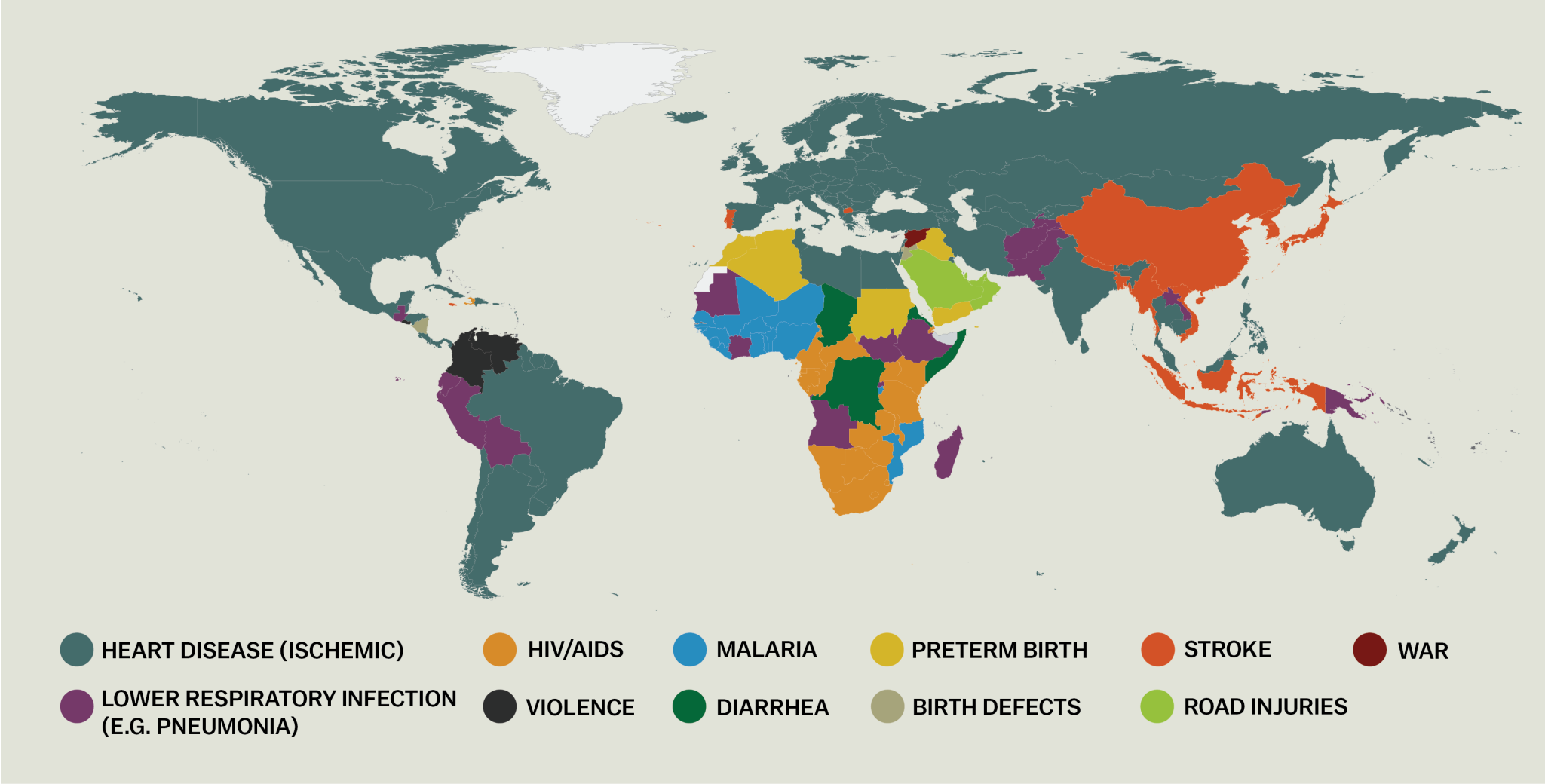
Introduction to Genetics and Genomics

5a. Evolution and Disease Risks

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

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

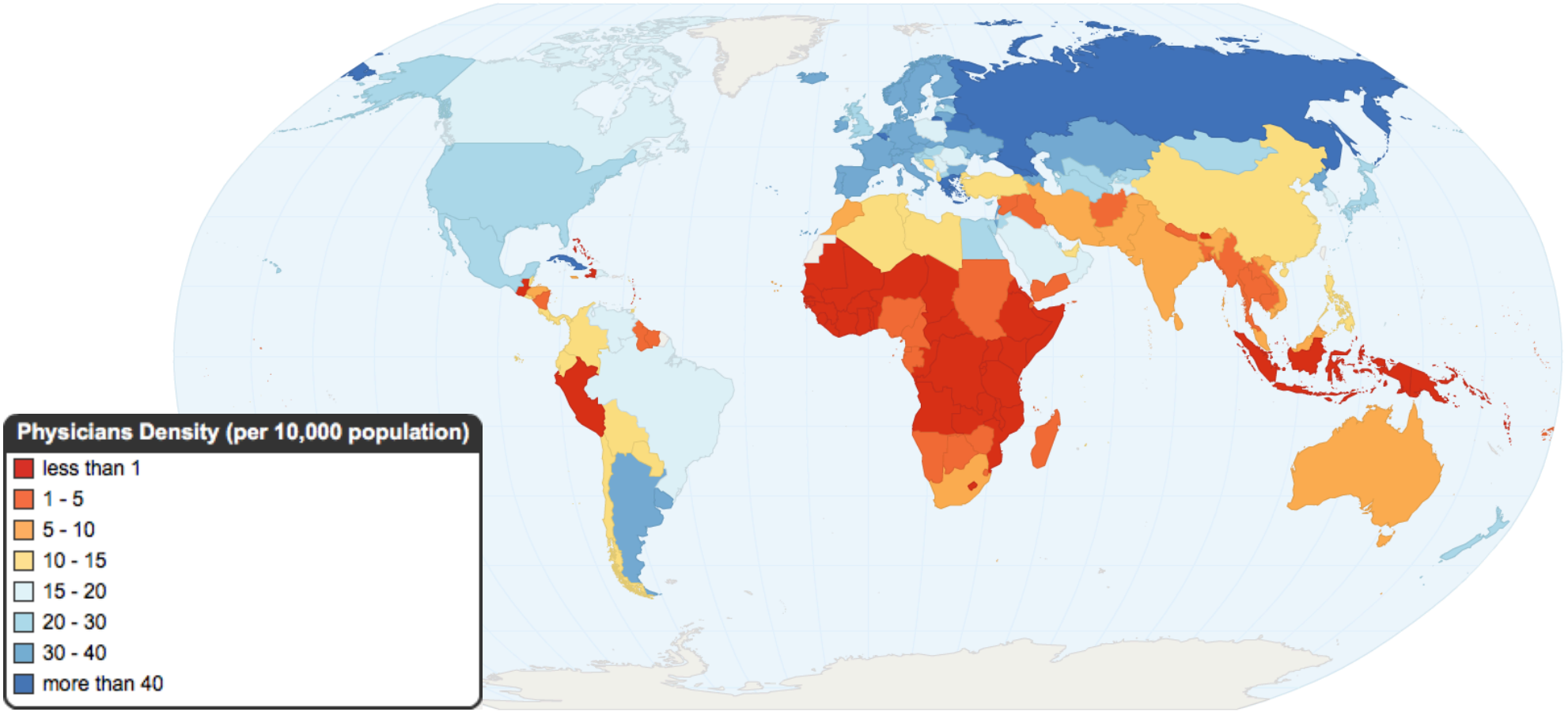
- Cases and controls need to be matched by ethnicity
- Odds ratios, risk allele frequencies, and LD can differ across populations
- Do you expect to find the same “hits” in each population?

