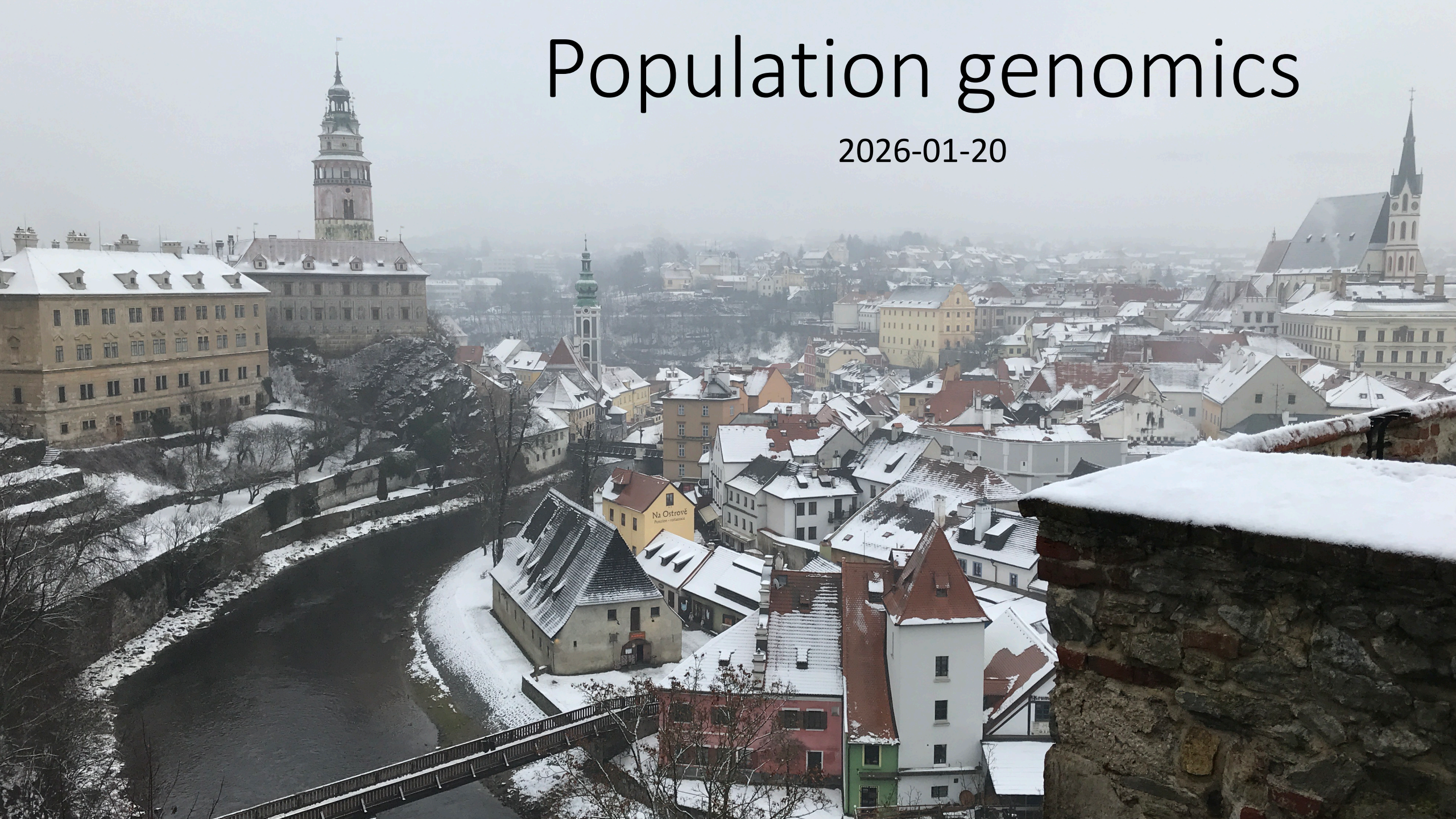
Population genomics

2026-01-20



UNIVERSITY OF OREGON

UPPSALA UNIVERSITET

VOGELWARTE.CH

2011 2018 2023

Bachelor/Masters

PhD

Postdoc










Outline for this evening

- Introduction to population genomics (~30 mins)

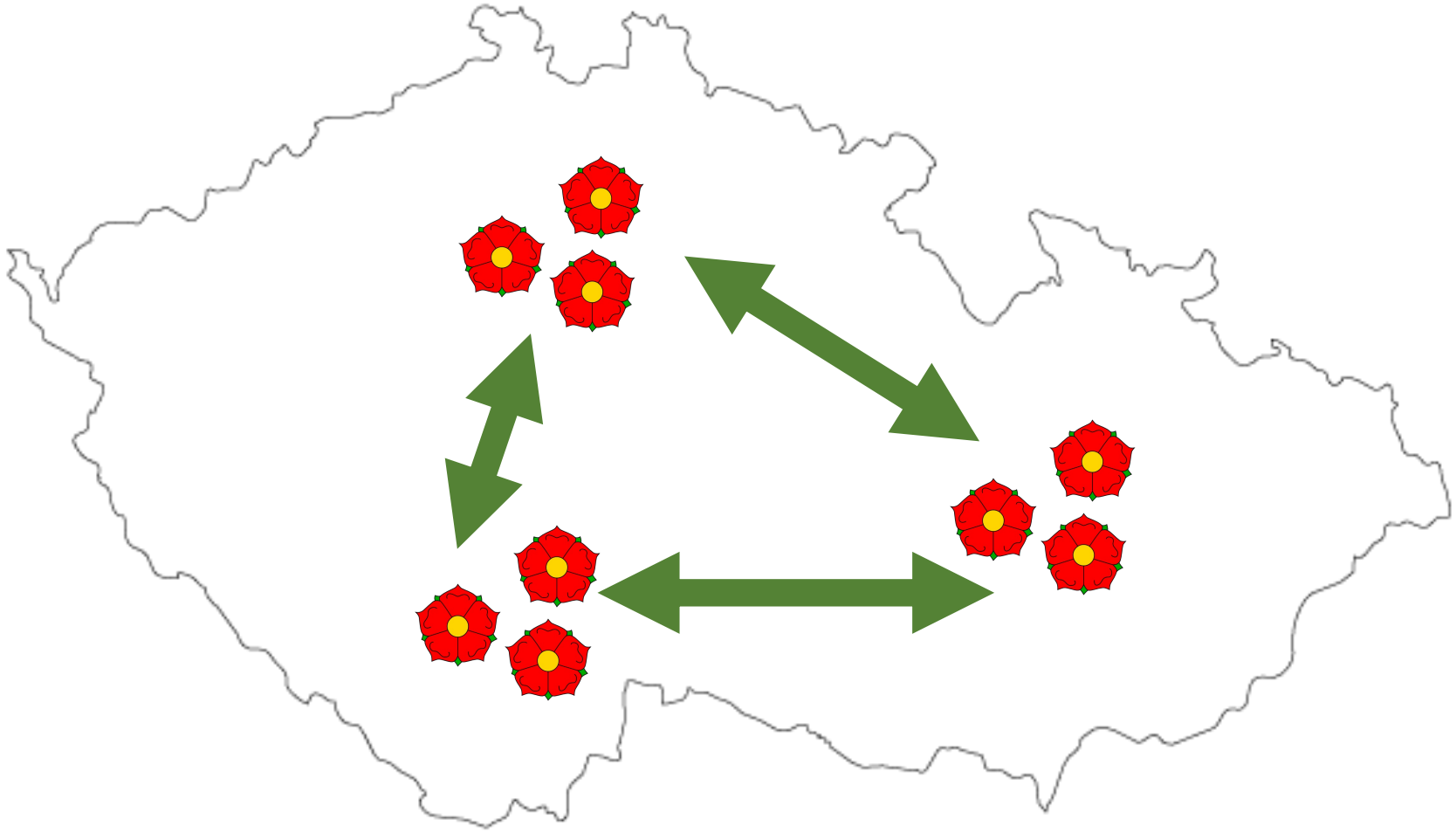
- Practical exercises
  - Part 1: Simulation based (~1 hour)

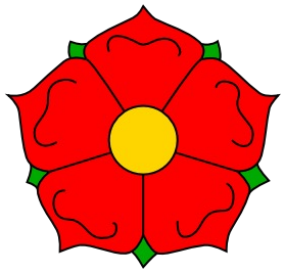
  - Part 2: Real world example (~1 hour)

What is population genomics?

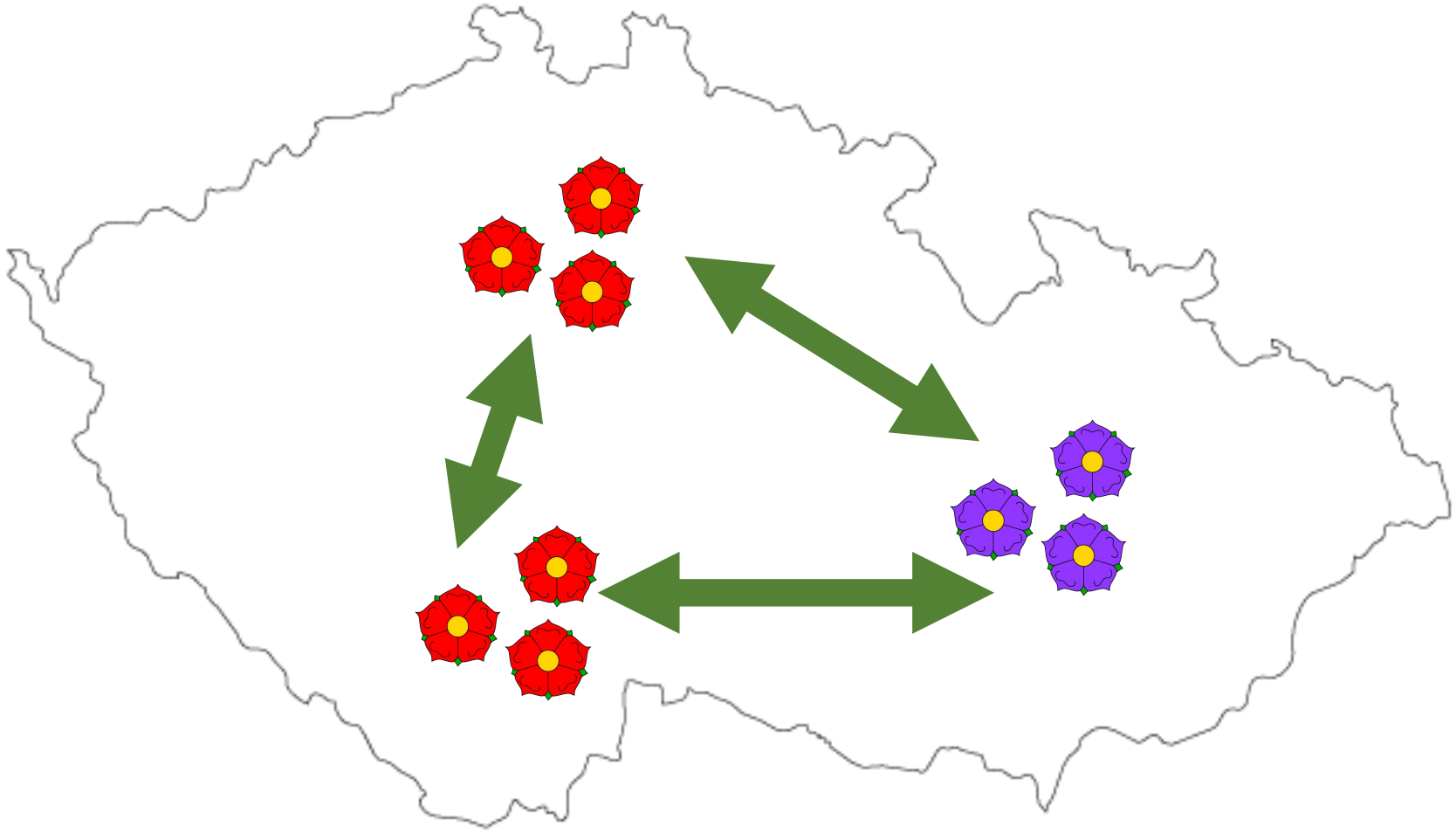
The study of how genomic variation is partitioned within and between populations, and the corresponding evolutionary causes and consequences

What is the genetic diversity of my species?

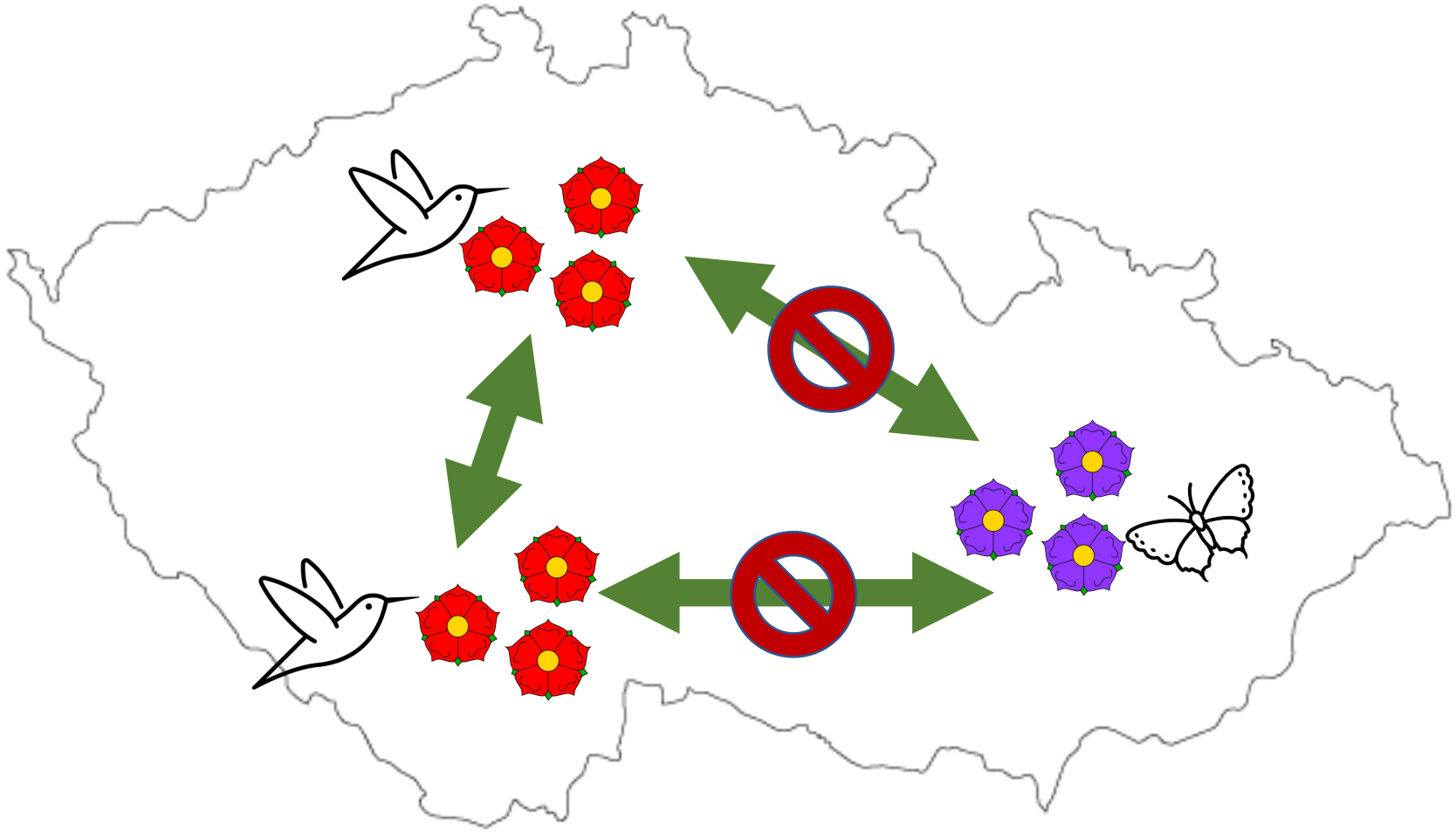




How much gene flow is there between populations?



