1 Bioinformatik Übung 2

Sarah Lange

1.1

 ${\bf I}={\bf prozentuale}$ Übereinstimmung | ${\bf M}={\bf Anzahl}$ der Übereinstimmungen | ${\bf L}={\bf L\ddot{a}n}$ ge der Sequenz

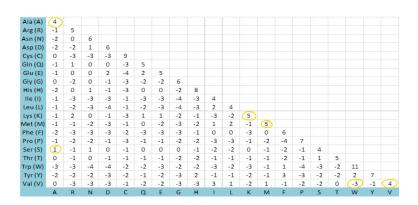
$$M=4$$
 $L=6$

$$I = \frac{100 \cdot M}{L}$$

$$I = \frac{100 \cdot 4}{6} = 66,7\%$$

Bei der prozentualen Berechnung gilt als Faustregel, wenn I > 30%, dann sind die Proteine wahrscheinlich homolog. Da bei der Berechnung ein Prozentwert von 66,7% heraus kam, kann also davon ausgegangen werden, dass die zwei Sequenzen homolog sind.

1.2



: Abbildung 1 BLOSUM62 Matrix

Quelle: https://students.aertslab.org/ebook/chapter $6_a lignments.html(5.5.22, 12:15 Uhr)$

A A K M W V
$$A~S~K~M~V~V~\Rightarrow +4~-1~+5~+5~-3~+4=14$$

1.3

siehe Code

1.4 Needleman-Wunsch-Algorithmus

σ= - λ	-	G	A	T	Τ	A
_	0	- 1	-2	-3	- ۷	-5
				- 1		-3
				1		-3
A				0		• 0
7	· .			+ 2		- 0
G	-6	-3	-1	+ 1	+1	• 0

Abbildung 2: Needleman-Wunsch-Algorithmus mit $\sigma=-1$

Optimales Alignment :

 ${\bf G} \ {\bf C} \ {\bf A} \ {\bf T} \ {\bf G}$ -

G G A T T A

O = 0	-	G	A	T	Τ	A	
1	0	0	0	0	0	0	
G		+1			+1	+1	
_		+1	1	- † +∧	+ 1	+1	_
Α	0	+1	+2	+2	+ 2	+2	
T	0	+1	+2	+3	+3	+3	
G	0	+ 1	+2	+3	+3	+3	

Abbildung 3: Needleman-Wunsch-Algorithmus mit $\sigma=0$

Optimales Alignment :

G C A T - - G

G - A T T A -

weitere mögliche Alignments:

G G A - T - G

G - A A T A -

GCATT-

- - AT - TA

G C A - T - G

G - A A T A -

1.5

1.5.1

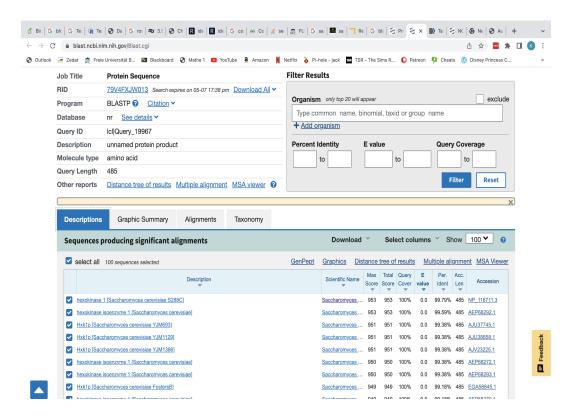


Abbildung 4: Ergebnisliste der BLAST-Suche

- 1. Welche BLAST Variante nutzen Sie? Warum?
 - Verwendung von blastp
 - vergleicht Proteinsequenz mit der Proteinen in Datenbank
 - kann daher anzeigen, zu welchem Protein diese Sequenz gehört, wonach in der Aufgabenstellung gefragt wird
 - kann zusätzlich Organismen mit ähnlichen Proteinen anzeigen, auch danach ist gefragt
- 2. Zu welchem Organismus gehört diese Sequenz wahrscheinlich?
 - Saccharomyces cerevisiae (Backhefe)

- 3.Zu welchem Protein gehört diese Sequenz wahrscheinlich?
 - Hexokinase I
- 4. Wie groß ist die Percentage Identity mit diesem Protein?
 - 99.79%
- 5. Wie lautet der E-Score und wie interpretieren Sie ihn?
 - E-Score ist hier 0.0
 - dies deutet darauf hin, dass die Proteine wahrscheinlich homolog sind

1.5.2

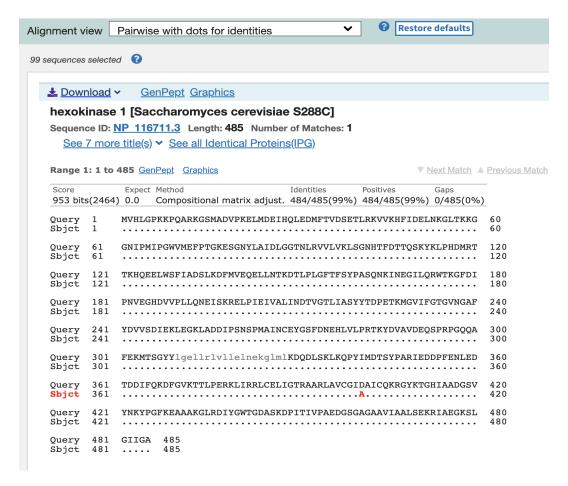


Abbildung 5: Alignment der vorgegebenen Sequenz mit der Vergleichssequenz, in rot ist die Mutation markiert. Wurde mithilfe von BLAST erstellt

Die Aminosäure A ist zur Aminosäure D mutiert. Wo genau, ist anhand der Grafik abzulesen.