



dashiell.massey@duke.edu

Simulating assortative mating by global ancestry in admixed populations

Dashiell J. Massey and Amy Goldberg

Department of Evolutionary Anthropology, Duke University, Durham NC 27708 USA

Non-random mating is a feature of real-world populations

Assortative mating describes any biased self-referential mate choice "rule"

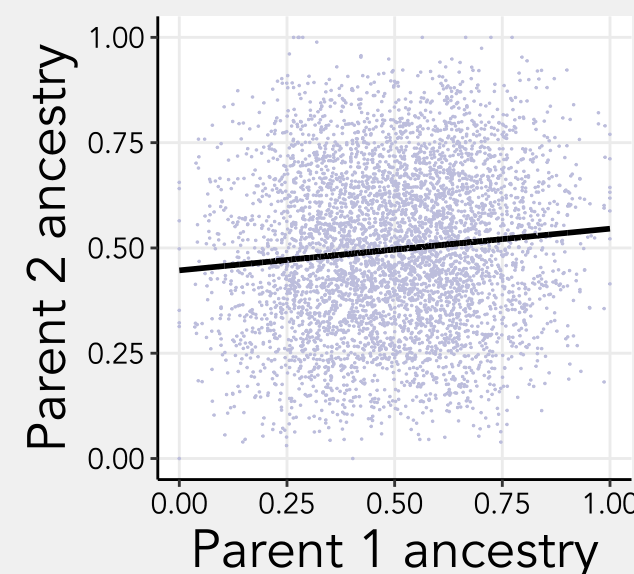
- Positive assortative mating: Prefer mate similar to self
- Negative assortative mating: Prefer mate different from self

Self-referential preferences can extend beyond "matching":
e.g., male-larger pairings in mountain chickadee (Welkin et al., 2023)

Ancestry-associated assortative mating has been reported in admixed human populations

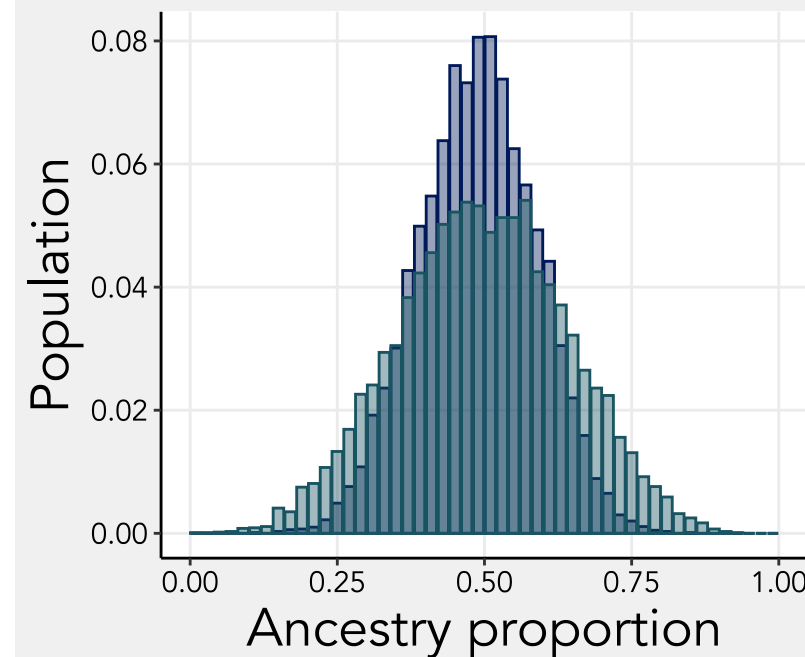
Typically measured by correlation in ancestry proportion within mate pairs

e.g., Risch et al., 2009; Sebro et al., 2010; Zou et al., 2015; Zaitlen et al., 2017; Avadhanam & Williams, 2022; Korunes et al., 2022

