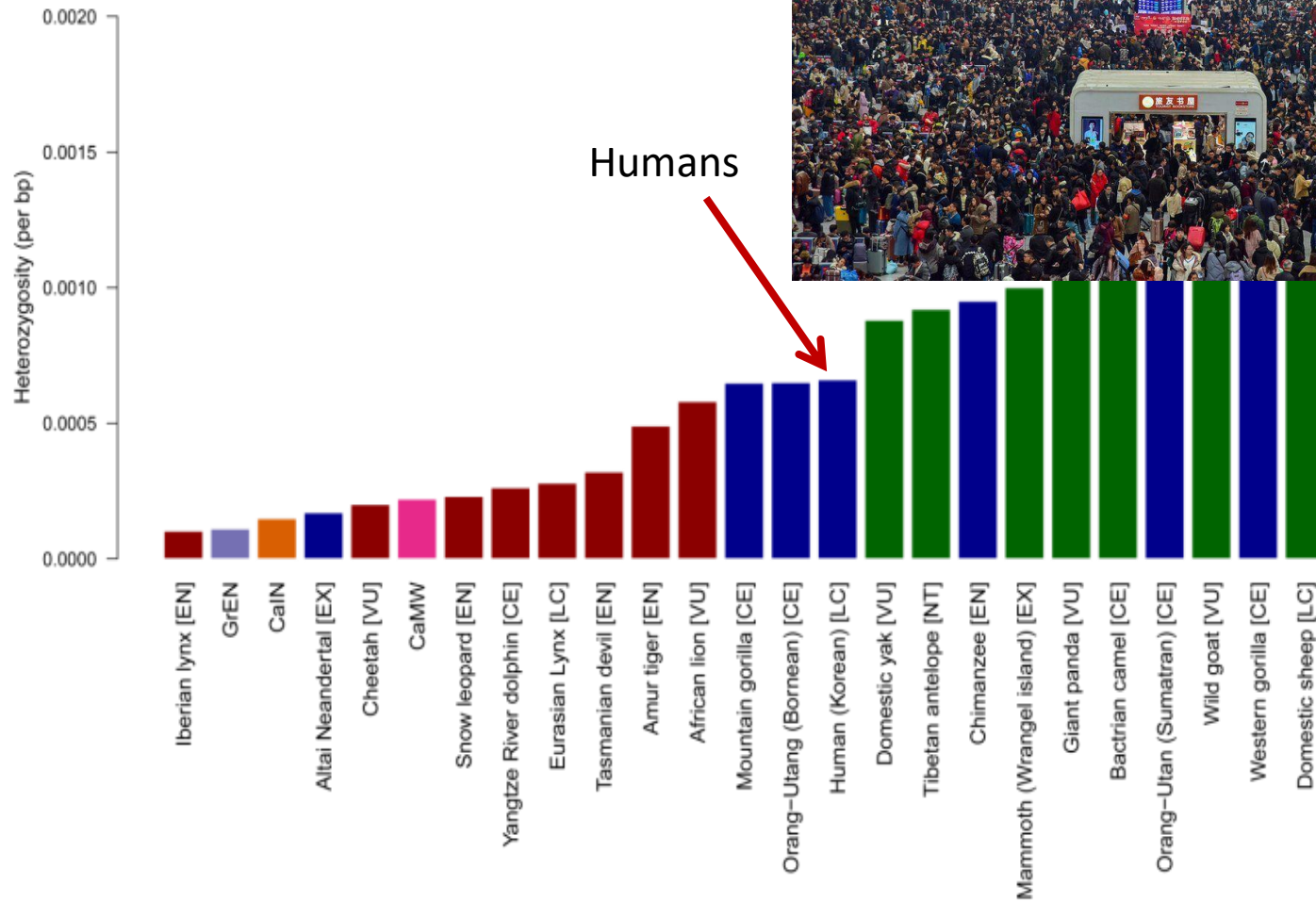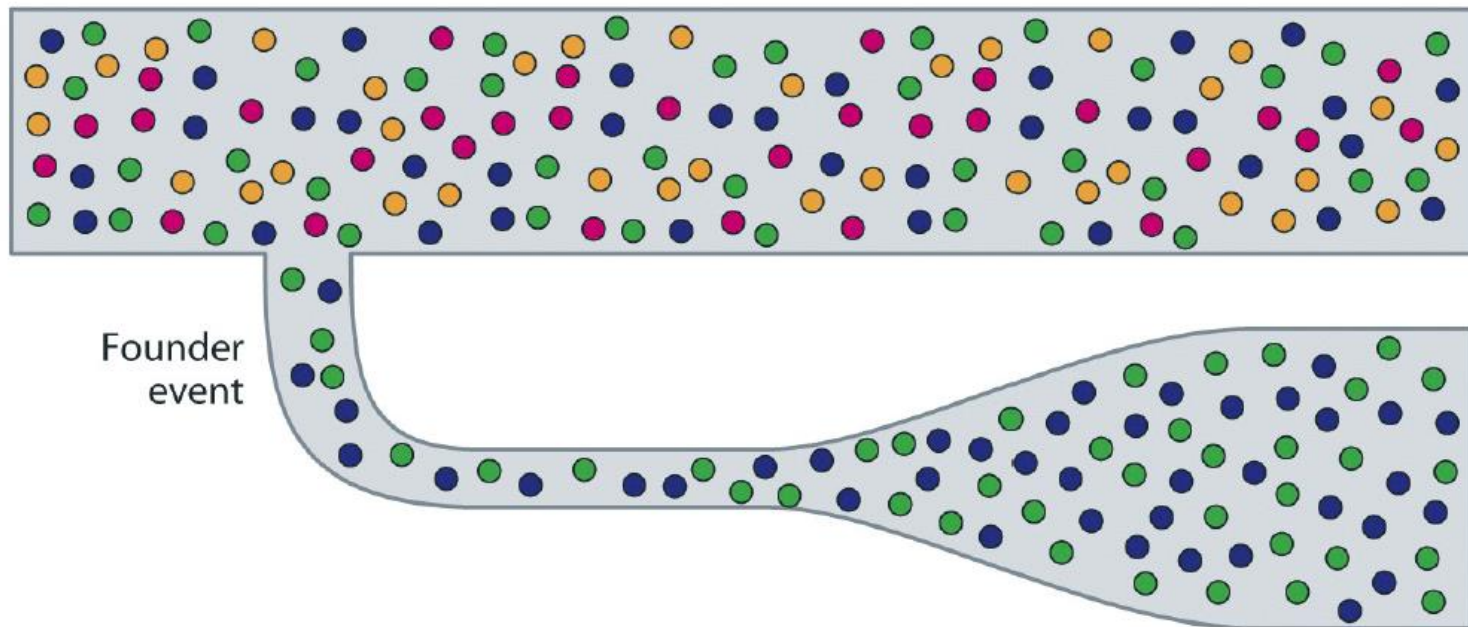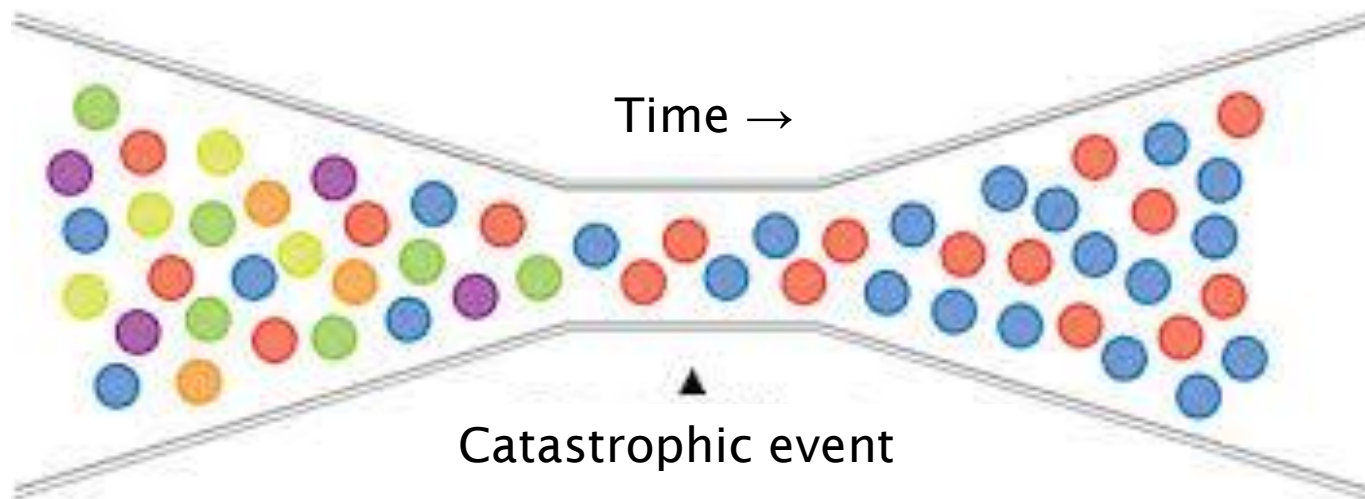# Genetic diversity and why it matters
