Introduction to Genetics and Genomics

3. Population and Evolutionary Genetics

lachance.joseph@gmail.com
https://popgen.gatech.edu/

3a. Introductory Concepts

Break

3b. Advanced Concepts

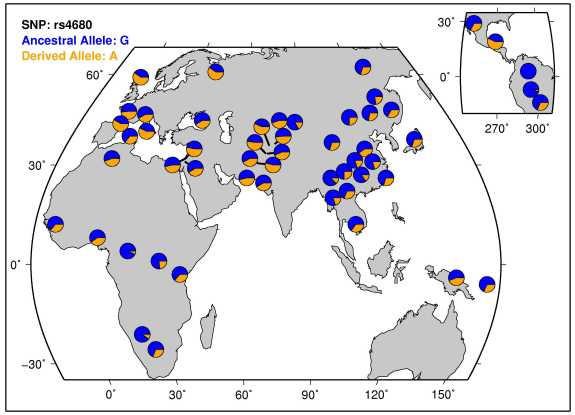
3c. Evolution and Disease

The Double Helix XX-XY
Sculpture by:
Franco Castellucio

Case Study #1

- *COMT* (catechol-O-methyltransferase) and test-taking anxiety

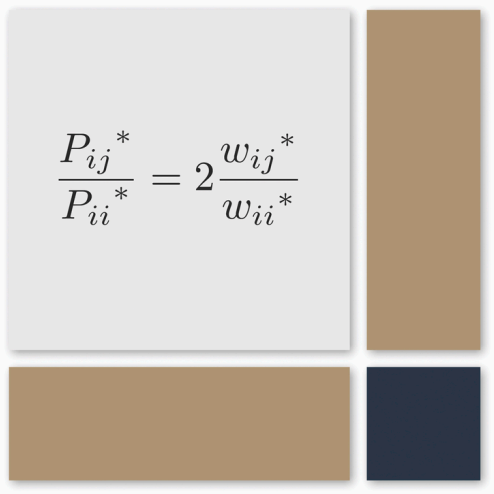
"Some scholars have suggested that we are all Warriors or Worriers. Those with fast-acting dopamine clearers are the Warriors, ready for threatening environments where maximum performance is required. Those with slow-acting dopamine clearers are the Worriers, capable of more complex planning. Over the course of evolution, both Warriors and Worriers were necessary for human tribes to survive. In truth, **because we all get one *COMT* gene from our father and one from our mother, about half of all people inherit one of each gene variation, so they have a mix of the enzymes and are somewhere in between the Warriors and the Worriers. About a quarter of people carry Warrior-only genes, and a quarter of people Worrier-only.**"



Why Can Some Kids Handle Pressure While Others Fall Apart?
Po Bronson and Ashley Merryman, New York Times, February 6, 2013

- What is wrong with this claim?

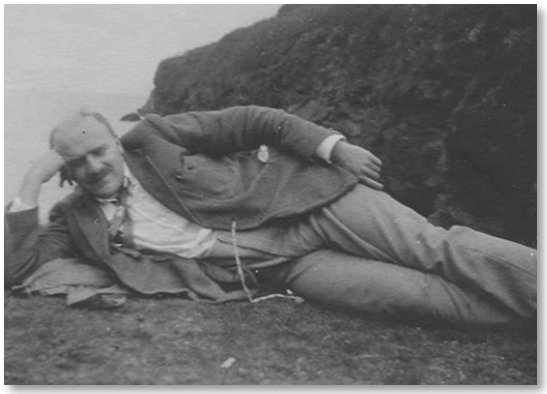
Clearing up some common misconceptions



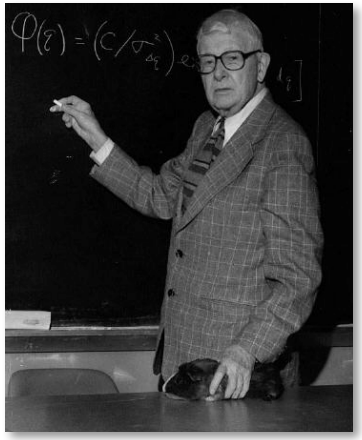
- Dominant alleles need not be the major (most common) allele
- Higher fitness alleles need not be major allele
- Higher fitness alleles are not always dominant (and vice versa)

Giants of population genetics

RA Fisher



JBS Haldane



Sewall Wright

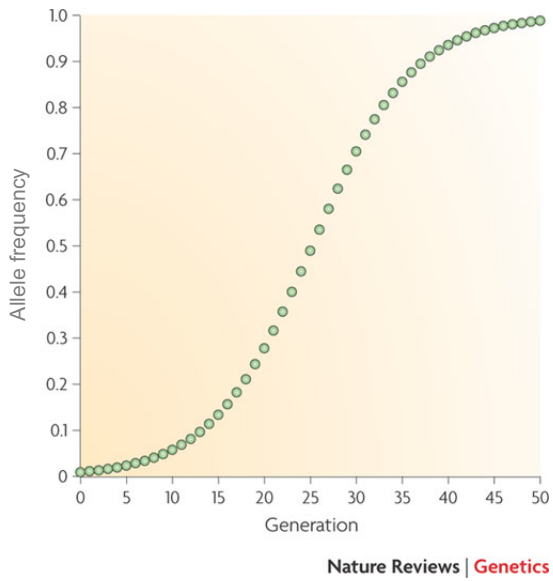
- Used mathematics to describe the genetics of populations
- Integrated evolutionary biology and Mendelian genetics
- *Neo-Darwinism* and the *Modern Synthesis*

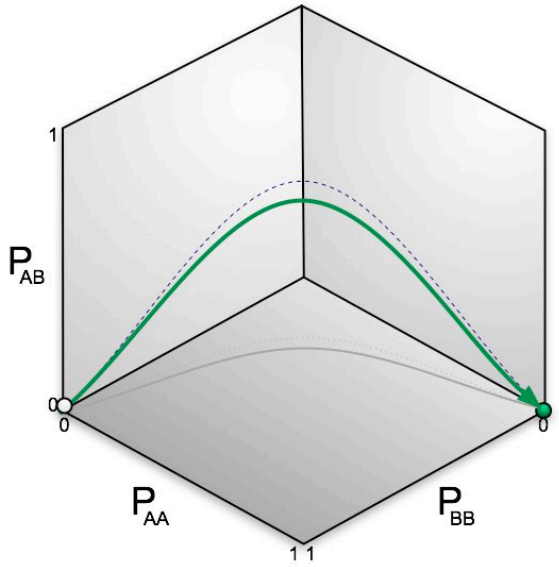
Gene pool

- Definition: the totality of the genes in a population
- Each individual contributes to a pool of gametes
- Contributions to the gene pool are weighted by fitness

- Genotypes next generation found by binomial sampling (w/ replacement)

Allele and genotype frequency space





- Allele and genotype frequencies sum to one
- A diploid population can be represented by a point in genotype frequency space
- Allele and genotype frequencies can be tracked over time
- When alleles are rare most copies are found in a heterozygous state

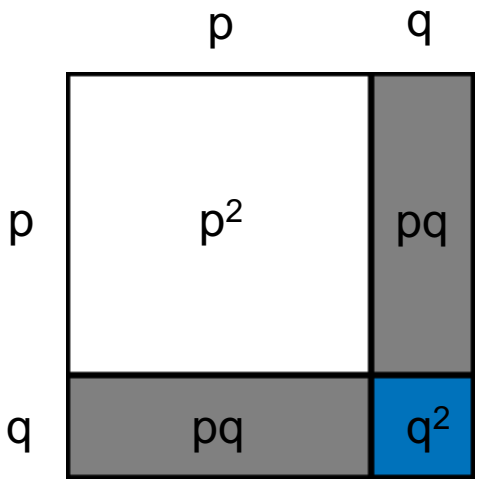
Hardy-Weinberg principle

- $p^2 + 2pq + q^2 = 1$

- p: frequency of *A* allele
- q: frequency of *a* allele

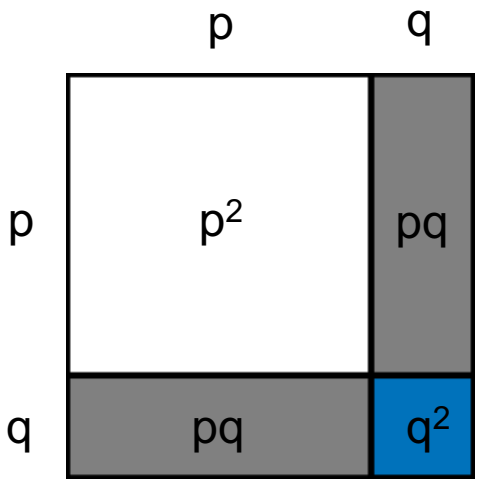
- p^2: frequency of *AA* homozygotes
- 2pq: frequency of *Aa* heterozygotes
- q^2: frequency of *aa* homozygotes

- Modified Punnett Square



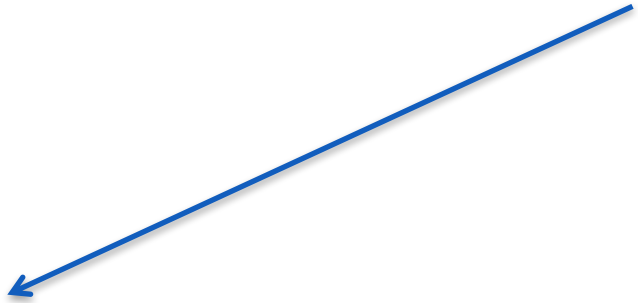
Hardy-Weinberg principle

- Allele frequencies used to calculate genotype frequencies
- Equilibrium reached in a single generation
  (so long as assumptions hold)
- Assumptions
  - Infinite population size
  - No selection
  - No mutation
  - No migraton
  - Random mating

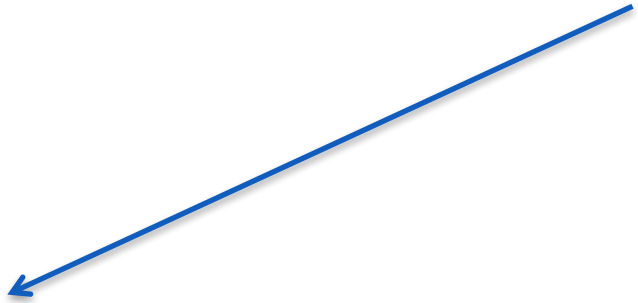


Hardy-Weinberg example

- Initial genotype frequencies: $P_{AA}=0.8$, $P_{AB}=0$, $P_{BB}=0.2$ Initial allele frequencies: p=0.8, q=0.2



- After one generation: $P_{AA}=0.64$, $P_{AB}=0.32$, $P_{BB}=0.04$ Allele frequencies: p=0.8, q=0.2



- After another generation: $P_{AA}=0.64$, $P_{AB}=0.32$, $P_{BB}=0.04$ Allele frequencies: p=0.8, q=0.2

Testing for departures from HW proportions

- Chi-square test with 1 degree of freedom
- $\chi^2 > 3.84$ indicates statistical significance (p-value < 0.05)
- Example:

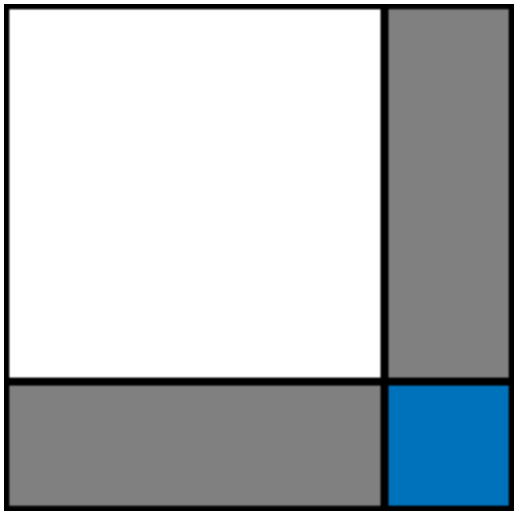
Genotype	Observed	Expected	χ^2
AA	145	131.31	1.426
AB	68	95.37	7.854
BB	31	17.32	10.815
Total	**244**	**244**	**20.095**

$$p = \frac{145 + 68/2}{145 + 68 + 31} = 0.7336$$

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

Major processes of population genetics

- Genetic drift
- Natural selection
- Mutation
- Migration (gene-flow)
- Mating structure



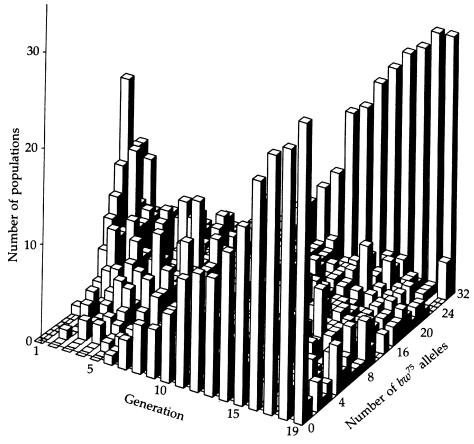
- These processes are mechanisms of evolution

- Additional factors:
  - Recombination (and linkage), gene conversion, ploidy, dominance, epistasis, developmental constraints

Random genetic drift

- In small populations there is a decay of heterozygosity:

$$H_t = H_0 \left(1 - \frac{1}{2N}\right)^t$$



Buri's 1956 experiment:
107 replicate population cages with segregating alleles at the *brown* locus (*D. melanogaster)*

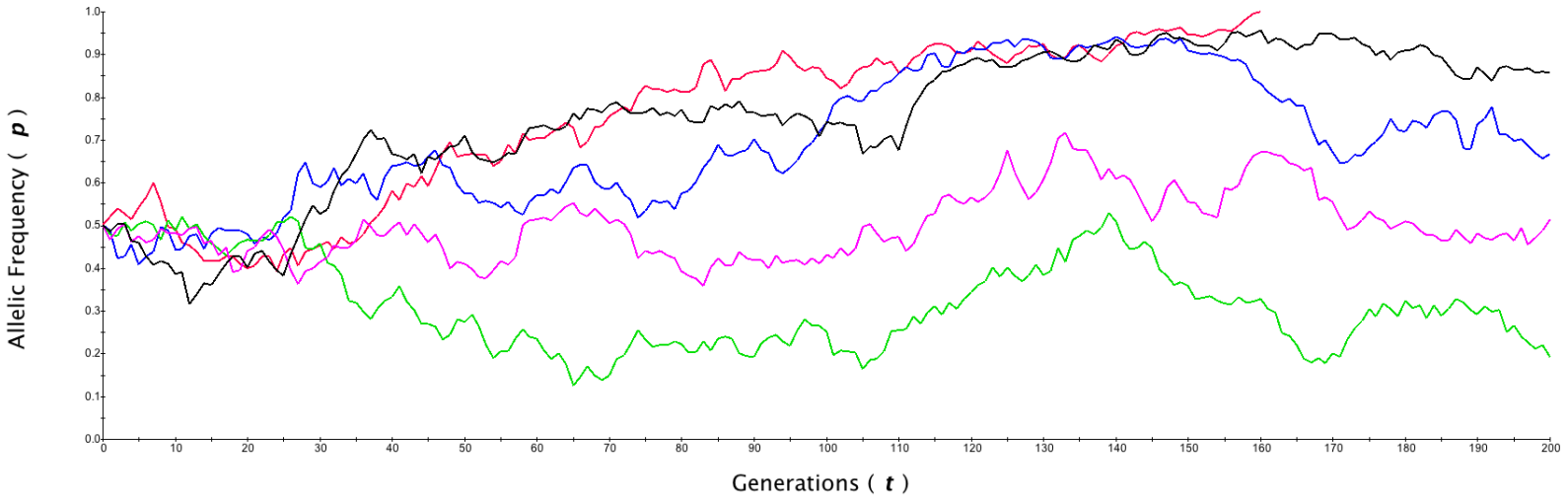
Figure from Hartl and Clark (1989)
Principles of Population Genetics
Sinauer, Sunderland, MA.

