How natural selection changes allele frequencies

Drift changes allele frequencies *randomly* (up or down) and *slowly* (proportional to 1/*N*).

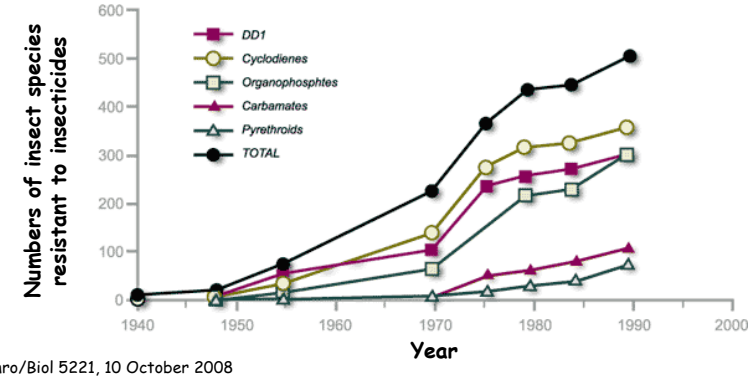
Selection *biases* the direction of allele-frequency change, and increases its speed.

Alleles change frequency at speeds proportional to their *difference in average fitness:*

$$\Delta p = pq(\bar{W}_1 - \bar{W}_2)/\bar{W}$$

Thus if selection is strong, it can change allele frequencies (and phenotypes) very quickly.

For example, alleles conferring resistance to insecticides and antibiotics have risen to high frequencies in just a few years, in many species of insects and bacteria.



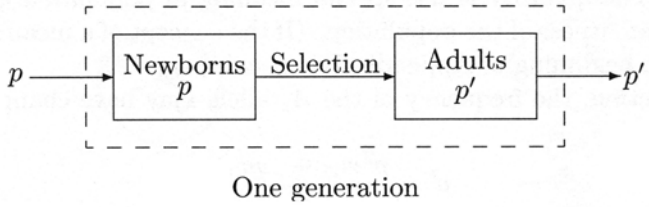


Figure 3.2: The simple life cycle used in the fundamental model of selection.

A general model of selection for two alleles at one locus

Let p be the frequency of allele $\mathbf{A_1}$ in the *present* generation [$q = 1-p = f(\mathbf{A_2})$].

Let p' be its frequency *next* generation.

Assume that the population mates at random with respect to genotypes at the A locus. (This does not require truly "random" mating!)

Let W_{11}, W_{12}, and W_{22} be the relative fitnesses (average surviving offspring) of the three diploid genotypes ($\mathbf{A_1A_1}$, $\mathbf{A_1A_2}$, $\mathbf{A_2A_2}$).

The population's average fitness is a weighted mean of the genotypic fitnesses.

genotype	$\mathbf{A_1A_1}$		$\mathbf{A_1A_2}$		$\mathbf{A_2A_2}$		
frequency at conception	p^2	+	$2pq$	+	q^2	=	1
fitness (*e.g.*, survival)	W_{11}		W_{12}		W_{22}		
population mean (total) fitness	p^2W_{11}	+	$2pqW_{12}$	+	q^2W_{22}	=	$\bar{W}$

Dividing the genotypic *components* of W-bar by W-bar gives the proportional reproductive contributions of the genotypes to next generation's gene pool.

genotype	$\mathbf{A_1A_1}$		$\mathbf{A_1A_2}$		$\mathbf{A_2A_2}$		
proportion of population's fitness	$p^2W_{11}/\bar{W}$	+	$2pqW_{12}/\bar{W}$	+	$q^2W_{22}/\bar{W}$	=	1

Next generation's allele frequency p' is the contribution of $\mathbf{A_1A_1}$ genotypes plus one half of the contribution of $\mathbf{A_1A_2}$ (since half of their gametes will be $\mathbf{A_1}$).

