

# 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

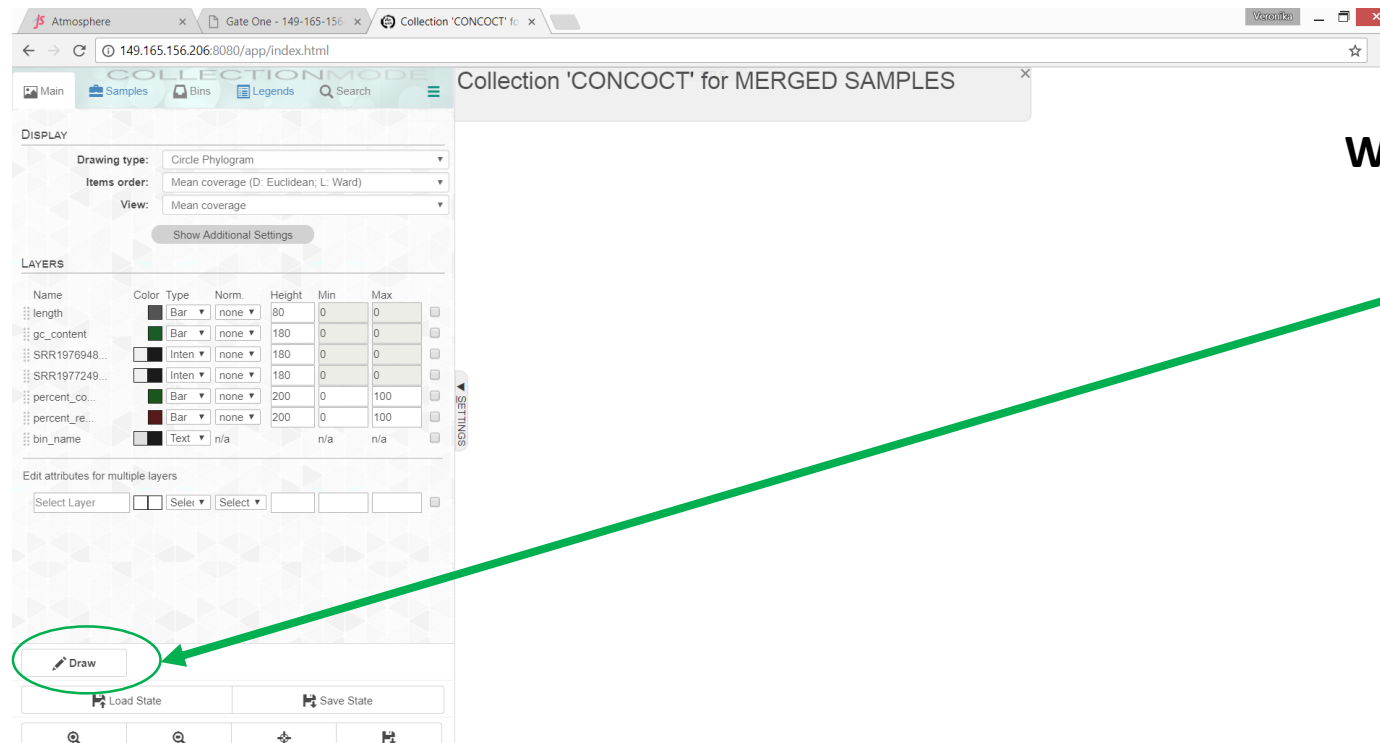
- The binning is based on—**factors** = **biological patterns** (e.g., tetranucleotide frequency/ GC content) and **differential coverage** (across samples) to generate genome bins
- How good is it?
- Can we look at what that means, what that looks like?

