Widespread Colonization of the Lung by *Tropheryma whipplei* in HIV Infection

Catherine Lozupone1, Adela Cota-Gomez2, Brent E. Palmer2, Derek J. Linderman2, Emily S. Charleston3, Erica Sodergren4, Makedonka Mitreva4, Sahar Abubucker4, John Martin4, Guohui Yao4, Thomas B. Campbell2, Sonia C. Flores2, Gail Ackerman1, Jesse Stombaugh1, Luke Ursell1, James M. Beck2,5, Jeffrey L. Curtis5, Vincent B. Young5, Susan V. Lynch6, Laurence Huang6, George M. Weinstock4, Kenneth S. Knox7, Homer Twigg8, Alison Morris9, Elodie Ghedin9, Frederic D. Bushman3, Ronald G. Colman3, Rob Knight1,10, and Andrew P. Fontenot2; for the Lung HIV Microbiome Project

1Department of Chemistry and Biochemistry and Biofrontiers Institute, University of Colorado, Boulder, Colorado; 2Department of Medicine, University of Colorado Denver, Aurora, Colorado; 3Departments of Medicine and Microbiology, University of Pennsylvania, Philadelphia, Pennsylvania; 4The Genome Institute, Washington University, St. Louis, Missouri; 5Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan; 6Department of Medicine, University of California San Francisco, San Francisco, California; 7Department of Medicine, University of Arizona, Tucson, Arizona; 8Department of Medicine, Indiana University, Indianapolis, Indiana; 9Departments of Medicine and Computational and Systems Biology, University of Pittsburgh, Pittsburgh, Pennsylvania; and 10Howard Hughes Medical Institute, Boulder, Colorado

**Rationale:** Lung infections caused by opportunistic or virulent pathogens are a principal cause of morbidity and mortality in HIV infection. It is unknown whether HIV infection leads to changes in basal lung microflora, which may contribute to chronic pulmonary complications that increasingly are being recognized in individuals infected with HIV.

**Objectives:** To determine whether the immunodeficiency associated with HIV infection resulted in alteration of the lung microbiota.

**Methods:** We used 16S ribosomal RNA targeted pyrosequencing and shotgun metagenomic sequencing to analyze bacterial gene sequences in bronchoalveolar lavage (BAL) and mouths of 82 HIV-positive and 77 HIV-negative subjects.

**Measurements and Main Results:** Sequences representing *Tropheryma whipplei*, the etiologic agent of Whipple’s disease, were significantly more frequent in BAL of HIV-positive compared with HIV-negative individuals. *T. whipplei* dominated the community (~50% of sequence reads) in 11 HIV-positive subjects, but only 1 HIV-negative individual (13.4 versus 1.3%; *P* = 0.0018). In 30 HIV-positive individuals sampled longitudinally, antiretroviral therapy resulted in a significantly reduced relative abundance of *T. whipplei* in the lung. Shotgun metagenomic sequencing was performed on eight BAL samples dominated by *T. whipplei* 16S ribosomal RNA. Whole genome assembly of pooled reads showed that uncultured lung-derived *T. whipplei* had similar gene content to two isolates obtained from subjects with Whipple’s disease.

**Conclusions:** Asymptomatic subjects with HIV infection have unexpected colonization of the lung by *T. whipplei*, which is reduced by effective antiretroviral therapy and merits further study for a potential pathogenic role in chronic pulmonary complications of HIV infection.

**Keywords:** human; microbiome; metagenome; 16S ribosomal RNA; bronchoalveolar lavage

HIV infection is characterized by impaired innate and adaptive immunity, resulting in an increased frequency of pneumonias caused by pathogenic and opportunistic micro-organisms. Antiretroviral therapy (ART) results in decreased viral replication and a concomitant increase in CD4+ T cells (2, 3). Although immune function does not completely normalize with ART, a decreased incidence of pneumonia and other opportunistic infections is seen. However, increased frequencies of noninfectious respiratory complications, such as chronic obstructive pulmonary disease (COPD), lung cancer, and pulmonary hypertension,
T. whipplei Dominates the Lung Microbiota of an Individual Chronically Infected with HIV and Decreases with ART

