Single cell profiling of human airway identifies tuft-ionocyte progenitor cells displaying cytokine-dependent differentiation bias in vitro

Viral S. Shah[1,2,3,4,13], Avinash Waghray[1,3,4,5,13], Brian Lin [1,3,4,6,13], Atharva Bhagwat [7,13], Isha Monga [7,13], Michal Slyper[8,9,13], Bruno Giotti[7,13], Sunghyun Kim[1,3], Dawei Sun[1,4], Ke Xu [7], Eric Park[7], Mohamad Bairakdar[7], Jiajie Xu[1,3], Julia Waldman[8], Danielle Dionne [8], Lan T. Nguyen[8], Wendy Lou[8], Peiwen Cai [7], Christoph Muus[8,10], Jiawei Sun[1,3], Manalee V. Surve[1,3,4], Lujia Cha Cha Yang[11], Orit Rozenblatt-Rosen [8], Toni M. Delorey [8], Srinivas Vinod Saladi [4,12], Aviv Regev [8,9,14] ✉, Jayaraj Rajagopal [1,2,3,4,8,14] ✉ & Alexander M. Tsankov [7,14] ✉

Human airways contain specialized rare epithelial cells including CFTR-rich ionocytes that regulate airway surface physiology and chemosensory tuft cells that produce asthma-associated inflammatory mediators. Here, using a lung cell atlas of 311,748 single cell RNA-Seq profiles, we identify 687 ionocytes (0.45%). In contrast to prior reports claiming a lack of ionocytes in the small airways, we demonstrate that ionocytes are present in small and large airways in similar proportions. Surprisingly, we find only 3 mature tuft cells (0.002%), and demonstrate that previously annotated tuft-like cells are instead highly replicative progenitor cells. These tuft-ionocyte progenitor (TIP) cells produce ionocytes as a default lineage. However, Type 2 and Type 17 cytokines divert TIP cell lineage in vitro, resulting in the production of mature tuft cells at the expense of ionocyte differentiation. Our dataset thus provides an updated understanding of airway rare cell composition, and further suggests that clinically relevant cytokines may skew the composition of disease-relevant rare cells.

Single cell profiling studies in human and mouse airways have revealed the existence of a triad of rare cells: ionocytes, tuft, and neuroendocrine (NE) cells[1–11]. Ionocytes express the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR), the causal gene whose disruption leads to cystic fibrosis (CF)[6,7]. Ionocyte ablation in a ferret model of CF resulted in abnormalities of airway physiology remarkably reminiscent of CF[12], and their loss in human large airway epithelial cultures disrupts ion transport[13,14]. In contrast, tuft cells and NE cells

have roles in environmental sensing and initiating inflammatory responses[7,15].

Alterations in tuft cell, neuroendocrine cell, and ionocyte numbers have been associated with diseases of the airway, but the mechanisms governing their abundance are poorly understood[13,16–19]. However, very few tuft cells have been identified in single-cell atlases of the normal human lung[5,20]. Tuft cells have been associated with Type 2 inflammation in the intestinal[21–24] and nasal[17] epithelia, but

surprisingly few tuft cells have been identified in asthma atlases[20,25]. Analogously, ionocytes have been reported to be absent in small airways, despite the prominent small airway pathology seen in CF[6,7,26]. This has led to the suggestion that ionocytes may not play a significant role in the small airways pathology characteristic of the early stages of CF[26].

In order to understand the underlying mechanisms governing the distribution and differentiation of lung ionocytes and tuft cells, we construct an anatomically-guided, deep single-cell RNA-Seq (scRNA-seq) atlas of 311,748 cells from 8 donor lungs. We identify a highly replicating bipotent progenitor cell population that gives rise to either tuft cells or ionocytes, hereafter referred to as Tuft-Ionocyte Progenitor (TIP) cells. We find that TIP cells are present in both the large and small airways. Consistently, in contrast to prior reports, we demonstrate the presence of ionocytes in the small airways, which suggests that ionocytes may indeed play a role in the small airway pathology associated with CF. We then demonstrate that both Type 2 and Type 17 cytokines skew TIP cell lineage and promote the differentiation of mature tuft cells at the expense of ionocytes, providing one explanation for the ionocyte loss recently reported in patients with both eosinophilic and non-eosinophilic asthma[18,19,27]. Thus, in addition to the classic airway basal stem cell that sits atop the airway epithelial hierarchy, we identify a second highly replicative human progenitor cell population that is dedicated to the production of rare epithelial cells. Furthermore, cytokine-mediated skewing of TIP cell lineage may provides a framework for re-examining alterations of rare cell numbers in airways disease.

Results

Deep cell atlas of the human lung

To build a deep human lung cell atlas, we profiled 311,748 cells from 8 normal explanted human lungs, from individuals of both genders (5 females, 3 males), a broad age range (3 months to 66 years old) and different ancestries (6 white, one African American, and one Hispanic/Latino), representing the largest single dataset of normal human lung cells profiled to date (Fig. 1A, B and Supplementary Figs. 1–3, Supplementary Data 1). From one individual (HU37), we profiled 71,263 single cells across 13 anatomic sites along the respiratory tree, including 6 proximal airway sites and 7 distal lung lobe sites (Fig. 1A, B and Supplementary Data 1). We complemented this dataset with cells obtained from another 7 donors to account for inter-individual variability (Supplementary Figs. 1–3, Supplementary Data 1). The cells partitioned into 17 epithelial (Fig. 1C, D and Fig. S1), 9 stromal/mesenchymal (Supplementary Fig. 2A–D), 5 endothelial (Supplementary Fig. 2E–H), and 32 immune (Supplementary Fig. 3) subsets, which were independently annotated and then matched to the Human Lung Cell Atlas (HLCA) ontology[20] (**Methods**).

In the proximal airway we identified all of the expected common epithelial cell types (basal, secretory, goblet, deutrosomal[10], and ciliated cells), along with cells from the glands (Supplementary Fig. 1A), and the three previously identified rare cell populations: neuroendocrine (NE) cells, ionocytes, and *POU2F3*[+] cells (Fig. 1B). Previously, these rare *POU2F3*[+] cells had been annotated as "tuft-like" cells[1,20], in part because *POU2F3* is required for tuft cell differentiation and maintenance[22,28]. In the distal region of the lung, we again found all the expected epithelial cell types, including alveolar type 2 (AT2) cells, alveolar type 1 (AT1) cells, mesothelial cells, as well as small airway epithelial cell types, including basal, secretory, goblet, and ciliated cells. Additionally, in alveolar regions we detected the previously described secretory cell subsets: secretory-ciliated (SecCil) and secretory-AT2 (SecAT2) cells[29,30] (Fig. 1D and Supplementary Fig. 1B). Despite the anatomical sampling and depth of our cell atlas, we did not detect rare airway epithelial cell types in the small airways, in contrast to the large airways, possibly due to the predominance of alveolar type 1 and type 2 cells.

Mature tuft cells are exceedingly rare in normal human lungs and in vitro epithelial models

In order to determine the abundance of mature tuft cells in human lungs, which we did not detect de novo in our deep scRNA-seq atlas, we leveraged the HLCA version 1 core dataset composed of harmonized data from 9 independent studies[31] to re-clustered the 885 annotated rare epithelial cells (Supplementary Fig. 4A–C). We used supervised scoring to discriminate NE cells, ionocytes, *POU2F3*[+] "tuft-like" cells, and mature tuft cells, using markers derived from the dataset[5] that contained the greatest number of mature tuft cells (10 cells) in a single study (**Methods**, Supplementary Data 2). Using this approach, we confirmed that the HLCA core dataset contains 561 ionocytes (0.096% of all cells; 0.199% of all epithelial cells) and 159 NE cells (0.027% of all cells; 0.056% of all epithelial cells) (Supplementary Fig. 4D, Supplementary Data 3). However, the 165 cells that were annotated as tuft cells in HLCA, were, in fact, composed of two subsets: A major subset of 152 *POU2F3*[+] "tuft-like" cells (0.026% of all cells; 0.054% of all epithelial cells), and a very small subset of 13 mature tuft cells (0.002% of all cells; 0.005% of all epithelial cells, Supplementary Fig. 4D, Supplementary Data 3).

To identify rare epithelial cells that may not have been detected in the HCLA extended data set, we used our scoring approach to detect 1829 rare epithelial cells (compared to 1649 previously annotated cells by HLCA[31]; Supplementary Fig. 4E, Supplementary Data 3). Next, we derived a more generalizable tuft cell signature based on rare cells from multiple studies (Supplementary Fig. 4F, Supplementary Data 2) and identified a total of 56 high confidence mature tuft cells. For ionocytes and neuroendocrine cells, we observed high correspondence between our de novo annotations and HCLA metadata labels (Supplementary Fig. 4G). In contrast, *POU2F3*[+] "tuft-like" cells were predominantly annotated as "tuft" (71%) or "Other" (27%), while most mature tuft cells were mislabeled as "Other" (68%) in HLCA (Supplementary Fig. 4G). Rare epithelial cells in HLCA were present across multiple anatomical regions, including the nose, respiratory airways, and distal lung lobes (Supplementary Fig. 4G, **right**). As noted before, mature tuft cells identified using our supervised approach were also extremely rare in the extended HLCA dataset (0.004% of all cells; 0.01% of all epithelial cells, Supplementary Data 3).

Applying the same scoring approach to our own deep atlas revealed 687 ionocytes (0.22% of all cells; 0.45% of airway epithelial cells), 363 NE cells (0.12% of all cells; 0.24% of airway epithelial cells), 725 *POU2F3*[+] "tuft-like" cells (0.23% of all cells; 0.47% of airway epithelial cells), and only 3 mature tuft cells (0.001% of all cells; 0.002% of airway epithelial cells; Supplementary Data 4). Thus, despite our very deep sampling of rare cells, mature tuft cells were indeed exceedingly rare in the normal human lung. These rare tuft cells do not form a distinct cluster using standard unsupervised community detection algorithms (Supplementary Fig. 4H). In total, our deep atlas added 1778 rare cell profiles to the 1649 profiles contained in HLCA, and our reanalysis added another 180 rare cells mislabeled as "Other" in HLCA and refined the annotation for another 128 cells, establishing that mature tuft cells are a remarkably scarce fraction (<0.02%) of an already scarce total rare epithelial cell pool.

Given the exceptionally low numbers of rare epithelial cells recovered from dissociated distal lungs and the predominance of alveoli in the tissue sample, we next compared the proportions of rare cells in epithelial air-liquid interface (ALI) cultures generated from large and microdissected small airway (LAE and SAE ALI, respectively). In contrast to prior reports, which reported that SAE ALI cultures contain fewer ionocytes than LAE ALIs[26], our scRNA-seq of LAE *vs*. SAE ALI cultures detected comparable numbers of ionocytes on average (per sampled ALI, **Methods**; Fig. 1E–G). Notably, mature tuft cells were not detected in either LAE or SAE ALI cultures.

We hypothesized that those physiological activities that are normally ascribed to mature tuft cells, such as environmental sensing, may be assumed by other cell types. As a precedent, genes related to bitter
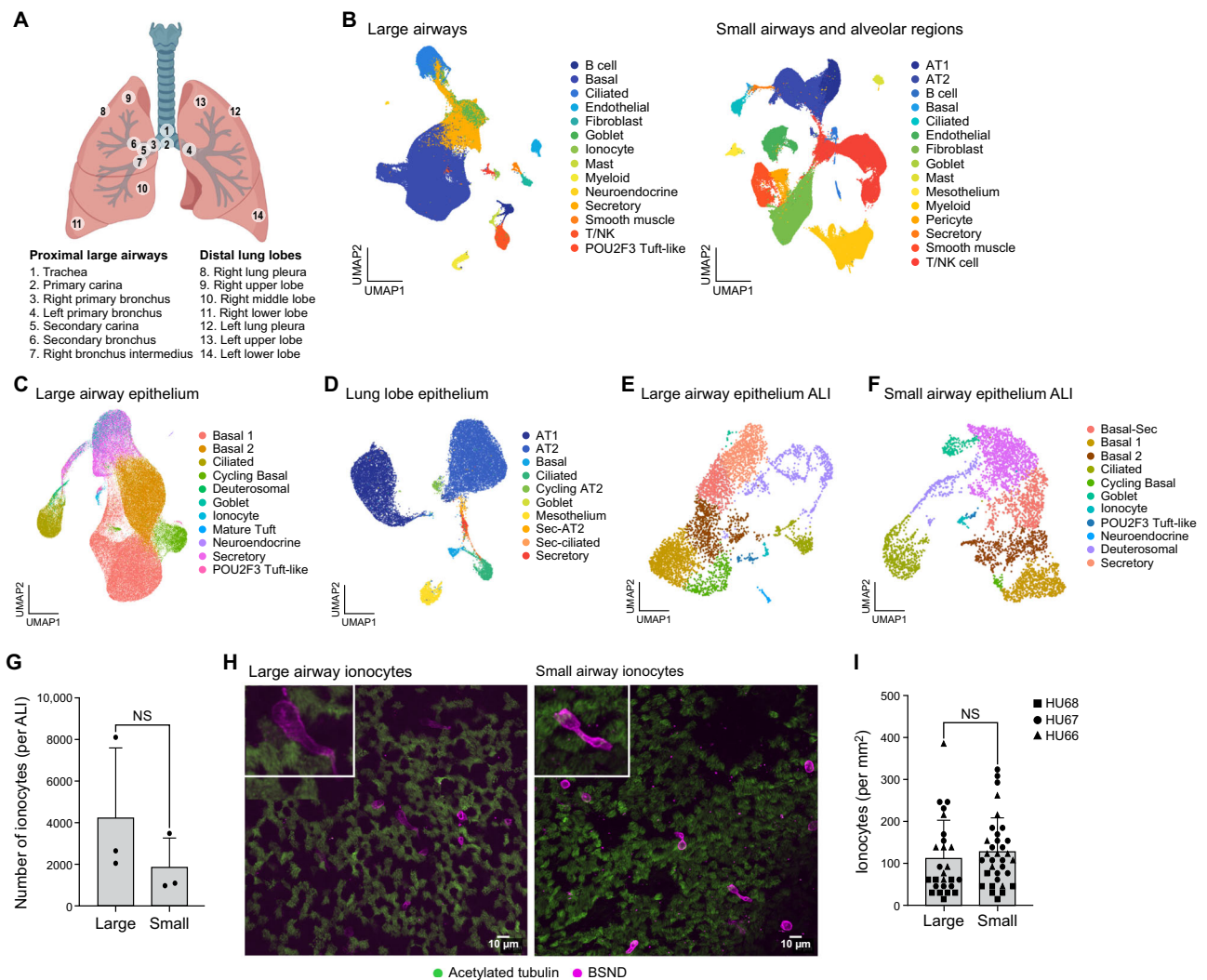
Fig. 1 | Deep cell atlas of the human lung reveals the presence of ionocytes in both human proximal and distal airways and proximal and distal ALI cultures. **A** Regional sampling for a deep lung cell atlas. Numbered circles represent sampled locations. **B** Lung cell atlas. Uniform manifold approximation and projection (UMAP) embedding of cell profiles (dots) from the large airways (left) and lung lobe regions (right) colored by cell type annotation. **C–F** Epithelial lung and ALI cell profiles. UMAP embeddings of epithelial cell profiles from the proximal airway (**C**), distal lung lobe (**D**), and ALI cultures generated from large airway basal cells isolated from primary bronchus (**E**) or from small airway basal cells isolated from microdissected small airway less than 2 mm in diameter (**F**). **G–I** Ionocyte abundance in the human proximal and distal airways and human proximal and distal ALI cultures. **G** Number of BSND+ mature ionocytes per ALI (y axis) in Large ALI and Small ALI cultures (x axis). $n = 3$ ALIs averaged from 3 separate donors, Two tailed unpaired T test. Error bars are standard deviation. **H** Whole mount images of dissected large (left) and small (right) airways stained for BSND (magenta) and acetylated Tubulin (green). Insets: Representative examples of BSND+ ionocytes (magenta). **I** Number of BSND+ mature ionocytes per mm2 (y axis) in microdissected large airways and small airways (x axis). $n = 26$ for large airways and 33 for small airways across three normal human lungs (Hu66, Hu67, and Hu68). One way ANOVA (Sidak's multiple comparisons). Error bars are standard deviation. Elements of 1 A was created with BioRender https://BioRender.com/0ruztco.

taste receptors (*TAS2R4*, *TAS2R43*) are expressed by human ciliated cells[32], genes related to olfactory receptor-mediated functions are expressed in NE[33] cells, and genes for succinate sensing (e.g., succinate receptor (*SUCNR1*)) are expressed in glandular NE cells[34]. Indeed, NE cells from our deep cell atlas and human LAE cultures expressed the succinate receptor (*SUCNR1*), whereas *POU2F3*+ "tuft-like" cells did not (Supplementary Fig. 5A, B). We confirmed SUCNR1 receptor protein expression in NE (TUJI+) cells in human LAE ALI culture in situ (Supplementary Fig. 5C) and further established that NE cells functionally respond to succinate with a rise in intracellular calcium (Supplementary Fig. 5D).

CFTR-rich ionocytes are comparably distributed in large and small airways in situ

Although CF is associated with small airways disease, prior reports claimed that small airways and SAE cultures contain few ionocytes[26,34],

leading to the suggestion that ionocytes do not play a role in CF-associated small airway pathology. In contrast, our LAE and SAE cultures contained a similar number of ionocytes (Fig. 1G). We further found that small airway ionocyte abundance, in particular, depended on the choice of culture media (Supplementary Fig. 6A). Consistent with these findings, a recent report demonstrates that media choice would account for the previously reported lack of small airway ionocyte differentiation in vitro[26,35]. In our cultures, both large airway and small airway ionocytes resemble their in vivo counterparts and comparably express ionocyte transcription factors (TFs) *FOXI1* and *ASCL3*[7], *CFTR*, and the CF-associated ion transporters *ATP1A1*, *ATP6VOD2* and *ATP6V1C2* (Supplementary Fig. 6B). Finally, to directly quantify the numbers of sparsely distributed ionocytes in situ we performed whole mount staining of microdissected airways[26] from three donors (Supplementary Fig. 6C, D, E)[26]. We detected mature ionocytes expressing BSND protein in both large and small airways (less than 2 mm in
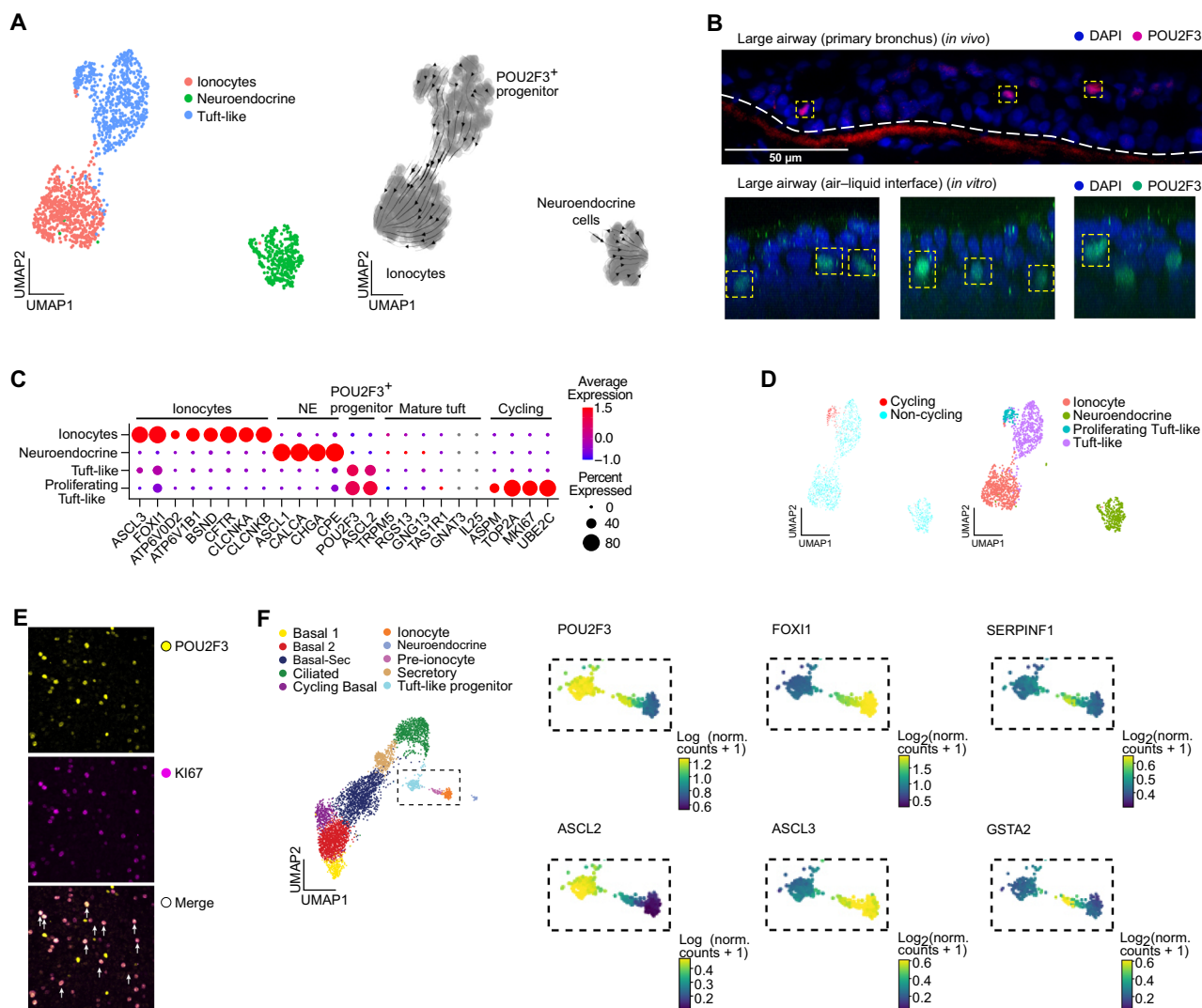
Fig. 2 | Transcriptional and chromatin accessibility profiles reveal a replicative rare cell progenitor and a pre-ionocyte state. A *POU2F3*+ tuft-like cells are predicted to be progenitors of mature ionocytes. UMAP embedding of scRNA-seq profiles (dots) of rare epithelial cells in our deep lung cell atlas, colored by cell annotation (left) and showing RNA velocity vectors (right) directed from tuft-like cells to ionocytes. **B** Human large airways and human Large ALI cultures both contain POU2F3+ cells. Antibody staining of a section of the right primary bronchus for POU2F3 (red) and DAPI (blue) and LAE ALI for POU2F3 (green) and DAPI (blue). This staining has been repeated in 3 separate samples. **C**–**E** *POU2F3*+ tuft-like cells include replicating and non-replicating cells. **C** Mean expression (dot color, relative diameter) (Fig. 1H) in similar numbers (Fig. 1I). Thus, small airway

expression) and percentage of cells (dot size) expressing selected cell identity and cell proliferation markers (columns) in different rare epithelial cell subsets (rows). **D** UMAP embedding of scRNA-seq profiles (dots) of rare cells, colored by cell cycle classification (left) or cell type annotation (right). **E** Large ALI cultures co-stained with POU2F3 (yellow, top) and MKI67 (purple, middle). The bottom panel shows cells expressing both markers (arrows). This staining was repeated in 3 samples. **F** Distinct chromatin state marks POU2F3+ progenitors. UMAP embedding of large airways epithelial cell scATAC-seq profiles (dots) colored by de novo cell type annotation. Zoom of boxed rare cells highlight chromatin accessibility at select gene loci associated with tuft cells, ionocytes, and progenitor cells.

diameter) (Fig. 1H) in similar numbers (Fig. 1I). Thus, small airway ionocytes expressing CF-related genes may yet play a role in CF-associated small airways disease[12].

Ionocytes differentiate from replicating *POU2F3*+ progenitor cells

Despite their prior annotation as "tuft-like" cells, RNA velocity analysis[36,37] on rare epithelial cells from our lung donors predicts that *POU2F3*+ progenitors give rise to ionocytes (Fig. 2A), consistent with prior observations that *POU2F3* knockout results in the loss of ionocyte differentiation in large airway ALI cultures[1]. We verified that POU2F3+ "tuft-like" cells were present in sections of primary bronchus and large airway ALI cultures (Fig. 2B). The *POU2F3*+ "tuft-like" cells express both ionocyte (*FOXI1*) and tuft cell (*POU2F3, ASCL2*) TFs (Fig. 2C), and do not express markers of mature tuft cells (*TRPM5, GNG13, RGS13, TAS1R1, GNAT3, IL25*), ionocytes (*BSND, CFTR, ATP6V0D2, ATP6V1B1, CLCNKA,*

CLCNKB), or NE cells (*ASCL1, CALCA, CHGA, CPE*) (Fig. 2C). Taken together, these findings suggest that *POU2F3*+ cells, previously annotated as "tuft-like" cells[1], act in fact as ionocyte progenitors during the course of normal homeostasis. RNA velocity analysis of cells profiled from SAE *vs.* LAE cultures also argue that both small and large airway ionocytes arise from *POU2F3*+ progenitor cells (Supplementary Fig. 6F).

