Lung-enriched Organisms and Aberrant Bacterial and Fungal Respiratory Microbiota after Lung Transplantation

Emily S. Charlson1,2, Joshua M. Diamond1, Kyle Bittinger2, Ayannah S. Fitzgerald1, Anjana Yadav1, Andrew R. Haas1, Frederic D. Bushman2*, and Ronald G. Collman1,2*

1Pulmonary, Allergy and Critical Care Division, Department of Medicine, and 2Department of Microbiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania

Rationale: Long-term survival after lung transplantation is limited by infectious complications and by bronchiolitis obliterans syndrome (BOS), a form of chronic rejection linked in part to microbial triggers. Objectives: To define microbial populations in the respiratory tract of transplant patients comprehensively using unbiased high-density sequencing.

Methods: Lung was sampled by bronchoalveolar lavage (BAL) and upper respiratory tract by oropharyngeal wash (OW). Bacterial 16S rDNA and fungal internal transcribed spacer sequencing was used to profile organisms present. Outlier analysis plots defining taxa enriched in lung relative to OW were used to identify bacteria enriched in lung against a background of oropharyngeal carryover.

Measurements and Main Results: Lung transplant recipients had higher bacterial burden in BAL than control subjects, frequent appearance of dominant organisms, greater distance between communities in BAL and OW indicating more distinct populations, and decreased respiratory tract microbial richness and diversity. Fungal populations were typically dominated by Candida in both sites or by Aspergillus in BAL but not OW. 16S outlier analysis identified lung-enriched taxa indicating bacteria replicating in the lower respiratory tract. In some cases this confirmed respiratory cultures but in others revealed enrichment by anaerobic organisms or mixed outgrowth of upper respiratory flora and provided quantitative data on relative abundances of bacteria found by culture.

Conclusions: Respiratory tract microbial communities in lung transplant recipients differ in structure and composition from healthy subjects. Outlier analysis can identify specific bacteria replicating in lung. These findings provide novel approaches to address the relationship between microbial communities and transplant outcome and aid in assessing lung infections.

Keywords: microbiome; 16S; ITS; bronchiolitis obliterans syndrome

The outcome for lung transplantation is much worse than other solid organ transplants, with only 53 and 30% survival at 5 and 10 years, respectively (1). Infection and primary graft dysfunction are the principal causes of death early after transplantation, whereas the major causes of later death are infection and bronchiolitis obliterans syndrome (BOS) (1). Lung shares with all other organ transplants an increased susceptibility to infection resulting from immunosuppressive therapy, but additional factors make it uniquely susceptible. These include direct communication with the external environment and microbe-rich upper respiratory tract (URT), defective lower respiratory tract (LRT) mechanical defense due to vagal denervation cough reflex defects and abnormal mucociliary clearance at anastomotic sites, and a high propensity for microaspiration (2, 3). Microbial triggers are also linked to BOS, the clinical manifestation of chronic rejection, based on epidemiological association with specific bacteria, fungi, or viruses detected in lung (4–8), linkage with genetic polymorphisms in innate microbial sensors (9–12), and responsiveness to azithromycin (13–15).

Traditional respiratory culture methods focus on individual organisms rather than whole populations, are only semiquantitative, and typically exclude organisms normally present in the upper respiratory tract due to the confounding issue of URT admixture in lung-derived specimens. Culture-independent molecular
approaches can identify entire microbial populations present and can be highly quantitative. These methods are starting to be applied to the respiratory tract (16–19). Despite the unique susceptibility of lung transplants to microbial invasion and its importance in outcome, the lung transplant microbiome has not been studied by deep sequencing.

We used bacterial 16S rRNA gene and fungal internal transcribed spacer (ITS) sequencing to define the microbial populations in lung allografts from transplant recipients sampled by bronchoalveolar lavage (BAL). Because a principal challenge to LRT sampling by bronchoscopy is instrument carryover from the upper tract, lung microbiota in BAL were compared with URT populations sampled by oropharyngeal wash (OW). We report here that the transplant recipients’ lung microbial communities were notably different from those found in healthy control subjects in composition and structure, were less closely related to matched URT samples, and showed frequent lung enrichment of specific bacteria. These methods help to define allograft colonization, may aid in the diagnosis of lung infections, and can provide insight into microbial contributors to BOS.

METHODS

Subjects and Sample Collection

Lung transplant recipients undergoing clinical bronchoscopy provided a 10-ml saline oropharyngeal wash/gargle (OW) (20) followed by transoral bronchoscopy using standard clinical procedures. BAL was performed using 60 to 120 ml of saline with the bronchoscope wedged in the right middle lobe for bilateral or single-right transplants or in the left lingula for single-left transplants. BAL fluid was submitted for bacterial and fungal respiratory culture, and an aliquot was used for microbiome evaluation. To control for microbial DNA that might be present in instruments or reagents, 10 ml of saline was washed through the bronchoscope immediately before each procedure.

