**Analysis question**: Is there difference in inflammatory urinary cytokines between MSM and MSW, and is that driven by   
 differences in penile microbiome composition?

**Initial Hypothesis:** Greater inflammation among MSM, driven by greater relative abundance of more anaerobic,   
 inflammation-associated bacteria.

**Step 1:** Characterize penile microbiome of MSM and MSW.

**Step 2**: Use rigorous method to identify taxa differences and importance of differences between MSM and MSW.

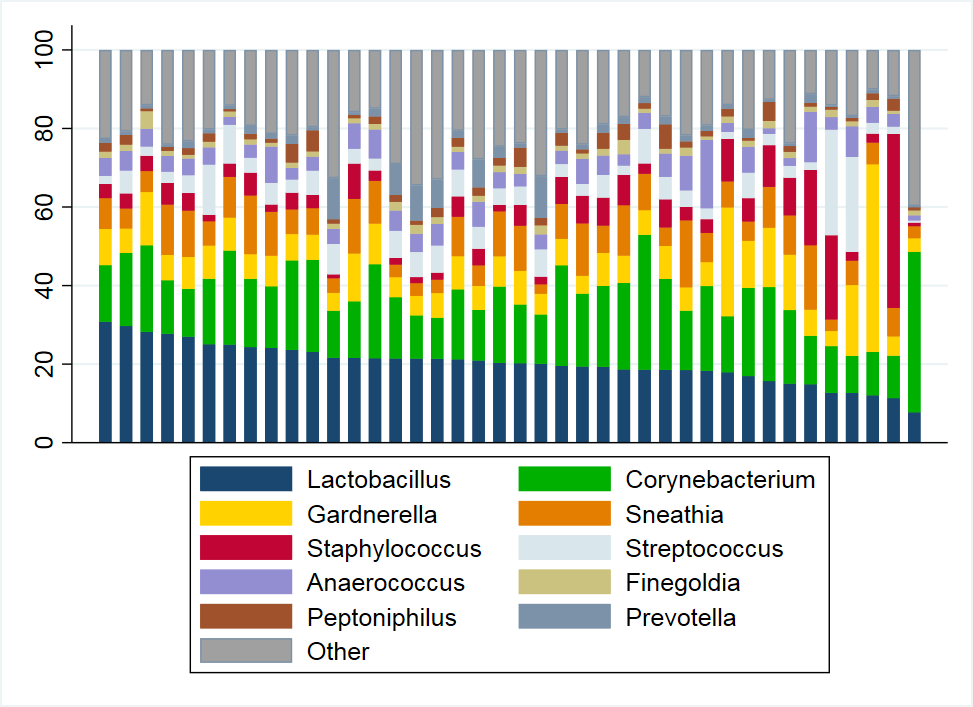
**Step 3**: Characterize urinary cytokines of MSM and MSW.

