Cell-intrinsic differences between human airway epithelial cells from children and adults

Elizabeth F. Maughan[1,2], Ersilia Nigro[1]*, Adam Pennycuick[1]*, Kate H.C. Gowers[1], Celine Denais[1], Sandra Gómez-López[1], Kyren A. Lazarus[1], Colin R. Butler[1,2,3], Dani Do Hyang Lee[4], Jessica C. Orr[1], Vitor H. Teixeira[1], Benjamin E. Hartley[3], Richard J. Hewitt[3], Chadwan Al Yaghchi[5], Gurpreet S. Sandhu[5], Martin A. Birchall[6], Christopher O'Callaghan[4], Claire M. Smith[4], Paolo De Coppi[2], Robert E. Hynds[1,7#] and Sam M. Janes[1#]

(1) Lungs for Living Research Centre, UCL Respiratory, University College London, London, U.K.
(2) Stem Cell and Regenerative Medicine Section, University College London Great Ormond Street Institute of Child Health, University College London, London, U.K.
(3) Tracheal Service, Great Ormond Street Hospital, London, U.K.
(4) Respiratory, Critical Care & Anaesthesia, University College London Great Ormond Street Institute of Child Health, University College London, London, United Kingdom.
(5) The National Centre for Airway Reconstruction, Department of Otolaryngology, Charing Cross Hospital, London, U.K.
(6) University College London Ear Institute, University College London, London, U.K.
(7) University College London Cancer Institute, University College London, London, U.K.

* These authors contributed equally.

 # Corresponding Authors
Dr. Robert E. Hynds (rob.hynds@ucl.ac.uk)
UCL Respiratory, University College London, 5 University Street, London, WC1E 6JF, U.K.

Professor Sam M. Janes (s.janes@ucl.ac.uk)
Lungs for Living Research Centre, UCL Respiratory, University College London, 5 University Street, London, WC1E 6JF, U.K.

Keywords: epithelial cells, epithelium, respiratory epithelium, biological ageing, SARS-CoV-2, COVID-19

Running title: Age-related expression changes in airway epithelium

Summary

The airway epithelium is a key protective barrier whose integrity is preserved by the self-renewal and differentiation of basal progenitor cells. Epithelial cells are central to the pathogenesis of multiple lung diseases. In chronic diseases, increasing age is a principle risk factor. In acute diseases, such as COVID-19, children suffer less severe symptoms than adults and have a lower rate of mortality. Few studies have explored differences between airway epithelial cells in children and adults to explain this age dependent variation in diseases. Here, we perform bulk RNA sequencing studies in laser-capture microdissected whole epithelium, FACS-sorted basal cells and cultured basal cells, as well as *in vitro* cell proliferation experiments, to address the intrinsic molecular differences between paediatric and adult airway basal cells. We find that, while the cellular composition of the paediatric and adult tracheobronchial epithelium is broadly similar, in cell culture, paediatric airway epithelial cells displayed higher colony forming ability, better *in vitro* growth and outcompeted adult cells in competitive proliferation assays. In RNA sequencing experiments, we observed potentially important differences in airway epithelial gene expression between samples from children and adults. However, genes known to be associated with SARS-CoV-2 infection were not differentially expressed between children and adults. Our results chart cell-intrinsic differences in transcriptional profile and regenerative capacity between proximal airway epithelial cells of children and adults.

Introduction

The human airways are lined by a pseudostratified epithelium from the trachea through most of the 16 generations of conducting airway branching. Functionally, the epithelium secretes a mucous layer and produces motile force to flow mucus proximally out of the lungs, providing protection against noxious particles and pathogens. These specialized functions are accomplished by luminal mucosecretory and ciliated epithelial cells, respectively. The differentiated cell types in this slow-turnover tissue are replenished by airway basal cells, which act as multipotent progenitors (Rock et al., 2010; Teixeira et al., 2013).

The cellular composition of the airway is being mapped in ever more detail through single cell RNA sequencing approaches. A novel CFTR-producing population, the ionocyte (Montoro et al., 2018; Plasschaert et al., 2018), has been described, and we have new insight into cellular differentiation through the resolution of differentiation intermediates; for example, basal luminal progenitor cells (Mori et al., 2015; Rock et al., 2011; Watson et al., 2015), which are defined by an intermediate keratin profile (TP63-/KRT5+/KRT8+/KRT4+/KRT13+) (Braga et al., 2019; García et al., 2019). Efforts in mapping pathologic states have identified mucous ciliated cells (ciliated cells which co-express a number of genes typically associated with goblet cells and are more frequent in asthmatic patients (Braga et al., 2019)), as well as novel pathological epithelial cell subtypes in idiopathic pulmonary fibrosis (Adams et al., 2019; Habermann et al., 2019; Reyfman et al., 2019). In addition, region-specific differences exist in airway epithelial cell phenotype within the bronchial tree such that basal, ciliated and secretory populations in the nasal epithelium differ phenotypically (and maybe functionally) from their counterparts in the large or small airways (Braga et al., 2019; Deprez et al., 2019; Kumar et al., 2011; Travaglini et al., 2019). However, despite well-characterized structural and functional consequences of ageing in the distal lung (Navarro and Driscoll, 2017; Thurlbeck and Angus, 1975; Turner et al., 1968), little is known about alterations in human airway epithelial cell composition or function between paediatric and adult tissue. The importance of age-related epithelial variation has been highlighted by the COVID-19 pandemic, where functional differences in airway predisposition to viral infection and response have had a major clinical impact.

Single cell RNA sequencing studies have shown that increased transcriptional noise and upregulation of a core group of age-associated molecular mechanisms – including protein

processing- and inflammation-associated genes – are correlated with ageing across mouse cell and tissue types, but additional processes are unique to particular cell types within specific organs, including the lungs (Angelidis et al., 2019; Kimmel et al., 2019). However, to date such studies have not profiled the trachea, which has distinct composition and stem cell biology to the more distal airways (Basil et al., 2020), in detail. During murine tracheal ageing, epithelial cell density is reduced and the proportion of basal cells within the epithelium is slightly decreased, but there is no obvious decline in their *in vitro* clonogenic potential or differentiation capacity (Wansleeben et al., 2014). However, microarray gene expression analysis shows changes consistent with the development of low-grade chronic inflammation in older mice, together with an increased presence of activated adaptive immune cells (Wansleeben et al., 2014).

There are striking differences in airway structure and composition between rodents and humans (Hogan et al., 2014), which suggests that the effects of ageing on airway epithelial regeneration might differ substantially between species. Careful characterization of human paediatric and adult airway epithelial cell composition and function will both inform lung regenerative medicine efforts and potential pathogenic mechanisms behind multiple chronic lung disease pathologies (Kicic et al., 2006; Prasse et al., 2018; Staudt et al., 2014), as age is a risk factor for COPD, pulmonary fibrosis, infection and lung cancer. In all such cases, basal stem cell dysfunction, perhaps accelerated by smoking, is likely to play a role in disease pathogenesis (Meiners et al., 2015).

Here, we compare paediatric and adult human tracheobronchial epithelia in terms of cellular composition, gene expression profiles and behaviour in primary cell culture. In light of the on-going COVID-19 pandemic, we compare expression of epithelial viral infection-associated genes, including those so far linked to SARS-CoV-2 infection, and make our data available as a resource for the community.

Results

Human tracheobronchial epithelium has comparable cellular composition in children and adults

In mouse trachea, the proportion of the epithelium that expresses the basal cell-associated protein keratin 5 (KRT5) decreases with age (Wansleeben et al., 2014), so we first compared the cellular composition of steady state normal human airway epithelium using haematoxylin and eosin (H&E) staining and immunohistochemistry for TP63 (basal cells), MUC5AC (mucosecretory cells) and FOXJ1 (ciliated cells) in tracheobronchial biopsies (Figure 1A/1B; donor characteristics are listed in Table S1). We found no significant differences in the proportion of cells in these three cellular compartments in paediatric and adult biopsies either by immunohistochemistry (Figure 1A/1B), or by assessing basal, mucosecretory or ciliated cell-associated gene expression (Table S2) in bulk RNA sequencing in which we had laser-capture microdissected the whole epithelium (Figure 1C; Figure S1).

Analysing this laser-capture microdissected whole epithelium RNA sequencing dataset using DESeq2 (Love et al., 2014) with a false discovery rate (FDR) of 1% and $\log_2$ fold change threshold of 1.2, we identified 37 genes with significant differential expression between paediatric and adult donors of which 17 were upregulated in adults and 20 were expressed at higher levels in children (Figure 2A; Table S3). To determine alterations in biologically functional gene groups, we performed gene set enrichment analysis (GSEA) using the Hallmark gene sets from MSigDB (Liberzon et al., 2015; Subramanian et al., 2005). In the paediatric airway epithelium, this demonstrated a higher expression of genes associated with interferon alpha and gamma responses. In adults, there was a higher expression of genes associated with TP53, TGFβ and Wnt-β-catenin signalling, as well as processes such as epithelial-mesenchymal transition and glycolysis (Figure 2B). Although broad transcriptional similarities are expected within the same tissue between childhood and adulthood, these differences in gene expression may be functionally important.

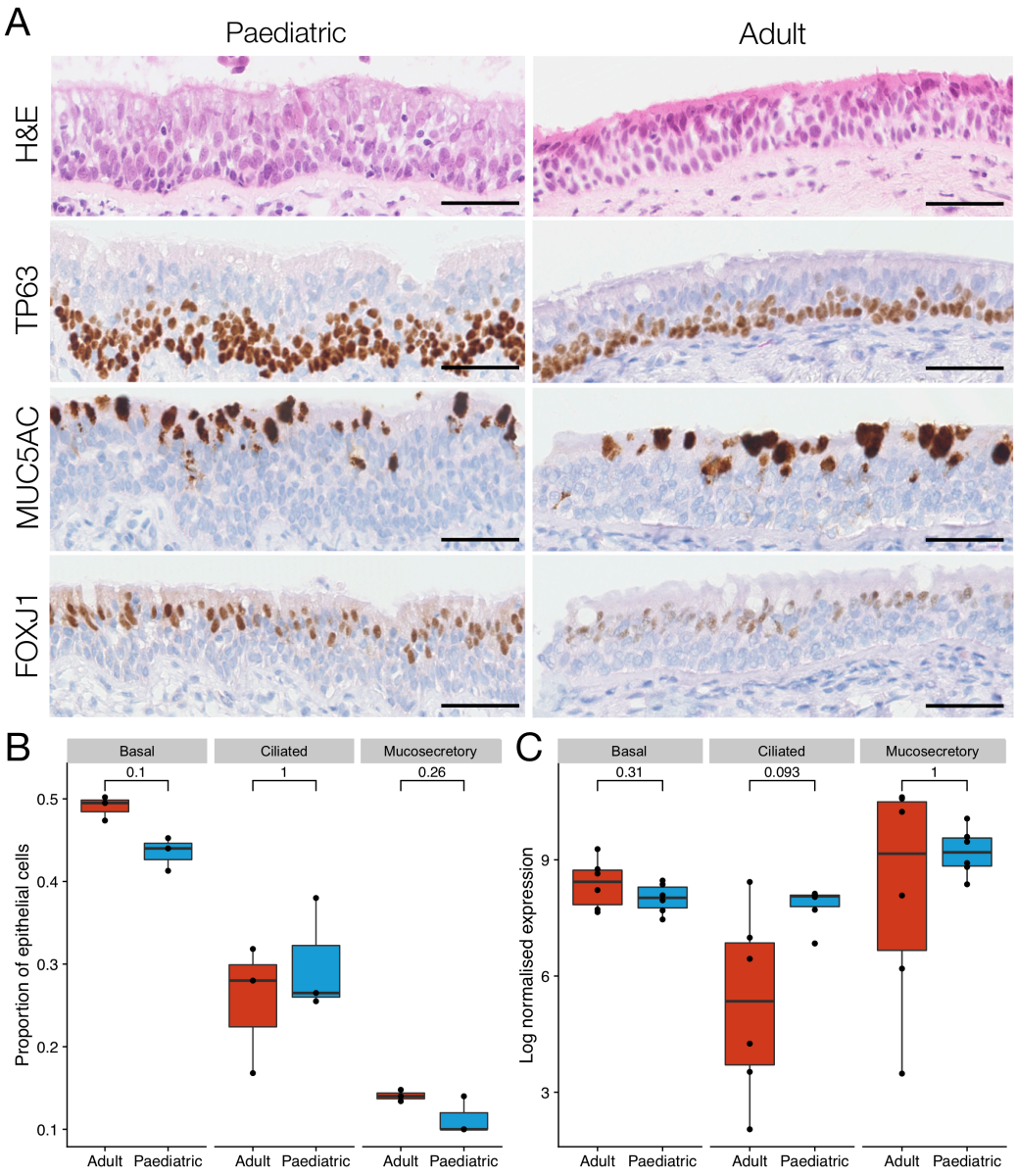

Figure 1: Cellular composition of paediatric and adult human tracheobronchial epithelium.

(A) Representative haematoxylin & eosin staining and immunohistochemistry comparison of TP63, MUC5AC and FOXJ1 protein expression in paediatric and adult tracheobronchial epithelium. Scale bars = 50 µm. **(B)** Quantification of TP63$^+$ basal cells, MUC5AC$^+$ mucosecretory cells and FOXJ1$^+$ ciliated cells in paediatric and adult tracheobronchial epithelium. Results are shown as a proportion of total cells within the epithelium (12,568 total cells for TP63, 9,651 for MUC5AC and 16,144 for FOXJ1). No significant differences were seen in a two-sided Wilcoxon rank sum test (n = 3 donors/age group; basal cells, p = 0.1; ciliated cells, p = 1; mucosecretory cells, p = 0.26). **(C)** Expression of basal, ciliated and mucosecretory cell markers in RNA sequencing data from paediatric and adult epithelium. For each sample, the geometric mean of a set of gene markers associated with a given cell type (Table S2) is shown. No significant differences were seen in a two-sided Wilcoxon rank sum test (n = 3 donors/age group; basal cells p = 0.31; ciliated cells p = 0.09, mucosecretory cells p = 1).

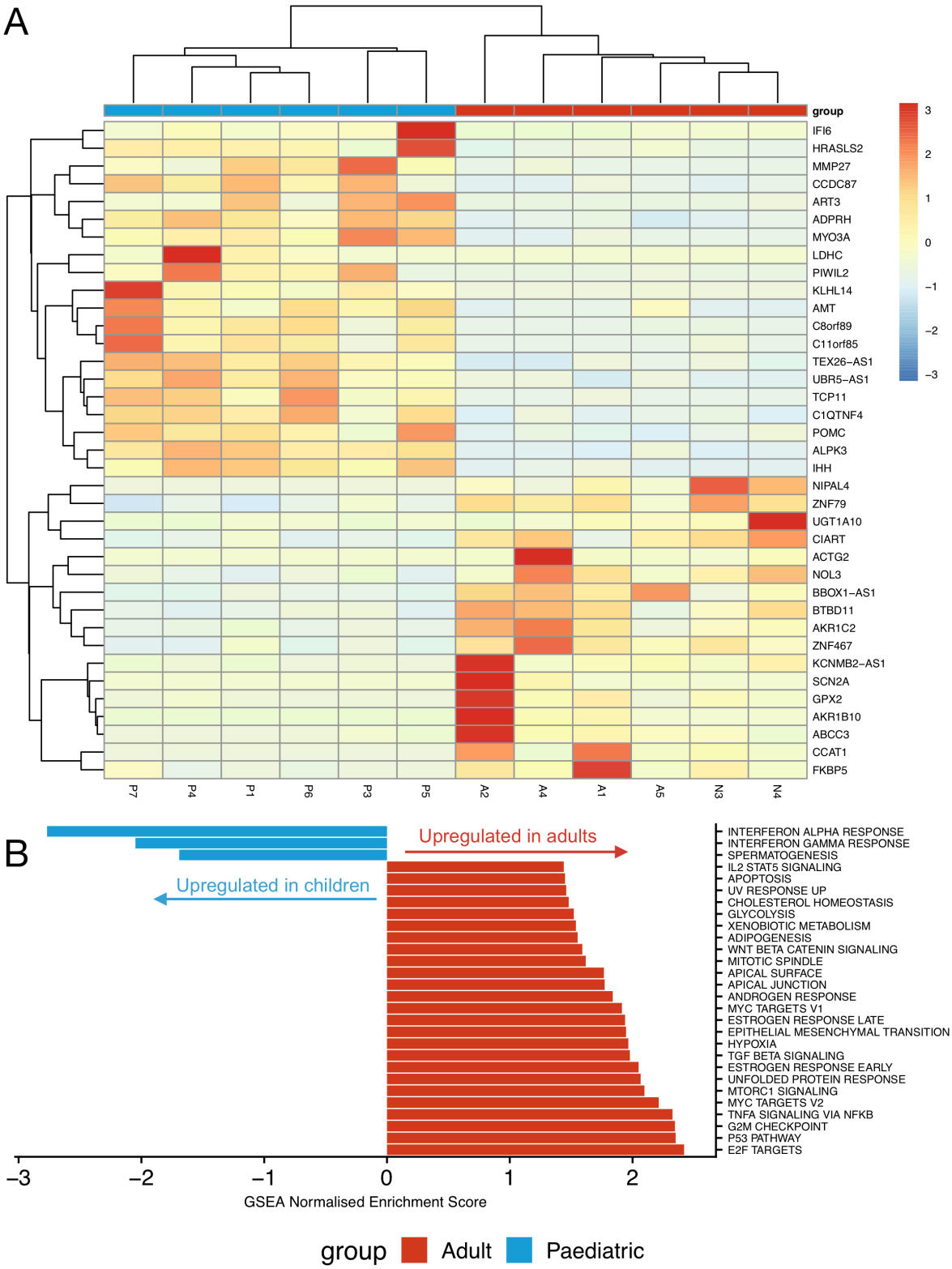
A
group
IFI6
HRASLS2
MMP27
CCDC87
ART3
ADPRH
MYO3A
LDHC
PIWIL2
KLHL14
AMT
C8orf89
C11orf85
TEX26−AS1
UBR5−AS1
TCP11
C1QTNF4
POMC
ALPK3
IHH
NIPAL4
ZNF79
UGT1A10
CIART
ACTG2
NOL3
BBOX1−AS1
BTBD11
AKR1C2
ZNF467
KCNMB2−AS1
SCN2A
GPX2
AKR1B10
ABCC3
CCAT1
FKBP5
P7
P4
P1
P6
P3
P5
A2
A4
A1
A5
N3
N4
3
2
1
0
−1
−2
−3
B
Upregulated in adults
Upregulated in children
INTERFERON ALPHA RESPONSE
INTERFERON GAMMA RESPONSE
SPERMATOGENESIS
IL2 STAT5 SIGNALING
APOPTOSIS
UV RESPONSE UP
CHOLESTEROL HOMEOSTASIS
GLYCOLYSIS
XENOBIOTIC METABOLISM
ADIPOGENESIS
WNT BETA CATENIN SIGNALING
MITOTIC SPINDLE
APICAL SURFACE
APICAL JUNCTION
ANDROGEN RESPONSE
MYC TARGETS V1
ESTROGEN RESPONSE LATE
EPITHELIAL MESENCHYMAL TRANSITION
HYPOXIA
TGF BETA SIGNALING
ESTROGEN RESPONSE EARLY
UNFOLDED PROTEIN RESPONSE
MTORC1 SIGNALING
MYC TARGETS V2
TNFA SIGNALING VIA NFKB
G2M CHECKPOINT
P53 PATHWAY
E2F TARGETS
−3
−2
−1
0
1
2
GSEA Normalised Enrichment Score
group
Adult
Paediatric

Figure 2: RNA sequencing of laser-capture microdissected whole epithelium from paediatric and adult proximal airways.