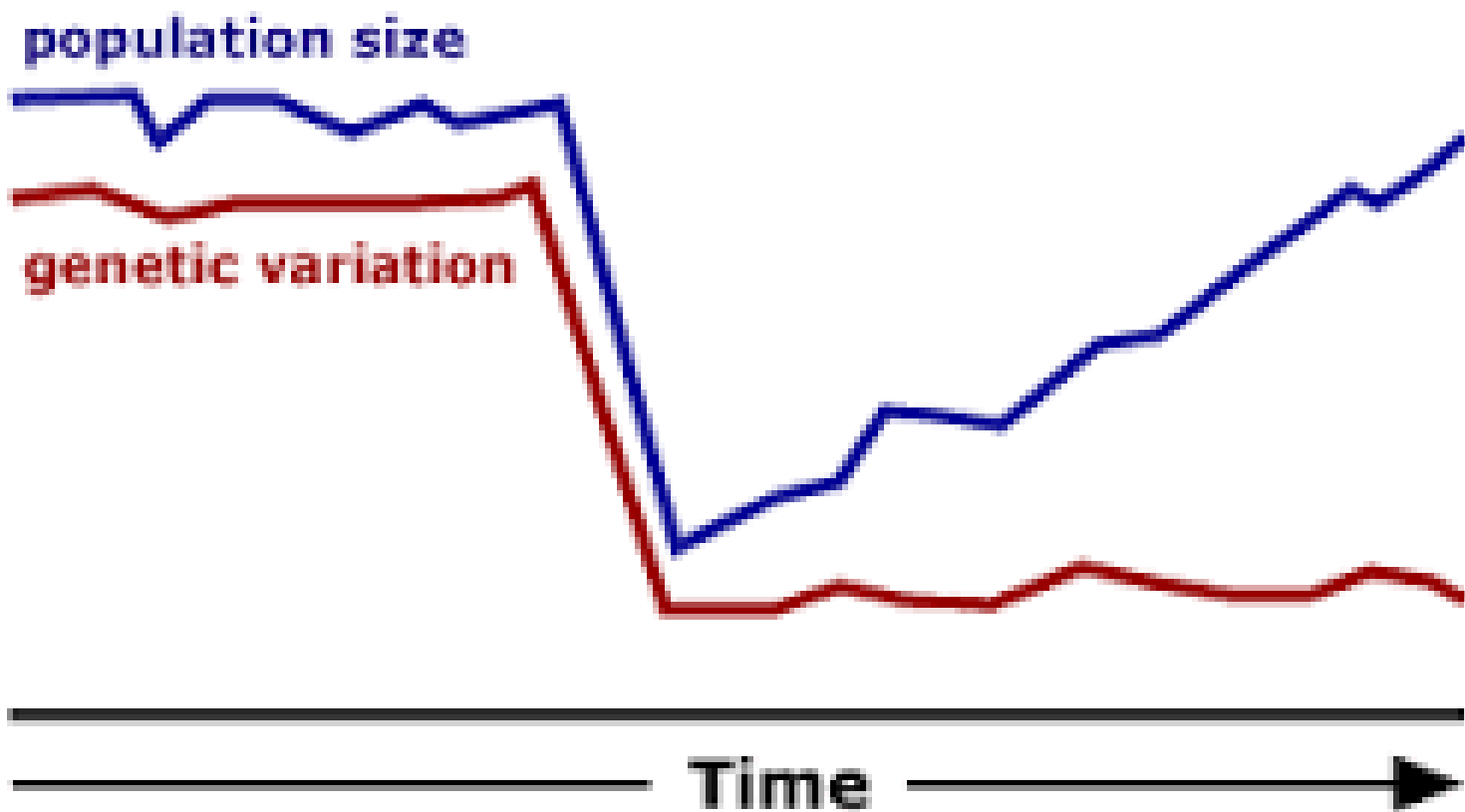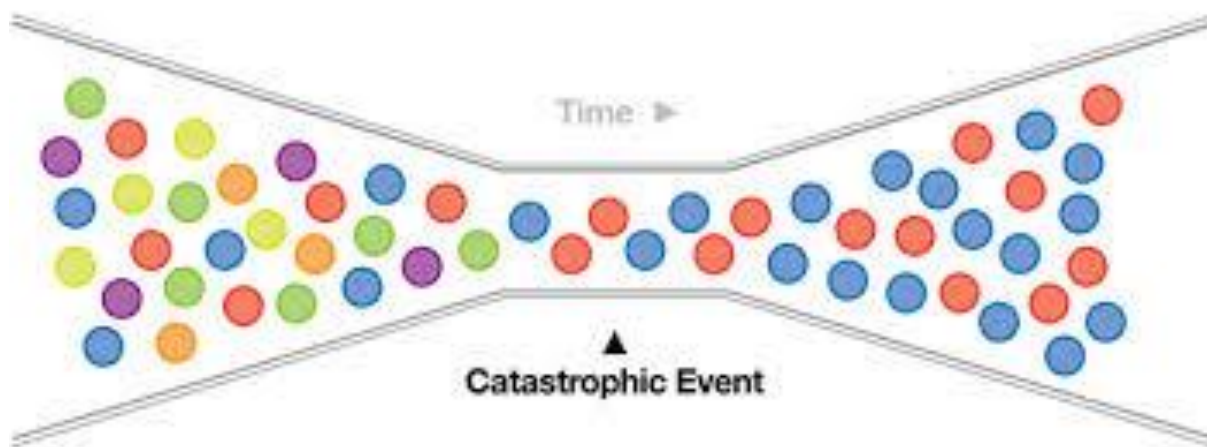
Day 1, Open Institute workshop, Kilifi, August 7th 2024

