**Advanced Practical Computer Concepts for Bioinformatics**

**Final Project Narrative**

**Bacterial gRNA Design Tool**

**Project Background**

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is a powerful gene-editing tool that has revolutionized genetic engineering. CRISPR uses guide RNAs (gRNAs) for guiding the Cas9 enzyme to specific locations in the genome where it can generate specific insertions or cuts. Effective gRNA design is a crucial step in the CRISPR gene-editing process. To drive the CRISPR effector (Cas) protein to its target, all CRISPR-Cas applications employ sgRNA. While there are existing gRNA design tools(5,6) such as [CRISPRscan](https://www.crisprscan.org/gene/), [CHOPCHOP](https://chopchop.cbu.uib.no/), and [Cas-offinder](http://www.rgenome.net/cas-offinder/), there is a need for a new tool that enhances gRNA target site specificity, customization, and visualization. In this project, I built user-friendly application, Bacterial gRNA design tool (for bacterial genome), that can assist users in designing gRNA using specific customization such as (PAM sequence, GC content, gRNA length), and provide a scatter plot to visualize the generated gRNAs with respect to the GC content.

**Revised Proposal**

My original project proposal was to build a user-friendly and highly efficient gRNA tool with customization and improved visualization. While working on the project, I realized that building a tool with more efficiency towards off-target sites requires more time, data, and machine learning models that are out of my scope at this time. Hence, in the revised version, I created a gRNA design tool for bacterial genome, where parameters such as GC content, PAM sequences, and gRNA length will be utilized to look for gRNAs corresponding to the sequence of interest.

**Project Technologies and Methods**

I focused on using most of the topics that were covered in the class for the project, such as UNIX OS and filesystem, Python CGI programming, CGI and HTML templating, HTML5 document markup, and page styling with CSS, relational database schemas and design, MySQL, python module to connect MySQL, Javascript and JQuery client-side interactions.

The gRNA design tool generates and stores gRNAs with PAM sequences, gRNA length, and GC\_content in a table within MySQL database. The mysql.connector Python module is used to connect to MySQL and query the database.

Python CGI programming was used to create the server-side scripts to process the data and user interactions. The application has the following modules: to define the Cas9 sites, identify gRNAs, store gRNAs, and retrieving the gRNAs, and module to generate a scatter plot.

The frontend of the tool was developed by using HTML, CSS, Javascript. JQuery was used to enable user-friendly interface for users to interact with the data where I leveraged AJAX and JQuery UI selectmenu and tooltip widget for PAM sequence selection. The HTML template was used to generate dynamic HTML pages to display the gRNA design results.

The scatter can provide an overview of gRNA distribution over different GC contents, which could be helpful in identifying any possible biases toward specific GC content ranges. This data can help in the selection of gRNAs for certain applications, such as targeting regions with a specific GC content range.

**Summary**

This "end-to-end" project was a valuable learning experience that has provided me with a solid foundation for expanding my skillset in the future. Throughout the project, I noticed potential limitations of the project that require more skillsets, but the exposure to all the necessary elements helped me feel confident about the steps required to make the project work. Even though the end-product was not very complex, it was still very exciting to figure out new problems and finally see the tool work as intended.

A screenshot of a computer

Description automatically generated with medium confidence