- The net effect of drift is to reduce the amount of genetic variation segregating in a population

Random genetic drift

- Random walks through allele frequency space
- Genetic drift is stronger in small populations
- Can lead to differentiation between isolated populations
- Relatively slow process (relative to selection)
  - Mean time for new mutation to reach fixation = $4N$ generations

Simulations of genetic drift



Populus
Simulations of Population Biology

Genetic drift and effective population size

- **Effective population size** (N_e): The idealized (haploid) population size that behaves the same way with respect to drift as a population of size N

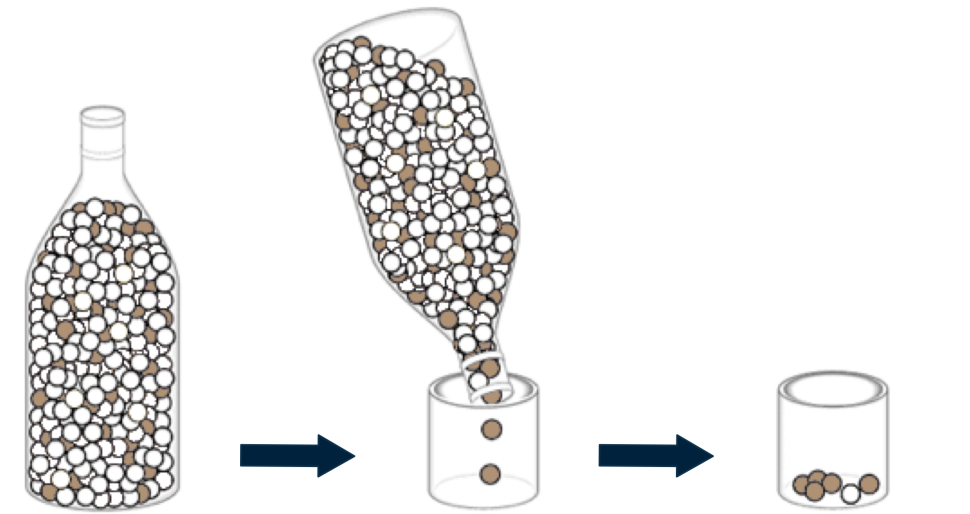
- N_e due to unequal sex ratio $$N_e = \frac{4N_m N_f}{N_m + N_f}$$

- N_e due to variance in reproductive success $$N_e = \frac{4N - 2}{V_k + 2}$$

- N_e due to changing population size $$N_e = \frac{t}{\sum_{i=1}^{t} \frac{1}{N_i}}$$

- Caveat: N_e is a descriptive term, and two populations with the same effective population size can have quite different dynamics

Population bottlenecks and founder effects



- **Population bottleneck**: A sharp reduction in the size of a population
- **Founder effect**: Bottleneck caused by the founding of a new population
- Random chance determines whether an allele increases or decreases in frequency

Genetic drift example

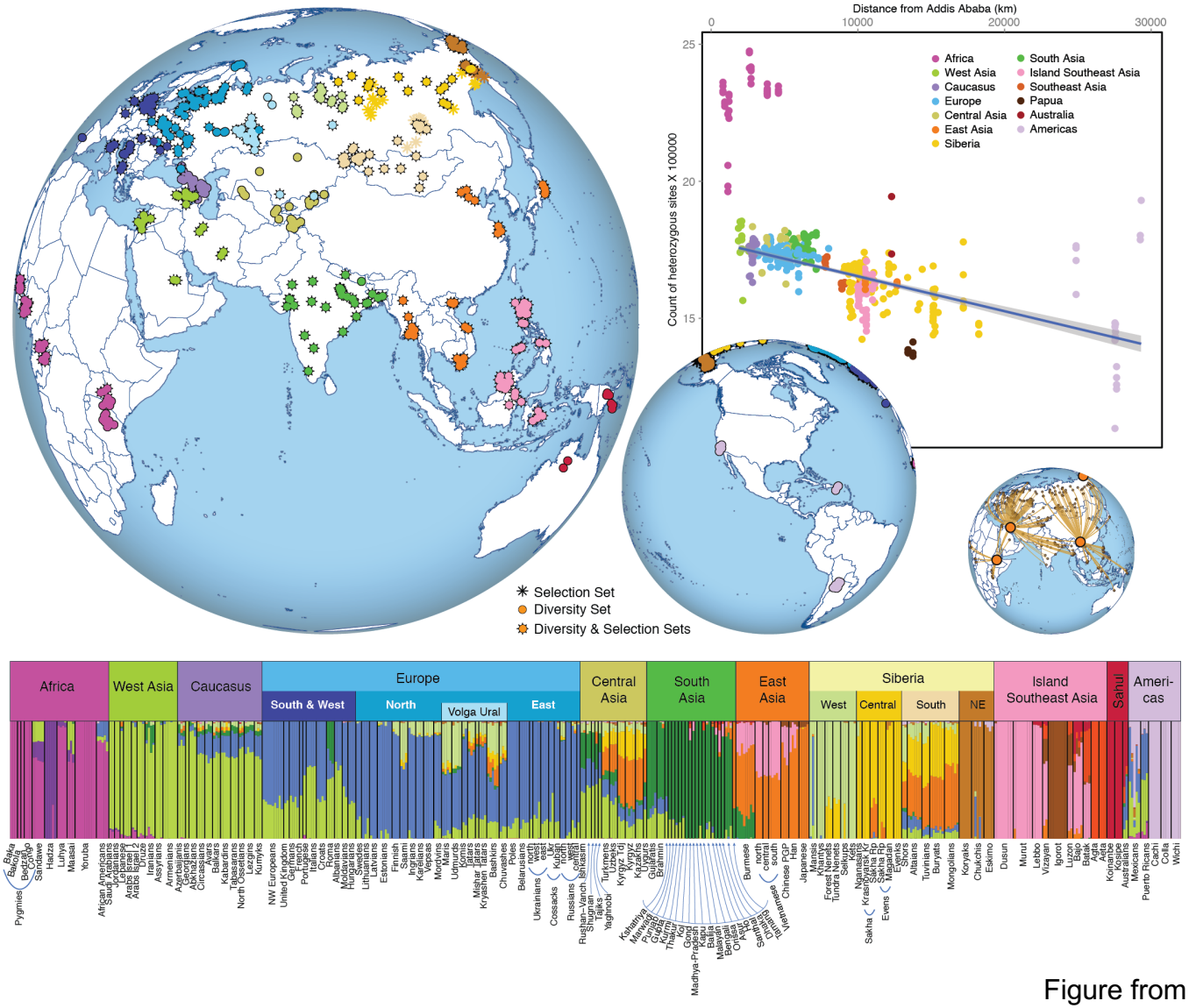
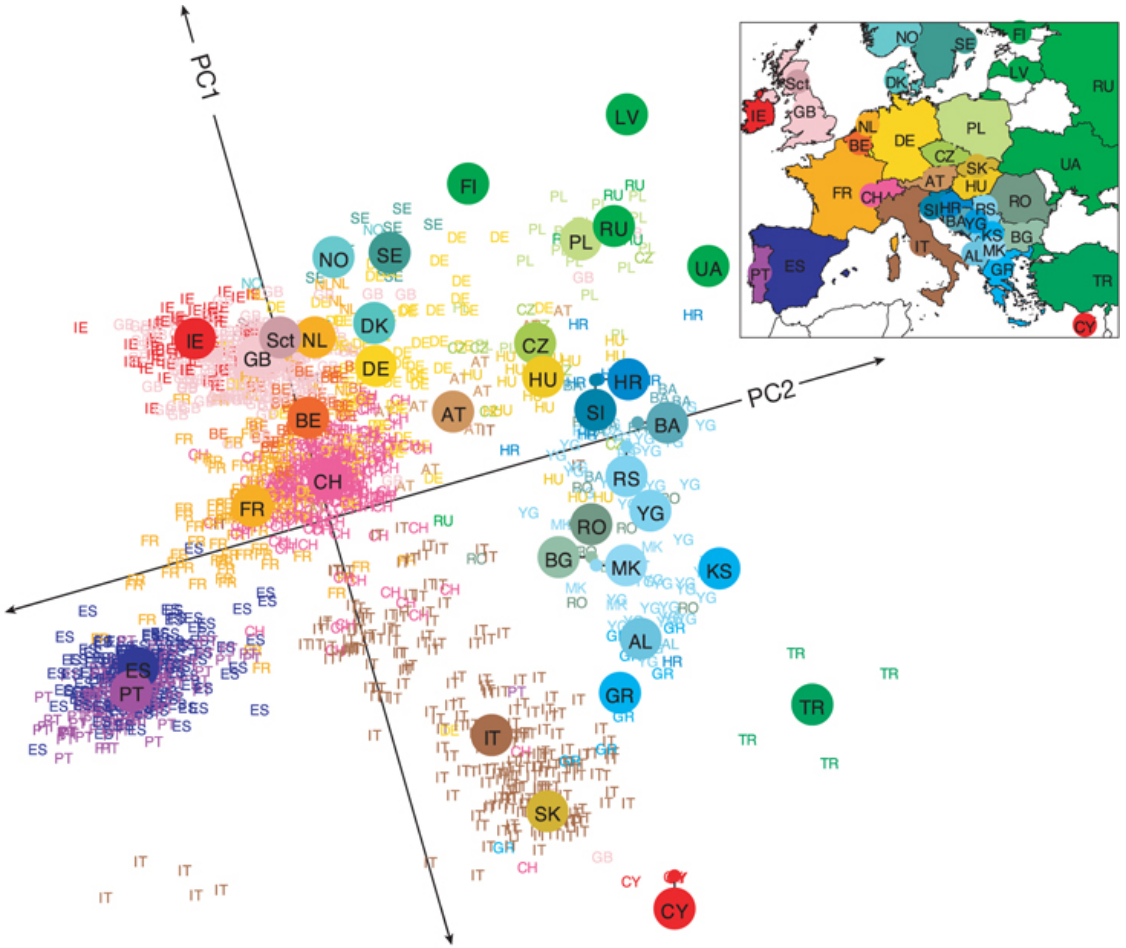


Figure from Pagani et al. 2016 (*Nature*)

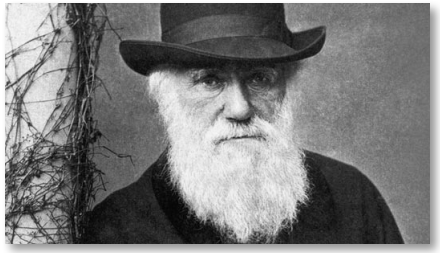
Genes mirror geography in Europe



Novembre et al. (2008, *Nature*)

Natural selection

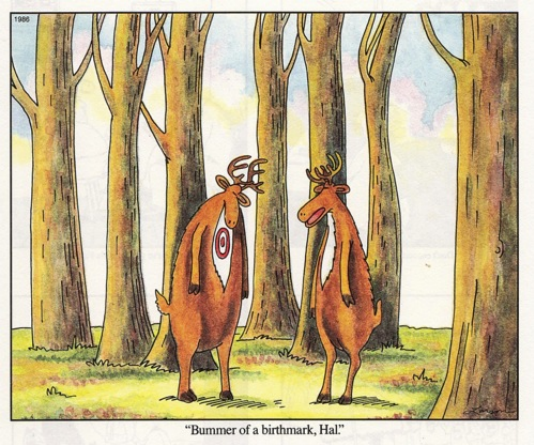
- **Natural selection**: The differential survival and/or reproduction of different genotypes due to unequal fitnesses
- Natural selection is not the same thing as evolution
- Selection coefficient (s)
  - s = 0.01 indicates a 1% fitness advantage
  - $|s|$ tends to be close to 0



- Operates on short time scales (~1/s generations)
- The outcome of natural selection depends on fitnesses and initial frequencies
- Probability of fixation: ~2s
  - Most advantageous mutations are not fixed

Natural selection: fitness

- Genotype-specific fitness is often represented by the parameter $\boldsymbol{w}$
- **Relative** fitness determines allele frequency changes over time
- Absolute fitness determines population growth rates

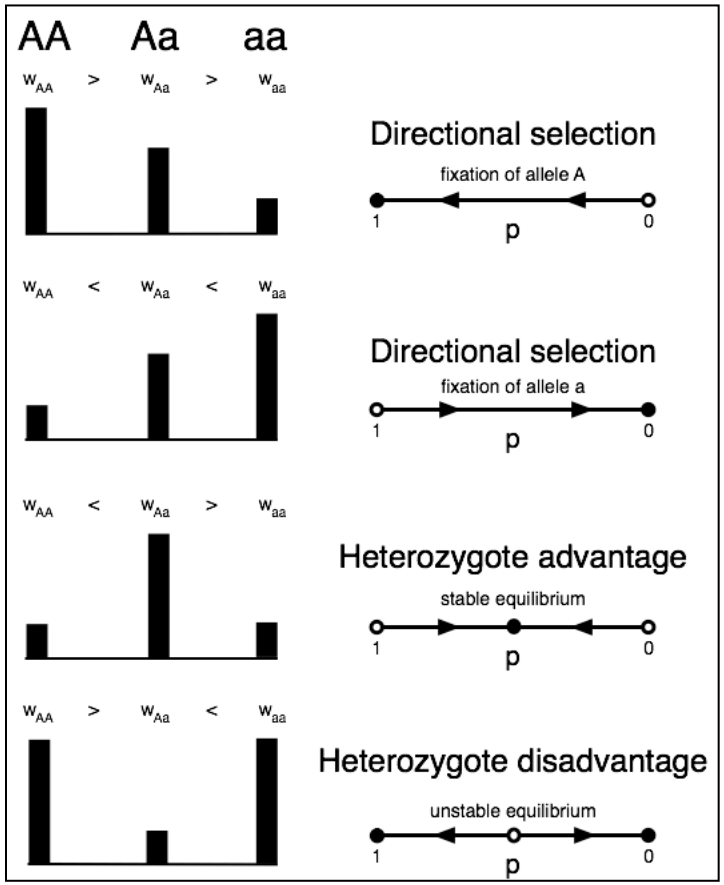


The Far Side
(Gary Larson)

- **Neutral** genotypes have a fitness of 1
- Advantageous genotypes have a fitness greater than 1
- Deleterious genotypes have a fitness less than 1

Types of natural selection

- Directional selection
- Overdominant selection
  - Heterozygte advantage
- Underdominant selection
  - Heterozygote disadvantage
- Frequency dependent selection



