

# Genome 540 Discussion

Jan. 13<sup>th</sup> 2026



# Office hours Logistics

- 3-4 pm Tuesdays
- 5-6 pm Wednesdays

Upstairs in S-334 (or similar)

Ping me for access into the building

# Office hours Logistics

! This week only !

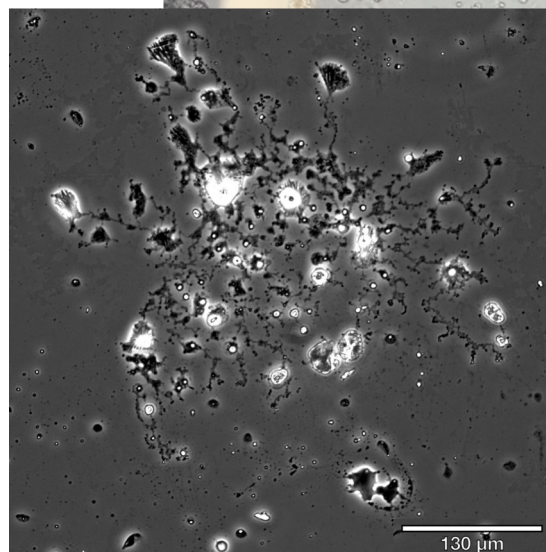
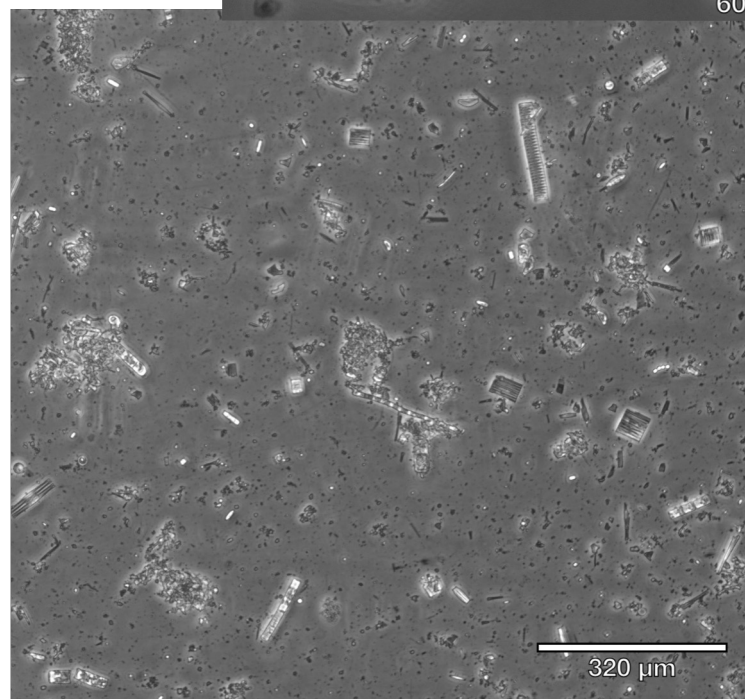
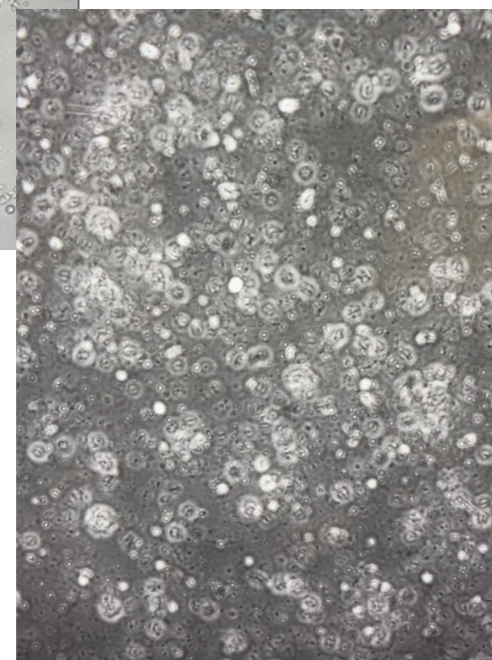
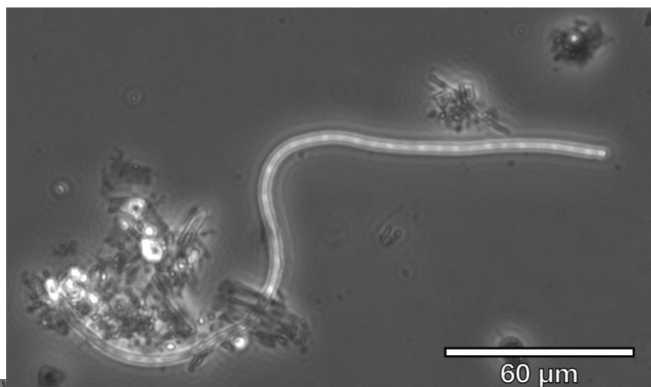
(Later today: 2-3 pm instead of 3-4 pm)

Times are still flexible: if we need to adjust hours that's ok.