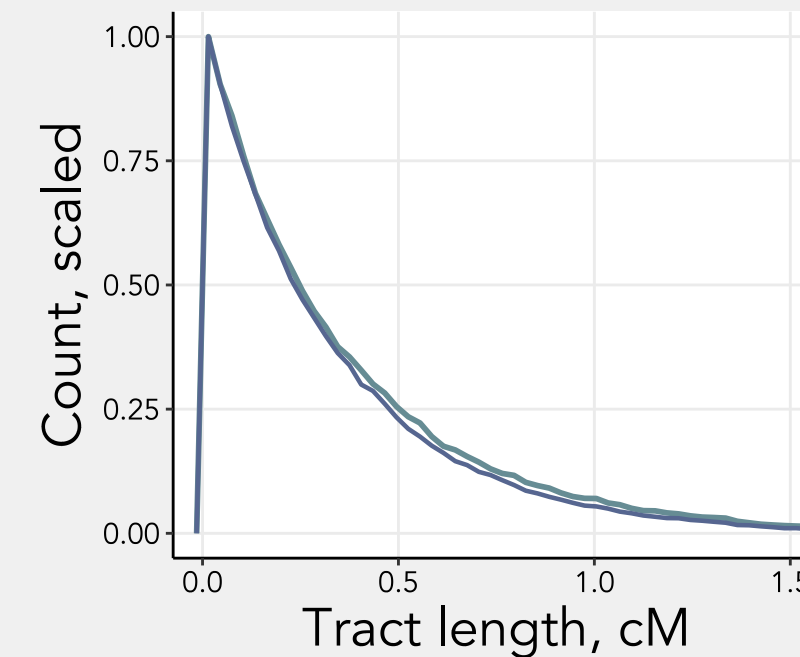


Assortative mating affects metrics of genetic diversity

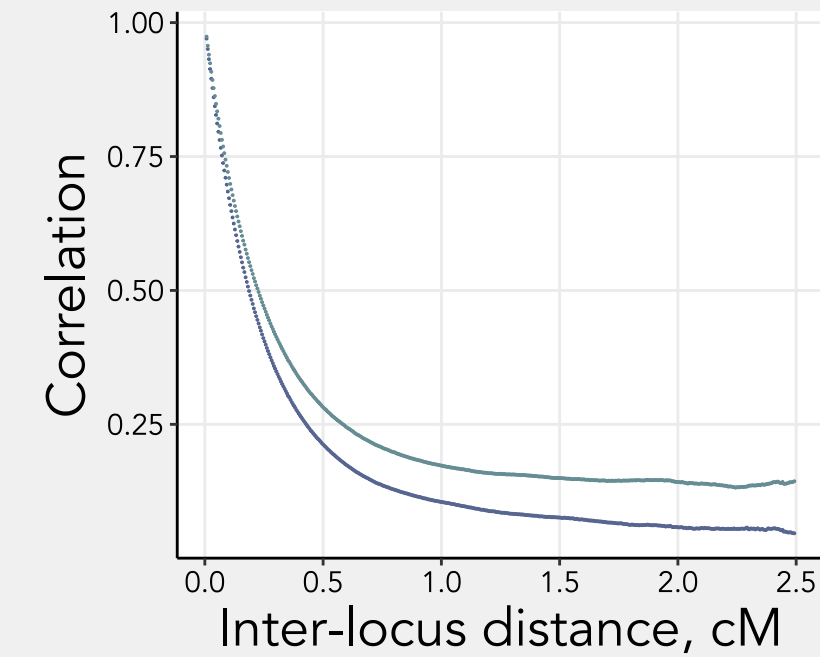
Greater variance in global ancestry



Longer tracts of local ancestry



Long-range linkage disequilibrium



These same metrics are used to infer demographic history and/or selection. Thus, assortative mating may bias such inference, if not accounted for.

The influence of assortative mating on genetic diversity is an under-appreciated potential confounder on demographic inference. Using forward-in-time population genetic simulations, we examine how assortative mating strength and mate-choice function shape expected patterns of global and local ancestry. We find that the effects of assortative mating strength on many population genetic parameters are consistent across replicate simulations. However, mate-choice functions differ substantially in how they impact within-pair correlation in ancestry, the primary metric used to infer assortative mating in humans.

