wiegman, National Heart and Lung Institute, Imperial College, London, UK; Siân Williams, International Primary Care Respiratory Group, Aberdeen, Scotland; Susan J. Wilson, Histochemistry Research Unit, Faculty of Medicine, University of Southampton, Southampton, UK; Ashley Woodcock, Centre for Respiratory Medicine and Allergy, Institute of Inflammation and Repair, University of Manchester and University Hospital of South Manchester, Manchester Academic Health Sciences Centre, Manchester, UK; Xian Yang, Data Science Institute, Imperial College, London, UK; Elizabeth Yeyasingham, UK Clinical Operations, GSK, Stockley Park, UK.

ACCEPTED MANUSCRIPT

Partner organisations: Novartis Pharma AG; University of Southampton, Southampton, UK; Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands; Imperial College London, London, UK; University of Catania, Catania, Italy; University of Rome 'Tor Vergata', Rome, Italy; Hvidore Hospital, Hvidore, Denmark; Jagiellonian Univ. Medi.College, Krakow, Poland; University Hospital, Inselspital, Bern, Switzerland; Semmelweis University, Budapest, Hungary; University of Manchester, Manchester, UK; Université d'Aix-Marseille, Marseille, France; Fraunhofer Institute, Hannover, Germany; University Hospital, Umea, Sweden; Ghent University, Ghent, Belgium; Ctr. Nat. Recherche Scientifique, Villejuif, France; Università Cattolica del Sacro Cuore, Rome, Italy; University Hospital, Copenhagen, Denmark; Karolinska Institutet, Stockholm, Sweden; Nottingham University Hospital, Nottingham, UK; University of Bergen, Bergen, Norway; Netherlands Asthma Foundation, Leusden, NL; European Lung Foundation, Sheffield, UK; Asthma UK, London, UK; European Fed. of Allergy and Airways Diseases Patients' Associations, Brussels, Belgium; Lega Italiano Anti Fumo, Catania, Italy; International Primary Care Respiratory Group, Aberdeen, Scotland; Philips Research Laboratories, Eindhoven, NL; Synairgen Research Ltd, Southampton, UK; Aerocrine AB, Stockholm, Sweden; BioSci Consulting, Maasmechelen, Belgium; Almirall; AstraZeneca; Boehringer Ingelheim; Chiesi; GlaxoSmithKline; Roche; UCB; Janssen Biologics BV; Amgen NV; Merck Sharp & Dohme Corp.

Third Parties to the project, contributing to the clinical trial: Academic Medical Centre (AMC), Amsterdam (In the U-BIOPRED consortium the legal entity is AMC Medical Research BV (AMR); AMR is a subsidiary of both AMC and the University of Amsterdam; AMC contribute across the U-BIOPRED project); University Hospital Southampton NHS Trust (third party of the University of Southampton and contributor to the U-BIOPRED clinical trial); South Manchester Healthcare Trust (third party to the University of Manchester,

South Manchester Healthcare Trust, contributor to the U-BIOPRED clinical trial and to the U-BIOPRED Biobank); Protisvalor Méditerranée SAS (third party to University of the Mediterranean; contributor to the U-BIOPRED clinical trial); Karolinska University Hospital (third party Karolinska Institutet (KI), contributor to the U-BIOPRED clinical trial); Nottingham University Hospital (third party to University of Nottingham, contributor to the U-BIOPRED clinical trial); NIHR-Wellcome Trust Clinical Research Facility.

Members of the ethics board: Jan-Bas Prins, biomedical research, LUMC, the Netherlands; Martina Gahlemann, clinical care, BI, Germany; Luigi Visintin, legal affairs, LIAF, Italy; Hazel Evans, paediatric care, Southampton, UK; Martine Puhl, patient representation (co-chair), NAF, the Netherlands; Lina Buzermaniene, patient representation, EFA, Lithuania; Val Hudson, patient representation, Asthma UK; Laura Bond, patient representation, Asthma UK; Pim de Boer, patient representation and pathobiology, IND; Guy Widdershoven, research ethics, VUMC, the Netherlands; Ralf Sigmund, research methodology and biostatistics, BI, Germany.

The patient input platform: Amanda Roberts, UK; David Supple (chair), UK; Dominique Hamerlijnck, The Netherlands; Jenny Negus, UK; Juliëtte Kamphuis, The Netherlands; Lehanne Sergison, UK; Luigi Visintin, Italy; Pim de Boer (co-chair), The Netherlands; Susanne Onstein, The Netherlands.

Members of the safety monitoring board: William MacNee, clinical care; Renato Bernardini, clinical pharmacology; Louis Bont, paediatric care and infectious diseases; Per-Ake Wecksell, patient representation; Pim de Boer, patient representation and pathobiology (chair); Martina Gahlemann, patient safety advice and clinical care (co-chair); Ralf Sigmund, bio-informatician.

This work was partially funded by the Engineering and Physical Sciences Research Council, UK (EP/N014189: Joining the Dots, from Data to Insight).

ACCEPTED MANUSCRIPT

568 Instrumentation in the Centre for Proteomic Research is supported by the BBSRC
569 (BM/M012387/1) and the Wessex Medical Trust.
570 We thank Ayasdi Inc. for use of, and support with, the Ayasdi TDA software.
571

ACCEPTED MANUSCRIPT

FIGURE LEGENDS

572

573 **Figure 1.** Asthma sputum proteomes/proteotypes/sub-phenotypes. The data network was
574 created by topological data analysis (TDA) of protein datasets from all asthmatic
575 (mild/moderate and severe) participants and consisted of 10 clusters that we termed
576 "proteotypes". Differently coloured nodes and edges denote different clusters. Dotted
577 boundaries represent clusters grouped according to shared granulocyte and atopy profiles,
578 while the proteome clusters delineated by continuous boundaries represent sub-phenotypes.
579 Connections (lines) between nodes represent overlap of patients between nodes/clusters. The
580 percentage of participants in each phenotype is displayed adjacent to the phenotype name.

581

582 **Figure 2**. Pathobiological, clinical and protein features associated with proteomes identified
583 in asthmatics. Eosinophilic phenotype patients, circled red in panel A, had elevated blood
584 periostin (D), IL13 protein (B), and sputum haptoglobin (C). The highly atopic phenotype,
585 circled in yellow in panel E, had high levels of sputum uteroglobin and clusterin proteins (E,
586 G), high total IgE and oral corticosteroid dose (F, H). The neutrophilic sub-phenotypes,
587 circled blue in panel I, had higher concentrations of sputum S100-A9 and matrix
588 metalloproteinase-9 (MMP-9) proteins (J, K). The number of ICU admissions in the last 12
589 months was highest in the neutrophilic phenotype. Colours denote the concentrations of the
590 individual variables, ranging from blue (low) to red (high) – see the vertical intensity bar
591 alongside each panel.

592

593 **Figure 3.** Sputum proteins shown by logistic regression and ROC analysis to be most
594 predictive of **A)** the eosinophilic, **B)** highly atopic and **C)** neutrophilic phenotypes.
595 Expression is normalised to mean expression in all asthma participants samples (set to 1).

596

597

598 **Figure 4.** Selected top upstream regulators of gene expression across the sub-phenotypes of
599 eosinophilic, highly atopic and neutrophilic phenotypes of asthma. The sequence of sub-
600 phenotypes shown in an approximate circular order, beginning with the highly eosinophilic
601 sub-phenotypes (clusters) 1, 2 and 3, moving through the highly allergic sub-phenotypes 5
602 and 6, and ending with the highly neutrophilic sub-phenotypes.

603

604 **Figure 5.** The pattern across the TDA structure of activation of IL-13 and IL-2, upstream
605 regulators of gene expression, representative of T2 and T1, respectively. Also shown are
606 arrows indicating increasing neutrophil and eosinophil cell counts.

607

ACCEPTED MANUSCRIPT

TABLE LEGENDS

608 **Table 1.** Participant demographics. Data are presented as mean±SD unless otherwise stated.
609 BMI: body mass index; FEV1: forced expiratory volume in 1 second; FVC: forced vital
610 capacity; ICU: intensive care unit; NA: not applicable. @Hydrocortisone and triamcinolone
611 doses were converted to equivalent prednisolone dose. ACQ: Asthma Control Questionnaire;
612 AQLQ: Asthma Quality of Life Questionnaire.
613
614 **Table 2.** Clinical features associated with the eosinophilic and neutrophilic asthma sub-
615 phenotypes. Highest variable value is in bold, lowest value in italics. ACQ5 or 7: asthma
616 quality questionnaire consisting of 5 or 7 questions. AQLQ: asthma quality of life
617 questionnaire.
618
619 **Table 3.** Sputum protein differences between sub-phenotypes within the eosinophilic, highly
620 atopic and neutrophilic asthma phenotypes identified by logistic regression.
621
622
623

ACCEPTED MANUSCRIPT

624 **TABLES**

Table 1

	Severe non-smoking asthma	*Smokers and ex-smokers with severe asthma*	*Mild/moderate non-smoking asthma*	*Healthy non-smoking controls*
Subjects (n)	117	45	38	40
Age (years)	52.88±13.07	55.2±10.25	42.37±15.23	37.2±13.11
Age at diagnosis (years)	20.32±16.42	33.53±19.87	16.23±16.85	NA
Sex ratio (M/F)	43/74	18/27	21/17	28/12
BMI kg.m^{-2}	28.98±6.61	30.1±6.47	25.44±4.61	25.61±3.19
BMI >30kg.m^{-2}	46	21	9	4
Serum IgE IU.ml^{-1}	274.88±440.53	385.09±1046.47	326.32±660.16	87.66±140.24
FEV1% pred.	65.44±21.68	65.33±17.2	90.68±18.48	101.89±12.72
FCV% pred.	87.67±19.89	89.67±16.8	107.91±18.15	108.71±12.91
FEV1/FVC ratio	60.11±13.45	59.35±11.1	70.06±10.87	NA
Exacerbations in previous year	2.27±1.79	2.6±2.46	0.45±0.93	NA
Smoking history pack-years	2.46 (0-5)	23.7 (5-70)	3.62 (1-5)	2.25 (0-5)
Intubation (ever)	13	1	0	NA
ICU admission (ever)	28	7	0	NA
Atopy test positive	83	26	39	14
Oral corticosteroid	48 (40.6%)	24 (50%)	0 (0%)	
Prednisolone (equ.) mg	11.81±6.94	13.75±8.95	0±0	
Inhaled corticosteroids	114 (96.6%)	44 (97.9%)	38 (100%)	
Long-acting β-agonist	113 (95.7%)	45 (93.7%)	1 (2.5%)	
Short-acting β-agonist	92 (77.9%)	35 (72.9%)	30 (75%)	
Injected corticosteroids[@]	8 (6.7%)	0 (0%)	0 (0%)	
Long-acting muscarinic antagonist	30 (25.4%)	13 (27%)	0 (0%)	
Short-acting muscarinic antagonist	51 (43.2%)	22 (45.8%)	0 (0%)	
ACQ				
Mean ACQ5	2.21±1.18	2.13±1.13	0.81±0.68	
Mean ACQ7	2.56±1.25	2.55±1.04	0.95±0.65	
AQLQ				
Total	4.55±1.22	4.54±1.14	5.86±1.05	
Symptoms	4.51±1.31	4.53±1.21	5.88±1.1	
Emotional	4.65±1.61	4.7±1.52	6±1.21	
Environmental stimuli	4.83±1.45	4.55±1.45	5.59±1.36	
Activity limitation	4.45±1.26	4.48±1.22	5.86±1	

Table 2

Asthma Phenotype	Eosinophilic					Highly atopic				Neutrophilic				All Asthma	Healthy
Asthma Sub-phenotype	**1**	**2**	**3**	**1, 2, 3**	**4**	**5**	**6**	**5, 6**	**7**	**8**	**9**	**10**	**8, 9, 10**		
Number of participants	13	38	11	**62**	22	18	32	50	*9*	27	19	11	59	200	40
Percentage of asthmatics	6.5%	19.0%	5.5%	**31.00 %**	11.0%	9.0%	16.0%	25.0%	*4.5%*	13.5%	9.50%	5.50%	29.5%		
Age (yr)	52.46 ±13.43	52.57 ±13.15	**59.36 ±10.84**	53.75 ±12.91	51.31 ±11.4	48.5 ±16.34	48.87 ±15.2	48.74 ±15.42	55.44 ±8.86	*47.62 ±15.31*	52.66 ±12.19	52.81 ±13.49	50.38 ±13.94	51.34 ±13.64	37.20 ±13.11
Smoking (pack-years)	3.35 ±8.52	6.72 ±15.87	2.88 ±4.75	5.33 ±13.19	9.68 ±17.59	*1.78 ±4.52*	4.9 ±9.64	3.78 ±8.19	**11.66 ±17.24**	9.11 ±15.32	6.64± 9.05	1.88 ±3.96	5.86 ±11.83	5.86 ±12.58	0.28 ±1.09
Mean ACQ 5	2.16 ±1.34	1.82 ±1.4	1.96 ±1.66	1.91 ±1.42	1.64 ±1.09	*1.16 ±0.96*	1.51 ±1.14	1.38 ±1.09	1.66 ±1.42	1.95 ±1.25	**2.3 ±1.48**	1.49 ±1.25	1.99 ±1.34	1.76 ±1.30	0.01 ±0.06
Mean ACQ 7	2.30 ±1.35	2.07 ±1.53	2.25 ±1.81	2.15 ±1.53	2.05 ±1.19	*1.35 ±1.08*	1.82 ±1.19	1.65 ±1.17	2.03 ±1.48	2.24 ±1.28	**2.78 ±1.66**	1.66 ±1.23	2.32 ±1.45	2.06 ±1.39	0.01 ±0.06
Mean AQLQ	4.76 ±1.26	4.3 ±2.09	*3.08 ±2.29*	4.18 ±2.03	4.78 ±1.62	**4.99 ±1.58**	4.02 ±2.25	4.37 ±2.14	3.87 ±2.1	4.55 ±1.51	3.89 ±1.88	4.33 ±1.77	4.28 ±1.69	4.31 ±1.90	6.95 ±0.08
Admitted to ICU (%)	7.69%	10.53 %	*0.00%*	8.06%	9.09%	22.22 %	18.75 %	20.00 %	11.11 %	22.22 %	21.05%	**63.64 %**	28.81%	17.00%	0.00%
Oral steroids (%)	30.76 %	28.95 %	27.27 %	29.03 %	36.36 %	*22.22 %*	28.13 %	26.00 %	*22.22 %*	33.33 %	**57.89%**	45.45 %	42.37%	33.00%	0.00%
Blood Periostin (ng/mL)	**53.97 ±31.72**	52.73 ±26.63	50.52 ±33.14	52.6 ±28.45	35.88 ±20.27	41.72 ±33.65	39.23 ±17.57	40.12 ±24.08	*31.81 ±24.41*	45.72 ±16.38	49.69± 11.15	39.51 ±21.89	40.94 ±20.82	43.36 ±24.98	38.85 ±19.61
FEV1 (% of predicted)	65.88 ±26.01	70.19 ±23.01	*52.16 ±13.15*	66.08 ±22.98	73.11 ±17.99	**86.51 ±25.24**	72.97 ±21.31	77.84 ±23.37	67.73 ±24.01	69.42 ±18.76	62.36 ±25.83	69.17 ±18.22	66.00 ±21.52	69.81 ±22.61	101.89 ±12.72
FVC (% of predicted)	87.36 ±20.23	92.41 ±21.33	*77.16 ±15.29*	88.64 ±20.68	94.73 ±21.41	**103.12 ±24.49**	92.18 ±18.52	96.12 ±21.15	89.05 ±14.63	95.26 ±19.04	86.37 ±22.28	90.99 ±15.56	90.69 ±19.78	91.77 ±20.47	108.71 ±12.91
Atopy (% positive)	53.84 %	*47.37 %*	81.82 %	54.84 %	50.00 %	**88.89 %**	71.88 %	78.00 %	55.56 %	77.78 %	21.05%	72.73 %	67.80%	63.86%	32.50%
Exhaled NO (parts per billion	37.48 ±39.12	**48.77 ±36.83**	37.63 ±27.47	44.43 ±35.75	42.11 ±45.76	44.58 ±43.17	*18.67 ±11.5*	28 ±29.54	27 ±20.33	25.87 ±13.65	29.72± 30.68	31.09 ±26.89	27.43 ±22.48	34.36 ±32.4	19.56 ±15.00
Sputum eosinophils (% of inflammatory cell count)	18.74 ±20.21	**23.08 ±23.77**	17.03 ±20.06	21.1 ±22.26	7.74 ±13.86	11.3 ±17.85	6.05 ±13.06	7.94 ±14.89	2.68 ±4.36	*1.57 ±1.49*	2.48 ±4.55	6.26 ±9.58	2.77 ±5.19	10.21 ±17.06	0.28 ±0.53
Sputum neutrophils (% of inflammatory cell count)	63.88 ±27.18	42.27 ±23.32	46.1 ±22.07	47.48 ±25.09	49.4 ±22.81	*40.69 ±23.37*	45.79 ±19.85	43.96 ±21.59	46.22 ±25.57	63.51 ±19.87	**82.03 ±23.71**	72.58 ±18.26	71.79 ±22.32	53.86 ±25.69	38.52 ±24.05
Sputum macrophages (% of inflammatory cell count)	16.63 ±21.58	33.09 ±24.1	34.68 ±16.89	29.92 ±23.21	41.45 ±25.23	47.37 ±21.63	46.58 ±21.48	46.86 ±22.12	**50.12 ±25.75**	33.38 ±19.29	*14.76 ±23.*3	20.45 ±15.61	24.34 ±21.68	34.64 ±24.37	59.6 ±24.36

