

# Evolutionary Genetics

(for *Encyclopedia of Biodiversity*)

Sergey Gavrilets

Departments of Ecology and Evolutionary Biology and Mathematics, University of Tennessee,  
Knoxville, TN 37996-1610 USA

**Evolutionary genetics** studies the patterns and mechanisms of genetic changes underlying evolutionary change. Because biological diversity at all levels (including population, species, and ecosystem diversity) has a genetic basis and is a result of evolution, evolutionary genetics provides insight into the ultimate mechanisms creating biodiversity. Evolutionary genetics utilizes empirical tools and mathematical methods and models. Evolutionary genetics has an important impact on many other fields including medicine, anthropology, psychology, agriculture, conservation, economics, and philosophy.

## Major questions of evolutionary genetics

Among the most important questions that evolutionary genetics attempts to answer are: What are the patterns of genetic variation within and among populations and species? What are the forces responsible for maintaining genetic variation within populations? What are the forces driving evolutionary change and, ultimately, the origin of new species? Why are there so many (or, perhaps, so few) species? How did they become so well adapted to their environments? What is the time scale for evolutionary change? What is the correct evolutionary tree for a set of species? What is the correct evolutionary history of life? Where do new genes and characters come from? Why do sex and recombination occur at all and what are their evolutionary consequences? Where did humans come from? To what extent are humans different from other organisms?

## Experimental tools of evolutionary genetics

The most common empirical tools and techniques of evolutionary genetics are (i) classical genetic analysis of the distribution of specific genes and traits among offspring of particular crosses, (ii) artificial selection experiments in which only organisms possessing certain characteristics contribute to the next generation, (iii) studies of variation in enzymes in natural populations (enzymes are proteins that act as catalysts in biochemical reactions), and (iv) studies of variation in molecular sequences (such as amino acid sequences, RNA and DNA sequences) between different organisms, populations and species. Molecular methods have become the most powerful tool of modern evolutionary genetics. Experimental evolutionary genetics allows one to both quantify the extent of genetic differences and estimate important evolutionary parameters (e.g. mutation rates, the distributions of effects of mutations on different characteristics including fitness components, strength of selection, rates of migration etc).

## Theoretical population genetics

The effects of different factors (such as mutation, random genetic drift, selection etc) on biological populations within one or few generations are typically very small, and the time-scale for most evolutionary change is rather long. This makes it very difficult to observe and/or measure these changes directly. Moreover even the simplest biological organisms such as bacteria have thousands of genes that can be in many different states (alleles). This results in an enormous potential for genetic diversity. For example, the number of different genetic combinations that are possible with  $L$  genes each with just two alleles is  $2^L$ . If  $L = 1000$  (which is a reasonable estimate for simple bacteria), there are  $2^{1000} \approx 10^{301}$  genetic combinations. This number is much larger than estimates of the number of elementary particles in our universe. In many cases biological intuition is not helpful in evaluating quantitative and/or qualitative

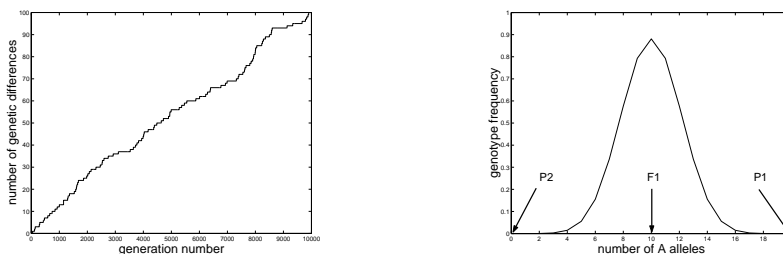


Figure 1: Effects of mutation and recombination. (A) Dynamics of genetic divergence between two isolated populations by mutation. (B) The distribution of the number of **A**-alleles among the offspring of  $F_1$  organisms achieved by recombination. Arrows mark the genotype  $F_1$  and its parents  $P_1$  and  $P_2$  (see text for details).

effects of different biological factors on genetic systems that are that complex. Difficulties in direct observations and enormous genetic complexity make it necessary to use mathematical models and methods to analyze evolutionary change at the genetic level.

Theoretical population genetics provides a mathematical foundation for the study of evolutionary genetics. The common procedure of theoretical population genetics is to start with some simple mathematical models that although not being fully realistic, can be completely analyzed and then refined into more realistic models which can be used to answer specific evolutionary questions. There are two general areas of theoretical population genetics: a prospective theory and a retrospective theory. The prospective theory takes the current state of a population (species, ecosystem) as given and tries to predict the relevant biological properties in the future. The prospective theory mostly uses methods developed within the mathematical theory of the dynamical systems (deterministic and stochastic). In the retrospective theory, one observes the current state of a population (species, ecosystem) and ask how it got here. The retrospective theory heavily relies on statistical methods. Theoretical population genetics is essential for interpreting genetic variation, for predicting evolutionary change, and for reconstructing evolutionary history. It also provides a foundation for understanding the evolution of different characteristics, such as life histories and genome structure.

