Listeria ASSemblyEs
Comparison

Sachin, Will, George, Sreenivasan AC, Jialin, Sid
Overview

- Background and Objectives
- Functionalities
- Technologies Used
- Questions
Background

- **Predictive Web Server**: Server that hosts web-pages that predict something
  - **Our Main Prediction**: Strands with least genetic distance

- Storing data, Processing data, and Delivering Results
- Raw fastq → Assembled Genomes → Predicted Genes → Functional Annotations → Comparative Genomics (Phylogenetic tree)
Objectives

● Create easy-to-use web-based tool to distinguish Listeria phenotypes
● Users input biological data (fastq files)
● Our system processes and returns assembled genome, genes and annotations, and phylogenetic tree to email
● Easy-to-use JBrowse (scalable genome browser)
Functionalities

- de-novo assembly tool
- Predict genes from assembled genome
- Feature to Annotate genes
- Create phylogenetic tree (showing genome and least distance strands)
- JBrowse - Genome Browser
Genome Assembly Tools

- Team 3 Genome Assembly Tools
  - Requiring pre-trimmed fastq files
  - SPAdes
  - QUAST
Team 3 Gene Prediction Tools

- Glimmer
- Prodigal
- GeneMarkS-2
Annotation Tools

- CARD
- VFDB
Genomic Distance Tools

- FastANI
- Mash
Web Server - Front End

- HTML
- Bootstrap (CSS)
- Javascript
- D3.js, phylogenetic tree visualization (tnt.tree)
- Genome browser (Jbrowse)
● Why Bootstrap?
  ○ Popular CSS framework
  ○ Less CSS styling work

● Why vanilla Javascript
  ○ JS frameworks (Vue.js, ReactJS) complex
  ○ Not required for our usecases
Why D3.js?

- Popular Data visualization framework
- Dynamic, interactive
- We have some experience
Web Server - Backend

- Why Flask?
  - Python based syntax
  - Lightweight server
  - for building Restful API for handling request from frontend
  - Simpler - without ORM or MVC or complex functionalities which we don’t need
Proposed Approach - Predictive Websvserver
Server Configuration and Automatic deployment

- Apache -> (Caddy -> gunicorn) -> Flask

  The Flask service will run in background. Restart upon push events on master branch at Git.

- Issues taken into consideration:
  - Virtualenv, Dependencies, Process Management
Task Delegation

- Front End (Jialin, George)
  - JBrowse (George)
  - UI (Jialin and George)
- Back End (Sachin, Siddhartha, Sreenivasan, Will)
  - Genome Assembly (Will, Sreenivasan)
  - Gene Prediction (Siddhartha)
  - Annotation (Sachin)
  - Phylogenetic Trees (Siddhartha)
Questions?