(A) Cluster diagram showing the normalized expression of all 37 genes differentially expressed with a false discovery rate (FDR) < 0.01 and a $\log_2$ fold change > 1.2 in six paediatric (months of age/sex; 12M, 14M, 18M, 41M, 83M, 106M) and six adult (years of age/sex; 58F, 60F, 63M, 68F, 69M, 72M) laser-capture microdissected tracheobronchial epithelial samples. Values are scaled by row. Gene order is based on hierarchical clustering based on the similarity in overall expression patterns. Red represents relative expression higher than the median expression and blue represents lower expression. **(B)** Pathway analysis was performed on the same paediatric and adult laser-capture microdissected tracheobronchial epithelial samples using gene set enrichment analysis (GSEA) to interrogate Hallmark pathways from MSigDB. For pathways with FDR < 0.05, normalized enrichment scores are shown. A negative score (blue) represents upregulation of the pathway in the paediatric samples; a positive score (red) represents upregulation in the adult samples.

Basal cells from children proliferate more readily in primary cell culture than those from adults

To investigate possible differences in regenerative potential between paediatric and adult proximal airway basal cells, we established primary cell cultures. First, we FACS-sorted single basal cells, identified by dual EpCAM (epithelial) and podoplanin (PDPN; basal (Miller et al., 2018; Weeden et al., 2017)) positivity, into individual wells of 96-well plates (range = 167 to 192 cells per donor) to compare the clonal potential of native paediatric and adult basal cells. After 7 days of culture in epithelial cell culture medium containing Y-27632, colony formation was significantly higher among basal cells derived from children than adults (Figure 3A), consistent with our previous work (Yoshida et al., 2020). At this timepoint, paediatric basal cells had often generated colonies that had become confluent to fill the well, whereas no adult colonies reached confluence (Figure 3B/3C).

When cells were isolated and cultured in epithelial cell culture medium without Y-27632 on 3T3-J2 feeder layers, expansion of paediatric and adult cells proceeded similarly at early passages but a growth advantage was observed in paediatric donors after 5 passages (Figure 3D). In MTT assays, cultured paediatric cells showed greater proliferation than adult cells after 5 and 7 days (Figure 3E). Likewise, differences in EdU uptake (Figure 3F) and Ki67-positivity (Figure 3G) were seen in passage 1 cell cultures in the presence of feeder cells. When cultured basal cells were assessed in colony formation assays, there was a trend towards paediatric cells forming more colonies than adult cells (Figure 3H).

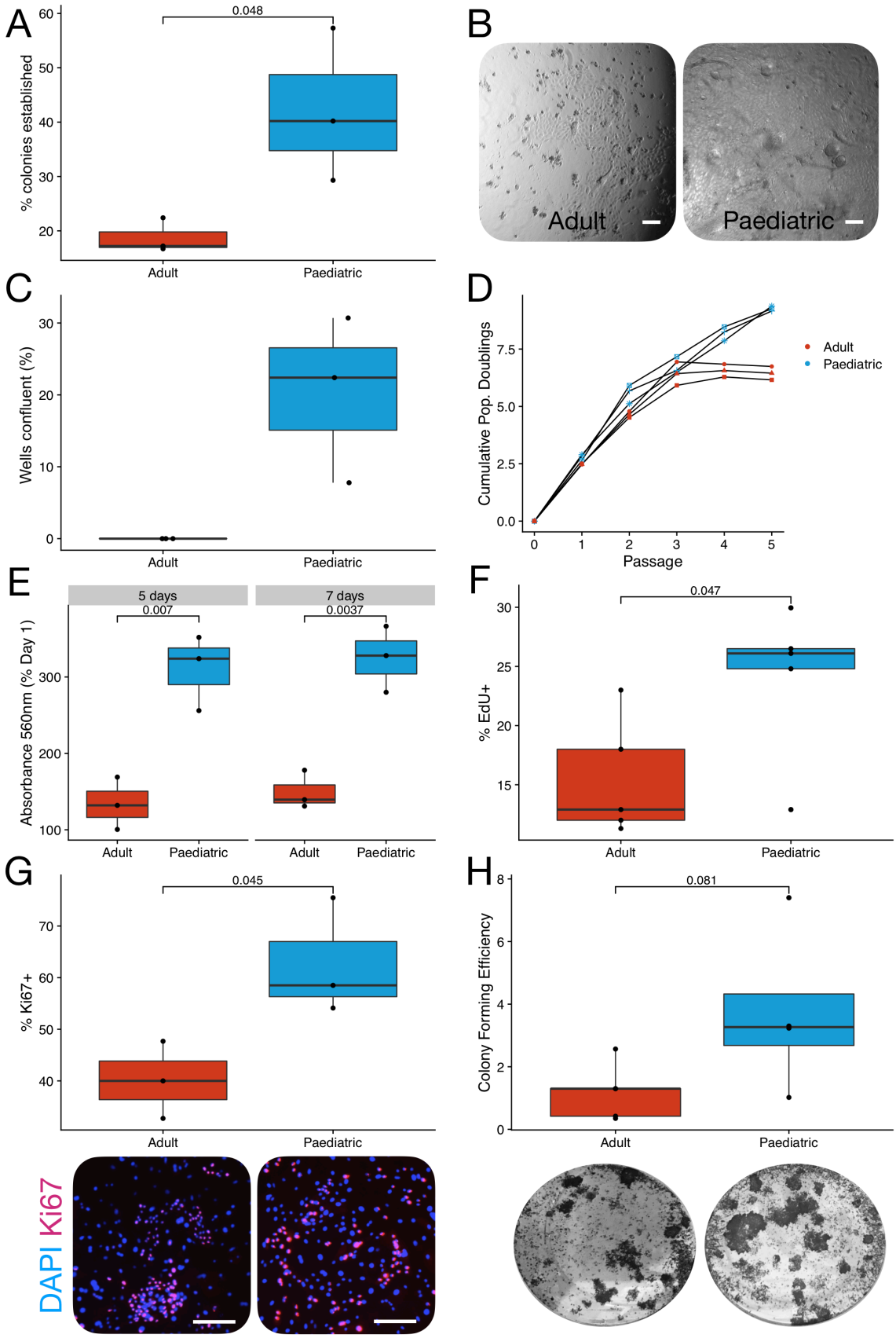
A
% colonies established
0.048
Adult
Paediatric
B
Adult
Paediatric
C
Wells confluent (%)
Adult
Paediatric
D
Cumulative Pop. Doublings
Passage
Adult
Paediatric
E
5 days
7 days
0.007
0.0037
Absorbance 560nm (% Day 1)
Adult
Paediatric
F
% EdU+
0.047
Adult
Paediatric
G
% Ki67+
0.045
Adult
Paediatric
DAPI Ki67
H
Colony Forming Efficiency
0.081
Adult
Paediatric

Figure 3: Cultured airway basal cells show cell-intrinsic differences in cell proliferation.

(A) Proportion of $EpCAM^+/PDPN^+$ basal cells that formed colonies within 7 days after sorting single cells into 96-well plates in epithelial cell culture medium containing Y-27632 (n = 3 donors/age group; n = 167-192 cells per donor; p = 0.048, two-tailed unpaired t-test). **(B)** Brightfield microscopy images showing a confluent paediatric well and an adult colony after 7 days of culture. Scale bar = 50 µm. **(C)** Quantification of wells that reached confluence in 96-well plates after 14 days of culture (n = 3 donors/age group). No wells achieved confluence in the adult group. **(D)** Population doubling analysis of the growth of paediatric and adult basal cell cultures in epithelial growth medium without Y-27632. Paediatric samples outperformed adults at passage 5 ($p = 9.8x10^{-5}$, two-tailed unpaired t-test). **(E)** MTT assay assessing the growth of paediatric and adult basal cell cultures in epithelial growth medium without Y-27632 and in the absence of feeder cells. Data are normalized to readings taken at day 1. Cells were P1. 5 days (n = 3 donors/age group; p = 0.007, two-tailed unpaired t-test) and 7 days (n = 3 donors/age group; p = 0.0037, two-tailed unpaired t-test). **(F)** Flow cytometric analysis of EdU uptake in passage 1 primary human airway epithelial cells from paediatric and adult donors grown in epithelial cell culture medium without Y-27632 at approximately 70% confluence (n = 3 donors/age group; p = 0.047, two-tailed unpaired t-test). **(G)** Ki67 immunofluorescence staining of passage 1 paediatric and adult basal cells cultured on feeder layers in epithelial cell culture medium without Y-27632 for 3 days before fixation. Ki67 positivity was quantified as a proportion of positive nuclei/total DAPI stained nuclei (n=3 donors/age group; mean of 1402 cells counted per donor, range 810 – 1600; p = 0.045, two-tailed unpaired t-test). **(H)** Colony formation assays comparing early passage (P1 or P2) cultured paediatric and adult basal cells (n = 4 paediatric donors, 5 adult donors; p = 0.81, two-tailed unpaired t-test).

Basal cells from children outcompete those from adults in mixed cultures

To better understand progenitor capacity, we next developed a competitive proliferation assay (Figure S2), using lentiviral cell labelling with fluorescent constructs (Eekels et al., 2012). After optimization in 293T cells to ensure that the two lentiviruses did not affect cell growth (Figure S3), we isolated and cultured patient basal cells in epithelial cell culture medium containing Y-27632 to facilitate lentiviral transduction (Horani et al., 2013) and transduced these with either with green fluorescent protein (GFP)- or mCherry-expressing lentiviral constructs (Figure 4A). When combining GFP^+ and $mCherry^+$ cells from the same donor in equal number, the ratio remained 1:1 over 7 days (Figure S4). After transducing three paediatric and four adult basal cell cultures in this manner, we combined each paediatric donor with each adult donor in both possible colour combinations so that we could monitor the growth dynamics of the two populations separately using fluorescence (Figure 4B/4C). There were no differences in lentiviral integration as determined by PCR targeting the puromycin resistance gene contained within both GFP and mCherry lentiviral vectors (Figure S4E). When cells were harvested at approximately 80% confluence, the growth differential between paediatric cells and adult cells was calculated for each paediatric/adult pair. In almost all donor pairs, the paediatric cells outgrew the adult cells (Figure 4D). Interestingly, the pairings that did not follow this pattern involved the youngest adult donor, who was 30 years of age.

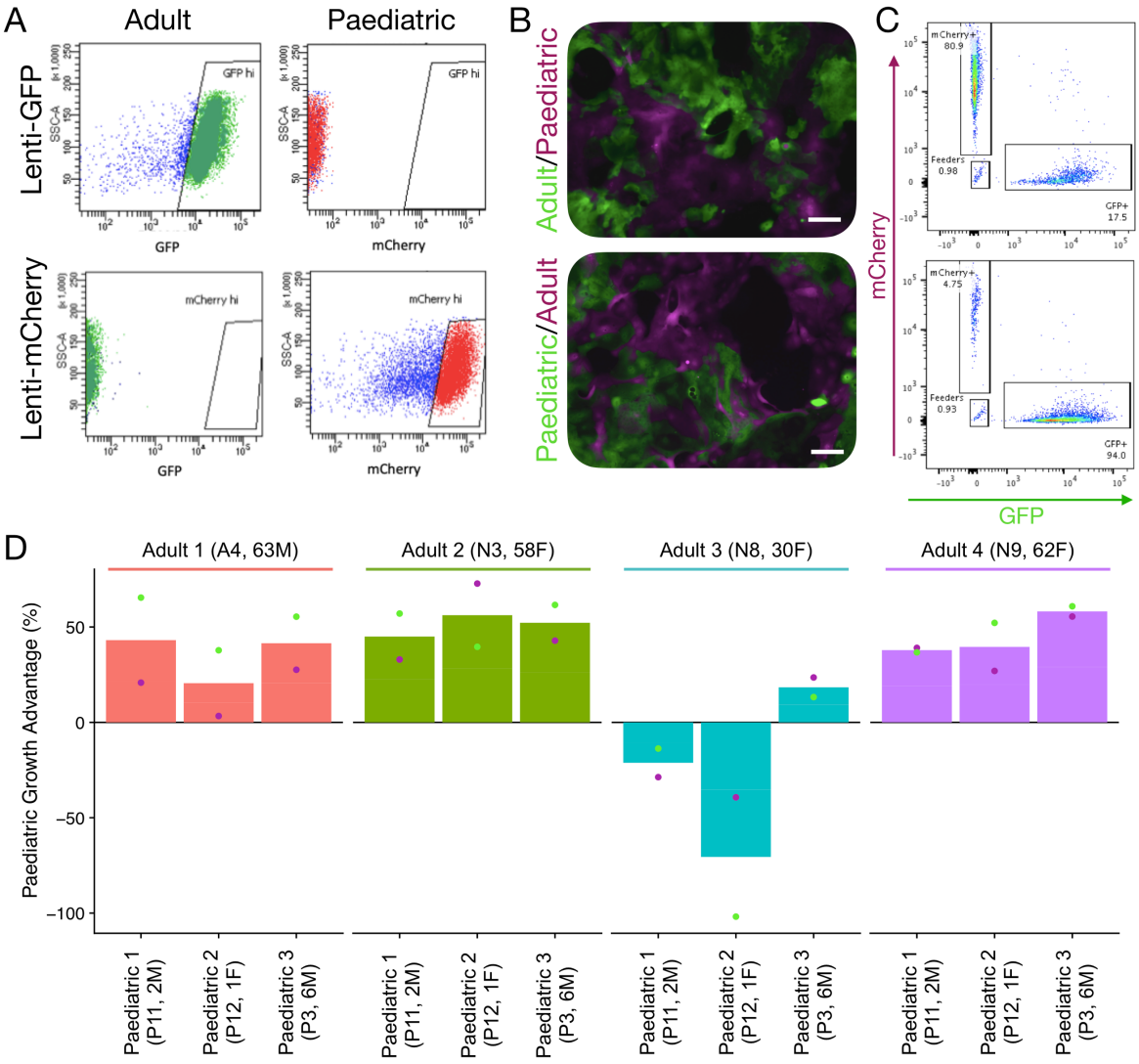

Figure 4: Paediatric airway basal cells out compete adult basal cells in mixed cultures.

(A) GFP$^+$ and mCherry$^+$ proximal airway basal cells were generated from paediatric (n=3) and adult (n=4) donors by lentiviral transduction and purified by fluorescence-activated cell sorting (FACS). **(B)** Representative immunofluorescence demonstrates how each paediatric donor was mixed with each adult donor in a 1:1 ratio by FACS in 24-well plates containing 3T3-J2 feeder cells and epithelial cell culture medium containing Y-27632. Scale bars = 200 µm. **(C)** Flow cytometric analysis was performed when cultures reached approximately 80% confluence. As shown, this allowed quantification of the abundance of GFP$^+$/mCherry$^+$ and GFP$^-$/mCherry$^-$ feeder cells in these co-cultures. **(D)** Summary data for the co-cultures between each paediatric and adult cell culture pair. The growth advantage of paediatric donors relative to adult donors is shown. Experiments were performed in technical triplicate and values for both GFP (paediatric) / mCherry (adult) and mCherry (paediatric) / GFP (adult) co-cultures are plotted as green circles and purple circles, respectively. Donor sex and age are as follows: P3 = M, 6 years old; P11 = M, 2 years old; P12 = F, 1 year old; A4 = M, 63 years old; N3 = F, 58 years old N8 = F, 30 years old; N9 = F, 62 years old.