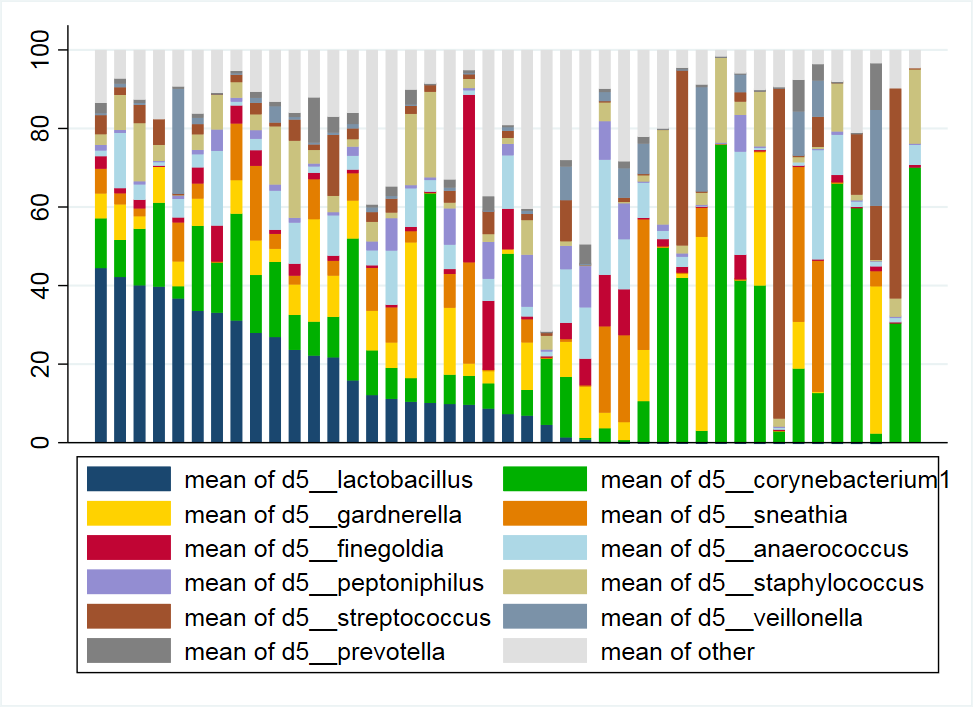
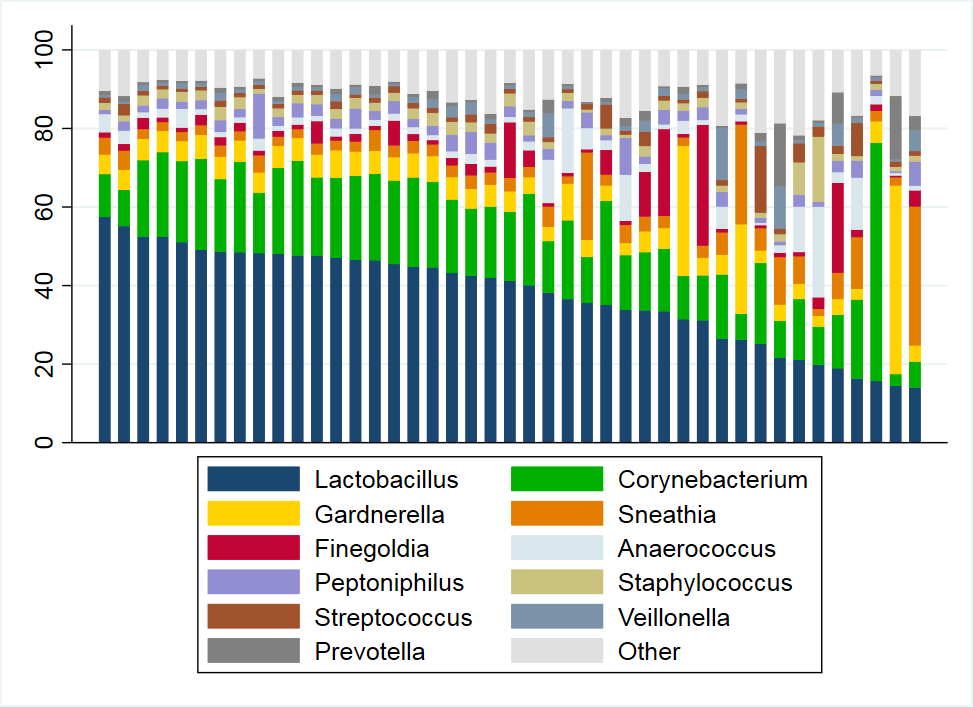
**Step 4**: Use multivariable regression to estimate the extent to which microbiome differences (identified in step 2)

account for differences in urinary cytokines between MSM and MSW.

**Step 1:** Characterize penile microbiome of MSM and MSW.

**Figure 1. Relative abundance of top 10 most abundant taxa by MSM or MSW status.**

****

****

**MSW Relative Abundance**

**MSM Relative Abundance**

**Table 1. Relative abundance of top 10 taxa by MSM or MSW status**

|  |  |  |
| --- | --- | --- |
| Taxa | **MSM**  **(%)** | **MSW**  **(%)** |
| *d5\_\_lactobacillus* | 37.32 | 11.47 |
| *d5\_\_corynebacterium1* | 17.62 | 22.33 |
| *d5\_\_gardnerella* | 7.82 | 8.95 |
| *d5\_\_sneathia* | 5.89 | 8.14 |
| *d5\_\_finegoldia* | 4.20 | 3.33 |
| *d5\_\_anaerococcus* | 4.00 | 7.17 |
| *d5\_\_peptoniphilus* | 2.93 | 2.51 |
| *d5\_\_staphylococcus* | 2.65 | 6.39 |
| *d5\_\_streptococcus* | 2.06 | 7.57 |
| *d5\_\_veillonella* | 1.99 | 4.75 |
| *OTHER* | 13.52 | 17.39 |

