

Konservierung

ASE Übung

31.05.2019

Sequenzalignment

ALYIC

|||

GLY-C

match = 1

mismatch = -1

gap = -2

$\Rightarrow -1 + 1 + 1 + -2 + 1 = 0$

Substitution Matrix (blosum62)

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	4	-1	-2	-2	0	-1	-1	0	-2	-1	-1	-1	-1	-2	-1	1	0	-3	-2	0
R	-1	5	0	-2	-3	1	0	-2	0	-3	-2	2	-1	-3	-2	-1	-1	-3	-2	-3
N	-2	0	6	1	-3	0	0	0	1	-3	-3	0	-2	-3	-2	1	0	-4	-2	-3
D	-2	-2	1	6	-3	0	2	-1	-1	-3	-4	-1	-3	-3	-1	0	-1	-4	-3	-3
C	0	-3	-3	-3	9	-3	-4	-3	-3	-1	-1	-3	-1	-2	-3	-1	-1	-2	-2	-1
Q	-1	1	0	0	-3	5	2	-2	0	-3	-2	1	0	-3	-1	0	-1	-2	-1	-2
E	-1	0	0	2	-4	2	5	-2	0	-3	-3	1	-2	-3	-1	0	-1	-3	-2	-2
G	0	-2	0	-1	-3	-2	-2	6	-2	-4	-4	-2	-3	-3	-2	0	-2	-2	-3	-3
H	-2	0	1	-1	-3	0	0	-2	8	-3	-3	-1	-2	-1	-2	-1	-2	-2	2	-3
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4	2	-3	1	0	-3	-2	-1	-3	-1	3
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4	-2	2	0	-3	-2	-1	-2	-1	1
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5	-1	-3	-1	0	-1	-3	-2	-2
M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5	0	-2	-1	-1	-1	-1	1
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6	-4	-2	-2	1	3	-1
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7	-1	-1	-4	-3	-2
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4	1	-3	-2	-2
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5	-2	-2	0
W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11	2	-3
Y	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	-1
V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4

Gewichtetes Sequenzalignment

ALYIC

|||

GLY-C

A -> G = 0

L -> L = 4

Y -> Y = 7

C -> C = 9

gap = -2

$\Rightarrow 0 + 4 + 7 + -2 + 9 = 18$

Multiples Sequenzalignment

ALYIC
GLY-C
GLR-C
-LY-I
-LY-I

Mit Lücken:

- Konsensus: GLY-C
- Konservierung: $[0.4, 1.0, 0.8, -, 0.6]$

Ohne Lücken:

- Konsensus: GLYIC
- Konservierung: $[0.\bar{6}, 1.0, 0.8, 1.0, 0.6]$

(Vereinfacht - unterschiedliche Ansätze zur Berechnung)

Frage 1

Welche Vorhersagen lassen sich mit Hilfe konservierter Sequenzen treffen?

Frage 2

Welche Tendenzen bzgl. Sequenzkonservierung erwarten Sie, wenn das Innere eines Proteins mit dessen Oberfläche verglichen wird und warum?

Frage 3

Im gegebenen Beispiel wurde die Konservierung als Verhältnis des Konsensuselements zur Größe der Spalte berechnet.

In welcher Hinsicht ist dieser einfache Ansatz problematisch? Was wäre eine bessere Alternative?

Frage 3 (cont.)

Col 1

Col 2

A

A

A

A

A

A

A

A

A

A

B

B

C

B

D

B

E

B

F

C

G

C

H

C

Frage 3 (cont.)

$$-\left(\frac{5}{12} \log_{20} \frac{5}{12} + 7 \times \left(\frac{1}{12} \log_{20} \frac{1}{12}\right)\right) \approx 0.6056$$

$$-\left(\frac{5}{12} \log_{20} \frac{5}{12} + \frac{4}{12} \log_{20} \frac{4}{12} + \frac{3}{12} \log_{20} \frac{3}{12}\right) \approx 0.3935$$