Random Genetic Drift & Gene Fixation

Arie Zackay

July 31, 2007

Abstract

Random Genetic Drift is one of the evolutionary forces that effects the distribution of alleles and changes their frequencies in populations. The main idea behind this theory is that genetic changes can happen randomly. This can be explained through the Bottleneck effect. There are two main factors that have an influence on genetic drift, the first one is the Effective Population Size N_e , which creates different genetic drift outcomes within different population sizes. The second factor that influences the rate of genetic drift is the selective advantage of a drifted allele. With these factors it is possible to calculate the time and probability of an allele fixation in any chosen population.

1 Introduction

In order to understand the phenomenon known to us as *Random Genetic Drift*, we must first of all define the meaning of *Population Genetics*. *Population Genetics* is the study of the changes in distribution of alleles in a certain population, considering the influence of the following evolutionary forces: *natural selection*, *genetic drift*, *mutation*, *gene flows* and *selective mating*. The main goals of this field of study is to determine how the frequency of a mutant gene will change over time under the influence of evolutionary forces, to examine how genetic variability is maintained, and to establish the probability that a new mutant completely replaces existing variants of genes, a phenomenon called *Fixation*. There are a few instigators that can cause changes in the evolution of a population. Such instigators might include:

- It can be caused by a mutation with a genetic advance.

- Environmental changes like a volcano event, or other events, like the hunting of a specific species, which drastically changes the amount of members of this species and therefore the distribution of alleles in this population, eventually changes its evolution.

In this essay I will clarify the *Random Genetic Drift Theory* and the main factors that influence the power of this drift, meaning the probability (known as *Fixation Probability*) and the duration of allele fixation (known as *Fixation Time*). In *Population Genetics*, *Fixation* occurs when every individual in a certain population has the same allele at a particular locus. The allele, for example a single point mutation or a whole gene, will be rare in the beginning, but it can spread through the population by processes of *Random Genetic Drift*

Genetic Drift—Bottleneck Effect

Parent population

Bottleneck (drastic reduction in population)

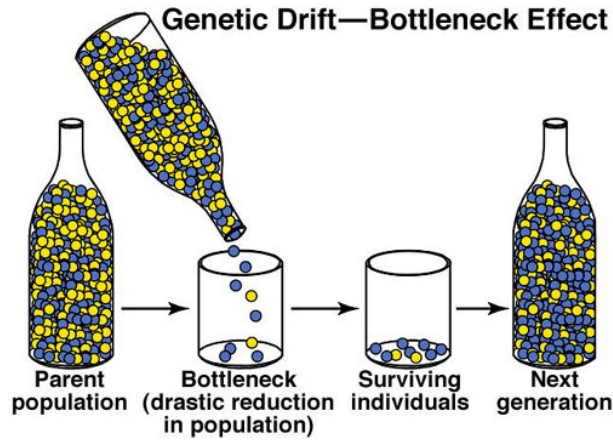Surviving individuals

Next generation

Figure 1: In a *Genetic Bottleneck*, allele frequency is changed due to a "crushing event". Before the tipping, there are more yellow balls in the bottle; After the tipping, or the "crushing event" in the life of the population, blue balls dominate the ball population in the bottle.

and/or *Positive Selection*. Once the frequency of the allele is at 100 %, meaning that every individual of this population has this gene at a particular locus, then this gene is fixed in the population. The opposite process can occur too - A certain allele can be diminished in a population through the *Random Genetic Drift*, or disappear altogether.

2 Random Genetic Drift

Unlike the *Natural selection Theory*, which prefers alleles with evolutionary advantage, the central idea of the *Random Genetic Drift Theory* is that the genetic distribution of a specific allele in a population happens by chance, a procedure we call *The Bottleneck Effect*. To simulate the idea of the Random Genetic Drift we imagine that there is a bottle with many little blue and yellow balls, and that the two colors stand for the two different types of the same allele (as shown in figure 1.). The bottle has a neck, and so the balls are able to spill out of it. It's important to remember that the proximity of the balls to the bottleneck is completely random, which means that when tipped over, random balls will fall out of the bottle. For the next step, let's tip the bottle on its open head for a few seconds, so that some of the balls are spilling out of it, an action that simulates a shocking event happening to a population.

A few factors determine the number of the balls flowing outside the bottle - The amount of time that the bottle was turned on its head, the width of the bottleneck, to name two of the most important ones. Due to the fact that the proportion between blue and yellow balls was random before the tipping of the bottle, and because a random number of them was spilling out of the bottle, the ratio between blue and yellow balls within the balls that spilled out, has now changed, depending on how many balls left the bottle, and on what color those
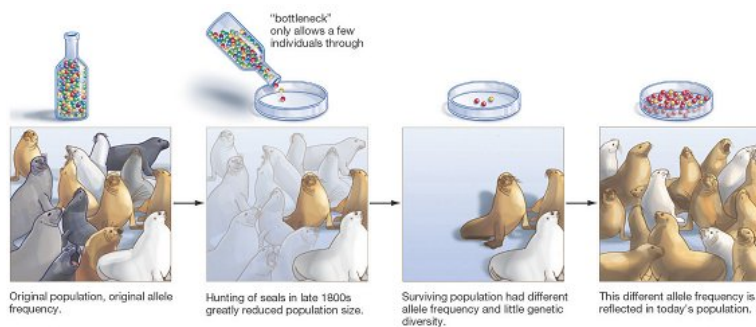
Figure 2: Before the massive hunting, these seals had a much wider diversity of allele's types then in their current population.

balls were. If there were only a few individuals who survived a certain event, (such a meteor hitting the earth or a major earthquake), meaning the width of the bottleneck was relatively narrow, and the bottle was tipped for a shorter period of time, then there is a chance that the ratio of yellow to blue balls that spilled out of the bottle is very different from the ratio that existed in the bottle before the tipping.

This is the core idea standing behind the *Random Generic Drift Theory*, which claims that those individuals in the population who left the bottle, did so by chance, without having any real advantage over the other members of their group. All they did was flow out of the bottleneck. If the situation was different, and some individuals would have survived because they were stronger or better evolutionarily speaking, and where one of the colors of balls would have an advantage over the other color, then we could have simulated the yellow balls, for example, to be large blue balls, and the blue balls would have the upper hand, because it would have been easier for them to leave the bottle. We would be left with more blue balls then yellow, after the tipping, this example correlates to the *Natural Distinction Theory*, and not to the *Random Genetic Drift*, which was displayed in the first example. A real example of a *Random Genetic Drift* is the change of alleles diversity and frequencies in the seal population in the North Pole (figure 2.). This population of seals was being hunted for the oil that could have been made from its fat. After the process of hunting the seal population dwindled to only a few hundreds. But then the hunting laws were changed in the early 20th century, and the hunting suddenly stopped. This allowed the population of seal to grow back to the impressive number of 100,000 individuals, but these individuals carry now a far smaller diversity of alleles, due to the *Random Genetic Drift*. They only carry the alleles of the seals that survived the hunting, and those alleles are random.

3 Effective Population Size

The *Effective Population Size N_e* is the number of breeding individuals in a population. It shows the same amount of dispersion of allele frequencies during the process of *Random Genetic Drift* or the same amount of inbreeding as the
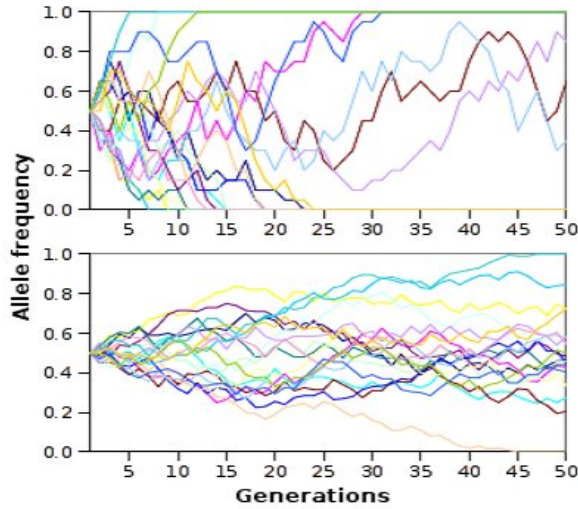
Figure 3: These two Plots demonstrate the drift of 20 alleles with initial frequency 0.5 for small (above) and small (bellow) population size (N=10 and 100 respectively). Every color represents one allele. In the big population there is only one allele fixation that occurs after about 45 generation. On the other hand there are relative large amount of alleles fixations in a short time (5-30 generations) in the small population. In general, alleles drift to fixation (frequency 1) significantly faster in smaller populations.

population under consideration. The *Effective Population Size* is usually smaller than the *Absolute Population Size N*.

Equation (1) calculates N_e, when there is a bias in the sex ratio:

$$N_e = \frac{4 * Nf * Nm}{Nf + Nm} \tag{1}$$

Where, N_f - number of females and N_m - number of males

According to Equation (1) a flock of Elephant seals (N=100), with 4 males and 96 females has the *Effective Population Size N_e* of about 15.

The influence of genetic drift is strongly related to the size of the parental population. This means that alleles drift is more rapid in small populations then in larger ones, as shown in figure 3.

To illustrate the Effect of Population Size we imagine that there are two populations of penguins (as shown in figure 4.). One population has 10,000 individuals, and the second has only 10 individuals. We are interested in following the distribution of a specific allele. Let's assume, that the frequency of this allele is 10% In both populations. In the bigger population there are 1000 penguins with this allele and in the smaller population only one penguin carries it. Now lets imagine that a big catastrophe occurs, and say that 50% of the each population survives (5000 und 5 respectively). But in the new bigger population
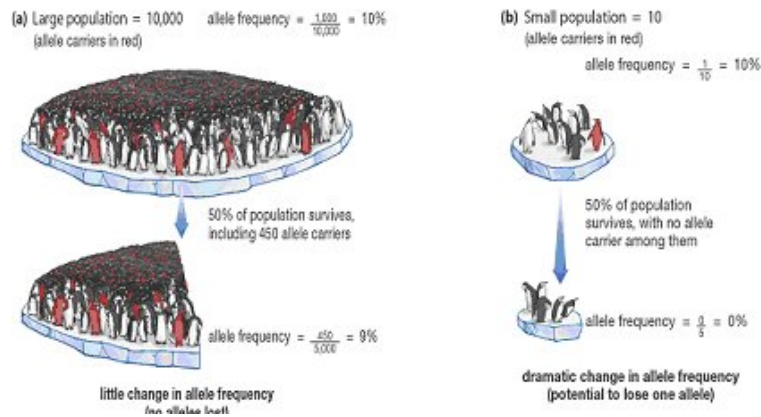
4

Figure 4: The way *Population Size* effects *Random Genetic Drift*. The observed allele is marked in red.

there are 450 individuals that have the specific allele, (9% of the new population), because randomly, out of the 5000 penguins lost in the catastrophe, 450 were carrying the allele. In the smaller population this allele was lost, because the only penguin carrying it was killed. This demonstrates how the size of the population effects the Rate of the Genetic Drift.

4 Fixation Time & Probability

Gene substitution is defined as the process whereby a mutant allele completely replaces the predominant or wild type allele in a population. In this process, a mutant allele arises in a population, and becomes fixed after certain number of generations. There are two values that are related to population genetic. The first one is the *Fixation Time*, means the time it takes for a new allele to become fixed in a certain population. The second term is the *Fixation Probability* witch defines the chance that a new mutant appears in individuals in the population. This two values depends strongly in the following parameters:

q the initial frequency of an allele
N_e effective population size
N absolute population size
S selective advantage

The *Fixation Probability* (P) is calculated with the following Equation :

$$P = (1 - e^{-(4N_e sq)/N})/(1 - e^{-4N_e s}) \qquad (2)$$

but for neutral mutation (without selective advantage):

$$P = 1/(2N) \qquad (3)$$

5

means that the probability of fixation is proportional to their frequency in the population. and for positive value of selective advantage (s) and large value of N:

$$P = 2s \tag{4}$$

The fixation time $\bar{t}$ is calculated with the following Equations:
For a mutation without selective advantage

$$\bar{t} = 4 * N_e * G \tag{5}$$

And for mutation with selective advantage,

$$\bar{t} = (2/s) * ln(2 * N_e) * G \tag{6}$$

Where G is the mean generation time.

There is a behavior difference between mutation with and without *Selective Advantage*. The time taken to reach *Fixation* is much slower, in case of no *Selective Advantage*, than a *Fixation* under mutation with *Selective Advantage*. To illustrate the difference between the two types of mutation, let us assume that there is a species with an *Effective Population Size* (N_e) of about 10^6 individuals and a mean generation time (G) of 2 years. Under these conditions, it will take a neutral mutation, on average 8 million years to become fixed in the population, but when we compare it to a mutation with *selective advantage* (s) of 1%, the mutation will become fixed in the same population in only 5803 years.

Literature

- Molecular Evolution. Wen-Hsiung Li, 1997, Chapter 2: Dynamics of genes in populations,Sinauer Associates, 31-52

- Development of Neutral and Nearly Neutral Theories. Tomoko Ohta & John H. Gillespie, 1996, theoretical population biology 49, 128-142, article no. 7

- www.wikipedia.com