Andy Halstead, Undecided

University: Northwest Missouri State University
Year in School: Sophomore
Hometown: Camden Point, MO
Faculty Mentor: Dr. Chi-Ren Shyu, Computer Science
Funding Source: Missouri Academy

Efficient organization of genetic data for easier statistical analysis

Andy Halstead, Jason Green, and Chi-ren Shyu

When biological experiments are run, specifically those examining genetics, huge amounts of data are produced. This data is hard to organize and even harder to analyze without organization. When faced with a vast amount of data, statistical patterns are hard to observe as associations between different types of data are nearly impossible to see. A relational database was designed to handle this problem. The database organizes data taken by large-scale, long-term genetic experiments into linked categories, and uses an interface system such that makes many approaches to analyze the data possible. The database is unique from other types of organization in that it can be readily applied to these different approaches, while also dealing with subject-specific phenotypic information. This information is often left out in other database structures concerning genetic experiments. To show an application of the database, the structure is being applied to a long-term Maize genetics experiment. Using PHP scripting to insert data taken in the field into a MySQL system, the database is being used to create a search engine for plant geneticists to use to get genotypical information from phenotypical observation. However, the structure is not limited to plant genetics nor search engines, but can be applied to a multitude of bioinformatic and statistical studies of experiment data.