Simulating assortative mating in SLiM in an admixed population

Four functions for defining the probability of two individuals mating:

- Exponential decay
- Normalized exponential decay ← Increases in strength over time
- Normal distribution
- Subpopulation ← Does not explicitly use genetic ancestry

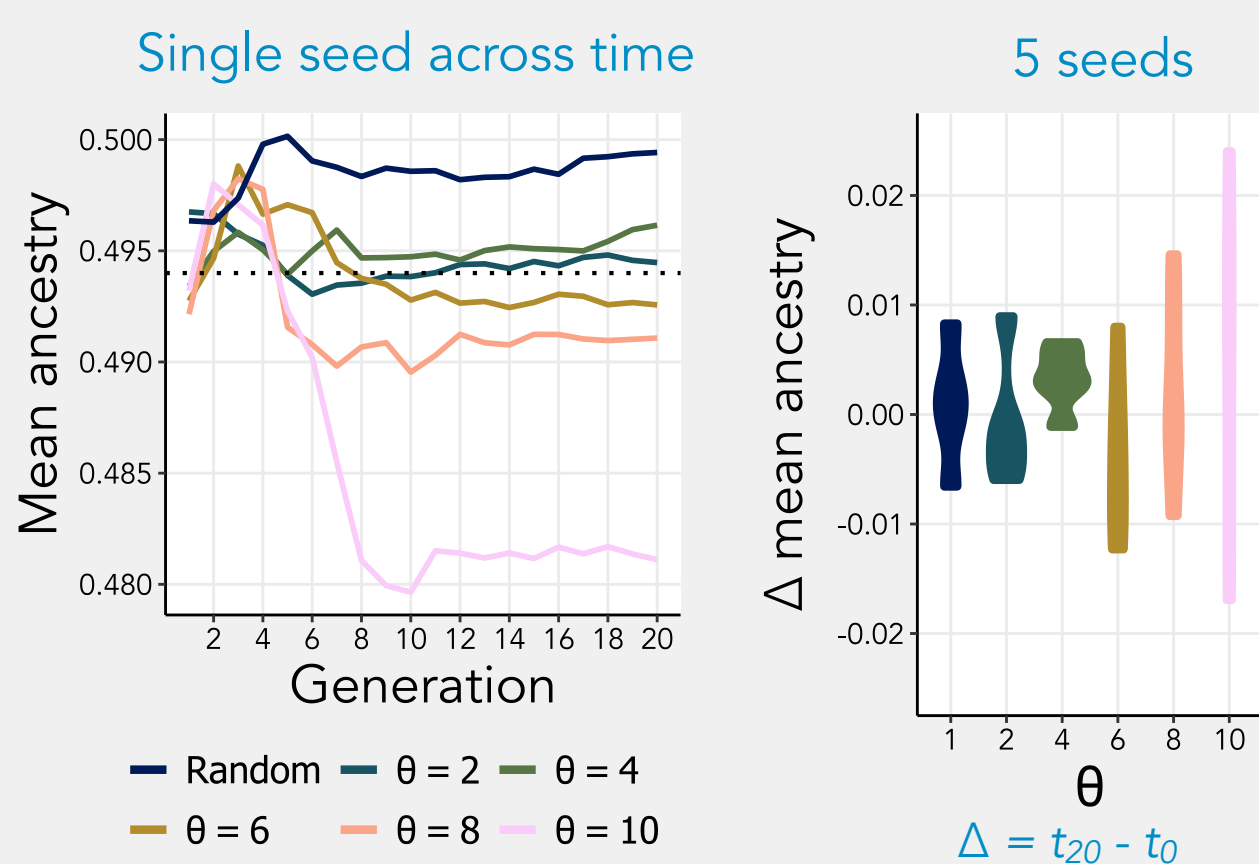
θ

