Exercises – Phylogenetics

Exercise Sheet 9 — 13.06.2019

Due: 27.06.2019

Task 1 Evolutionary Markov Processes. (3 points)

Use your own words to explain three of the five properties an *evolutionary Markov process (EMP)* has by definition. Write 2–5 sentences per property. Do not simply reproduce the formulas.

Task 2 Modeling of Sequence Evolution. (7 points)

Consider the following model of nucleotide substitution:

$$\pi = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$$

$$Q = \begin{pmatrix} -5/8 & 3/8 & 1/8 & 1/8 \\ 3/8 & -5/8 & 1/8 & 1/8 \\ 1/8 & 1/8 & -5/8 & 3/8 \\ 1/8 & 1/8 & 3/8 & -5/8 \end{pmatrix}$$

$$P(t) = \begin{pmatrix} 1-(x+2y) & x & y & y \\ x & 1-(x+2y) & y & y \\ y & y & 1-(x+2y) & x \\ y & y & x & 1-(x+2y) \end{pmatrix}$$

where

$$x = \frac{1 - 2\exp(-t) + \exp(-t/2)}{4}, \; y = \frac{1 - \exp(-t/2)}{4}$$

(a) How many PEM is the model calibrated to?

(b) Recalibrate the corresponding matrix to 1 PEM.
    (You are **not** supposed to calibrate **both** matrices to 1 PEM.)

(c) How many PAM is the model calibrated to?

(d) Recalibrate the corresponding matrix to 1 PAM.
    (You are **not** supposed to calibrate **both** matrices to 1 PAM.)

Task 3 Jukes-Cantor Correction. (5 points)

Consider an alignment with length 900 and 65 substitutions (no InDel included). We assume an evolution of the sequences following the Jukes-Cantor model.

(a) What is the Jukes-Cantor corrected distance d (in PEM) of the sequences?

(b) How many mutation events occured approximately during the evolution of the above sequence?

(c) How many mutation events occur and how many substitutions in the alignment would appear, if the distance d between the sequences is 20 PEM?