BLAST practice

Experiential Data science for Undergraduate Cross-disciplinary Education

Dr. Kim Dill-McFarland, U. of British Columbia

Contents

BLAST practice

These problems are designed to help you practice concepts and functions covered in the 'BLAST tutorial'.

Learning objectives

- Practice Unix command line
- Complete BLAST of a large amplicon sequencing data set via command line

Practice

1. In the BLAST tutorial, we viewed to first 3 'hits' for a subset of the sequence data. Try altering the `blastn` parameters to instead show 5 descriptions and 0 alignments.
2. While the subset data in the tutorial runs very quickly, it is not even close to the full data set! Try running `blastn` in your command line terminal on the full `Saanich_OTU_rep.fasta` file. How long does it take to complete?
    - You can use the `time` function to have your terminal track how long it runs. For example `time blastn -query ...`
3. Now, upload the full `fasta` file to BLAST online and run it there. How long does this take to run?
    - Be sure to change the online tool database to 16S and Program to 'Somewhat similar sequences (blastn)'