Cultured basal cells demonstrate greater age-related transcriptional differences than basal cells *in vivo*

Given the different behaviour and phenotype of cultured basal cells from children and adults observed *in vitro*, we assessed whether these changes correlate with gene expression differences in native basal cells. We used FACS to isolate EpCAM$^+$/PDPN$^+$ basal cells (Miller et al., 2018; Weeden et al., 2017) directly from tracheal biopsies and performed bulk RNA sequencing. Consistent with successful purification of basal cells, we saw enrichment for basal cell-associated gene expression in this dataset compared to the laser-capture microdissected whole epithelium (Figure S5A). Using DESeq2 with an FDR of 1% and $\log_2$ fold change threshold of 1.2, we identified 32 genes with significant differential expression between basal cells sorted from paediatric and adult donors, of which 7 were upregulated in children and 25 were more highly expressed in adults (Figure 5A; Table S3). *NTRK2* has previously been associated with basal cell function as it was upregulated in polyp versus non-polyp basal cells in human nasal basal epithelial cells (Ordovas-Montanes et al., 2018). Here, it was upregulated in adult compared to paediatric basal cells, consistent with a possible negative influence on basal cell progenitor function. However, the majority of differentially expressed genes do not have previously described roles in airway basal cells. GSEA suggested that pathways such as TNFα and MTORC1 signalling, as well as processes such as inflammation and apoptosis, were higher in paediatric basal cells although all pathways were of borderline statistical significance (Figure 5B).

Next, we asked whether cell culture differentially affected the transcriptome of paediatric and adult tracheobronchial basal cells in a manner that might result in functional differences. We performed bulk RNA sequencing on cultured basal cells that were isolated and expanded on 3T3-J2 mouse embryonic feeder cells in epithelial cell culture medium containing Y-27632. As expected, cultured cells were enriched for basal cell-associated gene expression compared to laser-capture microdissected whole epithelium (Figure S5A). Freshly sorted basal cells were more similar to laser-capture microdissected whole epithelium than cultured basal cells (Figure S5B), emphasising the significant impact of the artificial cell culture environment on the basal cell transcriptome. Using DESeq2 with an FDR of 1% and $\log_2$ fold change threshold of 1.2, we identified 983 genes with significant differential expression between basal cells sorted from

paediatric and adult donors, of which 554 were upregulated in children and 429 were more highly expressed in adults (Figure 5C; Table S3). Notably, the mucins *MUC2*, *MUC3A*, *MUC5AC*, *MUC5B*, and *MUC17*, as well as the secretory master regulator *SPDEF*, were all more highly expressed by adult cultured basal cells than paediatric basal cells. In adult basal cell cultures, we have previously observed upregulation of mucosecretory genes, such as *SCGB3A1*, in cells co-cultured with mouse embryonic feeder cells in medium containing Y-27632 compared to those in a serum-free alternative, bronchial epithelial growth medium (Butler et al., 2016). However, even if mucosecretory gene expression is favoured in these conditions, it is unclear why paediatric and adult epithelial cells differ in their response to culture. GSEA suggested alterations in multiple pathways, some of which, including TNFα signalling and the inflammatory response, were now in the opposite direction to what was seen in basal cells *in vivo* (Figure 5D).

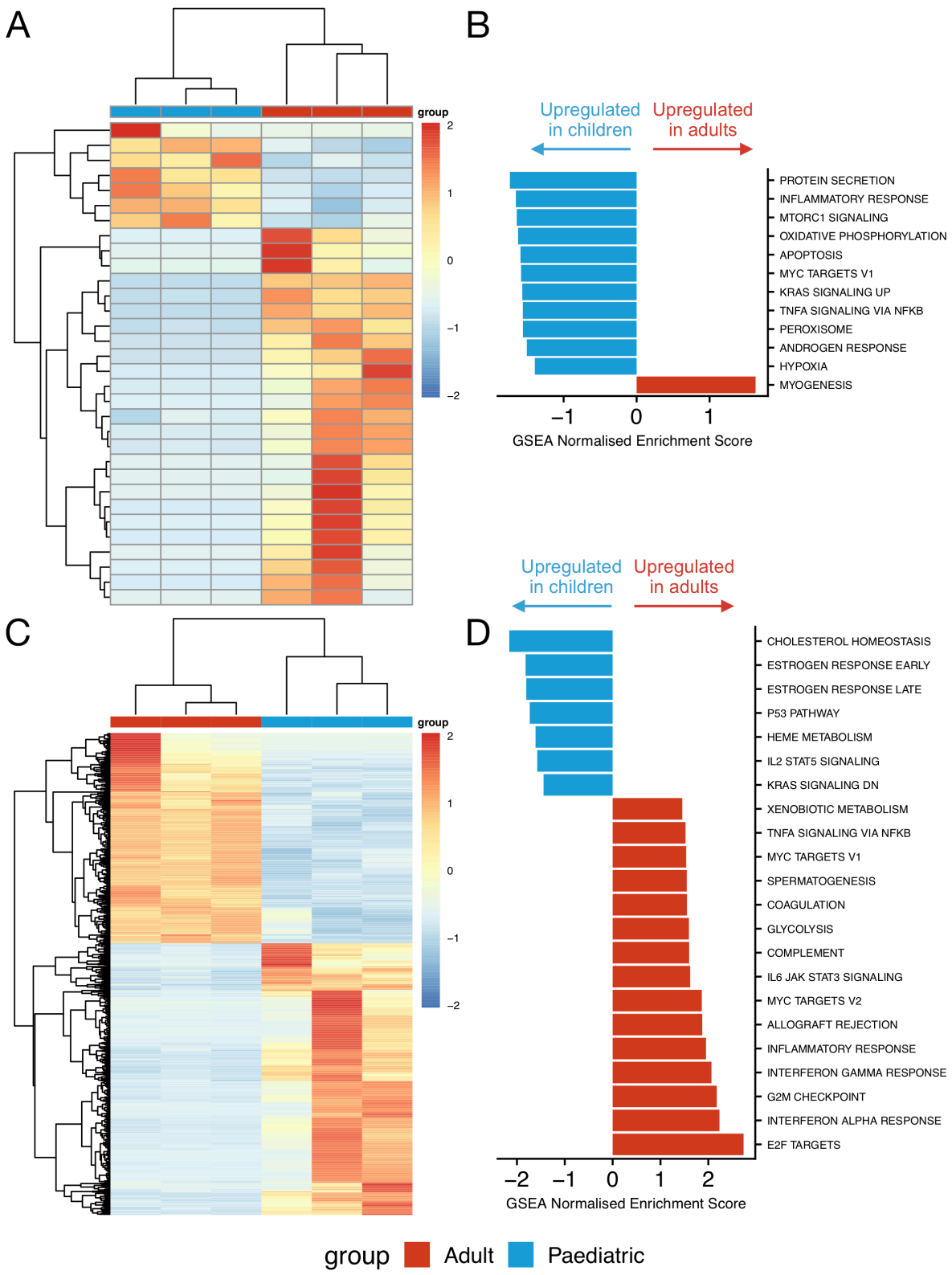
A
group
2
1
0
−1
−2
B
Upregulated in children
Upregulated in adults
PROTEIN SECRETION
INFLAMMATORY RESPONSE
MTORC1 SIGNALING
OXIDATIVE PHOSPHORYLATION
APOPTOSIS
MYC TARGETS V1
KRAS SIGNALING UP
TNFA SIGNALING VIA NFKB
PEROXISOME
ANDROGEN RESPONSE
HYPOXIA
MYOGENESIS
−1
0
1
GSEA Normalised Enrichment Score
C
group
2
1
0
−1
−2
D
Upregulated in children
Upregulated in adults
CHOLESTEROL HOMEOSTASIS
ESTROGEN RESPONSE EARLY
ESTROGEN RESPONSE LATE
P53 PATHWAY
HEME METABOLISM
IL2 STAT5 SIGNALING
KRAS SIGNALING DN
XENOBIOTIC METABOLISM
TNFA SIGNALING VIA NFKB
MYC TARGETS V1
SPERMATOGENESIS
COAGULATION
GLYCOLYSIS
COMPLEMENT
IL6 JAK STAT3 SIGNALING
MYC TARGETS V2
ALLOGRAFT REJECTION
INFLAMMATORY RESPONSE
INTERFERON GAMMA RESPONSE
G2M CHECKPOINT
INTERFERON ALPHA RESPONSE
E2F TARGETS
−2
−1
0
1
2
GSEA Normalised Enrichment Score
group
Adult
Paediatric

Figure 5: RNA sequencing of tracheobronchial basal cells from children and adults.

(A) Cluster diagram showing the expression of 32 differentially expressed genes with a false discovery rate (FDR) < 0.01 and a $\log_2$ fold change > 1.2 in three paediatric (months of age/sex; 2.5M, 11M and 14M) and three adult (years of age/sex; 58F, 61F, 62F) freshly FACS-sorted $EpCAM^+/PDPN^+$ basal cell samples. Gene order is based on hierarchical clustering based on the similarity in overall expression patterns. Red represents relative expression higher than the median expression and blue represents lower expression. **(B)** Pathway analysis was performed on the three paediatric and three adult freshly FACS-sorted $EpCAM^+/PDPN^+$ basal cell samples using gene set enrichment analysis (GSEA) to interrogate Hallmark pathways from MSigDB. For pathways with FDR < 0.05, normalized enrichment scores are shown. A negative score (blue) represents upregulation of the pathway in the paediatric samples; a positive score (red) represents upregulation in the adult samples. **(C)** Cluster diagram showing the expression of 983 differentially expressed genes with a false discovery rate (FDR) < 0.01 and a $\log_2$ fold change > 1.2 in three paediatric (months of age/sex; 3M, 30M, 83M) and three adult (years of age/sex; 58F, 60F 69M) cultured basal cell samples. Cells were cultured on 3T3-J2 mouse embryonic fibroblast feeder layers for two or three passages in epithelial cell culture medium containing Y-27632. Gene order is based on hierarchical clustering based on the similarity in overall expression patterns. Red represents relative expression higher than the median expression and blue represents lower expression. **(D)** Pathway analysis was performed on the three paediatric and three adult cultured basal cell samples using gene set enrichment analysis (GSEA) to interrogate Hallmark pathways from MSigDB. For pathways with FDR < 0.05, normalized enrichment scores are shown. A negative score (blue) represents upregulation of the pathway in the paediatric samples; a positive score (red) represents upregulation in the adult samples.

Transcription of SARS-CoV-2 infection-associated genes is similar between human paediatric and adult tracheobronchial epithelium

Airway epithelial cell types differ in their gene expression between the nasal and airway epithelium (Deprez et al., 2019), and indeed between the proximal and distal airways (Travaglini et al., 2019), potentially contributing to the region-variable infection pattern of the novel coronavirus SARS-CoV-2 (Sungnak et al., 2020; Wölfel et al., 2020). However, this virus also produces markedly less severe symptoms in paediatric than adult COVID-19 patients (Han et al., 2020; Su et al., 2020). Upregulated type I and II interferon-associated gene expression in children (Figure 2B) might be indicative of more effective anti-viral immunity whereas higher Wnt pathway-associated gene expression in adults (Figure 2B) might be relevant since Wnt activation has been shown to increase rates of influenza replication (More et al., 2018). To further assess epithelial-intrinsic factors that might explain the disparity in symptom severity, we compared the expression of a manually curated list of 99 genes associated with viral infection in epithelial cells (Table S2) in our laser-capture microdissected whole tracheobronchial epithelium dataset (Figure 6A). While there was heterogeneity of expression in some of these genes between individuals, clustering was not seen by donor age (Figure 6A).

The entry of coronaviruses into host cells depends upon the binding of the viral spike (S) protein to a host receptor and S protein priming by host proteases. It has been reported that SARS-CoV-2 uses ACE2 as a receptor and the TMPRSS2 protease for S priming, although cathepsin L activity may also have a role (Hoffmann et al., 2020). While *ACE2* and *TMPRSS2* transcripts are found in adult nasal and airway goblet and ciliated cells (Bertram et al., 2012; Sungnak et al., 2020), protein expression of ACE2 in particular seems to be low overall in airway epithelium and restricted to rare cells (Aguiar et al., 2020; Hikmet et al., 2020). Moreover, it is unclear whether these findings extend to paediatric nasal or airway epithelia. To address this, we looked at the expression of these, and other candidate SARS-CoV-2 infection-associated genes (Table S2), in laser-capture microdissected whole tracheobronchial epithelium. Expression was equivalent between the paediatric and adult tissue, although *ADAM17* – a gene that is involved in ectodomain shedding of both EGFR ligands (Vallath et al., 2014) and ACE2 in the context of SARS-CoV (Lambert et al., 2005) – neared statistical significance with higher expression in adult epithelium (Figure 6B). We also determined the relative expression of these

genes compared to all other genes in each dataset, and found that, although *ACE2* is detected in each, it is typically in the midrange of transcripts detected (Figure 6C). Expression of all candidate SARS-CoV-2 infection-associated genes was also comparable between children and adults in our FACS-sorted basal cell (Figure S6) and cultured basal cell datasets (Figure S7), although the lower n number in those experiments limits these analyses. We hope that these RNA sequencing datasets (GEO accession number GSE148818) will be a useful resource for investigations of COVID-19.

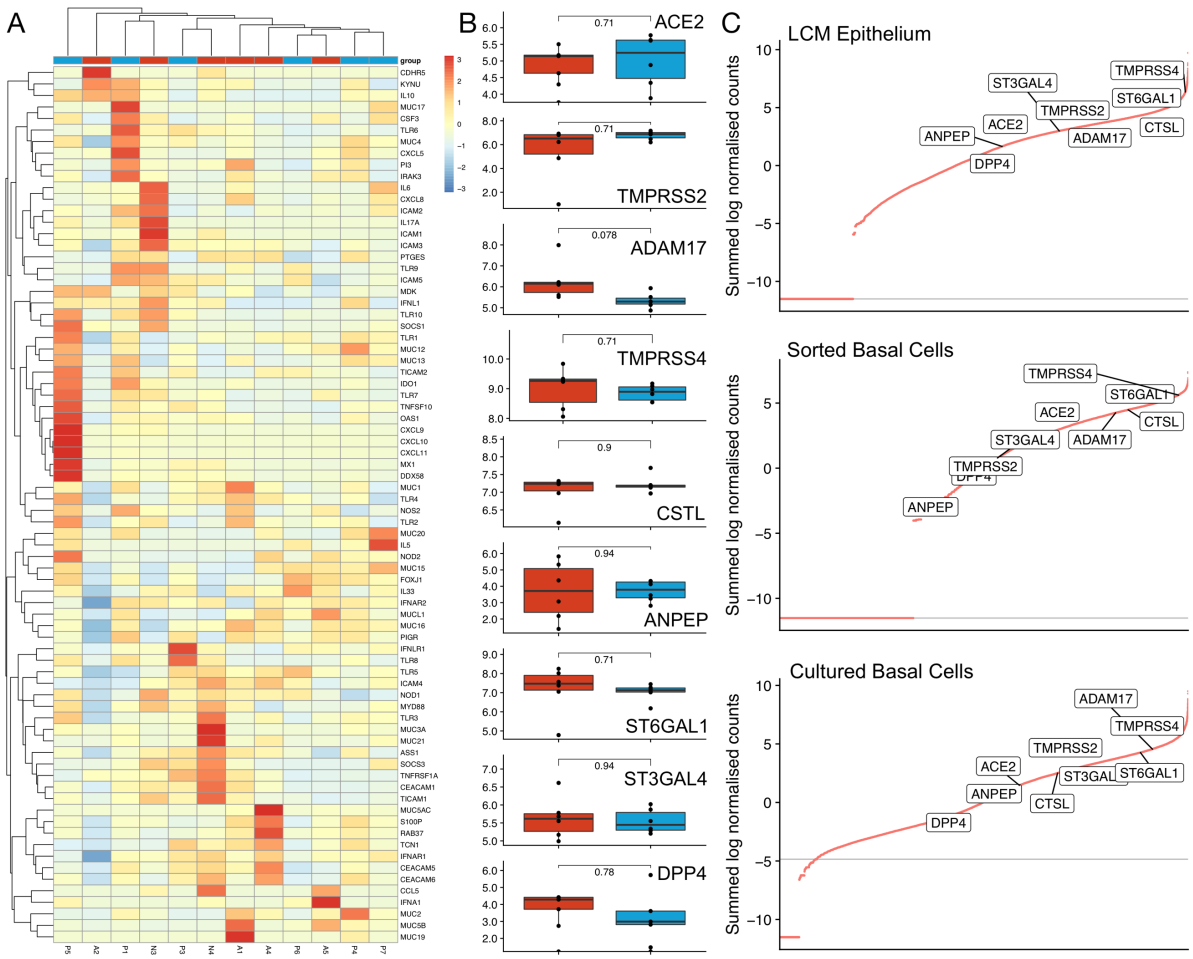

Figure 6: Expression of viral infection-associated epithelial cell genes in paediatric and adult proximal airway epithelial cells.