**Step 2**: Use rigorous method to identify taxa differences and importance of differences between MSM and MSW.

**Table 2**. **Random Forest Classifier Accuracy (msm=1, msw=0)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classifier | Accuracy  (CV=5) | TP | FP | FN | TN |
| Logistic Regression | 77% | 9 | 2 | 6 | 12 |
| Naive Bayes | 78% | 8 | 3 | 2 | 16 |
| Random Forest | 84% | 10 | 1 | 3 | 15 |

**Table 3. Taxa and importance from random forest classification of MSM vs. MSW status.**

|  |  |
| --- | --- |
| Taxa | IMPORTANCE |
| d5\_\_lactobacillus | 33.16 |
| d5\_\_corynebacterium1 | 4.48 |
| d5\_\_gardnerella | 3.22 |
| d5\_\_sneathia | 7.79 |
| d5\_\_finegoldia | 2.65 |
| d5\_\_anaerococcus | 0.73 |
| d5\_\_peptoniphilus | 5.98 |
| d5\_\_staphylococcus | 11.37 |
| d5\_\_streptococcus | 9.63 |
| d5\_\_veillonella | 16.56 |
| Other | 4.37 |

**Deb will :**

1. **Expand to use more taxa**
2. **Also compare with Elastic net**

**Step 3: Characterize urinary cytokines by MSM vs. MSW status**

**Figure 2. Summary of urinary cytokine concentrations by MSM or MSW status.**



**Table 4. Urinary cytokine concentration (pg/mL) compared by MSM vs. MSW status**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **MSM, N=43** | | **MSW, N=43\*** | | **Wilcoxon rank sum p-value** |
| **Mean** | **Median** | **Mean** | **Median** |
| **TNF-α** | 0.76 | 0.098 | 0.55 | 0.34 | 0.008 |
| **IL-1β** | 1.37 | 0.55 | 0.85 | 0.27 | 0.032 |
| **IL-8** | 18.0 | 11.1 | 19.3 | 8.34 | 0.397 |
| **IL-10** | 0.10 | 0.06 | 0.15 | 0.13 | <0.001 |
| **IP-10** | 24.3 | 16.5 | 44.9 | 33.6 | 0.001 |

**\*For MSW, TNFa has 3 missing values, and IP10 has 4 missing values**

Interpretation: Urinary measures of cytokine levels are generally **lower** for MSM than for MSW, and this is statistically significant for TNF-a, IL-10, and IP-10, but there is no difference in IL-8 between MSM and MSW, and IL-1b is statistically significantly lower for MSW.

