Unique nasal cell states induced by common pediatric respiratory viruses

Jaclyn M. L. Walsh[1,2,3,4*], Lillian J. Juttukonda[1,4,5,6*^], Ying Tang[1*], Ankita Chatterjee[7*], Marc Elosua-Bayes[1,4,8], Erica Langan[1,3,4,9], Amanda Frischmann[1,4], Faith Taliaferro[1,4,10,11], Hannah R. Matthews[1,4,12,13], Kyle Kimler[1,4,8,14,15,16,17], Chris M. Lent[1,4], Deb Purna Keya[18], Preonath Chondrow Dev[18], Apurba Rajib Malaker[18], Lubana Tanvia[18], Arif Mohammad Tanmoy[18], Supratim Ghosh[19], Arindam Maitra[19], Apurba Ghosh[20], Surupa Basu[20], Abigail S. Kane[10,11], Charles A. Coomer[1,21,22], Alex K. Shalek[3,4,14,15,16,23,24], Thushan I. de Silva[25,26,27], Abdul Karim Sesay[27], Jessica Edwards[28], Corrine Sin Quee[29], Wanda I. Gonzalez[10,11], Lael M. Yonker[10,11], Sarah C. Glover[30#], Partha Majumder[19,31#], Yogesh Hooda[18#], Senjuti Saha[18#], Jose Ordovas-Montanes[1,3,4,14,24#^@], Bruce H. Horwitz[1,3,22#]

*Co-first author
[#]Co-senior author
^Co-corresponding author
[@]Lead contact

1. Division of Gastroenterology, Hepatology, and Nutrition, Boston Children's Hospital, Boston, MA, USA;
2. Department of Immunology, Harvard Medical School, Boston, MA, USA;
3. Program in Immunology, Harvard Medical School, Boston, MA, USA;
4. Broad Institute of MIT and Harvard, Cambridge, MA, USA;
5. Division of Newborn Medicine, Boston Medical Center, Boston, MA, USA;
6. Department of Pediatrics, Chobanian and Avedisian School of Medicine, Boston University, Boston, MA, USA;
7. John.C. Martin Center for Liver Research & Innovations, Kolkata, India;
8. Cell Discovery Network, Boston Children's Hospital, Boston, MA, USA
9. Harvard-MIT MD/PhD Program, Boston, MA, USA;
10. Department of Pediatrics, Massachusetts General Hospital, Boston, MA, USA;
11. Harvard Medical School, Boston, MA, USA;
12. Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA;
13. Program in Biological Sciences in Public Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA;
14. Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA, USA;
15. Institute of Medical Engineering & Science, MIT, Cambridge, MA, USA;
16. Department of Chemistry, MIT, Cambridge, MA, USA;
17. Division of Pediatric Hematology/Oncology, Boston Children's Hospital, Boston, MA, USA
18. Child Health Research Foundation, Dhaka, Bangladesh;
19. BRIC-National Institute of Biomedical Genomics, Kalyani, India;
20. Institute of Child Health, Kolkata, India;
21. Boston Combined Residency Program, Boston, MA, USA;
22. Division of Emergency Medicine, Boston Children's Hospital, Boston, MA, USA;
23. Koch Institute for Integrative Cancer Research, MIT, Cambridge, MA, USA;
24. Harvard Stem Cell Institute, Boston, MA, USA
25. Department of Infection, Immunity and Cardiovascular Disease, Medical School, University of Sheffield, Sheffield, UK;
26. Florey Institute of Infection and NIHR Sheffield Biomedical Research Centre, University of Sheffield, Sheffield, UK;
27. MRC Unit the Gambia at London School of Hygiene & Tropical Medicine, Serrekunda, The Gambia;
28. Department of Pathology/Microbiology, Princess Margaret Hospital, Nassau, Bahamas;
29. Division of Haematology and Oncology, Department of Pediatrics, Princess Margaret Hospital, University of the West Indies, Nassau, Bahamas;
30. Department of Medicine, Section of Gastroenterology and Hepatology, Tulane University, New Orleans, LA, USA;
31. Indian Statistical Institute, Kolkata, India;

Summary

Respiratory viral infections in early childhood are major drivers of acute morbidity and long-term airway disease, yet how distinct viruses remodel the pediatric nasal mucosa at cellular resolution remains unresolved. Here, we generated a single-cell RNA sequencing atlas of 335,174 nasal epithelial and immune cells from 132 children under five years of age with SARS-CoV-2, rhinovirus, or respiratory syncytial virus (RSV) infection, alongside uninfected controls. Mapping viral transcripts to individual cells revealed virus-specific infected epithelial states: an NF-kB-responsive ciliated subset in SARS-CoV-2 and a previously undescribed *KRT17+* squamous-like subset in RSV. We delineated divergent mucosal response programs, including a robust interferon (IFN) response in SARS-CoV-2, an IL-13-responsive secretory program in rhinovirus, and heightened inflammatory and cytotoxic immune activation in RSV. In RSV, specific immune subsets and elevated IFN-response signatures were associated with disease severity, whereas rhinovirus-induced wheeze was marked by expansion of a *CST1+* goblet cell subset. Integration of asthma genome-wide association data with our atlas revealed a *KRT13+* hillock-like squamous epithelial subset enriched for expression of childhood-onset asthma risk loci. Finally, we demonstrate that this resource enables high-resolution annotation of independent pediatric cohorts in Kolkata, India and rural Bangladesh. Together, this atlas establishes a comprehensive view of antiviral immunity in the pediatric nasal mucosa and defines virus-specific mucosal immune programs relevant to disease severity and asthma risk in early life.

Introduction

Respiratory viral infections are exceedingly common in early childhood and represent a major cause of morbidity and hospitalization worldwide. Among these pathogens, respiratory syncytial virus (RSV) and rhinovirus account for a substantial fraction of hospitalizations, wheezing symptoms, and lower respiratory tract disease in infants and young children. RSV alone is responsible for millions of hospital admissions globally each year in children under five[1–5]. In addition to their acute clinical impact, respiratory viral infections in infancy and early childhood have been linked to long-term respiratory outcomes, including recurrent wheeze and childhood-onset asthma[6–11]. Despite this substantial disease burden and long-term risk, there is limited understanding of the cell states and immune programs engaged in the pediatric airway during natural respiratory viral infection.

The nasal mucosa is a primary site of entry for most respiratory viruses and the location where host antiviral responses are first initiated[12,13]. At this interface, epithelial sensing, interferon (IFN) signaling, and immune cell recruitment shape early host-virus interactions that may influence subsequent immune responses and clinical outcomes. Studies of adult SARS-CoV-2 infection have shown that analysis of nasal mucosal responses reveals epithelial infection states and immune correlates of disease severity that are not evident from blood-based measurements alone, highlighting the nasal mucosa as a uniquely informative tissue for studying antiviral immunity at the site of infection[14–19]. However, the cellular organization of antiviral responses in the pediatric nasal mucosa during natural infection remains poorly characterized.

Although respiratory viral infections are ubiquitous in early childhood, different viruses are associated with distinct clinical phenotypes and disease trajectories in young children. Acute SARS-CoV-2 infection is typically mild during early childhood and is less likely to lead to severe respiratory illness or hospitalization; however, further studies are needed to understand its potential long-term, post-acute effects on children[20–23]. RSV is the most common cause of severe bronchiolitis, particularly in children younger than 6 months old[4,24]. Rhinovirus infection can also lead to bronchiolitis, but is most notable for its propensity to trigger childhood asthma exacerbations in susceptible hosts[9,11]. These divergent outcomes suggest that different viruses engage distinct host response programs at the mucosal interface. However, how epithelial and immune responses in the pediatric nasal mucosa differ across common respiratory viruses, and whether this heterogeneity relates to disease severity and symptoms such as wheezing, remains incompletely understood.

101 Progress in defining pediatric antiviral immunity has benefited from a range of complementary
102 experimental approaches. *In vivo* single-cell studies of respiratory viral infection in children,
103 largely focused on SARS-CoV-2 to date, have revealed unique features of pediatric immune
104 responses, such as enhanced IFN responses and unique T cell populations[25–27]. However, most
105 cohorts primarily include children older than 5 years or adolescents. *Ex vivo* air-liquid interface
106 (ALI) cultures and organoid models have enabled further characterization of age-dependent
107 antiviral responses[28–30], identification of molecular features such as nucleolin and IGF1R as
108 candidate receptors for RSV[31,32], IL-1β-driven mucus production in the absence of IFN signaling
109 in rhinovirus[33], and investigation of viral effects on epithelial differentiation pathways[34–36]. While
110 these systems provide important mechanistic insights under controlled conditions, they do not
111 capture the full cellular heterogeneity or immune context present *in vivo*. In parallel, analyses of
112 pediatric blood samples and tracheal aspirates from children with severe disease have revealed
113 immune correlates of clinical outcomes, but provide limited information about epithelial infection
114 states and early mucosal responses at the initial site of viral entry and do not capture the full
115 spectrum of severity[37–41]. A recent upper-airway transcriptomic study of infants hospitalized with
116 bronchiolitis of multiple viral etiologies, including RSV, identified compartmentalized mucosal
117 immune signatures, including impaired interferon signaling and IL-36–associated neutrophilic
118 inflammation, that correlate with disease severity but are not reflected in peripheral blood[42].
119 However, this study does not resolve epithelial infection states or define how epithelial and
120 immune responses are coordinated within the nasal mucosa. Consequently, the *in vivo* identity of
121 infected epithelial cells and the relationship between early nasal responses, disease severity, and
122 wheezing outcomes remain incompletely defined.

123 Single-cell atlases provide a powerful framework for resolving cellular responses to infection *in*
124 *vivo*, yet most existing atlases have been generated from cohorts in high-income settings[43–45].
125 This reflects the substantial logistical and technical challenges of generating high-resolution
126 single-cell datasets in low- and middle-income countries, where the burden of severe pediatric
127 respiratory infection is highest[3,5,46]. Whether cellular programs identified in single-cohort atlases
128 are conserved across populations with distinct environments and disease burdens therefore
129 remains largely untested. Addressing this gap will require the development of adaptable
130 experimental and analytical approaches, as well as collaborative frameworks that support single-
131 cell data generation and interpretation across diverse settings.

132 Here, we present a single-cell atlas of the pediatric nasal mucosa, profiling epithelial and immune
133 cells from children under five years of age with SARS-CoV-2, rhinovirus, or RSV infection,

134 alongside uninfected controls. By directly detecting viral transcripts within individual cells, we
135 define virus-specific infected epithelial cell states, and through comparative analysis characterize
136 distinct epithelial and immune response programs associated with each virus. We further relate
137 nasal mucosal cell states to clinical outcomes, including hospitalization and wheezing illness, and
138 contextualize genetic risk for childhood-onset asthma at the cellular level. Finally, we demonstrate
139 the broader relevance of this atlas by applying it as a reference to independent pediatric cohorts
140 from Kolkata, India, comprising healthy children, and rural Bangladesh, including healthy and
141 infected children, revealing conserved cell states and virus-associated epithelial responses
142 across geographically distinct populations. Together, this work provides a framework for
143 understanding virus-specific antiviral immunity at the pediatric mucosal interface and establishes
144 a broadly applicable reference for studying early-life respiratory disease.

145

146 ## Results

147 ### Developing an atlas of pediatric nasal mucosa antiviral responses

148 Respiratory viruses such as SARS-CoV-2, rhinovirus, and RSV are exceedingly common in young
149 children and are known to cause disparate disease outcomes. We hypothesized that each virus
150 might induce unique cellular responses in the pediatric nasal mucosa. To explore this, we
151 collected research nasopharyngeal (NP) swabs from a convenience cohort of children visiting
152 Boston hospitals (USA-Boston cohort) and performed single-cell RNA-sequencing (scRNA-seq)
153 on 132 swabs from children younger than 5 years of age selected by viral positivity and
154 representative of a range of clinical severity. 103 of the participants presented to the Boston
155 Children's Hospital (BCH) Emergency Department (ED) with acute respiratory symptoms. They
156 were assigned to SARS-CoV-2, rhinovirus, RSV, or co-infection groups based on results from
157 clinically indicated testing, laboratory-based RT-PCR of viral transport media from clinical
158 samples, and/or RT-PCR of supernatants collected during processing of research NP swabs
159 **(Figure 1A-B, Methods)**. We also enrolled children younger than 5 years of age who presented
160 to the BCH ED or an outpatient clinic at Massachusetts General Hospital (MGH) without
161 respiratory symptoms as controls. When available, supernatants obtained during processing of
162 research NP swabs were tested for influenza, SARS-CoV-2, rhinovirus, and RSV by RT-PCR.
163 Controls with positive PCR values for any of these viruses and symptomatic patients with
164 evidence of co-infections (PCR positive for more than one of SARS-CoV-2, rhinovirus, and RSV)
165 were included in iterative clustering to define cell subsets, but were excluded from any virus-
166 specific analysis. Our main cohort included 24 controls (BCH n=19, MGH n=5), 17 children with

167 SARS-CoV-2, 39 children with rhinovirus, and 40 children with RSV **(Figure 1A, Supplementary**
168 **Tables 1-2)**.
169
170 There were no significant differences in age, sex, race, or ethnicity across viral groups, except for
171 a higher proportion of males in the RSV group **(Figure 1C-D, Supplementary Table 1)**. Two
172 participants (1 SARS-CoV-2 case and 2 RSV) received RSV prophylaxis with the monoclonal
173 antibody nirsevimab prior to their first RSV season and the mothers of 3 participants (1 SARS-
174 CoV-2 and 2 RSV) had received RSV vaccination during pregnancy, though the majority of
175 participants were enrolled prior to FDA approval of either preventative treatment **(Supplementary**
176 **Table 1)**. There were no significant differences in family history of atopy (asthma, eczema, or
177 seasonal allergies), history of SARS-CoV-2 infection, or SARS-CoV-2 vaccination
178 **(Supplementary Table 1)**.
179
180 Acute SARS-CoV-2 infection generally causes mild symptoms in children[22]. In our cohort, children
181 with SARS-CoV-2 had shorter duration of illness and less severe symptoms than those with
182 rhinovirus or RSV **(Figure 1E, Supplementary Table 1)**. RSV is the leading cause of bronchiolitis
183 in young children and can cause both upper and lower respiratory disease[47]. In the group selected
184 for scRNA-seq, children with RSV presented later in illness than children with SARS-CoV-2 or
185 rhinovirus, had higher median modified BROSJOD bronchiolitis severity scores[48] **(see Methods)**,
186 more frequent receipt of supplemental oxygen, and a higher rate of hospital admission **(Figure**
187 **1E-F, Supplementary Table 1)**. Both rhinovirus and RSV are associated with persistent wheeze
188 or asthma later in childhood[11]. Based on our interest in this pathology, we included children with
189 wheezing on exam and/or treatment with systemic steroids, bronchodilators, or both in the groups
190 selected for scRNA-seq **(Figure 1G, Supplementary Table 1)**.
191
192 Viable epithelial and immune cells were isolated from each nasal swab as previously described
193 by our group[15,19]. After iterative removal of low-quality cells, pool-specific clusters, and doublets,
194 we obtained 335,174 high-quality cells. Because we did not observe notable differences in sample
195 quality nor effects of technical or demographic factors on gene expression **(Supplementary**
196 **Figure 1A-F, Methods)**, and because we wanted to preserve biological variability across viral
197 groups, we proceeded without data integration.
198
199 We performed iterative clustering to define the cell types and states represented in our dataset
200 **(Supplementary Figure 1G, Supplementary Figures 2-3, Supplementary Table 3)**. We

201 identified diverse epithelial populations, including basal stem cells (*KRT5, KRT15*), intermediary
202 club cells (*SCGB1A1, SCGB3A1*), mucus-producing goblet cells (*MUC5AC, BPIFA1*), and
203 squamous cells (*SPRR3*, *SPRR2E*) important in barrier function **(Figure 1H, Supplementary**
204 **Figure 2B)**[49–52]. There were also robust populations of ciliated epithelial cells (*FOXJ1*, *CAPS,*
205 *DNAAF1*), deuterosomal cells (*CCNO*, *CDC20B*)[53,54], and ionocytes (*CFTR, RARRES2*)[55,56]
206 **(Figure 1H, Supplementary Figure 2C)**. Each major epithelial cell type contained multiple
207 transcriptionally distinct subsets **(Figure 1I, Supplementary Figure 1G, Supplementary Figure**
208 **2, Supplementary Table 3)**. For example, within goblet cells we identified a subset marked by
209 *CST1*, a gene relevant in asthma pathobiology[57–60]. Squamous cells included a *KRT13+HOPX+*
210 subset resembling hillock cells, an important cell subset in repair of injured airways[55,61].
211 Deuterosomal cells separated into populations that were more ciliated-like (*MORN2, PIFO,*
212 *CROCC2*) versus those resembling goblet cells (*MUC5AC, SERPINB3*). Among ciliated epithelial
213 cells, which are expected to be the main target of viral infections, we observed multiple interferon-
214 stimulated (IFN-stim) subsets (*RSAD1, IFIT1, CXCL11*), an NF-kB responsive subset (*NFKBIA,*
215 *TNFAIP3*), and a subset exhibiting mixed ciliated and squamous-like features (*KRT17, S100A9,*
216 *LAMB3).* We observed two ciliated subsets defined by genes involved in cellular growth and cilia
217 development (*GRM7+ ESRRG+* ciliated, *MORN5+FOXJ1+* ciliated) which were enriched in
218 controls, indicating a consistent decline in healthy ciliated populations in all three infections
219 **(Figure 1J)**.
220
221 We also recovered diverse immune cell types, including macrophages (*TYROBP, FCER1G*),
222 conventional type-1 dendritic cells (*CLEC9A*), plasmacytoid dendritic cells (pDCs) (*IRF7, GZMB*),
223 B cells (*MS4A1, BANK1*), plasma cells (*MZB1, JCHAIN*), T cells (*CD3D, TRAC*), and NK cells
224 (*GNLY, PRF1*) **(Figure 1H, Supplementary Figure 3)**. Subclustering revealed additional rare
225 subsets, including monocyte-derived dendritic cells (MoDC) (*GPR183, CD86*), mucosal-
226 associated invariant T (MAIT) cells (*ZBTB16, KLRB1*, *SCL4A10*)[62,63], and lymphoid tissue inducer
227 (LTi) cells (*KIT, LTB*)[64]. T cells further separated into CD4+, CD8+, gamma-delta (γδ), and cycling
228 subsets **(Supplementary Figure 3A, 3C, see Methods)**. Across immune lineages, we observed
229 both quiescent or ribosomal-high states (e.g., Ribo-hi B cells, resting CD4+ T cells) and more
230 activated subsets (e.g., pro-inflammatory macrophages, *HAVCR2+CD38+* CD4+ T cells,
231 cytotoxic CD8+ T cells), highlighting the diversity of activation states captured in our samples
232 **(Figure 1I, Supplementary Figure 1G, Supplementary Figure 3, Supplementary Table 3)**.
233 Control samples had higher frequencies of *IL7R+KLRB1+* CD4+ T cells with a naive / memory

234 phenotype and transcriptionally quiescent CD8+ T cells, suggesting that upon infection, the
235 phenotype of nasal immune cells shifts to a more active inflammatory state **(Figure 1J)**.
236
237 Additional rare populations, such as melanocytes (*PMEL*, *MLANA,* n=122, 0.036% of all cells),
238 sustentacular cells (*ERMN, GPX6, CYP2A13*, n=113, 0.034% of all cells)[65], and neuronal-like
239 cells (*RBFOX1, CTNAP2, PTPRD*, n=1,830, 0.55% of all cells), were also detected **(Figure 1H-**
240 **I)**. In total, we defined 18 cell types and 79 cell subsets across the dataset. This represents the
241 most comprehensive description of nasal cellular diversity in young children to date, enabling
242 detailed exploration of the shared and unique responses to each respiratory virus.
243
244 **RSV and SARS-CoV-2 each induce highly specific infected epithelial cell states**
245 Nasal epithelial cells are a primary target of respiratory viruses. Given the distinct transmission
246 kinetics and symptoms associated with SARS-CoV-2, rhinovirus, and RSV, we hypothesized that
247 the intrinsic transcriptional response of infected epithelial cells may differ for each virus. We first
248 quantified viral RNA within each cell. Using our previously described method[15,19], we aligned our
249 transcriptional data to a custom reference including human, SARS-CoV-2[66], numerous rhinovirus
250 genomes across species A, B, and C **(see Methods)**, RSV-A[67], and RSV-B[68] genomes and
251 assigned cells as viral RNA+ if the viral counts were above the expected background amount
252 **(Supplementary Figure 4A, "alignment method")**. For SARS-CoV-2 and RSV, we recovered
253 significant amounts of viral RNA among epithelial cells **(Supplementary Figure 4B)**. For
254 rhinovirus, very few transcripts were detected **(Supplementary Figure 4B)**, and only aligned to
255 Rhinovirus C[69]. We detected 4,465 SARS-CoV-2 RNA+, 150 rhinovirus RNA+, and 8,767 RSV
256 RNA+ cells across the dataset. For SARS-CoV-2 and RSV, viral UMI detected in the single cell
257 data and the frequency of viral RNA+ cells for each participant correlated strongly to the qRT-
258 PCR CT value for viral RNA detected in the swab supernatant **(Supplementary Figure 4C)**. We
259 did not observe a correlation between viral transcripts detected and severity score for any of the
260 viruses **(Supplementary Figure 4C)**. Transcripts were detected across the SARS-CoV-2 and
261 RSV viral genomes, and for RSV, participants were consistently found to have reads either from
262 RSV-A or RSV-B **(Supplementary Figure 4D)**.
263
264 Because very few rhinovirus RNA+ cells were identified by alignment to custom reference
265 genomes, we considered that detection by this method may be hampered by the extensive viral
266 diversity within rhinovirus, which has three different species and more than 170 different
267 serotypes. As such, we applied the Kraken2[70] metagenomics method on a single cell basis to

268 quantify microbial reads in an unbiased manner **(Supplementary Figure 4A, "Kraken method")**.
269 There was a strong correlation between SARS-CoV-2 and RSV viral+ cells between the Kraken
270 and alignment methods **(Supplementary Figure 4E)**. While a few more participants had
271 detectable rhinovirus+ cells using this method, the frequency of viral RNA+ cells for rhinovirus
272 was much lower than RSV or SARS-CoV-2 by either method **(Supplementary Figure 4E)**. This
273 method also permitted detection of other common respiratory viruses, including parainfluenza
274 (276 positive cells total) and human metapneumovirus (HMPV) (141 positive cells total), across
275 multiple participant groups **(Supplementary Figure 4F)**. However, these detections were sparse
276 and HMPV detection was distributed across groups, and therefore we did not reclassify or exclude
277 participants based on these findings. Future studies with greater numbers of parainfluenza and
278 HMPV infected participants are warranted to permit detailed analyses of infected cells and
279 responses.
280
281 Next, we mapped viral RNA+ cells to defined epithelial cell subsets to evaluate the specific
282 phenotypes of presumably infected cells **(Figure 2A-B)**. SARS-CoV-2 and RSV viral RNA+ cells
283 occupied distinct areas of the epithelial UMAP **(Figure 2B)**, and a discrete epithelial subset
284 comprised the majority of viral RNA+ cells for both viruses. In SARS-CoV-2, the primary viral
285 RNA+ subset was *NFKBIA*+*ATF3*+ ciliated cells **(Figure 2C)**, which were also enriched in SARS-
286 CoV-2 cases **(Figure 2D)**. *KRT17*+*S100A9*+ ciliated cells were the primary infected cell subset in
287 RSV **(Figure 2C)** and were specific to RSV cases **(Figure 2D)**. A *CXCL11*+ IFN-stim ciliated
288 subset included cells containing transcripts for either SARS-CoV-2 or RSV **(Figure 2C)**, and was
289 increased in frequency compared to controls for both viruses **(Figure 2D),** although only a minority
290 of cells in this subset were viral RNA+ (**Figure 2C**). These results suggest that while some RSV
291 and SARS-CoV-2 infected cells share a similar phenotype (*CXCL11*+ IFN-stim) the majority of
292 epithelial cells infected by each virus exhibit distinct and unique transcriptional phenotypes.
293 Assessment of subsets comprising viral RNA+ cells identified via the Kraken method
294 demonstrated similar results, with broader detection in SARS-CoV-2 **(Supplementary Figure 4G)**
295 and detection of rhinovirus+ cells in the *CXCL11*+ IFN-stim ciliated subset. We observed minimal
296 detection of viral transcripts across immune subsets with either method, with the highest amount
297 in macrophages in samples obtained from patients in the SARS-CoV-2 group **(Supplementary**
298 **Figure 4H)**. In keeping with this, SARS-CoV-2 and RSV viral UMIs were lower in immune than in
299 epithelial cells **(Supplementary Figure 4B)**. Interestingly, the distribution of viral RNA+ cells did
300 not correlate with expression of known viral receptors or host processing genes for SARS-CoV-2
301 and RSV **(Supplementary Figure 4I)**, which may indicate a later stage of infection.

302

303 We further explored the phenotypes of the viral RNA+ cells. In addition to the negative regulators
304 of NF-kB signaling *NFKBIA* and *ATF3*, the SARS-CoV-2 specific subset expressed many
305 regulators of cell growth and apoptosis, including *ATF3, PMAIP1, and WEE1* **(Figure 2E)**. These
306 cells resembled a SARS-CoV-2 RNA-high subset defined by the circadian clock gene *PER1* and
307 growth regulator *EGR1* that we previously identified as consistently enriched in viral RNA+ cells
308 across SARS-CoV-2 variants[19], as well as a "hyperinfected" ciliated cell subset described in a
309 SARS-CoV-2 human challenge study[18]. Differentially expressed genes identified in these
310 previously described subsets shared substantial overlap with those identified within
311 *NFKBIA+ATF3+* ciliated cells in this dataset, indicating a conserved signature of SARS-CoV-2
312 infected nasal epithelial cells **(Supplementary Figure 4J)**.
313

314 The RSV-specific subset was defined by specific cytokeratins (*KRT17, KRT23*), as well as genes
315 associated with squamous differentiation and adhesion (*SPRR1B, LAMB3, TPM4*) **(Figure 2E)**.
316 Since ciliated cells are thought to be the main target of RSV, the downregulation of ciliated
317 markers and upregulation of squamous markers was unexpected. When we evaluated the
318 differentially expressed genes between SARS-CoV-2 RNA+ and RSV RNA+ cells within the
319 shared *CXCL11+* IFN-stim subset, we saw that the RSV RNA+ cells upregulated genes from the
320 same squamous-like program, including *LAMB3, SPRR1B, S100A9, and IGFBP6*, suggesting
321 even within the same subset, RSV+ cells adopt a unique transcriptional phenotype **(Figure 2F,**
322 **Supplementary Table 4)**. Despite the differences in phenotypes between RSV and SARS-CoV-
323 2 infected cells, RNA velocity analyses were consistent with the hypothesis that ciliated epithelial
324 cells are the precursors to both viral-specific populations **(Figure 2G)**.
325

326 Investigation of transcription factor (TF) activity demonstrated unique patterns in each viral RNA+
327 subset. The SARS-CoV-2 specific subset was enriched for transcription factors controlling cell
328 differentiation and apoptosis (*ATF2, KLF6*) **(Figure 2H)**. Activity of canonical IFN-responsive
329 transcription factors (*IRF1, IRF3, IRF7, IRF9*) was highest in *CXCL11+* IFN-stim ciliated cells.
330 Differential enrichment of STAT family members was observed, with higher inferred *STAT2*
331 activity in *NFKBIA+ATF3+* ciliated cells and higher *STAT1* activity in *KRT17+S100A9+* ciliated
332 cells **(Figure 2H)**. Given that *STAT1* and *STAT2* combined are required for IFNα or IFNλ
333 signaling, while *STAT1* homodimers form downstream of IFNγ, this result could indicate stronger
334 type-II IFN activation in RSV infected epithelial cells than in SARS-CoV-2 infected cells. The RSV-
335 specific *KRT17+S100A9+* subset was also uniquely enriched for transcription factors promoting

336 cell growth and epithelial-mesenchymal transition (*EHF*, *YAP1, FOSL1*) **(Figure 2H)**. Together,
337 these analyses demonstrate that SARS-CoV-2 and RSV each induce a distinct infected ciliated
338 cell state defined by virus-specific transcriptional regulatory programs.
339
340 **SARS-CoV-2 induces a more robust IFN response compared to rhinovirus and RSV**
341 The detection of viruses through pattern recognition receptors and subsequent production of IFNs
342 to warn neighboring cells are essential steps in containing viral infections. Our group and others
343 have demonstrated that in adults, a robust upper airway IFN response is associated with milder
344 COVID-19 symptoms[15,71]. Multiple groups have compared IFN responses to SARS-CoV-2
345 between children and adults, and suggested that one reason children tend to have milder COVID-
346 19 symptoms is a stronger initial IFN response[25,26,28]. We therefore wondered whether a robust
347 nasal IFN response was a universal feature of childhood respiratory infections, or if it varied
348 across viruses. We first examined expression of genes encoding pattern recognition receptors
349 (PRRs) and associated signaling proteins. We found that these genes were expressed in all three
350 viral RNA+ subsets, though to a lower degree in *KRT17+S100A9+* ciliated cells **(Supplementary**
351 **Figure 5A)**.
352
353 Next, we defined the main producers of IFNs in our dataset. Among epithelial cells, the viral RNA+
354 subsets were the only subsets in which we detected type I (*IFNB1, IFNE*) or type III (*IFNL1, IFNL2,*
355 *IFNL3)* interferons **(Figure 3A)**. When we evaluated the differentially expressed genes between
356 RSV RNA+ and RSV RNA- cells in the RSV-specific *KRT17+S100A9+* ciliated subset, we
357 similarly found that type III IFN genes were upregulated among viral RNA+ cells **(Figure 3B,**
358 **Supplementary Table 4)**. This analysis was not possible in the *NFKBIA+ATF3+* ciliated subset
359 because there were not sufficient numbers of viral RNA- cells. Among immune cells, a subset of
360 plasmacytoid dendritic cells (IFN-high pDCs) were the primary producers of type I and type III
361 IFNs **(Figure 3C)**. This subset, though rare, was highest in frequency in SARS-CoV-2 cases
362 **(Figure 3D)** and SARS-CoV-2 participants for whom this subset was present had higher viral
363 loads **(Supplementary Figure 5B)**. We next evaluated the response to type I IFN signaling using
364 a signature derived from IFNα cytokine treatment of nasal epithelial cells **(Supplementary Table**
365 **5)**[72,73]. Scoring pseudobulked epithelial cells from each participant with this signature revealed
366 significantly higher IFNα response pathway activity in SARS-CoV-2 cases compared to both
367 controls and other viral groups **(Figure 3E, Supplementary Figure 5C)**. The epithelial IFNα
368 response was significantly correlated with viral load only in RSV cases, and was not associated
369 with age, duration of infection at time of sampling, or disease severity **(Figure 3F)**.

