

Aufgabenblatt 11

PAM&BLOSUM; BLAST

Abgabe: **25.01.2005** bis **17 Uhr** über Goya

Maximal: 20 Punkte

Namen nicht vergessen!

1. (Wie) ist es möglich, dass zwei Sequenzen einen Abstand von mehr als 100 PAM besitzen? **2 P**
2. Was bedeuten m und n in PAM m und BLOSUM n im Hinblick auf den evolutionären Abstand zweier Sequenzen? **1 P**
3. Substitutionsmatrizen können nicht nur auf Mutationsraten beruhen (wie PAM und BLOSUM). Nennt mindestens drei weitere vorstellbare Möglichkeiten für Bewertungen! **3 P**
4. Wir erinnern uns an das Aufgabenblatt 3. Nehmen wir die Sequenz von Aufgabe 3.1:

`tctcacagcaaattaggcaggagtgtctca`

 (Länge: 30; Organismus: Elefant *Loxodonta africana*; E-Value: 1e-07). Wieviele Nukleotide müssen wir (von hinten) entfernen, damit in BLAST
 - (a) die Zuordnung nicht mehr eindeutig ist (das bedeutet, die ersten Positionen bekommen (nahezu) identische Werte? **2 P**
 - (b) keine signifikanten Ähnlichkeiten mehr gefunden werden? **2 P**
 (BLAST findet Ihr im Internet auf
<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi?PAGE=Nucleotides>)
5. (a) In der letzten Aufgabe habt Ihr ein Programm für das Alignment mit globalen Mismatch/Replacementkosten entwickelt. In dieser Aufgabe wollen wir stattdessen eine Substitutionsmatrix verwenden. Erweitert Euer Programm entsprechend! Als Beispiel findet Ihr auf der Webseite die BLOSUM50 und zwei zu alignierende Stringpaare.

Programmaufruf:

```
java GruppeXAufgabe11 -A datei1 -B datei2 -X matrix
```

 (*datei1* und *datei2* enthalten die beiden zu alignierenden Strings, *matrix* die Substitutionsmatrix)
 Das Programm soll die Tabelle und das Alignment ausgeben, die Ausgabe kopiert Ihr bitte zusätzlich auf das Lösungsblatt. **6 P**
- (b) Nennt jeweils das optimale Alignment und dessen Score! **2×2=4 P**