Mathematics of natural selection

- Haploid scenario
- Allele frequency next generation can be found by weighting alleles by how much they contribute to the gene pool (fitness)

$$p' = \frac{pw_A}{pw_A + qw_B}$$

- Allele frequency at an arbitrary point in time:

$$p_t = \frac{p_0 {w_A}^t}{p_0 {w_A}^t + q_0 {w_B}^t}$$

Mathematics of natural selection

- Diploid scenario with fitness dominance
- Frequencies next generation can be found by weighting contributions to the gene pool

$$P_{AA}{}' = \frac{p^2 w_{AA}}{p^2 w_{AA} + 2pq w_{AB} + q^2 w_{BB}}$$

$$P_{AB}{}' = \frac{2pq w_{AB}}{p^2 w_{AA} + 2pq w_{AB} + q^2 w_{BB}}$$

$$P_{BB}{}' = \frac{q^2 w_{BB}}{p^2 w_{AA} + 2pq w_{AB} + q^2 w_{BB}}$$

$$p' = \frac{p^2 w_{AA} + pq w_{AB}}{p^2 w_{AA} + 2pq w_{AB} + q^2 w_{BB}}$$

Mathematics of natural selection

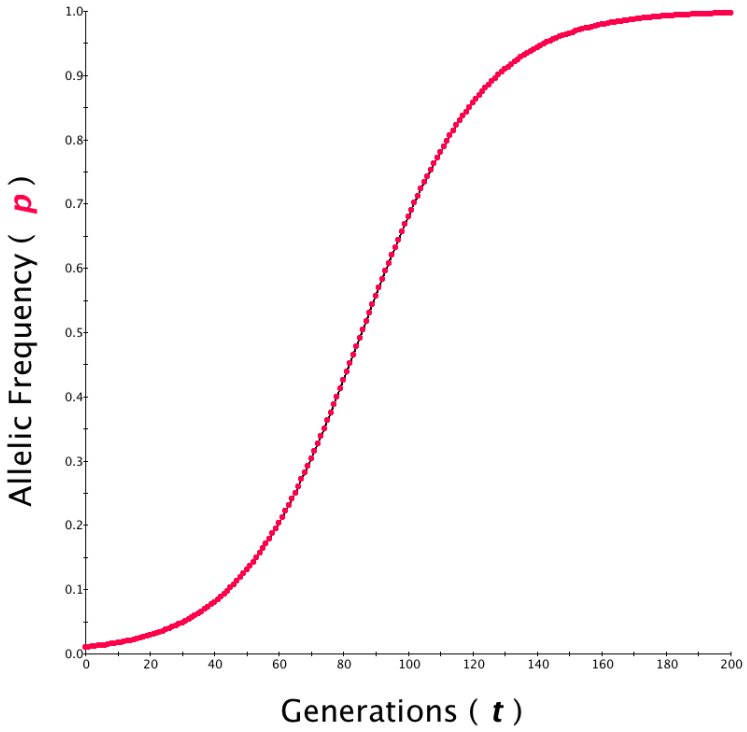
$$\Delta p = p' - p = \frac{pw_A}{\bar{w}} - p$$

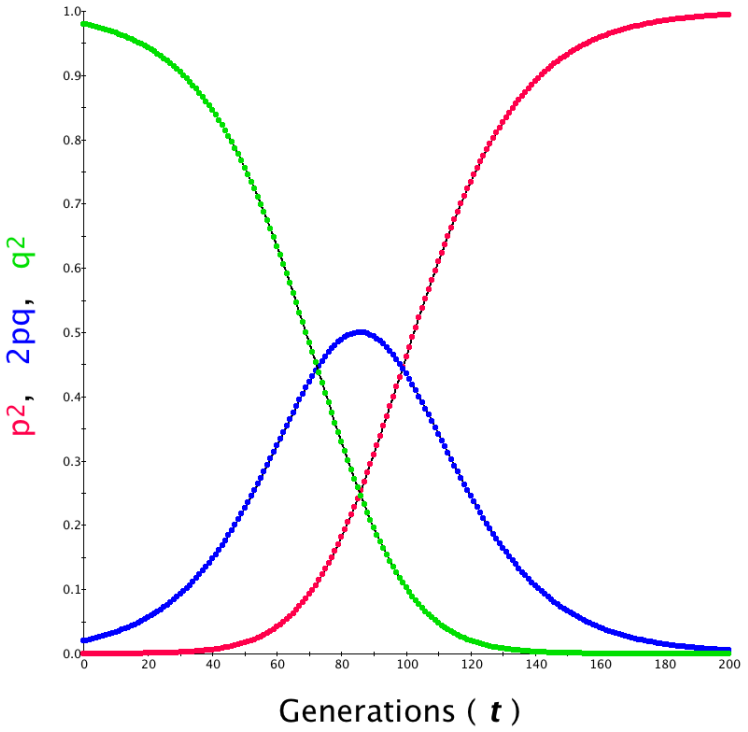
- General equation for single generation allele frequency change:

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

- Response to selection hinges on:
  - Allele frequencies
  - The relative fitness of an allele
  - Mean fitness of a population

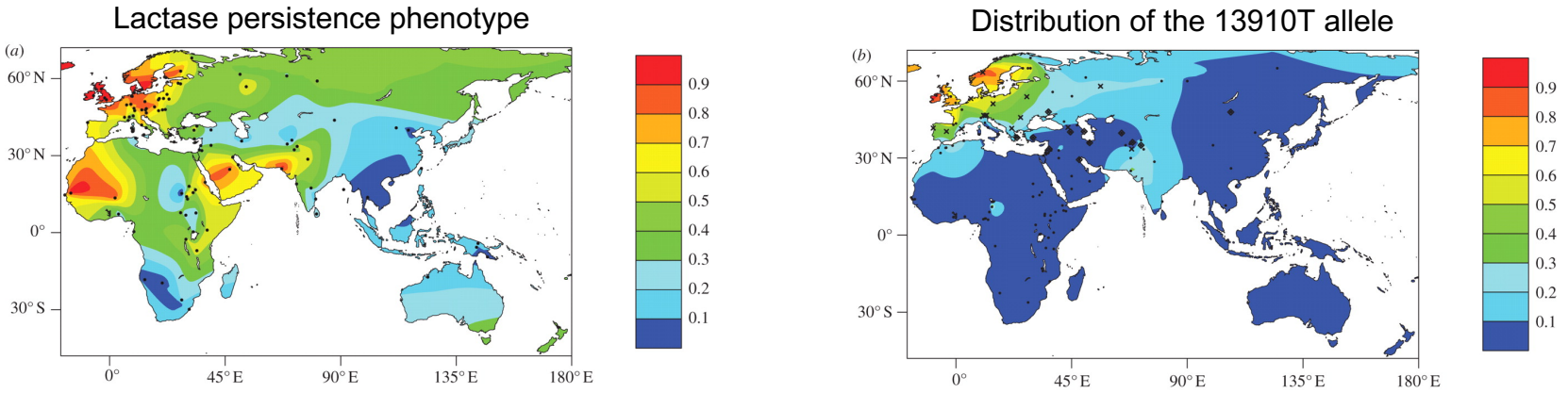
Simulations of directional selection





Populus
Simulations of Population Biology

Natural selection example



- Figures from Gerbault et al. 2011 (*Phil Trans Roy Soc B*)
- Lactase persistence alleles show evidence of positive selection
- Different causal alleles in Africa (convergent phenotypic evolution)

Mutation

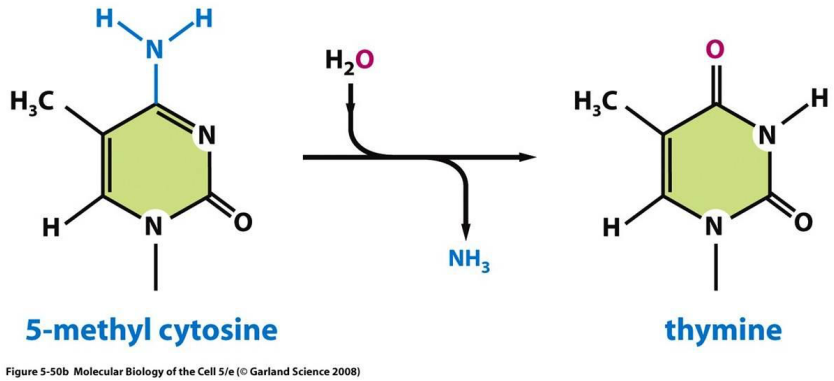


Figure 5-50b Molecular Biology of the Cell 5/e (© Garland Science 2008)

- A “Goldilocks” scenario: Too low a mutation rate and populations lack genetic diversity. Too high of a mutation rate and natural selection is unable to purge deleterious mutations.

- Evolutionary genetics tends to focus on *germline* mutations, as opposed to somatic mutations (most germline mutations occur during DNA replication)

- Mutation rates vary across the genome (much more common at CpG sites)

Human germline mutation rates

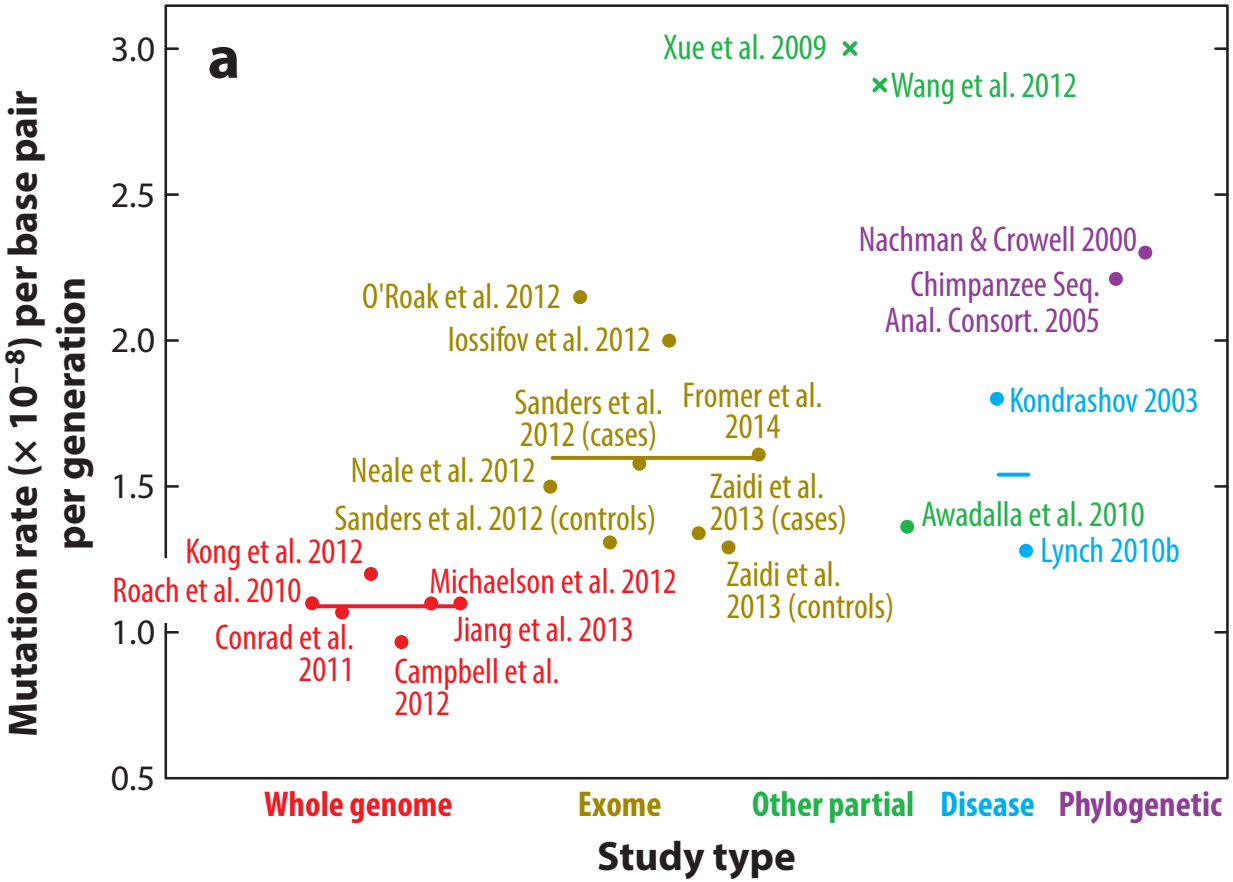
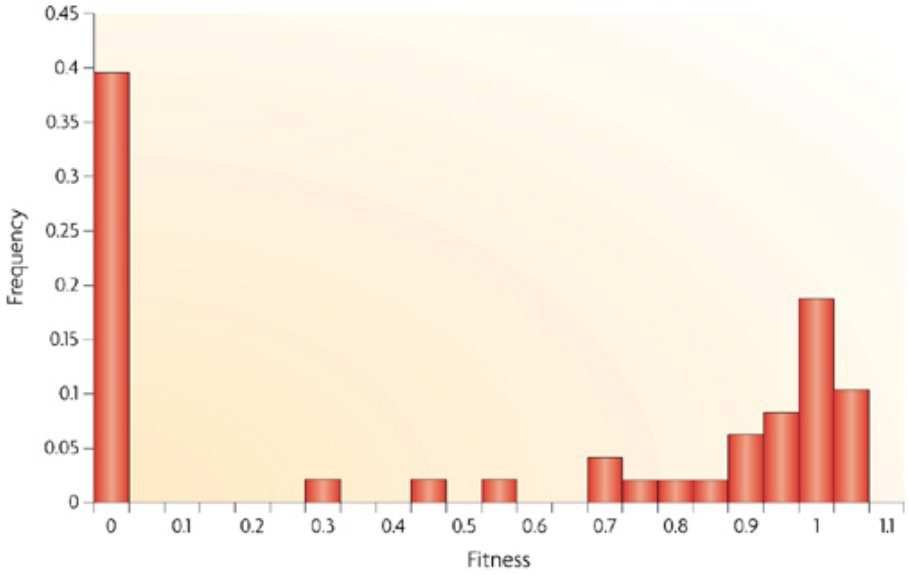


Figure from Ségurel et al 2015 (*Annual Review of Genomics and Human Genetics*)

Distribution of fitness effects (DFE)



Vesicular stomatitis virus data

Nature Reviews | Genetics

Marvel

- Most mutations are deleterious or neutral: they do not increase fitness

- Alas, most mutations don't result in *hopeful monsters* (a la Goldschmidt)

Mutation and molecular clocks

- The rate of neutral substitution depends on mutation rate alone (surprisingly it is independent of population size)

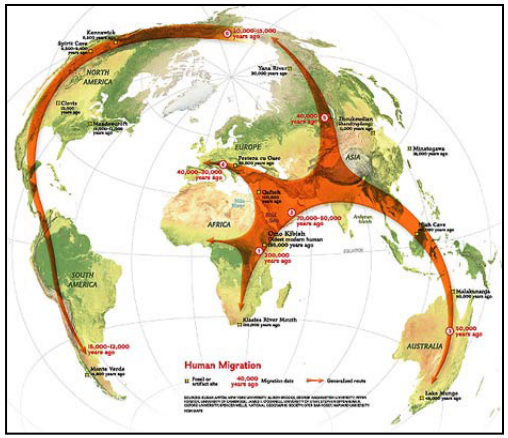
$$2N\mu \times \frac{1}{2N} = \mu \text{ substitutions per generation}$$

- Derivation:
  - A population of N diploid alleles
  - $2N\mu$ mutations per generation
  - Each of the 2N alleles present as an equal chance to be fixed
  - Rate of fixation=(population-level rate of mutation) × (probability of fixation)
  - Assumes that mutation rates are low ($4N\mu >> 1$)

Migration

- When population geneticists refer to migration they mean ***gene* flow**
- The parameter *m* equals the proportion of alleles in a population that are from immigrants
- Gene flow homogenizes populations
- Local differentiation occurs when there is

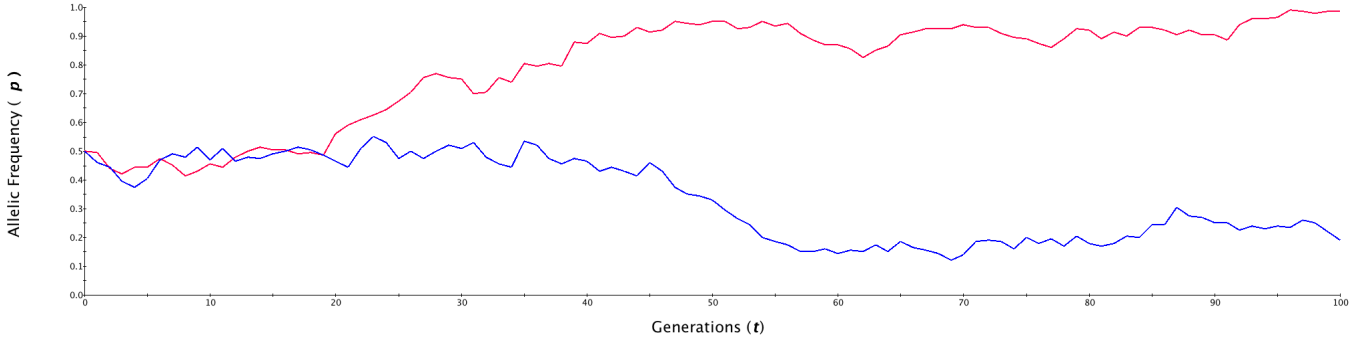
  < 1 migrant per generation (i.e. *Nm* < 1)



