POPULATION GENETICS

Brian Charlesworth
University of Edinburgh

GLOSSARY

allele frequency The frequency of a variant form of a genetic locus within a population.
genetic drift Evolutionary change caused by random sampling of genotype frequencies in a finite population.
genotype The state of an individual with respect to a defined genetic locus or set of loci.
heritability The proportion of the variance in a trait that is due to additive genetic effects.
inbreeding Matings between close relatives.
mutation rate The frequency with which new mutations arise per generation.
neutral mutations Mutations whose effects on fitness are either nonexistent or so small that their fate is controlled by genetic drift rather than selection.
phenotype The state of an individual with respect to a trait of interest.
polymorphism The existence at intermediate frequencies of two or more variants at a locus within a population.
selection The differential survival or reproductive success of individuals, associated with differences in phenotype or genotype.

DARWIN'S THEORY OF "DESCENT WITH MODIFICATION" implies that all of the stupendous diversity of life on Earth is ultimately traceable to genetic diversity within populations. The study of the nature and causes of within-population variation, and of the mechanisms by which it is transformed into differences between populations over space and time, is the province of population genetics. The subject involves both theoretical modeling of evolutionary processes, based on knowledge of the mechanisms of inheritance, and the testing of these models using data on variation and evolution in natural and artificial populations.

I. VARIATION WITHIN POPULATIONS

A. Types of Phenotypic Variation

Since evolutionary change depends on the existence of genetic variation within populations, measurement of the extent of such variation is crucial (Lewontin, 1974). Variation at the level of externally visible phenotypes can be divided into three categories.

1. Discrete Variation

This involves traits which can be divided into a small number of discrete categories, such as eye color in hu-

Encyclopedia of Biodiversity, Volume 4

mans or shell color and pattern in the land snail *Cepaea* (Ford, 1975). It is often controlled by one or a few genes, and usually it involves relatively superficial traits such as color patterns. Only a relatively small proportion of the phenotypic variation of interest to evolutionists is of this kind.

2. Quantitative Variation

Quantitative variation is all-pervasive. This can involve either meristic traits, such as bristle number in *Drosophila*, in which there is a large number of discrete categories, or continuously varying metrical traits such as body size. Variation in typical quantitative traits is known to be under the joint control of environmental effects, accidents of development, and sets of genes whose individual effects are small relative to the total range of variation in the traits (Falconer and Mackay, 1996). Statistical methods that utilize the degree of resemblance between close relatives enable the determination of the proportion of the total phenotypic variation that is contributed by additive genetic causes—the heritability, which controls the rate of response to selection on a trait (see Section II,D). Heritabilities for quantitative traits typically are between 20 and 80%, corresponding to the fact that artificial selection is highly effective in changing the mean value of almost every trait that has been examined (Falconer and Mackay, 1996).

3. Concealed Variability

A more subtle form of phenotypic variation is concealed variability, i.e., variability that is only exposed when homozygous genotypes are produced by close inbreeding. This is responsible for the increased variation among inbred lines when a set of such lines is made from a random-bred base population and for inbreeding depression, which is the decline in the mean values of fitness-related traits such as viability and fecundity with inbreeding (Falconer and Mackay, 1996). Both of these phenomena reflect, at least in part, the widespread occurrence of recessive or partially recessive rare alleles in random-mating populations (see Section II,F), whose phenotypic effects are only fully exposed when they are made homozygous and are therefore not evident in randomly mating populations.

In *Drosophila*, special breeding methods involving the use of genetically marked chromosomes with inversions that suppress crossing over have shown that up to 50% of haploid genomes carry recessive lethal genes. These contribute about half the inbreeding depression manifested when fully homozygous genotypes are produced; a similar magnitude of inbreeding depression is caused by genes of small effect (Crow, 1993). The net fitness of fully inbred *Drosophila* is only a few percent of that of outbred flies. Even more extreme effects of complete inbreeding are likely in vertebrates, which have much larger genomes. The deleterious fitness effects of inbreeding have probably played a major role in promoting the evolution of mechanisms of inbreeding avoidance, such as the self-incompatibility loci of flowering plants.

4. Interpreting Phenotypic Variation

The previously mentioned basic facts were established by the early 1950s and led to an active debate concerning the causes of natural variation (Lewontin, 1974). In one view, championed by H. J. Muller, variation is mostly due to rare deleterious alleles maintained by mutation pressure at a large number of loci; the coexistence of alleles at a locus at intermediate frequencies (polymorphism) is characteristic of only a small number of loci. In the other view, advocated by Dobzhansky, polymorphism is the norm, and it reflects variation that is actively maintained by selection. In the absence of any means of identifying loci without the prior existence of genetic variability, no unbiased survey of the extent of genetic variation at individual loci was possible with the methods of classical genetics, and so this question could not be answered.

B. Molecular Variation

1. Protein Electrophoresis

The previously mentioned situation was transformed by the development of molecular genetics. Gel electrophoresis of soluble enzymes and proteins provides a rapid and simple method for surveying populations for variants affecting the structure of a large number of different proteins and hence genes (Lewontin, 1974; Hartl and Clark, 1997). The results of such surveys reveal that a high fraction of loci coding for soluble proteins are polymorphic in the sense of having at least one rare variant whose frequency exceeds 5%; the average individual from a randomly mating population is typically heterozygous for a significant fraction (several percent) of such loci. Despite some biases in the methodology, particularly the inability of electrophoresis to detect many types of amino acid sequence changes and the restriction of the method to soluble proteins, it is clear that protein polymorphism is not an exceptional situation.

2. Measurement of Variation in DNA Sequences

The introduction of recombinant DNA technology has meant that population geneticists can now study variation at the level of the nucleotide sequence. Surveys of within-species DNA sequence variation of nuclear genes have been most intensively carried out in *Drosophila*, but comparable results are emerging from other species (Li, 1997). The basic conclusion is that variants due to single nucleotide changes are the most abundant source of variation in natural populations. For silent substitutions in third coding positions, which do not change the amino acid sequence, and for changes in introns and flanking sequences, the probability that two randomly chosen alleles from a *Drosophila* population differ at a given site (the nucleotide site diversity) is typically of the order of 1% or a few percent, depending on the species. The level of this type of variability is about 10 times higher in the bacterium *Escherichia coli* and one-tenth as high in humans (Li, 1997). For most genes, diversity is much lower for replacement changes, which alter the amino acid sequence. In addition to single nucleotide polymorphisms, DNA variability is contributed by small insertions and deletions of sets of nucleotides and by insertions of tranposable elements, mostly in noncoding regions. Other types of variability include variation in the sizes of tandem arrays of microsatellite and minisatellite loci, which are often highly polymorphic and provide useful genetic markers (Bruford and Wayne, 1993; Hartl and Clark, 1997). The density of such loci, however, is low in relation to the total size of the genome. The total level of variability at the level of DNA sequences is about two orders of magnitude greater than that revealed by electrophoresis because of the high degree of variability at silent and noncoding nucleotide sites relative to replacement sites.

3. Interpreting DNA Sequence Variation

Although variation at the level of DNA sequences must underly heritable phenotypic variation, it is difficult to relate the two, except for intensively studied human genetic diseases. The abundance of variation in both protein and DNA sequence might seem to vindicate Dobzhansky's view of the causes of natural variation. However, it is likely that much of the silent and noncoding variability is close to neutrality with respect to effects on fitness (see Section III,C). Nevertheless, there is a real possibility that selection also frequently influences variation and evolution in protein and regulatory sequences (Hartl and Clark, 1997; Li, 1997). The role of deterministic forces in variation and evolution within populations will thus be considered next.

II. DETERMINISTIC POPULATION GENETICS

A. Allele and Genotype Frequencies

If we focus on a given nucleotide position, the basic descriptor of the state of a population is the set of frequencies of the four alternative states, A, T, G, and C. If recombination within a gene is ignored, we can consider the set of all nucleotide sequences observed at a locus as alternative alleles, whose frequencies characterize the state of the population with respect to this locus. Mendelian inheritance implies that this state is not changed in the absence of evolutionary forces (the occurrence of intragenic crossing over and gene conversion at low frequencies means that in practice this is a good approximation rather than an exact description). This is enshrined in the Hardy–Weinberg principle, which states that the frequencies of diploid genotypes in a random mating population with a set of n alleles with frequencies $p_1, p_2, \cdots, p_n$ rapidly reach equilibrium values given by the multinomial expansion of $(p_1 + p_2 + \cdots, p_n)^n$. The importance of this result is that existing natural variation is preserved by Mendelian inheritance. This removes Darwin's difficulties over the rapid loss of variation under blending inheritance, which led him to adopt a theory of the inheritance of acquired characteristics for which there is no empirical foundation (Fisher, 1930).

B. Mutation

1. Types of Mutations and Their Rates

The ultimate source of natural variation is known to be spontaneous mutation, defined as a heritable change in the genetic material and which occurs without reference to the adaptive utility of the phenotypic consequences of the change in question. The most abundant mutations are nucleotide substitutions, but small deletions and insertions due to slippage during DNA replication are relatively common as well (Drake *et al.*, 1998). Insertions of transposable elements, large deletions and duplications, duplications and deletions of entire chromosomes or haploid genomes (aneuploidy and polyploidy), and chromosome rearrangements such as inversions and translocations also occur and contribute to evolutionary changes in genome structure. Rates of spontaneous mutation in organisms with DNA genomes are extremely low due to the operation of complex enzymatic systems which repair lesions in DNA; the rates of nucleotide changes per site per cell generation

in DNA-based microbes are between 1×10^{-10} and 5×10^{-10}. Similar values apply to higher eukaryotes such as *Drosophila* and humans, but the rate per organism generation is between 10 and 50 times higher owing to the many cell divisions that occur during the production of germ cells. Per locus rates of mutation to alleles with major phenotypic effects are substantially higher—on the order of 10^{-5} per generation in *Drosophila*, mice, and humans (Drake *et al.*, 1998). This is not surprising given the large number of nucleotides in the coding and regulatory regions of typical loci. Rates of change in copy numbers in microsatellite and minisatellite loci in mammals are much higher, however, up to 10^{-3} per locus per generation (Bruford and Wayne, 1993). Mutation rates in viruses with RNA genomes are also extremely high due to their lack of repair mechanisms (Drake *et al.*, 1998).

