Uncovering strain- and age- dependent differences in innate immune response to SARS-CoV-2 infection in nasal epithelia using combined short and long-read scRNA-seq

Jessie J.-Y. Chang[1], Samantha Grimley[1], Bang Tran[2], Georgia Deliyannis[1], Carolin Tumpach[1], Timothy P. Stinear[1], Miranda E. Pitt[1], Elizabeth Vincan[2,3] and Lachlan J. M. Coin[1,4]

1 Department of Microbiology and Immunology, University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Melbourne, Australia

2 Department of Infectious Diseases, University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Melbourne, Australia

3 Victorian Infectious Diseases Reference Laboratory (VIDRL), Department of Infectious Diseases, University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Melbourne, Australia

4 Department of Clinical Pathology, University of Melbourne

Abstract

Assessing the impact of SARS-CoV-2 variants on host immune systems is crucial with the continuous arrival of novel variants. However, there is a lack of reported studies utilizing single-cell RNA-sequencing (scRNA-seq) to elucidate the age-dependent host-viral interactions with different SARS-CoV-2 strains. Therefore, we employed Full-Length Transcriptome Sequencing (FLT-Seq) combining Illumina and Oxford Nanopore Technologies (ONT) with 10X 3' single cell sequencing to investigate host transcriptomic changes during wild-type (WT) and Alpha-strain SARS-CoV-2 infections within air-liquid-interface (ALI)-cultured human nasal epithelial cells from adults and adolescents. The results revealed increased innate immune responses in cells with lower viral loads which were lacking in cells with higher viral load. Alpha-infections showed heightened expression of genes related to interferon (IFN)-responses and protein-folding compared with WT, implying the increased immune response and endoplasmic reticulum stress with the variant of a higher transmission rate. *ACE2* expression was elevated in infected populations of secretory and goblet cells with high IFN-stimulation and a subpopulation of ciliated cells but was lacking in bystander cells and other infected-cell types, despite detectable IFN-stimulation and regardless of strain. We used long-read sequencing to identify differential transcript usage (DTU) of IFN-response and translational genes, including *IFI27*, *RPL4* and *RPL15* (padj < 0.05) in infected cells compared with control, and consistent DTU of *RPL15* and *MX1* between Alpha and WT strain-infected cells in various

cell-types. Together, these results reveal the dynamic transcriptional/translational/post-translational activity upon SARS-CoV-2 infection, which appeared to be age and strain-dependent. Overall, this study highlights the complexity of cell-type, age and viral-strain-dependent host epithelial responses to SARS-CoV-2.

Introduction

The single-stranded RNA virus, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is prone to mutations, contributing to the continuous emergence of novel variants. Within the past few years of the Coronavirus Disease 2019 (COVID-19) pandemic, waves of cases attributed to variants such as Alpha, Delta and Omicron have led to multiple phases of government restrictions such as lockdowns and travel restrictions. Although the dominant circulating variant will change over time, it is critical to gather information on earlier prevalent variants for addressing emerging variants (Consolazio et al., 2022; Esper et al., 2022).

The first of these variants to become a Variant of Concern (VOC) was the Alpha variant (B.1.1.7), initially detected in the United Kingdom from a sample taken in September 2020 (Public Health England, 2020). This variant has 23 mutations, including 14 non-synonymous mutations, 3 deletions, and 6 synonymous mutations (Mohammadi, Shayestehpour, & Mirzaei, 2021). The N501Y mutation, located within the Receptor Binding Motif (RBM) of the Receptor Binding Domain (RBD) of the Spike (S) protein has been shown to dramatically increase binding affinity to the host cell receptor Angiotensin-Converting Enzyme 2 (*ACE2*) in humans as well as other species such as mice (Niu et al., 2021), increasing the transmissibility and host range of the virus (Y. Liu et al., 2022). Despite this increased transmission rate, there is no clear consensus regarding increased disease severity with B.1.1.7 infections compared with wild-type (WT) SARS-CoV-2 infections (Challen et al., 2021; Frampton et al., 2021; Giles, Meredith, Robson, Smith, & Chauhan, 2021; Graham et al., 2021). More importantly, whether the Alpha variant causes different host responses within adults and children/adolescents is also still unclear (Brookman et al., 2021; Loenenbach et al., 2021; Meyer et al., 2021).

There have been a number of studies applying single-cell RNA-sequencing (scRNA-seq) studies to examine host responses to SARS-CoV-2 (Ahn et al., 2021; Gao et al., 2021; Loske et al., 2021; Ravindra et al., 2021; Zhang, Cui, Cao, Wang, & Zou, 2022). This method identifies cell type-specific responses to SARS-CoV-2 infection, providing an enhanced insight compared with bulk RNA-sequencing (RNA-seq) studies. Using

scRNA-seq, host responses between children and adults have also been compared due to the generally improved clinical outcomes in children vs adults (Loske et al., 2021; Yoshida et al., 2021). However, these studies have not investigated the effect of different viral strains. Moreover, a comprehensive analysis utilizing long-read technology is currently lacking. To this end, we sought to investigate the varied age-dependent effects of two different SARS-CoV-2 strains – the original WT and Alpha isolates, during infection of primary nasal epithelial cells derived from adolescents and adults at the single-cell level using both short and long-read RNA-seq. Here, we report the use of the previously described Full-Length Transcriptome sequencing (FLT-seq) protocol (Tian et al., 2021), to visualize host transcriptomic activity against SARS-CoV-2 by harnessing the combined power of Illumina and Oxford Nanopore Technologies (ONT) sequencing and 10X Genomics 3' technology.

Methods

Laboratory methods

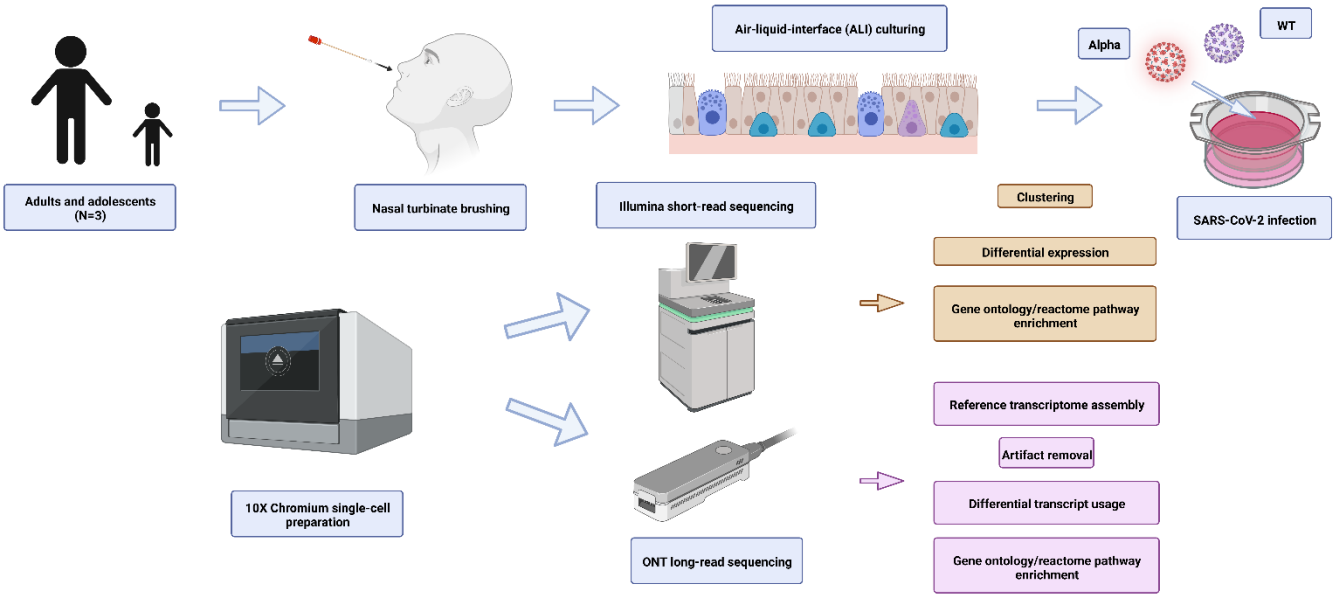


Figure 1. Overview of methods. Nasal turbinate brushings from three healthy adult and adolescent donors were acquired. The basal cells were cultured under ALI-culturing to form organoids via cell-differentiation. The organoids were infected with Alpha and WT strains of SARS-CoV-2. The RNA was harvested at 72 hpi for Alpha and 48 and 72 hpi for WT, with the inclusion of a mock-infected control sample. The FLT-seq method was carried out with a modified protocol of the Chromium Single Cell 3′ Reagent Kit v3.1, where 100% of cells were sequenced with Illumina sequencing, and ~15% of cells were also sequenced with ONT sequencing (SQK-LSK109). With short-read sequencing datasets, unsupervised clustering and pseudo-bulk differential expression (DE) and enrichment analyses were carried out. Long-read datasets were used to generate a

reference-guided *de novo* reference transcriptome, and estimated artifacts were removed. Subsequently, pseudo-bulk differential transcript usage (DTU) and enrichment analyses were performed with the count matrices generated at the isoform level. Illustration created using BioRender.com.

Cell culture

Human ethic permission was received from the Sydney Children's Hospital Network Ethics Review Board (HREC/16/SCHN/120) and the Medicine and Dentistry Human Ethics Sub-Committee, University of Melbourne (HREC/2057111) (Tran, Grimley, et al., 2022). Written consent was obtained from all participants (or participant's guardian) prior to collection of biospecimens. All samples were de-identified before tissue processing. Human nasal epithelial cell culture methods have been described previously (Awatade et al., 2021; Ferreira et al., 2021; X. Liu et al., 2012; Suprynowicz et al., 2012; Tran, Grimley, et al., 2022). Briefly, three healthy adults (Male (26Y), Female (32Y), and Male (32Y)) and children/adolescents (Female (12Y), Female (13Y), and Male (14Y)) were recruited for this study. Nasal turbinate brush samples were taken before the COVID-19 pandemic, ensuring no subject encounter with the SARS-CoV-2 virus before *in vitro* infection. To initiate differentiation at the air–liquid interface (ALI), cryo-preserved cells were thawed and seeded on to 6.5 mm Transwell inserts (Corning) pre-coated with PureCol-S collagen type I (Advanced BioMatrix). The cells were incubated at 37◦C and 5% v/v CO_2 until confluency in PneumaCult^{TM}-ExPlus media (STEMCELL Technologies) for 4-7 days before being switched to ALI culture by removing the apical media and feeding the basal side with PneumaCult^{TM} ALI medium (STEMCELL Technologies). Mucocilliary differentiation process would take 3-4 weeks and hallmarked by the presence of mucus and beating cilia.

SARS-CoV-2 propagation and ALI culture infections

Two strains of SARS-CoV-2 were utilized for this study – hCoV-19/Australia/VIC01/2020 (WT) and hCoV-19/Australia/VIC17991/2020 (Alpha). SARS-CoV-2 propagation and infection methods have been described previously (Tran, Grimley, et al., 2022). Briefly, propagation of the virus was carried out in Vero (African green monkey kidney epithelial – ATCC: CCL-81) cells cultured in MEM (MP Biomedicals), supplemented with 1 µg/mL TPCK-Trypsin (Trypsin-Worthington), penicillin (100 IU/mL), HEPES, Glutamax (Gibco), and streptomycin (100 IU/mL) under 37 ◦C and 5% v/v CO_2 incubation. Supernatant was harvested at 72 hpi, clarified via low-speed centrifugation before being filtered using a 0.45 µm syringe

filter, aliquoted and stored at -80 ◦C until use. Infectious titers were calculated by titration in Vero cells and the $TCID_{50}$/mL was calculated using the Reed and Muench formula (Reed & Muench, 1938). All viral *in vitro* infections were performed in a BSCII in the BSL-3 laboratories located at the Peter Doherty Institute.

ALI culture infections were carried out with a MOI of 0.014 in 30 µL of inoculum per insert (assuming ~300,000 cell at the surface) (Awatade et al., 2021). After virus adsorption for 2 h at 37 ◦C, the inoculum was washed off with PBS containing calcium and magnesium (PBS++). At each timepoint (0, 24, 48, 72 hours post infection (hpi)), 200 µL of PBS++ was added to the apical surface and harvested after 10 min at 37 ◦C before being stored at -80°C.

Immunofluorescence and Confocal Microscopy

Immunofluorescence and confocal microscopy imaging was performed as previously described (Tran, Grimley, et al., 2022). In brief, at the time of harvest, the cells were washed three times with PBS++. Cells were then fixed with 4% paraformaldehyde (#15710, Electron Microscopy Sciences, USA) for half an hour at room temperature. The fixative was removed and replaced with 100 mM glycine in PBS++ for 10 minutes to neutralize the remaining fixative. Cells were permeabilized with 0.5% Triton-X in PBS++ for half an hour on ice, before being washed 3 times with PBS++ at room temperature. At this stage, the membranes were carefully cut off from the Transwell inserts, excised into half, one for test antibodies and the other for control antibodies, and blocked for 90 minutes at room temperature in immunofluorescence (IF) buffer (PBS++ with 0.1% bovine serum albumin, 0.2% Triton, 0.05% Tween 20) supplemented with 10% goat serum. After this, the old block buffer was replaced by block buffer containing the primary antibodies which include antibody for acetylated α-tubulin (Sigma-Aldrich #T7451, diluted at 1:250) and antibody for SARS Nucleocapsid Protein (Novus Biologicals #NB100-56683, diluted at 1:200). After incubation for 48 hours at 4ºC, the primary antibodies were washed off with IF buffer for 3 times, fluorophore conjugated secondary antibodies, including goat-anti-mouse Alexa Fluor 488 (Invitrogen #A11001) and goat-anti-rabbit Alexa Fluor 647 (Invitrogen #21244) and Hoechst, were added and incubated for 3 hours at room temperature in the dark. Secondary antibodies were then washed off 5 times with IF buffer. The membranes were incubated with DAPI for half an hour, washed once with PBS++ and transferred to slides where they were mounted in FluoroSave Reagent (#345789 EMD Millipore). The confocal microscopy imaging was acquired on the Zeiss LSM 780 system. The acquired images were processed using *ImageJ* software.

10X Genomics single-cell preparation

The ALI-cultured human nasal epithelial cells were prepared for the 10X Chromium step according to the Single Cell Protocols Cell Preparation Guide General Sample Preparation RevC (10X Genomics). Briefly, cells were dissociated using trypsin and filtered through a 40 µm strainer and pipette-mixed to ensure a

single-cell suspension. The cells were washed with PBS with 0.04% BSA. Once cells were counted, they were harvested for mock-infected control, 48 hpi (WT infection) and 72 hpi (WT and Alpha variant infection) conditions for each donor (per age-group) and used as input for the 10X Chromium preparation. The Chromium Single Cell 3′ Reagent Kit v3.1 (10X Genomics) was utilized in conjunction with Dual Index kit TT Set A barcodes (10X Genomics) for multiplexing.

FLT-seq was carried out by modifying the Chromium Next GEM Single Cell 3′ Reagent Kits v3.1 (Dual Index) RevB protocol according to methods previously described (Jabbari & Tian, 2019; Tian et al., 2021). The modifications and the specific protocol changes used for this study are outlined in a protocols.io document https://www.protocols.io/edit/massively-parallel-long-read-sequencing-of-single-bsgunbww (to be published). Briefly, FLT-seq allows a subset of GEM-RTs (10-20%) to be sequenced by both short and long-read sequencing, and the remainder of the GEMs (80-90%) to be sequenced solely with short-read sequencing. This enables 100% of the GEMs to be sequenced via short-read sequencing for cell-type determination and differential gene expression and allows supervised cell barcode assignment of long-reads using the short-read data. In this study, ~15% of the cells were isolated following GEM-RT incubation (step 1.5 of the 10X protocol) and were subsequently treated as separate libraries throughout the remainder of the protocol. After amplification and clean-up, an aliquot of the resulting cDNA was reserved for ONT sequencing, while the rest of the cDNA was processed as per usual using the 10X short-read protocol. Therefore, 16 x Illumina libraries were prepared, consisting of 4 x adult main (~85% cells), 4 x adolescent main (~85% cells), 4 x adult subsample (~15% cells) and 4 x adolescent subsample (~15% cells) libraries. The PCR optimization and pooling steps for long-read sequencing are outlined in a subsequent section (**PCR optimization and pooling for ONT sequencing**).

Illumina sequencing

Each Illumina library was quantified with Qubit 4.0 Fluorometer via the Qubit 1X dsDNA HS Assay Kit (Invitrogen), and the fragment sizes were tested with Tapestation 4200 (Agilent Technologies) using the High Sensitivity D5000 ScreenTape (Agilent Technologies). All libraries were pooled according to respective molarities. The pooled libraries were split equally and sequenced on three NovaSeq S4 2x150bp lanes using the NovaSeq kit v1.5 (Illumina) with 0.5% of PhiX via the NovaSeq 6000 Sequencing System. The cycling parameters were as follows: Read 1 – 150 bp, Index 1- 10 bp, Index 2 – 10 bp, Read 2 – 150 bp. The sequencing was carried out by Ramaciotti Centre for Genomics at the University of New South Wales (UNSW). A total of 7.95 billion reads were acquired.

ONT sequencing

PCR optimization and pooling for ONT sequencing

To carry out ONT sequencing with the ~15% of the total number of cells, the aliquot of cDNA (described in section **10X Genomics single-cell preparation**) was amplified using PCR. The PCR was optimized by following the **PCR optimization and pooling for ONT sequencing** section of a previously described protocol (Jabbari & Tian, 2019; Tian et al., 2021), and outlined in the modified protocols.io document. Briefly, the PCR cycles for bulking for ONT sequencing were optimized to test the lowest number of PCR cycles to retain the peak distributions in fragment sizes as in the original cDNA. For this, 1 ng of the cDNA was amplified per reaction using PrimeSTAR GXL DNA Polymerase (Takara Bio) using multiple PCR cycles (**P= 8, 10, 12 and 14**). The general cycling conditions were: 1 cycle of 98°C (30 sec), **P** cycles of 98°C (10 sec), 65°C (15 sec), 68°C (8 min) and 1 cycle of 68°C (5 min). The primers used were identical to the primers used in the study by Tian et al. (2021) (F:5'-ACTAAAGGCCATTACGGCCTACACGACGCTCTTCCGATCT-3'; R:5'-TTACAGGCCGTAATGGCCAAGCAGTGGTATCAACGCAGAGTA-3'). The yield was quantified using Qubit 4.0 Fluorometer via the Qubit 1X dsDNA HS Assay Kit and peak distribution was assessed using BioAnalyzer 2100 (Agilent Technologies) via the High Sensitivity DNA Kit (Agilent Technologies). After the optimization, 5-6 x new reactions were carried out. 12 and 10 cycles were utilized for adults and adolescents, respectively. The amplicons were pooled and cleaned with 0.6X Ampure XP beads (Beckman Coulter).

ONT sequencing

200 fmoles of the amplicons were prepared into ONT libraries using SQK-LSK109 kits (ONT) via the 'Genomic DNA by Ligation for PromethION' protocol with minor modifications. The modifications included: adjusting volume of 200 fmols of DNA up to 47 µL instead of 49 µL, excluding DCS, addition of 1.8X Ampure beads for all bead cleans, using Short Fragment Buffer (SFB) for final wash, incubation steps in DNA repair and end-prep were extended to 20°C for 30 mins and 65°C for 30 mins, respectively, and the incubation step in the adapter ligation and clean-up step was extended to 20 mins at RT. 50 fmoles of the resulting libraries were sequenced using the PromethION sequencer (ONT) on R9.4.1 FLO-PRO002 flow cells for 72 hours with refueling by The Walter and Eliza Hall Institute (WEHI), Australia. A total of ~491.46 M reads were acquired.

Genotyping

Genotyping for each of the donors was required to accurately demultiplex the mixed population of cells used as inputs into the 10X Chromium preparation. DNA extraction for genotyping was carried out with

the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's guidelines (Purification of Total DNA from Animal Blood or Cells (Spin-Column Protocol)) with minor modifications. Briefly, ALI cell culture membranes were cut and placed into tubes and PBS and proteinase K was directly added to the membrane in the tube. Once Buffer AL was added, the sample was vortexed thoroughly and incubated at 56°C for 10 minutes with a Thermomixer C (Eppendorf) at 1000 rpm. 100% ethanol was added to the reaction and mixed thoroughly via vortexing. The reaction was loaded on to a spin-column and all subsequent spins were carried out at 12,000 rpm except for the step 6 of the protocol, where after adding AW2 buffer, the columns were spun at 14,000 rpm. 200 µL of buffer AE was used to elute the DNA and passed through the column in total of three times to concentrate the sample. Quality control of DNA was carried out using Qubit 4.0 Fluorometer via the Qubit 1X dsDNA HS Assay Kit, BioAnalyzer 2100 using the High Sensitivity DNA Assay and NanoDrop 2100 Spectrophotometer (ThermoFisher Scientific). Genotyping was carried out using DNA derived from each individual donor using the Infinium Global Screening Array (GSA) v2.0 BeadChip (Illumina) and performed by Macrogen (Korea). The reference annotation used was GRCh37.

