

Outline for today

1. Measures of variation
2. Population structure
3. Natural selection

Matthew Hahn
mwh@iu.edu
@3rdreviewer

All evolution begins as one mutation, on one chromosome, in one individual

TTACAATCCGATCGT

TTACGATGCGCTCGT

TCACAATGCGATGGA

TTACGATGCGCTCGT

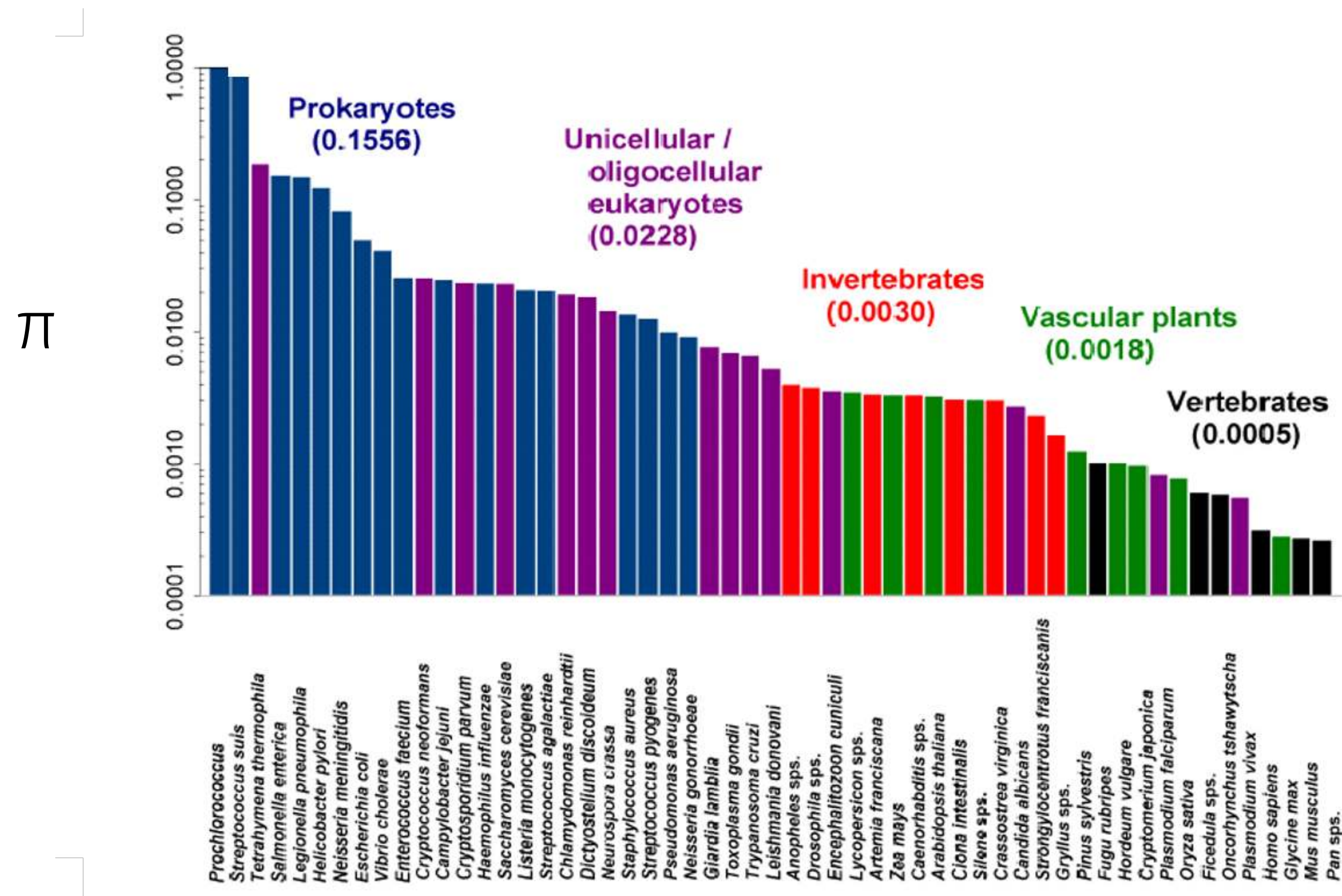
TTACAATCCGATCGT

--ACGATGCGCTCGT

TCACAATGCGATGGA

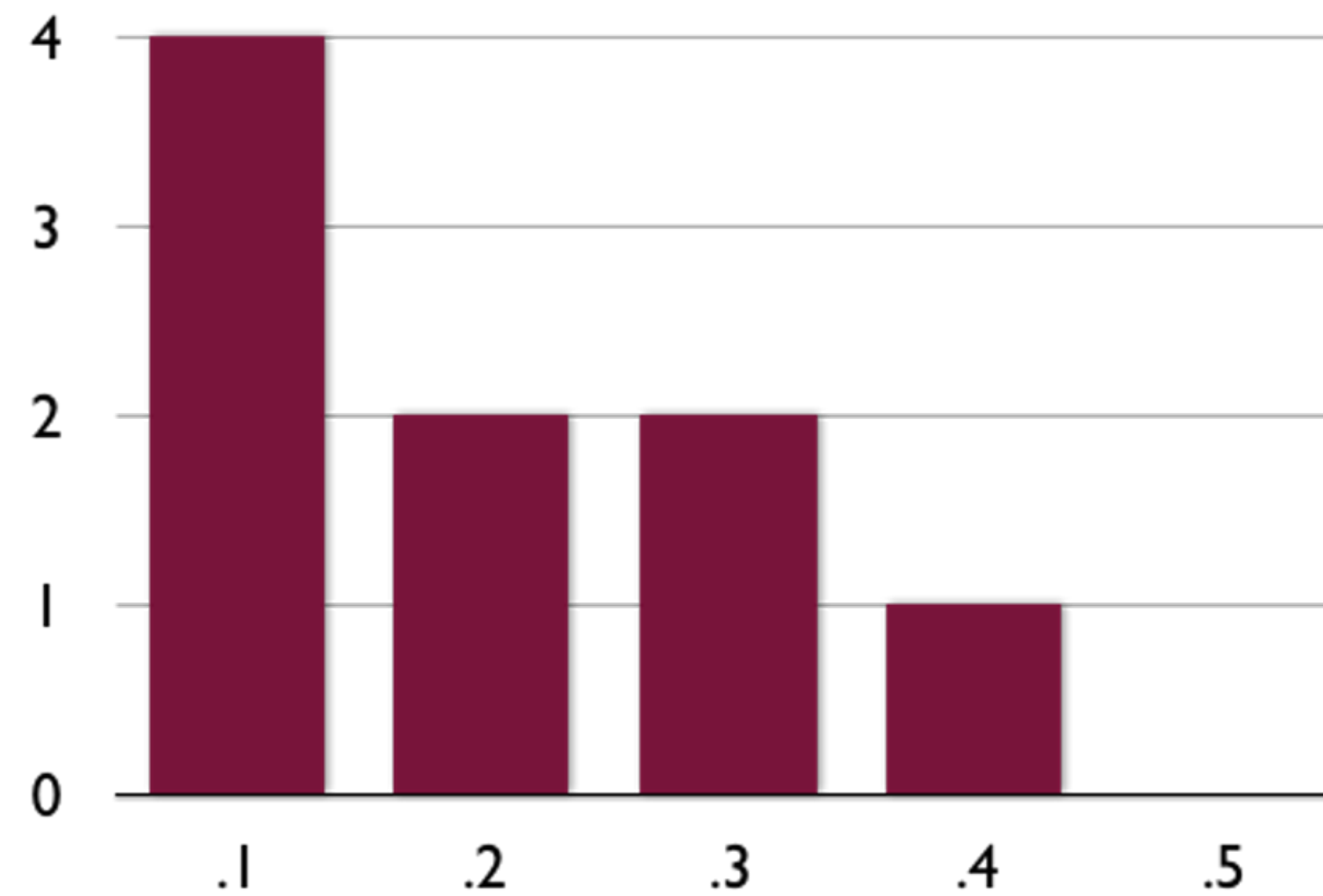
TTACGATG--CTCGT

Levels of diversity in different species



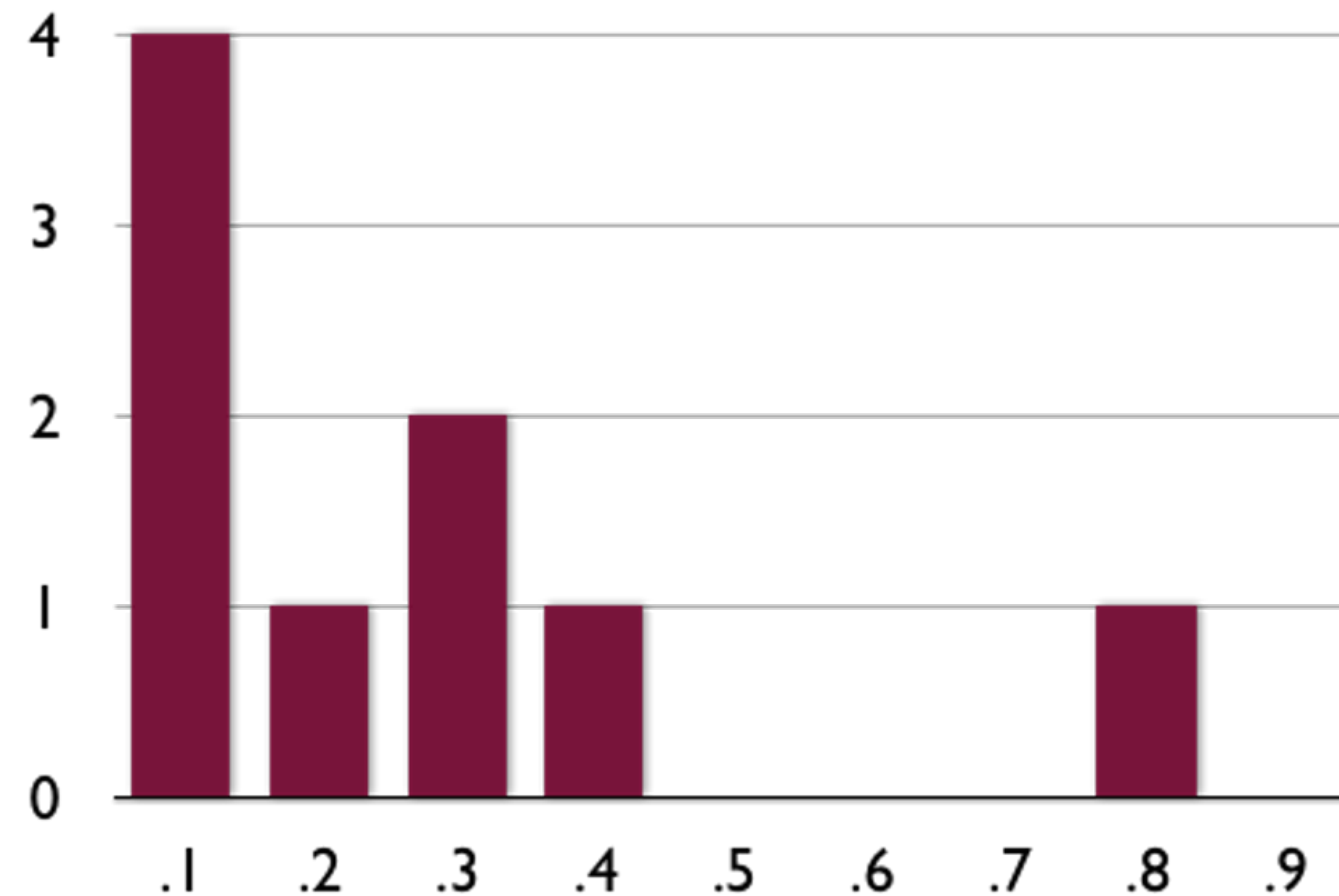
1 GCTCACCGGAATTATCCGATATGCTAGTA
2 GCTTACCGGAATTATGCGATATGCTTGTA
3 GCTCACCGGAATTATGCGATATGGTAGAA
4 GCTCACCGGAATTATGCGATATGGTAGAA
5 GCTCACCGGGATGATGCGATATGCTAGTA
6 GCTCACCGGAATTATGCGATATGCTAGAA
7 GCTTACCGGAATTATCCGATATGCTAGTA
8 GCTCACAGGGATTATGCGCTATGCTAGTA
9 GCTCACCGGAATTATGCGATATGGTAGAA
10 GCTCACCGGAATTATCCGATATGCTAGTA

