SNPVariants Program

About

The purpose of this program is to show the changed variants in each fasta sequence. The program manipulates the sequences in the inputted fasta file and changes the base based on the information obtained from a Run Variants excel file such as the run, barcode, chromosome, and position. The final output of the program is several fasta files, one fasta file per chromosome, with different sequences based on the Sample Id and Run.

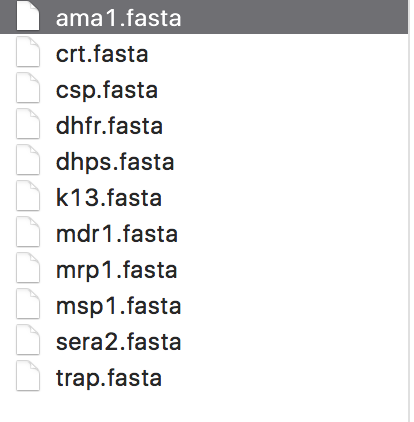
Example Output :

Figure 2: This picture shows the fasta file for chromosome mdr1, and shows how the fasta file is split up. Seen here is the modified sequence for each sample ID and run.

Figure : This picture shows a folder containing all the separate fasta file outputs generated by the program

Inputs and Outputs:

Inputs: A fasta file, an excel file contain all the snp variants, and a specific directory for the fasta files to be inputted

Output: A folder containing fasta files containing the modified sequences

Program Requirements:

Python Installer

A program called xslx2csv

Installation: sudo easy\_install xlsx2csv or pip install xlsx2csv

Questions/Concerns about the program:

* When inputting the directory for either the fasta or the excel file do not put any spaces before or after the input. An example directory would be:
  + /Users/Name/Folder/Folder/.fasta
* How to install xslx2csv?
  + Input in terminal either sudo easy\_install xlsx2csv or pip install xlsx2csv