Nontransplant control subjects included healthy volunteers who underwent research bronchoscopy using a two-scope procedure previously reported (20). In that protocol, one bronchoscope was used to anesthetize the glottis and provide URT samples, and a second clean scope was used to sample the LRT, to minimize and precisely define potential carryover from the URT by a bronchoscope entering the LRT. Their BAL first return was used to best approximate clinical (single-scope) conditions (20). Another group of healthy volunteers underwent research bronchoscopy using a standard single-scope procedure identical to clinical bronchoscopy. Samples were also collected from two nontransplant subjects undergoing diagnostic bronchoscopy for evaluation of sarcoidosis and a lung nodule found to be adenocarcinoma.

All subjects gave written informed consent under Investigational Review Board–approved protocols.

DNA Extraction and 16S rDNA Quantification

For 16S analysis, genomic DNA was isolated from unfractionated OW and BAL using the PowerSoil DNA isolation kit (MoBio, Carlsbad, CA) as previously described (20). For ITS analysis, samples were additionally incubated at 95°C for 10 minutes before beadbeating. Bacterial 16S copy number was quantified by real-time PCR (21).

Bacterial 16S rDNA and Fungal ITS Gene Amplification and Sequence Analysis

16S rDNA genes were amplified using barcoded V1V2 primers, and amplicons were purified, pooled, and 454-pyrosequenced as described (22). Using the QIIME pipeline (23), reads were denoised, clustered into 97% similar operational taxonomic units (OTUs), aligned to full-length 16S rDNA sequences, assigned taxonomy with RDP Classifier, and used to generate de novo phylogenetic trees with FastTree2 (24). Sequences were removed if unalignable or identified as chimeras (25). ITS genes were amplified using barcoded ITS1F/ITS2 primers (26). Amplicons were purified, pooled, and 454-pyrosequenced, and sequences were denoised, clustered into OTUs at 95.2% sequence identity, and assigned taxonomy with the BROCC software pipeline as described (S. Dowlive, unpublished observations). Additional details are provided in the online supplement.

Statistical Analysis

UniFrac (27) was used to measure β diversity as described by Losu-pone and colleagues (28) and visualized using principal coordinate analysis of the distance matrix. We tested for differences in UniFrac distance, within versus between groups, by comparing the T statistic against 10,000 random permutations (28). For Procrustes analysis, the goodness of fit (M² value) was measured by summing over the residuals, and significance was assessed by the Monte-Carlo label permutation method (29, 30). The Wilcoxon rank-sum test was used to compare the number of OTUs in healthy and transplant samples, rarefied to 3,000 sequence reads per sample. OTUs enriched in BAL versus OW were identified by computing the maximum likelihood parameters for a Dirichlet-multinomial distribution of OTU counts and then using the marginal form of this distribution to perform a one-sided test of OTU counts in the BAL sample (31, 32) (E.S. Charlson, unpublished observations).

RESULTS

Lower Respiratory Tract Sampling from Lung Transplant Recipients

Paired BAL and OW were obtained from 21 lung transplant subjects (Tables 1 and 2 and Table E1 in the online supplement). Seven subjects were transplanted for COPD or emphysema due to α-1 antitrypsin deficiency. Seven subjects were transplanted for interstitial diseases including IPF, collagen-vascular associated interstitial lung disease (ILD), or sarcoidosis. One subject was transplanted for congenital heart disease with pulmonary hypertension, and six subjects were transplanted for suppulsive lung diseases, mainly cystic fibrosis (CF). Most had bilateral transplants (81%), and two with severe pulmonary hypertension also received heart transplants. All bronchoecopies were routine surveillance procedures performed at regularly scheduled intervals during the first 15 months after transplant, except for Tx33, who had nonresolving clinical pneumonia despite treatment for culture-positive Pseudomonas aeruginosa, and Tx38, who was evaluated for declining FEV₁ 9.5 years after transplantation (late BOS). Two subjects also had early BOS identified by decreased FEV₁ (Tx25, Tx40). All subjects were on immunosuppressives and antimicrobials typical of posttransplant management (Table 1). Nontransplant control subjects included healthy volunteers who underwent research bronchoscopy using procedures identical to clinical bronchoscopy or using a previously described two-bronchoscope procedure (20) and two subjects without infection or immunosuppression undergoing clinical diagnostic bronchoscopy.

Bacterial Quantification in Lung Transplant Recipient Airways

Bronchoscope prewashes and control samples demonstrated levels of 16S DNA close to or below the lower limit of qPCR quantification (LOQ: 725 copies/ml; data not shown) as we previously described (20). OW showed high numbers of 16S copies, and values were similar between subject groups (Figure 1A). In contrast, BAL of lung transplant subjects had a mean copy number that was 44-fold higher than control subjects sampled by single-scope bronchoscopy and 91-fold higher than BAL first return from healthy subjects sampled with a two-scope protocol (Figure 1B). Bacterial 16S DNA levels in BAL did not differ based on underlying indication for transplant whether grouped by suppurrative versus nonsuppurrative or among the nonsuppurrative subgroups (Figures 1C and 1D).
### TABLE 1. CLINICAL FEATURES AND MICROBIOLOGICAL FINDINGS ASSOCIATED WITH 23 BAL SAMPLES FROM 21 LUNG TRANSPLANT RECIPIENTS STUDIED*

