



HAL
open science

The respiratory tract microbiome, the pathogen load, and clinical interventions define severity of bacterial pneumonia

Ana Elena Pérez-Cobas, Christophe Ginevra, Christophe Rusniok, Sophie Jarraud, Carmen Buchrieser

► **To cite this version:**

Ana Elena Pérez-Cobas, Christophe Ginevra, Christophe Rusniok, Sophie Jarraud, Carmen Buchrieser. The respiratory tract microbiome, the pathogen load, and clinical interventions define severity of bacterial pneumonia. *Cell Reports Medicine*, 2023, 4 (9), pp.101167. 10.1016/j.xcrm.2023.101167 . pasteur-04558958

HAL Id: pasteur-04558958

<https://pasteur.hal.science/pasteur-04558958>

Submitted on 25 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

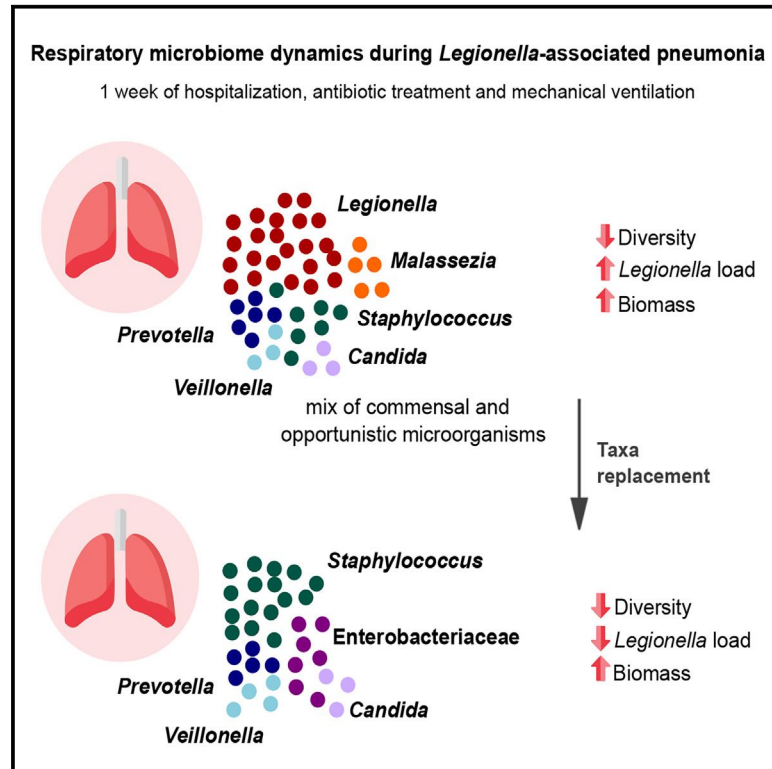
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

The respiratory tract microbiome, the pathogen load, and clinical interventions define severity of bacterial pneumonia

Graphical abstract



Authors

Ana Elena Pérez-Cobas,
 Christophe Ginevra, Christophe Rusniok,
 Sophie Jarraud, Carmen Buchrieser

Correspondence

anaelena.perez@salud.madrid.org
 (A.E.P.-C.),
 cbuch@pasteur.fr (C.B.)

In brief

Pérez-Cobas et al. show that inter-kingdom respiratory tract microbiomes during pneumonia and hospitalization are characterized by ecological species replacement. High bacterial and fungal loads are dysbiosis biomarkers. The equilibrium between a “commensal” and “pathogenic” cluster drives the microbiome to recovery or pathogenicity. Fungi, archaea, and protozoa contribute to progress of pneumonia.

Highlights

- A constant high biomass and species replacement characterizes the pneumonia microbiome
- High *L. pneumophila* load is associated with higher severity and co-morbidities
- High bacterial biomass appears as a biomarker for dysbiosis and feared co-infections
- Archaea, fungi, and protozoa within the respiratory microbiome need clinical research



Article

The respiratory tract microbiome, the pathogen load, and clinical interventions define severity of bacterial pneumonia

Ana Elena Pérez-Cobas,^{1,2,5,6,*} Christophe Ginevra,^{3,4} Christophe Rusniok,^{1,2} Sophie Jarraud,^{3,4} and Carmen Buchrieser^{1,2,7,*}

¹Institut Pasteur, Université Paris Cité, Biologie des Bactéries Intracellulaires, Paris, France

²CNRS UMR 6047, 75724 Paris, France

³Hospices Civils de Lyon, Centre National de Référence des Légionelles, Bron, France

⁴Centre International de Recherche en Infectiologie, Université Lyon 1, UMR CNRS 5308, U1111 Inserm, École Normale Supérieure de Lyon, Lyon, France

⁵Present address: Department of Microbiology, Ramón y Cajal Institute for Health Research (IRYCIS), Ramón y Cajal University Hospital, Madrid, Spain

⁶Present address: CIBER in Infectious Diseases (CIBERINFEC), Madrid, Spain

⁷Lead contact

*Correspondence: anaelena.perez@salud.madrid.org (A.E.P.-C.), cbuch@pasteur.fr (C.B.)

<https://doi.org/10.1016/j.xcrm.2023.101167>

SUMMARY

Bacterial pneumonia is a considerable problem worldwide. Here, we follow the inter-kingdom respiratory tract microbiome (RTM) of a unique cohort of 38 hospitalized patients ($n = 97$ samples) with pneumonia caused by *Legionella pneumophila*. The RTM composition is characterized by diversity drops early in hospitalization and ecological species replacement. RTMs with the highest bacterial and fungal loads show low diversity and pathogen enrichment, suggesting high biomass as a biomarker for secondary and/or co-infections. The RTM structure is defined by a “commensal” cluster associated with a healthy RTM and a “pathogen” enriched one, suggesting that the cluster equilibrium drives the microbiome to recovery or dysbiosis. *Legionella* biomass correlates with disease severity and co-morbidities, while clinical interventions influence the RTM dynamics. Fungi, archaea, and protozoa seem to contribute to progress of pneumonia. Thus, the interplay of the RTM equilibrium, the pathogen load dynamics, and clinical interventions play a critical role in patient recovery.

INTRODUCTION

Culture-independent techniques coupled with high-throughput sequencing have revolutionized the research on human-associated microbial communities.¹ In recent years, principally marker gene-based studies have revealed that airway surfaces are also inhabited by a resident microbial ecosystem, the respiratory tract microbiome (RTM).^{2,3} The bacterial composition of a healthy RTM mainly consists of the most common phyla of human microbiomes: Firmicutes, Bacteroidetes, and Proteobacteria.^{4,5} The primary source of microorganisms of the RTM is the oropharynx, from which they migrate through microaspiration and mucosal dispersion mechanisms to the lungs.^{6–8} Indeed, some of the common RTM genera, such as *Streptococcus*, *Prevotella*, or *Veillonella*, are shared with the oral microbiome.^{9,10} Although few studies have quantified the microbial load in the respiratory tract, it has been estimated that the biomass in healthy individuals decreases from the mouth to the lungs, where the load is minimum (up to 10^4 – 10^5 bacterial cells per milliliter of intraluminal fluid).¹¹ Despite the importance of non-bacterial in-

fections of the respiratory tract, only very few studies have focused on other domains of life in the RTM, such as eukaryotes, archaea, or viruses^{5,12,13}; thus, their role in the RTM ecology and human physiology remains unknown.

Research on the RTM has gained attention primarily because of its recently discovered role in host immune system stimulation.^{2,14,15} It has been also proposed that the RTM may protect against pathogens through colonization resistance, as seen for the gut microbiome.¹⁶ However, research on the RTM and its impact on infectious diseases such as pneumonia is still in its infancy, although pneumonia is a significant public health issue. To date, most research on the RTM and infection has been undertaken for specific chronic respiratory diseases, such as cystic fibrosis.^{17–19} These studies showed that cystic fibrosis severity indicators such as decreased lung function increase when a loss of microbiome diversity and a dominance of recognized cystic fibrosis pathogens are observed.¹⁸ Similarly, an ecological model of pneumonia proposed that the RTM equilibrium of pneumonia patients is altered to a state of dysbiosis characterized by low microbial diversity, a high burden of opportunistic species,



and particular host inflammatory responses.²⁰ Several clinical and ecological factors might locally affect the RTM ecology, shaping the outcome of a lung infection; however, the contribution of most of them still remains unknown.²¹

Among infectious diseases, acute lower respiratory tract infections comprising pneumonia, acute bronchitis, or bronchiolitis are a leading cause of morbidity and mortality worldwide.²² Pneumonia, an infection of the lung alveoli caused by bacteria, viruses, or fungi, represents a clinical and economic burden and a significant public health problem. *Legionella pneumophila*, which is also listed among the 33 pathogens responsible for an estimated 7.7 million deaths globally in 2019, is one of these bacteria, as it is responsible for a severe pneumonia called Legionnaires' disease (Global Burden of Disease Antimicrobial Resistance Collaborators, 2019, [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(22\)02185-7/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(22)02185-7/fulltext)). This infection can be fatal in approximately 5%–30% of cases despite timely and adequate treatment and in up to 40%–80% in untreated, immune-suppressed patients (World Health Organization, <https://www.who.int/news-room/fact-sheets/detail/legionellosis>). Legionellosis is mostly community acquired and occurs worldwide.²³ However, outbreaks are also reported frequently, the largest one occurring in 2014 in Portugal with 377 cases.²⁴ It is associated with high mortality, and patients often require hospitalization, which also represents an economic burden.²⁵ The main risk factors for legionellosis are old age, male gender, smoking, pulmonary-related conditions, immunosuppression, and chronic respiratory and renal illnesses.²⁶

In this study, we followed the inter-kingdom microbiome diversity and composition (bacteria, archaea, fungi, and protozoa) of 97 RTM samples from 38 patients with *Legionella*-associated pneumonia during their entire hospitalization period through marker gene high-throughput sequencing. The dynamics of the archaeal, bacterial, and fungal fractions were quantified based on a spike-in approach.²⁷ We show that in addition to the antibiotic treatment regimen, invasive ventilation, and patient risk factors, clinical and ecological variables related to pathogen load dynamics and microbiome ecology are important for the disease outcome. Our bacterial pneumonia model based on *L. pneumophila* infections sheds light on RTM during hospitalization and its relationship with ecological and clinical factors, which should be crucial in improving current clinical strategies to cure pneumonia and develop new microbiome-based therapies.

RESULTS

The respiratory microbiome of pneumonia patients is associated with high bacterial and fungal biomass and low species diversity

To investigate the composition of the inter-kingdom microbiome during Legionnaires' disease pneumonia and its possible role in disease severity, we studied a unique cohort of 38 patients hospitalized with pneumonia caused by *L. pneumophila* (Table S1). Those patients were sampled during the entire disease period, allowing us to analyze in total 97 samples and to follow the impact of clinical interventions on the microbiome. Samples were collected between 2017 and 2019, and all clinical param-

eters such as age, ventilation, antibiotic therapy, sequential organ failure assessment (SOFA) score, and *Legionella* PCR diagnostics were recorded for correlation analyses (for full patient-associated data, see Table S1). To quantify the microbial abundance of bacteria and fungi in these pulmonary samples, we used synthetic chimeric DNA spikes according to a method previously published²⁷ (detailed in STAR Methods [extended experimental procedures] and Figure S1). The spikes were co-isolated and PCR amplified, providing estimations of the amplicon abundance. After quality filtering, a total of 87 samples from 38 patients were used to characterize the diversity, composition, and biomass of bacteria and archaea (16S rRNA), 29 for the fungal diversity, composition, and load (internal transcribed spacer [ITS]), and 8 for diversity and composition of protozoa (18S rRNA) (the study design is summarized in Figure S2).