$$p' = p^2W_{11}/\bar{W} + pqW_{12}/\bar{W} = p[pW_{11} + qW_{12}]/\bar{W}$$

The term in square brackets is the average or "marginal" fitness of allele $\mathbf{A_1}$ (because a proportion p of all $\mathbf{A_1}$ alleles find themselves in $\mathbf{A_1A_1}$ homozygotes). Our updating rule or recurrence equation for p therefore reduces to

$$p' = p[\bar{W}_1]/\bar{W}$$

If we subtract p we get the *change* in p expected in one generation:

$$\Delta p = p' - p = p\bar{W}_1/\bar{W} - p = p(\bar{W}_1 - \bar{W})/\bar{W}$$

But the population's mean fitness can be rewritten as a weighted average of the allelic marginal fitnesses:

$$\bar{W} = p\bar{W}_1 + q\bar{W}_2$$

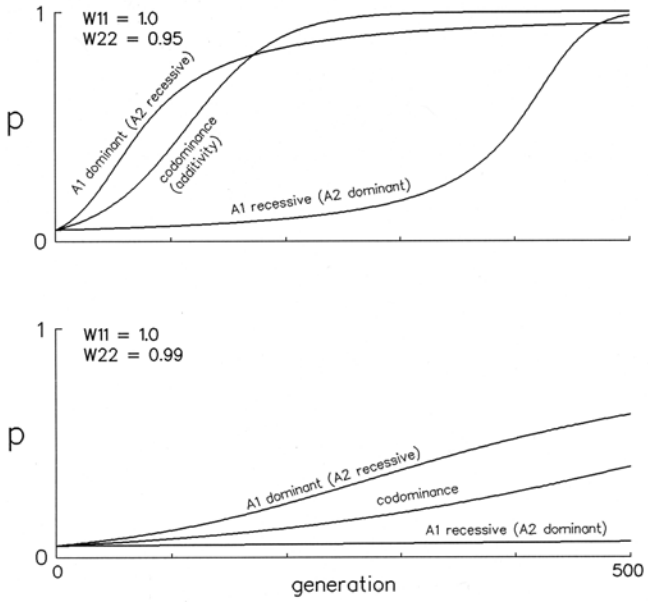
Thus the equation for Δp can be simplified to

$$\Delta p = pq(\bar{W}_1 - \bar{W}_2)/\bar{W}$$

The rate of allele-frequency change is fastest at intermediate allele frequencies (when *pq* is greatest).

Rare recessive alleles (whether advantageous or harmful) are almost "invisible" to selection.

Smaller fitness differences lead to proportionally slower rates of allele-frequency change.



Fitnesses are often described by the "selection differential" *s*

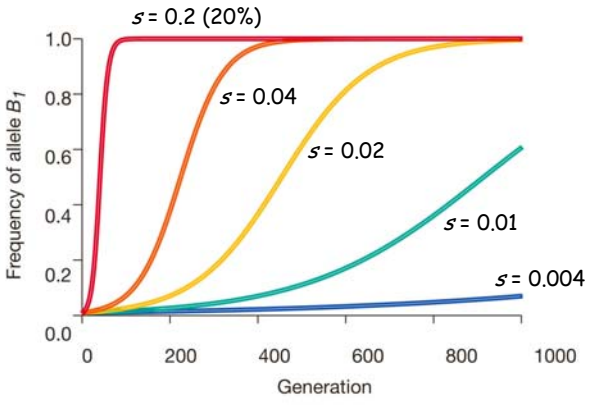
For example, it is common practice to write $W_{11} = 1$, $W_{22} = 1-s$, and $W_{12} = 1-hs$.

If $h = 0$, allele 1 is dominant.

If $h = 1$, allele 2 is dominant.

And if $h = 0.5$, the heterozygotes are intermediate in fitness (*codominance* or *additivity*).

In the examples below, $h = 0.5$; and in the first case, $s = 0.2$.



Selection scheme

	Percent surviving		
	B_1B_1	B_1B_2	B_2B_2
Strong			
	100	90.0	80.0
	100	98.0	96.0
	100	99.0	98.0
	100	99.5	99.0
	100	99.8	99.6
Weak			

Note: color code does not agree with graph!

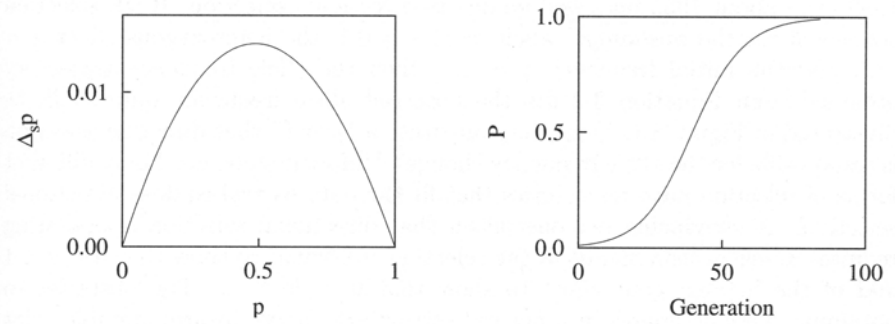


Figure 3.3: Properties of directional selection with $h = 0.5$ and $s = 0.1$. The left-hand graph shows the change in the allele frequency in a single generation. The right-hand graph shows the evolution of the allele frequency over 100 generations.