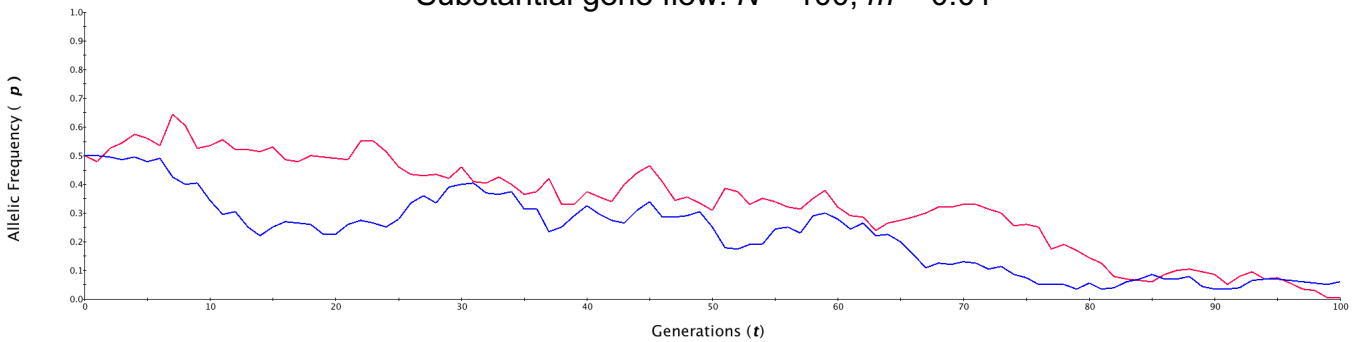
National Geographic

Simulations of migration (and genetic drift)

No gene flow*: N*=100, *m* =0



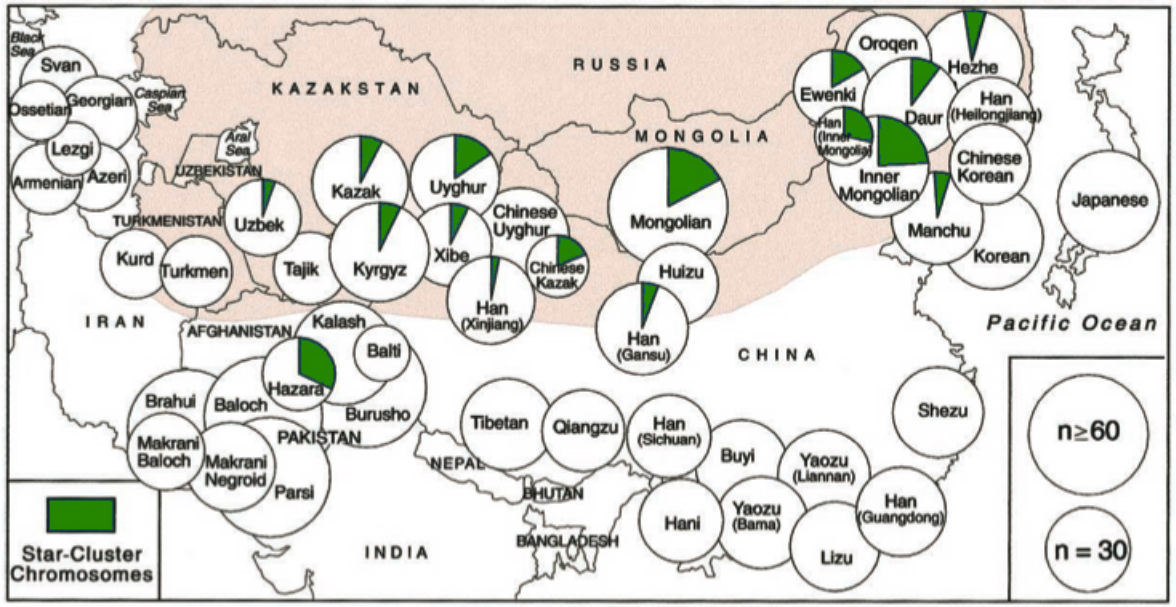
Substantial gene flow: *N* = 100, *m* = 0.01



Populus
Simulations of Population Biology

Migration example

- Geographic proximity results in genetic similarity



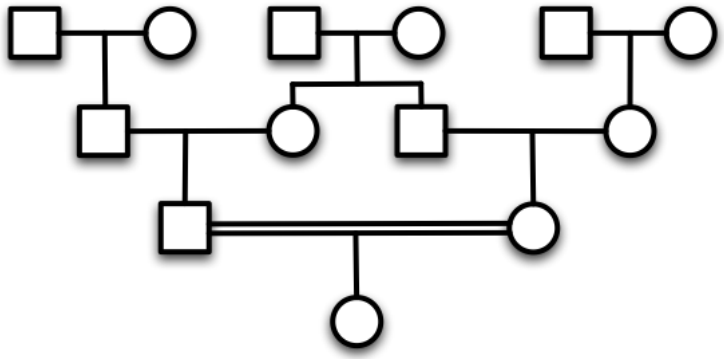
- The Y-chromosome legacy of Ghengis Khan
  (Zerjal et al. 2003, American Journal of Human Genetics)

Mating structure

- Panmixia: random-mating

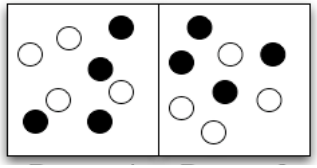
- Assortative mating
  - Non-random
  - Leads to departures from Hardy-Weinberg ***genotype*** frequencies
  - ***Allele*** frequencies can remain unchanged

- Inbreeding
  - Preferential mating with relatives

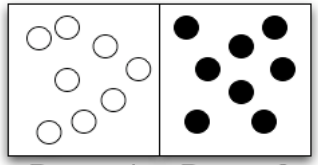


Mating structure: F_{ST}

$$F_{ST} = \frac{Var(p)}{\bar{p}(1-\bar{p})}$$



$F_{ST} = 0$



$F_{ST} = 1$

- F_{ST} measures how much genetic variation can be explained by sub-populations within the total population

- F_{ST} between divergent populations increases over time $F_{ST} = 1 - \left(1 - \frac{1}{2N}\right)^t$

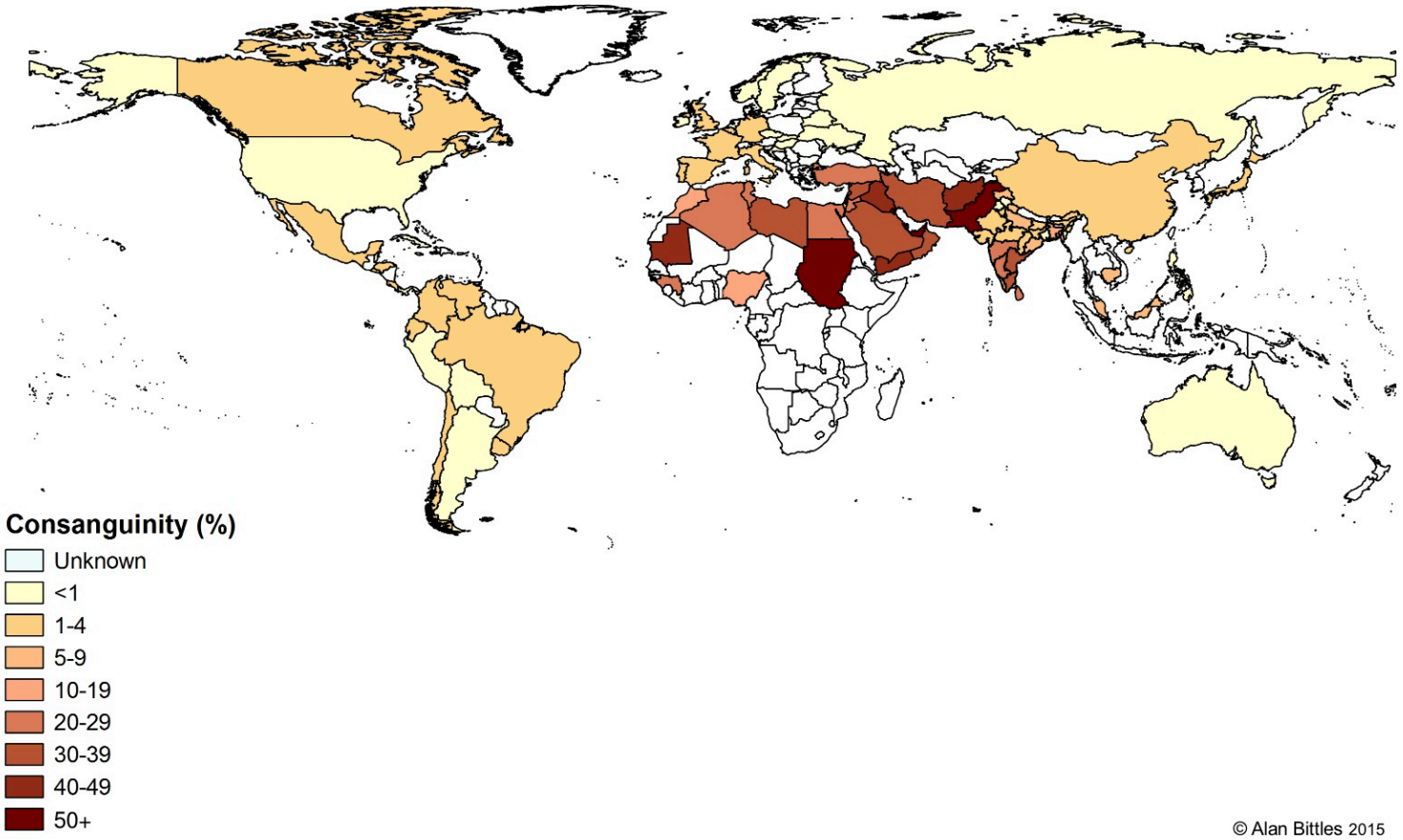
- Migration reduces F_{ST} (island model) $F_{ST} = \frac{1}{(4Nm+1)}$

Mating structure: inbreeding

$$F = 1 - \frac{H}{2pq}$$

- Inbreeding coefficient (*F*): Another F-statistic can be used to quantify the effects of inbreeding (the inbreeding coefficient
- Inbreeding results in an excess of homozygotes
- As many deleterious alleles are recessive this can result in adverse effects

Mating structure example (inbreeding)



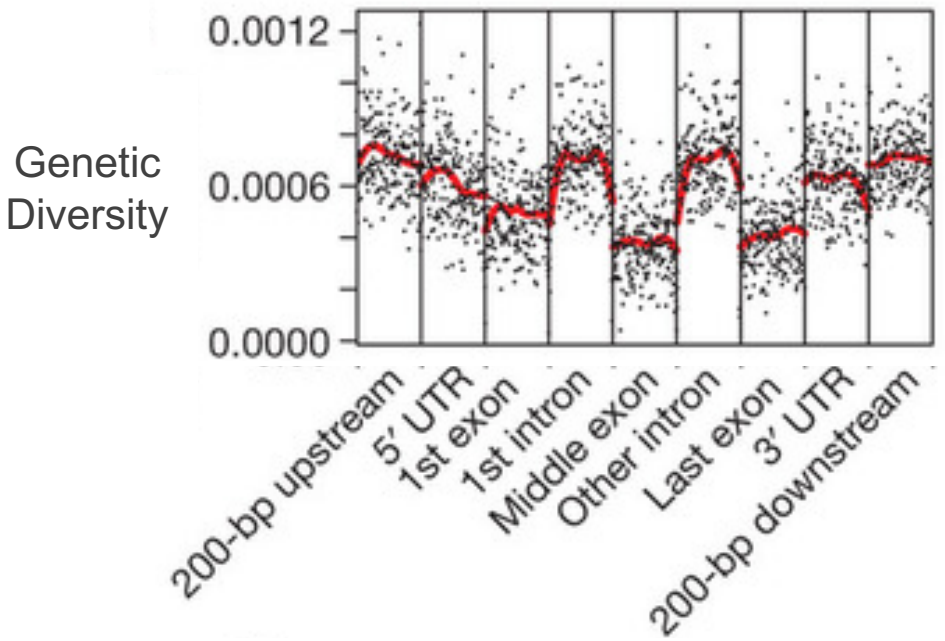
- Consanguinity: closer than 2nd cousin mating ($F > 0.015625$)

Effects of each major process

	Genetic Drift	Natural Selection	Mutation	Migration	Mating Structure
Time-scale	Medium	Fast	Slow	Medium	Fast
Effect on variation	Reduced	Mixed	Increased	Homogenized	Indirect

Case study #2

- Polymorphism data from the 1000 Genomes Project (*Nature,* 2010)



- What do you think causes these patterns?

Break

Advanced concepts in population genetics

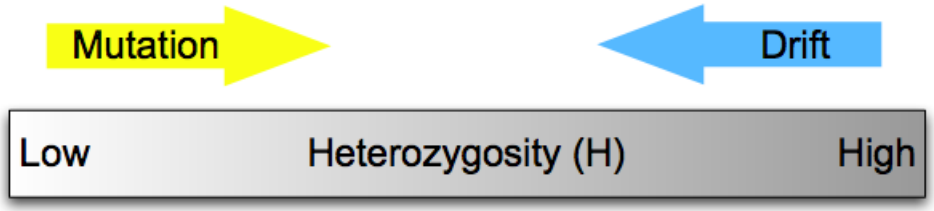
Genetic drift	**Natural selection**	**Mutation**	**Migration**	**Mating structure**	
Genetic drift	Nearly-neutral theory (Ohta)	Neutral theory (Kimura)	Gene flow	Inbreeding	**Genetic drift**
	Natural selection	Mutation-selection balance	Migration-selection balance	Sexual selection	**Natural selection**
		Mutation	Geographical genetics	Private alleles	**Mutation**
			Migration	Wahlund effect	**Migration**
				Mating structure	**Mating structure**

Neutral theory of evolution (Kimura)

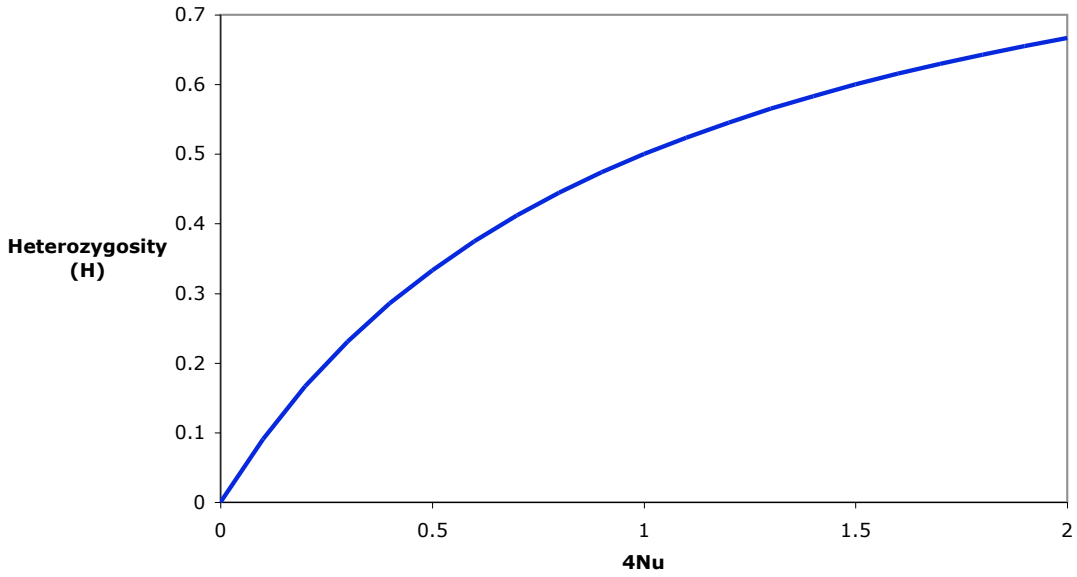
- Drift + mutation
- Most mutations are deleterious (bad)
- Most polymorphisms are neutral (neither good nor bad)
  - Synonymous changes (codon change, but same amino acid)
  - Pseudogenes: “dead genes” that are no longer expressed
  - Intergenic DNA

