

## Comparative study of epithelial entry gene expression of SARS-CoV-2 and other human viral species in asthma: differences by sex, airway location and disease endotype

**AUTHORS:** <sup>1</sup>Mackenzie E. Coden, <sup>1</sup>Lucas F. Loffredo, <sup>2</sup>Hiam Abdala-Valencia and <sup>1</sup>Sergejs Berdnikovs\*

<sup>1</sup>Division of Allergy and Immunology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA.

<sup>2</sup>Division of Pulmonary and Critical Care, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA.

\*Correspondence to: Sergejs Berdnikovs, Northwestern University Feinberg School of Medicine, Division of Allergy and Immunology, 240 East Huron Street, McGaw M-317, Chicago, IL 60611, USA; phone: 1-312-503-6924; fax number: 1-312-503-0078; e-mail address: s-berdnikovs@northwestern.edu

## ABSTRACT

Epithelial characteristics underlying the differential susceptibility of chronic asthma to SARS-CoV-2 (COVID-19) and other viral infections are currently unclear. By revisiting transcriptomic data from patients with Th2 low versus Th2 high asthma, as well as mild, moderate and severe asthmatics, we characterized the changes in expression of human coronavirus and influenza viral entry genes relative to sex, airway location and disease endotype. We found sexual dimorphism in expression of COVID-19 genes ACE2, TMPRSS2, TMPRSS4, and SLC6A19. ACE2 receptor downregulation occurred specifically in females in Th2 high asthma, while proteases broadly assisting coronavirus and influenza viral entry, TMPRSS2 and TMPRSS4, were highly upregulated in both sexes. Overall, changes in COVID-19 gene expression were specific to Th2 high molecular endotype of asthma, and different by asthma severity and airway location. The downregulation of ACE2 (COVID-19, SARS) and ANPEP (HCoV-229E) viral receptors correlated with loss of club and ciliated cells in Th2 high asthma, while the increase in DPP4 (MERS-CoV), ST3GAL4, and ST6GAL1 (influenza) associated with an increase in goblet and basal activated cells. Overall, this study elucidates sex, airway location, disease endotype and changes in epithelial heterogeneity as factors underlying asthmatic susceptibility, or lack thereof, to COVID-19.

## KEYWORDS

COVID-19, SARS, HCoV-229E, MERS, influenza, virus, epithelium, asthma, allergy, inflammation, sexual dimorphism, gene expression

## INTRODUCTION

As the COVID-19 pandemic continues to unfold worldwide, new insights are coming to light about the mechanisms and epidemiology of this disease. One of the most puzzling recent observations is an apparent lack of association of SARS-CoV-2 infection with asthma. Reports from Wuhan, China revealed that asthma was not a risk factor for severity and mortality in adult COVID-19 patients - the prevalence of asthma in patients with COVID-19 was only 0.9%, which is substantially lower than that in the adult population of Wuhan (6.4%)<sup>1</sup>. In another study of 140 patients infected with SARS-CoV-2 in Wuhan, none presented with asthma or other allergic diseases<sup>2</sup>. A study by Mahdavinia et al.<sup>3</sup> in a large cohort of adult COVID-19 cases in the United States showed that asthma was significantly associated with longer intubation but was not associated with higher rate of hospitalization, death, or acute respiratory distress syndrome among COVID-19 patients. This is surprising because respiratory viral infections in general are known to associate with asthma exacerbations and more serious illness. One study proposed an explanation for this by showing that respiratory allergies and controlled allergen exposures were each associated with significant reductions in epithelial expression of ACE2, the entry receptor for SARS-Cov-2<sup>4</sup>. Specifically, ACE2 expression was lowest in individuals with both high levels of allergic sensitization and asthma<sup>4</sup>. Another study confirmed this observation by showing that Type 2 cytokine IL-13 decreases expression of ACE2 in bronchial epithelium, thus providing one mechanistic link to Th2 inflammation<sup>5</sup>. In this study, we pursued the hypothesis that susceptibility of asthmatic epithelium to different viral species depends on differential expression of receptor entry genes in different sexes, airway locales and asthma endotypes.