To study the relationship between HIV infection and the lung microbiota, we initially sequenced bacterial 16S rRNA from BAL of 5 HIV-positive and 11 HIV-negative individuals recruited in Colorado. Given the high prevalence of oropharyngeal-related phylotypes in BAL and sputum samples (6, 13), we were interested in identifying phylotypes highly enriched in the lung compared with the mouth. The highest enrichment for a single phylotype occurred in the lung microbiota of an ART-naïve male smoker with chronic HIV infection (Table 1). This phylotype, which had a 16S rRNA gene sequence identical to T. whipplei, had a relative abundance of 67–99% in BAL fluid from two different lung sites, regardless of whether unfractionated BAL or the cell-free bacterial pellet was surveyed (Table 1). In contrast, T. whipplei 16S rRNA gene sequences were not found in the oral wash sample, and were only present in low abundance in a posterior pharyngeal swab sample (Table 1). Although T. whipplei is most frequently a gut pathogen, targeting the gastrointestinal tract before translocating to other body sites (14), we did not detect T. whipplei in the stool of this subject (Table 1).

The relative abundance of the T. whipplei phylotype in the lung of this individual significantly decreased after 6.5 months of successful ART (P = 0.027; paired t test; Table 1), associated with a decreased plasma HIV-1 RNA viral load from 136,945 to 40 copies/ml and an increased CD4+ T cell count from 386 to 547 cells/μL. Despite a decrease in relative abundance after ART, T. whipplei sequences remained highly enriched in lung samples compared with the upper respiratory tract (Table 1; see also Figure E1 in the online supplement).

Overall Patterns in Microbial Diversity Are Not Driven by Research Cohort

The dominance of the T. whipplei phylotype in this subject before ART initiation and the decrease with ART, as well as a known association between Whipple’s disease and immune dysfunction (14–21), led us to hypothesize that T. whipplei would be more prevalent in lungs of individuals with HIV infection. Therefore, we examined lung samples collected from 82 HIV-positive and 77 HIV-negative subjects from 8 LHMP cohorts.

Data were collected using similar, but distinct, research protocols (Table E2). To determine whether the different experimental
protocols employed across sites, such as DNA extraction, primer selection, or sequencing platform, biased the total observed bacterial diversity, we clustered samples from all sites using unweighted UniFrac and principal coordinate analysis (22). The first principal coordinate axis showed clear separation between oral wash samples and methodological controls, whereas BAL samples fell between these two sample types (Figure 1A). This pattern was stronger than clustering by research cohort, despite the differences in experimental protocols employed (Figure 1B). However, pairwise UniFrac distances from within-cohort comparisons of lung samples were significantly smaller than those for between-cohort comparisons, indicating that the overall cohort had some effect on the observed diversity.

### Higher Prevalence and Relative Abundance of T. whipplei in BAL in HIV-Positive Individuals

*T. whipplei* was detected in BAL samples from all eight cohorts (Table 2). Despite cohort effects on the overall observed diversity, *T. whipplei* relative abundance was not significantly different between cohorts within HIV-negative (Kruskal-Wallis test; P = 0.6098) or HIV-positive (P = 0.7182) individuals. Because the number of 16S rRNA sequences per sample varied widely (Table E2), and deeper sequencing results in higher detection rates of rare species, we measured carriage using all available data and using 500 randomly selected sequences from a single lung sample per subject.

The estimated carriage rate (using all 16S rRNA gene sequences) in HIV-negative individuals ranged from 12 to 40% within individual cohorts, and was 23.4% when considering all 77 individuals (Table 2). Carriage in HIV-positive individuals ranged from 20 to 70% within individual cohorts, and was 53.7% when considering all 82 individuals. Carriage was significantly higher in HIV-positive individuals when combining data from the different cohorts (31.7 versus 13.0% when standardized at 500 sequences from a single lung sample; G test for independence P = 0.002) (Table 2). Individuals infected with HIV also had a significantly higher relative abundance of *T. whipplei* compared with HIV-negative subjects when combining data from all cohorts (P = 0.002 G test; Table 2).

Of the eight cohorts, four had samples from both HIV-positive and HIV-negative individuals, but no cohort showed significant *T. whipplei* differences with HIV infection status in isolation (Table 2). To exclude the possibility that a significant overall result was driven by systematic differences induced by variability in the methodology across the different cohorts, we analyzed the P values from the four cohorts that had samples from HIV-positive and HIV-negative individuals using Fisher’s method, which combines the results from several independent tests bearing upon the same overall hypothesis. The combined P value indicated a significantly greater relative abundance of the *T. whipplei* phylotype in HIV-positive compared with HIV-negative individuals (P = 0.035), but the difference in carriage rate was not significant (P = 0.1; Table 2).

