Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2019
Dr. Daniel Dörr · Michel T. Henrichs

https://gi.cebitec.uni-bielefeld.de/teaching/2019summer/sa

Zusatztutorium vom 9.5.2019

Aufgabe 1 (Alignmentvarianten)
Gegeben seien Sequenzen $s = \texttt{ACTGGCTTTCG}$ und $t = \texttt{ACTCTAGCT}$, Gapkosten $\gamma = 1$ und folgende Scorefunktion

$$score(a, c) = \begin{cases} 1 & \text{wenn } a = c, \quad (a, c \in \Sigma) \\ -1 & \text{sonst.} \end{cases}$$

Berechne

1. das Globale Alignment von s und t,
2. das Semi-globale Alignment von $s[1..7]$ und t,
3. das Free-End-Gap Alignment von s und t,
4. das Lokale Alignment von s und t,
5. das Globale Alignment mit affinen Gapkosten, gegeben Gaperöffnungskosten $d = 1$ und Gaperweiterungskosten $e = 0.1$.

Globales Alignment

$s \setminus t$	ε	A	C	T	C	T	A	G	C	T
ε										
A										
C										
T										
G										
G										
C										
T										
T										
T										
C										
G										

Semi-globales Alignment

$s[1..7] \setminus t$	$\mathcal{E}$	A	C	T	C	T	A	G	C	T
$\mathcal{E}$										
A										
C										
T										
G										
G										
C										
T										

Free-End-Gap Alignment

$s \setminus t$	$\mathcal{E}$	A	C	T	C	T	A	G	C	T
$\mathcal{E}$										
A										
C										
T										
G										
G										
C										
T										
T										
T										
C										
G										

Lokales Alignment

$s \setminus t$	ε	A	C	T	C	T	A	G	C	T
ε										
A										
C										
T										
G										
G										
C										
T										
T										
T										
C										
G										

Globales Alignment mit affinen Gapkosten

H

$s \setminus t$	ε	A	C	T	C	T	A	G	C	T
ε										
A										
C										
T										
G										
G										
C										
T										
T										
T										
C										
G										

V

$s \setminus t$	ε	A	C	T	C	T	A	G	C	T
ε										
A										
C										
T										
G										
G										
C										
T										
T										
T										
C										
G										

S

$s \setminus t$	ε	A	C	T	C	T	A	G	C	T
ε										
A										
C										
T										
G										
G										
C										
T										
T										
T										
C										
G										

Lösung

1. Score des optimalen Globalen Alignments: 1

$s \setminus t$	ε	A	C	T	C	T	A	G	C	T
ε	0	-1	-2	-3	-4	-5	-6	-7	-8	-9
A	-1	1	0	-1	-2	-3	-4	-5	-6	-7
C	-2	0	2	1	0	-1	-2	-3	-4	-5
T	-3	-1	1	3	2	1	0	-1	-2	-3
G	-4	-2	0	2	2	1	0	1	0	-1
G	-5	-3	-1	1	1	1	0	1	0	-1
C	-6	-4	-2	0	2	1	0	0	2	1
T	-7	-5	-3	-1	1	3	2	1	1	3
T	-8	-6	-4	-2	0	2	2	1	0	2
T	-9	-7	-5	-3	-1	1	1	1	0	1
C	-10	-8	-6	-4	-2	0	0	0	2	1
G	-11	-9	-7	-5	-3	-1	-1	1	1	1

2. Score des optimalen Semi-globalen Alignments: 3

$s[1..7] \setminus t$	ε	A	C	T	C	T	A	G	C	T
ε	0	0	0	0	0	0	0	0	0	0
A	-1	1	0	-1	-1	-1	1	0	-1	-1
C	-2	0	2	1	0	-1	0	0	1	0
T	-3	-1	1	3	2	1	0	-1	0	2
G	-4	-2	0	2	2	1	0	1	0	1
G	-5	-3	-1	1	1	1	0	1	0	0
C	-6	-4	-2	0	2	1	0	0	2	1
T	-7	-5	-3	-1	1	3	2	1	1	3

3. Score des optimalen Free-End-Gap Alignments: 3

$s \setminus t$	ε	A	C	T	C	T	A	G	C	T
ε	0	0	0	0	0	0	0	0	0	0
A	0	1	0	-1	-1	-1	1	0	-1	-1
C	0	0	2	1	0	-1	0	0	1	0
T	0	-1	1	3	2	1	0	-1	0	2
G	0	-1	0	2	2	1	0	1	0	1
G	0	-1	-1	1	1	1	0	1	0	0
C	0	-1	0	0	2	1	0	0	2	1
T	0	-1	-1	1	1	3	2	1	1	3
T	0	-1	-2	0	0	2	2	1	0	2
T	0	-1	-2	-1	-1	1	1	1	0	1
C	0	-1	0	-1	0	0	0	0	2	1
G	0	-1	-1	-1	-1	-1	-1	1	1	1