Are differences between populations due to selection?



Do trait differences restrict gene flow?

How can we address these questions?

Sequence some genomes!

The genome contains the signature of a population's evolutionary history – we just need to decipher it

Evolutionary processes that shape diversity

Mutation

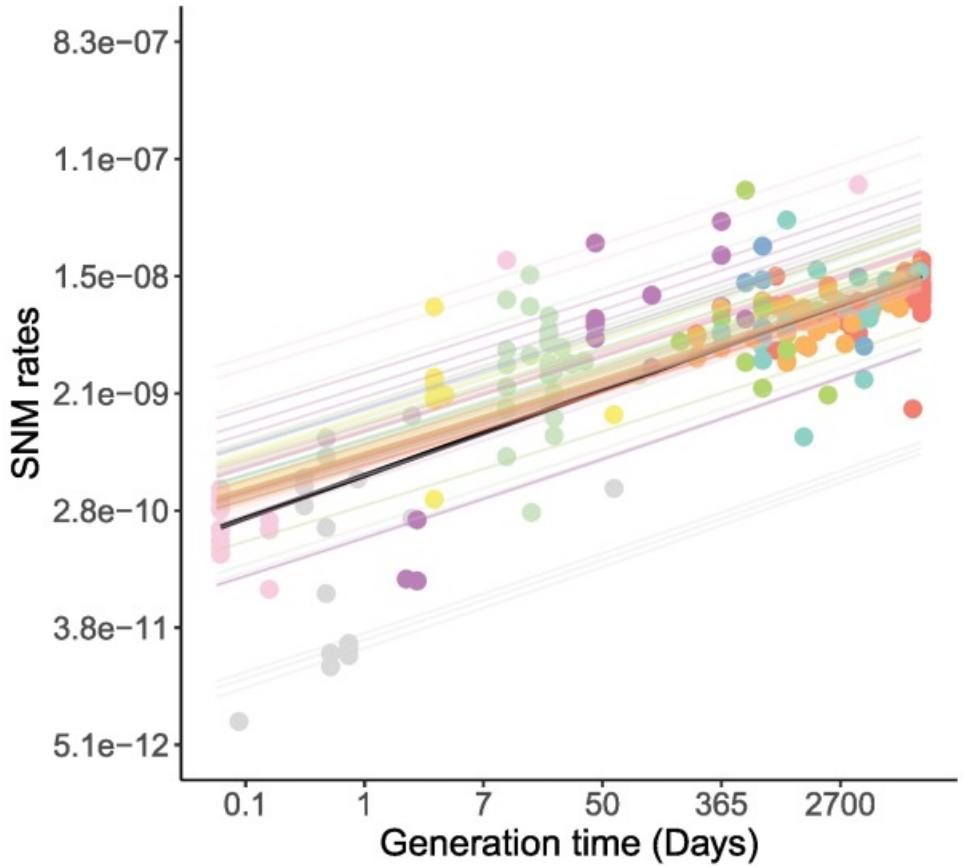
Genetic drift

Recombination

Natural selection

Migration

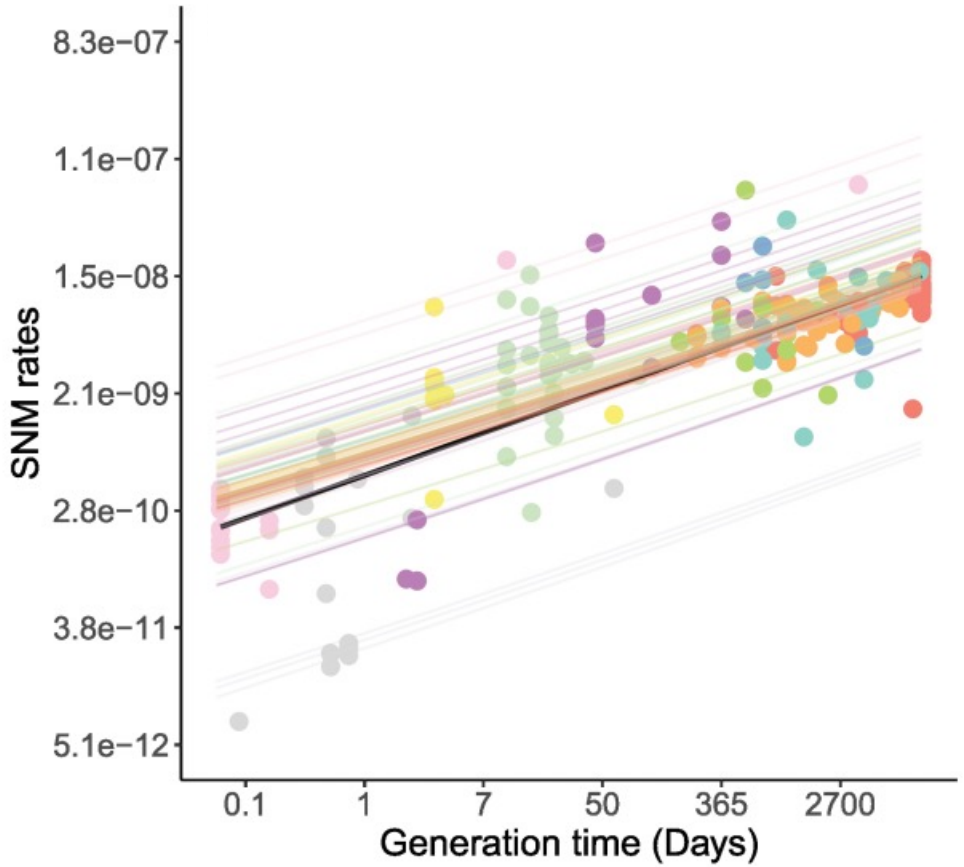
Mutation



Mutation is the original source of genetic variation

Wang & Obbard, 2023 Evo. Lett.

Mutation

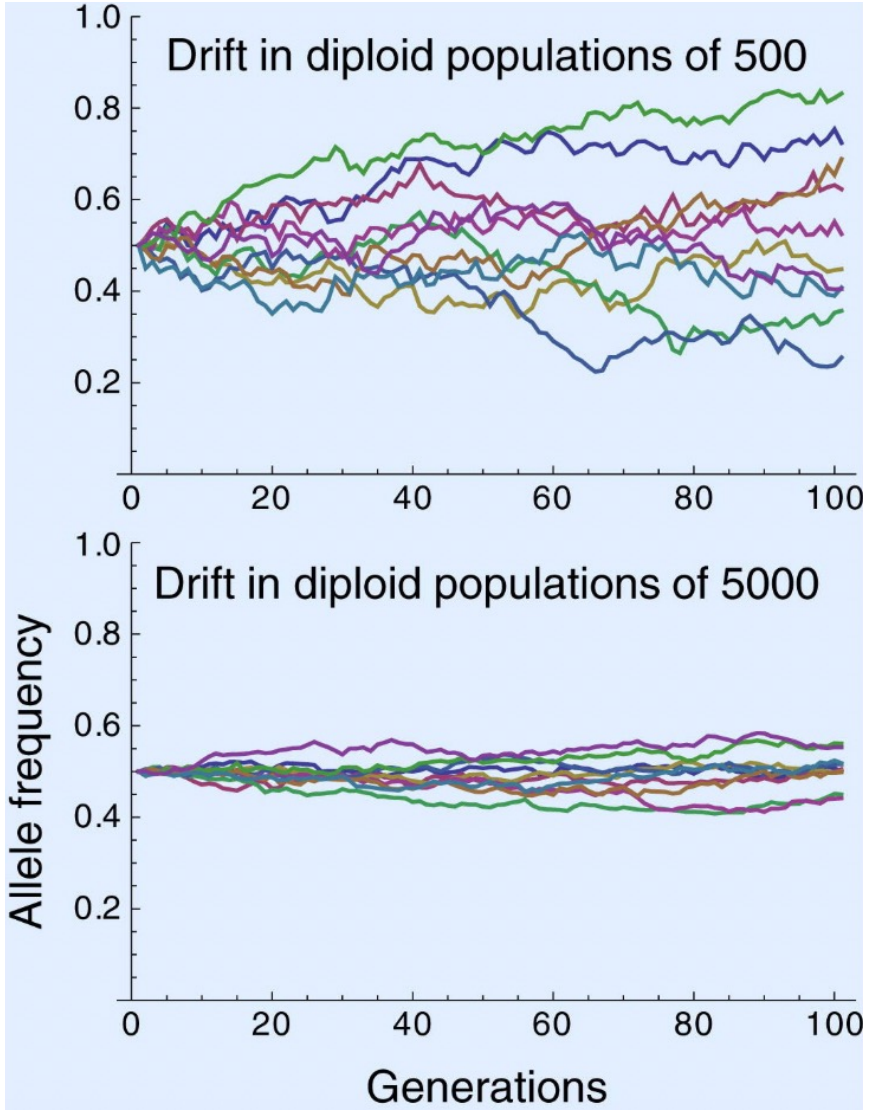


Mutation is the original source of genetic variation

Without imperfection in DNA repair machinery we wouldn't be here – variation is the substrate of evolution!

Wang & Obbard, 2023 Evo. Lett.

Genetic drift



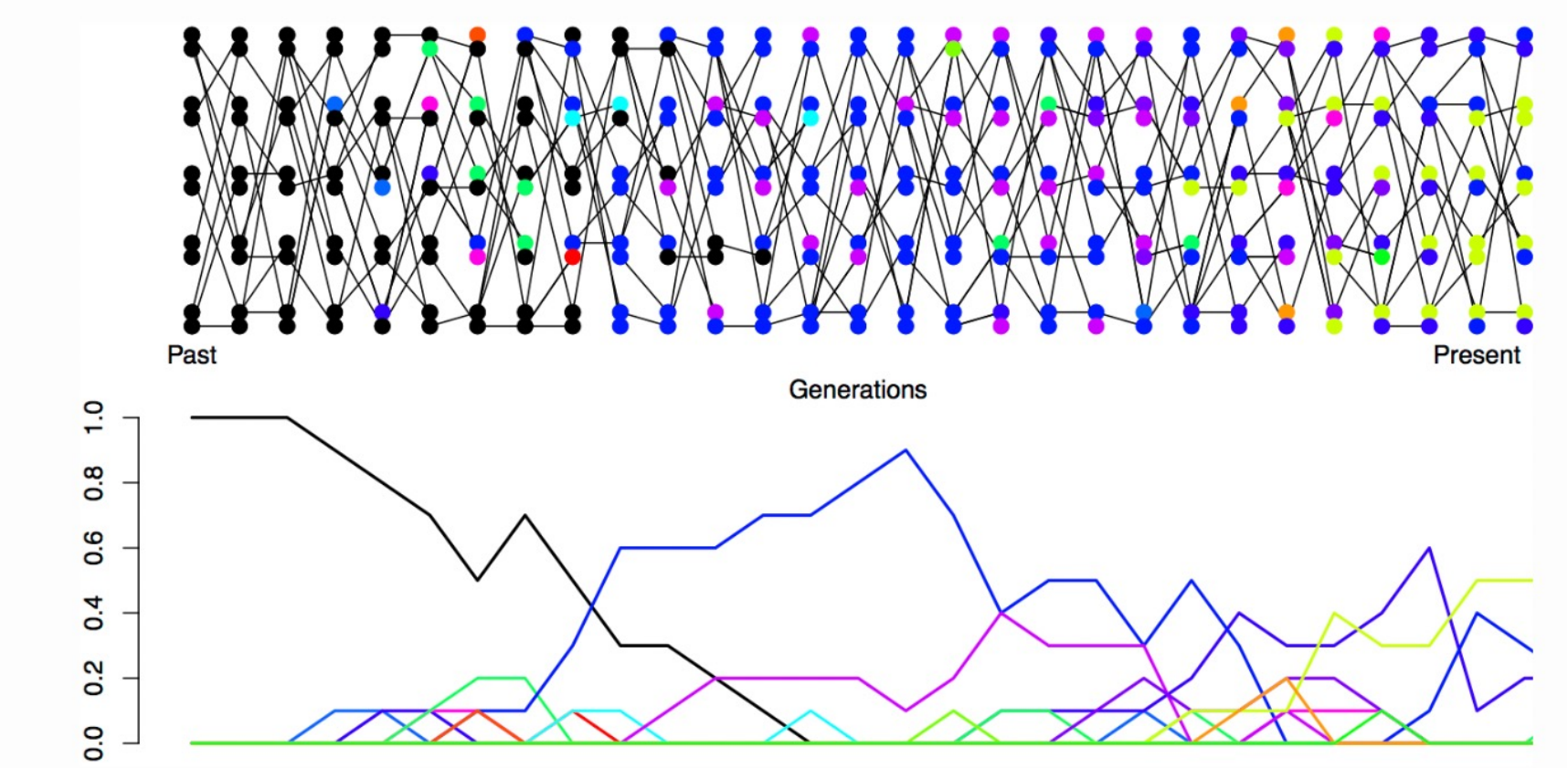
Drift leads to random fluctuations in allele frequencies

