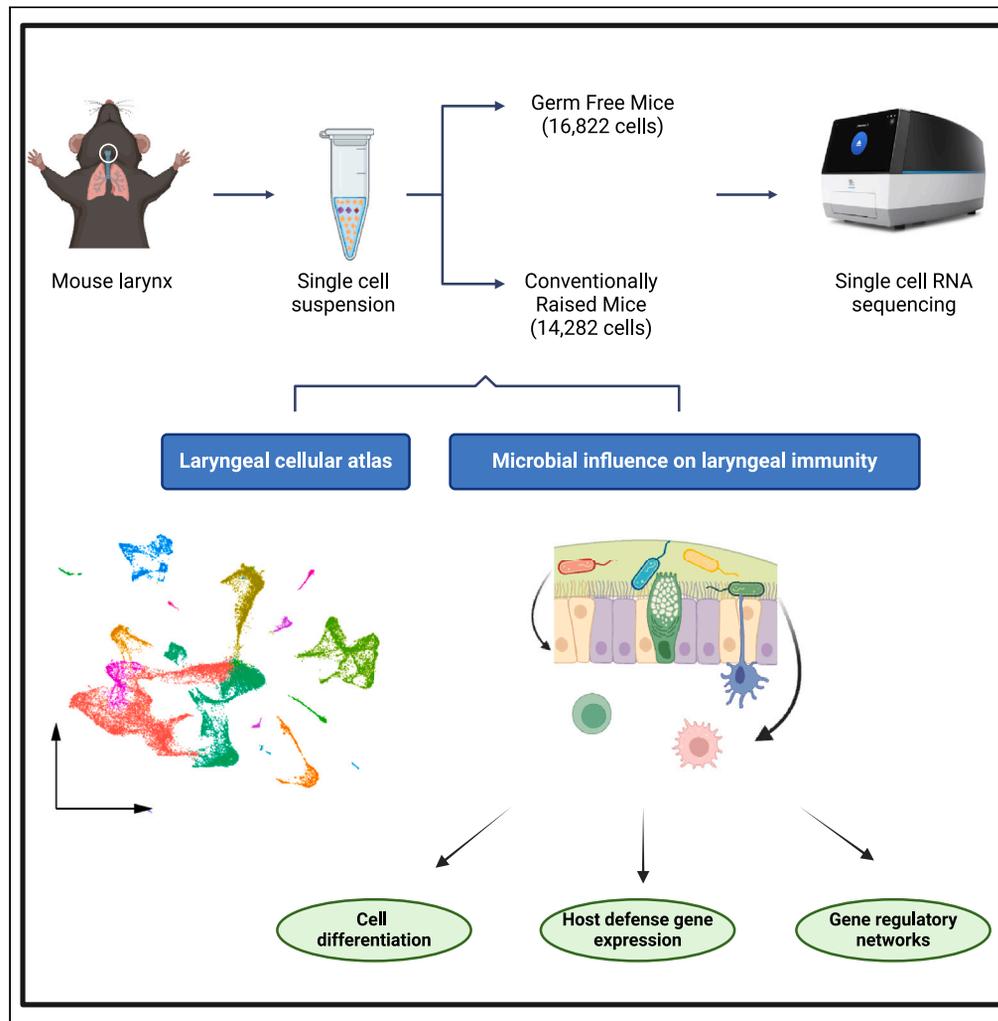**Article**

Single-cell view into the role of microbiota shaping host immunity in the larynx

Ran An, Zijian Ni, Elliott Xie, Federico E. Rey, Christina Kendziorski, Susan L. Thibeault

thibeault@surgery.wisc.edu

Highlights

The cellular landscape of the larynx was established in a mouse model

Multiple macrophage and secretory epithelial cell populations were found

Microbiota have extensive impact on the laryngeal immune system

Microbiota regulate cell differentiation and increase host defense genes expression

Article

Single-cell view into the role of microbiota shaping host immunity in the larynx

Ran An,[1] Zijian Ni,[2] Elliott Xie,[3] Federico E. Rey,[4] Christina Kendziorski,[3] and Susan L. Thibeault[1,5,*]

SUMMARY

Microbiota play a critical role in the development and training of host innate and adaptive immunity. We present the cellular landscape of the upper airway, specifically the larynx, by establishing a reference single-cell atlas, while dissecting the role of microbiota in cell development and function at single-cell resolution. We highlight the larynx's cellular heterogeneity with the identification of 16 cell types and 34 distinct subclusters. Our data demonstrate that commensal microbiota have extensive impact on the laryngeal immune system by regulating cell differentiation, increasing the expression of genes associated with host defense, and altering gene regulatory networks. We uncover macrophages, innate lymphoid cells, and multiple secretory epithelial cells, whose cell proportions and expressions vary with microbial exposure. These cell types play pivotal roles in maintaining laryngeal and upper airway health and provide specific guidance into understanding the mechanism of immune system regulation by microbiota in laryngeal health and disease.

INTRODUCTION

The larynx is a vital mucosal organ positioned at the intersection between the digestive and upper respiratory tracts (URT), with the vocal folds (VFs) serving as a valving mechanism for voicing and airway protection from ingested food and drink during swallowing.[1] Vocal fold (VF) disorders affect not only swallowing, breathing, and coughing but also lead to impaired voice production. It is estimated that voice disorders affect 1 in every 13 adults annually in the US, with infectious laryngitis being the most common diagnosis.[2] Total health care costs for these laryngeal disorders exceeds two hundred million dollars per year, causing far-ranging social and economic burden.[3] Many laryngeal diseases, such as acute/chronic laryngitis and benign vocal fold lesions, are associated with the inflammation of the mucosa, which is intimately linked to local immunity, based on studies in other part of the body.[4] However, a paucity exists in terms of a comprehensive understanding of the immune system in the larynx.

Current evidence suggests that the larynx possesses a distinct immunological architecture, embodied by the laryngeal epithelium as structural defense, lining its entire surface and professional immune cells, including macrophages, dendritic cells (DCs), T and B lymphocytes, natural killer (NK) cells, and granulocytes, are diffused throughout the epithelium and underlying lamina propria.[5] A small portion of immune cells are present in laryngeal-associated lymphoid tissue (LALT), a highly organized structure localized to the surface of the epiglottis and false VFs (for a detailed review see Thibeault et al.). The epithelium, covered by a thin mucus layer,[6] appears to be more than a physical barrier to external stimuli given the expression of Major Histocompatibility Complex (MHC) class I molecules, nonclassical MHC I molecule CD1d with unknown role in the superficial epithelial layer, as well as MHC II molecules through unknown mechanisms in the absence of cytokine induction.[5,7] Moreover, the presence of multiple lymphocyte subsets has been reported, including CD4[+] T cells, CD8[+] T cells, and a specialized subset of CD4[+] T lymphocytes T regulatory cells in the deeper layers of the epithelium, as well as natural killer T (NKT)– like lymphocytes co-expressing CD3[+] CD161[+] in the CD1d superficial layer of the epithelium.[8–10] Given the role of NKT cells in some bacterial infections, the presence of these immune cells in the superficial layer of the epithelium has been speculated to be associated with bacterial flora living in the mucosa.[9]

Indeed, the major driving force of local immune system development is the need to maintain homeostatic relationships with the microbiota that occupy the specific site.[11] Extensive microbiome research in other mucosal systems (i.e., intestine, lung, skin) has proven the impact of commensal and mutualistic microbes on host physiology in various aspects from development to metabolism, immunity, and behavior.[4,12,13] Laryngeal microbiota, though low in density, has been characterized over the past decade, in healthy and diseased populations with laryngeal pathologies spanning from chronic laryngitis, benign vocal fold lesions, and carcinoma, as well as multiple animal models.[14,15] Even though microbial profiles vary with health status, environmental exposure, age, and even genetics, *Prevotella*, *Streptococcus*, and *Veillonella* are the dominant genera in the larynx under selective pressure imposed by this specific environment.[16,17] Multiple lines of evidence suggest an

[1]Department of Surgery, School of Medicine and Public Health (SMPH), University of Wisconsin-Madison, Madison, WI, USA
[2]Department of Statistics, College of Letters and Sciences , UW-Madison, Madison, WI, USA
[3]Department of Biostatistics and Medical Informatics, School of Medicine and Public Health, UW-Madison, Madison, WI, USA
[4]Department of Bacteriology, College of Agriculture and Life Sciences, UW-Madison, Madison, WI, USA
[5]Lead contact
*Correspondence: thibeault@surgery.wisc.edu
https://doi.org/10.1016/j.isci.2024.110156

association between specific bacteria and VF pathology and rapid laryngeal immune responses to *in vitro* challenges with a single or mix of bacteria or viruses.[15,18] However, our knowledge of the laryngeal microbiota and its association with host immunity remains limited. To comprehensively define cell types/subtypes in the larynx and their immune response to commensal microbiota, we performed single-cell RNA sequencing on larynges collected from healthy conventionally raised (ConvR) and germ free (GF) mice. This study demarcates the cellular landscape of the larynx at a single cell resolution with global transcriptomic profiling of the immune responses to resident commensal microbiota.

RESULTS

Characterization of microbial profile in the mouse larynx

The commensal microbiota profile in the ConvR mouse larynx was characterized via 16S rRNA gene sequencing. Low microbial biomass samples are susceptible to contaminating DNA present in the environment, lab supplies, and reagents used for sample collection, DNA extraction, and library preparation.[19] Given the low microbial biomass nature of the larynx, false positive signals were expected in extraction negative controls ExNeg3&4 (Figures S1A and S1B). However, contaminant species in these extraction negative controls were negligible given their disproportionally low abundances compared with that of ConvR samples (Figure S1A). Dominant genera in ConvR included *Hae-mophilus, Streptococcus, Fusobacterium, Vellonella,* and *Prevotella* observed across all ConvR samples, with relative abundances of 17.1%, 13.0%, 12.7%, 11.2%, and 7.3%, respectively (Figure S1B). Considering that *Streptococcus* is underrepresented in positive controls of the oral microbiome whole cell mix, which is likely due to the potential bias of the current sample processing protocol toward certain groups of microorganisms, the relative abundance of *Streptococcus* in the larynx could be underestimated.

Establishing a single cell atlas of the murine larynx

Overview of cell type proportions

Larynges of male mice aged 8–10 weeks were collected fresh, pooled (ConvR or GF) and dissociated into single cells. We then performed scRNA-seq on a high-throughput, droplet-mediated scRNA-seq platform. After the removal of low-quality cells, a total of 16,822 (ConvR) and 14,282 (GF) cells were recovered for further analysis. Unsupervised graph-based clustering and cell type annotation based on *CellKb*, with manual refinement according to the expression of canonical markers in each cluster resulted in 16 cell types in both mouse groups (Figures 1A and 1B). The top 5 upregulated genes in each cluster were considered signature genes for each cell type (Figure 1C). Epithelial cells, including basal (BEC), suprabasal (SBEC), secretory (SEC), ciliated (CEC) epithelial cells, tuft cells (TC, aka brush cells), and cycling basal cells (CBC) dominated, representing 65.6% (62.0%) in ConvR (GF); fibroblasts, endothelial cells (EC), phagocytes, and lymphocytes represent 13% (14.5%), 8.9% (11.6%), 5.5% (5.6%), and 0.7% (0.7%) of total cells in ConvR (GF) (Figure S2A). Other cell types, including skeletal muscle cells, smooth muscle cells, thyroid follicular cells, lymphatic endothelial cells, Schwann cells, and neurons, constituted 6.4% (5.4%) of total cells in ConvR (GF) (Figures 1A and S2A). Given their low cell proportions, skeletal muscle cells, smooth muscle cells, thyroid follicular cells, lymphatic endothelial cells, Schwann cells, and neurons are not discussed in the present study. Previous literature suggests that microbes alter the cell quantity of the gut epithelium.[20,21] However, no significant difference was observed in the cell proportion of each major laryngeal cell type between ConvR and GF according to a scCODA analysis (FDR = 0.1, Figure 1C). It is worth noting that cell proportion is largely dependent upon the single-cell extraction protocol used.[22] Our data represents cell proportions in laryngeal tissues dissociated specifically following our established protocol.

Annotation of epithelial cells

We identified six cell types in the larynx that are of epithelial lineage. The largest epithelial cell population in the larynx, BEC (*Abi3bp, Ccdc3, Trp63,* and *Krt15*), accounted for one-third of all cells (Figure S2A). The high heterogeneity of BEC in both mouse groups was expected (Figures 1A and 1B), in that basal cells in the larynx are regionally distinct and form the basal layer of three major types of epithelia - stratified squamous epithelium (SSE) that covers the membranous VF and superficial surface of the epiglottis, respiratory epithelium (pseudostratified ciliated columnar epithelium (PCCE)) found throughout the supraglottic and subglottic regions,[1] and glandular epithelium (GE) in the sero-mucinous submucosal glands (SMGs) found throughout the larynx.[23] Two minor types of epithelia were found - transitional epithelium (stratified non-ciliated cuboidal or columnar epithelium) between areas of SSE and PCCE and the two layered stratified cuboidal epithelium lining the ventral pouch (Figure 2A). Subclustering of BEC in homeostasis yielded 6 distinct subpopulations in both mouse groups (Figures S2B1 and S2B2), including intermediate basal cells differentiating to club cells in PCCE (BEC1), basal cells with unknown function (BEC2), intermediate basal cells differentiating to goblet cells (BEC3), fast proliferating parabasal epithelial cells (BEC4), quiescent basal cells (BEC5), and hillock club cells (BEC6) expressing *Krt13/Krt4, Scgb1a1,* as well as immunomodulatory and barrier function related genes (*Anxa1, Lgals3, Cldn3*) (Figures S2B3 and S2C).[24,25] Further, we identified SBEC characterized by *Krt13, Krt17, Krt6a,* and *Ly6d*, with the low expression of *Trp63* and increased expression of *Notch3* that regulates the balance of basal and parabasal progenitor pools (Figures 1C and S2D).[26] SBEC are likely the suprabasal cells of the stratified squamous epithelium (SSE) unique to the VF and epiglottis. A cluster of cycling cells was detected around BEC and SBEC, defined by remarkably high expressions of *Stmn1, Top2a, Histh2ap, Hmgb2*, expressing many markers of the BEC and SBEC clusters (Figures 1A and 1C). As reported in earlier literature, these are likely cycling basal epithelial cells.[27,28]