- A balance exists between a decrease in variation due to drift and an increase in variation due to mutation

$$\Delta H_{mutation} = 2\mu(1 - H) \qquad \Delta H_{drift} = -\left(\frac{1}{2N}\right) H$$



Neutral theory of evolution (Kimura)

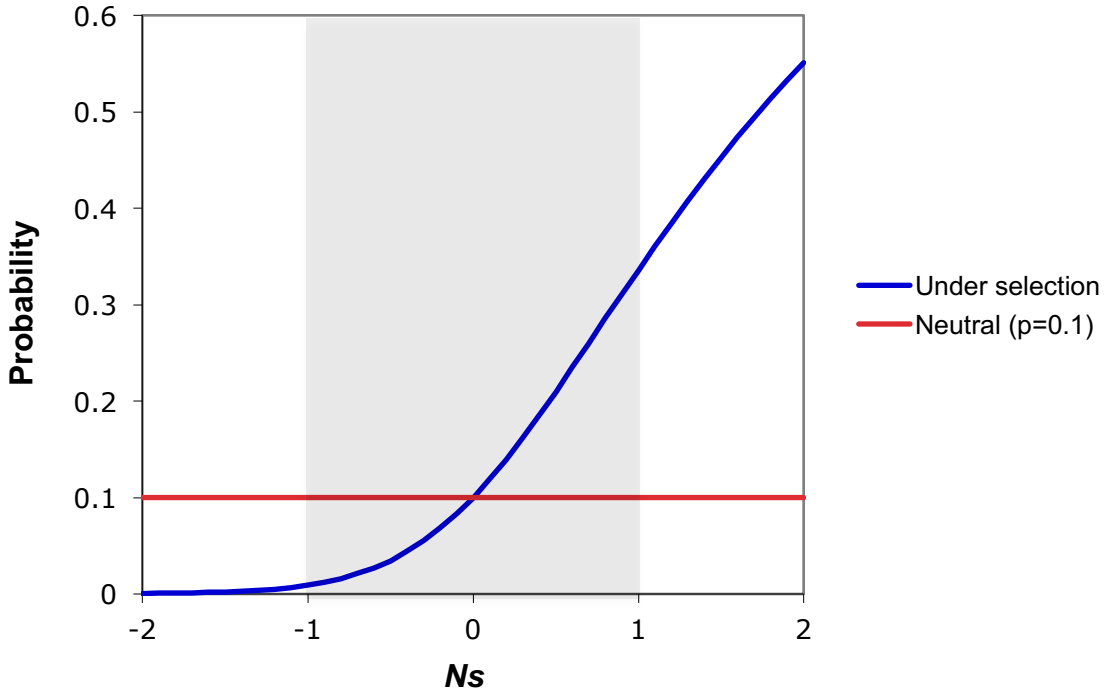


$$\hat{H} = \frac{4N\mu}{1+4N\mu}$$

- Substantial genetic variation is maintained if $4N\mu >> 1$
- Population-level mutational input $(2N\mu)$ is important
- $\theta = 4N\mu$ pervades population genetics and coalescent theory
- The neutral theory provides a null hypothesis for studies of molecular evolution

Nearly-neutral theory (Ohta)

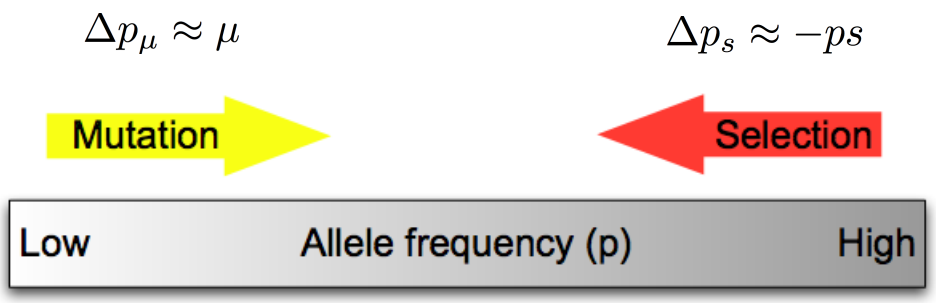
- The critical value is 4*Ns*
  - When |4*Ns*| >> 1, alleles undergo selection
  - When |4*Ns*| << 1, alleles are effectively neutral



$$Pr(fix) \approx \frac{1 - e^{-4Nsp}}{1 - e^{-4Ns}}$$

Mutation-selection balance

- Mutation + selection
- Deleterious mutants increase in frequency by mutation
- Deleterious mutants are reduced in frequency by selection
- There exists an equilibrium allele frequency where the magnitude of these two forces are balanced:
- Alleles under mutation-selection balance are rare

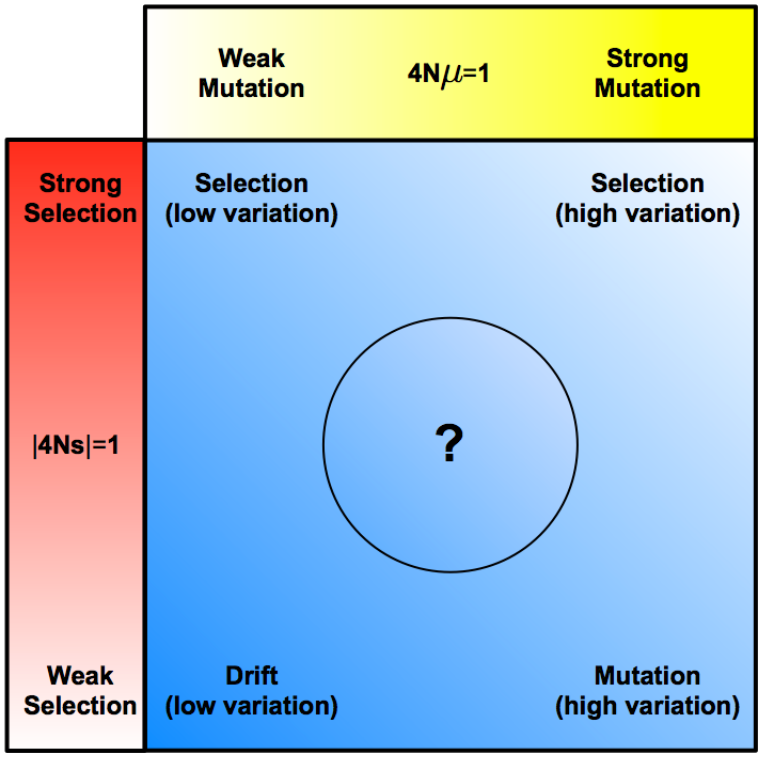


Mutation-selection balance

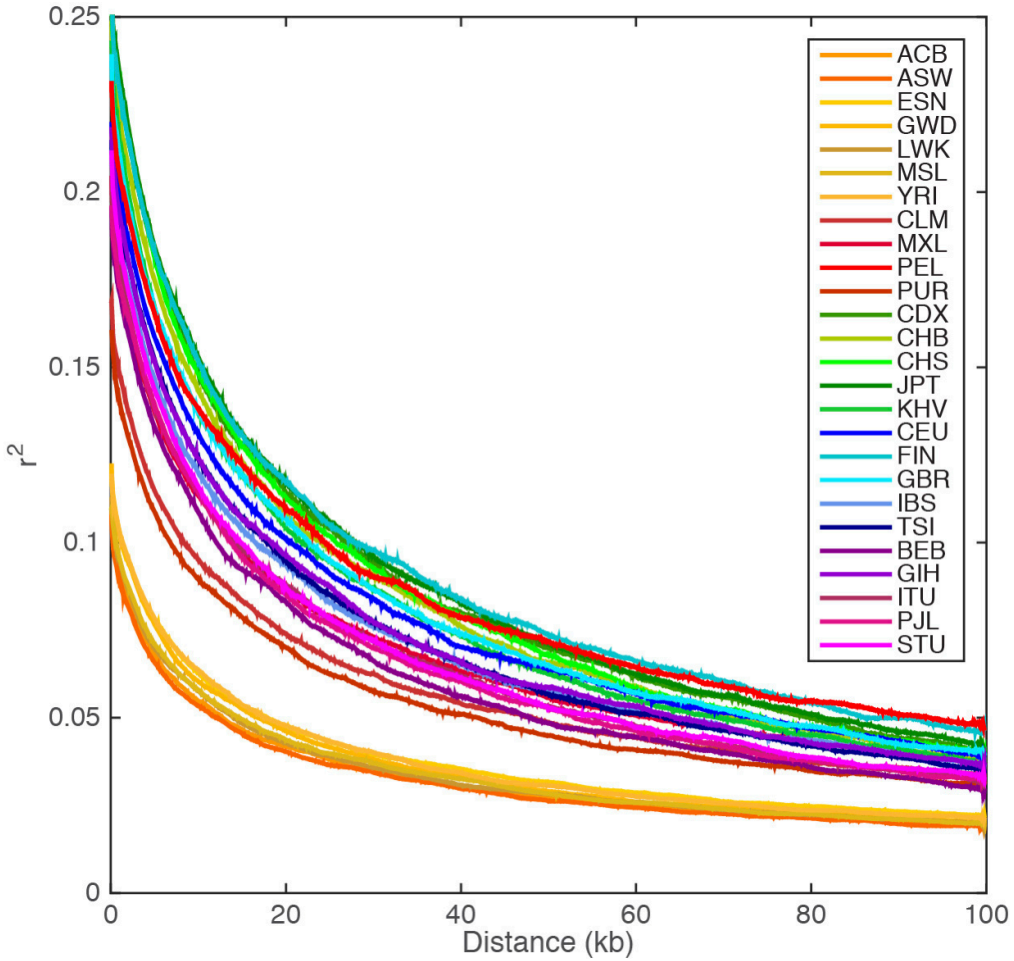
- Ploidy and dominance affect equilibrium allele frequencies
  - Haploid $\hat{p} \approx \frac{\mu}{s}$
  - Diploid, completely recessive $\hat{p} \approx \sqrt{\frac{\mu}{s}}$
  - Diploid, intermediate dominance $\hat{p} \approx \frac{\mu}{hs}$
- Deleterious alleles are more common when recessive

Selection, drift, and mutation

- Large populations are in the upper right and small populations are in the lower left
- Where in the blue part of this figure would you expect to find:
  - Protein coding genes?
  - Disease causing genes?
  - miRNA genes?
  - Pseudogenes?
  - MHC genes?
  - Transposons?
  - Microsatellites?
  - Cis-regulatory elements?



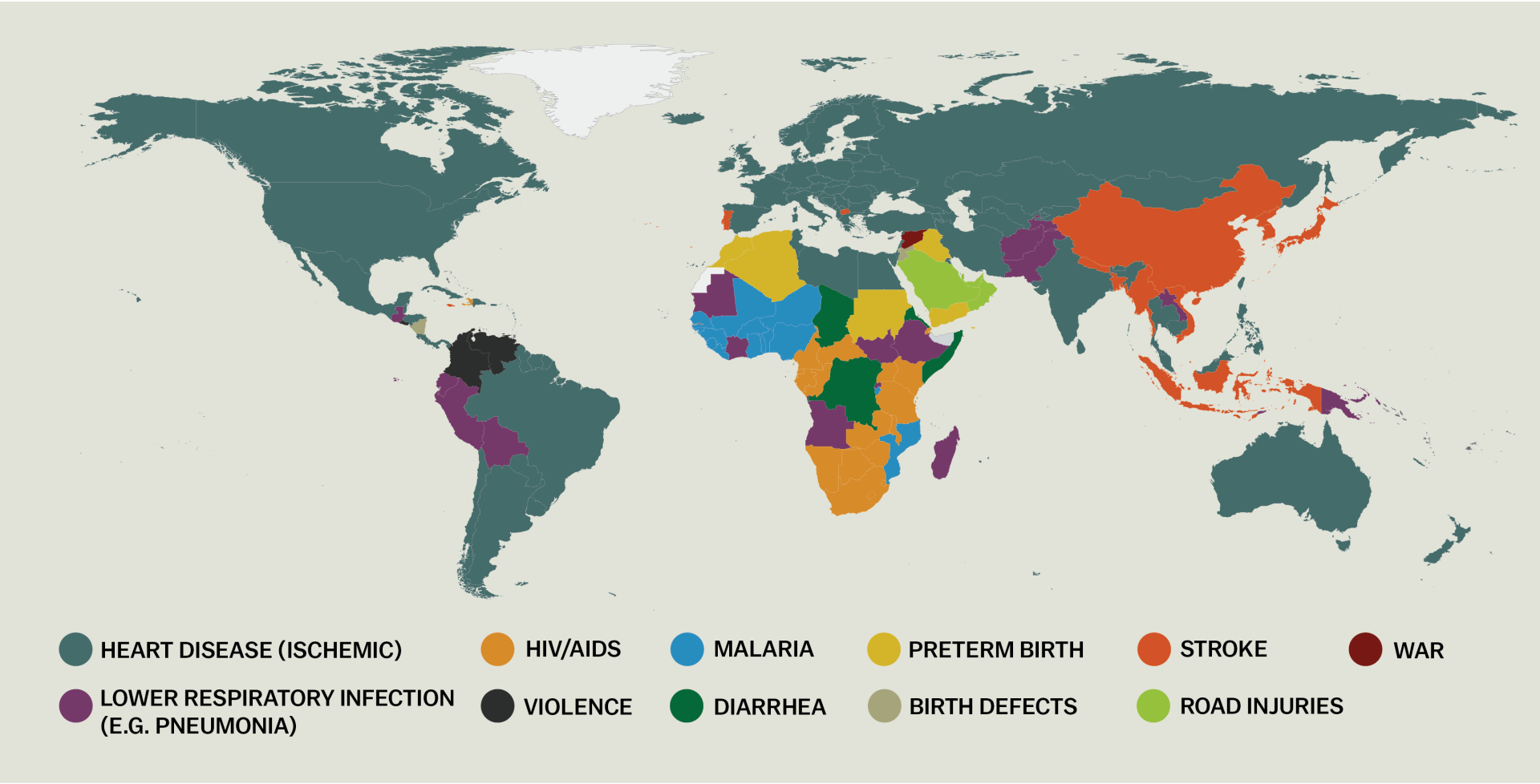
Linkage disequilibrium in human populations



Phase 3 data from the 1000 Genomes Project (*Nature*, 2015)

- Non-African populations have higher amounts of LD

Leading causes of lost years of life (2013)



Source: *Vox* and *The Lancet*

Replicating GWAS in multiple populations

		Direction relative to EA[a]			
		All Index SNPs	Index SNPs Replicated in EA	Index SNPs Not Replicated in EA	Strength Relative to EA
Pop.	N[b]	Same:Opposite[c]	Same:Opposite[c]	Same:Opposite[c]	Stronger:Weaker[d]
AA	14,492	57:11***	43:8***	14:3	0:12**
HA	8,202	60:8***	46:5***	14:3	0:0
AS	5,425	45:21**	34:15*	11:6	0:0
NA	6,186	45:10***	35:8***	10:2	0:2
PI	1,801	48:14***	34:12***	14:2	1:0

