Introduction to computer simulation of Population genetics using Populus.

I- Populus basic instructions:

Populus is a simulation of population biology software developed by Don Alstad at the University of Minnesota. You will need to download the program from their site: https://cbs.umn.edu/populus/download-populus. This program is a Java-Script that runs on Linus, MacOS and Windows. Read the instructions on Brightspace to download and run the program.

The simulations we will use during this exercise are accessible from the "Model" menu. Each simulation model consists of two windows: the **input windows** in which you enter values for various parameters of the simulation, and the **output window** where the results are displayed. Both windows are resizable.

You can zoom in any region of the display window by selecting the region of interest with the left mouse button (drag the selection rectangle around the region you want to zoom in). Right–clicking anywhere on the graph then choose "Auto Range > Both Axis" to reset the zoom.

You TA will demonstrate how to use Populous during the live session dedicated to this exercice.

Follow the instructions below and write your answers directly in this document:

II- Genetic drift:

In the **Model** menu, select **Mendelian Genetics** then **Genetic Drift**. The input window will appear. Make sure the **Monte Carlo** tab is selected.

By default, you can enter values for four parameters: population size (N), initial frequency (p), number of loci and number of generations.

Simulation 1-1:

First, let's simulate a population of **250 individuals** and observe what will happen to an allele with an **initial frequency of 0.5**.

In the input window enter these parameters: **N=250**, **p=0.5** and **number of loci=6**. The number of loci set to 6 allows us to observe 6 simulated populations at a time. Set the Runtime to **300 generations** (click on "other" in the runtime frame and type 300).

Next, click on "**View**" at the top of the input window to open the **Output window**. A new window appears, showing a graphic representation of the allele frequency over time (expressed in generation number). Each coloured broken line represents a locus (or, in our case, a population).

You will notice that a certain number of broken lines may have reached the value 0 or 1. This phenomenon is called **allele fixation**. If an allele frequency reaches 0, the allele is **lost** in the population. On the other hand, if the frequency reaches 1 it is said to be **fixed**.

Question 1: How many alleles reached **fixation** at 1? Which allele (blue, red…)?

Only a single allele reached fixation,which was the dark blue allele.

Question 2: What does it mean in term of genetic diversity for the population when an allele becomes fixed?

When an allele gets fixed, it means that every single individual within a population has that said allele at a locus, meaning that in terms of genetic diversity, there is none for that locus.

Run the simulation again (in the same conditions) by pressing the "**Iterate**" button in the output window.

Question3: Did you get the same result?

The result has changed.

Question 4: How many alleles have been **fixed** this time?

No alleles have been fixed this time, but the allele has been lost

Question 5: Are they the same than during the previous simulation?

No they are not.

Question 6: Why do you get different results when you run the same simulation several times?

The results differ since the possibility of adaptation varies. In smaller populations, the genetic drift that occurs is random.

Simulation 1-2

Now let's go back to the input window and **reduce the population size to N=100** (don't change the other parameters). Run the simulation.

Question 7: Did you observe the same number of fixed alleles compared to previous conditions?

No, there were 4 fixed alleles this time and one lost allele. (Black, red, blue, green – and pink respectively)

Question 8: Based on the two previous simulations, what would be your hypothesis regarding the effect of the population size on genetic drift:

Lower population size causes more alleles to become fixed. With a larger population, the chance of all individuals having allele that becomes fixed is smaller.

To test your hypothesis, run the simulation 10 times with 6 loci each time and write down the number of alleles that have been either lost or fixed after **100 generations** with the following population size: **N=25**, **75** and **150**.

You will use a quick statistical test to help you to compare your results for each population size: the 95% confidence interval. The 95% interval tells us that we can be 95% confident that the average is located between its lower and upper limits. Thus if the 95% intervals limits of two averages don't overlap, it is a good indication the averages are actually different. Use your calculator and/or Excel to calculate the 95% CI:

1- Calculate the average number of fixed allele ($\overline{m}$) for each population size.
2- Calculate the standard deviation (SD) of the mean by choosing the **STDEV** function in excel (**your TA will demonstrate how to do that in the live session**).
3- Calculate the standard error (SE) using this formula:

$$SE = \frac{SD}{\sqrt{n}}$$

where *SD* represents the standard deviation and *n* the sample size (number of measurements).