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sample</th>
<th>Pretransplant Disease</th>
<th>Transplant Type</th>
<th>Months after Transplant</th>
<th>Bronch indication/status</th>
<th>Immunosuppression</th>
<th>Antimicrobials</th>
<th>BAL Bacterial Culture</th>
<th>BAL Fungal Culture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tx24</td>
<td>24</td>
<td>IPF</td>
<td>S-L</td>
<td>2</td>
<td>Routine</td>
<td>MMF, Pred, Tac</td>
<td>TMP/S, Valc, Nystatin</td>
<td>Staphylococcus aureus</td>
<td>None</td>
</tr>
<tr>
<td>Tx25</td>
<td>25</td>
<td>IPF</td>
<td>S-L</td>
<td>6</td>
<td>Routine/early BOS</td>
<td>MMF, Pred, Tac</td>
<td>Azithro, TMP/S, Valgan, Nystatin</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Tx26</td>
<td>26</td>
<td>CF</td>
<td>B</td>
<td>9</td>
<td>Routine</td>
<td>Pred, Tac</td>
<td>TMP/S, Valc</td>
<td>Pseudomonas aeruginosa</td>
<td>None</td>
</tr>
<tr>
<td>Tx30</td>
<td>30</td>
<td>IPF</td>
<td>S-L</td>
<td>15</td>
<td>Routine</td>
<td>Aza, Pred, Tac</td>
<td>TMP/S, Valgan</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Tx31</td>
<td>31</td>
<td>CHD, PHTN</td>
<td>B-H</td>
<td>5</td>
<td>Routine</td>
<td>MMF, Pred, Tac</td>
<td>TMP/S, Valgan</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Tx32</td>
<td>32</td>
<td>CF</td>
<td>B</td>
<td>2</td>
<td>Routine</td>
<td>Aza, Pred, Tac</td>
<td>TMP/S, Valgan</td>
<td>Achromobacter, P. aeruginosa</td>
<td>None</td>
</tr>
<tr>
<td>Tx33</td>
<td>33</td>
<td>CF</td>
<td>B</td>
<td>9</td>
<td>Nonresolving</td>
<td>Aza, Pred, Tac</td>
<td>Levofloxacin, Meropenem, Valcys, Tobra</td>
<td>P. aeruginosa</td>
<td>None</td>
</tr>
<tr>
<td>Tx34</td>
<td>34</td>
<td>A1AT</td>
<td>B</td>
<td>4</td>
<td>Routine</td>
<td>MMF, Pred, Tac</td>
<td>TMP/S, Valgan, Vori</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Tx35</td>
<td>35</td>
<td>PCD</td>
<td>B</td>
<td>12</td>
<td>Routine</td>
<td>MMF, Pred, Tac</td>
<td>TMP/S</td>
<td>Mycobacterium avium-intraacellular</td>
<td>None</td>
</tr>
<tr>
<td>Tx36</td>
<td>36A</td>
<td>IPF</td>
<td>S-L</td>
<td>2</td>
<td>Routine</td>
<td>MMF, Pred, Tac</td>
<td>TMP/S, Valgan, Vori</td>
<td>Aspergillus flavus, C. albicans</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>36B</td>
<td>IPF</td>
<td></td>
<td>4</td>
<td>Routine</td>
<td>MMF, Pred, Tac</td>
<td>TMP/S, Valgan, Vori</td>
<td>Klebsiella oxytoca</td>
<td>None</td>
</tr>
<tr>
<td>Tx38</td>
<td>38</td>
<td>CF</td>
<td>B</td>
<td>114</td>
<td>Late BOS</td>
<td>Pred, Tac</td>
<td>Azithro, TMP/S, Ciprofloxacin</td>
<td>None</td>
<td>C. albicans</td>
</tr>
<tr>
<td>Tx39</td>
<td>39A</td>
<td>COPD</td>
<td>B</td>
<td>2</td>
<td>Routine</td>
<td>Aza, Pred, Tac</td>
<td>TMP/S, Valgan</td>
<td>None</td>
<td>C. albicans</td>
</tr>
<tr>
<td>39B</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>Routine</td>
<td>Aza, Pred, Tac</td>
<td>TMP/S, Valgan</td>
<td>None</td>
<td>C. albicans</td>
</tr>
<tr>
<td>Tx40</td>
<td>40</td>
<td>CF</td>
<td>B</td>
<td>18</td>
<td>Routine/early BOS</td>
<td>MMF, CSA, Pred pulse Methylpred</td>
<td>Azithro, TMP/S</td>
<td>None</td>
<td>P. aeruginosa</td>
</tr>
<tr>
<td>Tx41</td>
<td>41</td>
<td>Sarc, PHTN</td>
<td>B-H</td>
<td>1</td>
<td>Routine</td>
<td>MMF, Pred, Tac</td>
<td>TMP/S, Valc, Nystatin</td>
<td>S. aureus</td>
<td>C. albicans</td>
</tr>
<tr>
<td>Tx42</td>
<td>42</td>
<td>COPD</td>
<td>S-L</td>
<td>3</td>
<td>Routine</td>
<td>Aza, Pred, Tac</td>
<td>TMP/S, Valgan</td>
<td>None</td>
<td>C. albicans</td>
</tr>
<tr>
<td>Tx43</td>
<td>43</td>
<td>COPD</td>
<td>B</td>
<td>5</td>
<td>Routine</td>
<td>Aza, Pred, Tac</td>
<td>TMP/S, Valgan</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Tx44</td>
<td>44</td>
<td>Sarc, PHTN</td>
<td>B-H</td>
<td>12</td>
<td>Routine</td>
<td>MMF, Pred, Tac</td>
<td>TMP/S</td>
<td>Burkholderia cepacia, Aspergillus fumigatus, Paecilomyces Ilacinus</td>
<td>None</td>
</tr>
<tr>
<td>Tx45</td>
<td>45</td>
<td>ILD</td>
<td>B</td>
<td>4</td>
<td>Routine</td>
<td>MMF, Pred, Tac</td>
<td>TMP/S, Nystatin</td>
<td>M. avium-intraacellular</td>
<td>C. albicans</td>
</tr>
<tr>
<td>Tx46</td>
<td>46</td>
<td>COPD</td>
<td>B</td>
<td>4</td>
<td>Routine</td>
<td>Aza, Pred, Tac</td>
<td>Valgan</td>
<td>P. aeruginosa</td>
<td>None</td>
</tr>
<tr>
<td>Tx47</td>
<td>47</td>
<td>COPD</td>
<td>B</td>
<td>2</td>
<td>Routine</td>
<td>Aza, Pred, Tac</td>
<td>TMP/S, Valgan, Nystatin</td>
<td>Enterobacter cloacae</td>
<td>C. albicans</td>
</tr>
<tr>
<td>Tx49</td>
<td>49</td>
<td>ILD</td>
<td>B</td>
<td>2</td>
<td>Routine</td>
<td>Aza, Pred, Tac</td>
<td>TMP/S, Valgan</td>
<td>None</td>
<td>C. albicans</td>
</tr>
</tbody>
</table>

