MUTATIONS IN TIME:

SOME BASICS OF POPULATION GENETICS

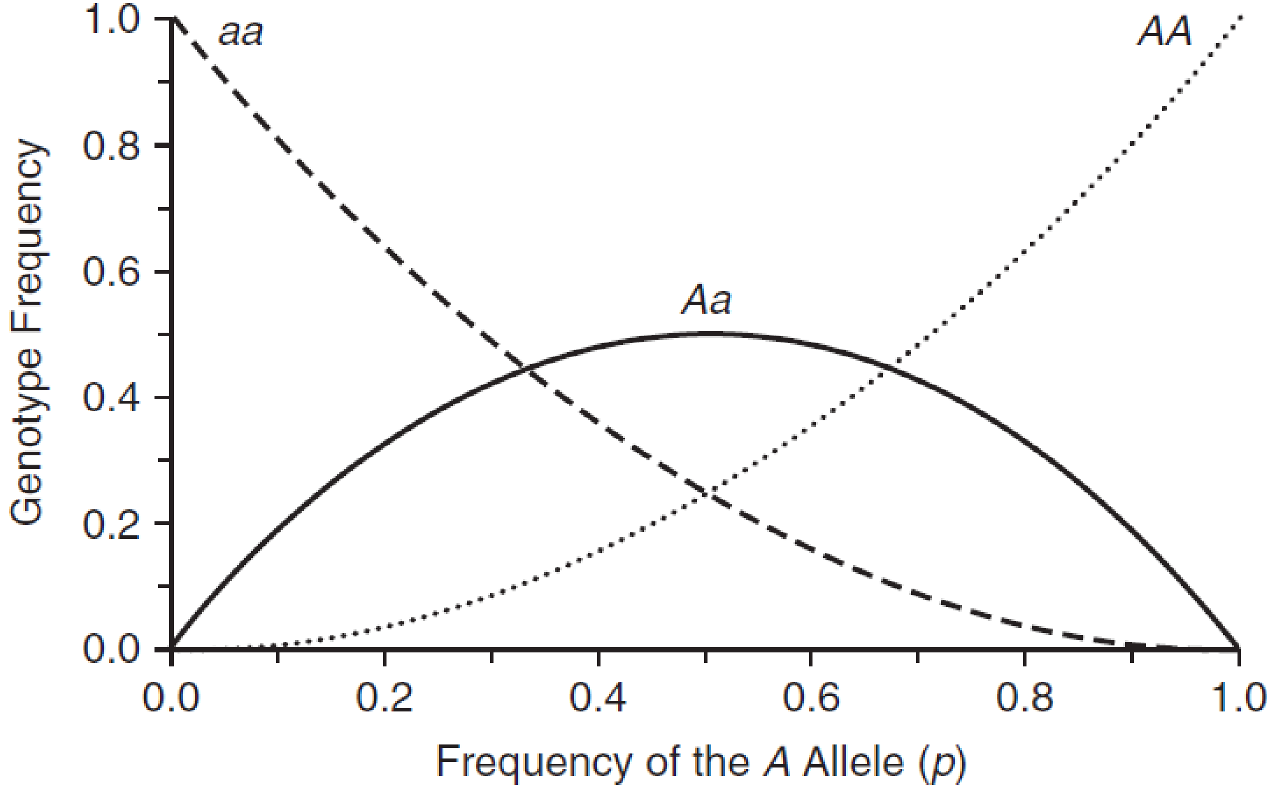
Lecture plan

- Hardy-Weinberg equilibrium
- Random genetic drift without mutations
- Effective population size
- Random genetic drift and mutations
- The coalescent theory
- Natural selection. Mutation-selection balance
- Random genetic drift, positive selection
- Selection coefficients, deleterious alleles
- Non-random mating, population subdivision, gene flow, admixture, adaptation

Hardy-Weinberg equilibrium (1908)

$$\text{Generation } N: \;\; f_A = p,\; f_a = q,\; p + q = 1$$

$$\text{Generation } N+1: \;\; F_{AA} = p^2,\; F_{Aa} = 2pq,\; F_{aa} = q^2$$



Hardy-Weinberg equilibrium

$$\text{Generation } N: \; f_A = p, \; f_a = q, \; p + q = 1$$

$$\text{Generation } N+1: \; F_{AA} = p^2, \; F_{Aa} = 2pq, \; F_{aa} = q^2$$

Implications:

1. The allele frequencies does not change:

$$p' = f'_A = F'_{AA} + F'_{Aa}/2 = p^2 + pq = p$$

Exercise: derive this

2. HWE frequencies are attained in one generation

Hardy-Weinberg equilibrium

Assumptions:

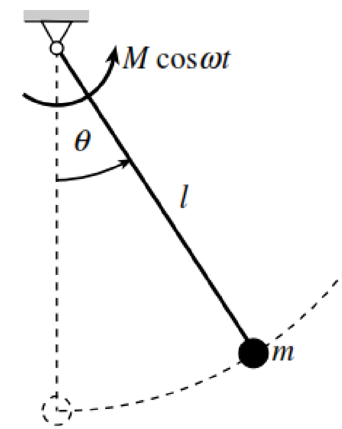
- Diploid species with sexual reproduction and random (not assortative) mating
- Same allele frequencies in males and females
- Non-overlapping generations
- Biallelic (autosomal) locus
- Population size is infinite
- No change in allele frequencies by migration, natural selection or mutation
- No genotyping errors

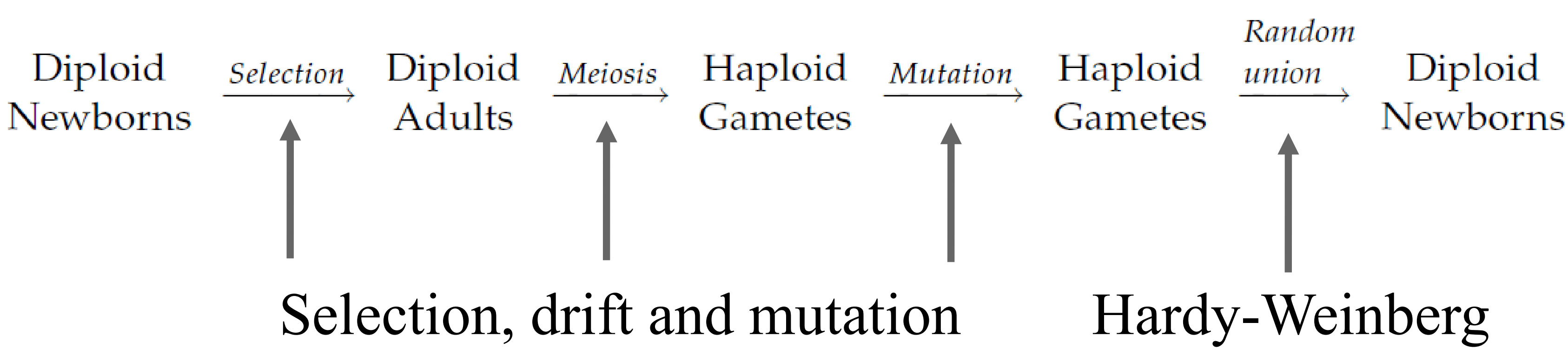
Hardy-Weinberg equilibrium

Does it still make sense with so many assumptions? Yes:

1. A baseline for more realistic models

2. The H-W model splits life history into two intervals: gametes → zygotes and zygotes → adults





Hardy-Weinberg equilibrium

Testing for HWE:

df = $n - k - 1$, where $n = 3$ is the number of classes and $k = 1$ is the number of independent parameters

Genotype	Observed Number (O)	Expected Number (E)	$(O-E)$	$(O-E)^2$	$(O-E)^2/E$
AA	90	83.2	6.8	46.24	0.5558
Aa	28	41.6	−13.6	184.96	4.4462
aa	12	5.2	6.8	46.24	8.8923

After performing the calculations in this table, we get a chi-square (χ^2) statistic of

$$\chi^2 = 0.5558 + 4.4462 + 8.8923 = 13.8943$$

This value is *much* larger than the critical value of 3.841, so we reject the hypothesis of Hardy–Weinberg equilibrium.

Exercise: do it yourself

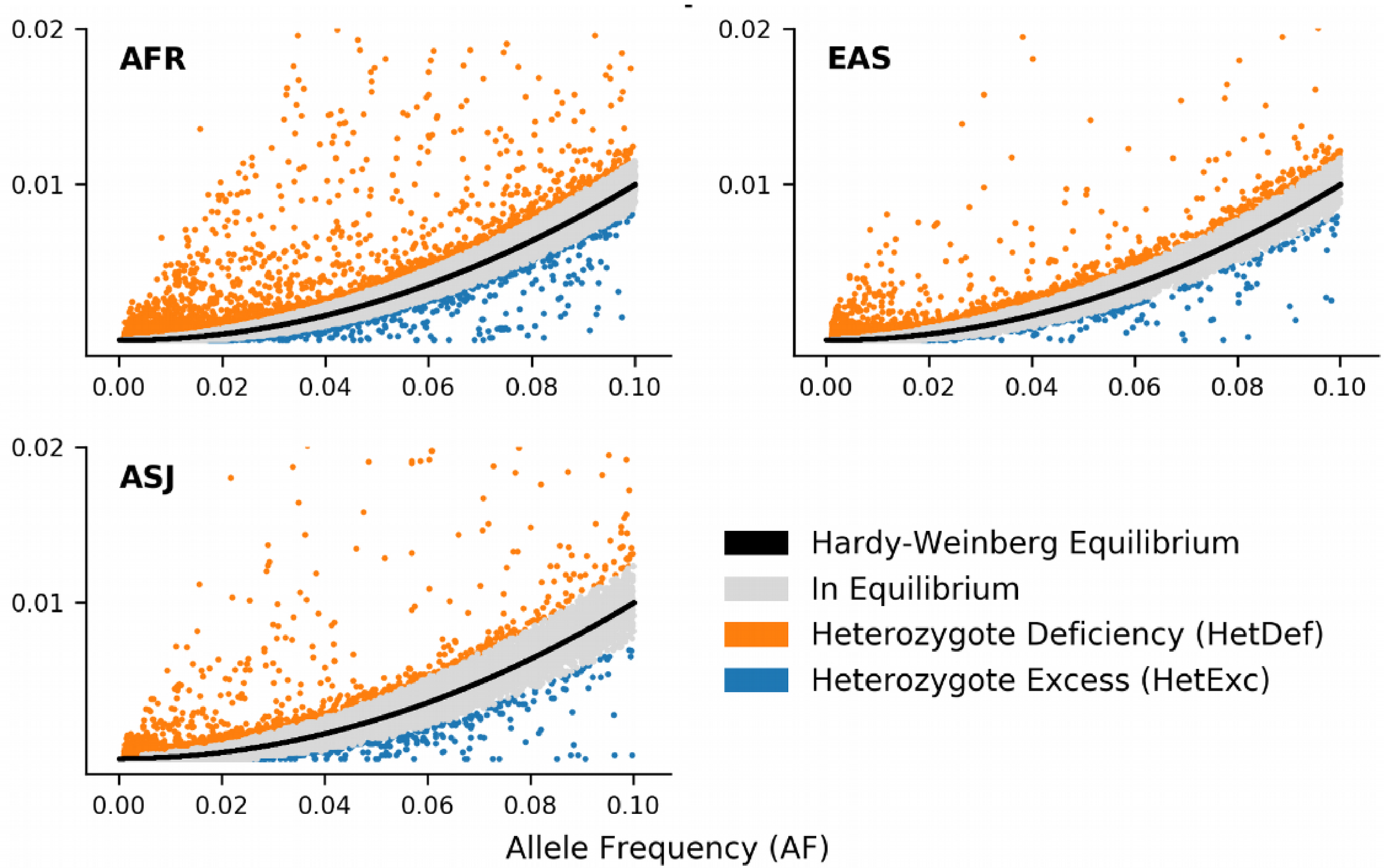
$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Relethford – *Human Population Genetics*

Hardy-Weinberg Equilibrium in the Large Scale Genomic Sequencing Era

Nikita Abramovs, Andrew Brass, May Tassabehji

doi: https://doi.org/10.1101/859462



gnomAD: 137,842 predominantly healthy individuals from 7 major ethnic populations

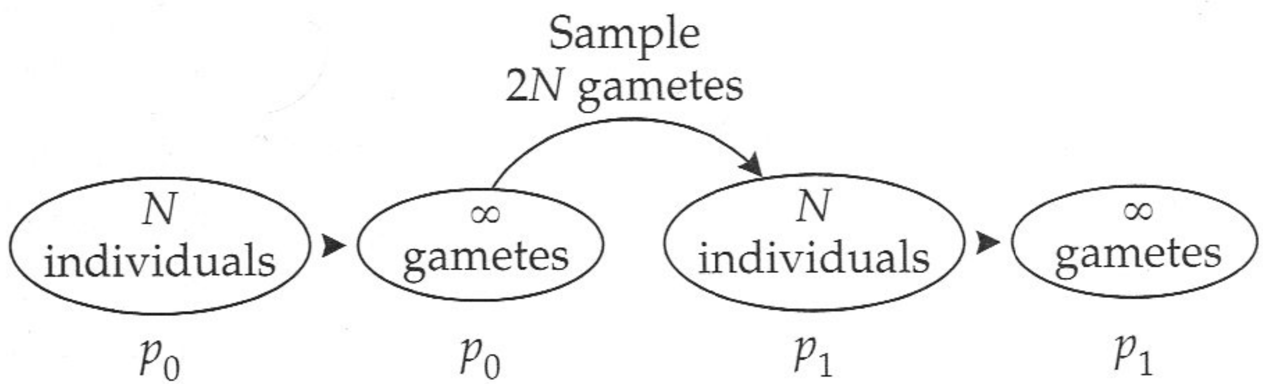
Random genetic drift (Wright-Fisher, 1930)