Data analysis

Genotyping analysis

PLINK v1.9 (Purcell et al., 2007) was utilized to convert the output of the genotyping data (from section **Genotyping**) to VCF files. Firstly, the sex of the samples was checked, and this information was incorporated into the data. Heterozygous haploid hardcalls, all female chrY calls were erased from the data. The resulting file was converted to VCF file with '–recode'. Variants with one or more multi-character allele codes and single-character allele codes outside of {'A', 'C', 'G', 'T', 'a', 'c', 'g', 't', <missing code>} were removed from the data. To match chromosome names with downstream processes, 'chr' was added to the chromosome names, all rows with 'chr0' was removed, then chromosomes were arranged lexicographically. The final file was gzipped by *bgzip* and indexed via *tabix*.

Illumina analysis

BCL files from Illumina sequencing were converted to FASTQ files using *Cellranger* v6.1.1 (10X Genomics) with the 'mkfastq' function. Count files were produced with *Cellranger* 'count' using the reference package 'refdata-gex-GRCh38-2020-A'. BAM files generated from GRCh38/hg38 were lifted with liftover tool *CrossMap* v0.5.4 (Zhao et al., 2014) to GRCh37/hg19 to enable incorporation of the donor genotype information which was analyzed using GRCh37. Similarly, a *Cellranger* reference was created for SARS-CoV-2 reference genome using *Cellranger* 'mkref' using the Ensembl reference ASM985889v3, INSDC Assembly GCA_009858895.3, Jan 2020 and a custom GTF file by setting the whole genome as an

exon. Viral counts were determined separately but similarly to host counts using *Cellranger* 'count'. For the viral counts, raw matrices instead of filtered matrices were utilized for downstream analysis as the effect of filtered matrices (i.e. filtering of artifactual cells) was not applicable to the viral counts. The VCF files and the sorted GRCh37 BAM files were used as inputs to *Demuxlet* v11022021 (Kang et al., 2018). This enabled the donor assignment to cell barcodes and estimated the number of doublets in the data.

Data filtering and unsupervised clustering

For downstream analysis of Illumina datasets, *Seurat* v4.0.5 was implemented. Seurat objects were created separately for both viral and host counts data. The demultiplexing information from *Demuxlet* was incorporated into the *Seurat* object using *importDemux* from *dittoSeq* v1.0.2. Firstly, the viral data was separated based on infection tier as follows; uninfected < 10, low = 10-99, medium =100-999, high = 1,000-9,999, very high ≥ 10,000. This information was incorporated into host meta data per cell-barcode. Then, the data was filtered for singlets, according to the *Demuxlet* results. Cells which had <20% mitochondrial RNA, >5% ribosomal RNA were kept for analysis. Also, cells with greater than 200 and less than 9000 detected genes were kept and genes expressed in at least 3 cells were kept for analysis. Each of the 16 libraries (i.e. 8 x main and 8 subsample) were analyzed separately. Data was normalized and scaled using *scTransform* v0.3.3 (using the original method) and differences in cell cycle were regressed out by the alternate workflow (regressing out the G2M – S phase scores). Then, these datasets were merged using *Seurat* 'merge'. Utilizing cell markers from scRNA-seq datasets and within the literature (Deprez et al., 2020; Ravindra et al., 2021; Ruiz García et al., 2019), unsupervised clustering was performed. 'FindAllMarkers' was used to detect cell markers, with default parameters (i.e. min.pct=0.1, logfc.threshold=0.25, Wilcoxon's test of ranks). Final parameters used were dims=1:20 for 'RunUMAP' and 'FindNeighbors' and resolution=0.3 for 'FindClusters' functions. Characteristics of sub-cell-types (e.g. ciliated 1-3 among ciliated cells) have been analysed by running the 'FindMarkers' function via *Seurat,* which uses the Wilcoxon's test of ranks with Bonferroni p-value adjustment. Parameters of min.pct=0.25 and min.diff.pct=0.25 were used. Subsequently, using *ShinyGO* v0.76 (Ge, Jung, & Yao, 2020), we assessed the enriched gene ontology (GO) biological pathways of significantly differentially expressed genes (padj < 0.05), via the option "Select by FDR, sort by Fold Enrichment" with the default background list.

Cell proportion change testing

The changes cell-type proportions for each treatment condition (Alpha-72hpi, WT-72hpi and WT-48hpi) were tested using *Propeller* via *Speckle* v0.99.7 (Phipson et al., 2022) via the 'propeller.ttest' function in *R* v4.2.0. Each donor was used as replicates, and the tests were carried out within the same age-group

(adult/adolescent) between the mock-infected control datasets and each infected dataset (Alpha-72hpi, WT-72hpi and WT-48hpi).

Differential gene expression analysis via pseudo-bulking

After identifying the different cell-types, DGE analysis was performed on the host via a pseudo-bulking method. Briefly, the data was separated, and the counts were aggregated by a unique combination of cell-type, treatment, age-group, infection status and donor information. All samples were filtered by a minimum of 15 cells and genes which had counts in less than 10 cells were removed from the analysis. For DGE analysis, each comparison was carried out by including samples which had at least two donor replicates on each side of the comparison. Following the results from Squair et al. (2021), the edgeR-likelihood ratio test (edger-LRT) method was carried out on the aggregated counts via *edgeR* v3.30.3. The effect of sex was added into the linear model to account for sex-effects. The p-values were adjusted via the Benjamini-Hochberg method. For comparisons between infected adolescent vs infected adult cells (by taking into count the baseline expression levels in the control cells), the limma-voom method via *limma* v3.44.3 (Ritchie et al., 2015) and *edgeR* was used. Here, the differences in genetic variability between adult and adolescent donors were regressed out blocking the 'age' (i.e. each adult donor = "adult", and child donor "child") factor variable as a random effect. The effect of sex was added as a fixed effect in the design matrix. Gene Ontology (GO) biological and reactome pathways were visualized using an in-house visualization tool – *multiGO* (**Table S1**). The parameters used were pv_thresh=0.05, enrichment pv_thresh=0.005 and logFC_thresh=1. Background lists of genes were curated from all genes tested for DGE in all groups which were displayed in each *multiGO* analysis via setting the 'Background set for DE' parameter as 'gene_list' (**Table S1**).

ONT analysis

Data from PromethION sequencing was re-basecalled using standalone GPU *Guppy* v5.0.7+2332e8d (ONT) on the University of Melbourne High Performance Computing (HPC) cluster (GPGPU). The basecalled data (FASTQ) along with filtered barcodes derived from the *Cellranger* analysis of the short-read datasets (subsamples) were input into *FLAMES* v11012023 (Tian et al., 2021). The 'match_barcode' function utilizes the *Cellranger* filtered barcode files to enable supervised cell barcode matching of the ONT reads. This process enables more accurate characterization of the barcodes and UMI's. Following this, *FLAMES* maps the reads using *minimap2* (Li, 2018) and generates a reference-guided consensus transcriptome assembly. Using the newly generated reference transcriptomes, the reads are re-mapped to the new transcriptome, allowing quantification of existing and novel transcripts. In our study, *minimap2* v2.17 was used. To further collapse the transcript models, we rounded up all exon start and end coordinates

per transcript to the nearest 10 bp and collapsed and renamed the isoforms to ensure compatibility between libraries. The transcript counts which were derived from the *FLAMES* were collapsed according to the new transcriptome. As long-read transcriptome assemblies are prone to artifacts, we analyzed the merged reference transcriptome with *SQANTI3* v 5.1 (Tardaguila et al., 2018). Statistics on isoforms in the reference transcriptome were identified using *SQANTI3* the 'QC' function. The *SQANTI3* 'filter' was run with the machine learning option with default parameters except with the POSML_prob threshold = 0.5. All known reference transcripts were manually retained in the 'isoform' group during the filtering of the count matrix derived from *FLAMES*. Due to the loss of reads during the *FLAMES* analysis via the cell-barcode assignment stage, sub-clusters of cells within one cell type were merged (e.g. Ciliated-1-3à Ciliated). The count matrix was pseudo-bulked per unique combinations of cell-type, treatment, age-group, infection-status, donor. Likewise with the DE analysis, we only utilized pseudo-bulk groups with at least 15 cells and performed analyses when at least two donor replicates were able to be included on each side of the analysis. The filtered pseudo-bulk groups were used as the input for DTU analysis via the bulk-RNA sequencing options with *DTUrtle* v1.0.2 (Tekath & Dugas, 2021). The genes with significant DTU (FDR < 0.05) at the first-stage were further analysed with GO biological and Reactome pathway analysis using the in-house pathway visualization tool *multiGO* (**Table S2**). The parameters used were pv_thresh=0.05, enrichment pv_thresh=0.005 and logFC_thresh=1, where pv_thresh is equivalent to 'gene_qvalue' and logFC_thresh is equivalent to 'number_tx' of the *DTUrtle* final output (dtu_table). Background lists of genes were manually curated from all genes tested for DTU in all groups which were displayed in each *multiGO* analysis by merging the tested list of genes in the 'FDR_table' outputs in each category (e.g. inf_vs_uninfected) and finding the unique genes.

Data availability

All code is available on Github: https://github.com/cjy-23/ALI_scRNA_seq_SC2. Protocols used for generating FLT-seq datasets are on protocols.io (will be published at later date). Raw sequencing data (FAST5 & FASTQ) will be released upon publication.

Results

Alpha variant generates higher viral titers and greater reduction of epithelial cilia in adults but not adolescents

To validate the magnitude of infection within our datasets, we determined the levels of infectious viral titer/load and visualized the effect of infection on host cells. The viral titers were measured by TCID50 of

apical washes (see **Methods**). Overall, contrasting results between adults and adolescents were observed. Adolescents showed higher titers overall in both WT and Alpha infections, compared with adults (**Figures 2a&b**). However, these titers were more comparable with Alpha-infections. In adults, lower viral titers within WT infections were observed compared with Alpha infections (**Figure 2c**). In contrast, adolescents revealed similar titers between both strains at the 72 hpi (**Figure 2d**). Interestingly, the viral titers peaked around 48 hpi, except for the child WT datasets (**Figure 2c**). The titers showed consistency between donors, except in child donor 3/donor 6, where the viral titers were generally lower compared with other donors (**Figure S1a-d**). Furthermore, confocal microscopy showed the reduction of cilia in Alpha-infected cells when compared with WT-infected cells in adults (**Figures 2e & S2a-c**), and lack of cilia loss in Alpha-infected adolescent cells (**Figures 2e & S2d-f**).

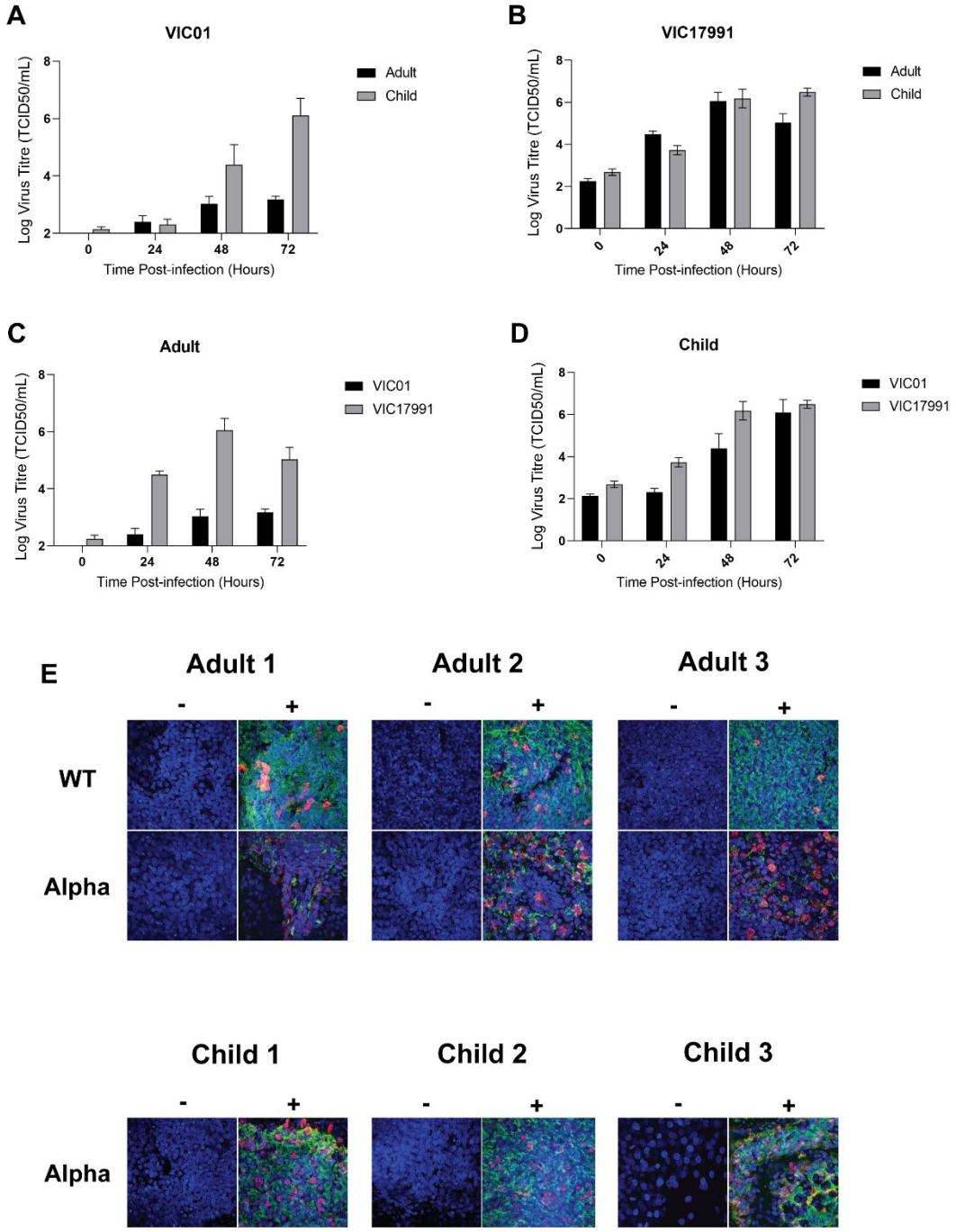


Figure 2. Viral load, titers and visualization of SARS-CoV-2 infected ALI-HNE. TCID50 results from apical washes at 0, 24, 48. 72 hpi comparing adults and adolescents with **A)** WT and **B)** Alpha infections. Comparison of WT and Alpha infections at each timepoint in **C)** adults and **D)** adolescents. **E)** Immunofluorescent confocal microscopy staining at 40X magnification for α-tubulin (AcTub, green), nucleoprotein (NP, red) and nuclei (DAPI, blue). Both WT and Alpha-infected cells are shown for adults and only Alpha-infected cells are shown with adolescents/children. Negative controls are indicated by '-' and complete stains are indicated with '+'.

A transitional cell type with secretory and ciliated properties is highly infected in the human nasal epithelia

The traditional landscape of the human nasal epithelium is mainly composed of ciliated, basal and secretory cells (Tran, Deliyannis, et al., 2022). Additionally, rarer cell types such as ionocytes (Lukassen et al., 2020; Ravindra et al., 2021), deuterosomal cells (Deprez et al., 2020; Ruiz García et al., 2019) and transitional cell types with cell signatures from more than one cell type may be present (Robinot et al., 2021). In our data, we observed ciliated, goblet, basal, suprabasal, secretory, cycling-basal, brush/tuft, deuterosomal and ionocyte cells – consistent with other scRNA-seq data of the human airway epithelium (Deprez et al., 2020; Ruiz García et al., 2019) (**Figures 3a & Table S3**). We additionally found the presence of cell types which were unable to be clearly classified into any one cell-type. These cells were assigned as transitional cell-types, which included Secretory-Ciliated and Goblet-Ionocyte cells (**Figure 3a**). Secretory-ciliated cells have been identified previously in human airway epithelial cells (Robinot et al., 2021). Additionally, we observed a small cluster of Goblet-Ionocyte cells with high expression of markers of Goblet (*MUC5AC*) and Ionocytes (*CFTR*), which have not been previously described to our knowledge.

We describe cells with fewer than 10 viral UMI counts as being uninfected, and cells with at least 10, 100, 1000, 1000 viral UMI counts as having a low, medium, high or very high level of infection respectively (**Figure 3b**). Of all the detected cell types, the largest group of cells were suprabasal cells with >20,000 cells overall, which were mostly lowly infected or uninfected. Ciliated and Secretory-Ciliated cell types had the highest proportion of medium, high or very high levels of infection. Within these groups we observe a clear separation of cells with mostly 'high'-'very-high' and have labelled these sub-clusters as "Ciliated+SC2" and "Secretory-Ciliated+SC2". These subclusters showed 50.7% and 34.8% of cells with 'very high' level of infection (**Figure 3c**). These results are consistent with the SARS-CoV-2 cellular tropism shown in the literature (Ahn et al., 2021; Ravindra et al., 2021; Robinot et al., 2021) (**Table S4**). Although the susceptibility of secretory-ciliated cells to SARS-CoV-2 has been noted in human airway epithelial cells (Robinot et al., 2021; Yoshida et al., 2021), the exceptionally high viral loads of infected secretory-ciliated cells has not been identified previously within human nasal epithelial cells, to our knowledge. We also found two subsets of goblet and secretory cells – "Goblet+IFN-stim", "Secretory+IFN-stim" exhibiting high interferon (IFN) stimulation, with elevated levels of IFN-stimulated genes (ISGs), which increased proportionally to infection level (**Figure S3 & Data S1-2**). While most infected cells were classified as lowly infected, we identified cells with high and very high levels of infection within Secretory-Ciliated (5.59%), Secretory-3+IFN-stim (4.59%), Goblet+IFN-stim (3.40%), Deuterosomal (2.34%), Ciliated-1 (1.40%), Goblet-Ionocyte (1.30%) and Ciliated-2 cells (1.06%) (**Figure 3c**). This highlights the SARS-CoV-2 susceptibility of secretory/goblet cells in addition to ciliated and transitional cell-types,

which again is consistent with the literature (N. Zhu et al., 2020). Alpha-infected datasets showed the highest proportion of infected cells, followed by WT 72 hpi and then WT 48 hpi-infected cells (**Figures S4a-e**). We note very few cells (0.33% of all WT 48 hpi data from adults and adolescents) were infected in WT-strain-infected cells harvested at 48 hpi, in comparison to 72 hpi with the WT and Alpha strains (**Figure 3c**).

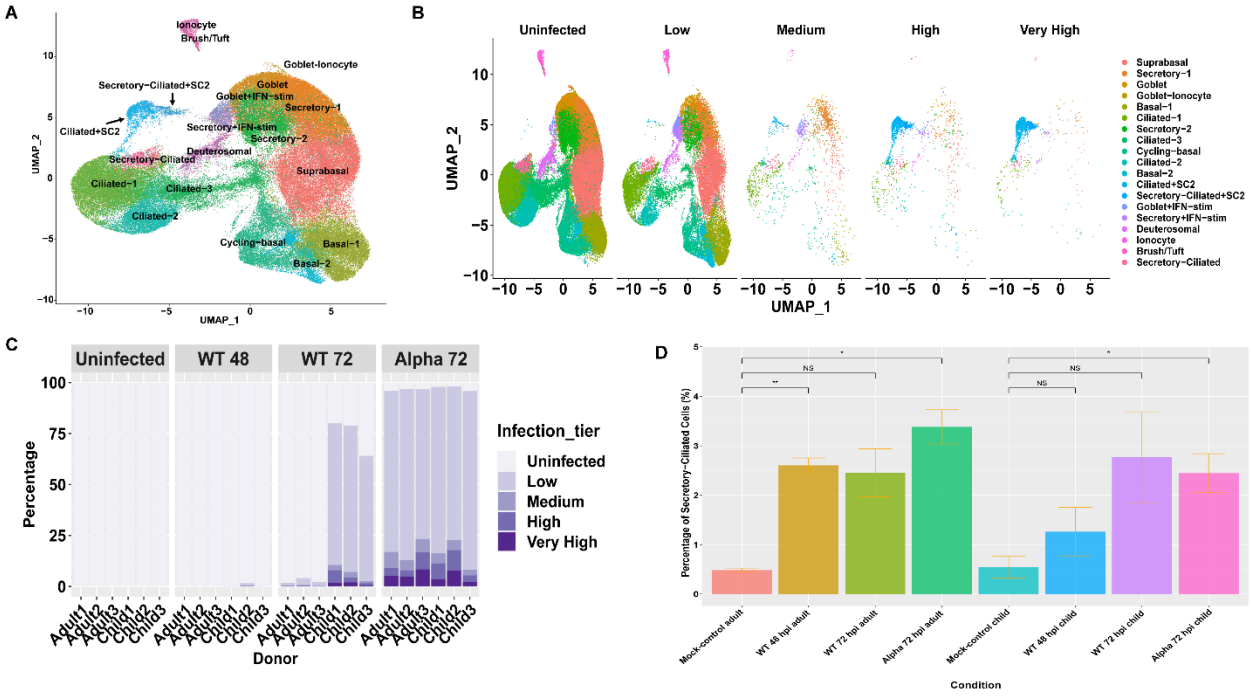


Figure 3. Infection levels per cell-type and condition in the human nasal epithelium. A) Uniform Manifold Approximation and Projection (UMAP) of cells from all samples after unsupervised clustering with known cell markers. **B)** UMAP of cells split by infection tier after filtering (see **Methods**), (uninfected (<10 viral counts per cell), low (<100 viral counts per cell), medium (<1000 viral counts per cell), high (<10,000 viral counts per cell), and very high infection levels (≥10,000 viral counts per cell)). **C)** Percentage of infected cells in adults and adolescents (n=3) in each condition per donor, stratified by infection tier (two-tailed T-test). **D)** Mean percentage of secretory-ciliated cells between each treatment condition in adults and adolescents (n=3) ± SEM, two-tailed T-test.

