Ryan W, Christina X, Saray F, Amitis E, Parsa S

Group-1
For Feb 5, 2026 meeting

Goal for today's meeting: align on analysis plan + feasibility + timeline for the three aims we identified last meeting.

What we did

- A Github repository was made and previous meetings were all uploaded in addition to the more important R files required for assignments.

What we plan to do

- In our last meetings, 3 aims were discussed:

Aim 1 - alpha/beta diversity across gastric cancer stages, with sex stratification:

A. Define groups clearly:

- Confirm which cancer progression stages we are comparing (e.g., control → gastritis → metaplasia → dysplasia → cancer, or whatever the dataset uses) or are we re-grouping them (decrease the number of groups)
- Decide how we will handle group sizes and whether we need to merge stages.
- Go back and find all the literature that has accessed the dataset before.

B. Alpha diversity

- Compute alpha diversity metrics

Compare:

- Across stages (all samples)
- Across stages within each sex (stratified analysis)
- Sex differences within each stage (M vs F within-stage comparisons)
- Decide on statistical approaches to test for significance and how to report effect sizes.

C. Beta diversity

- Choose metrics
- Test stage stratified by sex

Use PERMANOVA to test for

- stage effect
- sex effect
- stage × sex interaction (if possible)

Decisions we need from the meeting: Which alpha/beta metrics to prioritize for the final report (keep it focused).

1. Looking at the taxonomic differences between the stages/groups → indicator species
2. Going into functional analysis (if you see differences, maybe there are:)
    a. Pathways that relates to inflammation
    b. Sex differences where one sex is doing better
    c. Metabolism, enzyme, nitty gritty biologic function (k***egg*** terms)

Aim 2 - Taxonomic differences between stages

A. Identify taxa that differ by stage
    - Generate relative abundance summaries at multiple levels (phylum/genus/species depending on resolution).
    - Compare taxa across stages (and possibly within each sex if sample size supports it).

B. "Indicator species" / differential abundance approach
    - Propose an "indicator taxa" method (e.g., indicator value analysis OR a differential abundance method appropriate for the course).
    - Output:
        a. list of key taxa associated with each stage
        b. effect
        c. basic visualizations

C. Tie the taxonomic findings back to Aim 1
- If beta diversity differ between stages, identify which taxa drive that separation.
- If sex stratification changes patterns, identify taxa that differ by sex within stage.

Question/Decision: Which taxonomic level should we analyze (genus vs species)

Aim 3 - Functional Analysis

A. Determine what functional data we actually have/need
- If we have predicted function outputs already, we'll work directly from those

B. Functional comparisons

- Compare pathway abundances across stages (overall), then stratify by sex if feasible.
- Include the relevant pathways found in the table of definitions.
- Have ranked pathways by stage association

- a small set of biologically interpretable pathways to discuss in the final write-up

Question/Clarification: What scope is realistic? Should we aim for having a small pathway list? Focus on the most important functional differences?

What we want to discuss:

- What to do going forward and estimated deadlines to finish the project on time or to be ahead
- The question/clarifications written after each Aim