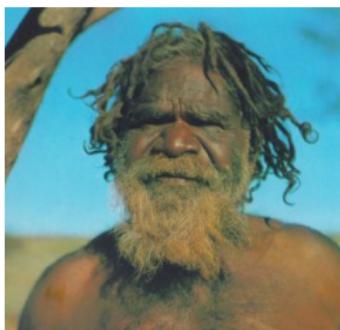


Natural selection

Populations tend to be polymorphic

Usually populations are polymorphic and the biological differences are partly heritable.



Aborigine morph of the hominin
Homo sapiens



Spaniard morph of the hominin
Homo sapiens

Evolutionary dynamics is the dynamics of polymorphic populations.

The language of population genetics

To understand evolutionary dynamics it is important to distinguish the four following fundamental population genetic concepts.

- **Phenotype.** This is any morphological, physiological or behavioral trait expressed by an individual.
- **Genotype.** Set of molecular transmitted determinants of the phenotype (it is the personal **genome** of an individual roughly the set of all the genes of an individual).¹
- **Gene.** A stable (non-recombinant) element of the genome that performs a particular function.
- **Allele.** A variant of a gene.

¹The genotype of a human depends on DNA inherited from Mom and Dad. In several organisms the genotype and thus personal genome of an individual also consists of much non-functional DNA.

Phenotype

Everything about an individual that is not its genotype is part of its phenotype.



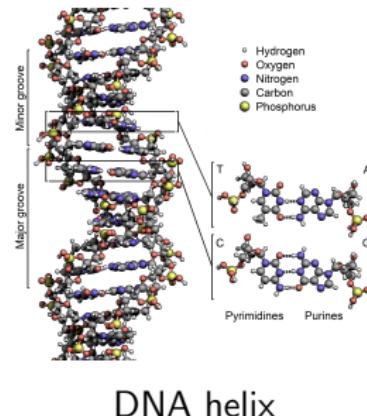
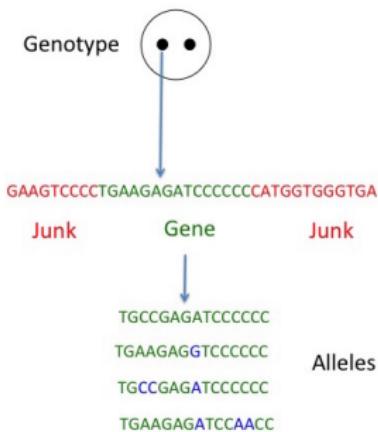
Organic phenotype and extended phenotype



Extended group phenotype

- Preferences and beliefs are phenotypes.
- Tools and technology are **extended phenotypes**.
- The words “phenotype” and “trait” will often be used interchangeably.

Genotype



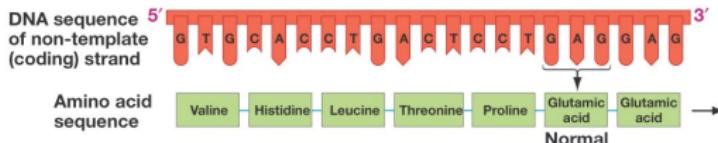
The genotype of an organism is its executive code.²

²While DNA is a double helix depending on four chemical (nucleo) bases [adenine (A), guanine (G), thymine (T), and cytosine (C)], hereditary information is determined by a single sequence owing to the complementary pairing of the bases.

From genotype to phenotype

Central dogma of biology: Genes → Proteins (intermediate phenotypes consisting of amino acids) → observed phenotype.

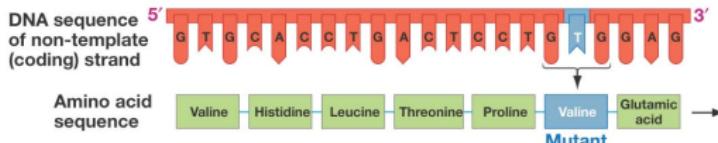
(a) DNA point mutation can lead to a different amino acid sequence.



(b) Phenotype



Normal red blood cells



Sickled red blood cells

Copyright © 2008 Pearson Benjamin Cummings. All rights reserved.