Assoc. Prof. Rasmus Heller, UCPH

Humans

Why do humans have low genetic diversity?

Time →

Catastrophic event

Founder
event

Time ▶

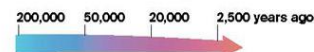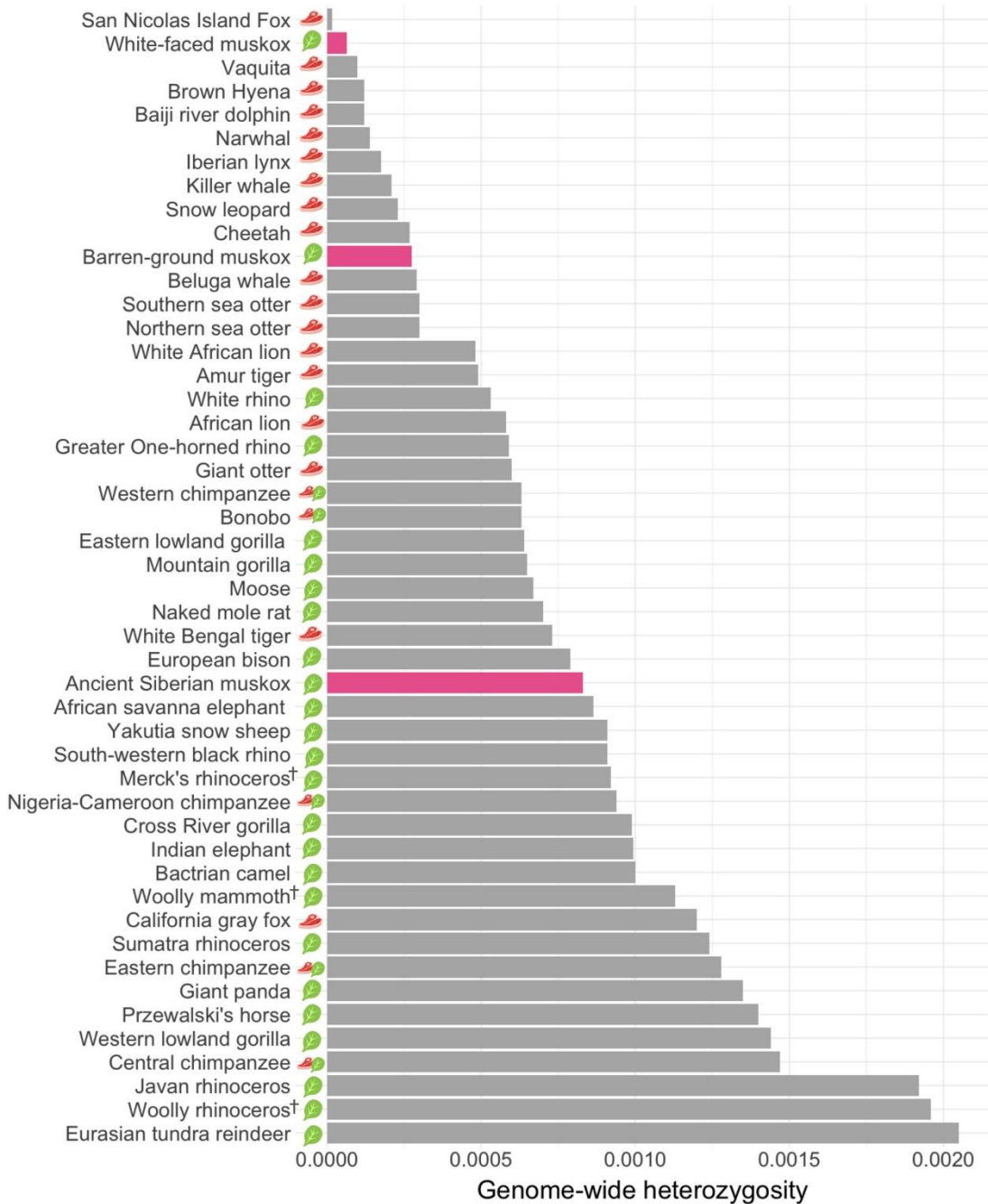Catastrophic Event