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

Contributing factors

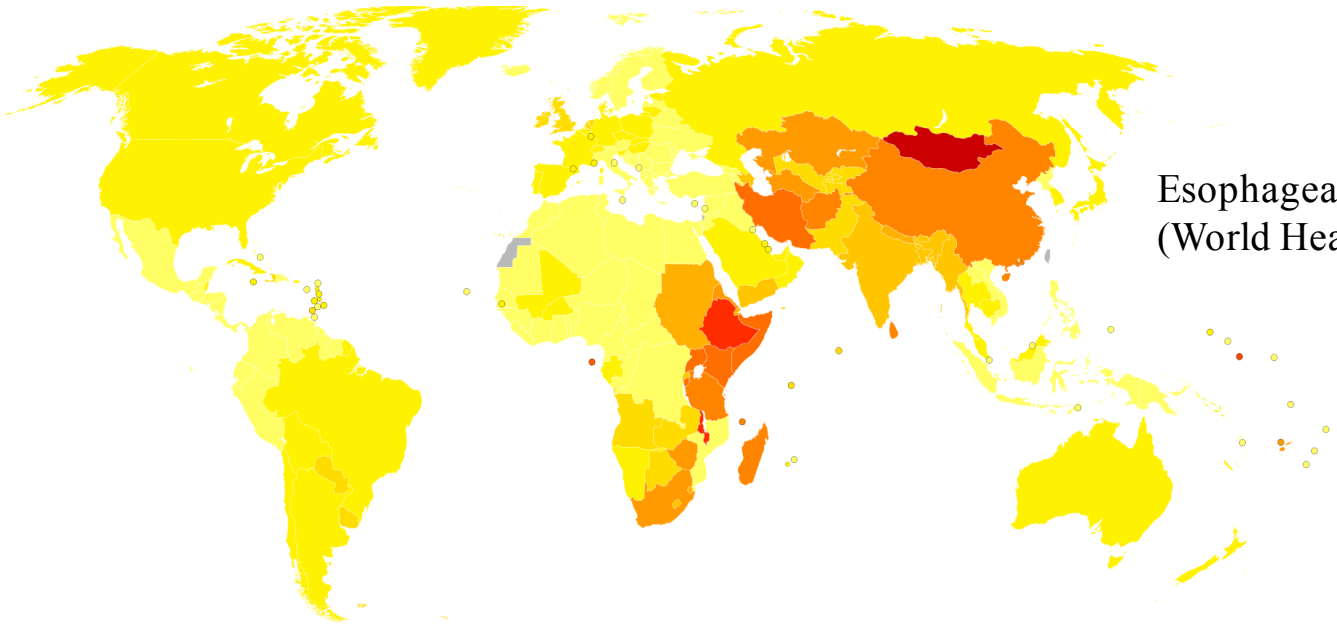
- Environment
- Genetic architecture
- Population bottlenecks
- Natural selection

Access to health care



Source data: World Health Organization (2010)

