Jerald Macachor  
AS.410.712.81

Advanced Practical Computer Concepts for Bioinformatics

Professor Joshua Orvis  
7/31/2017

Project Proposal.

Bioinformatics tool using MUSCLE – Multiple Sequence Alignment. <http://www.ebi.ac.uk/Tools/msa/muscle/>

I. How this project works.

1. User inputs a sequence of protein sequences in Fasta format into a form.  
2. User presses a button to make a MUSCLE alignment, they will receive a CLUSTAL W format Multiple sequence alignment. The alignment will be color coordinated where sequences match.  
For example.

|  |  |
| --- | --- |
|  | All perfect alignments will be the same color background. Those with 1 misalignment will be same color leaving out the misaligned. |

3. Sequences are automatically saved in a database unless they are completely redundant.  
The Fields will be Accession number, Name, Sequence, Date Saved.

4. To access protein sequences saved. There will be a Search Textbox with autocomplete. When the User starts typing, they can see a list of potential entries. Once the “search” textbox is clicked, if the input entry is valid, a form underneath the search will show the Fasta sequence of that search.

5. There will be an add button that adds the content of their searched protein onto the bottom of the MUSCLE form.

6. They can also save Protein Sequences on the MUSCLE form by clicking on save.

II. Implementation plan.

Create an HTML page with a form for the input of the MUSCLE sequence. Also a form for searching protein sequences within a MySQL database. The Submit button when clicked activates a CGI script that takes the input from the form and run it through the MUSCLE program. The muscle program will have an output file with Clustalw format. This will be parsed in such a way to make it readable in the html template generated. Javascript will be applied to change the background colors as done in the example above. When the submit button is clicked, a MySQL database will save the sequences if they are not redundant (same accession number). If redundant, then the old entry will be overiden. If the user simply wants to save the sequences into the database, there will be a save button for that.