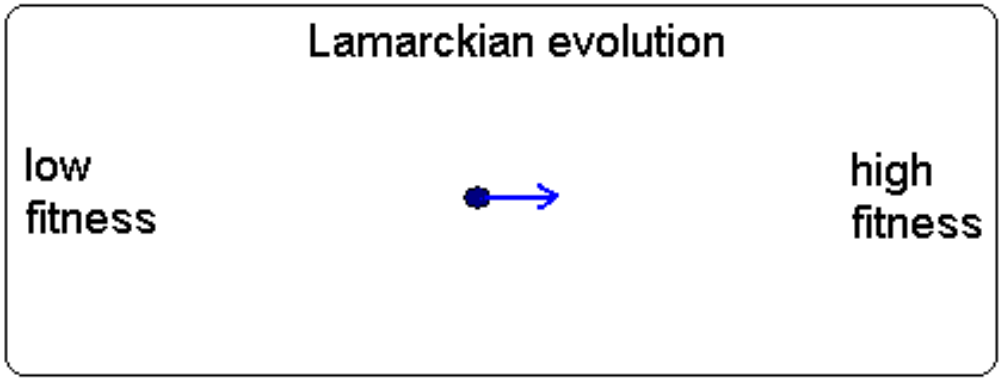
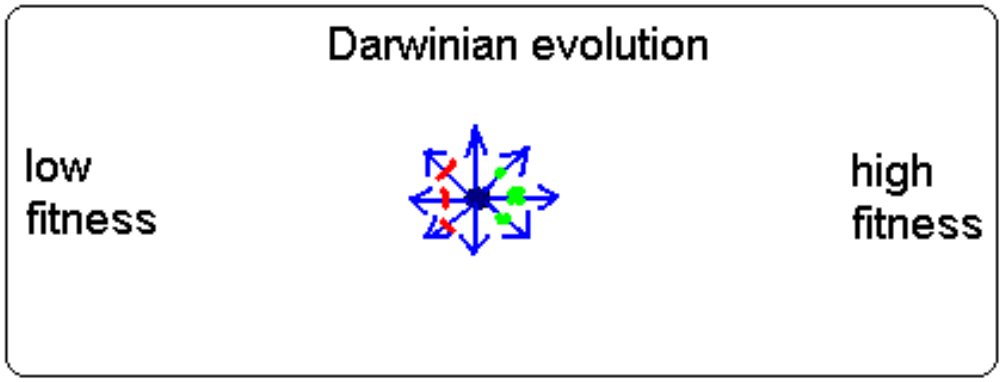
Natural Selection

1) The concept

Selection is one of the five factors of Microevolution (mutation, selection, mode of reproduction, population structure, genetic drift) and, together with mutation, one of the only two factors that are absolutely necessary for Darwinian evolution.





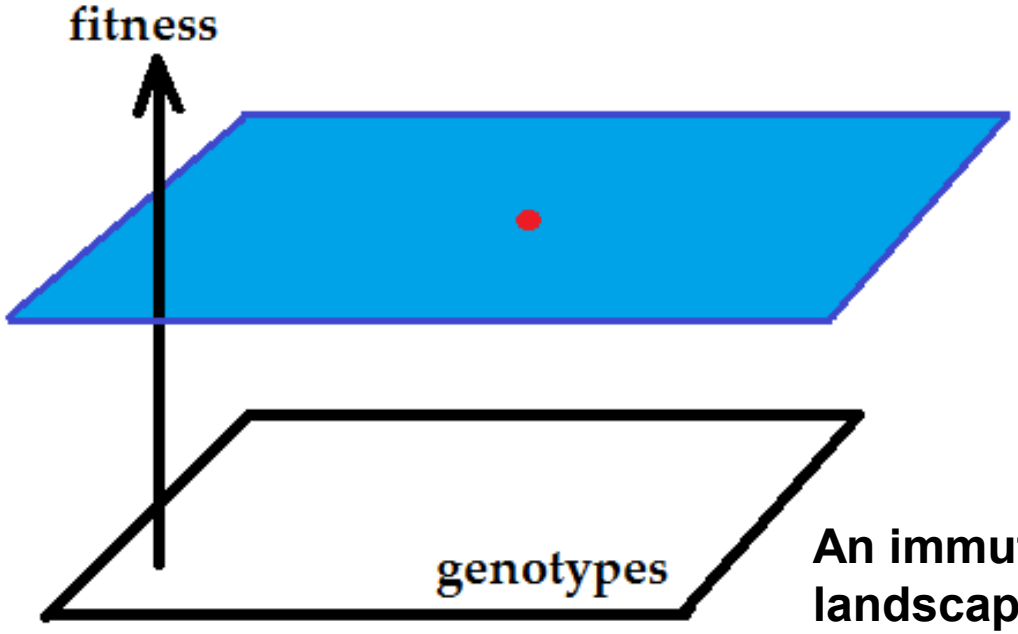
Selection within populations is the only feasible natural mechanism of evolution - as long as organisms cannot directly modify their DNA in the desired direction.

Selection is differential propagation of some entities. These entities may be individuals *per se*, or groups of individuals, defined by some rule, such as having a particular genotype, a particular phenotype, or being tightly related to each other. The most important form of selection is the selection among genotypes. Thus, selection usually means differential propagation of genotypes.

Selection operates within every population due to two universal properties of life.

First, unavoidable mutation makes every population genetically variable.

Second, genotype > phenotype maps are not flat, and the genotype of an individual affects all aspects of its phenotype, including the efficiency of reproduction or fitness.

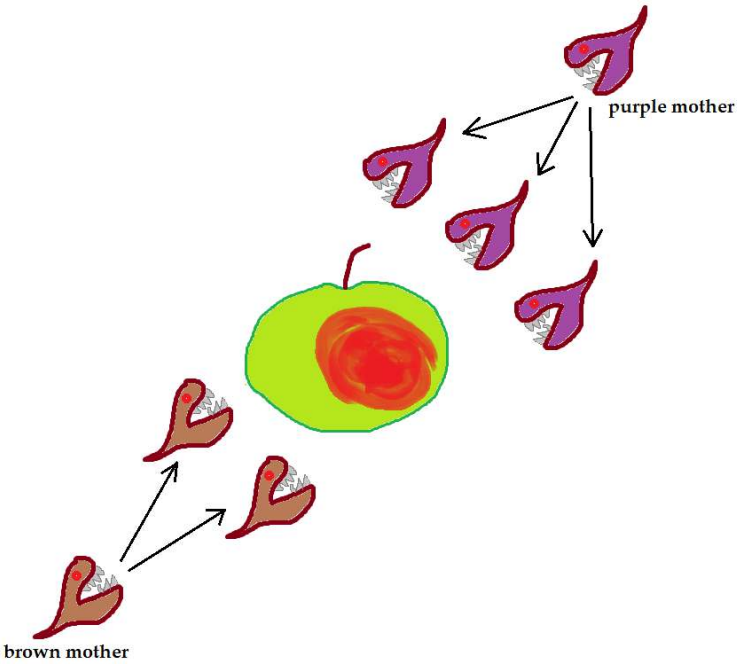


An immutable genotype on a flat fitness landscape; a double impossibility.

Two other properties of life make selection sustainable and efficient.

First, individuals are always capable of producing, under suitable conditions, more offspring than what is needed to just sustain the population, *i. e.* more than one offspring with apomixis or two with amphimixis (2 in unicellular organisms and, apparently, never less than 10 in multicellular organisms). Only due to these excessive offspring, a population can afford any selection, in the sense that variation in the efficiency of reproduction does not cause extinction.

Second, even a substantial change of the phenotype often does not adapt an organism to a new, distinct ecological niche. As a result, different genotypes do not form separate populations with independent regulation of their sizes but, instead, compete with each other.



A clonal apple-eater produces more than one offspring, making selection affordable for the population, and apple-eaters of different genotypes still compete for the same resource, apples.

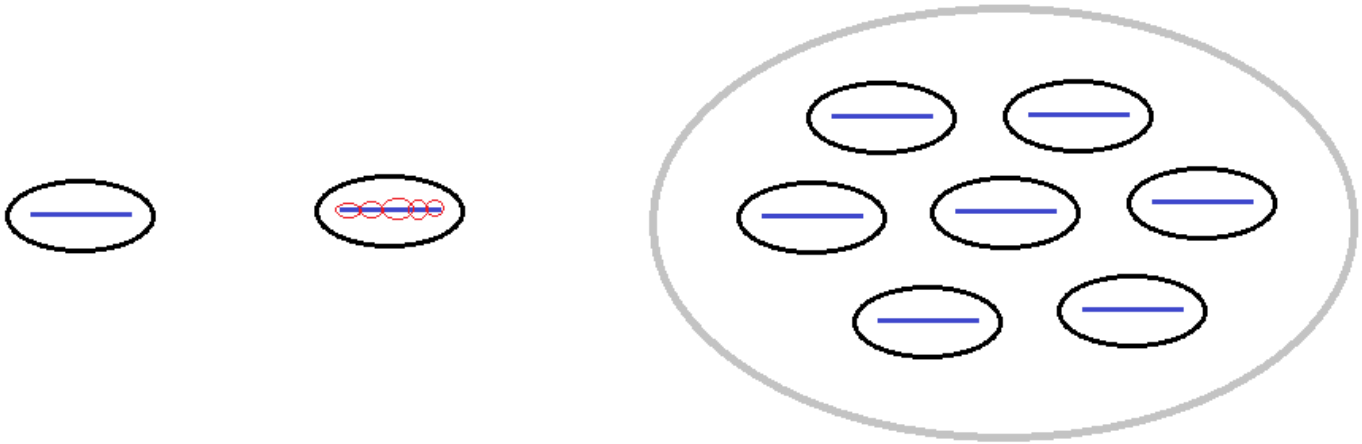
Units of selection

The simplest and the most important kind of selection is selection among individuals, or individual selection.

At the level of organisms, differences between efficiencies of reproduction of organisms with different genotypes are due to an endless variety of biological mechanisms, and can involve both very general and very specific adaptations.

At the populations level, the efficiency of reproduction of an individual is determined by its viability, the rate of development, the ability to obtain breeding partners, fecundity, and longevity. Still, treating fitness of an individual as a quantitative trait is usually an acceptable approximation.

Individual selection of genotypes means that individuals of different genotypes reproduce with different average efficiencies. Amphimixis can make loci, instead of complete genotypes, the units of selection. Groups of individuals can be units of selection in a structured population.



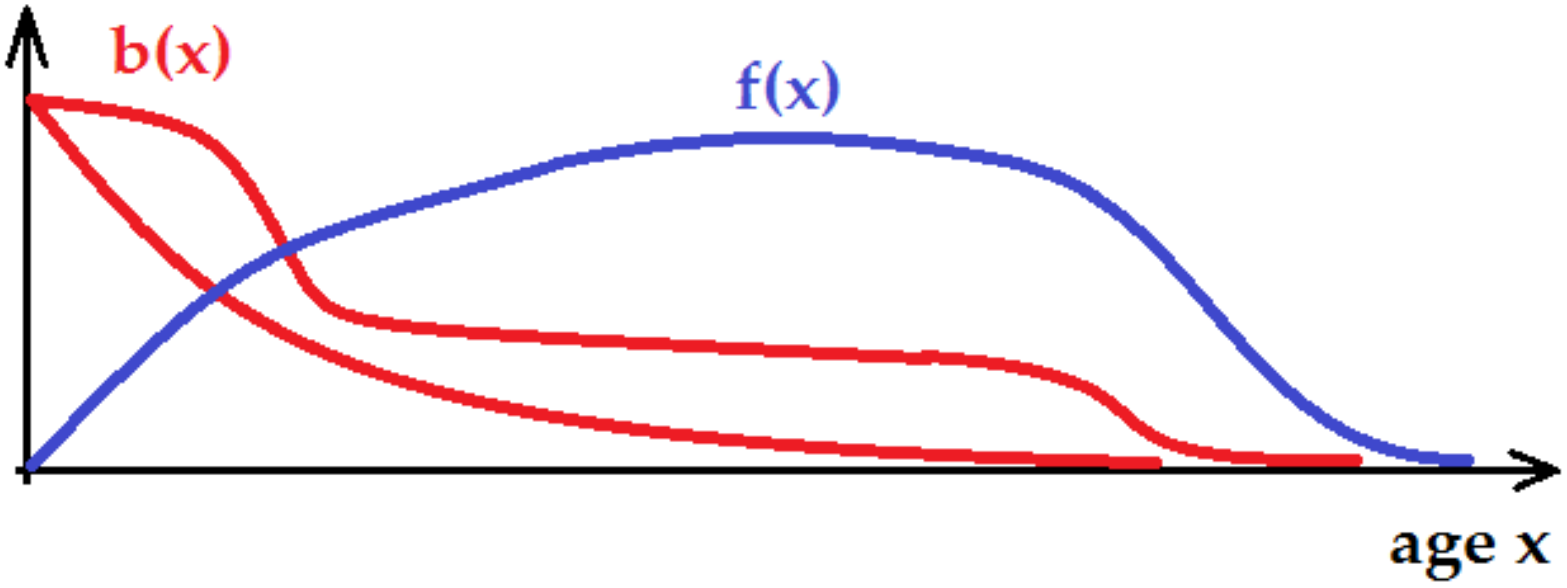
Fitness can be characterized by just one number

Within an age-structured population, describing the fitness by just one number if, strictly speaking, impossible. Instead, the efficiency of reproduction must be described by at least two functions:

b(x), the probability of reaching age x

and

f(x), fecundity at age x



Still, the age structure of a population usually does not affect its Microevolution strongly. This is because fitness still can be approximated by one number even when there is an age structure.

A simplistic way of doing this is by the life-time fecundity F:

$$F = \int_0^{x_{\max}} b(x) f(x) dx$$

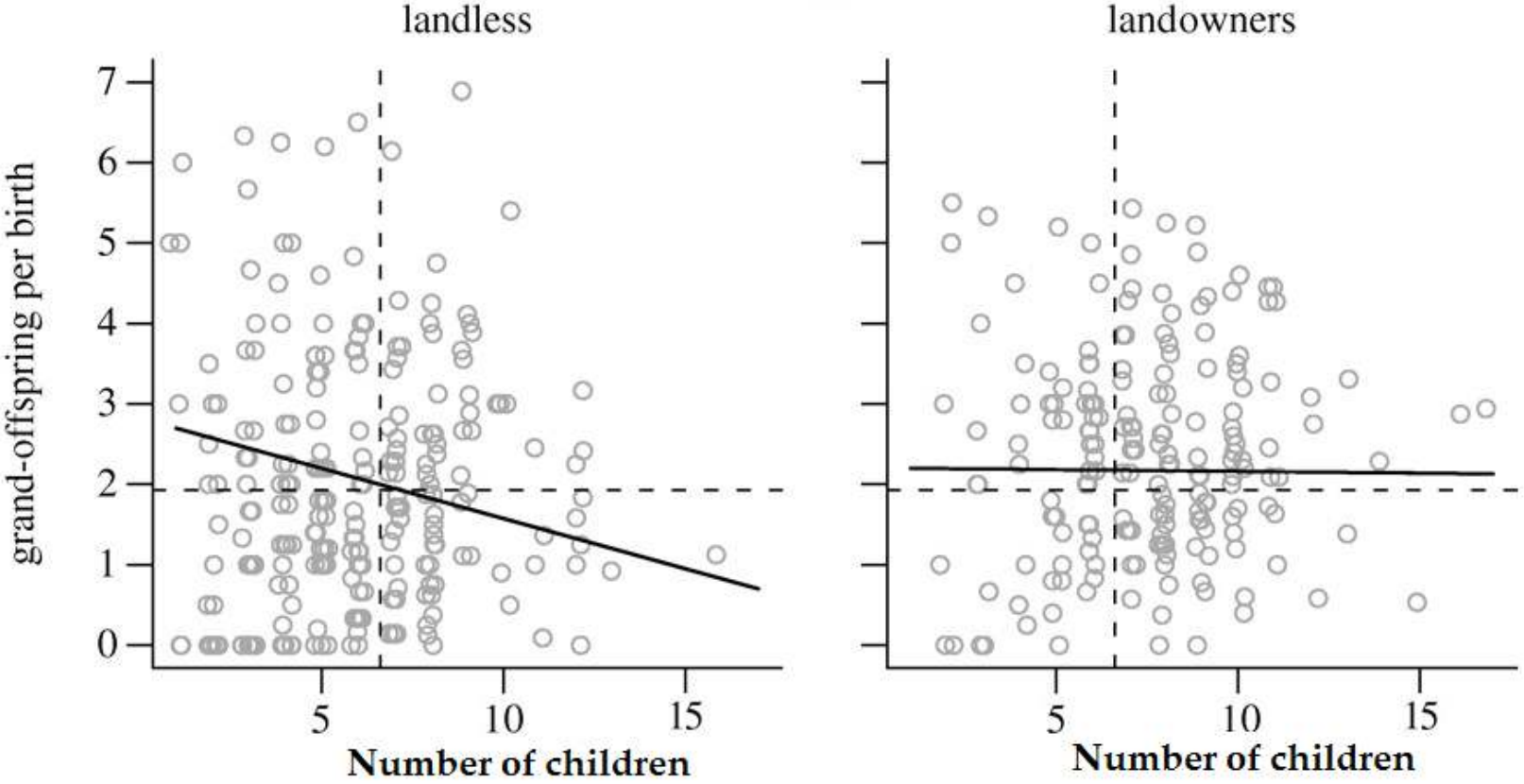
A better way of doing this is is through the Malthusian parameter r, which is the growth rate of a population of individuals with the same b(x) and f(x), after its age structure equilibrated:

$$\int_0^{x_{\max}} e^{-rx} b(x) f(x) dx = 1$$

Unless selection is very strong and fluctuates in time widely, individuals of different ages are characterized by approximately the same genotype frequencies. Thus, approximate description of an age-structured population which ignores this structure usually works. We can keep using discrete-generation models, and treat fitness as a scalar.

Complications

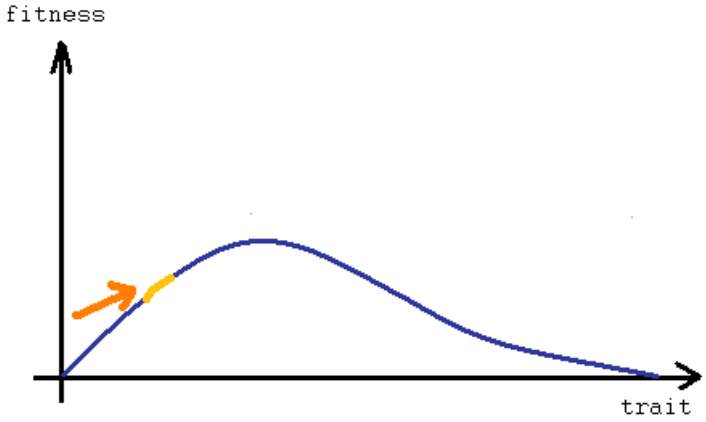
If individuals interact with each other in a way which affects their reproduction, considering their fitnesses separately is only an approximation. For example, the fitness of an offspring may, at least under some environments, depend on the amount of resources which parents were able to allocate to it.

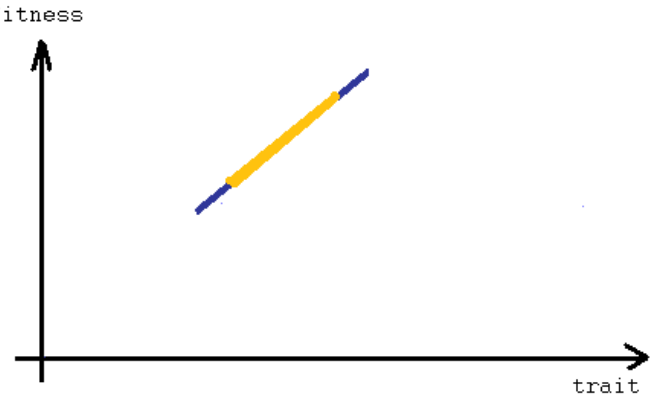


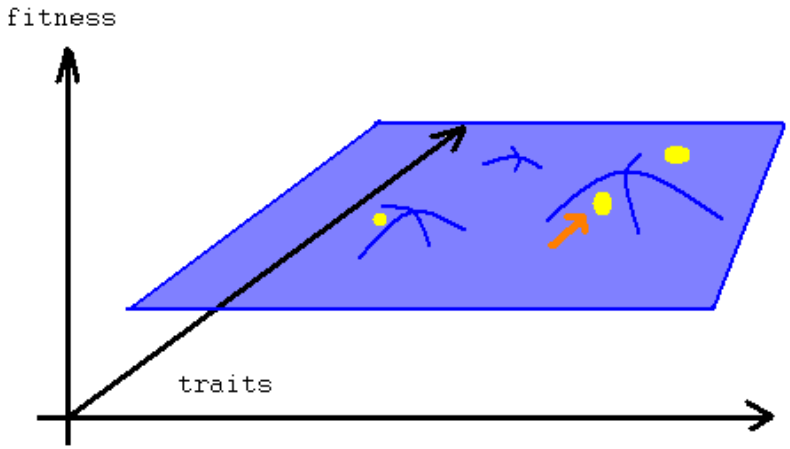
The regression of the number of grand-offspring per offspring for women from landless (left) and landowning (right) families in pre-industrial Finns on their number of children. This number declines in landless, but not in landowning families, presumably due to competition between offspring for limited resources.

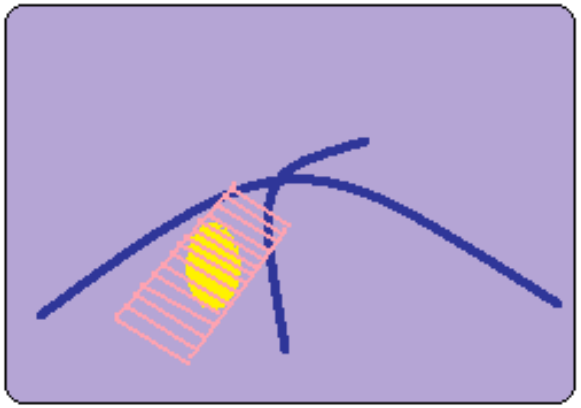
2) Populations on fitness landscapes

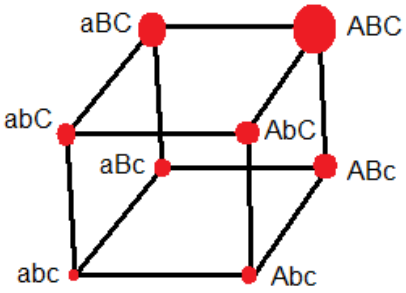
Natural selection acting within a population is defined by the fitness landscape and by how a population sits on it. The range of within-population variation is not wide, and here we care only about microscopic properties of fitness landscapes.









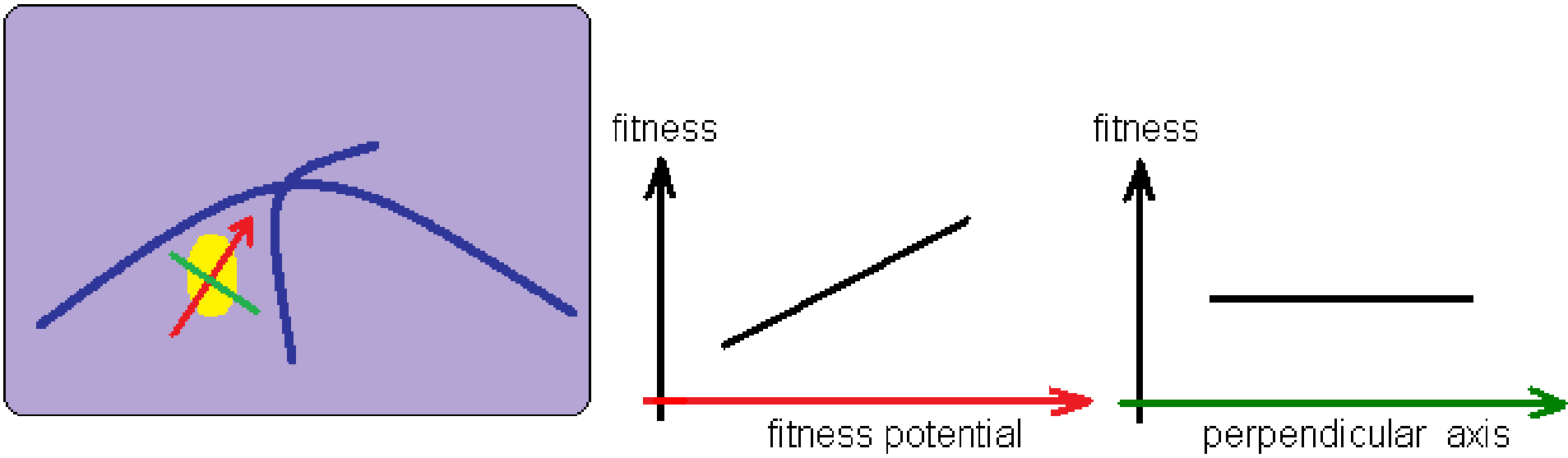


Under a strong enough magnification, every fitness landscape is close to linear.

