**VirLab Genome Project Journal**

March 12th, 2020

Today I edited k\_mer\_creator.py so that it would run successfully on my Mac. I clocked the running speed at

real 1m16.639s

user 1m20.279s

sys 0m02.689s

Next I would like to add an empty “results” folder to the github and edit the READ\_ME to reflect that ones working directory should be “src”

Then I would like to edit the randomreads.sh bash file so that the code runs correctly when placed inside a directory that’s name contains spaces. Placing the VirLab dir in a dir named Local Files demonstrated this bug.

After that I hope to make the code more efficient by replacing the parse\_dir function that utilizes os.walk with one that uses pathlib. I also hope that this reduces possible errors from running the code on different operating systems.

March 19th, 2020

At some point in the last few days we ran into an error with DS\_store files. These files exist on a macs to store visual information about things like the dynamic positions of items in directories. When moving around files in the Genomes folders, hidden DS\_store files were inadvertently created. These caused the k\_mer\_creator.py program to throw an error. We temporarily solved the problem by deleting the hidden files, but a more permanent solution should be implemented.

I’ve successfully updated the .gitignore files to allow a practically empty results folder to be part of the github repo. I have also made it so that git ignores os-generated files.

March 21st, 2020

I’ve successfully created a new branch of the git repo with my version of k\_mer\_creator.py

I’ve noticed that both the old and new versions of k\_mer\_creator.py do not create the test\_reads…fastq files on windows systems but do on osx.

Implemented Pathlib so that it

April 7th 2020

After much tinkering, I got analyze\_range and all of the functions it calls to run from a .py file in the VirLab dir.

This included making and importing a function that makes the paths to datasets into just the name of the dataset dir. It takes in str objects or object types used by PathLib and outputs a str object. This function should work across operating systems.

Now I am beginning work on making the randomreads.sh file handle paths that contain spaces.