"Overdominance" or heterozygote advantage (h < 0)

W_{11} = 1.0
W_{12} = 1-hs = 1-(-0.5)(0.1) = 1+0.05 = 1.05
W_{22} = 1-s = 1-0.1 = 0.9

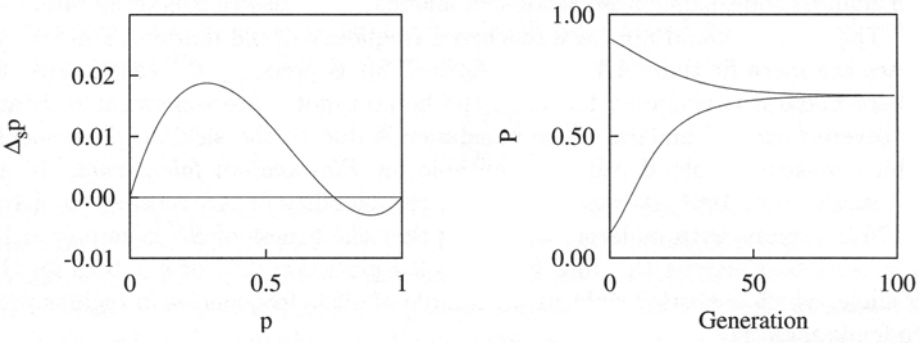


Figure 3.4: Properties of balancing selection with $h = -0.5$ and $s = 0.1$.

An important special case: lethal recessive alleles

$W(A_1A_1) = W(A_1A_2) = 1$
$W(A_2A_2) = 0$

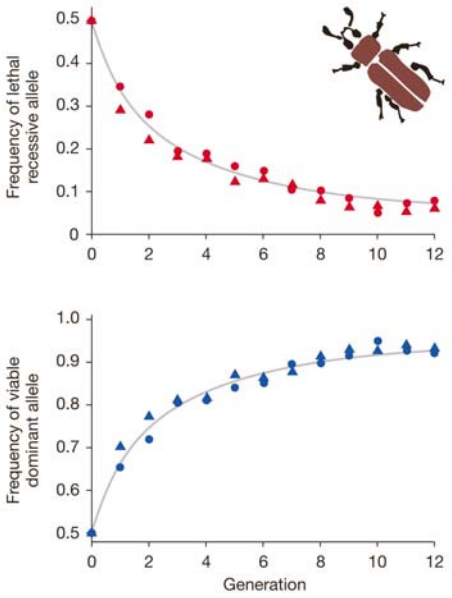
$W_1 = 1$
$W_2 = [(1)\frac{1}{2}2pq + (0)q^2]/[pq+q^2] = p$
$W = p(1) + q(p) = p(1+q) = p(2-p)$

$p' = p[W_1/W] = p[1/p(2-p)] = 1/(2-p)$

E.g., suppose p = 0.5

Then p' = 1/(2-0.5) = 1/1.5 = 0.67

Peter Dawson (1970) used a recessive lethal allele in the flour beetle (*Tribolium confusum*) to test this prediction of the model. His data are shown in the graphs on the right. The theoretical prediction is graphed as continuous gray lines. Amazing!



How and when do typical genes contribute to fitness?

A major puzzle: Most genes appear to be *unnecessary!*

Half or more can be "knocked out" (fully disabled) in yeast, worms, flies and even mice, without any obvious phenotypic effects (in the lab, anyway).

But these genes are maintained in evolution, so they must be useful. *How?*

Two hypotheses:

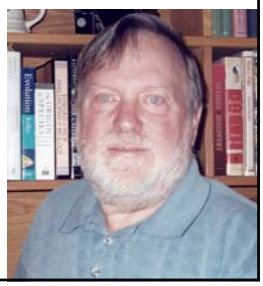
(1) Most are "special-purpose" genes needed only under certain circumstances (stresses that occur in nature but not in the lab).

