

Dynamics of Genes in Populations

Gute Ideen in der Systembiologie

Presented by: Arie Zackay & Thomas Spießer

Table of Contents

- Introduction and Background
- Natural Selection
 - Charles Darwin
 - Types of Selection
 - Hardy-Weinberg Equilibrium
- Random Genetic Drift
 - Genetic Bottleneck
 - Effective Population Size
- Gene Substitution
 - Fixation Probability
 - Fixation Time
- The Neo-Darwinian Theory and the Neutral Mutation Hypothesis

Introduction and Background

- What is Population Genetics?
 - The study of genetic changes that occur in populations
- What kind of problems does one seek to answer in Population Genetics?
 - Determine how the frequency of a mutant gene will change over time under the influence of evolutionary forces
 - How genetic variability is maintained
 - Determine the probability that new mutant completely replaces existing variants of genes
- What are the odds?
 - Chance effects in molecular evolution

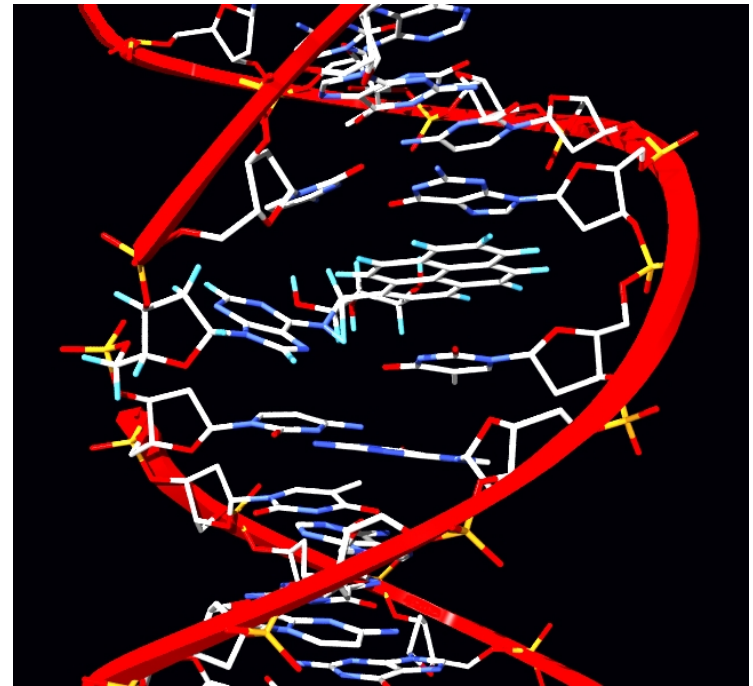
Introduction and Background

- Locus: genomic location of a gene
- Alleles: alternative forms of gene at given locus
- Allele frequency: is the fraction or percentage of loci that the allele occupies within the population
- Evolution (our case!): change in allele frequency with time
- Mutation: changes to the base pair sequence of genetic material

Introducing: Chance Effect Nr.1

THE MUTATION

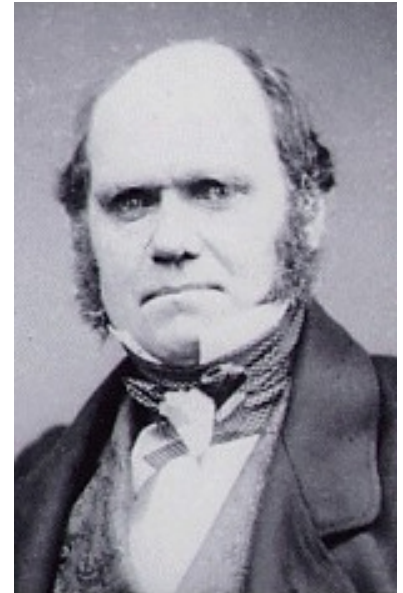
- Contributes to the genetic variability within a population
- Neutral Mutation: causes no change in protein sequence
- Fixed Mutant: its frequency increases with time and ultimately reaches 1
- For a mutant allele to increase in frequency, factors like natural selection, random genetic drift, recombination and migration must come into play!



