

Question. Note that simple selection model consider a single loci with two candidate alleles. This is likely to be too simple in real life. State the assumptions behind this model and list common modifications accordingly.

Answer. In the simplest selection model consisting of a single locus and two candidate alleles (alleles A and a , say), the frequency of allele A evolves according to

$$\frac{dq_A}{dt} = \sigma q_A(1 - q_A)$$

However, selection usually occurs over multiple loci. We can extend the basic to a model where we have two loci at i and j that can have the candidate alleles 1 and 0. Then, assuming that fitnesses are additive (i.e. if having allele 1 and locus i has a fitness advantage of σ_i and having allele 1 at locus j has a fitness advantage of σ_j , then $f_{ij}^{11} = \sigma_i + \sigma_j$), the frequency of allele 1 at locus i evolves according to:

$$\frac{dq_i^1}{dt} = \left[\sigma_i + \sigma_j \left(\frac{q_{ij}^{11}}{q_i^1} - \frac{q_{ij}^{01}}{q_i^0} \right) \right] q_i^1(1 - q_i^1)$$

We can further extend this model to reflect selection over more than one locus.

$$\frac{dq_i^1}{dt} = \left[\sigma_i + \sum_k \sigma_k \left(\frac{q_{ik}^{11}}{q_i^1} - \frac{q_{ik}^{01}}{q_i^0} \right) \right] q_i^1(1 - q_i^1)$$

Of course, genes that are being selected over often interact with each other (epistasis) and therefore fitness is not necessarily additive and selection cannot always be modeled this way.

When selection occurs over multiple loci in a haploid species, we begin to see phenomena such as clonal interference, i.e. when a beneficial allele is out-competed by another and genetic hitchhiking, i.e. when a beneficial allele carries a deleterious allele to fixation.

The basic two-allele model also assumes a haploid individual and therefore does not account factors present in diploid organisms, such as dominance in heterozygotes. For an single locus, two allele model where individuals with genotype AA have fitness $1 + 2s$, individuals with genotype Aa have fitness $1 + 2hs$ and individuals with genotype aa have fitness 1, the frequency of allele A evolves according to

$$\frac{dq_A}{dt} = 2\sigma q_A(1 - q_A)(q_A + h(1 - 2q_A))$$

Similar to the haploid case, this model can be extended to multiple loci to account for further complexity in gene selection.

We may also like to account for the effects of linkage disequilibrium, where there is non-random association of alleles at two or more loci. For example, in a driver-passenger model the evolution of a passenger mutation j evolves according to

$$\frac{dq_j^1}{dt} = \sigma D_{ij}$$

where the gene at allele i is the driver mutation.

Within a species, there also may be frequency-dependent selection, where the fitness of an individual is dependent on the proportion of other individuals in the population that are genetically similar.

Beyond selection, the simple model fails to account for other forms of genetic variation such as mutation and genetic drift.