* BAL samples were obtained at the indicated times after transplant for routine postsurveillance bronchoscopy, except for Tx33 and Tx38 as noted.

---

Characterization of Bacterial Lineages in the Transplanted Lung by 454 Pyrosequencing

We surveyed bacterial lineages using deep sequencing of 16S rRNA genes (Figure 2). Bronchoscope prewash and control specimens, which had low copy numbers, showed bacterial lineages typical of water and skin organisms (see Figure E1 in the online supplement) as previously described (20).

Oropharyngeal samples contained high abundances of taxa characteristic of the oral cavity, including Streptococcus, Prevotella, Veillonella, Porphyromonas, Neisseria, and Rothia (22, 33, 34). Most transplant recipients had similar oropharyngeal flora as healthy subjects, except for four subjects whose OW were dominated by Streptococcus spp.

BAL from healthy subjects shared a similar overall bacterial profile as OW samples, as we previously reported (20). Transplant subjects’ BAL largely shared bacterial profiles with OW samples, but a subset was clearly dominated by specific lineages that were absent in OW or detected at much lower abundances. Two BAL samples were dominated by P. aeruginosa (B26 and B33; both transplanted for CF) and revealed P. aeruginosa on culture. A high abundance of Staphylococcus aureus was seen in one BAL (B24; transplanted for IPF), also detected by culture. Similarly, one BAL showed dominant Achromobacter (B32; transplanted for CF), also identified by culture. In contrast, BAL from the subject with late-onset BOS (B38) was dominated by the unusual anaerobic lineage Sneathia (35), which was not identified by culture. Three BAL samples showed high levels of Streptococcus spp. (B25, B39B, B42; transplanted for COPD or α-1 antitrypsin deficiency) without any corresponding cultured organism, but Streptococcus spp. were also found at high or moderate abundance in OW from these subjects.

Characterization of Fungal Lineages in the Transplanted Lung and Matched OW

Fungi were analyzed by ITS gene sequencing (Figure 3; Figure E2). Although precise qPCR quantification is precluded by high length variation of the ITS1 gene, we used the amount of ITS PCR product as an estimate for the relative abundance of fungi (at top of the heat map).

OW of healthy subjects revealed varying ITS amplification levels and a distribution similar to that recently reported in oral samples (26), including environmental agents as well as Candida and Aspergillus (Figure 3). One healthy subject had relatively high ITS amplification in OW (3B08), comprised mainly of Candida albicans. OW of transplant subjects revealed some samples with scant fungal sequences and others with robust ITS amplification. Of the transplant subjects with high OW fungal amplification, all but one were dominated by Candida (Figure 3). Some transplant subjects with intermediate or low levels of OW fungal sequences had high relative abundances of...
Healthy volunteers’ BAL yielded scant fungal amplification. Although *Candida* was identified in several healthy subjects’ OW, it was absent in their BAL. The few fungal reads present in healthy BAL were comprised largely of environmental agents such as *Davidiellaceae* and *Cladosporium*, also present in pre-wash control samples (Figure E2), as well as low abundances of *Aspergillus*.