Additionally, the changes in cell-type distributions upon infection against the mock-control datasets were compared within the same age-groups. As expected, the increase in Ciliated+SC2, Secretory-Ciliated+SC2,

Secretory+IFN-stim, Goblet+IFN-stim cells in Alpha and WT-infected cells were observed at 72 hpi in both age-groups (**Data S2 & Figures S3a-d**, moderated T-test, FDR < 0.05). In all three adult infected datasets (Alpha 72 hpi, WT 72 hpi and WT 48 hpi), Secretory-Ciliated cells also increased in proportion. Alpha-infected adult datasets increased in Basal-1, WT-infected adults increased in Ciliated-1 and Secretory+IFN-stim cells at 72 and 48 hpi, respectively, compared with the mock-control datasets (**Data S2 & Figures S3a-d**, FDR < 0.05). No significant changes in cell-clusters within WT 48 hpi in adolescents were observed when compared with mock-controls. Interestingly, only adult cells showed significant decreases in certain cell-types, such as Secretory-2, Goblet and Brush/Tuft cells in all comparisons, with the additional decrease of Ciliated-2 cells in both WT-infected datasets. This showed the age-dependent differences in change in cell-type distributions with infection, regardless of viral strain or time since infection.

ACE2 and *TMPRSS2* transcriptional levels are low in human nasal epithelial cells

ACE2 mRNA expression levels were found to be low across all cell types in our data, consistent with previous studies (Ahn et al., 2021) (**Figure 4a-c**). Furthermore, *TMPRSS2* and *FURIN* expression was higher than *ACE2* levels across many different cell-types, with *TMPRSS2* expression being the lowest in basal and *FURIN* lacking in ciliated cell populations (**Figure 4a-c**). We noticed comparatively higher *ACE2* expression in a subset of secretory cells – Goblet+IFN-stim, and Secretory+IFN-stim (**Figures 4a**). These cells showed high-levels of IFN responses with elevated gene expression of ISGs and were only robustly present in conditions usually associated with higher levels of SARS-CoV-2 infection (Alpha-infected datasets in adults & adolescents and WT 72 hpi datasets in adolescents) (**Data S1**). The higher levels of *ACE2* in secretory/goblet cells has also been shown previously in the human nasal epithelia (Sungnak et al., 2020; Carly G. K. Ziegler et al., 2020). Interestingly, despite being associated with higher viral-load datasets, these clusters only involved low-levels of viral RNA (**Figure 3b**). We then performed DE between infected and mock-control datasets to investigate whether SARS-CoV-2 infection caused upregulation of *ACE2*. Significant upregulation of *ACE2* in only the same high IFN-stim clusters (Goblet+IFN-stim, and Secretory+IFN-stim) and Alpha-infected Ciliated-1 cells were observed, which were also largely lowly infected (**Figures 4d-k & Figure 3b, & Table S5** padj < 0.05). In contrast, *ACE2* was minimally expressed in highly-infected cell types such as Secretory-Ciliated+SC2 and Ciliated+SC2 (**Figure 4a**). Furthermore, other infected cell-clusters were not found to significantly upregulate *ACE2*.

These results led to the question whether infection or simply the exposure to IFN causes an upregulation of *ACE2* in various clusters. Bystander cells are cells which are exposed to the pathogen but have not been identified to be infected (Ravindra et al., 2021), but should be exposed to IFN through paracrine activity from neighboring infected cells. In our data, bystander cells did not show a significant increase in expression level of *ACE2* compared to mock-control (**Figures S5a-c**). Collectively, these results suggested that infection nor IFN-stimulation/ISG induction does not cause upregulation of *ACE2*, and perhaps low infection with high IFN-stimulation above a certain threshold is required for *ACE2* upregulation. Furthermore, *ACE2* expression correlated positively with levels of *STAT1* (**Figures 4d-k**). These results are in line with evidence showing that *ACE2* is stimulated by IFNs and has expression correlation with *STAT1* (Carly G. K. Ziegler et al., 2020). In the study by Carly G. K. Ziegler et al. (2020), the promoter of *ACE2* was found to contain two *STAT1* binding sites, increasing the importance of this relationship.

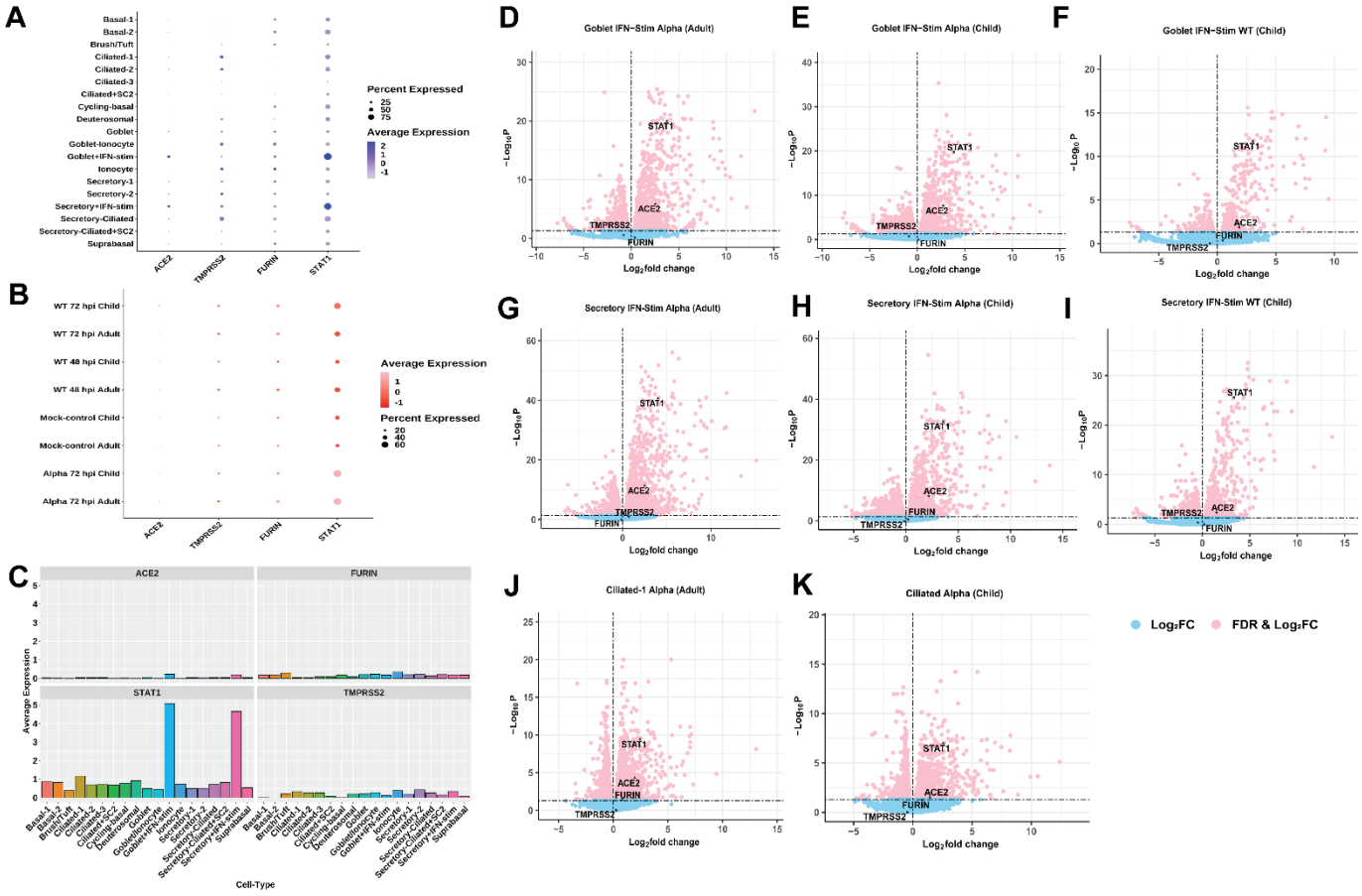
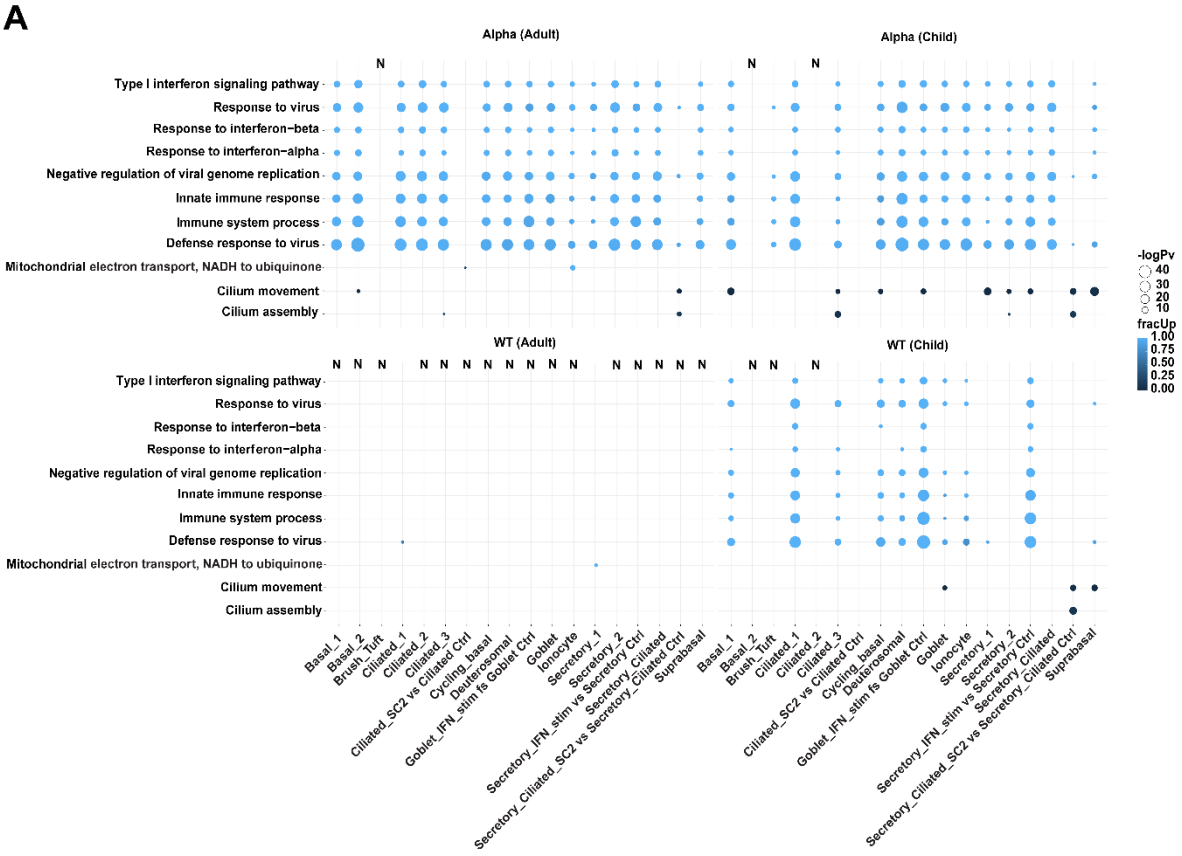


Figure 4. Expression of SARS-CoV-2 entry-related genes. (**A-B**) Relative/scaled average expression of *ACE2, TMPRSS2, FURIN* and *STAT1* **A)** overall and **B)** within different conditions. **C)** Average expression of *ACE2, TMPRSS2, FURIN* and *STAT1* in each cell-type. *ACE2* gene expression is generally low across all cell types but appear to be elevated in IFN-stimulated Goblet and Secretory cells (Goblet+IFN-stim,

Secretory+IFN-stim), providing support for *ACE2* being an ISG. (**D-K**) Volcano plots of showing the DE of *ACE2, TMPRSS2, FURIN* and *STAT1* in infected vs mock-control datasets where X-axis shows the log2 fold change and Y-axis shows the log10 padj. For ciliated Dots in blue show the genes which did not meet the logpadj threshold of padj = 0.05, and dots in pink show the genes which met the threshold. **D)** Goblet+IFN-stim Alpha vs Goblet mock-control (adult), **E)** Goblet IFN-stim Alpha vs Goblet mock-control (child) , **F)** Goblet+IFN-stim WT vs Goblet mock-control (child), **G)** Secretory+IFN-stim Alpha vs Secretory mock-control (adult), **H)** Secretory+IFN-stim Alpha vs Secretory mock-control (child), **I)** Secretory+IFN-stim WT vs Secretory mock-control (child), **J)** Ciliated-1 Alpha vs Ciliated-1 mock-control (adult) and **K)** Ciliated-1 Alpha vs Ciliated-1 mock-control (child).

Cell-types with low levels of infection show increased innate immune responses compared with cell-types with high levels of infection

We investigated whether the SARS-CoV-2-infected cells show any differences in host immune response compared with mock-control cells on a cell-type basis. Within both adults and adolescents, Alpha-infected cells showed strong enrichment of infection-specific GO terms such as *defense response to virus, response to virus, type I interferon signaling pathway, response to interferon-beta, response to interferon-alpha* and the *negative regulation of viral genome replication* except Brush/Tuft and Basal-2 and Ciliated-2 cells in adolescents (**Figure 5a**). Similarly, in WT-infected adolescents, IFN-response related terms were enriched, in similar groups of cells. The reactome pathway analysis largely agreed with the GO biological term analysis (**Figure S6a**). Comparison of expression between Secretory-Ciliated+SC2 cluster in infected samples with the Secretory-Ciliated cluster from mock-control; as well as between the Ciliated +SC2 cluster from the infected with the Ciliated cluster from mock control revealed a marked absence of significant enrichment of these GO terms, regardless of age or strain. Surprisingly, the only GO term enriched in Ciliated+SC2 cells versus Ciliated in mock control was *mitochondrial electron transport, NADH to ubiquinone*, despite many genes being significantly DE. Additionally, for Alpha-infections in both age groups and WT-infections in children, we note these highly infected Secretory-Ciliated+SC2 cell cluster showed downregulation of genes related to *cilium movement* and *cilium assembly* in comparison to mock control Secretory-Ciliated cells, consistent with loss of cilia observed from microscopy (**Figures 5a & 2e**).


A
Alpha (Adult)
Alpha (Child)
WT (Adult)
WT (Child)
Type I interferon signaling pathway
Response to virus
Response to interferon−beta
Response to interferon−alpha
Negative regulation of viral genome replication
Innate immune response
Immune system process
Defense response to virus
Mitochondrial electron transport, NADH to ubiquinone
Cilium movement
Cilium assembly
Basal_1
Basal_2
Brush_Tuft
Ciliated_1
Ciliated_2
Ciliated_3
Ciliated_SC2 vs Ciliated Ctrl
Cycling_basal
Deuterosomal
Goblet_IFN_stim fs Goblet Ctrl
Goblet
Ionocyte
Secretory_1
Secretory_2
Secretory_IFN_stim vs Secretory Ctrl
Secretory_Ciliated
Secretory_Ciliated_SC2 vs Secretory_Ciliated Ctrl
Suprabasal
-logPv
40
30
20
10
fracUp
1.00
0.75
0.50
0.25
0.00

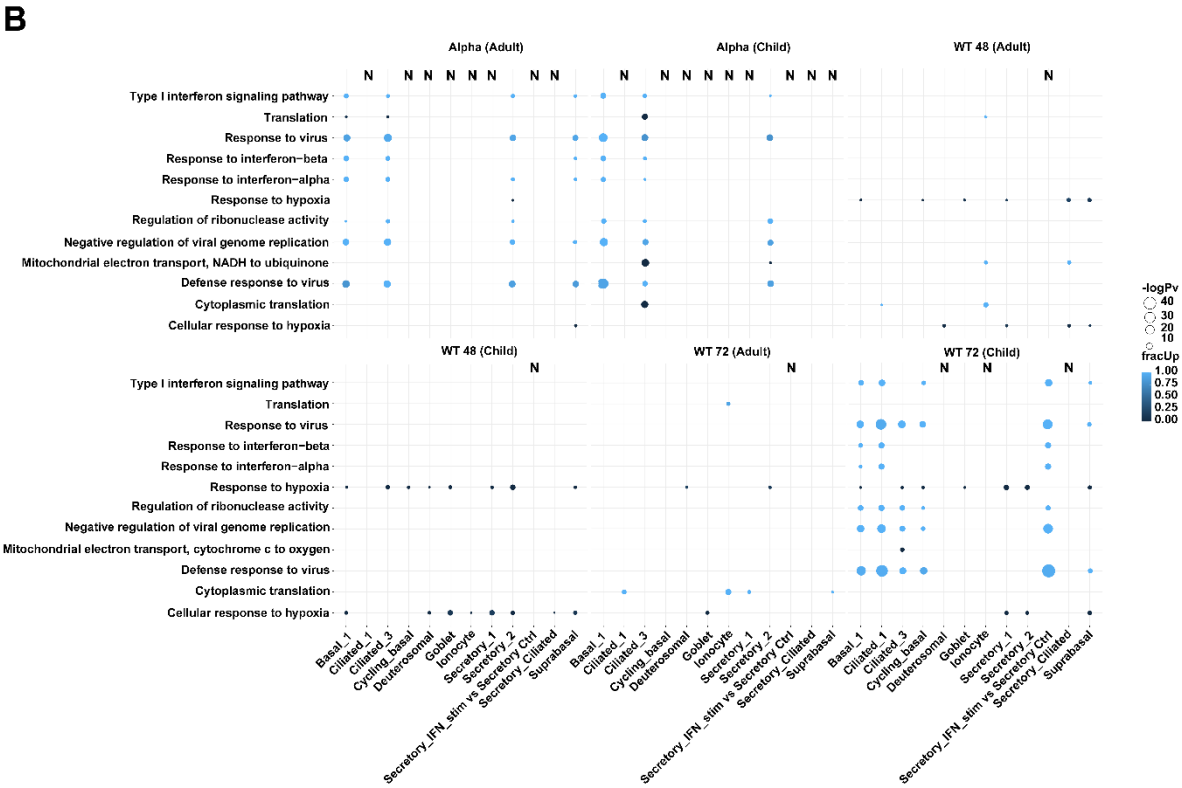

B
Alpha (Adult)
Alpha (Child)
WT 48 (Adult)
WT 48 (Child)
WT 72 (Adult)
WT 72 (Child)
Type I interferon signaling pathway
Translation
Response to virus
Response to interferon−beta
Response to interferon−alpha
Response to hypoxia
Regulation of ribonuclease activity
Negative regulation of viral genome replication
Mitochondrial electron transport, NADH to ubiquinone
Defense response to virus
Cytoplasmic translation
Cellular response to hypoxia
Mitochondrial electron transport, cytochrome c to oxygen
Basal_1
Ciliated_1
Ciliated_3
Cycling_basal
Deuterosomal
Goblet
Ionocyte
Secretory_1
Secretory_2
Secretory_IFN_stim vs Secretory Ctrl
Secretory_Ciliated
Suprabasal
-logPv
40
30
20
10
fracUp
1.00
0.75
0.50
0.25
0.00

Figure 5. Significantly enriched GO biological terms analyzed using *multiGO* using significant DGE results in infected cells and bystander cells (padj < 0.05, enrichment p-value < 0.005, |logFC| >1). **A)** Infected cells compared with mock-control, where mainly upregulated genes were involved in these processes. **B)** bystander vs mock-control cells. Bubble size indicates -log10 enrichment p-values, and the color of the bubble indicates the proportion of upregulated genes (i.e. fracUp). Columns with no matching DE data available are denoted with 'N'. High IFN-stim and high viral load populations (+SC2) have been compared with mock-control cells from other related cells.

