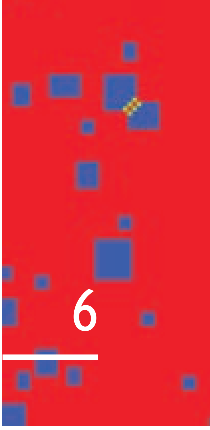

# 6

# FINITE POPULATIONS

**WE WILL NOW BEGIN** to analyze evolutionary dynamics in finite populations. The abundance of individuals is given by integers rather than by continuous variables. The resulting evolutionary dynamics are no longer described by deterministic differential equations, but require a stochastic formulation.

The best approach for studying a biological problem is to try a deterministic description first and then move to a stochastic analysis only when the deterministic one misses relevant aspects. Usually differential equations are easier to analyze and interpret than stochastic processes, but many important biological effects only arise in a stochastic context. One such effect is neutral drift. In this chapter, we will study neutral drift and constant selection in populations of finite size.

## 6.1 NEUTRAL DRIFT

Consider a population of fixed size $N$. There are two types of individuals, $A$ and $B$. They reproduce at the same rate. Therefore $A$ and $B$ are neutral variants with respect to selection. In any one time step, a random individual is chosen for reproduction and a random individual is chosen for elimination.

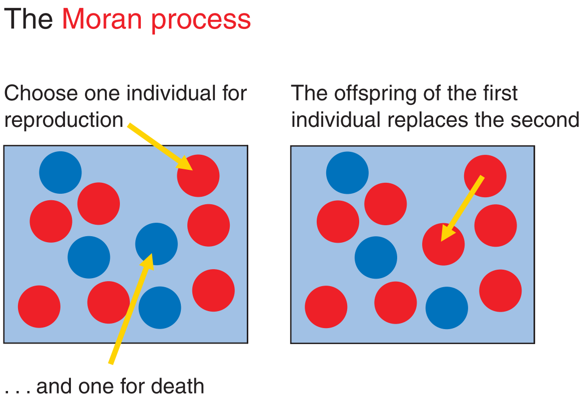

**Figure 6.1** The Moran process represents the simplest possible stochastic model to study selection in a finite population. In each time step, two individuals are chosen: one for reproduction and one for elimination. The offspring of the first individual will replace the second. Note that the two random choices could fall on the same individual; in this case, an individual will be replaced by its own offspring. The total population size is strictly constant.

We use sampling with replacement: the same individual could be chosen for reproduction and death. Reproduction occurs without mutation: $A$ produces $A$ and $B$ produces $B$.

This stochastic process is named after the Australian population geneticist P. A. P. Moran, who invented it in 1958. The feature that in each time step there is always one birth and one death event ensures that the total population size is strictly constant (Figure 6.1). The only stochastic variable is the number of $A$ individuals denoted by $i$. The number of $B$ individuals is $N - i$. Stochastic processes with one variable are much easier to investigate than stochastic processes with two or more variables.

The Moran process is defined on the state space $i = 0, \ldots, N$. The probability of choosing an $A$ individual (for birth or death) is given by $i/N$. The probability of choosing a $B$ individual is given by $(N - i)/N$. There are four possibilities of what could happen in any one time step.

(i) An $A$ individual could be chosen for reproduction and death. This event has probability $(i/N)^2$. After the event the number of $A$ individuals is the same as before; the variable $i$ has not changed.

(ii) A $B$ individual could be chosen for reproduction and death. This event has probability $[(N - i)/N]^2$. After the event the number of $B$ individuals is the same as before; the variable $i$ has not changed.

(iii) An $A$ individual could be chosen for reproduction and a $B$ individual for death. This event has probability $i(N-i)/N^2$. After the event there is one more $A$ individual than before; the variable $i$ has changed to $i+1$.

(iv) A $B$ individual could be chosen for reproduction and an $A$ individual for death. This event also has probability $i(N-i)/N^2$. After the event there is one less $A$ individual than before; the variable $i$ has changed to $i-1$.

The transition matrix, $P = [p_{ij}]$, determines the probabilities of moving from any one state $i$ to any other state $j$. $P$ is an $(N+1) \times (N+1)$ stochastic matrix. All entries are probabilities. The sum over each row is one. For our stochastic process, the transition matrix is given by

$$\begin{aligned} p_{i,i-1} &= i(N-i)/N^2 \\ p_{i,i} &= 1 - p_{i,i+1} - p_{i,i+1} \\ p_{i,i+1} &= i(N-i)/N^2 \end{aligned} \tag{6.1}$$

All other entries are zero. Therefore the transition matrix is tri-diagonal. This is the defining property of "birth-death" processes. In any elementary stochastic step, the state variable $i$ can only change by at most one (Figure 6.2).

For our particular birth-death process, we note that

$$p_{0,0} = 1 \qquad p_{0,i} = 0 \qquad \forall i > 0 \tag{6.2}$$

and

$$p_{N,N} = 1 \qquad p_{N,i} = 0 \qquad \forall i < N \tag{6.3}$$

The states $i = 0$ and $i = N$ are "absorbing states": once the process has reached such a state, it will stay there forever. The states $i = 1, \ldots, N-1$ are called transient. The process stays in the set of transient states only for some limited time. Eventually the population will consist of either all $A$ or all $B$ individuals. Although there is no selection, one of the two types will replace the other. Coexistence is not possible.

Since our stochastic process has two absorbing states, we can ask: starting in state $i$, what is the probability of reaching state $N$? In other words, given that

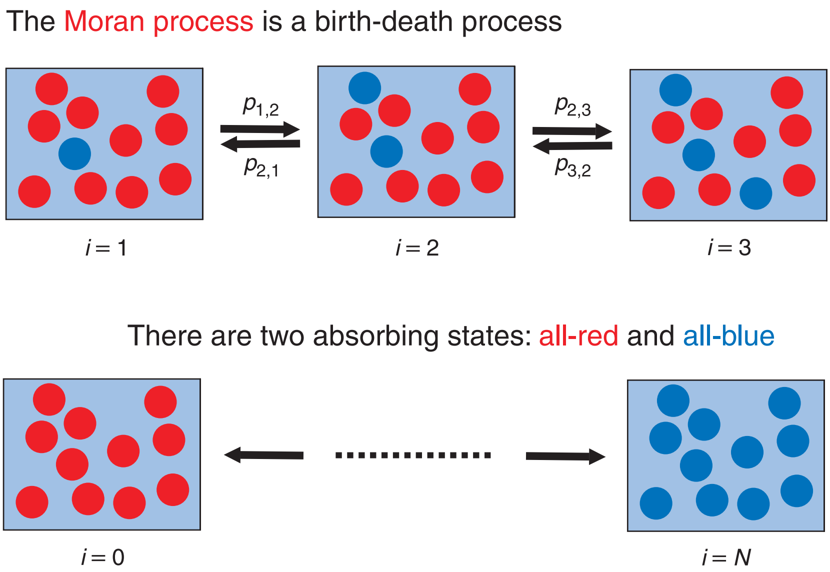

**Figure 6.2** The Moran process is a birth-death process. In each time step, the number, $i$, of blue individuals can only change by one at most. There are two absorbing states, $i = 0$ and $i = N$. In both cases, one type has taken over the entire population. No further change can occur (unless there is a new mutation).

we start with $i$ many $A$ individuals, what is the probability that eventually the whole population will consist of $A$ individuals?

Let us do a formal calculation, which will be generalized in the next section. Denote by $x_i$ the probability of ending up in state $N$ when starting from state $i$. The probability of ending up in state 0 when starting from state $i$ is given by $1 - x_i$, because there are no other absorbing states. We have

$$\begin{aligned} x_0 &= 0 \\ x_i &= p_{i,i-1}x_{i-1} + p_{i,i}x_i + p_{i,i+1}x_{i+1} \qquad \forall i = 1, \ldots, N-1 \\ x_N &= 1 \end{aligned} \tag{6.4}$$

The probability of being absorbed in state $N$ starting from $i$ is given by the sum of the following three terms: (i) the probability of going from $i$ to $i - 1$

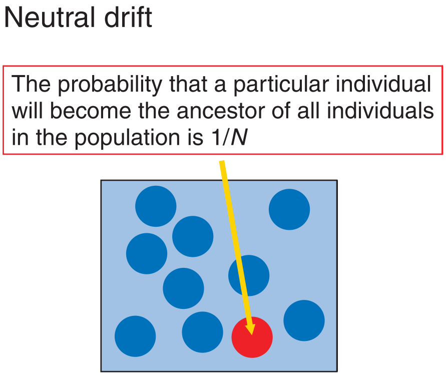

**Figure 6.3** In a finite population, if we wait long enough, the descendants of one particular individual will take over the entire population. If all individuals have the same fitness, then all individuals currently present in the population must have the same chance. Hence under neutral drift, the fixation probability (of the lineage) of any one individual is $1/N$.

multiplied by the probability of being absorbed from $i-1$; (ii) the probability of staying in $i$ multiplied by the probability of being absorbed from $i$; (iii) the probability of going from $i$ to $i+1$ multiplied by the probability of being absorbed from $i+1$. Thus we have a recursive equation for $x_i$. Note that $x_0 = 0$; from state 0 we can never reach state $N$. Moreover, we have $x_N = 1$; from state $N$, we will certainly reach state $N$, because we are already there.

Since $p_{i,i-1} = p_{i,i+1}$ and $p_{i,i} = 1 - 2p_{i,i+1}$, the solution of the linear system (6.4) is

$$x_i = i/N \qquad \forall i = 0, \ldots, N \tag{6.5}$$

The result is obvious. Since all individuals reproduce and die at the same rate, the chance that a particular individual will generate a lineage that will inherit the whole population must be $1/N$ (Figure 6.3). If there are $i$ many $A$ individuals, then the chance that one of them will make it (as opposed to one of the $N-i$ many $B$ individuals) is simply $i/N$.

For each trajectory of our stochastic process, there are only two final possibilities: the trajectory reaches either state 0 or state $N$. The probability of being absorbed in 0 is one minus the probability of being absorbed in $N$. The probability of ending up in all-$B$ when starting with $N-i$ many $B$ individuals is given by $(N-i)/N$.

## 6.2 BIRTH-DEATH PROCESSES

Let us now perform the same calculation for a general birth-death process. A birth-death process is a one-dimensional stochastic process on a discrete state space, $i = 0, \ldots, N$. In each stochastic event, the state variable $i$ can either remain unchanged or move to $i - 1$ or $i + 1$.

Denote by $\alpha_i$ the probability of a transition from $i$ to $i + 1$. Denote by $\beta_i$ the probability of a transition from $i$ to $i - 1$. We have $\alpha_i + \beta_i \leq 1$. The probability of remaining in state $i$ is given by $1 - \alpha_i - \beta_i$. Consider a birth-death process where $i = 0$ and $i = N$ are absorbing states. Therefore, we have $\alpha_0 = 0$ and $\beta_N = 0$. The transition matrix is of the form

$$P = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ \beta_1 & 1-\alpha_1-\beta_1 & \alpha_1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \beta_{N-1} & 1-\alpha_{N-1}-\beta_{N-1} & \alpha_{N-1} \\ 0 & 0 & 0 & \ldots & 0 & 0 & 1 \end{pmatrix} \tag{6.6}$$

Denote by $x_i$ the probability of reaching state $N$ when starting from $i$. Clearly, $1 - x_i$ denotes the probability of reaching state 0 when starting from state $i$. In analogy to (6.4) we have

$$\begin{aligned} x_0 &= 0 \\ x_i &= \beta_i x_{i-1} + (1 - \alpha_i - \beta_i)x_i + \alpha_i x_{i+1} \qquad i = 1, \ldots, N-1 \\ x_N &= 1 \end{aligned} \tag{6.7}$$

In vector notation, we can write

$$\vec{x} = P\vec{x} \tag{6.8}$$

The absorption probabilities are given by the right-hand eigenvector associated with the largest eigenvalue, which is one, because $P$ is a stochastic matrix.

Let us introduce the variables

$$y_i = x_i - x_{i-1} \qquad i = 1, \ldots, N \tag{6.9}$$

Note that $\sum_{i=1}^{N} y_i = x_1 - x_0 + x_2 - x_1 + \cdots x_N - x_{N-1} = x_N - x_0 = 1$. Let $\gamma_i = \beta_i/\alpha_i$. From equation (6.7) we find $y_{i+1} = \gamma_i y_i$. Therefore we have $y_1 = x_1$, $y_2 = \gamma_1 x_1$, $y_3 = \gamma_1\gamma_2 x_1$, and so on. If we sum all these expressions we get

$$x_1 = \frac{1}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} \gamma_k}. \tag{6.10}$$

From

$$x_i = x_1 \left( 1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} \gamma_k \right), \tag{6.11}$$

we obtain

$$x_i = \frac{1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} \gamma_k}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} \gamma_k}. \tag{6.12}$$

Consider a population of one $A$ individual and $N-1$ $B$ individuals. The probability that $A$ takes over the whole population is called the fixation probability of $A$. We denote this probability by $\rho_A$. The idea is that a homogeneous population of $B$ has produced a mutant of type $A$. We are interested in the probability of this mutant becoming fixed in the population, which means that it generates a lineage that takes over the whole population. Similarly, we denote by $\rho_B$ the probability that a single $B$ individual takes over a population that contains $N-1$ $A$ individuals. The fixation probabilities of $A$ and $B$ are respectively given by $\rho_A = x_1$ and $\rho_B = 1 - x_{N-1}$. Therefore, we have

$$\begin{aligned} \rho_A &= \frac{1}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} \gamma_k} \\ \rho_B &= \frac{\prod_{k=1}^{N-1} \gamma_k}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} \gamma_k} \end{aligned} \tag{6.13}$$

Note that the ratio of the fixation probabilities is simply given by the product over all $\gamma_i$

$$\frac{\rho_B}{\rho_A} = \prod_{k=1}^{N-1} \gamma_k. \tag{6.14}$$

If $\rho_B/\rho_A > 1$, then it is more likely that a single $B$ mutant becomes fixed in an $A$ population than the other way round.

The fixation probabilities that we have derived in this section hold for any selection scenario between $A$ and $B$, including neutral drift, constant selection, and frequency-dependent selection.

## 6.3 RANDOM DRIFT WITH CONSTANT SELECTION

Let us now study the same process as before, but assume that $A$ has fitness $r$ while $B$ has fitness 1. If $r > 1$, then selection favors $A$. If $r < 1$, then selection favors $B$. If $r = 1$, we are back to neutral drift. The fitness difference can be included in our process by modifying the probabilities of choosing $A$ or $B$ for reproduction.

The probability that $A$ is chosen for reproduction is given by $ri/(ri + N - i)$. The probability that $B$ is chosen for reproduction is given by $(N - i)/(ri + N - i)$. The probability that $A$ is chosen for elimination is $i/N$. The probability that $B$ is chosen for elimination is $(N - i)/N$. For the transition matrix, we obtain

$$\begin{aligned} p_{i,i-1} &= \frac{N-i}{ri+N-i}\frac{i}{N} \\ p_{i,i} &= 1 - p_{i,i+1} - p_{i,i+1} \\ p_{i,i+1} &= \frac{ri}{ri+N-i}\frac{N-i}{N} \end{aligned} \tag{6.15}$$

All other elements of the matrix are zero. Again we want to calculate the fixation probability, $x_i$, to reach state $N$ starting from state $i$. Note that

$$\gamma_i = \frac{p_{i,i-1}}{p_{i,i+1}} = \frac{1}{r}. \tag{6.16}$$

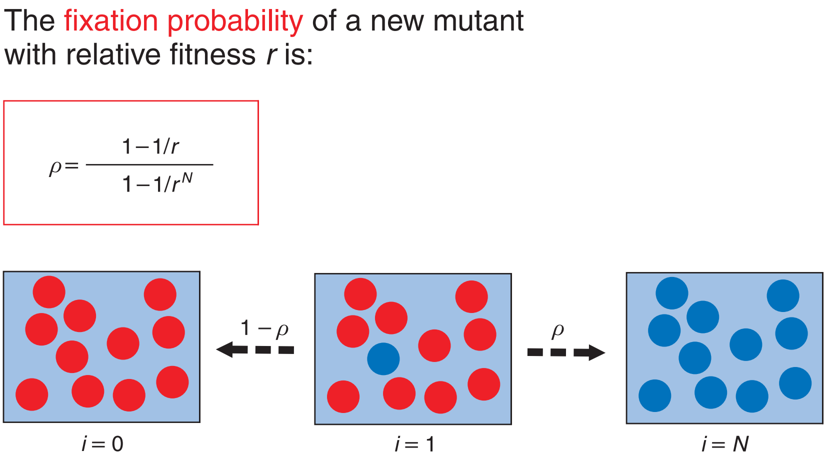

**Figure 6.4** Suppose that a new mutant (blue) arises in a population and has relative fitness $r$. The lineage starting from this mutant can either become extinct or take over the whole population. The probability that the mutants will take over the population is given by the "fixation probability" $\rho = (1 - 1/r)/(1 - 1/r^N)$.

Therefore the probability of being absorbed in state $N$ when starting in state $i$ is given by

$$x_i = \frac{1 - 1/r^i}{1 - 1/r^N}. \tag{6.17}$$

The fixation probability of a single $A$ individual in a population of $N - 1$ $B$ individuals (Figure 6.4) is

$$\rho_A = x_1 = \frac{1 - 1/r}{1 - 1/r^N}. \tag{6.18}$$

The fixation probability of a single $B$ individual in a population of $N - 1$ $A$ individuals is

$$\rho_B = 1 - x_{N-1} = \frac{1 - r}{1 - r^N}. \tag{6.19}$$

The ratio of the two fixation probabilities is given by

$$\frac{\rho_B}{\rho_A} = r^{1-N}. \tag{6.20}$$

For the fixation probability of an advantageous $A$ mutant, $r > 1$, in a large population, $N \gg 1$, we have the useful approximation

$$\rho_A = 1 - 1/r. \tag{6.21}$$

Even in the limit of an infinitely large population, $N \to \infty$, there is no guarantee that an advantageous mutant will take over. This is an important difference between deterministic and stochastic models of evolution. In a deterministic setting, an advantageous mutant is certain of victory regardless of how small $r$ is as long as $r > 1$. In a stochastic setting, the chance of extinction always remains, no matter how large the population size $N$.

Let us consider some numerical examples for a population of size $N = 100$:

A 100% selective advantage, $r = 2$, leads to $\rho = 0.5$.

A 10% selective advantage, $r = 1.1$, leads to $\rho = 0.09$.

A 1% selective advantage, $r = 1.01$, leads to $\rho = 0.016$.

For a neutral mutant, $r = 1$, we have $\rho = 1/N = 0.01$.

A 1% selective disadvantage, $r = 0.99$, leads to $\rho = 0.0058$.

A 10% selective disadvantage, $r = 0.9$, leads to $\rho = 0.000003$.

We can also ask how often must a mutant with relative fitness $r$ arise, before it has a probability 1/2 of taking over the population. The answer is $m = -\log 2/\log(1-\rho)$. Again we consider a population of size 100. A mutant with $r = 2$ must arise once. A mutant with $r = 1.1$ must arise 7 times, while a mutant with $r = 1.01$ must arise 44 times. A neutral mutant, $r = 1$, must arise 69 times. Disadvantageous mutants with $r = 0.99$ or $r = 0.9$ must arise 119 times and about 234,861 times, respectively.

## 6.4 THE RATE OF EVOLUTION

Imagine a population of $N$ reproducing individuals. All individuals are of the same type, $A$. Very rarely a mutation occurs which produces an individual of type $B$. Assume that mutation happens during reproduction. The mutation rate $u$ represents the probability that the reproduction of $A$ results in $B$. Thus $1-u$ is the probability that reproduction of $A$ occurs without mutation. For how long do we have to wait until a population of $N$ $A$ individuals will produce a $B$ mutant? The rate at which a $B$ mutant is being produced by the population is $Nu$. The time until the $B$ mutant arises is exponentially distributed with mean $1/(Nu)$.

Suppose type $B$ has a relative fitness $r$ compared to fitness 1 of type $A$. Thus the probability that the new $B$ mutant will take over the population is given by

$$\rho = \frac{1-1/r}{1-1/r^N}. \tag{6.22}$$

The rate of evolution from all-$A$ to all-$B$ is given by

$$R = Nu\rho. \tag{6.23}$$

The rate at which a $B$ mutant is being produced is $Nu$. The probability that a $B$ mutant reaches fixation is $\rho$. Hence the rate of transition from all-$A$ to all-$B$ is the product of these two terms.

