

Übungen zur Phylogenetik Vorlesung

Universität Bielefeld, WS 2013/2014, Dr. Roland Wittler, M.Sc. Kai Stadermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2013winter/Phylogenetik>

Blatt 12 vom 20.01.2014

Abgabe in einer Woche zu Beginn der Vorlesung oder vorab bei der Tutorin oder beim Veranstalter.

Aufgabe 1 Jukes-Cantor Korrektur.

(5 Punkte)

In einem Alignment der Länge 600 finden sich 40 Substitutionen (und keine Indels). Wir nehmen eine Entwicklung der Sequenzen nach dem Jukes-Cantor Modell an.

- Wie groß ist die Jukes-Cantor korrigierte Distanz d (in PEM) der beiden Sequenzen?
- Wie viele Mutationsereignisse haben demnach in etwa stattgefunden?
- Wie viele Mutationsereignisse und wie viele Substitutionen im Alignment würden sich ergeben, wenn die Sequenzen einen Abstand d von 20 PEM hätten?

Aufgabe 2 Modellierung von Aminosäure-Austausch.

(6 Punkte)

Zur Vereinfachung nehmen wir im Folgenden das Aminosäure-Alphabet $\mathcal{A} = \{A, B, C\}$ an. Betrachte das folgende Alignment der Sequenzen X und Y:

```
X = B C C B A A A B C B A C C B A B A C A C
Y = B C A C A A A B C B A C C A A B A C A A
```

- Gib die Werte für alle m_{ij} , alle f_i und N an und berechne daraus die Übergangsmatrix P .
- Ist P auf 1 PAM kalibriert? Begründe.
- Berechne aus P die Score-Matrix S (mit $\pi_i = f_i$). Wenn du keine Werte aus (a) hast, verwende

$$P = \begin{pmatrix} \frac{10}{13} & \frac{1}{13} & \frac{2}{13} \\ \frac{1}{10} & \frac{4}{5} & \frac{1}{10} \\ \frac{2}{17} & \frac{1}{17} & \frac{14}{17} \end{pmatrix} \text{ und } \pi = \left(\frac{13}{40}, \frac{1}{4}, \frac{17}{40} \right).$$

Aufgabe 3 Maximum-Likelihood Abschätzung von Distanzen.

(4 Punkte)

Nimm an, die Entwicklung der Sequenzen A und B aus Aufgabe 2 unterlag einem EMP mit der folgenden Übergangsmatrix:

$$P = \begin{pmatrix} 1 - \frac{2}{3}t & \frac{1}{3}t & \frac{1}{3}t \\ \frac{1}{3}t & 1 - \frac{2}{3}t & \frac{1}{3}t \\ \frac{1}{3}t & \frac{1}{3}t & 1 - \frac{2}{3}t \end{pmatrix}$$

Bestimme die Zeit t für die die Likelihood $Pr(A, B; t)$ maximal ist.