EA: European Americans, AA: African-Americans, HA: Hispanic Americans, AS: Asian Americans, NA: Native Americans, PI: Pacific Islanders
PAGE Study traits and diseases: BMI, lipid levels, and T2D

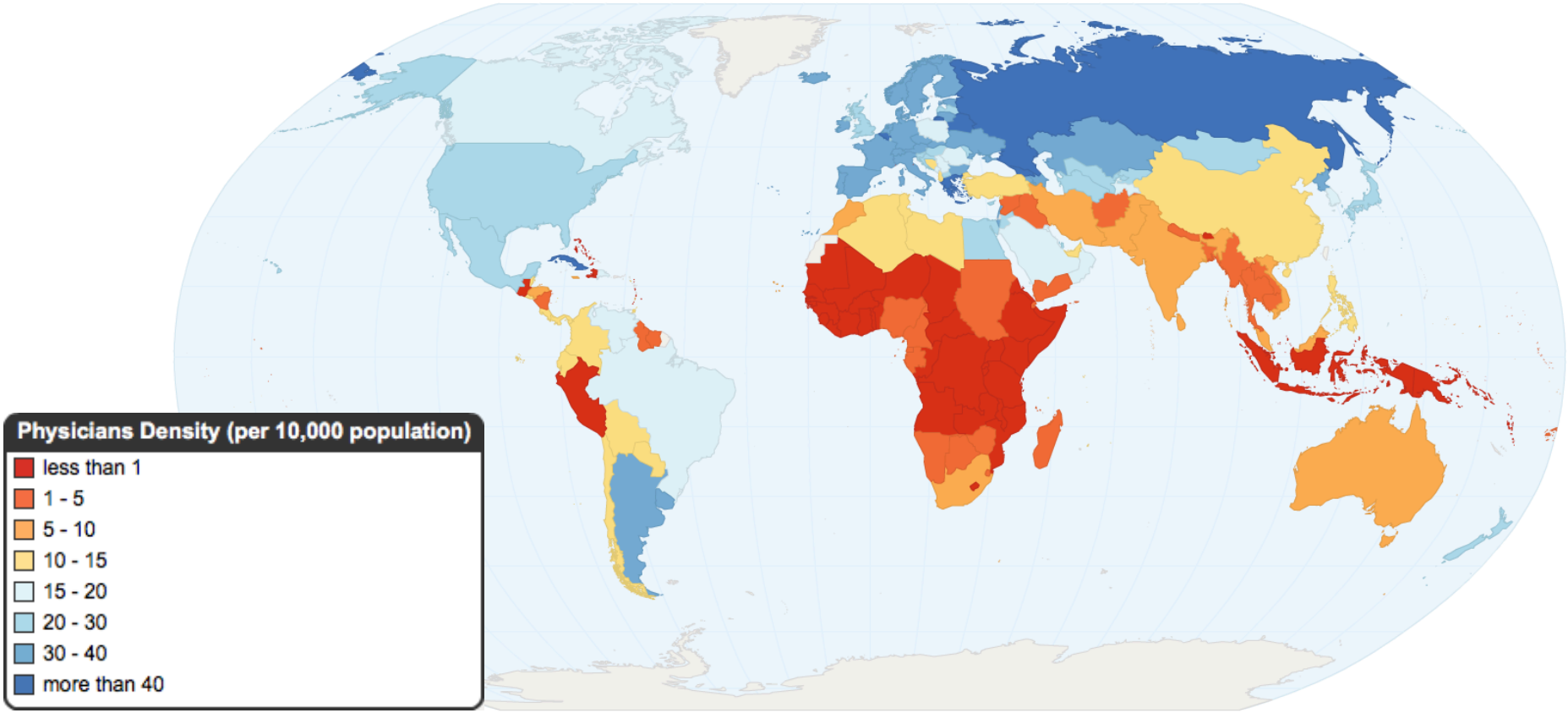
- Cases and controls need to be matched by ethnicity
- Odds ratios, risk allele frequencies, and LD can differ across populations
- Do you expect to find the same “hits” in each population?

Carlson et al. (2013, *PLoS Biology*)

Contributing factors

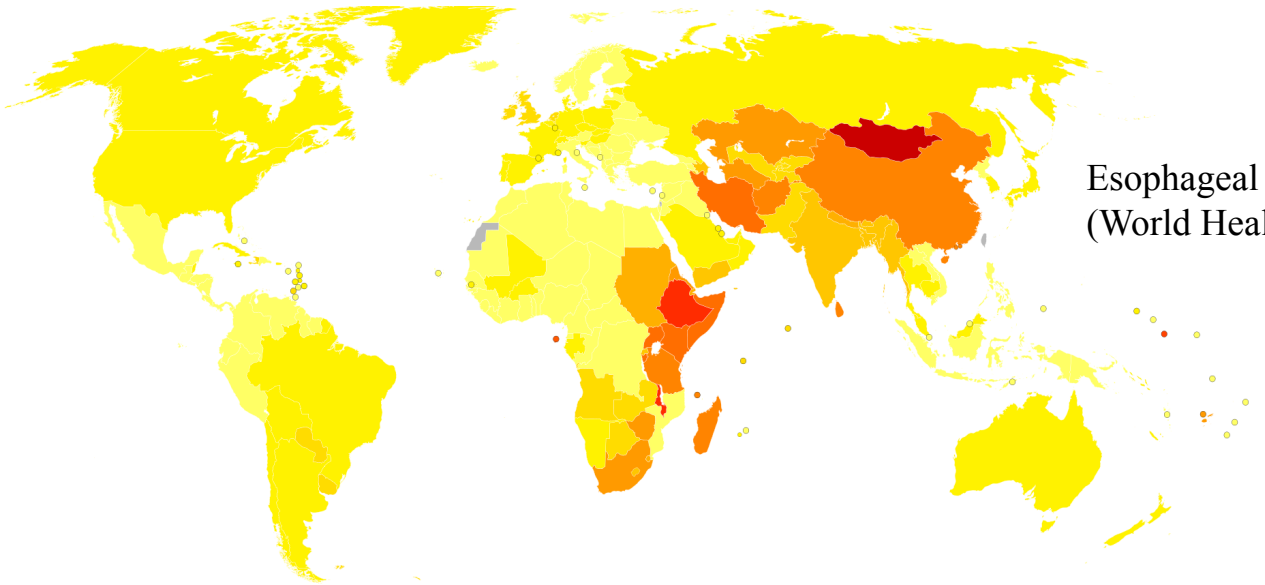
- Environment
- Genetic architecture
- Population bottlenecks
- Natural selection

Access to health care



Source data: World Health Organization (2010)

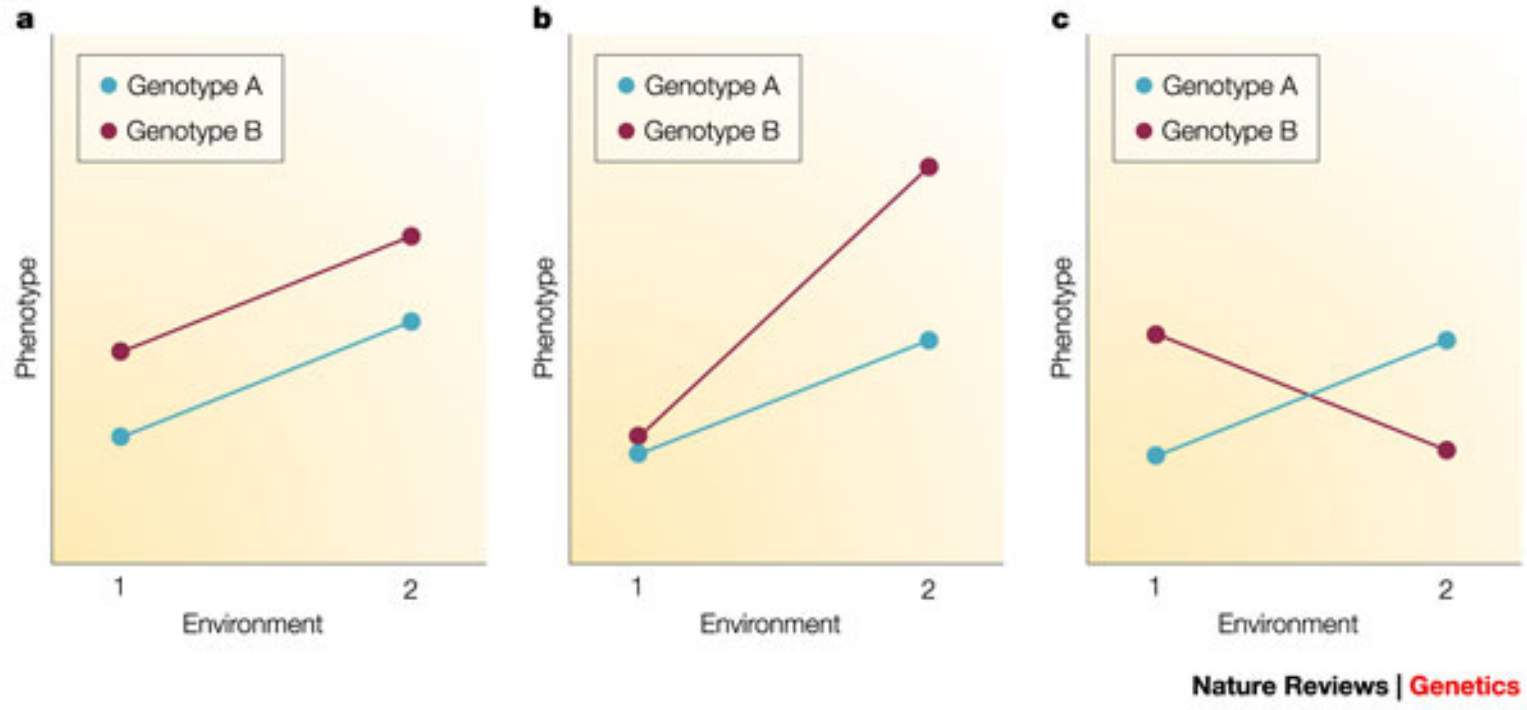
Environmental risk factors



Esophageal cancer death rates
(World Health Organization, 2004)

- Many different environmental risk factors exist
  (e.g. smoking, *Plasmodium falciparum*, famine - Dutch *Hongerwinter* of 1944)
- Environmental factors supply contexts in which natural selection acts
- Geographic patterns may help identify factors that contribute to diseases

Genotype-by-environment (GxE) interactions



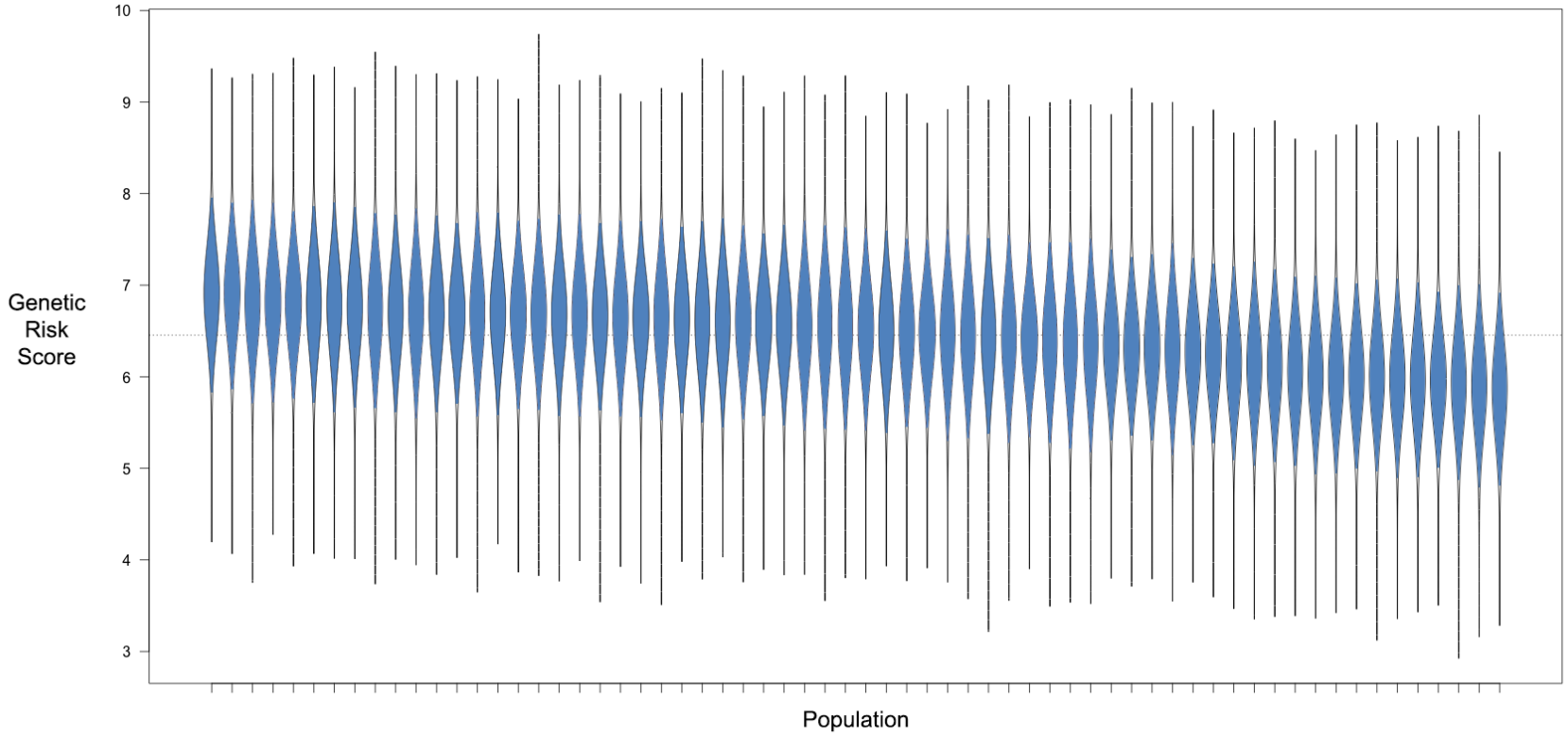
- **Reaction norms** describe the range of phenotypes produced by a genotype in different environment

Genetic architecture: monogenic disorders

- Single gene disorders are more likely to contribute to health disparities
- What are some evolutionary forces processes that can lead to large allele frequency differences across populations?