Central dogma of biology simplified: Genotype → phenotype.
The relation between genotype and phenotype is called the genotype-phenotype map.

Human genotype (or individual genome)

- 6.4×10^9 nucleobases long. About 3.2×10^9 inherited from Mom and a bit less from Dad, since we inherit mitochondria (organelle that supply chemical energy) from Mom.
- Approximately 1.5% is directly functionally relevant, that is consists of genes coding for proteins. We have about 20 000 genes. Each gene has potentially many alleles in the population. Genes are between 500 to 100000 nucleobases long.
- Hence, much of the genome is presumed to be non-functional and sometimes called **junk DNA**.
- Soon it will be possible to sequence the genome at the personal level.

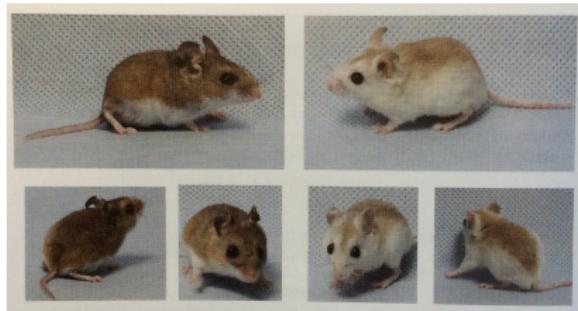
Two broad categories of phenotypes

- ① **Multifactorial phenotypes:** or **quantitative** traits depend on many genes and crucially on **environmental effects**.
- ② **Mendelian phenotypes:** is a trait that is controlled by a single gene. A single allele at a single gene can then cause a disease that is inherited like sickle-cell anemia. Mendelian traits tend to be inherited faithfully.

We will start to understand selection on mendelian traits as this is simpler and also allows to understand selection on multifactorial phenotypes.

Mendelian trait

The melanocortine-1 gene (MC1R) codes for the production of melanin, which partly determines skin colour in mice and humans.



Coat colour affects viability as it affects camouflage. An allele (say *AGTGGTAAA*) determines for light-coloured mice (more likely to be seen on beaches) and another (say *AGTGGTA***T***A*) determines for dark-coloured mice (living in the woods).

Mendelian trait

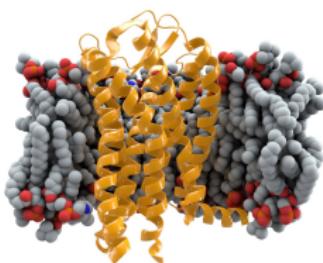
The endothelial PAS gene (EPAS1) codes for a protein that is induced as oxygen levels fall (hypoxia).



Tibetans carry a high proportion of an EPAS1 gene allele that improves oxygen transport, and which improves survival at high altitude.

Mendelian trait

Many diseases or resistance to pathogens are Mendelian traits. For instance, the C-C chemokine receptor type 5 gene (CCR5 gene) codes for a protein on the surface of white blood cells (immune system cell) that acts as a receptor for signaling molecules.



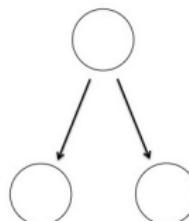
CCRA protein attached to a cell membrane.

Certain humans carry an allele (CCR5- Δ 32) protecting them against several HIV viral strains, and thus do not develop AIDS.

Selection on Mendelian traits

We will start to study evolution on Mendelian traits as this is the simplest.

Hence, we make the simplest genetic assumptions and further assume haploid reproduction.



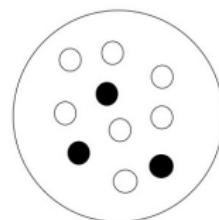
Asexual reproduction

- But this actually does not result in a loss of understanding and allows to apprehend evolution of multifactorial phenotypes.
- We now turn to introduce the main quantities to study evolutionary dynamics.

Two-allele system

Suppose we have two alleles in the population:

- ① An allele that we call A (say black type in the figure).
- ② An allele that we call B.



Think of A and B as being two different types of individuals.

