Фамилия\_\_\_\_\_\_\_Гарипова\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ имя\_\_\_\_\_\_\_Камила\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Дано выравнивание. GC генома равен 0.6. p(b) – частота буквы b. (b) = псевдоотсчёт для буквы b.

1. Напишите консенсус (Под каждой колонкой самую частую букву БОЛЬШОЙ, если такая не одна любую из них маленькой)
2. Напишите паттерн используя таблицу “ Ambiguous nucleotide codes”
3. Напишите тот же паттерн на комплементарной цепи
4. Постройте матрицу PWM для данного выравнивания
5. Вычислите вес W данной последовательности относительно полученной матрицы PWM

**Выравнивание 2**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| G | T | C | T | C | G | C | A | A | A | C | G | T | T | T | G |
| C | A | C | A | C | G | C | A | A | A | C | G | T | T | T | T |
| T | C | C | A | C | G | C | A | A | A | C | G | G | T | T | T |
| G | C | C | A | C | G | C | T | A | C | C | G | T | T | T | T |
| G | A | T | A | C | G | C | A | A | A | C | G | T | G | T | G |
| C | C | G | A | C | G | C | A | A | T | C | G | G | T | T | A |
| G | T | T | G | C | G | C | A | A | A | C | G | T | T | T | T |

Число букв по столбцам

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| **A** | **0** | 2 | 0 | 5 | 0 | 0 | 0 | 6 | 7 | 5 | 0 | 0 | 0 | 0 | 0 | 1 |
| **T** | **1** | 2 | 2 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 5 | 6 | 7 | 4 |
| **G** | **4** | 0 | 1 | 1 | 0 | 7 | 0 | 0 | 0 | 0 | 0 | 7 | 2 | 1 | 0 | 2 |
| **C** | **2** | 3 | 4 | 0 | 7 | 0 | 7 | 0 | 0 | 1 | 7 | 0 | 0 | 0 | 0 | 0 |
| **consensus** | **G** | C | C | A | C | G | C | A | A | A | C | G | T | T | T | T |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **pattern** | **G** | C | C | A | C | G | C | A | A | A | C | G | T | T | T | T |
| **complement** | **C** | G | G | T | G | C | G | T | T | T | G | C | A | A | A | A |

Матрица PWM **и в в**ес последовательности относительно PWM

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | p(b) | (b) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| A | 0,2 | 1 | -0,8 | 0,3 | -0,8 | 1,01 | -0,8 | -0,8 | -0,8 | 1,15 | 1,3 | 1,01 | -0,8 | -0,8 | -0,8 | -0,8 | -0,8 | -0,1 |
| T | 0,2 | 1 | -0,1 | 0,3 | 0,3 | -0,1 | -0,8 | -0,8 | -0,8 | -0,1 | -0,8 | -0,1 | -0,8 | -0,8 | 1,01 | 1,15 | 1,3 | 0,81 |
| G | 0,3 | 1 | 0,4 | -1,2 | -0,5 | -0,5 | -1,2 | 0,88 | -1,2 | -1,2 | -1,2 | -1,2 | -1,2 | 0,88 | -0,1 | -0,5 | -1,2 | -0,1 |
| C | 0,3 | 1 | -0,1 | 0,18 | 0,4 | -1,2 | 0,88 | -1,2 | 0,88 | -1,2 | -1,2 | -0,5 | 0,88 | -1,2 | -1,2 | -1,2 | -1,2 | -1,2 |
|  | 1 | 4 | -0,1 | 0,18 | 0,3 | 1,01 | 0,88 | 0,88 | 0,88 | 1,15 | 1,3 | 1,01 | 0,88 | 0,88 | 1,01 | 1,15 | 1,3 | 0,81 |
| W= | 13,52 | | C | C | T | A | C | G | C | A | A | A | C | G | T | T | T | T |