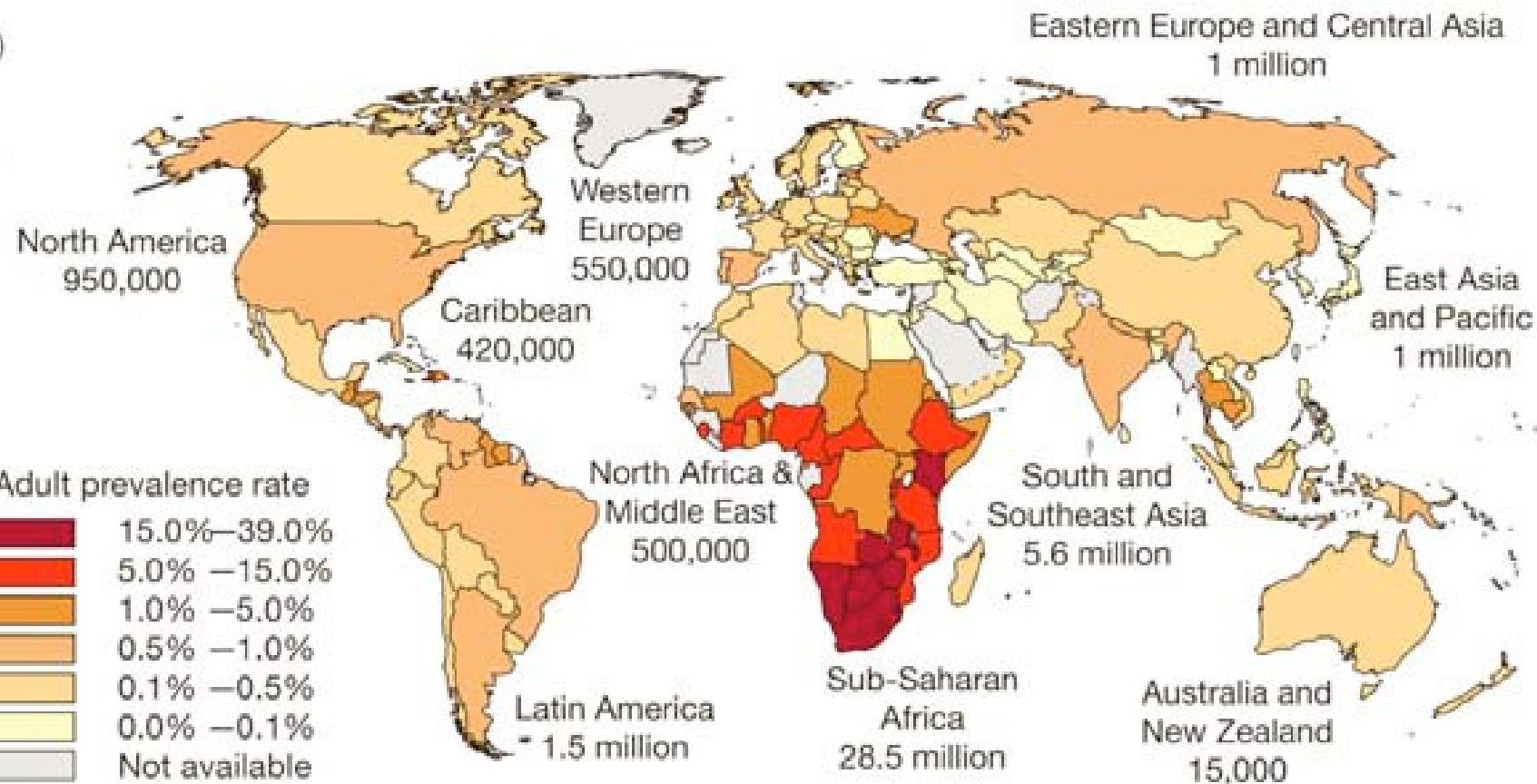


Evolutionary Biology

Aims

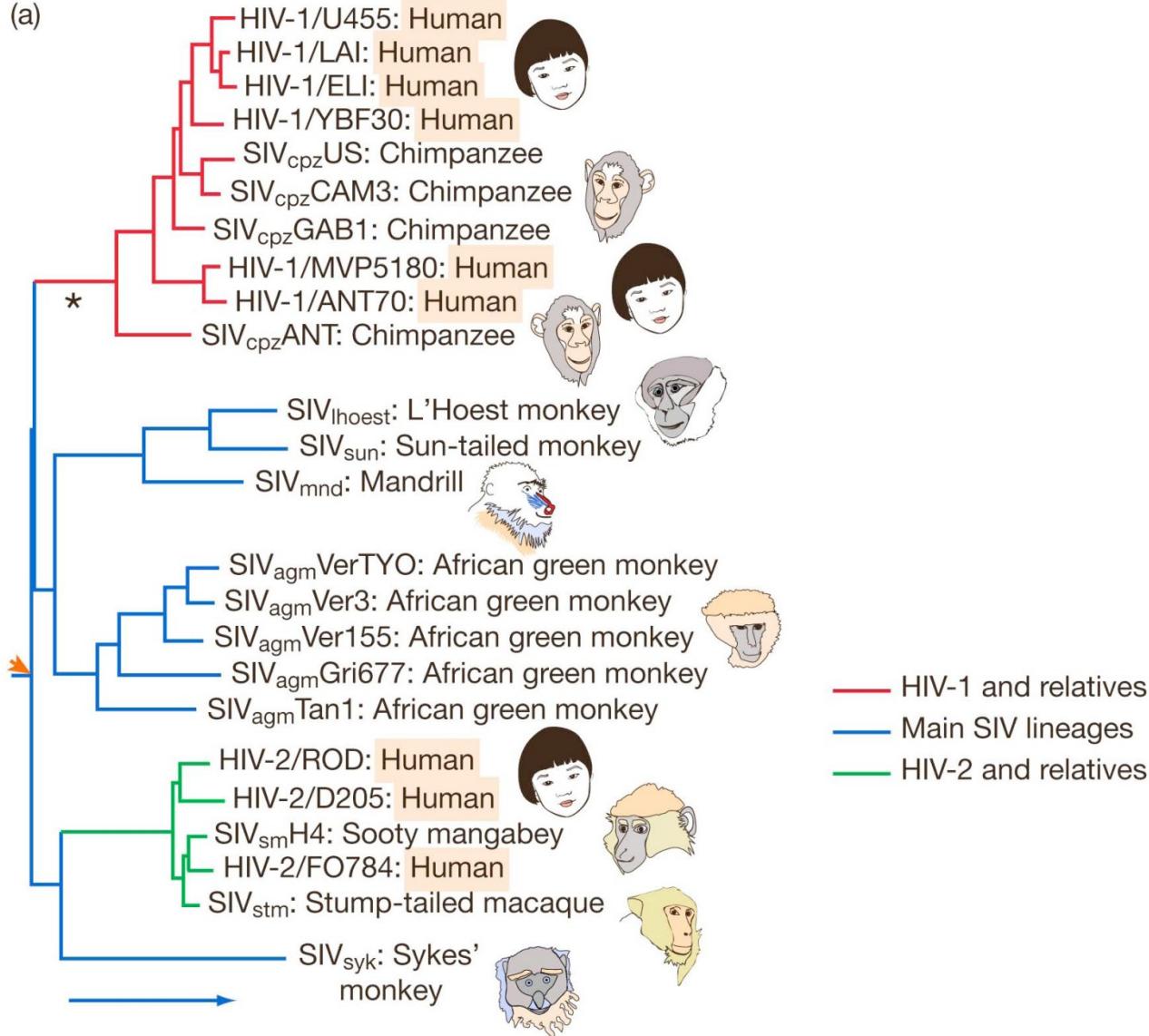
- The idea of phylogenetic tree
- To show the concept of natural selection
- The concept of co-evolution
- The idea of Trade-offs
- The nature of natural selection: Not goal oriented

(a)



Copyright © 2004 Pearson Prentice Hall, Inc.

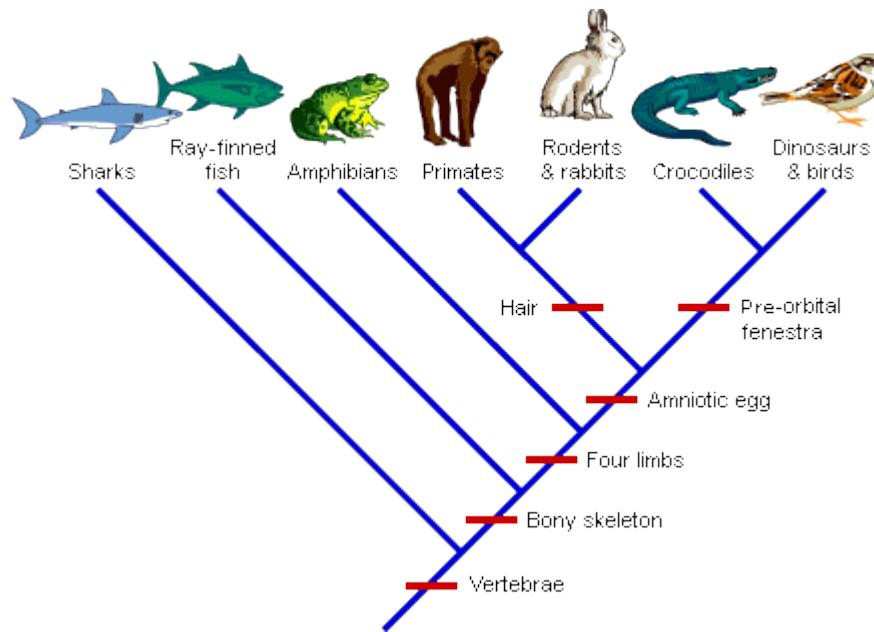
(a)



Copyright © 2004 Pearson Prentice Hall, Inc.

Reading phylogenetic trees: A quick review

(Adapted from evolution.berkeley.edu, Gloria Rendon)



- A phylogeny, or evolutionary tree, represents the evolutionary relationships among a set of organisms or groups of organisms, called taxa (singular: taxon) that are believed to have a common ancestor.

Leaves
/External
node/OTU

A and B are sister groups

C is the outgroup
to A and B

taxon A

taxon B

taxon C

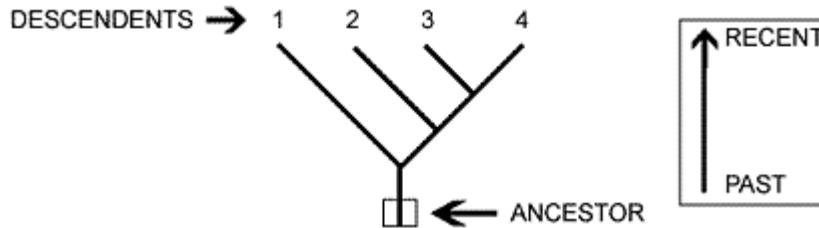
Edge/Branch/Lineage
In some trees,
Branch Length (Distance)
~ Time

common ancestor
of A and B

Internal Node
(Common
Ancestor)

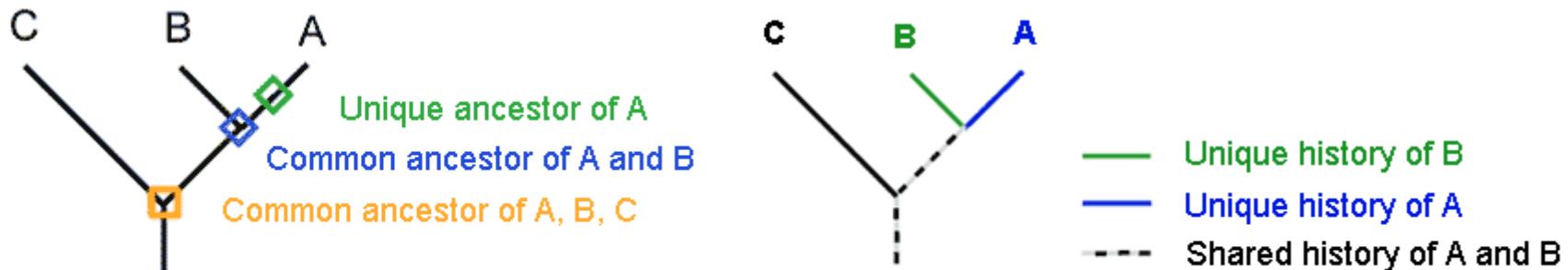
Root
(Common ancestor of all
Operational Taxonomic Unit- OTU)

Tips, Internal Nodes, Edges



- The tips of the phylogenetic tree represent groups of descendant taxa (often species)
- The internal nodes of the tree represent the common ancestors of those descendants.
- The tips are the present and the internal nodes are the past.
- The edge lengths in some trees correspond to time estimates - evolutionary time.

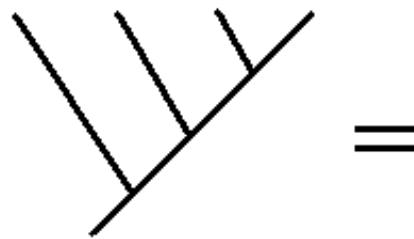
Sister Groups and a common ancestor



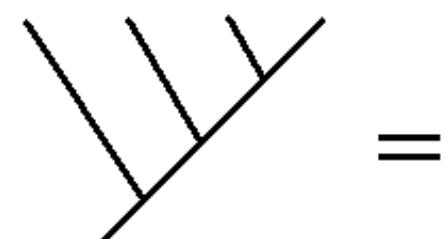
- Two descendants that split from the same node are called sister groups.
- In the trees above, species A & B are sister groups — they are each other's closest relatives; which means that:
 - i) they have a lot of evolutionary history in common and very little evolutionary history that is unique to either one of the two sister species and
 - ii) that they have a common ancestor that is unique to them.

Equivalent trees

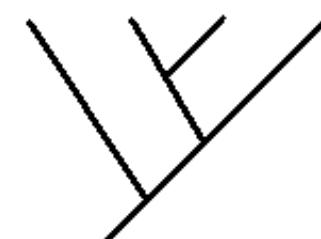
MOSS FERN PINE ROSE



MOSS FERN ROSE PINE

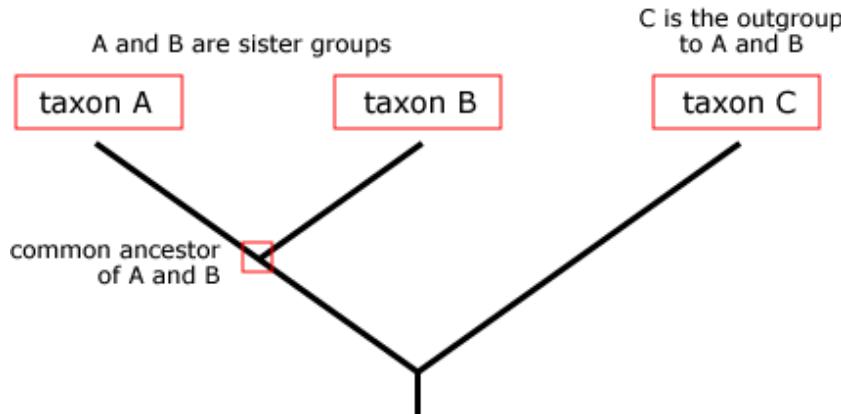


MOSS PINE ROSE FERN



- For any speciation event on a phylogeny, the choice of which lineage goes to the right and which one goes to the left is arbitrary.
- These three phylogenies are therefore equivalent.

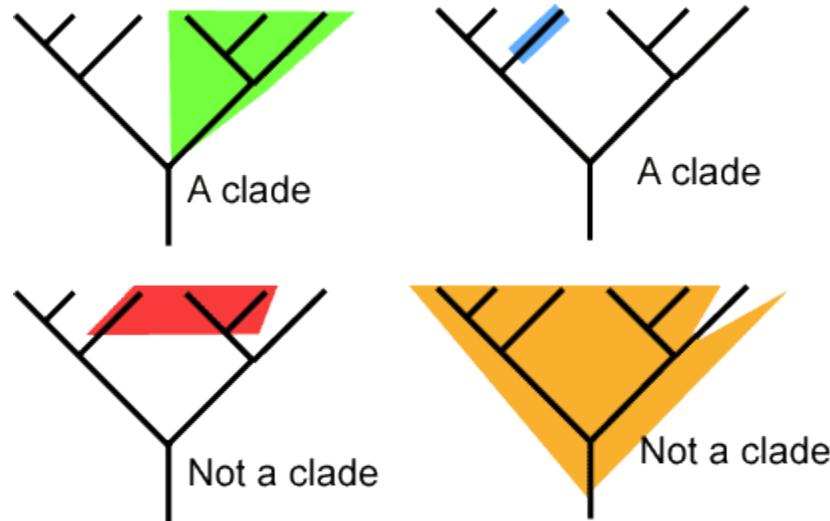
Outgroup



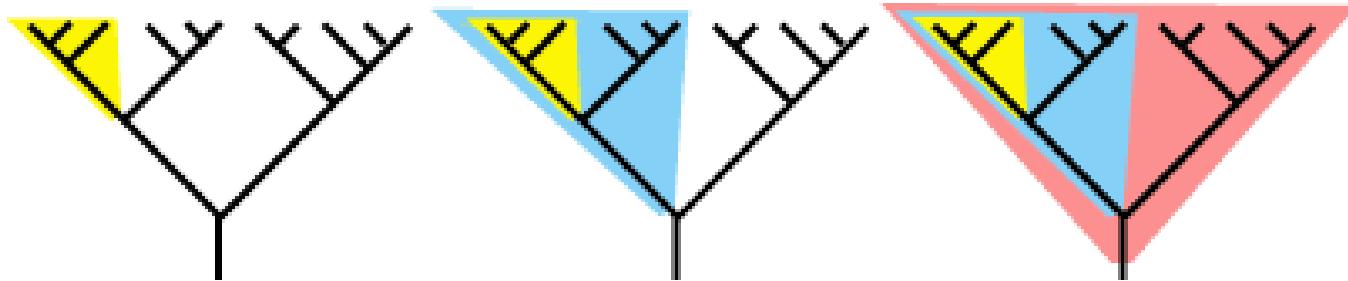
- Many phylogenies also include an outgroup — a taxon outside the group of interest.
- All the members of the group of interest are more closely related to each other than they are to the outgroup. Hence, the outgroup stems from the base of the tree.
- An outgroup can give you a sense of where on the bigger tree of life the main group of organisms falls. It is also useful when constructing evolutionary trees.

Branches and clades

- Evolutionary trees depict clades.
- A clade is a group of organisms that are all descended from a common ancestor; thus a clade includes an ancestor and all descendants of that ancestor.
- You can think of a clade as a branch on the tree of life.
- Some examples of clades and non-clades in a phylogenetic tree are shown here



More on clades. Nested clades



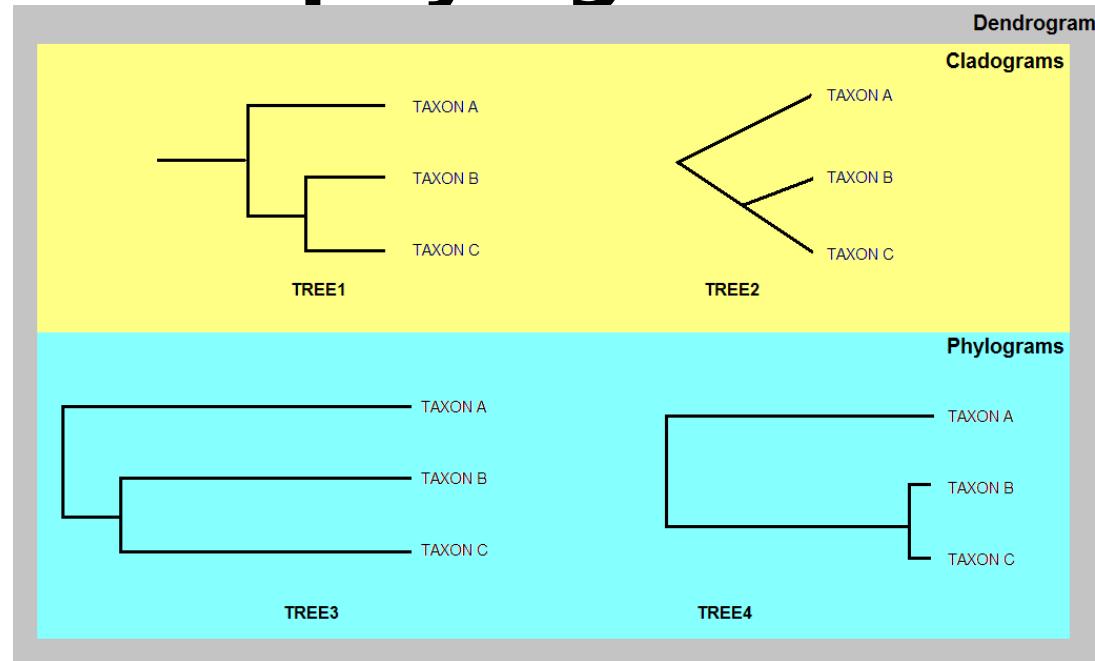
- Clades are nested within one another — they form a nested hierarchy.
- A clade may include many thousands of species or just a few.
- Some examples of clades at different levels are marked on the phylogenies above.
- Notice how clades can be nested within larger clades.

Types of trees: unrooted vs rooted



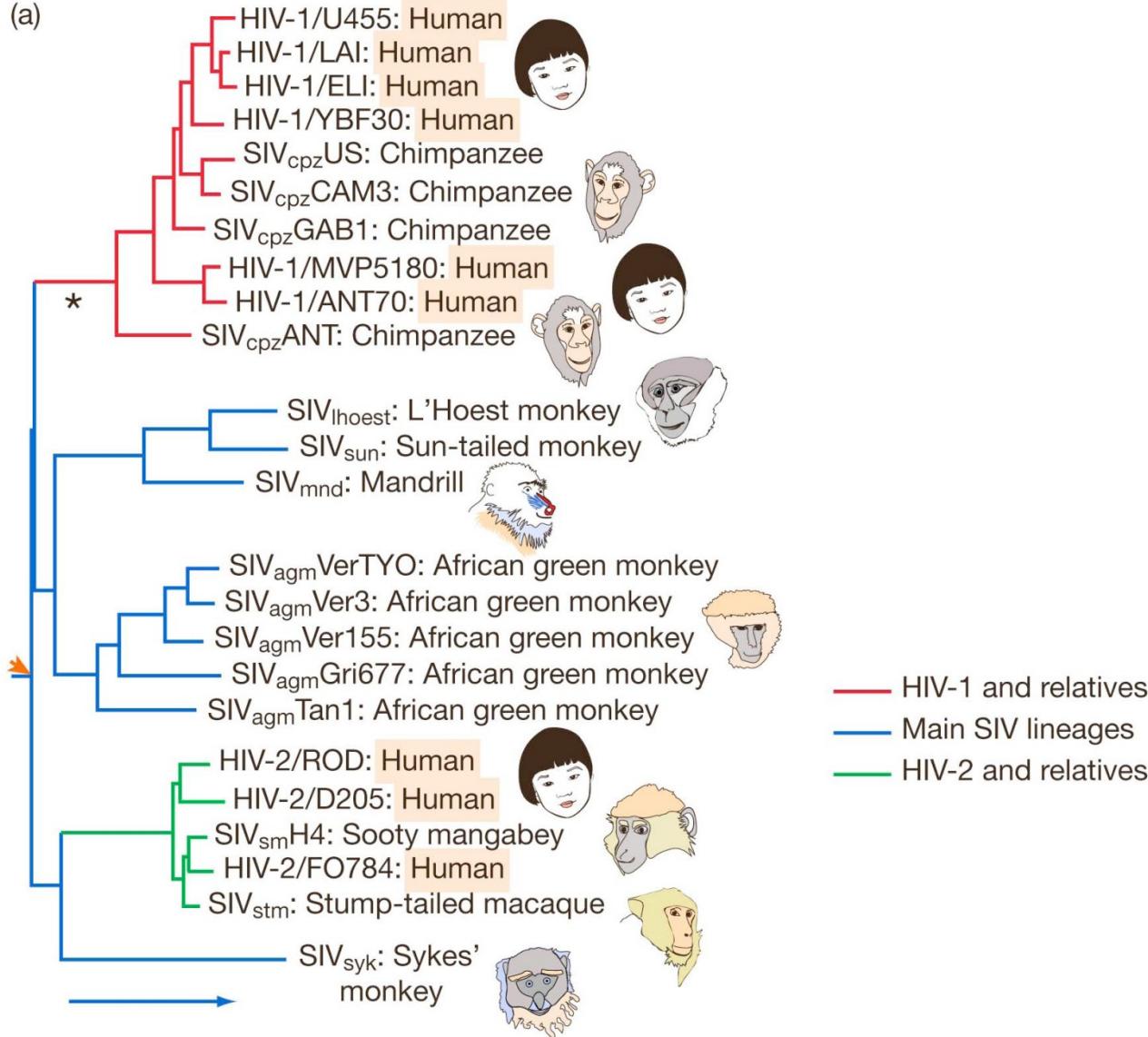
- A **rooted** phylogenetic tree is a tree with a unique root node corresponding to the (usually imputed) most recent common ancestor of all the entities at the leaves (aka tips) of the tree. A rooted tree is a binary tree.
- **Unrooted** trees illustrate the relatedness of the leaf nodes without making assumptions about common ancestry. An unrooted tree has a node with three edges; the rest of the nodes have up to two edges.

Dendrogram, cladogram, phylogram



- Dendrogram is the ‘generic’ term applied to any type of diagrammatic representation of phylogenetic trees. **All four trees depicted here are dendograms.**
- Cladogram (to some biologists) is a tree in which branch lengths DO NOT represent evolutionary time; clades just represent a hypothesis about actual evolutionary history
TREE1 and TREE2 are cladograms and TREE1 = TREE2
- Phylogram (to some biologists) is a tree in which branch lengths DO represent evolutionary time; clades represent true evolutionary history (amount of character change) **TREE3 and TREE4 are phylogenograms and TREE3 ≠ TREE4**

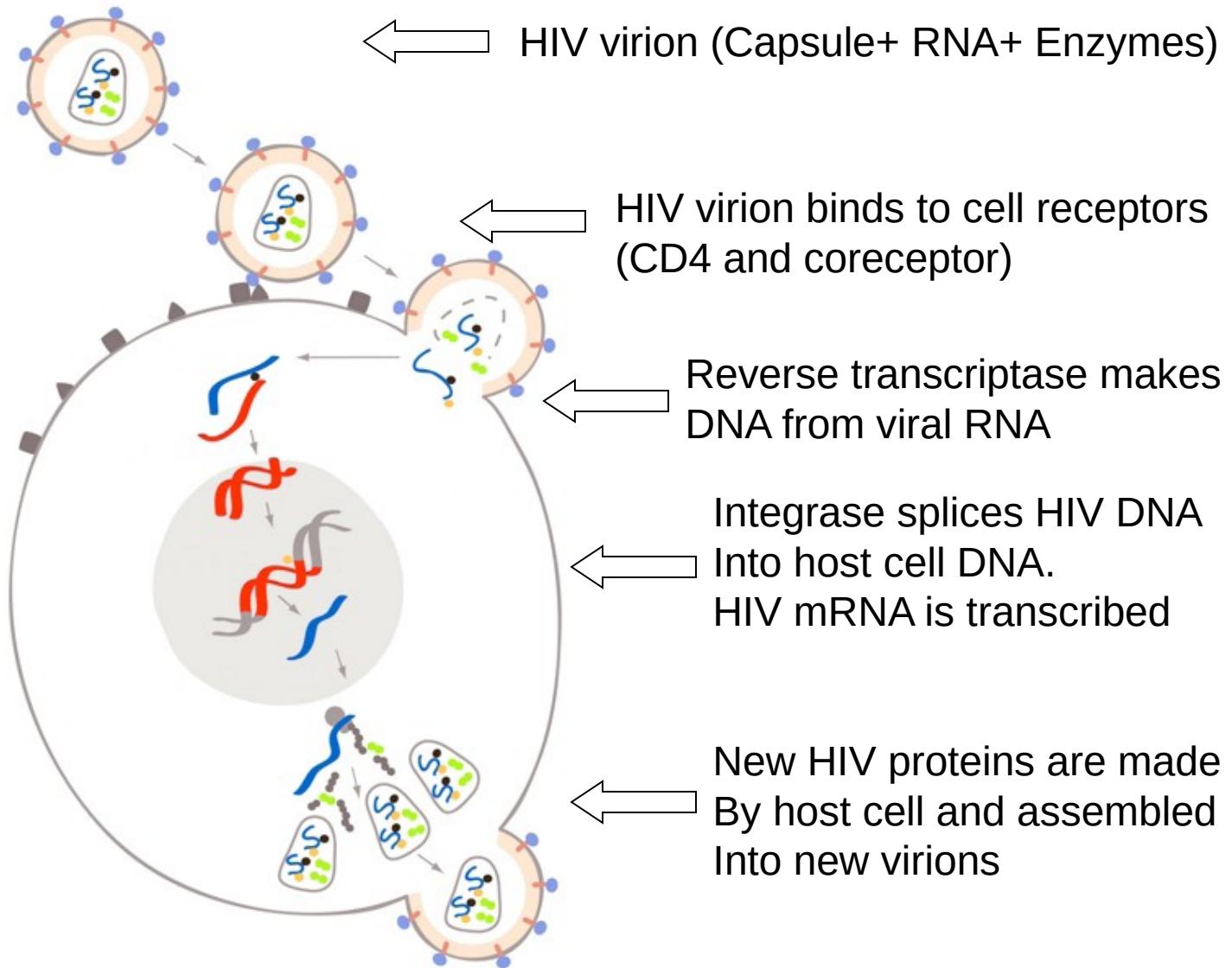
(a)



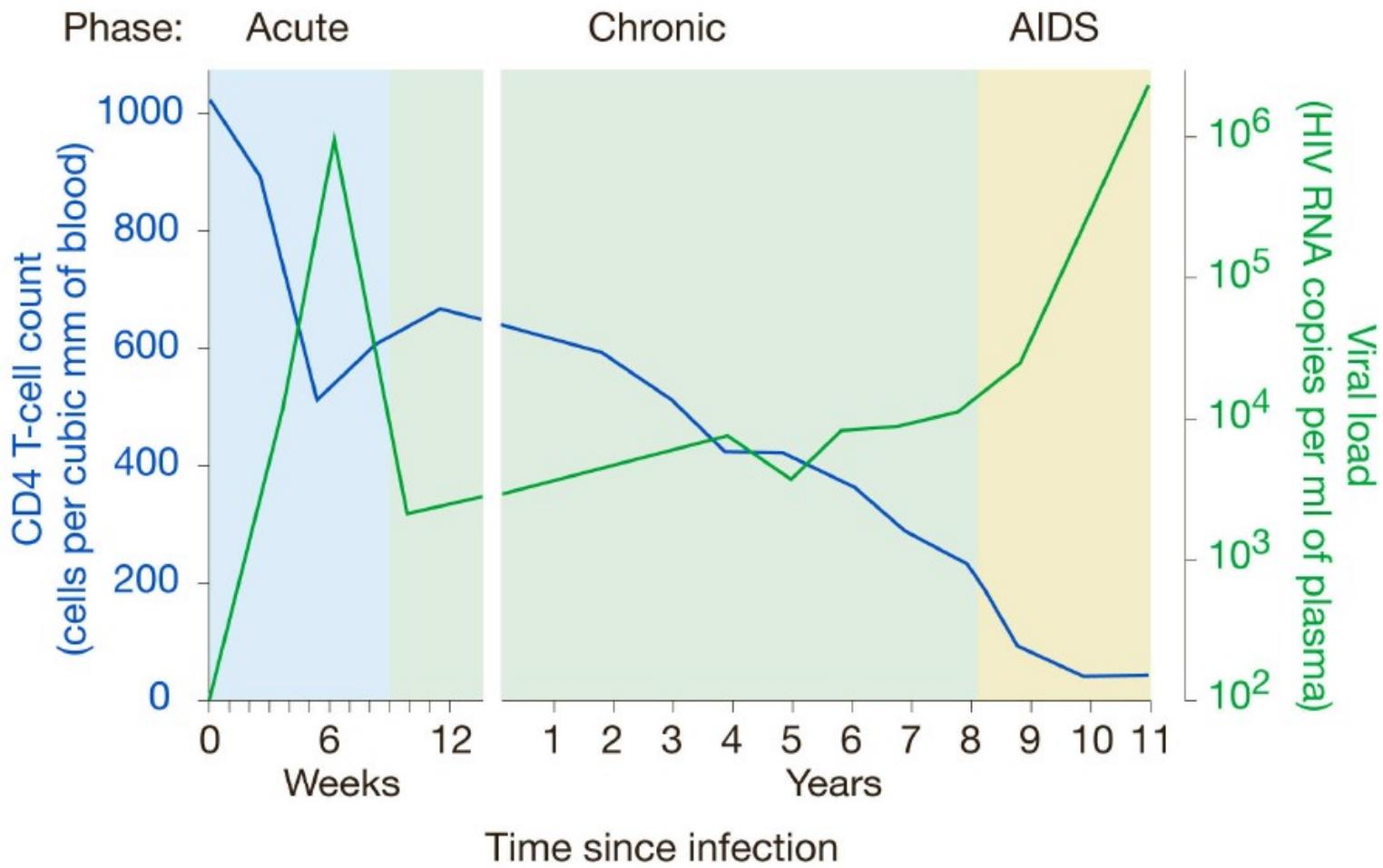
Copyright © 2004 Pearson Prentice Hall, Inc.

Retro Viruses

- HIV is an RNA based virus.
- Its genetic code needs to be translated into DNA before it can transcribe anything.
- It depends upon reverse transcription within the host cell to convert itself into DNA.



Progression of HIV infection

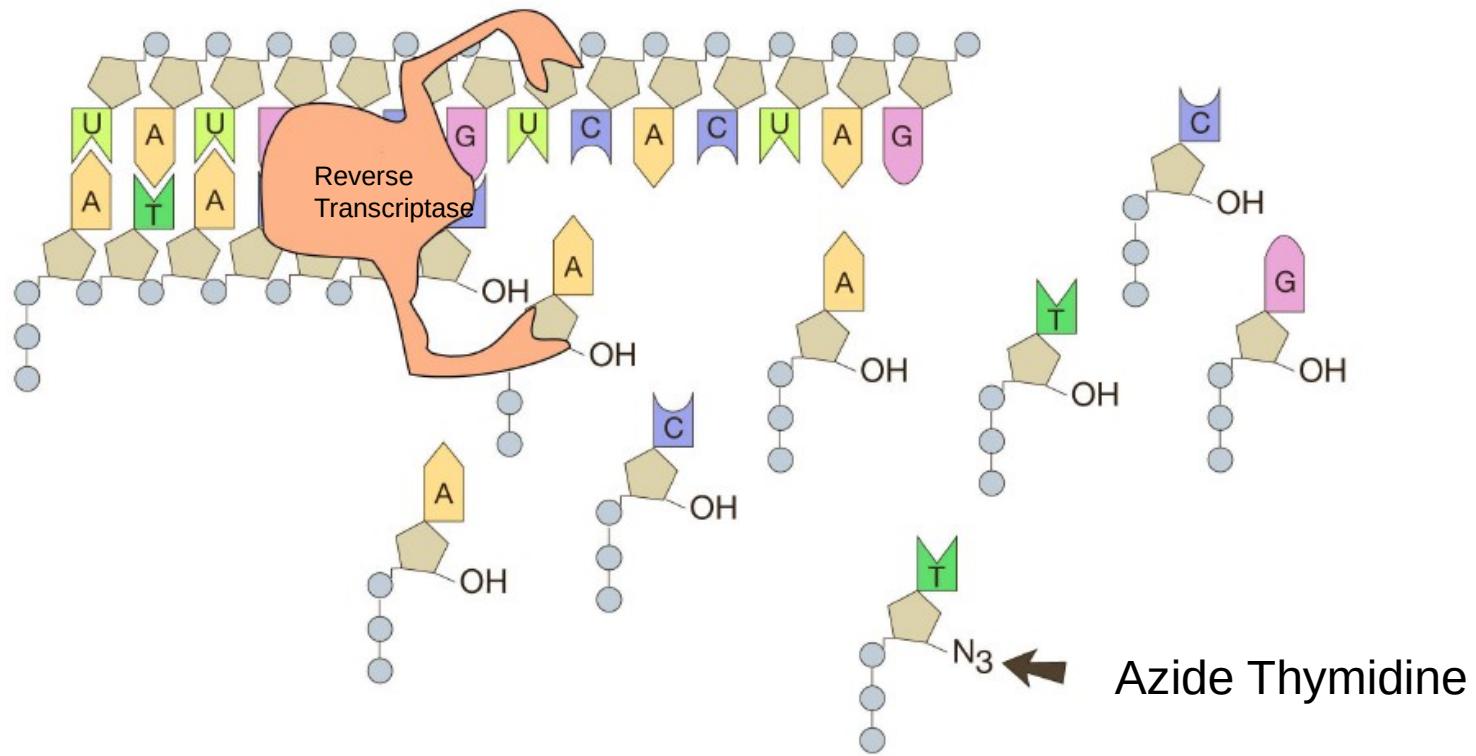


Copyright © 2004 Pearson Prentice Hall, Inc.

AZT

- A promising therapy based on striking reverse transcriptase into inserting a terminating base.
- Instead of normal Thymidine the enzyme inserts Azide Thymidine
- AZT worked in initial trials. Stopped the decline of macrophages.

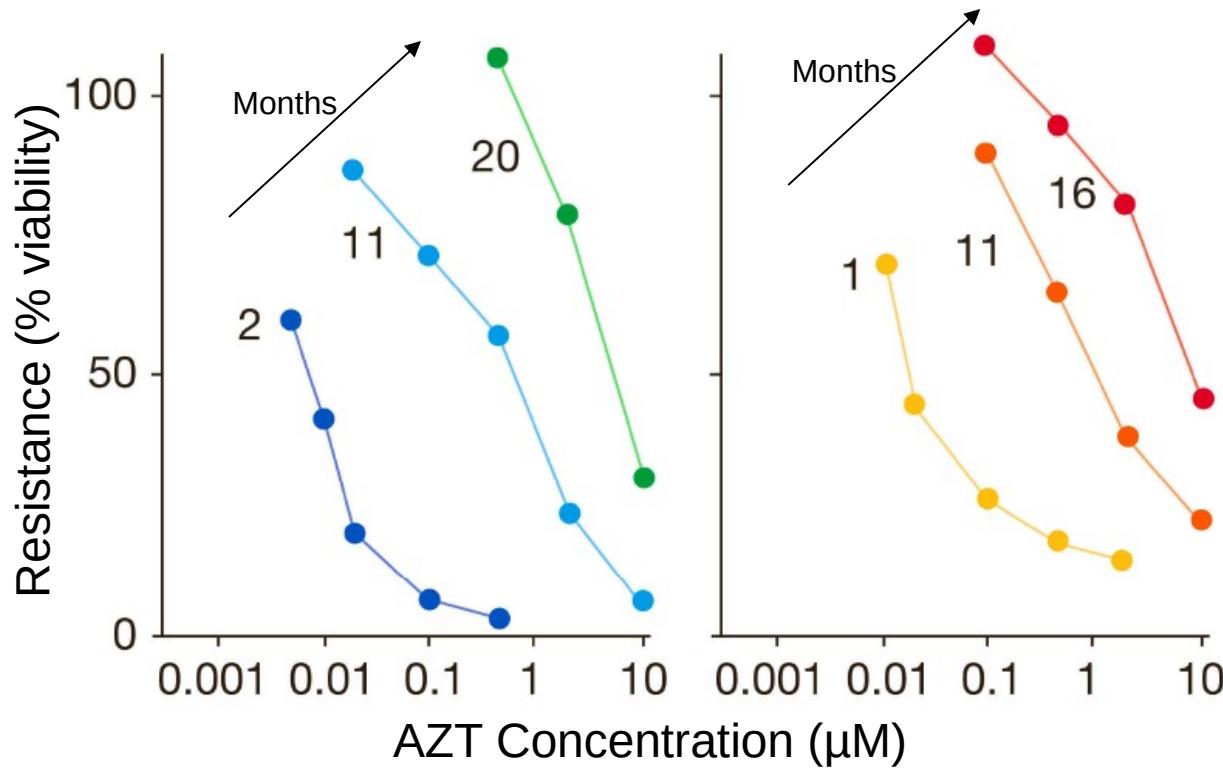
How AZT works



Copyright © 2004 Pearson Prentice Hall, Inc.

AZT stops working quickly

(a)

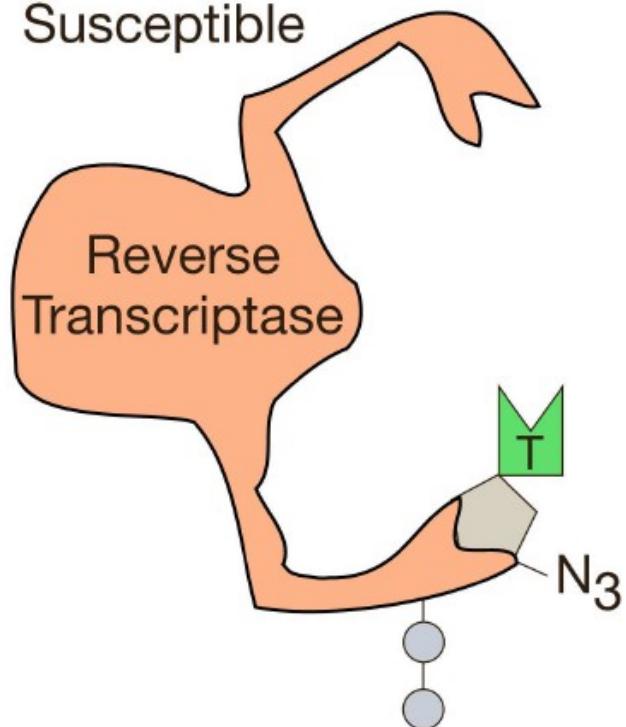


Why does resistance develop?

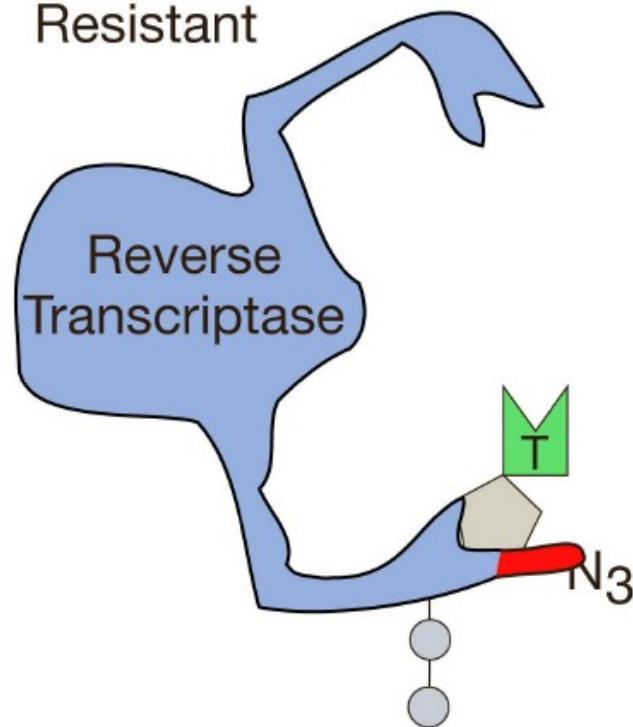
- Patient's own cells may change to avoid AZT
 - May exclude it from the cell or fail to phosphorylate it.
- The virus could evolve to avoid using AZT during reverse transcription.

Changes in the RT active site evolve very quickly

(a) Susceptible



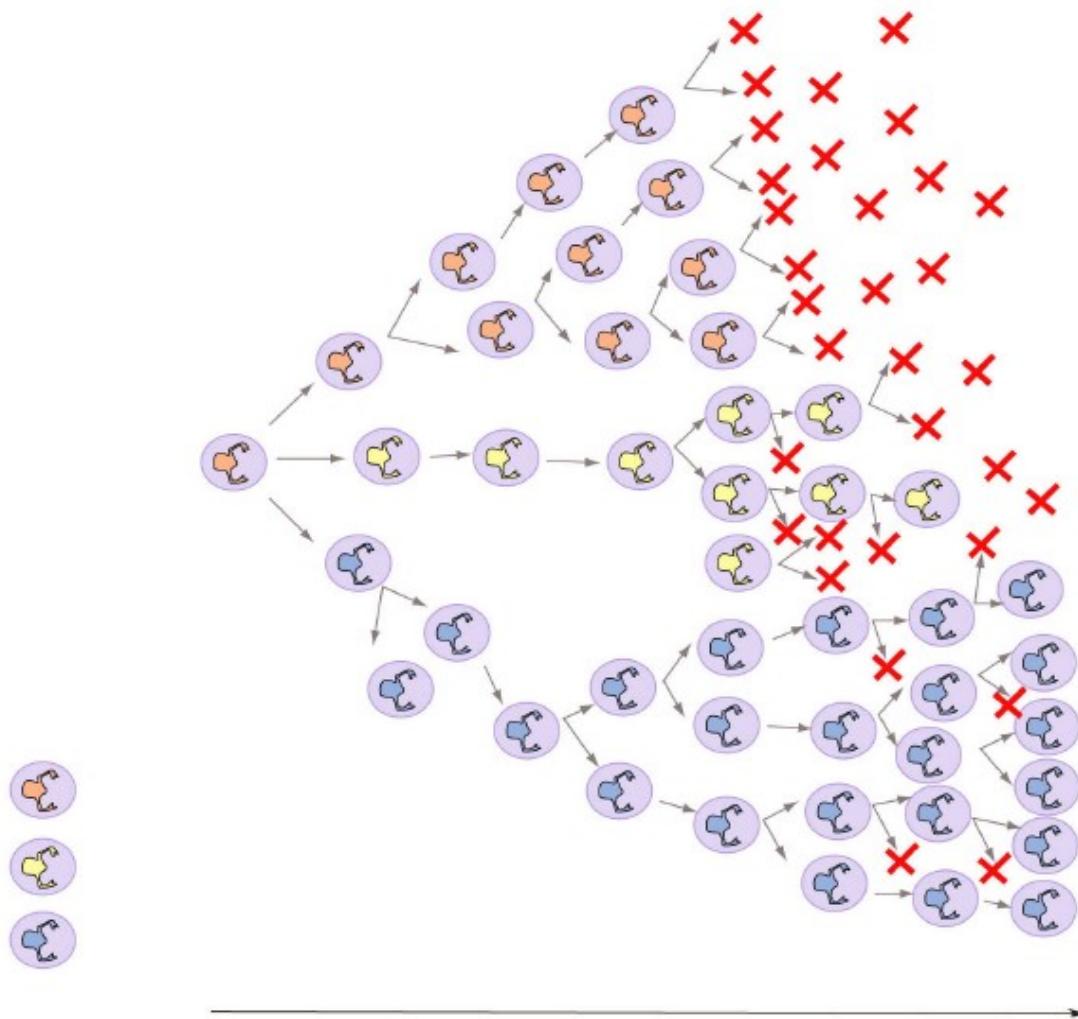
Resistant



Copyright © 2004 Pearson Prentice Hall, Inc.

Bad copy makes good evolvability

- RT highly error prone
 - No error correction mechanism
 - Highest rates of mutation
 - 50% of the DNA copies mutated
- Virion: Many thousands of generations with in a host (est generation time 1.2 days)



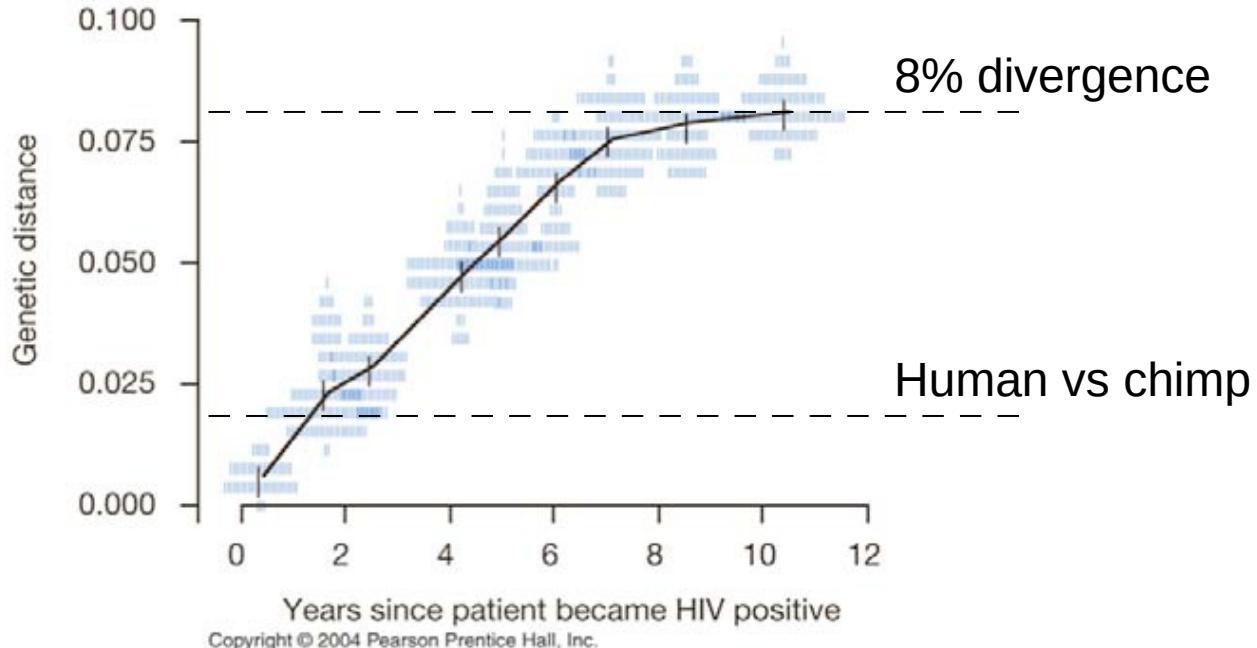
Copyright © 2004 Pearson Prentice Hall, Inc.

Lessons

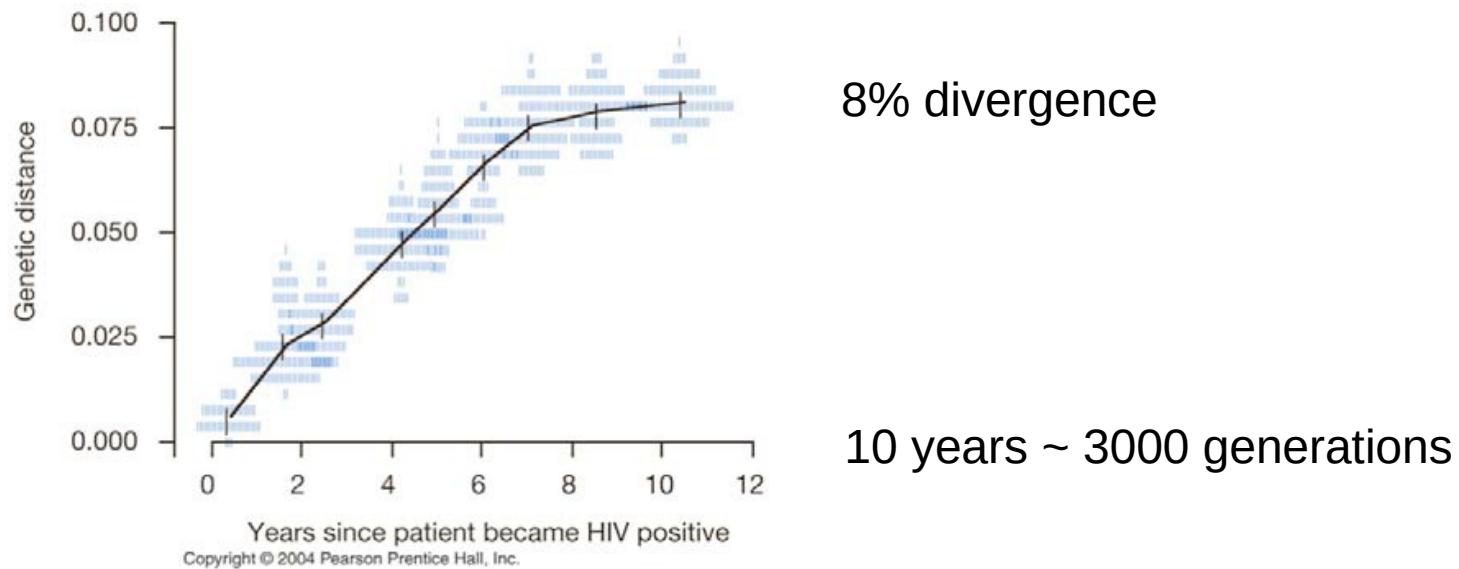
- Variation
- Heritable
- Non random survival and reproduction
 - That is natural selection in operation!

Divergence at binding site genes

(a) Divergence from founder population



(a) Divergence from founder population



- $4 \times$ the divergence between humans and chimp
- Humans and chimps diverged ~ 5.4 mya
 - $\sim 500,000$ generations
 - $\times 2$
 - $= 1,000,000$ generations

1,300 times faster sequence divergence in HIV RT

Compensatory evolution

- AZT-resistant RT will copy more slowly than susceptible RT
 - This fitness trade-off indicates that removal of therapy should lead to a reversion to non-selective forms.
- Would pulsed treatment be better?
 - Body retains copies of old virions
 - Resistance re-evolves very quickly
 - Resistant forms become more efficient with prolonged exposure

Why kill the host?

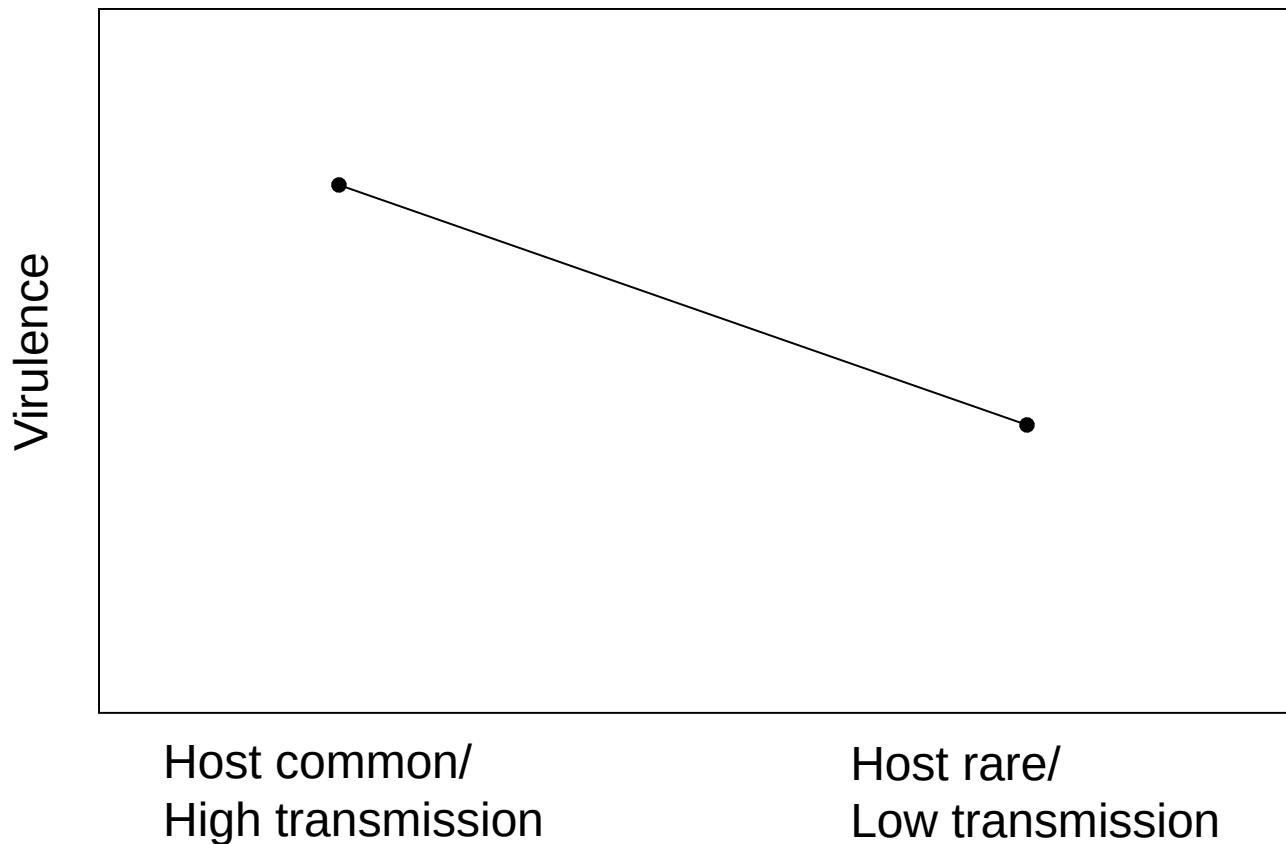
- Death of the host means the death of all virions carried by the host
- No further possibility of transmission
- Evolution does not think ahead
- Virulence usually correlated with host availability/ease of transmission

Levels of selection

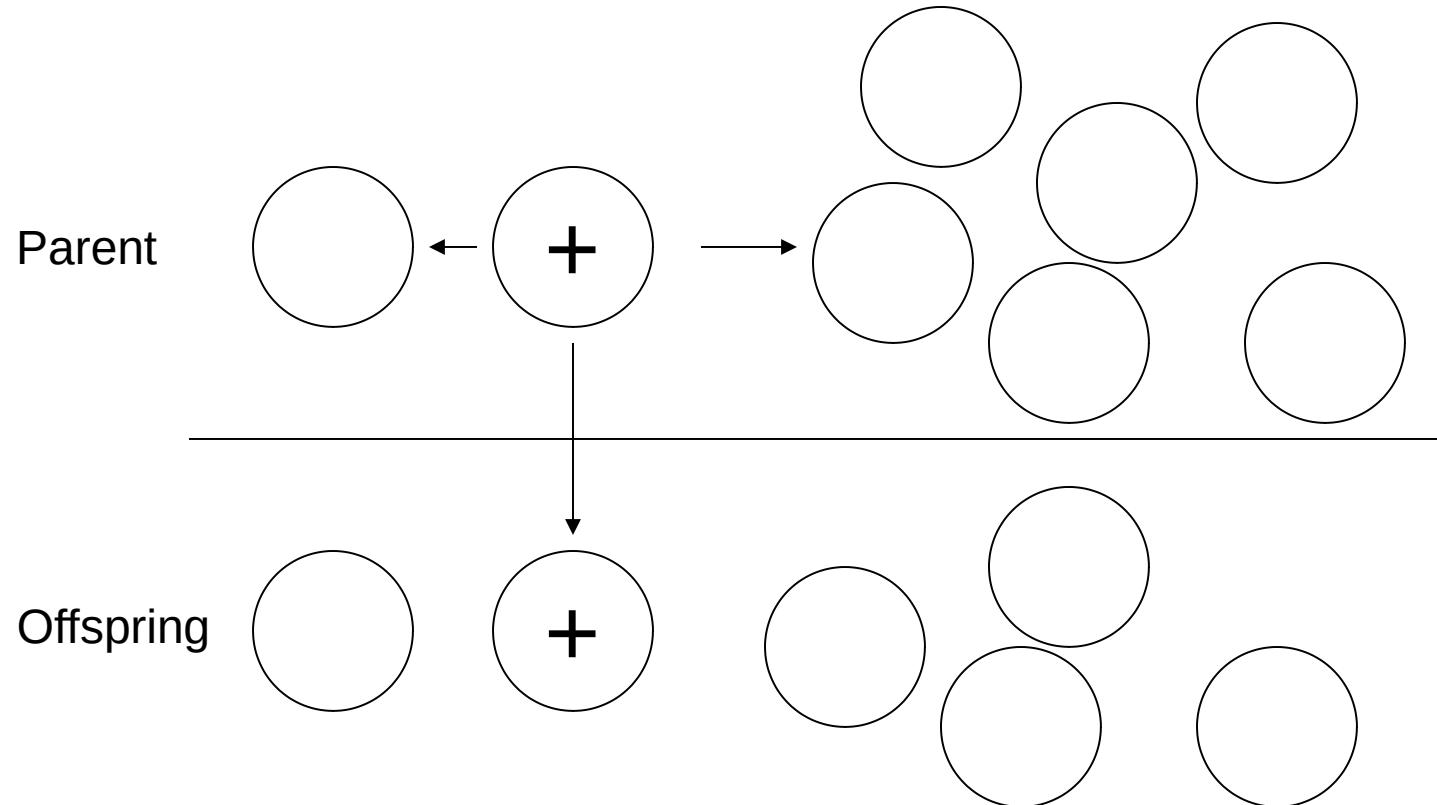
- Fitness of a virion
 - Intra host success¹ × inter host transmission²
 1. Selection within the infected host favours the viral particles that multiply most rapidly

These successful viral particles may rapidly degrade host performance and the probability of transmission
 2. Selection for transmission between hosts may favour lower virulence

Transmission effects on virulence

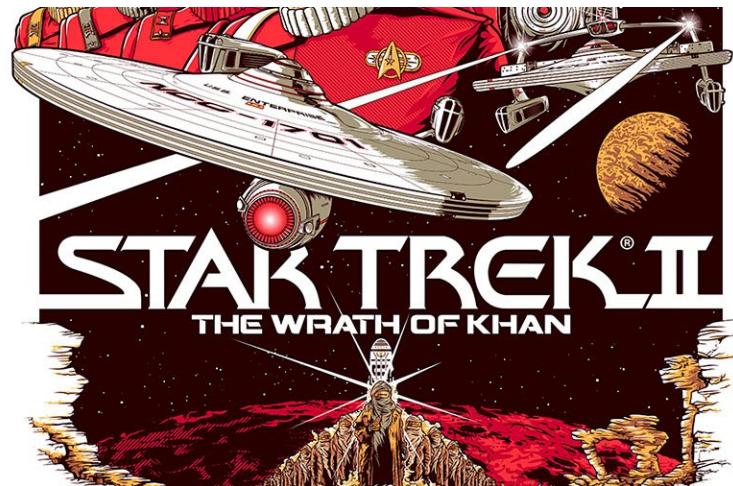


Modes of transmission



HIV is going both ways

- Current infections of HIV-1 creates very high viral titres before degradation of health
- Vertical transmission of HIV-1 a dead end
- Horizontal transmission and high virulence being selected for
- More benign strains (eg. HIV-2) spreading slowly but steadily

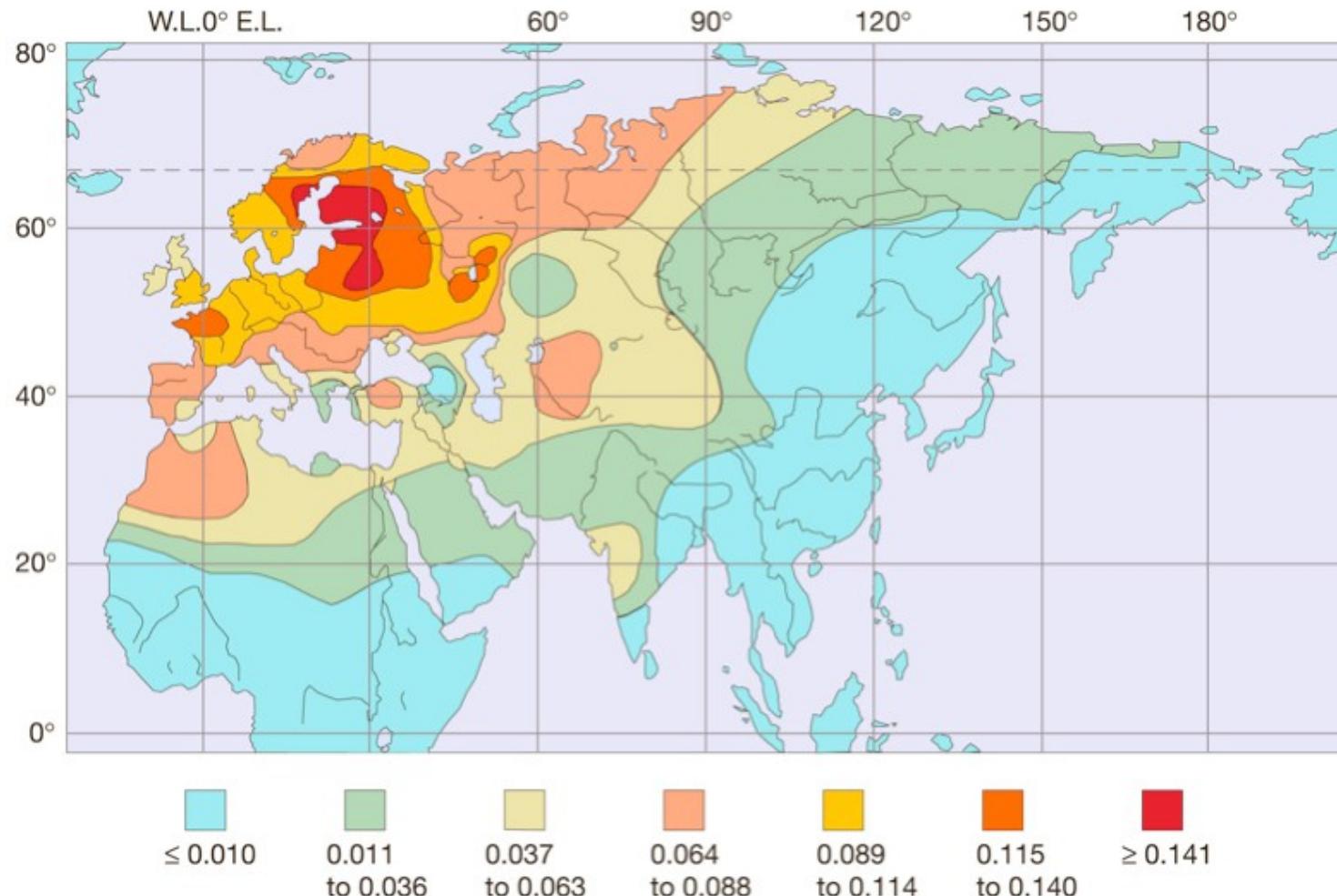


A truly **successful parasite** is commensal, living in amity with its host, or even giving it positive advantages.... A **parasite** that regularly and inevitably kills its host cannot survive long, in the evolutionary sense, unless it multiplies with tremendous rapidity.... It is not pro-survival.