Different alleles may result in different fitnesses:

w_A fitness of A

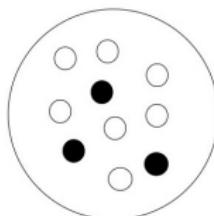
w_B fitness of B

Two-allele system

In a population where allele A and B segregate we have

- ① n_A individuals with allele A
(3 in the figure).

- ② n_B individuals with allele B
(7 in the figure).



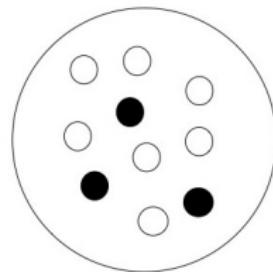
And the total number of individuals in the population is

$$n = n_A + n_B$$

Allele frequency

The **frequency** p of allele A in the population is

$$p = \frac{n_A}{n}$$

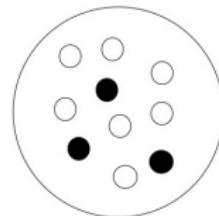


The frequency of A in the population of the figure is $p = 3/10$, since 3 individuals carry A (black dots) among 10 individuals.

Variance in allele frequency among individuals

The **variance in the frequency** of allele A among individuals in the population is

$$v = p(1 - p)$$



Since the frequency of A is $p = 3/10$, the variance is
 $v = (3/10)[1 - (3/10)] = 0.21$.

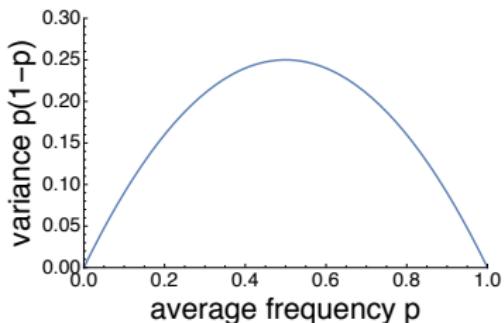
Intuitively, the variance³ among individuals can be thought of as the likelihood of sampling one individual that carries the A allele (probability p) and another one that does not carry it (probability $1 - p$).

³Technically, the variance is a quadratic measure of the “spread” of allele frequency in each individual in the population around the mean.

Variance in allele frequency among individuals

The variance in the frequency of allele A as a function of the average frequency in the population is

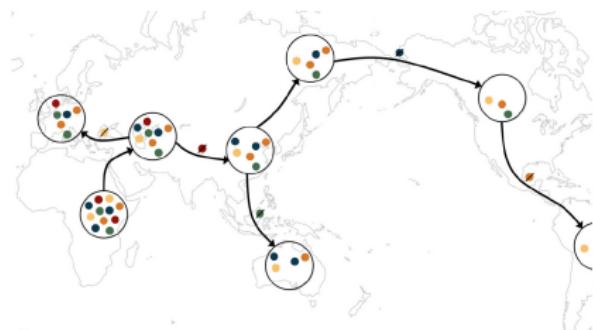
$$v = p(1 - p)$$



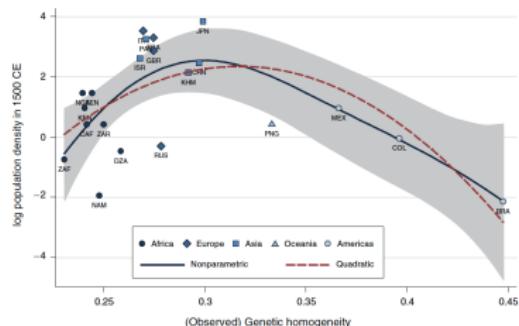
- If $p = 0$, everybody carries B and there is no variability.
- If $p = 1$, everybody carries A and there is no variability.

The variance is highest at intermediate frequency and gives a measure of the heterogeneity in the population.

Total (**neutral**) allelic variance among individuals in a population correlates worldwide with development



Serial founder model of human evolution.



Logarithm of population size in 1500 treated as a proxy of economic development.

