Rhinovirus infection of airway epithelial cells uncovers the non-ciliated subset as a likely driver of genetic susceptibility to childhood-onset asthma.

Sarah Djeddi[ε,1,2,3], Daniela Fernandez-Salinas[ε,1,2,3,4], George X. Huang[5,6], Vitor R. C. Aguiar[1,2,3], Chitrasen Mohanty[10], Christina Kendziorski[10], Steven Gazal[7,8], Joshua Boyce[5,6], Carole Ober[9], James Gern[10,11], Nora Barrett[5,6], Maria Gutierrez-Arcelus*[1,2,3]

ε These authors contributed equally to this work.

[1]Division of Immunology, Boston Children's Hospital, Boston, MA, USA

[2]Department of Pediatrics, Harvard Medical School, Boston, MA, USA

[3]Broad Institute of MIT and Harvard, Cambridge, MA, USA

[4]Licenciatura en Ciencias Genómicas, Instituto de Biotecnología, Universidad Nacional Autónoma de México (UNAM), Cuernavaca, Morelos, México.

[5]Department of Medicine, Harvard Medical School, Boston, MA, USA

[6]Jeff and Penny Vinik Center for Allergic Disease Research, Division of Rheumatology, Immunology, and Allergy, Brigham and Women's Hospital, Boston, MA, USA.

[7]Department of Quantitative and Computational Biology, University of Southern California

[8]Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California

[9]Department of Human Genetics, University of Chicago, Chicago, Ill, USA.

[10]Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI, USA.

[11]Departments of Pediatrics and Medicine, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA

*Corresponding author:

Maria Gutierrez-Arcelus
Boston Children's Hospital
Division of Immunology
Karp Family Research Laboratories, Room 102016.1
1 Blackfan Circle
Boston, MA, 02115
mgutierr@broadinstitute.org
857-559-3229

Abstract

Asthma is a complex disease caused by genetic and environmental factors. Epidemiological studies have shown that in children, wheezing during rhinovirus infection (a cause of the common cold) is associated with asthma development during childhood. This has led scientists to hypothesize there could be a causal relationship between rhinovirus infection and asthma or that RV-induced wheezing identifies individuals at increased risk for asthma development. However, not all children who wheeze when they have a cold develop asthma. Genome-wide association studies (GWAS) have identified hundreds of genetic variants contributing to asthma susceptibility, with the vast majority of likely causal variants being non-coding. Integrative analyses with transcriptomic and epigenomic datasets have indicated that T cells drive asthma risk, which has been supported by mouse studies. However, the datasets ascertained in these integrative analyses lack airway epithelial cells. Furthermore, large-scale transcriptomic T cell studies have not identified the regulatory effects of most non-coding risk variants in asthma GWAS, indicating there could be additional cell types harboring these “missing regulatory effects”. Given that airway epithelial cells are the first line of defense against rhinovirus, we hypothesized they could be mediators of genetic susceptibility to asthma. Here we integrate GWAS data with transcriptomic datasets of airway epithelial cells subject to stimuli that could induce activation states relevant to asthma. We demonstrate that epithelial cultures infected with rhinovirus significantly upregulate childhood-onset asthma-associated genes. We show that this upregulation occurs specifically in non-ciliated epithelial cells. This enrichment for genes in asthma risk loci, or ‘asthma heritability enrichment’ is also significant for epithelial genes upregulated with influenza infection, but not with SARS-CoV-2 infection or cytokine activation. Additionally, cells from patients with asthma showed a stronger heritability enrichment compared to cells from healthy individuals. Overall, our results suggest that rhinovirus infection is an environmental factor that interacts with genetic risk factors through non-ciliated airway epithelial cells to drive childhood-onset asthma.

Introduction

Asthma is a complex and heterogeneous disease that affects 300 million children and adults worldwide and represents a significant burden to healthcare ($82 billion for the US in 2013) [1]. It is characterized by inflammation of the airways leading to recurrent episodes of airflow obstruction and symptoms such as wheezing, shortness of breath and coughing. Asthma patients have impairments of epithelial barrier function, manifested by irregular disruption of the tight junctions, detachment of ciliated cells and reduced expression of cell-cell adhesion molecules (for instance, E-cadherin) [2]. This barrier disruption allows environmental substances like allergens, viruses, bacteria and toxic substances to penetrate the submucosa more easily. Allergens, viral infections, and type 2 inflammation have been shown to further damage the barrier integrity of the airway epithelium; moreover, they trigger asthma exacerbations [3]. Longitudinal epidemiological studies have shown that wheezing caused by rhinovirus infection in children is a risk factor for developing asthma later in childhood [4–6]. These observations have led to two hypotheses: (1) rhinovirus infection could be causal in asthma development or (2) rhinovirus-induced wheeze is a biomarker that identifies children at increased risk for asthma development. In some children, rhinovirus wheezing does not lead to asthma.

Genome-wide association studies (GWAS) have discovered more than 100 risk loci for asthma [7,8]. Similar to other complex diseases, the vast majority of the likely causal risk variants are non-coding. As a consequence, deciphering the mechanisms through which the risk alleles lead to disease is challenging; it has been achieved for only a small minority of risk loci. Using a suite of recently developed methods that integrate GWAS data with functional genomics datasets, investigators have discovered key cell types that mediate the genetic susceptibility to complex diseases. For example, risk variants for rheumatoid arthritis are enriched in regulatory elements specific for CD4 T cells, and studies in patients and mice have shown the relevance of these cells in the pathogenesis of this disease [9–13]. Additionally, GWAS integration with transcriptomics revealed that a significant proportion of the risk alleles for Alzheimer's disease act through the myeloid lineage rather than the brain [14]. Alzheimer's disease is now considered an immune-

mediated disease [14,15]. For asthma, T cell-specific regulatory elements and gene expression are enriched in genetic risk loci [12,13,16,17], with some highlighting particularly Th2 cells consistent with the role of type 2 inflammation in asthma pathogenesis [18–21].

The observations that non-coding risk variants affect gene regulation in cell types relevant to each disease have motivated large-scale transcriptomic studies to identify genetic variants that are associated with both gene expression (expression quantitative trait loci, eQTL) and disease risk. However, only 25-40% of risk variants for immune-mediated diseases co-localize with eQTLs in immune cells [22–24] . For asthma, alleles at only 47% of loci co-localize with leukocyte expression and/or splicing QTLs [22]. Hence, the regulatory effects of most non-coding risk variants remain unknown. More recent studies have highlighted that these “missing regulatory effects'' could be hidden in specific activation or differentiation cell states that haven’t been systematically ascertained [25–28].

GWAS enrichment studies have been highly biased towards annotations of blood immune cell types, with reduced resolution when using other tissues relevant to the context of asthma, such as GTEx tissues from post-mortem human organs [11,13,16,29,30]. Here we sought to define whether airway epithelial cell states could be driving genetic susceptibility to asthma. We analyzed 10 single-cell and bulk transcriptomic datasets of epithelial cells subject to different activation conditions. We integrated these datasets with GWAS summary statistics for childhood-onset asthma (COA), adult-onset asthma (AOA), unspecified-onset asthma (henceforth referred to as all asthma), and a genetically correlated type 2 inflammation trait: allergy/eczema (**Supplementary Table 1**) [31–33]. We additionally tested three control traits to assess the specificity of our findings. We used state of the art methods that control for linkage disequilibrium, take advantage of most ascertained genetic variants in the genome, and have been shown to work well for bulk and single cell datasets [13,17].

Results

We applied two methods that use GWAS data to identify relevant cell types for disease. Linkage Disequilibrium Score-regression in Specifically Expressed Genes (LDSC-SEG) identifies heritability enrichment in genomic annotations (such as genes or chromatin marks with specific presence in a particular cell type or cell state) [13]. Single-cell disease-relevance score (scDRS) identifies cells, from single-cell RNA-seq data, that significantly express genes in GWAS loci (weighted according to their strength of association with disease) relative to null sets of control genes in the same dataset [17]. We retrieved GWAS summary statistics from asthma related traits as described above. Throughout our analyses we included three complex traits as controls: height, as a non-immune control, Alzheimer's disease (AD) as a trait implicating myeloid cells, and rheumatoid arthritis (RA) as a lymphocyte-driven disease with a strong T cell component [11,13,16,17,29].

T cell validation

First, we sought to validate T-cell involvement in the genetic susceptibility to asthma, as previously reported in the literature [13,16,17,29]. Applying LDSC-SEG to bulk ATAC-seq data of human peripheral blood leukocyte populations, we confirmed that T-cell specific open chromatin regions are significantly enriched in heritability for asthma-related traits (**Supplementary Figure 1**)[16]. Furthermore, when comparing cell types between their resting and activated state, we confirmed that activation-specific open chromatin in T cells has significant heritability enrichment for all the asthma-related traits (**Supplementary Figure 1E**). Next, we applied scDRS to single-cell RNA-seq datasets to identify cells with significant over-expression of risk genes identified from GWAS studies (see Methods). In sinonasal mucosa tissue from healthy donors and chronic rhinosinusitis patients, we observed that 21-83% of cells with significant disease relevant score (10% FDR) for asthma-related traits are T cells (**Supplementary Figure 2**). In a dataset of house dust mite-activated T cells from asthma and allergic patients, we observed that among the cells with significant disease relevant score, most were T effector and Th2 cells (51-57% and 48-42% respectively, **Supplementary Figure 3**). Overall, these analyses confirm the validity of the

methods used for this study and confirm previous findings showing the relevance of T cells in the genetic susceptibility to asthma-related traits.

Rhinovirus infection induces upregulation of asthma-associated genes in epithelial cells from healthy donors. To assess the role that rhinovirus infection could play in asthma genetic susceptibility at the epithelial cell level, we analyzed a publicly available bulk RNA-seq data of basal airway epithelial cells from healthy donors (N=9) that were infected in vitro with RV-A16 (rhinovirus species RV-A, subtype 16) or treated with PBS vehicle control (**Figure 1A**) [34]. Using a linear mixed model, we performed differential expression analysis (DEA) testing for rhinovirus infection versus PBS vehicle. From the 14,883 tested genes, we selected the top 10% based on t-statistic (as recommended by LDSC-SEG) to select genes that are upregulated with rhinovirus infection(**Figure 1B**). We used LDSC-SEG to investigate whether this gene set showed an enrichment of asthma heritability. Our analysis showed significant heritability enrichment in rhinovirus-upregulated genes for all asthma (P = 0.033) and COA (P = 0.037), with a larger coefficient observed for COA (**Figure 1C**). Moreover, we did not observe any significant enrichment for any of the control traits (AD, RA, height) (**Supplementary Figure 4A**). The fact that we did not observe significant heritability enrichment in RV-upregulated genes for RA suggests that the signal observed for all asthma and COA is not due to a general immune transcriptional response, but rather a response that is specific for RV-infection of epithelial cells.

Next, we sought to validate these findings in an independent study using a different strain of RV. We reanalyzed a bulk RNA-seq time course dataset of epithelial cells from 3 healthy donors where cells were infected with RV-C15, and samples were collected before infection and at 12, 24, and 42 hours post-infection (HPI) (**Figure 1D**) [35]. We performed differential expression analysis to identify genes upregulated specifically in each time point (versus all others), and applied LDSC-SEG (**Supplementary Figure 4B**). We observed significant heritability enrichment for upregulated genes by rhinovirus infection specifically at 24 hours for all asthma-related traits ($P < 0.05$, **Figure 1E**). This enrichment was higher for COA and allergy/eczema, compared to all asthma and AOA (**Figure 1E**). Genes that were specifically expressed at 42 hours post-infection

also had positive enrichment for heritability in all traits, but this was only significant for allergy/eczema (P = 0.03). Once again, this enrichment was not present in any of the control traits (**Supplementary Figure 4C**). Together, these findings suggest that rhinovirus-infected epithelial cells represent a cell state that may mediate genetic susceptibility to asthma, with a greater contribution to COA than to AOA.

Asthma-associated genes after rhinovirus infection are specifically enriched in non-ciliated epithelial cells. We then asked whether there were specific epithelial cell subsets that may mediate asthma genetic risk after rhinovirus infection. To evaluate this, we used single-cell RNA-seq data of 3 healthy donors, where airway epithelial cell samples were infected with RV-C15 or resting and profiled at 24 hours (**Figure 2A**, **Supplementary Figure 5A-C**). We performed cell clustering and then annotated the clusters based on epithelial cell markers (**Supplementary Figure 5D**). We identified 2 ciliated cell subsets, and 5 non-ciliated cell subsets: basal, deuterosomal, neuroendocrine, secretory, and transitional (**Figure 2B**, **Supplementary Figure 5D**). We applied scDRS on this dataset to identify cells with significant over-expression of asthma-associated genes. The number of cells with significant disease relevant scores (10% FDR) varied per trait: 147 cells for all asthma, 850 cells for COA, 0 for AOA and 147 cells for allergy/eczema (**Figure 2C**). Notably, COA was the trait with the highest proportion of disease-relevant cells. Among the cells with a significant disease relevant score, 99% corresponded to the stimulated condition. Furthermore, the disease relevant cells were strongly over-represented in the non-ciliated cell subsets, representing 96-99% of the significant cells, while they make up 25% of the whole dataset (**Figure 2C-D**). As expected, we did not observe any significant cells for height, AD and RA (**Supplementary Figure 5E**).

Overall, we identified non-ciliated cells as the main epithelial cell subset with significant upregulation of childhood-onset asthma-associated genes after rhinovirus infection. However, only ciliated cells are known to be infected by RV-C15, which we confirmed by looking at the expression of the RV-15 receptor (*CDHR3*) and the presence of the viral sequence itself in the scRNA-seq data [36] (**Supplementary Figure 5F-G**). We therefore hypothesized that ciliated cells

may communicate with non-ciliated cells upon rhinovirus infection. We used CellphoneDB to investigate some of the possible ligand-receptor mechanisms through which cells may be communicating [37]. Specifically, we looked for ligand-encoding genes expressed in ciliated cells and their corresponding receptor-encoding gene expressed in non-ciliated cells. Furthermore, we required that the ciliated cell ligand-encoding gene is upregulated upon rhinovirus infection. We identified a potential pair of interactors consisting of *LGALS9*, which codes for a galectin from the beta-galactoside-binding protein family implicated in the modulation of the cell-cell and cell-matrix interactions, and *SORL1*, a gene encoding sortilin-related receptor, which may have a role in endocytosis and intracellular trafficking (**Supplementary Figure 5H-I**).

Enrichment of asthma-associated genes after rhinovirus infection is strong in epithelial cells from asthma patients. Having observed the enrichment of asthma-associated genes after RV infection in both bulk and single-cell level in healthy subjects, we asked whether we would observe the same enrichment in samples coming from asthma patients. To do this, we repeated the differential expression analysis between rhinovirus RV-A16 infection and treatment with PBS from the first dataset [34], this time using the asthma patient cohort (**Figure 3A**). We found 2,843 differentially expressed genes at 5% FDR, 1,353 of them upregulated and 1,481 downregulated after RV infection (**Figure 3B**). Of the 2,843 differentially expressed genes in patients, 1,834 were also differentially expressed in healthy controls. After selecting the top 10% genes by t-statistic (1,488) and running LDSC-SEG, we observed significant heritability enrichment for all asthma ($P = 0.003$) and COA ($P = 0.003$, **Figure 3C**). We did not observe heritability enrichment for down-regulated genes by RV ($P > 0.05$, **Supplementary Figure 6C**). While the results were consistent with what we observed in healthy controls (COA $\tau^*=0.17$, all asthma $\tau^*= 0.15$), the enrichment of RV-upregulated asthma-associated genes was more significant and had a larger enrichment coefficient in the asthma patients (COA $\tau^*=0.32$, all asthma $\tau^*=0.25$).

Based on these results, we hypothesized that asthma patients might have airway epithelial cells in a transcriptomic state that over-expresses asthma-risk genes in comparison to healthy controls, which might be linked to or independent of their response to RV. To test this, we

analyzed differentially expressed genes between asthma patients and healthy individuals taking all samples while controlling for RV/PBS treatment. We found 994 differentially expressed genes between patients and controls (5% FDR, **Figure 3D**). After selecting the top 10% genes upregulated in patients based on t-statistic (1,593), we found a suggestive significant enrichment for COA heritability (P = 0.05, **Figure 3E**). As expected, our control traits did not have any significant heritability enrichment for either of the annotations tested (**Supplementary Figure 6**). Overall, these results suggest that the epithelial cells from patients could be in a state that is over-expressing asthma-associated genes.

Genes at asthma risk loci upregulated with rhinovirus infection in airway epithelial cells. We then investigated which of the genes upregulated by rhinovirus infection are associated with COA and AOA. To do so, we retrieved the GWAS lead variants identified by Ferreira et al. [38]. We then linked risk variants to genes using three approaches: (1) selecting the likely target genes identified by the locus-to-gene (L2G) algorithm of Open Targets Genetics, herein called L2G genes, (2) selecting the closest gene to the lead variant, and (3) selecting genes within a 250kb window of the lead variant (see Methods). From these gene lists, we selected genes that were upregulated upon rhinovirus infection in epithelial cells at 5% FDR (**Figure 4**) [34,35]. For COA we identified 55 risk loci with genes upregulated upon rhinovirus infection in epithelial cells, 13 of which have L2G likely target genes (e.g. *IL1RL1*, *IL4R*, *GSDMB*, *OVOL1*, *MYC*), and 6 are the closest gene to the lead variant (e.g. *IRF1*, *GPR183*). For AOA only 19 risk loci have genes upregulated by RV in epithelial cells, among which 3 are likely target genes (e.g. *IL4R*, *HDAC7*, *IL1RL1*) and 3 are the closest gene to the lead variant (e.g. *RAPGEF3*, *IRF1*, *SSR3*). Few genes were shared between COA and AOA (*IRF1*, *IL4R*, *PDLIM4*, *IL1R2* and *IL1RL1*).

After having characterized the asthma-associated genes that are upregulated in RV-infected epithelial cells, we sought to define which of these genes are shared with T cells. To do so, we compared the levels of expression of the genes in the RV-datasets with a dataset we previously published consisting of 8 activation time points of human periphery memory CD4+ T cells stimulated with anti-CD3/CD28 microbeads [25]. One of the highlighted genes that also had an

increased expression in T cells was *MYC*, with an increase at 2 and 4 hours post-stimulation. Notably, this gene's expression was not only increased after rhinovirus infection within asthma patients (P=0.01) and within healthy controls (P=0.0009) but was also upregulated in patients compared to controls (P=0.01). On the other hand, *OVOL1*, another GWAS gene upregulated in RV-infected epithelial cells, shows the same pattern of expression as *MYC* in epithelial cells, but shows almost no expression in T cells (**Figure 4B**, **Supplementary Figure 7**). *OVOL1* is also associated with atopic dermatitis, another type 2 inflammatory disease and a recent meta-analysis study confirmed this susceptibility locus for eczema-associated asthma [39].

Other viral infections in epithelial cells and their association with asthma susceptibility. We next asked whether other viruses could potentially be inducing upregulation of asthma-associated genes in epithelial cells. First, we analyzed a bulk RNA-seq dataset of bronchial epithelial cells stimulated by influenza A virus at 48 hours or sham control (N = 3 healthy donors, **Figure 5A**). We identified differentially expressed genes between influenza A and the sham control and we selected the top 10% upregulated genes ranked by a t-statistic to run LDSC-SEG (**Supplementary Figure 8A**). The results show a significant enrichment of heritability for the four asthma-related traits ($P < 0.05$), suggesting that influenza A infection also significantly upregulates asthma and allergy-associated genes (**Figure 5B**). We did not observe any significant enrichment for the control traits (**Supplementary Figure 8B**).

Subsequently, we analyzed a single-cell RNA-seq dataset of immune and non-immune cells obtained by nasopharyngeal swabs from COVID-19 patients or healthy donors (N=58, **Figure 5C-D**) [40]. We identified cells with significant disease-relevant scores at 10% FDR, among which we found 174 cells for AOA, 795 cells for COA, 594 for allergy/eczema and 280 for all asthma (**Figure 5E**). COA yielded the highest number of disease relevant cells. We observed an increase in proportions of enriched cells for T cells in all traits and for squamous cells in COA, allergy/eczema and all asthma (**Figure 5F**). More precisely, T cells which represented 5% of the cells in the dataset, constituted 78% of significant cells for AOA, 11% for COA, 42% for allergy/eczema and 61% for all asthma (**Figure 5F**). The COVID-19 status did not have any significant impact on the

asthma-associated expression, as the disease-relevant cells were not overrepresented in any specific patient group (**Figure 5G**). As expected, we observed an enrichment for the macrophages cluster in Alzheimer's and T cells cluster in rheumatoid arthritis (**Supplementary Figure 8C**). Additionally, we analyzed a dataset of bronchial epithelial cells (BECs) infected with SARS-CoV-2 in vitro, and profiled with scRNA-seq after one, two or three days, along with non-infected cells (N = 1 healthy control, **Supplementary Figure 9A-B**). We clustered and annotated ciliated and non-ciliated cell subsets, and applied scDRS for the four asthma-related traits (**Supplementary Figure 9C-D**). We identified cells with significant disease-relevant scores at 10% FDR, finding 3 cells for AOA, 802 cells for COA, 243 cells for allergy/eczema and 621 cells for all asthma (**Supplementary Figure 9D-E**). The disease relevant cells were strongly over-represented in the non-infected cells subsets, constituting 62-89% of the significant cells (**Supplementary Figure 9F**). Overall, these findings indicate that the influenza virus significantly induces the expression of asthma-associated genes, whereas SARS-CoV-2 does not.

Epithelial cells activated with cytokines relevant for type 2 inflammation.

Both epithelial and immune cells respond to cytokines by upregulating signaling pathways that drive inflammation. Some pro-inflammatory cytokines relevant to asthma are IL-4, IL-13, IL-17 and interferon (IFNγ). These cytokines are upregulated in subsets of patients with severe or type 2 asthma [41–46]. Moreover, blockade of IL4Rα is a highly effective treatment for moderate to severe asthma [47]. Consequently, we asked whether epithelial cells stimulated with cytokines might induce a transcriptional program enriched for asthma-associated genes. First we used a bulk RNA-seq dataset consisting of human bronchial epithelial cells (HBECs) that were stimulated with either IFNα, IFNγ, IL-13 or IL-17 (N = 6 healthy donors, **Figure 6A**) [48]. We selected the top 10% upregulated genes by t-statistic for each stimulus (**Supplementary Figure 10A**). We used LDSC-SEG to analyze these 4 sets of genes in the 4 asthma-associated traits and the control traits. We did not identify any significant heritability enrichment for any of the stimuli gene sets for the asthma-associated traits tested here (**Figure 6B**), nor for the control traits (except for genes upregulated by IFNγ for rheumatoid arthritis, **Supplementary Figure 10B**).

Next, given that IL-13 might work synergistically with IL-4, we performed bulk RNA-seq of nasal airway epithelial cells from healthy donors (N = 5) co-stimulated in vitro with IL-4 and IL-13 (**Figure 6C**). We tested DE genes for the IL-4-IL-13 condition compared to the unstimulated control. In line with the results observed in the previous analysis, we found no significant enrichment for any of the asthma-associated (**Figure 6D**) or control traits (except for AD, **Supplementary Figure 10E**). Together, these results suggest that epithelial cells upregulate asthma-associated genes in a stimulus-specific manner, which to the extent of this study, is not caused by the stimulation with the cytokines tested here.

Discussion

While some genetic risk variants for asthma are enriched near genes with T cell-specific expression [13,17,22,28], the effects of most variants on gene regulation remain unknown. In this study, we asked whether some of these "missing regulatory effects" could be hidden in airway epithelial cells, given they are the first line of contact for respiratory viruses, including those that have been associated with asthma development or exacerbations. We analyzed ten transcriptomic datasets of human airway epithelial cells cultured under different stimuli and integrated them with genetic susceptibility data for asthma and related traits. We consistently showed that rhinovirus-activated epithelial cells significantly upregulate genes at childhood-onset asthma risk loci. We observed this in samples from healthy donors and even more so in cells from asthma patients. Notably, we discovered that non-ciliated cells are the subset driving these associations with asthma, indicating that non-ciliated airway epithelial cells activated with rhinovirus are key mediators of genetic susceptibility to childhood-onset asthma. While other respiratory viruses, such as influenza might also significantly upregulate genes at asthma risk loci, this is not likely a general virus response or epithelial cell activation signature, given that we did not detect asthma heritability enrichment for SARS-CoV-2 or cytokine-upregulated genes.