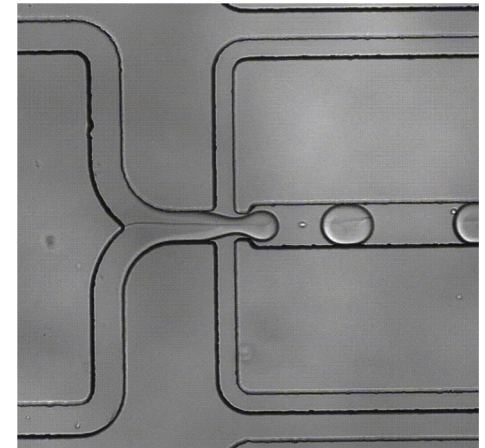
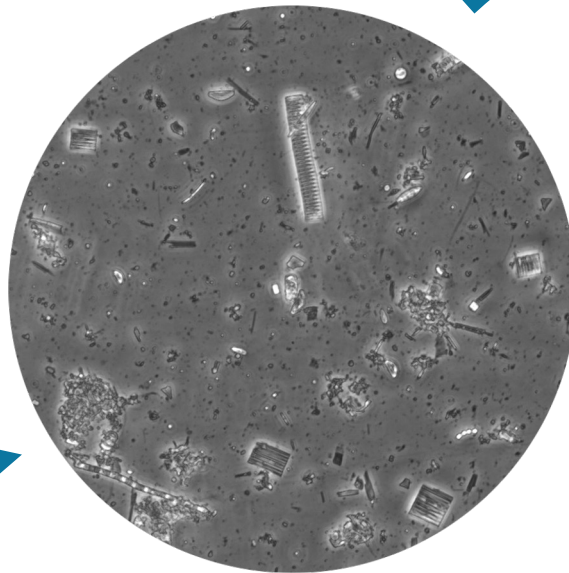
Still happy to meet by appointment, and always available over slack 😊

