

Übungen zur Vorlesung Sequenzanalyse II

Universität Bielefeld, SoSe 2008

Prof. Dr. Jens Stoye · Dipl.-Inform. Peter Husemann

<http://gi.cebitec.uni-bielefeld.de/teaching/2008summer/sequenzanalyse>

Blatt 10 vom 04.07.2008

Abgabe in einer Woche zu Beginn der Vorlesung.

Aufgabe 1 Berechne den generalisierten Suffixbaum der folgenden Sequenzen:

$$s_1 = \text{TATAGA}, s_2 = \text{TAGAGA} \text{ und } s_3 = \text{CAGATAG}$$

- Finde mit Hilfe des generalisierten Suffixbaums alle *maximal unique matches* (MUMs) und gib sie an.
- Ermittle alle *maximal multiple exact matches* (*multi-MEMs*) und gib sie an.

Aufgabe 2 Überlege Dir eine sinnvolle Scorefunktion für das Chaining-Problem beim multiplen Genomalignment.