Notably, a large fraction of human airway *POU2F3*+ progenitor cells were cycling in vivo (19%), expressing canonical markers of cell proliferation: *MKI67, ASPM, TOP2A,* and *UBE2C* (Fig. 2C, D). Remarkably, the cycling fraction of *POU2F3*+ progenitor cells is greater than the cycling fraction of canonical basal stem cells (11%) in in vivo human large airways (Supplementary Data 1). We confirmed the presence of proliferative *POU2F3*+ progenitor cells in ALI cultures (Fig. 2E), and demonstrated that they also express *ASCL2*, mimicking their in vivo counterparts (Supplementary Fig. 6G). Interestingly, a fraction of

POU2F3⁺ cells start to express the ionocyte TF *FOXI1* (Fig. 2C and Supplementary Fig. 6G). However, they do not express the full complement of mature ionocyte markers (e.g., *BSND, CLCNKA, ATP6VOD2*; Fig. 2C, Supplementary Fig. 6G).

Chromatin accessibility profiles in the lung distinguish a pre-ionocyte state

To delineate the chromatin accessibility landscape underpinning rare lung epithelial cell differentiation, we profiled 18,260 cells from lung donor Hu62 by single-cell Assay for Transposase-Accessible Chromatin with sequencing (scATAC-seq). Using unsupervised clustering, we identified 12 major large airway cell clusters (Supplementary Fig. 7A, B), including NE cells, ionocytes, and *POU2F3*⁺ progenitor cells (Fig. 2F), but not the exceedingly rare mature tuft cells (Supplementary Data 5). We also detected 12 major cell clusters in distal lung lobe regions that corresponded to expected cell classes typically found in the lung parenchyma (Supplementary Fig. 7C, D). Next, we de novo identified TF motifs enriched in chromatin accessible regions for each cell cluster, including *TP63* in basal and *POU2F3* in "tuft-like" cells, suggesting cell type specific regulatory activity for the corresponding TFs (Supplementary Fig. 7E, F; ChromVar[38] FDR ≤ 0.01, mean differential accessibility ≥0.5). We further defined differentially accessible regions across the annotated cell clusters that can be used for cell type specific enhancer discovery and the construction of cell type specific reporters (Supplementary Data 6).

Interestingly, *POU2F3*⁺ cells de novo partitioned into two clusters based on their scATAC-seq profiles: a '*POU2F3*⁺ progenitor' cluster and a 'pre-ionocyte' cluster (Fig. 2F). As expected, the *POU2F3*⁺ progenitors are characterized by high accessibility at the *POU2F3* and *ASCL2* loci (Fig. 2F **inset**, Supplementary Fig. 7G), suggesting a competence to differentiate into mature tuft cells. However, accessibility at the *POU2F3* and *ASCL2* loci gradually decreases as *POU2F3*⁺ progenitors transition into pre-ionocytes and then into ionocytes, alongside a concomitantly increasing accessibility at ionocyte TFs *FOXI1* and *ASCL3* loci (Fig. 2F **inset**, Supplementary Fig. 7G). This is consistent with commitment towards an ionocyte fate, prior to the opening of loci associated with mature ionocytes[39]. Indeed, while some mature ionocyte gene loci (*CLCNKB, PTGER3*) remain largely closed in pre-ionocytes, chromatin at other ionocyte loci (*BSND*) has already opened in some cells, and the *CFTR* locus is already decidedly open (Supplementary Fig. 7H). Interestingly, pre-ionocytes are also marked by unique open loci for genes such as *SERPINF1* and *GSTA2* that are not accessible in either *POU2F3*⁺ progenitor cells or mature ionocytes (Fig. 2F, **right**).

The type 2 cytokines IL-13 and IL-4 induce mature tuft cell differentiation

In order to assess whether we could identify pathways that govern the differentiation of *POU2F3*⁺ progenitors, we performed gene set enrichment analysis of differentially expressed genes (DEGs) between mature tuft cells and ionocytes detected in the HLCA (Supplementary Data 7). Unsurprisingly, ion transport was associated with ionocytes and neurotransmission was associated with mature tuft cell DEGs (Fig. 3A). Interestingly, interleukin signaling was the second most enriched REACTOME pathway for mature tuft cells (Fig. 3A). This finding is consistent with the documented role of IL-13 in tuft cell expansion in human and murine nasal epithelial cultures[40,41] and the presence of tuft cells in human allergic rhinosinusitis[17,41–45].

To assess the role of Type 2 cytokines on human rare cell differentiation, we treated mature large airway epithelial (LAE) ALI cultures with IL13 (20 ng/μl) for 96 h and then performed scRNA-seq. As expected[46], IL13 treatment resulted in the differentiation of secretory cells into goblet cells (Fig. 3B and Supplementary Fig. 8A). We continued to be able to identify *POU2F3*⁺ progenitors, ionocytes, and NE cells, but we also identified an additional new cluster of

TRPM5⁺SOX9⁺POU2F3⁺ mature tuft cells (Fig. 3C, D), expressing markers of both murine tuft 1[7] (chemosensory) airway epithelial cells (*GNAT3, TAS1R3*) and murine tuft 2[7] (leukotriene) (*ALOX5, ALOX5AP*) cells (Fig. 3D). In further support of a model in which Type 2 cytokines induce bipotent progenitor cells to differentiate into mature tuft cells, IL13 application results in the nascent expression of mature tuft cell genes (*TRPM5, GNAT3, ALOX5*) within the progenitor cluster itself (Fig. 3D). To verify that *POU2F3*⁺ progenitor cells had matured into bona fide tuft cells, we performed immunohistochemistry for the mature tuft cell marker GNAT3. IL13 treatment resulted in a significant increase in GNAT3⁺ cells, while the number of ionocytes (BSND⁺) remained unchanged (Fig. 3E).

IL4, another Type 2 pathway cytokine, also induces the expression of mature tuft cell genes (*TRPM5, GNAT3*) in LAE ALI cultures, as measured by qPCR of entire cultures, whereas treatment with the Type 2 cytokine IL5 did not show a pronounced effect (Supplementary Fig. 8B). The lack of an effect of IL5 is consistent with the lack of expression of the IL5 receptor (*IL5RA*) in progenitor cells; *IL5RA* expression is restricted to ciliated and deutrosomal cells (Supplementary Fig. 8C). Treatment of small airway epithelial (SAE) ALI cultures with IL13 (20 ng/μl) for 96 h also resulted in the differentiation of mature tuft cells, where we performed scRNA-seq as a readout (Fig. 3F, G).

Type 2 cytokines induce a lineage switch in *POU2F3*⁺ bipotent tuft-ionocyte progenitor (TIP) cells

Although *POU2F3*⁺ bipotent progenitors express both tuft cell TFs, *POU2F3* and *ASCL2*, we have shown above that at steady state, they normally give rise to ionocytes (Fig. 2A, C). Consistent with this ionocyte-biased lineage, chromatin accessibility at the *POU2F3* and *ASCL2* loci gradually decreases in pre-ionocytes, as *FOXI1* accessibility increases (Fig. 2F, Supplementary Fig. 7G). In order to define the lineage trajectory of these bipotent progenitors when they are exposed to IL13, we developed an enrichment strategy in order to obtain sufficient numbers of mature tuft cells to computationally infer a new lineage path, by sorting for CD45⁻CD31⁻EPCAM⁺KIT⁺NCAM1⁺ cells (Fig. 4A and Supplementary Fig. 9), which enriched for cells expressing *POU2F3* and *FOXI1* (Fig. 4B). Next, in order to specifically capture lineage transitions, we focused on differentiating epithelial cells. Because *POU2F3* expression first begins to increase on day 2 of ALI epithelial differentiation, we treated day 3 differentiating cultures with IL13 (10 ng/μl) and profiled cells using scRNA-seq at days 5, 6, and 8 (Fig. 4C, D). Following profiling, we detected all the expected common cells as well as differentiating rare cells (Supplementary Fig. 10A, B).

In aggregate, we obtained 572 rare cells from cultured control and IL-13-treated epithelial cells, and they partitioned into five rare cell clusters (Supplementary Fig. 10C, D). We next separately clustered the control cells (Fig. 4E–H) and the cells that were exposed to IL13 (Fig. 4I–L). These included cycling *POU2F3*⁺ progenitor cells enriched for markers of active proliferation (Fig. 4E–L) and non-cycling *POU2F3*⁺ progenitor cells that express *ASCL2, POU2F3*, and *FOXI1*, but not the ionocyte-specific TF *ASCL3* (Fig. 4E, H, I, L). We also identified *ASCL3*⁺ pre-ionocytes (Fig. 4E, H, I, L); mature ionocytes that further express *BSND, CFTR, ATP6V1C2, ATP6V0A4*, and *CLCNKB* (Fig. 4E, H, I, L); and mature tuft cells expressing a full suite of functional tuft cell genes (*IL25, CHAT, GNAT3, GNB3, GNG13, TAS1R1, TAS1R3, ALOX5, ALOX5AP*) (Fig. 4E, H, I, L).

Based on RNA velocity analysis, during differentiation, *POU2F3*⁺ progenitors once again give rise to ionocytes, as expected by the default lineage path (Fig. 4E, F). Following IL13 treatment, *POU2F3*⁺ progenitors now also differentiate into mature tuft cells (Fig. 4I, J). Having established that *POU2F3*⁺ progenitor cells give rise to both mature tuft and ionocyte lineages depending on the signaling milieu, we will hereafter refer to these cells as Tuft Ionocyte Progenitor cells (TIP cells). In ALI culture, many of these TIP cells proliferate, whereas
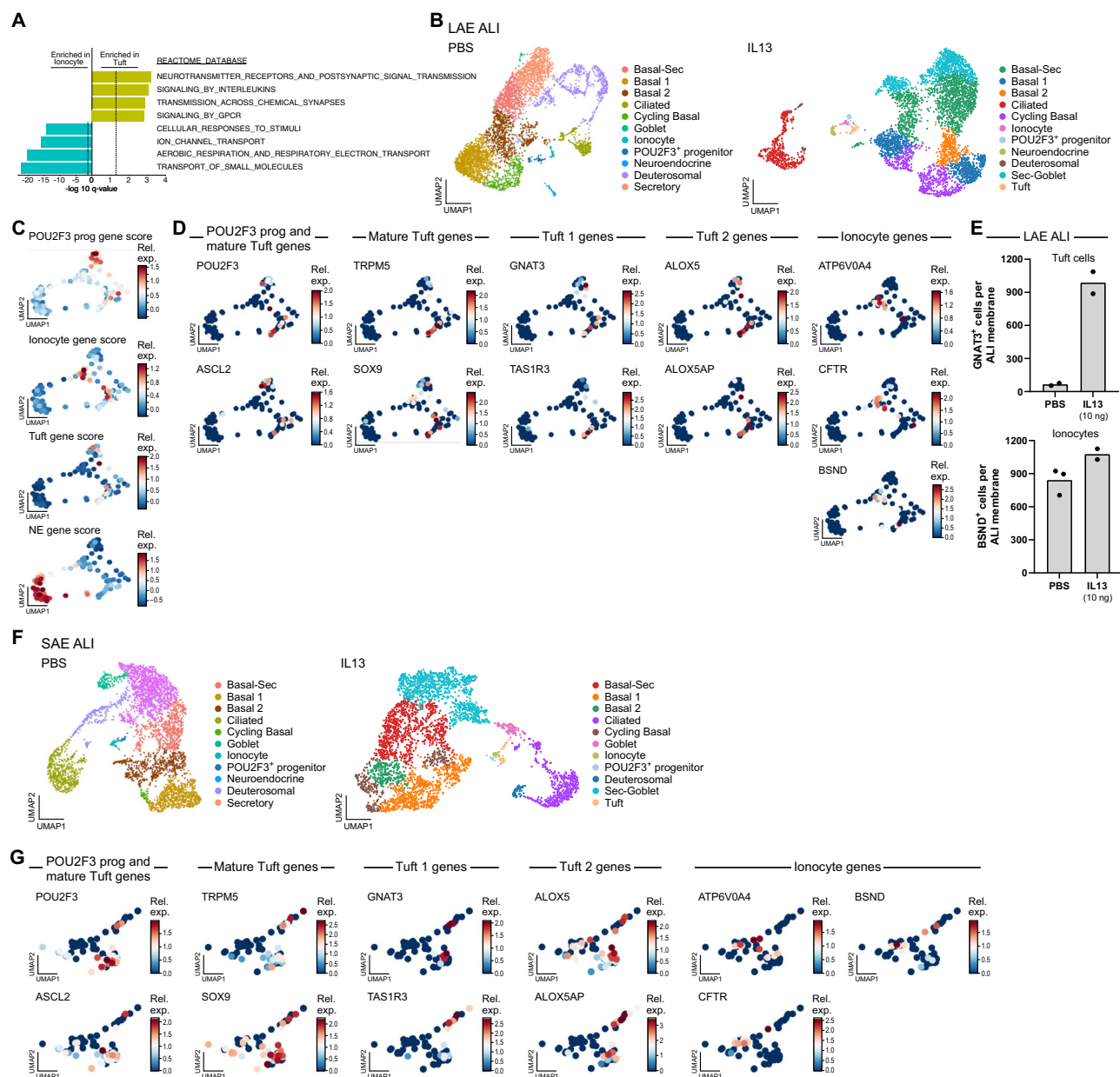
Fig. 3 | Type 2 pathway cytokines induce mature tuft cell differentiation.
A Neural and immune genes are induced in mature tuft cells. Significance (signed-log10 (q-value), x axis) of enrichment of the four functional gene sets (y axis) in the REACTOME database, most enriched in genes up-regulated in mature tuft cells (positive values) or ionocytes (negative values). **B** IL-13 treatment shifts the Large ALI cell composition. UMAP embedding of scRNA-seq profiles (dots) from control (left; same plot as in Fig. 1E, reproduced here for convenience) and IL-13-treated (right) LAE ALIs, colored by cell subset annotation. **C–E** Mature tuft cells are induced in IL-13-treated LAE ALIs. Zoom of a portion of the UMAP embedding in IL-13 treated LAE ALIs (from **B**, right) colored by scores for rare cell marker gene signatures (**Supplementary Data 2**) (**C**) or by expression of rare cell marker genes

