

README for Optional_Velvet_Prep_Wrapper.pl:

README generated by: Rishi R. Masalia, Burke Lab, Univ. of Georgia, August 2013
Corresponding email: jennifer.r.mandel@gmail.com or masalia@uga.edu

Please read the General_Workflow_README file first.

Create a main folder for this analysis to be run in. All raw sequence (fastq) read data must be in this main folder. In this main folder, create a ./programs and ./databases folder. Exactly named. In the ./programs folder put all downloaded scripts and outside programs (except the wrapper). If you are running this wrapper without running the Optional_Blast_Wrapper.pl, you will need to generate a ./Archived_Files and ./Output folder in the main directory. Programs to include are: Pairfq.pl and Shufflesequences.pl for this wrapper, placed in the ./programs folder.

Goal: To take the output from Top_Hits.pl (end of Optional_Blast_Wrapper.pl), and prepare both Taxon reads (R1 and R2) for Velvet Optimiser. Velvet Optimiser is not included in this wrapper since it can be dataset specific, though information on how to run Velvet Optimiser is located in the General Workflow README.

Script Location:

Place this wrapper in the main directory.

Output:

The output of this wrapper is TaxonName.together.paired.fasta and TaxonName.together.singletons.fasta. These files will be located in the ./Output folder, and should be used when running Velvet Optimiser.

Usage:

```
perl Optional_Velvet_Prep_Wrapper.pl <names_file>
```

Names File Generation:

Generation of a “names” file signified by the, the <names_file> argument. NOTE this is a DIFFERENT file than the previous (Optional_Blast_Wrapper.pl names file). As before, this is a separate file with just a list of names; sequence files you wish to analyze –generated from Top_Hits.pl (or from the previous wrapper).

Names file layout: <TaxonName_R1>=<TaxonName_R2>=<TaxonName>. It is absolutely crucial that the R1 and R2 names line up for the same Taxon; otherwise you’ll be generating mismatches. It is also crucial that R2 goes second. The TaxonName is the heading you wish to name the file when both R1 and R2 files are matched together.

Example:

TaxonName_R1=TaxonName_R2=TaxonName

NextTaxonName_R1=NextTaxonName_R2=NextTaxonName

An example of this is provided on the GitHub, under
“Optional_Velvet_Prep_Wrapper_Names_File_Example”

Note: Pairfq and Velvet Shufflesequences are ready-made programs we call in this workflow. As such, any questions regarding their operation or usage should be directed to their manuals, forums and help pages, or follow the necessary channels in contacting the authors of these programs.

Note: All program executables should be placed in the ./programs folder, in the main working directory.

Programs Needed Prior to Running:

- Pairfq.pl (<https://github.com/sestaton/Pairfq>)
- Shufflesequences.pl
(https://github.com/dzerbino/velvet/tree/master/contrib/shuffleSequences_fasta)