4. Score des optimalen Lokalen Alignments: 4

$s \setminus t$	ε	A	C	T	C	T	A	G	C	T
ε	0	0	0	0	0	0	0	0	0	0
A	0	1	0	0	0	0	1	0	0	0
C	0	0	2	1	1	0	0	0	1	0
T	0	0	1	3	2	2	1	0	0	2
G	0	0	0	2	2	1	1	2	1	1
G	0	0	0	1	1	1	0	2	1	0
C	0	0	1	0	2	1	0	1	3	2
T	0	0	0	2	1	3	2	1	2	4
T	0	0	0	1	1	2	2	1	1	3
T	0	0	0	1	0	2	1	1	0	2
C	0	0	1	0	2	1	1	0	2	1
G	0	0	0	0	1	1	0	2	1	1

5. Score des optimalen Globalen Alignments mit affinen Gapkosten: 2.6

H	ε	A	C	T	C	T	A	G	C	T
ε	$-\infty$	-1	-1.1	-1.2	-1.3	-1.4	-1.5	-1.6	-1.7	-1.8
A	$-\infty$	-2	0	-0.1	-0.2	-0.3	-0.4	-0.5	-0.6	-0.7
C	$-\infty$	-2.1	-1	1	0.9	0.8	0.7	0.6	0.5	0.4
T	$-\infty$	-2.2	-1.1	0	2	1.9	1.8	1.7	1.6	1.5
G	$-\infty$	-2.3	-1.2	-0.1	1	1	0.9	0.8	1.8	1.7
G	$-\infty$	-2.4	-1.3	-0.2	0.9	0.8	0.7	0.6	0.9	0.8
C	$-\infty$	-2.5	-1.4	-0.3	0.8	1.9	1.8	1.7	1.6	1.9
T	$-\infty$	-2.6	-1.5	-0.4	0.7	0.9	2.9	2.8	2.7	2.6
T	$-\infty$	-2.7	-1.6	-0.5	0.6	0.8	1.9	1.9	1.8	1.7
T	$-\infty$	-2.8	-1.7	-0.6	0.5	0.7	1.8	1.7	1.6	1.5
C	$-\infty$	-2.9	-1.8	-0.7	0.4	1.5	1.7	1.6	1.5	1.9
G	$-\infty$	-3	-1.9	-0.8	0.3	0.5	1.6	1.5	1.8	1.7

V	ε	A	C	T	C	T	A	G	C	T
ε	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
A	-1	-2	-2.1	-2.2	-2.3	-2.4	-2.5	-2.6	-2.7	-2.8
C	-1.1	0	-1	-1.1	-1.2	-1.3	-1.4	-1.5	-1.6	-1.7
T	-1.2	-0.1	1	0	-0.1	-0.2	-0.3	-0.4	-0.5	-0.6
G	-1.3	-0.2	0.9	2	1	0.9	0.8	0.7	0.6	0.5
G	-1.4	-0.3	0.8	1.9	1	0.8	0.7	1.8	0.8	0.7
C	-1.5	-0.4	0.7	1.8	0.9	0.7	0.6	1.7	0.8	0.6
T	-1.6	-0.5	0.6	1.7	1.9	0.9	0.8	1.6	1.9	0.9
T	-1.7	-0.6	0.5	1.6	1.8	2.9	1.9	1.8	1.8	2.9
T	-1.8	-0.7	0.4	1.5	1.7	2.8	1.9	1.7	1.7	2.8
C	-1.9	-0.8	0.3	1.4	1.6	2.7	1.8	1.6	1.6	2.7
G	-2	-0.9	0.2	1.3	1.5	2.6	1.7	1.5	1.9	2.6

S	ε	A	C	T	C	T	A	G	C	T
ε	0	-1	-1.1	-1.2	-1.3	-1.4	-1.5	-1.6	-1.7	-1.8
A	-1	1	0	-0.1	-0.2	-0.3	-0.4	-0.5	-0.6	-0.7
C	-1.1	0	2	1	0.9	0.8	0.7	0.6	0.5	0.4
T	-1.2	-0.1	1	3	2	1.9	1.8	1.7	1.6	1.5
G	-1.3	-0.2	0.9	2	2	1	0.9	2.8	1.8	1.7
G	-1.4	-0.3	0.8	1.9	1	1	0.7	1.9	1.8	0.8
C	-1.5	-0.4	0.7	1.8	2.9	1.9	1.8	1.7	2.9	1.9
T	-1.6	-0.5	0.6	1.7	1.9	3.9	2.9	2.8	2.7	3.9
T	-1.7	-0.6	0.5	1.6	1.8	2.9	2.9	1.9	1.8	3.7
T	-1.8	-0.7	0.4	1.5	1.7	2.8	1.9	1.9	1.7	2.8
C	-1.9	-0.8	0.3	1.4	2.5	2.7	1.8	1.6	2.9	2.7
G	-2	-0.9	0.2	1.3	1.5	2.6	1.7	2.8	1.9	2.6