population size

genetic variation

Time
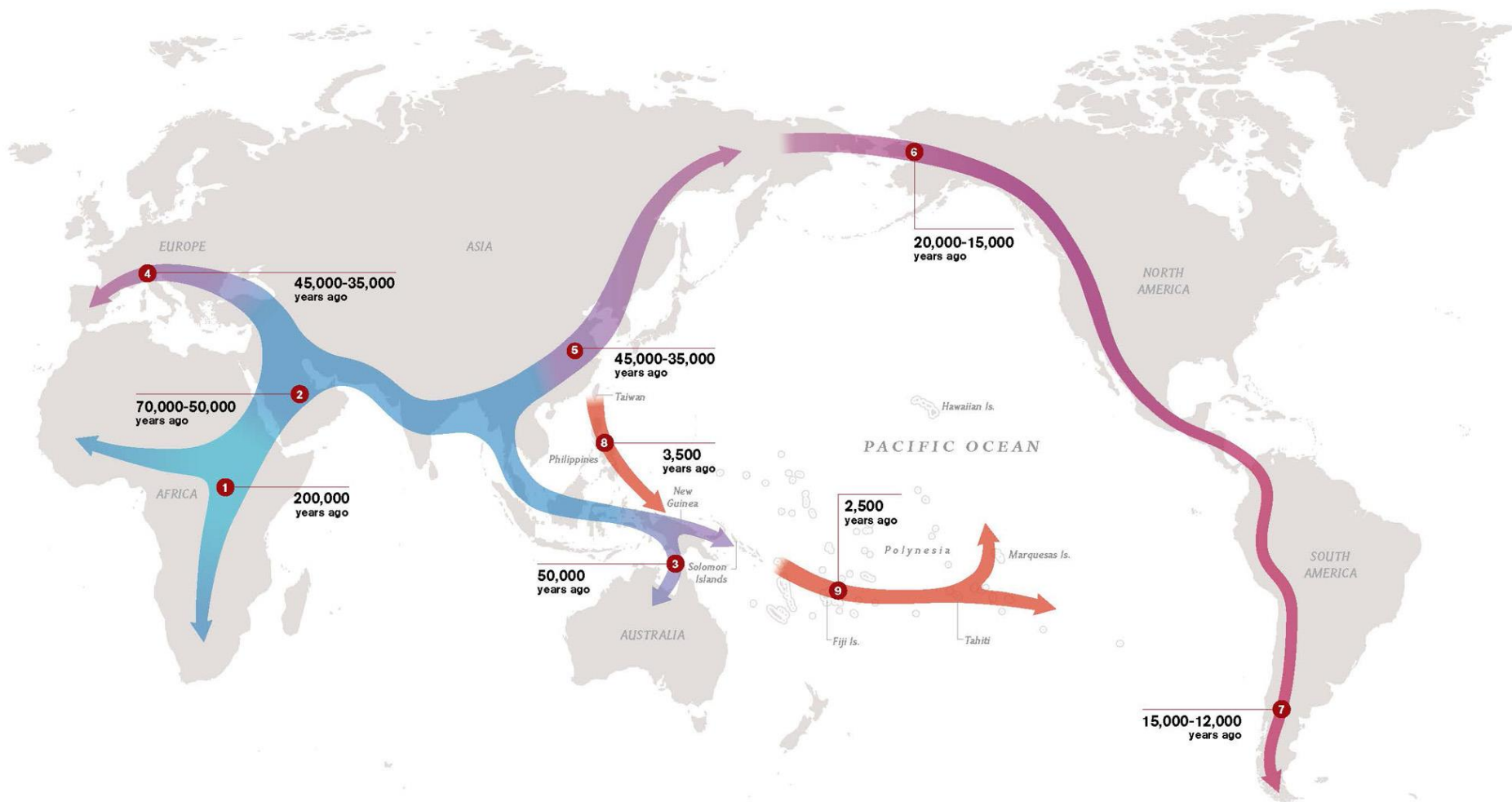
GLOBAL
JOURNEY

Once modern humans began their migration out of Africa some 60,000 years ago, they kept going until they had spread to all corners of the Earth. How far and fast they went depended on climate, the pressures of population, and the invention of boats and other technologies. Less tangible qualities also sped their footsteps: imagination, adaptability, and an innate curiosity about what lay over the next hill.