Of particular interest are individuals in whom *T. whipplei* dominated the lung microbiota. *T. whipplei* represented the majority of sequences (>50% of assigned sequence reads) in at least 1 lung sample in 11 of 82 HIV-positive subjects (13.4%), and comprised over 90% of sequences in at least 1 BAL sample in 8 HIV-positive individuals (9.8%). Conversely, *T. whipplei* comprised the majority of assigned reads in HIV-negative subjects only rarely (1 of 77 subjects), and the incidence of *T. whipplei* dominance significantly differed between the HIV-positive and HIV-negative groups (P = 0.0018 G test of independence).

The relative abundance of *T. whipplei* was significantly higher in HIV-positive subjects when controlling for cohort using a linear regression model (P < 0.05). In addition, a significant positive correlation was observed between *T. whipplei* relative abundance and BAL leukocyte counts (r = 0.33, P = 0.004). Although smokers were significantly enriched in our HIV-positive population compared with the HIV-negative population (P = 0.00071 G test of independence), the relative abundance of *T. whipplei* was significantly higher in HIV-positive subjects when controlling for smoking using a linear regression model (P = 0.020). In the HIV-negative population, however, smokers had a trend toward a higher relative abundance of *T. whipplei* compared with non-smokers (0.038 versus 0.0015; P = 0.08 G test).

None of the individuals in whom *T. whipplei* dominated the population (made up >50% of the sequence reads in at least one lung sample) reported any serious lung disorders at the time of sampling. One individual who reported *M. tuberculosis* infection had *T. whipplei* accounting for 25% of the sequence reads.

---

### Table 1. Percentage of Sequences within a 97% Identity Threshold to *Tropheryma whipplei* 16S Ribosomal RNA from a Human Immunodeficiency Virus-Infected Subject Before and After Antiretroviral Therapy

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>Before ART % (n)</th>
<th>After ART % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medial NeatBAL</td>
<td>99 (31,156)</td>
<td>34 (46,350)</td>
</tr>
<tr>
<td>Medial BALpellet</td>
<td>93 (31,508)</td>
<td>58 (47,552)</td>
</tr>
<tr>
<td>Lateral NeatBAL</td>
<td>67 (5,060)</td>
<td>11 (85,151)</td>
</tr>
<tr>
<td>Lateral BALpellet</td>
<td>ND</td>
<td>39 (82,028)</td>
</tr>
<tr>
<td>Posterior pharyngeal swab</td>
<td>0.045 (35,548)</td>
<td>0.0026 (38,079)</td>
</tr>
<tr>
<td>Oral wash</td>
<td>0 (25,730)</td>
<td>0 (45,108)</td>
</tr>
<tr>
<td>Stool</td>
<td>0 (42,010)</td>
<td>ND</td>
</tr>
</tbody>
</table>

---

**Definition of abbreviations:**
- ART = antiretroviral therapy
- BAL = bronchoalveolar lavage
- BALpellet = cell-free bacterial pellet
- ND = no data
- NeatBAL = unfractonated BAL

*Percentages are based on nonrarefied data.

1 The number of sequences evaluated in each case is in parentheses.
Selected sequences from one lung sample (rarefied; rare).

