



Published in final edited form as:

Cell Rep. 2021 April 13; 35(2): 108974. doi:10.1016/j.celrep.2021.108974.

High-dimensional profiling clusters asthma severity by lymphoid and non-lymphoid status

Matthew J. Camiolo^{1,2}, Xiaoying Zhou^{3,10}, Timothy B. Oriss^{1,10}, Qi Yan⁴, Michael Gorry¹, William Horne⁵, John B. Trudeau⁶, Kathryn Scholl¹, Wei Chen⁴, Jay K. Kolls⁷, Prabir Ray^{1,8}, Florian J. Weisel⁹, Nadine M. Weisel⁸, Nima Aghaeepour⁹, Kari Nadeau^{3,10}, Sally E. Wenzel^{1,6,8,10}, Anuradha Ray^{1,8,10,11,*}

¹Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

²Center for Systems Immunology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

³Sean N Parker Center for Allergy Research and Division of Pulmonary, Allergy, and Critical Care Medicine, Stanford University, Stanford, CA, USA

⁴Division of Pulmonary Medicine, Children's Hospital of Pittsburgh, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

⁵Richard King Mellon Foundation Institute for Pediatric Research, Children's Hospital of Pittsburgh, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

⁶Department of Environmental Medicine and Occupational Health, Graduate School of Public Health, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

⁷Department of Medicine and Center for Translational Research in Infection and Inflammation Tulane School of Medicine, New Orleans, LA, USA

⁸Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

⁹Departments of Anesthesiology, Pain, and Peri-operative Medicine and Biomedical Data Sciences, Stanford University, Stanford, CA, USA

¹⁰These authors contributed equally

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*Correspondence: raya@pitt.edu.

AUTHOR CONTRIBUTIONS

Conceptualization, A.R., S.E.W., K.N., and M.J.C.; methodology, M.J.C., X.Z., T.B.O., Q.Y., W.H., K.S., F.J.W., N.M.W., J.K.K., W.C., M.G., J.B.T., S.E.W., A.R., and P.R.; data analysis, M.J.C., T.B.O., Q.Y., S.E.W., N.A., and A.R.; writing, A.R., M.J.C., S.E.W., and T.B.O.; supervision, A.R., S.E.W., K.N., J.K.K., and W.C.; funding acquisition, A.R., S.E.W., J.K.K., M.J.C., P.R., T.B.O., and K.N.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2021.108974>.

DECLARATION OF INTERESTS

A.R. has a research agreement with Pieris Pharmaceuticals. S.E.W. is a consultant for AstraZeneca, Glaxo Smith-Kline, and Sanofi. She is also involved in clinical trials being run by Knopp, Sanofi, and AstraZeneca. She has a research agreement with Pieris Pharmaceuticals. M.J.C. is a consultant for Pieris Pharmaceuticals.

INCLUSION AND DIVERSITY

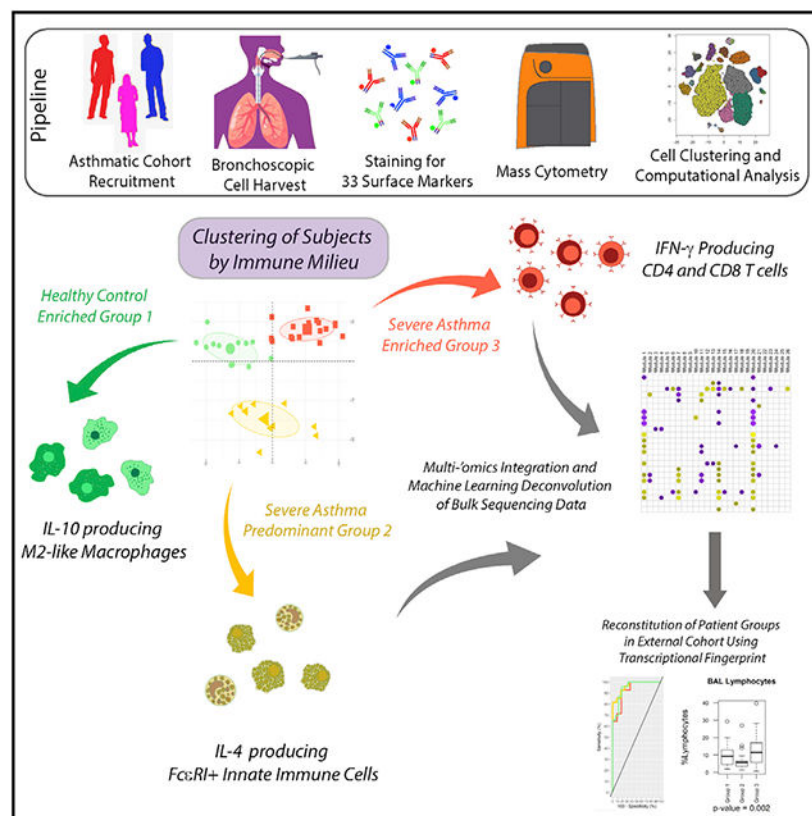
We worked to maintain gender balance and ethnic diversity in our selection of human subjects. We have also ensured gender balance in citations of previous work.

¹¹Lead contact

SUMMARY

Clinical definitions of asthma fail to capture the heterogeneity of immune dysfunction in severe, treatment-refractory disease. Applying mass cytometry and machine learning to bronchoalveolar lavage (BAL) cells, we find that corticosteroid-resistant asthma patients cluster largely into two groups: one enriched in interleukin (IL)-4⁺ innate immune cells and another dominated by interferon (IFN)- γ ⁺ T cells, including tissue-resident memory cells. In contrast, BAL cells of a healthier population are enriched in IL-10⁺ macrophages. To better understand cellular mediators of severe asthma, we developed the Immune Cell Linkage through Exploratory Matrices (ICLite) algorithm to perform deconvolution of bulk RNA sequencing of mixed-cell populations. Signatures of mitosis and IL-7 signaling in CD206⁻Fc ϵ RI⁺CD127⁺IL-4⁺ innate cells in one patient group, contrasting with adaptive immune response in T cells in the other, are preserved across technologies. Transcriptional signatures uncovered by ICLite identify T-cell-high and T-cell-poor severe asthma patients in an independent cohort, suggesting broad applicability of our findings.

Graphical abstract



In brief

Asthma is an immune-mediated heterogeneous disease. Current type-2 biomarkers provide limited understanding of underlying pathobiology or therapeutic guidance. Using a multi-omics approach,

Camiolo et al. identify two clusters of severe asthma patients with similar type-2 biomarkers but with strikingly distinct immune profiles and associated biological pathways.

INTRODUCTION

The term “asthma” encompasses a disease spectrum characterized by reversible airflow obstruction affecting more than 300 million people worldwide (Pavord et al., 2018; Ray et al., 2020; Wenzel, 2012). Clinical manifestations of asthma are heterogeneous, with 5%–10% of patients exhibiting poor response to inhaled corticosteroids (ICSs), the current standard of care, even when used at high doses or accompanied by systemic administration (Fahy, 2015; Gauthier et al., 2015; Ray et al., 2016, 2020; Wenzel, 2012). The economic impact of severe asthma (SA) in the U.S. is estimated at \$56 billion annually, driven primarily by frequent exacerbations and need for emergency care (Barnett and Nurmagambetov, 2011).

Initial studies of asthma heterogeneity clustered patients using combinations of clinical and biological variables such as sputum, bronchoalveolar lavage (BAL) fluid or blood cell counts, and age of onset (Haldar et al., 2008; Moore et al., 2007; Wu et al., 2019). While they did not identify specific biologic mechanisms, they paved the way for molecular phenotyping. Subsequent work focused on transcriptional profiling using platforms such as microarray and RNA sequencing of epithelial brushings, sputum, and, more recently, BAL cells (Baines et al., 2014; Hekking et al., 2017; Kuo et al., 2017; Modena et al., 2017). The concept of type-2 (T2)^{Hi} asthma and its associated biomarkers, including blood eosinophils and fraction exhaled nitric oxide (FeNO), was promoted in different reports (Hastie et al., 2013; Modena et al., 2014; Woodruff et al., 2009). Molecular signatures were associated with these biological variables using bioinformatic tools such as weighted gene correlation network analysis (WGCNA) and k-means clustering, confirming T2 immune pathway activation (Haldar et al., 2008; Modena et al., 2017; Weathington et al., 2019). Targeting these pathways using therapeutics directed at immunoglobulin E (IgE), interleukin (IL)-4R α , IL-5, or its receptor has confirmed their importance in many patients (Bleecker et al., 2018; Brusselle et al., 2017; Hanania et al., 2013; Ray et al., 2020; Wenzel et al., 2013). Response to targeted therapy remains variable, however, and does not fully correspond with biomarkers used in patient selection (Brightling et al., 2015; Flood-Page et al., 2007; Leckie et al., 2000; Panettieri et al., 2018). Possible explanations for this lack of correspondence include heterogeneity in the cellular sources of T2 inflammation, activation of concomitant signaling pathways, and limitations in currently available biomarkers. Taken together, many gaps remain in our understanding of immune and molecular characteristics that underlie the umbrella term asthma.

No study heretofore has conducted an in-depth analysis of lung immune cell populations as identified by cell-surface phenotype or the cytokines they produce in a range of rigorously phenotyped SA patients. We performed bronchoscopy to collect airway and alveolar immune cells and used mass cytometry and RNA sequencing to identify cells and pathways salient to disease in these participants. We show that treatment-refractory patients can be stratified into two distinct groups based on cellular immune profile and accompanying gene

signature. Despite broad differences in BAL cell count, sources of T2 cytokines, and coinciding type-1 (T1) pathway activation, these groups were inscrutable by clinically used markers of T2 inflammation. BAL cells in one patient group (PG) were enriched in IL-4⁺FcεRI⁺ innate immune cells that were distinct from gene expression signatures linked to mast cells or basophils, while those in the second were dominated by interferon (IFN)-γ⁺ T cells, including tissue-resident CD4⁺ and CD8⁺ memory (TRM) cells. Both PGs exhibited a reduction in IL-10⁺ alternatively activated M2-like macrophages as compared to the group that housed all healthy subjects.

The varying composition of mixed-cell samples may complicate bulk transcriptional analysis by introducing variance that obfuscates meaningful discovery. We have developed and validated a machine learning algorithm that performs deconvolution of mixed gene expression data called Immune Cell Linkage through Exploratory Matrices (ICLite), publicly available as a package for R Statistical Software (R Development Core Team, 2014; <https://github.com/camiolomj/ICLite/>). Compared to current technology, ICLite provided deeper insight into pathways linked to important cells from our mass cytometry analysis while also requiring less parameter tuning and providing integrated estimation of gene module number. Signatures uncovered by ICLite also identified groups of treatment refractory patients with impaired lung function from external cohorts. Importantly, these groups varied significantly in the lymphocyte content of their BAL, suggesting broad applicability of our findings to the study of SA.

RESULTS

Mass cytometry identifies diversity of BAL immune cell populations in asthma

Using clinical definitions, 7 healthy controls (HCs), 15 mild-moderate asthma (MMA), and 19 SA patients were recruited for analysis of bronchoscopically obtained distal airway and alveolar cells. SA was defined by European Respiratory Society-American Thoracic Society (ERS-ATS) criteria, which include treatment with high dose ICSs as well as an additional controller medication for at least 1 year, or systemic CSs for greater than half the preceding year. The majority of asthmatic patients in our cohort, including those with MMA, were treated with ICSs. Table S1 includes baseline characteristics of human subjects enrolled in this study.

Cells recovered in BAL fluid from HC, MMA, and SA participants were used for single-cell mass cytometry with a panel of antibodies (Method details) targeting lineage markers for the adaptive and innate immune systems. Unsupervised cell clustering was performed using the FlowSOM algorithm (Kohonen, 2013; Van Gassen et al., 2015) following gating for viable singlets and removal of artifacts. The resultant clusters comprised cells from HCs and MMA and SA patients (Figure 1A).

Examining surface marker staining intensity on t-stochastic neighbor embedding (t-SNE) dimension-reduced space revealed territories harboring mononuclear and lymphocytic cells (Figure 1B; Table S2). Using these data to guide manual validation, we assigned identities to FlowSOM cell clusters by parallel analysis with traditional biaxial gating (Figure 1C; Table S2). This included 7 CD4 T cell subsets, 3 CD8 T cells, as well as individual clusters of

natural killer (NK) cells and polymorphonuclear (PMN)- and $\gamma\delta$ T-cell-like cells. The remaining monocytic cell populations were identified as dendritic cells, CD206^{high} alternatively polarized macrophages with varying expression of FcεRIα, CD206^{low} M1-like macrophages, and three subsets of CD206⁻FcεRIα⁺ cells that occupied distinct territories of dimension-reduced space. As illustrated in Figure S1, the phenotypic uniqueness of FlowSOM cell clusters was determined using surface marker staining. As discussed below, we did not find a convincing link between expression of mast cell- or basophil-related genes with any FcεRIα⁺ cell population. Group 2 innate lymphoid cells (ILC2s) were not detectable in BAL as a discrete cell cluster.

Clustering of patients based on high-dimensional cell count

Seeking to identify immune response patterns that may better characterize patients than traditional evaluation of clinical disease severity, we next clustered cohort participants based on their BAL composition. We identified three groups distinguished by similarity in high-dimensional cell count (Figure 1D). Of these, one housed the entirety of HCs in the dataset as well as a third of MMAs (PG1). The other two comprised exclusively asthma cases (Figure 1E). PG2 accounted for 11 of the 19 total SAs and included only one MMA. PG3 was an admixture of both MMA and SA. Because these groupings were determined by BAL cell count, the 9 MMA patients in PG3 were more similar in immune composition to the SA patients within their group than to the MMA patients from others. Participants across groups varied in a number of clinical parameters (Table S1) including forced expiratory volume in one second (FEV1) (Figure 1F), asthma quality of life, and atopic dermatitis (eczema).

The growing availability of therapeutics targeted at T2 inflammation has led to the use of biomarker assays shown to predict response to these drugs during clinical assessment of asthma patients. Despite dramatic variation in immune cell profiles among PGs, we found no significant difference in the commonly used T2 biomarkers, absolute blood eosinophil count or FeNO, between them (Table S1). This inability to distinguish among PGs suggests that the complexity of underlying immune response in SA is not captured by the T2^{Hi} versus T2^{Low} paradigm alone. Supporting this intrinsic difference in immune response, participants of PG2 were much less likely to have eczema than PG3 despite similar T2 status.

Identifying cellular determinants of lung function

Because our PGs significantly varied in lung function (Figure 1F), we next identified cellular determinants of airway remodeling using elastic net (EN) regression, previously proven effective at handling mass cytometry data (Aghaeepour et al., 2017). Comparison of measured FEV1 to values predicted by high-dimensional cell count showed strong model performance (Figure 1G). Of the cells in our dataset, we found that 4 of the 6 CD206^{high} macrophage populations, all of them FcεRIα⁻, were positively associated with FEV1, indicating a possible protective role (Figure 1H). M2-like, or alternatively polarized, macrophages have been described as a continuum of phenotypes with differential roles in lung homeostasis and disease (Chen et al., 2012; Kambara et al., 2015; Wynn and Vannella, 2016). Cell clusters #8, a CD206^{high}ST2⁺FcεRIα⁺ population, and #19, another M2-like macrophage with high expression of FcεRI, were the only two macrophage populations with negative association to lung function. Numerous T cell subsets were also negatively

associated with lung function, including CD8 effector memory (EM) cells (cell cluster #11), and CD8TRM cells (cell cluster #15) (Figure 1H). The CD4⁺CD161⁺CCR5⁺ T cell cluster #16 showed the strongest negative association with FEV1, followed by the CD206⁻FcεRIα⁺CD127⁺ cell cluster #25 (Figure 1H).