Ultimate fate of alleles is loss/fixation due to drift

Probability of fixation for a neutral mutation = 1/2N

Masel 2011

Neutral diversity is a balance between drift and mutation

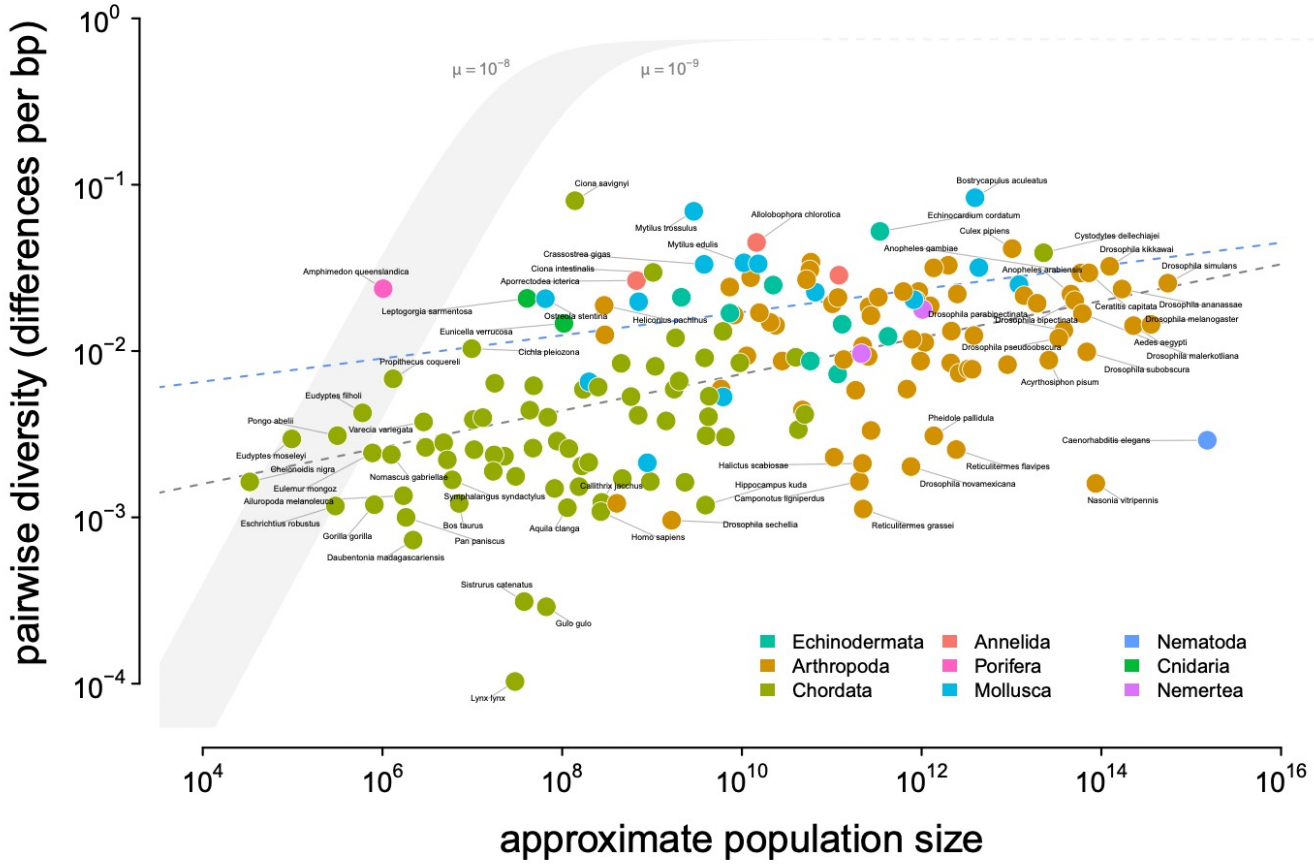


Coop 2019

Population scaled mutation rate = $4N\mu$

Effective population size: N_e

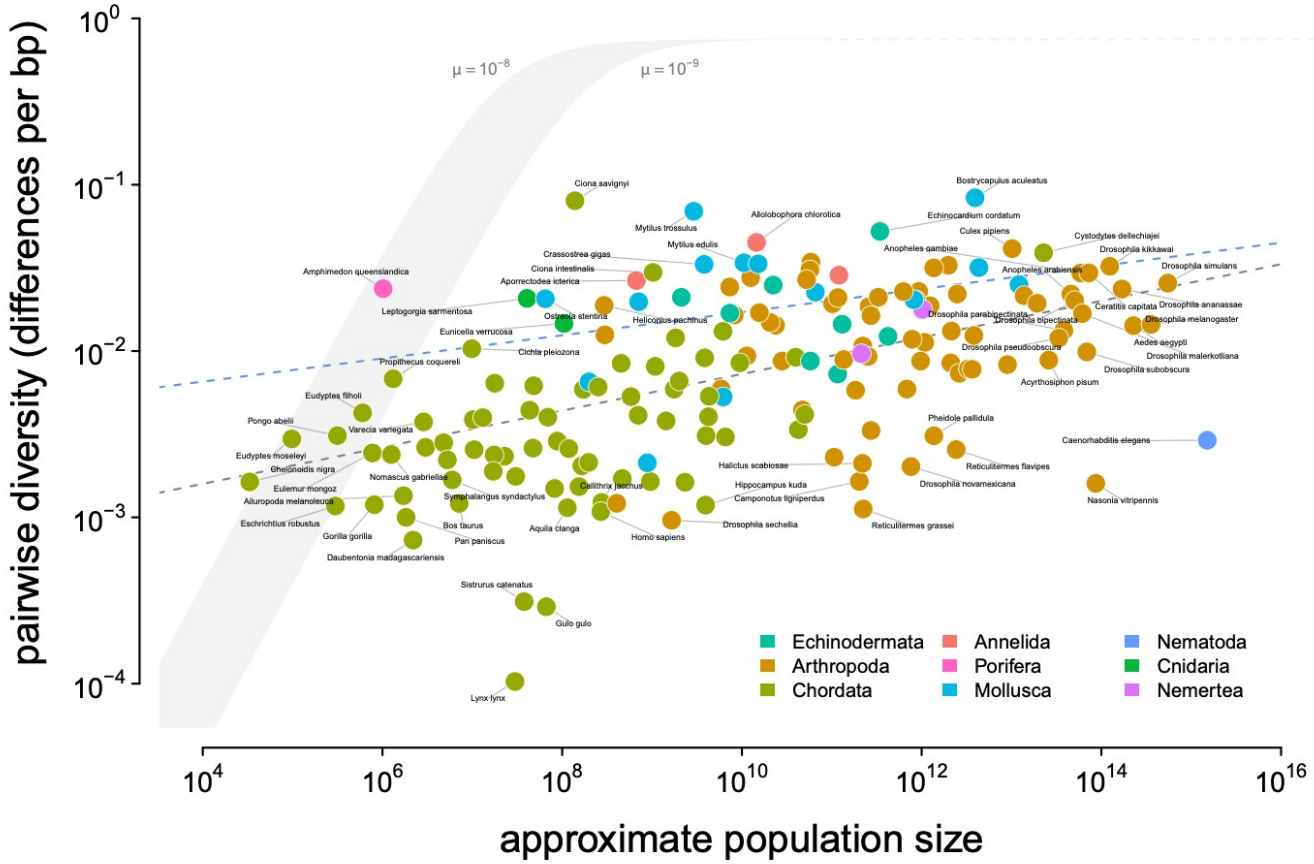
- Variation in census size is much greater than variation in genetic diversity



Buffalo 2021, eLife

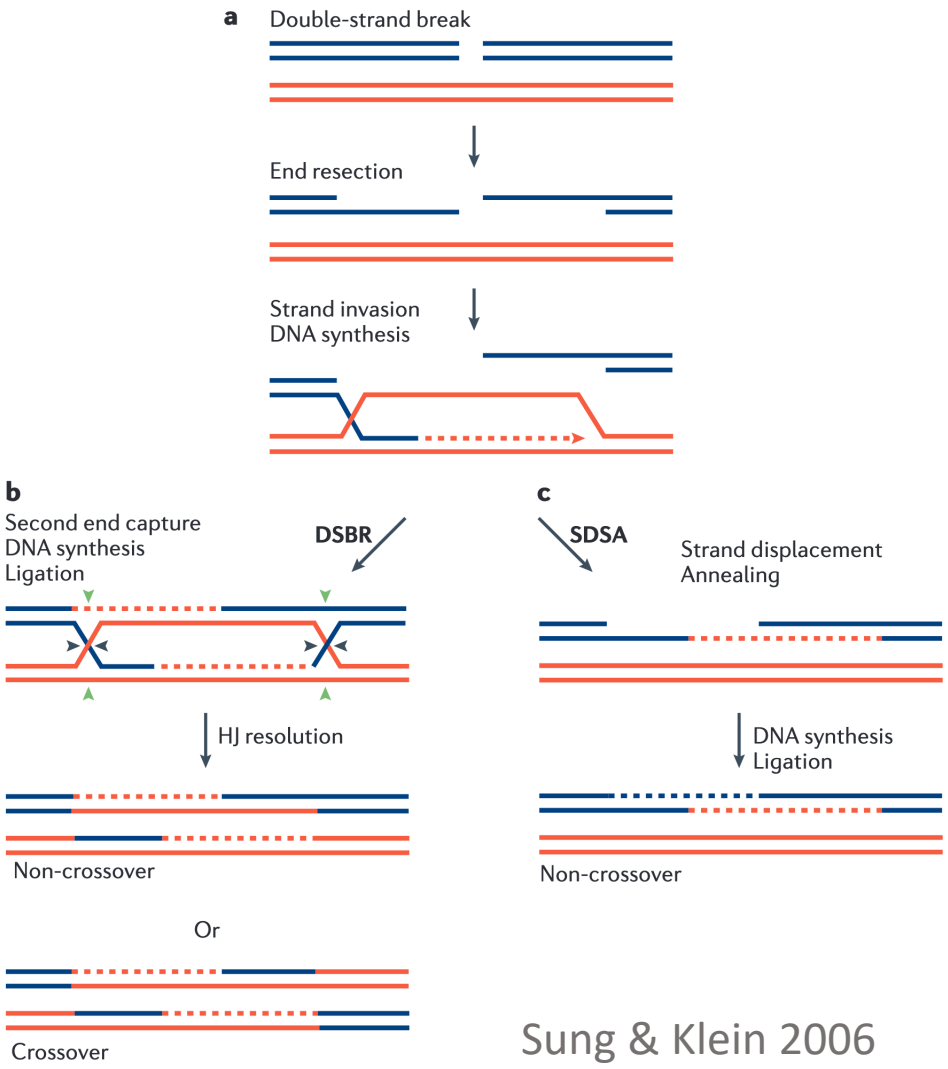
Effective population size: N_e

- Variation in census size is much greater than variation in genetic diversity
- Effective population size could explain this some of this disconnect
- N_e impacts the relative strengths of drift and selection



Buffalo 2021, eLife

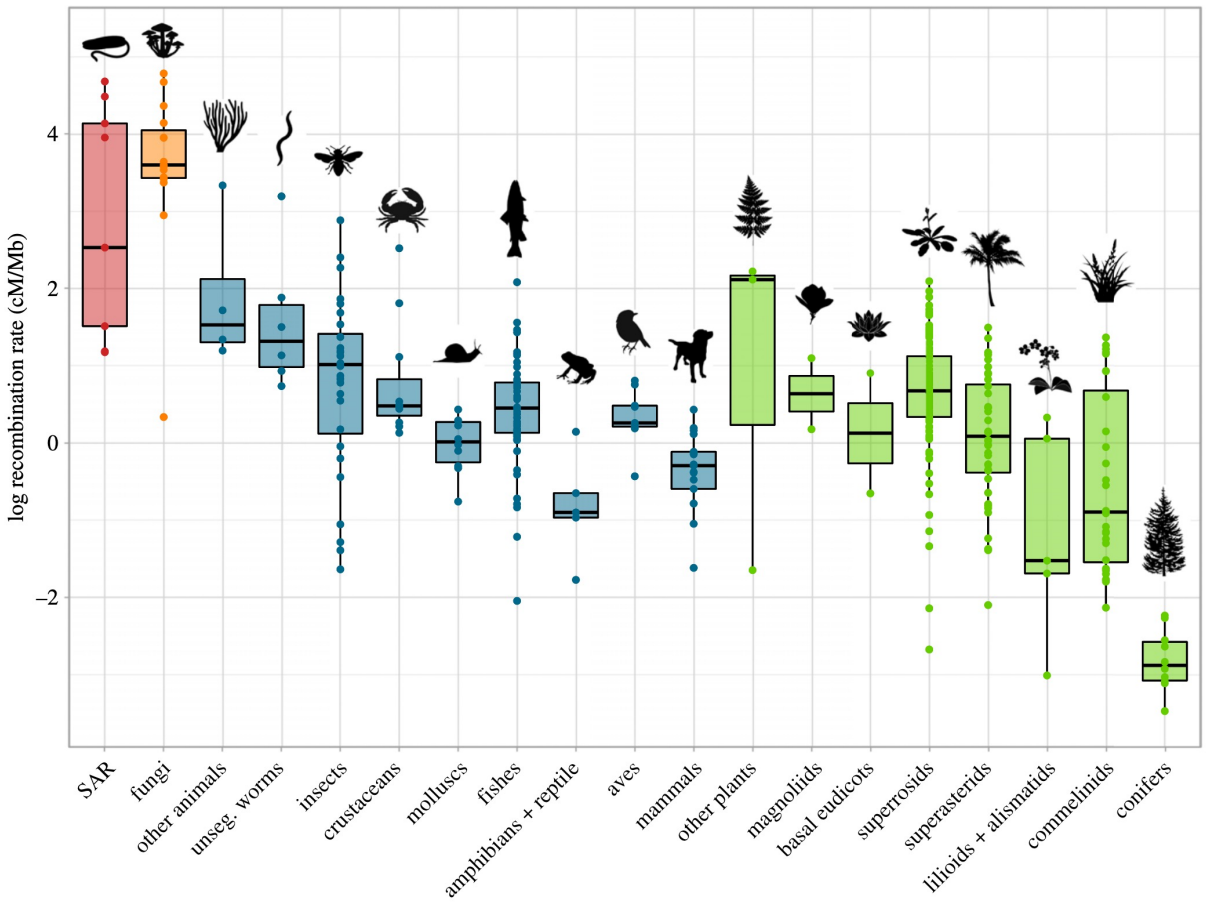
Recombination



Sung & Klein 2006

- Repair of DSBs can lead to crossover events → creates new combinations of alleles

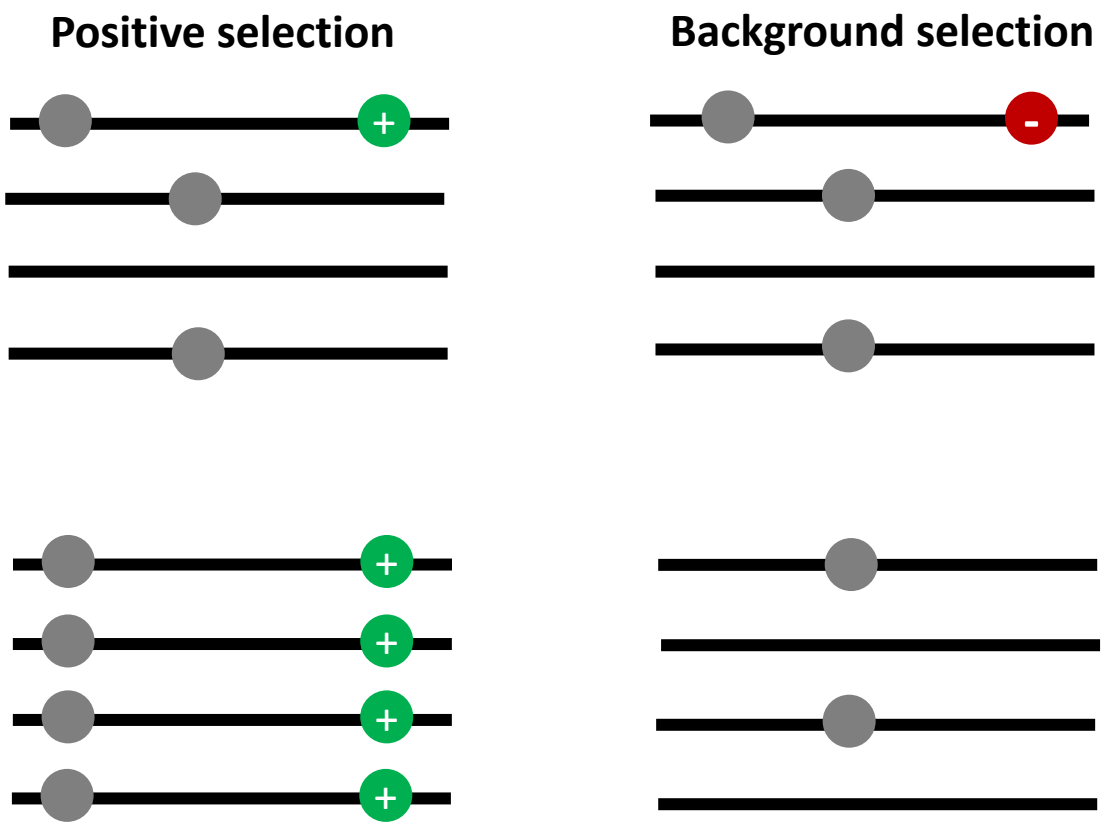
Recombination



- Repair of DSBs can lead to crossover events → creates new combinations of alleles
- Recombination rate varies across the genome and among species

Stapley et al. 2017

Natural selection



- Positive and background selection reduce variation in a population – both at target of selection and at linked sites
- Balancing selection maintains variation
- Methods for detecting selection often depend on this linked signature!