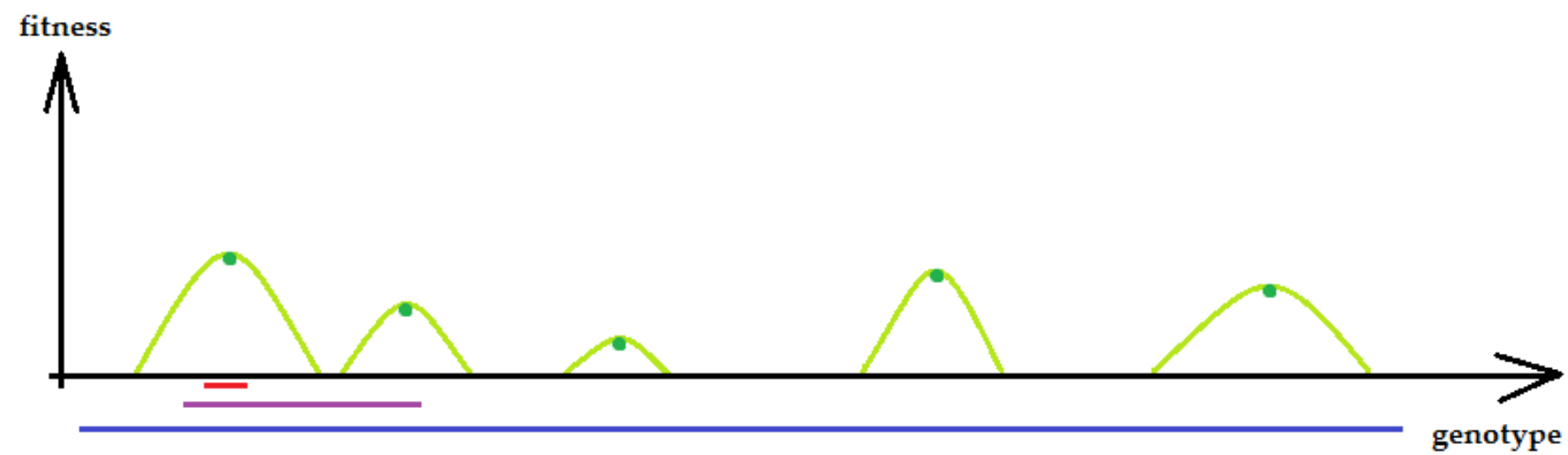
A linear fitness landscape has two key properties.

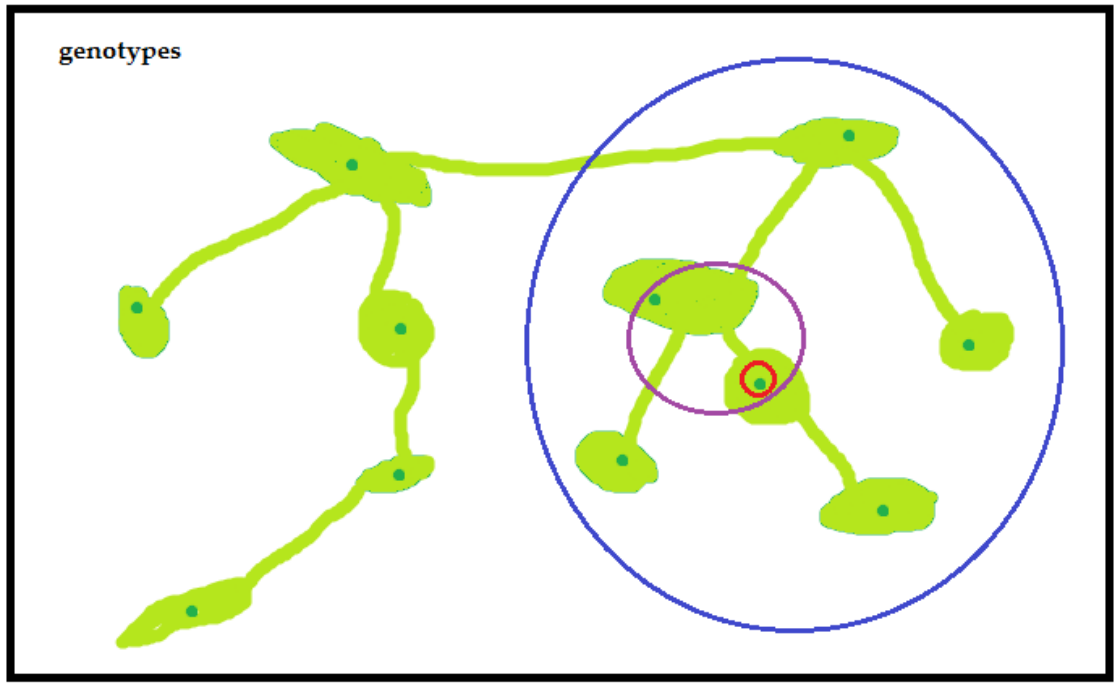
First, the fitness of a genotype can be represented by the sum of constant contributions from all its constituent allele.

Second, there is just one direction, known as gradient, in which fitness changes. Thus, the fitness of a genotype is determined by its fitness potential, the position alone the axis that points in the direction of gradient.



Fitness potential axis is in red, and the only perpendicular axis is in green. Fitness is an approximately linear function of the fitness potential of a genotype within the population and does not depend on its position along the perpendicular axis.

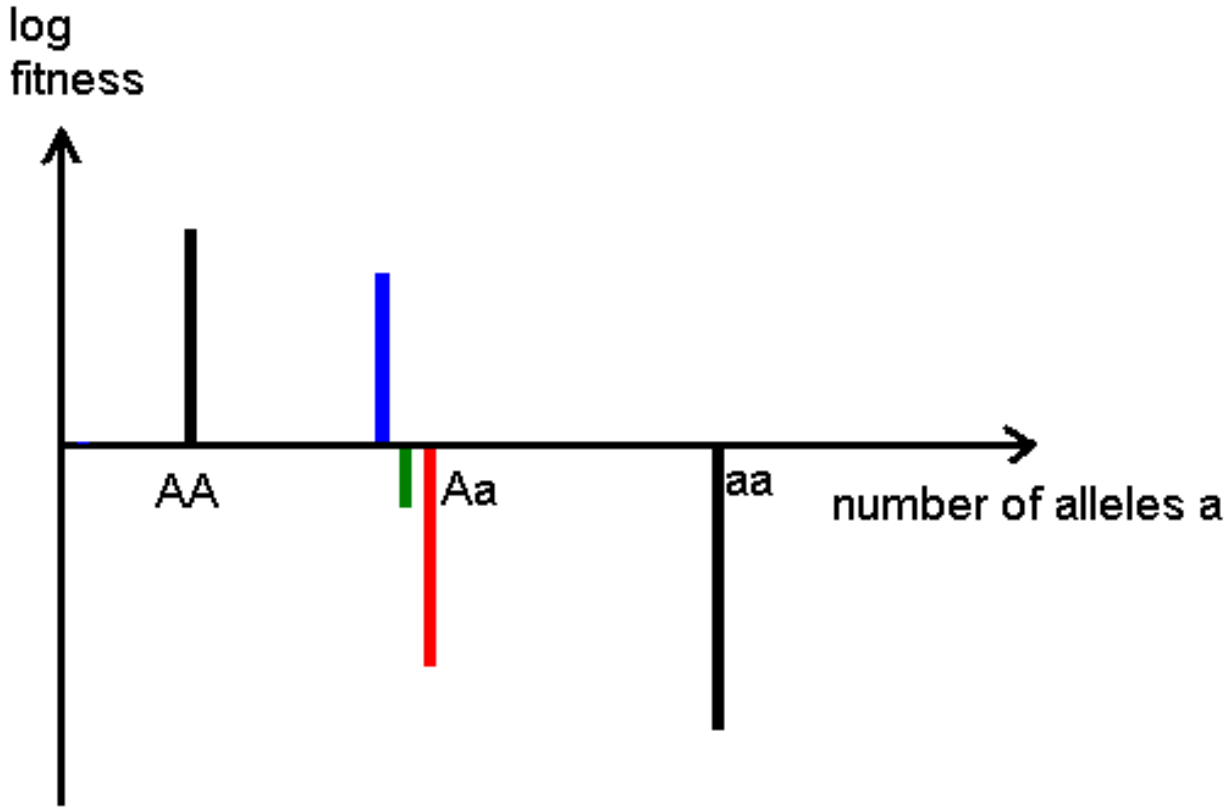




Micro, medium, and macro scales at which a fitness landscape can be considered. All viable genotypes are in light green, local fitness peaks are in dark greens, and the 3 scales are red, purple, and blue, repectively. Genotypes are arranged in one (top) or two (bottom) dimensions which is an oversimplification.

Even on the scale of within-population variation, fitness landscape as a linear function of fitness potential is not always a good approximation. Any deviation of fitness landscape from linearity on the logarithmic scale is called epistasis.

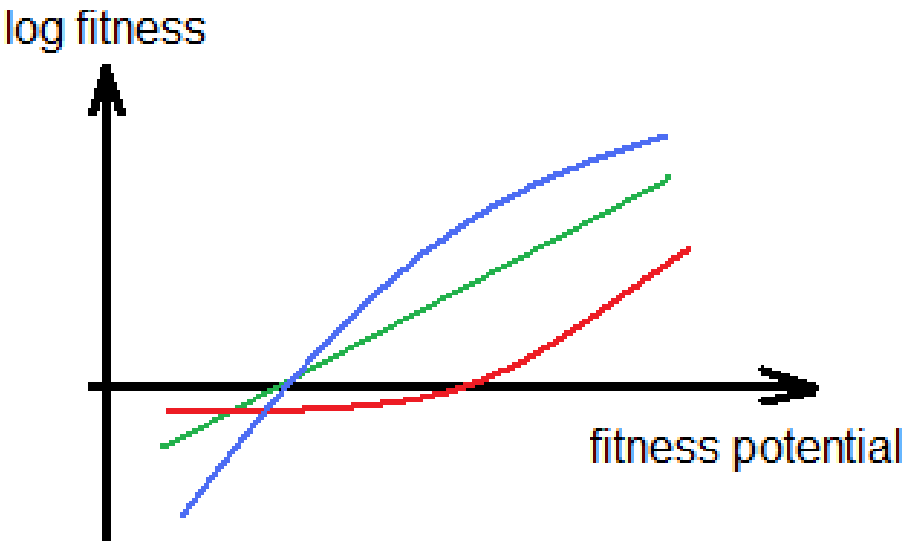
Example of epistasis: dominance and recessivity.

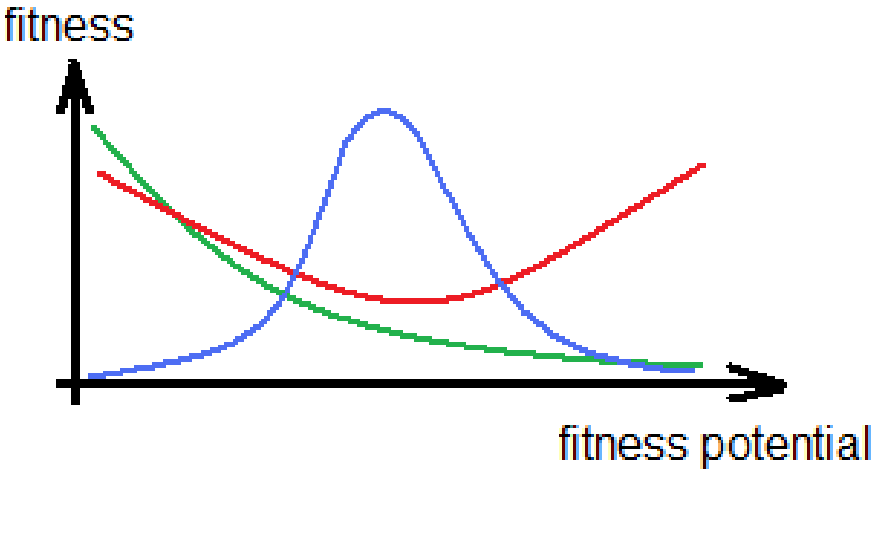


If we consider just one locus A, with alleles A and a, the log fitness of heterozygote Aa may be closer to the log fitness of AA (if A is dominant) or to the log fitness of aa (if A is recessive), or be the arithmetic mean of the log fitnesses of AA and aa (intermediate dominance).

Generic epistasis simply means an arbitrary fitness landscape, with multiple peaks, minima, etc. Such landscapes are necessary for consideration of Macroevolution. However, within a simpler context of Microevolution, it makes sense to consider three restrictive kinds of epistasis: one-dimensional, monotonic, and single-peak.

One-dimensional epistasis inherits, from the simplest linear case, the assumption that there is just one fitness-determining variable, fitness potential. However, now fitness can be an arbitrary function of this variable.

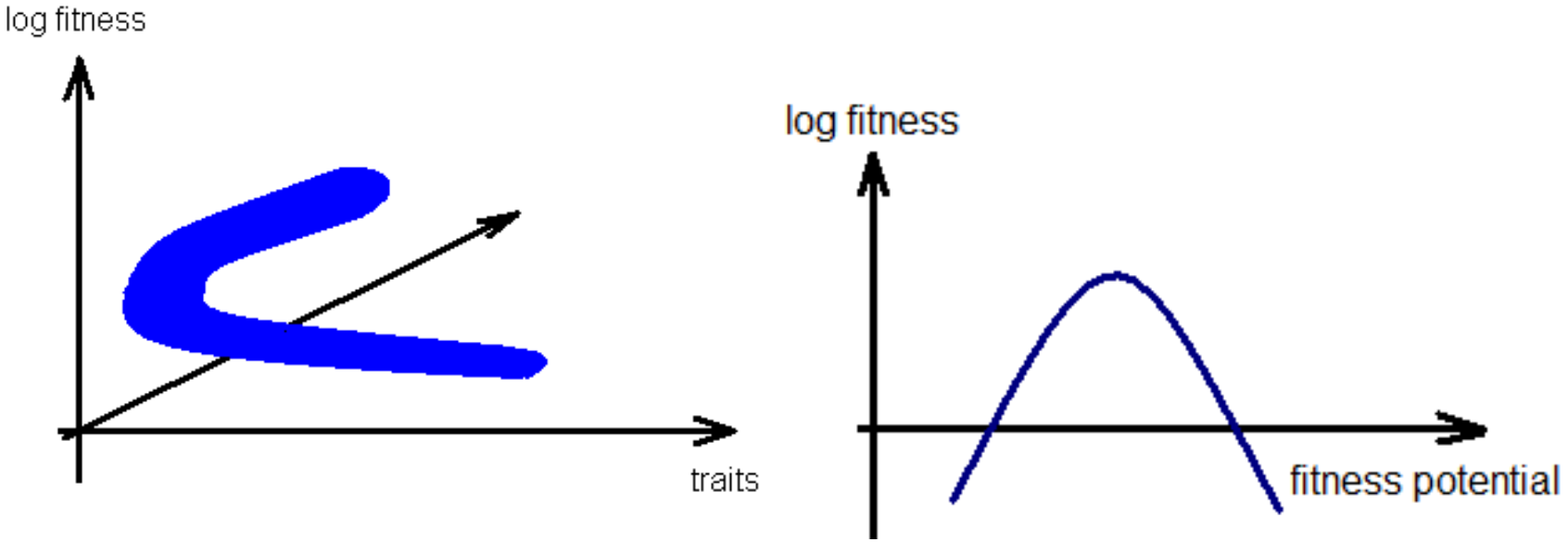




Examples of one-dimensional epistasis. (left) Log fitness is plotted; no epistasis (green), convex (blue), and concave (red) fitness functions. (right) Fitness is plotted; no epistasis (green), unimodal (blue), and bimodal (red) fitness functions.

The second restrictive kind of epistasis is monotonic, meaning that a particular genetic change never impacts fitness in the opposite directions.

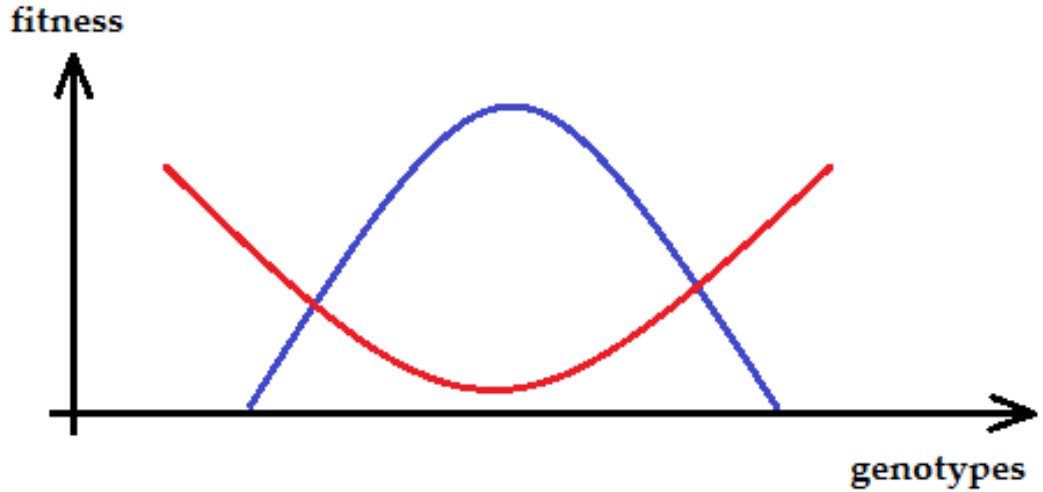
The two restrictive modes of epistasis are not equivalent: one-dimensional epistasis can be sign epistasis and monotonic epistasis can be multidimensional.



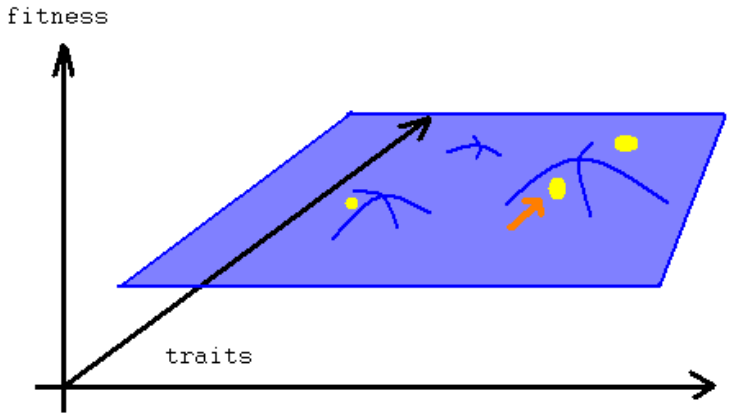
(left) Monotonic, multidimensional epistasis: high values of both traits are deleterious, and these deleterious effect reinforce each other.

(right) Sign, one-dimensional epistasis: intermediate values of the trait confer the highest fitness.

The third restrictive kind of epistasis can be called single-peak, meaning that all the genotypes under consideration belong to the domain of attraction of one fitness peak, so that evolutionary trajectories that climb the fitness landscape starting from every genotype will all end up on the same peak. Modes of epistasis that do not have this property are called multiple-peak epistasis.

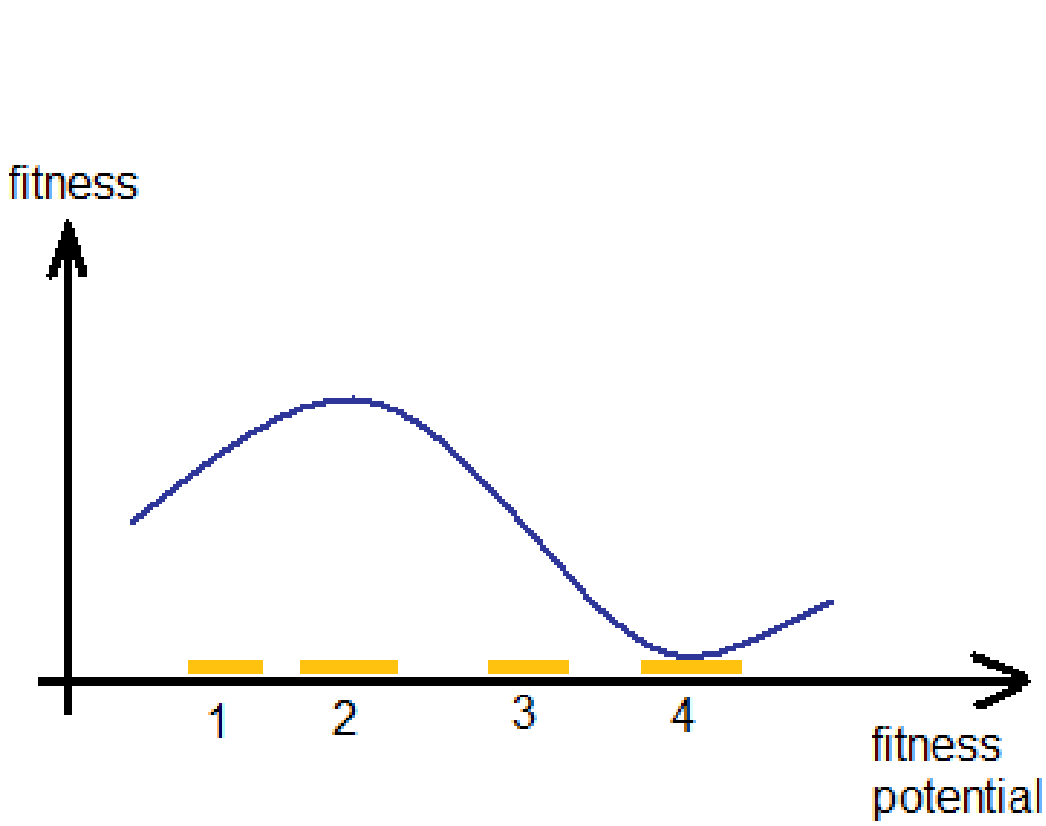


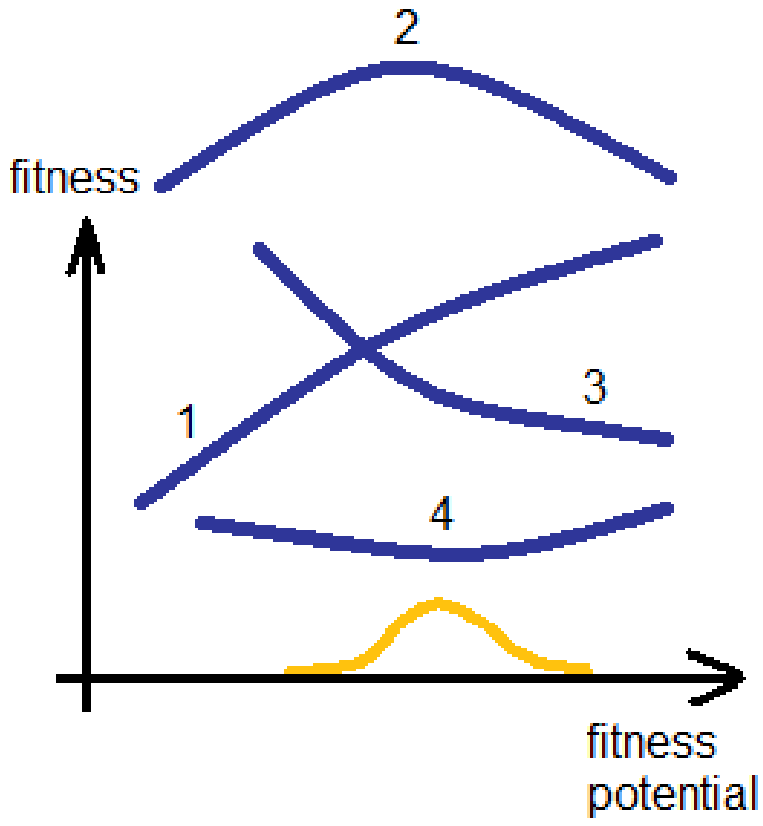
Two fitness landscapes with one-dimensional and sign epistasis: the blue one is single-peak, and the red one is multiple peak.



generic fitness landscape = generic epistasis

This analysis of fitness landscapes prepares us for considering selection. Of course, fitness landscape alone does not define selection - the position of the population is also essential.