We then estimated the alpha diversity of bacteria and archaea in pulmonary samples during pneumonia, based on adjusted spikes ($n = 87$) using the number of copies of the 16S rRNA gene (Figure 1). The median number of bacteria was 10^7 copies per milliliter of sample (Figure 1A). The alpha- and beta-diversity analyses were based on the amplicon single variant (ASV) level. The mean alpha-diversity values were Shannon index (3), richness estimator Chao 1 (220 ASVs), and Berger-Parker index, a dominance index (0.3). We found that the alpha-diversity metrics correlated significantly with bacterial abundance (Figure 1B). Diversity and richness decreases while the number of bacteria and the dominance index increased. In addition, the microbial composition changed along the gradient of the bacterial load toward enrichment in Microbacteriaceae, *Stenotrophomonas*, *Pseudomonas*, and *Mycoplasma* (Figure 1C). The fungal biomass (estimated by the number of ITS copies per milliliter) and alpha-diversity correlations ($n = 29$) showed the same trend as that of bacteria (Figure 1D), with a high load associated with a low diversity and opportunists such as *Candida albicans* growing with the load gradient (Figure 1E). Samples with high biomass have low diversity and are dominated by a few potentially pathogenic respiratory tract pathogens that might be considered markers for dysbiosis.

Legionnaires' disease-associated RTM is characterized by clusters of "commensal" oropharynx bacteria and of opportunistic respiratory "pathogens"

The most abundant phyla (relative and absolute abundances) identified in the patients were Firmicutes (mean 42%), Proteobacteria (36%), Bacteroidetes (11%), Actinobacteria (4%), and Fusobacteria (4%) (Figures S3A and S3B). However, we found significant variability between samples, which reveals the dramatic changes in the microbiome associated with disease and medical therapy. When the taxa prevalence among samples was analyzed (relative and absolute abundance), the most abundant genera were Enterobacteriaceae (closest species to the most abundant ASV *Escherichia coli*), *Haemophilus* (*H. parainfluenzae*), *Legionella* (*L. pneumophila*), *Prevotella* (*P. melaninogenica*), *Staphylococcus* (*S. aureus*), *Stenotrophomonas* (*S. maltophilia*), *Streptococcus* (*S. infantis*), and *Veillonella* (*V. dispar*) (Figure 2A). Notably, the profile obtained from the most abundant taxa based on absolute values pointed to bacterial groups that were not determinant

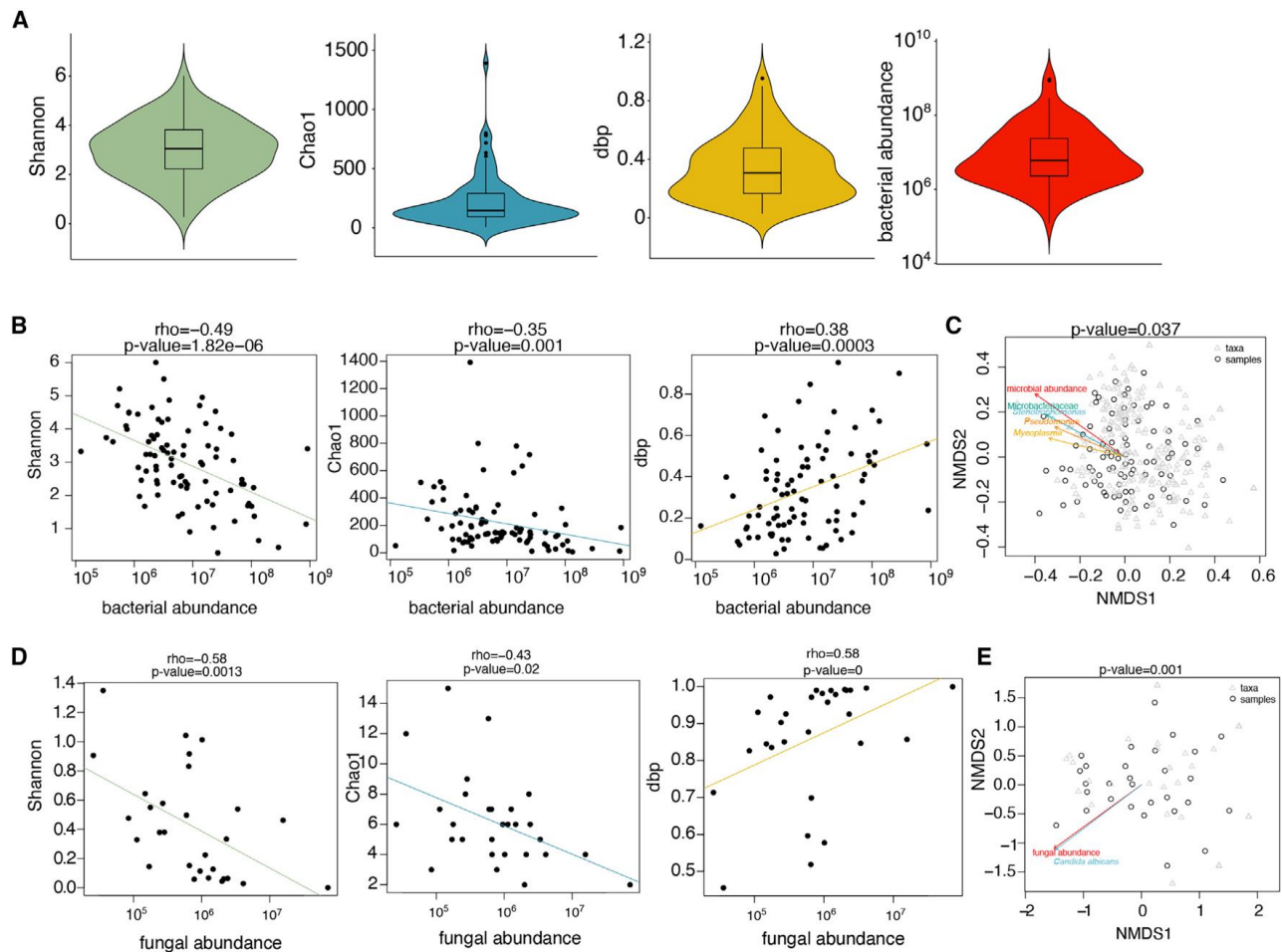


Figure 1. Diversity of the RTM of patients with pneumonia due to *L. pneumophila*

(A) Alpha-diversity metrics Shannon, Chao 1, Berger-Parker index (dbp), and biomass represented in violin box plots ($n = 87$). Bacterial abundance is measured as the number of 16S rRNA copies per milliliter of sample. The number of bacteria ranged from 10^5 to 10^9 , with a mean of 10^7 16S rRNA copies per milliliter of sample. The Shannon index distribution is variable (0.2–6), with most values between 3 and 4 (mean of 3). The richness estimator Chao 1 shows for most of the samples values between ~ 100 and 300 predicted ASVs (mean of 221), while the Berger-Parker index ranged mainly from 0.2 to 0.5 with a mean of 0.3.

(B) Spearman correlations of bacterial abundance with alpha-diversity metrics Shannon, Chao 1, and dbp ($n = 87$). The correlation coefficients (ρ) and p values are shown.

(C) Samples and taxa distribution represented in a non-metric multidimensional scaling (NMDS) plot based on the Bray-Curtis dissimilarity matrix ($n = 87$). The PERMANOVA-associated p value of the microbial abundance effect in shaping microbial composition is included above the graph. The bacterial abundance variable is plotted as a vector (red arrow), and the most significant taxa that change in correlation with the bacterial abundance are represented with colored arrows.

(D) Spearman correlations of fungal abundance with alpha-diversity metrics Shannon, Chao 1, and dbp ($n = 29$). The correlation coefficients (ρ) and p values are shown.

(E) Samples and taxa distribution represented in an NMDS plot based on the Bray-Curtis dissimilarity matrix ($n = 29$). The PERMANOVA-associated p value of the fungal abundance effect in shaping microbial composition is included above the graph. The fungal abundance variable is plotted as a vector (red arrow), and the most significant taxa that change in correlation with the bacterial abundance are represented with colored arrows.

n , number of samples.

when analyzing relative abundances (e.g., *Haemophilus* or *Stenotrophomonas*), highlighting the importance of quantification in microbiome studies.

To understand the role of the RTM in pneumonia, we need to shed light on ecological processes that define its functioning as microbial interactions. Thus, we performed a co-abundance network analysis to identify potential cooperative relationships between bacteria or bacteria-sharing niches (Figures 2B and

S3C). This allowed us to distinguish five clusters containing 41, 36, 23, 4, and 2 taxa, respectively. Considering the most prevalent taxa, *Prevotella*, *Streptococcus*, *Veillonella*, and *Haemophilus*, we identified a “commensal” cluster (red) enriched in beneficial bacteria generally considered members of the RTM of healthy individuals. On the other hand, *Staphylococcus*, *Stenotrophomonas*, *Enterobacteriaceae*, and *Legionella* constituted the “pathogenic” cluster (yellow). The highly connected nodes

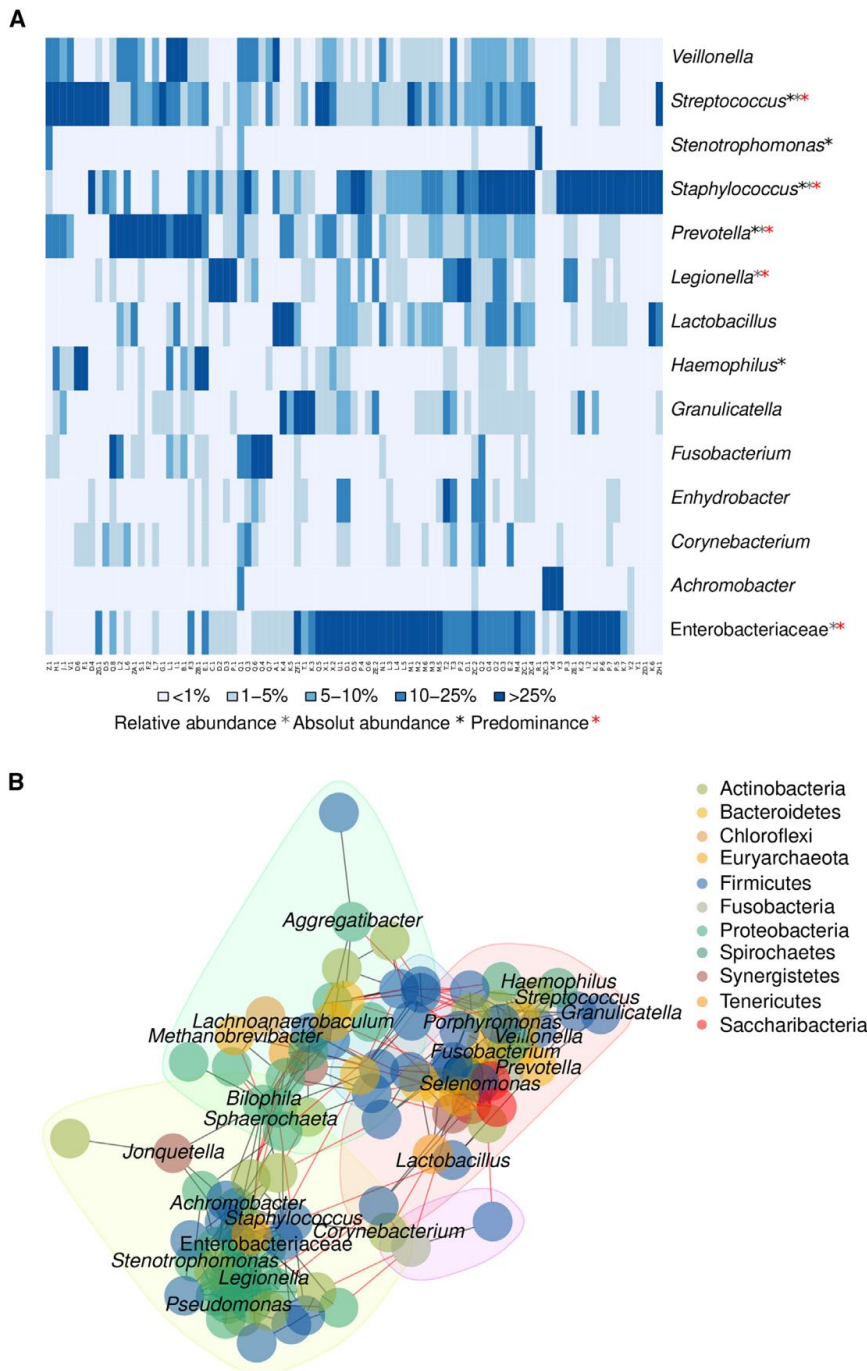


Figure 2. Putative interactions among dominant bacteria of the RTM of pneumonia patients

(A) Heatmap of the dominant bacterial taxa per sample ($n = 87$). The most significant taxa regarding predominance, relative abundance, and absolute abundance are marked with asterisks.