(A) Cluster diagram showing the expression of 76 epithelial cell genes related to viral infection in the six paediatric and six adult laser-capture microdissected epithelial samples. Gene order is based on hierarchical clustering based on the similarity in overall expression patterns. Red represents relative expression higher than the median expression and blue represents lower expression. Genes that are not expressed in this dataset are excluded (see Table S2). **(B)** Plots comparing the expression (log normalized counts) of selected host genes associated with SARS-CoV-2 (*ACE2*, *TMPRSS2*, *CTSL*), other coronavirus (*ADAM17, ANPEP*, *DPP4*) and influenza (*ST6GAL1, ST3GAL4, TMPRSS4*) infection in the paediatric and adult laser-capture microdissected epithelial samples (n = 6 donors/age group; all genes analysed were non-significant (FDR > 0.05) when comparing paediatric and adult samples using a Wilcoxon test with correction for multiple testing by the Benjamini-Hochberg method). **(C)** Plots showing the expression of the genes from (B) relative to the expression of all of the other genes in each of the three RNA sequencing datasets (laser-capture microdissected whole epithelium, "LCM epithelium"; FACS-sorted EpCAM^{+}/PDPN^{+} basal cells, "sorted basal cells"; cultured basal cells, "cultured basal cells".

Discussion

In this study, we explored differences between human tracheobronchial basal cells in children and adults in three bulk transcriptomic experiments; we compared laser-capture microdissected whole epithelium, FACS-sorted basal cells and cultured basal cells. At the level of the whole epithelium, there was broad conservation of airway epithelial transcriptional programmes but notable differences in the expression of genes associated with interferon responses and cell proliferation, which we speculate may lead to functional differences in cellular behaviour. In support of this, epithelial cell culture showed that basal cells from children have a greater colony forming capacity than those from adults. These data are consistent with studies in other epithelia, where ageing reduces the proportion of cells identified as stem cells using *in vitro* methodologies (Barrandon and Green, 1987). Such differences in clonal potential and proliferative capacity of basal cells between children and adults might be responsible for their differing repair responses following airway injury (Smith et al., 2013), and is of relevance for lung regenerative medicine. For example, paediatric cells might be more amenable to engraftment following transplantation, as is the case in bone marrow transplantation, where donor age significantly affects outcomes (Kollman et al., 2001). Although future airway epithelial cell therapies are likely to be predominantly required by older people and be autologous in nature, recognising age-related differences in regenerative capacity might allow the development of approaches that improve the culture and transplantation of aged basal cells. Further, corrective gene and cell therapies in the context of genetic diseases such as cystic fibrosis (Vaidyanathan et al., 2020) might be more efficient if performed early in life, when cultured cells have greater progenitor potential.

Basal cell-specific RNA sequencing revealed few differences in gene expression between paediatric and adult tracheobronchial basal cells *in vivo* that might explain the differences in their proliferative capacity once cultured, although the role of many differentially expressed genes has not been determined in respiratory epithelial cells. Upon culture, new differences between paediatric and adult basal cells emerged, suggesting that cells from donors of different ages might respond differently to cell culture. This would be consistent with paediatric cultures being more proliferative and typically proliferating for longer time periods than their adult counterparts in population doubling assays. The mechanisms by which airway

basal cells lose their *in vitro* proliferative capacity with age, and whether this reflects *in vivo* loss of progenitor capacity are important areas for further study.

During the current COVID-19 pandemic, it has been observed that children suffer less severe symptoms than adult patients. There has been speculation that this might be due to their lower expression levels of the viral entry receptor, ACE2. Our results imply that there are no major differences in the epithelial expression of viral infection-associated genes between children and adults in the proximal airways, and suggest that expression of *ACE2* and other genes implicated in SARS-CoV-2 cell entry are also comparable. Of course, there are multiple caveats: investigating nasal epithelium may reveal differences that are not present in the tracheobronchial epithelium since both COVID-19 symptomatic and asymptomatic patients show a greater viral yield in nasal samples than throat samples (Zhou et al., 2020), our RNA sequencing studies did not include donors over the age of 72 who have a higher COVID-19 morbidity and mortality risk, and future single cell RNA sequencing of blood or nasal epithelial brush biopsies might reveal differences in immune cell phenotype or viral response that we could not detect here. Thus far, human adult lung single cell RNA sequencing has been used in attempts to correlate ACE2 biology with clinical observations during the current outbreak of COVID-19 (Sungnak et al., 2020). However, the relative lack of symptoms in paediatric COVID-19 patients justifies accelerated efforts to map human paediatric nasal and lung cells to facilitate comparisons with these adult samples.

It is possible to culture large numbers of airway epithelial cells in these and other cell culture conditions (Butler et al., 2016; Peters-Hall et al., 2018; Zhang et al., 2018). Airway basal cells might be candidates for *in vitro* replication of SARS-CoV-2, which would minimize issues arising as a result of physiological differences with cell lines derived from other organs and/or species (Poon et al., 2005). Although we detected transcripts for genes associated with viral uptake in 2D primary cultured human tracheobronchial basal cells, other studies have shown that airway basal cells in another medium composition do not express ACE2 protein (Aguiar et al., 2020) so this requires validation at the protein level. Nevertheless, cell culture conditions alter basal cell gene expression more significantly than inter-patient variability (Butler et al., 2016), so comparing ACE2 protein expression across basal cell culture systems might be of interest. Basal cells can also be differentiated towards mucosecretory and ciliated lineages in scalable air-

liquid interface (Lee et al., 2020) or organoid (Sachs et al., 2019) cultures that contain the cellular lineages thought to be targeted by SARS-CoV-2 in patients, suggesting these for use in viral infection studies (Jonsdottir and Dijkman, 2016; Pizzorno et al., 2020).

Materials and methods

Patient samples

Ethical approval to obtain patient tracheobronchial biopsies was granted by the National Research Ethics Committee (REC references 11/LO/1522 and 06/Q0505/12) and patients (or their parents) gave informed, written consent. Luminal biopsies were obtained using cupped biopsy forceps from patients undergoing planned rigid laryngotracheobronchoscopy under general anaesthesia or flexible bronchoscopy under sedation. Patient characteristics, procedure indication and precise site of biopsy are included in Table S1.

Histology

Samples for histology were fixed overnight in 4% paraformaldehyde (PFA) before being dehydrated through an ethanol gradient using a Leica TP 1050 vacuum tissue processor. Samples were embedded in paraffin and sectioned at 5 µm thickness using a microtome. Haematoxylin and eosin (H&E) staining was performed using an automated system (Tissue-Tek DRS, Sakura). Immunohistochemistry was performed by The Queen Mary University of London Barts Cancer Institute Pathology Service using the Ventana DabMap Horseradish Peroxidase Kit. Stained slides were scanned using a Nanozoomer Whole Slide Imager (Hamamatsu Photonics) to create virtual slides using NDP.View2 software. For cell type quantification, images of sections that contained areas of intact epithelium with at least 300 cells were used (overall 1 – 6 slides were assessed per donor). Images were reviewed in Fiji software and positively stained cells counted using the cell count function. In total, 12,568 (TP63), 9,651 (MUC5AC) and 16,144 (FOXJ1) cells were assessed for expression of these proteins.

Bulk RNA sequencing of tracheobronchial epithelium

Bronchoscopic biopsies were frozen immediately in optimal cutting temperature compound (OCT; in liquid hexane or on dry ice) and transported to a histopathology laboratory (Great Ormond Street Children's Hospital, London, U.K.) within 2 hours on dry ice. Blocks and cut slides were stored at -80°C prior to use. 10 µm sections were mounted on MembraneSlide 1.0 PEN (D) membrane covered slides (Zeiss). One H&E slide was cut per block to aid navigation and identification of epithelium and basement membrane. Slides were prepared with serial washes in methanol, RNAse-free water, RNAse inhibitor and ethanol to remove residual OCT. Laser capture microdissection was performed using a PALM MicroBeam 4 Laser

Microdissection microscope at 10x and 20x magnification (Figure S1) to extract the epithelial portion (or all cells above the basement membrane) from each biopsy into microadhesive-capped tubes (Zeiss). Samples were suspended in a 2:1 mix of Arcturus PicoPure Extraction buffer: RNAlater (Life Technologies) and stored at -80°C until use.

For RNA extraction, samples were thawed, disrupted by lysis (incubation at 42°C for 30 mins followed by incubation at room temperature for 5 mins), vortexed and filtered using RNAeasy MinElute columns (Qiagen). RNA extraction was performed using the Arcturus PicoPure RNA Isolation Kit (Life Technologies Ltd; KIT0204) as per the manufacturer's instructions. RNA was quantified using the Qubit RNA HS Assay Kit (Thermo Fisher Scientific). Libraries were created by UCL Genomics Core Facility using the SMARTer Stranded Total RNAseq Kit (Clontech), cleaned using JetSeq (Bioline) and quality control analysis of RNA integrity was performed using High Sensitivity RNA ScreenTape and the TapeStation Analysis software (Agilent Technologies). RNA sequencing was performed using 0.5x NextSeq for 75 cycles (Illumina; 43PE, ~33M reads per sample).

RNA sequencing analysis

Following sequencing, run data were demultiplexed and converted to FASTQ files using Illumina's bcl2fastq Conversion Software v2.19. Quality control and adapter trimming were performed using fastp (Chen et al., 2018) version 0.20.1 with default settings. FASTQ files were then tagged with the UMI read (UMITools (Smith et al., 2017)) and aligned to the human genome UCSC hg38 using RNA-STAR (Dobin et al., 2013) version 2.5.2b. Aligned reads were UMI deduplicated using Je-suite (Girardot et al., 2016) version 1.2.1 and count matrices were obtained using featureCounts. Downstream analysis was performed using the R statistical environment version 3.5.0 with Bioconductor version 3.8.0 (Huber et al., 2015). Counts were compared between paediatric and adult groups using DESeq2 (Love et al., 2014) using the default settings. Only genes with non-zero counts in at least two samples were included in differential analysis. Pathways were assessed using an implementation of gene set enrichment analysis (GSEA) in the fgsea R package (Korotkevich et al., 2019), using Hallmark gene sets from MSigDB (Liberzon et al., 2015; Subramanian et al., 2005) as input. Heatmaps were plotted using the pheatmap (Kolde, 2012) package, implementing a complete linkage clustering method. All other plots were created using ggplot2. To avoid confounding by sex distribution, differences

between the paediatric and adult groups, genes on the X and Y chromosomes were removed prior to differential analysis.

Identification of gene lists

This study utilises several gene lists as included in Table S2: markers of basal, secretory and ciliated cells, viral response genes and COVID-19 genes of interest. In each case, potentially relevant genes were identified following expert review of the literature. For cell marker genes, we further refine these lists following the method of Danaher and colleagues (Danaher et al., 2017), based on the assumption that genes consistently associated with a cell type should correlate with each other. Using normal lung RNA sequencing data from the Gene Tissue Expression Project (GTEx Consortium, 2015), we constructed a similarity matrix for each cell type following the Danaher method and performed hierarchical clustering. In each case, a clear co-correlated cluster was observed (Figure S8). Genes within this cluster were taken forwards for further analysis (Figure S8; Table S2).

Fluorescence-activated cell sorting of basal cells

Samples were transported to the laboratory in transport medium consisting of αMEM containing penicillin/streptomycin, amphotericin B and gentamicin. Cell suspensions were generated by sequential enzymatic digestion using dispase 16 U/mL (Corning) for 20 minutes at room temperature followed by 0.1% trypsin/EDTA (Sigma) for 30 minutes at 37°C (both in RPMI medium, Gibco). Each enzyme step was quenched with medium containing FBS, placed on ice and combined following the second digestion. Biopsies were manually homogenized using sharp dissection between digest steps and by blunt homogenization through a 100 µm cell strainer (Miltenyi Biotec). Centrifugation steps were performed at 300 x g for 5 mins at 4°C. Cells were blocked in a fluorescence-activated cell sorting (FACS) buffer composed of PBS containing 1% FBS, 25 mM HEPES buffer and 1 mM EDTA for 20 minutes in 96-well V bottomed plates (Thermo Fisher Scientific). Cells were centrifuged as above before staining for 20 minutes on ice at 4°C. The antibodies used were CD31 (BV421; Biolegend; 303124), CD45 (BV421; Biolegend; 304031), EpCAM (APC; Biolegend; 324208), PDPN (PE-Cy7; Biolegend; 337013). Cells were resuspended in FACS buffer for sorting.

For colony formation assays, basal cells were sorted into collagen I-coated 96-well plates containing 3T3-J2 feeder cells at 20,000 cells per cm^2 using a BD FACSAria Fusion FACS sorter running BD FACSDiva 8.0 software at the UCL Cancer Institute Flow Cytometry Core Facility. Experiments lasted 7 days and at termination, the number of wells which had become confluent was counted manually using a light microscope. Brightfield images were taken using a Zeiss Axiovert A1 microscope.

For bulk RNA sequencing, basal cells were sorted into epithelial cell culture medium containing Y-27632 for transport to the laboratory before being centrifuged at 300 x g for 5 mins at 4°C and resuspended in RNA extraction buffer and processed as above for laser-captured samples.

Human airway epithelial cell culture

Primary human airway epithelial cells were isolated and expanded on mitotically inactivated 3T3-J2 feeder layers in two previously reported epithelial growth media, one containing Y-27632 (Butler et al., 2016; Liu et al., 2012) and one without (Hynds et al., 2018; Rheinwald and Green, 1975). Feeder layers were prepared as previously described (Hynds et al., 2019). Epithelial cell culture medium without Y-27632 consisted of Dulbecco's modified Eagle's medium (DMEM)/F12 in a 3:1 ratio containing 1x penicillin–streptomycin, 10% fetal bovine serum, 1% adenine, hydrocortisone (0.5 µg/ml), EGF (10 ng/ml), insulin (5 µg/ml), 0.1 nM cholera toxin, 2 x 10^{-5} T3 and gentamicin (10 mg/ml), as previously described (Hynds et al., 2018). Epithelial cell culture medium containing Y-27632 consisted of DMEM/F12 in a 3:1 ratio containing 1x penicillin–streptomycin (Gibco), 5% fetal bovine serum (Gibco) supplemented with 5 µM Y-27632 (Cambridge Bioscience), hydrocortisone (25 ng/ml; Sigma-Aldrich), epidermal growth factor (0.125 ng/ml; Sino Biological), insulin (5 µg/ml; Sigma-Aldrich), 0.1 nM cholera toxin (Sigma-Aldrich), amphotericin B (250 ng/ml; Thermo Fisher Scientific) and gentamicin (10 µg/ml; Gibco), as previously described (Butler et al., 2016). Where indicated, dishes were collagen I-coated by diluting rat tail collagen I (BD Biosciences) to 50 µg/ml in sterile 0.02N acetic acid and applying at 5 µg/cm^2 for one hour at room temperature in a tissue culture hood. Coated surfaces were washed once with sterile PBS before cell seeding. Population doublings were calculated as previously described (Butler et al., 2016).

RNA sequencing of cultured basal cells was performed after either two or three passages in

epithelial cell culture medium containing Y-27632. RNA was extracted using the RNeasy Mini Kit (Qiagen) following the manufacturer's instructions. RNA was submitted for library preparation by the Wellcome-MRC Cambridge Stem Cell Institute. Initial QC was performed using Qubit RNA HS Assay Kit (Thermo Fisher Scientific) and TapeStation Analysis software (Agilent Technologies). 600 ng RNA was used for library preparation using the NEBNext Ultra8482 II Directional RNA Library Prep Kit (Illumina) and the QIAseq FastSelect RNA Removal Kit (Qiagen). Libraries were subsequently measured using the Qubit system and visualized on a TapeStation D5000. Sequencing was performed using the NovaSeq6000 SP PE50 Standard (~58M Illumina reads per sample) at the CRUK-CI Genomics Core. Following quality control and adapter trimming with fastp (as described above), paired-end reads were aligned to the human genome UCSC hg38 using RNA-STAR (Dobin et al., 2013) version 2.5.2b. Count matrices were obtained using featureCounts. Downstream analysis was performed as described for epithelial and basal RNAseq datasets above.

Proliferation assays

For MTT assays, primary human airway epithelial cells that had been isolated and expanded in epithelial cell culture medium without Y-27632 were seeded in 96-well plates at a density of 5,000 cells/well for 24 hours without feeder cells. At the stated end-points, adherent cells were stained with MTT dye solution (10 µl of 1:10 diluted MTT stock solution in culture medium) for 3 hours at 37°C. After incubation, the medium was removed and 100 µl dimethyl sulfoxide (DMSO; Sigma-Aldrich) was added to dissolve the MTT crystals. The eluted specific stain was measured using a spectrophotometer (560 nm).

To analyse EdU uptake, passage 1 primary human airway epithelial cells that had been isolated and expanded in epithelial cell culture medium without Y-27632 were cultured until approximately 70% confluence. Cells were washed with PBS and feeder cells were removed by differential trypsinization. After washing in DMEM containing 10% FBS and then PBS again, the remaining epithelial cells were treated with 10 µM EdU (Life Technologies Click-iT EdU Alexa Fluor 488) for 1 hour. Cells were trypsinized to obtain single cell suspensions, stained according to the manufacturer's instructions and finally co-stained with DAPI. Cells were run on an LSRFortessa (BD Biosciences) flow cytometer and data were analysed using FlowJo 10.0.6 (TreeStar).

