

Softwarewerkzeuge der Bioinformatik

Prof. Dr. Volkhard Helms
PD Dr. Michael Hutter, Markus Hollander,
Marie Detzler
Wintersemester 2020/2021

Universität des Saarlandes
Zentrum für Bioinformatik

Übungsblatt 4

Sequenzanalyse: Motive

Lernziel: Sie sollen lernen, wie Sequenzmotive definiert sind, wie man sie darstellt und wie man entsprechende Datenbanken und Tools einsetzt. Weiterhin sollen Sie eine PSSM und die entsprechende Konsensussequenz berechnen können.

Aufgabe 4.1: PRINTS und PROSITE

- a) Suchen Sie das Protein Cytochrom C (**P00044**) in PRINTS (<http://bioinf.man.ac.uk/dbbrowser/PRINTS/index.php>). Hierfür könnte auch UniProtKB hilfreich sein.
 - i. Wie viele Motive gibt es?
 - ii. Aus wie vielen Proteinen wurden die Motive erstellt?
 - iii. Enthält **P00044** alle Motive?
 - iv. Suchen Sie die Motive in der Aminosäuresequenz.
- b) Vergleichen Sie Ihre potentiellen Motive aus dem Clustal Omega-Alignment von Aufgabe 3.3 mit den Einträgen für FOSB_MOUSE in den Datenbanken PROSITE (<http://www.expasy.org/prosite>) und PRINTS.
- c) Welches PRINTS-Motiv stimmt annähernd mit der PROSITE-Signatur überein?

Hinweis:

Accession Numbers:

Proteine haben unterschiedliche Accessions in UniProtKB, PROSITE and PRINTS.

Aufbau von PROSITE-Signaturen:

- [] Aminosäuren, die an dieser Position vorkommen können
- nächste Position
- x beliebige Aminosäure
- () Anzahl der Wiederholungen
- { } Aminosäuren, die an dieser Position nicht vorkommen dürfen

Aufgabe 4.2: MEME und MAST

<http://alternate.meme-suite.org>

- a) Verwenden Sie dieselben Proteine aus denen die PRINTS Motive für FOSB_MOUSE gebildet wurden zum Erstellen von MEME Motiven.
- b) Vergleichen Sie die MEME Motive mit den PRINTS Motiven.
- c) Testen Sie mithilfe von MAST ob FOSB_MOUSE diese MEME Motive enthält.

Aufgabe 4.3: Positions-spezifische Gewichtsmatrix (PSSM)

Berechnen Sie aus dem folgenden Alignment die PSSM und geben Sie die Konsensussequenz an. (X steht für alle nicht gelisten Aminosäuren.)

S E Q V E N C E	alignment matrix →	1	2	3	4	5	6	7	8		
S E Q Y E M C E		C	E	I	M	N	Q	S	V	Y	X
S I Q V I N S E		I	M	N	Q	S	V	Y	X	X	X
S E Q V E - C E		M	N	Q	S	V	Y	X	X	X	X
S E Q V E M C I		N	Q	S	V	Y	X	X	X	X	X
S E Q V E N S I		Q	S	V	Y	X	X	X	X	X	X
* * .		S	V	Y	X	X	X	X	X	X	X
consensus:		V	Y	X	X	X	X	X	X	X	X

	1	2	3	4	5	6	7	8
C								
E								
I								
M								
N								
Q								
S								
V								
Y								
X								

PSSM

$\forall i : p_i = 0.05$
a priori probability

Viel Spaß!