EUROPE

ASIA

4

45,000-35,000
years ago

2

70,000-50,000
years ago

AFRICA

1

200,000
years ago

5

45,000-35,000
years ago

Taiwan

8

Philippines

3,500
years ago

New
Guinea

3

Solomon
Islands

50,000
years ago

AUSTRALIA

20,000-15,000
years ago

6

NORTH
AMERICA

Hawaiian Is.

PACIFIC OCEAN

2,500
years ago

Polynesia

Marquesas Is.

9

Fiji Is.

Tahiti

SOUTH
AMERICA

15,000-12,000
years ago

7

Generalized route with migration dates

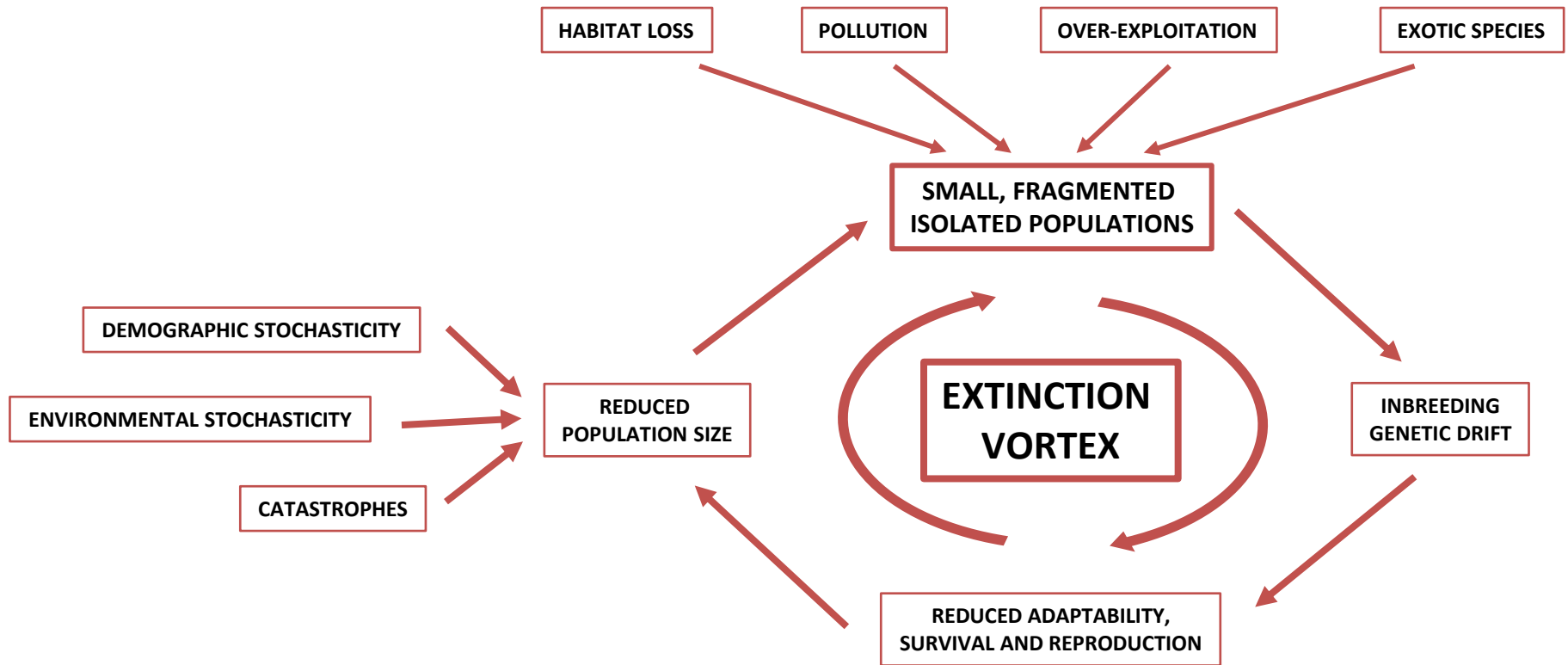200,000   50,000   20,000   2,500 years ago

