



\*   \*   \*   \*   \*

**GUCUUCGGCAC**  
**GCCUUCGGAGC**  
**GCAUUCGUCGC**  
**GCUUUCGAUGC**  
**GCGUUCGCUGC**  
**GUAUUCGUAAC**  
**GUGUUCGCGAC**  
**GUUUUCGAGAC**

**2 1 2 2 2 2 1 2**

Each column is independent

Maximum information is 2 bits per position, if column is 100% conserved.

$$\log_2 (4/1) = 2 \text{ bits}$$

sequence profile:  
14 bits

expect a match  
by chance:  
1 in  $2^{14}$  nt  $\approx$  4 Kb