SEC was the second largest epithelial cell population in the larynx regardless of microbial exposure status (Figures 1A, 1B, and S2A), characterized by multiple secretoglobin family genes (*Scgb3a1, Scgb3a2, Scgb1a1*) with cytokine-like functions potentially involved in
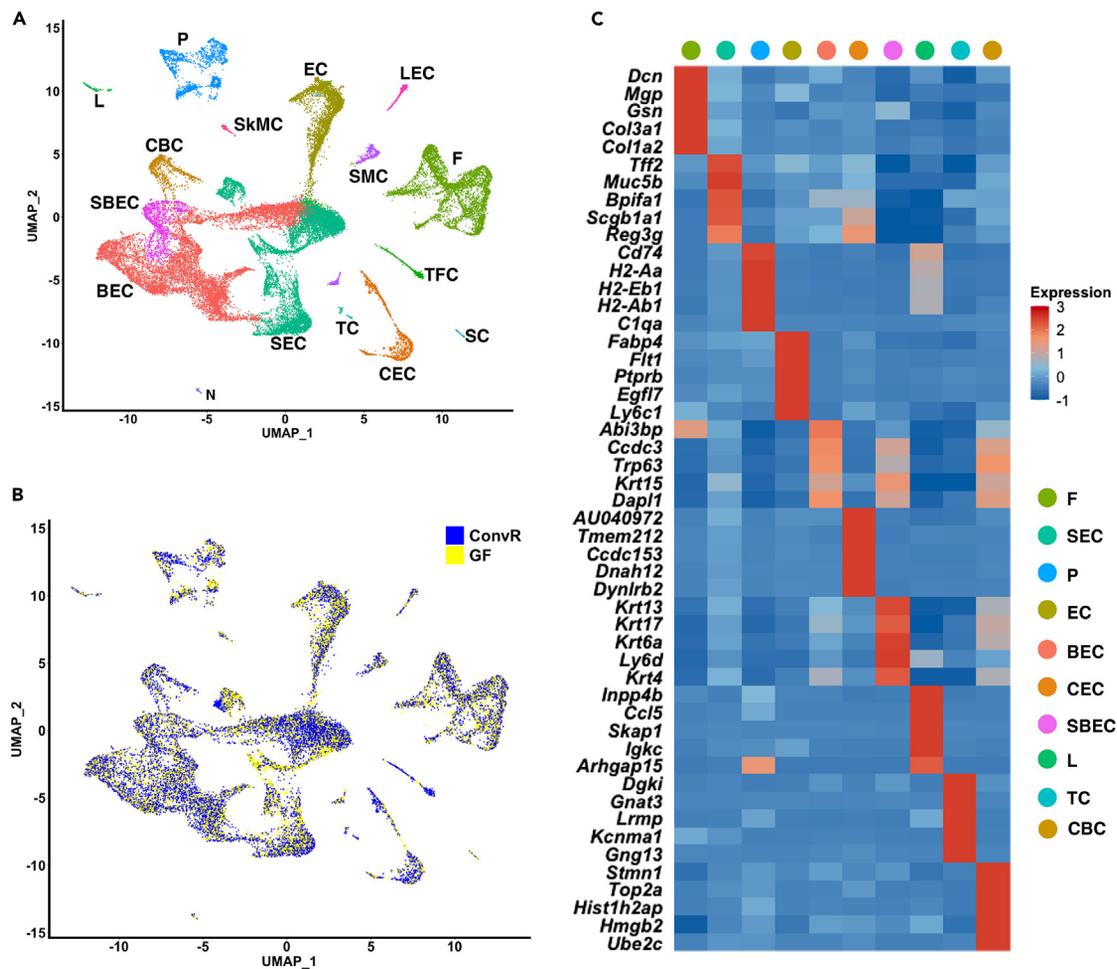
Figure 1. Laryngeal cell types in ConvR and GF mice

(A and B) UMAP plots of 15 laryngeal cell types found in ConvR and GF merged (A) and ConvR and GF mice (B). F, SEC, P, EC, BEC, CEC, SBEC, L, TC, CBC, LEC, SkMC, SMC, TFC, N, and SC represent fibroblast, secretory epithelial cell, phagocyte, endothelial cell, basal epithelial cell, ciliated epithelial cell, suprabasal epithelial cell, lymphocyte, tuft cell, cycling basal cell, lymphatic endothelial cell, skeletal muscle cell, smooth muscle cell, thyroid follicular cell, neuron, and Schwann cell, respectively.

(C) Averaged gene expression heatmap of top 5 differentially expressed (DE) genes in the larynx, ranked by average log$_2$ fold change for each major cell type for merged dataset. Color scale: red, high expression; blue, low expression. F, SEC, P, EC, BEC, CEC, SBEC, L, TC, and CBC represent fibroblast, secretory epithelial cell, phagocyte, endothelial cell, basal epithelial cell, ciliated epithelial cell, suprabasal epithelial cell, lymphocyte, tuft cell, and cycling basal cell, respectively.

anti-inflammation,[29–31] a list of antimicrobial genes (*Reg3g, Bpifb1, Ltf, Dmbt1, Lyz2, Bpifa1, Pglyrp1, Pigr*),[32–41] as well as mucus related genes (*Tff2, Muc5b, Agr, Gp2, Muc16, Tff1*)[42–45] (Table S1). SEC exhibited an average of 3- to 15-fold higher expression of these genes compared to the other cells in the larynx (Table S1). Likewise, their expression was highest in SEC when compared across epithelial cell types (Figure 2B). Functional enrichment of the upregulated genes in SEC (Table S2) indicated that these cells defend against external biotic or abiotic stimulants, capture and stabilize environmental antigens in mucus, and/or secret antimicrobial substances. Four distinct SEC subclusters in both mouse groups were detected and annotated as proliferating cells from basal cells (SEC1), club cells (SEC2), mucous cells (SEC3, both goblet cells in luminal epithelium and mucus producing cells in submucosal glands), and serous cells (SEC4) (Figures 2C1–2C3). A small subcluster SEC5 was found only in one of the ConvR samples (Figure 2C4). Functional annotation suggests the cells in each subtype are associated with cellular structure organization (SEC1), response to xenobiotic stimulus (SEC2), glycoprotein (mucin) biosynthesis (SEC3), and anti-bacterial humoral response (SEC4), further confirming their identities (Table S3).

We detected 2 clusters of CEC in the larynx, characterized by *AU040972, Ccdc153, Dnah12, Dynlrb2* genes (Figures 1A and 1C). Functions of *AU040972* and *Ccdc153* are not clear, while both *Dnah12* and *Dynlrb2* are related to cilia formation and microtubule motor activity.[46,47] Beyond the expression of CEC canonical markers, *Foxj1* and *Cdhr3*,[48] there were 18 genes exclusively upregulated in this population (average log$_2$ fold change >2, pct.1 > 0.7, pct.2 < 0.2, adjusted *p*-value <0.05), which could be larynx specific transcriptomic markers for CEC (Figures S2E and S2F). Two distinct CEC subtypes were identified in the larynx – *Bpifa1* (CEC1) and *Agr3* (CEC2), where the first was associated
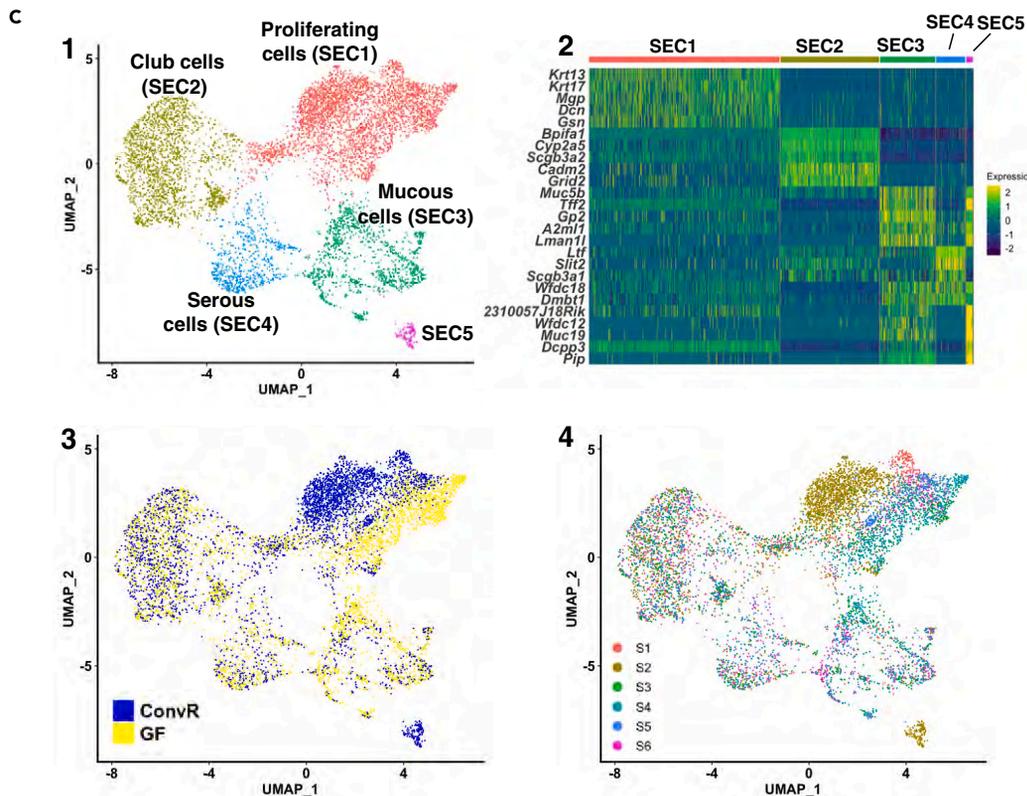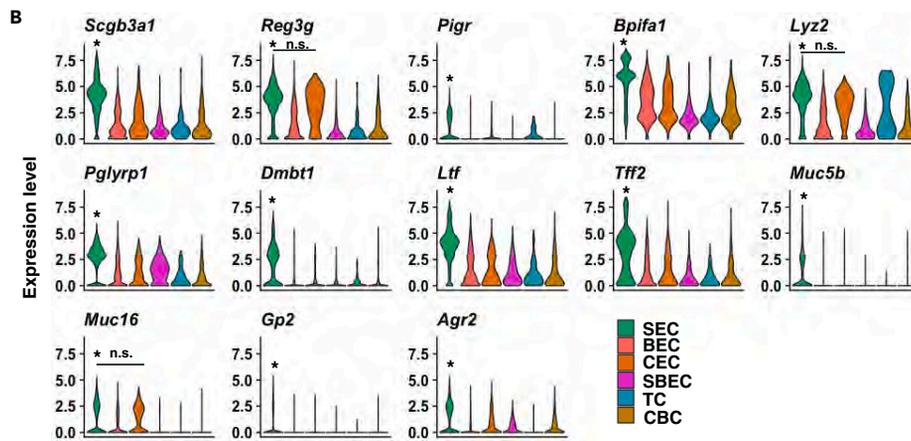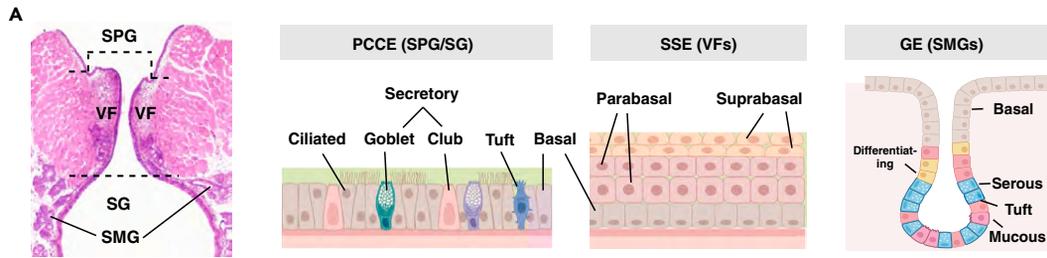
Figure 2. Epithelial cell heterogeneity in the mouse larynx

(A) Scheme of distribution of epithelial cell types in the coronal section of vocal folds. GE, SMGs, SPG, VF, SG, PCCE, and SSE represent glandular epithelium, submucosal glands, supraglottis, vocal folds, subglottis, pseudostratified columnar ciliated epithelium, and stratified squamous epithelium, respectively.

(B) Expression of antimicrobial and mucus related genes across four epithelial cell types. Asterisk indicates statistical significance of DE genes across cell types (average $\log_2$ fold change >0.25, adjusted p-value <0.05), unless specified as non-significant (n.s.) between designated cell types, for merged data. SEC, BEC, CEC, SBEC, TC, and CBC represent secretory epithelial cell, basal epithelial cell, ciliated epithelial cell, suprabasal epithelial cell, tuft cell, and cycling basal cell, respectively.

(C) 1) UMAPs of SEC subtypes identified in merged data; 2) Gene expression heatmap of top 5 differentially expressed (DE) genes (ranked by average $\log_2$ fold change) in each SEC subtype; 3) & 4) SEC subtype distribution across mouse groups (3) and biological replicates (4). SEC1-SEC5 represent secretory epithelial cell subtype 1–5. S1-S6 represent ConvR1, ConvR2, GF1, GF2, ConvR3, and GF3 samples, respectively.

with antimicrobial activities and the latter with cilia assembly and movement (Figures S3A and S3B). In addition, we detected tuft cells (TC) in the larynx, a rare epithelial cell type in the airway playing a key role in type 2 immunity.[49] The cluster was characterized by high expression of *Lrmp, Dclk1, Rgs13,* and *Cd24a* in both mouse groups (Figure 1C; Table S2).

Annotation of non-epithelial cells

Fibroblasts and endothelial cells were the two largest non-epithelial cell populations in the larynx, characterized by a group of collagen fibril associated genes (*Dcn, Mpg, Gsn, Col3a1, Col1a2*) and endothelial cell differentiation related genes (*Fabp4, Flt1, Ptprb, Egfl7, Ly6c1*) (Figures 1A and 1C). Functional enrichment of upregulated genes in these two clusters were associated with extracellular structure organization and blood vessel morphogenesis (Table S2). Unsupervised subclustering yielded 5 distinct fibroblast subtypes in both mouse groups, with each subtype fulfilling different roles in development and maintenance of lamina propria (Figures S3C1–S3C3; Table S3). Among the fibroblast subtypes, F1 and F2 were the two largest subpopulations, accounting for 45.4% (1939 cells) and 29.9% (1275 cells) of the total fibroblasts and fulfilling the major role of extracellular matrix organization and assembly (Table S3). Subtype F3 was likely involved in a defense response to bacteria, based on their gene signatures related to antimicrobial activity (Figure 3C3). Similarly, we found four distinct endothelial cell types, regardless of microbial colonization status, with EC2 (*Dach1, Cldn5, Cd36, Tcf15, Ccdc85a*) and EC3 (*Prkg1, Plvap, Ehd4, Mctp1, Insr*) being the two major subpopulations playing the role of endothelial cell development, proliferation, and migration (Figures S3D1–S3D3; Table S3).

Annotation of immune cells

Two major types of immune cells were identified in this study. A cluster of 1734 cells were identified as phagocytes, characterized by *Cd74, H2-Aa, H2-Eb1, H2-Ab1,* and *C1qb,* involved in innate immune system[50] (Figures 1A and 1C). Unsupervised subclustering yielded 7 distinct subclusters in both mouse groups, annotated as monocyte (P1), macrophage 1 (P2), dendritic cell (P3), macrophage 2 (P4), macrophage 3 (P5), dendritic cell (P6), neutrophil (P7) respectively, as per the mapping of upregulated genes of each subcluster to *CellKb*-immune database (Figures 3A1–3A3; Tables S5 and S6). This was further confirmed with a list of canonical immune cell markers (Figure 3B). Lymphocytes, the second immune cell type detected in the present study, were characterized by *Inpp4b, Ccl5, Skap1, Igkc, Arhgap15* (Figures 1A and 1C, Table S5). The four distinct subpopulations found within these cells of both mouse groups were likely innate lymphoid cells (L1), T cells (L2), B cells (L3), natural killer cells (L4), based on the mapping of the upregulated genes in each subtype to *CellKb*-immune database (Figures S4A and S4B; Table S7).

