A Single-Cell Atlas of the Human Healthy Airways

Marie Deprez1*, Laure-Emmanuelle Zaragosi1*, Marin Truchi1, Christophe Becavin1, Sandra Ruiz García1, Marie-Jeanne Arguel1, Magali Plaisant1, Virginie Magnone1, Kevin Lebrigand1, Sophie Abelanet1, Frédéric Brau1, Agnès Paquet1, Dana Pe’er2, Charles-Hugo Marquèt3, Sylvie Leroy1,3,‡ and Pascal Barbry1‡

1Université Côte d’Azur, CNRS, Institut Pharmacologie Moléculaire et Cellulaire, Sophia-Antipolis, France; 2Memorial Sloan Kettering Cancer Center, New York, New York; and 3Université Côte d’Azur, Centre Hospitalier Universitaire de Nice, Fédération Hospitalo-Universitaire OncoAge, CNRS, Inserm, Institute for Research on Cancer and Aging Nice Team 3, Pulmonology Department, Nice, France

Abstract

Rationale: The respiratory tract constitutes an elaborate line of defense that is based on a unique cellular ecosystem.

Objectives: We aimed to investigate cell population distributions and transcriptional changes along the airways by using single-cell RNA profiling.

Methods: We have explored the cellular heterogeneity of the human airway epithelium in 10 healthy living volunteers by single-cell RNA profiling. A total of 77,969 cells were collected at 35 distinct locations, from the nose to the 12th division of the airway tree.

Measurements and Main Results: The resulting atlas is composed of a high percentage of epithelial cells (89.1%) but also immune (6.2%) and stromal (4.7%) cells with distinct cellular proportions in different regions of the airways. It reveals differential gene expression between identical cell types (suprabasal, secretory, and multiciliated cells) from the nose (MUC4, PI3, SIX3) and tracheobronchial (SCGB1A1, TFF3) airways. By contrast, cell-type-specific gene expression is stable across all tracheobronchial samples. Our atlas improves the description of ionocytes, pulmonary neuroendocrine cells, and brush cells and identifies a related population of NRP+ positive cells. We also report the association of KRT13 with dividing cells that are reminiscent of previously described mouse “hillock” cells and with squamous cells expressing SCEL and SPRR1A/B.

Conclusions: Robust characterization of a single-cell cohort in healthy airways establishes a valuable resource for future investigations. The precise description of the continuum existing from the nasal epithelium to successive divisions of the airways and the stable gene expression profile of these regions better defines conditions under which relevant tracheobronchial proxies of human respiratory diseases can be developed.

Keywords: single-cell RNAseq; epithelium; nose; trachea; bronchus

The prevalence of chronic respiratory diseases is believed to arise in part because of exposure to diverse atmospheric contaminants (respiratory microbes, pollution, allergens, and smoking) that interact with the respiratory epithelium. The respiratory tract constitutes an elaborated line of defense based on a unique cellular ecosystem. Thus, secretory and multiciliated cells form a self-clearing mechanism that efficiently removes inhaled particles from the upper airways, impeding their transfer to

(Received in original form November 15, 2019; accepted in final form July 28, 2020)

*These authors contributed equally to this work.

‡These authors contributed equally to this work.

Supported by grants from Fondation pour la Recherche Médicale (DEQ20180339158), the association Vaincre la Mucoviscidose (RF20180502290), the Chan Zuckerberg Initiative (Silicon Valley Foundation, 2017-175159-5022), ANR SAHARRA (ANR-19-CE14-0027), and H2020-SC1-BHC-2018-2020 Discovair (grant agreement 874656). The UCAGenomix platform, a partner of the National Infrastructure France Génomique, is supported by Commissariat aux Grands Investissements (ANR-10-INBS-09-03, ANR-10-INBS-09-02), Conseil Départemental des Alpes Maritimes (2016-294DGADSH-CV), Vaincre la Mucoviscidose (RF20180502290), and Canceropôle Provence Alpes Côte d’Azur. P.B. is recipient of a chair from ANR-19-P1IA-0002-3IA Côte d’Azur–Nice–Interdisciplinary Institute for Artificial Intelligence.

Author Contributions: Conception and design: L.-E.Z., S.L., and P.B. Clinical management: C.-H.M. and S.L. Experimental work: L.-E.Z., M.T., C.B., K.L., and P.B. Biostatistical support: A.P. Bioinformatics auditing: D.P. Drafting the manuscript for important intellectual content: M.D., L.-E.Z., and P.B.