## Effects of major evolutionary factors on genetic variation

Evolutionary genetics has established that most natural populations have very high levels of genetic variation. A number of evolutionary factors control both the maintenance of genetic variation within populations (species, ecosystems) and the rates of genetic divergence between different populations (species, ecosystems). Among the most important factors are mutation, recombination, random genetic drift, gene flow, and natural and sexual selection. Mutation and recombination are the primary factors introducing genetic variation into the populations. Random genetic drift, gene flow, and natural and sexual selection act on this variation and control which genetic variants will be preserved and which will be removed from the population. Sufficient genetic divergence (cumulative or in specific traits or genes) results in the emergence of a new species.

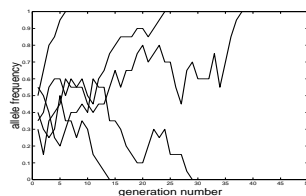


Figure 2: Effects of random genetic drift on allele frequencies in an asexual diploid population. Different lines describe the change in the frequency of allele **A** in five different simulated populations that had the same initial frequency of (equal to 0.4). The frequency of **A** quickly reaches one ( here, in populations 1, 2 and 3) or zero (here, in populations 4 and 5) meaning the disappearance from the population of allele **a** or **A**, respectively.

### Mutation and recombination

Mutation is a sudden heritable change in the genetic material, most often appearing as an alteration of a single gene by replacement, duplication or deletion of a number of DNA base pairs. Mutation can sometimes alter the structure or number of genes or entire chromosomes. Most mutations are harmful, but some are advantageous. Mutations occur naturally at low rates ( $10^{-5} - 10^{-6}$  per gene per generation). Mutation is considered to be the major factor limiting the speed of evolution. Isolated populations tend to accumulate different mutations diverging genetically. This is illustrated in Figure 1A which describes the dynamics of genetic divergence by mutation between two hypothetical populations that were identical genetically before their physical separation. It is assumed that a new mutation is fixed in one of the subpopulations with a small probability 0.001 per generation. The graph gives the number of mutations that differ between the populations as a function of time. Note that the number of genetic differences (“genetic distance”) increases approximately linearly with the time since separation and, thus, genetic distance can serve as a proxy for time (“molecular clock”) This property allows one to calculate the time to a common ancestor on the basis of genetic distance. This technique is widely used for reconstructing phylogenetic relationships between different groups of organisms.

Recombination is the exchange of genes between paternal and maternal chromosomes that occurs when reproductive cells are formed. Recombination results in offspring that have a combination of genes (characteristics) different from that of their parents. Recombination can potentially produce an enormous variety of new genotypes. This is illustrated in Figure 1B using a hypothetical example. It is assumed that one diploid parental organism,  $P_1$ , has alleles **A** in 10 specific loci (locus is a position on a chromosome occupied by a gene, “loci” is the plural form for “locus”). Because each diploid organism has two copies of each gene (inherited from its father and mother), the overall number of **A**-alleles is 20. Another parental organisms,  $P_2$ , has **a**-alleles in these loci. Thus the overall number of **A**-alleles here is zero. Their offspring, which we denote as  $F_1$ , will all have an intermediate number of **A**-alleles (which is 10). Figure 1B describes the distribution of the number of **A**-alleles among the offspring of  $F_1$  organisms assuming that these organisms mated among themselves and that the loci are unlinked. This distribution includes genotypes with many different numbers of **A**-alleles at relatively high frequencies.

### Random genetic drift

Random genetic drift is random changes in genotype frequencies that result from variation in the number of offspring between individuals and (in sexual organisms) from the inherent stochasticity of the process of production of reproductive cells. Random genetic drift affects biological evolution in two important ways.

First, it tends to remove genetic variation from populations. This is illustrated in Figure 2 using five hypothetical populations of diploid organisms each with one locus with two alleles (**A** and **a**). It is assumed that the population size is 20 individuals and that initial frequency of allele **A** is 0.4. In Figure 2 different lines show the dynamics of the frequency of allele **A** in each population. Allele frequencies fluctuate from generation to generation until genetic variation is completely lost (that is the allele frequency becomes equal to one or zero). Under random genetic drift, this happens in a time span comparable with the population size. Second, random genetic drift affects the probability of survival of new mutations in the population (see the section on natural selection below). The effects of random genetic drift are inversely proportional to the population size so that it is most important in small populations and is negligible in large populations. If mutation and random genetic drift are the only forces acting on a population, the population will continuously accumulate new mutations with the rate equal exactly to the mutation rate.

### Gene flow

Most species are composed of many local populations that occasionally exchange migrants. Gene flow between local populations is the exchange of genetic material by interbreeding between migrants and members of a local population. Gene flow increases the variation in the genetic composition of a population (because migrants bring “foreign” genes into the population) and decreases genetic differences between the populations exchanging migrants. In the absence of other factors, the time until the populations exchanging migrants become genetically similar is proportional to the inverse of the migration rate. Gene flow prevents continuous genetic divergence of populations that would happen by mutation and random genetic drift. The joint action of mutation, random genetic drift and gene flow results in complex genetic variation within geographically structured populations.