Commensal microbiota enhanced the expression of host defense associated genes within the larynx

To assess the transcriptional impact of commensal microbiota on laryngeal mucosa, we performed differential expression analysis between ConvR and GF for all laryngeal cell types and subtypes identified in our study. Compared with other epithelial cells, SEC were more responsive to long term commensal microbiota colonization and actively engaged in host-microbe interactions in the larynx. A total of 136 genes were differentially expressed in SEC between the two mouse groups (Figure 4A), of which 42 were upregulated in ConvR and 94 in GF. GO biological processes in our gene list ranked by average $\log_2$ fold change demonstrated that SEC in ConvR upregulated the expression of genes associated with organonitrogen compound biosynthetic process, regulation of lymphocyte proliferation, regulation of cytokine production, induction of bacterial agglutination, cellular response to bacterial lipopeptide and so forth, while the GF had a gene expression profile associated with protein glycosylation, signal peptide processing, peptide metabolic processes (Figure S5A; Table S8). Microbiota significantly increased the expression of host defense genes associated with epithelial barrier integrity (*Tcf7l1*), epithelial cell differentiation (*Notch2*), anti-inflammation (*Scgb1a1*), antimicrobial activity (*Sftpd, Sftpa1*), neutrophil mediated immunity (*Cxcl5*), and other host defense activities, such as iron homeostasis (*Ftl1*), xenobiotic metabolism (*Cbr2*), and serous demilune cell activity (*Dcpp1*) in ConvR (Figure 4B). We validated *Scgb1a1* expression with immunofluorescence staining (Figures 4C1 and 4C2), where expression was observed mainly in the supraglottic region, all the way down the larynx through to the trachea, with notably decreased expression on the membranous regions of VF (Figures 4C3 and 4C4). *Scgb1a1* expression, in GF mice, was remarkably reduced, observed only in the submucosal glands of subglottic region and its adjacent epithelial surface (Figures 4C5 and 4C6). We further performed differential expression analysis between ConvR and GF for SEC subtypes and found that ConvR had 10 genes upregulated in club cells (SEC2) and 3 in mucous cells (SEC3) (Figure 4D), of which *CD36* functions as a signaling receptor responding to pathogen associated molecular pattern (PAMP) molecules and *Iigp1* plays a role in resistance to
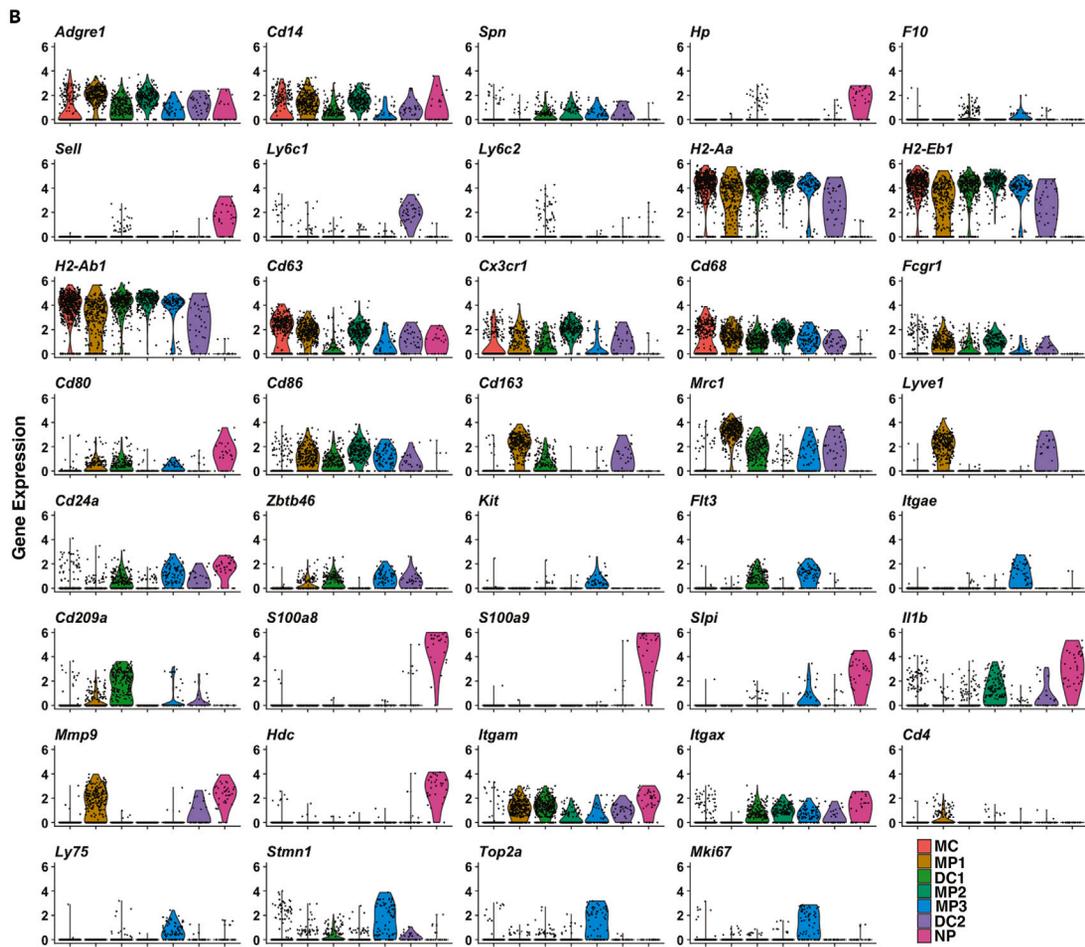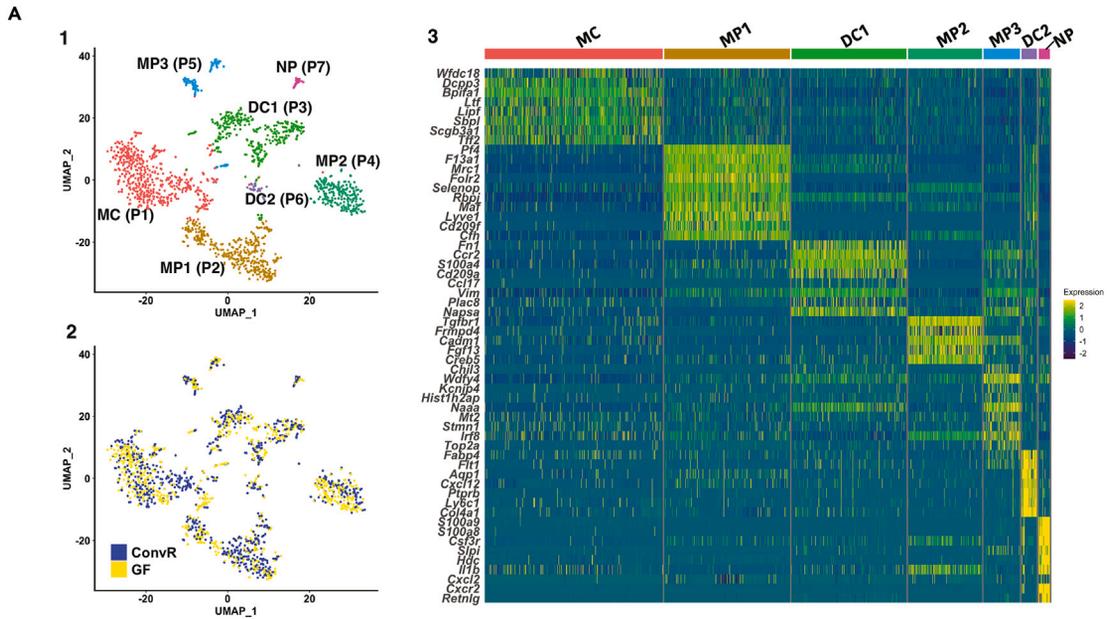
A

B

Figure 3. Phagocyte heterogeneity in the mouse larynx

(A) 1) & 2) UMAPs of the distribution of phagocyte subtypes in merged dataset (1) and across mouse groups (2); 3) Expression heatmap of phagocyte subtype markers based on top 5 differentially expressed (DE) genes (ranked by average $\log_2$ fold change) in merged dataset. MC, MP1-3, DC1-2, and NP represent monocyte, macrophage subtype 1–3, dendritic cell subtype 1–2, and neutrophil, respectively. P1-P7 represent phagocyte subtype 1–7.

(B) Violin plots of phagocyte canonical markers across subtypes in merged dataset. MC, DC, MP, and NP represent monocyte, dendritic cell, macrophage, and neutrophil, respectively.

intracellular pathogens.[51,52] The proportion of mucous (SEC3) cells was significantly elevated in GF (FDR = 0.1, Figure 4E). At a less stringent threshold (FDR = 0.3), serous cell (SEC4) proportion was significantly increased in GF mice. PAS staining and ImageJ quantification of mucous cells/mucus substances exhibited significantly increased mucous cell size and mucus production in submucosal glands of ConvR (p-value <0.05, Figure 4F), which correlated with findings in the small intestine, where intestinal goblet cells quickly increase in both size and number in GF mice upon colonization with commensal bacteria.[46] Cells in ConvR glands appeared plumper and more rounded, filled with increased mucus substances than that in GF (Figure 4F).

Phagocytes play a central role in defense against invading microorganisms, controlling whether protective or regulatory immune responses against pathogens or the commensal microbiota occur in host body.[53] A total of 95 genes were differentially expressed in phagocytes between ConvR and GF mice (Figure 5A). The largest GO biological process enrichment analysis revealed gene cluster was associated with immune system process, regulation of response to biotic stimulus, response to molecule of bacterial origin, antigen processing and presentation of exogenous peptide antigen via MHC class II, cell migration, and positive regulation of cell adhesion (Figure 5B). No significant GO biological process was found for genes upregulated in GF. We further performed differential expression analysis for monocytes (P1), combined macrophages (P2, P4, P5), and combined dendritic cells (P3, P6), considering cluster size of each subtype. No DE genes were found in monocytes or dendritic cells between ConvR and GF (adjusted p-value <0.05). A total of 185 DE genes (average $\log_2$ fold change >0.25, adjusted p-value <0.05) were upregulated in macrophages of ConvR (Figure 5C). These genes were associated with GO biological processes, including leukocyte migration, receptor-mediated endocytosis, response to interferon-γ, transcytosis, leukocyte chemotaxis, endosomal transport, monocyte chemotaxis, ERK1 and ERK2 cascade, and mononuclear cell migration (adjusted p-value <0.05) (Figure 5D; Table S7), which is likely due to microbial exposure in ConvR mice. A large gene cluster consisting of 122 genes was identified within the PPI built, with the genes upregulated in ConvR macrophages, where the top 20 hub genes, the most interactive gene nodes, identified in the cluster using the Maximal Clique Centrality (MCC) algorithm were *Egfr, Itgam, Igf1, Mmp9, Klf4, Kitl, Il15, Mrc1, Fcgr2b, Lyve1*. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis identified 29 significantly enriched pathways, where the top 5 pathways were *Rap1* signaling pathway, TNF signaling pathway, signaling pathway regulating pluripotency of stem cells, phagosome, and JAK-STAT signaling pathway (Figure 5E). By contrast, only 19 DE genes were upregulated in GF mice macrophages, and no GO terms were found for these genes. Among the 19 genes, however, it is worth noting that *Rgs1* has strong immune inhibitory effects,[54–56] which may have been upregulated in GF mice due to the lack of microbial antigenic stimulation to inhibit unwanted immune responses. *Tff2* is a negative regulator of gastrointestinal inflammation and immune cell cytokine responses, present in the gastrointestinal cells, macrophages, and lymphocytes.[57] High expression of *Tff2* in GF could be indicative of repressed macrophage responsiveness and blocked inflammatory cell recruitment to the larynx. At a less stringent threshold (FDR = 0.4), MP2 and MP3 populations were significantly elevated in GF (Figure 5F). Likewise, cell proportions of lymphocyte subtypes – T cells (L2) and B cells (L3) – were significantly increased in ConvR, while cell proportion of NK cells (L4) were increased in GF (FDR = 0.4). Given the number of lymphocyte cells recovered we did not perform differential expression analysis for this cell type. Microbial exposure seemed to have limited impact on BEC, SBEC, CEC, TC, CC, fibroblasts, and EC. Significant difference in cell proportion was observed only in BEC5, F5, EC1, and EC4 according to scCODA analysis (FDR = 0.4, Figure S5B). Several genes were differentially expressed between ConvR and GF for these cell types and subtypes (Tables S8 and S9).

Transcriptional map of pattern recognition receptors and adaptors in the larynx

To determine the capacity of each cell type to sense and respond to microbes in the larynx, we investigated the expression profile of a group of pattern recognition receptors (PRRs), including toll-like receptors (TLRs) 1–13, nucleotide oligomerization domain (NOD)-like receptors (NLRs) 1&2, and their adaptors *Nlrc4&6, Myd88, Ticam1, Ripk2,* and *Aim2*.[58] Expression of these genes, except *Tlr10* and *Nlrc6*, were detected in both ConvR and GF datasets. Across all major cell types, the PRRs, especially *Tlr5, 6, 7, 8, 9, 11, 12, 13*, were expressed at low levels in low proportion (<40%) of cells in each cell type, regardless of the microbial exposure status (Figure 6A1). While not differentially expressed in the presence of microbiota in ConvR mice, most PRRs and their adapters tended to have greater number of cells expressing these genes and higher average expressions in laryngeal cells of ConvR than GF, regardless of cell type. A similar observation was previously reported in epithelial cells of the small intestine in specific pathogen free (SPF) and GF mice.[49] These PRRs and adapters were primarily expressed in phagocytes, lymphocytes, and fibroblasts (Figure 6A1). Phagocytes exhibited higher expression of the PRRs, particularly TLRs, and their adapters, compared with other cell types, further confirming our previous observation that phagocyte is one of the most responsive cell types to bacterial stimulation in the larynx. Among phagocytes, macrophages were the major population expressing TLRs, with each macrophage subset carrying different types of TLRs (Figure 6A2). Across epithelial cell types, SBEC had the highest proportion of cells expressing *Tlr3, Tlr4, Nod1&2*, and adapters *Myd88, Ripk2*, more than likely due to these cells located at the apical side of the lumen constantly exposed to microbial stimulants. These genes were expressed in less abundance, but in greater proportion of cells in SBEC relative to phagocytes and all other cell types. *Reg3g* has previously shown to be induced in Paneth cells by bacteria in a *Myd88/Tlr*-dependent manner.[59,60] Given