All species carry the genetic legacy of their long-term history

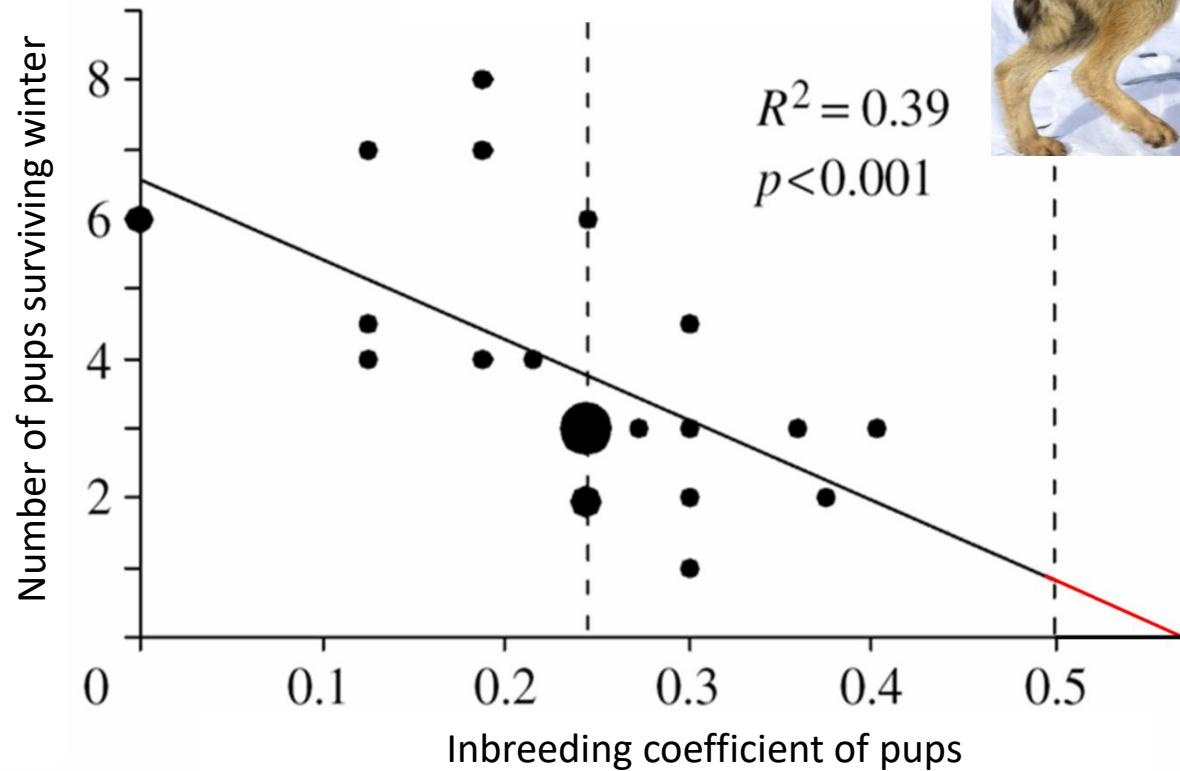# Wildebeest population size history
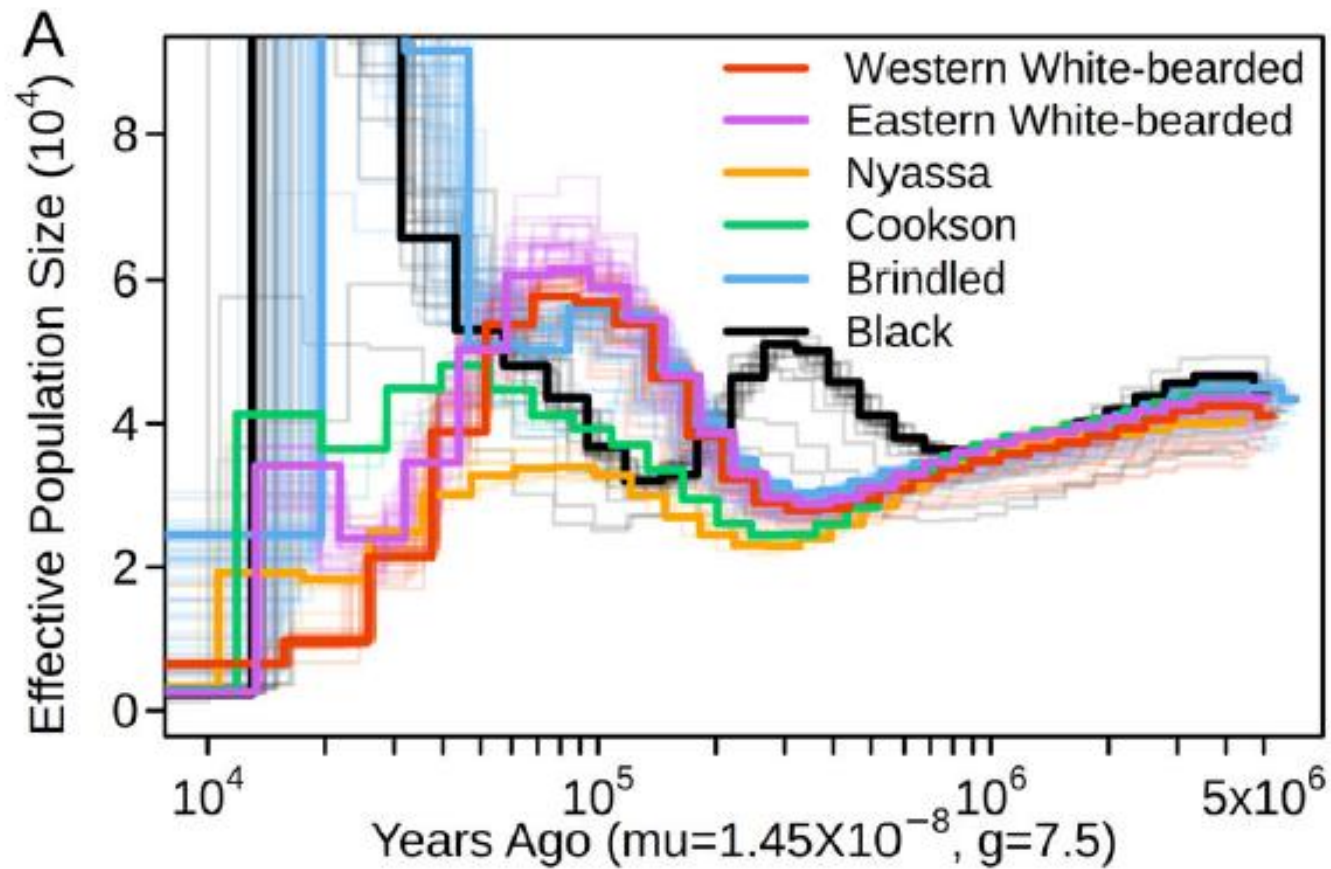
The predicament of small populations