Probability of mating with most-preferred mate over least-preferred mate

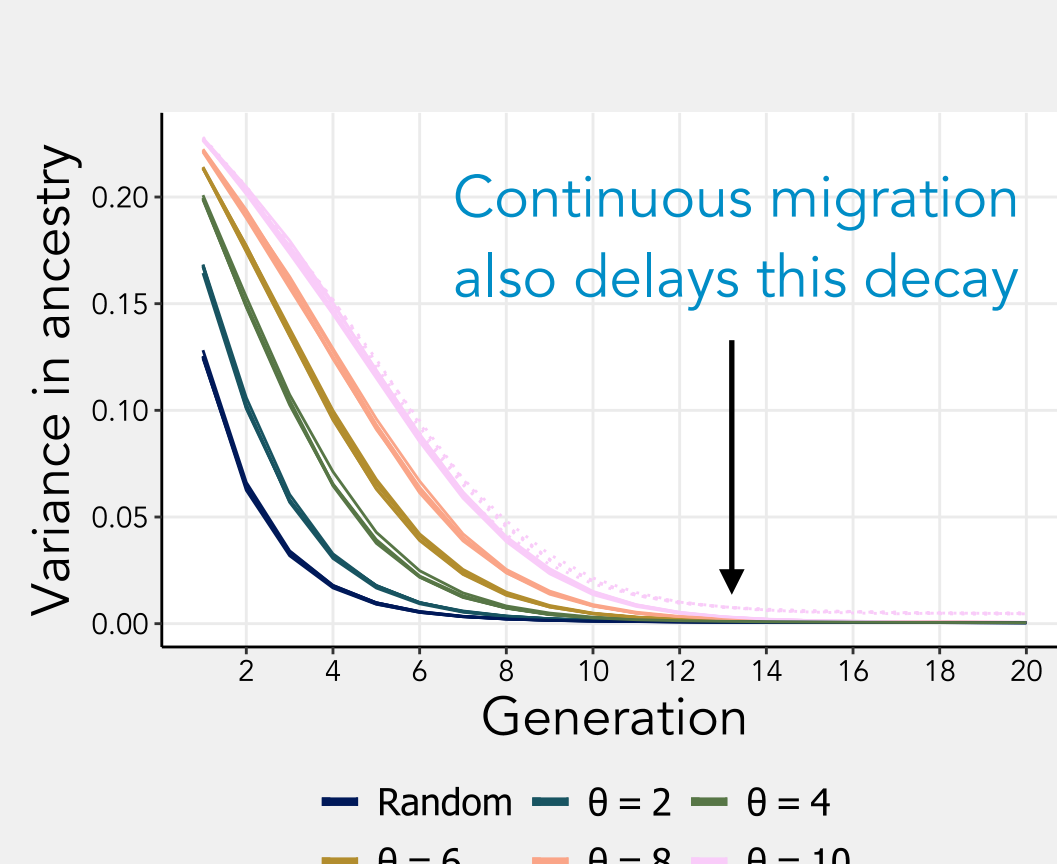
- Forward-in-time individual-based simulations in SLiM
- 22 chromosomes, modeled on sizes of the human autosomes
- Uniform recombination rate of 1×10^{-8}
- 10,000 individuals per generation, for 20 generations
- Single-pulse and continuous migration models