Our findings are consistent with epidemiological studies that have shown associations between wheezing illness caused by rhinovirus infection and asthma development in children [4,49–51]. Additionally, a previous birth cohort study identified genetic variants at the 17q21 locus that

were associated with asthma in children who had rhinovirus-associated wheezing illness in the first 3 years of life, but not in children who had RSV-associated wheezing illnesses at those same ages [52]. In that study, rhinovirus upregulated two genes at this locus, *ORMDL3* and *GSDMB*, in PBMCs [52]. Here, we observe that in non-ciliated airway epithelial cells rhinovirus induces upregulation of *GSDMB* as well as putative causal genes in 54 additional loci. This demonstrates a widespread interaction between *in vitro* rhinovirus infection and polygenic susceptibility to childhood-onset asthma, specifically mediated through airway epithelial cells. These findings are concordant with a previous study reporting that genes at COA-specific risk loci (as compared to AOA) have high expression in skin, which is a barrier tissue with an abundance of epithelial cells [30]. Overall, our findings support the hypothesis that rhinovirus could be causally linked to asthma development in children and not just be a biomarker of children destined to develop asthma. Not all children that get RV-wheezing develop asthma, and our findings suggest that the combination of preschool rhinovirus wheezing illnesses and a high genetic burden synergistically promote the development of childhood asthma.

We discovered that non-ciliated cells (basal, secretory, and transitional) are the specific cell subsets that overexpress genes at asthma risk loci. This suggests that an important fraction of the non-coding risk variants for asthma likely affect gene regulation in non-ciliated cells under specific viral activation states. In our study we looked at two different rhinovirus types. For the case of RV-C15, the receptor of the virus, CDHR3, is mainly expressed in the ciliated cells [53,54] and viral RNA quantification confirmed this subset is the one directly infected by the virus (**Supplementary Figure 5G**). This suggested that RV-infected ciliated cells efficiently transmit a signal to the non-ciliated cells, which then express genes in asthma risk loci. The time course experiments indicated this upregulation of asthma-associated genes occurred predominantly at 24 and 42 hours [35]. In our analyses of RV-A16 infected epithelial cells, the data came from bulk RNA-seq of basal cells (non-ciliated) treated for 24 hours. In contrast to RV-C15, RV-A16 binds to the ICAM receptor, which is expressed in ciliated cells and basal cells [55]. Strikingly, while the cell subsets that get directly infected differ between the two RV strains, both significantly upregulated asthma-associated genes in non-ciliated cells at 24 hours. By contrast, for SARS-CoV-

2, *in vitro* infection did not upregulate genes in asthma risk loci; rather, the non-infected cells presented a significant expression of asthma-associated genes. Although the data in this study came from only one individual, the cells with significant disease relevant scores were also predominantly non-ciliated cells (> 83%, **Supplementary Fig. 9**). Future single-cell studies with larger sample sizes and ascertaining infection by multiple types of viruses could point to additional epithelial cell subsets and cell states as candidate drivers of genetic susceptibility to asthma.

The observations in our study may also be relevant to virus-induced asthma exacerbation [56]. Here, we not only demonstrate that rhinovirus infection induces a transcriptional response enriched in childhood-onset asthma risk, but we also identified a heritability enrichment for genes upregulated in asthma patients compared to controls, even when controlling for RV infection (suggestive P = 0.051, **Figure 3E**). This result goes in line with a previous observation that, at the open chromatin level, airway epithelial cells of asthma patients have a large amount of open chromatin regions at baseline that are RV-response regions in healthy controls [34]. It is possible that over-expression of asthma-associated genes at baseline may increase the risk for acute virus-induced exacerbations in patients with asthma. Influenza infections, which can cause asthma exacerbations (especially in adults) [57–59], induced an enrichment of both adult-onset and childhood-onset asthma heritability in influenza upregulated genes in airway epithelial cells. Furthermore, SARS-CoV-2 seems less likely than other viruses to provoke asthma exacerbations, and asthma does not appear to be a risk factor for severe SARS-CoV-2 infection [60]. This could be due to multiple reasons, such as allergy-induced reduction in the ACE2 receptor [61], but it is also in line with our observations that SARS-CoV-2 infection itself does not induce a transcriptional program significantly enriched in asthma heritability [61].

Our study had some limitations. We were not able to ascertain all possible epithelial cell subsets and states. Most of the datasets analyzed involved *in vitro* infections, rather than *in vivo* infected samples. Additionally, all samples came from adults, which made it all the more striking that we detected heritability enrichments for childhood-onset asthma. Future studies in children are

important to validate these findings. Furthermore, we were limited by the cell sources and specific time points and experimental designs of each study. In particular, the absence of asthma heritability enrichment in cytokine-upregulated genes in epithelial cells could imply multiple scenarios. One possibility could be that even though pro-inflammatory cytokines (IL-4, IL-13, IFNα, IFNγ, IL-17) upregulate many genes in epithelial cells (684-2876 at 5% FDR in our analyses), they do not significantly interact with polygenic risk factors for asthma in epithelial cells. Other possibilities for the absence of signal could be that the cytokine-induce activation might interact with genetic risk factors acting in T cells or other non-epithelial cells, or that there are interactions with environmental conditions not present in the models included in this analysis.

Overall, our findings of asthma heritability enrichment in various epithelial cell states (resting versus virus-infected, patients versus healthy controls) could reflect variability in how risk variants contribute to disease onset versus progression. These results highlight the importance of studying the cellular context in which GWAS loci contribute to disease risk and will ultimately help to better understand the mechanisms through which those risk variants are acting. Moreover, the outcomes of our study could open the door to new therapeutic avenues. Indeed, drug targets that have genetic evidence are more likely to be approved and move forward to clinical trials than those without it [62]. Large-scale multi-omic studies (with comparable power to GWAS [63]) of non-ciliated airway epithelial cells activated with rhinovirus could help identify the target genes of non-coding asthma risk variants, together with functional validations with approaches such as base editing. Finally, if our observations are confirmed and further characterized by future studies, it would support the development of a rhinovirus vaccine or other protective intervention as a way to prevent childhood-onset asthma [64].

Material and methods

Dataset collection

We downloaded transcriptomic datasets from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO), Genome Sequence Archive (GSA) and from ImmPORT.

We also downloaded a chromatin accessibility dataset (ATAC-seq) from GEO (**Supplementary Table 2**).

Supplementary Table 2. Datasets used in this study.

Data type	Source	Accession number
ATAC-seq	Calderon et al., 2019	GSE118189
scRNA-seq	Wang et al., 2023	HRA000772 (Genome Sequence Archive)
scRNA-seq	Seumois et al., 2020	GSE146170
RNA-seq	Helling et al., 2020	GSE152550
RNA-seq	Basnet et al., 2023	SDY1882 (ImmPORT)
scRNA-seq	Basnet et al., 2023	SDY1882 (ImmPORT)
RNA-seq	Tao et al., 2022	GSE193164
scRNA-seq	Ravindra et al., 2021	GSE166766
RNA-seq	Koh et al., 2022	GSE185200
RNA-seq	Barret lab	Will be published as part of current study

Ethical approval

The Mass General Brigham Institutional Review Board gave ethical approval for the Barrett dataset.

GWAS collection

We downloaded pre-processed summary statistics for the four asthma-associated traits; adult-onset asthma, childhood-onset asthma, all asthma, allergy/eczema and for those in our control panel; height, Alzheimer's disease and rheumatoid arthritis (**Supplementary Table 1**).

Supplementary Table 1. Summary statistics used for heritability enrichment analysis.

	Studied trait	Reference
Asthma-associated traits	Adult-onset asthma	Ferreira et al., Am J Hum Genet, 2019 [38]
	Childhood-onset asthma	Ferreira et al., Am J Hum Genet, 2019 [38]
	Allergy/eczema	UK biobank [65]
	All asthma	UK biobank [65]
Control traits	Height	Lango, et al., Nature, 2010 [66]

	Alzheimer's disease	Lambert, et al., Nat Genet, 2013 [67]
	Rheumatoid arthritis	Okada, et al., Nature, 2014[68]

Summary statistics processing for visualization

We downloaded the childhood-onset asthma and adult-onset asthma [38] summary statistics from the GWAS catalog. We used the harmonized summary statistics in the GRCh38 version of the genome. We removed the MHC region (chr6:28510120-33480577).

Bulk RNA-seq data processing and quality check

FASTQ files were aligned to the GRCh38 or GRCh37 human genome using STAR (v2.7.9a) with standard parameters and two-pass mode, or the Salmon tool (v1.5.1). For BAM files generated with STAR, counts were calculated using RSEM (v1.3.3). We normalized the counts by transforming them to their log2(TPM+1) value, where TPM stands for transcripts per million. To detect outlier samples, we performed principal component analysis on scaled normalized expression for the top 1000 most variable genes that were expressed in at least 25% of the samples. For alignment and quantification, we used the ENSEMBL reference annotation release 105 which was downloaded from the ENSEMBL website.

Differential expression analyses

Differential gene expression was tested using a linear mixed model, similar to what we did in Gutierrez-Arcelus et al.[69]. Specifically, we used a likelihood ratio test between two nested models (*anova* function in R). In these models, gene expression levels (log2(TPM + 1)) represent the dependent variable. "Donor ID" was included as a predictor variable, treated as a random effect. To compare one condition against the others, we indicated with 1 the tested condition and 0 for the others (the test variable). We used the function "lmer" from the R package "lme4" to implement the model. For risk gene visualization in the miami plot, P-values were corrected for multiple hypothesis testing using the package "qvalue". Differentially expressed genes at 5% FDR are reported for depicting specific genes in risk loci. After each analysis, we calculated a t-statistic for each gene to rank them and chose the top 10% as annotations for heritability enrichment

473 analysis (see LDSC-SEG section below). The details of each analysis are divided by dataset and
474 described in the following section.
475
476 - *Helling Dataset*
477 To assess DE genes between rhinovirus treatment and PBS vehicle control within healthy
478 individuals, we tested genes that had a normalized count greater than 1 in at least 9 samples
479 which led to a total of 14,883 genes. The threshold for the minimum number of samples reflected
480 half of the biological replicates (9/18). In addition to having "donor ID" as a random effect, we
481 accounted for "sex" as a fixed effect. We repeated this process in asthma patients only, testing
482 14,888 genes for differential expression between rhinovirus infection and PBS vehicle control. To
483 find DE genes from asthma patients compared to healthy controls, we took all the samples and
484 recalculated the number of genes present in at least 9 samples. We tested 15,935 genes and
485 incorporated "treatment" with either PBS or rhinovirus as a fixed effect covariate in the model,
486 and tested for disease status (0/1).
487
488 *Tao Dataset*
489 We included 16,031 genes having a normalized count greater than 1 in half of the samples (3/6
490 samples). We tested for differential expression between influenza treatment and control (sham).
491
492 *Koh Dataset*
493 We tested 14,988 genes, to find differentially expressed genes specific to each condition
494 compared to all others: IFNα, IFNγ, IL-13, and IL-17, respectively. We selected genes having a
495 normalized count greater than 1 in at least 6 samples. This number reflected the smallest amount
496 of replicates found across conditions (6/36 samples).
497
498 - *Basnet Dataset*
499 We excluded the resting sample from donor B03 from these analyses (see Methods, Bulk RNA-
500 seq, QC, and analysis). To obtain differentially expressed gene profiles for each time point, we
501 performed four separate models in which we tested a single time point against the other two

and the resting condition. Our fourth model tested for DE genes in activation conditions versus the resting state through the same approach. Since we had 3 biological replicates for most time points, we tested genes with a normalized count greater than 1 in at least (3/11) samples, which yielded 14,380 genes. This gene set was used for all models.

- *Barrett Dataset*

Air-liquid interface (ALI) cultures were grown from nasal basal epithelial cells from 6 healthy adult donors. ALIs were allowed to mature for 14 days, then stimulated with 10 ng/mL of IL-4 and 10 ng/mL of IL-13 for an additional 7 days, and then lysed with TCL buffer (Qiagen 1031576) at the conclusion of the experiment. Lysates were stored at -80C and later submitted to the Broad for SmartSeq2 low input bulk RNA-seq (38bp paired-end sequencing).

To obtain genes DE under IL-4 and IL-13 co-stimulation we compared them against the non-stimulated cells. We tested 13,518 genes that had a normalized count greater than 1 in at least half of the samples (9/18) samples.

ATAC-seq data processing and differential accessibility analysis

We used the 829,942 consensus peaks called by Calderon et al. (peaks were called in each sample separately, then merged across samples, and then counts were re-calculated for all samples using the merged peak coordinates). We transformed counts into reads per kilobase per million (RPKM), then quantile normalized and finally scaled to their log2(normalized RPKM+1). To assess differentially accessible (DA) peaks, we first calculated the mean normalized count per cell type and then created cell-type accessible sets of peaks. We included a peak in the set if it had a normalized count greater than the mean of the cell type in at least half of the samples corresponding to subtypes of that cell type. We tested between 400 and 600 thousand peaks per cell-type for DA. To do so, we implemented a linear mixed model using the normalized counts as the response variable, and for the predictor, a bit flag system indicating 1 if the sample belonged to the tested cell type and 0 for the remaining cells. Peaks were sorted by t-statistic and we took the top 10% peaks for each cell-type-specific annotation. This process was replicated for a second selection model, implemented to divide peaks between stimulated and unstimulated categories.

Single-cell RNA-seq, QC and analysis

FASTQ files from the single-cell RNA-seq dataset from Wang et al., Seumois et al., Basnet et al., and Ravindra et al., [35,70–72] were downloaded as indicated in the "Dataset Collection" section. For each dataset, we aligned FASTQ files to the human reference human genome GRCh38 [73], using the GENCODE release 32 with cellranger count (v6.1.2), using default parameters. For Basnet et al, we added the RV-C15 sequence to the reference genome. Counts were then aggregated using the cellranger aggr (v6.1.2) function with the default parameters. The subsequent analyses were done for each of the datasets individually. We discarded cells with less than 500 genes expressed and cells expressing more than 20% of mitochondrial genes. We normalized raw counts with the "LogNormalize" method from Seurat package (v4.0.5). We used the normalized counts to perform a PCA with the 1000 most variable genes. We used the top 20 principal components to perform dimensionality reduction with UMAP to visualize the data. We identified clusters using the "FindClusters" function with the Louvain algorithm and a resolution parameter of 0.2, using the top 20 principal components (PCs). We corrected PCs with Harmony package [74] (v0.1.1) as indicated as follows, if nothing is indicated, no corrections were applied. For Wang et al., dataset we corrected for "donor ID" and "tissue". For Basnet et al., dataset we corrected for "donor ID" and virus infection. For Ravindra et al., dataset we corrected for virus infection.

To identify which cellular type was present in each cell cluster, we used the function "FindVariableFeatures" (parameter; test.use=wilcox) from Seurat package to identify differentially expressed genes in each cluster. If immune cells were present in the dataset we used the tool MCPcounter (v1.2.0) to annotate immune cells [75]. Based on these cellular markers we annotated the clusters with data from the literature [35,76,77].

For data from Ziegler et al.[40] we used the UMAP coordinates, and the cell annotation originally published by the authors.

LDSC-SEG

State-specific gene sets were generated using the top 10% of the genes tested ranked by t-statistic for each of our DE analyses. Genes coordinates were mapped from the human genome

reference GRCh37 GTF file and formatted into bed files. We repeated this process to generate control bed files containing all the genes tested for DE in each analysis. Symmetric windows of 100kb were added at each side of the genes using the bedtools (v2.31) "slop" function. For bed files containing ATAC-seq peaks, this window consisted of 225 bp at each side of the peak, to represent a similar genomic coverage. LD-Score files were generated using the LDSC pipeline along with data from HapMap 3 and Phase 3 of the European 1000 Genomes obtained from the LDSC-repository. The regression was run using the baseline model v1.2. We reported the P values of regression coefficients, and normalized regression coefficient as per-standardized-annotation effect sizes τ^* as in (Gazal et al. 2017 Nat Genet)[78] to allow for multi-trait comparisons. Regression P values were corrected using the "p.adjust" function from R using both FDR and bonferroni methods. The reference GTF file used to map the genes was obtained from GENCODE (v37).

MAGMA and scDRS

We used the adult-onset asthma, the childhood-onset asthma, the allergy/eczema, the all-asthma, the rheumatoid arthritis, and the Alzheimer's GWAS summary statistics as well as the corresponding set of 1000 putative disease genes (obtained with MAGMA) provided in the original publication of the scDRS method [17,79].

We used MAGMA (v1.10) to compute the gene-level association P-values and z-scores from GWAS summary statistics of Height. We transformed P-values to z-scores using this formula: 2*pnorm(abs(zscore), mean = 0, sd = 1, lower.tail = F) in R. To map SNPs to genes, we used magma with default parameters specified in the scDRS documentation. We retrieved the top 1000 genes based on MAGMA z-score as putative disease genes.

We used scDRS (v1.0.2) to quantify the expression of the putative disease genes derived from GWAS summary statistics using MAGMA in each cell of each single-cell RNA-seq dataset separately for the 8 GWAS tested and described previously (Table dataset). We used the function scDRS "compute-score" with default parameters (--flag-filter-data False).

Cell-cell interaction between ciliated epithelial cells and non-ciliated epithelial cells

To identify potential pairs of interactors between ciliated and non-ciliated epithelial cells we used the dataset of Basnet et al. [35]. We first evaluated which genes were differentially expressed between rhinovirus infected epithelial cells and non-ciliated non-infected epithelial cells. We identified the hypothetical pair of interactors (ligand-receptor) between ciliated and non-ciliated cells with CellPhoneDB (v3.1.0) with the method "deg_analysis", using the list of differentially expressed genes defined before and the default parameter. We filtered the results by identifying the ligand being expressed in ciliated cells and a receptor expressed in non-ciliated cells.

Linking variants to genes

For the Miami plot we used three different and complementary approaches to map variants to genes. First, we used data from Open Target Genetics website [80]. We queried the website to retrieve information for childhood-onset asthma (GCST007995) and adult-onset asthma (GCST007799). We then extracted the L2G gene and the Closest Gene for each variant when information was available. We also added a window of 250kb around each variant with the "bedtools slop" function and retrieved the genes falling in those regions with the "bedtools intersect" function. Finally, we obtained the intersect between this "snp-to-gene" list of genes and the genes upregulated upon rhinovirus infection (genes annotated on **Figure 4A**).

We also retrieved the lead variants from the original paper from Ferreira et al. [38]. We added a window of 250kb around each variant with the "bedtools slop" function and retrieved the genes falling in those regions with the "bedtools intersect" function. We identified the closest gene to the variant by using the function "bedclosest".

Software description (Plots, R, and Biorender)

All the plots were generated with R and graphic schematics were generated with Biorender.

Data availability

Bulk RNA-seq data on dataset with IL-4 and IL-13 activation of airway epithelial cells will be made available in GEO and dbGAP. Other datasets used are publicly available (details in **Supplementary Table 1**).

Acknowledgements

This study was supported by U19AI095219. CO was supported by U19AI162310. NAB was supported by AI134989 and U19AI095219. JG and NAB were supported by AADCRC RNA Sequencing Core for Airway Epithelial Cells. MGA was supported by P30AR070253. We thank Donata Vercelli, Luis Barrera, Peter Nigrovic and his laboratory, and the Gutierrez-Arcelus laboratory for feedback on this study.

Conflict of Interest Statement

JAB has served on scientific advisory boards for Siolta Therapeutics, Third Harmonic Bio, Sanofi/Aventis. NAB has served on scientific advisory boards for Regeneron. The rest of the authors declare that they have no relevant conflicts of interest.

References

1. Porsbjerg, C., Melén, E., Lehtimäki, L., and Shaw, D. (2023). Asthma. Lancet *401*, 858–873.

2. Heijink, I.H., Kuchibhotla, V.N.S., Roffel, M.P., Maes, T., Knight, D.A., Sayers, I., and Nawijn, M.C. (2020). Epithelial cell dysfunction, a major driver of asthma development. Allergy *75*, 1902–1917.

3. Hellings, P.W., and Steelant, B. (2020). Epithelial barriers in allergy and asthma. J. Allergy Clin. Immunol. *145*, 1499–1509.

4. Jackson, D.J., and Gern, J.E. (2022). Rhinovirus Infections and Their Roles in Asthma: Etiology and Exacerbations. J. Allergy Clin. Immunol. Pract. *10*, 673–681.

5. Holgate, S.T., Wenzel, S., Postma, D.S., Weiss, S.T., Renz, H., and Sly, P.D. (2015). Asthma. Nat Rev Dis Primers *1*, 15025.

6. Loxham, M., and Davies, D.E. (2017). Phenotypic and genetic aspects of epithelial barrier function in asthmatic patients. J. Allergy Clin. Immunol. *139*, 1736–1751.

7. Kim, K.W., and Ober, C. (2019). Lessons Learned From GWAS of Asthma. Allergy Asthma Immunol. Res. *11*, 170–187.

8. Vicente, C.T., Revez, J.A., and Ferreira, M.A.R. (2017). Lessons from ten years of genome-wide association studies of asthma. Clin Transl Immunology *6*, e165.

9. Banerjee, S., Webber, C., and Poole, A.R. (1992). The induction of arthritis in mice by the cartilage proteoglycan aggrecan: roles of CD4+ and CD8+ T cells. Cell. Immunol. *144*, 347–357.

10. Kobezda, T., Ghassemi-Nejad, S., Mikecz, K., Glant, T.T., and Szekanecz, Z. (2014). Of mice and men: how animal models advance our understanding of T-cell function in RA. Nat. Rev. Rheumatol. *10*, 160–170.

11. Trynka, G., Sandor, C., Han, B., Xu, H., Stranger, B.E., Liu, X.S., and Raychaudhuri, S. (2013). Chromatin marks identify critical cell types for fine mapping complex trait variants. Nat. Genet. *45*, 124–130.

12. Finucane, H.K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., Anttila, V., Xu, H., Zang, C., Farh, K., et al. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. *47*, 1228–1235.

13. Finucane, H.K., Reshef, Y.A., Anttila, V., Slowikowski, K., Gusev, A., Byrnes, A., Gazal, S., Loh, P.-R., Lareau, C., Shoresh, N., et al. (2018). Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nat. Genet. *50*, 621–629.

14. Gjoneska, E., Pfenning, A.R., Mathys, H., Quon, G., Kundaje, A., Tsai, L.-H., and Kellis, M. (2015). Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. Nature *518*, 365–369.

15. Bettcher, B.M., Tansey, M.G., Dorothée, G., and Heneka, M.T. (2021). Peripheral and central immune system crosstalk in Alzheimer disease - a research prospectus. Nat. Rev. Neurol. *17*, 689–701.

16. Calderon, D., Nguyen, M.L.T., Mezger, A., Kathiria, A., Müller, F., Nguyen, V., Lescano, N., Wu, B., Trombetta, J., Ribado, J.V., et al. (2019). Landscape of stimulation-responsive chromatin across diverse human immune cells. Nat. Genet. *51*, 1494–1505.

17. Zhang, M.J., Hou, K., Dey, K.K., Sakaue, S., Jagadeesh, K.A., Weinand, K., Taychameekiatchai, A., Rao, P., Pisco, A.O., Zou, J., et al. (2022). Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data. Nat. Genet., 1–9.

18. Robinson, D.S., Hamid, Q., Ying, S., Tsicopoulos, A., Barkans, J., Bentley, A.M., Corrigan, C., Durham, S.R., and Kay, A.B. (1992). Predominant TH2-like bronchoalveolar T-lymphocyte population in atopic asthma. N. Engl. J. Med. *326*, 298–304.

19. Fahy, J.V. (2015). Type 2 inflammation in asthma--present in most, absent in many. Nat. Rev. Immunol. *15*, 57–65.

20. Augustine, T., Al-Aghbar, M.A., Al-Kowari, M., Espino-Guarch, M., and van Panhuys, N. (2022). Asthma and the Missing Heritability Problem: Necessity for Multiomics Approaches in Determining Accurate Risk Profiles. Front. Immunol. *13*, 822324.

21. Busse, W.W., and Viswanathan, R. (2022). What has been learned by cytokine targeting of asthma? J. Allergy Clin. Immunol. *150*, 235–249.

22. Mu, Z., Wei, W., Fair, B., Miao, J., Zhu, P., and Li, Y.I. (2021). The impact of cell type and context-dependent regulatory variants on human immune traits. Genome Biol. *22*, 122.

23. Chun, S., Casparino, A., Patsopoulos, N.A., Croteau-Chonka, D.C., Raby, B.A., De Jager, P.L., Sunyaev, S.R., and Cotsapas, C. (2017). Limited statistical evidence for shared genetic effects of eQTLs and autoimmune-disease-associated loci in three major immune-cell types. Nat. Genet. *49*, 600–605.

24. GTEx Consortium, Laboratory, Data Analysis &Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NCI, NIH/NHGRI, NIH/NIMH, NIH/NIDA, Biospecimen Collection Source Site—NDRI, et al. (2017). Genetic effects on gene expression across human tissues. Nature *550*, 204–213.

