The healthy lung was long thought of as sterile, but recent advances using molecular sequencing approaches have detected bacteria at low levels. Healthy lung bacteria largely reflect communities present in the upper respiratory tract that enter the lung via microaspiration, which is balanced by mechanical and immune clearance and likely involves limited local replication. The nature and dynamics of the lung microbiome, therefore, differ from those of ecological niches with robust self-sustaining microbial communities. Aberrant populations (dysbiosis) have been demonstrated in many pulmonary diseases not traditionally considered microbial in origin, and potential pathways of microbe-host crosstalk are emerging. The question now is whether and how dysbiotic microbiota contribute to initiation or perpetuation of injury. The fungal microbiome and virome are less well studied. This Review highlights features of the lung microbiome, unique considerations in studying it, examples of dysbiosis in selected disease, emerging concepts in lung microbiome–host interactions, and critical areas for investigation.
The lung microbiome: progress and promise

Samantha A. Whiteside, John E. McGinniss, and Ronald G. Collman

Division of Pulmonary, Allergy and Critical Care Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA.

The healthy lung was long thought of as sterile, but recent advances using molecular sequencing approaches have detected bacteria at low levels. Healthy lung bacteria largely reflect communities present in the upper respiratory tract that enter the lung via microaspiration, which is balanced by mechanical and immune clearance and likely involves limited local replication. The nature and dynamics of the lung microbiome, therefore, differ from those of ecological niches with robust self-sustaining microbial communities. Aberrant populations (dysbiosis) have been demonstrated in many pulmonary diseases not traditionally considered microbial in origin, and potential pathways of microbe-host crosstalk are emerging. The question now is whether and how dysbiotic microbiota contribute to initiation or perpetuation of injury. The fungal microbiome and virome are less well studied. This Review highlights features of the lung microbiome, unique considerations in studying it, examples of dysbiosis in selected disease, emerging concepts in lung microbiome–host interactions, and critical areas for investigation.

Introduction

It is estimated that the human body contains roughly equal numbers of human and bacterial cells (1), and over the past decade studies of the human microbiome have expanded from the gut and other microbe-rich environments to sites previously considered sterile, such as the bladder (2), and the lung, the focus of this Review. Although the lungs were known to harbor microbes in acute infections and chronic suppurative diseases such as bronchiectasis and cystic fibrosis (CF), microbial communities in the healthy lung were only recently described through culture-independent methods (3–5). It is not surprising that the lung is not sterile, given its contiguity with the microbe-rich upper respiratory tract. However, the healthy lung possesses exceedingly low microbial biomass, and the nature of the lung microbiome differs markedly in quantity and dynamics from that of body habitats with robust microorganisms, such as the gut, skin, mouth, and vagina.

Many respiratory illnesses not traditionally considered microbially driven have now been linked to lung microbiota that differ from those of healthy people (“dysbiosis”), although research has only begun to address causality and potential mechanisms. While investigators differ in precisely how the term is interpreted and employed, “dysbiosis” refers broadly to disruption of microbial homeostasis with compositional or functional deviation from normal or optimal in a particular niche (6, 7). In the lung, consensus has not yet emerged on what defines dysbiosis; here we use the term to refer to deviation from microbiota seen in healthy people without lung disease, recognizing that there may be optimal microbial communities in certain disease states, such as CF, that nonetheless differ from health. Importantly, given the exceedingly low microbial burden of the normal lung, dysbiosis includes not just changes in relative abundances (proportions) of microbiome constituents within a community, but changes in total microbial content as well.

Lung microbiome publications are increasing exponentially, and, similarly to how microbiome research in other sites matured, the key questions now extend beyond “who is there?” to: What are they doing? How do they impact the host? What are the mechanisms? And can we harness them therapeutically?

Approach to studying the lung microbiome

Analytic approaches. In contrast to culture methods that focus on individual organisms of a priori interest, microbiome studies use sequence-based approaches that capture entire communities. The most widely used approach employs PCR amplification and sequencing of a region shared by all members of a group that can then be used to infer identity, such as the 16S rRNA gene of bacteria and archaea (the latter of uncertain importance in the lung) and the 18S rRNA gene or internal transcribed spacer (ITS) regions of fungi (8). Sequencing of the 16S rRNA gene typically targets one or more of the nine hypervariable regions (V1–V9) that differ across taxa. While efficient, this approach is limited in species-level identification, a deficit given the wide range in biological activities between species and even strains. Full-length 16S rRNA gene sequencing provides greater taxonomic definition (9, 10) and has been applied on a limited basis in the lung (11, 12), but short-read 16S rRNA gene variable region sequencing remains most common. Regardless of methodology, 16S rRNA gene profiling is largely restricted to descriptive correlations between taxonomy and host phenotype.