HIV suicide?

- HIV within host selection favours novel epitopes
- Virions switch for CCR5 to CXCR4 coreceptors over time with in hosts
- Host X4 T cells increase cell division with protracted infection
- Strongly decreases transmissibility
- Uninfected hosts have very low levels of CXCR4 cell division

Distribution of resistant CCR5 mutants



Frequency of CCR5-Δ32 allele
Copyright © 2004 Pearson Prentice Hall, Inc.

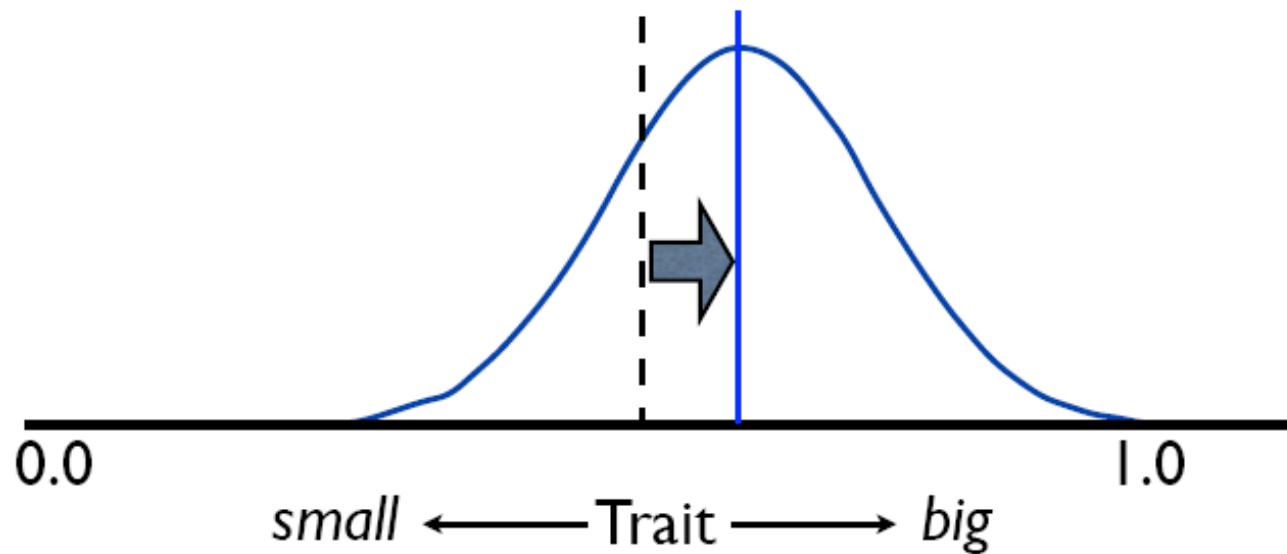
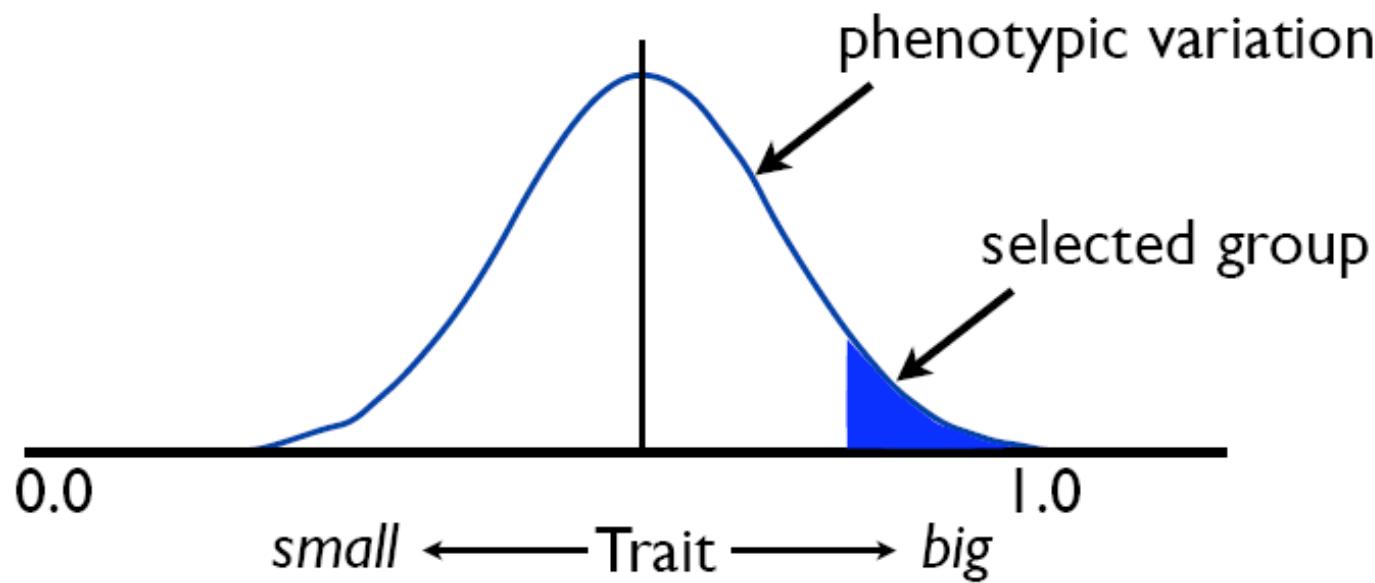
Genetic variation in resistance

- Some people are more resistant to HIV than others
- The Scandinavian mechanism involves the deletion of CCR5 receptor ($\text{CCR5}-\Delta 32$)
- Is this pattern due to selection?
- If so, can a recent origin of HIV explain it?

Take Home Messages

Darwin's Big Idea

1. There is variation within breeding groups (*populations*).
2. Part of the observed variation is *heritable*.
3. Some individuals are more successful in survival & reproduction (*fitness*) than others.
4. Variation influences fitness: those with the most favourable *phenotypes* are *naturally selected*.



Take Home Messages:

Evolution:

- Population level process.
 - “Descent with modification”
 - “Change in Phenotypic frequency over time”
 - “Change in allelic frequency over time”

Selection:

“Heritable, Differential Reproductive Success”

Take Home Messages:

- Evolution can occur very rapidly.
 - Mutation creates random variation.
 - Selection sorts this random variation leading to non-random survival and reproduction.
 - Adaptive evolution is NOT random.

“to Evolve” vs “to Adapt”

- Not all evolution is adaptive evolution!

Take Home Messages:

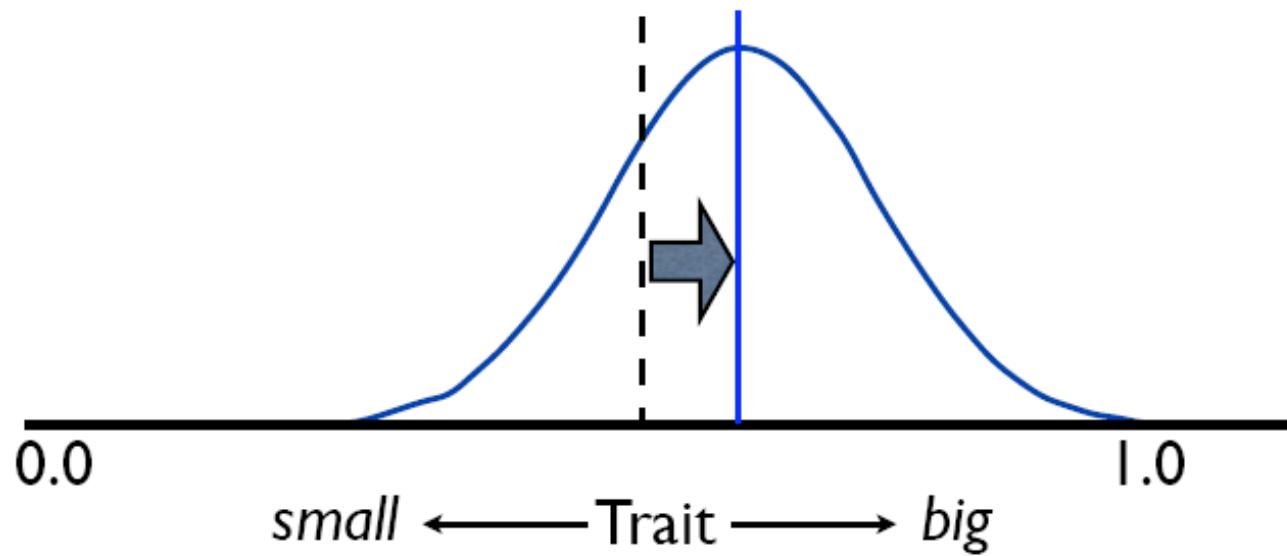
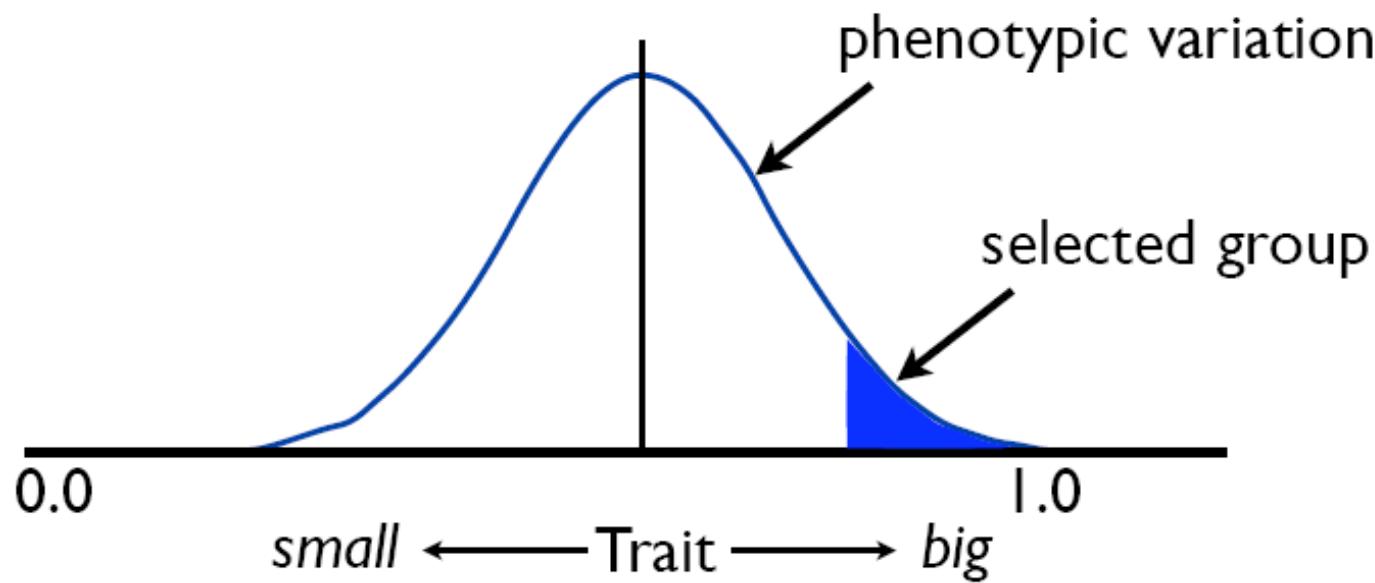
- Selection can lead to large divergence within short (by human standards) duration.
 - Divergence depends on variation, selection and number of generations.

Take Home Messages:

- Trade-offs are ubiquitous.
- Evolution does not “Plan Ahead”- it is “local”.
- Inferring selection and adaptive evolution from existing patterns is tricky.

Darwin's Big Idea

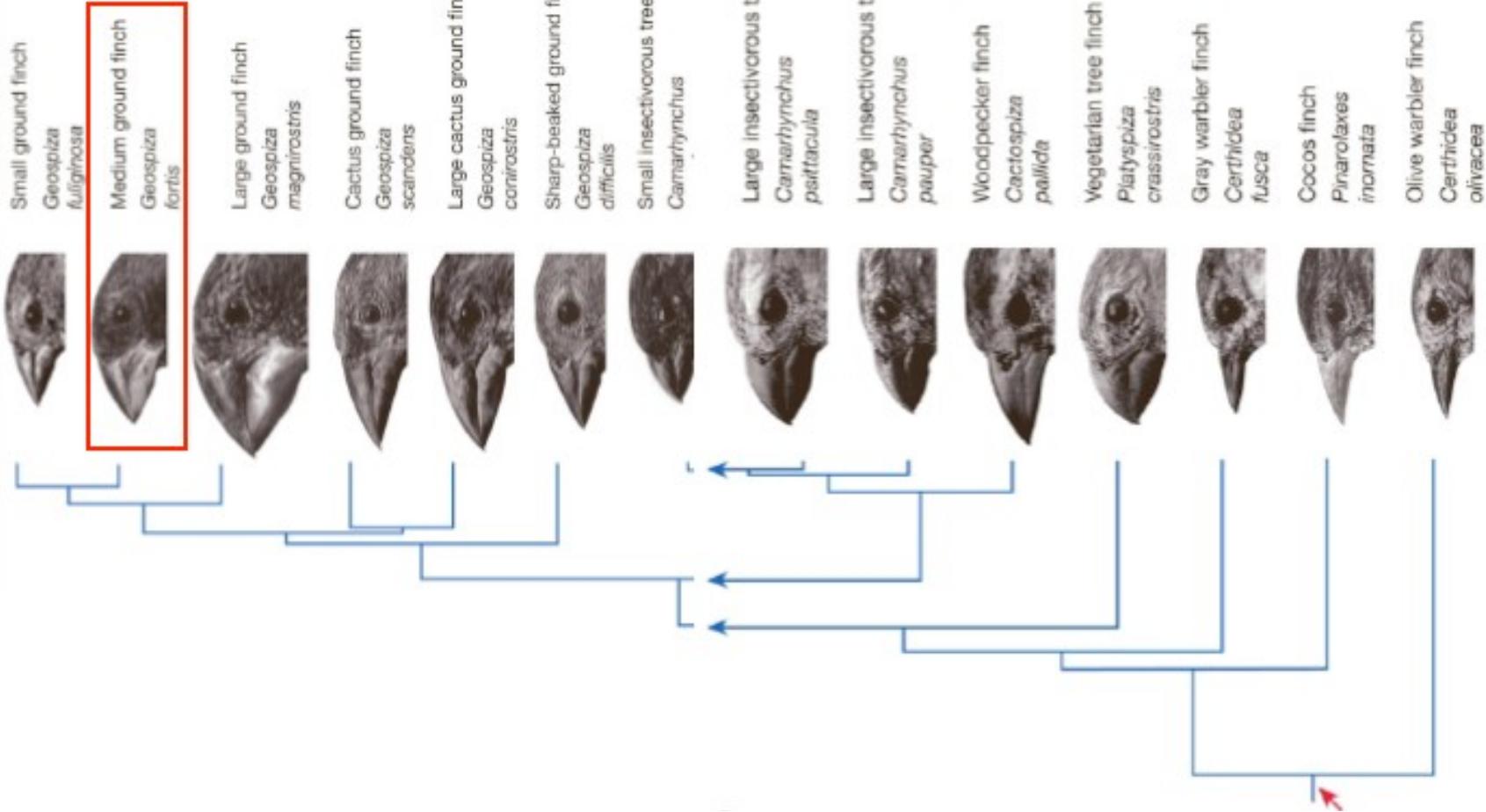
1. There is variation within breeding groups (*populations*).
2. Part of the observed variation is *heritable*.
3. Some individuals are more successful in survival & reproduction (*fitness*) than others.
4. Variation influences fitness: those with the most favourable *phenotypes* are *naturally selected*.



The famous finches

- 3 decades of research by Peter and Rosemary Grant and colleagues
- Galapagos is a “natural laboratory”
 - Replica populations relatively isolated on different islands.
 - Smaller islands = easy to census.
 - Survival, reproductive success, morphology can be measured on site.

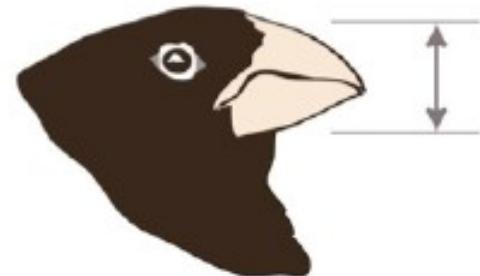
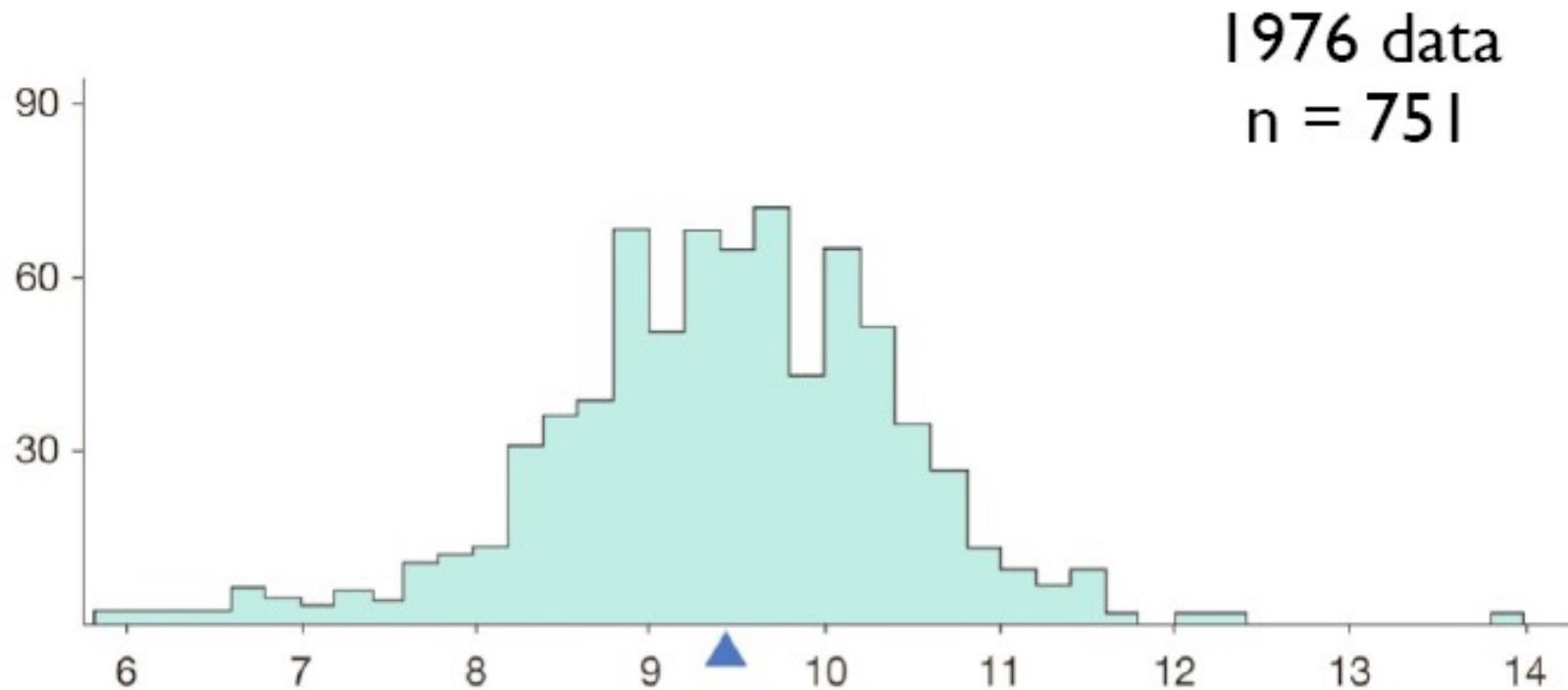
The Darwin's Finches



Prerequisites for selection

- **Genetic variation** in the population for a trait related to fitness.
Beak size important to finches in foraging
- **Agent of selection**
 - seed abundance and type believed to select for beak morphology
 - seeds available influenced by season, rainfall etc.

Phenotypic Variation: *G. fortis* beak depth



Is this variation heritable?

Approaches

Similarity between relatives

- Parents and offspring
- Siblings

Response to selection

selection will only change the population mean if the variation is heritable

Heritability

Broadsense Heritability: Proportion of total phenotypic variance that is explained by genetic variance.

$$V_P = V_G + V_E + 2\text{Cov}(GxE) \quad (\text{setting } 2\text{Cov}(GxE) = 0),$$

$$V_P = V_G + V_E$$

$$H^2 = V_G / V_P$$

$$H^2 = V_G / (V_G + V_E)$$

Heritability

Narrow-sense Heritability: Proportion of total phenotypic variance that is explained by additive genetic variance

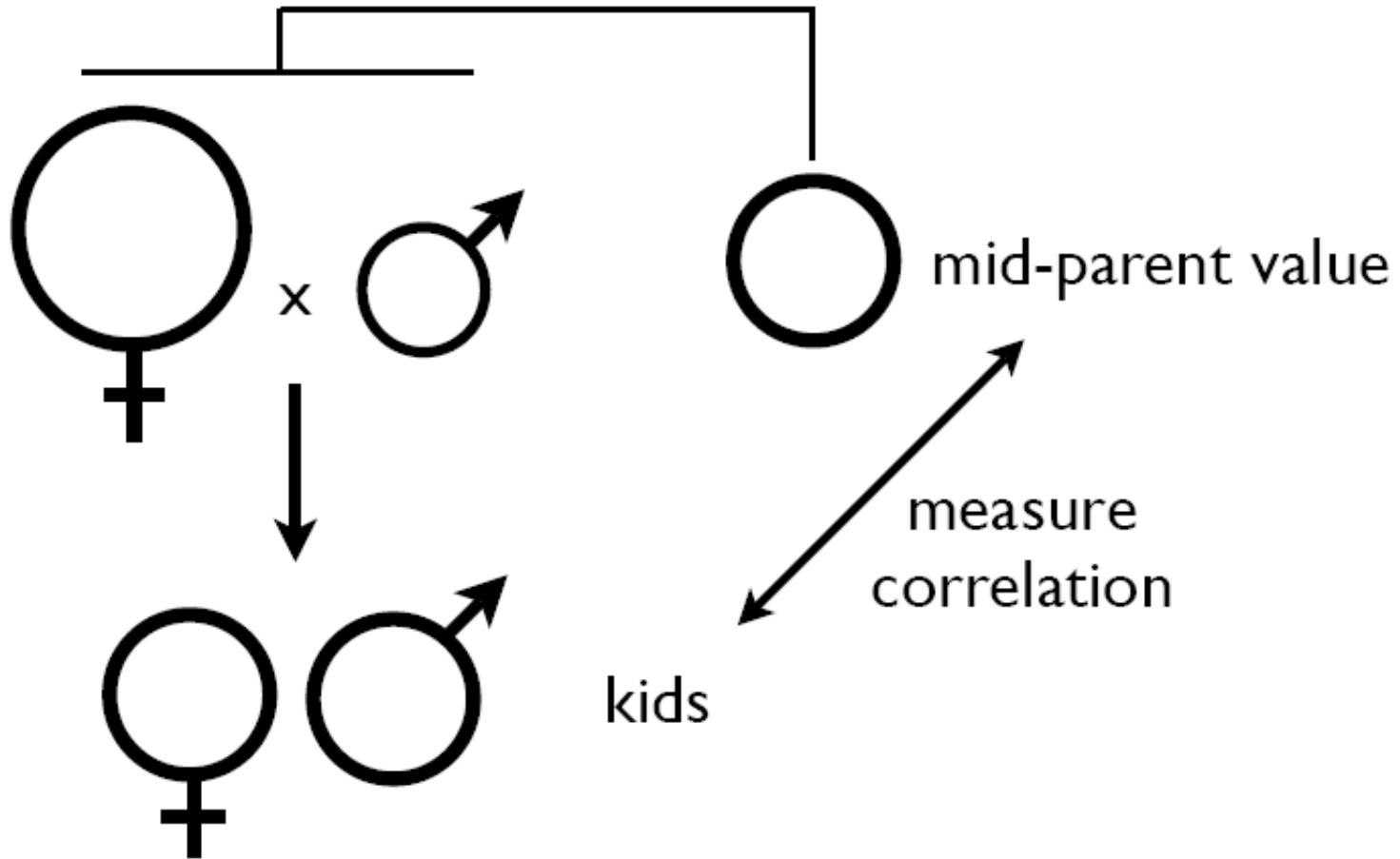
$$V_G = V_A + V_D + V_I$$

$$V_P = V_A + V_D + V_I + V_E \quad (\text{setting } 2\text{Cov}(GxE) = 0),$$

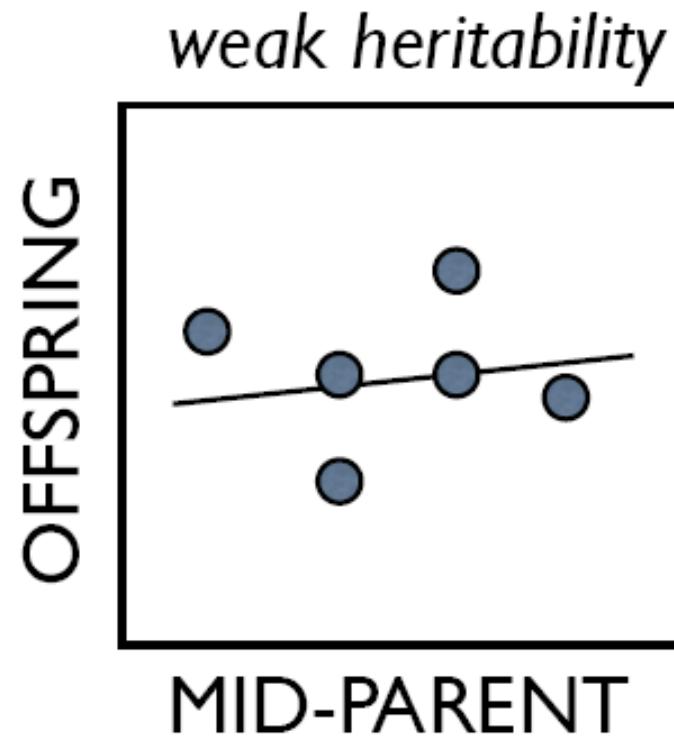
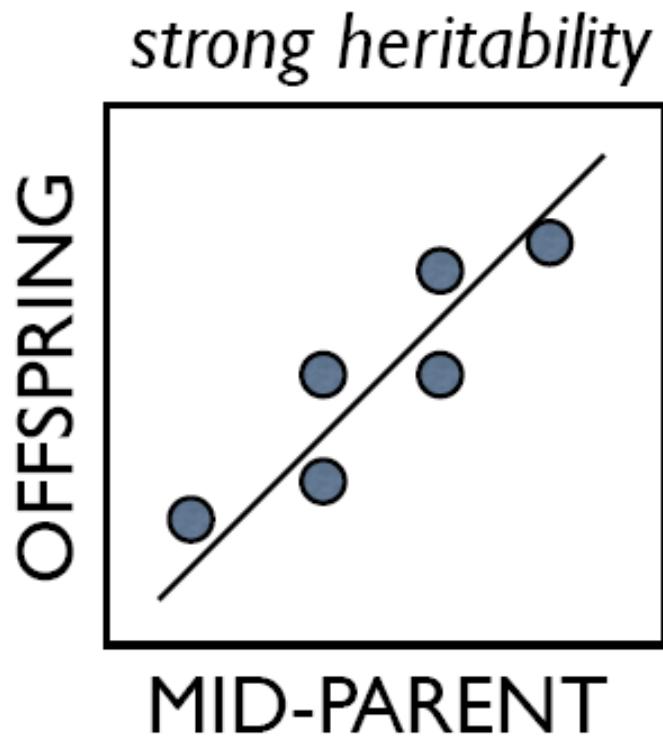
$$h^2 = VA/VP$$

$$h^2 = V_A / (V_A + V_D + V_I + V_E)$$

Basic QG

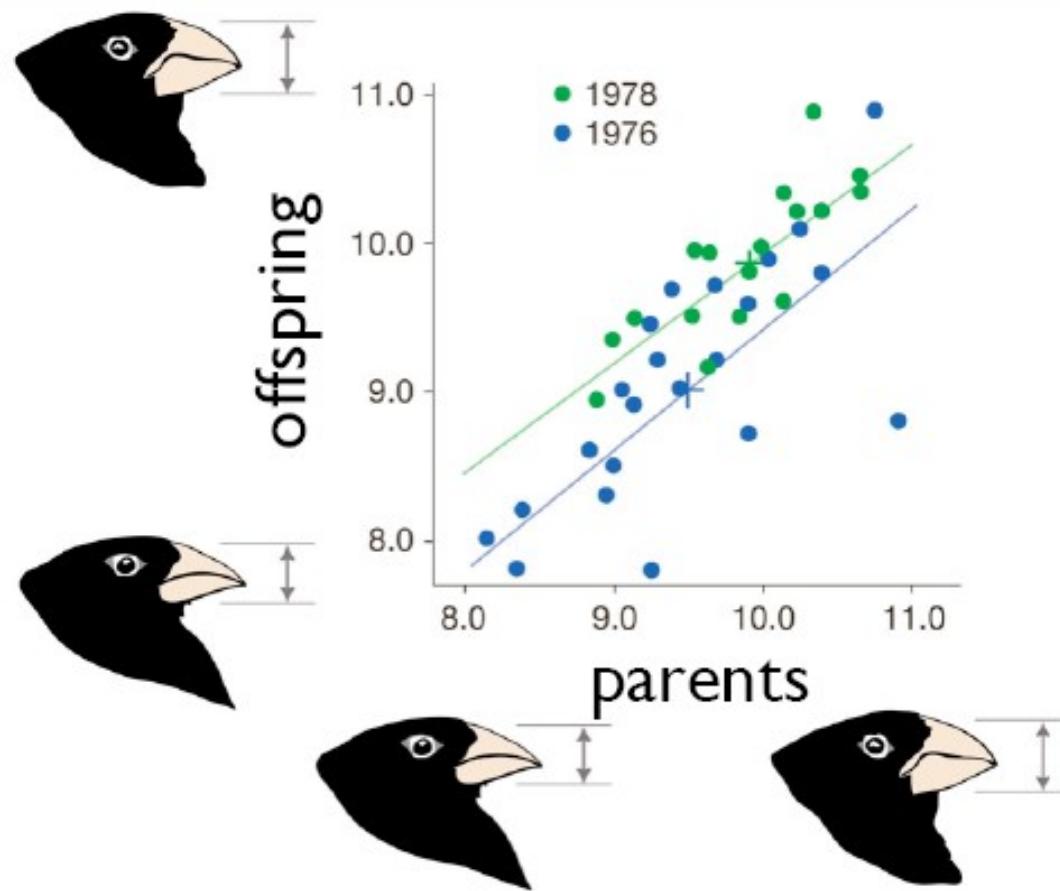


The correlation between parents and offspring is a measure of heritability



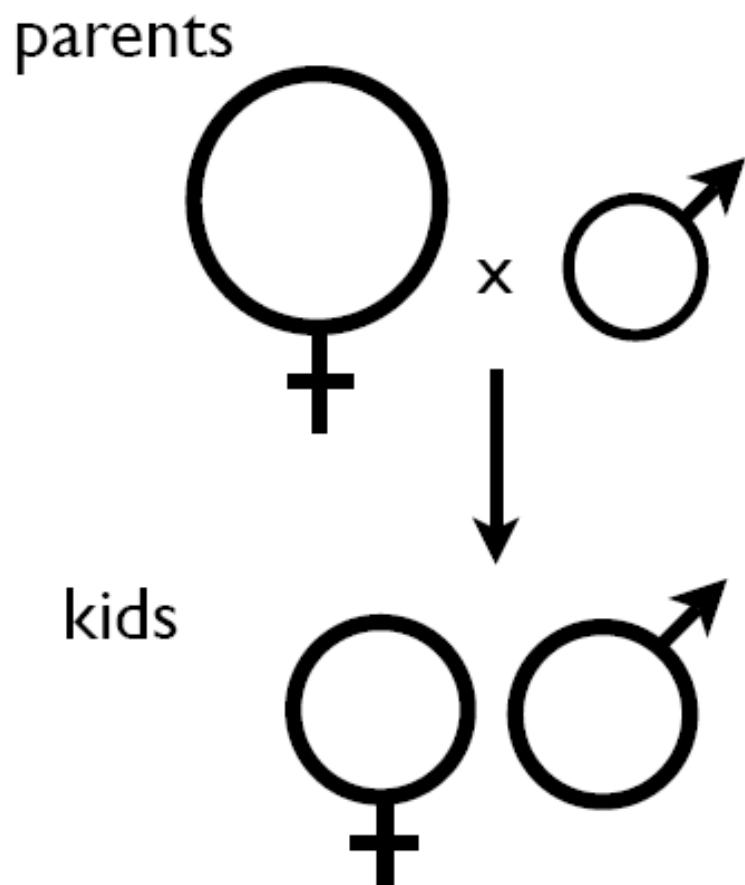
human height $h^2 = 0.6$

back to finches...



from Boag 1983

Problems with heritability



1) kids have same mother

- maternal effects conflated with heritability

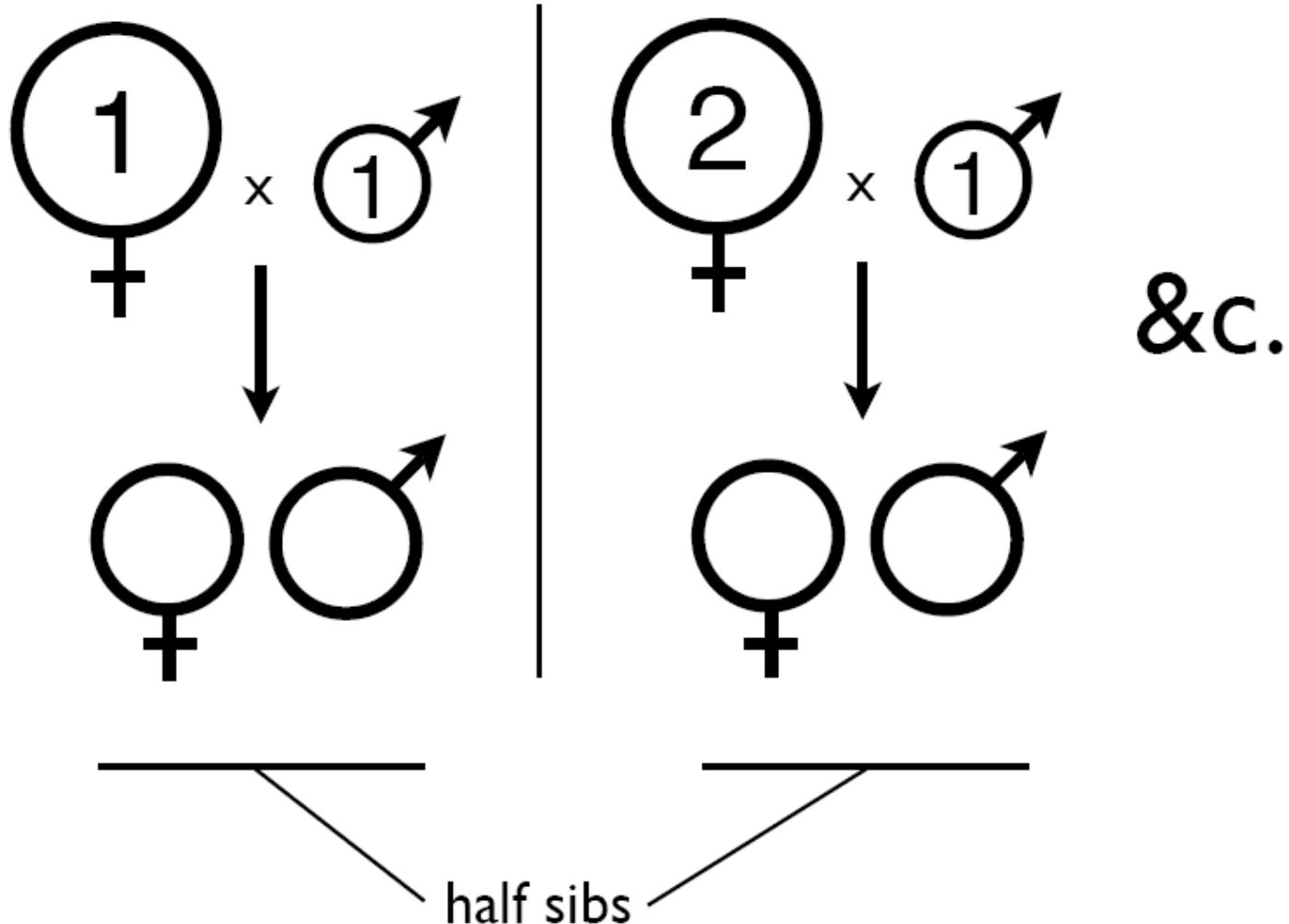
2) parents & kids reared in same environment

- $V_{\text{environmental}}$ conflated with V_{genetic}

3) Paternity uncertain

- extra-pair copulation reduces estimate of heritability.

Half Sib Analysis



Some heritability estimates

- Morphological traits = 0.51
- Behavioural traits = 0.37
- Physiological traits = 0.31
- Life-history traits = 0.27

The closer a trait is related to fitness, the lower the Vg we expect.

Life-history traits (fertility, survival etc.) are closely linked to fitness

Postulate 3:

Do individuals vary in
survivorship and reproduction?

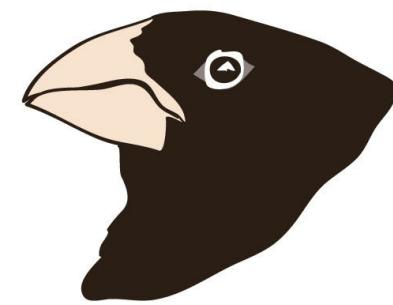
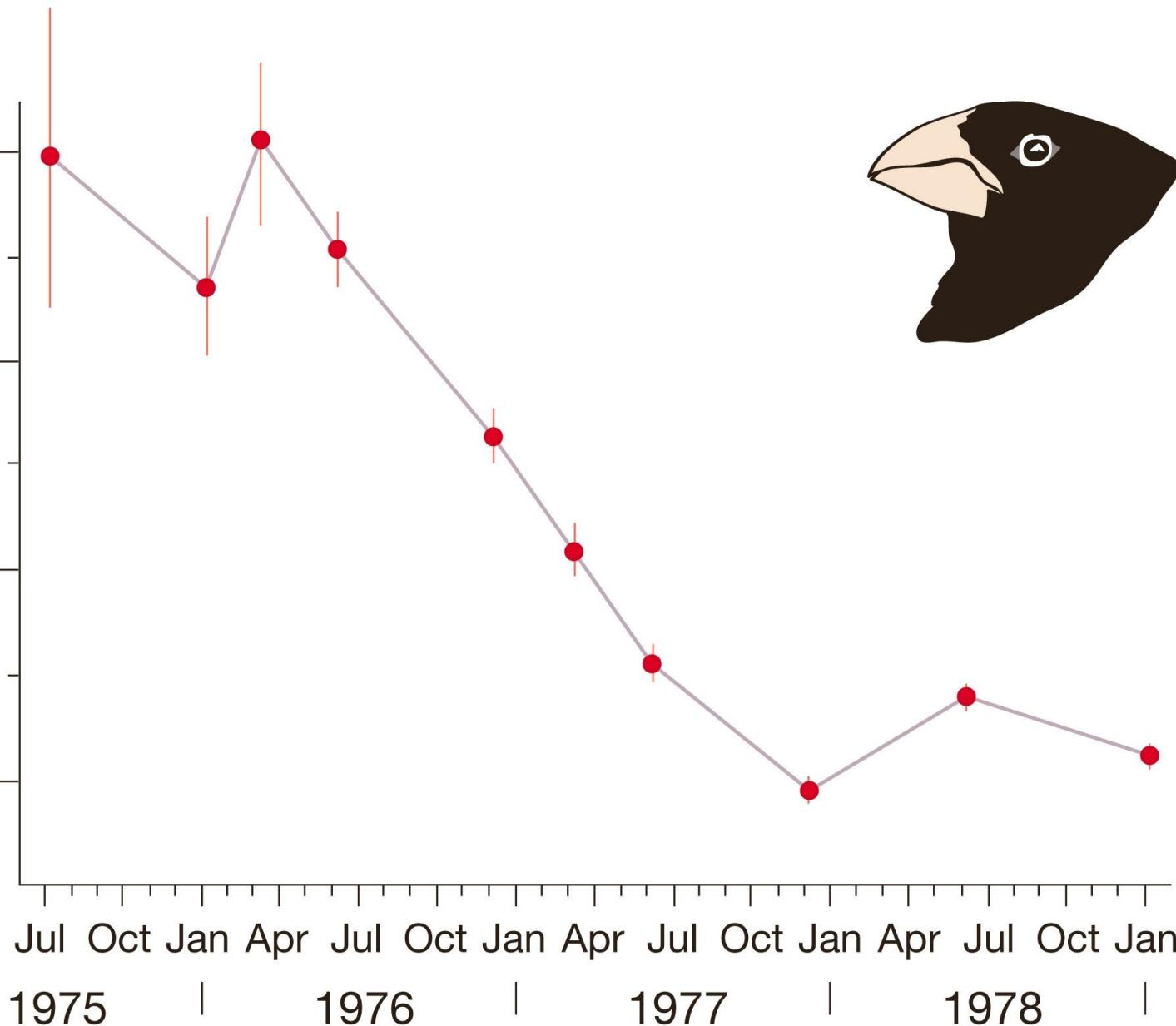


- Study on Daphne Major in 1977
- It was a major year of drought
 - <1/5 normal rainfall
- Plants failed to flower
- Seed production declined
- Birds died *en masse*

(a)

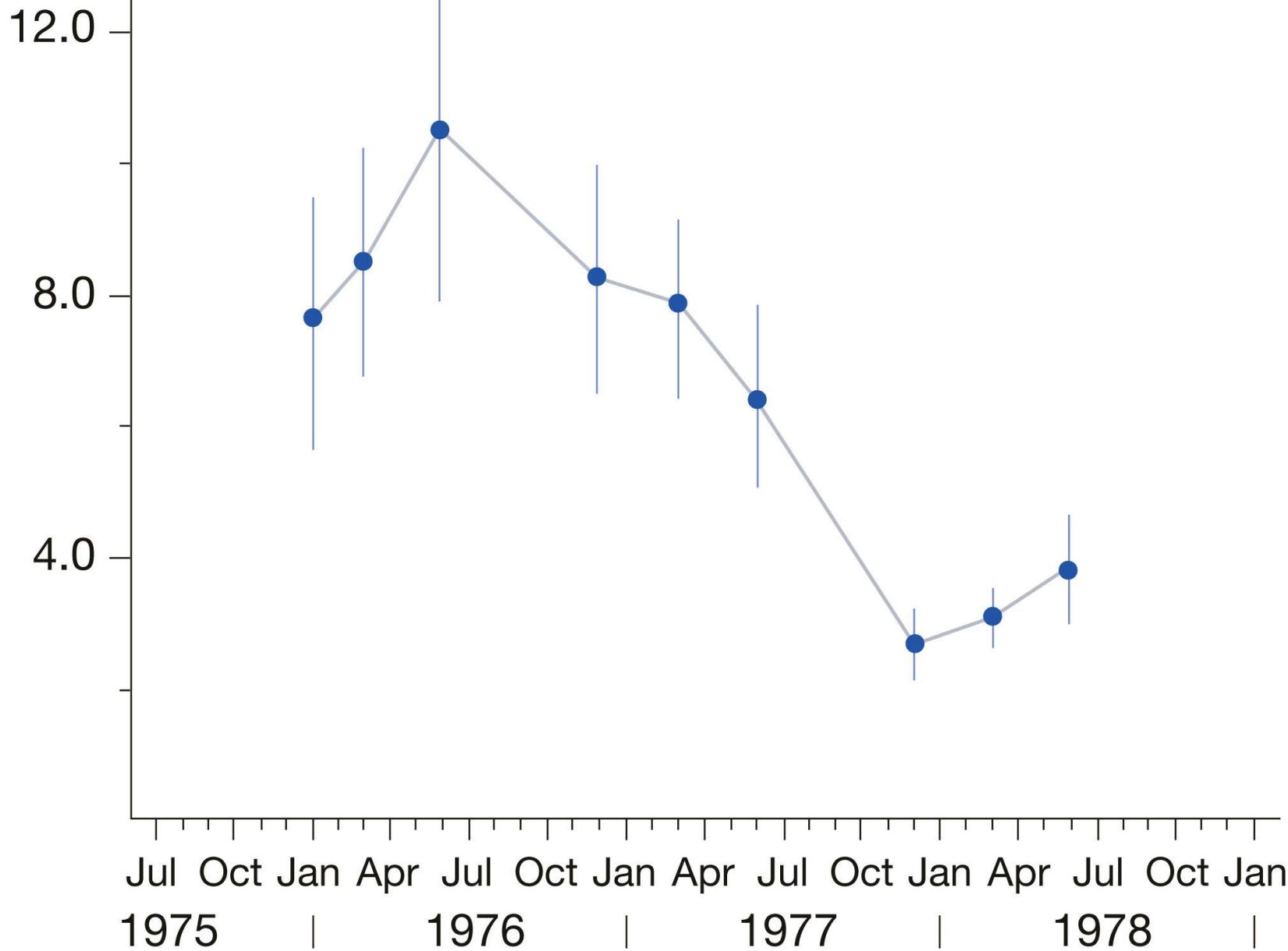
Number of finches

1400
1000
600
200



Copyright © 2004 Pearson Prentice Hall, Inc.

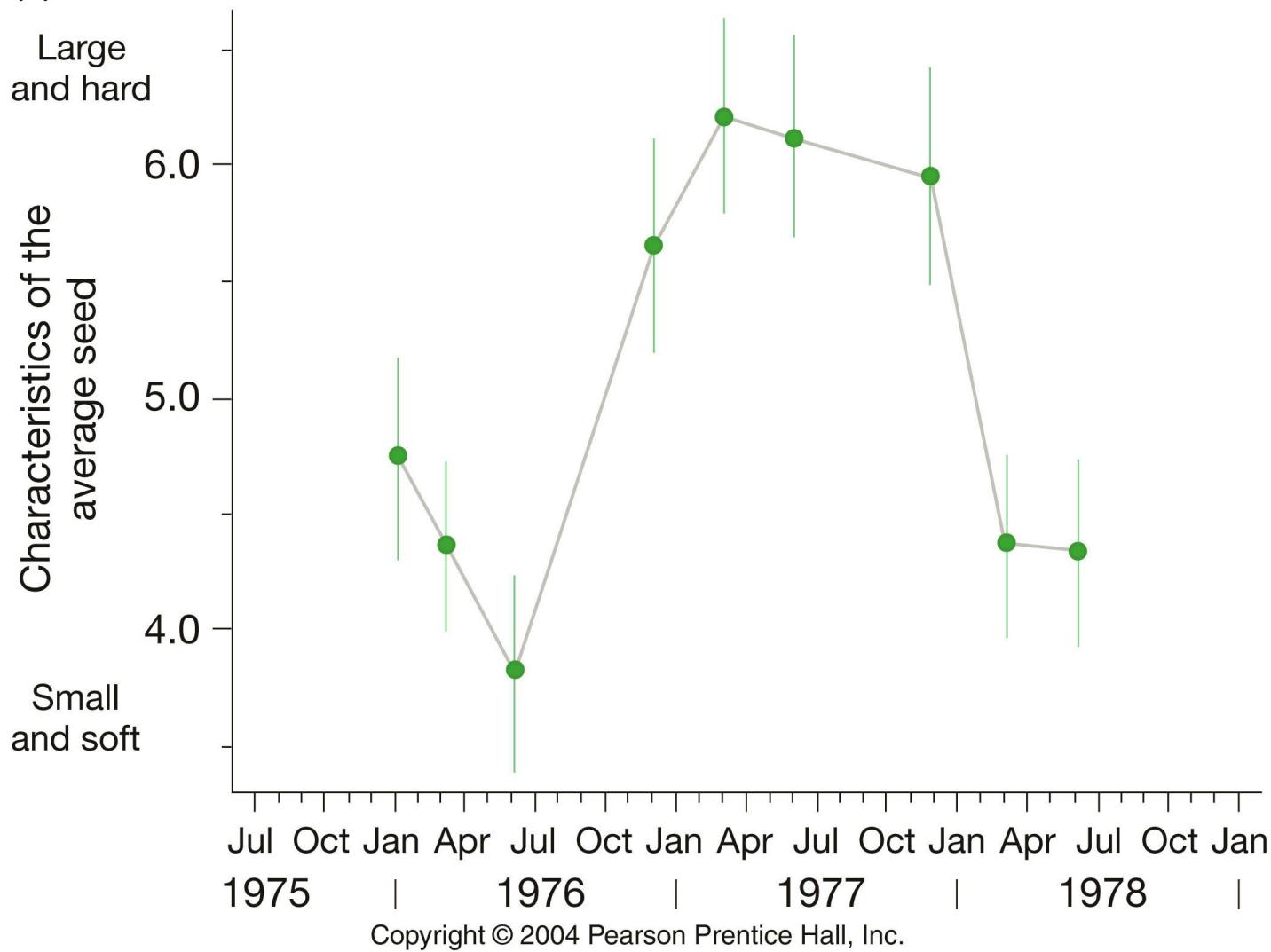
(b)

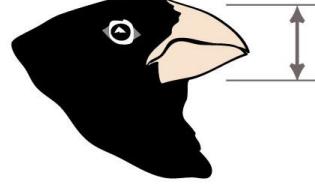
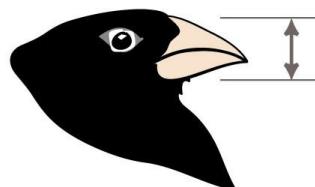
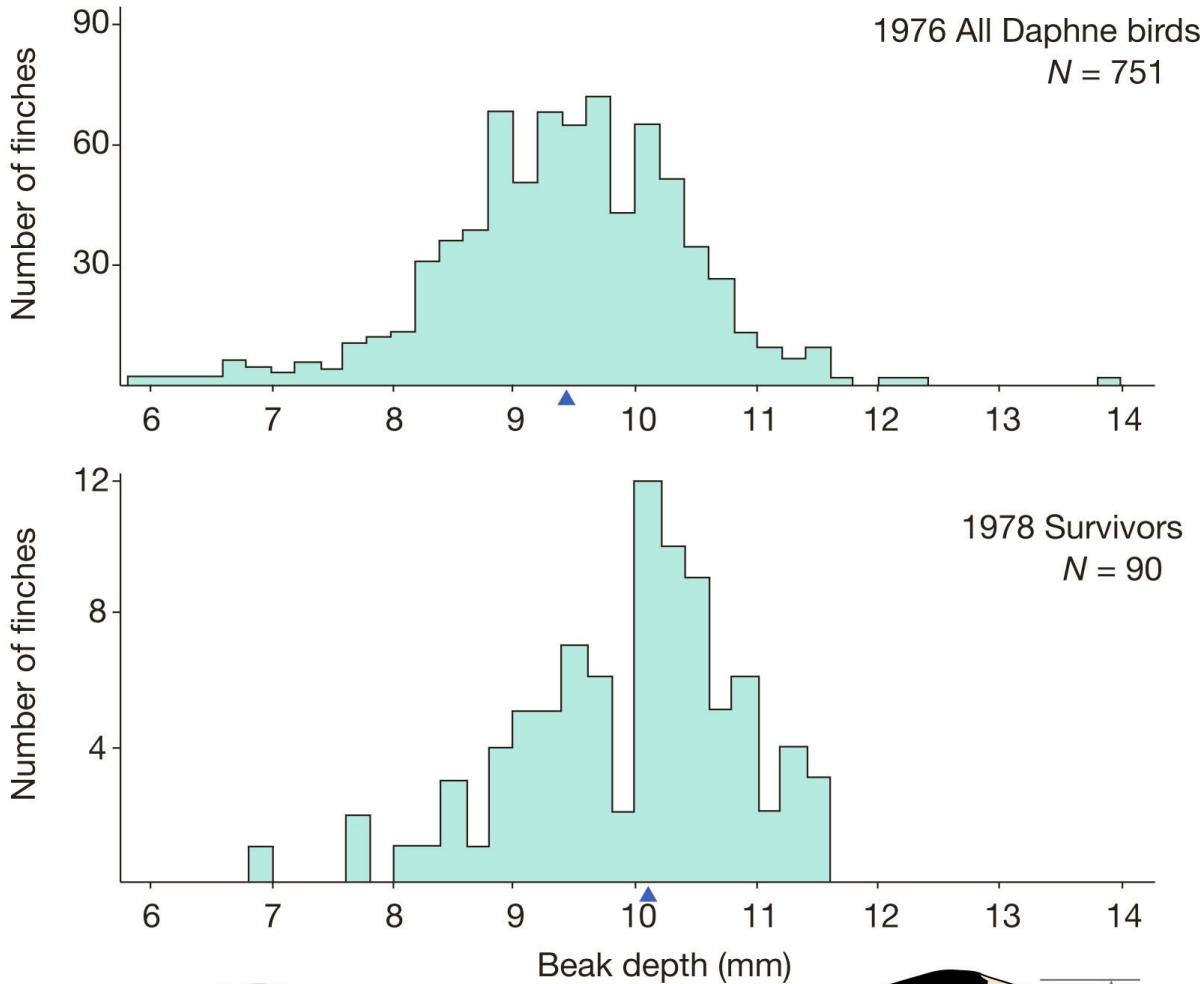
Abundance of seeds (g/m^2)

Postulate 4:

Is survivorship and reproduction
non-random?