Next, to investigate whether cells are affected by the infection of neighboring cells, we compared the gene expression of bystander cells compared with mock control cells. Alpha-infection-related bystander cells in both adolescents and adults and WT-infection-related bystander cells in adolescents showed an enrichment of GO biological terms associated with viral infection such as *defense response to virus, type I interferon signaling pathway, response to virus* and *response to interferon-alpha, response to interferon-beta*. Furthermore, enrichment of *regulation of ribonuclease activity* and *negative regulation of viral genome replication* was found (**Figure 5b**). These genes were upregulated in Alpha-associated Basal-1, Ciliated-3, Secretory-2, Suprabasal in adults and Basal-1, Ciliated-3, Secretory-2 in adolescents compared with control. Also, similar upregulation was observed in Basal-1, Ciliated-1, Ciliated-3, Cycling-basal, Secretory+IFN-stim and Suprabasal cells in WT 72 hpi-associated adolescent cells, with an absence of enrichment in *response to interferon-beta* and *response to interferon-alpha* in Ciliated-3, Cycling-basal and Suprabasal cells. Similarly, reactome pathways such as *interferon alpha/beta signaling* and *interferon signaling* were enriched in these datasets in the same cell-clusters (**Figure S6b**). This reciprocates the results observed in Alpha-strain-infected cells as discussed above. This suggests that consistent with existing evidence that bystander cells are affected by paracrine activity of cytokines which are released from infected neighboring cells (Czerkies, Kochanczyk, Korwek, Prus, & Lipniacki, 2022), the exposure but not infection of SARS-CoV-2 can still elicit an increase in anti-viral gene expression within these cells. *Translation* was also enriched but mostly composed of genes which were downregulated in bystander cells compared with control in both Alpha-associated adult and adolescent cells. Furthermore, *response to hypoxia* was enriched in WT-associated datasets in both adolescents and adults, and minimally in Alpha-associated adults in Secretory-2 cells.

Infected adolescent derived organoids have decreased oxidative phosphorylation and ribosomal gene expression levels compared with adults

Next, accounting for the baseline expression in the respective control cells, we compared the differences between infected adults and adolescents. Enriched GO biological terms included *mitochondrial electron transport, NADH to ubiquinone*, *translation, cytoplasmic translation* (**Figure 6a**). Secretory-1 and ciliated-1 cells were especially involved in significant gene set enrichment in both Alpha and WT infections (**Figure 6b**) and *cilium movement* and *cilium assembly* were only enriched in Alpha-infected datasets. Similarly, enriched reactome pathways included *respiratory electron transport, translation* and *influenza infection* (**Figure S6c**). Interestingly, aside from ciliated, secretory and secretory-ciliated cell-types, Ionocytes were also largely involved in gene set enrichment in Alpha infected cells (**Figure 6a**). Overall, adolescents showed overall downregulation of gene expression compared with adults in these enriched terms/pathways. This suggests either the lower requirement of these genes/pathways or the increased viral suppression of these pathways upon infection in adolescents and absent or diminished in adults, supporting the idea of an age-dependent response to SARS-CoV-2.

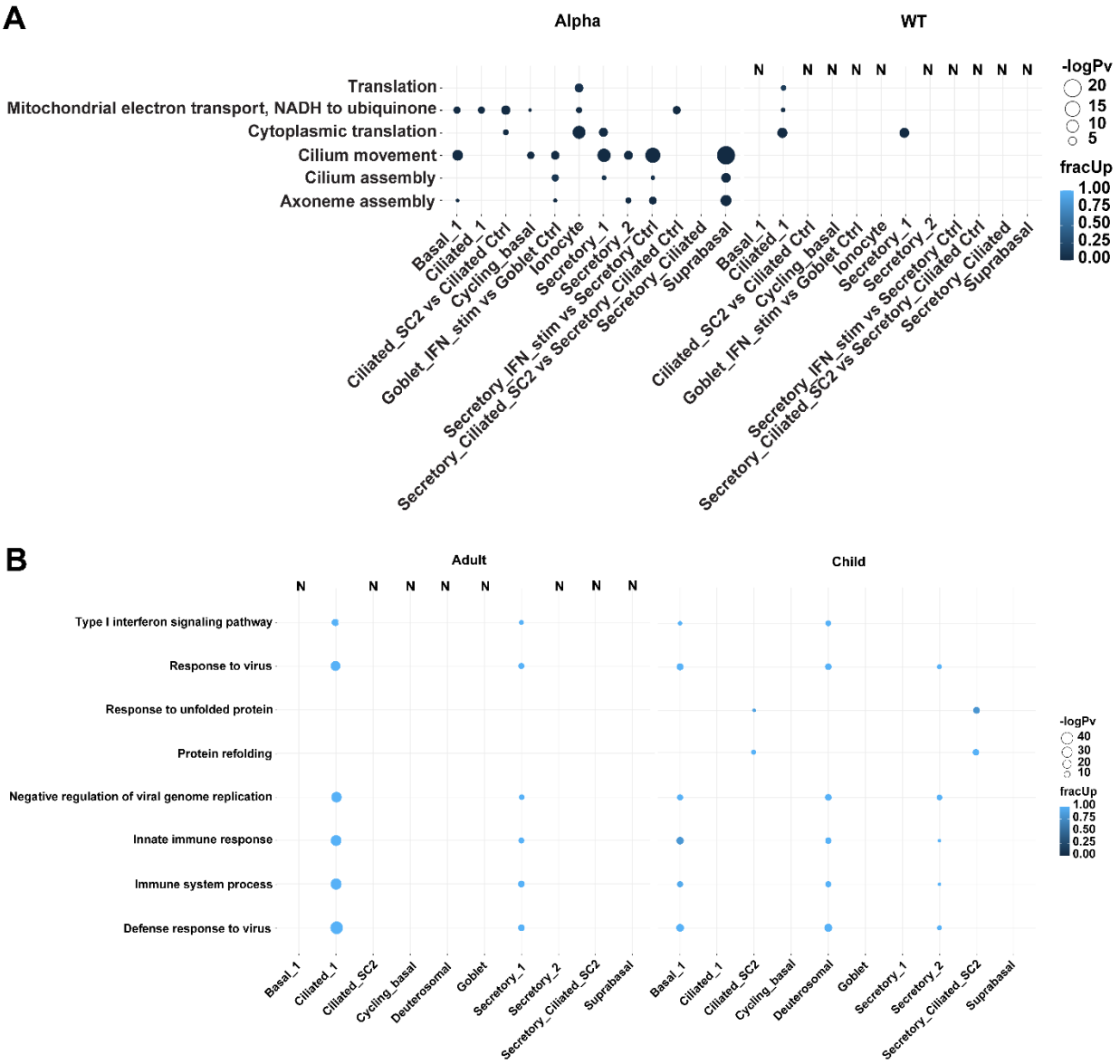


Figure 6. Significantly enriched GO biological terms analyzed using *multiGO* using significant DGE results between age-groups and viral strains (padj < 0.05, enrichment p-value < 0.005, |logFC| >1). **A)** Differences in infected adolescents vs adults, accounting for baseline expression in controls. Secretory-1 and ciliated cells with additional enrichment of translation-related GO terms. **B)** Alpha vs WT infected cells. Bubble size indicates -log10 enrichment p-values, and the color of the bubble indicates the proportion of upregulated genes (i.e. fracUp). Columns with no matching DE data available are denoted with 'N'. High IFN-stim and viral load populations (+SC2) have been compared with mock-control cells from other related cells.

Alpha variant induces increased protein folding and innate immune responses compared to WT strain

We next explored differences in host responses to the Alpha variant compared with the WT-strain of SARS-CoV-2. Anti-viral terms such *as response to virus, negative regulation of viral genome replication, innate immune response, immune system process and defense response to virus* were enriched in both adults (Ciliated-1, Secretory-1) and adolescents (Basal-1, Deuterosomal, Secretory-2) (**Figure 6d**). In terms of reactome pathways*, interferon signaling, interferon alpha/beta signaling, antiviral mechanism by IFN-stimulated genes* were enriched in the same cells (**Figure S6d**). The genes involved in these processes were upregulated in the Alpha-variant infections compared with WT. These results highlight the heightened antiviral host responses in Alpha vs WT infections. Furthermore, these results show that although similar processes are elicited in the two age-groups, there are differences within the groups of cells which are involved. Additionally, we observed similar upregulation of genes involved in *protein refolding* and *response to unfolded protein* GO biological terms in Ciliated+SC2 and Secretory-Ciliated+SC2 datasets in adolescents (**Figure 6d**). Therefore, this provides evidence that the Alpha variant elicits a greater post-translational activity related to refolding aberrantly folded/unfolded proteins in the most infected cluster of cells, at least within adolescents. We note that we did not have enough infected cells in those clusters in adults to compare with adolescents.

Translation and oxidative phosphorylation upregulation in infected cells vs bystander cells

Following the comparisons between infected and bystander cells in human bronchial epithelial cells (HBECs) by Ravindra et al. (2021), we wondered if HNECs cultured under ALI-conditions would also exhibit similar findings to HBECs. Firstly, we noted that Ciliated-3 cells had most enrichment of GO biological terms out of all datasets (**Figure 7a**). In Alpha-variant-infected Ciliated-3 cells, we observed an upregulation of genes involved in *translation, oxidative phosphorylation, NIK/NF-kappaB signaling* and *antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent* compared with bystander cells in both adults and adolescents (**Figure 7a**). Additionally, we observed *mitochondrial electron transport, NADH to ubiquinone* and *aerobic respiration* in Basal-1 cells and *mitochondrial electron transport, NADH to ubiquinone* in Secretory-2 Alpha-infected adults but not in adolescents, revealing age-dependent responses. Like the aforementioned results, the enriched GO terms overwhelmingly involved upregulated genes in infected cells compared with the bystander cells. Enriched reactome pathways included *the citric acid (TCA) cycle and respiratory electron transport, respiratory electron transport, ATP synthesis by chemiosmotic coupling and heat production by uncoupling proteins,*

infectious disease, influenza infection, interleukin-1 signaling, viral mRNA translation and *translation* in Ciliated-3 in Alpha-associated adults and adolescents (**Figure S6e**). Additionally, *the citric acid (TCA) cycle and respiratory electron transport, respiratory electron transport, ATP synthesis by chemiosmotic coupling and heat production by uncoupling proteins* were enriched in Basal-1 and Secretory-2 cells in Alpha-associated adults, Secretory-2 in Alpha-associated adolescents and Ciliated-3 cells in WT-associated adolescents. Furthermore, we then combined all cells from Ciliated-cell clusters (Ciliated1-3 & +SC2) and performed DE between infected and bystander cells. From these results, we noted the consistent upregulation of genes *NFKBIA, JUN* and *SOX4* in Alpha-infected cells when compared with bystander cells in both age-groups as well as WT-infected cells vs mock-control cells in adolescents (**Figures 7b-d**). This is consistent with results from the study by Ravindra et al. (2021). However, in contrast to the study by Ravindra et al., we did not find many downregulated genes (**Figures b-d**).

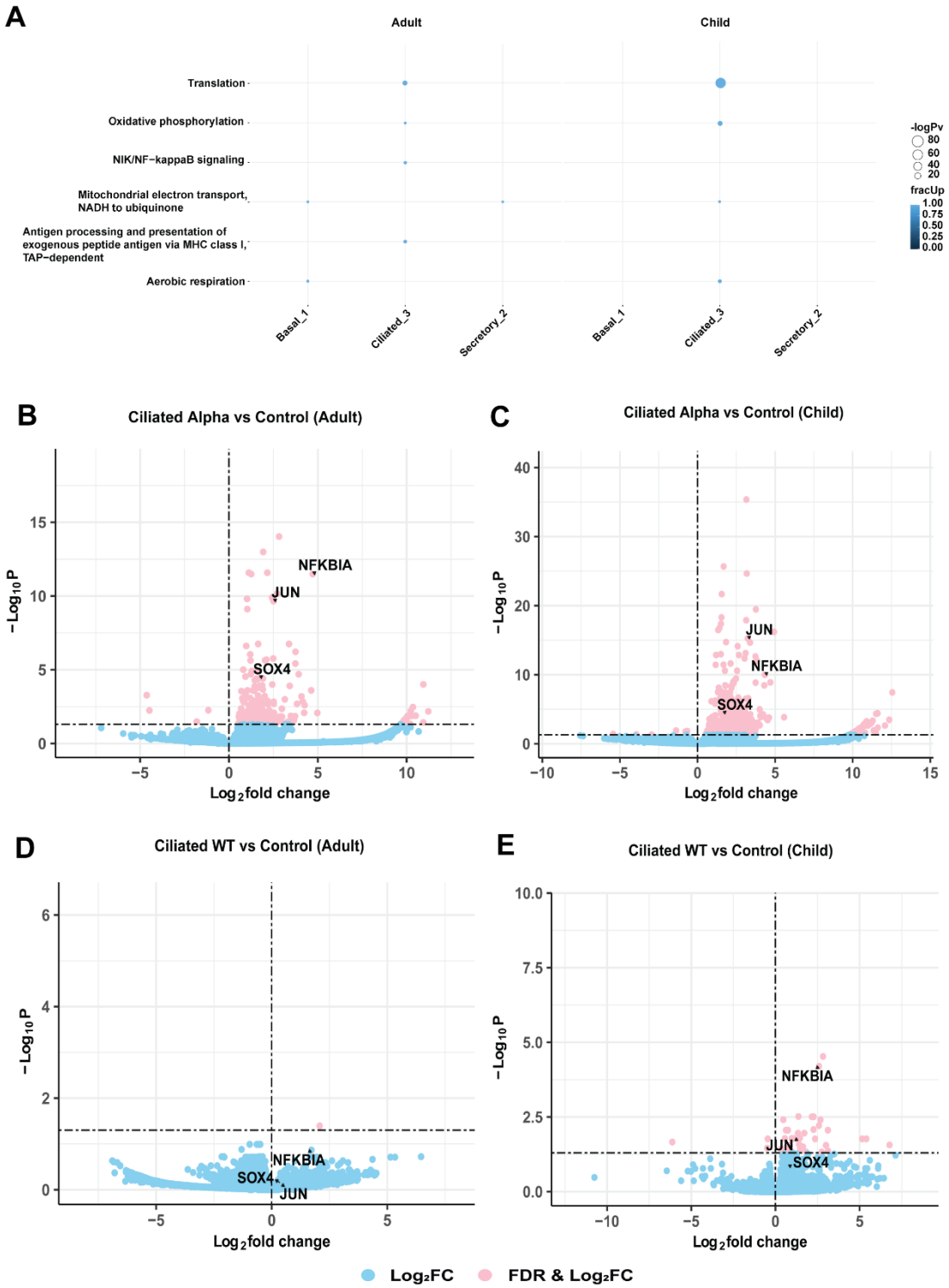


Figure 7. Differences in expression level in infected cells vs bystander cells. A) Enriched GO terms in Alpha and WT-infected cells in adults and adolescents vs bystander cells (pv_thresh=0.05, enrichment pv_thresh=0.005 and logFC_thresh=1). Bubble size indicates -log10 enrichment p-values, and the color of the bubble indicates the proportion of upregulated genes (i.e. fracUp). Volcano plots of significantly differentially expressed genes in ciliated cells, with upregulated *NFKBIA*, *JUN* and *SOX4* genes in infected cells vs bystander cells within **B)** Alpha-infected cells in adults, **C)** Alpha-infected cells in adolescents, **D)** WT-infected cells in adults, and **E)** WT-infected cells in adolescents. Thresholds of padj < 0.05 and logFC=0 were used. X-axis shows the log2FC in infected (Alpha/WT) vs bystander datasets, and Y-axis shows the -log10 padj from the DGE analysis. Blue dots indicate genes which meet only the log2FC threshold and pink dots indicate the genes which meet both the padj and log2FC thresholds.

Comparison of long and short read sequencing

We assessed whether the ONT long-read datasets re-capitulated the findings of the original FLT-seq paper. The SQK-LSK109 libraries on the PromethION yielded an average of ~61.4 M reads per library across the eight libraries we sequenced for this study, resulting in ~16,000 reads per cell. The original publication of the FLT-seq method showed the detection of 40-60% of long-reads to be assigned barcodes using a filtered barcode matrix via the Illumina short-read analysis via *Cellranger* (Tian et al., 2021). Similarly, we found that 48.7%-59.1% of the reads could be attributed to a barcode using the default parameters including an edit distance of 2 (**Figure 8a**). Furthermore, the number of detected $\log_{10}$ UMI counts per cell and genes were compared between Illumina and ONT datasets, which showed Pearson's correlations of R=0.89 ($p<0.001$) and R=0.87 ($p<0.001$), respectively, revealing strong significant correlations (**Figures 8b&c**). However, we note generally lower values per cell in the ONT datasets compared with the Illumina datasets, most likely owing to the lower sequencing depth. This can be attributed to the higher number of recovered cells in the subsample cells (~15 %) than expected, causing a lower number of reads/cell. Furthermore, the loss of ~40-50% of reads due to inability to match to a cell barcode likely contributed to this result (**Table S6**). However, each ONT read provides more information than a given Illumina read (such as splicing information) and therefore we proceeded with the downstream analysis by pseudo-bulking and merging the subsets of each cell-type (e.g. Ciliated-1-3 → Ciliated, see **Methods**). Additionally, long-read derived transcriptomes can contain many artifactual novel transcripts. Therefore, we employed *SQANTI* filtering to ascertain and exclude spurious transcripts (Tardaguila et al., 2018). Out of the 238,113 isoforms in the final merged reference, we identified 208,000 novel transcripts (~87.4%). After the filtering, out of the 208,000 novel transcripts, 164,503 isoforms (~79.1%) were classified as artifacts. Post-filtering, we observed a change in the proportion of different classes of transcripts (*SQANTI* classes), where Incomplete Splice Match (ISM) transcripts and Full Splice-Match (FSM) classes were most abundant before and after the filtering, respectively (**Figures 8d&e**).

Differential transcript usage occurs in IFN and translation-related genes

To determine per-cell-type DTU between different conditions, the count matrix derived from *FLAMES* was pseudo-bulked and used as the input for DTU analysis with *DTUrtle*. We compared the differences in transcript usage in infected vs uninfected cells, Alpha vs WT infected cells, infected vs bystander cells and bystander vs control cells (**Table S2**). Firstly, we noted that no significant DTU occurred between infected vs bystander cells. This suggested that transcript usage is similar between infected and bystander cells,

despite significant DE being found between these datasets (**Figures 7a-e**). Therefore, these results reiterate that DTU does not generally correlate with DE even within single-cell datasets. However, we note that significant DTU occurred between bystander and control cells, showing the need for further investigations into the role of bystander cells during SARS-CoV-2 infection.

Next, using the significant DTU genes between datasets which passed the first-stage omnibus test, we were able to determine the enriched GO terms and reactome pathways, using *multiGO* (**Figures 8f-g & S7a-d**). We scanned for commonly enriched GO terms and reactome pathways which were found in multiple analysis types (infected vs mock-control cells, Alpha vs WT infected cells, infected vs bystander cells and bystander vs control cells) to understand key pathways involved in SARS-CoV-2 infections. Interestingly, across the different analyses, we observed the repeated enrichment of translational processes such as *translation* and *cytoplasmic translation* in GO terms in less-susceptible cells such as basal, cycling-basal and goblet cells (**Figure 8f & S7a&c**). Similarly, commonly enriched reactome pathways showed *viral mRNA translation, translation, eukaryotic translational initiation* (**Figure 8g & S7b&d**). Furthermore, the lack of enrichment in highly-susceptible cells such as ciliated or secretory-ciliated cells was noted. This suggested that DTU is spread between genes involved in a wide variety of pathways/processes in highly-infected cells, compared with lowly-infected cells. In contrast, we noted that unlike the DE results, IFN-response-related GO terms were generally not found to be enriched (**Figures 8f-g & S7a-d**) . However, reactome pathways showed enrichment of *viral mRNA translation* (cycling-basal and goblet cells in Alpha-infected vs control child), *influenza viral RNA transcription and replication, influenza infection* and *infectious disease* (Alpha vs WT (Secretory-Ciliated), bystander vs control (Cycling-basal WT 48 child, Basal WT 72 child), Alpha-infected vs mock-control (Cycling-basal child, Goblet child)) and *SARS-CoV-2 infections* (Alpha vs WT (Secretory-ciliated child), WT 72 hpi-infected vs mock-control (Basal child)).