(B) The prokaryotic network of the RTM considers the most abundant taxa (mean over 10^3 16S rRNA copies per milliliter of the sample). The network is based on the ASV abundance tables collapsed at the genus level. Significant positive associations are shown ($p < 0.01$). Each node represents a taxon, color-coded by phylum. The main clusters of nodes are highlighted. Connections within a cluster are colored gray, and connections between clusters are in red. Dominant, highly connected, and articulation points of taxa are shown. A complete detailed network is shown in [Figure S3C](#).

n, number of samples.

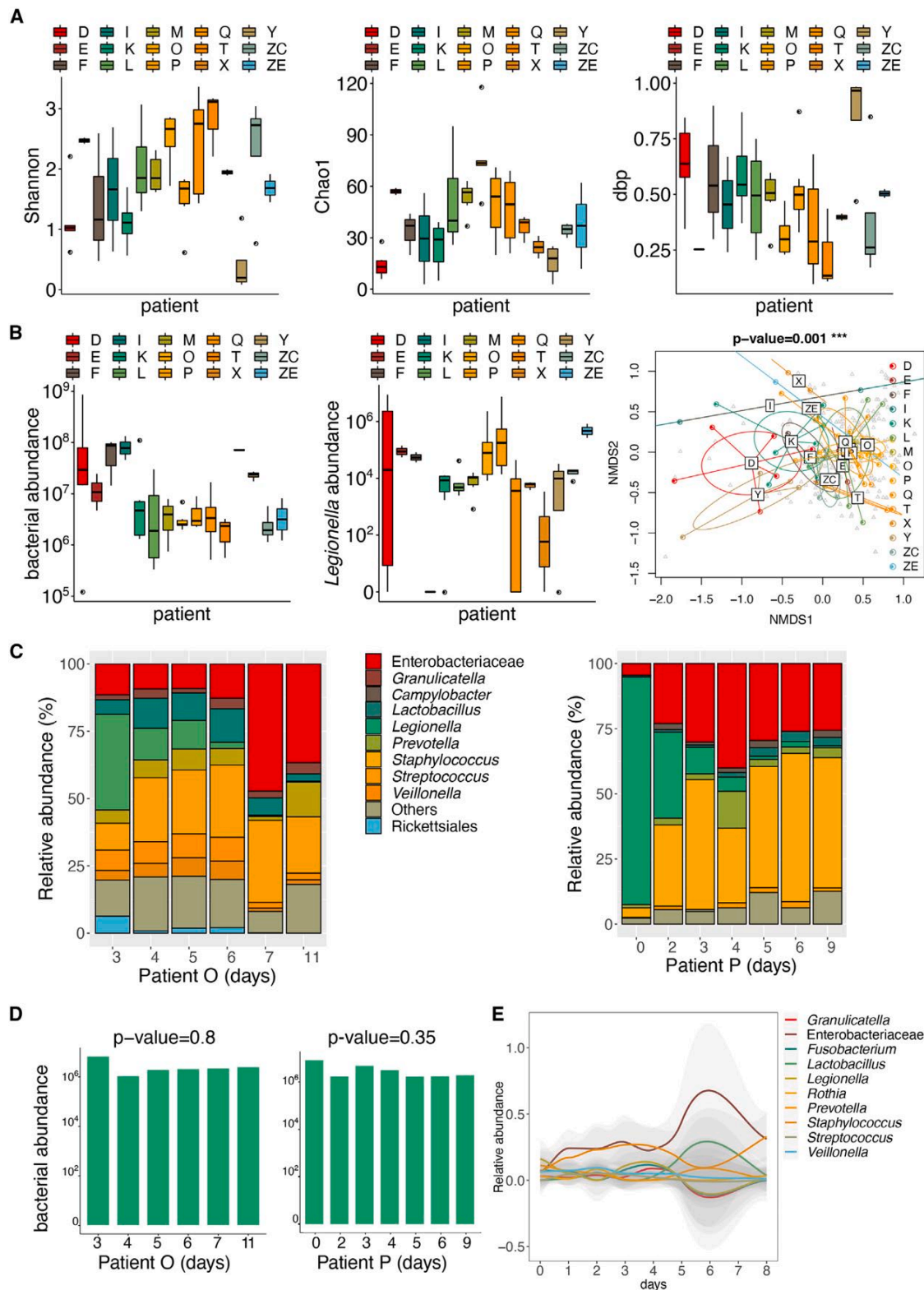
dances and the loss of such interconnected species in a cluster may determine the disappearance of the rest of the group as a “ripple effect.”

The patient's microbiome composition is rapidly shifted toward resistant, pathogenic species

To learn about the dynamics of the RTM during antibiotics treatment, we followed the alpha and beta diversity, bacterial biomass, and pathogen load of patients' samples over time ([Figure 3](#)). This shows that alpha diversity, bacterial biomass, and pathogen abundance variation exists within and between patients ([Figure 3A](#)). Interestingly, although great disparities in the microbiome composition were detected, the patient's individuality signal is well reflected in the bacterial composition of each patient ([Figures 3B and S3D](#)). The RTM dynamics of two patients (O and P) with similar sampling scheme (0–11 days) are shown in [Figures 3C and 3D](#), and full details of patients with longitudinal data are presented in [Figure S4](#). Both had an RTM

composition where a high abundance of *Legionella* was present, after which both microbiome compositions shifted rapidly (2–4 days post inclusion) toward an abundance of Enterobacteriaceae and *Staphylococcus* ([Figure 3C](#)). In both cases, the samples enriched in *Legionella* were those with the highest abundances (day 3 and day 0 for patients O and P, respectively) ([Figure 3D](#)). However, although *Legionella* was decreasing, there was still high biomass during the entire period examined ($\sim 10^6$ 16S rRNA copies per milliliter of sample). This

are *Staphylococcus*, Enterobacteriaceae, *Bilophila*, *Lachnoanaerobaculum*, and *Selenomonas* (buffering the network stability), and articulation points (nodes when absent disconnect the network) are *Aggregatibacter*, *Jonquetella*, and *Spirochaeta*. The stability of the most abundant RTM members during legionellosis might derive from the dynamics of the different network clusters. Those taxa more closely connected in the network might have a dependency on each other (i.e., cooperation) or might share the same niche (grow in the same conditions). Their abun-



(legend on next page)

suggests that during hospitalization and therapy, a “replacement” of *Legionella* by other potential pathogens that could occupy its niche in the lungs is taking place, thereby maintaining constant high biomass. Variable temporal trends occur for the most abundant taxa when clustering their relative abundances from patients sampled during the same days (Figure 3E). Evident changes in bacterial abundance were not observed. Nevertheless the composition changed significantly, primarily influenced by periods where *Legionella* increases in the short term and *Staphylococcus* in the mid to long term. Significant differences in the RTM dynamics during the hospitalization period were observed also when analyzing individual patients, similar to our overall findings (Figure S4). A similar trend was seen as some taxa such as *Staphylococcus* in patients D, K, and P, Enterobacteriaceae in patients O, or *Prevotella* in patients L and Q, grew when *Legionella* or other abundant taxa decreased. The increase in *Prevotella* in patients L and Q may indicate recovery of the microbiome in those patients as compared to the microbiome of healthy individuals (Figure S4). Taken together, our findings suggest that *Legionella* or other abundant bacteria are replaced, which leads to the maintenance of a relatively constant biomass.

Clinical factors are associated with the respiratory microbiome diversity, composition, and biomass

We analyzed the association of continuous and categorical metadata variables (Table S2) with the alpha and beta diversity, the biomass of the RTM, and the load of the pathogen *Legionella* (Figure 4). We first examined the effect of the sample type (sputum, tracheal aspirates, bronchoalveolar lavage [BAL]) on the alpha or beta diversity and on bacterial abundance. Figure 4 shows that no significant difference in the alpha diversity or abundance was observed among sample types with respect to their composition. The difference due to the sample type is much lower than the effects caused by the patient, the antibiotic usage, or invasive ventilation. Although the sample type is a relevant factor to be considered in respiratory microbiome studies, in strongly disturbed systems such as the lung microbiomes of hospitalized individuals with severe infection, the difference associated with the sample type is minor compared to the impact of other factors that introduce more significant perturbations in the microbial ecosystem. Indeed, based on PERMANOVA test coefficient R² (R function adonis2), the variable “patient” explained most of the variance of the microbiome composition with great differences compared to the rest of the factors, although there were no significant differences in alpha diversity or biomass. Several factors, such as immunosuppression or gender, are associated with composition, while

invasive ventilation and hospitalization periods are also related to alpha diversity. Moreover, *Legionella* abundance is associated with severity-related clinical factors such as SOFA scores. The following sections present further details of these associations.

Legionella pneumophila load influences the severity of the disease and is linked to gender and risk factors

L. pneumophila showed a variable distribution in 68 of the 87 samples with an average abundance of 10⁵ copies per milliliter of sample (range 10²–10⁶). When analyzing the qualitative and quantitative clinical factors significantly correlated with the *L. pneumophila* load (Table S3 and Figure 4), we found that the respiratory, coagulation, cardiovascular, and total SOFA scores were positively associated, indicating that elevated *Legionella* concentrations in the lungs are associated with systemic effects of the infection. Moreover, the abundance of *L. pneumophila* varies significantly between day 0 and the day of the diagnosis by urinary antigen test (between 0 and 4 days with a median of 2 days), the days between day 0 and the beginning of antibiotic treatment (period 0–5 days, a median of 2), and the duration of invasive ventilation (between days 3 and 20, a median of 12 days). A higher amount of *Legionella* is observed in men and individuals with hemopathies and immunosuppression (Table 1). Men are known to have a higher incidence and mortality rate of legionellosis than women,²⁶ but the differences in the microbiome have not been considered. Owing to the differential incidence of Legionnaires’ disease between men and women, the number of samples from women in our study was lower (n = 8) compared to men (n = 79). No differences in diversity, richness, or dominance between genders were observed (Figure S5A). However, women showed a higher bacterial burden than men, while men harbored a higher amount of *Legionella* per milliliter of sample, which could be a crucial factor explaining why the disease outcome is often worse in men than in women (Figure S5B). In addition, the microbiome composition of men and women is significantly different, as we identified 48 gender-associated features, 32 of which are associated with men and 16 with women. The microbiome of men is enriched in Proteobacteria containing known pathogenic species belonging to Enterobacteriaceae, *Pseudomonas*, *Enhydrobacter*, *Stenotrophomonas*, *Haemophilus*, and *Legionella* (Figure S5C and Table S3). Quantification of the *L. pneumophila* load and correlation with clinical features and gender revealed that higher amounts of *L. pneumophila* are associated with the patient health status defined by the SOFA score and co-morbidities such as male gender, underlying diseases, and immunosuppression.