# 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**

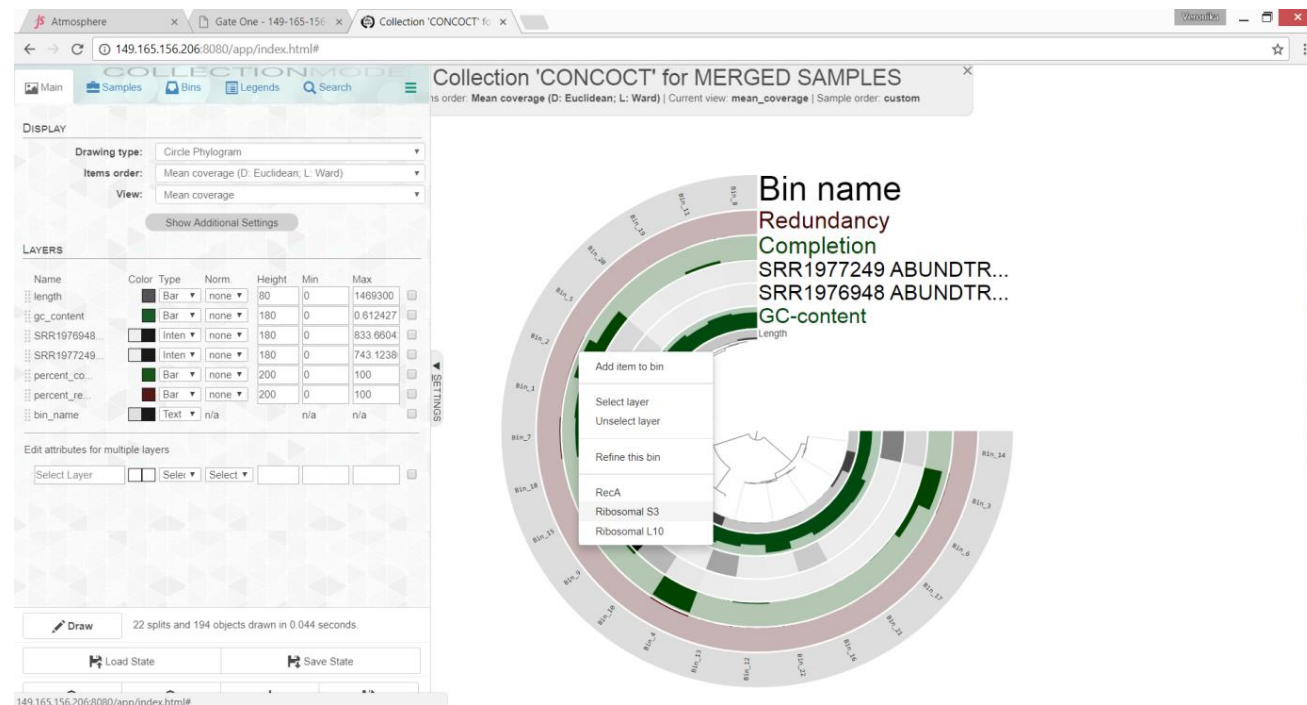
# Web-browser is used for interactive interface and refining- USE CHROME ONLY!

- [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>)