Fisher’s combined probability test 0.10 0.035

<table>
<thead>
<tr>
<th>Site</th>
<th>HIV⁺</th>
<th>HIV⁻</th>
<th>P Value (G Test)</th>
<th>HIV⁺</th>
<th>HIV⁻</th>
<th>P Value (t Test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorado–LHMP</td>
<td>5</td>
<td>40.0</td>
<td>11 36.4 18.2</td>
<td>0.20</td>
<td>0.20 ± 0.44</td>
<td></td>
</tr>
<tr>
<td>Colorado–San Francisco</td>
<td>5</td>
<td>40.0</td>
<td>0   ND ND</td>
<td>0.035</td>
<td>0.20 ± 0.44</td>
<td></td>
</tr>
<tr>
<td>Indiana–LHMP</td>
<td>10</td>
<td>70.0</td>
<td>15 40.0 20.0</td>
<td>0.12</td>
<td>0.096 ± 0.22</td>
<td></td>
</tr>
<tr>
<td>Indiana–long.</td>
<td>30</td>
<td>63.3</td>
<td>0   ND ND</td>
<td>0.12</td>
<td>0.087 ± 0.24</td>
<td></td>
</tr>
<tr>
<td>Michigan</td>
<td>0</td>
<td>ND</td>
<td>25 12.0 4.0</td>
<td>0.28</td>
<td>0.0013 ± 0.006</td>
<td></td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>15</td>
<td>40.0</td>
<td>20 20.0 15.0</td>
<td>0.11</td>
<td>0.20 ± 0.40</td>
<td></td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>12</td>
<td>50.0</td>
<td>6   16.7 4.0</td>
<td>0.16</td>
<td>0.037 ± 0.15</td>
<td></td>
</tr>
<tr>
<td>University of California, San Francisco</td>
<td>5</td>
<td>20.0</td>
<td>0   ND ND</td>
<td>0.20</td>
<td>0.002 ± 0.002</td>
<td></td>
</tr>
</tbody>
</table>

Total 82 53.7 31.7 77 23.4 13.0 0.0022 0.11

Pennsylvania 12 50.0 16.7 6 16.7 16.7 0.50 0.021

Michigan 0 ND ND 25 12.0 4.0 ND ND 0.0013

Indiana–long. 30 63.3 43.3 0 ND ND ND 0.087

Indiana–LHMP 10 70.0 50.0 15 40.0 20.0 0.12 0.096

Michigan 0 ND ND 25 12.0 4.0 ND ND 0.0013

Indiana–LHMP 10 70.0 50.0 15 40.0 20.0 0.12 0.096

Michigan 0 ND ND 25 12.0 4.0 ND ND 0.0013

Indiana–LHMP 10 70.0 50.0 15 40.0 20.0 0.12 0.096

Michigan 0 ND ND 25 12.0 4.0 ND ND 0.0013

The P value for carriage was calculated using the G test of independence and rarefied data.

Valuation of T. whipplei 16S rRNA Sequences Using hsp65-targeted qPCR

Although 16S rRNA sequences were clustered with T. whipplei with 97% sequence identity, we wished to confirm the presence of T. whipplei with an independent assay targeting a distinct genetic region. Therefore, we interrogated samples from the Colorado–LHMP, Indiana–LHMP, and Pennsylvania cohorts using qPCR targeting hsp65, which is highly conserved across all T. whipplei serotypes (12). There was no measurable amplification with these primers from purified Escherichia coli genomic DNA or from any of the methodological controls (Figure 2A). The sensitivity of the assay was reliably one copy per reaction; therefore, we could reliably detect T. whipplei in samples in which it was not very abundant.

When normalized to copies of 16S rRNA or from any of the methodological controls (Figure 2A).

The qPCR measurements were significantly correlated with relative abundance calculated from 16S rRNA sequences (Colorado: r = 0.90, P < 0.0001; Indiana: r = 1.0, P < 0.00001; and Pennsylvania: r = 0.50, P = 0.002), thus validating our T. whipplei 16S rRNA sequencing data.

Prevalence of T. whipplei in the Upper Respiratory Tract of HIV-Positive and HIV-Negative Individuals

Because T. whipplei has been previously detected in mouth samples (23–25), and the upper respiratory tract is the major source of overall diversity in our lung samples (Figure 1), we compared the relative abundance of T. whipplei in matched oral wash and BAL samples in individuals in whom T. whipplei was detected (n = 27) and found a significant enrichment of T. whipplei sequences in the lung (P = 0.0041, paired t test). T. whipplei was never detected in oral wash/posterior pharyngeal swab samples when 500 sequences per sample were evaluated (Table E3) and, in contrast to the lung, it was neither more frequently detected nor more abundant in oropharyngeal samples from HIV-positive than HIV-negative individuals.

Figure 2. Quantitative PCR analysis with the hsp65 gene. (A) Representative PCR of Tropheryma whipplei–specific hsp65 showing a base pair amplicon for a bronchoalveolar lavage (BAL) fluid sample (906–9006) from an individual with high T. whipplei relative abundance based on 16S ribosomal RNA (rRNA) in the Colorado–San Francisco (SF) cohort. (B) Copy numbers normalized to 16S rRNA expression in the Colorado, Indiana, and Pennsylvania cohorts is shown.

