BME 160
Professor David Bernick
By: Ziad Kedkad, Qudsi Aljabiri

# Restriction Endonuclease DNA Sequence Locator Write-Up

Restriction enzymes, also known as restriction endonucleases, are enzymes that cleave DNA at specific recognition sites. These enzymes are widely used in molecular biology for a variety of applications, including DNA sequencing, gene mapping, and recombinant DNA technology. Restriction enzymes have revolutionized the field of molecular biology by providing a powerful tool for the manipulation and analysis of DNA. In this review, we will discuss the applications and importance of restriction enzymes and the program we built to help locate these enzymes within a DNA sequence. We will also highlight the advantages of using our development of a novel restriction enzyme locator and its usage for potential future research. Comprehending the properties and functions of restriction enzymes is crucial for molecular biology researchers, and utilizing this program can help to save a significant amount of time and money, ultimately leading to important implications for biotechnology and medicine.

If a DNA sequence is known, scientists commonly use restriction enzymes to help verify, implement, or remove genes within a DNA sequence. They do this using restriction enzymes. Restriction enzymes recognize and bind to specific sequences of DNA known as restriction sites, and then cut the DNA at those sites. By choosing different restriction enzymes with different recognition sites, scientists can cut DNA in a controlled and precise manner, generating DNA fragments of specific sizes, shapes, and even further verifying the genetic makeup of that specimen. Due to how convenient these enzymes are, scientists frequently need a quick way and cost effective manner of finding what restriction enzymes they can use on a specific given DNA sequence. With a large database of restriction enzymes, it makes searching for types of restriction enzymes and restriction sites tedious. Developing an intuitive program that portrays this information in a far more digestible manner can save hours of research. Currently available restriction enzyme finders typically require prior knowledge to utilize powerful applications like SnapGene viewer, or navigating through unintuitive and difficult-to-digest website data pages.

The solution we have developed is a program that solves this issue of digestibility. With a known DNA sequence of interest, a scientist can retrieve an ordered and organized text file of every restriction enzyme found within that sequence. In addition, the text file produced has been organized to show the least to the most number of restriction enzyme hits. This allows the scientists to find enzymes that most likely bind in good positions with the least amount of nonspecific cuts. The database that is used comes from the largest and most comprehensive database on restriction enzymes to date (1). With more than 79,000 restriction enzymes, we have

personally organized the list to include 623 restriction enzymes that are commercially available to purchase from a source around the world. This enables our users to quickly identify enzymes that are available for purchase, thus minimizing the time spent searching for enzymes that are not obtainable. Finally, once an enzyme has been selected, we provide our clients with a website that lists all suppliers currently offering that enzyme for purchase.

The aim of this program is to enhance the accessibility and efficiency of existing databases containing information on restriction enzymes. By making information more accessible, time can be spent on performing methods rather than validating these methods. Further, current programs that have been found have outdated databases with unintuitive and lacking program interface. We've been able to patch these issues and further excel in creating a program with a scientist's experience and point of view. In addition, if more time was allotted, being able to directly connect the supplier information to the text file would've been added. Furthermore, we would've wanted to develop an intuitive and visually appealing front-end user interface that not only enhances the user experience but also attracts more scientists to use the browser. This traffic can be put through a paywall which eventually funds development of a browser with a plethora of other DNA sequence analysis tools. The biggest weakness of our program is that the output may be unattractive and enjoyable for the user to use as it is a .txt file rather than a decorated PDF or webpage interface.

Ultimately, we stand by our program and think that this has helped patch up and revolutionize the demand for intuitive and time efficient bioinformatic tools within the science world. Finding the right restriction enzymes is an essential step towards reducing the time and cost associated with scientific research. Utilizing the knowledge we had gained throughout this class, we utilized classes, for, if, and while statements. We also used dictionaries, lists, and leveraged previous assignments to help organize, name, docstring/comment, and build the entirety of the backend program.

References

1. Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D. (2009). Rebase—a database for DNA restriction and modification: Enzymes, genes and genomes. Nucleic Acids Research, 38(suppl_1). https://doi.org/10.1093/nar/gkp874