iScience, Volume 26

Supplemental information

The upper and lower respiratory tract

microbiome in severe aspiration pneumonia

Georgios D. Kitsios, Vi D. Nguyen, Khaled Sayed, Nameer Al-Yousif, Caitlin Schaefer, Faraaz A. Shah, William Bain, Haopu Yang, Adam Fitch, Kelvin Li, Xiaohong Wang, Shulin Qin, Heather Gentry, Yingze Zhang, Jack Varon, Antonio Arciniegas Rubio, Joshua A. Englert, Rebecca M. Baron, Janet S. Lee, Barbara Methé, Panayiotis V. Benos, Alison Morris, and Bryan J. McVerry

The upper and lower respiratory tract microbiome in severe aspiration pneumonia

Supplement

Table S1. Aspiration risk factors by clinical groups of aspiration diagnosis, related to Table 1 and STAR Methods.

Abbreviations: IL: interleukin; sTNFR1: soluble tumor necrosis factor receptor 1; ST2: suppressor of tumorigenicity-2; sRAGE: soluble receptor of advanced glycation end-products ; LBP: lipopolysaccharide binding protein (LBP); sCD14: soluble cluster of differentiation 14; SPD: surfactant protein D

Figure S1: Macro-aspiration (MAsP) and non-macro-aspiration pneumonia (NonMAsP) subjects had significantly worse 60-day survival (A) and liberation from mechanical ventilation (E) compared to uninfected intubated controls, related to Tables 1 and S1**.** Among MAsP subjects, those with high burden of chronic risk factors for aspiration had worse survival (B) and liberation (F) compared to those with high burden of acute risk factors for aspiration. Among NonMAsP subjects, there were no outcome differences by burden of risk factors for aspiration (C and G).

Figure S2: Daily Antibiotic spectrum intensity examined by the Narrow Antibiotic Treatment (NAT) score for MAsP (A), NonMAsP (B) and Intubated Control subjects (C), related to Figure 1**.** We calculated the daily NAT score from -5 days from sampling to post 10 days after sampling on day 1 and demonstrate median and interquartile range NAT scores for each day. D. MAsP and NonMAsP subjects had significantly higher NAT scores compared to Intubated Controls for the first 5 days post baseline sampling, but no significant differences in NAT scores between MAsP and NonMAsP subjects.

Figure S3: Histogram distributions of acute and chronic risk factors for aspiration, related to Tables 1 and S1.

Subjects with ≥2 risk factors for acute or chronic aspiration were assigned as those with high burden of acute and chronic risk factors for aspiration, respectively.

Figure S4: Ecological comparisons among all sample types, related to Figure 1 and STAR methods**.** Clinical samples from healthy volunteers for the upper and lower respiratory tract (oral washes [n=23] and bronchoalveolar lavage-BAL [n=24], respectively) and mechanically ventilated ICU subjects (oropharyngeal swabs [OPS, n=236] and endotracheal aspirates [ETA, n=216]) had significantly higher 16S rRNA gene sequencing yield (number of 16 reads, panel A), compared to experimental negative control samples (extraction negative and PCR negative controls), with the dashed line indicating the threshold of 300 reads used to exclude samples from further analyses (24 OPS and 29 ETA samples excluded). B-C: Alpha (Shannon index) and beta diversity (Manhattan distances) comparisons between clinical and experimental samples.

Figure S5: No difference in Bray Curtis dissimilarity indices between clinical groups for matched oropharyngeal (OPS) and endotracheal (ETA) samples, related to Figure 1.

Figure S6: Bacterial burden comparisons in oropharyngeal swab (OPS) samples for intubated patients with and

without specific aspiration risk factors, related to Figure 1.

