Anvi'o Interactive Interface and Refinement Tutorial

Intro

- Marine contaminated sediment at UCSB
  - vkivenson@gmail.com

- Ask me later if interested about:
  - **Using windows** to access a cluster
  - **NSF Supercomputer-** getting access, usage

  NOW: Exploring and refining data with Anvi'o visualization

Very useful tool for understanding workflow

- Recent Paper-

Metagenome sequencing and 98 microbial genomes from Juan de Fuca Ridge flank subsurface fluids

- We are now **doing the last step** → shows you the results of everything

Exploring your data using Anvio
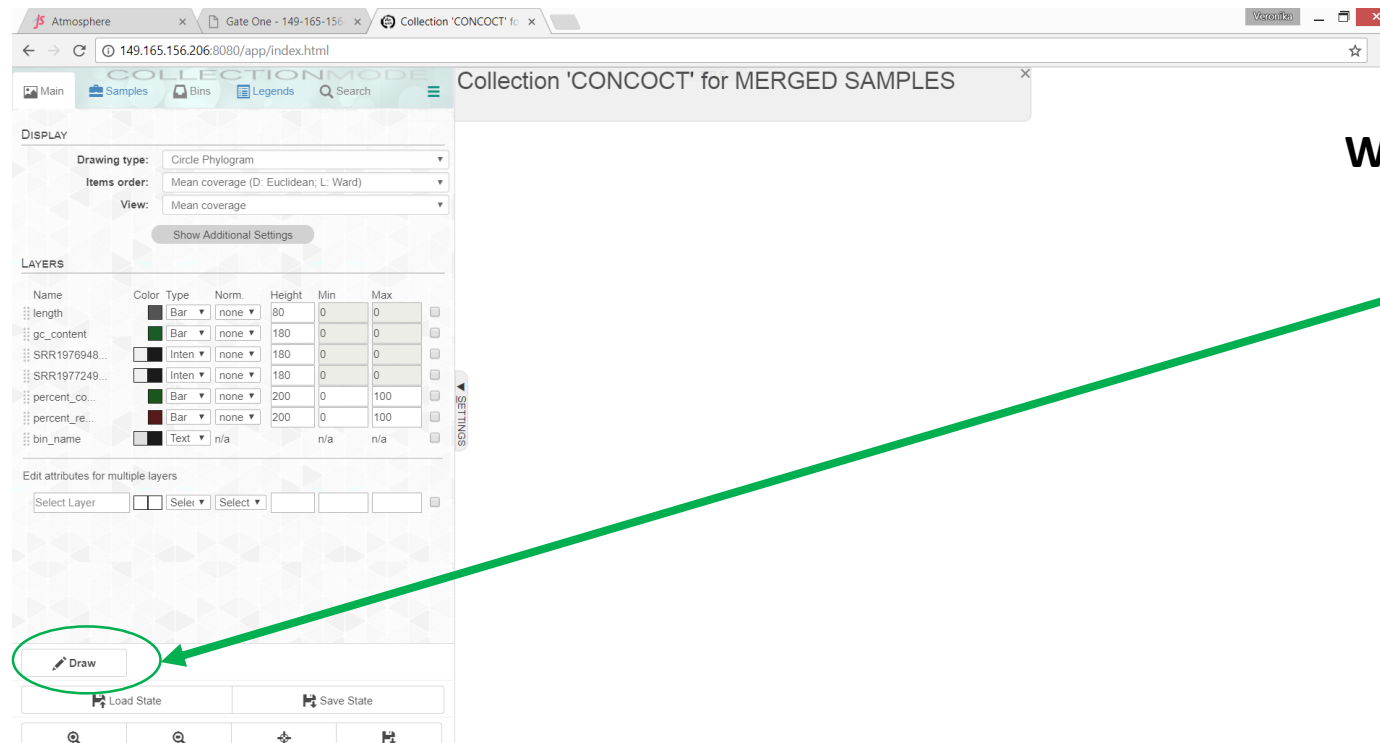
- The binning is based on—factors = **biological patterns** (e.g., tetranucleotide frequency/ GC content) and **differential coverage** (across samples) to generate genome bins

- How good is it?

- Can we look at what that means, what that looks like?

Review- Completion and Redundancy

- Single copy genes- metric we use to determine how complete a bin is
- (feedbacks)
    - _HOWEVER!!!_

- Completion and redundancy values are one of several aspects to look at

- **Even bins with great completion and redundancy estimates may need manual refinement**

Web-browser is used for interactive interface and refining- USE CHROME ONLY!