Assumptions:

- Diploid species with sexual reproduction and random (not assortative) mating
- Same allele frequencies in males and females
- Non-overlapping generations
- Biallelic (autosomal) locus
- ~~Population size is infinite~~
- No change in allele frequencies by migration, natural selection or mutation
- No genotyping errors

Random genetic drift

Finite population $\Longrightarrow$ Sampling variation $\Longrightarrow$
Allele frequency fluctuations $\Longrightarrow$ Random genetic drift



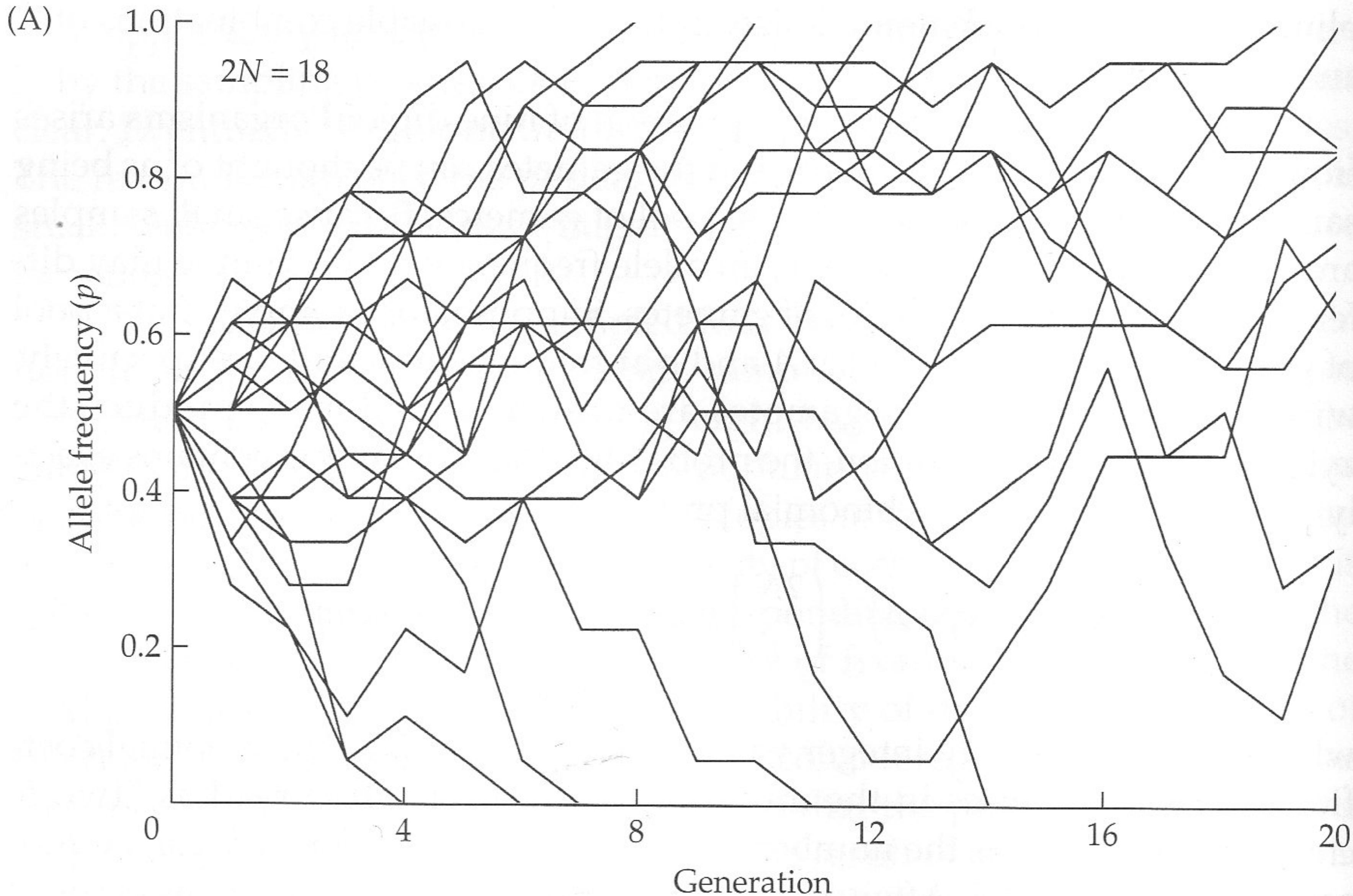
$$P(k) = \binom{2N}{k} p_0^k (1 - p_0)^{2N-k}$$

$$E(\Delta p \,|\, p) = E(k/2N - p \,|\, p) = 0$$

Exercise: derive

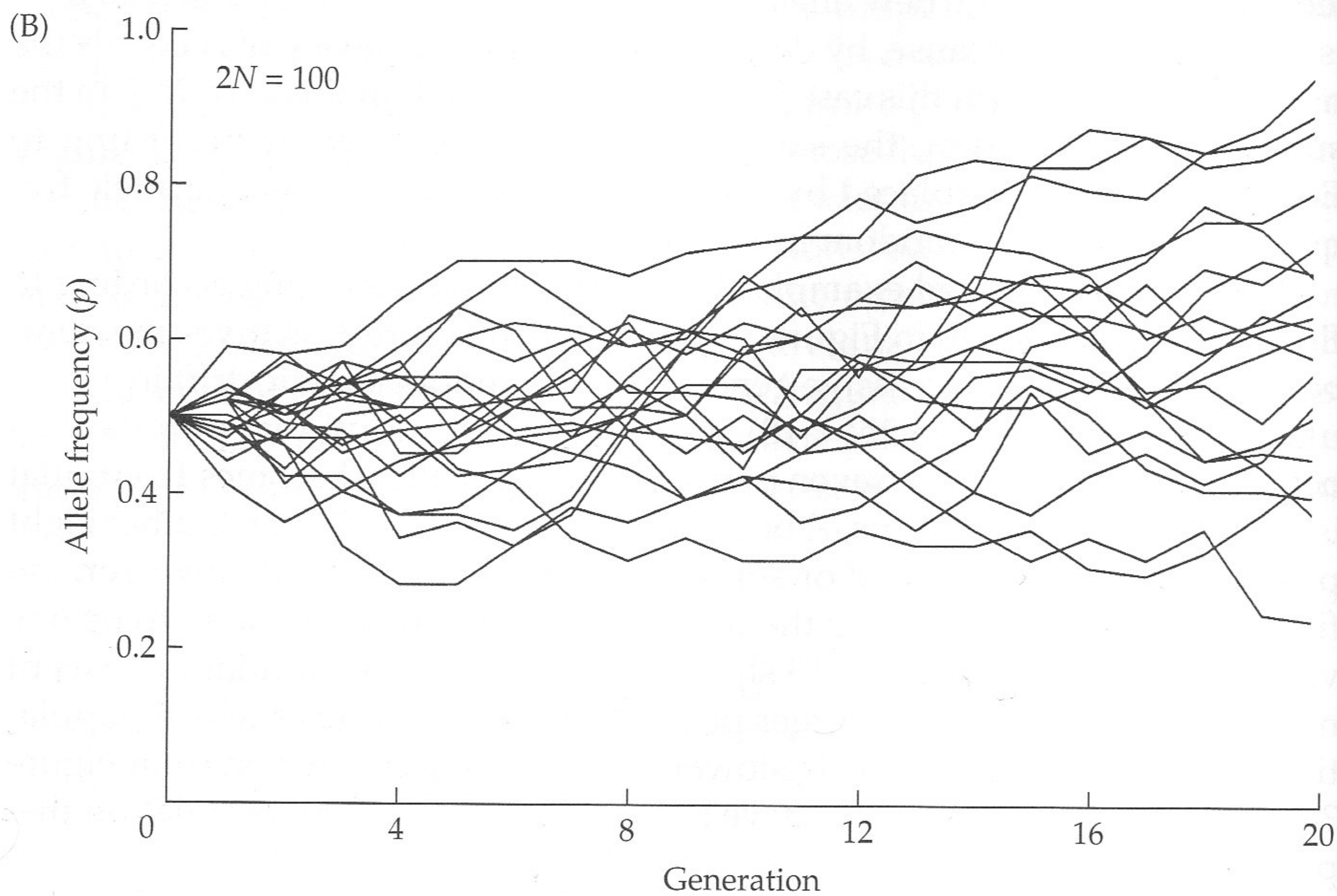
$$Var(\Delta p \,|\, p) = Var(k/2N - p \,|\, p) = p(1-p)/2N$$

Random genetic drift



Hartl & Clark – *Principles of population genetics*

Random genetic drift



Hartl & Clark – *Principles of population genetics*

Random genetic drift

The endpoint is allele fixation or loss: $P(F|p) = p$

Mean time to fixation, if fixed: $\overline{t_F}(p) = -4N\left(\frac{1-p}{p}\right) ln(1-p)$

Mean time to loss, if lost: $\overline{t_L}(p) = -4N\left(\frac{p}{1-p}\right) ln(p)$

Mean perstistence time: $\bar{t}(p) = p\overline{t_F}(p) + (1-p)\overline{t_L}(p) =$

$$= -4N[(1-p)ln(1-p) + p \cdot ln(p)]$$

Exercise: at which *p* persistence time is maximal and what is it?

Exercise: estimate $t_F(p)$ when $p \to 0$

Random genetic drift and genetic variation

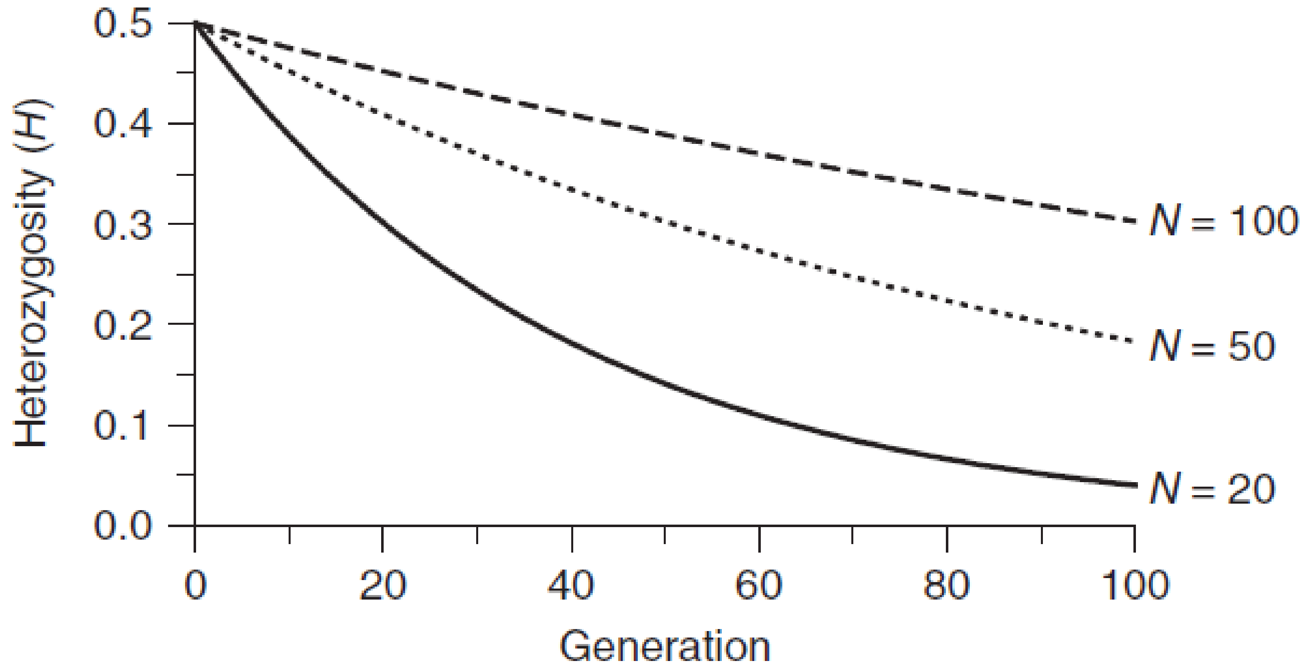
Heterozygosity: probability that an individuum is heterozygous at a locus: $H = 2pq$

$$H_{t+1} \simeq H_t - H_t/2N$$

Heterozygosity decay due to drift:

$$H_t = H_0(1 - 1/2N)^t$$

Decay is slow: $H_t = H_0/2 : \ t \approx 2N \ln(2) \ for \ N \gg 1$



Random genetic drift and genetic variation

Heterozygosity: probability that an individuum is heterozygous at a locus: $H = 2pq$