If $B$ is neutral, then $\rho = 1/N$ and the rate of neutral evolution is given by

$$R = u. \tag{6.24}$$

The rate of neutral evolution is independent of the population size and simply equals the mutation rate. This important result was derived by Motoo Kimura.

This insight is at the center of the so-called neutral theory of evolution. According to the neutral theory, the majority of mutations that can be observed—for example, when comparing genetic sequences of humans and chimpanzees—should be neutral. Advantageous mutations are extremely unlikely to occur in genes that have been optimized for millions of generations in the ancestors of these species. Deleterious mutations cannot be observed

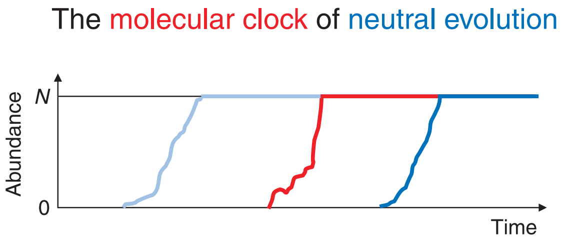

**Figure 6.5** The rate of producing neutral mutants is $Nu$, where $N$ is the population size and $u$ is the mutation rate. The fixation probability of a neutral mutant is $1/N$. Therefore the rate of neutral evolution is $R = Nu/N = u$: the rate of evolution equals the mutation rate. The population size cancels out. This relationship holds even if the total population size is changing over time. If the mutation rate is constant, then neutral mutants accumulate at a constant rate, giving rise to a "molecular clock." The figure shows three mutations that succeed in taking over the population. For each mutation that becomes fixed there are on average $N$ unsuccessful attempts.

because they would be eliminated with high probability. Hence the majority of observed mutations in any phylogeny should be neutral (or nearly neutral).

The rate of accumulating neutral mutations is simply given by the mutation rate and is independent of the population size and fluctuations in the population size. If the mutation rate depends mostly on the accuracy of DNA replication, which in turn is performed by a well-optimized system of enzymes that has not changed much in all eukaryotes, then the rate of evolution is constant. The neutral theory provides a "molecular clock" (Figure 6.5).

There was once a heated controversy between the supporters and the opponents of the neutral theory of evolution. The extreme neutralist would say: all observable mutations, say between human and chimp, are neutral; hence neutral variation alone can explain the evolutionary divergence between these two species; adaptation is unimportant. The extreme adaptationist would say: neutral evolution is unimportant; it is not even evolution, because it represents random variation without adaptation; evolution always requires adaptation.

The resolution of the controversy is obvious for those who were never involved. Most molecular variation is neutral. Therefore neutrality is an excellent model for studying genetic variation. Neutrality is often a good assumption for building mathematical tools that help to calculate phylogenetic relationships among species. Certainly the overwhelming majority of mutations that have been fixed in populations since the origin of life have been neutral. Very occasionally, however, advantageous mutations come into play. These mutations are extremely important for determining the trajectory of evolution.

## SUMMARY

- Evolutionary dynamics in populations of finite size require a stochastic theory.
- The Moran process is a birth-death process, which describes evolution in finite populations.
- If a finite population contains several different types, then without mutation eventually all but one type will be extinct. This is the case even if all types have the same fitness. This principle is called “neutral drift.”
- In a population of size $N$, a neutral mutant will reach fixation with probability $1/N$.
- A mutant with relative fitness, $r$, will reach fixation with probability $\rho = (1 - 1/r)/(1 - 1/r^N)$.
- The rate of evolution is given by the product of the population size, $N$, the mutation rate, $u$, and the probability of fixation $\rho$.
- The rate of neutral evolution is given by the mutation rate, $u$, and is independent of the population size (because $\rho = 1/N$).
- If the mutation rate is constant, then neutral mutations accumulate in genomes at a constant rate. This effect is called the “molecular clock.”
- The neutral theory of evolution recognizes the fact that the majority of mutations that become fixed in genomes are neutral.