Title

Lung Microbiota Is Related to Smoking Status and to Development of Acute Respiratory Distress Syndrome in Critically Ill Trauma Patients

Journal

American Journal of Respiratory and Critical Care Medicine, 197(5)

ISSN

1073-449X

Authors

Panzer, Ariane R
Lynch, Susan V
Langelier, Chaz
et al.

Publication Date

2018-03-01

DOI

10.1164/rccm.201702-0441oc

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

Lung Microbiota is Related to Smoking Status and to Development of ARDS in Critically Ill Trauma Patients

Ariane R. Panzer[1], Susan V. Lynch[1], Chaz Langelier,[2] Jason D. Christie,[3] Kathryn McCauley[1], Mary Nelson[4], Christopher K. Cheung[4], Neal L. Benowitz[5], Mitchell J. Cohen[6], Carolyn S. Calfee[7,8,9]

Affiliations:
[1] Division of Gastroenterology, Department of Medicine, University of California, San Francisco; San Francisco, CA.
[2] Division of Infectious Diseases, Department of Medicine, University of California, San Francisco; San Francisco, CA
[3] Division of Pulmonary and Critical Care Medicine, Department of Medicine, University of Pennsylvania; Philadelphia, PA
[4]Department of Surgery, University of California, San Francisco and Zuckerberg San Francisco General Hospital; San Francisco, CA
[5]Division of Clinical Pharmacology, Department of Medicine, University of California, San Francisco and Zuckerberg San Francisco General Hospital; San Francisco, CA
[6] Department of Surgery, Denver Health Medical Center; Department of Surgery, University of Colorado
[7] Division of Pulmonary and Critical Care Medicine, Department of Medicine, University of California, San Francisco; Department of Anesthesia, University of California, San Francisco; San Francisco, CA
[8] Cardiovascular Research Institute, University of California, San Francisco; San Francisco, CA
[9] Center for Tobacco Control Research and Education, University of California, San Francisco; San Francisco, CA

Corresponding author: Carolyn Calfee, MD, MAS
University of California, San Francisco
Pulmonary and Critical Care Division
505 Parnassus Avenue, Box 0111
San Francisco, CA 94143-0111, USA
Tel. (415) 476-5756
Fax (415) 476-5712
Email: carolyn.calfee@ucsf.edu

Keywords: Microbiota, cigarette smoking, trauma, 16S rRNA gene, acute respiratory distress syndrome, acute lung injury

Word Count: 4318

Funding: HL110969, HL133390 (CSC); HL115354, HL087115 (JDC). Analytical chemistry was supported in part by US Public Health Service Grant DA12393 from the

National Institute on Drug Abuse and the UCSF Helen Diller Comprehensive Cancer Center.

Author Contributions:

Conception and design: ARP, SVL, JDC, NB, MJC, CSC

Acquisition, analysis or interpretation of data: ARP, SVL, CL, JDC, MN, CKC, NLB, MJC, CSC

Drafting of work and/or revising for important intellectual content: ARP, SVL, CL, JDC, MN, CKC, NLB, MJC, CSC

Final approval of version to be published; agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved: ARP, SVL, CL, JDC, MN, CKC, NLB, MJC, CSC

At A Glance Commentary

Scientific knowledge on the subject: Cigarette smoking has been associated with ARDS development in multiple settings, but the mechanism of this association is not known. Smoking is associated with changes in the oropharyngeal and nasopharyngeal microbiota, but the impact of smoking on the lung microbiota in critically ill patients has not been studied.

What this study adds to the field: In this pilot study, we analyzed endotracheal aspirates from 74 critically ill, mechanically ventilated patients after severe blunt trauma and found that smoking status was associated with lung bacterial community composition at ICU admission, including enrichment for potentially pathogenic organisms. ARDS development was associated with lung bacterial community composition at 48 hours, and specific taxa enriched in smokers at baseline were found

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

to persist and be enriched at 48 hours in the lungs of patients who developed ARDS. To our knowledge, this is the largest cohort of critically ill patients to undergo lung microbiota analyses. These findings support further study of how smoking-related changes in lung microbiota taxa could be related to ARDS development after severe trauma.

ABSTRACT

Rationale: Cigarette smoking is associated with increased risk of acute respiratory distress syndrome (ARDS) in patients following severe trauma; however, mechanisms underlying this association are unknown.

Objectives: To determine whether cigarette smoking contributes to ARDS development following trauma by altering community composition of the lung microbiota

Methods: We studied lung microbiota of mechanically ventilated patients admitted to the intensive care unit (ICU) following severe blunt trauma using 16S rRNA gene amplicon sequencing of endotracheal aspirate (ETA) samples obtained upon ICU admission (n=74) and at 48 hours post-admission (n=30). Cigarette smoke exposure (quantified using plasma cotinine), ARDS development, and other clinical parameters were correlated with lung microbiota composition.

Results: Smoking status was significantly associated with lung bacterial community composition at ICU admission (PERMANOVA; p=0.007) and at 48 hours (PERMANOVA; p=0.03), and with significant enrichment of potential pathogens including *Streptococcus*, *Fusobacterium*, *Prevotella*, *Haemophilus* and *Treponema*. ARDS development was associated with lung community composition at 48 hours (PERMANOVA; p=0.04) and was characterized by relative enrichment of *Enterobacteriaceae* and of specific taxa enriched at baseline in smokers, including *Prevotella* and *Fusobacterium*.

Conclusion: In patients following severe blunt trauma, smoking is related to lung microbiota composition both at the time of ICU admission and at 48 hours. ARDS development is also correlated with respiratory microbial community structure at 48 hours and with taxa that are relatively enriched in smokers at ICU admission. The data from this pilot study suggest that smoking-related changes in the lung microbiota could be related to ARDS development after severe trauma.

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

INTRODUCTION

While supportive measures including lung protective ventilation and fluid conservative therapy have improved outcomes for acute respiratory distress syndrome (ARDS) over the past decade, there are still no known specific pharmacotherapies for ARDS. This lack of treatment options emphasizes the need for improved preventative approaches(1) and highlights the importance of understanding the role of environmental exposures in ARDS development. We previously reported that both active smoking and moderate to heavy passive smoking are associated with an increased risk of ARDS in patients following severe trauma, independent of other ARDS risk factors.(2) Similar associations with smoking have since been reported in non-pulmonary sepsis-related ARDS, transfusion-related ARDS, and in primary graft dysfunction (ARDS following lung transplantation).(3-5) While these studies provide strong epidemiologic evidence that cigarette smoke exposure may prime the lung for subsequent ARDS development, the mechanisms underlying this association remain unknown.

Cigarette smoke exposure has significant effects on lung epithelial permeability, endothelial injury, and lung inflammation, mimicking the pathologic abnormalities found in patients with ARDS.(6) Smoking may also prime the lung for ARDS by influencing the composition of the resident respiratory microbial communities (microbiota). Recently, a large study of respiratory samples from infants revealed that the existence of compositionally distinct airway microbiota was related to significant differences in susceptibility to lower respiratory tract infection (7). Cross-sectional comparisons of active smokers and non-smokers have demonstrated that smoking enriches the sub-gingival, oropharyngeal and nasopharyngeal microbial communities for potentially pathogenic species (8-10). There have been relatively few studies, however, examining the impact of smoking on lower airway microbiota, and the literature to date has been limited by relatively shallow community analyses and inconclusive results (10-11). One

study examining the lung microbiota of 3 healthy never-smokers, 7 smokers without active respiratory disease, and 4 smokers with chronic obstructive pulmonary disease (COPD) (11) found considerable overlap in bacterial community membership between all groups. In another study of 45 non-smokers and 19 current smokers, the mouth microbiota was found to differ in smokers versus non-smokers, but lung communities were not found to differ between these groups. (12)

In this pilot study, we used next generation sequencing technology to test the hypothesis that exposure to cigarette smoke, as quantified by plasma cotinine (the main proximate metabolite of nicotine and a biomarker of tobacco exposure), is associated with altered respiratory bacterial community composition. Specifically, we hypothesized that cigarette smoke exposure is associated with an increase in relative abundance of pathogenic genera in the lung microbiota, which may promote ARDS development. In addition, we evaluated compositional dynamics of the lung microbiota in mechanically ventilated trauma patients over the first 48 hours of ICU hospitalization. Some of the results of this study were previously presented in the form of an abstract.(13)

METHODS

Patients

We studied intubated and mechanically ventilated patients with severe blunt trauma admitted to the ICU at Zuckerberg San Francisco General Hospital (Level 1 trauma center for the City and County of San Francisco) between November 2012 and October 2013. Adult patients with blunt trauma who met criteria for full trauma team activation and were subsequently intubated and admitted to the ICU were eligible for inclusion. The Institutional Review Board of the University of California, San Francisco approved the research protocol for this prospective cohort study and granted a waiver of initial consent for the blood and endotracheal aspirate (ETA) sampling. Informed consent was

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

subsequently obtained from patients or their surrogates for continued study participation, as previously described.(2)

Sample collection

Blood was drawn within 10 minutes of emergency department arrival and again at 2, 4, 6, 12, 24 and 48 hours, as previously described.(2, 14) Samples were centrifuged and processed immediately, with plasma stored at -80°C for batch processing. ETA samples were obtained following standard ICU protocol using a new, sterile in-line suction catheter, both at the time of ICU admission and again approximately 48 hours later. Specimens were immediately stabilized in RNAlater, stored overnight at 4°C, and then at -80°C until further processing.

Data Collection and Outcome Measures

Clinical data were collected prospectively on patient demographics, time and mechanism of injury, comorbidities, injury severity as measured by the Injury Severity Score,(15) and prevalence and severity of organ dysfunction.(16) Data on smoking history and alcohol use were collected from the medical record and from the patient and/or surrogate, using a standardized questionnaire as in our prior work.(2) ARDS was defined using Berlin criteria(17) and adjudicated by two-physician review of all chest radiographs for each of the first 8 ICU days for all patients with any PaO2/FiO2 ratio < 300 mm Hg.