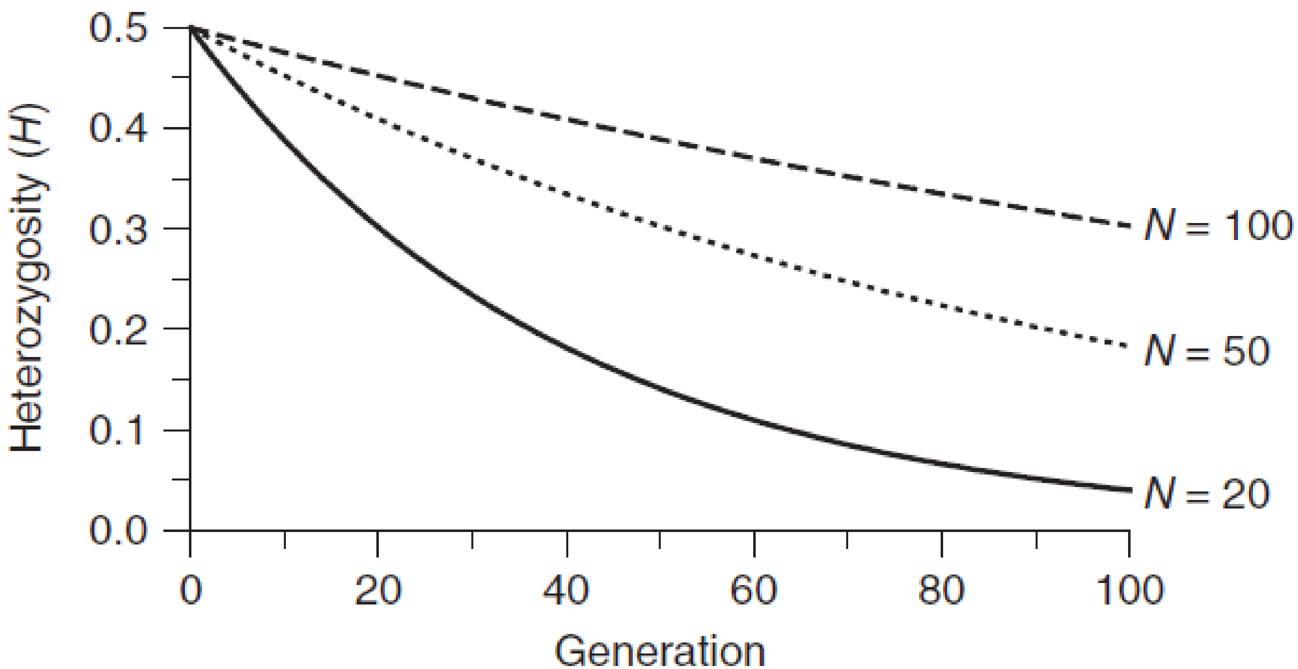
Drift strength is ≈1/2*N*

$$H_{t+1} \simeq H_t - H_t/2N$$

Heterozygosity decay due to drift:

$$H_t = H_0(1 - 1/2N)^t$$

Decay is slow: $H_t = H_0/2 : \; t \approx 2N \ln(2) \; for \; N \gg 1$



Effective population size

Effective population size of an actual population is the number of indivduals in a theoretically ideal population having the same magnitude of genetic drift as the actual population (Hartl & Clark, *Principles of population genetics*)

- Fluctuation in population size $\frac{1}{N_e} = \frac{1}{t}(\frac{1}{N_0} + \frac{1}{N_1} + \cdots + \frac{1}{N_{t-1}})$
- Unequal sex ratio: $N_e = \frac{4N_m N_f}{N_m + N_f}$

Exercise: bottleneck consequences for N_e

- Variance in offspring number: σ, ξ – offspring mean and variance $N_e = \frac{N-1}{(\sigma^2/\xi) + (\xi - 1)}$
- Subdivided population: *d* sub-populations of size *N*; *m*, migration $N_e = Nd\left(1 + \frac{1}{4Nm}\right)$

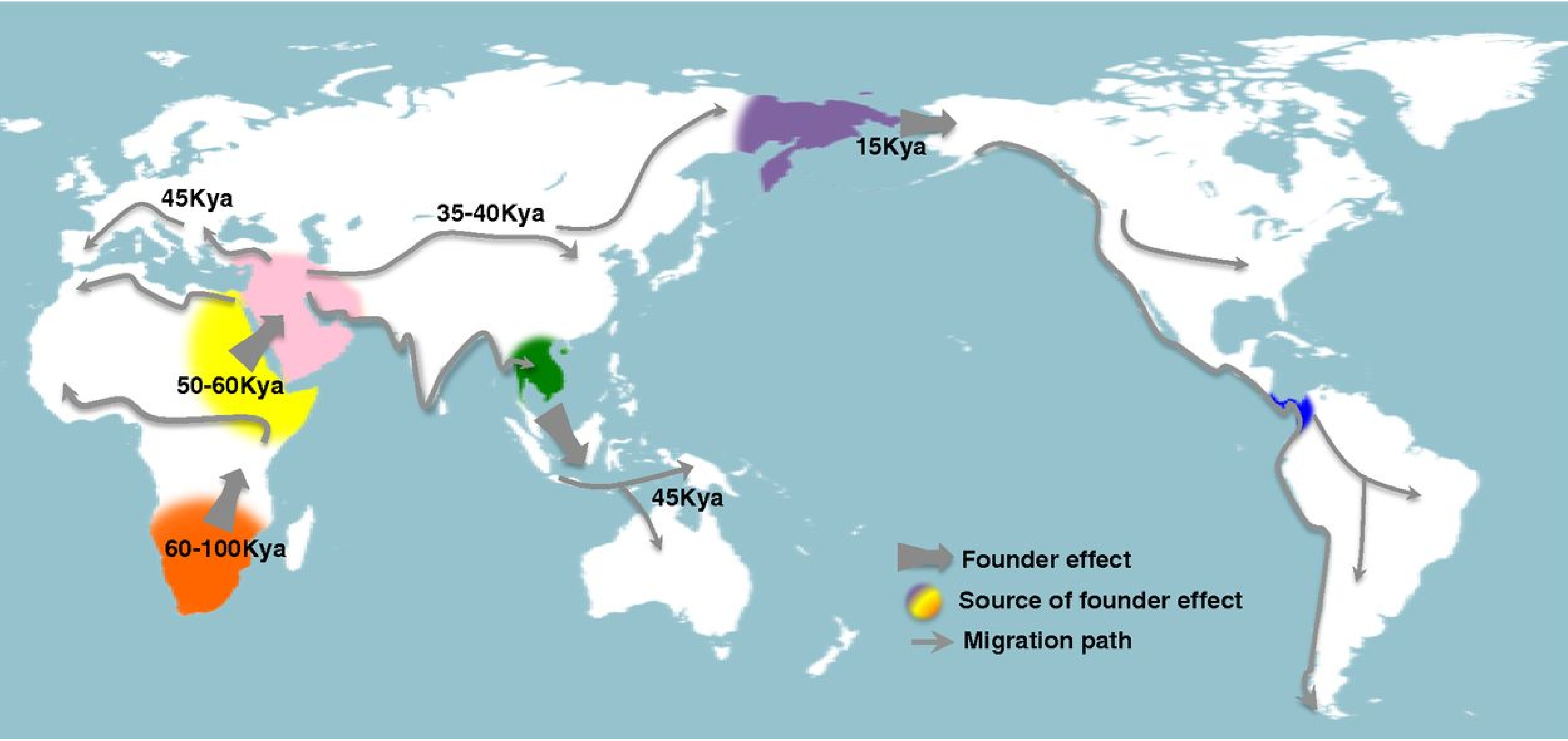
People living on Earth

7,849,058,679

All on this page, one by one

watch as we increase

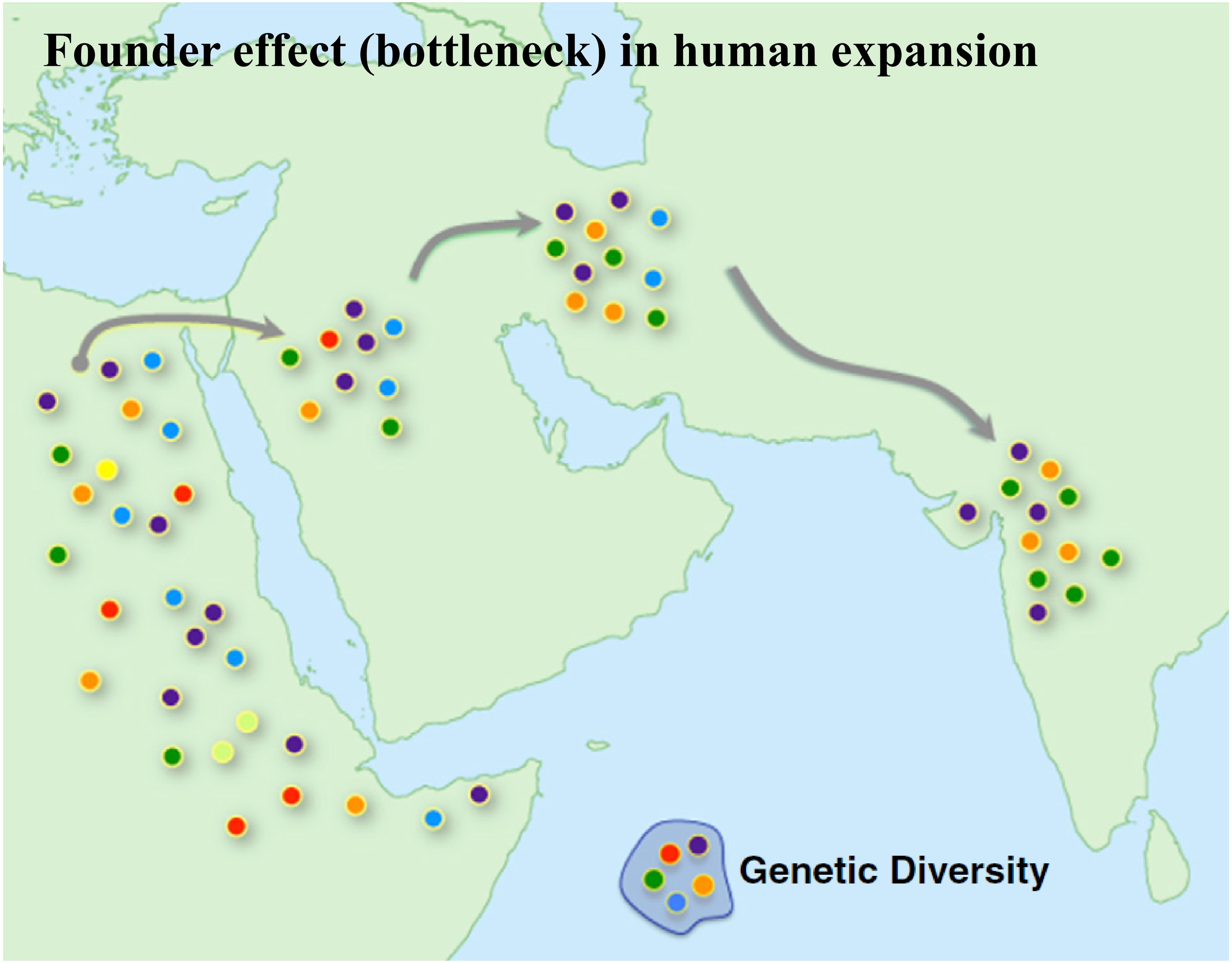
The great human expansion



Resequencing studies have estimated the ancestral effective population size at 12,800 to 14,400, with a 5- to 10-fold bottleneck beginning approximately 65,000 to 50,000 y ago (although see ref. 15 for a bottleneck to only 450 individuals).