Table 3

Phenotype	Sub-phenotype Comparison	Protein ID	Protein name	t-test p-value	Fold change (log2)
Eosinophilic	1 v 2 & 3	NGAL_HUMAN	Neutrophil gelatinase-associated lipocalin (NGAL)	8.00E-10	0.1
		G3P_HUMAN	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	2.00E-09	0.12
	2 v 1 & 3	HV320_HUMAN	Immunoglobulin heavy variable 3-20	2.00E-03	0.07
		TCO1_HUMAN	Transcobalamin-1 (TC-1)	3.00E-03	0.06
		MYH13_HUMAN	Myosin-13 (Myosin heavy chain 13)	7.00E-03	0.11
		LDHA_HUMAN	L-lactate dehydrogenase A chain (LDH-A)	1.00E-02	-0.06
		TPIS_HUMAN	Triosephosphate isomerase (TIM)	1.00E-02	-0.06
		MYH7_HUMAN	Myosin-7 (Myosin heavy chain 7)	2.00E-02	0.12
		CFAB_HUMAN	Complement factor B	3.00E-02	0.02
		AL3B1_HUMAN	Aldehyde dehydrogenase family 3 member B1	3.00E-02	0.05
		ILEU_HUMAN	Leukocyte elastase inhibitor (LEI) (Monocyte/neutrophil elastase inhibitor)	4.00E-02	-0.05
	3 v 1 & 2	CLUS_HUMAN	Clusterin (Complement cytolysis inhibitor)	2.00E-04	0.05
		LG3BP_HUMAN	Galectin-3-binding protein	9.00E-04	0.04
		PROL4_HUMAN	Proline-rich protein 4	2.00E-03	0.07
Highly atopic	5 v 6	LYSC_HUMAN	Lysozyme C	5.00E-03	-0.01
Neutrophilic	8 v 9 & 10	TKT_HUMAN	Transketolase (TK)	4.00E-11	-0.07
		CATA_HUMAN	Catalase	3.00E-10	-0.1
		PEDF_HUMAN	Pigment epithelium-derived factor (PEDF)	7.00E-08	0.08
		LG3BP_HUMAN	Galectin-3-binding protein (Basement membrane autoantigen p105)	2.00E-07	0.05
	9 v 8 & 10	BPIB1_HUMAN	BPI fold-containing family B member 1	5.00E-04	-0.01
		MUC1_HUMAN	Mucin-1 (MUC-1)	3.00E-03	-0.04
		CFAB_HUMAN	Complement factor B (C3/C5 convertase) (Glycine-rich beta glycoprotein) (GBG)	4.00E-03	0.04
		ALDOA_HUMAN	Fructose-bisphosphate aldolase A (Lung cancer antigen NY-LU-1)	4.00E-03	0.04
		ZG16B_HUMAN	Zymogen granule protein 16 homolog B	6.00E-03	-0.06
	10 v 8 & 9	IGHG1_HUMAN	Immunoglobulin heavy constant gamma 1 (Ig gamma-1 chain C region)	2.00E-06	0.03
		HEMO_HUMAN	Hemopexin (Beta-1B-glycoprotein)	4.00E-05	0.06
		A1AT_HUMAN	Alpha-1-antitrypsin	3.00E-04	0.03
		FIBG_HUMAN	Fibrinogen gamma chain	7.00E-04	0.09
		TKT_HUMAN	Transketolase (TK)	1.00E-03	0.05
		TALDO_HUMAN	Transaldolase	2.00E-03	0.06

REFERENCES

1. Pelaia G, Vatrella A, Maselli R. The potential of biologics for the treatment of asthma. Nat Rev Drug Discov. 2012;11(12):958.

2. Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. Nat Med. 2012;18(5):716.

3. Szefler SJ, Wenzel S, Brown R, Erzurum SC, Fahy J V, Hamilton RG, et al. Asthma outcomes: biomarkers. J Allergy Clin Immunol. 2012/03/14. 2012;129(3 Suppl):S9-23.

4. Price D, Wilson AM, Chisholm A, Rigazio A, Burden A, Thomas M, et al. Predicting frequent asthma exacerbations using blood eosinophil count and other patient data routinely available in clinical practice. J Asthma Allergy. 2016;9:1.

5. Green RH, Brightling CE, McKenna S, Hargadon B, Parker D, Bradding P, et al. Asthma exacerbations and sputum eosinophil counts: a randomised controlled trial. Lancet. 2002;360(9347):1715–21.

6. Shaw DE, Sousa AR, Fowler SJ, Fleming LJ, Roberts G, Corfield J, et al. Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. Eur Respir J. 2015;46(5):1308–21.

7. Burg D, Schofield JPR, Brandsma J, Staykova D, Folisi C, Bansal A, et al. Large-scale label-free quantitative mapping of the sputum proteome. J Proteome Res. 2018;17(6):2072–91.

8. Kuo C-HS, Pavlidis S, Loza M, Baribaud F, Rowe A, Pandis I, et al. T-helper cell type 2 (Th2) and non-Th2 molecular phenotypes of asthma using sputum transcriptomics in U-BIOPRED. Eur Respir J. 2017;49(2):1602135.

9. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8(1):118–27.

10. Kikkert M, Doolman R, Dai M, Avner R, Hassink G, van Voorden S, et al. Human HRD1 is an E3 ubiquitin ligase involved in degradation of proteins from the

endoplasmic reticulum. J Biol Chem. 2004;279(5):3525–34.

11. Sun S, Shi G, Sha H, Ji Y, Han X, Shu X, et al. IRE1α is an endogenous substrate of endoplasmic-reticulum-associated degradation. Nat Cell Biol. 2015;17(12):1546.

12. Chapple RH, Hu T, Tseng Y-J, Liu L, Kitano A, Luu V, et al. ERα promotes murine hematopoietic regeneration through the Ire1α-mediated unfolded protein response. Elife. 2018;7:e31159.

13. Dent J, El-Serag HB, Wallander M, Johansson S. Epidemiology of gastro-oesophageal reflux disease: a systematic review. Gut. 2005;54(5):710–7.

14. Bezkorovainy A. Antimicrobial properties of iron-binding proteins. In: Diet and Resistance to Disease. Springer; 1981. p. 139–54.

15. Azevedo EPC, Guimarães-Costa AB, Torezani GS, Braga CA, Palhano FL, Kelly JW, et al. Amyloid fibrils trigger the release of neutrophil extracellular traps (NETs), causing fibril fragmentation by NET-associated elastase. J Biol Chem. 2012;jbc-M112.

16. Martinho A, Gonçalves I, Costa M, Santos CR. Stress and glucocorticoids increase transthyretin expression in rat choroid plexus via mineralocorticoid and glucocorticoid receptors. J Mol Neurosci. 2012;48(1):1–13.

17. Martinho A, Goncalves I, Costa M, Santos CR. Stress and glucocorticoids increase transthyretin expression in rat choroid plexus via mineralocorticoid and glucocorticoid receptors. J Mol Neurosci. 2012/03/01. 2012;48(1):1–13.

18. Li H-N. Impact of cationic host defence peptide LL-37 on human neutrophil death and inflammatory responses. 2011;

19. Pereira HA. CAP37, a neutrophil-derived multifunctional inflammatory mediator. J Leukoc Biol. 1995;57(6):805–12.

20. Zoubeidi A, Ettinger S, Beraldi E, Hadaschik B, Zardan A, Klomp LWJ, et al. Clusterin facilitates COMMD1 and I-κB degradation to enhance NF-κB activity in

prostate cancer cells. Mol Cancer Res. 2010;1541–7786.

21. Pazdrak K, Young TW, Straub C, Stafford S, Kurosky A. Priming of eosinophils by GM-CSF is mediated by protein kinase CβII-phosphorylated L-plastin. J Immunol. 2011;1001868.

22. Castro-Giner F, Bustamante M, González JR, Kogevinas M, Jarvis D, Heinrich J, et al. A pooling-based genome-wide analysis identifies new potential candidate genes for atopy in the European Community Respiratory Health Survey (ECRHS). BMC Med Genet. 2009;10(1):128.

23. Strobeck MW, DeCristofaro MF, Banine F, Weissman BE, Sherman LS, Knudsen ES. The BRG-1 subunit of the SWI/SNF complex regulates CD44 expression. J Biol Chem. 2001;276(12):9273–8.

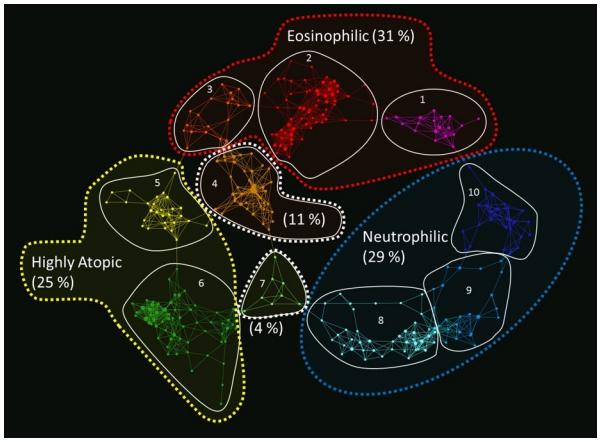
24. Wang T, Zou W, Niu C, Ding F, Wang Y, Cai S, et al. Brg1 inhibits E-cadherin expression in lung epithelial cells and disrupts epithelial integrity. J Mol Med. 2017;95(10):1117–26.

25. Guan H, Nagarkatti PS, Nagarkatti M. Role of CD44 in the differentiation of Th1 and Th2 cells: CD44-deficiency enhances the development of Th2 effectors in response to sheep RBC and chicken ovalbumin. J Immunol. 2009;183(1):172–80.

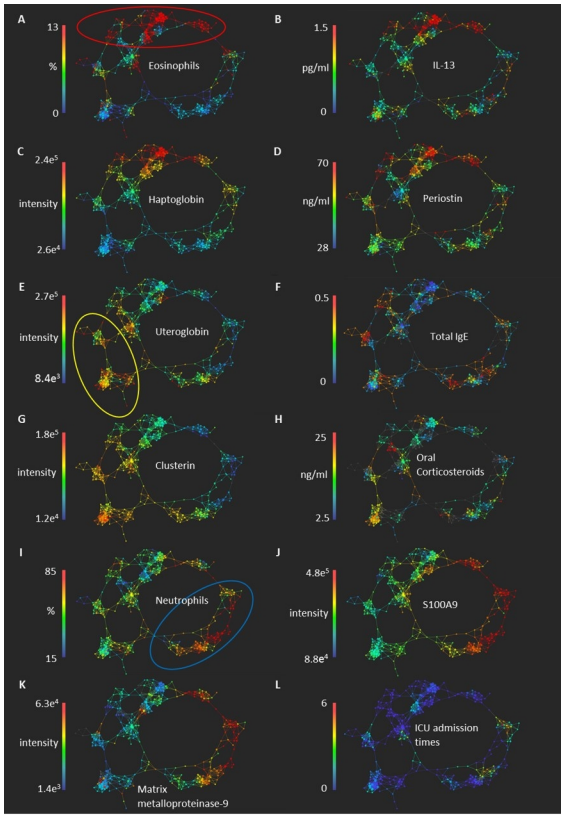
26. McNicholl DM, Stevenson M, McGarvey LP, Heaney LG. The utility of fractional exhaled nitric oxide suppression in the identification of nonadherence in difficult asthma. Am J Respir Crit Care Med. 2012;186(11):1102–8.

27. Farne HA, Wilson A, Powell C, Bax L, Milan SJ. Anti-IL5 therapies for asthma. Cochrane Libr. 2017;

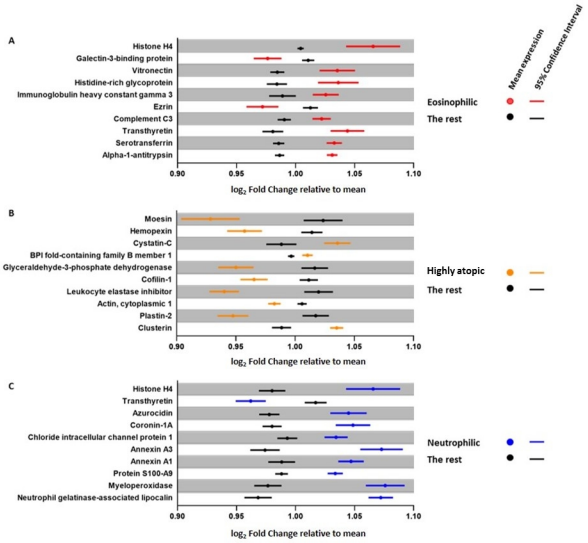
28. Walker S, Monteil M, Phelan K, Lasserson TJ, Walters EH. Anti-IgE for chronic asthma in adults and children. Cochrane Database Syst Rev. 2006;2(2).