Figure 3. RTM diversity and composition per patient

(A) Alpha-diversity metrics Shannon, Chao 1, and Berger-Parker index (dbp). Patients with temporal samples were included (15 patients, n = 69).
(B) Total bacterial and *Legionella* load. Abundance represents the number of 16S rRNA copies per milliliter of sample. Samples and taxa distribution are represented in an NMDS plot based on the Bray-Curtis dissimilarity matrix. Taxa distribution is represented with triangles. The PERMANOVA-associated p value of the variable “patient” is included above the graph.
(C and D) (C) Temporal variation of the microbiome composition and (D) bacterial load of two patients (O and P).
(E) Significant temporal trends of the most abundant bacteria (%). Significance is based on the “Trendyspliner” (test for a significant non-zero trend over time) implemented in the splinectomeR package.
n, number of samples.

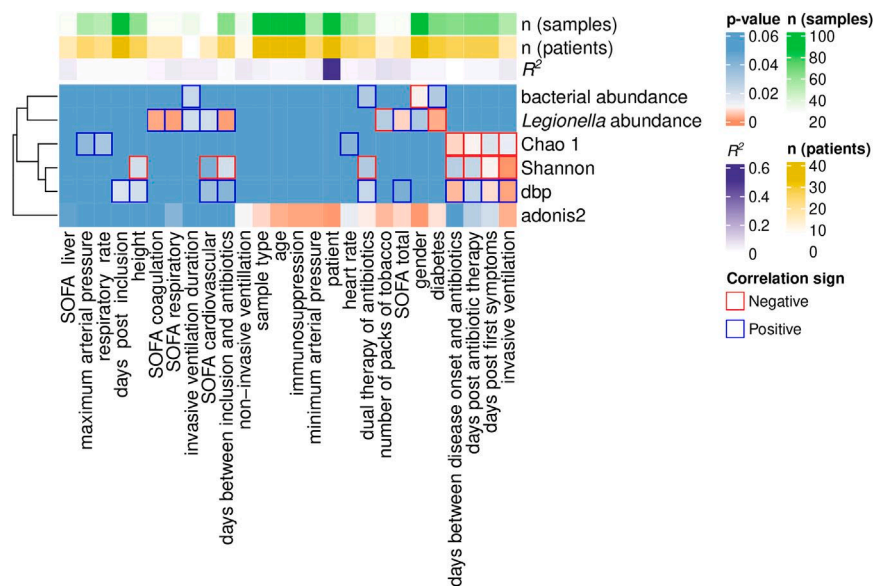


Figure 4. Statistical association of meta-data variables with the alpha and beta diversity, the biomass of the bacterial RTM, and the *Legionella* load

The heatmap shows the p values of the statistical analyses (Wilcoxon signed-rank test, Spearman correlation, PERMANOVA). The definition of the categorical variables and the alpha diversity (Shannon, Chao 1, and Berger-Parker index [dbp]) was based on the Wilcoxon signed-rank test with Bonferroni correction. The association with the alpha diversity was based on the Spearman correlation for continuous variables. The direction of the correlation coefficient is indicated. For the categorical variables “dual therapy of antibiotics,” the positive direction means higher abundance in treated samples; in “gender,” higher abundance in men; in “diabetes,” higher abundance in diabetic individuals; “invasive ventilation,” higher abundance in ventilated individuals. The association of the microbiome composition (beta diversity) with the variables was based on the PERMANOVA test (R adonis2 function). The number of samples and patients is shown at the top. The R^2 coefficient from the adonis2 result is also represented as a measure of the variance explained by the variable.

Invasive mechanical ventilation, hospitalization, and antibiotics alter the RTM of pneumonia patients

To investigate whether there is a connection between invasive ventilation and changes in the RTM microbiome, we analyzed all samples of ventilated patients ($n = 41$). This revealed a statistically significant association between the microbiome of ventilated patients with lower diversity and richness compared to that of non-ventilated ones and a higher Berger-Parker index indicating the dominance of few ASVs (Figure 5A). Generally, when analyzing ventilated patients over time (D, L, M, P, Y, ZC), we found a high abundance of only one or two taxa at the different points during hospitalization, including *Legionella*, Enterobacteriaceae, *Streptococcus*, or *Achromobacter* (Figure S4 and Table S2). Although no significant relationship between ventilation and total biomass or *Legionella* load was observed (Figure 5B), the microbiome composition differed between invasive-ventilated and non-ventilated patients (Figure 5B). Samples of patients with invasive ventilation showed a lower abundance of nine taxa and an enrichment in *Legionella*. Other clinically relevant pathogens, such as members of Enterobacteriaceae, *Staphylococcus*, or *Mycoplasma*, were associated with ventilation (Table S3). Invasive mechanical ventilation is associated with decreased alpha diversity and a composition shifted toward a source of pathogenic bacteria.

The evaluation of the effect of the most frequently administered antibiotics on the RTM (a dual therapy of macrolides and fluoroquinolones) ($n = 46$) showed that patients treated with this combination had a lower diversity and richness, and higher dominance of ASVs compared to patients who did not receive this combination therapy (Figure 5C). Interestingly, the total bacterial concentration was higher in treated individuals, but the *Legionella* amount was not statistically different among groups under different antibiotics treatment (Figure 5D). Furthermore, no global but specific compositional differences among both groups were identified

(Figure 5D). The samples from patients under dual therapy are associated with higher relative abundance of *Legionella* and a lower abundance of 17 taxa, including the predominant *Prevotella*, *Oribacterium*, or *Veillonella* (Table S3). Most patients who received this therapy were also under ventilation (D, M, P, Y, ZC), showing low diversity and dominance, and high abundance of one or two taxa, mainly *Legionella*, Enterobacteriaceae, *Streptococcus*, or *Achromobacter* (Figure S4; Table S2).

Antibiotic treatments are known to impact the human microbiome; however, here we showed that the standard dual therapy of macrolides and fluoroquinolones (the most common spiramycin + levofloxacin) led to a stronger decreasing diversity and richness than other dual therapies such as combinations of β -lactams (amoxicillin) and macrolides (pristinamycin, spiramycin, clarithromycin) or macrolides with β -lactams (cefotaxime + spiramycin; ceftriaxone + piramycin). We then compared the effect of specific combinations of antibiotics on the microbiome diversity and composition. No significant differences in alpha diversity related to the treatments were seen. However, the heterogeneity of the patient treatments is a limitation of our analyses, since it is strongly affected by the sample size. We found that significant composition changes depend on the antibiotic combinations administered (Figure S6). Further longitudinal studies with a larger number of samples would be critical to better understanding the impact of antimicrobial combinations.

Certain clinical periods could be determinants for the response of a patient to therapy, such as the number of days between diagnosis and antibiotic treatment or the number of days post antibiotic therapy. We thus evaluated the possible association of specific hospitalization periods with the microbiome. Most periods showed a significant negative correlation with diversity and richness and a trend in the dominance index to increase (Figure 4 and Table S3). These results suggest that RTM diversity decreases rapidly after hospitalization. Thus, the first days of

Table 1. Selected clinical parameters of patients included in this study

Patients	Days between day 0 and antibiotics
Total no. of patients 38; total samples 97	n = 31, median = 2, 1–3
Gender	Days post antibiotic therapy
male n = 30, %79	n = 31, median = 5, 2–8
Immunosuppression	Days post first symptoms
yes, n = 2, 5%	n = 31, median = 9, 7–14
Hemopathy	Days between symptoms and antibiotics
yes, n = 2, 5%	n = 31, median = 5, 3–7
Dual therapy macrolide-fluoroquinolone	SOFA cardiovascular
yes, n = 23, 60%	n = 33, median = 1, 0–4
Invasive ventilation	SOFA coagulation
yes, n = 13, 34%	n = 35, median = 0, 0–1
Invasive ventilation duration	SOFA total
n = 13, median = 12 days, 9–18	n = 35, median = 7, 3–10
Diabetes	SOFA respiratory
n = 2, 5%	n = 35, median = 3, 2–3
Days post inclusion (day 0)	Antibiotic therapy
n = 38, median = 2, 1–5	yes, n = 12, 32%

“n” refers to the number of patients. For complete detailed clinical information, see [Table S1](#).

hospitalization could be crucial to prevent RTM diversity drops typically associated with highly dysbiotic microbiome profiles. Further clinical studies are needed to test whether medical interventions at this time could avoid diversity loss and, in consequence, improve the patient’s health through modulating the host-microbiome interplay. Interestingly, the bacterial abundance generally remained constant over these periods, supporting the theory of species replacement (see extended experimental procedures in [STAR Methods](#) for statistics).

The non-bacterial fraction during *Legionella*-associated pneumonia is enriched in opportunistic pathogens

The fungal community was dominated by Ascomycota, mainly *Candida* species (median of 96%, 10^5), pointing to the presence of the opportunistic pathogen *C. albicans*, followed by *Malassezia* species (closest species *M. globosa* 0.6%, 10^3) during pneumonia. Unfortunately, 1.3% (10^3 copies) could not be allocated to specific taxa, indicating that a large part of the fungal diversity in the pulmonary microbiome remains undescribed and further in-depth analyses are necessary ([Figure 6A](#)). As for bacteria-bacteria, we performed a co-abundance network analysis to identify potential cooperative inter-kingdom relationships or shared niches ([Figure 6B](#)). Similar to our analysis of bacteria-bacteria interactions that had identified a cluster of pathogenic bacteria ([Figure 2B](#)), we identified a cluster where *Aspergillus*, a fungus associated with respiratory infections, co-occurs with bacterial pathogens such as *Legionella*, Enterobacteriaceae, *Acinetobacter*, or *Staphylococcus*. In contrast, *Candida* belongs to a cluster that is particularly enriched in oral anaerobic gram-positive commensals and various opportunistic pathogens such as *Bulleida*, *Atopobium*, *Orobacterium*, *Selenomonas*, or *Lachnoanaerobaculum* ([Figure S3E](#)). Furthermore, we identified a cluster of fungi in the RTM such as *Penicillium* and other Saccharomycetes

that are primarily associated with the environment. Additional research is necessary to determine whether their presence is transient in the RTM. Archaea constitute an important part of the RTM, as we identified them in 23% of the samples (n = 87) ([Table S3](#)). The genus *Methanobrevibacter* was the most common archaea, of which *Methanobrevibacter oralis* was identified in 14% of the samples (0.05% and 10^3 copies per milliliter of sample). Interestingly, a maximum of 10^5 copies were identified in a patient with an immune system condition (lymphoid leukemia). *Methanobrevibacter* is linked in the co-occurrence network with certain commensal bacteria, such as *Bifidobacterium*, as well as other bacteria that are not well known or not known to be significant in respiratory illnesses ([Figure S3C](#)); thus, their presence does not seem to be related to virulence. Although we could analyze only a small number of samples (n = 8) for the presence of protozoan species, we found that all were positive for the protozoa *Trichomonas tenax* in variable proportions ([Table S3](#)). *T. tenax* is a commensal in the oral cavity but it can also cause lung infections, particularly in immunodeficient patients.²⁸ Indeed, three of the five patients in whom *T. tenax* was identified suffered from cancer (one chronic lymphocytic leukemia and two multiple myeloma) associated with several alterations of the immune system, possibly favoring the presence of *Trichomonas*. Presence and diversity of protozoa in the RTM have been neglected. Despite marker gene and metagenomic analyses for this taxonomic group being challenging, the presence of protozoa in the lungs should be considered in future research, especially concerning patients with immunological conditions.