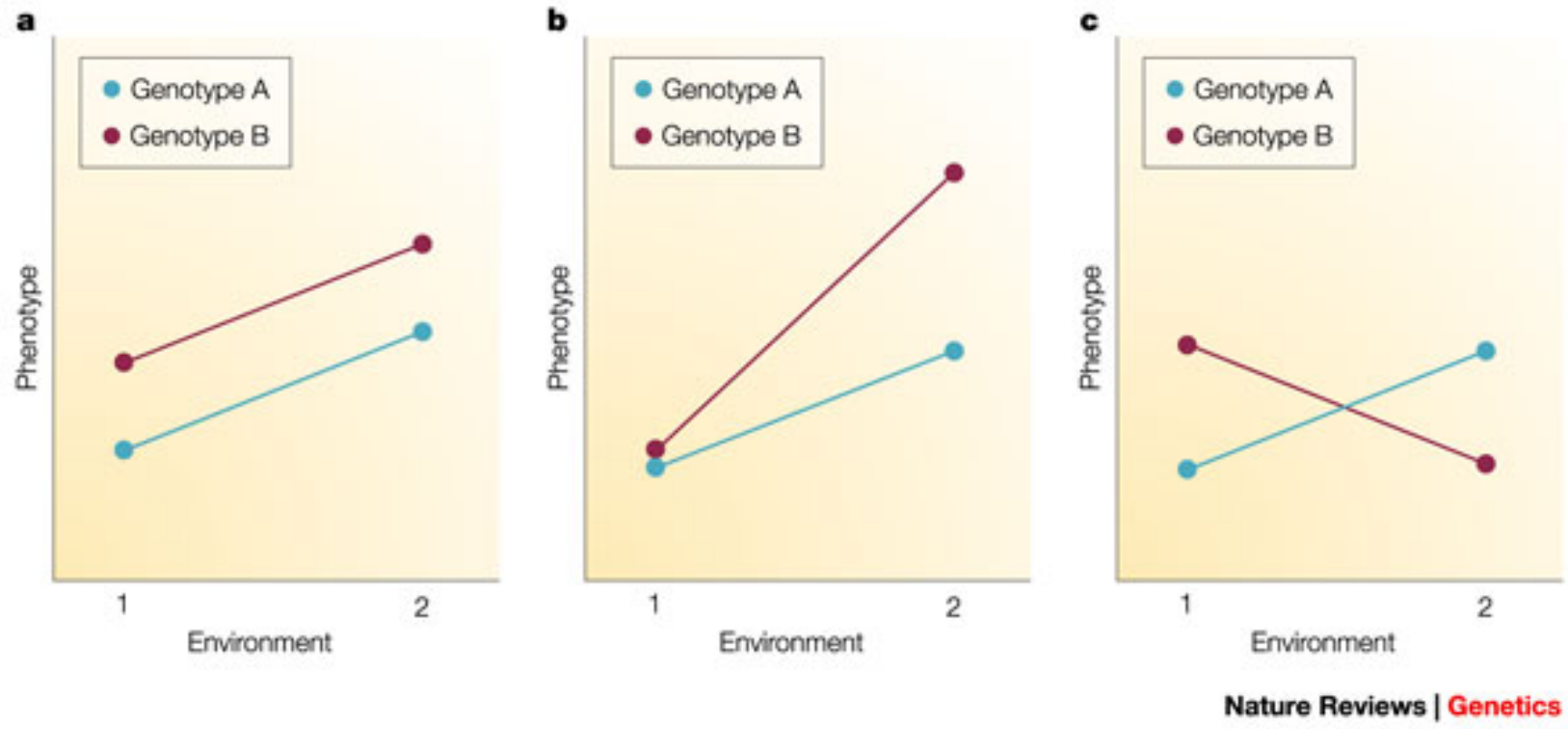
Environmental risk factors



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



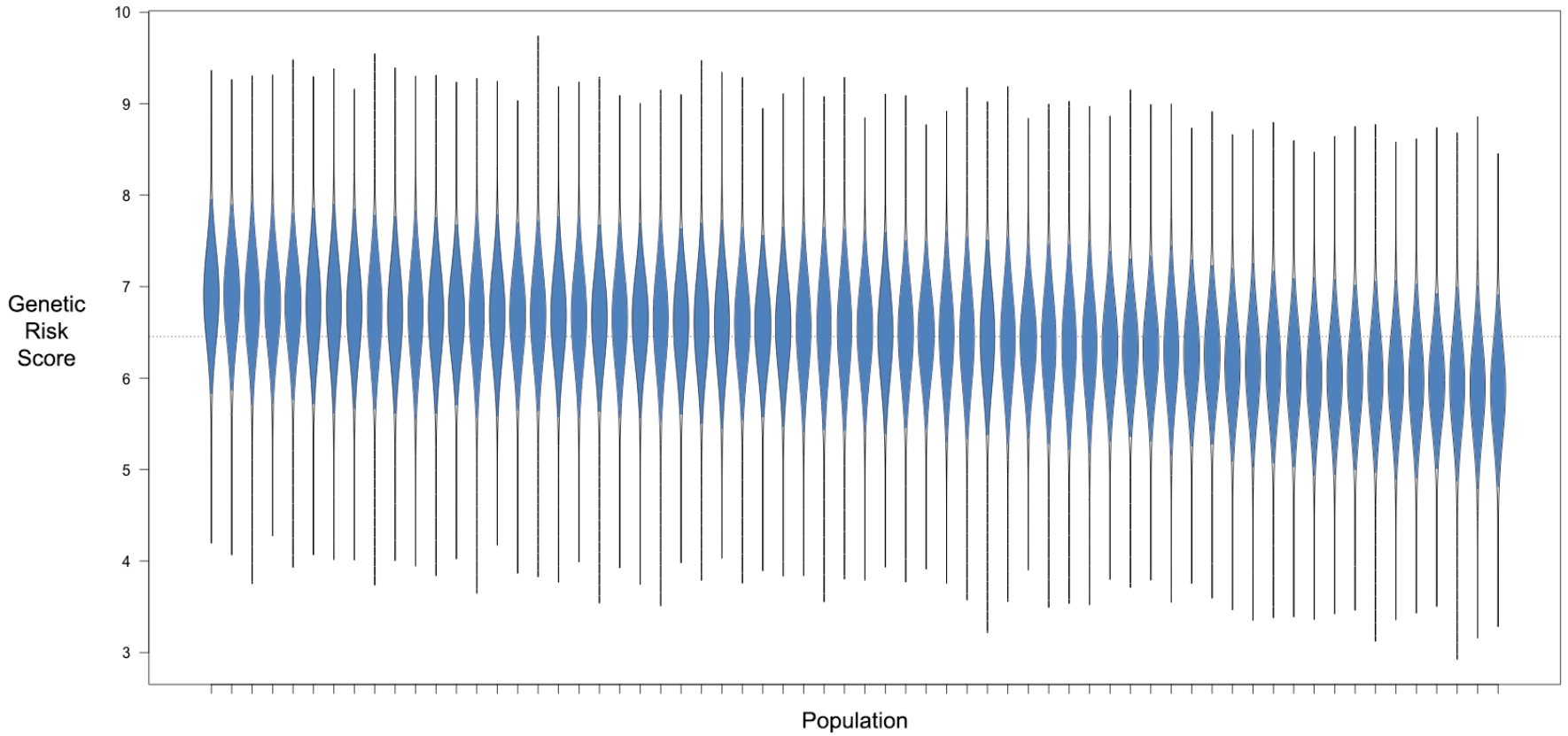
- **Reaction norms** describe the range of phenotypes produced by a genotype in different environment

