Understanding selection

The selection app gives a simple model of selection on the **phenotypes** of Mendelian trait, with phenotypes that are determined by a single gene with two alleles, R and S. It can be used to simulate selection under several different scenarios. This worksheet will guide you through each one.

The app is set up so that you can specify the frequency of the R allele (p), which will also set the frequency of the S allele (q, which is 1-p). You can also set the relative fitness of each genotype (as determined by the phenotype produced), and the number of generations to calculate. The steps done by the app each generation are numbered, and are:

1. Set the allele frequencies of the gametes that produce the initial genotype frequencies – the Generation 0 allele frequencies.
2. Calculate the genotype frequencies of the first parent generation at birth, as Hardy-Weinberg values.
3. Set the relative fitnesses of the phenotypes produced by the genotypes.
4. Multiply the genotype frequencies by the fitnesses – selection dictates what the frequencies of the breeders will be, and this step determines how the genotypes contribute to the next generation, given the fitnesses.
5. Calculate the mean fitness of the population, $\bar{w}$, by summing the frequency x fitness calculations from step 4.
6. Calculate the genotype frequencies of the breeders, which are the parental generation after selection has occurred. This is done by dividing the frequency x fitness values from step 4 by the mean fitness of the population in step 5.
7. Calculate the frequencies of the R and S alleles in the gametes produced by the breeders, using the genotype frequencies of the breeders.
8. Calculate the genotype frequencies of the next generation (Generation 1) from the gamete allele frequencies, using Hardy-Weinberg.

The full set of calculations to go from Generation 0 to Generation 1 are laid out in the Genotypes and Alleles tables. To get 1000 generations of selection these steps are repeated for each generation, using the parent genotype frequencies after selection from one generation to generate the next generation of offspring. This sort of process in which each step calculates the starting conditions for the next step is called **iteration**, and our model of natural selection is an iterative model – to see what the result of selection will be after 1000 generations of selection we need to iterate this process 1000 times.

The result of this iterative process is shown for genotypes (table “Change in genotype frequencies over time”) and alleles (table “Change in allele frequencies over time”). These tables of data are graphed, with generation on the x-axis and frequencies of genotypes or alleles on the y-axis.

With that background, you can simulate different selection scenarios and see how allele and genotype frequencies are expected to change over time. Six scenarios are given below.

1. **Selection against a deleterious recessive S allele, R dominant to S**.

The default settings are set up to simulate this situation, so no changes are needed. The fitness for the genotypes are set to 1 for RR and RS, and to 0.8 for SS. The allele frequency for R (p) is set to 0.5, and thus the allele frequency for S (q) is also 0.5 (because q = 1-p).

S is deleterious because the fitness of SS is (greater than, less than, equal to) the fitness for RR.

R is dominant to S because the fitness for RS is (greater than, less than, equal to) the fitness of RR.

Look at the graphs, and answer the following questions:

A. How does selection against SS over time affect the frequency of:

i. SS?

ii. S (q)?

iii. RR?

iv. RS?

v. R (p)?

B. Set the number of generations to 100 – what is the frequency of the S allele at the 100th generation (hover over the q line at generation 100)?

C. Set the number of generations to 1000 – what is the frequency of the S allele at the 1000th generation? Why doesn't q go to zero, even after 1000 generations? Hint: recessive alleles are shielded from selection in heterozygotes, where are most of the S alleles found when they become rare?

D. What is selected, genotypes, phenotypes, or alleles?

2. **Selection against a deleterious dominant allele, R**.

Start with the number of generations set to 1000.

Set the relative fitnesses to 0.8 for RR, 0.8 for RS, and 1 for SS.

R is deleterious because the fitness of RR is (greater than, less than, equal to) the fitness for SS.

R is dominant to S because the fitness for RS is (greater than, less than, equal to) the fitness of RR.

Look at the graphs, and answer the following questions:

A. How does selection against RR and RS over time affect the frequency of:

i. SS?

ii. S (q)?

iii. RR?

iv. RS?

v. R (p)?

B. Set the number of generations to 100 – what is the frequency of the R allele at the 100th generation (hover over the q line at generation 100)?

C. Set the number of generations back to 1000 – what is the frequency of the R allele at the 1000th generation? Why is selection against a dominant allele more effective at eliminating it than selection against the recessive allele was?

3. **Selection against a deleterious incompletely dominant allele, R**.