Patient groups are defined by specific cell lineages

We next used mixed-effects modeling of association of single cells (MASC) to identify PGs while controlling for contributions from sex, CS use, and CyTOF batch. Though CSs exert pleiotropic effects on immune cells, self-reported ICS or oral CS use did not differ between PG2 and PG3 (Figure S2A), suggesting BAL composition was not driven by CS response alone. We also found no relationship between cytometry by time of flight (CyTOF) batch and patient cluster (Figures S2B and S2C). PG1 was associated with lung-function-protective, CD206^{high}FcεRIα⁻ macrophages (cell clusters #2 and #17) (Figure S2D). PG2 exhibited a higher proportion of the three CD206⁻FcεRIα⁺ innate lineages (cell clusters #1, #24, and #25) (Figure S2D). PG3, on the other hand, was associated with different populations of T cells, including CD4 EM cells (cell clusters #4 and #16), CD4 central memory (CM) cells (cell cluster #7), CD8 CM cells (cell cluster #11), CD4⁺ TRM cells (cell clusters #12, #14, and #23), and CD8⁺ TRM cells (cell clusters #15 and #18) (Figure S2D).

To validate these findings, we next quantified manually gated cells across patients in our cohort (Figure 2A). PG1 most closely approximated a normal BAL cell differential, with ~90% of recovered cells being monocytic. These participants showed significant enrichment for alveolar macrophages, defined by a CD3⁻CD206⁺CD163⁺HLA-DR⁻ surface phenotype, as well as CD206⁺CD11b⁺CCR4⁺ macrophages akin to FlowSOM cell cluster #17 (Figures 1B and 2B). Patients in PG2 demonstrated greatly increased numbers of CD3⁻CD206⁻FcεRIα⁺ cells but very low numbers of T cells. Using FlowSOM clusters as guide, we identified subsets of CD3⁻CD206⁻FcεRIα⁺ cells with varying expression levels of the IL-7 receptor-α, CD127. CD3⁻CD206⁻FcεRIα⁺CD127⁺ cells could be further subdivided by CCR4 expression, yielding discrete populations of CD3⁻CD206⁻FcεRIα⁺CD127⁻, CD3⁻CD206⁻FcεRIα⁺CD127⁺ CCR4⁻, and CD3⁻CD206⁻FcεRIα⁺CD127⁺CCR4⁺ cells corresponding to FlowSOM clusters 1#, #24, and #25, respectively (Figure 2B). PG2 also showed enrichment for CD206⁺FcεRIα⁺ macrophages with varying levels of CCR4 expression, corresponding to FlowSOM clusters #8 (CCR4⁻) and #19 (CCR4⁺) (Figure 2B). We then gated out CD3⁺ populations, first defining CD4 lymphocytes including naive CD4 T cells (CD4⁺CD45RA⁺CD69⁻), CD4 TRM cells (CD4⁺CD45RA⁻CD69⁺), CD4⁺ EM cells (CD4⁺CD45RA⁻CD69⁻CCR7⁻CD27⁻), and CD4 CM cells (CD4⁺CD45RA⁻CD69⁻CCR7⁺CD27⁺). Of these populations, CD4 TRM cells were significantly increased in PG3, in agreement with our FlowSOM findings. CD4⁺CD45RA⁻CD69⁻CCR5⁺CD161⁺ cells, the phenotypic equivalent of FlowSOM cluster #16, were also enriched in PG3. We performed similar gating of CD8 T lymphocytes, identifying naive CD8 T cells (CD8⁺CD45RA⁺CD103⁻), CD8TRM cells (CD8⁺CD45RA⁻CD103⁺), CD8 EM cells (CD8⁺CD45RA⁻CD103⁻CCR7⁻CD27⁻), and CD8 CM cells (CD8⁺CD45RA⁻CD103⁻CCR7⁺CD27⁺). CD8⁺ CM cells, CD8⁺ EM cells, and CD8⁺ TRM cells were all significantly increased in PG3 (Figure 2B).

Satisfied that manual gating recapitulated findings of automated cell clustering, we next tested whether PGs could be successfully recapitulated with these counts (Figures S2E and S2F). Our supervised learning model for PG prediction used 18 different manually gated cells to identify PG with high accuracy. Mapping of correlative relationships between BAL cells summarizes the dominant immune signatures identified across cohort participants (Figure 2C).

Cytokine production by both T-cell and non-T-cell sources

Our data showed that cohort participants could be grouped based on similarity in their BAL cellular composition. As cytokines have been implicated in asthma pathogenesis and are being targeted in asthma therapy, we examined production of six cytokines associated with asthma, including interferon (IFN)- γ , IL-4, IL-5, IL-13, IL-17, and IL-10. As expected, stimulation with phorbol myristate acetate (PMA) plus ionomycin was required for cytokine detection in T cells, but not for non-T cells (Figure S3A). Projection of cytokine staining intensities across the cell clusters in the t-SNE plots revealed that signals for T2 cytokines such as IL-4, IL-5, and IL-13 localized primarily in non-T-cell regions (Figures 3A and 3B). IFN- γ most intensely overlaid T cell populations, suggesting its expression characteristics were distinct from the other five cytokines surveyed. Quantitation of cytokine production by traditional manual biaxial gating confirmed that only IFN- γ and IL-17A were predominantly produced by T cells (Figure 3C). The canonical T2 cytokines IL-4, IL-5, and IL-13, as well as IL-10, were derived from largely non-T-cell sources. While IL-4 decorated regions shared by IL-5 and IL-13, it also stained heavily within the territory harboring the CD206⁻Fc ϵ RI α ⁺ cells associated with PG2. Thus, although IL-4 is historically described as a product of T helper 2 cells (Th2, natural killer T [NKT]) in allergic diseases including asthma, our analysis revealed non-T cells to be its dominant source in SA.

Inflammatory signals are cellular phenotype specific

Having appreciated the cellular sources of cytokine production, we next sought to unravel how immune milieu may influence signaling. Regulation of maturation, chemotaxis, polarization, and tissue maintenance may be accomplished via cell-to-cell communication. We next evaluated how they may influence cytokine production across cell lineages and PGs.

IL-10 staining was predominantly found in CD206⁺ macrophage populations, including alveolar macrophages, CD206⁺CD11b⁺CCR4⁺ macrophages, and some CD206⁺Fc ϵ RI α ⁺ macrophages. Of these, alveolar macrophages and CD206⁺CD11b⁺CCR4⁺ macrophages showed greater relative (per cell) IL-10 production (Figure 3D) and a greater number of total IL-10⁺ cells per million BAL cells in PG1. Although CD4⁺ T regulatory cells (Tregs), defined by a CD3⁺CD4⁺FoxP3⁺CD25^{high} phenotype, also showed greater relative production of IL-10 in PG1, the number of IL-10⁺ Tregs per million BAL cells was greatest in PG3 (Figure 3D).

As described above, PG2 was enriched for Fc ϵ RI α ⁺ innate immune lineages that included discrete CD206⁺ and CD206⁻ populations. These cells were the major source of IL-4 across our dataset, with ~5-fold-greater relative staining in PG2 and PG3 (Figure 3E), indicating

heightened IL-4 production specifically in the context of treatment refractory asthma. IL-4⁺ CD206⁻FcεRIα⁺CD127⁻, CD206⁻FcεRIα⁺CD127⁺CCR4⁺, and CD206⁺FcεRIα⁺CCR4⁻ cells were all present at the highest counts per million BAL cells in PG2, suggesting possible roles as drivers of disease in these individuals. IL-4 production by CD206⁻FcεRIα⁺CD127⁻ and CD206⁻FcεRIα⁺CD127⁺ innate cells was independently verified using high dimensional fluorescence flow cytometry with relevant isotype control (Figure S4), confirming specificity of IL-4 staining and enrichment of IL-4⁺ cells in participants of PG2. Importantly, CD206⁻FcεRIα⁺ cells did not possess side scatter qualities typical of eosinophils when assessed by traditional flow cytometry. Though IL-5 shared sources with IL-4 (Figure 3B), there was no difference in relative production or cytokine positive cells per million between PGs (Figure 3F) among innate cells, suggesting uncoupling of upstream regulatory mechanisms governing these cytokines.

Dominant sources for IL-4 and other T2 cytokines varied in PG3, suggesting immune milieu may promote or suppress production in a lineage-specific manner. We found IL-4⁻, IL-5⁻, and IL-13-positive cells across numerous T cell subsets (Figure S5A), including CD4⁺ and CD8⁺ EM, CM, and TRM cells. T2 cytokine-positive T cells were increased in count per million among participants in PG3 (Figure S5B), though their number was modest compared to other sources. Individuals in PG3 did, however, demonstrate notably higher relative IFN-γ production from key cell populations such as CD4 and CD8 TRM cells, as well as a dramatic increase in IFN-γ⁺ cells per million BAL cells (Figure 3G), consistent with a prominent T1 response in PG3 not present in PG2.

Although IL-17A was also prominently derived from T cell sources (Figure 3C), we found the number of cytokine-positive cells per million in BAL to be an order of magnitude lower than IFN-γ for CD4 and CD8 TRM and EM populations (Figures 3G and S5C), consistent with a dominant T1 rather than Th17 signature. These data suggest that T1 polarization and antigen response may be intrinsically linked to disease in PG3, though these individuals did not exhibit lower levels of T2 biomarkers when compared to PG2 (Table S1).

PG3 featured MMA and SA patients who clustered together by cellular immune phenotype, suggesting similar enrichment for T cells in their BAL. We next examined whether cytokine production varied by clinical disease severity. The average numbers of IFN-γ⁺ and T2 cytokine⁺ T cells in the MMA versus the SA subjects of PG3 were not significantly different (data not shown). However, the intensity of IFN-γ per cell was higher in specific T cell clusters of SA compared to MMA patients, such as CD8 TRM cells (Figure S5D). Interestingly, although a similar trend was noted, this difference was not evident in CD4 TRMs or CD161⁺ CCR5⁺ EM cells, though the latter did display greater intensity per cell for IL-17A in SA (Figure S5E).

Cell-type-defined patient groups display distinct transcriptional programs

Given the profound differences in cellular immune phenotype identified among three PGs, we wondered how they might manifest at the gene expression level. To this end, we performed bulk RNA sequencing of BAL cells. Libraries for 38 of the 41 subjects included in the mass cytometry study passed through rigid quality control standards for inclusion (Table S3). Counts were normalized using rLog transformation and filtered to mitigate

contribution from genes expressed at low and highly variable levels. The remaining 17,373 genes were used for downstream analyses.

As mentioned previously, CSs have widespread and profound impact on immune cells. We found no difference in the expression of FKBP5, a commonly surveyed CS-inducible gene, between BAL cells of asthma patients taking CS across groups (Figure S6A), indicating comparable medication compliance. Pairwise analysis (Figure S6B) showed very few differentially expressed genes (DEGs) between PG1 and PG2. Comparison of PG1 and PG3 showed similar paucity in DEGs (Figure S6C). In contrast, the immune cells in PG3, when compared to PG2 (Figure S6D), showed numerous DEGs associated with a high T1 immune response. These included, but were not limited to, *IL12RB2*, *TBX21*, *CXCR3*, and *GZMK*.

To gain further insight into the biological processes that differ between our cellular immune phenotypes, we performed gene set enrichment analysis (GSEA) comparing each PG to all others in the dataset (Figure 4A). Among pathways enriched in PG1 were the citric acid cycle, mitochondrial respiration, and maintenance of cellular redox state, a profile generally associated with catabolic metabolism and immune quiescence (Olenchok et al., 2017; Rambold and Pearce, 2018). PG2 showed a surprising enrichment in pathways related to cell proliferation that included mediators of cell-cycle progression, DNA replication, and organization of the mitotic spindle (Figure 4A). This PG also exhibited enrichment for ribosomal synthesis, protein translation, and rapamycin-sensitive genes, consistent with activation of the well-described mechanistic target of rapamycin (mTOR) (Brown et al., 1995). Collectively, immune cells in PG2 subjects appeared to be highly proliferative, suggesting an increase in aerobic glycolysis and anabolic metabolism (Rambold and Pearce, 2018). As might be expected from mass cytometry measurement of cytokine, cells in PG3 were enriched in pathways related to T cell antigen response. Specific pathways such as IL-12 signaling and CD8⁺ T cell activation were identified in these patients, suggesting the importance of T cell maturation, proliferation, and polarization in this group.

Deconvolution of bulk transcriptional data using WGCNA

Differential expression and pathway analysis recapitulated some of the differences seen at a cellular level among PGs but failed at linking biological processes and molecular pathways to specific cell lineages in our study. To accomplish this, we first employed WGCNA, a software package that identifies modules of highly correlated genes and relates them to external sample traits (e.g., BAL cell frequencies measured by mass cytometry) (Langfelder and Horvath, 2008).

WGCNA identified 9 correlated gene modules from our BAL transcriptional data (Figure 4B; Table S4). Module eigengenes showed 16 total positive correlations and 13 negative to the 24 manually gated cell populations from mass cytometry, including several populations of interest from patient clustering. The Gene Ontology (GO) terms these modules represented included cell division and the adaptive immune response, processes that had been independently identified at the patient level (Figure 4C). CD206⁺CD11b⁺CCR4⁺ macrophages, which produce IL-10 and were increased in number in PG1, showed a link to the pink module (Figure 4D). Pink hub genes included components of the AP-1 transcription factor as well as cyclic AMP response element modulator (CREM) (Figure 4E), which was

recently shown to be downregulated in the BAL cells of SA patients (Weathington et al., 2019). The magenta module was correlated with PG2-enriched CD206⁻FcεRIα⁺CD127⁺CCR4⁻ cells (Figure 4F) and had hub genes associated with cytoskeletal arrangements during mitosis (Figure 4G). Magenta was associated with GO terms such as cell division and the IL-7 pathway (Figure 4C), suggesting a possible role for this signaling axis in proliferation. Multiple T cell populations, including CD4⁺ and CD8⁺ TRM cells, were correlated with the green module (Figure 4H), which had hub genes associated with T cell response and T1 polarization (Figure 4I).

ICLite uses high-dimensional cell count to improve module-to-cell mapping

While WGCNA was successful at uncovering some key relationships in our multi-omics dataset, it failed to provide insight into a number of important cell populations. To improve resolution of the deconvolution process, we built a machine learning platform tailored to the specific task of ascribing transcriptional signatures to cells in mixed composition bulk sequencing data. ICLite is a software package that utilizes the known abundance of cells in an admixture to identify the optimal conditions from an input set of gene clustering parameters (Method details; <https://github.com/camiolomj/ICLite/>). Using a scoring system that integrates maximization of integrated completed likelihood (ICL) for gene clustering efficacy with measures of gene module to cell connectivity, ICLite provides users with a framework that is built for the task of mixed population deconvolution. Compared to WGCNA, ICLite requires less parameter tuning and offers a simplified user experience for transcriptional deconvolution.

Using ICLite, we found 45 positive and 40 negative correlations among 24 cell populations and 26 gene modules (Figure 5A; Table S5). Though the ratio of connections made to total possible connections was comparable between WGCNA and ICLite (13.4% versus 13.6%, respectively), ICLite was able to describe relationships to a greater number of unique cells (21 versus 14 for WGCNA) and capture a greater depth of function. This included linkage of alveolar and CD206⁺CD11b⁺CCR4⁺ macrophages to genes related to mitochondrial health (module 13) (Figure 5B), more closely resembling patient level findings from GSEA (Figure 4A) than results from WGCNA. ICLite recapitulated the relationship between CD206⁻FcεRIα⁺CD127⁺CCR4⁻ innate cells and genes related to cell division and IL-7 signaling (Figure 5C). It also successfully associated both populations of CD206⁺FcεRIα⁺ macrophages with Fcγ and Fcε signaling (module 11). The relationship of these cells to neurotrophin or nuclear factor of activated T cells (NFAT) signaling highlights pathways that may be critical in SA pathogenesis for some. The correlation between T cell populations and gene modules representing terms such as adaptive immune response (module 1), T cell co-stimulation, and positive regulation of IFN-γ production (module 20) was similarly reproduced (Figure 5D). In addition, ICLite identified specific relationships between CD8⁺ T cells and modules 7 and 14, indicating a level of resolution not afforded by WGCNA.