370

371 To assess the response to each virus in a more holistic manner, we turned to compositional
372 methods that enable quantification of the cellular ecosystem of each sample[74,75]. Principal
373 component analysis based on the cellular compositions revealed that rhinovirus cases were most
374 similar to controls, RSV cases overlapped with both rhinovirus and controls but had some
375 separation along PC1, and SARS-CoV-2 cases were the most distinct from other samples
376 compositionally **(Figure 3G)**. The subsets defining control cases (positive PC1, negative PC2)
377 included basal cells, healthy ciliated cells, and resting immune populations as expected **(Figure**
378 **3H)**. The main drivers of variation separating the RSV cases (negative PC1) were the
379 *KRT17+S100A9+* ciliated cells, along with multiple macrophage and T cell subsets, suggesting a
380 heightened inflammatory response **(Figure 3H)**. The cellular composition of SARS-CoV-2 cases
381 (positive PC2) were defined by multiple IFN-responsive subsets **(Figure 3H)**, including the
382 *CXCL11+* IFN-stim and *NFKBIA+ATF3+* virally infected ciliated subsets. When we compared the
383 frequency of IFN-responsive subsets identified through the compositional analysis, we observed
384 that these subsets were either specific to SARS-CoV-2 (*GSTA2*+IFN-stim ciliated and
385 *SCGB1A1*+IFN-stim club), increased in frequency in all viruses but to the greatest extent in
386 SARS-CoV-2 (IFN-stim goblet, IFN-stim CD8+ T, IFN-stim γδ T), or in one case non-specifically
387 enriched in all three viruses (IFN-stim ionocytes) **(Figure 3I)**.
388

389 Noting that some IFN-responsive subsets were virus-specific while others were shared between
390 viruses, we hypothesized that different transcription factors were activated across subsets. We
391 next evaluated the use of key antiviral transcription factors across IFN-stim cell subsets.
392 Transcription factors downstream of viral pattern recognition and IFN signaling, including *IRF3,*
393 *IRF7, IRF1, IRF2,* and the ISGF3 complex (*STAT1/STAT2/IRF9*), showed the highest inferred
394 activity in the same cell subsets, particularly IFN-stim goblet cells, IFN-stim CD8+ T cells,
395 *CXCL11+* IFN-stim ciliated cells, and the SARS-CoV-2-high *NFKBIA+ATF3+* ciliated cells **(Figure**
396 **3J)**[76]. In contrast, the RSV-high subset *KRT17+S100A9+* ciliated cells and SARS-CoV-2-specific
397 bystander population *GSTA2+* ciliated cells had less activity of most IFN-response regulators
398 except for *STAT1*. IFN-stim CD8+ T cells showed specific enrichment of *STAT4*, *IRF5*, *IRF6*, and
399 *IRF8* activity, transcription factors previously implicated in IFN signaling, cytotoxic effector
400 differentiation, and antiviral T-cell responses **(Figure 3J)**[76,77]. *STAT3* activity, which upregulates
401 proinflammatory cytokine responses, was strongest in IFN-stim goblet and the RSV-high
402 *KRT17+S100A9+* ciliated cell subsets **(Figure 3J)**.
403

404 We observed positive correlations between the abundance of multiple IFN-stim cell subsets and
405 viral RNA levels in the sample, although the pattern differed by virus. In SARS-CoV-2, the viral
406 RNA+ *NFKBIA+ATF3+* and *CXCL11+* IFN-stim ciliated subsets correlated with viral load,
407 whereas in rhinovirus only IFN-stim goblet cells showed such an association **(Figure 3K)**. RSV
408 demonstrated the greatest number of significant correlations between IFN-stim subsets and viral
409 load **(Figure 3K),** further indicating that IFN activation in RSV is closely linked to the amount of
410 virus present. Taken together, our data indicate that SARS-CoV-2 induces a significantly more
411 robust productive IFN bystander response than either rhinovirus or RSV, which may reflect a more
412 amplified antiviral state that is not solely proportional to viral burden. These results suggest that
413 young children have a graded IFN response dependent on the specific viral challenge.
414
415 **Rhinovirus infection exhibits secretory-cell polarization of the epithelium**
416 Given the diminished IFN responses in rhinovirus in comparison to RSV and particularly in
417 comparison to SARS-CoV-2, we sought to further understand the epithelial response specific to
418 rhinovirus. We shifted from subset-level frequencies to analysis at the broader epithelial cell-type
419 level for each virus compared to controls, using centered log-ratio (CLR)-normalized abundance
420 to assess compositional changes[74]. We found that in all three viruses, there was an increase in
421 abundance in goblet cells **(Figure 4A)**. In rhinovirus cases, there was a notable depletion of
422 ciliated epithelial cells and an increase in deuterosomal cells **(Figure 4A)**. Because ciliated cells
423 are the primary targets of rhinovirus infection, we hypothesized that receptor-expressing cells
424 might be selectively lost. While *ICAM1* and *LDLR*, the receptors used by rhinovirus-A and
425 rhinovirus-B major and minor groups, respectively, were broadly expressed at low levels, CDHR3,
426 the receptor for rhinovirus-C, was most highly expressed in ciliated cells **(Figure 4B)**. Consistent
427 with this distribution, rhinovirus cases exhibited a significant reduction in *CDHR3+* epithelial cells
428 compared to controls and other viral infections **(Figure 4C)**. These findings suggest that
429 rhinovirus-infected CDHR3+ ciliated cells may be lost through cell death or are underrepresented
430 in our dataset, potentially explaining the limited detection of rhinovirus RNA+ cells.
431
432 Type 2 cytokines IL-4 and IL-13 are known to increase secretory and goblet cell differentiation.
433 Furthermore, rhinovirus induces a strong type 2 response in some individuals [78–83]. When we
434 evaluated response signatures for these cytokines **(Supplementary Table 5)**[72,73], we observed a
435 stronger IL-13 response in comparison to the IL-4 response **(Supplementary Figure 5C)**.
436 Comparison across viral groups revealed a modest but significant increase in IL-4 response
437 scores in rhinovirus cases compared to RSV cases and controls **(Figure 4D)**. IL-13 response

438 scores were notably higher in both rhinovirus and RSV cases compared to controls and SARS-
439 CoV-2 cases **(Figure 4D)**. To evaluate the balance between type 1 and type 2 immune responses
440 (represented by IFNα and IL-13, respectively) in each participant, we calculated the difference in
441 the z-scores for each response signature. We observed a gradient in this ratio, with a strong
442 enrichment in IFNα response in SARS-CoV-2, as noted in Figure 3, an even ratio for RSV, and
443 an enrichment in IL-13 responses for rhinovirus **(Figure 4E)**. Epithelial IL-13 response was not
444 significantly associated with viral load, age, duration of illness, or severity **(Supplementary**
445 **Figure 5D)**.
446
447 We next evaluated which subsets were the main producers of IL-13. Although detection in general
448 was quite low, we found that mast cells had the highest expression, followed by *TNFRSF4*+ CD4+
449 T cells **(Figure 4F)**. Ranking of epithelial cell subsets by their IL-13 response score showed that
450 *KRT13*+*HOPX*+ squamous cells, the subset we noted resembles hillock cells, had the strongest
451 score **(Figure 4G, Supplementary Figure 5C)**. These were followed by multiple goblet subsets
452 and *MUC5AC*+ deuterosomal cells, indicating that IL-13 is likely acting on secretory cells in the
453 epithelium. As deuterosomal cells are thought to be a differentiation intermediate towards the
454 ciliated cell[53,54], the expression of *MUC5AC* was unexpected and could indicate goblet cell
455 polarization within this population. When we compared the frequency of the subsets with the
456 highest IL-13 scores, we found that many were enriched in rhinovirus cases, including
457 *BPIFA1*+*MUC5B*+ goblet cells, developing goblet cells, *MUC5AC*+ deuterosomal cells, and
458 *CST1*+ goblet cells **(Figure 4H, Supplementary Figure 5E).** Together, these results indicate a
459 preferential secretory polarization during the immune response to rhinovirus infection.
460
461 **RSV is characterized by a heightened inflammatory and cytotoxic response**
462 Immune cell infiltration is a common feature following respiratory viral infection of the nasal
463 epithelium. However, particularly in RSV infection, the balance between protective and
464 pathogenic inflammation may be determined by the quantity and phenotype of immune cells
465 recruited to the airway[24]. Severe RSV infection has been linked to elevated numbers of
466 neutrophils, monocytes, NK cells, T cells, and inflammatory cytokines such as IL-6 and IL-8 in the
467 airways, as well as lower levels of type I and type II IFNs in blood and nasal secretions[40,84–88]. As
468 such, we hypothesized that the nasal cellular immune response to RSV may be distinct from that
469 of other viruses. We first assessed overall changes in CLR-normalized immune cell type
470 abundance, and observed that in both rhinovirus and RSV cases, plasmacytoid dendritic cells
471 (pDCs), NK cells, and cycling T cells were significantly increased **(Figure 5A)**. In contrast, there

472 were no significant increases in immune cell abundance in SARS-CoV-2 cases (**Figure 5A)**. Next,
473 we explored which cytokines, chemokines, and other immune signaling ligands **(see Methods,**
474 **Supplementary Table 6)** were significantly upregulated compared to controls for each virus. We
475 found that macrophages and CD8+ T cells were the strongest sources, and that despite limited
476 changes in immune cell abundance, SARS-CoV-2 cases still had significant upregulation of many
477 immune ligands **(Figure 5B, Supplementary Table 6)**.
478
479 We next sought to broadly understand differences in cell-cell communication between groups.
480 We applied multinichenet to rank the ligand-receptor interactions by 1) differential expression in
481 each group, 2) consistency of expression across samples, 3) ligand activity enrichment in the
482 receiver cell types, and 4) cell type specificity[89]. We found that of the top 50 ranked interactions,
483 the majority were in the RSV group **(Figure 5C, Supplementary Figure 6A, Supplementary**
484 **Table 6)** and that this was consistent across cutoffs **(Supplementary Figure 6B)**. SARS-CoV-2
485 cases had the second-highest number of significant interactions and consistently strong
486 prioritization scores **(Supplementary Figure 6C)**. To further characterize the RSV-specific
487 interactions, we selected those involving the top sender cell types, which included macrophages,
488 γδ T, NK, ciliated, CD4+ T, and CD8+ T cells **(Supplementary Figure 6D)**. Notably, macrophages
489 emerged as a central communication hub, participating in a large fraction of the top-ranked
490 interactions. The dominant signals included cytokines (*TNF, IL15, IL10, TNFSF13B*), IFN and
491 IFN-related genes (*IFNG, CXCL10, IFITM1-CD81* interactions), and immune regulatory or
492 adhesion-associated molecules (*SIRPA, ICAM1, SIGLEC1*). Although ligand activity scores were
493 often comparable across groups, RSV cases consistently showed the highest expression of the
494 corresponding ligand–receptor pairs **(Figure 5D)**. Together, these findings indicate that RSV
495 infection is marked by enhanced macrophage-centered inflammatory and interferon signaling,
496 with greater immune-immune and immune-epithelial communication than SARS-CoV-2 or
497 rhinovirus.
498
499 We next explored which cell subsets change in concert with each other. Correlation of the CLR-
500 normalized abundances of all cell subsets[75] revealed several clusters representing groups of
501 coordinated cell subsets **(Figure 5E, Supplementary Figure 6E)**. A cluster of interest contained
502 both RSV-infected ciliated cell subsets (*KRT17+S100A9+* and *CXCL11*+IFN-stim) as well as
503 multiple inflammatory macrophage, NK cell, and T cell subsets **(Figure 5F, Supplementary**
504 **Figure 6E)**, demonstrating that these immune cell subsets increased in relative abundance
505 together with RSV RNA+ epithelial cells. Many of these subsets were also upregulated in RSV

506 cases, including *ZFAT*+ pDCs, *S100A8*+*CCL2*+ macrophages, *NCAM1*+ NK cells, ribosomal-high
507 NK cells, activated cycling CD4+ T cells, cytotoxic CD8+ T cells, *CCL4*+ NK cells,
508 *PLCG2*+*KDM4B*+ macrophages, *SPP1*+*HK2*+ macrophages, and *HAVCR2*+*CD38*+ CD4+ T
509 cells **(Figure 5G, Supplementary Figure 6F)**. These results suggest that RSV is associated with
510 a distinct network of activated, cytotoxic, and inflammatory immune cells that is not observed in
511 SARS-CoV-2 or rhinovirus.
512
513 **Distinct cellular programs in severe RSV infection and rhinovirus-induced wheeze**
514 RSV and rhinovirus cause a broad spectrum of disease phenotypes in young children, while
515 SARS-CoV-2 is most commonly associated with a mild upper respiratory infection or febrile illness
516 in this population. RSV can cause severe lower respiratory tract infection in otherwise healthy
517 young children, while rhinovirus is strongly associated with viral-induced wheeze and asthma
518 exacerbations in genetically-susceptible individuals. However, it is not clear whether the nasal
519 responses to either virus correlate with disease severity or wheezing response. We identified
520 several significant correlations between nasal cell subset abundance and specific clinical
521 variables **(Figure 6A)**. Of note, when participants were split by viral group, most associations did
522 not retain significance **(Supplementary Figure 7)**. Resting and Trm-like CD8+ T cells positively
523 correlated with age, consistent with our previous findings[90]. In keeping with the known association
524 between IL-13 responses and asthma, two goblet cell populations that we identified as being IL-
525 13 responsive (*BPIFA1*+*MUC5B*+ and *CST1*+) were associated with wheezing symptoms and
526 treatments **(Figure 6A)**. Mast cells and a *CXCL13*+ CD8+ T cell subset also correlated with
527 wheezing symptoms and treatments **(Figure 6A)**. Multiple RSV-associated subsets were
528 associated with disease severity and respiratory support, including *KRT17*+*S100A9*+ ciliated
529 cells, activated cycling CD4+ T cells, cytotoxic CD8+ T cells, and ribosomal-high NK cells **(Figure
530 6A)**, in keeping with the RSV participants having generally more severe illness.
531
532 We next compared the nasal mucosa responses to RSV between participants who were
533 discharged from the ED, admitted to the hospital for general care on the floor, or admitted to the
534 ICU for intensive respiratory support. ICU cases had higher overall viral loads in the nasal swab
535 sample **(Figure 6B)** and trends towards higher amounts of viral RNA detected at the single-cell
536 level and frequency of viral RNA+ cells **(Supplementary Figure 8A-B)**. We next examined the
537 cellular composition PCA (as in **Figure 3F-G**) by the RSV admission groups. Participants who
538 were admitted to the ICU separated along PC2, suggesting increased abundance of IFN-
539 responsive subsets **(Figure 6C, Figure 3G)**. RSV cases who were admitted to the ICU were

540 characterized by increased frequencies of MT-high goblet cells, *CXCL11*+ IFN-stim ciliated cells,
541 *ZFAT*+pDCs, *S100A8*+ macrophages, and ribosomal-hi NK cells, as well as a decline in resting
542 CD8+ T cells **(Figure 6D)**. As the MT-high goblet cells and *CXCL11*+ IFN-stim ciliated cells both
543 contain some RSV RNA+ cells **(Figure 2C)**, this result suggests greater RSV dissemination
544 amongst the epithelium in more severe disease. Further investigation of differentially expressed
545 genes in T cells between RSV cases admitted to the floor and intensive care demonstrated
546 upregulation of many IFN-stimulated genes (*IFIT1, IFITM1, IFITM3, ISG15, IFI35*) in ICU cases
547 **(Figure 6E)**. Gene activity signatures **(Supplementary Table 5)** of IFNα and IFNγ were also
548 significantly higher in ICU cases compared to either discharged or admitted to floor participants
549 **(Figure 6F)**. We did not observe a particular pattern in cellular composition nor any changes in
550 IFN response scores in rhinovirus cases divided by admission status **(Supplementary Figure**
551 **8C-D)**. Together, these results indicate that severe RSV disease is characterized by a high nasal
552 viral load at the time of ED presentation, a strong type I / type II IFN response, and increases in
553 inflammatory immune subsets in the nasal mucosa.
554
555 Rhinovirus and RSV are both capable of eliciting wheezing, but rhinovirus is more strongly
556 associated with asthma exacerbations in children and adults[91]. Therefore, we hypothesized that
557 the nasal cellular responses in rhinovirus participants with wheezing symptoms would be distinct
558 from RSV participants with wheezing symptoms. To further understand cell subsets that may be
559 associated with wheeze, we first evaluated the cellular composition of rhinovirus and RSV cases
560 divided by wheeze group. We compared participants with rhinovirus or RSV who did not wheeze
561 on exam and did not receive bronchodilators or steroids during their ED visit (No wheeze) with
562 participants who had wheeze on exam and received bronchodilator and/or steroid therapy in the
563 ED (Wheeze + treatment). Of note, swabs were obtained at any point during the ED stay and the
564 specific time was not recorded relative to receipt of medical therapies. We noted that the
565 rhinovirus wheeze+treatment cases were the most similar to each other compositionally, with
566 greater variance in the other groups **(Figure 6G)**. We applied a machine learning method to build
567 a model to predict whether a rhinovirus sample belonged to the wheeze group and evaluated the
568 importance of each cell subset in this prediction. Both immune and epithelial subsets were
569 identified as important for the model, including subsets that were increased, decreased, or
570 unchanged in wheezing participants **(Figure 6H)**. A similar analysis in RSV cases revealed that
571 some subsets were important for distinguishing wheeze for both viruses, including
572 *XCL1/2*+*CTRAM*+ γδ T cells, neuronal cells, and *CXCL13*+ CD8+ T cells, but that the majority of
573 subsets identified for the RSV and rhinovirus wheezing models were specific to each virus

574 **(Supplementary Figure 8E).** We found that multiple IFN-responsive cell subsets, including IFN-
575 stim CD8+ cells, IFN-stim γδ T cells, and IFN-stim goblet cells, were decreased in frequency in
576 rhinovirus wheeze cases, but not for RSV wheeze cases **(Figure 6I)**. Furthermore, rhinovirus
577 participants with wheezing symptoms had overall lower scores for IFNα and IFNγ
578 **(Supplementary Figure 8F)**. This may indicate that the bias toward an IL-13-responsive
579 secretory state and away from an IFN-response epithelium observed in rhinovirus **(Figure 4)** is
580 particularly pronounced in participants with wheeze.
581
582 As many participants in the wheeze group received systemic corticosteroids, we sought to
583 determine the impact of corticosteroid treatment on nasal responses. In both rhinovirus and RSV
584 wheeze group cases, epithelial and immune cells scored higher for a corticosteroid response
585 signature **(Supplementary Table 5)**[92] compared to the non-wheeze group **(Supplementary**
586 **Figure 8G)**. We compared cell subsets from participants from all viral groups who received
587 corticosteroid treatment (systemic steroids in the ED or systemic or inhaled steroids in the week
588 prior to presentation) with participants who did not. While IFN-stim goblet cells decreased in
589 frequency in participants who had been treated with steroids, multiple cell subsets were enriched,
590 including *BPIFA1+MUC5B+* goblet cells, mast cells, and *CXCL13+* CD8+ T cells
591 **(Supplementary Figure 8H)**. As *CXCL13* is typically expressed in CD4+ T peripheral helper cells
592 (Tph), we confirmed that this subset expressed CD8+ T cell markers and was distinct from the
593 *CXCL13*-expressing Tph subset also identified in this dataset **(Supplementary Figure 8I)**. The
594 *CXCL13+* CD8+ subset was distinguished by *CXCL13* expression as well as upregulation of the
595 steroid receptor *FKBP5*, suggesting that corticosteroid treatment may upregulate this chemokine
596 in CD8+ T cells and alter B cell recruitment.
597
598 The steroid analysis suggests that some subset changes in wheeze groups may be due to
599 corticosteroid treatment in the wheeze groups. However, cycling basal cells, *MUC5AC+*
600 deuterosomal cells and *CST1+* goblet cells were uniquely enriched in the rhinovirus wheeze
601 cases **(Figure 6I)**. Intriguingly, *CST1* was recently identified as a nasal transcriptomic marker of
602 type 2-high asthma in adolescents[93], is increased in induced sputum from adults with eosinophilic
603 asthma[57], and is upregulated in the airways of children with atopy[94]. We anticipated that *CST1+*
604 goblet cells express a transcriptional program important for type 2-polarized responses to
605 rhinovirus. Indeed, the top three transcription factor activities were *STAT6*, *JUN*, and *AP1* **(Figure**
606 **6J)**. *STAT6* is a master regulator of type 2 inflammation[95], and *JUN* and *AP1* complexes regulate
607 type 2 inflammation with a particularly important role in airway reactivity[96]. Furthermore, multiple

608 markers of this cell subset (*CST1, ALOX15, TFF3, SERPINB3,* **Figure 6K**) are upregulated in
609 polyps from individuals with chronic rhinosinusitis[73]. Together, our results strongly implicate
610 *CST1*+ goblet cells as an important cell state connecting rhinovirus responses to wheezing
611 symptoms via local type 2 inflammation in the nose. Collectively, these findings highlight the
612 capacity of this approach to discern phenotype-specific epithelial and immune programs *in vivo*,
613 informing mechanistic hypotheses regarding the pathogenesis of severe RSV infection and
614 rhinovirus-associated wheeze.
615
616 **Pediatric nasal cell atlas enables contextualization of genetic risk factors for childhood-**
617 **onset asthma**
618 Early childhood viral infections, particularly with rhinovirus or RSV, have been linked to the onset
619 and later development of asthma[11]. At the population level, genome-wide association studies
620 (GWAS) have identified numerous genetic risk loci associated with asthma; however, these
621 studies do not indicate the specific cell types in which risk genes are active. This missing cellular
622 context limits our understanding of how the combination of genetic risk and viral infection
623 contributes to asthma development. We therefore asked whether our atlas of pediatric viral
624 infection, integrated with GWAS data, could help address this gap. We applied the single-cell
625 disease relevance score (scDRS) method, which scores individual cells based on their expression
626 of genes implicated in GWAS studies[97]. For all cells in our dataset, we calculated the scDRS for
627 childhood- and adult-onset asthma traits (COA and AOA, respectively), as well as height as a
628 control trait. Squamous epithelial cells had the highest disease relevance scores for childhood-
629 onset asthma, followed by goblet cells **(Figure 7A-B)**. Interestingly, very few immune cells had
630 significant enrichment of relevant genes for childhood-onset asthma in comparison with epithelial
631 cells. In contrast, fewer epithelial cells but a greater number of immune cells had significant
632 disease relevance scores for adult-onset asthma, with the highest scores in CD4+ T cells **(Figure**
633 **7A-B, Supplementary Figure 9A)**. These results highlight the importance of the airway
634 epithelium in childhood-onset asthma risk in comparison with adult-onset asthma risk.
635
636 Next, we investigated the specific cell subsets which had the highest disease relevance scores.
637 Among epithelial subsets, the score for childhood-onset asthma was strikingly enriched in
638 *KRT13+HOPX*+ squamous cells **(Figure 7C, Supplementary Figure 9B)**. This subset expressed
639 high levels of many cytokeratins (*KRT6A, KRT13, KRT78*), squamous markers (*SPRR2A*,
640 *SPRR3*), and genes involved in epithelial remodeling or barrier integrity (*TMPRSS11E*, *PRSS27,*
641 *CST6, ECM1*) **(Supplementary Figure 9C)**. Notably, this same subset had the highest IL-13

642 response score **(Figure 4G)**. The *KRT13+HOPX+* squamous cells in our dataset resemble
643 "hillock" cells described as important progenitors in mouse and human airway injury[55,61],
644 suggesting that epithelial repair programs may overlap with asthma risk genes. To understand
645 which genes were driving the disease relevance scores, we evaluated which genes had both a
646 correlation with the disease relevance score and a positive gene weight, which serves as a
647 correlate of GWAS risk. The genes contributing the most to the childhood onset asthma score in
648 epithelial cells included many of the marker genes for the *KRT13+HOPX+* squamous subset
649 **(Supplementary Figure 9D)**, demonstrating the power of this approach to identify genes with
650 individually weak contributions to genetic risk that when expressed in concert in a single cell may
651 have important contributions to disease pathogenesis. Additionally, many T cell-relevant genes,
652 including *ICOS, RORA,* and *CD247*, contributed the most to the adult-onset asthma score in
653 immune cells. Together, these results demonstrate the utility of our dataset in contextualizing
654 asthma GWAS data in terms of nasal cellular biology, and reveal that "hillock"-like squamous
655 epithelial cells may be of particular importance in childhood onset asthma, consistent with a recent
656 report[98].
657
658 **Pediatric nasal cell atlas can be used as a reference for independent datasets**
659 Geographic environment and genetic ancestry are known to contribute significantly to variation in
660 cellular and tissue biology, leading to concerted efforts among human cell atlases to diversify the
661 participants included. Additionally, while RSV infection is almost ubiquitous among children
662 globally, the vast majority of childhood mortality from RSV takes place in low- and middle-income
663 countries. Therefore, we sought to understand whether our atlas could be used as a viable
664 reference in scenarios where the population, environment, and disease burden may be very
665 different than in Boston, USA. To do this, we established independent participant cohorts in the
666 area surrounding Kolkata, India (IND-Kolkata cohort, n=35, all healthy controls) and rural
667 Bangladesh (BD-rural cohort, n=19, mix of controls and patients with respiratory infection) **(Figure**
668 **7D)**. scRNA-seq library capture and sequencing was done locally using Honeycomb technology
669 at each site, an important step in building capacity for future scRNA-seq studies without dedicated
670 instrumentation.
671
672 First, we compared nasal cells from healthy children in the IND-Kolkata cohort (metadata provided
673 in **Supplementary Table 2)** to the healthy controls in the USA-Boston cohort. After generating a
674 dataset of 41,721 quality cells **(Supplementary Figure 10A-D)**, we transferred labels separately
675 from epithelial and immune cells. Predicted epithelial labels included the majority of expected cell

types, including ciliated, goblet, basal, club, squamous, and deuterosomal cells, as well as ionocytes, and multiple cell subset labels were transferred for most cell types **(Figure 7E-F)**. We also recovered both common (T cells, B cells, macrophages) and rare (plasma cells, mast cells, pDCs, LTi cells) immune cell types **(Figure 7G)**, and many immune subsets were represented **(Figure 7G-H)**. The majority of cells in our dataset had prediction scores close to 1, though the distribution was slightly wider in immune cells and varied by cell type **(Supplementary Figure 10E-F)**. Analyzing the frequency of predicted cell types across the individual samples revealed consistent detection of both epithelial and immune cell types **(Supplementary Figure 10G)**. To compare the two cohorts on the cell subset level, we ranked the abundance of epithelial and immune subsets from each cohort **(Figure 7F, 7H)**. Healthy controls in the IND-Kolkata cohort had high frequencies of *GRM7+ESRRG+* ciliated and *MORN5+PIFO+*ciliated cells as well as multiple basal cell subsets **(Figure 7F)**, similar to the controls in the USA-Boston cohort **(Figure 1I, Figure 3F-G)**. There was less concordance in frequency of immune subsets, though quiescent populations such as resting CD8+ T cells, ribo-hi γδ T cells, and *MARCH5+PAX5+* B cells ranked highly in frequency in both cohorts **(Figure 7H)**. Together, these analyses indicate that our atlas can be used as a reference to find shared epithelial and immune phenotypes among healthy children from different environments.

Next, we evaluated similarities in infection responses between children in the BD-rural cohort and the virally-infected cases in the USA-Boston cohort. The BD-rural cohort was divided into children who had no respiratory symptoms (Healthy, n=8), had respiratory symptoms but were negative for RSV (Sick, n=4), and symptomatic RSV+ cases (n=7) **(Figure 7D, Supplementary Table 2)**. We generated a dataset of 17,981 high-quality cells **(Supplementary Figure 10H-K)** with single-cell detection of RSV RNA in 4/7 RSV samples **(Supplementary Figure 10L)**. We used the same label transfer approach to annotate epithelial and immune cells, and similarly found recovery of expected cell types, significant correlations between subset rankings, and strong prediction scores **(Figure 7I-L, Supplementary Figure 10M-N)**. Notably, *KRT17+S100A9+* ciliated cells were among the most frequent epithelial subsets in both the BD-rural RSV cases and USA-Boston RSV cases **(Figure 7J)**. Concordance in immune subset frequencies was improved when transferring between infected cohorts **(Figure 7L)**, indicating that the broader immune heterogeneity present during infection may facilitate more robust label transfer.

Investigation of cell type frequencies between groups in the BD-rural cohort revealed trends of enriched macrophages and goblet cells in RSV cases **(Figure 7M)**, as well as specific increases

710 in *BPIFA1+MUC5B+* goblet cells and a trending increase *KRT17+S100A9+* ciliated cells in RSV
711 cases **(Figure 7N)**, consistent with our previous results. Finally, we evaluated which of the
712 predicted epithelial subsets contained RSV RNA. Strikingly, we observe RSV RNA in the same
713 epithelial subsets as in the USA-Boston cohort, including *KRT17+S100A9+* ciliated cells and
714 *CXCL11*+IFN-stim ciliated cells, as well as in a few other subsets **(Figure 7O, Supplementary**
715 **Figure 10O)**. Together, these results confirm the ability of our atlas to serve as a reference for
716 both healthy and infected datasets, and highlight the finding of RSV-specific *KRT17+* ciliated cells
717 as a conserved phenotype across cohorts.
718

719 ## Discussion

720 Here, we present a comprehensive atlas of nasal epithelial and immune cell responses to three
721 common viral infections in young children. We sought to define how SARS-CoV-2, rhinovirus, and
722 RSV shape the cellular landscape of the pediatric nasal mucosa and how virus-specific cell states
723 relate to clinical outcomes. By systematically profiling epithelial and immune compartments, we
724 identified distinct infected epithelial cell states in SARS-CoV-2 and RSV and delineated broader
725 convergent and divergent changes in nasal cellular composition across infections. SARS-CoV-2
726 infection induced a robust IFN response, whereas rhinovirus infection instead drove a secretory
727 epithelial program marked by IL-13-responsive gene expression, and RSV elicited heightened
728 cytotoxic and inflammatory immune activity. We found that severe RSV cases exhibited increased
729 type I and type II IFN-responsive programs, while rhinovirus- associated wheezing showed
730 enrichment of goblet and deuterosomal cell subsets. By interpreting asthma GWAS data through
731 the lens of our atlas, we identified a hillock-like squamous epithelial subset with particular
732 relevance to childhood-onset asthma. Finally, we demonstrated that this atlas serves as a high-
733 resolution reference for annotating independent pediatric nasal datasets and uncovering shared
734 cellular states. Together, these findings suggest that common respiratory viruses engage distinct
735 epithelial response programs in the pediatric nasal mucosa, which in turn correspond to divergent
736 immune activation patterns and clinical outcomes.
737

738 SARS-CoV-2 infection is typically mild in young children, yet its prevalence and continued
739 circulation underscores the importance of defining pediatric mucosal responses to this virus. By
740 mapping viral transcripts to individual cells, we observed that the majority of SARS-CoV-2 RNA
741 in the pediatric nasal mucosa was found in a ciliated epithelial subset marked by strong NF-κB–
742 responsive gene expression. This transcriptional state closely mirrors a SARS-CoV-2 RNA-high
743 epithelial subset previously identified in the adult nasal mucosa as consistently enriched across

744 SARS-CoV-2 viral variants[19], as well as a “hyper-infected” ciliated epithelial population described
745 in an adult human SARS-CoV-2 challenge study[18]. Shared features of these populations include
746 induction of canonical NF-κB feedback regulators (*NFKBIA, NFKBIZ, TNFAIP3*), AP-1-associated
747 immediate-early genes indicative of epithelial activation (*ATF3, EGR1, FOS, JUN*), and markers
748 of cellular stress and metabolic adaptation (*DDIT3, DNAJB1, GDF15*). NF-κB-driven
749 transcriptional programs are functionally required for efficient SARS-CoV-2 replication in epithelial
750 cell lines and nasal organoid models[99,100]. Importantly, comparison with rhinovirus and RSV
751 infections demonstrated that this epithelial activation state is specific to SARS-CoV-2 in the
752 pediatric nasal mucosa. Together, these findings suggest that a conserved NF-κB-responsive
753 epithelial program may represent a relevant target for SARS-CoV-2 antivirals.
754
755 Although the SARS-CoV-2-infected epithelial cell subset is not primarily defined by a canonical
756 interferon-stimulated gene program, it expresses *IFNL*, which induces a robust interferon
757 response in surrounding epithelial cells in mucosal tissues[101,102]. Consistent with this, we observed
758 increased IFNα response signature scores in epithelial cells and enrichment of multiple interferon-
759 stimulated cell subsets - including goblet, club, ciliated epithelial cells and T cells - in SARS-CoV-
760 2 cases compared to rhinovirus and RSV. Prior scRNA-seq studies comparing adult and pediatric
761 SARS-CoV-2 infection have similarly reported heightened interferon responses and higher
762 baseline expression of interferon-stimulated genes in children[25,26,28], and it has been proposed
763 that this contributes to the generally mild clinical course of SARS-CoV-2 infection in children[103].
764 Our data are consistent with this model and further suggest that this pronounced interferon
765 landscape distinguishes SARS-CoV-2 from rhinovirus and RSV infection in early childhood.
766
767 In contrast to SARS-CoV-2, rhinovirus infection was characterized by a diminished interferon
768 response, loss of ciliated epithelial cells, and a shift toward secretory and goblet cell differentiation
769 accompanied by induction of IL-13-responsive genes. Consistent with these findings, prior studies
770 in both clinical samples and experimental infection models have reported reduced expression of
771 ciliary genes and increased goblet cell–associated programs during rhinovirus infection[104–107].
772 Recent work further demonstrated that epithelial interferon signaling is critical for restricting
773 rhinovirus replication, and that disruption of this response permits NF-κB-, NLRP1-, and IL-1β-
774 driven inflammatory signaling and mucus production[33]. Given that ciliated epithelial cells are a
775 primary target of rhinovirus infection[108,109], the pronounced reduction in ciliated cells we observed
776 may reflect preferential infection and subsequent sloughing of ciliated cells. In line with this model,
777 we identified multiple secretory cell subsets enriched in rhinovirus cases, including a *MUC5AC+*

778 deuterosomal population. Deuterosomal cells are thought to represent transient intermediates in
779 multiciliated cell differentiation[53]; however, their enrichment in rhinovirus infection may reflect
780 altered epithelial differentiation trajectories with potential diversification toward secretory or
781 ciliated fates. Together, these findings suggest that, *in vivo*, rhinovirus infection preferentially
782 engages epithelial differentiation pathways associated with mucus production downstream of
783 Type 2 cytokine signaling, rather than an interferon-mediated antiviral response.
784
785 The epithelial response to RSV infection *in vivo* remains poorly understood. RSV primarily infects
786 ciliated cells *in vitro*, although productive infection of basal cells has also been described[110,111]. In
787 our *in vivo* dataset, the epithelial subset harboring the majority of RSV RNA exhibited a striking
788 transcriptional profile characterized by programs related to squamous-like activation (*KRT17,*
789 *KRT23, SPRR1B*), wound repair and extracellular matrix remodeling (*LAMB3*, *PLAUR*, *TIMP2*),
790 and inflammation (*S100A9, C3, CXCL8*) alongside remarkably low expression of ciliated markers
791 and ISGs. RNA velocity analysis suggested that this subset most likely arose from ciliated cells,
792 consistent with viral injury-induced reprogramming following infection. Early work demonstrated
793 that RSV infection induces NF-kB dependent expression of cytokeratin 17 (KRT17) in respiratory
794 epithelial cells, with localization to sites of syncytia formation[112], providing initial experimental
795 evidence that RSV can drive cytoskeletal remodeling through inflammatory signaling. Our findings
796 extended these observations to an *in vivo* context, and showed that this *KRT17+S100A9+*
797 activated epithelial state is highly specific to RSV-infected cells. KRT17 has also been linked to
798 aberrant epithelial activation and dysregulated wound repair in other respiratory diseases,
799 including pro-fibrotic epithelial programs in idiopathic pulmonary fibrosis[113,114]. Together, these
800 observations suggest that this RSV-associated epithelial program reflects a cell intrinsic response
801 to viral infection, and we hypothesize that this state may promote viral persistence and dampen
802 effective intrinsic antiviral responses.
803
804 In addition to this unusual infected epithelial cell state, RSV cases exhibited a heightened
805 inflammatory and cytotoxic immune profile. Cell-cell communication analyses revealed
806 macrophage-centered signaling, and network-level analyses identified a coordinated circuit
807 linking RSV-infected epithelial cells with NK cell, T cell, and macrophage subsets. Several RSV-
808 enriched subsets, including *KRT17+S100A9+* ciliated cells, activated cycling CD4+ T cells, and
809 cytotoxic CD8+ T cells, correlated with clinical outcomes such as illness duration, need for
810 respiratory support, and hospital admission across all patients in the cohort. Consistent with this
811 observation, profiling in the blood and airway has demonstrated innate immune remodeling in

812 severe RSV, including monocyte activation, altered NK cell abundance and function, and IL-36
813 associated neutrophilic inflammation[24,40–42,87,88,115]. In our cohort, RSV participants admitted to the
814 ICU displayed a cellular composition driven by IFN-stimulated epithelial and immune subsets,
815 along with elevated type I and type II IFN response signatures and higher nasal viral loads. In
816 contrast, prior studies have generally reported diminished airway IFN responses in severe or
817 intubated RSV cases[39,42,86,116,117], despite associations between nasal viral load and IFN
818 signaling[118,119]. These differences likely reflect variation in disease stage and sampling site, as
819 our cohort was sampled in the emergency department and may represent an earlier phase of
820 illness. Together, these findings support a model in which RSV elicits a comparatively limited
821 epithelial-intrinsic antiviral program while driving robust innate and cytotoxic immune activation,
822 in contrast to SARS-CoV-2, where a stronger epithelial interferon response may help restrict viral
823 spread within the mucosa. In this heightened inflammatory context, *KRT17*+ epithelial remodeling
824 and elevated IFN signaling may reinforce inflammatory circuits, contributing to immune-mediated
825 pathology rather than effective viral containment.
826
827 Our dataset also enabled identification of cellular subsets associated with virus-induced wheeze.
828 Some of these phenotype-specific states, including mast cells and *CXCL13*+ CD8+ T cells, may
829 in part reflect transcriptional responses to therapeutic steroids, and disentangling treatment
830 effects from disease biology will require further study. Clinically, rhinovirus is more strongly linked
831 to subsequent asthma exacerbations, and rhinovirus-induced wheeze more often reflects an
832 asthma-like, type 2–skewed inflammatory phenotype that responds to standard asthma-directed
833 therapies, whereas RSV-associated wheeze typically arises from bronchiolitis-driven
834 inflammation and does not demonstrate the same treatment responsiveness[6,11,91]. In line with this,
835 wheeze-associated signatures in our cohort differed markedly between rhinovirus and RSV
836 cases. Rhinovirus-associated wheeze was characterized by further polarization toward a
837 secretory, IL-13-responsive epithelial program with reduced IFN-responsive signatures. These
838 findings suggest that robust interferon responses may mitigate wheezing severity, while type 2-
839 biased immunity may predispose to asthma-like symptoms. We identified a *CST1*+ goblet cell
840 subset enriched specifically in rhinovirus-induced wheeze, with features consistent with STAT6-
841 mediated transcriptional regulation downstream of IL-4 and IL-13 signaling[120]. CST1 expression
842 has been associated with eosinophilic asthma and correlates with Th2 cytokine activity in
843 adults[57,58,60]. The presence of this epithelial state in young children following infection raises the
844 possibility that type 2-skewed disease trajectories may begin early in life. It remains to be further
845 characterized whether rhinovirus drives such disease pathologies and subsequent risk for asthma

or if rhinovirus-associated wheezing in young children is a marker of susceptibility for a Th2-skewed response.