Colony formation assays

2,000 cultured human airway epithelial cells were seeded per well of a collagen I-coated six-well plate containing inactivated 3T3-J2 feeder cells. Medium was carefully changed on day 4 and day 8 of culture before the experiment was terminated on day 12. Colonies were fixed for 10 minutes in 4% PFA, stained using crystal violet (Sigma-Aldrich) at room temperature for 20 minutes and washed repeatedly in water. Colonies of more than 10 cells were counted manually using a light microscope. Colony forming efficiency was calculated as: (number of colonies formed/number of seeded cells) * 100.

Immunofluorescence

Cells were cultured in 8-well chamber slides (Ibidi) and fixed in 4% PFA for 20 minutes at room temperature. Slides were washed and stored in PBS until staining. Cells were permeabilized and blocked in PBS containing 10% FBS and 0.025% Triton X for 1 hour at room temperature. Primary anti-Ki67 antibody (Thermo Fisher Scientific, RM-9106) was incubated overnight in block buffer without Triton X at 4°C. After three 5-minute washes in PBS, anti-rabbit secondary antibodies (AlexaFluor dyes; Molecular Probes) were incubated at a 1:200 dilution in block buffer without Triton X for 2 hours at room temperature. Cells were washed in PBS, counter-stained using DAPI (1 µg/ml stock, 1:10,000 in PBS) and washed twice more in PBS. Images were acquired using a Zeiss LSM700 confocal microscope.

Ki67 staining was performed by seeding passage 1 cells on feeder cells for 3 days. Feeder cells were removed by differential trypsinization and basal cells were fixed in 4% PFA prior to staining as above. Ki67 positivity was assessed by manual counting of Ki67-stained nuclei as a proportion of all DAPI-stained nuclei in five images per donor (mean = 1402 cells per donor, range 810 – 1600).

GFP and mCherry lentiviral production

One Shot Stbl3 chemically competent E. coli bacteria (Thermo Fisher Scientific) were transformed using third generation lentiviral plasmids. The GFP-containing and mCherry-containing plasmids were pCDH-EF1-copGFP-T2A-Puro and pCDH-CMV-mCherry-T2A-Puro (gifts from Kazuhiro Oka; Addgene plasmids #72263 and #72264). The packaging plasmids used

were pMDLg/pRRE, pRSV-Rev and pMD2.G (gifts from Didier Trono; Addgene plasmids #12251, #12253 and #12259 (Dull et al., 1998)). Bacteria were plated on ampicillin-containing agar plates overnight, and colonies expanded from this in LB broth containing ampicillin. Plasmids were extracted using the PureLinkTM HiPure Plasmid Maxiprep Kit (Thermo Fisher Scientific) with the PureLinkTM HiPure Precipitator module (Thermo Fisher Scientific) as per the manufacturer's instructions. Plasmid DNA concentration was quantified using a NanoDrop system.

293T Human Embryonic Kidney (HEK) cells grown in DMEM containing 10% FBS and 1x penicillin/streptomycin were used to assemble and produce the viruses. Packaging and transfer plasmids were delivered into 293T cells by transfection with jetPEI (Polyplus; 101-01N) as per the manufacturer's instructions. Medium was changed at 4 hours and 24 hours post-transfection. Culture medium containing the shed virus was collected at 72 hours and incubated with PEG-it virus precipitation solution (System Biosciences) at 4°C for 48 hours. The precipitated virus was extracted following centrifugation of the mixture as per the manufacturer's instructions.

Viral stock titre was determined by incubating 293T HEK cells (50,000 cells/well) in 12- well plates at 1:100, 1:1000, 1:10,000 and 1:100,000 viral dilutions in complete DMEM medium containing 4 µg/mL polybrene for 4 hours before medium was refreshed. After 72 hours, cells were trypsinized, DAPI stained and GFP or mCherry positivity was determined by flow cytometry. Viral titre was calculated as:

Viral titre (particles/ml) = (% positive cells) x 50,000 / Volume of virus added (ml)

Competitive growth assay optimization in 293T HEK cells

293T HEK cells were transduced for 24 hours with either mCherry (multiplicity of infection (MOIs) of 0.15 or 0.75) or GFP viruses (MOIs of 0.05 or 0.5). Following incubation for 72 hours, cells were sorted by FACS into high- or low-expressing populations and expanded further. Population doubling times were calculated for each colour and MOI, and these values were compared to untransduced cells. Cells were then seeded as a 1:1 mix of GFP- and mCherry-expressing cells into quadruplicate wells of a 48-well plate (10,000 of each colour at either high

or low expression). These cells were passaged twice per week. At each passage, the contents of each well were trypsinized and the percentage of GFP and mCherry-expressing cells was analysed by flow cytometry. 80% of each well's cell volume was immediately re-plated into fresh 48-well plates whilst 20% of the sample from each well was stained with a live/dead fixable stain, fixed with 4% PFA and analysed for GFP and mCherry expression by flow cytometry. Reference wells containing 10,000 cells of a single colour were trypsinized at the same time points and counted manually to calculate a doubling time for each cell type.

Lentiviral transduction of cultured human basal cells

Primary human tracheobronchial basal cells were isolated and expanded in epithelial cell culture medium containing Y-27632 for two passages before transduction with either GFP- or mCherry-containing viruses (MOI = 100). Transduction was performed in culture medium plus 4 µg/mL polybrene and medium was exchanged for fresh culture medium after 16 hours. Purification of cultures to remove untransduced cells was performed by FACS for GFP or mCherry positivity after 7-10 days of further expansion. Since both the GFP and mCherry lentiviral plasmids carry a puromycin resistance cassette, PCR copy number analysis for the puromycin cassette was performed to quantify the number of lentiviral copies incorporated per donor cell (Taqman custom copy number assay). Genomic DNA (gDNA) was extracted from each transduced cell culture using a blood & cell culture DNA mini kit (Qiagen) as per the manufacturer's instructions. gDNA was quantified using a NanoDrop system and PCR performed using the TaqMan genotyping master mix and reference assay (Life Technologies). Stable transduction was confirmed by flow cytometry prior to competitive proliferation assays.

Human airway basal cell competitive proliferation assays

Feeder cells were removed from cultures of GFP+ and mCherry+ basal cells by differential trypsinization and basal cells were suspended in FACS buffer containing 5 µM Y-27632 for FACS. Cells were seeded by FACS as a 1:1 mix of GFP- and mCherry-expressing cells into triplicate wells of a 24-well plate containing 3T3-J2 feeder cells (5000 or 10,000 of each colour). All paediatric cultures were crossed with all of the adult cultures of the opposite label. Medium was refreshed after three and five days. At seven days or at 80-90% confluence as ascertained by fluorescence microscopy, the entire contents of the well were trypsinized for flow cytometry (i.e. feeder cells were not removed). The resulting cell suspensions were digested

enzymatically with 0.25 U / mL liberase (thermolysin medium; Roche) in serum-free medium to form a single cell suspension, which was quenched with medium containing 10% FBS. Zombie Violet Live/Dead fixable stain was applied as per the manufacturer's instructions and cell suspensions were fixed with 4% PFA for 15 mins at room temperature. GFP and mCherry expression was assessed by flow cytometry. Feeder cells were seen as an unlabelled population. Untransduced paediatric cells cultured in parallel were used as non-labelled controls and single colour wells containing 10,000 cells from each individual donor were used as single colour controls.

Cell growth advantages were calculated using the growth calculation of Eekels *et al.* for two cell populations, assuming exponential change in the ratio of the two populations over time (Eekels et al., 2012). To calculate the growth advantage of paediatric cells over adult cells, we used the formula:

$$G_a = T_{d(adult)} \times \log_2(y) / (x + T_{d(adult)} \times \log_2(y)) \times 100\%$$

Where G_a is the calculated growth advantage, $T_{d(adult)}$ is the doubling time in days of the adult donor cell culture, x is the number of days of the experiment and y is equal to % paediatric cells / % adult cells at the time point x divided by the % paediatric cells / % adult cells at the time point 0. $T_{d(adult)}$ values were calculated from population doubling curves for each individual donor in independent, single donor cultures under the same cell culture conditions. The experiment was performed in technical triplicate and for each donor pair we calculated values for both GFP (paediatric) / mCherry (adult) and mCherry (paediatric) / GFP (adult) crosses to rule out differential effects on proliferation from the two lentiviruses.

Acknowledgements

The authors thank Simon Broad and Professor Fiona Watt (Kings College London, U.K.) for providing the 3T3-J2 fibroblasts used in our study, Dr. Pascal Durrenberger (University College London, U.K.) for training and advice on laser capture microdissection, George Morrow and Dr. Barry Wilbourn (UCL Cancer Institute Flow Cytometry Core Facility, University College London, U.K.) for assistance with flow cytometry and FACS experiments, and Aimee Avery and Alex Virasami (Great Ormond Street Hospital Pathology Service, U.K.) for their assistance with cryosectioning for RNA sequencing. We also thank George Elia and his team (Pathology Service, Barts Cancer Institute, Queen Mary University of London, U.K.) for assistance with histology, Tony Brooks and Dr. Paola Niola (UCL Genomics; University College London, U.K.) for performing laser-capture microdissected whole epithelium and FACS-sorted basal cell RNA sequencing experiments and the CRUK-Cambridge Institute Genomics Core for performing cultured basal cell RNA sequencing, in particular Dr. Maike Paramor (Wellcome-MRC Cambridge Stem Cell Institute) who prepared the libraries. We further thank members of the Lungs for Living Research Centre (University College London) and Dr. Eva Grönroos (The Francis Crick Institute, U.K.) for critical reading of the manuscript.

E.F.M. (WT201265/Z/16/Z), A.P. (WT211161/Z/18/Z) and C.R.B. (WT097946MA) are Wellcome Trust Clinical Research Training Fellows. S.G.-L. was supported by a Royal Society Newton International Fellowship (NF161172). M.A.B. is an NIHR Senior Investigator (NIHR201360). P.D.C. is supported by the National Institute for Health Research (NIHR-RP-2014-04-046) and by the NIHR Great Ormond Street Hospital Biomedical Research Centre. R.E.H. is a Wellcome Trust Sir Henry Wellcome Postdoctoral Fellow (WT209199/Z/17) and S.M.J. is a Wellcome Trust Senior Fellow in Clinical Science (WT107963AIA). R.E.H. and S.M.J. receive funding as members of the UK Regenerative Medicine Platform (UKRMP2) Engineered Cell Environment Hub (MRC; MR/R015635/1) and P.D.C., R.E.H. and S.M.J. are members of the Longfonds BREATH lung regeneration consortium. R.E.H. and S.M.J. are supported by The Roy Castle Lung Cancer Foundation (2016/07/HYNDS) and S.M.J. also receives funding from The Rosetrees Trust and The UCLH Charitable Foundation. The National Institute for Health Research (NIHR) University College London Hospitals Biomedical Research Centre (UCLH BRC) supported the acquisition of tissue in our study.

Author Contributions

Conceptualization, E.F.M., K.H.C.G., C.R.B, R.E.H. and S.M.J.; Methodology, E.F.M., S.G.-L, C.D., C.R.B., V.H.T. and R.E.H.; Investigation, E.F.M., E.N., K.A.L., J.C.O. and D.D.H.L.; Formal Analysis, E.F.M, A.P. and R.E.H.; Resources, B.E.H., R.J.H., C.Y., G.S.S., M.A.B., C.O., C.M.S., P.D.C., S.M.J.; Visualization, E.F.M., A.P. and R.E.H.; Writing – Original Draft, E.F.M. and R.E.H.; Writing – Review & Editing, E.F.M., A.P., K.H.C.G., R.E.H. and S.M.J.; Funding Acquisition, E.F.M., R.E.H. and S.M.J.; Supervision, S.G.-L, M.A.B., P.D.C., R.E.H. and S.M.J.

Competing interests

The authors declare no competing interests relating to this manuscript. S.M.J. has attended advisory boards for Johnson and Johnson, BARD1 Life Sciences and AstraZeneca. S.M.J. receives grant funding from GRAIL Inc. and Owlstone Medical.

References

Adams, T.S., Schupp, J.C., Poli, S., Ayaub, E.A., Neumark, N., Ahangari, F., Chu, S.G., Raby, B.A., DeIuliis, G., Januszyk, M.*, et al.* (2019). Single Cell RNA-seq reveals ectopic and aberrant lung resident cell populations in Idiopathic Pulmonary Fibrosis. bioRxiv, 759902.

Aguiar, J.A., Tremblay, B.J.-M., Mansfield, M.J., Woody, O., Lobb, B., Banerjee, A., Chandiramohan, A., Tiessen, N., Dvorkin-Gheva, A., Revill, S.*, et al.* (2020). Gene expression and in situ protein profiling of candidate SARS-CoV-2 receptors in human airway epithelial cells and lung tissue. bioRxiv, 2020.2004.2007.030742.

Angelidis, I., Simon, L.M., Fernandez, I.E., Strunz, M., Mayr, C.H., Greiffo, F.R., Tsitsiridis, G., Ansari, M., Graf, E., Strom, T.M.*, et al.* (2019). An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. Nat Commun *10*, 963.

Barrandon, Y., and Green, H. (1987). Three clonal types of keratinocyte with different capacities for multiplication. Proceedings of the National Academy of Sciences *84*, 2302-2306.

Basil, M.C., Katzen, J., Engler, A.E., Guo, M., Herriges, M.J., Kathiriya, J.J., Windmueller, R., Ysasi, A.B., Zacharias, W.J., Chapman, H.A.*, et al.* (2020). The Cellular and Physiological Basis for Lung Repair and Regeneration: Past, Present, and Future. Cell Stem Cell *26*, 482-502.

Bertram, S., Heurich, A., Lavender, H., Gierer, S., Danisch, S., Perin, P., Lucas, J.M., Nelson, P.S., Pohlmann, S., and Soilleux, E.J. (2012). Influenza and SARS-coronavirus activating proteases TMPRSS2 and HAT are expressed at multiple sites in human respiratory and gastrointestinal tracts. PLoS One *7*, e35876.

Braga, F.A.V., Kar, G., Berg, M., Carpaij, O.A., Polanski, K., Simon, L.M., Brouwer, S., Gomes, T., Hesse, L., and Jiang, J. (2019). A cellular census of human lungs identifies novel cell states in health and in asthma. Nature medicine *25*, 1153-1163.

Butler, C.R., Hynds, R.E., Gowers, K.H., Lee Ddo, H., Brown, J.M., Crowley, C., Teixeira, V.H., Smith, C.M., Urbani, L., Hamilton, N.J.*, et al.* (2016). Rapid Expansion of Human Epithelial Stem Cells Suitable for Airway Tissue Engineering. Am J Respir Crit Care Med *194*, 156-168.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884-i890.

Danaher, P., Warren, S., Dennis, L., D'Amico, L., White, A., Disis, M.L., Geller, M.A., Odunsi, K., Beechem, J., and Fling, S.P. (2017). Gene expression markers of Tumor Infiltrating Leukocytes. J Immunother Cancer *5*, 18.

Deprez, M., Zaragosi, L.-E., Truchi, M., Garcia, S.R., Arguel, M.-J., Lebrigand, K., Paquet, A., Pee'r, D., Marquette, C.-H., Leroy, S.*, et al.* (2019). A single-cell atlas of the human healthy airways. bioRxiv, 2019.2012.2021.884759.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D., and Naldini, L. (1998). A third-generation lentivirus vector with a conditional packaging system. J Virol *72*, 8463-8471.

Eekels, J., Pasternak, A., Schut, A., Geerts, D., Jeeninga, R., and Berkhout, B. (2012). A competitive cell growth assay for the detection of subtle effects of gene transduction on cell proliferation. Gene therapy *19*, 1058-1064.

García, S.R., Deprez, M., Lebrigand, K., Cavard, A., Paquet, A., Arguel, M.-J., Magnone, V., Truchi, M., Caballero, I., and Leroy, S. (2019). Novel dynamics of human mucociliary differentiation revealed by single-cell RNA sequencing of nasal epithelial cultures. Development *146*, dev177428.

Girardot, C., Scholtalbers, J., Sauer, S., Su, S.Y., and Furlong, E.E. (2016). Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. BMC Bioinformatics *17*, 419.

GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science *348*, 648-660.

Habermann, A.C., Gutierrez, A.J., Bui, L.T., Yahn, S.L., Winters, N.I., Calvi, C.L., Peter, L., Chung, M.-I., Taylor, C.J., Jetter, C.*, et al.* (2019). Single-cell RNA-sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. bioRxiv, 753806.

Han, Y.N., Feng, Z.W., Sun, L.N., Ren, X.X., Wang, H., Xue, Y.M., Wang, Y., and Fang, Y. (2020). A comparative-descriptive analysis of clinical characteristics in 2019-Coronavirus-infected children and adults. J Med Virol.

Hikmet, F., Méar, L., Uhlén, M., and Lindskog, C. (2020). The protein expression profile of ACE2 in human tissues. bioRxiv, 2020.2003.2031.016048.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A.*, et al.* (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell.