## MATERIALS AND METHODS

### Transcriptomic data

In this study we used the following two independent gene expression datasets downloaded from Gene Expression Omnibus (GEO, NCBI): (1) GSE67472, providing data for gene expression in bronchial epithelial brushings from healthy controls and subjects with Th2 low and Th2 high asthma and (2) GSE41861, providing epithelial brushing data for mild, moderate and severe asthma in both upper and lower airways. The study comparing Th2 low and Th2 high asthma by Woodruff et al.<sup>6</sup> is the most highly cited study providing basis for molecular endotyping of asthma patients based on patterns of Type 2 inflammation. Clinical descriptions of asthma patients and airway sampling procedures can be found in GEO Datasets and in the original publication of this study<sup>6</sup>. Gene

annotations were updated using technology-corresponding annotation libraries in R. Each individual GEO dataset underwent consistent handling, implementing standard contrast fits on Series Matrix File data uploaded to GEO by the original study groups, preserving original normalization protocols to maintain consistency with gene signatures published by original studies.

### **Clustering of epithelial and viral entry genes**

Single cell analysis data used to generate epithelial subset gene list for unsupervised clustering of epithelial and viral genes is from Nature Medicine study by Vieira Braga et al.<sup>7</sup> providing a cellular census of human lung transcriptome in health and asthma. Exact genes and epithelial subset clustering structure in health and asthma can be found in Fig. 3c in the original Vieira Braga's publication. Viral genes were selected based on published works of human coronavirus and influenza entry genes in epithelial cells. Epithelial-viral gene expression data set was then assembled from GSE67472 data and included both males and females, and healthy, Th2 low and Th2 high individuals. The resulting data set was subjected to hierarchical clustering using One Minus Pearson correlation metric and average linkage method. Similarity matrix was derived from clustered data using Pearson correlation metric. All procedures were performed in MORPHEUS, <https://software.broadinstitute.org/morpheus> (Broad Institute), matrix visualization and analysis software. Clustering structure in Fig. 1d in this publication is nearly an exact match to the above mentioned study, demonstrating alignment in methodology and biological message in regards to epithelial heterogeneity.

