

Hausaufgabe 6

Aufgabe 6.1 Laplace Succession (10%)

Die letzten sieben Vorlesungen haben stattgefunden. Nutzen Sie die Laplace Succession um die Wahrscheinlichkeit, dass die nächste Vorlesung auch stattfindet, zu bestimmen.

Der Ergebnisraum lautet: $\Omega = \{VL, \overline{VL}\}$

Damit ist $P(VL) = \frac{7+1}{7+1\Omega} = \underline{\underline{\frac{8}{9}}}$

Aufgabe 6.2 Sequenzmotive (20%)

Sie haben die folgenden Motive gefunden

G A T A C
G A T C A
A T T G A
G C T C G

Prof.L

G	$\frac{3}{4}$	0	0	$\frac{1}{4}$	$\frac{1}{4}$
C	0	$\frac{1}{4}$	0	$\frac{1}{2}$	$\frac{3}{4}$
A	$\frac{1}{4}$	$\frac{1}{2}$	0	$\frac{1}{4}$	$\frac{1}{2}$
T	0	$\frac{1}{4}$	1	0	0

Prof.L mit Pseudocounts

G	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$
C	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{4}$
A	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$
T	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{5}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Aufgabe 6.3 Profile (20%)

Sie haben das folgende Profil gefunden

A	1/2	1/4	0	1
C	1/2	1/4	0	0
T	0	1/4	1/2	0
G	0	1/4	1/2	0

1. $P(A C T A) = \frac{1}{2} \cdot \frac{1}{4} \cdot \frac{1}{2} \cdot 1 = \frac{1}{16}$

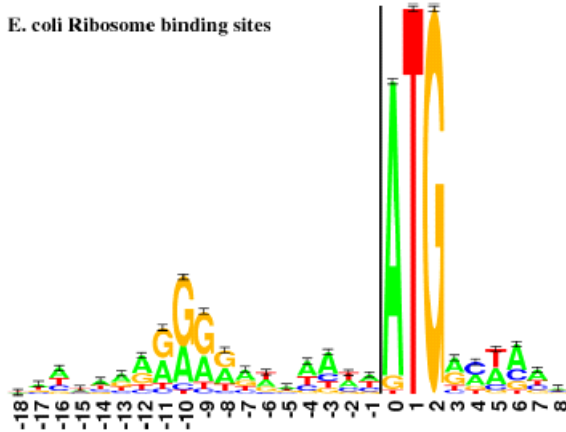
2. $P(A C T G) = \frac{1}{2} \cdot \frac{1}{4} \cdot \frac{1}{2} \cdot 0 = 0$

3. $P(C C T A A) = \frac{1}{2} \cdot \frac{1}{4} \cdot \frac{1}{2} \cdot 1 \cdot \frac{1}{4} = \frac{1}{64}$

Aufgabe 6.4 Sequence Logo (20%)

Interpretieren Sie das folgende Sequence Logo.

E. coli Ribosome binding sites



Diese in diesem Sequence Logo dargestellte Genabschnitt zeigt deutlich Unterschiede in der Informationsentropie und damit für die Konserviertheit zwischen den Basen. Sehr deutlich zu erkennen ist dabei ein sehr stark konserviertes Basentriplett ATG an den Stellen 0, 1 und zwei. An der Stelle 0 kommen noch in sehr kleinem Maße andere Basen vor, an den anderen beiden Stellen ist keine andere Base vorhanden. Dies lässt darauf schließen, dass dieses Triplett essentiell für die Bindung von Ribosomen in E. coli ist und eine Mutation einer dieser drei Basen die Funktion zerstört. Die Regionen um dieses Triplett sind deutlich unbestimmter und damit deutlich weniger relevant für die Funktion. Hier kommen alle Basen in vergleichbarer Häufigkeit vor.

Auffällig ist aber auch noch eine weitere Stelle ca. 10 Basenpaare vor dem konservierten Basentriplett. Hier ist eine Stelle deutlich kleinerer Informationsentropie, wo nahezu ausschließlich Guanin und Adenin vorkommen zu scheinen. Das lässt darauf schließen, dass auch diese Stelle konservierter ist als restliche Teile der Sequenz und somit eine wichtige Funktion für die Ribosombindung hat. Allerdings ist sie deutlich weniger stark konserviert als das zuvor erwähnte Basentriplett, weswegen eine Mutation an dieser Stelle die Funktion vermutlich einschränkt, nicht aber komplett zerstört.