BAL from lung transplant recipients showed a markedly different pattern. Four yielded high levels of ITS amplification; three of these were dominated by *Candida* (B38, B47, and B49), and one (B44) was dominated by *Aspergillus*. In each case, the corresponding fungus was found by culture. Several subjects with intermediate ITS levels revealed high abundance of *Candida* (B41, B25, B26, B39, B39B, and B42), and one was dominated by *Aspergillus* (B36, also found by culture). When *Candida* was seen in BAL, it was also always found at high abundance in OW, whereas for subjects with a high abundance of *Aspergillus* in BAL, this fungus was absent or at markedly lower abundance in OW. *Cryptococcus* was present in six transplant patients’ BAL but at relatively low abundance or in samples with low total ITS read numbers. *Cryptococcus* was not seen in BAL from the transplant subject who had this fungus at high abundance in OW (Tx46).

**Community Richness, Completeness of Microbial Identification, and Species Diversity**

To determine whether transplant and healthy subjects’ microbial communities differed in species richness, we calculated the number of observed OTUs at various sequencing depths for bacteria and fungi in OW and BAL specimens. Rarefaction analysis of bacterial oropharyngeal OTUs (Figure 4A) indicated that richness was lower in transplant than healthy OW (*P* < 0.05, Wilcoxon rank sum) and that sampling was near saturation in transplant subjects but not in healthy subjects. Richness was lowest in the subjects transplanted for suppurative lung disease and intermediate in those transplanted for nonsuppurative disease.

**TABLE 2. TRANSPLANT SUBGROUPS USED FOR ANALYSIS**

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonsuppurative (n = 15)</td>
<td>Age, yr (mean ± SD) 57.5 ± 7.2, Single/bilateral 4/11, Heart 2</td>
</tr>
<tr>
<td>Suppurative (n = 6)</td>
<td>Age, yr (mean ± SD) 35.0 ± 11.3, Single/bilateral 0/6, Heart 0</td>
</tr>
<tr>
<td>ILD/IPF (n = 7)</td>
<td>IPF (n) 3, Other ILD (n) 4, Age, yr (mean ± SD) 60.7 ± 6.2, Single/bilateral 3/4, Heart (n) 1</td>
</tr>
<tr>
<td>COPD (n)</td>
<td>COPD (n) 5</td>
</tr>
<tr>
<td>A1AT deficiency (n)</td>
<td>A1AT deficiency (n) 2</td>
</tr>
<tr>
<td>Age, yr (mean ± SD)</td>
<td>54.3 ± 7.0</td>
</tr>
<tr>
<td>Single/bilateral</td>
<td>1/6</td>
</tr>
<tr>
<td>Heart (n)</td>
<td>1</td>
</tr>
<tr>
<td>Suppurative (n = 6)</td>
<td>Cystic fibrosis (n) 5, Primary ciliary dyskinesia (n) 1, Age, yr (mean ± SD) 35.0 ± 11.3, Single/bilateral 0/6, Heart (n) 0</td>
</tr>
</tbody>
</table>

Definition of abbreviations: A1AT = α-1 antitrypsin deficiency; COPD = chronic obstructive pulmonary disease; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis. The top set of data shows group characteristics of subjects transplanted for nonsuppurative and suppurative indications. The lower data set shows underlying disease for the major subgroups of transplant subjects. One subject with congenital heart disease and pulmonary hypertension is included in the nonsuppurative group but not included in further subgroup analysis.

*Candida*, whereas others revealed no dominant lineage or environmental fungi. One transplant OW with an intermediate level of ITS amplification was dominated by *Cryptococcus* (OW46).
but was not different between the COPD/emphysema and ILD/IPF groups. Oropharyngeal fungal microbiota from healthy and transplant subjects showed flat rarefaction curves at high sequencing depth, indicating that further sequencing would yield few additional fungal OTUs (Figure 4B), although transplant patients had fewer fungal lineages identified than healthy subjects ($P < 0.05$, Wilcoxon rank sum). We also saw higher richness in BAL of healthy subjects compared with transplant subjects (data not shown), although URT carryover contributing to BAL precludes definitive estimates of richness in lung.

We then assessed bacterial community diversity using the Shannon Diversity Index (Figure 5). Transplant subjects had lower diversity in OW and BAL than healthy subjects ($P < 0.05$ Wilcoxon rank sum). Subjects transplanted for suppurative lung disease had the lowest community diversity, consistent

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Relative abundance of bacterial taxa derived from pyrosequencing data. Each column represents an individual sample. Sample type and subject group is indicated at the bottom of each group of columns. Each row corresponds to specific bacterial taxon, with their proportional representation in the given sample indicated by the color code at right. Bacterial operational taxonomic units (OTUs) were collected into genera, so some rows represent multiple OTUs. OTUs not assigned at the genera level and genera with less than 1,000 total reads are omitted. Hierarchical clustering groups the rows to emphasize lineages with similar abundance patterns.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Relative abundance of fungal taxa derived from pyrosequencing data. Fungal taxa identified in bronchoalveolar lavage and oropharyngeal wash samples are shown as described for bacterial lineages in Fig. 2. Fungal operational taxonomic units (OTUs) were collected into classes, so some rows represent multiple OTUs. OTUs not assigned at the class level and classes with less than 250 total reads are omitted. Along the top is the concentration of ITS DNA post-PCR for each sample, keyed by gray-scale at top right.}
\end{figure}
with outgrowth of particular bacteria. Other transplant sub-
jects were intermediate between the healthy and suppurative
(P, 0.05, Wilcoxon rank sum).