By interrogating which specific cell subsets in the pediatric nasal mucosa are enriched for genetic risk of asthma onset, we demonstrated the broader utility of this atlas beyond acute infection responses. Genetic risk for asthma intersects with susceptibility to rhinovirus-induced wheeze[7,9], and rhinovirus infection has been shown to upregulate childhood-onset asthma- associated genes, particularly in non-ciliated epithelial cells[121]. In our dataset, childhood-onset asthma GWAS enrichment scores were highest in squamous epithelial cells, and most pronounced in a *KRT13+HOPX+* squamous subset. This population resembles “hillock” cells, an epithelial subset implicated in airway repair and remodeling[50,55,61,122]. Emerging evidence further links KRT13+ hillock-like epithelial states and squamous differentiation programs to childhood-onset asthma, pediatric wheeze, and allergic airway inflammation[36,98,123]. Together, these findings suggest that infection-induced activation or expansion of this squamous repair program may intersect with inherited asthma risk and shape long-term airway remodeling trajectories.

Finally, we demonstrated that this dataset serves as a robust resource for annotating independent nasal scRNA-seq samples. Using our atlas as a reference, we successfully transferred annotations to both healthy and infected pediatric nasal samples generated in India and Bangladesh using a more portable technology, with strong concordance in both cell type identity and relative subset abundances across sites. Notably, *KRT17+S100A9+* ciliated cells identified in the BD-rural cohort similarly harbored RSV RNA, reinforcing the reproducibility of key epithelial states across geographic contexts. To our knowledge, this represents the first published single-cell RNA-sequencing dataset generated locally in Bangladesh, and one of the few locally produced pediatric airway scRNA-seq datasets from low- and middle-income countries. This work illustrates a scalable model in which deep reference profiling at one site can support streamlined, lower-depth profiling across additional settings, enabling greater flexibility in technology selection and study design. Advancing our understanding of nasal biology and respiratory infection requires inclusion of diverse geographies, ancestries, and populations, particularly given that the burden of childhood respiratory disease is greatest in low- and middle-income countries[3,5,43,44,46]. By prioritizing partnership and capacity building, we aimed to establish a sustainable framework for future studies. Moving forward, deeper integration of datasets across locations will allow us to assess both shared and region-specific cellular features of pediatric respiratory disease.

In summary, we establish a comprehensive single-cell atlas of the pediatric nasal mucosa during common respiratory viral infections, defining virus-specific epithelial and immune cell states and linking them to clinical outcomes. Beyond characterizing acute infection responses, this atlas enables hypothesis generation regarding mechanisms of disease severity and asthma risk, and provides a reference framework for annotation and comparison of future pediatric nasal cohorts. By integrating viral detection, cellular phenotypes, and genetic risk, our findings offer a foundation for understanding how early-life respiratory infections may influence longer-term airway biology. Future longitudinal studies and integration with functional experiments will be necessary to define the temporal dynamics and causal relationships underlying these responses across diverse populations.

Limitations of the study

This study has several limitations. First, each participant was sampled at a single time point, limiting our ability to assess the temporal dynamics of viral replication and host response. Children were sampled at presentation to the emergency department, representing different stages of illness, and illness duration was self-reported by the participant's parents. Second, our focus on early childhood restricts our ability to compare infants and toddlers directly or to fully evaluate age-dependent differences within the cohort. Third, cell numbers were limited in some samples, and certain cell types such as neutrophils were not detected, likely due to freeze-thaw processing. Fourth, single-cell clustering and annotation are inherently influenced by upstream analytical choices and remain partially subjective; we have therefore provided detailed rationale and methodological transparency to support reproducibility. Finally, detection of viral RNA within a cell does not definitively establish productive infection and should be interpreted as evidence of viral transcript presence rather than replication competence. In contrast to RSV and SARS-CoV-2, we detected rhinovirus transcripts only rarely in our single-cell data. This limited detection may reflect rapid sloughing or death of infected ciliated epithelial cells, lower levels or shorter duration of viral RNA persistence, or the substantial genomic diversity of rhinoviruses[91], which may reduce alignment efficiency to reference genomes. As a result, we were not able to define rhinovirus-infected epithelial cell states with the same resolution as for RSV and SARS-CoV-2. Dedicated approaches, including targeted viral enrichment or earlier sampling relative to symptom onset, will likely be required to more comprehensively characterize rhinovirus-infected epithelial populations *in vivo*.

Acknowledgements

We thank the study participants and their families for enabling this research, the clinical support staff at each site for their efforts in sample collection, and members of the Ordovas-Montanes and Horwitz labs for thoughtful discussion and feedback. From the BCH ED, we thank Tom Stivers for protocol management and oversight, and clinical research coordinators John Schultz, Michelle Du, Rebecca Wolf, Joseph Griffiths, Deirdre Flanagan, Abigail Bryant, Carolyn Drescher, Joseph Kanaan, and Aniyah Travis for patient recruitment and sample collection. For support of sample collection and data generation for the BD-rural cohort, we thank Naito Kanon, Himadree Sarkar, Mohammad Shameem Hassan, Shakiul Kabir and the Mirzapur Field Team. We thank the nurses and technicians of the Institute of Child Health, Kolkata, for help with recruitment of the children and collection of swabs from them. We thank all members of the BCH Cell Discovery Network for computational support and analytical insight. We also thank Maria Gutierrez-Arcelus and Vitor Rezende Da Costa Aguiar for helpful discussions related to scDRS analysis. We thank Erica Rutherford for support in data curation. We sincerely thank Arlene Sharpe and all members of the Sharpe laboratory for insightful discussion and feedback as well. We thank Juliane Weller for making the mg2sc pipeline for kraken2 single-cell analysis publically available. This has been made possible in part by CZI grant DAF2021-237665 and grant DOI https://doi.org/10.37921/881723wbyqpp from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation (funder DOI 10.13039/100014989), awarded to AKShalek, SCG, TIS, AKSesay, CSQB, LMY, PM, SS, BHH, and JOM, as well as a Grant DAF2023-330879 to JOM and Grant DAF 2023-330382 to JMLW. JOM is a New York Stem Cell Foundation – Robertson Investigator. JOM was supported by the AbbVie-Harvard Medical School Alliance, the Leona M. and Harry B. Helmsley Charitable Trust, The Pew Charitable Trusts Biomedical Scholars, The Broad Next Generation Award, The Chan Zuckerberg Initiative Pediatric Networks, The Mathers Foundation, The Kenneth Rainin Foundation, The Massachusetts Consortium for Pathogen Readiness, NIH R01 HL162642, NIH R01 DE031928, NIH P30 DK034854, the NIH CIDER Network, Project NextGen, and The Cell Discovery Network, a collaborative funded by The Manton Foundation and The Warren Alpert Foundation at Boston Children's Hospital. JMLW was supported by NIH Training Grant 5TL1TR002543. LJJ was supported by NIH 5T32HD098061 Neonatal Research Training Program to Boston Children's Hospital, the Marshall Klaus Perinatal Research Award, a Thrasher Foundation Early Career Award, and a Gerber Foundation Novice Award.

Declaration of Interests

JOM reports compensation for consulting services with Tessel Biosciences, Radera Biotherapeutics, and Passkey Therapeutics. AKShalek reports compensation for consulting and/or SAB membership from Honeycomb Biotechnologies, Cellarity, Conquest Technologies, Ochre Bio, Relation Therapeutics, IntrECate Biotherapeutics, Parabilis Medicines, Passkey Therapeutics, Danaher, and Dahlia Biosciences unrelated to this work. JMLW reports compensation for employment at Genentech unrelated to this work.

Author contributions

Contributions are listed according to CRediT criteria: https://credit.niso.org/

Conceptualization: JMLW, LJJ, YT, AC, AKShalek, PM, YH, SS, BHH, JOM;

Data curation: JMLW, LJJ, YT, AC, FT, CML, PCD, LT, AM, SG, ASK, YH,

Formal analysis: JMLW, LJJ, MEB, HRM;

Funding acquisition: AKShalek, SCG, TIS, AKSesay, CSQB, WG, LMY, PM, SS, BHH, JOM;

Investigation: JMLW, LJJ, YT, AC, EL, AF, FT, DPK, PCD, ARM, LT, AMT, AM;

Methodology: JOM, JMLW, LJJ, BHH, YT, SS, YH, PM, AC;

Project administration: JMLW, LJJ, YT, AC, AKShalek, PM, YH, SS, BHH, JOM;

Resources: YT, AC, AKShalek, PM, YH, SS, BHH, JOM;

Software: JMLW, LJJ, MEB, HRM, CML, PCD, SG

Supervision: JMLW, LJJ, PM, YH, SS, BHH, JOM;

Validation: JMLW, LJJ;

Visualization: JMLW, LJJ, MEB, KK;

Writing-original draft: JMLW, LJJ, JOM;

Writing-review & editing: JMLW, LJJ, YT, AC, AKShalek, PM, YH, SS, BHH, JOM

969 **Methods**

970 **Study Cohort Details**

971 Participants were recruited from 4 sites as described below. Participant metadata for BCH and
972 MGH cohorts is summarized in **Supplementary Table 1**, and information for all individual
973 participants is available in **Supplementary Table 2**.

974 **USA-Boston cohort: BCH**
975 A convenience sample of children under 5 years of age presenting to the Emergency Department
976 (ED) at Boston Children's Hospital (BCH) with either 1) signs of viral respiratory illness including
977 cough, nasal congestion, wheezing or difficulty breathing, or 2) minor traumatic injuries including
978 minor head trauma and simple lacerations without symptoms of respiratory viral infection within
979 the last 30 days were identified by screening of the electronic health record. Following informed
980 consent from a parent or legal guardian, a research nasopharyngeal (NP) swab was obtained
981 from each child. When available, discarded viral transport media from clinically indicated nasal
982 sampling were also collected. In addition, parents were administered a questionnaire regarding
983 past medical history, current medication usage, and current symptomatology. Following patient
984 disposition, review of the electronic health record was performed to obtain patient clinical
985 parameters related to the patients visit including vital signs, prescribed medications, ED
986 disposition, and hospital length of stay, as well as demographic data. All clinical and demographic
987 data was entered into a REDCap database. This study was approved by the Boston Children's
988 Institutional Review Board under protocol number P00028229.
989
990 Children who presented with symptoms consistent with respiratory viral infection were first
991 assigned to one of the viral groups (SARS-CoV-2, RSV, or rhinovirus) based on detection of one
992 of these viruses by clinically indicated testing and/or by qPCR performed on viral transport media.
993 Children who presented with minor traumatic injuries were assigned to the control group. For
994 126/132 samples, when processing the research NP swabs for single-cell RNA-sequencing, we
995 saved supernatants at multiple steps, combined and stored them for each sample, and repeated
996 the qPCR to confirm infection status (see assay details below). Any control samples with a positive
997 value for SARS-CoV-2, RSV-A, RSV-B, rhinovirus/enterovirus, influenza A, or influenza B were
998 re-classified as asymptomatic cases (n=5). Any samples in the viral groups with positive values
999 for more than one of the tested viruses were re-classified as coinfection cases (n=7).
1000 Asymptomatic and coinfection cases were included in cell clustering and annotation, but excluded
1001 from virus-specific comparisons.
1002
1003 Participants who were admitted to the hospital from the ED were categorized as either admitted
1004 to a standard pediatric inpatient unit (Floor) or admitted to the Intermediate Care Program,
1005 Medical Intensive Care Unit, or Medical-Surgical Intensive Care Unit (categorized together as
1006 ICU) at BCH. Groups of patients who were wheezing were defined by the following
1007 characteristics: (1) presence or absence wheeze documented on physical exam in ED physician
1008 note; (2) receipt of wheezing-directed therapy in the ED, including nebulized albuterol, IV
1009 magnesium sulfate, and/or dexamethasone. Participants were categorized in the "No wheeze"
1010 group if they did not have documented wheeze and did not receive wheezing-directed therapies

in the ED. Participants were categorized in the “Wheeze + treatment” group if they had documented wheeze and received wheezing-directed therapies in the ED.

Illness severity was quantified using a modified severity score based on the Bronchiolitis Score of Sant Joan de Deu (BROSJOD)[48]. We included the age-specific parameters for heart rate and respiratory rate, oxygen saturation, and accessory muscle use/retractions, all according to the BROSJOD criteria. We used the first recorded vital signs, obtained prior to medical therapy. We did not include the wheeze/rates or air entry components of the BROSJOD score as these were rarely reported in the BCH electronic medical record (EMR).

USA-Boston cohort: MGH

Healthy children aged from 1 month to 5 years receiving care at Massachusetts General Hospital (MGH) Chelsea Healthcare Center were enrolled during their visit under an MGH Institutional Review Board-approved study. Written informed consent was acquired from parents/legal guardians. Demographic and clinical variables were extracted from participants’ EMR.

IND-Kolkata cohort

Parental consent was obtained after providing adequate information about the research to obtain a nasal swab from children between 6 months and 5 years. None of the children was suffering from any infection. Most children had come for regular physical check-up or for immunization; swabs were collected before immunization. All sample collection procedures were minimally invasive and well tolerated, and no adverse events were reported. Ethical approvals for sample collection were provided by the Institutional Ethics Committees of the Institute of Child Health, Kolkata, and the National Institute of Biomedical Genomics, Kalyani, India.

BD-rural cohort

Eligible participants were recruited in a rural community located in the Mirzapur subdistrict in Tangail district, Bangladesh (BD-rural), where Child Health Research Foundation (CHRF) has maintained a demographic-based surveillance system since 2007[124]. Written informed consent was obtained from the parents or legal guardians of all participants prior to sample collection. Enrolled participants were children less than 2 years of age who were identified by village health workers as sick or healthy. A child was considered sick if they met the WHO inclusion criteria for RSV community surveillance (defined as having at least one of the following: shortness of breath, cough, sore throat, coryza)[125]. Age, sex and location-matched healthy participants were recruited as the control group; these were children who did not meet the RSV inclusion criteria, did not have any other illness at the time of recruitment and did not report use of any medications. Samples were collected during two time periods: October to November 2023 and September to October 2024. The Ethical Review Committee of the Bangladesh Shishu Hospital and Institute approved the study (Admin/BSHI/2023/923/1).

Method Details

Resource availability

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Jose Ordovas-Montanes. (jose.ordovas-montanes@childrens.harvard.edu).

Sample collection

Nasopharyngeal swabs were collected by a trained healthcare provider using FLOQSwabs (Copan). The swab was placed immediately into a cryogenic vial containing 2mL of heat inactivated fetal bovine serum (FBS, Corning) with 10% dimethyl sulfoxide (DMSO, Sigma). Cryovials were then placed into a room temperature Mr. Frosty Freezing Container (Thermo Fisher Scientific), and stored in a -80C freezer overnight. Samples were moved to individual boxes and kept at -80C.

Dissociation of single cells from nasal swabs

Swabs in freezing medium (90% FBS, 10% dimethylsulfoxide) were stored at -80C until immediately before dissociation. A detailed sample protocol can be found here: https://www.protocols.io/view/czi-pediatric-nasopharyngeal-swab-processing-for-1-6qpvr49jogmk/v1. Briefly, nasal swabs in freezing medium were thawed, and each swab was rinsed with RPMI (Thermo Fisher Scientific) before incubation in 1 ml RPMI, 10 mM DTT (Sigma) for 15 min at 37 °C with agitation. Next, the nasal swab was incubated in 1 ml Accutase (Sigma) for 30 min at 37 °C with agitation. The 1 ml of RPMI with 10 mM DTT from the nasal swab incubation was centrifuged at 400g for 5 min at 4 °C to pellet cells, the supernatant was saved, and the cell pellet was resuspended in 1 ml Accutase and incubated for 30 min at 37 °C with agitation. The original cryovial containing the freezing medium and the original swab washes were combined and centrifuged at 400g for 5 min at 4 °C. The supernatant was saved and the cell pellet was then resuspended in RPMI with 10 mM DTT and incubated for 15 min at 37 °C with agitation and centrifuged as above. The cell pellet was then resuspended in 1 ml Accutase and incubated for 30 min at 37°C with agitation. All cells were combined following Accutase digestion and filtered using a 40-µm nylon strainer. The filter and the swab were washed with RPMI with 10% FBS and 4 mM EDTA, and all washes were combined. Dissociated, filtered cells were centrifuged at 400g for 10 min at 4 °C and resuspended in 200 µl RPMI with 10% FBS for counting. 10,000 live cells from each sample were combined into pools of 3-4 samples. The pooled cells were centrifuged at 400g for 5 min at 4C, then resuspended in 500uL PBS+1% BSA and centrifuged again as above. The pellet was resuspended in 43.3 uL of PBS+1% BSA to proceed to scRNA-seq. 58/132 samples were additionally subjected to live cell enrichment on the LeviCell 1.0 instrument. After counting, cells were pooled at equal ratios. The pooled cells were centrifuged at 300g for 5 minutes and resuspended in 230 uL. of 150mM Levitation Agent (Levitas Bio) diluted in RPMI+10% FBS. The standard cell enrichment protocol was run. After collecting the enriched cell population, cell counting was repeated. 10,000 live cells times the number of samples in the pool were transferred to a new tube, washed and resuspended in PBS+1% BSA as above to proceed to scRNA-seq.

1093 **USA-Boston cohort: Single-cell RNA-seq**
1094 For participants in the USA-Boston cohort, pooled samples were processed using the Chromium
1095 NextGEM Single Cell 3' Kit v3.1 with dual indices (10X Genomics) per the manufacturer's
1096 instructions. Library quality was evaluated using the Agilent TapeStation 4200 (Agilent) and
1097 concentration was measured using the Qubit 1xdsDNA High-Sensitivity Assay Kit (Qubit).
1098 Sequencing was performed on the NovaSeq 6000 (Illumina) at the Broad Institute Sequencing
1099 Core with a targeted read depth of 25,000 reads per cell: read 1, 28; read 2; 90, index 1, 10; index
1100 2, 10.
1101
1102 **IND & BD cohorts: Single-cell RNA-seq**
1103 For participants in the IND-Kolkata and BD-rural cohorts, cells were isolated using the same
1104 dissociation protocol as above. scRNA-seq libraries were generated for 10,000 cells for each
1105 individual sample using the HIVE™ scRNAseq v1 Processing Kit (Honeycomb Biotechnologies)
1106 or the HIVE CLX scRNAseq Kit (Honeycomb Biotechnologies) and prepared according to the
1107 manufacturer's protocol. Libraries were sequenced on NextSeq 2000 or NovaSeq 6000 platforms
1108 (Illumina) with a targeted read depth of 25,000 reads per cell: read 1, 25; read 2, 50; index 1, 8;
1109 index 2, 8.
1110
1111 **USA-Boston cohort: Bulk RNA-seq**
1112 Aliquots of 5,000 cells from each sample were transferred to new tubes. These pellets were spun
1113 down at 400g for 5 min at 4C and resuspended in 50uL of RLT buffer (Qiagen) + 1% 2-
1114 mercaptoethanol (BME, Sigma). The lysed cells were snap frozen and stored at -80C. Libraries
1115 were prepared using the previously described SmartSEQ2 protocol[126]. Briefly, RNA was purified
1116 from 50 µL of cell lysate using 2.0X SPRIselect beads (Beckman-Coulter). Purified RNA was
1117 reverse transcribed into cDNA and whole transcriptome amplification was carried out. Libraries
1118 were fragmented and dual indexed using the Nextera XT Library Prep Kits (Illumina). Sequencing
1119 was performed on the NovaSeq 6000 or NextSeq 2000 at the Broad Institute Sequencing Core
1120 with a targeted read depth of 10 million reads per sample, read 1, 59; read 2; 59, index 1, 10;
1121 index 2, 10.
1122
1123 **USA-Boston cohort: Viral qRT-PCR**
1124 1mL RLT+BME was added to the combined supernatants collected during research NP swab
1125 processing to lyse any remaining cells. This lysate was then aliquoted into 1mL/cryovial, snap
1126 frozen, and stored at -80C until use. qRT-PCR was done on both scRNA-seq swab supernatants
1127 to confirm viral infection and on clinical swab samples to preliminarily assign groups. For both
1128 types of samples, RNA was extracted using QIAamp MinElute Virus Spin Kit (Qiagen) following
1129 the manufacturer's instructions. 1-step standard RT-qPCR was performed on QuantStudio™ 7
1130 Flex Real-Time PCR System (Thermo Fisher Scientific) using TaqMan assays to identify viral
1131 infections. Supernatants from 126/132 participants were tested for B-βCoV E gene, SARS-CoV-
1132 2 S gene, RSV-A, RSV-B, influenza A, influenza B, rhinovirus/enterovirus, and RPP30 as the
1133 internal control. All samples were negative for influenza A and B. Samples with a cycle threshold
1134 (CT) value <40 were considered positive. All CT values reported in figures refer to scRNA-seq
1135 swab PCR.
1136

BD cohort: RSV Detection

RNA was extracted from 200 µL of the specimen using Quick RNA Viral kit (Zymo Research) following the manufacturer's protocol. qPCR for RSV detection and subgrouping was performed using 6.25 µL RNA with Luna Universal Probe One-Step RT-qPCR Kit (New England Biolabs) and previously published RSV-specific primers and probes[127]. PCR cycling conditions were 50°C for 30 minutes, 95°C for 5 minutes, and 40 cycles of 95°C for 15 seconds and 60°C for 30 seconds. A true sigmoidal amplification curve of ct-value <35 confirmed RSV positivity.

Quantification & Statistical Analysis

Sample demultiplexing

Bulk RNA-seq libraries were demultiplexed using Illumina bcl2fastq (v2.20). SNPs specific to each participant were called using Genome Analysis Toolkit (GATK4) workflows provided by the Broad Institute using default settings[128]. Paired FASTQs were converted into unmapped BAM files and aligned to the GRCh38 genome. Duplicate marking and base recalibration was performed and haplotype caller was used to generate a variant call file (VCF) for each participant. A merged VCF file was then created by aggregating the VCFs from all participants in each pool.

SNP-based genetic demultiplexing of pooled scRNA-seq samples was performed as previously described[129–131]. Briefly, scRNA-seq pools were initially demultiplexed in a genotype-agonistic method through the Freemuxlet variation of the Popscle package. Freemuxlet performs unsupervised clustering of single cells based on common SNPs identified within transcripts and assigns each cell as a nameless donor or doublet[130]. In addition to assigning cells, a VCF is generated for identifying SNPs of each donor. These VCFs are then compared to the aggregated VCF from each pool and matched based on genotype similarity (vcf-match-sample-ids)[131]. Freemuxlet and participant labels were then added to each cell. Cells identified as inter-participant doublets were excluded from further analysis.

scRNA-seq pre-processing

To detect both host and viral transcripts, we built a custom reference genome combining human GRCh38-2020-A (Ensembl 93), SARS-CoV-2[66], RSV-A (GenBank: KY654518.1)[132], RSV-B (GenBank: MZ516105.1)[133], Rhinovirus C NAT001 (GenBank: EF077279.1) [69], Rhinovirus B104 (GenBank: FJ445137.1)[134], Rhinovirus B48 (GenBank: DQ473488)[69], and Enterovirus E19 (GenBank: AY302544.1)[135]. For rhinovirus/enterovirus genomes, these genomes were selected for the custom reference genome based on Kraken2[70] analysis of unmapped reads which categorized the greatest number of *Picornaviridae* k-mers to these genomes. Sequences for viral genomes were added to the GRCh38 FASTA and GTF files and processed using Cellranger's mkref() function. scRNA-seq libraries were then aligned and quantified using Cellranger (v7.2.0) via Cumulus tools[136]. To quantify ambient RNA, Cellbender (v0.2.0)[137] remove-background() function was applied to each pool with default parameters and epochs = 150, hardware_boot_disk_size_GB = 300, hardware_disk_size_GB = 300, learning_rate = 0.0001, low_count_threshold = 100 and total_droplets_included = 100,000. Using Seurat (v5.1.0)[138], an object for each pool was created with count matrices for human genes, cellbender corrected human genes, viral genes, and cellbender-corrected viral genes stored in separate assays.

1181 Assignment of RNA+ cells for each virus was done as previously described [15,19]. Briefly, we
1182 determined the amount of ambient RNA in each cell by comparing the raw human gene matrices
1183 with the cellbender-corrected human gene matrices. Then, we quantified pre-cellbender viral RNA
1184 in empty and low-quality droplets (defined as <500 UMI, <300 genes, >35% mitochondrial).
1185 Finally, we tested whether the abundance of post-cellbender viral RNA in each cell was
1186 significantly above the abundance in empty droplets given the ambient RNA profile of each pool
1187 using an exact binomial test (binom.test in R). We performed Benjamini–Hochberg correction to
1188 determine the FDR (p.adjust, method = 'BH' in R), and cells with FDR < 0.01 were assigned as
1189 SARS-CoV-2 RNA+. Following viral RNA+ assignment, non-cellbender human reads were used
1190 for all further analysis. Cells were filtered to remove those with <500 UMI, <300 genes, and >35%
1191 mitochondrial reads. This resulted in a dataset of 426,577 cells and 40,317 genes across 132
1192 study participants.
1193
1194 **Single-cell viral metagenomics**
1195 Due to challenges identifying Rhinovirus-infected cells in the setting of significant genetic diversity
1196 within Enteroviruses and the frequency of coinfections in pediatric patients, we sought to perform
1197 an unbiased taxonomic analysis of the data at a single-cell level. Following CellRanger alignment
1198 to the human genome, kraken2[70] was run on unmapped bam files using a previously published
1199 kraken2 database that includes all pathogenic human viruses accessible in GenBank as of
1200 2020[139] using the mg2sc pipeline[140]. The output cell-by-taxon k-mer counts matrix was collapsed
1201 to the species level using collapse_taxonomy.py from mg2sc[140]. In R, the .h5ad file was converted
1202 to an RDS using Zellkonverter (v.1.10.1)[141], merged with the GRCh38-aligned Cellranger h5 file,
1203 and a Seurat object was created. To create an input file for Cellbender correction, which was used
1204 to define true viral-positive cells as above, the Seurat object was saved as an h5 file using
1205 DropletUtils (v.1.24.0) Create_10X_h5() function[142]. For downstream analyses, a Seurat object
1206 was created with human and kraken reads in separate assays as above. To identify cells
1207 containing k-mers classified to specific viruses, species names were identified as follows: for
1208 SARS-CoV-2, "Severe acute respiratory syndrome-related coronavirus" and "Severe acute
1209 respiratory syndrome coronavirus 2"; for rhinovirus, "Rhinovirus A", "Rhinovirus B", and
1210 "Rhinovirus C"; for RSV, "Respiratory syncytial virus" and "Human respiratory syncytial virus". It
1211 was noted that a large number of viral k-mers in RSV+ participants were classified at the genus
1212 level instead of species level, so an additional classification was created for cells positive for
1213 "Human orthopneumovirus" and "Respiratory syncytial virus". Positive versus negative cells were
1214 classified by a binomial test comparing viral k-mers in cells to empty droplets identified by
1215 Cellbender, as described in more detail in the preceding section.
1216
1217 **Evaluation of batch effect**
1218 To determine if any batch effects in our dataset necessitated applying integration methods, we
1219 first performed sample-level gene expression principal component analysis (PCA). Pseudobulk
1220 gene expression profiles from each sample were created using the Seurat AggregateExpression()
1221 function and DESeq2 (v1.46.0)[143]. Variance stabilizing transformation was applied before the PCA
1222 analysis. To evaluate the impact of metadata variables on the overall variance in the dataset, we
1223 pre-processed the Seurat object and performed PCA on the 3,000 most variable human genes.
1224 For each principal component (PC), we calculated the proportion of variance explained as the

variance of that PC divided by the total variance across the top 20 PCs. We then fit a linear model between each PC and each metadata variable, recording the coefficient of determination (R^2). Finally, we estimated the variance contribution of each PC–metadata pair as the product of the PC's explained variance and the corresponding R^2. The maximum contribution was between PC1 and participant, explaining 5.1% of the dataset variance **(Supplementary Figure 1F)**. Because we did not want to obscure important biological variability between participants, and because we did not observe any significant batch effects based on technical and biological variables of interest, we chose not to integrate our data going forward.

Cell clustering and annotation

Non-cellbender-corrected matrices containing only human genes were used in cell annotation. Data were normalized using sctransform (v0.4.1)[144] with default parameters and method = 'glmGamPoi'. Cells were clustered using Louvain clustering. Unique marker genes for each cluster were identified using Seurat's FindAllMarkers() function with cutoffs min.pct = 0.25 and logfc.threshold = 0.25. To assign predicted doublet likelihood scores to each cell, we applied scrublet (v0.2.3)[145] (with python v3.7.12, scanpy (v1.9.3)[146,147], anndata (v0.8.0)[147]) to each pool.

We applied an iterative clustering process to remove low-quality cells and confidently determine cell annotations **(Supplementary Figures 2a, 3a)**. We first assigned preliminary clusters as either epithelial (n=294,285) or immune (132,292). In each following round, cells of interest were separated and sctransform normalization, PCA, and clustering were repeated. Clusters were removed if 1) the only marker genes were mitochondrial and the percentage of mitochondrial genes was significantly higher than in other clusters, 2) more than 75% of the cells in that cluster came from one pool, indicating a technical artifact, or 3) the cluster was suspected to be doublets based on the expression of marker genes for multiple cell types and higher predicted doublet scores than other clusters. Once objects were divided into cell-type specific objects, clusters were re-assigned if they expressed the marker genes for a different cell type to a greater extent than the marker genes for their original cell type annotation. This approach resulted in the removal of 53,331 epithelial cells (18%) and 40,016 immune cells (30%).

Major epithelial cell types were annotated using marker genes from previously published studies of the human airway[15,19,49–52]. Initial clustering of epithelial cells revealed small clusters of neuronal cells (*RBFOX1*, *CNTNAP2*, *PTPRD*)[148–150] and sustentacular cells (*ERMN*, *GPX6*, *SEC14L3*)[65]. We next clustered non-ciliated epithelial cells including basal cells (*KRT5, KRT15*), club cells (*SCGB1A1, SCGB3A1*), goblet cells (*MUC5AC, VMO1*), and squamous cells (*SPRR3, SPRR2E*). We identified 5 basal subsets defined by growth genes (*EGFR, SEMA5A*), classical basal markers (*KRT5, TP63*), proliferation genes (*MKI67, TOP2A*), stress-response genes (*EGR1, ATF3*), and secretory cell-like genes (*S100P, KRT7*). There were two club cell subsets alternately expressing interferon (IFN) response genes (*IFIT1*, *IFIT2*) and *BPIFA1*. Exploration of goblet cells revealed 7 subsets: a low-complexity subset we labeled "Developing", a mitochondrial-high subset, and 5 subsets marked by secretory / mucin genes (*BPIFA1, MUC5B)*, alarmin genes (*S100A8, S100A9*), IFN response genes, statherin (*STATH*), and the cysteine protease inhibitor *CST1*. Squamous cells similarly contained one low-complexity subset, as well as one high for *KRT13* and *HOPX* which resembled "hillock" cells[55,61] and a pro-inflammatory subset (*IL6, IL1A, CXCL8*).

We identified 7 subsets within ciliated cells (*CAPS, DNAH12)*. 2 ciliated subsets were defined by canonical ciliated markers or cellular growth genes (*MORN5, FOXJ1, GRM7, ESRRG*). The remaining ciliated subsets were defined by class-II antigen presentation genes (*HLA-DPB1*, *HLA-DQA1*), oxidative stress-response genes (*GSTA1, GSTA2*), IFN-response genes (*RSAD2, CXCL11*), keratins and tight junction genes (*KRT17, LAMB3)*, and NF-kB response genes (*NFKBIA, TNFAIP3*). Deuterosomal cells[53,54] (*CCNO, CDC20B*) contained two ciliated-like subsets (*CROCC, DLEC1, PIFO, MORN2*) and one subset expressing the goblet gene *MUC5AC*. We also identified a population of ionocytes[55,56] (*CFTR, SCNN1B*) which contained an IFN-responsive subset (*GBP1, CXCL10)*.