(Ashraf and Galor, 2013, "American Economic Review")

Evolution

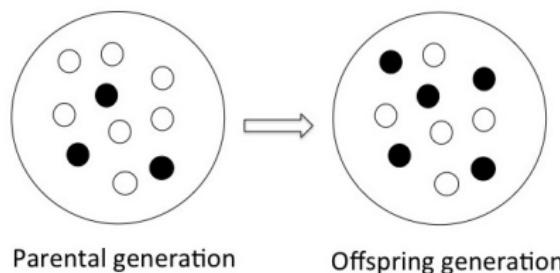
- Technically organic evolution is defined as the **change of allele frequency** in a population over time.
- **Question:** what are the forces resulting in allele frequency change?
- We will start with natural selection and mainly focus on Mendelian traits as it allows to assimilate many concepts.

To understand natural selection we start by focusing on the change in allele frequency over one demographic time period.

Change in allele frequency

Given frequency p in some parental generation, the change in the frequency of A over one demographic time period is

$$\Delta p = p' - p$$



- ① If $\Delta p > 0$, the frequency of A increases.
- ② If $\Delta p < 0$, the frequency of A decreases.
- ③ If $\Delta p = 0$, there is no change.

Allele frequency in the descendant generation

The frequency of A in the offspring generation for haploid reproduction under faithful transmission is

$$p' = \frac{\text{A number in offspring generation}}{\text{total number in offspring generation}} = \frac{n'_A}{n'}$$

The two unknowns here are n'_A and n' . The number of A in the descendant generation is

$$n'_A = w_A n_A = w_A np$$

which is the number of A individuals in the parental generation ($n_A = np$) times their fitness (w_A).

Allele frequency in the descendant generation

We can then also write the number of individuals in the descendant generation as

$$n' = \bar{w}n$$

where

$$\bar{w} = w_A p + w_B (1 - p)$$

is the average fitness in the population.

This gives population size in the offspring generation as a function of population size in the parental generation in a polymorphic population.⁴

⁴The equation $n' = \bar{w}n$ is similar to that we used when studying only population dynamic, with the only difference that in the latter case we had $\bar{w} = w$ since the fitness of everybody was the same.

Change in average allele frequency

The change in the frequency of allele A over one demographic time point in the population is

$$\Delta p = \frac{p(1-p)s}{\bar{w}}$$

which depends on two quantities:

- ① The selection coefficient on allele A

$$s = w_A - w_B$$

- ② The variance $p(1-p)$ in A frequency among individuals in the population.

Change in average allele frequency

The change in the frequency of A is

$$\Delta p = \frac{p(1-p)s}{\bar{w}} \quad \text{with} \quad s = w_A - w_B$$

- If $s > 0$, then natural selection is said to favor allele A since $\Delta p > 0$ and the allele increases in frequency in the population.⁵
- If $s < 0$, then allele A is counter-selected since $\Delta p < 0$ and it decreases in frequency in the population.

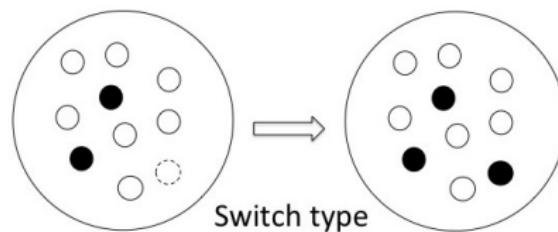
⁵We always have $p(1-p) \geq 0$ and $\bar{w} > 0$. Hence, s alone determines the direction of evolution. \bar{w} can be thought as a normalizer guaranteeing that the frequencies add up to one.

Natural selection as “gains from switching”

The selection coefficient

$$s = w_A - w_B$$

can be thought as the fitness gains to an individual from switching to expressing the A allele instead of allele B.



The selection coefficient s can thus be thought off as the net **marginal fitness benefit** to an individual from switching to A.

Replicator selection

The change in allele frequency

$$\Delta p = \frac{p(1-p)s}{\bar{w}} \quad \text{with} \quad s = w_A - w_B$$

is a formal encapsulation of Darwin's idea of natural selection. It describes how selection operate on any *replicator* defined as a unit that exhibits the following three features:

- ① Replication.
- ② Variation.
- ③ Faithful heredity.

This applies to any form of inheritance. It underlies the phrase "survival of the fittest".

Evolutionary equilibria

No change of allele frequency occurs when

$$\Delta p = 0 \implies p(1-p)s = 0$$

There are two type of evolutionary equilibria (or steady-states) with equilibrium frequency denoted p^* .

- ① **Boundary equilibria** (or trivial equilibria), namely $p^* = 0$ (extinction of A) and $p^* = 1$ (fixation of A).
- ② **Interior equilibria**, satisfying $s = 0$, where the fitness of each allele must be the same.

Both these equilibria can be stable or unstable depending on the selection regime.

Different regimes of selection

The change in allele frequency under natural selection encapsulates different modes of selection and in order to better understand the nature of evolutionary equilibria we now investigate different regimes of selection:

- ① Constant selection.
- ② Density-dependent selection.
- ③ Frequency-dependent selection.

Constant selection

We now assume that the selection coefficient is constant over time and that the fitness of B is one ($w_B = 1$), which is our point of reference. We can then set $w_A = 1 + s$ where s gives directly the fitness advantage. The change in frequency of A is then

$$\Delta p = \frac{p(1-p)s}{\underbrace{1+ps}_{\bar{w}}}$$

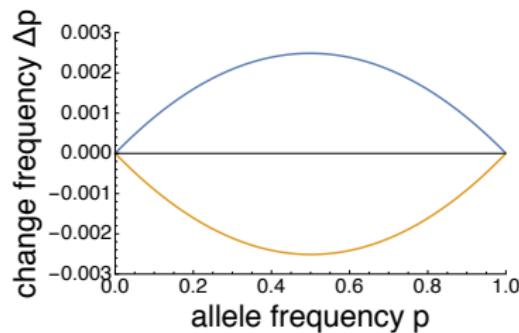
Since s is constant, the direction of allele frequency change is constant at all allele frequencies:

- If $s > 0$, then A is favored ($\Delta p > 0$) for all p .
- If $s < 0$, then A is counter-selected ($\Delta p < 0$) for all p .

Constant selection: change in allele frequency

The change in A frequency is

$$\Delta p = \frac{p(1-p)s}{1+ps}$$

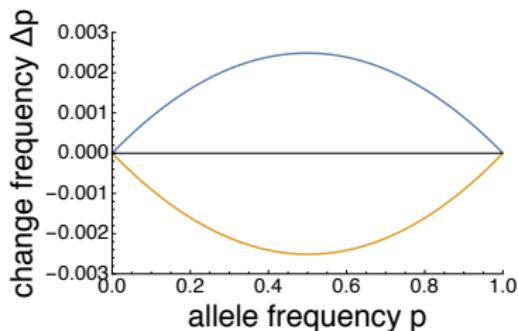


Parameters values $s = 0.01$ (blue) and $s = -0.01$ (yellow).

The sign of the change in A frequency is constant at all allele frequencies. A selection coefficient of 1% (i.e., $s = 0.01$ or $s = -0.01$) is a typical selection coefficient in natural population, for humans and other species.

Constant selection: equilibrium points

- No change when $\Delta p = 0$.
- This obtains when $p = 0$ and $p = 1$.
- There are only two equilibria in the model.



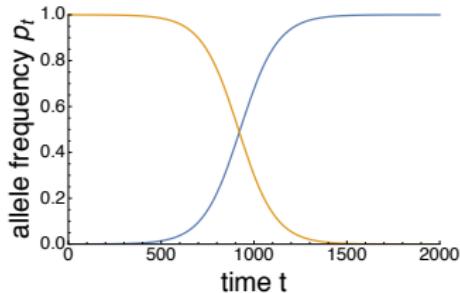
Parameters values $s = 0.01$ (blue) and $s = -0.01$ (yellow).

Hence, **invasion implies substitution**: the frequency of A not only increases when rare ("invasion") but goes to one so that A replaces B ("substitution").

Constant selection: time dynamics

One can solve explicitly for the frequency p_t ⁶ of A at time t .

$$p_t = \frac{1}{1 + \frac{1-p_0}{p_0} \left(\frac{1}{1+s}\right)^t}$$



where p_0 is the initial frequency.

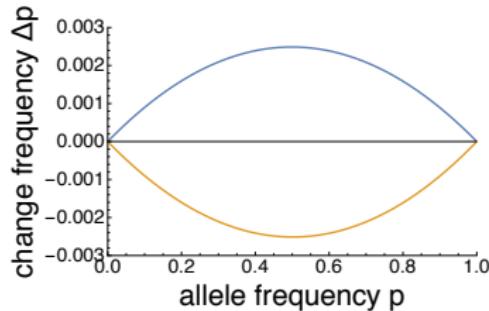
Parameters: $p_0 = 10^{-4}$ and $s = 0.01$ (blue), $p_0 = 1 - 10^{-4}$ and $s = -0.01$ (yellow).

Frequency change follows a sigmoid curve. Hence, invasion is initially geometrically fast. In fact, as long p_t remains close to zero (initial phase of invasion), we have $p_t \approx (1 + s)^t p_0$.

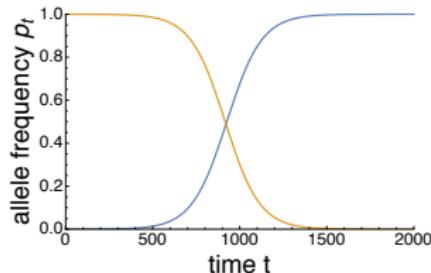
⁶Recall from Chapter 1 that $\Delta p = p' - p$ is equivalent to $\Delta p_t = p_{t+1} - p_t$.



Constant selection: state space and time dynamics



Parameters values $s = 0.01$ (blue) and $s = -0.01$ (yellow).



Parameters: $p_0 = 10^{-4}$ and $s = 0.01$ (blue), $p_0 = 1 - 10^{-4}$ and $s = -0.01$ (yellow).

The **state space** graph (on the right) has all the qualitative information about the time dynamics, as it tells for each frequency p what is the direction of evolutionary change.

Fitness is unlikely to be constant

- A model with constant fitness implies that there is no density-dependence.
- As we saw earlier this implies that the population either goes extinct or fill up the volume of the universe in a few days.
- We thus need that the population is regulated. Will this change the nature of natural selection?

Competition for material resources

Let us assume that the fitness of each allele follows a Beverton-Holt model of density-dependent competition

$$w_A(n) = \frac{f_A}{1 + \gamma_A n} \quad \text{and} \quad w_B(n) = \frac{f_B}{1 + \gamma_B n}$$

where n is the total number of individuals in the population that must satisfy the equation for population dynamics

$$n' = \underbrace{(pw_A(n) + (1 - p)w_B(n))}_\bar{w} n$$

Here, f_A and f_B are A's and B's fecundities (in the absence of density-dependence), and γ_A and γ_B are the strengths of density-dependent competition felt by the carrier of each allele.

Density-dependent selection: change in allele frequency

The change in allele frequency is given by

$$\Delta p = \frac{p(1-p)s(n)}{\bar{w}}$$

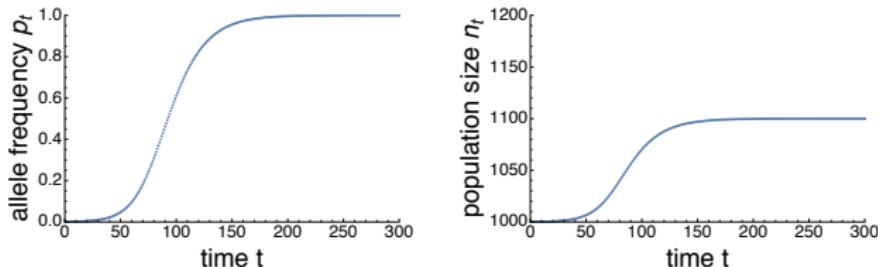
where the selection coefficient is

$$s(n) = w_A(n) - w_B(n) = \frac{f_A}{1 + \gamma_A n} - \frac{f_B}{1 + \gamma_B n}$$

- Since the coefficient of selection depends on n , the model exhibits **density-dependent selection**.
- How will this change the nature of natural selection?

