Lindsay Havens

Final project proposal

In order to further my research, I would like to write a program that carries out a pipeline in an iterative fashion. The purpose of the program will be to use transcriptomic libraries to improve a genome assembly. The program that I plan to write will take a starting assembly (using sys.argv) and will run through bwa index, seqtk, bwa mem, samtools, and besst\_rna. It will be written in a loop fashion so that the output of besst\_rna (the genome assembly) will be used as the input assembly for the following iterations. There has been evidence that conducting this pipeline in an iterative fashion will increase genomic assembly accuracy and lead to an increased NG50.

Alternatively, I could write a program that searches through my assembled genome to identify contigs that contain aphid contamination.