2. Evolution under Mutation Pressure

Given the low rate of mutation at the nucleotide level in DNA-based organisms, the time scale of mutational change in the frequencies of the four alternative states at a nucleotide site is very large—on the order of 1 billion generations. Mutation pressure at this level is thus an extremely weak force and is easily opposed by other evolutionary factors. For most purposes, therefore, mutation can be regarded simply as a source of new variation and as unimportant as a cause of directed evolutionary change (Fisher, 1930). This statement needs to be qualified when the aggregate effects of mutations affecting a particular phenotype are considered; the numbers of loci affecting a single quantitative trait are sufficiently large that increases in variability due to mutation are detectable in stocks that are initially genetically uniform. The rate of increase in variance per generation is typically on the order of 10^{-3}, relative to the nongenetic variance in the trait (Falconer and Mackay, 1996). Given that fitness-related traits are affected by many genes, and that most phenotypic changes caused by mutations are harmful to the organism, there is a tendency for the mean values of fitness components to decline under mutation pressure when selection is relaxed by experimental manipulations (Crow, 1993; Drake *et al.*, 1998).

3. Mutation and Selection

These considerations imply that a relatively weak force of selection, far smaller than is measurable experimentally, can prevent the spread of deleterious mutations at a locus. This explains the fact that amino acid polymorphisms are usually much less frequent than silent or noncoding polymorphisms. The same observation applies to comparisons of sequences between different species (Kimura, 1983). A very important form of selection is thus purifying selection, whereby deleterious mutations are constantly being eliminated from the population (see Section II,F). Similarly, quantitative traits are often subject to stabilizing selection, such that individuals with extreme trait values are less fit than individuals with intermediate values (Falconer and Mackay, 1996). Variability in quantitative traits is maintained, at least partially, by a balance between the input of new variation by mutation and its elimination by stabilizing selection (Falconer and Mackay, 1996).

C. Selection at a Single Locus

1. The Basic Model

A large body of theory has been developed to describe the action of natural selection on Mendelian variation. In the simplest case, a single locus with a pair of alternative alleles, A and a, is postulated. A randomly mating, infinitely large population is assumed. Ignoring sex differences, the relative fitnesses of the three possible genotypes AA, Aa, and aa in a diploid can be written as 1, $1 - hs$, and $1 - s$, respectively, where s is the selection coefficient and h is the dominance coefficient. s measures the strength of selection ($s > 0$, when a is disfavored by selection); h measures the extent to which the fitness of the heterozygote is reduced by the presence of a. The fitnesses are assumed to be constant over time.

Fitness in this context is most easily understood in terms of viability selection in a population with discrete generations, where the relative probabilities of the three genotypes surviving from egg to adult are equivalent to the three fitnesses. In general, selection may involve many different aspects of the life history, especially female fecundity and male mating success. More elaborate models have been developed to study these cases, including extensions of the theory to populations with overlapping rather than discrete generations, but the basic conclusions are similar (Crow and Kimura, 1970; Ewens, 1979; Hartl and Clark, 1997).

If the frequencies of the two alleles in one generation are p and q, respectively, the change in frequency of A over one generation is

$$\Delta p = spq\{q + h(1 - 2q)\}/\overline{w}, \qquad (1a)$$

where $\overline{w} = 1 - 2pq\ hs - q^2s$ is the population's mean fitness.

An interesting equivalent form, from Wright

(1977), is

$$\Delta p = \frac{1}{2} pq \frac{d \ln \overline{w}}{dp}. \qquad (1b)$$

Equation (1b) implies that gene frequency change is in the direction of the gradient of mean fitness with respect to gene frequency. A more detailed analysis of the dynamics of a single locus with an arbitrary number of alleles, and arbitrary but constant fitnesses, shows that mean fitness increases monotonically as allele frequencies change so that stable equilibria in gene frequency space correspond to local maxima in mean fitness: This is known as Wright's adaptive landscape (Crow and Kimura, 1970; Wright, 1977; Ewens, 1979; Hartl and Clark, 1997).

2. Directional Selection

Regarding the case of a pair of alleles, when $0 \leq h \leq 1$ (so that the fitness of the heterozygote is bounded by the fitnesses of the homozygotes), A is favored by selection and will progress to fixation. This case is referred to as directional selection. If A is initially rare so that second-order terms in p can be neglected. Eq. (1) can be approximated by $\Delta p = s\ (1 - h)p$. The initial rate of change of log gene frequency ($\approx \Delta p/p$) is therefore proportional to the fitness difference between the genotypes Aa and aa. This reflects the fact that, with random mating, rare alleles are overwhelmingly carried in heterozygotes (frequency $2pq$) since the frequency of homozygotes (p^2) is negligible. If A is completely recessive so that $h = 1$, Eq. (1a) for rare A is approximated by $\Delta p = sp^2$ so that the logarithmic rate of increase in frequency of A is proportional to sp, which tends to zero with decreasing p. This reflects the extreme rarity of the favored homozygotes, and it implies that rare recessive alleles are only weakly selected in randomly mating populations (Haldane, 1932).

3. Survival of Favorable Mutations

The previous conclusion is reinforced by calculation of the probability of survival of new mutations. Even in very large populations, a new mutation is likely to be represented initially in only one or a few individuals. Since reproduction is subject to random variation, described by the probability distribution of the numbers of surviving offspring per mated individual, there is a finite chance in each generation that all carriers of a rare mutant gene will fail to transmit it (Fisher, 1930; Haldane, 1932). The chance that a single copy of a new favorable mutation ultimately survives random loss from a large population of constant size is approximately $2s\ (1 - h)$, assuming a Poisson distribution of offspring number (Haldane, 1932; Crow and Kimura, 1970). Most favorable mutations are thus likely to be lost from the population a few generations after they arise; on average, 148 occurrences of a mutation with a heterozygous selective advantage of 1% are needed for a 95% chance that one will survive.

This implies that there is a considerable random element to adaptive evolution at the genetic level. If there are several different loci at which mutations that provide adaptations to a given pressure of selection can occur, it may be a matter of chance which locus actually responds to selection in a given population. The same pressure of selection can therefore result in the divergence of isolated populations at the level of the genotype, even if the same phenotype is evolving. The numerous different genes involved in the adaptation of different human populations to malaria parasites illustrate this principle (Hill and Weatherall, 1998).

The formula for survival probability implies that a recessive favorable mutation has a zero chance of survival in an infinite population; calculations based on diffusion theory (see Section IV,D) show that the probability for a randomly mating population of size N in this case is approximately $0.8\sqrt{s/N}$ (Crow and Kimura, 1970). This means that recessive autosomal mutations are unlikely to become established by selection in randomly mating populations of even moderate size. As first pointed out by Haldane (1932), it is thus no accident that nonrecessive alleles have been established by selection in cases of recent adaptation involving genes of major effect, such as industrial melanism and pesticide resistance, despite the fact that most spontaneous mutations are recessive with respect to their effects on the phenotype (Haldane, 1932; Ford, 1975).

4. Time Course of Gene Frequency Change

Once a favorable allele rises to a sufficiently high frequency that random loss is unlikely, its progress in a large population can be calculated by integration of the differential equation which approximates Eq. (1) when selection is weak. This procedure yields expressions for the time needed to change gene frequencies by a given amount (Haldane, 1932; Crow and Kimura, 1970; Hartl and Clark, 1997). These can be used to estimate selection intensities in experimental and natural populations by comparing observed trajectories of gene frequency change with the theoretical predictions. This method is particularly useful for microbes, which have short generation times and can be grown in very large artificial populations. Selection coefficients on the order of 0.5% can be measured in microbial experiments (Hartl and Clark, 1997).

The results of these calculations show that the time

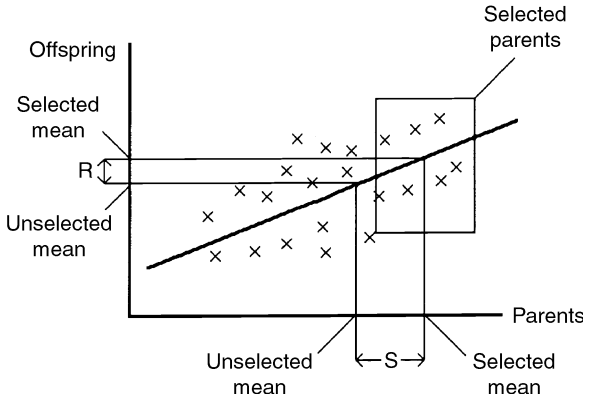

FIGURE 1 Each cross indicates the location of a point determined by the mean values of the phenotypic values of a pair of potential parents (*x* axis) and of their progeny (*y* axis). The straight line is the regression of *y* on *x*, whose slope is equal to the heritability. Families that are successful in reproducing are enclosed in the rectangular box.

required to cause a given amount of change in gene frequency is inversely proportional to *s* and depends only logarithmically on the initial frequency of the favorable allele, except when it is recessive. This implies that selection on nonrecessive alleles can cause evolutionary change on a timescale of the order of a few multiples of $1/s$, almost regardless of the initial conditions. Even a very weak force of selection can thus transform a population in a period of time that is negligible in geological terms (Fisher, 1930; Haldane, 1932). This conclusion is historically very important because it removes many of the objections to natural selection as a potent force in evolution.

D. Selection on Quantitative Traits

1. Predicting the Response to Selection

This question can also be considered in relation to quantitative traits, which are more relevant to phenotypic evolution than traits controlled by single genes. The standard model of quantitative genetics (which involves several simplifying assumptions) implies that the response to selection on a single quantitative trait is governed by the equation

$$R = h^2 S, \tag{2}$$

where the selection response, *R*, is the difference in mean value between one generation and the next; h^2 is the heritability of the trait (see Section I,A,2) (there is no connection with the dominance coefficient introduced previously); and *S*, the selection differential, measures the intensity of directional selection applied to the trait and is equal to the difference in mean between the parental individuals who have survived selection and the population mean before selection (Fig. 1). Given the fact that h^2 is usually substantially greater than zero, this implies that populations can respond rapidly to directional selection on quantitative traits, as is indeed the case (Falconer and Mackay, 1996).