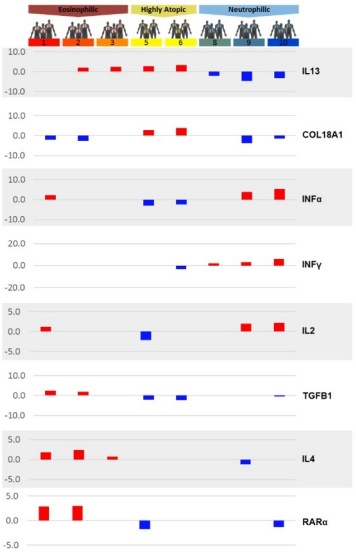
ACCEPTED MANUSCRIPT



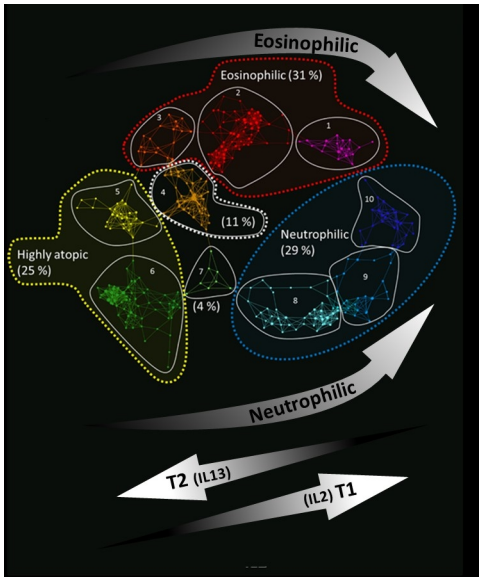
ACCEPTED MANUSCRIPT



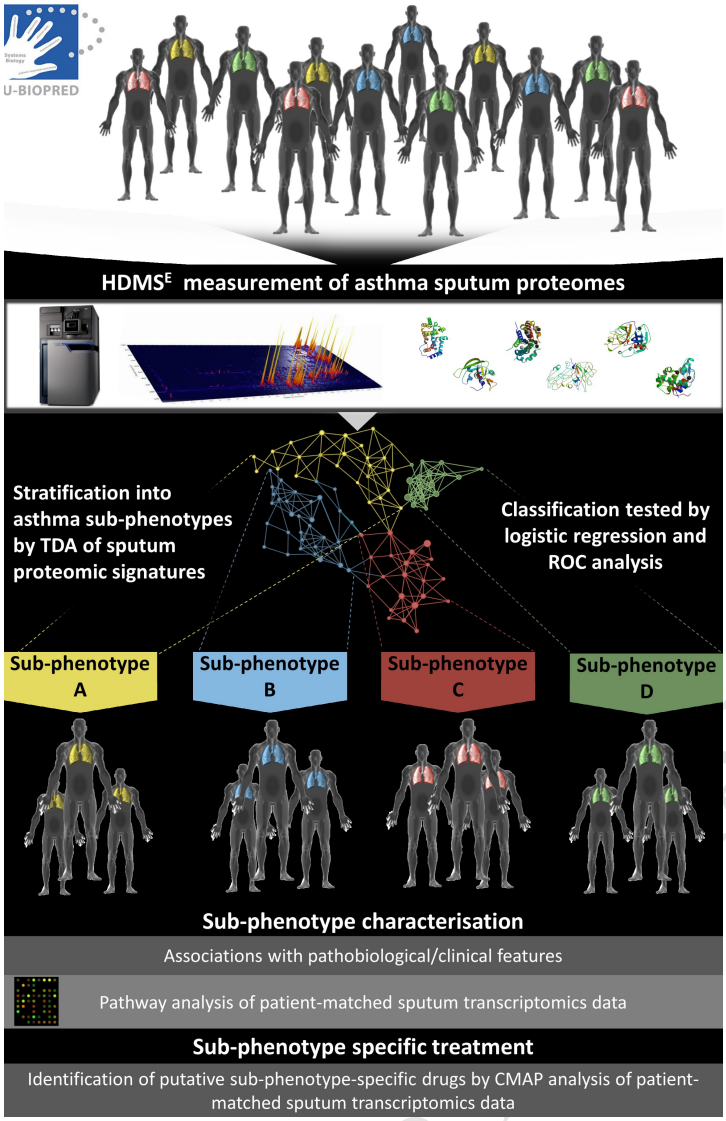
ACCEPTED MANUSCRIPT



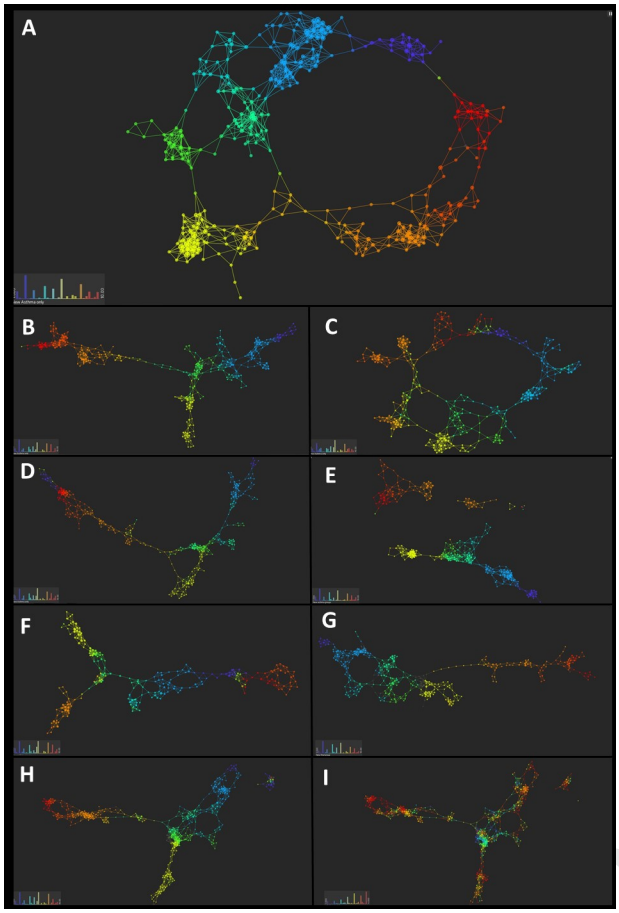
ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

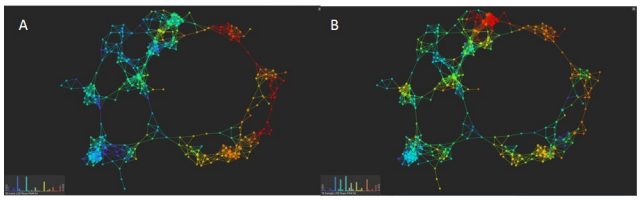


ACCEPTED MANUSCRIPT

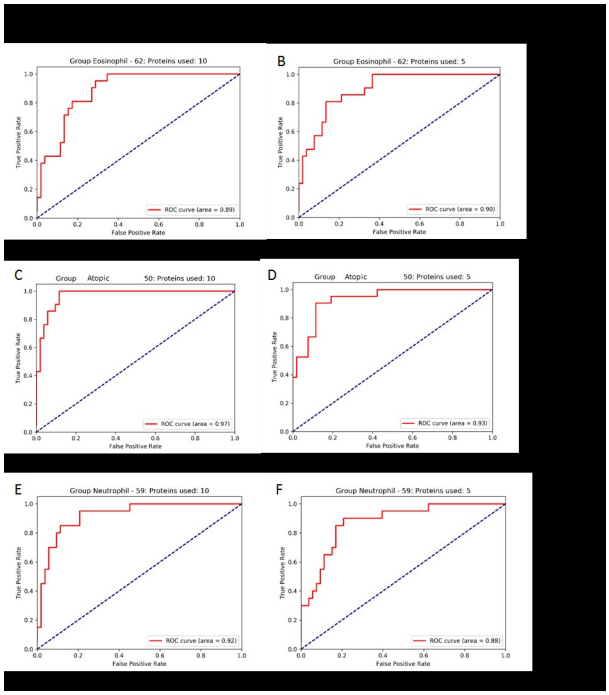


ACCEPTED MANUSCRIPT

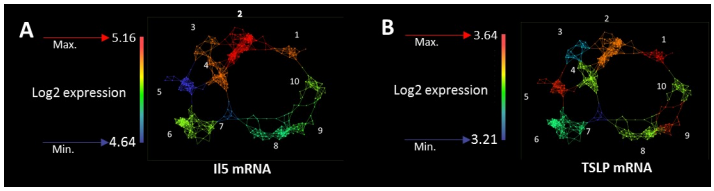
ACCEPTED MANUSCRIPT



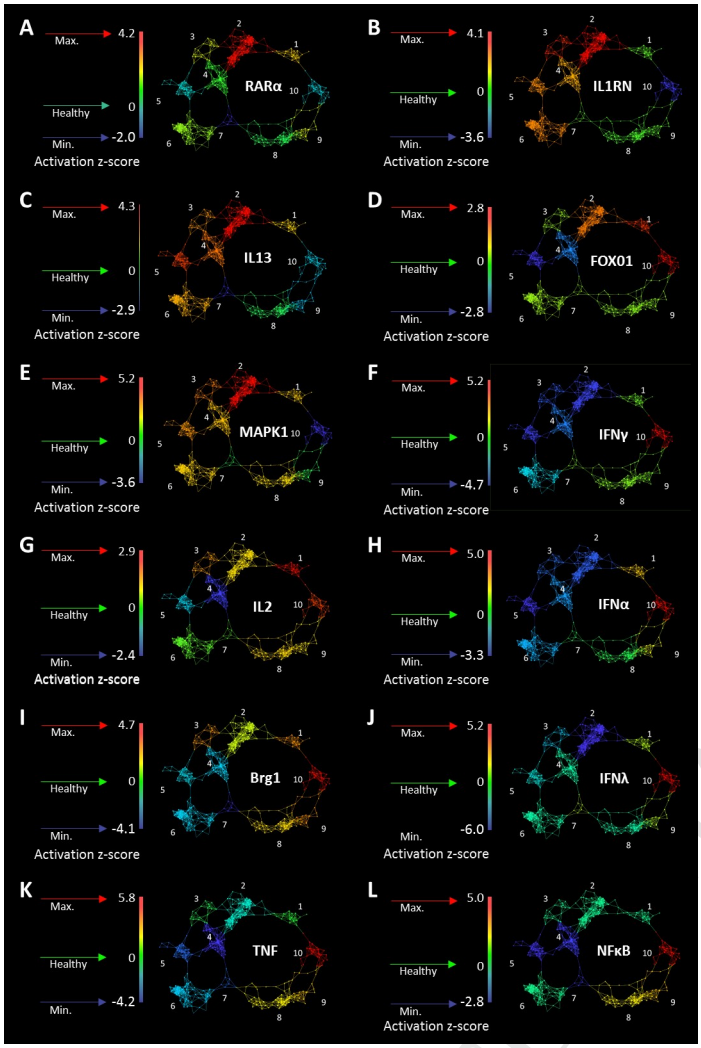
ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Stratification of asthma phenotypes by airway proteomic signatures

James P R Schofield, PhD[1, 2*], Dominic Burg, PhD[1, 2*], Ben Nicholas, PhD[2], Fabio Strazzeri, PhD[1, 3], Joost Brandsma, PhD[2], Doroteya Staykova, PhD[1], Caterina Folisi, PhD[1], Aruna T Bansal, PhD[4], Yang Xian, PhD[5], Yike Guo, PhD[5], Anthony Rowe, PhD[6], Julie Corfield, PhD[7], Susan Wilson, PhD[2], Jonathan Ward, Bsc[2], Rene Lutter, PhD[8, 9], Dominick E Shaw, MD[10], Per S Bakke, MD[11], Massimo Caruso, MD[12], Sven-Erik Dahlen, MD[13], Stephen J. Fowler, MD[14], Ildikó Horváth, MD[15], Peter Howarth, MD[2], Norbert Krug, MD[16], Paolo Montuschi, MD[17], Marek Sanak, MD[18], Thomas Sandström, MD[19], Kai Sun, PhD[5], Ioannis Pandis, PhD[5], John Riley, PhD[21], Charles Auffray, PhD[20], Bertrand De Meulder, PhD[20], Diane Lefaudeux, PhD[20], Ana R Sousa, PhD[21], Ian M Adcock, PhD[22], Kian Fan Chung, MD[22], Peter J Sterk, MD[9], Paul J Skipp, PhD[1, #], Ratko Djukanović, MD[2, #], on behalf of the U-BIOPRED Study Group

Supplementary material legends

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figure S1

Schematic of the workflow for asthma sub-phenotype stratification by TDA of sputum proteomes, through to analysis of pathobiological/clinical data and pathway analysis of sputum transcriptomics data associated with study participants of the identified sub-phenotypes.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figure S2

TDA structures of the UBIOPRED sputum proteome. **A)** Discretely coloured 10 clusters found in data of asthmatics (60% missing data, no imputation, natural distribution, normalised correlation and multidimensional scaling). The following figures **(B-H)** are examples of alternate structures, coloured by the group assignments from panel A, and highlight the local conservation of participant clustering despite treatment or dataset used. **B)** Asthma data, 60% missing data, no imputation, log2 transformation, normalised correlation and multidimensional scaling **C)** Asthma data, 10% missing data, zero imputation, log2 transformation, normalised correlation and multidimensional scaling. **D)** Asthma data, 10% missing data, zero imputation, natural data, normalised correlation and multidimensional scaling **E)** Severe Asthma data, 60% missing data, no imputation, natural data, normalised correlation and multidimensional scaling. **F)** Severe Asthma data, 10% missing data, zero imputation, natural data, normalised correlation and multidimensional scaling. **G)** All cohorts data, 60% missing data, no imputation, natural data, normalised correlation and multidimensional scaling. **H)** All cohorts data, 10% missing data, zero imputation, natural data, normalised correlation and multidimensional scaling. **I)** All cohorts data coloured by cohort with Red indicating severe asthma and blue indicating healthy. It is evident that healthy participants and mild asthmatics cluster towards the centre of the structure, with the flares being dominated by severe asthmatics.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figure S3

Consensus clustering group assignments overlaid as a colour scheme onto the TDA asthma structure. Regional conservation was observed in a number of consensus models. Shown here are **A)** 10% missing data with Pearsons correlation and PAM, K=4. **B)** 30% missing data with Pearson's correlation and PAM, k=6

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figure S4

ROC curves of logistic regression models of proteins predictive of phenotype classification. Classification success rate of the eosinophilic phenotype using A) 10 proteins and B) 5 proteins. Classification success rate of the atopic phenotype using C) 10 proteins and D) 5 proteins. Classification success rate of the neutrophilic phenotype using E) 10 proteins and F) 5 proteins.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figure S5

Varying patterns across the TDA structure for T2 cytokine mRNAs (IL5 in panel **A**, TSLP in panel **B**) in the individual sub-phenotypes: for each mRNA, the colour scale from blue to red denotes the levels expression.

ACCEPTED MANUSCRIPT

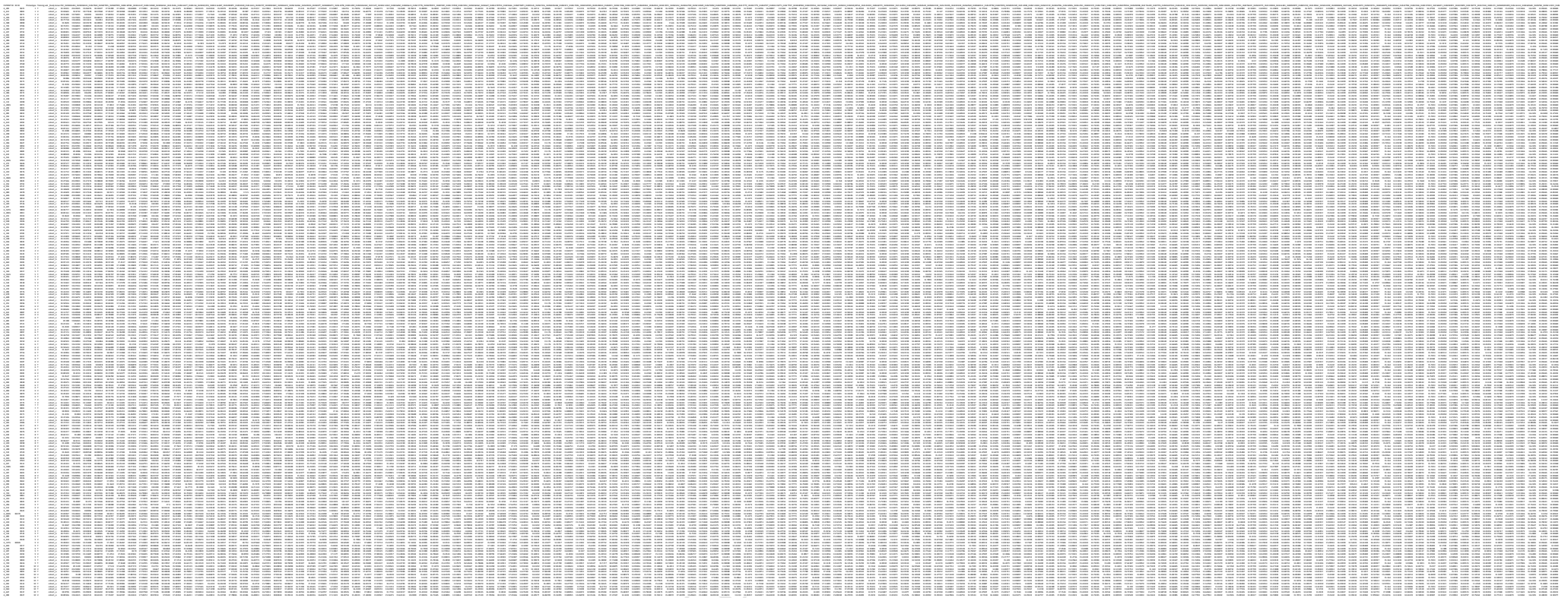
ACCEPTED MANUSCRIPT

Figure S6

Varying patterns across the TDA structure for selected upstream regulators (see white text in each panel) of gene expression signatures in the individual sub-phenotypes: for each upstream regulator, the colour scale from blue to red denotes the levels of activation (maximum shown in intense red) and inhibition (maximum shown in intense blue) relative to the healthy expression (shown in green).