## Genome background models: Failure of equal frequency assumption

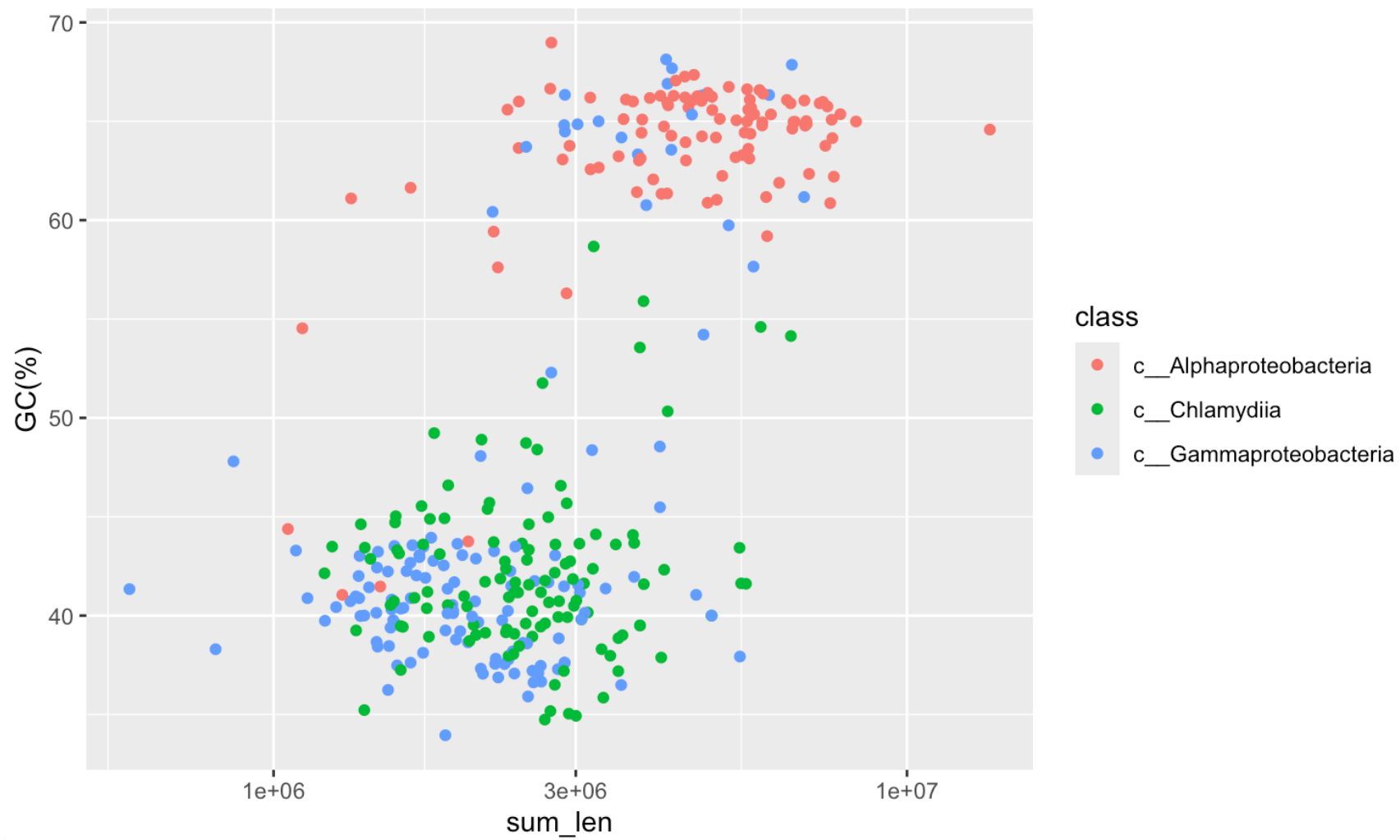
- For most organisms, the genomic nucleotide composition is significantly different from .25 for each nucleotide, e.g.:
  - *H. influenza* .31 A, .19 C, .19 G, .31 T
  - *P. aeruginosa* .17 A, .33 C, .33 G, .17 T