Effects of ART on T. whipplei Abundance in the Lung

We next explored whether T. whipplei decreased in HIV-positive individuals after ART. In a cross-sectional analysis, T. whipplei relative abundance and carriage were not significantly higher in ART-naive (n = 44) compared with ART-treated (n = 33) individuals (Table E4), nor was it significantly correlated with CD4 count or HIV viral load, giving an initial indication that ART did not help to control T. whipplei in the lung. Incomplete control of T. whipplei after ART is further supported by the observation that the relative abundance of T. whipplei was significantly higher in HIV-positive individuals than HIV-negative individuals when only ART-treated individuals were considered (P = 0.01, t test).

However, when comparing BAL samples collected longitudinally from 29 HIV-positive individuals before treatment and 4 weeks, 1 year, and 3 years after ART, treatment led to a significant reduction of the T. whipplei phylotype within individuals. We detected the T. whipplei phylotype in BAL samples from at least one time point in 19 of the 29 (66%) individuals. In six individuals, over 10% of the sequence reads belonged to the T. whipplei phylotype in at least one of the BAL samples (Figure 3). There was a significant decrease in the relative abundance of T. whipplei in BAL after 6 months to 3 years, but not 4 weeks of ART therapy. We measured significance with paired t tests comparing the average relative abundance of the T. whipplei phylotype at baseline and 4-week time points to the average from any 6-month, 1-year, and 3-year post-ART samples available. Significance was achieved when all 20 carriers (19 from individuals after ART). In a cross-sectional analysis, T. whipplei relative abundance and carriage were not significantly higher in ART-naive (n = 44) compared with ART-treated (n = 33) individuals (Table E4), nor was it significantly correlated with CD4 count or HIV viral load, giving an initial indication that ART did not help to control T. whipplei in the lung. Incomplete control of T. whipplei after ART is further supported by the observation that the relative abundance of T. whipplei was significantly higher in HIV-positive individuals than HIV-negative individuals when only ART-treated individuals were considered (P = 0.01, t test).

However, when comparing BAL samples collected longitudinally from 29 HIV-positive individuals before treatment and 4 weeks, 1 year, and 3 years after ART, treatment led to a significant reduction of the T. whipplei phylotype within individuals. We detected the T. whipplei phylotype in BAL samples from at least one time point in 19 of the 29 (66%) individuals. In six individuals, over 10% of the sequence reads belonged to the T. whipplei phylotype in at least one of the BAL samples (Figure 3). There was a significant decrease in the relative abundance of T. whipplei in BAL after 6 months to 3 years, but not 4 weeks of ART therapy. We measured significance with paired t tests comparing the average relative abundance of the T. whipplei phylotype at baseline and 4-week time points to the average from any 6-month, 1-year, and 3-year post-ART samples available. Significance was achieved when all 20 carriers (19 from individuals after ART). In a cross-sectional analysis, T. whipplei relative abundance and carriage were not significantly higher in ART-naive (n = 44) compared with ART-treated (n = 33) individuals (Table E4), nor was it significantly correlated with CD4 count or HIV viral load, giving an initial indication that ART did not help to control T. whipplei in the lung. Incomplete control of T. whipplei after ART is further supported by the observation that the relative abundance of T. whipplei was significantly higher in HIV-positive individuals than HIV-negative individuals when only ART-treated individuals were considered (P = 0.01, t test).
Indiana-longitudinal and 1 from Colorado) were considered (P = 0.032, paired t test), when only 7 individuals with high T. whipplei relative abundance were considered (>10% in at least one sample; P = 0.029, paired t test), and when only 9 low-abundance individuals were considered (<0.1% in all samples; P = 0.018, paired t test).

Similar to the cross-sectional analysis, successful ART did not always result in a decrease in the relative abundance of the T. whipplei phylotype. For example, subject 51,037 had no decrease after 1 year of treatment (Figure 3), despite a decrease in plasma HIV-1 RNA viral load from 10,600 to 400 copies/ml and an increase in CD4+ T cell count from 182 to 400 cells/μl during this period. Thus, although ART decreases the overall relative abundance of T. whipplei within individuals, ART does not completely restore the ability of HIV-positive individuals to control T. whipplei in the lung.