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



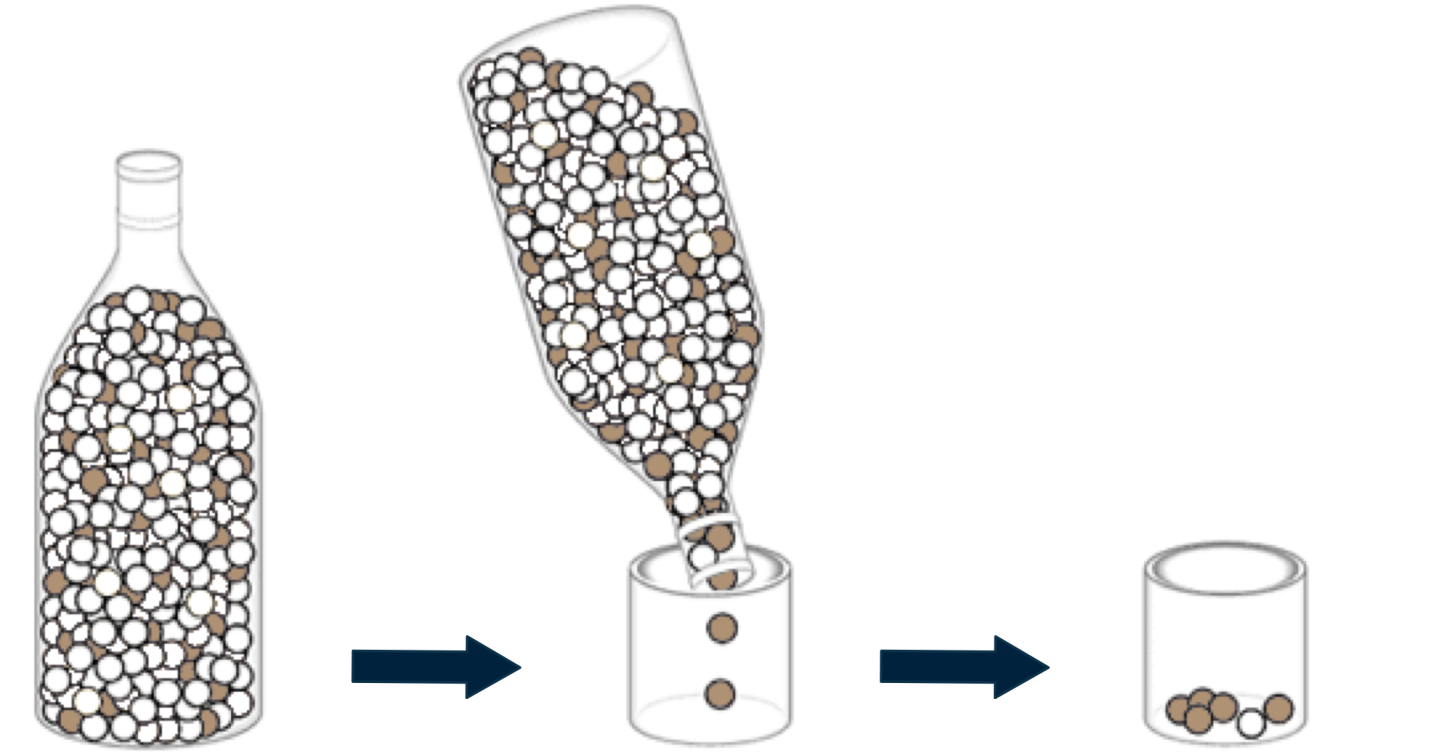
- If a large number of loci contribute to a disease… it is less likely that there will be large differences in genetic risk across populations

Dominance and recessivity

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

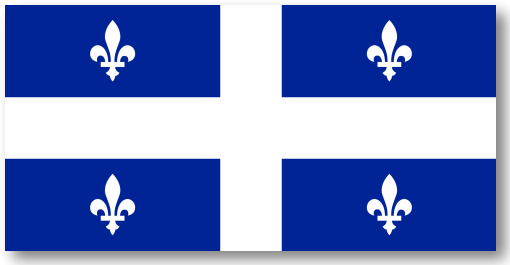
- Small differences in allele frequencies are magnified for recessive diseases

