Introduction

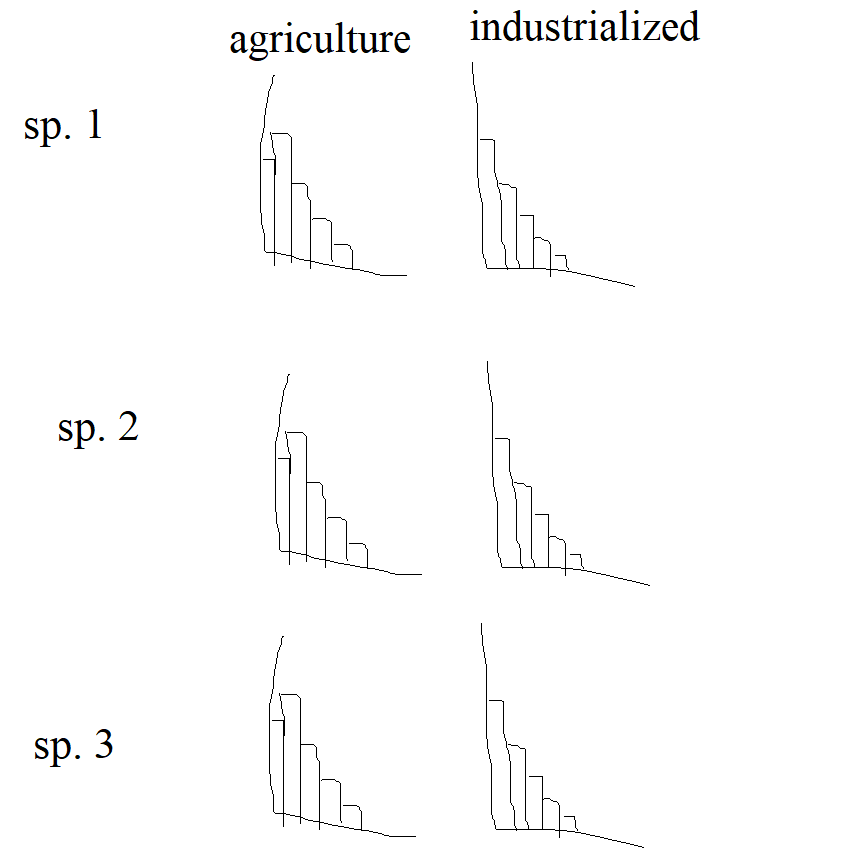
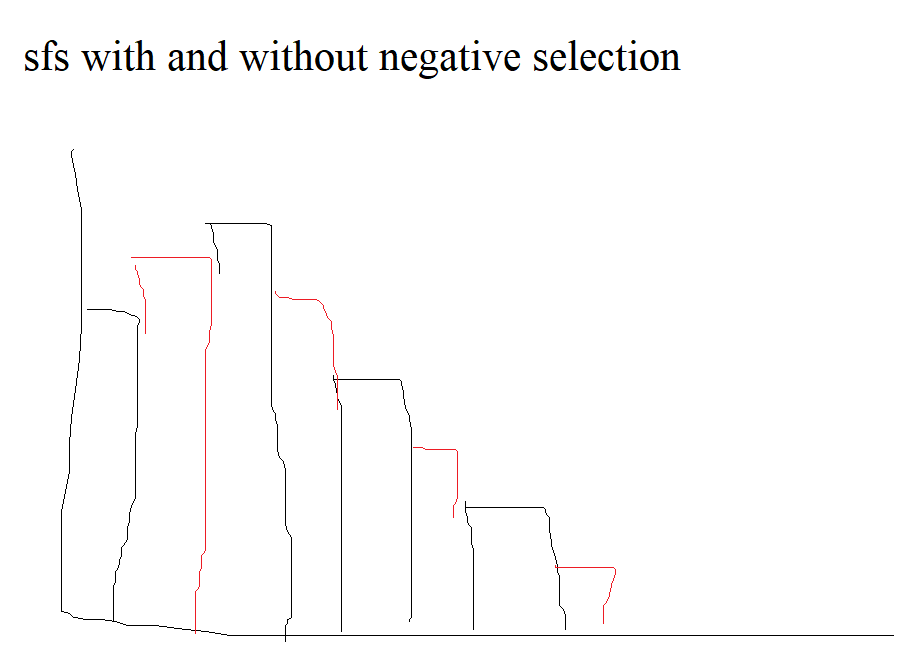
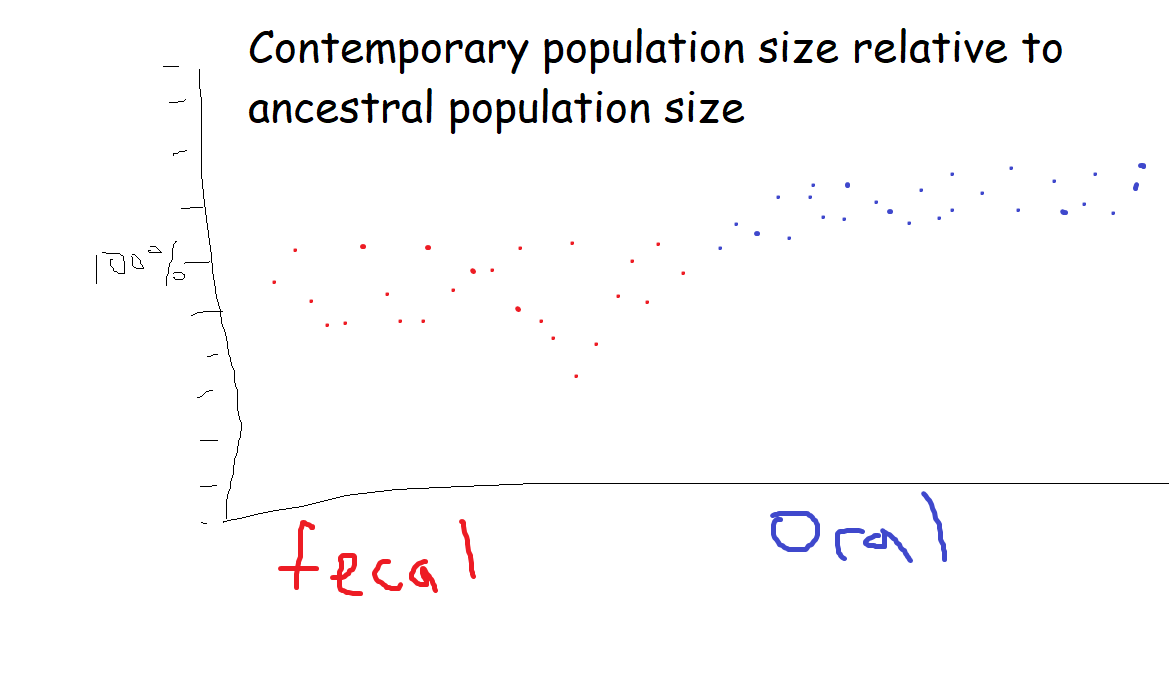
* Previous studies show broadly reductions in genetic diversity in the microbiome when “westernizing”, “industrializing”, or “modernizing” diets.
  + Not just diet, also sanitation (hygiene hypothesis)
  + Also Antibiotics -- Read (Missing Microbes)
* Studies show loss of genetic diversity (typically species-agnostic) when transitioning from hunter-gatherer to agricultural society (Sonnenburg)
  + Tie in with general health effects? E.g., “Vulnerability of the industrialized microbiota (Sonnenburg)
  + Elevated rate of HGT (Groussin)
* Also look at loss of clades (Tett 2019)
* Here we examine within-species changes in genetic diversity due to westernizing / industrialization etc. (Sonnenburg, Schnorr 2014, Wibowo)
* Mention with context of species-specific work, e.g., Cornejo (strep), Tett 2019 and 2021 (P. copri), Achtman and Zhou (Strep), Hershberg (M. tuberculosis) + Liu + Pepperell 2013
* Mention with context of linking human demographics with microbiome, e.g., Eisenhofer (Japan), Nandita’s 2019 paper,

