Supplemental Material for Parker et al.

Supplementary Methods

Comments on data aggregation and covariates

After the initial identification of studies, we noted that there was overlap in study populations examined in some studies, thus we chose one study from each set. For one group of such studies [1–3], we chose to use the largest study from this group because it contained most or all of the data from prior studies. For the second set of such studies[4–6], based on subjects in the Severe Asthma Research Program (SARP), we used the study for which airway epithelial data generation was best described[6].

Also, we did not include in vitro studies of cultured epithelia from cases vs. controls because preliminary analyses indicated substantially different gene expression patterns between in vivo and in vitro samples (not shown).

Case-control status

For the study by Giovanini-Chami et al. [7], we combined both types of asthma cases (controlled and uncontrolled) together. Likewise, for the study Modena et al. (GSE63142) [6], we combined data from severe asthmatics and mild-to-moderate asthmatics in the analysis of differential gene expression.

Smoking

For the majority of studies, subjects were non-smokers. However, for two studies, GSE85568 [8] and GSE89809 (unpublished), some study subjects smoked and hence smoking status was included in differential gene expression analyses for these studies.

Race

Not all studies reported data on the race/ethnicity of their study subjects. Rather than exclude these studies, we made the assumption that cases and controls from each study were reasonably well matched with respect to race/ethnicity. For GSE89809, we excluded one subject self identified as African because this was the only African subject in the study and analysis including this subject would not have allowed us to make meaningful inferences about the effect of race/ethnicity. For GSE63142, three subjects lacked race data and hence were excluded from the dataset.

Medication use

Most studies did not provide data on medication use in GEO. Therefore the effect of medications was not examined in the meta-analysis. Instead, we compared our list of DEGs to previous studies which examined the effect ICS on airway epithelia transcriptomes [9, 10]. From these two studies, we manually compiled a list of genes that were reported to be differentially expressed in either study, totaling 111 genes. We then examined overlap between these 111 genes and the genes from our meta-analysis.

Array probe annotation

In most cases, we used the array probe-gene annotation provided in GEO corresponding to the

array platform used. However, for the unpublished study by Yang et al. (NJH), we used Biomart to convert probes to genes using the "Agilent_sureprint_g3_ge_8x60k_v2" annotation. If multiple probes mapped to one gene, the probe with highest standard deviation was used. For the RNA-seq dataset (GSE85568), we used the Ensembl gene annotations provided in the GEO entry. These data are summarized in Supplementary Table 2.

Outlier identification

To identify outlier samples within each study, we used Bolstad's relative log expression (RLE) method [11]. We defined outliers as samples with an interquartile range (IQR) > mean(IQR)+2.5*SD(IQR). This resulted in the exclusion of 17 samples from six studies, as shown in Table 1 in the main text.

Enrichment analyses

We used Gene Set Variation Analysis (GSVA) [12] to transform gene expression data from each study to pathway level scores. Briefly, GSVA works by estimating a gene set score in an unsupervised way by calculating an enrichment statistic for each gene set in each sample. In this way the gene by sample matrix is transformed to a matrix of gene set by sample. The gene set by sample matrix was then used for association testing and meta-analysis using the same methods as described above for individual genes. Gene sets for analysis were collected from Enrichr (http://amp.pharm.mssm.edu/Enrichr/). We tested gene sets representing biological pathways (BioCarta_2016, KEGG_2016, NCI-Nature_2016, Panther_2016, Reactome_2016, Human_Gene_Atlas, WikiPathways_2016), sets representing regulatory targets (ChEA_2016, ENCODE_Histone_Modifications_2015, ENCODE_TF_ChIP-seq_2015,

ENCODE_and_ChEA_Consensus_TFs_from_ChIP-X, Epigenomics_Roadmap_HM_ChIP-seq, TargetScan_microRNA), protein-protein interactions (HumanCyc_2016, BioPlex_2017 Transcription_Factor_PPIs, PPI_Hub_proteins, huMAP) and transcription factor targets (Genome_Browser_PWMs, TRANSFAC_and_JASPAR_PWMs).

Supplementary results on overlap of asthma DEGs with other allergic disease states

We examined whether the top DEGs from the meta-analysis were also associated with other allergic disorders, including allergic rhinitis (AR), atopic dermatitis (AD), and eosinophilic esophagitis (EoE). We identified DEGs for each of these studies from published findings[13–15] and conducted hyper-geometric tests of enrichment. As expected, the overlap between gene sets identified in our meta-analysis with AR was quite strong statistically (hyper-geometric p-value = 6.52×10^{-5}), with nine significant genes shared in common, most prominently CST1. Notably, all nine genes in common were modulated in the same direction (Supplementary Table 9). Overlap with AD, while statistically significant, was not impressive in that only two genes were shared in common, one of which was discordant in terms of direction of expression. We did, however, find strong enrichment of asthma DEGs in EoE ($p = 4.8 \times 10^{-10}$), though the directionality of gene expression changes was not completely concordant (Supplementary Table 9). Three genes were shared in common between asthma, AR and EoE, namely CLC, CDH26, and HDC.

References

1. Woodruff PG, Modrek B, Choy DF, Jia G, Abbas AR, Ellwanger A, Arron JR, Koth LL, Fahy J V. T-helper type 2-driven inflammation defines major subphenotypes of asthma.

