Outline

- Quick platform/architecture intro, introduce main webserver - hosting with nginx (daniel)
- Screenshots of outputs embedded in powerpoint - show completed workflow outputs - gene prediction (Chloe), ANIclustermap (Charlotte), parsnp (joanne)
- Show screenshots of landing page, uploading, email, job tracking, etc - do upload demo (Bhavay)
- Eventually push to GitHub and have flowchart for whole pipeline

Email: team1.webserver@gmail.com
password - team1webserver@123

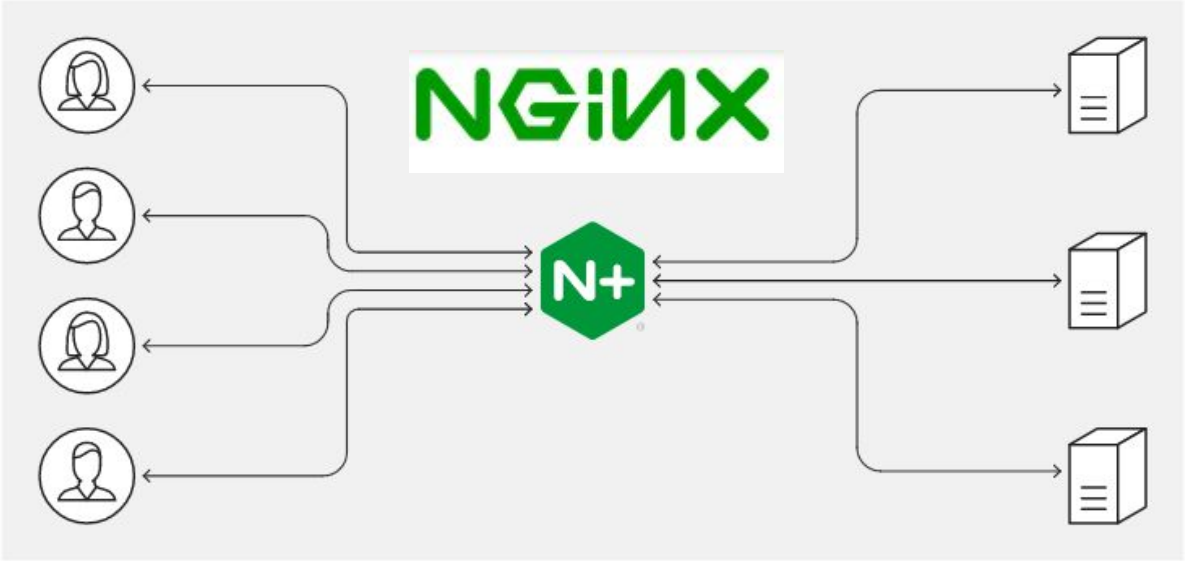
Predictive Webserver Results

Team 1
Bhavay Aggarwal, Joanne Boysen, Daniel Groves,
Chloe Pryor, & Charlotte Royer

NginX: Front-End

NginX

- Open source software for webservers
- Provides HTTP server features
- Mail proxy server features
- Supports our webserver architecture



Webserver Layers

NGINX

- Frontend server
- **Reduces load on backend servers**
- Reverse proxy between the internet and backend server
- High-performance
- Load-balancing and caching

