

# Computational Biology

## Assignment 2 - Report

Philip Hartout  
phartout@student.ethz.ch

October 28, 2019

1. I would expect the distribution of nucleotides not to change much given that the rate of change is slow compared to the time scale of the tree.
2. In that case, I expect the distribution of the nucleotides to approach the distribution of the equilibrium frequencies (0.25 for each nucleotide if there is no particular selective pressure) given the rate of change is fast relative to the size of the tree.
3. We see that each row of the transition matrix  $Q$  approaches the initial nucleotide distribution  $\pi$  after approximately 600 mya.
4. For a nucleotide  $i$ , randomly sampling from the exponential distribution with rate  $\lambda = -q_{ii}$  gives the time until the next substitution event.
5. Sample from uniform distribution from 0 until  $1 - q_{ii}$ , then given transition probabilities provided in  $Q$ , determine the interval in which the sample drawn ends up in, and derive the nucleotide name.