Supplementary

Cellular deconvolution

In this study, we used AutoGeneS[1] that automatically extracts informative genes and outperforms other methods for analyzing bulk RNA samples with closely correlated cell types and noisy single-cell reference profiles. The number of informative genes was manually set to 300 and 400 genes and we selected the one (n=400) with the most stable results across cohorts. Similarly, the proportions of major cell types--goblet, secretory, and ciliated cells--were consistent using different cell type resolutions.

scRNA-Seq data was used from bronchial biopsies[2]. Due to highly similar gene expression profiles, the scRNA-Seq signatures from the club and the 2 goblet cell clusters were combined to generate a uniform scRNA-Seq signature of secretory cells. The merged scRNA-seq count data was normalized to count per million (CPM) and highly variable (HV) genes (n=5,000) were selected. We used the method implemented in single-cell analysis in Python (SCANPY)[3] for selecting HV genesin which genes are binned by their mean expression and those with the highest variance-to-mean ratio are selected as HV genes in each bin. We then performed AutoGeneS[1] to filter 400 informative genes from the highly variable ones that differentiated the cell types. The informative genes minimized correlation and maximized distance between the clusters in the single-cell reference data. For visualisation, single-cell neighbourhood graph (kNN-graph) was computed on the first 50 principal components using 30 neighbours and low-dimensional uniform manifold approximation and projection

(UMAP) embedding was used. Bulk deconvolution was then conducted on all bulk samples using support vector regression (SVR) method[4] for samples measured by both RNA-Seq and microarray.

Interaction analysis

The interaction analysis was conducted on TAC signatures and Mast cell signatures using a linear mix effect (LME) model investigating the interaction between time and treatment with patient ID as the random factor.

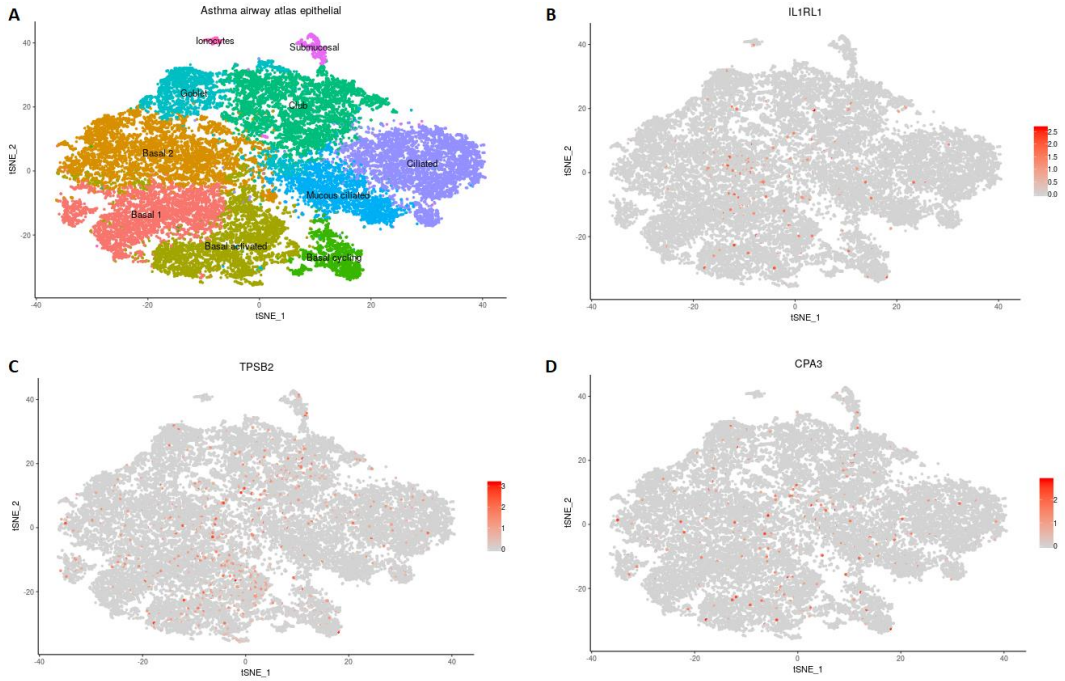

Figure S1. TSNE plots made from epithelial only cell/subtypes
TSNE plots for ICS sensitive TAC1 genes (IL1RL1, TPSB2 and CPA3), obtained from of single cell seq data obtained from asthmatic (n=4) and healthy controls (n=4) (A-D).