We next repeated this iterative clustering with the immune cells. After removing a small cluster of melanocytes (*PMEL*, *MLANA*)[151], we separated the cells into macrophages (*TYROBP, FCER1G*), plasmacytoid dendritic cells (pDCs) (*IRF7, IRF8, GZMB*), B cells (*MS4A1, BANK1*), mast cells (*CPA3, MS4A2*), and T/NK cells (*CD3D, TRAC, GNLY*). We identified 6 macrophage subsets marked respectively by complement genes (*C1QA, C1QB*), lipid metabolism genes (*APOE, APOC1*), alarmins and chemokines (*S100A8, CCL2*), pro-inflammatory genes (*CXCL8, IL1B*), the signaling molecule *PLCG2*, and a *SPP1*-high subset. We also identified subsets resembling monocyte-derived dendritic cells (*GPR183, CD86*) and conventional type 1 dendritic cells (*CLEC9A*). Plasmacytoid dendritic cells were divided into subsets expressing the transcriptional regulator *ZFAT*, the antimicrobial gene *LYZ* and the receptor tyrosine kinase *AXL*, and high amounts of type I IFNs (*IFNA, IFNB*). B cell subsets included one highly expressing canonical markers *MARCH1* and *PAX5*, a quiescent subset highest in ribosomal genes, an activated subset expressing *NR4A2* and *NR4A3*, an immunoglobulin lambda (*IGLC2, IGLC3*), a cycling subset (*MKI67, TOP2A*), and a distinct subset of plasma cells (*MZB1, JCHAIN, IGHA1*). We next separated natural killer (NK) cells from T cells based on their phenotype and lack of *CD3D* expression. We identified 3 NK cell subsets: a quiescent ribosomal-high subset, an *NCAM1*+ high subset, and a *CCL4*+ subset. We also noted a small cluster of innate lymphoid cells resembling lymphoid tissue inducer (LTi) cells based on the expression of *IL4I1, KIT, IL23R,* and *LTB*. We then defined gamma-delta (γδ) T cells as those that expressed both *TRGC2* and *TRDC* more highly than *TRAC* and *TRBC2*. There were 4 subsets of γδ T cells expressing ribosomal genes, the transcriptional repressor *ZBTB20*, immune signaling molecules *XCL1* and *XCL2*, and IFN-responsive genes respectively. When we sub-clustered alpha-beta (ab) T cells, we found that many phenotypic clusters (such as resident-memory T (Trm) cell-like, cycling, quiescent, and cytotoxic) contained both CD4+ and CD8+ T cells. To elucidate both the T cell lineage and the phenotype, we first separated the ab T cells into 7 phenotypic clusters. Next, we sub-clustered each of these to divide each one into CD4+ and CD8+ cells. Finally, we re-combined all identified CD4+ and CD8+ cells, respectively, and performed the final sub-clustering and annotation. Within CD4+ T cells, we found a quiescent subset, a naive-like *IL7R*+ *KLRB1*+ subset, a subset expressing exhaustion-related genes (*HAVCR2, CD38*), an activated *TNFRSF4*+ subset, Th17 cells (*IL17A, IL23R*), Tph cells highly expressing *CXCL13*, and Tregs (*FOXP3, IL2RA*). With CD8+ T cells, we similarly find quiescent and ribosomal-high subsets, a Trm-like subset (*ITGAE, ITGA1*), an IFN-responsive subset, chemokine-high *CCL4*+*IFNG*+ cells, conventional cytotoxic T cells (*GZMB, PRF1*) as well as a subset highly expressing granulysin (*GNLY*). In addition, we found two small subsets of CD8+ T cells expressing *IL17A* and *CXCL13*, respectively. Finally, we

identified a cluster of MAIT cells (*ZBTB16*, *SLC4A10*) as well as cycling CD4+ and CD8+ T cells. Marker genes used to annotate clusters in the final round for each cell type are provided in **Supplementary Table 3**.

Pseudobulk differential expression

To compare transcriptional signatures between different groups of participants or cells within a subset of interest, we created pseudobulk aggregates for each participant using Seurat's AggregateExpression(). A DESeqDataset (DESeq2 v.46.0) object was then constructed from the aggregated count matrix, and genes with <10 counts across all samples were filtered out. Differential expression testing was done using the DESeq() function with default parameters and contrasts specified per analysis. Genes with an adjusted p-value (FDR) <0.05 were considered statistically significant. RSV+ CXCL11+ IFN-stim ciliated cells vs. SARS-CoV-2+ CXCL11+ IFN-stim ciliated cells **(Figure 2F)**: participants with at least 20 CXCL11+IFN-stim cells were included (SARS-CoV-2: n=4; RSV: n=6); RSV+ KRT17+S100A9+ ciliated cells vs. RSV- KRT+S100A9+ ciliated cells **(Figure 3B)**: participants with at least 20 KRT17+S100A9+ ciliated cells were included (n=27); T cells in ICU admitted RSV participants vs. T cells in floor-admitted RSV participants **(Figure 6E)**: all admitted RSV participants were included (admitted to floor: n = 14, admitted to ICU: n =7). Complete lists of DE genes for each of these comparisons are provided in **Supplementary Table 4**.

RNA velocity

We compared potential infected cell fates using RNA velocity with scVelo (v0.3.3)[152]. Briefly, loom files containing unspliced and spliced RNA matrices were created by running velocyto (v0.17.17)[153] using Cellranger-aligned bam files, the GRCh38-2020-A reference genome (Ensembl 93), and GRCh38 with repetitive elements masked. The masked GRCh38 genome was obtained from the UCSC genome browser (https://genome.ucsc.edu/cgi-bin/hgTables) by downloading the gtf file of GRCh38 using track “Repeat Masker” for “All tracks”[154]. Aggregated loom files were merged with the processed, clustered, and annotated epithelial Anndata object and subset to high-quality cells from Control, SARS-CoV-2, and RSV groups. Following scVelo filtering, normalization, nearest neighbor calculation, and moments computation, splicing dynamics were calculated using the dynamical model. The velocity streams were visualized in UMAP space.

Transcription factor scores

Transcription factor scores were calculated by univariate linear modeling (ULM) of SCT-transformed counts with the collecTRI database[155] of transcription factor regulons using the decoupler package(v2.12)[156] in Python. Enriched transcription factor activities for subsets of interest compared to all other subsets were selected by t-test-overstim-var within decoupler.

Compositional analysis

For compositional analyses, we first created an abundance matrix of either cell types or cell subsets for each participant. Center-log-ratio (CLR) normalization was carried out using the setaCLR() function from the SETA (v1.0.0)[157] package. To assess changes in cell type abundance

relative to controls, we calculated the difference between the average CLR-normalized abundance for each viral group compared to the average CLR-normalized abundance for the control group. PCA on the CLR-normalized cell subset abundance matrix was computed using the setaLatent() function, method = "PCA", dims=10. To evaluate relationships between cell subsets, Spearman correlation between the CLR-normalized abundance for each combination of cell subsets was calculated.

Cytokine response signature scores

Response signatures were derived from gene expression of cytokine-treated airway basal cells[72,73] (**Figure 3E-F, Figure 4D-E, Supplementary Figure 5C–D**; IFNα, IL-4, and IL13 response signatures in epithelial cells), from the Hallmark database using msigdbr (v25.1.1)[158] **(Figure 6F, Supplementary Figure 8F**; IFNα and IFNγ response signatures in all cells), or gene response signatures from glucocorticoid treatment of airway epithelial cells[92] **(Supplementary Figure 8G)**. Gene signature scores per participant were calculated using univariate linear modeling (ULM) of pseudobulk-averaged SCT-transformed counts using the R implementation of decoupler[156] (v2.9.7). To score gene signatures on a per-cell level **(Supplementary Figure 5C)**, univariate linear modeling was applied to a downsampled object containing 10% of cells from each epithelial cell subset. Genes included in each signature are provided in **Supplementary Table 5**.

Cell-cell communication

To assess differences in cell-cell communication between groups, we applied multinichenetr (v2.1.0)[89], an extension of nichenetr (v2.2.1)[159] designed for multi-sample, multi-condition differential analysis. Pseudobulk expression was calculated for each sample and cell type combination. The contrast design compared the control group to all other samples and each viral group to controls only. We did not define any batch or covariate variables. When evaluating the differential expression of ligands of interest **(Figure 5B)**, we used more permissive cutoffs for the minimum number of cells per sample-cell type combination(min_cells = 5) and the minimum proportion of samples in which a gene was expressed (min_sample_prop = 0.25) and recommended cutoff for fraction of cells with nonzero expression per sample-cell type combination (fraction_cutoff = 0.05). Out of a comprehensive list of cytokines and chemokines **(Supplementary Table 6)**, we evaluated which were identified as differentially expressed (FDR<0.05) in each viral group compared to controls.

To prioritize receptor-ligand interactions, we applied standard cutoffs to filter cell types and genes included in the pseudobulk (min_cells =10, sample_prop = 0.5, fraction_cutoff = 0.05). The same contrast design described above was used to identify differentially expressed genes for each cell type. To assess ligand activities, we considered the top 250 target genes and standard parameters. All prioritization criteria were weighed equally. The top 50 ranked interactions were plotted in **Figure 5C** and **Supplementary Figure 6A**, though we evaluated the number of interactions, the prioritization scores, and the distribution of sender cell types for various cutoffs **(Supplementary Figure 6B-D)**. To visualize prioritization metrics, we calculated the average scaled pseudobulk expression product of the ligand and receptor for each group and interaction. We narrowed the interactions displayed in **Figure 5D** by 1) selecting the sender cell types that

accounted for at least 5% of interactions and 2) showing only the highest ranked interaction for repeated receptor-ligand-receiver combinations. The complete list of the top 500 ranked interactions can be found in **Supplementary Table 6**.

Predictive modeling of cell subset importance

We trained a penalized regression classifier with 5-fold cross-validation to distinguish patients by Wheeze_Group based on cell subset abundance using the glmnet model from the Caret package (v7.0.1)[160]. Variables (cell subsets) that contributed the most to predicting Wheeze_Group were extracted for plotting.

scDRS analysis

We linked our single cell transcriptomes with polygenic disease risk for asthma and height using scDRS (v1.0.4)[97]. GWAS summary statistics for childhood-onset asthma (GCST007800) and adult-onset asthma (GCST007799)[161] were downloaded from the NHGRI-EBI GWAS Catalog[162] on 2025-09-16. SNP-level summary statistics were converted to gene-level Z-scores using MAGMA version (v1.1.0)[163]. We removed HLA regions (chr6: 25-34 Mb) and used the 1000 Genomes European cohort as our reference panel. Gene-level scores were converted to scDRS .gs format with scDRS munge-gs (--weight zscore, --n-max 1000, and --n-min 100) and merged with pre-processed gene sets provided by the original scDRS study.

We then ran scDRS using a Nextflow workflow set up for this analysis using default parameters. During preprocessing, we corrected for Batch, Collection_Site, and LeviCell using scdrs.pp.preprocess (n_mean_bin=20, n_var_bin=20). Cell-level scores were obtained with scdrs.method.score_cell (n_ctrl=1000 and ctrl_match_key="mean_var”). For downstream analysis, we used scdrs.method.downstream_group_analysis, scdrs.method.downstream_corr_analysis, and scdrs.method.downstream_gene_analysis.

Label transfer

Data was aligned using BeeNetPlus (v1.0.X) software implemented on Terra (Honeycomb Biotechnologies). Seurat objects for each independent cohort were created from raw count matrices (tcm.tsv files) for each sample. For the BD-rural cohort, samples were aligned to a custom genome including both human and RSV genes designed for BeeNetPlus software. Count matrices for viral genes were stored in a separate assay within the Seurat object. Cells with <100 transcripts, <100 unique genes, or >35% mitochondrial reads were filtered out. Objects were first clustered independently at low resolution to assign cells as epithelial or immune. Each epithelial / immune object was then clustered once to ensure the correct identity and to remove any low-quality cells. This resulted in high-quality epithelial (IND-Kolkata cohort: n=16,523 cells; BD-rural cohort: n=12,715 cells) and immune (IND-Kolkata cohort: n=40,544 cells; BD-rural cohort: n=5,266 cells) objects.

Labels were then transferred within each lineage. For the IND-Kolkata cohort, only samples in the control group in the USA-Boston cohort were used as a reference. For the BD-rural cohort, labels were transferred from all samples in the USA-Boston cohort. UMAP embedding for the reference object was first re-calculated based on the cells and samples included. Label transfer anchors

were identified using the FindTransferAnchors() function with the union of variable genes as the features of interest and normalization.method = "SCT", then used as the anchorset in the TransferData() function. The anchorset was also used to project the UMAP embedding from the reference object to the query object via MapQuery(). Label transfer was evaluated by examining the distribution of prediction scores and consistency with manual annotations. Similarity in abundance of cell subsets between reference and query datasets was evaluated by 1) calculating the frequency of each subset out of all epithelial or immune cells per sample, 2) taking the average frequency within all samples from one site, 3) ranking the subsets by their average frequency such that the most abundant subset had the highest ranking and 4) correlating the rankings between each site.

References

1. Teoh, Z., Conrey, S., McNeal, M., Burrell, A., Burke, R.M., Mattison, C., McMorrow, M., Payne, D.C., Morrow, A.L., and Staat, M.A. (2023). Burden of respiratory viruses in children less than 2 years old in a community-based longitudinal US birth cohort. Clin. Infect. Dis. *77*, 901–909.

2. Jain, S., Williams, D.J., Arnold, S.R., Ampofo, K., Bramley, A.M., Reed, C., Stockmann, C., Anderson, E.J., Grijalva, C.G., Self, W.H., et al. (2015). Community-acquired pneumonia requiring hospitalization among U.S. children. N. Engl. J. Med. *372*, 835–845.

3. Wang, X., Li, Y., Shi, T., Bont, L.J., Chu, H.Y., Zar, H.J., Wahi-Singh, B., Ma, Y., Cong, B., Sharland, E., et al. (2024). Global disease burden of and risk factors for acute lower respiratory infections caused by respiratory syncytial virus in preterm infants and young children in 2019: a systematic review and meta-analysis of aggregated and individual participant data. Lancet *403*, 1241–1253.

4. Suh, M., Movva, N., Jiang, X., Bylsma, L.C., Reichert, H., Fryzek, J.P., and Nelson, C.B. (2022). Respiratory syncytial virus is the leading cause of United States infant hospitalizations, 2009-2019: A study of the national (nationwide) inpatient sample. J. Infect. Dis. *226*, S154–S163.

5. Li, Y., Wang, X., Blau, D.M., Caballero, M.T., Feikin, D.R., Gill, C.J., Madhi, S.A., Omer, S.B., Simões, E.A.F., Campbell, H., et al. (2022). Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in children younger than 5 years in 2019: a systematic analysis. Lancet *399*, 2047–2064.

6. Hasegawa, K., Mansbach, J.M., Bochkov, Y.A., Gern, J.E., Piedra, P.A., Bauer, C.S., Teach, S.J., Wu, S., Sullivan, A.F., and Camargo, C.A., Jr (2019). Association of Rhinovirus C bronchiolitis and immunoglobulin E sensitization during infancy with development of recurrent wheeze. JAMA Pediatr. *173*, 544–552.

7. Everman, J.L., Sajuthi, S., Saef, B., Rios, C., Stoner, A.M., Numata, M., Hu, D., Eng, C., Oh, S., Rodriguez-Santana, J., et al. (2019). Functional genomics of CDHR3 confirms its role in HRV-C infection and childhood asthma exacerbations. J. Allergy Clin. Immunol. *144*, 962–971.

8. Busse, W.W., Lemanske, R.F., Jr, and Gern, J.E. (2010). Role of viral respiratory infections in asthma and asthma exacerbations. Lancet *376*, 826–834.

9. 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.

10. De Leeuw, E., Justesen, J.F., Bosteels, C., Debeuf, N., Vanheerswynghels, M., Jonckheere, L., De Wolf, C., Wayenberg, A., Van Damme, K.F.A., Vanhee, S., et al. (2025). Maternal allergy and neonatal RSV infection synergize via FcR-mediated allergen uptake to promote the development of asthma in early life. Sci. Immunol. *10*, eadz4626.

11. Makrinioti, H., Hasegawa, K., Lakoumentas, J., Xepapadaki, P., Tsolia, M., Castro-Rodriguez, J.A., Feleszko, W., Jartti, T., Johnston, S.L., Bush, A., et al. (2022). The role of

1496 respiratory syncytial virus- and rhinovirus-induced bronchiolitis in recurrent wheeze and
1497 asthma-A systematic review and meta-analysis. Pediatr. Allergy Immunol. *33*, e13741.

1498 12. Gallo, O., Locatello, L.G., Mazzoni, A., Novelli, L., and Annunziato, F. (2021). The central
1499 role of the nasal microenvironment in the transmission, modulation, and clinical progression
1500 of SARS-CoV-2 infection. Mucosal Immunol. *14*, 305–316.

1501 13. Mettelman, R.C., Kaitlynn Allen, E., and Thomas, P.G. (2022). Mucosal immune responses
1502 to infection and vaccination in the respiratory tract. Immunity *55*, 749–780.

1503 14. Ziegler, C.G.K., Allon, S.J., Nyquist, S.K., Mbano, I.M., Miao, V.N., Tzouanas, C.N., Cao, Y.,
1504 Yousif, A.S., Bals, J., Hauser, B.M., et al. (2020). SARS-CoV-2 Receptor ACE2 Is an
1505 Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell
1506 Subsets across Tissues. Cell *181*, 1016–1035.e19.

1507 15. Ziegler, C.G.K., Miao, V.N., Owings, A.H., Navia, A.W., Tang, Y., Bromley, J.D., Lotfy, P.,
1508 Sloan, M., Laird, H., Williams, H.B., et al. (2021). Impaired local intrinsic immunity to SARS-
1509 CoV-2 infection in severe COVID-19. Cell *184*, 4713–4733.e22.

1510 16. Chua, R.L., Lukassen, S., Trump, S., Hennig, B.P., Wendisch, D., Pott, F., Debnath, O.,
1511 Thürmann, L., Kurth, F., Völker, M.T., et al. (2020). COVID-19 severity correlates with airway
1512 epithelium-immune cell interactions identified by single-cell analysis. Nat. Biotechnol. *38*,
1513 970–979.

1514 17. Mitsi, E., Diniz, M.O., Reiné, J., Collins, A.M., Robinson, R.E., Hyder-Wright, A., Farrar, M.,
1515 Liatsikos, K., Hamilton, J., Onyema, O., et al. (2023). Respiratory mucosal immune memory
1516 to SARS-CoV-2 after infection and vaccination. Nat. Commun. *14*, 6815.

1517 18. Lindeboom, R.G.H., Worlock, K.B., Dratva, L.M., Yoshida, M., Scobie, D., Wagstaffe, H.R.,
1518 Richardson, L., Wilbrey-Clark, A., Barnes, J.L., Kretschmer, L., et al. (2024). Human SARS-
1519 CoV-2 challenge uncovers local and systemic response dynamics. Nature *631*, 189–198.

1520 19. Walsh, J.M.L., Miao, V.N., Owings, A.H., Tang, Y., Bromley, J.D., Kazer, S.W., Kimler, K.,
1521 Asare, C., Ziegler, C.G.K., Ibrahim, S., et al. (2025). Variants and vaccines impact nasal
1522 immunity over three waves of SARS-CoV-2. Nat. Immunol. *26*, 294–307.

1523 20. Castagnoli, R., Votto, M., Licari, A., Brambilla, I., Bruno, R., Perlini, S., Rovida, F., Baldanti,
1524 F., and Marseglia, G.L. (2020). Severe acute respiratory syndrome Coronavirus 2 (SARS-
1525 CoV-2) infection in children and adolescents: A systematic review: A systematic review.
1526 JAMA Pediatr. *174*, 882–889.

1527 21. Koch, C.M., Prigge, A.D., Anekalla, K.R., Shukla, A., Do Umehara, H.C., Setar, L., Chavez,
1528 J., Abdala-Valencia, H., Politanska, Y., Markov, N.S., et al. (2022). Age-related Differences
1529 in the Nasal Mucosal Immune Response to SARS-CoV-2. Am. J. Respir. Cell Mol. Biol. *66*,
1530 206–222.

1531 22. Ho, R.M., Bowen, A.C., Blyth, C.C., Imrie, A., Kollmann, T.R., Stick, S.M., and Kicic, A.
1532 (2023). Defining the pediatric response to SARS-CoV-2 variants. Front. Immunol. *14*,
1533 1200456.

1534 23. Rao, S., Gross, R.S., Mohandas, S., Stein, C.R., Case, A., Dreyer, B., Pajor, N.M., Bunnell,
1535 H.T., Warburton, D., Berg, E., et al. (2024). Postacute sequelae of SARS-CoV-2 in children.

1536 Pediatrics *153*, e2023062570.

1537 24. Openshaw, P.J.M., Chiu, C., Culley, F.J., and Johansson, C. (2017). Protective and Harmful
1538 Immunity to RSV Infection. Annu. Rev. Immunol. *35*, 501–532.

1539 25. Loske, J., Röhmel, J., Lukassen, S., Stricker, S., Magalhães, V.G., Liebig, J., Chua, R.L.,
1540 Thürmann, L., Messingschlager, M., Seegebarth, A., et al. (2021). Pre-activated antiviral
1541 innate immunity in the upper airways controls early SARS-CoV-2 infection in children. Nat.
1542 Biotechnol. https://doi.org/10.1038/s41587-021-01037-9.

1543 26. Yoshida, M., Worlock, K.B., Huang, N., Lindeboom, R.G.H., Butler, C.R., Kumasaka, N.,
1544 Dominguez Conde, C., Mamanova, L., Bolt, L., Richardson, L., et al. (2022). Local and
1545 systemic responses to SARS-CoV-2 infection in children and adults. Nature *602*, 321–327.

1546 27. Magalhães, V.G., Lukassen, S., Drechsler, M., Loske, J., Burkart, S.S., Wüst, S., Jacobsen,
1547 E.-M., Röhmel, J., Mall, M.A., Debatin, K.-M., et al. (2023). Immune-epithelial cell cross-talk
1548 enhances antiviral responsiveness to SARS-CoV-2 in children. EMBO Rep. *24*, e57912.

1549 28. Woodall, M.N.J., Cujba, A.-M., Worlock, K.B., Case, K.-M., Masonou, T., Yoshida, M.,
1550 Polanski, K., Huang, N., Lindeboom, R.G.H., Mamanova, L., et al. (2024). Age-specific nasal
1551 epithelial responses to SARS-CoV-2 infection. Nat Microbiol. https://doi.org/10.1038/s41564-
1552 024-01658-1.

1553 29. Rajan, A., Nagaraj, D., Bomidi, C., Aloisio, G.M., Murray, A.M., Schultz, E.M., Kambal, A.,
1554 Estes, M.K., Nicholson, E., Avadhanula, V., et al. (2025). Single cell sequencing analysis of
1555 respiratory syncytial virus-infected pediatric and adult human nose organoids reveals age
1556 differences, proliferative diversity and identifies novel cellular tropism. J. Infect. *91*, 106617.

1557 30. Zhao, C., Bai, Y., Wang, W., Amonkar, G.M., Mou, H., Olejnik, J., Hume, A.J., Mühlberger,
1558 E., Lukacs, N.W., Fearns, R., et al. (2024). Activation of STAT3-mediated ciliated cell survival
1559 protects against severe infection by respiratory syncytial virus. J. Clin. Invest. *134*.
1560 https://doi.org/10.1172/JCI183978.

1561 31. Griffiths, C.D., Bilawchuk, L.M., McDonough, J.E., Jamieson, K.C., Elawar, F., Cen, Y., Duan,
1562 W., Lin, C., Song, H., Casanova, J.-L., et al. (2020). IGF1R is an entry receptor for respiratory
1563 syncytial virus. Nature *583*, 615–619.

1564 32. Tayyari, F., Marchant, D., Moraes, T.J., Duan, W., Mastrangelo, P., and Hegele, R.G. (2011).
1565 Identification of nucleolin as a cellular receptor for human respiratory syncytial virus. Nat.
1566 Med. *17*, 1132–1135.

1567 33. Wang, B., Amat, J.A.R., Mihaylova, V.T., Kong, Y., Wang, G., and Foxman, E.F. (2026).
1568 Rhinovirus triggers distinct host responses through differential engagement of epithelial
1569 innate immune signaling. Cell Press Blue *0*, 100001.

1570 34. Berg, K., Haid, S., Vafadarnejad, E., Carpentier, A., Geffers, R., Wiegmann, B., Saliba, E.,
1571 Erhard, F., and Pietschmann, T. (2025). Respiratory syncytial virus load-dependent ciliated
1572 cell dedifferentiation and downregulation of immune response genes. bioRxiv,
1573 2025.08.01.666372. https://doi.org/10.1101/2025.08.01.666372.

1574 35. Gay, A.C.A., Banchero, M., Carpaij, O., Kole, T.M., Apperloo, L., van Gosliga, D., Fajar, P.A.,
1575 Koppelman, G.H., Bont, L., Hendriks, R.W., et al. (2024). Airway epithelial cell response to

1576 RSV is mostly impaired in goblet and multiciliated cells in asthma. Thorax *79*, 811–821.

1577 36. Berdnikovs, S., Newcomb, D.C., Haruna, N.-F., McKernan, K.E., Kuehnle, S.N., Gebretsadik,
1578 T., McKennan, C., Ma, S., Cephus, J.-Y., Rosas-Salazar, C., et al. (2025). Single-cell profiling
1579 demonstrates the combined effect of wheeze phenotype and infant viral infection on airway
1580 epithelial development. Sci. Adv. *11*, eadr9995.

1581 37. Nehar-Belaid, D., Mejías, A., Xu, Z., Marches, R., Yerrabelli, R., Chen, G., Mertz, S., Ye, F.,
1582 Sánchez, P.J., Tsang, J.S., et al. (2025). SARS-CoV-2 induced immune perturbations in
1583 infants vary with disease severity and differ from adults' responses. Nat. Commun. *16*, 4562.

1584 38. Wimmers, F., Burrell, A.R., Feng, Y., Zheng, H., Arunachalam, P.S., Hu, M., Spranger, S.,
1585 Nyhoff, L.E., Joshi, D., Trisal, M., et al. (2023). Multi-omics analysis of mucosal and systemic
1586 immunity to SARS-CoV-2 after birth. Cell *186*, 4632–4651.e23.

1587 39. Flerlage, T., Crawford, J.C., Allen, E.K., Severns, D., Tan, S., Surman, S., Ridout, G., Novak,
1588 T., Randolph, A., West, A.N., et al. (2023). Single cell transcriptomics identifies distinct
1589 profiles in pediatric acute respiratory distress syndrome. Nat. Commun. *14*, 3870.

1590 40. Reilly, R.B., Ramdour, S.K., Fuhlbrigge, M.E., Tavares, L.P., Staffa, S.J., Booth, J.M.,
1591 Krishnamoorthy, N., Levy, B.D., and Duvall, M.G. (2024). An altered natural killer cell
1592 immunophenotype characterizes clinically severe pediatric RSV infection. Sci. Transl. Med.
1593 *16*, eado6606.

1594 41. Thibodeau, A., Mejias, A., Nehar-Belaid, D., Marches, R., Xu, Z., Eryilmaz, G., Josefowicz,
1595 S.Z., Jones, B., Wehenkel, M., Paust, S., et al. (2026). A systems immunology approach
1596 reveals divergent immune profiles of RSV and SARS-CoV-2 infections in infants. Sci. Transl.
1597 Med. *18*. https://doi.org/10.1126/scitranslmed.aea7097.

1598 42. Barnes, M.V.C., Stretch, S.W.H., Swieboda, D., Efstathiou, C., Culley, F.J., Hansel, T.T.,
1599 Strickson, S., Romero-Ros, X., Cohen, E.S., Nadel, S., et al. (2026). Mucosal IL-36 is a
1600 defining feature of severe paediatric bronchiolitis. Mucosal Immunol.
1601 https://doi.org/10.1016/j.mucimm.2026.01.012.

1602 43. Majumder, P.P., Mhlanga, M.M., and Shalek, A.K. (2020). The Human Cell Atlas and equity:
1603 lessons learned. Nat. Med. *26*, 1509–1511.

1604 44. Boakye Serebour, T., Cribbs, A.P., Baldwin, M.J., Masimirembwa, C., Chikwambi, Z.,
1605 Kerasidou, A., and Snelling, S.J.B. (2024). Overcoming barriers to single-cell RNA
1606 sequencing adoption in low- and middle-income countries. Eur. J. Hum. Genet. *32*, 1206–
1607 1213.

1608 45. World Health Organization (2025). Human genomics technologies in clinical studies - the
1609 research landscape: Report on the 1990-2024 period (World Health Organization).

1610 46. Mazur, N.I., Löwensteyn, Y.N., Willemsen, J.E., Gill, C.J., Forman, L., Mwananyanda, L.M.,
1611 Blau, D.M., Breiman, R.F., Madhi, S.A., Mahtab, S., et al. (2021). Global respiratory syncytial
1612 virus-related infant community deaths. Clin. Infect. Dis. *73*, S229–S237.

1613 47. Langedijk, A.C., and Bont, L.J. (2023). Respiratory syncytial virus infection and novel
1614 interventions. Nat. Rev. Microbiol. https://doi.org/10.1038/s41579-023-00919-w.

1615 48. Balaguer, M., Alejandre, C., Vila, D., Esteban, E., Carrasco, J.L., Cambra, F.J., and Jordan,
1616 I. (2017). Bronchiolitis Score of Sant Joan de Déu: BROSJOD Score, validation and
1617 usefulness. Pediatr. Pulmonol. *52*, 533–539.

1618 49. Vieira Braga, F.A., Kar, G., Berg, M., Carpaij, O.A., Polanski, K., Simon, L.M., Brouwer, S.,
1619 Gomes, T., Hesse, L., Jiang, J., et al. (2019). A cellular census of human lungs identifies
1620 novel cell states in health and in asthma. Nat. Med. *25*, 1153–1163.

1621 50. Deprez, M., Zaragosi, L.-E., Truchi, M., Becavin, C., Ruiz García, S., Arguel, M.-J., Plaisant,
1622 M., Magnone, V., Lebrigand, K., Abelanet, S., et al. (2020). A Single-Cell Atlas of the Human
1623 Healthy Airways. Am. J. Respir. Crit. Care Med. *202*, 1636–1645.

1624 51. Hewitt, R.J., and Lloyd, C.M. (2021). Regulation of immune responses by the airway epithelial
1625 cell landscape. Nat. Rev. Immunol. *21*, 347–362.

1626 52. Davis, J.D., and Wypych, T.P. (2021). Cellular and functional heterogeneity of the airway
1627 epithelium. Mucosal Immunol. *14*, 978–990.

1628 53. Ruiz García, S., Deprez, M., Lebrigand, K., Cavard, A., Paquet, A., Arguel, M.-J., Magnone,
1629 V., Truchi, M., Caballero, I., Leroy, S., et al. (2019). Novel dynamics of human mucociliary
1630 differentiation revealed by single-cell RNA sequencing of nasal epithelial cultures.
1631 Development *146*. https://doi.org/10.1242/dev.177428.

1632 54. Revinski, D.R., Zaragosi, L.-E., Boutin, C., Ruiz-Garcia, S., Deprez, M., Thomé, V., Rosnet,
1633 O., Gay, A.-S., Mercey, O., Paquet, A., et al. (2018). CDC20B is required for deuterosome-
1634 mediated centriole production in multiciliated cells. Nat. Commun. *9*, 4668.

1635 55. Montoro, D.T., Haber, A.L., Biton, M., Vinarsky, V., Lin, B., Birket, S.E., Yuan, F., Chen, S.,
1636 Leung, H.M., Villoria, J., et al. (2018). A revised airway epithelial hierarchy includes CFTR-
1637 expressing ionocytes. Nature *560*, 319–324.

1638 56. Plasschaert, L.W., Žilionis, R., Choo-Wing, R., Savova, V., Knehr, J., Roma, G., Klein, A.M.,
1639 and Jaffe, A.B. (2018). A single-cell atlas of the airway epithelium reveals the CFTR-rich
1640 pulmonary ionocyte. Nature *560*, 377–381.

1641 57. Du, L., Xu, C., Tang, K., Shi, J., Tang, L., Lisha, X., Lei, C., Liu, H., Liang, Y., and Guo, Y.
1642 (2023). Epithelial CST1 promotes airway eosinophilic inflammation in asthma via the AKT
1643 signaling pathway. Allergy Asthma Immunol. Res. *15*, 374–394.

1644 58. Wang, M., Gong, L., Luo, Y., He, S., Zhang, X., Xie, X., Li, X., and Feng, X. (2023).
1645 Transcriptomic analysis of asthma and allergic rhinitis reveals CST1 as a biomarker of unified
1646 airways. Front. Immunol. *14*, 1048195.

1647 59. Yao, L., Yuan, X., Fu, H., Guo, Q., Wu, Y., Xuan, S., Kermani, N.Z., Adcock, I.M., Zeng, X.,
1648 Liu, Y., et al. (2023). Epithelium-derived cystatin SN inhibits house dust mite protease activity
1649 in allergic asthma. Allergy *78*, 1507–1523.

1650 60. Hoyer, A., Chakraborty, S., Lilienthal, I., Konradsen, J.R., Katayama, S., and Söderhäll, C.
1651 (2024). The functional role of CST1 and CCL26 in asthma development. Immun. Inflamm.
1652 Dis. *12*, e1162.

1653 61. Lin, B., Shah, V.S., Chernoff, C., Sun, J., Shipkovenska, G.G., Vinarsky, V., Waghray, A.,

1654 Xu, J., Leduc, A.D., Hintschich, C.A., et al. (2024). Airway hillocks are injury-resistant
1655 reservoirs of unique plastic stem cells. Nature *629*, 869–877.

1656 62. Garner, L.C., Amini, A., FitzPatrick, M.E.B., Lett, M.J., Hess, G.F., Filipowicz Sinnreich, M.,
1657 Provine, N.M., and Klenerman, P. (2023). Single-cell analysis of human MAIT cell
1658 transcriptional, functional and clonal diversity. Nat. Immunol. *24*, 1565–1578.

1659 63. Bugaut, H., El Morr, Y., Mestdagh, M., Darbois, A., Paiva, R.A., Salou, M., Perrin, L.,
1660 Fürstenheim, M., du Halgouet, A., Bilonda-Mutala, L., et al. (2024). A conserved
1661 transcriptional program for MAIT cells across mammalian evolution. J. Exp. Med. *221*.
1662 https://doi.org/10.1084/jem.20231487.