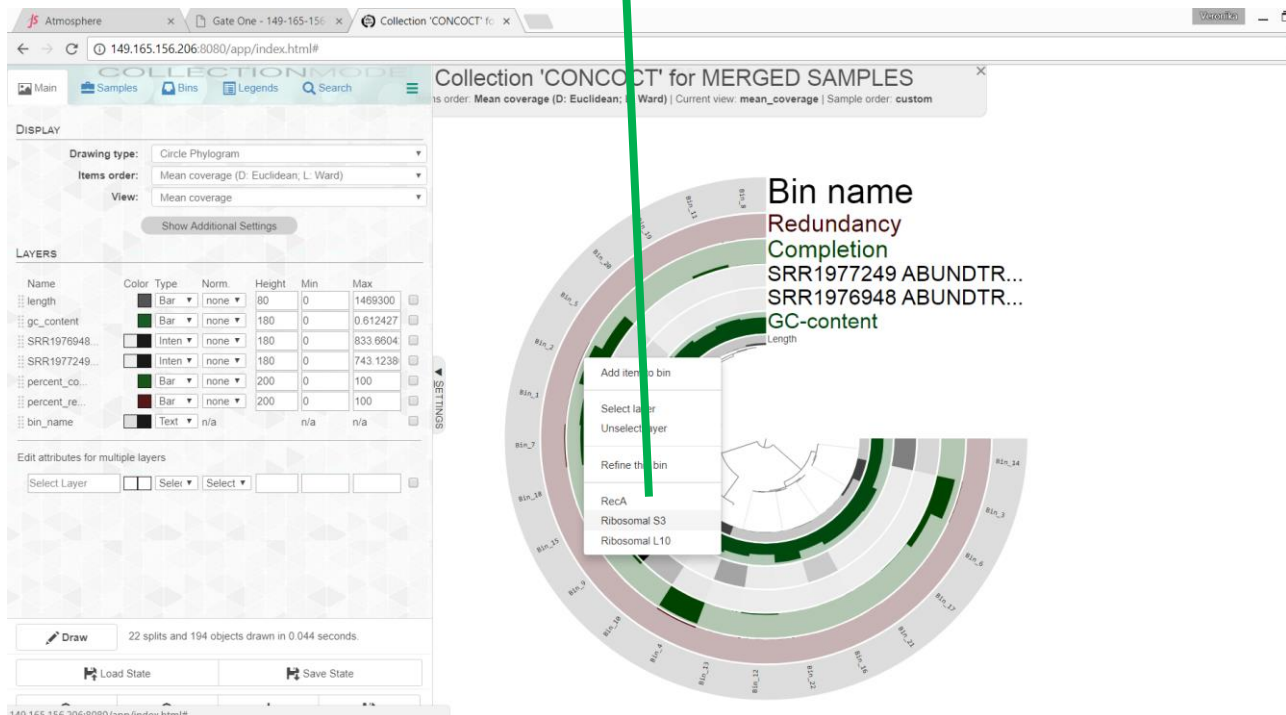
# Exploring your data

- 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



The 'Split Sequence' dialog box displays a DNA sequence for a specific bin. The sequence is as follows:

```
>Ribosomal_S3_C__Campbell_et_al__6969607bc3a97db5877cb07ed740224668
36b1f9e037896f3e1949a3|bin_id:Bin_2|source:Campbell_et_al|e_value:6.5e-
37|contig:c_000000005801|gene_callers_id:7368|start:6678|stop:7350|length:672
CTACCCCTCTCTTTTCTGCTCTGCATTACTGGAGGCCTGGGTTTCCGTTTGGGC
GATACCACGGTTTATCCAGACCTTTACTCCGATTACGCCGTACATGGTTTTCGCTT
CGGAGTAGCCGTAGTCAATATCGTTTCTGAGCGTCGAAAGCGGCAGTCGTCTT
CGAGATACTCGTCCCTGGCGATTTCGCTCCGCCGAGCCTGCCGCTGCAC
TGGATTTTATGCCTTCCGCTCCTGCTTTCATGGCCCGGAATATGGCCTGTTTCA
TCGCCGTCGAAACTCACCTGCGCTCAAGAGCCGAAGCAACGCCTTCGGAT
ACTATCTGGGCATCCTTATCCGGCGCCTTTATCTCCTGAACGTTGATCATGATCTT
GTGTCCGGTCATATTCTGAAGTTCTTCACGGACCGCCTGTATCTCCGTTCCGCC
CTTGCCTATGACTACACCGGGCCGGGCGGACCATACGGTAAACCTGATACTTT
CCCTATGCGCTCTATCTCGACCTTTGAAATCCCCGCGCGTGACCACCTTTTCATG
ATCCACTCGCGAAGCTTCAGGTCGTCATGAAGTTTTTCCGCGTAATTCTTCCGT
CGGCAAACCATTTTCGATTCCTGACGATTCCCAAGCGAAACCGACGG
GATGCACTTTTGGACCAA
```

A green circle highlights the 'Split Sequence' title, and another green circle highlights the 'bin\_id:Bin\_2' field in the sequence header. A green arrow points from the text 'See the sequence for it, try on several bins' to this dialog box.