Figure S7: Bacterial burden comparisons in endotracheal aspirate (ETA) samples for intubated patients with and

 $n, p = 0.87$ Wilcoxon, $p = 0.33$ Wilcoxon, $p = 0.064$ Wilcoxon, $p = 0.028$ $1e + 08$ þ. $1e + 08$ $\frac{1}{2}$ \bullet $1e + 08$ $1e + 08$ τš. Ŀ. $\frac{1}{2}$ × $\ddot{\bullet}$ \mathbb{R} $\overline{\mathbf{e}}$ 84 $\overline{\mathcal{F}_{\mathcal{G}}^{\mathcal{G}}}$:3 16S qPCR copies (log) - ETA
 $\frac{a}{\phi}$
G5 16S qPCR copies (log) - ETA
 $\frac{\ddot{\text{p}}}{\text{p}}$ ۵ò \mathbf{r} \bullet $\ddot{}$ 16S qPCR copies (log) - ETA
 $\frac{a}{\phi}$ E^K ϵ^+ $\ddot{}$ ÷. Ι. Þ 16S qPCR copies $\frac{\log 1}{10}$ - E $\ddot{\cdot}$ 441 \bullet ţ. Ñ se. ×ý $\ddot{\mathcal{L}}$ -0.5 ÷ $\overline{}$ ÷. $\mathbf{e}_{\mathbf{e}^{(1)}}$ $\ddot{\cdot}$ $\ddot{\cdot}$ $\ddot{}$ t. k. \cdot l, ī. 4 ti
A ł \bullet \bullet l. $1e + 02$ $1e + 02$ $1e + 02$ $1e + 02$ $\ddot{}$ l. $\ddot{}$ ار د $\ddot{}$ $\ddot{\bullet}$ $\ddot{}$ and \ddotsc \mathcal{A} $\ddot{}$ $\ddot{}$ \sim loss is \overline{N} No
Altered Mental Status on Intubation No
Seizure Prior to Intubation No. Yes Yes יס
Chronic Anti-Acid Use Chronic O₂ Use Willowson, $p = 0.32$
 $\therefore p = 0.32$ Wilcoxon, $p = 0.023$ $Wilqoxon$, $p = 0.042$ Wilcoxon, $p = 0.54$ $\mathbf{L}^{\frac{1}{2}}$ $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 &$ \mathbf{r} $\hat{\mathbf{r}}$ $1e + 08$ $1e + 08$ $1e + 08$ ŧ $\hat{\mathbf{v}}$ $1e + 08$ k. ٠. 44 ÷, y Ï, ų 16S qPCR copies (log) - ETA
 $\frac{a}{\phi}$

S \mathcal{D} : 16S qPCR copies (log) - ETA
.
.
. $\ddot{\bullet}$ $\ddot{\cdot}$ $\overline{\cdot}$ $-ETA$ ETA 薄 \bullet_{\bullet} 16S qPCR copies (log) - E $\frac{a}{c}$ Ř $\overline{8}$ 1e+05 $\ddot{\mathbf{z}}$ \ddotsc \bullet ¢ A l, $\ddot{}$ Ŀ. 16S qPCR copies ÷ $\ddot{}$ Ĵ, ٠, Ķ ķ, $\ddot{\cdot}$ ò. $\ddot{}$ \mathbf{t} $1e+02$ $1e + 02$ $1e + 02$ $1e + 02$. ÷, ú $\frac{1}{2}$ 66 No
Aspiration Reported by Radiologist \overline{No} No
ETOH intoxication or withdrawal No
Suspected Aspiration by Clinician Yes Drug Overdose

without specific aspiration risk factors, related to Figure 1.

Figure S8: Alpha diversity (Shannon index) comparisons in oropharyngeal swab (OPS) samples for intubated patients with and without specific aspiration risk factors, related to Figure 1.

Figure S9: Alpha diversity (Shannon index) comparisons in endotracheal aspirate (ETA) samples for intubated patients with and without specific aspiration risk factors, related to Figure 1.

Figure S10: Alpha diversity (Shannon index – A and B) and bacterial burden (C and D) comparisons in oropharyngeal swabs (OPS) endotracheal aspirate (ETA) samples for intubated patients, stratified to those with high burden for acute risk factors (AcuteRF, 26.5%), high burden for chronic risk factors (ChronicRF, 22.8%), and neither (50.8%), related to Figure 1.