Genomic Analysis of T. whipplei in Lung Samples

To determine whether T. whipplei in lung resembled isolates cultured from patients with Whipple’s disease (Twist, isolated from a cardiac valve [26] and TW08/27, isolated from cerebrospinal fluid [27]), we applied shotgun metagenomic sequencing directly to BAL (eight samples from six individuals) with a high relative abundance of the T. whipplei phylotype (Table E5). Although the majority of the sequenced reads aligned to the human genome, a comparison of 69,271,986 kb of nonhuman reads to nearly 6,000 reference genome sequences showed T. whipplei to be the most abundant bacterium (Table E6). Thus, the abundance of T. whipplei that was inferred from sequences of the 16S rRNA and hsp65 genes was confirmed at the whole-genome level.

The sequences that aligned to the T. whipplei TW08/27 reference genome were assembled either for each individual when coverage of T. whipplei was sufficient (≥2×) or from pooling of all eight samples. The best coverage (20×) and most contiguous assembly came from the pooled sequences (Table E7), which were used in all subsequent analyses. Alignment of the assembled contigs to two T. whipplei reference genomes showed 90 and 92% coverage for TW08/27 and Twist, respectively (Figure 4; Table E8). A total of 27 potential genes were novel to the lung isolates (Table E9). Similar to previous T. whipplei genome comparisons (26, 28), the novel genes included three members of the highly variable WiSP family of surface proteins, suggesting their potential role in host interactions and immune evasion. In addition, 20 and 27 genes in Twist and TW08/27 genomes were not identified in the pooled assembly from lung isolates (Figure 4; Table E8), including several WiSP proteins, although we cannot determine whether these were truly absent in the lung strains or not detected due to the level of coverage.

DISCUSSION

We found that approximately half of the individuals with HIV infection harbor T. whipplei bacterial sequences in their lungs, often at very high relative abundance, and that relative abundance decreases with effective ART. The identity of this organism was confirmed by extensive cross-validation using 16S rRNA sequencing, hsp65-targeted qPCR, and shotgun metagenomic whole-genome sequencing. This is the first description of widespread lung colonization in asymptomatic HIV infection, and the first application of emerging microbiome tools to identify, within
the lower respiratory tract, a specific organism shared among individuals in the absence of clinical respiratory tract disease.

*T. whipplei* is the causative agent of Whipple’s disease, a rare systemic infectious disease. Classic *T. whipplei* disease typically involves the gastrointestinal tract, and is associated with diverse clinical manifestations, including weight loss, diarrhea, and joint, neurological, and cardiac and/or pulmonary involvement (14). Localized infections without histological digestive lesions have also been described (14). Pulmonary manifestations of *T. whipplei* disease include dyspnea, pleuritic chest pain, chronic cough, reduced lung volumes, and pleural adhesions (29). Whipple’s disease is often associated with defects of innate immune activation, resulting in an inability to control the organism (14–21). However, classic *T. whipplei* disease has seldom been described in HIV-positive subjects (29–34). Nevertheless, our whole bacterial genome sequences from pooled reads in shotgun metagenomic data derived directly from BAL indicates that *T. whipplei* in lung represents strains extremely similar to those that cause Whipple’s disease.

Our results indicate a high frequency of carriage of *T. whipplei* in the lung of HIV-positive individuals; however, it was also identified in a minority of HIV-negative individuals. A potential niche for *T. whipplei* in the healthy lung has previously been suggested (6). *T. whipplei* has also been identified in the stool and mouth of healthy individuals, but asymptomatic carriers have significantly lower *T. whipplei* loads than patients with Whipple’s disease (23). Estimates of *T. whipplei* carriage in the gastrointestinal tract and mouth of healthy subjects vary considerably across studies (0.6–35% for saliva), which may reflect age, geography, and environmental exposures and/or methodological differences (23–25, 35). Our estimates of *T. whipplei* relative abundance in the lung, however, did not significantly differ across cohorts, despite the geographical distances and differences in DNA extraction and sequencing technologies. Genotyping of 39 *T. whipplei* DNA samples from patients and 10 from asymptomatic carriers revealed that *T. whipplei* genetic diversity is unrelated to bacterial pathogenicity (36).