Mast cells have long been implicated in asthma pathogenesis and are known to express FcεRI. To help resolve the identity of the CD206⁻FcεRI⁺ cells identified by mass cytometry, we compared the frequency of each of our manually gated cell types to the expression of 42 curated mast cell/basophil genes, finding multiple positive correlations (Figure S7A). The

majority of these relationships, however, were between canonical mast cell transcripts such as KIT, CPA3, and TPSAB1 (tryptase beta-1) and T cell populations such as CD8⁺ CM cells and EM cells rather than CD206⁻ FcεRI⁺ cells (Figure S7B). Hypergeometric overlap of our mast cell/basophil gene list and modules identified by deconvolution supported this relationship, with the T-cell-associated green module (WGNCA) and module 14 (ICLite) showing the strongest relationship (Figures 4B, 5A, and S7C). Indeed, most of the patients with very high expression of mast cell genes were in PG3 and relatively depleted of the CD206⁻FcεRI⁺ cells that defined PG2 (Figure S7D). This confirmed prior reports associating mast cells with T cells in asthma (Gauthier et al., 2017). CD127 expression by subsets of these cells further exclude a basophil identity. Collectively, the data showed that mast cells and basophils in BAL were present below a frequency that was detectable using our cell-based technology. As mentioned previously, the side scatter properties of CD206⁻FcεRI⁺ populations also exclude eosinophils as a possible identity of these cells (Figure S4).

ICLite recovers a broad array of ontological functions

Our algorithm ICLite confirmed relationships between genes and cells from BAL that were detected by WGCNA while also expanding information for a number of key populations. Inspecting gene module membership across technologies, we observed that several of the ICLite modules comprised genes from a single WGCNA module (Figure 6A). This included the T-cell-linked modules 1, 14, and 20, which consist of genes found in the WGCNA green module. Modules 11 and 25, which connected to CD206⁺FcεRI⁺ IL-4-producing macrophages and described biologically significant processes such as Fcε signaling, contained genes from the yellow module, which itself did not associate with any cells. Other ICLite modules, such as 10, were an admix of multiple WGNCA networks. Despite this, nearly all of the magenta (mitosis) module from WGNCA was included in module 10, which was enriched for similar GO terms as magenta and again associated with CD206⁺FcεRI⁺CD127⁺ cells in BAL. These data suggest that key relationships between genes and cells were robust to technology used and support validity of the additional findings uncovered by ICLite.

Our modules of correlated genes could be linked to ontological functions as described above. Because membership for each gene in our BAL dataset is exclusive and non-overlapping, we used GO semantic similarity to compare functional terms represented by each ICLite module (Figure 6B; Table S6). We found striking similarity in functional terms for modules that shared relation with correlated cells, reinforcing the biological plausibility of our solution and suggesting that our deconvolution process had been successful. Modules identified by ICLite could be linked to ~2-fold more GO terms than those from WGCNA (Figure 6C), with ~4-fold-greater GO terms effectively linked to cells. Using GO semantic similarity to summarize terms linked to cells by both modules, ICLite showed greater breadth in parent terms without losing key families such as cell division and T cell co-stimulation (Figure 6D).

Using BAL gene modules to classify patients in external cohorts

Persuaded that our modules showed a rational relationship with the cells in our study, we sought to validate our findings in another well-characterized asthma cohort. To this end, we employed BAL microarray data from the multi-center SA Research Program (SARP) cohort (Jarjour et al., 2012; Moore et al., 2007), which comprised 154 patients enrolled in phases I & II of the program. Though mass cytometry data were not available from SARP, we reasoned that expression of genes from our ICLite modules may be used as a surrogate. To accomplish this, we first trained a model for predicting PG based on expression of ICLite modules in the immune modulation in severe asthma (IMSA) (experimental) cohort (Figure 7A). Our model was highly effective in predicting patient grouping determined by high-dimensional cell count using strength of expression scoring for the 26 ICLite gene modules (Figure 7B).

After training on IMSA participants, clustering of SARP again yielded two SA-enriched groups and one predominantly comprising HCs (Figure 7D). Importantly, manual cell count differential of BAL showed the highest percentage of lymphocytes in SARP PG3 (Figure 7D), redemonstrating a key finding from IMSA. The four ICLite modules that correlated with T cells in IMSA showed a similar correlation with percent lymphocyte count from the BAL of SARP participants (Figure 7E), reaffirming their relationship with BAL T cell content. Clinical characteristics such as lung function were also similar across SARP PGs (Figure 7F). As further evaluation of our prediction model, we performed GSEA, again observing enrichment for terms related to proliferation in SARP PG2 (Figure 7G) and T cell antigen response in SARP PG3 (Figure 7F). We were thus able to reconstitute groups of treatment refractory asthma patients with impaired lung function that fundamentally varied in their BAL lymphocyte content, increasing confidence that our findings could be generalized beyond our cohort.

DISCUSSION

With the understanding that a diagnosis of asthma insufficiently reflects disease pathogenesis comes the need to define the disease with increased molecular precision. The rise in T2-targeted biologic therapy has had substantial impact on treatment, reducing systemic CS dependency. While targeted biologic therapy is highly efficacious for many, others fail to respond despite exhibiting elevated levels of T2 biomarkers. In this study, two groups of predominantly SA patients were identified that were poorly distinguished by T2 biomarkers (eosinophilis and FeNO) despite broad differences in lower airway immune milieu. Similarity in blood eosinophil counts may be attributed to roughly equivalent abundance of IL-5⁺ innate cells across PGs. The ability of both T2 cytokines (IL-4/IL-13) and IFN- γ to induce NOS2 may further explain why median FeNO levels were indistinguishable between groups (Spahn et al., 2016). Thus, neither biomarker revealed the striking and potentially treatment-modifying differences in immune profiles in the two PGs.

Our studies suggest that lung health, as observed in HC- and MMA-dominated PG1, associates with enrichment for IL-10-expressing macrophages, known to be regulatory in nature (Mantovani et al., 2004). These cells positively associate with lung function and are linked to genes related to mitochondrial fusion. For example, mitofusin2 mediates important

functions in macrophages, such as phagocytosis and cytokine production, in response to pathogen sensing (Tur et al., 2020). In contrast, IL-4-producing innate immune cells appear to replace these populations in PG2 and associate with lower lung function. Strikingly, a second PG (PG3) showed marked enrichment for IFN- γ^+ T cells, including TRM cells, also associated with lower lung function.

As a machine learning tool for deconvolution of bulk transcriptional data from mixed populations, ICLite was able to preserve important relationships between immune lineages in our dataset while capturing additional links to important cell types. The estimation of the underlying cluster number remains a field of ongoing research. One advantage to ICLite over WGCNA is the integration of a goodness-of-fit metric that removes manual parameter testing from the cluster estimation process. As the identification of plausible correlation modules has been previously inferred by successful GO analysis, we believe our algorithm was highly successful at functional term enrichment as well as linkage of those processes to cells. For the specific task it was built for, ICLite is a powerful complement to currently available packages.

BAL cells enriched in PG2 linked with signatures of cell division. BAL cells in this group were also associated with IL-7, Fc ϵ RI, Fc γ R, and FcR signaling, which may in turn stimulate calcineurin-NFAT signaling, also detectable in the cells. Notably, NFAT plays an important role in IL-4 gene transcription, though this relationship has been described predominantly in CD4 T cells (Wierenga and Messer, 2000), which may relate to the greater number of IL-4⁺ Fc ϵ RI⁺ cells in PG2. It is important to again note that these Fc ϵ RI⁺ cells do not associate with genes commonly expressed by mast cells.

As discussed above, IL-7 signaling was evident in the innate cells enriched in PG2 subjects. Since IL-7R α is a component of the thymic stromal lymphopoietin receptor (TSLPR), these Fc ϵ RI α^{bright} CD127⁺ cells may be targeted by anti-TSLP therapies, which have shown promise in the treatment of SA (Corren et al., 2017). As recently reported, the phase 3 trial of tezepelumab met its primary endpoint in reducing exacerbations in SA patients (Menzies-Gow et al., 2020; Wechsler et al., 2020). Enrichment of these cells in PG2 is evocative of *in situ* hematopoiesis, accomplished via expansion of tissue-resident pluripotent progenitors in mucosal tissue. *In situ* hematopoiesis has been associated with the epithelial cytokines IL-33 and TSLP (Li et al., 2018). Increased numbers of progenitor cells expressing TSLPR, CD127 (IL-7R), and ST2 were detected in the sputum of asthma patients in response to allergen challenge (Sergejeva et al., 2005). The combined signatures of cell proliferation and IL-7 signaling revealed by both WGCNA and ICLite invite the notion of cytokine-driven expansion of progenitor cells that differentiate into IL-4-secreting pathogenic cells. It will be interesting to determine whether polyclonal IgE (non-cross-linking) plays a role in the accumulation of these cells given the ability of non-cross-linking IgE to promote survival of Fc ϵ RI-expressing mast cells (Kawakami and Galli, 2002). While the exact identities of the IL-4-producing Fc ϵ RI α^{bright} CD127⁻, Fc ϵ RI α^{bright} CD127⁺CCR4⁺, and Fc ϵ RI α^{bright} CD127⁺CCR4⁻ cells that defined PG2 remain to be determined, it is possible that upstream mechanisms triggered by epithelial events play a role in their development.

In contrast, cells from PG3 patients were associated with gene modules tied to the adaptive immune response and positive regulation of IFN- γ , concordant with the preponderance of IFN- γ ⁺ T cells in PG3. We previously reported an increased T1 response in the airways of SA patients (Raundhal et al., 2015). A similar steroid-resistant T1 phenotype was also reported in a pediatric population (Wisniewski et al., 2018). One important insight gained in our study is the presence of high levels of CD4⁺ and CD8⁺ TRM cells expressing IFN- γ in the airways of PG3 subjects. CD8⁺ TRM cells have not been previously described within asthmatic airways, and their presence suggests a role for microbial pathogens (viruses) in disease etiology. A study of immune response in lung transplant patients has shown that numbers of TRM cells in BAL fluid closely correspond with those in the lung tissue (Snyder et al., 2019), suggesting our assessment of BAL is representative of adjacent tissue. CD8⁺ TRM cells provide host protection in the skin (Ariotti et al., 2014) and defend against respiratory viral infection (Jozwik et al., 2015). Their numbers have been found to peak during viral season and contract over time (Jozwik et al., 2015). For safety reasons, our cohort participants were not bronchoscoped within 6 weeks of infectious symptoms, suggesting their immune profiles were reflective of chronic changes. A recent study has documented higher circulating levels of rhinovirus-specific T1 cells in uninfected asthma patients with worse lung function (Muehling et al., 2020). It is interesting to note that DPP4/CD26 and SLAMF1 were present in modules associated with CD4⁺ and CD8⁺ TRM cells using both WGCNA and ICLite, given their well-established roles in T cell co-stimulation and IFN- γ production (Araki et al., 2009; Aversa et al., 1997; De Meester et al., 1999). Blood DPP4 was suggested and then abandoned as a T2 biomarker (Peters et al., 2016). Our findings suggest it may be more specific to the phenotype associated with PG3 patients. Notably, DPP4/CD26 expression has been also linked with virus-specific memory CD8⁺ T cells (Ibegbu et al., 2009). Thus, persistence of activated TRM cells following microbial clearance may be pathogenic, and mechanisms that prevent their reduction must be addressed in future studies. In this regard, it is important to remember that CD8⁺ T cells have also been associated with autoimmunity (Liblau et al., 2002).

The frequency of Th2 cells expressing IL-4, IL-5, or IL-13 was low in all patient subgroups, suggesting that Th2 cells in general may be a minor component in the T2 immune pool in the context of CS therapy. Similarly, ILC2 cells expressing T2 cytokines were not detectable in the BAL cells of the patients in our analysis. This is probably because of the responsiveness of both cell types to CSs, as previously shown (Jee et al., 2005; Nagakumar et al., 2019). It is worth noting that although IL-17⁺ T cells were detected in PG3, they were far fewer in number than IFN- γ ⁺ T cells across all T cell clusters. Of the subsets identified, CD4⁺CD161⁺CCR5⁺ EM cells showed the greatest skewing toward IL-17A production, though still demonstrating staining for T1 and T2 cytokines. Unlike other T cell populations, these cells expressed high levels of CD161, which has been associated with highly pathogenic Th17 cells in Crohn's disease (Kleinschek et al., 2009) and rheumatoid arthritis (Basdeo et al., 2017). CD4⁺CD161⁺ T cells also demonstrate plasticity enabling IFN- γ production (referred to as "ex-Th17" cells) and are refractory to suppression by Tregs and steroids (Basdeo et al., 2017). The relative median signal intensity of IL-17 and IFN- γ from the CD161⁺ cells also showed a higher trend in SA patients compared to MMA patients in PG3 (Figure S5). Thus, although fewer in number than TRMs, the ability of CD161⁺ T cells

to produce IL-17, IFN- γ , and T2 cytokines may contribute to poor lung function in some asthmatic patients.

We understand that the study, although the largest of its kind, is still small with respect to cohort size. This may have limited our ability to dissect additional cellular phenotypes, including those of patients with nasal polyps. We also appreciate that the lung is an anatomically complex organ with multiple compartments, of which our study addressed only the luminal compartment of the small airways and the alveoli.

In summary, our study used mass cytometry and RNA sequencing coupled with machine learning tools to identify two distinct molecular mechanisms, potentially driving disease pathogenesis in two SA-enriched PGs, who were largely indistinguishable by clinical measures of disease severity and T2 biomarkers. Importantly, the segregation of patients into a lymphocyte-poor (PG2) versus a lymphocyte-rich (PG3) group was also evident in an independent patient cohort, using a clustering approach based on gene module scoring. These findings go well beyond the broad and simplistic characterization of asthma as T2^{high} or T2^{low}, which provides limited help in determining treatment options. The presence of similar immune and molecular signatures in the airways of two independent cohorts suggests broad applicability of our findings to not only differentiate between health and disease, but also to uniquely distinguish two immune phenotypes of SA. Future studies to determine their responses to targeted therapies should move us closer to appreciating true endotypes of asthma.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Anuradha Ray (raya@pitt.edu).

Materials availability—This study did not generate new unique reagents.

Data and code availability—The GEO accession number for the RNA-seq data is GEO: GSE136587 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE136587>). FCS files of mass cytometry data are available through FlowRepository (<https://flowrepository.org/id/FR-FCM-Z395>). The ICLite package and documentation are available for download at <https://github.com/camiolomj/ICLite/>. Immune cell count and BAL transcriptional data from IMSA participants are included as part of the R package and are preformatted to run in ICLite. Additional code used in this analysis represents implementation of previously published work as noted below and will be available upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Study participants—All participants provided informed consent in accordance with an IRB protocol approved by the University of Pittsburgh. Male and female nonsmoking (< 10 pack year and no smoking in the previous year) subjects meeting the ERS/ATS definition of severe asthma (Chung et al., 2014) were recruited. Also recruited were milder asthma

subjects on no or lower doses of inhaled CS (Forced Expiratory Volume 1 > 60% predicted with no history of frequent or severe exacerbations of asthma in previous year) and healthy controls. Subjects comprised all racial/ethnic backgrounds and between the ages of 18 and 60 years old. Subjects underwent extensive baseline characterization, including physiologic (spirometry, diffusing capacity and PC20), allergic and clinical characterization (Moore et al., 2007). Subjects also underwent bronchoscopy per published protocols and consistent with procedures established at the University of Pittsburgh consistent with the severe asthma research program.