2. Rates of Change under Selection

In a sexually reproducing population, the ability of recombination to put together new combinations of alleles means that selection can eventually produce genotypes with trait values that far exceed the range of variation that existed in the original population, simply by causing the fixation of alleles which increase trait value and that were segregating in the original population. In the absence of new variability created by mutation, the selection response will eventually cease when all favorable alleles have been fixed by selection. Given the evidence for a significant input of new variation by mutation for typical quantitative traits, a relatively gentle pressure of selection in a large population will never fully deplete variation and therefore sustained responses to selection are possible (Falconer and Mackay, 1996). The breeder can produce substantial changes in a single trait over less than 100 generations, comparable in magnitude to geologically rapid changes that have taken thousands of generations in nature. It is therefore unlikely that lack of genetic variation is often a serious constraint on the ability of populations to undergo phenotypic evolution, unless there is intense selection for a novel combination of traits, which cannot immediately be produced by recombination among existing genotypes (Wright, 1977).

3. Selection on Multiple Characters

In general, natural selection acts on suites of characters, not on a single trait in isolation. A multivariate generalization of Eq. (2) has been derived by Lande (1988), and it is useful for the interpretation of data on selection on multiple traits:

$$\Delta\bar{\mathbf{z}} = \mathbf{G}\nabla \ln \bar{w} \tag{3}$$

where $\Delta\bar{\mathbf{z}}$ is the change per generation in the vector of mean values for the set of traits, $\nabla \ln \bar{w}$ is the gradient vector of log mean fitness with respect to the set of trait means, and $\mathbf{G}$ is the matrix of additive genetic variances and covariances among the set of traits. The similarity to Eq. (1b) is evident.

Equation (3) shows that, in general, we can only predict the effect of selection on a trait if we know the extent to which it is genetically correlated with other traits that are also the target of selection. Short-term evolutionary changes in a given trait may be due, at

least in part, simply to the fact that it covaries with another trait that is the target of selection. However, provided that the G matrix is nonsingular and fitnesses are constant in time, a stable equilibrium state corresponds to a maximum in the surface of mean fitness with respect to the mean trait values, indicating that selection will ultimately bring the population close to the optima with respect to each individual trait unless there are strong constraints on the structure of genetic variation. This implies that the concept of an adaptive landscape can be applied to quantitative traits (Lande, 1988).

E. Maintenance of Variation by Selection

The role of selection in preserving variation rather than destroying it, called balancing selection, is now discussed. This form of selection was unknown to Darwin, and its discovery is one of the most important contributions of population genetics.

1. Heterozygote Advantage

The simplest case is when the heterozygote at a locus with two alleles has a higher fitness than the two homozygotes, which was first investigated by Fisher in 1922. Let the fitnesses of AA and aa relative to Aa be written as $1 - s$ and $1 - t$; Eq. (1b) now yields the result that there is a unique stable equilibrium at which the frequency of A is $p^* = t/(s + t)$, to which the population converges from any starting point other than fixation for A or a (Crow and Kimura, 1970; Ewens, 1979; Hartl and Clark, 1997). The classic example of heterozygote advantage is the polymorphism for the β–globin variant in humans that causes sickle-cell anemia when homozygous, a disease which is effectively lethal under natural conditions. The maintenance of this allele in human populations subject to severe malarial infections is due to the selective advantage of heterozygous carriers, conferred by their resistance to malaria, as suggested by Haldane in 1948. Several human globin gene polymorphisms, as well as some other human polymorphisms, are also maintained by resistance to malaria (Hill and Weatherall, 1998).

2. Frequency-Dependent Selection

Balancing selection can also be caused by negative frequency-dependent selection, in which the relative fitness of a genotype decreases with its frequency in the population. This obviously acts to inhibit the fixation of an allele which may initially have a selective advantage over other alleles when introduced into a population. Genetically controlled resistance to parasitic disease is subject to this form of selection since the abundance of a parasite decreases as the number of susceptible hosts diminishes, thereby reducing the selective advantage to a host allele that causes resistance to a particular parasite genotype (Li, 1997). A similar frequency dependence may affect alleles controlling virulence in the parasite population. Genes thought to be involved in disease resistance, such as the major histocompatibility complex (MHC) loci of vertebrates, are often highly polymorphic and segregate for large numbers of alleles; molecular analyses show clear evidence for balancing selection (see Section IV,C). The role of frequency-dependent selection in these cases remains to be established, however. Frequency dependence is inherent in the dynamics of the self-incompatibility loci of flowering plants, which also have very high allele numbers (Hartl and Clark, 1997). Similarly, Batesian mimicry, in which an edible species mimics a distasteful model, exhibits frequency dependence since a predator is more likely to mistake a rare mimetic form for the model than a common one (Fisher, 1930; Ford, 1975). It should be noted that frequency-dependent selection does not necessarily lead to polymorphism; its outcome depends on a delicate balance of the selective parameters, and dynamic complexities such as limit cycles or even chaos are possible. With frequency-dependent selection, maxima in mean fitness do not necessarily correspond to stable equilibria.

3. Temporal Variation in Fitnesses

Temporal fluctuations in relative fitnesses can also lead to the maintenance of polymorphisms by selection; in the case of two alleles at a locus, a sufficient condition is that the geometric mean fitness of the heterozygote over generations exceeds that of both homozygotes (Crow and Kimura, 1970; Hartl and Clark, 1997). Similarly, spatial variation in the direction of selection can maintain variation. This can happen in two ways. The first requires strong density-dependent regulation within different environmental patches so that the number of adults emerging from a patch is largely independent of the genetic composition of the eggs laid in that patch. In this case, opposing directions of selection in different patches can maintain polymorphism even if there is complete random mating among patches (Crow and Kimura, 1970). Directional selection in opposite directions on alleles in males and females is a special case. Second, restricted migration among populations subject to different directions of selection can result in polymorphism within each population accompanied by genetic differentiation among populations. Clinal variation in allele frequencies or quantitative trait values

results from geographic gradients in selection pressures, coupled with the smoothing effect of migration, as in the case of Bergmann's rule, which states that the mean body sizes of mammal populations increase with higher latitudes (Hartl and Clark, 1997).

4. Meiotic Drive

Antagonism between the effects of selection at different levels may also maintain variation. The best studied cases involve the phenomenon of meiotic drive or segregation distortion, in which one allele (D) at a locus when heterozygous causes the destruction of gametes carrying the alternative allele (d). In animals, this is usually found to occur only in males. Provided that the fertility of Dd males is affected less than linearly by the destruction of the d sperm, D will gain a transmission advantage. It will spread through the population unless there is a countervailing selective disadvantage at the level of individuals (Hartl and Clark, 1997). In the best studied cases of this kind, the SD system of *Drosophila melaogaster* and the t-haplotype system of house mice, the primary disadvantage seems to come from sterility of DD males.

F. Mutation–Selection Balance

The balance between recurrent mutations to variants that impair the functions of gene products and selection against them is probably a major factor in maintaining genetic variation, given the fact that higher organisms have tens of thousands of genes (Drake *et al.*, 1998). The large number of changes in a coding sequence that can impair the function of a gene product implies that the process of mutation to deleterious alleles at a locus can be regarded as effectively irreversible, provided that the wild-type allele predominates in the population. If the deleterious alleles at a locus are completely recessive, with selection coefficient s and a rate of origination by mutation from wild-type of u, their equilibrium frequency in a randomly mating population (assuming $s \gg u$) is $q^* = \sqrt{u/s}$. Even lethal alleles ($s = 1$) can thus reach appreciable frequencies if they are completely recessive; with $u = 10^{-5}$, for example, $q^* = 3 \times 10^{-3}$. However, experimental studies of lethal mutations in *Drosophila* show that they usually impair the viability of heterozygotes with wild-type by 2 or 3%; detrimental alleles with more minor effects (s on the order of a few percent) appear to be much less recessive, with h values of approximately 0.25 (Crow, 1993). With random mating, the much greater frequency of heterozygotes than mutant homozygotes means that selection on heterozygotes controls the frequencies of mutations; in this case, $q^* = u/(hs)$. Indirect experimental evidence indicates that the mean of hs for detrimentals is on the order of 1% in *Drosophila* so that q^* is approximately 10^{-3} with $u = 10^{-5}$. The frequency of heterozygous carriers is approximately $2q^*$ so that the mean number of heterozygous detrimental mutations per individual in a genome such as that of *Drosophila* with approximately 15,000 genes is $30{,}000 \times 10^{-3} = 30$ with these assumptions. The total rate of mutation to lethal mutations per haploid genome in *Drosophila* is 0.01; assuming $hs = 0.02$, the mean number of heterozygous lethals per diploid individual is on the order of 1, in agreeement with the direct estimate mentioned previously.

G. Genetic Load

1. General Considerations

The previous discussion leads to the consideration of the effect of selection on the fitness of the population as whole; if there is a large amount of variability with respect to loci under selection, it is obvious that the mean fitness of the population must be much less than that of the best genotype. If fitnesses are measured relative to a value of 1 for the optimal genotype in the system under consideration, the reduction in fitness can be conveniently measured by the genetic load, defined as $L = 1 - \overline{w}$. In the case of viability selection, no genotype can have a survival probability greater than 1, so L provides an upper bound to the probability that a zygote survives to maturity. More generally, L measures the proportion of the population that dies or fails to reproduce as a result of selective differences among genotypes. The effects of multiple loci on mean fitness can be calculated by assuming that different loci have independent effects, so that the fitness of a multilocus genotype is given by the product of the fitnesses of all the single-locus genotypes which contribute to it. If L_i is the load contributed by the ith locus, the mean fitness with respect to m independent loci, relative to the value for the optimum genotype, is:

$$\overline{w} = \prod_{i=1}^{m} (1 - L_i) \approx \exp - \sum_{i=1}^{m} L_i. \qquad (4)$$

2. Mutational Load

With nonrecessive deleterious alleles maintained by mutation, the load for a single locus at equilibrium is approximated by $2q^*hs = 2u$. The total load is $1 - \exp - U$, where U is the mean number of new deleterious mutations in a diploid individual. Lethal

mutations contribute relatively little to this since their total mutation rate is very low, but detrimentals have a major effect: Assuming 15,000 genes with a mutation rate of 10^{-5}, the mutational load for a *Drosophila* population would be about 0.26 (Drake *et al.*, 1998). For a genome of 80,000 genes, as in mammals, the load would be 0.80, a considerable burden of selective loss.

The existence of a large mutational load suggests that there is an adaptive advantage to a reduction in the mutation rate; this can be studied theoretically by calculating the rate of spread of a rare modifier gene that reduces U by a small amount δU. If the modifier recombines freely with autosomal loci subject to mutation and selection, it has a selective advantage of $hs\ \delta U$. This raises the question of why mutation rates are not closer to zero; this probably reflects the fact that there is a fitness cost to the necessary repair systems so that U is adjusted to a level at which the costs and benefits of increased repair balance (Drake *et al.*, 1998).

3. Segregational Load

Similar calculations can be performed for models of balancing selection, yielding estimates of the segregational load. In the case of heterozygote advantage, the load due to a single locus is $st/(s + t)$ (Crow and Kimura, 1970). Equation (4) can be used to determine the segregational load contributed by a large number of polymorphic loci with independent effects. This can be considerable, even if selection is weak. For example, 10,000 loci each with $s = t = 0.001$ would yield a mean fitness of only 0.0067. This is so low that only a very high fecundity species would be able to produce the two surviving offspring per adult needed to maintain itself. This implies that either most molecular variation has very slight or no effects on fitness or the assumption of multiplicative fitnesses is unrealistic.

An extreme alternative to multiplicative fitnesses is truncation selection. Genotypes at a set of loci are assumed to be ordered with respect to their fitnesses as determined by the multiplicative fitness model; a fixed proportion of the population, containing the set of genotypes with the highest fitnesses, is allowed to survive. This is equivalent to assuming that individuals compete for a limiting resource, and that only the fittest succeed. Under these conditions, a much larger number of loci can be exposed to selection for a given total L than with multiplicative fitnesses, for the same selection intensity per locus (Crow and Kimura, 1970). Less extreme forms of departure from multiplicativity can have similar but smaller effects on the total load.

4. Substitutional Load

Genetic loads also apply to adaptive evolution. Consider the case of a biallelic locus, where A is initially deleterious and held at a low frequency. If there is a change in the environment so that A becomes favored by selection, it will start to increase. However, in any generation before it reaches fixation, the mean fitness of the population will be reduced below its final value of 1 because of the presence of the disfavored a allele. There is thus a load associated with the substitution of a by A, reflecting the fact that natural selection cannot instantly transform a predominantly a population into one which is fixed for A. The sum of the loads for each generation over the course of a gene substitution is Haldane's cost of selection, C; if the population size is N, the total number of individuals eliminated by selection is CN. Providing selection is not too strong, C is proportional to minus the logarithm of the initial frequency of A; a typical value is 30 (Crow and Kimura, 1970).

The effect of changes at multiple loci can be derived as follows. Assume that there is a steady rate of change in the environment so that each generation K loci start to experience gene substitutions of this kind. K is the rate of gene substitution in the genome as a whole such that after a long period of evolutionary time, T, the population will differ from its ancestral state by KT substitutions. If a gene substitution takes t generations to complete, Kt loci will be segregating in any given generation, each associated with an average load of C/t relative to a population which is fixed for the favorable alleles. The mean fitness under multiplicative fitness, relative to a population that is fixed for favorable alleles at all currently segregating loci, is then given by $(1 - C/t)^{Kt} \approx \exp - CK$.

Data on rates of protein evolution suggest that K for an average amino acid site is about 1.5×10^{-9} per year in mammals (Kimura, 1983). With 80,000 loci coding for proteins with average size of 300 amino acids, K for the genome is 0.036. With $C = 30$, the mean fitness is 0.34, assuming one generation per year. A much higher load would be found if changes at silent and noncoding sequences are also taken into account. This finding of a high substitutional load associated with molecular evolution was one of the main motivations for the development of the neutral theory of molecular evolution, which asserts that most evolution at the molecular level is caused by the random sampling of alleles in finite population size, genetic drift, and not by natural selection (Kimura, 1983). The substitutional load can be considerably reduced by modifications to the assumption of multiplicative fitnesses, such as truncation

selection, so that this argument loses much of its cogency. Nevertheless, there are other good reasons to take the neutral theory seriously (see Section III,C).

H. Multiple Loci

1. No Selection

So far, it has tacitly been assumed that evolutionary change involving more than one locus can be modeled by assuming that alleles at different loci are distributed independently of each other in the population and have independent effects on phenotypes and fitness. Although this may be a good approximation for many purposes, it is necessary to examine the consequences of relaxing these assumptions. Deviations from independence among loci in randomly mating populations can be described by linkage disequilibrium parameters, which measure the extent to which the frequencies of the different multilocus gamete types or haplotypes depart from the frequencies expected by randomly combining alleles at different loci. In the simplest case of a pair of loci, each with two alleles (A and a and B and b), there are four haplotypes: AB, Ab, aB, and ab. Let the frequencies of these be x_1, x_2, x_3 and x_4. If the allele frequencies at the two loci are p_A and p_B, we can write the haplotype frequencies as $p_Ap_B + D$, $p_A(1 - p_B) - D$, $(1 - p_A)p_B - D$, and $(1 - p_A)(1 - p_B) + D$, respectively, where D is the coefficient of linkage disequilibrium. It is easily seen that $D = x_1x_4 - x_2x_3$. If the frequency of recombination between the two loci is c ($0 \leq c \leq 0.5$), the value of D in the next generation in an infinitely large, randomly mating population with no selection is $D(1 - c)$.

In the absence of evolutionary forces other than recombination, the extent of nonrandom association between a pair of loci thus decays exponentially at a rate that is determined by the frequency of recombination. This result can be generalized to associations between multiple loci (Crow and Kimura, 1970). Unless populations are far from equilibrium, departures from linkage equilibrium at a set of loci require the operation of forces tending to generate nonrandom associations between alleles at the different loci. Their magnitude will be at least on the order of the recombination frequencies among them. One such force is genetic drift (see Section III), which can cause randomly generated linkage disequilibrium among loci for which c is on the order of the reciprocal of the population size (Ewens, 1979).

2. Selection on Several Loci

Another possible force causing linkage disequilibrium is epistatic selection, in which the difference in fitness between genotypes at one locus varies according to the genotypes at the other loci in the system. If the fitness effects of different loci combine additively, there is no epistasis, and it can be shown that polymorphic equilibria under random mating exhibit no linkage disequilibrium (Ewens, 1979). With epistasis, selection tends to preserve haplotypes which contain favorable combinations of alleles, whereas recombination breaks them down. Linkage disequilibrium is not necessarily present if linkage is sufficiently loose in relation to the strength of epistasis. Multiple alternative stable equilibria may occur in multilocus systems so that the fate of a population can be affected by the initial conditions from which evolution starts. Stable equilibria do not necessarily correspond to maxima in mean fitness in the space of haplotype frequencies, another violation of the adaptive landscape principle. However, if epistatic selection is weak in relation to the frequency of recombination, populations tend to converge to trajectories where linkage disequilibrium is nearly constant (quasi-linkage equilibrium), and mean fitness increases monotonically at a rate approximately equal to the additive genetic variance in fitness according to Fisher's fundamental theorem of natural selection (Ewens, 1979). This has provided a very useful tool for the analysis of the dynamics of multilocus systems (Barton and Turelli, 1991).

3. The Evolution of Close Linkage

There are two biologically important features of systems with strong epistatic selection. The first is that such selection may impose strong constraints on the degree of linkage between polymorphic loci. Suppose that the population is initially segregating for alleles A and a at one locus but is initially fixed for b at a second locus. If a mutation B arises at this locus which interacts with the alleles at the first locus, such that AB is selectively favored but aB is disfavored, B may be unable to invade the population unless c is below some threshold value. Only mutations at loci that are sufficiently closely linked to the first polymorphism in the system will be able to establish subsequent polymorphisms. This process has probably been important in the evolution of some of the classic examples of supergenes (systems of very closely linked loci held in strong linkage disequilibrium by selection), such as Batesian mimicry in butterflies (Ford, 1975) and sex chromosomes. Similarly, if ab and AB are both fitter than Ab and aB, a population fixed for ab may only evolve a two-locus polymorphism if there is a double mutation to AB and if c is sufficiently small. This probably occurred in the evolution of meiotic drive systems, which require combinations of alleles at several loci that are individually disfavored.

Second, there is a selective advantage to modifier alleles that reduce the frequency of genetic recombination between the two loci once a two-locus polymorphism has been established. If suitable genetic variation in recombination rates is available, this will eventually lead to such close linkage that the system has the appearance of a single locus (Fisher, 1930). This principle has wide generality; analysis of the conditions for spread of rare modifiers of recombination rates has shown that randomly mating populations under epistatic selection generating linkage disequilibrium at a system of loci will always tend to evolve closer linkage (Barton and Charlesworth, 1998). Since genetic recombination is a near-universal feature of living organisms, these findings have led to the search for situations that promote rather than repress recombination; these involve forces such as mutation and environmental change that perturb populations away from equilibria under selection (Barton and Charlesworth, 1998).

III. RANDOM GENETIC DRIFT

The discovery that random sampling of allele frequencies in finite populations may be a significant factor in evolution is another major contribution of population genetics. This process has two aspects: The first is the tendency for a population of finite size to become genetically uniform, owing to the fact that there is an increasing tendency as time passes for all the copies of a gene at a locus to be descended from a single ancestral allele (Fig. 2). The second is the tendency of isolated populations to diverge in allele frequencies over time, since independent trials of a population with the same initial state will arrive at different allele frequencies by chance.

The first process is closely related to the increase in homozygosity that accompanies the inbreeding of close relatives; both are conveniently studied by means of the concept of identity by descent. Two alleles at the same locus drawn from a population are said to be identical by descent if they trace their ancestry back to a single ancestral allele. The extent to which a population has progressed toward genetic uniformity can be measured by its inbreeding coefficient, defined as the probability that a pair of randomly sampled alleles are identical by descent (Hartl and Clark, 1997). This is always measured with respect to an initial generation, in which all the alleles at a locus are arbitrarily decreed to be nonidentical (Fig. 2).

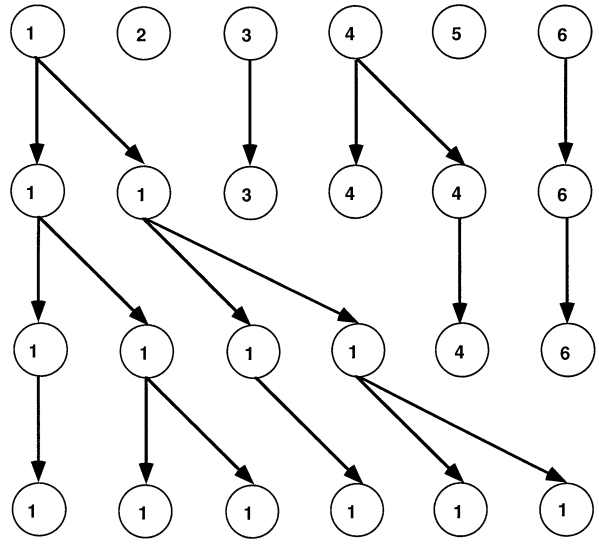

FIGURE 2 Each row of circles indicates the alleles at a locus present in a given generation. The distinct copies present in the initial generation are labelled 1–6. The arrows indicate the successful transmission of an allele to the next generation.

A. Increase in Homozygosity

The simplest Wright–Fisher model of genetic drift assumes a discrete-generation, randomly mating population of N hermaphroditic individuals with no selective differences among genotypes at the locus under consideration. New individuals are formed by random sampling (with replacement) of gametes produced by the parents. With diploid inheritance, such that there are $2N$ gene copies at a locus among the breeding adults, the following recursion relation for f_t, the inbreeeding coefficient in generation t, is obtained:

$$1 - f_t = \left(1 - \frac{1}{2N}\right)\left(1 - f_{t-1}\right). \qquad (5)$$

This follows from the fact that the probability that a pair of randomly sampled alleles are both derived from the same allele in the previous generation is $1/(2N)$, in which case they are identical by descent. The probability that they come from two different alleles is $1 - 1/(2N)$; their probability of identify is f_{t-1}.

Equations (5) shows that f_t tends asymptotically to 1; if N is moderately large, $1 - f_t$ decays exponentially with a half-life of $1.4N$ generations. This result can be generalized to more realistic types of breeding system by means of the concept of the inbreeding effective population number (N_e) (Crow and Kimura, 1970). This utilizes the fact that genetic drift in these more general cases can be described by matrix equations or

higher order difference equations, such that a constant rate of decay of $1 - f_t$ is reached asymptotically, yielding an expression of the same form as Eq. (5). The asymptotic decay constant can thus be equated to $1 - 1/(2N_e)$. For example, a population with N_f breeding females and N_m males has $N_e = 4\,N_fN_m/(N_f + N_m)$. In general, N_e is less than the census number of breeding individuals, often considerably so (Crow and Kimura, 1970).

B. Differentiation of Populations

The effect of drift in causing genetic differentiation among populations is described by the variance in allele frequencies among the set of populations in question. In the Wright–Fisher model, the sampling of allele frequencies at a neutral biallelic locus with alleles A and a and frequencies p and q will generate a binomial distribution of the new frequency of A, with mean p and variance $pq/(2N)$. Repetition of this sampling process results in a probability distribution of allele frequencies with steadily increasing variance (Crow and Kimura, 1970). The variance of allele frequency, σ_p^2, can be related to the inbreeding coefficient as follows. Assume that we have an infinitely large set of completely isolated populations, all founded with the same initial frequency, p_0. At some arbitrary time, the set of populations will have a mean gene frequency p_0 (since drift does not change the mean) and variance σ_p^2. If N is reasonably large, the genotype frequencies within each given population will be in Hardy–Weinberg proportions; the mean frequencies of AA, Aa, and aa over the set of populations are equal to $p_0^2 + \sigma_p^2$, $2(p_0q_0 - \sigma_p^2)$, and $q_0^2 + \sigma_p^2$, respectively.

This is the Wahlund effect: Genotype frequencies averaged over a set of populations that are individually in Hardy–Weinberg proportions show an excess of homozygotes and a deficiency of heterozygotes compared with Hardy–Weinberg expectation, whose value is determined by the variance in gene frequencies among the populations. This is a purely algebraic result, independent of the causes of the variation. It can be related as follows to the effects of genetic drift in causing increased homozygosity. Assume that a diploid individual is formed by sampling two random alleles from the same population. The probability that the alleles are identical by descent is f; they are then both A in state with probability p_0. The probability that they are nonidentical is $1 - f$, in which case the probability that the two alleles are both A is p_0^2. The net probability that the individual is AA is thus $fp_0 + (1 - f)p_0^2 = p_0^2 + fp_0q_0$. Comparison with the previous expression shows that

$$\sigma_p^2 = fp_0q_0. \tag{6}$$

This establishes that the variance in gene frequency is proportional to the inbreeding coefficient under the Wright–Fisher model, and therefore its change over time is governed by Eq. (5). This is often but not always true under more general models of population structure, with N_e replacing N in the binomial formula for the variance conditional on the current gene frequency. In some circumstances, particularly when the population size changes in time, the conditional variance in gene frequency requires a different denominator in order to be represented by the binomial formula. In this case, a variance effective population number is computed (Crow and Kimura, 1970).

C. Molecular Evolution and Variation

1. The Neutral Theory

These simple models of genetic drift can readily be applied to the study of molecular evolution and variation, assuming selective neutrality at the loci in question. Neutral theory allows for the possibility that many mutations are subject to purifying selection and are rapidly eliminated from the population (see Section IV,D,3); it is claimed that the fate of the bulk of the mutations that are not removed by purifying selection is determined by drift rather than selection (Kimura, 1983). This theory thus constitutes a useful null hypothesis, which can be tested against data on molecular evolution and variation by means of predictions of several types.

2. The Rate of Neutral Evolution

Consider first the rate of molecular evolution as measured by the rate of gene substitution, K (see Section II,G,4). In a population of N breeding adults, there are $2N$ allele copies at an autosomal locus. As shown in Section III,A, Eq. (5) implies that the population tends to homozygosity with probability 1. This means that the remote descendants of the current population will all trace their ancestry back to just one of these $2N$ alleles. Under neutrality, the probability that a given allele is the ancestor is thus $1/(2N)$. It follows that the probability of fixation of a new neutral mutation in a population of size N is $1/(2N)$; the probability that it is lost is $1 - 1/(2N)$. If the rate of mutation to neutral variants is u per generation, the expected number of new mutations that enter the population is $2Nu$, of which only $1/(2N)$ are destined for ultimate fixation; the expected number of mutations that ultimately become fixed is $2Nu/(2N) = u$.

If the process of mutation and drift has reached a stationary state, so that the expected number of new substitutions must equal the number of substitutions that go to completion each generation, we obtain the fundamental equation of neutral molecular evolution:

$$K = u \tag{7}$$

This relation holds for any level in the genetic hierarchy, from the nucleotide site through the locus to the genome as a whole, provided that K and u are defined appropriately. Since K can be determined by comparisons of DNA sequences among species with known divergence time (see Section II,G,4) and u values are known from molecular genetics, Eq. (7) can be used to test the neutral theory.

One prediction is that genomic regions whose sequences are essentially functionless, such as pseudogenes and the internal parts of introns, should evolve at the mutation rate since they are necessarily unconstrained by selection. If adaptive Darwinian evolution is a minor factor in molecular evolution, the rate for sites in these regions should be much higher than that for functionally significant regions, where selection is expected to eliminate most mutations [see Equation (10)]. These regions do indeed evolve at the rates expected from mutation rates, and other regions evolve more slowly (Kimura, 1983; Li, 1997). This does not, however, rule out a role for positive selection in fixing variants in selectively constrained regions; it could still be true, for example, that most replacement substitutions (see Section I,B,2) are deleterious, but a minority are advantageous rather than neutral, so that changes that are fixed in evolution are adaptive rather than neutral. There are also some exceptional cases of higher rates of replacement versus silent substitutions in coding regions; this is strong evidence for a positive role of selection on the amino acid sequences concerned (Li, 1997).

Another prediction of Eq. (7) is that the rate of molecular evolution should be constant over long periods of time since it depends only on the the mutation rate. In contrast, the rate of evolution under natural selection is expected to be highly variable since the theory described earlier suggests that populations will tend to adapt quickly to a new environment (or go extinct), after which change will be slow or nonexistent. This is borne out by the observations of comparative biology and paleontology, which show that evolution at the external phenotypic level is generally highly episodic and triggered by ecological opportunities such as the occupation of vacant niches. Studies of DNA and protein sequence evolution suggest that the rate of evolution of a given molecule is much less variable among different lineages, or within the same lineage at different times, than is true for the external phenotype, especially when noncoding or silent substitutions are considered (Kimura, 1983; Li, 1997). This has generated the concept of a molecular clock, which is used to estimate the times of divergence of taxa when paleontological data are absent. However, there is evidence for more variability in rates of evolution than predicted by the simplest form of the neutral theory, especially for amino acid sequences. This suggests a role for selection, although the interpretation of rate variability is controversial (Li, 1997).

3. Neutral Polymorphism

The process of fixation of neutral variants is a slow one; calculations based on diffusion theory (see Section IV,D) show that the mean time to fixation of a new neutral mutation in a random mating population (conditioned on ultimate fixation) is approximately $4N_e$ generations. While the variant is on the way to fixation, it causes a polymorphism. Similarly, variability is contributed by the large fraction of new mutations that are destined for ultimate loss; the mean time to loss is $(N_e/N)\ \ln\ (2N)$ generations, a much shorter time than the time to fixation. In the neutral theory, polymorphism is simply a phase of molecular evolution. Although there is a constantly shifting set of variants at any one locus, the mean amount of variability can be determined assuming a statistical equilibrium between drift and mutation, using the following argument (Kimura, 1983).

The simplest version of this applies to a single locus, which is assumed to have no recombination. Thus, these are many possible sequences; new neutral mutations are assumed to occur with probability u per generation, and each mutation represents a sequence that has not been observed before. This is the infinite alleles model. Let h_t be the probability that two randomly sampled alleles are distinct in sequence in generation t. An argument similar to that leading to Eq. (5) shows that, neglecting terms in $u/(2N)$, h_t for a Wright–Fisher population obeys the equation

$$h_t \approx \left(1 - \frac{1}{2N}\right)h_{t-1} + 2u(1 - h_{t-1}). \tag{8a}$$

At equilibrium, rearrangement of this equation yields

the relation

$$h \approx \frac{4Nu}{4Nu + 1}. \qquad (8b)$$

More generally, N_e can be substituted for N in Eq. (8b). The parameter $\theta = 4N_e u$ thus controls the equilibrium level of variability under the neutral model. A large amount of neutral variability can be maintained, provided θ is sufficiently large, e.g., with $N_e = 10^5$ and $u = 2 \times 10^{-7}$, $\theta = 0.08$, and $h = 0.074$. This value of h is similar to the mean per locus heterozygosity for electrophoretic alleles in mammalian populations (Lewontin, 1974; Kimura, 1983).

This model can be modified to predict the equilibrium level of diversity per nucleotide site by assuming that the units of observation are the individual sites, not the entire locus. If we assume that θ is much less than 1 so that at most one variant is segregating in the population at each site, Eq. (8) can be applied to yield the equilibrium value of π, the nucleotide site diversity (see Section I,B,2), such that $\pi = \theta$ to the assumed order of approximation. This is Kimura's infinite sites model, which is widely used in the interpretation of data on molecular variation.

D. The Coalescent Process

1. General Considerations

The growing body of data on DNA sequence variation within populations has stimulated interest in the development of statistical tests of the agreement of observed patterns of variation with the predictions of the neutral theory. In order to conduct such tests, it is essential to have predictions concerning the properties of statistics describing samples of alleles from populations and not merely of properties of the populations from which the samples are drawn since these cannot be observed directly. A powerful method for deriving properties of samples has recently been developed and is known as coalescent theory (Hartl and Clark, 1997; Li, 1997). It is based on the following principle. Consider a pair of alleles at an autosomal locus sampled from a Wright–Fisher population. As shown in Section III,A, there is a probability $1/(2N)$ that they are derived from a common ancestral allele in the previous generation, i.e., that they coalesce. If they fail to coalesce, which has probability $1 - 1/(2N)$, they have a probability $1/(2N)$ of coalescing in the next generation back, and so on.

There is thus a geometric distribution of the time back to the common ancestral allele, such that the probability of time t is $(1/2N)(1 - 1/[2N])^{t-1}$. From the well-known properties of this distribution, the mean time to the common ancestor is $2N$ and the variance is $(2N)^2$. Assume that mutations occur at rate u per site in the gene, such that each mutation that arises in the line of descent connecting the two sampled alleles is at a different site (the infinite sites assumption, see Section III,D,C,3). If there are m nucleotide sites in the sequence in question, the number of mutations conditioned on t has a mean and variance of $2tmu$ since the total time separating the alleles is $2t$, and the conditional number of mutations follows a Poisson distribution. The mean and variance of the number of differences between the two alleles are thus $m\theta$ and $m\theta + 0.5(m\theta)^2$, respectively. The result for the mean corresponds to that derived from Eq. (8) since the nucleotide site diversity is simply the number of differences between the pair of alleles divided by m.

This can be extended to a sample of n alleles from a population. If N is sufficiently large, the chance of more than one coalescent event per generation can be neglected. There are $n(n - 1)/2$ possible pairwise allelic combinations so that the probability of a coalescent event is $n(n - 1)/(4N)$. The time to the first coalescent event is therefore distributed approximately exponentially, with a mean of $4N/(n[n - 1])$ and variance equal to the square of the mean. The time from this to the next coalescent is also exponentially distributed, replacing n with $n - 1$, and so on. The process can be represented by a genealogical tree, whose nodes with k and $k - 1$ alleles are separated by a time t_k that is exponentially distributed with mean $2/(k[k - 1])$ on the coalescent timescale of $2N$ generations (Fig. 3). As before, N_e replaces N for more general models of breeding structure. The rate of coalescence evidently diminishes as the number of nodes decreases.

This representation can be used to derive many important results, and it also provides a rapid means of simulating genetic processes, since time can be rescaled to the coalescent timescale and the properties of a sample of alleles represented by generating genealogical trees from samples drawn from the relevant exponential distributions, with mutations scattered randomly over the branches of the tree. The expected pairwise difference between all $n(n - 1)/2$ pairs of alleles on the infinite sites assumption is readily seen to be $m\theta$. Its variance is also known (Li, 1997). This statistic provides the obvious means of estimating the nucleotide site diversity in the population (θ) by equating the observed mean pairwise difference to its expectation. However, another statistic, the total number of segregating sites in a sample of n alleles (S_n), has better statistical properties under the infinite sites model. These can be obtained as follows. The total size of a tree is the sum of kt_k over the entire tree. Application of the properties of the

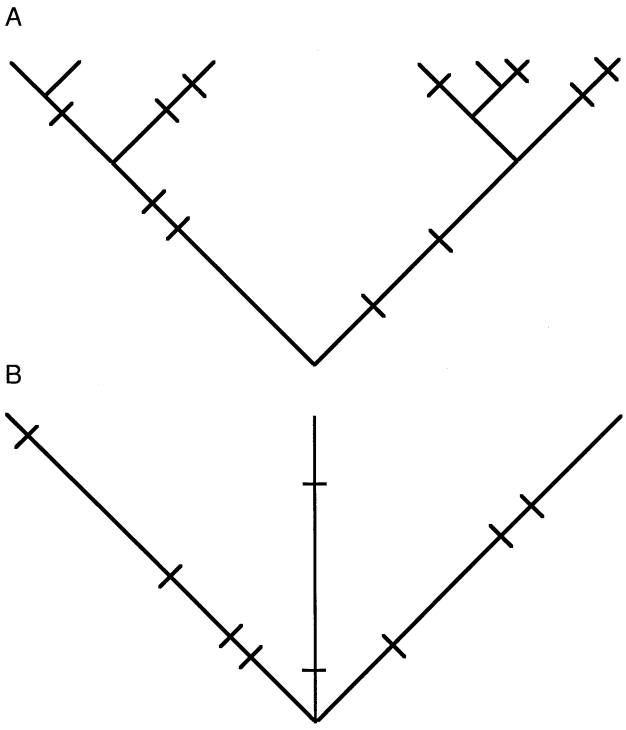

FIGURE 3 Two types of gene genealogies. (A) The genealogy of a set of alleles in a population of constant size. The tips of the tree represent six alleles sampled from the population. Each node corresponds to a coalescent event. Each cross line represents a new mutation in the DNA sequence of the gene in question, which has arisen since the common ancestor of the sample; mutations in internal branches of the tree give rise to at least two variants in the sample. (B) The genealogy of a set of three alleles in a population that experienced a recent reduction in size to one allele, followed by rapid expansion with no opportunity for coalescent events.

exponential distribution shows that the expectation of the tree size in units of coalescent time is simply $A = 2\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots \frac{1}{(n-1)}\right)$. The number of sites in the sample that are segregating for nucleotide variants is the number of mutations that appear on the tree. Its expectation is $2NmuA = m\theta$A; therefore, θ can be estimated by dividing S_n by mA. The variance of this estimator is substantially lower than that based on pairwise diversity.

2. Tests of Neutrality

Several tests for departures from neutrality have been devised based on the properties of the coalescent process (Li, 1997). One widely used method is Tajima's D test, which uses the fact that departures from neutrality have different effects on the pairwise difference estimate of θ than on the S_n estimate. For example, a population bottleneck of reduced size that completely removes variability at a locus will be followed by a long period of recovery, during which new mutations are all at low frequencies. This means that the pairwise diversity, which is weighted by the frequencies of variants, will be much lower in relation to the observed number of segregating sites than in an equilibrium population, reflecting the fact that the genealogical tree is like a "star phylogeny" in this case (Fig. 3). Purifying selection on the variants in question will have a similar but much smaller effect. The opposite pattern is expected if there has been a partial bottleneck, in which rare variants have been lost from the population, or if variation is affected by balancing selection. The Tajima test computes the ratio of the difference between the two estimates of θ to the standard deviation of their difference (calculated using the neutral equilibrium model) and compares the result with critical values obtained from coalescent simulations. Other tests have been derived from related principles (Li, 1997).

IV. THE INTERACTION OF DRIFT WITH DETERMINISTIC FORCES

A. Population Subdivision

1. General Considerations

One important complication of the neutral theory is population subdivision; species in nature are not simple randomly mating populations but instead are usually distributed over wide geographic areas, with limited migration between localities. Random genetic drift can thus produce significant genetic differentiation among local populations if migration is sufficiently restricted. This creates two problems: How to describe data on allele or nucleotide site variant frequencies when there is differentiation among populations, and how to relate models of the underlying evolutionary processes to the data.

2. Partitioning Variability

The basic method used to summarize data collected from a set of populations sampled from the same species is to partition variability into a within-population component that describes the mean level of variability within a sample and a between-population component that measures the mean difference between alleles sampled from different populations. This can be done either using data on allele frequencies, as in the case of electrophoretic loci or microsatellite loci, or using nucleotide site variant frequencies. In many cases, it is also possible to organize the populations sampled into a hierarchy of decreasing geographic scale, such as subspecies within a species, races within subspecies, and local populations within races. Variability can then be partitioned into variation between and within each level of the hierarchy (Hartl and Clark, 1997). A frequent practice is to nor-

malize each lower level measure of between-population variability relative to variability at the next higher level of the hierarchy, generating a set of Wright's F statistics or the related G statistics of Nei (Hartl and Clark, 1997).

The procedures will be illustrated here for the case of DNA sequence variation in a population divided into a set of equal-ranking subpopulations. The total nucleotide site diversity, π_T, is defined as the fraction of nucleotides that differ between a pair of alleles drawn randomly from the set of samples as a whole. The within-population nucleotide site diversity, π_S, is the mean fraction that differ between a pair of alleles drawn from the same subpopulation. The amount of between-population differentiation can be measured by the between-population component of nucleotide site diversity, $\pi_{T-S} = \pi_T - \pi_S$. Alternatively, a measure of nucleotide site divergence, π_B, can be defined as the mean fraction of nucleotide sites that differ between pairs of alleles drawn from a pair of different subpopulations, and another measure of between-population differentiation is defined as $\pi_D = \pi_B - \pi_S$. Two normalized measures of differentiation can be defined as $G_{ST} = \pi_{T-S}/\pi_T$ and $F_{ST} = \pi_D/(\pi_S + \pi_D)$. Both measures are widely used in the literature and are often similar numerically, especially if many subpopulations are sampled.

3. The Island Model

Although the previously mentioned measures of population differentiation are useful as descriptive tools, they can also be used to estimate the evolutionary parameters that determine the extent of population differentiation. This requires the development of models of the joint effects of genetic drift, mutation, and migration, one of the most complex problems in theoretical population genetics. The simplest and most widely used model is the island model, which assumes that the species is divided into n distinct subpopulations or demes, which each behave according to the Wright–Fisher model with population size N. After reproduction has occured within each deme, a fraction m of each deme's genes are replaced with genes drawn randomly from the other $n - 1$ demes. Coalescent theory can be used to determine the mean coalescence times of pairs of alleles sampled from the same population (t_0) or from different populations (t_1); $t_0 = 2N$ and $t_1 = 2Nn(1 + [n - 1]/[4Nnm])$. In the infinite sites model, the mean fraction of nucleotides that differ between a pair of alleles is equal to the product of the mutation rate per site and twice their coalescence time (see Section III,D) so that the expected nucleotide site differences between alleles can be derived directly from the corresponding coalescence times.

An important and somewhat counterintuitive conclusion is that the within-population nucleotide site diversity, π_S, is equal to $4Nnu$, i.e., it depends on the total number of individuals in the set of populations in the same way as does the diversity in a panmictic population under the infinite sites model, and it is independent of the migration rate (with the proviso that $m > 0$). As expected, the other diversity measures depend inversely on Nm, with large between-population divergence being possible only when $Nm < 1$; for large n, both F_{ST} and G_{ST} are approximately equal to $1/(1 + 4Nm)$. Values of these statistics that are close to zero are generally taken to indicate relatively little population differentiation, whereas values close to one imply considerable differences among local populations relative to within-population variability.

4. Other Models of Population Structure

Although simple to analyze, the island model is not very realistic biologically since dispersal is limited in most species so that local populations are most likely to acquire immigrants from nearby. Attempts to generate useful results from more realistic models have taken two directions. The first involves maintaining the assumption of a set of discrete demes but allowing for differences in deme sizes. Migration is described by a migration matrix $\boldsymbol{M}$, such that m_{ij} is the probability that an allele in deme i originated in deme j in the previous generation. However, it is difficult to obtain transparent general results for such a model. Under the infinite sites model, however, the result that $\pi_s = 4Nnu$ still holds if migration is conservative, i.e., migration does not change the sizes of local populations (Maruyama, 1977; Nagylaki, 1986). Results on genetic diversity within and between populations are only available for special cases, such as the stepping-stone model, in which demes of size N in a linear or planar array receive migrants with probability m only from immediately adjacent populations (Maruyama, 1977). This model also permits the analysis of the dependence of degree of genetic differentiation between populations on the distance between them. The results show that extensive differentiation between populations is possible with a one-dimensional array of populations, even if there is a considerable amount of migration, whereas the results for a two-dimensional array are similar to those for the island model.

The second approach assumes that individuals are distributed over a one or two-dimensional spatial continuum, with density ρ. Migration is represented by the probability density that an individual moves a given distance between birth and reproduction. If migration follows a random walk, the variance of the migration

function, σ^2, is sufficient to describe the process. This is Wright's isolation-by-distance model. A mathematically correct formulation of this model presents formidable difficulties, but it is possible to obtain explicit results for the case of a one-dimensional continuum by treating it as the limit of the corresponding discrete population model (Nagylaki, 1986).

B. Effects of Directional Selection at Linked Loci

Another complication in interpreting data on DNA sequence variation is that even neutral variation may be affected by selection at linked sites. The classic example is hitchhiking, whereby a new favorable mutation arises as a unique event and spreads through the population. In the absence of recombination, variants at linked loci will be dragged to fixation as the favorable allele sweeps through the population so that variability at these loci is eliminated. With recombination, the magnitude of this hitchhiking effect decreases with the ratio of the frequency of recombination between the selected and neutral locus (c) to the selection coefficient (s) at the selected locus, and it is negligible when this ratio is on the order of 1. The effects of hitchhiking by favorable mutations are thus only likely to be manifest at sites that are closely linked to the target of selection. There are several natural examples of such hitchhiking events, including an increased frequency of a DNA sequence variant linked to the sickle-cell anemia mutation.

However, hitchhiking effects can also be caused by deleterious mutations, a process termed background selection (Hartl and Clark, 1997; Li, 1997) A neutral variant which is tightly linked to a deleterious mutation that is in the process of elimination from the population will have a high chance of being eliminated before it can unhitch itself by recombination. As discussed previously, the genomes of higher organisms contain many loci subject to mutation to deleterious alleles; therefore, the relatively weak effects of selection against mutations at single loci may have a large cumulative effect on neutral variability.

Studies of genetic variation in *Drosophila* have shown a strong relationship between the amount of recombination in the region where a gene is located and the level of genetic variation that it exhibits (Hartl and Clark, 1997). Similarly, species or populations of hermaphroditic organisms with high rates of self-fertilization, in which the absence of heterozygotes means that genetic recombination is effectively absent, seem to also show reduced levels of molecular variation. Both hitchhiking by favorable mutations and background selection can account for these patterns; current research is attempting to distinguish between them (and other possible explanations).

C. Effects of Linkage to Balanced Polymorphisms

Selection at linked sites can also cause increased variability at neutral sites if selection is balancing rather than directional in nature. This can be understood in terms of a single locus with two alleles, A and a, maintained by strong balancing selection in a randomly mating population. This means that the population is effectively divided into two subpopulations represented by the two allelic classes. If the two alleles are equally frequent, a neutral site linked to this locus with recombination frequency c has a probability $0.5c$ of crossing over from one subpopulation to the other; this is equivalent to the migration rate m in the case of an island model with just two islands. This leads to an expansion of the coalescent time at the neutral site by a factor of $1 + 1/(4N_e c)$, where N_e is the effective population size. This implies that the maintenance of variability by selection is accompanied by a corresponding increase in nucleotide site diversity at sites that are very closely linked to the target of selection, as seen at the *Adh* locus in *D. melanogaster* and at the MHC loci in mammals (Li, 1997). Increases in neutral variation among subpopulations of a species may also occur at linked sites when there is local selection, causing large among-population differences in allele frequencies in different populations at sites closely linked to the targets of selection. Variation in the behavior of neutral variability along a chromosomal region can thus provide valuable evidence on the action of selection.

D. Diffusion Equations

1. General Considerations

A full treatment of genetic drift must deal with properties other than summary statistics of allele frequency distributions of the type considered so far, particularly if selection is to be studied. The difficulties involved in exact analytical treatments of the properties of genotype frequency distributions are great, so resort is usually made to approximations involving the use of diffusion equations, which treat genotype frequencies and time as continuous variables and assume that all evolutionary forces are sufficiently weak that second-order terms in their effects on frequencies are negligible. The standard approach is to assume a single biallelic locus. Using the

continuity assumptions, the state of the population can be described by either of the following two partial differential equations, writing $\phi(x, p, t)$ for the probability density of the frequency, x, of allele A at time t, given initial frequency p:

$$\frac{\partial \phi}{\partial t} = \frac{1}{2}\frac{\partial^2(\phi V_{\delta x})}{\partial x^2} - \frac{\partial(\phi M_{\delta x})}{\partial x} \tag{9a}$$

and

$$\frac{\partial \phi}{\partial t} = \frac{V_{\delta p}}{2}\frac{\partial^2 \phi}{\partial p^2} + M_{\delta p}\frac{\partial \phi}{\partial p}, \tag{9b}$$

where $M_{\delta x}$ is the expected change in gene frequency and $V_{\delta x}$ is the variance of the change in gene frequency, both conditioned on x. $M_{\delta x}$ can be equated to the change in gene frequency in an infinite population; under random sampling of allele frequencies, $V_{\delta x} = x(1 - x)/(2N_e)$.

Equation (9a) is the forward Kolmogorov equation and Eq. (9b) is the backward Kolmogorov equation (Crow and Kimura, 1970; Ewens, 1979). Multidimensional versions of these equations describe systems with multiple alleles and multiple loci but are usually difficult to use.

2. Stationary Distributions

Equation (9a) is most useful for describing the probability density function. However, even for the simple case of a biallelic locus, a full general solution of this equation to yield the density as a function of time has been obtained only for some special cases, such as no selection. It is most useful for studying the properties of stationary distributions of gene frequency, when drift comes into statistical equilibrium with mutation, migration, and selection.

The study of such distributions has led to some important conclusions, most notably that selection is effective at countering the effects of drift in a randomly mating population when $N_e s$ is $\gg 1$ but is ineffective when $N_e s$ is $\ll 1$. When the first condition holds, there is little scatter around the mode of the gene frequency distribution; when the second condition holds, there is a high probability that the population is close to fixation, or fixed, for the disfavored allele. If $N_e s$ is on the order of 1, both drift and selection are significant forces.

3. Fixation Probabilities

The previous conclusion can also be derived by consideration of the probability of fixation of an allele, $P(p)$, for which a general formula was found by Kimura using the backward equation (Crow and Kimura, 1970; Ewens, 1979). With intermediate dominance, such that $h = 0.5$ in Eq. (1), and assuming that a single copy of the mutation is present initially ($p = 1/[2N]$), this formula simplifies as follows, first discussed by Fisher (1930):

$$P\left(\frac{1}{2N}\right) = \frac{1 - \exp\left(-\frac{N}{N_e}s\right)}{1 - \exp\left(-2N_e s\right)}. \tag{10}$$

When $s > 0$, the fixation probability tends to $(N/N_e)s$ as N tends to infinity. When $N_e = N$ (so that the Wright–Fisher model applies), this is equivalent to the branching process result for the survival probability of a favorable mutation in a very large population (see Section II,C,3). The asymptote is approached when $N_e s \gg 1$ so that a selection coefficient larger than $1/N_e$ is sufficient to ensure that a new favorable mutation behaves as though the population is infinite. Conversely, a selection coefficient of this order ensures that a deleterious mutation ($s < 0$) is almost certain to be eliminated from the population.