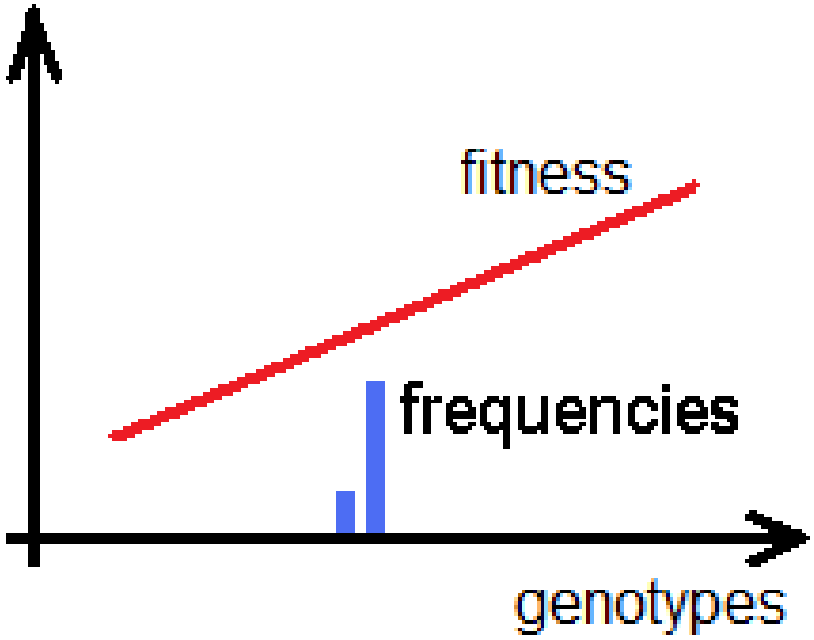


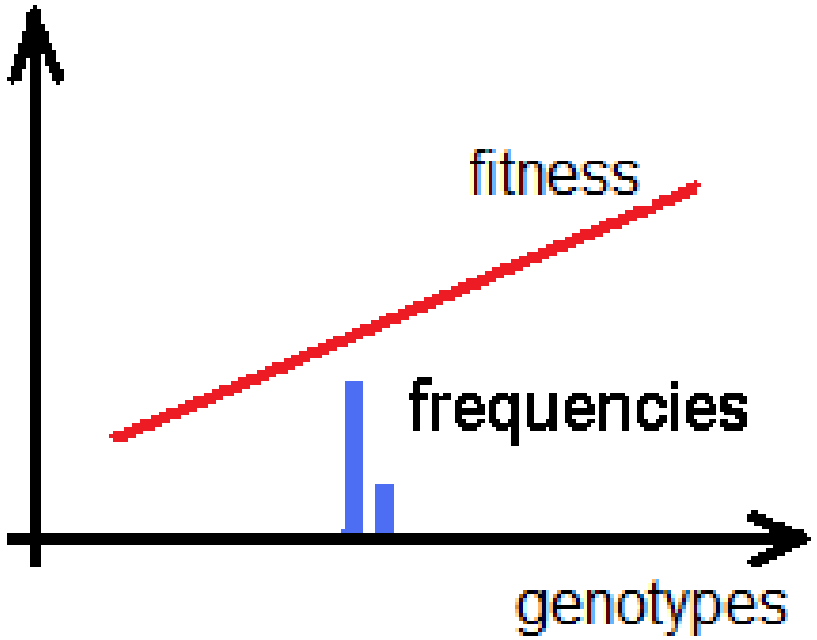
Selection favors high, intermediate, low, and extreme values of the trait in populations 1, 2, 3, and 4, respectively.

Let us start from the simplest case of unordered genotypes. Then, just two key modes of selection are possible, which do not depend on subtle features of the fitness landscape:

i) Negative selection - the most fit of the available genotypes is common in the population, and less fit genotypes are rare,

ii) Positive selection - the most fit of the available genotypes is rare, and the most common genotype is less fit.

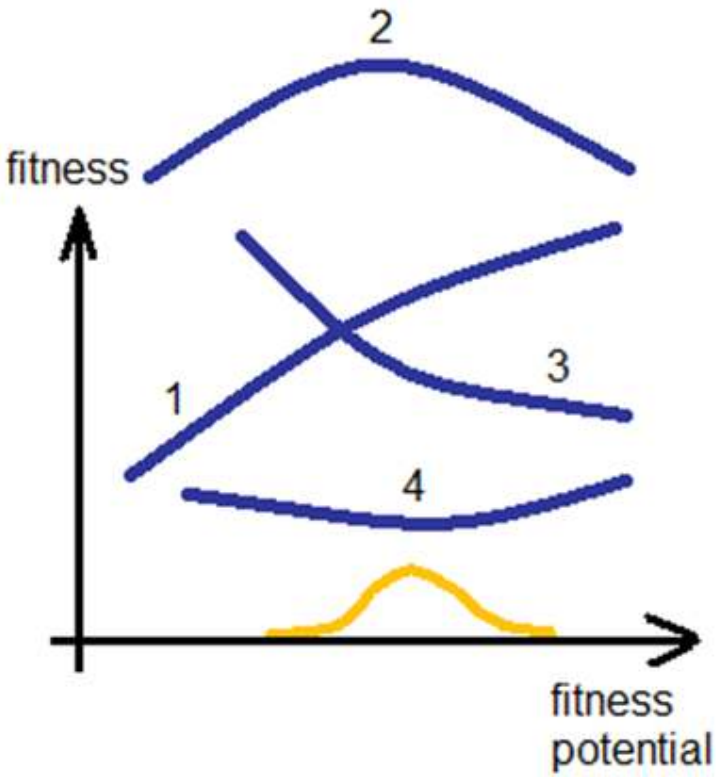




The same fitness landscape induces negative selection in a population with two genotypes if the common genotype is superior (left) and positive selection if it is inferior (right).

Let us now consider genotypes arranged by their values of a quantitative trait, which can be a genotype-level trait such as fitness potential or a phenotypic trait such as body size.

First, we can classify selection on such genotypes into:
i) directional: fitness increases or decreases monotonously, favoring genotypes with one of rare extreme values of the trait,
ii) stabilizing: fitness has one maximum, favoring genotypes possessing intermediate, common values of the trait, and
iii) disruptive: fitness has two maxima, favoring genotypes with either of the two extreme values of the trait.

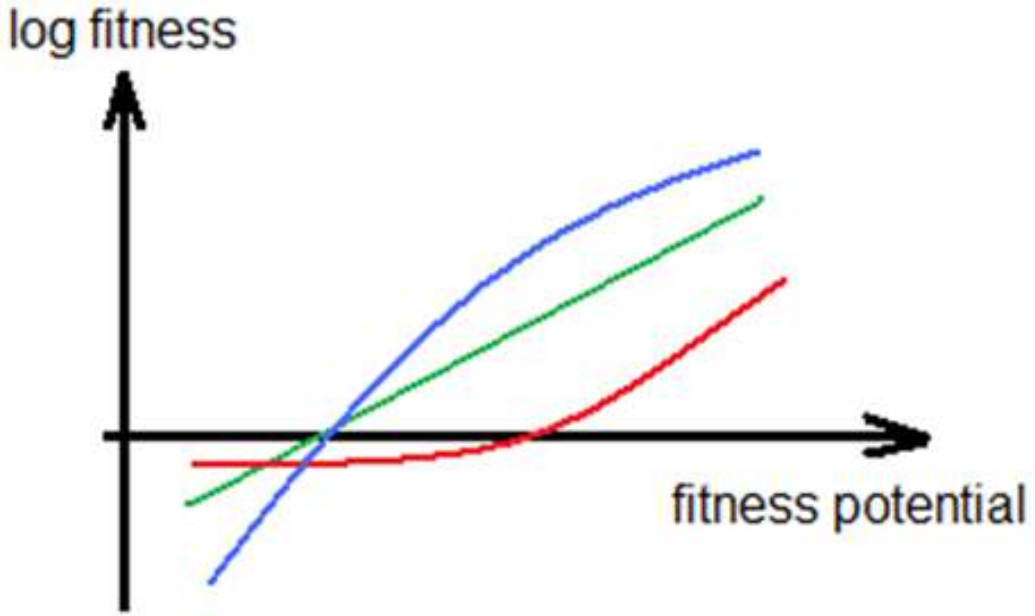


Directional (1 and 3), stabilizing (2), and disruptive (4) selection.

Second, we can classify selection on such genotypes into:

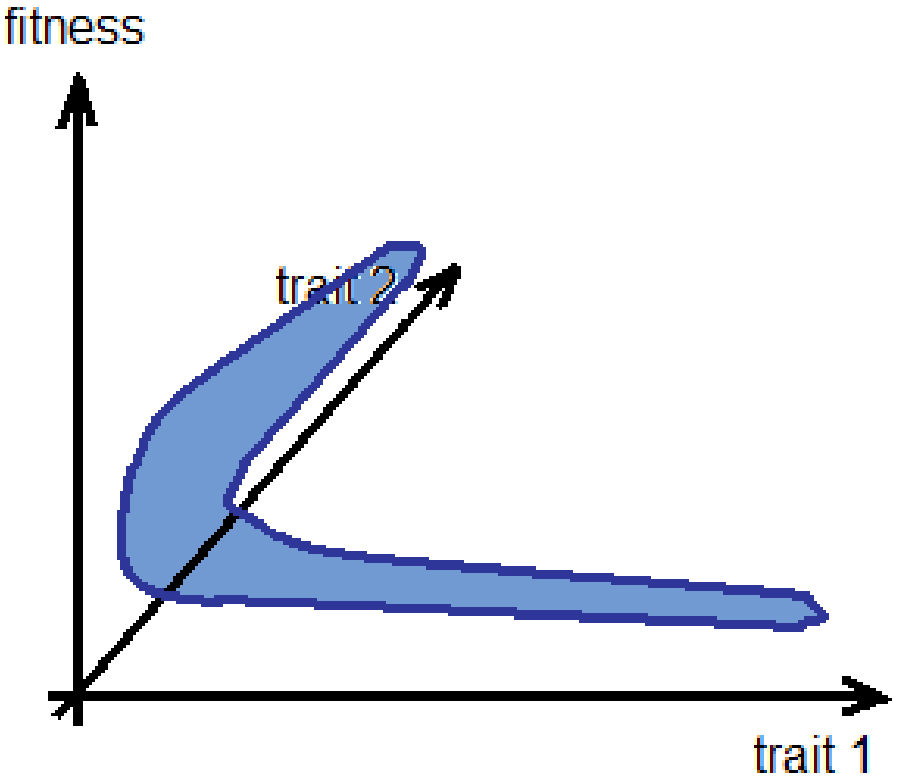
1) narrowing, which reduces the variance of a trait with Gaussian distribution and
2) widening, which increases this variance.

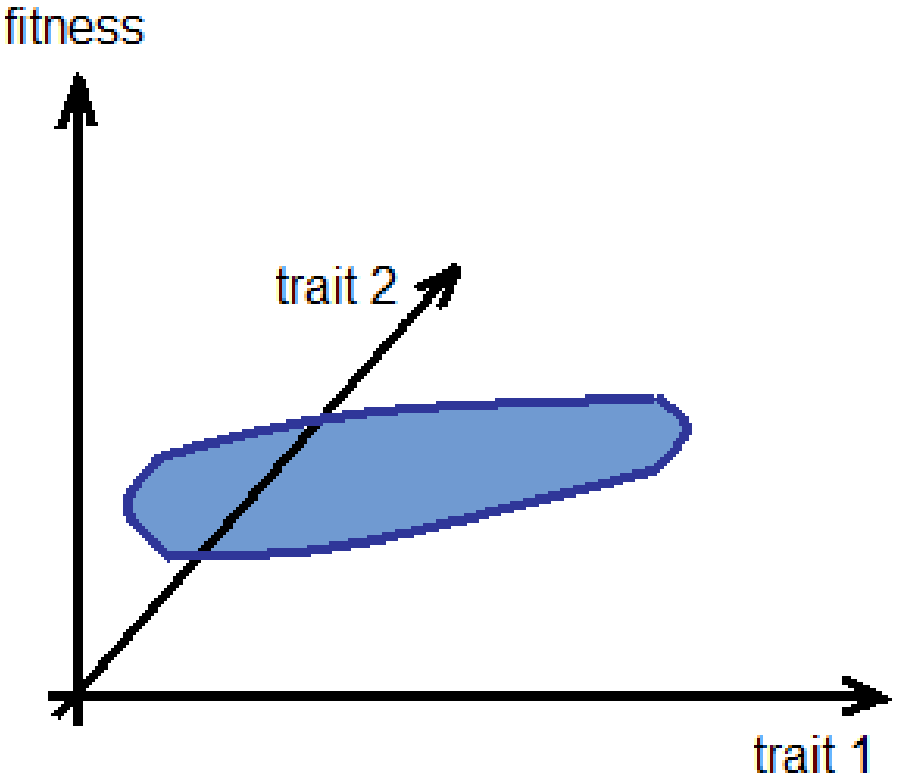
If log fitness is concave (its second derivative is negative everywhere) selection is narrowing, and if log fitness is always convex (positive second derivative) selection is widening. When the log fitness is linear, so that fitness is exponential, the variance of the Gaussian trait does not change.



Narrowing (blue), widening (red) and exponential (log-linear, red) selection. Stabilizing selection is narrowing, disruptive selection is widening, and directional selection and be both narrowing and widening.

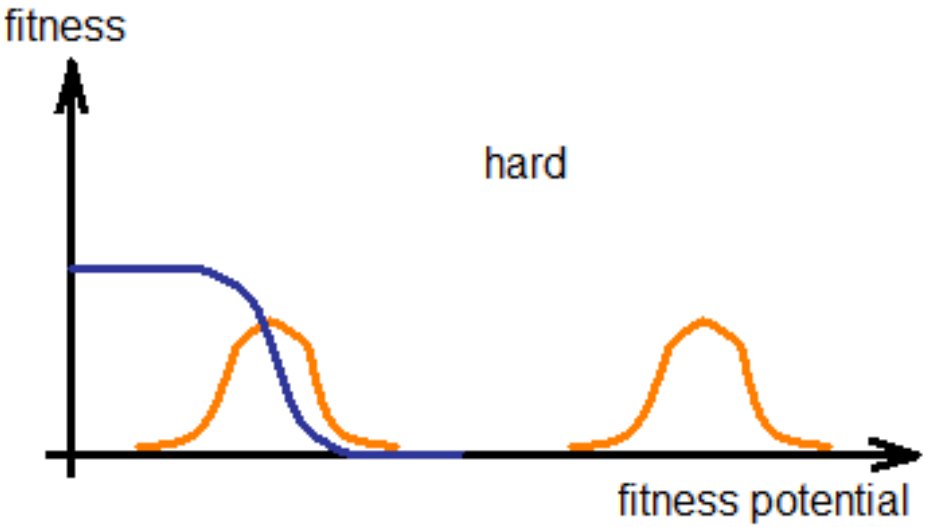
Two opposite kinds of selection are possible with monotonous but multidimensional fitness landscapes: incompatibility (left) and complementation (right) selection. Incompatibility selection can lead to speciation.

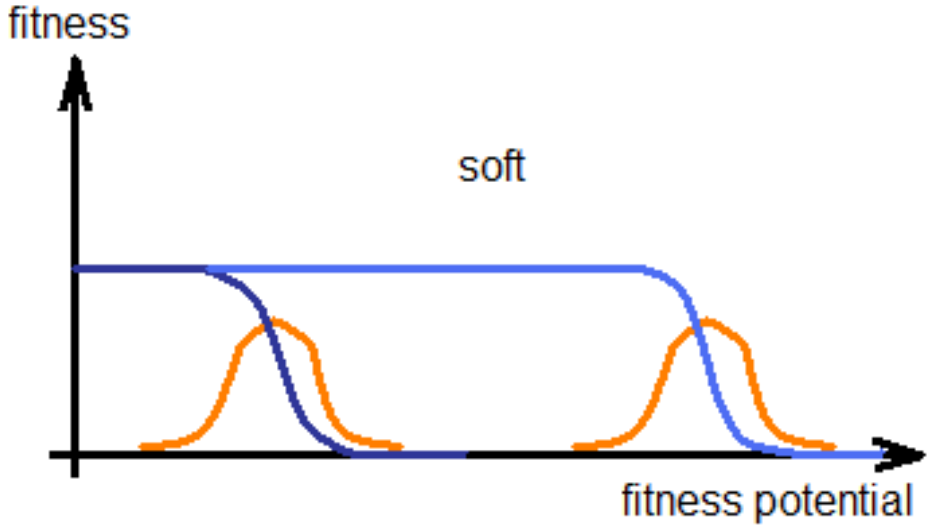




Fitness landscape may well be affected by the population sitting on it.

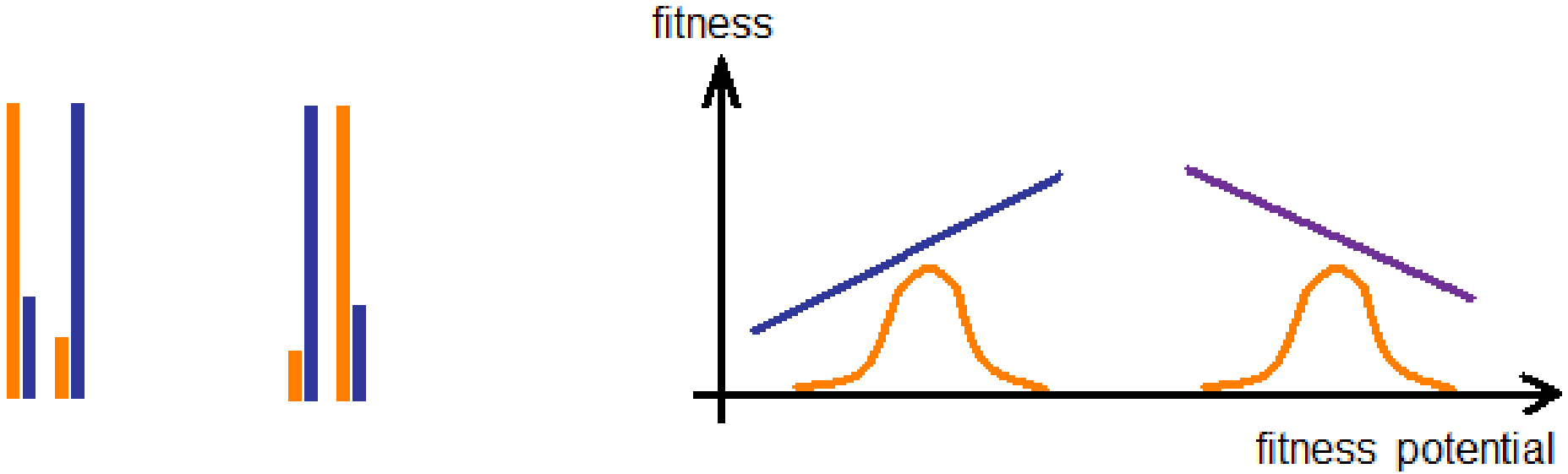
If features of fitness landscape are aligned to position of the population, selection is called soft, as opposed to hard.





Soft selection can naturally result from competition.

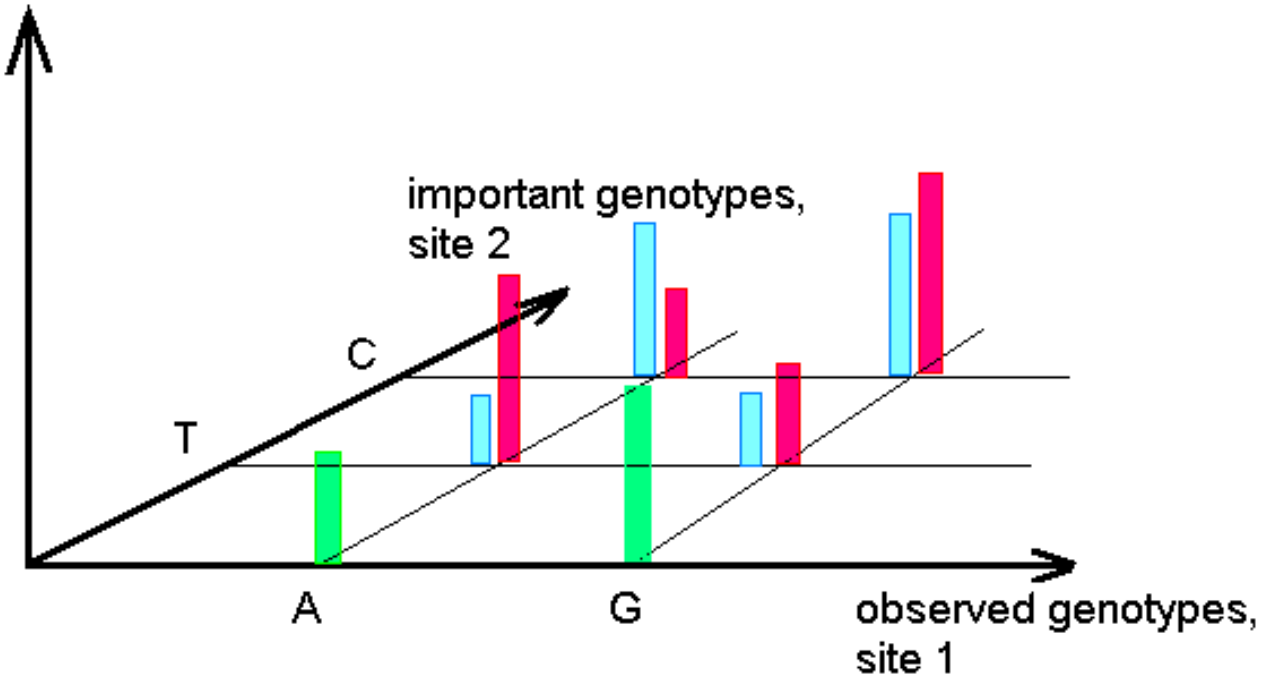
If direction of selection depends on the position of the population, selection is called frequency-dependent.



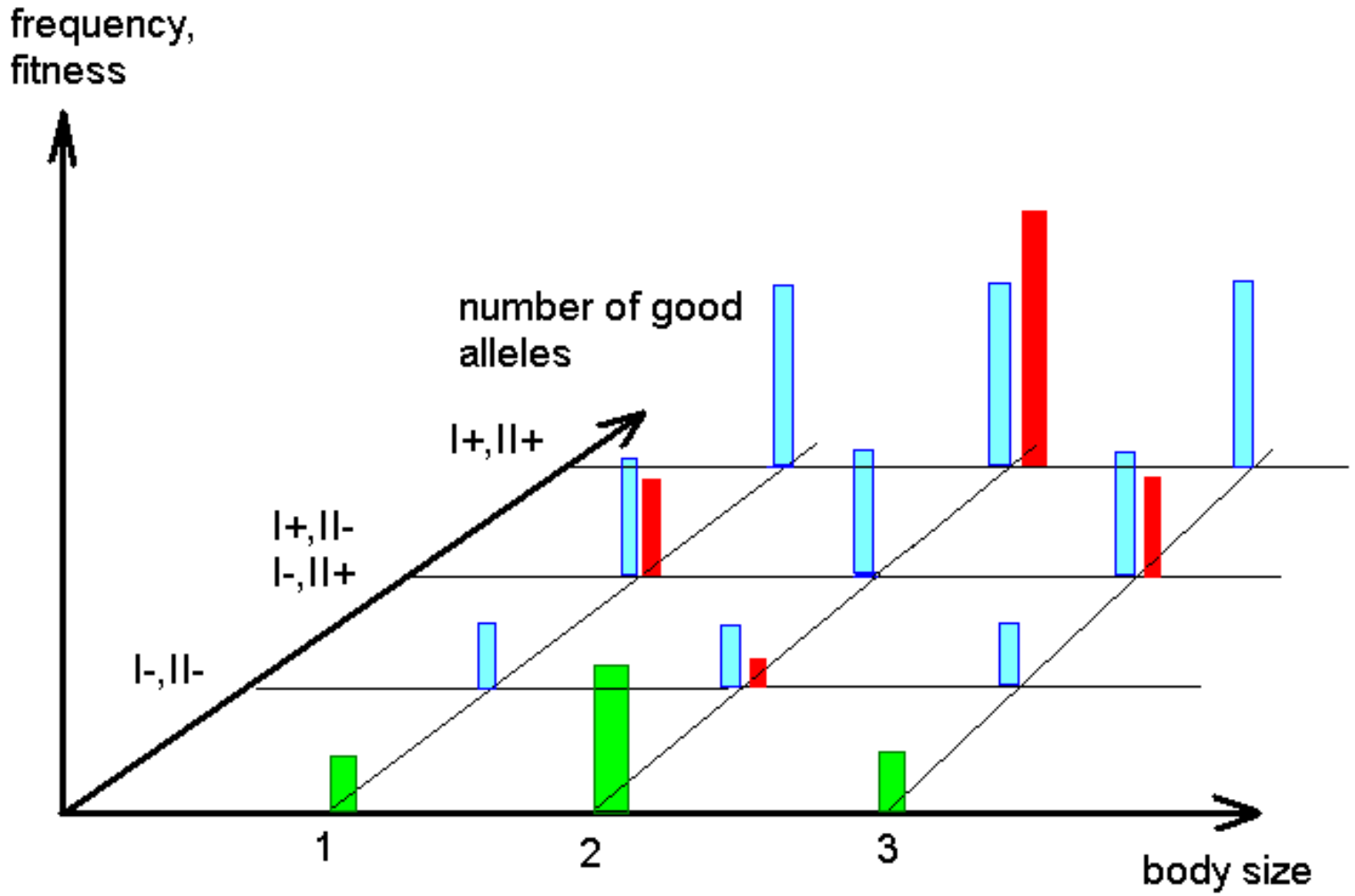
Frequency-dependent selection in the case of two genotypes (left) and a quantitative trait (right).

Frequency-dependent selection can naturally result from different genotypes using different resources. With apomixis such genotypes would form different populations, but with amphimixis they may still interbreed.

Finally, selection acting on a trait can be either real - phenotypes which we watch really affect fitness - or only apparent - phenotypes which we watch do not affect fitness but are connected to some variation which does. This connection can be due to non-independent distribution of variation of different traits or to pleiotropy.