Correspondence and requests for reprints should be addressed to Pascal Barbry, Ph.D., Université Côte d’Azur and CNRS, Institut de Pharmacologie Moléculaire et Cellulaire, 660 route des lucioles, F06560 Sophia-Antipolis, France. E-mail: barbry@ipmc.cnrs.fr; or Laure-Emmanuelle Zaragosi, Ph.D., Université Côte d’Azur and CNRS, Institut de Pharmacologie Moléculaire et Cellulaire, 660 route des lucioles, F06560 Sophia-Antipolis, France. E-mail: zaragosi@ipmc.cnrs.fr.

This article has a related editorial.

This article has an online supplement, which is accessible from this issue’s table of contents at www.atsjournals.org.

Am J Respir Crit Care Med Vol 202, Iss 12, pp 1636–1645, Dec 15, 2020
Copyright © 2020 by the American Thoracic Society
Originally Published in Press as DOI: 10.1164/rccm.201911-2199OC on July 29, 2020
Internet address: www.atsjournals.org
At a Glance Commentary

Scientific Knowledge on the Subject: Single-cell RNA profiling has already been applied to nearly normal airway samples, but so far, no data set includes biopsies collected from young healthy adults at distinct and well-identified macroanatomical regions in the airways.

What This Study Adds to the Field: Our data set provides the first picture of the heterogeneity of gene expression at a single-cell level across different sites of biopsies located between the nose and the 12th division of the airway tree.

deeper lung zones. Several mechanical filters (the nose, pharynx, and ramified structure of the lung airways) further limit the influx of pathogens and inhaled particles downward within the bronchial tree. Although the nose and bronchus share many cellular properties, which has led to the definition of a pathophysiological continuum in allergic respiratory diseases (1, 2), they differ by features such as host defense against viruses, oxidative stress (3), or antibacterial mechanisms (4, 5). In the framework of the Human Cell Atlas (HCA) consortium, we have now established a precise airway epithelium cell atlas in a population of 10 healthy living volunteers. Minimally invasive methods were set up to collect biopsies and brushings using bronchoscopy. A high-quality data set of 77,969 single cells comprising a large panel of epithelial cell subtypes was generated from 35 distinct samples taken at precise positions in the nose, trachea, and bronchi. Data integration and analysis provide a unique view of the cell type proportions and gene signatures from the first to approximately the 12th division of the airways. The resulting picture defines a relatively stable cellular composition and gene expression across the first 12 successive generations of the tracheobronchial tree. The largest differences were found between nasal and tracheobronchial samples.

Some of the results of these studies have been previously reported in the form of a preprint (https://doi.org/10.1101/2019.12.21.884759).

Methods

The atlas of the airway epithelium was obtained from biopsies and brushings from 10 healthy nonsmoking volunteers. Each donor was sampled 4–5 times in different regions of the upper (nose) and lower airways (tracheal, intermediate, and distal bronchi), located in different lobes (Figure E1 in the online supplement and Table E1). Single-cell capture was performed using the 10× Genomics Chromium device (3’ V2) (10X Genomics). Large integrative analysis of the 35 samples composing the atlas was done using fastMNN (6), and analysis was performed using Scannpy (7). Cell-type annotation was based on hg19, but we also mapped the 35 samples on the human genome Grch38 3.0.0 (version 3.0.0 of Genome Reference Consortium Human Build 38) using CellRanger 3.0.2 (10X Genomics). After concatenation with Scannpy, cells and genes were filtered based on hg19 quality control. Additional differential gene expression analysis was undertaken using edgeR (8) to investigate both cell distributions and gene expression heterogeneity along the airways. Differences between nasal and tracheobronchial compartments (suprabasal, secretory, and multiciliated cells) were specifically analyzed after creating pseudobulk samples for each cell cluster. The method, detailed in the online supplement, summed gene expression from equal numbers of randomly picked cells in each sample. This ensured an equivalent gene expression background among all bulk samples. Trajectory inference (partition-based graph abstraction, PAGA) (7) and gene network inference (GRNBoost2) (7) were also performed to further characterize the identified cell populations. Results were validated using RNAscope and immunostainings. Additional details on the methods are provided in the online data supplement.

Results

Building a Molecular Cell Atlas of the Airways in Healthy Volunteers

Data collection. Cells were analyzed by single-cell RNA sequencing (scRNA-seq) after isolation from four distinct locations using two sampling methods: 1) nasal biopsies (3 samples) and 2) nasal brushings (4 samples), 3) tracheal biopsies (carina, 1st division, 9 samples), 4) intermediate bronchial biopsies (5th–6th divisions, 10 samples), and 5) distal brushings (9th–12th divisions, 9 samples) in 10 healthy volunteers (Figures 1A and 1B and E1 and Table E1). Optimized protocols allowed the profiling of 77,969 single cells that were collected at 35 distinct positions of the airways, resulting in the detection of an average of 1,892 expressed genes per cell with 7,070 unique molecular identifiers per cell (Figures E2A and E2B).