Assortative mating strength alters global ancestry dynamics

Mean ancestry tends to drift more under stronger assortative mating

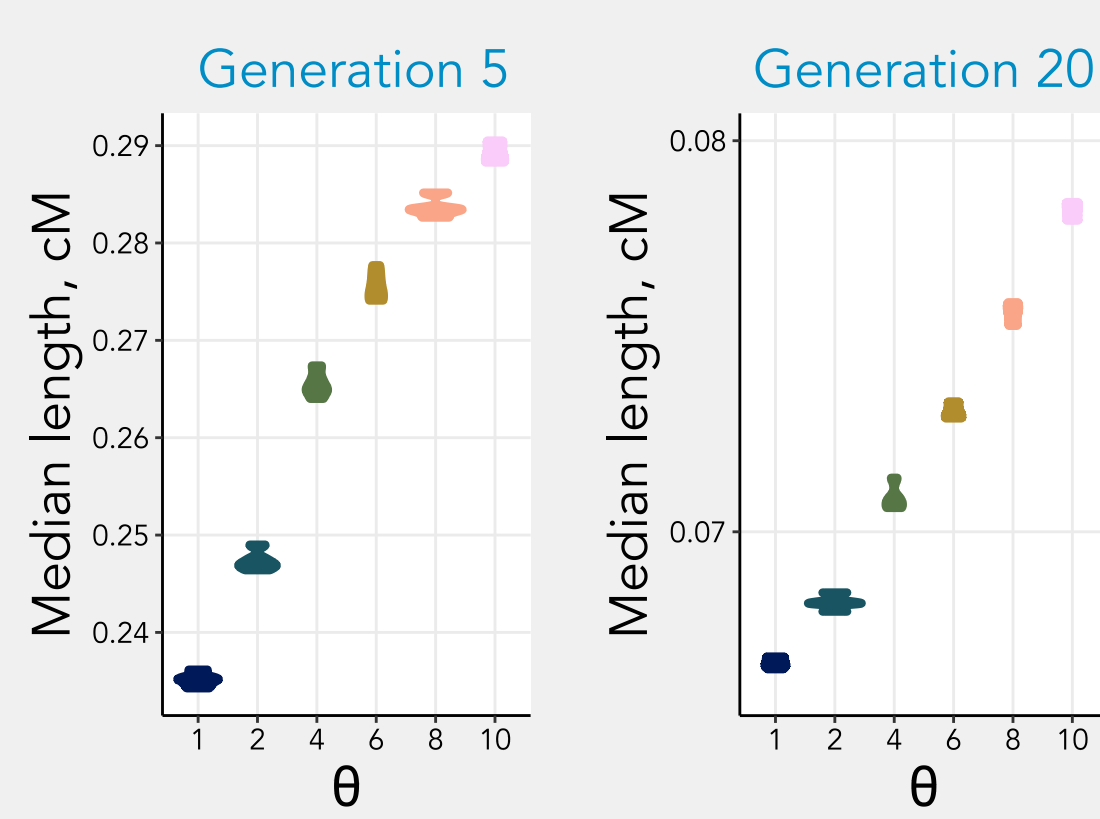


Assortative mating attenuates the decay in ancestry variance



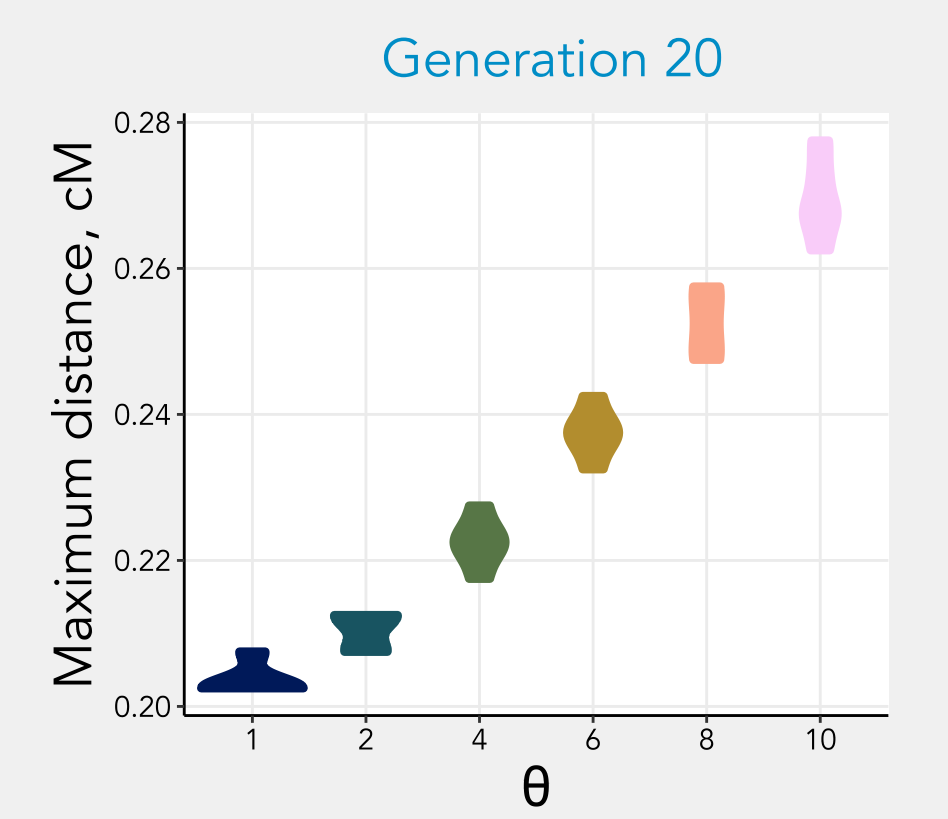
Local ancestry tracts scale with assortative mating strength

Median local ancestry tract length scales linearly with θ



Differences between θ values is greatest ~7 generations

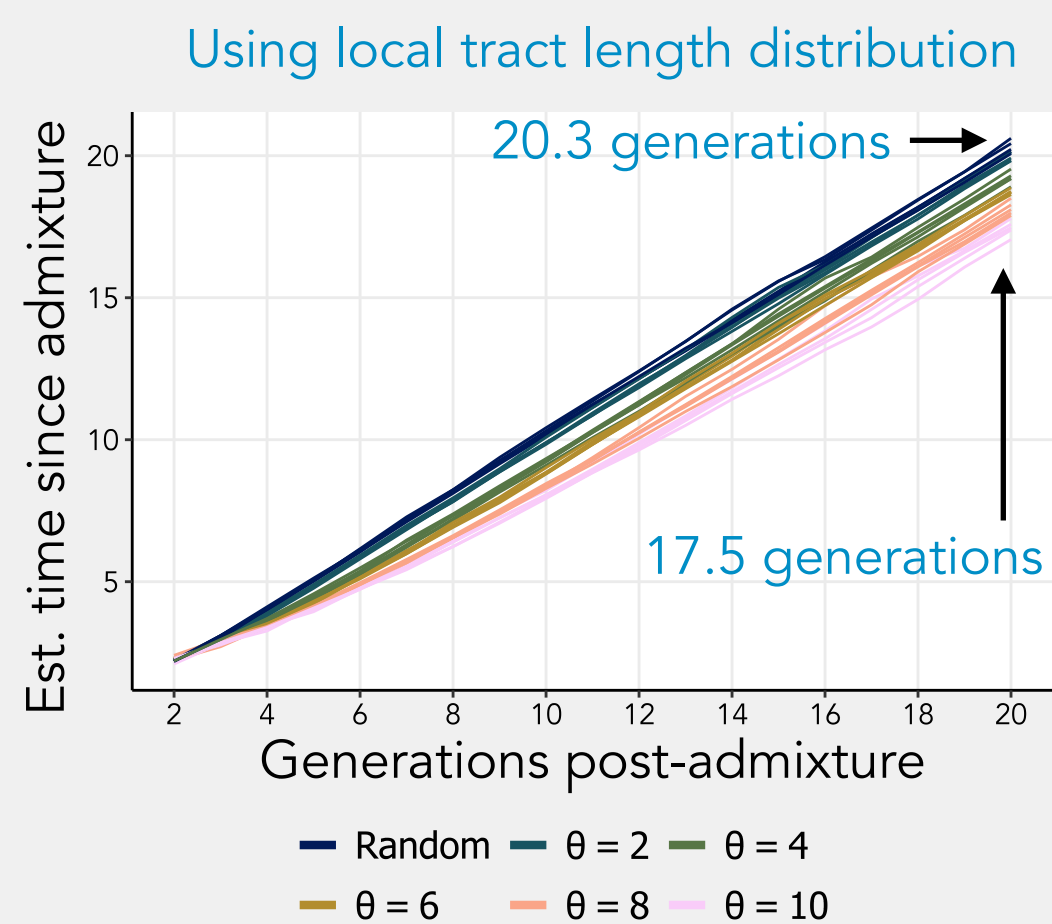
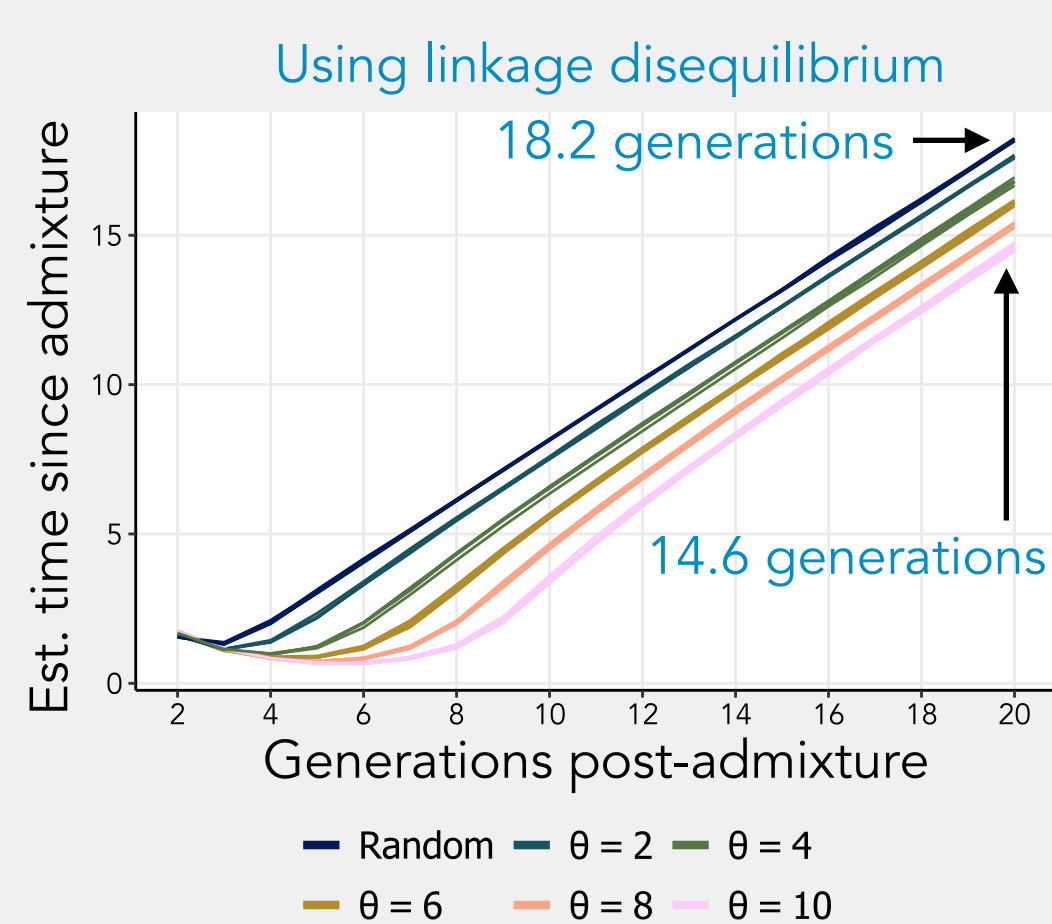
Maximum distance of alleles in linkage disequilibrium scales with θ