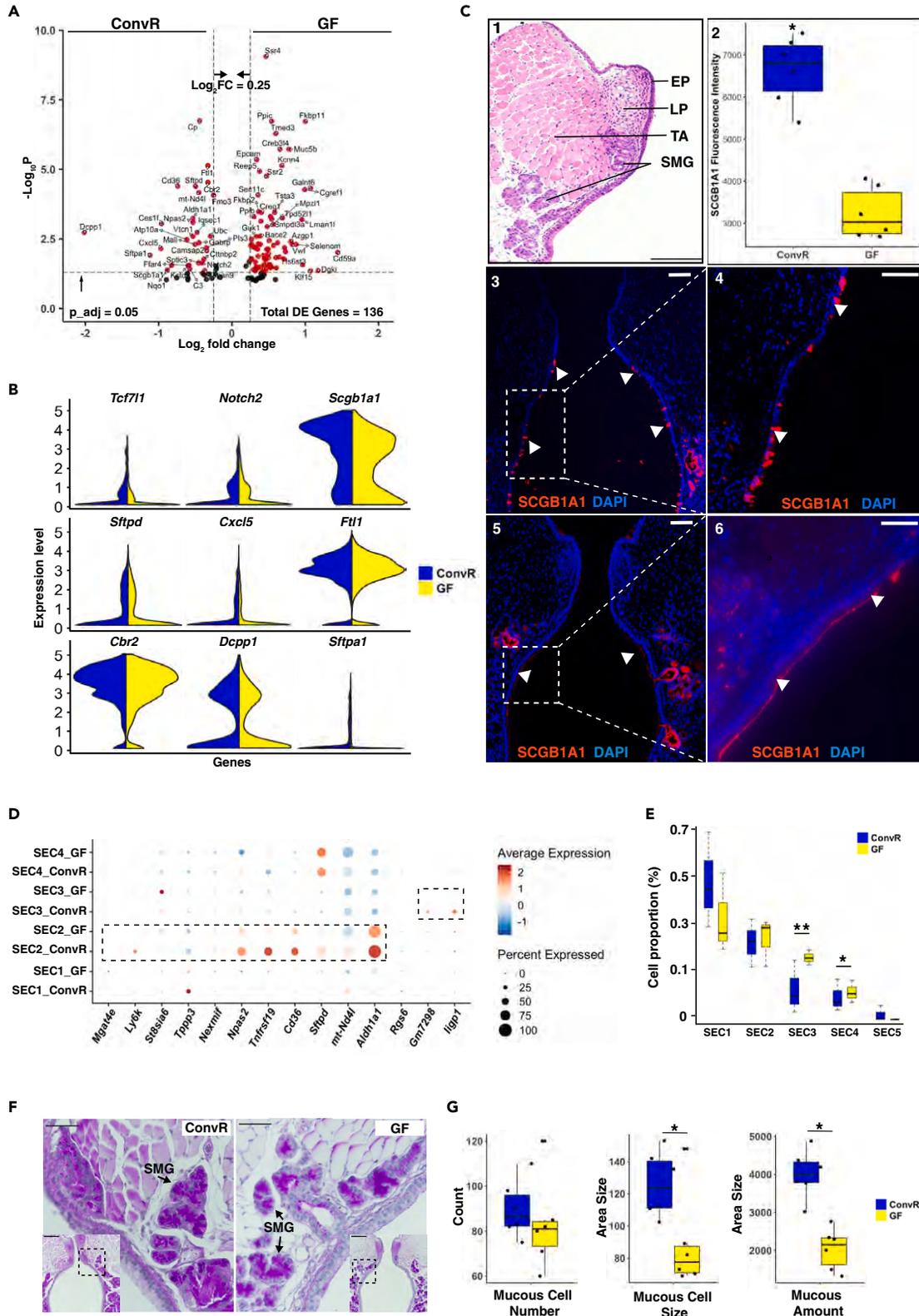**Figure 4. Secretory Epithelial Cell (SEC) differential gene expression in ConvR and GF mice larynges**

(A) Volcano plot of differentially expressed (DE) genes in SEC of ConvR and GF mouse larynges (average $\log_2$ fold change >0.25, adjusted *p*-value <0.05).

(B) Expression of host defense associated DE genes in SEC of ConvR versus GF larynges (average $\log_2$ fold change >0.25, adjusted *p*-value <0.05).

Figure 4. *Continued*

(C) Antibody staining of SCGB1A1 in paraffin-embedded vocal fold coronal sections for ConvR and GF mice. 1) Hematoxylin and Eosin stained vocal fold coronal sections demonstrating morphology of vocal fold (VF) mucosa; EP, LP, TA, and SMG represent epithelium, lamina propria, thyroarytenoid muscle, and submucosal glands, respectively; 2) Quantification of SCGB1A1 fluorescence intensity in immunostained images; Asterisk indicates that SCGB1A1 fluorescence intensity was significantly higher in ConvR based on a two tailed t-test with equal variance (p-value <0.05); the horizontal black line in the center of the box represents the median value, the top and bottom whiskers represent maximum and minimum value, the upper and lower lines of the box represent the upper and lower quartile values; confidence interval 95%; 3) & 4) Antibody staining of SCGB1A1 in ConvR (3) and GF (4) vocal folds; 5) & 6) Magnification of the regions in white dotted rectangles in panels 3) and 4). SCGB1A1 protein and DAPI (nucleus) were stained in red and blue, respectively. Scale bar in 1 = 250 μm; scale bars in panels 3 and 4 = 500 μm; scale bars in panels 5 and 6 = 100 μm; white arrows indicate the location of cells with high SCGB1A1 protein expression; white dotted rectangles indicate the magnified regions in 5) and 6).

(D) Dotplot of DE genes between ConvR and GF larynges for SEC subtypes. SEC1-SEC4 represent SEC subtypes 1–4. Genes upregulated in each subtype of ConvR mice were framed in dotted rectangles. Color scale - red, high expression; blue, low expression.

(E) Proportions of SEC subtypes (SEC1-SEC5) in ConvR and GF larynges. Asterisk indicates significant difference in cell proportion between ConvR and GF according to scCODA analysis at FDR = 0.1 (**) and FDR = 0.3 (*). The horizontal black line in the center of the box represents the median value, the top and bottom whiskers represent maximum and minimum value, the upper and lower lines of the box represent the upper and lower quartile values; confidence interval 95%.

(F) Periodic acid – Schiff staining of mucus and mucus producing cells in paraffin-embedded sections of ConvR (left) and GF (right) mouse larynx (magnification = 40x); scale bar = 500 μm. Black dotted rectangles indicate zoomed-in regions with higher magnification; magnification in zoomed-out images = 20x, scale bar = 500 μm. Black arrows denote submucosal glands (SMG).

(G) Quantification of mucous cell number, size, and mucus amount in submucosal glands; asterisk indicates significant difference between mouse groups according to a two-tailed Student's t test with equal variance (p-value <0.05). The horizontal black line in the center of the box represents the median value, the top and bottom whiskers represent maximum and minimum value, the upper and lower lines of the box represent the upper and lower quartile values; confidence interval 95%.

the expression of multiple antimicrobial genes in SBEC and the localization of these cells in the VF mucosa (Figures 2A and 2B), SBEC are likely the first line of defense against microbes entering body by sensing microbial antigens and secreting TLR-dependent antimicrobial substances and other inflammatory mediators for local and systemic control of infection. The proportion of cells in SBEC expressing *Tlr3, Tlr4, Nod1, Myd88,* and *Ripk2* was relatively higher in ConvR than GF (Figure 6A).

Gene regulatory network analysis reveals altered metabolic gene expression in response to microbial colonization in the larynx

The transcriptional state of a cell is governed by a gene regulatory network (GRN) where transcription factors (TFs) and cofactors work together to regulate the expression of target genes. To better understand the GRN involved in the development of laryngeal immunity in response to microbial exposure, we used *SCENIC* (single cell regulatory network inference and clustering) to infer the activity of regulons – TFs and their target genes - that could regulate the differentiation and function of the major cell types in the larynx. Overall, we found 89 and 111 regulons for each mouse group that govern transcriptional profiles of these cell populations. The top 5 regulons in the major cell types of GF mice were notably different from those in ConvR due to the absence of commensal microbiota (Figure 6B). We further identified differential regulons across mouse groups based on the AUC scores for each major cell type. A varying number (18–48) of differentially expressed regulons were detected across the major cell types, with more regulons highly active in ConvR than GF, in all cell types (Figure 6C). Interestingly, differential regulons with greater activity in GF mice were dominated by genes from *Fos* and *Jun* families, which are previously identified as nutrient starvation stress response TFs.[61] These TFs appeared in multiple cell types of GF mice, including fibroblasts, SEC, EC, BEC, and SBEC. Further evaluation of the expression level of additional TFs involved in nutrient starvation stress response found *Fos, Fosl2, Atf3, Atf4, Jun, Junb,* and *Jund* exhibited increased expression in the GF across majority of cell types (Figure S5C). Although ConvR and GF were fed the same diets, it is likely that they were exposed to different nutrients due to the lack of microbially derived metabolites in GF mice, since commensal microbiota mediate host metabolism, providing nutrients such as vitamin K and short chain fatty acids that cannot be produced by the host body. As mTOR signaling serves as a central regulator of cell metabolism, growth, proliferation, survival and is repressed by nutrient starvation stress, we further assessed the expression of mTORC1components (*Mtor* and *Rptor*) that mainly regulates cell growth and metabolism in mammals. We observed no significant difference in the expression of *Mtor* or *Rptor* between ConvR and GF. Interestingly, mTORC1 target gene ribosomal protein S6 (*Rps6*) was not significantly expressed between the mouse groups, while its kinase *Rps6kb1* trended for greater expression in fibroblasts, phagocytes, SBEC, and lymphocytes of ConvR (Figure 6A). This suggests that the GF larynx represses *Rps6kb1* expression and mTORC1 activity, resulting in reduced S6 phosphorylation within the GF larynx due to nutrient starvation stress. Meanwhile, the expression of *Deptor*, a well-known negative regulator of mTOR, demonstrated increased expression in ConvR phagocytes and SBEC, indicating the potential negative regulation of *Deptor* present in ConvR. Further investigation is needed for corroboration.

DISCUSSION

We investigated the cellular landscape of the larynx and impact of microbial exposure on the development of laryngeal immunity, in a murine model. We established a reference single-cell atlas of the normal larynx and dissect the role of microbiota in laryngeal cell development and functions at single-cell resolution. Our data suggests commensal microbiota have an extensive impact on the laryngeal immune system, manifested by the regulation of cell differentiation, elevated expression of host defense genes, PRRs, and metabolic genes. In addition,
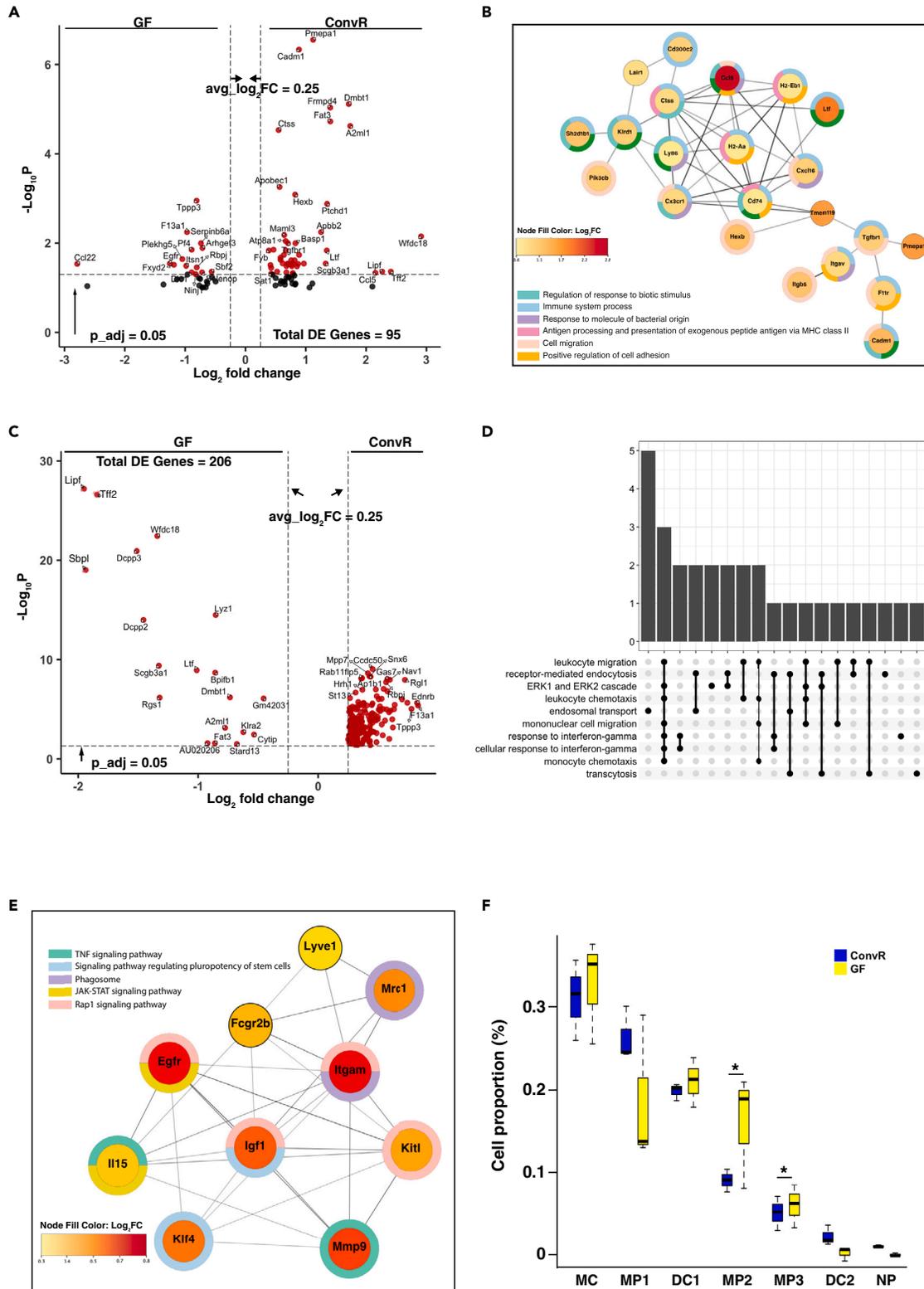
Figure 5. Functional analysis of differentially expressed (DE) genes in macrophages of ConvR mice larynges

(A) Volcano plots of DE genes in phagocytes between ConvR and GF mouse groups (average log$_2$ fold change >0.25, adjusted *p*-value <0.05).

(B) Largest protein-protein interaction (PPI) network found among the DE genes of phagocytes significantly upregulated in ConvR larynges. GO biological process-based enrichment analysis of the StringDB PPI revealed 6 significantly enriched biological processes shown as split donut of different colors

Figure 5. Continued

enclosing nodes. Each node represents a gene. Black line linking two nodes represents the known or predicted interaction between the two genes. Color of a node indicate average log$_2$ fold change of the gene.

(C) Volcano plots of DE genes in combined macrophage subsets (P2,4,5) between ConvR and GF mice groups (average log$_2$ fold change >0.25, adjusted *p*-value <0.05).

(D) Upset plot of top 10 enriched terms for macrophages in ConvR mice based on their top 50 DE genes (average log$_2$ fold change >0.25, adjusted *p* < 0.05). Gray bars indicate the number of genes associated with GO terms; black connected dots indicate the associated GO terms.

(E) Top 10 PPI hub genes identified by the Maximal Clique Centrality algorithm using cytoHubba plugin of Cytoscape software, demonstrating the most interactive 10 genes found within the DE genes upregulated in macrophages of ConvR larynges. KEGG pathway-based enrichment analysis of the StringDB PPI revealed 5 innate inflammation associated pathways as indicated by split donut of different colors enclosing nodes. Each node represents a gene. Black line linking two nodes represents the known or predicted interaction between the two genes. Color of a node indicate average log$_2$ fold change of the gene.

(F) Comparison of macrophage subtype proportions between ConvR and GF mice; asterisk indicate significant difference in cell proportion between mouse groups based on scCODA analysis (FDR = 0.4). The horizontal black line in the center of the box represents the median value, the top and bottom whiskers represent maximum and minimum value, the upper and lower lines of the box represent the upper and lower quartile values; confidence interval 95%.

our investigation uncovers the presence of hillock club cells, club cells, serous cells, tuft cells, innate lymphoid cells, and multiple subpopulations of macrophages in the larynx, as well as several potential transcriptional markers for multiple cell types, that have not previously been reported, in the larynx.