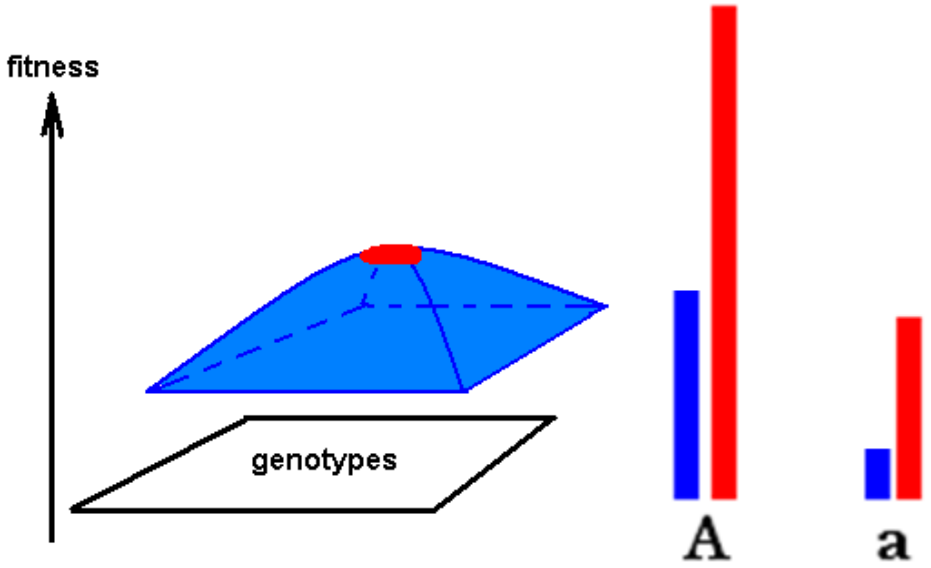


Apparent selection at site 1 (observed) due to non-independent distribution of alleles at sites 1 and 2 (which is under real selection). Allele A at site 1 preferentially occurs, within the studied population, with allele T at locus 2, and allele G at locus 1 is associated with allele C at locus 2. C confers fitness that is higher than fitness conferred by T, and locus 1 does not affect fitness at all. Green bars - apparent fitnesses of A and G, blue bars - real fitnesses of the four genotypes, pink bars - their frequencies.

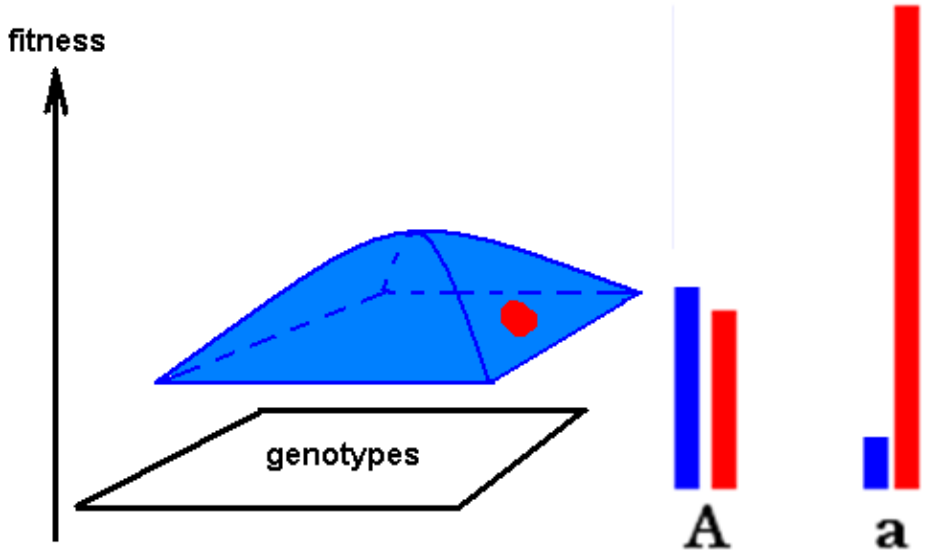


Apparent selection due to pleiotropy. Bad alleles reduce fitness and randomly affect body size (which *per se* do not affect fitness). Individuals with average body size possess, on average, a smaller number of deleterious alleles that those with extreme body sizes, leading to apparent selection.

3) Dual role of selection in evolution: negative vs. positive selection



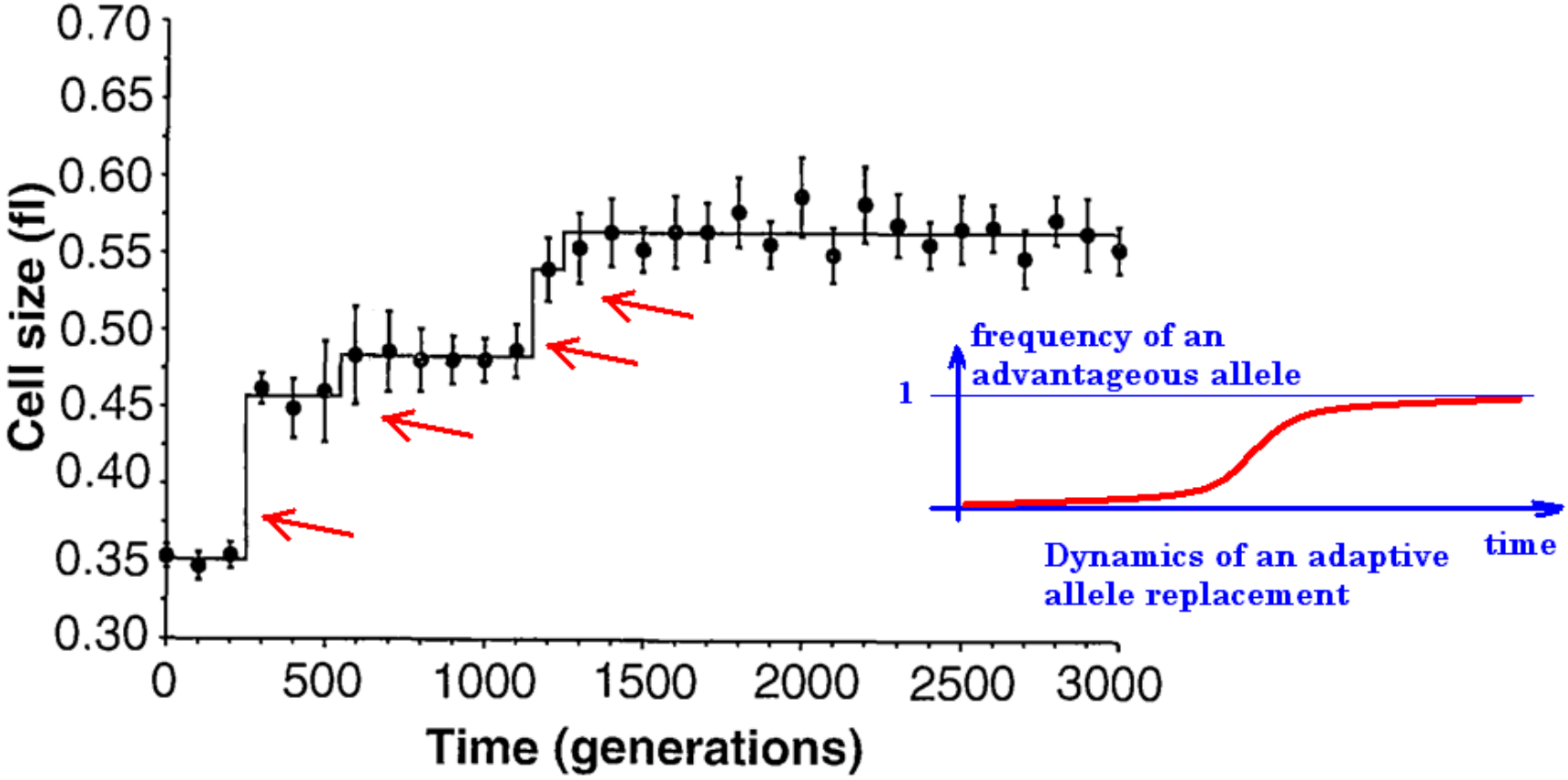
Population sits on top of a fitness peak, so that common genotypes have the highest fitness. Such selection is called negative or purifying (blue bars - genotype frequencies, red bars - fitnesses).



Population sits on a slope of a fitness peak, so that rare genotypes have the highest fitness. Such selection is called positive or Darwinian.

Negative selection maintains *status quo* and prevents changes. Positive selection promotes changes. After positive selection completes its job, the highest-fitness genotype becomes common, and selection becomes negative, on the same fitness landscape. Thus, at any given moment, negative selection is more common than positive selection. Looking for sites of ongoing or recent positive selection is a difficult, exciting, and controversial area of research.

Different modes of selection affect populations differently. Still, the most important outcome of selection from the point of view of Microevolution is an adaptive allele replacement, the replacement of an old, initially common inferior allele (trait state) with an a new, initially rare advantageous allele which confers a higher fitness.



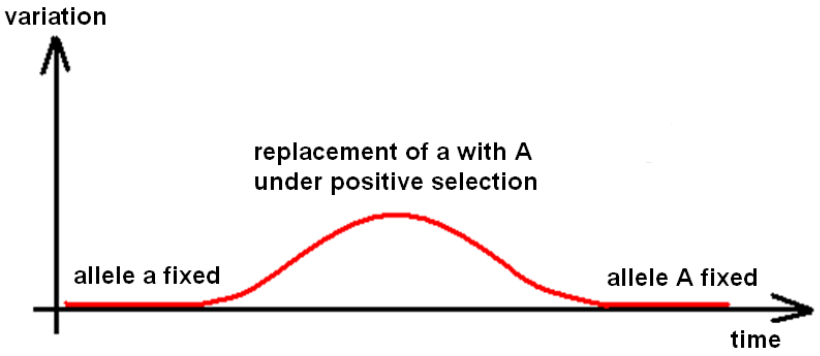
Changes in average cell volume (in femtoliters) of *Escherichia coli* cells in the course of 3000 generations of experimental evolution. There were 4 episodes of sharp increase of this volume. Very likely, each one of them was due to an adaptive allele (genotype) replacement.

The impact of selection on within-population variation

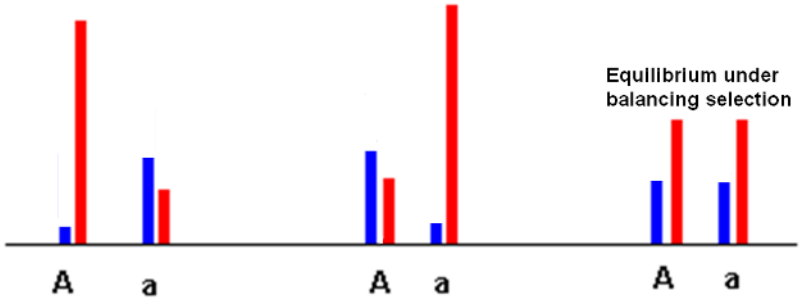
Natural selection is survival of the fittest. One can expect selection acting alone to always remove variation - because only the fittest survive! This is mostly true, but not exactly.

1) Negative selection always reduces genetic variation - that is why it is also called purifying selection.

2) Positive selection increases variation temporarily (creating a transitive polymorphism), after which selection becomes negative and population again becomes monomorphic.



3) Only if selection is frequency-dependent and favors rare genotypes, genetic variation can be maintained indefinitely.



4) Two ways of characterizing the action of selection

Selection is a complex process, and it is impossible to fully describe it by just a single number. Instead, it is described by f(w) - the density of genotypes' fitness w. Realistically, f(w) is confined between 0 and some maximal value w_{max}.

The mean value of f(w), M[f(w)],

$$W = \int_0^{w_{max}} w f(w) dw$$

is called the mean population fitness. However, W in itself is not a useful characteristic of selection, because it changes if w is multiplied by a constant. Instead, let us introduce the following two characteristics:

the load L: $L = (w_{max}-W)/w_{max} = 1-W/w_{max}$

and

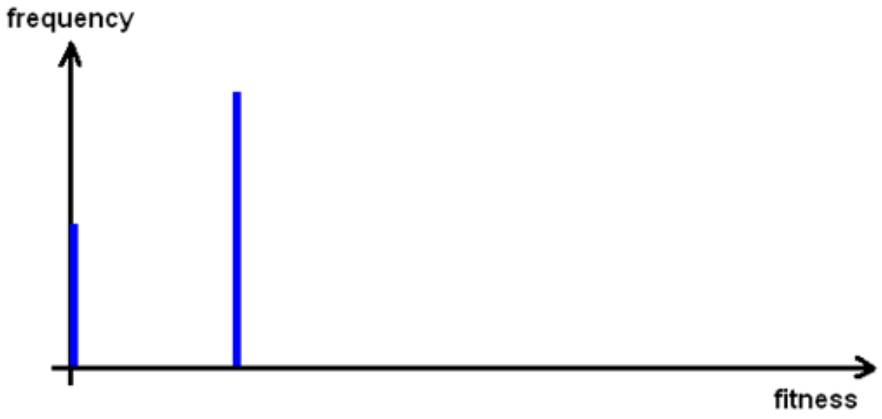
the variance of relative fitness V: $V = Var[f(w)/W] = \int_0^{w_{max}} (w/W-1)^2 f(w) dw$

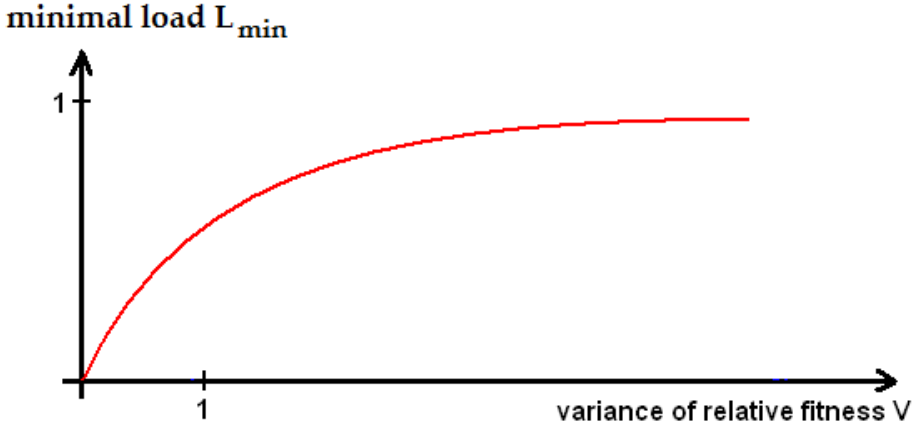
Both L and V depend only on relative fitnesses, which is good. Still, they are quite different - L compares all fitnesses to its highest value, and V - to its mean value.

A distribution of fitness, which produces a low value of V, because almost all genotypes have the same fitness, but a high value of L, because of a very small proportion of exceptionally fit genotypes.

Thus, if we know only L (but not f(w)!), we cannot recover V (and *vice versa*).

Still, knowing V (which is much easier to measure than L) is sufficient to determine the minimal boundary of L.





f(w) which produces the minimal L, consistent with a particular V, is confined to just two points, 0 and w_{max}.

Miinimal L = V/(V+1), as a function of V.

V is of crucial importance, due to Fisher Fundemental Theorem.

Although harder to measure, L is also an important characteristic. It determines the minimal maximal fecundity which is necessary to sustain the population under particular selection. For example, if L = 0.8, the most fit individuals must, on average, produce at least 5 offspring.

5) Action of selection: Fisher's Fundamental Theorem

Let us study the impact of selection, acting alone, on a variable population.

Suppose that there are n different genotypes a_i: a_1, ... a_n. Their frequencies are $[a_i]$, and the total population size is N. Thus, there are $N[a_i]$ individuals of genotype a_i. Fitness of an a_i individual is w_i. Then, the number of a_i individuals in the next generation is $N[a_i]w_i$. If the genotypes breed true (apomixis), the frequency of a_i in the next generation, $[a_i]_{t+1}$, is:

$$[a_i]_{t+1} = N[a_i]w_i / \sum_{j=1}^{n} N[a_j]w_j = [a_i]w_i / W \quad \text{where} \quad W = \sum_j [a_j]w_j$$

is the mean population fitness.

This is the key equation describing the impact of selection on heritable variation.

We already made 2 discoveries:

1) Population size does not affect the dynamics of genotype frequencies - N disappeared from the equation. Can you explain in words, why?

2) If we multiply all fitnesses by the same positive constant, the dynamics of genotype frequencies would not be affected. Can you explain in words, why?

Let us now calculate the mean population fitness in the next generation. We have to take genotype frequencies, as they will be in the next generation, multiply each one by its fitness, and add all the results:

$$W_{t+1} = \sum_j [a_j]_{t+1} w_j = \sum_j [a_j] w_j^2 / W$$

Now let us calculate the relative increment of the mean population fitness after one generation of selection, $\Delta W/W = (W_{t+1} - W)/W$:

$$\Delta W/W = W^{-1}\left(\sum_j [a_j] w_j^2 / W - W\right) = W^{-2}\left(\sum_j [a_j] w_j^2 - W^2\right) = V[w_j/W]$$

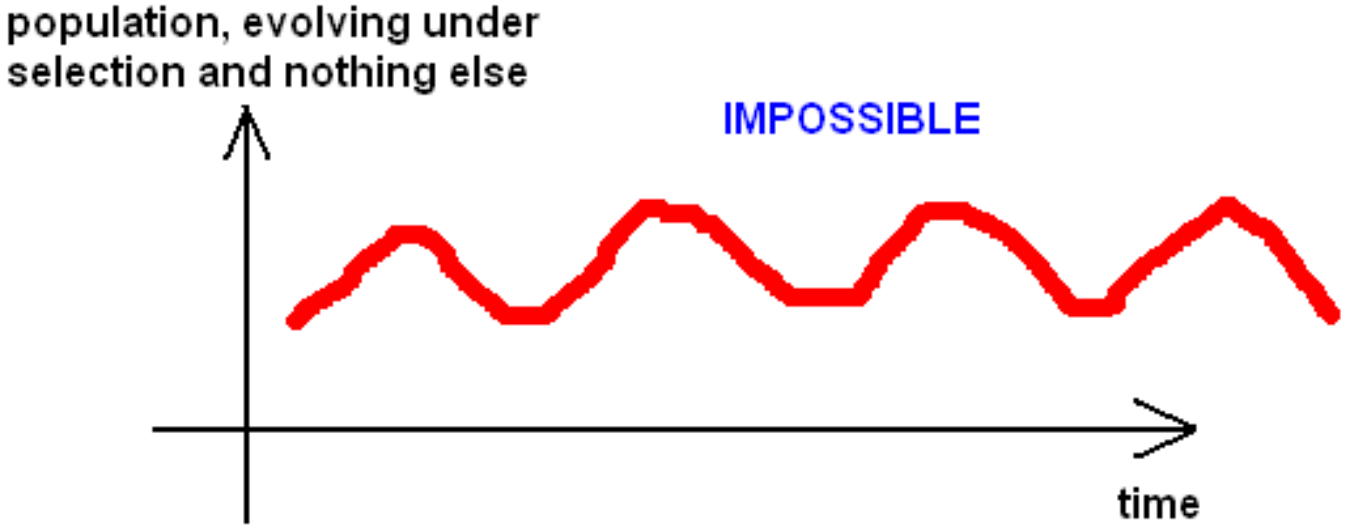
where $V[w_j/W]$ is the variance of relative fitness. Let us prove the last equality. By definition of variance:

$$V[w_j/W] = \sum_j (w_j/W - 1)^2 [a_j] = W^{-2}\left(\sum_j w_j^2 [a_j] - 2W \sum_j w_j [a_j] + W^2 \sum_j [a_j]\right) =$$

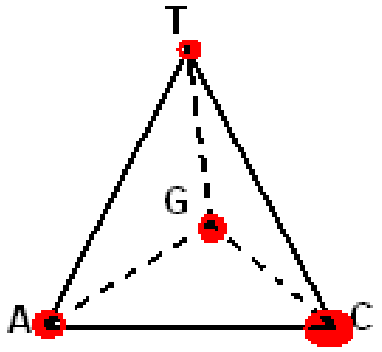
$$= W^{-2}\left(\sum_j [a_j] w_j^2 - 2W^2 + W^2\right) = W^{-2}\left(\sum_j [a_j] w_j^2 - W^2\right)$$

In words, we showed that the relative increment of the mean population fitness equals to the within-population variance of relative fitness. This result is known as Fisher's Fundamental Theorem of Natural Selection.

Why FFT is so important? Because it captures the essence of what selection does, and leads to important insights. In particular, FFT implies that changes under selection are irreversible. Indeed, mean fitness W of an evolving population always increases, and, thus, population cannot return to where it once was.



Instead, the population climbs, on the fitness landscape, until it reaches the highest peak, which corresponds to fixation of the most fit genotype, among those available.



Thus, selection acting alone is a remarkably simple force. Only when combined with other forces, it can lead to more complex dynamics. Still, dynamics of genetic variation within a population are much simpler than, for example, dynamics of sizes of interacting populations, where cycling and other complex phenomena are common.

Of course, FFT also applies when fitness varies continuously, and is described by density f(w). In this case, density after selection f'(w) is provided by:

$$f'(w) = wf(w)/W$$

where W, as before, is the mean population fitness:

$$W = \int_0^{w_{\max}} wf(w)\,dw$$

Then, relative increment of mean population fitness is equal to variance of relative fitness:

$$(W'-W)/W = \int_0^{w_{\max}} wf'(w)\,dw/W^2 - 1 = \int_0^{w_{\max}} (w/W)^2 f(w)\,dw - 1 = V$$

where

$$V = Var[f(w)/W] = \int_0^{w_{\max}} (w/W - 1)^2 f(w)\,dw$$

6) Action of selection: the dynamics of an allele replacement

Let us now consider just two alleles (genotypes), A and a, but treat the action of selection more quantitatively. FFT tells us that the more fit allele (say, A), will eventually replace a - but how exactly will this occur?

If A and a individuals leave, on average, w_A and w_a offspring, respectively, in the next generation the frequency of A, x, will be

$$x_{t+1} = w_A x/[w_A x + w_a(1-x)]$$

Let us define selective advantage of A over a as $s = (1- w_a/w_A)$. Clearly, $s = 0$ if fitnesses of A and a are equal, $s > 0$ if $w_a < w_A$, and $s < 0$ if $w_a > w_A$. Then, after dividing all terms over w_A, we obtain:

$$x_{t+1} = x/[x + (1-s)(1-x)] = x/[1 - s(1-x)]$$

Assuming that selection is weak, so that s is small and $x/[1 - s(1-x)] \approx x + sx(1-x)$, we obtain:

$$x_{t+1} = x + sx(1-x) \text{ or } \Delta x = x_{t+1} - x = sx(1-x)$$

or, if time is continuous,

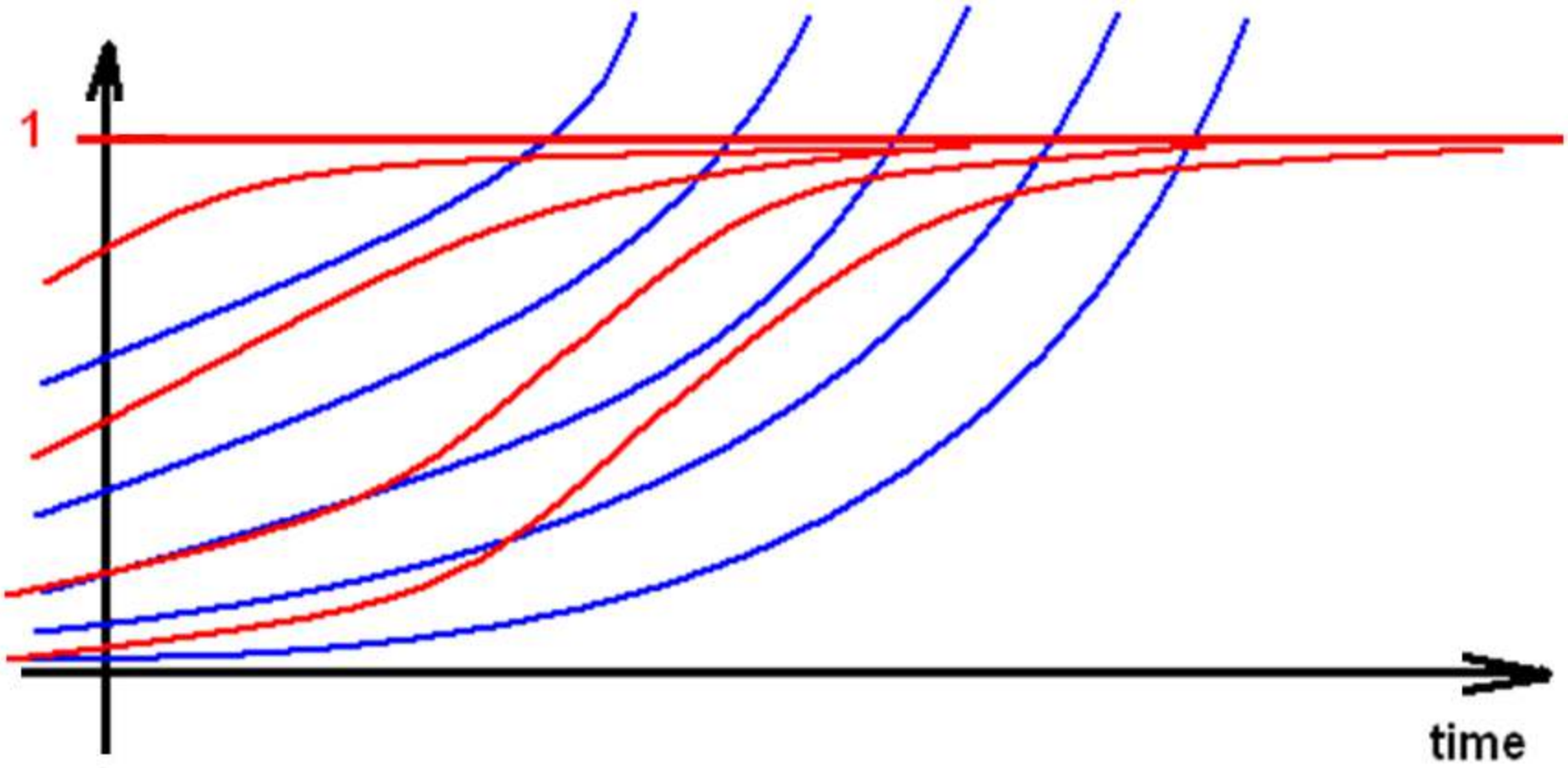
$$dx/dt = sx(1-x)$$

This is the key dynamical equation of Microevolution, describing an allele replacement driven by positive selection.