Am. J. Respir. Crit. Care Med. 2009; 180: 388–395.

2. Choy DF, Modrek B, Abbas AR, Kummerfeld S, Clark HF, Wu LC, Fedorowicz G, Modrusan Z, Fahy J V., Woodruff PG, Arron JR. Gene Expression Patterns of Th2 Inflammation and Intercellular Communication in Asthmatic Airways. *J. Immunol.* 2011; 186: 1861–1869.
3. Christenson SA, Steiling K, Van Den Berge M, Hijazi K, Hiemstra PS, Postma DS, Lenburg ME, Spira A, Woodruff PG. Asthma–COPD Overlap. Clinical Relevance of Genomic Signatures of Type 2 Inflammation in Chronic Obstructive Pulmonary Disease. *Am. J. Respir. Crit. Care Med.* American Thoracic Society - AJRCCM; 2015; 191: 758–766.
4. Li X, Hawkins GA, Moore WC, Hastie AT, Ampleford EJ, Milosevic J, Li H, Busse WW, Erzurum SC, Kaminski N, Wenzel SE, Bleecker ER, Meyers DA. Expression of asthma susceptibility genes in bronchial epithelial cells and bronchial alveolar lavage in the Severe Asthma Research Program (SARP) cohort. *J. Asthma* Taylor & Francis; 2016; 53: 775–782.
5. Voraphani N, Gladwin MT, Contreras AU, Kaminski N, Tedrow JR, Milosevic J, Bleecker ER, Meyers DA, Ray A, Ray P, Erzurum SC, Busse WW, Zhao J, Trudeau JB, Wenzel SE. An airway epithelial iNOS-DUOX2-thyroid peroxidase metabolome drives Th1/Th2 nitrative stress in human severe asthma. *Mucosal Immunol.* Nature Publishing Group; 2014; 7: 1175–1185.
6. Modena BD, Tedrow JR, Milosevic J, Bleecker ER, Meyers DA, Wu W, Bar-Joseph Z, Erzurum SC, Gaston BM, Busse WW, Jarjour NN, Kaminski N, Wenzel SE. Gene expression in relation to exhaled nitric oxide identifies novel asthma phenotypes with

unique biomolecular pathways. *Am. J. Respir. Crit. Care Med.* 2014; 190: 1363–1372.

7. Giovannini-Chami L, Marcet B, Moreilhon C, Chevalier B, Illie MI, Lebrigand K, Robbe-Sermesant K, Bourrier T, Michiels JF, Mari B, Crénesse D, Hofman P, De Blic J, Castillo L, Albertini M, Barbry P. Distinct epithelial gene expression phenotypes in childhood respiratory allergy. *Eur. Respir. J.* European Respiratory Society; 2012; 39: 1197–1205.
8. Nicodemus-Johnson J, Naughton KA, Sudi J, Hogarth K, Naurekas ET, Nicolae DL, Sperling AI, Solway J, White SR, Ober C. Genome-wide Methylation Study Identifies an IL-13 Induced Epigenetic Signature in Asthmatic Airways. *Am. J. Respir. Crit. Care Med.* 2015; 193: rccm.201506-1243OC.
9. Woodruff PG, Boushey H a, Dolganov GM, Barker CS, Yang YH, Donnelly S, Ellwanger A, Sidhu SS, Dao-Pick TP, Pantoja C, Erle DJ, Yamamoto KR, Fahy J V. Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc. Natl. Acad. Sci. U. S. A.* 2007; 104: 15858–15863.
10. Leigh R, Mostafa MM, King EM, Rider CF, Shah S, Dumonceaux C, Traves SL, McWhae A, Kolisnik T, Kooi C, Slater DM, Kelly MM, Bieda M, Miller-Larsson A, Newton R. An inhaled dose of budesonide induces genes involved in transcription and signaling in the human airways: enhancement of anti- and proinflammatory effector genes. *Pharmacol. Res. Perspect.* 2016; 4: e00243.
11. Bolstad BM, Collin F, Brettschneider J, Simpson K, Cope L, Irizarry RA, Speed T. Quality Assessment of Affymetrix Genechip data. *Bioinforma. Comput. Biol. Solut. Using R Bioconductor* 2005. p. 33–47.
12. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* 2013; 14: 7.

13. Imoto Y, Tokunaga T, Matsumoto Y, Hamada Y, Ono M, Yamada T, Ito Y, Arinami T, Okano M, Noguchi E, Fujieda S. Cystatin SN Upregulation in Patients with Seasonal Allergic Rhinitis. *PLoS One* 2013; 8: e67057.

14. De Benedetto A, Rafaels NM, McGirt LY, Ivanov AI, Georas SN, Cheadle C, Berger AE, Zhang K, Vidyasagar S, Yoshida T, Boguniewicz M, Hata T, Schneider LC, Hanifin JM, Gallo RL, Novak N, Weidinger S, Beaty TH, Leung DYM, Barnes KC, Beck LA. Tight junction defects in patients with atopic dermatitis. *J. Allergy Clin. Immunol.* Elsevier Ltd; 2011; 127: 773–786.

15. Blanchard C. Eotaxin-3 and a uniquely conserved gene- expression profile in eosinophilic esophagitis. *JCI* 2006; 116.