METHOD DETAILS

CyTOF sample preparation—Cryopreserved BAL cells were thawed and incubated in RPMI 1640 media containing 10% FBS at 37°C for overnight resting. Cells were incubated in the absence or presence of PMA (20ng/ml) and ionomycin (1 ug/ml) for 4 hours. A protein transport inhibitor-Brefeldin A (BFA, 5ug/ml) was added to the culture of unstimulated (treated with BFA alone) and stimulated (treated with BFA, PMA and ionomycin) samples for 4 hours. Cells were harvested and stained for viability with 5uM cisplatin (Fluidigm) for 5 minutes at room temperature before quenching with CyTOF Buffer (Fluidigm). Cells were then incubated with Fc Receptor Blocking solution (Human TruStain FcX, Biolegend) for 10 min at room temperature and stained with cell surface antibody cocktail for 45 minutes on ice. Samples were washed with CyTOF Buffer and fixed at 4°C overnight with 2% paraformaldehyde. On the following day, cells were permeabilized (permeabilization solution from eBiosciences) and stained with antibodies for 6 cytokines for 45 minutes at room temperature. Cells were washed and incubated in 0.25µM Ir intercalator (Fluidigm) for 20 min at room temperature, and stored in PBS at 4°C until acquisition. Antibodies were purchased from DVS Biosciences or conjugated in-house using purified antibodies and Fluidigm conjugation kit following manufacture’s instruction. All the antibodies were titrated prior to use.

CyTOF data acquisition—Samples were washed twice with PBS and de-ionized water consecutively and then resuspended at a concentration of 0.5 million cells/ml in deionized water containing a 1/4 dilution of Equation 4 Element Calibration Beads (Fluidigm). The samples were acquired on a Helios CyTOF instrument (Fluidigm) at an event rate of < 350 events/second. To minimize the effect of experimental variability on mass cytometry measurements between samples from clinical status groups, cryopreserved BAL cells from severe asthma, mild to moderate asthma and healthy control samples were processed, stained, and run within the same batch for a total of 5 batches of samples, with 2 days of acquisition for each batch. Instrument tuning was performed daily to minimize oxide-induced signal crosstalk during data acquisition and to find the balance between detection sensitivity and oxide formation. Following instrument tuning, Fluidigm’s EQ Four Element Calibration Beads were used to track daily instrument performance. EQ beads contain four natural abundance elements including cerium (Ce), europium (Eu), holmium (Ho), and lutetium (Lu) with a total of seven isotopes distributed across the CyTOF mass window. The median signal intensities of Eu151 and Eu153 on the EQ bead singlet population were determined following instrument tuning. The coefficient of variance (CV) of median intensities of Eu151 and Eu153 was collected on each of the 10 running days covering all of

the BAL samples included in this study. Given that beads have identical ion components, the instrument performance was monitored for precision and repeatability by comparison of CV of the bead signal intensities. The CVs of Eu151 and Eu153 were 10.9% and 10.5%, respectively across the 10 days of sample acquisition. In addition, in order to account for the variability of instrument performance and therefore represent the real biological difference among samples, each sample was spiked with the EQ beads as described above as a constant reference prior to data acquisition, which allowed for normalization of each sample's signal intensity. After acquisition, the raw data files were normalized using bead-based normalization using CyTOF 6.7 software (Fluidigm) with a "Bead passport" (EQ-P13H2302_ver2).

The intact and viable cells were gated in order to exclude normalization beads, dead cells, cell debris and cell doublets using FlowJo software (Tree Star, Inc.) prior to downstream computational analysis or/and additional analysis with FlowJo. FlowJo analysis for cell types identifiable by primary gating using cell surface phenotype (e.g., total CD3 cells) was performed essentially according to conventional flow cytometry practice. An additional level of stringency for cytokine expression in particular was that steps were taken to place thresholds of positivity to ensure parity with machine-based determination of expression as described below. For analysis of post-FlowSOM cell clustering solution files, the primary gate was set on the cell cluster of interest on plots of cluster number designation versus CD45 or another relevant cell type-specific surface marker (e.g., CD4 or CD8). From there, conventional analysis was performed using the criteria described above.

Immune cell identification using machine learning—The R implementation of FlowSOM through the cytofkit package (version 3.7) was used to cluster cells together based on expression of 33 surface markers. Cell clustering was performed using 20,000 arcsin transformed events per participant. A maximum number of clusters of 40 was tested using ConsensusClusterPlus with an accepted solution of 28 returned by "consensus" output. As the purpose of this study was identify patient groupings by cell count, our results could be skewed by overclustering of cells leading to Type I error driven by cell clusters with minimal difference in cell surface phenotype. To combat the effect of redundant cell clusters on our downstream analysis, we compressed all clusters that demonstrated a cosine similarity in mean surface marker intensity greater than 0.80. The clusters effected were 6 monocytic/macrophage populations that were compressed to 3. This resulting 25 clusters from FlowSOM analysis were then used for downstream analysis. Biological veracity and homogeneity of cell clusters was investigated using traditional biaxial gating via FlowJo software as described above.

t-SNE plot preparation—Staining intensity for 33 surface markers included in our mass cytometry experiments were used for optimized T-distributed stochastic neighbor embedding (opt-SNE) plot generation (Belkina et al., 2019). Scaled staining intensity is presented for individual surface markers as indicated.

Handling of compositional data—The resultant distribution of each subject's BAL across the 25 cell clusters identified by FlowSOM represented a compositional simplex in which each sample added up to a total of 100%. Extensive work has demonstrated that

spurious correlation or anti-correlation may occur when using this type of data. To address this issue, FlowSOM cluster percentage data was converted into real vector space for subsequent analysis using log-ratio transformation. The R package “compositions” (version 1.40-3) was used for this purpose.

Adjustment of high-dimensional cell count for CyTOF batch effect—To mitigate the contribution of batch effect to recovered cell cluster percentages, adjustment of log transformed data was performed via parametric empirical Bayesian framework using the function “ComBat” from R package “sva” (version 3.20.0).

Estimation of patient group number—The determination of three patient groups was arrived at by gap statistic calculation as well as Gaussian mixture modeling using log-ratio transformed, batch corrected data. These fundamentally different methods for patient cluster estimation both arrived at a consensus of $k = 3$. Gap statistic was calculated using the R package “NbClust” (version 3.0). Gaussian mixture modeling was done using the R package “mclust” (version 5.4.5).

Patient clustering based on high-dimensional cell count—An input of $k = 3$ clusters as determined above was used for patient clustering with the robust sparse k-means (RSKM) algorithm using the “RSKC” package in R (version 2.4.2). Log ratio cell cluster percentages with adjustment for CyTOF batch were used as described above. Resultant patient cluster designations are projected on principal component analysis (PCA) reduced space using FactoMineR (version 2.0). Ellipses represent the 75% confidence interval around the patient cluster centroid.

Modeling of cellular determinants of FEV1—To address the limitations of compositional data and small sample size, we employed elastic net modeling to assess relationship of cells to FEV1. This approach allows for consideration of multi-collinearity within the dataset and does not require correction for multiple testing. It has also been previously shown to successfully handle mass cytometry data without additional transformation (Aghaeepour et al., 2017). The R package “glmnet” (version 3.0-2) was used for this analysis, with 41-fold model cross validation. Model prediction was cross-validated using leave-one-out to estimate FEV1 based on an individual’s cell count. Model performance was evaluated via comparison of actual measured FEV1% predicted and values derived from EN modeling by Spearman’s rank correlation.

Mixed-effects association testing for single cell determination of patient group specific cells—To identify cell clusters variant between patient groups, we employed Mixed-effects Association testing for Single Cells (MASC) (“immunogenomics/masc,” version 0.1.0). Using a reverse association strategy, mixed-effects logistic regression is used to predict individual cell membership in previously defined populations. In our case, patient groupings as determined from RSKM clustering were used for comparison. The model used for odds-ratio calculation took biological sex, any steroid use (ICS or oral CS) as well as CyTOF batch into account. Patient group specificity was determined based off odds ratio comparison, with cells showing at least 2-fold greater likelihood to in one group compared to any other being designated as group specific.

Testing for variance in cell count across patient clusters—Following identification of cell populations of interest using automated cell clustering, we performed validation by manual gating of cell populations using the corresponding surface phenotypes as described in Results. Cell counts across participants were log ratio transformed as described above and subjected to Kruskal-Wallis testing.

Modeling patient groups using manually gated cell counts—To ensure patient clusters identified by FlowSOM cell counts could be reproduced using manual gating, we used sparse partial least-squares discriminant analysis (sPLS-DA) from the “mixOmics” package in R (version 6.10.8) to build a prediction model for group prediction using this data. Following model tuning, we used log-ratio transformed manual cell count data for prediction of patient group based on the 25 FlowSOM cell clusters that are detailed above. MixOmics offers leave-one-out or fold cross validation as options for these steps and models assembled using both protocols performed similarly in AUROC curve generation and model error rate measurement. For the AUROC demonstrated in Figure S2, 5-fold cross validation was used.

Network correlation mapping of BAL cells—Network map was assembled using Spearman rank-order correlation of unscaled cell count data following log ratio transformation of data. Nodes represent all cell clusters found to have at least one significant correlation as determined by a p value of < 0.05 and a rho of > 0.32 . Correction for multiple testing was performed using a FDR of < 0.05 . Nodes were colored for clinical disease specificity using comparison of odds ratios as described above. Edges of the network represent correlations that have passed through filtering as above. Network assembly was performed using the Fruchterman-Reingold force-directed layout algorithm.

Cytokine quantification—Scaled intensity of staining for the indicated cytokine is projected on tSNE dimension reduced space. Biaxial gating of cell events included in the FlowSOM clustering solution was performed for the indicated cytokines using Flowjo as described above. Relative distribution of T cell versus non-T cell cytokine is demonstrated by pie charts. Differences in distribution of staining intensity for each cytokine across patient groups was determined by pairwise Kolmogorov-Smirnov (KS) test on down-sampled data to allow for comparison of equal event numbers. The process of downsampling and comparison by KS test was repeated ten thousand times, with a meaningful difference defined as a test statistic (D) greater than 0.5 and p value less than $1e-10$ in greater than 99% of test iterations. Staining distributions that passed testing are plotted using the R “density” function with illustration of the 85th quantile of staining intensity and quantification of the proportion of cell events for each patient group that fall above that threshold in plot area (Figures 3D-3F). These cutoffs were then used to quantify cytokine positive cells per million across patient groups.

High dimensional fluorescence flow cytometry—Certain key aspects of mass cytometry were validated using conventional fluorescence flow cytometry. Specifically, the ability to identify innate immune cell populations (FlowSOM clusters 1 and 25) and to validate IL-4 expression in them using isotype control antibodies were examined. Due to the

number of analytes measured in the context of severe autofluorescence inherent with human BAL innate immune cells, particularly macrophages, high dimensional fluorescence flow cytometry was employed using an Aurora instrument (Cytek). One primary feature of the Aurora platform is the ability to collect data for unstained cells such that autofluorescence in all channels can then be treated as a parameter in the algorithm for minimizing fluorochrome spectral overlap. The process is analogous to conventional compensation, but is termed “unmixing” by Cytek to reflect proprietary methodology for accomplishing it.

Staining of cryopreserved BAL cells for fluorescence flow cytometry was carried out to the greatest extent possible to match the reagents and methodology for mass cytometry, which were described above. A panel consisting of 14 monoclonal antibodies constructed using mass cytometry-derived phenotypes was used with all but two being identical clones as those used for mass cytometry (see Key resources table). Importantly, the critical antibodies used to identify the cell populations of interest (FlowSOM clusters 1 & 25) and the anti-cytokine antibodies (anti-IL-4 and anti-IL-13) were matched clones. Fluorochromes were chosen for anti-IL-4, anti-IL-13, and key cell surface markers to identify/distinguish clusters 1 & 25, such as CD127, to coincide with Aurora channels with the least amount of BAL cell autofluorescence in order to minimize its impact. Additionally, relevant rat anti-human IgG1 isotype control antibodies from the same manufacturer (Biolegend) were used at the same optimized concentrations. Stimulation and staining was methodologically identical to the mass cytometry experiments with the following exceptions. Monensin (Biolegend) was used in addition to Brefeldin A, viability was determined using fixable viability dye eFluor 506 (eBioscience) instead of cisplatin, FACS buffer (2% FBS in PBS) was used instead of CyTOF buffer, and data acquisition was performed immediately upon final washings.

Controls for Aurora instrument set up and calibration included unstained cells serving as the autofluorescence control, cells stained with only fixable viability dye, and individual single color controls with each specific antibody bound to UltraComp eBeads™ compensation control beads. Spectral unmixing (analogous to conventional compensation) and data acquisition were carried out using SpectroFlo software (Cytek). Verification of successful unmixing was made by the SpectroFlo software and by examining single color-stained cells in pilot experiments as was practical. Importantly, autofluorescence was set in the SpectroFlo software as a parameter for unmixing in order to reliably detect specific signal for cell surface markers on innate (non-T) cells. Controls for cytokine expression were of the fluorescence minus one (FMO) variety but with isotype control anti-cytokine antibodies included. All relevant cell surface markers were included plus rat anti-human IgG1 labeled with each necessary specific fluorochrome. Available cryopreserved BAL cell samples consisted of 5×10^6 cells each such that minimal numbers of cells were used for unmixing controls and data for essentially all of the remaining cells was acquired in order to be able to visualize rare subpopulations such as those approximating FlowSOM clusters 1 & 25. Analysis was carried out by conventional methods using FlowJo software (TreeStar) with doublet discrimination (FSC-Area versus FSC Height) being employed throughout.

BAL transcriptional profiling—RNA isolated from BAL cells was used to generate cDNA libraries using the Illumina TruSeq RNA Library Prep Kit. The libraries were sequenced using the Illumina NextSeq500 System. Quality control for raw FASTQ files was

performed using FastQC and the low-quality reads and 3' adapters were trimmed with Trim Galore! (Babraham Bioinformatics). The RNA sequence aligner, STAR (Dobin et al., 2013), was used to align the trimmed reads to the reference human genome (hg19). Gene expression was subsequently quantified by counting the number of read fragments uniquely mapped to genes using featureCounts (Liao et al., 2014). Gene expression data was normalized using the rLog package from the DESeq2 suite (Love et al., 2014). Samples were controlled for batch effect using the COMBAT algorithm (Hu et al., 2004; Johnson et al., 2007). Duplicate gene identities were resolved using maximum values. Stochastic variation in gene expression was mitigated using a median cutoff of 5 across patient samples. Pairwise differential expression analysis was performed using t test with p value adjustment for a FDR < 0.05. “Volcano” plotting was accomplished based off log₂ fold expression change between indicated groups.

Gene set enrichment analysis across patient groupings—Gene set enrichment analysis (GSEA) was performed using the Java implementation of the application on gene expression data filtered to include probes with a median value greater than 1 across all patients. Comparison was made to the Hallmark (H), curated (C2) and GO sets (C5) libraries with illustrated gene sets meeting a p value of < 0.05 and an FDR < 0.25. Actual p value, ES and NES scores are illustrated in Figures 5B, 7G, and 7H.

WGCNA analysis of BAL—The R implementation of Weighted Gene Correlation Network Analysis (WGCNA version 1.69) (Langfelder and Horvath, 2008) was used for comparison of gene network expression to the abundance of immune cells identified by mass cytometry. BAL gene network construction was performed according to best practices as outlined by the package authors. Briefly, a soft thresholding of 4 for network topology construction was selected as the lowest power for which scale-free topology fit index exceeded a cutoff of 0.9. Dynamic tree cutting of the gene dendrogram was performed with a minimum module size of 30. This resulted in 10 gene modules which were subsequently collapsed to 9 based on module eigengene correlation value greater than 0.8. We tested a number of input parameters for method (e.g., hierarchical versus hybrid), PAM integration and deepSplit integer and found only minor variation (range 8-10 modules) in number following the compression of highly correlated modules as recommended by the WGCNA authors. Immune cell log ratios were used as comparative traits to these modules and significant interactions were defined by gene-trait correlations with p value less than 0.05. Gene network hubs were visualized using VisANT (Hu et al., 2004). Gene ontology analysis of gene modules was performed using the TopGO package (Alexa and Rahnenfuhrer, 2020).

