

2. Übung zur Vorlesung “Einführung in die Bioinformatik I, 2. Teil”

Sommersemester 2021

Prof. Peter Dittrich, Emanuel Barth, Marcus Ludwig

Ausgabe: 22. April 2021,
Abgabe: 29. April 2021 zu Beginn der Übung

Aufgabe 1 (3 Punkte): Gegeben seien die Sequenzen $u = \text{GATCAGGT}$ und $v = \text{GACC}$. Bestimmen Sie die Matrix D und alle optimalen **globalen** Alignments mit Einheitskosten.

Aufgabe 2 (6 Punkte): Gegeben seien die beiden Sequenzen aus Aufgabe 1.

Wir definieren die Ähnlichkeitsfunktion S mit $S(a, a) = 1$, $S(a, b) = -1$ für $a \neq b$ und $S(a, -) = S(-, b) = -1$. Bestimmen Sie die Matrix H für ein **lokales** Alignment von u, v . Wie lauten die optimalen Alignments (Traceback in die Matrix einzeichnen), wie ihre Ähnlichkeit?

Aufgabe 3 (6 Punkte): Gegeben seien wiederum die Sequenzen aus Aufgabe 1. Wir definieren die Ähnlichkeitsfunktion $S' : \{A, C, G, T\} \times \{A, C, G, T\} \rightarrow \mathbb{R}$ mit

- $S'(A, A) = 2$, $S'(C, C) = S'(G, G) = S'(T, T) = 1$,
- $S'(C, G) = S'(G, C) = 1$, $S'(a, b) = -1$ für alle anderen $a \neq b$, und
- $S'(a, -) = S'(-, b) = -1$.

Bestimmen Sie die Matrix H' für ein **lokales** Alignment von u, v . Was sind diesmal die optimalen Alignments (Traceback in die Matrix einzeichnen) und deren Ähnlichkeit?

Aufgabe 4 (5 Punkte): Gegeben sei die Sequenz $u = \text{CCAT}$ sowie das Profil:

Pos.	1	2	3	4
A	0,2	0	0,7	0
C	0,6	0,4	0,2	0
G	0	0,5	0	0
T	0,1	0	0,1	1
-	0,1	0,1	0	0

Bestimmen Sie die Ähnlichkeit des Profil-Sequenz-Alignments für die Ähnlichkeitsfunktion S' aus Aufgabe 3, wenn keine zusätzlichen Gaps in u eingefügt werden.