Benzopyrene, the major mutagen in tobacco smoke, in an adduct to DNA. Produced from PDB 1JDG
http://en.wikipedia.org/wiki/Image:Pyrene_adduct.jpg

Natural Selection

- He defined natural selection as the "principle by which each slight variation [of a trait], if useful, is preserved"
- As long as there is some variation between individuals, there will be an inevitable selection of individuals with the most advantageous variations
- Introduction of the term Fitness: Ability to survive and reproduce is called *Fitness* of an individual

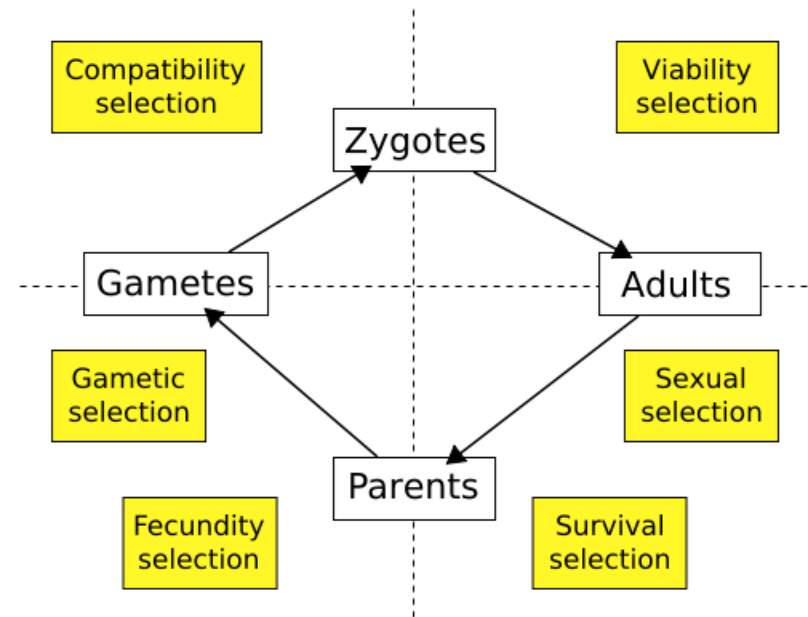


Charles Darwin

http://en.wikipedia.org/wiki/Image:Charles_Darwin.jpg

'Natural' Selection

- Defined as the differential reproduction of genetically distinct individuals within a population
- Differences among individuals such as mortality, fertility, ...
- Selection leads to change in allele frequency with time



Released under public domain, <http://en.wikipedia.org/wiki/User:Wykis>
http://en.wikipedia.org/wiki/Image:Life_cycle_of_a_sexually_reproducing_organism.svg

Types of Selection

- There are new mutants in the population, that reduce the fitness of their carriers and are therefore removed from it, the so called *negative/purifying selection*
- A mutant as fit as the best allele in the population is *selectively neutral* and its fate is determined by chance
- Should a mutation confer a selective advantage on its carrier, it is called a *positive selection*

Hardy-Weinberg Equilibrium

- Theoretical, simplifying concept of genetics
- Predicts how gene frequencies will be transmitted from generation to generation given a specific set of assumptions
 - **If** an infinitely large, random mating population is free from outside evolutionary forces (i.e. mutation, natural selection)
 - **then** on average the gene frequencies will not change over time and the frequencies in the next generation will be p^2 for the AA genotype, $2pq$ for the Aa genotype and q^2 for the aa genotype.

Hardy-Weinberg Equilibrium

Genotype:	A_1A_1	A_1A_2	A_2A_2
Fitness:	w_{11}	w_{12}	w_{22}
Frequency:	p^2	$2pq$	q^2

- Relationship between the frequency of alleles and the genotype of the population
- 2 Alleles: A_1 and A_2 of a locus
- 3 different diploid genotypes A_1A_1 , A_1A_2 , A_2A_2
- Their fitness values: w_{11} , w_{12} and w_{22}
- Frequency of A_1 in population is p and
- Frequency of A_2 in population is $q = 1 - p$
- Frequency of A_2 in next generation is q'
- Extent of change in A_2 allele frequency per generation denoted as $q = q' - q$

Modes of Selection

- In the *codominant* mode of *selection*, or *genic selection*, the two homozygotes have different fitness values, while the fitness of the heterozygote is the mean of the fitness of the two homozygous genotypes.