ICLite gene module assembly—BAL gene module assembly was performed as outlined below using a machine learning process that sought to link cells to correlated genes through an iterative process designed to maximize interaction while accounting for quality of gene clustering results. This was accomplished by creating sparse binary matrices from the input gene expression data using varying thresholds for inclusion and transformation which are detailed below as well as in package documentation available at <https://github.com/camiolomj/ICLite/>. Optimal condition testing was performed using varied thresholds for gene inclusion, binary transformation and assumed number of underlying gene clusters.

The first step in the ICLite algorithm involves construction of a sparse binary matrix based on user input values for minimum connectivity and Spearman rho cutoff. After calculation of pairwise Spearman correlation across the entire transcriptome, values that fall over the input cutoff are converted to 1 on a binary scale while those under the threshold are reduced to 0. Any genes that do not demonstrate a total number of interactions above the user-defined set minimum connectivity threshold are then removed from the matrix. This process is repeated for all starting conditions, resulting in a varied set of matrices for downstream testing.

Clustering of the binary matrices is accomplished via the blockcluster package in R (Bhatia et al., 2017). Afterward, module strength of expression is calculated for genes in a module using geometric mean expression of scaled normalized values for each cohort participant and correlation between module scoring and immune cell log ratio is performed. Solutions are scored on the ratio of gene clusters to successful connections, defined as positive correlations between gene modules and cells with p value of less than 0.05. This data is used to inform optimal gene clustering solution selection, along with integrated completed likelihood (ICL), a measure of assessing how appropriate the assumed number of clusters was for the expression data (Biernacki et al., 2000). The solution that scores highest in a weighted system is then used for comparison to WGCNA results and downstream analysis. For this study, input rho cutoff values were tested from 0.4 to 0.65, minimum connectivity was varied between 25 and 200 and number of clusters ranged from 12 to 30. It is important to note that the blockclustering algorithm is sensitive to seed. For consistency between runs, a fixed value in the R environment should be set beforehand.

Cellular phenotype prediction modeling—In order to build a predication model for cellular phenotype using transcriptional data, we employed sparse partial least-squares discriminant analysis (sPLS-DA) within the “mixOmics” package in R (version 6.10.8). sPLS-DA has been previously shown to perform well on datasets with multiple related variables (Le Cao et al., 2011), making it well suited for our analysis. Training on the IMSA (experimental) cohort was performed using module strength of expression scores which were calculated for all 26 modules as detailed above. We then used these scores to train a sPLS-DA model to predict cell count-based patient cluster using gene expression data for the 39 participants that had both datasets available. MixOmics offers leave-one-out or fold cross validation as options for these steps and models assembled using both protocols performed similarly in AUROC curve generation and model error rate measurement. For the AUROC demonstrated in Figure 7A, 5-fold cross validation was used.

Cross validation using BAL gene expression from the SARP cohort—BAL transcriptional data from the Severe Asthma Research Project (SARP) I & II cohorts was prepared as previously described in the literature and subjected to patient scoring using mean scaled expression values. These patient scores were used for class prediction by the sPLS-DA model trained on IMSA as detailed above.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analysis was performed using the R computing environment (Version 3.5.3 or 3.6.1 contingent on package dependencies) unless otherwise noted. Statistical testing and methodology is described above within context specific methodologies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by NIH grants P01AI106684 (to A.R. and S.E.W.), R01HL113956 (to A.R. and P.R.), R01AI048927 (to A.R.), U10HL109152 (to S.E.W.), 1F32HL147415 (to M.J.C), R01AI140134, R01HL118612, R01ES020926, and P01ES022849 (to K.N.). We thank Shyamal Peddada for advice on biostatistical analyses, Renee Wunderley for human subject recruitment and obtaining informed consent, Marc Gauthier for bronchoscopic assistance, and Rachel Hannum for processing of the BAL samples. We thank Drs. Harinder Singh, Garry Nolan, and Amanda Poholek for critical reading of our manuscript.

REFERENCES

- Aghaeepour N, Ganio EA, Mcilwain D, Tsai AS, Tingle M, Van Gassen S, Gaudilliere DK, Baca Q, McNeil L, Okada R, et al. (2017). An immune clock of human pregnancy. *Sci. Immunol* 2, ean2946. [PubMed: 28864494]
- Alexa A, and Rahnenfuhrer J (2020). topGO: Enrichment analysis for gene ontology (Bioconductor). <https://bioconductor.org/packages/release/bioc/html/topGO.html>.
- Araki Y, Wang Z, Zang C, Wood WH 3rd, Schones D, Cui K, Roh TY, Lhotsky B, Wersto RP, Peng W, et al. (2009). Genome-wide analysis of histone methylation reveals chromatin state-based regulation of gene transcription and function of memory CD8⁺ T cells. *Immunity* 30, 912–925. [PubMed: 19523850]
- Ariotti S, Hogenbirk MA, Dijkgraaf FE, Visser LL, Hoekstra ME, Song JY, Jacobs H, Haanen JB, and Schumacher TN (2014). T cell memory. Skin-resident memory CD8⁺ T cells trigger a state of tissue-wide pathogen alert. *Science* 346, 101–105. [PubMed: 25278612]
- Aversa G, Chang CC, Carballido JM, Cocks BG, and de Vries JE (1997). Engagement of the signaling lymphocytic activation molecule (SLAM) on activated T cells results in IL-2-independent, cyclosporin A-sensitive T cell proliferation and IFN-gamma production. *J. Immunol* 158, 4036–4044. [PubMed: 9126961]
- Baines KJ, Simpson JL, Wood LG, Scott RJ, Fibbens NL, Powell H, Cowan DC, Taylor DR, Cowan JO, and Gibson PG (2014). Sputum gene expression signature of 6 biomarkers discriminates asthma inflammatory phenotypes. *J. Allergy Clin. Immunol* 133, 997–1007. [PubMed: 24582314]
- Barnett SB, and Nurmagambetov TA (2011). Costs of asthma in the United States: 2002–2007. *J. Allergy Clin. Immunol* 127, 145–152. [PubMed: 21211649]
- Basdeo SA, Cluxton D, Sulaimani J, Moran B, Canavan M, Orr C, Veale DJ, Fearon U, and Fletcher JM (2017). Ex-Th17 (Nonclassical Th1) Cells Are Functionally Distinct from Classical Th1 and Th17 Cells and Are Not Constrained by Regulatory T Cells. *J. Immunol* 198, 2249–2259. [PubMed: 28167631]
- Belkina AC, Ciccolella CO, Anno R, Halpert R, Spidlen J, and Snyder-Cappione JE (2019). Automated optimized parameters for T-distributed stochastic neighbor embedding improve visualization and analysis of large datasets. *Nat. Commun* 10, 5415. [PubMed: 31780669]
- Bhatia PS, Lovleff S, and Govaert G (2017). blocbuster: an R package for model based co-clustering. *J. Stat. Soft* 76 (9).
- Biernacki C, Celeux G, and Govaert G (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Trans. Pattern Anal. Mach. Intell* 22, 719–725.

- Bleecker ER, Wechsler ME, FitzGerald JM, Menzies-Gow A, Wu Y, Hirsch I, Goldman M, Newbold P, and Zangrilli JG (2018). Baseline patient factors impact on the clinical efficacy of benralizumab for severe asthma. *Eur. Respir. J* 52, 1800936. [PubMed: 30139780]
- Brightling CE, Chanez P, Leigh R, O'Byrne PM, Korn S, She D, May RD, Streicher K, Ranade K, and Piper E (2015). Efficacy and safety of tralokinumab in patients with severe uncontrolled asthma: a randomised, double-blind, placebo-controlled, phase 2b trial. *Lancet Respir. Med* 3, 692–701. [PubMed: 26231288]
- Brown EJ, Beal PA, Keith CT, Chen J, Shin TB, and Schreiber SL (1995). Control of p70 s6 kinase by kinase activity of FRAP in vivo. *Nature* 377, 441–46. [PubMed: 7566123]
- Brusselle G, Germinaro M, Weiss S, and Zangrilli J (2017). Reslizumab in patients with inadequately controlled late-onset asthma and elevated blood eosinophils. *Pulm. Pharmacol. Ther* 43, 39–45. [PubMed: 28159511]
- Chen F, Liu Z, Wu W, Rozo C, Bowdridge S, Millman A, Van Rooijen N, Urban JF Jr., Wynn TA, and Gause WC (2012). An essential role for TH2-type responses in limiting acute tissue damage during experimental helminth infection. *Nat. Med* 18, 260–266. [PubMed: 22245779]
- Chung KF, Wenzel SE, Brozek JL, Bush A, Castro M, Sterk PJ, Adcock IM, Bateman ED, Bel EH, Bleecker ER, et al. (2014). International ERS/ATS guidelines on definition, evaluation and treatment of severe asthma. *Eur. Respir. J* 43, 343–373. [PubMed: 24337046]
- Corren J, Parnes JR, Wang L, Mo M, Roseti SL, Griffiths JM, and van der Merwe R (2017). Tezepelumab in Adults with Uncontrolled Asthma. *N. Engl. J. Med* 377, 936–946. [PubMed: 28877011]
- De Meester I, Korom S, Van Damme J, and Scharpé S (1999). CD26, let it cut or cut it down. *Immunol. Today* 20, 367–375. [PubMed: 10431157]
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. [PubMed: 23104886]
- Fahy JV (2015). Type 2 inflammation in asthma—present in most, absent in many. *Nat. Rev. Immunol* 15, 57–65. [PubMed: 25534623]
- Flood-Page P, Swenson C, Faiferman I, Matthews J, Williams M, Brannick L, Robinson D, Wenzel S, Busse W, Hansel TT, and Barnes NC; International Mepolizumab Study Group (2007). A study to evaluate safety and efficacy of mepolizumab in patients with moderate persistent asthma. *Am. J. Respir. Crit. Care Med* 176, 1062–1071. [PubMed: 17872493]
- Gauthier M, Ray A, and Wenzel SE (2015). Evolving Concepts of Asthma. *Am. J. Respir. Crit. Care Med* 192, 660–668. [PubMed: 26161792]
- Gauthier M, Chakraborty K, Oriss TB, Raundhal M, Das S, Chen J, Huff R, Sinha A, Fajt M, Ray P, et al. (2017). Severe asthma in humans and mouse model suggests a CXCL10 signature underlies corticosteroid-resistant Th1 bias. *JCI Insight* 2, e94580.
- Haldar P, Pavord ID, Shaw DE, Berry MA, Thomas M, Brightling CE, Wardlaw AJ, and Green RH (2008). Cluster analysis and clinical asthma phenotypes. *Am. J. Respir. Crit. Care Med* 178, 218–224. [PubMed: 18480428]
- Hanania NA, Wenzel S, Rosén K, Hsieh HJ, Mosesova S, Choy DF, Lal P, Arron JR, Harris JM, and Busse W (2013). Exploring the effects of omalizumab in allergic asthma: an analysis of biomarkers in the EXTRA study. *Am. J. Respir. Crit. Care Med* 187, 804–811. [PubMed: 23471469]
- Hastie AT, Moore WC, Li H, Rector BM, Ortega VE, Pascual RM, Peters SP, Meyers DA, and Bleecker ER; National Heart, Lung, and Blood Institute's Severe Asthma Research Program (2013). Biomarker surrogates do not accurately predict sputum eosinophil and neutrophil percentages in asthmatic subjects. *J. Allergy Clin. Immunol* 132, 72–80. [PubMed: 23706399]
- Hekking PP, Loza MJ, Pavlidis S, De Meulder B, Lefaudeux D, Baribaud F, Auffray C, Wagener AH, Brinkman P, Lutter R, et al.; U-BIOPRED Study Group (2017). Transcriptomic gene signatures associated with persistent airflow limitation in patients with severe asthma. *Eur. Respir. J* 50, 1602298. [PubMed: 28954779]
- Hu Z, Mellor J, Wu J, and DeLisi C (2004). VisANT: an online visualization and analysis tool for biological interaction data. *BMC Bioinformatics* 5, 17. [PubMed: 15028117]

- Ibegbu CC, Xu YX, Fillos D, Radziejewicz H, Grakoui A, and Kourtis AP (2009). Differential expression of CD26 on virus-specific CD8(+) T cells during active, latent and resolved infection. *Immunology* 126, 346–353. [PubMed: 18657205]
- Jarjour NN, Erzurum SC, Bleecker ER, Calhoun WJ, Castro M, Comhair SA, Chung KF, Curran-Everett D, Dweik RA, Fain SB, et al.; NHLBI Severe Asthma Research Program (SARP) (2012). Severe asthma: lessons learned from the National Heart, Lung, and Blood Institute Severe Asthma Research Program. *Am. J. Respir. Crit. Care Med* 185, 356–362. [PubMed: 22095547]
- Jee YK, Gilmour J, Kelly A, Bowen H, Richards D, Soh C, Smith P, Hawrylowicz C, Cousins D, Lee T, and Lavender P (2005). Repression of interleukin-5 transcription by the glucocorticoid receptor targets GATA3 signaling and involves histone deacetylase recruitment. *J. Biol. Chem* 280, 23243–23250. [PubMed: 15826950]
- Johnson WE, Li C, and Rabinovic A (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8, 118–127. [PubMed: 16632515]
- Jozwik A, Habibi MS, Paras A, Zhu J, Guvenel A, Dhariwal J, Almond M, Wong EHC, Sykes A, Maybeno M, et al. (2015). RSV-specific airway resident memory CD8+ T cells and differential disease severity after experimental human infection. *Nat. Commun* 6, 10224. [PubMed: 26687547]
- Kambara K, Ohashi W, Tomita K, Takashina M, Fujisaka S, Hayashi R, Mori H, Tobe K, and Hattori Y (2015). In vivo depletion of CD206+ M2 macrophages exaggerates lung injury in endotoxemic mice. *Am. J. Pathol* 185, 162–171. [PubMed: 25447055]
- Kwakami T, and Galli SJ (2002). Regulation of mast-cell and basophil function and survival by IgE. *Nat. Rev. Immunol* 2, 773–786. [PubMed: 12360215]
- Kleinschek MA, Boniface K, Sadekova S, Grein J, Murphy EE, Turner SP, Raskin L, Desai B, Faubion WA, de Waal Malefyt R, et al. (2009). Circulating and gut-resident human Th17 cells express CD161 and promote intestinal inflammation. *J. Exp. Med* 206, 525–534. [PubMed: 19273624]
- Kohonen T (2013). Essentials of the self-organizing map. *Neural Netw.* 37, 52–65. [PubMed: 23067803]
- Kuo CS, Pavlidis S, Loza M, Baribaud F, Rowe A, Pandis I, Sousa A, Corfield J, Djukanovic R, Lutter R, et al.; U-BIOPRED Study Group (2017). T-helper cell type 2 (Th2) and non-Th2 molecular phenotypes of asthma using sputum transcriptomics in U-BIOPRED. *Eur. Respir. J* 49, 1602135. [PubMed: 28179442]
- Langfelder P, and Horvath S (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559. [PubMed: 19114008]
- Le Cao K-A, Boitard S, and Besse P (2011). Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics* 12, 10.1186/1471-2105-12-253.
- Leckie MJ, ten Brinke A, Khan J, Diamant Z, O'Connor BJ, Walls CM, Mathur AK, Cowley HC, Chung KF, Djukanovic R, et al. (2000). Effects of an interleukin-5 blocking monoclonal antibody on eosinophils, airway hyper-responsiveness, and the late asthmatic response. *Lancet* 356, 2144–2148. [PubMed: 11191542]
- Li Y, Wang W, Lv Z, Li Y, Chen Y, Huang K, Corrigan CJ, and Ying S (2018). Elevated Expression of IL-33 and TSLP in the Airways of Human Asthmatics In Vivo: A Potential Biomarker of Severe Refractory Disease. *J. Immunol.* 200, 2253–2262. [PubMed: 29453280]
- Liao Y, Smyth GK, and Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. [PubMed: 24227677]
- Liblau RS, Wong FS, Mars LT, and Santamaria P (2002). Autoreactive CD8 T cells in organ-specific autoimmunity: emerging targets for therapeutic intervention. *Immunity* 17, 1–6. [PubMed: 12150886]
- Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. [PubMed: 25516281]
- Mantovani A, Sica A, Sozzani S, Allavena P, Vecchi A, and Locati M (2004). The chemokine system in diverse forms of macrophage activation and polarization. *Trends Immunol.* 25, 677–686. [PubMed: 15530839]
- Menzies-Gow A, Colice G, Griffiths JM, Almqvist G, Ponnarambil S, Kaur P, Ruberto G, Bowen K, Hellqvist Å, Mo M, and Garcia Gil E (2020). NAVIGATOR: a phase 3 multicentre, randomized,

double-blind, placebo-controlled, parallel-group trial to evaluate the efficacy and safety of tezepelumab in adults and adolescents with severe, uncontrolled asthma. *Respir. Res* 21, 266. [PubMed: 33050934]