Another approach is shotgun metagenomic sequencing, in which all the DNA in a sample is sequenced. Although the high proportion of host-derived DNA present in lung samples is a challenge, shotgun sequencing is being increasingly used in lung studies (13–17). Shotgun metagenomics provides taxonomic composition and microbiome gene content, which enumerates the community’s functional potential. It can annotate bacterial strains and includes information regarding carriage of antibiotic resistance or virulence.
genes. Shotgun metagenomics concurrently identifies bacterial and fungal constituents, and also viruses, which do not share any sequence tags comparable to bacteria and fungi.

Both 16S rRNA gene sequencing and shotgun sequencing are associated with reads that do not align to reference databases (18). Furthermore, there may be low-abundance taxa with functional impact, such as Erysipelotrichaceae, that are present below the limits of detection or difficult to distinguish from background contaminant sequences (19). Notably, neither targeted sequencing nor shotgun metagenomics differentiates between live and dead bacteria. Distinguishing viability is particularly relevant in the lung, where a balance exists between microbial entry and clearance, and some diseases are associated with abundant relic DNA, such as CF (20, 21). Also, shotgun metagenomics details the functional potential of a microbial community based on gene content but not its actual activity.

Beyond shotgun metagenomics, metatranscriptomics (RNA sequencing and metabolomics—small-molecule analysis) depend on living cells and may better reflect functional activity (13). Finally, sequence-based methods are complementary to culture, which has high sensitivity for targeted organisms (22). Thus, molecular approaches are not inherently better than culture, and combining culture and sequencing approaches may provide optimal community resolution in specimens dominated by human DNA and/or low-abundance taxa (23), and also provide an opportunity to study the specific organisms.

Sequence-based studies typically report taxa in terms of relative abundances, that is, proportionality within a community that totals 100%. However, it is essential to consider relative enrichment of specific taxa in the context of the total microbial burden, particularly in the lung, where the microbial burden varies over many orders of magnitude. Indeed, bacterial burden may be as important as composition, as described in studies where total bacterial burden in bronchoalveolar lavage fluid (BALF) predicted chronic lung allograft dysfunction or death following lung transplantation (24), progression of interstitial pulmonary fibrosis (25, 26), and outcomes during critical illness (27).

**Sampling the lung microbiome.** Sputum and BALF are common approaches to lung microbiome sampling and are associated with distinct biases. Sputum is noninvasive and can be collected serially but represents a mix of upper and lower respiratory tract (URT and LRT). This admixture is especially problematic in states in which the lung microbial burden is low, though sputum may better reflect the lung in conditions like CF in which the LRT microbial burden is high (28). There is little consensus regarding how well microbial communities in spontaneous and induced sputum correlate (29, 30). While bronchoscope passage through the URT means that BALF may contain carryover (particularly of periglottic material; refs. 3, 31), it is less influenced by contamination than sputum but requires an invasive procedure. Still, careful analysis is necessary to distinguish lung versus other sources. The use of whole versus host cell–depleted BALF varies by study, though cell-depleted BALF often shows loss of α-diversity, a marker of how many different taxa are present and how evenly they are distributed, likely from removal of cell-associated bacteria (32). Other sampling options with both advantages and limitations include endotracheal aspirates, used extensively in intubated patients (33-35); protected specimen brushings (36, 37); bronchial biopsies (38, 39); exhaled breath condensates (40, 41); and tissue from explanted lungs (42-44).

The low microbial biomass of lung specimens except in highly colonized or purulent conditions makes lung samples exquisitely susceptible to contamination from the URT during collection, and also from background sequences derived from equipment, kits, and reagents, which are sterile but not necessarily DNA free (3, 45, 46). Artificial sequences can also result from the sequencing process itself (47). Many BALF samples from healthy people have very low bacterial content and a high relative abundance of taxa characteristic of background controls (e.g., bronchoscope channel wash prior to the procedure), indicating a paucity of authentic lung microbiota (48, 49). A point of some controversy is whether taxa suspected as contaminants should be removed before analysis (8, 45). In shotgun sequencing, a large portion of reads in low-microbial-content samples will correspond with human as opposed to microbial sequences, impeding the capture of microbial constituents. Thus, executing and interpreting lung microbiome studies requires careful attention to stringent sampling and the distinguishing of authentic signals from both URT contamination and background artifact.

