

Homework 3: due Monday, Nov. 4th, 6:00PM

In this assignment set up a BLAST search using a locally installed BLAST version. Using the sequences in human.fa find the most homologous sequences to each human sequence in the mouse data base mouse.fasta. You need to:

1. set up a Python code that parses the human sequences,
2. iterates through each sequence,
3. searches the mouse database with BLAST and
4. prints the human sequence ID, the mouse ID of the most similar homolog, the corresponding alignment, E-value and bitscore to a file of your choice.

Furthermore, you need to explain the choice of the (i) blast program, (ii) substitution matrix and (iii) choice of parameters.