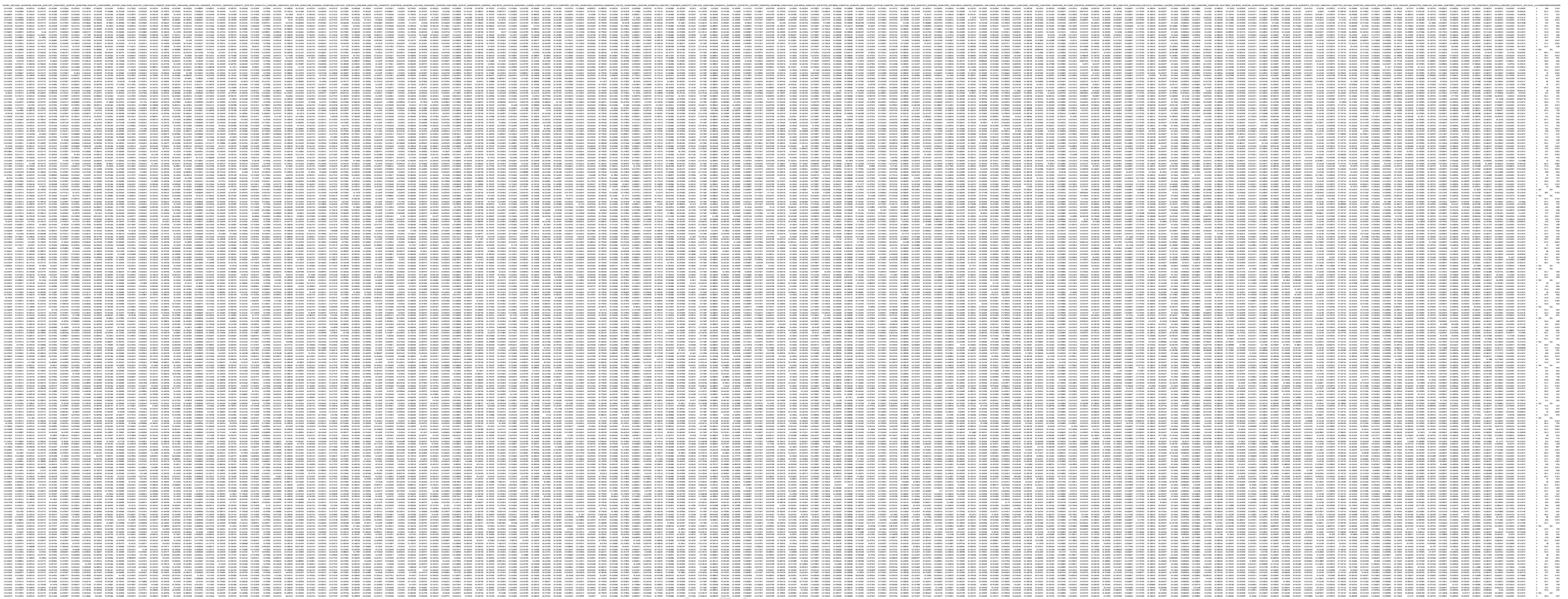
ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



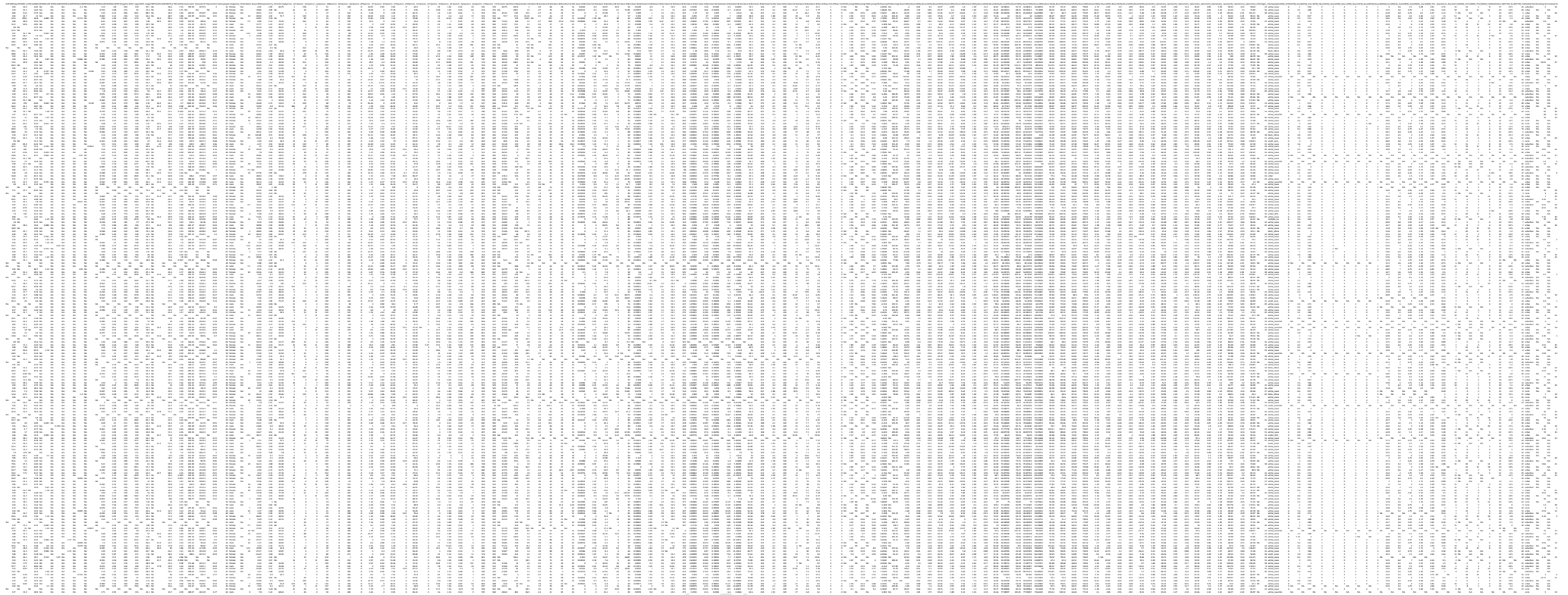
ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



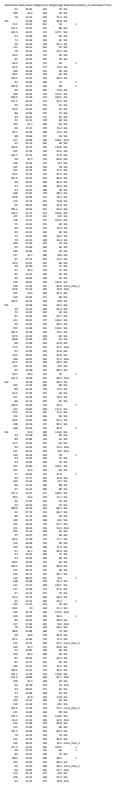
ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table S2 Eosinophilic Different Proteins and Healthy

Proteins different between Eosinophilic phenotype and healthy		
Protein	KS Sign	T-test (FDR corrected)
TCO1_HUMAN	-	0.965989
CK5P2_HUMAN	-	0.968612
MYL6_HUMAN	-	0.77806
IGHM_HUMAN	+	0.830141
TOPZ1_HUMAN	-	0.852421
MYH3_HUMAN	-	0.938371
PPIA_HUMAN	-	0.901239
CO4A_HUMAN	-	0.96441
A1ATR_HUMAN	-	0.938371
HSP76_HUMAN	-	0.852421
LRIQ1_HUMAN	-	0.766236
TBB4B_HUMAN	+	0.901239
GSTP1_HUMAN	+	0.719349
VP13C_HUMAN	+	0.727994
KV106_HUMAN	-	0.762164
MYH9_HUMAN	+	0.809308
CNOT1_HUMAN	-	0.812819
TRFL_HUMAN	-	0.924745
NFM_HUMAN	-	0.852421
H90B2_HUMAN	-	0.630423
1433Z_HUMAN	+	0.962468
PARK7_HUMAN	+	0.671219
HV305_HUMAN	-	0.964138
KV304_HUMAN	-	0.968612
SPB3_HUMAN	+	0.965989
EF1A1_HUMAN	-	0.511427
MYH7B_HUMAN	+	0.635916
HS71L_HUMAN	+	0.916521
PDIA1_HUMAN	+	0.77806
KV306_HUMAN	+	0.924745
FAM3D_HUMAN	-	0.665705
ACTB_HUMAN	-	0.796026
BD1L1_HUMAN	-	0.968612
MYH2_HUMAN	+	0.968612
CE350_HUMAN	+	0.964138
CAPG_HUMAN	+	0.968612
UTRO_HUMAN	+	0.965989
HV320_HUMAN	-	0.809308
CROCC_HUMAN	-	0.77806
HPTR_HUMAN	-	0.830141
MYH4_HUMAN	-	0.852421

ACCEPTED MANUSCRIPT

ANXA5_HUMAN	-	0.404203
HV304_HUMAN	-	0.901239
LYSC_HUMAN	+	0.882775
FIBB_HUMAN	+	0.494127
CENPE_HUMAN	-	0.948065
TTYH2_HUMAN	+	0.991877
TEN2_HUMAN	-	0.372982
KV302_HUMAN	-	0.579381
HV303_HUMAN	+	0.454794
AKAP9_HUMAN	-	0.782968
H31T_HUMAN	+	0.645393
DYH7_HUMAN	+	0.719349
ANXA2_HUMAN	-	0.330076
LV302_HUMAN	-	0.676835
HV316_HUMAN	+	0.97368
H2B1A_HUMAN	+	0.237884
K1C19_HUMAN	-	0.98165
MYH7_HUMAN	+	0.665705
LV102_HUMAN	-	0.676835
ACTN4_HUMAN	+	0.852421
DYH3_HUMAN	-	0.961071
ANR12_HUMAN	+	0.852421
CAMP_HUMAN	+	0.762164
GDIR1_HUMAN	-	0.852421
RBP2_HUMAN	+	0.645393
CO3_HUMAN	+	0.809308
VIME_HUMAN	-	0.719349
TRRAP_HUMAN	-	0.748826
1433E_HUMAN	-	0.657138
CLAP2_HUMAN	-	0.812819
CEAM5_HUMAN	+	0.727994
ENOB_HUMAN	-	0.558565
KV101_HUMAN	-	0.321976
1433F_HUMAN	-	0.968612
CE290_HUMAN	-	0.412004
ACTN2_HUMAN	+	0.310767
CENPF_HUMAN	+	0.852421
DYH1_HUMAN	-	0.412004
K2C8_HUMAN	+	0.643787
CNTRL_HUMAN	-	0.572684
MYH13_HUMAN	+	0.576585
SLPI_HUMAN	+	0.61133
HSP7C_HUMAN	-	0.314543
TIMP1_HUMAN	-	0.968612
PERM_HUMAN	+	0.34226
NHRF1_HUMAN	+	0.865576
TRY1_HUMAN	-	0.852421
GSTA1_HUMAN	+	0.372982

ACCEPTED MANUSCRIPT

MYH10_HUMAN	+	0.94886
CATG_HUMAN	+	0.582253
ACTN3_HUMAN	-	0.727994
DAPLE_HUMAN	+	0.533931
FIBG_HUMAN	+	0.61133
MINT_HUMAN	-	0.454794
ITPR1_HUMAN	+	0.250508
MSMB_HUMAN	-	0.454794
CNTLN_HUMAN	+	0.665705
DYH12_HUMAN	-	0.558565
ANK2_HUMAN	-	0.412004
SMG1_HUMAN	+	0.61133
TRNK1_HUMAN	-	0.328723
ANK3_HUMAN	-	0.741334
PRKDC_HUMAN	+	0.36871
LAC2_HUMAN	+	0.454794
POTEE_HUMAN	-	0.852421
BPIB2_HUMAN	-	0.279869
LKHA4_HUMAN	-	0.61133
IGHG3_HUMAN	+	0.091991
K1C10_HUMAN	-	0.454794
CIB1_HUMAN	-	0.479448
SKT_HUMAN	-	0.968612
IGLL5_HUMAN	+	0.237626
K22E_HUMAN	-	0.568068
CYTN_HUMAN	-	0.262487
SETX_HUMAN	-	0.568068
A2GL_HUMAN	-	0.852421
CLAP1_HUMAN	+	0.82208
1433G_HUMAN	-	0.184391
LV301_HUMAN	+	0.61133
CAP7_HUMAN	-	0.741334
EZRI_HUMAN	-	0.093592
1433B_HUMAN	-	0.398355
CAYP1_HUMAN	-	0.561801
FIBA_HUMAN	+	0.43667
K2C1B_HUMAN	-	0.214479
WDR87_HUMAN	-	0.61133
K1C18_HUMAN	-	0.561801
PERE_HUMAN	+	0.411578
MYH6_HUMAN	-	0.61864
IGHG2_HUMAN	+	0.266623
HSP71_HUMAN	+	0.865576
H2A1H_HUMAN	+	0.159355
PCNT_HUMAN	+	0.628153
BPIB1_HUMAN	-	0.425061
NGAL_HUMAN	+	0.47582
6PGD_HUMAN	+	0.481036

ACCEPTED MANUSCRIPT

NUCB2_HUMAN	+	0.313142
THIO_HUMAN	-	0.77806
LDHB_HUMAN	-	0.549378
NUMA1_HUMAN	-	0.285019
ANXA1_HUMAN	+	0.207898
RADI_HUMAN	+	0.61133
LC1L1_HUMAN	-	0.741334
GOGB1_HUMAN	+	0.313142
TMPSD_HUMAN	-	0.204526
XIRP2_HUMAN	-	0.103337
ANXA3_HUMAN	+	0.314687
H2B1B_HUMAN	+	0.06469
LDH6A_HUMAN	+	0.366476
B2MG_HUMAN	-	0.127738
GDIB_HUMAN	+	0.830764
RHG07_HUMAN	-	0.669977
K2C6B_HUMAN	+	0.063491
IGKC_HUMAN	+	0.678805
NR6A1_HUMAN	-	0.379881
MYH11_HUMAN	-	0.549378
S10A8_HUMAN	+	0.852421
CLIC1_HUMAN	-	0.121189
MYH1_HUMAN	-	0.61864
H3C_HUMAN	+	0.211902
GELS_HUMAN	+	0.120066
BASP1_HUMAN	-	0.235623
K1C13_HUMAN	-	0.353152
KV310_HUMAN	-	0.297394
VP13B_HUMAN	-	0.063491
LCN1_HUMAN	-	0.178645
VTDB_HUMAN	+	0.184391
ANGT_HUMAN	+	0.121189
SFPA2_HUMAN	-	0.285019
ACTBL_HUMAN	-	0.186045
ANXA6_HUMAN	+	0.099244
IGHA2_HUMAN	-	0.111335
DMBT1_HUMAN	-	0.029125
DEF1_HUMAN	+	0.240534
KV309_HUMAN	-	0.014511
GOGA4_HUMAN	-	0.028612
HS90A_HUMAN	-	0.100526
LPPL_HUMAN	+	0.163235
CATD_HUMAN	-	0.589108
ACTA_HUMAN	+	0.314543
PIP_HUMAN	-	0.036074
PERL_HUMAN	-	0.014205
S10A9_HUMAN	+	0.186045
ARP3_HUMAN	+	0.121189

ACCEPTED MANUSCRIPT

MMP9_HUMAN	+	0.103337
HSP72_HUMAN	+	0.110969
MLL1_HUMAN	-	0.372982
KPYM_HUMAN	+	0.178645
IGHA1_HUMAN	-	0.025235
PLSL_HUMAN	+	0.315869
K1671_HUMAN	-	0.0387
K1C9_HUMAN	-	0.314543
ENOA_HUMAN	+	0.61864
PEDF_HUMAN	-	0.052933
ILEU_HUMAN	+	0.118025
CFAH_HUMAN	+	0.024703
TTHY_HUMAN	+	0.021448
LDHA_HUMAN	+	0.035869
GOLM1_HUMAN	-	0.079197
GDIR2_HUMAN	+	0.010344
BPIA2_HUMAN	-	0.014511
MOES_HUMAN	+	0.117324
COR1A_HUMAN	+	0.005898
CAP1_HUMAN	+	0.052933
COF1_HUMAN	+	0.026309
ANT3_HUMAN	+	0.022505
LAC7_HUMAN	+	0.150332
DYHC2_HUMAN	-	0.041695
TPIS_HUMAN	+	0.141636
CAH6_HUMAN	-	0.025265
LV106_HUMAN	-	0.014511
WFDC2_HUMAN	-	0.024703
TBA1C_HUMAN	-	0.041695
KNG1_HUMAN	+	0.480352
AL1A1_HUMAN	-	0.015222
ZG16B_HUMAN	-	0.00128
KV204_HUMAN	-	0.040247
PROL4_HUMAN	-	0.310767
K1C14_HUMAN	+	0.022505
MUC1_HUMAN	-	0.041695
PGK1_HUMAN	+	0.099244
IGHG1_HUMAN	+	0.000598
K2C1_HUMAN	-	0.041695
PRDX1_HUMAN	-	0.006345
ALDOA_HUMAN	+	0.025837
ACTN1_HUMAN	+	0.01537
A1AG1_HUMAN	+	0.00205
PROF1_HUMAN	+	0.00875
CFAB_HUMAN	+	0.014281
ZA2G_HUMAN	-	0.00511
ECP_HUMAN	+	6.64E-05
CYTS_HUMAN	-	0.010526

ACCEPTED MANUSCRIPT

IGHG4_HUMAN	+	0.000563
PSPB_HUMAN	-	0.003261
CERU_HUMAN	-	0.001219
HS90B_HUMAN	-	0.142561
ALBU_HUMAN	+	0.000835
H4_HUMAN	+	0.010526
VTNC_HUMAN	+	0.002377
IGJ_HUMAN	-	0.000891
RL40_HUMAN	-	0.040212
LG3BP_HUMAN	-	0.00205
CYTC_HUMAN	-	0.014511
SBP1_HUMAN	-	0.00018
G6PI_HUMAN	+	0.003261
G3P_HUMAN	+	0.122869
IC1_HUMAN	+	0.078831
HRG_HUMAN	+	0.000373
CATA_HUMAN	+	0.00511
AL3B1_HUMAN	-	0.000127
QSOX1_HUMAN	-	0.001219
TKT_HUMAN	+	0.00875
PIGR_HUMAN	-	0.000667
CF058_HUMAN	-	0.000335
APOH_HUMAN	+	0.009595
AACT_HUMAN	+	4.33E-05
TRFE_HUMAN	+	3.88E-05
AMY1_HUMAN	-	0.000292
CYTT_HUMAN	-	0.001219
A1BG_HUMAN	+	0.000127
TALDO_HUMAN	+	0.001003
CLUS_HUMAN	-	0.002144
FETUA_HUMAN	+	1.12E-05
A2MG_HUMAN	+	6.37E-06
UTER_HUMAN	-	0.000108
APOA1_HUMAN	+	1.52E-07
HPT_HUMAN	+	9.21E-12
HEMO_HUMAN	+	1.44E-09
A1AT_HUMAN	+	7.18E-10