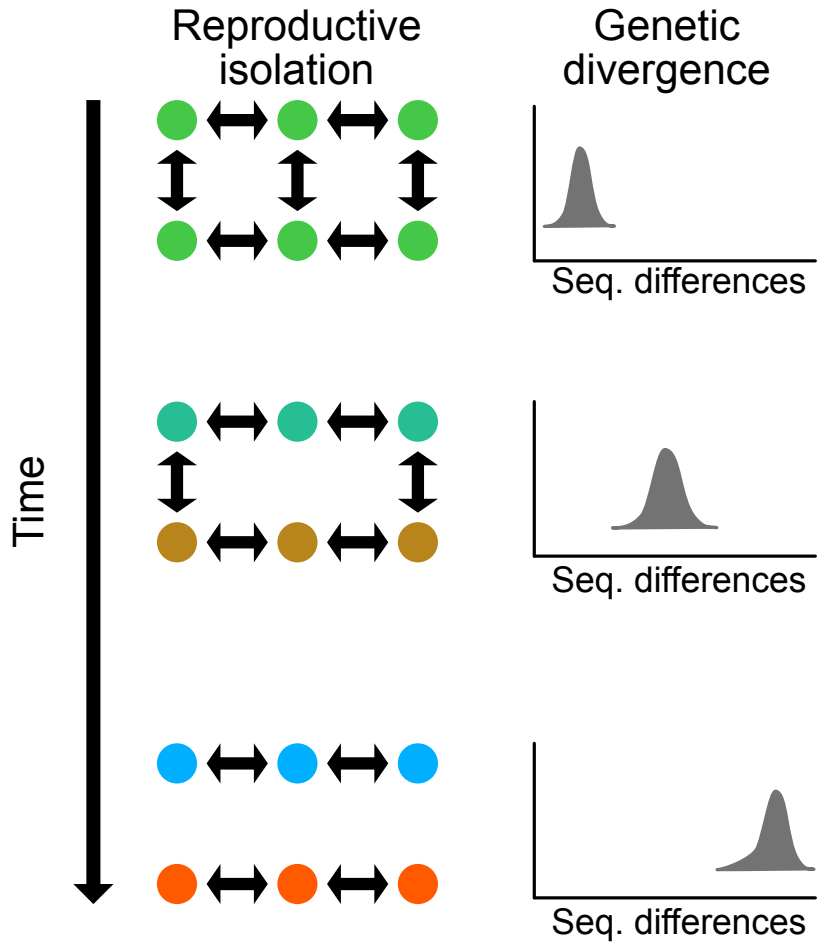
Lower RR creates **stronger** signature of selection

Gene flow

- Gene flow homogenizes populations

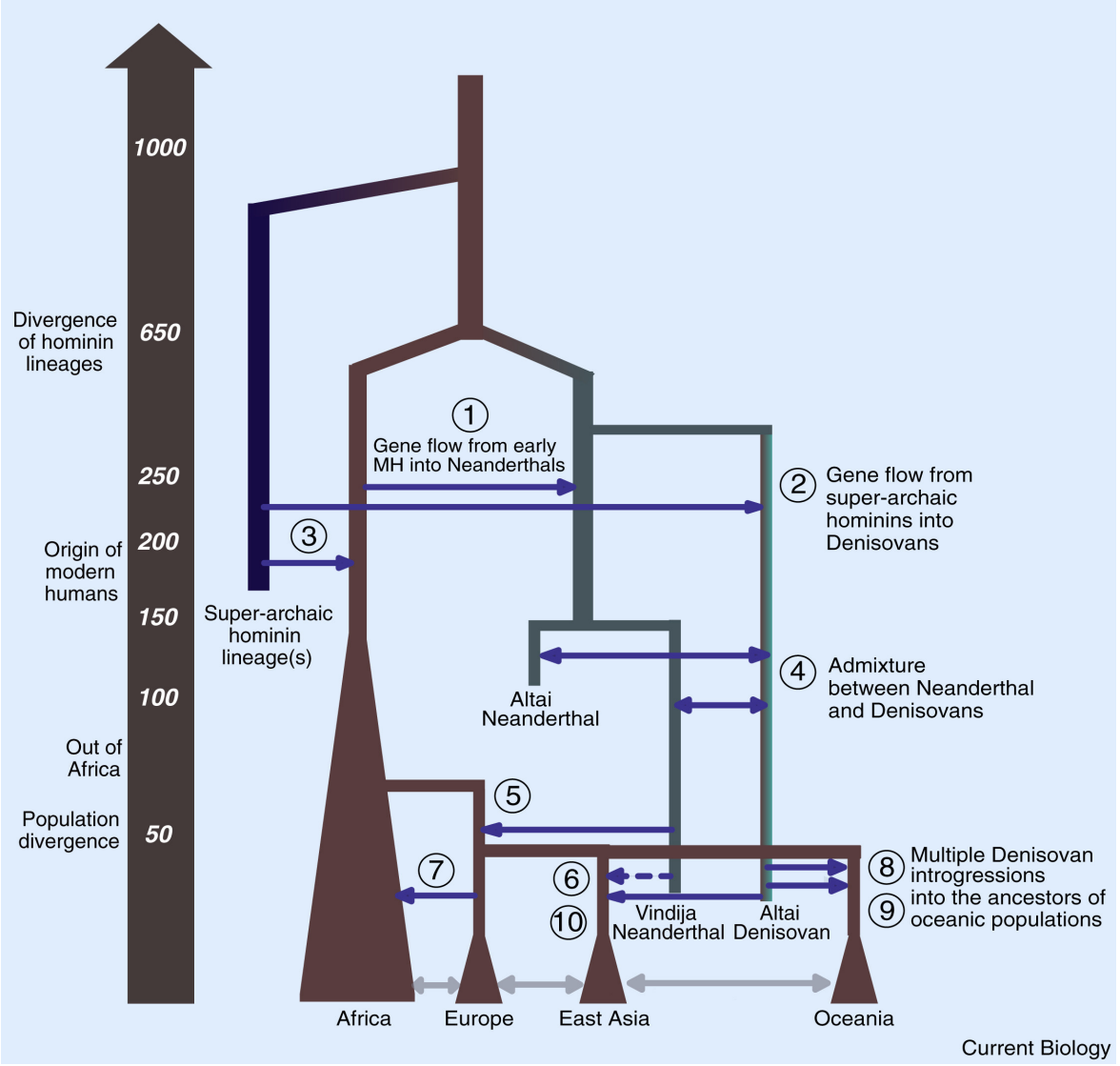
Gene flow

- Gene flow homogenizes populations
- Reduction in gene flow (reproductive isolation) can lead to speciation



Gene flow

- Gene flow homogenizes populations
- Reduction in gene flow (reproductive isolation) can lead to speciation
- Gene flow can also be an important source of adaptive mutations!



Reilly et al. 2022

Determinants of genetic diversity

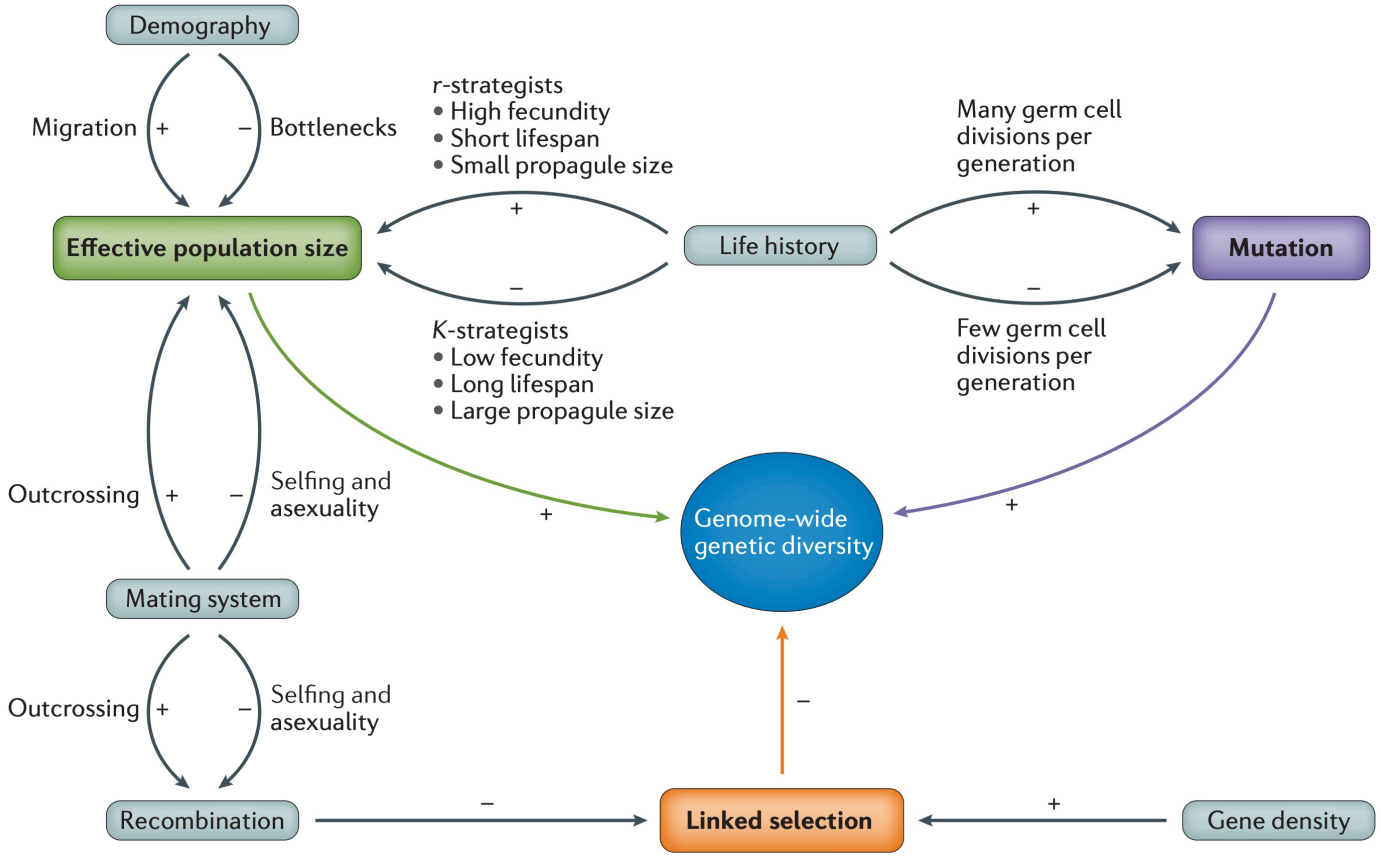
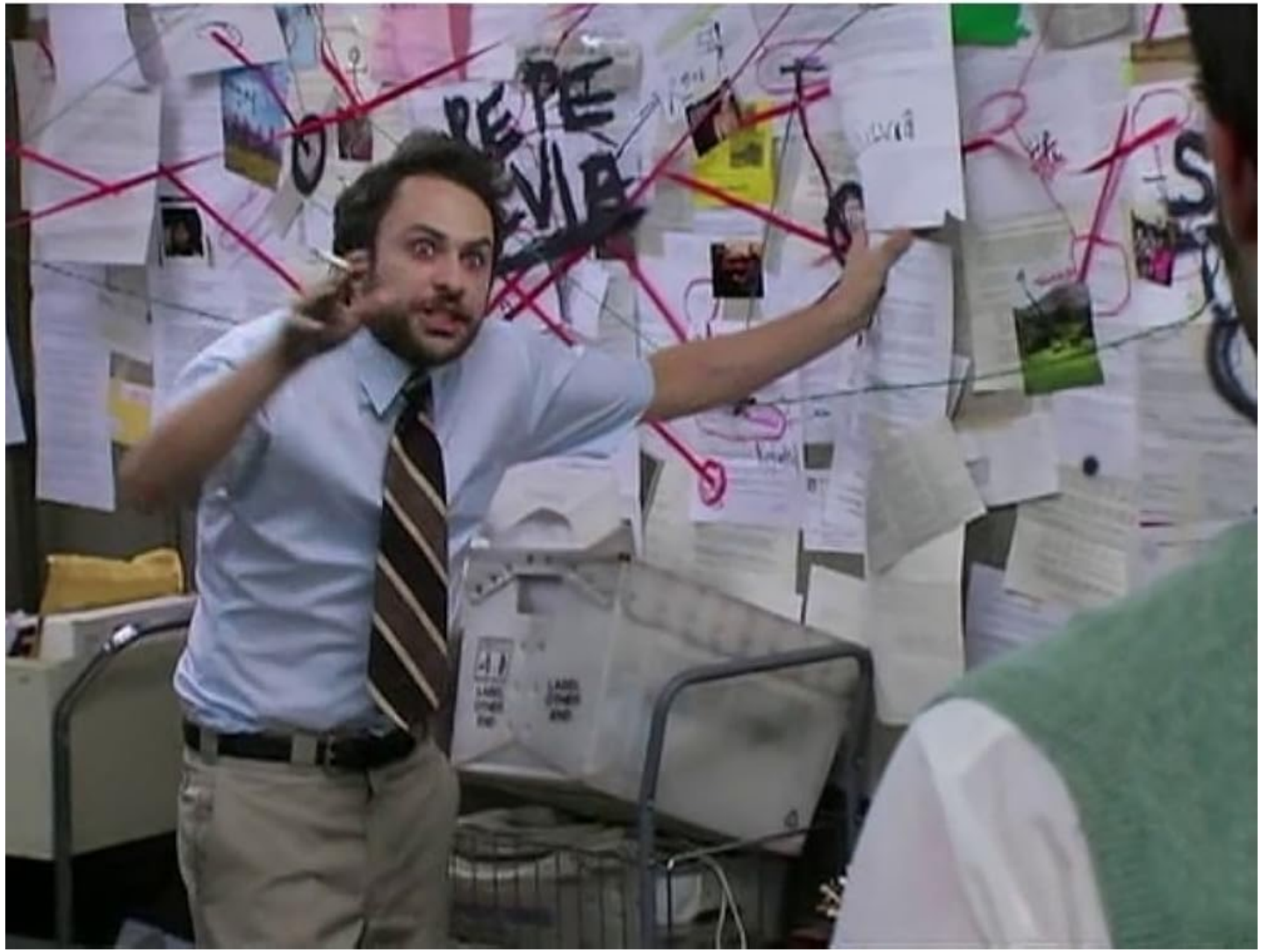


Figure 2 | **Overview of determinants of genetic diversity.** Effective population size, mutation rate and linked selection are the main factors affecting diversity. These factors are in turn governed by several other parameters. The direction of correlation is indicated by the + and – symbols. Selfing, self-fertilization.