Allele frequency spectrum (Minor allele frequency)

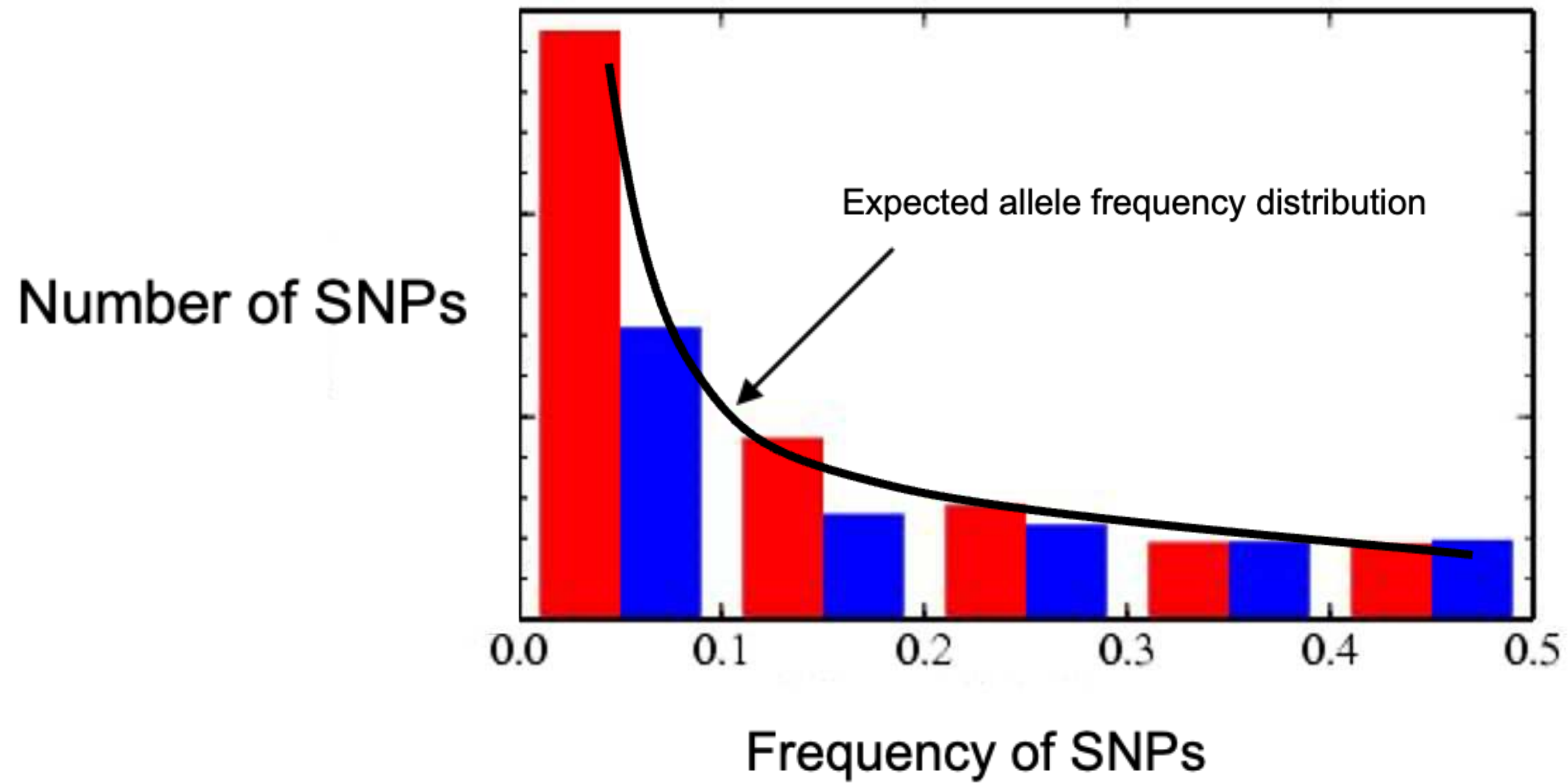


1 GCTCACCGGAATTATCCGATATGCTAGTA
2 GCTTACCGGAATTATGCGATATGCTTGTA
3 GCTCACCGGAATTATGCGATATGGTAGAA
4 GCTCACCGGAATTATGCGATATGGTAGAA
5 GCTCACCGGGATGATGCGATATGCTAGTA
6 GCTCACCGGAATTATGCGATATGCTAGAA
7 GCTTACCGGAATTATCCGATATGCTAGTA
8 GCTCACAGGGATTATGCGCTATGCTAGTA
9 GCTCACCGGAATTATGCGATATGGTAGAA
10 GCTCACCGGAATTATCCGATATGCTAGTA

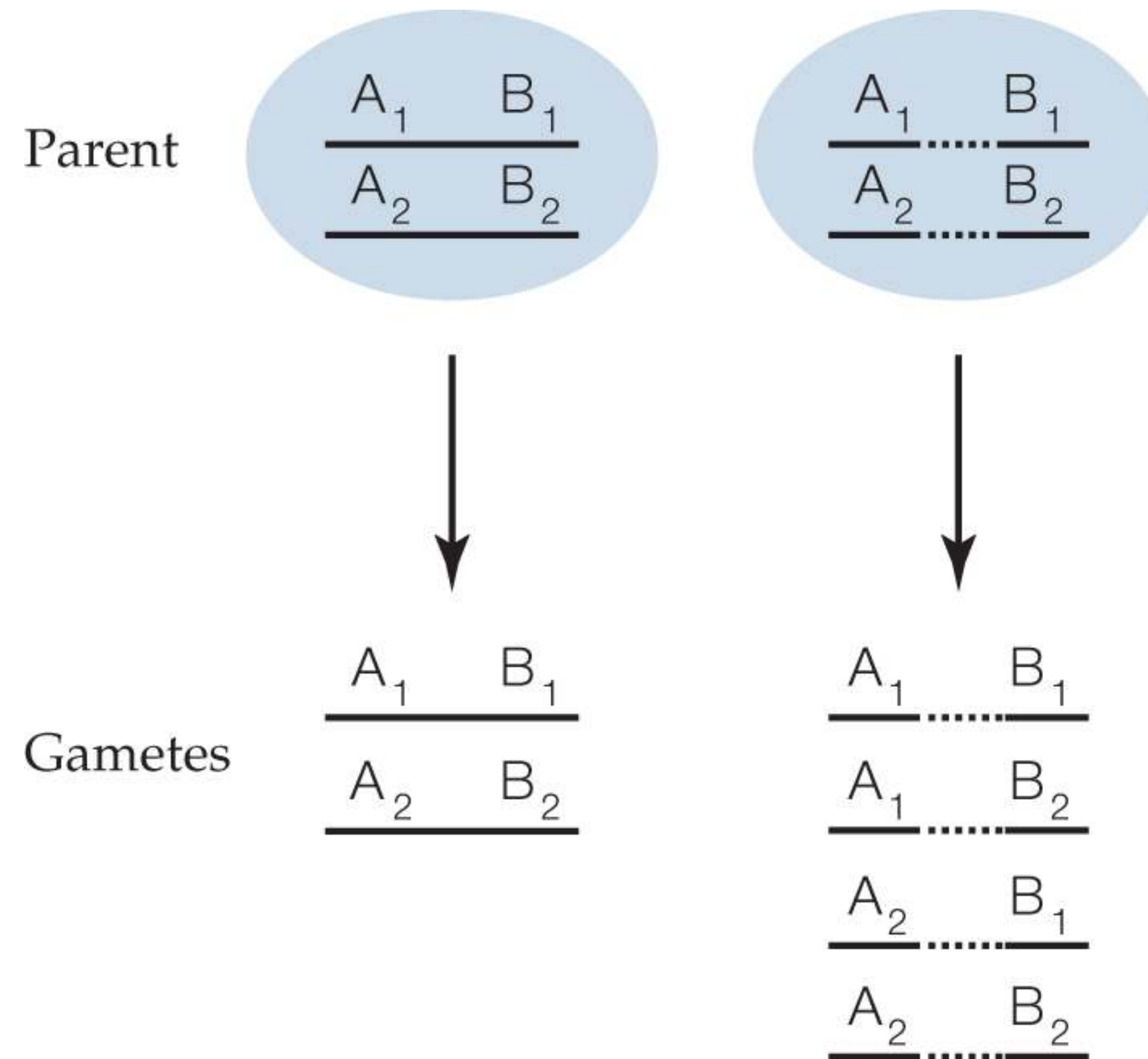
Allele frequency spectrum (Derived allele frequency)