**The healthy lung microbiome**

*Nature of the lung microbiome in health.* Traditional teaching was that healthy lungs were sterile, meaning the airways below the vocal cords (glottis) and distal to the alveoli, but studies using rigorous sampling revealed that bacterial sequences are usually detectable in the lung (3, 37, 50). However, bacterial density is orders of magnitude lower in the lung than in the URT, and the community composition closely matches the URT, especially the periglottic region. Firmicutes and Bacteroidetes are the most common phyla, while Prevotella, Streptococcus, and Veillonella are the most common genera (3, 51). The nearly identical composition of two contiguous communities, one of which has high microbial biomass (URT) and one of which has low (lung), suggests that the lung microbiota is derived passively from the URT through microaspiration, which occurs even in health (52). The lung microbiome, therefore, is largely transient and composed of URT-derived immigrants that are cleared through mucociliary and innate immune mechanisms. Some of these bacteria are viable and transcriptionally and metabolically active (13, 53). Bacterial levels are greater in more proximal lung regions, and modest regional differences exist as well, suggesting differential clearance and, potentially, limited local replication (37). Whether there are true long-term, self-sustaining bacterial residents is unclear. Thus, the nature and the dynamics of the healthy lung microbiome differ from those in sites such as gut, skin, and vaginal and oral cavities, which have robust, self-sustaining microbial communities.

This work has led to a healthy lung microbiome model that accounts for a steady-state low-level microbial community as a consequence of immigration from the URT, clearance, and limited local replication, sometimes termed the “adapted island model” (37). This concept is supported by murine studies modeling aspiration, in which the bacterial RNA metatranscriptome signal disappeared from the lung within a day while DNA sequences persisted for several days (13). This model is useful as a basis for understand-
In lung disease, structural disturbances and alterations to the local environment favor microbial growth, which predominantly determines lung microbial replication in the airways and lung parenchyma is limited. In healthy lungs, the local microbial replication via immune and mechanical processes. In healthy lungs, the controlled nature of animal models. The impact of smoking alters the lung microbiome in mice (56); this may be detectable because of the controlled nature of animal models. The impact of other environmental exposures on the lung microbiome is understudied. For example, lung microbiota in people highly exposed to indoor air pollution are enriched in Streptococcus and Neisseria (57). Further, URT dysbiosis may be an undervalued driver of lung dysbiosis, as recently described in patients with end-stage lung disease, and following lung transplantation (58).

In contrast, hematogenous sources for the lung microbiome seem limited to disease states. Lung infection from hematogenous seeding occurs commonly in bacteremia and fungemia. Sequences characteristic of gut bacteria are found in BALF in acute respiratory distress syndrome (ARDS), likely reflecting gut translocation and hematogenous spread; whether this process contributes to local inflammation in ARDS is unknown (59). Lung enrichment with Tropheryma whipplei, typically considered a gut bacteria, is seen in asymptomatic people with HIV infection, though its presence has unclear consequences (60).

**Functional impact of the lung microbiome in health.** The gut microbiome is central in regulating both gut mucosal and systemic immune development and function (61–63). In the lung, emerging data suggest that, although scant, the microbiota even in health regulates immune tone (Figure 2). Healthy people can be grouped into individuals whose BALF microbiota resembles oral communities, indicating a URT-derived microbiome, and individuals with BALF that resembles background controls, suggesting an absence of (or below-detectable) authentic lung microbiota (64). People whose BALF was enriched in oral commensals exhibited a Th17/neutrophilic mucosal immune profile with less robust innate immunity. This phenotype was associated with distinct metabolic profiles, suggesting a potential mechanism of regulation. Similarly, community composition and bacterial diversity in mouse lungs correlated inversely with the inflammatory cytokines IL-1α and IL-4 (65). While inflammation is an obvious component of acute infection, these data indicate that a key frontier is to understand how the microbiome regulates lung immune function in health and nonpurulent diseases. Both live and dead bacteria could modulate immune function, via metabolite production by live organisms, and pattern recognition receptors by both.