(**D**). **E** Number of antibody-stained cells (y axis) for GNAT3 expressing tuft cells ($n = 2$ ALIs (Hu19, Hu67)) and BSND expressing ionocytes $n = 3$ ALIs (Hu19, Hu62, Hu67) in LAE ALI treated with PBS or IL13 (10 ng/ml) (x axis). (All experimental treatments were done in parallel; **Methods**). **F** Mature SAE ALIs are treated for 96 h with PBS (control) or IL13 (20 ng/ml). IL-13 treatment shifts SAE ALI cell composition. UMAP embedding of scRNA-seq profiles (dots) from control (left; as in Fig. 1F) and IL-13-treated (right) SAE ALIs, colored by cell type annotation. **G** Mature tuft cells are induced in IL-13-treated SAE ALIs. Zoom of a portion of the UMAP embedding in IL-13 treated SAE ALIs (from **G**, right) colored by expression of rare cell marker genes.

pre-ionocytes, mature ionocytes, and mature tuft cells do not cycle (Fig. 4E, G, H, I, K, L). As is the case in vivo, the fraction of cycling TIP cells (49%) is once again higher than that of cycling canonical basal stem cells (22%) (Supplementary Data 8). TIP cells express both ionocyte (*FOXI1*) and tuft cell (*ASCL2, POU2F3*) TFs, in both control and IL-13 conditions (Fig. 4E, H, I, L). During differentiation towards a mature tuft cell or ionocyte, the expression of TFs of the opposing lineage are suppressed (Fig. 4E, H, I, L). Of particular note, *SOX9* expression is induced in the tuft cell lineage and absent in the default ionocyte

lineage, while conversely *ASCL3* expression presages ionocyte commitment and is absent in the mature tuft cells (Fig. 4E, H, I, L). Ultimately, functional tuft and ionocyte genes are expressed when the lineage path is completed (Fig. 4E, H, I, L). Interestingly, prolonged IL-13 treatment of differentiating LAE ALI cultures resulted in a decreased in the number of ionocytes per ALI membrane (Fig. 4M), presumably because ionocytes could not be fully replenished by TIP cells, which were being diverted towards a mature tuft cell fate (Fig. 4N). Thus, RNA velocity analysis supports a model in which TIP cell descendants
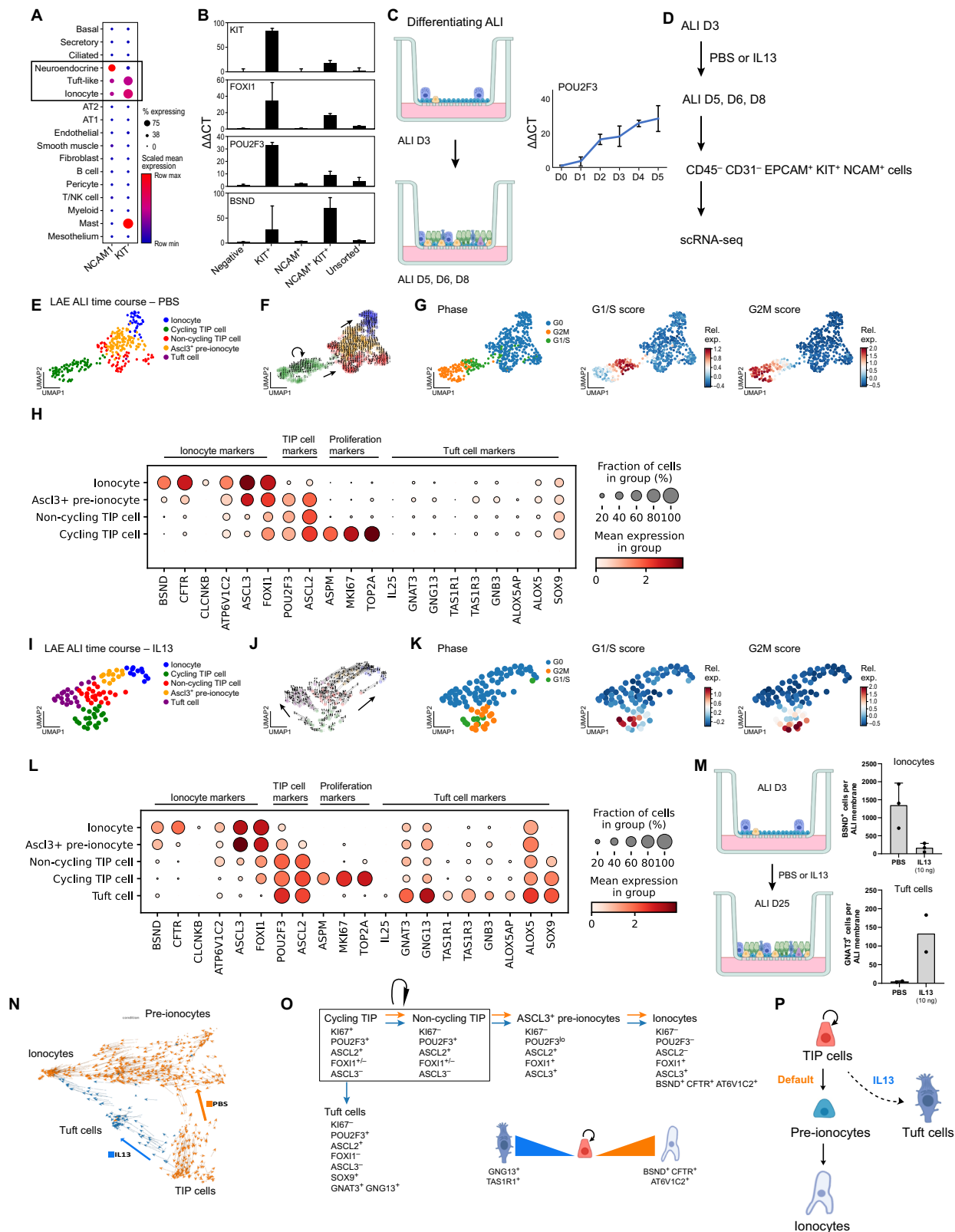
proceed through an orderly process of differentiation associated with the sequential expression of distinct cell fate-determining transcription factors based on the presence or absence of IL13 (Fig. 4N, O). In aggregate, these findings establish a TIP cell lineage hierarchy where the default lineage path is biased towards ionocyte fate, while Type 2 cytokines redirect TIP cells towards a mature tuft cell fate (Fig. 4O, P).

IL17A induces TIP cell differentiation towards a mature tuft cell lineage

Tuft cell differentiation is generally associated with Type 2 inflammation, but tuft cells are also found in human SARS-Cov2 pneumonia[47] and in mouse models of influenza infection[48]. Interestingly, the differentiation of murine influenza-induced tuft cells has been shown to

Fig. 4 | Type 2 pathway cytokines redirect the lineage of bipotent Tuft-Ionocyte Progenitor (TIP) cells. A, B Sorting strategy to enrich rare cells. **A**. Mean expression (dot color, relative expression) and proportion of cells (dot size) expressing genes encoding the cell surface proteins NCAM1 and KIT in human in vivo scRNA-seq data from Fig. 1B. **B** Expression level ($\partial\partial$CT, qPCR, y axis) of key rare cell marker genes (marked on top left) in sorted cell populations from human ALI cultures (x axis, labeled by sorting marker). $n = 3$ technical replicates. Error bars are standard deviation. Based on this expression data, dissociated cells were stained for anti-human CD45–BV421 (1:100; BioLegend 368522), anti-human CD31–BV421 (1:100; BioLegend 303124), anti-human CD326(EPCAM)–APC (1:100 BioLegend 324208), CD117(KIT)–FITC (1:100; BioLegend 313231) and anti-human CD56(NCAM)–BV711 (1:100; BD Biosciences 563169). A negative sort was performed for CD45 (immune cells) and CD31 (endothelial cells), with positive selection for CD326 (epithelial cells), CD56 (NCAM – rare cell marker), and CD117 (KIT – rare cell marker). Please see Supplementary Fig. 9 for gating strategy. **C, D** Experimental strategy. **C** Left: Model of differentiating ALI. Right: POU2F3 mRNA expression ($\partial\partial$CT, qPCR, y axis) at different time points (x axis) during ALI differentiation. $n = 3$ technical replicates. Error bars are standard deviation. **D** Schematic of experimental time course, where starting at ALI D3 (top; the time point at which TIP cells are first present) cultures were treated with IL13 (10 ng/ml) or PBS control for 5 days and then CD45- CD31- EPCAM + NCAM1 + , CD45- CD31- EPCAM + KIT+ and CD45- CD31- EPCAM + KIT + NCAM1+ rare cells were collected, pooled, and profiled using scRNA-seq. **E–H** TIP cells give rise to ionocytes via defined transition states in control LAE ALI cultures. UMAP embedding of scRNA-seq profiles (dots) from PBS-treated (control) LAE ALI cultures colored by cell type annotation (**E**), overlaid RNA velocity vectors (**F**), cell cycle phase classification (**G**, left), G1/S (**G**, middle), and G2/M (**G**, right) gene signature scores. **H** Mean expression (dot color) and fraction of cells (dot size) expressing different lineage markers (columns) in each cell subset along the default lineage transition of TIP cells towards ionocytes (rows). **I–L** TIP cells are diverted towards mature tuft cell fate following IL-13 treatment. UMAP embedding of scRNA-seq profiles (dots) from IL-13-treated LAE ALI cultures colored by cell subset annotation (**I**), overlaid RNA velocity vectors (**J**), cell cycle phase classification (**K**, left), G1/S (**K**, middle), and G2/M (**K**, right) gene signature scores. **L** Mean expression (dot color) and fraction of cells(dot size) expressing different lineage markers columns, same genes as in H) in each cell subset (rows) along the IL13-induced lineage transition of TIP cells towards mature tuft cells. **M** Left: Experimental setup schematic of differentiating ALI treated continuously with IL-13 for 25 days. Right: Number of BSND+ cells (ionocytes, y axis, $n = 3$ ALIs (Hu19, Hu60, Hu67)), error bars are standard deviation, and GNAT3+ cells (tuft cells, y axis, $n = 2$ ALIs (Hu60, Hu67)) in PBS and IL-13 conditions (x axis). **N** RNA velocity analysis of the pooled PBS and IL13-treated ALIs (clustering shown in Supplementary Fig. 10C, D) demonstrates that IL13 redirects TIP cell differentiation towards mature tuft cells and away from the default pathway of ionocyte differentiation. **O** Schematic depicting the observed expression of lineage-specifying TFs in TIP cell descendants that are differentiating towards either ionocyte or mature tuft cell fate. Orange arrows indicate default differentiation (PBS) and blue arrows indicate IL13-induced differentiation. **P** Proposed model of cytokine-mediated TIP cell lineage switching.

be independent of Type 2 cytokine signaling[49]. We therefore hypothesized that Type 17 cytokine signaling also biases TIP cell lineage towards mature tuft cell differentiation. Indeed, IL17A receptors (*IL17RA* and *IL17RC*) are expressed in TIP cells (Supplementary Fig. 8C). To test our hypothesis, we treated fully differentiated (day 39) ALI cultures with IL17A (50 ng/μl) for five days (Fig. 5A). IL17A administration resulted in the maturation of bona fide tuft cells as evidenced by an increase in the numbers of mature tuft cells that express GNAT3 per ALI membrane (Fig. 5B). Similarly, continuous administration of IL17A for 21 days during the course of ALI epithelial differentiation from day 3 to 24 also resulted in an increase in the number of mature GNAT3 expressing tuft cells per membrane (Fig. 5C, D). To our knowledge, this represents the first direct evidence for a role of IL-17 signaling in airway tuft cell differentiation.