Henn *et al* (2012) *PNAS*

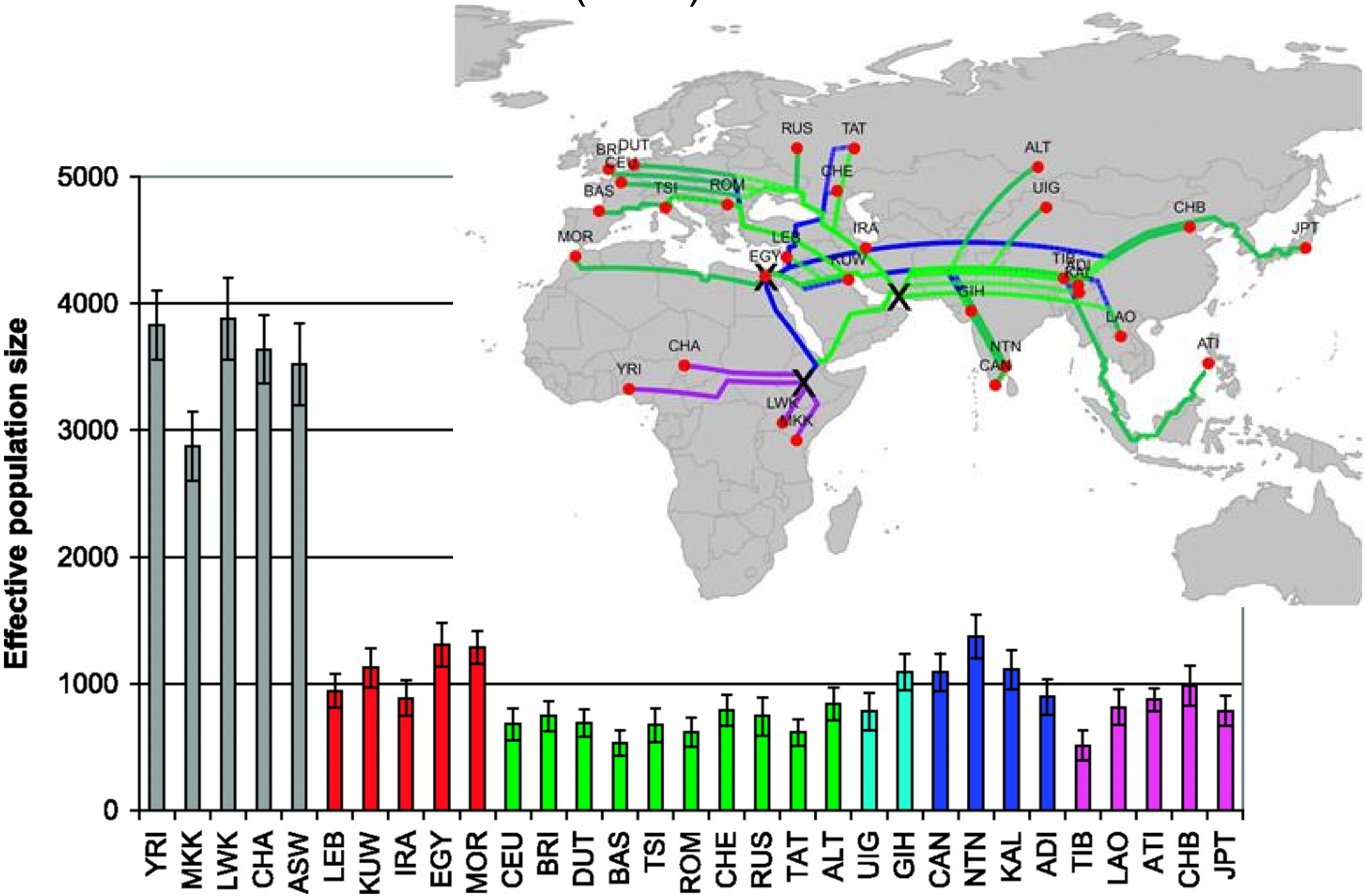
The great human expansion



Henn *et al* (2012) *PNAS*

Recombination Gives a New Insight in the Effective Population Size and the History of the Old World Human Populations

Mele *et al* (2011) *Mol Biol Evol*



Random genetic drift and mutations

The neutral theory: most mutations are selectively neutral with allele frequency determined by random genetic drift (Kimura 1968)

$2N$ gametes $\Longrightarrow$ $2N\mu$ mutations in each generation, where μ = mutations per gamete per generation

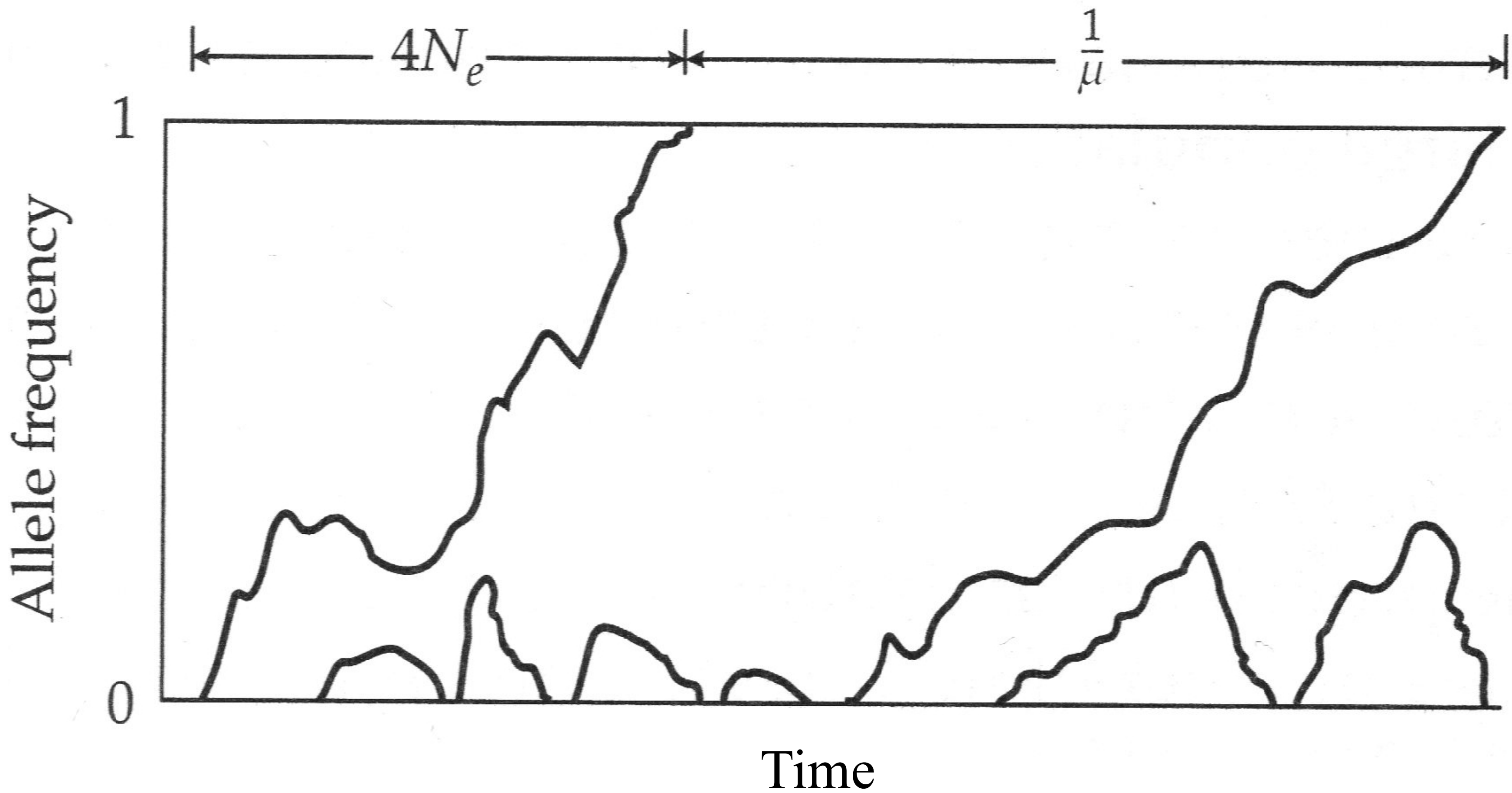
Each mutation $p_0 = 1/2N \Longrightarrow P_{\mathrm{Fix}} = 1/2N$

The steady-state rate at which neutral mutations are fixed in a population: $k = 2N\mu P_{\mathrm{Fix}} = \mu$

Q: What is the average time between fixation events?

Mean time to fixation, if fixed: $t_{\mathrm{F}}(p) = 4N_{\mathrm{e}}$ for $p \approx 0$

Random genetic drift and mutations



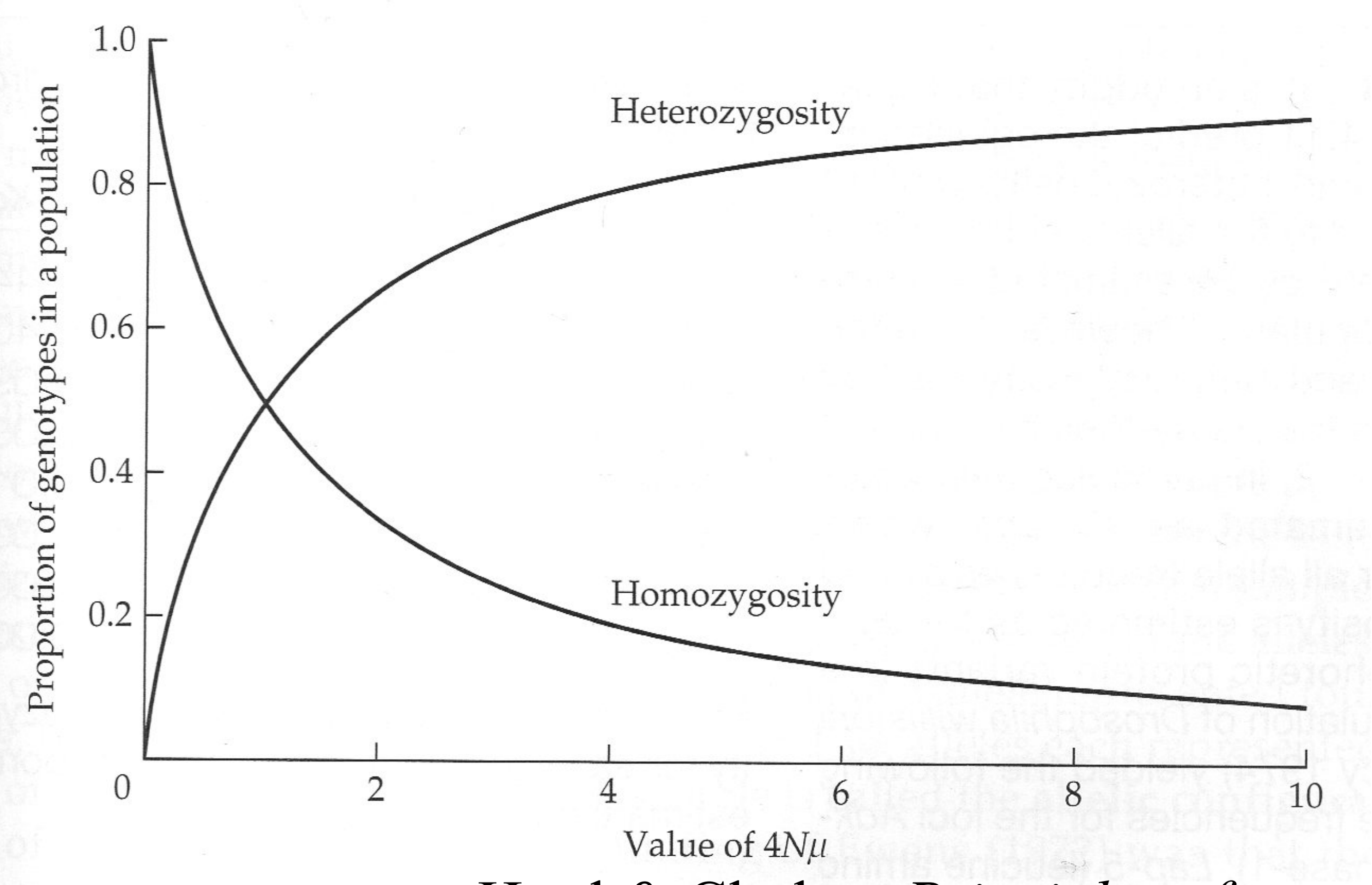
Exercise: estimate fixation time for a new neutral allele

Hartl & Clark – *Principles of population genetics*

Random genetic drift and mutations

The infinite-alleles model: each mutation creates a new allele in the population

$$\text{Heterozygosity } H = \frac{\theta}{1+\theta}, \; where\ \theta = 4N_e\mu$$



Hartl & Clark – *Principles of population genetics*

Random genetic drift and mutations

The infinite-alleles model: each mutation creates a new allele in the population

$$\text{Heterozygosity } H = \frac{\theta}{1+\theta},\ \textit{where } \theta = 4N_e\mu$$

N_e: effective population size, **~10,000**

μ: mutation rate per site per generation, **~1.2×10^{-8}**

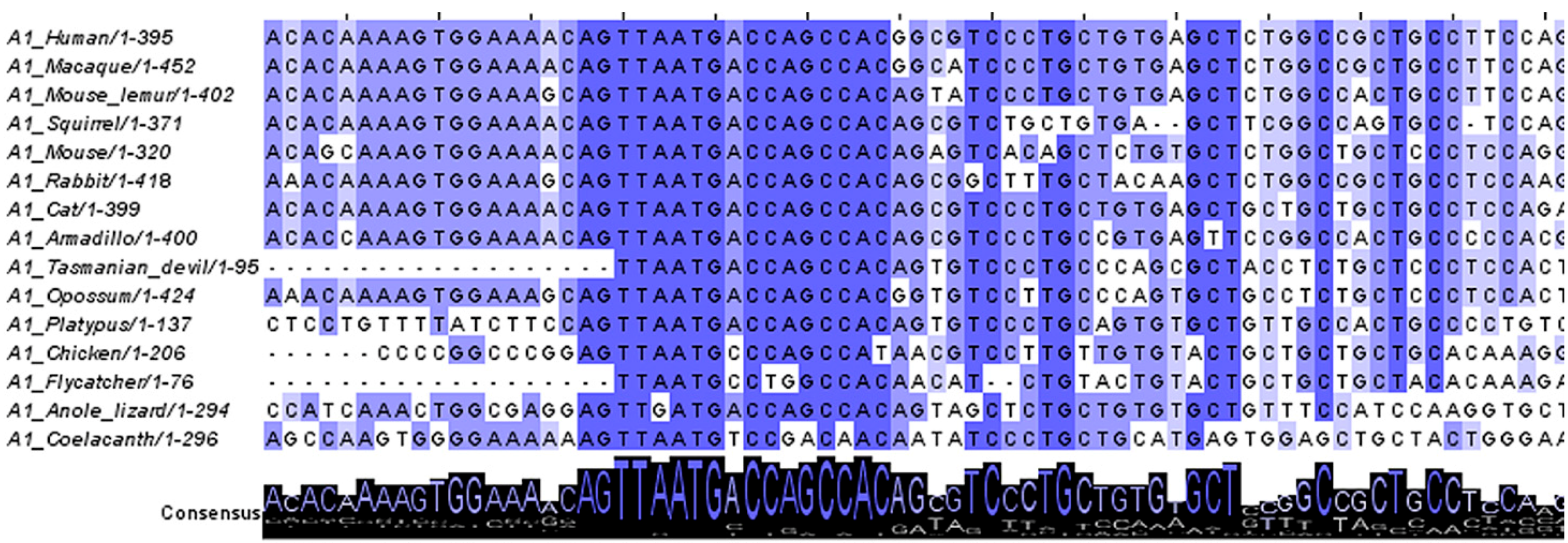
$$\theta = 4\times10^{4}\times1.2\times10^{-8} \approx 5\times10^{-4}$$

$$\theta << 1 \Longrightarrow H \approx \theta = 1/2000$$

Random genetic drift and mutations

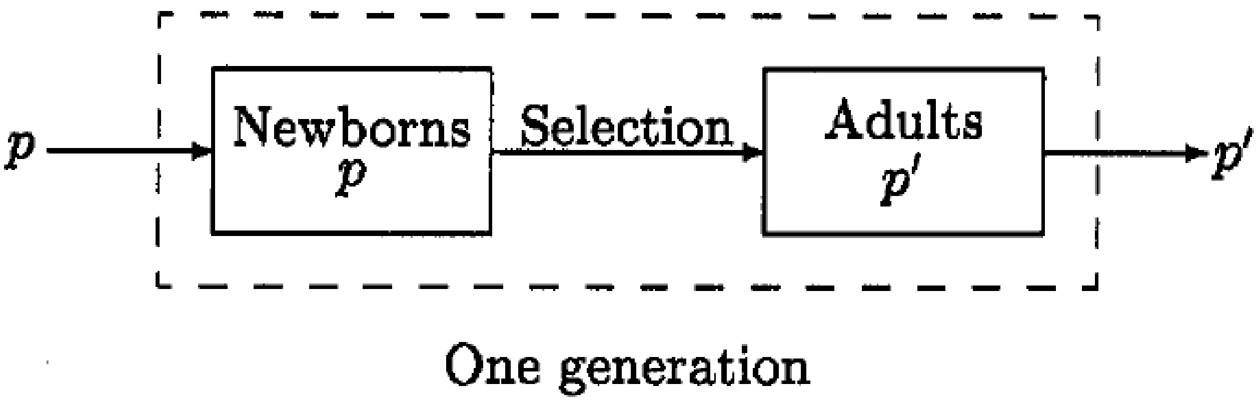
The neutral (Motoo Kimura) and nearly neutral (Tomoko Ohta) theory of molecular evolution (1960-70):

- Random genetic drift of [nearly] neutral alleles is the source of polymorphism, not balancing selection.
- Most substitutions (fixations) are due to random drift of neutral mutants, not advantageous mutations
- Missing substitutions are then evolutionary forbidden



Natural selection

Natural selection is the differential survival and reproduction of individuals resulting from differences in phenotype.
Natural selection changes allele frequencies:



Fitness is an individual's ability to propagate its alleles
≈ viability [+fertility+developmental time+mating, …]

Deleterious alleles reduce fitness (≠ pathogenic, damaging)

Gillespie – *Population genetics. A concise guide*

Natural selection

TABLE 5.2 Diploid Selection for Survivorship (Viability)

	Genotype			Total
Generation $t-1$	AA	Aa	aa	
Frequency before selection	p^2	$2pq$	q^2	$1 = p^2 + 2pq + q^2$
Relative fitness (viability)	w_{11}	w_{12}	w_{22}	
After selection	$p^2 w_{11}$	$2pq w_{12}$	$q^2 w_{22}$	$\bar{w} = p^2 w_{11} + 2pq w_{12} + q^2 w_{22}$
Normalized	$\frac{p^2 w_{11}}{\bar{w}}$	$\frac{2pq w_{12}}{\bar{w}}$	$\frac{q^2 w_{22}}{\bar{w}}$	$w_{11} > 0,\ w_{12} \geqslant 0,\ w_{22} \geqslant 0$
Generation t	$p' = \frac{p^2 w_{11} + pq w_{12}}{\bar{w}}$			
	$q' = \frac{pq w_{12} + q^2 w_{22}}{\bar{w}}$			

$$\Delta p = \frac{pq[p(w_{11} - w_{12}) + q(w_{12} - w_{22})]}{\bar{w}}$$

Exercise: derive

Hartl & Clark – *Principles of population genetics*

Natural selection

Genotype	A_1A_1	A_1A_2	A_2A_2
Viability (fitness)	w_{11}	w_{12}	w_{22}
Relative fitness	1	w_{12}/w_{11}	w_{22}/w_{11}

Natural selection

Genotype	A_1A_1	A_1A_2	A_2A_2
Viability (fitness)	w_{11}	w_{12}	w_{22}
Relative fitness	1	w_{12}/w_{11}	w_{22}/w_{11}
Relative fitness	1	$1-hs$	$1-s$

where $0 \le s \le 1$ is the **selection coefficient**,
h is the **heterozygous effect** and measures **dominance**

$h = 0$	A_1 dominant, A_2 recessive // 1, 1, $1-s$
$h = 1$	A_1 recessive, A_2 dominant // 1, $1-s$, $1-s$
$0 < h < 1$	incomplete dominance
$h = ½$	additivity // 1, $1-s/2$, $1-s$
$h < 0$	overdominance
$h > 1$	underdominance

Exercise: h<0, h>1

Natural selection

$$\Delta p = \frac{pq[p(w_{11}-w_{12})+q(w_{12}-w_{22})]}{\bar{w}}$$

Switch to relative fitness: $w_{12}/w_{11} = 1 - hs, \;\; w_{22}/w_{11} = 1 - s$

$$\Delta p = \frac{pqs[ph + q(1-h)]}{\tilde{w}}$$

$$\tilde{w} = 1 - 2pqhs - q^2 s$$

Exercise: derive

Gillespie – *Population genetics. A concise guide*

Natural selection

1. Directional (positive, negative, purifying) selection

Recessive allele: $w_{11}=1$, $w_{12}=1$, $w_{22}=1-s$ // $w_{12}=1$

Dominant allele: $w_{11}=1$, $w_{12}=1-s$, $w_{22}=1-s$ // $w_{12}=w_{22}$

Incomplete dominance: $w_{11}=1$, $w_{12}=1-hs$, $w_{22}=1-s$, $0<h<1$

// $w_{11}>w_{12}>w_{22}$

Exercise: derive

2. Balancing selection

Overdominance: $w_{11}=1$, $w_{12}=1-hs$, $w_{22}=1-s$, $h<0$ // $w_{12}>w_{11,22}$

3. Disruptive selection

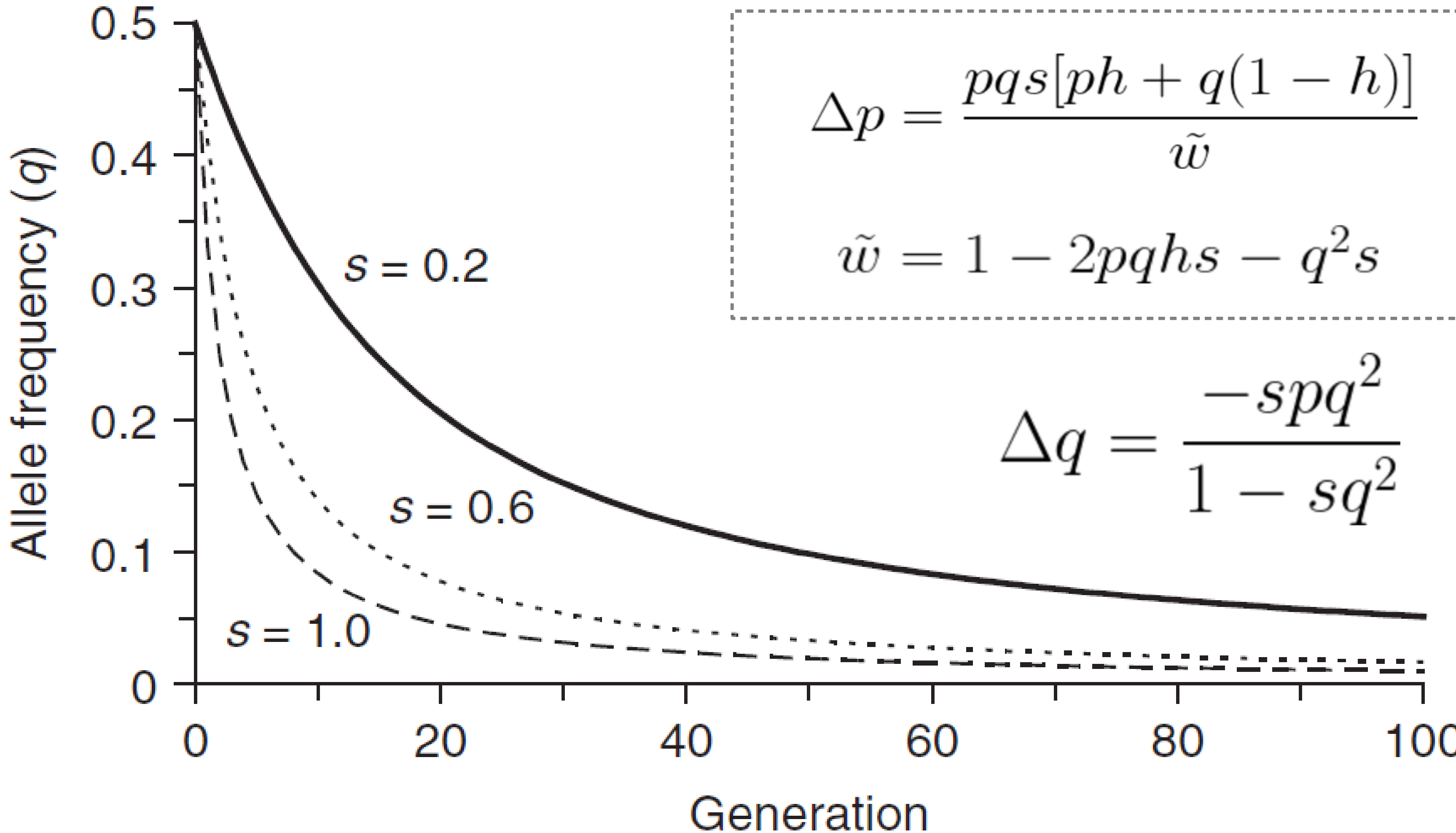
Underdominance: $w_{11}=1$, $w_{12}=1-hs$, $w_{22}=1-s$, $h>1$

Exercise: valid range for *h* ?

Natural selection

Directional selection against a recessive allele:

$w_{11} = w_{12} = 1,\ \ w_{22} = 1 - s$



$$\Delta p = \frac{pqs[ph + q(1-h)]}{\tilde{w}}$$

$$\tilde{w} = 1 - 2pqhs - q^2 s$$

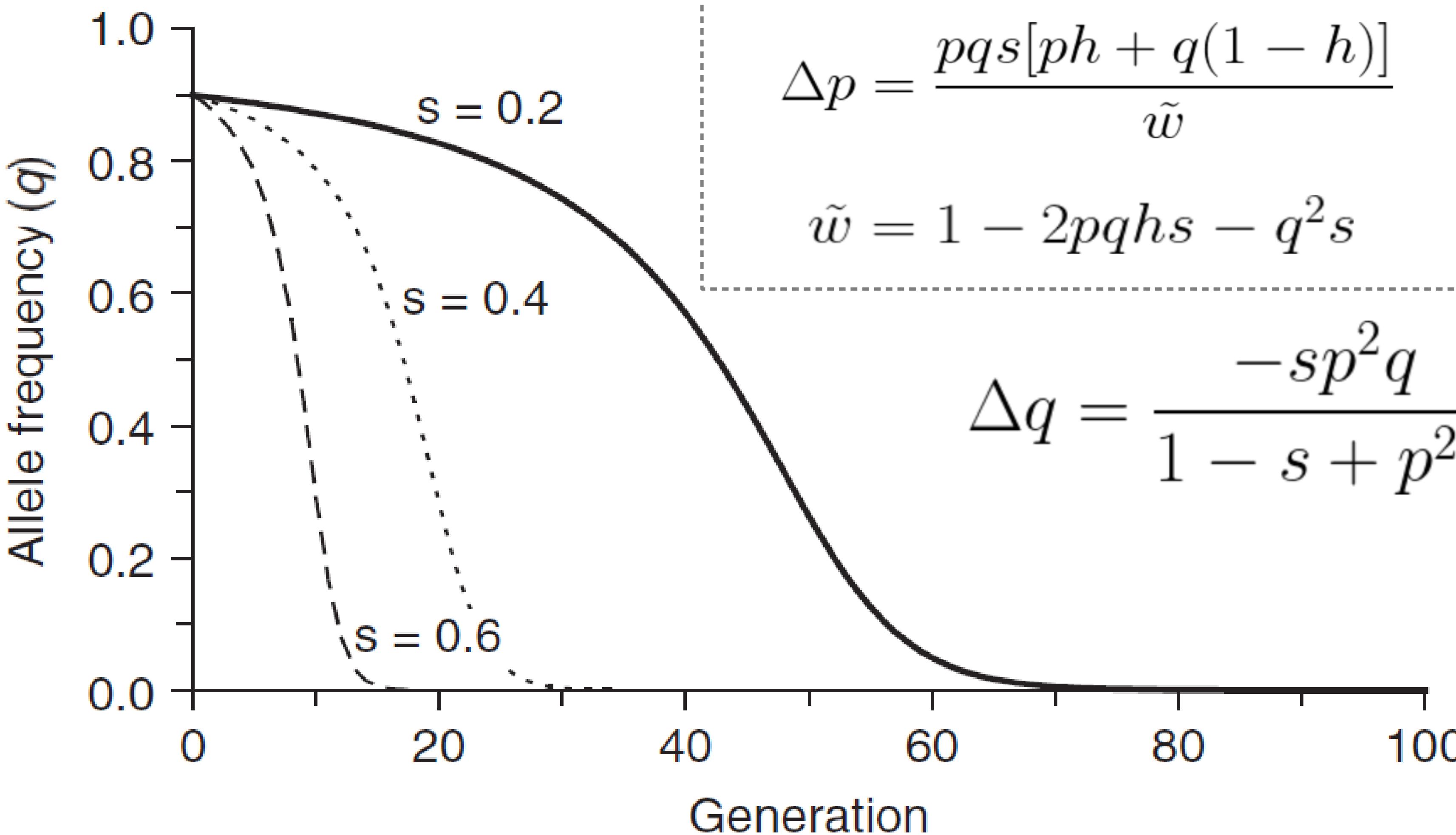
$$\Delta q = \frac{-spq^2}{1 - sq^2}$$

Relethford – *Human Population Genetics*

Natural selection

Directional selection against a dominant allele:

$w_{11} = 1, \;\; w_{12} = w_{22} = 1 - s$



$$\Delta p = \frac{pqs[ph + q(1-h)]}{\tilde{w}}$$

$$\tilde{w} = 1 - 2pqhs - q^2 s$$

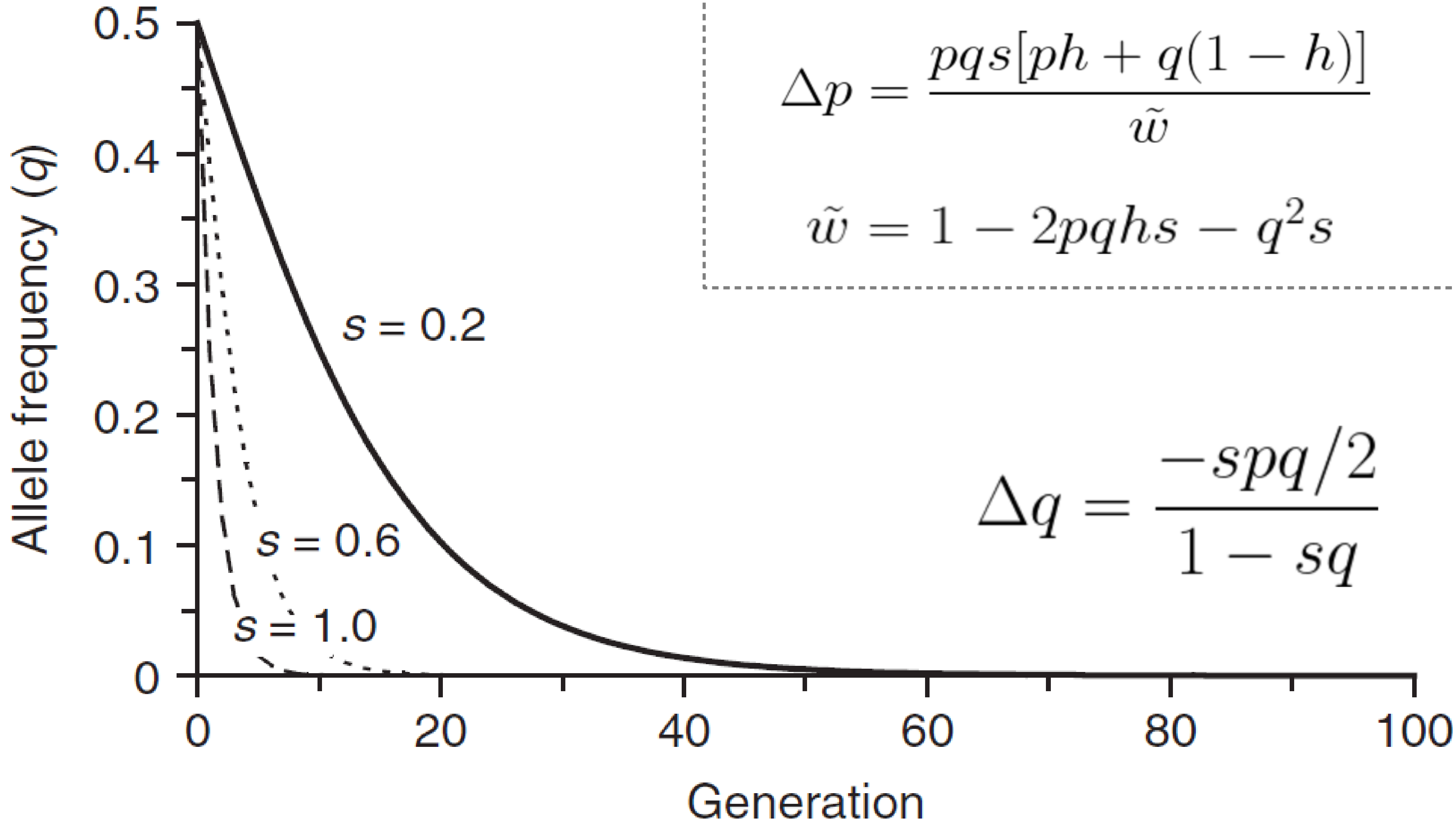
$$\Delta q = \frac{-sp^2 q}{1 - s + p^2 s}$$

Relethford – *Human Population Genetics*

Natural selection

Directional selection against a codominant additive allele:

$w_{11} = 1$, $w_{12} = 1 - s/2$, $w_{22} = 1 - s$ // incomplete dominance



$$\Delta p = \frac{pqs[ph + q(1-h)]}{\tilde{w}}$$

$$\tilde{w} = 1 - 2pqhs - q^2 s$$

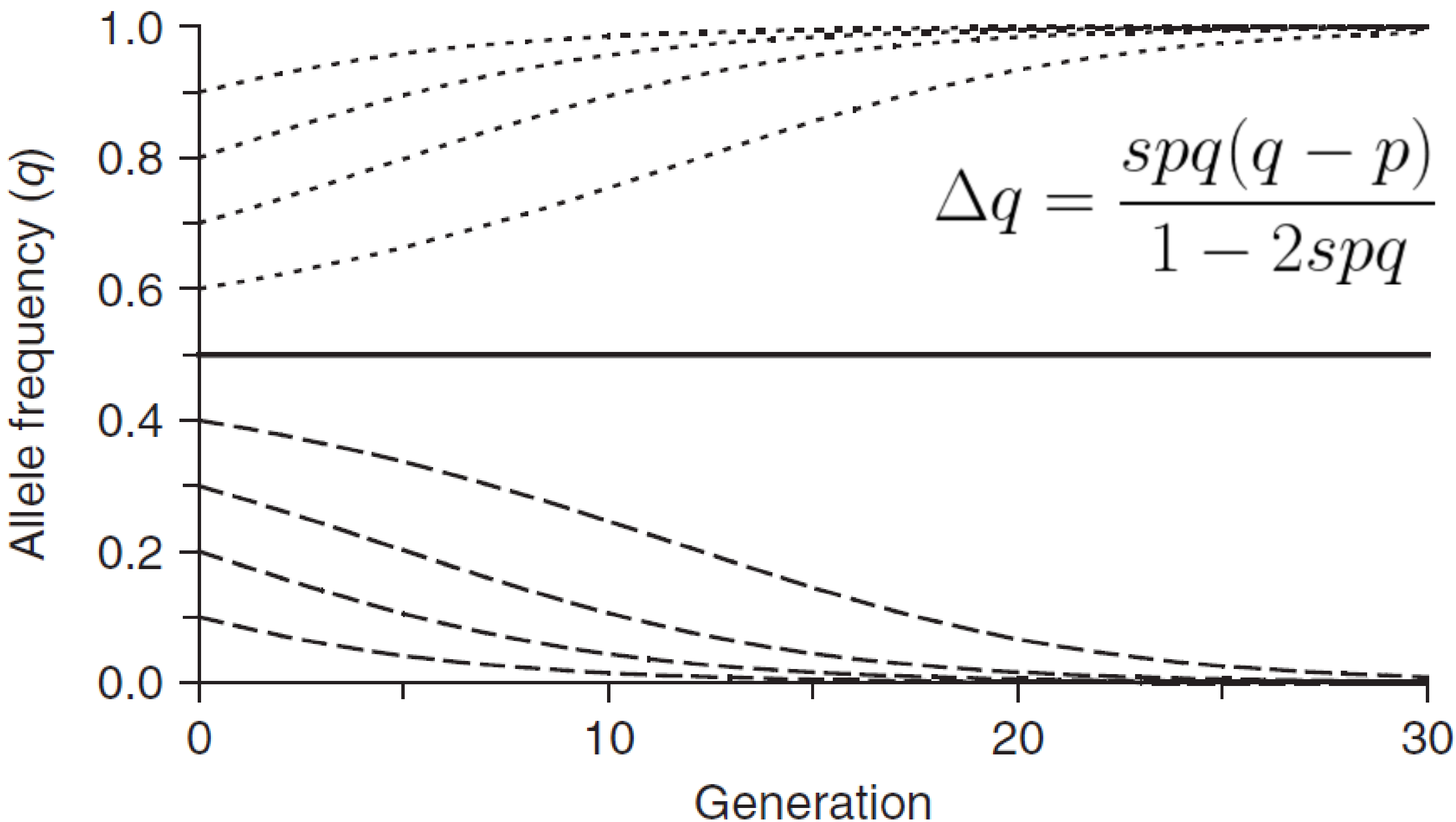
$$\Delta q = \frac{-spq/2}{1 - sq}$$

Relethford – *Human Population Genetics*

Natural selection

Disruptive selection against a heterozygote:

$w_{11} = 1$, $w_{12} = 1 - s$, $w_{22} = 1$ // underdominance



Relethford – *Human Population Genetics*

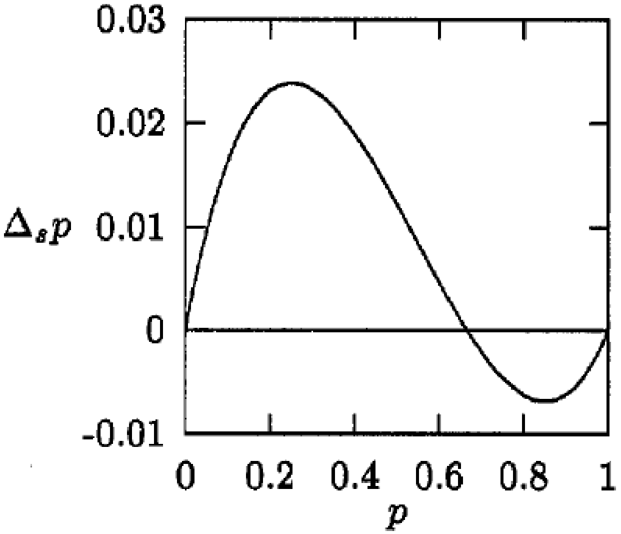
Natural selection

Balancing selection for a heterozygote:

$w_{11}=1$, $w_{12} = 1 - hs$, $w_{22} = 1 - s$, $h < 0$ // overdominance

$$\Delta p = \frac{pqs[ph + q(1-h)]}{\tilde{w}}$$

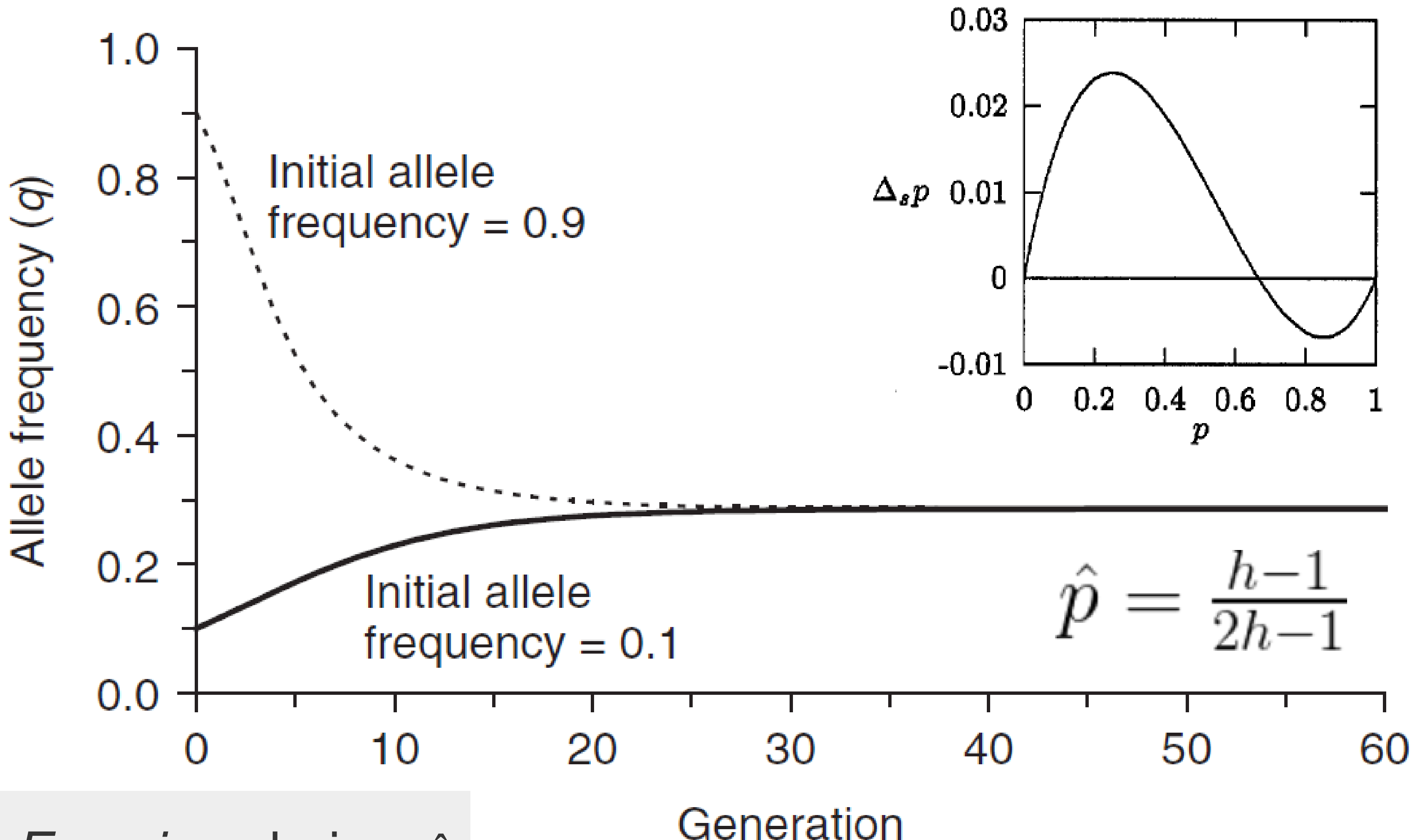
$$\tilde{w} = 1 - 2pqhs - q^2 s$$



Natural selection

Balancing selection for a heterozygote:

$w_{11}=1$, $w_{12} = 1 - hs$, $w_{22}= 1 - s$, $h < 0$ // overdominance



Relethford – *Human Population Genetics*

Exercise: derive p̂

Balancing selection: the case of CF

BOX 3.7 SELECTION IN FAVOR OF HETEROZYGOTES FOR CYSTIC FIBROSIS

For CF, the disease frequency in Denmark is about one in 2000 births.

Phenotypes:	Unaffected		Affected
Genotypes:	*AA*	*Aa*	*aa*
Frequencies:	p^2	$2pq$	$q^2 = 1/2000$

q^2 is 5×10^{-4}; therefore $q = 0.022$ and $p = 1 - q = 0.978$.
$p/q = 0.978/0.022 = 43.72 = s_2/s_1$.
If $s_2 = 1$ (affected homozygotes never reproduce), $s_1 = 0.023$.
The present CF gene frequency will be maintained, even without fresh mutations, if *Aa* heterozygotes have on average 2.3% more surviving children than *AA* homozygotes.

Exercise: express heterozygous advantage h as a function of $\hat{p}$, verify estimate above

Strachan, Read – *Human Molecular Genetics*

Balancing selection: the case of ß-hemoglobin

The most thoroughly studied example of overdominance is the sickle-cell hemoglobin polymorphism found in many human populations in Africa. Hemoglobin, the oxygen-carrying red protein found in red blood cells, is a tetramer composed of two alpha chains and two beta chains. In native West and Central African populations, the S allele of beta hemoglobin reaches a frequency as high as 0.3 in some areas. The more common A allele is found at very high frequency in most other areas of the world. The two alleles differ only in that the S allele has a glutamic acid at its sixth amino position while the A allele has a valine. The glutamic acid causes the hemoglobin to form crystal aggregates under low partial pressures of oxygen, as occur, for example, in the capillaries. As a result, SS homozygotes suffer from sickle-cell anemia, a disease that is often fatal.

The S allele could not have reached a frequency of 0.3 unless AS heterozygotes are more fit than AA homozygotes. This is precisely the case in regions where malaria is endemic, for there the heterozygotes are somewhat resistant to severe forms of malaria. The resistance is due to the sickling phenomena, which makes red blood cells less suitable for *Plasmodium falciparum*. In an old study from 1961, it was shown that the viability of AS relative to AA is 1.176 in regions with malaria. Assuming that the fitness of SS is zero ($s = 1$), $h = -0.176$. Plugging this into Equation 3.4 gives $\hat{p} = 0.87$ or $\hat{q} = 0.13$ for the S allele, which is nestled right in the middle of allele frequencies in regions with endemic malaria.

Gillespie – *Population genetics. A concise guide*

Natural selection

$$\Delta p = \frac{pqs[ph + q(1-h)]}{\tilde{w}} \qquad \tilde{w} = 1 - 2pqhs - q^2 s$$

Sewall Wright: $$\Delta p = \frac{pq}{2\tilde{w}(p)} \frac{\mathrm{d}\tilde{w}(p)}{\mathrm{d}p}$$

"Natural selection always increases the mean fitness and does so at a rate that is proportional to the genetic variation"

Gillespie – *Population genetics. A concise guide*

Mutation-selection balance

- Many new alleles are deleterious and incompletely dominant.
- They enter the population by mutation and are removed by negative selection.

$$A_1(p \approx 1) \xrightarrow{\mu} A_2(q \approx 0)$$

- Balance: the rate of introduction of mutations equals rate of loss due to selection

$$\Delta_{mut}\, p = -\mu p \approx -\mu$$

$$\Delta_{sel}\, p = \frac{pqs[ph + q(1-h)]}{1 - 2pqhs - q^2 s} \approx qhs$$

$$\Delta_{mut}\, p + \Delta_{sel}\, p = 0$$

$$\hat{q} \approx \frac{\mu}{hs}$$

Gillespie – *Population genetics. A concise guide*

Mutation-selection balance

- Many new alleles are deleterious and incompletely dominant.
- They enter the population by mutation and are removed by negative selection.

$$A_1(p \approx 1) \xrightarrow{\mu} A_2(q \approx 0)$$

- Balance: the rate of introduction of mutations equals rate of loss due to selection

$$\Delta_{mut}\, p = -\mu p \approx -\mu$$

$$\Delta_{sel}\, p = \frac{pqs[ph + q(1-h)]}{1 - 2pqhs - q^2 s} \approx qhs$$

$$\Delta_{mut}\, p + \Delta_{sel}\, p = 0$$

$$\hat{q} \approx \frac{\mu}{hs}$$

Large effect → Low frequency

Exercise: derive $\hat{q}$ for a recessive allele

Gillespie – *Population genetics. A concise guide*

Random drift and advantageous allele

Selection in finite population is very weak for *de novo* alleles:
New allele: $\Delta p \approx (1+s)p - p = sp = s/2N << 1/2N$ (drift),
unless $s \approx 1$

$$P_F(p) = \frac{1 - e^{-2Nsp}}{1 - e^{-2Ns}}, \text{ if } h = 1/2$$

$$P_F(1/2N) = \frac{1 - e^{-s}}{1 - e^{-2Ns}}$$

$P_F \approx s$ if $s \approx 0$ and $2Ns >> 1$

Gillespie – *Population genetics. A concise guide*

Random drift and advantageous allele

Selection in finite population is very weak for *de novo* alleles:
New allele: $\Delta p \approx (1+s)p - p = sp = s/2N << 1/2N$ (drift),
unless $s \approx 1$

$$P_F(p) = \frac{1 - e^{-2Nsp}}{1 - e^{-2Ns}}, \text{ if } h = 1/2$$

$$P_F(1/2N) = \frac{1 - e^{-s}}{1 - e^{-2Ns}}$$

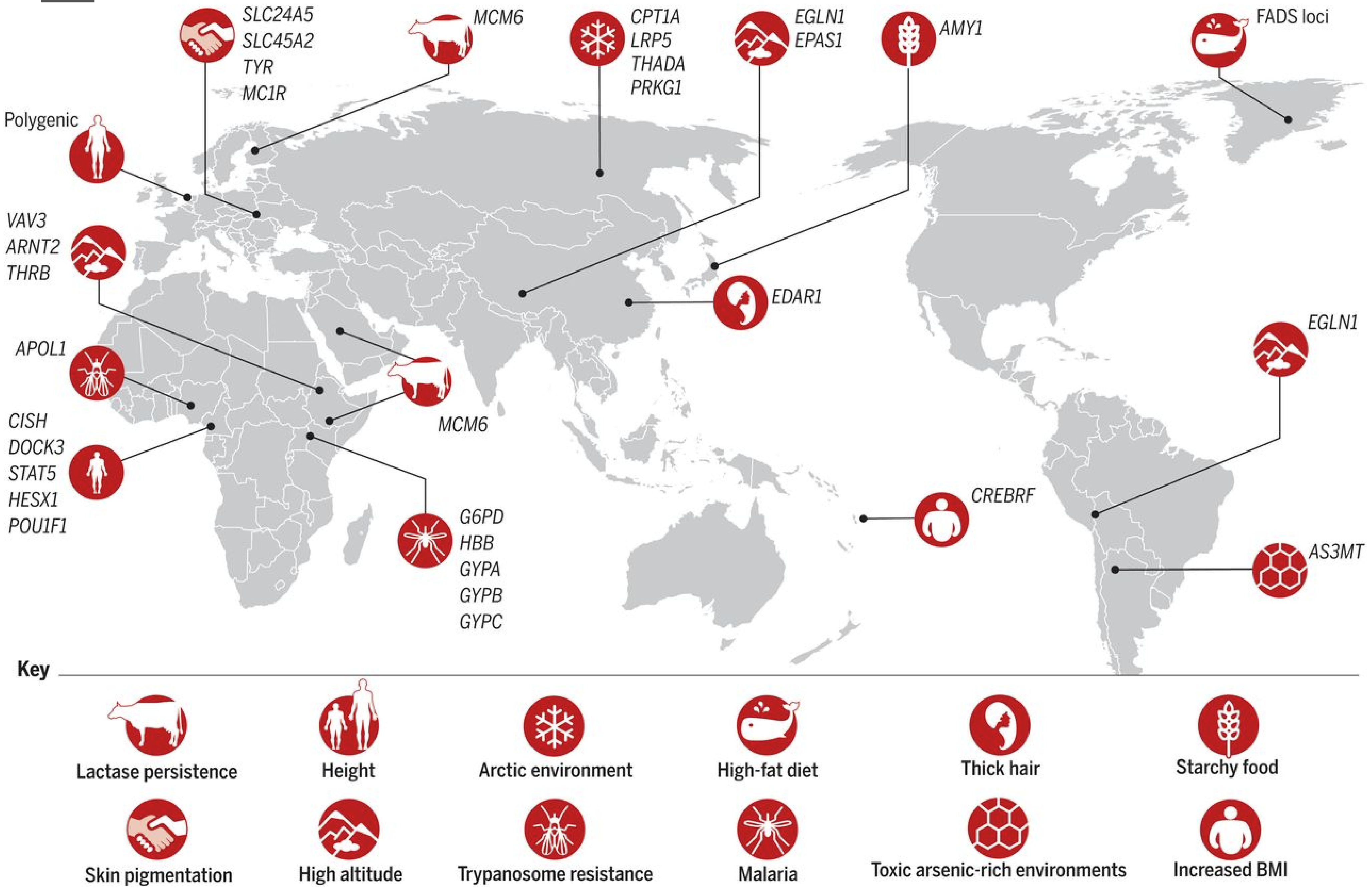
$P_F \approx s$ if $s \approx 0$ and $2Ns >> 1$

- Most advantageous alleles are lost.
- Adaptive evolution is random

Exercise: P_F for s, $2Ns \approx 0$

Gillespie – *Population genetics. A concise guide*

Examples of human local adaptations



Fan (2016) *Science*

Random drift and deleterious allele

Can a deleterious allele fix in a finite population?

$$P_F(q) = 1 - P_F(1-q) = \frac{e^{2Nsq} - 1}{e^{2Ns} - 1}$$

$$P_F(1/2N) \approx \frac{s}{e^{2Ns} - 1}$$

$P_F \approx 0$ **if** $2Ns >> 1$

Gillespie – *Population genetics. A concise guide*

Random drift and deleterious allele

Can a deleterious allele fix in a finite population?

$$P_F(q) = 1 - P_F(1-q) = \frac{e^{2Nsq} - 1}{e^{2Ns} - 1}$$

$$P_F(1/2N) \approx \frac{s}{e^{2Ns} - 1}$$

$P_F \approx 0$ if $2Ns >> 1$

Fixation rate for deleterious alleles:

$$k = 2N\mu P_F(1/2N) = \frac{2N\mu s}{e^{2Ns} - 1}$$

Exercise: P_F for $s \rightarrow 0$

Exercise: k for $s \rightarrow 0$?