Studies in mice also suggest that lung microbiota play a role in local immune development and tolerance. In germ-free mice, lung inflammatory responses are exaggerated, and neonatal animals exhibit robust responses to allergens that subsequently dampen as the bacterial burden increases and regulatory T cells appear (66, 67). Mice treated with antibiotics have increased susceptibility to respiratory infection following pathogen exposure (68). While it can be difficult to distinguish direct effects of the lung microbiome from pathways involving gut microbiota in these models, a recent study showed that intratracheal instillation of oral commensals in a mouse model of microaspiration induced a Th17 response and resistance to Streptococcus pneumoniae infection (69). Thus, physiological levels of lung bacteria likely shape local immune responses.

**The lung mycobiome and virome.** Compared with bacteria, the fungal and viral communities (termed “mycobiome” and “virome”) of the lung are understudied. Fungal analysis employs targeted ITS or 18S rRNA gene or shotgun metagenomic sequencing. Mycobiome studies are challenging because of the low fungal biomass of the lung, the low number of fungal taxa present, difficulty in DNA extraction from fungi, biases in 18S rRNA gene and ITS amplification, inconsistent nomenclature, and suboptimal annotation of fungal databases, as well as frequently robust fungal populations in the URT (70, 71). Ascomycota and Basidiomycota are the most commonly identified taxa, followed by the genera Candida, Saccharomyces, Penicillium, Cladosporium, and Fusarium (72–75). Whether the lung mycobiome contributes to normal physiology remains unknown.
Viruses do not share any conserved regions comparable to bacterial 16S or fungal 18S genes. Thus, most studies of viruses in the lungs use targeted detection of specific viruses, whereas comprehensive delineation of the virome requires shotgun sequencing. Studies of the lung virome have found Anelloviridae as the predominant eukaryotic virus, with sporadic detection of herpesviruses, papillomaviruses, retroviruses, and other respiratory viruses (76–78). Anelloviridae are DNA viruses with greater than 90% prevalence in adults and wide distribution in blood and tissues (79). They are under tonic immune control, since levels rise with immune dysfunction (76), but have not been linked to physiological effects on the host or disease (80). Bacteriophages are also common, particularly those associated with ororespiratory tract bacterial hosts.

Virome studies are further challenged by the high proportion of shotgun sequences that fail to align with anything in current databases (called “metagenomic dark matter”), raising the possibility of yet-undiscovered viruses. A newly recognized virus family was identified in metagenomic sequences from the airway of critically ill compared with healthy people (81). As with Anelloviridae, it is unknown whether Redondoviridae play a role in lung health or pathogenesis.

The lung microbiome in disease

Dysbiosis of the lung microbiome can result from disruption in entry, clearance, or local replication. Greater entry can result from increased aspiration, and dysbiotic URT microbiota may have downstream impact on the lung. Indeed, it has long been known that colonization of the URT with respiratory pathogens including Streptococcus pneumoniae or Gram-negative bacteria often precedes pneumonia. Decreased clearance can result from immune dysfunction, sputum defects like in CF, architectural distortion as in bronchiectasis and fibrosis, and cough defects or physical barriers as in lung transplantation. Increased local replication may result from augmented nutrients due to mucous or edema fluid, as well as from the increase in steady-state bacterial populations caused by greater entry or deficient clearance. Lung dysbiosis is now described in many pulmonary diseases, but the fundamental unknown is whether dysbiosis initiates or perpetuates inflammation and/or injury, or whether it is a consequence and marker of disease. This section highlights insights in selected lung diseases not traditionally considered microbially driven; advances in other conditions (CF, lung transplantation, ARDS) have been addressed in recent disease-specific reviews (82–84).

Chronic obstructive pulmonary disease. Chronic obstructive pulmonary disease (COPD) affects over 300 million people globally and is characterized by progressive irreversible airflow limitation associated with chronic bronchitis, small airway occlusion, inflammation, and emphysematous destruction of parenchyma (85–87). It is associated primarily with smoking, though additional factors are indoor smoke exposure, pollution, early-life respiratory infections, and genetics (87). Bacteria can be isolated by culture from sputum and BALF during exacerbations (88–90), and prophylactic macrolide antibiotics reduce exacerbations in advanced disease (91), implicating microbial drivers.