Human asthmatic epithelium contains mature GNAT3⁺ tuft cells

Since cytokine signaling resulted in differentiation of mature GNAT3 expressing tuft cells in epithelial ALI cultures, we hypothesized that airway epithelial cultures derived from asthmatic individuals would similarly contain mature tuft cells, although such cells have not heretofore been found in the airways of asthmatic individuals[3,25]. To test our hypothesis, we isolated basal cells from the large airways of two individuals who died during an asthma exacerbation and generated fully differentiated for 28 days LAE ALI cultures without added cytokines. We found a 2-4 fold increase in the number of GNAT3⁺ mature tuft cells per airway membrane compared to two ALIs from two healthy lungs (Fig. 5E). We next performed whole-mount staining of microdissected human airway from the one patient from whom intact tissue was available and identified GNAT3 expressing mature tuft cells (Fig. 5F). Because mature tuft cells have not been previously reported in the intact asthmatic airway, we performed immunohistochemistry for GNAT3 and ALOX5AP, demonstrating the expression of functional proteins in both taste transduction and the synthesis of asthma-associated leukotrienes (Fig. 5F). Of note, ALOX5AP staining reveals the characteristic bipolar morphology associated with mature tuft cells. Conversely, we could not detect GNAT3⁺ mature tuft cells in human airway whole-mounts from individuals without asthma. To our knowledge, this represents the first direct evidence for the existence of mature tuft cells in situ in asthmatic airways.

Discussion

We describe the identity and physiological regulation of the TIP cell, a new replicative rare cell progenitor that has the potential to differentiate into either ionocytes or mature tuft cells. Despite their prior annotation as "tuft-like" cells based on their expression of *POU2F3*, the default pathway of TIP cell differentiation leads to the steady-state production of ionocytes. In contrast, cytokines induce a deviation from the default TIP cell lineage pathway, resulting in the differentiation of mature tuft cells. Remarkably, ~19% of TIP cells are replicating, exceeding the cycling fraction of basal stem cells. This high rate of turnover suggests that ionocytes, the default products of TIP cell differentiation, may have a short half-life and therefore require constant replenishment. Indeed, the application of IL13 to patient samples decreased ionocyte numbers within days and this led to abnormal airway surface physiology[19]. Thus, we speculate that episodes of acute inflammation may result in rapid changes in disease pathophysiology on the order of days to weeks.

We named the cells "TIP cells" following the simple naming convention used to label hematopoietic progenitors by their spectrum of differentiation. Mechanistically, TIP cell bipotency is associated with the dual expression of TFs of both the tuft and ionocyte lineages. Prior studies were based on the analysis of scant cells and had annotated these rare cell progenitors as "tuft-like"[1], "tuft"[20], or "undefined rare"[5], while the actual tuft cells were identified as "other" cells. Our revised annotation and inference of the role of these cells illustrates how analysis of a very large compendium of scRNA-seq data from multiple studies can resolve confusing lineage relationships. Notably, genetic deletion of *POU2F3* leads to a loss of ionocytes[1], consistent with the TIP cell mechanism we describe. Although the existence of multipotent rare cell progenitors has been suggested, we only find evidence for a bipotent progenitor of ionocytes and tuft cells. Furthermore, basal stem cells are known to directly differentiate into the third rare cell lineage, neuroendocrine cells[16]. Provocatively, two recent studies suggested that basal cells, and possibly TIP cells, foreshadow subclasses of small cell lung cancer[50,51].

Prior studies argued that ionocytes were not relevant to CF-associated small airways disease because they could not be identified in the small airways. However, our demonstration of small airway ionocytes, accompanied by their small airway TIP cell parents, leaves open the door for a substantive role for ionocytes in CF-related small
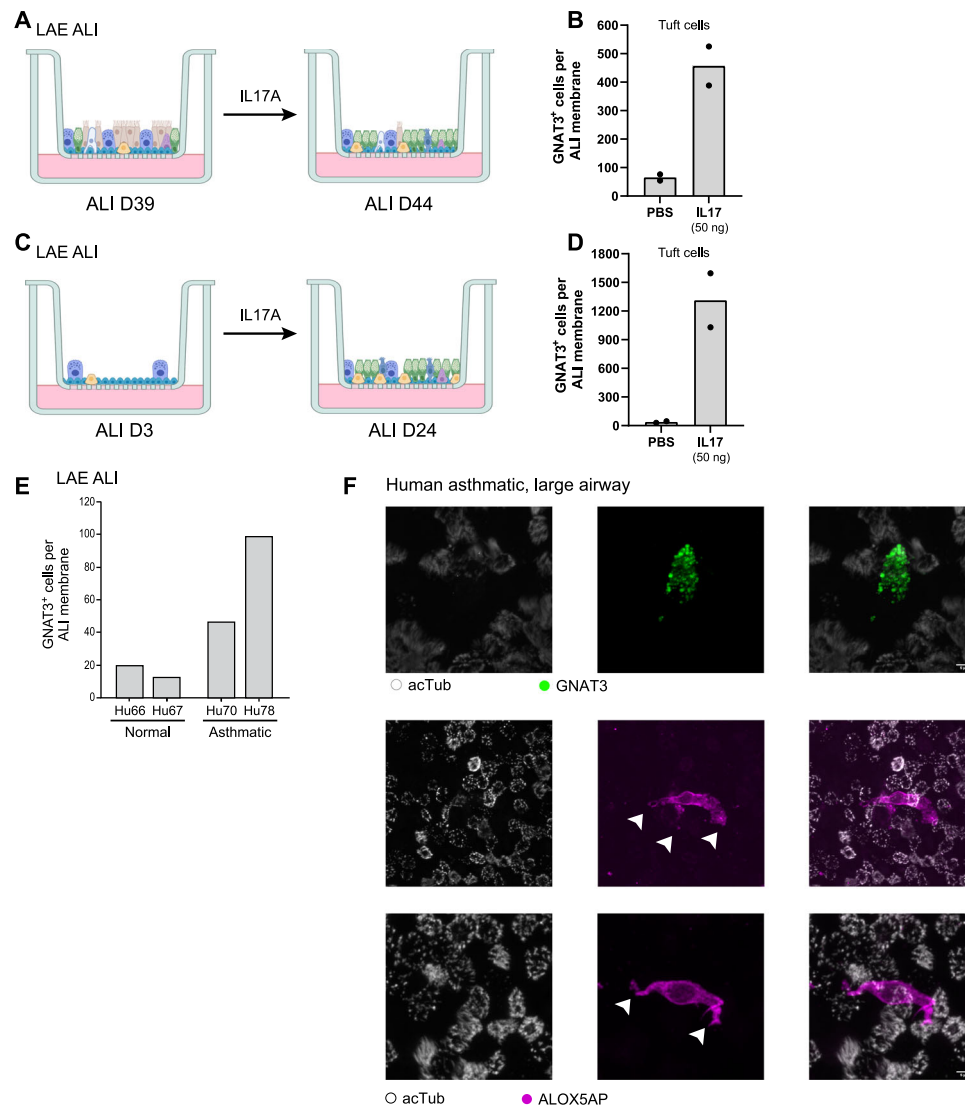
Fig. 5 | IL17A promotes tuft cell differentiation and asthmatic airway epithelium contains mature tuft cells. A–D Tuft cell abundance increases following IL17A treatment in LAE ALI. **A, C** Overview schematic of IL17A treatment experiments in mature LAE ALI treated with PBS (control) or IL17A (50 ng/ml; from D39 to D44) (**A**) and differentiating ALI at D3 (the first time point at which TIP cells are present) treated cells with IL17A (50 ng/ml) or PBS from D3 to D24 (**C**). **B** Number of GNAT3+ mature tuft cells (quantified by immunohistochemistry, y axis, $n = 2$ ALIs (Hu19, Hu67)), when mature LAE ALIs are treated with PBS or IL17A (50 ng/ul; from D39 to D44) (x axis) **D** Number of GNAT3+ mature tuft cells (quantified by immunohistochemistry, y axis, $n = 2$ ALIs, (Hu19, Hu67)), in LAE ALIs when treated with PBS or IL17A from D3 to D24. **E, F** Tuft cells are increased in ALIs and airways from asthmatic patients. **E** Number of GNAT3+ mature tuft cells (y axis) in LAE ALIs derived from two asthmatic individuals (Hu70 and Hu78) and in patients with no history of lung diseases (Hu66, Hu67) (x axis), $n = 1$ ALI per donor. **F** Whole mount staining of dissected airways from a patient with asthma exacerbation. Staining in the top panel with GNAT3 (green) shows a mature tuft cell, surrounded by ciliated cells (Atub, white). Staining with additional mature tuft cell marker ALOX5AP (magenta) reveals the characteristic bipolar morphology (arrows) associated with mature tuft cells (bottom panels). Due to the availability of tissue, immunostaining was performed on only 1 donor. Illustrations 4B and D were created with BioRender: https://BioRender.com/g37wh81.

airways disease. Furthermore, the TIP cell model suggests that inflammation may worsen CF-related airway pathology due to a relative depletion of ionocytes.

The TIP cell model provides an explanatory framework for considering the mechanisms that drive alterations in the frequency of ionocytes and mature tuft cells that have been reported in eosinophilic and non-eosinophilic asthma[18,19]. Indeed, the application of asthma-associated cytokines to airway epithelial cultures resulted in a loss of ionocytes associated with pathologic alterations of CFTR current and CFTR-mediated fluid absorption[18,19]. We now identify mature tuft cells in the asthmatic airway that express ALOX5AP, necessary for asthma-associated leukotriene production. However, one wonders whether such rare cells are truly pathogenic or whether these rare cells are merely a byproduct of TIP cell diversion and

serve as markers of disease. Alternatively, cytokines as well as growth factors may result in pathologic proportions of rare cell numbers[13,52]. In aggregate, the pathologic regulation of TIP cells may explain the reported alterations in the frequency of ionocytes and mature tuft cells in asthma, cystic fibrosis, and viral pneumonia. That said, true normative values for rare cell numbers have not yet been established. Only then can we assess whether and how alterations in rare cell numbers are physiologically compensatory, incidental, or whether they are pathologic.

Methods

The research complies with all relevant ethical regulations and work is approved by the Massachusetts General Brigham Institutional Review Board.

Human lung tissue

The use of all human material has been approved by the Massachusetts General Brigham Institutional Review Board [Protocol # 2010P00 1354]. Human lung tissue was received from New England Donor Services. New England Donor Services obtained informed consent, and the donor's information was de-identified as per IRB protocol. Tissue was collected from individuals without any reported history of lung disease, smoking or drug use.