Figure S11: Alpha diversity (Shannon index – A and B) and bacterial burden (C and D) comparisons in oropharyngeal swabs (OPS) endotracheal aspirate (ETA) samples for pneumonia patients, stratified by microbiologic culture results into Culture-positive (CxPos) vs. Culture-negative (CxNeg), related to Figure 1.

Figure S12: Density distributions for the relative abundances of the top 10 most common genera in oropharyngeal swab (OPS) samples (left panel) and endotracheal aspirate (ETA) samples (right panel), with density curves colored by clinical group diagnosis, related to Figures 1 and 2.

Figure S13: Volcano plots highlighting differentially abundant genera in oropharyngeal swab (OPS) samples (top panel) and endotracheal aspirate (ETA) samples (bottom panel) between macro-aspiration pneumonia subjects and healthy controls, related to Figures 1 and 2.

Figure S14. MAsP and NonMAsP subjects had lower abundance of oral commensals and higher abundance of plausible pathogens compared to intubated, uninfected controls, related to Figure 2**.** We classified all component bacteria by their membership in the typical commensal population of oral origin in the healthy lung microbiome (oral commensals) vs. plausible pathogens implicated in lower respiratory tract infections. Bacteria that did not fall clearly in either category were classified as other. A: Relative abundance barplot for endotracheal aspirate (ETA) samples, stratified in facets by clinical group categories: macro-aspiration pneumonia (MAsP), non-aspiration pneumonia (NonMAsP), intubated controls and healthy control subjects. B-D: Comparisons of relative abundance for the three main categories of bacteria (oral commensals, pathogens and other bacteria) by clinical groups.

Figure S15: Nanopore metagenomics results, related to Figure 2 and STAR methods**.** A. No difference in number of microbial reads for bacteria, fungi and viruses between the three clinical groups: macro-aspiration pneumonia (MAsP), non-aspiration pneumonia (NonMAsP) and intubated controls. B. N of bacterial reads for species with >1000 reads stratified by clinical group. C. No significant differences in relative abundance of *Streptococcus spp* between clinical groups. D. N of fungal reads for most common fungal species. *Candida spp* accounted for 95% of all fungal reads.

Figure S16: Laplace approximations showed optimal fit for two clusters in oropharyngeal swab (OPS) samples

(left) and three clusters in endotracheal aspirate (ETA) samples (right panel), related to Figure 3.

Figure S17: Density distributions for the relative abundances of the top 10 most common genera in oropharyngeal swab (OPS) samples (left panel) and endotracheal aspirate (ETA) samples (right panel), with density curves colored by DMM clusters, related to Figure 3.

Figure S18: Volcano plots highlighting differentially abundant genera in oropharyngeal swab (OPS) samples (top panel) and endotracheal aspirate (ETA) samples (bottom panel) between the low-diversity and high-diversity DMM clusters, related to Figure 3.

Figure S19: Comparisons of relative abundance for each category of bacteria by oxygen requirements, stratified by DMM clusters for oropharyngeal swab samples (OPS) and endotracheal aspirates (ETA), related to Figure 3. We classified all component bacteria in endotracheal aspirate specimens (ETA) by oxygen requirement into obligate anaerobes, aerobes, facultative anaerobes, microaerophiles, bacteria of variable oxygen requirement and unclassifiable. A-B: Relative abundance barplots for OPS (top) and ETA (bottom) samples, stratified in facets by DMM clusters. C-H: Comparisons of relative abundance for the three main categories of bacteria (obligate anaerobes, aerobes and facultative anaerobes) by DMM clusters in OPS (top) and ETA (bottom) samples. High diversity clusters had significantly higher abundance of obligate anaerobes, and lower abundance of facultative anaerobes in both OPS and ETA samples.