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

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

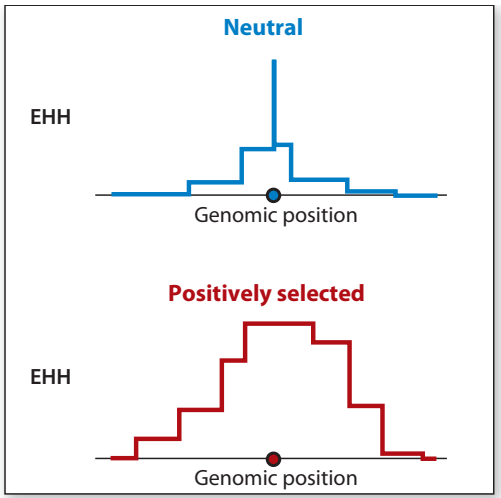
Neutral

	Nonsynonymous	Synonymous
Fixed	4	4
Polymorphic	3	3

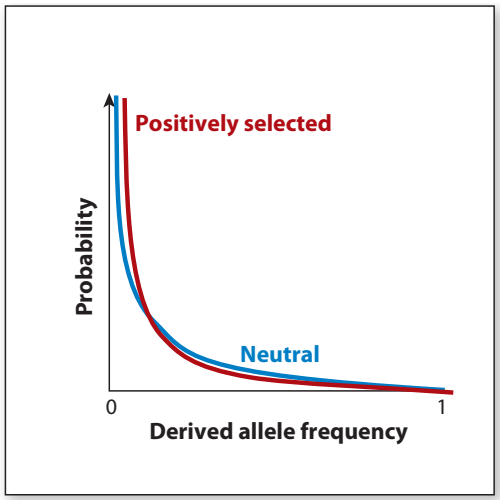
Positively selected

	Nonsynonymous	Synonymous
Fixed	8	4
Polymorphic	3	3

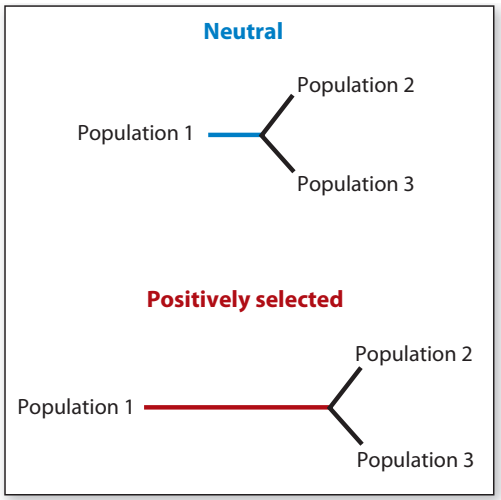
Haplotype statistics



Allele frequencies



Multiple populations

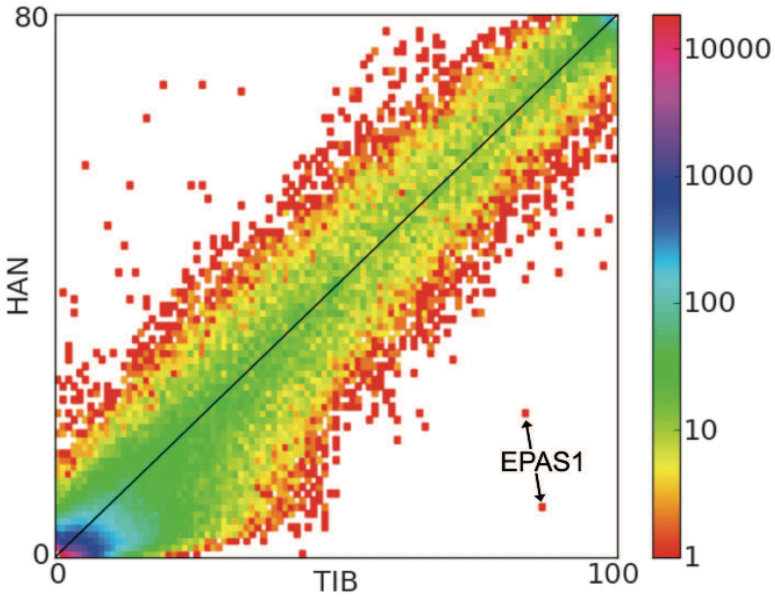


Modified from: Lachance and Tishkoff (2013, *AREES*)

EPAS1 and high-altitude

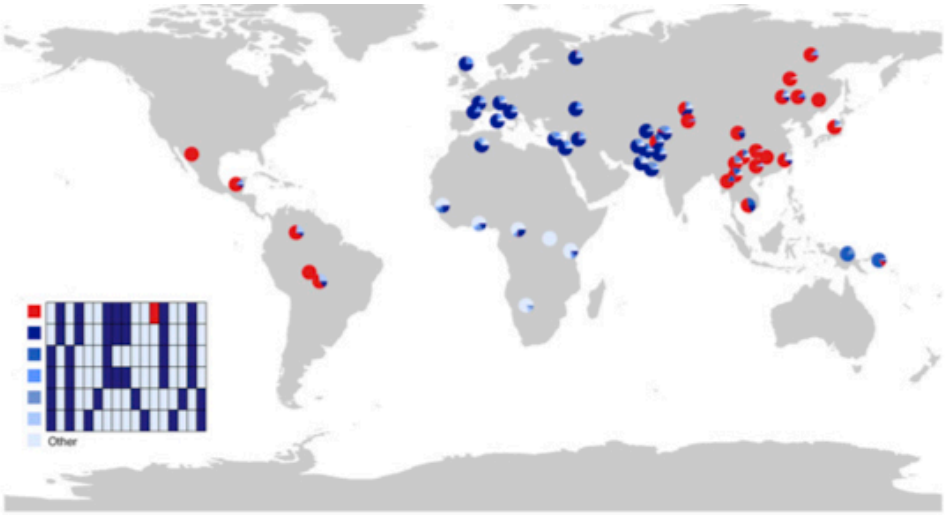
- Reduced $[O_2]$ is a strong selective pressure
- Allele frequencies compared between Tibetans (TIB) and Han Chinese from Beijing (HAN)
- Outlier SNPs are located near *EPAS1*, a hypoxia-induced transcription factor
- The Tibetan EPAS1 haplotype comes from Denisovans (Huerta-Sanchez et al. 2014)**!!!**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

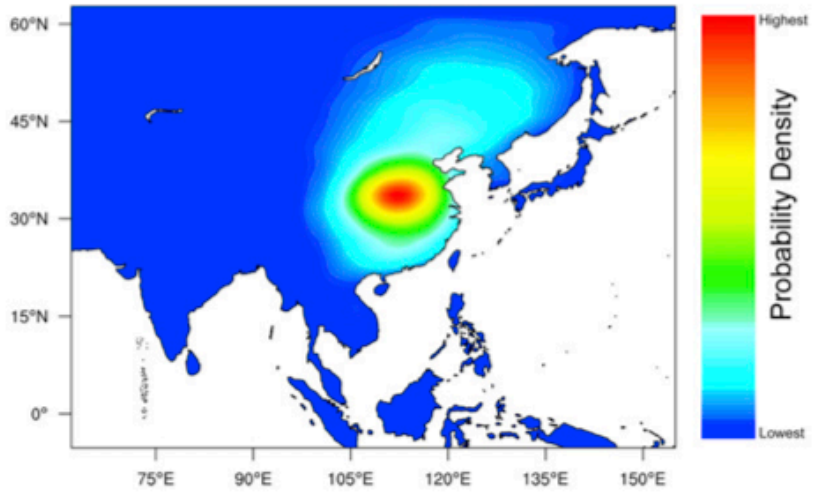
Image rights: EasyTourChina



Yi et al. (2010, *Science*)

EDAR and eccrine glands

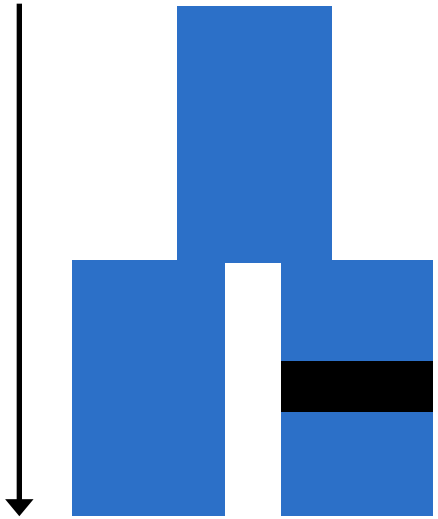


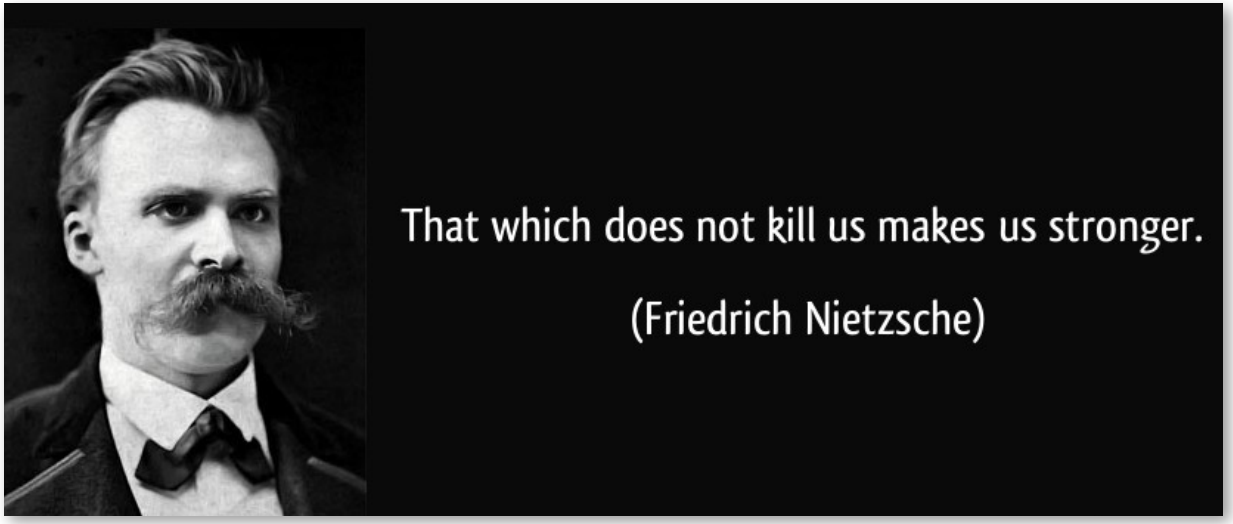


- CMS scans reveal that the *EDAR V370A* allele is a target of selection
- *EDAR* encodes the Ectodysplasin receptor
- Relevant phenotypes in humans and mice
  - Increased hair thickness
  - Increased eccrine (sweat) gland density

Kamberov et al. (2013, *Cell*)

The benefits of a challenging past

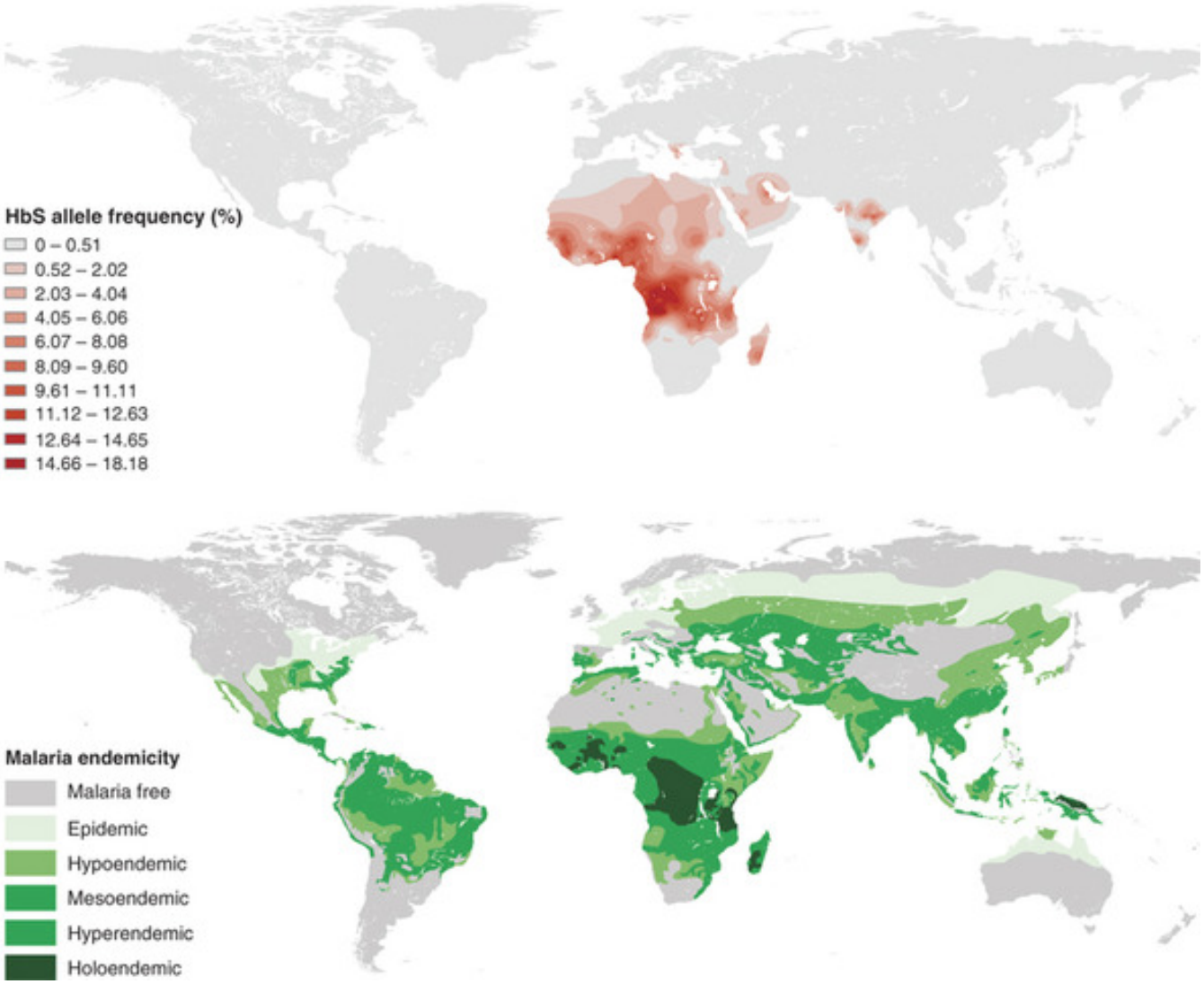




- Multiple mechanisms
  - Positive selection increases the frequency of protective alleles
  - Negative selection decreases the frequency of risk alleles
  - High environmental risks can coincide with lower genetic risks

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

Trade-offs



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

The thrifty gene hypothesis

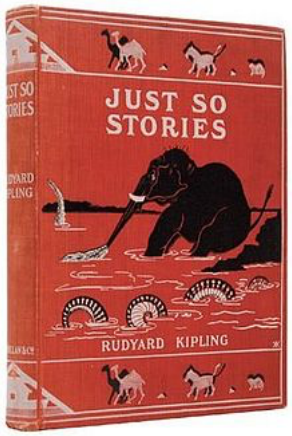
- Type 1 diabetes (T1D)
  - Early onset and insulin deficiency
- Type 2 diabetes (T2D)
  - Adult onset and insulin resistance

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



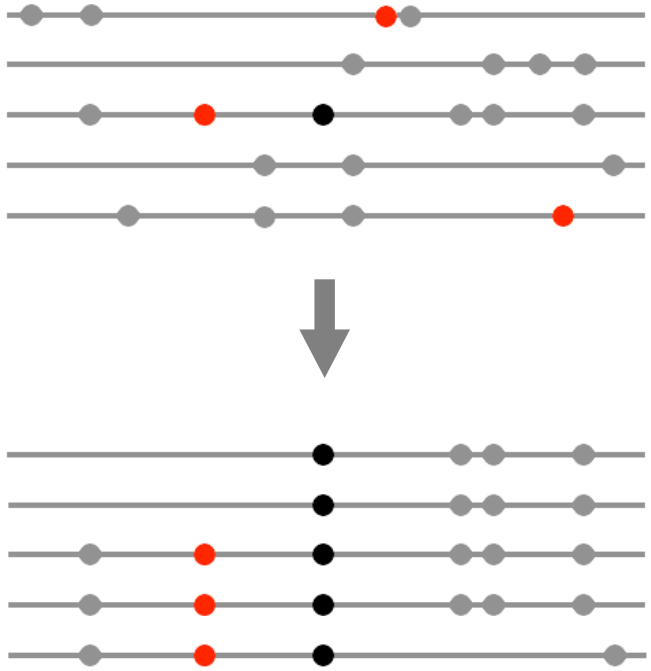
- It is a little too easy to make up stories of adaptive evolution
- Be careful when identifying traits that have been under selection in the past
- Allele surfing and gene conversion can mimic signatures of positive selection
- Convincing narratives of selection can be made for random sets of loci (Pavlidis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