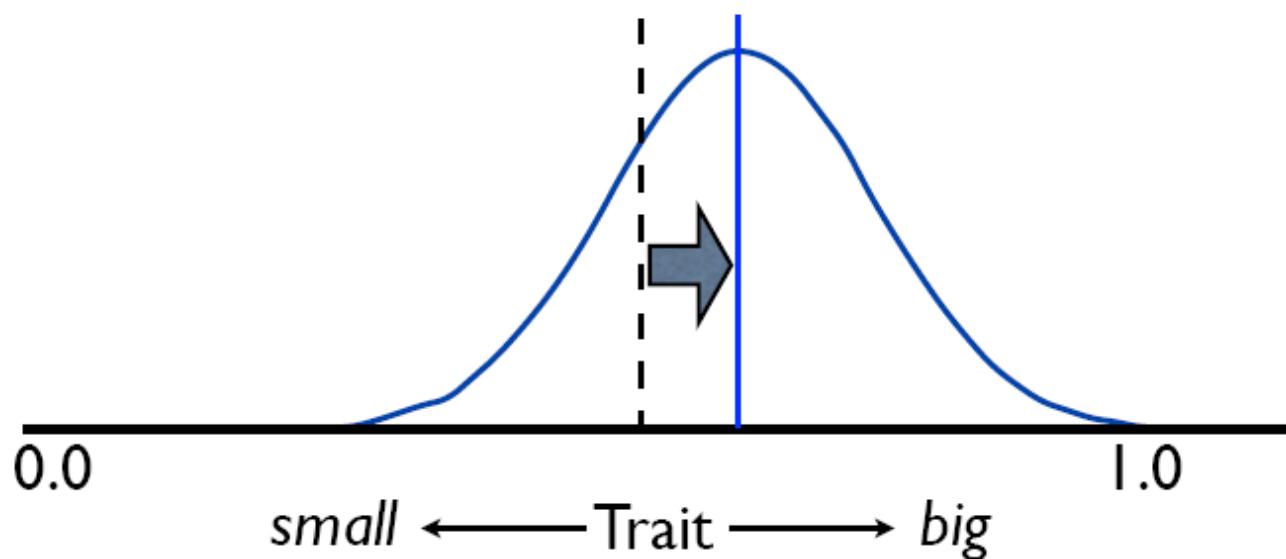
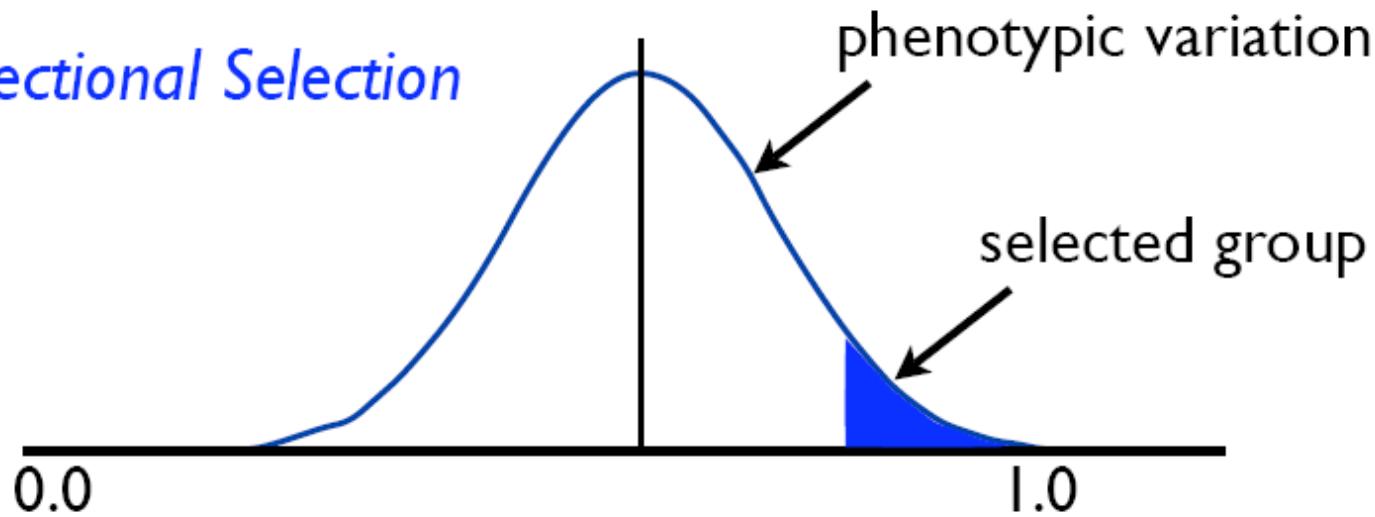
(c)



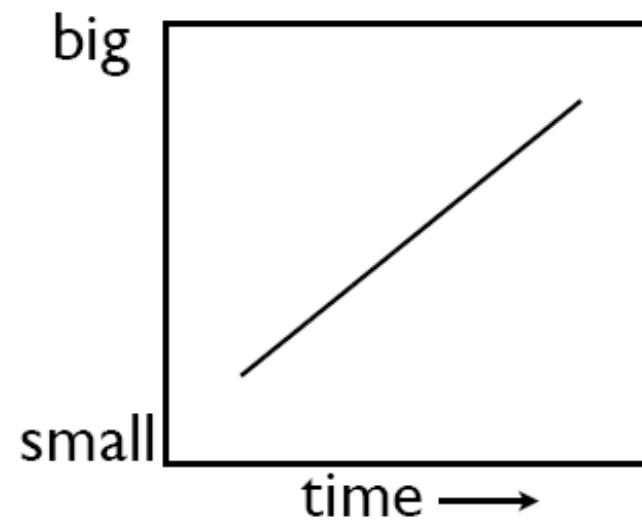
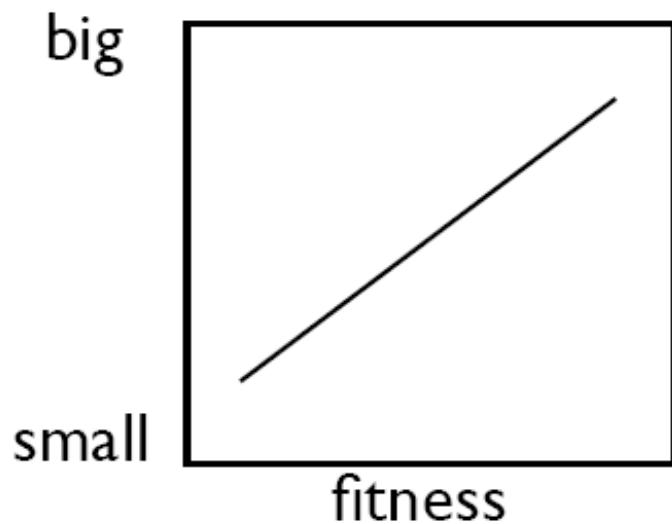
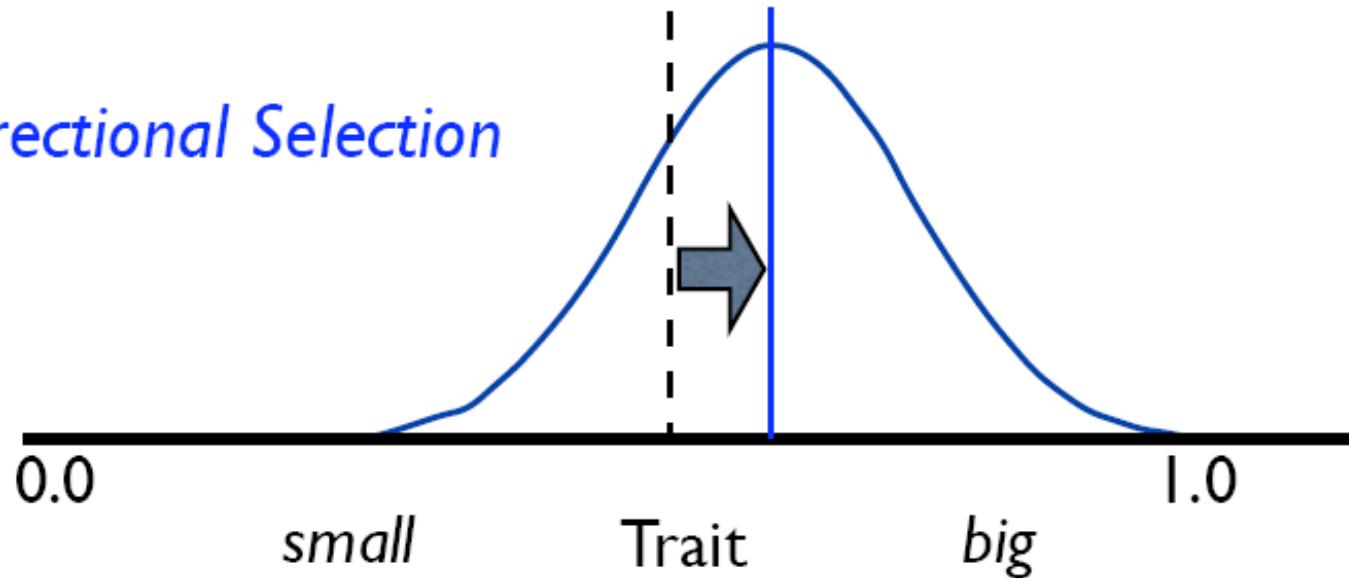


Copyright © 2004 Pearson Prentice Hall, Inc.

Directional Selection



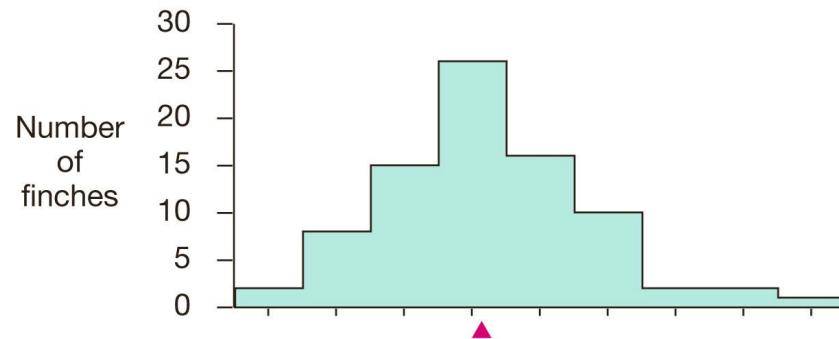
Directional Selection



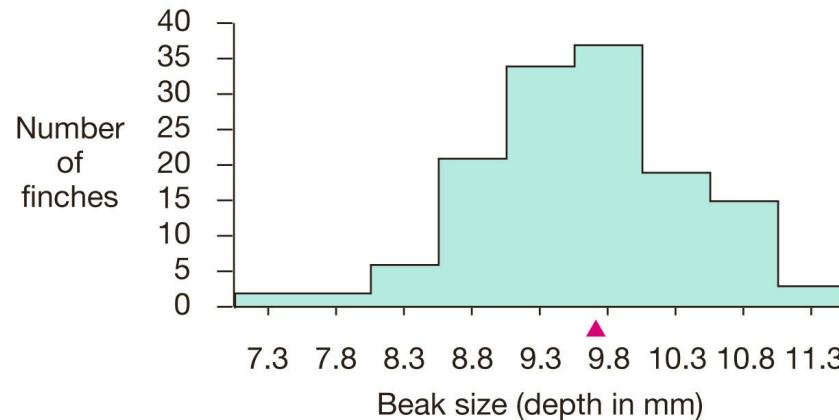
Therefore,

Did the finch population Evolve?

Finches hatched in 1976, the year before the drought



Finches hatched in 1978, the year after the drought



Copyright © 2004 Pearson Prentice Hall, Inc.



Bigger was better....

In 1977

1. Better access to large, hard seeds.
2. More calories to resist starvation?
3. Lower surface area / volume ratio = greater drought resistance?

So will the beak size increase
always?

(Or

“Can one become as tall as they
want?”)

Bigger isn't always better

- May be poorer at manipulating smaller objects.
- More energetically expensive to produce.
 - Slower growth? Less energy for other parts?
- Affects vocal communication, other aspects of mate choice.
- Genetically correlated with other traits also under selection?

Variation begets variation

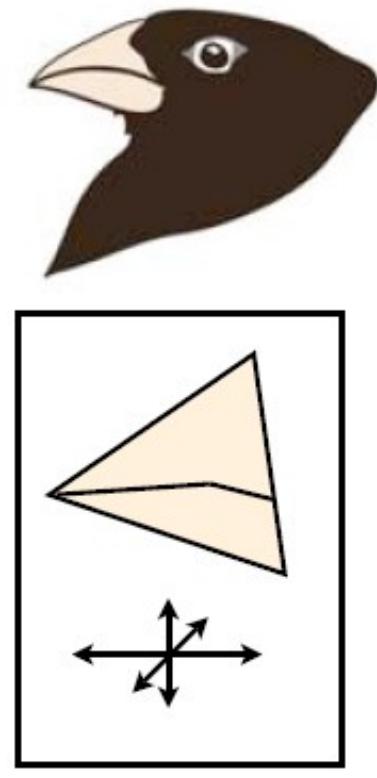
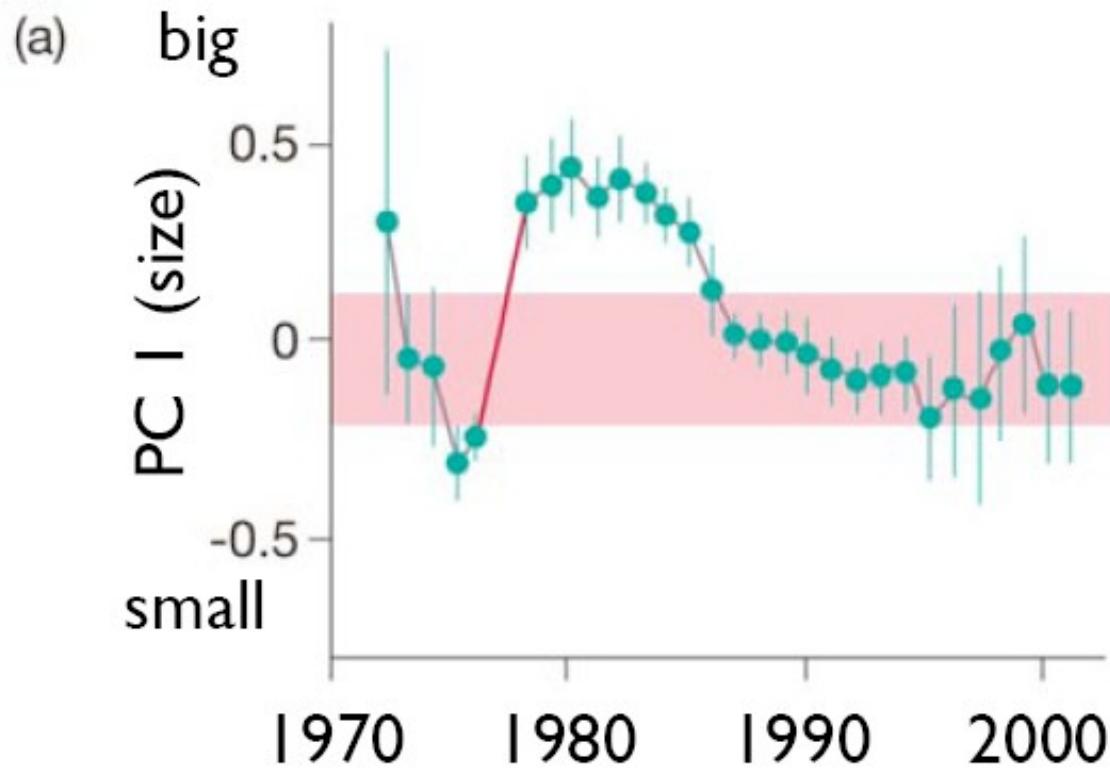
- Major shift of 1977 towards big, deep beaks
- Moist years produce more soft small seeds
 - Selection favours small beaks under these conditions
 - ‘El Nino’ (very wet) years reveal strong selection towards smaller beaks and bodies.



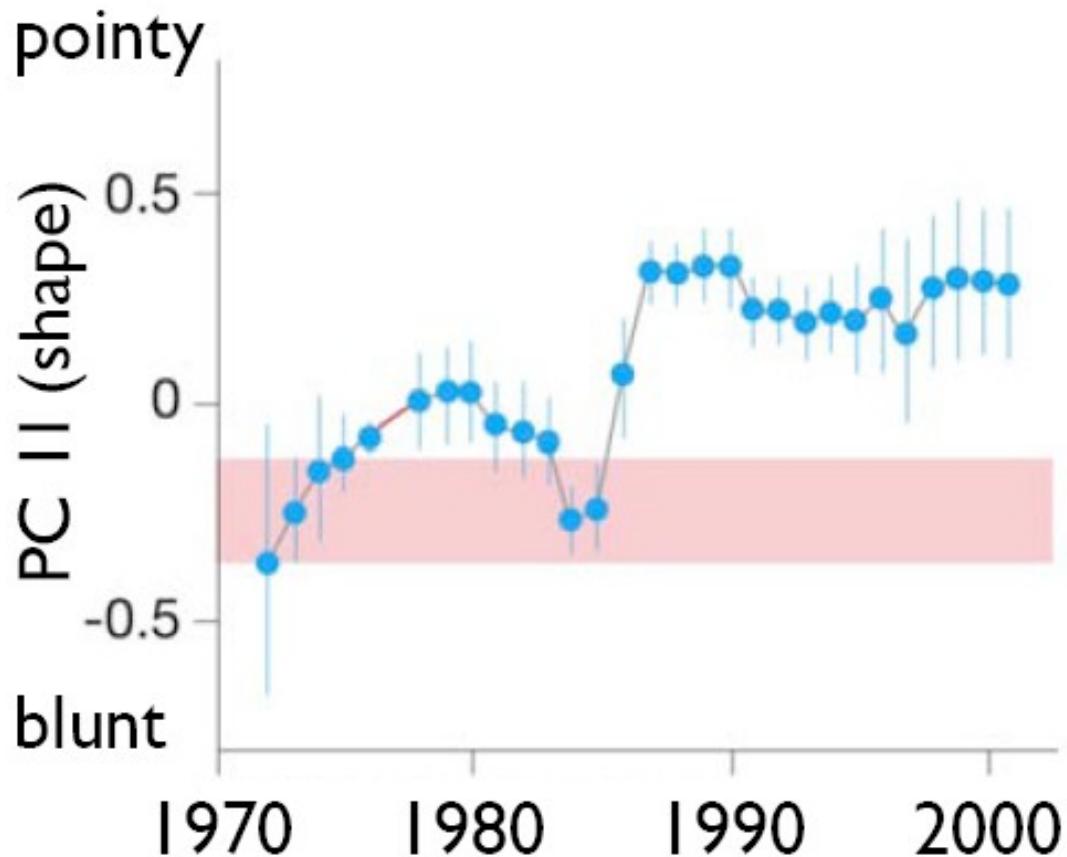
Tribulus trouble

- Opening *Tribulus* seeds requires good downward force and twist
- Deep beaks provide force but narrow ones are good at twist action
- Beak depth and width are positively **genetically correlated**
- This places an **evolutionary constraint** on adaptation

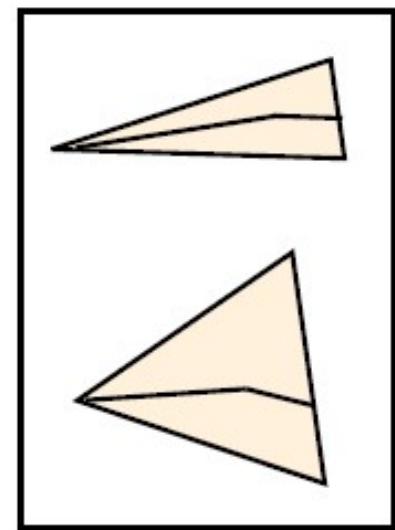
3 decades of change



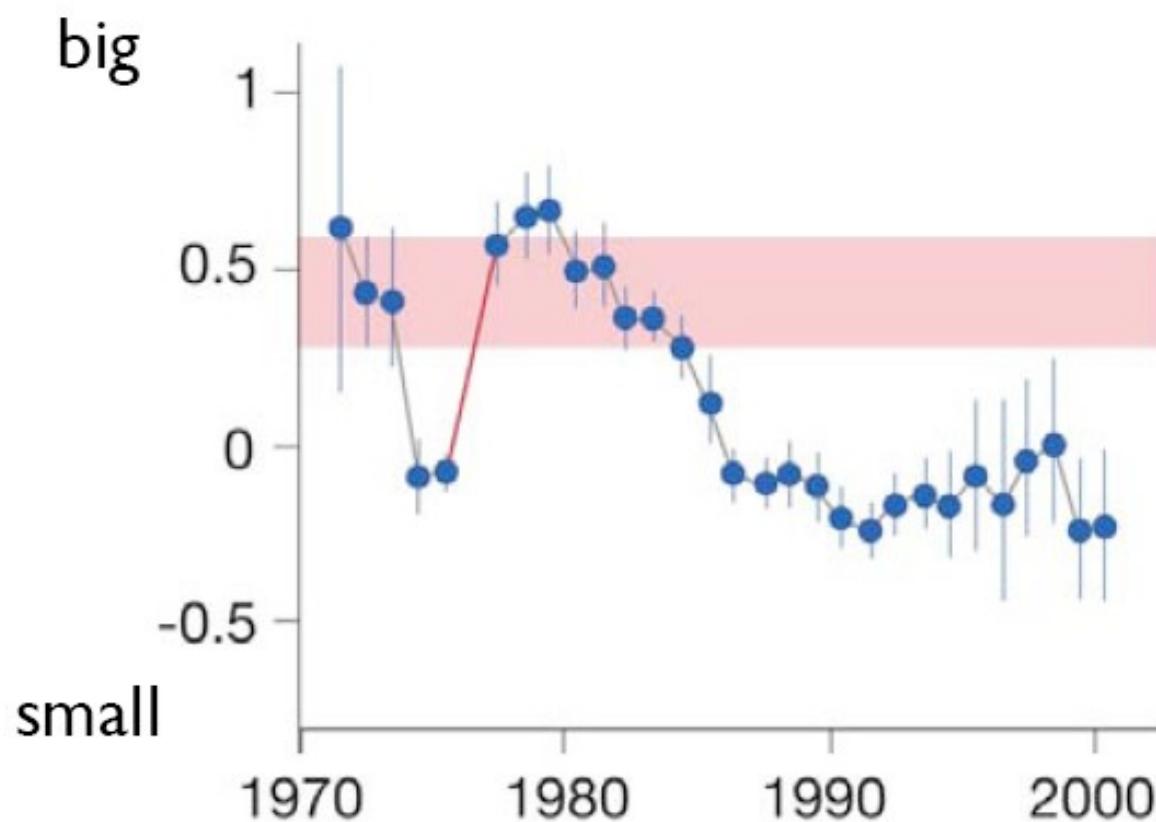
Beak Shape



PC 2



Body Size



(Daphne) Major lessons

1. Evolution can occur very rapidly and dramatically.
2. Sources of natural selection often shift, so evolution reverses itself or takes new directions.
3. The see-saw of selection in changing environments may give the impression of stately gradual change.
 - Typically the fossil record does not have the resolving power to see zig-zaggy evolution.

The Nature of Selection

- Natural Selection is a population level process
 - We define selection by changes in allele frequency.
 - *Individuals* vary in survival and reproduction
 - When the genes that they carry contribute to differences in fitness, then, evolution occurs.
- Selection acts on the *phenotype of individuals* but evolution can only happen when it acts upon *heritable variation*.

Selection is always behind

- NS changes gene frequencies in populations *after* exposure to selection.
 - Avg. size of *G. fortis* beaks changed because of changes in allelic frequencies in the population.
 - Allele frequencies shifted because big beaks were good *that* year.
- Selection cannot anticipate future changes in the environment.
 - If the environment changes too rapidly, population extinction may result.

Selection is not “Progressive” (as in - it is good for the species)

- Although organisms have become more specialised and complex, this does not imply some ultimate goal or direction.
 - Organisms ‘devolve’ as often as they evolve.
 - Parasitism is a common way of life.
- Adaptation occurs through character gain, character loss and character modification.

Alternative views of Evolution and “Progress”

- Stephan Jay Gould
 - Defined progress as increase in complexity
 - Evolution is not progressive
 - Unicellular organisms dominate
 - A “left wall of complexity”
 - Hence “Progress” is a mere increase in variance.
 - **Must Read- “Life's Grandeur or Full House”**

Alternative views of Evolution and “Progress”

- Richard Dawkins
 - Defined progress as increase in adaptedness
 - Evolution is progressive (well..to a limited degree)

Richard Dawkins contd...

- Over short term- organisms adapt to an environment
- Over medium term- Hosts and pathogens coevolve
- Over longer term- new innovations-cells, sex etc.
- **Must Read- “Selfish Gene”, “Blind Watch Maker”**

Selection and ‘Found Objects’

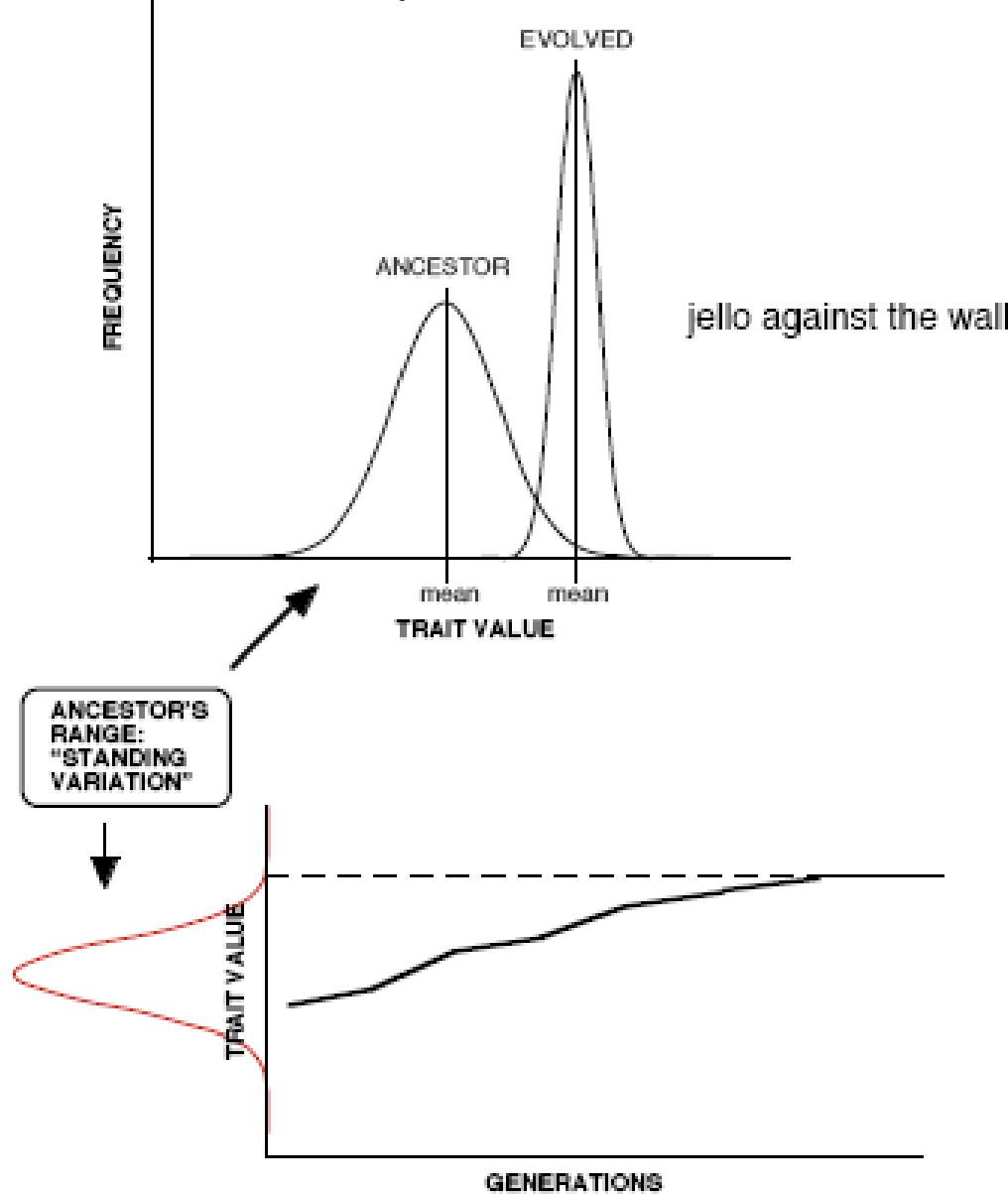
- Selection operates on preexisting variation.
 - Often the tools at hand are less than perfect solutions to the problem.
 - Characters that are coopted for new purposes are called *preadaptations*.
 - Evolution is not forward-thinking; new mutations do not arise to solve problems.

Fitness is environment-specific

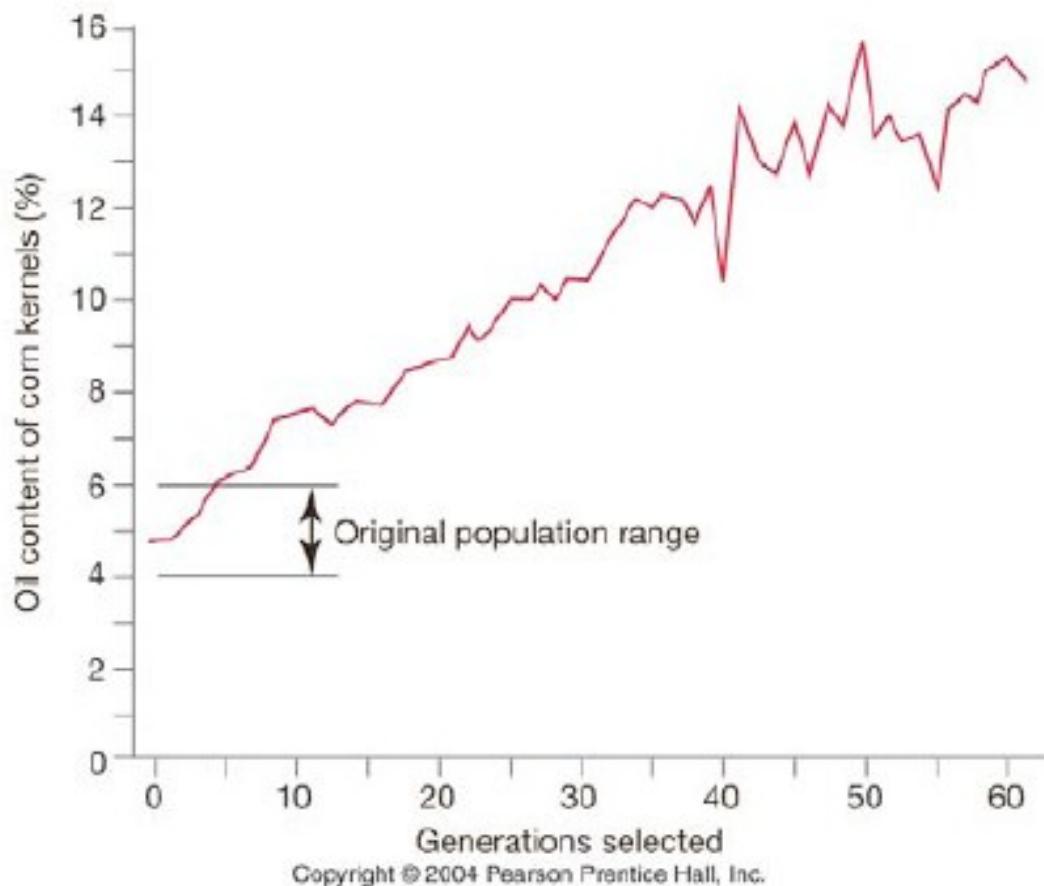
- Evolved changes often make an organism more specialised.
 - Specialized organisms are more susceptible to environmental changes.
 - Environments do not always change randomly.
 - Biotic environments often themselves evolve, often in competition or as enemies of a species.
 - Van Valen called the evolution / counter-evolution between species the Red Queen principle.



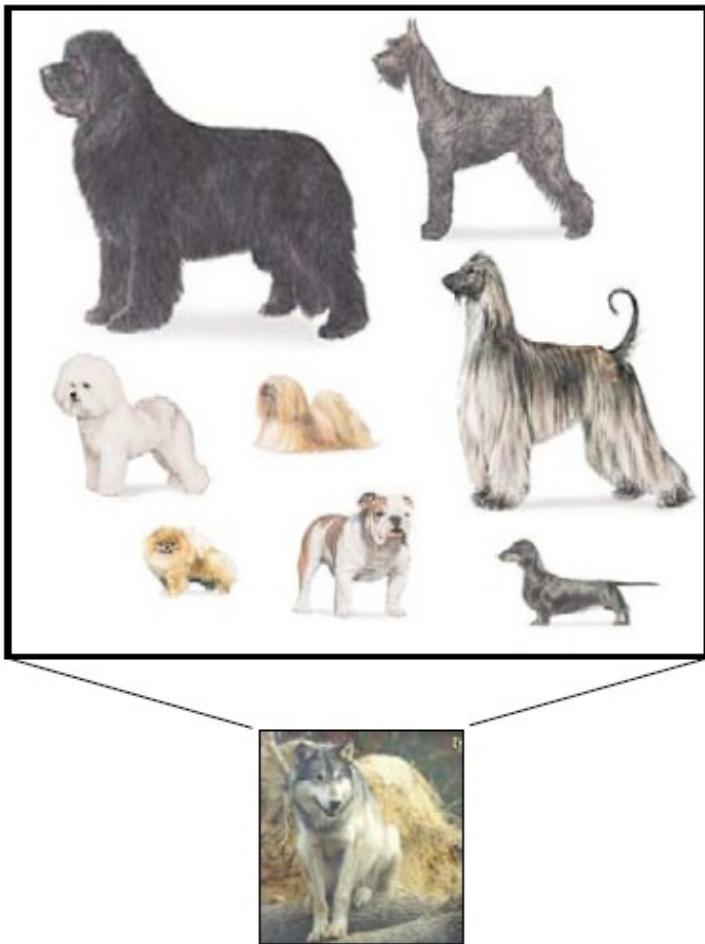
The limits of selection: The relationship between variation and selection



Sustained selection

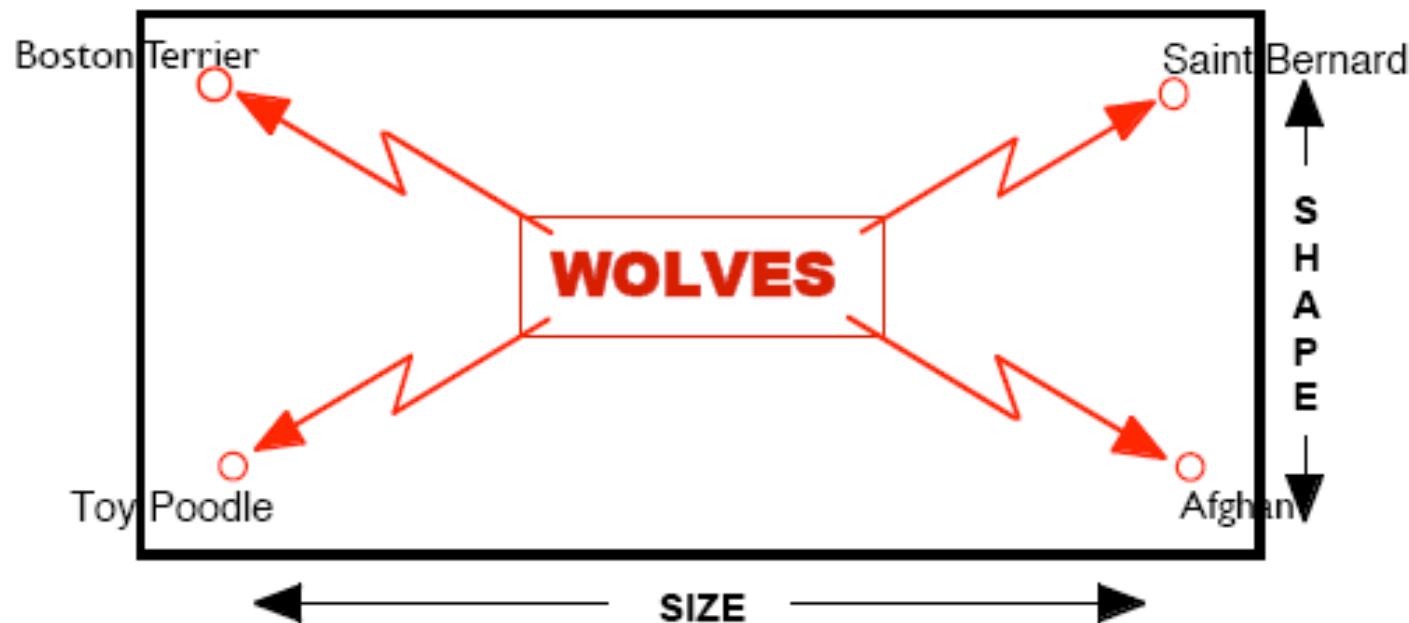


Pet Store Evolution

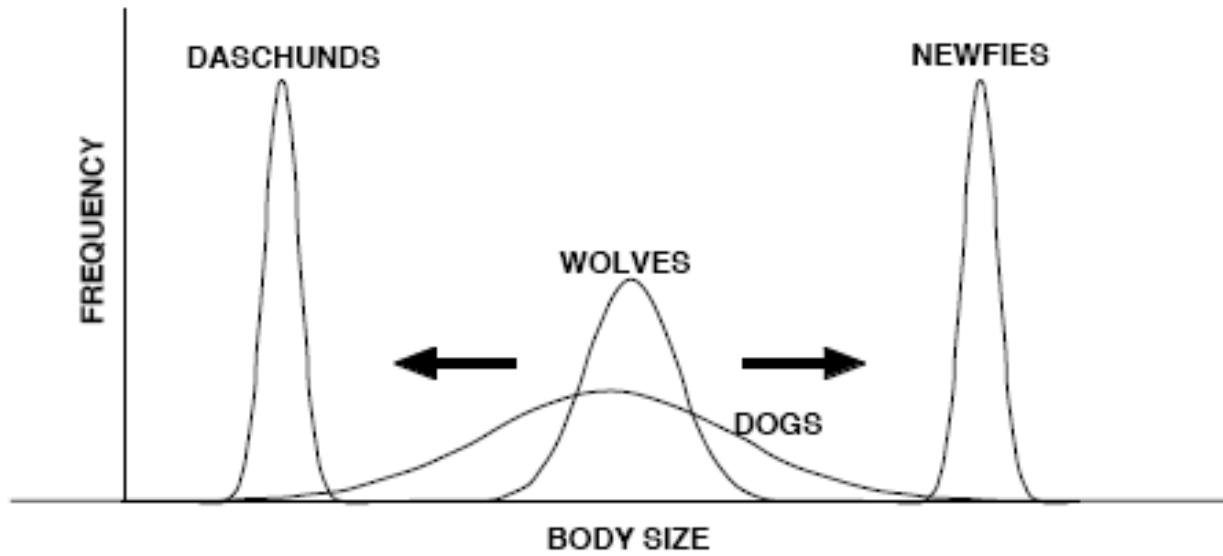


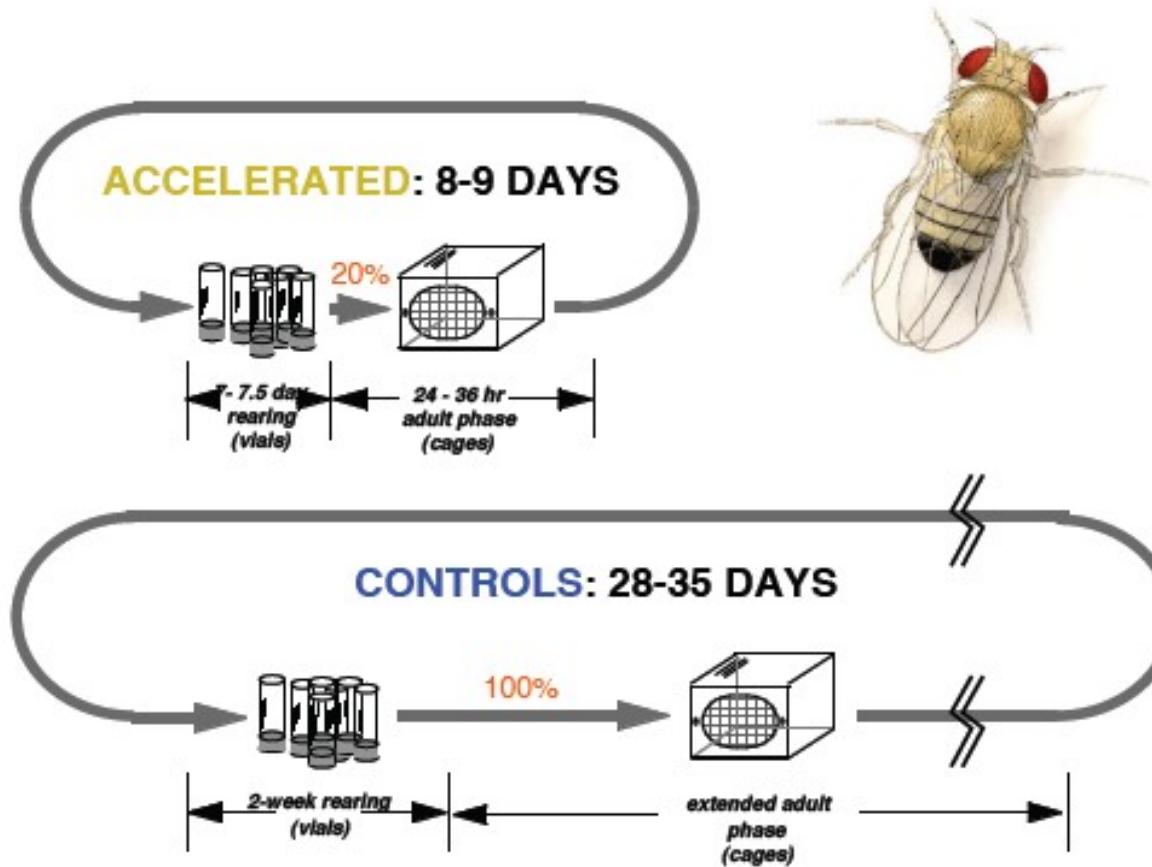
All domestic dog breeds derived from a single a few domestication events @ 10KYA from the Asian grey wolf.

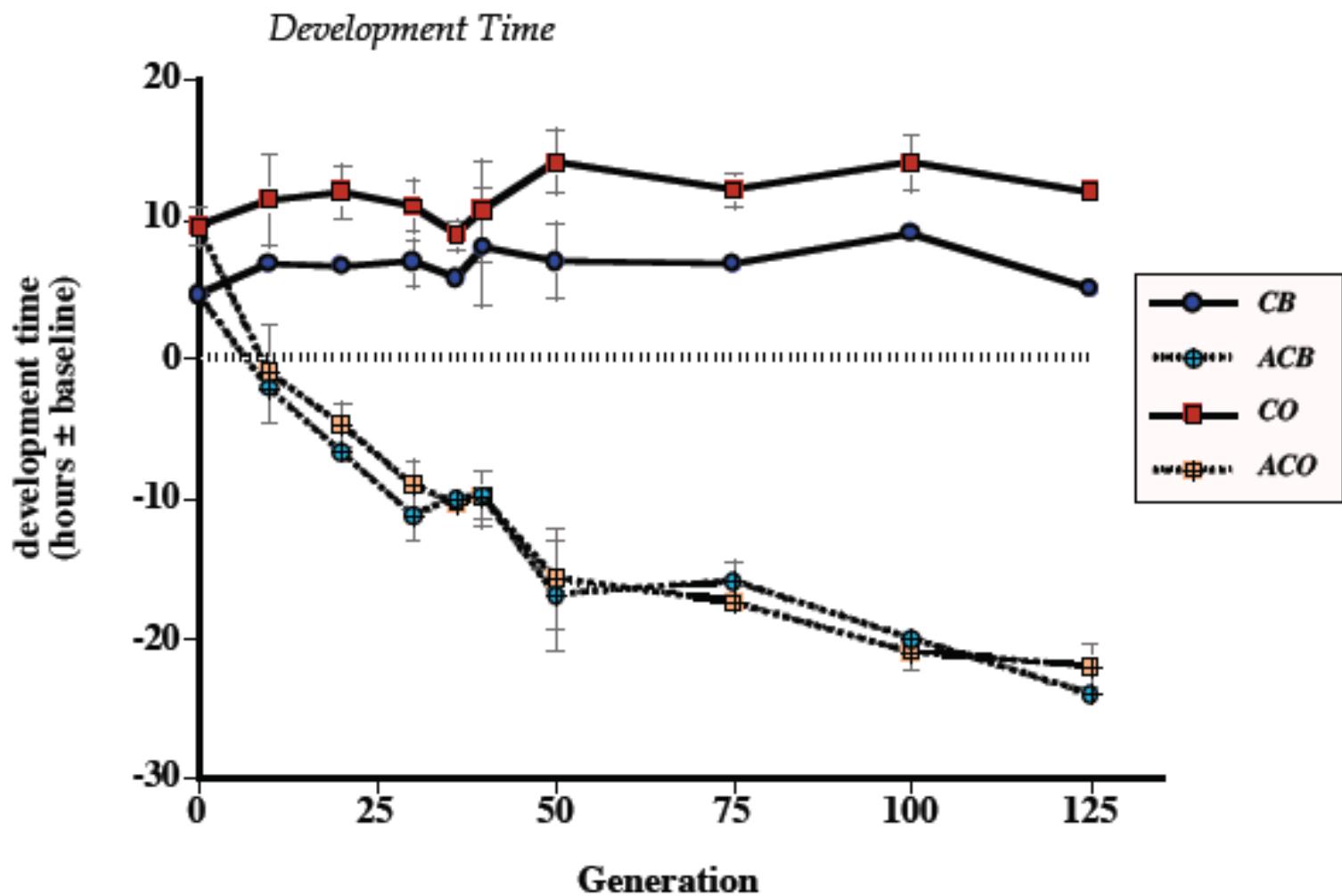
Expansion of Grey Wolf Morphospace



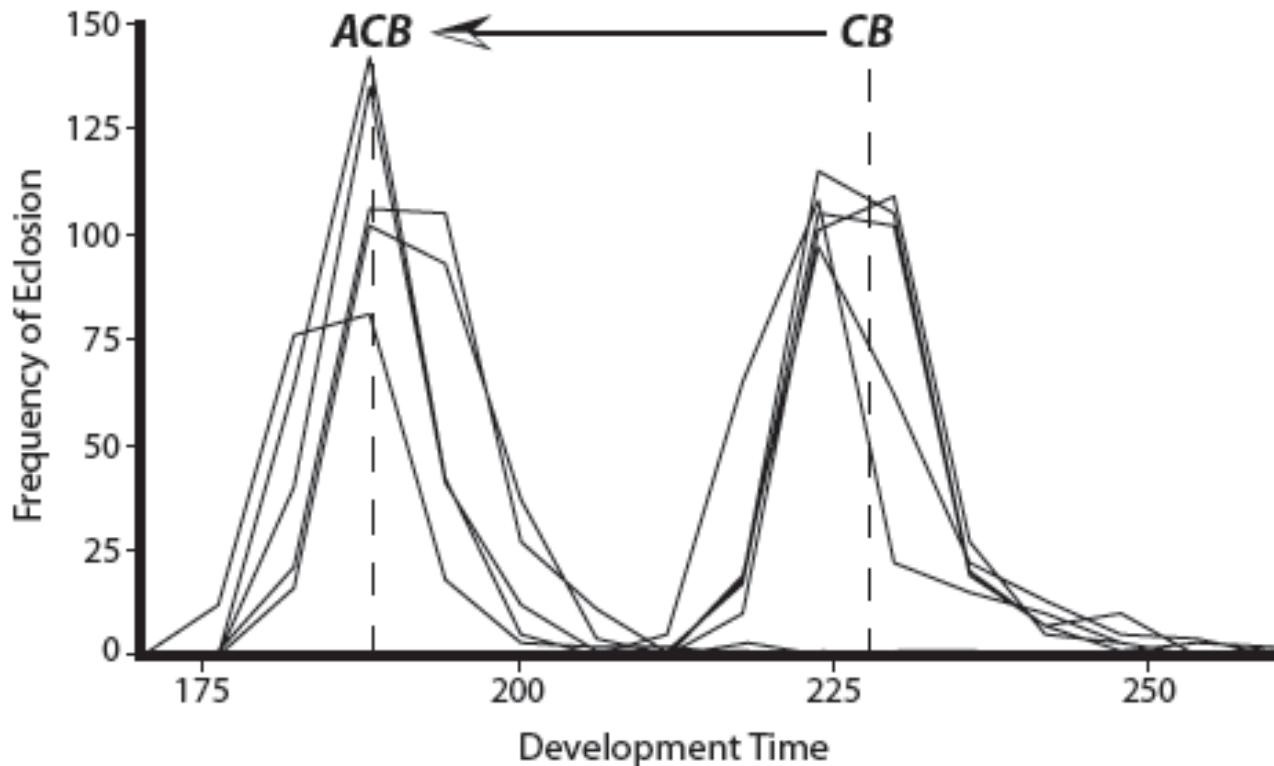
Pushing the Envelope

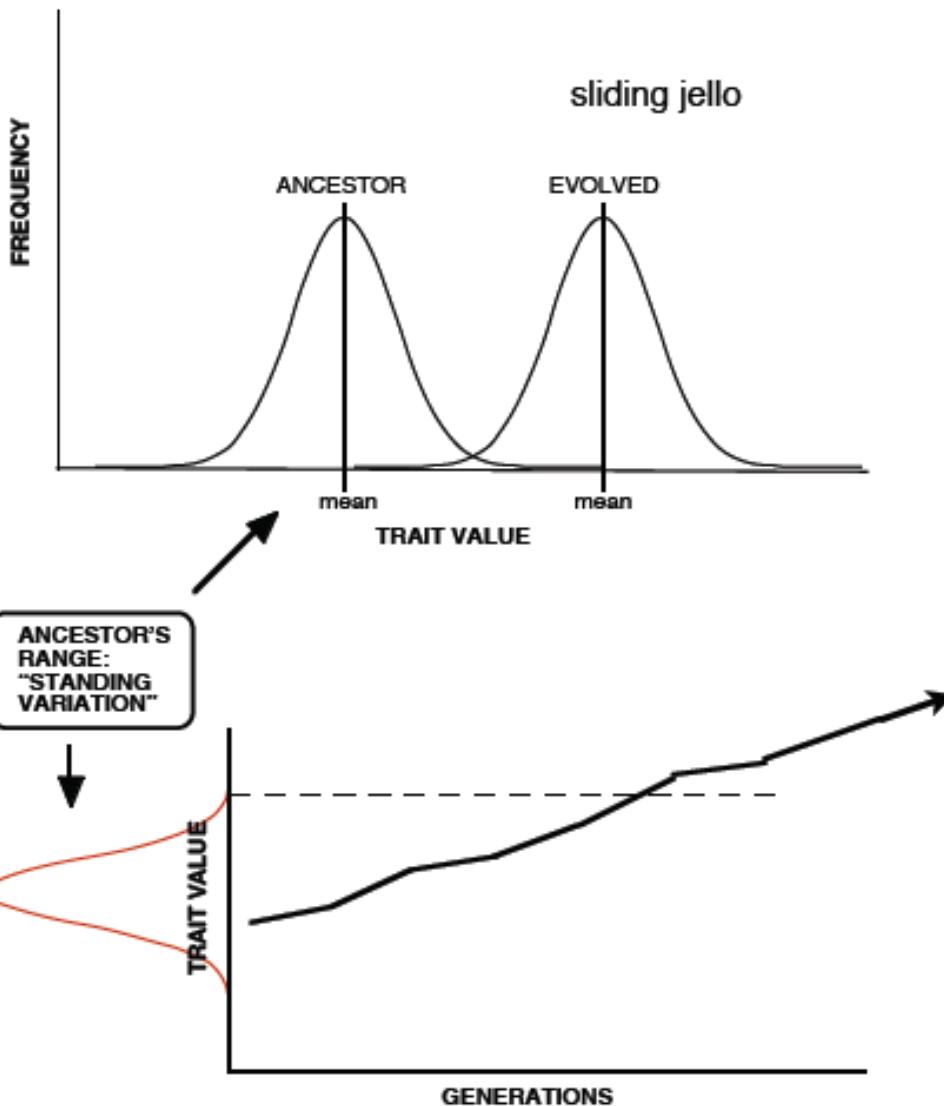






Selection on development time: No erosion of variation

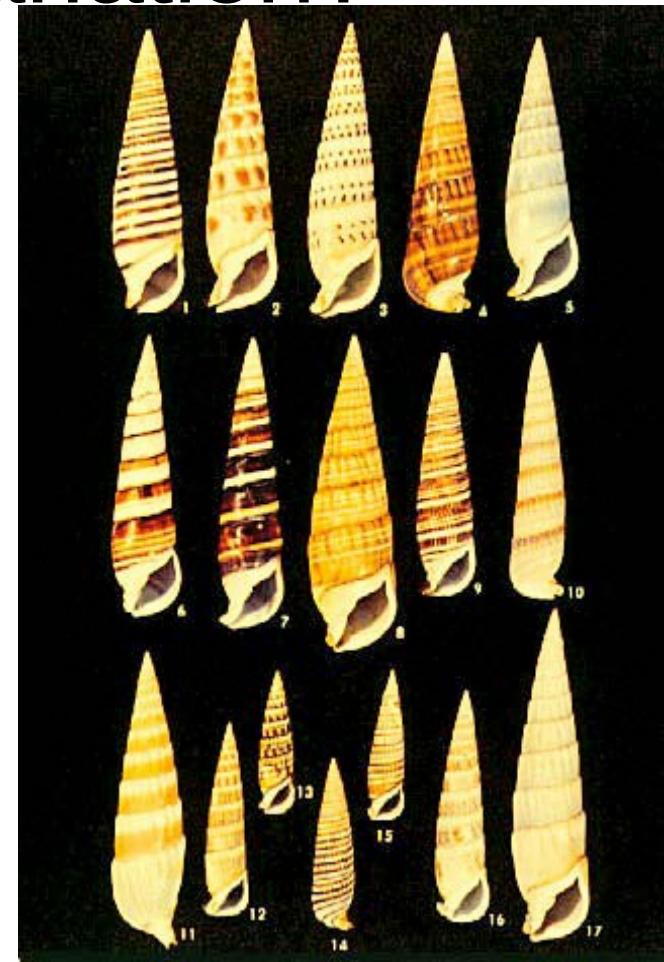




Early Problems for Darwinism:

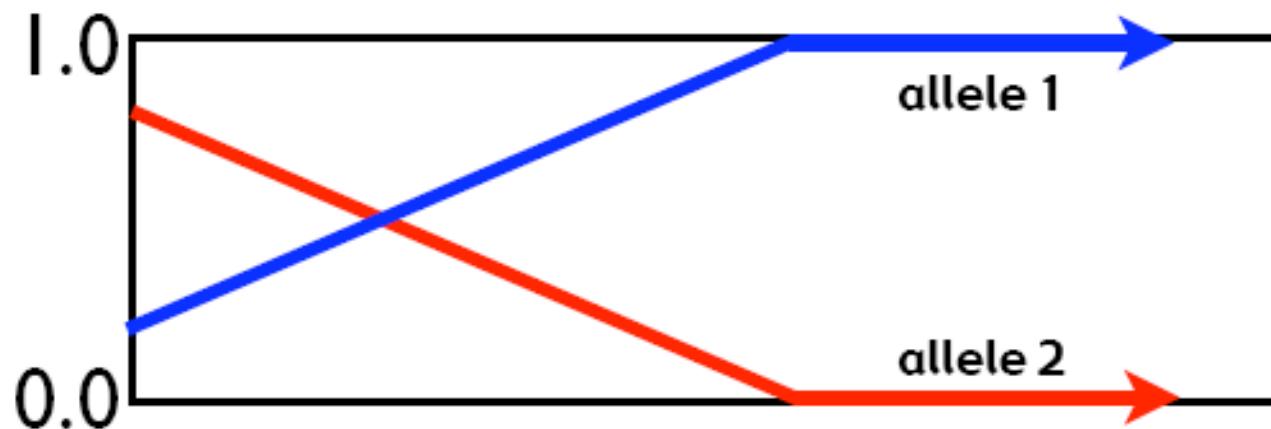
1. A paradox of variation?

- Selection removes variation.
- Natural observations and lab breeding experiments reveal rapid evolutionary responses = abundant variation in fitness.
- What sustains genetic variation in characters that affect fitness?



Selection is a purging process (?)

- Long term directional selection should eliminate less-favoured variants.
 - Fixation of favoured alleles should lead to reduced genetic variation.
 - Frequencies of 1.0 and 0.0 are absorbing boundaries.



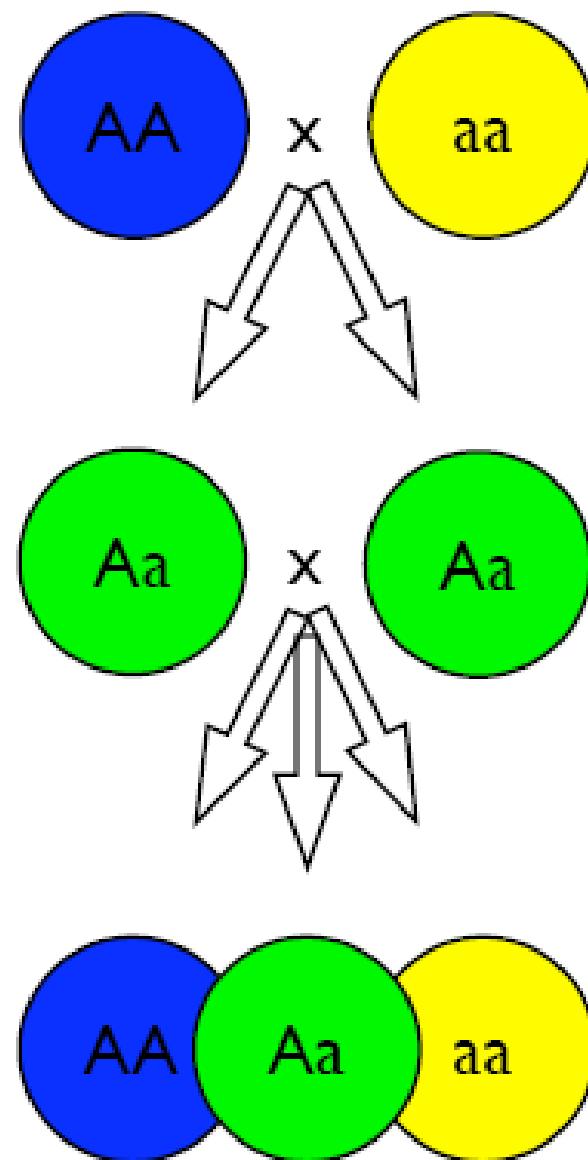


Early Problems for Darwinism: 2. The problem of Blending

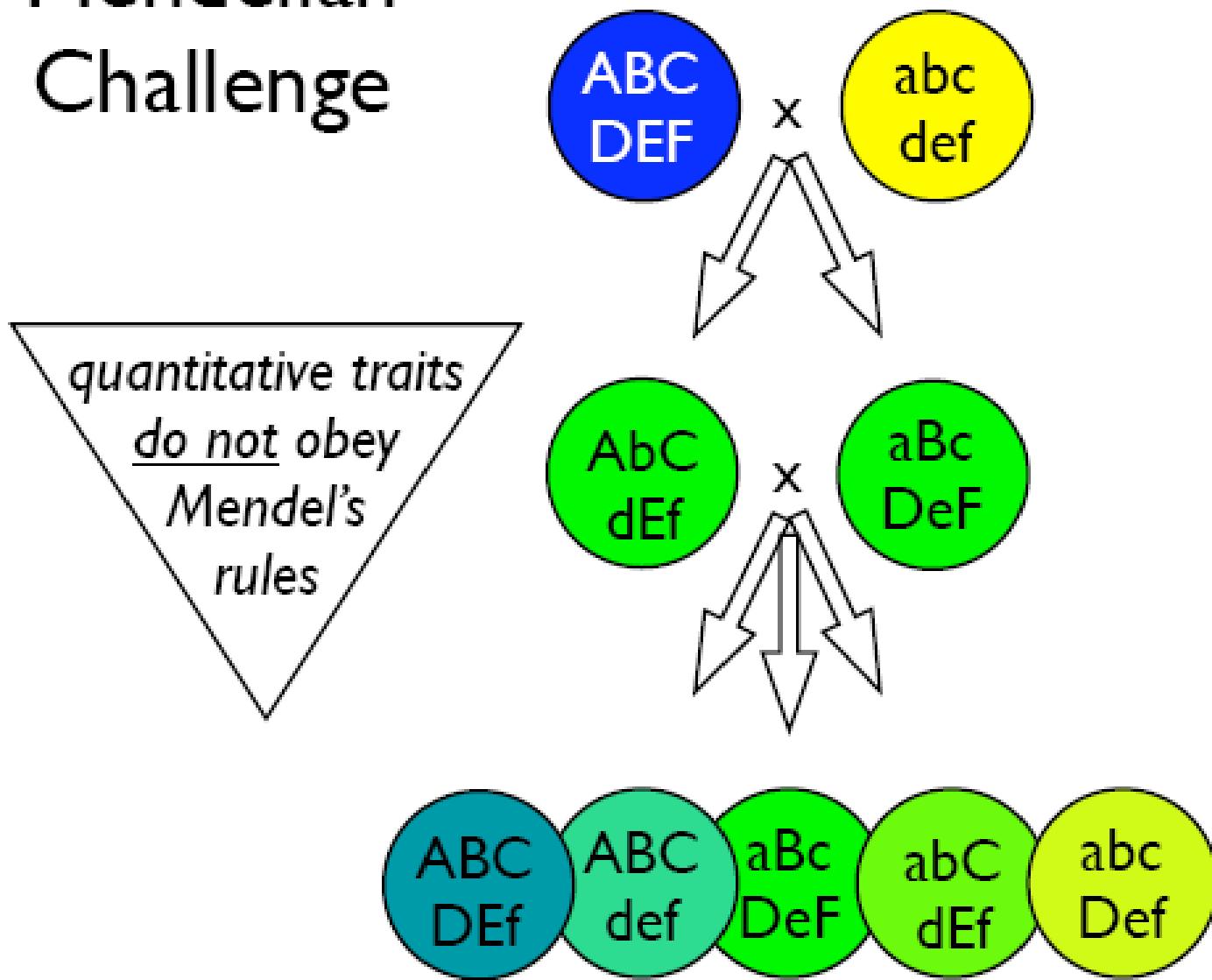
- Darwin and contemporaries did not understand genetics and inheritance.
 - Darwin followed Lamarck's thinking!
- Rediscovery of Mendelian genetics presented a major challenge initially but was ultimately crucial for evolutionary theory.
- Few traits are discrete and Mendelian (e.g. one locus with two alleles).