Hogan, B.L., Barkauskas, C.E., Chapman, H.A., Epstein, J.A., Jain, R., Hsia, C.C., Niklason, L., Calle, E., Le, A., Randell, S.H., *et al.* (2014). Repair and regeneration of the respiratory system: complexity, plasticity, and mechanisms of lung stem cell function. Cell Stem Cell *15*, 123-138.
Horani, A., Nath, A., Wasserman, M.G., Huang, T., and Brody, S.L. (2013). Rho-associated protein kinase inhibition enhances airway epithelial Basal-cell proliferation and lentivirus transduction. Am J Respir Cell Mol Biol *49*, 341-347.
Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., *et al.* (2015). Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods *12*, 115-121.
Hynds, R.E., Butler, C.R., Janes, S.M., and Giangreco, A. (2019). Expansion of Human Airway Basal Stem Cells and Their Differentiation as 3D Tracheospheres. Methods Mol Biol *1576*, 43-53.
Hynds, R.E., Gowers, K.H.C., Nigro, E., Butler, C.R., Bonfanti, P., Giangreco, A., Prele, C.M., and Janes, S.M. (2018). Cross-talk between human airway epithelial cells and 3T3-J2 feeder cells involves partial activation of human MET by murine HGF. PLoS One *13*, e0197129.
Jonsdottir, H.R., and Dijkman, R. (2016). Coronaviruses and the human airway: a universal system for virus-host interaction studies. Virol J *13*, 24.
Kicic, A., Sutanto, E.N., Stevens, P.T., Knight, D.A., and Stick, S.M. (2006). Intrinsic biochemical and functional differences in bronchial epithelial cells of children with asthma. Am J Respir Crit Care Med *174*, 1110-1118.
Kimmel, J.C., Penland, L., Rubinstein, N.D., Hendrickson, D.G., Kelley, D.R., and Rosenthal, A.Z. (2019). Murine single-cell RNA-seq reveals cell-identity- and tissue-specific trajectories of aging. Genome Res *29*, 2088-2103.
Kolde, R. (2012). Pheatmap: pretty heatmaps. R Package Version 61.
Kollman, C., Howe, C.W., Anasetti, C., Antin, J.H., Davies, S.M., Filipovich, A.H., Hegland, J., Kamani, N., Kernan, N.A., and King, R. (2001). Donor characteristics as risk factors in recipients after transplantation of bone marrow from unrelated donors: the effect of donor age. Blood, The Journal of the American Society of Hematology *98*, 2043-2051.
Korotkevich, G., Sukhov, V., and Sergushichev, A. (2019). Fast gene set enrichment analysis. bioRxiv, 060012.
Kumar, P.A., Hu, Y., Yamamoto, Y., Hoe, N.B., Wei, T.S., Mu, D., Sun, Y., Joo, L.S., Dagher, R., Zielonka, E.M., *et al.* (2011). Distal airway stem cells yield alveoli in vitro and during lung regeneration following H1N1 influenza infection. Cell *147*, 525-538.
Lambert, D.W., Yarski, M., Warner, F.J., Thornhill, P., Parkin, E.T., Smith, A.I., Hooper, N.M., and Turner, A.J. (2005). Tumor necrosis factor-alpha convertase (ADAM17) mediates regulated ectodomain shedding of the severe-acute respiratory syndrome-coronavirus (SARS-CoV) receptor, angiotensin-converting enzyme-2 (ACE2). J Biol Chem *280*, 30113-30119.
Lee, D.D.H., Cardinale, D., Nigro, E., Butler, C.R., Rutman, A., Fassad, M.R., Hirst, R.A., Moulding, D., Agrotis, A., Forsythe, E., *et al.* (2020). High-content screening for rare respiratory diseases: readthrough therapy in primary ciliary dyskinesia. bioRxiv, 2020.2002.2028.959189.
Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst *1*, 417-425.
Liu, X., Ory, V., Chapman, S., Yuan, H., Albanese, C., Kallakury, B., Timofeeva, O.A., Nealon, C., Dakic, A., Simic, V., *et al.* (2012). ROCK inhibitor and feeder cells induce the conditional reprogramming of epithelial cells. Am J Pathol *180*, 599-607.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.
Meiners, S., Eickelberg, O., and Konigshoff, M. (2015). Hallmarks of the ageing lung. Eur Respir J *45*, 807-827.
Miller, A.J., Yu, Q., Czerwinski, M., Tsai, Y.-H., Conway, R.F., Wu, A., Holloway, E.M., Walker, T., Glass, I.A., Treutlein, B., *et al.* (2018). Basal stem cell fate specification is mediated by SMAD signaling in the developing human lung. bioRxiv, 461103.
Montoro, D.T., Haber, A.L., Biton, M., Vinarsky, V., Lin, B., Birket, S.E., Yuan, F., Chen, S., Leung, H.M., Villoria, J., *et al.* (2018). A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. Nature *560*, 319-324.
More, S., Yang, X., Zhu, Z., Bamunuarachchi, G., Guo, Y., Huang, C., Bailey, K., Metcalf, J.P., and Liu, L. (2018). Regulation of influenza virus replication by Wnt/beta-catenin signaling. PLoS One *13*, e0191010.
Mori, M., Mahoney, J.E., Stupnikov, M.R., Paez-Cortez, J.R., Szymaniak, A.D., Varelas, X., Herrick, D.B., Schwob, J., Zhang, H., and Cardoso, W.V. (2015). Notch3-Jagged signaling controls the pool of undifferentiated airway progenitors. Development *142*, 258-267.
Navarro, S., and Driscoll, B. (2017). Regeneration of the aging lung: a mini-review. Gerontology *63*, 270-280.

Ordovas-Montanes, J., Dwyer, D.F., Nyquist, S.K., Buchheit, K.M., Vukovic, M., Deb, C., Wadsworth, M.H., Hughes, T.K., Kazer, S.W., Yoshimoto, E.*, et al.* (2018). Allergic inflammatory memory in human respiratory epithelial progenitor cells. Nature *560*, 649-654.
Peters-Hall, J.R., Coquelin, M.L., Torres, M.J., LaRanger, R., Alabi, B.R., Sho, S., Calva-Moreno, J.F., Thomas, P.J., and Shay, J.W. (2018). Long-term culture and cloning of primary human bronchial basal cells that maintain multipotent differentiation capacity and CFTR channel function. Am J Physiol Lung Cell Mol Physiol *315*, L313-L327.
Pizzorno, A., Padey, B., Julien, T., Trouillet-Assant, S., Traversier, A., Errazuriz-Cerda, E., Fouret, J., Dubois, J., Gaymard, A., Lescure, F.-X.*, et al.* (2020). Characterization and treatment of SARS-CoV-2 in nasal and bronchial human airway epithelia. bioRxiv, 2020.2003.2031.017889.
Plasschaert, L.W., Zilionis, R., Choo-Wing, R., Savova, V., Knehr, J., Roma, G., Klein, A.M., and Jaffe, A.B. (2018). A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. Nature *560*, 377-381.
Poon, L.L., Leung, C.S., Chan, K.H., Yuen, K.Y., Guan, Y., and Peiris, J.S. (2005). Recurrent mutations associated with isolation and passage of SARS coronavirus in cells from non-human primates. J Med Virol *76*, 435-440.
Prasse, A., Binder, H., Schupp, J.C., Kayser, G., Bargagli, E., Jaeger, B., Hess, M., Rittinghausen, S., Vuga, L., Lynn, H.*, et al.* (2018). BAL Cell Gene Expression is Indicative of Outcome and Airway Basal Cell Involvement in IPF. Am J Respir Crit Care Med.
Reyfman, P.A., Walter, J.M., Joshi, N., Anekalla, K.R., McQuattie-Pimentel, A.C., Chiu, S., Fernandez, R., Akbarpour, M., Chen, C.I., Ren, Z.*, et al.* (2019). Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. Am J Respir Crit Care Med *199*, 1517-1536.
Rheinwald, J.G., and Green, H. (1975). Serial cultivation of strains of human epidermal keratinocytes: the formation of keratinizing colonies from single cells. Cell *6*, 331-343.
Rock, J.R., Gao, X., Xue, Y., Randell, S.H., Kong, Y.-Y., and Hogan, B.L. (2011). Notch-dependent differentiation of adult airway basal stem cells. Cell stem cell *8*, 639-648.
Rock, J.R., Randell, S.H., and Hogan, B.L. (2010). Airway basal stem cells: a perspective on their roles in epithelial homeostasis and remodeling. Disease Models and Mechanisms *3*, 545-556.
Sachs, N., Papaspyropoulos, A., Zomer-van Ommen, D.D., Heo, I., Bottinger, L., Klay, D., Weeber, F., Huelsz-Prince, G., Iakobachvili, N., Amatngalim, G.D.*, et al.* (2019). Long-term expanding human airway organoids for disease modeling. EMBO J *38*.
Smith, L.S., Zimmerman, J.J., and Martin, T.R. (2013). Mechanisms of acute respiratory distress syndrome in children and adults: a review and suggestions for future research. Pediatr Crit Care Med *14*, 631-643.
Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res *27*, 491-499.
Staudt, M.R., Buro-Auriemma, L.J., Walters, M.S., Salit, J., Vincent, T., Shaykhiev, R., Mezey, J.G., Tilley, A.E., Kaner, R.J., Ho, M.W.*, et al.* (2014). Airway Basal stem/progenitor cells have diminished capacity to regenerate airway epithelium in chronic obstructive pulmonary disease. Am J Respir Crit Care Med *190*, 955-958.
Su, L., Ma, X., Yu, H., Zhang, Z., Bian, P., Han, Y., Sun, J., Liu, Y., Yang, C., Geng, J.*, et al.* (2020). The different clinical characteristics of corona virus disease cases between children and their families in China - the character of children with COVID-19. Emerg Microbes Infect *9*, 707-713.
Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.
Sungnak, W., Huang, N., Bécavin, C., Berg, M., and Network, H.L.B. (2020). SARS-CoV-2 Entry Genes Are Most Highly Expressed in Nasal Goblet and Ciliated Cells within Human Airways. arXiv *2003.06122*.
Teixeira, V.H., Nadarajan, P., Graham, T.A., Pipinikas, C.P., Brown, J.M., Falzon, M., Nye, E., Poulsom, R., Lawrence, D., and Wright, N.A. (2013). Stochastic homeostasis in human airway epithelium is achieved by neutral competition of basal cell progenitors. Elife *2*, e00966.
Thurlbeck, W.M., and Angus, G.E. (1975). Growth and aging of the normal human lung. Chest *67*, 3S-6S.
Travaglini, K.J., Nabhan, A.N., Penland, L., Sinha, R., Gillich, A., Sit, R.V., Chang, S., Conley, S.D., Mori, Y., and Seita, J. (2019). A molecular cell atlas of the human lung from single cell RNA sequencing. bioRxiv, 742320.
Turner, J.M., Mead, J., and Wohl, M.E. (1968). Elasticity of human lungs in relation to age. Journal of applied physiology *25*, 664-671.
Vaidyanathan, S., Salahudeen, A.A., Sellers, Z.M., Bravo, D.T., Choi, S.S., Batish, A., Le, W., Baik, R., de la, O.S., Kaushik, M.P.*, et al.* (2020). High-Efficiency, Selection-free Gene Repair in Airway Stem Cells from Cystic Fibrosis Patients Rescues CFTR Function in Differentiated Epithelia. Cell Stem Cell *26*, 161-171 e164.
Vallath, S., Hynds, R.E., Succony, L., Janes, S.M., and Giangreco, A. (2014). Targeting EGFR signalling in chronic lung disease: therapeutic challenges and opportunities. Eur Respir J *44*, 513-522.

Wansleeben, C., Bowie, E., Hotten, D.F., Yen-Rei, A.Y., and Hogan, B.L. (2014). Age-related changes in the cellular composition and epithelial organization of the mouse trachea. PloS one *9*, e93496.
Watson, J.K., Rulands, S., Wilkinson, A.C., Wuidart, A., Ousset, M., Van Keymeulen, A., Gottgens, B., Blanpain, C., Simons, B.D., and Rawlins, E.L. (2015). Clonal Dynamics Reveal Two Distinct Populations of Basal Cells in Slow-Turnover Airway Epithelium. Cell Rep *12*, 90-101.
Weeden, C.E., Chen, Y., Ma, S.B., Hu, Y., Ramm, G., Sutherland, K.D., Smyth, G.K., and Asselin-Labat, M.L. (2017). Lung Basal Stem Cells Rapidly Repair DNA Damage Using the Error-Prone Nonhomologous End-Joining Pathway. PLoS Biol *15*, e2000731.
Wölfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Müller, M.A., Niemeyer, D., Jones, T.C., Vollmar, P., Rothe, C.*, et al.* (2020). Virological assessment of hospitalized patients with COVID-2019. Nature.
Yoshida, K., Gowers, K.H.C., Lee-Six, H., Chandrasekharan, D.P., Coorens, T., Maughan, E.F., Beal, K., Menzies, A., Millar, F.R., Anderson, E.*, et al.* (2020). Tobacco smoking and somatic mutations in human bronchial epithelium. Nature *578*, 266-272.
Zhang, C., Lee, H.J., Shrivastava, A., Wang, R., McQuiston, T.J., Challberg, S.S., Pollok, B.A., and Wang, T. (2018). Long-Term In Vitro Expansion of Epithelial Stem Cells Enabled by Pharmacological Inhibition of PAK1-ROCK-Myosin II and TGF-beta Signaling. Cell Rep *25*, 598-610 e595.
Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L.*, et al.* (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature *579*, 270-273.

Supplementary Data

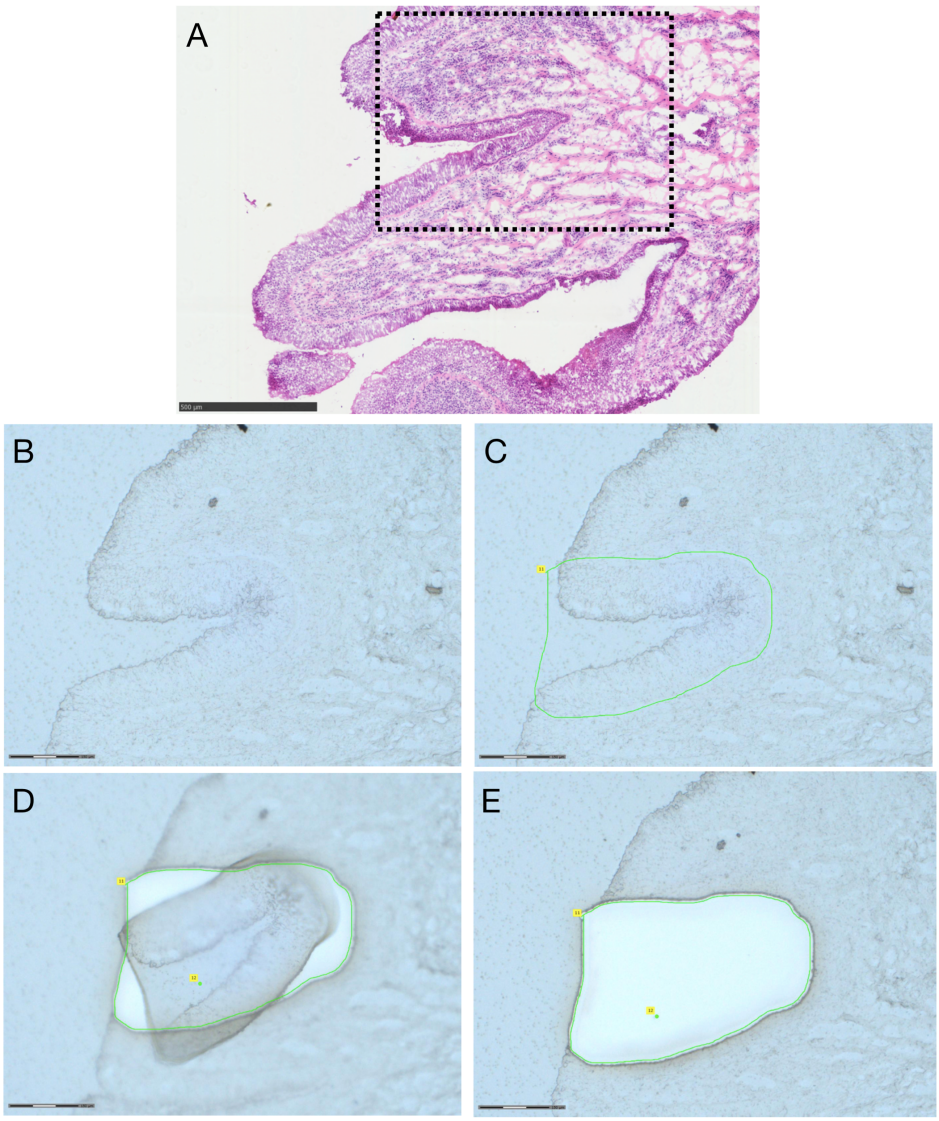

Figure S1: Laser-capture microdissection of whole tracheobronchial epithelium from fresh frozen paediatric and adult donor tissue.