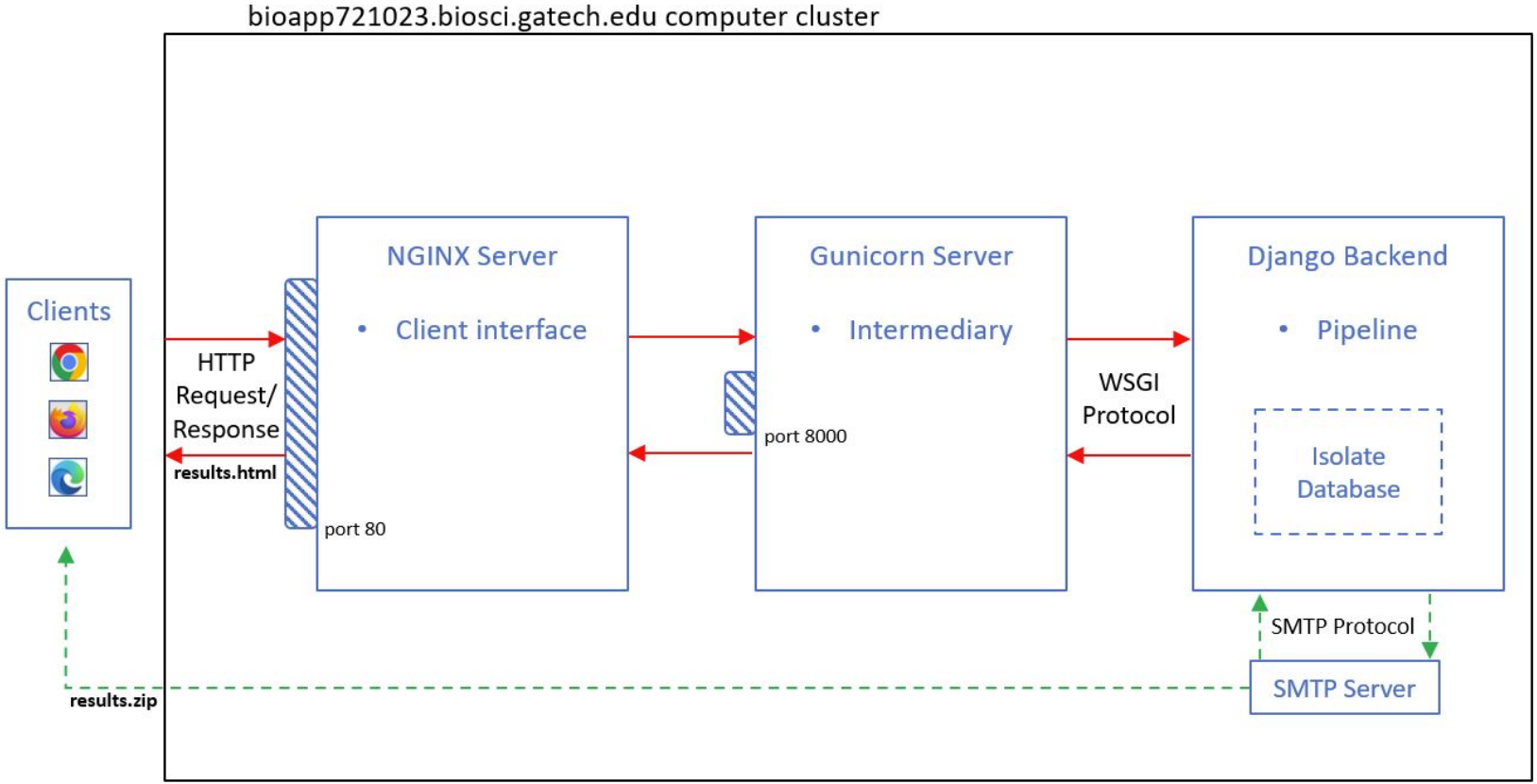
Gunicorn

- Application server
- Necessary when using Python-based frameworks
- **Forwards HTTP requests to the web application**
- Communicates via WSGI protocol