Mendelian Challenge

*simple
polymorphism
obeys
Mendel's
rules*

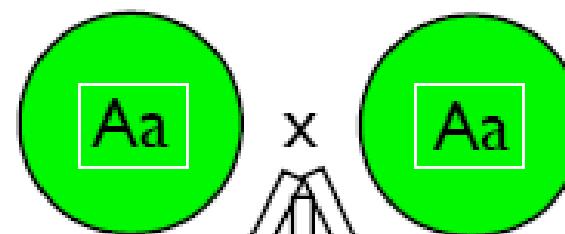
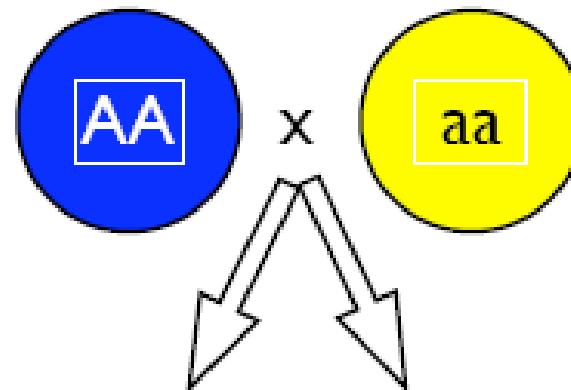


Mendelian Challenge



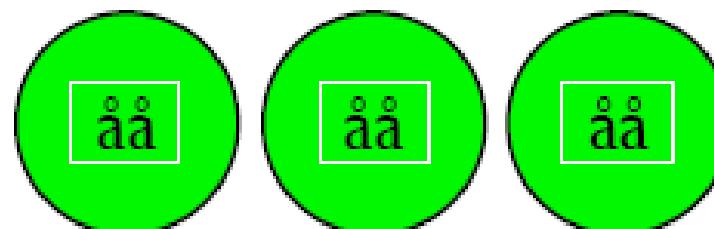
Mendelian Challenge

blending doesn't provide for genetic variation



average of AA
+
average of aa

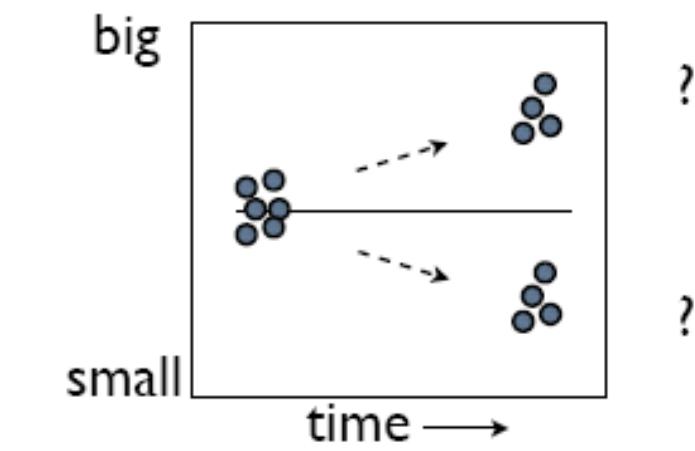
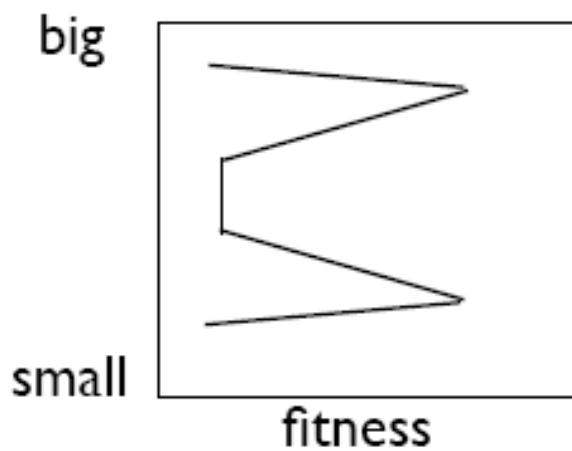
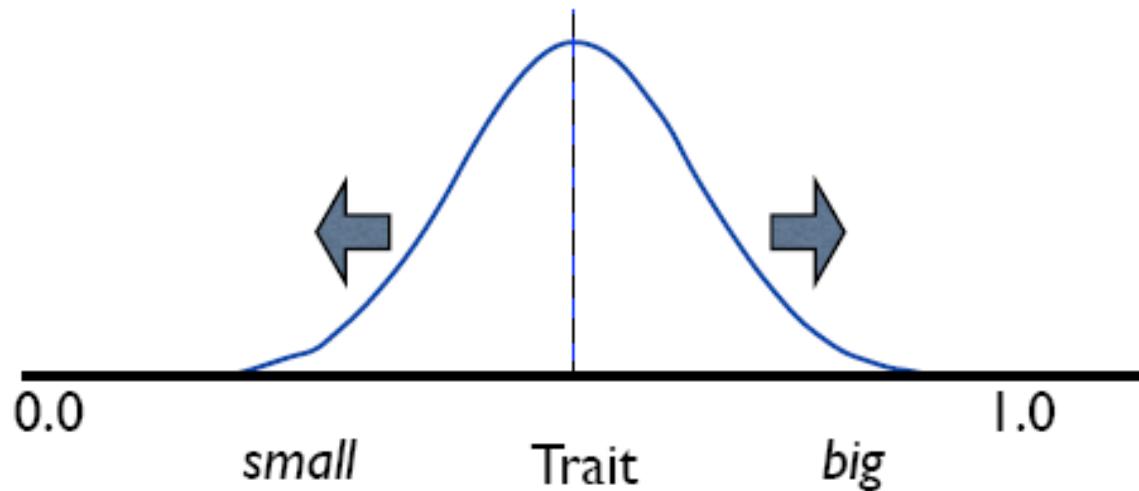
blended allele \ddot{a}



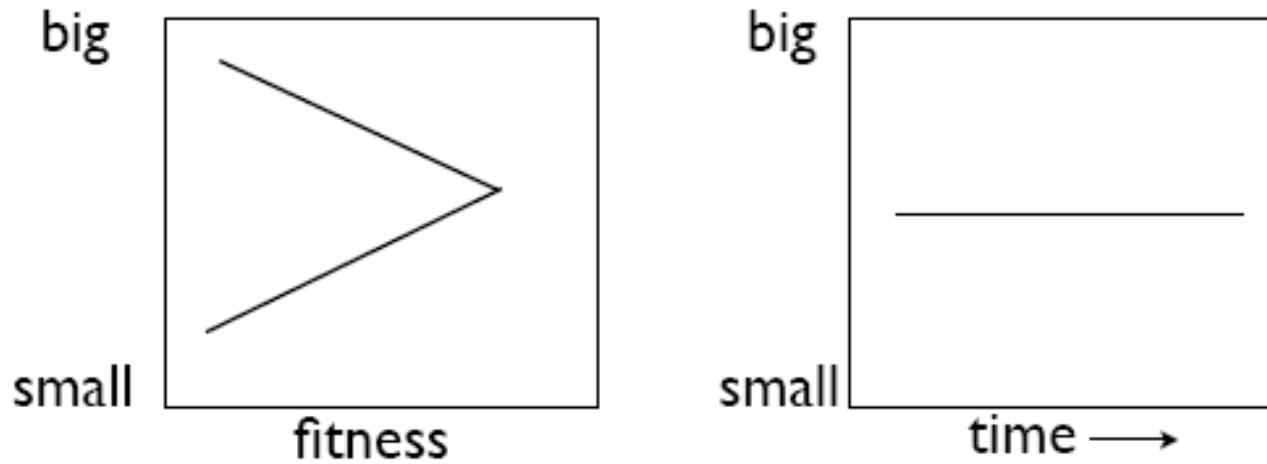
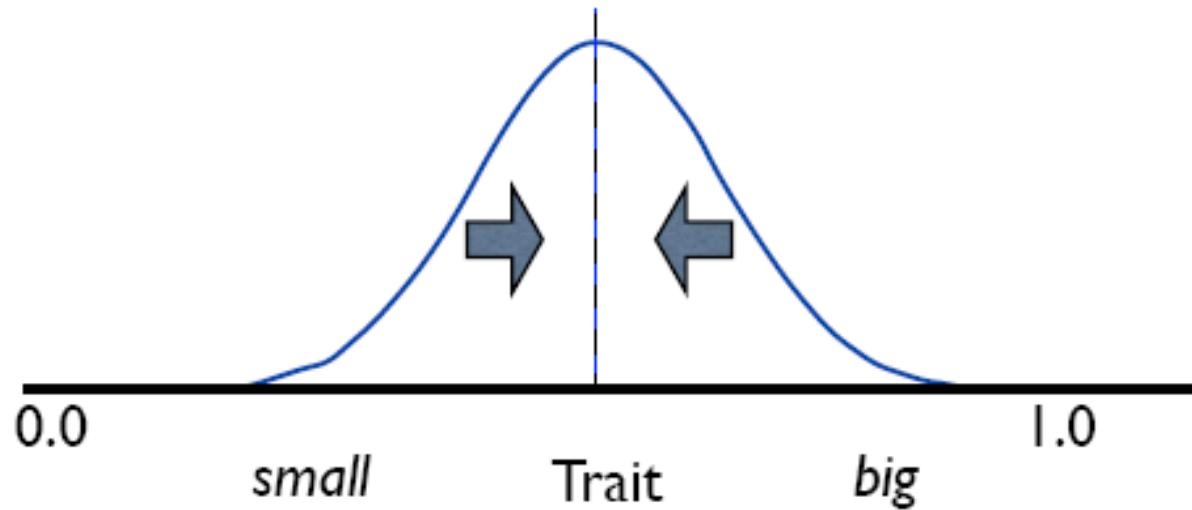
So what maintains variation for fitness?

- Trade-offs: improvement in one character result in decline of another.
- Immigration and spatial variation.
- Fluctuating selection pressure over time.
- Disruptive selection:
 - Over time, space or use of different elements of habitat.
- Mutation: the ultimate source of variation.

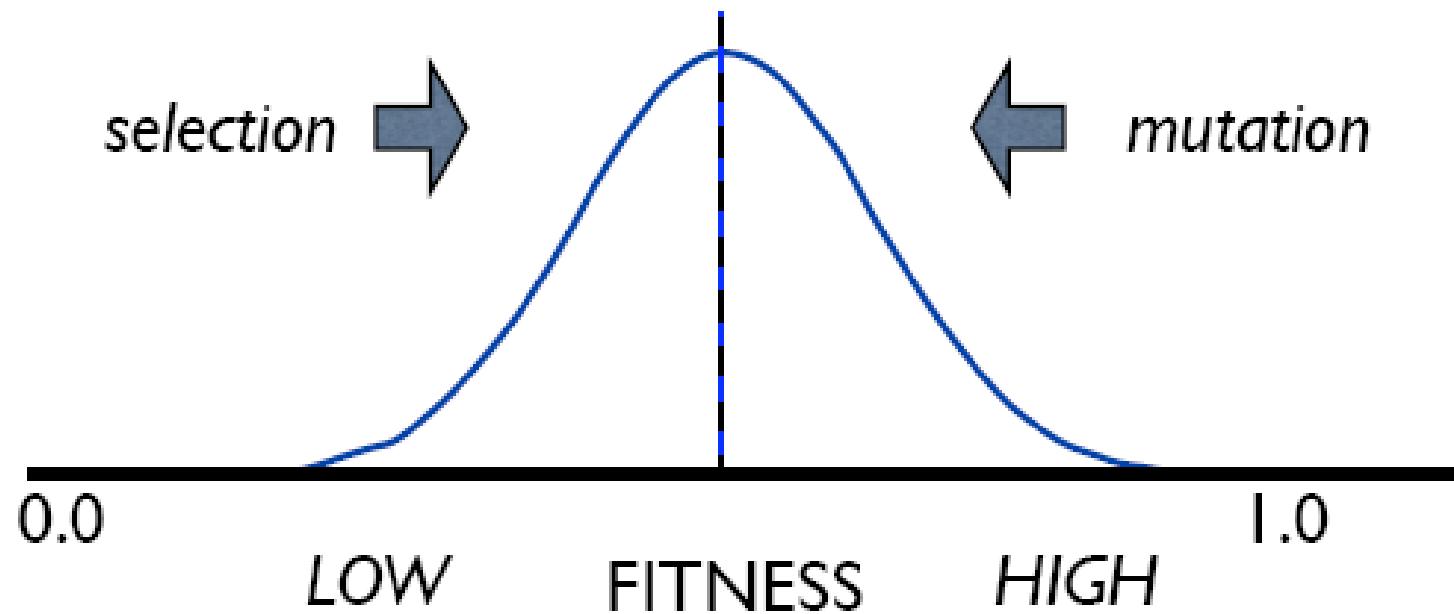
Disruptive selection



Stabilizing Selection



Mutation-Selection balance





Epidemiological Section.

February 28, 1908.

Dr. NEWSHOLME, President of the Section, in the Chair.

Mendelism in Relation to Disease.

By R. C. PUNNETT, M.A.

IT was with some trepidation that I accepted an invitation to read a paper bearing upon the inheritance of disease before a distinguished body like the Royal Society of Medicine; but I recollect the motto



were more blues than browns. The same applied to the examples of brachydactyly. The author said that brachydactyly was dominant. In the course of time one would then expect, in the absence of counteracting factors, to get three brachydactylous persons to one normal, but that was not so. There must be

Udny Yule



must lead to equilibrium. On my return to Cambridge I at once sought out G. H. Hardy with whom I was then very friendly. For we had acted as joint secretaries to the Committee for the retention of Greek in the Previous Examination and we used to play cricket together. Knowing that Hardy had not the slightest interest in genetics I put my problem to him as a mathematical one. He replied that it was quite simple and soon handed to me the now well-known formula $pr = q^2$.* Naturally pleased at getting so neat and prompt an answer I promised him that it should be known as "Hardy's Law"—a promise fulfilled in the next edition of my *Mendelism*. Whether the battle of Waterloo was won on the playing fields of Eton is still, I gather, a matter for conjecture : certain it is, however, that "Hardy's Law" owed its genesis to a mutual interest in cricket.

DISCUSSION AND CORRESPONDENCE

Mendelian Proportions in a Mixed Population

To The Editor of Science: I am reluctant to intrude in a discussion concerning matters of which I have no expert knowledge, and I should have expected the very simple point which I wish to make to have been familiar to biologists. However, some remarks of Mr. Udny Yule, to which Mr. R. C. Punnett has called my attention, suggest that it may still be worth making.

In the *Proceedings of the Royal Society of Medicine* (Vol. I., p. 165) Mr. Yule is reported to have suggested, as a criticism of the Mendelian position, that if brachydactyly is dominant "in the course of time one would expect, in the absence of counteracting factors, to get three brachydactylous persons to one normal."

It is not difficult to prove, however, that such an expectation would be quite groundless. Suppose that *Aa* is a pair of Mendelian characters, *A* being dominant, and that in any given generation the numbers of pure dominants (*AA*), heterozygotes (*Aa*), and pure recessives (*aa*) are as $p:2q:r$. Finally, suppose that the numbers are fairly large, so that the mating may be regarded as random, that the sexes are evenly distributed among the three varieties, and that all are equally fertile. A little mathematics of the multiplication-table type is enough to show that in the next generation the numbers will be as

$$(p+q)^2 : 2(p+q)(q+r) : (q+r)^2,$$

or as $p_1:2q_1:r_1$, say.

The interesting question is — in what circumstances will this distribution be the same as that in the generation before? It is easy to see that the condition for this is $q^2 = pr$. And since $q = p_1r_1$, whatever the values of p , q , and r may be, the distribution will in any case continue unchanged after the second generation.

Suppose, to take a definite instance, that *A* is brachydactyly, and that we start from a population of pure brachydactylous and pure normal persons, say in the ratio of 1:10,000. Then $p = 1$, $q = 0$, $r = 10,000$ and $p_1 = 1$, $q_1 = 10,000$, $r_1 = 100,000,000$. If brachydactyly is dominant, the proportion of brachydactylous persons in the second generation is 20,001:100,020,001, or practically 2:10,000, twice that in the first generation; and this proportion will afterwards have no tendency whatever to increase. If, on the other hand, brachydactyly were recessive, the proportion in the second generation would be 1:100,020,001, or

practically 1:100,000,000, and this proportion would afterwards have no tendency to decrease.

In a word, there is not the slightest foundation for the idea that a dominant character should show a tendency to spread over a whole population, or that a recessive should tend to die out.

I ought perhaps to add a few words on the effect of the small deviations from the theoretical proportions which will, of course, occur in every generation. Such a distribution as $p_1:2q_1:r_1$, which satisfies the condition $q = p_1r_1$, we may call a *stable* distribution. In actual fact we shall obtain in the second generation not $p_1:2q_1:r_1$ but a slightly different distribution $p:2q:r$, which is not "stable." This should, according to theory, give us in the third generation a "stable" distribution $p_2:2q_2:r_2$, also differing from $p_1:2q_1:r_1$; and so on. The sense in which the distribution $p_1:2q_1:r_1$ is "stable" is this, that if we allow for the effects of casual deviations in any subsequent generation, we should, according to theory, obtain at the next generation a new "stable" distribution differing but slightly from the original distribution.

I have, of course, considered only the very simplest hypotheses possible. Hypotheses other than [sic] that of purely random mating will give different results, and, of course, if, as appears to be the case sometimes, the character is not independent of that of sex, or has an influence on fertility, the whole question may be greatly complicated. But such complications seem to be irrelevant to the simple issue raised by Mr. Yule's remarks.

G. H. Hardy
Trinity College, Cambridge,
April 5, 1908

P. S. I understand from Mr. Punnett that he has submitted the substance of what I have said above to Mr. Yule, and that the latter would accept it as a satisfactory answer to the difficulty that he raised. The "stability" of the particular ratio 1:2:1 is recognized by Professor Karl Pearson (*Phil. Trans. Roy. Soc. (A)*, vol. 203, p. 60).

Reprinted from

Hardy, G. H. 1908. Mendelian proportions in a mixed population, *Science*, N. S. Vol. XVIII:49-50. (letter to the editor)

In the *Proceedings of the Royal Society of Medicine* (Vol. I., p. 165) Mr. Yule is reported to have suggested, as a criticism of the Mendelian position, that if brachydactyly is dominant "in the course of time one would expect, in the absence of counteracting factors, to get three brachydactylous persons to one normal."

It is not difficult to prove, however, that such an expectation would be quite groundless. Suppose that *Aa* is a pair of Mendelian characters, *A* being dominant, and that in any given generation the numbers of pure dominants (*AA*), heterozygotes (*Aa*), and pure recessives (*aa*) are as $p:2q:r$. Finally, suppose that the numbers are fairly large, so that the mating may be regarded as random, that the sexes are evenly distributed among the three varieties, and that all are equally fertile. A little mathematics of the multiplication-table type is enough to show that in the next generation the numbers will be as

$$(p+q)^2 : 2(p+q)(q+r) : (q+r)^2,$$

or as $p_1:2q_1:r_1$, say.

The interesting question is — in what circumstances will this distribution be the same as that in the generation before? It is easy to see that the condition for this is $q^2 = pr$. And since $q = p_1r_1$, whatever the values of p , q , and r may be, the distribution will in any case continue unchanged after the second generation.

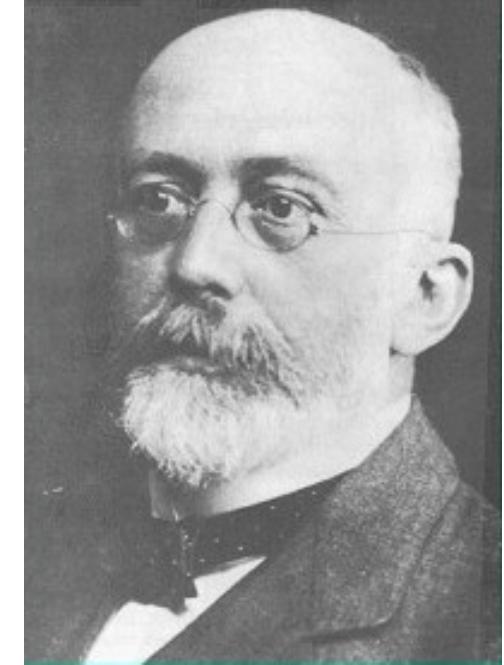
Suppose, to take a definite instance, that *A* is brachydactyly, and that we start from a population of pure brachydactylous and pure normal persons,



Godfrey Hardy



The equilibrium principle

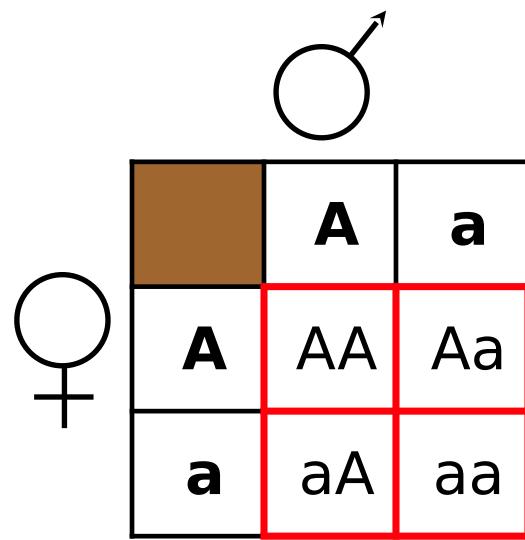


Wilhelm Weinberg

“A fundamental principle in population genetics stating that the genotype frequencies and gene frequencies of a large, randomly mating population remain constant providing that mutation, immigration, and selection do not take place”

-- American Heritage Dictionary

Allele vs. Genotype



Allele
frequencies

$$a=0.5$$
$$A=0.5$$

Genotype
Frequencies

$$aa=0.25$$
$$Aa=0.50$$
$$AA=0.25$$

		A	a
♀	A	AA	Aa
♂	a	aA	aa

Allele frequencies

$$a=0.5=q$$

$$A=0.5=p$$

$$p + q = 1.0$$

	A	a
♀	A	p ²
♂	a	pq

Genotype Frequencies

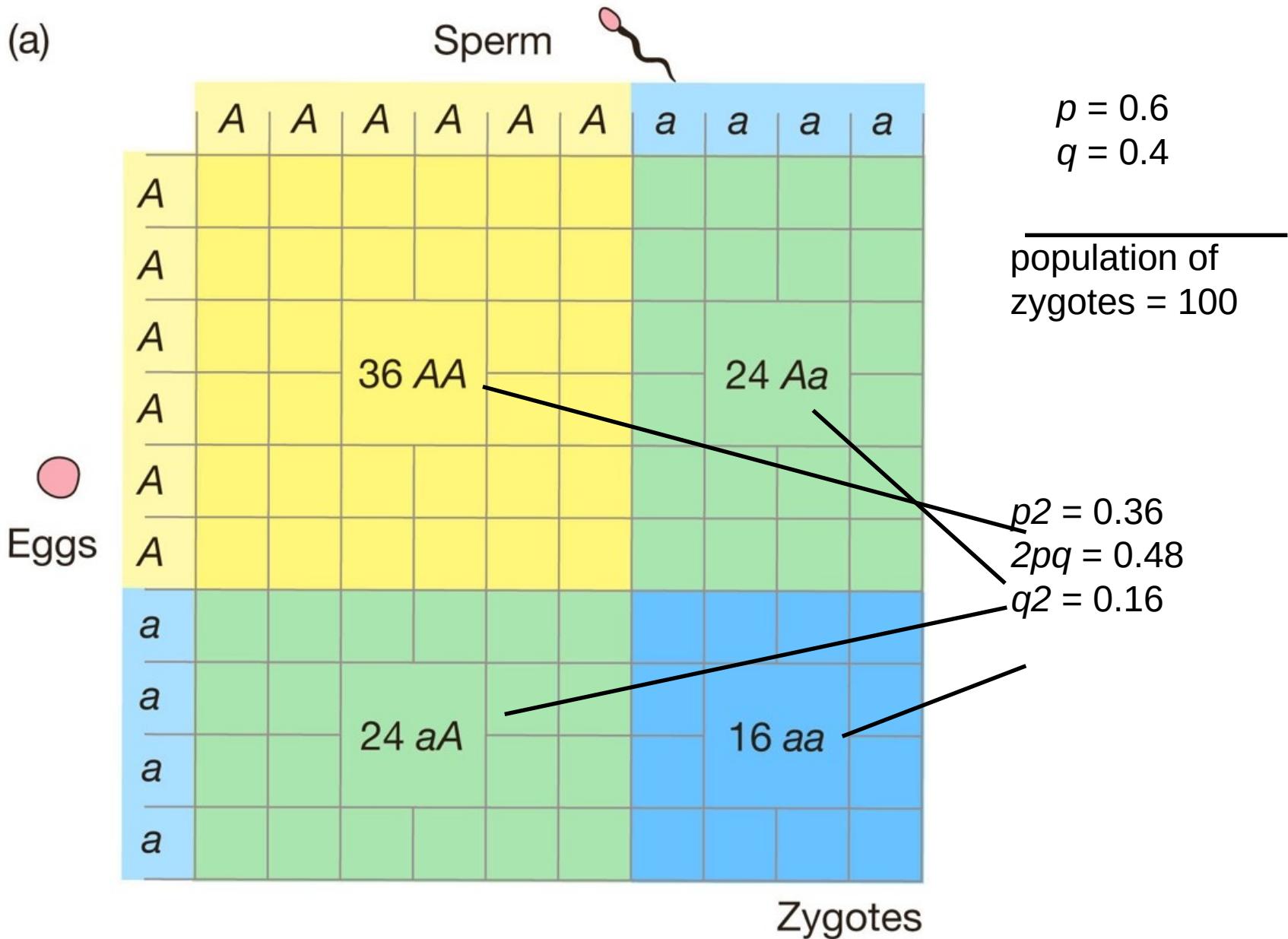
$$AA=p^2$$

$$Aa=2pq$$

$$aa=q^2$$

$$p^2 + 2pq + q^2 = 1.0$$

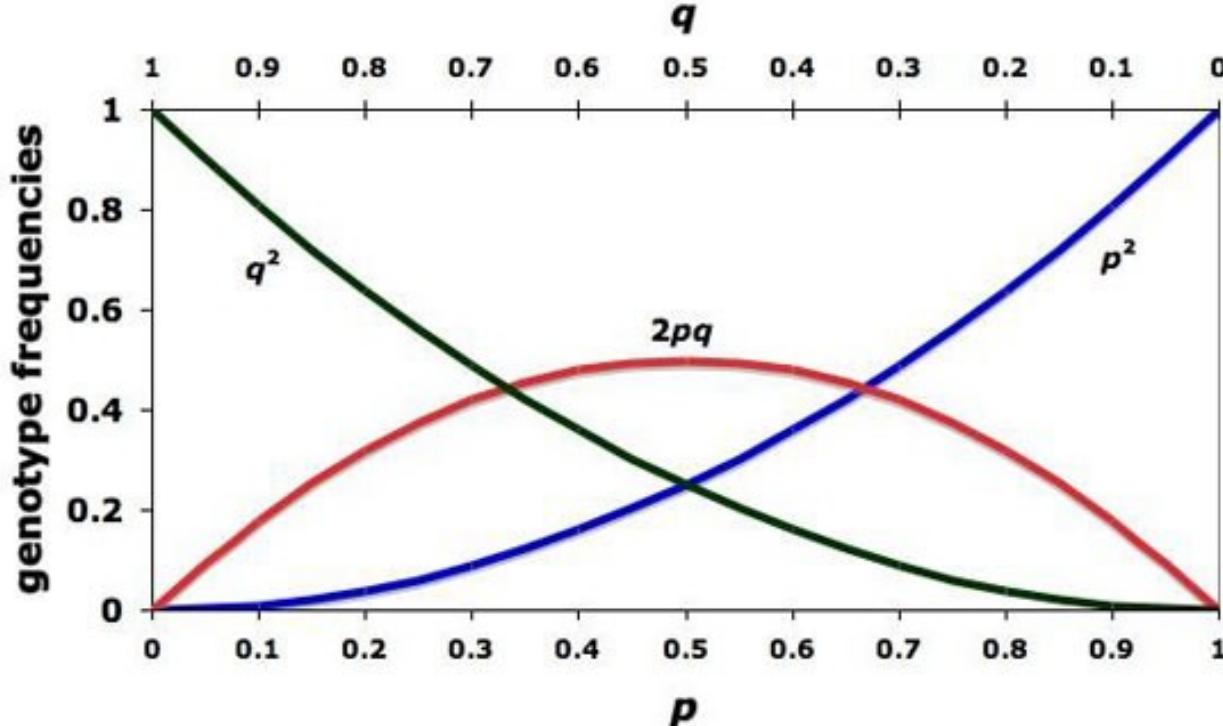
(a)



The next generation

- Assume each zygote grows up, produces 10 gametes and mates randomly.
 - $AA = 36 \times 10 = 360$
 - $Aa = 48 \times 10 = 480$
 - $aa = 16 \times 10 = 160$
- The frequency of A = $360 + 1/2 \times 480 = 600/1000 = 0.6$
- The frequency of a = $160 + 1/2 \times 480 = 400/1000 = 0.4$

The equilibrium is
reached after one
generation of random
mating.



- Hardy-Weinberg equilibrium point is a **neutral equilibrium point (not a stable equilibrium point)**, which means that a population perturbed from its Hardy-Weinberg genotype frequencies will reach a ***NEW equilibrium point*** after a single generation of random mating
- Population heterozygosity (the frequency of heterozygotes) is highest when $p = q = 0.5$.
- Rare alleles are found primarily in heterozygotes.

Using H-W

- we may use the H-W principle to determine if the population is in equilibrium.
 - count allele (gene) frequencies
 - calculate expected genotype frequencies
 - compare observed number of homozygotes with expected number.

Human Blood Groups

	MM	MN	NN	Σ	freq. M	freq. N
African American	Obs.	79	138	61		
	Exp. %				278	0.532
	Exp. #					0.468
Euro. American						
Amerind						
	$\text{freq. M} = (79 + 0.5 \cdot 138) / 278 = 0.532$					
	$\text{freq. N} = (61 + 0.5 \cdot 138) / 278 = 0.468$					

Human Blood Groups

		MM	MN	NN	Σ	freq. M	freq. N
African American	Obs.	79	138	61	278	0.532	0.468
	Exp. %	0.28	0.5	0.22			
	Exp. #	78.8	138.7	60.8			
Euro. American							
American Indian							
expected # of MM	2	$M = 0.532 * 278 = 0.28 * 278 = 78.8$					
expected # of NN	2	$N = 0.468 * 278 = 0.22 * 278 = 60.8$					

expected # of MN = $2 * 0.532 * 0.468 * 278 = 138.7$

Human Blood Groups

		MM	MN	NN	Σ	freq. M	freq. N
African American	Obs.	79	138	61	278	0.532	0.468
	Exp. %	0.28	0.5	0.22			
	Exp. #	78.8	138.7	60.8			
Euro. American	Obs.	1787	3039	1303	6129	0.540	0.460
	Exp. %	0.292	0.497	0.211			
	Exp. #	1787.2	3044.9	1296.9			
Amerind	Obs.	123	72	10	205	0.776	0.224
	Exp. %	0.602	0.348	0.05			

So who cares?

- H-W is important for three major reasons:
 - **practically**, deviations from H-W equilibrium point to interesting problems for investigation.
 - **conceptually**, it solved a major dilemma for early population genetics: how variation can be maintained.
 - **theoretically**, it helps simplify the calculations needed to detect nonequilibrium.

Sources of Non-Equilibrium

- “... the genotype frequencies and **gene frequencies** of a **large¹**, **randomly mating²** population **remain constant** providing that **mutation³**, **immigration⁴**, and **selection⁵** do not take place”
- deviation from equilibrium helps point out evolution and its sources

Sources of Non-Equilibrium: drift & sexual selection

- Genetic Drift: in finite populations, rare genotypes lost by chance death, infertility.
- Mate Choice: some males are more successful at breeding than others (gene frequency change may come about by selection or drift)
- Assortative Mating: pairing by like with like or by opposites.
 - e.g., positive assortative mating creates a deficit of heterozygotes -- may lead to reproductive isolation between groups.

Sources of Non-Equilibrium: Mutation & Migration

- mutation will rarely be detectable by H-W calculation due to its rarity.
- migration -- the movement of genes between populations -- a major topic of interest.

Sources of Non-Equilibrium: Selection

- deviations from H-W equilibrium often signal the operation of selection.
 - survival differences among genotypes
 - fertility differences among genotypes

H-W often violated

- The interest value of the H-W calculation is rarely predictive.
- Populations will often violate one or more of its assumptions.
 - the *mechanism of deviation* generally of much greater interest to researchers.

Conceptual Importance of H-W Law

- The H-W rule asserts that, all else being equal, allelic diversity is maintained.
- It expands easily to multiple loci.
- Helped save Mendelian principles from the problem of blending.

Analytical Importance of the H-W Law: General Model

P generation

adult genotype frequencies

1



frequency of mating between genotypes

2



frequency of genotypes in offspring from each type of mating

3



frequency of genotypes at birth

4

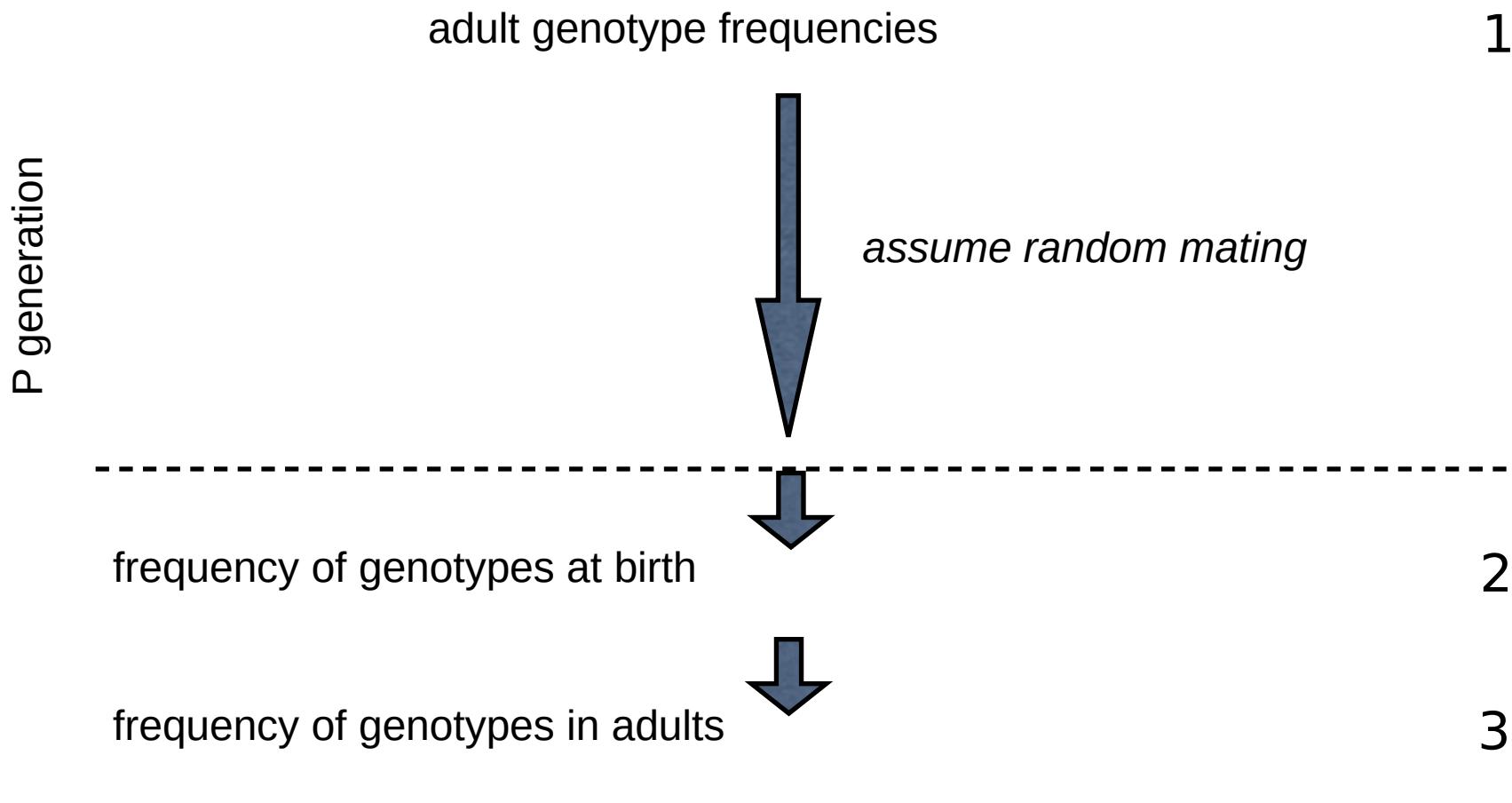


frequency of genotypes in adults

5

F1

Analytical Importance of the H-W Law: Simplification



A Special Case: X-Linked loci

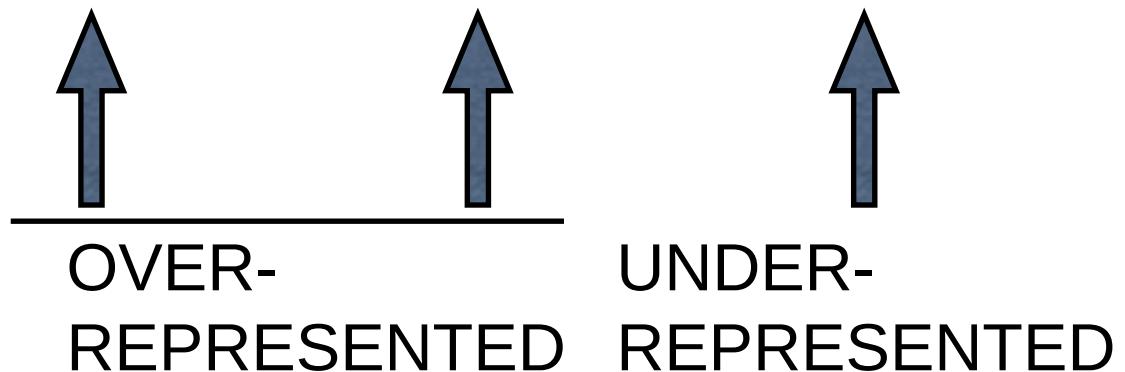
Is Selection Happening?

- Jaeken Syndrome = carbohydrate-deficient glycoprotein syndrome associated with c16 gene *PMM2* (*Phospomannomutase*)
- results in psychomotor dysfunction, severe retardation, numerous morphological abnormalities.
- all known affected individuals are homozygous deficient at *PMM2*, but there are 24 different mutations known.

Simple H-W

Other = 0.6, R141H = 0.4

	Other / Other	Other / R141H	R141H / R141H
observed	11	43	0
expected	0.36×54	0.48×54	0.16×54
<i>expected</i>	<i>19.4</i>	<i>25.9</i>	<i>8.6</i>



from Matthijs et al.

A simple test: the χ^2

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

= 23.56 for the example used

df = 1

p<0.001

Conclusions

- *R141H* is not in H-W equilibrium with other disease-causing alleles.
- suggests that this is a particularly strong allele, underrepresented in the population.
 - *R141H* homozygotes die before birth
- further survey of over 1000 Danish and Dutch individuals confirms these conclusions.
 - functional PMM expression necessary for life.

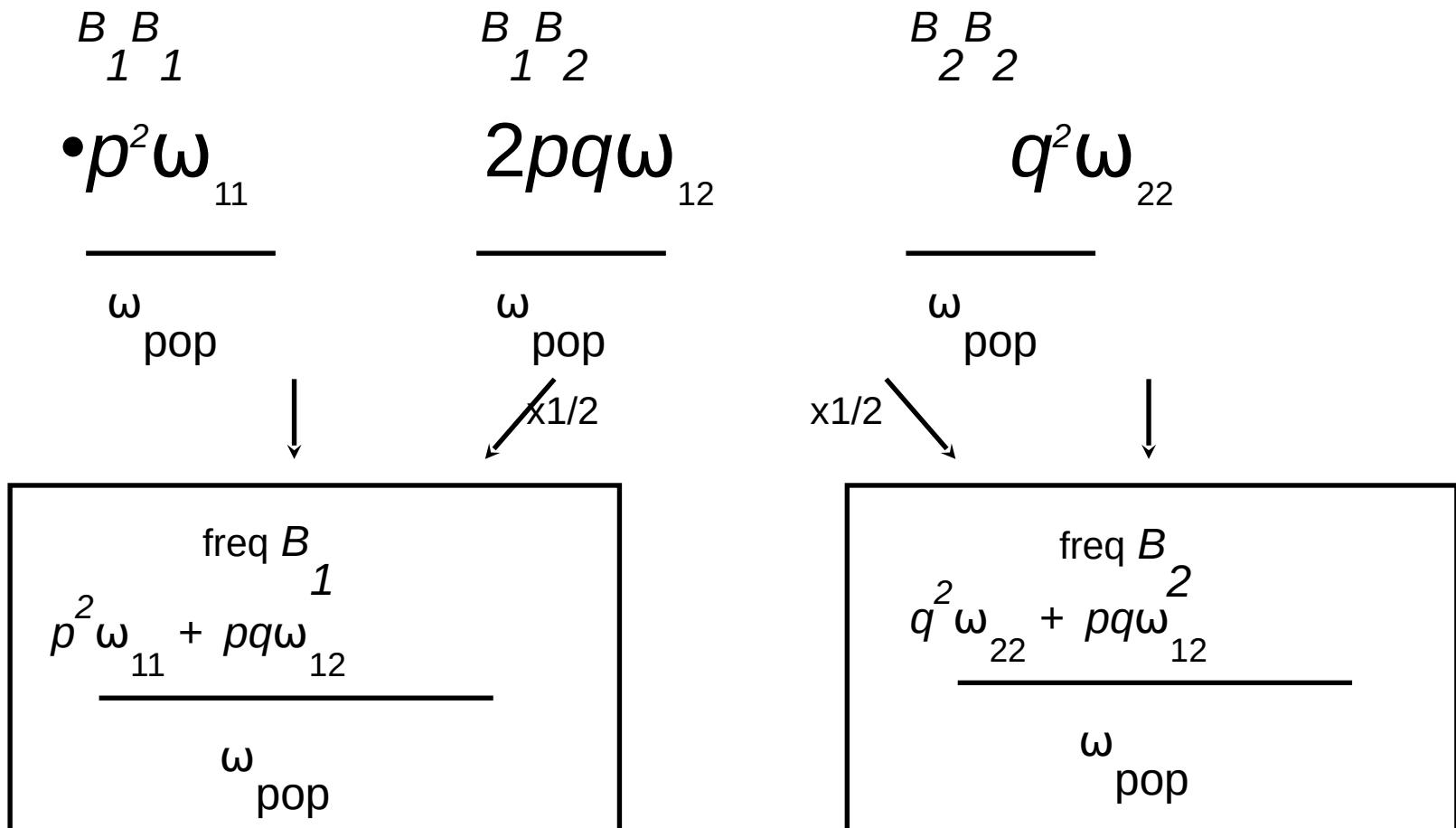
THE MATH

- ω used to denote fitness
- the average fitness of the population is simply:

$$-\omega_{\text{pop}} = p^2\omega_{11} + 2pq\omega_{12} + q^2\omega_{22}$$

—this is the number of offspring the population will produce in a given generation.

Expected genotype & gene frequencies



Changing frequencies

The change in frequency (Δ) for each allele
= (freq. generation x) - (freq. generation x-1)

$$\Delta B_1 = \frac{p^2 \omega_{11} + pq \omega_{12}}{\omega_{\text{pop}}} - p$$

$$\Delta B_1 = \frac{p^2 \omega_{11} + pq\omega_{12}}{\omega_{\text{pop}}} - p$$

$$\Delta B_1 = \frac{p^2 \omega_{11} + pq\omega_{12}}{\omega_{\text{pop}}} - \frac{p\omega_{\text{pop}}}{\omega_{\text{pop}}}$$

$$\Delta B_1 = \frac{p}{\omega_{\text{pop}}} \{ p\omega_{11} + q\omega_{12} - \omega_{\text{pop}} \}$$

Average Excess: predicts rate of change in frequency

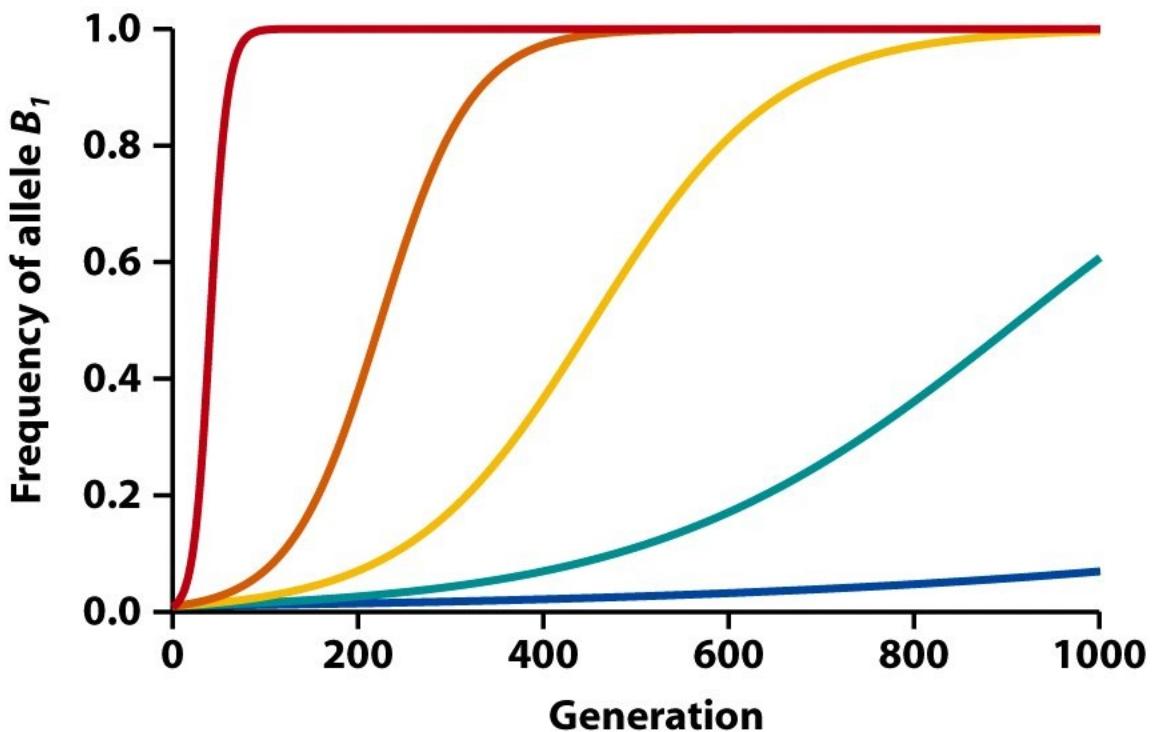
$$\Delta B_1 = \frac{p}{\omega_{\text{pop}}} \left\{ p\omega_{11} + q\omega_{12} - \omega_{\text{pop}} \right\}$$

$$\frac{\left\{ p\omega_{11} + q\omega_{12} - \omega_{\text{pop}} \right\}}{\text{average fitness of } B1} \quad \frac{\text{average fitness of population}}$$

General Behaviour

$$\Delta B_1 = \frac{p}{\omega_{\text{pop}}} \{ p\omega_{11} + q\omega_{12} - \omega_{\text{pop}} \}$$

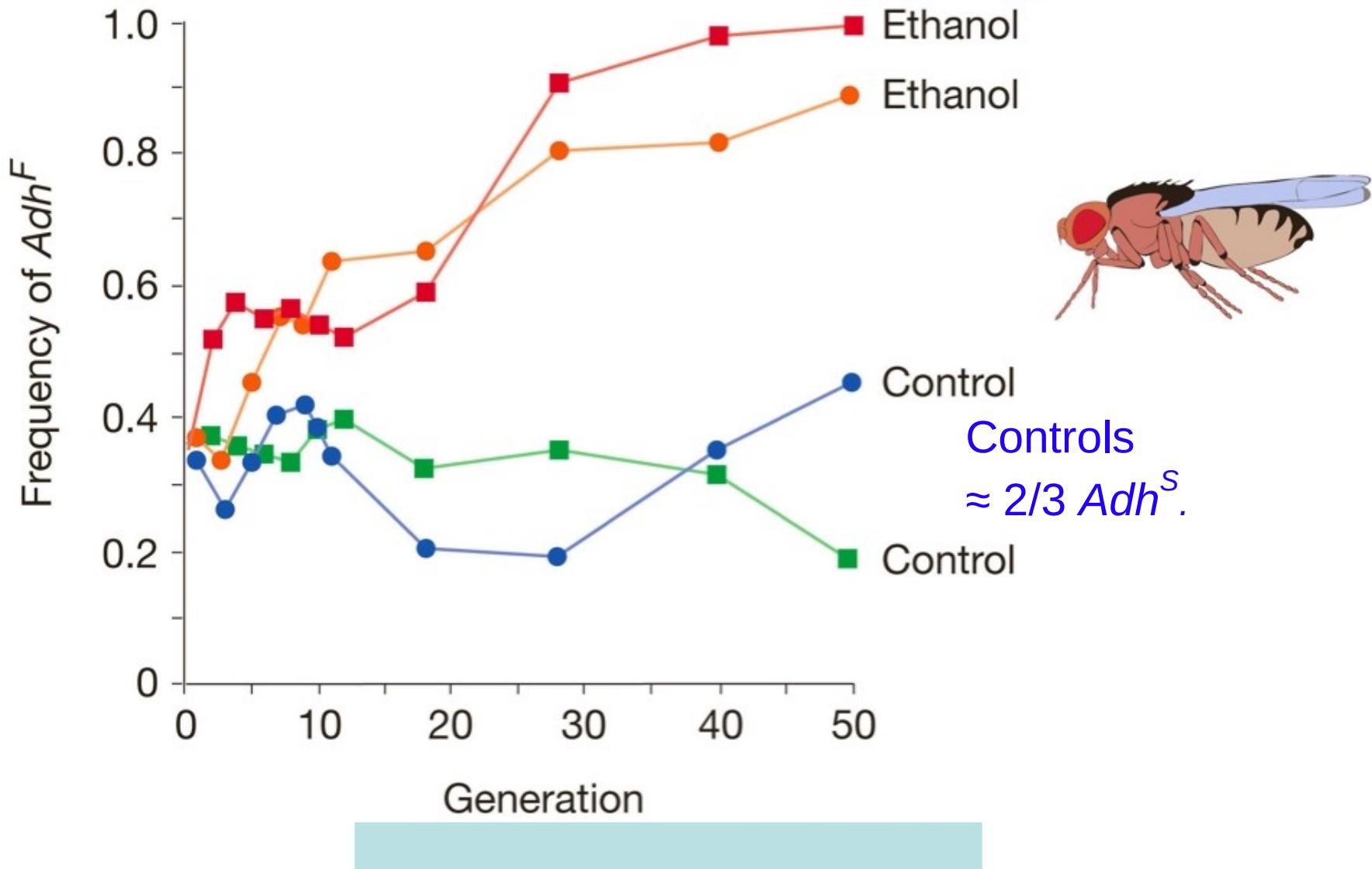
- the greater the fitness advantage, the faster the fixation.
- when rare:
 - B_1 advantage accelerates with frequency.
- as B_1 becomes more common:
 - population average fitness increases and its relative advantage declines.



Selection scheme

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

Figure 6-12 Evolutionary Analysis, 4/e
 © 2007 Pearson Prentice Hall, Inc.

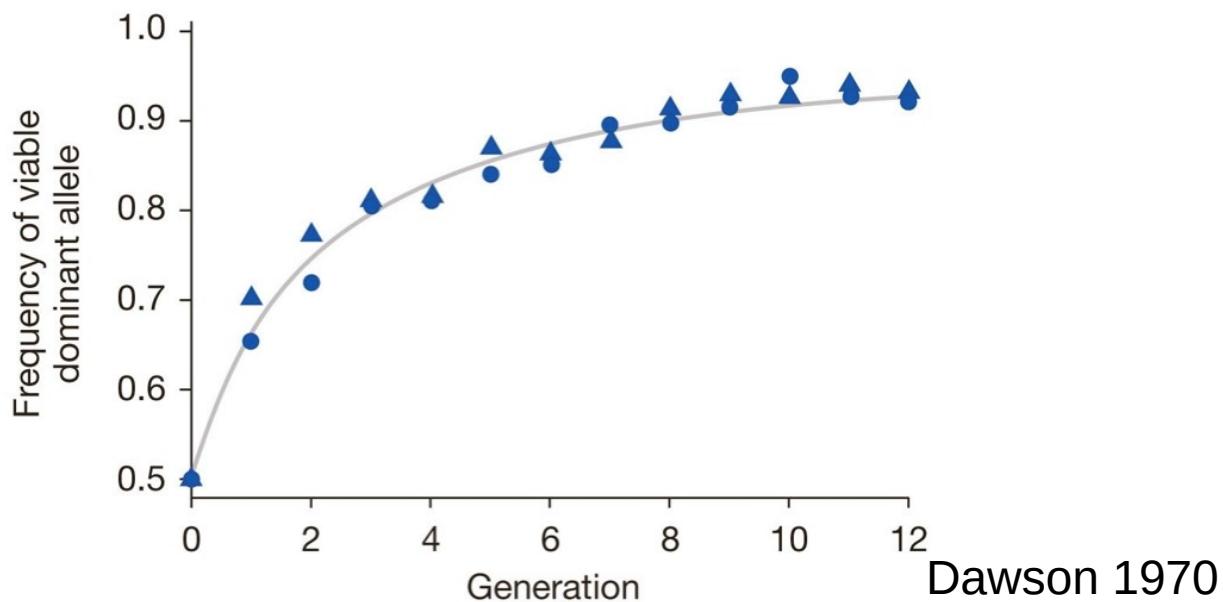
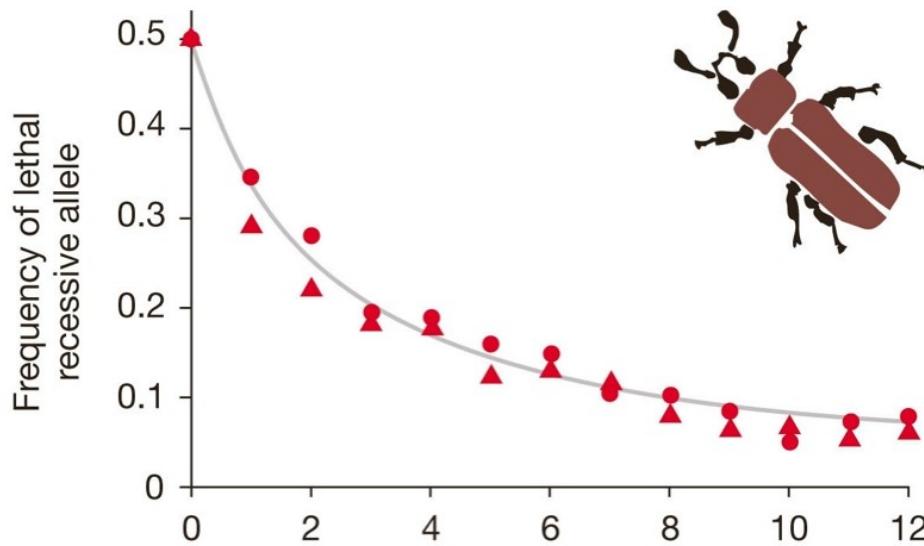


Selection for Adh^F allele in *D. melanogaster*
by Cavener & Clegg

1. Dominance

- Selection for a dominant allele is also selection against its recessive counterpart.
 - A disfavoured recessive becomes rarer, but its expression rarer still.
 - A favoured dominant rapidly becomes more common, but population fitness increases as it does so.
 - *Q: how quickly does a recessive lethal exit a population?*

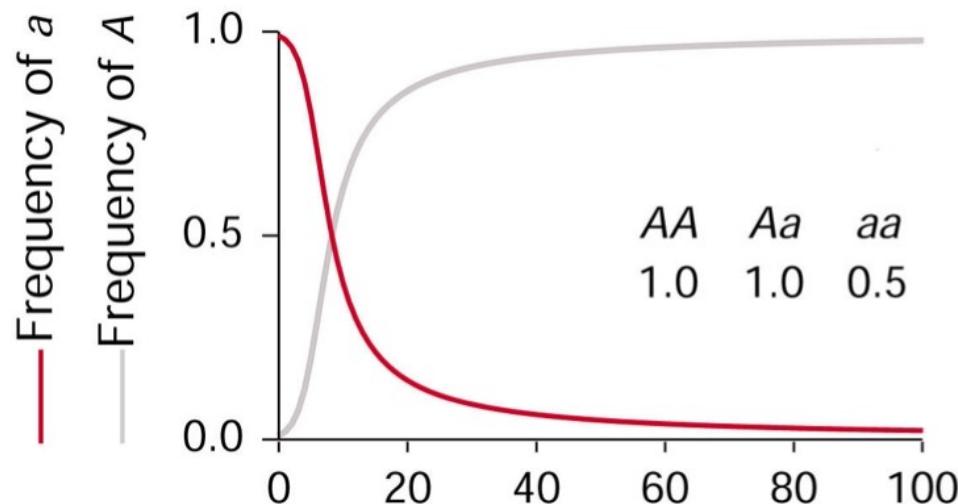
Example: recessive lethal.



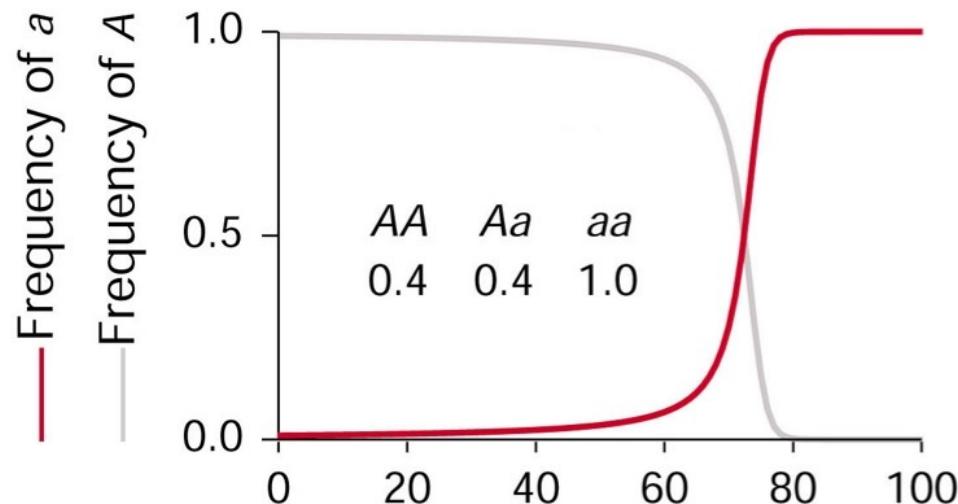
Rates of Change

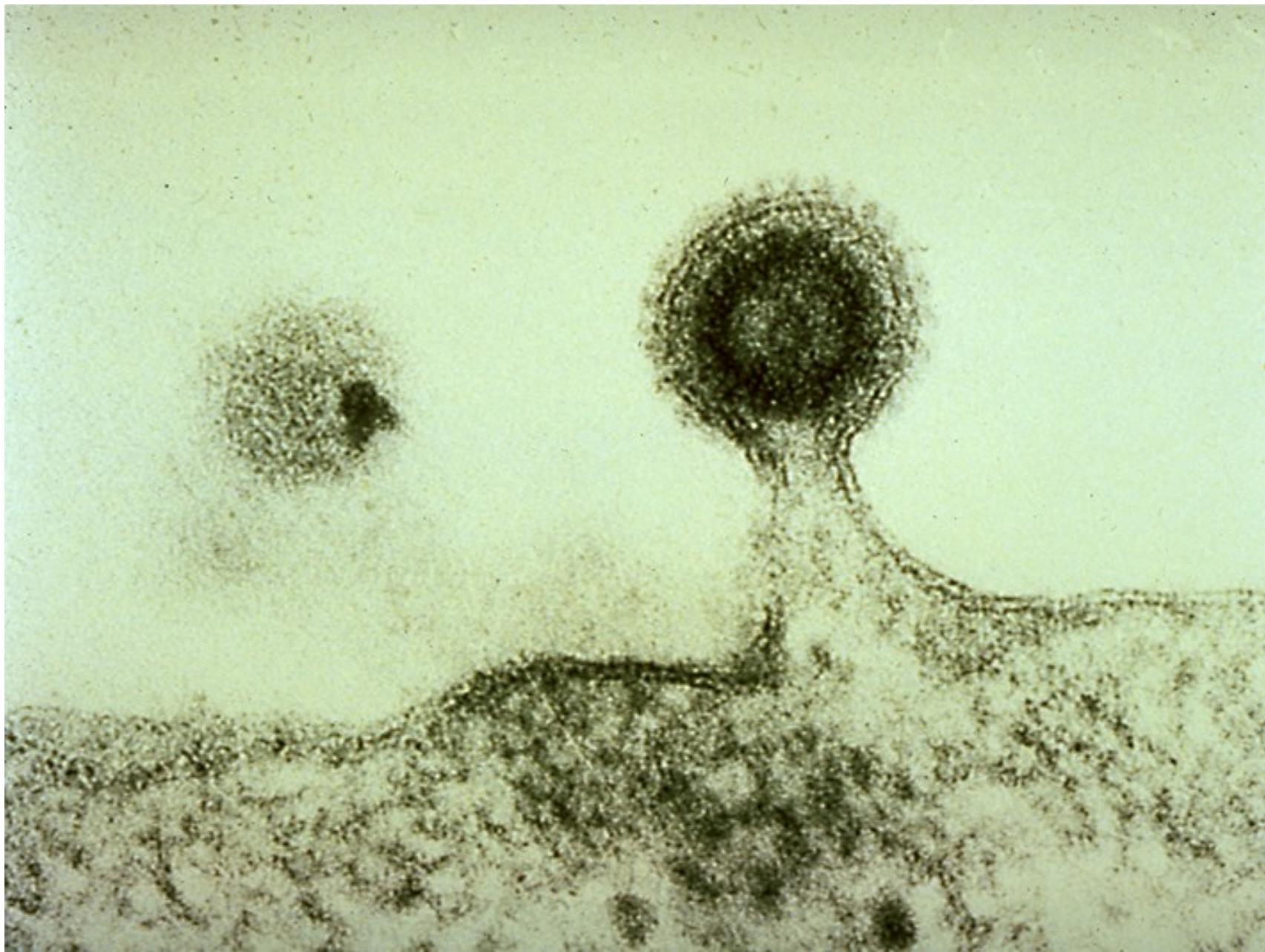
- favoured dominant alleles quick to take off, slow to finish.
- favoured recessives slow to take off, quick to finish.
- fastest change occurs at intermediate frequencies (i.e., 0.5:0.5)

(a)



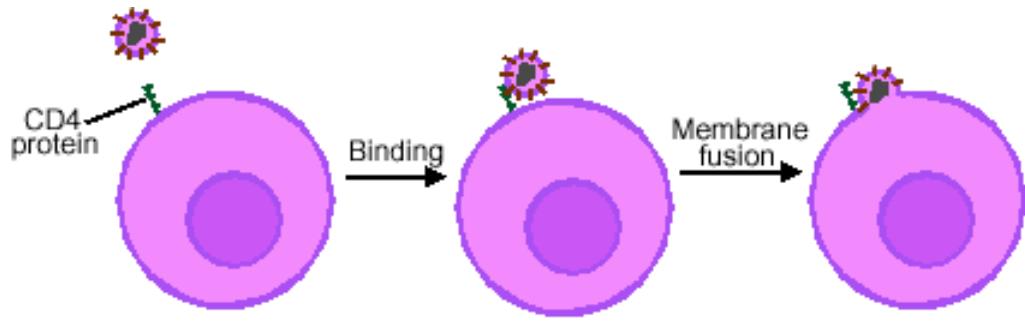
(b)



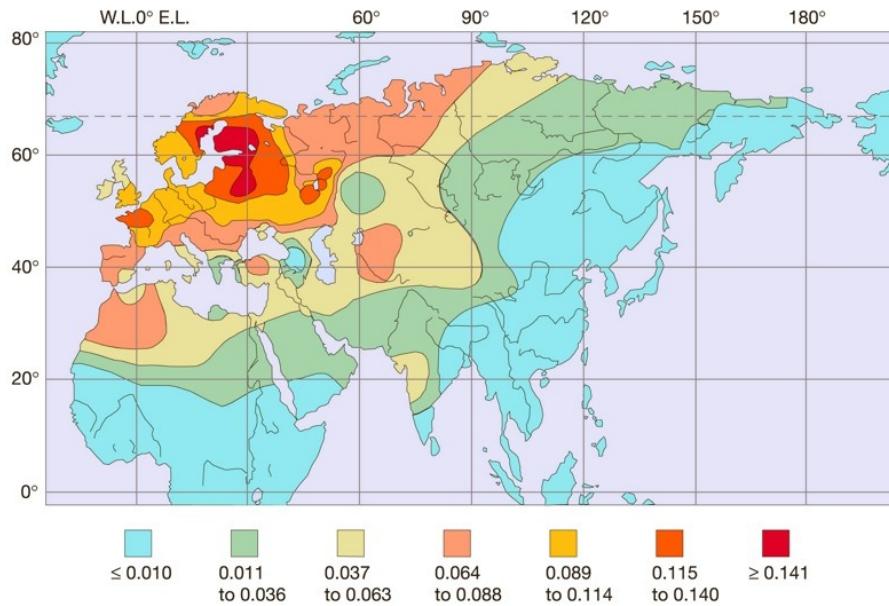


© H.R. Gelderblom, Robert-Koch-Institut, Berlin

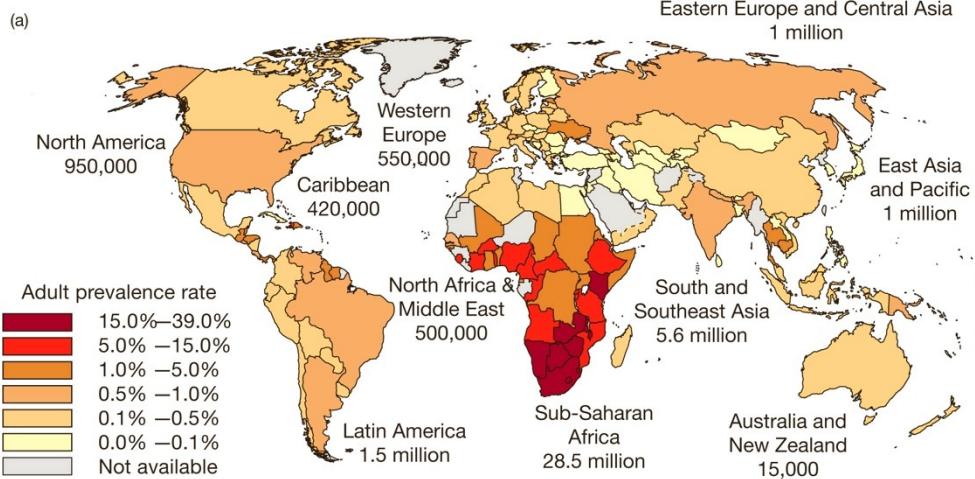
CCR5-Δ32 revisited



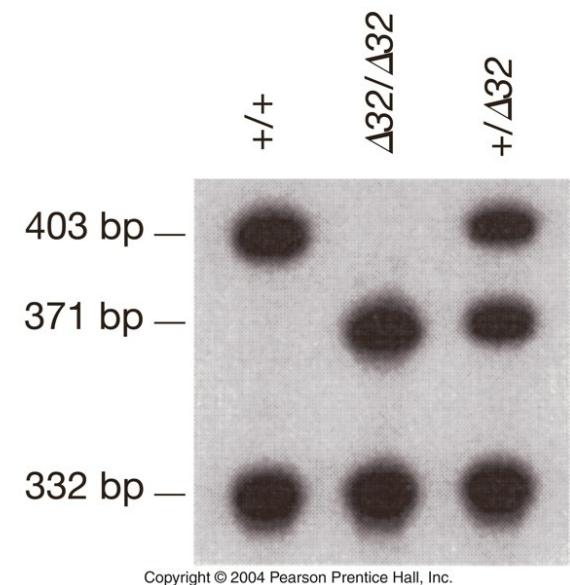
- This mutant allele of the chemokine receptor confers significant protection against HIV
 - will it sweep?
 - The answer depends upon:
 1. the frequency of the allele.
 2. the fitness advantage conferred when exposed
 - note: there may be a cost too, in absence of HIV
 3. the probability of exposure to disease.