Modena BD, Tedrow JR, Milosevic J, Bleecker ER, Meyers DA, Wu W, Bar-Joseph Z, Erzurum SC, Gaston BM, Busse WW, et al. (2014). Gene expression in relation to exhaled nitric oxide identifies novel asthma phenotypes with unique biomolecular pathways. *Am. J. Respir. Crit. Care Med* 190, 1363–1372. [PubMed: 25338189]

Modena BD, Bleecker ER, Busse WW, Erzurum SC, Gaston BM, Jarjour NN, Meyers DA, Milosevic J, Tedrow JR, Wu W, et al. (2017). Gene Expression Correlated with Severe Asthma Characteristics Reveals Heterogeneous Mechanisms of Severe Disease. *Am. J. Respir. Crit. Care Med* 195, 1449–1463. [PubMed: 27984699]

Moore WC, Bleecker ER, Curran-Everett D, Erzurum SC, Ameredes BT, Bacharier L, Calhoun WJ, Castro M, Chung KF, Clark MP, et al.; National Heart, Lung, Blood Institute's Severe Asthma Research Program (2007). Characterization of the severe asthma phenotype by the National Heart, Lung, and Blood Institute's Severe Asthma Research Program. *J. Allergy Clin. Immunol* 119, 405–13. [PubMed: 17291857]

Muehling LM, Heymann PW, Wright PW, Eccles JD, Agrawal R, Carper HT, Murphy DD, Workman LJ, Word CR, Ratcliffe SJ, et al. (2020). Human T_H1 and T_H2 cells targeting rhinovirus and allergen coordinately promote allergic asthma. *J. Allergy Clin. Immunol* 146, 555–570. [PubMed: 32320734]

Nagakumar P, Puttur F, Gregory LG, Denney L, Fleming L, Bush A, Lloyd CM, and Saglani S (2019). Pulmonary type-2 innate lymphoid cells in paediatric severe asthma: phenotype and response to steroids. *Eur. Respir. J* 54, 1801809. [PubMed: 31164437]

Olenchock BA, Rathmell JC, and Vander Heiden MG (2017). Biochemical Underpinnings of Immune Cell Metabolic Phenotypes. *Immunity* 46, 703–713. [PubMed: 28514672]

Panettieri RA Jr., Sjöbring U, Péterffy A, Wessman P, Bowen K, Piper E, Colice G, and Brightling CE (2018). Tralokinumab for severe, uncontrolled asthma (STRATOS 1 and STRATOS 2): two randomised, double-blind, placebo-controlled, phase 3 clinical trials. *Lancet Respir. Med* 6, 511–525. [PubMed: 29792288]

Pavord ID, Beasley R, Agusti A, Anderson GP, Bel E, Brusselle G, Cullinan P, Custovic A, Ducharme FM, Fahy JV, et al. (2018). After asthma: redefining airways diseases. *Lancet* 391, 350–400. [PubMed: 28911920]

Peters MC, Nguyen ML, and Dunican EM (2016). Biomarkers of Airway Type-2 Inflammation and Integrating Complex Phenotypes to Endotypes in Asthma. *Curr. Allergy Asthma Rep* 16, 71. [PubMed: 27613654]

Rambold AS, and Pearce EL (2018). Mitochondrial Dynamics at the Interface of Immune Cell Metabolism and Function. *Trends Immunol.* 39, 6–18. [PubMed: 28923365]

Raundhal M, Morse C, Khare A, Oriss TB, Milosevic J, Trudeau J, Huff R, Pilewski J, Holguin F, Kolls J, et al. (2015). High IFN- γ and low SLPI mark severe asthma in mice and humans. *J. Clin. Invest* 125, 3037–3050. [PubMed: 26121748]

Ray A, Raundhal M, Oriss TB, Ray P, and Wenzel SE (2016). Current concepts of severe asthma. *J. Clin. Invest* 126, 2394–2403. [PubMed: 27367183]

Ray A, Camiolo M, Fitzpatrick A, Gauthier M, and Wenzel SE (2020). Are we meeting the promise of endotypes and precision medicine in asthma? *Physiol. Rev* 100, 983–1017. [PubMed: 31917651]

R Development Core Team (2014). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

Sergejeva S, Malmhäll C, Lötval J, and Pullerits T (2005). Increased number of CD34+ cells in nasal mucosa of allergic rhinitis patients: inhibition by a local corticosteroid. *Clin. Exp. Allergy* 35, 34–38. [PubMed: 15649263]

Snyder ME, Finlayson MO, Connors TJ, Dogra P, Senda T, Bush E, Carpenter D, Marboe C, Benvenuto L, Shah L, et al. (2019). Generation and persistence of human tissue-resident memory T cells in lung transplantation. *Sci. Immunol* 4, eaav5581. [PubMed: 30850393]

Spahn JD, Malka J, and Szeffler SJ (2016). Current application of exhaled nitric oxide in clinical practice. *J. Allergy Clin. Immunol* 138, 1296–1298. [PubMed: 27664377]

- Tur J, Pereira-Lopes S, Vico T, Marín EA, Muñoz JP, Hernández-Alvarez M, Cardona PJ, Zorzano A, Lloberas J, and Celada A (2020). Mitofusin 2 in Macrophages Links Mitochondrial ROS Production, Cytokine Release, Phagocytosis, Autophagy, and Bactericidal Activity. *Cell Rep.* 32, 108079. [PubMed: 32846136]
- Van Gassen S, Callebaut B, Van Helden MJ, Lambrecht BN, Demeester P, Dhaene T, and Saeys Y (2015). FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A* 87, 636–645. [PubMed: 25573116]
- Weathington N, O'Brien ME, Radder JE, Whisenant TC, Bleecker ER, Busse WW, Erzurum SC, Gaston B, Jarjour NN, Meyers DA, et al. (2019). BAL Cell Gene Expression in Severe Asthma Reveals Mechanisms of Severe Disease and Influences of Medications. *Am. J. Respir. Crit. Care Med* 200, 837–856. [PubMed: 31161938]
- Wechsler ME, Colice G, Griffiths JM, Almqvist G, Skärby T, Piechowiak T, Kaur P, Bowen K, Hellqvist Å, Mo M, and Garcia Gil E (2020). SOURCE: a phase 3, multicentre, randomized, double-blind, placebo-controlled, parallel group trial to evaluate the efficacy and safety of tezepelumab in reducing oral corticosteroid use in adults with oral corticosteroid dependent asthma. *Respir. Res* 21, 264. [PubMed: 33050928]
- Wenzel SE (2012). Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat. Med* 18, 716–725. [PubMed: 22561835]
- Wenzel S, Ford L, Pearlman D, Spector S, Sher L, Skobieranda F, Wang L, Kirkesseli S, Rocklin R, Bock B, et al. (2013). Dupilumab in persistent asthma with elevated eosinophil levels. *N. Engl. J. Med* 368, 2455–2466. [PubMed: 23688323]
- Wierenga EA, and Messer G (2000). Regulation of interleukin 4 gene transcription: alterations in atopic disease? *Am. J. Respir. Crit. Care Med* 162, S81–S85. [PubMed: 10988157]
- Wisniewski JA, Muehling LM, Eccles JD, Capaldo BJ, Agrawal R, Shirley DA, Patrie JT, Workman LJ, Schuyler AJ, Lawrence MG, et al. (2018). TH1 signatures are present in the lower airways of children with severe asthma, regardless of allergic status. *J. Allergy Clin. Immunol* 141, 2048–2060, e2013. [PubMed: 28939412]
- Woodruff PG, Modrek B, Choy DF, Jia G, Abbas AR, Ellwanger A, Koth LL, Arron JR, and Fahy JV (2009). T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am. J. Respir. Crit. Care Med* 180, 388–395. [PubMed: 19483109]
- Wu W, Bang S, Bleecker ER, Castro M, Denlinger L, Erzurum SC, Fahy JV, Fitzpatrick AM, Gaston BM, Hastie AT, et al. (2019). Multiview Cluster Analysis Identifies Variable Corticosteroid Response Phenotypes in Severe Asthma. *Am. J. Respir. Crit. Care Med* 199, 1358–1367. [PubMed: 30682261]
- Wynn TA, and Vannella KM (2016). Macrophages in Tissue Repair, Regeneration, and Fibrosis. *Immunity* 44, 450–462. [PubMed: 26982353]

Highlights

- Multi-omics and machine learning reveal two immune phenotypes of severe asthma
- Severe asthma stratified by FcεRI⁺IL-4⁺ innate immune and IFN-γ⁺ T cells
- ICLite identifies pathways potentially related to disease pathogenesis
- Gene expression modules fingerprint disease severity in independent cohorts

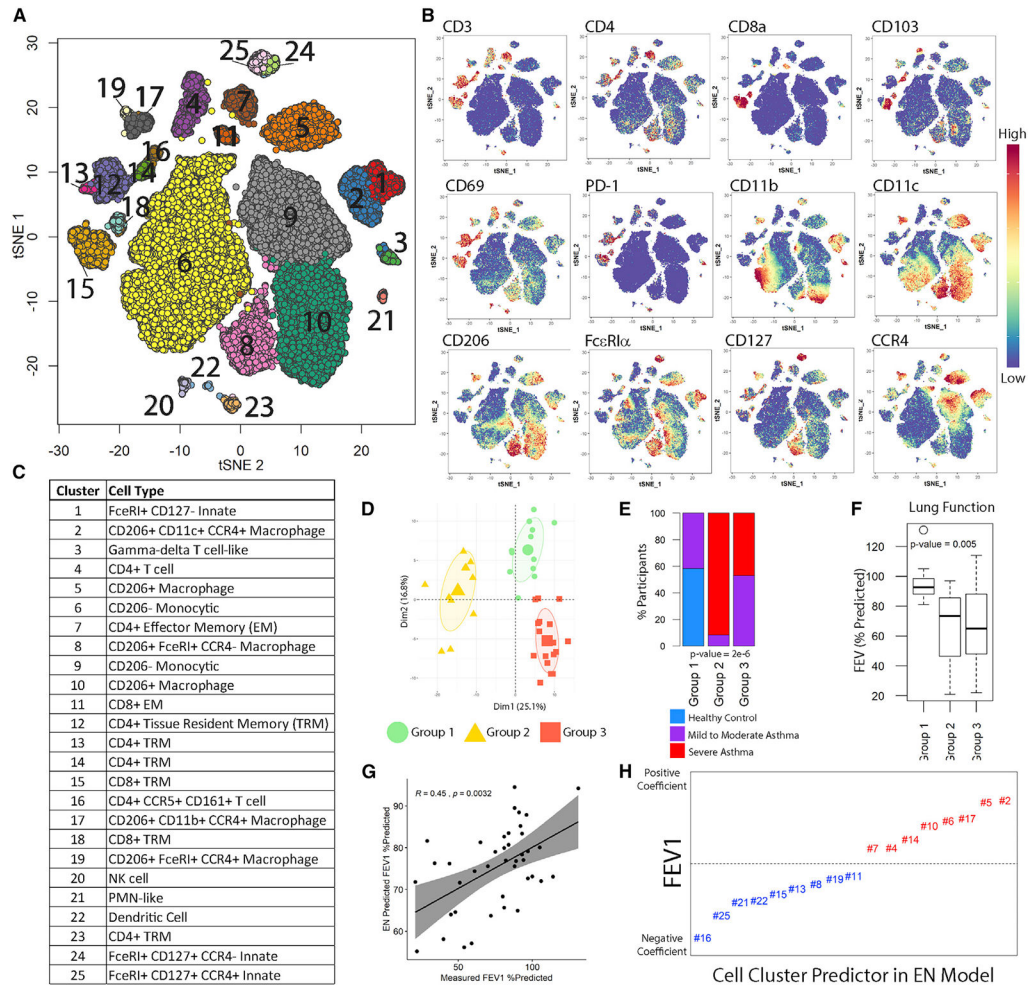


Figure 1. Unsupervised clustering of BAL cells identifies immune populations

(A) Results of FlowSOM projected on t-stochastic neighbor embedding (t-SNE) space.

(B) Relative staining intensity for indicated surface markers across t-SNE space.

(C) Summary of cell types identified by FlowSOM.

(D) Results of patient clustering projected on principal component analysis (PCA) plot. Patients are colored by group, with ellipses representing 90% of the confidence interval around group centroid.

(E) Distribution of clinical disease severities across groups is represented by stacked bar chart of proportion with p value determined by Pearson’s chi-square testing of raw values.

(F) Boxplot of FEV1 measured by spirometry across PGs with p value of variance calculated using Kruskal-Wallis testing.

(G) Elastic net (EN) predicted lung function based on high-dimensional cell count versus measured values of FEV1% predicted. Gray area indicates the 95% confidence bounds around a linear regression model comparing the two. Spearman’s rho and p value are indicated in plot area.

(H) Graphical representation of EN modeling of cellular determinants of lung function (FEV1). Coefficients are plotted in order of ascending value from left to right, with distance

from the hashed line indicating magnitude of contribution to the model. Blue coloration of cluster ID denotes a negative coefficient, and red indicates positive.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

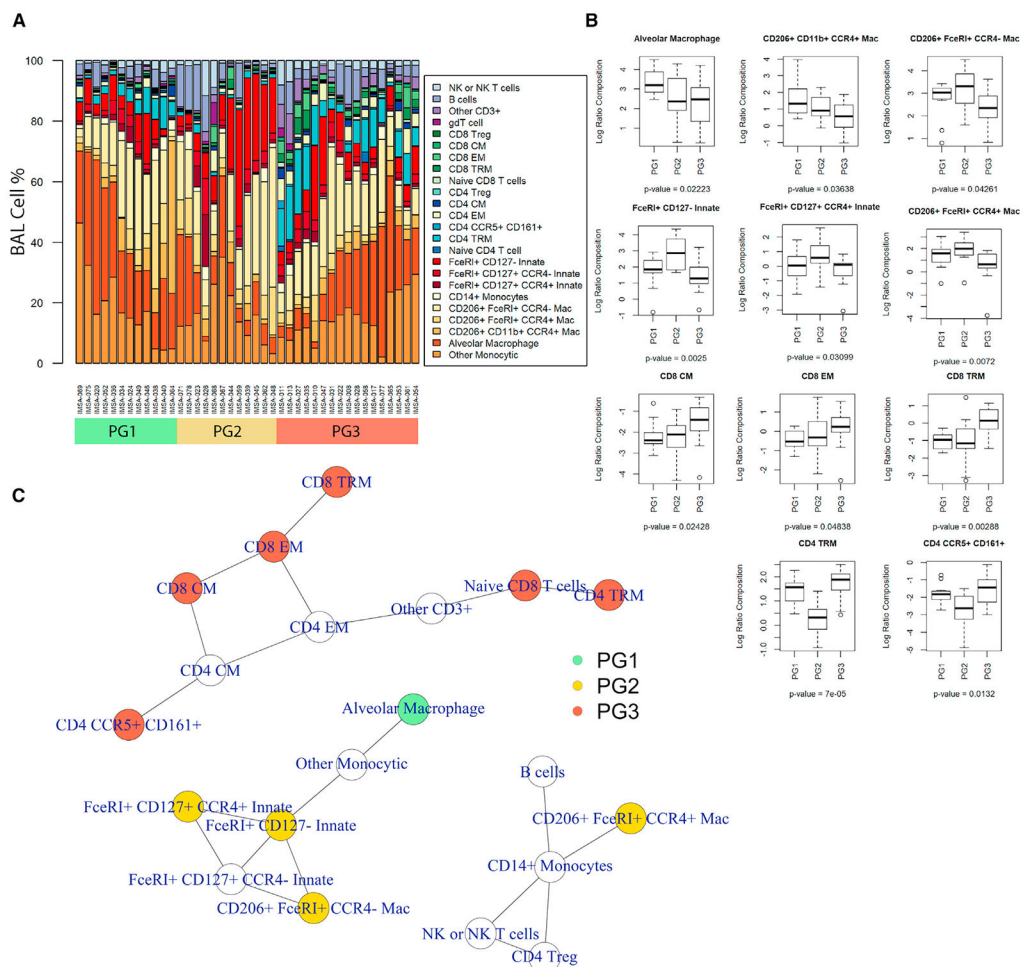


Figure 2. Patient groups are defined by divergent cell lineages

(A) Stacked bar plot of BAL immune cell composition of cohort participants from manual hierarchical gating, arranged by PG.