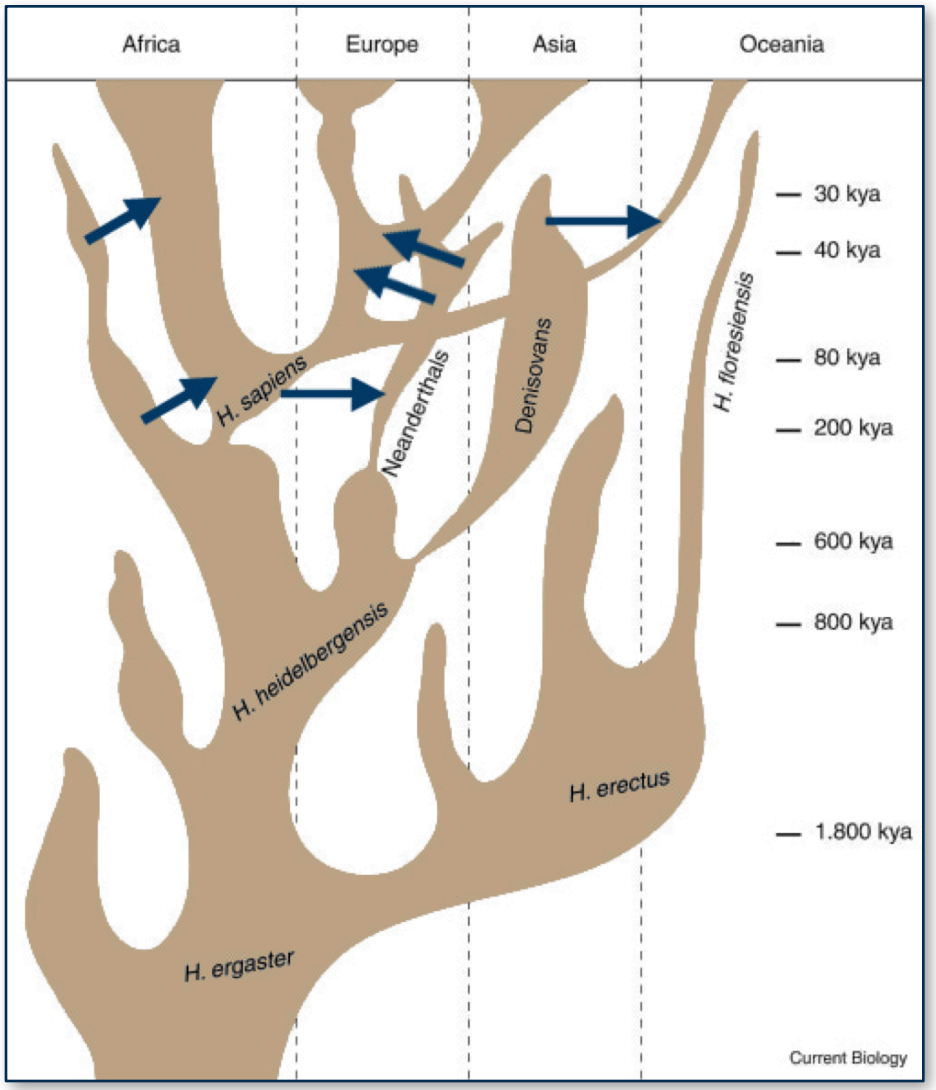


Figure modified from Lalueza-Fox and Gilbert (2011, *Current Biology*)

Introgression of disease and resistance alleles

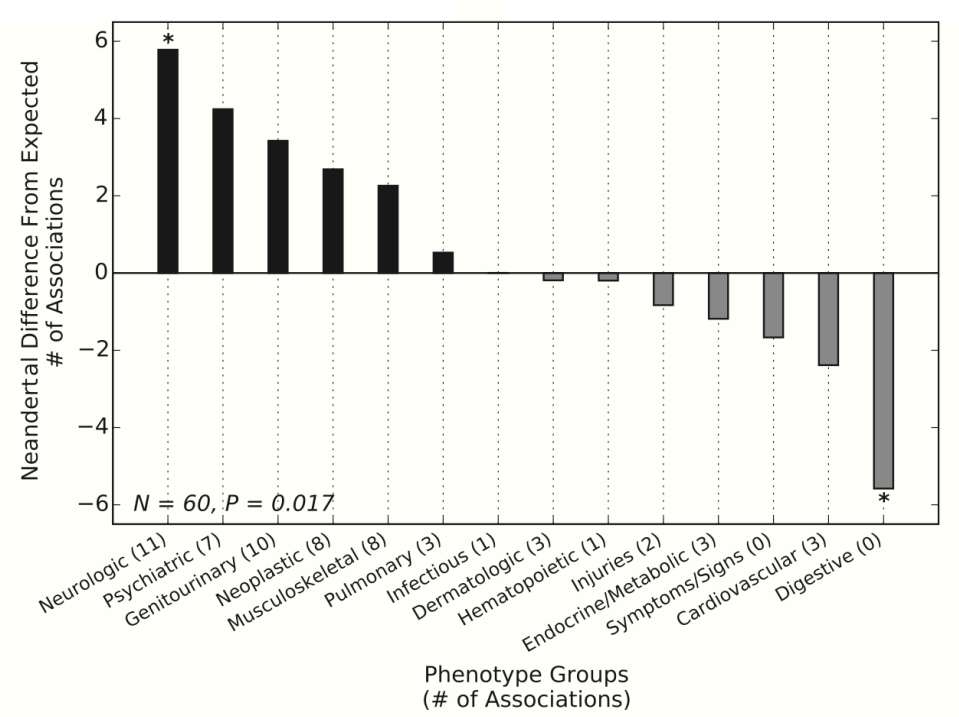


Figure from Simonti et al (2016, *Science*)

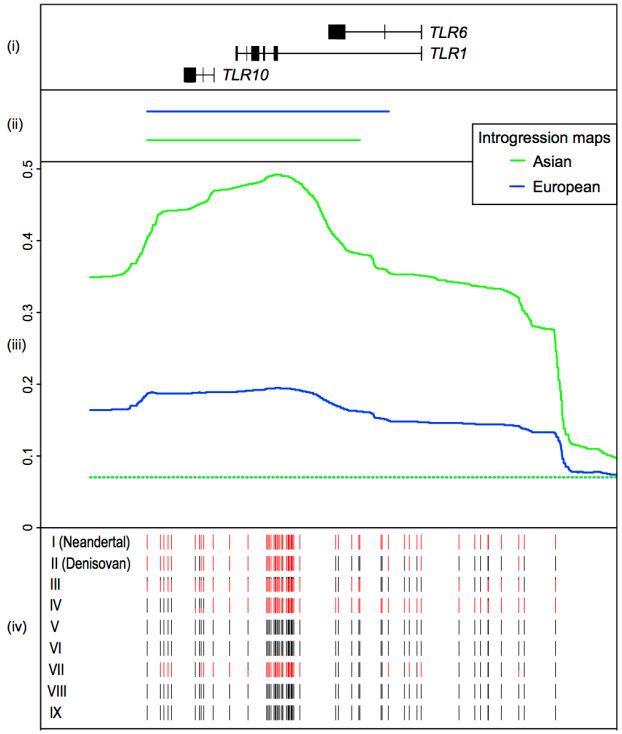


Figure from Danneman et al (2016, *AJHG*)

- 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

Allele surfing