Adv. Practical Comp. Concept for Bioinformatics Final Project Documentation – Sebastian Kyllmann

**SmellBase: Design and Development of a Web Tool to Visualize Drosophila Olfactory Genes**

Tool Background

I am currently gearing up to begin a PhD in Biological Sciences at the University of Nebraska Lincoln. The potential thesis project I might undertake would use machine learning software to analyze drosophila body patterns from video data. These behavioral data would then be related to protein function, with an emphasis on olfactory proteins. I have chosen to use Drosophila data from FlyBase for this final project so that I may familiarize myself with FlyBase, with olfactory genes, and with the bioinformatics tools we have been studying this semester.

A dataset of approximately 128 olfactory genes from D. Melanogaster was selected from FlyBase. The GO ontology “sensory perception of smell”, with accession number 0007608 was chosen to obtain the results. These data will provide the foundation for the final project. Additional data from other Drosophila species may be included as cross references, pulled from NCBI databases.

Tool Description

Data may be configured into an SQL table to host on the class server. When users choose a simplified data source, this SQL table will be used. The table will only contain the FlyBaseID, FeatureType, and Description of the gene. When users want more thorough data, the web tool will access the data using the FlyBase API. Given a list of olfactory genes and their FlyBaseID, the data may be fetched in XML form from a URL such as: <https://api.flybase.org/api/v1.0/chadoxml/FBgn0004514>.

A python CGI script file will handle the backend of the web application. The script will query either the SQL table hosted on the class server or the FlyBase API. The CGI script will then connect with the frontend parts of the web application. These front end parts will include an HTML template, a CSS file used to improve the design of the HTML template, and a JavaScript file to provide further interactive components.

When using simplified data, the web application would display one random olfactory gene at a time. The presentation of the gene would emulate a visually appealing index card. The frontend programming would then allow “swipe” motion or other means for the user to generate new index cards. This web tool would seek to emulate the feeling of flipping through a book to get an overview of olfactory genes. A button feature would allow switching to the more advanced features which would add search functionality and include more data fields for each gene entry. The image below displays a rapid prototype of the frontend GUI.

A picture containing graphical user interface

Description automatically generated