(B) Boxplot of cell lineages differentially presented across PGs with p value of variance calculated using Kruskal-Wallis testing. Bars represent median values, with bounds of boxes representing interquartile range (IQR) and whiskers representing 1.5× the upper or lower IQR.

(C) Network of correlated cell types across BAL of all patients in cohort. Nodes represent cells, and edges represent a Spearman correlation $\rho \geq 0.45$ with $p < 0.05$. Coloring of nodes for PG specificity is based on Dunn post hoc testing of cell lineages identified as variant by Kruskal-Wallis.

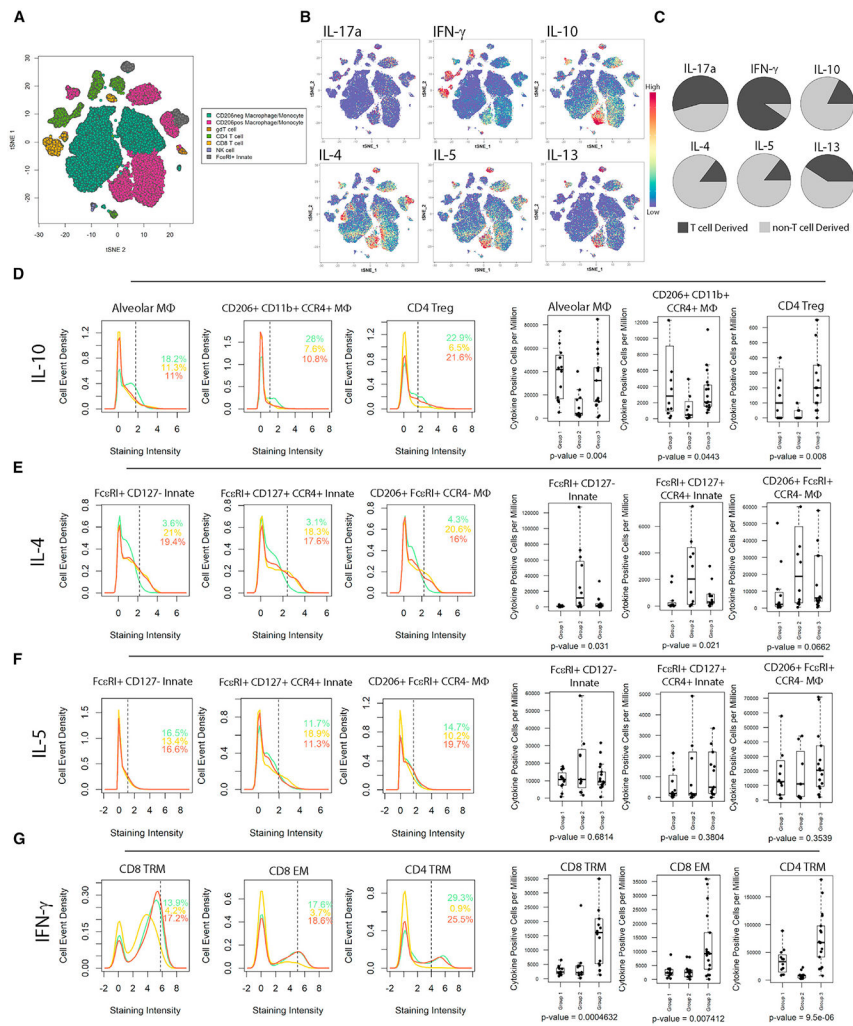


Figure 3. Cellular immune phenotype relates to distinct cytokine expression patterns
 (A) Graphical summary of cell lineages in t-SNE space.
 (B) Relative staining intensity of six cytokines across t-SNE-reduced space.
 (C) Pie charts demonstrating distribution of cytokine-positive cells in T-cell and non-T-cell compartments.
 (D) Density plots for IL-10 staining intensity of indicated cells demonstrate the relative distribution of events for a PG. Hashed line indicates the 85th quantile. Numeric values presented in plot area indicate the percentage of events in each PG falling above this cutoff. Cell types presented have passed Kolmogorov-Smirnov testing for variance in distribution with p value < 1e-10. Boxplots demonstrate cytokine-positive cells per million BAL cells per patient in respective groups with p value of variance calculated using Kruskal-Wallis testing. Bars represent median values, with bounds of boxes representing IQR and whiskers representing 1.5× the upper or lower IQR.
 (E) Plotting for IL-4 within indicated cell lineages as described in (D).
 (F) Plotting for IL-5 within indicated cell lineages as described in (D).
 (G) Plotting for IFN-γ within indicated cell lineages as described in (D).

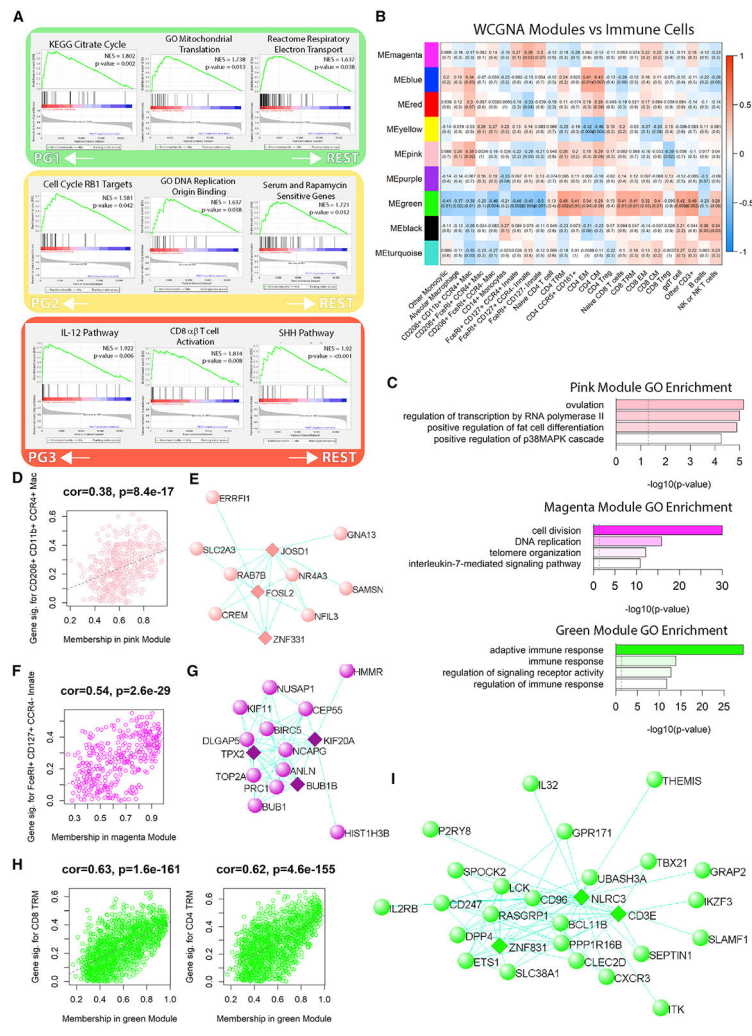


Figure 4. WGCNA of BAL links global transcriptional signatures to specific cell lineages
 (A) GSEA using Hallmark and curated gene sets. The p values, normalized enrichment score (NES), and enrichment scores (ES) are listed in the figure.
 (B) Heatmap of correlation between WGCNA module eigengenes and immune cell log ratios from mass cytometry evaluation of BAL. The p value and Spearman’s rho are indicated in the figure.
 (C) Barplot of $-\log_{10}(p \text{ values})$ for gene ontology (GO) term enrichment of WGCNA modules of interest based on correlation to cells from BAL.
 (D) Plotting pink module absolute gene significance (GS) for correlation to $CD206^+ CD11b^+ CCR4^+$ macrophages versus module membership (MM).
 (E) Visualization of pink module network hub genes (diamonds) with next closest network members (circles).
 (F) Plotting magenta module GS for correlation to $Fc\epsilon R1^+ CD127^+ CCR4^-$ innate cells versus MM.
 (G) Visualization of magenta module network hub genes (diamonds) with next closest network members (circles).
 (H) Plotting green module GS for correlation to CD8 or CD4 T cells versus MM.

(I) Visualization of green module network hub genes (diamonds) with next closest network members (circles).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

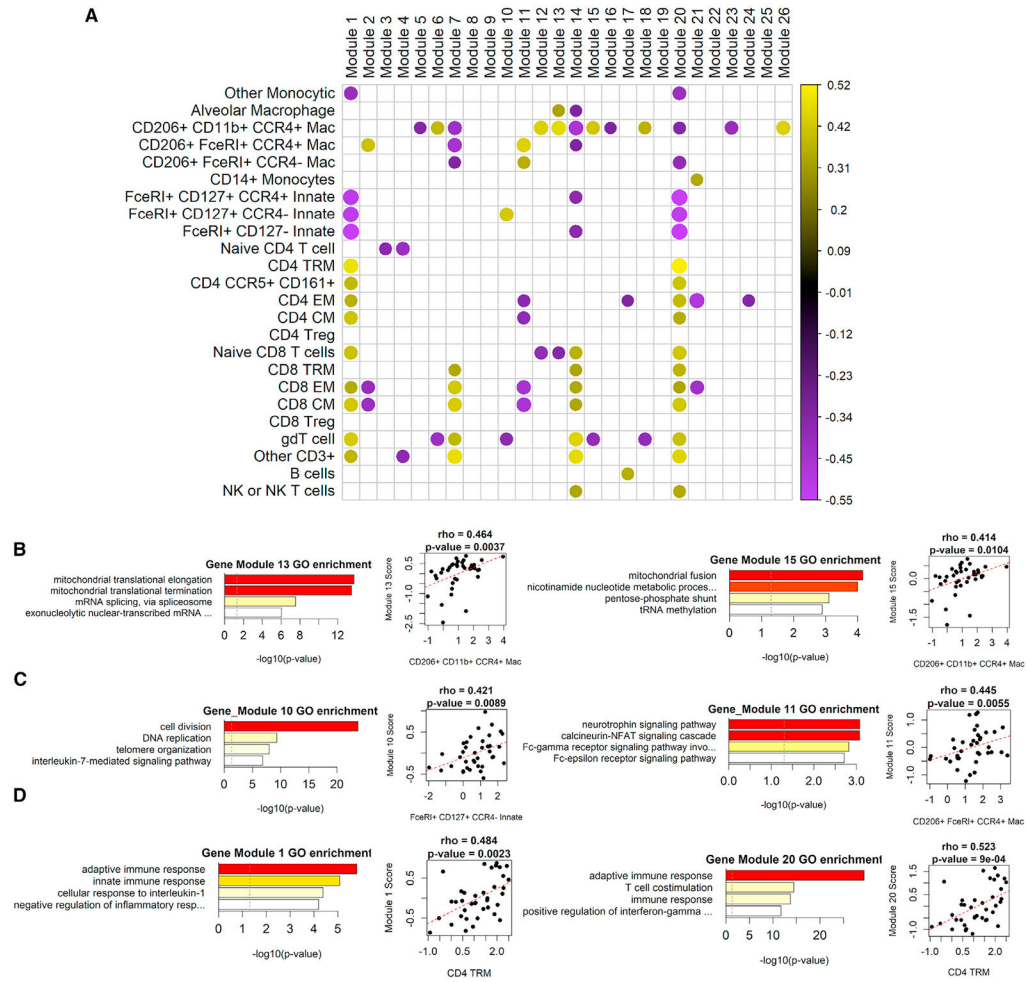


Figure 5. ICLite deconvolution of BAL sequencing improves GO term linkage

(A) Correlogram of gene module to cell cluster interactions using the ICLite algorithm.

Spearman correlation of module scoring with immune cell log ratio was used to construct a correlogram of BAL gene modules (x axis) versus cell lineages (y axis). Only associations with false discovery rate (FDR) corrected $p < 0.05$ are illustrated.

(B) Barplot of $-\log_{10}(p\text{ values})$ for GO term enrichment of ICLite modules 13 and 15, which correlate with CD206⁺ CD11b⁺ CCR4⁺ macrophages. Plotting of patient module score versus immune cell log ratio is adjacent, with Spearman's rho and p value of comparison as indicated in the figure. Red hashed line indicates linear regression model for data.

(C) Barplot of $-\log_{10}(p\text{ values})$ for GO term enrichment of ICLite modules 10 and 11, which correlate with FcεRI⁺ CD127⁺ CCR4⁻ innate cells and CD206⁺ FcεRI⁺ macrophages, respectively. Plotting of respective patient module score versus immune cell log ratio is adjacent, with Spearman's rho and p value of comparison as indicated in the figure. Red hashed line indicates linear regression model for data.