After batch correction and graph-based clustering, cell types were assigned to each cluster using well-established sets of marker genes (Figures 1C and E3A and E3B). We identified 14 epithelial cell types, including 12 for the surface epithelium and 2 for submucosal glands, which collectively represented 89.1% of total cells (Figures 1C–1E and Table E2). A similar cell typing was found when data was mapped on either hg19 (Figure 1C) or hg38 (Figure E3C). All data (hg19 and hg38) can be accessed through our interactive web tool (https://www.genomique.eu/cellbrowser/HCA/). Stromal and immune cells represented 4.7% and 6.2%, respectively, of all cells (Figure 1E).

Annotation of epithelial cells. Basal cells (KRT5, TP63, and DLK2-high) accounted for one-third of all cells (Figures 1D and 1E). We also identified suprabasal cells, characterized by low TP63 expression, decreasing gradients of KRT5 expression, and increasing gradients of KRT19 and NOTCH3 expression (9–12) (Figure 1D). We grouped club and goblet cells as “secretory cells” because these two populations could not be clustered separately and essentially differed by the level of expression of MUC5AC and MUC5B (Figure E4) (12). We detected clusters of multiciliated cells (expressing high levels of FOXJ1, TPPP3, and SNTN) and deuteromosal cells, which correspond to precursors of multiciliated cells and express several specific markers: DEU1P, FOXN4, and CDC20B (Figures 1C and 1D) (12, 13). The suprabasal, secretory, and multiciliated clusters each comprised a subcluster of cells that could only be detected in nasal samples. These clusters were labeled “Suprabasal N,” “Secretory N,” and “Multiciliated N” and will be described later in the manuscript. Two cell types were associated with submucosal glands: serous cells (expressing high levels of LTF, LYZ, and PIP) and mucous cells (expressing high levels of MUC5B but no MUC5AC).
Figure 1. A molecular cell atlas of the healthy human airways. (A) Schematic representation of the sampled anatomical regions. (B) Experimental design of the study, detailing the anatomical regions, sampling methods, number of donors, biopsies, and cells after data curation. (C) Uniform Manifold Approximation and Projection visualization of the whole human healthy airway data set. Each distinct cell type is defined by a specific color. (D) Heatmap of expression for top marker genes of each cell type. (E) Pie chart of the total proportion of each cell type identified in human airways. (F) Bar plot of the relative abundance of each cell type collected by two distinct modes of biopsies at four macroanatomical locations.
(Figures 1C and 1D). Finally, we identified 222 cells belonging to clusters of rare epithelial cells (0.3% of the cells) (Figures 1C and 1D). We also detected the presence of some alveolar cells: 10 type I (AT1) and 11 type II (AT2) pneumocytes, which were all derived from a unique distal brushing (Table E2 and Figure E5A). AT1 expressed HOPX, AGER, and SPOCK2; AT2 expressed SFTP Alpha, SFTP Beta, SFTPC, and SFTP Delta (Figure E5B).

**Immune cells: annotation and distribution along the respiratory tree.** We clustered the 4,891 immune cells into seven distinct cell types (Figure E6A). Four clusters of myeloid cells were found: 1) macrophages and 2) monocytes, mostly detected in distal brushings; 3) mast cells, mostly detected in distal brushings and, to a lesser extent, in tracheal and intermediate bronchial biopsies; and 4) dendritic cells, found everywhere. We also identified the following three clusters of lymphoid cells: T cells were found in all samples; plasma cells were exclusively found in biopsies, in line with an interstitial localization; and B cells were mostly detected in distal airway brushings (Figures E6B and E6C and Table E2). The gene regulatory network was further characterized with GRNboost2, a program that infers regulatory unit activity (14) (Figure E6D). In the lymphoid lineage, we were able to discriminate B cells (expressing high levels of MS4A1 and LTβ, and high PAX5 inferred activity) from plasma cells (expressing high levels of IGI and MZB1, and high IRF4 inferred activity) (Figures 1D and E6D). T cells and related subtypes, which our analysis did not separate well, were characterized by a high and specific transcriptional activity of the XCL1 and CD3D regulatory units (Figures E6D and E6E).