Let us compare this equation the equation that describes exponential population growth (s > 0) or radioactive decay (s < 0).

dx/dt = sx - population growth (blue)

dx/dt = sx(1-x) - allele replacement (red).



Solution of the equation for exponential growth or decay:

$$\frac{dx}{dt} = sx$$

$$\frac{dx}{x} = s\,dt \quad \Rightarrow \quad \int_{x_0}^{x(t)} \frac{dy}{y} = s\int_{t_0}^{t} d\tau \quad \Rightarrow \quad \ln\frac{x(t)}{x_0} = s(t - t_0)$$

$$\Rightarrow \quad x(t) = x_0\, e^{s(t-t_0)}$$

Now, let us solve the equation of an allele replacement.

$$\frac{dx}{dt}=sx(1-x) \quad \Rightarrow \quad \frac{dx}{x(1-x)}=sdt \quad \Rightarrow \quad \int_{x_0}^{x(t)}\frac{dy}{y(1-y)}=s\int_{t_0}^{t}d\tau$$

$$\Rightarrow \quad \ln\frac{x(t)/(1-x(t))}{x_0/(1-x_0)}=s(t-t_0)$$

Let us make sure that the last transformation is correct:

$$\int_{x_0}^{x(t)}\frac{dy}{y(1-y)}=\int_{x_0}^{x(t)}\frac{dy}{y}+\int_{x_0}^{x(t)}\frac{dy}{1-y}=\ln x(t)-\ln x_0-\ln(1-x(t))+\ln(1-x_0)=\ln\frac{x(t)}{1-x(t)}-\ln\frac{x_0}{1-x_0}$$

Above, we used this formula $\int\frac{dy}{ay+b}=\frac{1}{a}\ln|ay+b|$

To simplify formulae, let us define $C_0=\frac{x_0}{1-x_0}$ **. Then,**

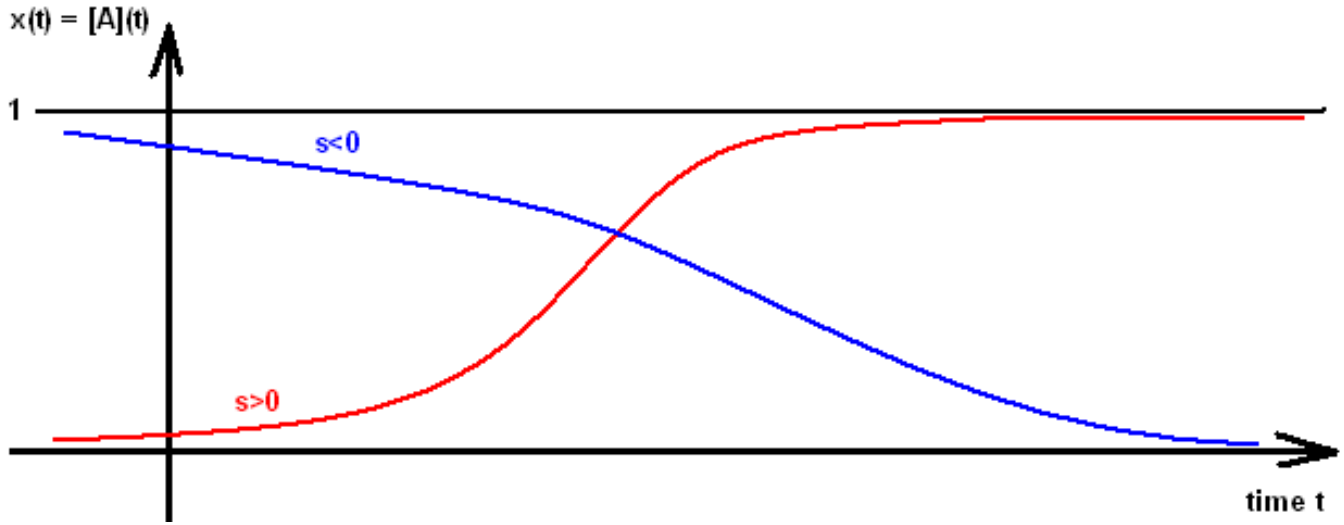
$$\ln\frac{x(t)/(1-x(t))}{x_0/(1-x_0)}=s(t-t_0) \quad \Rightarrow \quad \ln x(t)/(1-x(t))=C_0 s(t-t_0) \quad \Rightarrow$$

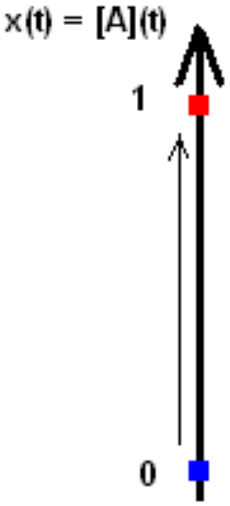
$$\Rightarrow \quad x(t)/(1-x(t))=C_0 e^{s(t-t_0)}$$

Finally, we can recover x(t), the function that shows how the frequency of allele A changes when time flows:

$$x(t)=C_0 e^{s(t-t_0)}/(1+C_0 e^{s(t-t_0)})=1/(1+(1/C_0)e^{-s(t-t_0)})=1/(1+\frac{(1-x_0)}{x_0}e^{-s(t-t_0)})$$

Naturally, x increases with time if s > 0 and decreases if s < 0. What if s = 0? The exact trajectory of x depends on both s and the initial allele frequency x_0.





There are also two exceptional initial frequencies $x_0 = 0$ and $x_0 = 1$, such that x does not change. They are called equilibria. With s > 0, equilibrium $x_0 = 0$ is unstable (small deviations from it will increase) and equilibrium $x_0 = 1$ is stable (small deviations from it will decrease) , and it is the other way around with s < 0.

Why x = 0 and x = 1 are equilibria, biologically?

This brief review is sufficient to understand how selection operates alone on unstructured variation.

The same in detail

Studies of any changes are based on investigating dynamical models. A dynamical model consists of a sufficiently detailed description of the changing object at a particular moment of time and of a transformation law that describes how these changes occur. Time can be treated as continuous or as discrete. The description of an object is provided by variables, and all possible combinations of their values constitute its phase space. Changes of the object can be represented by trajectories within the phase space. In addition to variables, a model usually contains parameters.

Let us build and study a simple deterministic dynamical model with one variable.

Consider a population of N individuals with two possible genotypes, A and a. Individuals breed true. Generations do not overlap, so that time is discrete. The expected numbers of offspring of an individual of genotypes A and a are w_A and w_a, respectively. Unless w_A and w_a are identical, selection operates within the population.

The numbers of offspring with genotypes A and a will be almost precisely $N[A]w_A$ and $N[a]]w_a$, respectively, as long as N is large enough.

The frequency of A in the next generation, $[A]_{n+1}$, is provided by the ratio of the number of offspring of genotype A over the total number of offspring. Because [a] = 1 - [A], full description of our population consists of just one number, for example [A], and we obtain the following transformation law:

$$[A]_{n+1} = w_A[A]/\{w_A[A]+w_a(1-[A])\}$$

It appears that dynamics of the population depends on two parameters, w_A and w_a. However, if we divide both the numerator and the denominator of the right-hand side by, say, w_A, we can see that this is not so:

$$[A]n+1 = [A]/\{[A]+(w_a/w_A)(1-[A])\}$$

The following four statements summarize what we achieved so far:

1) if N is very large, the model is deterministic,
2) the phase space of our model is one-dimensional,
3) the transformation law of our model does not depend of N,
4) only the ratio of fitnesses of the two genotypes, w_a/w_A, is important.

Let us create a continuous-time version of the model. Dynamics with discrete and continuous time can be very different. Still, if selection is weak, *i. e.* that w_a and w_A are close to each other, there will be no long jumps and time can be treated in either way. Let us define selective advantage of A over a as $s = 1 - w_a / w_A$. $s = 0$ if fitnesses of A and a are equal, $s > 0$ if $w_a < w_A$, and $s < 0$ if $w_a > w_A$. Then:

$$[A]_{n+1} = [A]/\{[A] + (1-s)(1-[A])\} = [A]/\{1 - s(1-[A])\}$$

Selection is weak if s is small, so that w_a/w_A is close to 1. Then, we can use an approximation $1/(1-e) \approx 1 + e + O(e^2)$ (e means a small number):

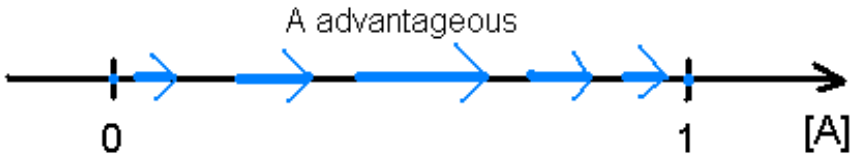
$$[A]n+1 = [A] + s[A](1-[A])$$

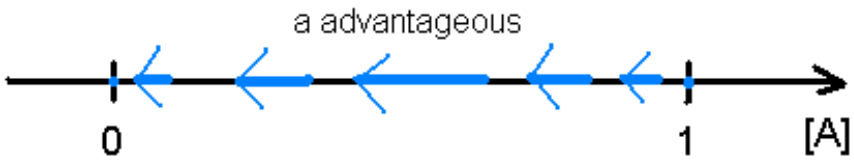
Assume that velocity of [A], d[A]/dt, is equal to its increment between two successive moments in discrete-time treatment, $[A]_{n+1} - [A]$:

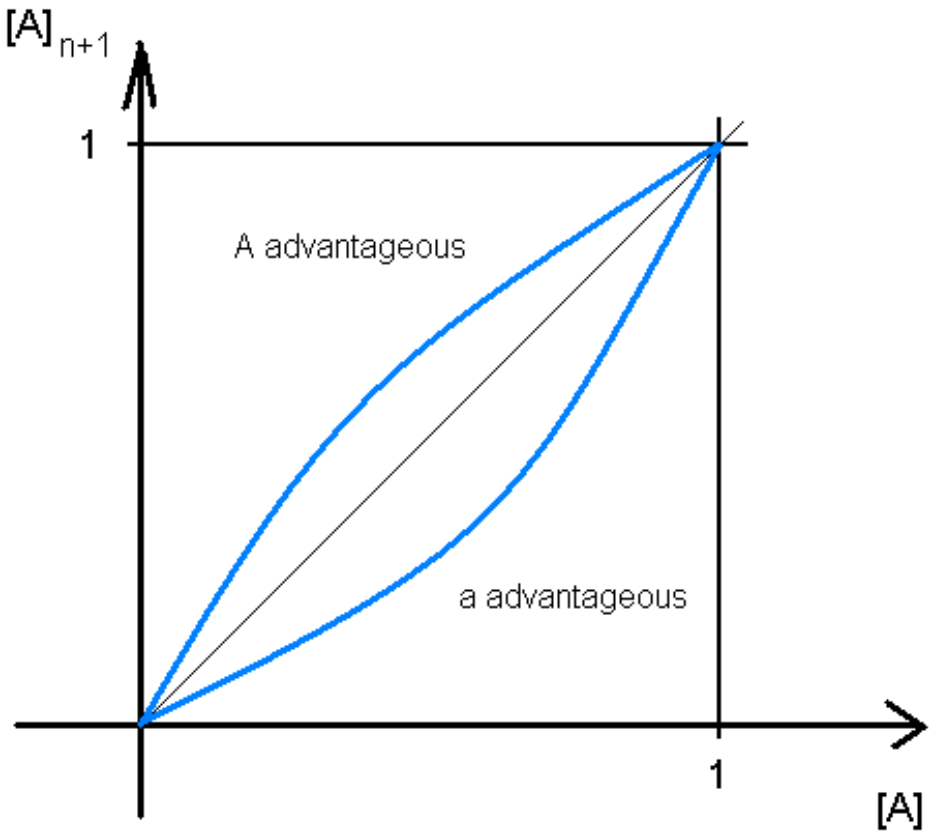
$$d[A]/dt = s[A](1-[A])$$

This differential equation describes the most important process in Microevolution, an allele replacement driven by natural selection. Essentially the same equation also plays the key role in population ecology, where it describes population growth with self limitation (r is per capita growth rate and K is carrying capacity):

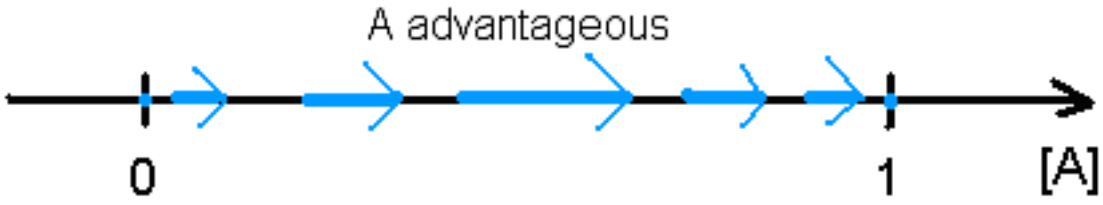
$$dN/dt = rN(1-N/K)$$

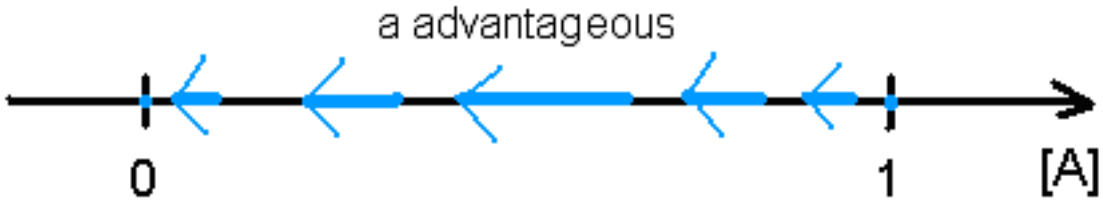


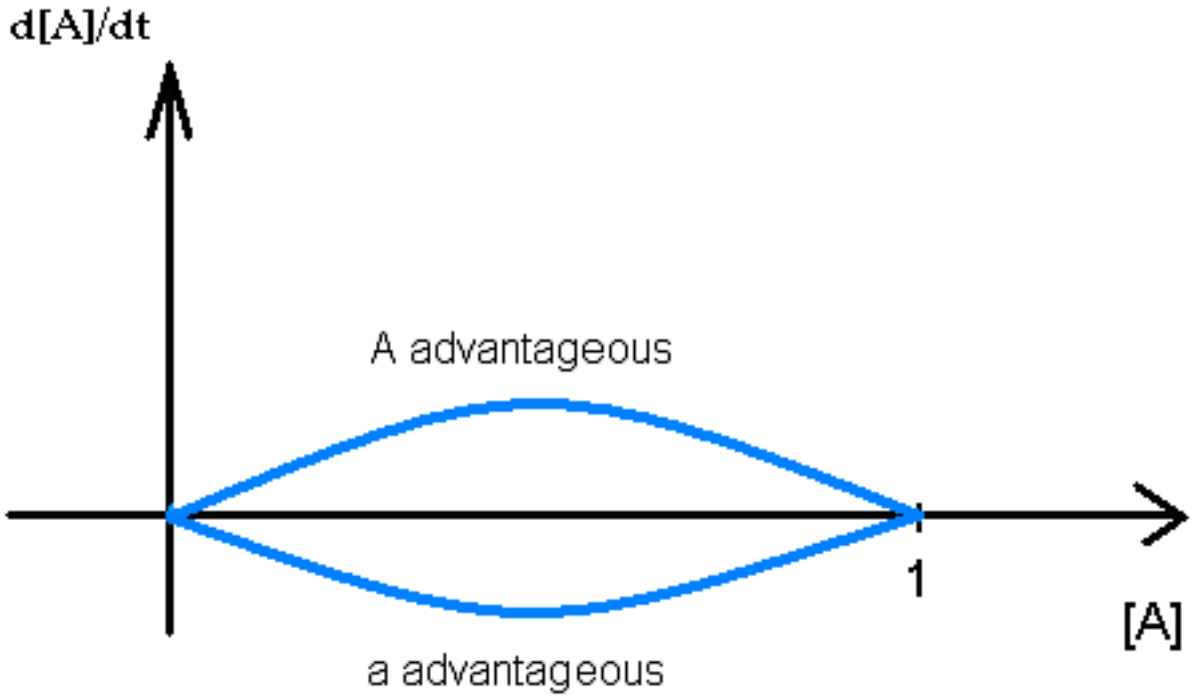




Two ways of presenting the same model with discrete time graphically: by vectors that describe jumps from a value [A] to the corresponding value of $[A]_{n+1}$ (two top figures) and by a function $[A]_{n+1} = f([A])$ (bottom).







Two ways of presenting the same model with continuous time graphically: by vectors that describe velocities of [A] corresponding to its current values (two top figures) and by a function d[A]/dt = f([A]) (bottom).

Comprehensive solution of a dynamical model is a family of trajectories which show, for all possible initial values, how the variables will changes in the future. Model of selection-driven allele replacement is simple enough to be solved explicitly (x = [A]):

$$\frac{dx}{dt} = sx(1-x)$$

Gather different variables at different sides (useful mneumonics):

$$\frac{dx}{x(1-x)} = sdt$$

Rewrite the differential equation in integral form:

$$\int_{x_0}^{x(t)} \frac{dy}{y(1-y)} = s\int_{t_0}^{t} d\tau$$

The right-hand side integral is simply $s(t-t_0)$, and the left-side integral is:

$$\int_{x_0}^{x(t)} \frac{dy}{y(1-y)} = \int_{x_0}^{x(t)} \frac{dy}{y} + \int_{x_0}^{x(t)} \frac{dy}{1-y} = \ln x(t) - \ln x_0 - \ln(1-x(t)) + \ln(1-x_0)$$

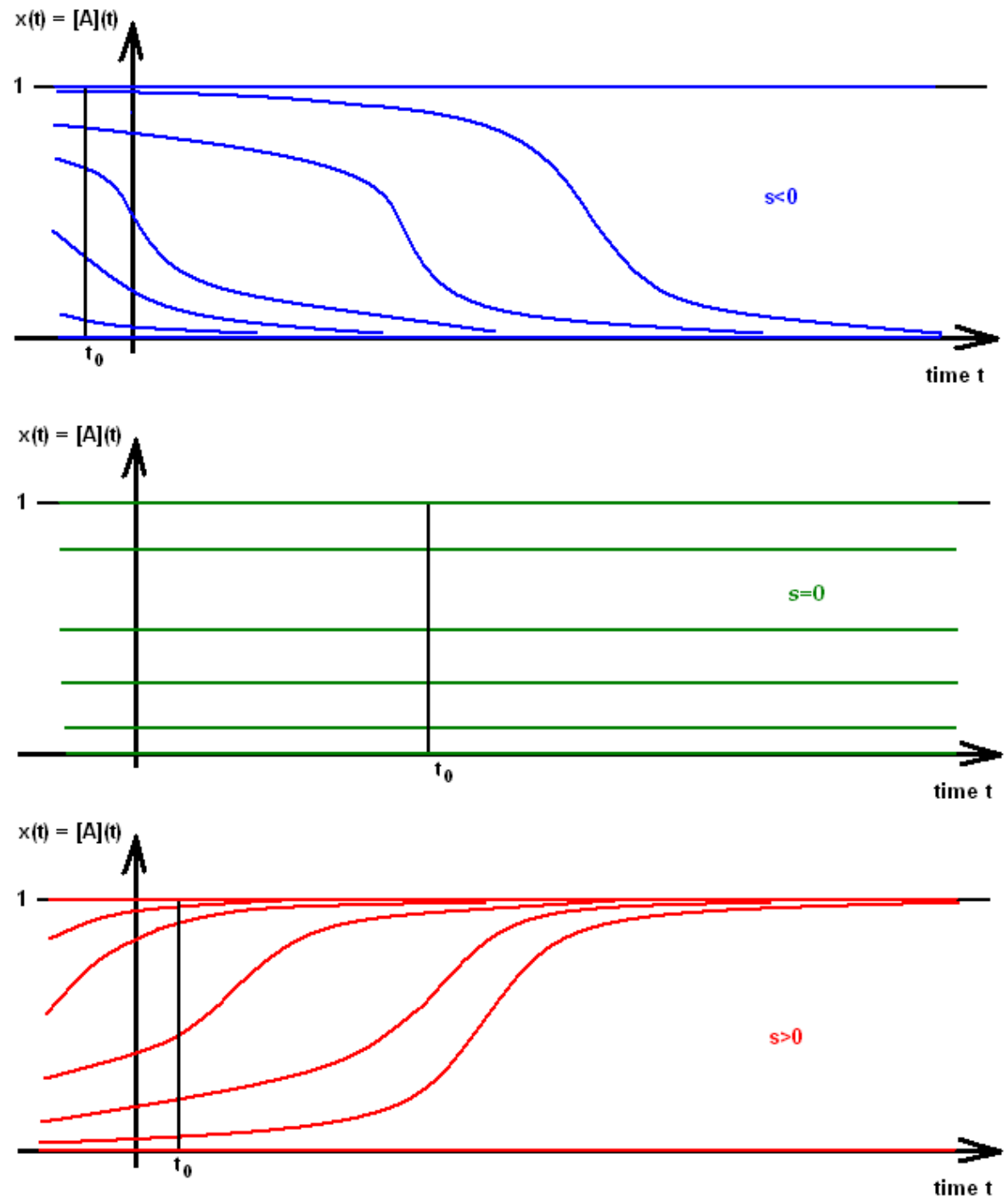
because: $\int \frac{1}{ay+b} dy = \frac{1}{a} \ln|ay+b|$ **. Further, the right-habd side is:**

$$\ln \frac{x(t)/(1-x(t))}{x_0/(1-x_0)}$$

Thus, we now need to recover x(t) from:

$$\ln \frac{x(t)/(1-x(t))}{x_0/(1-x_0)} = s(t-t_0)$$

$$x(t) = C_0 e^{s(t-t_0)} / (1 + C_0 e^{s(t-t_0)}) = 1/(1 + (1/C_0) e^{-s(t-t_0)}) = 1/(1 + \frac{(1-x_0)}{x_0} e^{-s(t-t_0)})$$

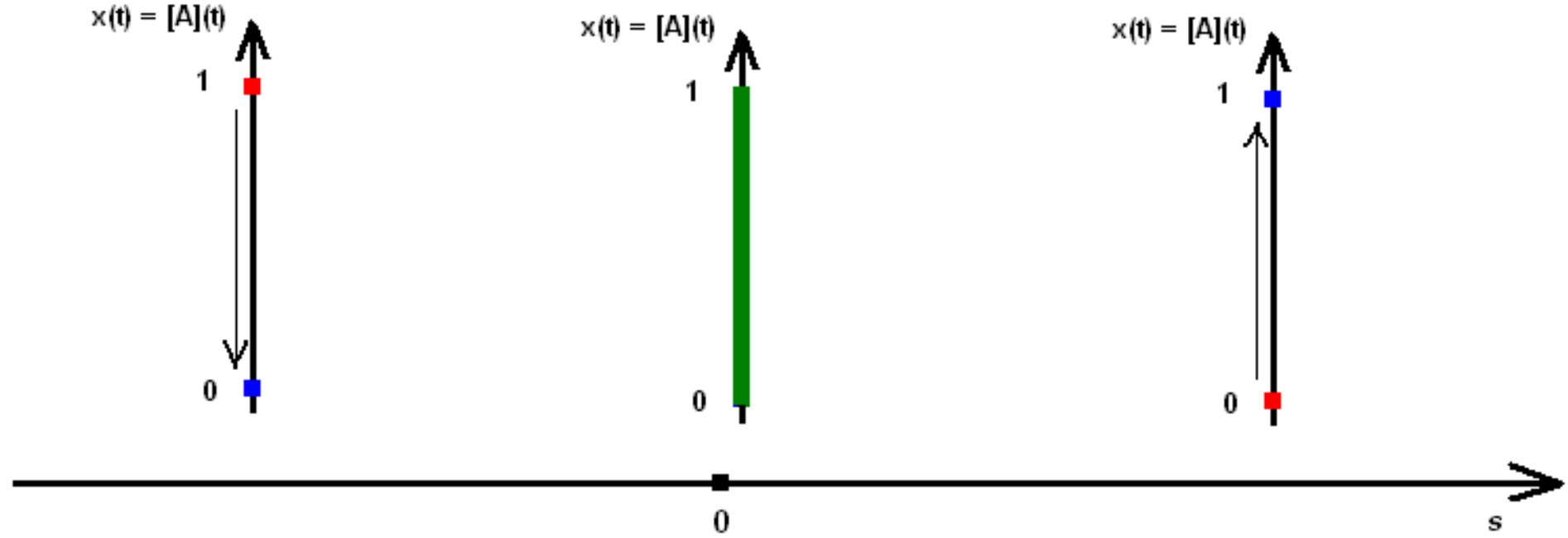


This family of x(t) is a comprehensive solution of our model. Each trajectory corresponds to its own initial frequency of A, with the value x_0 at time t_0. Moreover, the dynamics of the frequency of A also depend on s, and for each value of s there exists its own family of trajectories. Naturally, x(t) increases with time if $s > 0$, decreases if $s < 0$, and does not change if $s = 0$.

