Nov 19, 2024 | Meeting w Lydia

Attendees: Lydia Abels Robert Alvarez-Quinto Ronan Keener

Notes

- Goal is pipeline
    - Raw data → trimming → assembly → blast → virus diagnostics
- Step 1 - trimming raw data in MSI
    - Been using BBDuk, but MSI does not offer that
    - MSI has trimmomatic - look at website for more info
- Helpful resource - https://github.com/matthewmoscou/Emo

Virology Pipeline

1. Raw data, 2 files (150 paired-end reads) 40M reads
2. Trim adapters and remove low-quality reads (BBDuk, Trimmomatic)
3. De novo Assembly (SPAdes or Trinity)
    a. Might try Haploflow? Did research and it's reported to be good for viral genomes
4. Remove host reads (Blastx against Arabidopsis genome)
    a. Bin hits and no hits
5. Analyze the "no hits" file using Blastx against the full nr database.

Action items

- ☑ ~~Trim with trimmomatic in MSI~~
- ☑ ~~Look into next flow (nf-co.re)~~
    - denovotranscript: Introduction

12/13/24 - Notes

- Really nice work!!!
- Everything is working as expected?
- Next step would be assembly with trinity
    - Need to use scratch.global

Robert Notes:

- Invite Lydia to Lab meetings every two weeks (in-person)
    - To share progress with teams
    - Learn about what everyone is working on
    - Journal Club discussion

- Trello Board
    - Install trello
    - Team List fill out
    - Card for this project

Problem with Trinity