User Manual & Support Documentation

# Files for HTML and CSS for Summary and Detail Page From Summary Results.

The database layer is made up of the following files / source code:

1. createdb.sql – this is a sql script that creates the tables for the database
2. indexes.sql – this is a script that creates the indexes for each tables
3. GenBankParser.pm – this is the main perl code which parses the genbank file
4. main.pl – this is the main program which calls the GenBankParser package and then inserts all the database into the database tables.

# Webpage Functions

# Summary Page

## Tab on Menu Bar

By Clicking the summary tab on the menu bar the CGI script get Summary will run this will populate a table which will contain the gene identifiers, protein product names, Genbank accession and chromosomal location.

# Detail Page

## Page layout and detail boxes

By clicking on a genebank accession number this will open up the detail page. The detail page has four boxes with each individual boxes contain the following information;

1. Box 1 contains the complete DNA sequence with the coding regions highlighted
2. Box 2 contains the amino acid sequence displayed with the coding DNA sequence
3. Box 3 contains codon usage frequencies within the coding region and the overall codon usage information for all sequences in the database.
4. Box 4 contains the ability to identify sticky-end restriction enzyme sites in the genomic DNA with the ability to search for EcoRI, BamHI and BsuMI.

## Detail Boxes

Each box offers a quick preview of the information inside them which can be viewed by scrolling up and down. By clicking on the plus button (+) next to the ID name e.g. Complete Sequence this will maximise the page making it easier to view the information.

#### Box 1 Complete Sequence Results

This box displays the complete DNA sequence for the gene selected in the summary page with the coding region highlighted in red

#### Box 2 Full Amino Acid Sequence

This Box displays the full amino acid sequence for the gene selected is the summary page. Below this it displays the coding sequence which will code for protein.

#### Box 3 Codon Usage

Box 3 displays the codon usage as a percentage from the coding region for the sequence selected. It also displays the codon usage as a percentage from the coding region for all the sequence in database. This allows the user to compare how frequent a given codon is in comparison to other sequences in the database. The results are displayed in a tabular format with the three nucleotides (triplets)and the specific amino acid residue or stop codon these triplets encode displayed.

#### Box 4 Sticky End Display

Box 4 allows the user to identify sticky-end restriction enzyme sites in the genomic DNA by searching using a unique identifier with the ability to search for EcoRI, BamHI and BsuMI. Foor more details please refer to ............Yolanda Part