We can also investigate our model only qualitatively, finding attractors and their stability. This this only way, when there is not explicit comprehensive solution.

There are two exceptional initial frequencies of A, $x_1 = 0$ and $x_2 = 1$. Trajectories with such initial frequencies are flat, *i. e.* if x is equal to 0 or to 1 at some moment, it never changes and retains this value forever. Biologically, this result is obvious.

Values of variables that do not change are called equilibria. With $s > 0$, equilibrium $x_1 = 0$ is unstable, in the sense that a small deviation from it will increase and equilibrium $x_2 = 1$ is stable, because a small deviation from it will decrease. It is the other way around with $s < 0$. With $s = 0$, every value of x is an equilibrium, and all these equilibria are neutral.



Stable, unstable, and neutral equilibria are blue, red, and green, respectively.

To find equilibria, we replace a dynamical equation with an algebraic equation. With discrete time, we ask that the next state is identical to the current state:

$$[A] = [A]/\{[A]+(w_a/w_A)(1-[A])\}$$

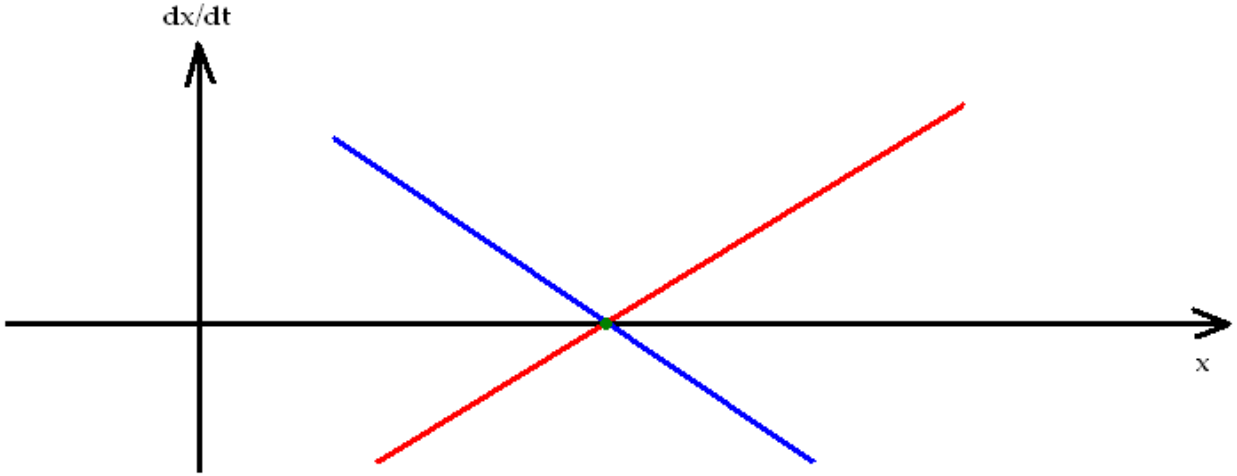
which is a quadratic equation with two roots, $[A]_1 = 0$ and $[A]_2 = 1$. Only if $w_a/w_A = 1$, every value of [A] satisfies this equation.

With continuous time, we ask that the velocity is zero:

$$0 = s[A](1-[A])$$

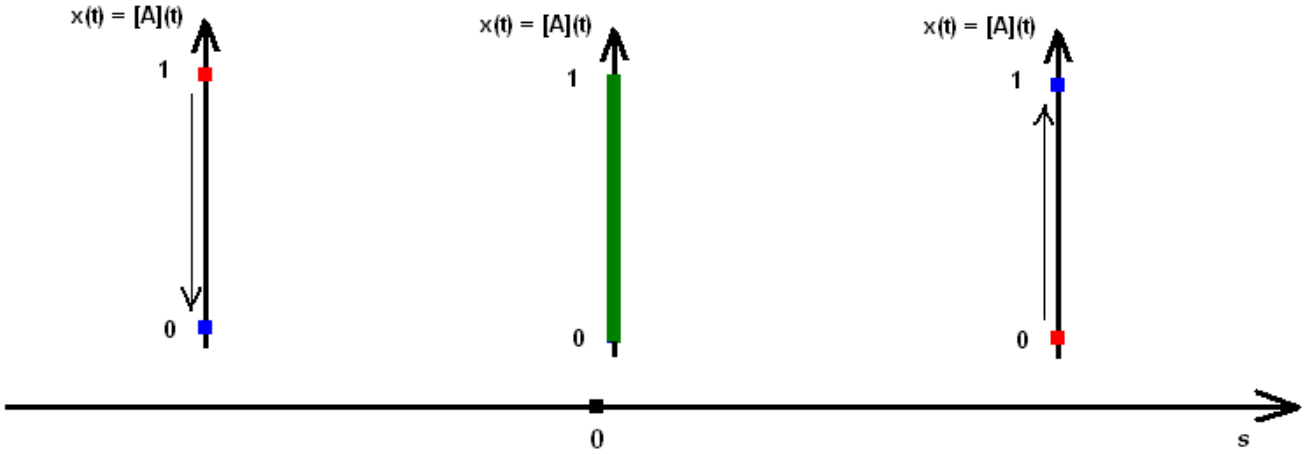
which has the same roots.

Local stability of an equilibrium is determined by whether small deviations from it increase or decrease.

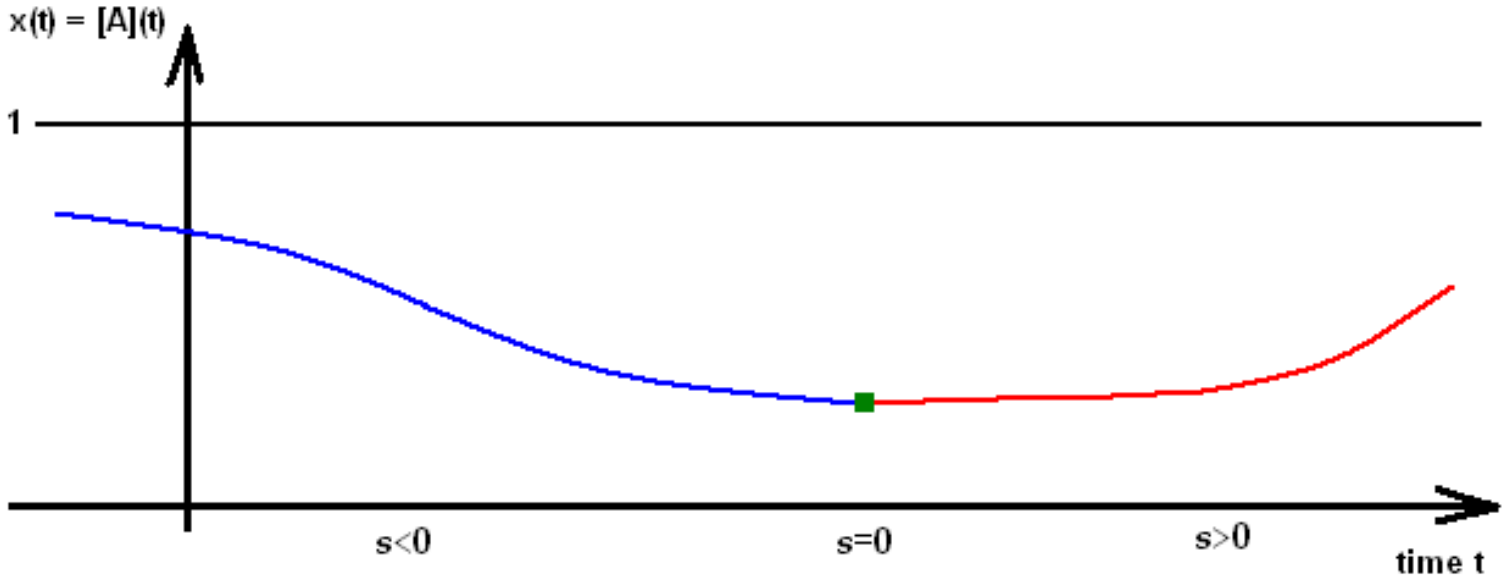


With continuous time, equilibrium is stable if dx/dt is a decreasing function at it, and unstbale if it is increasing.

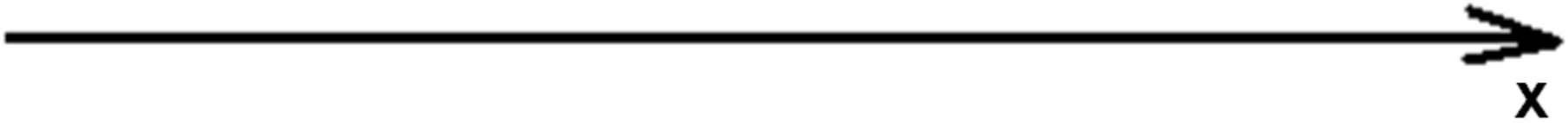
To complete qualitative investigation of the model, we need to understand transitions between its qualitatively different modes of dynamics. Here, there are three such modes: s < 0, s = 0, and s > 0.



When, for example, a negative s starts increasing very slowly, the rate of decline of allele A frequency will diminish, until everything freezes at s = 0, after which the frequencies will start growing slowly.

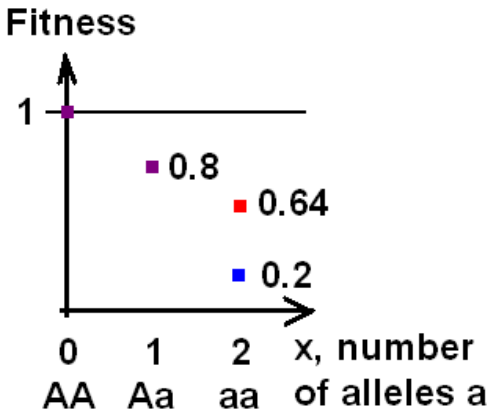


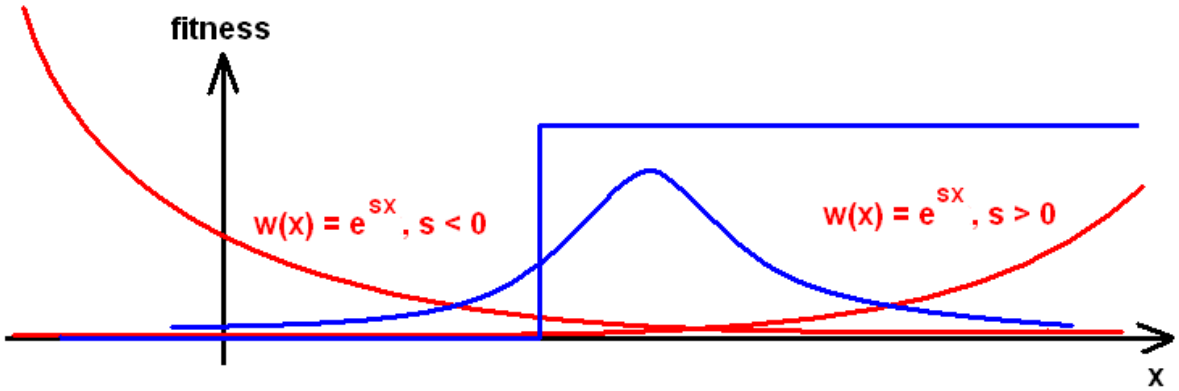
7) Action of selection on a quantitative trait



x can be a discrete (number of vertebrae, number of nucleotides G and C within a short sequence, ...) or a continuous (body weight) variable.

In this case, selection is described by fitness landscape w(x). Selection acts on all variable factors that contribute to x independently if each of them always causes the same increase or decline of fitness. If x is discrete, this means that incrementing x by 1 always leads to the same relative change in fitness: $w(x) = (1-s)^x$. With continuous x, this means that $w(x) = e^{sx}$. In all other cases, selection acting on x is epistatic.





If $w(2) = w_{aa} = 0.64$, selection acts against maternal and paternal a independently (intermediate dominance). If $w(2) = w_{aa} = 0.2$, negative effects of these two a's reinforce each other (epistasis).

Red curves show independent selection, and blue curves show two important modes of epistatic selection on a continuous quantitative trait x.

The impact of selection on the mean value of the trait is called selection differential.

$$D = M[\tilde{p}] - M[p]$$

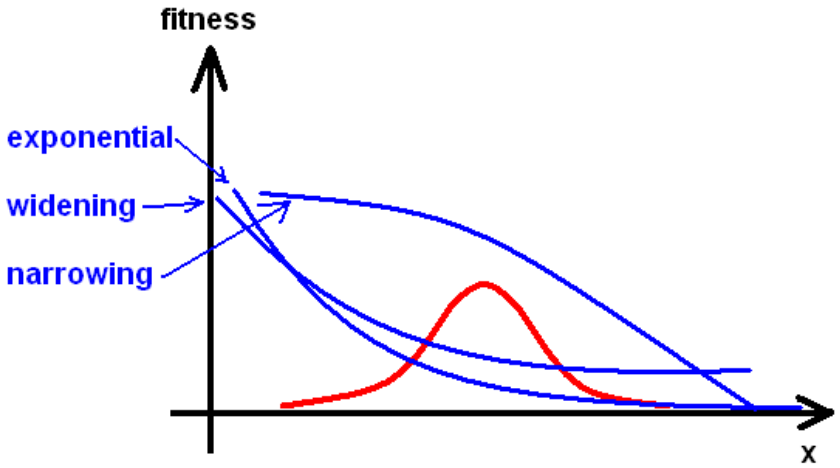
The impact of selection on the variance of the trait, does not have a common name, surprisingly.

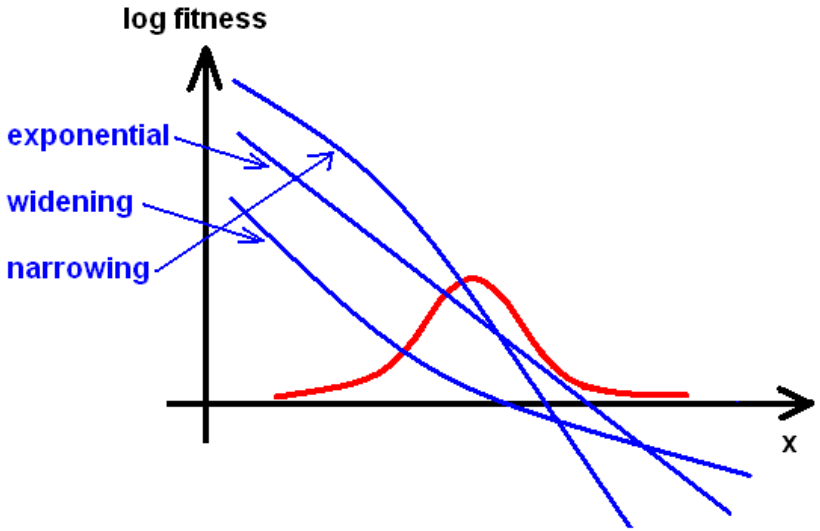
$$R = V[\tilde{p}]/V[p]$$

Why D is a difference, but R is a ratio?

It is easy to understand how selection affects the mean value of the trait. In particular, under directional selection, if w(x) increases (decreases), D > 0 (D < 0).

The impact of selection on the variance of the trait is less intuitive, but very important. If p(x) is Gaussian, exponential selection does not change the variance. If relative fitness decreases faster than exponentially, variance declines (narrowing selection), and if it decreases slower than exponentially, variance increases (widening selection).





Naturally, stabilizing selection is always narrowing, and disruptive selection is always widening. Almost any selection eventually becomes narrowing, due to survival of the fittest.

Truncation is the most efficient form of selection on a quantitative trait

We have already introduced four characteristics of selection. Two of them, genetic load and variance of relative fitness, are applicable to selection acting on any kinds of traits. In the case of quantitative traits, they are defined as follows:

genetic load $L = 1-W/w_{max}$, **where** $W=\int_{-\infty}^{\infty} p(x)w(x)dx$ **is the mean population fitness**

and

variance or relative fitness $V[w(x)/W]=\int_{-\infty}^{\infty}[p(x)(w(x)/W-1)^2dx]$

Two other characteristics,

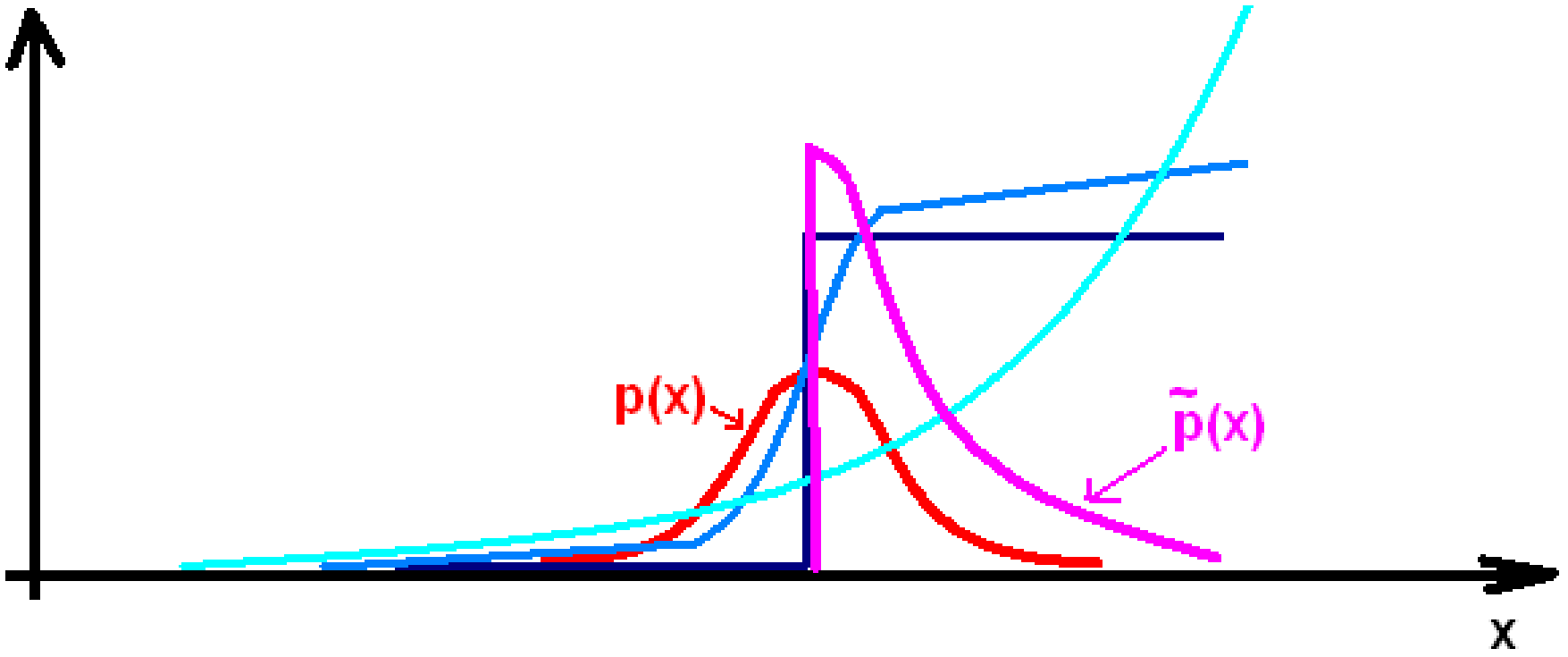
selection differential $D=M[\tilde{p}]-M[p]$

and

the impact of selection on the trait variance $R=V[\tilde{p}]/V[p]$

are applicable only to selection acting on a quantitative trait. In addition to the relationship between the genetic load and the variance of relative fitness, considered before, the relationship between the genetic load and selection differential is also very important, and makes it possible to claim that truncation is the most efficient form of selection.

Indeed, it is easy to prove the following theorem: truncation selection leads to the minimal value of the genetic load among all possible modes of selection that result in a particular selection differential.



For example, in order to shift the mean value of the quantitative trait by ~0.8 of its standard deviation, truncation selection has to impose the genetic load of 0.5 (cull 50% of individuals). Any other mode of selection needs to impose a larger genetic load in order to achieve the same result.

Robertson-Price theorem

Selection differential is equal to covariance between the trait and relative fitness.

Proof. Covariance between trait x and relative fitness w(x)/W in a population with density p(x) is (mean relative fitness is 1):

$$Cov(x, w(x)/W) = \int_{x_{min}}^{x_{max}} (x - \int_{x_{min}}^{x_{max}} xp(x)dx)(w(x)/W - 1)p(x)dx =$$

$$\int_{x_{min}}^{x_{max}} xw(x)p(x)dx/W - \int_{x_{min}}^{x_{max}} w(x)p(x)dx \int_{x_{min}}^{x_{max}} xp(x)dx)/W -$$

$$\int_{x_{min}}^{x_{max}} xp(x)dx + \int_{x_{min}}^{x_{max}} p(x)dx \int_{x_{min}}^{x_{max}} xp(x)dx$$

Here, the first term is the mean after selection (because p'(x) = w(x)p(x)/W), the second term is mean before selection (because W in the numerator and in the denominator cancel each other), and 3rd and the 4th terms cancel each other, because the integral of density is 1. Thus, this expression is D! Indeed, the more a trait covaries with fitness, the more its mean must be affected by selection!!!

What if our trait coincides with fitness (and, thus, IS fitness)? In this case w(x) = x and $W=\int_{x_{min}}^{x_{max}} xp(x)dx$ **. Thus,** $Cov(x,w(x)/W)=\int_{x_{min}}^{x_{max}}(x-\int_{x_{min}}^{x_{max}} xp(x)dx)(x/\int_{x_{min}}^{x_{max}} xp(x)dx-1)p(x)dx=$

$$\int_{x_{min}}^{x_{max}} x^2 p(x)dx/\int_{x_{min}}^{x_{max}} xp(x)dx-\int_{x_{min}}^{x_{max}} xp(x)dx\int_{x_{min}}^{x_{max}} xp(x)dx)/\int_{x_{min}}^{x_{max}} xp(x)dx-$$

$$\int_{x_{min}}^{x_{max}} xp(x)dx+\int_{x_{min}}^{x_{max}} p(x)dx\int_{x_{min}}^{x_{max}} xp(x)dx=$$

$$\int_{x_{min}}^{x_{max}} x^2 p(x)dx/\int_{x_{min}}^{x_{max}} xp(x)dx-\int_{x_{min}}^{x_{max}} xp(x)dx)$$

Which is selective differential of x, D[x]. And selective differential (= increment!) of relative fitness is D[x]/ $\int_{x_{min}}^{x_{max}} xp(x)dx$ **. When the last formula is divided over** $\int_{x_{min}}^{x_{max}} xp(x)dx$ **,**

we obtain variance of relative fitness. In other words, FFT is a special case of Robertson-Price theorem, when our trait is fitness itself.

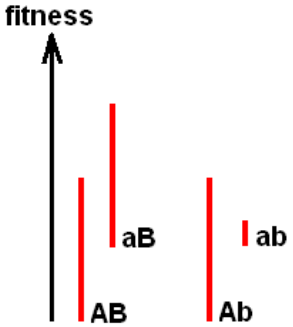
8) Independent vs. epistatic selection

Suppose now that two (or more) traits (loci) are variable within the population. Selection can act on them independently. In the simplest case of two loci, A and B, with alleles (trait states) A and a, and B and b, respectively, this means that the fitness a genotype is the product of "contributions" from different loci:

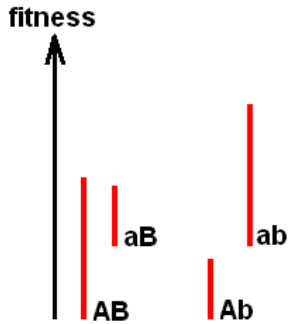
$$w_{AB} = w_A \times w_B,$$
$$w_{Ab} = w_A \times w_b,$$
$$w_{aB} = w_a \times w_B,$$
$$w_{ab} = w_a \times w_b,$$

where w_{AB}, w_{Ab}, w_{aB}, w_{ab} are genotype fitnesses, and w_A, w_a, w_B, w_b are allele contributions.