ACCEPTED MANUSCRIPT

Proteins different between Eosinophilic phenotype and rest

Protein	KS Sign	T-test (FDR corrected)
LG3BP_HUMAN	-	5.49E-05
CLIC1_HUMAN	-	5.49E-05
EZRI_HUMAN	-	5.49E-05
PIGR_HUMAN	-	0.000576
ZA2G_HUMAN	-	0.000576
NGAL_HUMAN	-	0.001827
QSOX1_HUMAN	-	0.003298
RL40_HUMAN	-	0.003608
TIMP1_HUMAN	-	0.009722
CLUS_HUMAN	-	0.015106
AL3B1_HUMAN	-	0.019706
PEDF_HUMAN	-	0.022746
1433Z_HUMAN	-	0.02481
DEF1_HUMAN	-	0.033544
ENOA_HUMAN	-	0.036008
IGJ_HUMAN	-	0.041692
S10A8_HUMAN	-	0.057727
DMBT1_HUMAN	-	0.057727
TRFL_HUMAN	-	0.059164
PERL_HUMAN	-	0.059308
VP13B_HUMAN	-	0.066791
B2MG_HUMAN	-	0.07112
SLPI_HUMAN	-	0.090698
PIP_HUMAN	-	0.118213
ANXA3_HUMAN	-	0.121491
WFDC2_HUMAN	-	0.137596
CF058_HUMAN	-	0.137596
GOLM1_HUMAN	-	0.160341
KV302_HUMAN	-	0.16902
ACTBL_HUMAN	-	0.172141
ACTN4_HUMAN	-	0.188519
MSMB_HUMAN	-	0.188519
ANXA2_HUMAN	-	0.188519
LV106_HUMAN	-	0.188519
UTER_HUMAN	-	0.198287
HSP7C_HUMAN	-	0.198287
H31T_HUMAN	-	0.221122
PERM_HUMAN	-	0.221122
CAH6_HUMAN	-	0.225868
BPIB2_HUMAN	-	0.225868
HS90A_HUMAN	-	0.225868
TBA1C_HUMAN	-	0.229527
MUC1_HUMAN	-	0.229527
PROL4_HUMAN	-	0.229527

ACCEPTED MANUSCRIPT

LKHA4_HUMAN	-	0.229527
ILEU_HUMAN	-	0.24244
CAP7_HUMAN	-	0.253485
MYH11_HUMAN	-	0.255611
PSPB_HUMAN	-	0.285529
KV309_HUMAN	-	0.285529
CYTC_HUMAN	-	0.316915
PLSL_HUMAN	-	0.316915
GOGA4_HUMAN	-	0.325975
CE290_HUMAN	-	0.325975
ANXA1_HUMAN	-	0.37058
CAMP_HUMAN	-	0.416949
ACTN3_HUMAN	-	0.416949
LDHB_HUMAN	-	0.4381
CATG_HUMAN	-	0.492262
CYTT_HUMAN	-	0.492262
LV301_HUMAN	-	0.492262
HS90B_HUMAN	-	0.513158
CERU_HUMAN	-	0.513158
CYTS_HUMAN	-	0.540874
SFPA2_HUMAN	-	0.540874
GSTP1_HUMAN	-	0.540874
FAM3D_HUMAN	-	0.541458
ACTN1_HUMAN	-	0.553888
NR6A1_HUMAN	-	0.578456
CEAM5_HUMAN	-	0.578456
DYH1_HUMAN	-	0.580271
CIB1_HUMAN	-	0.68453
SETX_HUMAN	-	0.688757
K2C1B_HUMAN	-	0.697349
H90B2_HUMAN	-	0.702879
MINT_HUMAN	-	0.748989
K1671_HUMAN	-	0.770711
PRDX1_HUMAN	-	0.770711
G3P_HUMAN	-	0.810103
NHRF1_HUMAN	-	0.81572
K2C1_HUMAN	-	0.81572
LRIQ1_HUMAN	-	0.81572
H4_HUMAN	-	0.830802
MYL6_HUMAN	-	0.847645
DYHC2_HUMAN	-	0.898769
S10A9_HUMAN	-	0.907456
H2A1H_HUMAN	-	0.907456
GDIB_HUMAN	-	0.907456
G6PI_HUMAN	-	0.917917
K22E_HUMAN	-	0.917917
ENOB_HUMAN	-	0.945979
LYSC_HUMAN	-	0.959478

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

BPIB1_HUMAN	-	0.959478
BASP1_HUMAN	-	1
SBP1_HUMAN	-	1
BPIA2_HUMAN	-	1
ANR12_HUMAN	-	1
1433B_HUMAN	-	1
CROCC_HUMAN	-	1
RADI_HUMAN	-	1
POTEE_HUMAN	-	1
1433E_HUMAN	-	1
LV102_HUMAN	-	1
CNTRL_HUMAN	-	1
IGHA2_HUMAN	-	1
KV101_HUMAN	-	1
PPIA_HUMAN	-	1
ANXA5_HUMAN	-	1
LV302_HUMAN	-	1
MYH6_HUMAN	-	1
MYH4_HUMAN	-	1
KV310_HUMAN	-	1
IGHA1_HUMAN	-	1
LDHA_HUMAN	-	1
LC1L1_HUMAN	-	1
CLAP2_HUMAN	-	1
HV304_HUMAN	-	1
CAYP1_HUMAN	-	1
ITPR1_HUMAN	-	1
1433G_HUMAN	-	1
PCNT_HUMAN	-	1
DYH12_HUMAN	-	1
NUMA1_HUMAN	-	1
TPIS_HUMAN	-	1
IGLL5_HUMAN	-	1
MOES_HUMAN	-	1
K1C10_HUMAN	-	1
LAC2_HUMAN	-	1
ZG16B_HUMAN	-	1
PROF1_HUMAN	-	1
AL1A1_HUMAN	-	1
MYH9_HUMAN	-	1
IGHM_HUMAN	-	1
H2B1A_HUMAN	-	1
SMG1_HUMAN	-	1
MMP9_HUMAN	-	1
HSP76_HUMAN	-	1
CAPG_HUMAN	-	1
CENPF_HUMAN	-	1
K1C18_HUMAN	-	1

ACCEPTED MANUSCRIPT

KV204_HUMAN	-	1
LCN1_HUMAN	-	1
UTRO_HUMAN	-	1
AMY1_HUMAN	-	1
ARP3_HUMAN	-	1
HPT_HUMAN	+	3.94E-27
A1AT_HUMAN	+	5.32E-24
ALBU_HUMAN	+	7.77E-18
TRFE_HUMAN	+	7.36E-17
HEMO_HUMAN	+	1.63E-16
FETUA_HUMAN	+	2.80E-14
APOA1_HUMAN	+	6.48E-13
A2MG_HUMAN	+	7.66E-13
A1BG_HUMAN	+	1.13E-09
TTHY_HUMAN	+	1.24E-08
CO3_HUMAN	+	7.98E-08
CFAH_HUMAN	+	1.98E-07
VTNC_HUMAN	+	5.23E-07
ANT3_HUMAN	+	9.99E-07
HRG_HUMAN	+	1.65E-06
VTDB_HUMAN	+	1.65E-06
IGHG1_HUMAN	+	8.83E-06
APOH_HUMAN	+	1.78E-05
LPPL_HUMAN	+	5.40E-05
IGHG2_HUMAN	+	8.23E-05
KNG1_HUMAN	+	8.73E-05
PERE_HUMAN	+	0.000103
IGHG3_HUMAN	+	0.000208
IGHG4_HUMAN	+	0.0005
DYH3_HUMAN	+	0.00297
HSP72_HUMAN	+	0.006776
KV304_HUMAN	+	0.00828
ECP_HUMAN	+	0.020559
IC1_HUMAN	+	0.026128
FIBA_HUMAN	+	0.036302
COF1_HUMAN	+	0.039514
ANGT_HUMAN	+	0.040121
RBP2_HUMAN	+	0.040121
ACTA_HUMAN	+	0.081288
K1C14_HUMAN	+	0.110949
IGKC_HUMAN	+	0.158514
HV303_HUMAN	+	0.164959
TRRAP_HUMAN	+	0.164959
CFAB_HUMAN	+	0.172764
CO4A_HUMAN	+	0.19035
HV305_HUMAN	+	0.19035
MYH7_HUMAN	+	0.199138
ALDOA_HUMAN	+	0.201781

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

PRKDC_HUMAN	+	0.247663
A1AG1_HUMAN	+	0.247663
DAPLE_HUMAN	+	0.248441
CNTLN_HUMAN	+	0.27595
CE350_HUMAN	+	0.316237
MYH7B_HUMAN	+	0.357964
AKAP9_HUMAN	+	0.36889
K2C8_HUMAN	+	0.390923
CYTN_HUMAN	+	0.390923
H2B1B_HUMAN	+	0.420778
FIBG_HUMAN	+	0.420778
HV316_HUMAN	+	0.461943
LAC7_HUMAN	+	0.511405
TMPSD_HUMAN	+	0.513316
FIBB_HUMAN	+	0.732082
PGK1_HUMAN	+	0.745351
GSTA1_HUMAN	+	0.787395
HPTR_HUMAN	+	0.862118
KPYM_HUMAN	+	0.989854
EF1A1_HUMAN	+	1
GDIR2_HUMAN	+	1
CNOT1_HUMAN	+	1
MYH2_HUMAN	+	1
K2C6B_HUMAN	+	1
CATA_HUMAN	+	1
BD1L1_HUMAN	+	1
6PGD_HUMAN	+	1
XIRP2_HUMAN	+	1
K1C19_HUMAN	+	1
K1C13_HUMAN	+	1
CK5P2_HUMAN	+	1
TTYH2_HUMAN	+	1
A2GL_HUMAN	+	1
WDR87_HUMAN	+	1
MLL1_HUMAN	+	1
THIO_HUMAN	+	1
HV320_HUMAN	+	1
TCO1_HUMAN	+	1
ANK2_HUMAN	+	1
ANXA6_HUMAN	+	1
COR1A_HUMAN	+	1
LDH6A_HUMAN	+	1
TRNK1_HUMAN	+	1
PDIA1_HUMAN	+	1
NUCB2_HUMAN	+	1
CAP1_HUMAN	+	1
TKT_HUMAN	+	1
VIME_HUMAN	+	1

ACCEPTED MANUSCRIPT

SPB3_HUMAN	+	1
A1ATR_HUMAN	+	1
TRY1_HUMAN	+	1
ANK3_HUMAN	+	1
GOGB1_HUMAN	+	1
AACT_HUMAN	+	1
MYH13_HUMAN	+	1
CLAP1_HUMAN	+	1
SKT_HUMAN	+	1
PARK7_HUMAN	+	1
H3C_HUMAN	+	1
NFM_HUMAN	+	1
VP13C_HUMAN	+	1
K1C9_HUMAN	+	1
GDIR1_HUMAN	+	1
TALDO_HUMAN	+	1
TBB4B_HUMAN	+	1
HSP71_HUMAN	+	1
ACTN2_HUMAN	+	1
MYH3_HUMAN	+	1
GELS_HUMAN	+	1
DYH7_HUMAN	+	1
RHG07_HUMAN	+	1
1433F_HUMAN	+	1
ACTB_HUMAN	+	1
CATD_HUMAN	+	1
KV106_HUMAN	+	1
TOPZ1_HUMAN	+	1
KV306_HUMAN	+	1
TEN2_HUMAN	+	1
MYH1_HUMAN	+	1
HS71L_HUMAN	+	1
CENPE_HUMAN	+	1
MYH10_HUMAN	+	1

ACCEPTED MANUSCRIPT

Table S4 Highly atopic Different Proteins

Proteins different between Highly atopic phenotype and rest		
Protein	KS Sign	T-test (FDR corrected)
A2GL_HUMAN	-	0.628721
MYH1_HUMAN	-	0.709151
MYH4_HUMAN	+	0.735717
MYH3_HUMAN	+	0.742472
KV204_HUMAN	+	0.837383
TBB4B_HUMAN	+	0.402362
MYH6_HUMAN	+	0.548802
1433G_HUMAN	+	0.297414
CIB1_HUMAN	+	0.671933
EF1A1_HUMAN	+	0.92005
TRNK1_HUMAN	+	0.402362
GOGA4_HUMAN	-	0.735717
VIME_HUMAN	-	0.402362
MYH11_HUMAN	+	0.663162
A1ATR_HUMAN	-	0.942596
UTRO_HUMAN	+	0.402957
IGLL5_HUMAN	-	0.809171
PCNT_HUMAN	+	0.809171
CENPF_HUMAN	+	0.564668
WDR87_HUMAN	+	0.797281
VP13C_HUMAN	+	0.297173
NUMA1_HUMAN	+	0.655223
KV106_HUMAN	+	0.657836
BASP1_HUMAN	+	0.850379
GOGB1_HUMAN	-	0.490601
HSP76_HUMAN	+	0.493706
CNTLN_HUMAN	-	0.735717
HV320_HUMAN	-	0.383465
THIO_HUMAN	-	0.416095
LV102_HUMAN	+	0.596354
PSPB_HUMAN	+	0.709151
MYH7B_HUMAN	-	0.557198
LC1L1_HUMAN	+	0.555679
RL40_HUMAN	-	0.941228
VP13B_HUMAN	+	0.95981
CLAP2_HUMAN	+	0.663162
CEAM5_HUMAN	+	0.892909
HV316_HUMAN	-	0.450535
TBA1C_HUMAN	+	0.530125
TEN2_HUMAN	+	0.493706
DYH3_HUMAN	+	0.472791
ANR12_HUMAN	+	0.201277

ACCEPTED MANUSCRIPT

RADI_HUMAN	-	0.614374
CROCC_HUMAN	+	0.891081
IGHM_HUMAN	+	0.709151
TRRAP_HUMAN	+	0.55414
HV304_HUMAN	+	0.691227
ANK3_HUMAN	+	0.290268
K1C18_HUMAN	+	0.971976
NHRF1_HUMAN	+	0.461487
SETX_HUMAN	+	0.638484
PARK7_HUMAN	-	0.450697
SFPA2_HUMAN	+	0.209092
CK5P2_HUMAN	+	0.190184
GSTA1_HUMAN	-	0.358028
K1C13_HUMAN	+	0.942596
KV306_HUMAN	+	0.5404
MYH13_HUMAN	+	0.286856
CNTRL_HUMAN	+	0.809171
MUC1_HUMAN	+	0.647043
1433F_HUMAN	+	0.184559
HS90B_HUMAN	+	0.565037
KV310_HUMAN	+	0.366343
LV302_HUMAN	+	0.276461
SKT_HUMAN	+	0.892909
HPTR_HUMAN	-	0.176738
ITPR1_HUMAN	+	0.096938
CAYP1_HUMAN	+	0.628721
IGHG3_HUMAN	-	0.358028
BD1L1_HUMAN	+	0.636951
TMPSD_HUMAN	+	0.140408
RBP2_HUMAN	-	0.687665
ANGT_HUMAN	-	0.757445
K1671_HUMAN	+	0.19821
HS71L_HUMAN	+	0.676856
LAC7_HUMAN	-	0.450535
LCN1_HUMAN	+	0.16872
CFAH_HUMAN	-	0.053017
CYTN_HUMAN	+	0.709151
NFM_HUMAN	+	0.265941
1433E_HUMAN	-	0.955617
IGJ_HUMAN	+	0.160103
KV101_HUMAN	-	0.269806
K2C8_HUMAN	-	0.850379
TOPZ1_HUMAN	+	0.905751
CATD_HUMAN	+	0.748116
TCO1_HUMAN	+	0.007394
CENPE_HUMAN	-	0.456537
MINT_HUMAN	+	0.104752
MYH2_HUMAN	+	0.333821

ACCEPTED MANUSCRIPT

HV305_HUMAN	-	0.36098
GDIR1_HUMAN	-	0.00983
CE350_HUMAN	+	0.171333
ANK2_HUMAN	+	0.317227
SBP1_HUMAN	+	0.098399
H90B2_HUMAN	+	0.166677
PRKDC_HUMAN	-	0.209092
NUCB2_HUMAN	+	0.973245
XIRP2_HUMAN	-	0.965192
HS90A_HUMAN	-	0.882673
AKAP9_HUMAN	-	0.709151
SMG1_HUMAN	+	0.0784
CNOT1_HUMAN	-	0.636951
ANXA5_HUMAN	-	0.663162
HV303_HUMAN	-	0.358028
LV106_HUMAN	+	0.046349
ACTA_HUMAN	-	0.596354
MYL6_HUMAN	-	0.385759
ENOB_HUMAN	+	0.899729
MYH10_HUMAN	+	0.018043
DAPLE_HUMAN	+	0.809171
DYH12_HUMAN	+	0.018043
NR6A1_HUMAN	+	0.339853
DYH1_HUMAN	+	0.02753
LRIQ1_HUMAN	+	0.159222
KNG1_HUMAN	-	0.647043
AL1A1_HUMAN	+	0.286856
HSP72_HUMAN	-	0.029794
PDIA1_HUMAN	-	0.648965
POTEE_HUMAN	-	0.447612
IGKC_HUMAN	-	0.204302
CO3_HUMAN	-	0.074297
MLL1_HUMAN	+	0.053017
K2C1B_HUMAN	+	0.049631
MYH7_HUMAN	-	0.040594
TTYH2_HUMAN	-	0.166677
ANXA2_HUMAN	+	0.236513
LDH6A_HUMAN	+	0.18331
K1C19_HUMAN	-	0.171333
VTDB_HUMAN	-	0.007394
RHG07_HUMAN	+	0.00222
TTHY_HUMAN	-	0.171333
EZRI_HUMAN	+	0.159222
H3C_HUMAN	-	0.005279
DYH7_HUMAN	-	0.566864
ACTN2_HUMAN	+	0.891081
LAC2_HUMAN	+	0.277703
KV302_HUMAN	+	0.022235