(D) Barplot of $-\log_{10}(p\text{ values})$ for GO term enrichment of ICLite modules 1 and 20, which correlate with CD4 and CD8 EMs, CMs, and TRMs. Plotting of respective patient module score versus CD4 TRM log ratio is adjacent, with Spearman's rho and p value of

comparison as indicated in the figure. Red hashed line indicates linear regression model for data.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

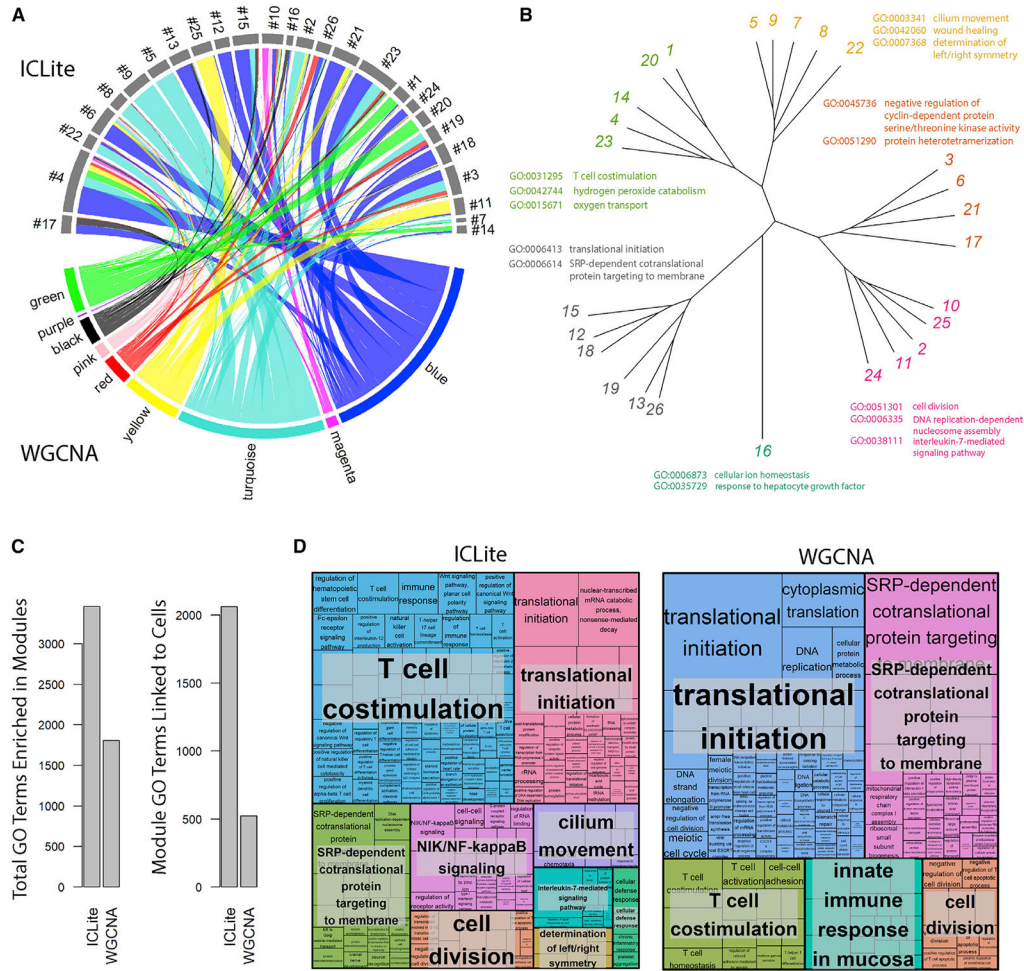


Figure 6. ICLite ascribes greater breadth of functional terms to BAL cells than WGCNA
 (A) Circos plot demonstrating gene module membership across trait-association technologies. WGCNA modules (bottom) are connected to ICLite modules based off commonality in membership, with chords of diagram colored by WGCNA module.
 (B) Phylogram of GO term semantic similarity for ICLite modules. Distance is independent of cell associations and based only on functional enrichment from transcriptional data. Module coloring is based off hierarchical clustering of semantic similarity.
 (C) Barplots quantifying the total number of GO terms enriched in modules or the number of GO terms effectively linked to cells by respective technology.
 (D) Tree plot of GO term semantic clustering results for all modules effectively linked to BAL cells using either ICLite or WGCNA. Size of box and text indicates $-\log_{10}(p \text{ value})$ of enrichment. Families are colored according to parent-child relationship in term clustering.

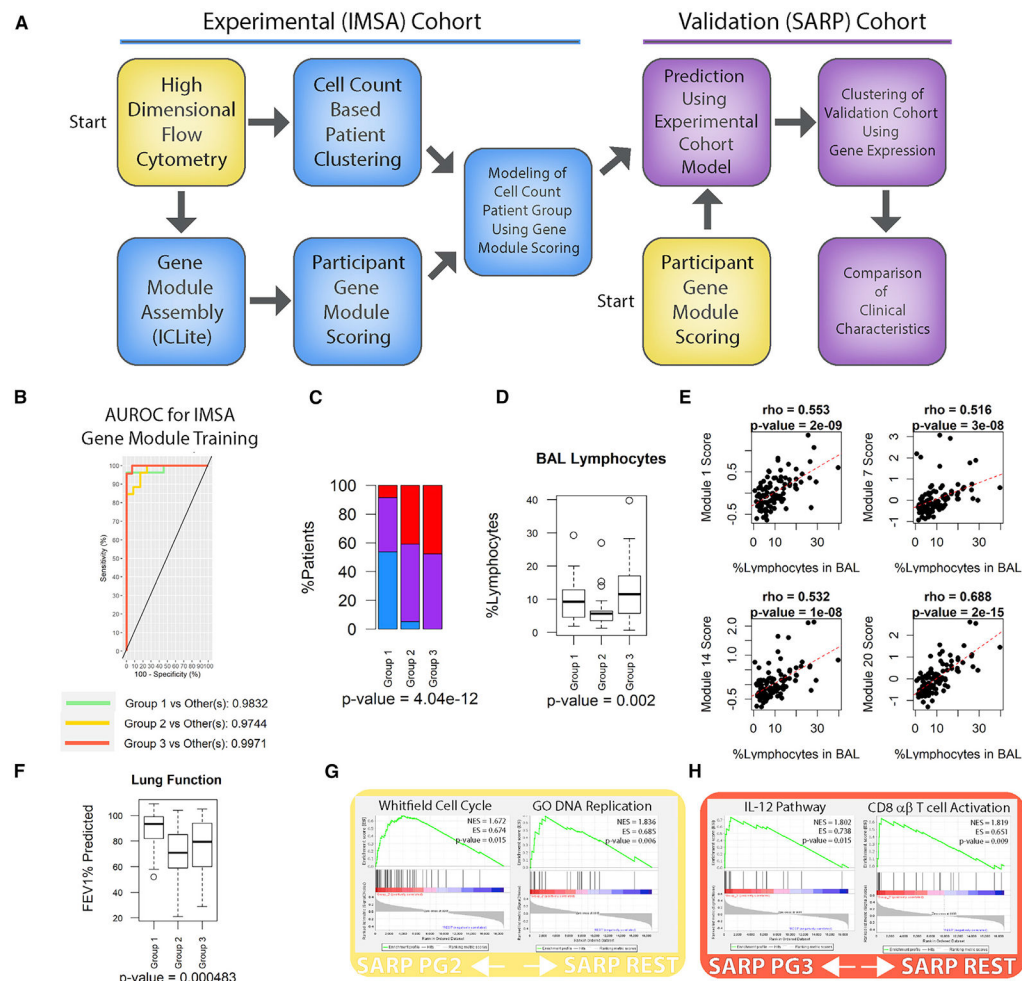


Figure 7. Gene modules predict cellular phenotype in SARP cohort

(A) Schematic for machine learning model of patient classification in external cohort using training based on cell count and transcriptional profile

(B) Receiver operating characteristics (ROC) curve of a sparse-partial least-squares discriminant analysis (sPLS-DA) model for cellular immune phenotype prediction using BAL gene module scoring within the IMSA (experimental) cohort. ROC curves were calculated as one class versus the others using 5-fold validation on the original training set. Reported area under the curve (AUC) values (right of plot) are based on comparison of predicted scores of one class versus the others using a two-component model. Wilcoxon test of predicted scores for one class versus the others met a significance threshold of $p < 0.05$ for all groups.

(C) Distribution of clinical disease severities across predicted SARP PGs is represented as proportion in stacked bar chart with overall p value determined by Pearson's chi-square testing of raw values.

(D) Boxplot of manual differential cell counts from BAL across predicted SARP PGs with p value of variance calculated using Kruskal-Wallis testing and represented on plot. Bars represent median values with bounds of boxes representing IQR and whiskers representing $1.5 \times$ the upper or lower IQR.

(E) Plotting of patient module scores versus percent lymphocyte of participant BAL from the SARP cohort, with Spearman's rho and p value of comparison as indicated in the figure. Red hashed line indicates linear regression model for data.

(F) Boxplot of measured spirometry across predicted SARP PGs with p value of variance calculated using Kruskal-Wallis testing and represented on plot.

(G) SARP BAL gene expression data was used for GSEA using Hallmark and curated gene sets. The p values, NES, and ES for gene sets enriched in predicted SARP PG2 are listed in the figure.

(H) SARP BAL gene expression data were used for GSEA. The p values, NES, and ES for gene sets enriched in predicted PG3 are listed in the figure.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Purified anti-human CD206 (MMR) Antibody	Biolegend	Cat# 321102; RRID:AB_571923
Purified anti-human CD163 Antibody	Biolegend	Cat# 333602; RRID:AB_1088991
Purified anti-human FcεR1α Antibody	Biolegend	Cat# 334602; RRID:AB_1227649
Anti-human CD45-89Y (clone HI30)	Fluidigm	Cat# 3089003; RRID:AB_2661851
Anti-Human CD196/CCR6 (G034E3)-141Pr	Fluidigm	Cat# 3141003A; RRID:AB_2687639
Anti-Human CD4 (RPA-T4)-145Nd	Fluidigm	Cat# 3145001; RRID:AB_2661789
Anti-Human CD8 (RPA-T8)-146Nd	Fluidigm	Cat# 3146001; RRID:AB_2687641
Anti-Human CD11c (Bu15)-147Sm	Fluidigm	Cat# 3147008; RRID:AB_2687850
Anti-CD278/ICOS (C398.4A)-148Nd	Fluidigm	Cat# 3148019B; RRID:AB_2756435
Anti-Human CD25 (2A3)-149Sm	Fluidigm	Cat# 3149010B; RRID:AB_2756416
Anti-Human CD103 (Ber-ACT8)-151Eu	Fluidigm	Cat# 3151011B; RRID:AB_2756418
Anti-Human TCRgd (11F2)-152Sm	Fluidigm	Cat# 3152008B; RRID:AB_2687643
Anti-Human CD194/CCR4 (L291H4)-153Eu	Fluidigm	Cat# 3158006A; RRID:AB_2687647
Anti-Human CD3 (UCHT1)-154Sm	Fluidigm	Cat# 3154003; RRID:AB_2687853
Anti-Human CD45RA (HI100)-155Gd	Fluidigm	Cat# 3155011B; RRID:AB_2810246
Anti-Human CD195/CCR5 (NP-6G4)-156Gd	Fluidigm	Cat# 3156015; RRID:AB_2661814
Anti-Human CD197/CCR7 (G043H7)-159Tb	Fluidigm	Cat# 3159003; RRID:AB_2714155
Anti-Human CD28 (CD28.2)-160Gd	Fluidigm	Cat# 3160003B
Anti-Human CD183/CXCR3 (G025H7)-163Dy	Fluidigm	Cat# 3156004B; RRID:AB_2687646
Anti-Human CD161 (HP-3G10)-164Dy	Fluidigm	Cat# 3164009B; RRID:AB_2687651
Anti-Human CD16 (3G8)-165Ho	Fluidigm	Cat# 3165001B; RRID:AB_2802109
Anti-Human CD27 (L128)-167Er	Fluidigm	Cat# 3167006B
Anti-Human CD127/IL-7Ra (A019D5)-168Er	Fluidigm	Cat# 3168017B; RRID:AB_2756425
Anti-Human HLA-DR (L243)-170Er	Fluidigm	Cat# 3170013B

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Anti-Human CD38 (HIT2)-172Yb	Fluidigm	Cat# 3172007B; RRID:AB_2756288
Anti-Human CD279/PD-1 (EH12.2H7)-174Yb	Fluidigm	Cat# 3175008; RRID:AB_2687629
Anti-Human CD14 (M5E2)-175Lu	Fluidigm	Cat# 3175015B
Anti-Human CD56 (NCAM16.2)-176Yb	Fluidigm	Cat# 3176008B
Anti-Human CD11b/Mac-1 (ICRF44)-209Bi	Fluidigm	Cat# 3209003 RRID:AB_2687654
Anti-Human CD69 (FN50)-144Nd	Fluidigm	Cat# 3144018 RRID:AB_2687849
Anti-Human IL-10 (JES3-9D7)-166Er	Fluidigm	Cat# 3166008B
Anti-Human IL-4 (MP4-25D2)-142Nd	Fluidigm	Cat# 3142002B
Anti-Human/Mouse IL-5 (TRFK5)-143Nd	Fluidigm	Cat# 3143003B
Anti-Human IL-22 (22URTI)-150Nd	Fluidigm	Cat# 3150007B
Anti-Human IFN-g (B27)-158Gd	Fluidigm	Cat# 3158017B
Anti-Human IL-17A (BL168)-161Dy	Fluidigm	Cat# 3161008B
Anti-Human FoxP3 (259D/C7)-162Dy	Fluidigm	Cat# 3162024A
Anti-Human IL-13 (JES10-5A2)-169Tm	Fluidigm	Cat# 3169016B
Human/Primate ST2/IL-33 R Antibody	R&D Systems	Cat# MAB523-100
CD19 Antibody, Qdot® 605 (Monoclonal, SJ25-C1)	ThermoFisher	Cat# Q10306; RRID:AB_10374734
Anti-Human IL-4 (MP4-25D2) Brilliant Violet 421	Biolegend	Cat# 500825 RRID: AB_10898316
Anti-Human IL-13 (JES10-A2) PE-Cy7	Biolegend	Cat# 501914 RRID:AB_2616746
Anti-Human FcεR1α (AER-37; CRA-1)Alexa Fluor 488	Biolegend	Cat# 334639 RRID:AB_2721289
Anti-Human Mannose Receptor CD206 (15-2) PE-Cy5	Biolegend	Cat# 321108 RRID:AB_571919
Anti-Human CD127 (A019D5) Brilliant Violet 605	Biolegend	Cat# 351334 RFID:AB_2562022
Anti-Human CD25 (BC96) PE-Dazzle 594	Biolegend	Cat# 302646 RRID:AB_2734260
Anti-Human CD194 (L291H4) PE	Biolegend	Cat# 359412 RRID:AB_2562433
Anti-Human CD183 (G025H7) Alexa Fluor 700	Biolegend	Cat# 353742 RRID:AB_2616920
Anti-Human HLA-DR (L243) Brilliant Violet 570	Biolegend	Cat# 307638 RRID:AB_2650882

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Anti-Human CD11c (3.9) Brilliant Violet 650	Biolegend	Cat# 301638 RRID:AB_2563797
Anti-Human CD11b (ICRF44) Brilliant Violet 711	Biolegend	Cat# 301344 RRID:AB_2563792
Anti-Human CD3 (UCHT1) APC-Cy7	Biolegend	Cat# 300426 RRID:AB_439780
Anti-Human CD45 (HI30) Brilliant Violet 785	Biolegend	Cat# 304048 RRID:AB_2563129
Rat Anti-Human IgG1 Isotype Control Brilliant Violet 421	Biolegend	Cat#401911
Rat Anti-Human IgG1 Isotype Control PE-Cy7	Biolegend	Cat#401908
Critical commercial assays		
Maxpar® X8 Antibody Labeling Kit, 173Yb-4 Rxn	Fluidigm	Cat# 201173A
Maxpar® X8 Antibody Labeling Kit, 171Yb-4 Rxn	Fluidigm	Cat# 201171A
Deposited data		
RNA-seq GEO accession	This paper	GEO: GSE136587
Flow cytometry FCS files	This paper	https://flowrepository.org/id/RvFr2BjGjcsd9iMYBm2whCXRWrdZpdwd99WkYjqmfxcFJsgTNG2f5mKHTu0dLcVH
Software and Algorithms		
FlowSOM	Cytobank, Inc.	RRID:SCR_016899
FlowJo	Tree Star, Inc.	RRID:SCR_008520
ICLite	This paper	https://github.com/camiolomj/ICLite/
Other		
Maxpar Cell Staining Buffer	Fluidigm	Cat# 201068
eBioscience IC Fixation Buffer	ThermoFisher	Cat# 00-8222-49
eBioscience Permeabilization Buffer	ThermoFisher	Cat#00-8333-56
Cell ID-intercalator-Ir 500uM	Fluidigm	Cat# 201192B
Cell-ID Cisplatin 100 uL	Fluidigm	Cat# 201064
Lanthanum (III) chloride heptahydrate 99.999% trace metals basis	Sigma	Cat# 203521
Fixable Viability Dye eFluor 506	eBioscience ThermoFisher	Cat # 65-0866-14
Human TruStain FcX Fc Receptor Blocking Solution	Biolegend	Cat# 422302 RRID:AB_2818989
Brefeldin A (1000X)	Biolegend	Cat# 420601
Monensin (1000X)	Biolegend	Cat# 420701
UltraComp eBeads Compensation Beads	ThermoFisher	Cat# 01-2222-41

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Invitrogen UltraComp eBeads Compensation Beads	ThermoFisher	Cat# 01-3333-42

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript