

## Unessay Project: Simulating the Forces of Evolution through an Interactive Web App

Created by Patrick Murphy

Interactive Website Link: <https://patlm.github.io/anth-242-unessay-project/>

Code Repository (if interested): <https://github.com/patlm/anth-242-unessay-project>

### Introduction to Program

For my unessay project, I created a web application allowing the configuration and simulation of a population with three different traits. Furthermore, for the four forces of evolution, I added a way that each force can be directly impacted through a configuration setting.

### Simulation Algorithm

At the core of the simulation is a genetic algorithm based algorithm. The first step of this algorithm is to generate a random initial population. The way individuals are represented in the simulation is through 6 letter sequences where each letter can either be an A (dominant allele), a (recessive allele), or o (missing allele). The first two alleles correspond to the eye color trait, the second two alleles correspond to the climate adaptability trait, and the final two alleles correspond to the reproducibility trait of the individual. Below, is a list of the characteristics and what each allele set means:

#### *Eye Color*

AA: Green

Aa: Blue

aa: Brown

If contains an o (missing allele) in any position: Gray

#### *Climate*

AA: Better adapted for warm climate

Aa: Better adapted for warm climate

aa: Better adapted for cold climate

If contains an o (missing allele) in any position: Dies (result of missing allele is deadly)

#### *Able to Reproduce*

AA: Yes

Aa: Yes

aa: No

If contains an o (missing allele) in any position: Dies (result of missing allele is deadly)

For the initial population, all alleles are complete in that they do not contain missing alleles. This initial population is considered generation 1. To move on to the second generation, first, a list of parent pairs are generated taking into account climate adaptability and reproducibility measures of the parents. After selecting parent pairs, the punnet square operation is performed. This operation takes the two alleles from each of the parents and gives the child a 25% chance of inheriting the alleles expressed in each of the four corners of the punnet square. The final step in generating the second generation is performing mutations. For each allele an individual has, there is the probability specified in the configuration section chance that it has a mutation and a "probability \* 0.01" chance that it has an allele replaced with an empty allele (o). A mutation, when not replacing an allele with an empty allele, flips which allele is at the specified location for the individual. So if an allele was originally an "A", the mutation would replace the allele with an "a".

### Configurations

#### *Natural Selection*

In the natural selections portion of the configuration, the current climate type can be selected from the options of warm, normal, or cold. This relates to natural selection because when this is configured to a

warm or cold climate, individuals who are adapted for the opposite type of climate will be 20% less likely to be a parent of an offspring for the next generations. This means that overtime, when there are few impacts from other forces of evolution, the general population will trend towards having a higher percentage of individuals adapt for the current climate condition. This trend is a perfect illustration of natural selection and how populations will become more adapt to a climate over time.

### *Mutation*

For mutations, you can set the frequency of mutations as a percent. What the percent represents is for each allele, what is the probability that it will be mutated. As noted on the website, an un-configurable part of the mutation force is that 1% of the mutations will result in an allele being replaced with an empty allele which is typically fatal. For the non-fatal allele mutations, a dominant allele will be replaced with a recessive allele or vice versa.

### *Gene Flow*

For gene flow, you are able to set the gene flow type to one-way gene flow, none, or multi-way gene flow. In one-way gene flow, one-way gene flow will be simulated through bringing in random individuals that have no missing alleles. Specifically, the number of individuals brought in with one-way gene flow is equal to 10% of the current generation's population.

For the none option, there will be no gene flow occurring meaning that all individuals created for the next generation have parents from the current generation.

For multi-way gene flow, multi-way gene flow is simulated through removing 10% of the current population. Then, that 10% of the population is crossed over with randomly generated individuals (who have no missing alleles). The children of these cross overs are then contributed back to the population meaning that this gene flow has no impact on the population size but will tend to have a slight impact on the number of mutated individuals.

### *Genetic Drift*

For the genetic drift configuration, the attribute I chose that could be configured is being able to apply the founder's effect or population bottleneck effect.

The founder's effect is illustrated by the idea that a subsection of the current population picks up and moves to somewhere else. After clicking the button to apply the founder's effect, the smaller generation that has moved is what will be shown in the simulation. Specifically, for populations greater than 1,000, 100 individuals will be selected as the "founders." For populations smaller than 1,000, 10% of the current population size will be selected as founders. This helps capture the founders effect because randomly selecting these individuals can lead to a different distribution of alleles and phenotypes.

The population bottleneck effect is illustrated in a similar manner. Accept for the population bottleneck, the population that we are monitoring remains the same. We simply remove all but 100 or 10% of the individuals, whatever is smaller, from the population and continue to monitor this population that now has a different distribution of alleles and phenotypes.

### *Advanced Configurations*

For the advanced configurations, there are three different configurations that are mostly utilitarian in nature. First, you can set how many generations per second will run. This setting is available to make it easier to speed up or slow down the simulation. With large population sizes, however, the program may slow down below what is set.

The second configuration is for the initial population size. This simply sets how large the initially generated population is.

The last configuration is the offspring multiple, this multiple is used to set how much the population of the next generation should be increased or decreased. Simply put, this number is multiplied by the current population size to determine how many parent pairs should be generated as parents for the next generation. This was added as a configuration because increasing this multiple can help combat running the population size to 0 when configuration settings preference recessive traits such as colder climates. Overall, this configuration helps provide tools to the user of the website to keep the population size around a desired number.

### **Phenotypes in Current Generation**

The phenotypes in current generation section of the website simply provides the number of individuals who have each phenotype in the current population. This is interesting to monitor as the simulation runs and helps you visualize the impact your configuration settings are having on the simulation.

### **Current Generation Stats**

The current generation stats portion of the website allows the website user to see the generation number, current population size, and Hardy-Weinberg statistics for each of the traits. The Hardy-Weinberg equilibrium is the idea that in a large population with random mating and no disruptive circumstances, genotype and allele frequencies will remain constant. Using the website I created, it is cool to see that this holds true for the most part (the able to reproduce trait may have slightly harmed this illustration since homozygous recessive individuals for that trait will be unable to pass on their genes). Furthermore, when adding back different forces of evolution, it is evident that the genotypes and allele frequencies begin to move away from the original equilibrium.