

The genetic theory of natural selection

1 Fitness

An individual's absolute fitness is the number of offspring produced over its lifetime. It is often useful to consider the average fitness of individuals of specific genotype/phenotype, and it is useful to simplify absolute fitness (denoted by W) as the product of two components: the probability that the individual survives to maturity, and the expected number of offspring if the individual does survive.

The strength of selection is determined by fitness differences. Therefore, it is convenient to work with relative fitness, which is the absolute fitness divided by a fitness of reference.

2 Positive selection: the spread of beneficial mutations and the elimination of variation

Whenever one allele has higher fitness than others, natural selection will favour its spread through the population, which is called positive selection. Consider a single locus with two alleles A and a under Hardy-Weiberg equilibrium and the allele frequency of a is p_0 . Assume a increases fitness: for each a copy that an individual carries, its fitness increases by a proportion of s . The relative fitness for AA , Aa and aa , are 1, $1 + s$ and $1 + 2s$, respectively. The genotype of AA , Aa and aa in the first generation is $\frac{(1-p_0)^2}{\bar{w}}$, $\frac{2p_0(1-p_0)(1+s)}{\bar{w}}$ and $\frac{p_0^2(1+2s)}{\bar{w}}$, respectively, where $\bar{w} = (1 - p_0)^2 + 2p_0(1 - p_0)(1 + s) + p_0^2(1 + 2s)$ is the mean fitness of the 0th population. The allele frequency of a in the first generation is $p_1 = \frac{p_0[1+s(1+p_0)]}{\bar{w}}$, and $\Delta p = p_1 - p_0 = \frac{sp_0(1-p_0)}{\bar{w}} \approx sp_0(1 - p_0)$ since s is often very small (say, below 0.1). The key implications here are: (1) the rate of evolution caused by selection is proportional to the strength of selection measured by s , which is called selection coefficient; (2) the rate of evolution caused by selection is proportional to $p(1 - p)$, which represents the genetic variation at the locus. (3) the beneficial allele a will spread through the population and fixed, which is called selective sweep. As a rough estimation, a beneficial allele will increase frequency from 0.1 to 0.9 in about $\frac{4}{s}$ generations.

Dominance of beneficial allele speeds its spread driven by natural selection, as heterozygotes share the full fitness benefit that the homozygotes have, especially when it is rare. However, when the beneficial

allele is common, almost all individuals have high fitness and there is very little variation in fitness, and selection has little power to increase the frequency of beneficial allele further.

3 Evolutionary side effect

Genetic correlations occur when two traits tend to be inherited together and cause evolutionary side effect: one phenotype can be preferred by natural selection because of its correlation with another beneficial phenotype, while itself can be neutral or even deleterious. Therefore, an allele that increase fitness through its effect on one trait can decrease fitness due to its effect on another trait, a phenomenon called evolutionary trade-off. Also, an allele that has no effect on fitness can also spread by natural selection by association with a beneficial allele of a second locus, which is called hitchhiking. In hitchhiking, a beneficial mutation first appear, and it is in perfect linkage equilibrium with all other alleles on its chromosome. As the mutation spreads, recombination breaks down the linkage disequilibrium. However, on the chromosome, there are more recombinations between distant sites than neighboring sites. So sites close to the selected locus do not have a chance to recombine before the mutation is fixed. As a result, these sites carry the same alleles that were on the original chromosome where the mutation appeared. Variation at these sites is eliminated.

4 Balancing selection: the maintainance of variation

Balancing selection maintains genetic variation within a population. One form of balancing selection is overdominance, which occur when the heterozygote have higher fitness than both homozygotes. Consider a locus with two alleles A and S , and the relative fitness of AA , AS and SS are $w_{AA} < 1$, 1 and $w_{aa} < 1$, respectively, and assume that natural selection is the only cause of evolution. The population approaches to polymorphic equilibrium, with the frequency of S given by $p = \frac{1-w_{AA}}{2-w_{AA}-w_{SS}}$.

Another form of balancing selection is frequency-dependent selection, which occur when the fitnesses of alleles change depending on their own frequencies. In some cases, an allele gets a fitness advantage when it is rare, which is called negative frequency dependence.

A third form of balancing selection occur when different genotypes specialize on different ecological niches. In such cases, each genotype is partly shielded from competition with other genotypes and has its own ecological capacity.

52 **5 Selection with historical contingency**

53 Positive selection leads to the fixation of beneficial allele, while balancing selection drives the population
54 evolve towards a constant equilibrium allele frequency no matter where it begins. In selection with
55 historical contingency, variation is eliminated, but which allele spreads to fixation depends on the initial
56 allele frequency.

57 One form of historical contingency is underdominance, which occur when the heterozygote has lower
58 fitness than homozygotes. The scenario is: if the initial frequency of an allele is below a threshold, almost
59 all of its copies are in low fitness heterozygotes and selection drives it out of the population. The threshold
60 is determined by the relatibe fitness of the two homozygotes.

61 Another form is positive frequency-dependent selection, which favours alleles that have high frequency.
62 As a result, it eliminates genetic variation within population.

63 **6 Mean fitness of population (\bar{w})**

64 The mean fitness of a population evolves as the allele frequencies changes. By natural selection, mean
65 fitness of a population increases, and the increase per generation depends on genetic variation of the
66 population (fundamental theorem of natural selection).

67 A complementary perspective on the evolution of fitness is adaptive landscape, in which mean fitness
68 \bar{w} is plotted against allele frequency p . By selection, the population evolve uphill on the landscape, and
69 the allele frequency changes at rate $\Delta p = \frac{1}{2}p(1-p)\frac{d\ln\bar{w}}{dp}$.

70 **7 Purifying selection: removal of deleterious mutations**

71 Deleterious mutations are much more common than beneficial ones. Selection that acts to remove dele-
72 terious mutations from a population is called purifying selection.

73 Purifying selection often fails to eliminate deleterious mutations, as they are being continually intro-
74 duced. This flowof new mutations into the population is offset by natural selection that acts to eliminate
75 them. This situation is called mutation-selection balance. Let the mutated-allele free homozygotes have
76 relative fitness 1, and the heterozygotes have relative fitness $1-s$, and the homozygotes with two mutated
77 alleles have relative fitness $1-2s$. The probability that a copy of normal allele mutates to a deleterious
78 allele in a given generation is μ . When the input of the deleterious allele by mutation balances its removal
79 by selection, the deleterious mutation reaches an equilibrium frequency of $\hat{p} \approx \frac{\mu}{s} (\mu \ll s)$.

80 Mutation load is the proportion by which the mean fitness of the population is reduced by deleterious
81 mutations compared with a hypothetical population without mutations. For a single locus, mutation load

82 is given by $L = 2\frac{\mu}{s}(1 - \frac{\mu}{s})s + (\frac{\mu}{s})^2 2s = 2\mu$.

83 Mutation rates at individual loci are often very small, but there are many loci in a genome. Let U be
84 the average number of new deleterious mutations that are added to the genome each generation. Assume
85 that mutations have independent effects on fitness, the mutation load is given by $L = 1 - e^{-U}$.