Allele frequency spectrum



Linkage



Linkage disequilibrium

$A_1 \quad B_1$

$A_1 \quad B_1$

$A_1 \quad B_1$

$A_1 \quad B_1$

$A_1 \quad B_1$

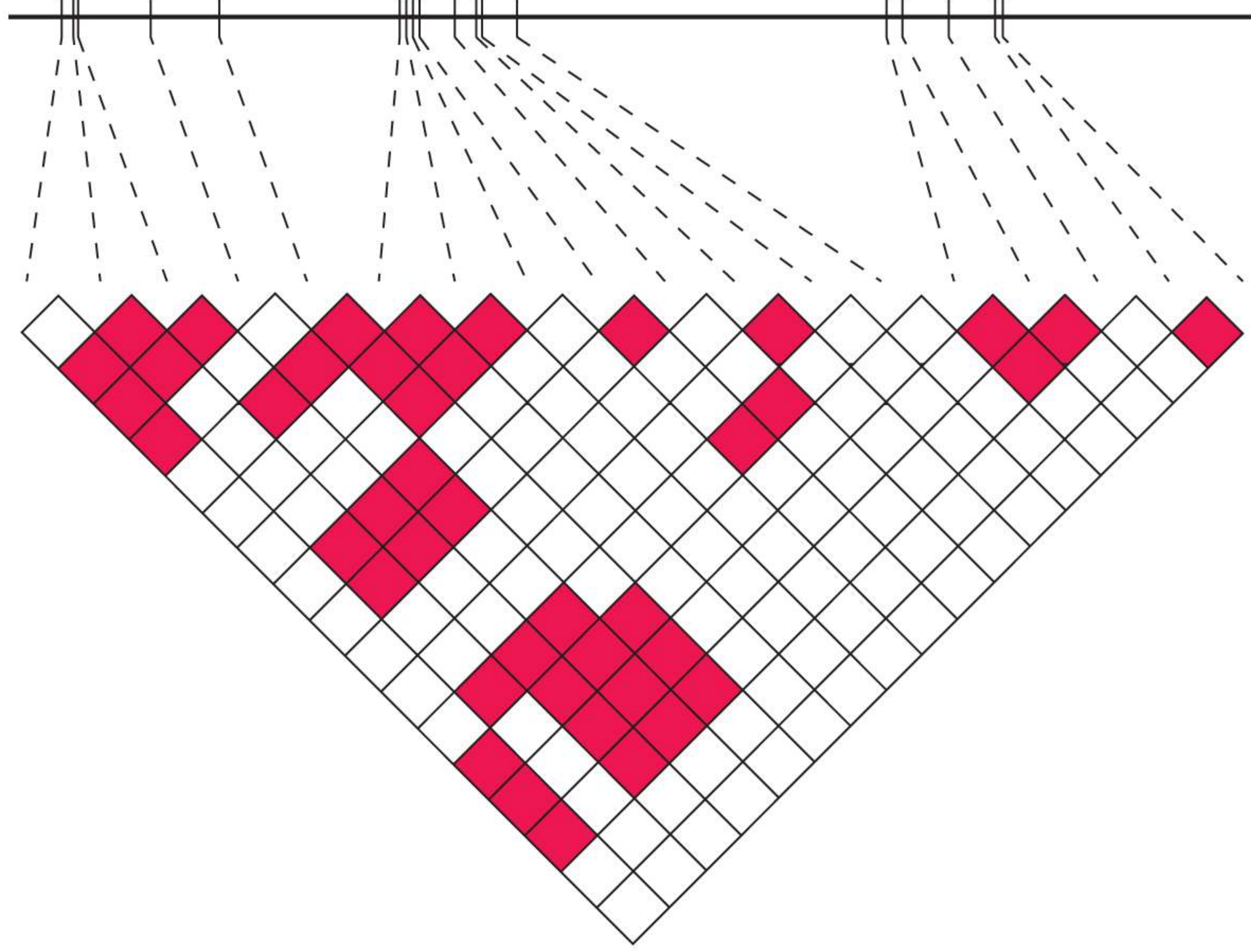
$A_2 \quad B_1$

$A_2 \quad B_1$

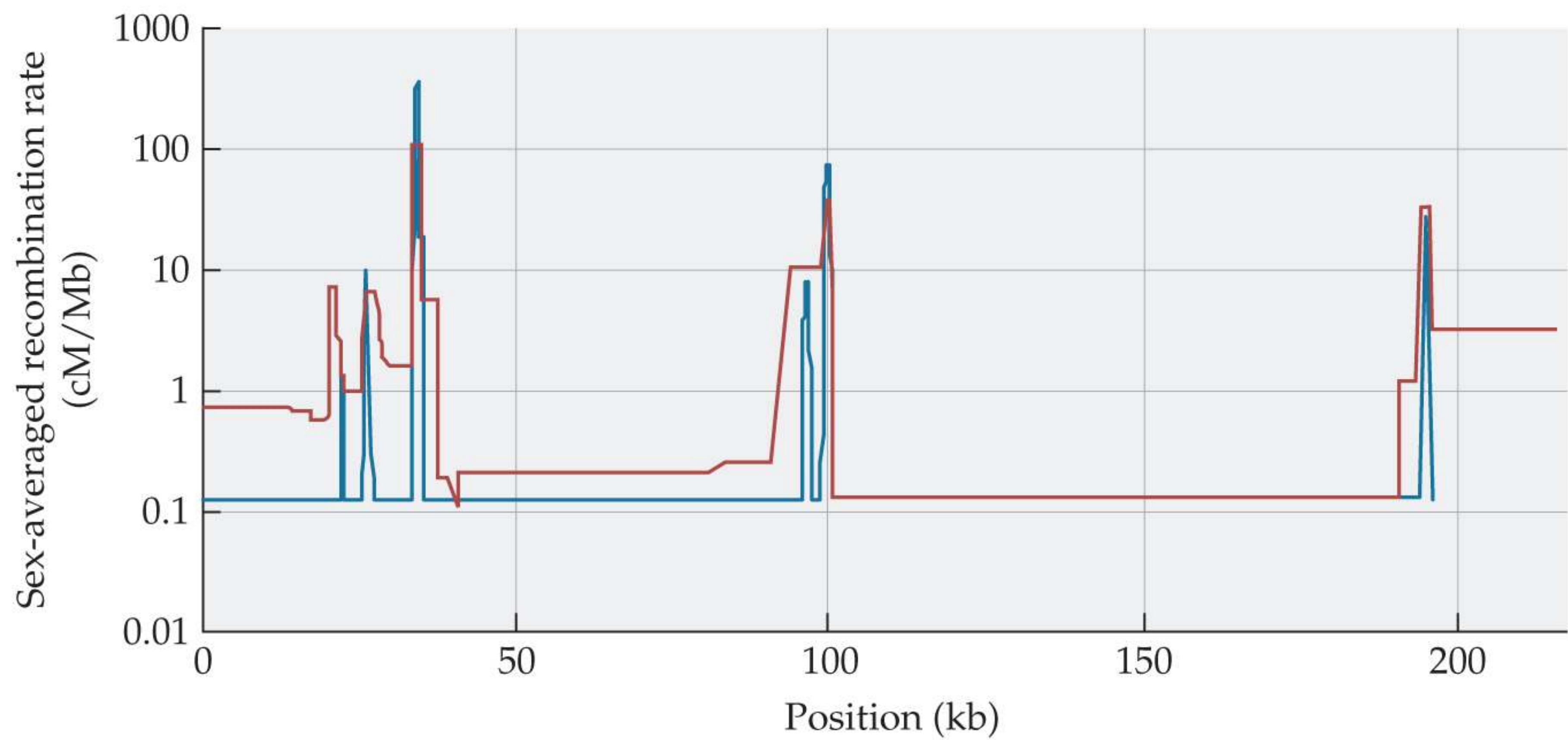
$A_2 \quad B_1$

$A_2 \quad B_1$

$A_2 \quad B_2$



From Langley et al. 2000. *Genetics* 156:1837-1852.



After McVean et al., 2004. *Science* 304:581-584.

Outline for today

1. Measures of variation
- 2. Population structure**
3. Natural selection

Matthew Hahn
mwh@iu.edu
@3rdreviewer

Evolution:

change in allele frequencies over time

Forces that change allele frequencies:

- mutation
- genetic drift
- migration
- natural selection

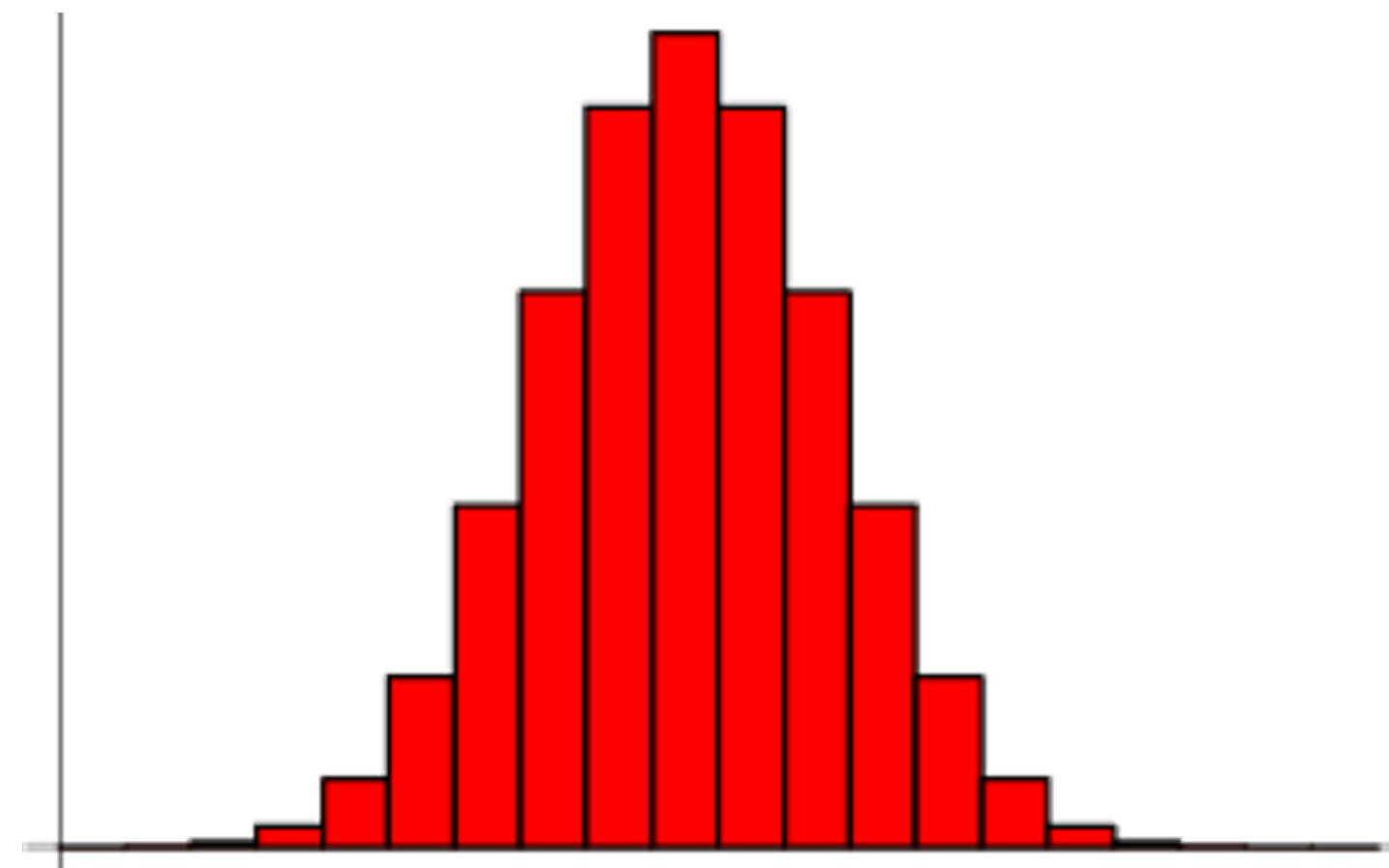
Forces that change genotype/haplotype frequencies:

- mating
- recombination

Genetic drift:

change in allele frequencies due to sampling in a finite population

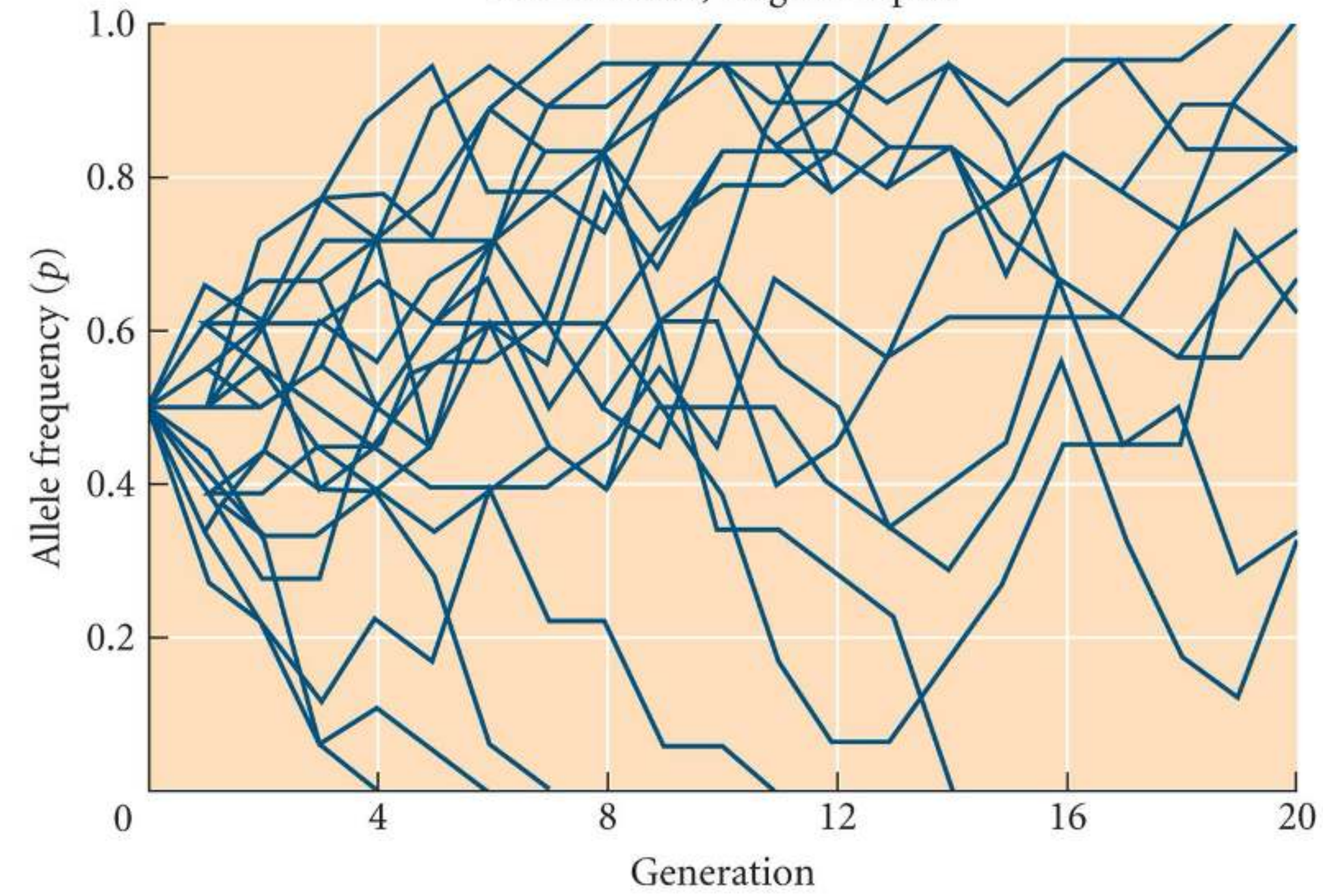
$$\text{---} V = \frac{2pq}{2N_e} \text{---}$$



$2N=20$ and $p=0.5$

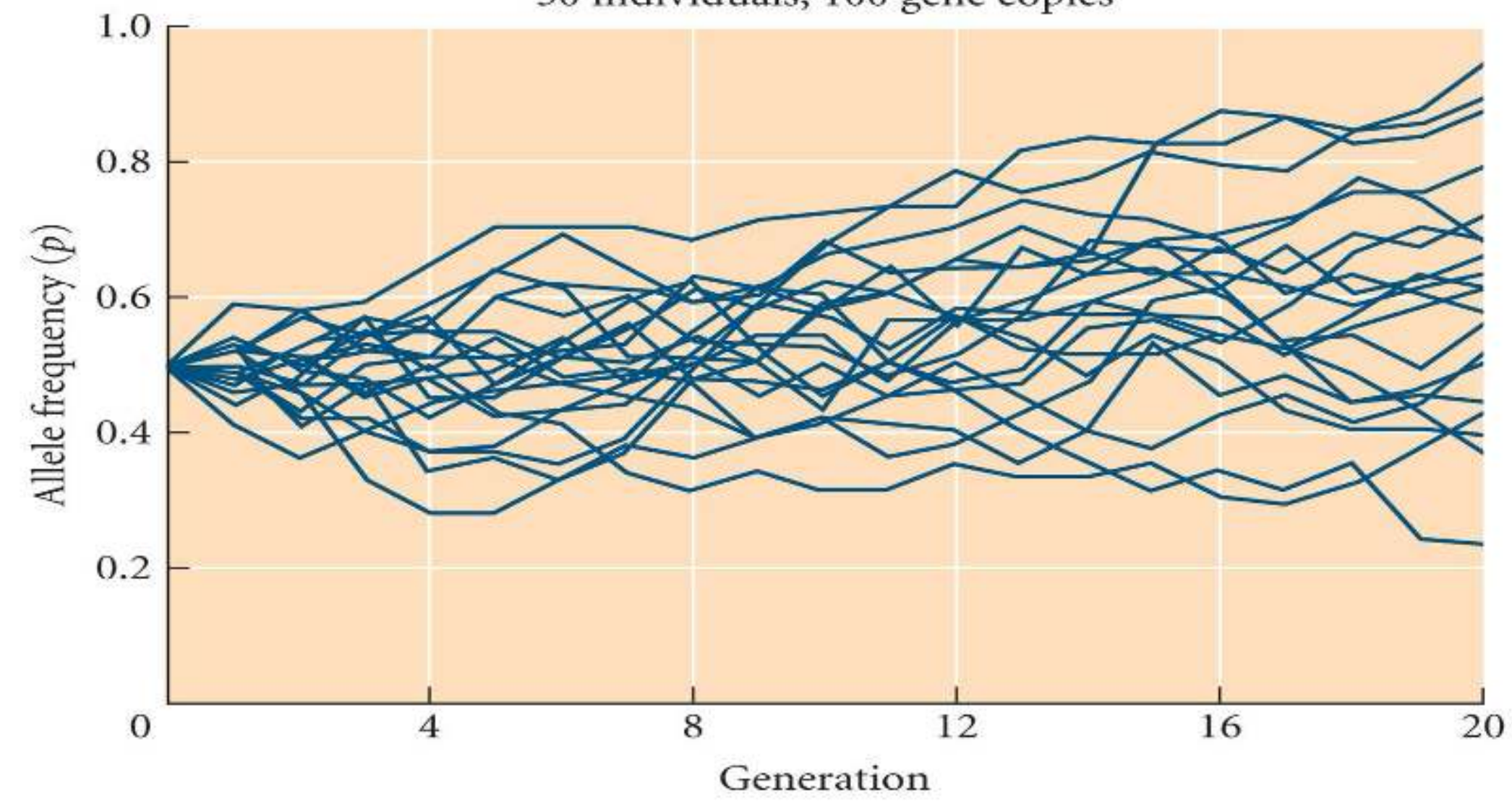
(A)

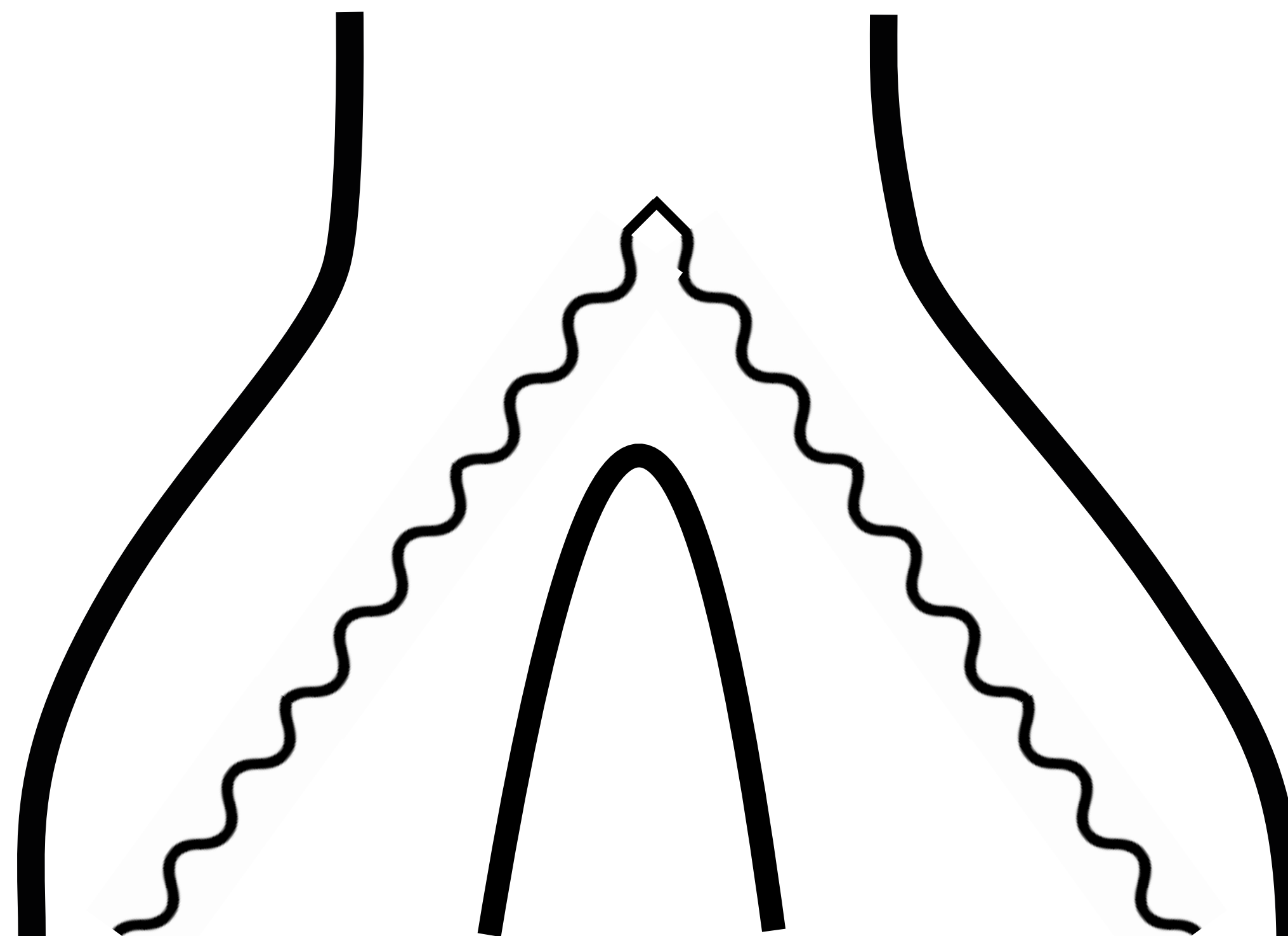
9 individuals, 18 gene copies



(B)