Copyright © 2004 Pearson Prentice Hall, Inc.

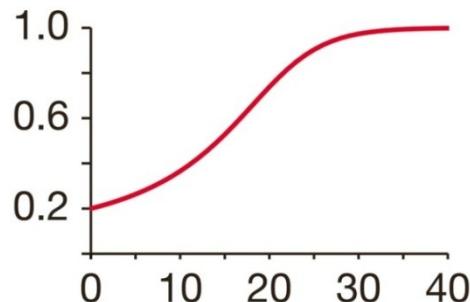


Copyright © 2004 Pearson Prentice Hall, Inc.



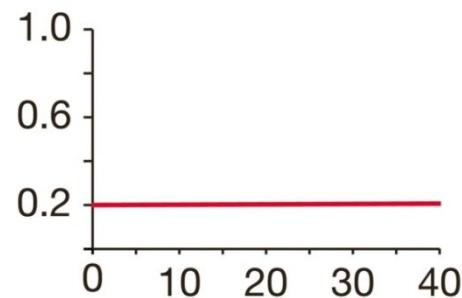
PCR used to amplify fragment of *cIII*

$+/+$	$+/\Delta 32$	$\Delta 32/\Delta 32$
0.75	0.75	1.0



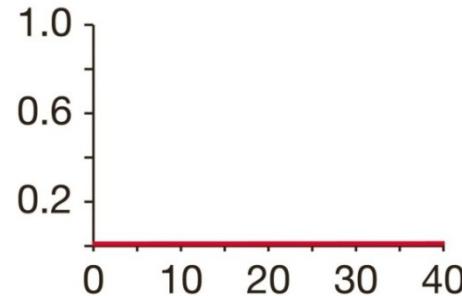
- strong selection
- high initial frequency (20%)

$+/+$	$+/\Delta 32$	$\Delta 32/\Delta 32$
0.995	0.995	1.0



- weak selection
- high initial frequency (20%)

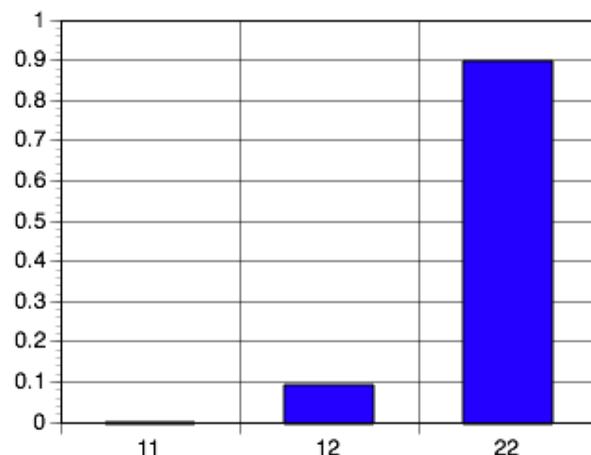
$+/+$	$+/\Delta 32$	$\Delta 32/\Delta 32$
0.75	0.75	1.0



- strong selection
- low initial frequency (1%)

Freq. 1 = 5%

Freq. 2 = 95%

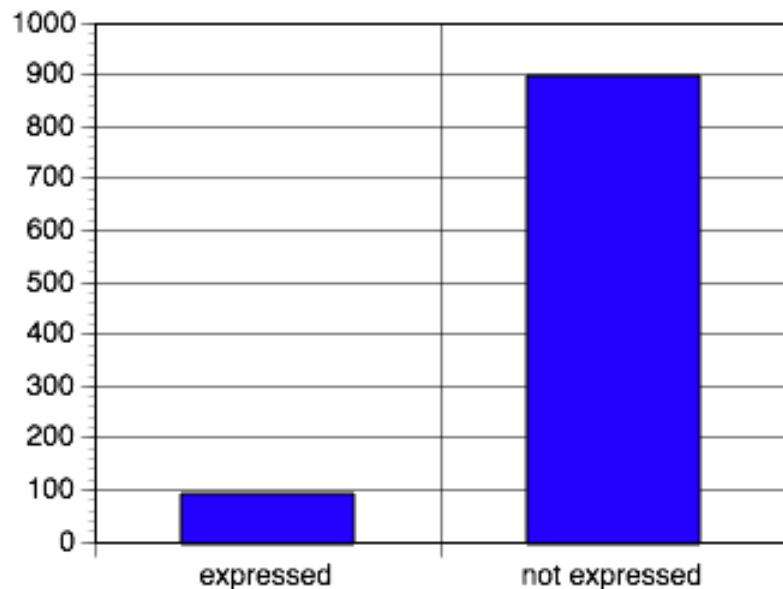


rare
homozygote
 $0.05 \times 0.05 = 0.025$

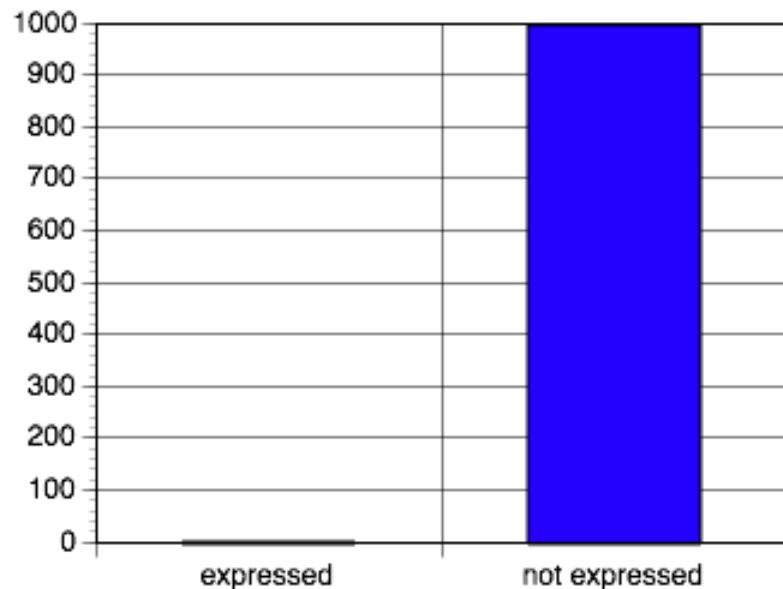
common
homozygote
 $0.95 \times 0.95 = 0.9025$

heterozygote
 $2(0.05 \times 0.95) = 0.095$

allele 1 dominant



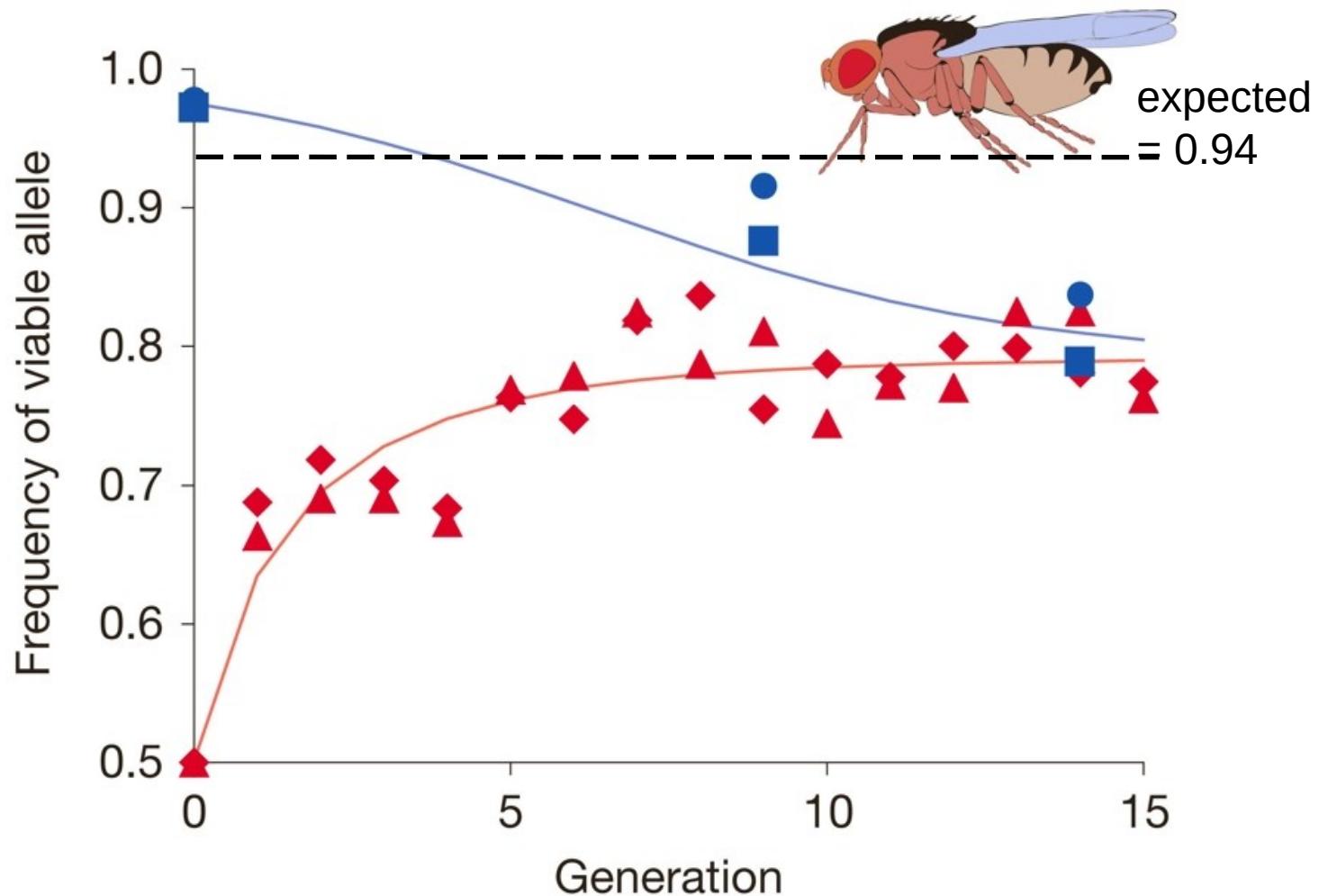
allele 2 dominant



II. Overdominance

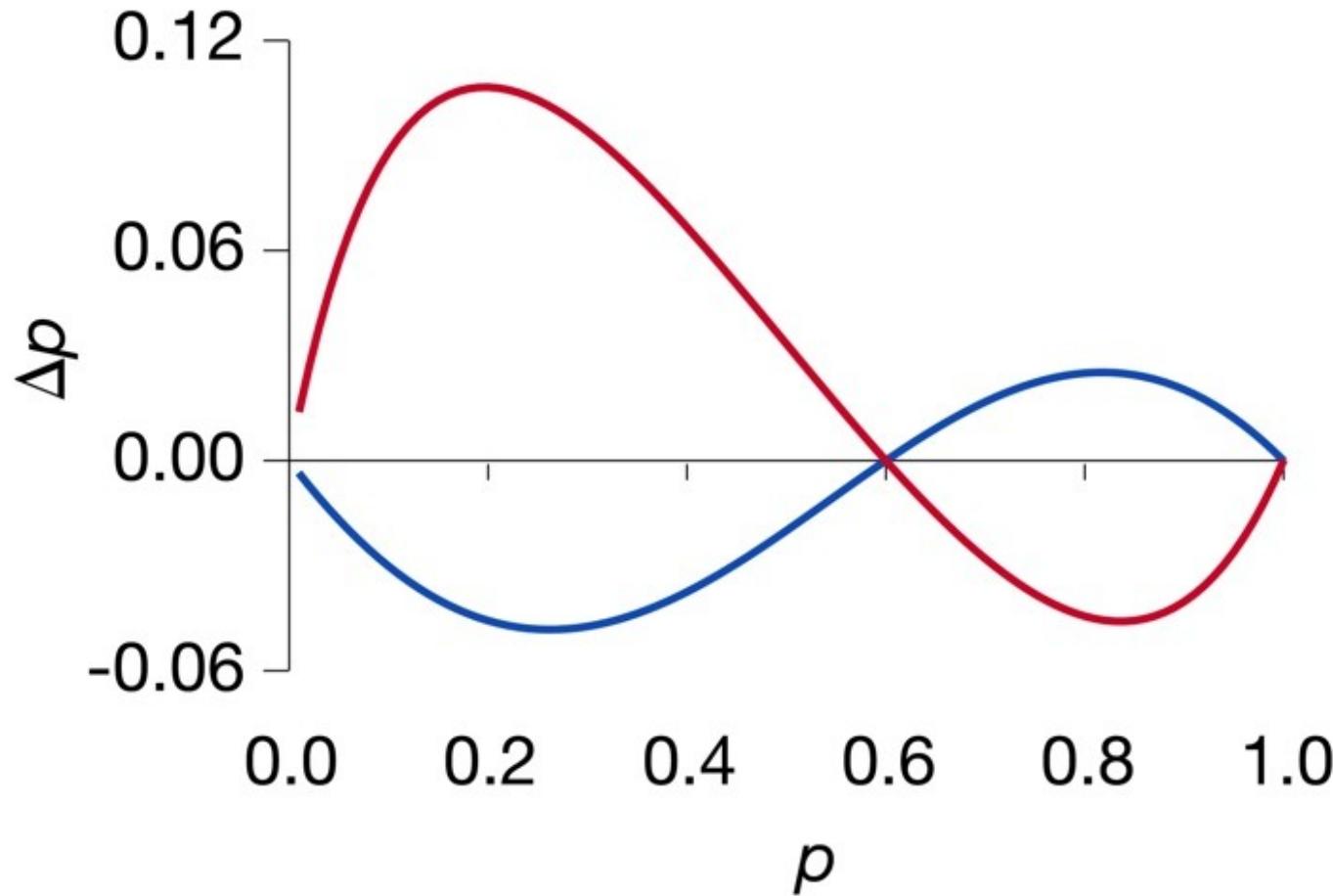
- Sickle cell Anaemia
 - S allele maintained by the advantage of heterozygotes.
- Mukai & Burdick
 - Selection against a recessive lethal allele.

Mukai & Burdick Data



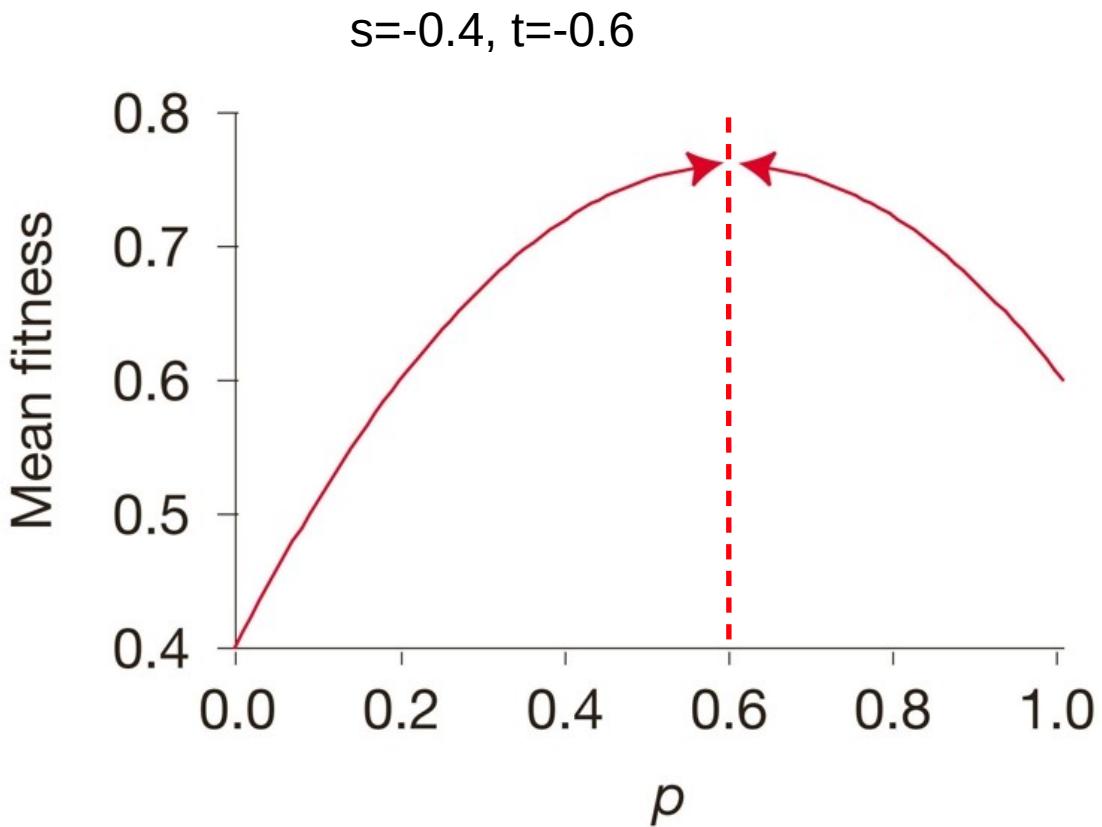
Selection on Homozygotes/ Heterozygotes

(a) Δp as a function of p



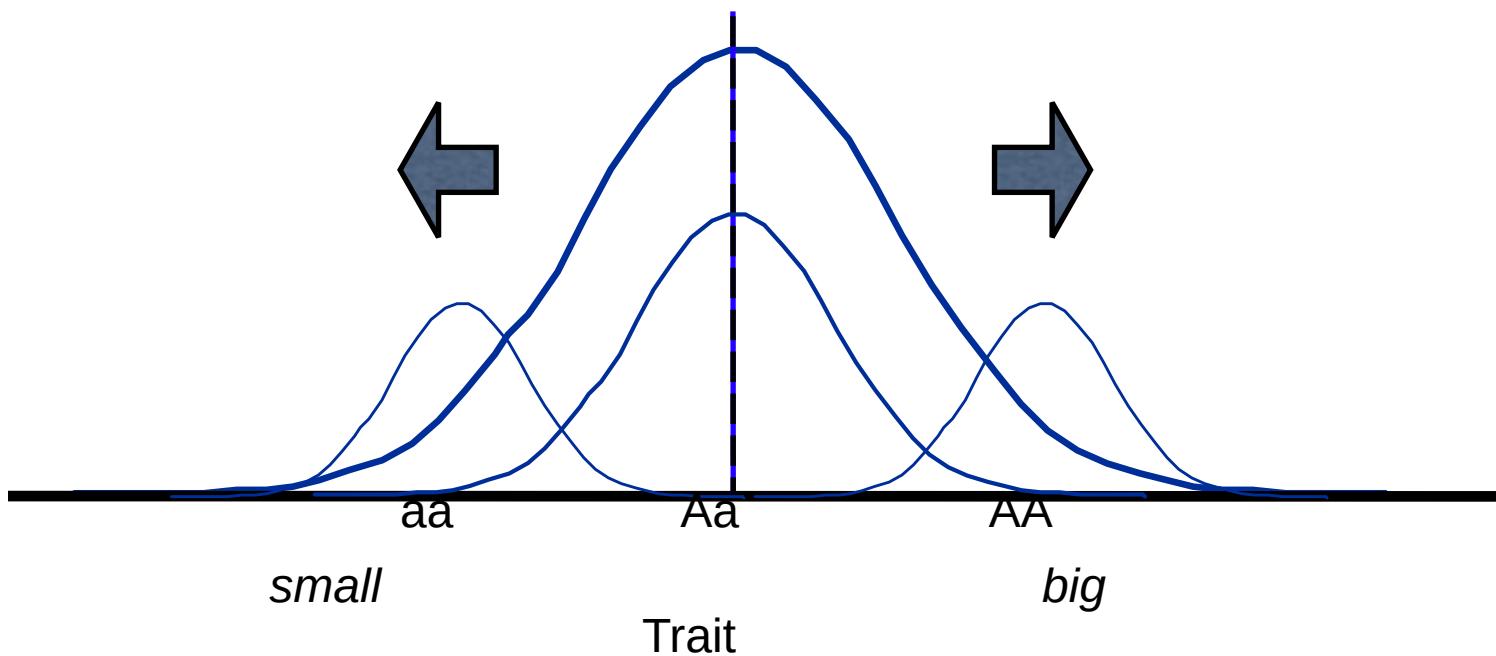
Copyright © 2004 Pearson Prentice Hall, Inc.

- a simple *fitness landscape*
- stable equilibrium of alleles reached at $p = 0.6$
- both alleles retained



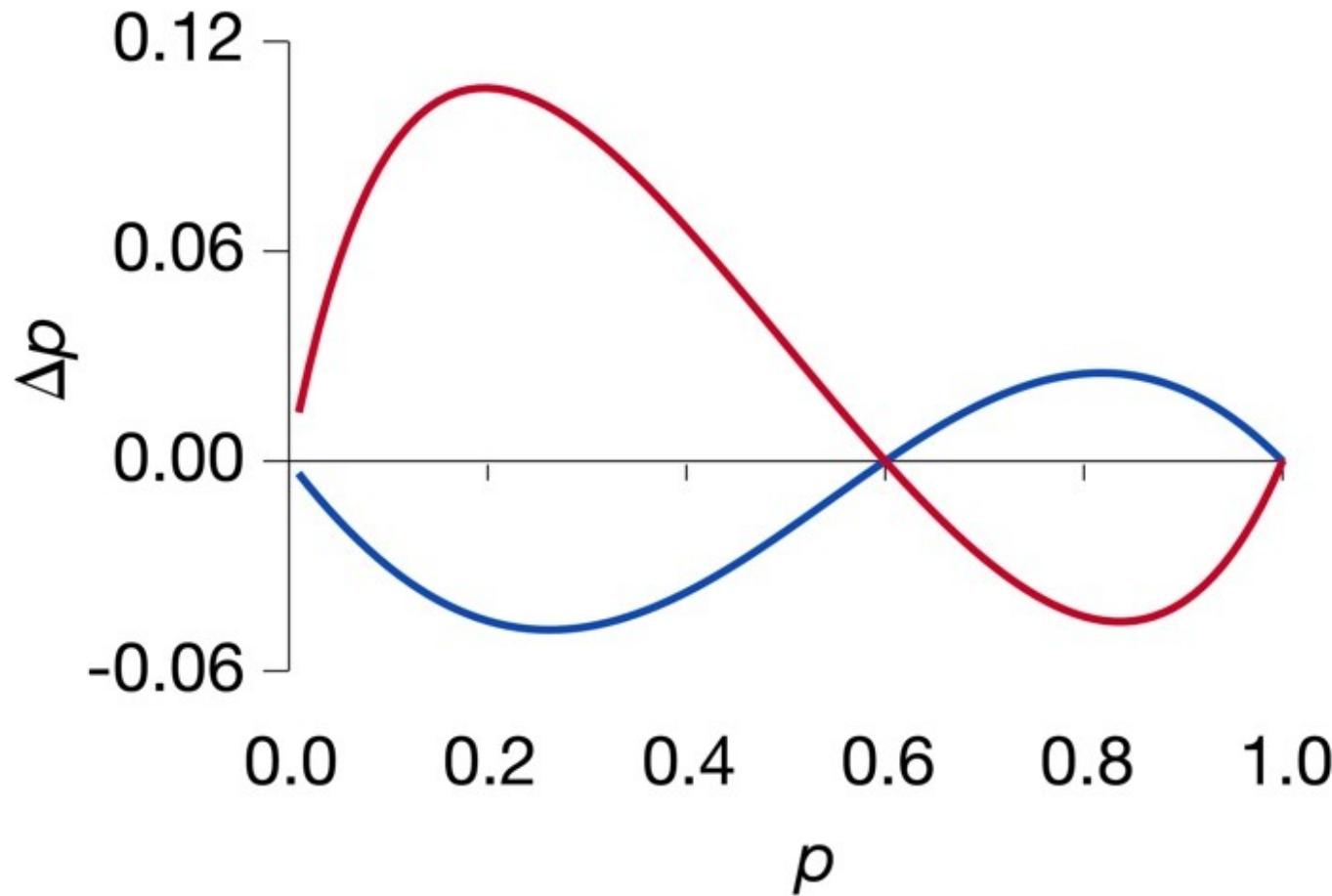
III. Underdominance

- selection against the heterozygote
- may be created by disruptive selection for two extreme phenotypes



Selection on Homozygotes/ Heterozygotes

(a) Δp as a function of p



Copyright © 2004 Pearson Prentice Hall, Inc.

Frequency-dependence

- the fitness of an allele is dependent upon its frequency in the population.
 - rare-male advantage
 - exploitation of different resources
 - cheaters in pollination systems

The Rare Male

- If females choose mates based on ‘exotic’ characteristics then a rare type may have disproportionately greater mating success.
- If females have ‘resistance’ to male charms, then they are poorly ‘defended’ against the rare male.

Resource-Specialization

- increase in frequency of an allele coding for a specialized resource creates tougher competition.
- favours types specialized on alternative resource.
- Example: ‘early’ and ‘late’ morphs in crowded fly populations.

Drosophila resource partitioning

(Borash et al.)

- “Earlies” feed rapidly, metabolize large quantities of food inefficiently.
 - create pool of toxic waste metabolites (ammonia) at surface.
- “Lates” feed and grow slowly, metabolize slowly.
 - slower metabolism and low food uptake = higher ammonia tolerance.

Elder Flower Orchid



© 2004 B.D. Ripley

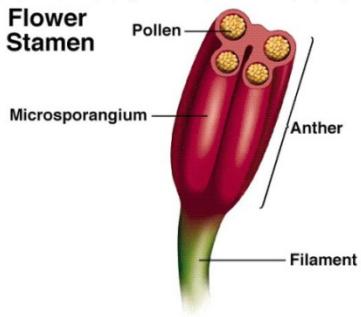


Cheating Orchids

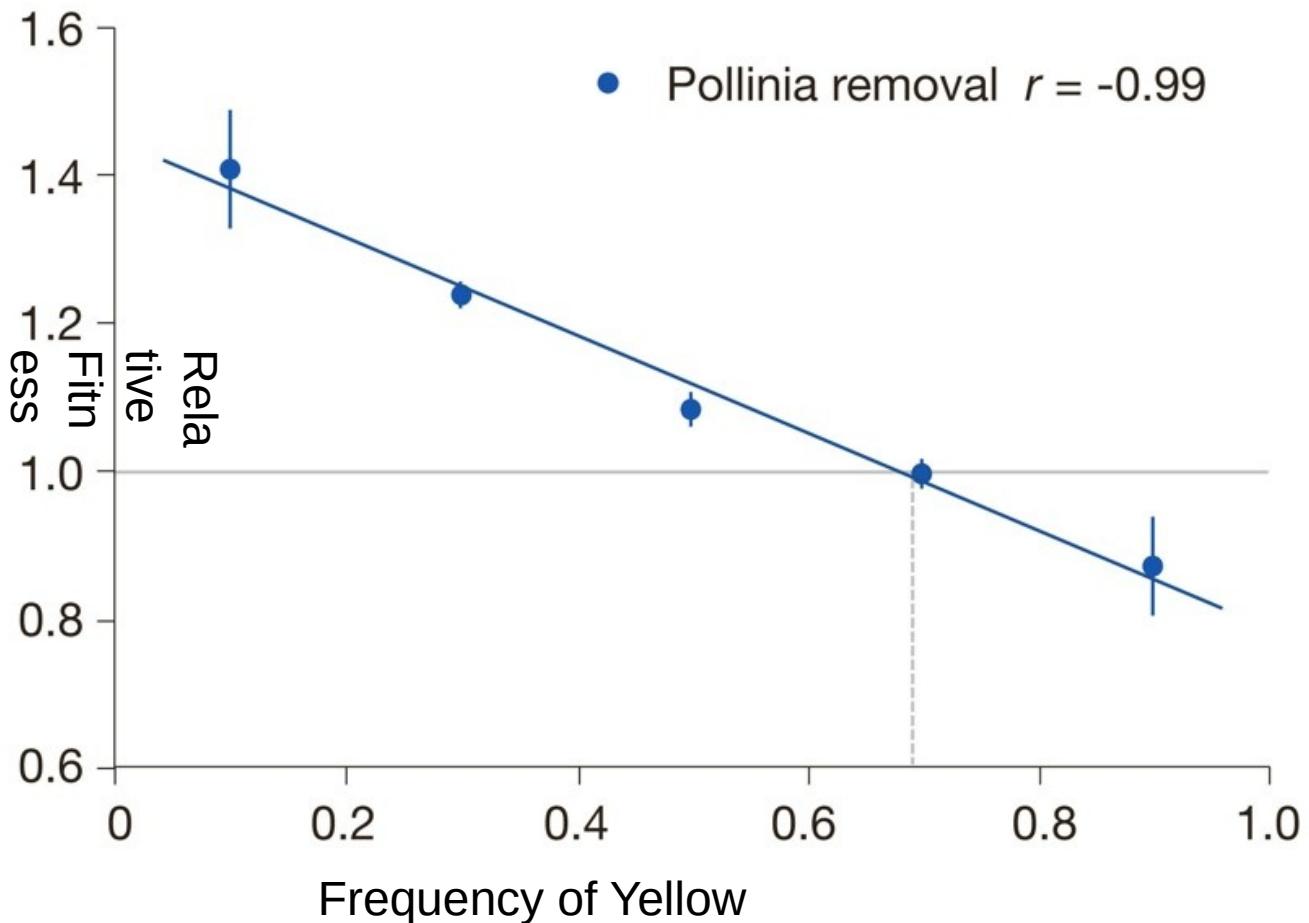


- orchid offers no nectar reward to bumble bees.
- flower has both purple and yellow forms.
- bees visit flowers and are disappointed
 - they avoid that colour next time, going to the other one.

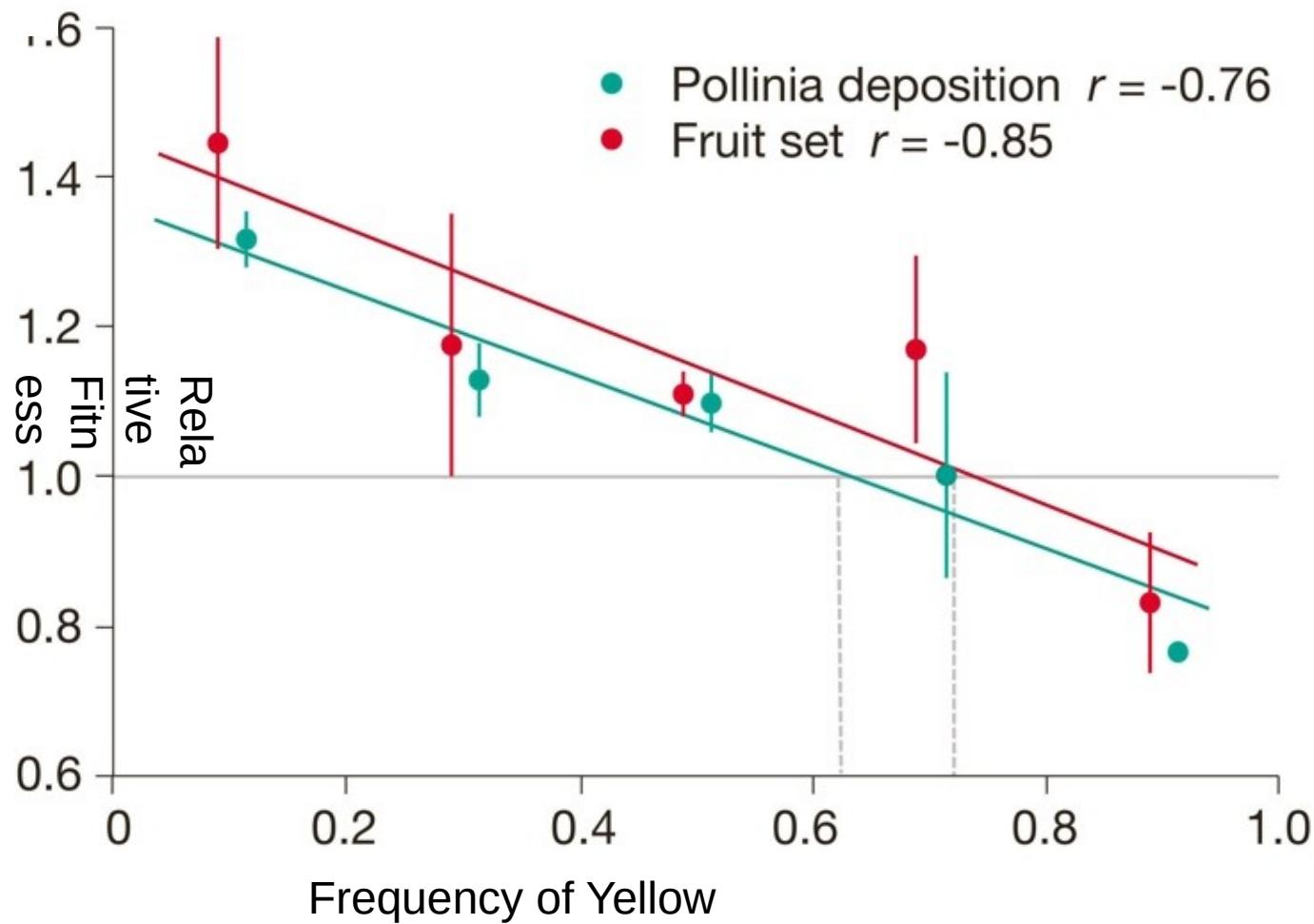




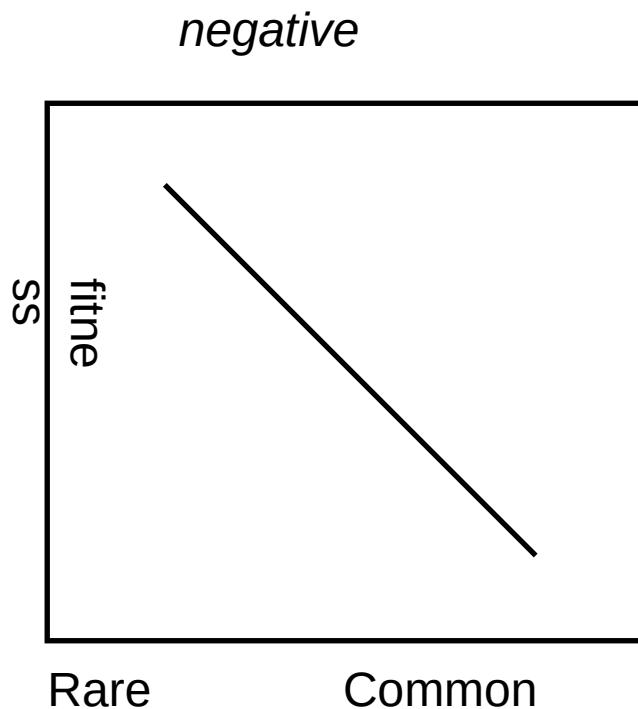
Male Fitness



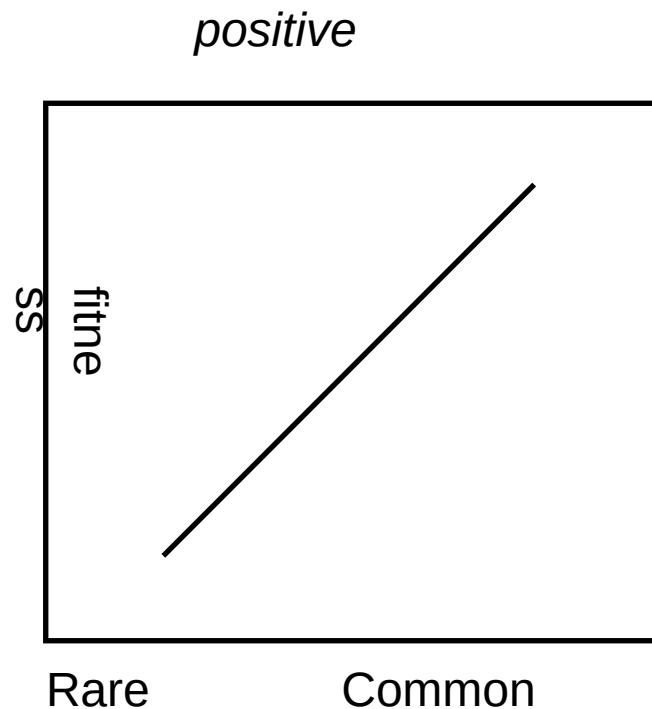
Female Fitness



Negative Frequency-Dependence can Maintain Polymorphism

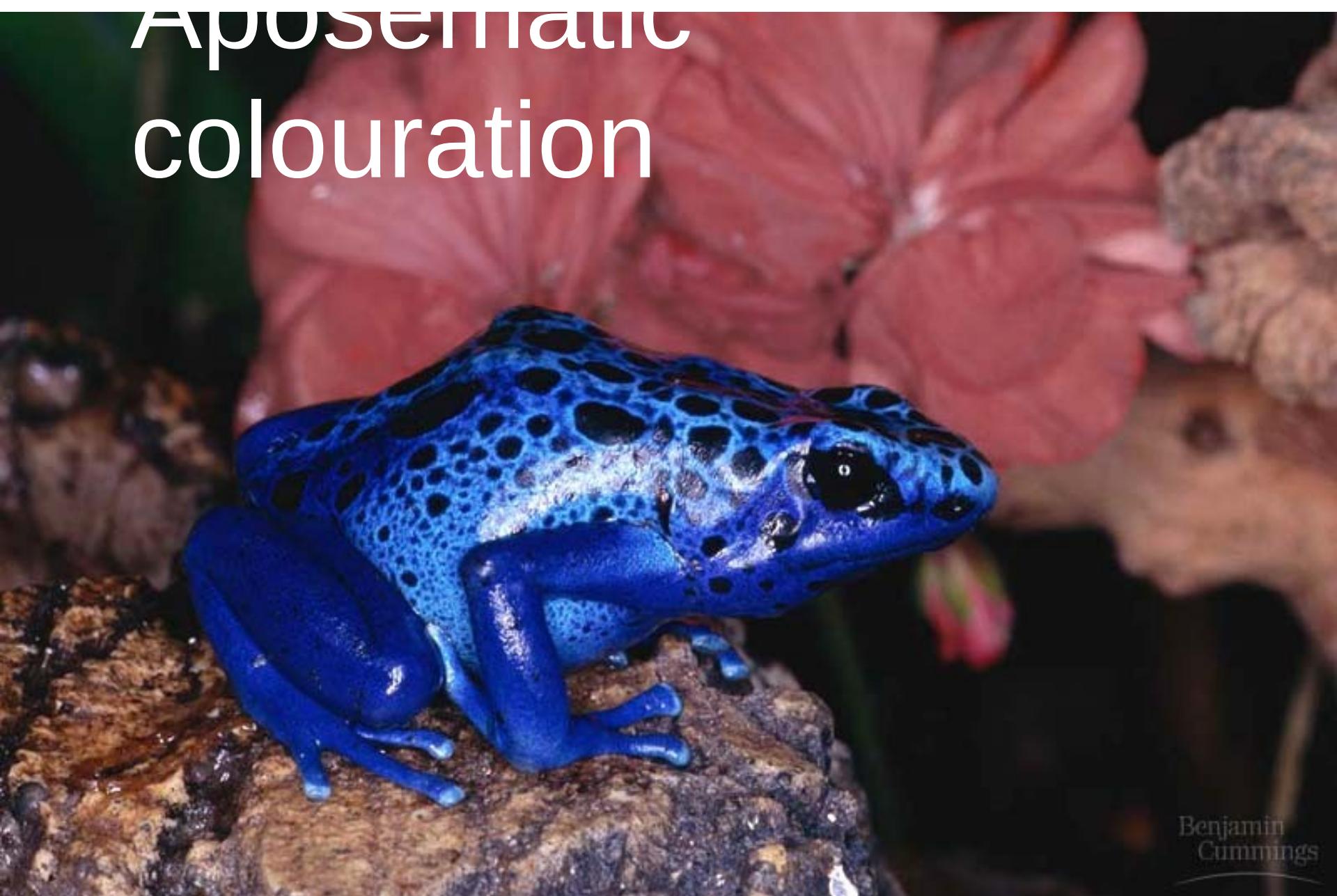


drives frequency
up when rare; down
when common

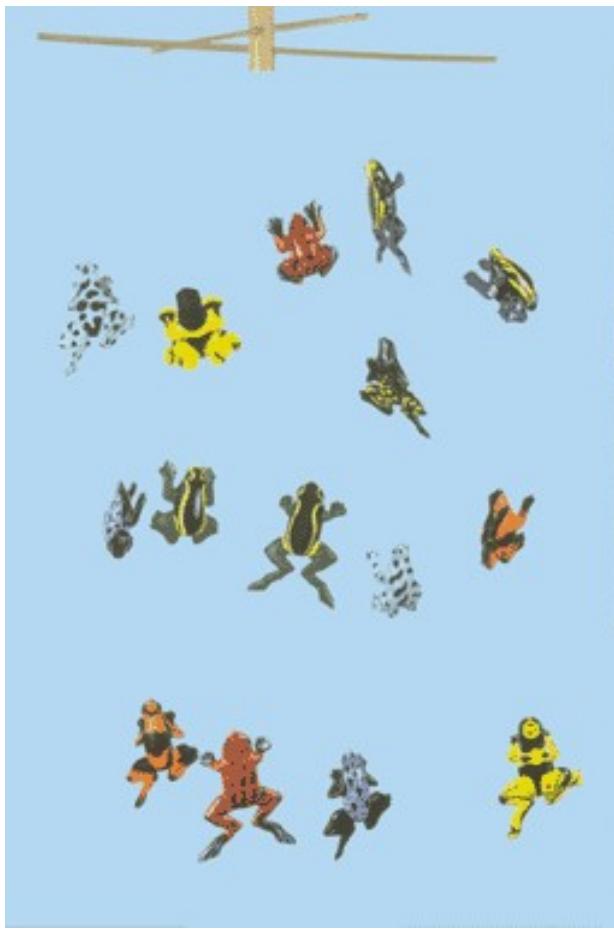


selection reinforces
frequency

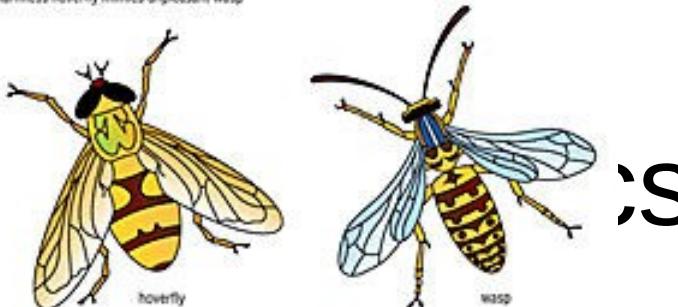
Aposematic colouration



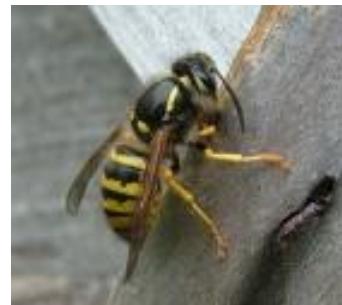
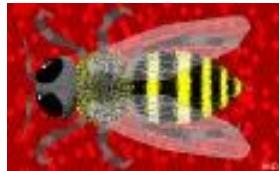
Benjamin
Cummings



harmless hoverfly mimics unpleasant wasp



- Batesian mimicry: the mimic is palatable; the model is toxic.
- Mullerian mimicry: species cooperate to form a unified 'search image' for predators





Monarch / Viceroy species pair considered a classic case of Batesian mimicry (J. Brower,
Evolution 1958)
(i.e., *monarch distasteful to toxic / viceroy yummy*)

But more recent work by Ritland and Brouwer (1991) suggests a Mullerian association.

Industrial Melanism

- Peppered Moth (*Biston betularia*)
- Commonly Found in England
- Two colour morphs
 - “White” morph or Typica
 - “Black” morph or Carbonaria
- Genetics- One major locus with Carbonaria dominant to Typica

Industrial melanism contd..

- Prior to 1850s, most moths caught in England were white morphs
- By 1900s, most moths found were the black variety
- WHY?

A conspicuous correlation with Industrialization

- Moths typically rest on lichen covered trees
- White morph well camouflaged



Industrial melanism contd..

- With industrialization, soot from the factories covered the trees.
- Lichens were killed
- Tree trunks became black
- The white morph was no longer camouflaged
- The black morph was camouflaged

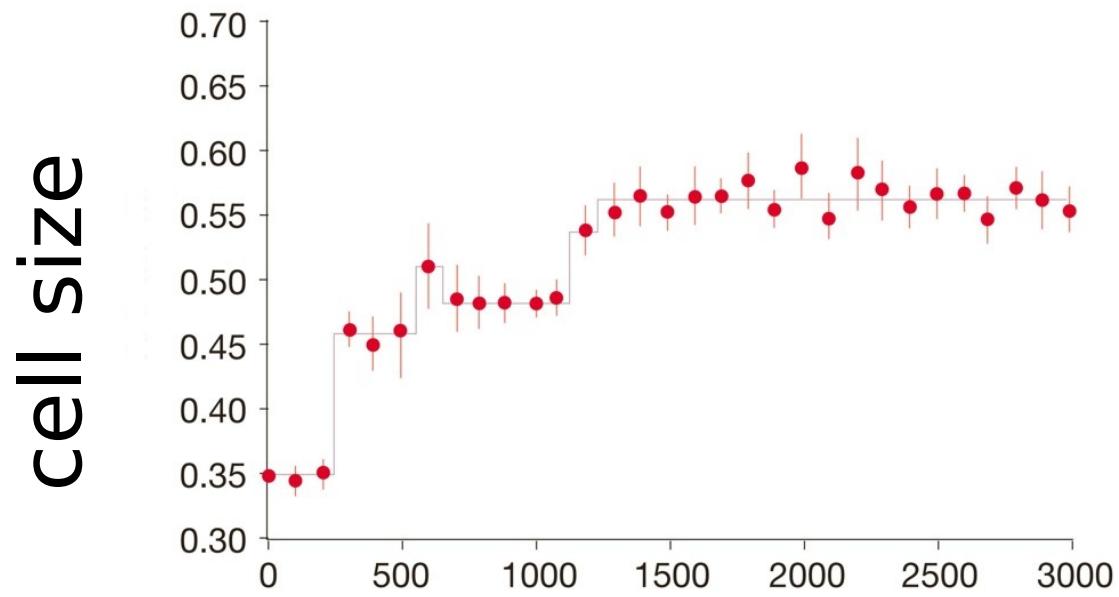


Kettlewell's experiments

1. The frequency of black and white morphs should differ in industrial and non-industrial areas.
2. Lab experiments using birds and different coloured hiding places for the moths.
3. Mark-recapture experiments in the wild using the two morphs.



Mutation



Copyright © 2004 Pearson Prentice Hall, Inc.

generations

- Experimental evolution in asexual bacteria relies on mutation as the sole source of evolutionary change.

Recurrent Mutation

- selection will eliminate unfavourable alleles, ultimately.
- but mutation continuously regenerates them.
- what is the equilibrium frequency of a recurring mutation?
 - i.e., the balance point between mutation and selection

Calculating the $\mu:s$ balance

- suppose A is a dominant harmful mutation arising at a rate μ and found at frequency p .
- selection acts against A at a value s .

genotype	AA	Aa	aa
fitness	ω_{11}	ω_{12}	ω_{22}
	$1-s$	$1-s$	1.0

- a proportion of the genes in the population are a , which is $= 1 - p$
- the rate of creation of new mutants is $\mu(1-p)$
- the rate of loss of mutants is ps .

at equilibrium (p^*), gain = loss

$$\mu(1-p^*) = p^*s$$

or approximately...

$$p^* \approx \mu / s$$

- if the allele is dominant and lethal, then $s=1$ and $p^* = \mu$.
- for a recessive mutation, the equations simplify to $p^* = \sqrt{(\mu / s)}$

note: the equation for a recessive is a very crude

- the frequency of the mutant is a rough guide to the mutation rate and selection against it.
- rough calculations:

dominant

$$\begin{aligned} p^* &= 10^{-6} / 10^{-2} \\ &= 10^{-4} \end{aligned}$$

recessive

$$\begin{aligned} p^* &= \sqrt{10^{-6} / 10^{-2}} \\ &= 10^{-2} \end{aligned}$$

for haploid, asexual organisms, like *E. coli*, all variation

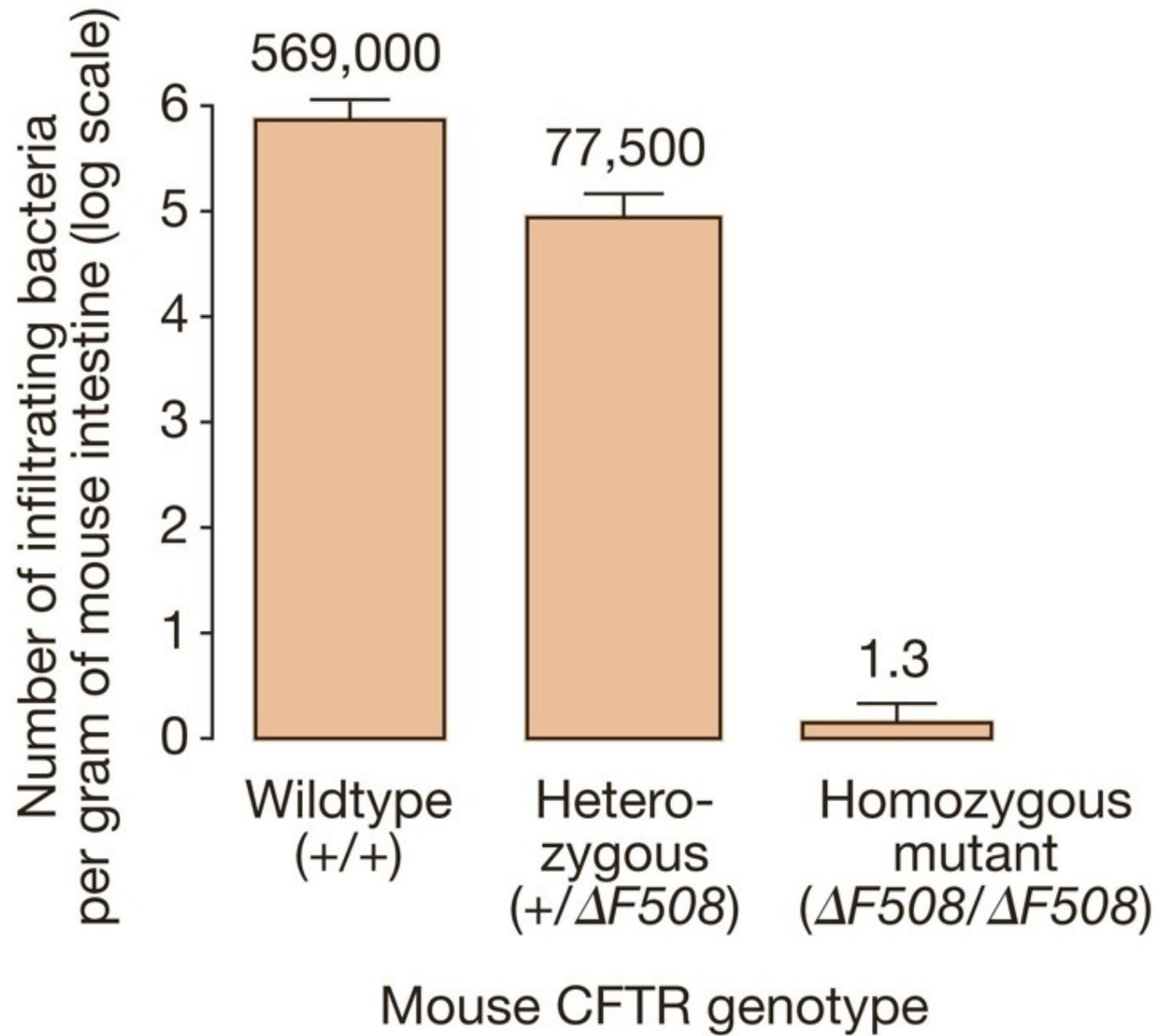
Mutants

- Spinal muscular atrophy
 - a cV mutation at *telSMN* at frequency about 0.01 in Caucasian populations.
 - devastating neuromuscular disease; *s* est'd at 0.9
 - if the equilibrium frequency = $\text{SQRT}(\mu/s)$, then
 - $0.0001 = \mu / 0.9, \dots$ or $\mu = 9 \times 10^{-5}$
- Wirth et al. estimated the spontaneous mutation rate from 340 individuals, based on 7 *de novo* cases at 1.1×10^{-4} .

CFTR

- cystic fibrosis occurs at a frequency of 0.01%
 - cVII disease requires a mutation rate of 4×10^{-4} .
 - observed rate is 7×10^{-7}
- Pier et al. (1998) suggested that overdominance might be the root of high CFTR occurrence:
 - heterozygotes for CFTR deficiency may be better at resisting typhoid fever (a *Salmonella* disease).
 - using the mouse model, wildtype, heterozygote, and homozygous $\Delta F 508$ (a CFTR allele), they showed a higher degree of resistance in mutant homozygotes and pa

(a)



Eugenics

- Eugenics seeks to alter fertility patterns for the betterment of the society, race, or humankind.
 - typically this means sterilizing less desirable individuals through some means.
 - in the extreme, e.g., Nazi Germany, this meant genocide.
 - in milder forms, this has been done through ‘incentives’.
- Many major nations practiced compulsory sterilization of the ‘infirm’ or undesirable until quite recently.

A Liberal Thing?



- The U.S.A: 65,000 sterilized people in 33 states until '70s.
 - mentally retarded, gay, blind, deaf, epileptic, alcoholic, homeless, first nation.
- Sweden: compulsory sterilization of 63,000 until 1976.
 - mentally retarded, mixed race (e.g., gypsy), socially undesirable, criminal.
- Japan: 16,500 forced sterilizations up to 1992.
 - schizophrenia, manic depression, epilepsy, alcoholism, drug addiction, physical handicaps

Courtesy of Marty Pernick

Printed for Alexander H. Revell & Co., New York.

JEWEL, LEVINE & MERRIFIELD
LA SALLE NOW | ALL Seats **25c**
9 A.M. TO 11 P.M. CONTINUOUS

Dr Harry J. Haiselden

in

"THE BLACK STORK"

— By Jack Bait —

A vivid pictorial drama that tells you why Dr. Haiselden is opposed to operating to save the lives of defective babies.

"The law of heredity winds like a red thread through the family history of every criminal, of every epileptic, moronic and insane person. Should we sit still and witness our civilization go into decay and fall to pieces without raising the cry of warning and applying the remedy?" — Dr. August Forel, Zurich.

NO CHILDREN ADMITTED.



Courtesy of Marty Rennick

ENGLISH HERITAGE

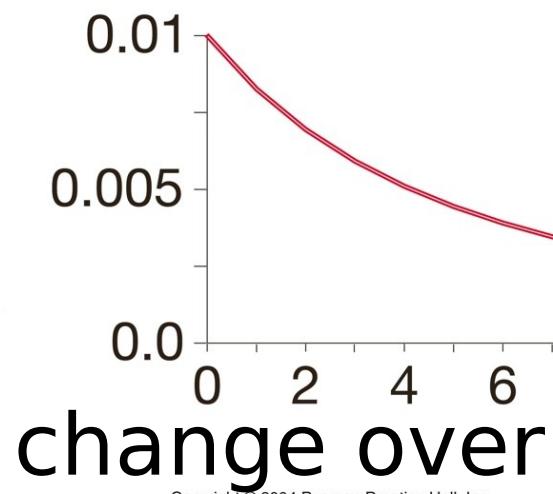
Sir
RONALD
AYLMER
FISHER
1890-1962

Statistician and
Geneticist
lived here
1896-1904

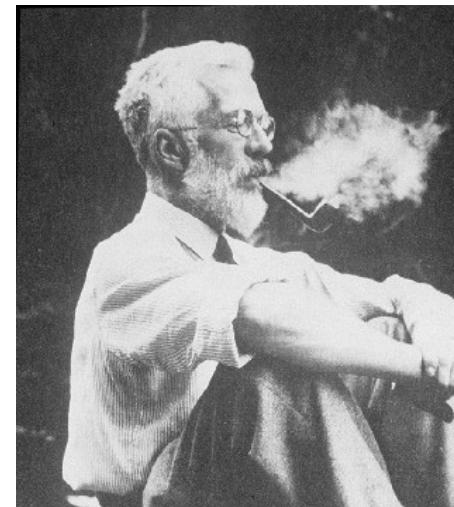


The Punnett / Fisher Calculation

- The offending trait is recessive, a frequency 0.01 homozygous carriers
- Selection is against homozygotes at $s=1$ (because they are sterilized), then the frequency would decline from 100 per 10,000 affected to:
 - 83 per 10,000 in one round of sterilization.
 - 25 per 10,000 after 10 generations.



Copyright © 2004 Pearson Prentice Hall, Inc.



Migration & Drift



©Disney Enterprises, Inc./ Pixar Animation Studios





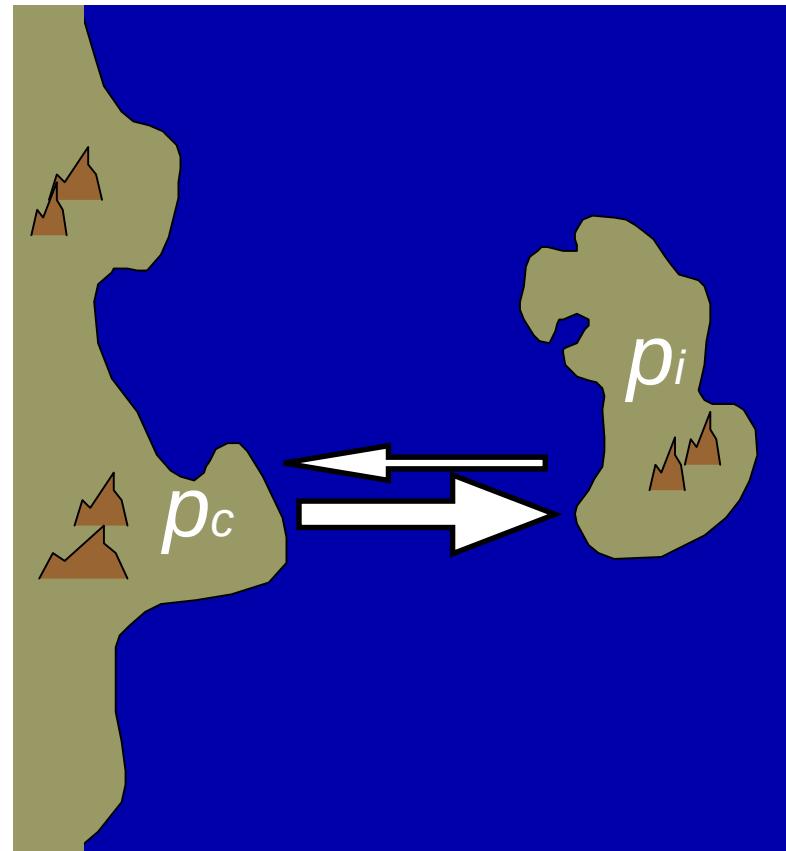
- **Migration = Gene Flow**
 - the movement of alleles between putative populations.
 - gene flow binds biological species together.
- **Genetic Drift = Sampling Error**
 - an inevitable feature of finite populations.
- Both forces are agents of evolutionary change.
- Both impart deviations from H-W equilibrium.