**Stromal cells: annotation and distribution along the respiratory tree.** We annotated four stromal cell types (Figure E8A) found only in biopsies, especially in the intermediate samples (Figures E8B and E8C), including endothelial cells expressing high levels of ACKR1, fibroblasts expressing high levels of FBNL1, and smooth muscle cells, characterized by high levels of DES (desmin) and high activity of the HOXA4 regulatory unit (Figures 1D and E8D). Based on specific expression of markers such as RERGL, MCAM, and PDGFRB, we also identified pericytes, a population of periendothelial mesenchymal cells with contractile properties that are located on the vascular basement membrane of capillaries (15, 16). Pericytes also share markers such as ACTA2 and MYL9 with smooth muscle cells (Figures 1D and E8E).

**Large Variations in the Composition of Epithelial Cells Distinguish Nasal and Tracheobronchial Airways**

We then compared the epithelial composition in each of the five types of samples. We noticed that the sampling mode produced a large effect on the distribution of cells; brushings collected more luminal cell types, such as multiciliated or secretory cells, whereas forceps biopsies collected cells located deeper in the tissue, such as basal,stromal, and submucosal gland cells (Figures 1F and E7 and Table E2). All subsequent comparisons were then performed on samples obtained with similar sampling methods.

Tracheal and intermediate bronchial biopsies shared very similar cell type distributions, with few differences between biopsies taken from upper, middle, and lower lobes (Figures 1F and E7). The most striking variation was for submucosal gland cells (serous and mucous cells). Their detection in three out of three nasal biopsies, and zero out of nine intermediate biopsies (Figure E7) suggests a larger density of glands in the nose and a progressive decline in smaller airways, as previously described (17–20). Comparison between nasal and distal brushing samples also showed a clear enrichment of secretory cells in nasal samples and an enrichment of multiciliated cells in distal samples (Figure 1F and E7). To characterize qualitative differences between nasal and tracheobronchial compartments, we assessed the correlations in average gene expression between each epithelial cell type. We found stronger correlations (>0.9) between cells belonging to the same cell type, in a donor-independent manner, than between cells belonging to distinct cell types (Figure 2A), confirming that cell type identity was well conserved across samples (Figures E9A and E9C). This analysis also revealed nasal-specific and tracheobronchial-specific subclusters for suprabasal, secretory, and multiciliated cells (Figures 1C and 1D and 2A and E9A and Tables E3A–E3C), characterized by differentially expressed genes. Twenty overlapping genes between suprabasal, secretory, and multiciliated cell types were associated with the nasal epithelium (Figures 2B and 2C). Among the top 16 genes shared by all three nasal cell types were SIX3 and PAX7 (Figure 2C and Table E3D), which have well-reported roles in the eye, neural, and/or neural crest-derived development (21–23) (Table E4).

Nasal and tracheobronchial gene expression were compared in suprabasal, secretory, and multiciliated cells (Tables E3A–E3C). Among multiciliated cells, LYPD2, SPRR3, and C15orf48 were enriched in nasal cells as well as ACE2, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) receptor (24) (Figure 2D and Table E3C). Among secretory cells, we noticed an enriched expression of SCGB1A1, which was absent from both the surface and submucosal glands epithelium in the nose, as well as the nasal-enriched expression of PI3 and MUC4 (Figure 2E). Thirty-seven additional transcripts were confirmed based on a comparison with the Protein Atlas database (25) (Table E5). Functional properties were inferred by gene set enrichment analysis and GRNboost2. Several regulatory units were associated with nasal cells, such as MESP1, reported as a regulator of somitic mesoderm epithelialization (26). IRF1, IRF27, and STAT1 (i.e., transcription factors related to IFN pathways) were enriched in the nasal tissue (Figure E10C and Table E6). FOXA3 regulatory unit, which promotes goblet metaplasia in mouse and induces MUC5AC and SPDEF expression (27, 28), was enriched in tracheobronchial samples (Figures E10C and E10D).

Intriguingly, dissociated nasal cells appeared larger. There was a proximodistal gradient of cell size, with the largest average size in the nose (12.56 ± 0.71) and the smallest size in the distal airways (8.77 ± 0.71) (Figure E10E). This difference correlated with the number of detected genes and unique molecular identifiers (Figures E2A and E2B).