DISCUSSION

The RTM during pneumonia has been poorly analyzed, with only very few longitudinal studies that have mainly focused

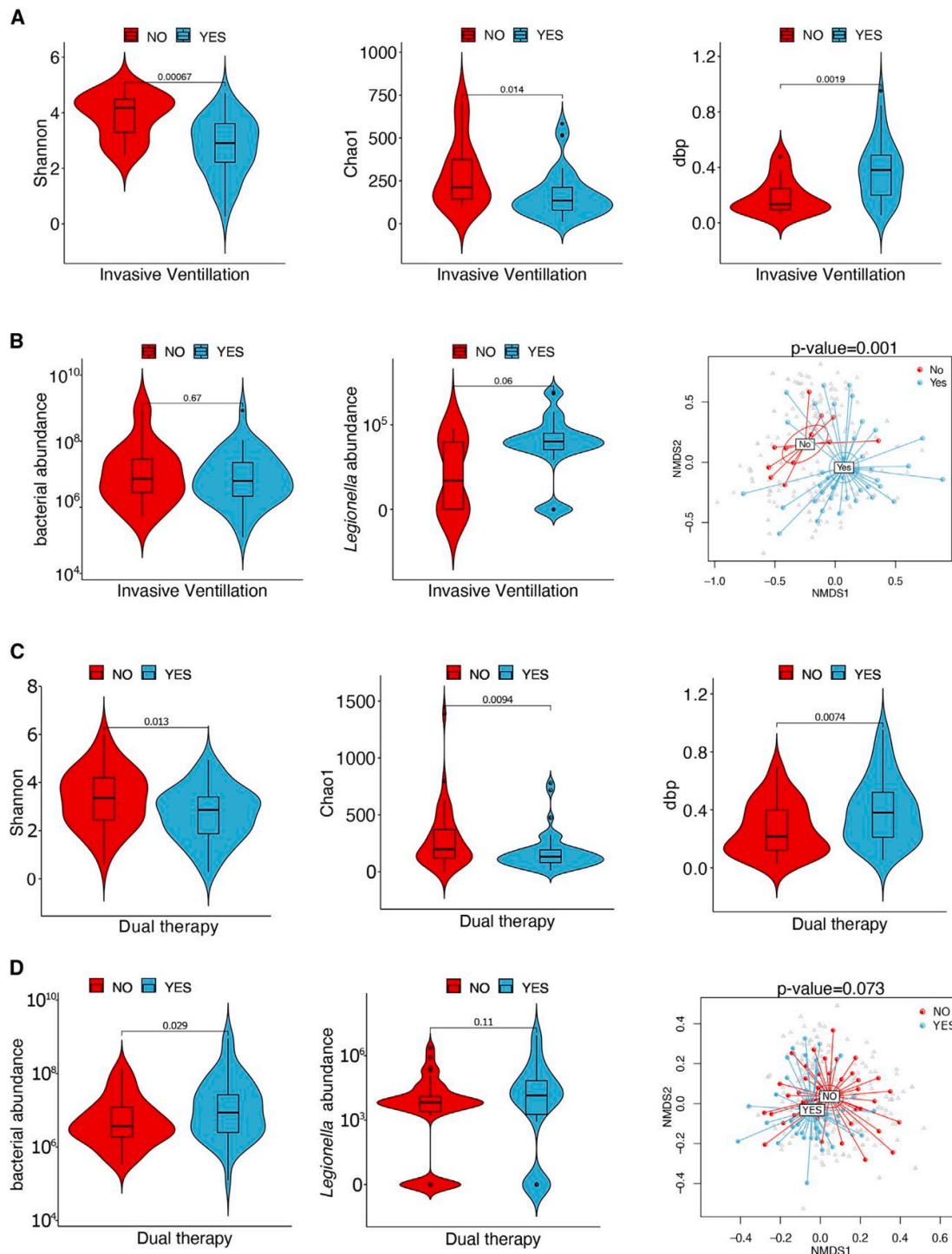


Figure 5. Impact of clinical interventions on RTM diversity and composition

(A) Effect of mechanic ventilation on alpha-diversity metrics Shannon, Chao 1, and Berger-Parker index (dbp) (n = 53).
 (B) Total bacterial and *Legionella* load (n = 53). Abundance: number of 16S rRNA copies per milliliter of sample. The comparison is based on the Wilcoxon signed-rank test with Bonferroni correction. p values are shown. Samples and taxa distribution are represented in an NMDS plot based on the Bray-Curtis dissimilarity matrix. Taxa distribution is represented with triangles. The PERMANOVA-associated p value of the variable "mechanical ventilation" is included above the graph.
 (C) Effect of dual therapy (macrolide and fluoroquinolone) on alpha-diversity metrics Shannon, Chao 1, and dbp (n = 50).

(legend continued on next page)

on bacterial communities characterized through qualitative approaches.^{29,30} In this study, we performed a comprehensive analysis of the RTM in a unique cohort of 38 patients hospitalized with severe pneumonia caused by the bacterium *L. pneumophila* as model of infection. We investigated the inter-kingdom microbiome composition (archaea, bacteria, fungi, protozoa) and analyzed its composition, diversity, and biomass as well as the interplay of clinical variables, pathogen load dynamics, microbiome diversity, composition, and biomass.

Our results support the ecological model of pneumonia, which proposes that the RTM equilibrium is disrupted during infection, leading to dysbiosis characterized by low microbial diversity, high microbial burden, and host inflammation.²⁰ The pulmonary microbiome of cystic fibrosis patients associated with infections shows similar biomass and alpha-diversity ranges.^{31,32} In contrast, analyses of healthy individuals revealed lower biomass (10^4 – 10^5) and higher diversity than present in samples from pneumonia patients.^{11,33} Here we found that pulmonary microbiomes with the highest biomasses had the lowest diversity and were enriched in opportunistic bacterial and fungal pathogens such as *Pseudomonas*, *Stenotrophomonas*, or *Candida*. This phenomenon is well known in other human-associated microbial communities such as the gut microbiome, where low diversity and compositional profiles different from healthy ones are linked to disease.³⁴ Although high biomass is expected to occur in dysbiotic microbiomes with low diversity and enrichment in pathogens, few studies include quantification in this regard and even less in the analyses of the RTM. Our results thus suggest that high biomass could be used as a risk index for secondary infections or co-infections.

We discovered that the microbiome of Legionnaires' disease pneumonia patients is characterized by an outgrowth of Firmicutes and Proteobacteria, a typical characteristic of lung dysbiosis.³⁵ At the genus level, the RTM contained a combination of oral commensals and respiratory or opportunistic pathogens such as Enterobacteriaceae, *Haemophilus*, *Legionella*, *Prevotella*, *Staphylococcus*, *Stenotrophomonas*, *Streptococcus*, and *Veillonella*. This is similar to what we found previously in the analysis of persistent legionellosis, where *Legionella* and *Staphylococcus* were also abundant pathogens in BAL samples.³³ We found that the dominant microorganisms are primarily members of the oral and respiratory microbiome that are known to cause lung infections during hospitalization and immunosuppression (certain Enterobacteriaceae species, *Haemophilus*, Parainfluenzae, or *Stenotrophomonas maltophilia*).^{36–38} *S. maltophilia* is a nosocomial pathogen due to its inherent and acquired resistance to many broad-spectrum antibiotics,³⁷ and *S. aureus* is a critical respiratory pathogen associated with high antibiotic resistance.^{39,40} Moreover, commensals of the RTM, such as *Prevotella melaninogenica*, *Streptococcus infantis*, and *Veillonella dispar*,^{41–43} are also abundant. A correlation network approach showed that abundant bacteria are distributed mainly in a “path-

ogenic” and a “commensal” cluster. Indeed, some of the taxa associated with *Legionella*-caused pneumonia are not common members of a healthy RTM¹⁴ and are worrying antibiotic-resistant bacteria^{44,45} such as *Kocuria*, *Curvibacter*, *Acinetobacter*, *Paracoccus*, *Pseudomonas*, and *Enhydrobacter*. Importantly, besides pathogenic and/or resistant bacteria, clinically relevant fungi involved in respiratory infections such as *Aspergillus* are also found to be associated with *Legionella*-caused pneumonia.⁴⁶ This analysis identified intra- and inter-kingdom associations that might be critical for the microbiome equilibrium, which deserve further research as they may represent a target for microbiome-based therapeutic strategies. The pathogenic species identified other than *Legionella* might participate and cause co-infections or lead to secondary infections; thus, the cluster of these bacteria should be considered in medical interventions as a target. Instead of targeting individual pathogens, strategies disrupting the microbial interaction network might be helpful to eliminate the entire cluster that is enriched in pathogens. Supporting this idea, a study of bronchiectasis, a chronic airway disease often complicated by microbial infections, showed that beyond antibiotic killing the common pathogen (*Pseudomonas*) it also helps to eliminate *Pseudomonas* by significantly affecting its microbial interactions network.⁴⁷ In contrast, the cluster of beneficial bacteria might be used for strategies to restore the microbiome equilibrium and colonization resistance capacity, for example through prebiotics, probiotics, or synbiotics.⁴⁸

The longitudinal analyses of the RTM of 15 patients disclosed significant differences in alpha diversity, bacterial biomass, and pathogen abundance within and between patients. This resembles what was observed for the gut microbiome and reported in other studies of the RTM.^{49,50} Although the microbial community equilibrium within patients is strongly altered by infection, hospitalization, and treatment for pneumonia, high loads of bacteria persist in the RTM over time due to a replacement of species. For example, *L. pneumophila* was seen to be replaced by *Staphylococcus* or Enterobacteriaceae. The reduction of *Legionella* due to antibiotic treatment probably generates an empty niche that can now be occupied by bacteria that are resistant to the antibiotics used.⁵¹ Indeed, *Staphylococcus* and Enterobacteriaceae are known to possess various resistance mechanisms against several classes of antibiotics.^{52,53} The quantification of the pathogen biomass showed that a high *L. pneumophila* load is associated with more severe disease, and it is also higher in patients with co-morbidities such as male gender, underlying disease, or immunosuppression. In men, we observed a higher burden of *Legionella* and a microbiome enriched in Proteobacteria, known to be associated with dysbiosis and inflammation,^{33,35,54} and we provide an explanation for the higher severity and mortality of Legionnaires' disease pneumonia in men.

Among clinical factors, invasive mechanical ventilation was significantly associated with low-diversity microbiome and

(D) Total bacterial and *Legionella* load (n = 50). Abundance represents the number of 16S rRNA copies per milliliter of sample. The comparison is based on the Wilcoxon signed-rank test with Bonferroni correction. p values are shown. Samples and taxa distribution are represented in an NMDS plot based on the Bray-Curtis dissimilarity matrix. Taxa distribution is represented with triangles. The PERMANOVA-associated p value of the variable “dual therapy” is included above the graph.

n, number of samples.