Genotype:	A_1A_1	A_1A_2	A_2A_2
Fitness:	1	$1+s$	$1+2s$

$$\Delta q = \frac{spq}{1 + 2spq + 2sq^2}$$

$$\frac{dq}{dt} = spq = sq(1 - q)$$

Modes of Selection

$$q_t = \frac{1}{1 + \left(\frac{1 - q_0}{q_0} \right) e^{-st}}$$

- q_0 and q_t are the frequencies of A_2 at time 0 and t , respectively
- Genic selection is a type of directional selection, because eventually one allele is eliminated

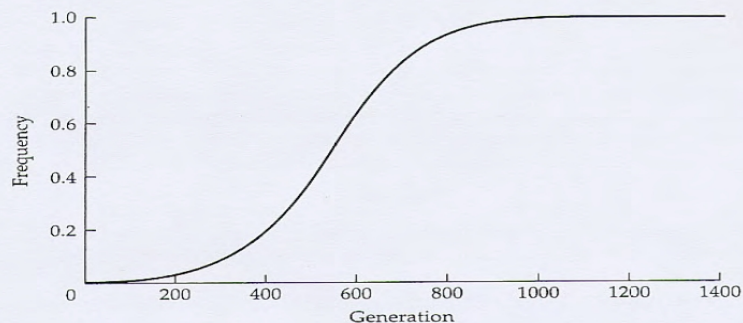


Figure 2.1 Frequency of a codominant advantageous allele with $s = 0.01$ following its appearance as a result of mutation in generation 0. From Li and Graur (1991).

Molecular Evolution by Li (1997). Chapter 2: Dynamics of genes in populations

Modes of Selection

- In the *overdominant mode of selection*, the heterozygote has the highest fitness

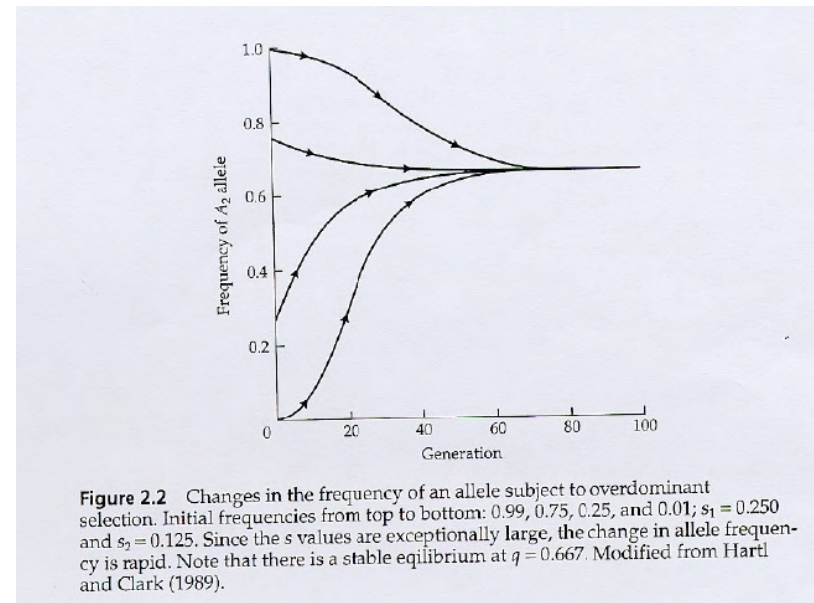
Genotype:	A_1A_1	A_1A_2	A_2A_2
Fitness:	1	$1+s_1$	$1+s_2$

- With $s_1 > 0$ and $s_1 > s_2$
- Change in allele frequencies is expressed as:

$$\Delta q = \frac{-pq(2s_1q - s_2q - s_1)}{1 + 2s_1pq + s_2q^2}$$

Modes of Selection

- Belongs to the class of *balancing or stabilizing selections*, because the population will sooner or later reach an equilibrium in which the two alleles will coexist
- Equilibrium is obtained by solving equation for $q = 0$
- When $s_2=0$ the equilibrium frequencies of both alleles will be 50%



Molecular Evolution by Li (1997). Chapter 2: Dynamics of genes in populations

Random Genetic Drift

- Natural selection is not the only factor that cause changes in allele frequency. Allele Frequency changes can also occur by chance.
- Genetic drift is random flow in allele frequency between generation. Random sampling in transmission of copies of genes to offspring
- This effect may cause allele to be more common or more rare, but may remove the allele or all others
- Drifting rate depends strongly on population size -Small populations drift more rapidly than large ones