The larynx has a unique ecological niche with selective pressure against colonizing bacteria. While the role of commensal microbiota in regulating host immune system has been well documented in other organs,[4] there is a paucity of research in the larynx due to the unavailability of healthy human tissue, challenge of sampling, and the low microbial biomass nature of samples. Laryngeal epithelium, as an initial site of microbial exposure, was influenced by commensal microbiota according to our study, and secretory cells were most responsive among all epithelial cell types toward microbial exposure (Figure 6A). The leading immune mechanism that governs asymptomatic epithelial colonization in the gut is the secretion of antimicrobial peptides (AMPs) by secretory epithelial cells.[32] The larynx shares many AMPs with the gut, including but not limited to *Pglyrp1, Pigr, Reg3g, Bpifa1,* and so forth.[62] (Figure 2). We measured an elevated expression of surfactant proteins A1 and D (*Sftpa1, Sftpd*) in SEC of ConvR mice compared to GF mice (Figures 4A and 4B). Both proteins are pattern recognition molecules of the collection family of C-type lectins, maintaining host homeostasis through their dual roles as anti-infectious and immunomodulatory agents.[63] Given the considerable difference in microbiota diversity and abundance between the two systems, selection of AMPs for host defense reinforcement appears to be largely dependent on microbial composition. Laryngeal microbiota also induced expressions of other host defense genes associated with epithelial barrier integrity, cell differentiation, anti-inflammation (Figure 4B). Expressions of multiple *SCGB* family members, dominated by *Scgb1a1*, a pulmonary surfactant protein with anti-inflammatory function in SEC, were notably increased in SECs of ConvR larynges (Figure 4C). Our data also suggests club cells as one of the secretory epithelial cell populations, enriched for a quantity of genes in ConvR associated with the regulation of DCs (*Lgals3, Tmem176b, Tmem176a, Hmgb1*), T cells (*Lgals3, Cebpb, Anxa1, Sdc4, Sftpd*), keratinocytes (*Notch2, Reg3g, Zfp36L1*), and fibroblasts (*Gstp1, Fth1, S100a6, Ctnnb1*) proliferation and differentiation, implying while facilitating immune response, epithelial cells interact with innate and adaptive immune system cells to maintain laryngeal homeostasis.

Ciliated cells play a pivotal role in airway homeostasis by trapping and expelling microorganisms, mucus, and other debris through mucociliary clearance (MCC). It is noteworthy that a heat shock protein *Hsp90aa1* and antiviral protein *Ifitm1* were expressed broadly across all epithelial cell types, with significantly elevated expressions in CEC (Figure S2E). *Hsp90aa1* could bind bacterial lipopolysaccharide (LPS) and mediate LPS-induced inflammatory response and *Ifitm1* restricts cellular entry of diverse viral pathogens such as influenza A virus, Ebola, and *Sars-Cov-2*, implying that trapping and expelling mucus, microorganisms, and other debris through mucociliary clearance may not be the only way that CEC maintains airway homeostasis{Citation}. The moderate expression of *Reg3g, Bpifa1, Lyz2, Pglyrp1, Ltf,* and *Muc16* in CEC, also demonstrates the potential involvement of CEC in antimicrobial activity through the secretion of these substances (Figure 2B). Direct evidence showing specific microorganisms or microbiota, as a whole, contributing to the increased expression of these genes is still lacking, suggesting an urgent demand for further investigation to corroborate.

Microbes regulate cell differentiation in the larynx. Epithelial cell subsets identified in human lungs suggest that subtle variation in expression profiles may result in distinct functional or phenotypical change.[28] Here, we observed increased number of serous cells in GF (FDR = 0.3) and increased size of mucous cells in ConvR in submucosal glands, suggesting potential impact of commensal microbiota on SEC differentiation and subtype allocation during long-term host-microbe interactions (Figures 4E and 4F). We also observed a secretory epithelial subtype SEC5 present only in ConvR mice, and the expression pattern of proliferating cells (SEC1) varied with mouse groups based on the cell distribution distinctly divided in the UMAP (Figure 2C3), which implies the potential influence of commensal microbiota on laryngeal cell differentiation. However, the expression of mucus related genes was not significantly altered by colonization status in mucous cells (Table S9), implying the possibility of potential post-transcriptional regulation in the larynx that promotes mucus production.

Long-term adaptation to commensal microbiota alters transcriptional and differentiation status of professional immune cells in healthy states. Our data suggests macrophages are the most responsive immune cells in the larynx to bacterial exposure. Macrophage subsets MP2 and MP3 were significantly increased in GF mice (FDR = 0.3), which is suggestive of the potential role of commensal microbiota in regulating immune cell differentiation. Macrophages in the laryngeal tissues are likely from the circulating monocyte pool, given the high expression of monocyte signatures genes, exhibiting consistency with the findings in human lungs.[64] Classical monocytes in the blood are thought to
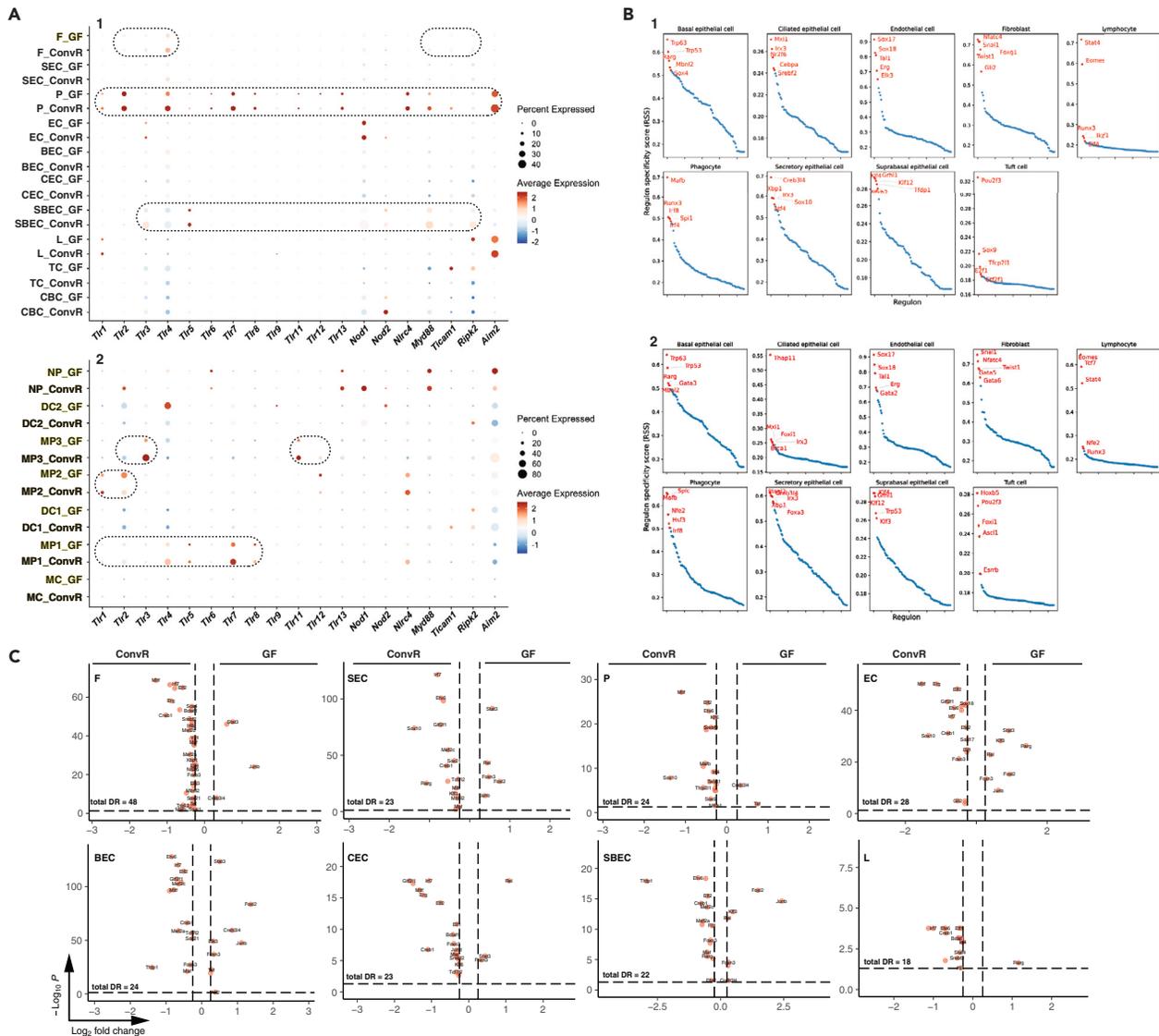
Figure 6. Expression of pattern recognition receptors (PRRs) and their adapters and regulon activity in ConvR and GF mice larynges

(A) Dotplot of average expression of PRRs and their adapters across major cell types of ConvR and GF mice larynges. Dot size represent the percent of cells expressed in a cell type; Color scale: red, high expression; blue, low expression. Genes that have higher average expression in ConvR were framed in dotted rounded rectangles.

(B) Dotplot of top 5 regulon in major laryngeal cell types of ConvR (top) and GF (bottom) ranked by regulon specificity score (RSS).

(C) Volcano plot of significant differentially expressed regulons in major laryngeal cell types of ConvR and GF mice (average log$_2$ fold change >0.25, adjusted *p*-value <0.05). F, SEC, P, EC, BEC, CEC, SBEC, and L represent fibroblast, secretory epithelial cell, phagocyte, endothelial cell, basal epithelial cell, ciliated epithelial cell, suprabasal epithelial cell, and lymphocyte, respectively.

be MHCII⁻.[65] Given the expression of *Adgre1*^{lo}, *Cd14*$^+$, *Spn/Cd43*^{lo} (an optimal marker than *Ly6c* to divide monocytes in all species to date), MHCII$^+$, P1 cells were likely to be the one of the five blood monocyte subsets reported by Jakubzick et al., 2015.[65] While these cells only represent 1% of the total blood monocytes, they extravasate into both mildly inflamed (ConvR) and noninflamed (GF) tissues continuously and form the largest phagocyte population in the larynx (Figures 3A1, 3A2, and 3B). Expression level of *Cx3cr1* has been linked to monocyte maturation, where classical circulatory monocytes express low levels of *Cx3cr1* and non-classical monocytes have high *Cx3cr1* expression,[66] which further supports the P1 cluster as extravasated monocytes from the blood. After populating the larynx, the expression of the blood monocytes markers (*Hp*$^+$, *F10*$^+$, *Hdc*$^+$, *Ly6c2*$^+$, *Sell/Cd62L*$^+$) were downregulated to almost zero in the monocyte[67] (Figure 3B). Monocytes in both conditions appeared to be patrolling monocytes based on the low expression of *Ly6c1/2*, which aligns with the findings that the mobilization of monocyte into tissues (skin, skin-draining lymph nodes, mesenteric lymph nodes) is not affected by the presence or absence of microbial colonization.[65] *Adgre*/F4/80-expressing macrophages in P2 and P4 subsets were *Cd163*$^+$, *Mrc1/Cd206*$^+$ and Cd86$^+$, *Cd80*$^+$ respectively,

suggesting that they are more than likely M2-macrophages and M1-macrophages (Figure 3B). Subset P5, is likely another lineage of macrophages, a subpopulation recently reported by two other single cell studies[68,69] (Figures 3A1, 3A2, and 3B). Both dendritic cell populations were likely mature dendritic cells given the expression of *Itgax/Cd11c, Cd80, Cd86, MHCII, Cd24a*, and *Flt3*. However, each displayed unique expression of multiple subtype markers, including *Cd4+, Itgam/Cd11c+, Mrc1/Cd206+* in P3 and *Ly75/Cd205+, Itgae/Cd103+, Zbtb46+, Kit+* in P6 (Figure 3B). Each phagocyte subset was enriched for a class of specialized biological functions associated with the immune response, despite overlap between subsets, possibly fulfilling different roles, in the larynx (Table S3). Limited discussion was given to P5-P7 subsets given their small cluster sizes.

Further, macrophages express elevated Tlrs for the detection and removal of pathogens through phagocytosis (Figure 6A2). Enrichment of non-inflammatory signaling pathways, such as extracellular signal related kinase (ERK) 1 and ERK2, in ConvR mice is a strong indication of the interaction between immune cells and commensals and their role in maintaining laryngeal homeostasis[70] (Figure 5D). In addition, *Egfr, Itgam, Igf1, Mmp9, Klf4, Kitl, Il15, Mrc1, Fcgr2b, Lyve1* genes in macrophages appear to be playing an important role in the interaction with commensal microbiota in the larynx, modulating *Rap1* signaling pathway, TNF signaling pathway, signaling pathway regulating pluripotency of stem cells, phagosome, and JAK-STAT signaling pathway (Figure 5E). Commensal microbiota in the mouse colon drives the functional diversification of macrophages, with a proportional decrease of *Cd11c+, Cd206int, Cd121b+* and *Cd11c−, Cd206hi, Cd121b−* subsets in GF.[67] Conversely, macrophage subsets (MP2 and MP3) in the larynx appear to have decreased number of cells in ConvR than GF as per the scCODA analysis at a less stringent threshold (FDR = 0.4). These evidence suggest a potential larynx-specific host-microbe interaction with unknown mechanisms that need further investigation. Interestingly, *Aim2* was primarily expressed in phagocytes and lymphocytes. It regulates caspase-1-dependent processing of pro-IL-1β and pro-IL-18 in response to double-stranded DNA in both dendritic cells and macrophages, playing important role in sensing both bacterial and viral pathogens and in triggering innate immunity.[71] Taken together, these findings suggest that macrophages and SBEC are the two major cell types associated with commensal microbiota sensing and balancing and that the expression of PRRs in the larynx is not significantly altered by microbial colonization.

Limitations of the study

In this article, we demonstrate the cell atlas of the larynx and dissect the role of commensal microbiota in cell differentiation and function at single-cell resolution. However, this study could have benefited from the inclusion of both male and female mice to improve rigor and reproducibility. While sex difference in immune responses change throughout life and are influenced by age, it can occur in both innate and adaptive immune responses and are evolutionarily conserved across diverse species. Inclusion of female mice also increases sample size and therefore the power of statistical analysis. Despite these limitations, we conclude that laryngeal microbiota play a pivotal role in maintaining laryngeal innate and adaptive immunity through regulating cell differentiation, increasing the expression of genes associated with host defense, and altering gene regulatory networks.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - 16S rRNA gene sequencing
  - Dissociation of laryngeal tissues
  - Droplet-based scRNA-seq
  - Immunofluorescence (IF) staining
  - Periodic acid – Schiff (PAS) staining
  - Preprocessing of scRNA-seq data
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Dimensionality reduction using PCA and UMAP
  - Merged analysis of the scRNA-seq data for ConvR and GF
  - Clustering and subclustering
  - Differential expression and enrichment analysis
  - Cell type annotation
  - Protein-protein interaction (PPI) network analysis
  - Gene regulatory network analysis
  - Comparison of cell proportions
  - Comparison of cell features

AUTHOR CONTRIBUTIONS

RA, SLT, FER, and CK designed the research; RA performed all experiments; RA, ZN, and EX analyzed scRNA-seq data; RA wrote the original draft and all authors reviewed, edited, and approved the final draft.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Levendoski, E.E., Leydon, C., and Thibeault, S.L. (2014). Vocal Fold Epithelial Barrier in Health and Injury: A Research Review. J. Speech Lang. Hear. Res. 57, 1679–1691. https://doi.org/10.1044/2014_JSLHR-S-13-0283.

2. Bhattacharyya, N. (2014). The prevalence of voice problems among adults in the United States. Laryngoscope 124, 2359–2362. https://doi.org/10.1002/LARY.24740.