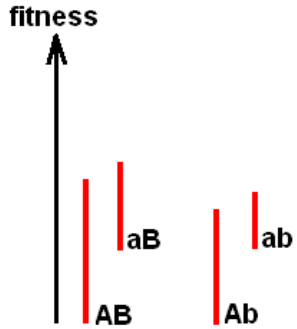
However, selection can also act on different loci non-independently, or epistatically. The following modes of epistasis are particularly important:



Incompatibility epistasis: alleles A and B are OK separately, but bad together. For example, $w_{AB} = w_{Ab} = w_{aB} = 1$ but $w_{ab} = 0.2$.



Sign epistasis: B is better than b in the presence of A, but b is better than B in the presence of a, for example $w_{AB} = 1$, $w_{Ab} = 0.5$, $w_{aB} = 0.5$, $w_{ab} = 1.0$.



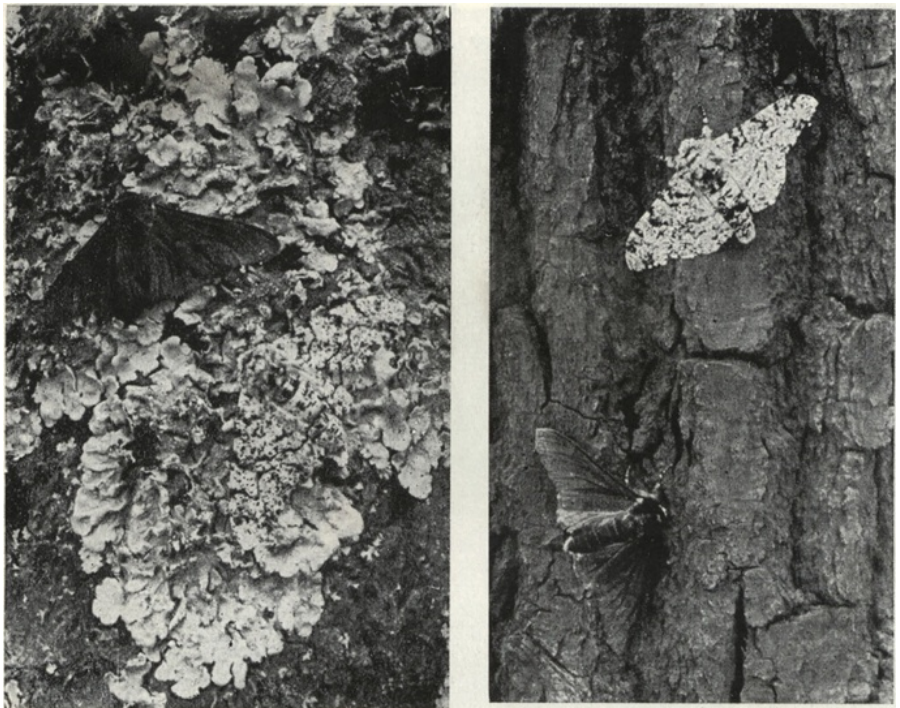
Non-epistatic selection, for comparison, for example $w_{AB} = 1$, $w_{Ab} = 0.8$, $w_{aB} = 0.6$, $w_{ab} = 0.48$.

9) Measuring natural selection

Direct problem of dynamical theory: we know what forces affect our object, how will it change?

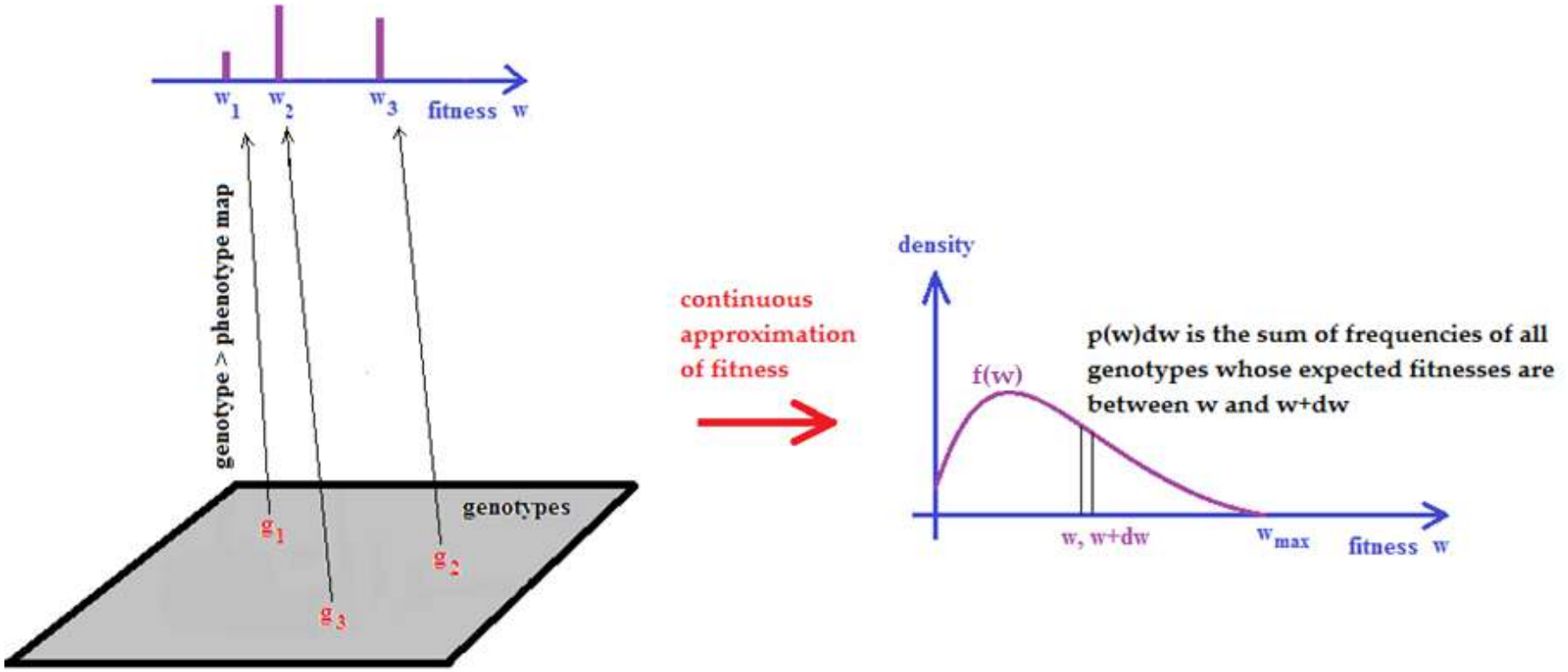
Inverse problem of dynamical theory: we know ow our object changes, what forces are affecting it?

Measuring fitnesses directly is very difficult (it is essentially impossible to measure fitness of a multicellular organism with an error less than 1-3%), and the results obtained in the laboratory cannot be applied to wild populations. Thus, indirect methods based on inverse theory are crucial.



Direct measurements of individual selection

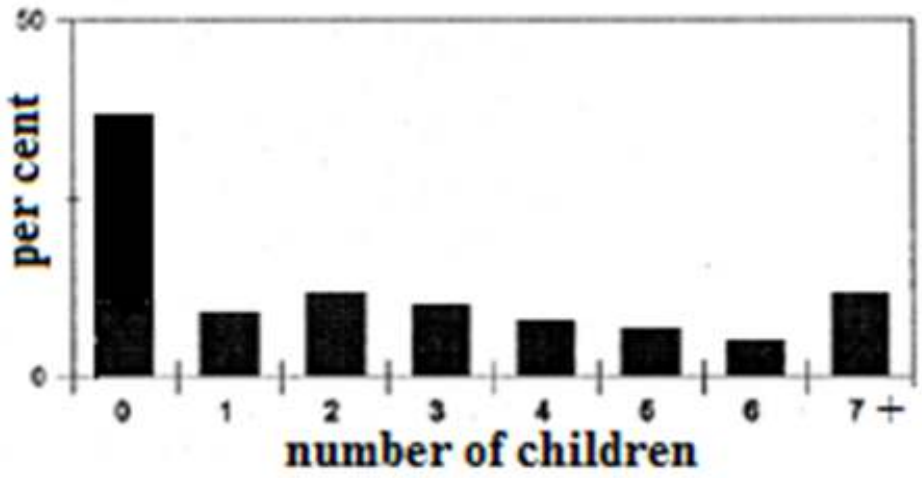
Individual selection is completely described by the expected fitnesses (the word expected may be omitted) of all the present genotypes, together with their frequencies. This information could be condensed into the density of fitnesses of genotypes, f(w), which provides an overall description of individual selection.

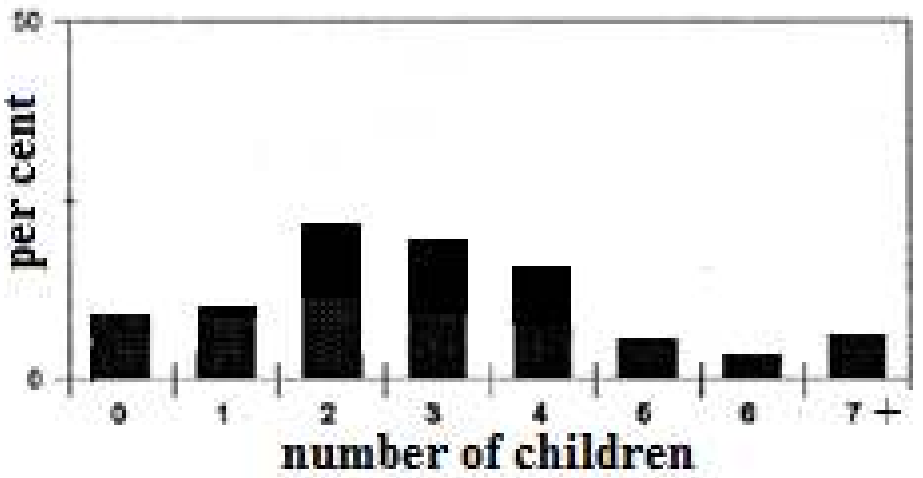


Schematic representation of the relationship between frequencies and fitnesses of genotypes and the density of fitness f(w). Within a population, each genotype g_i is characterized by its expected fitness w_i and its frequency f_i (shown by heights of purple bars). If there are many genotypes with different values of w_i, we can treat fitness of a genotype as a continuousvariable.

However, expected fitnesses of genotypes and/or alleles within a population are very hard to assay with precision sufficient for detecting all biologically important selection. There is no way we can show directly that the expected fitness of one allele is 1.001 of the expected fitness of the alternative allele.

Alternatively, we can study the density of fitness of individuals g(v). Clearly, Var[g(v)] > Var[(f(w)], so that data on fitnesses of individuals overestimate selection on genotypes.





Data on life-time fecundity of Russian women born in 1905 (taking into account that ~30% of counceptions result in a spontaneous abortion)/

A likely density of the expected fitnesses of genotypes f(w), in the same population - in, fact, its variance may be even less.

Thus, indirect estimates of selection are also very important.

Indirect detection of negative selection

This is a relatively easy task - because negative selection is very common. Negative selection affects evolving sequences in two ways:

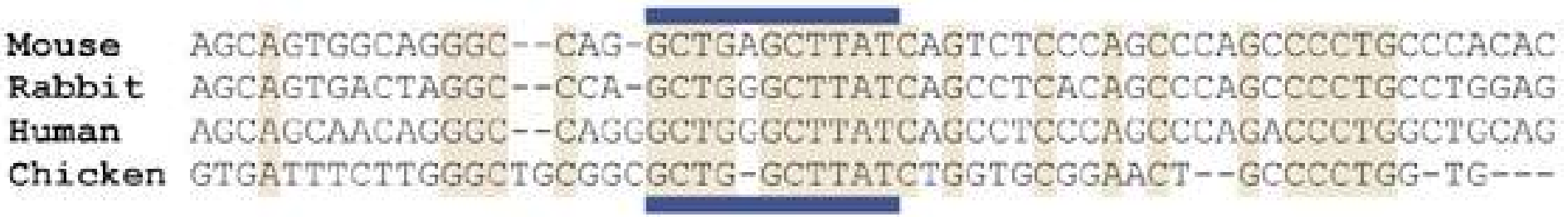
1) it reduces the probability of fixation of a mutation with s < 0

2) it reduces the time until elimination of a mutation with s < 0

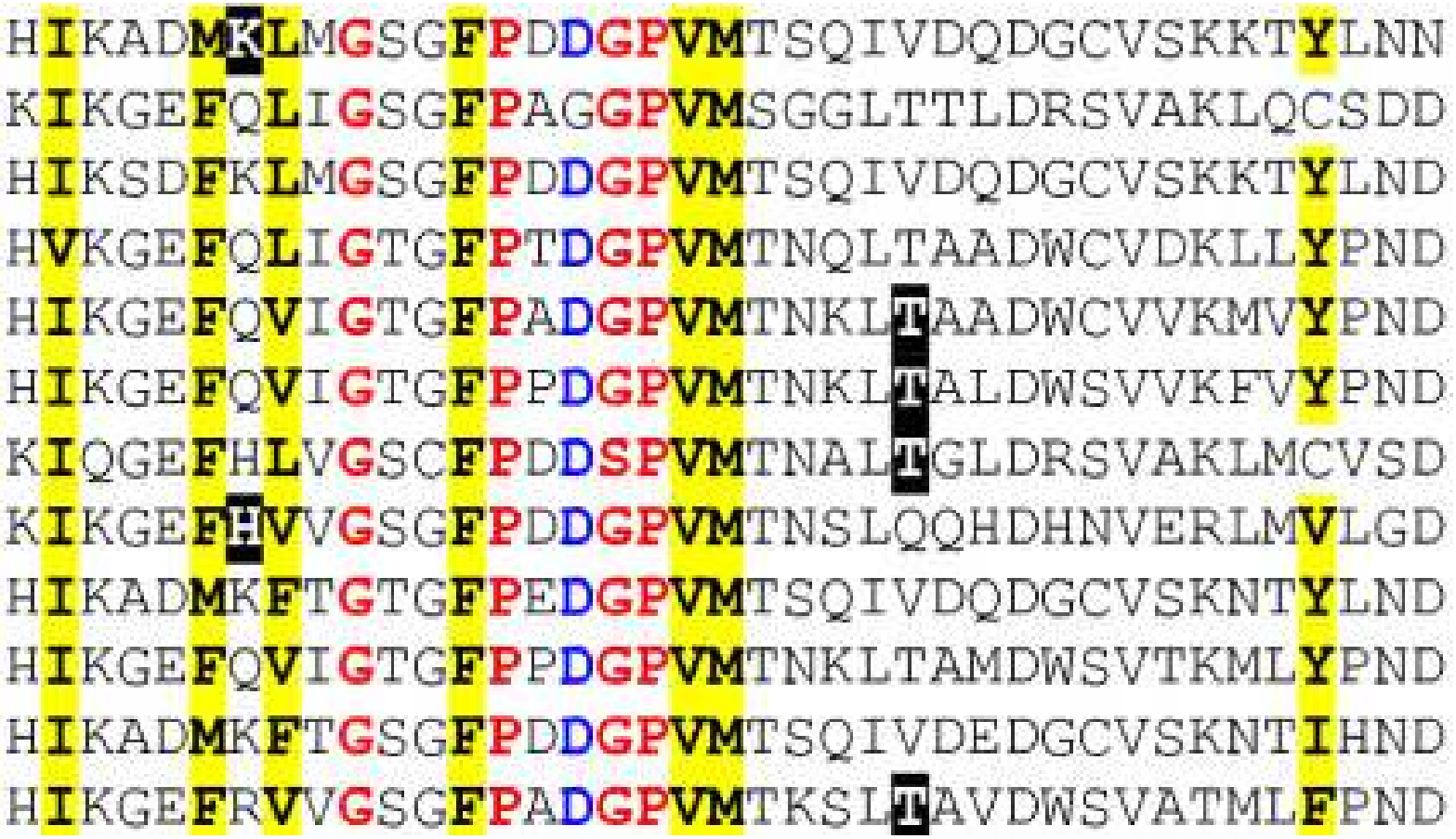
As a result, negative selection leaves two kinds of footprints:

1) reduced rate of evolution and the level of within-population variation

Reduced relative to what? - to the rate of evolution at selectively neutral sites. According to the fundamental theorem of neutral evolution, neutral sites evolve at the mutation rate (this is intuitively obvious). Practically, negative selection is detected by comparing the amount of interspecies divergence or within-population polymorphism to that at plausibly neutral sequence sites.



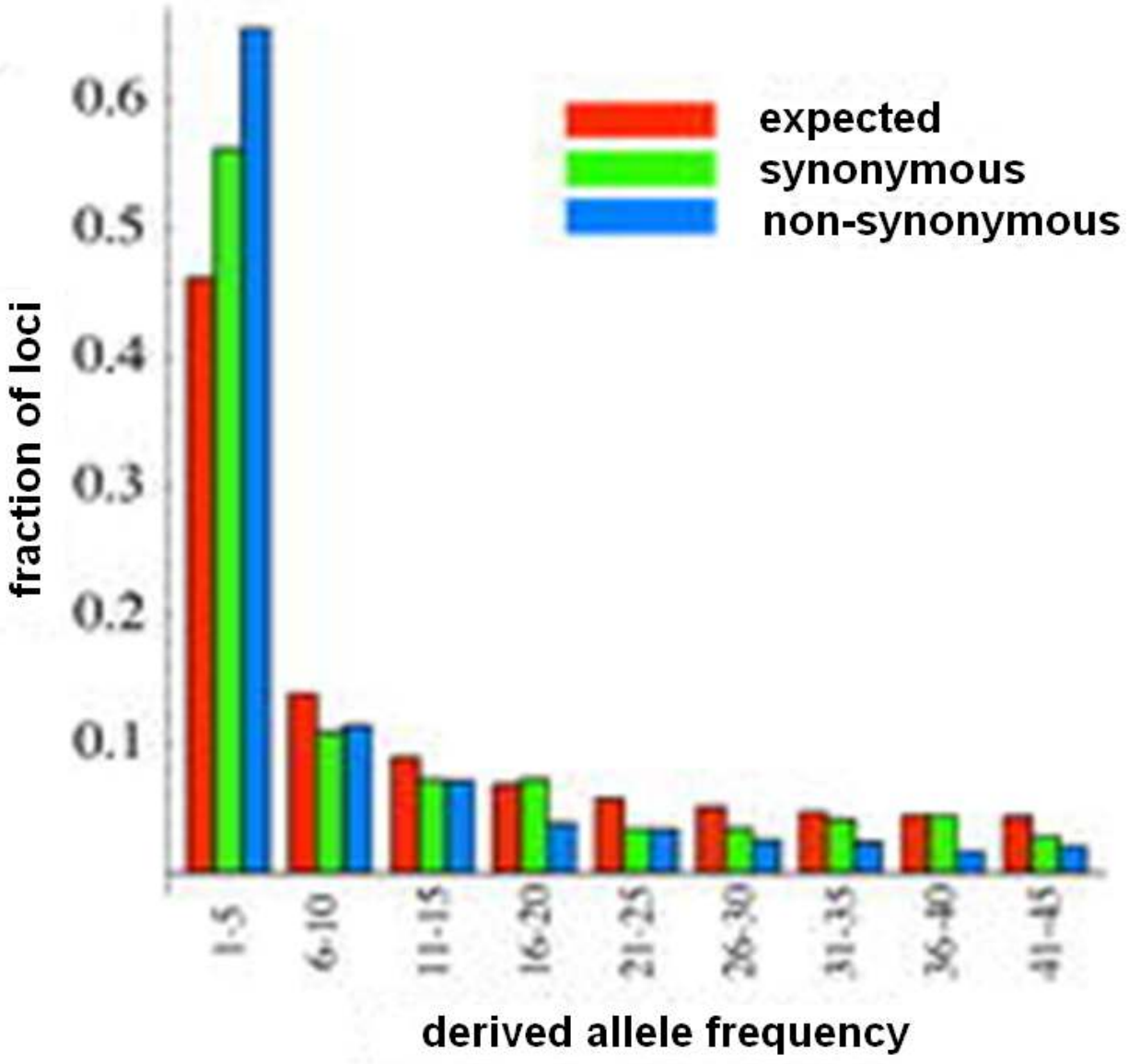
Alignment of orthologous regulatory regions of 4 mammals. A transcription factor-binding site with low divergence is marked by blue. If the alignment includes only a few sequences, we can only detect substantial segments with reduced divergence rates (never call them mutation rates!) - for example, using Hidden Markov Model technique.



A typical segment of an alignment of orthologous proteins from different species. Here the number of sequences makes it possible to detect negative selection even at individual sites.

Data on within-population variation usually allow us only to detect negative selection in wide classes of sites, for example to show that non-synonymous coding sites are under stronger selection than synonymous sites. However, with high H making inferences about individual sites may become possible. We badly need 100 genotypes of *Ciona savignyi*.

2) An excess of rare alleles



Distribution of allele (nucleotide) frequencies in *Arabidopsis thaliana*. *PLoS Biology* 3, 1289-1299, 2005.

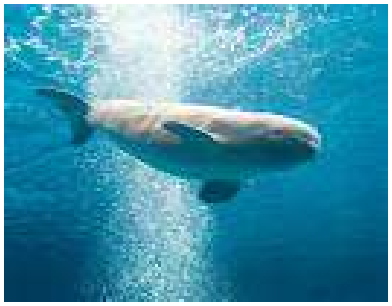
At non-synonymous sites an excess of rare alleles, relative to the neutral expectation, is higher. Of course, here we cannot make inferences about individual sites.

However, we can make inferences about the strength of negative selection - because only alleles with small s are observed as rare polymorphisms.

In contrast, reduced rate of evolution tells us very little about the strength of selection: s = -0.001 is enough to stop evolution.

Detection of positive selection

This is a difficult and important problem - because positive selection is rare, relatively to negative selection (this was proposed in 1935 by Ivan Schmalhausen) and because positive selection is the only driving force of adaptive evolution.


Positive selection affects evolving sequences in two ways:

1) it increases the probability of fixation of a mutation with s > 0

2) it reduces the time until fixation of a mutation with s > 0

Footprint of positive selection looks rather differently depending on its age.

1) Positive selection accomplished a long time ago - interspecies comparisons

In contrast to negative selection, positive selection accelerates evolution (not the rate of evolution!). Thus, it makes sites or segments to evolve faster than neutrally. As a result, we can detect positive selection only from comparing relatively close species, such that the number of accepted substitutions between them per neutral site, K_{neu}, is ~1-3. Ancient actions of positive selection, that occurred more than 1/m generations ago (m is the per nucleotide mutation rate) could never be detected.

One can try to directly estimate the strength of selection from the rate of changes of the genotype frequencies:

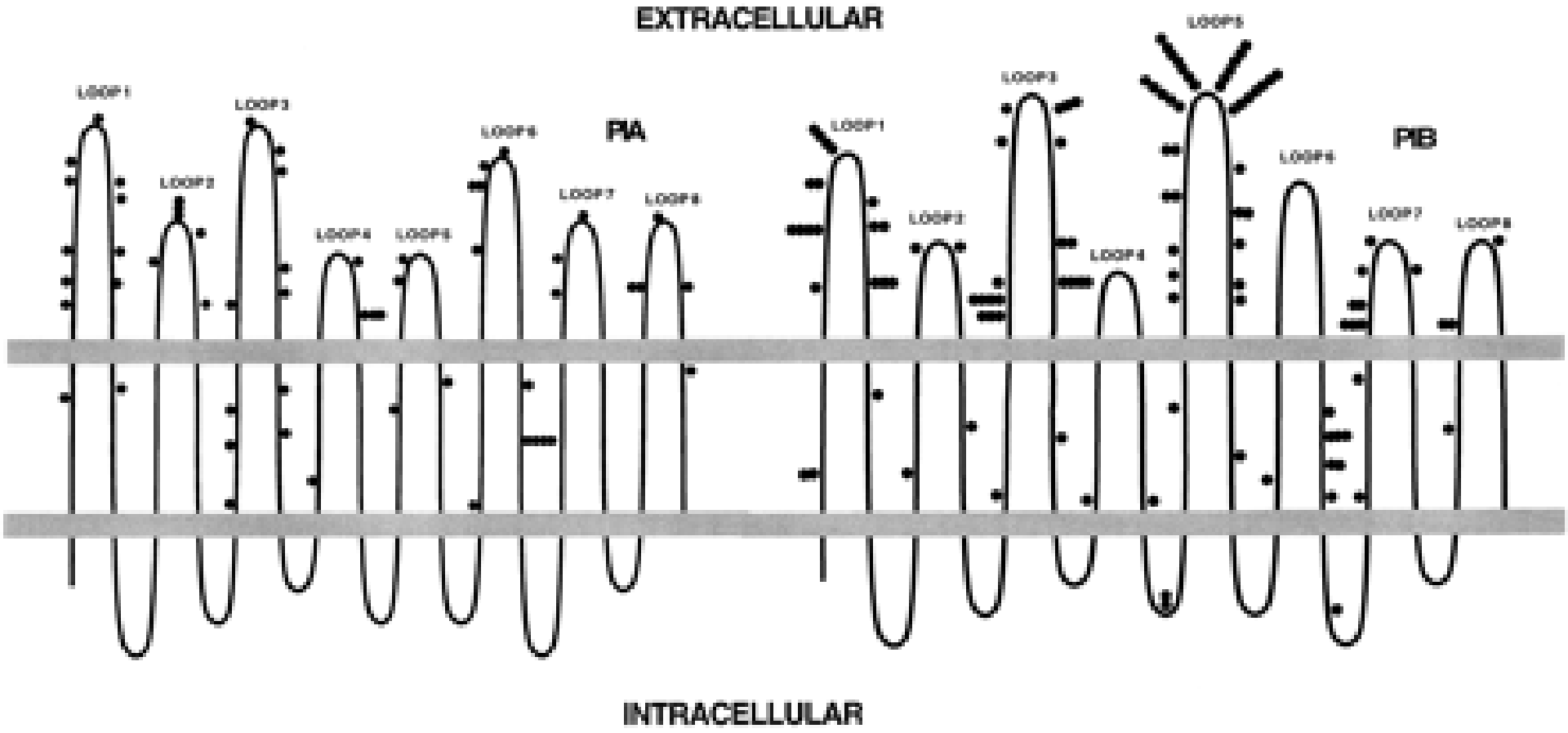
$$s = \frac{dx/dt}{x(1-x)}$$

If the frequency of allele A changed from 50% to 51% in one generation, its selective advantage must be 0.04.