Our data reveal a new ecological niche within the human body for *T. whipplei* colonization, and suggest that a lack of immunologic control due to HIV coinfection increases prevalence. Although *T. whipplei* carriage in the mouth (23–25, 35) and a specific niche for *T. whipplei* in the subgingival and gingival sulcus plaque has been suggested (35), the dominance of *T. whipplei* in the lung compared with matched mouth samples suggests that the true niche of *T. whipplei* detected in saliva may often be the lung. Alternatively, the oral cavity or small bowel may serve as a reservoir for *T. whipplei*, and HIV-positive individuals may have a reduced ability to clear this microbe once introduced into the lung. A niche in the lung, however, is consistent with a high abundance in the lung of macrophages, the main target cell for *T. whipplei* elsewhere in the body, and the correlation between *T. whipplei* prevalence and BAL leukocyte counts.

Genomic sequencing has been performed on two isolates of *T. whipplei* from patients with Whipple’s disease: Twist (26) and TW08/27 (27). Sequenced after prolonged *in vitro* culture, these genomes were highly similar to one another, displaying 99% identity at the nucleotide sequence level (26). This remarkable degree of conservation was confirmed by comparative genomic hybridization of 16 clinical isolates to the Twist strain (28). Both studies, however, observed variability specifically in the WiSP membrane protein family, which was hypothesized to be associated with changes in surface-exposed bacterial proteins (26, 28). Our analysis, based on metagenomic sequences obtained directly from primary human material without *in vitro* culture, similarly identified members of the WiSP family among both the putative lung-specific genes and genes absent in our pooled assembly. The strong genomic similarity that we observed between the *T. whipplei* in our BAL samples and isolates from patients with Whipple’s disease, such as a reduced genome size, lack of mobile DNA elements, lack of genes encoding essential metabolic capabilities, and an elaboration of mechanisms for the variation of surface structures potentially for immune evasion, are consistent with a highly host-restricted and host-adapted intracellular lifestyle (26, 28).

The clinical consequences of *T. whipplei* growth in the lung of HIV-positive individuals are not known. Individuals infected with HIV have a much higher prevalence of noninfectious lung
diseases, including pulmonary hypertension and COPD, compared with control subjects (4). Whipple’s disease is associated with pulmonary manifestations (37), and a complete resolution of pulmonary hypertension after antibiotic treatment has been reported in an individual with Whipple’s disease (37, 38). Another study using 16S rRNA gene sequencing reported that T. whippelii comprised 66% of sequences from a subject with interstitial lung disease (39). Similarly, T. whippelii was the only identified bacterium in BAL from an immunocompromised patient who presented with community-acquired pneumonia and septic shock (40). In addition, respiratory symptoms can precede the development of gastrointestinal manifestations of Whipple’s disease, suggesting that the primary route of infection is not always the gastrointestinal tract (41). The possibility that T. whippelii overgrowth could lead to lung disease is supported by the pulmonary manifestations of Whipple’s disease and the finding that T. whippelii is abundant in the lung of subjects with other pulmonary diseases. Because T. whippelii is resistant to growth with standard culture techniques (42), its importance in HIV-associated lung disease may be underappreciated. Alternative methods to cultural techniques for detecting T. whippelii in clinical samples have been extensively developed (14), but are not typically used for patients with acquired immune deficiency syndrome.

Further clinical and immunologic studies are needed to elucidate the functional consequences of T. whippelii expansion in the lung of individuals infected with HIV, and its relationship to long-term pulmonary sequelae. Because progression to classic systemic Whipple’s disease rarely occurs in HIV-positive individuals (29), HIV-associated T. whippelii overgrowth may have unique, but still important, clinical implications.

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

Acknowledgment: A subset of subjects in this cohort was recruited from the University of California, Los Angeles (Roger Detels) and the University of Pittsburgh (Charles R. Rinaldo, Lawrence Kingsley), as part of the Multicenter AIDS Cohort Study (AI35042, UL1-RR025005, AI35043, AI35039, AI35040, AI35041). A subset of subjects in this cohort was recruited from the Connie Wofsy Study Consortium of Northern California (Ruth Greenblatt) of the Women’s Interagency HIV Study Collaborative Study Group (AI35004, AI35834, AI34994, AI34989, AI34993, U01AA42590 and HD33623). The WHS Collaborative Study Group was cofunded by the National Cancer Institute, the National Institute on Drug Abuse, and the National Institute on Deafness and Other Communication Disorders. Funding was also provided by National Center for Research Resources University of California San Francisco–CTSI grant UL1 RR024131.

References