**Identification of Rare Epithelial Cells along the Human Airways**

We identified 13 brush/tuft cells according to their high expression of LMP and RGS13 (12, 29, 30) (Figures 3A–3C). We also noticed in these cells a specific activity
Figure 2. Distinct gene expression signatures are detected between nasal and tracheobronchial airways. (A) Unsupervised hierarchical clustering of gene expression correlation between sample-specific cell types. (B and C) Venn diagrams indicating the number of specific transcripts of each cell type (secretory, suprabasal, and multiciliated cells) in tracheobronchial (B) and nasal airway epithelia (C). The size of the different subgroups is indicated after hg19 and hg38 (in brackets) mapping (derived from Table E3), together with a list of 16 nasal and 8 tracheobronchial genes expressed in common in suprabasal, secretory, and multiciliated cells. (D and E) Mean average plot of differential expression between nasal and tracheobronchial airways in multiciliated (D) and secretory cells (E). Red and blue dots indicate nasal and tracheobronchial airways overexpressed genes, respectively. Black-circled dots indicate genes that are expressed in common in suprabasal, secretory, and multiciliated cells in nasal samples. Yellow gene names indicate gene expression that has been validated at the protein level. (F) Detection by immunofluorescence of proteins that are more specifically associated with a nasal or a tracheobronchial expression in biopsies and brushings. Images are representative of three distinct subjects. Scale bars, 20 μm and 50 μm.

logCPM = log of count per million; logFC = log2 fold change; SMG = submucosal gland.
of HOXC5, HMX2, and ANXA4 regulatory units (Figure E11A). A cluster of 29 pulmonary neuroendocrine cells (PNECs) (Figure 3A) was found, mostly in tracheal and intermediate biopsies (Figure 3B). PNECs expressed the neurotransmitter-associated genes PCSK1N, SGCN, and NEB (Figure 3C), and we identified HOX6B, ASCL1, and FOXA2 as PNEC-specific regulatory units (Figure E11A). A cluster of 117 ionocytes was also identified (Figure 3A), mostly in nasal and distal brushings (Figure 3B). Ionocytes were characterized by markers such as ASCL3 and CFTR (30) (Figure 3C), and we identified ASCL3, FOXI1, and DMRT2 as ionocyte-specific regulatory units (Figure E11A). A cluster of 63 cells, labeled as “undefined rare” cells, was sampled evenly across all macroanatomical locations (Figures 3A and 3B). Relative to the other rare populations, these cells expressed more specifically NREI, STMN1, and MDD (Figure 3C) but shared the expression of HEPACAM2, HES6, AZGP1, CRYM, and LRMP with ionocytes, brush cells, and PNECs. When we searched for a correlation with the other epithelial cell types, we found a high correlation with ionocytes (>0.85), PNECs, and brush cells (>0.80) as well as with basal and suprabasal populations (>0.85) (Figure 3D). This profile appears to be intermediate between basal cells and the other rare cells. We named the last group of rare cells, multiciliated-goblet cells, a cell type that has already been described in primary cultures (12) and in patients with asthma (31). These cells express both goblet and multiciliated cell markers. In our data set, around 60 cells were found positive for both FOXIJ1 and MUC5AC. They were equally distributed between the secretory and the multiciliated cell clusters (Figure 3E). We used SoupX to remove gene counts that may emerge from cell-free mRNA contamination, thus avoiding interference with the quantification of multiciliated-goblet cells. We confirmed the presence of these cells by MUC5AC and cilia immunostaining of freshly dissociated nasal epithelium (Figure 3F) and by using RNAscope in situ RNA hybridization on nasal epithelium sections (Figure 3G). When these cells were superimposed in a PAGA representation of tracheobronchial cell lineages, they were located close to multiciliated cells, while they were located between secretory and deuterosalominal nasal cells, nearer to these latter (Figures E11B–E11E). This result supports our previous description of goblet cells as precursors of multiciliated cells in homeostatic and healthy epithelium and additionally suggests that transition through this stage may have slightly distinct dynamics between nasal and tracheobronchial epithelia (12).

Cell Proliferation within Homeostatic Airways

Before batch correction, we identified a cluster of cycling cells defined by the expression of MKI67, TOP2A, CDC20 (Figure 4A). After batch correction, these cells spread between the basal and suprabasal clusters (Figure 4B). A cell cycle analysis of all cell types identified two clusters with positive cell cycle scores. One corresponds to cycling cells (MKI67-positive) and the other to deuterosomal cells (MKI67-negative) (Figure 4C), in agreement with Ruiz Garcia and colleagues (12). Figure 4D shows Uniform Manifold Approximation and Projection graphs for the subgroup of cells that belonged to the bona fide cycling cluster with a superimposition of the cell cycle scores for G1, S, and G2/M phases, which delineates each phase of the cell cycle inside the circular embedding (Figure 4D). We noticed that the marker genes of this cycling population largely overlap with those of suprabasal cells (Figure 4E), suggesting that in the homeostatic and healthy epithelium, suprabasal cells may be the main proliferating population in the epithelium. Labeling of bronchial epithelium sections with MKI67 antibody confirmed the presence of MKI67+/KRT13+ cells that were located in a para/suprabasal position (Figure 4F).

