

Bioinformatics Computing
CSE40532/60532
Homework #4

Problems: (due 11/4)

1. Visit the MUMmer website (<http://mummer.sourceforge.net/>).
2. Read the Online Tutorial. Specifically, focus on the section on “nucmer” and on genome alignments in general.
3. Download and install the MUMmer package from the site above.
4. Download two Anthrax strains from GenBank, the gold standard “Ames ancestor” that is virulent (NC_007530) and the non-virulent lab strain “Ames” (NC_003997).
5. Run nucmer on the two strains using default parameters and with the “Ames ancestor” as the reference. Place the resulting “.delta” file in your drop box (10 points).
6. Summarize the alignments using the “show-coords” utility with the following options: “-r -c -l”. Save the “.coords” file in your drop box (5 points)
7. Summarize SNPs and indels between these two strains using the “show-snps” utility with the “-C” option. Save the “.snps” file in your drop box (5 points)
8. Submit a brief write up in your drop box summarizing the results. Hypothesize what the differences may be between the strains? Are there a lot of SNPs? Potential structural differences? How could this technique be used for other tests related to bacteria or other species? (5 points)