



Einführung in die Bioinformatik

Prof. Dr. Enno Ohlebusch, Dr. Karlheinz Holzmann,
Tobias Badura

WS 15/16

Übungsblatt 4

Abgabe bis 13.01.2016, 10:00. Lösungen bitte elektronisch (an tobias.badura@uni-ulm.de) mit Namen in Text und Quelltextdateien abgeben.

1. Aufgabe (10): Progressives Multiples Sequenzalignment

Berechnen Sie ein progressives multiples Alignment nach der Feng-Doolittle Methode der folgenden 4 Aminosäuresequenzen (benutzen Sie die PAM250 Matrix und -8 als Gap-Penalty):

NFS
NFLS
NYLS
NKYLS

Legen Sie ihrem Vorgehen den Guide Tree aus dem Skript zugrunde. Geben Sie die Dynamic Programming Matrizen und Scores aller benötigten paarweisen Alignments an.

2. Aufgabe (5)

Laden Sie die Datei `sequenz.fas` von der Webseite der Veranstaltung herunter. Ermitteln Sie die Funktion dieses Gens. Suchen Sie dazu mit dem `blastn` Programm von der Webseite <http://blast.ncbi.nlm.nih.gov/Blast.cgi> nach ähnlichen Sequenzen.

3. Aufgabe (5): UPGMA

Gegeben ist eine Matrix mit den paarweisen Distanzen von 4 Organismen. Bestimmen Sie den Baum, der sich mit dem UPGMA-Verfahren ergibt.

D	a	b	c	d
a	0			
b	3	0		
c	6	7	0	
d	5	6	3	0

- Welche paarweisen Distanzen ergeben sich durch die Kantenlängen?
- Zeichnen Sie einen Baum der die paarweisen Distanzen korrekt widerspiegelt.