

Genetic distance is a standard concept, of which the most common measure is Nei's index D . It is calculated like this. We take two populations and measure the frequency of both of the alleles at several loci. Consider an allele A_i at a locus in population 1 and 2, and define its frequency as p_i in population 1 and q_i in population 2. (By the usual convention, i has as many values as there are alleles at a locus: if there are two alleles, $i = 1$ for one of them and $i = 2$ for the other.) First calculate $j_1 = \sum p_i^2$, $j_2 = \sum q_i^2$, and $j_{12} = \sum p_i q_i$, where j_1 is the probability that two randomly chosen genes in population 1 are identical, j_2 is the same probability for population 2, and j_{12} is the probability that two genes, one drawn randomly from population 1 and the other from population 2, are identical. (j_1 is more easily recognized as the frequency of all homozygotes in population 1, j_2 for population 2.) The same set of three numbers can be calculated for every locus; the next step is to calculate the average for all loci in each of the three cases. Define the averages as follows: $J_1 = j_1$, $J_2 = j_2$, $J_{12} = j_{12}$. Finally, Nei's formula for genetic distance is:

$$D = -\log_e \frac{J_{12}}{\sqrt{J_1 J_2}}$$

D decreases as more alleles are shared between the two populations, making the product J_{12} higher. In the extreme case of two samples from one population, $J_1 = J_2 = J_{12}$, and $D = -\log_e 1 = 0$. As shared alleles are lost, D increases. D has no mathematic upper limit (in the extreme case of two samples that share no alleles, $J_{12} = 0$, $D = -\log_e 0 = \infty$); but in practice, it is found to have values of about 0.1–2 for different species, and of about 1 for different genera.

The formula can also be expressed $D = -\log_e I$, where $I = J_{12}/\sqrt{(J_1 J_2)}$. I is called the genetic identity between the two populations (see Table 16.1).

Because many proteins evolve at least

approximately in the manner of a molecular clock (see chapter 7), genetic distance increases approximately in proportion to the time since the two species diverged.

Postzygotic isolation. If two species can be crossed in the laboratory, some or all of their offspring may be inviable or sterile. We saw in chapter 14 that it is often the heterogametic sex that is inviable or sterile, a generalization known as Haldane's rule. There are two sorts of cross between two species: either a male of species 1 and a female of species 2, or a male of species 2 and a female of species 1. According to its sex and the type of cross, an offspring can be one of four types: male and from male species 1 \times female species 2; female and from male species 1 \times female species 2, etc. For a species pair we count the number of the four types that are sterile or inviable and divide the result by four. The index can therefore have values 0, 0.25, 0.5, 0.75, 1 in order of increasing isolation. For a pair of species that fit Haldane's rule, the index is 0.5.

Prezygotic isolation. Prezygotic isolation is generally measured in a mate choice experiment in which an individual of one species has a choice of mates from its own, and another, species. If it mates with its own species, the mating is called homospecific; if with the other, heterospecific. As an index of prezygotic isolation, Coyne and Orr used

$$\text{Prezygotic isolation} = 1 -$$

$$\frac{\text{Frequency of heterospecific matings}}{\text{Frequency of homospecific matings}}$$

The index can vary from 1 (complete isolation) to 0 (no isolation) through, in theory, to $-\infty$ for complete disassortative mating. In practice the index varies between 0 and 1 for most pairs of species. (In Table 16.3 a slightly different isolation index was used.)