1663 64. Cupedo, T., Crellin, N.K., Papazian, N., Rombouts, E.J., Weijer, K., Grogan, J.L., Fibbe,
1664 W.E., Cornelissen, J.J., and Spits, H. (2009). Human fetal lymphoid tissue-inducer cells are
1665 interleukin 17-producing precursors to RORC+ CD127+ natural killer-like cells. Nat. Immunol.
1666 *10*, 66–74.

1667 65. Finlay, J.B., Brann, D.H., Abi Hachem, R., Jang, D.W., Oliva, A.D., Ko, T., Gupta, R.,
1668 Wellford, S.A., Moseman, E.A., Jang, S.S., et al. (2022). Persistent post-COVID-19 smell
1669 loss is associated with immune cell infiltration and altered gene expression in olfactory
1670 epithelium. Sci. Transl. Med. *14*, eadd0484.

1671 66. Kim, D., Lee, J.-Y., Yang, J.-S., Kim, J.W., Kim, V.N., and Chang, H. (2020). The Architecture
1672 of SARS-CoV-2 Transcriptome. Cell *181*, 914–921.e10.

1673 67. Human respiratory syncytial virus A isolate TBp-13-044, complete genom - Nucleotide - NCBI
1674 https://www.ncbi.nlm.nih.gov/nuccore/KY654518.1.

1675 68. Human respiratory syncytial virus B isolate OX02-8128-V01, complete ge - Nucleotide - NCBI
1676 https://www.ncbi.nlm.nih.gov/nuccore/MZ516105.1.

1677 69. Kistler, A., Avila, P.C., Rouskin, S., Wang, D., Ward, T., Yagi, S., Schnurr, D., Ganem, D.,
1678 DeRisi, J.L., and Boushey, H.A. (2007). Pan-viral screening of respiratory tract infections in
1679 adults with and without asthma reveals unexpected human coronavirus and human rhinovirus
1680 diversity. J. Infect. Dis. *196*, 817–825.

1681 70. Wood, D.E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken
1682 2. Genome Biol. *20*, 257.

1683 71. Sposito, B., Broggi, A., Pandolfi, L., Crotta, S., Clementi, N., Ferrarese, R., Sisti, S.,
1684 Criscuolo, E., Spreafico, R., Long, J.M., et al. (2021). The interferon landscape along the
1685 respiratory tract impacts the severity of COVID-19. Cell *184*, 4953–4968.e16.

1686 72. Giovannini-Chami, L., Marcet, B., Moreilhon, C., Chevalier, B., Illie, M.I., Lebrigand, K.,
1687 Robbe-Sermesant, K., Bourrier, T., Michiels, J.-F., Mari, B., et al. (2012). Distinct epithelial
1688 gene expression phenotypes in childhood respiratory allergy. Eur. Respir. J. *39*, 1197–1205.

1689 73. Ordovas-Montanes, J., Dwyer, D.F., Nyquist, S.K., Buchheit, K.M., Vukovic, M., Deb, C.,
1690 Wadsworth, M.H., 2nd, Hughes, T.K., Kazer, S.W., Yoshimoto, E., et al. (2018). Allergic
1691 inflammatory memory in human respiratory epithelial progenitor cells. Nature *560*, 649–654.

1692 74. Quinn, T.P., Erb, I., Richardson, M.F., and Crowley, T.M. (2018). Understanding sequencing

1693 data as compositions: an outlook and review. Bioinformatics *34*, 2870–2878.

1694 75. Kazer, S.W., Match, C.M., Langan, E.M., Messou, M.-A., LaSalle, T.J., O'Leary, E.,
1695 Marbourg, J., Naughton, K., von Andrian, U.H., and Ordovas-Montanes, J. (2024). Primary
1696 nasal influenza infection rewires tissue-scale memory response dynamics. Immunity *57*,
1697 1955–1974.e8.

1698 76. Yanai, H., and Taniguchi, T. (2026). The IRF transcription factor family in type I interferon-
1699 mediated antiviral immunity and associated diseases. Immun. Inflamm. *2*, 1.

1700 77. Yang, C., Mai, H., Peng, J., Zhou, B., Hou, J., and Jiang, D. (2020). STAT4: an
1701 immunoregulator contributing to diverse human diseases. Int. J. Biol. Sci. *16*, 1575–1585.

1702 78. Turi, K.N., Shankar, J., Anderson, L.J., Rajan, D., Gaston, K., Gebretsadik, T., Das, S.R.,
1703 Stone, C., Larkin, E.K., Rosas-Salazar, C., et al. (2018). Infant viral respiratory infection nasal
1704 immune-response patterns and their association with subsequent childhood recurrent
1705 wheeze. Am. J. Respir. Crit. Care Med. *198*, 1064–1073.

1706 79. Jackson, D.J., Makrinioti, H., Rana, B.M.J., Shamji, B.W.H., Trujillo-Torralbo, M.-B., Footitt,
1707 J., Jerico Del-Rosario, Telcian, A.G., Nikonova, A., Zhu, J., et al. (2014). IL-33-dependent
1708 type 2 inflammation during rhinovirus-induced asthma exacerbations in vivo. Am. J. Respir.
1709 Crit. Care Med. *190*, 1373–1382.

1710 80. Fedele, G., Schiavoni, I., Nenna, R., Pierangeli, A., Frassanito, A., Leone, P., Petrarca, L.,
1711 Scagnolari, C., and Midulla, F. (2018). Analysis of the immune response in infants
1712 hospitalized with viral bronchiolitis shows different Th1/Th2 profiles associated with
1713 respiratory syncytial virus and human rhinovirus. Pediatr. Allergy Immunol. *29*, 555–557.

1714 81. Hansel, T.T., Tunstall, T., Trujillo-Torralbo, M.-B., Shamji, B., Del-Rosario, A., Dhariwal, J.,
1715 Kirk, P.D.W., Stumpf, M.P.H., Koopmann, J., Telcian, A., et al. (2017). A comprehensive
1716 evaluation of nasal and bronchial cytokines and chemokines following experimental
1717 Rhinovirus infection in allergic asthma: Increased interferons (IFN-γ and IFN-λ) and type 2
1718 inflammation (IL-5 and IL-13). EBioMedicine *19*, 128–138.

1719 82. Beale, J., Jayaraman, A., Jackson, D.J., Macintyre, J.D.R., Edwards, M.R., Walton, R.P.,
1720 Zhu, J., Man Ching, Y., Shamji, B., Edwards, M., et al. (2014). Rhinovirus-induced IL-25 in
1721 asthma exacerbation drives type 2 immunity and allergic pulmonary inflammation. Sci.
1722 Transl. Med. *6*, 256ra134.

1723 83. Dhariwal, J., Cameron, A., Wong, E., Paulsen, M., Trujillo-Torralbo, M.-B., Del Rosario, A.,
1724 Bakhsoliani, E., Kebadze, T., Almond, M., Farne, H., et al. (2021). Pulmonary innate lymphoid
1725 cell responses during Rhinovirus-induced asthma exacerbations in vivo: A clinical trial. Am.
1726 J. Respir. Crit. Care Med. *204*, 1259–1273.

1727 84. Bont, L., Heijnen, C.J., Kavelaars, A., van Aalderen, W.M., Brus, F., Draaisma, J.T., Geelen,
1728 S.M., van Vught, H.J., and Kimpen, J.L. (1999). Peripheral blood cytokine responses and
1729 disease severity in respiratory syncytial virus bronchiolitis. Eur. Respir. J. *14*, 144–149.

1730 85. Bont, L., Heijnen, C.J., Kavelaars, A., van Aalderen, W.M., Brus, F., Draaisma, J.M.,
1731 Pekelharing-Berghuis, M., van Diemen-Steenvoorde, R.A., and Kimpen, J.L. (2001). Local
1732 interferon-gamma levels during respiratory syncytial virus lower respiratory tract infection are
1733 associated with disease severity. J. Infect. Dis. *184*, 355–358.

1734 86. Thwaites, R.S., Coates, M., Ito, K., Ghazaly, M., Feather, C., Abdulla, F., Tunstall, T., Jain,
1735 P., Cass, L., Rapeport, G., et al. (2018). Reduced Nasal Viral Load and IFN Responses in
1736 Infants with Respiratory Syncytial Virus Bronchiolitis and Respiratory Failure. American
1737 Journal of Respiratory and Critical Care Medicine. https://doi.org/10.1164/rccm.201712-
1738 2567OC.

1739 87. Heinonen, S., Velazquez, V.M., Ye, F., Mertz, S., Acero-Bedoya, S., Smith, B., Bunsow, E.,
1740 Garcia-Mauriño, C., Oliva, S., Cohen, D.M., et al. (2020). Immune profiles provide insights
1741 into respiratory syncytial virus disease severity in young children. Sci. Transl. Med. *12*,
1742 eaaw0268.

1743 88. Zivanovic, N., Öner, D., Abraham, Y., McGinley, J., Drysdale, S.B., Wildenbeest, J.G.,
1744 Crabbe, M., Vanhoof, G., Thys, K., Thwaites, R.S., et al. (2023). Single-cell immune profiling
1745 reveals markers of emergency myelopoiesis that distinguish severe from mild respiratory
1746 syncytial virus disease in infants. Clin. Transl. Med. *13*, e1507.

1747 89. Browaeys, R., Gilis, J., Sang-Aram, C., De Bleser, P., Hoste, L., Tavernier, S.J., Lambrechts,
1748 D., Seurinck, R., and Saeys, Y. (2023). MultiNicheNet: a flexible framework for differential
1749 cell-cell communication analysis from multi-sample multi-condition single-cell transcriptomics
1750 data. bioRxiv. https://doi.org/10.1101/2023.06.13.544751.

1751 90. Cosgrove, P.R., Redhu, N.S., Tang, Y., Monuteaux, M.C., and Horwitz, B.H. (2021).
1752 Characterizing T cell subsets in the nasal mucosa of children with acute respiratory
1753 symptoms. Pediatr. Res. https://doi.org/10.1038/s41390-021-01364-2.

1754 91. Jartti, T., Bønnelykke, K., Elenius, V., and Feleszko, W. (2020). Role of viruses in asthma.
1755 Semin. Immunopathol. *42*, 61–74.

1756 92. Mostafa, M.M., Rider, C.F., Shah, S., Traves, S.L., Gordon, P.M.K., Miller-Larsson, A., Leigh,
1757 R., and Newton, R. (2019). Glucocorticoid-driven transcriptomes in human airway epithelial
1758 cells: commonalities, differences and functional insight from cell lines and primary cells. BMC
1759 Med. Genomics *12*, 29.

1760 93. Yue, M., Gaietto, K., Han, Y.Y., Rosser, F.J., Xu, Z., Qoyawayma, C., Acosta-Perez, E.,
1761 Canino, G., Forno, E., Chen, W., et al. (2025). Transcriptomic profiles in nasal epithelium and
1762 asthma endotypes in youth. JAMA *333*, 307–318.

1763 94. Kicic, A., de Jong, E., Ling, K.-M., Nichol, K., Anderson, D., Wark, P.A.B., Knight, D.A.,
1764 Bosco, A., Stick, S.M., WAERP, et al. (2020). Assessing the unified airway hypothesis in
1765 children via transcriptional profiling of the airway epithelium. J. Allergy Clin. Immunol. *145*,
1766 1562–1573.

1767 95. Karpathiou, G., Papoudou-Bai, A., Ferrand, E., Dumollard, J.M., and Peoc'h, M. (2021).
1768 STAT6: A review of a signaling pathway implicated in various diseases with a special
1769 emphasis in its usefulness in pathology. Pathol. Res. Pract. *223*, 153477.

1770 96. Yamashita, M., Onodera, A., and Nakayama, T. (2007). Immune mechanisms of allergic
1771 airway disease: regulation by transcription factors. Crit. Rev. Immunol. *27*, 539–546.

1772 97. Zhang, M.J., Hou, K., Dey, K.K., Sakaue, S., Jagadeesh, K.A., Weinand, K.,
1773 Taychameekiatchai, A., Rao, P., Pisco, A.O., Zou, J., et al. (2022). Polygenic enrichment
1774 distinguishes disease associations of individual cells in single-cell RNA-seq data. Nat. Genet.

54, 1572–1580.

98. Kersten, E.T.G., Pett, J.P., Malmström, K., Chun, Y., Jonker, M.R., Wilbrey-Clark, A., Worlock, K.B., van den Berge, M., Vermeulen, R.C.H., Vonk, J.M., et al. (2023). Childhood-onset asthma is characterized by airway epithelial hillock-to-squamous differentiation in early life. bioRxiv. https://doi.org/10.1101/2023.07.31.549680.

99. Nilsson-Payant, B.E., Uhl, S., Grimont, A., Doane, A.S., Cohen, P., Patel, R.S., Higgins, C.A., Acklin, J.A., Bram, Y., Chandar, V., et al. (2021). The NF-κB transcriptional footprint is essential for SARS-CoV-2 replication. J. Virol. *95*, e0125721.

100. Simoneau, C.R., Chen, P.-Y., Xing, G.K., Hayashi, J.M., Chen, I.P., Khalid, M.M., Meyers, N.L., Taha, T.Y., Leon, K.E., Suryawanshi, R.K., et al. (2024). NF-κB inhibitor alpha controls SARS-CoV-2 infection in ACE2-overexpressing human airway organoids. Sci. Rep. *14*, 15351.

101. Lazear, H.M., Schoggins, J.W., and Diamond, M.S. (2019). Shared and Distinct Functions of Type I and Type III Interferons. Immunity *50*, 907–923.

102. Walker, F.C., Sridhar, P.R., and Baldridge, M.T. (2021). Differential roles of interferons in innate responses to mucosal viral infections. Trends Immunol. *42*, 1009–1023.

103. Svensson Akusjärvi, S., and Zanoni, I. (2024). Yin and yang of interferons: lessons from the coronavirus disease 2019 (COVID-19) pandemic. Curr. Opin. Immunol. *87*, 102423.

104. Wong, J.Y.W., Rutman, A., and O'Callaghan, C. (2005). Recovery of the ciliated epithelium following acute bronchiolitis in infancy. Thorax *60*, 582–587.

105. Hewson, C.A., Haas, J.J., Bartlett, N.W., Message, S.D., Laza-Stanca, V., Kebadze, T., Caramori, G., Zhu, J., Edbrooke, M.R., Stanciu, L.A., et al. (2010). Rhinovirus induces MUC5AC in a human infection model and in vitro via NF-κB and EGFR pathways. Eur. Respir. J. *36*, 1425–1435.

106. Ong, H.H., Andiappan, A.K., Duan, K., Lum, J., Liu, J., Tan, K.S., Howland, S., Lee, B., Ong, Y.K., Thong, M., et al. (2022). Transcriptomics of rhinovirus persistence reveals sustained expression of RIG-I and interferon-stimulated genes in nasal epithelial cells in vitro. Allergy *77*, 2778–2793.

107. Li, Y., Singh, S., Briggs, H.L., Kreger, J.E., Sliwicki, A.L., Eberhardt, E.L., Kuo, S., Czapla, J.A., Bentley, J.K., Flori, H.R., et al. (2025). Rhinovirus C15 infection induces airway epithelial cell remodeling and robust inflammatory responses: Potential implications for airway obstruction in children. Mucosal Immunol. *0*. https://doi.org/10.1016/j.mucimm.2025.09.002.

108. 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.

109. 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. *16*, 386–398.

110. Zhang, L., Peeples, M.E., Boucher, R.C., Collins, P.L., and Pickles, R.J. (2002). Respiratory

1814 Syncytial Virus Infection of Human Airway Epithelial Cells Is Polarized, Specific to Ciliated
1815 Cells, and without Obvious Cytopathology. Journal of Virology.
1816 https://doi.org/10.1128/jvi.76.11.5654-5666.2002.

1817 111. Persson, B.D., Jaffe, A.B., Fearns, R., and Danahay, H. (2014). Respiratory Syncytial Virus
1818 Can Infect Basal Cells and Alter Human Airway Epithelial Differentiation. PLoS ONE *9*,
1819 e102368.

1820 112. Domachowske, J.B., Bonville, C.A., and Rosenberg, H.F. (2000). Cytokeratin 17 is expressed
1821 in cells infected with respiratory syncytial virus via NF-kappaB activation and is associated
1822 with the formation of cytopathic syncytia. J. Infect. Dis. *182*, 1022–1028.

1823 113. Habermann, A.C., Gutierrez, A.J., Bui, L.T., Yahn, S.L., Winters, N.I., Calvi, C.L., Peter, L.,
1824 Chung, M.-I., Taylor, C.J., Jetter, C., et al. (2020). Single-cell RNA sequencing reveals
1825 profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. Sci.
1826 Adv. *6*, eaba1972.

1827 114. Jaeger, B., Schupp, J.C., Plappert, L., Terwolbeck, O., Artysh, N., Kayser, G., Engelhard, P.,
1828 Adams, T.S., Zweigerdt, R., Kempf, H., et al. (2022). Airway basal cells show a
1829 dedifferentiated KRT17highPhenotype and promote fibrosis in idiopathic pulmonary fibrosis.
1830 Nat. Commun. *13*, 5637.

1831 115. Chappin, K., Besteman, S.B., Hennus, M.P., Wildenbeest, J.G., Mokry, M., Bont, L.J., van
1832 der Vlist, M., Calis, J.J.A., and PROMISE Investigators (2024). Airway and blood monocyte
1833 transcriptomic profiling reveals an antiviral phenotype in infants with severe respiratory
1834 syncytial virus infection. J. Infect. Dis. *229*, S100–S111.

1835 116. Chu, C.-Y., Qiu, X., McCall, M.N., Wang, L., Corbett, A., Holden-Wiltse, J., Slaunwhite, C.,
1836 Grier, A., Gill, S.R., Pryhuber, G.S., et al. (2021). Airway Gene Expression Correlates of
1837 Respiratory Syncytial Virus Disease Severity and Microbiome Composition in Infants. J.
1838 Infect. Dis. *223*, 1639–1649.

1839 117. Taveras, J., Garcia-Maurino, C., Moore-Clingenpeel, M., Xu, Z., Mertz, S., Ye, F., Chen, P.,
1840 Cohen, S.H., Cohen, D., Peeples, M.E., et al. (2022). Type III interferons, viral loads, age,
1841 and disease severity in young children with respiratory syncytial virus infection. J. Infect. Dis.
1842 *227*, 61–70.

1843 118. Piedra, F.-A., Mei, M., Avadhanula, V., Mehta, R., Aideyan, L., Garofalo, R.P., and Piedra,
1844 P.A. (2017). The interdependencies of viral load, the innate immune response, and clinical
1845 outcome in children presenting to the emergency department with respiratory syncytial virus-
1846 associated bronchiolitis. PLoS One *12*, e0172953.

1847 119. Koch, C.M., Prigge, A.D., Setar, L., Anekalla, K.R., Do-Umehara, H.C., Abdala-Valencia, H.,
1848 Politanska, Y., Shukla, A., Chavez, J., Hahn, G.R., et al. (2022). Cilia-related gene signature
1849 in the nasal mucosa correlates with disease severity and outcomes in critical respiratory
1850 syncytial virus bronchiolitis. Front. Immunol. *13*, 924792.

1851 120. Goenka, S., and Kaplan, M.H. (2011). Transcriptional regulation by STAT6. Immunol. Res.
1852 *50*, 87–96.

1853 121. Djeddi, S., Fernandez-Salinas, D., Huang, G.X., Aguiar, V.R.C., Mohanty, C., Kendziorski,
1854 C., Gazal, S., Boyce, J.A., Ober, C., Gern, J.E., et al. (2024). Rhinovirus infection of airway

1855 epithelial cells uncovers the non-ciliated subset as a likely driver of genetic risk to childhood-
1856 onset asthma. Cell Genom. *4*, 100636.

1857 122. Otto, Y., Shipkovenska, G., Liu, Q., Sun, J., Hariri, L.P., Shah, V.S., and Rajagopal, J. (2025).
1858 The identification and characterization of hillocks in the postnatal mouse and human airway.
1859 bioRxiv, 2025.04.15.648833. https://doi.org/10.1101/2025.04.15.648833.

1860 123. Alladina, J., Smith, N.P., Kooistra, T., Slowikowski, K., Kernin, I.J., Deguine, J., Keen, H.L.,
1861 Manakongtreecheep, K., Tantivit, J., Rahimi, R.A., et al. (2023). A human model of asthma
1862 exacerbation reveals transcriptional programs and cell circuits specific to allergic asthma.
1863 Sci. Immunol. *8*, eabq6352.

1864 124. Saha, S., Kanon, N., Islam, M.S., Hassan, M.S., Hooda, Y., Tanvia, L., Rahman, Q.S.-U.,
1865 Das, R.C., Saha, S., Bari, S., et al. (2025). Cohort profile: Health and demographic
1866 surveillance system in mirzapur, tangail, Bangladesh. Int. J. Epidemiol. *54*.
1867 https://doi.org/10.1093/ije/dyaf184.

1868 125. World Health Organization RSV Surveillance Case Definitions. WHO Global Influenza
1869 Programme. https://www.who.int/teams/global-influenza-programme/global-respiratory-
1870 syncytial-virus-surveillance/case-definitions.

1871 126. Trombetta, J.J., Gennert, D., Lu, D., Satija, R., Shalek, A.K., and Regev, A. (2014).
1872 Preparation of single-cell RNA-Seq libraries for next generation sequencing. Curr. Protoc.
1873 Mol. Biol. *107*, 4.22.1–4.22.17.

1874 127. Hu, A., Colella, M., Tam, J.S., Rappaport, R., and Cheng, S.M. (2003). Simultaneous
1875 detection, subgrouping, and quantitation of respiratory syncytial virus A and B by real-time
1876 PCR. Journal of clinical microbiology *41*. https://doi.org/10.1128/jcm.41.1.149-154.2003.

1877 128. Van der Auwera, G., and O'Connor, B. (2020). Genomics in the cloud (O'Reilly Media, Inc.).

1878 129. Kang, H.M., Subramaniam, M., Targ, S., Nguyen, M., Maliskova, L., McCarthy, E., Wan, E.,
1879 Wong, S., Byrnes, L., Lanata, C.M., et al. (2018). Multiplexed droplet single-cell RNA-
1880 sequencing using natural genetic variation. Nat. Biotechnol. *36*, 89–94.

1881 130. Hartoularos, G.C., Si, Y., Zhang, F., Kathail, P., Lee, D.S., Ogorodnikov, A., Sun, Y., Song,
1882 Y.S., Kang, H.M., and Ye, C.J. (2023). Reference-free multiplexed single-cell sequencing
1883 identifies genetic modifiers of the human immune response. bioRxiv, 2023.05.29.542756.
1884 https://doi.org/10.1101/2023.05.29.542756.

1885 131. van der Wijst, M.G.P., Vazquez, S.E., Hartoularos, G.C., Bastard, P., Grant, T., Bueno, R.,
1886 Lee, D.S., Greenland, J.R., Sun, Y., Perez, R., et al. (2021). Type I interferon autoantibodies
1887 are associated with systemic immune alterations in patients with COVID-19. Sci. Transl. Med.
1888 *13*, eabh2624.

1889 132. Malasao, R., Furuse, Y., Okamoto, M., Dapat, C., Saito, M., Saito-Obata, M., Tamaki, R.,
1890 Segubre-Mercado, E., Lupisan, S., and Oshitani, H. (2018). Complete genome sequences of
1891 13 human respiratory syncytial virus subgroup A strains of genotypes NA1 and ON1 isolated
1892 in the Philippines. Genome Announc. *6*, e00151–18.

1893 133. Lin, G.-L., Drysdale, S.B., Snape, M.D., O'Connor, D., Brown, A., MacIntyre-Cockett, G.,
1894 Mellado-Gomez, E., de Cesare, M., Bonsall, D., Ansari, M.A., et al. (2021). Distinct patterns

of within-host virus populations between two subgroups of human respiratory syncytial virus. Nat. Commun. *12*, 5125.

134. Palmenberg, A.C., Spiro, D., Kuzmickas, R., Wang, S., Djikeng, A., Rathe, J.A., Fraser-Liggett, C.M., and Liggett, S.B. (2009). Sequencing and analyses of all known human rhinovirus genomes reveal structure and evolution. Science *324*, 55–59.

135. Oberste, M.S., Maher, K., and Pallansch, M.A. (2004). Evidence for frequent recombination within species human enterovirus B based on complete genomic sequences of all thirty-seven serotypes. J. Virol. *78*, 855–867.

136. Li, B., Gould, J., Yang, Y., Sarkizova, S., Tabaka, M., Ashenberg, O., Rosen, Y., Slyper, M., Kowalczyk, M.S., Villani, A.-C., et al. (2020). Cumulus provides cloud-based data analysis for large-scale single-cell and single-nucleus RNA-seq. Nat. Methods *17*, 793–798.

137. Fleming, S.J., Chaffin, M.D., Arduini, A., Akkad, A.-D., Banks, E., Marioni, J.C., Philippakis, A.A., Ellinor, P.T., and Babadi, M. (2023). Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. Nat. Methods *20*, 1323–1335.

138. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., et al. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat. Biotechnol. *42*, 293–304.

139. Lemieux, J.E., Siddle, K.J., Shaw, B.M., Loreth, C., Schaffner, S.F., Gladden-Young, A., Adams, G., Fink, T., Tomkins-Tinch, C.H., Krasilnikova, L.A., et al. (2021). Phylogenetic analysis of SARS-CoV-2 in Boston highlights the impact of superspreading events. Science *371*, eabe3261.

140. Weller, J. mg2sc: Analyzing metagenomics of single cell RNA-seq (Github).

141. zellkonverter Bioconductor. https://www.bioconductor.org/packages/release/bioc/html/zellkonverter.html.

142. DropletUtils Bioconductor. https://bioconductor.org/packages/release/bioc/html/DropletUtils.html.

143. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

144. Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. *20*, 296.

145. Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: Computational identification of cell Doublets in Single-cell transcriptomic data. Cell Syst. *8*, 281–291.e9.

146. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15.

147. Virshup, I., Rybakov, S., Theis, F.J., Angerer, P., and Wolf, F.A. (2024). anndata: Access and store annotated data matrices. J. Open Source Softw. *9*, 4371.

148. Gehman, L.T., Stoilov, P., Maguire, J., Damianov, A., Lin, C.-H., Shiue, L., Ares, M., Jr, Mody,

I., and Black, D.L. (2011). The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. Nat. Genet. *43*, 706–711.

149. Lu, Z., Reddy, M.V.V.V.S., Liu, J., Kalichava, A., Liu, J., Zhang, L., Chen, F., Wang, Y., Holthauzen, L.M.F., White, M.A., et al. (2016). Molecular architecture of contactin-associated protein-like 2 (CNTNAP2) and its interaction with contactin 2 (CNTN2). J. Biol. Chem. *291*, 24133–24147.

150. Tomita, H., Cornejo, F., Aranda-Pino, B., Woodard, C.L., Rioseco, C.C., Neel, B.G., Alvarez, A.R., Kaplan, D.R., Miller, F.D., and Cancino, G.I. (2020). The protein tyrosine phosphatase receptor delta regulates developmental neurogenesis. Cell Rep. *30*, 215–228.e5.

151. Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196.

152. Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat. Biotechnol. *38*, 1408–1414.

153. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498.

154. Fernandes, J.D., Zamudio-Hurtado, A., Clawson, H., Kent, W.J., Haussler, D., Salama, S.R., and Haeussler, M. (2020). The UCSC repeat browser allows discovery and visualization of evolutionary conflict across repeat families. Mob. DNA *11*, 13.

155. Müller-Dott, S., Tsirvouli, E., Vazquez, M., Ramirez Flores, R.O., Badia-I-Mompel, P., Fallegger, R., Türei, D., Lægreid, A., and Saez-Rodriguez, J. (2023). Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities. Nucleic Acids Res. *51*, 10934–10949.

156. Badia-I-Mompel, P., Vélez Santiago, J., Braunger, J., Geiss, C., Dimitrov, D., Müller-Dott, S., Taus, P., Dugourd, A., Holland, C.H., Ramirez Flores, R.O., et al. (2022). decoupleR: ensemble of computational methods to infer biological activities from omics data. Bioinform. Adv. *2*, vbac016.

157. Kyle Kimler, M.E.-B. (2025). SETA: Single Cell Ecological Taxonomic Analysis. Preprint.

158. Dolgalev, I. (2025). msigdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format. Preprint.

159. Browaeys, R., Saelens, W., and Saeys, Y. (2020). NicheNet: modeling intercellular communication by linking ligands to target genes. Nat. Methods *17*, 159–162.

160. Kuhn, M. (2008). Building Predictive Models inRUsing thecaretPackage. J. Stat. Softw. *28*. https://doi.org/10.18637/jss.v028.i05.

161. 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.

162. Cerezo, M., Sollis, E., Ji, Y., Lewis, E., Abid, A., Bircan, K.O., Hall, P., Hayhurst, J., John, S.,

1972 Mosaku, A., et al. (2025). The NHGRI-EBI GWAS Catalog: standards for reusability,
1973 sustainability and diversity. Nucleic Acids Res. *53*, D998–D1005.

1974 163. de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized
1975 gene-set analysis of GWAS data. PLoS Comput. Biol. *11*, e1004219.

Main Figures

Figure 1: Cellular diversity in the pediatric nasal mucosa response to common respiratory viruses

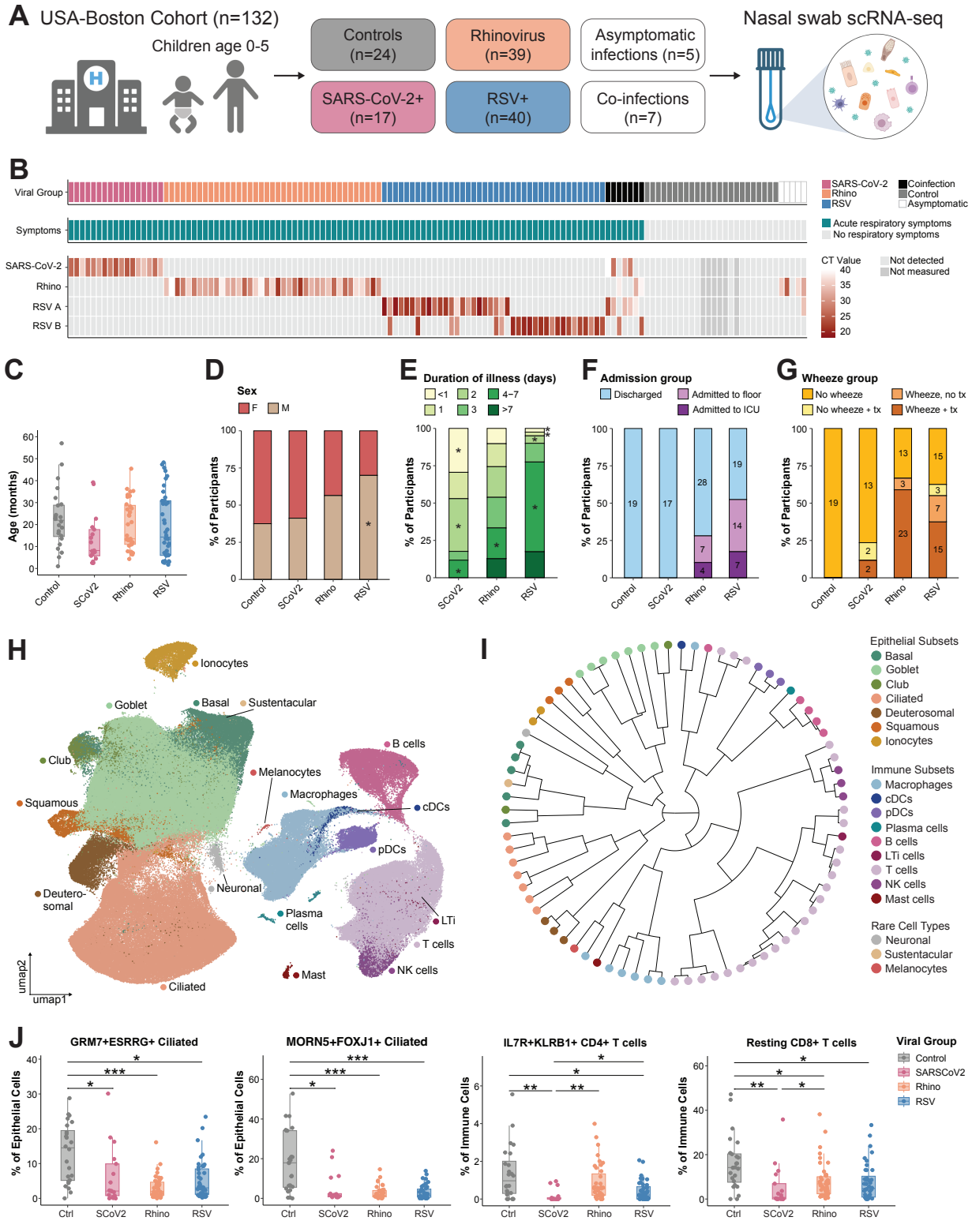
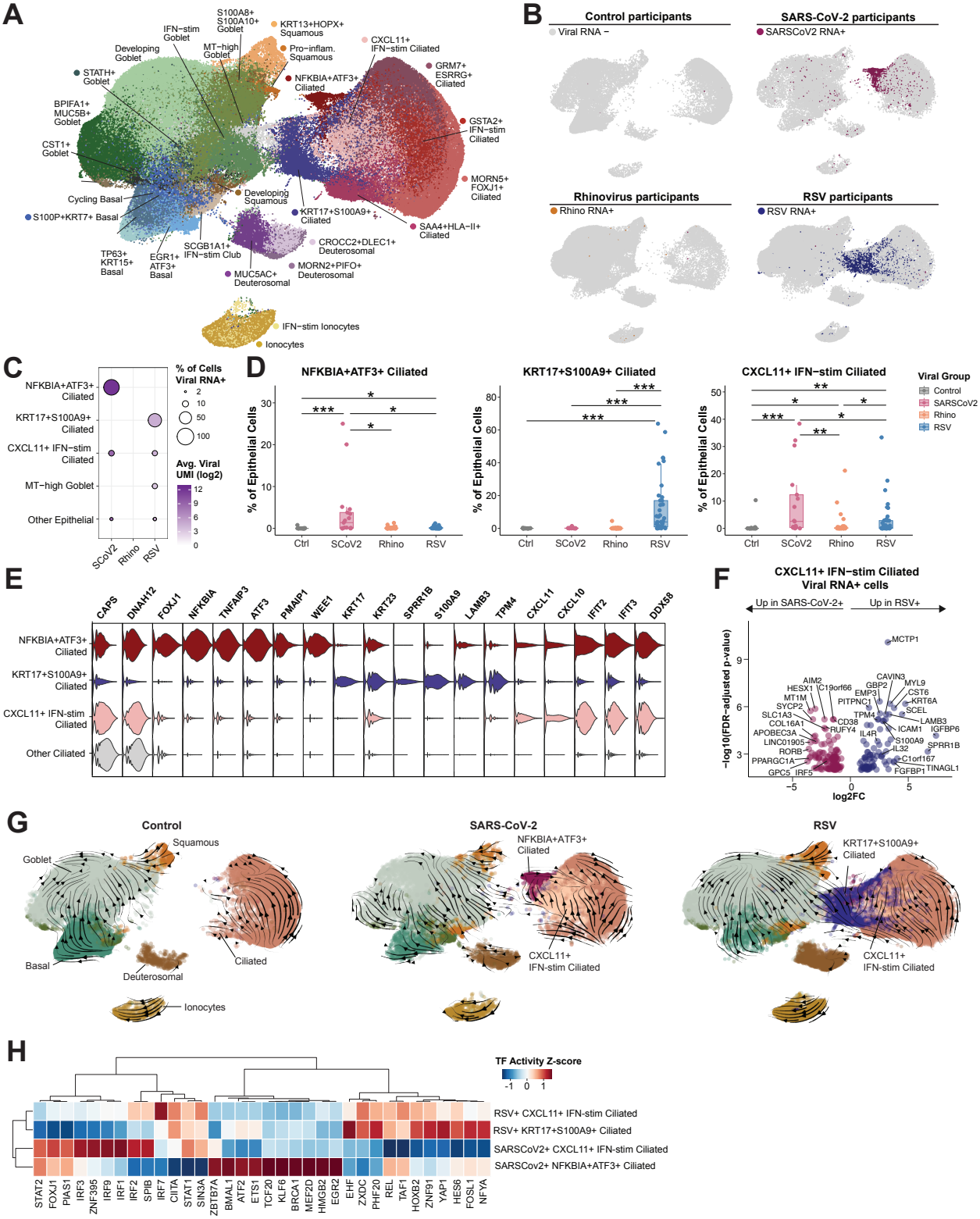


Figure 1: Cellular diversity in the pediatric nasal mucosa response to common respiratory viruses

- **A.** Schematic of participant cohort and study design. RSV = respiratory syncytial virus; scRNA-seq = single-cell RNA sequencing. Participant metadata is summarized in **Supplementary Table 1** and provided in full in **Supplementary Table 2**.
- **B.** Heatmap of participant cohort colored by assigned viral group (top), the presence of acute respiratory symptoms (middle), and viral RNA cycle threshold (CT) value from supernatant collected during swab processing measured by qPCR (bottom). 1 asymptomatic participant was assigned based on a positive PCR from clinically sampled viral transport media. Not detected: CT value >40 or undetectable; Not measured: supernatant was not available; Rhino = rhinovirus.
- **C.** Age of each participant in months by viral group. SCoV2 = SARS-CoV-2.
- **D.** Distribution of sex within each viral group. *chi-sq. standardized Pearson residual > 2.
- **E.** Distribution of reported duration of illness within each viral group. *chi-sq. standardized Pearson residual > 2.
- **F.** Distribution of hospital admission within each viral group. Control samples are restricted to those collected at Boston Children's Hospital (BCH). ICU = intensive care unit.
- **G.** Distribution of wheezing groups (see **Methods** for assignment criteria) within each viral group. Control samples are restricted to those collected at BCH. Tx = wheezing-directed medication.
- **H.** UMAP of all cells (n=335,174) from all participants colored by cell type. Cell type labels were assigned after iterative clustering. cDCs = conventional dendritic cells; pDCs = plasmacytoid dendritic cells; LTi = lymphoid tissue inducer cells; NK = natural killer.
- **I.** Hierarchical tree of all 79 cell subsets identified after iterative clustering. End nodes are colored by cell type of origin. Hierarchical distances were calculated from pseudobulk aggregates of each cell subset. See **Supplementary Figure 1G** for subset labels.
- **J.** Frequency of specified cell subsets as a percentage of either all epithelial or all immune cells per sample, compared across viral groups. Samples with <100 epithelial or immune cells were excluded. Ctrl = Control.

