Bi/Ge105: Evolution
Homework 5
Due Date: Wednesday, February 19, 2020

"We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes."

–Pierre Simon Laplace, A Philosophical Essay on Probabilities

1. The Wright-Fisher Model and Genetic Drift

In the previous homework, you had a chance to explore genetic drift by direct numerical simulation of the effects of finite population sizes. What we found was that because of the simple properties of binomial partitioning, there are fluctuations of allele frequencies that over long times can result in the fixation of some alleles over others in a population, thus resulting in neutral evolution (as opposed to selection). In this problem we are going to flesh out the approach described in class using the transition matrix $\mathbf{P}$ in conjunction with the evolution law $\mathbf{X}_n = \mathbf{P}\mathbf{X}_{n-1}$. Like in class, here we use the notation $\mathbf{X}_n$ to refer to the vector that carries the probabilities of all the different possible allele counts that could exist in the population at generation n. For example, in Buri's experiment on fly eye color, this is a 33-dimensional vector since there are 16 flies and hence 32 total alleles and the number of red alleles in a given vial can be 0, 1, 2, $\cdots$ 32, giving 33 entries in our vector. Your job is to use the transition matrix

$$P_{i\to j} = \frac{(2N)!}{j!(2N-j)!}\left(\frac{i}{2N}\right)^j\left(1-\frac{i}{2N}\right)^{2N-j} \tag{1}$$

to generate and plot the time series of $\mathbf{X}$ from one generation to the next. Make three plots, one with $N = 16$ as in the Buri case, one with $N = 10$ (5 male, 5 female) and one with $N = 30$ (15 male, 15 female) and use your graphs to comment on the time to fixation.

2. The Eocyte Hypothesis

(A)**Reading questions** To prepare for the eocyte hypothesis vs. three domain hypothesis debate next week, we would like you to read several papers and reflect on them. The papers can be found at https://tinyurl.com/Eocyte3D.

(B)**Summarize** Please write a 1-2 sentence summary of the eocyte hypothesis and the three domain hypothesis. What are the key differences between them. Also, write a few sentences about each of the papers summarizing their main points. Remember that the Woese (1981) article is optional reading, but it may provide helpful context and background for the other papers.

(C)**Reflect** Please wait to do this section until after you have read all of the papers and written their summaries. For each of the papers, please write at least one question that you had about that paper and discuss at least one connection you noticed to one or more of the other papers.