ACCEPTED MANUSCRIPT

IGHG2_HUMAN	-	0.106789
CLAP1_HUMAN	+	0.019876
K2C6B_HUMAN	-	0.791968
KV304_HUMAN	-	0.045859
SPB3_HUMAN	+	0.345407
FETUA_HUMAN	-	0.036314
GSTP1_HUMAN	-	0.0061
CERU_HUMAN	+	0.038793
ACTN3_HUMAN	-	0.809171
KV309_HUMAN	+	0.014498
6PGD_HUMAN	-	0.002908
FAM3D_HUMAN	+	0.020509
CO4A_HUMAN	+	0.65086
CYTT_HUMAN	+	0.004904
MYH9_HUMAN	-	0.061176
CE290_HUMAN	+	0.166677
LV301_HUMAN	+	0.035745
H2B1A_HUMAN	-	0.005316
AL3B1_HUMAN	+	0.058607
TRY1_HUMAN	+	0.493706
APOH_HUMAN	-	0.032468
1433B_HUMAN	-	0.215828
LPPL_HUMAN	-	0.035006
ANXA6_HUMAN	+	0.614374
DYHC2_HUMAN	+	0.026887
ZG16B_HUMAN	+	0.00244
ACTN4_HUMAN	-	0.730612
IGHA1_HUMAN	+	0.00945
CAH6_HUMAN	+	0.004014
LDHB_HUMAN	-	0.033936
CLIC1_HUMAN	-	0.056231
ECP_HUMAN	-	0.104712
IGHA2_HUMAN	+	0.040594
ALBU_HUMAN	-	0.001663
ARP3_HUMAN	-	0.004704
IGHG4_HUMAN	-	0.035745
VTNC_HUMAN	-	0.008941
HRG_HUMAN	-	0.039972
ANT3_HUMAN	-	0.106789
PPIA_HUMAN	-	0.149746
AACT_HUMAN	+	0.000791
K1C14_HUMAN	-	0.000369
FIBG_HUMAN	-	0.018043
K2C1_HUMAN	+	0.003573
PERE_HUMAN	-	0.548802
A2MG_HUMAN	-	0.053963
LKHA4_HUMAN	-	0.007394
ACTBL_HUMAN	-	0.233196

ACCEPTED MANUSCRIPT

GDIB_HUMAN	-	0.020872
CAMP_HUMAN	-	0.003442
UTER_HUMAN	+	1.06E-05
MSMB_HUMAN	+	0.053757
FIBB_HUMAN	-	0.001631
QSOX1_HUMAN	+	4.61E-05
G6PI_HUMAN	-	0.461487
IC1_HUMAN	-	0.026598
CAPG_HUMAN	-	0.001281
GELS_HUMAN	-	0.013033
APOA1_HUMAN	-	0.001281
KPYM_HUMAN	-	1.61E-05
K22E_HUMAN	+	0.0483
AMY1_HUMAN	+	4.95E-05
HSP7C_HUMAN	-	0.01012
CATA_HUMAN	-	0.090958
H2B1B_HUMAN	-	0.317227
H31T_HUMAN	-	0.018316
PIGR_HUMAN	+	1.82E-05
ENOA_HUMAN	-	0.000813
CAP7_HUMAN	-	4.95E-05
FIBA_HUMAN	-	0.007265
CYTS_HUMAN	+	0.176893
PEDF_HUMAN	+	2.98E-05
A1BG_HUMAN	-	0.00168
PRDX1_HUMAN	+	0.000985
LG3BP_HUMAN	+	4.29E-05
CFAB_HUMAN	-	6.14E-05
HSP71_HUMAN	-	1.03E-05
H2A1H_HUMAN	-	0.000791
TRFE_HUMAN	-	6.83E-05
MMP9_HUMAN	-	4.67E-05
B2MG_HUMAN	+	0.006301
LDHA_HUMAN	-	7.97E-06
BPIA2_HUMAN	+	0.000833
A1AG1_HUMAN	-	0.01327
WFDC2_HUMAN	+	0.000562
CAP1_HUMAN	-	0.000479
NGAL_HUMAN	-	0.00019
ALDOA_HUMAN	-	4.61E-05
1433Z_HUMAN	-	0.000535
TPIS_HUMAN	-	1.64E-05
K1C9_HUMAN	+	0.000813
ANXA3_HUMAN	-	6.28E-07
PROL4_HUMAN	+	8.18E-09
TKT_HUMAN	-	6.94E-07
BPIB1_HUMAN	+	4.44E-06
ANXA1_HUMAN	-	1.65E-05

ACCEPTED MANUSCRIPT

SLPI_HUMAN	+	3.61E-06
CATG_HUMAN	-	0.005316
K1C10_HUMAN	+	0.005069
BPIB2_HUMAN	+	0.000101
ACTN1_HUMAN	-	7.45E-05
COF1_HUMAN	-	1.90E-08
IGHG1_HUMAN	-	2.29E-05
TALDO_HUMAN	-	0.000391
PERM_HUMAN	-	4.67E-08
LYSC_HUMAN	+	9.16E-08
H4_HUMAN	-	1.55E-07
A1AT_HUMAN	-	9.63E-08
DEF1_HUMAN	-	3.03E-08
HEMO_HUMAN	-	3.03E-08
HEMO_HUMAN	-	3.03E-08
HPT_HUMAN	-	1.56E-08
MOES_HUMAN	-	6.19E-08
PGK1_HUMAN	-	3.69E-07
G3P_HUMAN	-	1.84E-09
S10A8_HUMAN	-	9.13E-08
TIMP1_HUMAN	+	3.65E-10
PLSL_HUMAN	-	9.08E-12
TRFL_HUMAN	+	7.96E-08
CYTC_HUMAN	+	8.05E-07
CF058_HUMAN	+	1.58E-14
ACTB_HUMAN	-	1.44E-09
GDIR2_HUMAN	-	1.52E-11
GOLM1_HUMAN	+	8.98E-05
PERL_HUMAN	+	4.64E-12
COR1A_HUMAN	-	4.61E-12
ILEU_HUMAN	-	1.94E-14
S10A9_HUMAN	-	1.05E-11
DMBT1_HUMAN	+	2.52E-18
PIP_HUMAN	+	2.77E-13
ZA2G_HUMAN	+	4.84E-15
CLUS_HUMAN	+	7.26E-16
PROF1_HUMAN	-	2.52E-18

ACCEPTED MANUSCRIPT

Proteins different between Highly atopic phenotype and healthy

Protein	KS Sign	T-test (FDR corrected)
VTNC_HUMAN	+	0.877963
VP13C_HUMAN	+	0.877963
NUMA1_HUMAN	+	0.944076
LC1L1_HUMAN	+	0.877963
HV304_HUMAN	-	0.944076
NFM_HUMAN	-	0.941532
CO4A_HUMAN	-	0.947423
AL1A1_HUMAN	-	0.877963
A1ATR_HUMAN	-	0.877963
PERE_HUMAN	-	0.877963
HRG_HUMAN	+	0.877963
K1C14_HUMAN	+	0.999825
SKT_HUMAN	+	0.877963
UTRO_HUMAN	+	0.877963
CNTLN_HUMAN	+	0.877963
K1C10_HUMAN	-	0.941532
H90B2_HUMAN	-	0.877963
ACTN4_HUMAN	+	0.877963
GDIB_HUMAN	+	0.956711
A1AG1_HUMAN	-	0.797322
IGHM_HUMAN	+	0.890663
LRIQ1_HUMAN	-	0.890663
ALBU_HUMAN	+	0.907291
HPT_HUMAN	-	0.877963
NR6A1_HUMAN	+	0.828437
HSP76_HUMAN	-	0.944076
TRRAP_HUMAN	+	0.947423
TBA1C_HUMAN	-	0.947423
LV302_HUMAN	-	0.877963
MYH13_HUMAN	-	0.877984
ANXA2_HUMAN	+	0.94721
HV303_HUMAN	+	0.877963
CENPF_HUMAN	-	0.713752
SMG1_HUMAN	+	0.907291
CNTRL_HUMAN	-	0.809214
SETX_HUMAN	-	0.947423
ACTA_HUMAN	+	0.877963
EF1A1_HUMAN	-	0.818106
TOPZ1_HUMAN	-	0.877963
THIO_HUMAN	-	0.877963
DYH3_HUMAN	+	0.877984
PARK7_HUMAN	-	0.714407
TRNK1_HUMAN	-	0.963653
MYH7_HUMAN	-	0.713752

ACCEPTED MANUSCRIPT

K2C8_HUMAN	+	0.982226
IGHA2_HUMAN	+	0.877963
POTEE_HUMAN	-	0.807446
HEMO_HUMAN	-	0.877963
CFAH_HUMAN	-	0.755695
HV320_HUMAN	-	0.472557
K1C19_HUMAN	-	0.513579
SPB3_HUMAN	+	0.817004
FAM3D_HUMAN	+	0.745442
CE290_HUMAN	-	0.713752
DYH12_HUMAN	+	0.877963
PDIA1_HUMAN	+	0.890663
ENOB_HUMAN	+	0.877963
DEF1_HUMAN	-	0.877963
IGHG4_HUMAN	-	0.92628
KNG1_HUMAN	-	0.877963
CE350_HUMAN	-	0.877963
CROCC_HUMAN	+	0.982226
A1AT_HUMAN	-	0.581158
QSOX1_HUMAN	+	0.877963
TMPSD_HUMAN	+	0.967924
MLL1_HUMAN	-	0.944076
K1671_HUMAN	-	0.748772
A2MG_HUMAN	+	0.939574
MYH11_HUMAN	-	0.714407
ENOA_HUMAN	-	0.821931
ANXA5_HUMAN	-	0.832028
H2B1A_HUMAN	+	0.999825
ARP3_HUMAN	-	0.877963
ANXA6_HUMAN	+	0.562726
PEDF_HUMAN	+	0.499639
TTHY_HUMAN	+	0.877963
FIBA_HUMAN	-	0.809214
CF058_HUMAN	+	0.676976
EZRI_HUMAN	+	0.947423
G3P_HUMAN	-	0.536116
LV102_HUMAN	-	0.877963
1433G_HUMAN	-	0.877963
KV310_HUMAN	-	0.828437
FIBG_HUMAN	-	0.832028
MYH3_HUMAN	-	0.877963
MYH1_HUMAN	-	0.877963
K22E_HUMAN	-	0.484104
GSTA1_HUMAN	-	0.872255
ANK2_HUMAN	-	0.807446
MYH9_HUMAN	-	0.877963
CIB1_HUMAN	-	0.595174
FETUA_HUMAN	+	0.491926

ACCEPTED MANUSCRIPT

LCN1_HUMAN	-	0.797322
WFDC2_HUMAN	+	0.877963
DYHC2_HUMAN	-	0.885285
APOH_HUMAN	+	0.952132
ZG16B_HUMAN	-	0.525187
GOLM1_HUMAN	+	0.890663
DYH7_HUMAN	+	0.832028
TBB4B_HUMAN	+	0.982226
DAPLE_HUMAN	+	0.76282
NHRF1_HUMAN	+	0.748772
LAC7_HUMAN	-	0.877963
ACTN2_HUMAN	+	0.525187
ACTBL_HUMAN	-	0.810987
RADI_HUMAN	+	0.822776
VP13B_HUMAN	-	0.907291
PIGR_HUMAN	-	0.92628
A1BG_HUMAN	-	0.822776
LG3BP_HUMAN	+	0.877963
HSP72_HUMAN	-	0.54245
H2B1B_HUMAN	-	0.877963
ANT3_HUMAN	-	0.809214
MYH6_HUMAN	+	0.821931
RBP2_HUMAN	-	0.821931
CATG_HUMAN	-	0.517866
IGLL5_HUMAN	-	0.944076
ANR12_HUMAN	+	0.877963
PPIA_HUMAN	-	0.499639
MYH4_HUMAN	+	0.821931
HS90B_HUMAN	-	0.877963
H2A1H_HUMAN	-	0.877963
IGHG3_HUMAN	-	0.54245
BD1L1_HUMAN	-	0.877963
GOGA4_HUMAN	-	0.363437
G6PI_HUMAN	+	0
ANK3_HUMAN	-	0.552822
H31T_HUMAN	-	0.676976
KV106_HUMAN	+	0.714407
CK5P2_HUMAN	+	0.877963
KV309_HUMAN	-	0.360952
DYH1_HUMAN	+	0.809214
K2C1B_HUMAN	+	0.75621
WDR87_HUMAN	-	0.809214
1433F_HUMAN	+	0.755695
MYH2_HUMAN	+	0.676976
HPTR_HUMAN	-	0.54245
CATA_HUMAN	+	0
GELS_HUMAN	+	0.340181
MINT_HUMAN	-	0.525187

ACCEPTED MANUSCRIPT

TTYH2_HUMAN	-	0.525187
CLIC1_HUMAN	-	0.809214
HV316_HUMAN	-	0.267867
CENPE_HUMAN	-	0.807446
MUC1_HUMAN	-	0.877242
CYTC_HUMAN	+	0.54245
HS71L_HUMAN	+	0.676976
A2GL_HUMAN	-	0.499639
TEN2_HUMAN	-	0.267867
CERU_HUMAN	-	0.342062
COF1_HUMAN	-	0.484104
H4_HUMAN	-	0.622887
UTER_HUMAN	-	0.713752
1433B_HUMAN	-	0.652608
FIBB_HUMAN	-	0.484104
IGHA1_HUMAN	-	0.877963
S10A8_HUMAN	-	0.428895
CATD_HUMAN	-	0.525187
BASP1_HUMAN	-	0.788346
MYL6_HUMAN	-	0.755695
K1C9_HUMAN	-	0.472557
K2C6B_HUMAN	-	0.872803
TALDO_HUMAN	+	0.890663
GOGB1_HUMAN	+	0.396181
K1C18_HUMAN	-	0.88914
SFPA2_HUMAN	+	0.605575
APOA1_HUMAN	+	0.341415
PRKDC_HUMAN	-	0.213526
CYTN_HUMAN	-	0.472557
TKT_HUMAN	-	0.420798
1433E_HUMAN	-	0.499639
IC1_HUMAN	-	0.497413
1433Z_HUMAN	-	0.365792
PCNT_HUMAN	+	0.491926
CO3_HUMAN	-	0.134083
CYTS_HUMAN	-	0.525187
ANXA1_HUMAN	-	0.524942
MYH7B_HUMAN	-	0.54245
CLUS_HUMAN	+	0.939574
RHG07_HUMAN	+	0.103239
ITPR1_HUMAN	+	0.622887
GDIR1_HUMAN	-	0.321129
LDH6A_HUMAN	+	0.473419
NGAL_HUMAN	+	0.877963
KV304_HUMAN	-	0.420961
CLAP2_HUMAN	-	0.5516
LDHB_HUMAN	-	0.420798
CFAB_HUMAN	+	0.877963

ACCEPTED MANUSCRIPT

TCO1_HUMAN	+	0.472557
ACTN3_HUMAN	-	0.0339
PRDX1_HUMAN	-	0.491926
MSMB_HUMAN	+	0.513579
GSTP1_HUMAN	-	0.605575
KV302_HUMAN	+	0.342062
CLAP1_HUMAN	+	0.363437
AKAP9_HUMAN	+	0.54245
XIRP2_HUMAN	-	0.484104
IGKC_HUMAN	-	0.472557
HS90A_HUMAN	-	0.877963
VTDB_HUMAN	-	0.363437
NUCB2_HUMAN	+	0.525187
ANGT_HUMAN	+	0.75621
TRFE_HUMAN	-	0.321129
VIME_HUMAN	-	0.581158
K2C1_HUMAN	-	0.213526
TRY1_HUMAN	-	0.464338
BPIA2_HUMAN	-	0.472557
LPPL_HUMAN	+	0.525187
IGJ_HUMAN	-	0.303185
IGHG1_HUMAN	-	0.304008
GDIR2_HUMAN	-	0.128097
DMBT1_HUMAN	+	0.048513
LDHA_HUMAN	-	0.590458
CEAM5_HUMAN	+	0.231012
PERL_HUMAN	+	0.289447
H3C_HUMAN	-	0.209448
BPIB2_HUMAN	+	0.125822
LV301_HUMAN	+	0.133079
B2MG_HUMAN	+	0.54245
KV204_HUMAN	-	0.138014
6PGD_HUMAN	-	0.134083
HSP7C_HUMAN	-	0.159038
MMP9_HUMAN	-	0.127415
CAMP_HUMAN	-	0.316236
CYTT_HUMAN	-	0.133079
KV101_HUMAN	-	0.128097
RL40_HUMAN	-	0.10518
CAYP1_HUMAN	-	0.472557
CAP1_HUMAN	-	0.019584
S10A9_HUMAN	-	0.321129
ANXA3_HUMAN	-	0.025102
ECP_HUMAN	+	0.267867
PERM_HUMAN	-	0.133079
SBP1_HUMAN	-	0.240075
CNOT1_HUMAN	-	0.267867
TPIS_HUMAN	-	0.267867