25. Gutierrez-Arcelus, M., Baglaenko, Y., Arora, J., Hannes, S., Luo, Y., Amariuta, T., Teslovich, N., Rao, D.A., Ermann, J., Jonsson, A.H., et al. (2020). Allele-specific expression changes dynamically during T cell activation in HLA and other autoimmune loci. Nat. Genet. *52*, 247–253.

26. Elorbany, R., Popp, J.M., Rhodes, K., Strober, B.J., Barr, K., Qi, G., Gilad, Y., and Battle, A. (2022). Single-cell sequencing reveals lineage-specific dynamic genetic regulation of gene expression during human cardiomyocyte differentiation. PLoS Genet. *18*, e1009666.

27. Umans, B.D., Battle, A., and Gilad, Y. (2021). Where Are the Disease-Associated eQTLs? Trends Genet. *37*, 109–124.

28. Soskic, B., Cano-Gamez, E., Smyth, D.J., Ambridge, K., Ke, Z., Matte, J.C., Bossini-Castillo, L., Kaplanis, J., Ramirez-Navarro, L., Lorenc, A., et al. (2022). Immune disease risk variants regulate gene expression dynamics during CD4+ T cell activation. Nat. Genet. *54*, 817–826.

29. Soskic, B., Cano-Gamez, E., Smyth, D.J., Rowan, W.C., Nakic, N., Esparza-Gordillo, J., Bossini-Castillo, L., Tough, D.F., Larminie, C.G.C., Bronson, P.G., et al. (2019). Chromatin activity at GWAS loci identifies T cell states driving complex immune diseases. Nat. Genet. *51*, 1486–1493.

30. Pividori, M., Schoettler, N., Nicolae, D.L., Ober, C., and Im, H.K. (2019). Shared and distinct genetic risk factors for childhood-onset and adult-onset asthma: genome-wide and transcriptome-wide studies. Lancet Respir Med *7*, 509–522.

31. Zhu, Z., Lee, P.H., Chaffin, M.D., Chung, W., Loh, P.-R., Lu, Q., Christiani, D.C., and Liang, L. (2018). A genome-wide cross-trait analysis from UK Biobank highlights the shared genetic architecture of asthma and allergic diseases. Nat. Genet. *50*, 857–864.

32. Guibas, G.V., Mathioudakis, A.G., Tsoumani, M., and Tsabouri, S. (2017). Relationship of Allergy with Asthma: There Are More Than the Allergy “Eggs” in the Asthma “Basket.” Front Pediatr *5*, 92.

33. Khan, A.H., Gouia, I., Kamat, S., Johnson, R., Small, M., and Siddall, J. (2023). Prevalence and Severity Distribution of Type 2 Inflammation-Related Comorbidities Among Patients with Asthma, Chronic Rhinosinusitis with Nasal Polyps, and Atopic Dermatitis. Lung *201*, 57–63.

34. Helling, B.A., Sobreira, D.R., Hansen, G.T., Sakabe, N.J., Luo, K., Billstrand, C., Laxman, B., Nicolae, R.I., Nicolae, D.L., Bochkov, Y.A., et al. (2020). Altered transcriptional and chromatin responses to rhinovirus in bronchial epithelial cells from adults with asthma. Commun Biol *3*, 678.

35. Basnet, S., Mohanty, C., Bochkov, Y.A., Brockman-Schneider, R.A., Kendziorski, C., and Gern, J.E. (2023). Rhinovirus C Causes Heterogeneous Infection and Gene Expression in Airway Epithelial Cell Subsets. Mucosal Immunol. 10.1016/j.mucimm.2023.01.008.

36. Bochkov, Y.A., Watters, K., Ashraf, S., Griggs, T.F., Devries, M.K., Jackson, D.J., Palmenberg, A.C., and Gern, J.E. (2015). Cadherin-related family member 3, a childhood asthma susceptibility gene product, mediates rhinovirus C binding and replication. Proc. Natl. Acad. Sci. U. S. A. *112*, 5485–5490.

37. Efremova, M., Vento-Tormo, M., Teichmann, S.A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat. Protoc. *15*, 1484–1506.

38. Ferreira, M.A.R., Mathur, R., Vonk, J.M., Szwajda, A., Brumpton, B., Granell, R., Brew, B.K., Ullemar, V., Lu, Y., Jiang, Y., et al. (2019). Genetic Architectures of Childhood- and Adult-Onset Asthma Are Partly Distinct. Am. J. Hum. Genet. *104*, 665–684.

39. Esparza-Gordillo, J., Weidinger, S., Fölster-Holst, R., Bauerfeind, A., Ruschendorf, F., Patone, G., Rohde, K., Marenholz, I., Schulz, F., Kerscher, T., et al. (2009). A common variant on chromosome 11q13 is associated with atopic dermatitis. Nat. Genet. *41*, 596–601.

40. Ziegler, C.G.K., Miao, V.N., Owings, A.H., Navia, A.W., Tang, Y., Bromley, J.D., Lotfy, P., Sloan, M., Laird, H., Williams, H.B., et al. (2021). Impaired local intrinsic immunity to SARS-CoV-2 infection in severe COVID-19. Cell *184*, 4713-4733.e22.

41. Lambrecht, B.N., Hammad, H., and Fahy, J.V. (2019). The Cytokines of Asthma. Immunity *50*, 975–991.

42. Hammad, H., and Lambrecht, B.N. (2021). The basic immunology of asthma. Cell *184*, 1469–1485.

43. Camiolo, M.J., Zhou, X., Oriss, T.B., Yan, Q., Gorry, M., Horne, W., Trudeau, J.B., Scholl, K., Chen, W., Kolls, J.K., et al. (2021). High-dimensional profiling clusters asthma severity by lymphoid and non-lymphoid status. Cell Rep. *35*, 108974.

44. Woodruff, P.G., Boushey, H.A., Dolganov, G.M., Barker, C.S., Yang, Y.H., Donnelly, S., Ellwanger, A., Sidhu, S.S., Dao-Pick, T.P., Pantoja, C., et al. (2007). Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. Proc. Natl. Acad. Sci. U. S. A. *104*, 15858–15863.

45. Woodruff, P.G., Modrek, B., Choy, D.F., Jia, G., Abbas, A.R., Ellwanger, A., Koth, L.L., Arron, J.R., and Fahy, J.V. (2009). T-helper type 2-driven inflammation defines major subphenotypes of asthma. Am. J. Respir. Crit. Care Med. *180*, 388–395.

46. Krishnamoorthy, N., Douda, D.N., Brüggemann, T.R., Ricklefs, I., Duvall, M.G., Abdulnour, R.-E.E., Martinod, K., Tavares, L., Wang, X., Cernadas, M., et al. (2018). Neutrophil cytoplasts induce TH17

differentiation and skew inflammation toward neutrophilia in severe asthma. Sci Immunol *3*. 10.1126/sciimmunol.aao4747.

47. Castro, M., Corren, J., Pavord, I.D., Maspero, J., Wenzel, S., Rabe, K.F., Busse, W.W., Ford, L., Sher, L., FitzGerald, J.M., et al. (2018). Dupilumab Efficacy and Safety in Moderate-to-Severe Uncontrolled Asthma. N. Engl. J. Med. *378*, 2486–2496.

48. Koh, K.D., Bonser, L.R., Eckalbar, W.L., Yizhar-Barnea, O., Shen, J., Zeng, X., Hargett, K.L., Sun, D.I., Zlock, L.T., Finkbeiner, W.E., et al. (2023). Genomic characterization and therapeutic utilization of IL-13-responsive sequences in asthma. Cell Genom *3*, 100229.

49. Rubner, F.J., Jackson, D.J., Evans, M.D., Gangnon, R.E., Tisler, C.J., Pappas, T.E., Gern, J.E., and Lemanske, R.F., Jr (2017). Early life rhinovirus wheezing, allergic sensitization, and asthma risk at adolescence. J. Allergy Clin. Immunol. *139*, 501–507.

50. Jartti, T., and Gern, J.E. (2011). Rhinovirus-associated wheeze during infancy and asthma development. Curr. Respir. Med. Rev. *7*, 160–166.

51. Bønnelykke, K., Sleiman, P., Nielsen, K., Kreiner-Møller, E., Mercader, J.M., Belgrave, D., den Dekker, H.T., Husby, A., Sevelsted, A., Faura-Tellez, G., et al. (2013). A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. Nat. Genet. *46*, 51–55.

52. Calışkan, M., Bochkov, Y.A., Kreiner-Møller, E., Bønnelykke, K., Stein, M.M., Du, G., Bisgaard, H., Jackson, D.J., Gern, J.E., Lemanske, R.F., Jr, et al. (2013). Rhinovirus wheezing illness and genetic risk of childhood-onset asthma. N. Engl. J. Med. *368*, 1398–1407.

53. Basnet, S., Bochkov, Y.A., Brockman-Schneider, R.A., Kuipers, I., Aesif, S.W., Jackson, D.J., Lemanske, R.F., Jr, Ober, C., Palmenberg, A.C., and Gern, J.E. (2019). CDHR3 Asthma-Risk Genotype Affects Susceptibility of Airway Epithelium to Rhinovirus C Infections. Am. J. Respir. Cell Mol. Biol. *61*, 450–458.

54. Griggs, T.F., Bochkov, Y.A., Basnet, S., Pasic, T.R., Brockman-Schneider, R.A., Palmenberg, A.C., and Gern, J.E. (2017). Rhinovirus C targets ciliated airway epithelial cells. Respir. Res. *18*, 84.

55. Zhou, X., Zhu, L., Lizarraga, R., and Chen, Y. (2017). Human Airway Epithelial Cells Direct Significant Rhinovirus Replication in Monocytic Cells by Enhancing ICAM1 Expression. Am. J. Respir. Cell Mol. Biol. *57*, 216–225.

56. Hammond, C., Kurten, M., and Kennedy, J.L. (2015). Rhinovirus and asthma: a storied history of incompatibility. Curr. Allergy Asthma Rep. *15*, 502.

57. Veerapandian, R., Snyder, J.D., and Samarasinghe, A.E. (2018). Influenza in Asthmatics: For Better or for Worse? Front. Immunol. *9*, 1843.

58. Hasegawa, S., Hirano, R., Hashimoto, K., Haneda, Y., Shirabe, K., and Ichiyama, T. (2011). Characteristics of atopic children with pandemic H1N1 influenza viral infection: pandemic H1N1 influenza reveals “occult” asthma of childhood. Pediatr. Allergy Immunol. *22*, e119-23.

59. Obuchi, M., Adachi, Y., Takizawa, T., and Sata, T. (2013). Influenza A(H1N1)pdm09 virus and asthma. Front. Microbiol. *4*, 307.

60. Sunjaya, A.P., Allida, S.M., Di Tanna, G.L., and Jenkins, C. (2022). Asthma and risk of infection, hospitalization, ICU admission and mortality from COVID-19: Systematic review and meta-analysis. J. Asthma *59*, 866–879.

61. Jackson, D.J., Busse, W.W., Bacharier, L.B., Kattan, M., O'Connor, G.T., Wood, R.A., Visness, C.M., Durham, S.R., Larson, D., Esnault, S., et al. (2020). Association of respiratory allergy, asthma, and expression of the SARS-CoV-2 receptor ACE2. J. Allergy Clin. Immunol. *146*, 203-206.e3.

62. Ochoa, D., Karim, M., Ghoussaini, M., Hulcoop, D.G., McDonagh, E.M., and Dunham, I. (2022). Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs. Nat. Rev. Drug Discov. *21*, 551.

63. Mostafavi, H., Spence, J.P., Naqvi, S., and Pritchard, J.K. (2023). Systematic differences in discovery of genetic effects on gene expression and complex traits. Nat. Genet. *55*, 1866–1875.

64. Stone, C.A., Jr, and Miller, E.K. (2016). Understanding the Association of Human Rhinovirus with Asthma. Clin. Vaccine Immunol. *23*, 6–10.

65. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. *12*, e1001779.

66. Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., Raychaudhuri, S., et al. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature *467*, 832–838.

67. Lambert, J.C., Ibrahim-Verbaas, C.A., Harold, D., Naj, A.C., Sims, R., Bellenguez, C., DeStafano, A.L., Bis, J.C., Beecham, G.W., Grenier-Boley, B., et al. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat. Genet. *45*, 1452–1458.

68. Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., Yoshida, S., et al. (2014). Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature *506*, 376–381.

69. Gutierrez-Arcelus, M., Teslovich, N., Mola, A.R., Polidoro, R.B., Nathan, A., Kim, H., Hannes, S., Slowikowski, K., Watts, G.F.M., Korsunsky, I., et al. (2019). Lymphocyte innateness defined by transcriptional states reflects a balance between proliferation and effector functions. Nat. Commun. *10*, 687.

70. Wang, W., Xu, Y., Wang, L., Zhu, Z., Aodeng, S., Chen, H., Cai, M., Huang, Z., Han, J., Wang, L., et al. (2022). Single-cell profiling identifies mechanisms of inflammatory heterogeneity in chronic rhinosinusitis. Nat. Immunol. *23*, 1484–1494.

71. Seumois, G., Ramírez-Suástegui, C., Schmiedel, B.J., Liang, S., Peters, B., Sette, A., and Vijayanand, P. (2020). Single-cell transcriptomic analysis of allergen-specific T cells in allergy and asthma. Sci Immunol *5*. 10.1126/sciimmunol.aba6087.

72. Ravindra, N.G., Alfajaro, M.M., Gasque, V., Huston, N.C., Wan, H., Szigeti-Buck, K., Yasumoto, Y., Greaney, A.M., Habet, V., Chow, R.D., et al. (2021). Single-cell longitudinal analysis of SARS-CoV-2 infection in human airway epithelium identifies target cells, alterations in gene expression, and cell state changes. PLoS Biol. *19*, e3001143.

73. Frankish, A., Diekhans, M., Ferreira, A.-M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res. *47*, D766–D773.

74. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296.

75. Becht, E., Giraldo, N.A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., Selves, J., Laurent-Puig, P., Sautès-Fridman, C., Fridman, W.H., et al. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. *17*, 218.

76. Hewitt, R.J., and Lloyd, C.M. (2021). Regulation of immune responses by the airway epithelial cell landscape. Nat. Rev. Immunol. *21*, 347–362.

77. Ordovas-Montanes, J., Dwyer, D.F., Nyquist, S.K., Buchheit, K.M., Vukovic, M., Deb, C., Wadsworth, M.H., 2nd, Hughes, T.K., Kazer, S.W., Yoshimoto, E., et al. (2018). Allergic inflammatory memory in human respiratory epithelial progenitor cells. Nature *560*, 649–654.

78. Gazal, S., Finucane, H.K., Furlotte, N.A., Loh, P.-R., Palamara, P.F., Liu, X., Schoech, A., Bulik-Sullivan, B., Neale, B.M., Gusev, A., et al. (2017). Linkage disequilibrium–dependent architecture of human complex traits shows action of negative selection. Nat. Genet. *49*, 1421–1427.

79. de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput. Biol. *11*, e1004219.

80. Ghoussaini, M., Mountjoy, E., Carmona, M., Peat, G., Schmidt, E.M., Hercules, A., Fumis, L., Miranda, A., Carvalho-Silva, D., Buniello, A., et al. (2021). Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. Nucleic Acids Res. *49*, D1311–D1320.

Graphical abstract

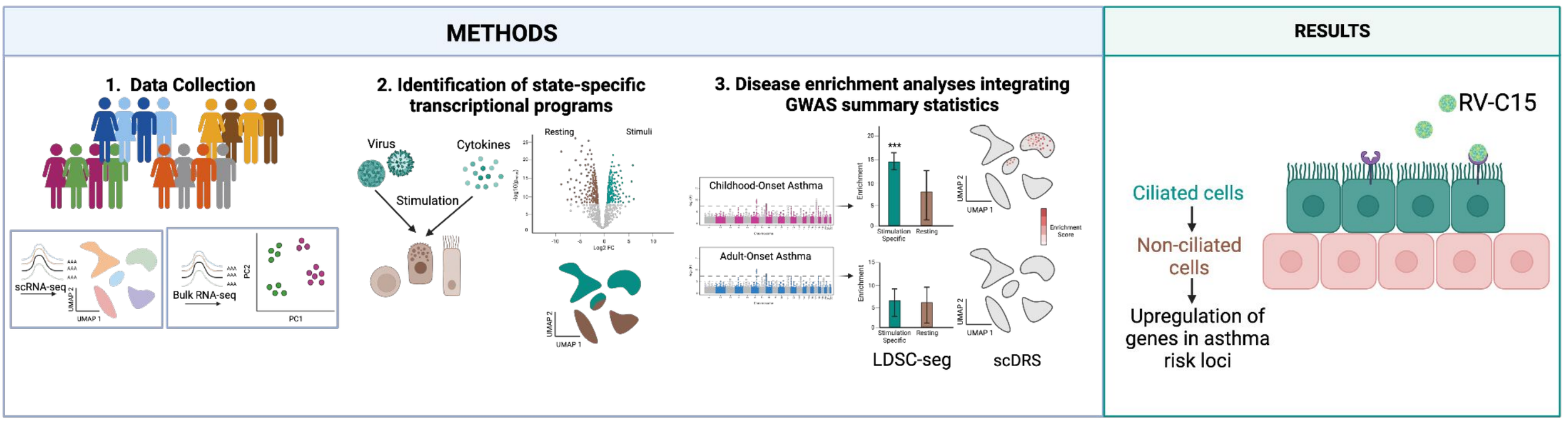


Figure 1

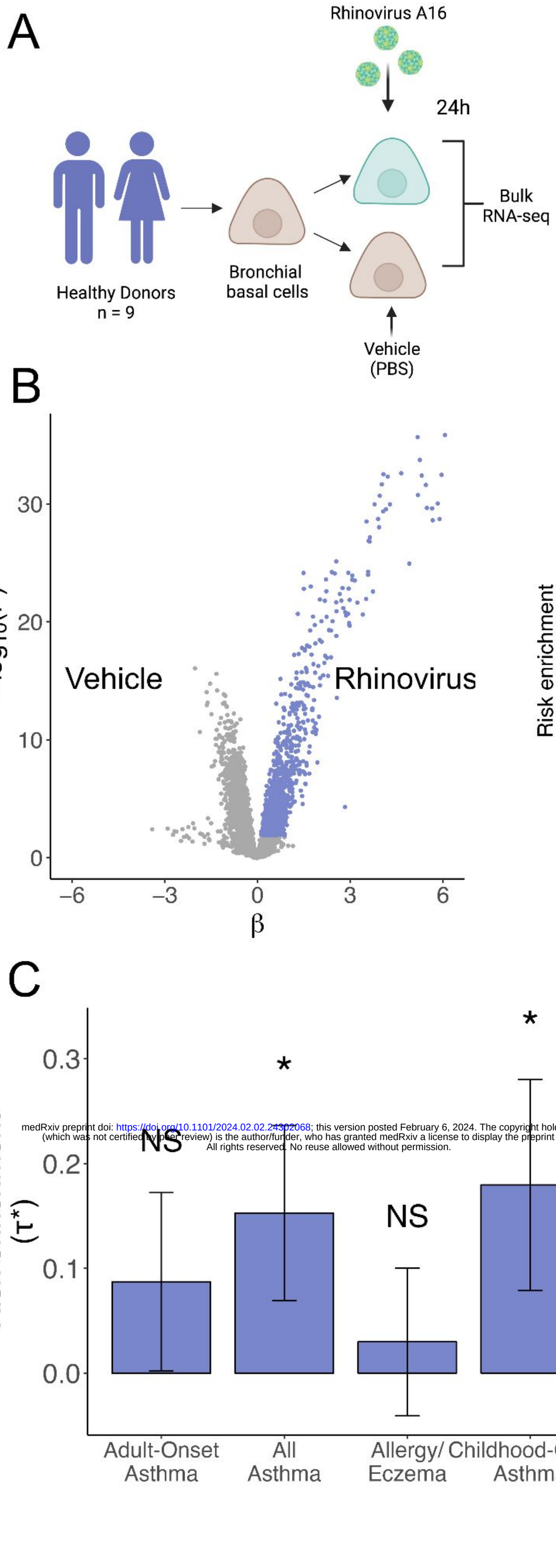


Figure 1. Bronchial epithelial cells infected with rhinovirus upregulate genes associated with asthma susceptibility.

(A) Experimental design of the *Helling et al.* dataset consisting of basal bronchial epithelial cells from healthy donors stimulated with PBS or RV-A16.

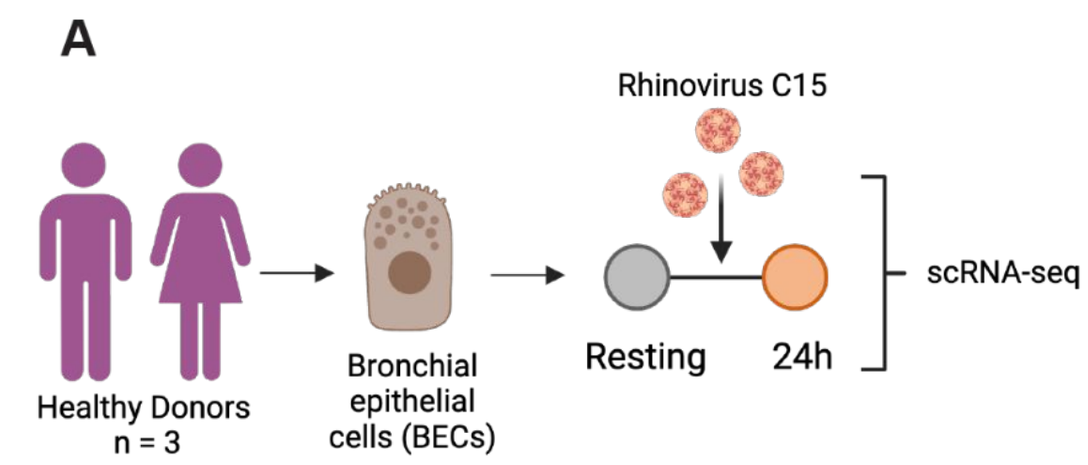
(B) Volcano plot showing differentially expressed genes after rhinovirus infection, genes selected based on t-statistic are colored in purple.

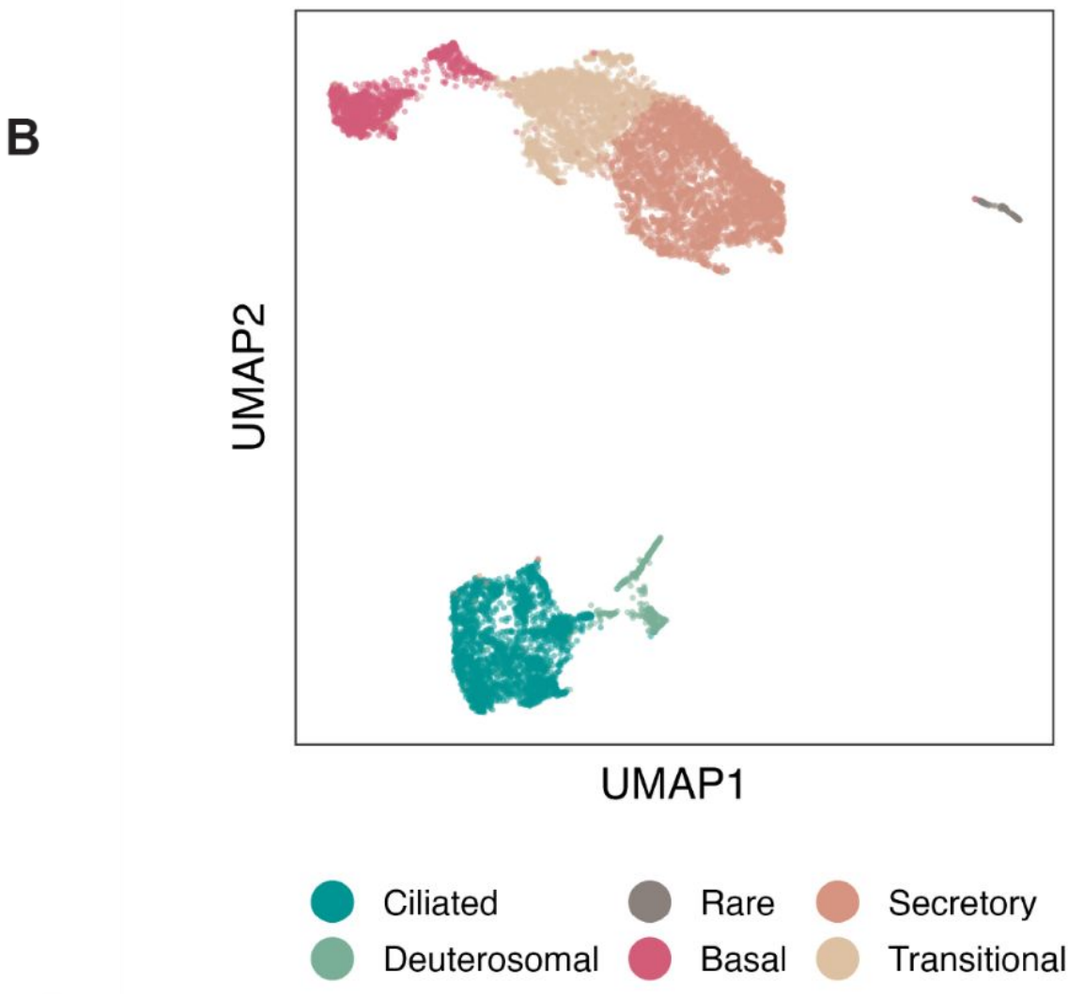
(C) Bar plot showing LDSC-SEG heritability enrichment coefficient (τ*) for each of the asthma-associated GWAS studies. Error bars represent τ* +/- standard error. Asterisk denotes $P < 0.05$ and NS denotes nonsignificant ($P > 0.05$).