Then, some key genes which were found to be commonly DTU across and within different comparisons were investigated. Within infected vs control datasets, we observed significant DTU of IFN response and ribosomal protein genes. For instance, DTU of *IFI27* was observed in Basal Alpha (adult), Ciliated Alpha (adult & child), Cycling-basal Alpha (adult & child), Goblet Alpha (adult), Secretory Alpha (adult & child), Suprabasal Alpha (adult) datasets. Similarly, host ribosomes are manipulated by SARS-CoV-2 for viral protein translation (Eriani & Martin, 2022). We observed ribosomal genes such as *RPL4* (Basal WT (child), Ciliated Alpha (adult), Goblet Alpha (child), Secretory Alpha (adult), Suprabasal Alpha (adult), Suprabasal WT (child)) and *RPL15* (Basal Alpha (adult & child), Basal WT (child), Ciliated Alpha (adult & child),

Ciliated WT (child), Cycling-basal Alpha (child), Cycling-basal WT (child), Goblet Alpha (child), Secretory Alpha (child), Secretory WT (child), Suprabasal Alpha (child), Suprabasal WT (child)) as being commonly DTU. *RPL15* and an ISG *MX1* (which has been shown to be increased in COVID-19 patients (Bizzotto et al., 2020)), were also found to be DTU in all Alpha vs WT comparisons (**Figure S8 & Table S7**), further highlighting the implication of the remodeling of ribosomal protein genes as well as some IFN-response-related genes (**Table S7**), which appeared to be strain-dependent.

Interestingly, we noticed that although DTU was observed at the gene level (first-stage), individual transcript differences between conditions were generally lacking. This indicated that the overall proportional distributions of all transcripts differ between conditions, but not the proportion of individual transcripts. Furthermore, this did not appear to be due to the lack of counts across different transcripts due to pseudo-bulking, which may lead to lower power of significance testing. However, we found some second-stage significant DTU of transcripts belonging to *RPL15*, a ribosomal protein gene. We found that in Alpha-infected Ciliated adolescent cells had a significantly increased usage of *ENST00000413699/RPL15-204* compared with mock-control cells, but not in adults (**Figures 8h&i**), revealing an age-dependent usage. This transcript is a reference protein-coding transcript with three coding exons and a transcript length of 2,221 bps. Surprisingly, the same was found in Secretory cells and Basal cells in adolescents, where *RPL15-204* showed an increase in transcript usage in Alpha-infected cells compared with mock-control but not in adults.

We were then curious whether our scRNA-seq data revealed any similar results to our earlier study with mammalian continuous cell lines infected with WT-strain of SARS-CoV-2 sequenced by bulk RNA-seq (Chang et al., 2022) (**Table S8**). Of the tested genes, *RPL4, PKM, ATRAID, ADK, WARS1, TRA2A, NACA* and *GSDMB* were among the DTU genes in the bulk RNA-seq study and also appeared to be DTU in the infected vs uninfected datasets. Of note, while the other genes were found in multiple comparisons across the infected vs uninfected datasets, *GSDMB* was only found to be DTU within Ciliated Alpha child vs control datasets, highlighting a cell, age and strain-specific response. However, no transcripts passed the second-stage filtering for *GSDMB* within this dataset (**Figure 7j**). Additionally, all Alpha vs WT comparisons involved the DTU of *ATRAID*, suggesting that this is a key gene which interacts differently with the two SARS-CoV-2 strains used in this study.

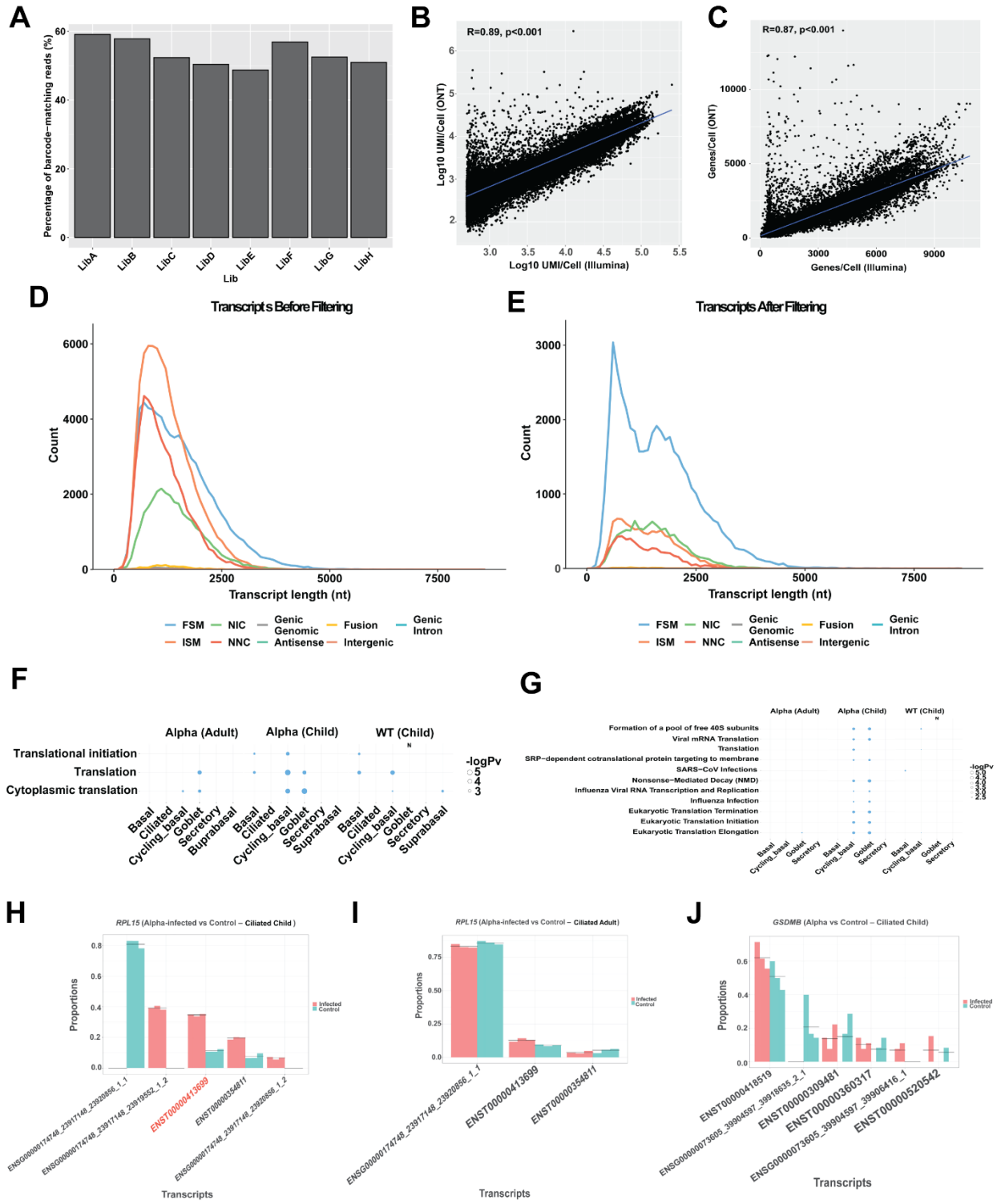


Figure 8. Differential transcript usage occurs in SARS-CoV-2 infected cells compared with control cells. A) Percentage of barcode-matching reads according to FLAMES results per PromethION library. **B**) Log*10* UMI counts per cell in ONT and Illumina datasets. **C)** Number of genes detected per cell in ONT and Illumina datasets. **(D-E**) Transcripts in the reference-guided transcriptome assembly **D)** before and **E)** after *SQANTI3*

filtering of artifactual isoforms. **(F-G)** *multiGO* outputs showing enriched **F)** GO biological terms and **G)** reactome pathways from infected vs control datasets. Secretory-ciliated cells were unable to be included in this analysis due to lack of control cells to meet minimum cells for pseudo-bulking and minimum donor replicates. Bubble size indicates -log10 enrichment p-values. Columns with no matching DE data available are denoted with 'N'. (**H-I)** *DTUrtle* outputs of *RPL15* transcript proportions in Alpha-infected ciliated cells vs mock-control in **H)** adolescents and **I)** adults. **J)** *DTUrtle* outputs of *GSDMB* transcript proportions in Alpha-infected ciliated cells vs mock-control in adolescents. X-axis shows the transcript/isoforms belonging to the gene of interest and Y-axis shows the proportion of each transcript.

Discussion

ALI-cultures are effective *in vitro* models for interrogating host-pathogen interactions and have been shown to be useful for imitating SARS-CoV-2 infections (Ravindra et al., 2021; Tran, Deliyannis, et al., 2022; Tran, Grimley, et al., 2022). In the dynamic heterogeneous differentiation process that is ALI-culturing, transitional cell-types should be expected in such models. While many studies have identified ciliated cells as being the most SARS-CoV-2-susceptible cell-type in the human nasal epithelia (Ahn et al., 2021; Robinot et al., 2021), we found high levels of infection within cells which show ciliated cells with secretory properties, with some notable infection in secretory cells (**Figures 3b**). The infection of such transitional cells have been noted before (Robinot et al., 2021). We speculate that this phenomenon can be attributed to **1)** secretory cells being precursors to ciliated cells (Rawlins et al., 2009), and therefore de-differentiation of ciliated cells following infection proposed by Robinot et al., (2021) may convert ciliated cells back into the transitional epithelial state; OR **2)** secretory-ciliated cells are present prior to infection and are also infected due to their ciliated properties. If in the case of **1)** as suggested by Robinot et al., we would expect a significant decrease in proportion of ciliated cells upon infection. However, *Propeller* analysis revealed that while a significant increase of Secretory-Ciliated+SC2 cells was observed in Alpha and WT 72hpi-infected datasets, the significant decrease of ciliated cells was only observed in Ciliated-2 cells within WT 48 and 72 hpi datasets in adults (**Data S2**) (Phipson et al., 2022). If in the case of **2)**, we would expect a decrease in proportion of Secretory-Ciliated (low CS2) cells with infected datasets. However, contrary to expectations, Secretory-Ciliated cells (low SC2), either significantly or non-significantly increased in proportion upon infection compared with mock-control cells. These results implied that perhaps there are other unknown mechanisms occurring during the infection process, and further work will be required to fully understand the dynamics of these transitional cells, with the involvement of microscopy studies. One speculation for the expansion of Secretory-Ciliated cells may be that as increase of mucus production is favorable for trapping viral particles during infection, and perhaps ciliated cells may acquire more secretory properties to facilitate this activity.

The main objective of this study was to understand the age and strain-dependent responses to SARS-CoV-2. At the 72 hpi time point, the TCID50 (**Figures 2a-b**) and short-read sequencing data (**Figure 3c**) showed that the Alpha variant infected adults and adolescents similarly and the WT strain resulted in much higher viral load in adolescents compared with adults. Also, generally the Alpha-variant infections in general yielded higher viral titers (**Figures 2c-d**) and viral reads (**Figure 3c**) than WT strain-infections. The elevated viral titers with Alpha-infections compared with WT was expected, due to the increased transmissibility in the variant (Y. Liu et al., 2022; Loenenbach et al., 2021). However, the distinctly increased viral titers and reads in WT-infected adolescents compared with adults in all replicates was unexpected. This is because it has been thought that children, especially with WT-infections are less susceptible to SARS-CoV-2 (Silverberg et al., 2022). One of the reasons is perhaps the lower abundance of *ACE2* receptors in the upper airways of children compared with adults. However, there are mixed reports regarding age-dependent *ACE2* mRNA levels, showing either that *ACE2* mRNA expression can increase with age in the nasal epithelia (Bunyavanich, Do, & Vicencio, 2020), or there are no age-dependent effects (Berni Canani et al., 2021). In this study, we did not observe significant differences in *ACE2* mRNA expression level between infected adults and adolescents in all tested cell-clusters (FDR ≥ 0.05), perhaps due to this study involving a single-cell method, preserving the heterogeneity in the sample. Therefore, our data appeared to be in line with previous studies suggesting that *ACE2* transcriptional levels do not correlate with susceptibility to SARS-CoV-2 (Ahn et al., 2021). Despite the lack of differences between *ACE2* at the mRNA level, we note that the protein level may be different, as staining of *ACE2* protein has been shown previously in ALI-HNECs (Gamage et al., 2020), which could cause these differences in viral load. However, another similar study using ALI-HNECs derived from adults and children using multiple strains of SARS-CoV-2 has shown WT-strain infects less in children vs adults, which is directly in contradiction to our results (Y. Zhu et al., 2021). Although unlikely, there is also a possibility that the MOI used for infecting the adolescents may have been unintentionally increased compared with the adults with the WT infections, e.g. due to overestimation of the number of host cells.

ACE2 mRNA expression was low across all cell types (**Figures 4a&c**). However, we noticed that the gene was particularly upregulated in IFN-stimulated populations of cells such as Secretory+IFN-stim and Goblet+IFN-stim cells and Ciliated-1 cells (**Figures 4d-k**). This was in line with evidence that *ACE2* is an ISG (Carly G. K. Ziegler et al., 2020). Interestingly, a similar study involving ALI-HAECs infected with SARS-CoV-2 did not show increased levels of *ACE2* mRNA after infection (Ravindra et al., 2021) and similarly between HNECs derived from COVID-19 patients vs healthy controls (Ahn et al., 2021). We

speculate that perhaps if an IFN-stimulated cluster was also separated from the main body of cells, these datasets may also show similar results. Furthermore, *ACE2* was not found to be upregulated in bystander cells when compared with mock-control cells, although induction of ISGs occurred in these cells (**Figures S5a-c**). In the literature, bystander cells appeared to be stimulated upon exposure to SARS-CoV-2 virus in HBECs (Ravindra et al., 2021) and increases in ISG expression in IAV infections (Hamele, Russell, & Heaton, 2022). Hence, the ISG-like properties of *ACE2* maybe age, strain, transmission or viral load dependent, and further work will be required to fully understand this mechanism.

Cells with comparatively lower viral loads compared with Ciliated+SC2 and Secretory-Ciliated+SC2 showed enrichment of terms and pathways related to innate immune responses, and an absence of these responses were observed in highly-infected cell-clusters (**Figure 5a**). We speculate that these cells did not mount a robust IFN-response during early infection, leading to a higher viral load. Otherwise, certain ORFs have antagonistic properties to IFN-responses (Shemesh et al., 2021). Multiple virion particles could have infected these cells, leading to increased antagonization of IFN-responses and therefore high viral load. Finally, due to the low MOI applied, multiple rounds of viral infection may occur, which would lead to some cells infecting earlier than others. Hence, these high viral-load cells may be part of this early-infection group. The challenges of identifying the source of cell-to-cell heterogeneity in virus-infected scRNA-seq datasets have been reviewed by Suomalainen and Greber (2021).

We also assessed the differences in responses between infected adolescents and infected adults. Loske et al. (2021) showed the pre-primed immunity to SARS-CoV-2 as well as increased ISG induction in children compared with adults. Contrary to this finding, we did not observe strong enrichment of IFN-response-related GO terms when in infected adolescent organoids were compared with infected adult organoids (**Figure 6a**). However, we did find the enrichment of translational, oxidative phosphorylation and cilia-related terms, where most genes were found to be downregulated in the infected adolescents compared with the infected adults. Similarly, with reactome pathways, the level of gene expression of genes involved in *translation, respiratory electron chain* and *influenza infection* pathways were decreased compared with infected adults (**Figure S7c**). This suggested that in the infected state, adults require higher levels of these genes compared with adolescents. Considering that cilia-related GO terms was found to be enriched with the significant genes when expression levels of COVID-19 airways was compared with healthy controls (Yoshida et al., 2021), potentially the higher requirement of these genes in adults may be indicative of a greater damage caused by SARS-CoV-2 in adults compared with adolescents. Overall, these results contrast

with the evidence of a stronger IFN response gene expression in children compared with adults (Loske et al., 2021; Yoshida et al., 2021).

The Alpha-variant infected cells showed increase expression of genes involved in *protein refolding* and *response to unfolded protein*, compared with WT-infected cells in adolescents, in the clusters with the highest level of infection (i.e. Ciliated+SC2 and Secretory-Ciliated+SC2) (**Figure 6b**). In normal conditions, protein refolding response is not activated, and is switched on after an accumulation of unfolded/misfolded proteins occurs under endoplasmic reticulum (ER) stress (Read & Schröder, 2021). The aggregation of unfolded proteins may occur as a host defense mechanism, or viral manipulation to increase replication or viral immune evasion (Hinte, Van Anken, Tirosh, & Brune, 2020; Muscolino, Luoto, & Brune, 2021). The induction of the unfolded protein response (UPR) due to ER stress has been documented with the spike protein of SARS-CoV-1 (Versteeg, Van De Nes, Bredenbeek, & Spaan, 2007) and also SARS-CoV-2 (Hsu et al., 2020). Overall, this increased activity in adolescents may be in part due to the increased transmissibility of the Alpha strain increasing the build-up of misfolded/unfolded proteins. We note the absence of enrichment of these processes in the adult datasets, suggesting an age-dependent effect. However, we note that the matching data with Ciliated+SC2 and Secretory-Ciliated+SC2 clusters in adults were unavailable and these processes could have been reciprocated in these datasets had we been able to analyse these datasets.

Long-read single-cell sequencing has gained attention in recently with the involvement of ONT (Philpott et al., 2021; Wang et al., 2021; You et al., 2022) and Pacific Biosciences (PacBio) (Hazzard et al., 2022) technologies. However, currently, most published single-cell RNA-seq studies still involve short-read sequencing, especially with SARS-CoV-2 infection studies (Puray-Chavez et al., 2020; Purkayastha et al., 2020; Carly G.K. Ziegler et al., 2021). In this study, we have utilized long-read single-cell sequencing in the context of *in vitro* SARS-CoV-2 infections to investigate cell-type-dependent DTU with a transcriptome involving reference and novel transcripts. Largely, translation-related genes were mostly involved between infected vs control datasets, showing the importance of the ribosome (**Figures S8f-i**). Specifically, we found the importance of *RPL15*, where RPL15 has been identified to be a part of the top 10 Level 1 hub proteins interacting with SARS-CoV-2 (Ghosh, Saha, & Sharma, 2022). This gene was DTU in many datasets, including many infected vs uninfected datasets e.g. Alpha-infected Ciliated, Secretory and Basal cells vs control where it showed increase in the usage of the *RPL15-204* transcript of in children but not adults (**Figure 8h**), showing an age-dependent effect. Furthermore, *RPL4* was also commonly DTU in infected vs

control datasets, which was also found to be DTU in bulk RNA-seq of Caco-2 cells infected with WT SARS-CoV-2 (Chang et al., 2022). *RPL4* has been found to be differentially expressed in nasopharyngeal samples of COVID-19 patients vs controls with bulk RNA-seq (Hoque et al., 2022). The protein of this gene has also been shown to be a primary interactor of SARS-CoV-2 proteins (Biji et al., 2021). This highlights the importance of *RPL4* upon SARS-CoV-2 infection, regardless of WT or Alpha strain. Together, these results highlight the advantages of long-read analysis in addition to short-read analysis in transcriptomic studies and implies the importance of ribosomal protein genes during SARS-CoV-2 infections in the human nasal epithelia.

Limitations

While we have utilized an organoid model approach for our study which is superior compared with continuous cell lines, we acknowledge that these results may not directly translate to *in vivo* situations. Particularly, the absence of immune cells in the model may distort the relevance of these results within this study. However, via this model, we were able to determine the epithelial immune responses to SARS-CoV-2 without confounding effects from immune-epithelial cell interactions. Additionally, we have also used a low MOI of < 0.02, which may be more clinically relevant as low numbers of virions will initiate an *in vivo* infection, but this means that the infection stage of each cell will be temporally asynchronized. Furthermore, we acknowledge the lower coverage in the ONT datasets and therefore, the DTU analysis will be required to undergo further investigation to gain an improved understanding of the transcript usage activity within SARS-CoV-2 infections. However, we have attempted to minimize this effect by the utility of pseudo-bulking and have shown the strong correlation between the number of genes and $\log_{10}$ UMIs between Illumina and ONT datasets. Finally, one donor (child donor 6/donor 3) showed some differences compared with other donors. However, most cell-types aligned with other donors, as shown by the clustering analysis and this effect has been minimized by applying filters such as minimum number of donor replicates and minimum number of cells for pseudo-bulking. Furthermore, due to ambient RNA being able to be encapsulated into 10X droplets, there is a potential overestimation of true viral RNA load in each cell. We have applied a threshold of 10 UMI viral counts per cell for a cell to be deemed as infected to account for the potential contamination, according to the empirical threshold found by Ravindra et al.

Acknowledgements

We thank the Biological Optical Microscopy Platform (BOMP) at the University of Melbourne for their assistance. This research was supported by The University of Melbourne's Research Computing Services and the Petascale Campus Initiative. This research was undertaken using the LIEF HPC-GPGPU Facility

hosted at the University of Melbourne. This Facility was established with the assistance of LIEF Grant LE170100200. We also thank Dr. Jafar Jabbari for guidance with the FLT-seq protocol and Dr. Josh Lee and Dr. Jan Schroeder for guidance regarding single-cell analysis. We thank Dr. Quentin Gouil from the Walter and Eliza Hall Institute for sequencing all PromethION libraries generated for this study and guidance on preparing the libraries. Additionally we wish to thank the Ramaciotti Center for Genomics for sequencing of all Illumina libraries and Macrogen for running the microarrays for genotyping analysis.