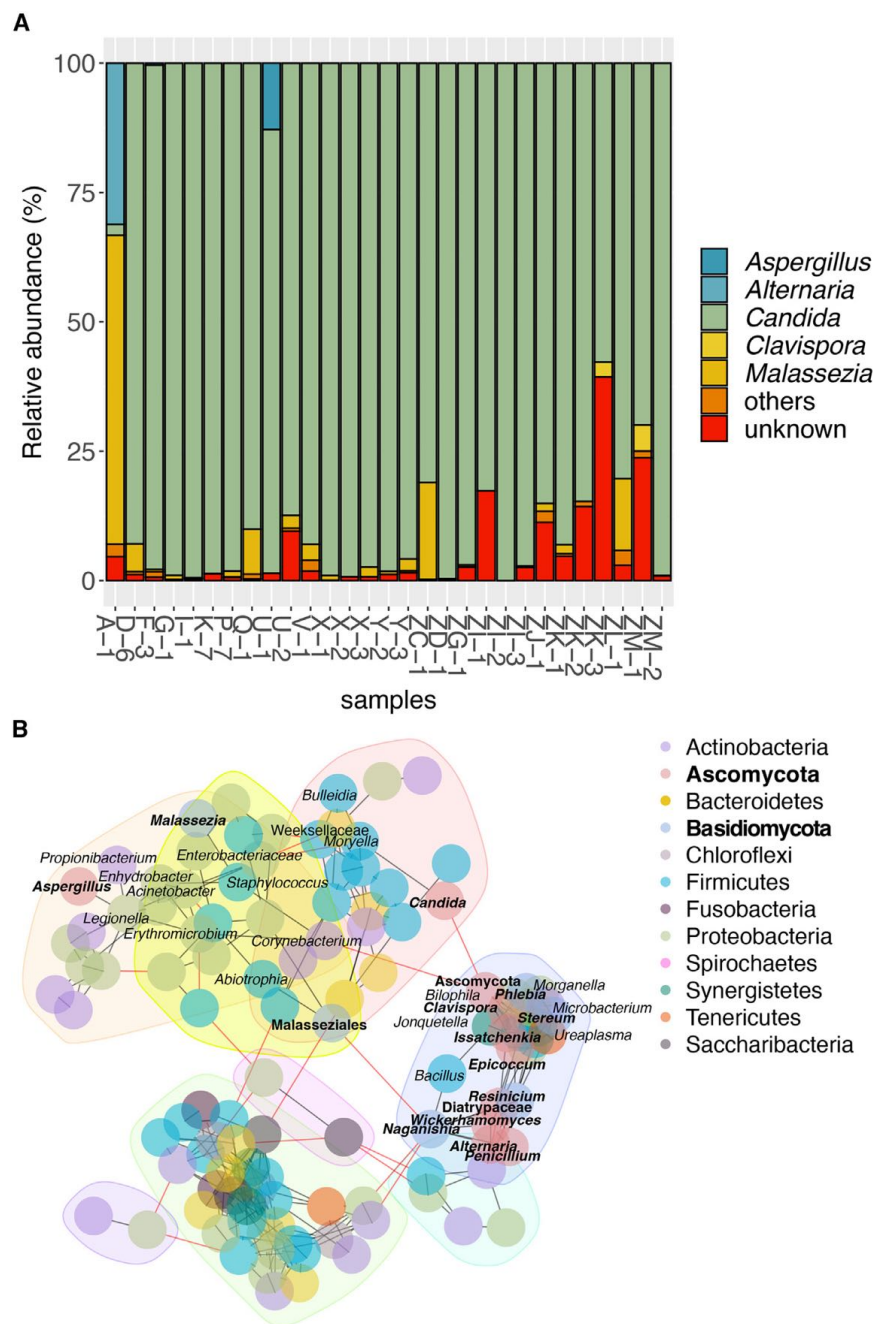


Figure 6. Predicted inter-kingdom RTM interactions

(A) Relative abundance (%) of the overall most prevalent fungal genera of the RTM. The taxonomy is based on the UNITE database from QIIME software (n = 29).
 (B) The inter-kingdom (bacteria, archaea, and fungi) network considers the most abundant taxa (mean over 10³ 16S rRNA copies per milliliter of the sample and 10 for the ITS) (n = 17). The network is based on the ASV abundance tables collapsed at the last identified taxonomic level. Significant positive associations are shown (p < 0.05). Each node represents a taxon color-coded by phylum. The main clusters of nodes are highlighted. Gray, connections within a cluster; red, connections between clusters; bold, fungal taxa. Only the closest taxa to fungi are shown. A complete, detailed network is shown in Figure S3E.
 n, number of samples.

specific compositional profiles enriched in Enterobacteriaceae and *Staphylococcus*.⁵⁵ In this regard, it has been shown that the microbiome in ventilated intensive care unit (ICU) patients is more perturbed than in non-ventilated individuals, exhibiting a lower alpha diversity that decreases with the duration of mechanical ventilation and dominance of only one or few species.^{29,30,56,57} A study analyzing the RTM of children admitted to the pediatric ICU showed that the RTM composition shifted after mechanical ventilation and antibiotics to profiles dominated by Enterobacteriaceae and *Staphylococcus*, which supports our results.⁵⁸ However, only few studies have evaluated the ef-

fect of mechanical ventilation on the microbiome, making it challenging to establish conclusions.⁵⁵ We also cannot rule out that the decrease in diversity is caused by co-factors such as the patient's health state. Most patients with the lowest diversity, high dominance of one or few species, and presence of a pathogen(s) required ventilation and dual therapy. These results can be the basis for further studies to search microbiome-based biomarkers of severity.

Here we revealed that a significant drop in diversity associated with invasive mechanical ventilation and its duration occurs, that commensals are under-represented, and that enrichment in harmful pathogens, including resistant bacteria, took place, which might promote polymicrobial and/or secondary infections. Another significant clinical factor associated with RTM during pneumonia is hospitalization and treatment duration. We show that the diversity and richness rapidly decrease with time, favoring

higher dominance and changes in the composition generally toward resistant bacteria such as *S. aureus* and a replacement of bacteria over time, as we observed by analyzing patients individually. Indeed, it has been shown that microbiome dysbiosis increases with hospitalization time in ICU patients.⁵⁹ Taken together, these data indicate that pneumonia and medical interventions could favor a decrease in alpha diversity and changes in the composition of the RTM that are critical during the first days of hospitalization.

Antibiotic treatments are known to impact the human microbiome. However, here we showed that the standard dual therapy

of macrolides and fluoroquinolones (the most commonly used ones are spiramycin + levofloxacin) led to a stronger decrease in diversity and richness than the dual therapy β -lactams (amoxicillin) and macrolides (pristinamycin, spiramycin, clarithromycin) or macrolides combined with cephalosporins (spiramycin + cefotaxime, spiramycin + ceftriaxone). Each combination of antibiotic treatments is associated with different, specific microbiome profiles but all antibiotics have a strong short-term and long-term effect on the microbiome composition. However, treatment variations between patients do not allow us to draw definitive conclusions, as too many variables are under consideration. Large cohorts of patients need to be analyzed in combination with animal models and *in vitro* studies to understand the specific effects associated with the use of different antibiotics and different antibiotic combinations.

Recent studies suggest that interactions of organisms of different kingdoms, in addition to bacteria-bacteria interactions, play a significant role in the ecology of the RTM.⁵ Interestingly, archaea (*M. oralis*), protozoa (*T. tenax*), and fungi (highlighting *C. albicans*) were present in various samples, patients, or time points analyzed, suggesting that they might be resident members and not transitory in the respiratory tract of hospitalized individuals. The potential inter-kingdom interactions identified support the theory that these taxonomic groups are relatively stable in the respiratory microbiome during pneumonia, possibly impacting human health. *M. oralis* has also been detected recently in sputum and BAL samples of hospitalized individuals.⁵⁰ *C. albicans* is an opportunistic fungus that is often promoted by extensive antibiotic usage and mechanical ventilation in hospitalized patients.^{61,62} Despite the difficulties associated with the identification of these taxonomic groups due to experimental procedures, classification methodologies, and their under-representation in databases, the presence of domains of life other than bacteria need to be considered in clinical microbiology, in particular in immunosuppressed patients where their abundance seems to be higher and might have an important influence on the success of the treatment and disease progression.

Taken together, here we showed that: (1) pneumonia caused by *L. pneumophila* in hospitalized individuals is characterized by two main clusters of “commensal” and “pathogenic” bacteria that may be the base for future microbiome-based therapies; (2) a constant biomass of bacteria during hospitalization and antibiotics treatment suggests a species niche replacement occurring after clinical interventions that needs further research; (3) high biomass of bacteria and fungi ($>10^7$, $>10^6$) might be risk biomarkers of co-infection and/or secondary infections by opportunistic pathogens; (4) high *Legionella* biomass correlates with disease severity and is linked to co-morbidities, suggesting that quantification of the pathogen should be included in medical monitoring of patients; (5) invasive mechanical ventilation, long hospitalization periods, and combined antimicrobial therapies strongly alter the RTM, an effect that should be considered in the clinic; and (6) the inter-kingdom microbiome may play a more critical role than considered so far and should be studied further. Therefore, the interaction of the RTM equilibrium and the pathogen load dynamics, together with clinical interventions, are crucial factors in the severity and outcome of bacterial pneumonia.

Limitations of the study

As RTM research is a new field, the study faces limitations in design, sampling, experimental, and computational approaches. According to the incidence of legionellosis, we analyzed an exceptionally large cohort with respect to sample size. However, the invasiveness of the sampling procedure (i.e., bronchoscopy) challenges the sample collection in respiratory studies, and thus the study is primarily observational. Another limitation is that microbiome analyses are based on marker gene sequencing. Metagenomics is more informative but is technically challenging because of the low microbial biomass and high cell content of respiratory samples. Future studies using novel methods based on microbial DNA enrichment and host DNA depletion should enable the application of omics for low-biomass samples.⁶³ Data analysis is also challenging, as most species remain unidentified and in particular significantly affect non-bacterial microbial communities. Finally, separating specific effects of clinical or environmental variables of the human RTM is challenging because of several confounding factors. More studies combining omics applied to large-size and/or longitudinal clinical samples, experimental models (*in vitro*, animals), and advanced computational and statistical approaches (i.e., machine learning) will shed light on RTM ecology in health and disease.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Human participants
- METHOD DETAILS
 - DNA extraction and marker genes sequencing
 - Quantification of bacterial and fungal load
 - Controls
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Marker gene data processing, ecological and statistical analysis
 - Extended experimental procedures

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.xcrm.2023.101167>.

ACKNOWLEDGMENTS

This work was supported by the Institut Pasteur, and the Agence Nationale de Recherche grant numbers ANR-15-CE17-0014-03 and ANR-10-LABX-62-IBEID to C.B. A.E.P.-C. was financed by ANR-10-LABX-62-IBEID. Work in the S.J. laboratory is financed by the Agence Nationale de la Recherche grant number ANR-15-CE17-0014-01, the DGOS, and Santé Publique France. A.E.P.-C. is a recipient of a Grant for the Attraction of Talent within the Community of Madrid (grant No. 2019-T2/BMD-12874). We thank the Institut

Pasteur Biomics pole (Christiane Bouchier and Laurence Ma) for sequencing the samples reported in this study.

AUTHOR CONTRIBUTIONS

A.E.P.-C. and C.B. designed the experiments. A.E.P.-C. conducted the experiments. A.E.P.-C. and C.R. analyzed and interpreted the data. C.G. and S.J. contributed to sample and metadata collection. The manuscript was written by A.E.P.-C. and C.B. with input from co-authors. The project was conceived, planned, and supervised by C.B.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 5, 2022