Start with the number of generations set to 1000.

Set the relative fitnesses to 0.8 for RR, 0.9 for RS, and 1 for SS.

R is deleterious because the fitness of RR is (greater than, less than, equal to) the fitness for SS.

R is incompletely dominant to S because the fitness for RS is (greater than, less than, equal to, between) the fitness of RR and SS.

Look at the graphs, and answer the following questions:

A. How does strong selection against RR and weaker selection againt RS over time affect the frequency of:

i. SS?

ii. S (q)?

iii. RR?

iv. RS?

v. R (p)?

B. Set the number of generations to 100 – what is the frequency of the R allele at the 100^{th} generation (hover over the q line at generation 100)?

C. Set the number of generations back to 1000 – what is the frequency of the R allele at the 1000^{th} generation?

D. Selection is most effective at eliminating a deleterious allele when it is (dominant, recessive, incompletely dominant), and least effective when it is (dominant, recessive, incompletely dominant)

4. **Selection for a beneficial recessive mutation, S**.

A new mutation will happen in a single individual, and will thus be at a low frequency. Because it is beneficial it should spread, but we will see how effective selection is at spreading it when it is recessive.

Start with the number of generations set to 1000.

Set the relative fitnesses to 1 for RR, 1 for RS, and 1.2 for SS.

Set the frequency of R (p) to 0.99 (which will set q to 0.01).

S is beneficial because the fitness of SS is (greater than, less than, equal to) the fitness for RR.

S is recessive to R because the fitness for RS is (greater than, less than, equal to) the fitness of RR.

Look at the graphs, and answer the following questions:

A. How does positive selection for SS affect the frequency of:

i. SS?

ii. S (q)?

iii. RR?

iv. RS?

v. R (p)?

B. There is a distinct lag of several hundred generations with little change in the frequency of S, and then the frequency increases rapidly – why? Hint: most of the S alleles are in heterozygotes during this lag phase, do heterozygotes get the benefit of carrying the S allele?

C. By what generation is S fixed in the population (that is, it becomes 100% of the alleles present)?

5. **Selection for a beneficial dominant mutation, R**.

A beneficial dominant allele will be expressed in both the homozygote and heterozygote – we will see how this changes the pattern of change due to selection.

Start with the number of generations set to 1000.

Set the relative fitnesses to 1.2 for RR, 1.2 for RS, and 1 for SS.

Set the frequency of R (p) to 0.01 (which will set q to 0.99).

R is beneficial because the fitness of RR is (greater than, less than, equal to) the fitness of SS.

S is recessive to R because the fitness for RS is (greater than, less than, equal to) the fitness of RR.

Look at the graphs, and answer the following questions:

A. How does positive selection for RR and RS affect the frequency of:

i. SS?

ii. S (q)?

iii. RR?

iv. RS?

v. R (p)?

B. Is there a lag in the spread of the R allele when it is dominant and beneficial?

C. Does R become fixed in the population (that is, does it becomes 100% of the alleles present)? Why would it not go to fixation when R did as a beneficial recessive? Hint: S is acting as a deleterious recessive in this example – is it being effectively elimnated when it becomes rare?

6. **Selection for a beneficial incompletely dominant mutation, R**.

A beneficial incompletely dominant allele will be most beneficial in the homozygote, and will have intermediate fitness in the heterozygote.

Start with the number of generations set to 1000.

Set the relative fitnesses to 1.2 for RR, 1.1 for RS, and 1 for SS.

Set the frequency of R (p) to 0.01 (which will set q to 0.99).

R is beneficial because the fitness of RR is (greater than, less than, equal to) the fitness of SS.

R is incompletely dominant because the fitness for RS is (greater than, less than, equal to, between) the fitness of RR and SS.

Look at the graphs, and answer the following questions:

A. How does positive selection for RR and RS affect the frequency of:

i. SS?

ii. S (q)?

iii. RR?

iv. RS?

v. R (p)?

B. Why does R go to fixation when it is incompletely dominant, but not wen it was completely dominant to S? Hint: is selection against S still ineffective for the heterozygotes?

C. In all three cases of the spread of beneficial mutations, heterozygotes first increase in frequency, peak around 0.5, and then decline. Why? Hint: when an allele is rare, is it more likely to be matched with the same allele to form a homozygote, or is it more likely to be matched with the other allele to form a heterozygote?