**Table 5. Multivariable linear regression: Association of urinary cytokines with MSM vs. MSW status.**

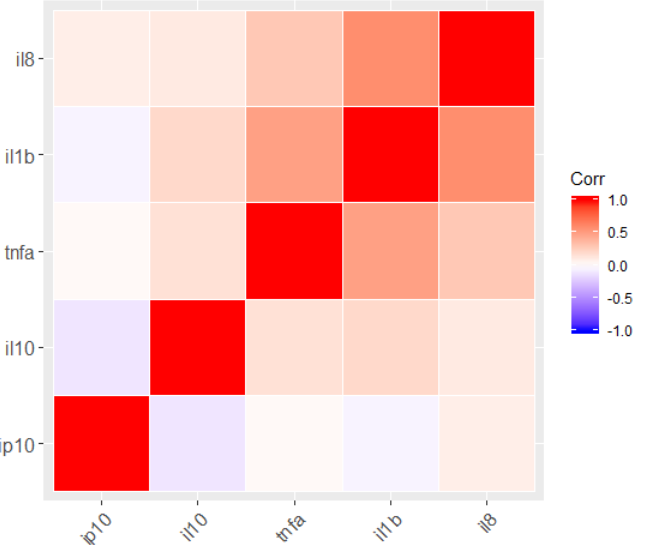
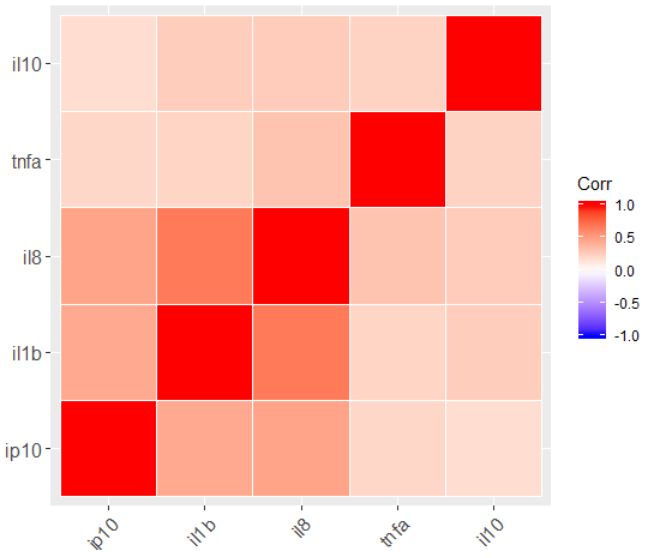
|  |  |  |  |
| --- | --- | --- | --- |
|  | **Estimate** | **P-value** | **Interpretation** |
| **TNF-a** | **-0.45** | **0.050** | Lower for MSM vs. MSW |
| **IL-8** | **0.04** | **0.860** | No difference MSM vs. MSW |
| **IL-1b** | **0.73** | **0.011** | Higher for MSM vs. MSW |
| **IL-10** | **-2.03** | **<0.001** | Lower for MSM vs. MSW |
| **IP-10** | **-0.43** | **0.077** | Lower for MSM vs. MSW |

Because cytokines are positively correlated with each other, we examine whether cytokines differ significantly by MSM vs. MSW status when simultaneously adjusted. So, these differences in cytokines by MSM vs. MSW status are independent of their inter-correlation. Note: cytokines are natural log transformed for linear regression analysis.

**Figure 3. Spearman Correlation plots**

The cytokines are more strongly positively correlated with each other among MSW than MSM. Among MSM, IP-10 is weakly negatively correlated with IL-1b and IL-10, with minimal correlation with TNFa and IL-8, while among MSW there are no negative correlations.

**MSM** **MSW**



**Step 4**: Use multivariable regression to estimate the extent to which microbiome differences (identified in step 2)

account for differences in urinary cytokines between MSM and MSW.