50 individuals, 100 gene copies





$p=0.8$

A

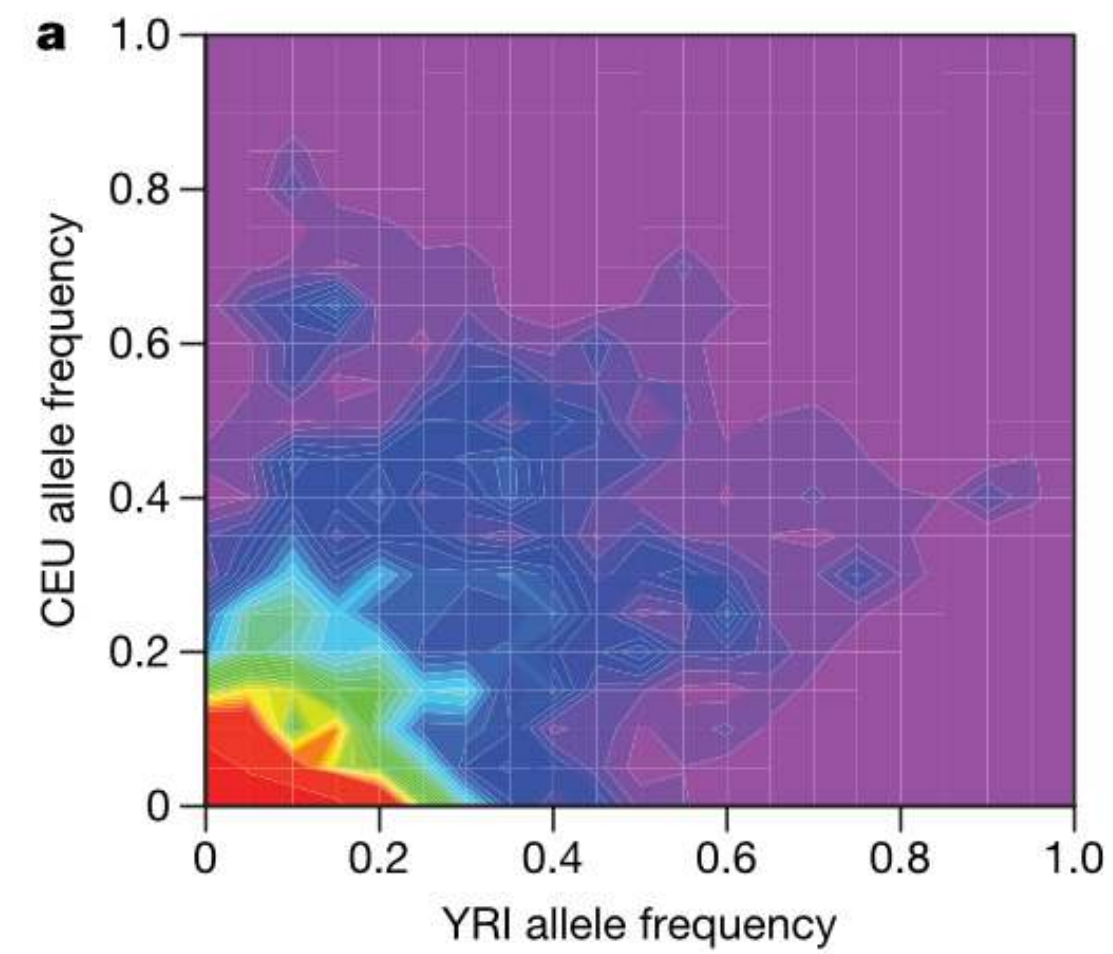
$p=0.2$

B

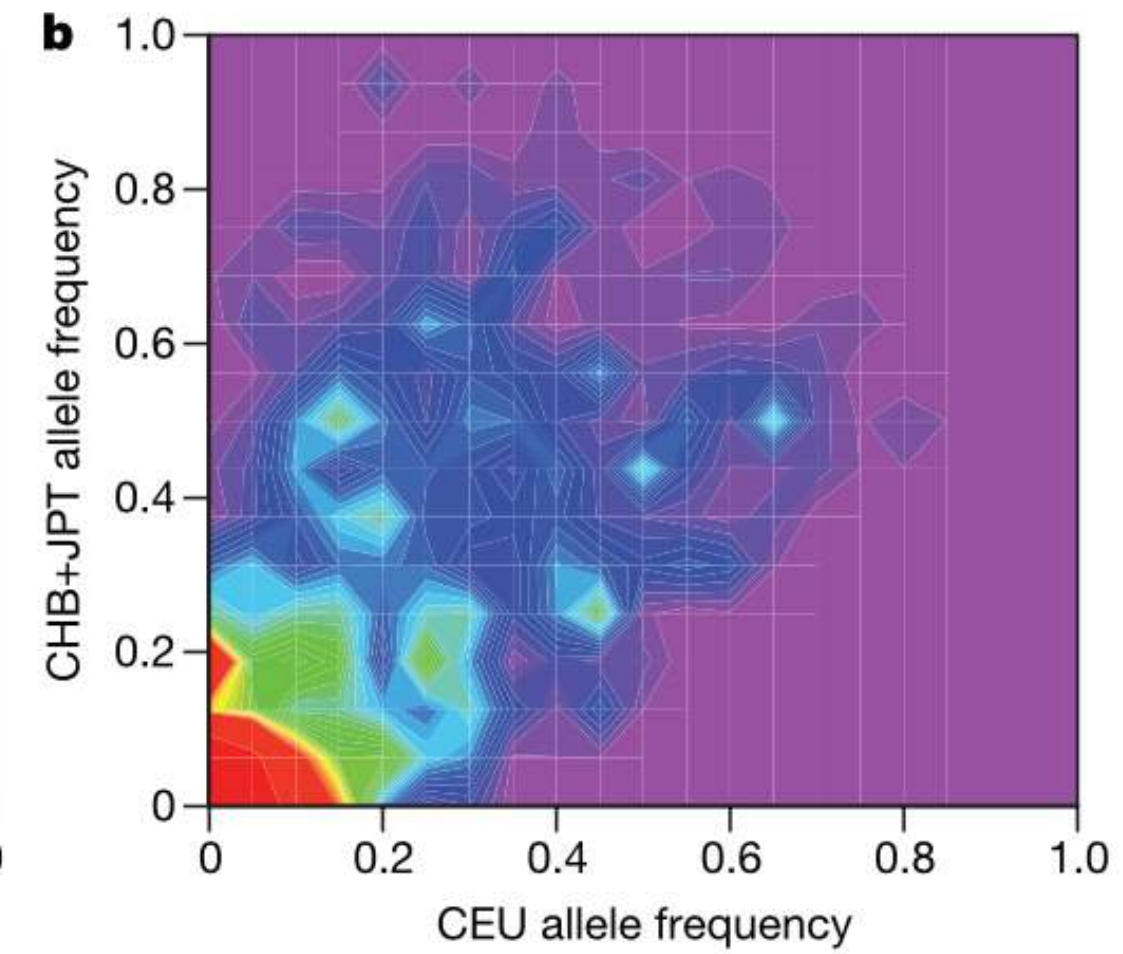
$T=t/2N_e$

1	GCTCACCGGAATTATCCGATATGCTAGTA	}	A
2	GCTTACCGGAATTATGCGATATGCTTGTA		
3	GCTCACCGGAATTATGCGATATGGTAGAA		
4	GCTCACCGGAATTATGCGATATGGTAGAA		
5	GCTCACCGGGATGATGCGATATGCTAGTA	}	B
6	GCTCACCGGAATTATGCGATATGCTAGAA		
7	GCTTACCGGAATTATCCGATATGCTAGTA		
8	GCTCACAGGGATTATGCGCTATGCTAGTA		
9	GCTCACCGGAATTATGCGATATGGTAGAA		
10	GCTCACCGGAATTATCCGATATGCTAGTA		