Density-dependent selection: time dynamics

The frequency p_t of A and total population size n_t at time t when A is introduced as a single copy at the carrying capacity of B are



Parameter values are $f_A = 6.5$, $\gamma_A = 0.005$, $f_B = 2$, $\gamma_B = 0.001$, $\bar{p}_0 = 1/n_0$, and $n_0 = (f_B - 1)/\gamma_B = 1000$ (carrying capacity of the B allele).

Hence, **invasion implies substitution** and by invading A raises the carrying capacity.

Density-dependent selection: increase in carrying-capacity

Allele A invades when rare and goes to fixation when

$$w_A(n_B^*) > w_B(n_B^*) = 1 \quad \text{and} \quad n_B^* = \underbrace{\frac{f_B - 1}{\gamma_B}}_{B's \text{ carrying capacity}}$$

- Hence, A **invades when its carrying capacity is larger** than that of B, otherwise its fitness could not be larger. This also implies that its fitness at all frequencies is larger than that of B.
- The nature of natural selection under density-dependent selection is thus not changed relative to constant selection.

Competition here increases population “productivity”, as measured by population size. This is the closest analogue in biology of the idea that competitive systems (e.g., markets) increase efficiency.

Social interactions

Individuals interact with each other and this affects survival and reproduction. All organisms thus “play games” in the game theory sense.



Social interactions result in frequency-dependent selection as we will next investigate.

The Hawk-Dove game

Suppose there are two pure actions (or strategies or phenotypes): play "Hawk" or "Dove". Individuals play the following pairwise game

		Player 2	
		"Hawk"	"Dove"
Player 1	"Hawk"	$\frac{V}{2} - C$	V
	"Dove"	0	$\frac{V}{2}$

V is the value of a resource for which individuals compete, while C is the cost of an aggressive interaction. The entries in the matrix are the payoffs to the individuals.

A key assumption of the game is that $C > V/2$.

Hawk-Dove game: fitness

Suppose an individual carrying allele A is programmed to play “Hawk”, and a B individual to play “Dove”. Alleles thus code for **pure strategies** and assuming random pairwise interactions (in a large population) we can write the fitnesses as

$$w_A(p) = 1 + p \left(\frac{V}{2} - C \right) + (1 - p)V$$
$$w_B(p) = 1 + p \times 0 + (1 - p)\frac{V}{2}$$

	"Hawk"	"Dove"
"Hawk"	$\frac{V}{2} - C$	V
"Dove"	0	$\frac{V}{2}$

Here, the fitnesses are averages over the interactions: a Hawk (Dove) interacts with a Hawk with probability p and with a Dove with probability $1 - p$. Each individual has baseline fitness of one. The “1” in front of the fitnesses stems for a baseline fitness that each individual has.

Hawk-Dove game: change in allele frequency

The change in the frequency of allele A is given by

$$\Delta p = \frac{p(1-p)s(p)}{\bar{w}}$$

where the selection coefficient is

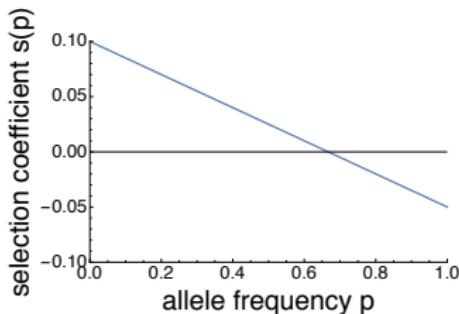
$$s(p) = w_A(p) - w_B(p) = \frac{V}{2} - Cp$$

Since $s(p)$ depends on the frequency p , the model exhibits **frequency-dependent selection**.