### **Statistical analysis and data visualization**

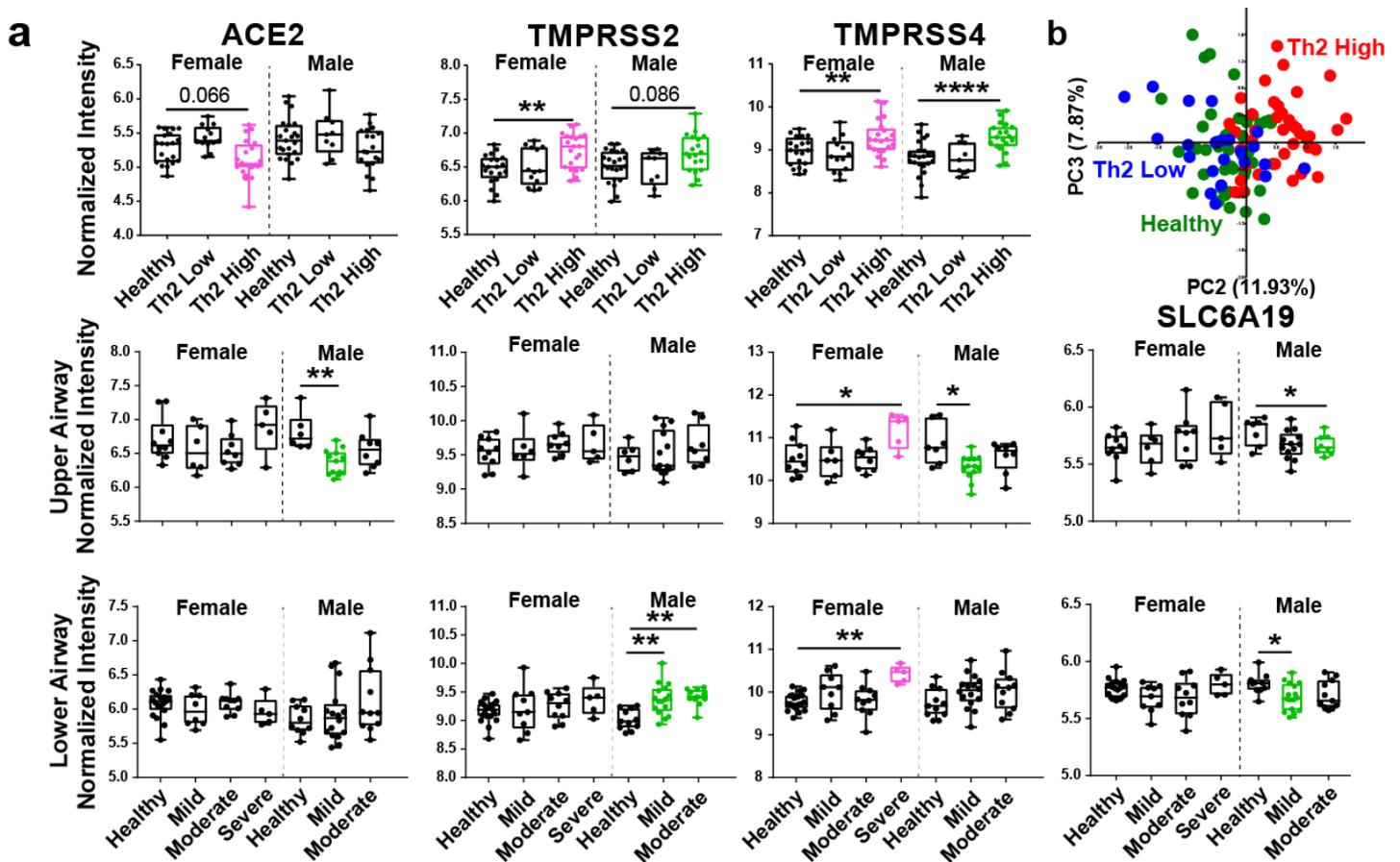
All graphs and statistical analyses were carried out in GraphPad Prism (Graphpad) and Systat 13 (Systat Software, Inc.). Principal Component Analysis (PCA) was executed in PAST 2.17 software using covariance matrices and singular value decomposition. Statistical significance was determined by ANOVA followed by Tukey's post-hoc pairwise testing. All data are represented as mean  $\pm$  S.E.M. An alpha level of 0.05 was used as a significance cut-off. Marginal values ( $0.05 > p < 0.1$ ) are also shown.

### **Data availability**

All raw data used in this study, including series matrix file data, can be accessed in GEO under the accession numbers GSE67472 and GSE41861. scRNA-seq data that support the findings of this study can be found in the above mentioned study by Vieira Braga et al<sup>7</sup>.