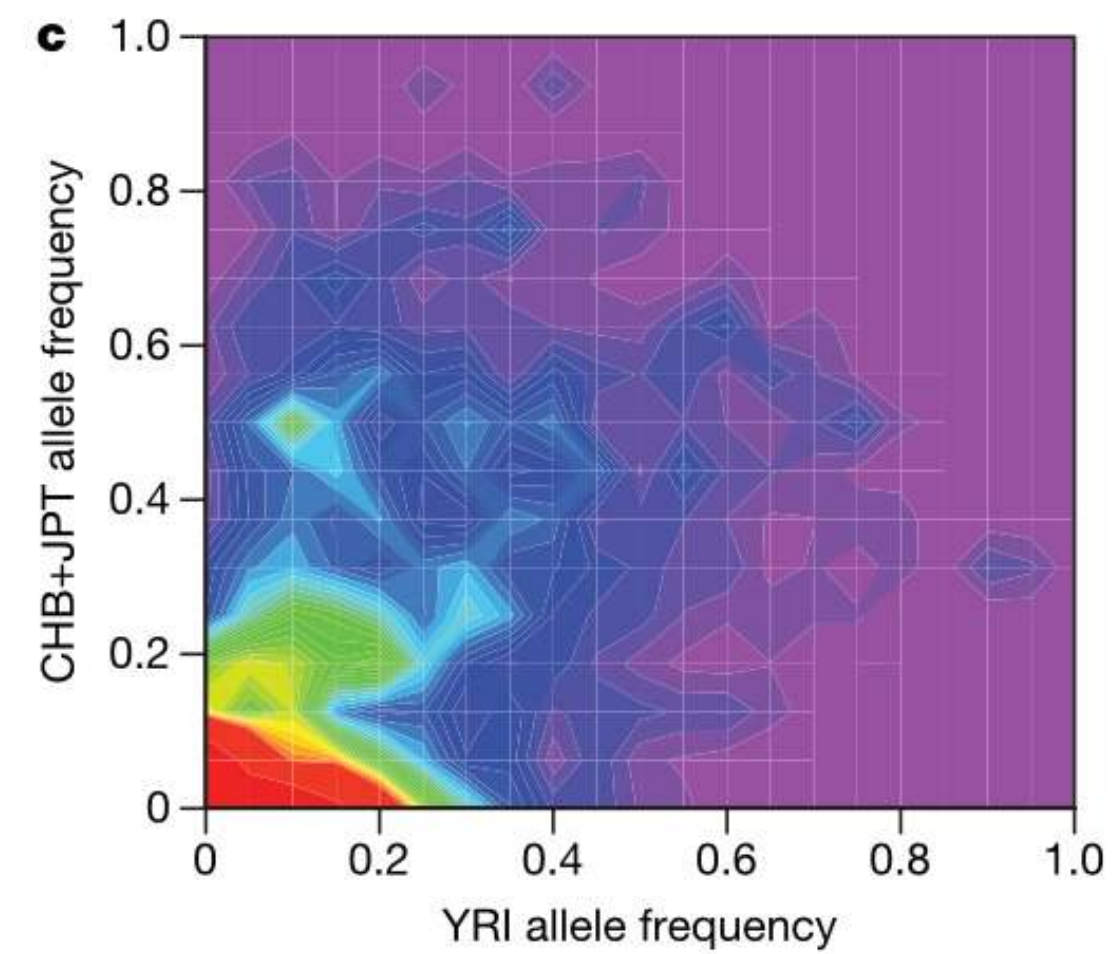
$$F_{ST}=0.11$$



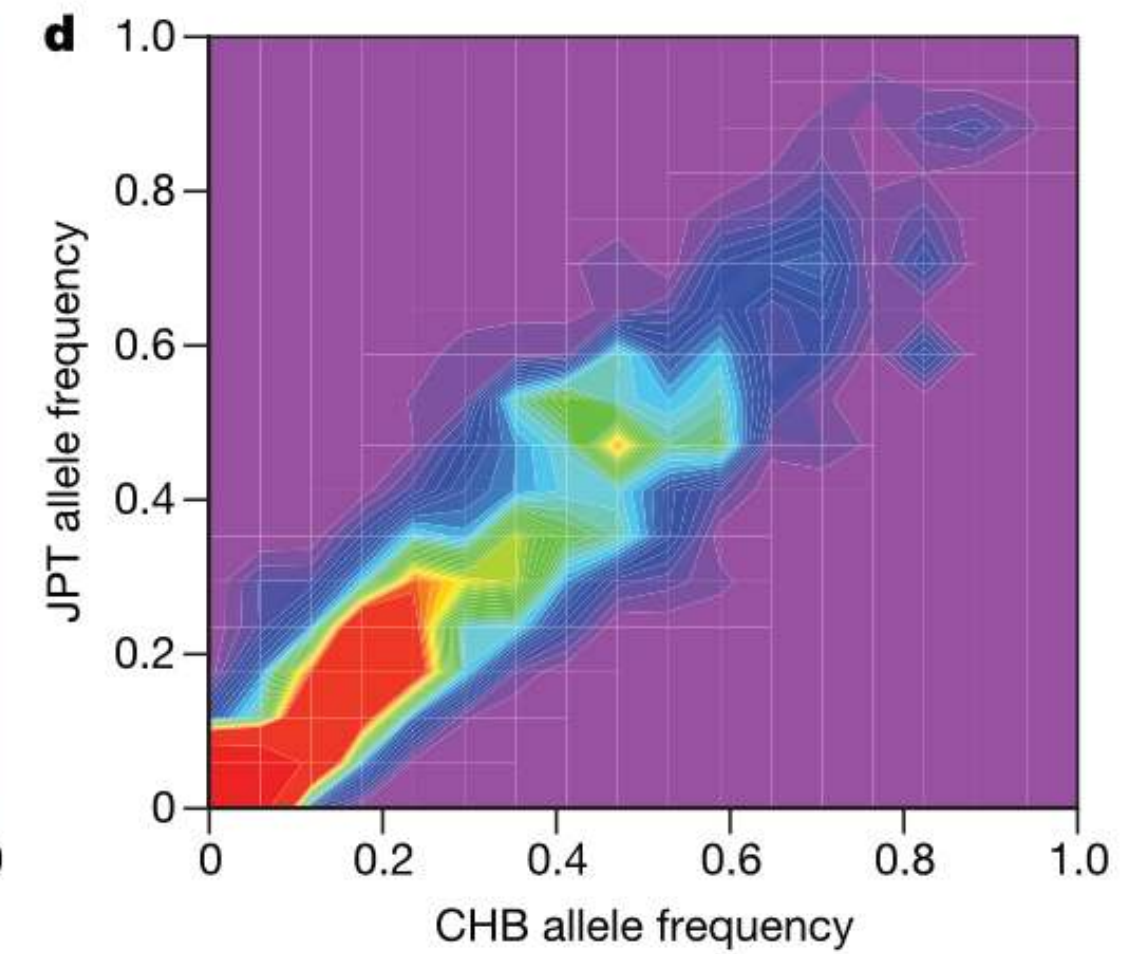
$$F_{ST}=0.07$$



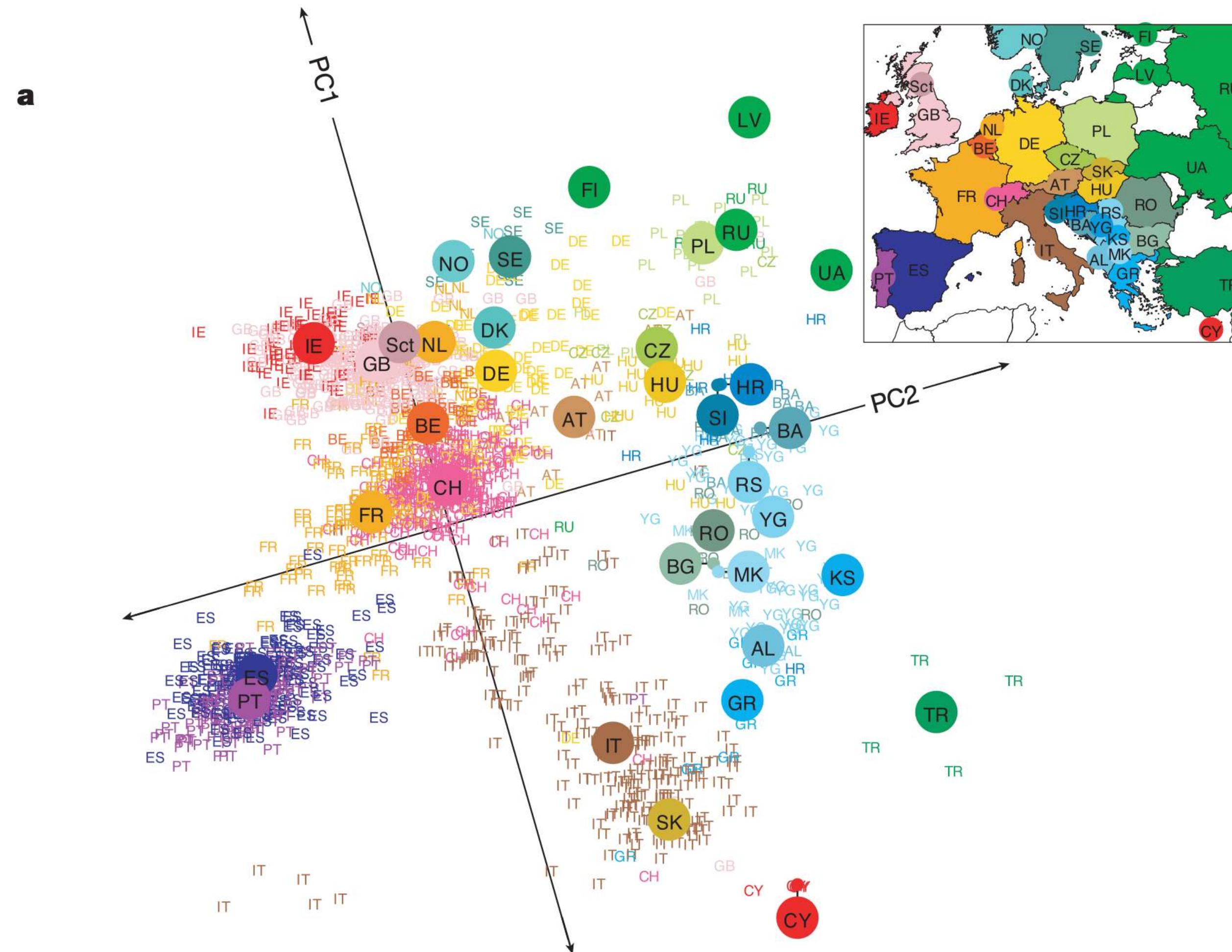
$$F_{ST}=0.12$$



$$F_{ST}=0.01$$

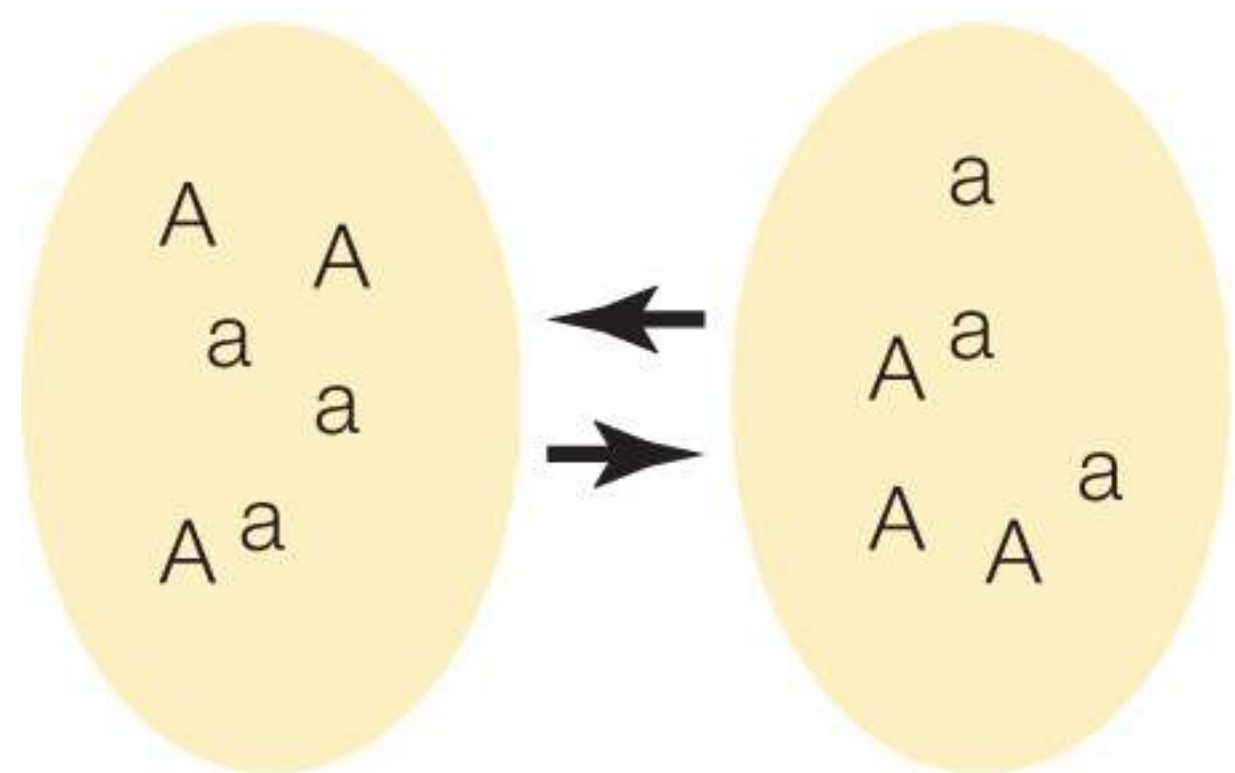


HapMap (2005)



Novembre et al. (2008)