Ellegren and Galtier 2016

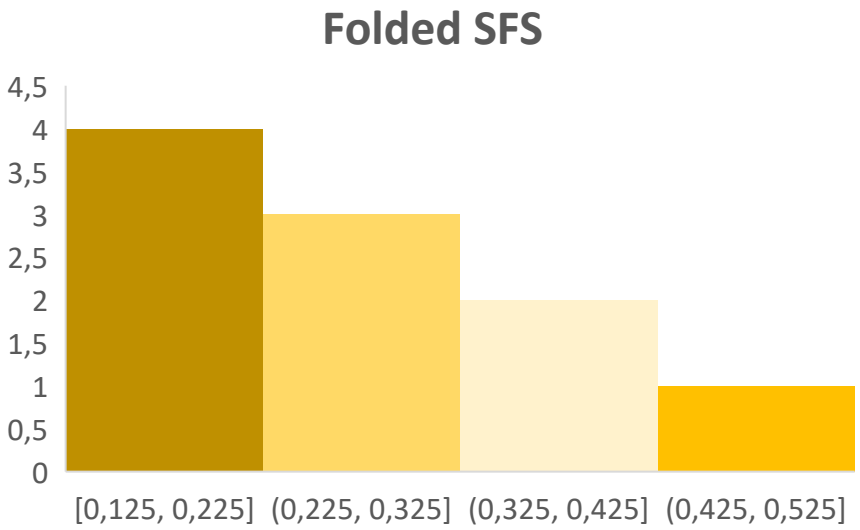
Determinants of genetic diversity

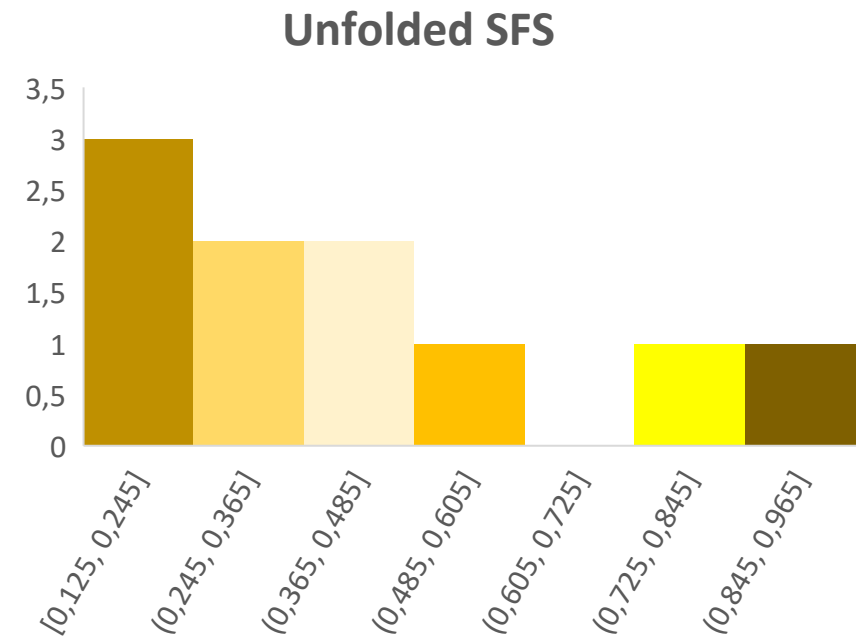


Ellegren and Galtier 2016

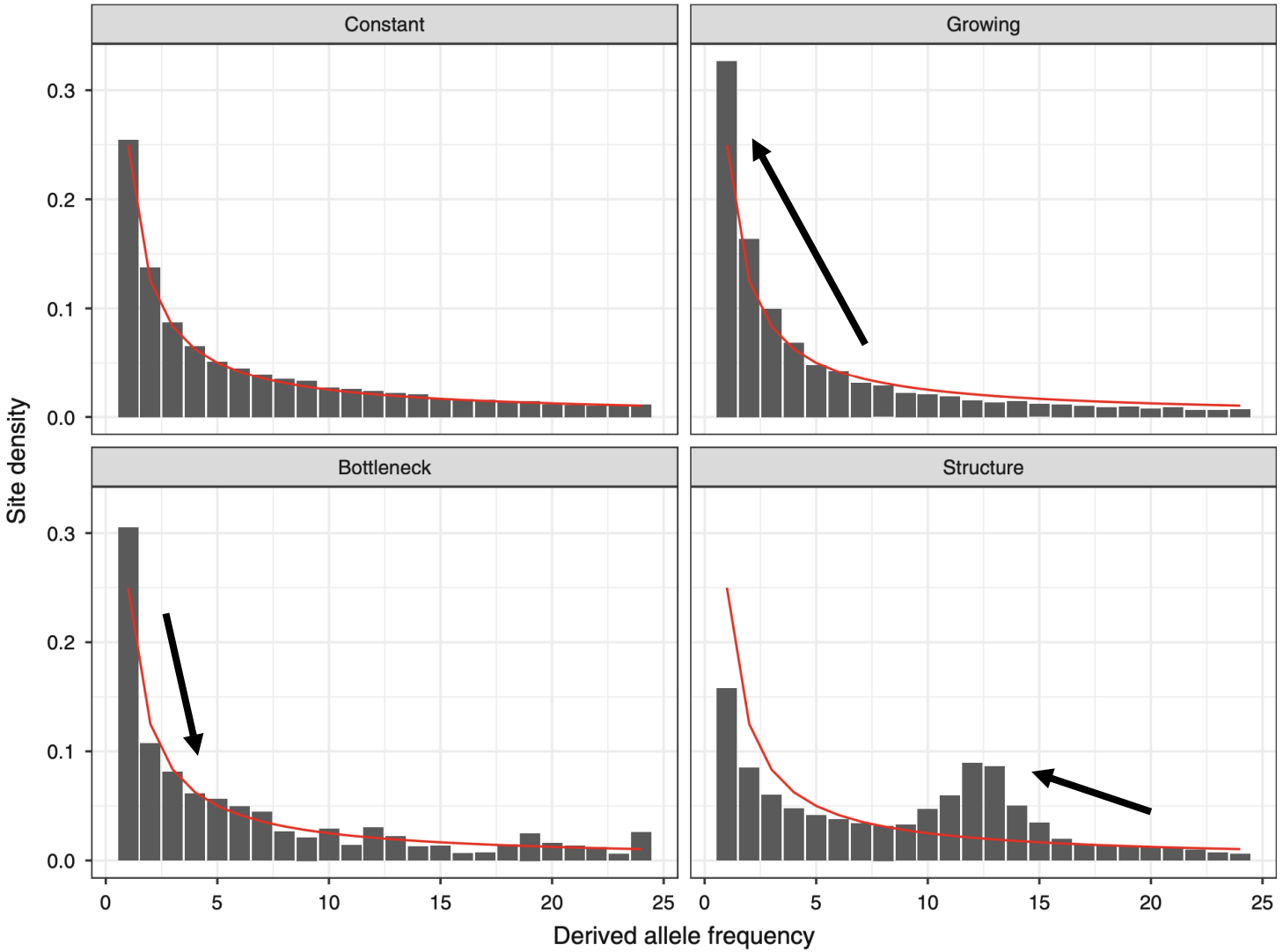
The site frequency spectrum

Outgroup	C	C	T	G	T	C	A	G	T	T
Sample 1	C	C	T	G	T	C	A	G	T	T
Sample 2	C	T	T	G	T	A	A	C	T	A
Sample 3	G	T	T	C	C	C	A	C	T	T
Sample 4	C	T	T	G	T	C	A	C	A	A
Sample 5	G	T	T	C	C	A	G	C	T	T
Sample 6	C	T	T	G	C	A	A	C	A	T
Sample 7	C	T	A	G	T	C	A	G	A	T
Sample 8	C	T	T	G	T	A	A	C	T	T





