Comparison of the Respiratory Microbiome in Healthy Nonsmokers and Smokers

Weiguang (Wayne) Mao

Background & Goal

Compare the microbiome of the upper and lower respiratory track in healthy HIV-uninfected nonsmokers and smokers. (Lung HIV microbiome project)

• Culture-independent methods are not reliable. -> methodological challenges

• Prior work has detected bacterial DNA in cigarettes. -> direct impact
Participants ➔ Sample Collection ➔ Processing

Exclusion criteria

- Many criterions...

- Nonsmokers: having smoked less than 100 cigarettes in their lifetime with none/illicit inhaled drugs/cigar/pipe in the past year

- Smokers: currently smoking at least 6 cigarettes per day for at least six months and might also be smoking illicit drugs, cigars and/or pipes.

<table>
<thead>
<tr>
<th>TABLE 1. CHARACTERISTICS OF PARTICIPANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
</tr>
<tr>
<td>Age, yr</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Ethnicity</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Not Hispanic</td>
</tr>
<tr>
<td>Race</td>
</tr>
<tr>
<td>White</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>

Data are presented as n (%) or mean ± SD.
Participants → Sample Collection → Processing

Oral washes (OW), 10ml sterile 0.9% saline

Bronchoalveolar lavage (BAL)

10ml to 50ml of sterile 0.9% saline washed through bronchoscope

• Gargle with 10ml to 50ml of sterile before topical anesthesia
• The bronchoscope is then inserted (a maximum of 300 ml 0.9% saline)

Sterile saline in a sample Collection cup
Participants ➔ Sample Collection ➔ Processing

6 centers for DNA extraction
1 center for DNA sequencing at Washington University

• DNA extraction validation (5 BAL specimens/center)
  
The amount of amplified material doesn’t correlate with the center but correlate with the BAL sample.

• **Negative control** (reagent –derived contamination)
  V1-V3: Several BAL samples have similar community structures compared with samples
  V3-V5: No significant overlap
Participants ➔ Sample Collection ➔ Processing

Non-metric multidimensional scaling
Participants ➔ Sample Collection ➔ Processing

• 16S rRNA gene sequences (highly conserved)

Two variable regions (separate) are amplified, which are V1-3 (base 27 to 534) and V3-5 (base 357 to base 926)

• OTU (Operational Taxonomic Units)
  Definition: A cluster of reads with 97% similarity.
  Cutoff: 0.03 distance

• Mothur package
Analysis

Confounding factors

- **Sex**
  Repeat analysis comparing OW and BAL in nonsmokers and smokers, with women excluded

- **The degree of smoking**
  Split the participants based on median pack-year smoking history

- **Body mass index (BMI)**
  Compare diversity measure between participants categorized by BMI

- **Systematic differences between centers**
Is mouth a source for the microbial community in the lung?

- Neural model

With probability $m$

A individual must leave or die at $\delta$

(uniformly distributed)

Initial abundance $N_i$

Saturated with a total of individuals

Reproduction probability $1 - m$

Source community/mouth

Lung

Result

1. Mouth as a source explains much of the microbial community in the lungs

V1-3

nonsmokers

smokers
Result

1. Mouth as a source explains much of the microbial community in the lungs

V3-5

nonsmokers

smokers
Result

2. Particular OTUs are differentially represented in BAL compared with OW communities

Fig E2

V1-3

V3-5
2. Particular OTUs are differentially represented in BAL compared with OW communities

**Result**

V1-3

V3-5
Result

3. OUT-Level Comparisons between Nonsmokers and smokers (OW)

V1-3

V3-5
3. OUT-Level Comparisons between Nonsmokers and smokers (BAL)
Result

4. α diversity (number of different bacterial sequences in a sample)

<table>
<thead>
<tr>
<th>Region</th>
<th>Smoker</th>
<th>Site</th>
<th>Samples</th>
<th>Observed Richness</th>
<th>Shannon Index</th>
<th>Inverse Simpson Index</th>
<th>Phylogenetic Diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1-3</td>
<td>No</td>
<td>BAL</td>
<td>37</td>
<td>58.4 (18.8)</td>
<td>2.90 (0.35)</td>
<td>10.9 (3.9)</td>
<td>2.26 (1.34)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>BAL</td>
<td>13</td>
<td>63.2 (28.0)</td>
<td>2.89 (0.49)</td>
<td>11.3 (5.5)</td>
<td>2.36 (1.52)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>OW</td>
<td>44</td>
<td>57.6 (13.9)</td>
<td>2.77 (0.34)</td>
<td>9.9 (3.4)</td>
<td>2.08 (1.41)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>OW</td>
<td>18</td>
<td>68.9 (24.0)</td>
<td>2.92 (0.44)</td>
<td>11.6 (4.4)</td>
<td>1.83 (1.67)</td>
</tr>
<tr>
<td>V3-5</td>
<td>No</td>
<td>BAL</td>
<td>16</td>
<td>54.0 (17.3)</td>
<td>2.60 (0.70)</td>
<td>8.8 (3.9)</td>
<td>2.75 (2.02)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>BAL</td>
<td>7</td>
<td>43.5 (18.1)</td>
<td>2.24 (1.01)</td>
<td>7.7 (5.6)</td>
<td>2.13 (1.63)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>OW</td>
<td>39</td>
<td>55.2 (12.8)</td>
<td>2.75 (0.30)</td>
<td>10.0 (2.9)</td>
<td>2.21 (1.44)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>OW</td>
<td>16</td>
<td>65.5 (18.4)</td>
<td>2.86 (0.46)</td>
<td>11.0 (4.7)</td>
<td>2.29 (1.39)</td>
</tr>
</tbody>
</table>

Definition of abbreviations: BAL = bronchoalveolar lavage; OW = oral wash; V1-3 = variable regions 1 through 3; V3-5 = variable regions 3 through 5. All metrics are based on the average of rarefying samples to 1,000 sequences. BAL samples excluded at V1-3 and V3-5 if community structure resembled that of controls. Samples were also excluded at V3-5 if there were insufficient sequences.

- Using V1-3 regions, there are no significant effects in comparisons of smoking status or OW to BAL on any of the α diversity measure
- Using V3-5 regions, a significantly higher number of OTUs measured by V3-5 in smokers’ BAL and OW than nonsmokers’ (p = 0.02)
Result

5. Structures

V1-3

Significant overlaps between OW and BAL.

V3-5

Fig E4
Result

5. Structures

Significant differences Among oral community, But not in the lung Community.
Smoking disrupt the normal structure community structure in mouth. For example, *porphyromonas*, a bacteria linked to periodontal disease, is depleted in OW of smokers.

Take care of the different 16S regions for amplification V1-V3 (more reads), V3-V5 (better detection)
Some considerations brought by authors

- Mouth is the only source community (nose, throat, gastrointestinal tract)
- Lack power to measure significant differences considering some factors (race, sex)
- Possibility of contamination
- Different methods used at difference centers
- Two-bronchoscope method?
- Neural model -> dead bacteria may be still clinically important