Revised: June 18, 2023

Accepted: August 2, 2023

Published: August 25, 2023

REFERENCES

- Pérez-Cobas, A.E., Gomez-Valero, L., and Buchrieser, C. (2020). Metagenomic approaches in microbial ecology: an update on whole-genome and marker gene sequencing analyses. *Microb. genomics* 6, 1–22. <https://doi.org/10.1099/mgen.0.000409>.
- Man, W.H., De Steenhuisen Piters, W.A.A., and Bogaert, D. (2017). The microbiota of the respiratory tract: Gatekeeper to respiratory health. *Nat. Rev. Microbiol.* 15, 259–270. <https://doi.org/10.1038/nrmicro.2017.14>.
- Kumpitsch, C., Koskinen, K., Schöpf, V., and Moissl-Eichinger, C. (2019). The microbiome of the upper respiratory tract in health and disease. *BMC Biol.* 17, 1–20. <https://doi.org/10.1186/s12915-019-0703-z>.
- Faner, R., Sibila, O., Agustí, A., Bernasconi, E., Chalmers, J.D., Huffnagle, G.B., Manichanh, C., Molyneaux, P.L., Paredes, R., Brocal, V.P., et al. (2017). The microbiome in respiratory medicine: Current challenges and future perspectives. *Eur. Respir. J.* 49, 1–12. <https://doi.org/10.1183/13993003.02086-2016>.
- Pérez-Cobas, A.E., Rodríguez-Beltrán, J., Baquero, F., and Coque, T.M. (2023). Ecology of the respiratory tract microbiome. *Trends Microbiol.* <https://doi.org/10.1016/j.TIM.2023.04.006>.
- Bassis, C.M., Erb-Downward, J.R., Dickson, R.P., Freeman, C.M., Schmidt, T.M., Young, V.B., Beck, J.M., Curtis, J.L., and Huffnagle, G.B. (2015). Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. *mBio* 6, 000377–e110. <https://doi.org/10.1128/mBio.00037-15>.
- Venkataraman, A., Bassis, C.M., Beck, J.M., Young, V.B., Curtis, J.L., Huffnagle, G.B., and Schmidt, T.M. (2015). Application of a neutral community model to assess structuring of the human lung microbiome. *mBio* 6, 1–9. <https://doi.org/10.1128/mBio.02284-14>.
- Dickson, R.P., Erb-Downward, J.R., Freeman, C.M., McCloskey, L., Falkowski, N.R., Huffnagle, G.B., and Curtis, J.L. (2017). Bacterial topography of the healthy human lower respiratory tract. *mBio* 8, e02287-16. <https://doi.org/10.1128/mBio.02287-16>.
- Araghi, A. (2020). The lung microbiome and pneumonia: Where precision medicine meets pulmonology. *Pulmonology* 26, 333–334. <https://doi.org/10.1016/j.pulmoe.2020.04.005>.
- Natalini, J.G., Singh, S., and Segal, L.N. (2022). The dynamic lung microbiome in health and disease. *Nat. Rev. Microbiol.* 21, 222–235. <https://doi.org/10.1038/s41579-022-00821-x>.
- Schneeberger, P.H.H., Prescod, J., Levy, L., Hwang, D., Martinu, T., and Coburn, B. (2019). Microbiota analysis optimization for human bronchoalveolar lavage fluid. *Microbiome* 7, 141–216. <https://doi.org/10.1186/s40168-019-0755-x>.
- Young, J.C., Chehoud, C., Bittinger, K., Bailey, A., Diamond, J.M., Cantu, E., Haas, A.R., Abbas, A., Frye, L., Christie, J.D., et al. (2015). Viral metagenomics reveal blooms of anelloviruses in the respiratory tract of lung transplant recipients. *Am. J. Transplant.* 15, 200–209. <https://doi.org/10.1111/ajt.13031>.
- Koskinen, K., Pausan, M.R., Perras, A.K., Beck, M., Bang, C., Mora, M., Schilhabel, A., Schmitz, R., and Moissl-Eichinger, C. (2017). First insights into the diverse human archaeome: Specific detection of Archaea in the gastrointestinal tract, lung, and nose and on skin. *mBio* 8, 008244–17–e917. <https://doi.org/10.1128/mBio.00824-17>.
- Segal, L.N., Clemente, J.C., Tsay, J.C.J., Koralov, S.B., Keller, B.C., Wu, B.G., Li, Y., Shen, N., Ghedin, E., Morris, A., et al. (2016). Enrichment of the lung microbiome with oral taxa is associated with lung inflammation of a Th17 phenotype. *Nat. Microbiol.* 1, 16031–16111. <https://doi.org/10.1038/nmicrobiol.2016.31>.
- Brown, R.L., Sequeira, R.P., and Clarke, T.B. (2017). The microbiota protects against respiratory infection via GM-CSF signaling. *Nat. Commun.* 8, 1512. <https://doi.org/10.1038/s41467-017-01803-x>.
- Thibeault, C., Suttrop, N., and Opitz, B. (2021). The microbiota in pneumonia: From protection to predisposition. *Sci. Transl. Med.* 13, eaba0501. <https://doi.org/10.1126/scitranslmed.aba0501>.
- O'Toole, G.A. (2018). Cystic fibrosis airway microbiome: Overturning the old, opening the way for the new. *J. Bacteriol.* 200, e00561-17–e00568. <https://doi.org/10.1128/JB.00561-17>.
- Cuthbertson, L., Walker, A.W., Oliver, A.E., Rogers, G.B., Rivett, D.W., Hampton, T.H., Ashare, A., Elborn, J.S., De Soya, A., Carroll, M.P., et al. (2020). Lung function and microbiota diversity in cystic fibrosis. *Microbiome* 8, 45–13. <https://doi.org/10.1186/s40168-020-00810-3>.
- Françoise, A., and Héry-Arnaud, G. (2020). The microbiome in cystic fibrosis pulmonary disease. *Genes* 11, 536. <https://doi.org/10.3390/genes11050536>.
- Dickson, R.P., Erb-Downward, J.R., and Huffnagle, G.B. (2014). Towards an ecology of the lung: New conceptual models of pulmonary microbiology and pneumonia pathogenesis. *Lancet Respir. Med.* 2, 238–246. [https://doi.org/10.1016/S2213-2600\(14\)70028-1](https://doi.org/10.1016/S2213-2600(14)70028-1).
- Pérez-Cobas, A.E., Baquero, F., de Pablo, R., Soriano, M.C., and Coque, T.M. (2022). Altered Ecology of the Respiratory Tract Microbiome and Nosocomial Pneumonia, 12, p. 4295.
- GBD 2016 Lower Respiratory Infections Collaborators; Blacker, B., Khalil, I.A., Rao, P.C., Cao, J., Zimsen, S.R.M., Albertson, S.B., Deshpande, A., Farag, T., Abebe, Z., et al. (2018). Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect. Dis.* 18, 1191–1210. [https://doi.org/10.1016/S1473-3099\(18\)30310-4](https://doi.org/10.1016/S1473-3099(18)30310-4).
- Mondino, S., Schmidt, S., Rolando, M., Escoll, P., Gomez-Valero, L., and Buchrieser, C. (2020). Legionnaires' Disease: State of the Art Knowledge of Pathogenesis Mechanisms of Legionella. *Annu. Rev. Pathol.* 15, 439–466. <https://doi.org/10.1146/annurev-pathmechdis-012419-032742>.
- Shivaji, T., Sousa Pinto, C., San-Bento, A., Oliveira Serra, L.A., Valente, J., Machado, J., Marques, T., Carvalho, L., Nogueira, P.J., Nunes, B., et al. (2014). A large community outbreak of Legionnaires disease in Vila Franca de Xira, Portugal. *Euro Surveill.* 19. <https://doi.org/10.2807/1560-7917.ES2014.19.50.20991>.
- Collier, S.A., Deng, L., Adam, E.A., Benedict, K.M., Beshearse, E.M., Blackstock, A.J., Bruce, B.B., Derado, G., Edens, C., Fullerton, K.E., et al. (2021). Estimate of burden and direct healthcare cost of infectious waterborne disease in the United States. *Emerg. Infect. Dis.* 27, 140–149. <https://doi.org/10.3201/eid2701.190676>.
- Farnham, A., Alleyne, L., Cimini, D., and Balter, S. (2014). Legionnaires' disease incidence and risk factors, New York, New York, USA, 2002–2011. *Emerg. Infect. Dis.* 20, 1795–1802. <https://doi.org/10.3201/eid2011.131872>.

- microbiota during and following mechanical ventilation for respiratory infections in children. *Eur. Respir. J.* 57, 2002652. <https://doi.org/10.1183/13993003.02652-2020>.
59. McDonald, D., Ackermann, G., Khailova, L., Baird, C., Heyland, D., Kozar, R., Lemieux, M., Derenski, K., King, J., Vis-Kampen, C., et al. (2016). Extreme Dysbiosis of the Microbiome in Critical Illness. *mSphere* 1, e00199-16. <https://doi.org/10.1128/mSphere.00199-16>.
60. Yasmine, H., Fabienne, B., Gérard, A., Michel, D., and Ghiles, G. (2020). Detection of methanobrevobacter smithii and methanobrevibacter oralis in lower respiratory tract microbiota. *Microorganisms* 8, 1–9. <https://doi.org/10.3390/microorganisms8121866>.
61. Jensen, J.-U.S., Hein, L., Lundgren, B., Bestle, M.H., Mohr, T., Andersen, M.H., Loken, J., Tousi, H., Soe-Jensen, P., Lauritsen, A.Ø., et al. (2015). Invasive Candida Infections and the Harm From Antibacterial Drugs in Critically Ill Patients. *Crit. Care Med.* 43, 594–602. <https://doi.org/10.1097/CCM.0000000000000746>.
62. Gutierrez, D., Weinstock, A., Antharam, V.C., Gu, H., Jasbi, P., Shi, X., Dirks, B., Krajmalnik-Brown, R., Maldonado, J., Guinan, J., and Thangamani, S. (2020). Antibiotic-induced gut metabolome and microbiome alterations increase the susceptibility to Candida albicans colonization in the gastrointestinal tract. *FEMS Microbiol. Ecol.* 96, fiz187. <https://doi.org/10.1093/FEMSEC/FIZ187>.
63. Nelson, M.T., Pope, C.E., Marsh, R.L., Wolter, D.J., Weiss, E.J., Hager, K.R., Vo, A.T., Brittnacher, M.J., Radey, M.C., Hayden, H.S., et al. (2019). Human and Extracellular DNA Depletion for Metagenomic Analysis of Complex Clinical Infection Samples Yields Optimized Viable Microbiome Profiles. *Cell Rep.* 26, 2227–2240.e5. <https://doi.org/10.1016/j.celrep.2019.01.091>.
64. Pérez-Cobas, A.E., and Buchrieser, C. (2019). Analysis of the pulmonary microbiome composition of Legionella pneumophila-infected patients. In *Methods in Molecular Biology*. https://doi.org/10.1007/978-1-4939-9048-1_27.
65. Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2. <https://doi.org/10.1038/s41587-019-0209-9>.
66. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. <https://doi.org/10.1038/nmeth.3869>.
67. Lahti, L., Shetty, S., and Blake, T. (2017). Tools for Microbiome Analysis in R. *Microbiome Packag. Version 0.99*.
68. Oksanen, J. (2015). *Multivariate Analysis of Ecological Communities in R: Vegan Tutorial*.
69. Lin, H., and Peddada, S.D. (2020). Analysis of compositions of microbiomes with bias correction. *Nat. Commun.* 11, 3514–3611. <https://doi.org/10.1038/s41467-020-17041-7>.
70. Friedman, J., and Alm, E.J. (2012). Inferring Correlation Networks from Genomic Survey Data. *PLoS Comput. Biol.* 8, 10026877–e1002711. <https://doi.org/10.1371/journal.pcbi.1002687>.
71. Csardi, G., and Nepusz, T. (2006). The Igraph Software Package for Complex Network Research (*InterJournal Complex Sy*), p. 1695.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
ZymoBIOMICS Microbial Community Standard	ZymoBIOMICS	#CATD6300
pSpike-P	addgene	#CAT101172
pSpike-F	addgene	#CAT101174
Chemicals, peptides, and recombinant proteins		
Phusion™ High-Fidelity DNA Polymerase	ThermoFisher	#CATF-530XL
DMSO	Merck	#CATD9170
dNTPs mix	Merck	#CATDNTPM-RO
agarose	ThermoFisher	#CATA16693-06
LightCycler TaqMan Master	Roche Diagnostics	#CAT04535286001
Critical commercial assays		
DNeasy PowerSoil Pro Kits	Qiagen	Cat#47014
Deposited data		
Marker genes sequencing fastq files	This paper	European Bioinformatics Institute: PRJEB55553
Oligonucleotides		
515F	GTGCCAGCMGCCGCGGTAA	N/A
806R	GGACTACHVGGGTWTCTAAT	N/A
ITS1	CTTGGTCATTTAGAGGAAGTAA	N/A
ITS2	GCTGCGTCTTCATCGATGC	N/A
Vahl730F	TAATACTGCTGTAG TAAAACGCCC	N/A
R-1200	CCCGTGTGAGTCAAATTAAGC	N/A
Software and algorithms		
QIIME2	Qiime2 website	http://qiime2.org/
R	CRAN	https://cran.r-project.org/
Rstudio	posit	https://posit.co/download/rstudio-desktop/
sparCC	github	https://github.com/bio-developer/sparcc

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to the lead contact, Carmen Buchrieser (cbuch@pasteur.fr).

Materials availability

This study did not generate new unique reagents.