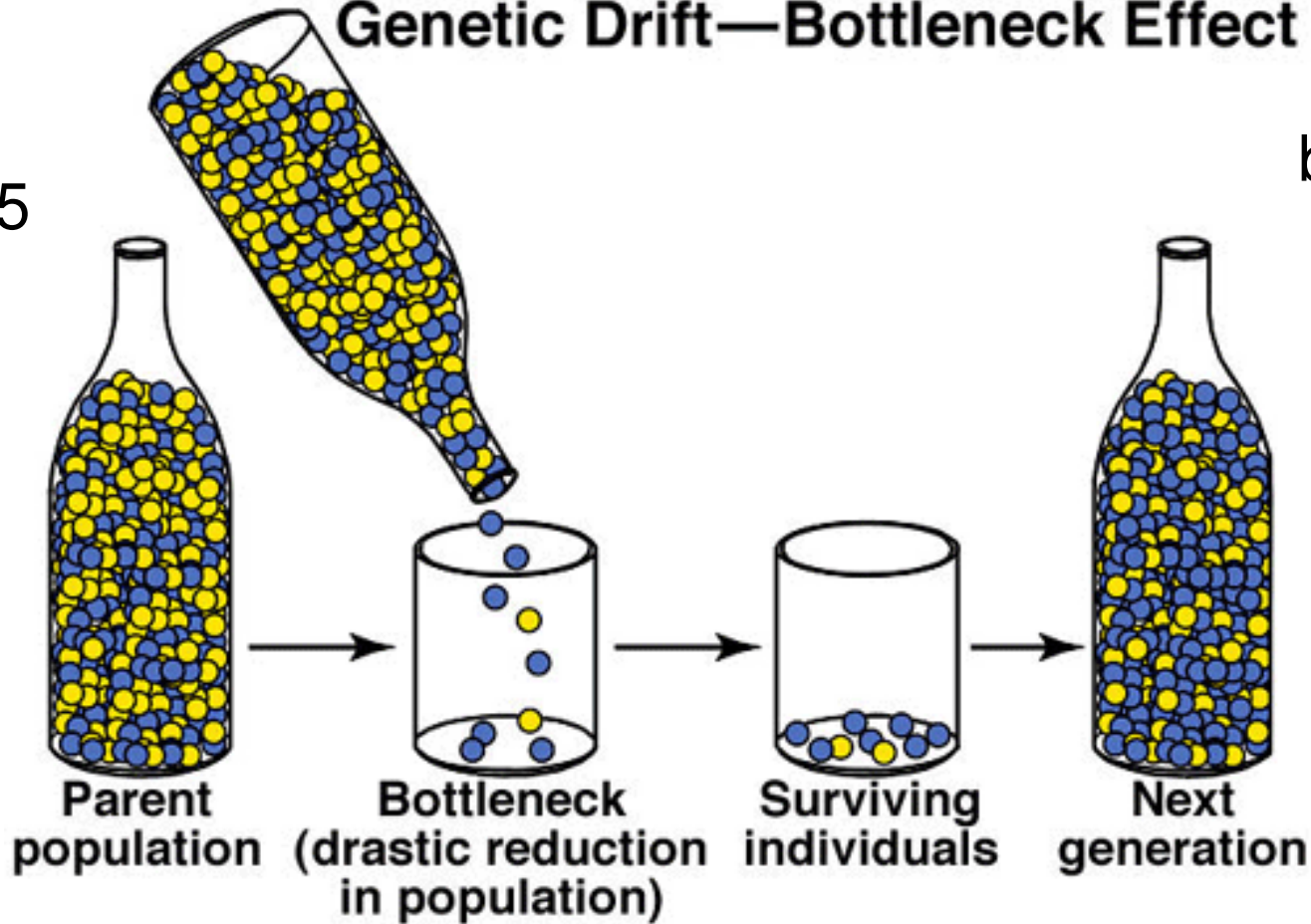
Random Genetic Drift

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

Genetic Drift—Bottleneck Effect