3. Cohen, S.M., Kim, J., Roy, N., Asche, C., and Courey, M. (2012). Direct health care costs of laryngeal diseases and disorders. Laryngoscope 122, 1582–1588. https://doi.org/10.1002/LARY.23189.

4. Belkaid, Y., and Hand, T.W. (2014). Role of the microbiota in immunity and inflammation. Cell 157, 121–141. https://doi.org/10.1016/j.cell.2014.03.011.

5. Thibeault, S.L., Rees, L., Pazmany, L., and Birchall, M.A. (2009). At the crossroads: Mucosal immunology of the larynx. Mucosal Immunol. 2, 122–128. https://doi.org/10.1038/mi.2008.82.

6. An, R., Robbins, D., Rey, F.E., and Thibeault, S.L. (2022). Vocal fold mucus layer: Comparison of histological protocols for visualization in mice. Laryngoscope Investig. Otolaryngol. 7, 444–453. https://doi.org/10.1002/LIO2.743.

7. Rees, L.E.N., Gunasekaran, S., Sipaul, F., Birchall, M.A., and Bailey, M. (2006). The isolation and characterisation of primary human laryngeal epithelial cells. Mol. Immunol. 43, 725–730. https://doi.org/10.1016/J.MOLIMM.2005.03.011.

8. Barker, E., Haverson, K., Stokes, C.R., Birchall, M., and Bailey, M. (2006). The larynx as an immunological organ: Immunological architecture in the pig as a large animal model. Clin. Exp. Immunol. 143, 6–14. https://doi.org/10.1111/j.1365-2249.2005.02950.x.

9. Hobbs, C.G.L., Rees, L.E.N., Heyderman, R.S., Birchall, M.A., and Bailey, M. (2006). Major histocompatibility complex class I expression in human tonsillar and laryngeal epithelium. Clin. Exp. Immunol. 145, 365–371. https://doi.org/10.1111/j.1365-2249.2006.03125.x.

10. Jetté, M.E., Seroogy, C.M., and Thibeault, S.L. (2017). Laryngeal T regulatory cells in the setting of smoking and reflux. Laryngoscope 127, 882–887. https://doi.org/10.1002/lary.26223.

11. Eberl, G. (2010). A new vision of immunity: homeostasis of the superorganism. Mucosal Immunol. 3, 450–460. https://doi.org/10.1038/mi.2010.20.

12. Sommariva, M., Le Noci, V., Bianchi, F., Camelliti, S., Balsari, A., Tagliabue, E., and Sfondrini, L. (2020). The lung microbiota: role in maintaining pulmonary immune homeostasis and its implications in cancer development and therapy. Cell. Mol. Life Sci. 77, 2739–2749. https://doi.org/10.1007/s00018-020-03452-8.

13. Hooper, L.V., Littman, D.R., and Macpherson, A.J. (2012). Interactions between the microbiota and the immune system. Science 336, 1268–1273. https://doi.org/10.1126/science.1223490.

14. Gong, H., Shi, Y., Xiao, X., Cao, P., Wu, C., Tao, L., Hou, D., Wang, Y., and Zhou, L. (2017). Alterations of microbiota structure in the larynx relevant to laryngeal carcinoma. Sci. Rep. 7, 5507–5509. https://doi.org/10.1038/s41598-017-05576-7.

15. Hanshew, A.S., Jetté, M.E., and Thibeault, S.L. (2014). Characterization and comparison of bacterial communities in benign vocal fold lesions. Microbiome 2, 1–9. https://doi.org/10.1186/2049-2618-2-43.

16. An, R., Gowda, M., Rey, F.E., and Thibeault, S.L. (2020). Selective Bacterial Colonization of the Murine Larynx in a Gnotobiotic Model. Front. Microbiol. 11, 594617. https://doi.org/10.3389/fmicb.2020.594617.

17. Walker, M.J., Barnett, T.C., McArthur, J.D., Cole, J.N., Gillen, C.M., Henningham, A., Sriprakash, K.S., Sanderson-Smith, M.L., and Nizet, V. (2014). Disease manifestations and pathogenic mechanisms of Group A Streptococcus. Clin. Microbiol. Rev. 27, 264–301. https://doi.org/10.1128/CMR.00101-13.

18. Jetté, M.E., Dill-Mcfarland, K.A., Hanshew, A.S., Suen, G., and Thibeault, S.L. (2016). The human laryngeal microbiome: Effects of cigarette smoke and reflux. Sci. Rep. 6, 6–16. https://doi.org/10.1038/srep35882.

19. Glassing, A., Dowd, S.E., Galandiuk, S., Davis, B., and Chiodini, R.J. (2016). Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. Gut Pathog. 8, 24. https://doi.org/10.1186/s13099-016-0103-7.

20. Yu, Y., Lu, L., Sun, J., Petrof, E.O., and Claud, E.C. (2016). Preterm infant gut microbiota affects intestinal epithelial development in a humanized microbiome gnotobiotic mouse model. Am. J. Physiol. Gastrointest. Liver Physiol. 311, G521–G532. https://doi.org/10.1152/AJPGI.00022.2016.

21. Schoenborn, A.A., von Furstenberg, R.J., Valsaraj, S., Hussain, F.S., Stein, M., Shanahan, M.T., Henning, S.J., and Gulati, A.S. (2019). The enteric microbiota regulates jejunal Paneth cell number and function without impacting intestinal stem cells. Gut Microb. 10, 45–58. https://doi.org/10.1080/19490976.2018.1474321.

22. Jovic, D., Liang, X., Zeng, H., Lin, L., Xu, F., and Luo, Y. (2022). Single-cell RNA sequencing technologies and applications: A brief overview. Clin. Transl. Med. 12, e694. https://doi.org/10.1002/ctm2.694.

23. Maynard, R.L., and Downes, N. (2019). Anatomy and Histology of the Laboratory Rat in Toxicology and Biomedical Research (Elsevier). https://doi.org/10.1016/C2016-0-02030-2.

24. Reid, A.T., Sutanto, E.N., Chander-Veerati, P., Looi, K., Li, N.F., Iosifidis, T., Loo, S.L., Garratt, L.W., and Kicic, A. (2019). Ground zero—the

airway epithelium. In Rhinovirus Infections: Rethinking the Impact on Human Health and Disease, pp. 61–98. https://doi.org/10.1016/B978-0-12-816417-4.00003-2.

25. Ruysseveldt, E., Martens, K., and Steelant, B. (2021). Airway Basal Cells, Protectors of Epithelial Walls in Health and Respiratory Diseases. Front. Allergy 2, 787128. https://doi.org/10.3389/FALGY.2021.787128.

26. Mori, M., Mahoney, J.E., Stupnikov, M.R., Paez-Cortez, J.R., Szymaniak, A.D., Varelas, X., Herrick, D.B., Schwob, J., Zhang, H., and Cardoso, W.V. (2015). Notch3-Jagged signaling controls the pool of undifferentiated airway progenitors. Development 142, 258–267. https://doi.org/10.1242/DEV.116855.

27. Carraro, G., Langerman, J., Sabri, S., Lorenzana, Z., Purkayastha, A., Zhang, G., Konda, B., Aros, C.J., Calvert, B.A., Szymaniak, A., et al. (2021). Transcriptional analysis of cystic fibrosis airways at single-cell resolution reveals altered epithelial cell states and composition. Nat. Med. 27, 806–814. https://doi.org/10.1038/S41591-021-01332-7.

28. Vieira Braga, F.A., Kar, G., Berg, M., Carpaij, O.A., Polanski, K., Simon, L.M., Brouwer, S., Gomes, T., Hesse, L., Jiang, J., et al. (2019). A cellular census of human lungs identifies novel cell states in health and in asthma. Nat. Med. 25, 1153–1163. https://doi.org/10.1038/S41591-019-0468-5.

29. Joshi, S., Kumar, S., Bafna, S., Rachagani, S., Wagner, K.U., Jain, M., and Batra, S.K. (2015). Genetically engineered mucin mouse models for inflammation and cancer. Cancer Metastasis Rev. 34, 593–609. https://doi.org/10.1007/s10555-015-9549-1.

30. Xu, M., Yang, W., Wang, X., and Nayak, D.K. (2020). Lung Secretoglobin Scgb1a1 Influences Alveolar Macrophage-Mediated Inflammation and Immunity. Front. Immunol. 11, 2471. https://doi.org/10.3389/FIMMU.2020.584310.

31. Reynolds, S.D., Reynolds, P.R., Pryhuber, G.S., Finder, J.D., and Stripp, B.R. (2002). Secretoglobins SCGB3A1 and SCGB3A2 define secretory cell subsets in mouse and human airways. Am. J. Respir. Crit. Care Med. 166, 1498–1509. https://doi.org/10.1164/rccm.200204-285OC.

32. Lokken-toyli, K.L., Piters, W.A.A.S., Zangari, T., Martel, R., Kuipers, K., Shopsin, B., Loomis, C., Bogaert, D., and Weiser, J.N. (2021). Decreased production of epithelial-derived antimicrobial molecules at mucosal barriers during early life. Mucosal Immunol. 14, 1358–1368. https://doi.org/10.1038/s41385-021-00438-y.

33. Bakshani, C.R., Morales-Garcia, A.L., Althaus, M., Wilcox, M.D., Pearson, J.P., Bythell, J.C., and Burgess, J.G. (2018). Evolutionary conservation of the antimicrobial function of mucus: A first defence against infection. NPJ Biofilms Microbiomes 4, 14. https://doi.org/10.1038/s41522-018-0057-2.

34. Osanai, A., Sashinami, H., Asano, K., Li, S.J., Hu, D.L., and Nakane, A. (2011). Mouse peptidoglycan recognition protein PGLYRP-1 plays a role in the host innate immune response against Listeria monocytogenes infection. Infect. Immun. 79, 858–866. https://doi.org/10.1128/IAI.00466-10.

35. Johansen, F., and Kaetzel, C.S. (2011). Regulation of the Polymeric Immunoglobulin Receptor and IgA Transport: New Advances in Environmental Factors that Stimulate pIgR Expression and its Role in Mucosal Immunity.

Mucosal Immunol. 4, 598–602. https://doi.org/10.1038/mi.2011.37.

36. Tsou, Y.A., Tung, M.C., Alexander, K.A., Chang, W.D., Tsai, M.H., Chen, H.L., and Chen, C.M. (2018). The Role of BPIFA1 in Upper Airway Microbial Infections and Correlated Diseases. BioMed Res. Int. 2018, 2021890. https://doi.org/10.1155/2018/2021890.

37. Haber, A.L., Biton, M., Rogel, N., Herbst, R.H., Shekhar, K., Smillie, C., Burgin, G., Delorey, T.M., Howitt, M.R., Katz, Y., et al. (2017). A single-cell survey of the small intestinal epithelium. Nature 551, 333–339. https://doi.org/10.1038/nature24489.

38. Travaglini, K.J., Nabhan, A.N., Penland, L., Sinha, R., Gillich, A., Sit, R.V., Chang, S., Conley, S.D., Mori, Y., Seita, J., et al. (2020). A molecular cell atlas of the human lung from single-cell RNA sequencing. Nature 587, 619–625. https://doi.org/10.1038/s41586-020-2922-4.

39. Samarrai, R., Frank, S., Lum, A., Woodis, K., Weinstock, G., and Roberts, D. (2022). Defining the microbiome of the head and neck: A contemporary review. Am. J. Otolaryngol. 43, 103224. https://doi.org/10.1016/j.amjoto.2021.103224.

40. Watt, A.P., Sharp, J.A., Lefevre, C., and Nicholas, K.R. (2012). WFDC2 is differentially expressed in the mammary gland of the tammar wallaby and provides immune protection to the mammary gland and the developing pouch young. Dev. Comp. Immunol. 36, 584–590. https://doi.org/10.1016/J.DCI.2011.10.001.

41. Hewitt, R.J., and Lloyd, C.M. (2021). Regulation of immune responses by the airway epithelial cell landscape. Nat. Rev. Immunol. 21, 347–362. https://doi.org/10.1038/s41577-020-00477-9.

42. Nikolaidis, N.M., Wang, T.C., Hogan, S.P., and Rothenberg, M.E. (2009). Allergen induced TFF2 is expressed by mucus-producing airway epithelial cells but is not a major regulator of inflammatory responses in the murine lung. Exp. Lung Res. 32, 483–497. https://doi.org/10.1080/01902140601059547.

43. Livraghi-Butrico, A., Grubb, B.R., Wilkinson, K.J., Volmer, A.S., Burns, K.A., Evans, C.M., O'Neal, W.K., and Boucher, R.C. (2017). Contribution of mucus concentration and secreted mucins Muc5ac and Muc5b to the pathogenesis of muco-obstructive lung disease. Mucosal Immunol. 10, 395–407. https://doi.org/10.1038/mi.2016.63.

44. Park, S.W., Zhen, G., Verhaeghe, C., Nakagami, Y., Nguyenvu, L.T., Barczak, A.J., Killeen, N., and Erle, D.J. (2009). The protein disulfide isomerase AGR2 is essential for production of intestinal mucus. Proc. Natl. Acad. Sci. USA 106, 6950–6955. https://doi.org/10.1073/PNAS.0808722106.

45. Kesimer, M., Ehre, C., Burns, K.A., Davis, C.W., Sheehan, J.K., and Pickles, R.J. (2013). Molecular organization of the mucins and glycocalyx underlying mucus transport over mucosal surfaces of the airways. Mucosal Immunol. 6, 379–392. https://doi.org/10.1038/mi.2012.81.

46. Church, D.M., Goodstadt, L., Hillier, L.W., Zody, M.C., Goldstein, S., She, X., Bult, C.J., Agarwala, R., Cherry, J.L., DiCuccio, M., et al. (2009). Lineage-specific biology revealed by a finished genome assembly of the mouse. PLoS Biol. 7, e1000112. https://doi.org/10.1371/JOURNAL.PBIO.1000112.

47. Jiang, J., Yu, L., Huang, X., Chen, X., Li, D., Zhang, Y., Tang, L., and Zhao, S. (2001).

Identification of two novel human dynein light chain genes, DNLC2A and DNLC2B, and their expression changes in hepatocellular carcinoma tissues from 68 Chinese patients. Gene 281, 103–113. https://doi.org/10.1016/S0378-1119(01)00787-9.

48. Davis, J.D., and Wypych, T.P. (2021). Cellular and functional heterogeneity of the airway epithelium. Mucosal Immunol. 14, 978–990. https://doi.org/10.1038/s41385-020-00370-7.

49. Tsang, D.K.L., Wang, R.J., De Sa, O., Ayyaz, A., Foerster, E.G., Bayer, G., Goyal, S., Trcka, D., Ghoshal, B., Wrana, J.L., et al. (2022). A single cell survey of the microbial impacts on the mouse small intestinal epithelium. Gut Microb. 14, 2108281. https://doi.org/10.1080/19490976.2022.2108281.

50. Gordon, S., and Plüddemann, A. (2017). Tissue macrophages: Heterogeneity and functions. BMC Biol. 15, 1–18. https://doi.org/10.1186/s12915-017-0392-4.

51. Uthaiah, R.C., Praefcke, G.J.K., Howard, J.C., and Herrmann, C. (2003). IIGP1, an interferon-gamma-inducible 47-kDa GTPase of the mouse, showing cooperative enzymatic activity and GTP-dependent multimerization. J. Biol. Chem. 278, 29336–29343. https://doi.org/10.1074/JBC.M211973200.

52. Chen, Y., Zhang, J., Cui, W., and Silverstein, R.L. (2022). CD36, a signaling receptor and fatty acid transporter that regulates immune cell metabolism and fate. J. Exp. Med. 219, e20211314. https://doi.org/10.1084/JEM.20211314/213166.

