

## **Documentation and Design for an Antibiotic Resistance Gene Finder for Mycobacterium tuberculosis**

### Project Background

Antibiotic resistance is a significant public health concern, particularly in the treatment of infectious diseases like tuberculosis. Mycobacterium tuberculosis, the causative agent of tuberculosis also commonly referred to as TB, has developed resistance to various antibiotics, complicating treatment regimens. The ability to identify resistance genes is paramount for effective treatment strategies for patient health, outcomes, and future security through new therapies. The aim of this project is to develop a user-friendly tool that allows researchers and clinicians to quickly determine if specific genes within TB have both known and predicted antibiotic resistance associated with them.

### Tool Functionality

The Antibiotic Resistance Gene Finder will enable users to:

- Input specific gene names or sequences to check against a database of confirmed and predicted antibiotic-resistance genes
- Search using a FASTA sequence to retrieve the top 5 matching genes along with any associated resistance information
- View results indicating whether the queried gene has a matching resistance gene
- View other relevant data? (If possible through time and ability, there is more data on the classes/compounds of antibiotics the specific genes are resistant to as well as external database references e.g. GenBankID that can be included)

### Tool Description

This project will utilize the following key technologies:

#### *MySQL Relational Database:*

Stores gene information, including:

- Gene IDs from NCBI
- Gene sequences from NCBI
- Experimentally confirmed resistance genes from the BacMet Resistance Genes Database
- Predicted resistance genes from the BacMet Resistance Genes Database

#### *Python-based CGI and Sequence Matching:*

Handles the backend logic for:

- Receiving user input and executing database queries
- Performing sequence alignments for FASTA inputs to identify the top 5 matching genes
- Returning results to the user interface in an accessible format

#### HTML/CSS/JavaScript-based GUI:

Provides a user-friendly interface for:

- Inputting gene names or sequences
- Displaying the results clearly, indicating the presence or absence of resistance genes and if searching by sequence, its relevant matches
- Ensuring a smooth user experience through responsive design and interactive elements

### Tool Design and Development

The following steps outline the initial design and development process for the Antibiotic Resistance Gene Finder:

#### *Database Setup:*

1. Create a schema for the database that includes tables for storing gene information and resistance data
2. Have a table storing all of the genes of Mycobacterium tuberculosis from NCBI
3. Have a table storing all of the experimentally confirmed resistance genes from BacMet
4. Have a table storing all of the predicted resistance genes from BacMet
5. Populate the database with additional sources as needed

#### *User Interface Development:*

1. Develop a GUI using HTML, CSS, and JavaScript to facilitate user inputs and present results
2. Ensure the interface is intuitive, with clear instructions for users on where to enter gene names or FASTA sequences (and probably have separate search boxes for the different searches)
3. Make it look semi-professional/on par with other bioinformatic tools

#### *Backend Implementation:*

1. Write a Python script to parse user inputs and query the database
2. Implement functionality to handle both direct gene queries and FASTA sequence submissions, including sequence alignment and match scoring

#### *Validation and Testing:*

1. Confirm the HTML-friendly CGI script through the server and proper presentation on a browser
2. Validate that the database queries and sequence alignments return expected results, making adjustments as needed

**References:**

NCBI. (2015). Mycobacterium tuberculosis H37Rv complete genome. *NCBI Nucleotide*.

<https://www.ncbi.nlm.nih.gov/nuccore/AL123456.3>

BacMet. (2018). *BacMet: Antibacterial Biocide and Metal Resistance Genes Database*.

Biomedicine.gu.se. <http://bacmet.biomedicine.gu.se/index.html>