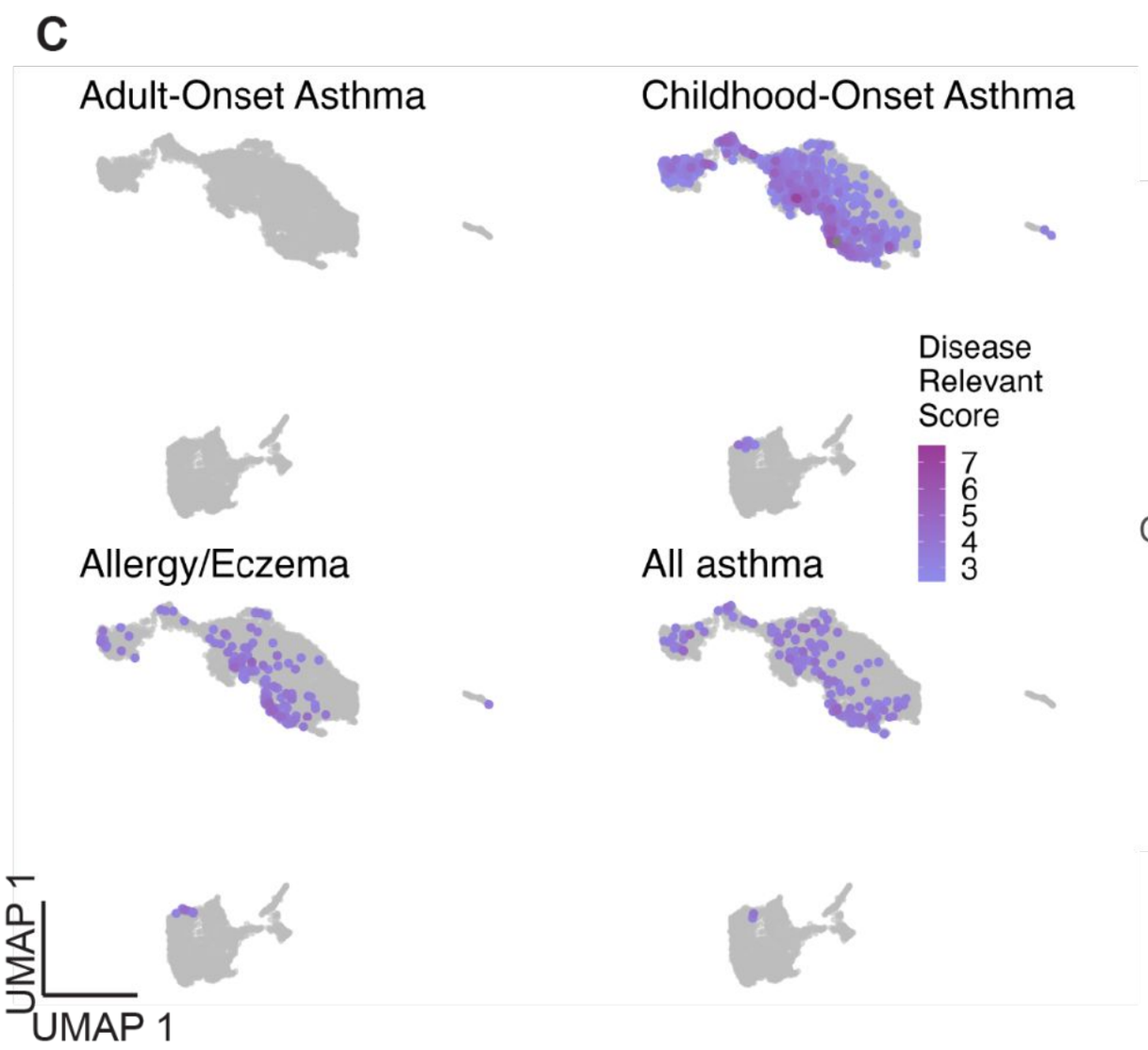
(D) Experimental design of the Basnet et al. time course dataset consisting of BECs stimulated with RV-C15.

(E) Bar plot showing LDSC-SEG heritability enrichment coefficient (t*) for differentially expressed genes at each time point when compared against all others. Error bars represent τ* +/- standard error. Asterisks denote significance as * $P < 0.05$, *** Bonferroni-adjusted $P < 0.05$ and NS denotes nonsignificant ($P > 0.05$).

Figure 2







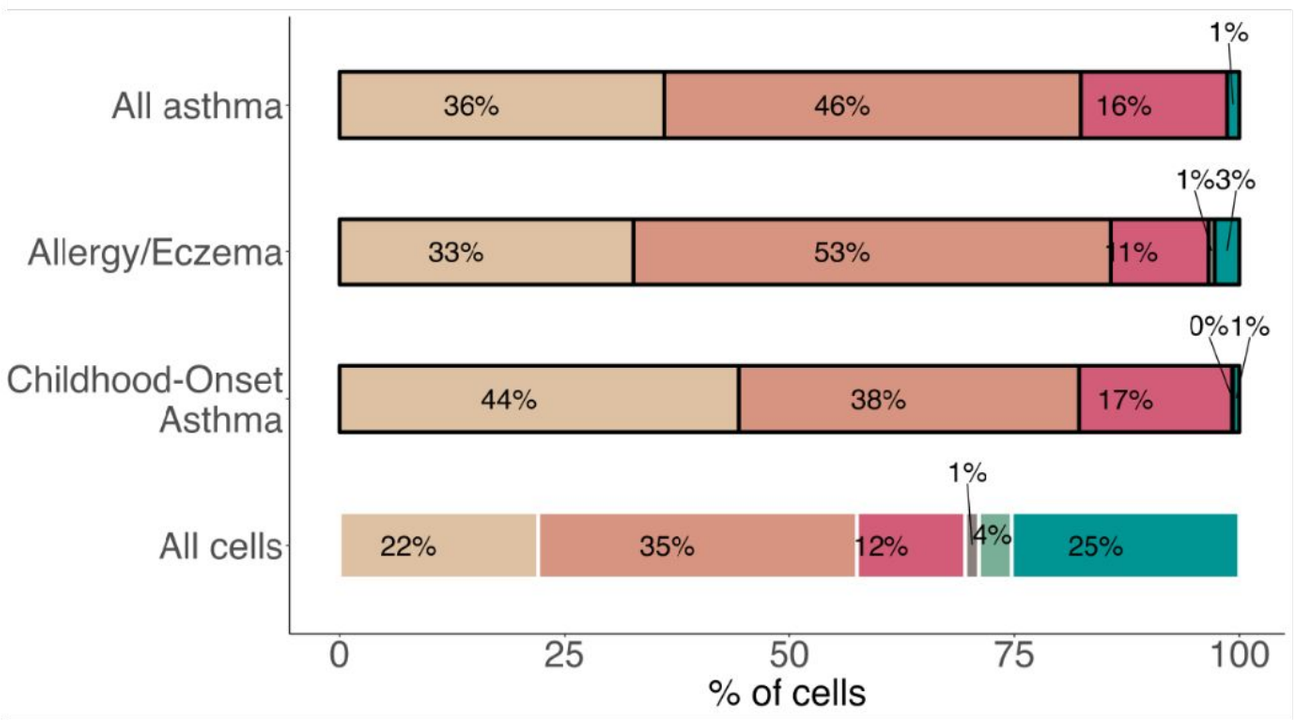


Figure 2. scRNA-seq uncovers non-ciliated epithelial cells as potential mediators for asthma risk.

(A) Experimental design of the *Basnet et al.* scRNA-seq dataset of BECs from healthy donors infected with RV-C15.

(B) UMAP visualization of the 10,721 airway epithelial cells colored by cell type.

(C) scDRS results represented on the UMAP for the 4 asthma-associated GWAS tested. The intensity of the color represents the disease relevant score, the lighter purple represents a less intense score whereas a more intense purple represents cells associated with a stronger score. nonsignificant cells with a FDR higher than 10% are depicted in gray.

(D) Bar plot representing the percentage of each cell type in all cells followed by the significant cells at 10% FDR for scDRS in COA, Allergy/Eczema, All asthma. AOA is not represented on this barplot because no cells were significant.

Figure 3

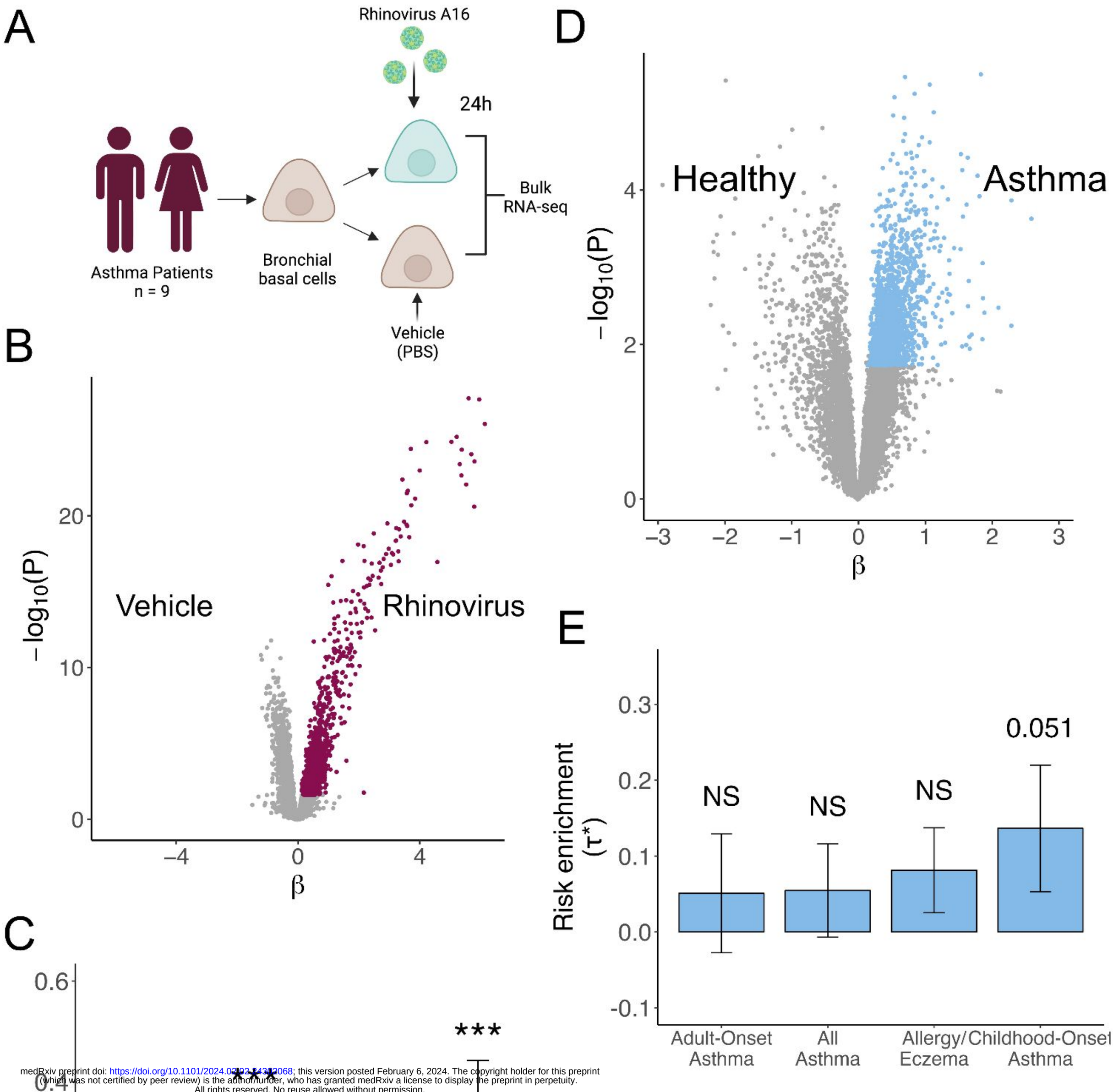


Figure 3. RV-infected BECs from asthma patients showed a stronger enrichment for asthma risk compared to those of healthy individuals.

(A) Experimental design of *Helling et al.* dataset consisting of BECs from asthma patients stimulated with PBS or RV-A16.

(B) Volcano plot showing differentially expressed genes after rhinovirus infection in patient samples. Genes upregulated after infection were selected based on t-statistic and are colored in burgundy.

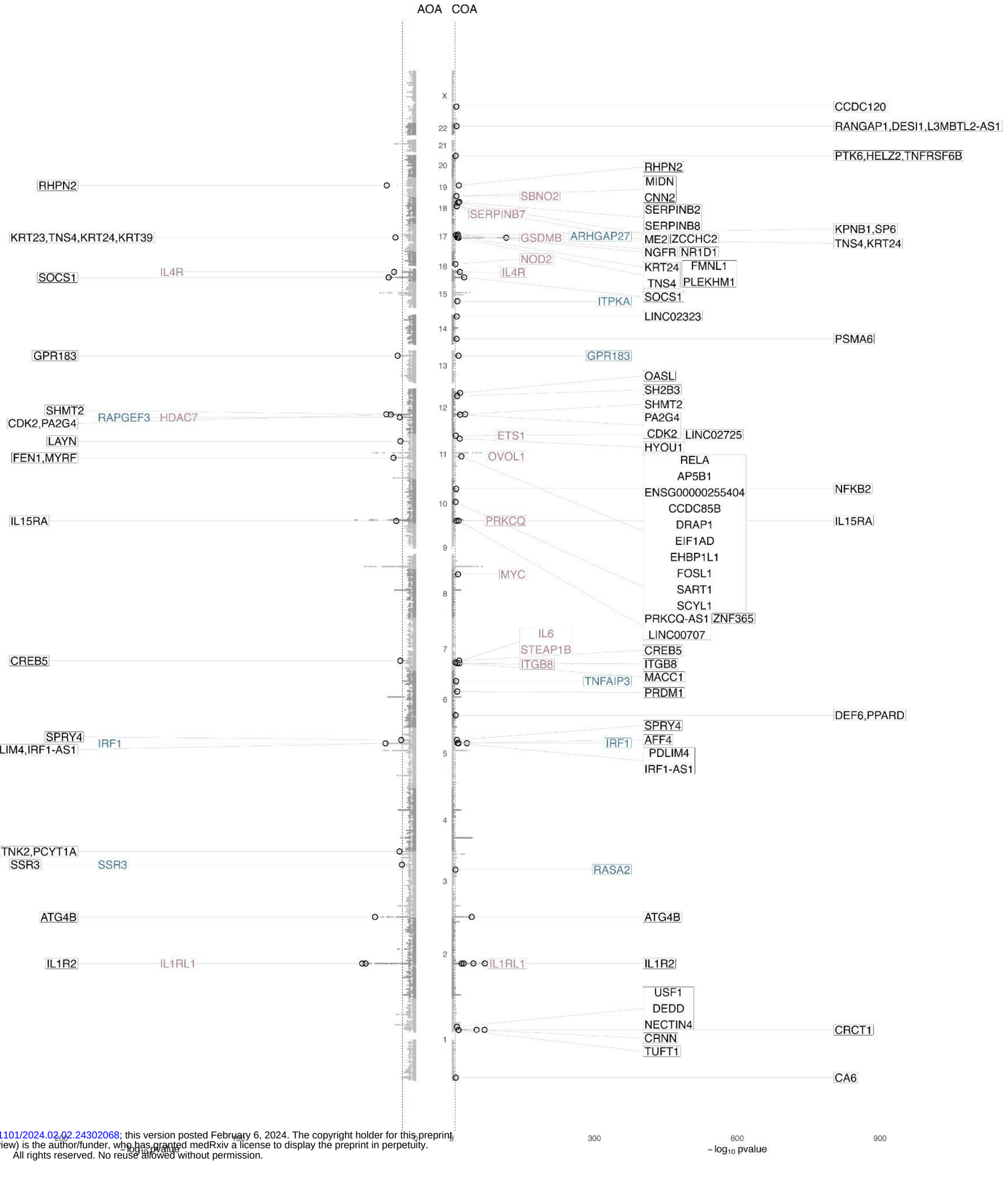
(C) Bar plot showing LDSC-SEG heritability enrichment coefficient (τ*) across GWAS studies. Error bars represent τ* +/- standard error. Asterisks denote significance as *** Bonferroni-adjusted $P < 0.05$ and NS denotes nonsignificant ($P > 0.05$).

(D) Volcano plot showing differentially expressed genes between asthma patients and healthy donors. Genes upregulated in patients were selected based on t-statistic and are colored in blue.

(E) Bar plot of LDSC-SEG heritability enrichment coefficient (τ*), suggestively significant enrichment is labeled. Error bars represent τ* +/- standard error. NS denotes nonsignificant ($P > 0.05$).

Figure 4

A



B

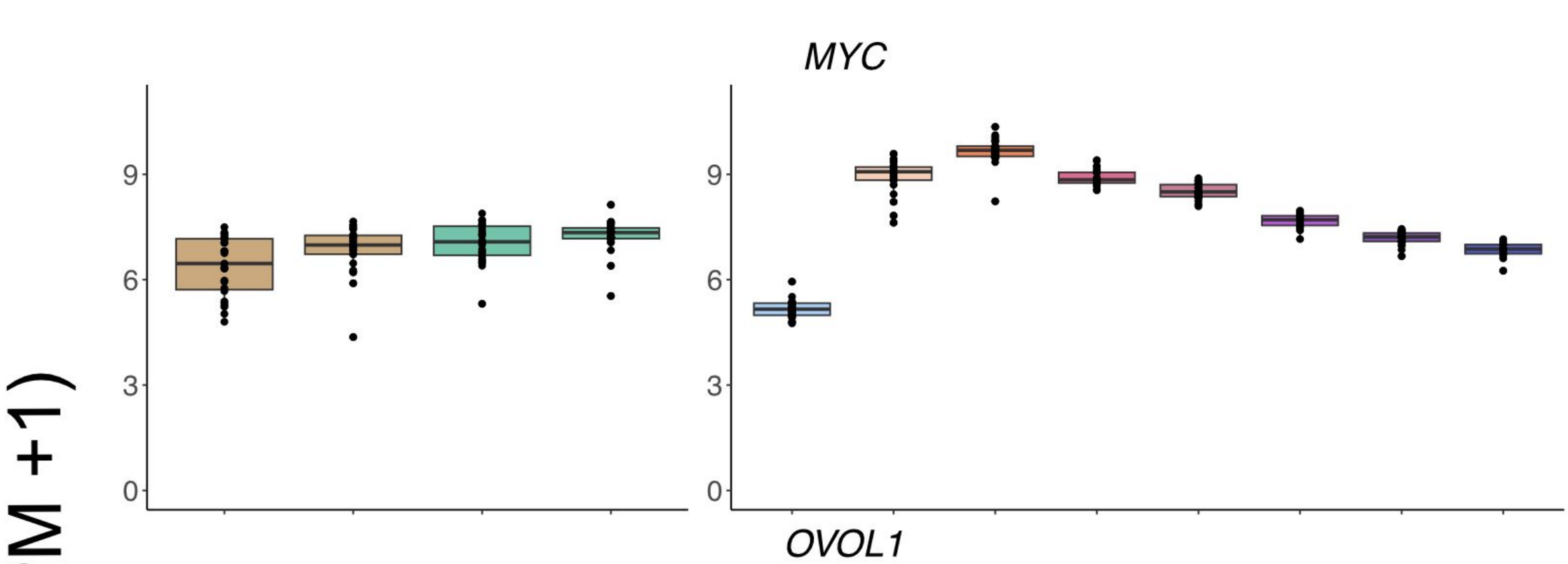


Figure 4. Rhinovirus induced genes that are found in COA and AOA GWAS.

(A) Miami plot of COA and AOA GWAS. Each gray dot shows the SNP found in the *Ferreira et al.*, GWAS. The black circles represent SNPs being found as lead variants in either Open Targets Genetics or in Ferreira study. Highlighted genes are upregulated upon rhinovirus infection, in purple the genes being L2G genes, in blue the closest one to the transcription start site of the variant (Open Target Genetics), and in black genes found in a window of 250kb around the SNP. The blue dashed line represents the P-value threshold of -log10($5x10^{-8}$).

(B) Box plots depicting gene expression levels for *MYC* (top) and *OVOL1* (bottom). Left panel shows gene expression in epithelial cells from the *Helling et al.* dataset; samples infected with RV-A16 are shown in green and non-infected samples are in brown. Right panel shows gene expression from *Gutierrez-Arcelus et al.* in activated CD4 memory T cells; colors depict time points after activation. In each plot every point represents an individual sample.

Figure 5

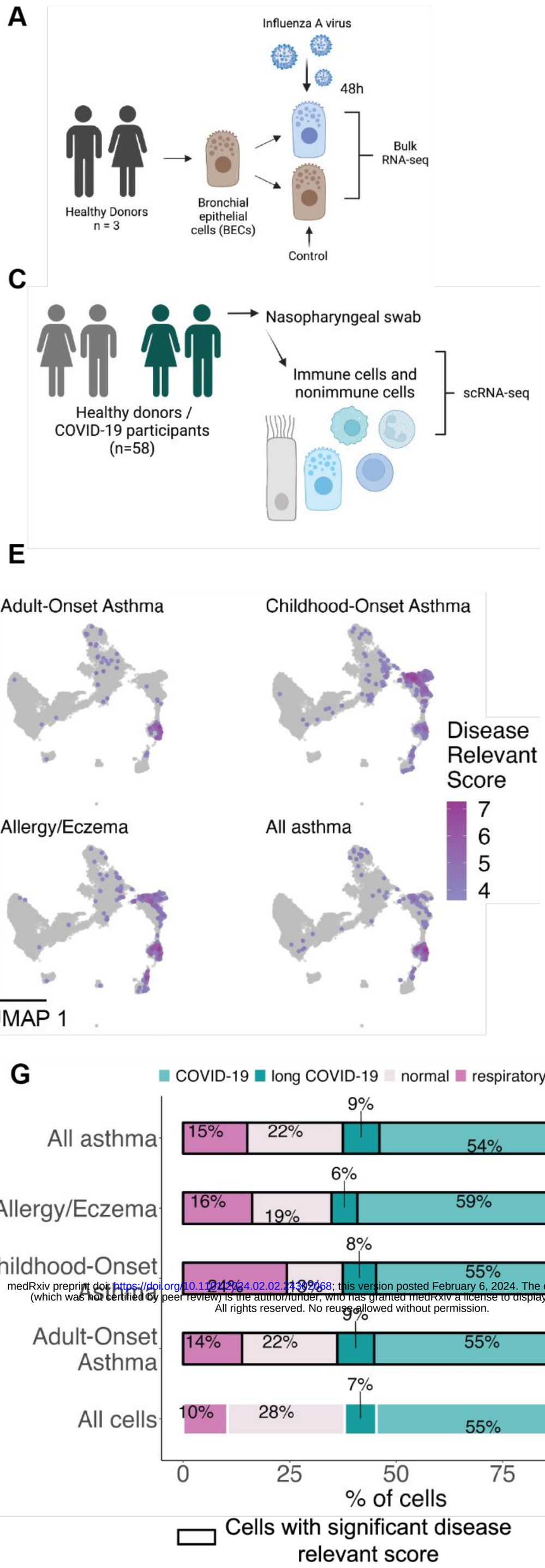


Figure 5. Influence from other viruses on asthma genetics.

(A) Experimental design of *Tao et al.* bulk RNA-seq dataset.

(B) Bar plot representing LDSC-SEG enrichment. Error bars represent τ* +/- standard error. Asterisk denotes significance as * P < 0.05.

(C) Experimental design of *Ziegler et al.* single-cell RNA-seq dataset of epithelial cells infected with SARS-CoV-2 or not.