## RESULTS AND DISCUSSION

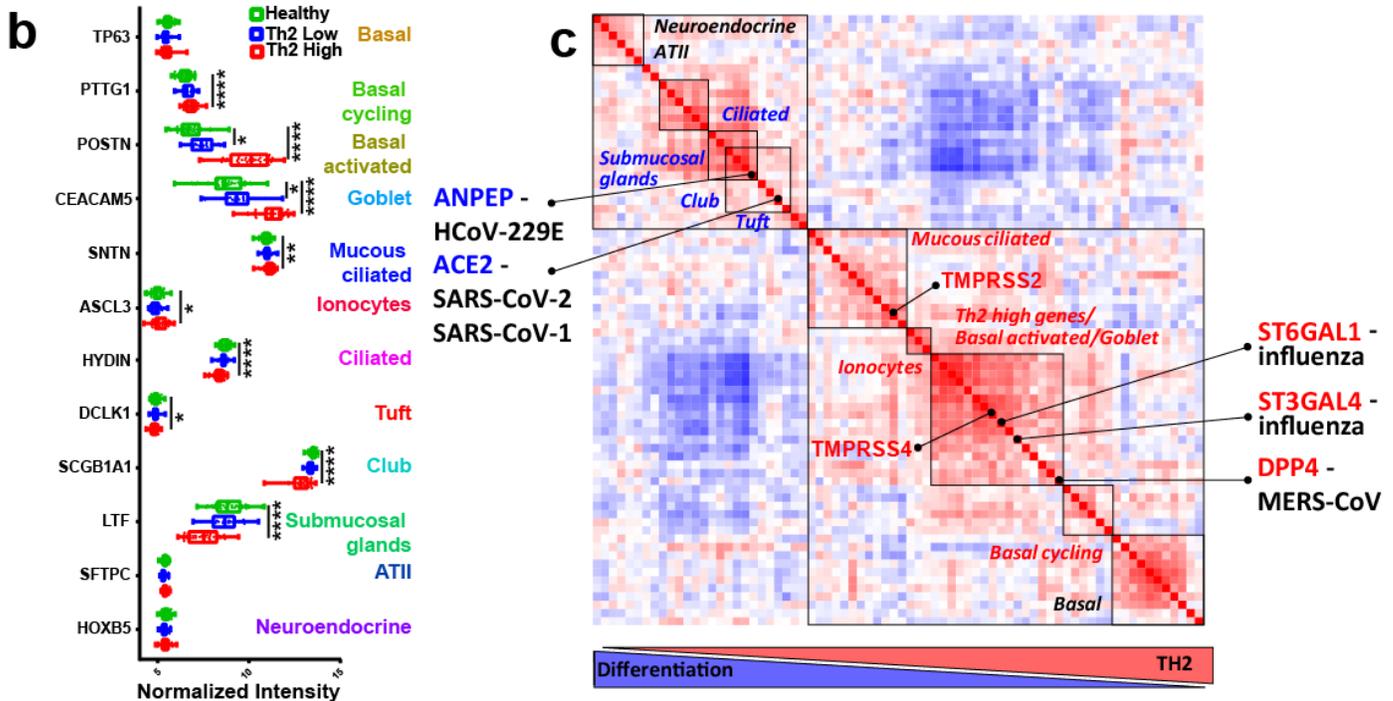
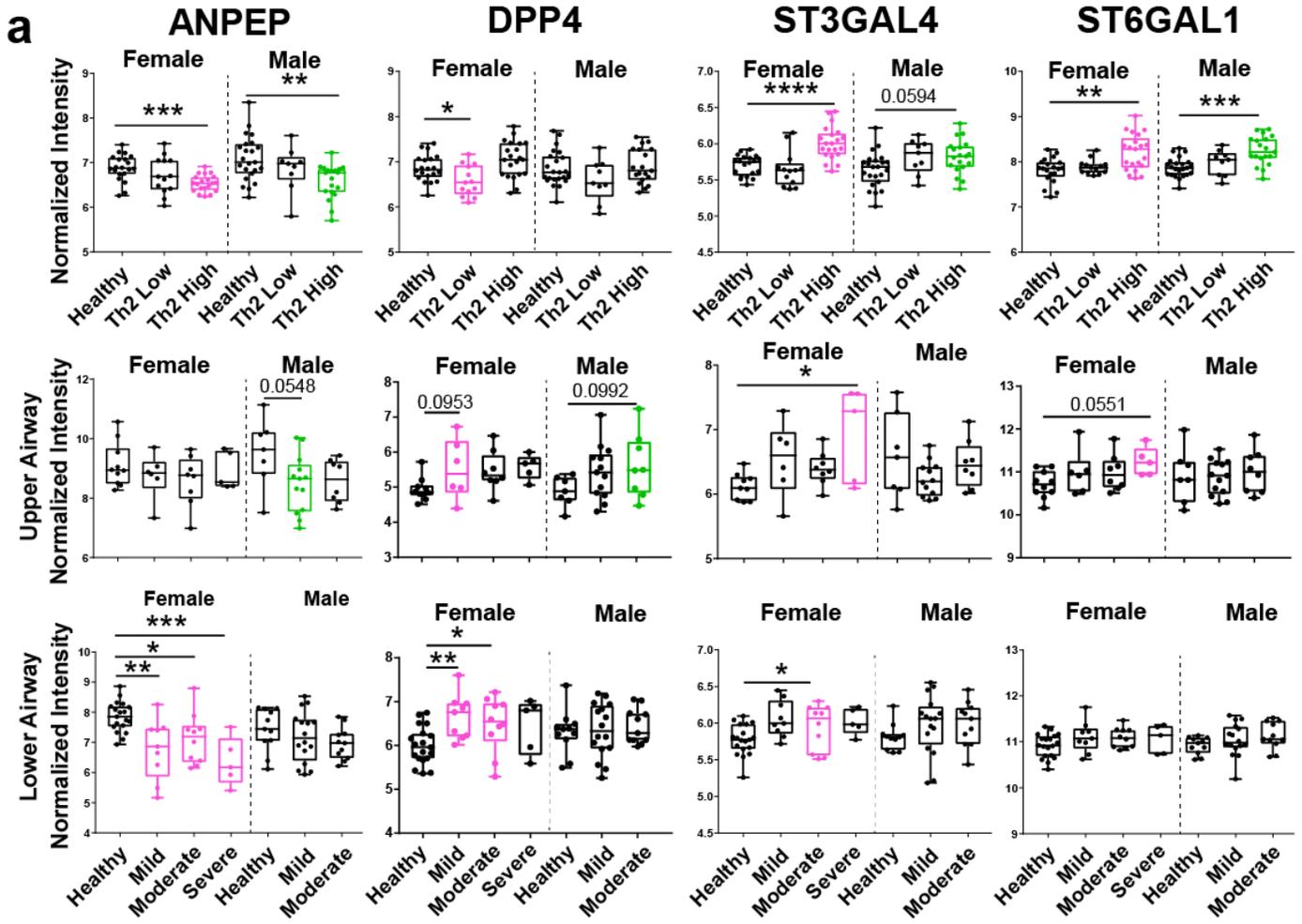
We revisited publicly available transcriptomic data from two independent studies of bronchial and nasal epithelium for an unbiased assessment of epithelial expression of COVID-19, influenza, and other human coronavirus related genes in (1) Th2 low versus Th2 high asthma<sup>6</sup> in bronchial epithelial brushings (GSE67472) and (2) expression of these genes in mild, moderate, and severe asthma in both upper and lower airways (GSE41861). We found that dysregulation of these genes in asthma is both sexually dimorphic and airway location dependent. For SARS-CoV-2 genes, we found marginally significant downregulation of viral receptor ACE2 in females in Th2 high asthma and significant downregulation in males in mild asthma (upper airway) (Fig. 1a). Expression of specific ACE2-partnering membrane protein SLC6A19 (B<sup>0</sup>AT1) was also found to be downregulated in males in mild asthma (lower airway) and moderate asthma (upper airway) (Fig. 1a). Interestingly, we found an opposite pattern of expression for two related membrane-bound serine proteases, TMPRSS2 and TMPRSS4, which promote SARS-CoV-2 virus epithelial entry by processing the spike glycoprotein resulting in the release of viral contents into the host cell cytosol. Expression of both proteases was significantly upregulated specifically in Th2 high asthma (Fig. 1a). We also found upregulation of