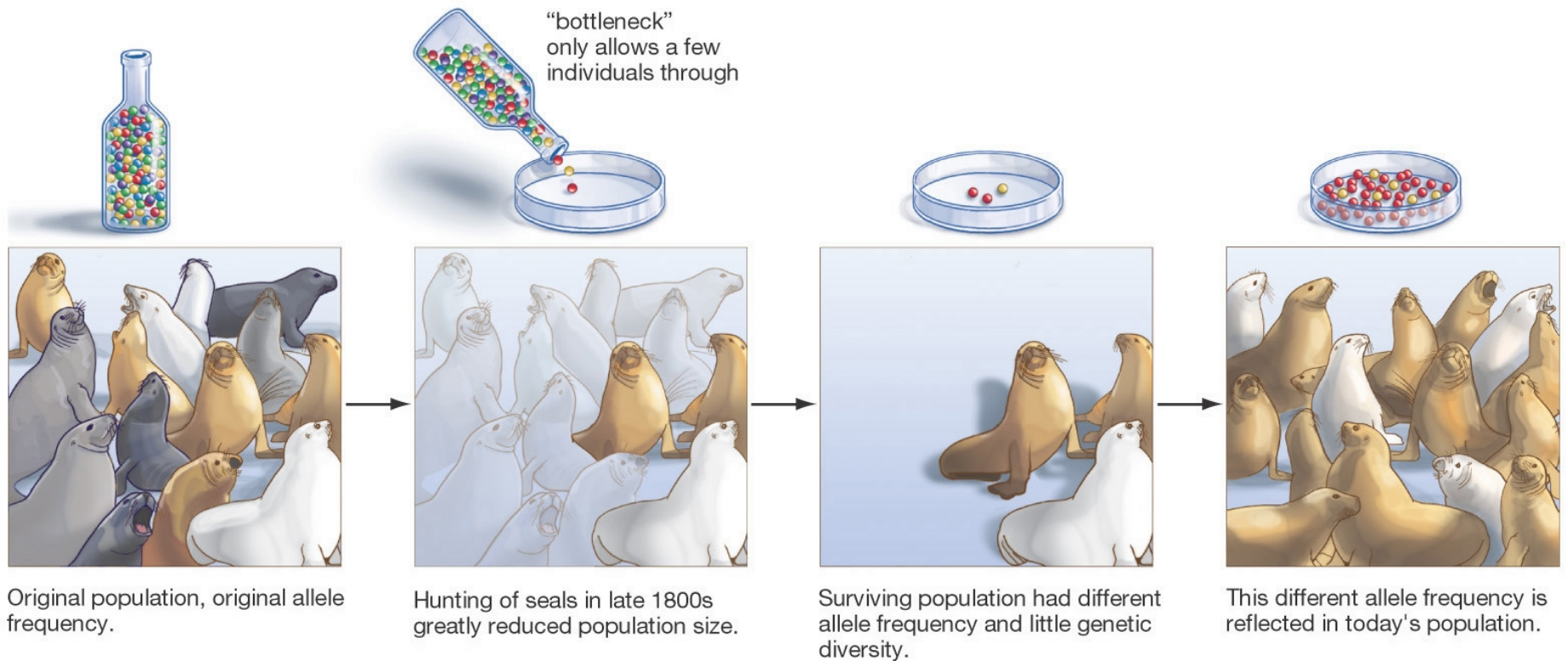
$b=y=0.5$

$b>y$



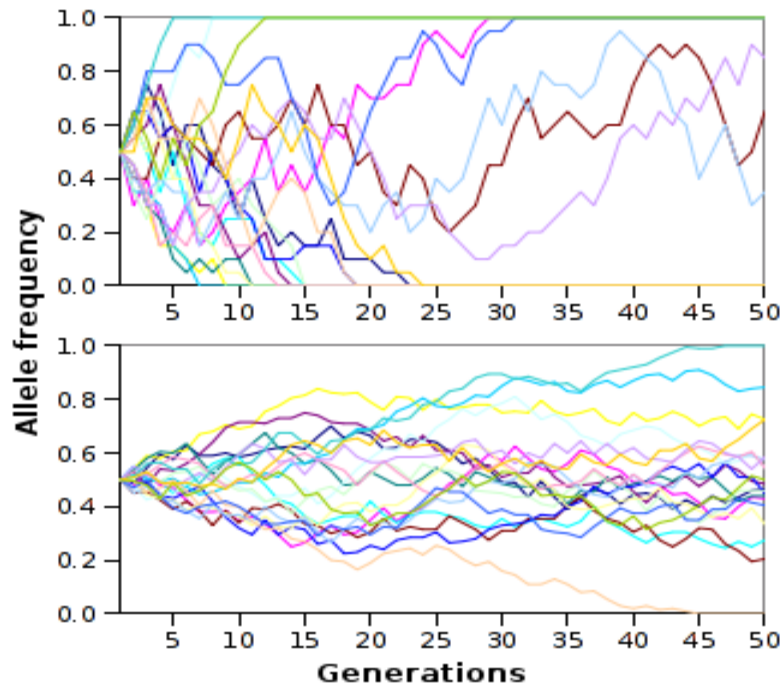
Bottleneck

A severe genetic bottleneck occurred in northern elephant seals.