Cotinine Measurements and Protein Biomarker Assays

Plasma cotinine was measured on samples drawn in the emergency department using liquid chromatography/tandem mass spectrometry, as previously described.(18) The limit of quantitation was 0.08 ng/ml. For the few patients who did not have a plasma sample from 0h, plasma from 2h, 4h or 6h was used. Patients were classified as

smokers, passive smokers, or non-smokers using previously validated cotinine cutpoints; these cut-points were selected a priori as part of the main study hypothesis.(2, 19) Enzyme linked immunoassays (ELISA) were used to measure protein biomarkers of lung injury in plasma from 0h and 48h. Additional details regarding ELISA measurements are in the online supplement.

16S rRNA gene-based ETA Microbiota Profiling using Illumina MiSeq

ETA samples were maintained at -80°C until processing, full details of which are provided in the online supplement. Following total DNA extraction, PCR amplification of the 16S rRNA gene variable region 4 (V4) was performed in 25-µl reactions using barcoded 515F/806R primers and either 30 ng/ul of DNA template or DNA extraction buffer as negative controls. All samples amplified, and amplicons were purified using SPRI beads (Beckman Coulter), analyzed for purity using the Agilent 1000 DNA Kit and Bioanalyzer, quantified using the Qubit HS dsDNA kit (Invitrogen), and pooled at 50 ng per sample. The pooled library, including negative controls, was quantified using the Qubit HS dsDNA kit, diluted to 2 nM, denatured, and 5pM was loaded onto the Illumina MiSeq cartridge (V3) in combination with a 15% (v/v) of denatured 12.5pM PhiX for sequencing (50,000 reads per sample). Paired-end sequence reads were combined using Fast Length Adjustment of SHort reads (FLASh) version 1.2.7. Additional details regarding operational taxonomic unit (OTU; taxa) picking and data analyses are provided in the online supplement.

Alpha diversity (within sample) indices were calculated using Quantitative Insights Into Microbial Ecology.(20) Beta-diversity (between sample composition dissimilarity) was assessed using Canberra, Unweighted Unifrac, and Weighted Unifrac distance matrices for 0 hour and 48 hour samples.(21, 22) Distance matrices were visualized using Emperor.(23) Cigarette smoke exposure, development of ARDS, and

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

other clinical data were analyzed using the Vegan version 2.0.10 (24) in R version 3.1.1 and Permutational Multivariate Analysis of Variance (PERMANOVA) for relationships with lower respiratory bacterial beta diversity (composition).(25) To test for significant differences in taxon abundance between smokers and non-smokers and patients with and without ARDS (ARDS+ and ARDS-), we simultaneously applied Poisson, negative binomial, and zero-inflated negative binomial regressions (three model approach) to taxon count data and tested for best fit on a taxon-by-taxon basis using a multiply rarefied OTU table. Following correction for false discovery (Benjamini-Hochberg), taxa that exhibited q value < 0.1 and a p value < 0.05 were considered significant and these data were illustrated on phylogenetic trees using iTOL version 3.0.(26) Procrustes and Mantel test analyses, combining the 0 and 48 hour coordinate matrices, were performed in QIIME to examine relative relationships in community composition over time.

RESULTS

Patient Demographics

We studied 74 mechanically ventilated patients admitted to the ICU following severe blunt trauma using ETA samples obtained at the time of ICU admission (n=74), and at 48 hours post admission (n=30, including 25 patients who were studied at both time points). Plasma cotinine levels of the 74 patients with 0 hour samples indicated that 21 subjects were non-smokers, 15 were passive smokers, and 38 were smokers. Smokers were younger than non-smokers and more commonly had a history of alcohol abuse, compared to non-smokers and passive smokers. Further demographic details about these patients are provided in **Table 1**.

Lung Microbial Composition at ICU Admission

We first examined 0 hour samples to characterize the baseline lung microbiota of trauma patients upon ICU admission. Alpha-diversity measures including bacterial community richness (number of taxa detected), evenness (distribution of taxa within a community), and diversity (Shannon's, Simpsons or Faith's phylogenetic index) did not significantly differ between non-smokers, passive smokers, and smokers (Kruskal-Wallis $p>0.05$).

We next analyzed whether smoking co-varied with lower airway microbial community composition (beta-diversity; **Table 2**). At baseline, smoking status based on cotinine level categorized into three groups (i.e. smokers, passive smokers, non-smokers) was significantly associated with microbial community composition in the lower airways (Univariate PERMANOVA; Canberra distance; $r^2=0.03$, $p=0.007$; **Figure 1A**). Cotinine categorized into two groups (i.e. smokers and passive smokers vs. non-smokers) was also significantly associated with microbial community composition in the lower airways (Univariate PERMANOVA; Canberra distance; $r^2=0.02$, $p=0.003$; **Table 2**). Additionally, the presence or absence of congestive heart failure, alcohol abuse based on chart history, and age were significantly associated with bacterial community beta-diversity, while season of ETA sample collection and cirrhosis diagnosis trended towards a statistically significant relationship (**Table 2**). These findings indicate that the presence or absence of lower abundance taxa represent a primary discriminatory feature of the microbial communities associated with these clinical factors. ARDS development during the first 8 days in the ICU was not significantly associated with overall lung bacterial community composition at 0 hours (**Table 2**).

Since smoking, age and alcohol abuse may be co-associated, we performed analyses of the association between smoking and microbiota composition adjusted for age and for alcohol abuse, using the varpart function in the Vegan package in R. Due to missing alcohol abuse data for a number of patients, the sample size for these analyses

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

was reduced to n=54. With this smaller sample size and both smoking and alcohol abuse in the model, neither were significantly associated with microbiota composition (**Table S1**). Since alcohol abuse data was frequently missing, reducing sample size and statistical power, and is likely not randomly missing, we also carried out a sensitivity analysis in which missing data was treated either as all "yes" responses or all "no" responses. In both the "yes" response model and "no" response model, the adjusted r^2 value for smoking by cotinine status was larger than the adjusted r^2 for alcohol abuse (**Table S1**). The smoking by cotinine r^2 value adjusted for age (r^2=0.00312 (p-value = 0.026)) was also higher than the age r^2 value adjusted for smoking by cotinine (r^2=0.0011 (p-value = 0.141)). Hence, we concluded that smoking is related to airway microbiota composition, but that additional factors such as alcohol abuse and age may also shape the microbial communities in the airways.

To identify specific bacterial groups enriched in the lower airways of smokers, we compared taxon relative abundance of 0-hour samples obtained from smokers and non-smokers (**Fig. 1B; Table S2**). Smokers exhibited significant enrichment of putative pathogens including taxa assigned to *Haemophilus* (*Proteobacteria*), *Prevotella* (*Bacteroidetes*), *Streptococcus (Firmicutes)*, *Fusobacterium* (*Fusobacteria*) and *Treponema* (*Spirochaetes*). Non-smokers exhibited enrichment of distinct taxa belonging to *Streptococcus* (*Firmicutes*), *Lactobacillus* (*Firmicutes*), *Rothia* (*Actinobacteria*) and *Neisseria* (*Proteobacteria*). These findings indicate that the composition of the endotracheal microbiota is related to smoking status and is characterized by the enrichment of specific bacterial taxa, some of which may possess pathogenic potential.

We also identified taxa enriched at 0 hours in patients with a "yes" response to alcohol abuse by chart history. Several taxa enriched in alcohol abusers were shared with those enriched in smokers at this time point including *Haemophilus* (*Proteobacteria*), *Prevotella* (*Bacteroidetes*), *Streptococcus (Firmicutes)*, *Fusobacterium*

(*Fusobacteria*), and *Porphyromonas* (*Bacteroidetes*) (**Table S2**; highlighted orange); however, the majority of taxa enriched in smokers were not enriched in alcohol abusers. Taxa enriched at 0 hours in both patients under 40 years of age (median age in this cohort) and smokers included distinct *Haemophilus* (*Proteobacteria*), *Streptococcus (Firmicutes)*, *Prevotella* (*Bacteroidetes*), *Fusobacterium* (*Fusobacteria*) (**Table S2;** highlighted light blue). Taxa enriched in all three categories in patients at 0 hours similarly included specific *Streptococcus (Firmicutes), Fusobacterium* (*Fusobacteria*), and *Prevotella* (*Bacteroidetes*) taxa (**Table S2**: highlighted pink).

Change in Lung Microbiota Composition Between 0 and 48 Hours

To examine lung microbiota dynamics over the first 48 hours post-ICU admission, we next studied community beta-diversity in 25 patients from whom paired ETA samples were obtained at 0 and 48 hours. At 48 hours, all patients, regardless of smoking status, exhibited a substantial compositional divergence relative to their baseline bacterial community composition (**Fig. 2A**). This observation was supported by both *Procrustes* and *Mantel* statistical analyses, which tested whether a significant relationship existed between bacterial community composition of paired 0 and 48 hour samples. Both of these analyses indicated that 0 and 48-hour overall bacterial community composition were not significantly related (Monte Carlo p=0.332; Mantel test p=0.251, **Fig 2A**). Calculated mean community distances between 0 and 48-hour samples further confirmed this finding; 48-hour samples exhibited significantly greater community composition variation (as indicated by greater mean distance), compared to those samples collected at the time of hospital admission ($p<0.0001$; **Fig. 2B**).