**Fig. 1. SARS-CoV-2 genes are dysregulated specifically in Th2 high asthma and are different by sex and airway location.** **a**, box and whisker plots showing normalized intensity values of COVID-19 related genes in healthy and asthmatic samples broken down by sex. Top row compares healthy, Th2 low, and Th2 high asthma. Middle row compares healthy, mild, moderate, and severe asthma in the upper airway. Bottom row compares healthy, mild, moderate, and severe asthma in the lower airway. **b**, PCA plot summarizes variation in virus entry genes and associated viral immunity genes (ANPEP, DPP4, CTSL, CTSB, ST3GAL4, ST6GAL1, IRAK3, IDO1, NOS2, OAS1, MX1, TNFSF10). Statistical significance was calculated using ANOVA with Tukey post-hoc analysis. In box and whisker plots, all data points are shown and the box represents the second and third quartiles and median. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ .

TMPRSS4 in females in severe asthma (lower and upper airway) and TMPRSS2 in males in mild and moderate asthma (lower airway) (Fig. 1a). Expression of another protease suspected in assisting viral entry, FURIN, was not found to be significantly different in any dataset (data not shown). Such differences due to sex deserve further investigation, as COVID-19 resulted in higher mortality and severity in males<sup>8</sup>. Downregulation of ACE2 expression predominantly in mild asthma is also interesting, as asthma is a very heterogeneous disease. The Center for Disease Control (CDC) currently lists moderate to severe asthma as a risk factor for severe COVID-19 illness (last updated on January 7). Whether the downregulation of ACE2 confers a protective effect in mild asthma as opposed to more severe forms of disease remains to be investigated. Another study showed that the use of inhaled corticosteroids in asthma was associated with lower expression of ACE2 and TMPRSS2<sup>9</sup>. Of note, severe asthma patients are the ones most likely to use corticosteroids, and yet we found ACE2 downregulation only in mild asthma. Principal Component Analysis (PCA) analysis of covariance in several human virus entry genes (ANPEP, DPP4, CTSL, CTSB, ST3GAL4, ST6GAL1) and associated viral immunity genes (IRAK3, IDO1, NOS2, OAS1, MX1, TNFSF10) demonstrated that dysregulation of viral immunity in general is characteristic of the Th2 high but not Th2 low subset of asthma patients (Fig. 1b).

We further performed a comparative analysis of SARS-CoV-2 genes with genes promoting the entry of other human coronaviruses and influenza. We found that ANPEP, the entry receptor for HCoV-229E, was significantly downregulated specifically in Th2 high asthma, as well as downregulated in females in mild, moderate, and severe asthma (lower airway) (Fig. 2a). Interestingly, HCoV-229E appears to be the human coronavirus least associated with asthma exacerbations. Studies have failed to find HCoV-229E in either adult or childhood episodes of asthma exacerbation<sup>10</sup>. However, asthma has been shown to be a risk for MERS-CoV (Middle Eastern Respiratory Syndrome) coronavirus infection. In one study, 31% of those with MERS had asthma<sup>11</sup>.



**Fig. 2. A comparative analysis of human coronavirus and influenza entry genes in Th2 high**

**inflammation. a**, box and whisker plots showing normalized gene expression intensity values of several viral species in healthy and asthmatic samples broken down by sex. Top row compares healthy, Th2 low, and Th2 high asthma. Middle row compares healthy, mild, moderate, and severe asthma in the upper airway. Bottom row compares healthy, mild, moderate, and severe asthma in the lower airway. **b**, box and whisker plots showing normalized intensity values of key epithelial subset marker genes in healthy, Th2 low, and Th2 high asthma patients. **c**, similarity matrix demonstrating unsupervised hierarchical clustering of SARS, COVID-19, HCoV-229E, MERS and influenza entry genes together with top epithelial subset genes from Nature Medicine study by Vieira Braga et al.<sup>7</sup> Higher levels of Th2 inflammation are associated with lower levels of epithelial differentiation (diagram below matrix). Significance in **a**. was calculated using ANOVA with Tukey post-hoc analysis. In box and whisker plots, all data points are shown and the box represents the second and third quartiles and median. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ .

DPP4, the entry receptor for MERS-CoV, was downregulated in females in Th2 low asthma (trending up in Th2 high asthma) and significantly upregulated in females in mild and moderate asthma (lower airway) (Fig. 2a). Two genes that promote the synthesis of receptors for human influenza strains, ST3GAL4 and ST6GAL1, were specifically upregulated in Th2 high inflammation in both sexes (Fig. 2a). SARS-Cov-1 (SARS) also relies on the ACE2 receptor as an entry mechanism. The SARS (SARS-CoV-1) outbreak also did not affect asthma exacerbations in children<sup>12</sup>.