C,J: Boxplots represent the median (center line), interquartile range (box), and 1.5x the interquartile range (whiskers). Statistical test is Kruskal-Wallis test with Benjamini-Hochberg correction for multiple comparisons. Dunn's post-hoc test, *p<0.05, **p<0.001, ***p<0.0001.

Figure 2: RSV and SARS-CoV-2 each induce highly specific infected epithelial cell states



2014 **Figure 2: RSV and SARS-CoV-2 each induce highly specific infected epithelial cell states**

2015 **A.** UMAP of epithelial cells (n=224,431) from participants in control and monoinfection groups
2016 colored by cell subset. Pro-inflam = pro-inflammatory; MT-high = mitochondrial-high; IFN-
2017 stim = interferon -stimulated.

2018 **B.** Split UMAP of epithelial cells from each viral group (Control: n=34,031; SARS-CoV-2:
2019 n=39,836; Rhinovirus: n=61,198; RSV: n=89,366) colored by viral RNA status.

2020 **C.** Dotplot of viral RNA+ cells in each epithelial subset for SARS-CoV-2, rhinovirus, and RSV.
2021 Dot size represents the percentage of cells in the subset that are viral RNA+. Color
2022 represents the average viral transcripts (UMI) across all cells in the subset. Epithelial
2023 subsets with >5% of cells viral RNA+ were included, other subsets were collapsed. SCoV2
2024 = SARS-CoV-2.

2025 **D.** Frequency of specified viral RNA+ cell subsets as a percentage of epithelial cells per
2026 sample, compared across viral groups. Samples with <100 epithelial cells were excluded.
2027 Boxplots represent the median (center line), interquartile range (box), and 1.5x the
2028 interquartile range (whiskers). Statistical test is Kruskal-Wallis test with Benjamini-
2029 Hochberg correction for multiple comparisons. Dunn's post-hoc test, *p<0.05, **p<0.001,
2030 ***p<0.0001. Ctrl = Control; Rhino = Rhinovirus.

2031 **E.** Violin plot of marker genes for select viral RNA+ epithelial subsets.

2032 **F.** Differentially expressed (DE) genes between RSV RNA+ cells and SARS-CoV-2 RNA+
2033 cells within the *CXCL11*+IFN-stim ciliated subset. Participants with at least 20 RNA+
2034 *CXCL11*+IFN-stim ciliated cells (SARS-CoV-2: n=4; RSV: n=6) were included in
2035 pseudobulk DE analysis. Wald significance test with Benjamini-Hochberg correction.
2036 Genes with adjusted p<0.05 are shown. Full list of DE genes is provided in
2037 **Supplementary Table 4**.

2038 **G.** RNA velocity streams for all epithelial cells split by viral group for Control, SARS-CoV-2,
2039 and RSV groups, colored by cell type with *NFKBIA*+*ATF3*+ ciliated, *CXCL11*+IFN-stim
2040 ciliated, and *KRT17*+*S100A9*+ ciliated subsets highlighted.

2041 **H.** Heat map of inferred transcription factor (TF) activity, z-score normalized between groups
2042 of cells, for select viral RNA+ cells, split by *NFKBIA*+*ATF3*+ ciliated, *CXCL11*+IFN-stim
2043 ciliated, and *KRT17*+*S100A9*+ ciliated subsets, showing the top 10 transcription factors
2044 for each group of cells.

Figure 3: SARS-CoV-2 induces a more robust IFN response compared to rhinovirus and RSV

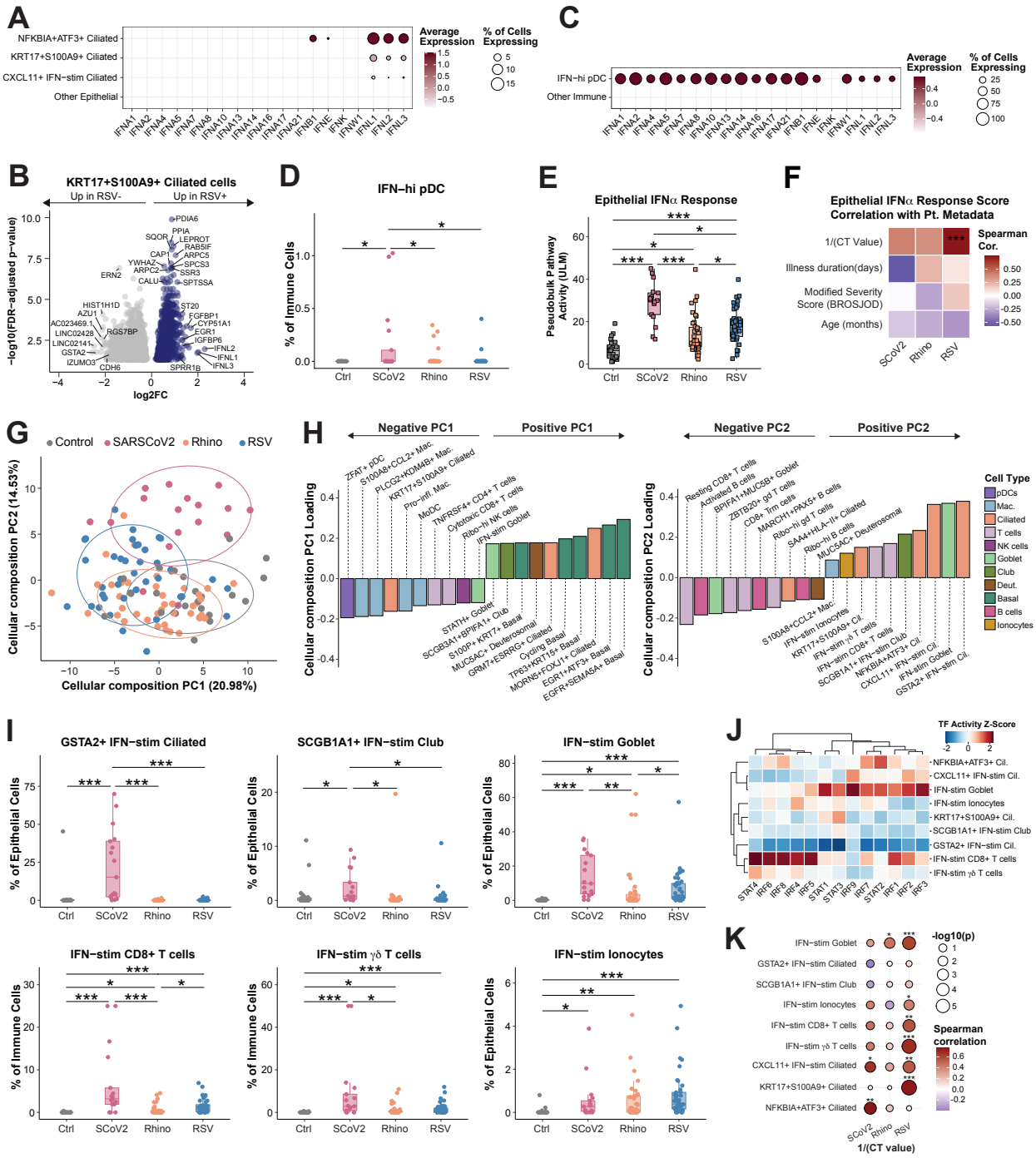


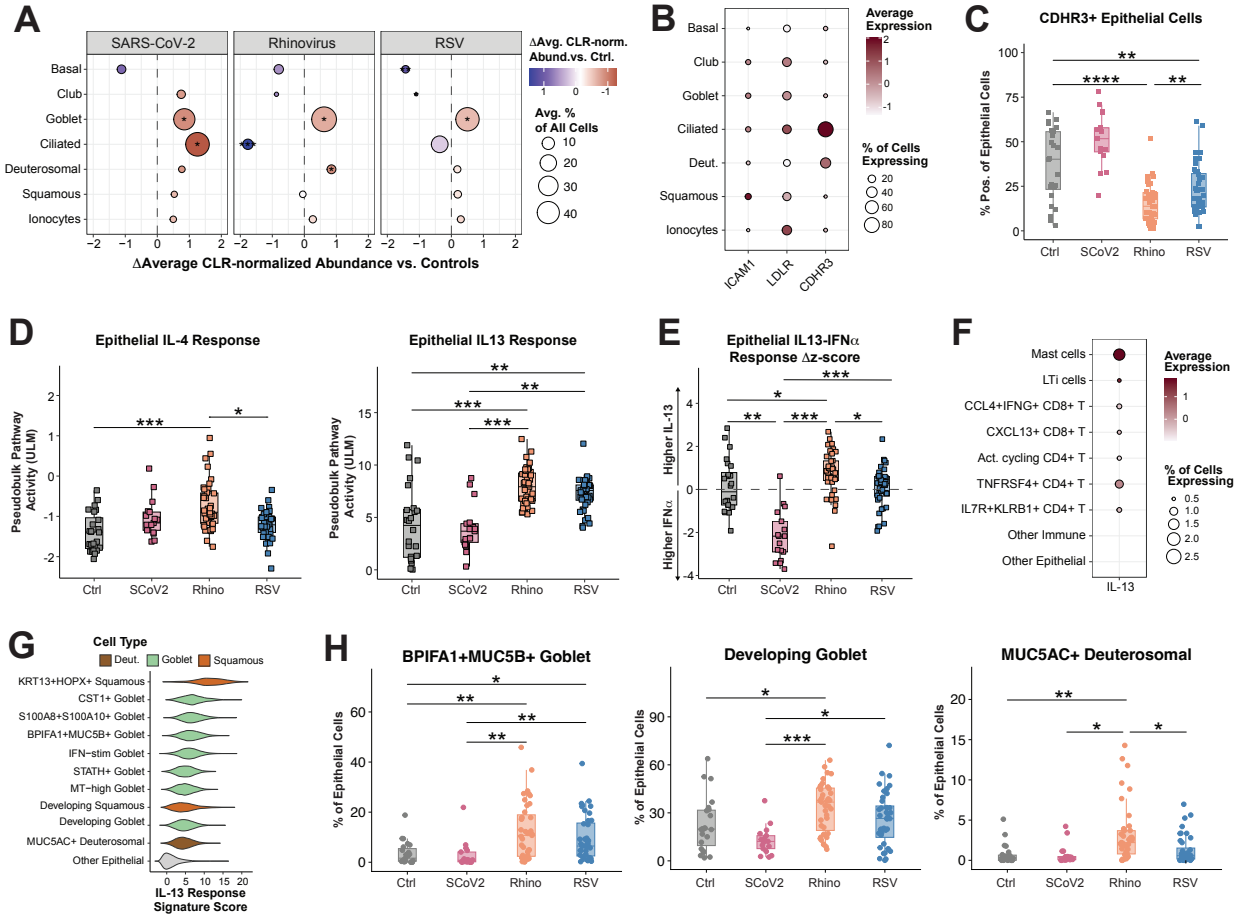
Figure 3: SARS-CoV-2 induces a more robust IFN response compared to rhinovirus and RSV

- **A.** Dotplot of IFN gene expression by epithelial subset. Epithelial subsets with >1% of cells expressing any IFN gene were included, other subsets were collapsed.
- **B.** Differentially expressed (DE) genes between RSV RNA+ cells and RSV RNA- cells within the *KRT17+S100A9+* ciliated subset. Participants with at least 20 *KRT17+S100A9+* ciliated cells (n=26) were included in pseudobulk DE analysis. Wald significance test with Benjamini-Hochberg correction. Genes with adjusted $p<0.05$ are shown. Full list of DE genes is provided in **Supplementary Table 4**.
- **C.** Dotplot of IFN gene expression by immune subset. Immune subsets with >2% of cells expressing any IFN gene were included, other subsets were collapsed.
- **D.** Frequency of IFN-hi plasmacytoid dendritic cells (pDCs) as a percentage of immune cells per sample, compared across viral groups. Ctrl = Control; SCoV2 = SARS-CoV-2; Rhino = Rhinovirus.
- **E.** Pseudobulk pathway activity score per participant for interferon-alpha (IFNα) response in epithelial cells, compared across viral groups. ULM = univariate linear model. Signature provided in **Supplementary Table 5**.
- **F.** Spearman correlation between pseudobulk epithelial IFNα pathway activity score and select participant metadata variables within each viral group. CT value represents viral RNA PCR from scRNA-seq swab supernatants.
- **G.** Principal component analysis (PCA) of cellular composition for each sample. PCA was run on center-log-ratio (CLR) normalized cell subset abundances with percent of variance explained by each PC labeled on x and y axes. Ellipses for each viral group are centered at centroid and represent 80% confidence region.
- **H.** Cell subsets ranked by loadings for principal components 1 (left) and 2 (right). Top 10 positive and negatively ranked subsets are shown. Mac. = macrophage; Cil. = ciliated; Deut. = deuterosomal.
- **I.** Frequency of specified IFN-responsive epithelial and immune subsets as a percentage of immune or epithelial cells per sample, compared across viral groups.
- **J.** Transcription factor (TF) activity z-scores for select TFs across select IFN-responsive epithelial and immune subsets.
- **K.** Spearman correlation between CLR-normalized abundance of select cell subsets and inverse CT value from viral RNA PCR of scRNA-seq swab supernatants for each virus.

D,I: Participants with <100 immune or epithelial cells were excluded.

D,E,I: Boxplots represent the median (center line), interquartile range (box), and 1.5x the interquartile range (whiskers). Statistical test is Kruskal-Wallis test with Benjamini-Hochberg correction for multiple comparisons. Dunn's post-hoc test, *$p<0.05$, **$p<0.001$, ***$p<0.0001$.

Figure 4: Rhinovirus is associated with secretory-cell polarization of the epithelium.



2087 **Figure 4: Rhinovirus is associated with secretory-cell polarization of the epithelium.**

2088 **A.** Difference in average (Avg.) center-log-ratio (CLR)-normalized cell type abundance
2089 (Abund.) compared to controls (Ctrl.) for epithelial cell types. Dot size represents
2090 frequency of the cell type out of all cells in the sample, averaged per viral group. Statistical
2091 test is Wilcoxon test with Benjamin-Hochberg correction for multiple comparisons.
2092 *p<0.05, **p<0.01, ***p<0.001.

2093 **B.** Dotplot of rhinovirus receptor gene expression by epithelial cell type.

2094 **C.** Percentage of epithelial cells expressing the rhinovirus receptor gene *CDHR3* per
2095 participant, compared across viral groups. Cells were defined as CDHR3+ if they
2096 contained >1 *CHDR3* UMI. Statistical test is Wilcoxon test with Benjamin-Hochberg
2097 correction for multiple comparisons. *p<0.05, **p<0.01, ***p<0.001. Ctrl = Control; SCoV2
2098 = SARS-CoV-2; Rhino = Rhinovirus.

2099 **D.** Pseudobulk pathway activity score per participant for IL-4 (left) and IL-13 (right) response
2100 genes in epithelial cells, compared across viral groups. ULM = univariate linear model.
2101 Signatures provided in **Supplementary Table 5**.

2102 **E.** Difference in scaled Z-scores of pseudobulk IL-13 and IFNα activity scores in epithelial
2103 cells per participant, compared across viral groups.

2104 **F.** Dotplot of IL-13 gene expression by cell subset. Subsets with >0.5% of cells expressing
2105 IL-13 or average scaled expression >2 were included, other subsets were collapsed. Act.
2106 = activated.

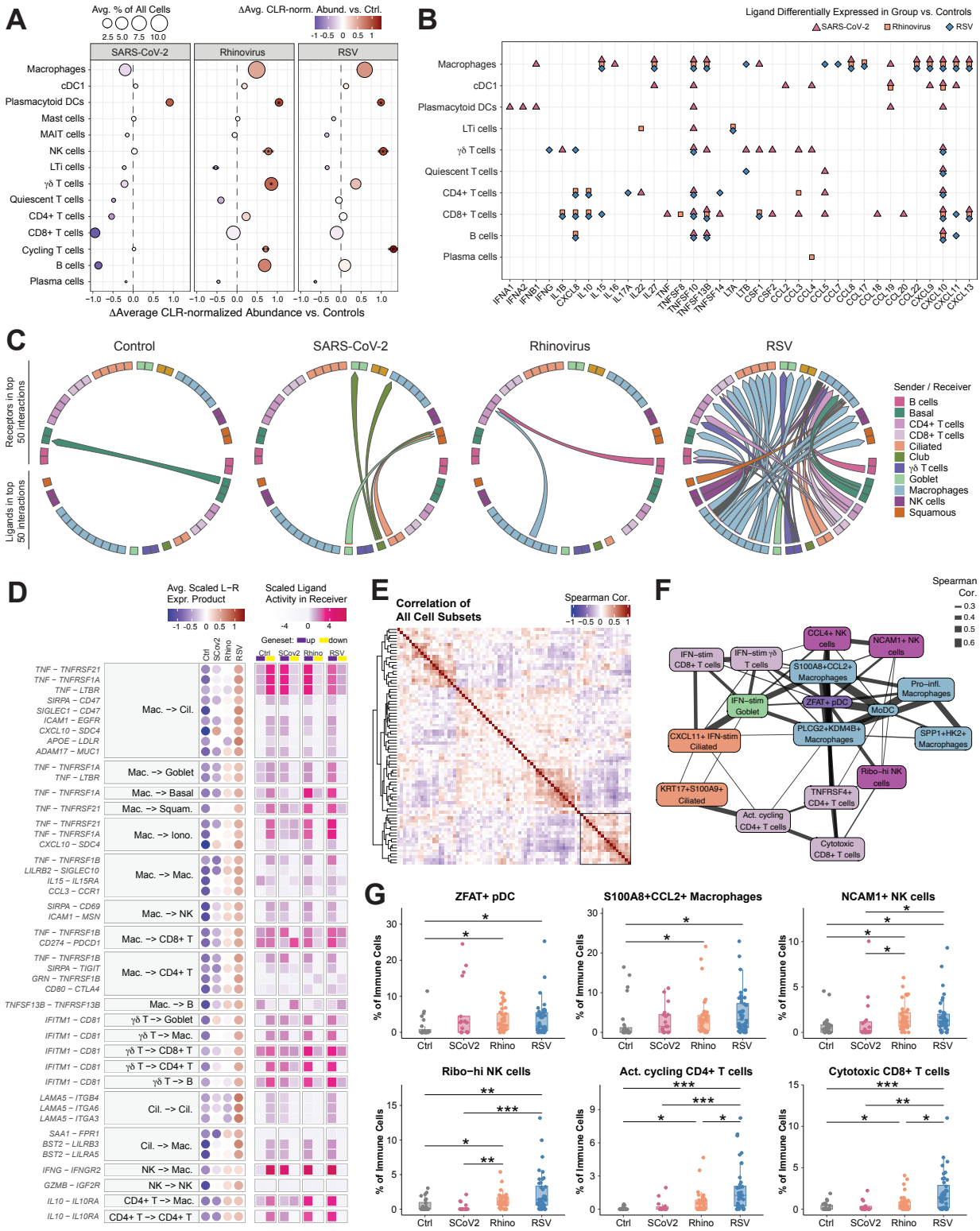
2107 **G.** Top 10 epithelial cell subsets by IL-13 response signature. Scoring was done on a
2108 downsampled object containing 10% of cells from each epithelial cell subset.

2109 **H.** Frequency of specified epithelial subsets as a percentage of epithelial cells per sample,
2110 compared across viral groups. Participants with <100 epithelial cells were excluded.

2111 **C,D,E,H:** Boxplots represent the median (center line), interquartile range (box), and 1.5x the
2112 interquartile range (whiskers).

2113 **D,E,H:** Statistical test is Kruskal-Wallis test with Benjamini-Hochberg correction for multiple
2114 comparisons. Dunn's post-hoc test, *p<0.05, **p<0.001, ***p<0.0001.

Figure 5: Nasal immune cell responses to viral infection reveal a heightened inflammatory and cytotoxic response in RSV.



2117

Figure 5: Nasal immune cell responses to viral infection reveal a heightened inflammatory and cytotoxic response in RSV.

A. Difference in center-log-ratio (CLR)-normalized cell type abundance compared to controls for immune cell types. Dot size represents frequency of the cell type out of all cells in the sample, averaged per viral group. cDC1 = conventional type 1 dendritic cell; MAIT = mucosal-associated invariant T; NK = natural killer; LTi = lymphoid tissue inducer.

B. Immune-signaling ligands per immune cell type differentially expressed (DE) relative to controls for each viral group. Cell types where at least 2 samples in every group had at least 5 cells were included. Cell type - ligand combinations with significant results (FDR <0.05) for at least one viral group are shown. See **Supplementary Table 6** for list of ligands evaluated.

C. Chord diagram of top 50 overall ranked cell-cell interactions (multinichenet), divided by group specificity. Starting point of chord represents ligand and endpoint represents receptor. Chords are colored by sender cell type. See **Supplementary Figure 6A** for ligand and receptor labels and **Supplementary Table 6** for list of top 500 ranked interactions.

D. Prioritization criteria for select cell-cell interactions in the RSV group within the top 100 overall. Interactions with sender cell types in the top 5% percent of interactions are shown. In cases where ligand-receptor-receiver combinations are not unique, interaction with the highest prioritization score is shown. Circles: scaled product of pseudobulk expression of ligand in sender and receptor in receiver, averaged across samples in the group. Squares: scaled ligand activity score in receiver cell type, representing enrichment of ligand target genes in differentially expressed (DE) genes; up: upregulated genes in that group, down: downregulated genes in that group. Ctrl = Control; SCoV2 = SARS-CoV-2; Rhino = Rhinovirus, Avg. Scaled L-R Expr. Product = average scaled ligand-receptor expression product.

E. Spearman correlation (cor.) of CLR-normalized cell subset abundances among all samples. Both rows and columns were clustered using hierarchical clustering with Euclidean distance and complete linkage.

F. Network plot of cell subsets highlighted in the box in **E**. Edge weight represents spearman correlation. Nodes are colored by cell type of origin.

G. Frequency of specified immune subsets as a percentage of immune cells per sample, compared across viral groups. Participants with <100 immune cells were excluded. Boxplots represent the median (center line), interquartile range (box), and 1.5x the interquartile range (whiskers). Statistical test is Kruskal-Wallis test with Benjamini-Hochberg correction for multiple comparisons. Dunn's post-hoc test, *p<0.05, **p<0.001, ***p<0.0001.

Figure 6: Cell states associated with severe RSV infection and rhinovirus-induced wheeze.

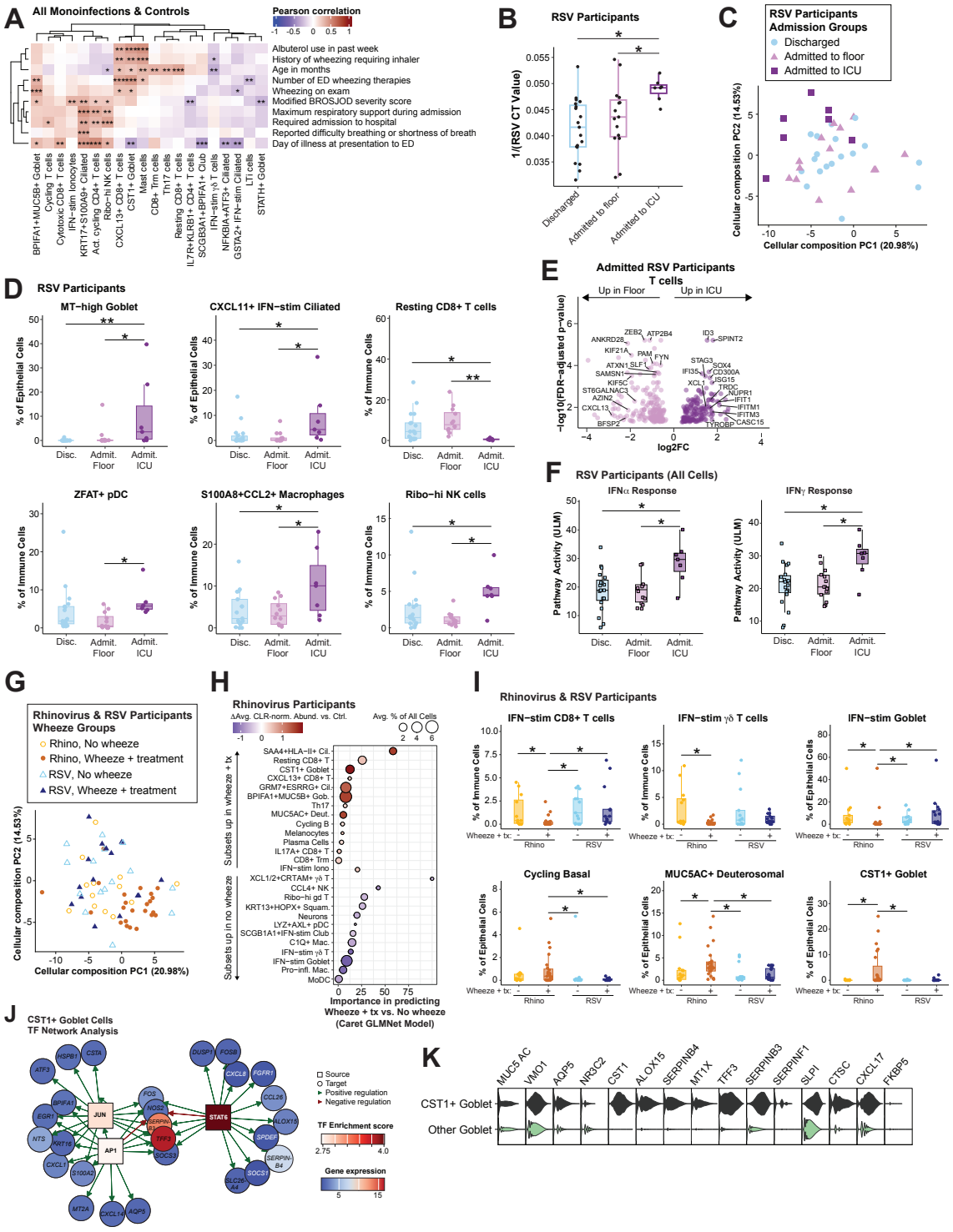


Figure 6: Cell states associated with severe RSV infection and rhinovirus-induced wheeze.

A. Pearson correlation of CLR-normalized cell subset abundances among control, SARS-CoV-2, RSV, and rhinovirus participants. Both rows and columns were clustered using hierarchical clustering with Euclidean distance and complete linkage. P values from correlation test adjusted with Benjamini-Hochberg correction for multiple comparisons. *p<0.05, **p<0.01, ***p<0.001. ED = emergency department.

B. Viral RNA level (1/CT value) from PCR of scRNA-seq swab supernatants for each sample in the RSV group divided by hospital admission status. ICU = intensive care unit.

C. Principal component analysis (PCA) of cellular composition for each sample in RSV participants colored by admission groups. The percent of variance explained by each PC is labeled on x and y axes. PCA was run on center-log-ratio (CLR) normalized cell subset abundances. See **Figure 3G** for PC loadings.

D. Frequency of specified subsets as a percentage of epithelial or immune cells in RSV participants compared across admission groups. Disc. = Discharged, Admit. floor = Admitted to floor, Admit ICU = Admitted to ICU.

E. Differentially expressed (DE) genes in all T cells between RSV participants admitted to ICU or admitted to the floor by pseudobulk differential expression. Wald significance test with Benjamini-Hochberg correction. Genes with adjusted p value of <0.05 are displayed. Full list of DE genes is provided in **Supplementary Table 4**.

F. Pseudobulk pathway activity score for IFNα (left) and IFNɣ (right) response genes per RSV participant in all cells, compared across admission groups. ULM = univariate linear model. Signatures provided in **Supplementary Table 5**.

G. PCA of cellular composition for RSV and Rhinovirus participants colored by wheezing group (see **Figure 1E** for group breakdown). Treatment included nebulized bronchodilator and/or corticosteroids.

H. Machine-learning-derived subset importance for distinguishing no wheeze vs wheeze + treatment for Rhinovirus participants by Caret's penalized generalized linear (GLMNet) model. Color of dot indicates the difference in center-log-ratio (CLR)-normalized cell subset abundance between patients with wheeze+tx compared to no wheeze. Dot size represents frequency of the cell subset, averaged across all included rhinovirus participants. All subsets that contributed to model importance are shown. Cil. = ciliated; Gob. = goblet; Deut. = deuterosomal; Iono. = ionocytes; Squam = squamous; Mac. = macrophages.

I. Frequency of specified subsets as a percentage of epithelial or immune cells for RSV and Rhinovirus participants categorized as having no wheeze/no treatment or wheeze + treatment. Tx = treatment; Rhino = rhinovirus.