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

AL3B1_HUMAN	-	0.605575
CAP7_HUMAN	-	0.128097
MYH10_HUMAN	+	0.058159
MOES_HUMAN	-	0.524942
ALDOA_HUMAN	-	0.125822
LKHA4_HUMAN	-	0.016334
CAH6_HUMAN	-	0.525187
ZA2G_HUMAN	+	0.121595
COR1A_HUMAN	-	0.135053
HV305_HUMAN	-	0.277165
KPYM_HUMAN	-	0.080617
PSPB_HUMAN	-	0.277165
ILEU_HUMAN	-	0.133079
CAPG_HUMAN	-	0.000414
LAC2_HUMAN	+	0.069133
PGK1_HUMAN	-	0.134083
KV306_HUMAN	-	0.877963
ACTN1_HUMAN	-	0.213526
BPIB1_HUMAN	+	0.009178
LV106_HUMAN	-	0.029733
TRFL_HUMAN	+	0.019584
AMY1_HUMAN	-	0.134083
PROF1_HUMAN	-	0.133079
HSP71_HUMAN	-	0.038424
PIP_HUMAN	+	0.036162
IGHG2_HUMAN	-	0.303185
K1C13_HUMAN	-	0.360952
LYSC_HUMAN	+	0.002416
PLSL_HUMAN	-	0.004207
ACTB_HUMAN	-	0.004207
AACT_HUMAN	+	0.000168
SLPI_HUMAN	+	9.83E-05
PROL4_HUMAN	+	0.000736
TIMP1_HUMAN	+	6.11E-06
PROF1_HUMAN	-	2.52E-18

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Proteins different between Neutrophilic phenotype and rest

Protein	KS Sign	T-test (FDR corrected)
G3P_HUMAN	+	7.14E-08
PROF1_HUMAN	+	7.14E-08
ILEU_HUMAN	+	7.14E-08
ENOA_HUMAN	+	1.14E-06
DEF1_HUMAN	+	1.14E-06
PLSL_HUMAN	+	1.54E-06
ANXA3_HUMAN	+	6.06E-06
COR1A_HUMAN	+	8.31E-06
K1C9_HUMAN	-	8.31E-06
AMY1_HUMAN	-	8.71E-06
TALDO_HUMAN	+	8.71E-06
NGAL_HUMAN	+	8.71E-06
LDHA_HUMAN	+	8.99E-06
ANXA1_HUMAN	+	9.33E-06
S10A9_HUMAN	+	9.72E-06
S10A8_HUMAN	+	1.13E-05
ACTN1_HUMAN	+	1.73E-05
MMP9_HUMAN	+	1.73E-05
K1C10_HUMAN	-	6.39E-05
GDIR2_HUMAN	+	6.72E-05
K2C1_HUMAN	-	9.58E-05
MOES_HUMAN	+	0.00015
TPIS_HUMAN	+	0.000174
H4_HUMAN	+	0.00027
CF058_HUMAN	-	0.000275
TKT_HUMAN	+	0.000275
IGHG2_HUMAN	-	0.000319
UTER_HUMAN	-	0.000338
CFAB_HUMAN	+	0.000623
DYHC2_HUMAN	-	0.001229
PGK1_HUMAN	+	0.001683
MINT_HUMAN	-	0.001722
CYTT_HUMAN	-	0.001722
CATA_HUMAN	+	0.001982
CYTC_HUMAN	-	0.002095
PRDX1_HUMAN	-	0.002846
ANK3_HUMAN	-	0.003008
SBP1_HUMAN	-	0.00318
H2A1H_HUMAN	+	0.004287
CO3_HUMAN	-	0.004529
KPYM_HUMAN	+	0.005057
GOLM1_HUMAN	-	0.005057
CLUS_HUMAN	-	0.005224
PERM_HUMAN	+	0.005224

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

G6PI_HUMAN	+	0.006459
HPT_HUMAN	+	0.007217
CLIC1_HUMAN	+	0.007217
ZG16B_HUMAN	-	0.008065
ACTN4_HUMAN	+	0.008065
GDIB_HUMAN	+	0.010677
CYTS_HUMAN	-	0.011057
H3C_HUMAN	+	0.011057
ZA2G_HUMAN	-	0.012336
BPIA2_HUMAN	-	0.012336
DMBT1_HUMAN	-	0.01376
CERU_HUMAN	-	0.01376
H2B1A_HUMAN	+	0.01428
CAMP_HUMAN	+	0.01428
K22E_HUMAN	-	0.015083
PERL_HUMAN	-	0.01593
AL1A1_HUMAN	-	0.01655
TRY1_HUMAN	-	0.01655
A1AG1_HUMAN	+	0.017479
PSPB_HUMAN	-	0.018457
CAH6_HUMAN	-	0.019487
HEMO_HUMAN	+	0.020264
GSTP1_HUMAN	+	0.020264
HSP7C_HUMAN	+	0.021392
MLL1_HUMAN	-	0.022579
KV309_HUMAN	-	0.023827
PIP_HUMAN	-	0.028756
1433Z_HUMAN	+	0.02949
COF1_HUMAN	+	0.02949
H2B1B_HUMAN	+	0.02949
CATD_HUMAN	-	0.032766
CAP7_HUMAN	+	0.032766
ECP_HUMAN	+	0.036845
H31T_HUMAN	+	0.0388
CATG_HUMAN	+	0.04354
WFDC2_HUMAN	-	0.045244
K1671_HUMAN	-	0.045244
KV306_HUMAN	-	0.050674
KV304_HUMAN	-	0.063289
CE290_HUMAN	-	0.063289
LAC7_HUMAN	+	0.066451
IGHA1_HUMAN	-	0.069754
CAP1_HUMAN	+	0.073201
XIRP2_HUMAN	-	0.076799
ANK2_HUMAN	-	0.080551
GELS_HUMAN	+	0.088541
GOGA4_HUMAN	-	0.088541
CNOT1_HUMAN	-	0.098271

IGHA2_HUMAN	-	0.10075
VTDB_HUMAN	-	0.10075
MYH9_HUMAN	+	0.10075
BPIB1_HUMAN	-	0.104418
ACTBL_HUMAN	+	0.104418
ACTB_HUMAN	+	0.113289
AACT_HUMAN	+	0.113289
LV106_HUMAN	-	0.113289
CLAP1_HUMAN	-	0.132289
SFPA2_HUMAN	-	0.138287
ARP3_HUMAN	+	0.15244
IGLL5_HUMAN	+	0.159191
CAPG_HUMAN	+	0.16619
ANXA5_HUMAN	-	0.17928
QSOX1_HUMAN	-	0.17928
NUCB2_HUMAN	+	0.17928
HSP76_HUMAN	-	0.187003
PEDF_HUMAN	-	0.19324
TTHY_HUMAN	-	0.19324
TMPSD_HUMAN	-	0.197918
MYH6_HUMAN	-	0.197918
EF1A1_HUMAN	-	0.197918
KNG1_HUMAN	-	0.204517
MYH7B_HUMAN	-	0.204517
ALBU_HUMAN	-	0.213119
AL3B1_HUMAN	-	0.222011
LCN1_HUMAN	-	0.231197
CE350_HUMAN	-	0.238694
LPPL_HUMAN	+	0.238694
KV204_HUMAN	-	0.248421
TRFE_HUMAN	-	0.258459
IGHG3_HUMAN	-	0.26666
HSP71_HUMAN	+	0.26666
1433G_HUMAN	-	0.275081
K1C13_HUMAN	-	0.275081
CROCC_HUMAN	-	0.279392
VP13B_HUMAN	-	0.279392
HV316_HUMAN	-	0.279392
CEAM5_HUMAN	+	0.279392
HS71L_HUMAN	-	0.288189
CAYP1_HUMAN	-	0.288189
6PGD_HUMAN	+	0.292871
VP13C_HUMAN	-	0.292871
K2C6B_HUMAN	-	0.292871
K2C1B_HUMAN	-	0.292871
A1AT_HUMAN	-	0.302046
SLPI_HUMAN	-	0.302046
EZRI_HUMAN	+	0.321071

ACCEPTED MANUSCRIPT

CYTN_HUMAN	-	0.321071
CO4A_HUMAN	-	0.321071
WDR87_HUMAN	-	0.321071
DYH1_HUMAN	-	0.333222
ANXA6_HUMAN	+	0.361
IGHG4_HUMAN	+	0.371687
BD1L1_HUMAN	-	0.371687
CK5P2_HUMAN	-	0.382603
ANR12_HUMAN	-	0.382603
PIGR_HUMAN	-	0.39637
RADI_HUMAN	+	0.402482
PARK7_HUMAN	+	0.402482
DYH3_HUMAN	-	0.402482
CLAP2_HUMAN	-	0.402482
DYH12_HUMAN	-	0.414034
MYH10_HUMAN	-	0.414034
ENOB_HUMAN	-	0.428518
MUC1_HUMAN	-	0.45849
K1C14_HUMAN	+	0.45849
PERE_HUMAN	-	0.473975
TIMP1_HUMAN	-	0.486764
THIO_HUMAN	-	0.486764
K2C8_HUMAN	-	0.499755
HRG_HUMAN	+	0.499755
HS90B_HUMAN	-	0.50987
KV310_HUMAN	-	0.50987
ANGT_HUMAN	-	0.50987
PROL4_HUMAN	-	0.526316
CFAH_HUMAN	-	0.563426
BPIB2_HUMAN	-	0.580842
LG3BP_HUMAN	-	0.588216
POTEE_HUMAN	+	0.588216
APOA1_HUMAN	+	0.588216
PRKDC_HUMAN	+	0.588216
LV102_HUMAN	-	0.602485
MYH11_HUMAN	-	0.602485
1433B_HUMAN	+	0.610028
GOGB1_HUMAN	+	0.610028
TRRAP_HUMAN	-	0.610028
MYH4_HUMAN	-	0.610028
ANXA2_HUMAN	-	0.624455
1433E_HUMAN	-	0.624455
IC1_HUMAN	-	0.635529
AKAP9_HUMAN	+	0.635529
RHG07_HUMAN	-	0.635529
K1C19_HUMAN	-	0.671851
NFM_HUMAN	-	0.671851
LYSC_HUMAN	-	0.679432

ACCEPTED MANUSCRIPT

ALDOA_HUMAN	+	0.679432
MYH13_HUMAN	+	0.679432
CIB1_HUMAN	-	0.679432
LAC2_HUMAN	+	0.694153
TOPZ1_HUMAN	+	0.694153
CENPE_HUMAN	-	0.712538
IGKC_HUMAN	-	0.727294
GDIR1_HUMAN	+	0.727294
CNTLN_HUMAN	-	0.745779
HS90A_HUMAN	-	0.782862
HPTR_HUMAN	-	0.782862
RL40_HUMAN	+	0.789555
LC1L1_HUMAN	-	0.789555
NUMA1_HUMAN	-	0.789555
A1BG_HUMAN	-	0.789555
HV304_HUMAN	-	0.822015
TBB4B_HUMAN	-	0.822015
FIBB_HUMAN	+	0.822015
PCNT_HUMAN	-	0.828025
FAM3D_HUMAN	-	0.828025
MYH2_HUMAN	-	0.828025
TEN2_HUMAN	-	0.828025
VIME_HUMAN	-	0.837743
SPB3_HUMAN	-	0.837743
K1C18_HUMAN	-	0.837743
IGHM_HUMAN	-	0.85627
LKHA4_HUMAN	+	0.85627
FETUA_HUMAN	-	0.85627
ACTN2_HUMAN	+	0.85627
DAPLE_HUMAN	-	0.85627
MYH1_HUMAN	-	0.85627
TRFL_HUMAN	-	0.86893
HV320_HUMAN	-	0.86893
IGHG1_HUMAN	+	0.881207
PPIA_HUMAN	+	0.881207
PDIA1_HUMAN	+	0.893064
MYH3_HUMAN	-	0.893064
HV305_HUMAN	-	0.908464
KV106_HUMAN	-	0.919378
FIBG_HUMAN	+	0.919378
NHRF1_HUMAN	-	0.921739
TTYH2_HUMAN	-	0.921739
SKT_HUMAN	-	0.921739
A1ATR_HUMAN	-	0.921739
MSMB_HUMAN	-	0.931526
1433F_HUMAN	-	0.931526
CENPF_HUMAN	-	0.940708
VTNC_HUMAN	-	0.940708

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

LDH6A_HUMAN	-	0.945277
NR6A1_HUMAN	-	0.945277
SMG1_HUMAN	-	0.945277
KV101_HUMAN	-	0.9828
TRNK1_HUMAN	-	0.9828
B2MG_HUMAN	-	0.9844
GSTA1_HUMAN	+	0.9844
RBP2_HUMAN	-	0.9844
BASP1_HUMAN	+	0.985274
APOH_HUMAN	-	0.985274
DYH7_HUMAN	+	0.985274
TBA1C_HUMAN	-	0.987356
LV302_HUMAN	-	0.987356
ACTN3_HUMAN	+	0.987356
LDHB_HUMAN	+	0.987356
H90B2_HUMAN	-	0.987356
ITPR1_HUMAN	+	0.987356
MYL6_HUMAN	-	0.987356
LRIQ1_HUMAN	-	0.987356
IGJ_HUMAN	-	0.989088
A2MG_HUMAN	-	0.989088
KV302_HUMAN	-	0.990029
SETX_HUMAN	+	0.990029
LV301_HUMAN	-	0.990193
ANT3_HUMAN	+	0.990193
TCO1_HUMAN	-	0.997192
A2GL_HUMAN	+	0.997192
HSP72_HUMAN	-	0.997192
CNTRL_HUMAN	+	0.997192
UTRO_HUMAN	-	0.997192
MYH7_HUMAN	-	0.997192
FIBA_HUMAN	+	0.997192
ACTA_HUMAN	-	0.997192
HV303_HUMAN	-	0.997192

ACCEPTED MANUSCRIPT

Table S7 Neutrophilic Different Proteins and Healthy

Proteins different between Neutrophilic phenotype and healthy		
Protein	KS Sign	T-test (FDR corrected)
TCO1_HUMAN	-	0.997192
A2GL_HUMAN	+	0.997192
HSP72_HUMAN	-	0.997192
CNTRL_HUMAN	+	0.997192
UTRO_HUMAN	-	0.997192
MYH7_HUMAN	-	0.997192
FIBA_HUMAN	+	0.997192
ACTA_HUMAN	-	0.997192
HV303_HUMAN	-	0.997192
LV301_HUMAN	-	0.990193
ANT3_HUMAN	+	0.990193
KV302_HUMAN	-	0.990029
SETX_HUMAN	+	0.990029
IGJ_HUMAN	-	0.989088
A2MG_HUMAN	-	0.989088
TBA1C_HUMAN	-	0.987356
LV302_HUMAN	-	0.987356
ACTN3_HUMAN	+	0.987356
LDHB_HUMAN	+	0.987356
H90B2_HUMAN	-	0.987356
ITPR1_HUMAN	+	0.987356
MYL6_HUMAN	-	0.987356
LRIQ1_HUMAN	-	0.987356
BASP1_HUMAN	+	0.985274
APOH_HUMAN	-	0.985274
DYH7_HUMAN	+	0.985274
B2MG_HUMAN	-	0.9844
GSTA1_HUMAN	+	0.9844
RBP2_HUMAN	-	0.9844
KV101_HUMAN	-	0.9828
TRNK1_HUMAN	-	0.9828
LDH6A_HUMAN	-	0.945277
NR6A1_HUMAN	-	0.945277
SMG1_HUMAN	-	0.945277
CENPF_HUMAN	-	0.940708
VTNC_HUMAN	-	0.940708
MSMB_HUMAN	-	0.931526
1433F_HUMAN	-	0.931526
NHRF1_HUMAN	-	0.921739
TTYH2_HUMAN	-	0.921739
SKT_HUMAN	-	0.921739

