

3. Übungsblatt “Einführung in die Bioinformatik” SS 2008

Prof. Dr. Daniel Huson und Juliane D. Klein

25. April 2008

1 PAM und BLOSUM Matrix (4 Punkte)

1. Was ist eine PAM Matrix und wie wird sie berechnet?
2. Die BLOSUM Matrix wird im Unterschied zur PAM Matrix aus lokalen statt globalen Alignments erstellt. Auf welche Regionen der Proteinsequenz beschränkt sich die BLOSUM Matrix dabei? Welche anderen Regionen betrachtet die PAM zusätzlich und welcher Nachteil könnten daraus entstehen?
3. Welche Eigenschaft haben die Sequenzen, die für eine BLOSUM Matrix mit einer kleineren Zahl, verwendet wurden? Was für eine PAM Matrix würden Sie dagegen wählen, wenn Sie in nah verwandten Sequenzen suchen wollen?
4. Was ist der grundlegende Unterschied bei der Berechnung der verschiedenen BLOSUM und PAM Matrizen (mit unterschiedlicher Nummerierung)?

2 Lokales Alignment (2 Punkte)

Welche drei Schritte müsste man im Needleman-Wunsch Algorithmus ändern, damit ein lokales statt einem globalen Alignment gefunden wird?

3 Smith-Waterman Algorithmus (4 Punkte)

Finden Sie für die beiden Protein-Sequenzen ISALIGNED und THISLINE ein optimales lokales Alignment. Füllen Sie dafür folgend dem Smith-Waterman Algorithmus eine Matrix aus. Verwenden Sie bitte die BLOSUM-62 als Substitutionsmatrix und eine Gap-Penalty von $d = 6$. Geben Sie dann ein optimales Alignment an, was Sie durch Traceback gefunden haben. Markieren Sie es bitte auch in der Matrix. Wenn die verkürzten Sequenzen ISALIGN und THISLIN gegeben sind, wieviele optimale lokale Alignments gibt es dann?