References

Ahn, J. H., Kim, J., Hong, S. P., Choi, S. Y., Yang, M. J., Ju, Y. S., . . . Koh, G. Y. (2021). Nasal ciliated cells are primary targets for SARS-CoV-2 replication in the early stage of COVID-19. *Journal of Clinical Investigation, 131*(13). doi:10.1172/jci148517

Awatade, N. T., Wong, S. L., Capraro, A., Pandzic, E., Slapetova, I., Zhong, L., . . . Waters, S. A. (2021). Significant functional differences in differentiated Conditionally Reprogrammed (CRC)- and Feeder-free Dual SMAD inhibited-expanded human nasal epithelial cells. *J Cyst Fibros, 20*(2), 364-371. doi:10.1016/j.jcf.2020.12.019

Berni Canani, R., Comegna, M., Paparo, L., Cernera, G., Bruno, C., Strisciuglio, C., . . . Castaldo, G. (2021). Age-Related Differences in the Expression of Most Relevant Mediators of SARS-CoV-2 Infection in Human Respiratory and Gastrointestinal Tract. *Frontiers in Pediatrics, 9*. doi:10.3389/fped.2021.697390

Biji, A., Khatun, O., Swaraj, S., Narayan, R., Rajmani, R. S., Sardar, R., . . . Tripathi, S. (2021). Identification of COVID-19 prognostic markers and therapeutic targets through meta-analysis and validation of Omics data from nasopharyngeal samples. *EBioMedicine, 70*, 103525. doi:10.1016/j.ebiom.2021.103525

Bizzotto, J., Sanchis, P., Abbate, M., Lage-Vickers, S., Lavignolle, R., Toro, A., . . . Gueron, G. (2020). SARS-CoV-2 Infection Boosts MX1 Antiviral Effector in COVID-19 Patients. *iScience, 23*(10), 101585. doi:10.1016/j.isci.2020.101585

Brookman, S., Cook, J., Zucherman, M., Broughton, S., Harman, K., & Gupta, A. (2021). Effect of the new SARS-CoV-2 variant B.1.1.7 on children and young people. *The Lancet Child & Adolescent Health, 5*(4), e9-e10. doi:10.1016/s2352-4642(21)00030-4

Bunyavanich, S., Do, A., & Vicencio, A. (2020). Nasal Gene Expression of Angiotensin-Converting Enzyme 2 in Children and Adults. *JAMA, 323*(23), 2427. doi:10.1001/jama.2020.8707

Challen, R., Brooks-Pollock, E., Read, J. M., Dyson, L., Tsaneva-Atanasova, K., & Danon, L. (2021). Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: matched cohort study. *BMJ*, n579. doi:10.1136/bmj.n579

Chang, J. J.-Y., Gleeson, J., Rawlinson, D., De Paoli-Iseppi, R., Zhou, C., Mordant, F. L., . . . Pitt, M. E. (2022). Long-Read RNA Sequencing Identifies Polyadenylation Elongation and Differential Transcript Usage of Host Transcripts During SARS-CoV-2 In Vitro Infection. *Front Immunol, 13*. doi:10.3389/fimmu.2022.832223

Consolazio, D., Murtas, R., Tunesi, S., Lamberti, A., Senatore, S., Faccini, M., & Russo, A. G. (2022). A Comparison Between Omicron and Earlier COVID-19 Variants' Disease Severity in the Milan Area, Italy. *Frontiers in Epidemiology, 2*. doi:10.3389/fepid.2022.891162

Czerkies, M., Kochanczyk, M., Korwek, Z., Prus, W., & Lipniacki, T. (2022). Respiratory Syncytial Virus Protects Bystander Cells against Influenza A Virus Infection by Triggering Secretion of Type I and Type III Interferons. *J Virol, 96*(22), e0134122. doi:10.1128/jvi.01341-22

Deprez, M., Zaragosi, L.-E., Truchi, M., Becavin, C., Ruiz García, S., Arguel, M.-J., . . . Barbry, P. (2020). A Single-Cell Atlas of the Human Healthy Airways. *American Journal of Respiratory and Critical Care Medicine, 202*(12), 1636-1645. doi:10.1164/rccm.201911-2199oc

Eriani, G., & Martin, F. (2022). Viral and cellular translation during SARS-CoV-2 infection. *FEBS Open Bio, 12*(9), 1584-1601. doi:10.1002/2211-5463.13413

Esper, F. P., Adhikari, T. M., Tu, Z. J., Cheng, Y. W., El-Haddad, K., Farkas, D. H., . . . Rubin, B. P. (2022). Alpha to Omicron: Disease Severity and Clinical Outcomes of Major SARS-CoV-2 Variants. *J Infect Dis*. doi:10.1093/infdis/jiac411

Ferreira, I. A. T. M., Kemp, S. A., Datir, R., Saito, A., Meng, B., Rakshit, P., . . . Gupta, R. K. (2021). SARS-CoV-2 B.1.617 Mutations L452R and E484Q Are Not Synergistic for Antibody Evasion. *The Journal of Infectious Diseases, 224*(6), 989-994. doi:10.1093/infdis/jiab368

Frampton, D., Rampling, T., Cross, A., Bailey, H., Heaney, J., Byott, M., . . . Nastouli, E. (2021). Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B.1.1.7 lineage in London, UK: a whole-genome sequencing and hospital-based cohort study. *The Lancet Infectious Diseases, 21*(9), 1246-1256. doi:10.1016/s1473-3099(21)00170-5

Gamage, A. M., Tan, K. S., Chan, W. O. Y., Liu, J., Tan, C. W., Ong, Y. K., . . . Wang, L.-F. (2020). Infection of human Nasal Epithelial Cells with SARS-CoV-2 and a 382-nt deletion isolate lacking ORF8 reveals similar viral kinetics and host transcriptional profiles. *PLoS Pathogens, 16*(12), e1009130. doi:10.1371/journal.ppat.1009130

Gao, K. M., Derr, A. G., Guo, Z., Nündel, K., Marshak-Rothstein, A., Finberg, R. W., & Wang, J. P. (2021). Human nasal wash RNA-Seq reveals distinct cell-specific innate immune responses in influenza versus SARS-CoV-2. *JCI Insight, 6*(22). doi:10.1172/jci.insight.152288

Ge, S. X., Jung, D., & Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics, 36*(8), 2628-2629. doi:10.1093/bioinformatics/btz931

Ghosh, N., Saha, I., & Sharma, N. (2022). Palindromic target site identification in SARS-CoV-2, MERS-CoV and SARS-CoV-1 by adopting CRISPR-Cas technique. *Gene, 818*, 146136. doi:https://doi.org/10.1016/j.gene.2021.146136

Giles, B., Meredith, P., Robson, S., Smith, G., & Chauhan, A. (2021). The SARS-CoV-2 B.1.1.7 variant and increased clinical severity—the jury is out. *The Lancet Infectious Diseases, 21*(9), 1213-1214. doi:10.1016/s1473-3099(21)00356-x

Graham, M. S., Sudre, C. H., May, A., Antonelli, M., Murray, B., Varsavsky, T., . . . Gunson, R. N. (2021). Changes in symptomatology, reinfection, and transmissibility associated with the SARS-CoV-2 variant B.1.1.7: an ecological study. *The Lancet Public Health, 6*(5), e335-e345. doi:10.1016/s2468-2667(21)00055-4

Hamele, C. E., Russell, A. B., & Heaton, N. S. (2022). In Vivo Profiling of Individual Multiciliated Cells during Acute Influenza A Virus Infection. *J Virol, 96*(14), e0050522. doi:10.1128/jvi.00505-22

Hazzard, B., Sá, J. M., Ellis, A. C., Pascini, T. V., Amin, S., Wellems, T. E., & Serre, D. (2022). Long read single cell RNA sequencing reveals the isoform diversity of Plasmodium vivax transcripts. *PLOS Neglected Tropical Diseases, 16*(12), e0010991. doi:10.1371/journal.pntd.0010991

Hinte, F., Van Anken, E., Tirosh, B., & Brune, W. (2020). Repression of viral gene expression and replication by the unfolded protein response effector XBP1u. *eLife, 9*. doi:10.7554/elife.51804

Hoque, M. N., Sarkar, M. M. H., Khan, M. A., Hossain, M. A., Hasan, M. I., Rahman, M. H., . . . Islam, T. (2022). Differential gene expression profiling reveals potential biomarkers and pharmacological compounds against SARS-CoV-2: Insights from machine learning and bioinformatics approaches. *Front Immunol, 13*, 918692. doi:10.3389/fimmu.2022.918692

Hsu, A. C.-Y., Wang, G., Reid, A. T., Veerati, P. C., Pathinayake, P. S., Daly, K., . . . Wark, P. A. (2020). *SARS-CoV-2 Spike protein promotes hyper-inflammatory response that can be ameliorated by Spike-antagonistic peptide and FDA-approved ER stress and MAP kinase inhibitors in vitro*. Cold Spring Harbor Laboratory.

Jabbari, J. S., & Tian, L. (2019). Massively parallel long-read sequencing of single cell RNA isoforms. *protocols.io*.

Kang, H. M., Subramaniam, M., Targ, S., Nguyen, M., Maliskova, L., McCarthy, E., . . . Ye, C. J. (2018). Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat Biotechnol, 36*(1), 89-94. doi:10.1038/nbt.4042

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics, 34*(18), 3094-3100. doi:10.1093/bioinformatics/bty191

Liu, X., Ory, V., Chapman, S., Yuan, H., Albanese, C., Kallakury, B., . . . Schlegel, R. (2012). ROCK Inhibitor and Feeder Cells Induce the Conditional Reprogramming of Epithelial Cells. *The American Journal of Pathology, 180*(2), 599-607. doi:10.1016/j.ajpath.2011.10.036

Liu, Y., Liu, J., Plante, K. S., Plante, J. A., Xie, X., Zhang, X., . . . Weaver, S. C. (2022). The N501Y spike substitution enhances SARS-CoV-2 infection and transmission. *Nature, 602*(7896), 294-299. doi:10.1038/s41586-021-04245-0

Loenenbach, A., Markus, I., Lehfeld, A.-S., An Der Heiden, M., Haas, W., Kiegele, M., . . . Buchholz, U. (2021). SARS-CoV-2 variant B.1.1.7 susceptibility and infectiousness of children and adults deduced from investigations of childcare centre outbreaks, Germany, 2021. *Eurosurveillance, 26*(21). doi:10.2807/1560-7917.es.2021.26.21.2100433

Loske, J., Röhmel, J., Lukassen, S., Stricker, S., Magalhães, V. G., Liebig, J., . . . Lehmann, I. (2021). Pre-activated antiviral innate immunity in the upper airways controls early SARS-CoV-2 infection in children. *Nature Biotechnology*. doi:10.1038/s41587-021-01037-9

Lukassen, S., Chua, R. L., Trefzer, T., Kahn, N. C., Schneider, M. A., Muley, T., . . . Eils, R. (2020). SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. *The EMBO Journal, 39*(10). doi:10.15252/embj.20105114

Meyer, M., Holfter, A., Ruebsteck, E., Gruell, H., Dewald, F., Koerner, R. W., . . . Weber, L. T. (2021). The Alpha Variant (B.1.1.7) of SARS-CoV-2 in Children: First Experience from 3544 Nucleic Acid Amplification Tests in a Cohort of Children in Germany. *Viruses, 13*(8), 1600. doi:10.3390/v13081600

Mohammadi, M., Shayestehpour, M., & Mirzaei, H. (2021). The impact of spike mutated variants of SARS-CoV2 [Alpha, Beta, Gamma, Delta, and Lambda] on the efficacy of subunit recombinant vaccines. *The Brazilian Journal of Infectious Diseases, 25*(4), 101606. doi:https://doi.org/10.1016/j.bjid.2021.101606

Muscolino, E., Luoto, L.-M., & Brune, W. (2021). Viral Induced Protein Aggregation: A Mechanism of Immune Evasion. *International Journal of Molecular Sciences, 22*(17), 9624. doi:10.3390/ijms22179624

Niu, Z., Zhang, Z., Gao, X., Du, P., Lu, J., Yan, B., . . . Sun, Q. (2021). N501Y mutation imparts cross-species transmission of SARS-CoV-2 to mice by enhancing receptor binding. *Signal Transduction and Targeted Therapy, 6*(1). doi:10.1038/s41392-021-00704-2

Philpott, M., Watson, J., Thakurta, A., Brown, T., Brown, T., Oppermann, U., & Cribbs, A. P. (2021). Nanopore sequencing of single-cell transcriptomes with scCOLOR-seq. *Nature Biotechnology, 39*(12), 1517-1520. doi:10.1038/s41587-021-00965-w

Phipson, B., Sim, C. B., Porrello, E. R., Hewitt, A. W., Powell, J., & Oshlack, A. (2022). propeller: testing for differences in cell type proportions in single cell data. *Bioinformatics, 38*(20), 4720-4726. doi:10.1093/bioinformatics/btac582

Public Health England. (2020). *Investigation of novel SARS-CoV-2 variant: variant of concern 202012/01, technical briefing 3.* Retrieved from London, United Kingdom: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/950823/Variant_of_Concern_VOC_202012_01_Technical_Briefing_3_-_England.pdf

Puray-Chavez, M., Lee, N., Tenneti, K., Wang, Y., Vuong, H. R., Liu, Y., . . . Kutluay, S. B. (2020). *The translational landscape of SARS-CoV-2 and infected cells*. Cold Spring Harbor Laboratory.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., . . . Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet, 81*(3), 559-575. doi:10.1086/519795

Purkayastha, A., Sen, C., Garcia, G., Langerman, J., Shia, D. W., Meneses, L. K., . . . Gomperts, B. N. (2020). Direct Exposure to SARS-CoV-2 and Cigarette Smoke Increases Infection Severity and Alters the Stem Cell-Derived Airway Repair Response. *Cell Stem Cell, 27*(6), 869-875.e864. doi:10.1016/j.stem.2020.11.010

Ravindra, N. G., Alfajaro, M. M., Gasque, V., Huston, N. C., Wan, H., Szigeti-Buck, K., . . . Wilen, C. B. (2021). Single-cell longitudinal analysis of SARS-CoV-2 infection in human airway epithelium identifies target cells, alterations in gene expression, and cell state changes. *PLoS Biology, 19*(3), e3001143. doi:10.1371/journal.pbio.3001143

Rawlins, E. L., Okubo, T., Xue, Y., Brass, D. M., Auten, R. L., Hasegawa, H., . . . Hogan, B. L. M. (2009). The Role of Scgb1a1+ Clara Cells in the Long-Term Maintenance and Repair of Lung Airway, but Not Alveolar, Epithelium. *Cell Stem Cell, 4*(6), 525-534. doi:10.1016/j.stem.2009.04.002

Read, A., & Schröder, M. (2021). The Unfolded Protein Response: An Overview. *Biology, 10*(5), 384. doi:10.3390/biology10050384

Reed, L. J., & Muench, H. (1938). A simple method of estimating fifty per cent endpoints. *American Journal of Epidemiology, 27*(3), 493-497. doi:10.1093/oxfordjournals.aje.a118408

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research, 43*(7), e47-e47. doi:10.1093/nar/gkv007

Robinot, R., Hubert, M., De Melo, G. D., Lazarini, F., Bruel, T., Smith, N., . . . Chakrabarti, L. A. (2021). SARS-CoV-2 infection induces the dedifferentiation of multiciliated cells and impairs mucociliary clearance. *Nature Communications, 12*(1). doi:10.1038/s41467-021-24521-x

Ruiz García, S., Deprez, M., Lebrigand, K., Cavard, A., Paquet, A., Arguel, M.-J., . . . Zaragosi, L.-E. (2019). Novel dynamics of human mucociliary differentiation revealed by single-cell RNA sequencing of nasal epithelial cultures. *Development*, dev.177428. doi:10.1242/dev.177428

Shemesh, M., Aktepe, T. E., Deerain, J. M., Mcauley, J. L., Audsley, M. D., David, C. T., . . . Harari, D. (2021). SARS-CoV-2 suppresses IFNβ production mediated by NSP1, 5, 6, 15, ORF6 and ORF7b but does not suppress the effects of added interferon. *PLoS Pathogens, 17*(8), e1009800. doi:10.1371/journal.ppat.1009800

Silverberg, S. L., Zhang, B. Y., Li, S. N. J., Burgert, C., Shulha, H. P., Kitchin, V., . . . Sadarangani, M. (2022). Child transmission of SARS-CoV-2: a systematic review and meta-analysis. *BMC Pediatr, 22*(1), 172. doi:10.1186/s12887-022-03175-8

Squair, J. W., Gautier, M., Kathe, C., Anderson, M. A., James, N. D., Hutson, T. H., . . . Courtine, G. (2021). Confronting false discoveries in single-cell differential expression. *Nature Communications, 12*(1). doi:10.1038/s41467-021-25960-2

Sungnak, W., Huang, N., Bécavin, C., Berg, M., Queen, R., Litvinukova, M., . . . Barnes, J. L. (2020). SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nature Medicine, 26*(5), 681-687. doi:10.1038/s41591-020-0868-6

Suomalainen, M., & Greber, U. F. (2021). Virus Infection Variability by Single-Cell Profiling. *Viruses, 13*(8), 1568. doi:10.3390/v13081568

Suprynowicz, F. A., Upadhyay, G., Krawczyk, E., Kramer, S. C., Hebert, J. D., Liu, X., . . . Schlegel, R. (2012). Conditionally reprogrammed cells represent a stem-like state of adult epithelial cells. *Proceedings of the National Academy of Sciences, 109*(49), 20035-20040. doi:10.1073/pnas.1213241109

Tardaguila, M., De La Fuente, L., Marti, C., Pereira, C., Pardo-Palacios, F. J., Del Risco, H., . . . Conesa, A. (2018). SQANTI: extensive characterization of long-read transcript sequences for quality control in full-length transcriptome identification and quantification. *Genome Research, 28*(3), 396-411. doi:10.1101/gr.222976.117

Tekath, T., & Dugas, M. (2021). Differential transcript usage analysis of bulk and single-cell RNA-seq data with DTUrtle. *Bioinformatics, 37*(21), 3781-3787. doi:10.1093/bioinformatics/btab629

Tian, L., Jabbari, J. S., Thijssen, R., Gouil, Q., Amarasinghe, S. L., Voogd, O., . . . Ritchie, M. E. (2021). Comprehensive characterization of single-cell full-length isoforms in human and mouse with long-read sequencing. *Genome Biology, 22*(1). doi:10.1186/s13059-021-02525-6

Tran, B. M., Deliyannis, G., Hachani, A., Earnest, L., Torresi, J., & Vincan, E. (2022). Organoid Models of SARS-CoV-2 Infection: What Have We Learned about COVID-19? *Organoids, 1*(1), 2-27. doi:10.3390/organoids1010002

Tran, B. M., Grimley, S. L., Mcauley, J. L., Hachani, A., Earnest, L., Wong, S. L., . . . Vincan, E. (2022). Air-Liquid-Interface Differentiated Human Nose Epithelium: A Robust Primary Tissue Culture Model of SARS-CoV-2 Infection. *International Journal of Molecular Sciences, 23*(2), 835. doi:10.3390/ijms23020835

Versteeg, G. A., Van De Nes, P. S., Bredenbeek, P. J., & Spaan, W. J. M. (2007). The Coronavirus Spike Protein Induces Endoplasmic Reticulum Stress and Upregulation of Intracellular Chemokine mRNA Concentrations. *Journal of Virology, 81*(20), 10981-10990. doi:10.1128/jvi.01033-07

Wang, Q., Bönigk, S., Böhm, V., Gehring, N., Altmüller, J., & Dieterich, C. (2021). Single-cell transcriptome sequencing on the Nanopore platform with ScNapBar. *RNA, 27*(7), 763-770. doi:10.1261/rna.078154.120

Yoshida, M., Worlock, K. B., Huang, N., Lindeboom, R. G. H., Butler, C. R., Kumasaka, N., . . . Meyer, K. B. (2021). Local and systemic responses to SARS-CoV-2 infection in children and adults. *Nature*. doi:10.1038/s41586-021-04345-x

You, Y., Prawer, Y. D. J., De Paoli-Iseppi, R., Hunt, C. P. J., Parish, C. L., Shim, H., & Clark, M. B. (2022). *Identification of cell barcodes from long-read single-cell RNA-seq with BLAZE*. Cold Spring Harbor Laboratory.