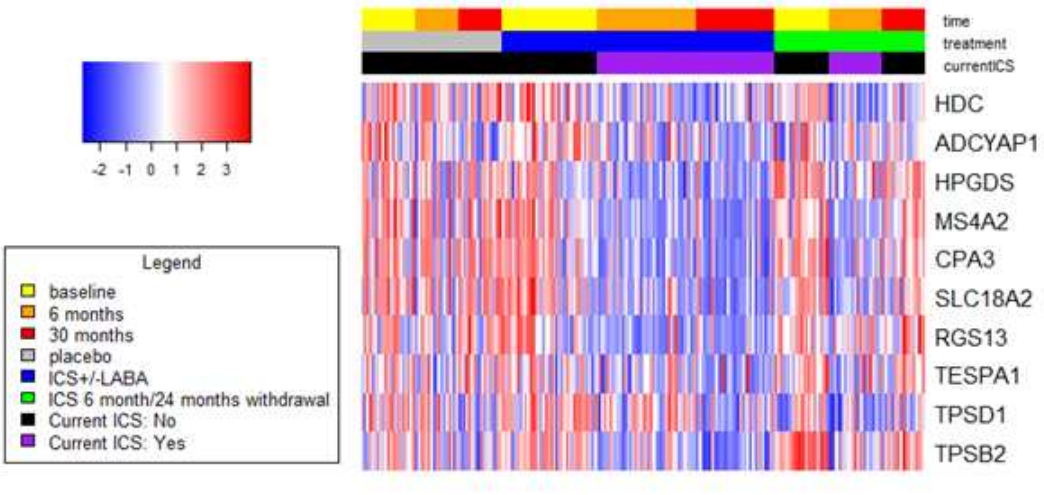

Figure S2 Individual expression of genes from the mast cells signature. Heatmap of the genes in the mast cell signature.

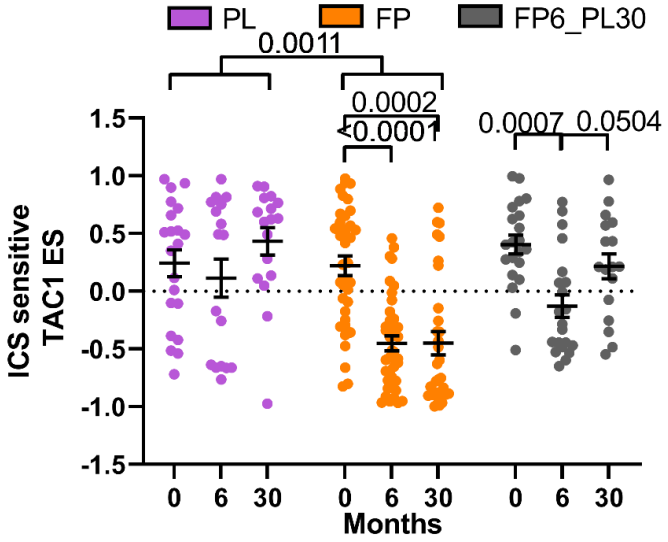

Figure S3. Relationship between corticosteroid sensitive signature and the influence of inhaled corticosteroid (ICS) treatment Mean ± SEM is presented in the figure. The interaction analysis was conducted on TAC signature using a LME model investigating the interaction between time and treatment with patient ID as the random factor. For two-way anovas a Benjamin Hochberg adjusted pvalue<0.05 was considered significant.