Following the SARS outbreak in 2002-2003, multiple studies examined mechanisms of SARS infectivity in epithelial cells. Multiple studies confirmed that ACE2 receptor expression positively correlated with the differentiation state of bronchial epithelial cells. Importantly, undifferentiated human airway epithelial cells expressing low levels of ACE2 receptor were poorly infected with SARS-CoV-1, while differentiated cells with higher expression of ACE2 were more readily infected<sup>13</sup>. Given that asthma and other allergic diseases associate with loss of epithelial differentiation<sup>14</sup>, this prompted us to examine the possibility that Th2 inflammation modulates viral susceptibility by affecting the differentiation of epithelial cells. Few studies have looked at SARS-Cov-2 gene expression in airway epithelial cells by single cell RNA-seq, mostly at healthy baseline. At the protein and expression level, ACE2 primarily localizes to club (Clara) cells in bronchial epithelium and ATII in alveoli in the mouse lung<sup>15</sup>. An excellent study by the Human Cell Atlas Consortium<sup>16</sup> summarized ACE2 expression data across multiple single cell studies of human upper and lower airway. It showed that in nasal epithelium, ACE2 expression was the highest in ciliated and goblet cells; in the lower airway, the highest expression was found in club cells; and in lung parenchyma (alveoli), it was primarily found in ATII cells. Another human study reports the expression of ACE2 in the lung in club

and ciliated cells<sup>17</sup>. In a single cell study in *Macaca mulatta* and humans, ACE2 was also detected primarily in ATII cells but also in club cells<sup>18</sup>. Together, this shows that viral gene expression is restricted to specific epithelial cell subsets.

Unfortunately, there are no studies comparing viral genes in regards to epithelial heterogeneity across Th2 low and Th2 high asthma phenotypes. By comparing known markers of epithelial cell subsets between healthy, Th2 low, and Th2 high asthma, we found evidence for a remarkable reshaping of epithelial heterogeneity in Th2 high asthma (Fig. 2b). We found an increase in markers of basal cycling, basal activated, goblet, mucous ciliated cells, and ionocytes in allergic inflammation. At the same time, there was a significant decrease in markers of ciliated cells, submucosal glands, club (Clara) and tuft cells. This pattern of epithelial dysregulation was highly specific to Th2 high inflammation, confirming at high resolution loss of epithelial differentiation and specialized cell subsets expressing COVID-19 receptor (club cells, ciliated cells). To gain further insights whether expression of viral genes may correspond to epithelial heterogeneity pattern, we clustered top markers describing each epithelial subset (guided by a recently published single cell study characterizing novel airway epithelial cell subsets in asthma<sup>7</sup>) with entry genes of different viral species in “Th2 high vs. healthy control” gene expression analysis (Fig. 2c). Our analysis suggested that genes corresponding to different viral species may co-express with markers of different epithelial subsets. In particular, ACE2 and ANPEP were co-expressed with markers of club, tuft, ciliated cells, and submucosal glands, all downregulated during Th2 high (but not Th2 low) inflammation (Fig. 2c). TMPRSS2 expression very closely correlated with mucous ciliated cells, while TMPRSS4, DPP4 and influenza genes correlated with goblet and basal activated cells, all increasing in Th2 high inflammation (Fig. 2c). Although we acknowledge that this is only an exploratory correlation analysis, which requires further rigorous validation at the single cell analysis level, it does show importance of relating viral gene expression to epithelial heterogeneity. Such localization is also supported by single cell studies. This also explains why there is a complete discrepancy in the expression of ACE2 and TMPRSS2/TMPRSS4 in Th2 high inflammation, since these genes differentially follow epithelial subset divergence during inflammation. It is also very important to acknowledge that while ACE2 is a specific to certain human coronavirus species (SARS-CoV-1, SARS-CoV-2, and HCoV-NL63), TMPRSS2 and TMPRSS4 are certainly not. These proteases are important for cleavage of the spike protein of multiple pathogenic viral species, including MERS-CoV<sup>19</sup> as well as influenza strains<sup>20</sup>.

In summary, changes in epithelial specialization during Th2 high inflammation, as well sex, location and disease endotype may underlie differential susceptibility to SARS-CoV-2 and other

human coronavirus species. This warrants further investigation at the single cell level in asthma and other chronic inflammatory conditions.