(2) Most are "fine-tuning" genes that increase the efficiency or accuracy of some physiological or developmental process in most environments.

Experimental test devised by Joe Dickinson:

Compete "no-phenotype knockouts" against genotypes that are identical except for the knockout, and let natural selection measure their relative fitnesses.

Dickinson talked Janet Shaw (a yeast cell biologist) and John Thatcher (an undergraduate) into helping him try to do this with yeast.

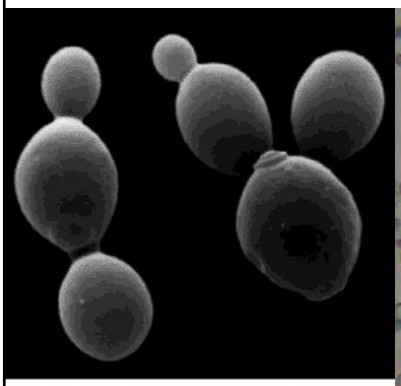


How to ask cells if they miss a (random) gene

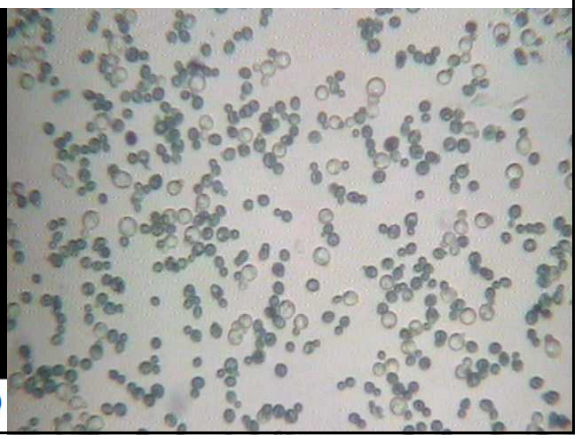
Mark either the random, "no-phenotype" knockout, or the wild-type parent, with *lac Z* so that you can score their relative numbers on indicator plates.

Start populations with equal numbers of wild-type and knockout cells; grow them for many generations in complete (rich) liquid media.

Sample the populations every 10-20 generations and score the relative numbers of marked and unmarked cells.

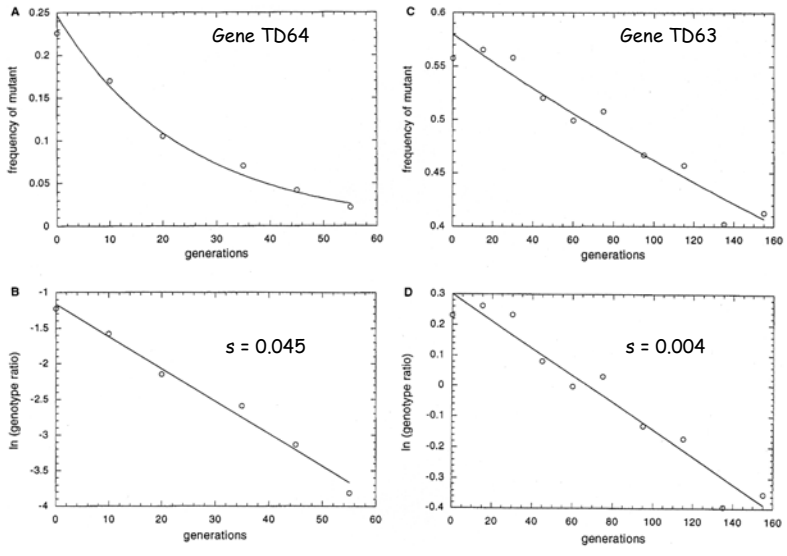


Yeast (*Saccharomyces cerevisiae*)



Plot the frequency of the knockout as a function of generations (A, C)

Also plot the log of the ratio of the allele frequencies [ln(q/p)] versus generation (B, D). The slope of this (straight) line is an estimate of the selection coefficient (*s*).



Summary of results for 27 "no-phenotype" knockouts

Nineteen mutations (70%) showed statistically significant fitness defects ranging from 0.3% (s = 0.003) to 23% (s = 0.23).