Global Relationship between Lung and URT Bacterial
Communities in Transplant and Healthy Subjects

We next investigated the relationship of lung and URT commu-
nities, following the hypothesis that due to the growth of microbes in
transplanted lung, BAL would be more different from URT samples
in transplant subjects than in healthy subjects. For this comparison
we used UniFrac, which measures the similarity of bacterial commu-
nities based on the extent to which members of the communities are
phylogenetically related (27). UniFrac distances were calculated for
all pairs of samples, and then relationships among communities
were plotted using principal coordinate analysis (PCoA).

UniFrac analysis showed that each subject’s BAL was more
closely related to their own OW than to the BAL of different
individuals (P < 0.05 for abundance-weighted and unweighted
UniFrac). An additional nonphylogenetic measure of shared
OTUs, the Jaccard Index (36, 37), also revealed that BAL from
each individual was most tightly linked to their own OW (P < 0.05
binary/abundance Jaccard), consistent with BAL communities
originating largely in the URT (20).

We then quantified the resemblance between lung and URT
communities. The BAL and OW weighted UniFrac data were
plotted using PCoA, and then the two PCoA plots were aligned
using Procrustes (29, 30), which rotates and scales two matrices
to maximize overlap (Figure 6). For healthy subjects, BAL and
OW formed a tight cluster, separate from transplant subjects (P <
0.05 UniFrac and Jaccard), with small distances between commu-
nities from the same individual. In contrast, transplant subjects’
BAL and OW microbiota were often divergent and separated by
greater distances. Thus, healthy subjects show close similarity
between their BAL and OW microbiota, whereas distances are
greater for lung transplant subjects, consistent with outgrowth of
bacterial lineages in transplanted lungs.

**Global Relationship between Lung and URT Bacterial
Communities in Transplant and Healthy Subjects**

We next investigated the relationship of lung and URT commu-
nities, following the hypothesis that due to the growth of microbes in
transplanted lung, BAL would be more different from URT samples
in transplant subjects than in healthy subjects. For this comparison
we used UniFrac, which measures the similarity of bacterial commu-
nities based on the extent to which members of the communities are
phylogenetically related (27). UniFrac distances were calculated for
all pairs of samples, and then relationships among communities
were plotted using principal coordinate analysis (PCoA).

UniFrac analysis showed that each subject’s BAL was more
closely related to their own OW than to the BAL of different

![Figure 4. Rarefaction analysis of bacterial and fungal operational taxo-
nomic units (OTUs) in the oropharynx of healthy and lung transplant
subjects. The y axis denotes the number of OTUs detected by pyrosequencing of bacterial 16S genes (A) or fungal internal transcribed
spacer genes (B) at the corresponding sequencing depths shown along
the x axis. The subject group is indicated by the color key on the right. Error bars denote the standard deviation within subject groups.](image)

![Figure 5. Alpha diversity of bacterial communities in healthy and trans-
plant subjects. Diversity was calculated using the Shannon Index for
oropharyngeal wash (A) and bronchoalveolar lavage (B). *P < 0.05;
**P < 0.01 (Wilcoxon rank sum).](image)

![Figure 6. Relationship between bronchoalveolar lavage (BAL) and oro-
pharyngeal bacterial communities within individuals. Weighted UniFrac
distances were calculated between all pairs of samples within BAL or
oropharyngeal wash (OW), and then each sample type was plotted
separately in 3D space by principal coordinate analysis. The two plots
(BAL and OW) were then transformed by Procrustes analysis to achieve
maximum alignment. Each point corresponds to a bacterial commu-
nity, with transplant subjects’ communities shown in blue, healthy sub-
jects’ communities shown in red, and the two communities from each
subject connected by a bar. The orange end of each bar connects to the
OW sample data; the black end connects to the BAL sample data from
the same individual. If BAL and OW plots are similar, then the relative
distance between connected points (residuals) will be small. The overall
similarity is summarized by the M² value, and statistical goodness of fit
is measured by a Monte Carlo label permutation approach (10,000
iterations).](image)
We then determined whether the relationship between URT and BAL communities differed based on underlying indications for transplant. When BAL-OW distances within individuals were compared across groups, there was a trend for greater BAL-OW distances in the suppurative transplant group compared with all nonsuppurative or COPD/emphysema and ILD/IPF groups (Figure E3), suggesting that the greater BAL-OW distance among transplant subjects was driven mainly by the subgroup of subjects transplanted for suppurative diseases.