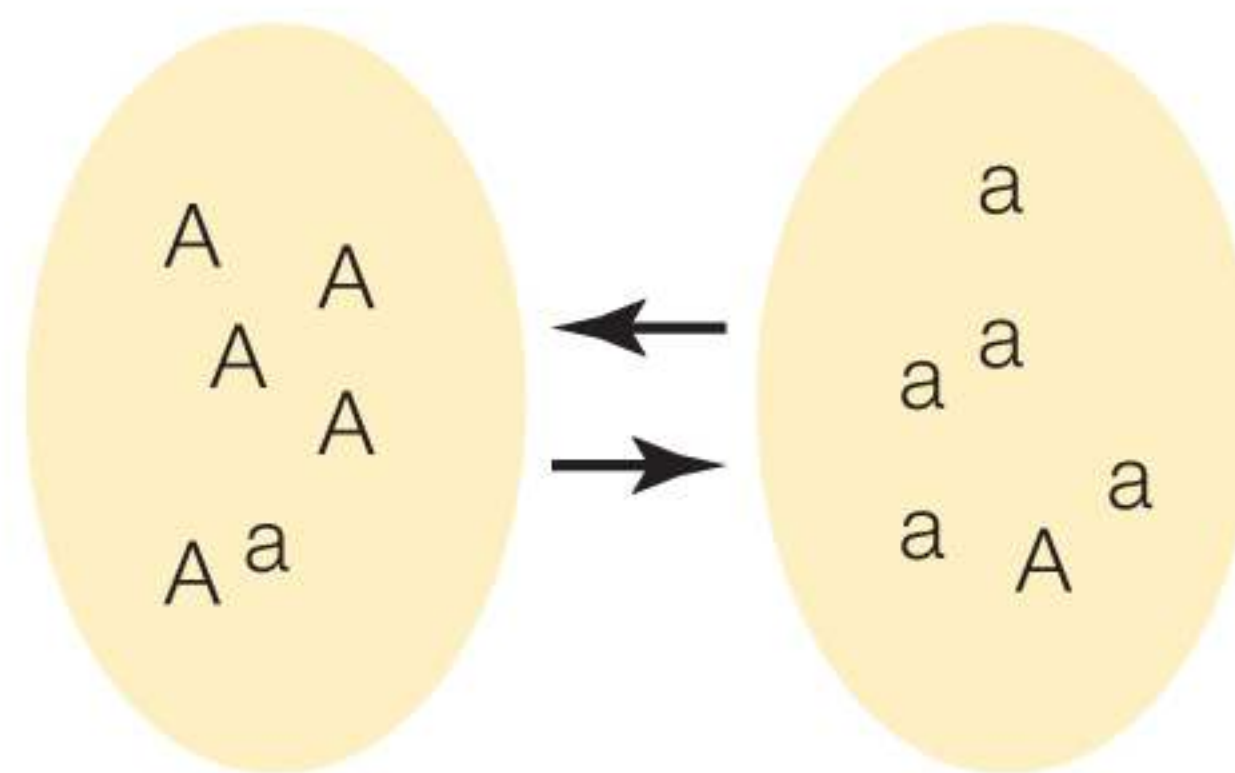
(A) $F_{ST} = 0$



Population 1

Population 2

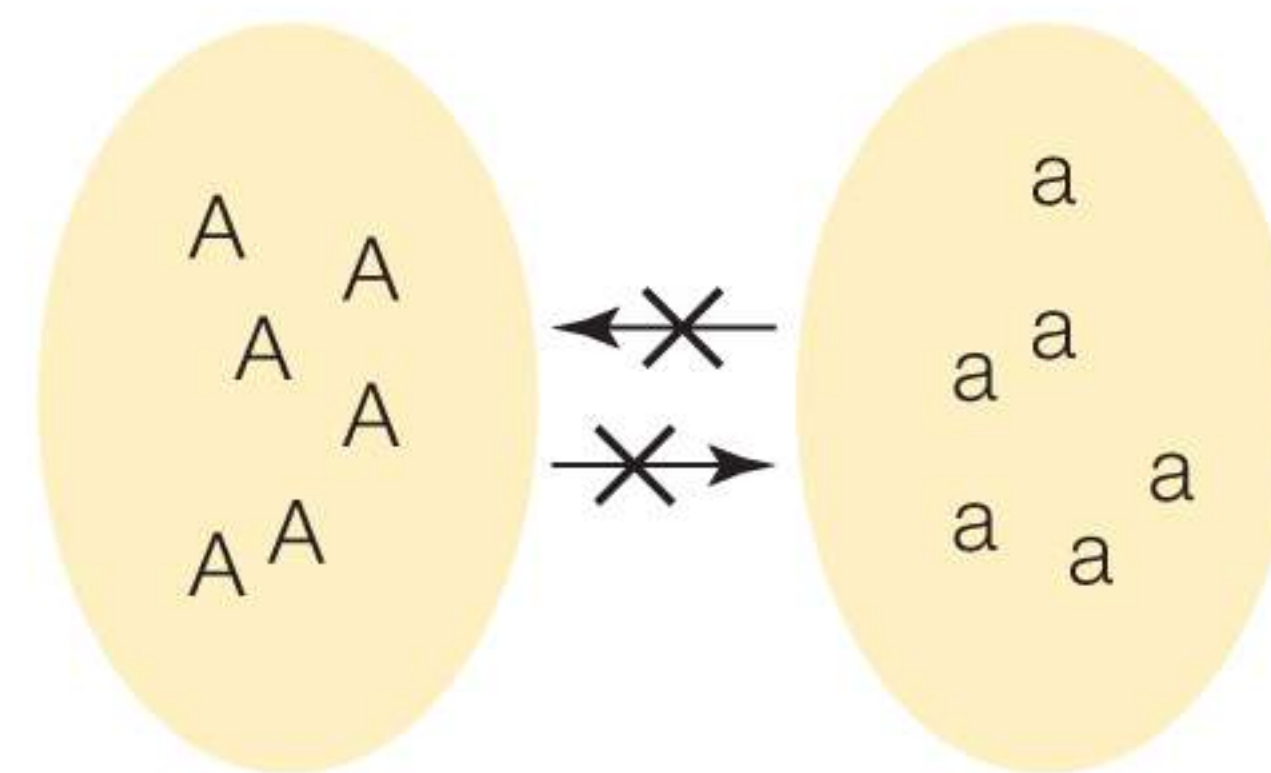
(B) $F_{ST} = 0.44$



Population 1

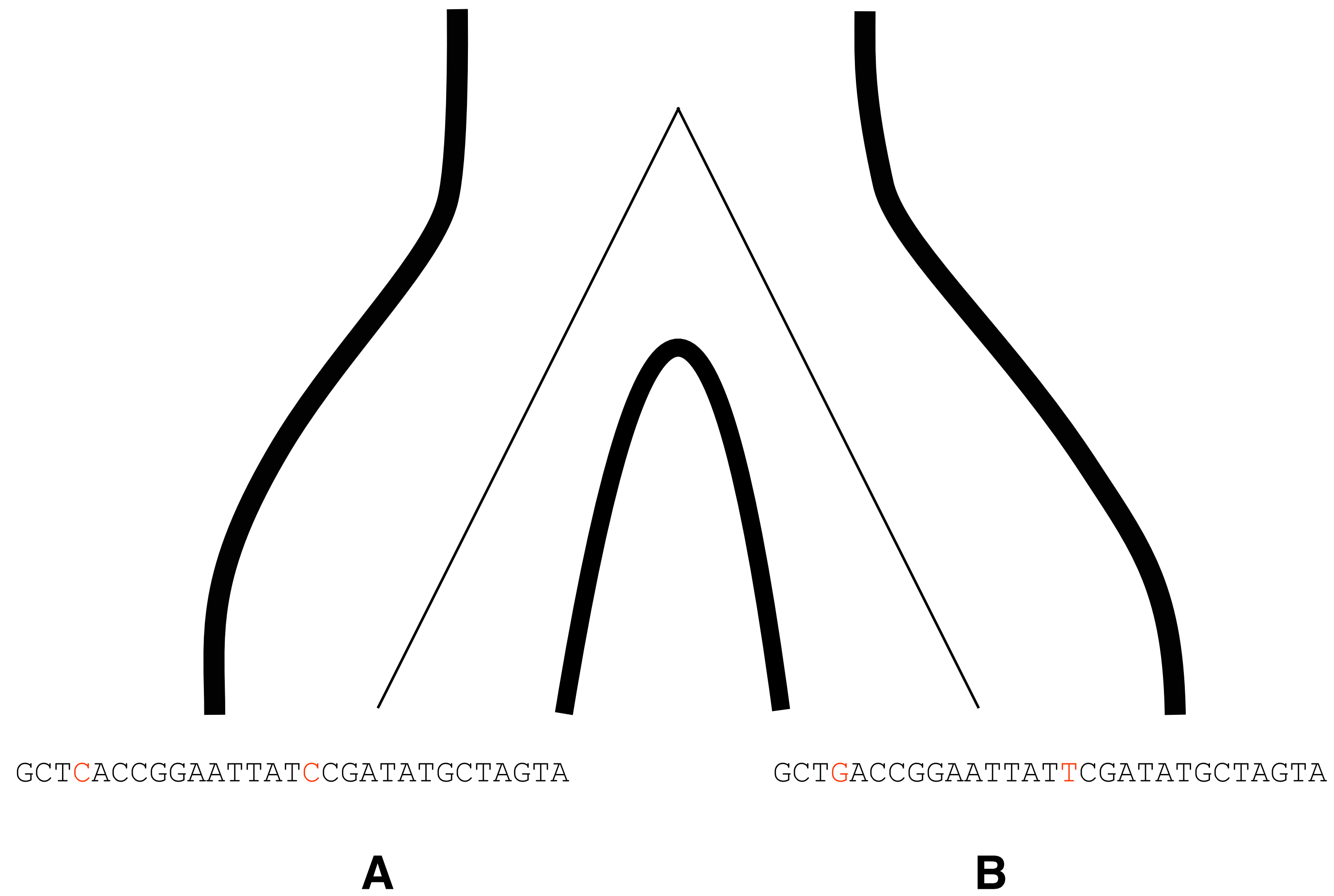
Population 2

(C) $F_{ST} = 1$

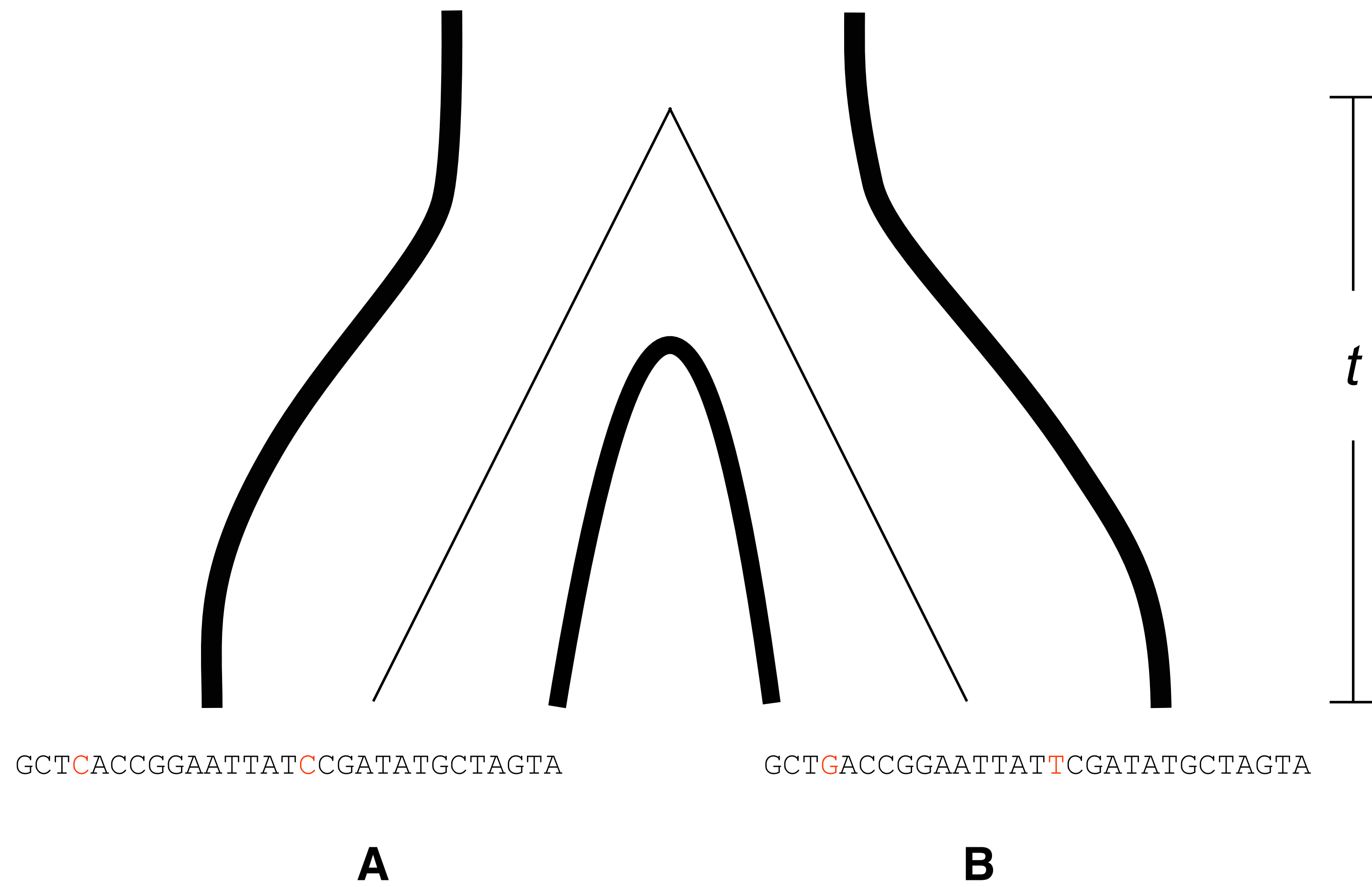


Population 1

Population 2



1	GCTCACCGGAATTATCCGATATGCTAGTA	}	A
2	GCTTACCGGAATTATGCGATATGCTTGTA		
3	GCTCACCGGAATTATGCGATATGGTAGAA		
4	GCTCACCGGAATTATGCGATATGGTAGAA		
5	GCTCACCGGGATGATGCGATATGCTAGTA	}	B
6	GCTCACCGGAATTATGCGATATGCTAGAA		
7	GCTTACCGGAATTATCCGATATGCTAGTA		
8	GCTCACAGGGATTATGCGCTATGCTAGTA		