# 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

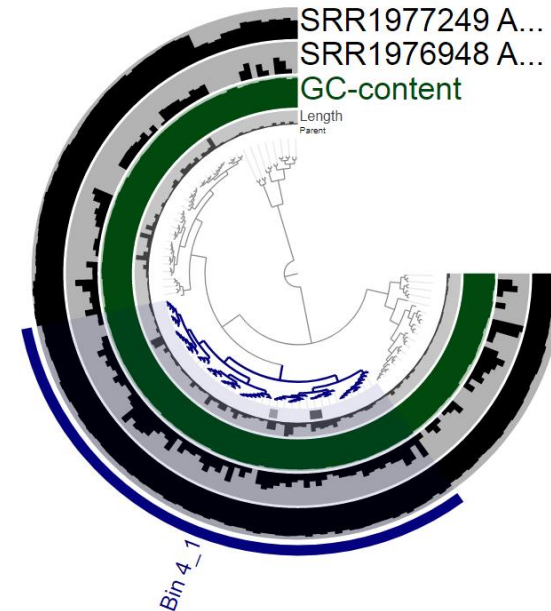
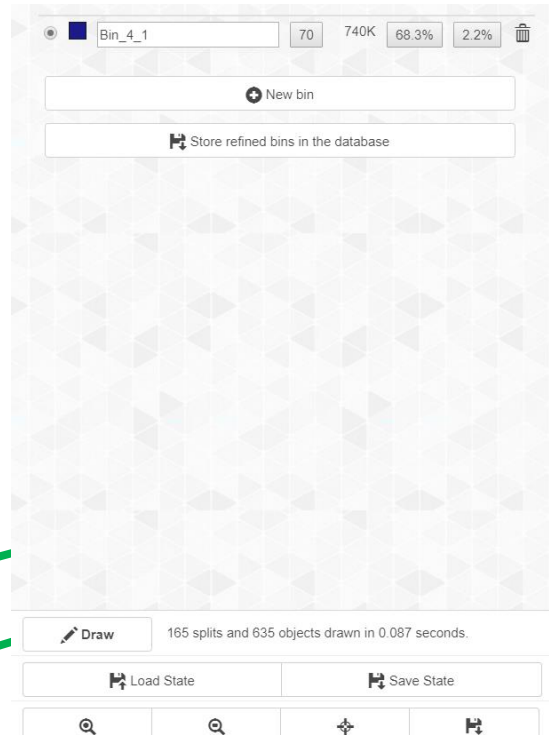
- **Press s** for settings
- **Press m** and see mouse options appear

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

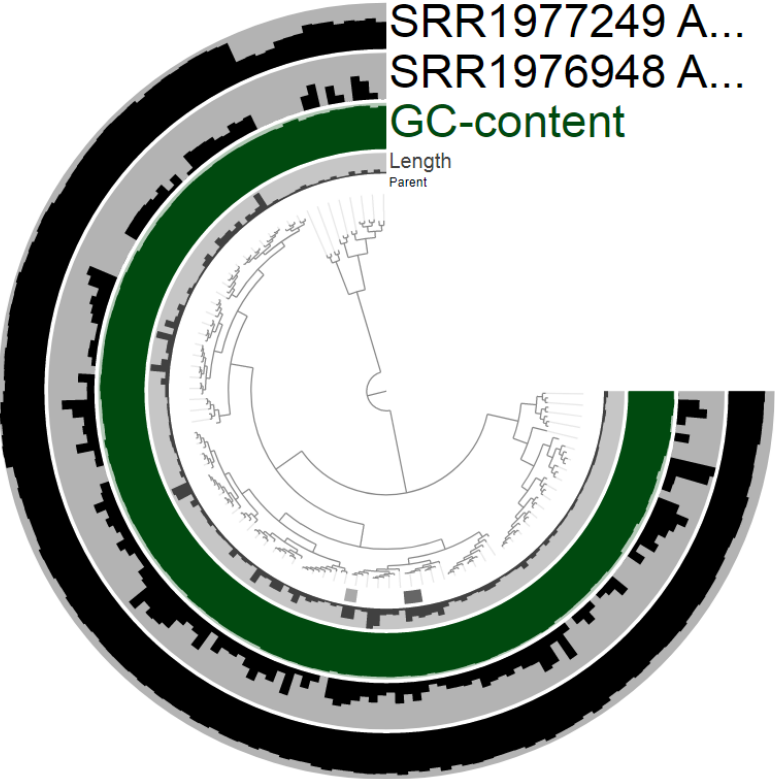
- Move your cursor over Bin\_4
- From the innermost ring:

Length,  
GC Content,  
**coverage**

completion %,  
redundancy %



split_name	c_000000003155_split_00001
parent	n/a
num_genes	34
avg_gene_length	946.0588235294117
ratio_coding	0.9839706332211685
length	32690
gc_content	0.45500152951973083
SRR1976948_ABUN DTRIM_SUBSET_PE _ANVIO	0.12285102477821964
SRR1977249_ABUN DTRIM_SUBSET_PE _ANVIO	21.207647598654024
bin	<input type="checkbox"/> n/a



- # 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`