- [anvi-interactive -p MERGED-SAMPLES/PROFILE.db -c anvio-contigs.db -C CONCOCT]
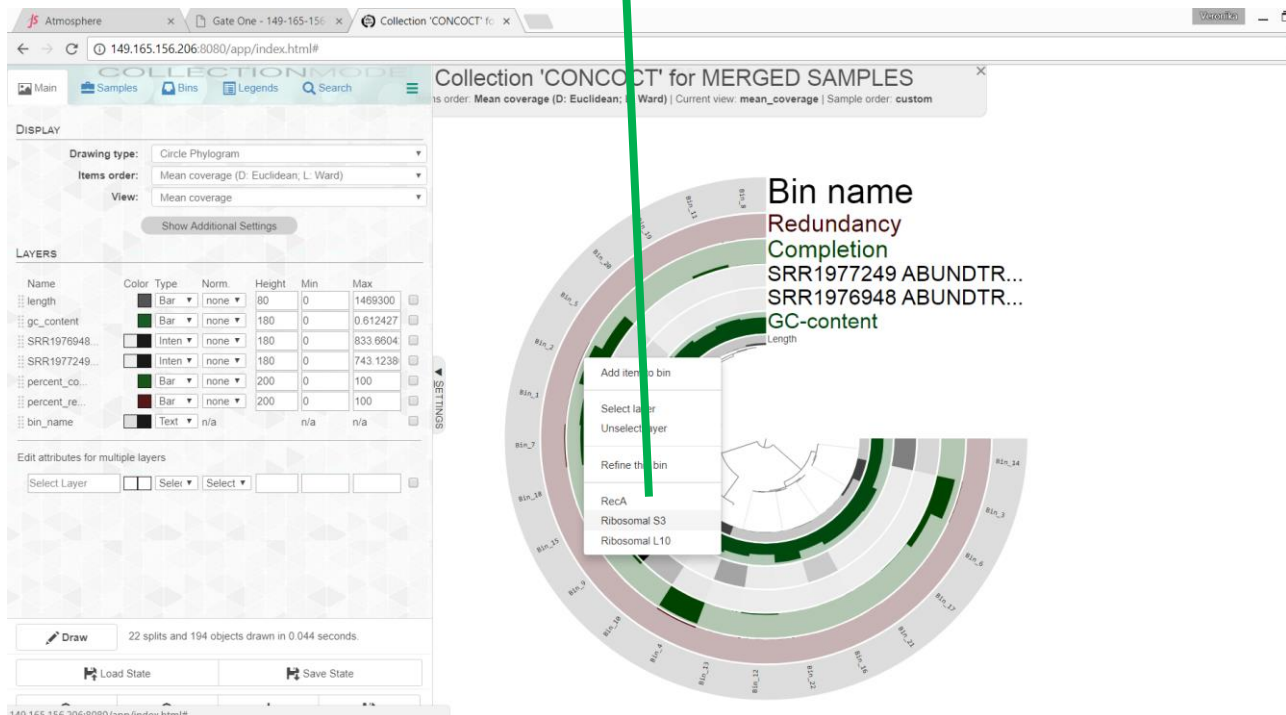
- Type in your IP address, here's mine: (http://149.165.156.206:8080/app/index.html

When you first open it, click "draw")

CONCOCT is the binning tool we used here, it made a collection

Exploring your data

- After drawing- you see all the bins; concoct is the binning program….

- ***THE DIFFERENT LAYERS- this is what we've been talking about; and bins***

- *Factors contributing* to bin: abundance, tetranucleotide frequency, GC content

- *Feedbacks*: redundancy, completion
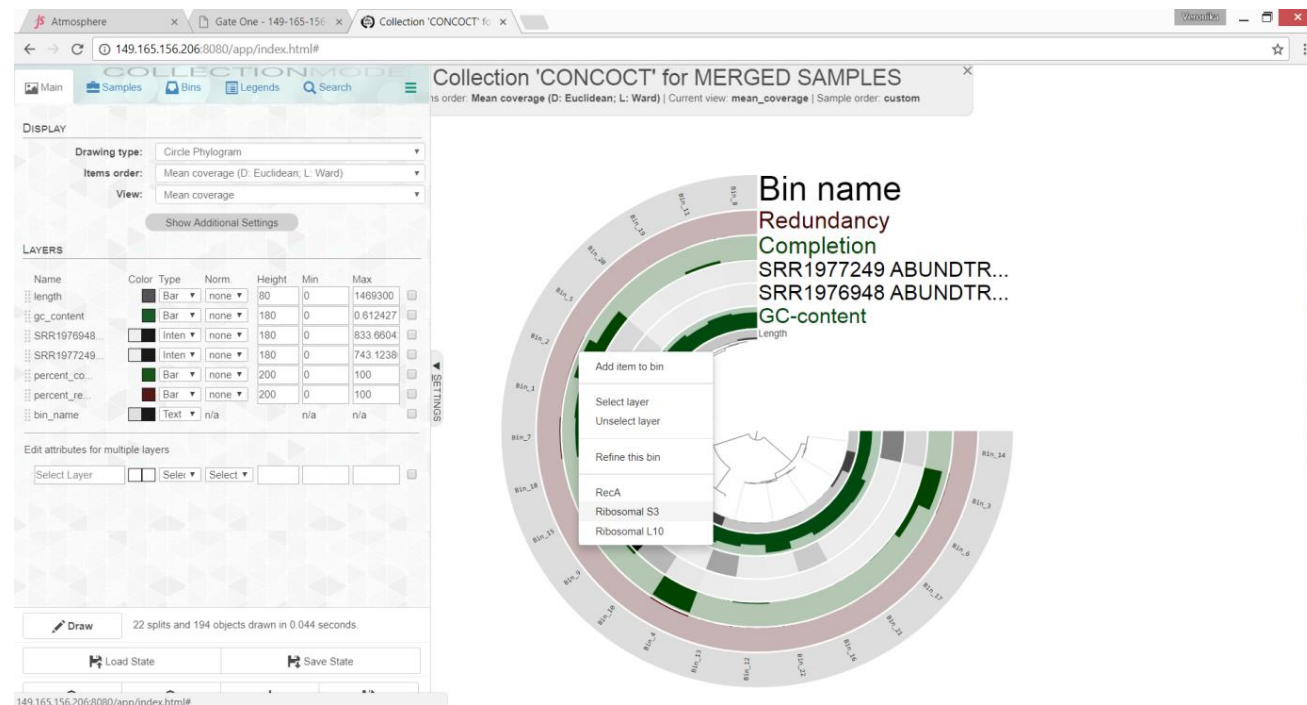
Exploring your data