### Natural selection

Natural selection is the non-random and consistent differences between different genotypes in fitness (i.e. in survival and reproduction). Natural selection acts to preserve genotypes that have higher fitness and to eliminate genotypes that have lower fitness. Natural selection is viewed as a creative force that directs the course of evolution by preserving genotypes or traits best adapted in the face of environmental conditions and competition. Natural selection causes adaptation that is improvement in function. Natural selection can act on tiny differences in fitness. This is illustrated in Figure 3a that shows the changes in the frequency of a slightly advantageous mutant that is initially rare in a hypothetical asexual population. It is assumed that the ancestral genotype has relative fitness one. [In this example, “fitness” means “viability” that is the probability to survive to the age of reproduction.] The mutant genotype has slightly higher fitness ( $w = 1.001$ ) and a very low initial frequency ( $p = 0.00001$ ). Figure 3a shows that by generation *ca.* 18000 mutant genotype replaces the ancestral genotype. Another illustration is provided in Figure 3B using a hypothetical asexual population that initially has five different genotypes at equal frequencies. In this figure, five different lines describe the frequencies of five different genotypes with fitnesses  $w_1 = 0.96$ ,  $w_2 = 0.97$ ,  $w_3 = 0.98$ ,  $w_4 = 0.99$  and  $w_5 = 1.00$ . Here, the differences in fitness between different genotypes are larger than in the previous example meaning that natural selection is stronger. The least fit genotypes (genotypes one and two) practically disappear from the population by generation 100. The fittest genotype (genotype five) dominates by generation 500. Figure 3B shows that if natural selection is strong, biological evolution can be very rapid.

In large populations, natural selection will quickly establish advantageous genotypes and eliminate disadvantageous genotypes (as illustrated in Figure 3). In contrast, in small populations random genetic drift will interfere with natural selection resulting in the possibility of a stochastic loss of advantageous genotypes and establishment of slightly deleterious genotypes. Accumulation of deleterious mutations in small populations may result in their extinction.

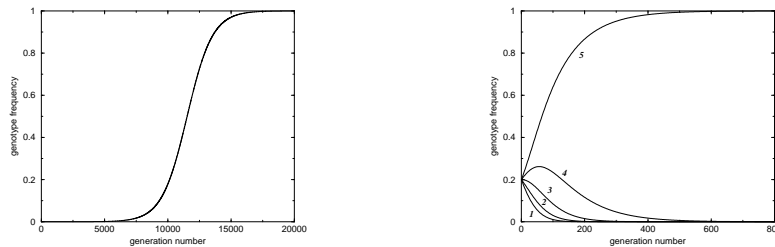


Figure 3: Effects of natural selection on allele frequencies in an asexual population. (A) Establishment of a slightly advantageous mutant that is initially rare. (B) Strong selection in a polymorphic population.

### Sexual selection

Sexual selection is the non-random and consistent differences between different genotypes of the same sex in their ability to mate with (or fertilize) the organisms of the other sex. Sexual selection can be as strong as natural selection and can result in rapid genetic divergence of different populations. Sexual selection is the cause of the evolution of many secondary sexual characteristics, particularly in male animals, such as bright colors, elaborate ornaments and conspicuous songs that are not useful or even deleterious for survival. Sexual selection is thought to be a major cause of speciation in some groups of organisms.

## Implications for biodiversity

Two major evolutionary mechanisms creating biodiversity are speciation (that is the origin of new species) and divergent evolution. Evolutionary genetics has identified the most important factors affecting these mechanisms and their time-scales. Mutation, recombination, spatial structure and gene flow, natural and sexual selection are all very important in generating and maintaining biodiversity. The level of biodiversity depends on a delicate balance of deterministic and stochastic forces acting on biological systems. Evolutionary genetics has shown that although under certain conditions biodiversity can be created rapidly, under the most likely conditions the processes of generating biodiversity are very slow.

## References

1. Bell, Graham. 1997. *Selection: the Mechanism of Evolution*. New York: Chapman & Hall.
2. Hartl, Daniel L. and Andrew G. Clarke. 1997. *Principles of Population Genetics*. Sunderland: Sinauer Associates.
3. Futuyma, Douglas J. 1998. *Evolutionary Biology*. Sunderland: Sinauer Associates.
4. Li, Wen-Hsiung. 1997. *Molecular Evolution*. Sunderland: Sinauer Associates.
5. Maynard Smith, John. 1998. *Evolutionary Genetics*. Oxford: Oxford University Press.
6. Nei, Masatoshi. 1987. *Molecular Evolutionary Genetics*. New York: Columbia University Press.

7. Ridley, Mark. 1993. *Evolution*. Boston: Blackwell Scientific Publications.
8. Stearns, Stephen C. and Rolf F. Hoekstra. *Evolution: an Introduction*. Oxford: Oxford University Press.