Migration

- a byproduct of population subdivision
- for gene flow to occur:
 - migrants must disperse to a new population at some rate (m)
 - they must reproduce
 - be suitably adapted to the conditions in the new population.
 - not be selected out (e.g. by predators)

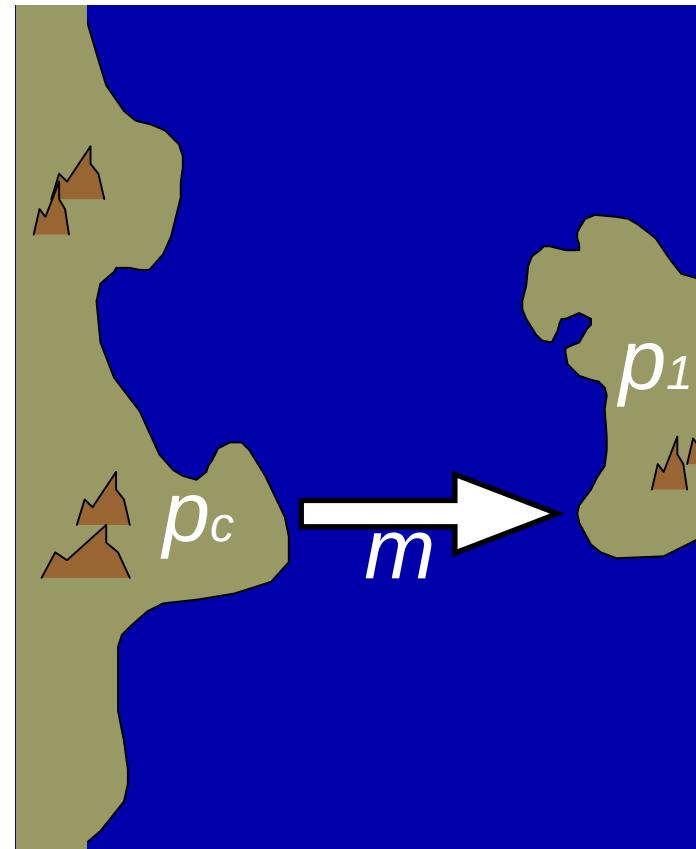
The Island Model

- population is polymorphic at the A locus.
- how does frequency of A_1 change due to migration?



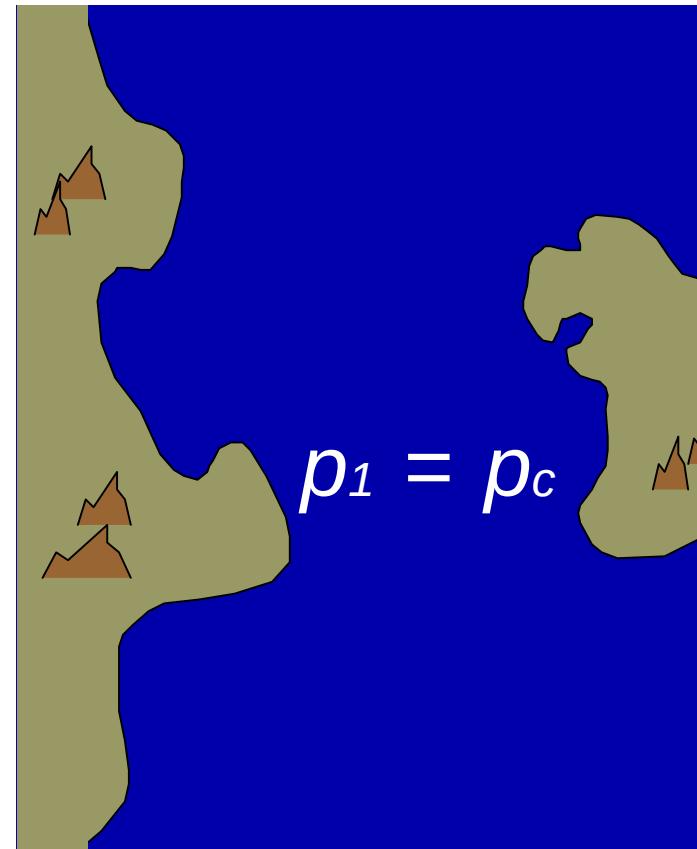
- A_1 is at frequency p_i before migration
- m = fraction of migrant individuals.
- $1-m$ = fraction of resident individuals
- frequ. of A_1 in next generation is:

$$(1-m)(p_i) + (m)(p_c)$$

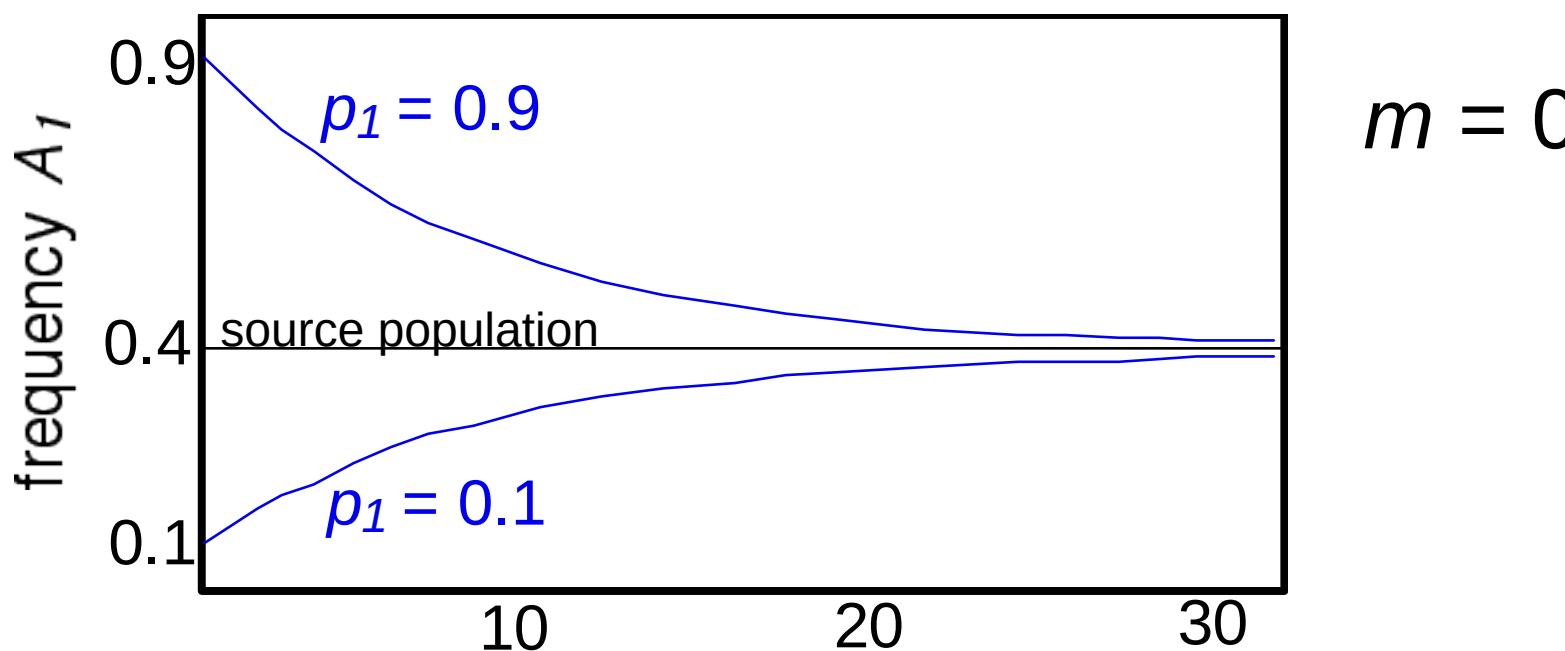


$$\Delta p_i = m(p_c - p_i)$$

- at equilibrium ($\Delta p_i = 0$) the frequ. of A_1 is always the same in both populations.
- the rate of homogenization depends upon:
 - 1) the level of migration.
 - 2) the difference in frequency of A_1 between the populations.
 - 3) the absence of selection



Gene flow rapidly homogenizes populations



at $m=10\%$, it takes about 30 generations to equalize gene frequency

The Lake Erie Water Snake



Northern Water Snake
Nerodium sipedon sipedon



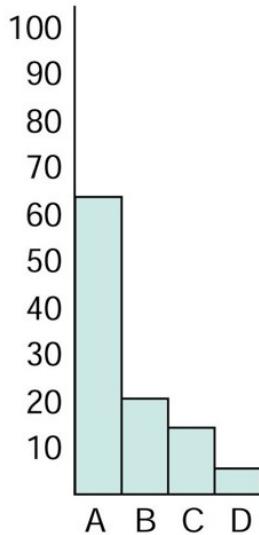
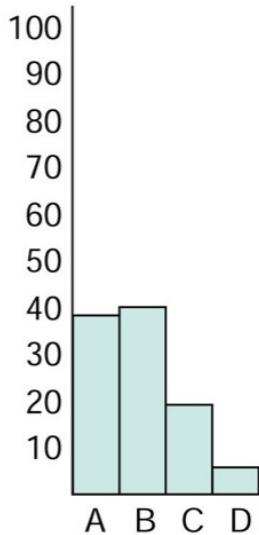
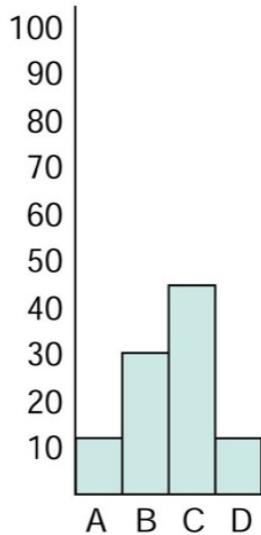
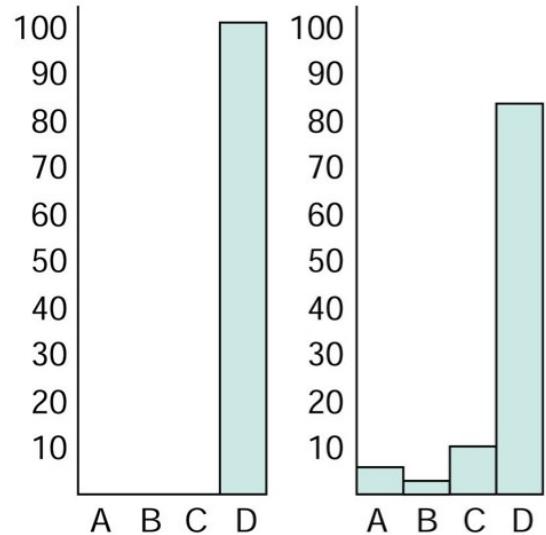
L. Erie Water Snake
Nerodium sipedon insulare



Selection on Banding Pattern

mainland

islands



↑
no bands ↑
strong bands

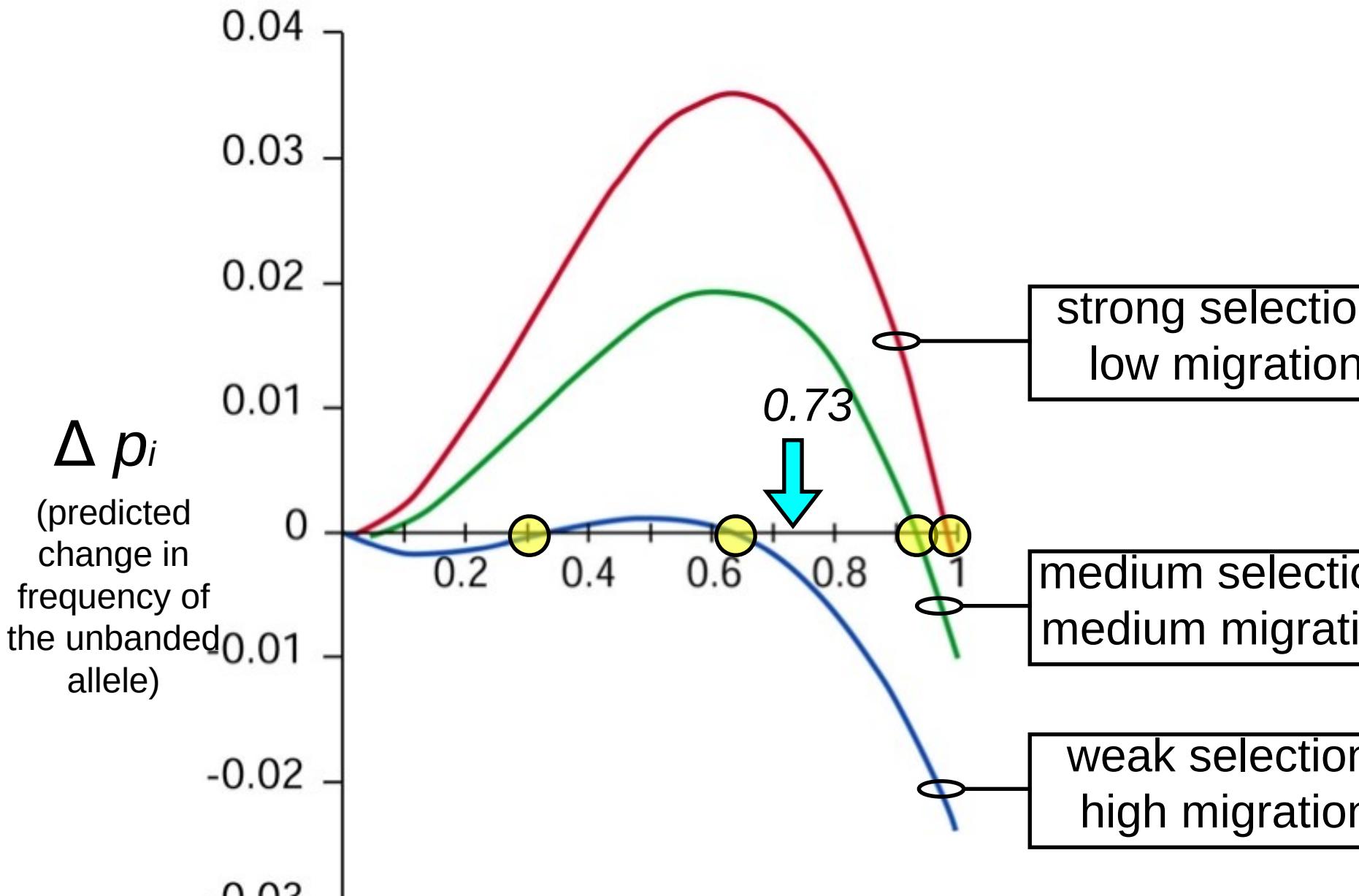


- banding determined by a single locus two allele system.
- King (1993) used mark & recapture experiments to document selection differentials.

Assignment: Using real data

- relative fitness of banded snakes on islands due to predation pressure = 0.78 to 0.90
- island population $\approx 10^3$
- molecular genetic estimate of 13 migrants / year (about 1% of population)
- calculate equilibrium frequency of the unbanded (recessive) allele.

Modeling Migration, Selection

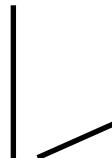


Possible Reasons for Inaccuracy

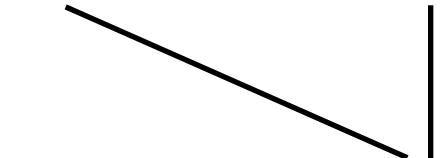
1. mark-recapture underestimates selection against banded (other characteristics important in natural or sexual selection).
2. parameter estimates variable, based on limited sampling.
3. nature & human intervention changing the habitat.
4. population sizes fluctuate.

Population Subdivision

	AA	Aa	aa
subpop. 1	$(0.3)^2 = 0.09$	0.42	$(0.7)^2 = 0.49$
subpop. 2	$(0.7)^2 = 0.49$	0.42	$(0.3)^2 = 0.09$
Mean	0.29	0.42	0.29



frequ. A = 0.5



frequ. a = 0.5

The Wahlund Effect

	AA	Aa	aa
Mean Observed	0.29	0.42	0.29
H-W Frequencies	0.25	0.50	0.25

- a population with subdivision always displays a deficit of heterozygotes in the equilibrium calculation.
- occurs even though the two populations are themselves at

Wahlund Effect

Cont'd

- one needs to know the population structure before believing equilibrium frequencies.
 - *or, a deficit of heterozygotes may signal subdivision.*
- when two distinct populations fuse, the proportion of homozygotes declines.
 - fitness loss due to deleterious homozygous recessive variation (e.g., many genetic diseases) will decline.



Wiarton Willy:
did not see shadow

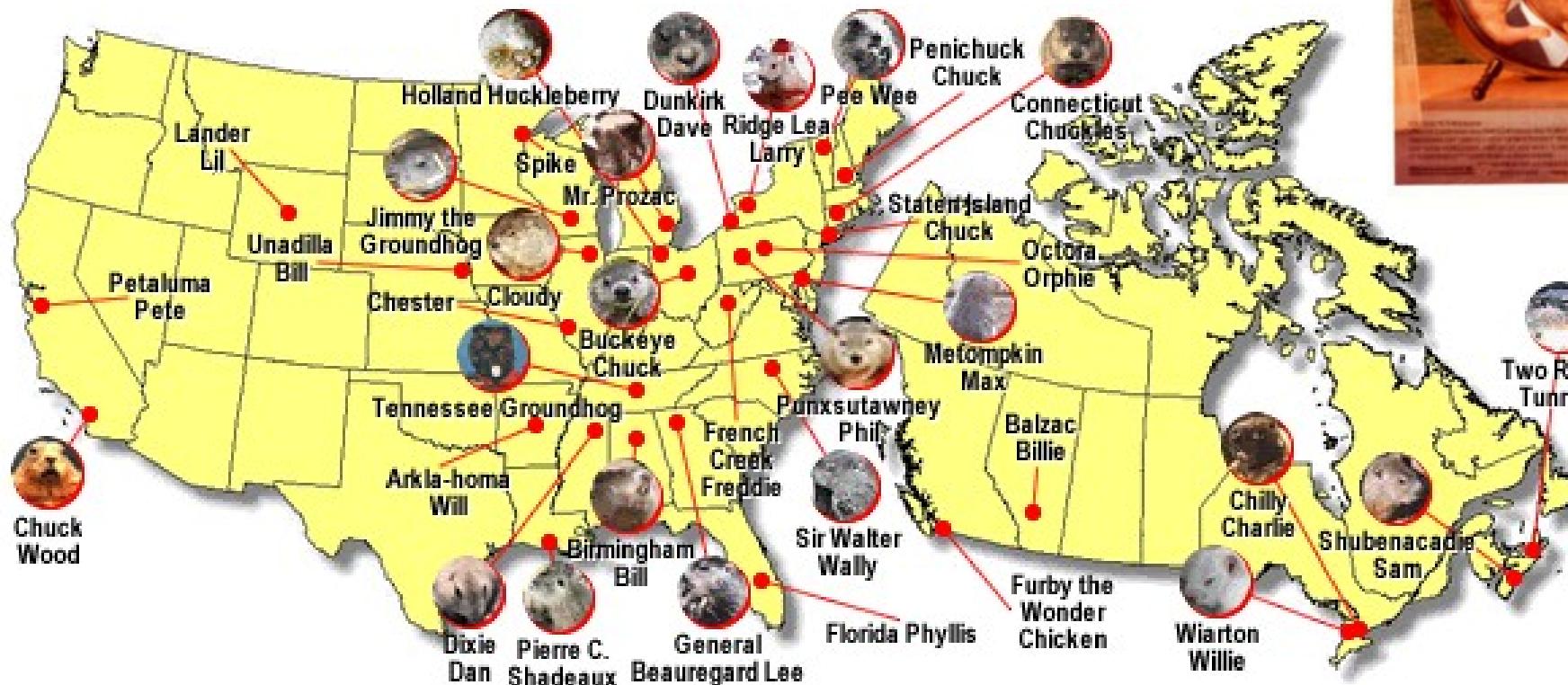


Punxsutawney Phil:
did not see shadow



Shubenacadie
Sam:
did not see shadow

Who you gonna trust?



<http://www.groundhogsday.com>

[NEWS](#)[BUSINESS](#)[SPORTS](#)[ARTS & ENTERTAINMENT](#)[WEATHER](#)[HEALTH & SCIENCE](#)[CBC ARCHIVES](#)[KIDS](#)[TEENS](#)[PROGRAM GUIDE](#)[E-MAIL NEWSLETTERS](#)[SERVICES](#)[CONTACT US](#)[ABOUT CBC](#)[!\[\]\(8ccac51df515ea8bd3083833c2b714ea_img.jpg\) RADIO-CANADA](#)[shop](#)[!\[\]\(af2ec10aad8d2ccaf9fbfd9fe81a7655_img.jpg\) CANADA VOTES 2006](#)

Wiarton Willie suspected in double murder

Last Updated Tue, 23 Sep 2003 20:35:26

WIARTON, ONT. - An official in small-town Ontario is being accused of covering up a murder, but she says she was protecting the town from bad publicity.

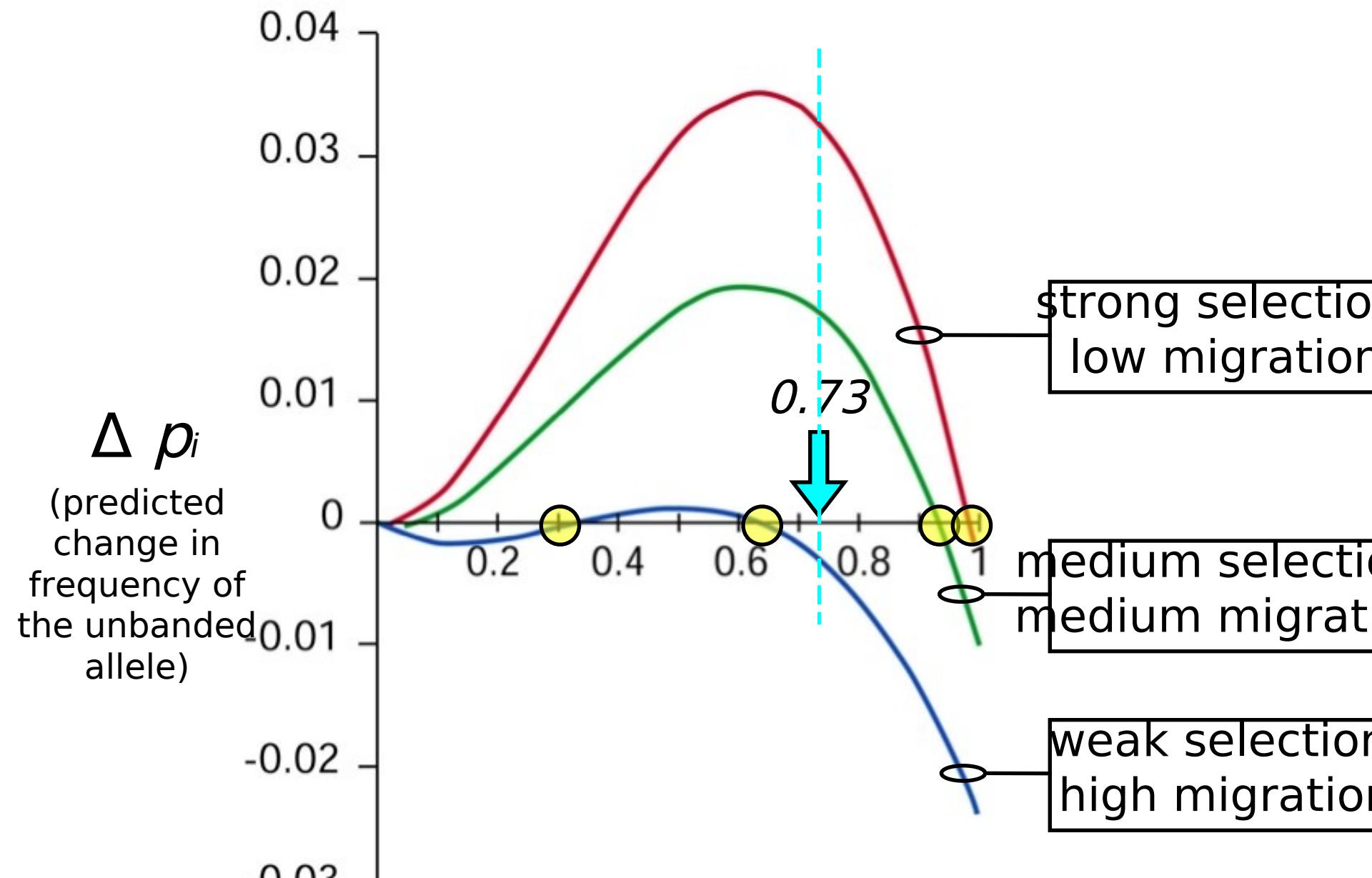
Francesca Dobbyn told Wiarton town council on Monday that she knew Wiarton Willie might have killed his two understudies, but hid the facts so the town's summer tourist season wouldn't be hurt.

Wiarton Willie is an albino groundhog that is brought out every Groundhog Day to predict the coming end of winter. He is also the town's mascot and main tourist attraction.



Wiarton Willie

The Wiarton Willie Festival held every February generates about \$750,000 for the town. That makes Willie a powerful rodent.



Population Subdivision

	AA	Aa	aa
subpop. 1	$(0.3)^2 =$ 0.09	0.42	$(0.7)^2 =$ 0.49
subpop. 2	$(0.7)^2 =$ 0.49	0.42	$(0.3)^2 =$ 0.09
Mean	0.29	0.42	0.29



frequ. A = 0.5



frequ. a = 0

The Wahlund Effect

	<i>AA</i>	<i>Aa</i>	<i>aa</i>
Mean Observed	0.29	0.42	0.29
H-W Frequencies	0.25	0.50	0.25

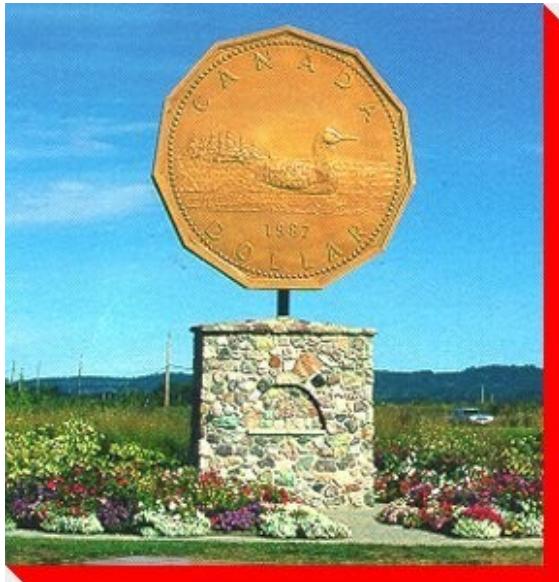
- a population with subdivision

always displays a deficit of heterozygotes in the equilibrium calculation.

- occurs even though the two populations are themselves at

Wahlund Effect

- one needs to know the population structure before believing equilibrium frequencies.
 - *or, a deficit of heterozygotes may signal subdivision.*
- when two distinct populations fuse, the proportion of homozygotes declines.
 - fitness loss due to deleterious homozygous recessive variation (e.g., many genetic diseases) will decline.



Genetic Drift

- toss a coin 10 times
- odds of a head = $0.5 / \text{toss}$
- odds of 10 heads = $0.5^{10} = 1/1000$
- odds of 6 heads $\approx 1/5$
- toss a coin 1000 times, the chance of flipping 600 heads is a very

The Founder Effect

- part one: a new population is established from a small number of colonizers or survivors.
- part two: their gene frequencies are unrepresentative of those in their predecessor population.
- what is the chance a subsample of the originating population will *lose* an allele?



The Founder Effect

- high frequencies of otherwise rare diseases in populations founded by small number of colonists.
 - achromatopsia (total colour blindness / rod monochromy)
 - recessive, occurs at frequency <0.0001 (carriers ≈ 1/200) -- a cone defect found on cVIII.
 - 5-10% occurrence; carriers found > 30% in Pingelapese.
 - 3000 Pingelaps founded from 20 individuals from the Marshall Islands



- chance an individual founder is AA
 $= p^2$
- chance two individuals are $AA = (p^2)^2$
- chance n individuals are $AA = (p^2)^n$
- chance of total homozygosity in founder population: $(p^2)^n + (q^2)^n$

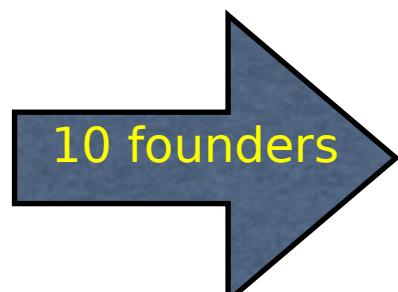
e.g., chance of losing an allele that has frequency of 10%:

- in founder population of 2: $(0.81)^2 = 0.6561$
- in population of 10: $(0.81)^{10} = 0.122$
- in population of 20: $(0.81)^{20} = 0.015$

Many alleles

- The odds of fixing any single allele by drift are simply its (frequency²)^N
- The probability of losing any allele will be the sum of the other individual fixation probabilities.

Alleles
1 @ 0.6
2 @ 0.3
3 @ 0.1



probability of losing allele 1 is
 $= 1 \times 10^{-8}$

probability of losing allele 3 is

General Consequences

- chance of losing a common allele very small, even with a strong bottleneck.
- drift is adirectional.
- chance of changing gene frequencies very large.
 - Dutch Afrikaaners arrived in S. Africa in 1652 on one ship.
 - 50% of current 2.5 million population have 20 names traceable to that ship. 1/3 white South Africans descended from 40 founders.

>8000 cases are
traceable to German
(settler) or African
names like Jacobs (wife)

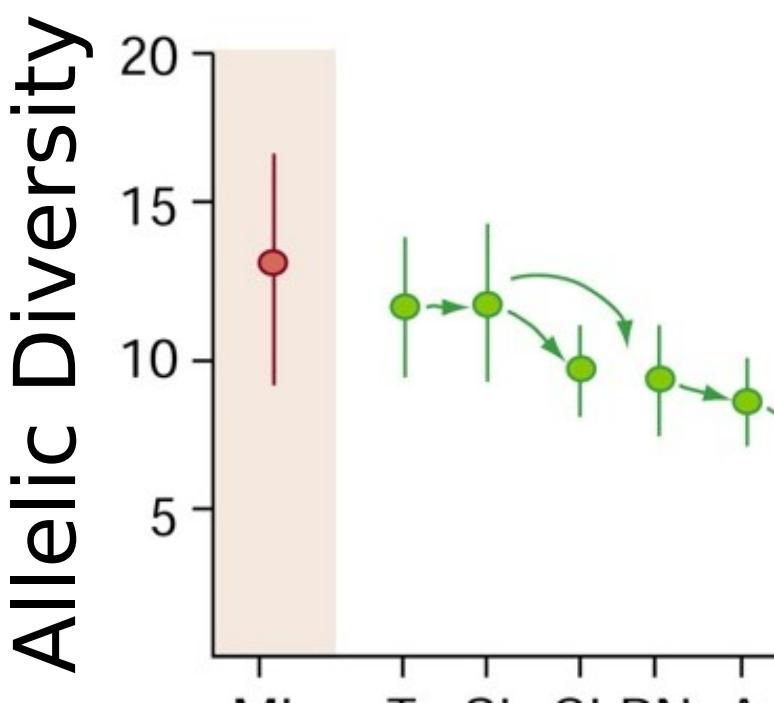
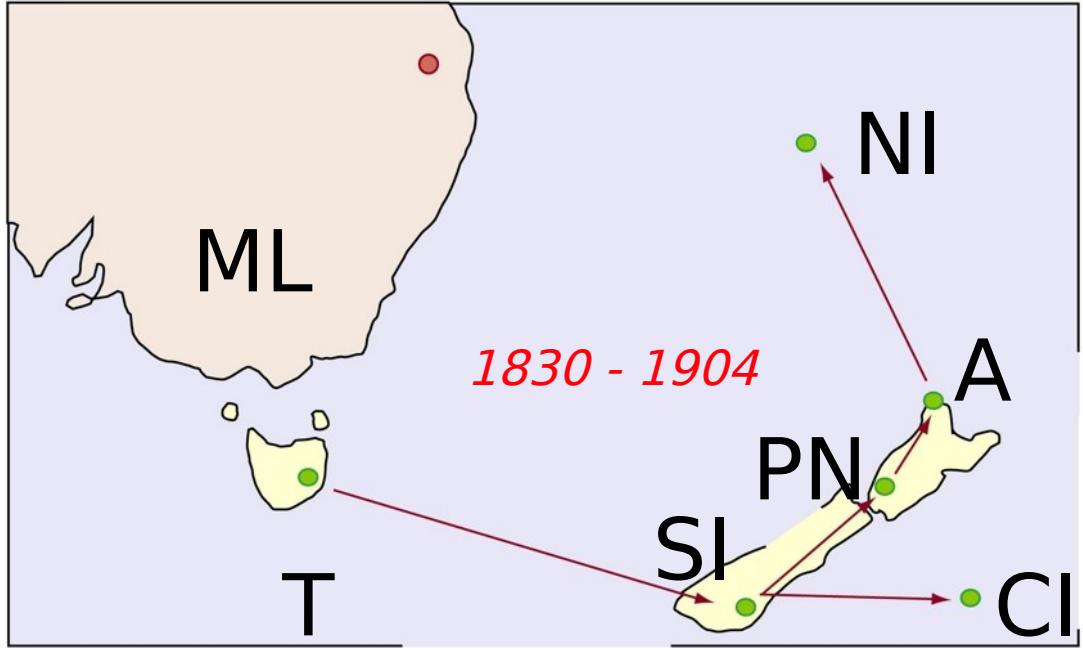
Measures of Diversity

Components of Heterozygosity

1. Allele Richness: the average number of alleles per locus in the genome.
2. Genetic Polymorphism: the fraction of loci in the genome with 2+ alleles at frequency of >0.01.

- chance of losing some rare alleles is high

(b)





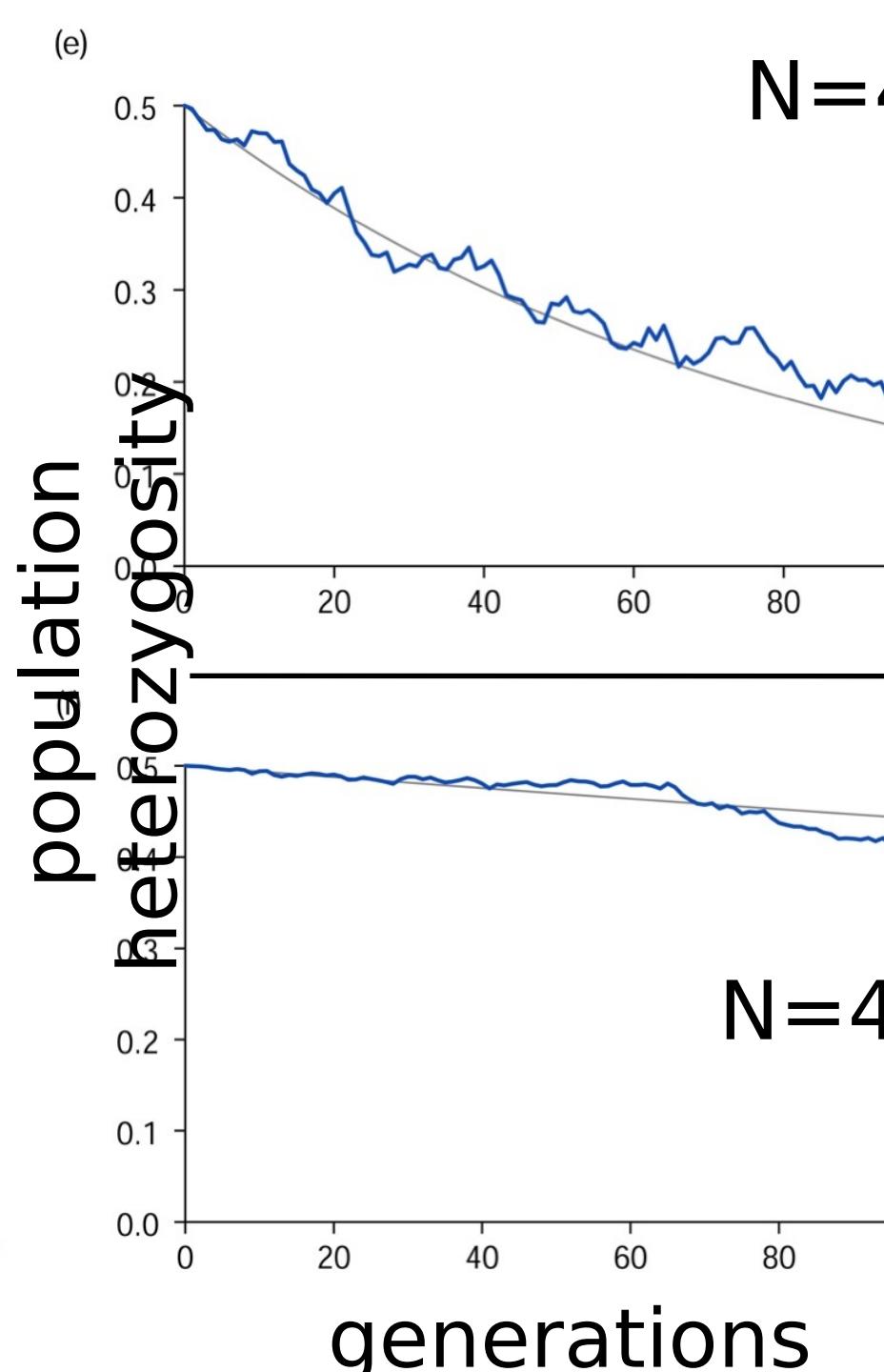
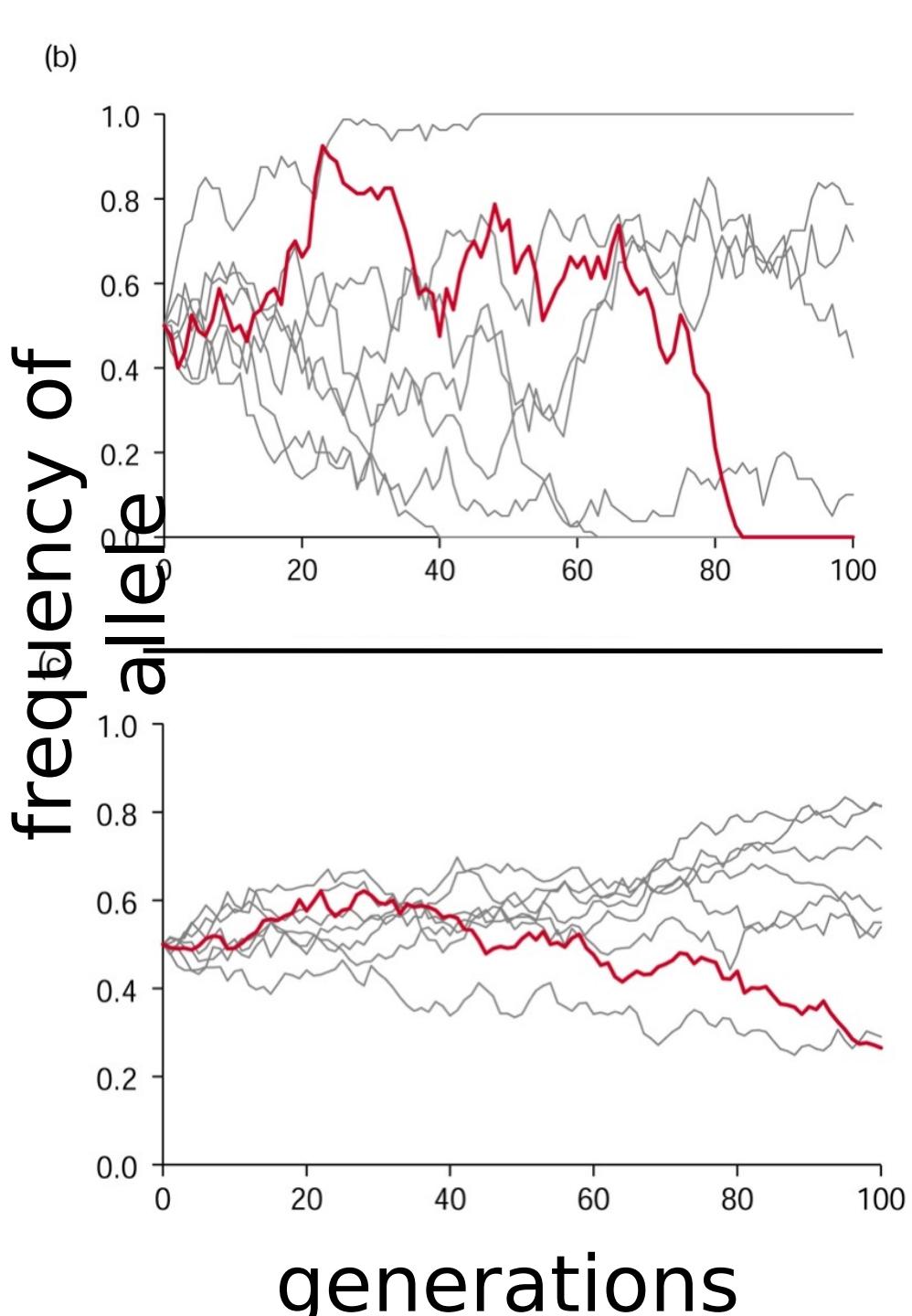
Fixation by Drift

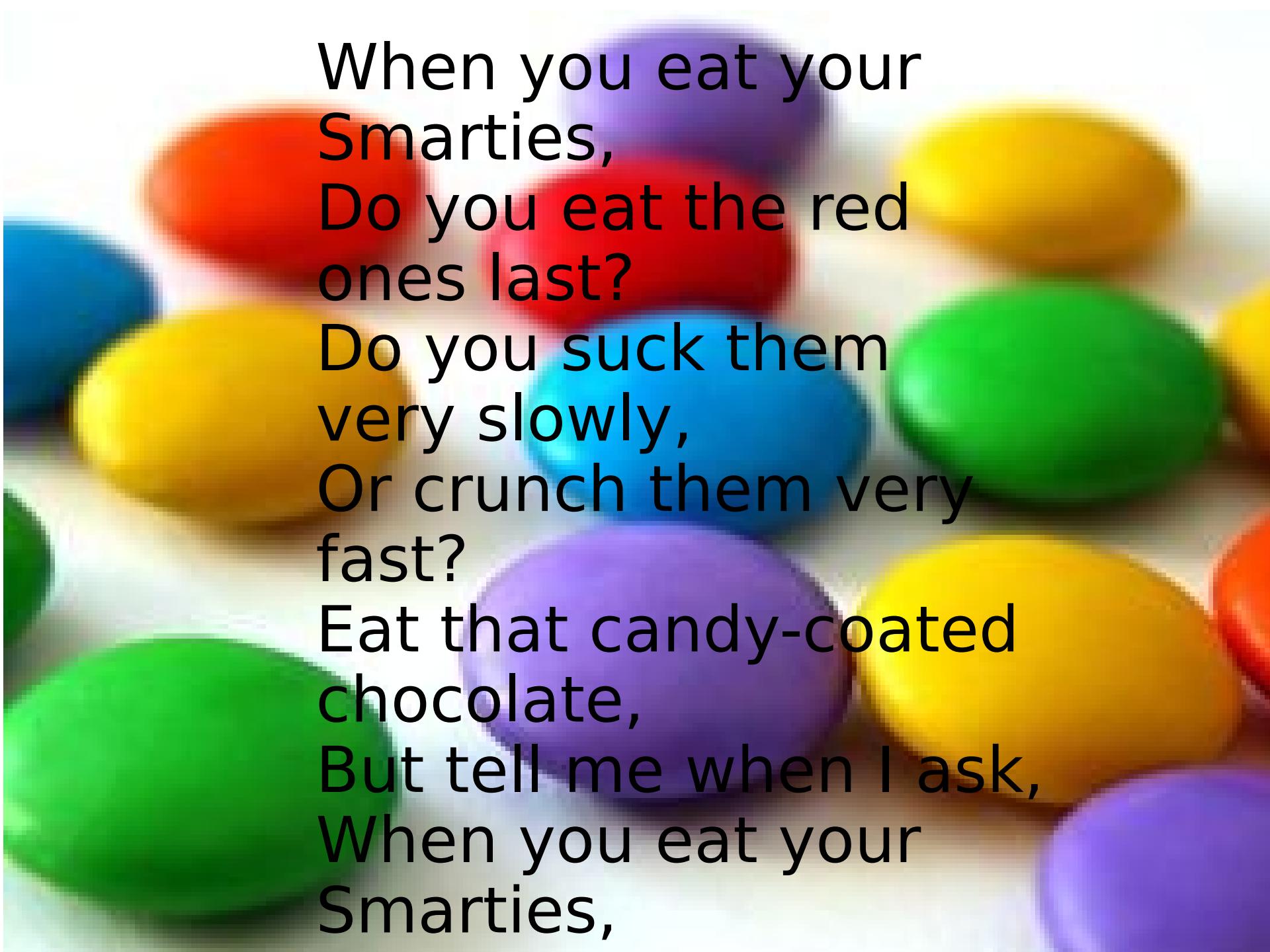
- drift is integrally related to population size.
- due to drift alone, the chance of an individual copy of a gene fixing is **$1/2N$** (if population is diploid).
- assuming there are multiple copies of the same gene (allele) then the probability of an allele fixing is: **its copy number / $2N$.**



It follows that...

- sooner or later one allele will fix due to selection or drift.
- heterozygosity -- the frequency of heterozygotes -- will decline with time.
- by drift, the number of heterozygotes in the next generation will be $H^* \{1-1/2N\}$





When you eat your
Smarties,
Do you eat the red
ones last?
Do you suck them
very slowly,
Or crunch them very
fast?
Eat that candy-coated
chocolate,
But tell me when I ask,
When you eat your
Smarties,



Smarties Trivia

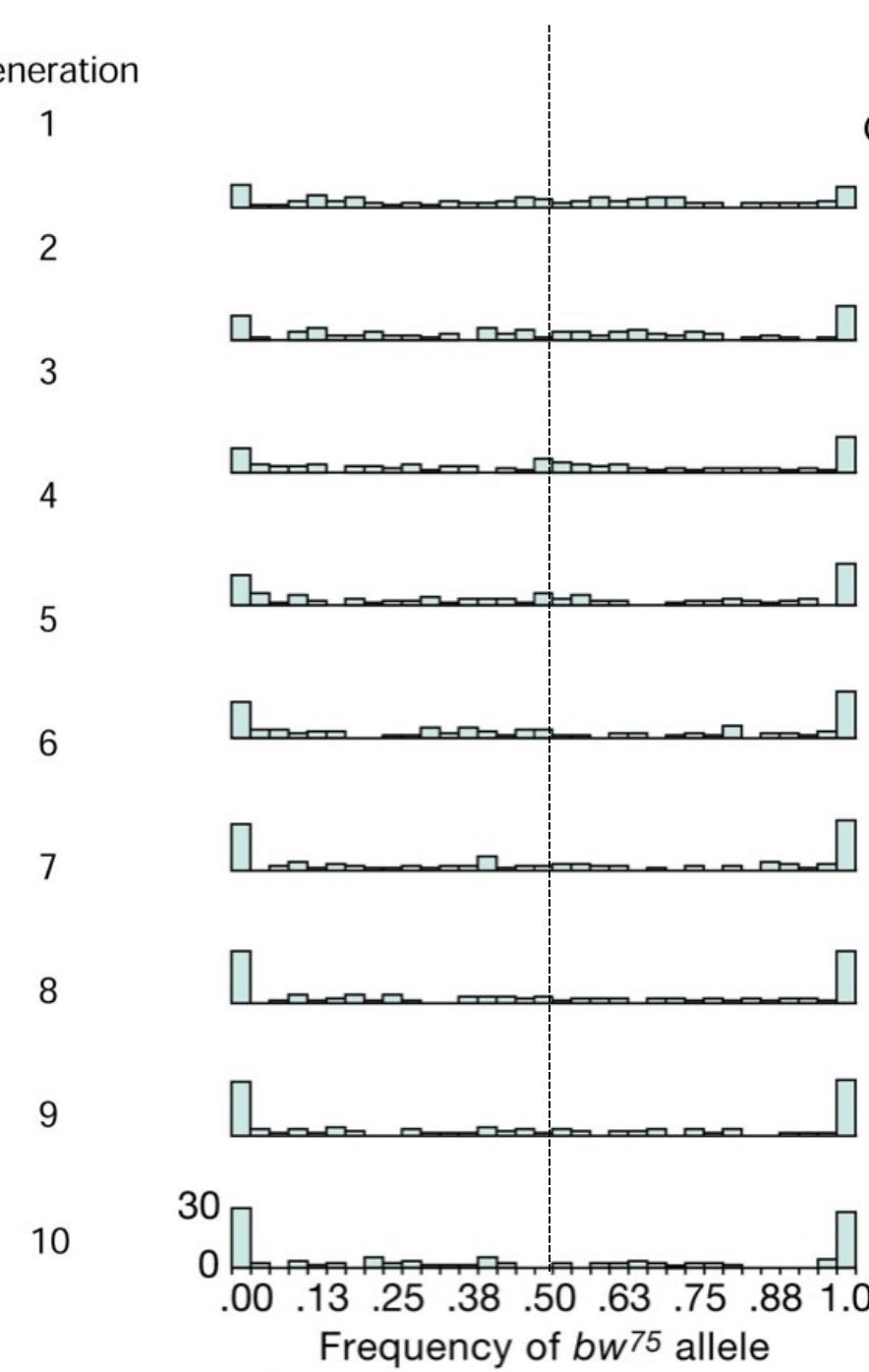
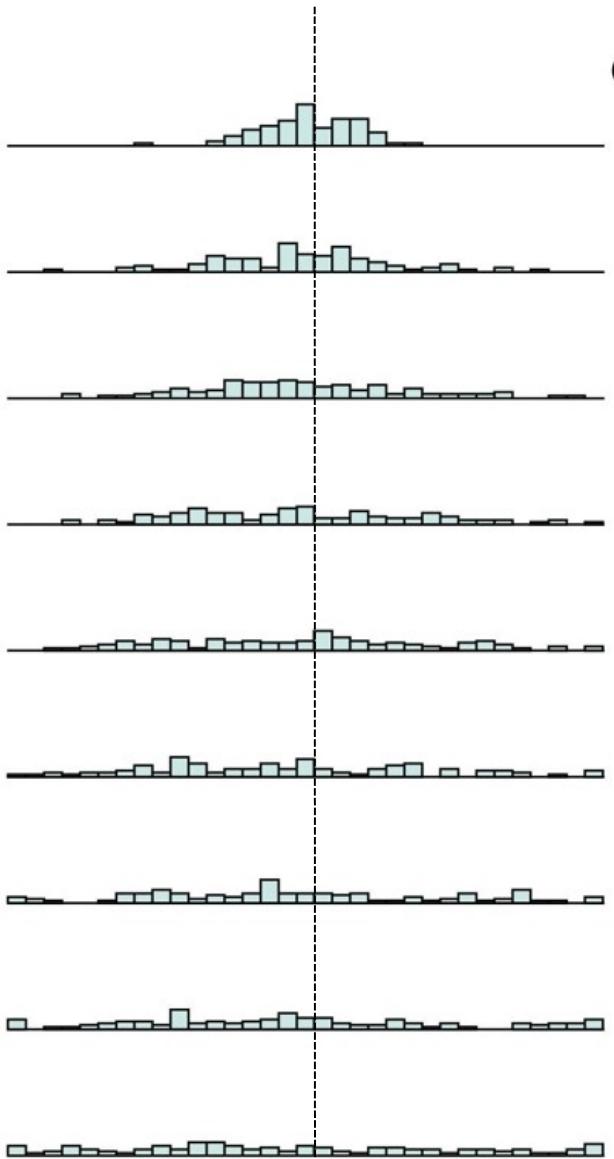
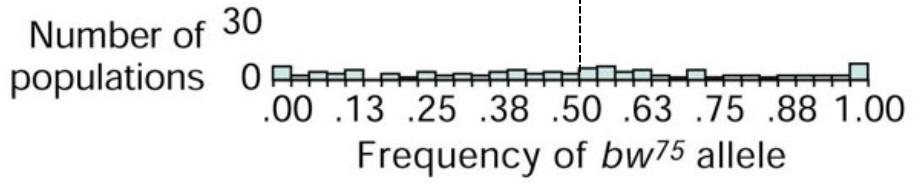
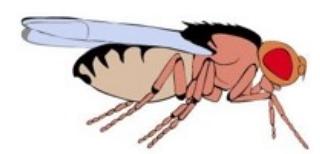
- 1937 “Chocolate Niblet Beans” Rowntree of York introduced.
- Light brown replaced by blue in 1989.
- Both orange and light brown Smarties’ chocolate flavoured in UK (never in Canada).
- 570,000 tubes / day (48/box; UK)
- over 300 tubes consumed every minute in UK.
- Nestlé claims that Canadians eat enough Smarties each year to circle Earth 350x.
- On [25 October](#), 2003, Kathryn Ratcliffe set a [Guinness World Record](#) by eating 138 Smarties in three minutes using [chopsticks](#).

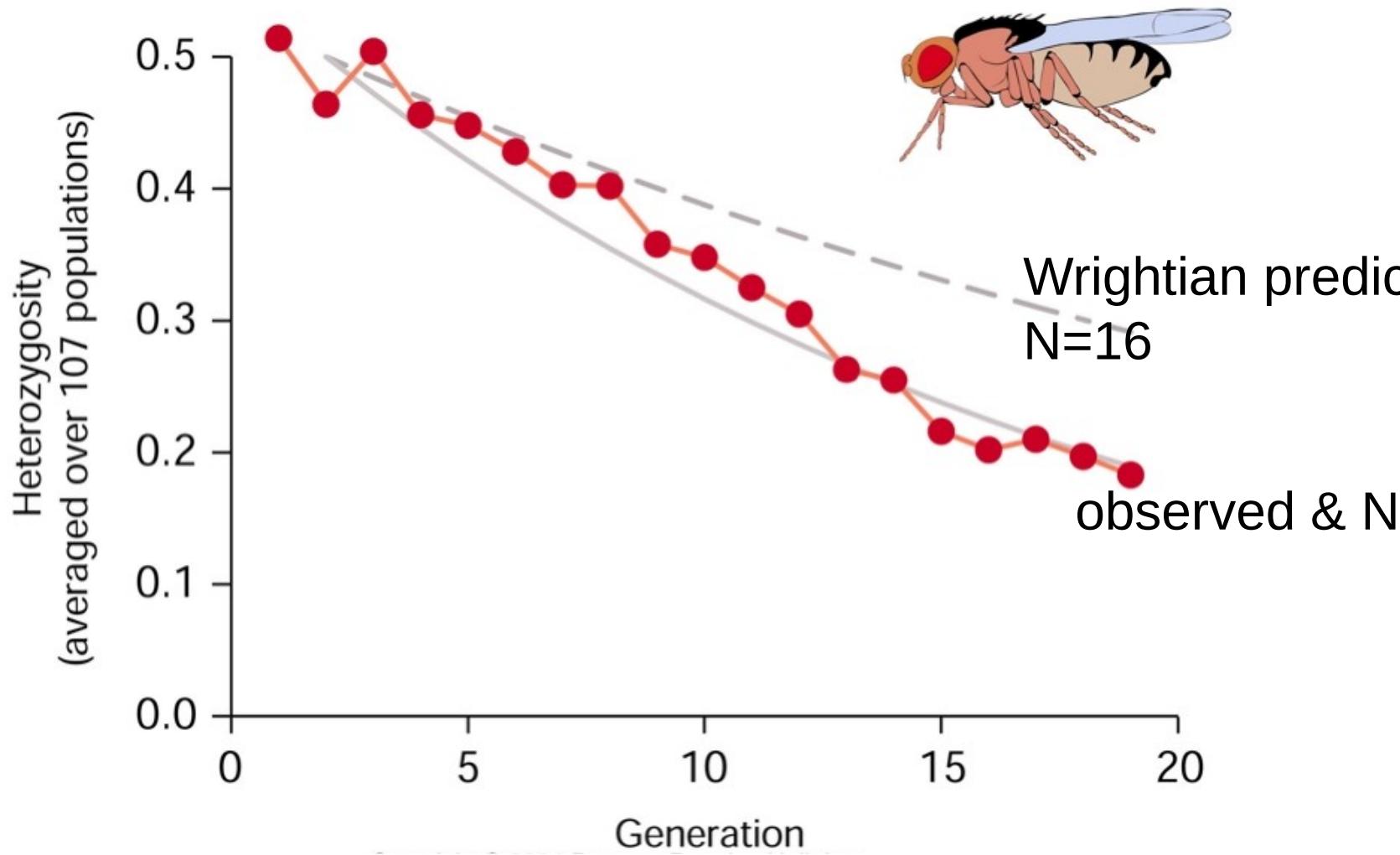
Assignment

- Collect genotype frequency data for a population of Smarties
 - count each of seven morphs (red, yellow, orange, brown, green, purple, blue)
 - count any mutant Smarties with unusual phenotypes and note with email
 - email genotype frequency to me for computation of estimated global genotype frequencies.
 - eat Smarties (red ones last)

Buri's Drift Exp't

- founded 107 populations of *D. melanogaster* with 8 pairs each.
- all founders heterozygous for bw^+ / bw^{75}
- no known (or measured) fitness effect of alleles.
- ran experiment for 19 generations.





A population's EFFECTIVE SIZE is never as large as its census size.

Effective Population Size

- The number of individuals in an idealised population that shows the same magnitude of drift as the real population.
- populations are genetically never as large as their census size.
- differences in survival and repro success lead to unequal contributions of gametes to next generation.
 - variance in male mating success may be particularly high.
- skewed sex ratios strongly affect N_e

Calculating N_e

a crude* approximation of effective size is given by:

$$N_e = (4N_m * N_f) / (N_m + N_f)$$

*this is approximate because relative fitness is not estimated.

Sexual Selection & N_e

- a population with 100 breeding males and 900 females has an effective size of 360.
- in lekking species or those with extreme dominance heirarchies, whole social groups of females may mate with the same male.

$$N_e = (4N_m * N_f) / (N_m + N_f) \approx 4$$

note: this calculation assumes only one round of breeding
and no migration between groups.

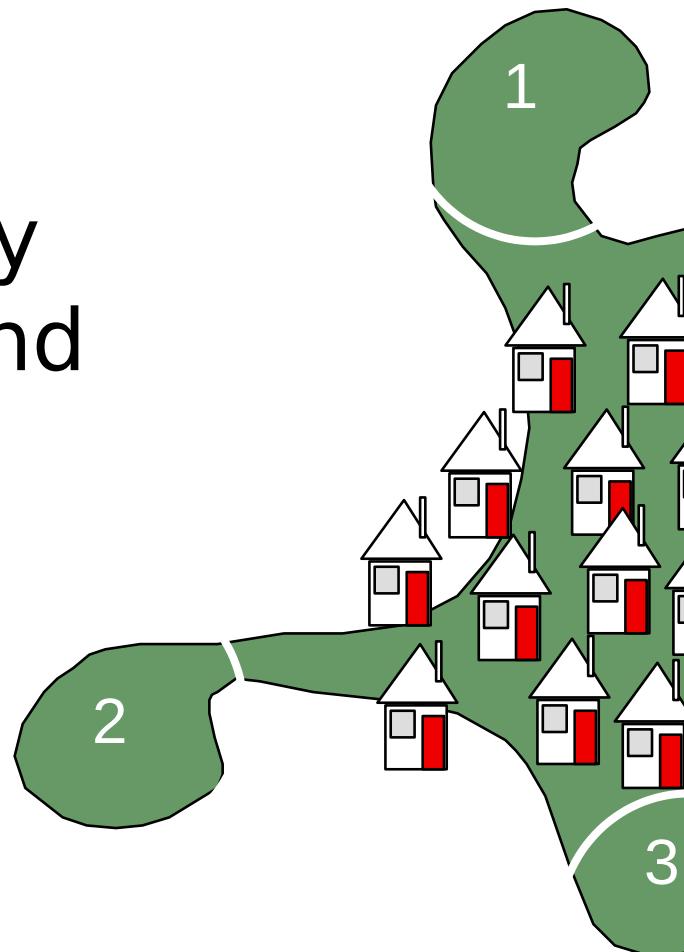
N_e and population fluctuation

- effective population size is extremely sensitive to population bottlenecks
- calculated based on *Harmonic mean*

(do not worry about how to calculate this!)

Drift is often the product of reduced gene-flow

- when populations become fragmented, interrupted gene flow will often lead to:
 - increased homozygosity via drift and the Wahlund Effect.
 - inbreeding depression
 - reduced adaptability
 - less variation to resist environmental challenges, disease, parasitism.



