**Nov 4 – 10**

**PICRUST**

I attempted to download PICRUSt on the Comsol computer according to the instructions on the website provided by Kathleen. -   
<https://huttenhower.sph.harvard.edu/picrust/>

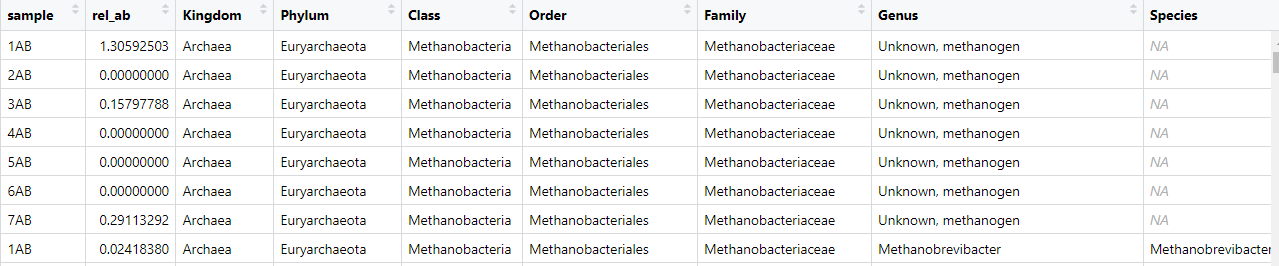
I was running into a lot of issues using the conda installer, so I followed the ‘from the source’ instructions. The general process was as follows :

1. Download the miniconda installer from the anaconda website.
2. Run it to install conda.
3. Open quest in PUTTY because the instructions are in unix commands, windows powershell doesn’t speak unix ( I think)
4. Follow the wget command from the github tutorial
5. Transfer the file from quest to the computer using WINSCP
6. Follow the instruction line to create a conda environment called picrust2. This step only needs to be done once. (I think)
7. Activate picrust in this environment using the unix command in the tutorial
8. Run the final pip- install command.
   1. This is the stage where I got stuck. After running the command, I got a number of errors saying that the computer was unable to “build wheels for biom format” I do not know where this error came from, what it means or which or if any of the previous steps were done wrong.

**Classifying unknowns**

A large portion of the genera for the community at large (bacteria and archaea together) were either unable to be defined using the MIDAS classifier we fed to QIIME, or were classified as not yet being named or classified. These two groups were reperesneted by NA or midas\_#### (number used only for identification on MIDAS) I coded a few lines that identified and combined these genera. This code is generalized, meaning it can be used for future data. It helps the charts and tables be less cluttered.

This screenshot shows how the code changes the table:

****

And here is a screenshot of the commented code. A close-up of a computer screen

Description automatically generated

This code uses the dplyr package to perform vector operations, or operations that affect every row in the data frame. Ifelse() is used within the mutate() function to replace

**Genes of Interest for qPCR**

**For Acetogens:**

* Acetogenesis in anaerobic digestors is a result of the Wood-Ljungdall pathway. There are multiple enzymes involved in this pathway that could be used as gene markers to detect acetogens.
* George suggested a paper that developed primers for **formyltetrahydrofolate synthetase,** and I found another paper that corroborates.
  + <https://www.sciencedirect.com/science/article/pii/S1075996409000559?via%3Dihub#sec1>
  + [**https://www.jstage.jst.go.jp/article/jsme2/26/4/26\_ME11123/\_pdf**](https://www.jstage.jst.go.jp/article/jsme2/26/4/26_ME11123/_pdf)
  + However, this paper also talks about the same gene potentially being exhibited in syntrophic acetate oxidizing bacteria; I do not yet know what this means o
    - [**https://www.sciencedirect.com/science/article/pii/S0306261916308364#bi005**](https://www.sciencedirect.com/science/article/pii/S0306261916308364#bi005)
* There are doubtless other enzymes unique to this pathway or other genes that could be used, but this has been done before and has been shown to work

**Acetoclastic Methanogens**

* Chat GPT said that acetate kinase is a gene of interest. Acetate Kinase is a significant component of the acetoclastic enzyme pathway
  + Ferry, J. G. (1992). *Biochemistry of Methanogenesis. Critical Reviews in Biochemistry and Molecular Biology, 27(6), 473–503.* doi:10.3109/10409239209082570
  + <https://www.sciencedirect.com/science/article/pii/S0959652622033352#sec2>
* Primers for this gene are reported/used in the following
  + <https://link.springer.com/article/10.1007/s00253-020-10494-2>
  + This study also reports a variety of other genes involved in the acetoclastic pathway, including
    - *cs* (acetyl-CoA synthase subunit δ), *mtrH* (methyl-H4SPT/CoM methyltransferase subunit H), *cdhA* (CO dehydrogenase/acetyl-CoA synthase complex subunit α), *mtaA* (methyl-CoM reductase subunit α), *fwdB* (formylmethanofuran dehydrogenase subunit β), *fdhB* (formate dehydrogenase subunit β), *gapdh* (glyceraldehyde-3-phosphate dehydrogenase), and *gyrB* (DNA gyrase subunit β)
* I need determine what makes a gene of interest a good candidate for an RNA primer; all of these genes along with many others are involved in the process, so it is important to learn why some are commonly used as primers and some are not.

**Formate Reducers**

* As I have learned, there are a third category of methanogens: Those that use hydrogen and formate as electron donors but cannot use acetic acid.
* Formate dehydrogenase is used for this process; although I do not know if this is specific to this kind of microbe.

**Week 8**

10 hrs 20 mins

**Identifying Acetogens**

All sources indicate that the Wood-Ljungdahl pathway is the only relevant mechanism for acetate production in anaerobic environments. This pathway is well studied and well characterized. I have started to try and find a study that lists taxonomic groups that possess this pathway or to compile a list myself.

* Acetobacterium, Clostridium, Morella, Eubacterium, Sporomusa – acetogenic genera - <https://www.mdpi.com/2076-2607/10/2/397>

**PICRUSt** –

I spent some time trying to determine the issue with the PICRUSt download. I tried again a few times after moving some files around, but still got the same error. Today (11/18) I have an appointment with McKormick IT to help figure out what’s wrong.

**Coding**

In the meantime, I continued to analyze the data based on the functionality groupings created from the Bergley’s manual. In previous weeks I compiled a chart of all archaea present in the reactor that identified if each was hydrogenotrophic, acetolactic, formate reducing, or a combination of hydrogenotrophic and acetolactic. This week I coded a portion that imported this as an excel file and combined it with the existing charts in the R file. The information was converted into a single column that contained a string value which identified the metabolic characteristic of the taxonomic group of interest. This code is generalizable for any functionality metadata chart you could upload as long as its in the same format as the one I made.

These functional groups were then plotted in the same way as the previous taxnommic groups, with the relative abundance in the y axis and either time or sample number in the y axis, with each functional group stacked in different colors.

A graph with different colored squares

Description automatically generated

A screenshot of a computer

Description automatically generated

**PCR**

I followed Emily’s instructions to do the PCR; I created a new thermocycle method and saved it in the machine. Later today (11/18) I’ll run the samples on a gel to see if it worked.