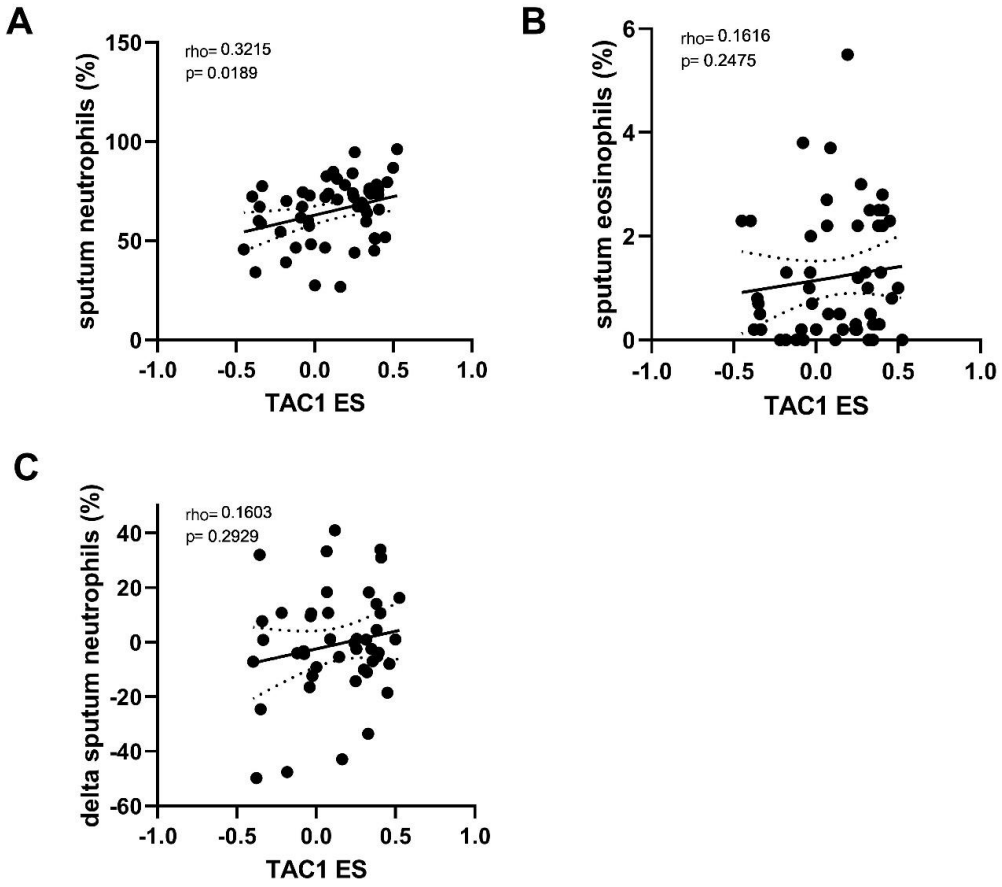

Figure S4. Percentage of sputum cell correlations with TAC1. Correlation with TAC1 and baseline A) sputum neutrophils and B) sputum eosinophils and C) delta between 30 months ICS+/-LABA compared to baseline. Spearman correlations were conducted.

Table S1. Gene set variation analysis (GSVA) datasets

Dataset	Reference
TAC1	[5]
TAC2	[5]
TAC3	[5]

IL13 IVS array	[6]
Inflammasome. Gibson	[7]
OXPHOS	[8]
Lung.brushings,cigarette.irreversible.UP.HS	[9]

Table S2. GLUCOLD patient characteristics

	Fluticasone±salmeterol for 30 months			Placebo for 30 months			Fluticasone for 6 months followed by placebo between 6 and 30 months		
	Baseline	**6 months**	**30 months**	**Baseline**	**6 months**	**30 months**	**Baseline**	**6 months**	**30 months**
Number of included patients	45			23			21		
Number of biopsies available at each time point	37	39	31	21	17	17	21	21	17
Male/female, n	41/4			19/2			19/2		
Age, years	62.4±7.2			60.2±7.8			63.1±7.4		
BMI	25.5±3.7			24.2±3.9			25.4±3.6		
Current smokers, n (%)	22 (59)	20 (51)	14 (45)	14 (67)	10 (59)	8 (47)	10 (48)	9 (43)	8 (47)
RIN score	3.3±1.5	3.5±1.3	4.8±1.5**	3.5±1.3	3.9±1.6	5.2±1.8**	3.3±1.7	3.7±1.7	3.7±1.5
FEV_1, %predicted	62.6±9.0	63.6±10.7	64.2±12.3	61.3±8.80	62.3±9.20	57.0±8.3	64.7±8.62	64.9±9.0	64.2±12.5

Table S3. Clinical characteristics separated by TAC grouping

	TAC1	TAC2	TAC3
n	26	18	14
gender male n(%)	20(76.9)	16(88.9)	12(85.7)
current smokers n(%)	18(69.2)*	4(22.2)*	14(100)
age	61(9)*	62(7)*	55(6)
FEV1pp	62.5(8.4)	61.2(10.3)	65.3(8.1)
RV/TLCpp	127.3(20.7)	120.5(14.7)	121.9(16.8)
log sputum eos	0.52(0.45)	0.49(0.4)	0.25(0.27)
log sputum neutro	2.1(0.4)*	2.28(0.4)*	1.55(0.45)
log sputum macro	1.6(0.4)*	1.73(0.36)	1.41(0.37)
log sputum lymphocytes	0.7(0.3)*	0.86(0.39)*	0.37(0.26)

*=p<0.05 compared to TAC3. t-test conducted

Table S4. Interaction analysis of Time and treatment for TAC signatures and Mast cell signatures

Test	Beta	Standard Error	P value
TAC1 signature	-0.00550736	0.00328515	0.0970
TAC2 signature	-0.00207232	0.00379294	0.5859
TAC3 signature	-0.00308084	0.00480443	0.5226
ICS sensitive TAC1 signature	-0.02409835	0.00719126	0.0011

Mast cell proportions (deconvolution)	0.000037159	0.0000315923	0.2420
Mast Cell signature (GSVA)	0.01518509	0.00599422	0.0127
Mast cell counts (mm2)	-0.0184948	0.00462402	0.0001

Table S5. TAC1 differential gene expression analysis comparing bronchial expression profiles of placebo and treatment arms of the GLUCOLD study at baseline, 6 months, and 30 months.

