

Übung der angewandten Bioinformatik: Homology

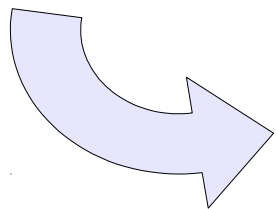
LXR - Sequenzprofil

- Sequenzprofile des Transkriptionsfaktors:
LXR

	1	2	3	4	5	6	7	8
A	0	0	0	7	0	0	2	1
C	0	0	4	1	6	2	2	2
G	8	0	0	0	1	1	3	4
T	0	8	4	0	1	5	1	1

Wahrscheinlichkeiten (p)

	1	2	3	4	5	6	7	8
A	0	0	0	7	0	0	2	1
C	0	0	4	1	6	2	2	2
G	8	0	0	0	1	1	3	4
T	0	8	4	0	1	5	1	1

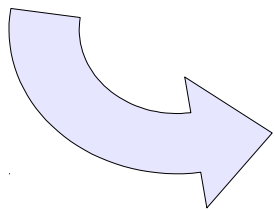


	1	2	3	4	5	6	7	8
A	0	0	0	7/8	0	0	1/4	1/8
C	0	0	1/2	1/8	3/4	1/4	1/4	1/4
G	1	0	0	0	1/8	1/8	3/8	1/2
T	0	1	1/2	0	1/8	5/8	1/8	1/8

Wahrscheinlichkeiten

Informationsgehalt H

	1	2	3	4	5	6	7	8
A	0	0	0	7/8	0	0	1/4	1/8
C	0	0	1/2	1/8	3/4	1/4	1/4	1/4
G	1	0	0	0	1/8	1/8	3/8	1/2
T	0	1	1/2	0	1/8	5/8	1/8	1/8



$$H_{i,a} = -p_{i,a} \log_2 p_{i,a}$$

	1	2	3	4	5	6	7	8
A				0.17			0.5	0.375
C			0.5	0.375	0.311	0.5	0.5	0.5
G	0				0.375	0.375	0.53	0.5
T		0	0.5		0.375	0.424	0.375	0.375

Rechenhilfe

$$\log_2 x/y = \log_2 x - \log_2 y$$

$$\log_2 1 = 0$$

$$\log_2 2 = 1$$

$$\log_2 4 = 2$$

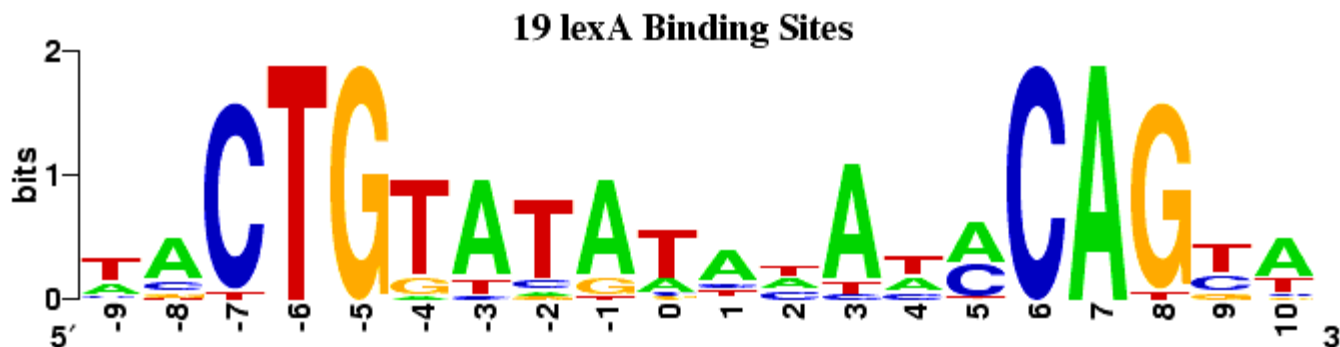
Beispiel:

$$\begin{aligned}\log_2 \frac{1}{4} &= \log_2 1 - \log_2 4 \\ &= 0 - 2 \\ &= -2\end{aligned}$$

$$\begin{aligned}&\rightarrow -\frac{1}{4} \cdot -2 \\ &= 0,5\end{aligned}$$

Logoplot / Sequence plot

- Grafische Darstellung
- Höhe aller Buchstaben an einer Position: Konserviertheit
- Höhe der einzelnen Buchstaben: relatives Verhältnis der Buchstaben zueinander



Regeln für Logoplot

- Gesamthöhe R an Position i

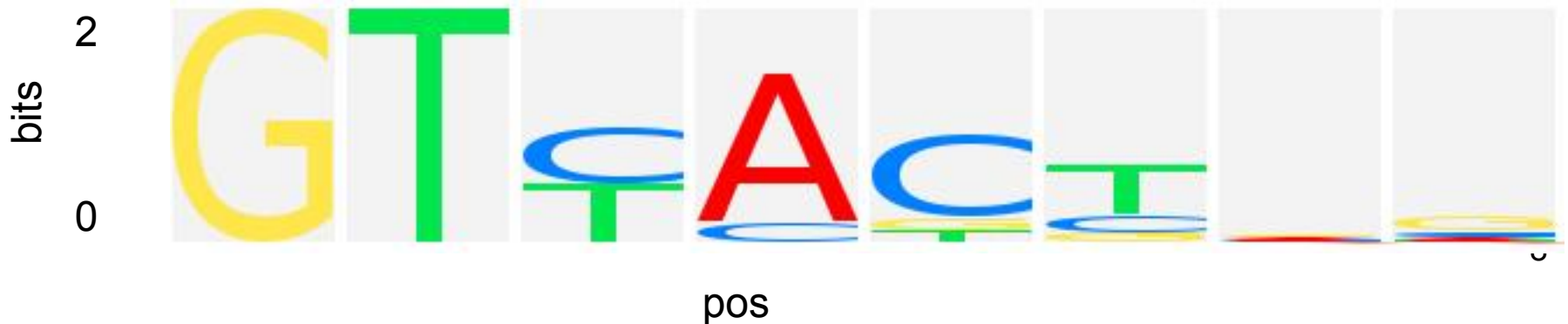
$$R_i = 2 - H_i$$

- Höhe der Buchstaben a an Position i

$$h = p_{ai} * R_i$$

Logo plot

	1	2	3	4	5	6	7	8
A				0.17			0.5	0.375
C			0.5	0.375	0.311	0.5	0.5	0.5
G	0				0.375	0.375	0.53	0.5
T		0	0.5		0.375	0.424	0.375	0.375
H _{pos}	0	0	1	0.545	1.061	1.299	1.905	1.75



Übungsseite

- <http://gobics.de/lectures/appliedBI/homology/>

- oder -

- <http://gobics.de>
- Teaching
- Bioinformatik im BSc-Studiengang Biologie
- Einführung in die Angewandte Bioinformatik
- Homology