Studies applying sequence-based methods in COPD differ in subject and specimen type, and describe dysbiosis during both periods of stable COPD and acute exacerbations. During stable disease, relative abundance of Pseudomonas was higher and of Prevotella lower in bronchial wash in comparison with healthy smokers and nonsmokers (92). A large meta-analysis of sputum data sets reported enrichment of Haemophilus, Streptococcus, Moraxella, and Lactobacillus, and depletion of several genera, relative to healthy controls (93). In a cross-sectional study of 72 COPD patients, airflow limitation correlated with increased relative abundance in sputum of Pseudomonas and decreased Treponema (94). A prospective study of 101 patients found that greater disease severity was associated with increased Haemophilus and decreased Prevotella and Veillonella (95). Microbial profiles have also been associated with bronchodilator responsiveness and peak expiratory flow rate in the SPIROMICS cohort (96). Diversity is lower in stable COPD compared with healthy smokers and nonsmokers (92, 97) and is further decreased during acute exacerbations (98). Lower diversity of LRT specimens also correlated with decreased lung function and exacerbation frequency in stable COPD (5, 12, 99).

In 200 COPD patients followed prospectively, acute exac-
erations were correlated strongly with viral infections and only modestly with bacterial microbiota, but exacerbation frequency was linked with higher relative abundances of several taxa (Bacillus, Escherichia, Lactobacillus, Moraxella, and Staphylococcus) as well as greater longitudinal variability (100). Sputum bacterial communities during exacerbations have also predicted outcome; the presence of Staphylococcus and the absence of Veillonella, along with decreased diversity, were associated with higher mortality in a cohort of 102 subjects (101). Similarly, Haemophilus- and Streptococcus-dominated sputa have been associated with greater mortality in a cross-sectional study of 252 subjects (102).

However, findings are not always consistent across studies. As highlighted in a recent study of 122 patients, interindividual differences may outweigh group differences, and the complexity of lung microbiota presents substantial challenges for use as a biomarker or for simple mechanistic interpretation (103).

Distinct lung microbiota have also been described across COPD inflammatory endotypes. An analysis of 1706 sputa from 510 subjects reported that neutrophilic inflammation was associated with two major community classes during both stability and exacerbations, one characterized by an URT-like bacterial community and one characterized by pathogens including Haemophilus, Moraxella, and Streptococcus (104). Neutrophilia correlated with Moraxella abundance in sputa in another study of 87 COPD patients (105) and with Haemophilus and Neisseria in a 43-subject cohort (97). Moreover, Haemophilus-predominated neutrophilic inflammation correlated with increased sputum IL-1β and TNF-α (104, 106) and with neutrophil extracellular trap formation (107). Neutrophilic inflammation also associated with decreased diversity relative to other inflammatory endotypes (108). These findings suggest that a pathogen-predominant community type may drive neutrophilic inflammation in COPD. In contrast, eosinophilic inflammation in COPD is associated with fewer clinical bacterial infections, lower bacterial burden, increased diversity, and lower prevalence of Haemophilus and Moraxella (94, 102, 104, 105, 108, 109). Compared with patients who had neutrophilic inflammation, patients with eosinophilic inflammation demonstrated relative enrichment in Gemella, Campylobacter, Capnocytophaga, and Granulicatella in an analysis of three cohorts (104), and a higher proportion of Bacteroidetes in a cohort of 73 COPD and 32 asthma patients (109). Thus, distinct mechanisms likely link the lung microbiome to inflammation in COPD endotypes.

Finally, provocative data implicate changes in metabolic capabilities of the lung microbiome during exacerbations using both inferred metagenomics (bacterial gene content predicted from taxonomic data) and shotgun metagenomic sequencing (16, 93). Inferred metagenomics (bacterial gene content predicted from taxonomic data) and shotgun metagenomic sequencing (16, 93). This common inflammation-microbiome association is consistent with either a particular microbiome type driving mucosal inflammation or, conversely, microbiome profiles reflecting a response to inflammation (118). Finally, beyond the scope of this Review, abundant evidence supports a key role for the gut microbiome in aberrant immune responses in asthma (119).

**Lung cancer.** Lung cancer accounts for almost one-quarter of all cancer deaths (120), with 82% attributed to cigarette smoking (121). Among never-smokers, lung cancer is associated with diseases characterized by chronic airway inflammation, including COPD, Chlamydia pneumoniae, Mycobacterium tuberculosis, and respiratory viruses (122–124), suggesting that the lung microbiome may play a role in cancer development and progression. Supporting this theory, germ-free rodents develop lung cancer less frequently (125, 126). Interest in the airway microbiome in lung cancer has also been spurred by advances in understanding the gut microbiome and colon cancer; it is now recognized that microbial composition and ecological patterns differ in patients with cancer, local mucosal biofilms are associated with genetic and epigenetic abnormalities, and animal models suggest that specific bacteria play key roles in the molecular process of colorectal carcinogenesis (127, 128).