Random Genetic Drift

Changing in frequencies



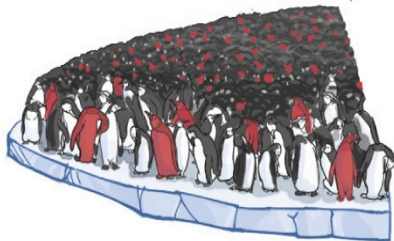
Simulations of genetic drift for 20 alleles with initial frequency 0.5, for populations of size $N=10$ and 100 , respectively. In general, alleles drift to fixation (frequency of 0 or 1) significantly faster in smaller populations.

Bottleneck

(a) Large population = 10,000 allele frequency = $\frac{1,000}{10,000} = 10\%$
(allele carriers in red)



50% of population survives, including 450 allele carriers



allele frequency = $\frac{450}{5,000} = 9\%$

little change in allele frequency
(no alleles lost)

(b) Small population = 10 allele carriers in red

allele frequency = $\frac{1}{10} = 10\%$



50% of population survives, with no allele carrier among them



allele frequency = $\frac{0}{5} = 0\%$

dramatic change in allele frequency
(potential to lose one allele)

Large populations = small effects. Small populations = large effects.

Population size

How big is that population?

Effective population size = number of adults contributing gametes to the next generation.

Formula for calculating N_e when there is a bias in the sex ratio

$$N_e = \frac{4 * N_f * N_m}{N_f + N_m}$$

Population size

Examples:

A flock of Geese:

50 males

50 females

(N = 100)

$$N_e = \frac{4 * 50 * 50}{50 + 50} = 100$$

A harem of Elephant seals:

4 males

96 females

(N = 100)

$$N_e = \frac{4 * 96 * 4}{96 + 4} = 15.36$$

Fixation

Gene substitution is defined as the process whereby a mutant allele completely replaces the predominant or wild type allele in a population. In this process, a mutant allele arises in a population, and becomes **fixed** after certain number of generations.

- The time it takes for a new allele to become fixed is called the **fixation time**
- Not all new mutants, however, reach fixation. In fact, the majority of them are lost after a few generations. Thus, we also need to address the issue of **fixation probability**.

Fixation Probability

Probability depends on:

- Initial frequency
- Selective advantage or disadvantage
- The effective population size N_e

$$P = (1 - e^{-(4N_e s q) / N}) / (1 - e^{-4N_e s})$$

Where:

q - the initial frequency of an allele

N_e - effective population size

N - absolute population size

S - selective advantage

Fixation Probability

There are some cases....

Example:

For neutral mutation $\rightarrow P = 1/(2N)$

For positive value of s
and large value of N $\rightarrow P = 2s$

Fixation time

The Time required for the fixation or loss of an allele depends on the frequency of the allele and the size of the population.

For a new mutation **without** selective advantage:

$$\bar{t} = 4 * N_e * G$$

For a new mutation **with** selective advantage:

$$\bar{t} = (2 / s) * \ln(2 * N_e) * G$$

Fixation time

Example:

For a new mutation **without** selective advantage:

A species with an effective population size $N_e = 10^6$ and a mean generation time $G = 2_{\text{years}}$

$$\rightarrow 4 * 10^6 * 2 = 8,000,000_{\text{years}}$$

For a new mutation **with** selective advantage:

$$S = 0.01$$

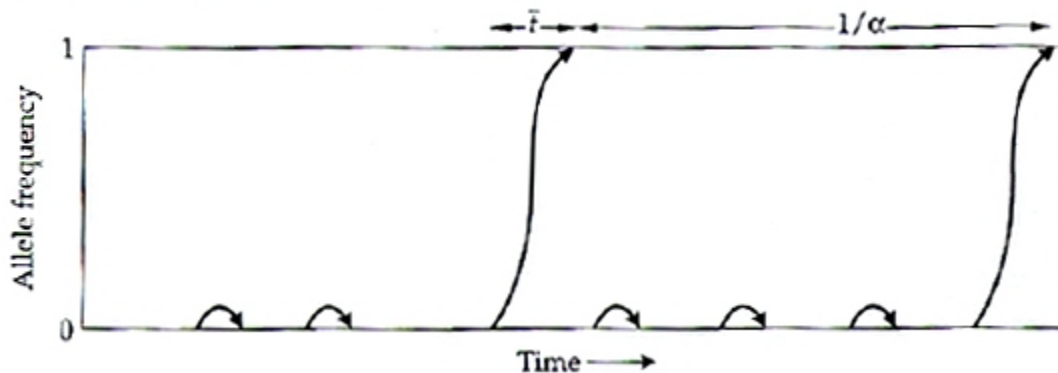
$$\rightarrow (2/0.01) * \ln(2 * 10^6) * 2 = 5803_{\text{years}}$$