Human lung tissue dissection, dissociation and cell enrichment sorting

Large airways were manually dissected and washed with HBSS media (Sigma, 55021 C) containing Y-27632 (5 µM, Selleck Chemicals S1049) inhibitor in order to remove any mucus and debris. Single cells were dissociated in HBSS media containing pronase (1 mg/ml Roche 11 459 643 001), DNAse (1 unit/ml, Worthington LK003170) and Y-27632 inhibitor (5 µM, Selleck Chemicals S1049) on a 37 °C shaker for 30–60 min. The dissociated epithelial layer was collected by centrifugation. Cell pellets were treated with TrypLEExpress solution (ThermoFisher Scientific 12605036) for 5 min at 37 °C with rotation. Cell suspensions were collected by centrifugation and treated with ACK lysis buffer (ThermoFisher Scientific A1049201) for 1 min on ice. Cell pellets were resuspended in phosphate-buffered saline (PBS ThermoFisher Scientific 14190250) with 10% FBS (Thermo Fisher Scientific A3160502) and Y-27632 inhibitor (5 µM, Selleck Chemicals S1049) and then sequentially filtered through 100 µM and 40 µM filters. Cell viability was assessed using a hemocytometer, and single cell suspensions were placed on ice for further experiments. To further enrich large airway rare epithelial cells we used FACS sorting. We stained cells with anti-human CD45−BV605 (1:100; BioLegend 368524) or anti-human CD45−BV421 (1:100; BioLegend 368522), anti-human CD31−BV421 (1:100; BioLegend 303124), anti-human EPCAM−APC (1:100 BioLegend 324208) or EPCAM−PE (1:100; BioLegend 324205), anti-human CD117−FITC (1:100; BioLegend 313231) or CD117−APC-Cy7 (1:100; BioLegend 313227) for 20 min at room temperature. Cells were washed and suspended in PBS (ThermoFisher Scientific 14190250) containing 10% FBS (Thermo Fisher Scientific A3160502) and Y-27632 inhibitor (5 µM, Selleck Chemicals S1049) with added DAPI solution (ThermoFisher Scientific 62248) to assess cell viability. After washing, cells were sorted on a FACS sorter, and analysis was performed using FlowJo software.

Distal lung lobe regions were manually dissected, avoiding all airways larger than 2 mm, and washed with HBSS (Sigma, 55021 C) containing Y-27632 inhibitor (5 µM, Selleck Chemicals S1049) to remove any mucus and debris. Single cells were dissociated in HBSS media (Sigma, 55021 C) containing collagenase (225 units/ml Worthington LS004188), dispase (2.5 units/ml Worthington Biochemical LS02100), elastase (2 units/ml Worthington Biochemical LS002279), pronase (1 mg/ml Roche11 459 643 001), DNAse (1 unit/ml, Worthington LK003170) and Y-27632 inhibitor (5 µM, Selleck Chemicals S1049). Tissue fragments were injected with the above dissociation solution multiple times with a syringe and then divided into smaller pieces that were placed in a 37 °C shaker for 60 min. Cell suspensions were treated with ACK lysis buffer (ThermoFisher Scientific A1049201) for 2 min on ice to remove red blood cells. Cell pellets were resuspended in PBS with 10% FBS and Y-27632 inhibitor (5 µM) and sequentially filtered through 100 µM and 40 µM filters. Cell viability was assessed using a hemocytometer and single cell suspensions were placed on ice for further experiments. Additionally, in order to enrich for rare cell types in distal lung lobe regions, we used sequential magnetic cell separation using MACS cell separation columns on three human lung samples (Hu37, Hu49, and Hu62). Specifically, for Hu37 and Hu49 distal lung lobe samples, we used human CD45 MicroBeads (Miltenyi Biotec 130-045-801) to enrich immune cells in the CD45 positive fraction and further loaded the CD45 negative fraction onto CD31 MicroBeads

(Miltenyi Biotec 130-091-935). We collected the CD31-positive fraction to enrich endothelial cells, and also collected and loaded the CD31-negative fraction. For Hu62, distal lung lobe samples, we used human CD45 MicroBeads (Miltenyi Biotec 130-045-801) to enrich immune cells in the CD45 positive fraction and further loaded the CD45 negative fraction onto CD31 MicroBeads (Miltenyi Biotech 130-091-935). We collected the CD31-positive fraction to enrich endothelial cells and the CD31 negative fraction was then loaded onto human anti-IgM MicroBeads conjugated to anti HTII-280 antibody (Terrace Biotech TB-27AHT2-280) to enrich AT2 cells. We collected the HTII-280 negative fraction to increase the number of captured small airway epithelial cells.

Dissection of large and small airways and whole mount staining

Large and small airway samples were handled identically. Within 24 h after explanting the lungs, airways were dissected at various indicated areas in the right and left mainstem bronchi and the right and left distal airways (<2 mm in diameter; as in Fig. S7). Tissue explants were then fixed in 4% PFA (Santa Cruz Biotechnology sc-281692) for 1 h and washed in PBS for 1 h. Tissues were preserved in PBS until staining. Explants were then permeabilized in PBS-0.3% TritonX-100 (PBST, Sigma Aldrich T8787-100ML for 30 min. Samples were stained with primary antibody at 4 °C overnight, diluted in 1% BSA-0.3% PBST (Sigma Aldrich A7906). The following antibodies and dilutions were used as per Supplementary Data 9.

For each donor, 3 large airway and 3 small airway samples were obtained. From these, a number of airways (each 0.06480mm^2 surface area region) were obtained for each donor based on the size of tissue dissected, as follows: HU66: 8 large airway regions and 9 small airway regions; HU67: 10 (large) and 17 (small); HU68: 8 (large) and 7 (small). Each explant was mounted onto an O-ring (McMaster Carr 1283N264) fixed into a 60 mm dish (Cellstar 628160) with Sylgard 184 silicone elastomer (EMS 24236-10). Each region was imaged with a 25X objective, and cells were counted on a maximal intensity projection. Only regions in which there was concomitant staining for acetylated tubulin were included to ensure imaging of the airway epithelial layer. For BSND staining, 1.68 mm^2 of large airway and 2.14 mm^2 of small airway were imaged in total. One way ANOVA (Sidak's multiple comparisons) was done. Each dot represents individual fields imaged on respective donors. Error bars are the standard deviation.

Air liquid interface cultures

Cell were dissociated from large airways or distal lung lobe regions and expanded in basal cell media containing PneumaCult-Ex Plus medium (Stem Cell Tech 5040) containing 0.5 µM DMH-1 (Tocris 4126), 1 µM CHIR 99021 (Tocris 4423), 1µM A-8301 (Tocris 2939), 5 µM Y-27632 (Selleckbio S1049), and Primocin 500x (Invivogen ant-pm-2). To initiate air−liquid interface (ALI) cultures, basal cells were seeded onto transwell membranes with basal cell media in the upper and lower chambers. After reaching confluence, the media was removed and replaced with PneumaCult-ALI Medium (Stem Cell Tech 05001) in both the upper and lower chambers. After 24 h, media was removed from the upper chamber to begin basal cell differentiation for the indicated number of days. ALI cultures were treated with either 10 or 20 ng/ml of recombinant human IL13 (Peprotech 200-13-10UG), 10 ng/ml of recombinant human IL4 (Peprotech AF-200-04 100ug), 10 ng/ml of recombinant human IL5 (Peprotech 200−05 100ug), or 50 ng/ml of recombinant human IL17A (Peprotech 200−17 50ug) as indicated (all diluted in PneumaCult-ALI Medium). We initially used 20 ng/ml of IL13 in ALI cultures as previously reported[7,53], but subsequently found that a lower concentration of 10 ng/ml IL13 was sufficient to induce mature tuft cell and goblet cell differentiation, so this concentration was employed in the data presented in Fig. 4. Control cultures received an equal volume of PBS or DMSO (Sigma Aldrich D2650-100ML).

Staining and quantification of cells in Air liquid interface cultures

The luminal side of the ALI membranes were washed with 200ul of PBS or 10 mM DTT (Thermo Fisher Scientific R0861) in PBS for 1 min followed by an additional wash with PBS. ALI membranes were fixed in 4% PFA for 15 min and washed in PBS for 5 min. ALI membranes were preserved in PBS until staining. ALI membranes were permeabilized in PBS-0.3% TritonX-100 (PBST) for 60 min. Samples were stained with primary antibody at 4 °C overnight, diluted in 3% BSA 0.5% Tween20 5% Donkey Serum in PBS. Samples were stained with secondary antibodies at room temperature for 1 h, diluted in 3%BSA 0.5%Tween20 5% Donkey Serum in PBS. The primary and secondary antibodies used and their concentrations are listed in Supplementary Data 9.

The ALI membranes were then mounted using a mounting medium containing DAPI (Southern Biotech DAPI Fluoromount G 0100-20). Confocal images were obtained using an Olympus FV10i confocal laser-scanning microscope with a 10× objective, and the images were processed using Image J and cells from the ALI membrane were manually counted. Statistical analysis for Fig. 1G was a 2 tailed unpaired T test. Each dot represents the counts on ALI membrane for individual donors. Statistical analysis for Supplementary Fig. 6A one way ANOVA. $P = 0.0001$.

Generation of GCaMP6s cell line and live visualization of calcium signals

Human airway basal cell lines were transfected with lentivirus carrying CMV-GCaMP6s and a puromycin selection marker (Vectorbuilder VB900088-2805ckn). Positively transfected cells were selected with 1 µg/ml puromycin (Invivogen ant-pr-1) for 3 consecutive days. Basal cells were then grown in air-liquid interface cultures as described above. After maturation, ALI membranes were cut from the transwell and mounted on a glass bottom imaging chamber (Warner Instruments 64-0228). GCaMP signals were visualized using an Olympus IX81 inverted epifluorescence microscope fitted with a blue excitation filter.

ALI membrane cell dissociation and FACS

ALI membranes were cut and washed twice with PBS. Membranes were incubated in TrypLE Express solution (ThermoFisher Scientific) and incubated at 37 °C for 5 min with rotation. Dissociation was stopped with PBS containing 10% FBS and Y-27632 inhibitor (5 µM). Dissociated cells were passed through a 40 µM filter. Cells were then pelleted and stained with anti-human CD45−BV605 (1:100; BioLegend 368524), anti-human CD31−BV421 (1:100; BioLegend 303124), anti-human EPCAM−APC (1:100 BioLegend 324208) or EPCAM−PE (1:100; BioLegend 324205), anti-human CD117−FITC (1:100; BioLegend 313231) or CD117−APC-Cy7 (1:100; BioLegend 313227) and anti-human CD56−BV711 (1:100; BD Biosciences 563169) for 20 min at room temperature. Cells were washed and suspended in PBS (ThermoFisher Scientific 14190250) containing 10% FBS (Thermo Fisher Scientific A3160502) and Y-27632 inhibitor (5 µM, Selleck Chemicals S1049) with added DAPI solution (ThermoFisher Scientific 62248) to assess cell viability. After washing, cells were sorted on a FACS sorter and analysis was performed using FlowJo software.

q-RT PCR

Total RNA was extracted from ALI using Direct-zol RNA minipre (Zymo Research, R2051). cDNA was generated using Goscript reverse transciption mix with random primers (Promega A2801). A 10uL q RT-PCR reaction was performed using using primer pairs in Supplementary Data 10, cDNA, Luna universal qPCR master mix (NEB Biolabs M3003E) and dH$_2$O. A BioRad C1000 touch Thermal Cycler (CFX 384) was used. Reaction protocol was as follows: A) 95 °C for 1 min, B) 95 °C for 15 seconds, C) 60 °C for 30 seconds, B and C were repeated 44 times, D) 65 °C for 5 seconds, E) 95 °C, and held at 4 °C. Fold changes using

ΔΔCT methodology was calculated with GAPDH as housekeeping gene (see Source Data file for Fig. 4B).

scRNA-seq library generation and sequencing

Droplet-based scRNA-Seq 10x Genomics V2 (donor samples from Hu28, Hu30, Hu32, Hu37, Hu39, and Hu49) or V3 (donor samples from Hu52 and Hu62) 3' gene expression technology was used according to manufacturer's recommendations. In short, single cells were washed with 0.4% BSA in PBS, and loaded onto a Chromium single-cell 3' Chip. Single cells were partitioned into droplets with gel beads in the Chromium Controller. After emulsions were formed, barcoded reverse transcription of RNA took place. This was followed by cDNA amplification, fragmentation and adapter and sample index attachment. Libraries were pooled together and sequenced on an Illumina NextSeq 500, or an Illumina Nova-Seq, with paired end reads.

Pre-processing of human lung scRNA-seq data

Raw 3' scRNAseq sequencing BCL files were demultiplexed using the CellRanger pipeline (10X Genomics) version 7.0.1. *'mkfastq'* command using default parameters to generate *fastq* files followed by *'count'* command to align the reads to the human GRCh38 reference genome.