Zhang, Z., Cui, F., Cao, C., Wang, Q., & Zou, Q. (2022). Single-cell RNA analysis reveals the potential risk of organ-specific cell types vulnerable to SARS-CoV-2 infections. *Computers in Biology and Medicine, 140*, 105092. doi:10.1016/j.compbiomed.2021.105092

Zhao, H., Sun, Z., Wang, J., Huang, H., Kocher, J. P., & Wang, L. (2014). CrossMap: a versatile tool for coordinate conversion between genome assemblies. *Bioinformatics, 30*(7), 1006-1007. doi:10.1093/bioinformatics/btt730

Zhu, N., Wang, W., Liu, Z., Liang, C., Wang, W., Ye, F., . . . Tan, W. (2020). Morphogenesis and cytopathic effect of SARS-CoV-2 infection in human airway epithelial cells. *Nature Communications, 11*(1). doi:10.1038/s41467-020-17796-z

Zhu, Y., Chew, K. Y., Karawita, A. C., Yamamoto, A., Labzin, L. L., Yarlagadda, T., . . . Short, K. R. (2021). *Pediatric nasal epithelial cells are less permissive to SARS-CoV-2 replication compared to adult cells*. Cold Spring Harbor Laboratory. Retrieved from https://dx.doi.org/10.1101/2021.03.08.434300

Ziegler, C. G. K., Allon, S. J., Nyquist, S. K., Mbano, I. M., Miao, V. N., Tzouanas, C. N., . . . Zhang, K. (2020). SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells

and Is Detected in Specific Cell Subsets across Tissues. *Cell, 181*(5), 1016-1035.e1019. doi:10.1016/j.cell.2020.04.035

Ziegler, C. G. K., Miao, V. N., Owings, A. H., Navia, A. W., Tang, Y., Bromley, J. D., . . . Ordovas-Montanes, J. (2021). Impaired local intrinsic immunity to SARS-CoV-2 infection in severe COVID-19. *Cell, 184*(18), 4713-4733.e4722. doi:10.1016/j.cell.2021.07.023

Supplementary Items

Supplementary Data

Data S1. Output from FindAllmarkers to determine cell-clusters. Default parameters were used.

Data S2. Outputs from *Propeller* analysis of changes in cell-type distributions.

Supplementary Figures

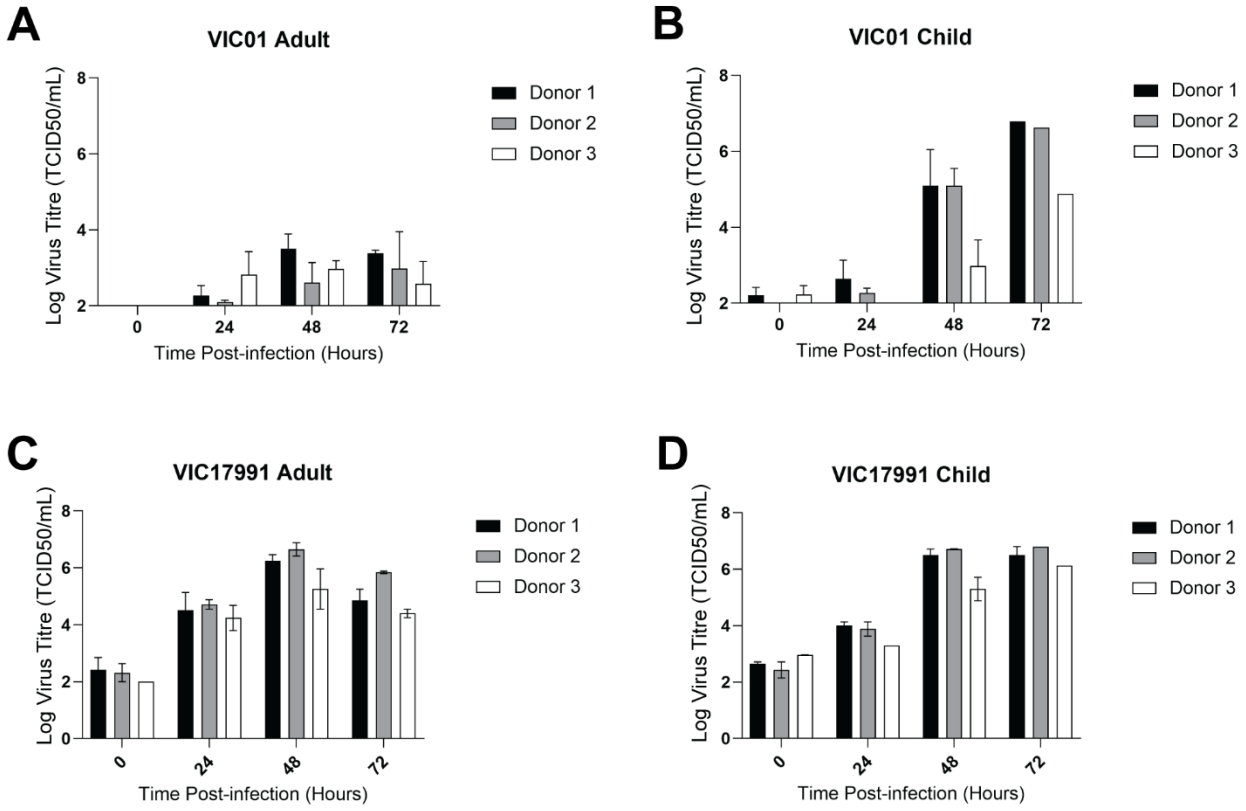


Figure S1. Viral load, titers and visualization of SARS-CoV-2 infected ALI-HNE per donor. TCID50 results from apical washes at 0, 24, 48. 72 hpi comparing WT-infections in **A)** adults and **B)** adolescents. Comparison of Alpha-infections in **C)** adults and **D)** adolescents.

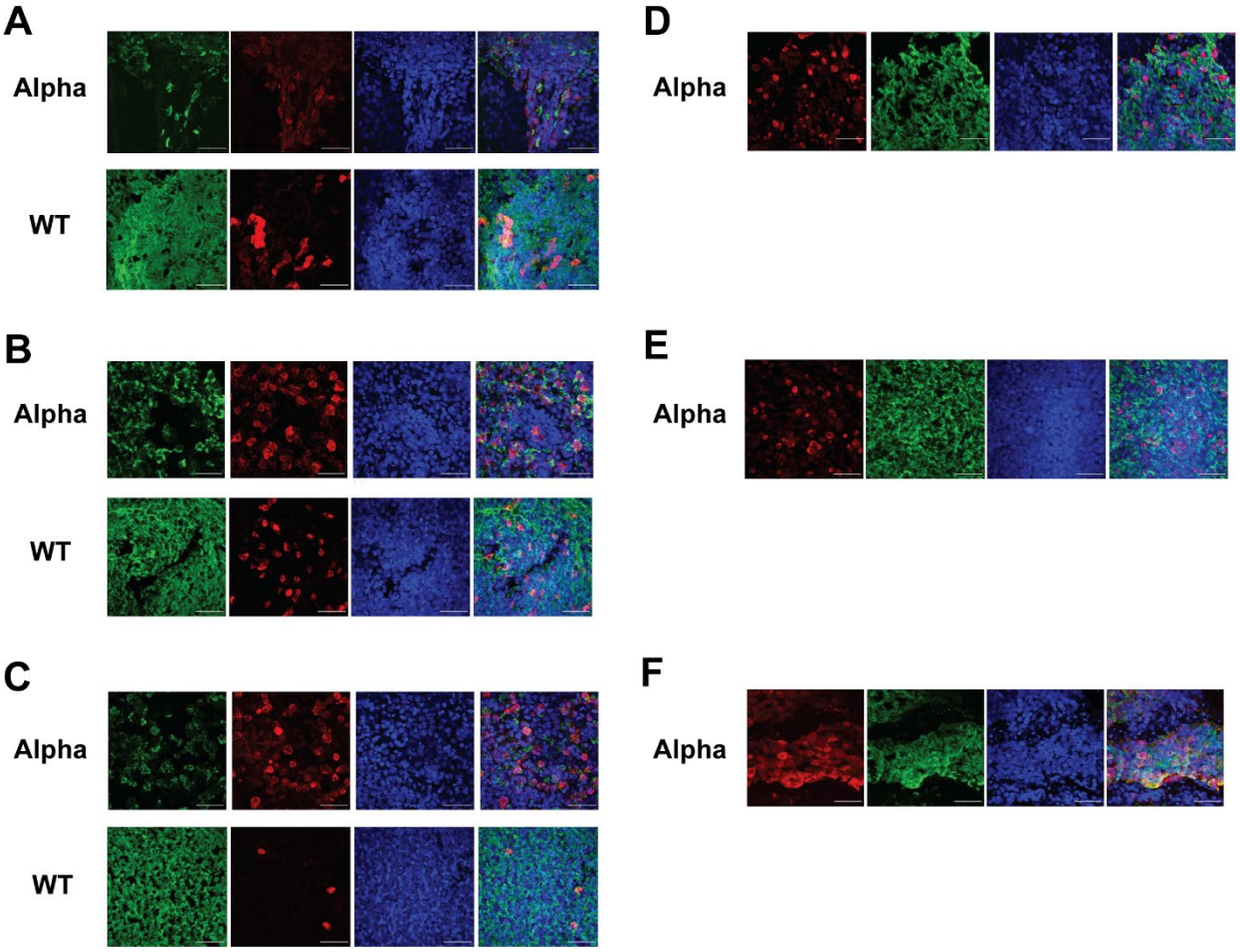


Figure S2. Immunofluorescent confocal microscopy staining at 40X magnification of ALI-HNE. Stains in adult donor cells infected with WT and Alpha in **A)** adult 1, **B)** adult 2, **C)** adult3 and child donors infected with Alpha in **D)** child 1, **E)** child 2 **and F)** child 3. α-tubulin (AcTub, green), nucleoprotein (NP, red) and nuclei (DAPI, blue).

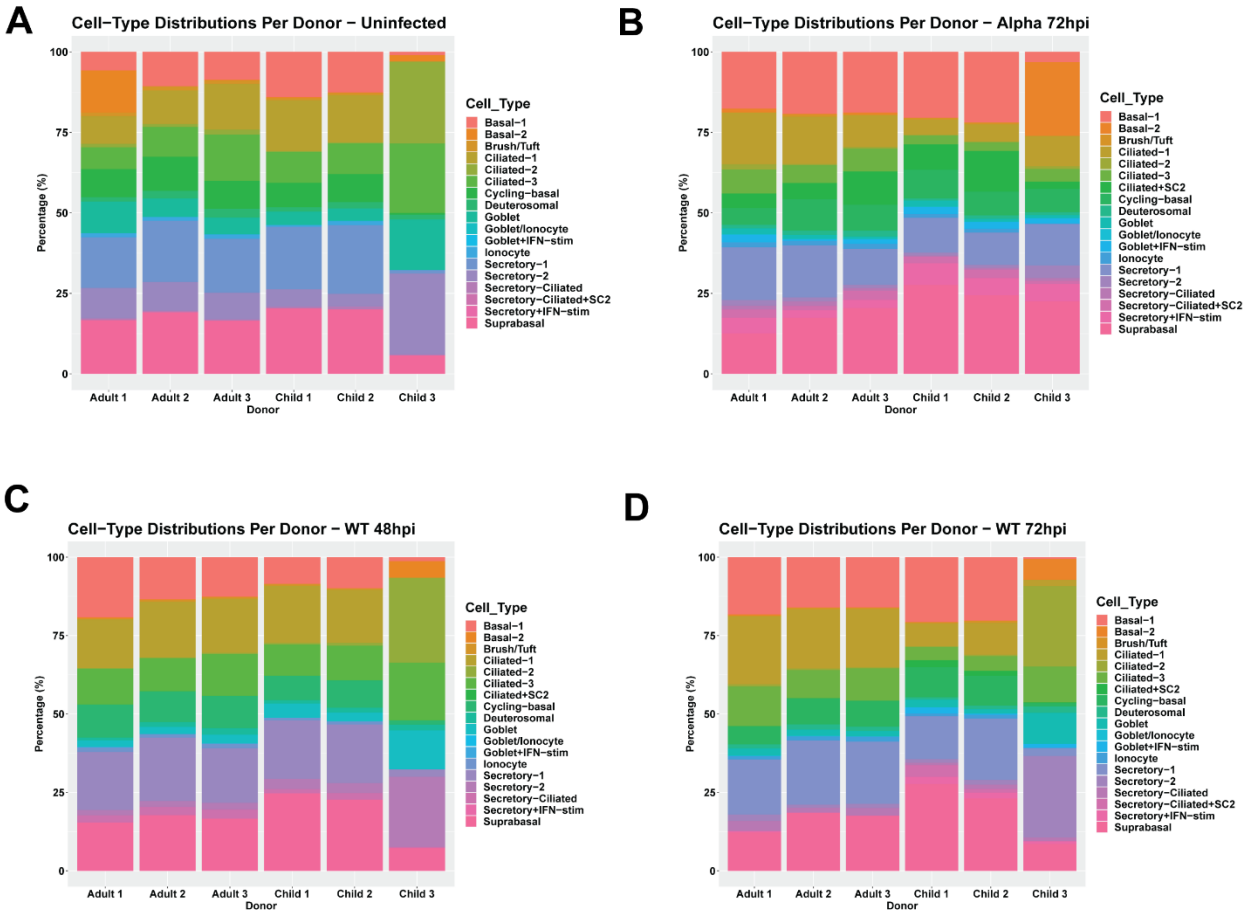


Figure S3. Cell-type distributions across donors per condition. A) Uninfected mock-control, **B)** Alpha-72hpi, **C)** WT-48hpi and **D)** WT-72hpi. X-axis indicates donors and Y-axis indicates the contributions of each cell-type in percentages.

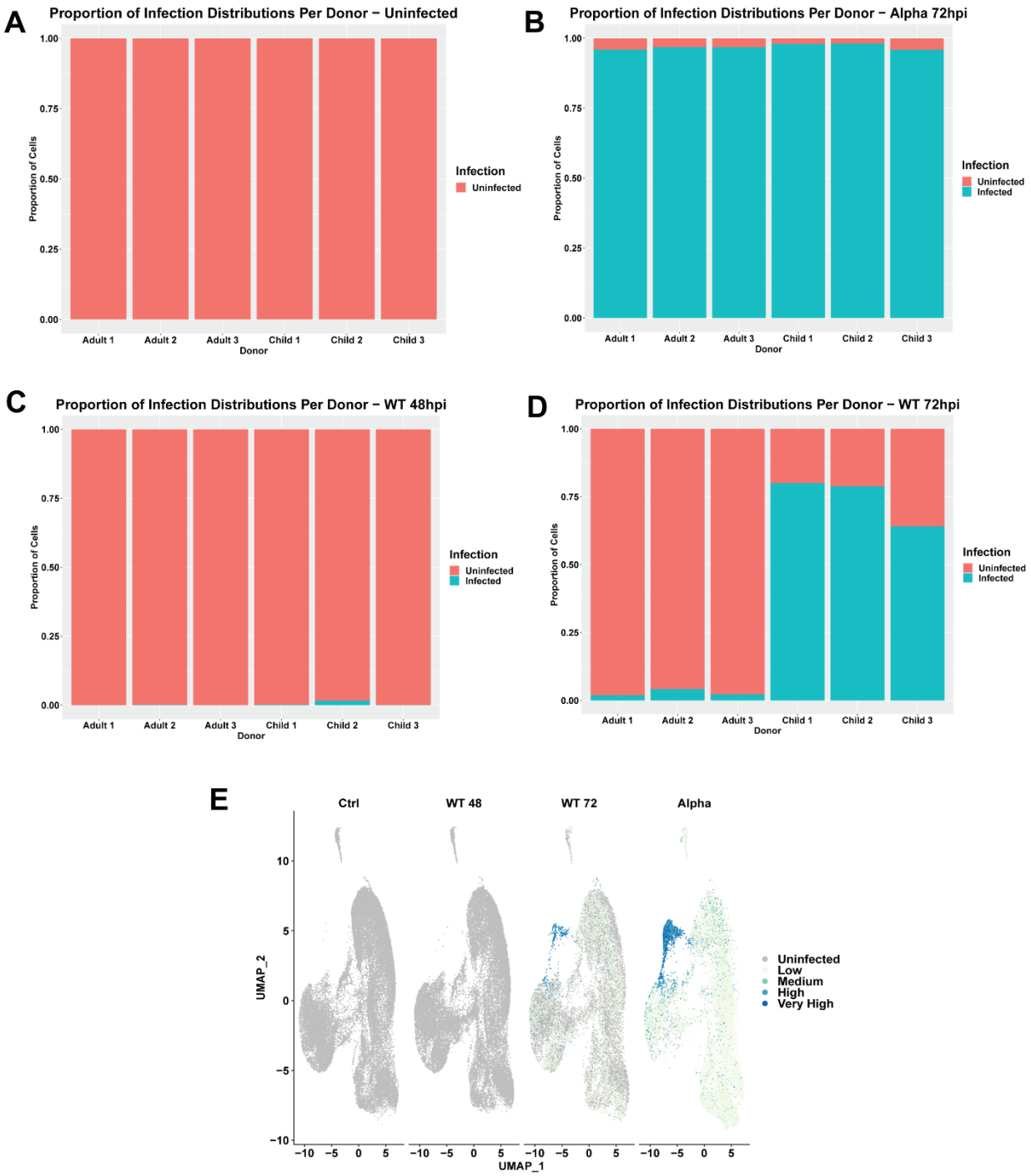


Figure S4. Infected vs uninfected cells. Proportion of infected vs uninfected cells per donor per condition in **A)** uninfected mock-control, **B)** Alpha-72hpi, **C)** WT-48hpi and **D)** WT-72hpi. X-axis indicates donors and Y-axis 6indicates the proportion of cells where red indicates proportion of infected cells and blue indicates proportion of uninfected cells. Alpha-infected datasets show highest proportion of infected cells compared with WT. **E)** UMAP plot of viral counts per cell, split based on mock-control, WT-infected or Alpha-infected datasets.

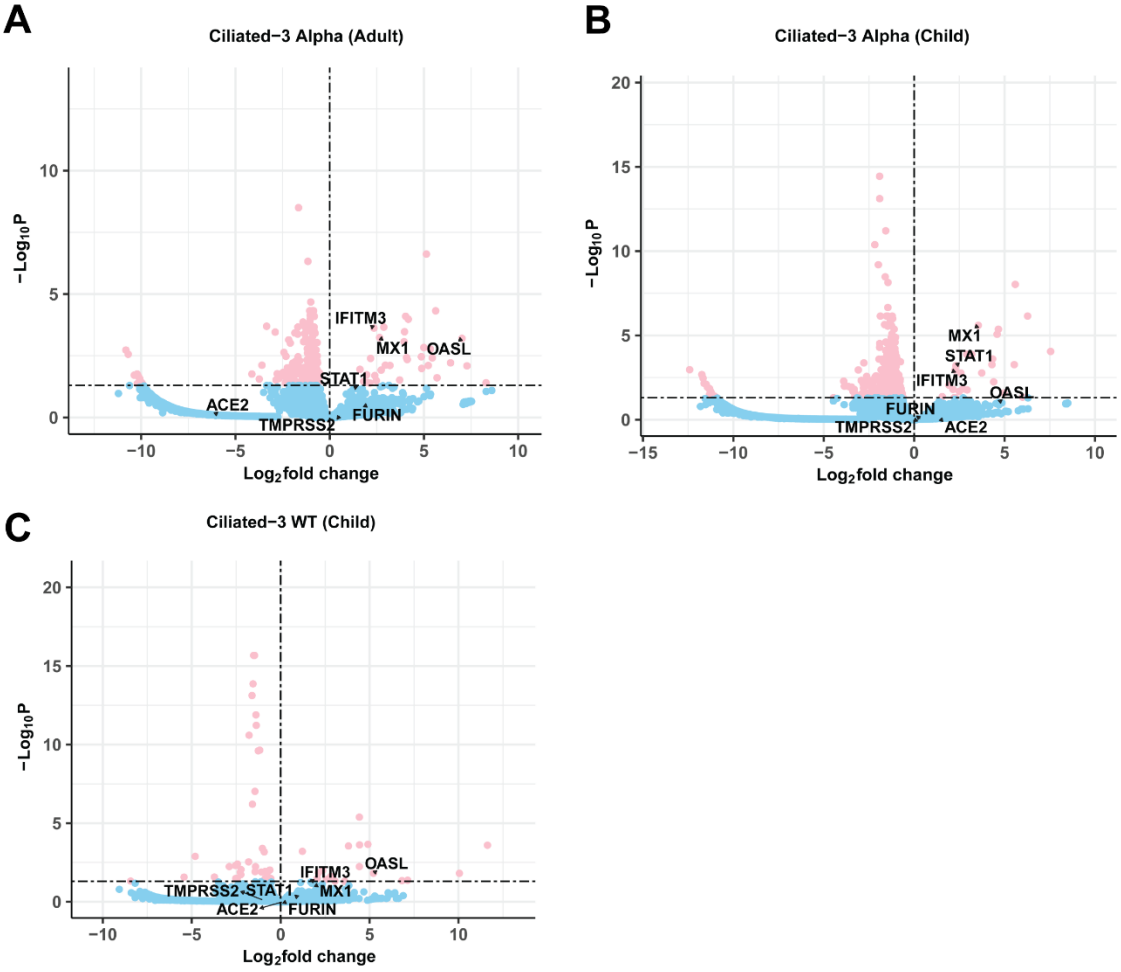


Figure S5. DGE results between bystander cells and mock-control (padj < 0.05). **A)** Ciliated-3 bystander cells exposed to Alpha SARS-CoV-2 (adults), **B)** Ciliated-3 bystander cells exposed to Alpha SARS-CoV-2 (child) and **C)** Ciliated-3 bystander cells exposed WT SARS-CoV-2 (child). X-axis shows the log2FC change between bystander and control cells and Y-axis shows the padj.