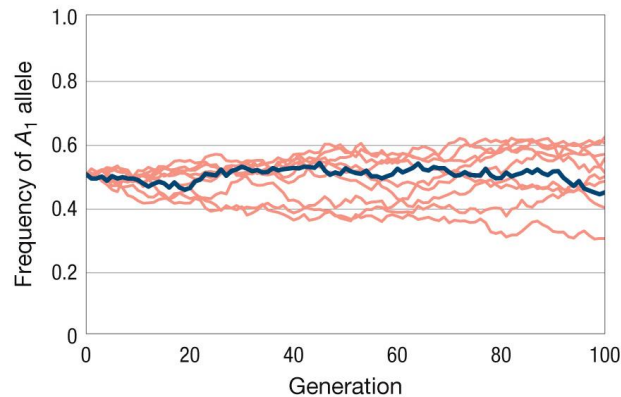Small populations easily get in trouble

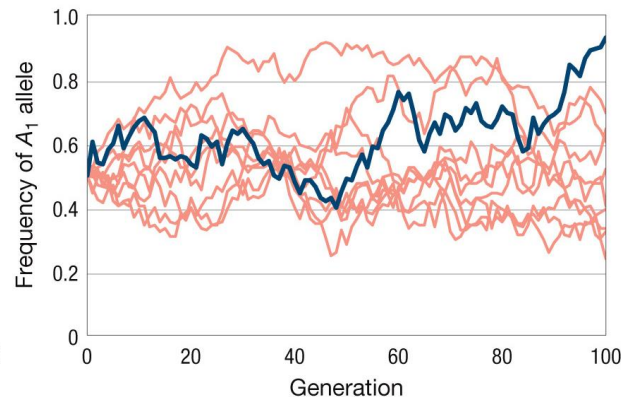# Inbreeding depression

More inbred individuals have lower fitness

Genetic drift in small populations

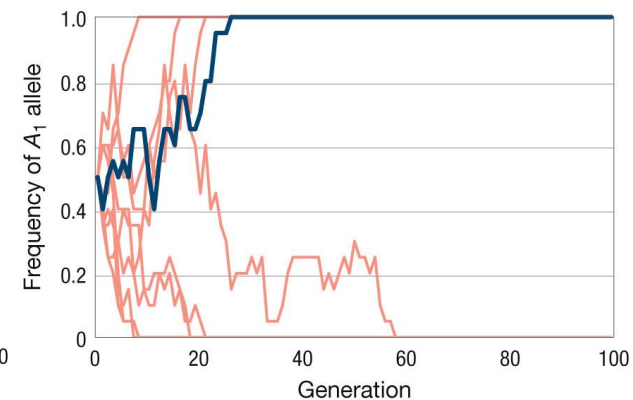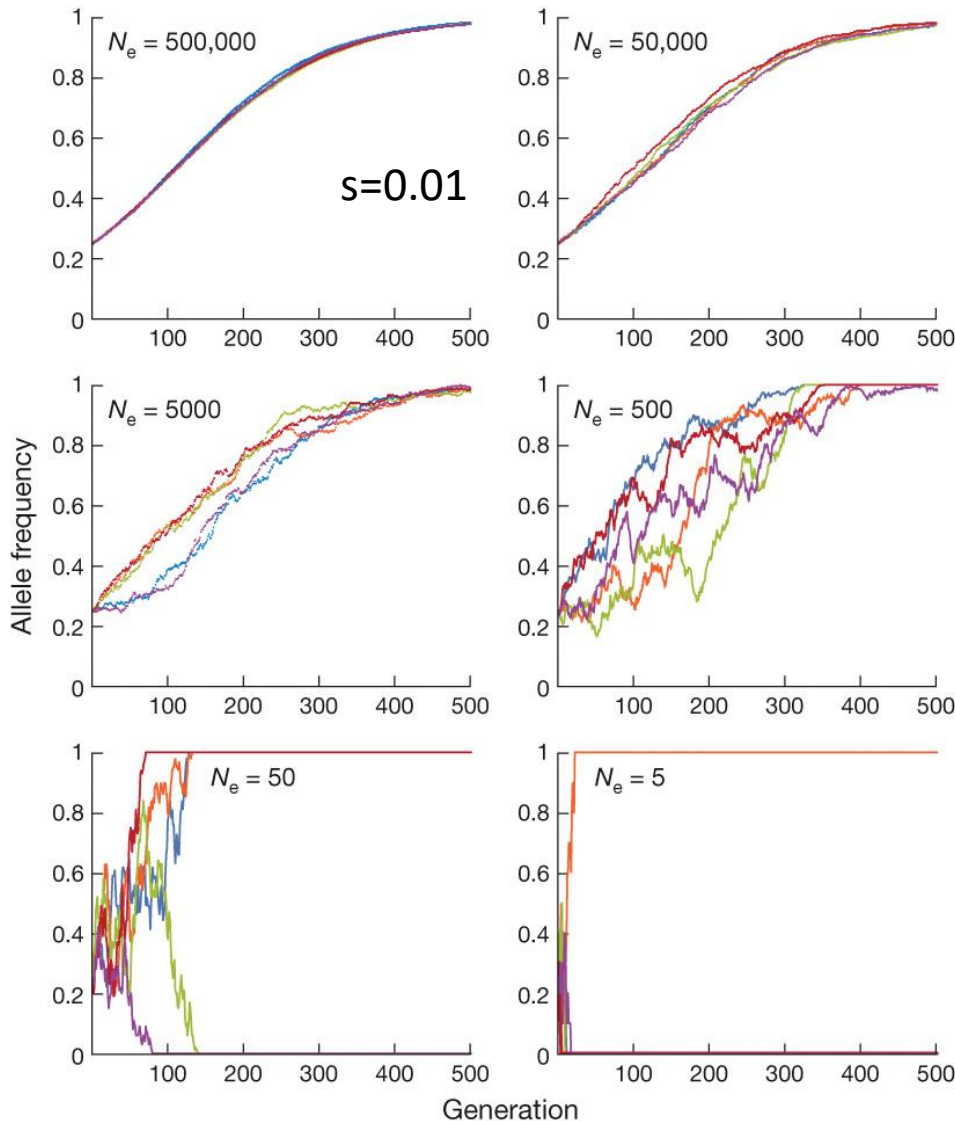Population size 1000

Population size 100

Population size 10

Small populations (e.g. *N* = 10) lose many genetic variants quite quickly. Some of these are beneficial.

