

## Contents of Supplementary Data Package

- Python Scripts
- SRA Tables - Study Metadata
- all.16S.sequence.lengths.by.library.tsv
- all.hfaA.metadata.jgi.csv
- all.hfaB.metadata.jgi.csv
- ANiB\_percentage\_identity.tab
- BLOCK01 - Importing 16S Libraries.R
- BLOCK02a - Figure 2a - Rank Abundance.R
- BLOCK02b - Figure 2b - Sample Totals.R
- BLOCK02c - Figure 3ab - Soil vs. Water Sources.R
- BLOCK02d - Caulobacteraceae at Large.R
- BLOCK03a - Figure 4ab - Holdfast Genes.R
- BLOCK03b - Figure SX - Map of hfaB.R
- BLOCK04a - Create Phyloseq Object from QIIME Output.R
- BLOCK04b - Figure 5a - Rank Abundance of OTUs.R
- BLOCK04c - Figure 5b - OTU Richness.R
- BLOCK04d - Figure 5c - Unifrac.R
- BLOCK05a - Comparative Genomics - ANI.R
- BLOCK05a - Comparative Genomics - MLSA Distances.R
- BLOCK05b - Comparative Genomics - Tanglegram.R
- BLOCK05c - Comparative Genomics - Mantel Test.R
- BLOCK05d - Comparative Genomics - Functional Genes.R
- BLOCK05e - Comparative Genomics - CAZymes.R
- BLOCK05f - Comparative Genomics - Lauro 2009.R
- BLOCK06a - LTSP.R
- BLOCK06b - HAKAI.R
- BLOCK06c - Henan Dry-Wet Experiment.R
- BLOCK06d - Lake Bacterioplankton Source Study.R
- BLOCK06e - fullCyc - SIP.R
- BLOCK06f - Inconclusive Case Studies.R
- classifying.hfaA.by.blast.tsv
- classifying.hfaB.by.blast.tsv
- COG.categories.csv
- genome.cazy.dist.rds
- genome.COG.BC.distance.mtarix.rds
- genome.info.2.mlsa.info.csv
- genome.info.csv
- genome.info.rds
- genomes.Caulobacter.JGI.functional.gene.count.csv
- genomes.cazy.filtered.out
- genomes.sample.data.csv
- genomes.vs.lauro.oligotroph.assay.rds
- hakai.final.Caulobacter.count\_table
- hakai.sample.totals.csv
- hfaA.Caulobacter.abundance.rds
- hfaB.Caulobacter.abundance.rds
- hfaB.genomes.faa
- img.environments.csv
- img.taxonomy.csv
- KEGG.2.COG.tsv
- lovK.genomes.faa.fasta
- lovK.genomes.mas
- main.normalized.counts.Domain.rds
- main.normalized.counts.Family.rds
- main.normalized.counts.Genus.rds
- metagenome.hfaAB.list.csv
- MLSA.Caulobacteraceae.genome.tree.nwk
- mlsa.distance.matrix.rds
- mlsa.genome.info.csv
- normalized.counts.caulobacter.final.rds
- otu\_table\_0.99.biom
- otu\_table\_0.99.pairwise.qiime+mothur.tsv
- OTUs.2.genomes.csv
- p.fullCyc.bact.normalized.rds
- p.QIIME.99.norm.rds
- p.QIIME.99.rds
- p\_Total.LTSP.Bacteria.13.8\_0.01.filtered.normalized.per1000.rds
- pangenome\_matrix\_t0.tab
- qiime.otus.caulobacter.tree
- sources.of.sizeable.population.rds