Measures of Diversity

Components of Heterozygosity

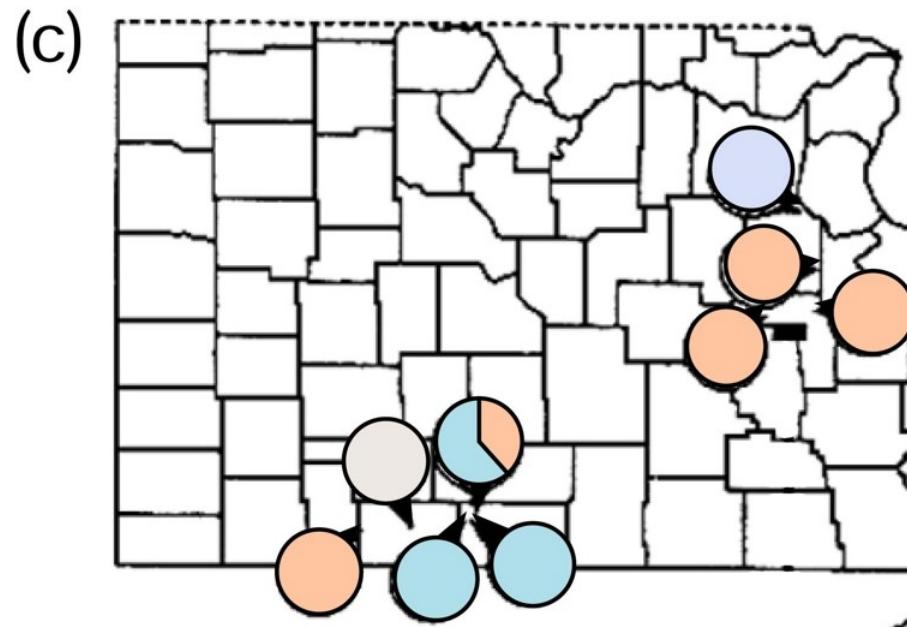
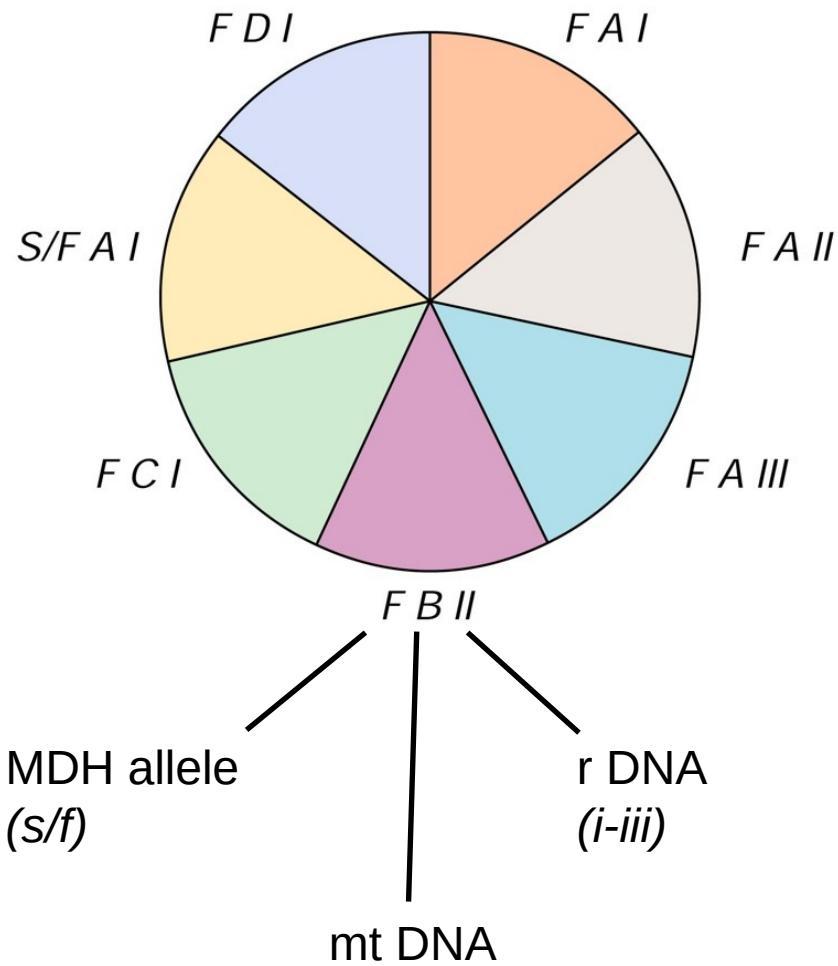
1. Allele Richness: the average number of alleles per locus in the genome.
2. Genetic Polymorphism: the fraction of loci in the genome with 2+ alleles at frequency of >0.01.

Collared Lizards

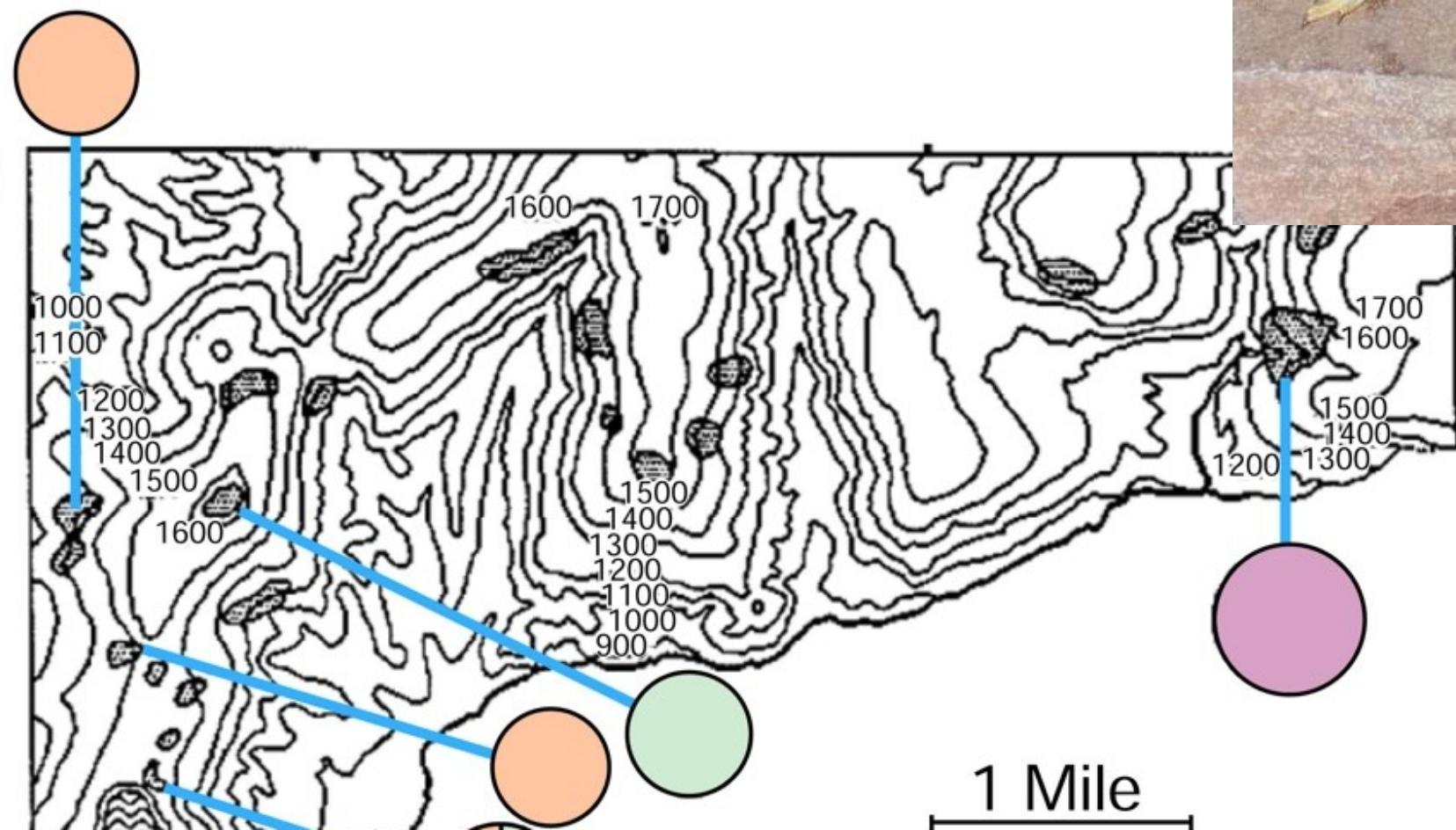


- Relict populations in the Ozark Mountains
 - occupy small glades (remnants of SW deserts) that were once isolated by savannah lands.
 - savannahs burned periodically.
- human intervention:
 - clear cutting & fire extinguishing
 - allowed oak-hickory forest to take over.
 - allowed red cedar to grow into glades.

Ozark Collared Lizards



Habitat Fragmentation in Glade Populations



Perils of Fragmentation

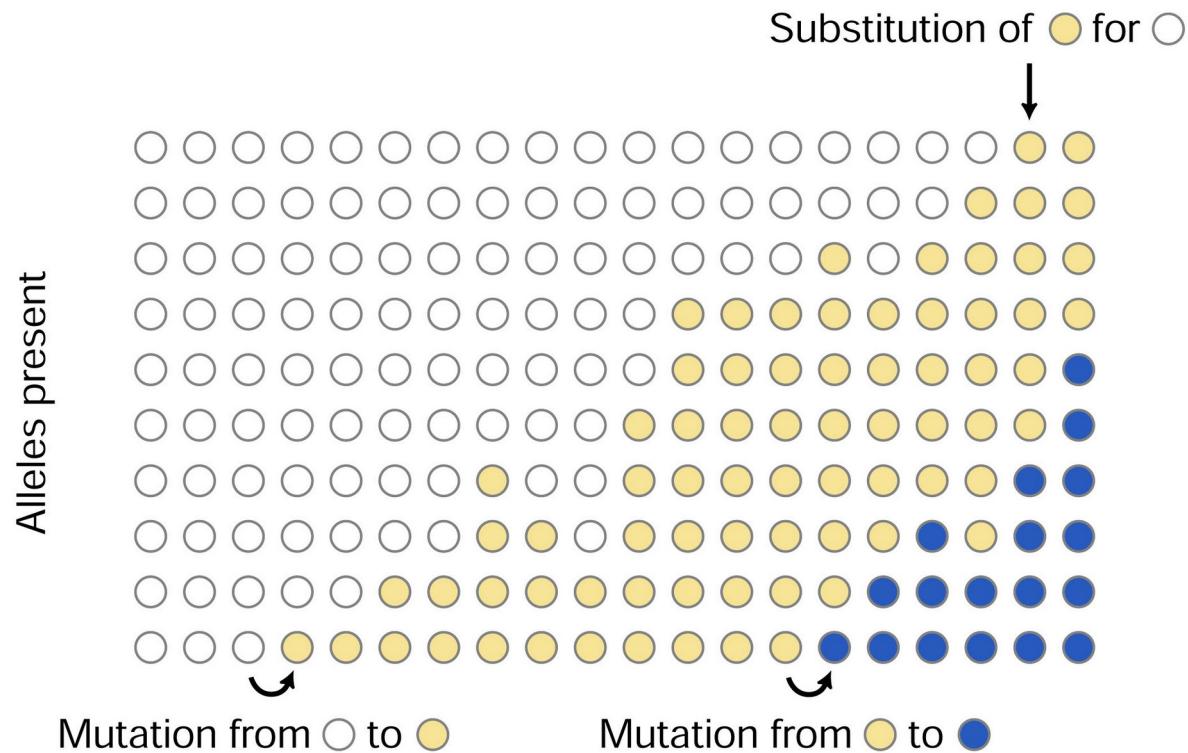
- Occupants of any given glade genetically homogeneous
 - unable to adapt to further changes in the environment
 - sitting ducks for diseases
 - ever more sickly due to inbreeding depression.
- Remediate via restoration of empty glade populations, creation of migration corridors with controlled burns.

Wright's F statistics

(Has nothing to do with the F statistics used in ANOVA)

F (Fixation index) = the reduction in heterozygosity expected with random mating at any one level of a population hierarchy, relative to another more inclusive level.

Neutral theory of evolution



Time (Generations) →

Reminder: substitution vs. polymorphism

What happen after a mutation changes a nucleotide in a locus

Polymorphism: mutant allele is one of several present in population

Substitution: the mutant allele fixes in the population. (New mutations at other nucleotides may occur later.)

Substitution schematic

Individual 1 2 3 4 5 6 7

Time 0: aaat aaat aaat aaat aaat aaat

Time 10: aaat aaat aaat aaat a**cat** aaat aaat

Time 20: aaat aaat a**cat** aaat a**cat** a**cat** a**cat**

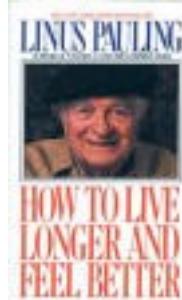
Time 30: a**cat** a**cat** a**cat** a**cat** a**cat** a**cat** a**cat**

Time 40: a**cat** a**cat** a**ctt** a**cat** a**cat** a**cat** a**cat**

Times 10-29: polymorphism

Time 30: mutation fixed -> substitution

Time 40: new mutation: polymorphism

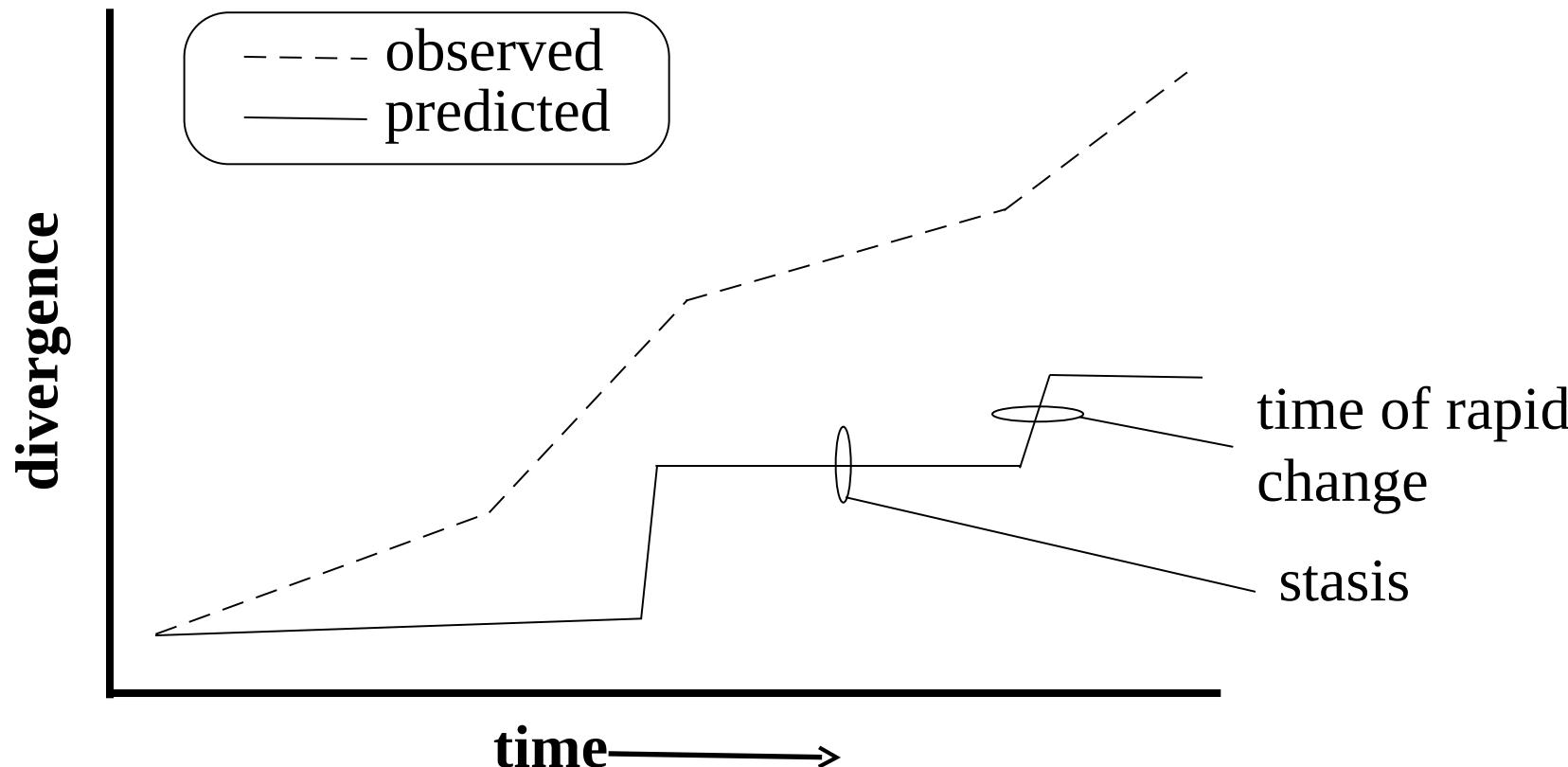


Molecular Evidence

- Early data on amino acid sequences in vertebrate proteins:
 - Kimura: mutations causing amino acid substitution appeared to be occurring at a surprising pace.
 - Zuckerkandl & Pauling: amino acid subs appear to occur at a steady pace
 - could be used as a 'molecular clock' to estimate divergence of species.



- Problem: steady, clock-like change did not fit predictions of evolutionary theory



Neutralist vs. selectionist view

Are most substitutions due to drift or natural selection?

“Neutralist” vs. “selectionist”

Agree that:

Most mutations are deleterious and are removed.

Some mutations are favourable and are fixed.

Dispute:

Are most replacement mutations that fix beneficial or neutral?

Is observed polymorphism due to selection or drift?

Reminder: substitution rates for neutral mutations

Most *neutral* mutations are lost

Only 1 out of $2N$ fix

Most that are lost go quickly (< 20 generations for population sizes from 100 - 2000)

Most *replacement* mutations are lost since deleterious: rate of loss is faster than neutral

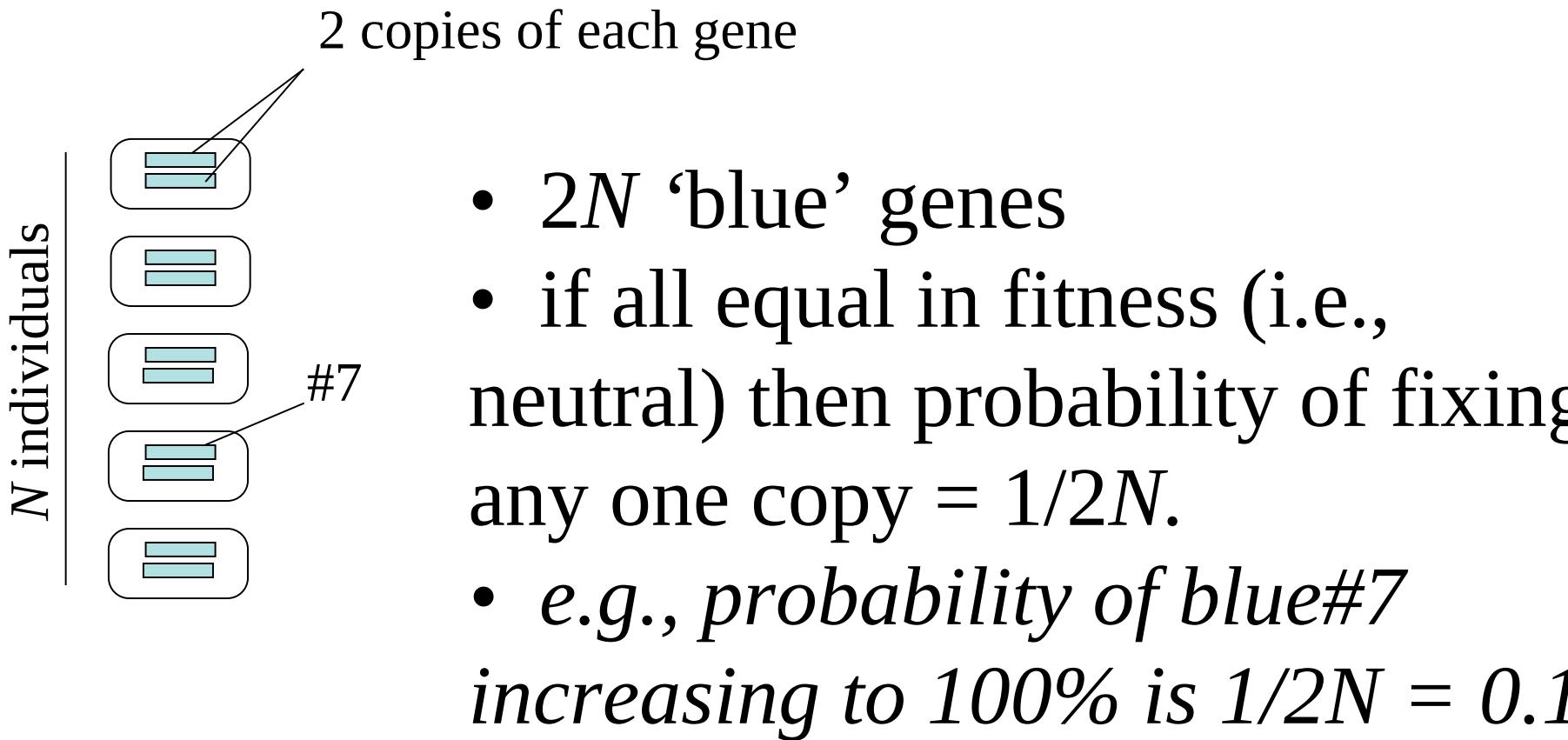


Kimura & Neutral Theory

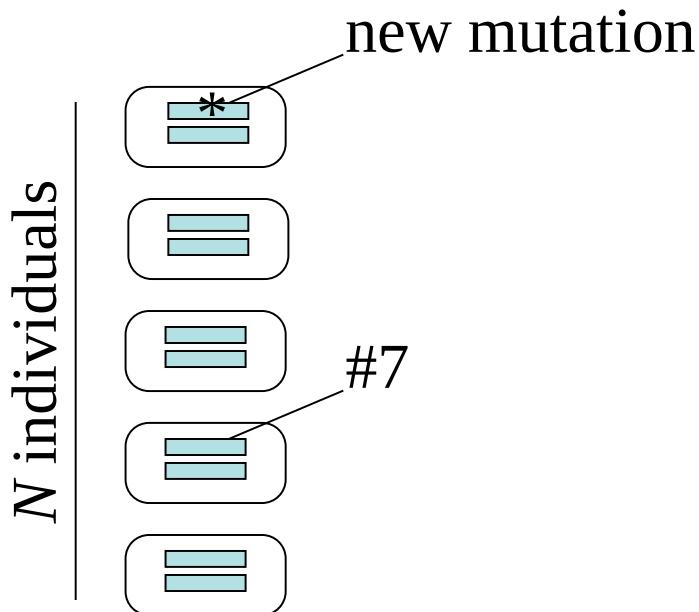


- the vast majority of mutations are selectively neutral.
- rate of evolutionary change = mutation rate for most genes in a population.
- due to drift alone, the chance of an individual copy of a gene fixing is $1/2N$
- if the rate of new mutation at a locus is v , then the number of new mutants in the population is $2Nv$ per generation.

Kimura Model



Importance of mutation



- v is mutation rate per gene generation.
- $2Nv$ mutations (new copies happen per generation).
- *if $v=0.1$, then in our population $2Nv = 0.02$.*
- rate of fixation =
- $(1/2N) \times (2Nv) = v$

probability rate of new
by drift mutations.

Neutral Theory & The Clock

- Episodes of selection should be followed by periods of less change, once a protein is optimized for a particular function.
- Kimura proposed that most molecular evolution is driven by selectively neutral variation:
 - alleles have no selection coefficient
 - genetic drift & mutation the dominant forces in evolution

Neutral Theory-O-Rama

- 1) Population size does not matter
 - drift stronger in small populations
 - *but* mutation equally less common!
- 2) Selection for new beneficial genes does not matter
 - most new mutations harmful, eliminated by natural selection.
 - v therefore approximates the maximum rate of evolutionary change.

Neutralists vs. Selectionists

- the neutral view for molecular evolution spread in the ‘80s with the finding of clock-like sequence divergence.
- view was rebutted by John Gillespie and others.
 - selectionists view the incidence of positive mutation as non-trivial, the general effect of mutation to be deleterious and both to be exposed to selection.
 - selection is progressive, even if ‘progress’ is blind

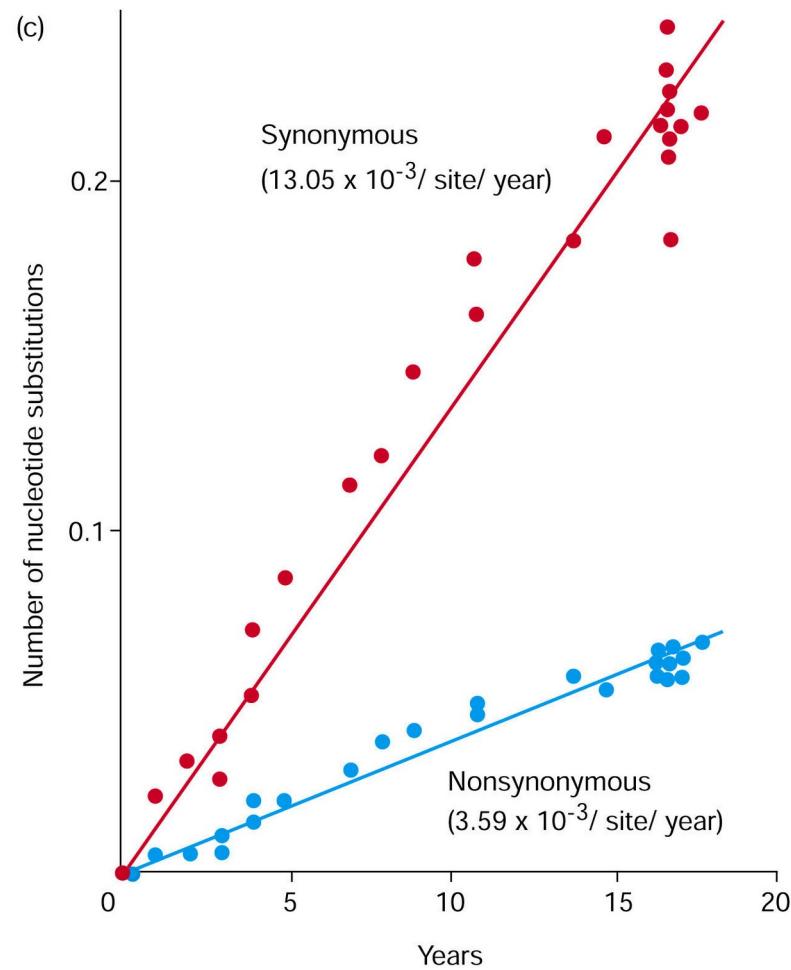
Testing the Neutral Theory

- pseudogenes should undergo neutral evolution:
 - non-coding sequences & junk DNA have the highest rates of sequence divergence.
- codons
 - silent-sites in codons undergo neutral evolution.
 - replacements at silent sites called SYNONYMOUS replacements because AA product unchanged.

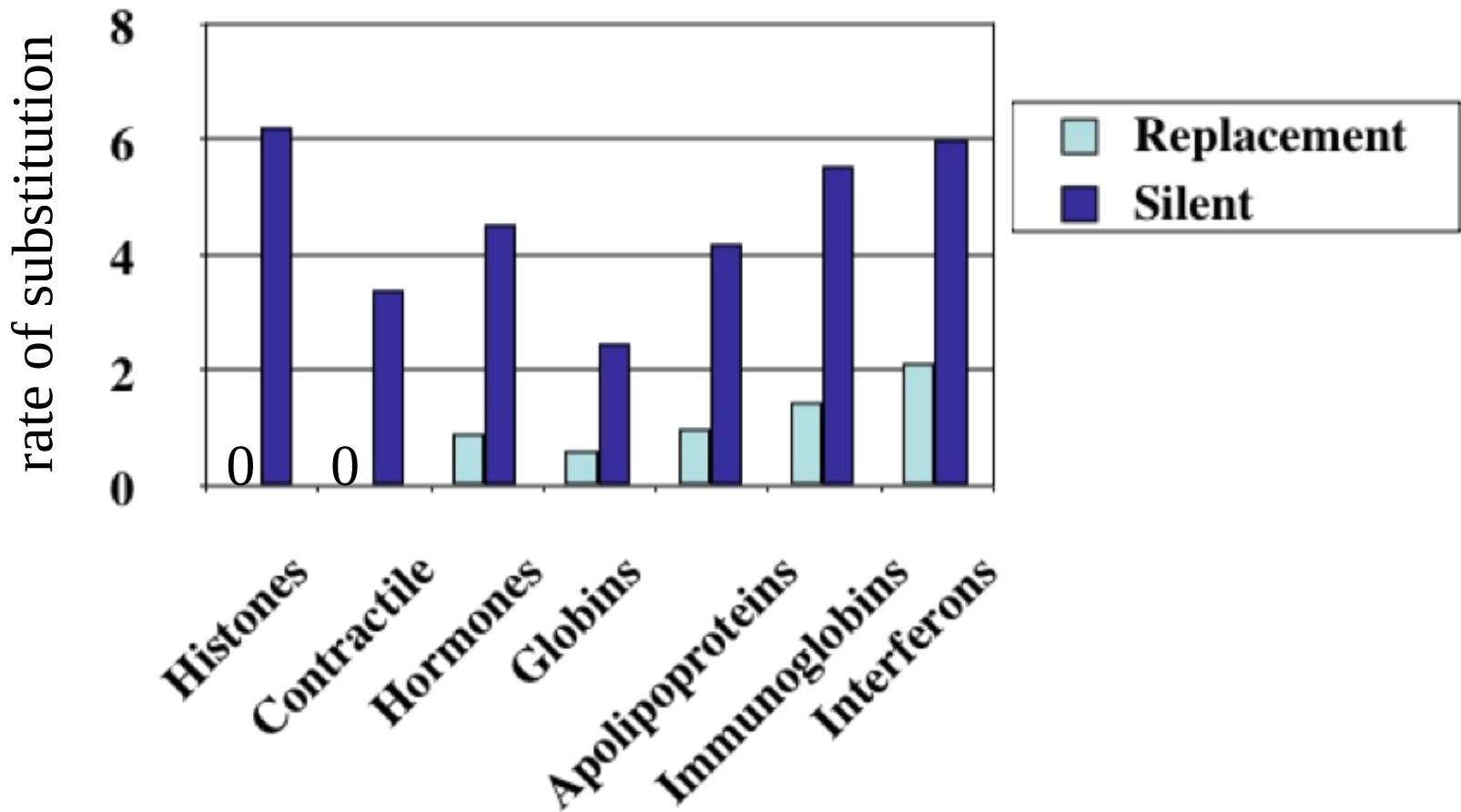
Replacement or Non synonymous Silent or synonymous

GLYCINE	GGG	G
	GGU	U
	GGA	A
	GGC	C

Molecular evolution in influenza virus



However,
Codon position: Human vs. Mouse



from Li & Graur 1991

Rates vary by locus

- Frequency of nonsynonymous (replacement) substitutions related to constraints on gene function.
 - EXAMPLE: Histones are vital in cell functioning and show very little replacement substitution; immunoglobins show higher rate of replacement.
- Predictions exactly the same as for coding vs. non-coding sequences, but graduated.

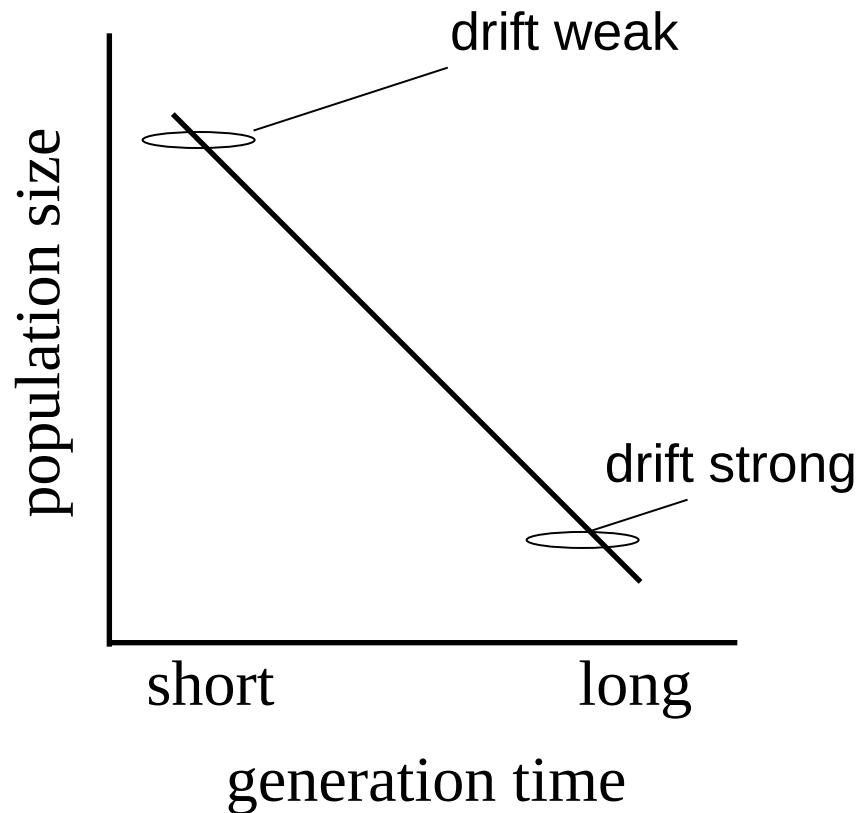
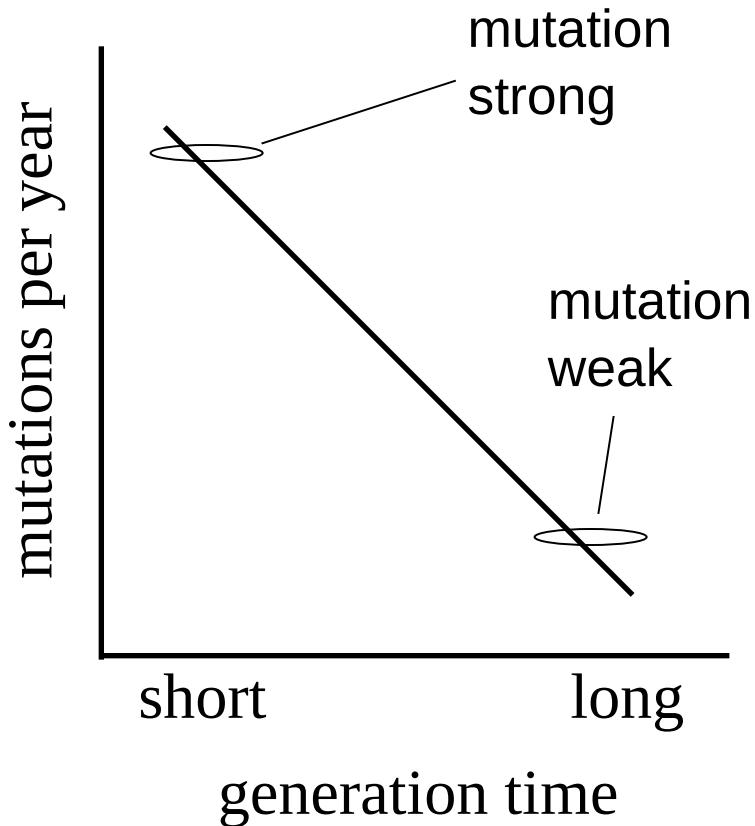


More Trouble for Neutrality

- why do protein sequences change in a steady clocklike fashion among species with very different generation times?
 - i.e., why change with absolute time when organisms have different reproductive rates?
 - fly (11d) vs. albatross (9y): fly should have many more mutations per year than the albatross!
- Ohta & Kimura modified theory to include mutations with slightly deleterious effects
 - makes drift more important in small populations
 - Mutations effectively neutral when $s \leq 1/2Ne$



Reconciling the time paradox



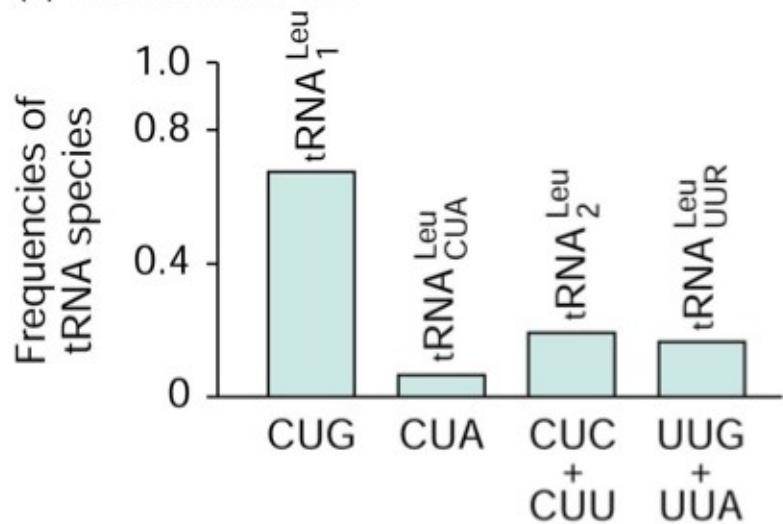
FLY: strong v x weak drift
ALB: weak v x strong drift

= rate of change

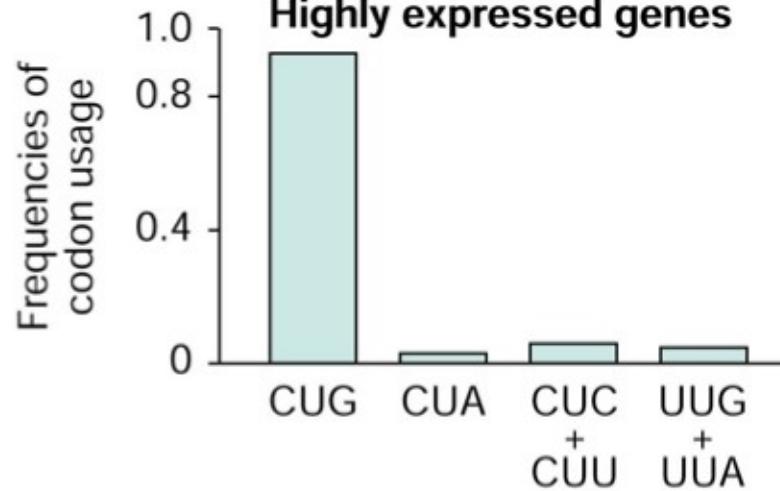
Codon Bias

- most amino acids have multiple codons (20 AAs, 64 combos of bases).
- example: leucine has 6 codons:
- UUA, **UUG**, CUU, CUC, CUA, **CUG**
 - favoured in *S. cerevisiae*
 - favoured in *E. coli*,
D. melanogaster.

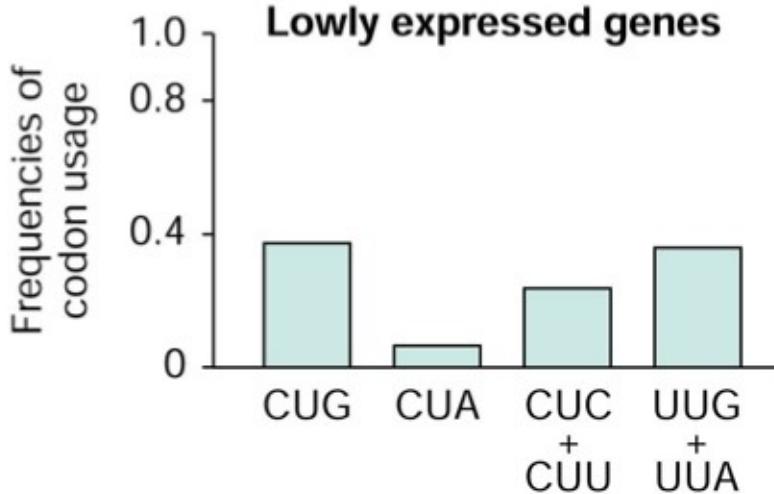
(a) *Escherichia coli*



Highly expressed genes



Lowly expressed genes



Implications of codon bias

- 1) Strongest in most highly expressed genes.
- 2) Selection is *not* silent at the third position after all.
- 3) Transcriptional efficiency is probably the driving factor in codon bias:
 - tRNAs match most common mRNAs produced in cells.
 - mismatch (i.e., mutation to a new codon) reduces translational efficiency of transcript -> selected against.

Neutrality in the Docket

- silent site changes and pseudogenes support a neutral model.
- The variation in functional positions and coding loci do not support a strictly neutral view.
- Conclusion: parts of the genome will evolve neutrally.
 - these portions will prove very useful in some population genetic analyses.
 - estimates suggest 95% of human genome is non-coding (junk) DNA.

More Uses of Neutral Theory

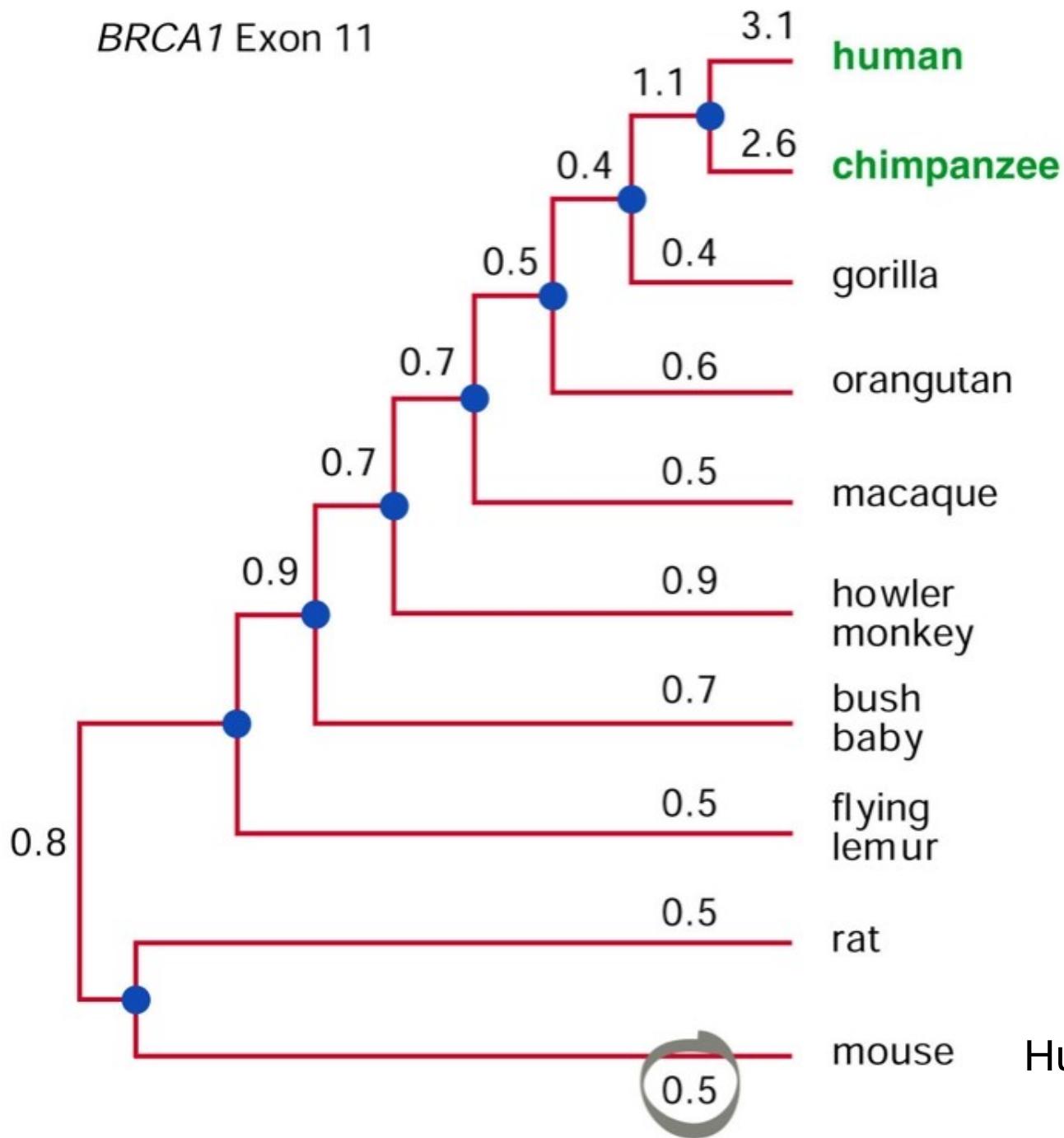
- Explains several important patterns in molecular evolution.
- Like Hardy-Weinberg Equilibrium, provides a null hypothesis for testing adaptive evolution:
 - *compare change to neutral expectation to determine if natural selection has shaped the sequence.*

Looking for Positive Evolution

- Hughes & Nei looked at major histocompatibility complex (MHC) gene sequences.
 - membrane proteins important in immune system recognition of infected cells.
- synonymous substitution rate used to estimate the mutation rate, ν .
 - provides a benchmark for gaugeing non-synonymous (replacement) rate of change.
- H & N found *higher* Non synonymous rates in the antigen recognition system than predicted by the Synonymous site benchmark.

- Higher replacement rate can only be explained by positive selection.
- i.e.,
- $d_{NON} / d_{SYN} > 1$ only when replacements advantageous.
- $d_{NON} / d_{SYN} < 1$ when replacements disadvantageous

BRCA1 Exon 11



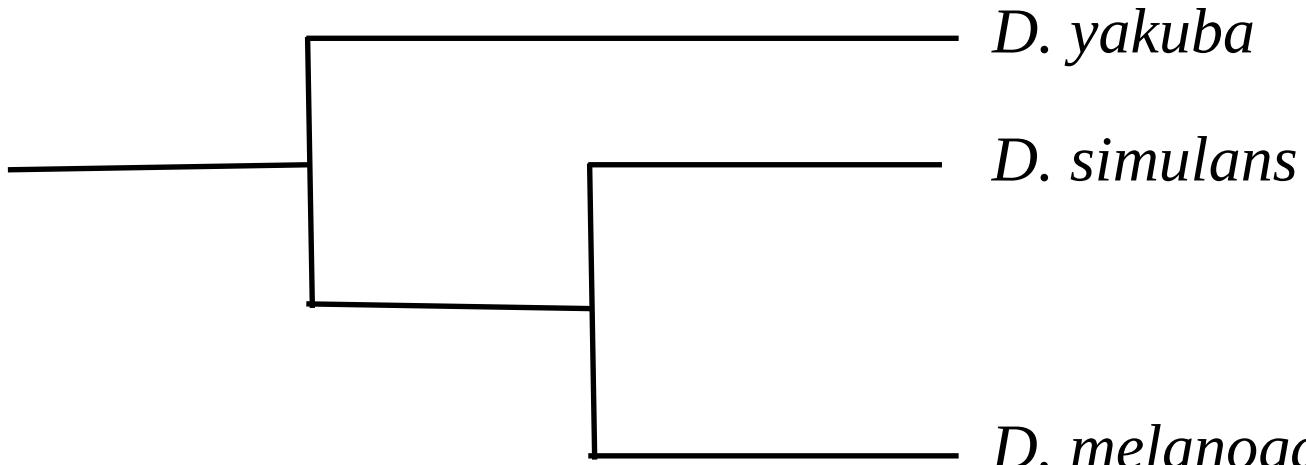
Hutley G et al. 2

$$d_{NON} / d_{SYN} > 1$$

- Unduly conservative index of positive selection (Sharp 1997).
- McDonald & Kreitman refined the silent vs. replacement hypothesis:
 - ratio silent:substitution is constant through time according to neutral theory.
 - use ratio to compare within-species change to between-species change.



MK test: *Adh* in *Drosophila*



- *Adh* (Alcohol dehydrogenase) important for fruit flies because they live on rotting fruit.
- M&K scored number of fixed vs. polymorphic sites based on sequence data.
- Null (neutral prediction): within species ratio of silent:replacement substitutions $\approx 20:1$

The Genes to Watch: Evolving at Warp Speed

- genes recruited to new functions (e.g., duplications)
- sex-determination genes
- fertilization interactions (sperm-egg; pollen-stamen)
- some enzymes, regulatory proteins
- immune-system genes
- sexual conflict genes?

Inbreeding

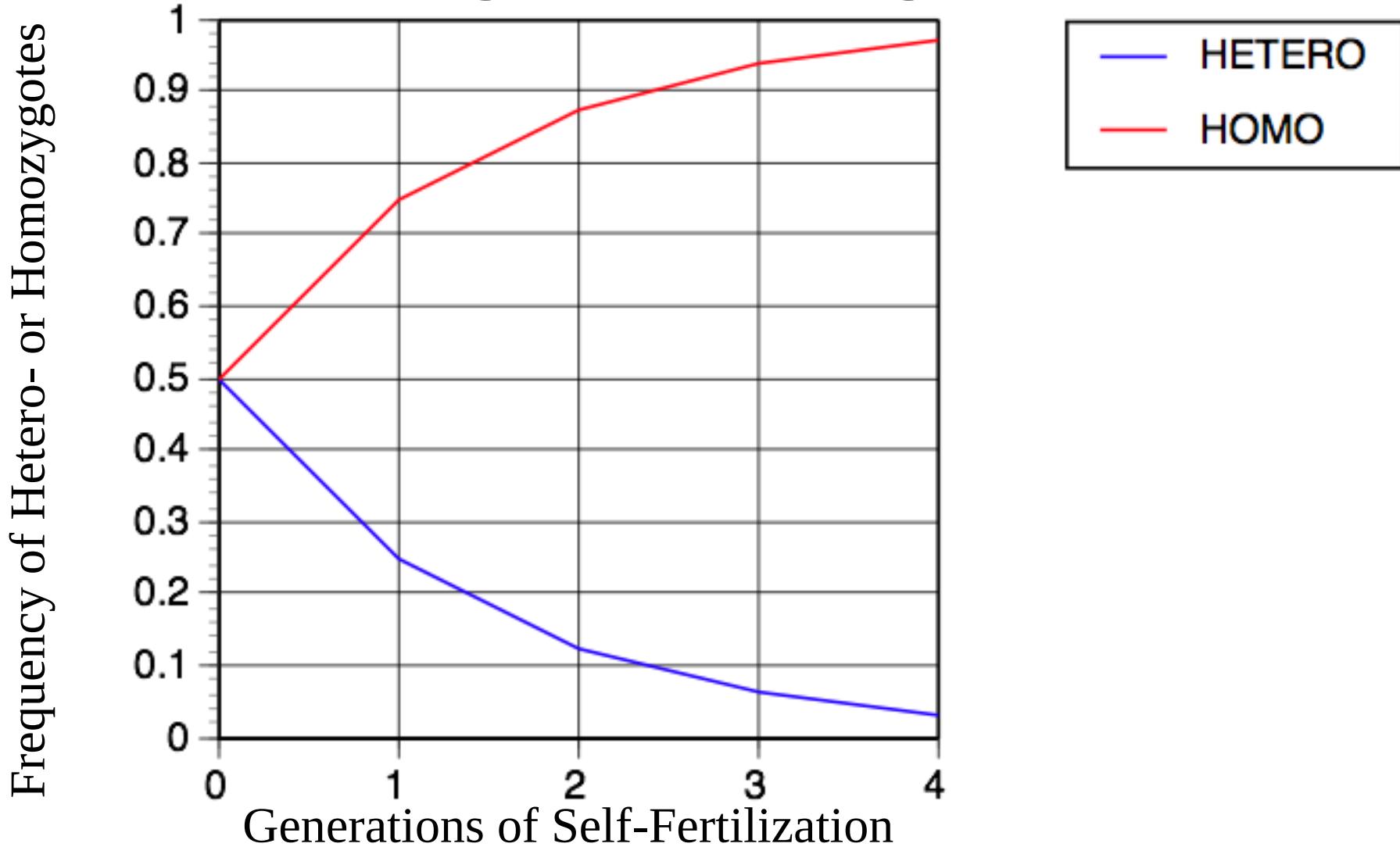
- Mating between related individuals
- F = inbreeding coefficient (probability that two homologous alleles in an individual are identical by descent)

- Identity by Descent: Two alleles created from the same ancestral copy of the gene
- Identity by Type: Two alleles sharing the same nucleotide sequence but not a common ancestor.
 - Some more terms-
 - Autozygous (Always homozygous)
 - Allozygous (Homozygous or heterozygous).

Genotype frequencies with Inbreeding

- $AA = p^2(1-F) + pF$
- $aa = q^2(1-F) + qF$
- $Aa = 2pq(1-F)$
- $H = H_0(1-F)$
- Pedigrees are used to compute F
- Inbreeding INCREASES genotypic variance
- Inbreeding does not change allelic frequencies

Inbreeding via Selfing

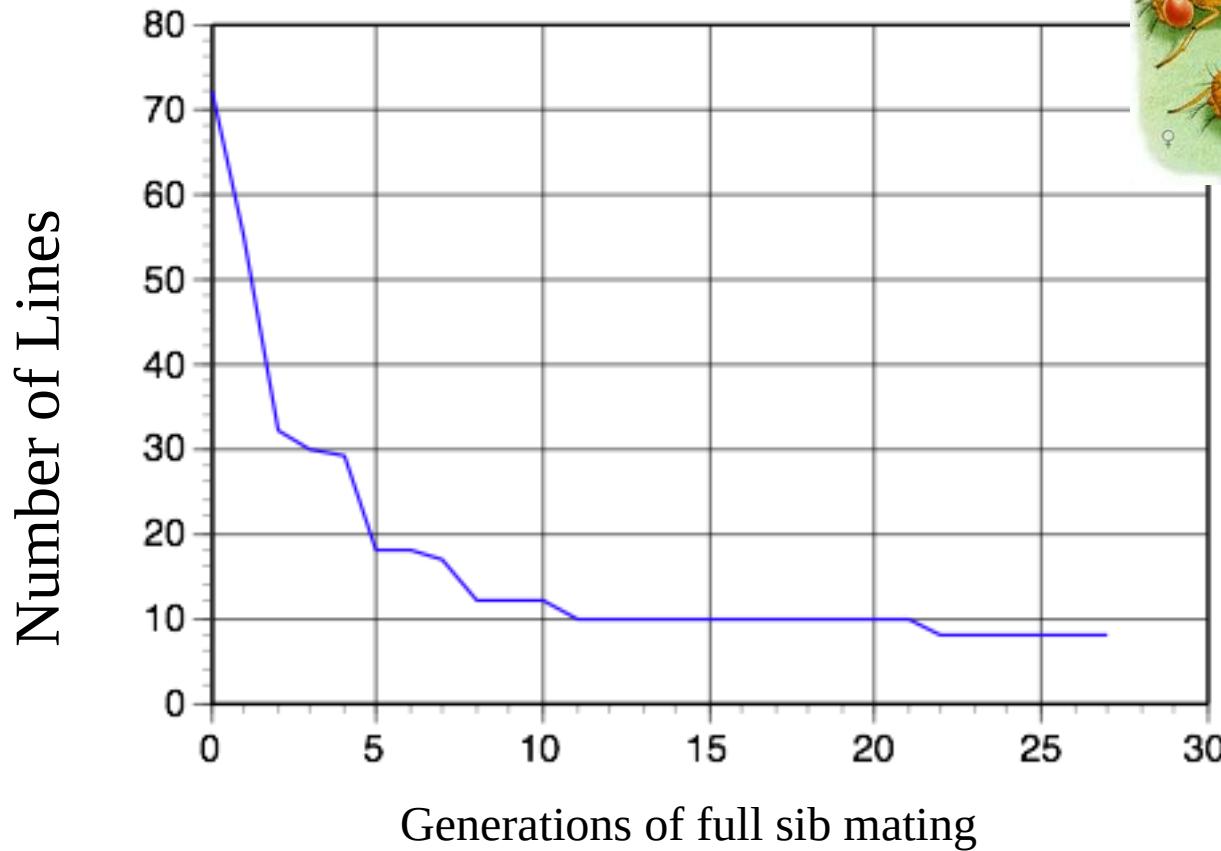


Despite changes in GENOTYPE FREQUENCIES,
inbreeding has no effect on ALLELE FREQUENCIES.

Melting down the Vortex

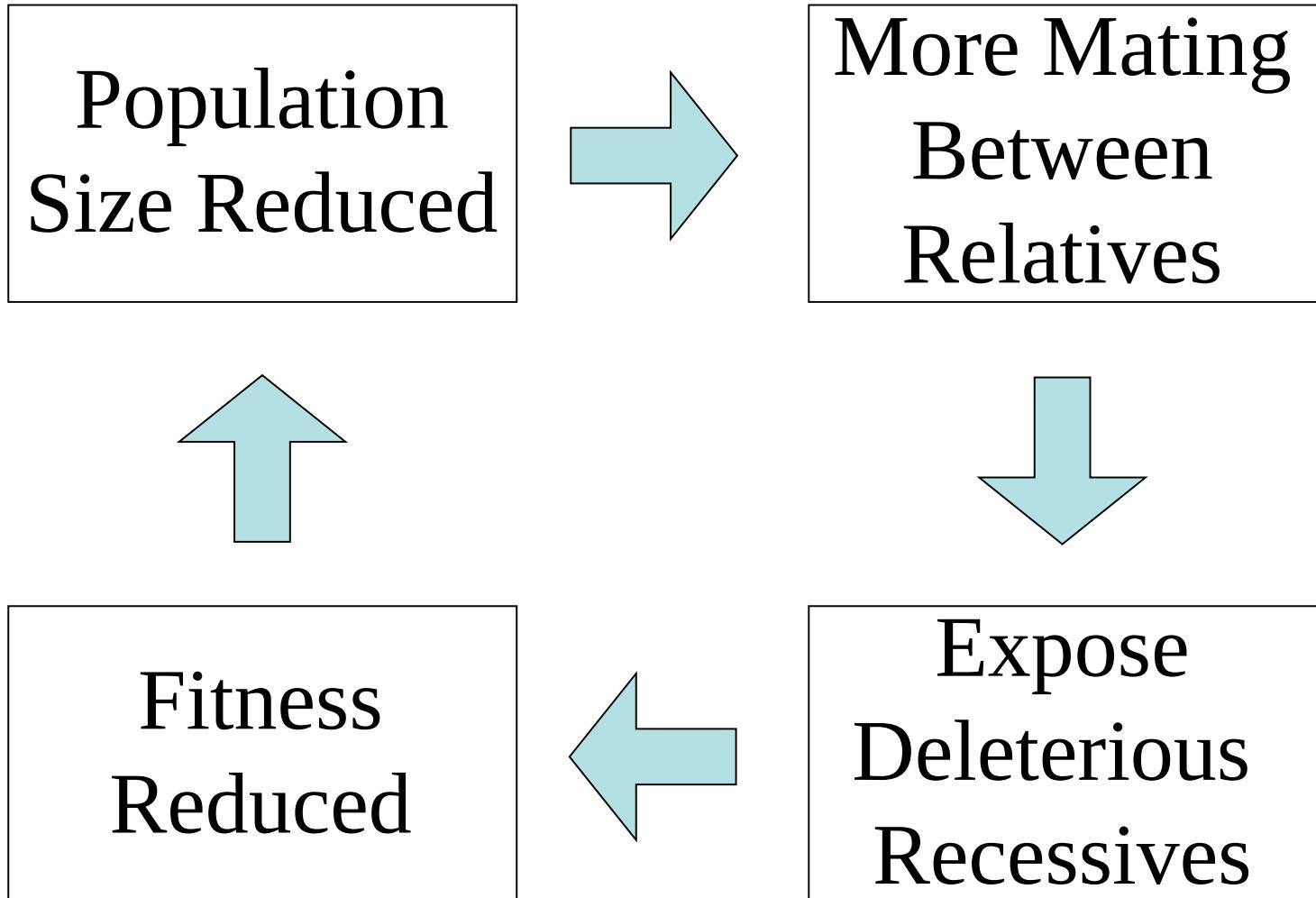
- Inbreeding and drift have similar effects – increased homozygosity.
- Genetic Drift & Mutation have profound implications for the conservation of populations.
- recall the Wahlund Effect:
 - as continuous populations become fragmented, the frequency of homozygotes increases due to drift.
 - homozygosity increases; relatives mate with relatives to an ever greater degree.

Inbreeding depression



- exposure of recessive deleterious variation leads to a syndrome of low performance:
 - = low survival & low fertility

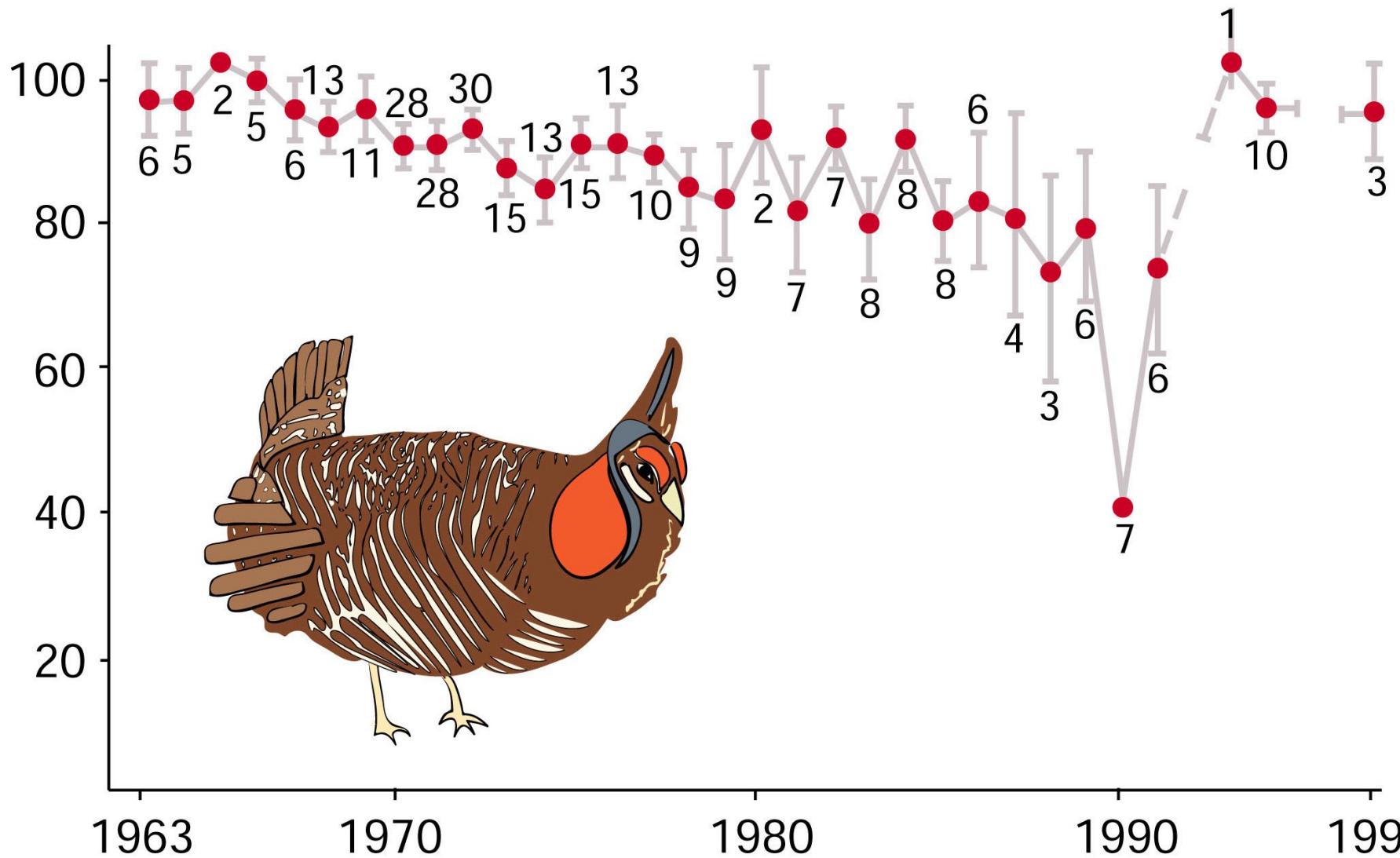
Mutational Meltdown



genetic load increases

Meltdown or Vortex?

