Methods:

If the abundance of only few microbes vary rather than the whole microbial profile, the global methods may lose power and multivariate methods that model multiple taxa jointly are more adequate. Dirichlet-multinomial (DM) distribution is commonly-used to model multiple taxon counts (La Rosa et al 2012). However, DM fails to accommodate complex correlation structure and excessive zero observations, hence, the fitting of DM model to the data is often poor and the discovery power of association tests based on DM-based regression is low (Tang and Chen 2018). In our analysis, we perform the state-of-the-art quasi-conditional association tests (QCAT) that recently developed by Tang et al. 2017. The QCAT has been modified to handle correlated repeated microbiome measures. The QCAT methods do not make any distributional assumptions on the data, hence, allow complex covariance structures among taxa and across time points. The two-part version of the method has higher discovery power in the presence of the rare taxa with low abundance.

A microbiome dataset usually involves hundreds and thousands of species represented by Operational Taxonomic Units (OTUs). A recent focus on tackling this issue of high-dimensional

compositional data involves decomposing the full composition into subcompsitions according to the structure of a taxonomic tree (Tang et al., 2017; Shi and Li, 2017; Wang and Zhao, 2017;

Tang et al., 2018). Following this strategy, we apply our tests to the subcompositions on a taxonomic tree to identify differential microbial lineages. Specifically, we visited every internal node on the tree; for each node, we applied the tests to the subcomposition defined as the vector of taxon counts on its immediate child nodes. The False Discovery Rate (FDR) can be controlled using the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001). Because taxonomically near taxa tend to have similar biological functions, this testing strategy leverages taxonomic structure to define the unit of the tests in a biologically meaningful way and effectively detects the local association signal. We adopt this testing strategy in our data analysis of the CFS study.

We have considered different formulations of linear predictors in the QCAT test (see the five models in the tables below). We also repeated the analysis by considering time points after exercise and adjusting the abundance before exercise (baseline abundance) as a covariate.

Filter: remove OUTs that present in less than 10 samples

Software: miLineage

Perform test on the effects of the disease-status-related terms in the model that are highlighted in red

======================= Stool Samples =======================

|  |  |
| --- | --- |
| Model | differential lineage |
| Model1: status | N/A |
| Model2: status + time | N/A |
| Model3: status + time + time\*status | N/A |
| Model4: status + time + time^2 | N/A |
| Model5: status + time + time\*status + time^2 + time^2 \* status | o\_\_Clostridiales 0.04  f\_\_Dehalobacteriaceae <0.001 |

Adjust for baseline:

|  |  |
| --- | --- |
| Model | differential lineage |
| Model1: status | f\_\_Actinomycetaceae 0.02 |
| Model2: status + time | N/A |
| Model3: status + time + time\*status | N/A |

======================= Blood Samples =======================

|  |  |
| --- | --- |
| Model | differential lineage |
| Model1: status | N/A |
| Model2: status + time | N/A |
| Model3: status + time + time\*status | f\_\_Moraxellaceae 0.005 |
| Model4: status + time + time^2 | N/A |
| Model5: status + time + time\*status + time^2 + time^2 \* status | N/A |

Adjust for baseline:

|  |  |
| --- | --- |
| Model | differential lineage |
| Model1: status | N/A |
| Model2: status + time | N/A |
| Model3: status + time + time\*status | N/A |
| Model4: status + time + time^2 | N/A |
| Model5: status + time + time\*status + time^2 + time^2 \* status | N/A |