

Übungen zum Sequenzanalyse-Praktikum

Universität Bielefeld, SoSe 2015

Dr. Roland Wittler · M.Sc. Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2015summer/SequaPrak>

praktikum-seqan@CeBiTec.Uni-Bielefeld.DE

Übungsblatt 10 vom 22.06.2015

Abgabe bis Donnerstag, 24:00 Uhr.

Benutze für diesen Aufgabebettel die *T-Coffee*-Versionen von der Seite www.tcoffee.org.

Aufgabe 1 (MSA mit T-Coffee und Clustal W2)

1. Berechne mit den Hämoglobinsequenzen von letzter Woche, die du nun auf der Veranstaltungshomepage findest, ein multiples Alignment mit *T-Coffee* mit Standardeinstellungen. Berechne auch mit *Clustal W2* mit Standardeinstellungen ein multiples Alignment mit diesen Sequenzen.
2. Vergleiche das *T-Coffee*-Layout mit dem von *Clustal W2*. Was ist gleich, was ist anders? Erkläre, was die Unterschiede bedeuten.
3. *Clustal W2* berechnet aus einer paarweisen Distanzmatrix einen *Guide Tree* und baut ein multiples Alignment auf, indem es die Sequenzen progressiv dem *Guide Tree* folgend aligniert. Dabei wird ein zuvor festgelegtes Scoring-Schema benutzt. *T-Coffee* benutzt auch einen *Guide Tree*, besitzt für jedes zu alignierende Basenpaar jedoch einen spezifischen Score, der aus Konsistenzberechnungen in einer früheren Phase resultiert. So werden bei der Alignmentberechnung homologe Bereiche schneller gefunden und nicht so schnell Fehler gemacht.
Wie erklärst du dir aber nun, dass *Clustal W2* das gleiche Alignment wie *T-Coffee* gefunden hat?

Aufgabe 2 (T-Coffee Flavors)

Der Algorithmus von *T-Coffee* wurde dir im Seminar vorgestellt. Es gibt auch noch einige Erweiterungen, zum Beispiel *M-Coffee*. Dieser benutzt gleich zu Beginn mehrere MSA-Algorithmen, um ein besseres multiples Alignment zu erhalten.

1. Benutze das Programm *T-Coffee* mit den Beispielsequenzen auf der Eingabeseite und Standardeinstellungen. Wie hoch ist der Score des Alignments?
2. Benutze nun *M-Coffee* mit den gleichen Sequenzen und ebenfalls Standardeinstellungen. Wie hoch ist der Score hier? Welches der beiden Alignments ist besser? Kannst du das auch an der Farbdarstellung erkennen?
3. Welche multiplen Alignmentmethoden benutzt *M-Coffee* standardmäßig?
4. Schau nach, ob es diese oder ähnliche Methoden auch bei *T-Coffee* gibt. Wenn ja, wähle sie alle aus. Wie sieht nun das Ergebnis aus, das du erhältst?

Aufgabe 3 (T-Coffee auf der Kommandozeile)

1. Lade dir über www.tcoffee.org *T-Coffee* für dein System herunter und installiere es. Führe *T-Coffee* mit dem einfachsten Aufruf für die Hämoglobinsequenzen aus und gib diesen auch in deinem Protokoll an.
2. Vergleiche das entstandene Alignment auf der Konsole mit dem aus Aufgabe 1. Wie unterscheidet sich die Darstellung? Ist das Alignment identisch?
3. Ein Teil der Ausgabe beschreibt den *Guide Tree*. Stelle ihn in deinem Protokoll dar. (Zeichne ihn entweder per Hand oder am Computer.)
4. Finde mithilfe der technischen Dokumentation heraus, was der Parameter *-matrix* bedeutet und mit welchem Standardwert er ausgeführt wird. Bei welchem Schritt von *T-Coffee* findet er Anwendung?
5. Nenne einen Vorteil, den die Benutzung des Kommandozeilentools gegenüber der Onlineversion bietet. Was ist ein Nachteil? (ca. vier Sätze)