Tutorial 8 – Anvi'o

06 NOVEMBER 2020

MICB405

Workflow

Everyone:
- Install Anvi'o on your profile in the servers OR your laptop
- Run example dataset through simple Anvi'o metagenomic workflow on the servers OR your laptop
  - Requires Linux, Windows Subsystem for Linux, or macOS terminal
  - May work for non-WSL Windows users if Anaconda is installed in PowerShell
  - Won't work with PuTTY or WinSCP, as these are not full local terminal emulators

Optional:
- Visualize merged data on local Anvi'o
  - This can be done by SCP-ing merged data folder output from the servers to your laptop that has Anvi'o installed, or simply visualizing the output you produced on your laptop if you took that route

Breakout room team members should share screen so all can view the visualization in case any member doesn't have Linux/WSL/macOS, or is having technical issues

Your data (what anvi'o needs for metagenomics)

A FASTA file of your contigs

- Contigs.fa contains contigs from a co-assembly
  - This may be produced by running a larger set of high-quality samples through megahit
- May also have been a reference genome from NCBI, a metagenome-assembled genome (MAG), or a bunch of genes you are interested in profiling

BAM files for your samples.

- These are pre-aligned for simplicity, but ultimately come from FASTQ sequences unique to every *sample* you took from the environment
- e.g.
  - Day 1, Day 2, Day 3
  - Location 1, Location 2, Location 3