Supplement Description:

otu\_table.csv OTU Table from SRA Data

Atacama\_mash\_dists.txt Mash Distances between Samples in SRA Data

Untrim\_Atacama\_Dists.txt Mash Distances without quality trimming

metadata.csv Metadata from SRA

clusters.K3.csv Generated clusters where K=3

clusters.K4.csv Generated clusters where K=4

k3\_jaccard\_sim\_between\_groupings.csv K3 Jaccard Similarities

k4\_jaccard\_sim\_between\_groupings.csv K4 Jaccard Similarities

final.tx.1.k3.5000nsNMF.nrun250.predict\_samples.log\_transform.nmf\_results.zip

* Sake run at K=3 with log transform

final.tx.1.k3.5000nsNMF.nrun250.predict\_samples.nmf\_results.zip

* Sake run at K=3 without log transform

final.tx.1.k4.5000nsNMF.nrun350.predict\_samples.nmf\_results.zip

* Sake run at K=4 with log transform