

## Exercise sheet 8: Substitution Scoring

---

### Exercise 1 - Point Accepted Mutation (PAM)

We want to calculate the  $PAM_1$  matrix based on the following two sequence alignments of the DNA sequences a, b, c and d.

Tip: In order to solve a) and b) create a combined alignment comprised of two combined sequences a' and b' (based on the two initial alignments and their symmetric counterparts)

$$a' = a + c + b + d$$

$$b' = b + d + a + c$$

The order does not matter, as the frequency identification is position-insensitive.

Unless otherwise stated round all results to 4 decimal places.

**1a)**

Calculate the nucleotide frequencies  $r_x$

**Hide**

**Hint : Formulae**

**Answer**

$$r_A = 0.3333 \quad (1)$$

$$r_C = 0.1667 \quad (2)$$

$$r_T = 0.3333 \quad (3)$$

$$r_G = 0.1667 \quad (4)$$

$$(5)$$

**1b)**

Calculate the symmetric mutation matrix  $E(x, y)$ .

**Hide**

**Hint : Formulae**

**Intermediate Values** Non normalized values. Further multiplied by  $|a'|$

**Answer**

**1c)**

Calculate the non-normalized PAM matrix  $S$  with  $10 * \log_{10}(odds)$ , using the previously determined  $r$  values and  $E$  matrix. (round to integers)

**Hide**

**Hint : Formulae**

**Answer**

**1d)**

Given the sequences  $a = ACC$  and  $b = ATT$ , compute the optimal Needleman-Wunsch alignments using:

1. The general similarity scoring function.
2. The PAM1-based similarity scoring function.

**Hide**

**Hint: Possible Answers**

**Answer** B

**1e)**

Calculate the normalization factor  $\gamma$  based on  $E$  .

**Hide**

**Hint: Formulae**

**Answer**  $\gamma = 0.027$

1f)

Calculate the mutation rate matrix  $P$ .

**Hide**

**Hint: Formulae**

**Answer**

1g)

Calculate the normalized mutation rate matrix  $P'$  using  $P$  and the normalization factor  $\gamma$ .

**Hide**

**Hint: Formulae**

**Answer**

1h)

Determine  $PAM_1$  based on the normalized mutation rate matrix  $P'$  with  $10 * \log_{10}(odds)$  (round to integer)

**Hide**

**Hint: Formulae**

**Answer**

1i)

Determine  $PAM_2$ . (round to integer)

**Hide**

**Hint: Formulae**

**Answer**

## **Exercise 2 - Programming assignment**

Programming assignments are available via Github Classroom and contain automatic tests.

We recommend doing these assignments since they will help you to further understand this topic.

Access the Github Classroom link: [Programming Assignment: Sheet 08](#).