Cycling cells were distributed all across the 35 samples, although with a highly variable distribution, which was reminiscent to the expression profile of KRT13 in suprabasal cells (Figures 4G and 4H and E12A and E12B). These KRT13-high samples displayed the highest cycling cell proportion (>20% of cycling cells, Figure 4G). In situ RNA hybridization in nasal epithelium sections confirmed an association of MKI67 RNA with cells located at a suprabasal position, some of which expressing KRT13 (Figure E12C). This association between KRT13 expression and proliferation, together with the variability of detection of these cells, is highly reminiscent of the recent description of hillocks in mouse airway epithelium (30). We confirmed the presence of KRT13+ cell clusters in nasal epithelium, with patterns very similar to those previously found in mouse (Figure 4I). It is, however, important to notice that KRT13 was also detected in an additional group of cells, located between nasal suprabasal and secretory cells in the scRNA-seq data (Figure E12B). This group of cells was devoid of MKI67 but expressed SCEL, SPRR1A, and SPRR1B (Figure E12D) (i.e., known markers of squamous/cornified epithelial cells) (32, 33).

Discussion

We have established a reference single-cell atlas of normal human airways after analyzing 35 fresh tissue samples collected by bronchoscopy in 10 healthy volunteers, resulting in a large-scale gene expression profiling that also integrated spatial information for each sample. This approach was well adapted to collect samples from the nose to the midairways but excluded the bronchiolar compartment and the parenchyma, for which alternative experimental approaches have already been proposed (31, 34). The combination of our atlas with these other data sets will enable the establishment of a comprehensive airway atlas. Our approach provides a unique opportunity to build a single-cell gene expression resource based on well-characterized healthy volunteers, who are rarely accessible in most large-scale studies. The use of bronchoscopy, a minimally invasive approach in the airways, creates a real opportunity to rapidly transfer novel information generated in the context of the HCA project to new clinical practices.

In our workflow, a critical analytical step led to robust cell type annotation of 35 single-cell RNA-sequencing experiments. Specifically, integration was performed sequentially after quantification of individual samples, merging, and batch correction. The quality of our sampling and analysis resulted in nonsignificant donor-related effects and in very high epithelial cell proportions, two important quality criteria that had not been systematically reached by the other lung atlas reports, making our resource particularly more reliable. Our conclusions were all based on observations that were made on several donors and independently confirmed. Future integration of our data set into a larger atlas with many more individuals and anatomical.
Figure 3. Detection of rare epithelial cells across human airways. (A) Focused Uniform Manifold Approximation and Projection (UMAP) visualization on the group of ionocytes, neuroendocrine, brush cells, and undefined rare cells. (B) Pie charts of the anatomical distribution of each cell type according to location (top line) or mode of sampling (bottom line). Corresponding numerical values are listed in Table E2. (C) Dot plot of the top gene markers identified per cell type of interest. (D) Unsupervised hierarchical clustering of gene expression correlation between position-specific epithelial cell types. (E) UMAP visualization of double-positive FOXJ1\(^{+}\) MUC5AC\(^{+}\) cells (purple) relative to FOXJ1\(^{+}\) cells (blue) and MUC5AC\(^{+}\) cells (green). (F) Immunostaining for MUC5AC and acetylated \(\alpha\)-tubulin showing multiciliating-goblet cells from dissociated nasal epithelium in a healthy subject. (G) RNAscope detection of a mucous-multiciliated cell in nasal tissue. Red indicates FOXJ1\(^{+}\) RNA; green indicates MUC5AC\(^{+}\) RNA. The arrowhead points at a cell expressing both FOXJ1 and MUC5AC. Images are representative of two distinct subjects. Scale bars, 10 µm. PNECs = pulmonary neuroendocrine cells.
Figure 4. Characterization of cycling cells and KRT13 expression in the healthy airway epithelium. (A and B) Highlights of cycling basal cells in global Uniform Manifold Approximation and Projection (UMAP) representations without (A) or with (B) batch correction of the embedding. (C) Violin plot of the cell cycle phase score in all cell types detected in the whole data set. (D) Focused UMAP visualizations on the subset of cycling cells, colored by cell cycle phase scores at G1, S, G2/M stages. (E) Dot plot of marker gene expression in cycling, basal, and suprabasal cells. (F) Immunostaining for MKI67 and KRT5 in a bronchial biopsy section. Scale bar, 10 μm. (G) Bar plot of the percentage of cycling cells per sample. (H) violin plots of the expression of KRT13 in suprabasal cells. (I) Immunostainings for KRT13 (green) and acetylated α-tubulin (red) in nasal turbinate whole mount (top view). Scale bars, 100 μm and 1,000 μm. Images are representative of three distinct subjects. AT1 = alveolar pneumocyte, type 1; AT2 = alveolar pneumocyte, type 2; LT = T lymphocyte; NK = natural killer cell.
locations will allow a more precise definition of regional and interindividual idiosyncrasies.

