

Algorithms in Bioinformatics 2, SoSe2007

Assignment sheet # 1

Tobias Kloeppe

April 16, 2007

This weeks assignment is to generate two profile HMMs and to search through an online database with your profiles. Finally you are asked to analyzes your results as well as the profiles. For this assignment please download the provided files located at <http://www-ab.informatik.uni-tuebingen.de/teaching/ss07/albi2/assign/files/assign01.zip>. All assignments can either be performed using a local installation or a web service of your choice.

1 Generate profile HMMs from domain alignments (2 points)

Extract the two files in the zip archive and generate two hmm profiles from these alignments using HMMER *hmmbuild*. After you generated your models calibrate both of them using HMMER *hmmcalibrate*. Why is it important to use *hmmcalibrate*?

2 Search the *uniprot* database (2 points)

Search through the *uniprot* database with your generated models.

3 Analyze your results (4 points)

- Report a list of the accession numbers found by each model.
- Draw a graph that shows the total number of proteins found by each model and the number of sequences that where identified by both of them in dependence to their best e-Value.
- What kind of sequences where found and how are they related among each other?

4 Description of the profiles (2 points)

Use the software *LogoMat-M* at <http://www.sanger.ac.uk/Software/analysis/logomat-m> to visualize the profiles and give a detailed description of both HMMs. Are there positions specific to each of the models?

Assignments due: **Monday, April 23, 10h**