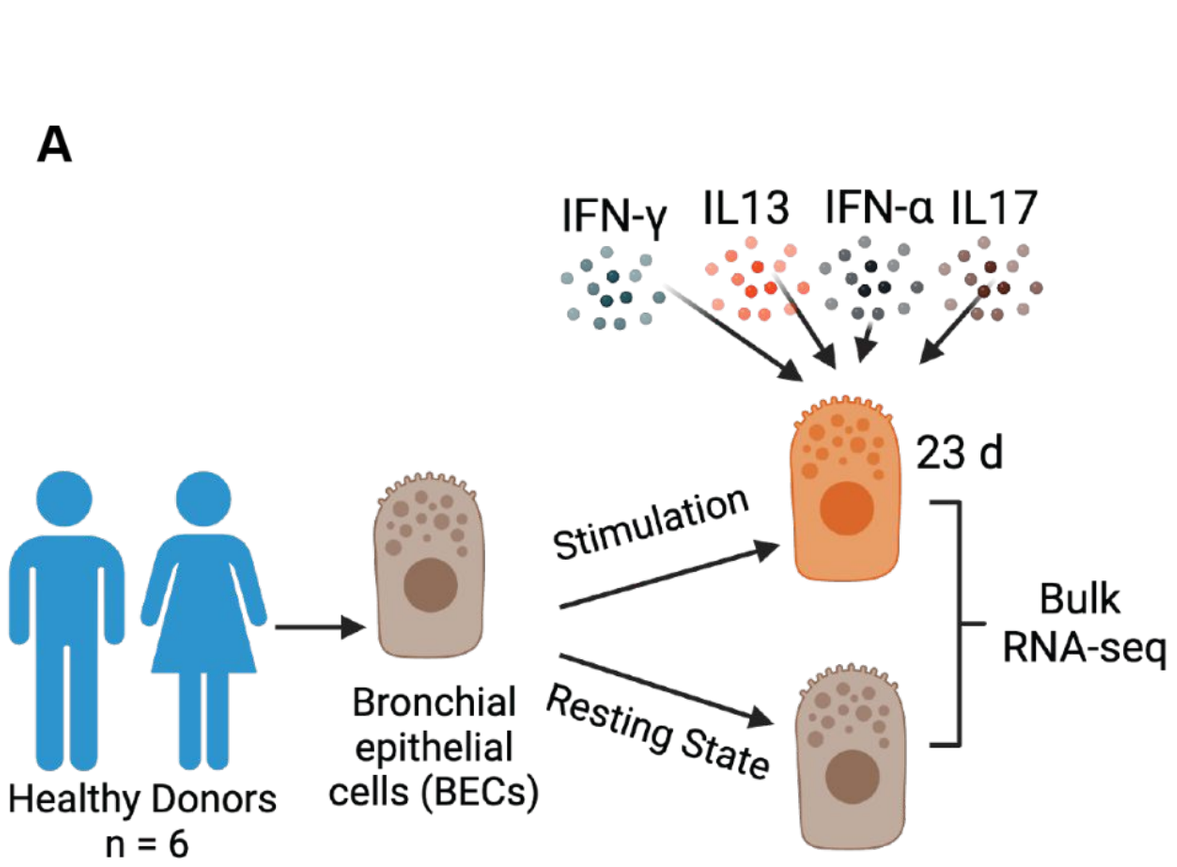
(D) UMAP visualization of the 32,588 cells colored by cell type.

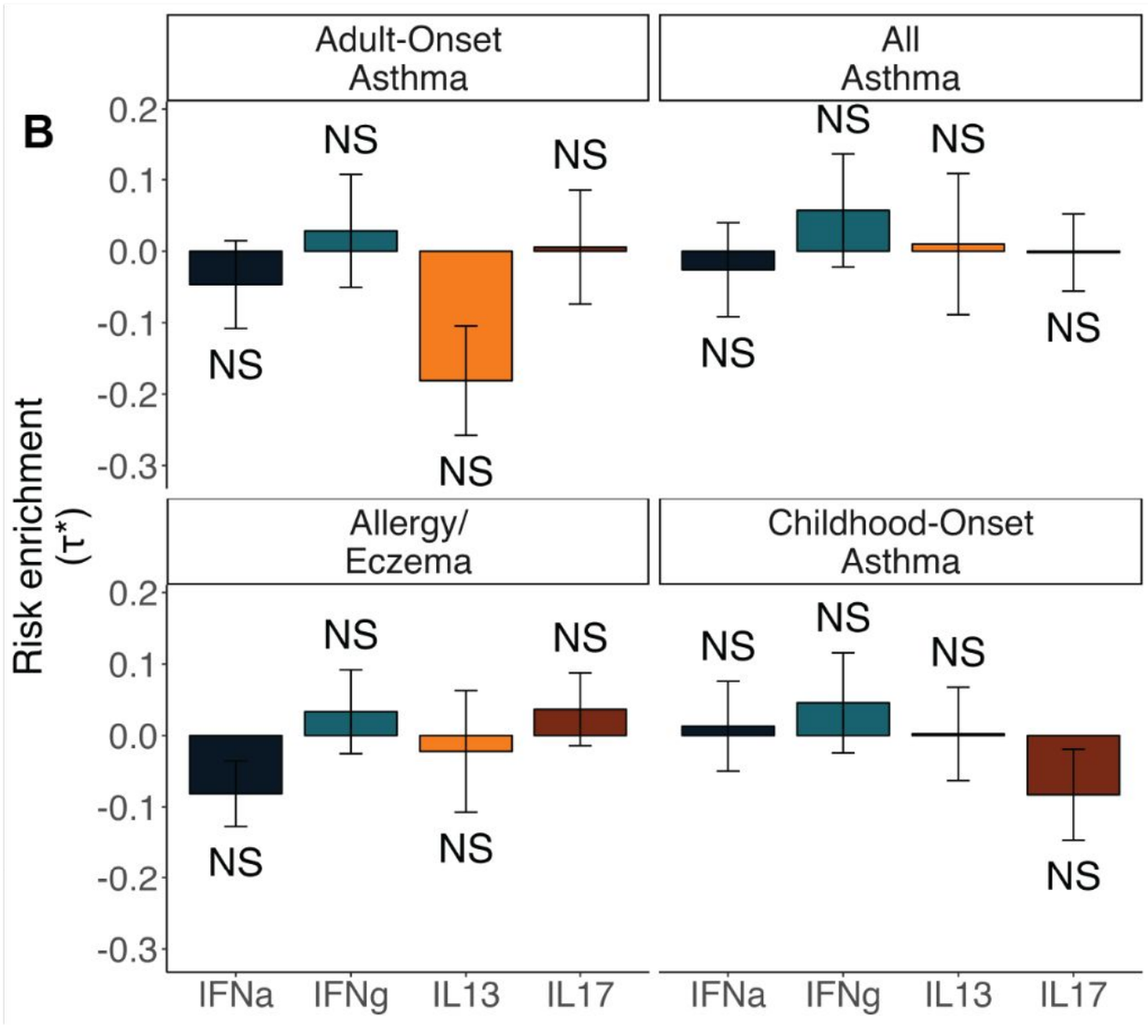
(E) scDRS results represented on the UMAP for the 4 asthma-associated GWAS tested. The intensity of the color represents the disease relevant score, the lighter purple represents a less intense score whereas a more intense purple represents cells associated with a stronger score. nonsignificant cells with a FDR higher than 10% are depicted in gray.

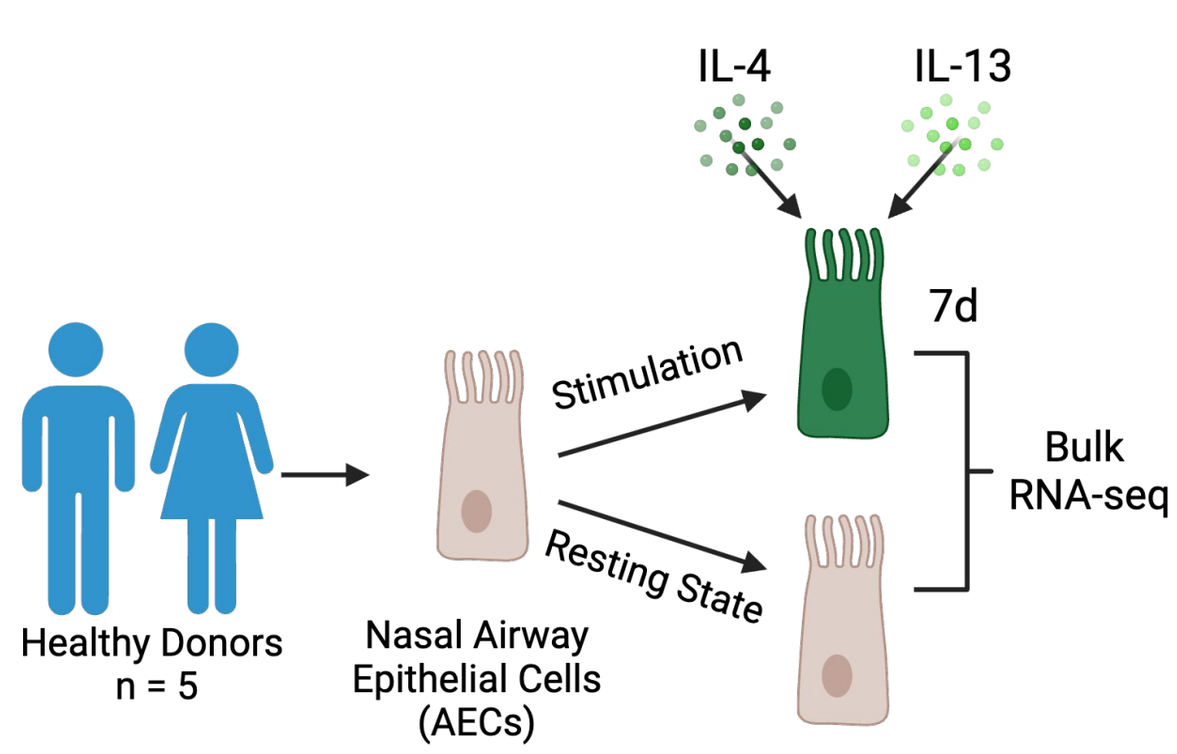
(F) Bar plot representing the percentage of each cell type in all cells followed by the significant cells at 10% FDR for scDRS in AOA, COA, Allergy/Eczema and All asthma.

(G) Bar plot representing the percentage of cells in the full dataset coming from patients grouped by COVID severity categories (COVID-19, long COVID-19, respiratory failure) or from healthy donors (normal). Followed by percentage of cells passing significance at 10% FDR for AOA, COA, Allergy/Eczema, All asthma, by disease severity category or control.

Figure 6







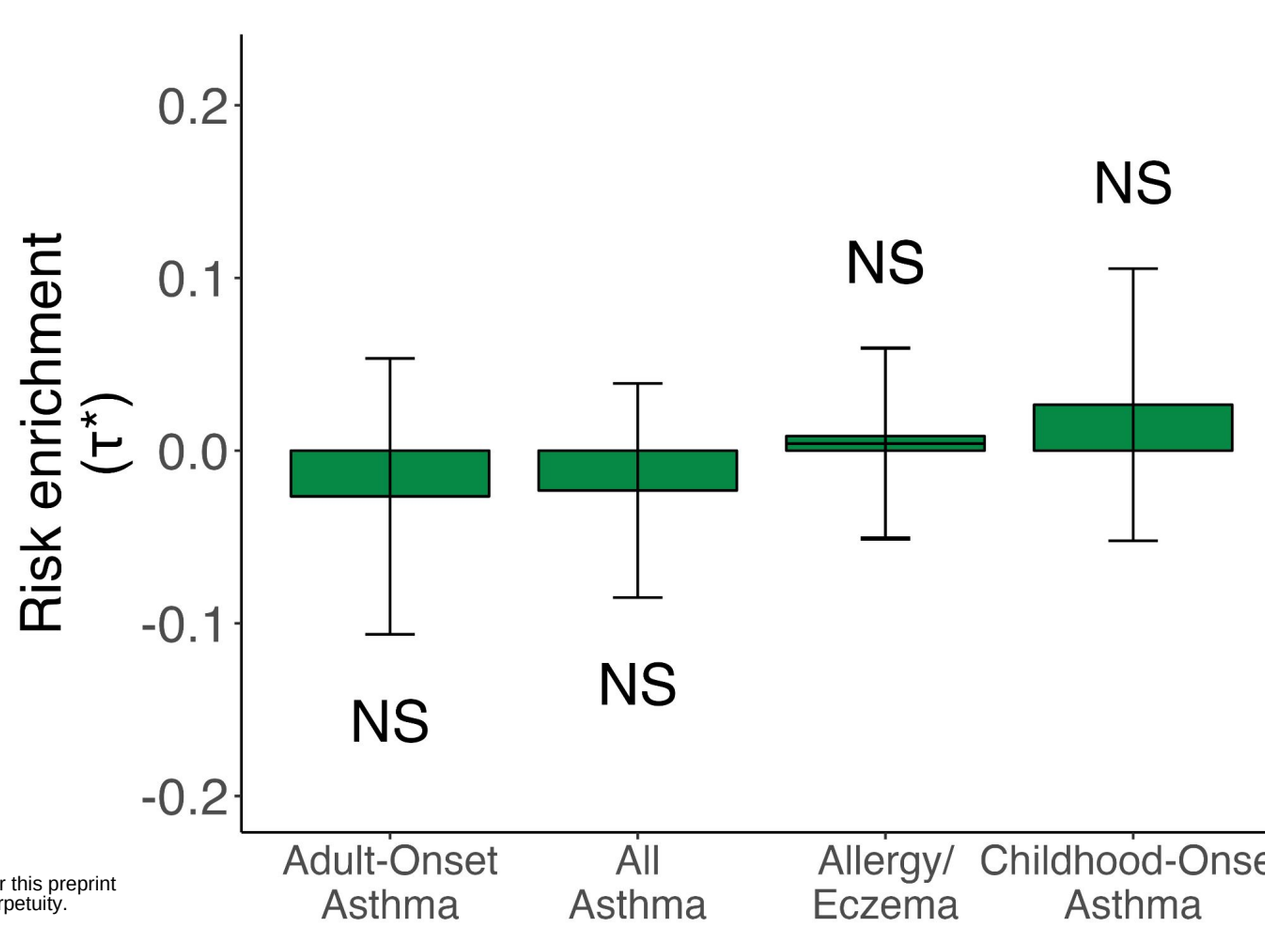


Figure 6. Cytokines impact on asthma-associated genes.

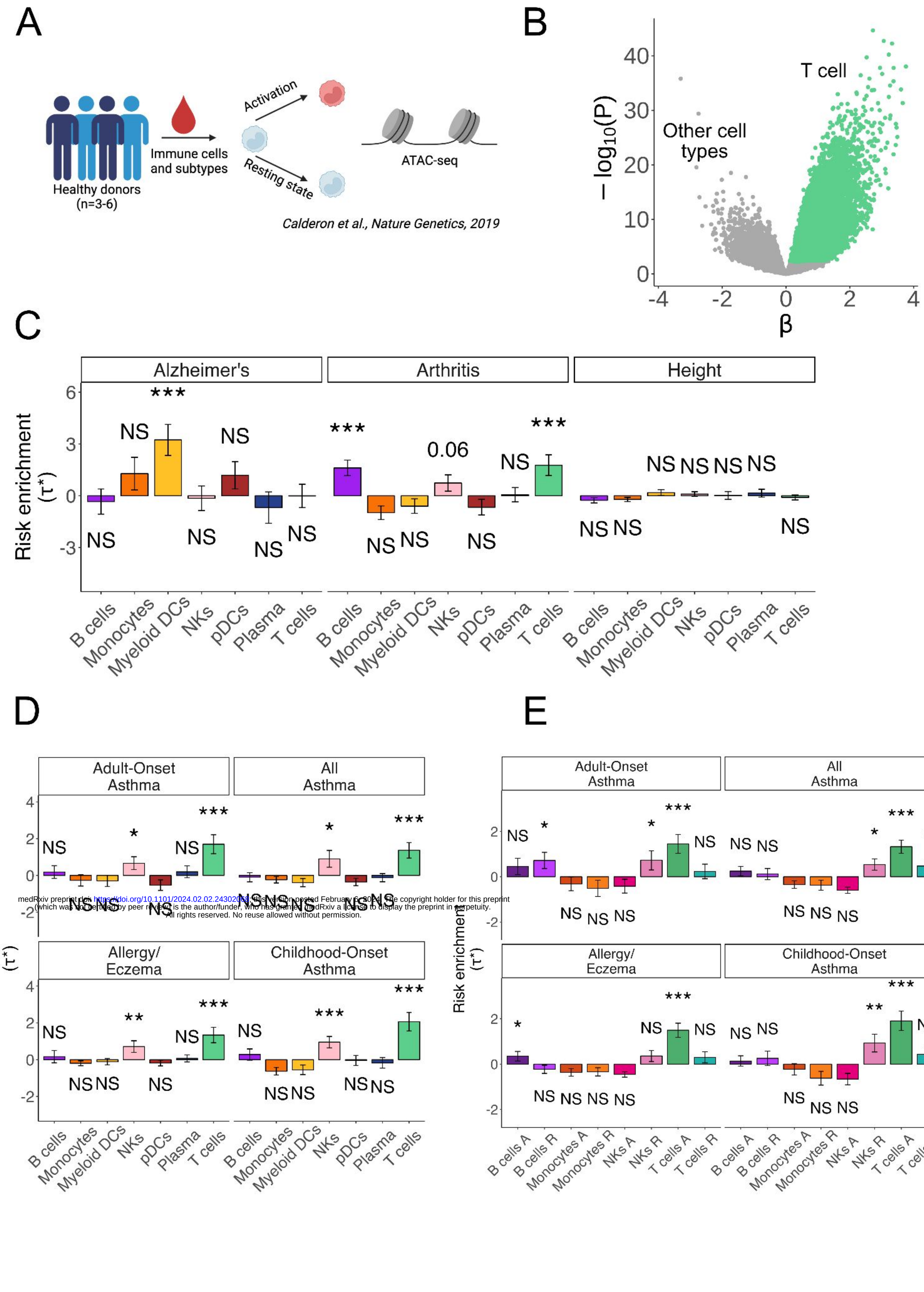
(A) Experimental design of the *Koh et al.* bulk RNA-seq dataset of BECs from healthy donors that were stimulated or not with different cytokines (IFNγ, IFNα, IL-13, IL-17).

(B) Bar plot representing LDSC-SEG heritability enrichment coefficient (τ*) for genes up-regulated in each stimuli against all others, for each of the asthma-associated GWAS.Error bars represent τ* +/- standard error. NS denotes nonsignificant (P > 0.05).

(C) Experimental design of bulk RNA-seq of AECs from healthy donors co-stimulated with IL-4 and IL-13.

(D) Bar plot showing LDSC-SEG enrichment for each of the asthma-associated GWAS. Error bars represent τ* +/- standard error. NS denotes nonsignificant (P > 0.05)

Supplementary Figure 1



Supplementary Figure 1. Validation of T cells enrichment in asthma-associated loci using differentially accessible peaks.

(A) Experimental design of the *Calderon et al.* ATAC-seq dataset of immune cell types that were activating or not *in vitro* .

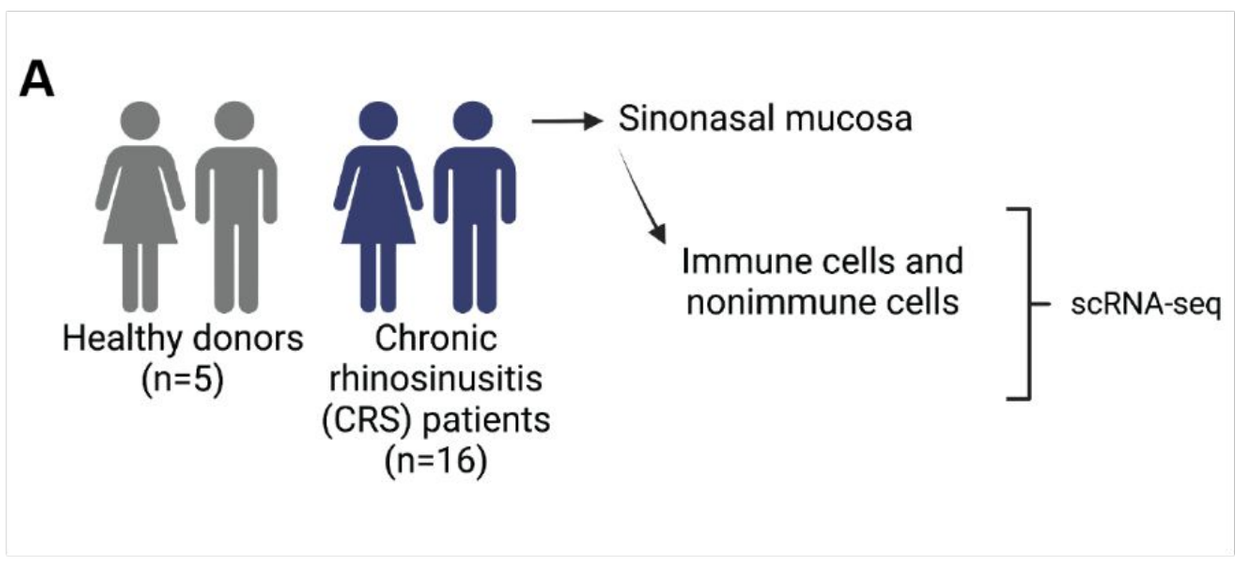
(B) Volcano plot showing differentially expressed peaks between T cells and all other cell types. Peaks upregulated in T cells were selected based on t-statistic and are colored in green.

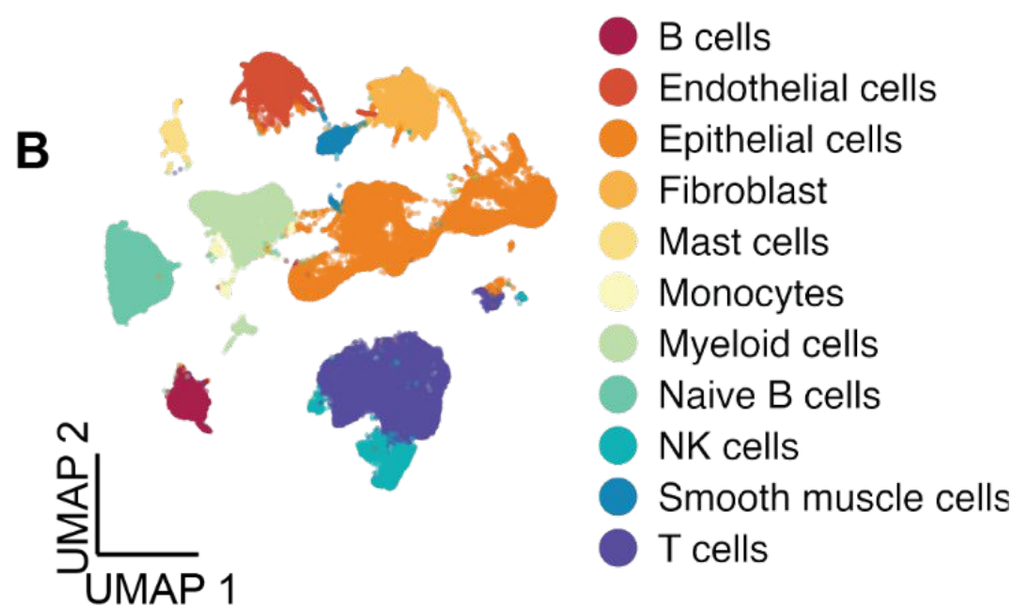
(C) Bar plots representing LDSC-SEG heritability enrichment coefficient (τ*) for each cell type for the 3 control traits tested.