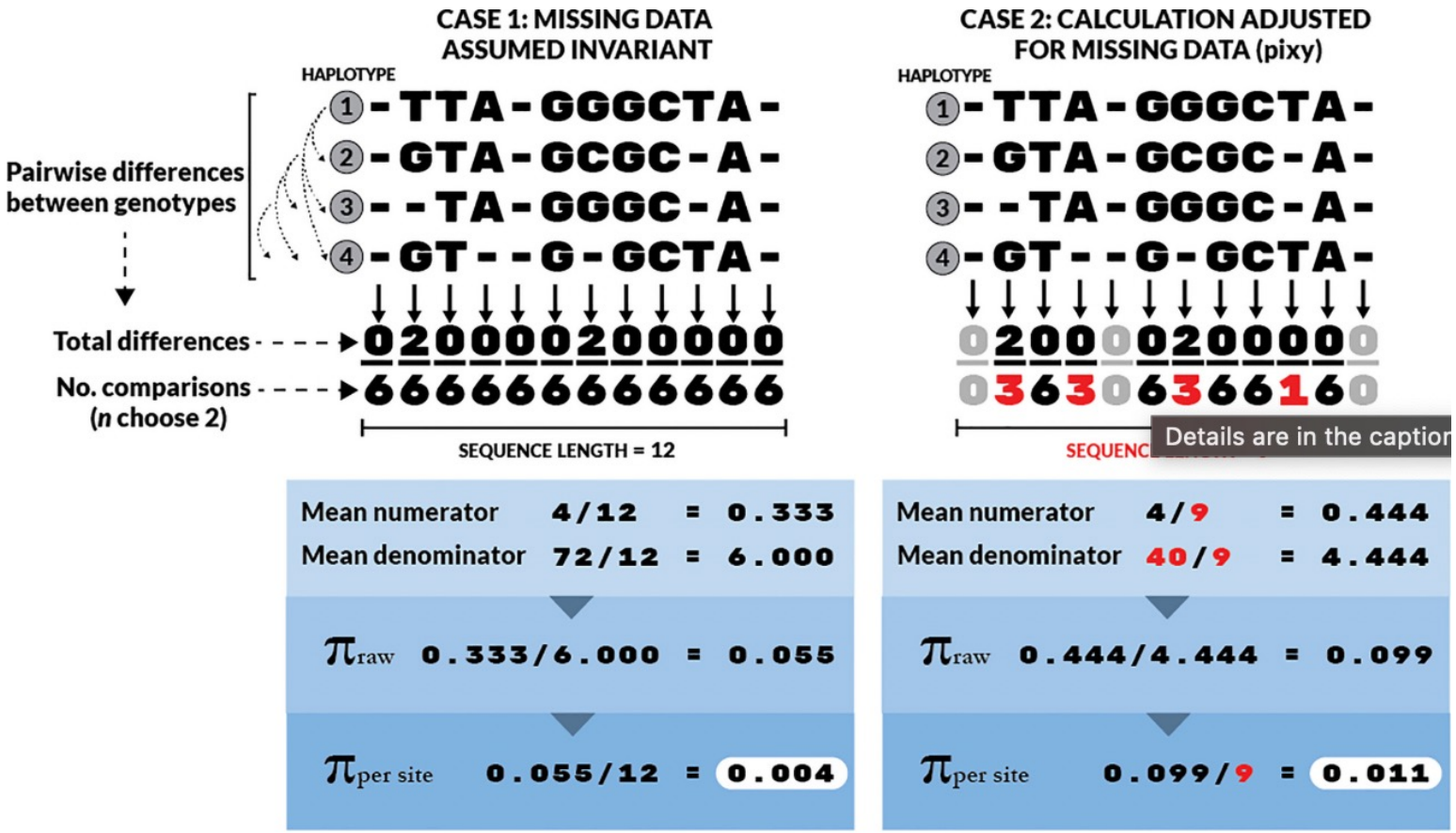
The site frequency spectrum



Neutral expectation for the density of sites in a frequency class = 1/f

The genome-wide SFS is impacted by the demographic history of a population

Estimating genetic diversity



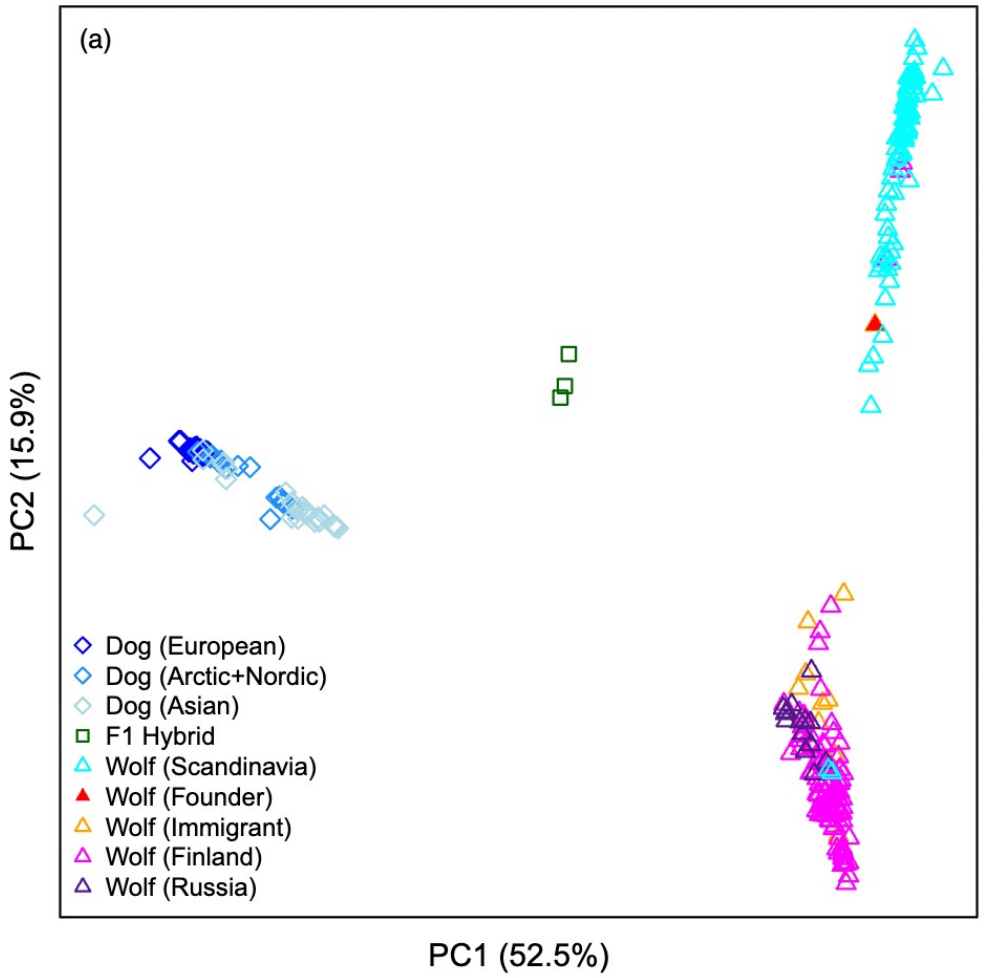
Pairwise nucleotide diversity: π

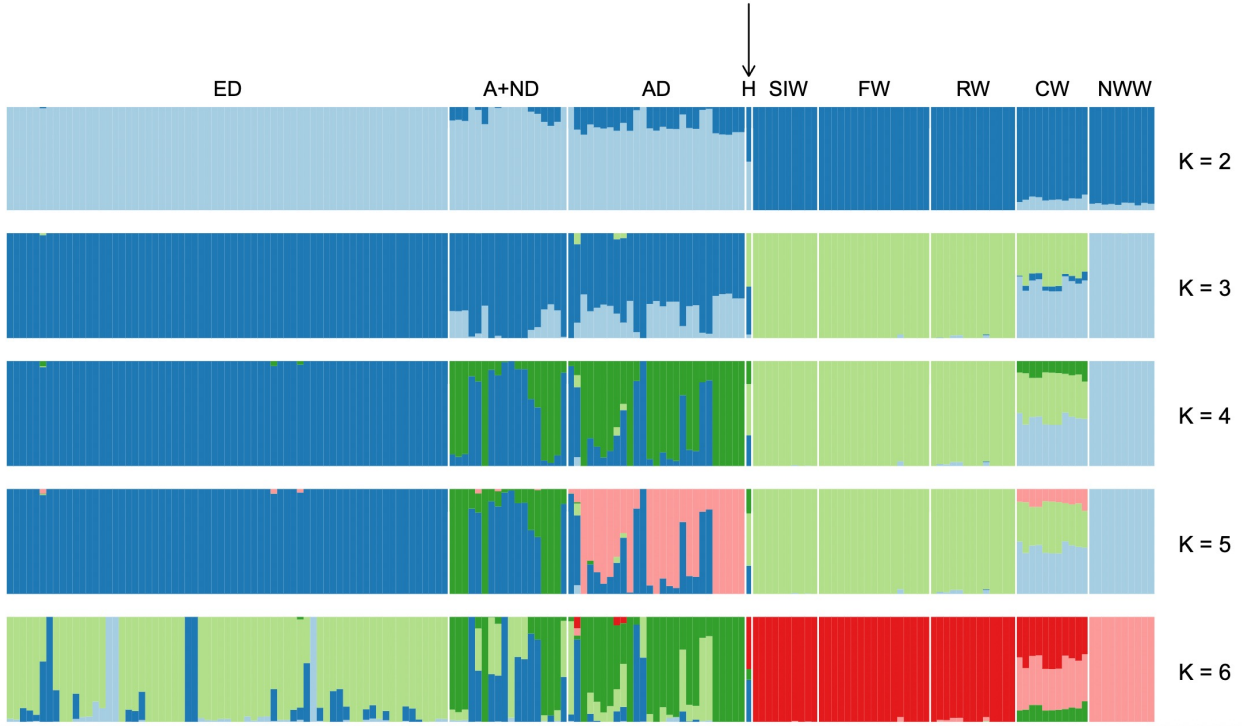
Mean number of pairwise differences in a population

$\pi = 4N_e \mu$

Korunes & Samuk 2021, Mol. Ecol. Res.

Population structure

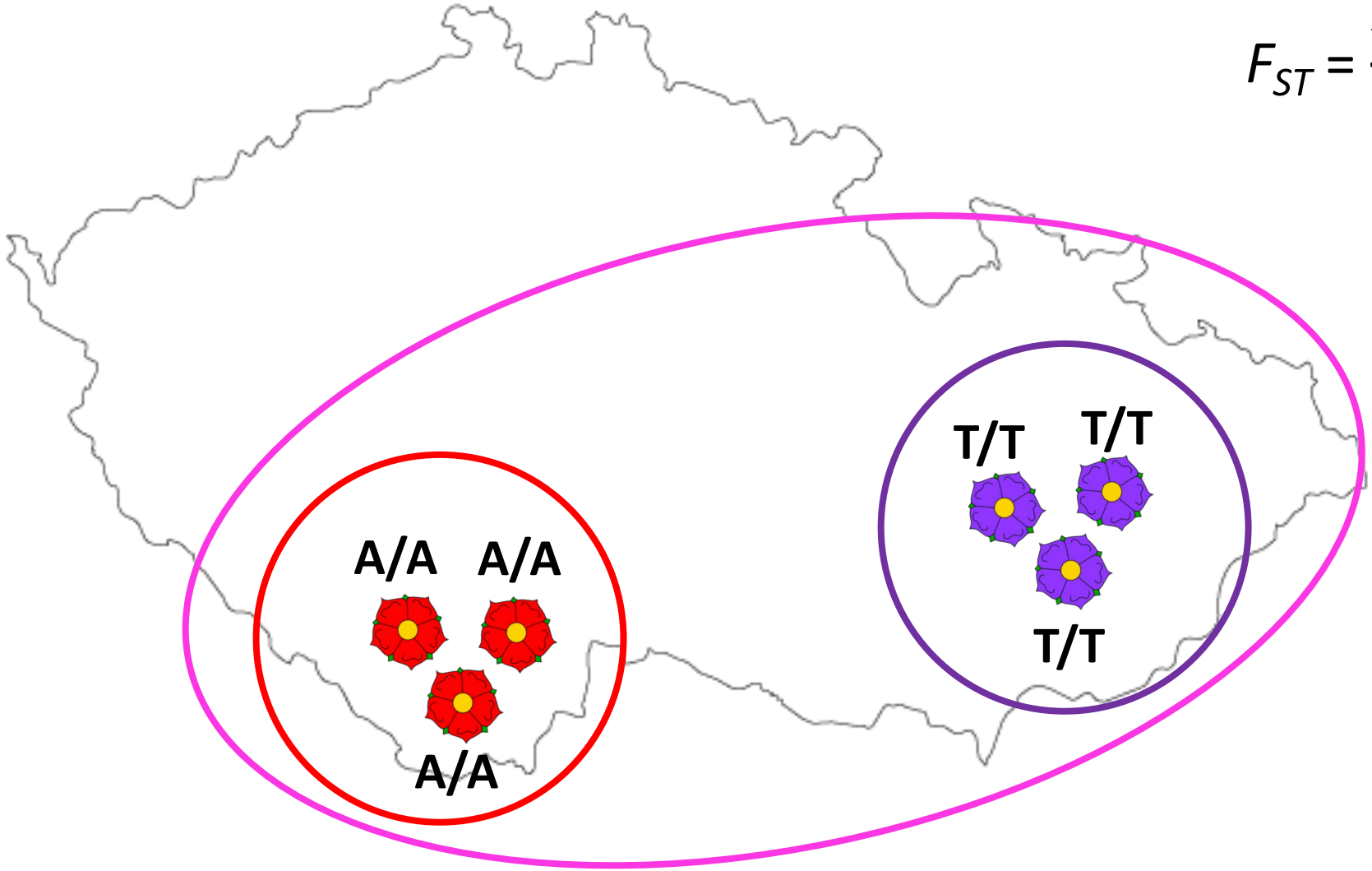




Smeds et al. 2020

Measures of genetic differentiation

$$F_{ST} = \frac{\pi_{BETWEEN} - \pi_{WITHIN}}{\pi_{BETWEEN}}$$

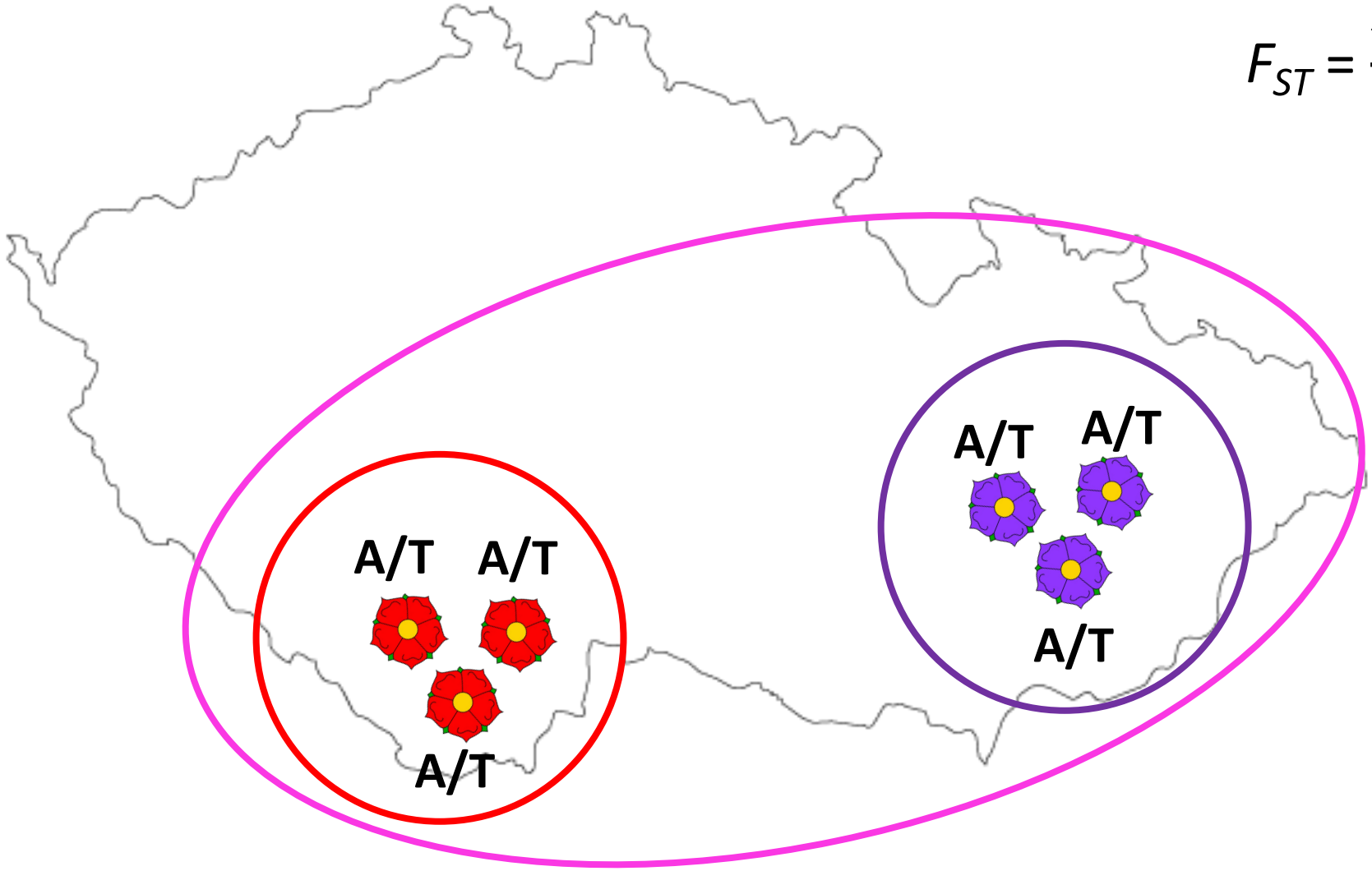


$$F_{ST} = \frac{0.5 - 0}{0.5}$$

$$F_{ST} = 1$$

Measures of genetic differentiation

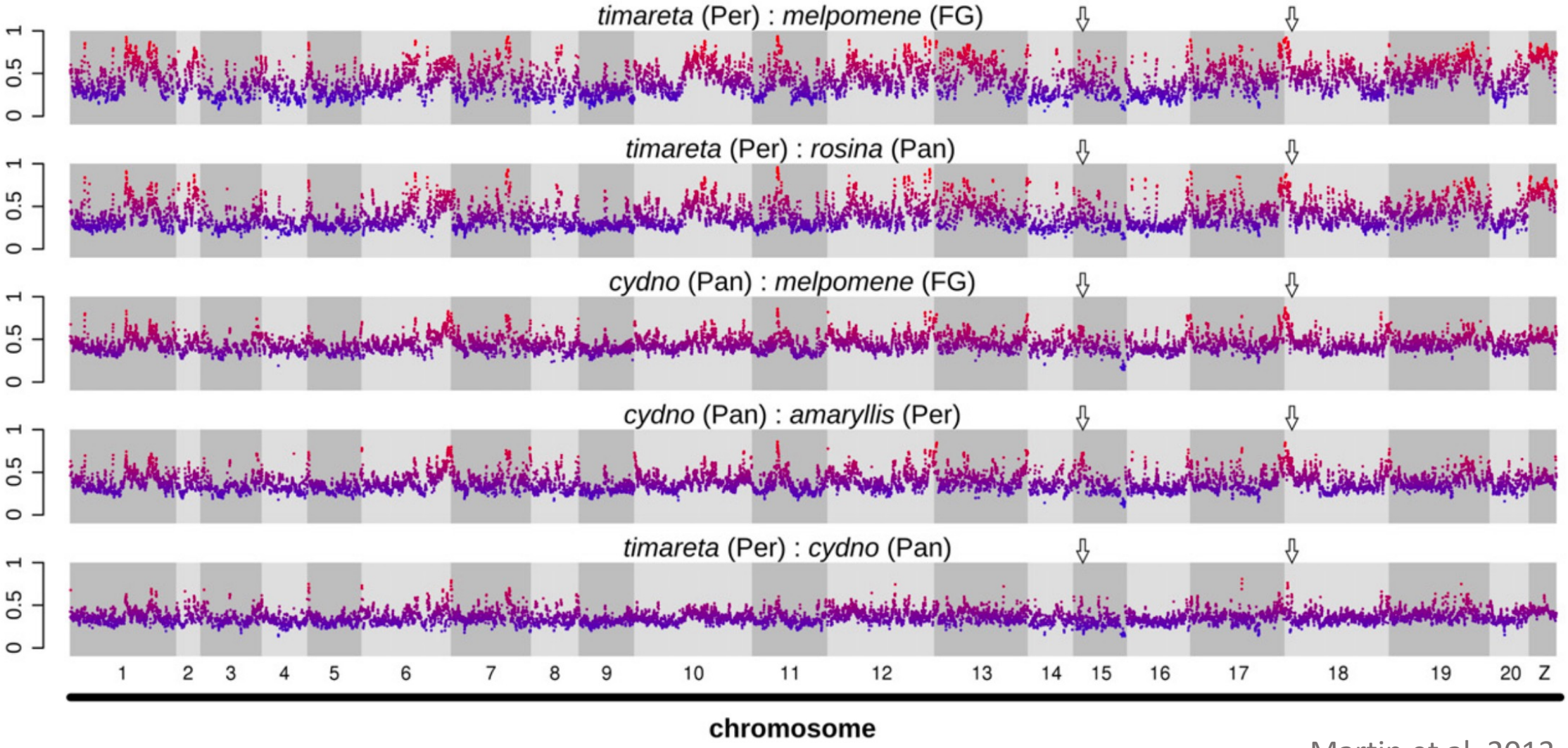
$$F_{ST} = \frac{\pi_{BETWEEN} - \pi_{WITHIN}}{\pi_{BETWEEN}}$$