  - *M. janaschii* .34 A, .16 C, .16 G, .34 T
  - *S. cerevisiae* .31 A, .19 C, .19 G, .31 T
  - *C. elegans* .32 A, .18 C, .18 G, .32 T
  - *H. sapiens* .29 A, .21 C, .21 G, .29 T



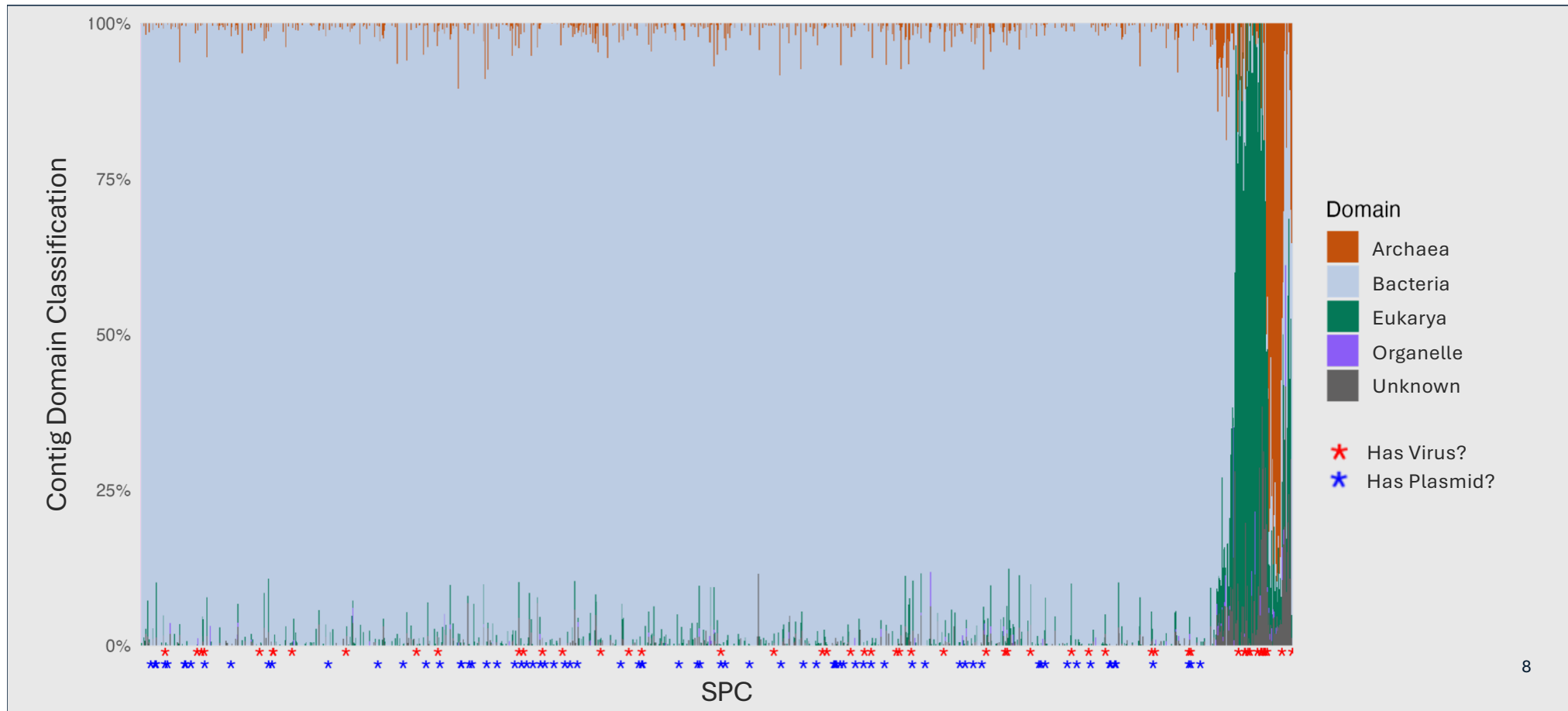
# Generating Individual Bacterial Genomes with Semi-permeable capsules