Image from GATTACA (Columbia Pictures)

Genetic architecture: polygenic disorders



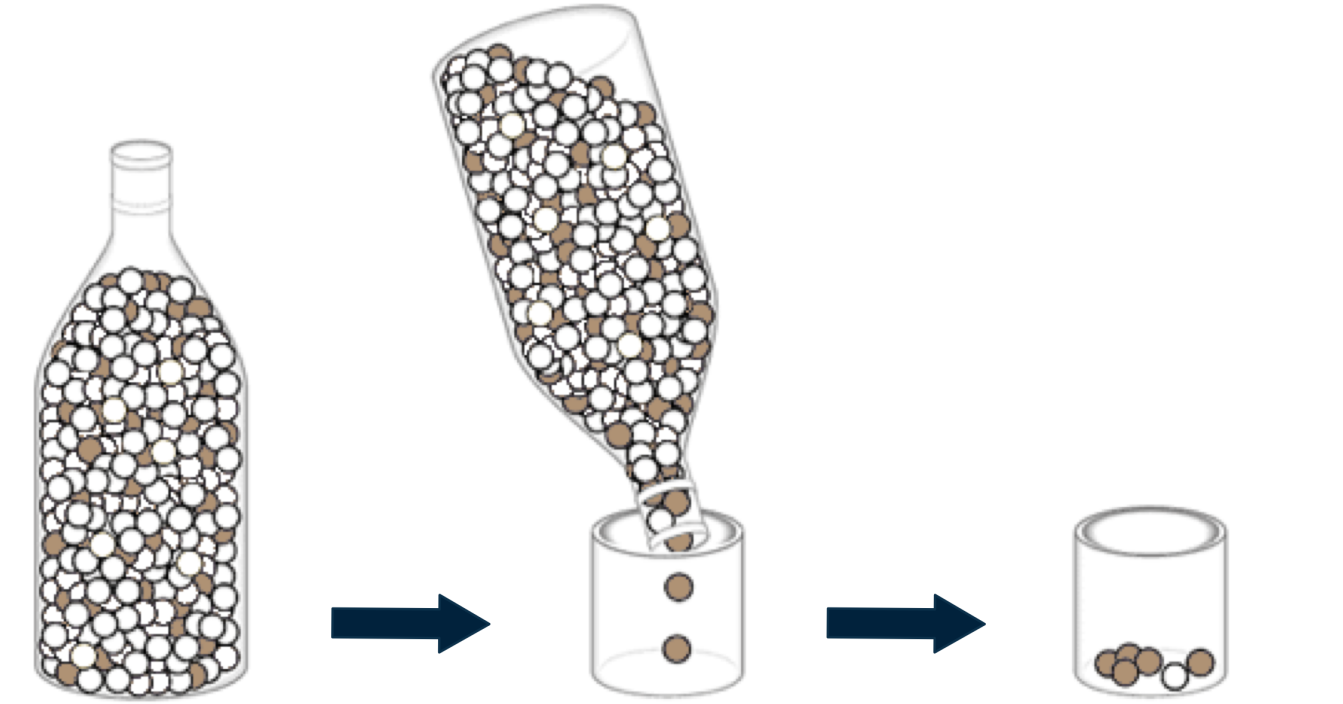
- If a large number of loci contribute to a disease… it is less likely that there will be large differences in genetic risk across populations

Dominance and recessivity

Population	Allele frequency	Homozygote frequency
Population A	0.1	0.01
Population B	0.2	0.04

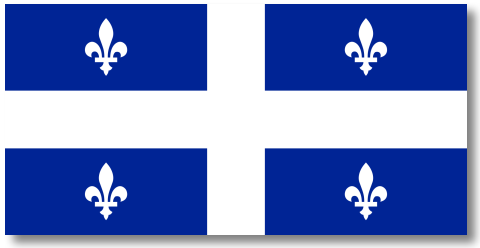
- Small differences in allele frequencies are magnified for recessive diseases

Population bottlenecks and founder effects



Examples of founder effects

- French Canadians (Québécois)



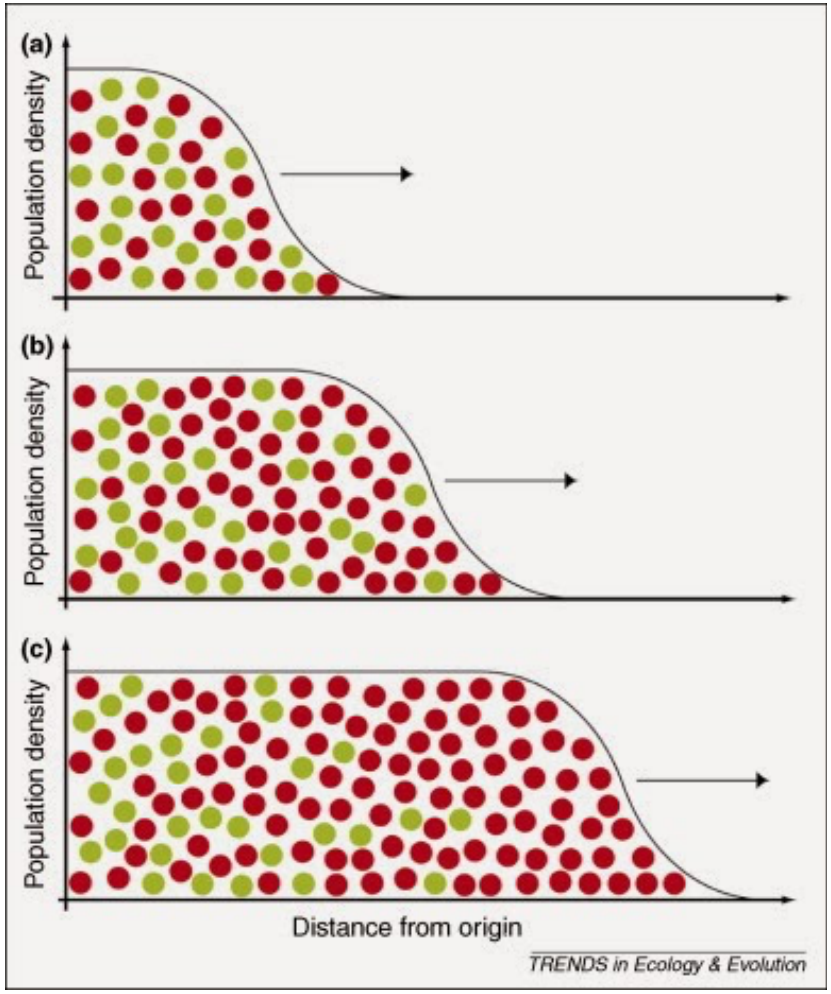
- Old Order Amish

- HMS Bounty mutineers and Pitcairn Island

Diseases associated with founder effects

Population	Disease
Afrikaners in South Africa	Fanconi anemia
Ashkenazi Jews	Tay-Sachs disease
Lake Maracaibo area, Venezuela	Huntington's disease
Island of Tristan de Cunha	Retinitis pigmentosa

Allele surfing



Genetic load

$$L = \frac{w_{max} - w}{w_{max}}$$

- Natural selection eficiently eliminates deletious alleles when $|4N_e s| > 1$
- Since non-African populations have experienced population bottlenecks in the last 75,000 years, they have a lower effective population size
- This means that purging of mildly deleterious alleles is likely to have been lest effective in non-African populations
- Non-African genomes also have increased homozygosity (which can be an issue if deleterious alleles are recessive)

Do non-African populations have greater load?

- Simons et al. (*Nature Genetics*, 2014) state that human demographic history has "probably had little impact on the average burden of deleterious mutations."
- Do et al. (*Nature Genetics*, 2015) find little difference in the efficacy of natural selection across different human populations.
- But see Lohmueller (*Current Opinion in Genetics and Development*, 2014)…

Table 1

Statistically significant differences in patterns of deleterious variants in African and non-African populations

	Number heterozygous genotypes per individual	Number homozygous derived genotypes per individual	Number derived alleles per individual	Number synonymous variants in a sample	Number nonsynonymous variants in a sample	Proportion of variants in a sample that are nonsynonymous
African	Higher	Lower	Approximately equal	Higher	Higher	Lower
Non-African	Lower	Higher	Approximately equal	Lower	Lower	Higher
Mechanism	Bottleneck in non-African population reduced number of heterozygous variants	Bottleneck in non-African population led to increase in high-frequency derived variants	Different effects may cancel and/or lack of power[a]	Bottleneck in non-African population reduced number of variants	Bottleneck in non-African population reduced number of variants	Recovery from a bottleneck; spatial expansion
Reference	[17,19,43,44]	[17,19,43,44]	[17,27•,32,37•]	[14,17]	[14,17]	[17,36••,43,45]

[a] Lack of a significant difference in the number of deleterious alleles per individual in African and non-African populations may be due to a lack of power to detect slight differences. Recent growth and population bottlenecks are predicted to only slightly increase this quantity [31,37•] (also see Section 'efficacy of natural selection').

Local adaptation

Approaches used to detect adaptation

Comparative genomics

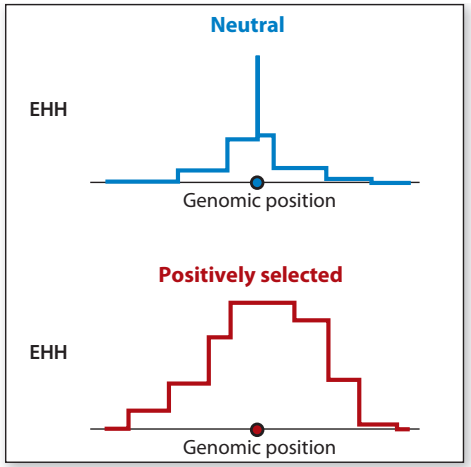
Neutral

	Nonsynonymous	Synonymous
Fixed	4	4
Polymorphic	3	3

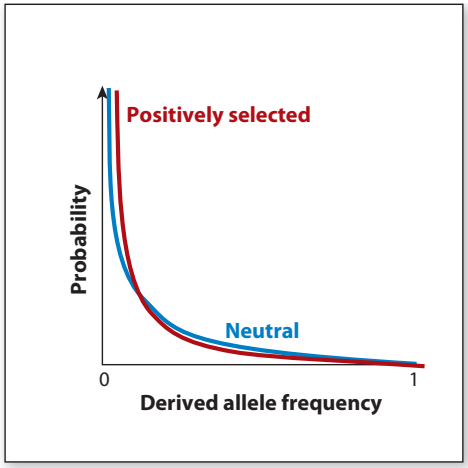
Positively selected

	Nonsynonymous	Synonymous
Fixed	8	4
Polymorphic	3	3

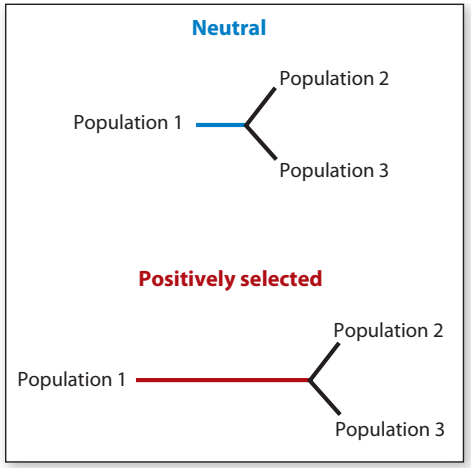
Haplotype statistics



Allele frequencies



Multiple populations

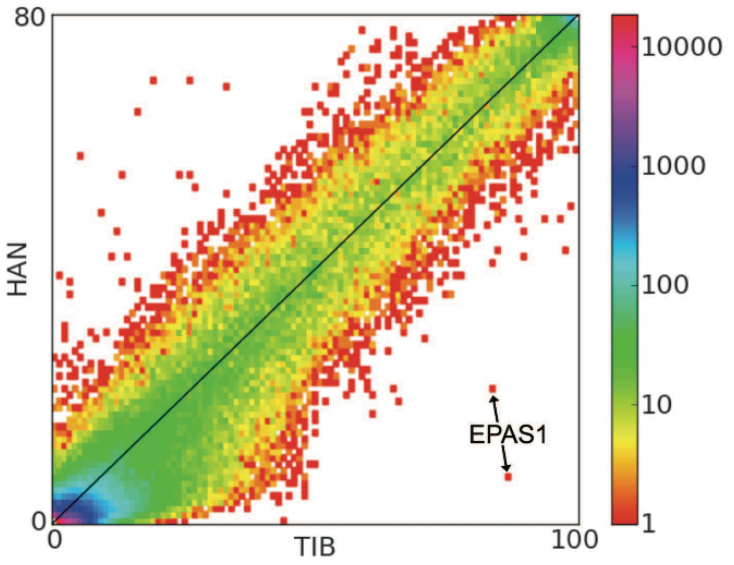


Modified from: Lachance and Tishkoff (2013, *AREES*)

EPAS1 and high-altitude

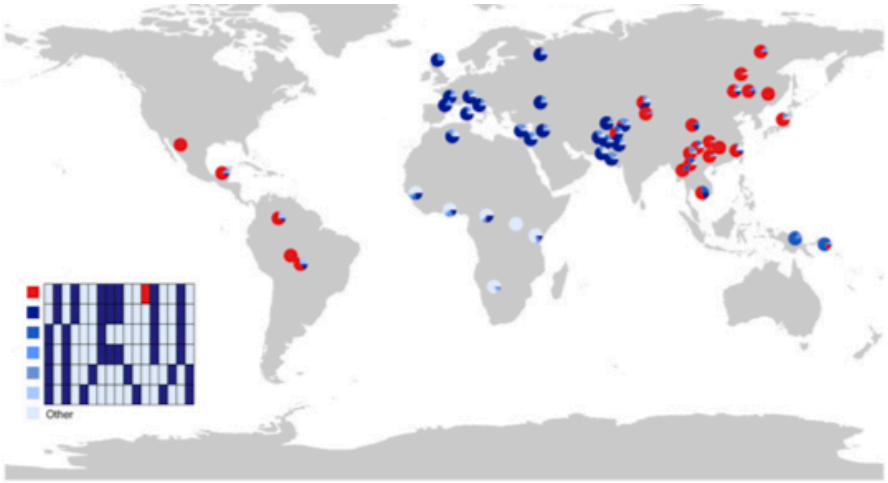
- Reduced [O_2] is a strong selective pressure
- Allele frequencies compared between Tibetans (TIB) and Han Chinese from Beijing (HAN)
- Outlier SNPs are located near *EPAS1*, a hypoxia-induced transcription factor
- The Tibetan EPAS1 haplotype comes from Denisovans (Huerta-Sanchez et al. 2014)**!!!**
- Positively selected *EPAS1* haplotype contains a deletion that occurred 12kya (Lou et al. 2015)

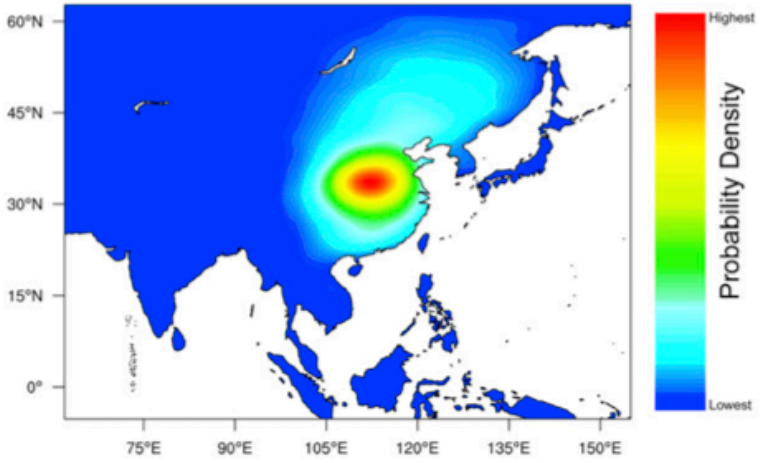
Image rights: EasyTourChina



Yi et al. (2010, *Science*)