Lung Microbial Composition at 48 Hours After ICU Admission

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

We next examined all patients from whom 48-hour samples were collected (n=30) to identify factors that explained the observed variation in bacterial community composition at this time point. At 48 hours, the cohort consisted of 8 non-smokers, 6 passive smokers, and 16 smokers. The demographics of these patients are described in **Table S3**. At 48 hours, smoking by cotinine level categorized as three groups or two groups (as described above) was again significantly associated with the variance in lung bacterial community composition (Univariate PERMANOVA; Unweighted UniFRAC distance; r^2=0.093, p=0.03 and r^2=0.052, p=0.03 respectively; **Fig. S1**), explaining larger proportions of variance (9.3% and 5.2% respectively). This finding indicates that phylogeny and the presence or absence of specific taxa are the primary discriminatory features of 48-hour lung bacterial communities associated with smoking by cotinine level. Several taxa which were significantly enriched in smokers at 0 hours remained significantly enriched in smokers (versus non-smokers) at 48 hours post-ICU admission; these included specific *Haemophilus*, *Streptococcus*, *Prevotella*, *Porphyromonas* and *Campylobacter* taxa (**Table 3**). These data suggest that smoking exerts a persistent effect on the lower airway microbiota, characterized by the enrichment of specific taxa upon ICU admission that remain in high relative abundance at 48 hours post-admission.

Univariate PERMANOVA analysis based on a weighted UniFrac distance matrix (which considers both phylogenetic relationships and relative abundance) showed that ARDS development during the first 8 days in the ICU (**Fig. 3A**; r^2=0.08; p=0.04), Injury Severity Score (r^2=0.09; p=0.03) and presence or absence of asthma (r^2=0.08; p=0.04) were significantly related to lung bacterial community composition at 48 hours. This finding indicates that lung injury, severity of traumatic injury, and chronic airway disease are related to both bacterial phylogeny and taxon relative abundance in the lower airways following 48 hours of hospitalization. It should be noted that by 48 hours, 80% of patients had received parenteral or enteral antimicrobial therapy. However, antibiotic

administration (categorized by antibiotic class or by spectrum of activity) was not significantly associated with lung community composition (r^2 for antibiotic class 0.04, p=0.08; r^2 for antibiotic spectrum 0.04; p=0.29; analyses by unweighted Unifrac; similar findings for weighted Unifrac and Canberra), and patients who developed ARDS did not differ in the antibiotic class (p = 0.47) or spectrum of antibiotic activity (p = 0.1) administered over the initial 48 hours, compared to those who did not develop ARDS.

Lung Microbial Community Composition Associated with ARDS Development

We next determined taxonomic differentials that characterized the 48-hour lung microbiota of patients who developed ARDS during the first 8 days in the ICU (ARDS+; n=17) and those who did not develop the syndrome during that time period (ARDS-; n=13; **Fig 3B; Table S4**). We hypothesized that some of these ARDS-associated taxa would be enriched in smokers versus non-smokers at baseline. ARDS+ patients were most significantly enriched for a specific taxon belonging to the *Enterobacteriaceae* **(**OTU 2119418; **Figure 3B)**; others exhibiting relative expansion in ARDS+ patients included specific *Prevotella* (OTU 4447248) and *Fusobacterium* (OTU 4319899) **(Figure 3B).** These latter two taxa were also amongst those significantly enriched in smokers at baseline (**Figure 1B; Table S2**).

Plasma Biomarkers and Lung Microbiota Composition

As an intermediate phenotype of lung injury, we measured selected plasma biomarkers of inflammation, lung epithelial injury and endothelial injury at 0 and 48 hours. These biomarkers (complete list in supplementary materials) have previously been associated with lung injury pathogenesis and/or outcomes. Using paired microbiota and biomarker data at 0 hours, soluble intercellular adhesion molecule-1 (ICAM), Interleukin 8 (IL-8), and vascular endothelial growth factor (VEGF) categorized into

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

quartiles were all found to be significantly associated with lung bacterial community variation (**Table 4**). Using paired microbiota and biomarker data generated from 48 hour samples, IL-6 and IL-8 categorized into quartiles were also significantly associated with community variation at this time-point (**Table 4**). Interestingly, 0 hour microbiota composition was significantly related to 48-hour levels of VEGF, receptor for advanced glycation end-products (RAGE), angiopoietin-2 (ANG-2), pentraxin 3 (PENT3), and IL-8, and these markers explained the largest proportion of variation in the baseline microbiota (**Table 4**). These data indicate that microbiota composition is related to contemporary markers of endothelial and epithelial injury and inflammation, and that baseline microbiota composition upon ICU admission is also related to markers of lung injury and inflammation observed following 48 hours of hospitalization.

DISCUSSION

This pilot study finds that cigarette smoking is associated with differences in lung microbial community composition with enrichment of specific taxa both at the time of ICU admission and at 48 hours post-admission. In addition, lung microbiota composition at 48 hours was associated with ARDS development and with injury severity. We also find that for all patients, bacterial community composition in the lung significantly diverged at 48 hours post-ICU admission from that observed upon admission to the ICU. To our knowledge, this study represents the largest cohort of critically ill patients to undergo sequence-based analyses of the lung microbiota, as well as the first study examining the impact of smoking on lung microbiota in the setting of critical illness.

Lung bacterial community composition at 0 hours and 48 hours was significantly associated with smoking status based on cotinine level. These relationships were observed using distinct distance matrices (Canberra at 0 hours and Unweighted UniFrac

at 48 hours), and were robust following adjustment for confounding factors. This finding indicates that while at baseline, smoking is related to lung bacterial communities that primarily differ based on the relative abundance of specific taxa, at 48 hours, this relationship is driven by the presence or absence of specific phylogenetic groups. Specific taxonomic enrichments associated with smokers at baseline included *Haemophilus*, *Streptococcus, Prevotella*, and *Campylobacter*. A small number of these taxa were also enriched in patients who abused alcohol and were younger, indicating that these factors may also shape airway community composition. At 48 hours, these same taxa were significantly enriched in smokers. These data suggest smoke exposure leads to increased relative abundance of specific bacterial taxa in the lung microbiota, which remain persistently enriched in smokers over time. In a recent study in infants, *Haemophilus*-dominated airway microbiota were associated with an odds ratio (OR) of 11 for acute respiratory infection (ARI), while *Streptococcus* dominated communities were associated with an (OR) of 2.0 for ARI (7). Thus, it is plausible that smoke exposure shapes the lung microbiota – or alternatively, shapes the oropharyngeal microbiota that subsequently migrates lower in the respiratory tract in the setting of intubation and critical illness - and enriches for specific potentially pathogenic organisms that increase susceptibility toward ARI.

Patients who developed ARDS over the first 8 days in the hospital also exhibited taxonomic differences in their lung microbial communities at 48 hours, compared to those who did not develop ARDS during that time period. ARDS+ patients were most significantly enriched for *Enterobacteriaceae* (OTU 2119418), which are found as both commensals and pathogens in the human gut. Although controversial,(27) some studies have suggested that trauma can lead to bloodstream translocation of gut microbiota, which offers a potential explanation for these findings.(28, 29) Alternatively, micro-aspiration (either prior to or despite endotracheal intubation) could potentially lead to

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

increased presence of gut microbiota members in the lungs of ARDS patients at 48 hours. Notably, these findings are also consistent with a recently published report that gut-specific bacteria were the dominant microbiota members in bronchoalveolar lavage fluid from humans with ARDS.(30) Other taxa associated with ARDS development at 48 h included specific *Fusobacterium* and *Prevotella* taxa also found to be significantly enriched in smokers at 0 hours. Members of both these genera common in oral microbiota have been found in anaerobic respiratory infections, in some cases associated with development of severe ARDS.(31, 32) Furthermore, both *Fusobacterium* and *Prevotella* have been shown to induce increased production of the pro-inflammatory cytokine IL-6.(33, 34) Lung microbial community composition at 0 hours, based on a distance metric that weights relative abundance, was related to endothelial injury (plasma ICAM-1), epithelial injury (VEGF), and inflammation (IL-8), whereas at 48 hours microbiota variation was associated with inflammation (IL-6 and IL-8). At 48 hours, the observed relationship between microbiota and inflammation was driven by the presence or absence of specific taxa, suggesting that loss or outgrowth of specific bacteria following the first 48 hours in the ICU, may contribute to inflammation observed at this time point. Additionally, bacterial beta-diversity upon ICU admission was related to severity of both lung injury and inflammation, as reflected by plasma biomarkers, at 48 hours. Taken together, these data suggest that smoke exposure is associated with a distinct bacterial community, enriched for potential pathogens that could increase susceptibility to acute respiratory infection and contribute to pulmonary inflammation and injury during hospitalization.

Compared to the relatively modest variation in lung community composition observed at admission, lung microbiota in patients after 48 hours of hospitalization significantly diverged. While we did not detect a significant association between antibiotic administration and composition of the lung microbiota at 48 hours, this analysis

was likely underpowered due to the wide variety of antimicrobials administered and the relatively small number of patients at the 48 hour time point. Studies that have observed a relationship between antibiotic administration and microbial community composition have had far larger patient numbers, as much as six times the number of patients as in this study.(35)

This study also identified taxon-level variation (i.e. different OTUs), but genus level conservation (e.g. *Neisseria* or *Streptococcus*) in smokers versus non-smokers. This finding suggests that smoking may select for specific species or strains within airway-adapted genera, potentially resulting in replacement of commensal strains with phylogenetically related but genomically distinct members of the same genera. Further exploration of this possibility will require metagenomic or targeted isolate sequencing efforts, as the 16S rRNA gene V4 region sequenced in this study cannot confidently assign identity at the species or strain level.