Comparing BAL and Oropharyngeal Bacterial Lineages to Identify Taxa Enriched in Lung

All bacterial and fungal organisms found by BAL culture were identified via 16S and ITS sequencing, except for *Mycobacterium avium-intracellulare*, which was found by culture in two subjects (Tx35, Tx45) but not by 16S sequencing (Tables 1 and E1). Respiratory pathogens identified by culture were sometimes the most abundant lineage identified by sequencing (B24, B26, B32, B33, B40, B44), but in many cases they were exceeded in abundance by taxa typical of upper respiratory tract flora (B36, B41, B46, B47, B50). Thus, potentially significant pathogens in the lung often represent a minority population against the background of oropharyngeal organisms in BAL. Furthermore, lung infection can result from bacteria that are normal inhabitants of the URT.

Therefore, to identify sequences that represent organisms genuinely replicating in lung, we developed a single-sided outlier test to compare each taxon’s abundance in BAL with its abundance in OW. This strategy assumes that microbes replicating in the LRT will be enriched in BAL compared with OW, whereas URT taxa present in BAL due to passive admixture such as bronchoscopic carryover or aspiration will be represented at a relative abundance no greater than that in OW. The model also accounts for uncertainty due to stochastics of sampling (E.S. Charlson, unpublished observations).

Figure 7 shows six outlier plots selected from among the 21 subjects to exemplify different patterns revealed by this analysis (full set in Figure E4). Each OTU is represented by a dot, with its relative abundance in BAL shown on the y axis and in OW on the x axis, and significant enrichment in BAL indicated by color-coding. 

OTUs enriched in BAL sample

- Not significant
- p < 0.05
- p < 0.01
- p < 0.05 with FDR correction

Figure 7. Identification of lung-enriched bacteria using single-sided outlier plots to compare taxa abundances in bronchoalveolar lavage (BAL) and oropharyngeal wash (OW) samples. Six representative outlier plots are shown (A–F). Each dot represents an operational taxonomic unit (OTU), with its abundance in OW plotted on the x axis and in BAL on the y axis. Taxa found in equal proportions in lung and oropharynx are close to the diagonal line of identity from lower left to upper right, whereas those enriched in BAL are above the diagonal and to the left. Taxa enriched in BAL compared with OW at a level that reaches statistical significance (P < 0.05 after correction for false discovery rate) are highlighted in red.

We then determined whether the relationship between URT and BAL communities differed based on underlying indications for transplant. When BAL-OW distances within individuals were compared across groups, there was a trend for greater BAL-OW distances in the suppurative transplant group compared with all nonsuppurative or COPD/emphysema and ILD/IPF groups (Figure E3), suggesting that the greater BAL-OW distance among transplant subjects was driven mainly by the subgroup of subjects transplanted for suppurative diseases.
TX32 (Figure 7B) grew Achromobacter and P. aeruginosa in BAL culture. The outlier plot showed highly significant enrichment of Achromobacter, whereas Pseudomonas was above the line of identity but did not reach statistical significance. Achromobacter comprised more than 80% of bacteria in BAL and was 17-fold more abundant than Pseudomonas, indicating the relative quantity of each organism in lung.

TX38 (Figure 7C) had late-onset BOS. Outlier analysis revealed high-level BAL enrichment of the anaerobic bacterium Sneathia, which represented more than 50% of all sequences in BAL but less than 0.5% in the oropharynx and indicates robust lung growth. No respiratory pathogens were identified in BAL culture, which would not detect anaerobes such as Sneathia.

TX47 (Figure 7D) yielded Enterobacter cloacae in BAL culture. The outlier plot revealed clustering of many taxa along the line of identity, suggesting abundant oropharyngeal admixture with BAL. However, even in the background of high carryover, two Enterobacteriaceae taxa are significantly enriched in BAL relative to OW, confirming outgrowth in lung.

TX49 (Figure 7E) was negative for respiratory pathogens by culture. However, outlier analysis revealed significant enrichment in BAL of several taxa normally found in the URT (Prevotella, Fusobacterium), suggesting outgrowth of mixed oropharyngeal flora not detected by standard respiratory culture.

TX46 (Figure 7F) BAL culture revealed P. aeruginosa. However, outlier analysis revealed no taxa enriched in lung compared with URT. P. aeruginosa was present in BAL but represented only 0.01% of sequences and had a similar abundance in OW. This result suggests detection by culture in BAL likely from bronchoscopic carryover or passive aspiration without replication and enrichment in the lung.

Additional examples of clinical utility (Figure E4) include robust single-taxa lung enrichment (TX24, TX26, TX44), lung outgrowth of mixed URT/anaerobic bacteria or other organisms not found on culture (TX25, TX41), and lack of outliers indicating absence of significant lung-enriched taxa (TX30, TX31).

Figure E4 shows a concise clinical report developed to summarize molecular microbiome findings, including the outlier analysis, culture, and clinical data for each subject. Statistical and software tools for report generation are embedded in a "reproducible" format, which allows repeated analysis over user-specified data sets (E. S. Charlson, unpublished observations).