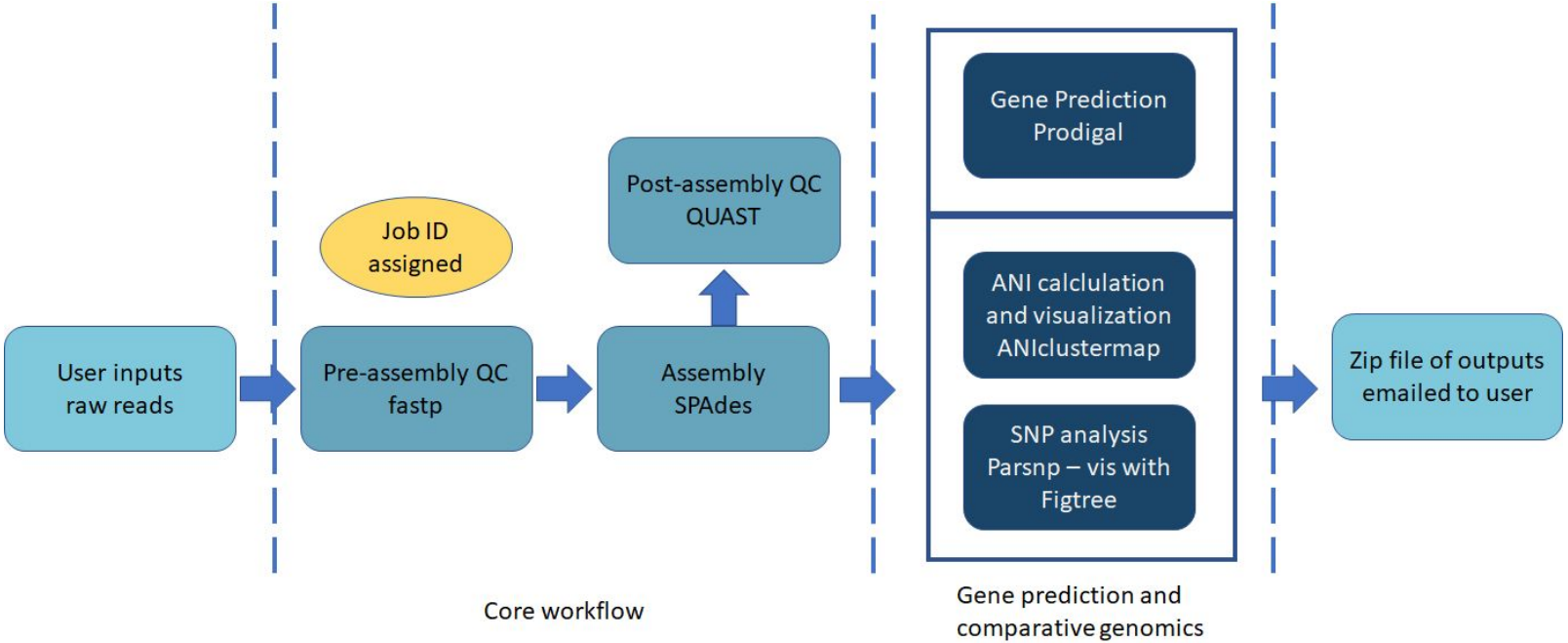
Django

- Backend server
- Processes data, application logic, etc
- Python-based web application framework
- Follows MTV structure for easy data storage, job organization, and presentation
- **Processes the raw reads and performs our pipeline**

Webserver Architecture



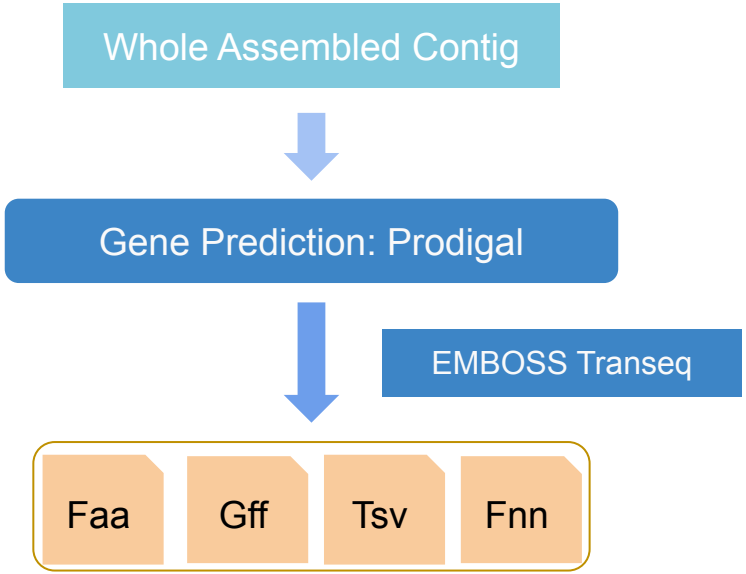
Workflow


User inputs raw reads
Job ID assigned
Pre-assembly QC fastp
Assembly SPAdes
Post-assembly QC QUAST
Core workflow
Gene Prediction Prodigal
ANI calclulation and visualization ANIclustermap
SNP analysis Parsnp – vis with Figtree
Gene prediction and comparative genomics
Zip file of outputs emailed to user