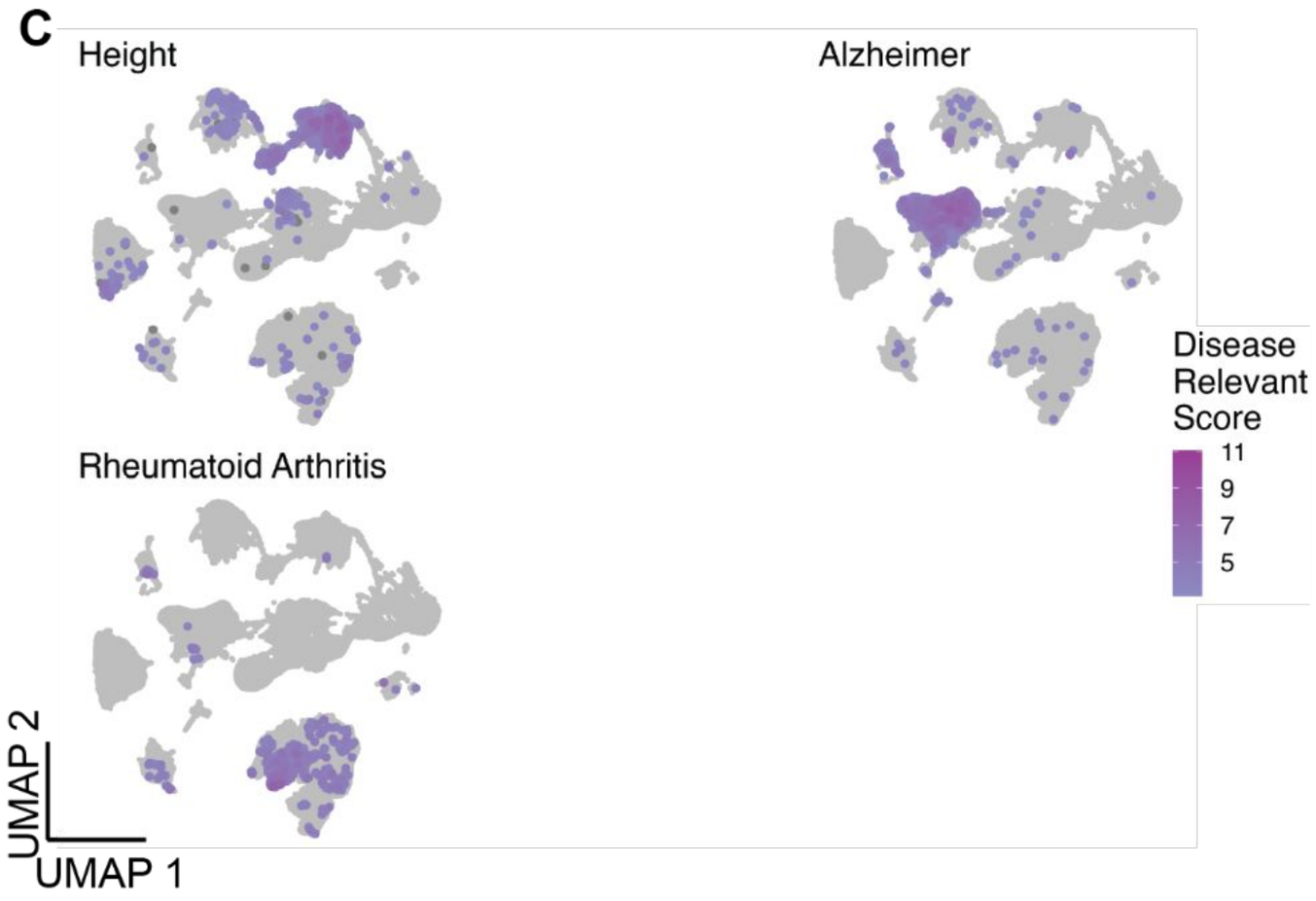
(D) Bar plot representing LDSC-SEG heritability enrichment coefficient (τ*) for each set of cell-type-specific differentially accessible peaks for each of the asthma-associated traits.

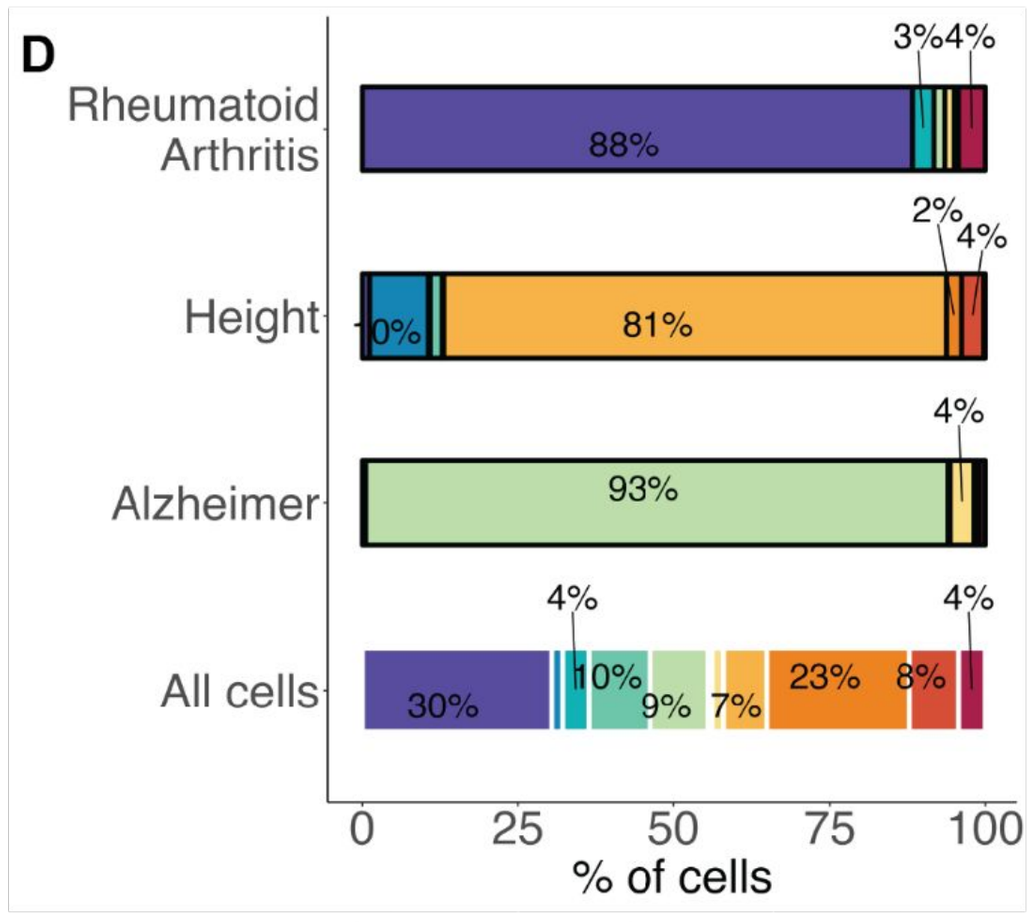
(E) Bar plots showing LDSC-SEG heritability enrichment coefficient (τ*) for each of the asthma-associated traits in cell-state-specific DA peaks of immune cells divided in either resting (light colors) or activated (dark colors) condition. In all bar plots, error bars represent τ* +/- standard error and asterisks denote significance as *** Bonferroni-adjusted $P < 0.05$, ** FDR 5% , * $P < 0.05$, and NS denotes nonsignificant ($P > 0.05$).

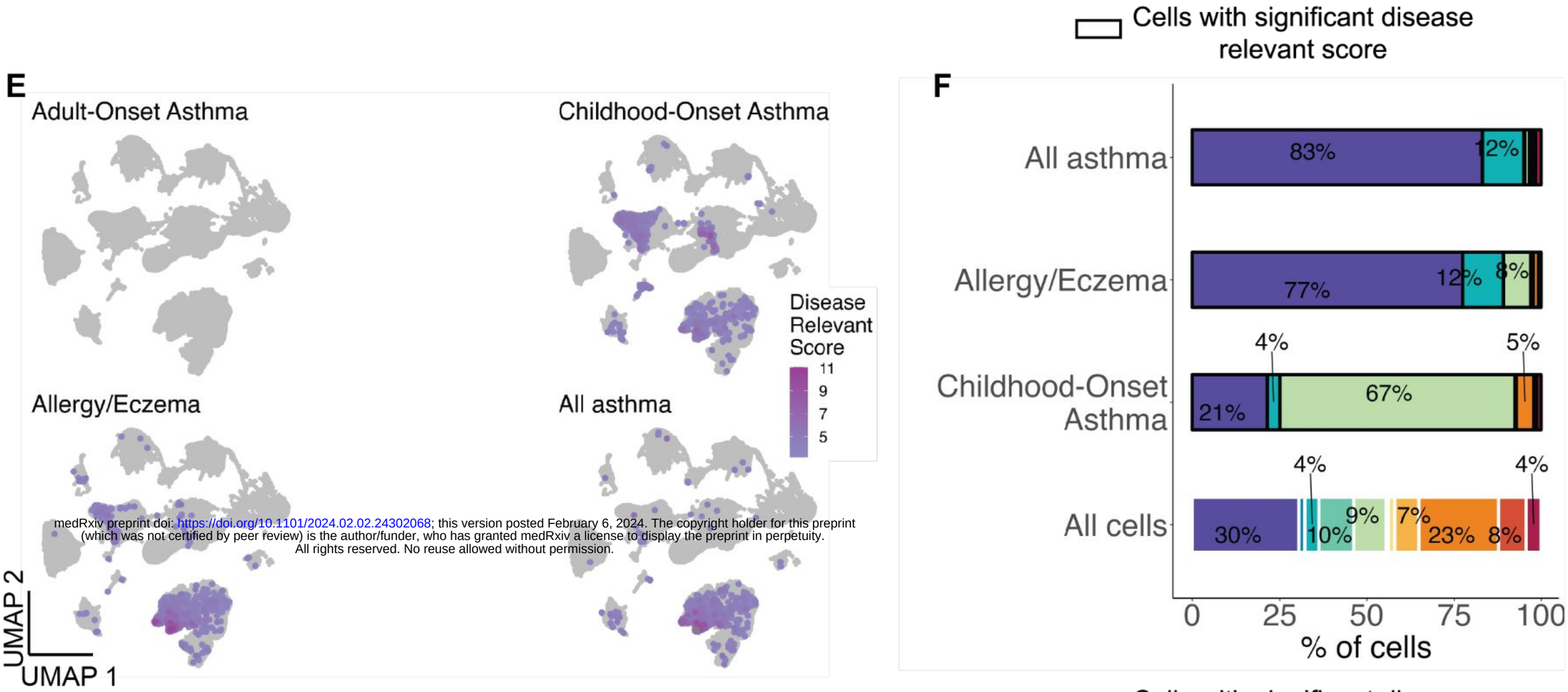
Supplementary Figure 2











Supplementary Figure 2. T cell validation at the single-cell RNA-seq level.

(A) Experimental design of the *Wang et al.* scRNA-seq dataset of sinonasal mucosa from healthy donors and chronic rhinosinusitis patients (CRS).

(B) UMAP visualization of the 1,115,856 immune and non-immune cells colored by cell type.

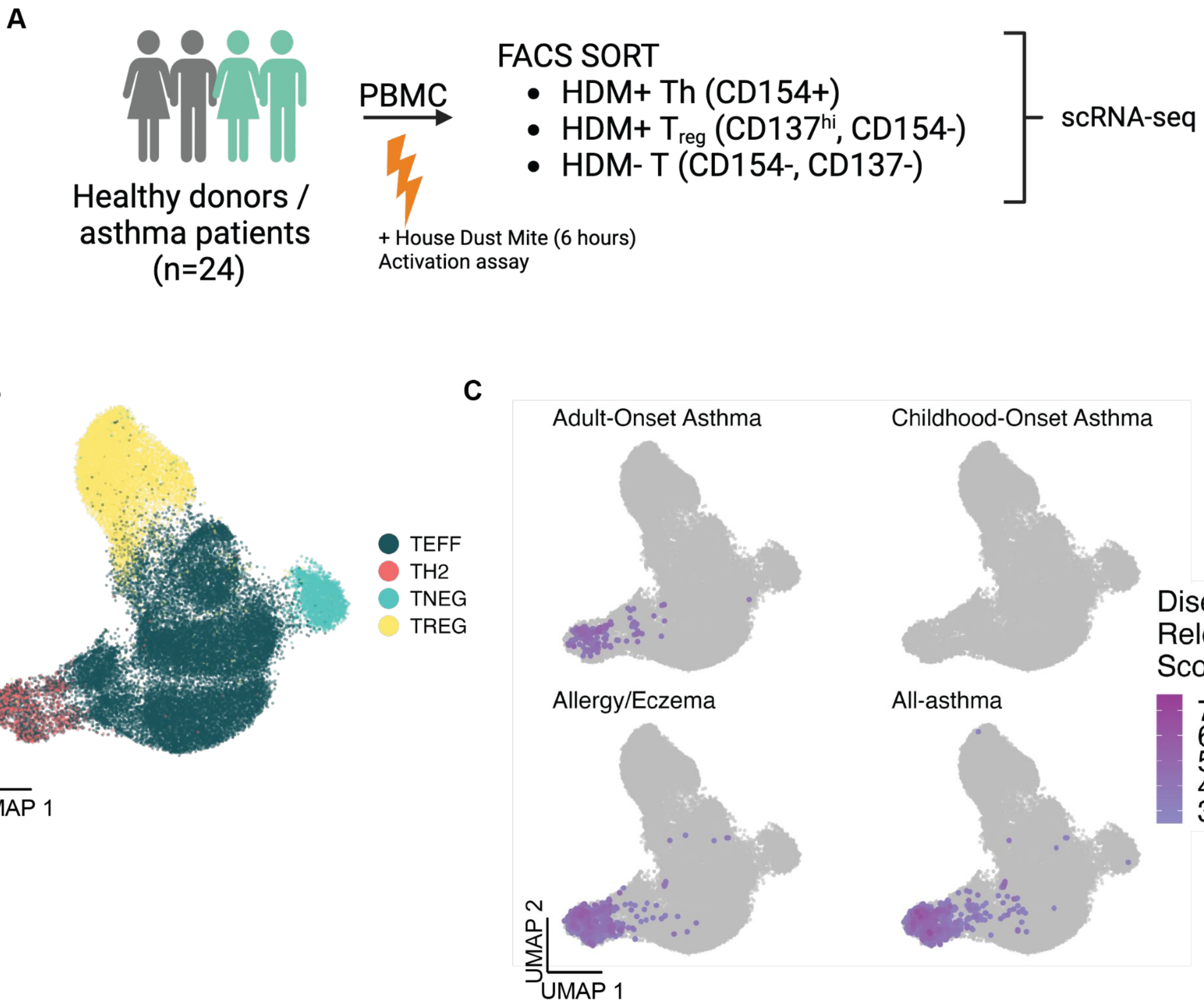
(C) scDRS results represented on the UMAP for the 3 control GWAS tested. The intensity of the color represents the disease relevant score, the lighter purple represents a less intense score whereas a more intense purple represents cells associated with a stronger score. Non-significant cells with a FDR higher than 10% are depicted in gray.

(D) Bar plot representing the percentage of each cell type in all cells followed by the significant cells at 10% FDR for scDRS in Alzheimer's Disease, Height and Rheumatoid Arthritis.

(E) scDRS results represented on the UMAP for the 4 asthma-associated traits tested. The intensity of the color represents the disease relevant score, the lighter color represents a less intense score whereas a more intense color represents cells associated with a stronger score.

(F) Bar plot representing the percentage of each cell type in all cells followed by the significant cells at 10% FDR for scDRS in COA, Allergy/Eczema and All asthma.

Supplementary Figure 3



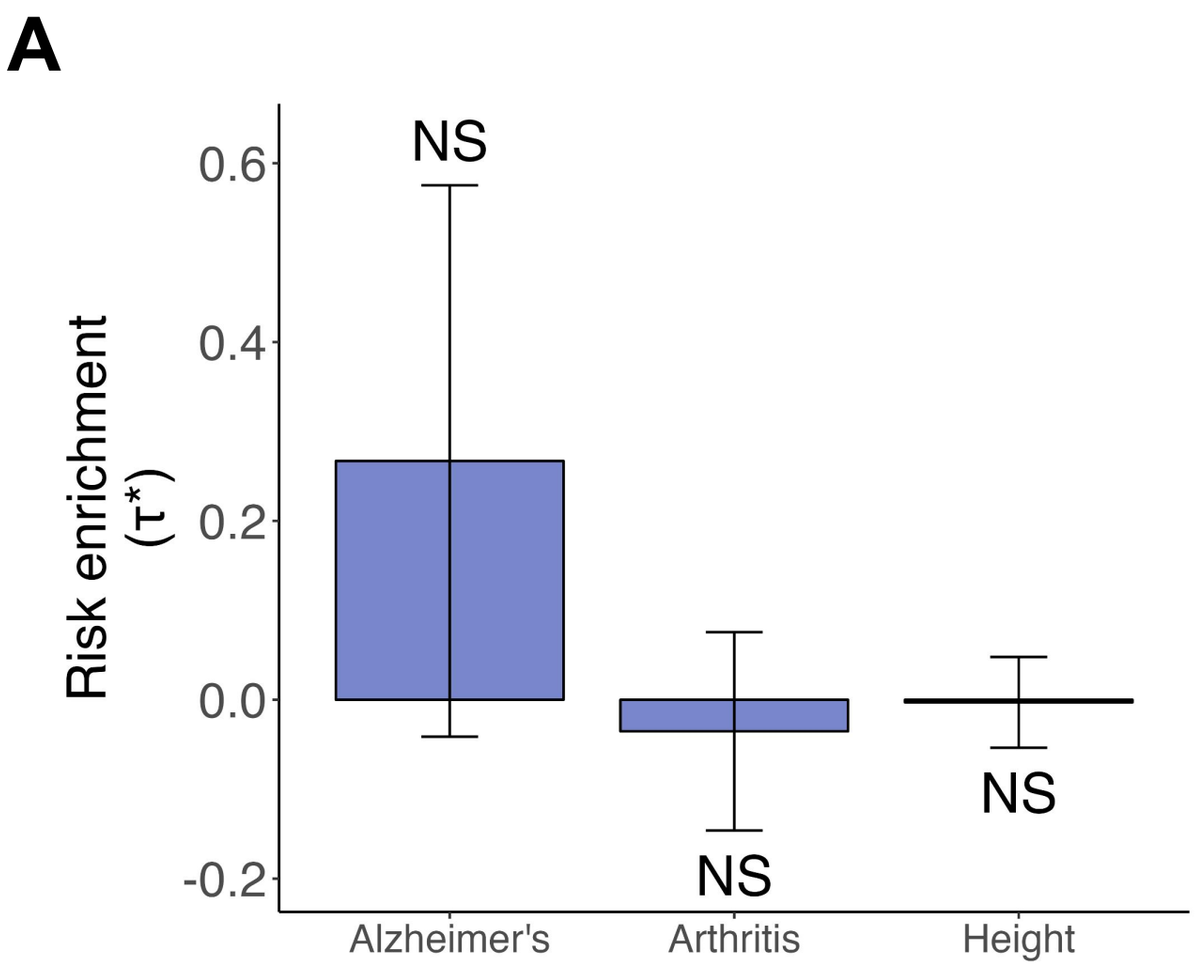
Supplementary Figure 3. Th2 validation at the single-cell RNA-seq level.

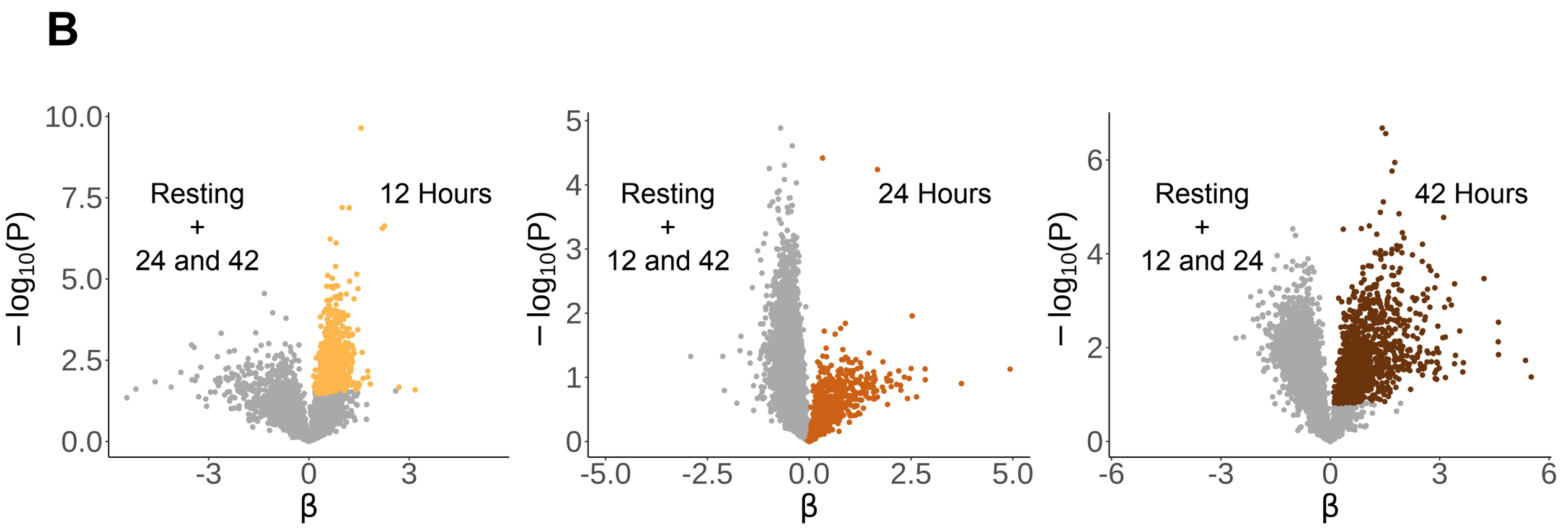
(A) Experimental design of the *Seumois et al.* scRNA-seq dataset consisting of T cells.

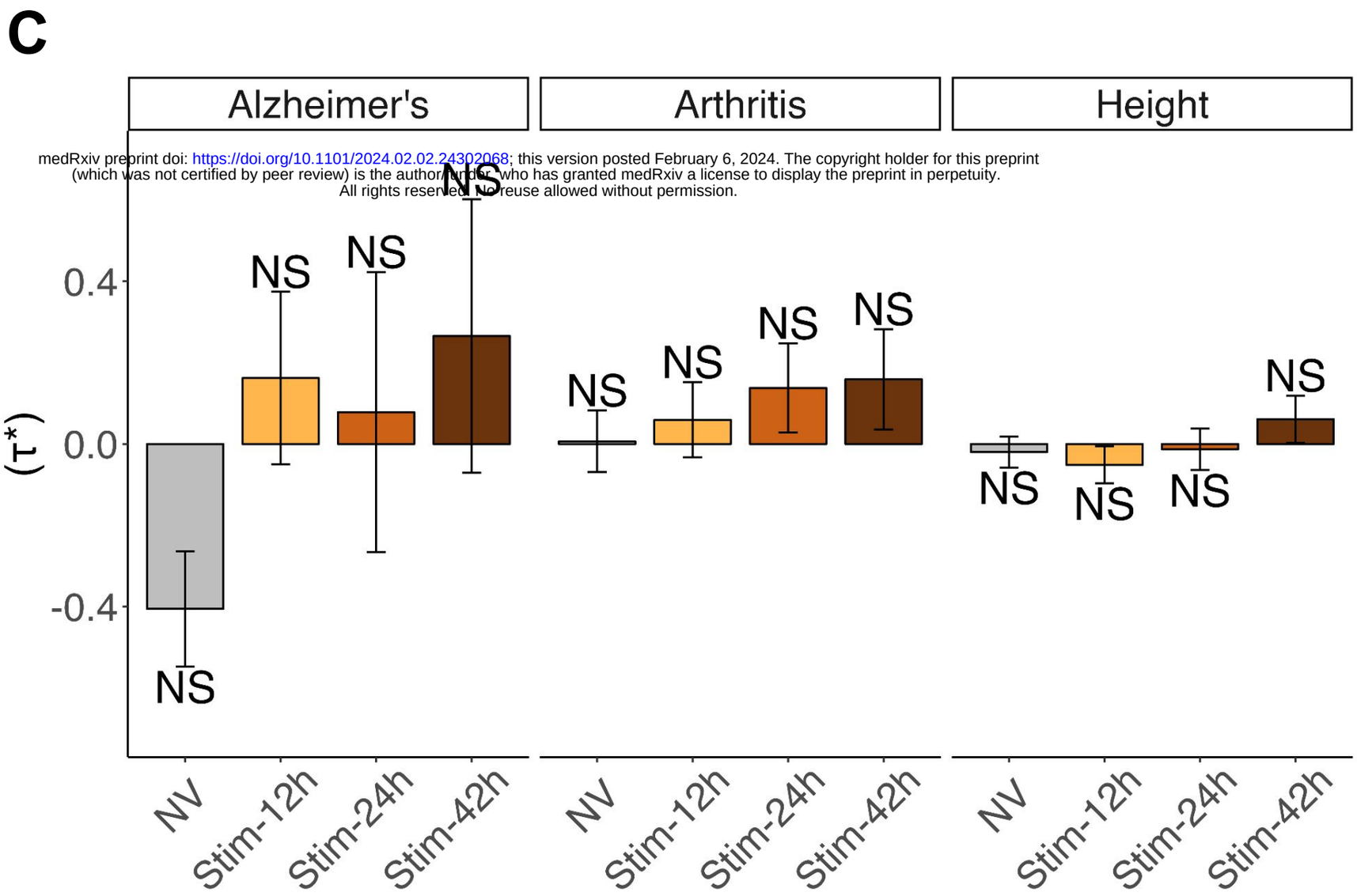
(B) UMAP visualization of the 38,559 T cells colored by subtypes.