**DISCUSSION**

Approximately 24% of lung transplant recipients die from posttransplant infections, most of which are bacterial or fungal pulmonary infections (1, 38, 39). The major cause of later death is BOS (1, 39), a form of chronic rejection that is not well understood but is linked to microbial triggers within the allograft (4–12, 40). Here we applied broad-range 16S bacterial and ITS fungal rRNA gene profiling to define microbial populations in the lungs of 21 transplant recipients. Transplant subjects showed marked differences from healthy subjects in airway microbial burden, community composition and structure, lower richness, and greater distances between lung and upper respiratory communities. We then used a novel approach to visualize taxa enriched in bronchoscopic specimens relative to oropharyngeal samples, allowing identification of specific bacteria replicating in allografts of individual subjects.

The richness and diversity of airway communities were reduced in transplant subjects compared with control subjects. Although richness and diversity was most reduced in subjects with CF; it was also reduced in subjects transplanted for other indications (Figure 5). Decreased airway community richness and diversity in CF has been correlated with deteriorating lung function (41, 42) but has not been investigated in relation to outcome from lung transplant either among subjects with CF or subjects with non-suppurative lung disease.

Although molecular profiling can identify all taxa present and can quantify abundances, the source of those sequences is critical in determining their significance (43). By comparing the relative abundance of each bacterial taxon in BAL and oropharyngeal samples, we can identify enrichment in lung (E.S. Charlson, unpublished observations). Enrichment in BAL compared with the URT suggests replication within the lung and thus can distinguish such organisms from those entering BAL through bronchoscopic carryover or passive aspiration. A large proportion of transplant recipients had specific lung-enriched bacteria. Enriched taxa often confirmed organisms found by culture but also identified bacteria not amenable to culture, such as unusual anaerobes or normal oropharyngeal residents. Molecular profiling also determined proportional abundances of organisms, adding quantitative information to culture results. In some cases we showed that bacteria identified by culture were equally represented in BAL and URT, providing evidence against authentic LRT replication.

Transplant subjects had markedly higher levels of bacterial 16S DNA in BAL than controls subjects. This finding was true regardless of underlying indication for transplant, with no difference between subjects transplanted for COPD/emphysema, interstitial lung disease, or CF suppurative lung disease. Although taxa present in BAL at abundances similar to the URT may result from bronchoscopic carryover, they may also reflect bacteria genuinely present in lung due to aspiration without replication and enrichment. The high 16S DNA levels seen even in subjects without outliers (where outliers reflect a minority of BAL sequences) is consistent with such a source, which may result from increased propensity for aspiration or inability to clear aspirated bacteria effectively due to cough and mucociliary clearance defects (3, 44, 45). Passively aspirated or even dead bacteria can be potent immunological activators, and aspiration has been implicated in BOS (46).

BOS is associated with polymorphisms in genes that regulate innate microbial responses (9–12), specific bacteria, fungi, or viruses in the LRT (4–8) and may be more responsive to the antibiotic azithromycin than to intensified immunosuppression (13–15). It is hypothesized that aberrant microbial populations in the lung, among other factors, trigger inflammatory pathways, resulting in graft injury (40). Our study was not designed to address BOS specifically. However, one subject had late-onset BOS, and 16S sequencing revealed high-level lung enrichment of the anaerobic lineage Sneathia. This organism is a component of the vaginal flora and a rare cause of obstetrical infections (35, 47–49) but has not previously been reported in the respiratory tract and would not be identified by standard culture. Two subjects with early BOS also had taxa enriched in lung that were not detected by culture (TX25, TX40). Longitudinal studies using broad-range profiling are needed to determine whether overall microbial burden or specific lung microbiome features, either early after transplant or at the time of BOS, have a role in late graft failure.

Our study has several limitations. (1) To not "over-call" lung-enriched taxa, we used a 95% confidence threshold (E.S. Charlson, unpublished observations). There may be organisms authentically replicating in the lung that do not reach statistical significance for enrichment, although they are likely to be present at low abundance and/or with modest fold-increase over OW. (2) Although OW is a clinically tractable way to sample the URT, there may be subtle differences in outliers identified when BAL is compared with the most accurate representation of potential carryover, direct peri-glottic specimens obtained with high stringency two-scope sampling (20; E.S. Charlson, unpublished observations). (3)
Although most pathogens found by culture were detected by molecular analysis, two subjects had *M. avium-intracellulare* infection on culture that were not found by 16S sequencing. Although MAI is occasionally isolated as a laboratory contaminant, this was not found in any of the 16S rDNA surveys with the primers used (50). A relatively small number of subjects was included in each transplant group here, which should be expanded to larger numbers of subjects and serial specimens.

In summary, aberrant microbial communities and specific lung-enriched taxa are common in lung after transplantation. Future studies are needed to determine whether particular community structures or individual microorganisms identified by microbiome analysis are associated with late graft failure and BOS and whether broad-range 16S bacterial or ITS fungal profiling added to standard culture improves diagnosis and management of lung colonization and infection in transplant and other lung diseases.

**Author disclosures** are available with the text of this article at [www.atjournals.org](http://www.atjournals.org).

**Acknowledgment.** The authors thank the research subjects who volunteered for this study and the clinicians who assisted with specimen collection; W. Russell and D. Frame for critical study assistance; the Penn Center for AIDS Research (P30-AI045008) for technical support; J. Christie for valuable discussion and insights; and members of the Bushman and Collman laboratories for helpful discussions.

**References**


