

Vorlesung: Spezielle Algorithmen auf Sequenzen
Sommersemester 2005

Übungen

Übung 6, Diskussion: 16.06.2005.

1. Baumalignment.

Die Benutzung der **PAM250** Ähnlichkeitsmatrix ist nicht für alle Schritte eines Alignments entlang eines Baumes empfehlenswert. Warum?

2. Carrillo-Lipman Heuristik.

- (a) Charakterisieren Sie Sequenzen, für die die Carrillo-Lipman Heuristik schlecht bzw. gut funktioniert.
- (b) Wie viele Carrillo-Lipman Schranken $L_{x,y}$ werden für k Sequenzen berechnet?
- (c) Durchsuchen Sie das Internet für eine herunterladbare MSA-Implementation (das NCBI hat eine, MSA 2.1). Laden Sie die Software herunter und installieren Sie diese. Berechnen Sie ein multiples Alignment für die Sequenzen auf der Webseite mit
 - i. *default* Optionen und
 - ii. *fast heuristic alignment*.

Vergleichen Sie die Laufzeiten, den benutzten Speicherplatz und die Alignmentqualität. Seien Sie in der Lage, ihre Ergebnisse in den Übungen zu diskutieren.

3. Multiples Sequenzalignment mit affinen Gapkosten.

Welche Probleme treten auf, wenn Sie versuchen, den *Sum-of-Pairs* Score bei affinen Gapkosten zu berechnen?

Wie löst Altschul dieses Problem?

4. Center-Star Methode.

Was ist die Zeitkomplexität der Center-Star Methode zur Berechnung einer 2-Approximation des Sum-of-Pairs multiplen Sequenzalignmentproblems?