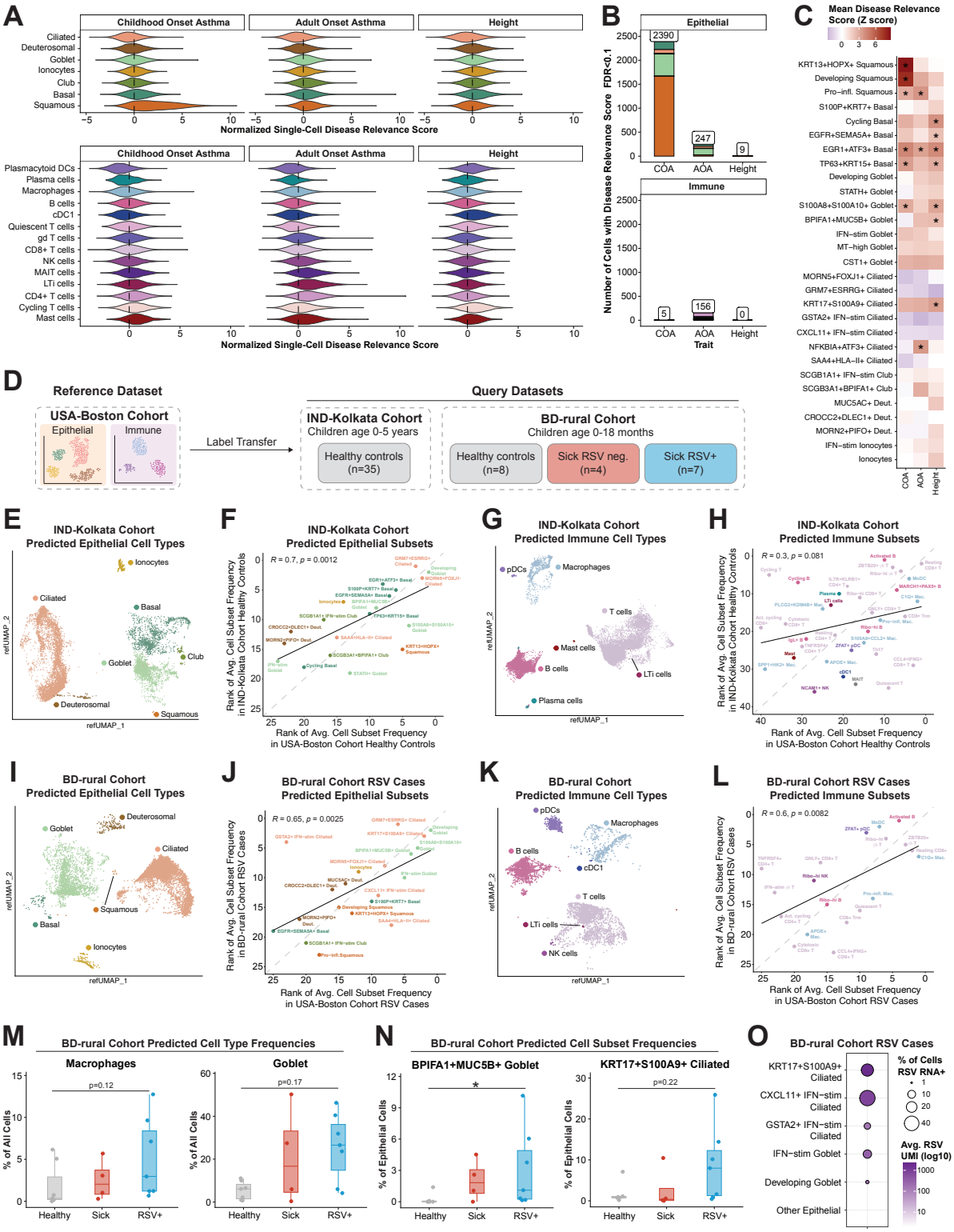
J. Transcription factor (TF) network analysis for the top 3 transcription factors within the *CST1*+ goblet cell subset. Boxes: transcription factors colored by enrichment score, circles: target genes colored by gene expression relative to all epithelial cells.

K. Violin plot of marker genes for *CST1*+ goblet compared to all other goblet cells.

D,I: Participants with <100 epithelial or immune cells were excluded.

B,D,F,I: Boxplots represent the median (center line), interquartile range (box), and 1.5x the interquartile range (whiskers). Statistical test is Kruskal-Wallis test with Benjamini-Hochberg correction for multiple comparisons. Dunn's post-hoc test, *p<0.05, **p<0.001, ***p<0.0001

Figure 7: Applying the nasal cell atlas to understand genetic and geographical contributors to disease.

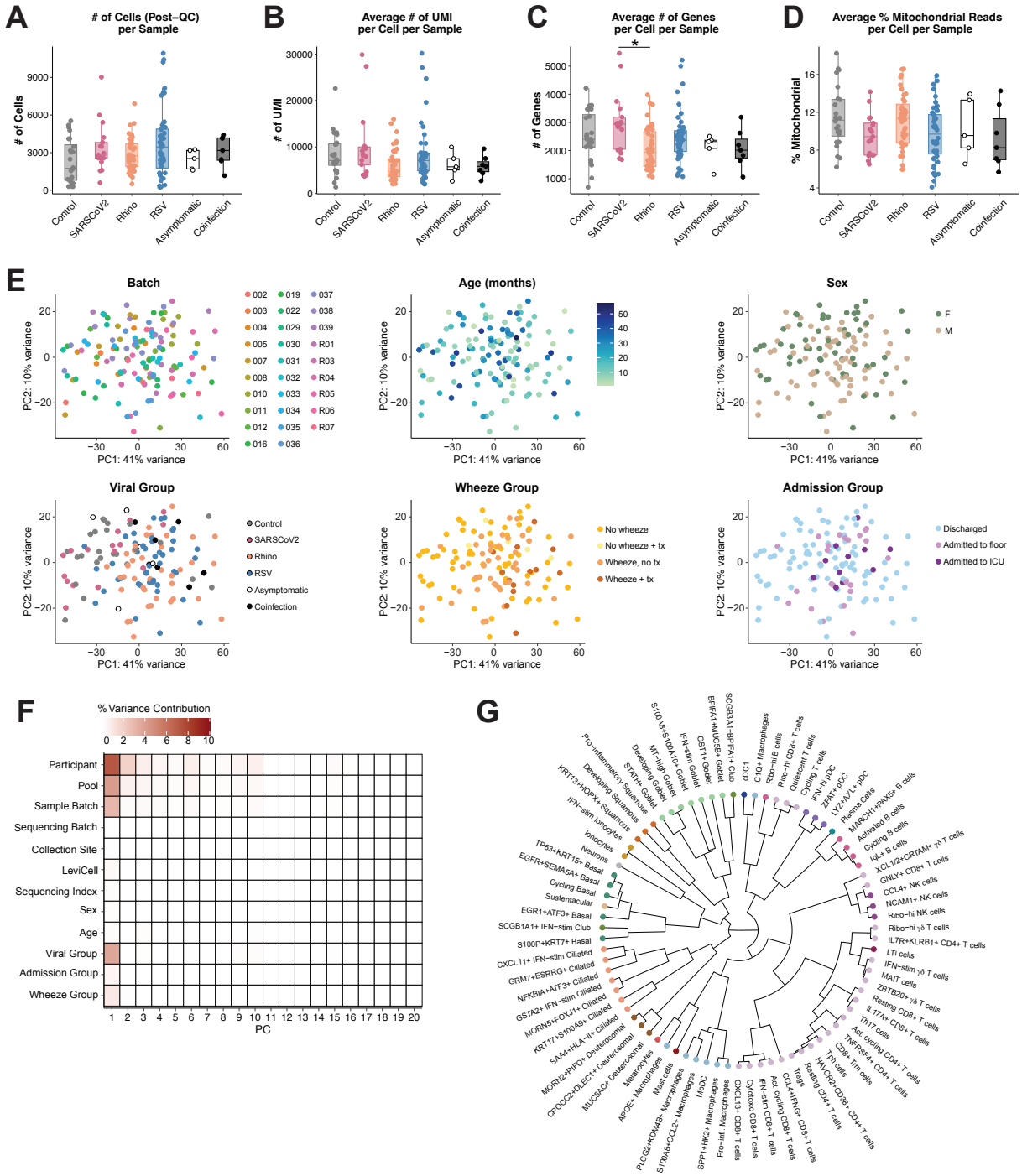


2205 **Figure 7: Applying the nasal cell atlas to understand genetic and geographical**
2206 **contributors to disease.**

2207 **A.** Violin plot of normalized single-cell Disease Relevance Scores (scDRS) for childhood
2208 onset asthma, adult onset asthma, and height genome-wide association study (GWAS)
2209 traits, split by epithelial (top) and immune (bottom) cell type.
2210 **B.** Number of cells with a significant scDRS score at 10% FDR for childhood onset asthma
2211 (COA), adult onset asthma (AOA), and height. Cell types are colored as in **(A)**.
2212 **C.** Cell type-trait association Z-score for childhood onset asthma (COA), adult onset asthma
2213 (AOA), and height across epithelial subsets. *FDR adjusted p<0.05.
2214 **D.** Schematic of label transfer approach. Cell type and cell subset annotations were
2215 transferred separately from epithelial and immune cells from the USA-Boston cohort to
2216 two separate query cohorts. See **Methods** and **Supplementary Table 2** for cohort details.
2217 **E,G**. UMAP of epithelial cells (n=13,514) **(E)** and immune cells (n=28,342) **(G)** from all
2218 samples in the IND-Kolkata cohort colored by predicted cell type. Cells with prediction
2219 score >0.8 are shown.
2220 **F, H.** Spearman correlation of cell subsets from controls ranked by average frequency
2221 among epithelial cells **(F)** and immune cells **(H)**. X axis, rank in healthy controls in USA-
2222 Boston cohort. Y axis, rank in healthy controls in IND-Kolkata cohort. Subsets are colored
2223 by cell type of origin as in **E, G**.
2224 **I, K**. UMAP of epithelial cells (n=12,715) **(I)** and immune cells (n=5,266) **(K)** from all
2225 samples in the BD-rural cohort colored by predicted cell type. Cells with prediction score
2226 >0.8 are shown.
2227 **J,L.** Spearman correlation of cell subsets from RSV cases ranked by average
2228 frequency among epithelial cells **(J)** and immune cells **(L)**. X axis, rank in RSV cases in
2229 USA-Boston cohort. Y axis, rank in RSV cases in BD-rural cohort. Subsets are colored by
2230 cell type of origin as in **I,K.**
2231 **M**. Frequency of select predicted cell types out of all cells per sample in BD-rural cohort,
2232 divided by disease group. Participants with <100 total cells were excluded.
2233 **N.** Frequency of select predicted epithelial cell subsets out of all epithelial cells per sample
2234 in BD-rural cohort, divided by disease group. Participants with <100 epithelial cells were
2235 excluded.
2236 **O.** Dotplot of RSV RNA+ cells in each epithelial subset in all RSV participants in BD-rural
2237 cohort. Dot size represents the percentage of cells in the subset that are viral RNA+. Color
2238 represents the average viral transcripts (UMI) across all cells in the subset.
2239 **M,N:** Boxplots represent the median (center line), interquartile range (box), and 1.5x the
2240 interquartile range (whiskers). Statistical test is Kruskal-Wallis test with Benjamini-
2241 Hochberg correction for multiple comparisons. Dunn's post-hoc test, *p<0.05, **p<0.001,
2242 ***p<0.0001.
2243

Supplementary Figures

Supplementary Figure 1, Related to Figure 1

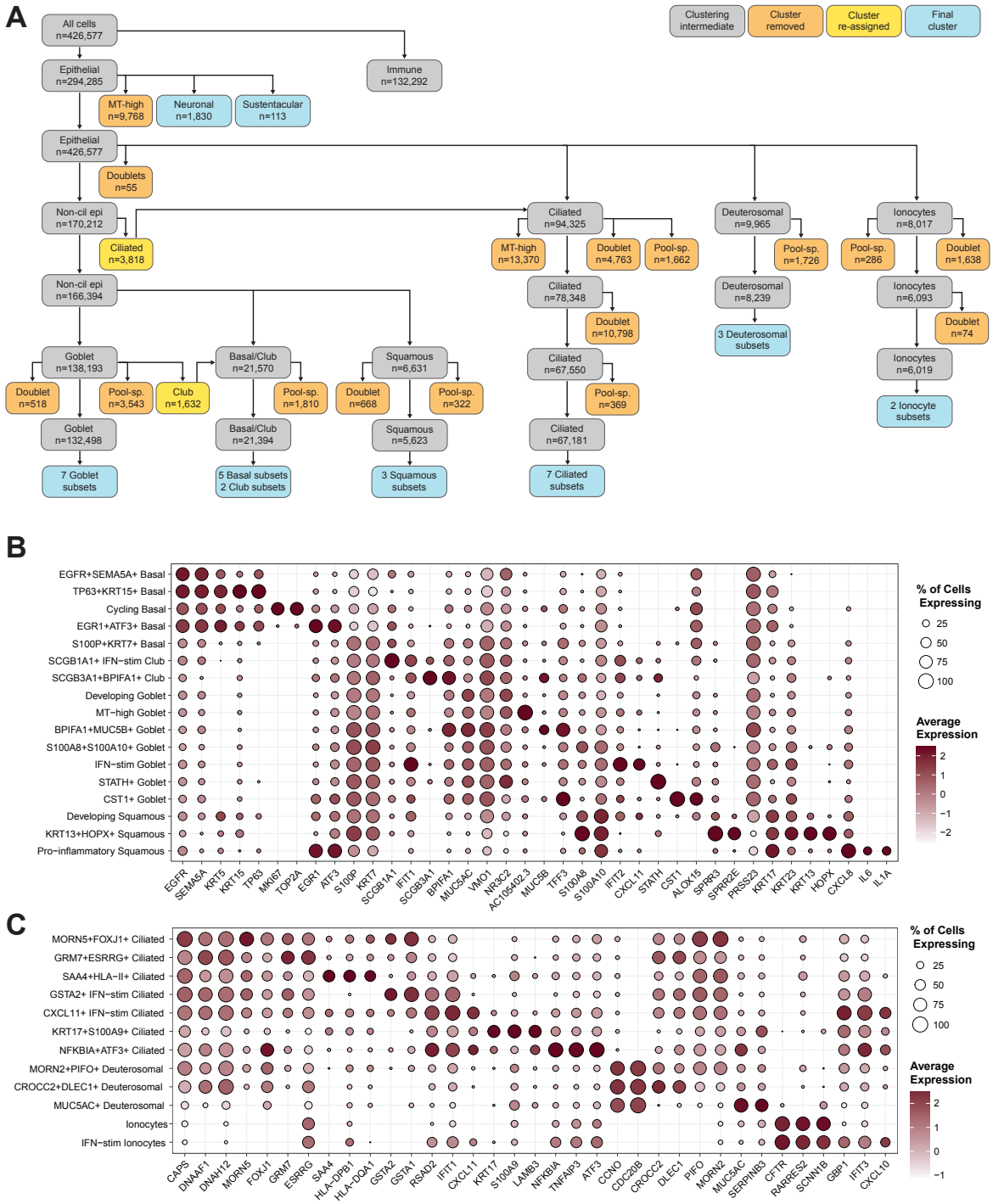


2246

2247 **Supplementary Figure 1, Related to Figure 1**

2248 **A.** Number of cells per sample after selecting high-quality cells (defined as >500 unique
2249 molecular identifiers (UMI), >300 unique genes, and <35% mitochondrial reads) by viral
2250 group. Metrics in **B-D** are restricted to high-quality cells. QC = quality control; Rhino =
2251 rhinovirus.
2252 **B.** Average number of UMI per cell per sample by viral group.
2253 **C.** Average number of unique genes per cell per sample by viral group.
2254 **D.** Average percentage of mitochondrial reads per cell per sample by viral group.
2255 **E.** Principal component analysis (PCA) of pseudobulk aggregated gene expression for each
2256 sample, colored by metadata variables. F = female; M = male; tx = wheezing-directed
2257 medication; ICU = intensive care unit.
2258 **F.** Variance contribution matrix for metadata variables of interest and top 20 principal
2259 components (PC) from single-cell gene expression data. Color represents the percentage
2260 of variance in the dataset explained by each variable through each PC.
2261 **G.** Hierarchical tree of cell subsets as in **Fig. 1I** with subset labels. End nodes are colored by
2262 cell type of origin. Hierarchical distances were calculated from pseudobulk aggregates of
2263 each cell subset. IFN-stim = interferon- stimulated; MT = mitochondrial; MoDC =
2264 monocyte-derived dendritic cell; cDC1 = conventional type 1 dendritic cell; pDC =
2265 plasmacytoid dendritic cell; Ribo. = ribosomal; LTi = Lymphoid tissue inducer cell; MAIT =
2266 mucosal-associated invariant T cell; Th17 = T helper cell type 17; Act. = activated; Trm =
2267 resident memory T cell; Tph = T peripheral helper cell; Treg = regulatory T cell.
2268 **A-D:** Boxplots represent the median (center line), interquartile range (box), and 1.5x the
2269 interquartile range (whiskers). Statistical test is Kruskal-Wallis test with Benjamini-Hochberg
2270 correction for multiple comparisons. Dunn's post-hoc test, *$p<0.05$

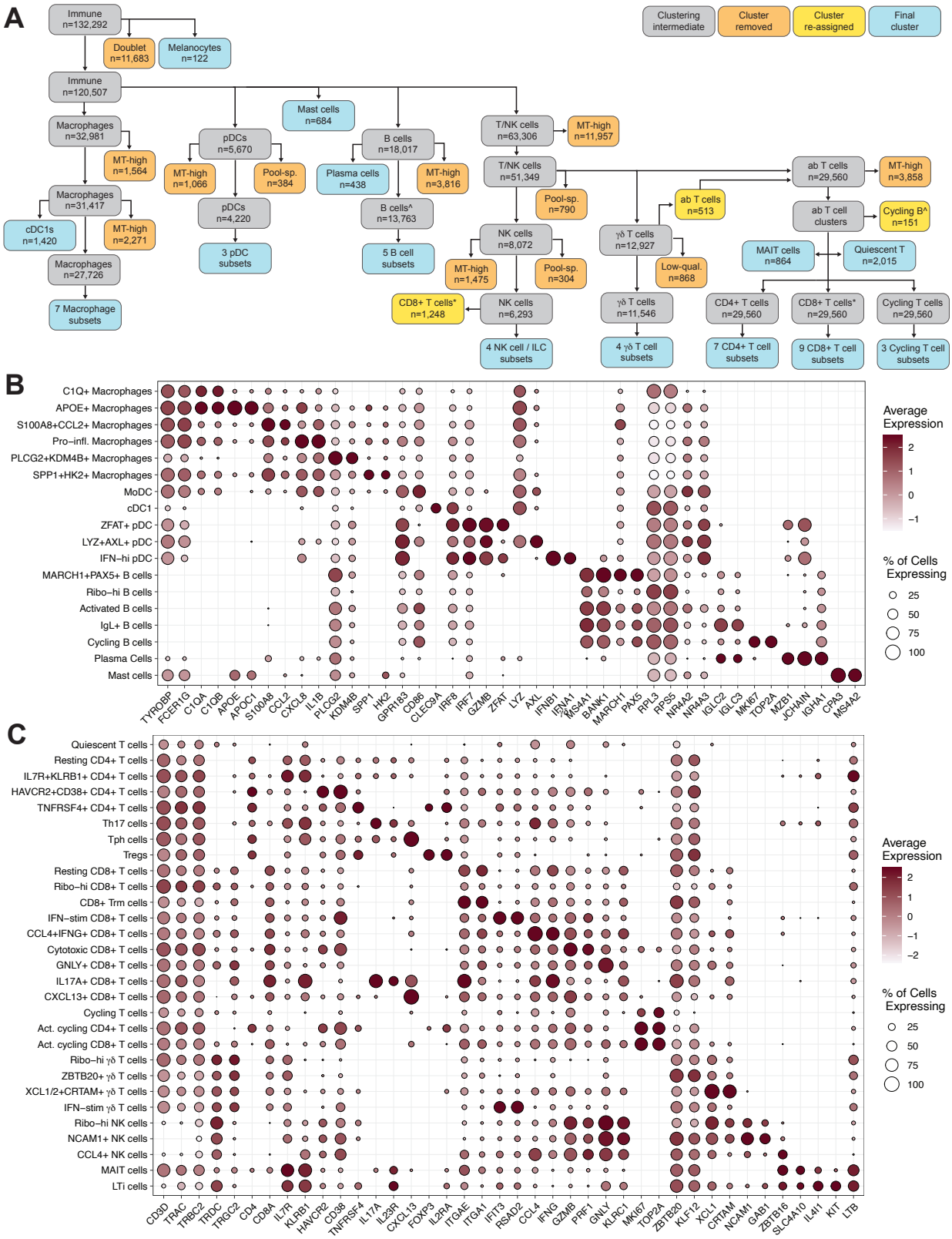
Supplementary Figure 2, Related to Figure 1



2273 **Supplementary Figure 2, Related to Figure 1**

2274 **A.** Schematic of epithelial cell clustering approach. Clusters were removed at each level if
2275 they were defined by only mitochondrial genes, had >75% of cells from one sample pool,
2276 or were suspected doublets (see **Methods**). Clusters were re-assigned if they expressed
2277 marker genes for a different cell type to a higher degree than the marker genes for the cell
2278 type of interest. Normalization was repeated at each level. Final clusters were determined
2279 after comparing multiple resolutions. Pool-sp = pool-specific, Non-cil = non-ciliated.
2280 **B.** Marker genes for basal, club, goblet, and squamous cell subsets. IFN-stim = interferon-
2281 stimulated.
2282 **C.** Marker genes for ciliated, deuterosomal, and ionocyte subsets.
2283 **B-C.** See **Supplementary Table 3** for complete lists of marker genes for final cell subsets.

Supplementary Figure 3, Related to Figure 1



2286 **Supplementary Figure 3, Related to Figure 1**

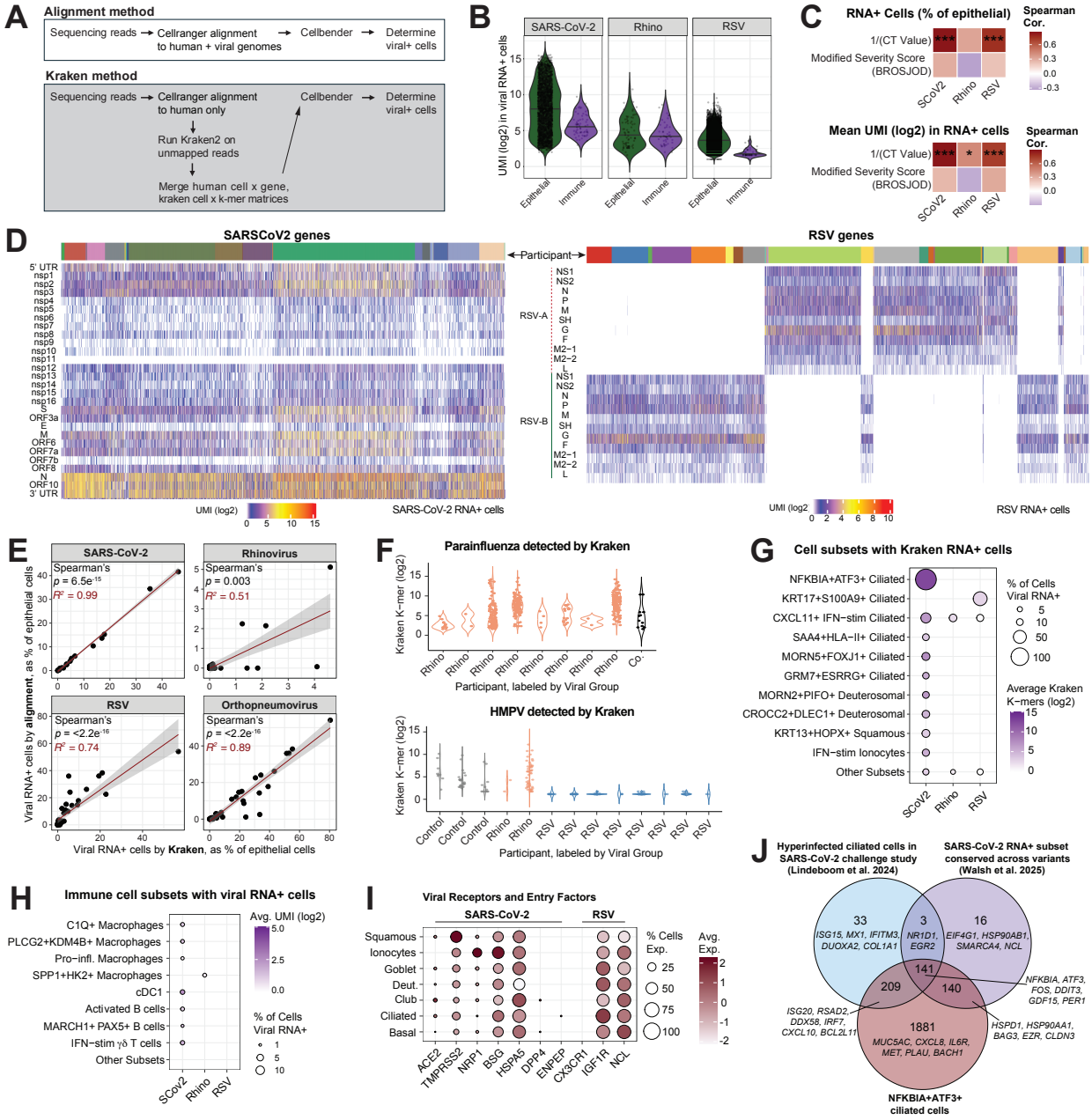
2287 **A.** Schematic for immune cell clustering approach. See legend for **Supplementary Figure 2**
2288 for further information. Pool-sp = pool-specific.^* denotes where re-assigned clusters were
2289 included if not shown with arrows.

2290 **B.** Marker genes for macrophage, dendritic cell, B cell, and mast cell subsets. MoDC =
2291 monocyte-derived dendritic cell; cDC1 = conventional type 1 dendritic cell; pDC =
2292 plasmacytoid dendritic cell.

2293 **C.** Marker genes for T cell and natural killer (NK) cell subsets. Th17 = T helper cell type 17;
2294 Tph = T peripheral helper cell; Treg = regulatory T cell; Ribo. = ribosomal; Trm = resident
2295 memory T cell; Act. = activated; IFN-stim = interferon- stimulated; MAIT = mucosal-
2296 associated invariant T cell; LTi = Lymphoid tissue inducer cell.

2297 **B-C.** See **Supplementary Table 3** for complete lists of marker genes for final cell subsets.

Supplementary Figure 4, Related to Figure 2



2300 **Supplementary Figure 4, Related to Figure 2**

2301 **A.** Schematic for identifying viral RNA+ cells by alignment or Kraken2 metagenomics
2302 approach. See Methods for details.

2303 **B.** Violin plot demonstrating viral UMI/cell for epithelial compared to immune viral RNA+ cells
2304 for SARS-CoV-2, rhinovirus, and RSV. Rhino = rhinovirus. RSV = respiratory syncytial
2305 virus.

2306 **C.** Spearman correlation between viral RNA+ cells as a percentage of epithelial cells and
2307 inverse CT value from qRT-PCR of supernatant collected during swab processing and
2308 disease severity (see **Methods** for details) for each virus (top) and modified BROSJOD
2309 disease severity score (see **Methods** for details, bottom). SCoV2 = SARS-CoV-2.
2310 Spearman correlation test with Benjamini-Hochberg correction for multiple comparisons.
2311 *p<0.05, ***p<0.001.

2312 **D.** Heatmap of SARS-CoV-2 (left) and RSV (right) gene expression in viral RNA+ cells. Top
2313 bar depicts the individual participant for each cell. UMIs shown are post removal of
2314 background reads by Cellbender.

2315 **E.** Spearman correlation between viral RNA+ cells as a percentage of epithelial cells between
2316 alignment and Kraken methods for respective viral groups. R^2 and Spearman *p* values are
2317 shown. For RSV, two different Kraken2 levels of classification are depicted. The plot
2318 labeled "RSV" indicates the custom kraken database species-level identifier "Respiratory
2319 Syncytial Virus" compared to RSV+ cells by alignment method. The plot labeled
2320 "Orthopneumovirus" indicates the custom kraken database genus-level identifier
2321 "Orthopneumovirus" compared to RSV+ cells by alignment method.

2322 **F.** Violin plot of kraken k-mers for parainfluenza or human metapneumovirus (HMPV) per
2323 participant for participants with >1 viral RNA+ cell for parainfluenza or HMPV. Participants
2324 are labeled by viral group. Co. = coinfection.

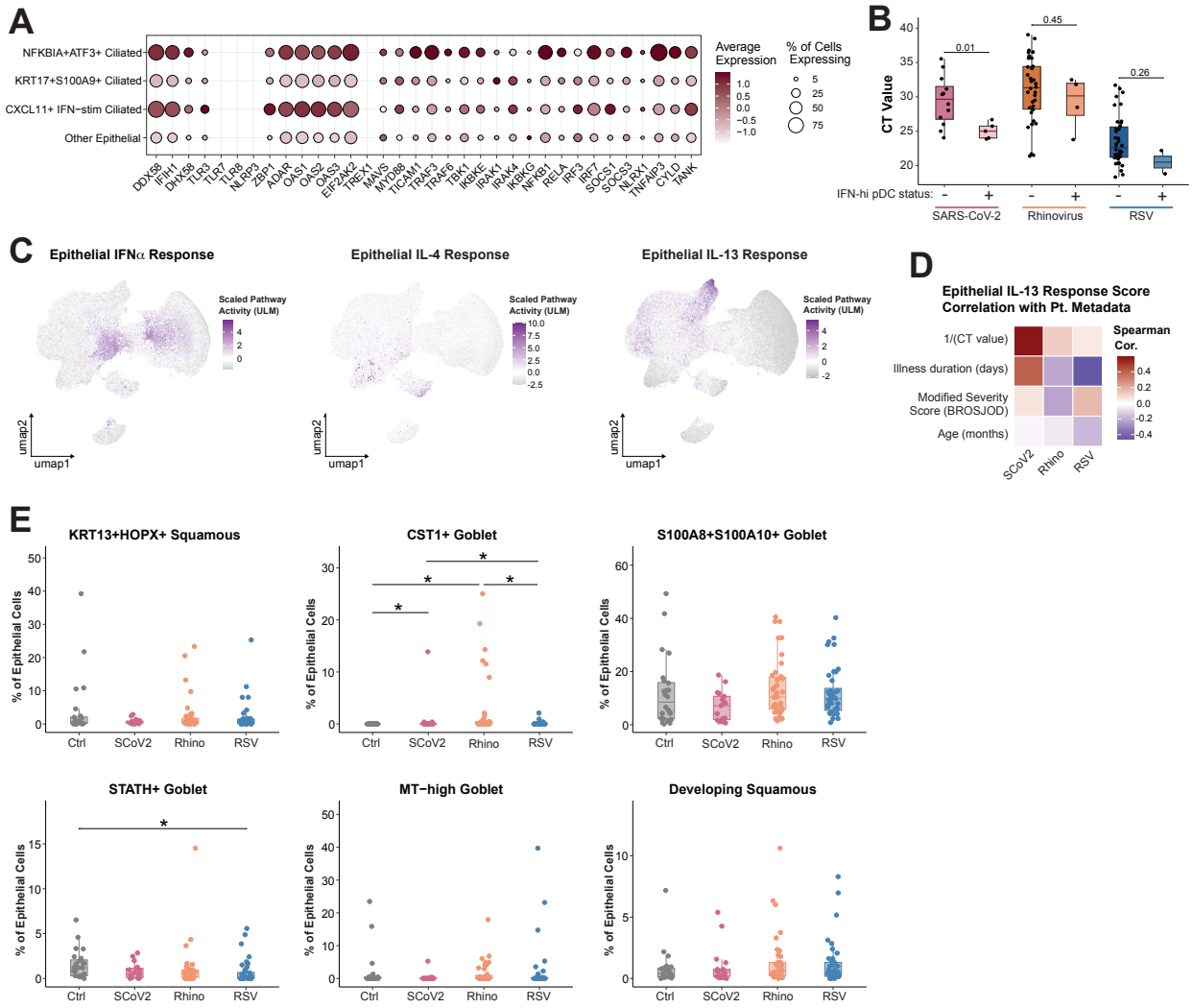
2325 **G.** Dotplot of viral RNA+ cells identified by Kraken method in all subsets for SARS-CoV-2,
2326 rhinovirus, and RSV. Dot size represents the percentage of cells in the subset that are
2327 viral RNA+. Color represents the average viral k-mers across all cells in the subset. Cell
2328 subsets with <5% viral RNA+ cells are collapsed. Cell subsets with fewer than 100
2329 cells/group were excluded.

2330 **H.** Dotplot of viral RNA+ cells by alignment method in each immune subset for SARS-CoV-
2331 2, rhinovirus, and RSV. Dot size represents the percentage of cells in the subset that are
2332 viral RNA+. Color represents the average viral transcripts (UMI) across all cells in the
2333 subset. Cell subsets with <1% viral RNA+ cells are collapsed. Cell subsets with fewer than
2334 100 cells/group were excluded.

2335 **I.** Dotplot of SARS-CoV-2 and RSV viral receptor and entry factors within each epithelial cell
2336 type. Includes all epithelial cells from all participants.

2337 **J.** Venn diagram comparing marker genes for *NFKBIA+ATF3+* ciliated cells in this dataset,
2338 "hyperinfected" ciliated cells identified in a SARS-CoV-2 human challenge study[18] and a
2339 PER2+EGR1+GDF15+ ciliated subset consistently enriched in viral RNA+ cells across
2340 SARS-CoV-2 variants[19].

Supplementary Figure 5, Related to Figures 3 and 4



2343 **Supplementary Figure 5, Related to Figures 3 and 4**

2344 **A.** Dotplot of pattern recognition receptor and antiviral genes in viral RNA+ epithelial cell
2345 subsets.

2346 **B.** Viral RNA CT value from supernatant collected during swab processing divided by viral
2347 group and IFN-hi plasmacytoid dendritic cell status (+ = >0 IFN-hi pDCs in the sample).
2348 Statistical test is a Mann-Whitney U test.