	Baseline			6 months			30 months		
	logFC	P Value	FDR	logFC	P Value	FDR	logFC	P Value	FDR
CPA3	-0.205	4.159E-01	9.221E-01	-1.142	2.871E-05	4.593E-04	-1.702	3.421E-09	5.473E-08
IL1RL1	-0.036	7.815E-01	9.221E-01	-0.384	5.495E-03	4.396E-02	-0.749	3.050E-07	2.440E-06
TARP	-0.330	1.322E-01	5.286E-01	-0.329	1.558E-01	2.769E-01	-0.958	8.398E-05	4.479E-04
TPSB2	0.032	8.073E-01	9.221E-01	-0.293	3.914E-02	1.566E-01	-0.391	7.909E-03	3.164E-02
ATP2A3	0.071	6.903E-01	9.221E-01	0.029	8.763E-01	8.763E-01	0.453	2.040E-02	6.529E-02
LGALS12	0.158	6.693E-02	5.246E-01	0.201	2.871E-02	1.531E-01	0.185	5.122E-02	1.366E-01
PRSS33	-0.017	8.142E-01	9.221E-01	0.135	8.707E-02	2.322E-01	0.128	1.170E-01	2.675E-01
OLIG2	-0.015	8.645E-01	9.221E-01	0.071	4.445E-01	5.926E-01	0.134	1.624E-01	3.248E-01
CCR3	0.071	3.443E-01	9.221E-01	0.136	8.581E-02	2.322E-01	0.070	3.935E-01	6.995E-01
ALOX15	0.428	2.522E-02	4.035E-01	-0.060	7.650E-01	8.160E-01	-0.092	6.587E-01	8.782E-01
FAM101B	-0.056	4.705E-01	9.221E-01	0.058	4.846E-01	5.965E-01	0.042	6.185E-01	8.782E-01
CD24	0.016	9.292E-01	9.292E-01	-0.156	4.209E-01	5.926E-01	-0.117	5.582E-01	8.782E-01
CLC	-0.040	6.655E-01	9.221E-01	0.153	1.158E-01	2.647E-01	0.025	8.065E-01	9.926E-01
SOCS2	-0.130	9.837E-02	5.246E-01	0.053	5.243E-01	5.992E-01	-0.001	9.951E-01	9.951E-01
HRH4	0.029	7.453E-01	9.221E-01	0.133	1.544E-01	2.769E-01	0.004	9.647E-01	9.951E-01
VSTM1	0.012	8.530E-01	9.221E-01	0.093	1.934E-01	3.095E-01	0.008	9.160E-01	9.951E-01

Abbreviation FC= Fold Change, FDR=False Discovery Rate

References

1. Aliee H, Theis F. AutoGeneS: Automatic gene selection using multi-objective optimization for RNA-seq deconvolution. *bioRxiv* 2020.
2. Braga FAV, Kar G, Berg M, Carpaij OA, Polanski K, Simon LM, Brouwer S, Gomes T, Hesse L, Jiang J. A cellular census of human lungs identifies novel cell states in health and in asthma. *Nature medicine* 2019: 25(7): 1153-1163.
3. Li J, Yu C, Ma L, Wang J, Guo G. Comparison of Scanpy-based algorithms to remove the batch effect from single-cell RNA-seq data. *Cell Regeneration* 2020: 9(1): 1-8.
4. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, Khodadoust MS, Esfahani MS, Luca BA, Steiner D. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nature biotechnology* 2019: 37(7): 773-782.
5. Kuo C-HS, Pavlidis S, Loza M, Baribaud F, Rowe A, Pandis I, Sousa A, Corfield J, Djukanovic R, Lutter R, Sterk PJ, Auffray C, Guo Y, Adcock IM, Chung KF. T-helper cell type 2 (Th2) and non-Th2 molecular phenotypes of asthma using sputum transcriptomics in U-BIOPRED. *European Respiratory Journal* 2017: 49(2): 1602135.
6. Alevy YG, Patel AC, Romero AG, Patel DA, Tucker J, Roswit WT, Miller CA, Heier RF, Byers DE, Brett TJJTJoci. IL-13–induced airway mucus production is attenuated by MAPK13 inhibition. 2012: 122(12): 4555-4568.
7. Simpson JL, Phipps S, Baines KJ, Oreo KM, Gunawardhana L, Gibson PGJERJ. Elevated expression of the NLRP3 inflammasome in neutrophilic asthma. 2014: 43(4): 1067-1076.
8. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JPJB. Molecular signatures database (MSigDB) 3.0. 2011: 27(12): 1739-1740.
9. Spira A, Beane J, Shah V, Liu G, Schembri F, Yang X, Palma J, Brody JS. Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proceedings of the National Academy of Sciences* 2004: 101(27): 10143-10148.