Quantitative Biology / Bioinformatics Homework 2
Fall 2010

Due Date: Email mschatz@cshl.edu by 11:59pm on Nov 10

1. BLAST Sequence Identification

Download the sequence whoami.fa from:
http://schatzlab.cshl.edu/teaching/2010/whoami.fa

Assume the sequence is contig from a metagenomics sequencing project. Identify
the sequenced organism using BLAST: which organism has the most significant
match and what is the e-value? For this task, you should execute a nucleotide BLAST
(blastn) using the Nucleotide Collection (nr) as your database on the NCBI website:
http://blast.ncbi.nlm.nih.gov/Blast.cgi

Hint 1: whoami has an *extremely* significant nucleotide match, but you will need
to increase the sensitivity of the search – review the slides

After you find the correct organism, inspect the alignment and report why the
default parameters don't detect the otherwise very significant alignment