Fixation

What happens to alleles that **have** an advantage?

- The probability of fixation (100% frequency) is close to 1.0
- the evolutionary time taken to reach fixation can be 'short'
- The time taken to reach fixation is much slower than under directional selection, and is proportional to s .

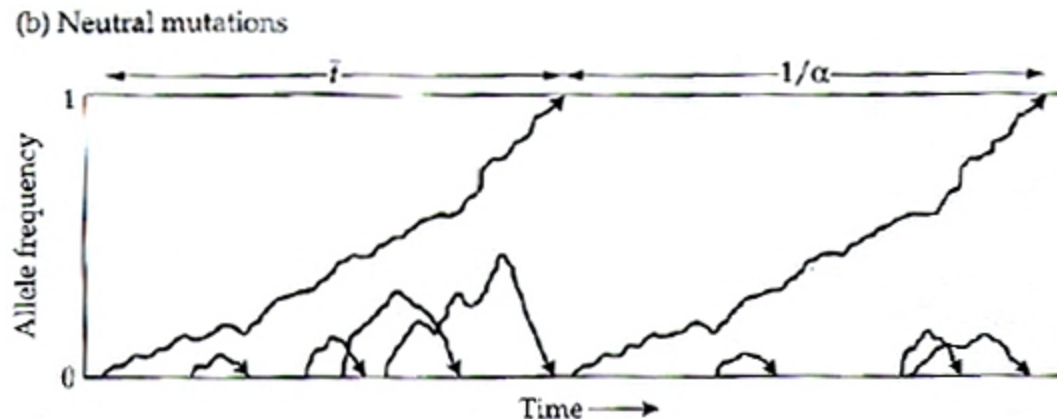
(a) Advantageous mutations



Fixation

What happens to alleles that have *no* advantage?

- Their probability of fixation is proportional to their frequency in the population.
- The time taken to reach fixation is much slower than under directional selection, and is proportional to the population size.



The Neo-Darwinian Theory versus The Neutral Mutation Hypothesis

- Neo-Darwinism: fusion of Darwin's theory of Natural Selection and the Mendelism, that states:
- Mutation is recognized as the ultimate source of genetic variation, but natural selection is given the dominant or 'creative' role in shaping the genetic makeup of populations
- Gene substitution occurs as consequence of selection for advantageous mutations
- Polymorphism is maintained by balancing selection
- ➔ Substitution and polymorphism are 2 separate phenomena driven by different evolutionary forces

The Neo-Darwinian Theory versus The Neutral Mutation Hypothesis

- Neutral Theory of Molecular Evolution: on molecular level, the majority of evolutionary changes and much of the variability within species are caused by random genetic drift of mutant alleles that are selectively neutral or nearly so
- Selection might operate, but is too weak to offset the influences of chance effects
- ➔ Substitution and polymorphism are 2 facets of the same phenomenon

The Neo-Darwinian Theory versus The Neutral Mutation Hypothesis

Agreement:

- Most new mutations in Proteins are deleterious and are quickly removed from the population

Disagreement:

- Shape of the genome
- About proportion of neutral mutations among nondeleterious mutations

Controversy had strong impact on research in mol. evolution:

- General recognition of chance effect of random drift
- Synthesis of molecular biology and population genetics

➔ Controversy will continue, but it is recognised, that both aspects must be consistent in adequate theories

Thank You!

Any Questions???

Literature

- Molecular Evolution by Li (1997). Chapter 2: Dynamics of genes in populations
- A review on the 'Nearly Neutral Theory of Evolution': Ohta (1992) The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Syst. 23:263-86
- www.wikipedia.com
- www.wikipedia.de
- <http://www.ndsu.nodak.edu/instruct/mcclean/p>