$$F_{ST} = \frac{0.5 - 0.5}{0.5}$$

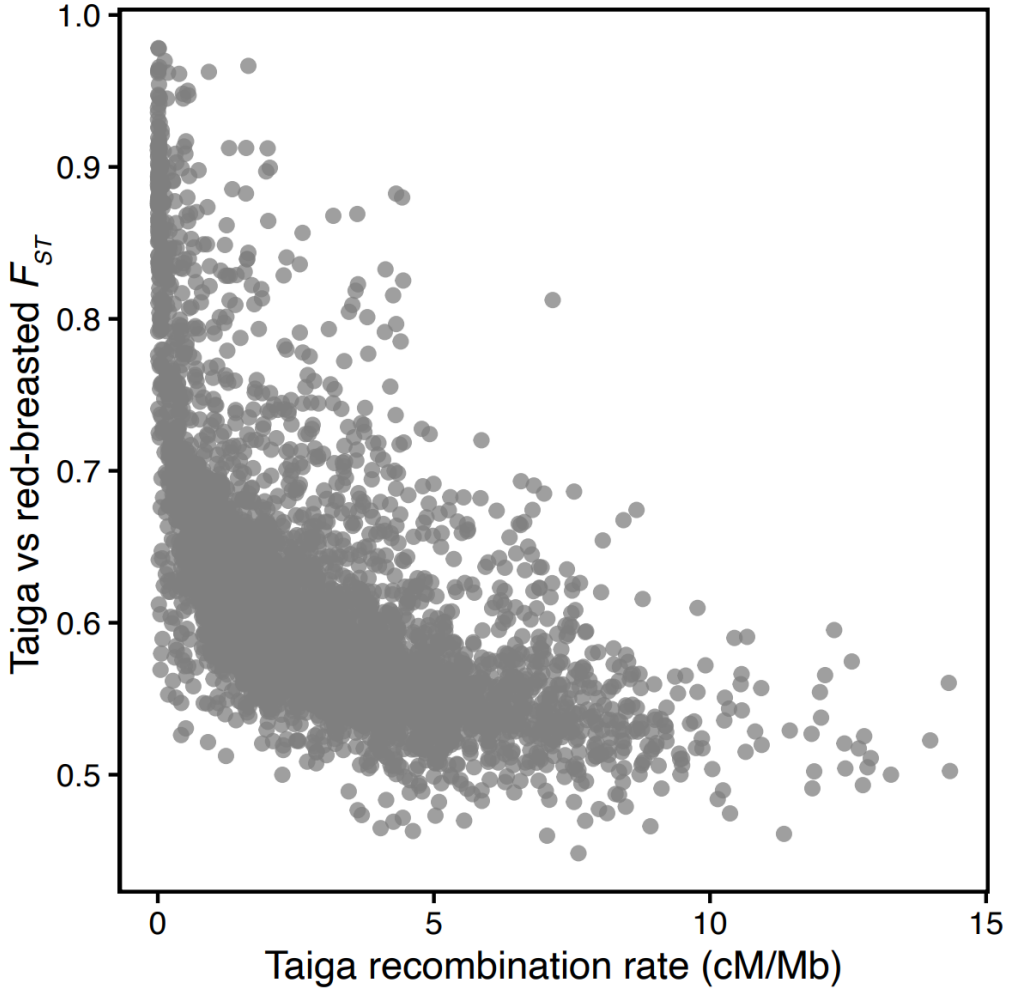
$$F_{ST} = 0$$

Differentiation varies across the genome



Martin et al. 2013

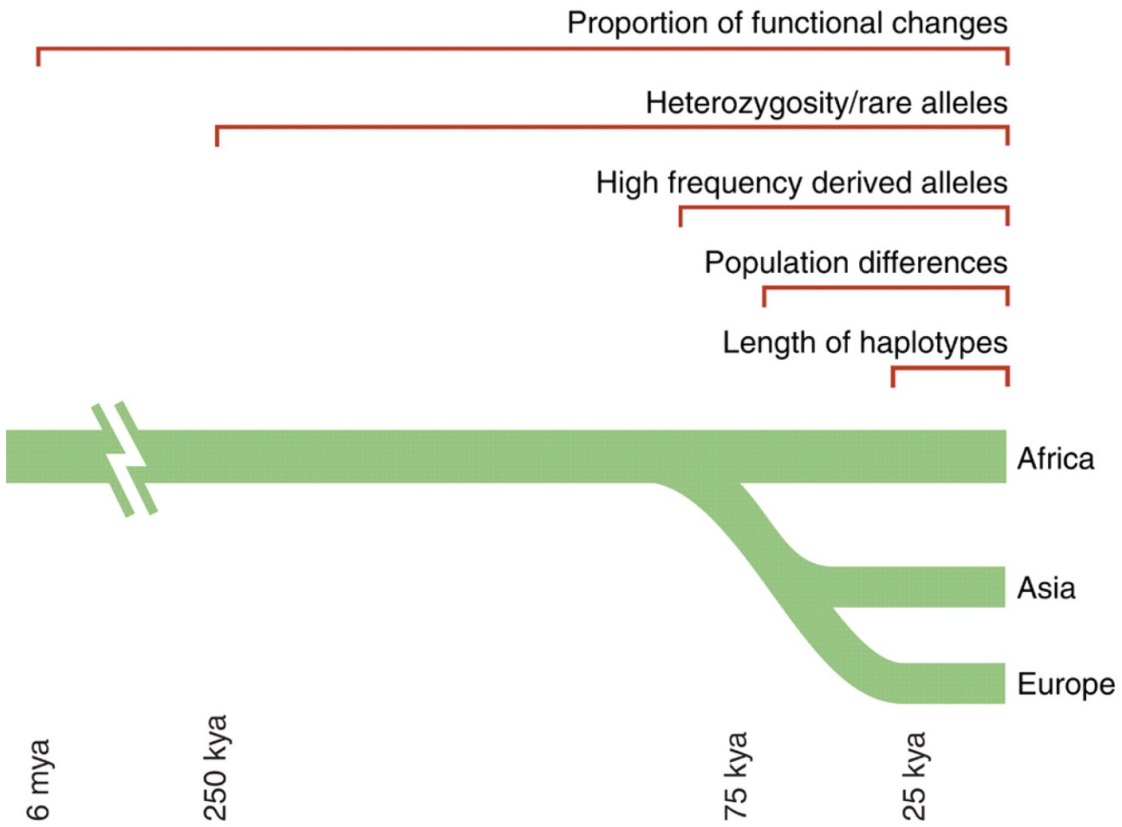
Differentiation varies across the genome



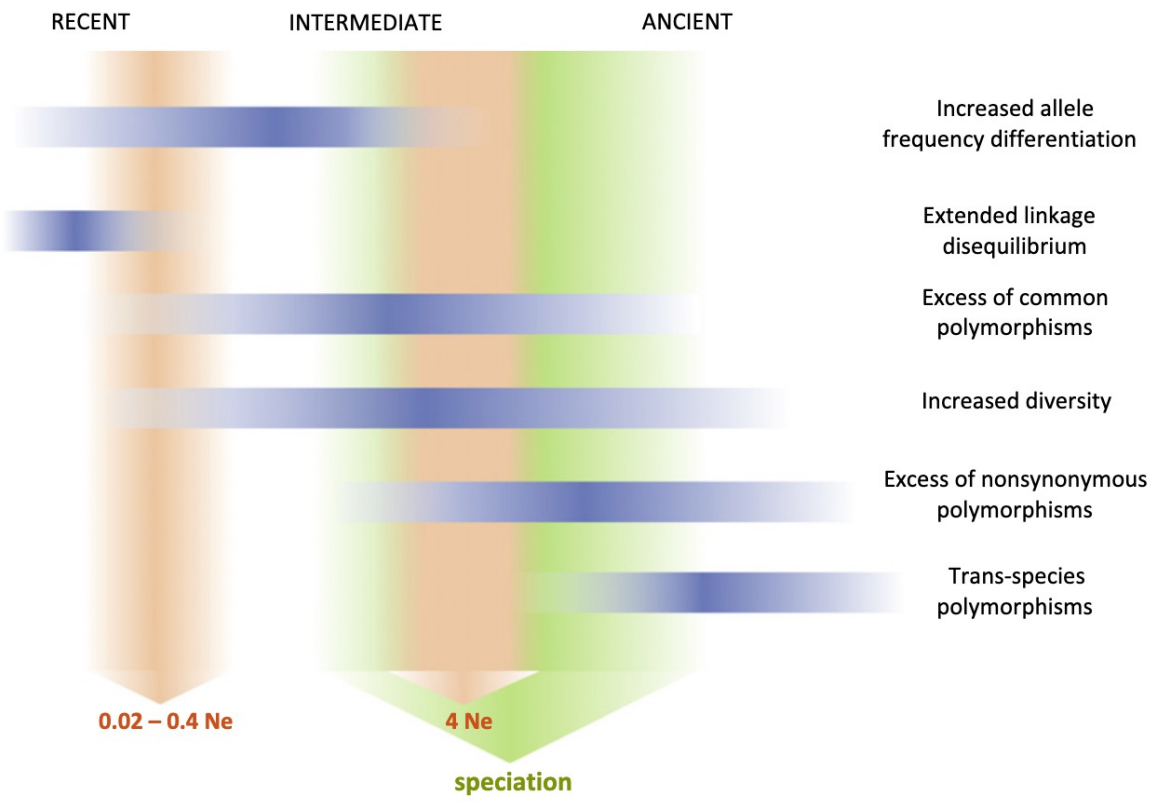
Differentiation is strongly correlated with recombination rate across a wide range of organisms

Chase et al. 2024, Proc. B

Detecting natural selection



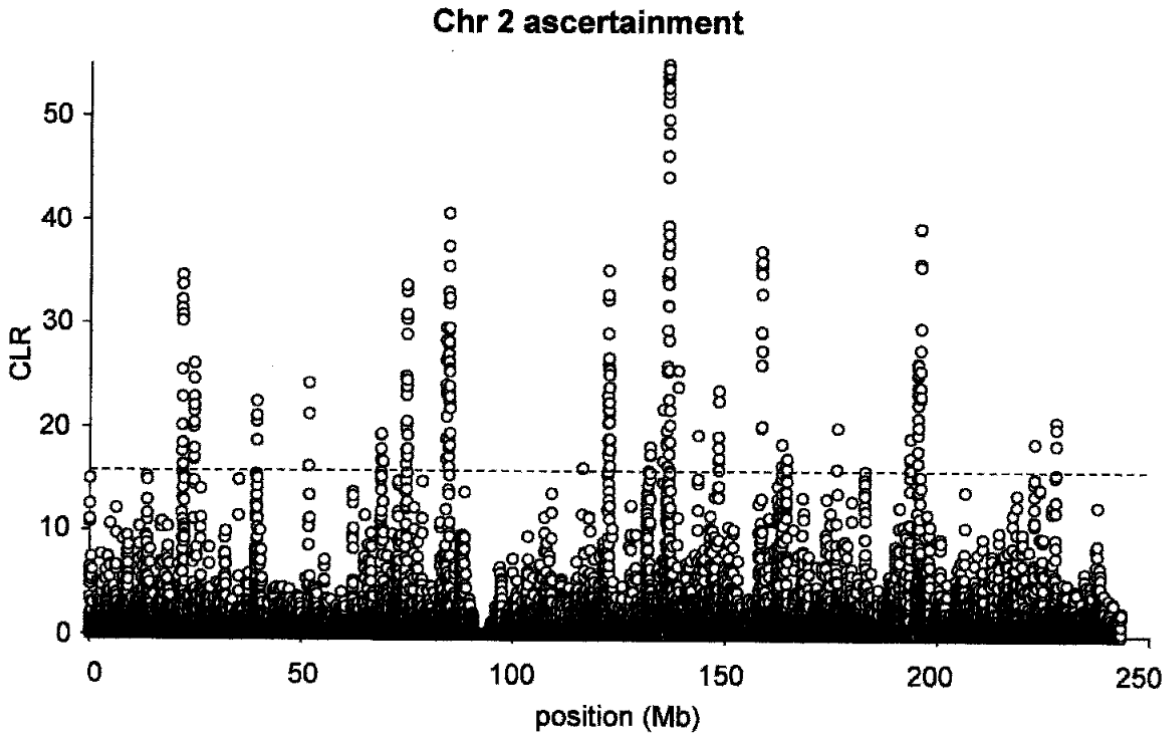
Sabeti et al. 2006, Science



Fijarczyk & Babik 2015, Mol. Ecol.