(C) scDRS results represented on the UMAP for the 4 asthma-associated GWAS tested. The intensity of the color represents the disease relevant score, the lighter color represents a less intense score whereas a more intense color represents cells associated with a stronger score. nonsignificant cells with a FDR higher than 10% are depicted in gray.

Supplementary Figure 4







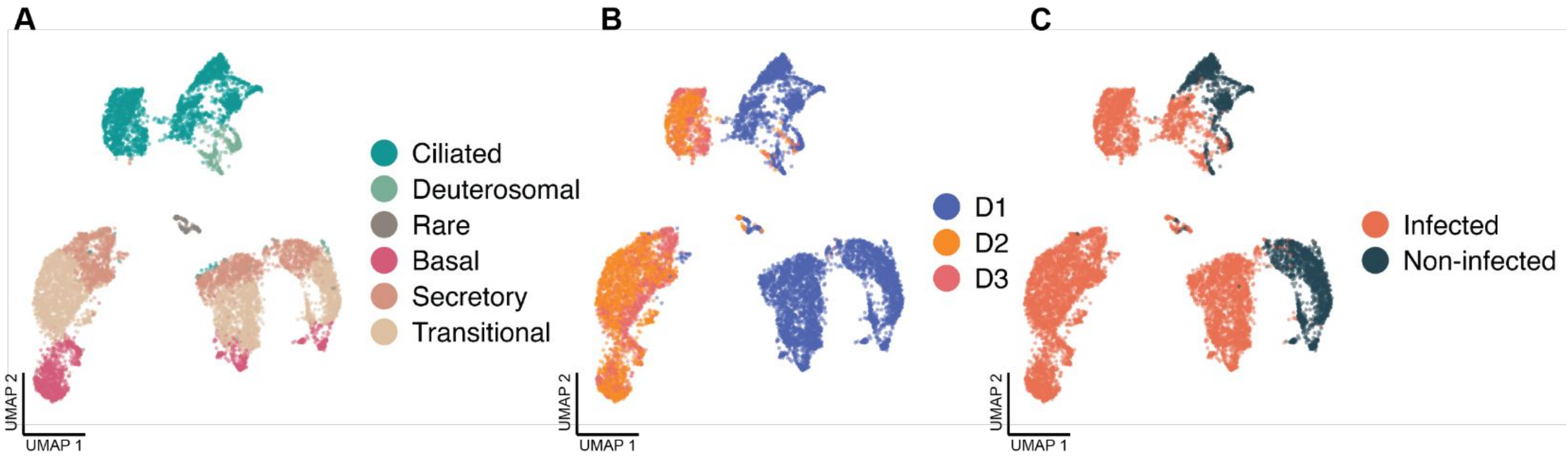
Supplementary Figure 4. Control traits of BECs from healthy individuals infected with RV.

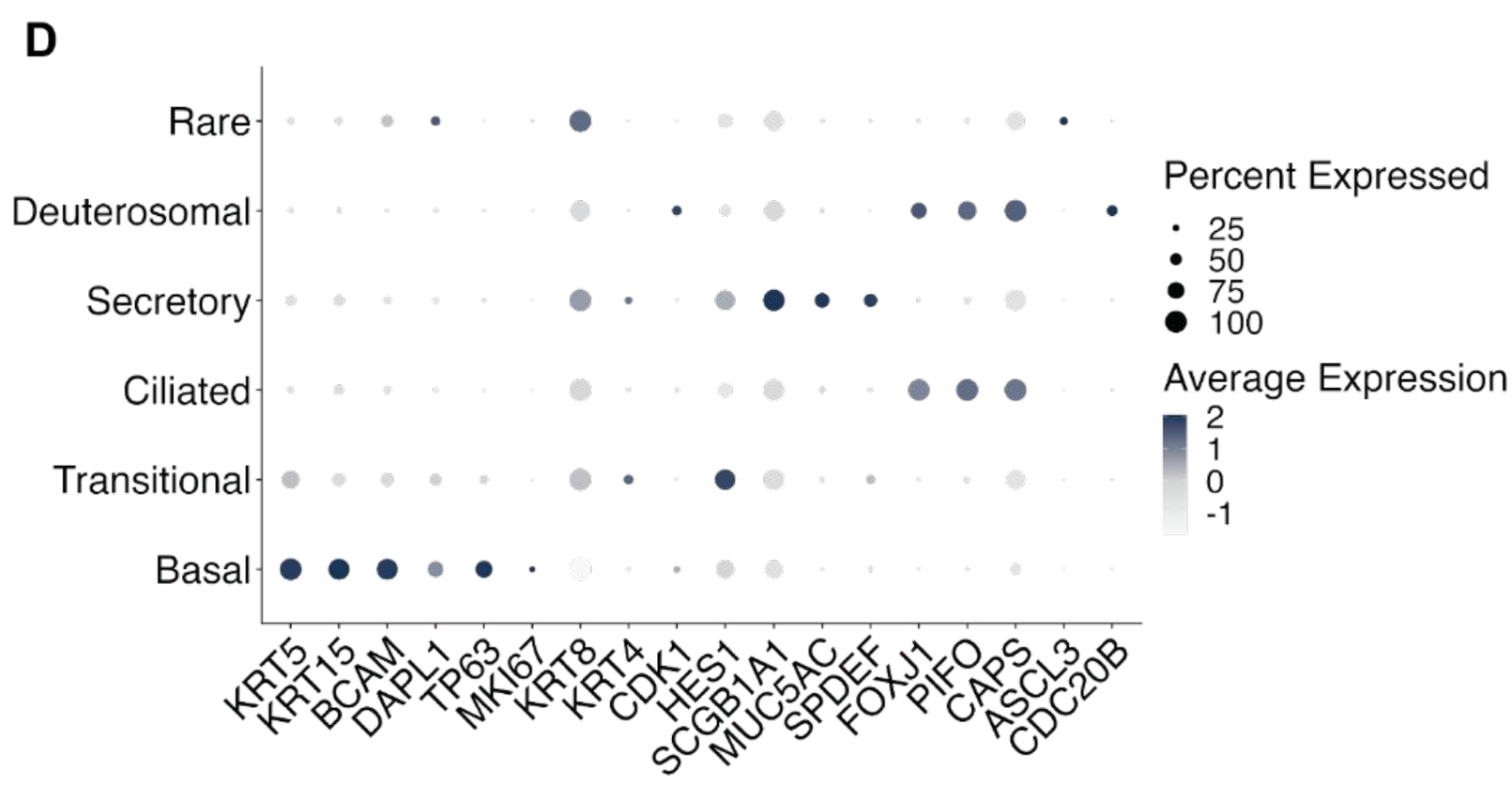
(A) Bar plots representing LDSC-SEG heritability enrichment coefficient (τ*) for the 3 control traits. NS denotes nonsignificant (P > 0.05).

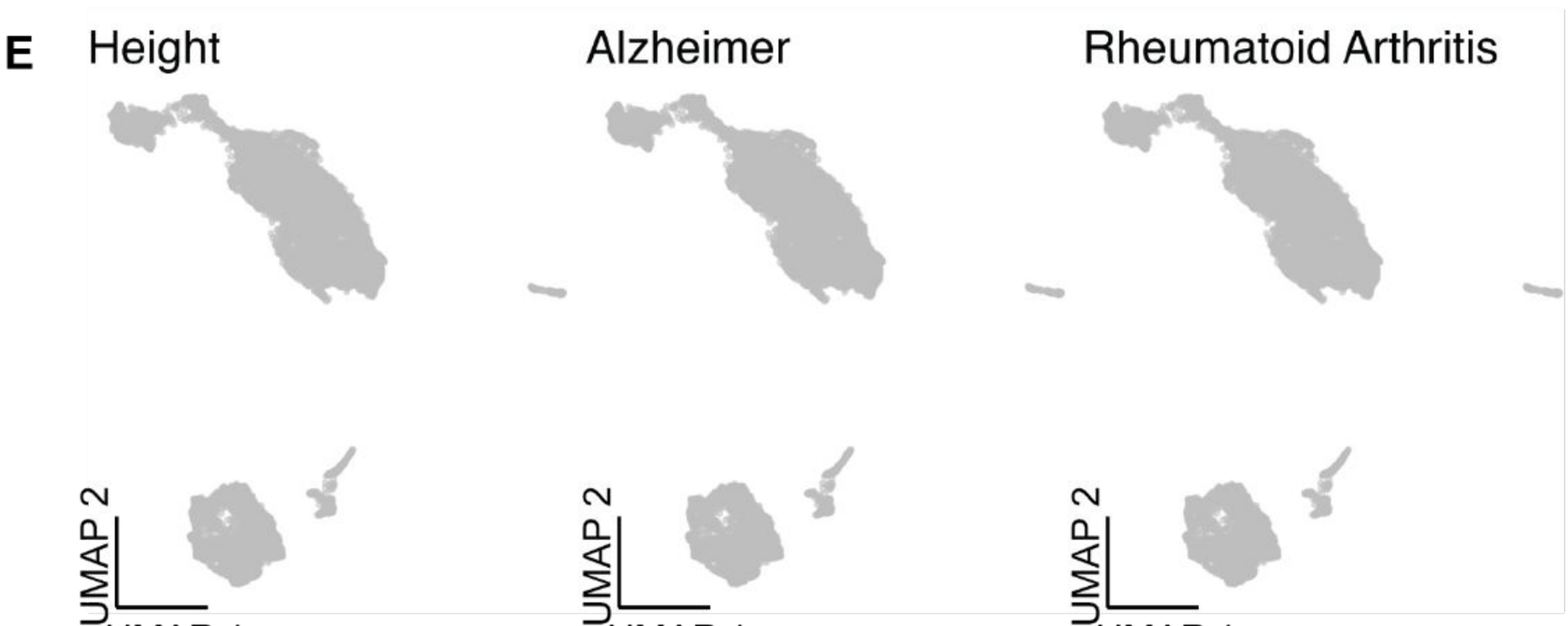
(B) Volcano plots showing differentially expressed genes at each time point (12,24,42) compared to all others in RV-infected epithelial cells colored in yellow, orange and brown respectively.

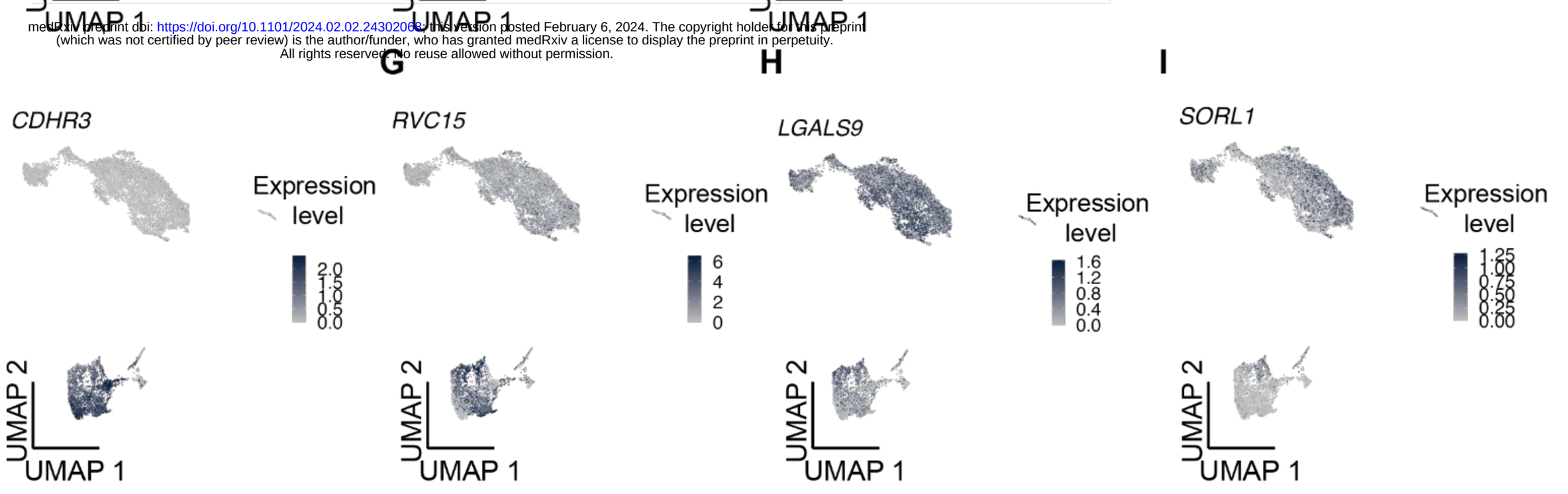
(C) Bar plot representing LDSC-SEG heritability enrichment coefficient (τ*) for each time point against all others for each of the control trait tested. NS denotes nonsignificant (P > 0.05). In all bar plots, error bars represent τ* +/- standard error.

Supplementary Figure 5









Supplementary Figure 5. Additional informations related to Figure 2.

UMAP visualization of the 10,721 epithelial cells colored by **(A)** cell type. **(B)** donor.

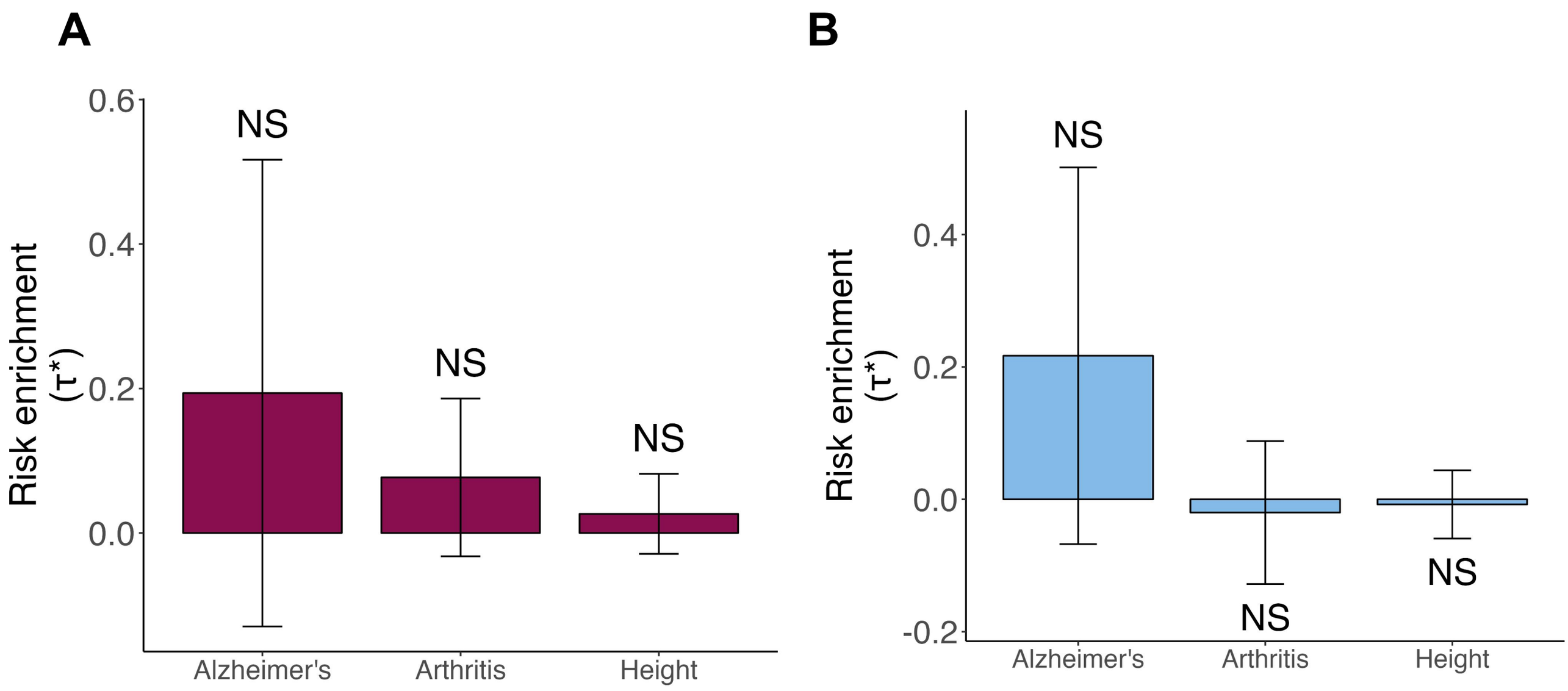
(C) infection status (Infected / Non-infected).

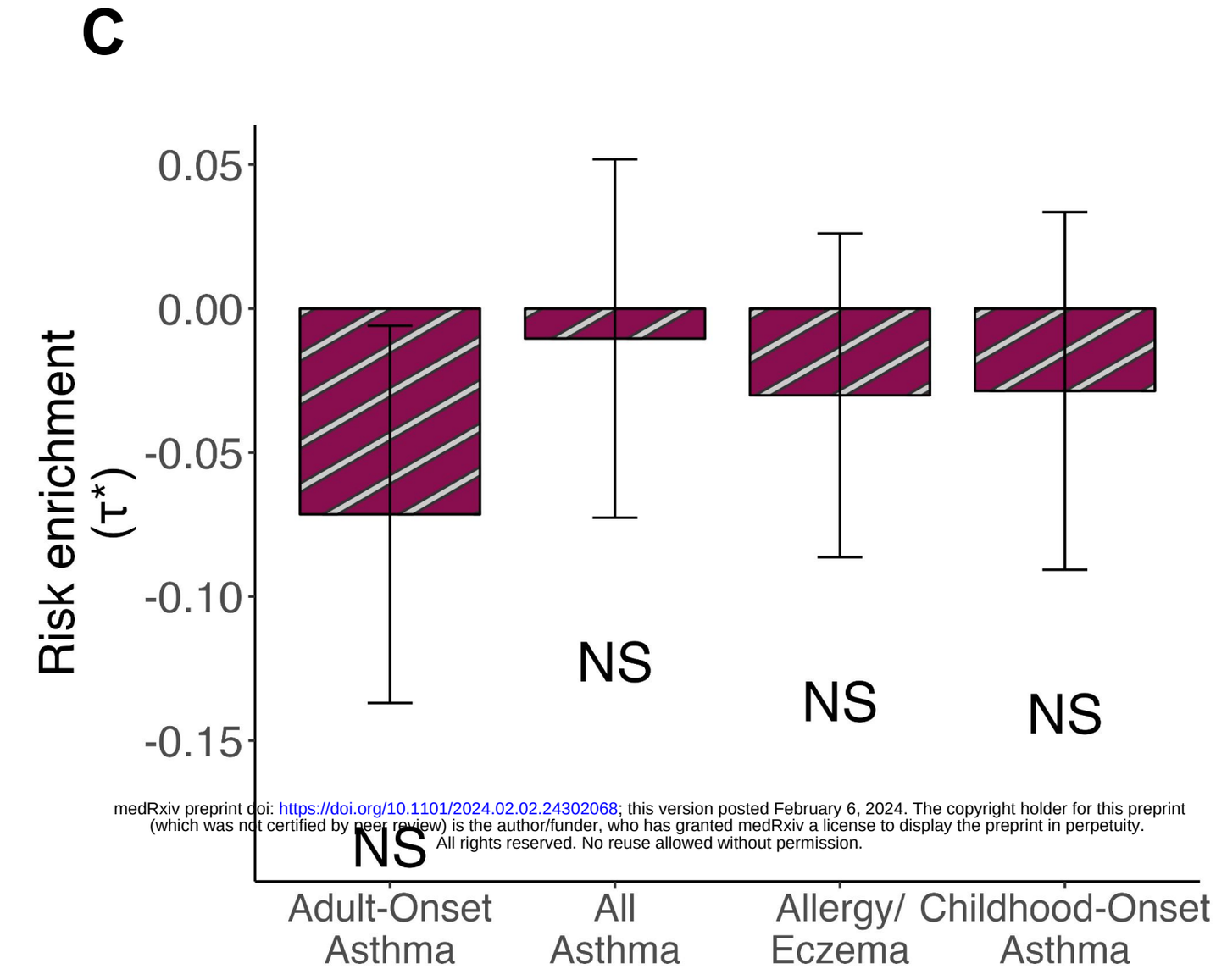
(D) Dot plot representing the normalized average expression and the percent of cells expressing a given gene for epithelial cell markers.

(E) scDRS results represented on the UMAP for the 3 control GWAS tested. nonsignificant cells with a FDR higher than 10% are depicted in gray.

(F-I) Harmonized UMAP representing *CDHR3, RVC15, LAGLS9 and SORL1* normalized expressions.

Supplementary Figure 6



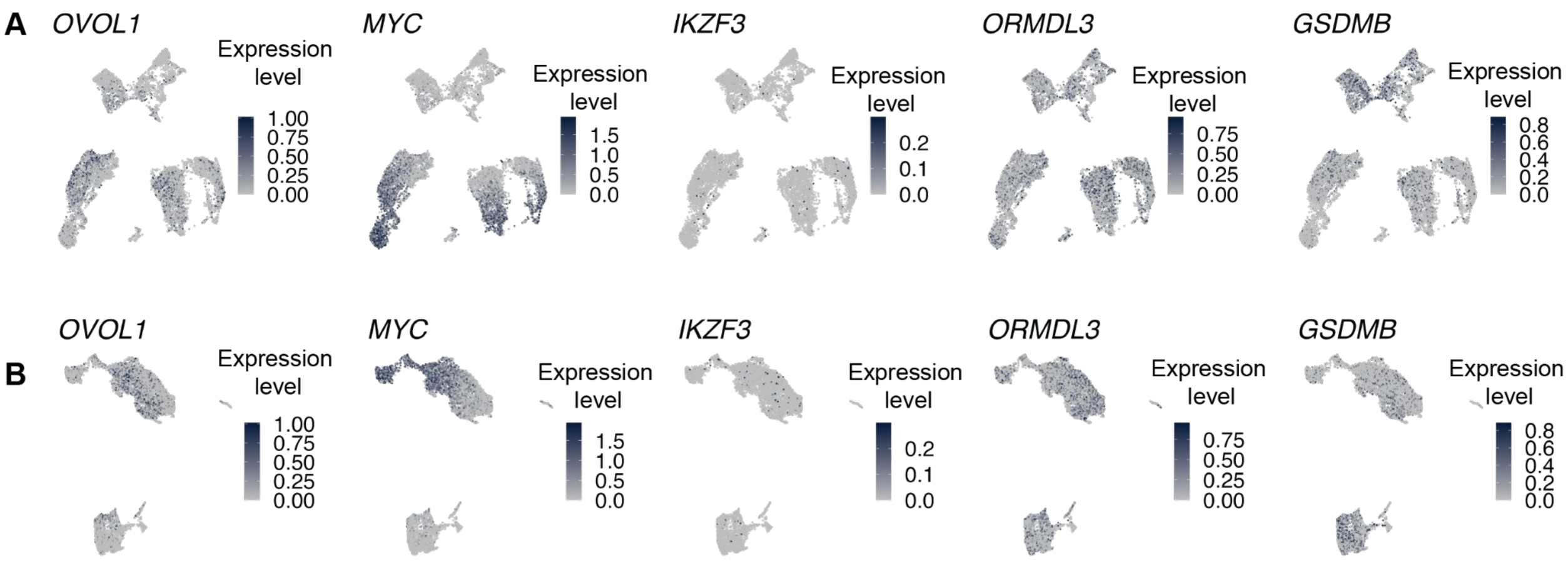


Supplementary Figure 6. Control traits of BECs from asthma patients infected with RV.

Bar plots representing LDSC-SEG heritability enrichment coefficient (τ*) for the 3 control traits tested. NS denotes nonsignificant (P > 0.05). **(A)** Using DE genes after RV-infection in epithelial cells from patients.

(B) Using genes differentially expressed in asthmatics when compared to healthy individuals. **(C)** Bar plot showing LDSC-SEG heritability enrichment coefficient (τ*) across traits for genes downregulated upon RVC-15 infection. NS denotes nonsignificant (P > 0.05). In all bar plots, error bars represent τ* +/- standard error.