53. Scott, N.A., and Mann, E.R. (2020). Regulation of mononuclear phagocyte function by the microbiota at mucosal sites. Immunology 159, 26–38. https://doi.org/10.1111/IMM.13155.

54. Rashidi, M., Bandala-Sanchez, E., Lawlor, K.E., Zhang, Y., Neale, A.M., Vijayaraj, S.L., O'Donoghue, R., Wentworth, J.M., Adams, T.E., Vince, J.E., and Harrison, L.C. (2018). CD52 inhibits Toll-like receptor activation of NF-κB and triggers apoptosis to suppress inflammation. Cell Death Differ. 25, 392–405. https://doi.org/10.1038/cdd.2017.173.

55. Tiribuzi, R., D'Angelo, F., Berardi, A.C., Martino, S., and Orlacchio, A. (2012). Knock-down of HEXA and HEXB genes correlate with the absence of the immunostimulatory function of HSC-derived dendritic cells. Cell Biochem. Funct. 30, 61–68. https://doi.org/10.1002/CBF.1819.

56. Zhang, S., Wang, H., Liu, J., Tao, T., Zeng, Z., and Wang, M. (2022). RGS1 and related genes as potential targets for immunotherapy in cervical cancer: computational biology and experimental validation. J. Transl. Med. 20, 334. https://doi.org/10.1186/S12967-022-03526-0.

57. Kurt-Jones, E.A., Cao, L., Sandor, F., Rogers, A.B., Whary, M.T., Nambiar, P.R., Cerny, A., Bowen, G., Yan, J., Takaishi, S., et al. (2007). Trefoil family factor 2 is expressed in murine gastric and immune cells and controls both gastrointestinal inflammation and systemic immune responses. Infect. Immun. 75, 471–480. https://doi.org/10.1128/IAI.02039-05.

58. Kaur, D., Patiyal, S., Sharma, N., Usmani, S.S., and Raghava, G.P.S. (2019). PRRDB 2.0: a comprehensive database of pattern-recognition receptors and their ligands. Database 2019, baz076. https://doi.org/10.1093/DATABASE/BAZ076.

59. Cash, H.L., Whitham, C.V., Behrendt, C.L., and Hooper, L.V. (2006). Symbiotic bacteria direct expression of an intestinal bactericidal

lectin. Science *313*, 1126–1130. https://doi.org/10.1126/SCIENCE.1127119.

60. Vaishnava, S., Behrendt, C.L., Ismail, A.S., Eckmann, L., and Hooper, L.V. (2008). Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host-microbial interface. Proc. Natl. Acad. Sci. USA *105*, 20858–20863. https://doi.org/10.1073/PNAS.0808723105.

61. Galves, M., Sperber, M., Amer-Sarsour, F., Elkon, R., and Ashkenazi, A. (2023). Transcriptional profiling of the response to starvation and fattening reveals differential regulation of autophagy genes in mammals. Proc. Biol. Sci. *290*, 20230407. https://doi.org/10.1098/RSPB.2023.0407.

62. Gallo, R.L., and Hooper, L.V. (2012). Epithelial antimicrobial defence of the skin and intestine. Nat. Rev. Immunol. *12*, 503–516. https://doi.org/10.1038/NRI3228.

63. Watson, A., Madsen, J., and Clark, H.W. (2020). SP-A and SP-D: Dual Functioning Immune Molecules With Antiviral and Immunomodulatory Properties. Front. Immunol. *11*, 622598.

64. Byrne, A.J., Powell, J.E., O'Sullivan, B.J., Ogger, P.P., Hoffland, A., Cook, J., Bonner, K.L., Hewitt, R.J., Wolf, S., Ghai, P., et al. (2020). Dynamics of human monocytes and airway macrophages during healthy aging and after transplant. J. Exp. Med. *217*, e20191236. https://doi.org/10.1084/JEM.20191236.

65. Jakubzick, C., Gautier, E.L., Gibbings, S.L., Sojka, D.K., Schlitzer, A., Johnson, T.E., Ivanov, S., Duan, Q., Bala, S., Condon, T., et al. (2013). Minimal Differentiation of Classical Monocytes as They Survey Steady-State Tissues and Transport Antigen to Lymph Nodes. Immunity *39*, 599–610. https://doi.org/10.1016/J.IMMUNI.2013.08.007.

66. Geissmann, F., Jung, S., and Littman, D.R. (2003). Blood Monocytes Consist of Two Principal Subsets with Distinct Migratory Properties. Immunity *19*, 71–82. https://doi.org/10.1016/S1074-7613(03)00174-2.

67. Kang, B., Alvarado, L.J., Kim, T., Lehmann, M.L., Cho, H., He, J., Li, P., Kim, B.H., Larochelle, A., and Kelsall, B.L. (2020). Commensal microbiota drive the functional diversification of colon macrophages. Mucosal Immunol. *13*, 216–229. https://doi.org/10.1038/s41385-019-0228-3.

68. Ibrahim, A.M., Moss, M.A., Gray, Z., Rojo, M.D., Burke, C.M., Schwertfeger, K.L., dos Santos, C.O., and Machado, H.L. (2020). Diverse Macrophage Populations Contribute to the Inflammatory Microenvironment in Premalignant Lesions During Localized Invasion. Front. Oncol. *10*, 1812. https://doi.org/10.3389/FONC.2020.569985.

69. Hassel, C., Gausserès, B., Guzylack-Piriou, L., and Foucras, G. (2021). Ductal Macrophages Predominate in the Immune Landscape of the Lactating Mammary Gland. Front. Immunol. *12*, 754661. https://doi.org/10.3389/FIMMU.2021.754661.

70. Wentworth, C.C., Alam, A., Jones, R.M., Nusrat, A., and Neish, A.S. (2011). Enteric Commensal Bacteria Induce Extracellular Signal-regulated Kinase Pathway Signaling via Formyl Peptide Receptor-dependent Redox Modulation of Dual Specific Phosphatase 3. J. Biol. Chem. *286*, 38448–38455. https://doi.org/10.1074/JBC.M111.268938.

71. Rathinam, V.A.K., Jiang, Z., Waggoner, S.N., Sharma, S., Cole, L.E., Waggoner, L., Vanaja, S.K., Monks, B.G., Ganesan, S., Latz, E., et al. (2010). The AIM2 inflammasome is essential for host defense against cytosolic bacteria and DNA viruses. Nat. Immunol. *11*, 395–402. https://doi.org/10.1038/ni.1864.

72. Patil, A., and Patil, A. (2020). CellKb Immune: a manually curated database of mammalian immune marker gene sets optimized for rapid cell type identification. Preprint at bioRxiv. https://doi.org/10.1101/2020.12.01.389890.

73. Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. *37*, 852–857. https://doi.org/10.1038/s41587-019-0209-9.

74. van Rossum, G., and Drake, F.L. (2001). Python Reference Manual.

75. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res. *13*, 2498–2504. https://doi.org/10.1101/gr.1239303.

76. Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J., et al. (2017). SCENIC: Single-cell regulatory network inference and clustering. Nat. Methods *14*, 1083–1086. https://doi.org/10.1038/nmeth.4463.

77. Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl. Environ. Microbiol. *79*, 5112–5120. https://doi.org/10.1128/AEM.01043-13.

78. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High resolution sample inference from Illumina amplicon data. Nat. Methods *13*, 581–583. https://doi.org/10.1038/NMETH.3869.

79. Callahan, B.J., McMurdie, P.J., and Holmes, S.P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J. *11*, 2639–2643. https://doi.org/10.1038/ISMEJ.2017.119.

80. McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., Desantis, T.Z., Probst, A., Andersen, G.L., Knight, R., and Hugenholtz, P. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. *6*, 610–618. https://doi.org/10.1038/ISMEJ.2011.139.

81. Sekiguchi, R., and Hauser, B. (2019). Preparation of Cells from Embryonic Organs for Single-Cell RNA Sequencing. Curr. Protoc. Cell Biol. *83*, e86. https://doi.org/10.1002/cpcb.86.

82. Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res. *27*, 491–499. https://doi.org/10.1101/GR.209601.116.

83. Lungova, V., Verheyden, J.M., Herriges, J., Sun, X., and Thibeault, S.L. (2015). Ontogeny of the mouse vocal fold epithelium. Dev. Biol. *399*, 263–282. https://doi.org/10.1016/j.ydbio.2014.12.037.

84. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675. https://doi.org/10.1038/nmeth.2089.

85. McGinnis, C.S., Murrow, L.M., and Gartner, Z.J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. Cell Syst. *8*, 329–337.e4. https://doi.org/10.1016/J.CELS.2019.03.003.

86. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., and Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat. Biotechnol. *42*, 293–304. https://doi.org/10.1038/s41587-023-01767-y.

87. Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. *37*, 38–44. https://doi.org/10.1038/nbt.4314.

88. Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502. https://doi.org/10.1038/nbt.3192.

89. Carlson, M. (2022). org.Mm.eg.db: Genome wide annotation for Mouse. R package version 3.8.2.

90. Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. *44*, W90–W97. https://doi.org/10.1093/NAR/GKW377.

91. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation *2*, 100141. https://doi.org/10.1016/j.xinn.2021.100141.

92. Aaron Lun, Assigning Cell Types with SingleR. R version 4.3.2 Patched (2023-11-13 r85521), https://github.com/LTLA?SingleRBook-Base.

93. Aran, D., Looney, A.P., Liu, L., Wu, E., Fong, V., Hsu, A., Chak, S., Naikawadi, R.P., Wolters, P.J., Abate, A.R., et al. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat. Immunol. *20*, 163–172. https://doi.org/10.1038/s41590-018-0276-y.

94. Büttner, M., Ostner, J., Müller, C.L., Theis, F.J., and Schubert, B. (2021). scCODA is a Bayesian model for compositional single-cell data analysis. Nat. Commun. *12*, 1–10. https://doi.org/10.1038/s41467-021-27150-6.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-Scgb1a1	Abcam	RRID: AB_2650558
goat anti-rabbit AF584	ThermoFisher Scientific	RRID: AB_143165
Vectashield with DAPI	Vector Laboratories	RRID: AB_2336788
Chemicals, peptides, and recombinant proteins		
Periodic Acid Schiff (PAS) Stain Kit	Newcomer	Part # 9162B
Accutase	Stem Cell Technologies	Catalog # 07920
Accumax	Stem Cell Technologies	Catalog # 07921
Bacillus lichenformis protease	MilliporeSigma	P5380
DNase I	MilliporeSigma	11284932001
Critical commercial assays		
DNeasy Blood & Tissue Kit	Qiagen	Catalog#69504
Qubit High Sensitivity DNA Kit	Thermo Fisher Scientific	Catalog#Q32851
Chromium Single Cell 3' Reagent Kit (v3.1)	10x Genomics	PN-1000268
Deposited data		
Single cell RNA-seq data for mouse larynges	Gene Expression Omnibus	[GeneExpressionOmnomics]: [GSE203233]
16s rRNA gene sequencing data for laryngeal microbiome	GenBank	[GenBank]: [PRJNA839147]
Experimental models: Organisms/strains		
Wildtype C57BL/6J conventionally raised mice (ConvR)	Jackson laboratory	Stock No. 000664
Wildtype C57BL/6J Germ-free mice (GF)	BRMS Gnotobiotic Core (UW-Madison)	N/A
Oligonucleotides		
515F: GTGCCAGCMGCCGCGGTAA	Integrated DNA Technologies	N/A
806R: GGACTACHVGGGTWTCTAAT	Integrated DNA Technologies	N/A
Software and algorithms		
RStudio (R4.1.2)	RStudio Team	http://www.rstudio.com/
R (R4.1.2)	R Core Team	https://www.R-project.org/
Cell Ranger (v7.1.0)	10x Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger
CellKb (v2.0)	Patil & Patil, 2020[72]	https://cellkb.combinatics.com/
QIIME2 (v2021.8)	Bolyen et al. 2019[73]	https://doi.org/10.1038/s41587-019-0209-9
Python3 (v3.10.1)	van Rossum & Drake, 2001[74]	http://www.python.org
Cytoscape (v3.8.2)	Shannon et al. 2003[75]	https://cytoscape.org/
SCENIC (v0.2.0)	Aibar, S. et al. 2017[76]	https://scenic.aertslab.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Susan Thibeault (thibeault@surgery.wisc.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- ScRNA-seq data are accessible in Gene Expression Omnibus: GSE203233. 16S rRNA gene sequencing data are accessible in GenBank: PRJNA839147. Microscopy data in this paper will be shared by the lead contact upon request.
- All original code is available from the lead contact upon request.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All animal studies were performed in accordance with the approved protocol (M005669) from the Animal Care and Use Committee at the University of Wisconsin–Madison. All experiments were performed on male mice of 8-10 weeks old. Considering that little is known regarding the impact of microbiota on laryngeal immunity, we used male mice only to avoid potential sex effect. Wild-type C57BL/6J mice (Stock no. 000664) were used for all experiments. ConvR mice were obtained from the Jackson Laboratory and housed under specific-pathogen-free conditions at the Biomedical Research Model Services (BRMS) breeding core for two weeks until use; C57BL/6J GF mice were obtained from the BRMS Gnotobiotic Core adjacent to the breeding core and housed there until use. Sterility of GF mice was confirmed by checking the most densely populated microbial niche cecum contents through PCR (16S rRNA gene) and culture-dependent methods. All mice were fed with autoclaved mouse breeder diet Labdiet 5021 (Purina, St. Louis, MO) and sterilized reverse osmosis water.

METHOD DETAILS

16S rRNA gene sequencing

Larynges were collected from ConvR (n=6) mice, minced into 0.5 mm small pieces on a sterile petri dish under a dissection scope. Bacterial cells were collected by washing the minced tissue in 500 µl sterile PBS in a 2 ml microcentrifuge using a vortex with flat bed at the maximum speed for three times. Lavages were combined and centrifuged at 15,000 rpm for 5 min. Cell pellet was digested in enzymatic lysis buffer consisting of 20 mM Tric-Cl, 2 mM EDTA, 1.2% Triton-x 100, and 20 mg/ml lysozyme, for 1 hour at 37°C. DNA extraction was subsequently performed using Qiagen DNeasy Blood & Tissue kit (Qiagen, Germantown, MD) following the manufacture instructions. V4 regions of the 16S rRNA gene were amplified using MolTaq 16S/18S DNA-free Taq Polymerase (Molzym GmbH & Co. KG, Bremen, Germany) in a 25 µl reaction containing 10 ng of DNA template and 200 nM of 515F/806R primers fused with Illumina sequencing adapters.[77] Non-template control (NTC, n=3) and a PCR positive control with Helicobacter pylori genomic DNA as template was included for each PCR run. PCR cycling conditions were as follows: one cycle of enzyme activation at 95°C for 3 min followed by 35 cycles of denaturation at 95°C for 30 sec, annealing at 55°C for 30 sec, and extension at 72°C for 30 sec, and a final extension at 72°C for 5 min. Each sample was amplified in triplicate. Resulting PCR products were identified on 2% agarose gel, and the successful reactions were pooled, purified with NucleoSpin Gel and PCR cleanup kit (Takara, San Jose, CA), and quantified with Qubit dsDNA HS Assay Kit (Invitrogen, Eugene, OR). Samples were equimolar pooled and sequenced on the Illumina MiSeq 2x250 bp platform.

