NucleoDepot: A Novel Approach to Genetic Depositories for the Novice Researcher

Dani, Raunak (School: West Lafayette Junior/Senior High School)
Slamovich, Aaron (School: West Lafayette Junior/Senior High School)

Current genetic databases provide a wealth of information accessible to the general population, but are primarily designed for usage by professionals already experienced in the fields of genetics and bioinformatics. Due to the great variety and specificity of these depositories, it can be difficult for novice researchers to find the tools necessary to access certain genetic information, as many simply do not know where to look. We sought to remedy this by creating NucleoDepot, a database focusing on breadth and variety of information rather than depth and specificity. By doing this, we aim to provide a more generalized “stepping stone” database from which users may begin research before determining which specific datasets will need to be collected elsewhere. Using input from fellow students and advice from college professors, we developed a list of commonly used types of information to include in our database. We then extracted the data from a variety of more specialized depositories and tools, including the GDC, GEO, GEO2R, Cytoscape, DGIdb, String-DB, COSMIC, and cBioPortal. Once all data had been extracted, we wrote programs using Ruby to load the data into the database, as well as scripts to serve queries for information. To help users identify trends in data, we added functionality using R to generate data visualizations. As a result of the described effort, we have created a database with a rich set of data on over 20,000 genes across 28 types of cancer.

Awards Won: