Genetic drift

1. The Nature of Genetic Drift

To date, we have assumed that populations are infinite in size. This assumption enabled us to easily calculate the expected frequencies of alleles and genotypes without considering stochastic processes. Random changes in allelic frequencies arising from stochastic processes in populations of finite size. Today, we consider the effect of such processes on evolution.

Stochastic processes that contribute to genetic drift include:

1) independent assortment of alleles
2) random variation in survival
3) random variation in mating success
4) random variation in fecundity

Note that when we say that these processes are random, what we do not mean that they are not determined by some causal mechanisms. Rather, we mean that these process proceed randomly with respect to the identity of the alleles or genotypes.

The effect of genetic drift can be seen from a very simple example. Assume a population of a single individual who is heterozygous at a locus. The individual reproduces once and then dies. The probability that this individual produces a heterozygous offspring is 0.5. The probability that the next generation also contains a heterozygous individual is 0.5 x 0.5 (= 0.5^2). Thus, the probability that both alleles will be still remain in any generation in the future is 0.5^t, where t is the number of generations.

Just continuing this calculation for a few more generations shows that the both alleles are not expected to be maintained for long when stochastic processes occur. The probability that a particular allele will be fixed by genetic drift is equal to its frequency. However, the very fact that this process is stochastic means we cannot know for certain which allele will become fixed.

Today, we shall demonstrate the following:

1) Genetic drift tends to erode genetic variation of a population, and does so most rapidly when genetic variation is high and the population is small.

2) Genetic drift can easily eliminate new alleles (mutations) because they are present in low frequencies; this statement is even true for beneficial alleles when the population is small.

2. The Wright-Fisher Model

Both Wright and Fisher independently derived models of evolution by genetic drift that applies to diploid organisms with non-overlapping generations. In the Wright-Fisher model,
stochasticity arise from the independent assortment of alleles, but not from random variation in survival, mating success, or fecundity. In fact all assumptions are identical to those of the Hardy-Weinberg model, except that population size is assumed to be finite and constant.

The model treats genetic drift as a Markov Process; that is, the state of the population at time \( t+1 \) depends only on the state of the population at time \( t \) and a set of probabilities that define the change in state. The details of the model can be found in Rice (2004, pp. 74-76). The important result of the model is that the variance in allelic frequency \( (p) \) is expected to be

\[
\text{var}(p_{t+1}) = \frac{p_t(1 - p_t)}{2N} \quad [1]
\]

where \( N \) is the population size. Recall that this is the general function for calculating the variance in a binomial sampling process (see lecture on viability selection). From this result, we can conclude that an increase in sample size reduces the variance of the expected allelic frequency. Although \( p \) will increase or decrease with equal probabilities, the expected result after multiple generations is that \( p \) will equal 1 or 0. This expectation comes from the fact that \( p \) is fixed or eliminated at these values, and cannot change thereafter; in mathematical terms, we call these absorbing boundaries.

3. The Loss of Heterozygosity by Genetic Drift

We can calculate the loss of heterozygosity caused by genetic drift for our simple genetic system of two alleles at one locus.

Let

- \( G \) = probability that two alleles differing in origin are identical in state (i.e., probability of being AA given one copy of A)
- \( H = 1-G \) = probability that two alleles differing in origin also differ in state (i.e., probability of being heterozygous)

Then

\[
G_{t+1} = \frac{1}{2N} + (1 - \frac{1}{2N})G_t \quad [2]
\]

and

\[
H_{t+1} = 1 - \left[ \frac{1}{2N} + (1 - \frac{1}{2N})(1 - H_t) \right], \quad [3]
\]

which simplifies to

\[
H_{t+1} = (1 - \frac{1}{2N})H_t. \quad [4]
\]
Rearranging Eqn. 4, we can solve for the change in $H$:

$$\Delta H = H_{t+1} - H_t = \frac{H_t}{2N}.$$  \[5\]

From this result, we can draw two important conclusions. The first conclusion is logically consistent with the one drawn from the Wright-Fisher model (see above): an increase in the population size will decrease the rate at which heterozygosity will be lost. The second conclusion is a little more hard to understand; the higher the level of heterozygosity in a population, the faster it will be lost. This result is analogous to a previous result, which suggests that evolution by natural selection proceeds more rapidly in populations with more genetic variation (see lecture on viability selection). It also bolsters one’s confidence in the saying “if you have little to lose, you might as well take a chance.”

4. Mutation and Genetic Drift

While genetic drift erodes heterozygosity, mutation will increase it. We can calculate the equilibrium value of $H$ ($\hat{H}$) when both genetic drift and mutation occur.

Let

$$P_c = \text{the probability that in a given generation that alleles at a given locus will coalesce (i.e., the two alleles will become identical in state)}$$

Then, the equilibrium value of $H$ is the probability of a mutation of one of the two alleles divided by the probability that either mutation or coalescence will occur.

$$\hat{H} = \frac{2u}{2u + \frac{1}{2N}}.$$  \[6\]

which can be simplified to yield

$$\hat{H} = \frac{4Nu}{1 + 4Nu}.$$  \[7\]

One conclusion from that result is that population size greatly affects the maintenance of genetic variation caused by mutation. For example, if we assume a population size of $10^6$ and a mutation rate of $10^{-6}$, we should expect $H$ to be 0.80. However, if we reduce the population size to $10^4$, the same mutation rate would maintain $H$ at only 0.04.
5. Rates of Fixation & Neutral Theory

We have eluded to the Neutral Theory of Evolution (sometimes referred to as Non-Darwinian Evolution), which states that the majority of molecular evolution can be explained by the combined effects of mutation and drift. Now, let us examine one of the major factors that inspired this theory: the predicted and observed rates of allelic substitution.

In a population of size \(N\), the probability that a mutant allele will arise is simply equal to the probability of a mutation in any allele (\(u\)) multiplied by the number of alleles (\(2N\)). Assuming this new mutant is selectively neutral, the probability that it will become fixed is simply equal to its frequency; because the mutant is the only copy of that allele, its frequency is \(1/2N\). Thus, the rate of substitution of alleles by mutants is

\[
 k = 2Nu \cdot \frac{1}{2N},
\]

which simplifies to yield

\[
 k = u.
\]

Amazingly, the rate of fixation of a neutral mutant allele by genetic drift is not a function of population size! This phenomenon results because population size increase the rate of mutation but decreases the magnitude of genetic drift.

Kimura and Ohta (1971) noted that rates of fixation for proteins appear to be constant, and are very close to the rates expected to be result from mutation and genetic drift. These data were used as evidence for the Neutral Theory.

More recently, the rates of fixation of coding versus non-coding regions have been shown to differ, indicating that selection is an important mechanism of evolution. In light of these data, Ohta revised her view of Neutral Theory, claiming that mutations that affect protein structure are detrimental but mutations that are do not affect protein structure (silent mutations) are neutral. The role of selection versus the roles of mutation and drift in evolution remain controversial to this day.

6. References