This pilot study has several strengths, including the use of next-generation sequencing to identify members of the lung microbiota, early sampling of patients at the time of ICU admission and subsequent sampling 48 hours later, meticulous clinical phenotyping for ARDS, and quantitative measurement of tobacco exposure using a well-validated biomarker. This pilot study also has some limitations. First, the sample size, though the largest to our knowledge published on this topic, remains relatively modest, both at time 0 and at 48 hours, limiting the insights gained from analyses adjusted for potential confounders. Future studies with a larger sample size and more complete data on alcohol abuse must be carried out to further discern the relationship between these variables, as it remains possible that alcohol abuse and/or age may be contributing to some of the observed differences between smokers and non-smokers. Similarly, the analyses of association between lung microbiota and both clinical variables and plasma biomarkers were not controlled for multiple comparisons due to the sample size and, as

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

such, merit replication in larger cohorts. Second, we were unable to obtain 48-hour specimens on all patients enrolled at baseline, as some patients died or were extubated between the two time points. Third, we used ETA rather than BAL specimens, which some investigators have argued may be superior for capturing the distal lung microbiota, although this is still controversial. For logistical reasons, it was not possible to obtain immediate BAL specimens on severely injured patients at the time of ICU admission. While there may be regional differences in the microbiota of the upper and lower airways and the alveolar compartments in the lung,(36) these microbial communities likely represent a continuum;(37, 38) thus, we think that ETA is a practical and useful approach for microbial community analysis in critically ill patients in whom more invasive procedures may be prohibitively risky or temporally impossible. Similarly, while ETA samples may be affected by colonization of the endotracheal tube (more so at 48 hours than at baseline), these samples have the advantage of bypassing the oropharynx, and there are no feasible alternative approaches to studying the lung microbiota in ventilated patients that do not pass through the endotracheal tube. Finally, although ARDS developed after 48 hours in most patients (11 out of 16), because 5 patients developed ARDS between 0 and 48h, it is possible that the link between lung microbiota and ARDS could be driven in part by ARDS development, either via antibiotic administration spurred by the development of radiographic opacities or by dysfunction of the alveolar-capillary barrier.

In conclusion, we identified significant differences in the lung microbiota composition of critically ill patients related to both smoking status and subsequent development of ARDS, as well as a marked divergence in bacterial community composition between 0 and 48 hours of ICU admission in critically ill blunt trauma patients. Further investigations with larger sample sizes, and ideally with concurrent

experimental models, will be needed to confirm the findings of this pilot and further evaluate the role of the lung microbiota in ARDS pathogenesis.

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

Table 1: Patient Demographics By Smoking Status As Determined by Plasma Cotinine (n=74)

	Non-Smokers (Cotinine < LOQ; n=21)	Passive Smokers (Cotinine > LOQ and ≤ 3.08 ng/ml; n=15)	Smokers (n=38; Cotinine > 3.08 ng/ml)	ANOVA p-value
Age, years	59 ± 23	43 ± 23	39 ± 13	**0.01**
Male gender, n (%)	11 (52)	10 (67)	30 (79)	0.11
Race, n (%)				0.08
Caucasian	12 (57)	13 (87)	28 (74)	
African American	0 (0)	0 (0)	5 (13)	
Asian	6 (29)	2 (13)	1 (3)	
N/A	3 (14)	0 (0)	4 (11)	
Hispanic ethnicity, n (%)	1 (5)	5 (33)	6 (16)	0.09
ARDS, n (%)	4 (19)	5 (33)	7 (18)	0.32
Injury Severity Score	29 ± 13	24 ± 21	20 ± 15	0.12
Medical History, n (%)				
Asthma	0 (0)	2 (13)	5 (13)	0.24
Cirrhosis	1 (5)	2 (13)	0 (0)	0.09
Congestive Heart Failure	0 (0)	1 (7)	0 (0)	0.14
COPD	0 (0)	0 (0)	2 (5)	0.40
Coronary Artery Disease	0 (0)	1 (7)	1 (3)	0.50
Diabetes	6 (29)	3 (20)	3 (8)	0.09
HIV	0 (0)	0 (0)	3 (8)	0.2
Malignancy	2 (10)	2 (13)	1 (3)	0.32
Alcohol abuse Composite*, n (%)				**0.001**
Yes	1 (5)	5 (33)	23 (61)	
No	14 (67)	7 (47)	13 (34)	
N/A	6 (29)	3 (20)	2 (5)	
ETA Sample Season, n (%)				0.93
Autumn	4 (19)	2 (13)	10 (26)	
Winter	8 (38)	3 (20)	10 (26)	
Spring	6 (29)	1 (7)	10 (26)	
Summer	3 (14)	9 (60)	8 (21)	

***** Defined as a chart history of alcohol abuse or an AUDIT score ≥ 8 for men, ≥ 5 for women (39, 40)
LOQ = Limit of Quantitation

Table 2: Significance of the relationship between clinical parameters and time-zero lung microbiota (n = 74)

Parameter	R^2	p value
Significant Meta Data		
Smoking status by plasma cotinine level (non-smoker or passive smoker/smoker)	0.02	0.003
Smoking status by plasma cotinine level (non-smoker, passive smoker, smoker)	0.03	0.007
Congestive Heart Failure	0.02	0.02
Alcohol Abuse by Chart History	0.02	0.03
Age	0.02	0.04
Demographics		
Race (Caucasian, African-American, Asian)	0.03	0.11
Insurance Source	0.04	0.51
BMI	0.01	0.66
Gender (Male/Female)	0.01	0.69
Hispanic ethnicity	0.01	0.99
Comorbidities		
Cirrhosis	0.02	0.06
Diabetes	0.02	0.13
Asthma	0.02	0.20
Chronic Obstructive Pulmonary Disease	0.01	0.55
HIV	0.01	0.62
Patient History		
ETA Sample Season	0.05	0.051
ETA Sample Month	0.16	0.07
Alcohol Abuse Composite (Y/N)*	0.02	0.07
Alcohol Use by Chart History (Y/N)	0.02	0.08
Antibiotics Prior to ETA Sample (Y/N)	0.01	0.21
Injury Severity Score	0.01	0.47
Pack Years	0.08	0.66
Outcomes		
Acute Respiratory Distress Syndrome during first 8 days in ICU (Y/N)	0.01	0.68

All analyses are univariate PERMANOVA.
* Defined as a chart history of alcohol abuse or an AUDIT score ≥ 8 for men, ≥ 5 for women (39, 40)

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

Table 3: Taxa significantly enriched in smokers compared to non-smokers at both 0 hours and at 48 hours

OTU	Phylum	Family	Genus	Average Relative Abundance[a] (%) Smokers 0 hr	Average Relative Abundance[a] (%) Non-Smokers 0 hr	Mean Read Count Smoker 0 hr	Mean Read Count Non-Smoker 0 hr	p-value[b]	Average Relative Abundance[a] (%) Smoker 48 hrs	Average Relative Abundance[a] (%) Non-Smokers 48 hrs	Mean Read Count Smoker 48 hr	Mean Read Count Non-Smoker 48 hr	p-value[b]
956702	*Proteobacteria*	*Pasteurellaceae*	*Haemophilus*	1.22	0.0024	732.89	1.43	1.48E-15	0.762	0.0008	456.59	0.5	1.60E-04
4471251	*Proteobacteria*	*Pasteurellaceae*	*Haemophilus*	6.85	0.1727	4108.19	103.57	4.39E-14	6.219	0.1514	3728.91	90.75	5.63E-05
4469359	*Proteobacteria*	*Pasteurellaceae*	*Haemophilus*	0.58	0.0188	346.98	11.29	1.53E-12	0.588	0.0140	352.27	8.38	1.11E-04
4416265	*Proteobacteria*	*Pasteurellaceae*	*Haemophilus*	0.05	0.0017	29.4	1	1.88E-11	0.044	0.0019	26.32	1.13	2.98E-03
4443574	*Proteobacteria*	*Pasteurellaceae*	*Haemophilus*	0.10	0.0060	62.85	3.62	6.24E-08	0.109	0.0019	65.23	1.13	1.11E-03
4407433	*Proteobacteria*	*Pasteurellaceae*	*Haemophilus*	0.01	0.0005	7.87	0.29	1.57E-05	0.013	0.0004	8.09	0.25	4.30E-03
1059655	*Firmicutes*	*Streptococcaceae*	*Streptococcus*	5.93	0.7744	3557.62	464.33	2.83E-05	0.365	0.0527	218.59	31.63	0.014
3678349	*Firmicutes*	*Streptococcaceae*	*Streptococcus*	0.33	0.0458	199.13	27.48	3.61E-05	0.568	0.0492	340.73	29.5	1.28E-03
4428060	*Bacteroidetes*	*Prevotellacea*	*Prevotella*	0.06	0.0054	38.47	3.24	0.0005	1.609	0.0225	964.95	13.5	1.39E-04
747980	*Bacteroidetes*	*Prevotellacea*	*Prevotella*	0.21	0.0325	124.68	19.48	0.002	0.355	0.0015	212.82	0.88	6.65E-09
4423790	*Bacteroidetes*	*Porphyromonadaceae*	*Porphyromonas*	0.18	0.0357	109.53	21.43	0.003	0.238	0.0240	142.55	14.38	0.005
32546	*Proteobacteria*	*Campylobacteraceae*	*Campylobacter*	0.02	0.0027	11.53	1.62	0.012	0.009	0.0002	5.55	0.13	0.01

[a] Relative abundance calculated as follows: (read count of taxa from sample/read depth samples were normalized to) * 100. The average of these values is presented in the table.

[b] p-value calculated via three model approach which considers mean read counts of OTUs for smokers compared to non-smokers.

Table 4: Plasma biomarkers significantly associated with lung microbiota composition

Lung Microbiota Time Point	Plasma Biomarker, Time Point	R^2	p-value
Baseline (0 hours)	ICAM-1, 0 hour†	0.02	0.02
	IL-8, 0 hour†	0.02	0.04
	VEGF, 0 hour†	0.02	0.04
	VEGF, 48 hours†*	0.07† 0.09*	0.03† 0.02*
	RAGE, 48 hours*	0.08	0.04
	Ang-2, 48 hours**	0.17	0.02
	Pentraxin-3, 48 hours**	0.15	0.02
	IL-8, 48 hours**	0.13	0.04
48 hours	IL-6, 48 hours*	0.08	0.01
	IL-8, 48 hours*	0.07	0.02

All biomarkers analyzed in quartiles, using univariate PERMANOVA.

