Effects of Alcohol Exposure on *D. melanogaster*

Austen Miller

June 14, 2024

Diverse communities of microorganisims live within the guts of all animals. The diet of the hosts of these microscopic communities can affect them in a variety of ways. Studying the composition of these communities in model organisms can provide insights into how they might function within humans.

* (going to remove the red text) Define full-length 16s rRNA sequencing and its uses for microbiome composition

Full-length 16s rRNA sequencing uses the ribosomal sequence of bacteria to analyze the taxonomic composition of bacterial communities, such as the microbiomes of animals. This is possible because ribosomal sequences mutate relatively slowly due to their importance in biological functioning. The similarity of the 16s gene within different bacteria with only a few changes then makes it useful as a tool to analyze community composition when sequenced.

In this particular case we will be looking at the effects on the microbiome of the fruit fly species *Drosophila melanogaster* after being exposed to alcohol. D. melanogaster is useful as a model organism due to its relative ease to work with, short generation times, and many analogous genes to humans. These flies will be exposed to 50% ethanol vapor, disected, and their intestinal communities analyzed.

* + Pros and cons of using full-length 16s (~1500bp) vs “normal” 16s (where only ~500 bp are sequenced)

Oxford Nanopore sequencing allows the use of full-length 16s reads (~1500 bp), allowing the entire gene to be sequenced. This is in contrast to the norm where much shorter sequences (~500 bp) must be used. These longer reads allow for increased specificity in taxonomic classification of bacteria. However, full-length 16S rRNA sequencing generates large amounts of data, which can be expensive to process and requires substantial computational resources for analysis. Standardization across different laboratories and studies is essential but challenging to achieve. Additionally, Nanopore sequencing exhibits a relatively high error rate on raw sequences compared to standard Next-Generation Sequencing (NGS) devices such as Illumina (Delahaye and Nicolas 2021).

* State the problem: lack of bioinformatic pipelines that analyze full-length 16s sequencing

Unfortunately, the existing pipelines for full-length 16s sequence data analysis are not always consistently updated. The rapid advance of technology can make the old analysis tools obsolete and difficult to replicate.

* NanoRTax and others?
  + Run NanoRTazx and identify new pipelines

One of the pipelines we will be trying is theNanoRTax pipeline NanoRTax allows “...a nextflow-based pipeline for bacterial taxonomy classification and sample diversity analysis of nanopore full-length 16S rRNA amplicon reads” (Rodríguez-Pérez, Ciuffreda, and Flores 2022). We will also be working on developing our own pipelines using known tools for bacterial taxonomy classification and R for sample diversity analysis. The computational skills gained in the IQBio program’s carpentries will be quite useful in working with these programs and understanding their results.

* Del Rubric:
  + - What is the main research question that you will be addressing this summer? (Question)\*
  + - What is already known about it? (Background)\*
  + - Why is it an interesting and/or important question to address? (Purpose)\*
  + - What experiments/analysis will you carry out to address this question? (Experimental Plan)\*
  + - How will you analyze your results? (Proposed strategy for analysis)\*
  + - How can you incorporate your newly acquired computational skills to improve your research project?
  + - What obstacles do you foresee encountering?
  + - How will you overcome these obstacles? (Who you gonna call?)

Possible challenges that we will be presented with are trouble with the computational tools being worked with, the details of *D. melanogaster* biology, and missing technical terms while working in Spanish and English. To overcome these obstacles, I will communicate with the PI, lab mates l and look for resources using the internet.

References

Delahaye, Clara, and Jacques Nicolas. 2021. “Sequencing DNA with Nanopores: Troubles and Biases.” *PLoS ONE* 16 (10): e0257521.<https://doi.org/10.1371/journal.pone.0257521>.

Rodríguez-Pérez, Héctor, Laura Ciuffreda, and Carlos Flores. 2022. “NanoRTax, a Real-Time Pipeline for Taxonomic and Diversity Analysis of Nanopore 16S rRNA Amplicon Sequencing Data.” *Computational and Structural Biotechnology Journal* 20:5350–54.<https://doi.org/10.1016/j.csbj.2022.09.024>.