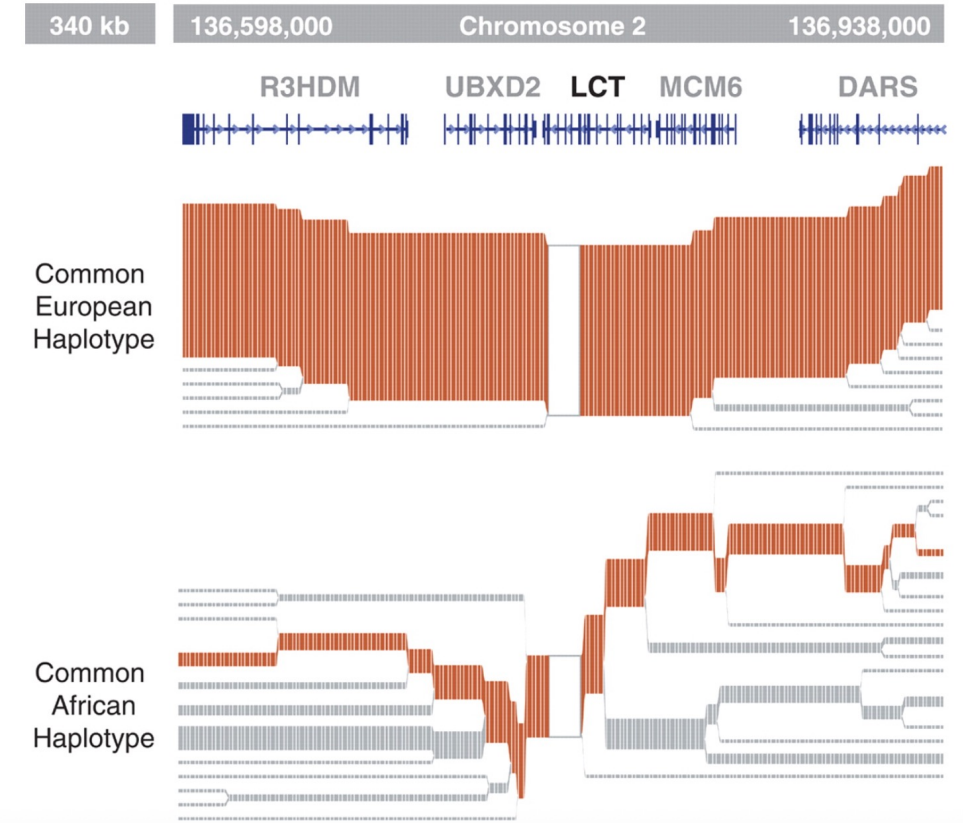
Detecting natural selection

SFS-based methods



Nielsen et al. 2005, Genome Research

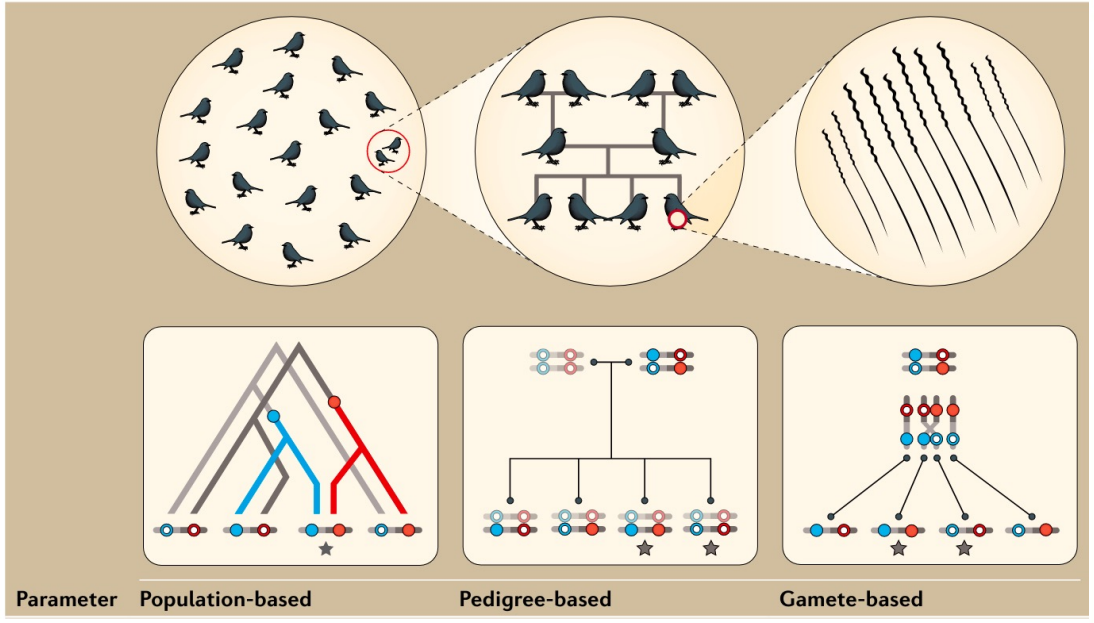
Haplotype-based methods



Sabeti et al. 2006, Science

Estimating recombination rates

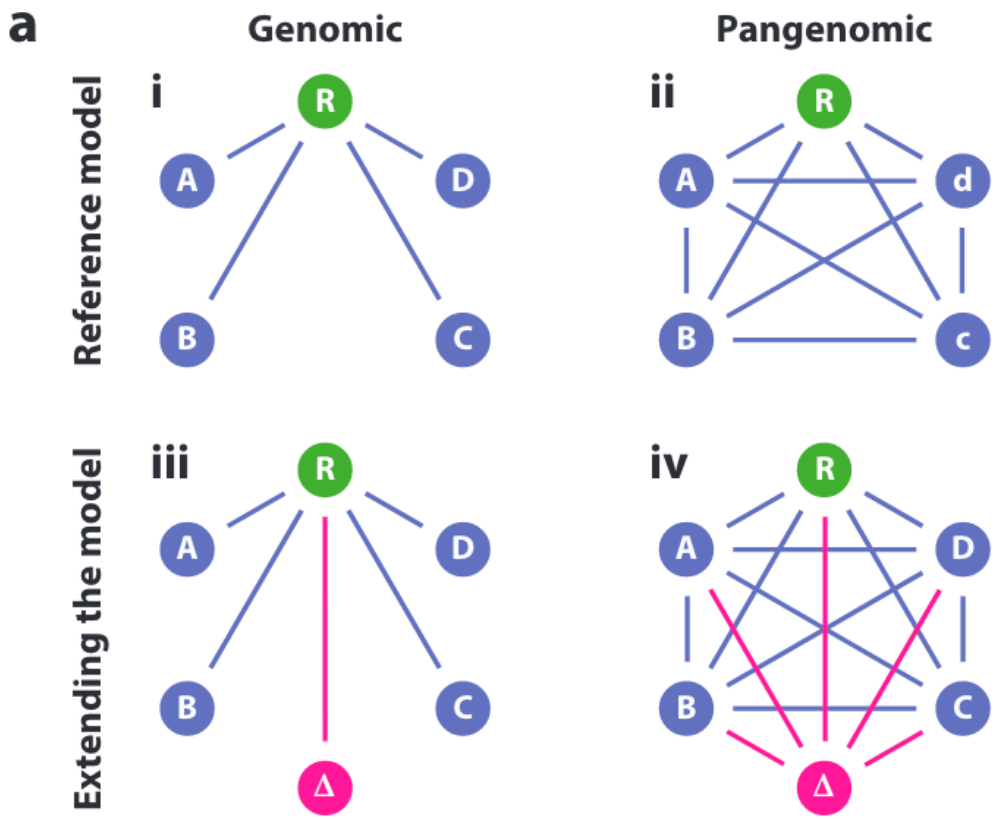
Table 1 | **Summary of the three genomic-based approaches to infer the recombination landscape**

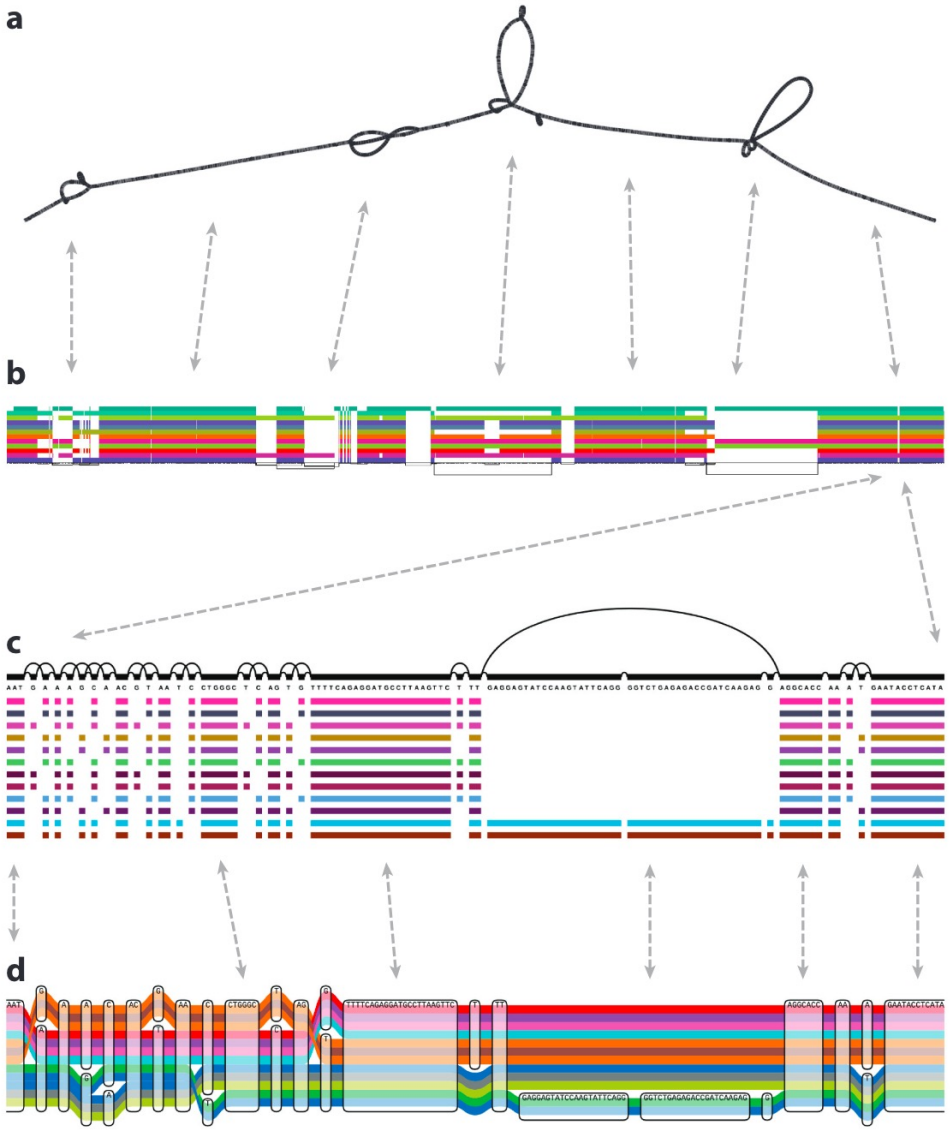


Parameter	Population-based	Pedigree-based	Gamete-based
Estimate	Population recombination rate	Genetic distance	Crossover frequency
Sample size	10–30	>100	≥1
Pros	Low to moderate sampling Moderate to high resolution Simple study design	Sex-specific recombination maps Unbiased by population level processes Variation between individuals and heritability can be measured Can improve genome assemblies	High resolution Low sampling required Variation between individuals can be compared
Cons	Average between sexes and through time Biased by effective population size, demographic history, selection and mutation Dependent on genome assembly quality	Large sampling required Complex and time-intensive study design required Low resolution (~1 Mb, 0.5–2 cM) Biased by SNP density in mapping population Only contemporary snapshot of recombination	Single-sex recombination map (usually males) Biased by individual SNP density Dependent on genome assembly quality Only contemporary snapshot of recombination

Peñalba & Wolf 2020

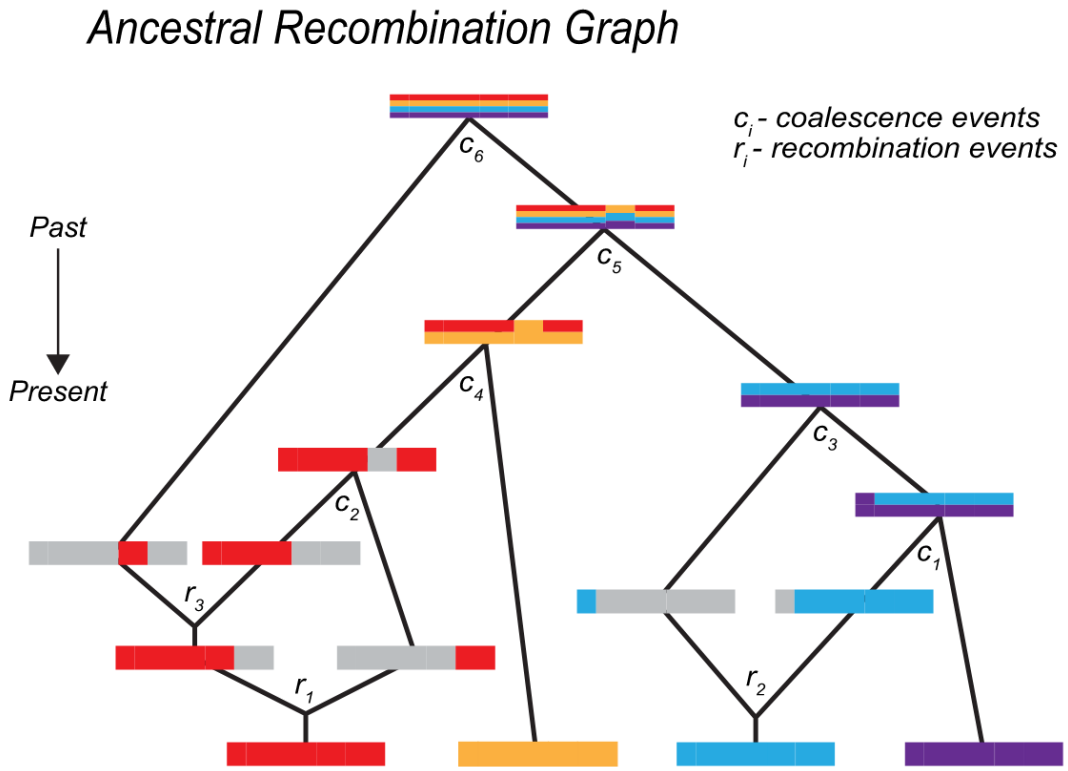
Advances in population genomics: Pan-genomics



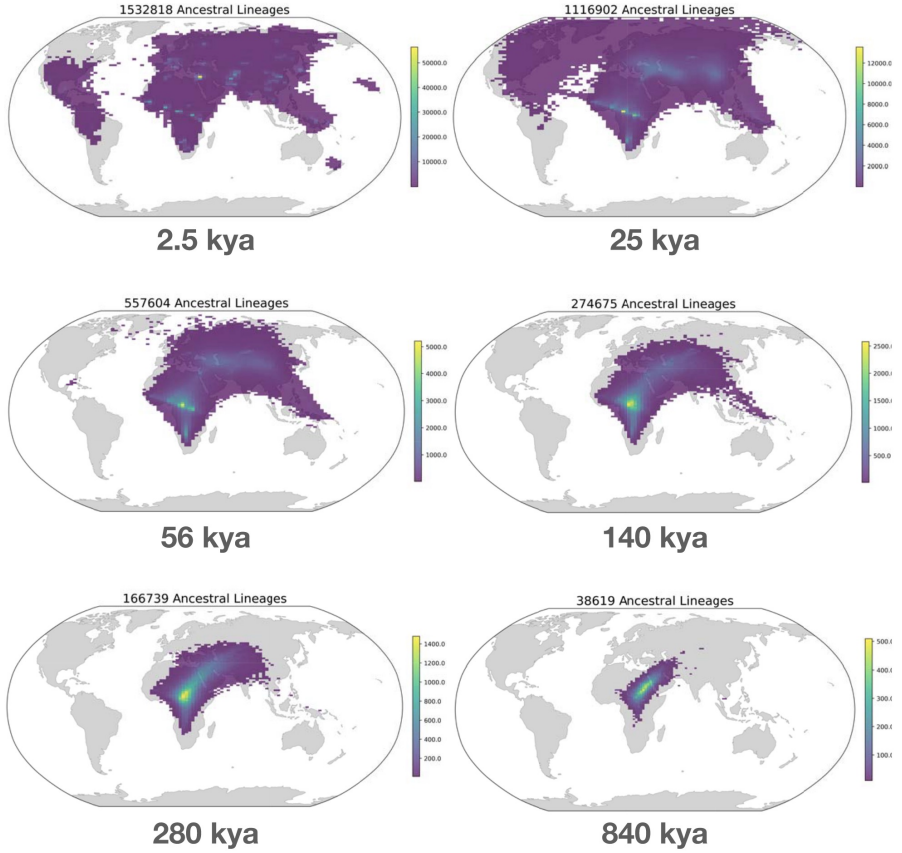


Eizenga et al. 2020

Advances in population genomics: Ancestral recombination graphs



Shipilina et al. 2022



Wohns et al. 2022

Population genetics and scientific racism

POPULATION GENETICS

A geographic history of human genetic ancestry

Michael C. Grundler[1], Jonathan Terhorst[2], Gideon S. Bradburd[1]*

Genetic variation in humans is often summarized with discrete labels, but these can be inaccurate and misleading (*5*, *6*). Even when based on geographic history, genetic ancestry labels oversimplify a complex picture because they implicitly focus on only a single point in time. For example, based on our current understanding of human origins, all living individuals are "African" (regardless of the geography of their recent ancestors) when considering their ancestry ~200,000 years before present. Advances in the study of ancient DNA have revealed a lack of genetic continuity within geographic regions (*7*–*10*), further highlighting the shortcomings of genetic ancestry labels. The fact that these labels are generated using statistical genetics approaches gives them the veneer of authenticity, further reifying problematic notions of race and ancestry in society (*6*, *11*).

Reviewing Rushton's magnum opus book, Barash (1995) averred that combining "numerous little turds of variously tainted data" does not yield results, just a "pile of shit." It is a good line, but to our argument, it speaks to the cognitive style of the human biodiversity movement: accumulate a giant corpus of decontextualized mainstream science, flawed race science, and political writing, ignore or dismiss any critiques, and then treat genetically determined racial differences in behavior as forbidden knowledge—obviously true, but suppressed by the academic left.

Panofsky et al. 2020

Further resources

- Statistical population genomics

  doi.org/10.1007/978-1-0716-0199-0

- Population and quantitative genetics
  - https://cooplab.github.io/popgen-notes/

- Work on scientific racism and genetics:
  - https://doi.org/10.1002/hast.4925
  - https://doi.org/10.3389/fgene.2024.1345631
  - https://doi.org/10.1002/ajpa.24150
  - https://doi.org/10.1038/s41436-021-01109-w