Quality control and data analysis of human lung scRNA-seq datasets

The resulting raw barcode, gene, and count matrices from individual samples were individually loaded in R(4.0.5) for downstream analysis using Seurat[54] version 4.0.6 with the *read 10X* function. We removed low quality cells with fewer than 1000 unique molecular identifiers (UMIs), fewer than 400 detected genes or greater than 25% mitochondrial genes from all downstream analysis. The remaining cells were 'LogNormalized' using NormalizeData() function and 2000 variable genes were identified using the default 'vst' method in the 'FindVariableFeatures' function. The normalized data was scaled using the ScaleData function, which was then used to perform principal component (PCA) analysis on the variable gene expression space. Next, dimensionality reduction of the first 20 principal components was performed using Uniform Manifold Approximation and Projection (UMAP) for visualization and annotation of each donor lung scRNA-seq data prior to performing donor-wise integration.

Cell clustering and annotation of human lung scRNA-seq

Cell type identification was performed by assigning cell clusters to cell subsets at different resolutions. The 'FindNeighbors' function was applied to each donor's normalized gene-expression data matrix to calculate the pairwise distances between the cells and construct KNN and shared nearest neighbor (SNN) graphs. Then de novo clustering was performed using the Louvain algorithm at different resolutions (0.2, 0.4, 0.8, 1.2, 2.4, and 5) on the SNN graph space. For each cluster, we used the findAllMarkers() function to identify the positive markers for that cluster with a minimum fraction of 0.25, log2 fold-change of 0.25, where we used the Wilcoxon rank-sum test for finding upregulated differentially expressed genes (DEGs) for each cluster and computed Bonferroni corrected p-values. We overlapped de novo discovered DEGs with well-established cell type markers to perform high-level cell type annotation in an iterative and semi-supervised way to assign clusters to broad classes based on consistent expression of known markers in a given cluster. We used the following markers for broad cell class annotation: epithelial (*EPCAM, KRT8, KRT18, TP63, S100A2*), mesenchymal (*COL1A1, COL1A2, DCN, FBLN1*), myeloid (*LYZ, C1QA, C1QB, APOC1*), T (*CD3D, CD3E, GZMK*) & NK (*GNLY, GZMB, KLRD1*), B/plasma (*CD79A, MS4A1, MZB1, JCHAIN, IGHG1*), mast (*TPSAB1, TPSB2, CPA3*), and endothelial (*VWF, PECAM1, CLDN5*). We then subset cells within each of the broad classes and further subset samples into two groups (proximal and distal) based on whether they were acquired

from the large airways (proximal) or parenchyma (distal). To annotate cell subsets within each broad cell class from either proximal or distal region samples, the whole process of log normalization, selection of "variable" genes, scaling, dimensionality reduction (PCA/UMAP), clustering at different resolutions, and finding DEGs was repeated, as described above. Doublet enriched clusters were identified as clusters expressing canonical markers for two cell types and containing higher numbers of genes per UMI (genes/UMI) than expected from each individual cell type. These doublets were removed from downstream analysis.

Data integration of human lung scRNA-seq

There was a noticeable batch effect in our human lung data set primarily driven by using either V2 or V3 chemistry 3' gene expression kits for different donors, which required robust data integration. We chose to integrate the data across donors using the anchor-based approach as implemented in Seurat v3[54]. Prior to integration, count matrices for individual donors were normalized and mitochondrial gene content per cell was regressed out using the SCTransform() function. Next, we selected 3000 integration features via the SelectIntegrationFeatures() function followed by donor-wise integration using PrepSCTIntegration(), FindIntegrationAnchors(), and IntegrateData() function respectively. Post-integration, the resulting integrated Seurat objects were subjected to downstream analysis, as described above, including data log normalization, selection of variable features, data scaling, dimensionality reduction (PCA/UMAP), clustering, finding DEGs, and cell annotation.

scRNA-seq cell cycle analysis

To identify cycling cells, we used the Seurat CellCycleScoring() function to compute, for each cell, the enrichment score for the expression of genes linked to either G1/S or G2/M phase of the cell cycle[55]. Cells with either a G1/S- or G2/M-score greater than 0.1 were classified as cycling and all other cells were considered non-cycling (Fig. 2D). Finally, the resulting cycling status or cell cycle phase of cells were visualized using Seurat's DimPlot() function for different UMAP embedding. To quantify cycling cell stages in the ALI scRNA-seq data, raw FASTQ files were aligned to the human reference build GRCh38 using CellRanger 7.0.1 'count' function, input into Scanpy, and scored with the integrated list of cell cycle genes used in the approach implemented in Seurat[55] via the score_genes_cell_cycle() function with the default n_bins=25.

scRNA-seq module score analysis and detection of mature tuft cells

Sub-clustering of epithelial cells in our scRNA-seq data allowed us to identify clusters containing ionocytes, neuroendocrine (NE) cells, and TIP cells, but not mature tuft cells. To detect mature tuft cells in our data, we utilized published scRNA-seq data of rare lung epithelial cells[5] and identified the top ten differentially expressed genes for ionocytes, NE, TIP, and mature tuft cells (Supplementary Data 2). Signature scores for rare epithelial cells were assigned per cell using the AddModuleScore() function in Seurat v4, which compares the average expression of genes in an input gene set to the aggregate expression of control sets of genes randomly selected from the same expression level bins as the input gene set[55].

HLCA rare epithelial cell analysis

To characterize the number and proportion of rare lung epithelial cells in other published data sets, we subset the rare epithelial cells in the main HLCA v1 data set using Seurat v4. We log-normalized the data using the NormalizeData() function and identified 2000 variable features using the FindVariableFeatures() function. Next, we performed data scaling, PCA, batch correction using Harmony (v1.0), and UMAP dimensionality reduction to visualize the data in two dimensions

(Supplementary Fig. 4A–C). Finally, we scored all cells for rare lung epithelial cell signatures (Supplementary Data 2) using the AddModuleScore() function and assigned clusters that were specifically enriched for one of the 4 rare cell type signatures of the corresponding rare cell identity (Supplementary Fig. 4D).

To de novo identify rare lung epithelial cells in HLCA, all "Normal" condition data sets from HLCA v1.0 extended (full) were subsetted. Next, all cells were scored for the original rare lung epithelial cell signatures (Supplementary Data 2) using the Scanpy v1.9.6 score_genes() function with parameter settings n_bins=24, ctrl_size=100, and use_raw=False, recovering 41 cells with tuft cell signature score > 1 after doublet removal, which were labeled as mature tuft cells. The 41 cells were used to identify more generalizable markers of lung mature tuft cells, using the Scanpy function rank_genes_group(). Refined (new) mature tuft marker genes were selected manually based on having a logFC > 7.5 and highly specific expression in mature tuft cells across the HLCA cohort (Supplementary Fig. 4F) by inspection of dotplots generated using the Scanpy dotplot() function. Next, the original ionocyte, neuroendocrine, tuft-like, and refined mature tuft cell signatures were used to score all HLCA cells (Supplementary Data 2) with the Scanpy score_genes() function with parameter settings n_bins=24, ctrl_size=100, and use_raw=False (Supplementary Fig. 4G), using a threshold of 1 for ionocytes and neuroendocrine cells, threshold of 1.6 for POU2F3+ tuft-like cells, and a threshold of 0.5 for mature tuft cells. For batch effect correction in Supplementary Fig. 4E, scANVI model with n_layers = 3 and n_latent = 32 from scvi-tools v0.20.3 was used on raw counts to integrate the data across samples with default parameters when training.

Differential expression and gene set enrichment analysis

Differential expression analysis between 56 mature tuft cells and 1,419 ionocytes that were de novo scored in the HLCA normal extended data sets, was performed using Scanpy rank_genes_groups() function, retaining 94 and 257 genes with a log fold change > 0.7 and q-value < 0.05 as differentially up regulated in tuft cells and ionocytes, respectively (Supplementary Data 7). Gene Set Enrichment Analysis (GSEA) on these two gene sets was performed against the REACTOME database at https://www.gsea-msigdb.org/gsea/ using default parameters, where the top 4 most enriched functional categories were displayed in Fig. 3A and the rest included in Supplementary Data 7.

RNA velocity analysis of scRNA-Seq

Aligned scRNA-seq reads were further processed using the Velocyto 0.17[37] read counting pipeline in Python 3.6.0. The GRCh38 expressed repeat mask was used. The output loom file was processed using Scanpy 1.6.0 with leidenalg 0.8.3. For in vivo human lung scRNA-seq data, quality control and integrated UMAP plots were obtained as described above. For ALI scRNA-seq data, high-quality cells were retained, those that had >200 genes detected, <15% mitochondrial read count, and > . 45 doublet prediction score from scrublet (2020 version) (automatic detection threshold, manually verified). scVelo 0.2.2 was used to predict RNA velocity with the following parameters: n_neighbors=20 (for full data set) and n_neighbors=5 (for subset computation). Stochastic, deterministic, and dynamical modeling were all tested, and quality control plots of known key marker genes were plotted to select stochastic estimation as the final modeling method.

Single-cell ATAC-seq (scATAC-seq) library generation and sequencing

scATAC-seq was performed from primary carina and subpleural parenchymal tissue of one donor (Hu62). Libraries were generated using the 10x Chromium Controller and the Chromium Single Cell ATAC Library & Gel Bead Kit (Chromium Single Cell ATAC Library & Gel Bead Kit 1000111) according to the manufacturer's instructions

(CG000169-Rev C; CG000168-Rev B) with the following modifications in cell handling and processing. Briefly, human lung primary cells were processed in 1.5 ml DNA LoBind tubes (Eppendorf), washed in PBS via centrifugation at $400 \times g$ for 5 min at 4 °C and lysed for 3 min on ice before washing via centrifugation at $500 \times g$ for 5 min at 4 °C. Supernatant was discarded and lysed cells were diluted in 1× diluted nucleus buffer (10x Genomics, 2000153) before counting using trypan blue and a Countess II FL Automated Cell Counter to validate lysis. If large cell clumps were observed, a 40-μm Flowmi cell strainer was used before the tagmentation reaction, followed by Gel Bead-In-Emulsion generation and linear PCR as described in the manufacturer's protocol. After breaking the emulsion, barcoded tagmented DNA was purified and further amplified to enable sample indexing and enrichment of scATAC-seq libraries. Final libraries were quantified using a Qubit dsDNA HS Assay kit (ThermoFisher Scientific Q32851) and a High Sensitivity DNA chip run on a Bioanalyzer 2100 system (Agilent). All libraries were sequenced using NextSeq High Output Cartridge kits and a NextSeq 500 sequencer (Illumina), and 10x scATAC-seq libraries were characterized by paired-end sequencing (2 × 72 cycles).

scATAC-seq data analysis

Fastq files were aligned to human genome reference hg19 using cellranger-atac (v1.2). Fragment files were parsed with ArchR (v1.0.2) and initial QC on cells was applied based on sequencing quality and depth, discarding cells with <1,000 fragments or a transcription start site (TSS) enrichment score <4. Dimensionality reduction and clustering was applied on cells passing filters using the standard ArchR analysis workflow. Cell type annotation of scATAC-seq data was guided using cell annotation labels from scRNA-seq data, which were transferred to the scATAC-seq data using the function addGeneIntegrationMatrix. De novo marker discovery was run with the function getMarkerFeatures using the GeneScoreMatrix assay. Final scATAC-seq cell annotation was performed by assigning the de novo discovered clusters to cell types predicted by the RNA or to cell types associated with the most enriched scATAC-seq features in that cluster when no analog cell type was found in the RNA data. TF activity was inferred using ChromVAR (v1.16) implemented by ArchR via addMotifAnnotations, addBgdPeaks and addDeviationsMatrix functions. For visualization, ArchR built-in plotting functions and ComplexHeatmap were used (v2.10).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data are included in the Supplementary Information or available from the authors, as are unique reagents used in this Article. The raw numbers for charts and graphs are available in the Source Data file whenever possible. The scRNAseq and scATACseq data generated from donor lungs in this study have been deposited in the NIH SRA database: PRJNA1044083. The analyzed cell counts and differential gene analysis by cell type of the scRNAseq and scATACseq data generated from donor lungs data are available in the Supplementary Data 1-10 of this work. The ALI scRNAseq data generated in this study have been deposited GEO with accession numbers: GSE240168. Source data are provided with this paper.