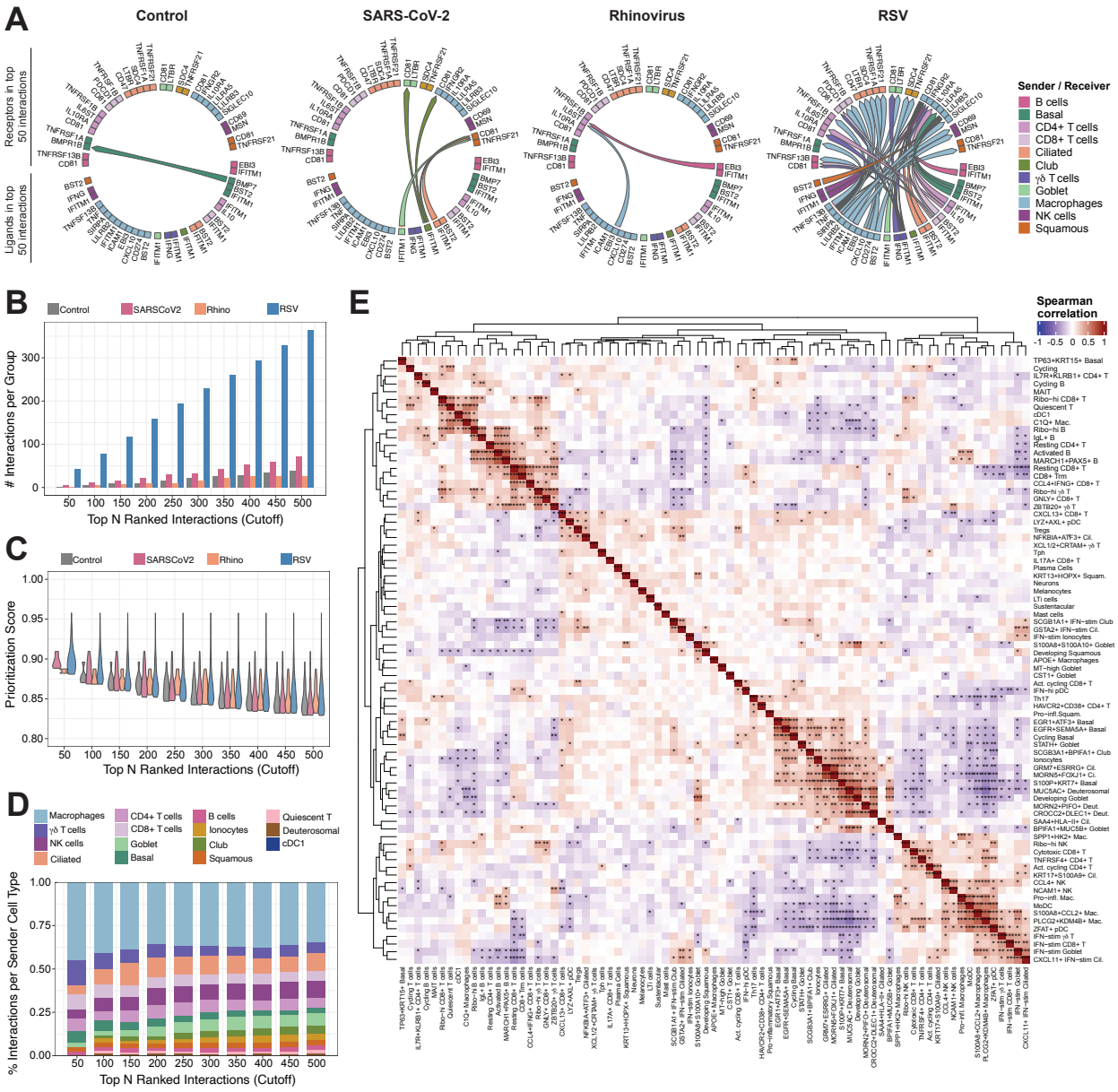
2349 **C.** UMAP of all epithelial cells from participants in control and mono-infection groups colored
2350 by scaled pathway activity score for IFNα (left), IL-4 (center), and IL-13 (right) response
2351 genes. Scoring was done on a downsampled object containing 10% of cells from each
2352 epithelial cell subset. Signatures provided in **Supplementary Table 5**.

2353 **D.** Spearman correlation between pseudobulk epithelial IL-13 pathway activity score and
2354 select participant metadata variables within each viral group. CT value represents viral
2355 RNA CT value from supernatant collected during swab processing. Spearman correlation
2356 test with Benjamini-Hochberg correction for multiple comparisons. *$p<0.05$, otherwise n.s.

2357 **E.** Frequency of specified epithelial subsets as a percentage of epithelial cells per sample,
2358 compared across viral groups. Participants with <100 epithelial cells were excluded.
2359 Statistical test is Kruskal-Wallis test with Benjamini-Hochberg correction for multiple
2360 comparisons. Dunn's post-hoc test, *$p<0.05$.

2361 **B, E:** Boxplots represent the median (center line), interquartile range (box), and 1.5x the
2362 interquartile range (whiskers).

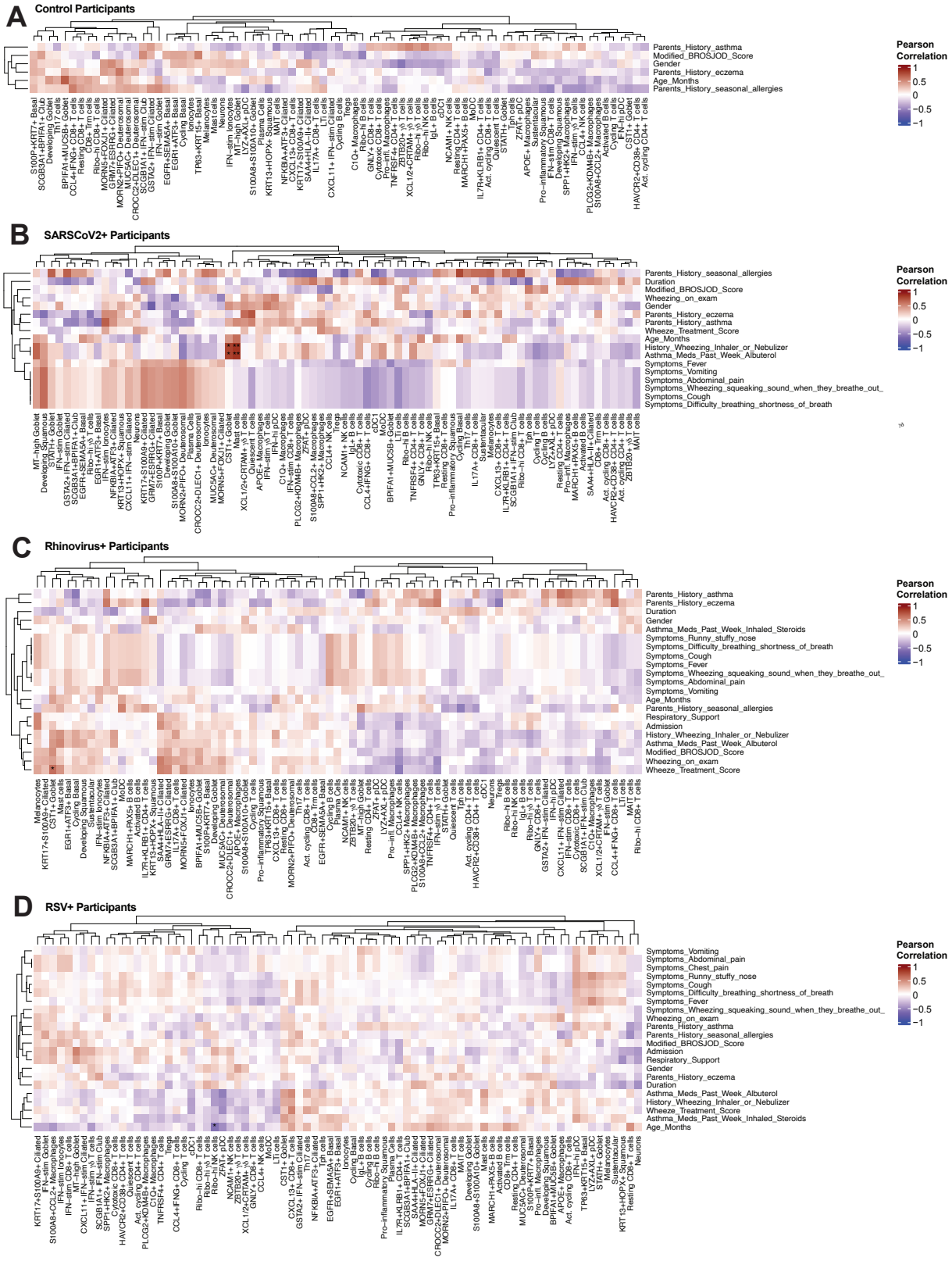
Supplementary Figure 6, Related to Figure 5



Supplementary Figure 6, Related to Figure 5

- **A.** Chord diagrams of top 50 overall ranked cell-cell interactions (multinichenet) as in **Figure 5C**, including ligand receptor labels. Starting point of chord represents the ligand, endpoint represents the receptor. Chords are colored by sender cell type.
- **B.** Number of interactions specific to each viral group in top interactions across various cutoffs. See **Supplementary Table 6** for list of top 500 prioritized interactions.
- **C.** Prioritization scores of interactions specific to each viral group in top interactions across various cutoffs.
- **D.** Distribution of most frequent sender cell types in top interactions across various cutoffs.
- **E.** Spearman correlation of CLR-normalized cell subset abundances among all samples as in **5E**, including subset labels. Both rows and columns were clustered using hierarchical clustering with Euclidean distance and complete linkage. P values from correlation test adjusted with Benjamini-Hochberg correction for multiple comparisons. *$p<0.05$, **$p<0.001$, ***$p<0.0001$.
- **F.** Frequency of specified immune subsets as a percentage of immune cells per sample, compared across viral groups. Participants with <100 immune cells were excluded. Boxplots represent the median (center line), interquartile range (box), and 1.5x the interquartile range (whiskers). Statistical test is Kruskal-Wallis test with Benjamini-Hochberg correction for multiple comparisons. Dunn's post-hoc test, *$p<0.05$, **$p<0.001$.

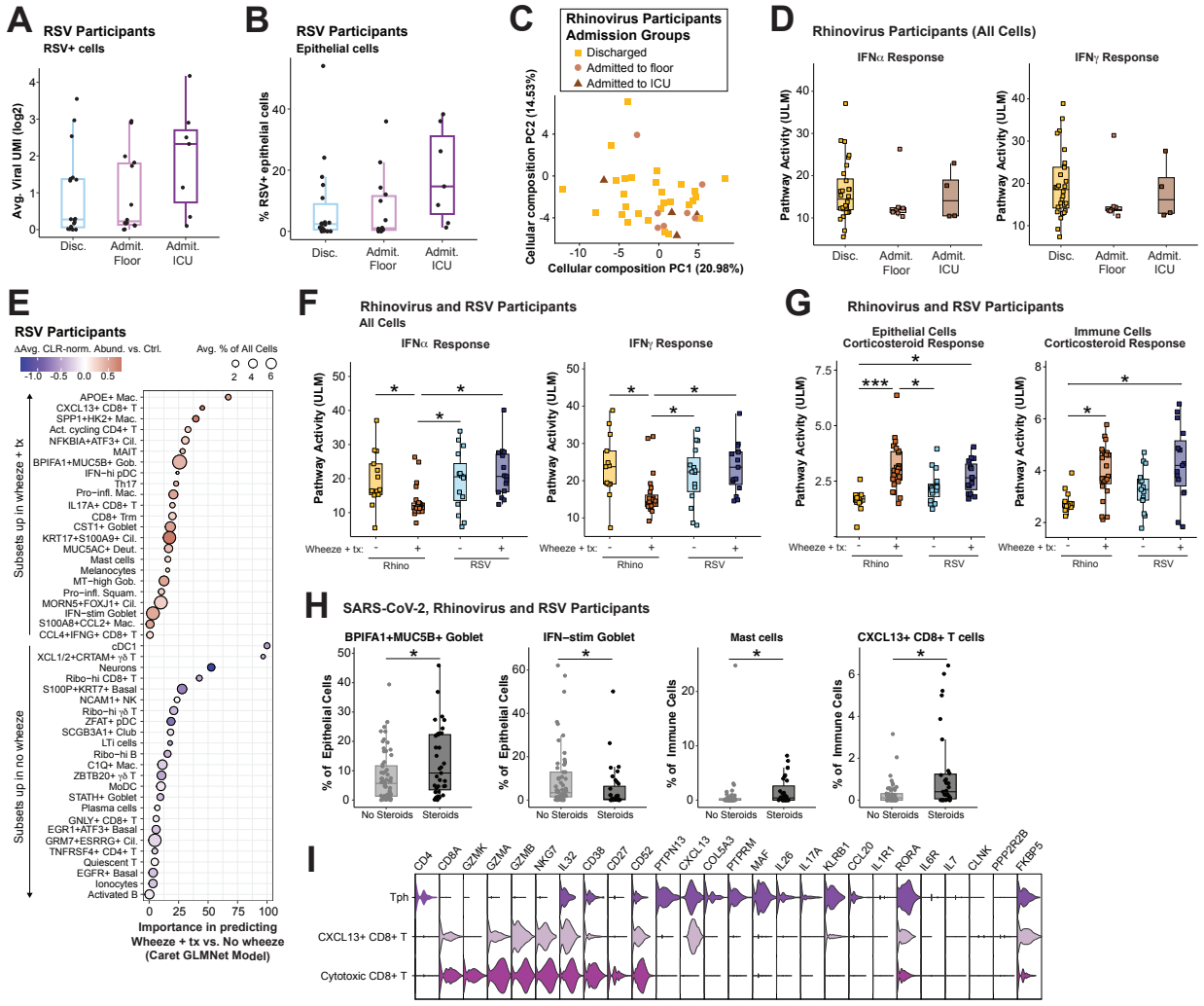
Supplementary Figure 7, Related to Figure 6



2386 **Supplementary Figure 7, Related to Figure 6**

2387 **A-D.** Pearson correlation of CLR-normalized cell subset abundances among **(A)** control,
2388 **(B)**, SARS-CoV-2, **(C)** Rhinovirus participants, and **(D)** RSV participants. Both rows and
2389 columns were clustered using hierarchical clustering with Euclidean distance and
2390 complete linkage. Pearson correlation test with Benjamini-Hochberg correction for multiple
2391 comparisons. *$p < 0.05$, **$p < 0.01$.

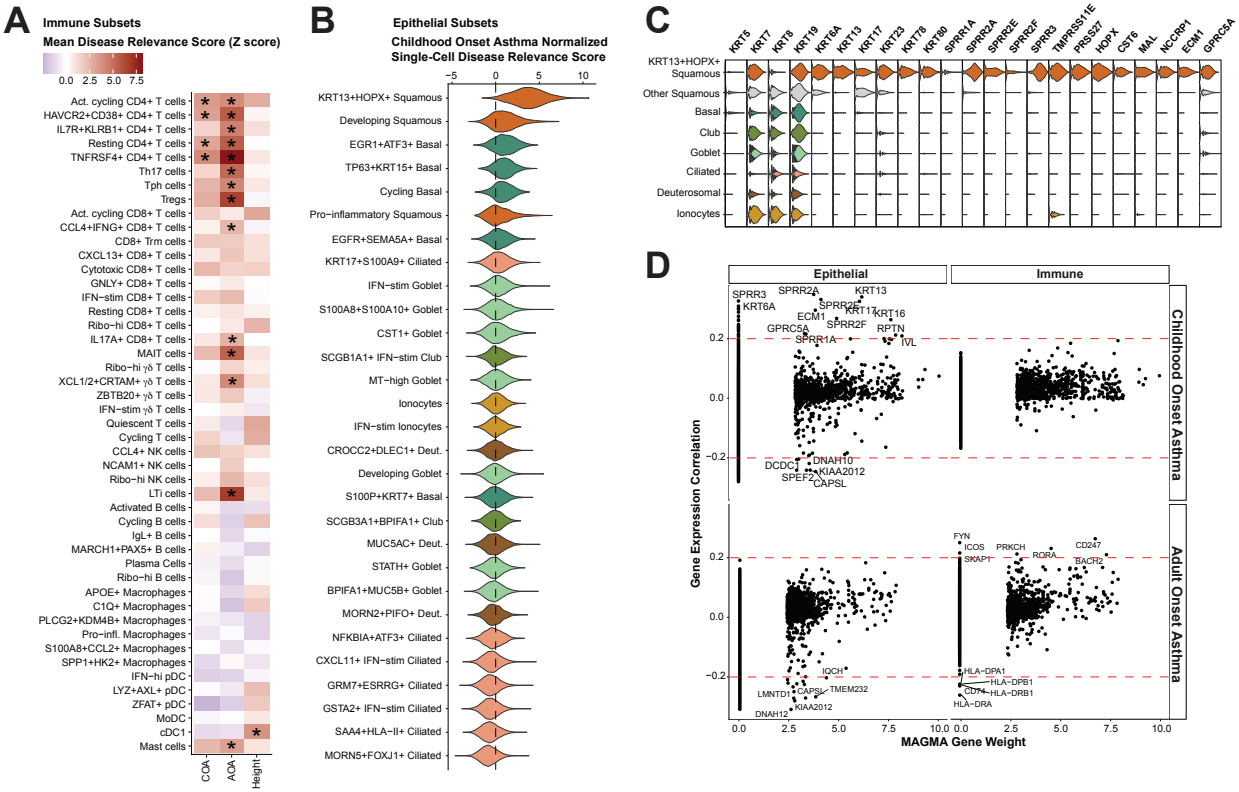
Supplementary Figure 8, Related to Figure 6



2394 **Supplementary Figure 8, Related to Figure 6**

2395 **A.** Average number of viral UMI in RSV+ cells for each RSV participant, compared across
2396 admission groups. Disc. = discharged; Admit. Floor = admitted to floor, Admit. ICU =
2397 admitted to intensive care unit.
2398 **B.** Frequency of RSV+ cells out of all epithelial cells per sample, compared across admission
2399 groups.
2400 **C.** Principal component analysis (PCA) of cellular composition for each sample in rhinovirus
2401 group colored by admission groups. PCA was run on center-log-ratio (CLR) normalized
2402 cell subset abundances. See **Figure 3G** for PC loadings.
2403 **D.** Pseudobulk pathway activity score for IFNα (left) and IFNɣ (right) response genes per
2404 Rhinovirus participant in all cells, compared across admission groups. Signatures
2405 provided in **Supplementary Table 5**.
2406 **E.** Machine-learning-derived subset importance for distinguishing no wheeze vs wheeze +
2407 treatment for RSV participants by Caret's penalized generalized linear (GLMNet) model.
2408 All subsets that contributed to model importance are shown. Color of dot indicates the
2409 difference in center-log-ratio (CLR)-normalized cell subset abundance between patients
2410 with wheeze+tx compared to no wheeze. Dot size represents frequency of the cell subset,
2411 averaged across all included RSV participants. Tx = treatment. Mac. = macrophage; Cil.
2412 = ciliated; MAIT = mucosal-associated invariant T; Gob. = goblet; pDC = plasmacytoid
2413 dendritic cell; Th17 = T-helper 17; Trm = T resident memory; Deut. = deuterosomal;
2414 Squam. = squamous; cDC = conventional dendritic cell; NK = natural killer; LTi = lymphoid
2415 tissue inducer; MoDC = monocyte-derived dendritic cells.
2416 **F.** Pseudobulk pathway activity score for IFNα (left) and IFNɣ (right) response genes per
2417 participant in Rhinovirus and RSV groups in all cells, compared across wheeze groups
2418 (no wheeze/no treatment (-) or wheeze + treatment (+)). Tx = treatment.
2419 **G.** Pseudobulk pathway activity score for corticosteroid response genes per-participant score
2420 in Rhinovirus and RSV groups in epithelial (left) and immune cells (right), compared across
2421 wheeze groups (no wheeze/no treatment (-) or wheeze + treatment (+)). Signature
2422 provided in **Supplementary Table 5**.
2423 **H.** Frequency of specified subsets as a percentage of epithelial or immune cells for all
2424 participants in SARS-CoV-2, Rhinovirus, and RSV groups, compared by treatment with
2425 corticosteroids in the week prior to presentation or in the ED.
2426 **I.** Violin plot of marker genes for *CXCL13*+ CD8 T cells, T peripheral helper (Tph) cells, and
2427 cytotoxic CD8 T cells.
2428 **A,B,D,F,G,H:** Boxplots represent the median (center line), interquartile range (box), and 1.5x the
2429 interquartile range (whiskers).
2430 **A,B,D,F,G,H:** Statistical test is Kruskal-Wallis test with Benjamini-Hochberg correction for multiple
2431 comparisons. Dunn's post-hoc test, *$p<0.05$, **$p<0.001$, ***$p<0.0001$, otherwise n.s.

Supplementary Figure 9, Related to Figure 7



2433

2434 **Supplementary Figure 9, Related to Figure 7**

2435 **A.** Cell type-trait association Z-score for childhood onset asthma (COA), adult onset asthma
2436 (AOA), and height across immune subsets. *FDR adjusted $p<0.05$. NK = natural killer cell.
2437 Th17 = T helper cell type 17; Tph = T peripheral helper cell; Treg = regulatory T cell; Ribo.
2438 = ribosomal; Trm = resident memory T cell; Act. = activated; IFN-stim = interferon-
2439 stimulated; MAIT = mucosal-associated invariant T cell; LTi = Lymphoid tissue inducer
2440 cell, MoDC = monocyte-derived dendritic cell; cDC1 = conventional type 1 dendritic cell;
2441 pDC = plasmacytoid dendritic cell.

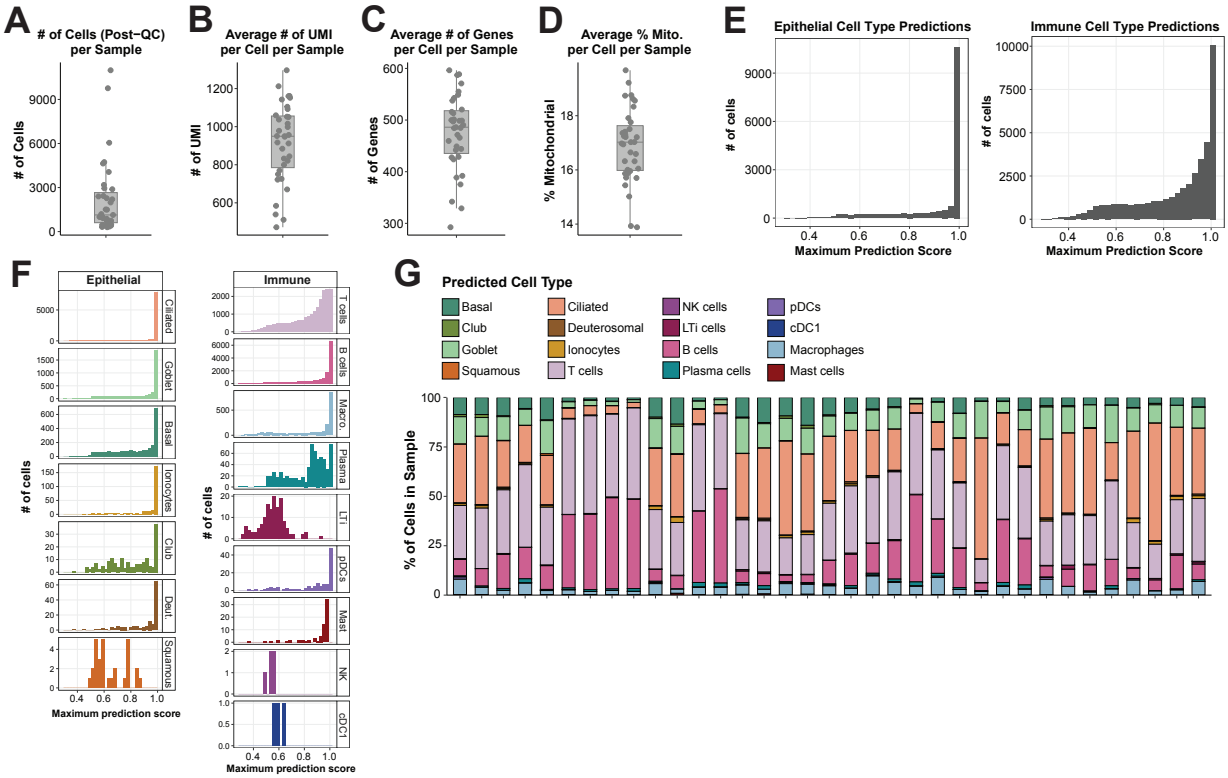
2442 **B.** Violin plot of scDRS z-scores for childhood onset asthma by epithelial cell subset. Colors
2443 represent epithelial cell types. Deut. = deuterosomal.

2444 **C.** Marker genes for *KRT13+HOPX+* squamous epithelial subset.

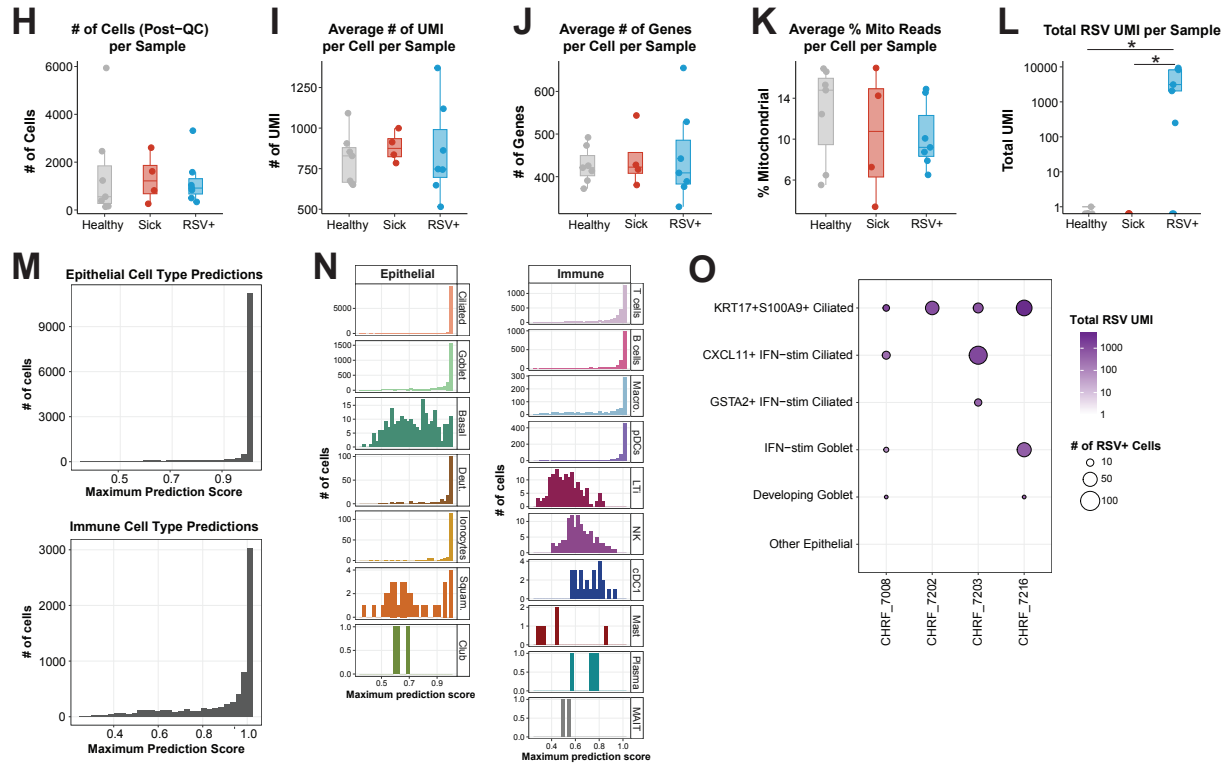
2445 **D.** Relationship between gene expression-trait correlation and MAGMA gene weights for that
2446 trait, split by epithelial (left) and immune (right), childhood onset asthma (top) and adult
2447 onset asthma (bottom). Dotted lines mark absolute correlations greater than 0.2. Genes
2448 with a 0 MAGMA gene weight had no GWAS association with the disease trait.

Supplementary Figure 10, Related to Figure 7

IND-Kolkata Cohort



BD-rural Cohort



2451 **Supplementary Figure 10, Related to Figure 7**

2452 **A-H.** Quality-control (QC) and prediction metrics for samples in IND-Kolkata cohort.

2453 **A.** Number of cells per sample after selecting high-quality cells (defined as >150 unique
2454 molecular identifiers (UMI), >100 unique genes, and <35% mitochondrial reads). Metrics
2455 in B-D are restricted to high-quality cells.

2456 **B.** Average number of UMI per cell per sample.

2457 **C.** Average number of unique genes per cell per sample.

2458 **D.** Average percent of mitochondrial reads per cell per sample.

2459 **E.** Distribution of maximum cell type prediction scores within all epithelial cells (left) and
2460 immune cells (right).

2461 **F.** Distribution of maximum cell type prediction scores divided by cell type. Cell types are
2462 colored as in **(G)**.

2463 **G.** Distribution of predicted cell types as a percentage of all cells. Each column represents
2464 one sample.

2465 **H-O.** Quality-control and prediction metrics for samples in BD-rural cohort.

2466 **H.** Number of cells per sample after selecting high-quality cells (defined as >150 unique
2467 molecular identifiers (UMI), >100 unique genes, and <35% mitochondrial reads), divided
2468 by disease group. Metrics in I-K are restricted to high-quality cells.

2469 **I.** Average number of UMI per cell per sample by disease group.

2470 **J.** Average number of genes per cell per sample by disease group.

2471 **K.** Average percent of mitochondrial reads per cell per sample by disease group.

2472 **L.** Total number of RSV transcripts (UMI) per sample by disease group.

2473 **M.** Distribution of maximum cell type prediction scores within all epithelial cells (top) and
2474 immune cells (bottom).

2475 **N.** Distribution of maximum cell type prediction scores divided by cell type. Cell types are
2476 colored as in **(G)**.

2477 **O.** Dotplot of RSV RNA+ cells in each epithelial subset, divided by sample. Samples with
2478 nonzero RSV RNA counts are shown. Dot size represents the number of cells that are
2479 viral RNA+. Color represents the total viral transcripts (UMI).

2480 **H-L:** Boxplots represent the median (center line), interquartile range (box), and 1.5x the
2481 interquartile range (whiskers). Statistical test is Kruskal-Wallis test with Benjamini-Hochberg
2482 correction for multiple comparisons. Dunn's post-hoc test, *p<0.05, **p<0.001, ***p<0.0001.

2483 **Supplementary Tables**

2484 **Supplementary Table 1, related to Figure 1:** Summarized participant metadata for USA-Boston
2485 cohort.

Variable	Control	SARS-CoV-2	Rhino	RSV
Participants	24 (20.0%)	17 (14.2%)	39 (32.5%)	40 (33.3%)
BCH ED participants	19 (79.2%)	17 (100%)	39 (100%)	40 (100%)
Outpatient clinic participants	5 (20.8%)	0	0	0
Demographics				
Age (months)				
Minimum	1.0	2.5	4.3	1.7
Median (IQR)	22.0 (14.3)	8.17 (11.94)	13.5 (17.8)	14.3 (24.9)
Maximum	57.0	39.2	45.4	48.3
Sex				
Male*	9 (37.5%)	7 (41.2%)	22 (56.4%)	$28 (70.0%)
Race				
White	10 (41.7%)	5 (29.4%)	9 (23.1%)	16 (40.0%)
Black	4 (20.0%)	2 (11.8%)	7 (17.9%)	6 (15.0%)
Asian	2 (8.3%)	1 (5.9%)	0 (0%)	4 (10.0%)
Other	5 (20.8%)	3 (17.6%)	7 (17.9%)	6 (15.0%)
Unknown	3 (12.5%)	6 (35.3%)	17 (43.6%)	9 (22.5%)
Ethnicity				
Hispanic or Latino	10 (41.7%)	4 (23.5%)	8 (20.5%)	8 (20.0%)
Medical history				
Premature#	0 (0%)	0 (0%)	1 (2.6%)	4 (10.0%)
History of wheezing#*	0 (0%)	$1 (5.9%)	16 (41.0%)	13 (32.5%)
Breastfed >1 month*	$13 (54.2%)	9 (52.9%)	$9 (23.1%)	22 (55.0%)
Vaccine/Exposure history				
SARS-CoV-2 vaccine	3 (12.5%)	2 (11.8%)	6 (15.4%)	9 (22.5%)
Prior history of SARS-CoV-2 infection	12 (50.0%)	6 (35.3%)	8 (20.5%)	13 (32.5%)
RSV antibody**	0	1 (5.9%)	0	$1 (2.5%)
Maternal RSV vaccine during pregnancy	0	1 (5.9%)	0	2 (5%)
Family history				
Parental atopy#	12 (50.0%)	10 (58.8%)	17 (43.6%)	25 (62.5%)
Symptoms/History of present illness				
Duration of illness				
0-1 Days**		$8 (47.1%)	10 (25.6%)	$2 (5%)
2-3 Days*		7 (41.2%)	16 (41.0%)	$7 (17.5%)
More than 4 days***	N/A	$2 (11.8%)	$13 (33.3%)	$31 (77.5%)
Medication in past week				
Albuterol*	N/A	$1 (5.9%)	14 (35.9%)	10 (25.0%)
Inhaled steroids	N/A	0	3 (7.7%)	3 (7.5%)
Severity of illness				
Wheezing on exam***	N/A	$2 (11.8%)	$26 (66.7%)	22 (55%)
Modified BROSJOD score#				
Minimum	0	1	0	1
Medium (IQR)	1 (2)	3 (2)	3 (3.5)	^4 (3)
Maximum	4	5	8	8
Treatment with systemic steroids in ED*	N/A	4 (23.5%)	$18 (46.2%)	12 (30%)
Treatment with albuterol in ED***	N/A	$0 (0%)	$23 (59%)	15 (37.5%)
Respiratory support				
Simple oxygen*	N/A	$0	9 (23.1%)	$16 (40%)
Non-invasive ventilation	N/A	0	0	3 (7.5%)
Mechanical ventilation	N/A	0	0	1 (2.5%)
Admitted to hospital				
Floor*	N/A	$0	7 (17.9%)	$14 (35%)
Intermediate/intensive care	N/A	0	4 (10.2%)	7 (17.5%)

P-values for continuous variables were generated using Kruskal Wallis rank test. ^p<0.05 Dunn post-hoc test relative to SARSCoV2 group, otherwise non-significant. P-values for categorical variables were generated by chi-square test. *p < 0.05, ** p < 0.001, *** p < 0.0001, otherwise non-significant. $chi-square standardized Pearson residual >2. IQR: Inter-quartile range. #Definitions: Premature = Participants for whom gestational age of 36 completed weeks or fewer was recorded in the physician's note. Information was not recorded for the majority of participants; History of wheezing = History of wheezing requiring a nebulizer or inhaler at any point in the past (parent-reported); Parental atopy = history of asthma, eczema, or seasonal allergies in either parent; Maximum severity score on modified BROSJOD scale is 10.

2486

Supplementary Table 2, related to Figures 1, 7 and Supplementary Figure 10: Full participant metadata for USA-Boston, IND-Kolkata, and BD-rural cohorts.

Supplementary Table 3, related to Supplementary Figures 2-3: Marker genes for final cell subsets defined within each cell type.

Supplementary Table 4, related to Figures 2F, 3B, and 6E: Differentially expressed genes displayed on volcano plots.

Supplementary Table 5, related to Figures 3E, 4D, 4E, 6F and Supplementary Figures 8F, 8G: Gene signatures used to calculate cytokine response scores.

Supplementary Table 6, related to Figure 5 and Supplementary Figure 6: Results from receptor-ligand analysis.