* Unweighted UNIFRAC (distance matrix weighted by phylogenetic relationships)
** Weighted UNIFRAC (distance matrix weighted by both phylogenetic relationships and relative abundance)
† Canberra (distance matrix that weights rare taxa)

Abbreviations: ICAM (intercellular adhesion molecule); IL (interleukin), VEGF (vascular endothelial growth factor), RAGE (receptor for advanced glycation end-products), Ang (angiopoietin)

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

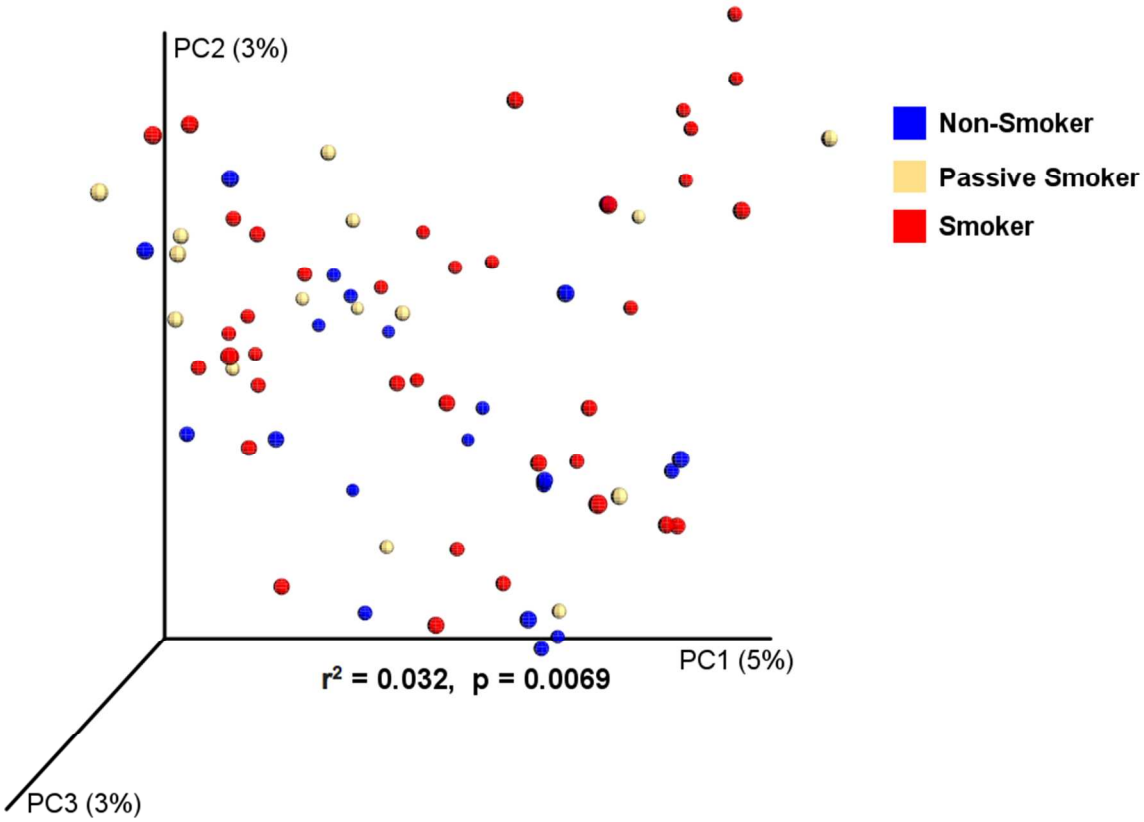
Figure Legends

Fig. 1A. PCoA plot (Canberra distance matrix) of 0-hour samples colored by smoking status based on cotinine level. The r^2 value represents the percent variance explained, and the p-value denotes the significance of this percentage. **B.** Discriminatory taxa (0 hour 3 model analysis comparing smokers vs non-smokers) labeled by genus. * *Fusobacterium* OTU4319899, ** *Prevotella* OTU4447248

Fig. 2A. Procrustes analysis comparing 0-hour and 48-hour sample Weighted Unifrac coordinate matrices (n = 25). **2B**. A comparison of mean community distance between all 0 hour samples on the left and between all 48 hour samples on the right, indicating that 48 hour microbiota exhibit significantly greater mean distance between samples, i.e. greater microbiota compositional variability at this time point.

Figure 3A. PCoA plot (weighted Unifrac distance) of 48-hour samples colored by ARDS status. The r^2 value represents the percent variance explained and the p-value denotes the significance of this percentage **3B**. Discriminatory taxa (48 hour 3 model analysis comparing ARDS+ to ARDS- patients) labeled by genus.

Fig 1A.



AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

Fig 1B.

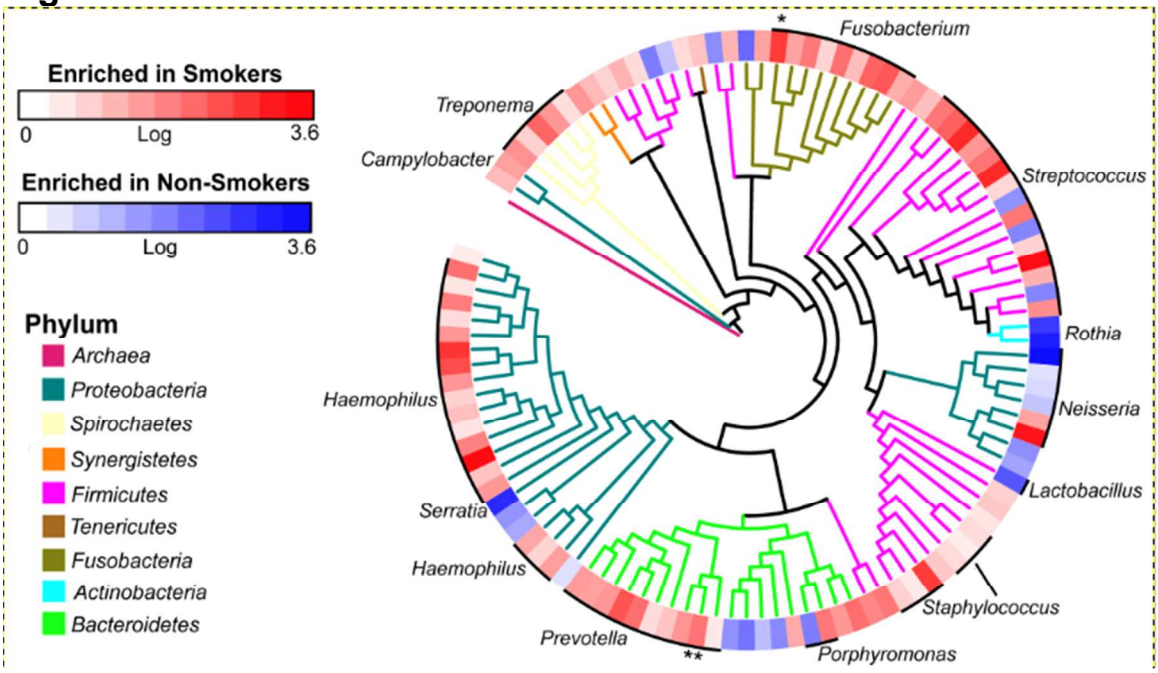
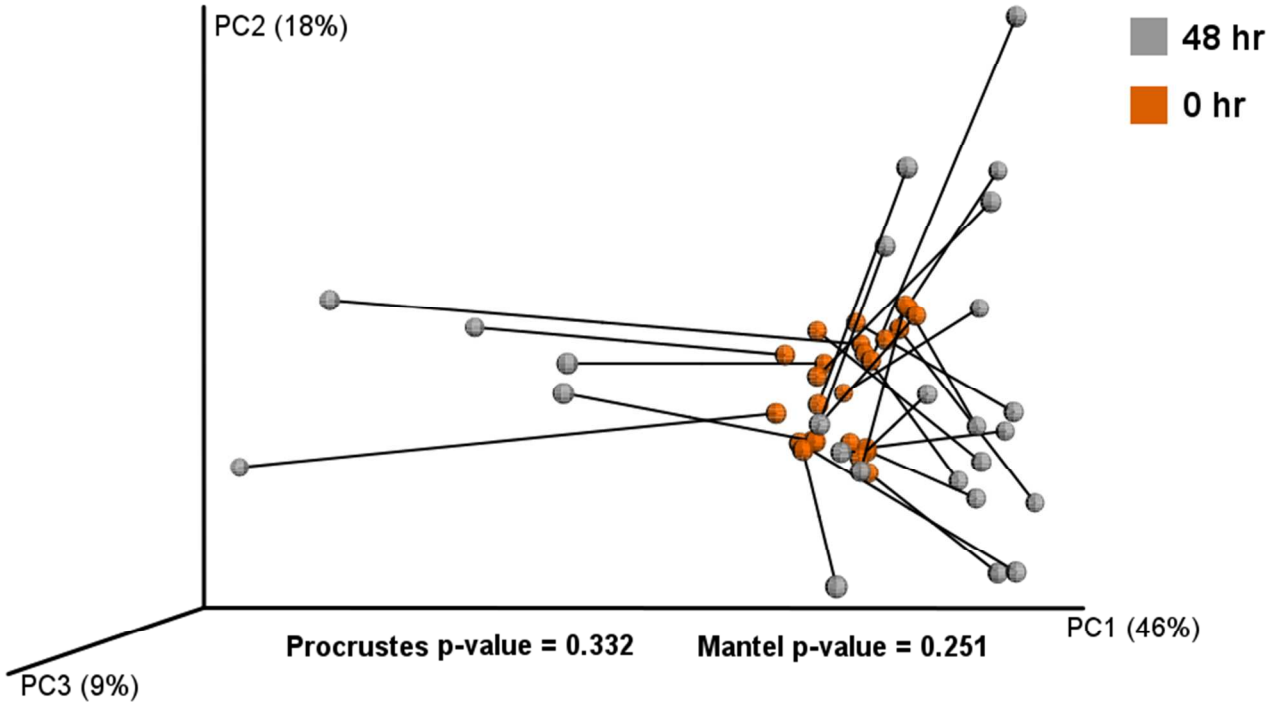


Figure 2A.



AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

Figure 2B

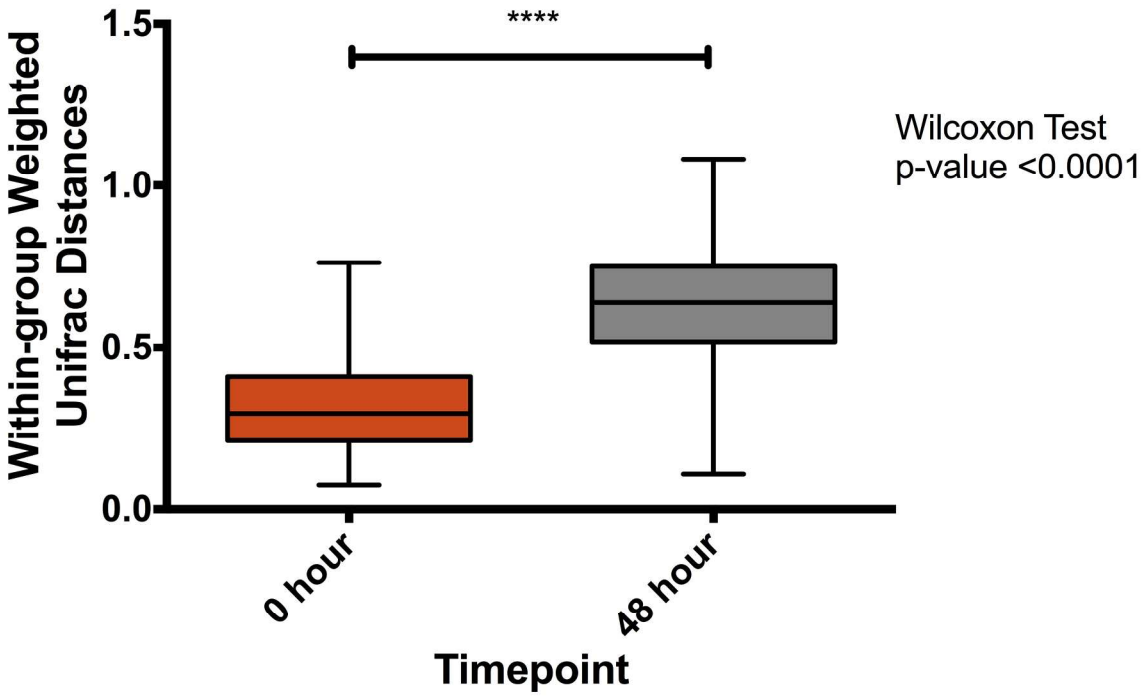
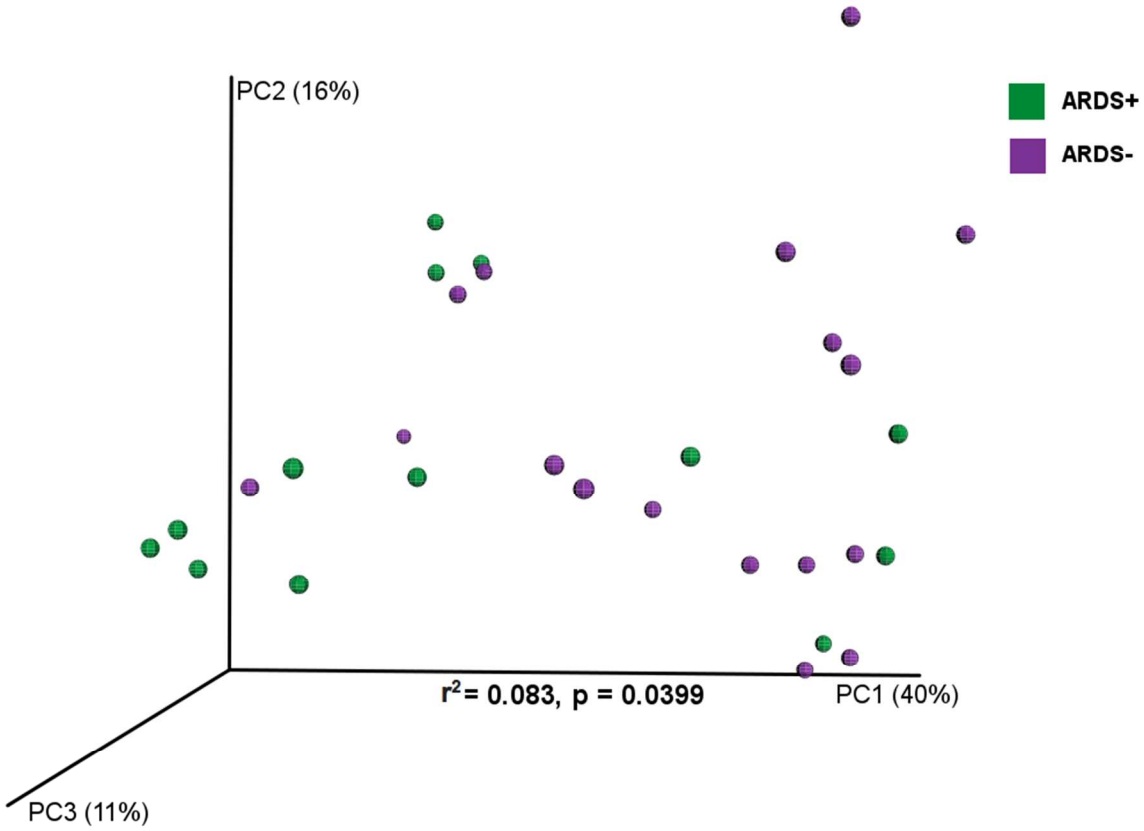
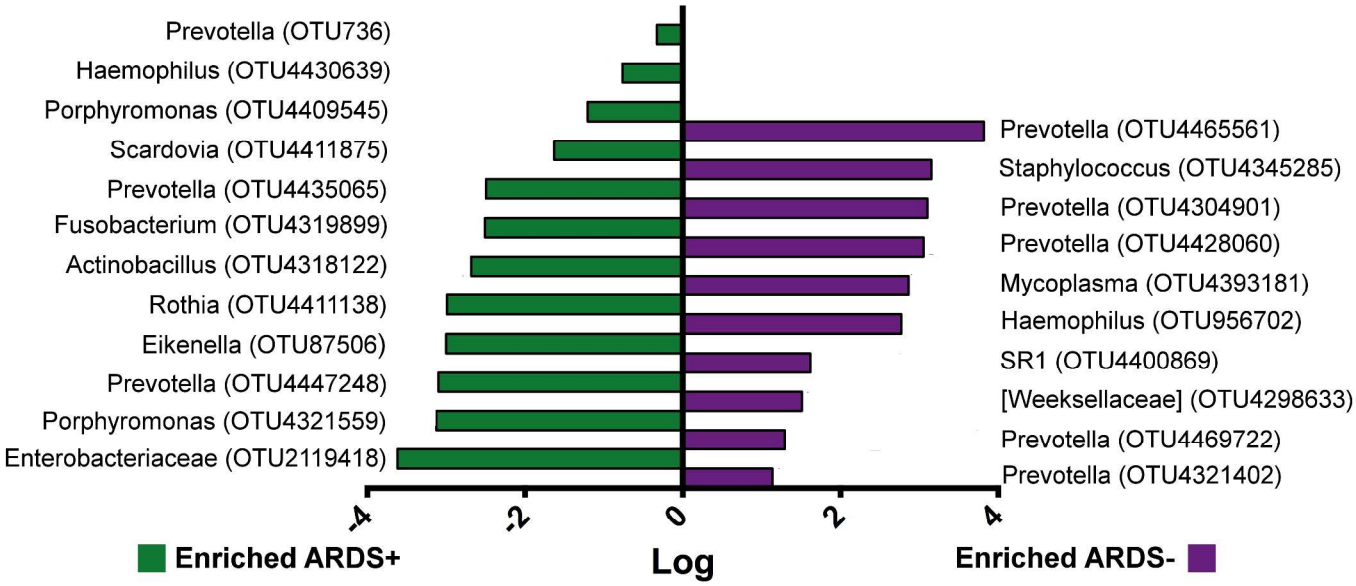


Figure 3A



AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

Fig 3B.



REFERENCES

1. NHLBI Clinical Trials Research Network for the Prevention and Treatment of Acute Lung Injury (PETAL Network). 2012 [cited 2014 March 20]. Available from: https://www.fbo.gov/index?s=opportunity&mode=form&tab=core&id=a0cadf303527630e423f3bd2af2d2f5e&_cview=0.

2. Calfee CS, Matthay MA, Eisner MD, Benowitz NL, Call M, Pittet JF, Cohen MJ. Active and passive cigarette smoking and acute lung injury after severe blunt trauma. *American Journal Of Respiratory And Critical Care Medicine* 2011;183:1660-1665.

3. Toy P, Gajic O, Bacchetti P, Looney MR, Gropper MA, Hubmayr R, Lowell CA, Norris PJ, Murphy EL, Weiskopf RB, Wilson G, Koenigsberg M, Lee D, Schuller R, Wu P, Grimes B, Gandhi MJ, Winters JL, Mair D, Hirschler N, Sanchez Rosen R, Matthay MA. Transfusion-related acute lung injury: incidence and risk factors. *Blood* 2012;119:1757-1767.