ACCEPTED MANUSCRIPT

A1ATR_HUMAN	-	0.921739
KV106_HUMAN	-	0.919378
FIBG_HUMAN	+	0.919378
HV305_HUMAN	-	0.908464
PDIA1_HUMAN	+	0.893064
MYH3_HUMAN	-	0.893064
IGHG1_HUMAN	+	0.881207
PPIA_HUMAN	+	0.881207
TRFL_HUMAN	-	0.86893
HV320_HUMAN	-	0.86893
IGHM_HUMAN	-	0.85627
LKHA4_HUMAN	+	0.85627
FETUA_HUMAN	-	0.85627
ACTN2_HUMAN	+	0.85627
DAPLE_HUMAN	-	0.85627
MYH1_HUMAN	-	0.85627
VIME_HUMAN	-	0.837743
SPB3_HUMAN	-	0.837743
K1C18_HUMAN	-	0.837743
PCNT_HUMAN	-	0.828025
FAM3D_HUMAN	-	0.828025
MYH2_HUMAN	-	0.828025
TEN2_HUMAN	-	0.828025
HV304_HUMAN	-	0.822015
TBB4B_HUMAN	-	0.822015
FIBB_HUMAN	+	0.822015
RL40_HUMAN	+	0.789555
LC1L1_HUMAN	-	0.789555
NUMA1_HUMAN	-	0.789555
A1BG_HUMAN	-	0.789555
HS90A_HUMAN	-	0.782862
HPTR_HUMAN	-	0.782862
CNTLN_HUMAN	-	0.745779
IGKC_HUMAN	-	0.727294
GDIR1_HUMAN	+	0.727294
CENPE_HUMAN	-	0.712538
LAC2_HUMAN	+	0.694153
TOPZ1_HUMAN	+	0.694153
LYSC_HUMAN	-	0.679432
ALDOA_HUMAN	+	0.679432
MYH13_HUMAN	+	0.679432
CIB1_HUMAN	-	0.679432
K1C19_HUMAN	-	0.671851
NFM_HUMAN	-	0.671851
IC1_HUMAN	-	0.635529
AKAP9_HUMAN	+	0.635529
RHG07_HUMAN	-	0.635529
ANXA2_HUMAN	-	0.624455

ACCEPTED MANUSCRIPT

1433E_HUMAN	-	0.624455
1433B_HUMAN	+	0.610028
GOGB1_HUMAN	+	0.610028
TRRAP_HUMAN	-	0.610028
MYH4_HUMAN	-	0.610028
LV102_HUMAN	-	0.602485
MYH11_HUMAN	-	0.602485
LG3BP_HUMAN	-	0.588216
POTEE_HUMAN	+	0.588216
APOA1_HUMAN	+	0.588216
PRKDC_HUMAN	+	0.588216
BPIB2_HUMAN	-	0.580842
CFAH_HUMAN	-	0.563426
PROL4_HUMAN	-	0.526316
HS90B_HUMAN	-	0.50987
KV310_HUMAN	-	0.50987
ANGT_HUMAN	-	0.50987
K2C8_HUMAN	-	0.499755
HRG_HUMAN	+	0.499755
TIMP1_HUMAN	-	0.486764
THIO_HUMAN	-	0.486764
PERE_HUMAN	-	0.473975
MUC1_HUMAN	-	0.45849
K1C14_HUMAN	+	0.45849
ENOB_HUMAN	-	0.428518
DYH12_HUMAN	-	0.414034
MYH10_HUMAN	-	0.414034
RADI_HUMAN	+	0.402482
PARK7_HUMAN	+	0.402482
DYH3_HUMAN	-	0.402482
CLAP2_HUMAN	-	0.402482
PIGR_HUMAN	-	0.39637
CK5P2_HUMAN	-	0.382603
ANR12_HUMAN	-	0.382603
IGHG4_HUMAN	+	0.371687
BD1L1_HUMAN	-	0.371687
ANXA6_HUMAN	+	0.361
DYH1_HUMAN	-	0.333222
EZRI_HUMAN	+	0.321071
CYTN_HUMAN	-	0.321071
CO4A_HUMAN	-	0.321071
WDR87_HUMAN	-	0.321071
A1AT_HUMAN	-	0.302046
SLPI_HUMAN	-	0.302046
6PGD_HUMAN	+	0.292871
VP13C_HUMAN	-	0.292871
K2C6B_HUMAN	-	0.292871
K2C1B_HUMAN	-	0.292871

ACCEPTED MANUSCRIPT

HS71L_HUMAN	-	0.288189
CAYP1_HUMAN	-	0.288189
CROCC_HUMAN	-	0.279392
VP13B_HUMAN	-	0.279392
HV316_HUMAN	-	0.279392
CEAM5_HUMAN	+	0.279392
1433G_HUMAN	-	0.275081
K1C13_HUMAN	-	0.275081
IGHG3_HUMAN	-	0.26666
HSP71_HUMAN	+	0.26666
TRFE_HUMAN	-	0.258459
KV204_HUMAN	-	0.248421
CE350_HUMAN	-	0.238694
LPPL_HUMAN	+	0.238694
LCN1_HUMAN	-	0.231197
AL3B1_HUMAN	-	0.222011
ALBU_HUMAN	-	0.213119
KNG1_HUMAN	-	0.204517
MYH7B_HUMAN	-	0.204517
TMPSD_HUMAN	-	0.197918
MYH6_HUMAN	-	0.197918
EF1A1_HUMAN	-	0.197918
PEDF_HUMAN	-	0.19324
TTHY_HUMAN	-	0.19324
HSP76_HUMAN	-	0.187003
ANXA5_HUMAN	-	0.17928
QSOX1_HUMAN	-	0.17928
NUCB2_HUMAN	+	0.17928
CAPG_HUMAN	+	0.16619
IGLL5_HUMAN	+	0.159191
ARP3_HUMAN	+	0.15244
SFPA2_HUMAN	-	0.138287
CLAP1_HUMAN	-	0.132289
ACTB_HUMAN	+	0.113289
AACT_HUMAN	+	0.113289
LV106_HUMAN	-	0.113289
BPIB1_HUMAN	-	0.104418
ACTBL_HUMAN	+	0.104418
IGHA2_HUMAN	-	0.10075
VTDB_HUMAN	-	0.10075
MYH9_HUMAN	+	0.10075
CNOT1_HUMAN	-	0.098271
GELS_HUMAN	+	0.088541
GOGA4_HUMAN	-	0.088541
ANK2_HUMAN	-	0.080551
XIRP2_HUMAN	-	0.076799
CAP1_HUMAN	+	0.073201
IGHA1_HUMAN	-	0.069754

ACCEPTED MANUSCRIPT

LAC7_HUMAN	+	0.066451
KV304_HUMAN	-	0.063289
CE290_HUMAN	-	0.063289
KV306_HUMAN	-	0.050674
WFDC2_HUMAN	-	0.045244
K1671_HUMAN	-	0.045244
CATG_HUMAN	+	0.04354
H31T_HUMAN	+	0.0388
ECP_HUMAN	+	0.036845
CATD_HUMAN	-	0.032766
CAP7_HUMAN	+	0.032766
1433Z_HUMAN	+	0.02949
COF1_HUMAN	+	0.02949
H2B1B_HUMAN	+	0.02949
PIP_HUMAN	-	0.028756
KV309_HUMAN	-	0.023827
MLL1_HUMAN	-	0.022579
HSP7C_HUMAN	+	0.021392
HEMO_HUMAN	+	0.020264
GSTP1_HUMAN	+	0.020264
CAH6_HUMAN	-	0.019487
PSPB_HUMAN	-	0.018457
A1AG1_HUMAN	+	0.017479
AL1A1_HUMAN	-	0.01655
TRY1_HUMAN	-	0.01655
PERL_HUMAN	-	0.01593
K22E_HUMAN	-	0.015083
H2B1A_HUMAN	+	0.01428
CAMP_HUMAN	+	0.01428
DMBT1_HUMAN	-	0.01376
CERU_HUMAN	-	0.01376
ZA2G_HUMAN	-	0.012336
BPIA2_HUMAN	-	0.012336
CYTS_HUMAN	-	0.011057
H3C_HUMAN	+	0.011057
GDIB_HUMAN	+	0.010677
ZG16B_HUMAN	-	0.008065
ACTN4_HUMAN	+	0.008065
HPT_HUMAN	+	0.007217
CLIC1_HUMAN	+	0.007217
G6PI_HUMAN	+	0.006459
CLUS_HUMAN	-	0.005224
PERM_HUMAN	+	0.005224
KPYM_HUMAN	+	0.005057
GOLM1_HUMAN	-	0.005057
CO3_HUMAN	-	0.004529
H2A1H_HUMAN	+	0.004287
SBP1_HUMAN	-	0.00318

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ANK3_HUMAN	-	0.003008
PRDX1_HUMAN	-	0.002846
CYTC_HUMAN	-	0.002095
CATA_HUMAN	+	0.001982
MINT_HUMAN	-	0.001722
CYTT_HUMAN	-	0.001722
PGK1_HUMAN	+	0.001683
DYHC2_HUMAN	-	0.001229
CFAB_HUMAN	+	0.000623
UTER_HUMAN	-	0.000338
IGHG2_HUMAN	-	0.000319
CF058_HUMAN	-	0.000275
TKT_HUMAN	+	0.000275
H4_HUMAN	+	0.00027
TPIS_HUMAN	+	0.000174
MOES_HUMAN	+	0.00015
K2C1_HUMAN	-	9.58E-05
GDIR2_HUMAN	+	6.72E-05
K1C10_HUMAN	-	6.39E-05
ACTN1_HUMAN	+	1.73E-05
MMP9_HUMAN	+	1.73E-05
S10A8_HUMAN	+	1.13E-05
S10A9_HUMAN	+	9.72E-06
ANXA1_HUMAN	+	9.33E-06
LDHA_HUMAN	+	8.99E-06
AMY1_HUMAN	-	8.71E-06
TALDO_HUMAN	+	8.71E-06
NGAL_HUMAN	+	8.71E-06
COR1A_HUMAN	+	8.31E-06
K1C9_HUMAN	-	8.31E-06
ANXA3_HUMAN	+	6.06E-06
PLSL_HUMAN	+	1.54E-06
ENOA_HUMAN	+	1.14E-06
DEF1_HUMAN	+	1.14E-06
G3P_HUMAN	+	7.14E-08
PROF1_HUMAN	+	7.14E-08
ILEU_HUMAN	+	7.14E-08
PROF1_HUMAN	-	2.52E-18

ACCEPTED MANUSCRIPT

Table S8

Phenotype	Sub-phenotype	Upstream Regulator	Protein name	Molecule Type	P Value of Overlap	Activation z-Score
Eosinophilic	**1, 2 & 3**	**SYVN1 (Hdr1)**	**Synoviolin 1**	**Transporter**	**$3.86E^{-04}$**	**-1.84**
	1	SYVN1 (Hdr1)	Synoviolin 1	Transporter	$3.25E^{-05}$	-1.55
	2	MLH1	mutL homolog 1	Ligand-dependent nuclear receptor	$2.75E^{-04}$	N/A
	3	AIP	Aryl hydrocarbon receptor interacting protein	Transcription regulator	$2.16E^{-03}$	N/A
	4	TNF	Tumor necrosis factor	Cytokine	$1.61E^{-06}$	-3.85
Highly atopic	**5 & 6**	**TCR**	**T-cell receptor**	**complex**	**$9.83E^{-07}$**	**1.38**
	5	TNF	Tumor necrosis factor	Cytokine	$4.27E^{-06}$	-4.93
	6	LONP1	lon peptidase 1, mitochondrial	Peptidase	$1.67E^{-03}$	-2.70
	7	Calpain	N/A	Complex	$3.38E^{-05}$	N/A
Neutrophilic	**8, 9 & 10**	**IFNγ**	**Interferon gamma**	**Cytokine**	**$3.07E^{-18}$**	**5.64**
	8	AIMP2	Aminoacyl tRNA synthetase complex interacting multifunctional protein 2	Other	$5.39E^{-04}$	-0.45
	9	IFNγ	Interferon gamma	Cytokine	$7.84E^{-10}$	3.18
	10	IFNα	Interferon alpha	Cytokine	$6.98E^{-16}$	5.46

ACCEPTED MANUSCRIPT

Table S9

Upstream regulator	Proteotype (asthma sub-phenotype) 1	6	4	7	3	8	2	9	1	10
SMARCA4	-2.5	-2.8	-2.9	-4.1	2.4	1.4	0.7	2.7	2.6	4.7
IFNL1	-3.3	-2.4	-2.3	-	-4.2	-2.4	-6.0	0.0	-	5.2
TNF	-3.5	-2.6	-4.2	-3.5	-0.2	1.8	-1.5	2.6	-	5.8
IFNG	-4.3	-2.7	-4.1	-	-4.3	-	-4.7	-	-	5.2
IFNA2	-3.6	-2.3	-2.8	-	-3.4	-3.2	-5.1	-	-	4.5
IL13	3.8	2.4	3.2	-2.9	2.8	-	4.3	-1.9	1.9	-1.7
MAPK1	3.1	2.2	2.5	-	3.7	2.0	5.2	-	2.3	-3.6
STAT1	-3.3	-2.7	-2.3	-	-3.1	-	-4.2	-	2.1	5.1
Interferon alpha	-3.3	-2.3	-2.5	-	-2.9	-	-3.0	1.3	2.2	5.0
TGM2	-3.3	-	-3.4	-1.8	-3.5	-	-3.4	1.1	-	4.7
TLR7	-2.7	-	-3.2	-	-	2.4	-2.4	2.5	2.8	3.3
IL1RN	2.5	2.7	2.0	-	3.8	-	4.1	-	-	-3.6
PRL	-2.9	-	1.2	-	-3.9	-	-4.6	-0.4	-	5.2
NFkB (complex)	-2.8	-	-2.6	-2.6	-	2.0	-	2.7	-	5.0
TREM1	-	-	-1.8	-	2.2	2.4	3.4	1.2	3.8	2.1
EIF2AK2	-	-1.9	-2.5	-	-2.9	-2.3	-2.7	1.2	-	2.5
TLR7/8	-3.0	-2.5	-3.5	-2.2	-	-	-	2.0	-	2.3
CEBPA	-2.2	-	-3.2	-1.9	-	-	-2.5	3.1	-	2.6
RARA	-0.6	1.1	0.8	-2.0	1.6	0.5	4.2	1.6	2.1	-0.9
TLR4	-2.6	-	-2.8	-1.8	2.4	-	-	2.2	-	3.1
MGEA5	-	2.6	-	2.8	-2.6	-	-3.0	-0.9	-0.9	-1.5
TLR9	-3.2	-	-2.8	-	-	-	-2.2	-	2.1	3.6
ECSIT	-2.0	-2.2	-3.2	-	-	-	-	3.2	-	3.3
CD40LG	-1.9	-	-3.3	-2.0	-	2.1	-	1.6	-	2.9
FOXA1	-	1.8	2.4	1.8	-	-0.6	0.5	-2.4	-2.2	-1.9
NUPR1	-2.0	-	-4.4	-	-	1.6	0.8	-	3.2	1.6
JUN	-	-2.6	-1.3	-	1.2	-2.1	2.1	2.4	-	1.9
IL2	-1.5	-	-2.4	-	1.5	1.0	0.8	0.9	2.9	2.4
CD3	-	-	1.0	-	-	-	-2.9	-2.8	-2.9	-3.4
FOXO3	-	-	-1.9	-2.0	-	-	2.4	2.2	2.4	2.0
Cg	-1.8	-	-0.9	-2.7	-	-1.0	2.3	-	1.9	1.5
TCR	0.4	-	-2.3	-2.4	-	-	0.9	2.8	2.5	0.9
FOXO1	-2.8	-	-2.3	-	-	-	1.8	-	2.4	2.8
SAFB	2.0	1.9	2.5	-	-	-	-	-2.1	-	-3.1
TAB1	1.9	2.2	2.2	-	-	-	1.9	-	-	-3.2
FOXL2	-	-	-3.0	-2.2	1.8	-	-	2.0	-	1.9
PGR	-	-2.0	-2.9	-	-	-	-2.0	1.2	-	2.7
CD28	-	-	0.3	-	-	-	-2.6	-2.0	-2.7	-3.0
PDGF BB	-	-	-3.3	-	-	-	-	2.3	1.9	3.1
MYOCD	-	-	-2.0	-2.8	-	-2.8	-	-	3.0	-
ETS1	-2.2	-2.0	-2.2	-	-0.5	-	-1.0	2.4	-	-
EGR1	-2.0	-	-2.4	-2.0	-	-	-	-	1.6	2.1
IFNB1	-	-	-	-	-2.8	-	-3.6	-	-	3.6
RELA	-	-	-2.7	-2.1	-	-	-	1.2	0.8	3.0
PPP2R5C	-	-	1.6	-	-	-	-2.0	-1.6	-2.0	-1.9
TCF7L2	-	-	-	-2.4	2.1	-	2.1	2.2	0.4	-

ACCEPTED MANUSCRIPT