Results

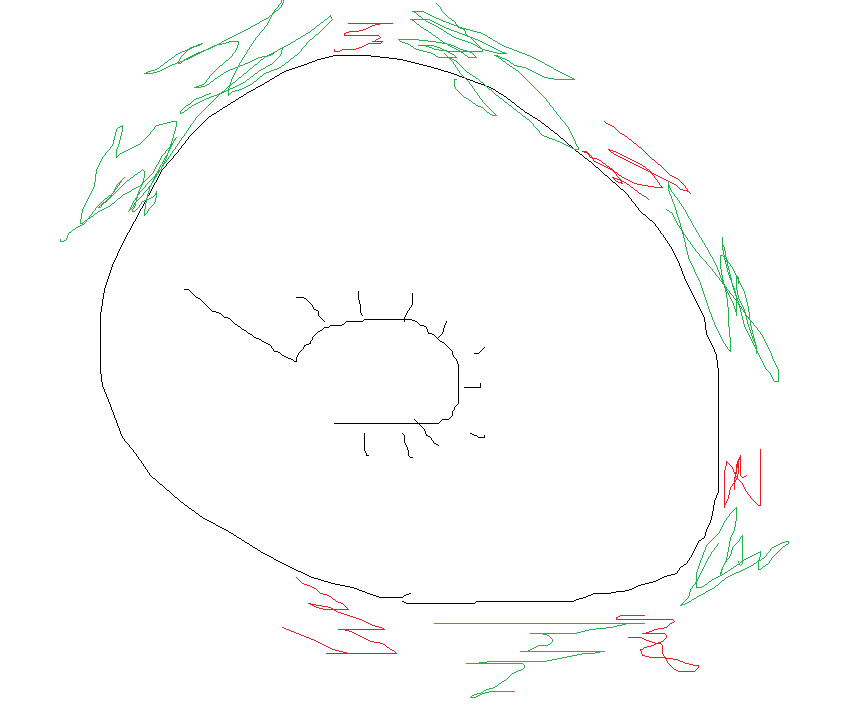
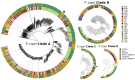
1. Pi between Madagascar / Ethiopian populations vs. HMP data
   1. A priori expectation is that there is a difference in pi between these two populations  
        
      Possibly some form of table which shows the values (placeholders relative to difference between two populations)  
        
      I expect there to be more nucleotide diversity in the oral microbiome when transitioning to an industrialized microbiome as shown in Cornejo. I expect the opposite trend in the gut microbiome based on our preliminary data

|  |  |  |
| --- | --- | --- |
|  | Agricultural | Industrialized |
| Oral | Expect lower pi | Expect higher pi |
| Gut | Expect higher pi | Expect lower pi |

* 1. If possible, it’d be nice if we could show Agricultural vs. Industrialized microbiome for both oral and gut

1. SFS of prevalent microbial species (maybe multiple species), perhaps *Prevotella*?  
   
   1. A priori I expect that in industrialized populations there is an increase in rare variants in the gut microbiome due to demographic contractions. I think it’d be cool to display this as some form of panel where we show the same array of species in different human populations but with an (ideally) visual difference in trend.
2. SFS with negative selection  
     
   I think we should frame this similarly to how negative selection is addressed in Cornejo. In summary, when accounting for negative selection in our model, we are able to fit a model SFS which has better(?) model fit than an sfs without negative selection, which suggests that our data has undergone negative selection 🡪 further evidence for population bottleneck.
   1. We can contrast the sfs with negative selection for agricultural and industrialized societies, maybe industrialized has more negative selection, or incorporation of negative selection has better fit (haven’t checked this out yet)
3. Demographic models
   1. Example: 
   2. We can do this kind of figure for other summary statistic.
   3. Ideally some sort of large summary figure which shows a clear difference between fecal and oral bacteria
   4. Prioritize data for industrialized populations as demographic inference will tell us what happened since migration out of Africa.
4. Model comparison or evaluation metric, e.g., LD
   1. If this does not work correctly, we can describe how to improve models
   2. This can probably be shown with a table for each species, or something along those lines. Alternatively, this is really data heavy so maybe something for the supplement?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Species 1 | | | | | |
| Model | LL | N\_a | N\_e | T1 | T2 |
| 2 epoch |  |  |  |  |  |
| 3 epoch |  |  |  |  |  |
| Bottleneck |  |  |  |  |  |
| Exp. Growth |  |  |  |  |  |

1. Possibly phylogeny or some different method to validate model findings  
     
   I’m thinking something like a phylogenetic tree diagram (like this one):  
   where we show in green clades which have persisted from agricultural to industrial and then in red we show clades which only exist in agricultural microbiomes. In this way we can demonstrate that there is loss of clade and/or loss of genetic diversity associated with industrialization of the human microbiome.

Discussion

* Pending on results
* In summary, I think our big takeaway is that we can:
  + Perform within-species population genetics and demographics analysis on a wide array of microbiome species (oral and gut)
  + When doing so, we find that (RESULTS PENDING) there is an environment-wide increase in genetic diversity in the oral microbiome when industrializing the diet / microbiome, and vice versa for the gut microbiome
  + This has *implications* (I need to read more about the implications) for human health

Look at GTDP

Phylogenetic regression