Webserver Output 1

Gene Prediction

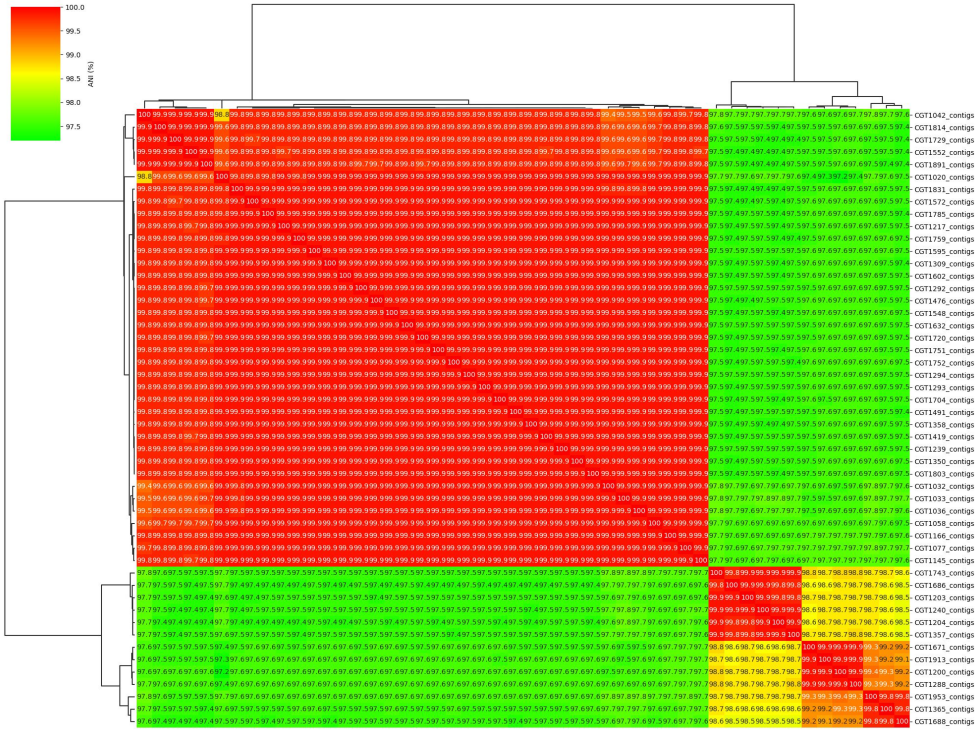
- Chosen tool: Prodigal
- Predicts genes from the assembled contigs (generated with SPAdes)
- Output: fasta and gff files bundled with other outputs in the zip file that is emailed to the user



Webserver Output 2

Comparative Genomics: Whole Genome Analysis

- Chosen tool : ANIclustermap
- Calculates all-against-all ANI between samples using fastANI
- Constructs heatmap visualization for user
- User input genomes are added to the database, and ANIclustermap is run on the entire folder
- Output compares user's input genomes to entire database
- Output is generated as .png and bundled with outputs in zip file