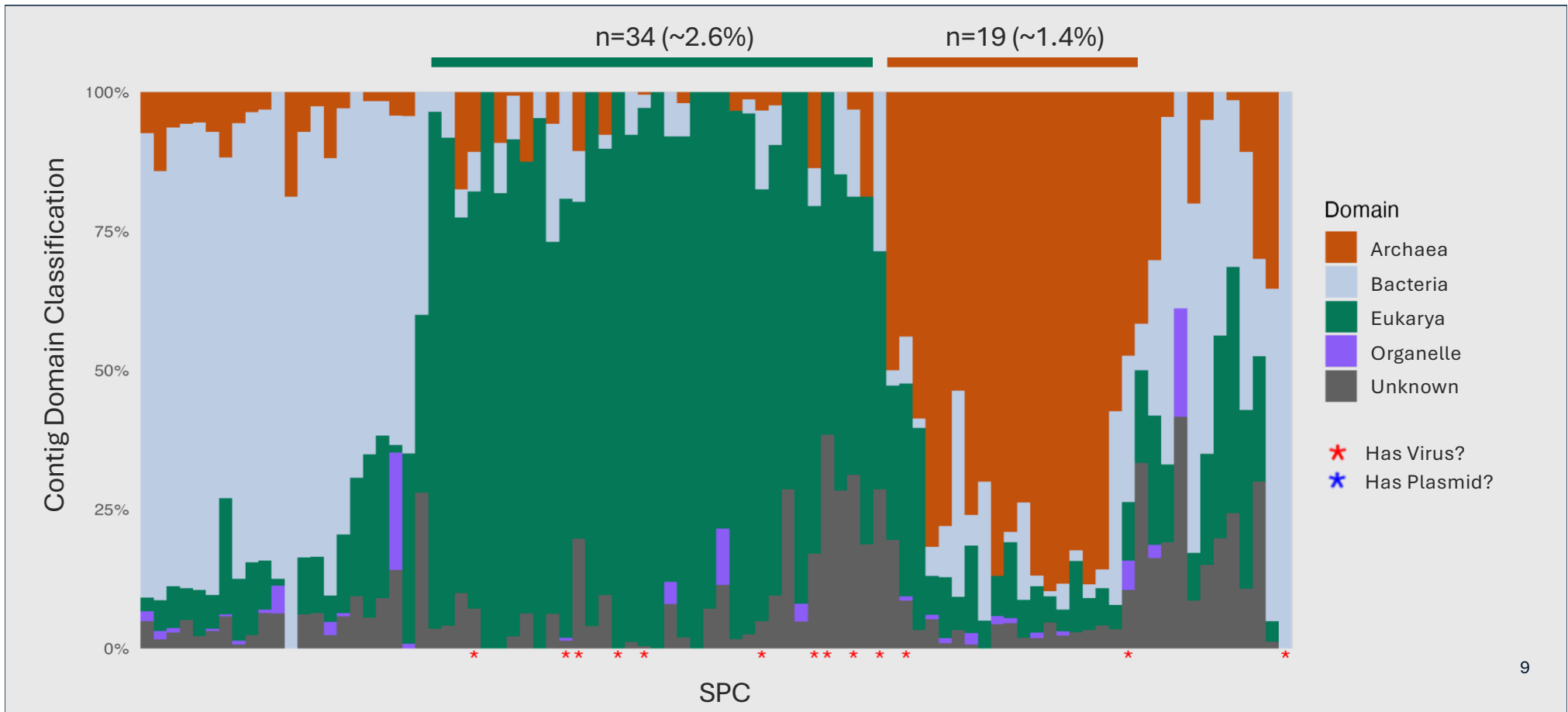
with Sanjay Srivatsan



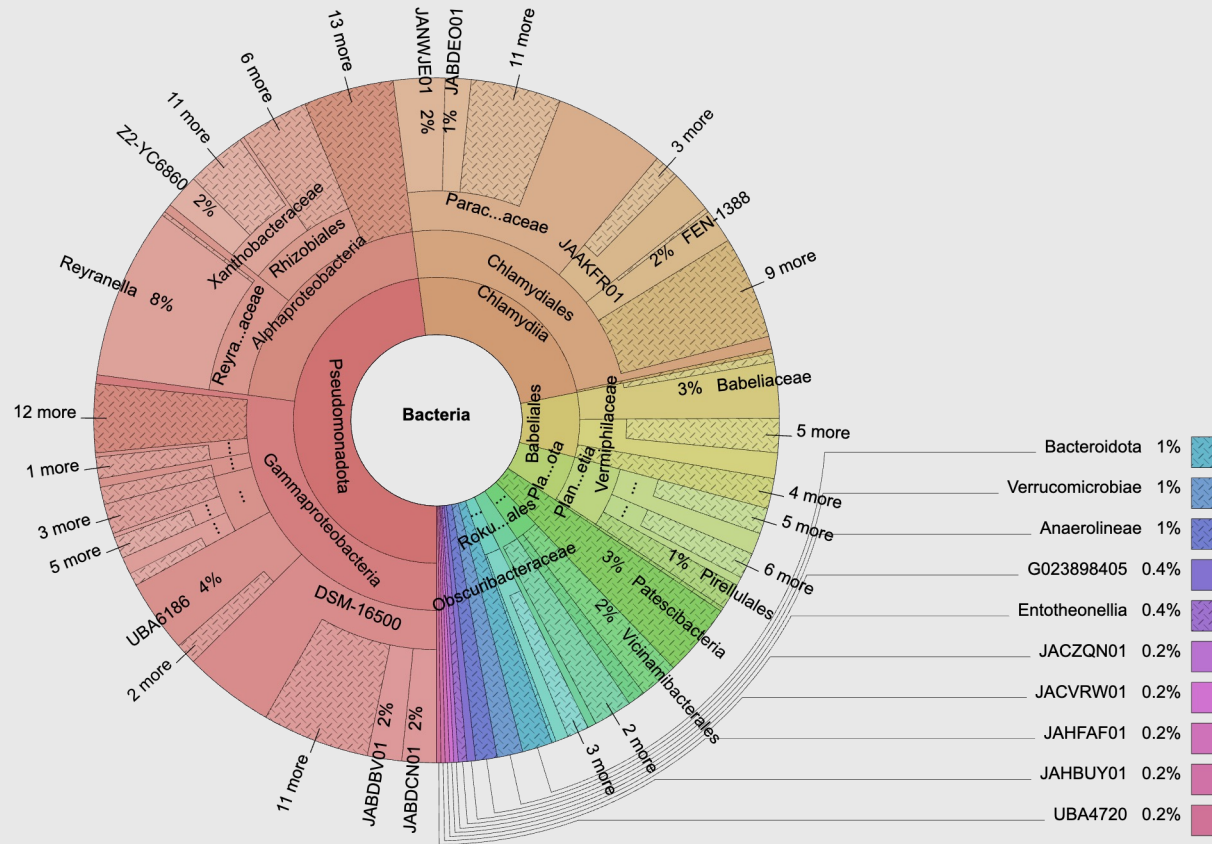
# What's in the Tube? Mostly Bacteria



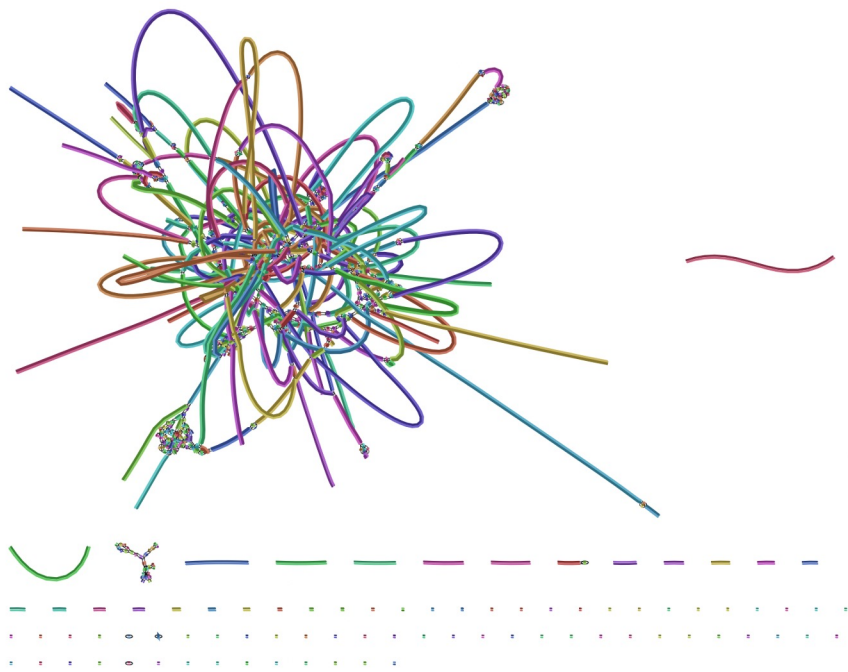
# What's in the Tube? Mostly