Notes with Ron

Scott Teresi

19 February, 2019

Questions for Ron:

In a class about random walks, which revolves around modeling and learning about random processes in biological systems. I am interested in transposons and their insertion locations. We are tasked with doing a semester-long project and I thought, given my data on the strawberry, that it would be feasible to compare the location of transposons against a null model. Obviously things are much more complex than what we are trying to do, but at best I think it would be cool to see that **x** and **y** TEs prefer certain areas over the other. Maybe make that cycling graph where over the length of the entire chromosome you can say "There is a gene desert here, and here there is a TE desert".

Notes with Ron:

You an think of allele frequencies as transposon distributions. Ron works on Monte Carlo simulations. Some lack of dependence with TEs. Has an array of 1s and 0s for 2000 lines representing 2000 chromosomes, where each 1 is a TE. He has some copy rate, for each generation per TE. But he also has a probability of excision, the chance that a TE gets lost during copying. Can't have the excision rate vastly outnumber the copy rate. He also has a raining in value where if you think fo a TE insertion from a virus you add in a TE at a random place. If I only have these things and nothing else going on, my TEs will all either go to extinction or fill up the entire chromosome.

Recombination, two arrays of the same length. Generate some random number somewhere along the array and swap the arrays. Best way to fix the fact that everything will go extinct or fill up, without getting computationally hard:

$$Fitness = \frac{1}{1 + TotalTELoad}$$

Fitness:

We can also add in a fitness value to help manage TEs fixing in the population at high rates. Organism with twice the fitness of something else should be twice as likely to survive or reproduce. And then do a Monte Carlo from 0 to 1, Add up the fitnesses. and then divide by that number to scale everything to 1. This gives sme a fitness score and we can select organisms with probability proportional to their fitness.

Construct a number line with the share of the number line equivalent to the organisms fitness so more fit individuals have a higher share of the number line, and then use a random number (Monte Carlo) to pick a spot on the number line where they choose which ones repreoduce. This can be used to keep a steady amount while not having the excision rate (he doesn't consider the excision rate here) to make sure that you don't crash the TEs.

Another Option:

Copy rate as a function of load. Saying effective copy rate is inserting into a non TE and so effective copy rate goes down as transposon load or density increases.

$$PropbabiltyOfCopy = \frac{Probability}{1+n}$$