WEB - SERVER

BACKGROUND & STRATEGY

Manu Tej Sharma
Arrojwala
Mingyue Zhang
Sachin Kumar
Dipro Chakraborty
Priyam Raut
Nishant Gerald
INTRODUCTION

- Salmonella Enterica affects 17 million people worldwide each year
- Difference in the various serotypes of Salmonella are caused by their genetic makeup
- Has a large stable core, but has abundance of accessory genes (Pathogenicity islands, TE etc.)
- Visualize conservation and variation across genomes wrt chromosomal location and structural features


OBJECTIVES

● Assemble the input reads
● Analyze the assembly and predict annotated genes
● Identifying the strain as a phylogenetic tree (or heatmap)
● Calculate distance from the strain in the existing database
● Virulence factor and antimicrobial resistance profiling
● Visualize results in an effective manner
Why web-server?

- Resource invariant(ish)
- Ready to go
- Global Access
- Information rich
- Easy on eyes
Welcome to STRING

Protein-Protein Interaction Networks

ORGANISMS: 5090
PROTEINS: 24.6 mio
INTERACTIONS: >2000 mio

SEARCH
<table>
<thead>
<tr>
<th>node1</th>
<th>node2</th>
<th>node1 accession</th>
<th>node2 accession</th>
<th>node1 annotation</th>
<th>node2 annotation</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>DD95_04165</td>
<td>pabB</td>
<td>CY43_08805</td>
<td>CY42_09315</td>
<td>Antithranilate synthase component 1;</td>
<td>Para-aminobenzoate synthase component</td>
<td>0.845</td>
</tr>
<tr>
<td>DD95_04165</td>
<td>trpA</td>
<td>CY43_08805</td>
<td>CY43_08825</td>
<td>Antithranilate synthase component 1;</td>
<td>Tryptophan synthase alpha chain;</td>
<td>0.957</td>
</tr>
<tr>
<td>DD95_04165</td>
<td>trpB</td>
<td>CY43_08805</td>
<td>CY43_08820</td>
<td>Antithranilate synthase component 1;</td>
<td>Tryptophan synthase beta chain;</td>
<td>0.943</td>
</tr>
<tr>
<td>DD95_04165</td>
<td>trpC</td>
<td>CY43_08815</td>
<td>CY43_08815</td>
<td>Antithranilate synthase component 1;</td>
<td>Multifunctional fusion protein;</td>
<td>0.993</td>
</tr>
<tr>
<td>glyA</td>
<td>ilvA</td>
<td>CY43_13635</td>
<td>CY43_20455</td>
<td>Serine hydroxymethyltransferase; Catalyzes threonine</td>
<td>L-threonine dehydratase; Catalyzes threonine</td>
<td>0.949</td>
</tr>
<tr>
<td>glyA</td>
<td>trpA</td>
<td>CY43_13635</td>
<td>CY43_08825</td>
<td>Serine hydroxymethyltransferase; Catalyzes threonine</td>
<td>Tryptophan synthase alpha chain;</td>
<td>0.941</td>
</tr>
<tr>
<td>glyA</td>
<td>trpB</td>
<td>CY43_20455</td>
<td>CY43_08820</td>
<td>Serine hydroxymethyltransferase; Catalyzes threonine</td>
<td>Tryptophan synthase beta chain;</td>
<td>0.939</td>
</tr>
<tr>
<td>ilvA</td>
<td>trpA</td>
<td>CY43_13635</td>
<td>CY43_08825</td>
<td>L-threonine dehydratase; Catalyzes threonine</td>
<td>Tryptophan synthase alpha chain;</td>
<td>0.949</td>
</tr>
<tr>
<td>ilvA</td>
<td>trpB</td>
<td>CY43_20455</td>
<td>CY43_08820</td>
<td>L-threonine dehydratase; Catalyzes threonine</td>
<td>Tryptophan synthase beta chain;</td>
<td>0.949</td>
</tr>
<tr>
<td>ilvA</td>
<td>trpC</td>
<td>CY43_20455</td>
<td>CY43_08815</td>
<td>L-threonine dehydratase; Catalyzes threonine</td>
<td>Multifunctional fusion protein;</td>
<td>0.949</td>
</tr>
<tr>
<td>pabB</td>
<td>DD95_04165</td>
<td>CY43_09315</td>
<td>CY43_08805</td>
<td>Para-aminobenzoate synthase comp...</td>
<td>Para-aminobenzoate synthase comp...</td>
<td>0.845</td>
</tr>
<tr>
<td>pabB</td>
<td>trpA</td>
<td>CY43_09315</td>
<td>CY43_08825</td>
<td>Para-aminobenzoate synthase comp...</td>
<td>Tryptophan synthase alpha chain;</td>
<td>0.958</td>
</tr>
<tr>
<td>pabB</td>
<td>trpB</td>
<td>CY43_09315</td>
<td>CY43_08820</td>
<td>Para-aminobenzoate synthase comp...</td>
<td>Tryptophan synthase beta chain;</td>
<td>0.995</td>
</tr>
<tr>
<td>trpA</td>
<td>DD95_04165</td>
<td>CY43_08825</td>
<td>CY43_08805</td>
<td>Tryptophan synthase alpha chain;</td>
<td>Antithranilate synthase component 1;</td>
<td>0.957</td>
</tr>
<tr>
<td>trpA</td>
<td>glyA</td>
<td>CY43_08825</td>
<td>CY43_13635</td>
<td>Tryptophan synthase alpha chain;</td>
<td>Serine hydroxymethyltransferase; Catalyzes threonine</td>
<td>0.941</td>
</tr>
<tr>
<td>trpA</td>
<td>ilvA</td>
<td>CY43_08825</td>
<td>CY43_20455</td>
<td>Tryptophan synthase alpha chain;</td>
<td>L-threonine dehydratase; Catalyzes threonine</td>
<td>0.916</td>
</tr>
<tr>
<td>trpA</td>
<td>pabB</td>
<td>CY43_08825</td>
<td>CY43_09315</td>
<td>Tryptophan synthase alpha chain;</td>
<td>Para-aminobenzoate synthase comp...</td>
<td>0.949</td>
</tr>
<tr>
<td>trpA</td>
<td>trpB</td>
<td>CY43_08825</td>
<td>CY43_08820</td>
<td>Tryptophan synthase alpha chain;</td>
<td>Tryptophan synthase beta chain;</td>
<td>0.958</td>
</tr>
</tbody>
</table>
EXPECTED FEATURES

