Meeting #4 Agenda
Maddy to take minutes

1. **Dataset overview portion of the proposal**
   a. Are we meant to explain what we did with our dataset in each (major) step of the QIIME2 pipeline, or only provide the info asked for in the checklist?
   b. Should we describe how we **plan** to filter/control for birth method in R, or just stick to how and why we **already have** filtered our samples to only include 4 month old infants?
2. **Confirm filtering of samples to only include 4 month old infants**
   a. We filtered table in QIIME2 for future steps
   b. Is there a difference between that and filtering the manifest file in R first and then re-running the QIIME2 pipeline?
3. **Hypotheses:** how do you state a hypothesis without making a prediction (as stated in the rubric)?
4. **Experimental Aims:**
   a. **General formatting** → run through the aims we have so far and their logical order.
   b. **What is the grouped filtered table in the practice proposal?** Not quite sure the equivalent in our study →
   c. **General Steps for PieCrust analysis? → on canvas**
   d. **Would we be looking at impact of infant feeding type and maternal BMI factors affecting gene expression, or narrow in on one?**
5. **Any feedback on how our proposal is looking? (See attached pdf).**

Minutes

Divide BMI into low and high categories and then further parse out feeding mode - Avril said that is fine
   - As long as we can provide justification/rationale
If we'd prefer to create grouped/filtered table in R rather than qiime that's ok

PieCrust is done in Qiime - use abundance data.
   - "This bacteria is present in this proportion so this function should be present in the same proportion"
   - Output will be similar to taxonomy output so you can apply the same downstream steps

R package that runs DeSeq and PieCrust data automatically - Avril will email it to us later tonight, otherwise reach out to him to remind him

Just follow Qiime checklist - not necessary to touch on steps from Qiime not included in checklist (eg. training the classifier)
   - Review rubric

- If you plan to **remove** samples entirely then mention it but we don't plan to do that for birth mode we just plan to control
- Mention that you plan to **control** for birth mode eventually in R in the 4 month age group

It's fine that we filtered the table then ran from there rather than filtering the manifest and re-running everything

"Hypotheses are not written as predictions" - Based on literature, this is what's been seen before, it is likely that this will occur/justifies our exploration of this concept
- It is not a random guess, it is a hypothesis based on fact - that's all Dr. Sun really wants
- Based on literature, this is the justification for us doing this analysis

Justification for how we are dividing BMI into two categories can come from just what our dataset looks like rather than a full literature review
- BMI under 26 is low and over 26 is high

Continuous BMI for first two experimental aims and then switching to low and high for the rest

You can use a continuous variable for the stats but beta diversity is usually best graphically represented using categories rather than a colour gradient for points on the graph

Comments on proposal:
Well done in general, easy read, things make sense
Title: "With respect to" is better than "through" (feeding type)
Intro: Contradiction where mom BMI does and does not affect infant microbiota - fix that
Intro: Be careful with how strongly you word things - instead use this is "possibly" the mechanism…. Remember that we don't know things for sure or even likely
Natural composition of breast milk: be more specific about what we mean
Impact of research - this is extremely important
- Add something like this to the intro as well - think of it like you are trying to convince someone to invest in this project
Predicted functional phenotype
Gene expression data is not accurate - avoid this term
Explain more about what we mean by interactive effects in last sentence of research objective - pretend it's not Avril reading it but someone unfamiliar with the entire project
Experimental aims: a, b and possibly c could actually be combined???
- Definitely combine a and b - here is why bmi matters, here's why feeding method matters
- c: could be it's own aim but probably can be combined with a and b
- One aim for diversity, one aim for DeSeq? - might be easier for chart?
- Be more specific in a and b about alpha and beta diversity - specify
- Be more specific and say diversity rather than composition
- d collaboratively affect - "use in a manner dependant on…"
- D in middle of paragraph: female gut microbiota (adults or babies?)

- Adult women might be a hard sell - so many things could affect microbiota at that point but if it's babies you can say possibly affected by oligosaccharide diet
- DADA2 citation - there should be a paper associated with this tool and that's what you're supposed to be citing - ask avril if you can't find it

Lab notebook
PDF not necessary?
Get this fully updated by next week because he'll be checking