

IC									
	1	2	3	4	5	6	7	8	9
A	0,421	1,878	0,594	0,010	2,373	0,000	0,000	0,000	0,263
T	0,000	0,000	0,421	0,000	0,000	2,373	0,000	0,124	-0,102
G	0,000	-0,160	-0,147	-0,160	0,000	0,000	1,704	0,921	-0,160
C	0,580	0,000	-0,147	0,746	0,000	0,000	0,000	0,000	0,282
IC (колонки)	1,001	1,718	0,721	0,596	2,373	2,373	1,704	1,046	0,282
IC(выравнивания)	11,813								

Итоговые информационные содержания отдельных букв и колонок.



Картинка с лого.

Формула, которая использовалась для расчетов:

$$IC(b,j) = f(b,j) \cdot \log_2[f(b,j)/p(b)] = f(b,j) \cdot w(b,j)$$

Для полученного выравнивания были посчитаны наблюдаемые частоты встречаемости нуклеотидов. Затем была вычислена матрица PWM без псевдокаунтов (как двоичный логарифм наблюдаемой частоты на ожидаемую). Ожидаемая частота встречаемости нуклеотидов была взята из данных по *Danio Rerio*.

Затем был посчитан информационный вес каждой буквы (с помощью формулы), информационный вес по позициям (как сумма) и суммарный вес.

С помощью сервиса <http://weblogo.berkeley.edu/logo.cgi> была получена лого-картинка.