This led Fisher to conclude that random genetic drift is unlikely to be effective as an evolutionary force on the grounds that (i) most species have numbers of breeding individuals at least in the tens of thousands, and usually in the hundreds of thousands or even millions, and (ii) it is unlikely that a gene would have such a small effect on the phenotype that its average effect on fitness over evolutionary time would be on the order of 10^{-4} or less. Although a compelling argument with regard to genetic changes that affect the phenotype, this view has been challenged by the neutral theory of molecular evolution. As already noted, the causes of protein sequence evolution remain controversial, but studies of the statistical properties of between- and within-species patterns of silent substitutions at third coding positions support the idea that this is affected by both drift and selection in bacteria, yeast, and *Drosophila*. There is a tendency for certain triplets that code for the same amino acid to be favored by selection, but there is also evidence that this selection is so weak that disfavored nucleotide changes can drift to fixation within a species. Estimates of the intensity of selection can be obtained by comparing observed distributions of frequencies of silent-site variants with those predicted by solutions of the diffusion equations (Hartl and Clark, 1997).

E. Muller's Ratchet

Selection may become ineffective if recombination is restricted among the loci subject to selection as a

result of linkage disequilibrium generated by random drift (the Hill–Robertson effect; Barton and Charlesworth, 1998). One process that has been much studied in this context is Muller's ratchet, which operates in a finite population subject to mutation to deleterious alleles at many loci. Consider, for example, the case of a haploid asexual population in which mutations occur exclusively from wild-type to deleterious alleles but not in the opposite direction (see Section II,F) and where $N_e s \gg 1$. If the selective effects of mutations at different loci are identical, a population can be characterized by the frequencies of genomes containing 0, 1, 2 . . . mutations. If the frequency of the mutational class containing the lowest number of mutations (the least-loaded class) is sufficiently small, it will be lost from the population after a finite number of generations. Given the assumed irreversible nature of mutation and the lack of opportunity for genetic recombination, the least-loaded class cannot be reconstituted and will be replaced by the class with one more mutation. This class is now vulnerable to stochastic loss in the same way. There is thus a repetitive process of loss of successive least-loaded classes, in which the loss of each class can be regarded as a turn of Muller's ratchet. The ratchet can only operate at an evolutionarily significant rate if the equilibrium number of individuals in the least-loaded class in an infinite population (n_0) is relatively small (e.g., <100); if it does operate, there will be an approximately exponential decline in the mean fitness of the population. Given the results discussed previously, which suggest that a typical *Drosophila* individual may carry 30 or more detrimental alleles so that the frequency of mutation-free individuals is approximately $\exp(-30) = 9 \times 10^{-14}$, even a very large asexual population of a higher organism is likely to be vulnerable to the operation of the ratchet, leading to its eventual extinction. Asexual prokaryotes, with their much smaller genomes and enormous population sizes, are unlikely to suffer this fate. This may account for the fact that very few species of higher organisms, with their large genomes, are asexual, whereas prokaryotes and mitochondria have very low levels of recombination (Barton and Charlesworth, 1998). It may also contribute to the evolutionary degeneration of Y chromosomes, which are usually largely devoid of active genes.

F. Group Selection

Another possibility that violates the supremacy of individual selection occurs when species are subdivided into local populations, among which migration is so restricted in relation to the reciprocal of local population size that drift can overcome its homogenizing influence (see Section IV,A). An allele that is deleterious within its local population may drift to fixation locally in opposition to selection since N_e for the local population is much smaller than for the species as a whole. This raises the possibility of group or interdeme selection. If the allele in question causes its carriers to be altruistic in the sense of conferring increased fitness on the members of their deme at the expense of suffering a loss in fitness to themselves, demes in which the allele reaches high frequency or fixation will achieve a higher mean fitness. This may render them less susceptible to extinction and more able to contribute to the pool of migrants or to found new demes. Selection among demes can therefore result in the spread of an allele that causes a loss in individual fitness but benefits the population as a whole (Haldane, 1932; Hartl and Clark, 1997).

Although this is a theoretically viable mechanism, there are reasons to doubt that it is widely applicable to evolution in nature. First, studies of molecular genetic variation within and among populations indicate that many species lack extensive differentiation among local populations (see Section IV,A,2), suggesting that migration is so effective that the necessary conditions for group selection are often not met. Second, it does not seem capable of producing a stable evolutionary outcome: If an altruistic allele becomes fixed in a set of populations, a "selfish" opponent that is reintroduced by migration or mutation will have a good chance of spreading through the species since the outcome of the conflict between group and individual selection is probabilistic and not deterministic.

G. Kin Selection

Altruistic behavior, such as the sterility of the worker castes in social insects, is generally believed usually to be due to kin selection. This is based on the fact that an altruistic genotype that aids relatives may experience a selective advantage, even in a large, randomly mating population, if the fitness benefit b to the recipients of the altruism is sufficiently large in relation to the cost c to the altruists (Fisher, 1930; Haldane, 1932; Hartl and Clark, 1997). This is because the genetic similarity of related individuals means that the relatives of an altruist have a greater chance than average of carrying an allele that promotes altruism; according to Hamilton's rule, there can be an increase in the frequency of an altruistic allele

when $br > c$, where r is a measure of the relatedness of the altruists to the recipients. There is an extensive theoretical literature on the correct way to calculate r under different conditions; this result has played an important role in the evolutionary interpretation of social behavior in animals.

H. The Shifting Balance Theory

The second objection to group selection is overcome by the related model of Wright, the shifting balance theory of evolution (Wright, 1977; Hartl and Clark, 1997). Wright postulated that epistatic interactions in fitness among alleles at different loci are widespread, resulting in multiple stable equilibria under selection. The simplest case is a haploid two-locus, two-allele model, with ab and AB both fitter than Ab and aB. Fixation for ab or AB is stable against introduction of Ab and aB, whereas fixation for Ab or aB is unstable to the introduction of ab and AB. With constant fitnesses and loose linkage, locally stable equilibria are approximated by the peaks in the surface of mean fitness as a function of the allele frequencies at the loci concerned (see Sections II,C,1 and II,H,2). A population will approach the peak that is the nearest attractor rather than the highest peak in the landscape. Genetic drift can cause a local population to travel down the adaptive valley separating the current equilibrium to the zone of attraction of a neighboring peak, and selection can then bring it to the new equilibrium. If this is associated with a higher mean fitness than the surrounding peaks, the process of interdeme selection can cause the species as a whole to acquire the genotype associated with this peak, and hence improve in mean fitness. In contrast to the group selection model of altruism, the new equilibrium is dynamically stable, and so there is only a low probability of reversal. This process has the attractive feature that it allows the species to acquire adpatively superior genotype or character combinations that would require multiple simultaneous mutations to be produced in a large population. Its drawback is that it requires a delicate balance between restricted migration, local population size, and the nature and intensity of selection if it is to operate with any frequency. It is difficult to distinguish the end products of the shifting balance process from ordinary individual selection, which does not have such stringent requirements, and it is unclear to what extent it has played an important role in adaptive evolution (Kimura, 1983).

V. CONCLUSIONS

This article necessarily omitted many important topics. It focused on the basic general principles governing evolutionary change in populations; space did not permit more than a brief mention of the application of these principles to wider biological problems, including life history evolution, the evolution of genetic and sexual systems, the evolution of social behavior, speciation, and the interpretation of macroevolution. Very little useful can be said about natural variation and evolution at almost any level without taking population genetic concepts into account.

See Also the Following Articles

DIVERSITY, MOLECULAR LEVEL • EVOLUTION, THEORY OF • INBREEDING AND OUTBREEDING • PHENOTYPE, A HISTORICAL PERSPECTIVE

Bibliography

Barton, N. H., and Charlesworth, B. (1998). Why sex and recombination? *Science* **281**, 1986–1990.

Barton, N. H., and Turelli, M. (1991). Natural and sexual selection on many loci. *Genetics* **127**, 229–255.

Bruford, M. W., and Wayne, R. K. (1993). Microsatellites and their application to population genetic studies. *Curr. Opin Genet. Dev.* **3**, 939–943.

Crow, J. F. (1993). Mutation, mean fitness, and genetic load. *Oxford Surv. Evol. Biol.* **9**, 3–42.

Crow, J. F., and Kimura, M. (1970). *An Introduction to Population Genetics Theory*. Harper & Row, New York.

Drake, J. W., Charlesworth, B., Charlesworth, D., and Crow, J. F. (1998). Rates of spontaneous mutation. *Genetics* **148**, 1667–1686.

Ewens, W. J. (1979). *Mathematical Population Genetics*. Springer-Verlag, Berlin.

Falconer, D. S., and Mackay, T. F. C. (1996). *An Introduction to Quantitative Genetics,* 4th ed. Longman, London.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford Univ. Press, Oxford. (2nd ed., 1958, Dover, New York)

Ford, E. B. (1975). *Ecological Genetics,* 4th ed. Chapman & Hall, London.

Haldane, J. B. S. (1932). *The Causes of Evolution*. Longmans Green, London.

Hartl, D. L., and Clark, A. G. (1997). *Principles of Population Genetics,* 3rd ed. Sinauer, Sunderland, MA.

Hill, A. V. S., and Weatherall, D. J. (1998). Host genetic factors in resistance to malaria. In *Malaria Parasite Biology, Pathogenesis and Protection* (I. W. Sherman, Ed.), pp. 445–455. ASM Press, Washington, D.C.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge Univ. Press, Cambridge, UK.

Lande, R. (1988). Quantitative genetics and evolutionary theory. In *Proceedings of the Second International Conference on Quantitative Genetics* (B. S. Weir, E. J. Eisen, M. M. Goodman, and G. Namkoong, Eds.), pp. 71–84. Sinauer, Sunderland, MA.

Lewontin, R. C. (1974). *The Genetic Basis of Evolutionary Change*. Columbia Univ. Press, New York.

Li, W.-H. (1997). *Molecular Evolution*. Sinauer, Sunderland, MA.

Maruyama, T. (1977). *Stochastic Problems in Population Genetics*, Lecture Notes in Biomathematics No. 17. Springer-Verlag, Berlin.

Nagylaki, T. (1986). Neutral models of geographical variation. In *Stochastic Spatial Processes* (P. Tautu Ed.), Lecture Notes in Mathematics No. 1212, pp. 216–237. Springer-Verlag, Berlin.

Wright, S. (1977). *Evolution and the Genetics of Populations. Vol. 3. Experimental Results and Evolutionary Deductions*. Univ. of Chicago Press, Chicago.