Gillespie – *Population genetics. A concise guide*

Mildly deleterious *vs* neutral mutations

Mutations can be placed in three main categories:

- those that are selected (either positively or negatively);
- those that are neutral (i.e. have no effect on fitness) and
- those that have low selection coefficients, and thus behave as neutral in small populations (where the effects of drift dominate) or are selected in large populations, where the deterministic effects of selection prevail

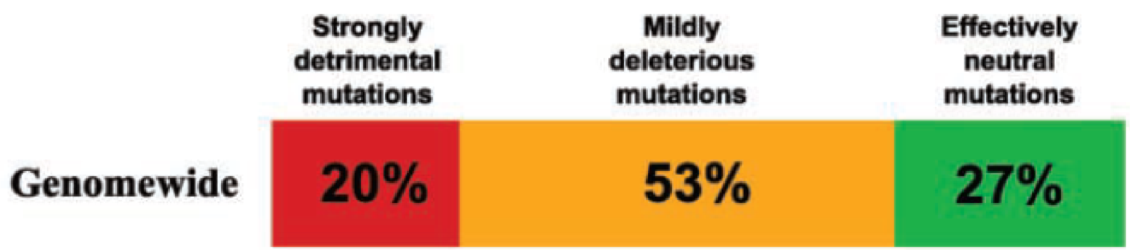
Meyer, Diogo; and, Harris, Eugene E (March 2008) Selection Operating on Protein-coding Genes in the Human Genome. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0020791

Mildly deleterious *vs* neutral mutations

Most Rare Missense Alleles Are Deleterious in Humans: Implications for Complex Disease and Association Studies

Gregory V. Kryukov, Len A. Pennacchio, and Shamil R. Sunyaev

The American Journal of Human Genetics Volume 80 April 2007



0.003 … 0.001

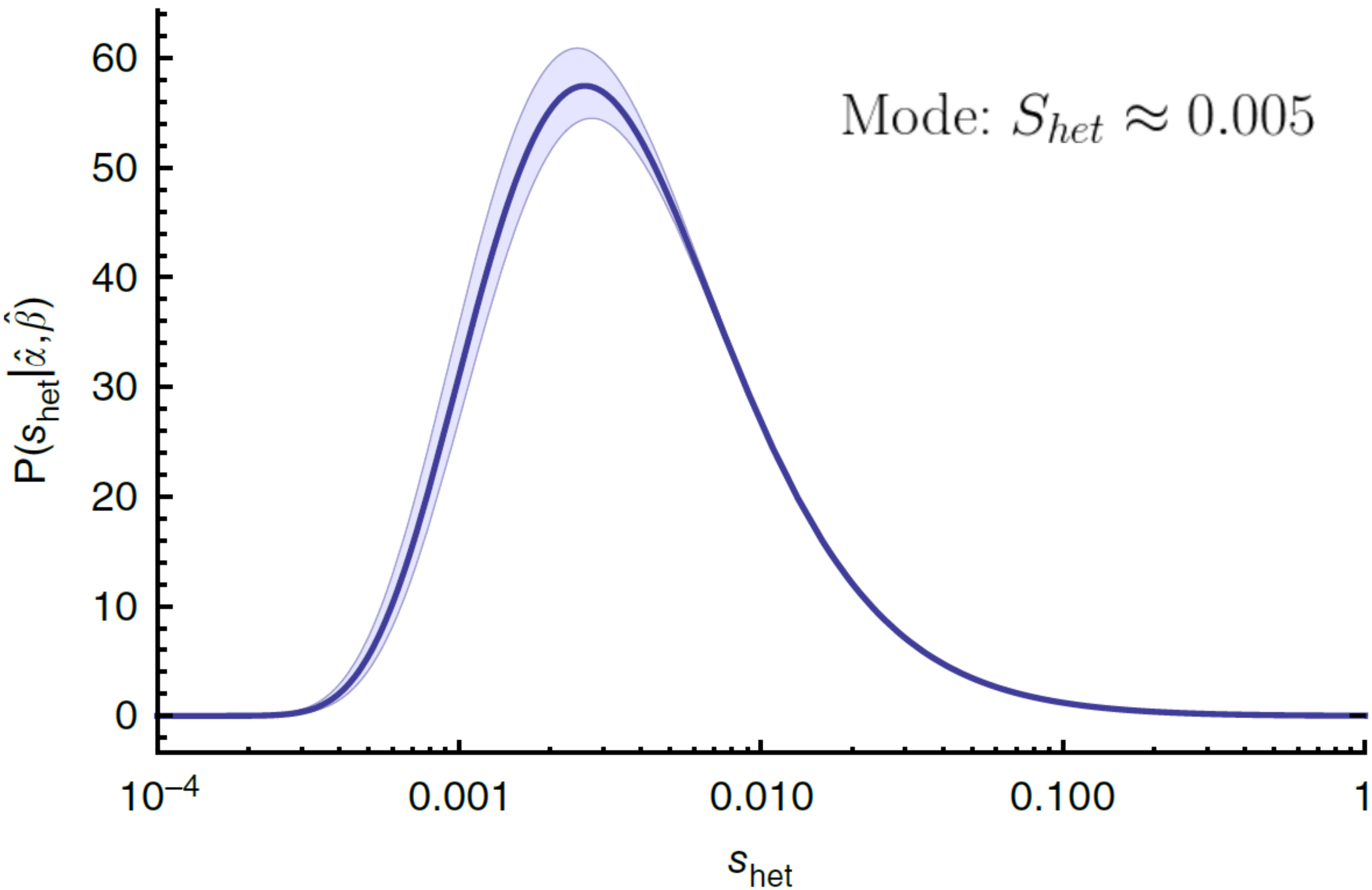
We combined analysis of mutations causing human Mendelian diseases, of human-chimpanzee divergence, and of systematic data on human genetic variation and … estimated that >50% of *de novo* missense mutations in an average human gene and 70% of missense SNPs detected only once among 1,500 chromosomes are mildly deleterious. Such **mildly deleterious mutations are associated with selection coefficients within a surprisingly narrow range of 0.001–0.003**

Kryukov (2007) *Am J Hum Genet*

Estimating the selective effects of heterozygous protein-truncating variants from human exome data

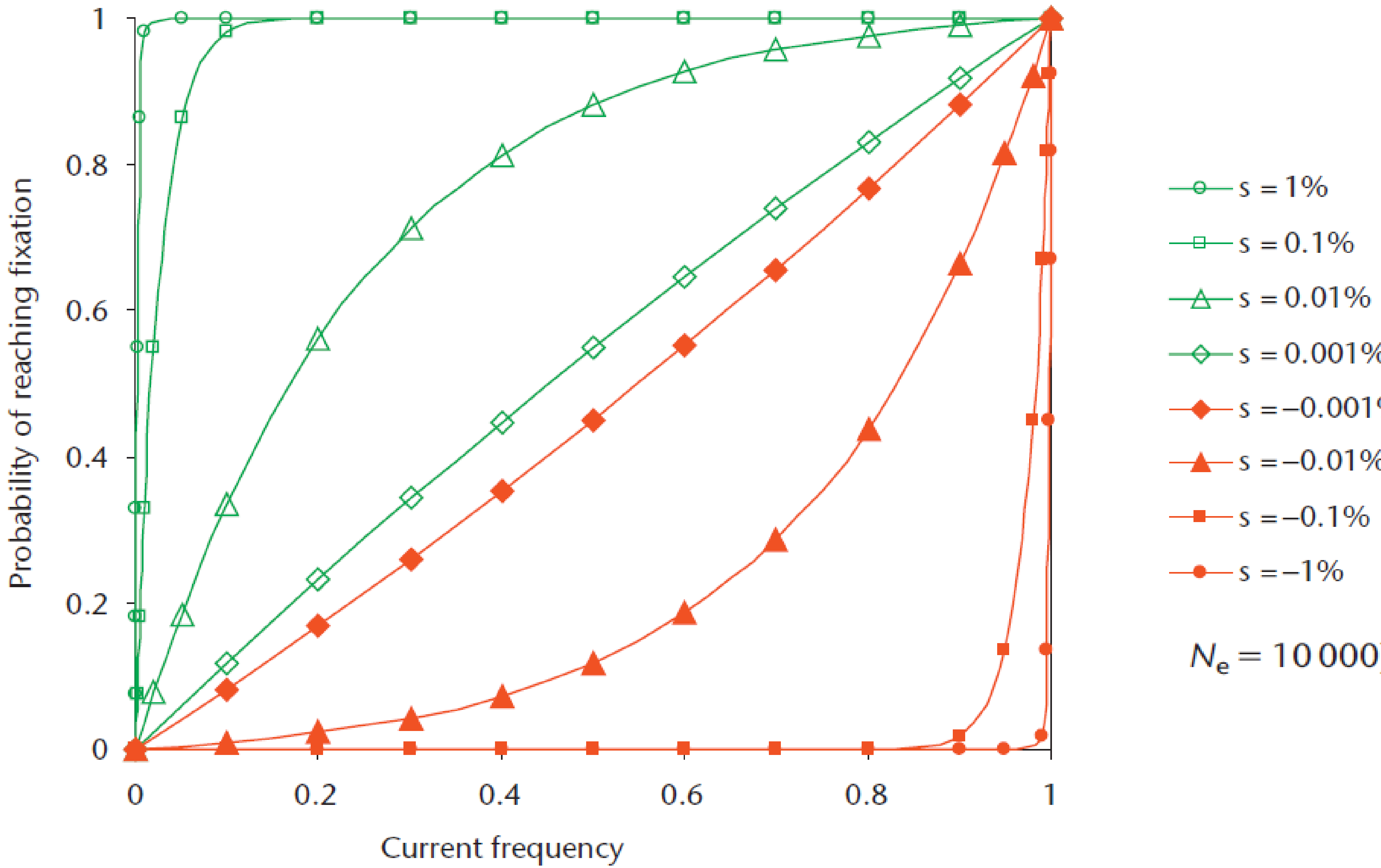
Christopher A Cassa[1,2,9], Donate Weghorn[1,9], Daniel J Balick[1,9], Daniel M Jordan[3,9], David Nusinow[1], Kaitlin E Samocha[4,5], Anne O'Donnell-Luria[4,6], Daniel G MacArthur[2,4], Mark J Daly[2,4], David R Beier[7,8] & Shamil R Sunyaev[1,2]

VOLUME 49 | NUMBER 5 | MAY 2017 NATURE GENETICS



Cassa (2017) *Nat Genet*

Fixation probabilities for all alleles



Thomas, Paul D (July 2008) Single Nucleotide Polymorphisms in Human Disease and Evolution: Phylogenies and Genealogies. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0020763

Summary

What changes allele/genotype frequencies?

- **Mutation:** introduction of new alleles into a population
- **Genetic drift:** sampling variation of transmitted alleles
- **Selection:** different probabilities of survival/reproduction depending on genotypes
- **Gene flow:** movement of alleles due to migration
- **Non-random mating** of individuals in a population

Summary

- Hardy-Weinberg equilibrium describes how zygotes originate from gametes
- Random genetic drift drives alleles to loss or fixation and reduces heterozygosity
- Neutral theory postulates that most inter- and intra-species changes are due to negative selection and random drift
- A coalescent is the lineage of alleles in a sample traced backward in time to their common ancestor allele
- Natural selection changes allele frequencies. It always increases the mean fitness and does so at a rate that is proportional to the genetic variation
- Most new alleles are deleterious and incompletely dominant. They appear by mutation and are subject to negative selection (mutation-selection balance).
- In a finite population, a new advantageous mutation is usually lost because of random drift. On the other hand, a deleterious allele can fix.

Further reading

- Meyer, D., Harris, E. (2008) Selection Operating on Protein-coding Genes in the Human Genome. In: *Encyclopedia of Life Sciences* (ELS). doi:10.1002/9780470015902.a0020791

- Nei, M., Suzuki, Y., and Nozawa, M. (2010). The neutral theory of molecular evolution in the genomic era. *Annu Rev Genomics Hum Genet* 11, 265–289

- Hurst, L.D. (2009). Genetics and the understanding of selection. *Nature Reviews Genetics* 10, 83–93.

- Fan, S., Hansen, M.E.B., Lo, Y., and Tishkoff, S.A. (2016). Going global by adapting local: A review of recent human adaptation. *Science* 354, 54–59.

- John H. Gillespie – Population Genetics. A concise guide

- John H. Relethford – Human population genetics