

Basic Genome Analysis \2017

Practical 1 - Comp Gen 2017

Assistants: Daniel Morgan and Mateusz Kaduk

PURPOSE	1
OBJECTIVES	1
ACTIVITY	1
CHECK-POINT	2
Report structure	2

PURPOSE

Learn two popular sequence similarity search methods:

- [BLAST](#)
- HMMER

These are cornerstone sequence analysis tools. Knowing the differences between them will enrich your sequence related skillset. In this exercise their usefulness is demonstrated by genome classification problem.

OBJECTIVES

- Learn and compare algorithms
- Perform exercise understanding general rules of use and applicability
- Write report

ACTIVITY

Perform the following steps in this order

1. Familiarize yourself with BLAST and HMMER algorithms
2. Compare BLAST to HMMER, where do the methods overlap and where are they unique. What are advantages of each?
3. Characterize your genomes with BLAST using NCBI website
<https://blast.ncbi.nlm.nih.gov/>
 - a. What organisms genome belongs to?
 - b. How do the genomes differ (size, kingdom, number of genes)?
 - c. Does anything interesting stand out?
 - d. Why aren't we asking you to run HMMER in parallel?

CHECK-POINT

Steps to check for student understanding

1. What is the difference between Needleman-Wunsch and Smith-Waterman algorithms? Why- what are advantages of each?
2. Is BLAST different from Smith-Waterman algorithm? In what ways?
3. What is the Viterbi algorithm ?
4. How do these methods compare in speed ?

Report structure

The report should be formatted in one PDF file and sent in by the end of the week. It should cover all steps performed and why, answer all check-point and activity questions, include referencing citations and interpreting results. Answers pertinent for follow up practicals will be provided.

Answers need not be exact, but rather should convey that you have completed and understood the assignment. Feel free to expound on problems you encountered, how you overcame them, and anything you found interesting that we have not asked for outright.

1. A short summary of what you have done as well as a short discussion of the methods and parameters chosen.
2. Discuss the questions in the lab specification as well as any questions and thoughts of your own.
3. Discussion of your results.

4. Well commented code (if applicable)