Profiling of identical cell types across many sites of the airways has allowed us to quantify the frequencies of epithelial, submucosal gland, immune, and stromal cells and has revealed an influence of the mode of sampling. However, this did not prevent us from defining stable core cell type signatures for each epithelial, stromal, and immune cell types, irrespective of their anatomical location. In contrast, important variations of gene expression were found when comparing the same populations of suprabasal, secretory, and multiciliated cells from the surface epithelium between nasal and tracheobronchial compartments. These results fit well with previous work reporting differential gene expression signatures between nasal and bronchial brushings (5, 35). Interestingly, SIX3, PAX6–7, and OTX1/2, which we found to be specific of the nasal epithelium, are all associated with gene ontology terms such as “regionalization” and “morphogenesis involved in neuron differentiation” (Table E4) and have well-described functions during embryonic patterning of the head (22, 36–38). Expression of SIX3 in murine ependymocytes, which are radial glia-derived multiciliated cells, is necessary for the maturation of these cells during postnatal stages of brain development (36). Hence, nasal-specific expression of developmental patterning genes might be the consequence of head versus trunk differential developmental origins and may not necessarily confer specific functions to nasal epithelial cells. The underlying mechanisms that confer a persistence in the expression of these developmental hallmarks remain to be elucidated. We also found an enrichment in ACE2 expression in nasal multiciliated cells, a finding that may have clinical implications in the course of infection of SARS-CoV-2. A tonic activation of IFN pathways may contribute to the increased nasal expression of specific genes, such as ACE2 (39), which fits well with our finding of enriched IFN-related genes in nasal secretory cells. These specific data have recently been included in two collaborative studies by the HCA Lung Biological Network (39, 40).

A focus on secretory cells demonstrates that nasal cells contain few SCGB1A1+ and SCGB3A1+ cells. Despite this low secretoglobin content, they display the core gene signature of secretory cells, suggesting that secretoglobins may not be sufficient marker genes to identify all secretory/club cells. These differences are important to consider when using nasal samplings as a proxy to assess bronchial status.

Our atlas sheds some light on two novel cell types, namely, the multiciliating-goblet and the undefined rare cells. Because the two populations were also found in pig trachea (12) (S. Ruiz Garcia, M. Truchi, and colleagues, unpublished results), we are convinced of the validity of these two cell categories. Even though much caution should be taken when performing trajectory inference performed with few cells, multiciliating-goblet cells may be facultative multiciliated cell precursors, a notion that is consistent with our previous in vitro work (12). Regarding the undefined rare cells, considering their intermediate profile between basal and other rare cells, it is tempting to speculate that they may be precursors for the ionocytes, PNECs, and brush cells. However, further work is clearly needed to describe more comprehensively their function and establish hierarchical lineages. As it is, our atlas already provides the first detailed identification of human PNECs and brush cells at an sRNA-seq level. Finally, our work contributes at some point to the description of airway hillocks that was initially made by Montoro and colleagues (30). We indeed found a population of KRT13+ cells that are highly reminiscent of these cells (that we confirmed by immunolabeling). We also report a second population of KRT13+ cells that express markers of squamous/cornified epithelium. The balance and the identification of the respective functions of these two populations will require further work.

Altogether, our atlas provides a significant contribution to resolve the cellular stratification of gene expression profiles in the healthy human airway epithelium. It now makes possible an extensive exploration of the various situations involved in homeostasis and regeneration of normal and pathological airways.

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank the UCAGenomix platform for fruitful discussions and technical help on single-cell RNA sequencing, Julie Cazareth from the Institut Pharmacologie Moléculaire & Cellulaire imaging platform for fruitful discussions and technical help on imaging, and Jennifer Griffonnet for her invaluable help in building an experimental protocol that respects ethical rules in collaboration with the French health authorities (Agence Nationale de Sécurité du Médicament et des produits de santé et Comité de Protection des Personnes). They also thank Professor Colin D. Bingley (University of Sheffield) for his careful reading of the manuscript.