EDAR and eccrine glands

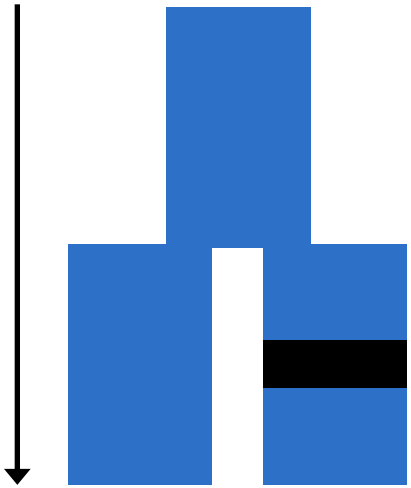


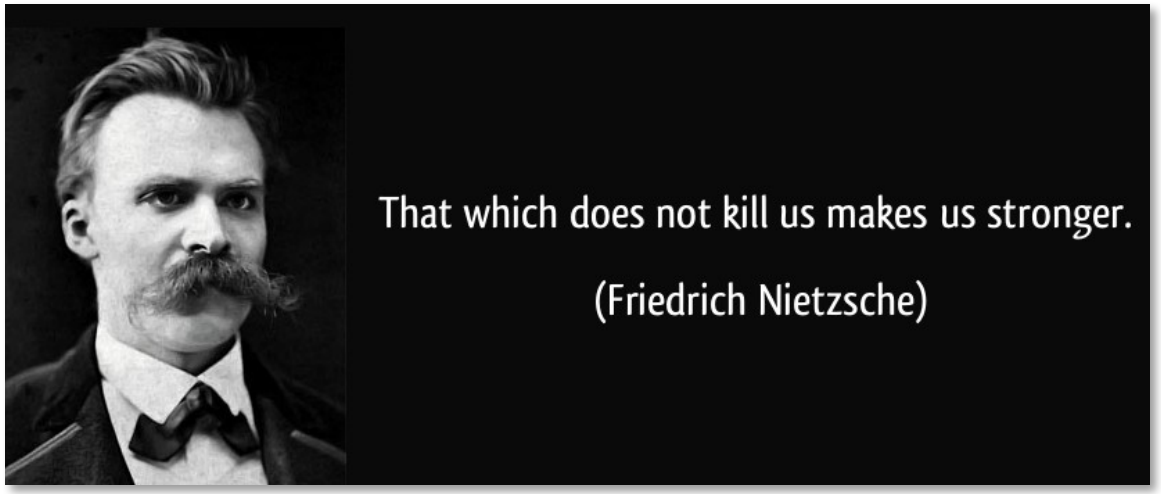


- CMS scans reveal that the *EDAR V370A* allele is a target of selection
- *EDAR* encodes the Ectodysplasin receptor
- Relevant phenotypes in humans and mice
  - Increased hair thickness
  - Increased eccrine (sweat) gland density

Kamberov et al. (2013, *Cell*)

The benefits of a challenging past

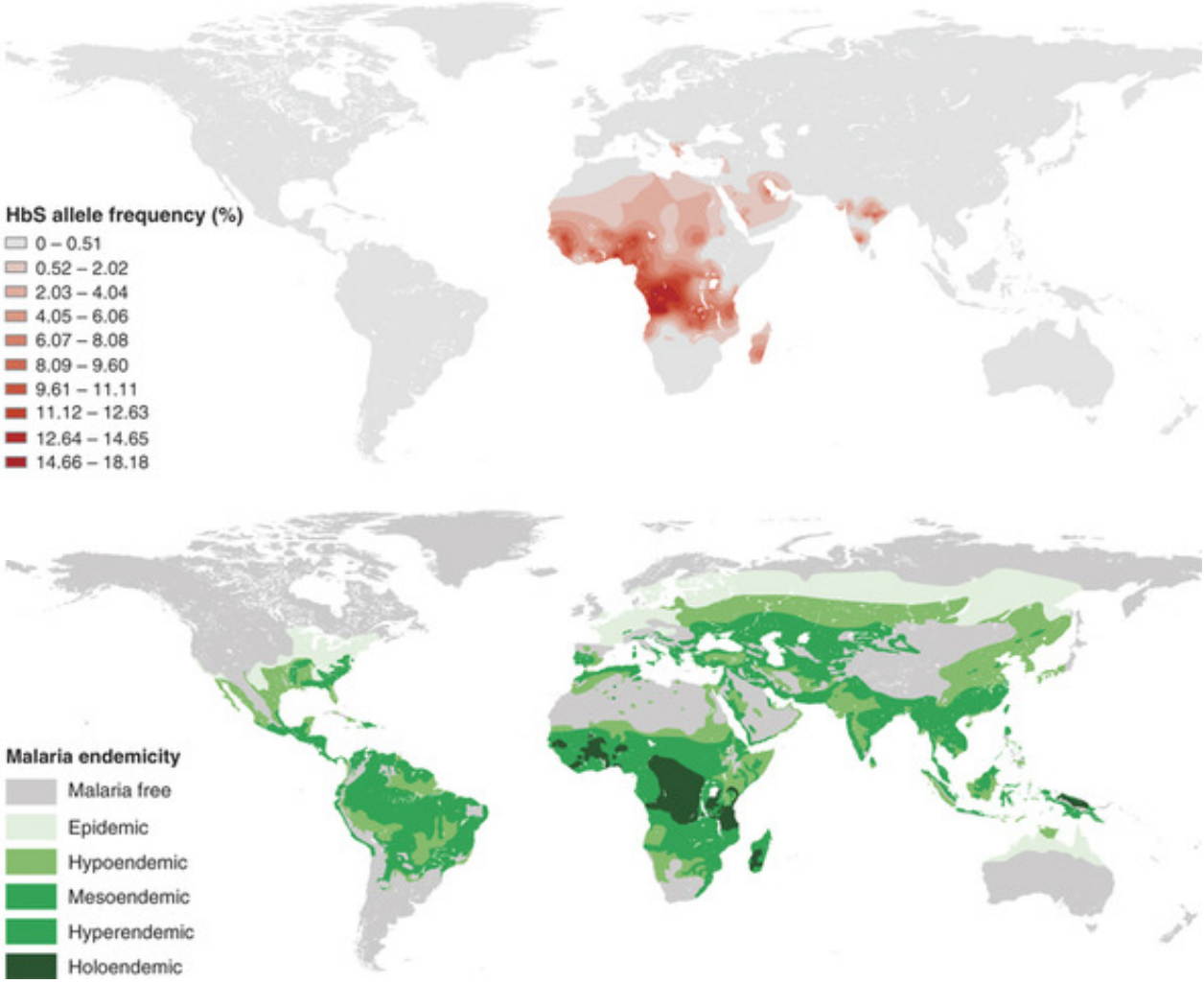




- Multiple mechanisms
  - Positive selection increases the frequency of protective alleles
  - Negative selection decreases the frequency of risk alleles
  - High environmental risks can coincide with lower genetic risks

- Example: *CCR5* Δ32 and HIV resistance in Europe

Trade-offs



Piel et al. (2010, *Nature Communications*)

The thrifty gene hypothesis

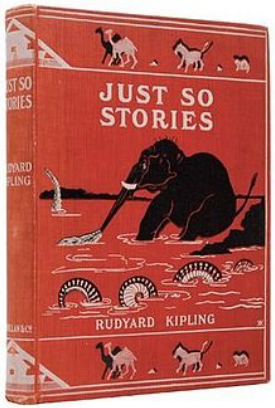
- Type 1 diabetes (T1D)
  - Early onset and insulin deficiency
- Type 2 diabetes (T2D)
  - Adult onset and insulin resistance

Art by Banksy

- James Neel (1962): Paleolithic feast-famine cycles may have selected for the ability to fatten rapidly. “Thrifty genes” confer a predisposition to diabetes.

- How much support is there for this hypothesis? Ayub et al. (2014, *AJHG*) found only minimal support for positive selection at T2D loci.

The dangers of story telling



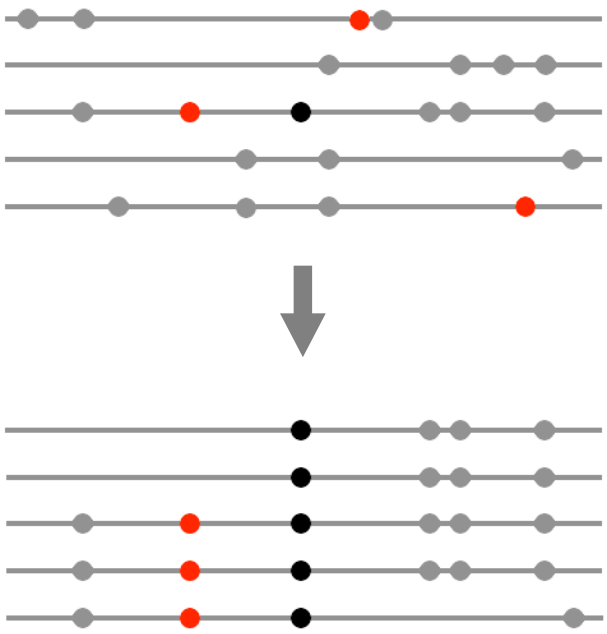
- It is a little too easy to make up stories of adaptive evolution
- Be careful when identifying traits that have been under selection in the past
- Allele surfing and gene conversion can mimic signatures of positive selection
- Convincing narratives of selection can be made for random sets of loci (Pavlidis et al. 2012)

Mismatch diseases

Acid reflux/heartburn	Endometriosis	Lactose intolerance
Acne	Flat feet	Lower back pain
Asthma	Glaucoma	Metabolic syndrome
Athlete's foot	Gout	Myopia
Carpal tunnel syndrome	Hemorrhoids	OCD
Cavities	High blood pressure	Osteoporosis
Coronary heart disease	Iodine deficiency	Pre-eclampsia
Crohn's disease	Impacted wisdom teeth	Rickets
Diabetes (Type 1)	Insomnia	Scurvy
Eating disorders	Inflammatory bowel disease	Stomach ulcers

Table modified from *Evolutionary Medicine* by Stearns and Medzhitov

Genetic hitchhiking



- Disease alleles can hitchhike to high frequency if they are linked to locally adaptive alleles

- This can lead to large allele frequency differences if selection pressures differ across populations

Many opportunities for archaic introgression?

Ars Technica

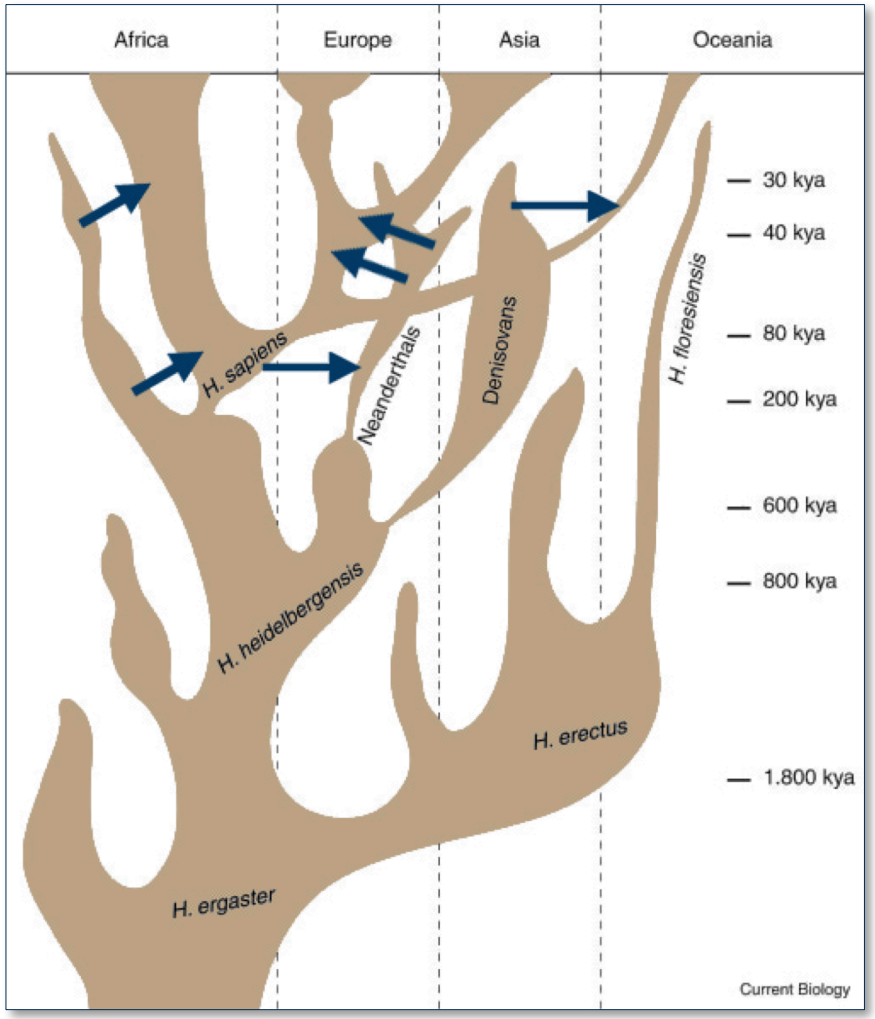


Figure modified from Lalueza-Fox and Gilbert (2011, *Current Biology*)

Introgression of disease and resistance alleles

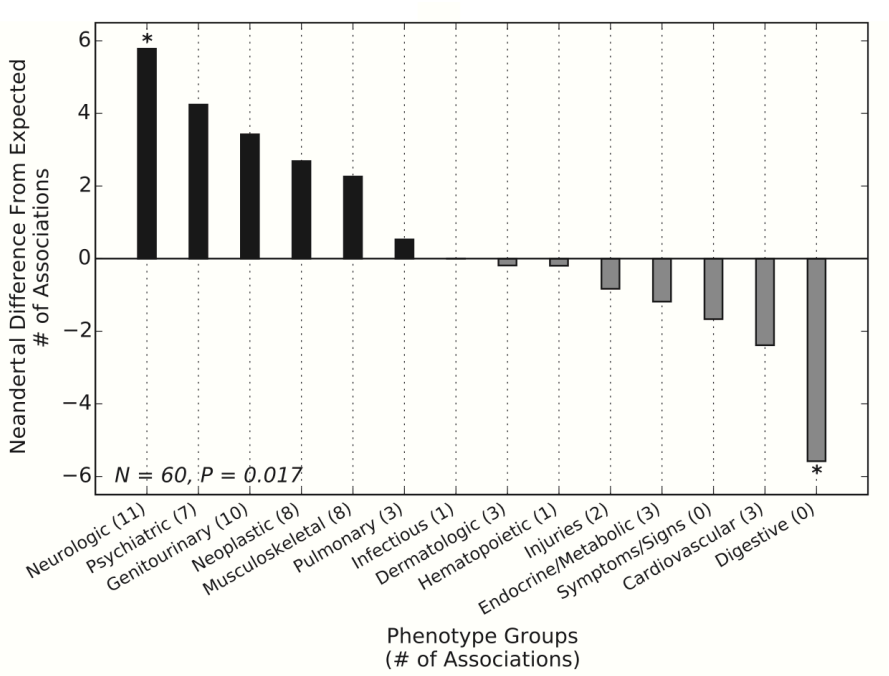


Figure from Simonti et al (2016, *Science*)

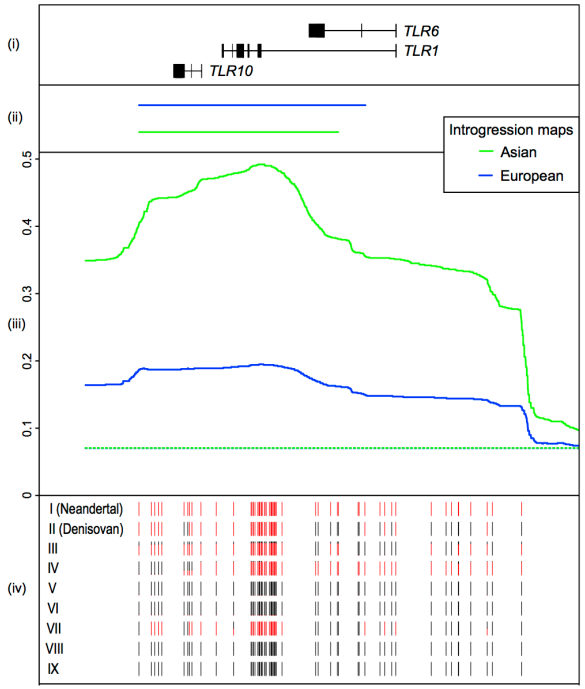


Figure from Danneman et al (2016, *AJHG*)

- Electronic health records and SNP data: Neanderthal DNA contributes to depression and skin lesions in humans (1 to 2% of risk explained)

- Introgressed Neanderthal and Denisovan TLR genes contribute to innate immunity, including antimicrobial and inflammatory response