- Interactive and easy-to-use
- Consistency in design
- Scalability to devices and browsers
- Visually informative
Pipeline

BASH

- Integrate pipelines from previous groups into bash script
- Generate data for Visualization
- Prepare downloadable outputs
ADDRESSING UI NEEDS

Provide customization and staging of the process converting raw genomic data to knowledge

Users should be able to
- Go through the entire process
- Execute individual stages of the process
- Go through the entire process after a specified stage
COMPONENTS

PHP

- Libraries are compatible with pipeline
- Integrates HTML to display views of MVC architecture
- Widespread documentation
- Mostly used server-side and serves HTML to visitors
COMPONENTS

LARAVEL

- PHP Framework
- Follows MVC architecture pattern
- Rapid prototyping framework - ideal for projects!
Why MVC? One word: Maintainability
What is a “Model”? 

- Constitutes the computation and execution part of an application.

- Completely hidden from user and is activated through validated inputs from the controller.

- Also responsible for generating visualizations.
What is a “View”?
- A view in MVC represents the component which is available for User Interaction
- Each stage of the pipeline would be expected to take user input and provide appropriate outputs, either as downloadable files or visualizations
- Separates user from backend processing
Welcome to the Salmonella Enterica Predictive Web-server

Created by Team-2 Web-server Group

Proceed to View Available Pipelines
What is a “Controller”?

- Bridging component responsible for interaction between view and model.

- Responsible for validations of inputs from View and outputs from Model before sending out data to either of them.

- Also responsible for invocation of specific programs with respect to the data received.
DATABASE

- Technology used: MySQL

- Things to store:
  - File paths
  - Data from intermediary steps for easy visualizations
VISUALIZATIONS

Tools: D3.js

What we’ll be visualizing?
- Genome Assembly — Post-QC results, Assembly Scores, etc.
- Prediction & Annotation — # of genes, gene lengths, etc.
- Comparative genomics — Phylogenetic trees (Using phylotree.js)
Additional Functionalities

- Structural variant analyzer for gene phylogenies.
- Batch mode
- Local deployment via GUI or command-line
- Cloud storage to web-server pipeline
- Add links to relevant external databases
- Email functionality
THANK YOU