

Bioinformatics Computing
CSE40532/60532
Homework #5

Problems: (due 11/13)

1. Visit the AMOS website (<http://amos.sourceforge.net/>).
2. Read the main page to get a feel for the project, the players and API, and corresponding modules.
3. Download and install AMOS (#8 on the main page, or click “Download”). For this homework, you do not need to have Qt installed. Minimus will work without it.
4. Read the minimus documentation, and look at the example projects in `amos-2.0.8/test/minimus`. Get a basic idea of the input, what the bank is, and reperform an assembly. Submit contigs for the influenza virus in your dropbox (5 points)
5. Download the test datasets from the course website, and assemble each using minimus. Submit the resulting contigs in your dropbox for the maize data (5 points)
6. Compare your minimus assemblies to the references available from the course webpage using `nucmer`. Place the resulting delta file for the bacteria genome in your dropbox (5 points).
7. Provide a brief summary of the results (how many contigs, average contig size, etc.) for grading and do the same sort of comparisons in Hw #4 (`show-coords`, `show-snps`, etc). Do you think these are good or bad assemblies and why? Discuss some strategies you could do to make each one “better” (hint: talk about sequencing or other mummer utilities)? Submit in your dropbox this summary and any relevant supporting files (10 points)
8. For fun: Install Qt and look at the test assembly based on minimus in Hawkeye.