Data and code availability

DNA sequencing data have been deposited at European Bioinformatics Institute database and are publicly available as of the date of publication. Accession numbers are listed in the [key resources table](#).

This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human participants

Written informed consent was obtained for all participants. Ethics approval was obtained from the national review board for biomedical research (Comité de Protection des Personnes Sud Est IV, France; ID-RCB 2016-A01021-50) in September 2016 and from Agence Sanitaire de Sécurité du médicament et des produits de santé (ANSM) in August 2016, and the study was registered on [ClinicalTrials.gov](https://www.clinicaltrials.gov) (NCT03064737). The study included hospitalized patients suffering from Legionnaires' disease pneumonia caused by *L. pneumophila* serogroup 1 (Lp1) diagnosed under medical and biological criteria. 69 tracheal aspirates, 23 sputum and 5 bronchoalveolar lavage fluids were taken from patients with clinical and/or radiological signs of pneumonia combined with isolation of *L. pneumophila* and/or positive *Legionella* PCR from broncho-pulmonary secretions and/or positive *Legionella* urinary antigen tests (UAT). For each respiratory sample, the presence of *Legionella* was tested by PCR using primers and probes of the R-DiaLeg kit (Diagenode, Belgium). PCRs were prepared in a final volume of 20 μ L by adding the Master Mix 5X (TaqMan Probe LC2.0, Roche Diagnostics, France), 4 μ L of primers and probe Lsp&Lp (R-DiaLegTM) or the internal control (DICD-YDL100) 2 μ L, H₂O (PCR grade) 2 μ L and 10 μ L of sample DNA. The amplification was performed on an LC2 system with the program: 95°C for 10 min followed by 45 cycles of 95°C for 10 s, 60°C for 40 s and 72°C for 1 s, and a final step of 30 s at 40°C. The samples were stored at -80°C until processing. A total of 38 patients were included and followed over time, considering the inclusion day as day 0. [Table S1](#) details the main clinical information of the patients (related to [Table 1](#)).

METHOD DETAILS

DNA extraction and marker genes sequencing

DNA was extracted from 300 μ L of the sample using a PowerSoil DNA isolation kit (Qiagen) following the manufacturer's instructions. To analyze the bacterial composition, the V4 region of the 16S rRNA gene was amplified by PCR using the 16SrRNA Illumina sequencing standard primers with adaptors. Forward primer 515F with adaptor (5'-TCGTCGGCAGCGTCAGATGTGTA TAAGAGACAG-3') and reverse primer 806R with adaptor (5'-GTCTCGTGGGCTCGGAGATG TGTATAAGAGACAG-3'). For each sample, a 20 μ L PCR mix was prepared, containing 5 μ L of Buffer Taq (10X), 1 μ L of 25 mM MgCl₂, 0.5 μ L of dNTPs (10 mM), 1.25 μ L of each primer (10 mM), 0.25 μ L of Phusion High-Fidelity DNA Polymerase (5u/ μ L), 0.5 μ L of DMSO, 8.25 μ L of nuclease-free water and 1 μ L of DNA template. PCR conditions: 95°C for 5 min followed by 25 cycles of 95°C for 30 s, 55°C for 1 min and 72°C for 1 min, and a final extension step of 7 min at 72°C. For fungal composition, we amplified the ITS region with the ITS1/ITS2 primers. The same adaptors and reaction mix and PCR conditions as described above were used, with 30 cycles.

To analyze the protozoa composition Vahl730F_C and R-1200 primers were used. The same sequencing adaptors and reaction mix described above were used for the PCR with the conditions: 95°C for 5 min followed by 30 cycles of 95°C for 30 s, 57°C for 1 min and 72°C for 1 min, and a final extension step of 7 min at 72°C. All amplicons were checked by electrophoresis in agarose gels (1.4%). The Illumina libraries were prepared following the manufacturer's instructions. High-throughput sequencing was performed with a MiSeq Illumina sequencer (2 \times 300 bp) by the Biomics Pole (Institut Pasteur).

Quantification of bacterial and fungal load

To quantify microbial abundance (bacteria and fungi), we set up the protocol proposed by Tkacz and colleagues for pulmonary samples²⁷ (see details in the [Figure S1](#) and extended experimental procedures). Briefly, the method consists of adding synthetic, chimeric DNA spikes to the samples (pSpike-P for prokaryotes and pSpike-F for fungi). The spikes then are co-isolated and PCR-amplified, providing estimations of the absolute abundance of the amplicons analyzed. The microbial abundance in pulmonary samples was estimated following microbial rRNA abundance = (number of microbial-origin reads/number of synthetic-origin reads) \times synthetic spike copies added to sample before DNA isolation for 16S rRNA with the prokaryotic synthetic spike and ITS with a synthetic fungal spike.

Controls

Kit and PCR negative controls were included for all primers used. It also included a positive extraction control based on a mock community (ZymoBIOMICS Microbial Community). See control results in [extended experimental procedures](#).

QUANTIFICATION AND STATISTICAL ANALYSIS

Marker gene data processing, ecological and statistical analysis

The microbiome data analysis was following previously published guidelines.⁶⁴ Specifically, the Quantitative Insights into Microbial Ecology (QIIME2) pipeline (version qiime2-2020.8) was used to perform the quality filtering, discard chimeric sequences, merge the paired-end reads and estimate the ASVs.⁶⁵ The ASV calculation was based on DADA2 implemented in QIIME2.⁶⁶ We used the SILVA database for 16S rRNA and 18S rRNA taxonomic assignment and the UNITE for ITS data, both in the QIIME2 pipeline. Samples with less than 1000 reads were excluded from the ecological analyses.

The alpha-diversity prediction was based on the microbial abundance, the richness (Chao 1), the diversity (Shannon Index), and dominance (Berger–Parker index) of the samples (see [Figures 1, 3, and 5](#) legends). The analyses were performed through the “microbiome” R package.⁶⁷ To statistically compare the metrics between sample groups, we used the Wilcoxon signed-rank test with Bonferroni correction to adjust for multiple comparisons for qualitative variables and Spearman correlation analyses for quantitative ones, both implemented in R core (see [Figures 1, 3, and 5](#) legends). To explore differences in composition between sample groups we used the non-metric multidimensional scaling (NMDS) based on Bray–Curtis dissimilarity (see [Figures 1, 3, and 5](#) legends). The multivariate analysis of variance based on dissimilarities (PERMANOVA test, `adonis2` R function) was used to test the impact of external variables in explaining differences between groups regarding microbial composition (see [Figure 4](#) legend). Both functions are implemented in the R package “Vegan”.⁶⁸ Only clinical factors with a minimum of 3 patients with metadata collected were analyzed. To identify those taxa associated with different groups when comparing among conditions, we used the ANCOM-BC method adjusted by patient as main co-variate⁶⁹ (see [Table S3](#)). The co-variate patient was the most relevant in explaining variance in composition; thus, the ANCOM-BC comparisons need to be adjusted by the patient.

We used SparCC software⁷⁰ to predict a co-occurrence network based on the most abundant bacteria (minimum of 10^3 copies per mL on average) and 10 copies per mL when the fungal community was included (see [Figures 2 and 6](#) legends). Correlation coefficients were estimated from the abundance table with 100 iterations. Only positive coefficients were considered for the plot (>0.5). To consider a correlation significant, we selected p values <0.01 with 500 bootstraps. We used the `igraph` package implemented in R software to plot the networks (force-directed layout option) and clusters based on the Newman–Girvan algorithm.⁷¹

Extended experimental procedures

Set-up of quantification method in pulmonary samples for 16S rRNA and ITS

We adjusted the protocol proposed by Tkacz and colleagues to quantify microbial abundance (bacteria and fungi) in pulmonary samples.²⁷ The method was tested for 16S rRNA gene analysis by spiking 30 samples (300 μ L) with 3.28×10^7 copies of plasmid “synthetic spikes” before DNA extraction to quantify bacteria. The estimated samples showed abundances ranging mainly from 10^6 to 10^{10} copies per mL of sample ([Figure S1A](#)). 41% of the samples belonged to a range of 10^7 copies per mL, while 21% have around 10^6 copies and 38% more than 10^8 . In three samples, we did not find spike-in reads what indicates a high microbial abundance impossible to estimate with this spike-in amount or experimental manipulation error from the last group. It is useful to test spike-in quantities to find a level that favors quantifying accurately as many samples as possible. Out test showed that in samples with abundances in the range of 10^6 copies per mL representing 21% of the analyzed set, the microbial reads represented less than 8% of the total, compromising further diversity analysis. Hence, we spiked-in the rest of the study samples with a lower number of copies of the plasmid (3.16×10^6).

Quantitative real-time PCR (qPCR), based on the 16S rRNA gene, was performed to quantify the total bacteria to compare with the spike-in estimates. We used the primers: forward 515F (5'-GTGCCAGCMGCCGCGGTAA-3') with adaptor (5'-TCGTCGGCAGCGT CAGATGTGTA TAAGAGACAG-3') and reverse 806R (5'-GGACTACHVGGGTWTCTAAT-3') with adaptor (5'-GTCTCGTGGGC TCGGAGATG TGTATAAGAGACAG-3'). Standard curves were estimated by using serial 10-fold dilutions of a plasmid with 1 copy of the gene. The PCR mixture was prepared in a final volume of 20 μ L by adding 10 μ L of $5 \times$ SYBR green qPCR master mix (Applied Biosystems), 0.8 μ L of primers (10 mM), 3.4 μ L of H₂O, and 5 μ L of the DNA sample. The amplification was performed on a Bio-Rad CFX qPCR instrument using the following program: 95°C for 3 min, followed by 40 cycles of 95°C for 5 s, 55°C for 30 s, and 72°C for 1 s. All reactions, including negative controls, were run in triplicate.

We found no significant differences between the qPCR and spike-in-based predictions (Wilcoxon signed-rank test paired test; p value >0.05). It indicates that measures in the predicted ranges are comparable to qPCR predictions commonly applied in microbiome studies. Thus, the spike-in method quantifies precisely microbial abundance from pulmonary samples when adding before DNA extraction.

For ITS, we tested by spiking 30 samples (300 μ L) with 4.16×10^3 copies of plasmid “synthetic spikes” before DNA extraction to quantify fungi. The estimated samples showed abundances ranging from 10^4 to 10^8 copies per mL of the sample with a median of 10^5 ([Figure S1B](#)). Out test showed that the fungal reads represented as a minimum the 70% of the total in all samples allowing to perform the diversity analyses. Then, we spiked-in the rest of the studied samples with the tested number of copies of the plasmid (4.16×10^3). We found no significant differences between the qPCR and spike-in-based predictions (Wilcoxon signed-rank test paired test; p value >0.05). It indicates that as for the 16S rRNA, spike-in measures are comparable to qPCR predictions. The qPCR was performed as for the 16S rRNA but with ITS primers (5'-CTTGGTCATTTAGAGGAAGTAA-3') and reverse primer (5'-GCTGCGTTCTTCATC GATGC-3').

Control analysis in pulmonary samples

All negative controls produced negative bands in agarose gel. Although their quantities for the library were low, we included them for sequencing. Kit and PCR controls of 16SrRNA and ITS carried out in parallel produced a low number of reads (less than 50 on average), indicating that contaminants did not represent enough biomass to affect the samples.

The kit and PCR reagents control for the 18S rRNA showed less than 200 reads each, all of them classified as unknown eukaryotes or from human origin but not *Trichomonas*, indicating that its presence in the samples is not a result of contamination.

The positive control was based on a mock community (ZymoBIOMICS Microbial Community): *Pseudomonas*, *Escherichia*, *Salmonella*, *Lactobacillus*, *Enterococcus*, *Staphylococcus*, *Listeria*, and *Bacillus*. We found all the taxa and the expected relative abundance per each taxon (12%) compared to our prediction was not statistically significantly different (Wilcoxon, p value >0.05).