4. Diamond JM, Lee JC, Kawut SM, Shah RJ, Localio AR, Bellamy SL, Lederer DJ, Cantu E, Kohl BA, Lama VN, Bhorade SM, Crespo M, Demissie E, Sonett J, Wille K, Orens J, Shah AS, Weinacker A, Arcasoy S, Shah PD, Wilkes DS, Ware LB, Palmer SM, Christie JD. Clinical risk factors for primary graft dysfunction after lung transplantation. *Am J Respir Crit Care Med* 2013;187:527-534.

5. Calfee CS, Matthay MA, Kangelaris KN, Siew ED, Janz DR, Bernard GR, May AK, Jacob P, Havel C, Benowitz NL, Ware LB. Cigarette Smoke Exposure and the Acute Respiratory Distress Syndrome. *Critical Care Medicine* 2015;43:1790-1797.

6. Moazed F, Calfee CS. Environmental risk factors for acute respiratory distress syndrome. *Clin Chest Med* 2014;35:625-637.

7. Teo SM, Mok D, Pham K, Kusel M, Serralha M, Troy N, Holt BJ, Hales BJ, Walker ML, Hollams E, Bochkov YA, Grindle K, Johnston SL, Gern JE, Sly PD, Holt PG, Holt KE, Inouye M. The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell Host Microbe* 2015;17:704-715.

8. van Winkelhoff AJ, Bosch-Tijhof CJ, Winkel EG, van der Reijden WA. Smoking affects the subgingival microflora in periodontitis. *J Periodontol* 2001;72:666-671.

9. Shchipkova AY, Nagaraja HN, Kumar PS. Subgingival microbial profiles of smokers with periodontitis. *J Dent Res* 2010;89:1247-1253.

10. Charlson ES, Chen J, Custers-Allen R, Bittinger K, Li H, Sinha R, Hwang J, Bushman FD, Collman RG. Disordered microbial communities in the upper respiratory tract of cigarette smokers. *PLoS One* 2010;5:e15216.

11. Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, Young VB, Toews GB, Curtis JL, Sundaram B, Martinez FJ, Huffnagle GB. Analysis of the lung microbiome in the "healthy" smoker and in COPD. *PLoS One* 2011;6:e16384.

12. Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, Flores SC, Fontenot AP, Ghedin E, Huang L, Jablonski K, Kleerup E, Lynch SV, Sodergren E, Twigg H, Young VB, Bassis CM, Venkataraman A, Schmidt TM, Weinstock GM. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. *Am J Respir Crit Care Med* 2013;187:1067-1075.

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

13. Panzer AR, Lynch SV, Cohen MJ, Nelson MF, Cheung C, Matthay MA, Calfee CS. Lung microbiome at 48 hours of hospitalization is associated with the development of acute respiratory distress syndrome (ARDS) in severe blunt trauma patients. *Am J Respir Crit Care Med* 2015;191:A3957.
14. Brohi K, Cohen MJ, Ganter MT, Matthay MA, Mackersie RC, Pittet JF. Acute traumatic coagulopathy: initiated by hypoperfusion: modulated through the protein C pathway? *Ann Surg* 2007;245:812-818.
15. Baker SP, O'Neill B, Haddon W, Jr., Long WB. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma* 1974;14:187-196.
16. Bernard G. The Brussels Score. *Sepsis* 1997;1:43-44.
17. Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, Fan E, Camporota L, Slutsky AS. Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012;307:2526-2533.
18. Dempsey D, Tutka P, Jacob P, 3rd, Allen F, Schoedel K, Tyndale RF, Benowitz NL. Nicotine metabolite ratio as an index of cytochrome P450 2A6 metabolic activity. *Clin Pharmacol Ther* 2004;76:64-72.
19. Benowitz NL, Bernert JT, Caraballo RS, Holiday DB, Wang J. Optimal serum cotinine levels for distinguishing cigarette smokers and nonsmokers within different racial/ethnic groups in the United States between 1999 and 2004. *Am J Epidemiol* 2009;169:236-248.
20. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;7:335-336.
21. Lozupone C, Hamady M, Knight R. UniFrac--an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* 2006;7:371.
22. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 2005;71:8228-8235.
23. Vazquez-Baeza Y, Pirrung M, Gonzalez A, Knight R. EMPeror: a tool for visualizing high-throughput microbial community data. *Gigascience* 2013;2:16.
24. Oksanan J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. vegan: Community Ecology Package. R package version 2.0.10. R version 3.1.1 ed; 2014.
25. Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 2001;26:32-46.
26. Letunic I, Bork P. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* 2011;39:W475-478.
27. Moore FA, Moore EE, Poggetti R, McAnena OJ, Peterson VM, Abernathy CM, Parsons PE. Gut bacterial translocation via the portal vein: a clinical perspective with major torso trauma. *J Trauma* 1991;31:629-636; discussion 636-628.
28. Yao YM, Sheng ZY, Tian HM, Wang YP, Yu Y, Fu XB, Lu LR, Wang DW, Ma YY. Gut-derived endotoxemia and multiple system organ failure following gunshot wounds combined with hemorrhagic shock: an experimental study in the dog. *J Trauma* 1995;38:742-746.

29. Kale IT, Kuzu MA, Berkem H, Berkem R, Acar N. The presence of hemorrhagic shock increases the rate of bacterial translocation in blunt abdominal trauma. *J Trauma* 1998;44:171-174.
30. Dickson RP, Singer BH, Newstead MW, Falkowski NR, Erb-Downward JR, Standiford TJ, Huffnagle GB. Enrichment of the lung microbiome with gut bacteria in sepsis and the acute respiratory distress syndrome. *Nat Microbiol* 2016;1:16113.
31. Litmathe J, Zardo P, Dickgreber N, May G, Sucker C, Fischer S. Severe ARDS induced by fusobacterial infections: a rare clinical presentation of Lemierre syndrome. *Thorac Cardiovasc Surg* 2013;61:754-757.
32. Bartlett JG. Anaerobic bacterial infections of the lung and pleural space. *Clin Infect Dis* 1993;16 Suppl 4:S248-255.
33. Matsushita K, Tajima T, Tomita K, Abeyama K, Maruyama I, Takada H, Nagaoka S. Inflammatory cytokine production and specific antibody responses against possible causative bacteria in patients with multilesional periapical periodontitis. *J Endod* 1998;24:817-821.
34. Gemmell E, Seymour GJ. Interleukin 1, interleukin 6 and transforming growth factor-beta production by human gingival mononuclear cells following stimulation with Porphyromonas gingivalis and Fusobacterium nucleatum. *J Periodontal Res* 1993;28:122-129.
35. Shenoy MK, Iwai S, Lin DL, Worodria W, Ayakaka I, Byanyima P, Kaswabuli S, Fong S, Stone S, Chang E, Davis JL, Faruqi AA, Segal MR, Huang L, Lynch SV. Immune Response and Mortality Risk Relate to Distinct Lung Microbiomes in Patients with HIV and Pneumonia. *Am J Respir Crit Care Med* 2017;195:104-114.
36. Dickson RP, Erb-Downward JR, Huffnagle GB. Homeostasis and its Disruption in the Lung Microbiome. *Am J Physiol Lung Cell Mol Physiol* 2015:ajplung 00279 02015.
37. Bassis CM, Erb-Downward JR, Dickson RP, Freeman CM, Schmidt TM, Young VB, Beck JM, Curtis JL, Huffnagle GB. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. *MBio* 2015;6:e00037.
38. May AK, Brady JS, Romano-Keeler J, Drake WP, Norris PR, Jenkins JM, Isaacs RJ, Boczko EM. A pilot study of the noninvasive assessment of the lung microbiota as a potential tool for the early diagnosis of ventilator-associated pneumonia. *Chest* 2015;147:1494-1502.
39. Reinert DF, Allen JP. The Alcohol Use Disorders Identification Test (AUDIT): a review of recent research. *Alcohol Clin Exp Res* 2002;26:272-279.
40. Donovan DM, Dunn CW, Rivara FP, Jurkovich GJ, Ries RR, Gentilello LM. Comparison of trauma center patient self-reports and proxy reports on the Alcohol Use Identification Test (AUDIT). *J Trauma* 2004;56:873-882.

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

Lung Microbiota is Related to Smoking Status and to Development of ARDS in Critically Ill Trauma Patients: Online Supplementary Materials

Ariane R. Panzer, Susan V. Lynch, Chaz Langelier, Jason D. Christie, Kathryn McCauley, Mary Nelson, Christopher K. Cheung, Neal L. Benowitz, Mitchell J. Cohen, Carolyn S. Calfee

SUPPLEMENTARY METHODS

Plasma Biomarker Assays

Using plasma from 0 and 48 hours, we measured selected protein biomarkers of endothelial injury (angiopoietin-2, soluble intercellular adhesion molecule-1), epithelial injury (receptor for advanced glycation end-products [RAGE], vascular endothelial growth factor [VEGF]), and inflammation (interleukins 6 and 8, soluble tumor necrosis factor receptor-1, pentraxin 3). All biomarkers were measured in duplicate, using commercially available immunoassays (R&D Systems, Minneapolis, MN).

Endotracheal Aspirate (ETA) Nucleic Acid Extraction

Samples were thawed at room temperature, mixed until homogenous using a 10 mL pipette, and aliquoted. A 2.5 to 3 mL aliquot was centrifuged at 4,500 rcf for 25 min at 4°C to pellet the sample. Total DNA was extracted using the AllPrep DNA/RNA Mini Kit (Qiagen). 600 µl of Buffer RLT with 0.1% 2-Mercaptoethanol was added to the pelleted ETA samples or loaded directly into a Lysing Matrix E tube to serve as a negative control. One negative control was included for each batch of samples extracted, for a total of four negative controls. ETA sample pellets were resuspended and then transferred to Lysing Matrix E tubes. Samples and controls were bead-beaten using MPBio FastPrep-24 at 5.5 m/s for 30 s then centrifuged at 400 rcf for 1 min at room temperature. Samples

and controls were then transferred to DNA spin columns and manufacturer's instructions were followed for the remainder of the procedure. DNA was eluted in 100 µl EB Buffer and stored at -20°C.