(A) A slide was haematoxylin & eosin stained at the start of sectioning as a guide for identification of the epithelium. The approximate area identified in (B)-(E) is indicated by a black box. Scale bar = 500 µm. **(B)** The equivalent area of epithelium was identified (scale bar = 150 µm) and **(C)** captured using the 'freehand' tool (Zeiss RoboSoftware; green line; scale bar = 150 µm). Care was taken to cut immediately below the basement membrane. **(D)** The laser cut around this outline to release the tissue (scale bar = 150 µm). **(E)** If samples did not immediately lift onto the adhesive surface of the collection tube cap, the 'dot' tool was used to capture the tissue (Zeiss RoboSoftware; scale bar = 150 µm).

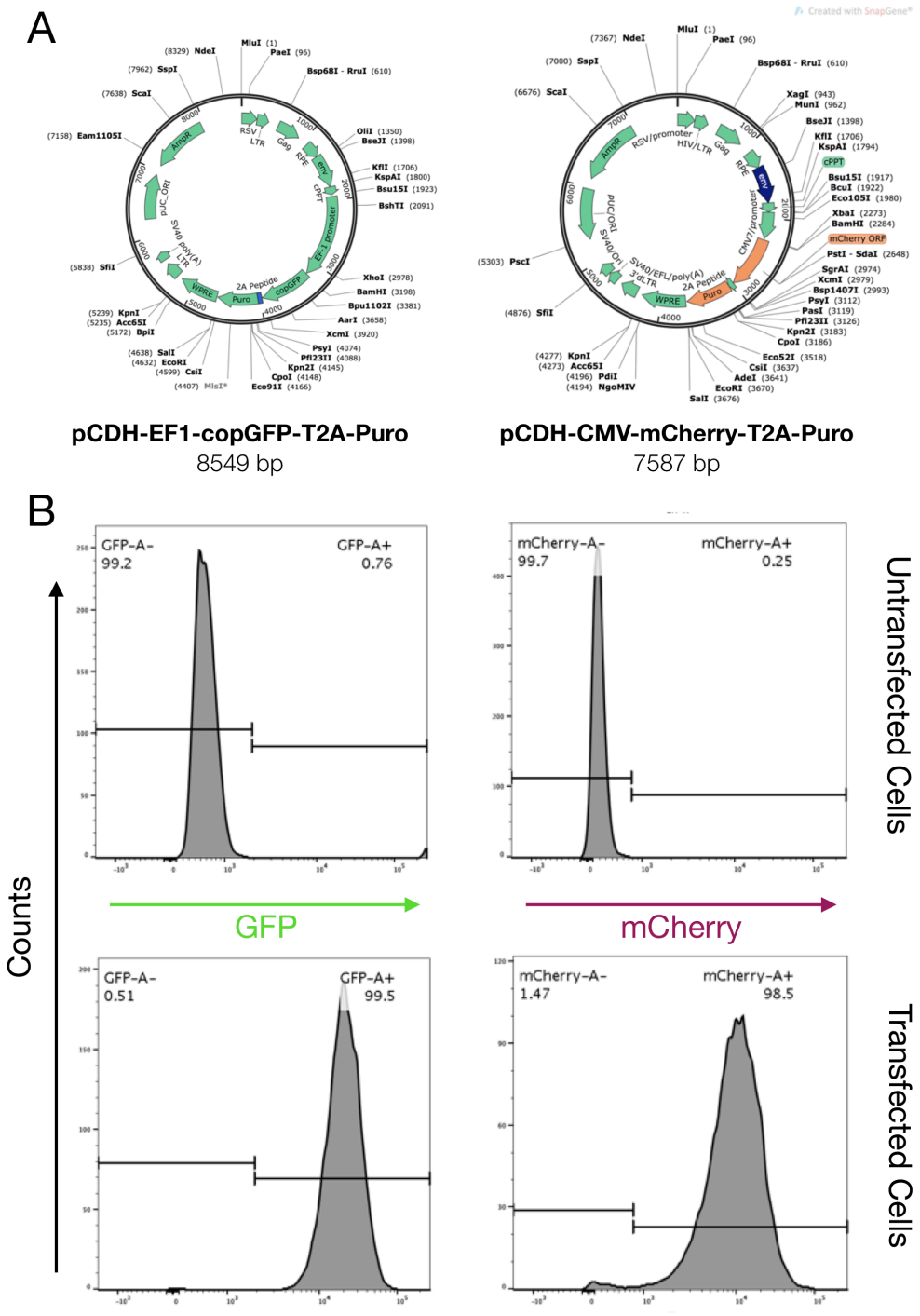

Figure S2: Lentiviral expression of GFP or mCherry for competitive proliferation assay.

(A) Plasmid maps for GFP- and mCherry-containing lentiviral plasmids (created using SnapGene software). **(B)** Flow cytometric analysis of transfection efficiency in 293T HEK cells.

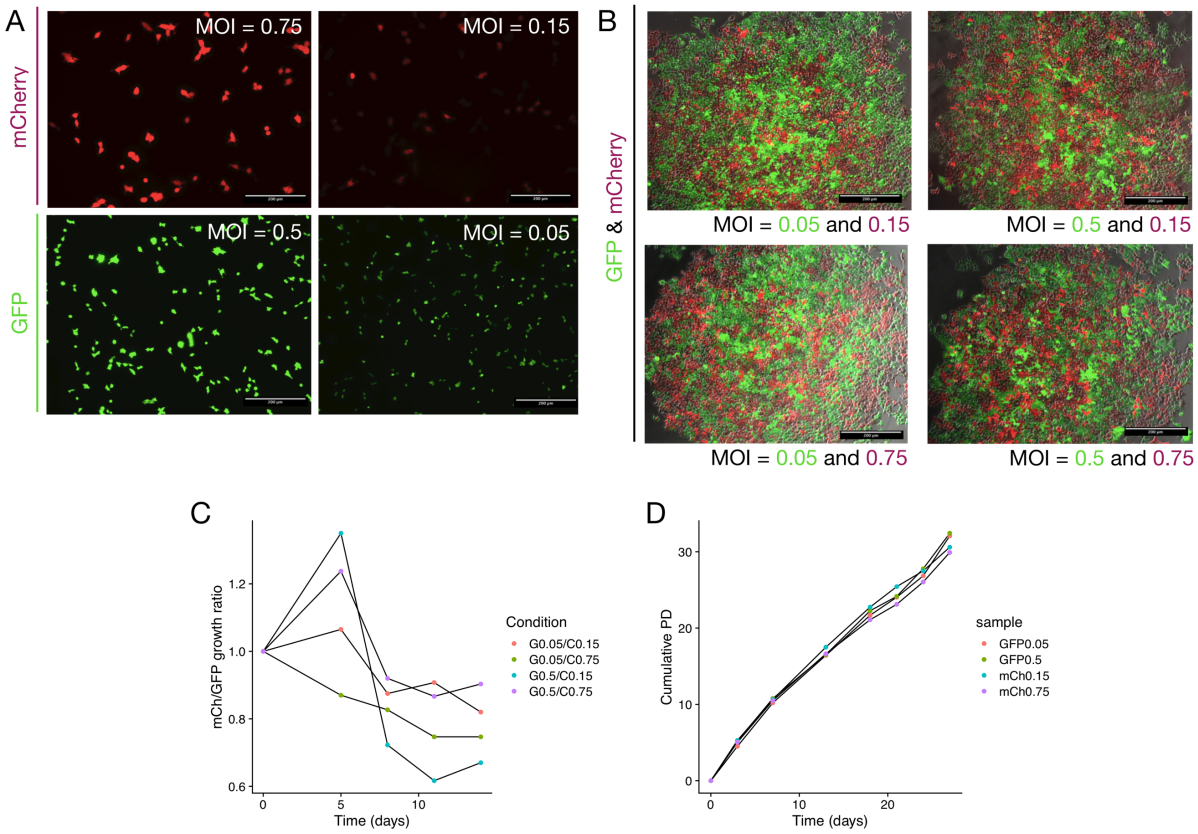

Figure S3: Validation of competitive proliferation assay approach using 293T human embryonic kidney (HEK) cells.

(A) Following lentiviral transduction at two different multiplicities of infection (MOI), 293T HEK cells were FACS-sorted to generate pure cell populations of either GFP+ or mCherry+ cells. Immunofluorescence imaging showed that the fluorescence intensity was higher after transductions at the higher MOI. Cells were imaged 24 hours after sorting. Cells remained stably transduced even after transduction using the lower MOI. Scale bars = 200 µm. **(B)** Fluorescence imaging of mixed cultures of GFP+ and mCherry+ 293T HEK cells after 5 days. Scale bars = 200 µm. **(C)** Flow cytometry analysis of the ratio of mCherry+ and GFP+ cells in mixed cultures of transduced 293T HEK cells. Cells were initially combined in a 1:1 ratio and analysed at the time points indicated. The cells remained close to a 1:1 ratio over two weeks of culture. **(D)** Population doubling analysis of the growth of GFP- and mCherry-expressing 293T HEK cells.

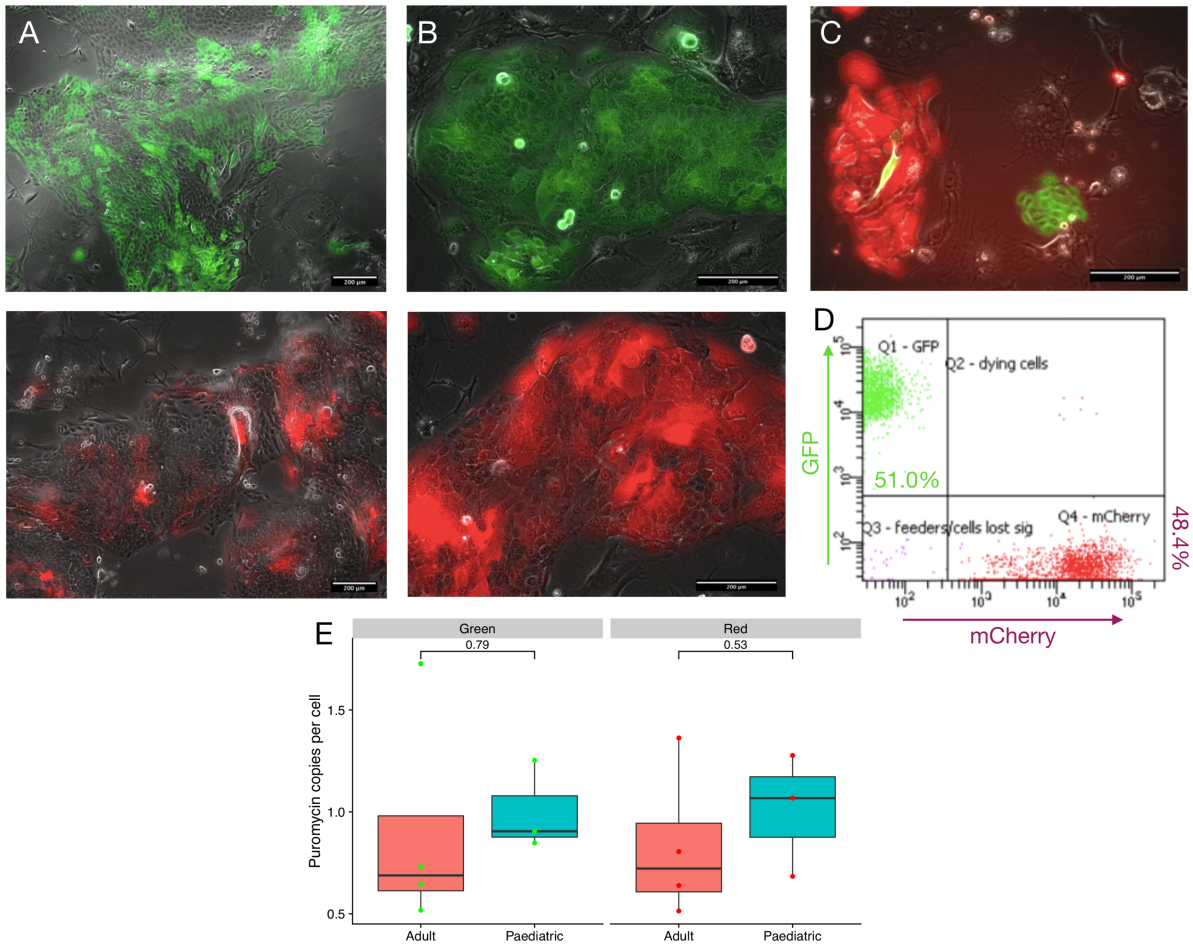

Figure S4: Validation of competitive growth assay approach in primary human proximal airway basal cells.

(A) Immunofluorescence imaging 7 days post-transduction of human basal cells with GFP (above) and mCherry (below) lentiviral constructs. **(B)** Immunofluorescence imaging 10 days after fluorescence-activated cell sorting to purify GFP+ (above) and mCherry+ (below) populations from untransduced cells. **(C)** Immunofluorescence imaging of cultures 4 days after mixing of primary human airway basal cells in a 1:1 ratio. **(D)** Flow cytometry experiment in which GFP+ and mCherry+ cells from the same donor were mixed in equal ratio and analysed after 7 days. **(E)** Copy number determination for the puromycin resistance gene, which is present in both the GFP- and mCherry-containing lentiviruses. There were no differences in viral integration between either paediatric and adult cell cultures or between GFP and mCherry lentiviruses ($p = 0.79$ for GFP, $p = 0.53$ for mCherry, two-tailed unpaired t-test).

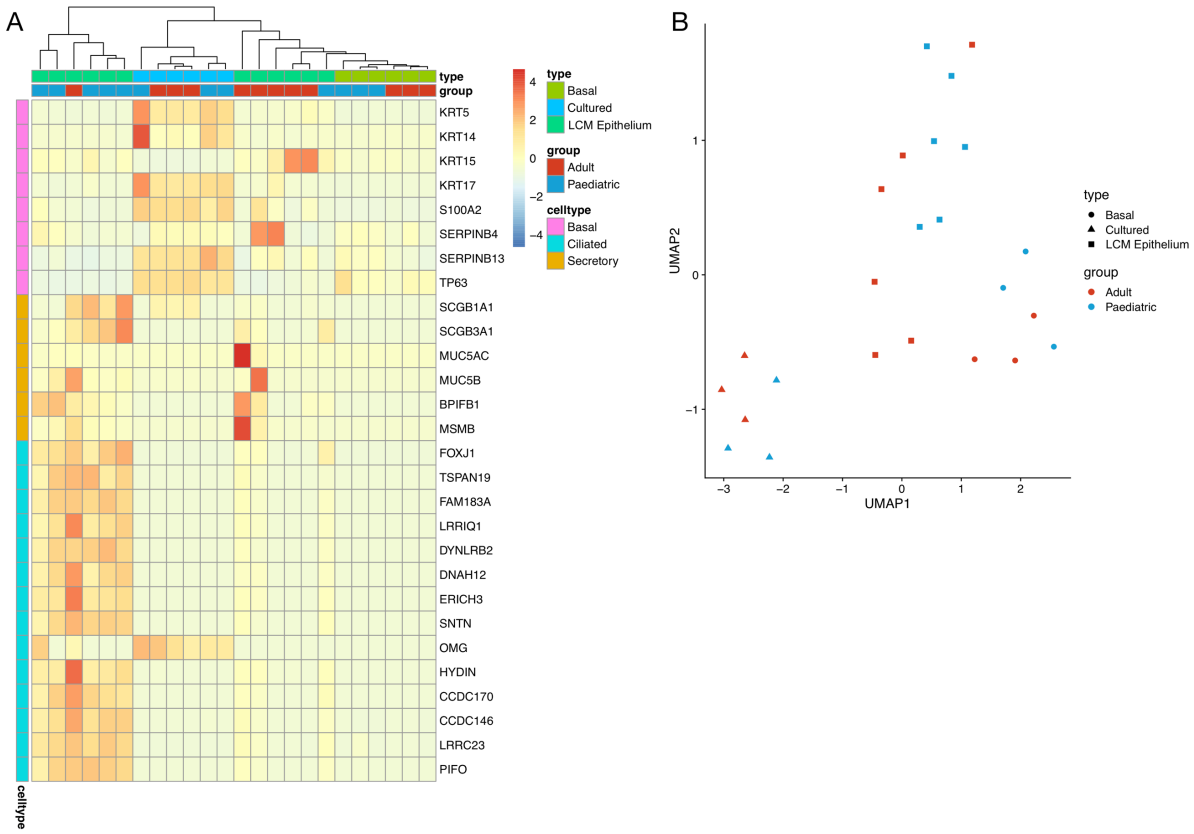

Figure S5: Comparison of three proximal airway epithelial cell RNA sequencing datasets.

(A) Heatmap showing the expression of genes in our basal, mucosecretory and ciliated gene lists (see Table S2) across the three RNA sequencing datasets (laser-capture microdissected whole epithelium, "LCM epithelium"; FACS-sorted EpCAM^{+}/PDPN^{+} basal cells, "sorted"; cultured basal cells, "cultured"). **(B)** UMAP plot visualising laser-capture microdissected whole tracheobronchial epithelium ("LCM epithelium", squares), FACS-sorted basal cell ("basal", circles) and cultured basal cell ("cultured", triangles) datasets. Blue colour indicates paediatric samples and red colour indicates adult samples.