4- The 95% confidence interval is centered on the mean value and has two limits:

Upper limit L1= $\overline{m} + 2 \times SE$

Lower limit L2= $\overline{m} - 2 \times SE$

Enter your results in table 5 below.

Table 5: 95% Confidence interval results:

Population size	Average number of fixed alleles ($\overline{m}$)	Standard error (SE)	95% interval L1	95% interval L2
N=25	4.7	0.260342	5.220684	4.179316
N=75	1.8	0.326599	2.453198	1.146802
N=150	0.3	0.213438	0.726876	-0.126875

Question 9: Was your hypothesis correct? (Briefly explain why)

Yes, when the population was smaller (25), the number of fixed alleles was much higher, whereas with a larger population (150), there were very few fixed alleles.

Simulation 1-3

Set the population size to N=10 and run the simulation for **50 generations**. Try different values of p (initial frequency) and run the simulation several times for each value tested. Count the number of fixed alleles, and the number of lost alleles.

Question 10: What is your conclusion regarding the effect of initial frequency on allele fixation in a small population?

The initial frequency on a small population heavily dictates whether or not the allele will become fixed or lost. When the initial frequency is small(closer to 0), there is a higher chance of the allele becoming lost, whereas when the initial frequency is large (closer to 1), the chance of the allele becoming fixed is higher.

III- Drift and selection

In previous simulations, we observed how genetic drift can affect a population by changing allele frequencies. We will see now what is happening if we introduce a

new parameter in our simulation: **selection**. Selection is controlled in the simulation by assigning relative fitness values to the different genotypes.

The relative fitness of each genotype is the combination of the survival rate until reproductive age and the reproductive success (=how many offspring they produce) of individuals of a certain genotype relative to the **maximum** survival and reproductive rate of other genotypes in the population.

In summary, the relative fitness of a genotype is indicative of the success of individuals with this genotype to successfully reproduce and transmit their genotype.

In the real world, both **selection and genetic drift** act simultaneously. The goal of this exercise is to observe the effect of selection on allele frequency and the combined effect of both genetic drift and selection.

Close the previous window(s) and now select Model>Mendelian Genetics>Drift and Selection.

Simulation 2-1

In the input window, enter the following parameters:

Set the fitness values to w_{AA}**=0.8,** w_{Aa}**=1 ,** w_{aa}**=0.8**, population size to **N=500**. Start with an initial frequency **p=0.5** and run the simulation for **100 generations**.

Question 11: What do you observe? Has one of the alleles been fixed?

None of the alleles have become fixed, the trend remains consistently in the middle of the graph.

Simulation 2-2

Change the population size to **N=250**, then to **N=50**

Question 12: What do you observe with these parameters? Is there any change in the probability of the allele to be fixed? How would you measure the genetic drift strength?

When changing the population size, the probability of an allele to become fixed remains the same, it does not become fixed. In order to measure the genetic drift strength, we would analyze the variation from the middle (the 0.5 initial frequency) in either allele directions.

Simulation 2-3

Now let's give a selective advantage to one of the genotypes and set fitness values to w_{AA}=**0.6,** w_{Aa}=**0.6 ,** w_{aa}=**0.55** in a population of **300 individuals** and an initial frequency **p=0.5**

Question 13: Which allele possesses a selective advantage in this case?

The dominant allele (AA) would posses a selective advantage.

Question 14: Run the simulation for **500 generation** several times (at least 10 times). What do you observe?

The allele becomes fixed well before 500 generations is reached, always the wAA allele becoming fixed.

Simulation 4

Change the population size to 200, 100, 50, 25, 10 and run the simulation several times (at least 10 times for each population size).

Question 15: What do you observe when population size decreases?

When the population size decreases, the probability of the allele (wAA) becoming fixed becomes smaller. While still trending towards becoming fixed, more often than not, the allele does not become fixed.

Question 16: What are the evolutionary forces in action?

The evolutionary forces in action are random genetic drift and natural selection.

Question 17: What can you tell about the relative strength of each of these evolutionary forces in populations of various sizes?

Previously, when testing the effect of a smaller population size, the smaller it was, the more effect of genetic drift, whereas it is the opposite in this scenario. When decreasing the population, the amount of times the allele became fixed became less, meaning that selection is a much stronger evolutionary force than genetic drift.

This exercise (part II) was modified from a laboratory proposed by Dr J. Brown at Grinnel College in 1998.