Selection is less efficient in small populations

- In large populations, selection always "wins".
- In small populations, drift usually wins.
- Drift makes fixation or loss occur faster.
- Even beneficial mutations often get lost!

Measures of genetic diversity

The following are the most commonly used measures for sequencing data.

- *Heterozygosity*: the proportion of loci that are heterozygous (carry two different alleles) out of all the examined loci.

- *Nucleotide diversity*: the average number of pairwise differences between two or more DNA sequences, taken per site examined.

- *Segregating sites*: the number of sites in which at least one individual has a different allele than the remaining individuals.

Measures of genetic diversity

The following are the most commonly used measures for sequencing data.

- *Heterozygosity*: the proportion of loci that are heterozygous (carry two different alleles) out of all the examined loci.

- *Nucleotide diversity*: the average number of pairwise differences between two or more DNA sequences, taken per site examined.

- *Segregating sites*: the number of sites in which at least one individual has a different allele than the remaining individuals.

Heterozygosity is widely used, and it can also inform us about e.g. inbreeding and genetic load

What is genome-wide heterozygosity?
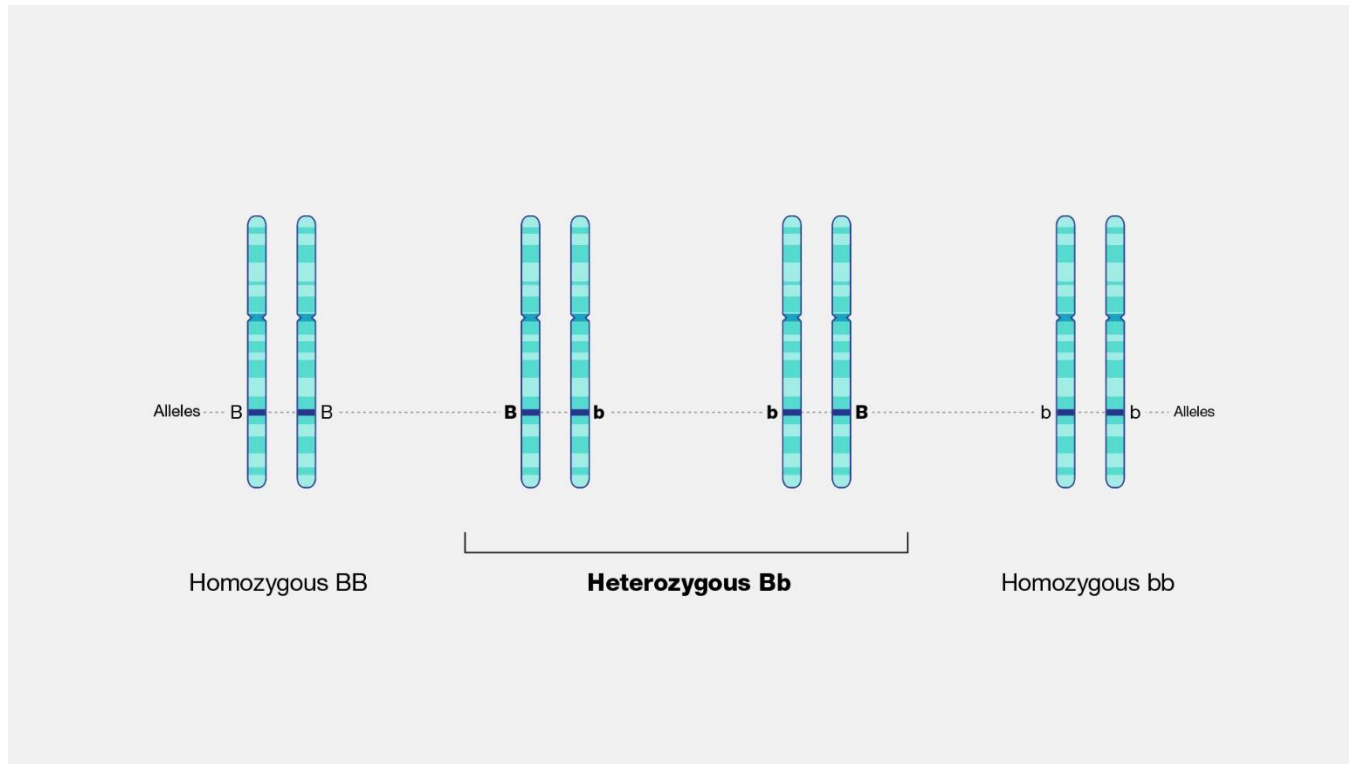
We move along the chromosomes and calculate how often we encounter a heterozygous locus, out of all the loci we encounter.

$$\widehat{Het} = \frac{\#heterozygous\ loci}{\#heterozygous\ loci + \#homozygous\ loci}$$

Does it matter?

● PERSPECTIVE

The inflated significance of neutral genetic diversity in conservation genetics

João C. Teixeira[a,b,1] and Christian D. Huber[a,1]

● PERSPECTIVE

The crucial role of genome-wide genetic variation in conservation

Marty Kardos[a,1], Ellie E. Armstrong[b], Sarah W. Fitzpatrick[c,d,e], Samantha Hauser[f], Philip W. Hedrick[g], Joshua M. Miller[h,i,j], David A. Tallmon[k], and W. Chris Funk[l]

Genetic diversity tells us many things

Map of global human genetic diversity

Genetic Variability

- 1.3 to 1.4 **LOWER**
- 1.4 to 1.5
- 1.5 to 1.6
- 1.6 to 1.7
- 1.7 to 1.8
- 1.8 to 1.9
- 1.9 to 2.0 **HIGHER**

Data Source: Samuels, David C et al. (Genetics - 2016) reanalysis of the 1000 Genomes Project dataset