LD = correlation between alleles ≥ 0.03

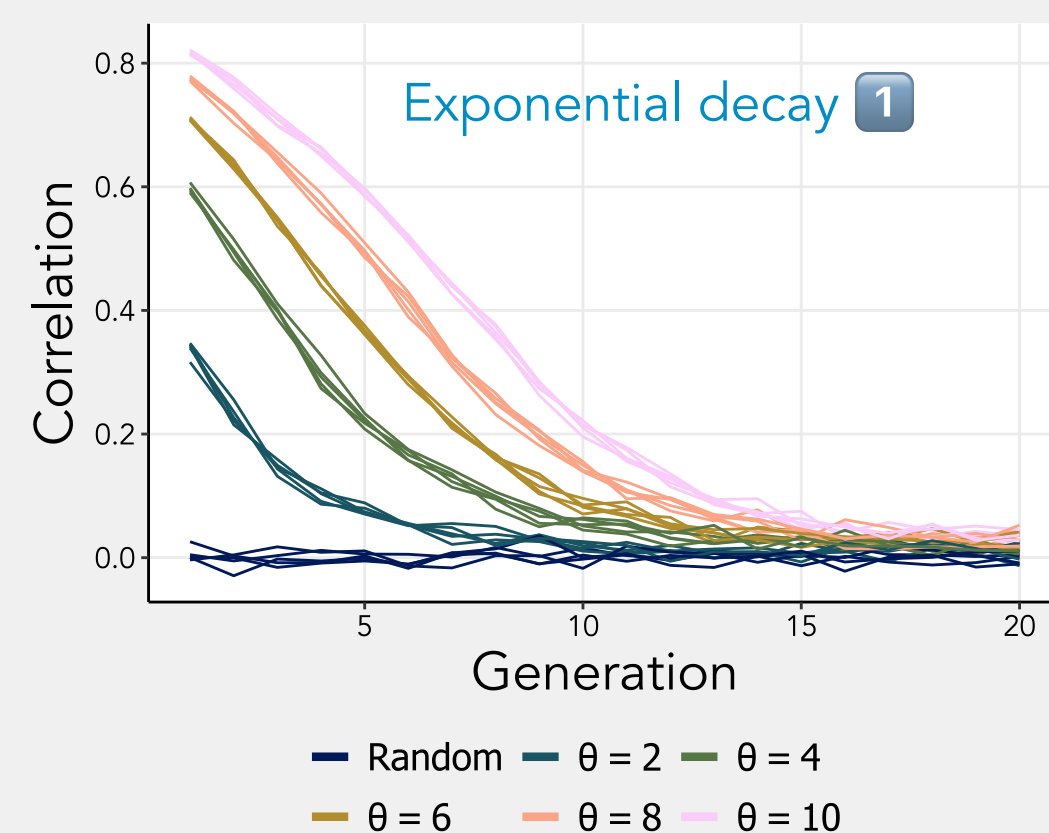
Assortative mating biases estimated time since admixture