Bacteria



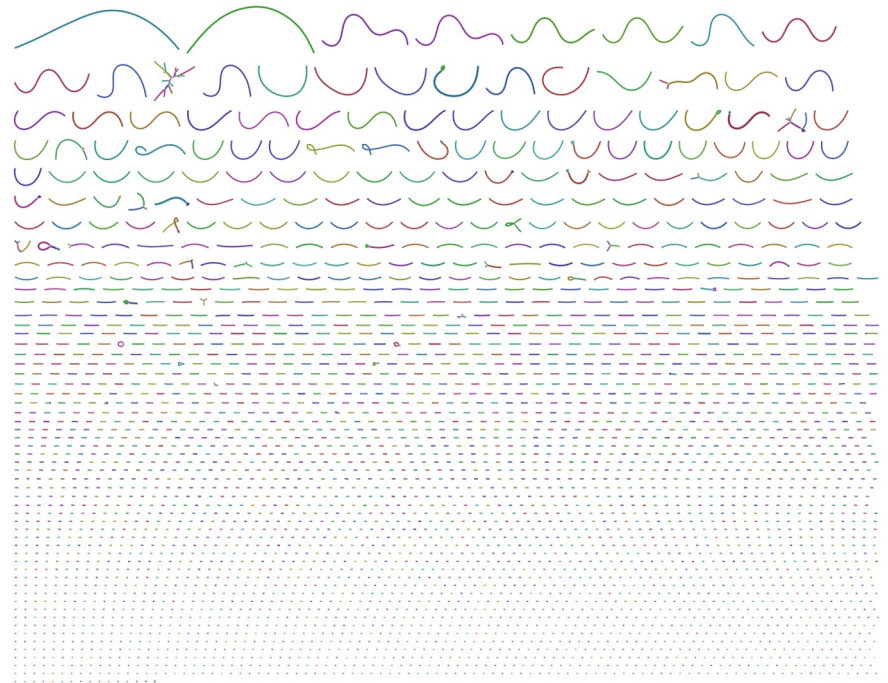
# What kinds of bacteria?



# Genome Assembly - Soil Examples



3.5 Mb - 795 contigs - 921 edges  
*soil\_glandiform*



4.2 Mb - 3.5k contigs - 160 edges  
*soil\_radular*