Hawk-Dove game: frequency-dependent selection

The selection coefficient as a function of allele frequency is

$$s(p) = \frac{V}{2} - Cp$$

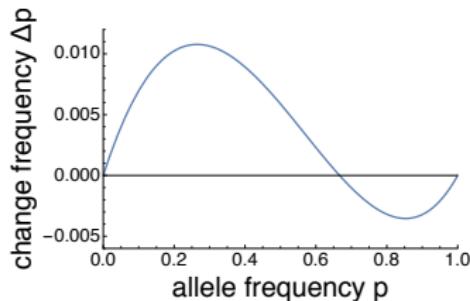


Parameters: $V = 0.2$ and $C = 0.15$.

- Near $p \approx 0$, the selection coefficient is positive, which implies $\Delta p > 0$ (Hawks are favored).
- Near $p \approx 1$, the selection coefficient is negative, which implies $\Delta p < 0$ (Doves are favored).

Hawk-Dove game: equilibrium points

The change in allele frequency $\Delta p = p(1 - p)s(p)/\bar{w}$ as a function of frequency is



Parameters: $V = 0.2$ and $C = 0.15$.

There are thus two type of equilibria here:

- Boundary equilibria $p^* = 0$ (extinction of A) and $p^* = 1$ (fixation of A), which are **unstable**.
- Interior equilibrium satisfying $s(p^*) = 0$, which is **stable**.

Hawk-Dove game: stable interior equilibrium

For the Hawk-Dove game the interior equilibrium \bar{p}^* satisfies

$$s(p^*) = \frac{V}{2} - Cp^* = 0$$

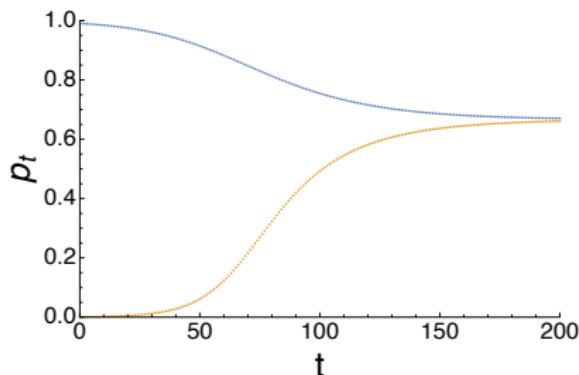
whereby solving for p^* gives

$$p^* = \frac{V}{2C}$$

Whenever $V/2 > 0$ and $V/2 - C < 0$ we have a **protected polymorphism**. There is thus a mixture between two morphs, Hawks and Doves.

Hawk-Dove game: time dynamics

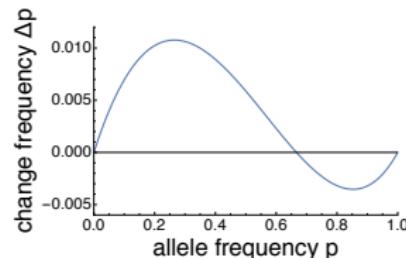
The frequency p_t of A as a function of time t is



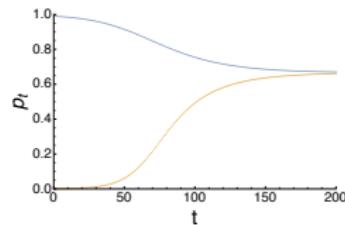
Parameters: $V = 0.2$ and $C = 0.15$, and $p_0 = 0.001$ (blue) and $p_0 = 0.99$ (blue).

Hence, the interior equilibrium $p^* = V/(2C) = 0.66$ is reached from different initial conditions (it is a globally attracting point).

Hawk-Dove game: state space and time dynamics



Parameters: $V = 0.2$ and $C = 0.15$.



Parameters: $V = 0.2$ and $C = 0.15$, and $p_0 = 0.001$ (blue) and $p_0 = 0.99$ (blue).

The **state space** graph (on the right) has all the qualitative information about the time dynamics, as it tells for each frequency p what is the direction of evolutionary change.

Summary

- Natural selection changes the frequency of a particular allele proportionally to the current variance in that frequency and the individual fitness benefits gained from expressing it.
- Constant selection.
- Density-dependent selection.
- Frequency-dependent selection: results in stable or unstable polymorphism.