Admixture is estimated to be more recent under stronger assortative mating



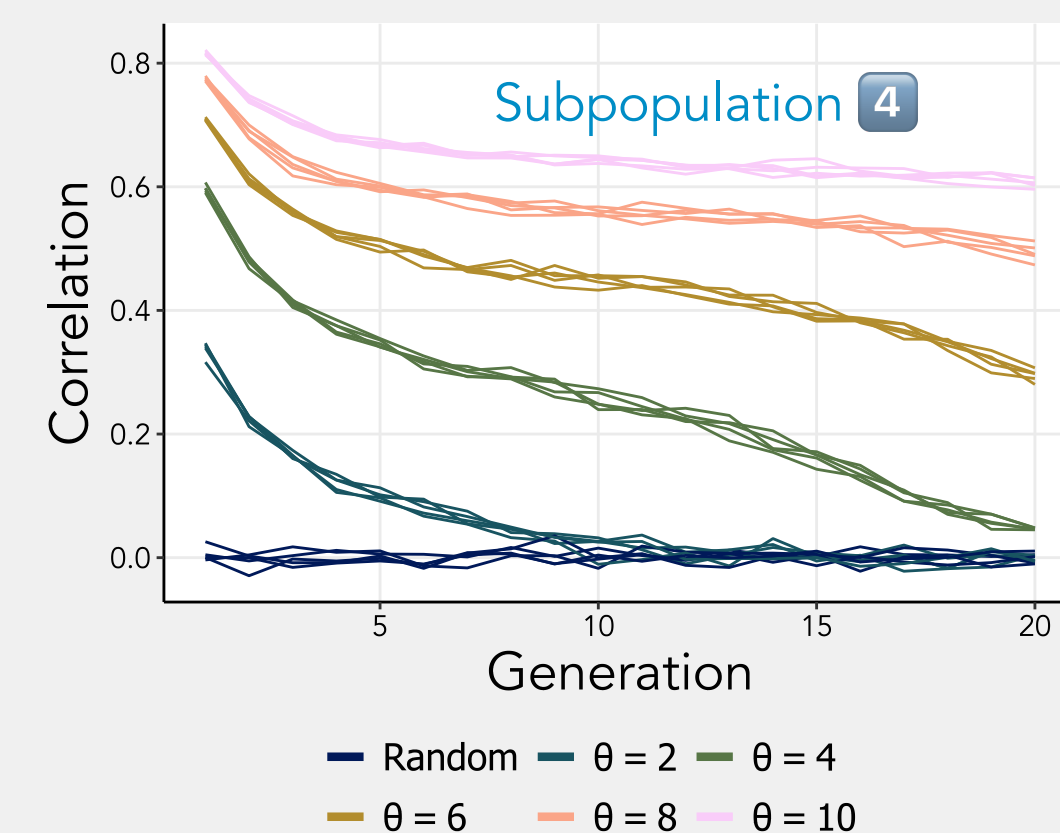
Correlation in ancestry maintained only in some models

Correlation decays to 0



Same result for normal distribution 3

Correlation depends on θ



Same result for normalized exponential decay 2

Exponential decay 1

$$\Psi(x_1, x_2) = \alpha^{-2|x_1 - x_2|}$$

Normalized exponential decay 2

$$\Psi(x_1, x_2) = e^{-\frac{2 \ln \alpha |x_1 - x_2|}{\alpha}}$$

Normal distribution 3

$$\Psi(x_1, x_2) = f(x_2 | x_1, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_2 - x_1)^2}{2\sigma^2}}$$

Subpopulation 4

$$\Psi(s_1, s_2) = \begin{cases} 1 - \frac{1}{\theta + 1} & : s_1 = s_2 \\ \frac{1}{\theta + 1} & : s_1 \neq s_2 \end{cases}$$