Webserver Output 3

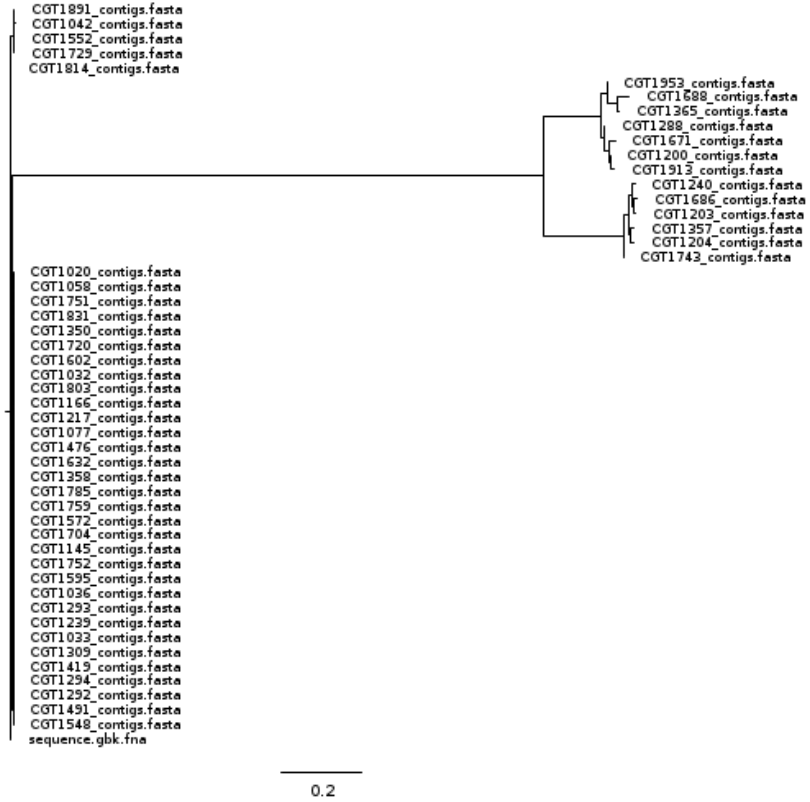
Comparative Genomics: SNP Analysis

Chosen tools: Parsnp and Figtree

- Parsnp excels at rapid core genome mulit-alignment
- Alignment based → fast SNP analysis and microbial core genome phylogeny
- With NCBI Reference Sequence: NZ_CP120944.1, EDL933
- Figtree is a graphical viewer of phylogenetic trees

Outputs included in zip file:

- Figtree phylogenetic tree PNG
- parsnp.tree
  - In case the user would like to visualize the data differently



Webserver Demonstration!

Webserver

Usage Instructions:

Attach a ZIP file contains the raw fastq files

Choose File | No file chosen

Email ID

abc@example.com

Make Predictions

Enter valid Job ID to fetch results for

Job ID

Fetch Results

Made by Team 1.

Webserver

File uploaded. Results will be available shortly. Job ID: "95743385"

Made by Team 1.

Webserver

Fastp

Started

View Results

Spades

Pending

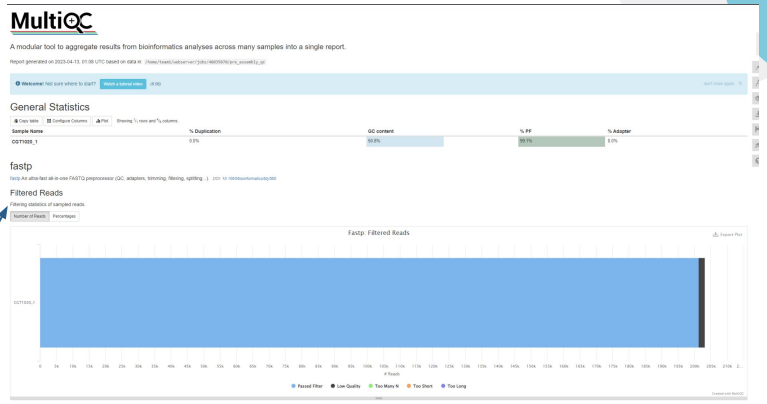
View Results

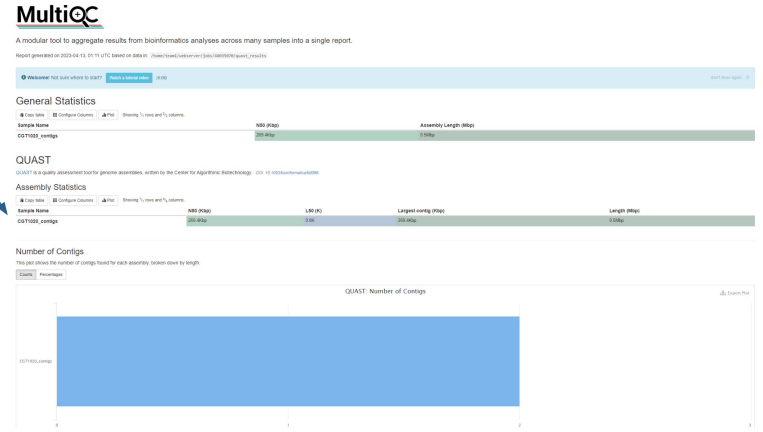
Quast

Pending

View Results

Made by Team 1.





Thank you!
Questions?