**Table 6. Association between taxa and urinary cytokines, crude and adjusted for MSM vs. MSW status.**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Taxa | **TNFa** | | **IL-1b** | | **IL-8** | | **IL-10** | | **IP-10** | |
| Taxa Only | MSM/MSW Adjusted | Taxa Only | MSM/MSW Adjusted | Taxa Only | MSM/MSW Adjusted | Taxa Only | MSM/MSW Adjusted | Taxa Only | MSM/MSW Adjusted |
|  | Estimate  (p-value) | Estimate  (p-value) | Estimate  (p-value) | Estimate  (p-value) | Estimate  (p-value) | Estimate  (p-value) | Estimate  (p-value) | Estimate  (p-value) | Estimate  (p-value) | Estimate  (p-value) |
| MSM vs. MSW |  | -0.96 (0.04) |  | 0.08 (0.85) |  | -0.18 (0.71) |  | -0.62 (<0.01) |  | -0.70 (0.11) |
| Lactobacillus | -0.10 (0.68) | 0.12 (0.63) | 0.40 (0.07) | 0.38 (0.12) | 0.51 (0.04) | 0.56 (0.04) | -0.25 (0.02) | -0.11 (0.34) | 0.13 (0.57) | 0.29 (0.24) |
| Corynebacterium | -0.44 (0.18) | -0.19 (0.57) | 0.56 (0.06) | 0.54 (0.10) | 0.96 (<0.01) | 1.00 (<0.01) | -0.20 (0.19) | -0.04 (0.79) | 0.26 (0.40) | 0.44 (0.18) |
| Gardnerella | 0.00 (0.99) | 0.13 (0.66) | 0.41 (0.13) | 0.40 (0.15) | 0.87 (<0.01) | 0.90 (<0.01) | -0.14 (0.29) | -0.06 (0.65) | 0.30 (0.27) | 0.40 (0.16) |
| Sneathia | -0.07 (0.79) | 0.07 (0.80) | 0.33 (0.16) | 0.32 (0.19) | 0.43 (0.10) | 0.46 (0.10) | -0.26 (0.03) | -0.17 (0.14) | 0.14 (0.58) | 0.24 (0.35) |
| Finegoldia | -0.12 (0.67) | -0.01 (0.97) | 0.09 (0.71) | 0.08 (0.75) | 0.51 (0.07) | 0.53 (0.07) | -0.21 (0.09) | -0.14 (0.25) | -0.11 (0.68) | -0.02 (0.92) |
| Anaerococcus | -0.80 (0.80) | -0.05 (0.87) | 0.31 (0.29) | 0.31 (0.30) | 0.69 (0.04) | 0.70 (0.04) | -0.11 (0.46) | -0.09 (0.53) | 0.46 (0.13) | 0.49 (0.11) |
| Peptoniphilus | 0.12 (0.72) | 0.39 (0.28) | 0.61 (0.05) | 0.59 (0.08) | 0.72 (0.04) | 0.77 (0.04) | -0.28 (0.07) | -0.11 (0.48) | 0.16 (0.60) | 0.36 (0.29) |
| Staphylococcus | 0.40 (0.26) | 0.46 (0.18) | 0.22 (0.49) | 0.21 (0.50) | 0.31 (0.37) | 0.33 (0.36) | -0.22 (0.17) | -0.18 (0.24) | 0.17 (0.60) | 0.22 (0.50) |
| Streptococcus | 0.09 (0.76) | 0.25 (0.40) | 0.44 (0.10) | 0.43 (0.13) | 0.42 (0.16) | 0.45 (0.15) | -0.19 (0.17) | -0.08 (0.52) | 0.35 (0.21) | 0.47 (0.10) |
| Veillonella | 0.07 (0.81) | 0.28 (0.36) | 0.38 (0.16) | 0.36 (0.20) | 0.60 (0.05) | 0.63 (0.05) | -0.19 (0.16) | -0.06 (0.66) | 0.16 (0.56) | 0.31 (0.28) |

TNFa: No taxa are independently associated with TNF-a at the p<0.05 level

IL1b: Increasing values of Peptoniphilus are associated with increasing value of IL-1b.

IL-8: Increasing relative abundance of most common taxa is associated with increasing value of IL-8.

IL-10: Increasing relative abundance of Lactobacillus, Sneathia are associated with increasing value of IL10.

IP-10: No taxa are independently associated with IP-10 at the p<0.05 level

**Repeat linear regression with new results union from EN and RF**

To do:

1. **AIC Select: for parsimonious models**
2. **Test of two way interaction with status**
3. **ABH for multiple testing**

Stepwise Selection (AIC)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | IL10 | Il-8 | tnfa | ip10 | il1b |
|  | *Status* | *d5\_veillonella* | *Status* | *Status* | *d5\_lacto* |
|  |  | *d5\_corynebacterium1* | *d5\_corynebacterium1* | *d5\_lacto* | *d5\_coryne* |
|  |  | *d5\_gardnarella* | *d5\_staphylococcus* | *d5\_coryne* | *d5\_gard* |
|  |  | *d5\_peptoniphilus* |  | *d5\_gardnare* | *d5\_sneat* |
|  |  |  |  | *d5\_anaero* | *d5\_pepto* |
|  |  |  |  | *d5\_strepto* | *d5\_strep* |
|  |  |  |  | *d5\_veillonella* | *d5\_veillonella* |