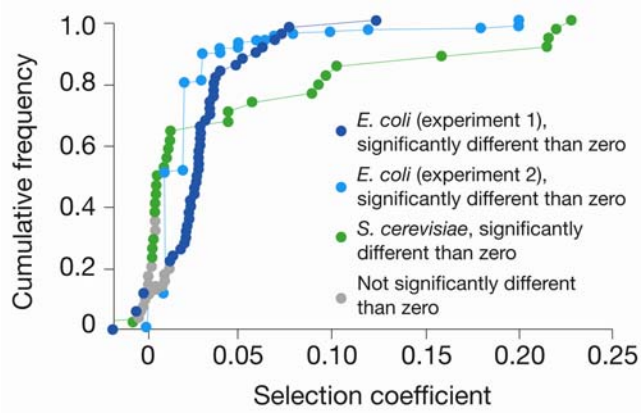
Among these, the typical (median) selection coefficient was 1-2%.

Six mutations (22%) were not statistically distinguishable from neutral. (Five of the six appeared to be weakly deleterious, and one appeared to be beneficial.)

A more sensitive experimental design (with larger populations and allele-frequency assays) would probably show most of these to be significant, raising the fraction of deleterious no-phenotype knockouts to 85-90%.

Two of the 27 knockouts (7%) were significantly *advantageous*, with "negative" coefficients of s = -0.005 and s = -0.007.

Conclusion: Most genes make *modest* contributions to fitness



This finding (in bacteria, too) supports the "fine-tuning" hypothesis.

Such small effects *could not be detected except by natural selection.*

Read the paper: Thatcher, Shaw & Dickinson, *PNAS* **95**, 253-257 (1998).

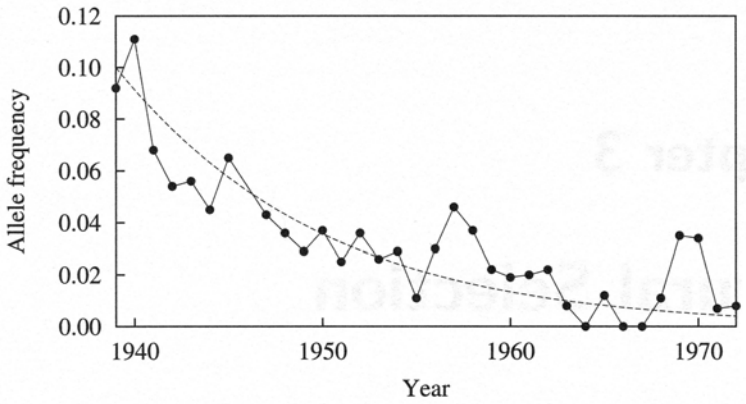
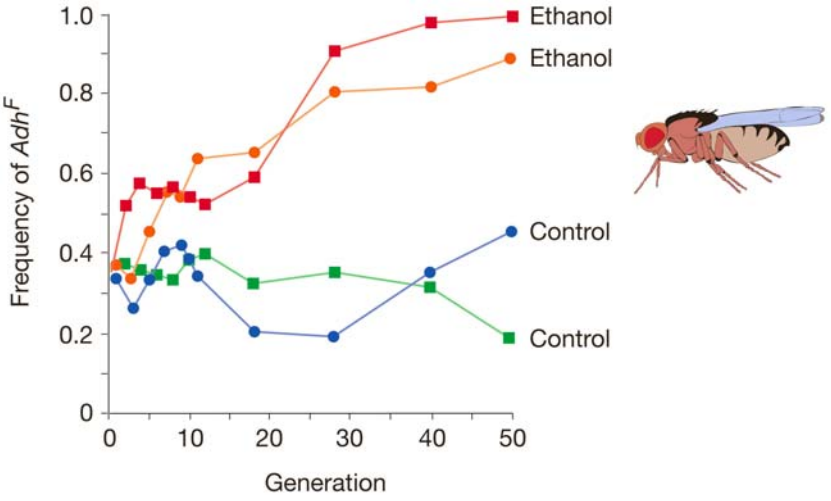


Figure 3.1: The observed frequency of the *medionigra* allele in the scarlet tiger moth population compared to the expected frequency assuming a 10 percent disadvantage.

Adh^F beats *Adh*^{S} on laboratory fly food soaked in EtOH



Homework problem: Let FF homozygotes have a fitness of 1.0 (W_{FF}=1), and assume F/S heterozygotes have *intermediate* fitness (h = ½). Very roughly, what is the fitness of the SS homozygotes, in the ethanol-soaked environment?