NTC were confirmed to have no visible amplicon bands and were excluded from subsequent analysis. Given that the larynx is low microbial biomass sample, extraction negative control (ExNeg, n=4) containing only sterile PBS were processed in parallel with laryngeal tissues to control the potential environmental contamination. Extremely faint bands were observed for extraction negative controls; therefore, the amplicons were gel-purified and sequenced as well. Oral Microbiome Whole Cell Mix (ATCC, Manassas, VA) (Oral, n=4) was included as positive control, processed, and sequenced in parallel with tissue samples, to examine the effectiveness of sample/data processing methods. The mock community was a defined synthetic community comprised of equal portion of *Schaalia odontolytica* (ATCC 17982), *Prevotella melaninogenica* (ATCC 25845), *Fusobacterium nucleatum* subsp. *nucleatum* ( ), *Streptococcus mitis* (ATCC 49456), *Veillonella parvula* (ATCC 17745), and *Haemophilus parainfluenzae* (ATCC 33392), with each representing 16.7% of the total cells.

Sequences were processed using *QIIME2* pipeline.[73] Demultiplexed 250 bases paired-end sequences were imported using Casava 1.8 format and denoised using *DADA2* to obtain amplicon sequence variant (ASV) table.[78,79] Singletons (ASV present < 2 times) and ASVs that are present in less than 10% of the samples were discarded. Given that larynx is low microbial biomass tissue, additional filtering was performed, including 1) the removal of any ASV whose total abundance across all samples was above 1% of the total abundance in the ExNeg, and 2) the removal of any ASV whose total abundance across all samples is below 1% of the total abundance of all ASVs across all samples. Greengenes reference sequencescs (clustered at 99% similarity) were used to train a naïve Bayes taxonomy classifier to further annotate ASVs taxonomically.[80] ASVs were then collapsed based on genus level. ASV files was imported and visualized in RStudio (v4.1.2).

Dissociation of laryngeal tissues

Mouse larynges were collected from ConvR and GF mice and minced into 0.5 mm small pieces on a sterile petri dish under a dissection scope, then digested in a 1 ml modified dissociation solution for 40 min in a cold room (4°C). We started from the dissociation protocol by Sekiguchi & Hauser (2019) and determined the optimal dissociation condition by optimizing the enzyme composition & concentration and dissociation temperature & time.[81] Specifically, 1 ml of modified dissociation solution contained 5 mM $CaCl_2$, 90 µl Accutase (Stem Cell Technologies,

Cambridge, MA), 90 μl Accumax (Stem Cell Technologies, Cambridge, MA), 10 mg/ml Bacillus lichenformis protease (MilliporeSigma, Burlington, MA), and 125 U/ml DNase I (MilliporeSigma, Burlington, MA) dissolved in sterile calcium/magnesium-free PBS.[81] Samples were gently triturated every 5 min using a wide-bore pipette tip during dissociation. After dissociation, cell suspension was filtered through 40 μm cell strainer to remove aggregates. Undigested cartilage was removed before cell straining to prevent clogging cell strainer and further impact dissociation efficiency. Dissociation was performed for each larynx separately. Single-cell suspensions from three larynges per mouse group were pooled as one sample (3 biological replicates per mouse group). Pooled suspensions were added to 5 ml Ca/Mg-free PBS supplemented with 10% FBS, filtered with a 40 μm cell strainer (MilliporeSigma, Burlington, MA), then centrifuged at 300 g, for 10 min, at 4°C. Cell pellet was resuspended in 30 μl ACK lysing buffer (Lonza, Bend, OR) briefly to remove erythrocytes, following manufacture's instruction.

Droplet-based scRNA-seq

ACK treated single cells were resuspended in 500 μl of Ca/Mg-free PBS supplemented with 2% FBS. Libraries were prepared according to Chromium Single Cell 3′ Reagent Kit v3.1 (10x Genomics, Pleasanton, CA). Briefly, single cell suspensions were transported on ice to the University of Wisconsin-Madison Biotechnology Center, where cell concentration and viability were quantified on the Countess II (Thermo Fisher Scientific, Waltham, MA) using 0.4% Trypan Blue (Invitrogen, Carlsbad, CA). Cells were then partitioned into Gel Beads in Emulsion in the GemCode instrument, where cell lysis and barcoded reverse transcription of RNA occurred, followed by amplification, shearing and 5′ adaptor and sample index attachment. Library preparation was performed to generate a single pool of cDNA containing both ConvR and GF samples to minimize batch effect between samples. Libraries were quantified using the Qubit High Sensitivity DNA Kit (Thermo Fisher Scientific, Waltham, MA), uniquely indexed, pooled, balanced for the number of estimated reads per cell with a *MiSeq* (Illumina, San Diego, CA) sequencing run, and subsequently sequenced on a *NovaSeq SP* flowcell for the first batch of two samples (1 ConvR and 1 GF) and S1 flowcell for the second batch of 4 samples (2 ConvR and 2 GF) (Illumina, San Diego, CA) using a 150 x 150 bp sequencing reaction targeting > 60,000 reads/cell, yielding a calculated depth of over 100K mean reads per cell for samples of the first batch and over 70K mean reads per cell for samples of the second batch. MiSeq balancing run was quality controlled using calculations based on UMI-tools.[82] Sequencing of the first batch of samples was performed first to serve as a pilot experiment for the second batch.

Immunofluorescence (IF) staining

Larynges were fixed in 4% paraformaldehyde at 4°C overnight and remained in 70% ethanol until standard tissue processing procedures for paraffin sections. Paraffin-embedded sections were cut coronally (n=6) into 5 μm sections and stored at 4°C until use. Paraffin-embedded were deparaffinized, rehydrated, and stained using a standard IF protocol.[83] Antigen retrieval was performed by heating sections in sodium citrate pH= 6 at 80°C water bath for 2 hours. Primary antibodies anti-Scgb1a1 (Abcam, Cambridge, CA) was diluted at 1:100; goat anti-rabbit AF584 (ThermoFisher Scientific, Waltham, MA) diluted at 1:500 was used as secondary antibodies for SCGB1A1 staining. Sections were incubated with primary antibodies at 4°C overnight followed with secondary antibodies for 1.5 hours. Slides were mounted using Vectashield with DAPI (Vector Laboratories, Burlingame, CA). Images were taken with a Nikon Eclipse Ti2 inverted microscope with a DS-Qi2 high-sensitivity monochrome camera (Nikon Instruments Inc., Melville, NY) and adjusted for brightness using the NIS Elements software. Fluorescence intensity of SCGB1A1 was measured with ImageJ based on 30 images (6 biological replicate, 5 images per replicate) per mouse group following the instruction (https://kpif.umbc.edu/image-processing-resources/imagej-fiji/determining-fluorescence-intensity-and-positive-signal/).

Periodic acid – Schiff (PAS) staining

Larynges of ConvR (n=6) and GF (n=6) were fixed overnight in methacarn solution consisting of 60% (vol/vol) absolute methanol, 30% chloroform, and 10% glacial acetic acid, washed in methanol (2 x 30 min) and 100% ethanol (2 x 20 min), cleared in xylene (2 x 15 min), and infiltrated with hot paraffin (60°C) for 2 hours. Paraffin-embedded tissues were cut coronally into 5-μm sections and stored at 4°C until use. PAS staining was performed following manufacture's instruction (Newcomer, Middleton, WI). Glycogen was digested with 1% α-amylase solution for 30 min at room temperature prior to PAS staining. Images were taken with a Nikon Eclispe Ti2 inverted microscope (Nikon Instruments Inc., Melville, New York) with a DS-Ri2 high-speed color camera (Nikon Instruments Inc., Melville, New York) and were adjusted for brightness with a corresponding analysis software NIS Elements (Nikon Instruments Inc., Melville, New York). A total of 30 PAS-stained images (5 images per biological replicate, 6 biological replicates) were opened in ImageJ to quantify mucous cell number, mucus cell size, and mucus substances.[84]

Preprocessing of scRNA-seq data

For each sample of ConvR and GF, raw reads obtained from the 10x Genomics' Chromium single-cell RNA-seq (scRNA-seq) platform were demultiplexed and mapped to the mouse reference genome (GRCm38) using Cell Ranger software (v7.1.0) with default parameters following standard protocols. Cells with less than 500 total UMI counts, over 20% mitochondrial gene ratio, and less than 250 detected genes/features were considered low-quality and filtered out. Cells with $\log_{10}$GenesPerUMI < 0.8 were removed from downstream analysis, where the metric was determined by taking the $\log_{10}$ of the total UMI counts and dividing that by the log base 10 of the feature. Genes that were not expressed in at least 10 cells were also excluded from downstream analysis. Normalization was then conducted using the Seurat NormalizeData function. Detection of technical artifact doublets was performed using DoubletFinder (v2.0.3),[85] under default settings. Few barcodes were predicted

as suspicious doublets and they were distributed randomly across all cells without forming their own cluster, therefore they are likely singlets mislabeled as doublets and were not filtered out.

QUANTIFICATION AND STATISTICAL ANALYSIS

Subsequent analysis of the scRNA-seq dataset was mainly performed using Seurat v4.3.0.1.[86]

Dimensionality reduction using PCA and UMAP

Variable genes selection, principal component analysis (PCA), and uniform manifold approximation and projection (UMAP) were conducted using the Seurat (v 4.0.5) pipeline.[87,88] Specifically, selection of variable genes was implemented with the function *FindVariableFeatures* with method='vst'. The top 2000 highly variable genes were centered and scaled with function *ScaleData*. The first 25 PCs were calculated via function *RunPCA*, after which the function *RunUMAP* was applied to generate UMAP embeddings for 2-dimensional visualization.

Merged analysis of the scRNA-seq data for ConvR and GF

The two scRNA-seq datasets, for ConvR and GF samples, were merged by combining cell matrices using Merged function from Seurat. Merged data was then processed following the Seurat pipeline, as mentioned above. Batch effects were well-controlled between the six samples as demonstrated in Figure 1B, where cells in each cell type displayed similar patterns in each group of mice.

Clustering and subclustering

To cluster single cells by their expression profiles, we used an unsupervised clustering approach based upon the Louvain algorithm implemented in Seurat. Briefly, a shared nearest neighbor (SNN) graph was constructed by calculating neighborhood overlap between every cell and its nearest neighbors. Clusters were then identified by optimizing the modularity. Seurat function *FindNeighbors* was applied by adjusting the dim =1:25, and *FindClusters* was applied under default parameters. For subclustering of cell types of interest, namely basal epithelial cells, secretory epithelial cells, ciliated epithelial cells, fibroblasts, and macrophages, we applied the same clustering process on the subset of cells, in each cell type. Resolution parameters in *FindClusters* were manually specified to provide a reasonable number of clusters.

Differential expression and enrichment analysis

We performed differential expression analysis between two groups of cells using the *FindMarkers* function in Seurat to identify cell type specific markers. A non-parametric Wilcoxon rank-sum test was used with adjusted p-values calculated using the Bonferroni correction to adjust for multiple tests. Genes with adjusted p-values < 0.05 and an average $\log_2$ fold change (avg_$\log_2$FC) > 0.25 or < -0.25 were deemed differentially expressed (DE) between cell (sub)types and used in downstream analysis. The top 5 upregulated, non-ribosomal, non-mitochondrial DE genes ranked by average $\log_2$ fold change were selected as putative marker genes for each cluster. For each sample of ConvR and GF, DE genes for each cell type were determined by comparing a cell type versus remaining cells in the sample. Similarly, DE genes of each cell subtype were determined by comparing a cell subtype versus the remaining cells within that cell type. At least 1 top upregulated non-ribosomal DE gene(s) ranked by average $\log_2$ fold change was selected as the putative signature gene(s) for each subcluster. In order to elucidate the differentially expressed (DE) genes between conditions in a specific cell type, a pseudobulk method was conducted using the integrated dataset. This approach took into account the inherent correlation structure present within the data. We aggregate the cellular expression profiles within each sample to mitigate the effects of individual cellular variability. Subsequently, the DESeq2 framework was utilized as the primary tool for comparing expression levels. This methodological choice allowed for a robust and statistically sound identification of DE genes. In this comparison, DE genes with avg_$\log_2$FC > 0.25 or < 0.25 and adjusted p-value < 0.1 were deemed as significant. Mitochondrial genes (genes that begin with "mt-") and ribosomal genes (genes that begin with "Rps-") were removed, prior to further analysis, from the DE genes given that these genes are not of our interest in the present study. A list of the DE genes (avg_$\log_2$FC > 0.25, adjusted p-values < 0.05, unless otherwise specified) were analyzed for enrichment for common biological process (BP), molecular function (MF), and cellular components (CC) using the R package *enrichR* (v3.0). Enrichment results were visualized with tables or upset plots using defined functions provided in the R script.[89–91]

Cell type annotation

SingleR (v1.4.0) was used to annotate clusters into cell types using the pre-annotated built-in mouse bulk RNA-seq data as a ref.[92] Specifically, *SingleR* trains a classification model based on the reference data and assigns a cell type label based on the highest predicted probability. We also performed additional manual annotation by identifying cell type-specific marker genes among the top differentially expressed genes in each cluster using external literature, since *SingleR* is unable to annotate cell types not existing in the reference data. Specifically, for each cluster, we mapped the top 30 DE genes ranked by the average $\log_2$ fold change against cell-type ontology in the database *CellKb Basic*[72,93] (v2.0). Final cell types were determined by refining *SingleR* annotations using manual annotations.

Protein-protein interaction (PPI) network analysis

The interaction of DE genes at the protein level was analyzed using StringDB (string-db.org). Medium confidence PPI clustered networks with a minimum score of 0.4 and Markov Cluster Algorithm inflation parameter as 2 were built. KEGG pathway enrichment analysis was performed,

with FDR < 0.05 and minimum gene count as 3 considered to be statistically significant. Within the PPIs, hub genes (the most interactive genes) were identified using the cytoHubba plugin in the open-source software Cytoscape (v3.8.2). MCC algorithm was used to identify the top 10 hub genes.[75]

Gene regulatory network analysis

To construct gene regulatory networks (GRN), single-cell regulatory network inference (SCENIC, v0.2.0) was used to identify TFs and associated target genes with Python3 (v3.10.1), referred to as regulons, in each cell type of ConvR and GF groups.[74,76] A regulon activity score (RAS) was calculated for each regulon in each single cell by summing up the area under the recovery curve. In short, SCENIC calculates enrichment of a regulon as an area under the recovery curve, where the recovery curve plots gene expression of all genes in a cell (x-axis) vs. number of genes recovered from the regulon (y-axis). The regulon specificity score (RSS) for a cell type was then calculated according to the entropy of RAS of cells within the cell type compared to other cell types. An RSS ranges from 0 to 1, with a higher value representing greater specificity of a regulon in the cell type. To reduce computation time and effectively increase the data quality, for cell (sub)types with more than 10 cells, we randomly average every 10 cells, so that each group produced a pseudo-cell by the average gene expression profile of cells within the group, and SCENIC was applied to the pseudo-cells. RAS was used to identify differential regulons between conditions use Pseudobulk method as described before.

Comparison of cell proportions

Differences in cell proportions for each cell (sub)type between ConvR and GF were evaluated using scCODA with the false discovery rate cutoff set to 0.1 or 0.4.[94] Asterisk (*) was used to indicate significant difference between groups.

Comparison of cell features

A non-parametric Wilcoxon rank-sum test ($p = 0.05$) was performed for the comparisons of SCGB1A1 fluorescence intensity, mucous cell/size, and mucus amount between ConvR and GF mice. Asterisk (*) was used to indicate significant difference between groups.