Figure S20: Comparisons of relative abundance for each category of bacteria by plausible pathogenicity requirements, stratified by DMM clusters for endotracheal aspirates (ETA), related to Figure 3. We classified all component bacteria by their membership in the typical commensal population of oral origin in the healthy lung microbiome (oral commensals) vs. plausible pathogens implicated in lower respiratory tract infections. Bacteria that did not fall clearly in either category were classified as other. A: Relative abundance barplot for endotracheal aspirate (ETA) samples, stratified in facets by DMM clusters. B-D: Comparisons of relative abundance for the three main categories of bacteria (oral commensals, pathogens and other bacteria) by DMM clusters.

Figure S21: Cross-correlation analyses for the top 20 taxa abundances in oropharyngeal swab (OPS) samples

(top panel) and endotracheal aspirate (ETA) samples (bottom panel) with 13 plasma biomarkers of host response, related to Figure 4. Significant univariate analyses are indicated with a "+" sign; however, no results remained significant after adjustment for multiple comparisons.

Figure S22: Correlograms depicting significant correlations between plasma biomarkers and OPS and ETA

bacterial taxa abundances classified by oxygen requirements, related to Figure 4. Significant correlations are shown

in color.

Figure S23: Taxa category comparisons between survivors and non-survivors, related to Figure 5.

Figure S24: Membership in the low diversity cluster in both the upper and lower respiratory tract (URT and LRT) was associated with worse 60-day survival, related to Figure 5. A: Waffle plot indicating the distribution of subjects by DMM clusters when considering both URT and LRT specimens. B. Kaplan-Meier curve for patients with low diversity cluster in both URT/LRT (combined dysbiosis group) vs. patients with high diversity clusters in both URT/LRT or patients with low diversity in one compartment.

Figure S25: Comparisons of bacterial burden (A), Shannon indices (B), and beta-diversity by Principal coordinates analyses (C) for OPS samples at different follow-up intervals, stratified by aspiration diagnosis, related to Figure 1. For bacterial burden and Shannon index, we constructed mixed regression models with random patient intercepts and adjusted for the number of days post-intubation that each sample was taken (as a proxy for the exposure to the hyperoxic environment of the ventilator) and the antibiotic exposure score (convex model) by the day of sampling. The bacterial burden over time was significantly associated with antibiotic exposure (p=0.04), whereas the Shannon index was associated with time of sample post intubation (p=0.003). In plots A and B, MAsP subjects are shown in black, NonMAsP subjects in orange and Intubated Controls in blue. Sampling intervals: baseline (within 48hrs of intubation), middle (days 3-6) and late (days 7-11 post intubation).

Figure S26: Comparisons of bacterial burden (A), Shannon indices (B), and beta-diversity by Principal coordinates analyses (C) for ETA samples at different follow-up intervals, stratified by aspiration diagnosis, related to Figure 1**.** For bacterial burden and Shannon index, we constructed mixed regression models with random patient intercepts and adjusted for the number of days post-intubation that each sample was taken (as a proxy for the exposure to the hyperoxic environment of the ventilator) and the antibiotic exposure score (convex model) by the day of sampling. The bacterial burden over time was not associated with either time post intubation or antibiotic exposure, whereas the Shannon index was associated with time of sample post intubation (p=0.002). In plots A and B, MAsP subjects are shown in black, NonMAsP subjects in orange and Intubated Controls in blue. Sampling intervals: baseline (within 48hrs of intubation), middle (days 3-6) and late (days 7-11 post intubation).

Figure S27: Change in composition of bacteria by oxygen requirement in OPS and ETA samples during follow-up,

related to Figure 2. In all plots, MAsP subjects are shown in black, NonMAsP subjects in orange and Intubated Controls in blue. Each plot is stratified by whether subjects were exposed to antibiotics with anaerobic coverage at the time of baseline sampling. Sampling intervals: baseline (within 48hrs of intubation), middle (days 3-6) and late (days 7-11 post intubation).

Table S3. Clinical characteristics and biomarker levels stratified by oral DMM clusters, related to Figures 3, 4 and 5.

Table S4: Clinical characteristics and biomarker levels stratified by the three DMM clusters in ETA samples, related to Figures 3, 4 and 5.

Figure S28. Enrollment and patient selection flowchart, related to STAR methods.