Several reports have associated lung cancer with enrichment of specific taxa in sputum and BAL, including Granulicatella, Abiotrophia, Streptococcus, Veillonella, Megasphaera, and Selenomonas (17, 129, 130). Altered lung microbiome diversity in cancer specimens relative to nonmalignant matched tissue and cancer-free patient controls has also been described, although without uniformity in the directionality of that association (39, 131, 132). Unique bacterial profiles have also been associated with recurrence- and disease-free survival (39, 133), and with specific mutations in squamous cell carcinoma (131). Compositional differences in lower airway microbiota have also been seen in lower airway brushings between small cell and non–small cell lung cancer (134). Finally, altered bacterial diversity in lung tissue associates with advanced cancer stage, as well as with air pollution and pack-years of tobacco smoking, which are lung cancer risk factors (38).

Beyond associations, the questions of whether and how microbiome dysbiosis contributes to lung cancer pathogenesis remain unresolved. Enrichment with Streptococcus and Veillonella in airway brushings of cancer patients was associated with increased inflammatory cell infiltration and bronchiol epithelial upregulation of PI3K/ERK signaling pathways (135), an early
pathogenic event in non-small cell lung carcinoma (136). These taxa also upregulated these signaling pathways in airway epithelial cells in vitro. Exposure of mice to bacterial consortia isolated from a late-stage murine lung tumor accelerated tumor growth, which was mediated by myeloid cell cytokine secretion and activation of IL-17-producing γδ T cells in the lung mucosa (125). Thus, associations exist between lung cancer and lung microbiome communities, and provocative findings suggest the possibility of mechanistic contributions of the microbiome.

**Interstitial lung disease.** The interstitial lung diseases comprise multiple entities, the most common being idiopathic pulmonary fibrosis (IPF). IPF is a progressive fibroproliferative process with poor survival and few effective treatments, characterized by proliferation of mesenchymal cells, collagen and extracellular matrix deposition, fibrosis, and distortion of pulmonary architecture. Studies report that BALF from IPF patients has higher bacterial load than that from healthy people; furthermore, both total bacterial burden and specific taxa (*Streptococcus* and *Staphylococcus*) associate with more rapid disease progression (25, 26, 137, 138). A mechanistic pathway was suggested by findings that a peptide fragment of *Staphylococcus* was elevated in IPF patients compared with controls, and intrapulmonary instillation of this peptide induced lung epithelial cell injury and features of IPF acute exacerbations in mice (139). Another rodent model employed bleomycin-induced fibrosis in conventional and germ-free mice, and showed that bacterial colonization did not affect fibrosis directly; rather, germ-free mice exhibited altered Th1/Th2 immune responses and increased survival, suggesting that microbiota may impact IPF progression through inflammatory rather than fibrotic pathways (26). However, human studies have not uniformly supported findings of increased bacterial content or have suggested considerable interpatient and intrapatient geographic heterogeneity in lung explants of advanced IPF (140, 141). Also, a recent interventional trial found no benefit from the antibiotic trimethoprim-sulfamethoxazole, though the impact on microbiota was not assessed (142). Thus, it is unclear whether the microbiome has a causal role in lung injury, versus dysbiosis resulting from architectural distortion of IPF.

In one of the few lung virome studies, Anelloviridae were elevated in the BALF of patients with acute exacerbations compared with stable IPF (143), although cause versus consequence of inflammation in IPF exacerbations remains unclear.

**From microbiome association to function**

A principal pathway of gut microbiome influence on the host is through small-molecule metabolites, including short-chain fatty acids (SCFAs) (144). Many of these molecules are not produced by human metabolism and are thus markers of bacterial activity as well as potential mediators. SCFAs can directly modulate cell proliferation and differentiation, hormone secretion, metabolic homeostasis, and immune and inflammatory responses, and gut bacteria–derived butyrate even serves as a primary energy source for colon epithelial cells (145–148). Studies have begun to explore the role of SCFAs in microbiome-host crosstalk in the lung and relevant model systems. In vitro, butyrate impairs reactive oxygen species production, phagocytosis, and microbial killing by neutrophils (149). The presence of butyrate during differentiation of alveolar macrophages can induce cellular reprogramming and long-lasting antimicrobial activity (147). Similarly, the combination of propionate or butyrate with TNF-α synergistically increases IL-6 and CXCL8 production by primary human lung fibroblasts and airway smooth muscle cells, while acetate elicits a synergistic response to CXCL8 in airway smooth muscle cells (150). Thus, SCFAs may act as bioactive mediators of the lung microbiome, although relationships may be complex, as different types of lower airway cells respond differently to SCFAs (151).