References

1. McDougall CM, Blaylock MG, Douglas JG, Brooker RJ, Helms PJ, Walsh GM. Nasal epithelial cells as surrogates for bronchial epithelial cells in airway inflammation studies. Am J Respir Cell Mol Biol 2008;39:560–568.
2. Samitas K, Carter A, Kariyawasam HH, Xanthou G. Upper and lower airway remodelling mechanisms in asthma, allergic rhinitis and chronic rhinosinusitis: the one airway concept revisited. Allergy 2018;73:993–1002.
3. Mihaylova VT, Kong Y, Fedorova O, Sharma L, Dela Cruz CS, Pyle AM, et al. Regional differences in airway epithelial cells reveal tradeoff between defense against oxidative stress and defense against rhinovirus. Cell Rep 2018;24:3000–3007, e3.
4. Roberts N, Al Mubarak R, Francisco D, Kraft M, Chu HW. Comparison of paired human nasal and bronchial airway epithelial cell responses to rhinovirus infection and IL-13 treatment. Clin Transl Med 2018;7:13.
5. Giovannini-Chami L, Paquet A, Sanfionocco C, Pons N, Cazareth J, Magnone V, et al. The “one airway, one disease” concept in light of Th2 inflammation. Eur Respir J 2018;52:1800437.
6. Haghverdi L, Lun ATL, Morgan M, Marioni JC. Batch effects in single-cell RNA-seq data are corrected by matching mutual nearest neighbors. Nat Biotechnol 2018;36:421–427.
7. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 2018;19:15–20.
8. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139–140.
9. Boers JE, Ambergen AW, Thunissen FB. Number and proliferation of basal and parabasal cells in normal human airway epithelium. Am J Respir Crit Care Med 1998;157:2000–2006.
10. Mori M, Mahoney JE, Stupnikov MR, Paez-Cortez JR, Szymaniak AD, Varelas X, et al. Notch3-Jagged signaling controls the pool of undifferentiated airway progenitors. Development 2015;142:258–267.
11. Pardo-Saganta A, Law BM, Tata PRR, Villoria J, Saez B, Mou H, et al. Injury induces direct lineage segregation of functionally distinct airway basal stem/progenitor cell subpopulations. Cell Stem Cell 2015;16:184–197.

12. Ruiz García S, Deprez M, Lebrigand K, Cazavats A, Paquet A, Arguel M-J, et al. Novel dynamics of human mucociliary differentiation revealed by single-cell RNA sequencing of nasal epithelial cultures. Development 2019;146:dev177428.

13. Revinska Z, Zagarosi L, Boutin C, Ruiz-García S, Deprez M, Thomé V, et al. CDC20B is required for deuterosome-mediated centriole production in multiciliated cells. Nat Commun 2018;9:4686–4683.

14. Albar S, González-Bias CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods 2017;14:1083–1086.

15. Rock JR, Barkauskas CE, Cronce MJ, Xue Y, Harris JR, Liang J, et al. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. Nature 2018;560:377–381.

16. Angelidis I, Simon LM, Fernandez IE, Strunz M, Mayr CH, Greiffo FR, et al. DMAD and rostro-caudal patterning of somites. Development 2005;132:3969.

17. Widdicombe JH. Airway gland structure and function. Anatom Rec 1968;123:376–389.

18. Jeffrey PK. The development of large and small airways. Am J Respir Crit Care Med 1998;157:5174–5180.

19. Tos M. Development of the mucous glands in the human main bronchus. Anat Anz 1968;123:376–389.

20. Zilionis R, Choo-Wing R, Savova V, Knehr J, Roma G, et al. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. Nature 2018;560:377–381.

21. Zeisel A, Hochgerner H, Lönnerberg P, Johnsson A, Memic F, van der Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods 2017;14:1083–1086.

22. Angelidis I, Simon LM, Fernandez IE, Strunz M, Mayr CH, Greiffo FR, et al. An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. Nat Commun 2019;10:963.

23. Widdicombe JH, Wine JJ. Airway gland structure and function. Physiol Rev 2015;95:1241–1319.

24. Angelidis I, Simon LM, Fernandez IE, Strunz M, Mayr CH, Greiffo FR, et al. A Single Cell Atlas of the Human Healthy Airways. Nature 2018;560:377–381.