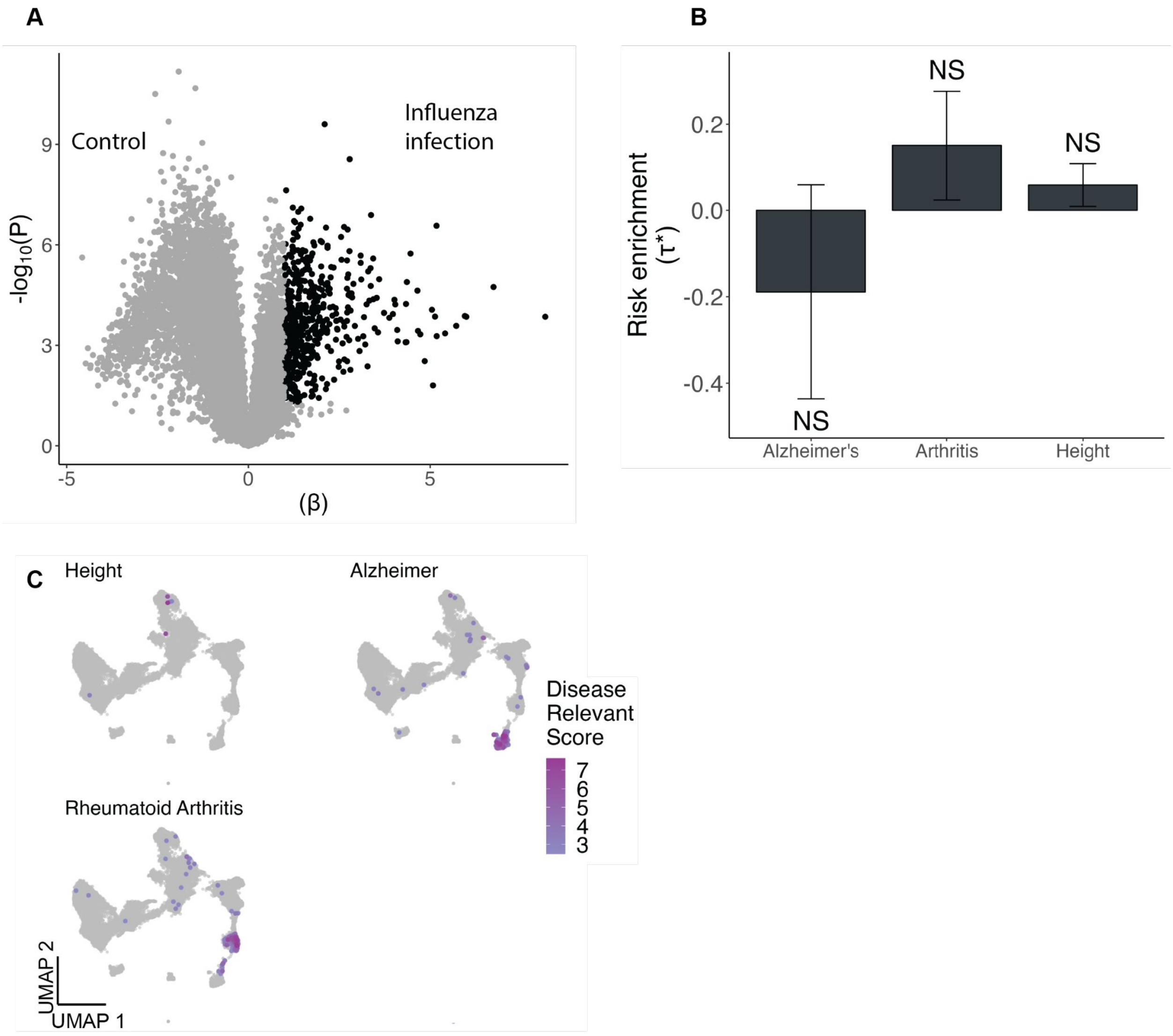
Supplementary Figure 7



Supplementary Figure 7. Normalized expressions of gene of interest.

(A) Normalized expression of genes of interest represented on the UMAP and **(B)** on the harmonized UMAP.

Supplementary Figure 8



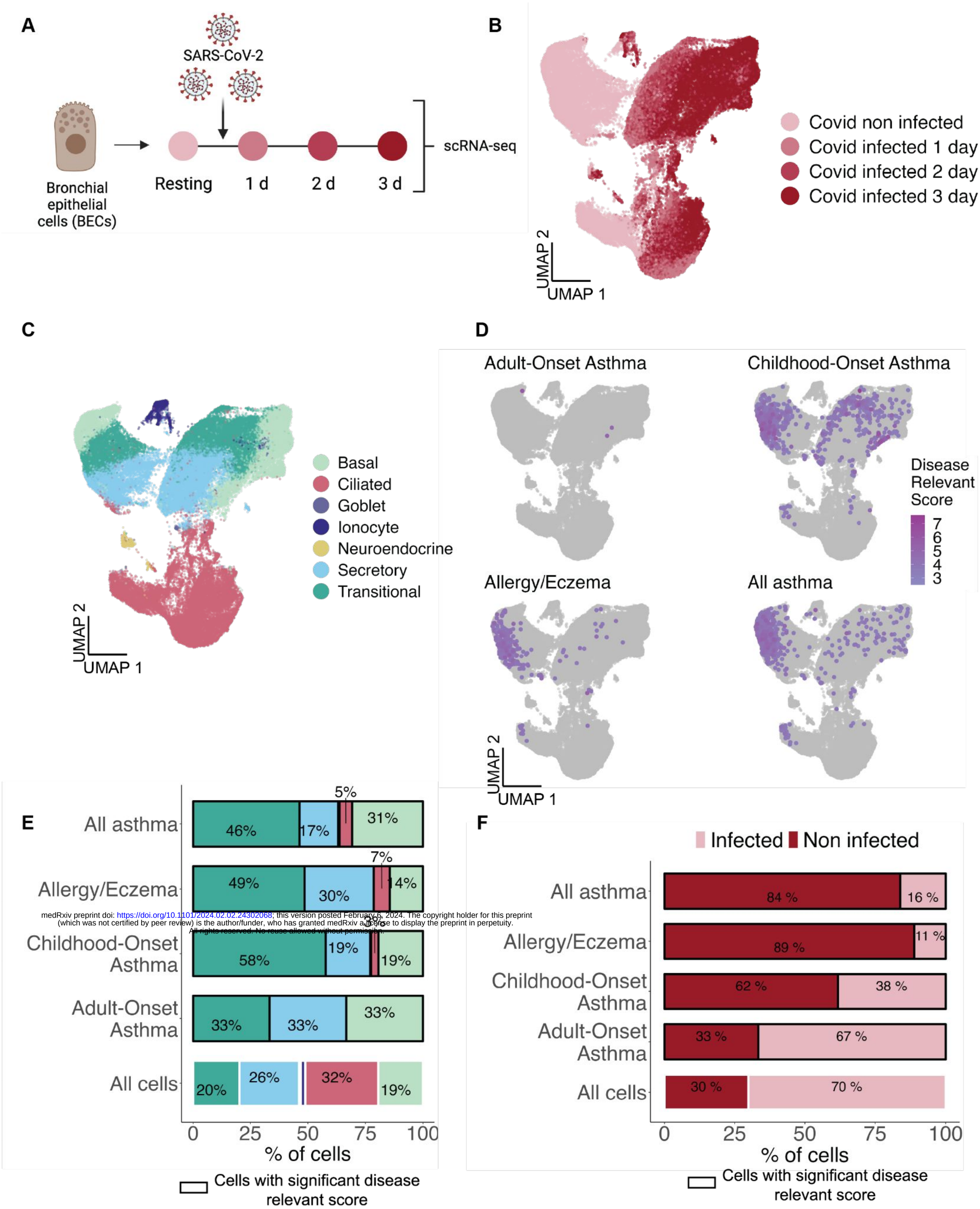
Supplementary Figure 8. Volcano plot and controls traits related to Figure 5.

(A) Volcano plot showing differentially expressed genes between influenza infection and control. Genes upregulated upon influenza infection were selected based on t-statistic and are colored in black.

(B) Bar plots representing LDSC-SEG heritability enrichment coefficient (τ*) for the 3 control GWAS tested. Error bars represent τ* +/- standard error. NS denotes nonsignificant (P > 0.05).

(C) scDRS results represented on the UMAP for the 3 control traits tested. The intensity of the color represents the disease relevant score, the lighter color represents a less intense score whereas a more intense color represents cells associated with a stronger score. nonsignificant cells with a FDR higher than 10% are depicted in gray.

Supplementary Figure 9



Supplementary Figure 9. scRNA-seq analysis of BECs infected or not with SARS-CoV-2.

(A) Experimental design of the *Ravindra et al.* scRNA-seq dataset of bronchial epithelial cells infected with SARS-CoV-2 or not, from one healthy donor.

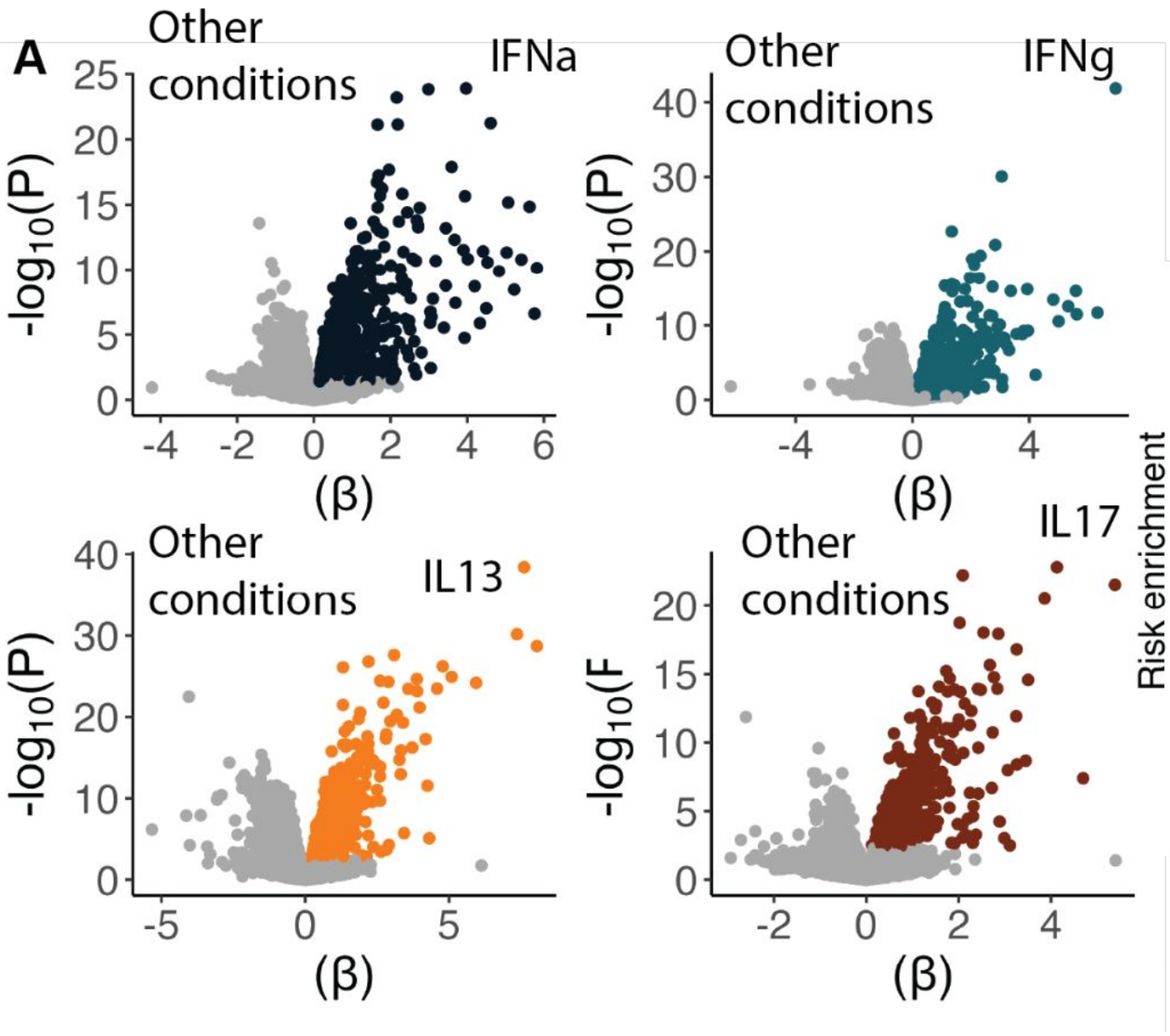
UMAP visualization of the 74,088 cells colored by **(B)** days after virus infection and **(C)** by cell type.

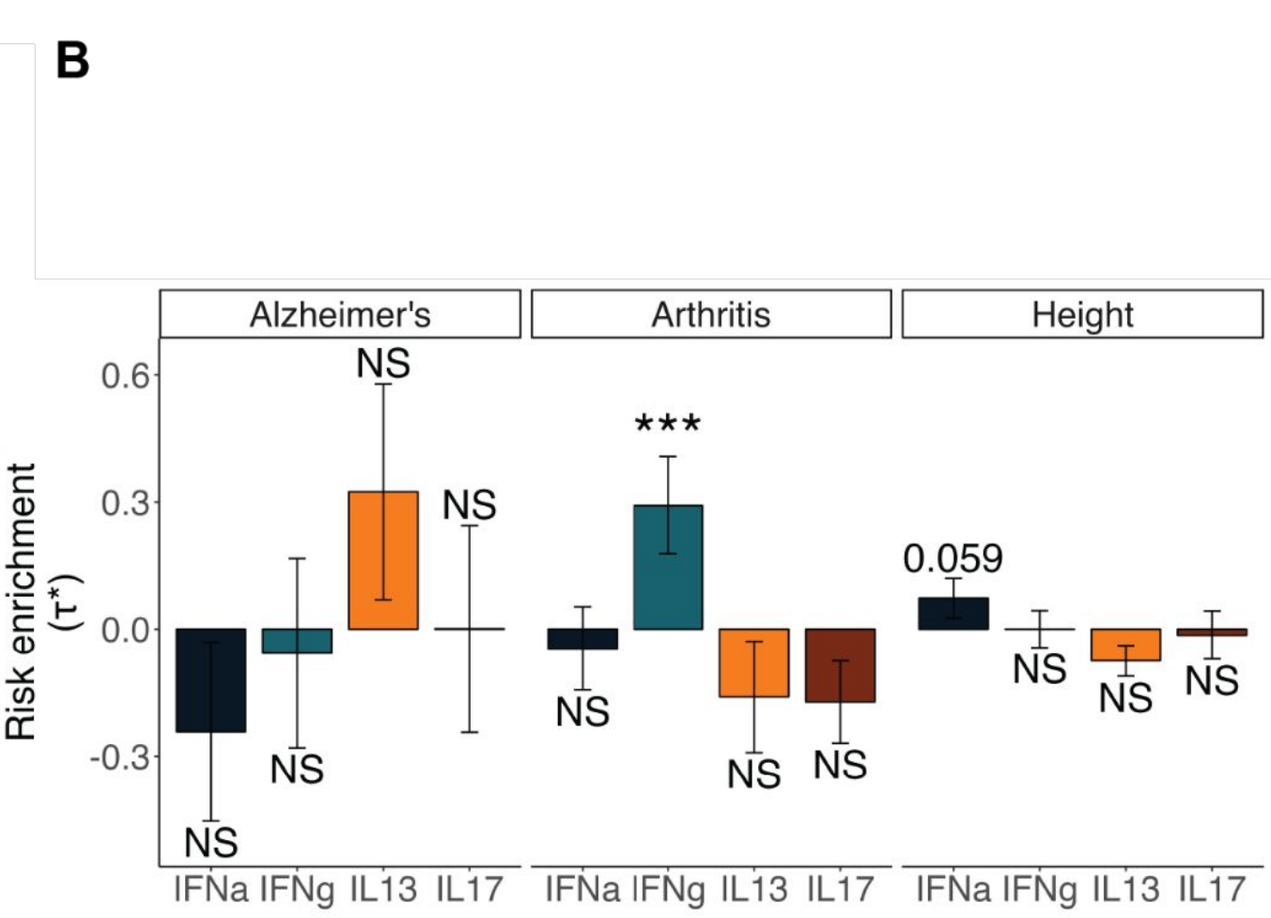
(D) scDRS results represented on the UMAP for the 3 control traits tested. The intensity of the color represents the disease relevant score, the lighter color represents a less intense score whereas a more intense color represents cells associated with a stronger score. nonsignificant cells with a FDR higher than 10% are depicted in gray.

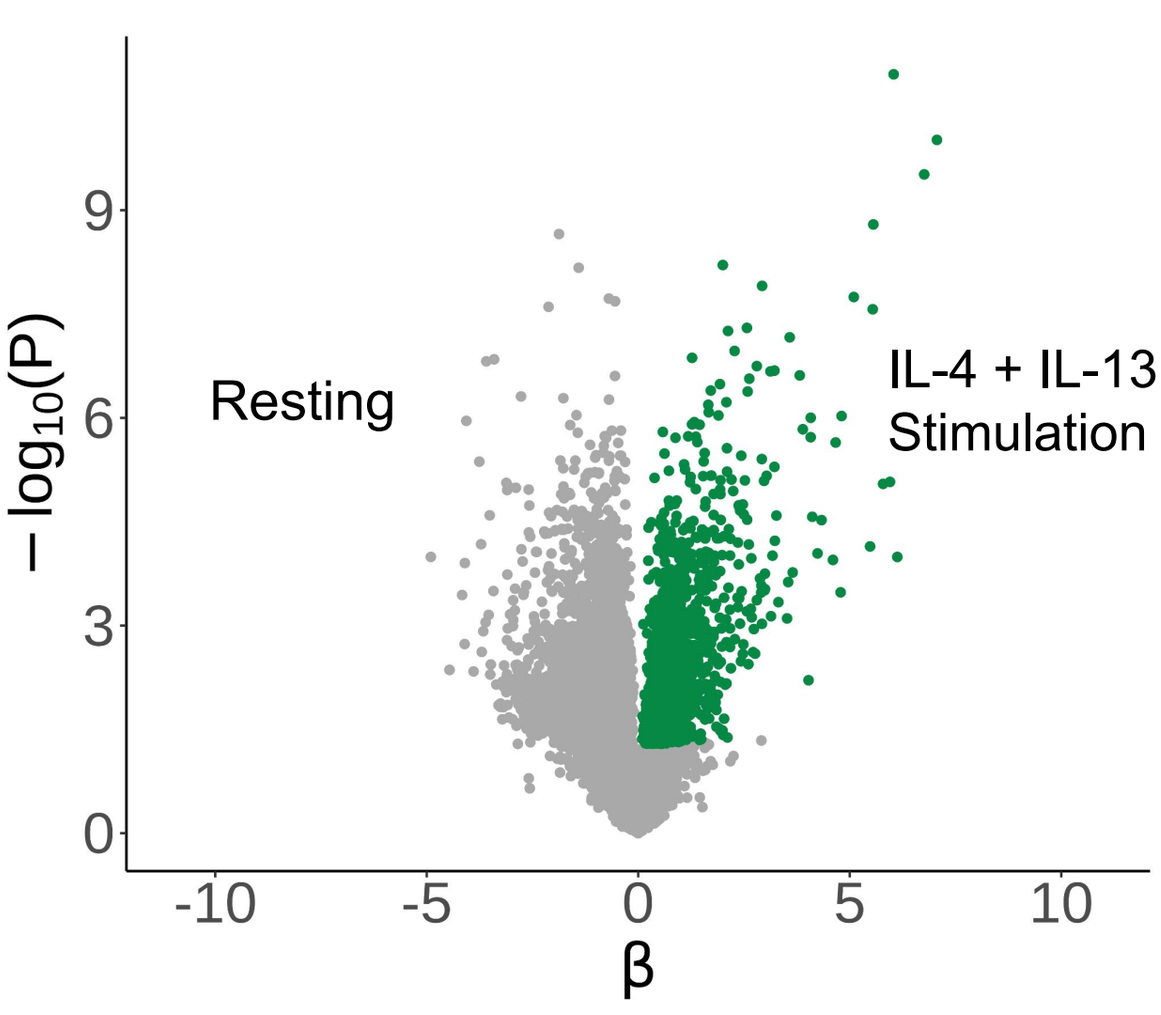
(E) Bar plot representing the percentage of each cell type in all cells followed by the significant cells at 10% FDR for scDRS in AOA, COA, Allergy/Eczema and All Asthma.

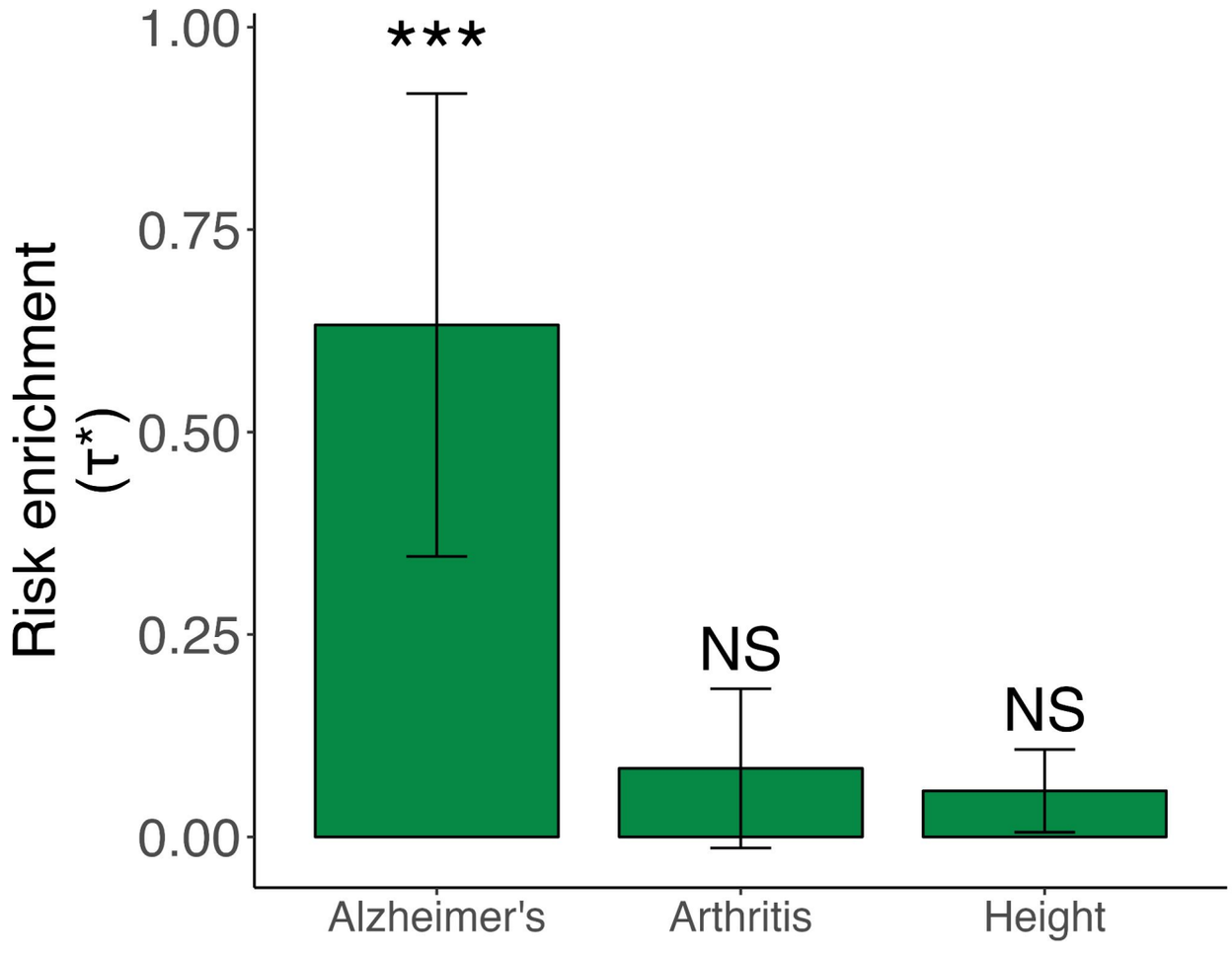
(F) Bar plot representing the percentage of cells in the full dataset classified in SARS-CoV-2 infection or non infected, followed by percentage of cells passing significance at 10% FDR for AOA, COA, Allergy/Eczema, All asthma, by SARS-CoV-2 infection or not.

Supplementary Figure 10









Supplementary Figure 10. Control traits for Figure 6.

(A) Volcano plots showing differentially expressed genes for each stimuli (IFNα, IFNγ, IL-13 and IL-17) compared to all others in bronchial epithelial cells colored in black, teal, orange and brown respectively.

(B) Bar plot representing LDSC-SEG enrichment for each stimuli against all others for each of the 3 control traits. Error bars represent τ* +/- standard error. Asterisk denotes significance as * P < 0.05 and NS denotes nonsignificant (P > 0.05).

(D) Volcano plot showing differentially expressed genes between IL-4/IL-13 stimulation and resting condition. Genes upregulated upon IL-4/IL-13 stimulation were selected based on t-statistic and are colored in green.

(E) Bar plots representing LDSC-SEG heritability enrichment coefficient (τ*) for the 3 control traits. Error bars represent τ* +/- standard error. Asterisk denotes significance as * P < 0.05 and NS denotes nonsignificant (P > 0.05).