In the sputum and BALF of CF patients, SCFAs are present in micromolar concentrations (145, 151). Concentrations of acetate correlate with IL-8 in CF BALF, and in vitro experiments suggest that SCFAs may mediate this increase in IL-8 via G protein–coupled receptors on airway epithelial cells (145). Similarly, the inferred metagenome of bronchial microbiota from people with asthma, sampled by protected bronchial brushings, was enriched in genes for pathways involved in the metabolism of SCFAs and amino acids (36). In contrast, inferred metagenomics of sputum indicated that butyrate production pathways were depleted in COPD (12).

Azithromycin is a broad-spectrum antibacterial that is used to prevent exacerbations or delay progression in lung diseases not traditionally considered microbe-driven, including COPD, asthma, and chronic lung allograft dysfunction following transplantation (152, 153). Recently, several studies investigated whether azithromycin might act by modifying lung microbiome communities, but results on community composition have been inconclusive (154–156). Alternatively, the antiinflammatory properties of azithromycin have been linked to modulation of bacterial metabolism, wherein production of glycolic acid and indole-3-acetate was increased following azithromycin treatment. These SCFAs reduced ex vivo secretion of TNF-α, IL-12, and IL-13, and CXCL1 by alveolar macrophages, and these cytokines were also reduced in BALF following azithromycin treatment of patients (157). Consequently, therapeutic interventions might modulate the composition of the lung microbiome or target its metabolic activity. Alternatively, SCFAs or other small-molecule microbially derived mediators might themselves be directly administered therapeutics, as is beginning to be investigated in the gut (158).

Notably, the microbiome as profiled taxonomically through 16S rRNA gene or shotgun metagenomic sequencing does not fully reflect the population of live, metabolically active organisms. Recent studies show that the microbial metagenome inferred from 16S rRNA gene taxonomy or assessed by shotgun sequencing differs from that derived from direct metatranscriptome analysis, which better correlated with bacterial metabolites (13). Animal studies suggest that this discrepancy reflects the mix of metabolically active (live) bacteria and dead (relic) bacteria caught in a snapshot of the dynamic entry/replication/clearance process. Of course, even nonviable bacteria may impact the lower respiratory tract through pattern recognition receptors that recognize the physical constituents of the organisms.

Finally, while this Review focuses on the respiratory tract microbiome, the gut microbiome likely also shapes the lung immune environment. Just as the gut microbiome plays a central role in the development and regulation of systemic immunity,
emerging data suggest that a gut-lung axis also exists (159).

Knowledge gaps and future directions

The lung bacterial microbiome is altered in many diseases that are not traditionally considered infectious or microbial in etiology. However, the lung microbiome field is relatively young and requires more longitudinal, human intervention, and mechanistic animal studies to address causality. Another knowledge gap involves the lung mycobiome and virome, which are greatly understudied. Robust animal models will be needed to move the field forward and address the key question of directionality: Is the microbiome a cause (perpetuator) or a consequence (marker) of disease, or both? Can microbiome-targeted sequence-based technologies improve discrimination between health and disease or predict disease trajectory? Do environmental or other factors known to cause lung disease influence the lung microbiome, and is that a component of their mechanism of injury? What are the specific mechanisms by which the microbiome affects the host? Finally, this knowledge will provide groundwork for the most compelling questions: Can the microbiome be therapeutically manipulated in lung diseases, either by supplementation of communities with beneficial taxa; by selective reduction of proinflammatory or deleterious taxa or modification of their activities; or by direct use of microbially derived small molecules that have beneficial host effects?

Acknowledgments

This work was supported by NIH grants R33-HL137063 (to RGC) and KL2-TR001879 (to JEM), and a grant from the Cystic Fibrosis Foundation (to RGC and SAW).

Address correspondence to: Ronald G. Collman, 522 Johnson Pavilion, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA. Phone: 215.898.0913; Email: collmanr@pennmedicine.upenn.edu.

REVIEW

The Journal of Clinical Investigation

8