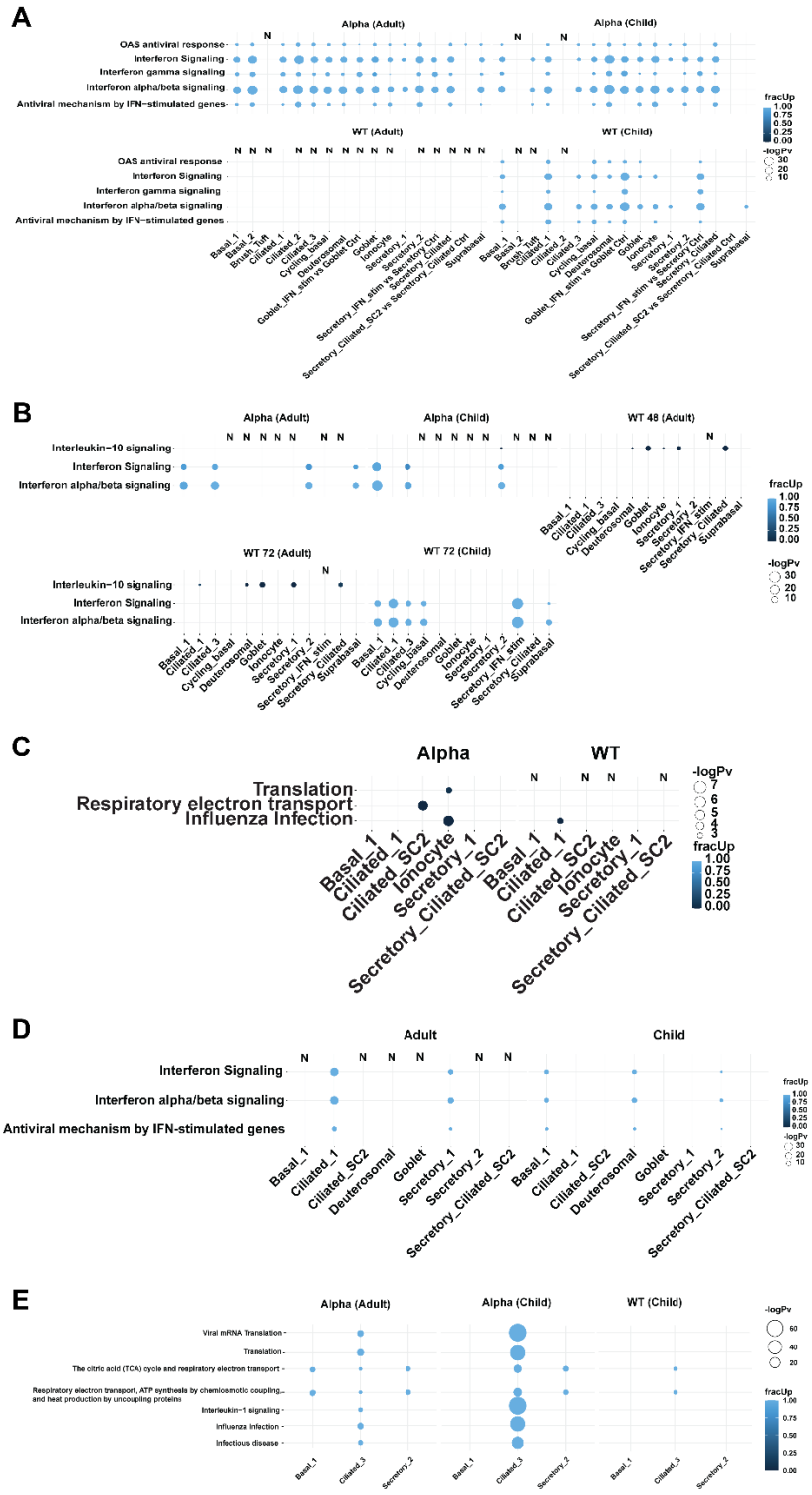


Figure S6. Significantly enriched reactome pathways analyzed using *multiGO* using significant DGE results (pv_thresh=0.05, enrichment pv_thresh=0.005 and logFC_thresh=1). **A)** Infected vs control (adolescents). Upregulated genes were involved in these processes. **B)** Bystander vs mock-control cells in ciliated and secretory-1 cells. **C)** Differences in Alpha-infected adolescents vs adults accounting for baseline expression in controls. **D)** Alpha vs WT infected adolescents in ciliated, secretory-ciliated-3 and secretory-ciliated-4 cells. **E)** Infected vs bystander cells. Bubble size indicates -log10 enrichment p-values, and the color of the bubble indicates the proportion of upregulated genes (i.e. fracUp).

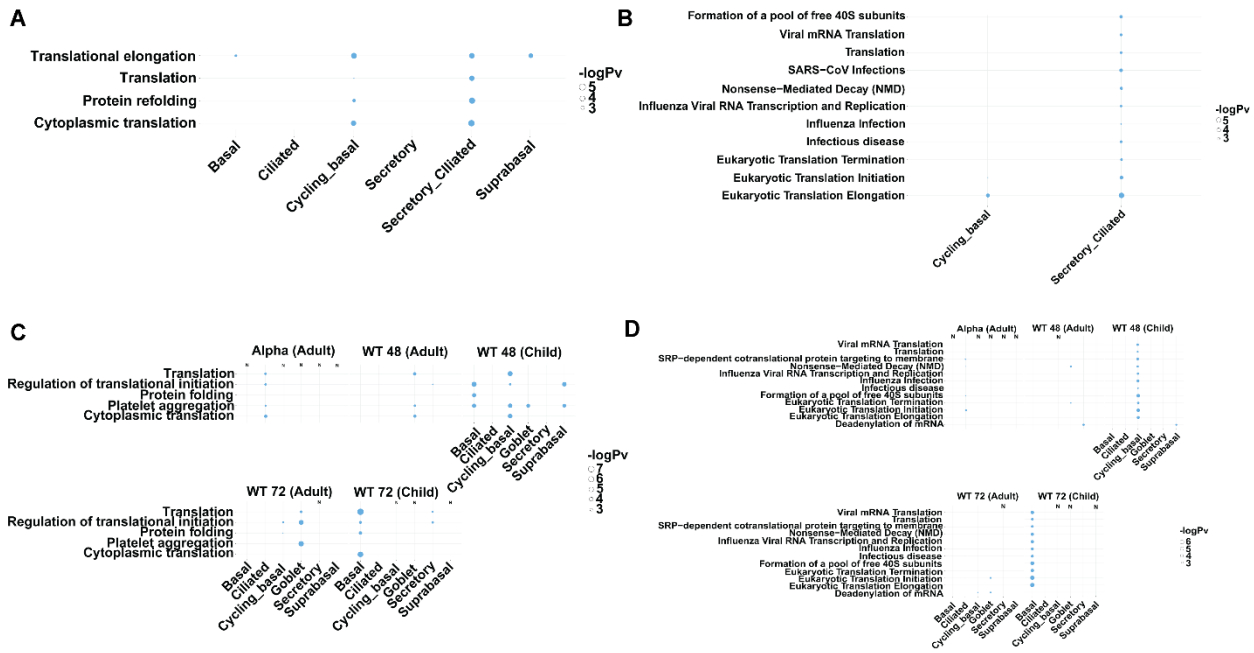


Figure S7. Significantly enriched reactome pathways analyzed using *multiGO* using significant DTU results (pv_thresh=0.05, enrichment pv_thresh=0.005 and logFC_thresh=1). **(A-B)** Alpha vs WT. **A)** GO biological terms and **B)** Reactome pathways. **(C-D)** Bystander vs control. **C)** GO biological terms and **D)** reactome pathways. Bubble size indicates -log10 enrichment p-values, color of the bubble is meaningless.

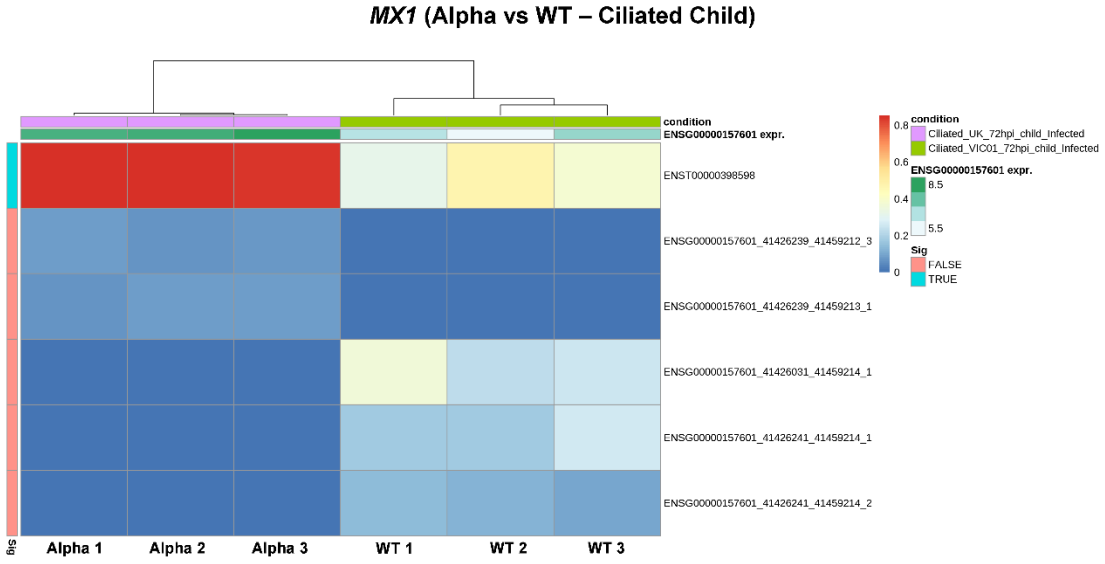


Figure S8. DTU analysis of *MX1* between Alpha vs WT-infected cells in child donors within Ciliated cells.

Supplementary Tables

Table S1. *multiGO* links to GO biology terms/reactome pathway enrichment analysis for DGE results.

multiGO analysis	**Type**	Link
Alpha vs WT	GO	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_edgeR_LRT_2/Alpha_vs_WT.zip&go_type=biological_process&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Alpha vs WT	Reactome	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_edgeR_LRT_2/Alpha_vs_WT.zip&go_type=ReactomePathways&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Child_vs_adult	GO	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_JAN_23_limma/limma_child_vs_adult_mixed_model.zip&go_type=biological_process&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Child_vs_adult	Reactome	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_JAN_23_limma/limma_child_vs_adult_mixed_model.zip&go_type=ReactomePathways&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Inf_vs_bystander	GO	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_edgeR_LRT_2/inf_vs_bystander.zip&go_type=biological_process&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35

Inf_vs_bystander	Reactome	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_edgeR_LRT_2/inf_vs_bystander.zip&go_type=ReactomePathways&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=100
Bystander_vs_control	GO	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_edgeR_LRT_2/bystander_vs_control.zip&go_type=biological_process&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Bystander_vs_control	Reactome	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_edgeR_LRT_2/bystander_vs_control.zip&go_type=ReactomePathways&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Inf_vs_uninf	GO	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_edgeR_LRT_2/inf_vs_uninf.zip&go_type=biological_process&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Inf_vs_uninf	Reactome	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_edgeR_LRT_2/inf_vs_uninf.zip&go_type=ReactomePathways&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Immune_profiles_LRT	GO	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_edgeR_LRT_2/immune_profiles_LRT.zip&go_type=biological_process&reorder=FALSE&go_thresh=0.005&pvthresh=0.005&reorder=FALSE&fc_thresh=1&max_go=35

Table S2. *multiGO* links to GO biology terms/reactome pathway enrichment analysis for DTU results.

multiGO analysis	**Type**	Link
DTU_inf_uninf	GO	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_DTU/inf_vs_uninf.zip&go_type=biological_process&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
DTU_inf_uninf	Reactome	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_DTU/inf_vs_uninf.zip&go_type=ReactomePathways&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
DTU_child_vs_adult	GO	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_DTU/child_vs_adult.zip&go_type=biological_process&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
DTU_child_vs_adult	Reactome	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_DTU/child_vs_adult.zip&go_type=ReactomePathways&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Uninfected	GO	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_DTU/uninfected.zip&go_type=biological_process&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Uninfected	Reactome	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_DTU/uninfected.zip&go_type=ReactomePathways&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Alpha vs WT	GO	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_DTU/Alpha_vs_WT.zip&go_type=biological_process&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Alpha vs WT	Reactome	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_DTU/Alpha_vs_WT.zip&go_type=ReactomePathways&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35

Bystander vs Control	GO	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_DTU/bystander_vs_control.zip&go_type=biological_process&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35
Bystander vs Control	Reactome	http://coinlab.mdhs.unimelb.edu.au/multigo2/?dir=multiGO_sc/recluster_Jan_23_DTU/bystander_vs_control.zip&go_type=ReactomePathways&reorder=FALSE&go_thresh=0.005&pvthresh=0.05&reorder=FALSE&fc_thresh=1&max_go=35

Table S3. Characteristics of each cell type with multiple sub-clusters.

Cell-type	Characteristics
Secretory-1	Ethanol oxidation Retinoic acid metabolic proc. Reg. of neural precursor cell proliferation
Secretory-2	Decreased: Ethanol oxidation Detoxification of copper ion Retinoic acid metabolic proc.
Secretory-3/IFN-stim	Neg. reg. of viral genome replication Type I interferon signaling pathway Cellular response to type I interferon
Basal-2	High *KRT14*;
Ciliated_1	CC phase: G1> G2M/S Epithelial cilium movement involved in extracellular fluid movement Extracellular transport Axoneme assembly
Ciliated-2	G2M/S > G1 phase (ciliated) PROLIFERATING, Reg. of cilium beat frequency Epithelial cilium movement involved in extracellular fluid movement Extracellular transport

Ciliated_3	G2M/S > G1 phase (low mitochondrial content) (ciliated) – PROLIFERATING Mitochondrial ATP synthesis coupled proton transport Purine ribonucleoside triphosphate metabolic proc. Oxidative phosphorylation

Table S4. Number and percentages of infected vs uninfected cells per cluster.

Clusters	Uninfected	Infected	Percent_inf
Ciliated+SC2	44	1881	97.7
Secretory-Ciliated+SC2	132	599	81.9
Secretory-Ciliated-1	1372	93	6.3
Secretory-3	2068	84	3.9
Deuterosomal	1797	44	2.4
Ciliated-1	13584	264	1.9
Secretory-1	20428	237	1.1
Ciliated-2	6177	71	1.1
Ionocyte	1137	9	0.8
Basal-2	3044	14	0.5
Suprabasal	20939	86	0.4
Secretory-2	8811	36	0.4
Ciliated-3	13177	45	0.3
Goblet	3155	9	0.3
Cycling-basal	10340	28	0.3
Basal-1	12757	28	0.2
Brush/Tuft	569	1	0.2

Table S5. Cell-clusters with increased *ACE2* levels with low SARS-CoV-2 infection compared with mock-control cells. Pseudo-bulked datasets (per combination of cell-type, condition, age and infection status) were compared using edgeR-LRT method with comparing groups by accounting for effect of sex information in the design matrix (padj < 0.05, see **Methods**).

Cell-cluster	Condition	Age	Log2FC	Padj
Secretory+IFN-stim	Alpha 72 hpi	Adult	2.58	2.52E-12
Secretory+IFN-stim	Alpha 72 hpi	Child	2.11	1.16E-8
Secretory+IFN-stim	WT 72 hpi	Child	1.41	0.006
Ciliated-1	Alpha 72 hpi	Adult	2.05	3.96E-5
Ciliated-1	Alpha 72 hpi	Child	1.35	0.036
Goblet+IFN-stim	Alpha 72 hpi	Adult	2.62	1.04E-6
Goblet+IFN-stim	Alpha 72 hpi	Child	2.79	1.78E-8
Goblet+IFN-stim	WT 72 hpi	Child	1.79	0.017

Table S6. Statistics on the ONT reads after running through *FLAMES* pipeline

Library	Total Fastq before cellbarcode matching	Fastq barcodehm match	Fastq fuzzy match	Fastq matching to barcode	Percentage of reads with matching barcode	Cells	Reads per cell after barcode	Reads per cell before barcode
LibA	67163409	34497993	5181472	39679465	59.07899196	3078	12891.31	21820.47
LibB	64471399	31780739	5464100	37244839	57.76955298	2692	13835.38	23949.26
LibC	61759028	28491395	3827382	32318777	52.33044957	7843	4120.716	7874.414
LibD	59548465	25781812	4214929	29996741	50.37365951	4642	6462.03	12828.19
LibE	62466754	26434866	4016441	30451307	48.74802203	3735	8152.96	16724.7
LibF	71703806	36681659	4096208	40777867	56.86987801	5515	7393.992	13001.6
LibG	53294978	24083354	3876707	27960061	52.46284368	3290	8498.499	16199.08
LibH	51054268	22150519	3858461	26008980	50.94379181	3078	8449.961	16586.83

Table S7. Common DTU genes in each dataset.

Infected vs Uninfected	Alpha vs WT	Bystander vs Control
GNAS	***SYPL1***	***GNAS***
EIF3L	***TYMP***	***EIF3L***
MSLN	***HEBP2***	***HSPB1***
IGFBP2	***ZC3H15***	***HMGB1***
EBNA1BP2	***PABPC1***	
F3	***PAFAH1B3***	
EPHA2	***HSP90AA1***	
ATP1B1	***FXYD3***	
COX5A	***COMT***	
HMGB1	***VTI1B***	
	SRSF5	
	MYL12A	
	EIF3E	
	EEF1D	
	KRT23	
	HSPA8	
	PTGES3	
	EEF1B2	
	IGFBP2	
	F3	
	PRDX6	
	RPS25	
	HDHD3	
	ENY2	
	RPL21	
	ARL4A	
	HNRNPR	
	SAT1	

NDUFA10
TXN
RPLP1
RTCA
ATRAID
EPHA2
ATP1B1
UFC1
SELENBP1
SNRNP200
RPL10
NAPRT
ITGB1
EEF1A1
MX1
PTMS
RSPH1
AP2M1
ELF3
RPL22L1
H2AZ1
AQP3
SPINT1
B2M
SRRM2
SLC3A2
RPL15
KRT5
CMC1
PTMA
HMGB1
ADH7
FAM3C

S100A6
HMGN2
KRT6A
ATP5MF
NCOA4

Table S8. DTU transcripts from bulk RNA-sequencing data of WT-strain-infected vs control in human cell lines Caco-2 and Calu-3 (Chang et al., 2022).

TRANSCRIPT NAME	ENSEMBL GENE ID	ENSEMBL TRANSCRIPT ID
CACO-2 2 HPI		
NACA-221	ENSG00000196531	ENST00000552540
SERF2-201	ENSG00000140264	ENST00000249786
CACO-2 48 HPI		
PKM-204	ENSG00000067225	ENST00000449901
RPL4-201	ENSG00000174444	ENST00000307961
RPL4-215	ENSG00000174444	ENST00000568588
CALU-3 2 HPI		
ATP13A3-201	ENSG00000133657	ENST00000429136
STIL-201	ENSG00000123473	ENST00000337817
STIL-202	ENSG00000123473	ENST00000360380
TRA2A-202	ENSG00000164548	ENST00000392502
TRA2A-211	ENSG00000164548	ENST00000538367
CALU-3 48 HPI		
SPTBN1-207	ENSG00000115306	ENST00000615901
GSDMB-208	ENSG00000073605	ENST00000477054
KPNA2-201	ENSG00000182481	ENST00000330459

KPNA2-202	ENSG00000182481	ENST00000537025
RAE1-202	ENSG00000101146	ENST00000395840
RAE1-203	ENSG00000101146	ENST00000395841
ATIC-205	ENSG00000138363	ENST00000435675
ATIC-207	ENSG00000138363	ENST00000443953
ADK-202	ENSG00000156110	ENST00000372734
ADK-206	ENSG00000156110	ENST00000541550
ATRAID-201	ENSG00000138085	ENST00000380171
ATRAID-207	ENSG00000138085	ENST00000606999
ANKRD12-201	ENSG00000101745	ENST00000262126
ANKRD12-203	ENSG00000101745	ENST00000400020
WARS1-202	ENSG00000140105	ENST00000355338
WARS1-204	ENSG00000140105	ENST00000392882