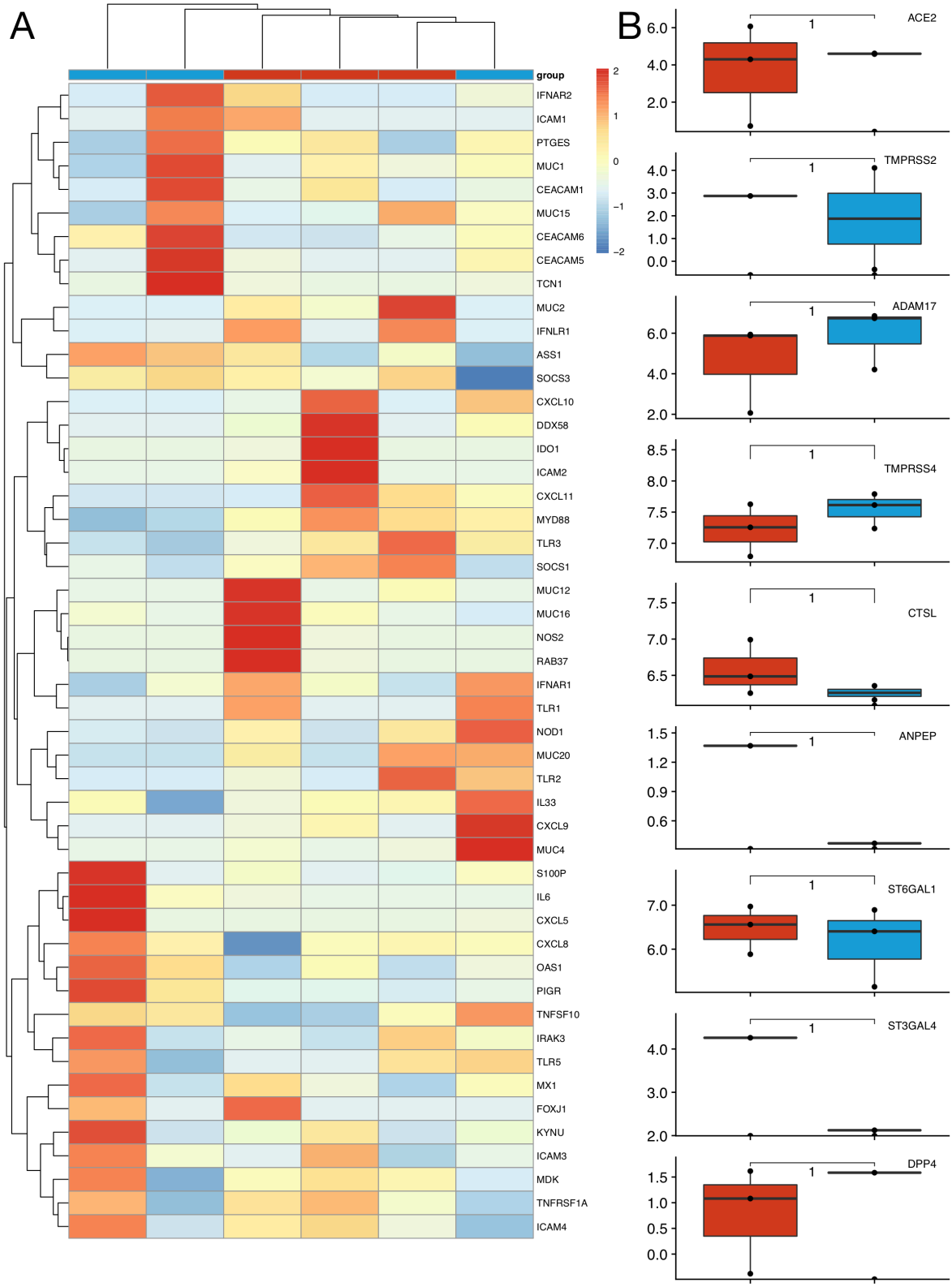
A
group
IFNAR2
ICAM1
PTGES
MUC1
CEACAM1
MUC15
CEACAM6
CEACAM5
TCN1
MUC2
IFNLR1
ASS1
SOCS3
CXCL10
DDX58
IDO1
ICAM2
CXCL11
MYD88
TLR3
SOCS1
MUC12
MUC16
NOS2
RAB37
IFNAR1
TLR1
NOD1
MUC20
TLR2
IL33
CXCL9
MUC4
S100P
IL6
CXCL5
CXCL8
OAS1
PIGR
TNFSF10
IRAK3
TLR5
MX1
FOXJ1
KYNU
ICAM3
MDK
TNFRSF1A
ICAM4
2
1
0
−1
−2
B
ACE2
TMPRSS2
ADAM17
TMPRSS4
CTSL
ANPEP
ST6GAL1
ST3GAL4
DPP4
1

Figure S6: Expression of viral infection-associated epithelial cell genes in paediatric and adult FACS-sorted, non-cultured proximal airway basal cells.

(A) Cluster diagram showing the expression of 49 genes related to viral infection of epithelial cells in the three paediatric and three adult FACS-sorted EpCAM$^+$/PDPN$^+$ basal cell samples. Gene order is based on hierarchical clustering based on the similarity in overall expression patterns. Red represents relative expression higher than the median expression and blue represents lower expression. Genes that are not expressed in this dataset are excluded (see Table S2). **(B)** Plots comparing the expression (log normalized counts) of selected host genes associated with SARS-CoV-2 (*ACE2*, *TMPRSS2*, *CTSL*), other coronavirus (*ADAM17, ANPEP*, *DPP4*) and influenza (*ST6GAL1, ST3GAL4, TMPRSS4*) infection in the three paediatric and three adult FACS-sorted EpCAM$^+$/PDPN$^+$ basal cell samples. All genes analysed were non-significant (FDR > 0.05) when comparing paediatric and adult samples using a Wilcoxon test with correction for multiple testing by the Benjamini-Hochberg method.

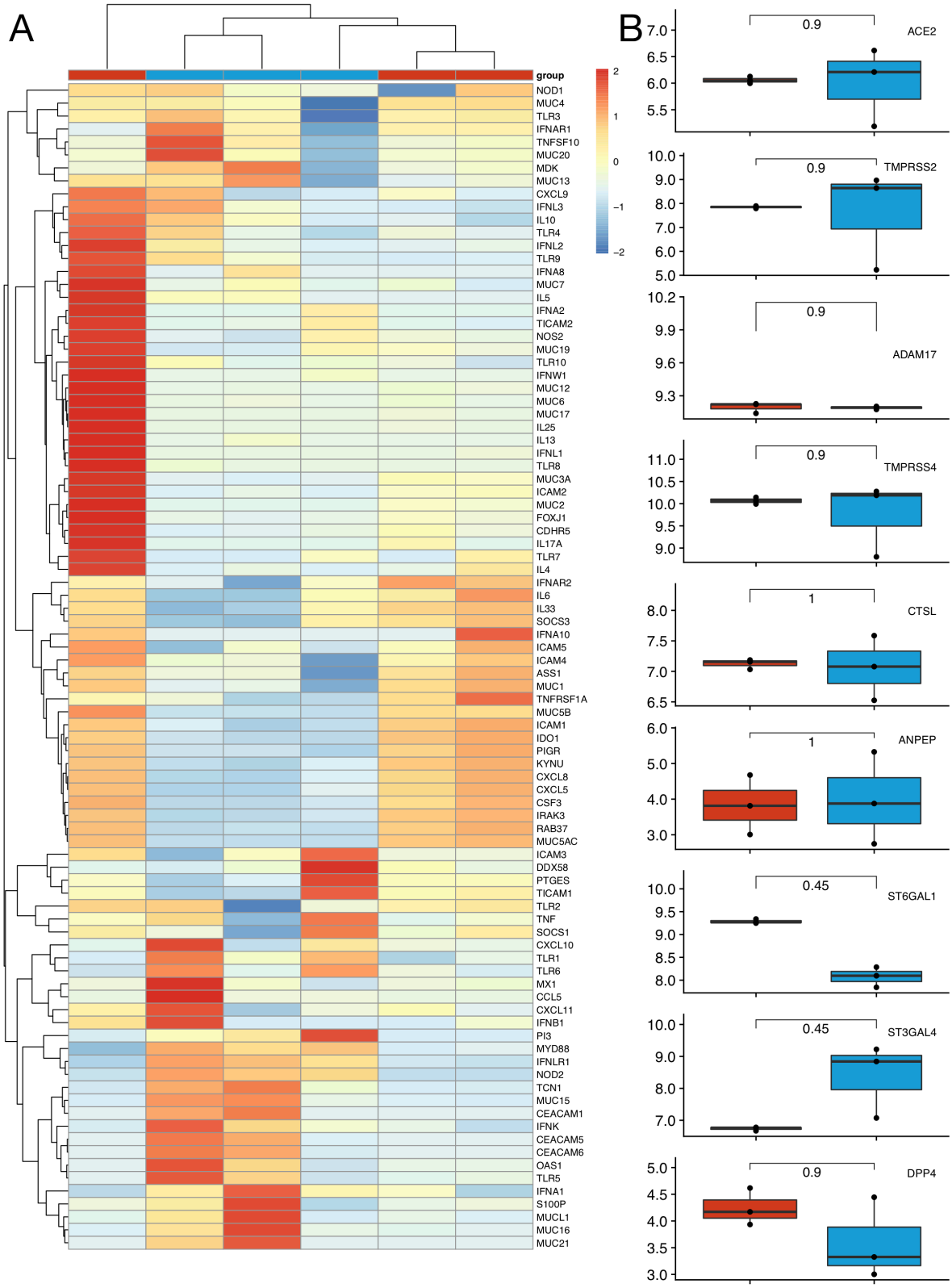
A
group
NOD1
MUC4
TLR3
IFNAR1
TNFSF10
MUC20
MDK
MUC13
CXCL9
IFNL3
IL10
TLR4
IFNL2
TLR9
IFNA8
MUC7
IL5
IFNA2
TICAM2
NOS2
MUC19
TLR10
IFNW1
MUC12
MUC6
MUC17
IL25
IL13
IFNL1
TLR8
MUC3A
ICAM2
MUC2
FOXJ1
CDHR5
IL17A
TLR7
IL4
IFNAR2
IL6
IL33
SOCS3
IFNA10
ICAM5
ICAM4
ASS1
MUC1
TNFRSF1A
MUC5B
ICAM1
IDO1
PIGR
KYNU
CXCL8
CXCL5
CSF3
IRAK3
RAB37
MUC5AC
ICAM3
DDX58
PTGES
TICAM1
TLR2
TNF
SOCS1
CXCL10
TLR1
TLR6
MX1
CCL5
CXCL11
IFNB1
PI3
MYD88
IFNLR1
NOD2
TCN1
MUC15
CEACAM1
IFNK
CEACAM5
CEACAM6
OAS1
TLR5
IFNA1
S100P
MUCL1
MUC16
MUC21
2
1
0
−1
−2
B
ACE2
0.9
TMPRSS2
0.9
ADAM17
0.9
TMPRSS4
0.9
CTSL
1
ANPEP
1
ST6GAL1
0.45
ST3GAL4
0.45
DPP4
0.9

Figure S7: Expression of viral infection-associated epithelial cell genes in paediatric and adult cultured proximal airway basal cells.

(A) Cluster diagram showing the expression of 90 genes related to viral infection of epithelial cells in the three paediatric and three adult cultured basal cell samples. Gene order is based on hierarchical clustering based on the similarity in overall expression patterns. Red represents relative expression higher than the median expression and blue represents lower expression. Genes that are not expressed in this dataset are excluded (see Table S2). **(B)** Plots comparing the expression (log normalized counts) of selected host genes associated with SARS-CoV-2 (*ACE2*, *TMPRSS2*, *CTSL*), other coronavirus (*ADAM17, ANPEP*, *DPP4*) and influenza (*ST6GAL1, ST3GAL4, TMPRSS4*) infection in the three paediatric and three adult cultured basal cell samples. All genes analysed were non-significant (FDR > 0.05) when comparing paediatric and adult samples using a Wilcoxon test with correction for multiple testing by the Benjamini-Hochberg method.

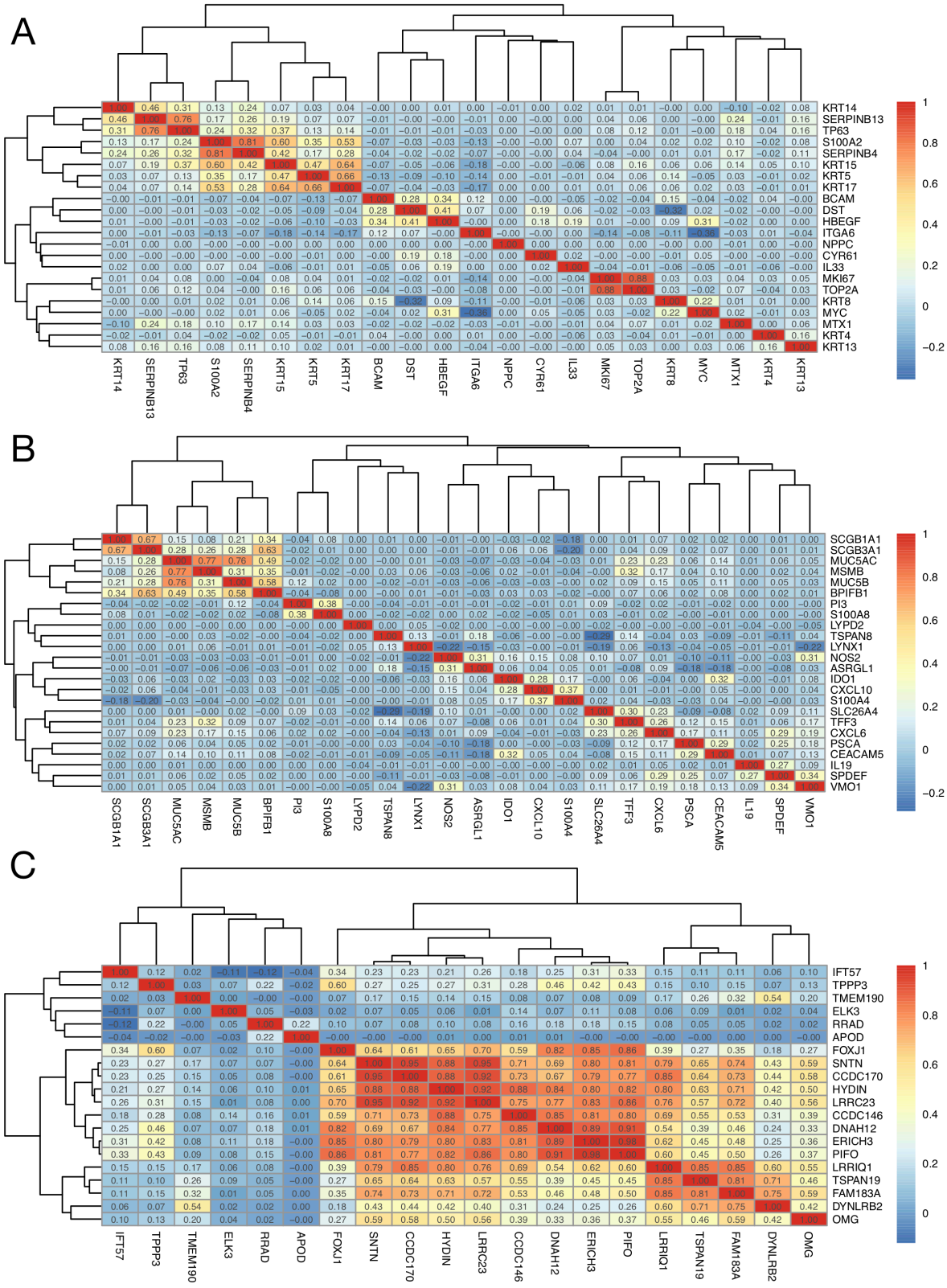
A
KRT14
SERPINB13
TP63
S100A2
SERPINB4
KRT15
KRT5
KRT17
BCAM
DST
HBEGF
ITGA6
NPPC
CYR61
IL33
MKI67
TOP2A
KRT8
MYC
MTX1
KRT4
KRT13
B
SCGB1A1
SCGB3A1
MUC5AC
MSMB
MUC5B
BPIFB1
PI3
S100A8
LYPD2
TSPAN8
LYNX1
NOS2
ASRGL1
IDO1
CXCL10
S100A4
SLC26A4
TFF3
CXCL6
PSCA
CEACAM5
IL19
SPDEF
VMO1
C
IFT57
TPPP3
TMEM190
ELK3
RRAD
APOD
FOXJ1
SNTN
CCDC170
HYDIN
LRRC23
CCDC146
DNAH12
ERICH3
PIFO
LRRIQ1
TSPAN19
FAM183A
DYNLRB2
OMG

Figure S8: Refining tracheobronchial cell type-specific gene lists by co-correlation.

(A) A similarity matrix, based on the method of Danaher et al. (Danaher et al., 2017), showing co-correlation of the expression of a manually curated list of genes expected to be expressed by airway basal cells (see Table S2) in the Gene Tissue Expression Project normal lung RNA sequencing data (GTEx Consortium, 2015). Hierarchical clustering revealed a sub-cluster of genes whose expression was correlated. These were taken forward for further analyses as cell type-specific gene signatures and applied to our proximal airway epithelial cell RNA sequencing datasets. **(B)** As for (A) but using a manually curated list of genes expected to be expressed by airway mucosecretory cells. **(C)** As for (A) and (B) but using a manually curated list of genes expected to be expressed by airway ciliated cells.