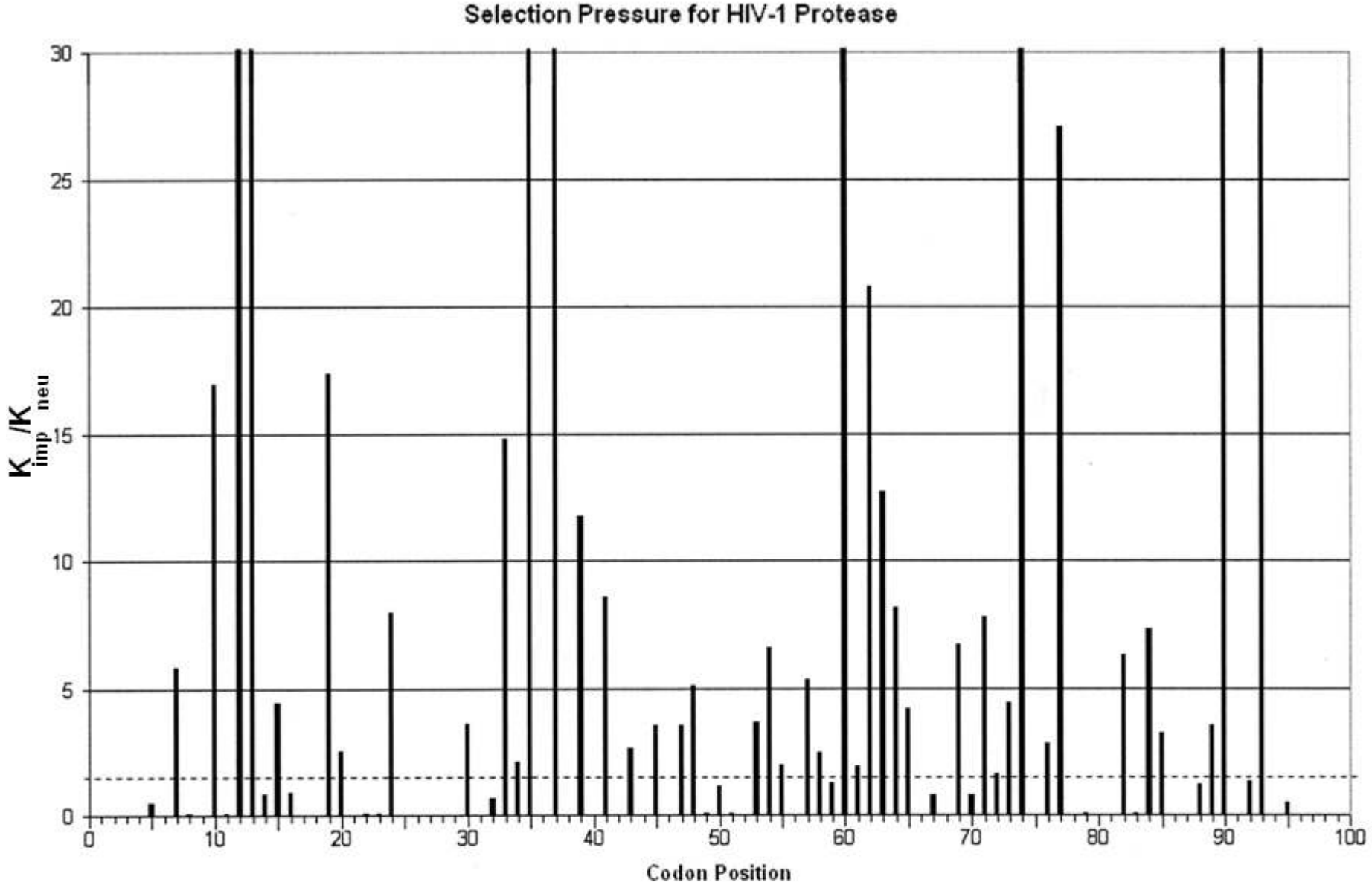
Factors that can affect dynamics of within-population variation are mutation, selection, mode of reproduction, population structure, and drift. Microevolution is due to their joint action.

However, before dealing with this joint action, we will first need to consider variation and each of the 5 factors acting separately.

So, if we have a large number of close enough sequences, even individual sites where K > K_{neu} (K_{neu} is measured for sites that are probably under no selection) can be detected. This approach works well for pathogens, with multiple moderately different strains.



Distribution of amino acid replacements along the *Neisseria gonorrhoeae* transmembrane porin sequence. Each dot represents one replacement. Obviously, sequence segments exposed outside the cell evolve much faster, probably due to positive selection. *Molecular Biology and Evolution* 17, 423-436, 2000.



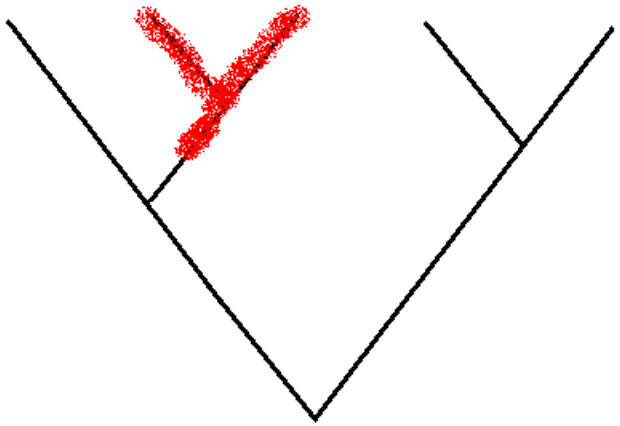
Positive selection in HIV-1 protease, detected on samples from 40,000 patients. For each codon site, the ratio of the rate of the most common allele replacement over the neutral rate is shown (*Journal of Virology* 78, 3722-3732, 2004).

However, there are two problems with this approach:

1) Positive selection can act only within one clade, with negative selection acting at the same site in the rest of the phylogeny. Then, overall K will be low at the site.

2) There may be not enough species to measure K for individual sites. If so, all probably important sites are treated together, and their average per site number of changes, K_{imp}, is calculated. Trouble is, sites under positive are generally scattered between numerous sites under negative selection, leading to $K_{imp} < K_{neu}$. Only very rarely, there are long enough segments with a majority of sites under positive selection.

Positive selection acting in one clade, on a sparse phylogenetic tree.



Sophisticated statistical methods can be used to analyze such data - but, in my opinion, they reliably detect positive selection only if a substantial fraction of sites to $K_{imp} > K_{neu}$. at least within a large clade - and this is generally very rare. Most of "important" sites are, most of the time, under negative, and not positive selection.

A clever idea of MacDonald and Kreitman can offer some help. They realized that the condition $K_{imp} > K_{neu}$ (or $K_{imp}/K_{neu} > 1$) can be relaxed. If negative selection is strong, "important" sites under it will not be polymorphic in the population. Sites under positive selection also make only minimal contribution to polymorphism (because polymorphism in the course of an allele replacement is very short-lived). Thus, instead of asking for

$$K_{imp}/K_{neu} > 1$$

as a signature of positive selection it is enough to ask for

$$K_{imp}/K_{neu} > H_{imp}/H_{neu}$$

H_{imp}/H_{neu} can be as low as 0.2-0.3 (due to a large fraction of sites under negative selection among "important" sites), so this is a much less stringent condition.

One problem with this approach is that slightly deleterious variants with $-s \sim 1/N_e$ can segregate within the population, but are only rarely fixed, and thus inflate H_{imp}/H_{neu}. A possible way of dealing with this problem is to ignore rare variants.

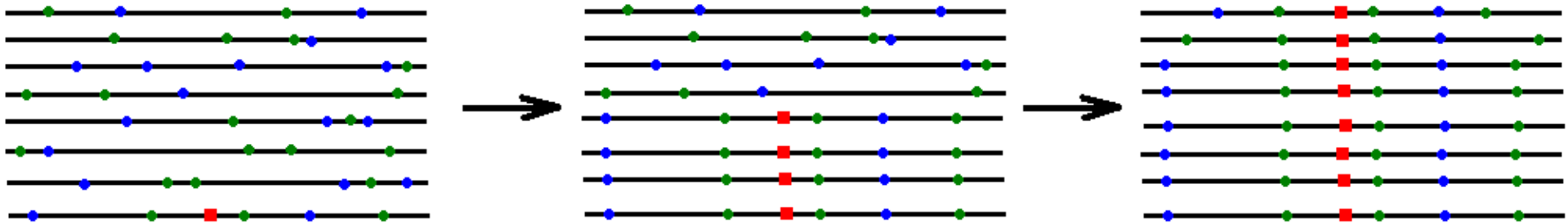
Some applications of MacDonald-Kreitman test to *Drosophila* species suggest that as many as 50% of allele replacements in fly evolution were driven by positive selection, because

$$K_{imp}/K_{neu} = 2H_{imp}/H_{neu}$$

In contrast, in mammals $K_{imp}/K_{neu} < H_{imp}/H_{neu}$, suggesting no positive selection. The reasons for such contrast are unclear. Anyway, MK test could never establish identities of individual sites under positive selection.

2) Positive selection accomplished recently - within-population variation

A recent allele replacement driven by positive selection produces a region of very low variation, flanked by regions with some high-frequency derived alleles. Such a scar of an allele replacement is due to an effect called hitch-hiking, and it remains visible for << $1/N_e$ generations, where N_e is the effective population size per nucleotide mutation rate.

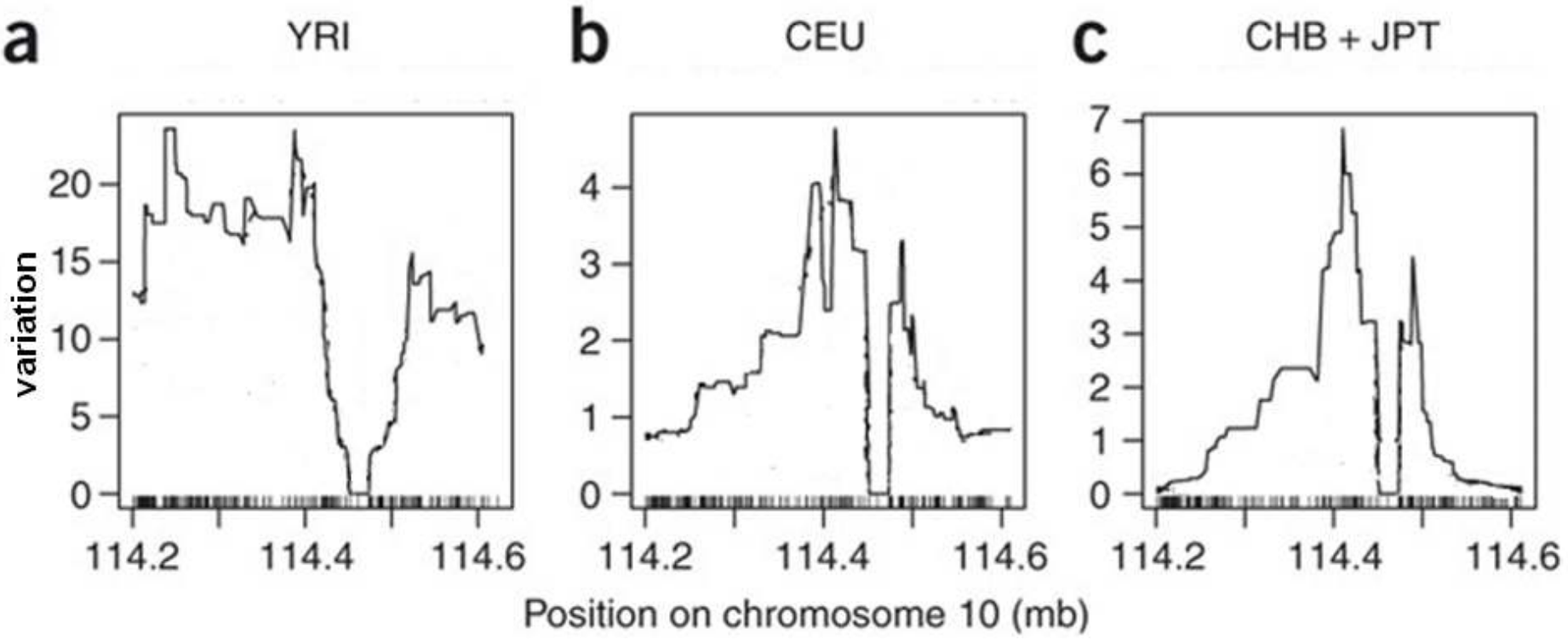


A beneficial mutation (red) in a population with many segregating neutral (green) and slightly deleterious (blue) variants.

Half-way towards fixation, the beneficial mutation carries with it the close-by variants.

Some of these variants become detached, due to crossing-over, by the time of the fixation.

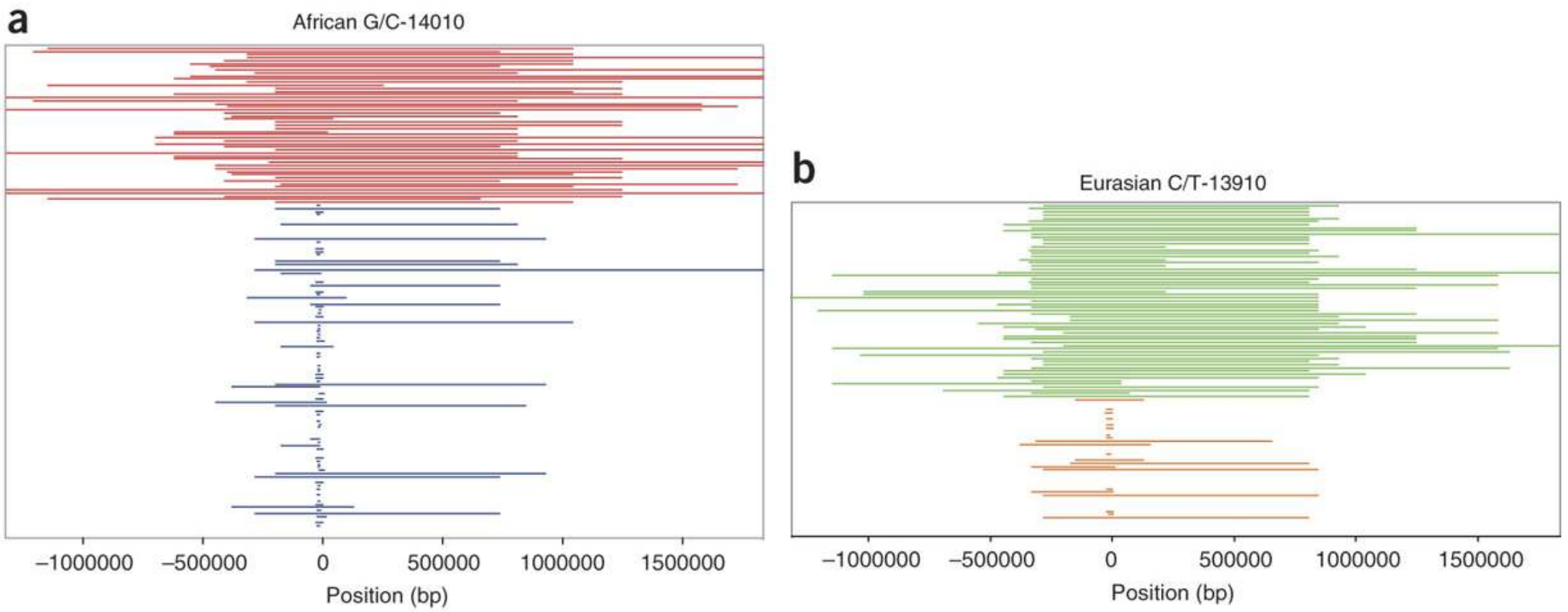
There are several definite known cases of recently accomplished selective sweeps.



Reduced levels of genetic variation around the site of recent positive selection-driven allele replacement (selective sweep) in human populations from Africa (a), Europe (b), and East Asia (c) (*Nature Genetics* 39, 218 - 225, 2007).

3) Ongoing positive selection - within-population variation

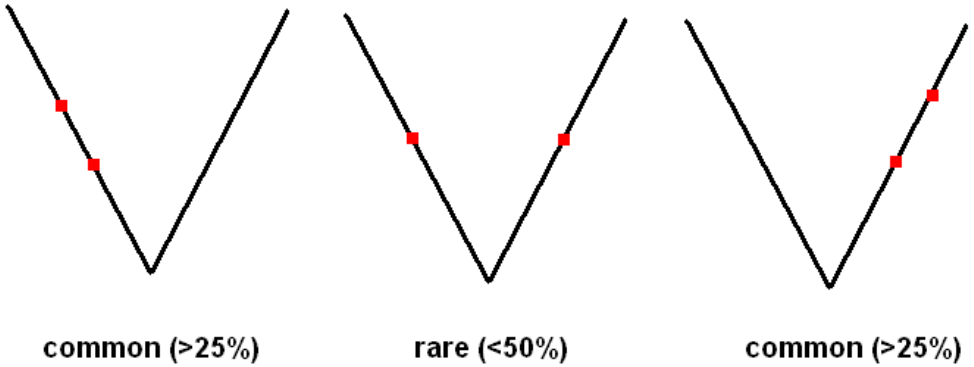
One must be lucky to study the right population at the right time. Still, there are some definite cases of ongoing allele replacements driven by strong positive selection. One of them is parallel acquisition the ability of adults to digest milk (due to persistent expression of lactase) in Africans and non-Africans. These ongoing sweeps left clear-cut signatures.



(a) Kenyan and Tanzanian C-14010 lactase-persistent (red) and non-persistent G-14010 (blue) homozygosity tracts. (b) European and Asian T-13910 lactase-persistent (green) and C-13910 non-persistent (orange) homozygosity tracts. Positions are relative to the start codon of lactase locus (*Nature Genetics* 39, 31 - 40, 2006).

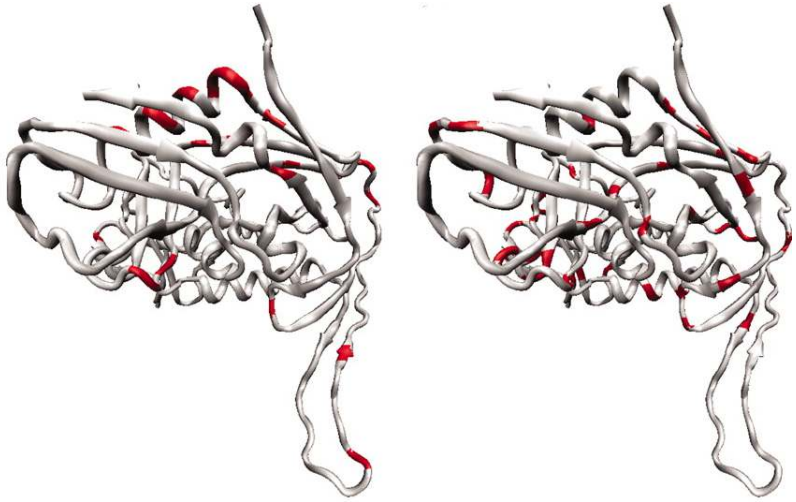
4) A different approach - detecting positive selection by bursts of substitutions

Suppose that at a codon site fitness landscape was suddenly changed. The new optimal amino acid may not be reachable from the old one by a single nucleotide substitution. Then, a clump of two or even three non-synonymous substitutions may follow. Such clumps were observed in evolution of mammals and HIV-1 (*PNAS* 103, 19396-19401, 2006).



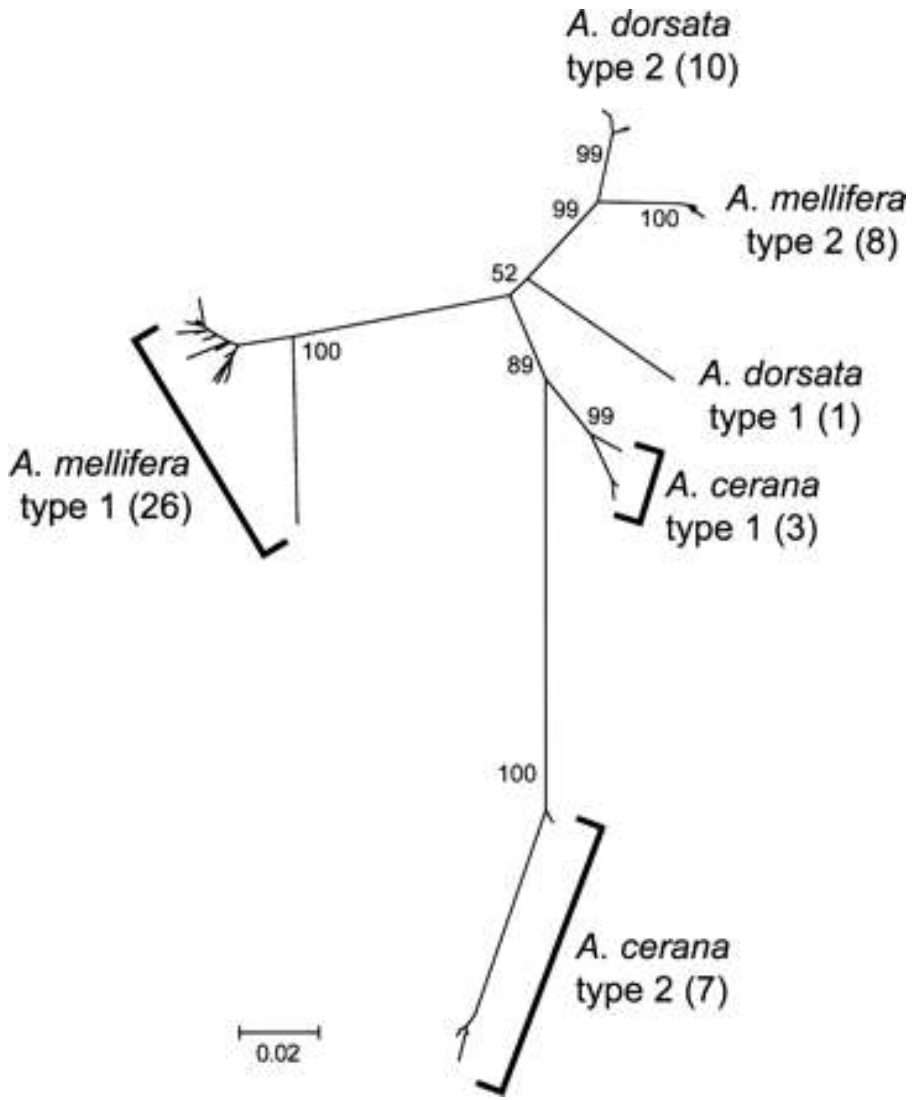
Clumping of nonsynonymous substitutions is the strongest in conservative regions of proteins, where the 1:1 situations occur only in ~20% of codons. Indeed - if an important amino acid is replaced, this must be beneficial. This approach reveals a number of slowly-evolving sites that occasionally undergo positive selection.

Amino acid sites inferred to be under positive selection in HIV-1 gp120. Left: rapidly evolving sites previously inferred to be under positive selection. Right: conservative sites with strongly clumped substitutions.



Detecting balancing selection

Balancing selection, which requires changing fitness landscapes, favors rare alleles. It prevents fixations and losses of the alleles involved, leading to durable polymorphisms.



In the extreme case this can lead to transspecies polymorphisms, persisting from the time of species divergence. This is the case for sad *csd* (complementary sex determination) locus in bees. Female must be heterozygous at this locus, and homozygotes develop into sterile males, causing strong selection against common alleles (*Genome Res.* 16, 1366-1375, 2006*).*