- After drawing- you see all the bins- **can change color**
- Press s for settings
- Right click on bin 2
- Select "Ribosomal S3" (marker gene like 16s)
- See the sequence for it, try on several bins

Refining a Bin

- Let's pick a bin and take a closer look at it

- Press Ctrl+C to close the interactive and type in anvi-refine:

anvi-refine -p MERGED-SAMPLES/PROFILE.db -c anvio-contigs.db -b Bin_4 -C CONCOCT
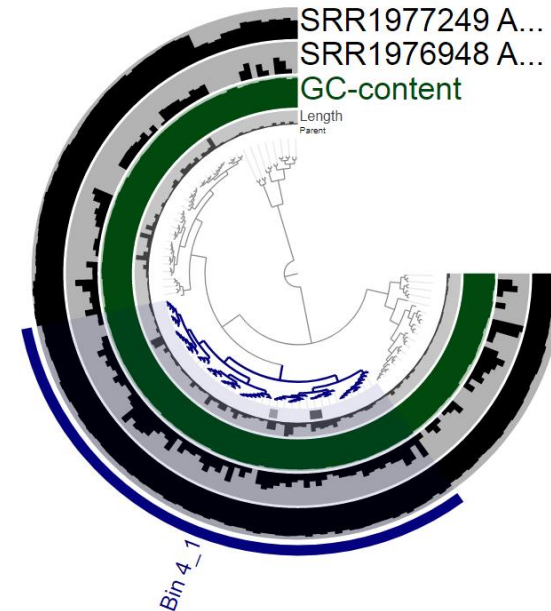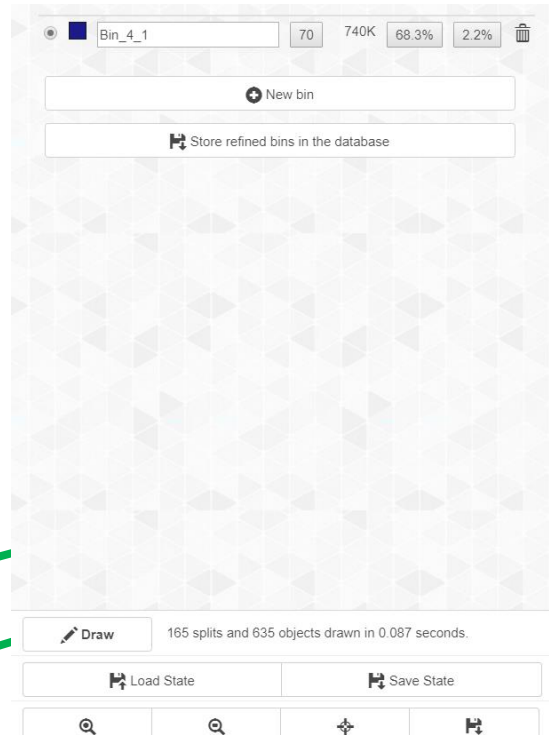
Exploring your data

- **Press s** for settings
- **Press m** and see mouse options appear

- Move your cursor over Bin_4
- From the innermost ring:

Length,

GC Content,

coverage

completion %, redundancy %

PRESS M TO SEE NUMBERS!
Draw coverage on the board!!! 40X EXAMPLE

- Try exploring with these options
- Click to toggle and then click draw again to see how things change

SAVE BIN AND WAIT FOR CONFIRMATION

- Press Ctrl+C to close it after saving the bin

- Then press UP button and type in the underscore after Bin_4_1

- Now you see your refined bin and you can continue refining or just keep it

- anvi-refine -p MERGED-SAMPLES/PROFILE.db -c anvio-contigs.db -b Bin_4_1 -C CONCOCT