Code availability

Analyses in this manuscript largely used publicly available code/functions from Scanpy and Seurat R package as described in the Methods, except for the code used for rare cell detection in HLCA, which is available at the following link: [https://github.com/TsankovLab/LungRareCells] and [https://doi.org/10.5281/zenodo.15298362][56].

References

1. Goldfarbmuren, K. C. et al. Dissecting the cellular specificity of smoking effects and reconstructing lineages in the human airway epithelium. *Nat. Commun.* **11**, 2485 (2020).
2. Adams, T. S. et al. Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci. Adv.* **6** eaba1983 (2020).
3. Braga, F. A. V. et al. A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat. Med.* 10.1038/s41591-019-0468-5 (2019).
4. Travaglini, K. J. et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619–625 (2020).
5. Deprez, M. et al. A Single-Cell Atlas of the Human Healthy Airways. *Am. J. Respir. Crit. Care Med.* **202**, 1636–1645 (2020).
6. Plasschaert, L. W. et al. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* **560**, 377–381 (2018).
7. Montoro, D. T. et al. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* https://doi.org/10.1038/s41586-018-0393-7 (2018).
8. Madissoon, E. et al. scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biol.* **21**, 1 (2019).
9. Carraro, G. et al. Single-Cell Reconstruction of Human Basal Cell Diversity in Normal and Idiopathic Pulmonary Fibrosis Lungs. *Am. J. Respir. Crit. Care Med.* **202**, 1540–1550 (2020).
10. Ruiz García, S. et al. Novel dynamics of human mucociliary differentiation revealed by single-cell RNA sequencing of nasal epithelial cultures. *Development* **146**, https://doi.org/10.1242/dev.177428 (2019).
11. Wang, A. et al. Single-cell multiomic profiling of human lungs reveals cell-type-specific and age-dynamic control of SARS-CoV2 host genes. *Elife* **9**, https://doi.org/10.7554/eLife.62522 (2020).
12. Yuan, F. et al. Transgenic ferret models define pulmonary ionocyte diversity and function. *Nature* **621**, 857–867 (2023).
13. Cai, Q. et al. Sonic Hedgehog signaling is essential for pulmonary ionocyte specification in human and ferret airway epithelia. *Am. J. Respir. Cell Mol. Biol.* https://doi.org/10.1165/rcmb.2022-0280OC (2023).
14. Lei, L. et al. CFTR-rich ionocytes mediate chloride absorption across airway epithelia. *J. Clin. Invest.* **133**, https://doi.org/10.1172/JCI171268 (2023).
15. Žilionis, R. et al. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* https://pubmed.ncbi.nlm.nih.gov/31209336/ (2018).
16. Shivaraju, M. et al. Airway stem cells sense hypoxia and differentiate into protective solitary neuroendocrine cells. *Science* **371**, 52–57 (2021).
17. Sell, E. A., Ortiz-Carpena, J. F., Herbert, D. R. & Cohen, N. A. Tuft cells in the pathogenesis of chronic rhinosinusitis with nasal polyps and asthma. *Ann. Allergy Asthma Immunol.* **126**, 143–151 (2021).
18. Chen, L. et al. Inflammation-induced loss of CFTR-expressing airway ionocytes in non-eosinophilic asthma. *Respirology.* https://doi.org/10.1111/resp.14833 (2024).
19. Romano Ibarra, G. S. et al. IL-13 induces loss of CFTR in ionocytes and reduces airway epithelial fluid absorption. *J. Clin. Invest.* **134**. https://doi.org/10.1172/JCI181995 (2024).
20. Sikkema, L. et al. An integrated cell atlas of the lung in health and disease. *Nat. Med.* **29**, 1563–1577 (2023).
21. Schneider, C. et al. A Metabolite-Triggered Tuft Cell-ILC2 Circuit Drives Small Intestinal Remodeling. *Cell* **174**, 271–284.e14 (2018).
22. Gerbe, F. et al. Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth parasites. *Nature* **529**, 226–230 (2016).
23. Howitt, M. R. et al. Tuft cells, taste-chemosensory cells, orchestrate parasite type 2 immunity in the gut. *Science* **351**, 1329–1333 (2016).

24. von Moltke, J., Ji, M., Liang, H.-E. & Locksley, R. M. Tuft-cell-derived IL-25 regulates an intestinal ILC2–epithelial response circuit. *Nature* **529**, 221–225 (2015).

25. Alladina, J. et al. A human model of asthma exacerbation reveals transcriptional programs and cell circuits specific to allergic asthma. *Sci. Immunol.* **8**, eabq6352 (2023).

26. Okuda, K. et al. Secretory Cells Dominate Airway CFTR Expression and Function in Human Airway Superficial Epithelia. *Am. J. Respir. Crit. Care Med.* **203**, 1275–1289 (2021).

27. Shah, V. S. & Rajagopal, J. The TIPping point: Inflammation and Ionocyte loss. *Respirology* https://doi.org/10.1111/resp.14857 (2024).

28. Yamashita, J., Ohmoto, M., Yamaguchi, T., Matsumoto, I. & Hirota, J. Skn-1a/Pou2f3 functions as a master regulator to generate Trpm5-expressing chemosensory cells in mice. *PLoS One* **12**, e0189340 (2017).

29. Basil, M. C. et al. Human distal airways contain a multipotent secretory cell that can regenerate alveoli. *Nature* **604**, 120–126 (2022).

30. Kadur Lakshminarasimha Murthy, P. et al. Human distal lung maps and lineage hierarchies reveal a bipotent progenitor. *Nature* **604**, 111–119 (2022).

31. Sikkema, L. et al. An integrated cell atlas of the human lung in health and disease. bioRxiv, 2022.03.10.483747. https://doi.org/10.1101/2022.03.10.483747 (2022).

32. Shah, A. S. Motile cilia of human airway are chemosensory. *Science* **325**, 1131–1134 (2009).

33. Gu, X. et al. Chemosensory Functions for Pulmonary Neuroendocrine Cells. *Am. J. Respir. Cell Mol. Biol.* **50**, 637–646 (2014).

34. Yu, W., Moninger, T. O., Rector, M. V., Stoltz, D. A. & Welsh, M. J. Pulmonary neuroendocrine cells sense succinate to stimulate myoepithelial cell contraction. *Dev. Cell* **57**, 2221–2236.e5 (2022).

35. Saint-Criq, V. et al. Choice of differentiation media significantly impacts cell lineage and response to CFTR modulators in fully differentiated primary cultures of cystic fibrosis human airway epithelial cells. *Cells* **9**, 2137 (2020).

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

37. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).

38. Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975–978 (2017).

39. Ma, S. et al. Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin. *Cell* **183**, 1103–1116.e20 (2020).

40. Kotas, M. E. et al. IL-13-programmed airway tuft cells produce PGE2, which promotes CFTR-dependent mucociliary function. *JCI Insight* **7**. https://doi.org/10.1172/jci.insight.159832 (2022).

41. Kohanski, M. A. et al. Solitary chemosensory cells are a primary epithelial source of IL-25 in patients with chronic rhinosinusitis with nasal polyps. *J. Allergy Clin. Immunol.* **142**, 460–469.e7 (2018).

42. Beale, J. et al. Rhinovirus-induced IL-25 in asthma exacerbation drives type 2 immunity and allergic pulmonary inflammation. *Sci. Transl. Med.* **6**, 256ra134 (2014).

43. Tojima, I. et al. Evidence for the induction of Th2 inflammation by group 2 innate lymphoid cells in response to prostaglandin D2 and cysteinyl leukotrienes in allergic rhinitis. *Allergy* **74**, 2417–2426 (2019).

44. Lee, R. J. et al. Bacterial d-amino acids suppress sinonasal innate immunity through sweet taste receptors in solitary chemosensory cells. *Sci. Signal.* **10**, https://doi.org/10.1126/scisignal.aam7703 (2017).

45. Lee, R. J. et al. Bitter and sweet taste receptors regulate human upper respiratory innate immunity. *J. Clin. Invest.* **124**, 1393–1405 (2014).

46. Jackson, N. D. et al. Single-Cell and Population Transcriptomics Reveal Pan-epithelial Remodeling in Type 2-High Asthma. *Cell Rep.* **32**, 107872 (2020).

47. Melms, J. C. et al. A molecular single-cell lung atlas of lethal COVID-19. *Nature* **595**, 114–119 (2021).

48. Rane, C. K. et al. Development of solitary chemosensory cells in the distal lung after severe influenza injury. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **316**, L1141–L1149 (2019).

49. Barr, J. et al. Injury-induced pulmonary tuft cells are heterogenous, arise independent of key Type 2 cytokines, and are dispensable for dysplastic repair. *Elife* **11**. https://doi.org/10.7554/eLife.78074 (2022).

50. Ireland, A. S. et al. Basal cell of origin resolves neuroendocrine-tuft lineage plasticity in cancer. bioRxivorg. https://doi.org/10.1101/2024.11.13.623500 (2024).

51. Bairakdar, M. D. et al. Learning the cellular origins of cancer using single-cell chromatin landscapes. bioRxiv 2025.04.06.646131; https://doi.org/10.1101/2025.04.06.646131 (2025).

52. Guseh, J. S. et al. Notch signaling promotes airway mucous meta-plasia and inhibits alveolar development. *Development* **136**, 1751–1759 (2009).

53. Feldman, M. B., Wood, M., Lapey, A. & Mou, H. SMAD Signaling Restricts Mucous Cell Differentiation in Human Airway Epithelium. *Am. J. Respir. Cell Mol. Biol.* **61**, 322–331 (2019).

54. Stuart, T. et al. Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888–1902.e21 (2019).

55. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).

56. Shah, V. S. et al. Tuft-Ionocyte Progenitor (TIP) cells. TsankovLab/LungRareCells: v1.0.0. https://doi.org/10.5281/zenodo.1529836. (2025).

Acknowledgements

Author contributions

V.S.: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. A.W.: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. B.L.: Formal analysis, Investigation, Writing – original draft, Writing – review & editing. A.B.: Data curation, Formal analysis, Investigation, Software, Visualization, Writing – review & editing. I.M.: Data curation, Formal analysis, Investigation, Software, Visualization, Writing – review & editing. M.S.: Project administration, Supervision, Investigation. B.G.: Data curation, Formal analysis,

Investigation, Software, Visualization, Writing – review & editing. S.K.: Formal analysis. D.S.: Formal analysis. K.X.: Investigation, Software. E.P.: Investigation, Software. M.B.: Investigation, Software. J.X.: Investigation. J.W.: Investigation. D.D.: Investigation. L.N.: Investigation. W.L.: Investigation. P.C.: Investigation. C.M.: Investigation. J.S.: Investigation. M.V.S.: Investigation. L.C.C.Y.: Investigation. O.R.-R.: Project administration, Supervision. T.D.: Investigation. S.V.S.: Formal analysis, Resources. A.R.: Funding acquisition, Resources, Project administration, Supervision, Writing – original draft, Writing – review & editing. J.R.: Funding acquisition, Resources, Project administration, Supervision, Writing – original draft, Writing – review & editing. A.T.: Funding acquisition, Conceptualization, Formal analysis, Investigation, Methodology, Resources, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Competing interests

Additional information

Correspondence and requests for materials should be addressed to Aviv Regev, Jayaraj Rajagopal or Alexander M. Tsankov.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[1]Simches 4, Northeast Corridor, Massachusetts General Hospital, Boston, MA, USA. [2]Department of Internal Medicine, Division of Pulmonary and Critical Care Medicine, Massachusetts General Hospital, Boston, MA, USA. [3]Harvard Stem Cell Institute, Cambridge, MA, USA. [4]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [5]Immunology and Inflammation Research Therapeutic Area, Sanofi, Cambridge, MA, USA. [6]Developmental, Molecular and Chemical Biology, Tufts University School of Medicine, Boston, MA, USA. [7]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai (ISMMS), New York, USA. [8]Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [9]Genentech, South San Francisco, CA, USA. [10]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. [11]Harvard College, Boston, MA, USA. [12]Department of Cell and Cancer Biology, University of Toledo College of Medicine and Life Sciences, Toledo, OH, USA. [13]These authors contributed equally: Viral S. Shah, Avinash Waghray, Brian Lin, Atharva Bhagwat, Isha Monga, Michal Slyper, Bruno Giotti. [14]These authors jointly supervised this work: Aviv Regev, Jayaraj Rajagopal, Alexander M Tsankov. ✉e-mail: aviv.regev.sc@gmail.com; jrajagopal@mgh.harvard.edu; alexander.tsankov@mssm.edu