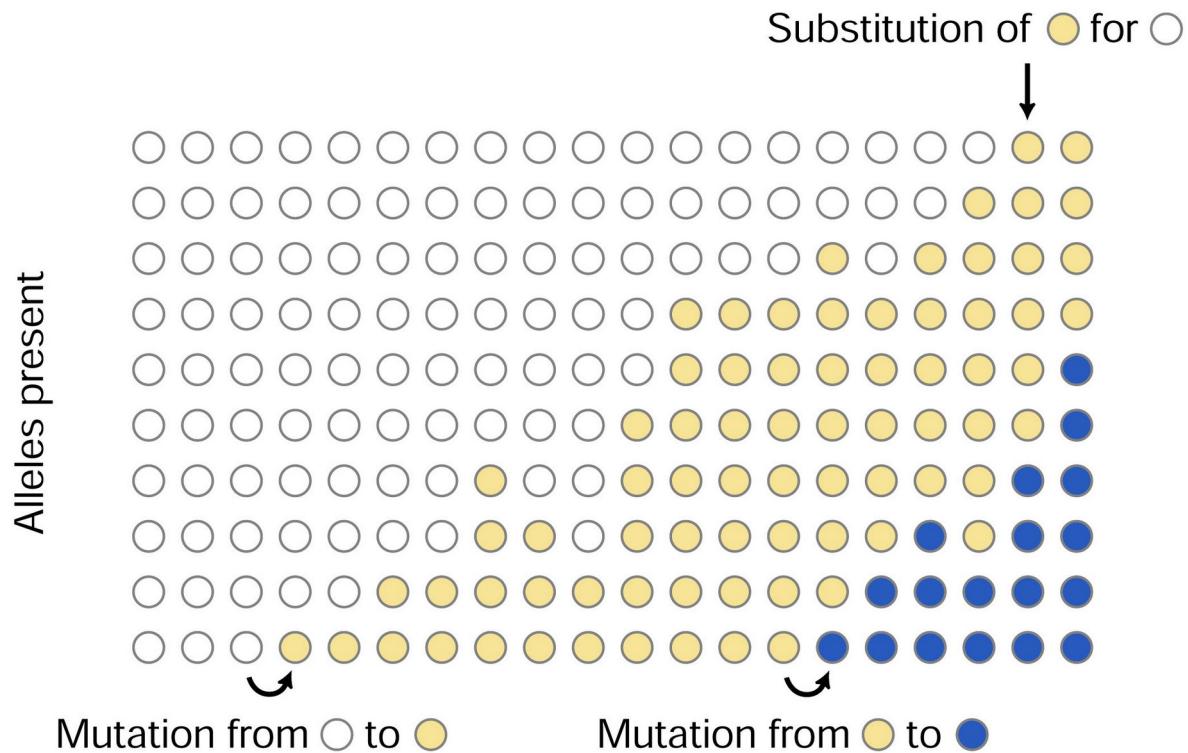
- Inbreeding is a strong form of genetic drift
- Causes exposure of deleterious recessive variation hidden in large populations.
- As fitness declines because of increased Genetic Load, population size shrinks further.
- This intensifies drift (i.e., further inbreeding)
- Synergism of effects leads to extinction.



Conservation of the Greater Prairie Chicken

- allelic richness had declined to about 60% of neighbouring populations (or since the 1930s).
- strategies based on habitat recovery were largely unsuccessful.
- importation of neighbouring stock from other states lead to dramatic reversal
 - a little bit of gene flow goes a long way
 - conservation also needs to be cognizant of the breeding system and effective population size

Neutral theory of evolution



Time (Generations) →

Copyright © 2004 Pearson Prentice Hall, Inc.

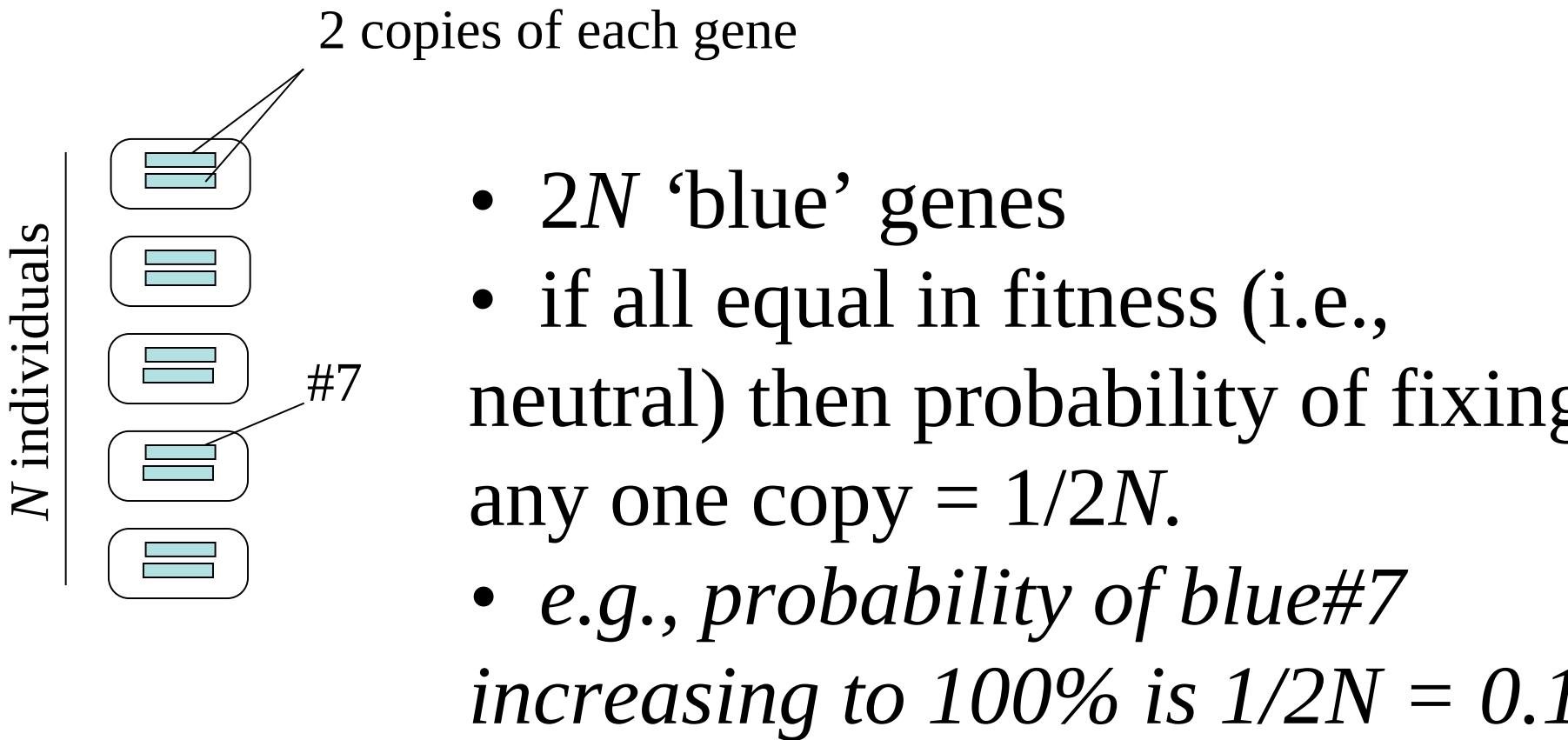


Kimura & Neutral Theory

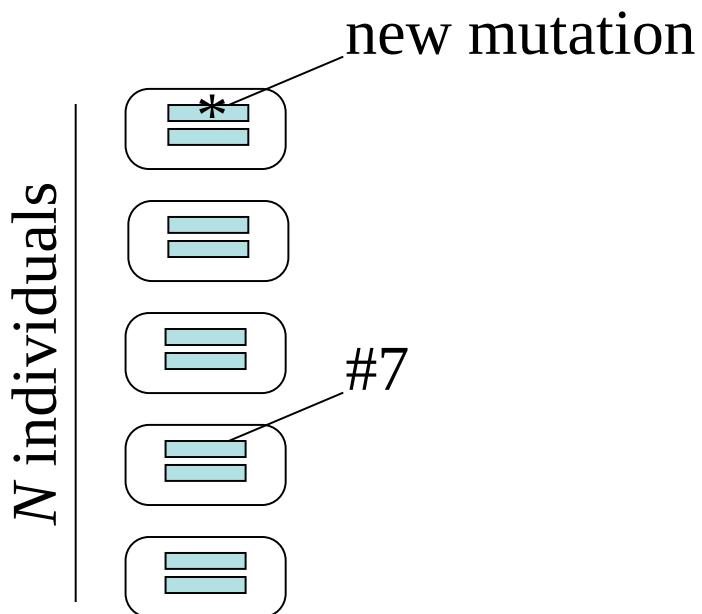


- the vast majority of mutations are selectively neutral.
- rate of evolutionary change = mutation rate for most genes in a population.
- due to drift alone, the chance of an individual copy of a gene fixing is $1/2N$
- if the rate of new mutation at a locus is v , then the number of new mutants in the population is $2Nv$ per generation.

Kimura Model

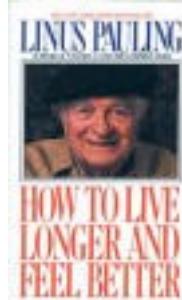


Importance of mutation



- v is mutation rate per gene generation.
- $2Nv$ mutations (new copies happen per generation).
- *if $v=0.1$, then in our population $2Nv = 0.02$.*
- rate of fixation =
- $(1/2N) \times (2Nv) = v$

probability rate of new
by drift mutations.

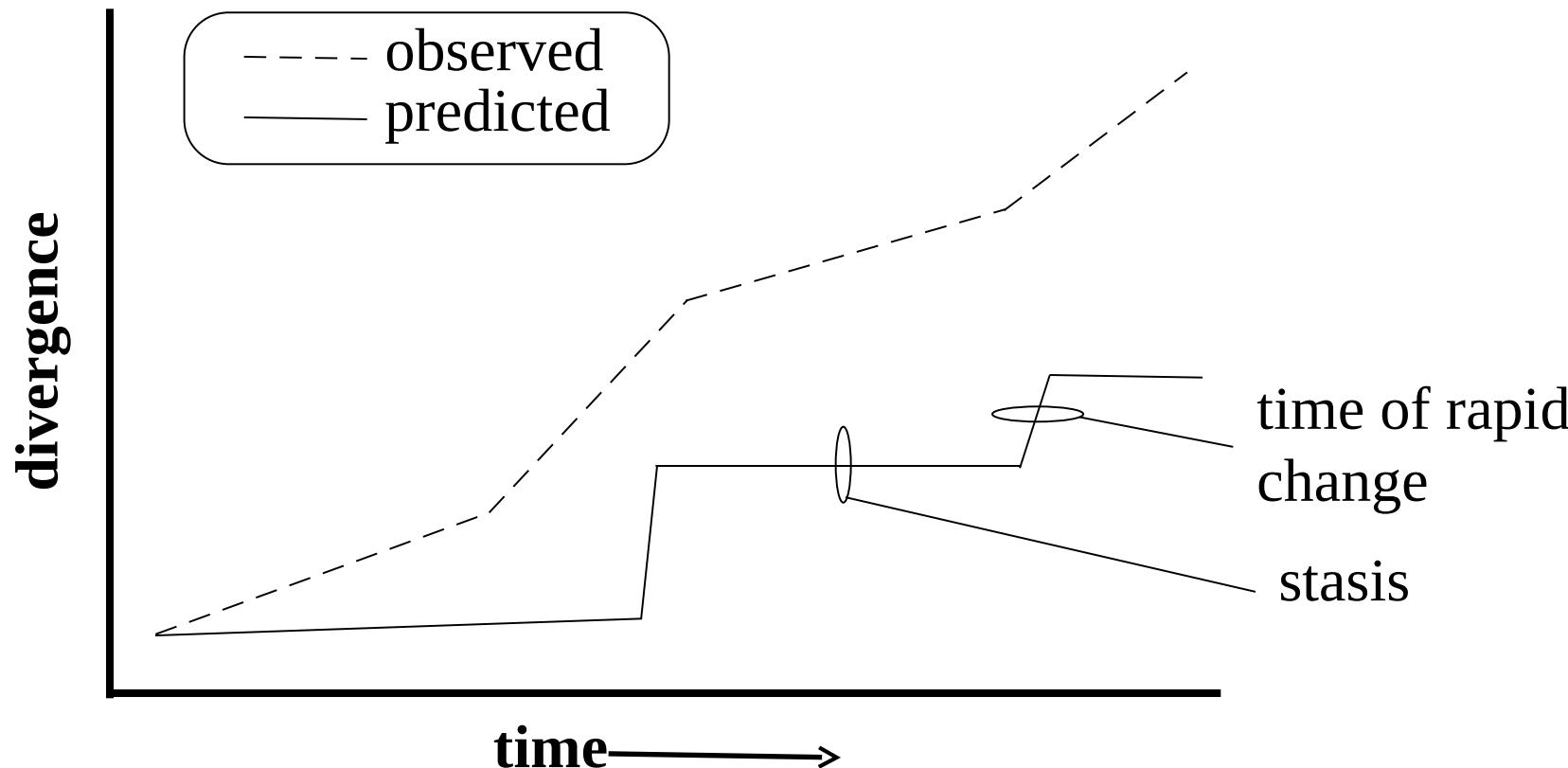


Molecular Evidence

- Early data on amino acid sequences in vertebrate proteins:
 - Kimura: mutations causing amino acid substitution appeared to be occurring at a surprising pace.
 - Zuckerkandl & Pauling: amino acid subs appear to occur at a steady pace
 - could be used as a 'molecular clock' to estimate divergence of species.



- Problem: steady, clock-like change did not fit predictions of evolutionary theory



Neutral Theory & The Clock

- Episodes of selection should be followed by periods of less change, once a protein is optimized for a particular function.
- Kimura proposed that most molecular evolution is driven by selectively neutral variation:
 - alleles have no selection coefficient
 - genetic drift & mutation the dominant forces in evolution

Neutral Theory-O-Rama

- 1) Population size does not matter
 - drift stronger in small populations
 - *but* mutation equally less common!
- 2) Selection for new beneficial genes does not matter
 - most new mutations harmful, eliminated by natural selection.
 - v therefore approximates the maximum rate of evolutionary change.

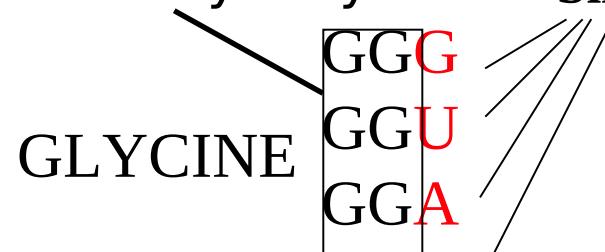
Neutralists vs. Selectionists

- the neutral view for molecular evolution spread in the ‘80s with the finding of clock-like sequence divergence.
- view was rebutted by John Gillespie and others.
 - selectionists view the incidence of positive mutation as non-trivial, the general effect of mutation to be deleterious and both to be exposed to selection.
 - selection is progressive, even if ‘progress’ is blind

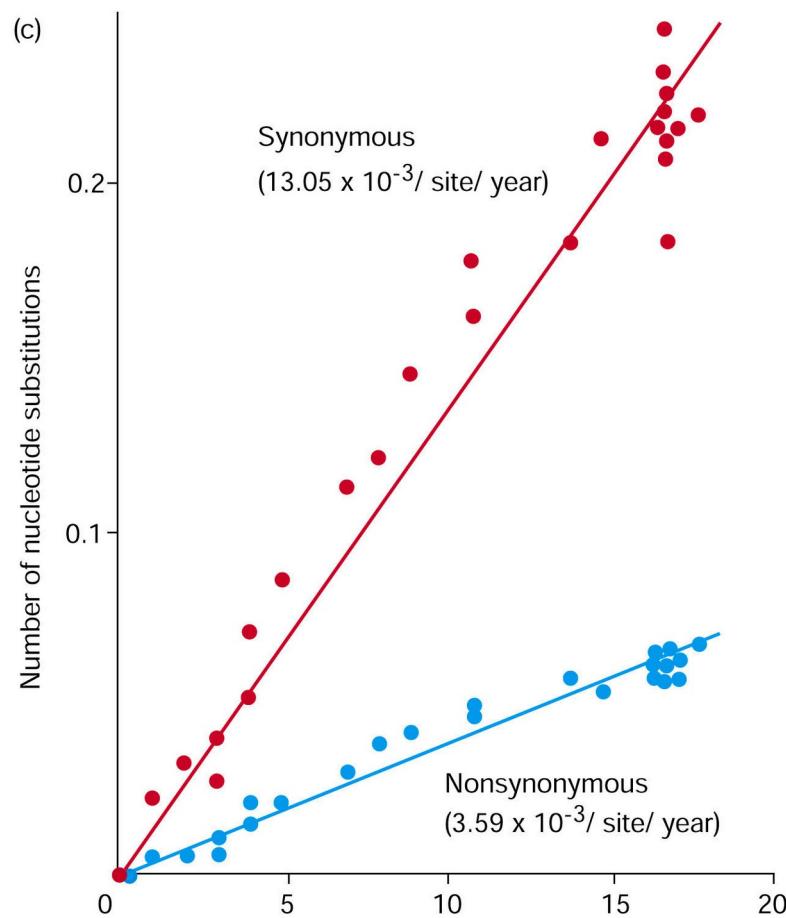
Testing the Neutral Theory

- pseudogenes should undergo neutral evolution:
 - non-coding sequences & junk DNA have the highest rates of sequence divergence.
- codons
 - silent-sites in codons undergo neutral evolution.
 - replacements at silent sites called SYNONYMOUS replacements because AA product unchanged.

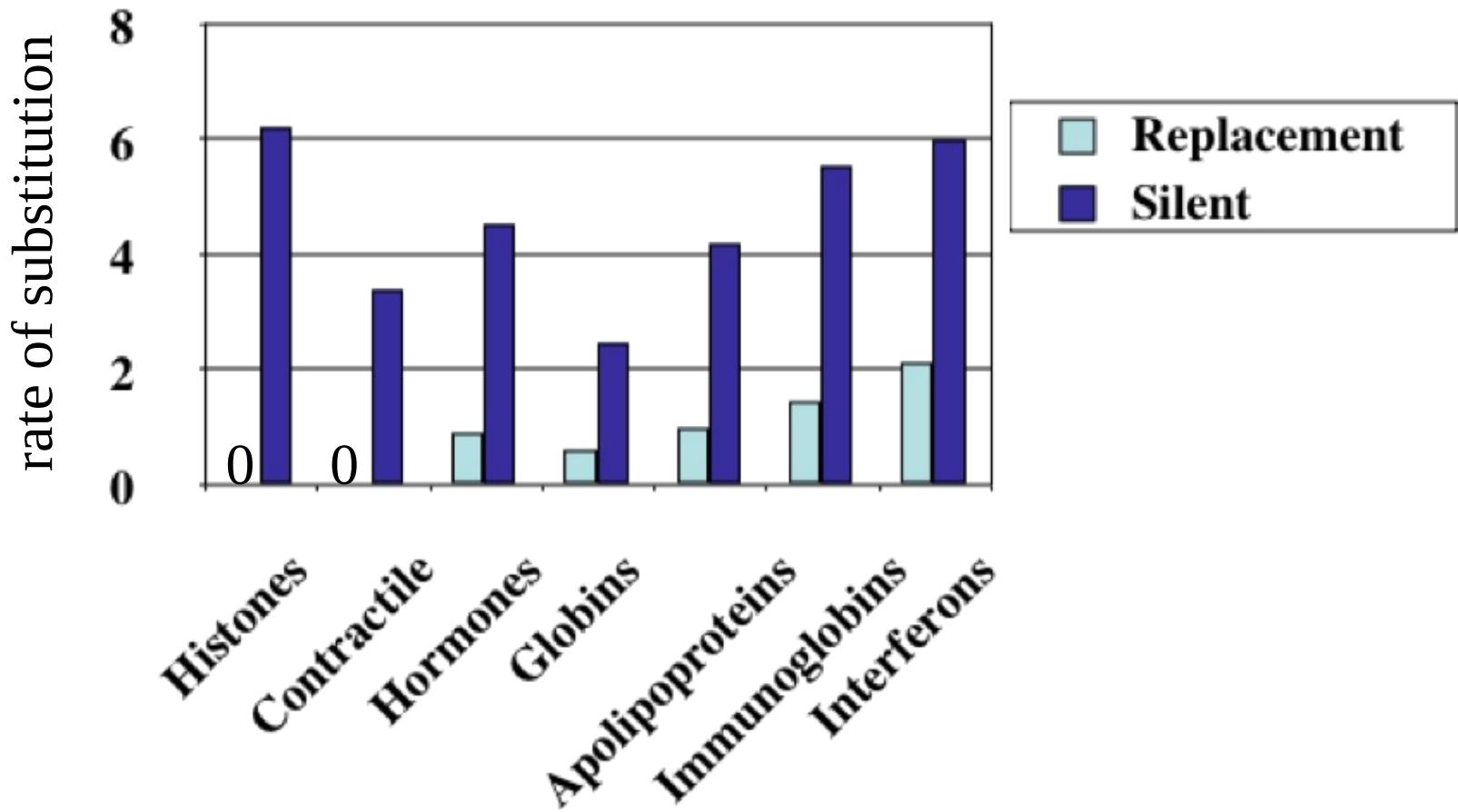
Replacement or Non synonymous Silent or synonymous



Molecular evolution in influenza virus



However,
Codon position: Human vs. Mouse



from Li & Graur 1991

Rates vary by locus

- Frequency of nonsynonymous (replacement) substitutions related to constraints on gene function.
 - EXAMPLE: Histones are vital in cell functioning and show very little replacement substitution; immunoglobins show higher rate of replacement.
- Predictions exactly the same as for coding vs. non-coding sequences, but graduated.

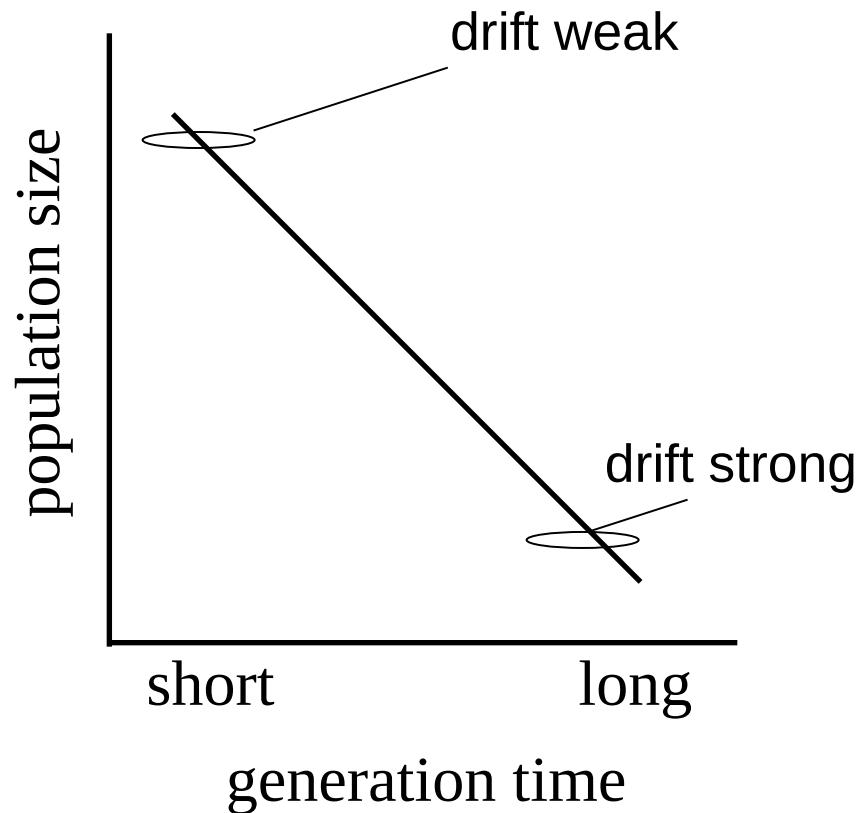
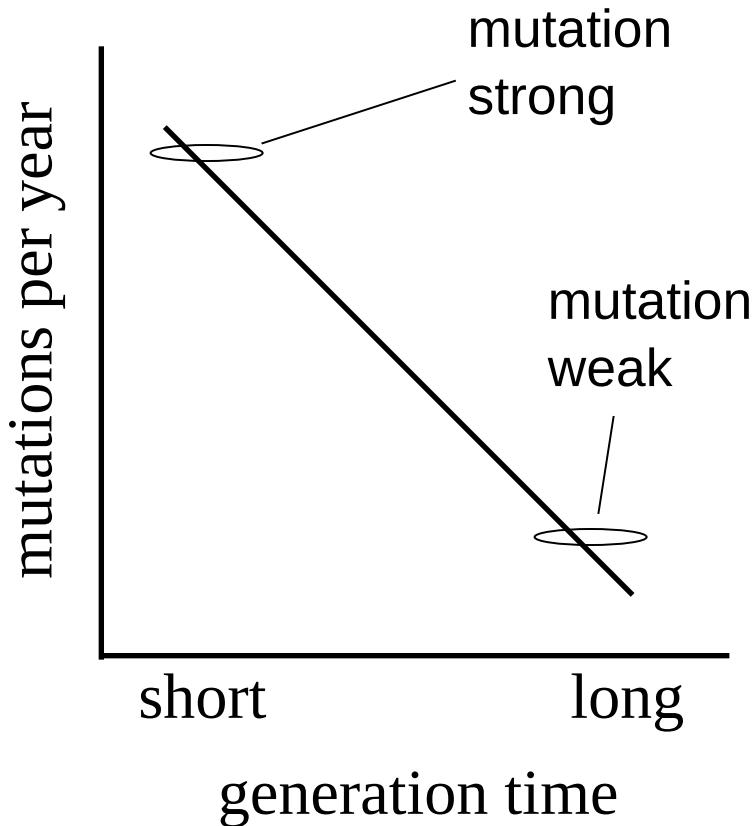


More Trouble for Neutrality

- why do protein sequences change in a steady clocklike fashion among species with very different generation times?
 - i.e., why change with absolute time when organisms have different reproductive rates?
 - fly (11d) vs. albatross (9y): fly should have many more mutations per year than the albatross!
- Ohta & Kimura modified theory to include mutations with slightly deleterious effects
 - makes drift more important in small populations
 - Mutations effectively neutral when $s \leq 1/2Ne$



Reconciling the time paradox



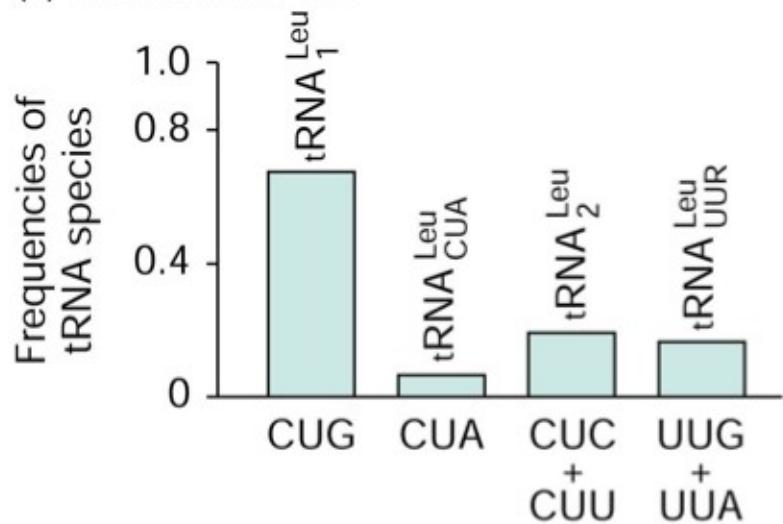
FLY: strong v x weak drift
ALB: weak v x strong drift

= rate of change

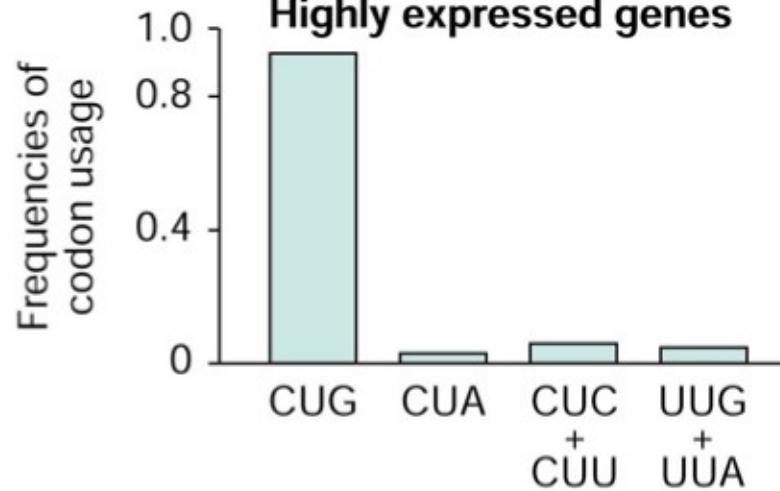
Codon Bias

- most amino acids have multiple codons (20 AAs, 64 combos of bases).
- example: leucine has 6 codons:
- UUA, **UUG**, CUU, CUC, CUA, **CUG**
 - favoured in *S. cerevisiae*
 - favoured in *E. coli*,
D. melanogaster.

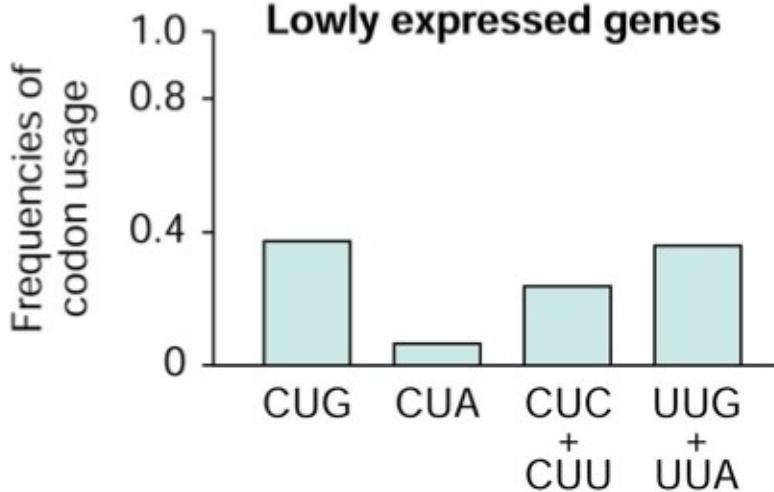
(a) *Escherichia coli*



Highly expressed genes



Lowly expressed genes



Implications of codon bias

- 1) Strongest in most highly expressed genes.
- 2) Selection is *not* silent at the third position after all.
- 3) Transcriptional efficiency is probably the driving factor in codon bias:
 - tRNAs match most common mRNAs produced in cells.
 - mismatch (i.e., mutation to a new codon) reduces translational efficiency of transcript -> selected against.

Neutrality in the Docket

- silent site changes and pseudogenes support a neutral model.
- The variation in functional positions and coding loci do not support a strictly neutral view.
- Conclusion: parts of the genome will evolve neutrally.
 - these portions will prove very useful in some population genetic analyses.
 - estimates suggest 95% of human genome is non-coding (junk) DNA.

More Uses of Neutral Theory

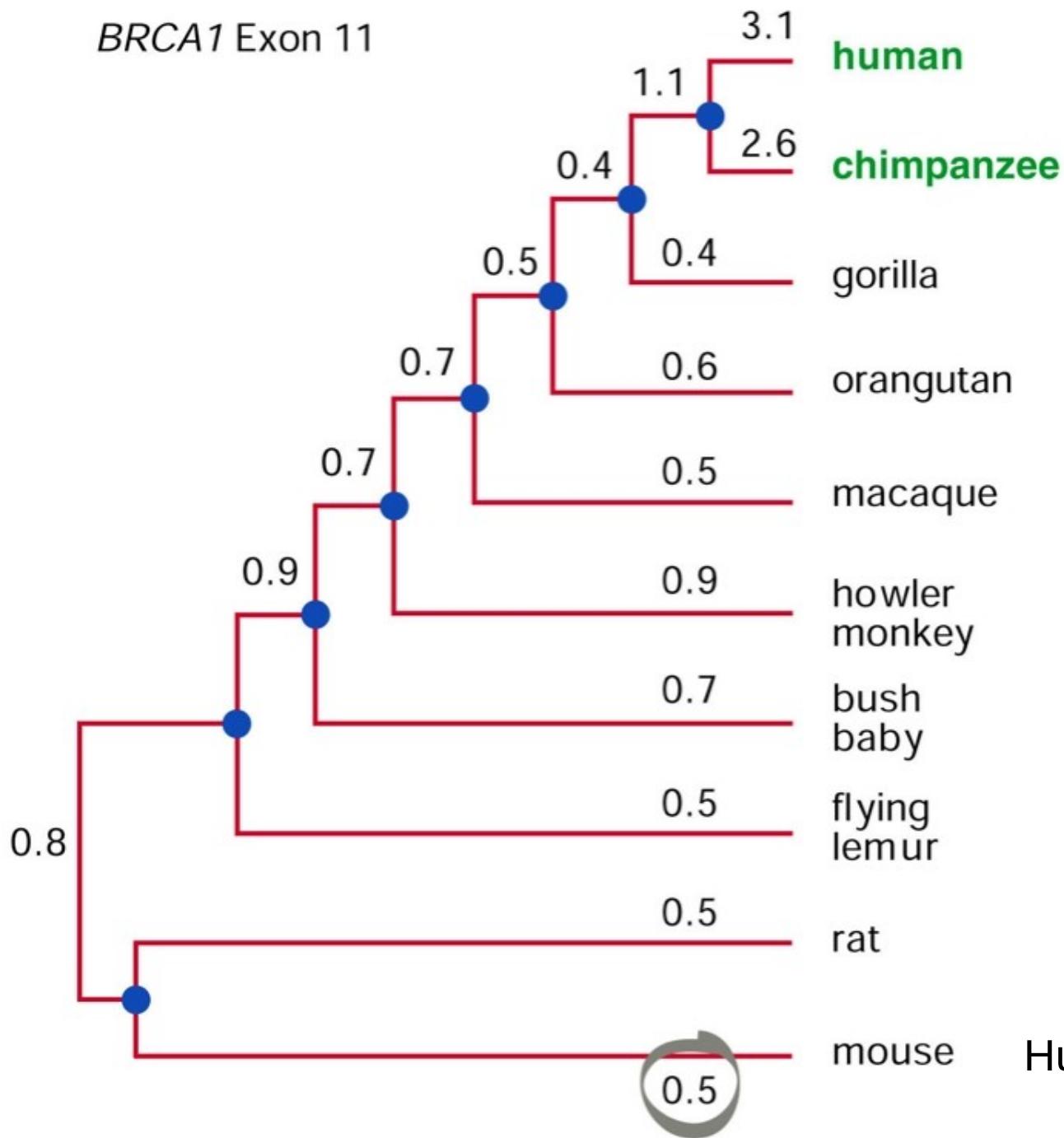
- Explains several important patterns in molecular evolution.
- Like Hardy-Weinberg Equilibrium, provides a null hypothesis for testing adaptive evolution:
 - *compare change to neutral expectation to determine if natural selection has shaped the sequence.*

Looking for Positive Evolution

- Hughes & Nei looked at major histocompatibility complex (MHC) gene sequences.
 - membrane proteins important in immune system recognition of infected cells.
- synonymous substitution rate used to estimate the mutation rate, ν .
 - provides a benchmark for gaugeing non-synonymous (replacement) rate of change.
- H & N found *higher* Non synonymous rates in the antigen recognition system than predicted by the Synonymous site benchmark.

- Higher replacement rate can only be explained by positive selection.
- i.e.,
- $d_{NON} / d_{SYN} > 1$ only when replacements advantageous.
- $d_{NON} / d_{SYN} < 1$ when replacements disadvantageous

BRCA1 Exon 11



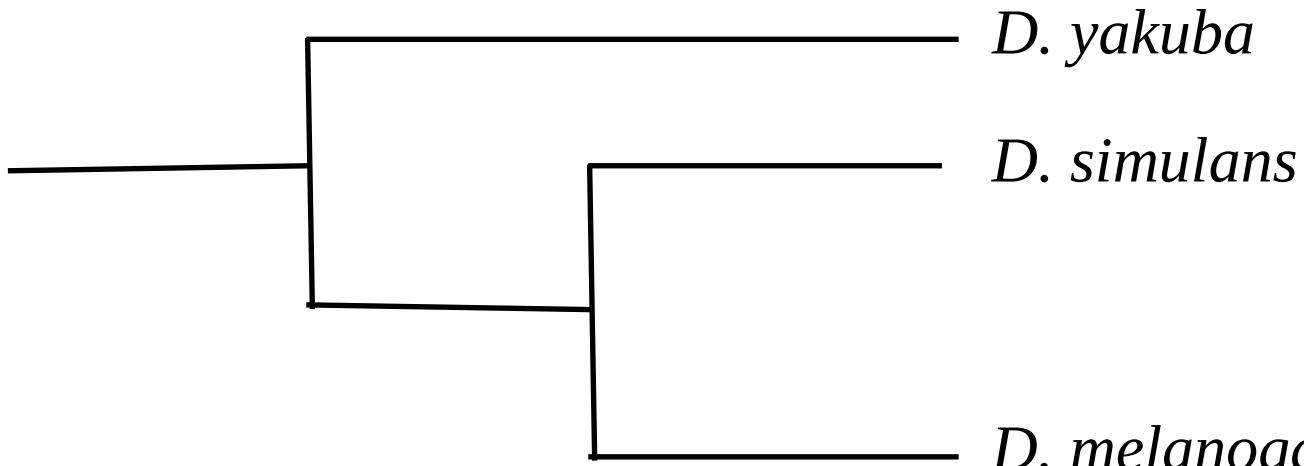
Hutley G et al. 2

$$d_{NON} / d_{SYN} > 1$$

- Unduly conservative index of positive selection (Sharp 1997).
- McDonald & Kreitman refined the silent vs. replacement hypothesis:
 - ratio silent:substitution is constant through time according to neutral theory.
 - use ratio to compare within-species change to between-species change.



MK test: *Adh* in *Drosophila*



- *Adh* (Alcohol dehydrogenase) important for fruit flies because they live on rotting fruit.
- M&K scored number of fixed vs. polymorphic sites based on sequence data.
- Null (neutral prediction): within species ratio of silent:replacement substitutions $\approx 20:1$

The Genes to Watch: Evolving at Warp Speed

- genes recruited to new functions (e.g., duplications)
- sex-determination genes
- fertilization interactions (sperm-egg; pollen-stamen)
- some enzymes, regulatory proteins
- immune-system genes
- sexual conflict genes?

Inbreeding

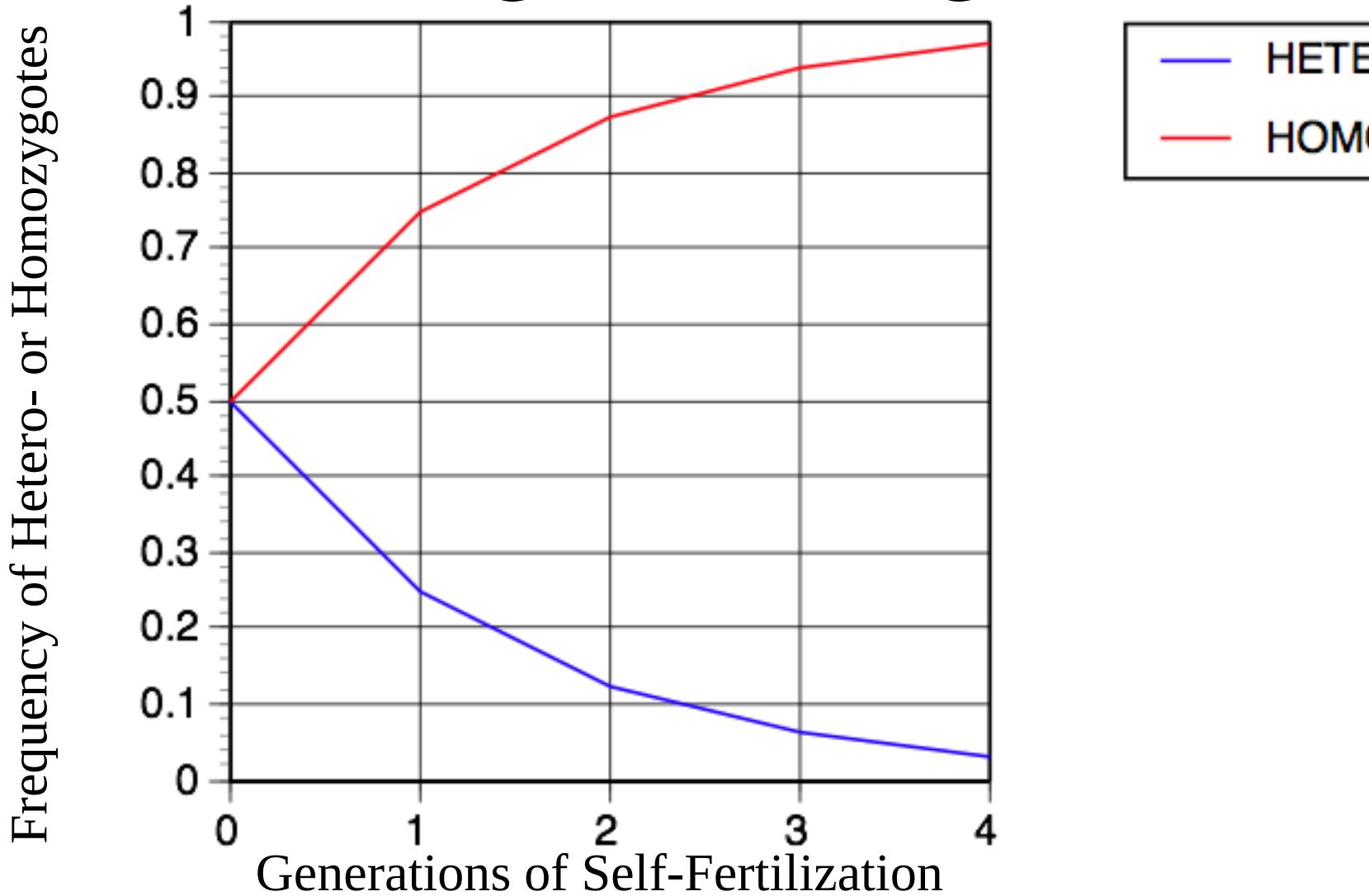
- Mating between related individuals
- F = inbreeding coefficient (probability that two homologous alleles in an individual are identical by descent)

- Identity by Descent: Two alleles created from the same ancestral copy of the gene
- Identity by Type: Two alleles sharing the same nucleotide sequence but not a common ancestor.
 - Some more terms-
 - Autozygous (Always homozygous)
 - Allozygous (Homozygous or heterozygous).

Genotype frequencies with Inbreeding

- $AA = p^2(1-F) + pF$
- $aa = q^2(1-F) + qF$
- $Aa = 2pq(1-F)$
- $H = H_0(1-F)$
- Pedigrees are used to compute F
- Inbreeding INCREASES genotypic variance
- Inbreeding does not change allelic frequencies

Inbreeding via Selfing

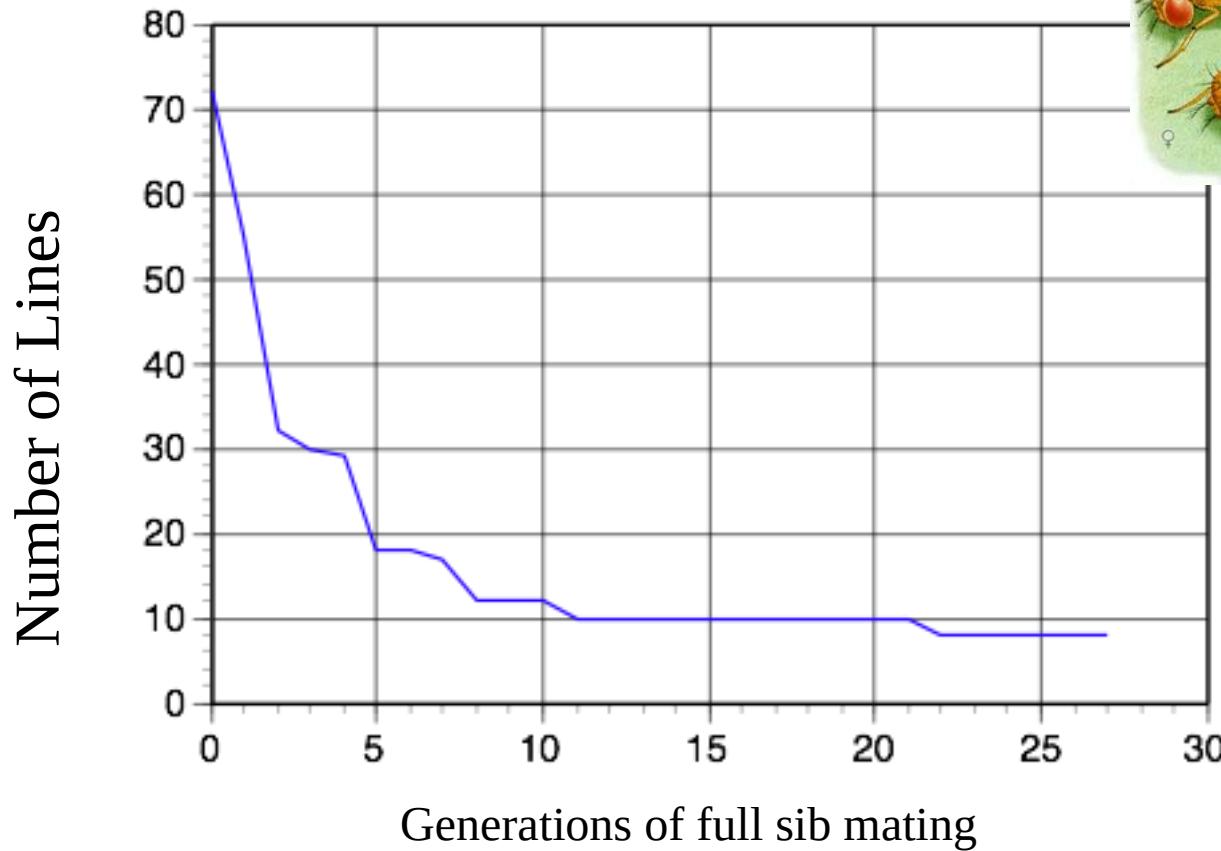


Despite changes in GENOTYPE FREQUENCIES,
individuals have the same ALLELE FREQUENCIES

Melting down the Vortex

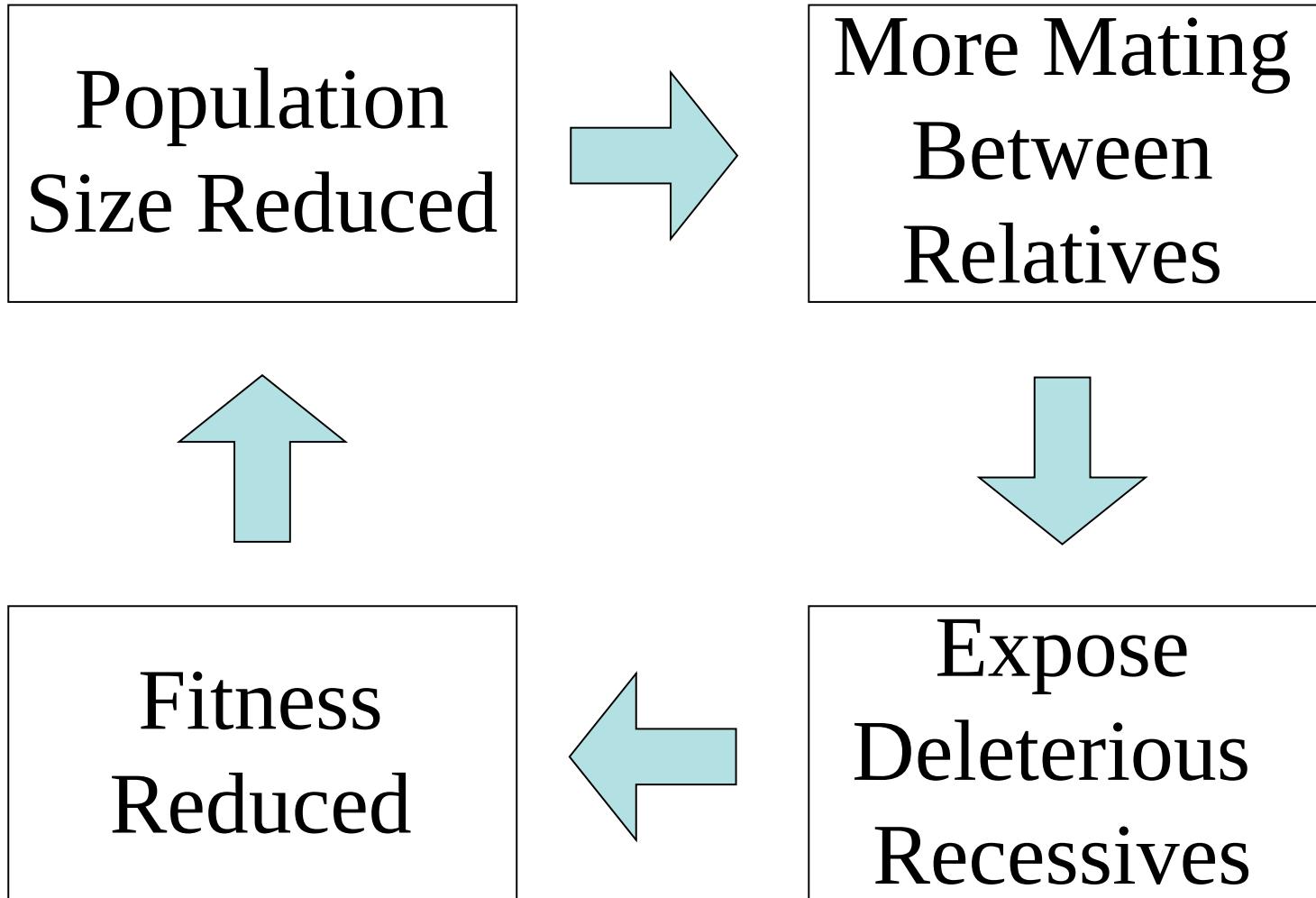
- Inbreeding and drift have similar effects – increased homozygosity.
- Genetic Drift & Mutation have profound implications for the conservation of populations.
- recall the Wahlund Effect:
 - as continuous populations become fragmented, the frequency of homozygotes increases due to drift.
 - homozygosity increases; relatives mate with relatives to an ever greater degree.

Inbreeding depression



- exposure of recessive deleterious variation leads to a syndrome of low performance:
 - = low survival & low fertility

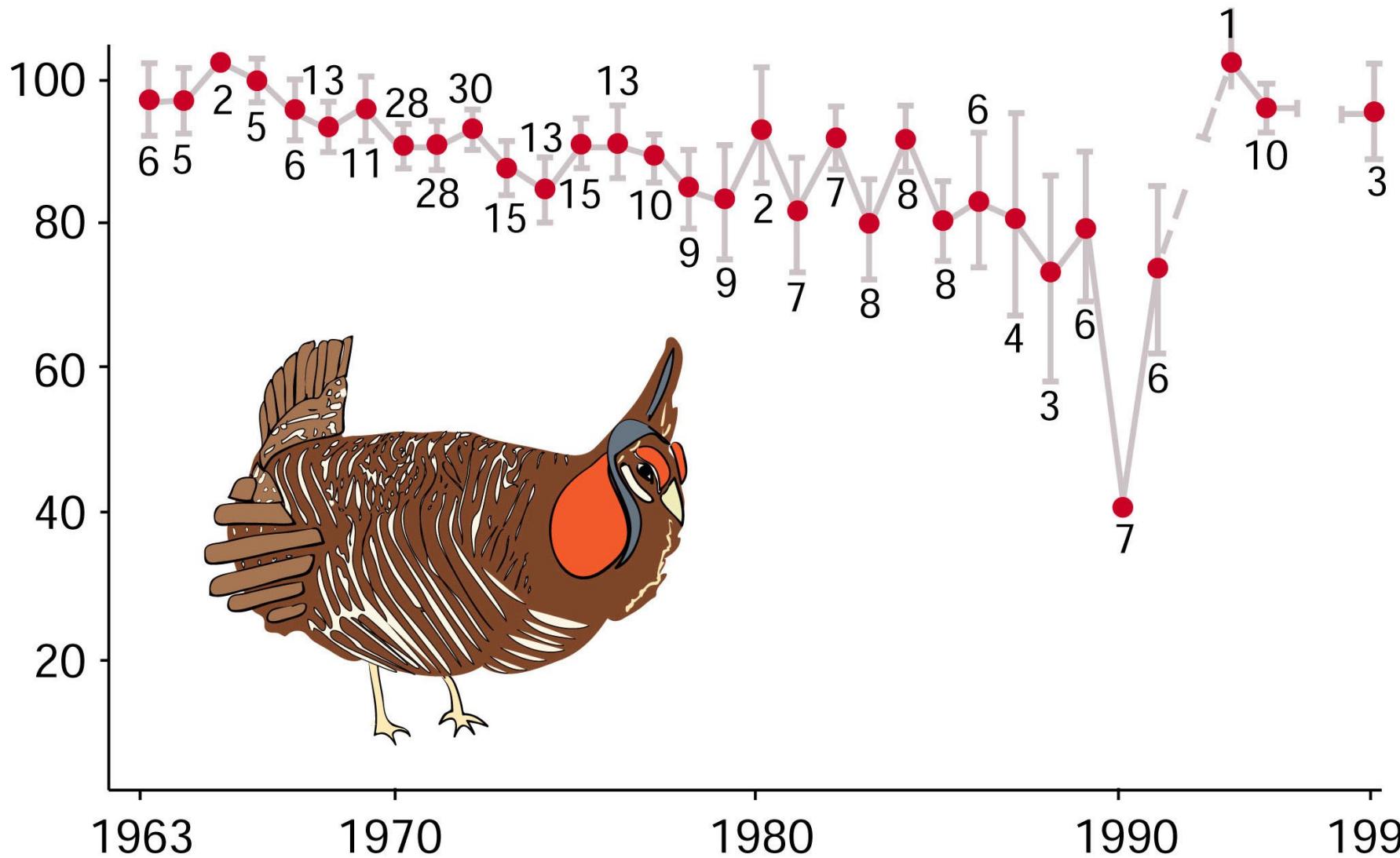
Mutational Meltdown



genetic load increases

Meltdown or Vortex?

- Inbreeding is a strong form of genetic drift
- Causes exposure of deleterious recessive variation hidden in large populations.
- As fitness declines because of increased Genetic Load, population size shrinks further.
- This intensifies drift (i.e., further inbreeding)
- Synergism of effects leads to extinction.

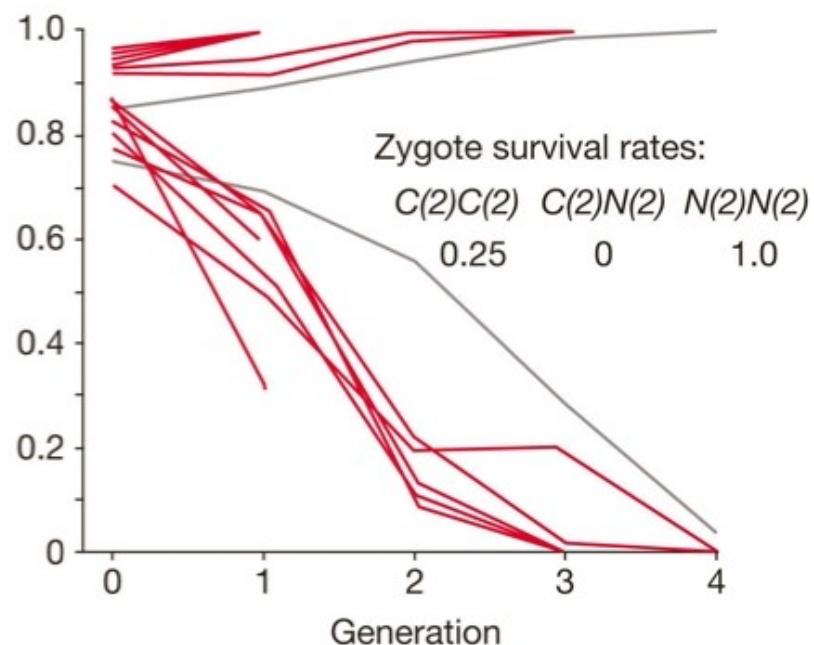
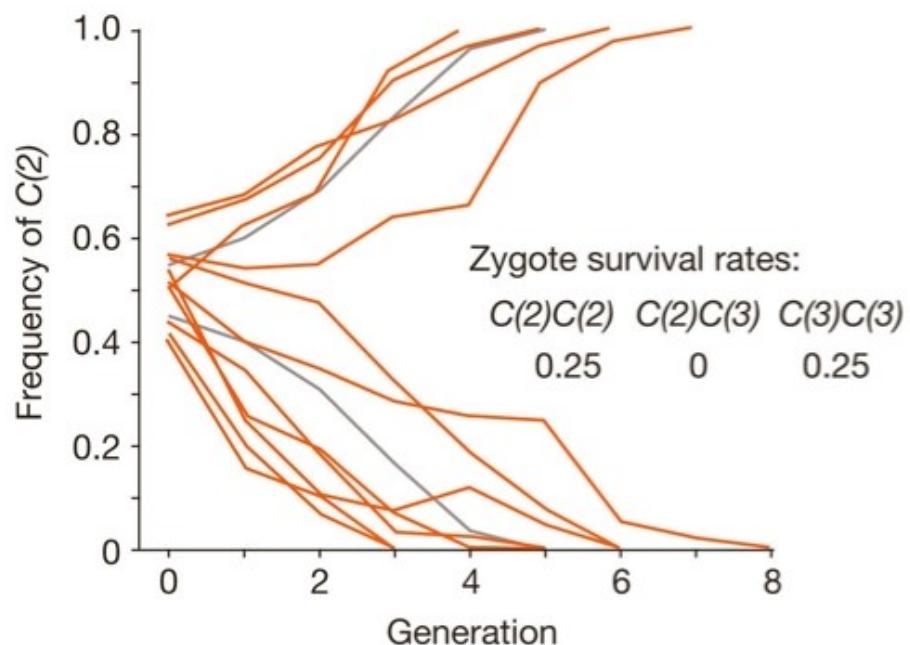


Conservation of the Greater Prairie Chicken

- allelic richness had declined to about 60% of neighbouring populations (or since the 1930s).
- strategies based on habitat recovery were largely unsuccessful.
- importation of neighbouring stock from other states lead to dramatic reversal
 - a little bit of gene flow goes a long way
 - conservation also needs to be cognizant of the breeding system and effective population size



(f) Left: Evolution in 11 populations of *Drosophila melanogaster* containing compound second chromosomes [C(2)] and compound third chromosomes [C(3)]. The initial frequency of C(2) ranged from 0.4 to 0.65. Right: Evolution in 13 populations of *Drosophila melanogaster* containing a mixture of compound second chromosomes [C(2)] and normal second chromosomes [N(2)]. The initial frequency of C(2) ranged from 0.71 to 0.96.



Selection against

- Homozygotes: $\Delta q = pq(tp-sq)/(1-tp^2-sq^2)$
- Heterozygotes: $\Delta q = pq(sq-tp)/(1+tp^2+sq^2)$
- $q_e = t/(t+s)$

Adding Selection to H-W Equilibrium

Genotype Frequencies
(population n=100)

- $B_1 = 0.60; p^2 = 0.36$ (expect 36 B_1/B_1)
- $B_2 = 0.40; q^2 = 0.16$ (expect 16 B_2/B_2)
- then we expect $2pq$ heterozygotes = 0.48; (expect 48 B_1/B_2)

$$\text{FREQUENCY } B_1 = 60 / 100 = 60\%$$

$$\text{FREQUENCY } B_2 = 40 / 100 = 40\%$$

Strong selection against an allele

additive variation: B1 favoured:

- B1/B1 homozygotes survive at 100%
- B1/B2 heterozygotes survive at 75%
- B2/B2 homozygotes survive at 50%

After selection

- each genotype produces 10 gametes
- # of B1/B1 = $36 * 1.0 * 10 = 360$
- # of B1/B2 = $48 * 0.75 * 10 = 360$
- # of B2/B2 = $16 * 0.5 * 10 = 80$

FREQUENCY B1 = $540 / 800 = 67.5\%$

FREQUENCY B2 = $260 / 800 = 32.5\%$