16S rRNA-gene Analytic Methods

Sequence analysis of 16S rRNA-gene data was performed using the Quantitative Insights Into Microbial Ecology (QIIME) pipeline version 1.8.0 in the terminal environment. Raw sequences were de-multiplexed by barcode and base quality was assessed. If three or more consecutive bases had a Q score < 30, the sequence was truncated. Additionally, if the length of the sequence was less than 75% of the original 251 bp read length the whole sequence was discarded. Putative chimeric sequences were identified using USearch and removed from the sequencing data. Sequences were aligned using PyNAST (1) and operational taxonomic units (OTUs) were picked at 97% sequence identity using UCLUST against the Greengenes database (13_5).(2) Reads that failed to hit the reference sequence collection were retained and clustered de novo. A phylogenetic tree was built using FastTree.(3) Following this, negative controls yielded a read depth of 9,549 reads or less, although one negative control yielded 37,476 reads. The most abundant taxa across all negative controls were *Psuedomonadaceae*, *Commonadaceae*, and *Enterobacteriaceae*, with the negative control with 37,476 reads generally having an increased abundance of these taxa compared to the other negative controls. Importantly, the OTUs present in negative controls (**Table S5**) were not amongst the OTUs significantly enriched in smokers versus non-smokers or ARDS- versus ARDS+ patients at the 0 or 48 hour time points. To normalize read depth across samples, data were rarefied 100 times to a read depth of 59,959 reads, which was the minimum read depth amongst ETA samples. The distribution of sequencing depth for samples was as follows: 59,959 minimum read depth, 829,448 maximum read depth, 164,783 mean read depth, 105,395 standard deviation. The representative community composition for each sample was defined as that which exhibited the minimum average Euclidean distance to all other OTU vectors generated from all sub-samplings for

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

that particular sample. No samples were excluded from the study due to lack of amplification or inadequate read depth, but one patient was removed due to substantial missing clinical data. All sequencing data, including data on negative controls, has been uploaded to the European Nucleotide Archive with accession number PRJEB20913; the study unique name is: ena-STUDY-UCSF-17-05-2017-17:47:19:685-10.

Table S1: Analyses of Association Between Smoking, Alcohol Abuse, and Lung Microbiota Composition

Subset of Cohort With Chart Data for Alcohol Abuse (n = 54)

Individual Fractions	Adj. R. Squared	p-value of fraction
X1|X2	-0.00016	0.528
X2|X1	0.00155	0.157

Full cohort, missing data treated as "No" (n = 74)

Individual Fractions	Adj. R. Squared	p-value of fraction
X1|X2	0.00305	**0.044**
X2|X1	0.00179	0.061

Full cohort, missing data treated as "Yes" (n = 74)

Individual Fractions	Adj. R. Squared	p-value of fraction
X1|X2	0.00344	**0.028**
X2|X1	-0.00041	0.628

X1: Smoking status by cotinine
X2: Alcohol abuse by chart history

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

Table S2: Taxa enriched in smoker (white) and non-smoker (gray) endotrachael aspirates at 0 hours. Taxa highlighted in orange were also enriched at 0 hours in patients with a “yes” response to alcohol abuse by chart history; taxa highlighted in blue were also enriched in patients under 40 years of age (median age in the cohort). Taxa enriched in all three categories are highlighted in pink. A three model approach [Poisson (P), negative binomial (NegBin) and zero-inflated negative binomial (ZINegBin) model] were applied to the data. The best model fit was determined using a Bayseian information criterion value and is indicated.

ATTACHED AS SEPARATE FILE

Table S3: Patient demographics at 48 hours (n = 30)

	Non-Smokers (n=8)	Passive Smokers (n=6)	Smokers (n=16)	ANOVA p-value
Age, years	61 ± 20	43 ± 31	39 ± 15	0.06
Male gender, n (%)	4 (50)	3 (50)	12 (75)	0.39
Race, n (%)				0.91
Caucasian	5 (63)	4 (66)	10 (62)	
African American	0 (0)	0 (0)	4 (25)	
Asian	3 (37)	1 (17)	0 (0)	
N/A	0 (0)	1 (17)	2 (13)	
Hispanic ethnicity, n (%)	1 (13)	2 (33)	3 (19)	0.54
ARDS, n (%)	2 (25)	4 (66)	7 (44)	0.32
Injury Severity Score	30 ± 6	30 ± 11	31 ± 13	0.95
Medical History, n (%)				
Asthma	0 (0)	1 (17)	4 (25)	0.32
Cirrhosis	0 (0)	1 (17)	0 (0)	0.13
Congestive Heart Failure	0 (0)	0 (0)	0 (0)	--
COPD	0 (0)	0 (0)	0 (0)	--
Coronary Artery Disease	0 (0)	0 (0)	2 (13)	0.42
Diabetes	3 (37)	0 (0)	1 (6)	0.06
HIV	0 (0)	0 (0)	1 (6)	0.64
AIDS	0 (0)	0 (0)	0 (0)	--
Malignancy	0 (0)	1 (17)	1 (6)	0.49
Alcohol abuse Composite*, n (%)				0.03
Yes	0 (0)	1 (17)	9 (56)	
No	5 (63)	4 (66)	6 (38)	
N/A	3 (37)	1 (17)	1 (6)	
Antibiotic class administered				0.39
Beta Lactam Combo	1 (13)	0 (0)	2 (13)	
Carbapenem	0 (0)	0 (0)	1 (6)	
Cephalosporin	4 (50)	2 (33)	8 (50)	
Cephalosporin/ Aminoglycoside	1 (13)	1 (17)	3 (19)	
Cephalosporin/Lincosamide	1 (13)	0 (0)	0 (0)	
None	1 (13)	3 (50)	2 (13)	

* Defined as a chart history of alcohol abuse or an AUDIT score ≥ 8 for men, ≥ 5 for women (4, 5)

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

Table S4. Taxa enriched in ARDS+ (white) and ARDS- (gray) endotrachael aspirates at 48 hours post-hospitalization. A three model approach [Poisson (P), negative binomial (NegBin) and zero-inflated negative binomial (ZINegBin) model] were applied to the data. The best model fit was determined using a Bayseian information criterion value and is indicated.

ATTACHED AS SEPARATE FILE

Table S5: Raw Read Counts of Taxa Present in Negative Controls

ATTACHED AS SEPARATE FILE

AJRCCM Articles in Press. Published on 16-October-2017 as 10.1164/rccm.201702-0441OC

Fig. S1

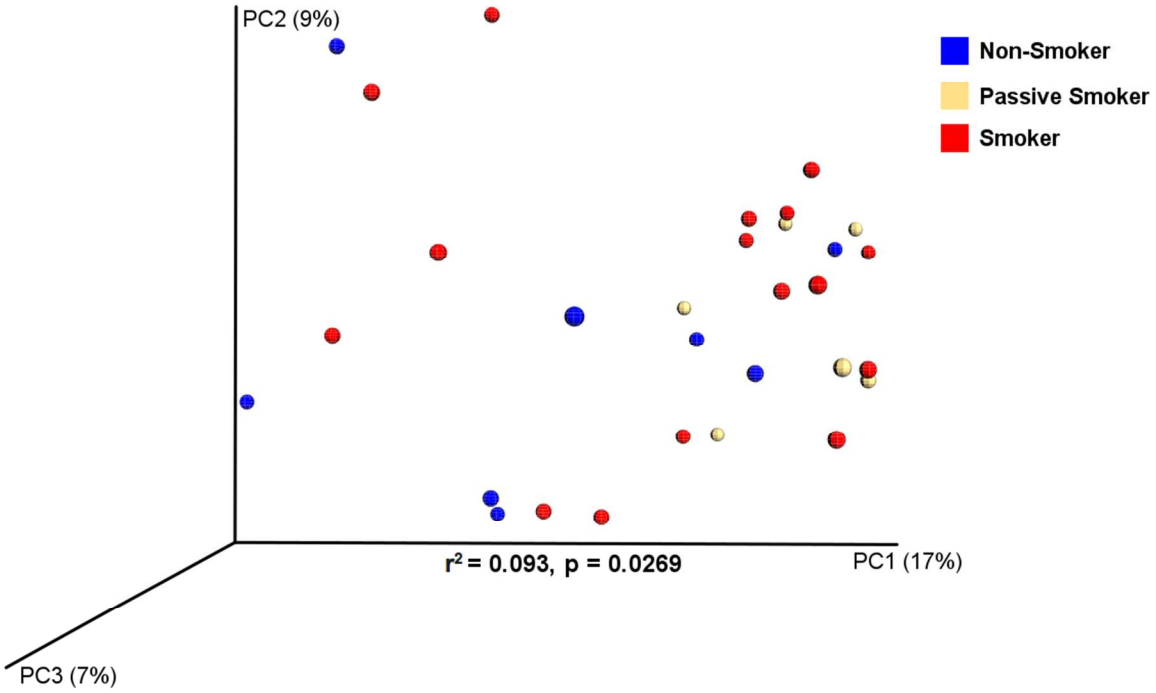


Figure S1. PCoA plot (unweighted Unifrac distance) of 48-hour samples colored by smoking status based on cotinine level. The r^2 value represents the percent variance explained and the p-value denotes the significance of this percentage.

REFERENCES

1. Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 2010;26:266-267.
2. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *Isme j* 2012;6:610-618.
3. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009;26:1641-1650.
4. Reinert DF, Allen JP. The Alcohol Use Disorders Identification Test (AUDIT): a review of recent research. *Alcohol Clin Exp Res* 2002;26:272-279.
5. Donovan DM, Dunn CW, Rivara FP, Jurkovich GJ, Ries RR, Gentilello LM. Comparison of trauma center patient self-reports and proxy reports on the Alcohol Use Identification Test (AUDIT). *J Trauma* 2004;56:873-882.