## REFERENCES

1. Li, X., *et al.* Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. *J Allergy Clin Immunol* (2020).
2. Zhang, J.J., *et al.* Clinical characteristics of 140 patients infected with SARS-CoV-2 in Wuhan, China. *Allergy* (2020).
3. Mahdavinia, M., *et al.* Asthma prolongs intubation in COVID-19. *J Allergy Clin Immunol Pract* (2020).
4. Jackson, D.J., *et al.* Association of respiratory allergy, asthma, and expression of the SARS-CoV-2 receptor ACE2. *J Allergy Clin Immunol* (2020).
5. Kimura, H., *et al.* Type 2 Inflammation Modulates ACE2 and TMPRSS2 in Airway Epithelial Cells. *J Allergy Clin Immunol* (2020).
6. Woodruff, P.G., *et al.* T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am J Respir Crit Care Med* **180**, 388-395 (2009).
7. Vieira Braga, F.A., *et al.* A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat Med* **25**, 1153-1163 (2019).
8. Jin, J.M., *et al.* Gender Differences in Patients With COVID-19: Focus on Severity and Mortality. *Front Public Health* **8**, 152 (2020).
9. Peters, M.C., *et al.* COVID-19 Related Genes in Sputum Cells in Asthma: Relationship to Demographic Features and Corticosteroids. *Am J Respir Crit Care Med* (2020).
10. Talbot, H.K., *et al.* The pediatric burden of human coronaviruses evaluated for twenty years. *Pediatr Infect Dis J* **28**, 682-687 (2009).
11. Van Kerkhove, M.D., *et al.* Transmissibility of MERS-CoV Infection in Closed Setting, Riyadh, Saudi Arabia, 2015. *Emerg Infect Dis* **25**, 1802-1809 (2019).
12. Van Bever, H.P., Chng, S.Y. & Goh, D.Y. Childhood severe acute respiratory syndrome, coronavirus infections and asthma. *Pediatr Allergy Immunol* **15**, 206-209 (2004).
13. Jia, H.P., *et al.* ACE2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia. *J Virol* **79**, 14614-14621 (2005).
14. Schleimer, R.P. & Berdnikovs, S. Etiology of epithelial barrier dysfunction in patients with type 2 inflammatory diseases. *J Allergy Clin Immunol* **139**, 1752-1761 (2017).

15. Wiener, R.S., Cao, Y.X., Hinds, A., Ramirez, M.I. & Williams, M.C. Angiotensin converting enzyme 2 is primarily epithelial and is developmentally regulated in the mouse lung. *J Cell Biochem* **101**, 1278-1291 (2007).
16. Sungnak, W., *et al.* SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat Med* **26**, 681-687 (2020).
17. Zhang, H., *et al.* Expression of the SARS-CoV-2 ACE2 Receptor in the Human Airway Epithelium. *Am J Respir Crit Care Med* (2020).
18. Ziegler, C.G.K., *et al.* SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. *Cell* (2020).
19. Iwata-Yoshikawa, N., *et al.* TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection. *J Virol* **93**(2019).
20. Hatesuer, B., *et al.* Tmprss2 is essential for influenza H1N1 virus pathogenesis in mice. *PLoS Pathog* **9**, e1003774 (2013).

## ACKNOWLEDGEMENTS

We thank Drs. Tina Hartert (Vanderbilt University, Center for Asthma and Environmental Sciences Research) and Fernando Holguin (University of Colorado School of Medicine, Pulmonary Sciences and Critical Care) for thoughtful discussions. FUNDING: This work was supported by NIH NIAID R01AI127783 grant to Sergejs Berdnikovs and by the Ernest S. Bazley Foundation.

## AUTHOR CONTRIBUTIONS

S.B. conceived and designed the study. L.F.L. downloaded and prepared transcriptomic datasets for analysis. S.B. and M.E.C. analyzed data. S.B. and M.E.C. prepared the figures. S.B., H.A.V. and M.E.C. interpreted the results. The manuscript was written by S.B. and edited by H.A.V. and M.E.C. Final version of the manuscript was approved by S.B.

## COMPETING INTERESTS

The authors declare no competing interests.