9	GCTCACCGGAATTATGCGATATGGTAGAA		
10	GCTCACCGGAATTATCCGATATGCTAGTA		



Outline for today

1. Measures of variation
2. Population structure
- 3. Natural selection**

Matthew Hahn
mwh@iu.edu
@3rdreviewer

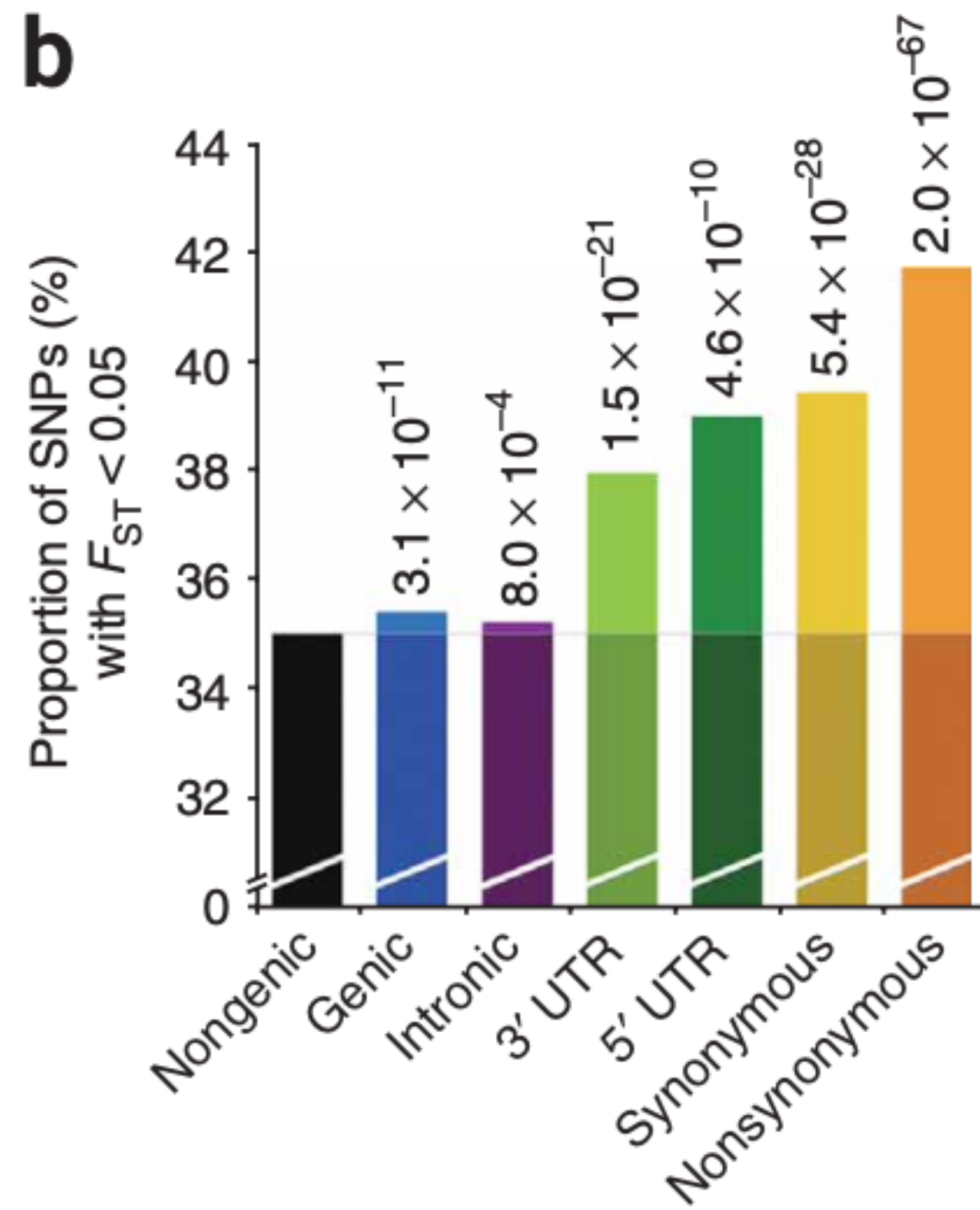
Natural selection:

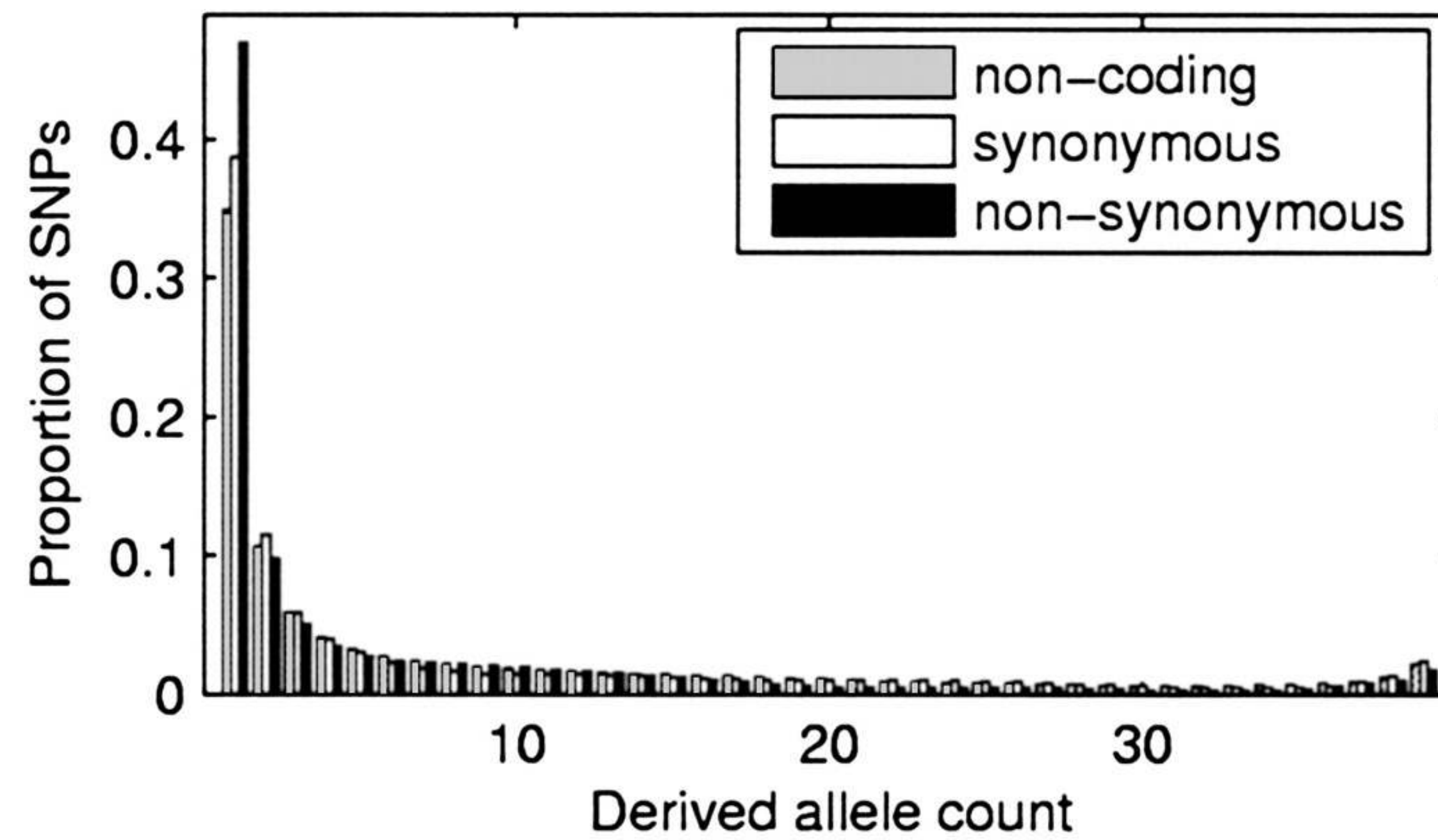
A consistent difference in survival and/or reproduction among individuals that differ in one or more heritable traits

GCTTACCGGAATTATGCG
GCTTACAGGAATTATGCG
GCTCACCGGATTATGCG
GCTTACAGGATTATGCG
GCTTACCGGAATTATGG
GCTTACCGGATTATGCG
CCTTACCGGAATTATTCG
CCTTACCGGAATTATTCG
CCTTACAGGAATTATGCG
GCTTACCGGATTATGCG

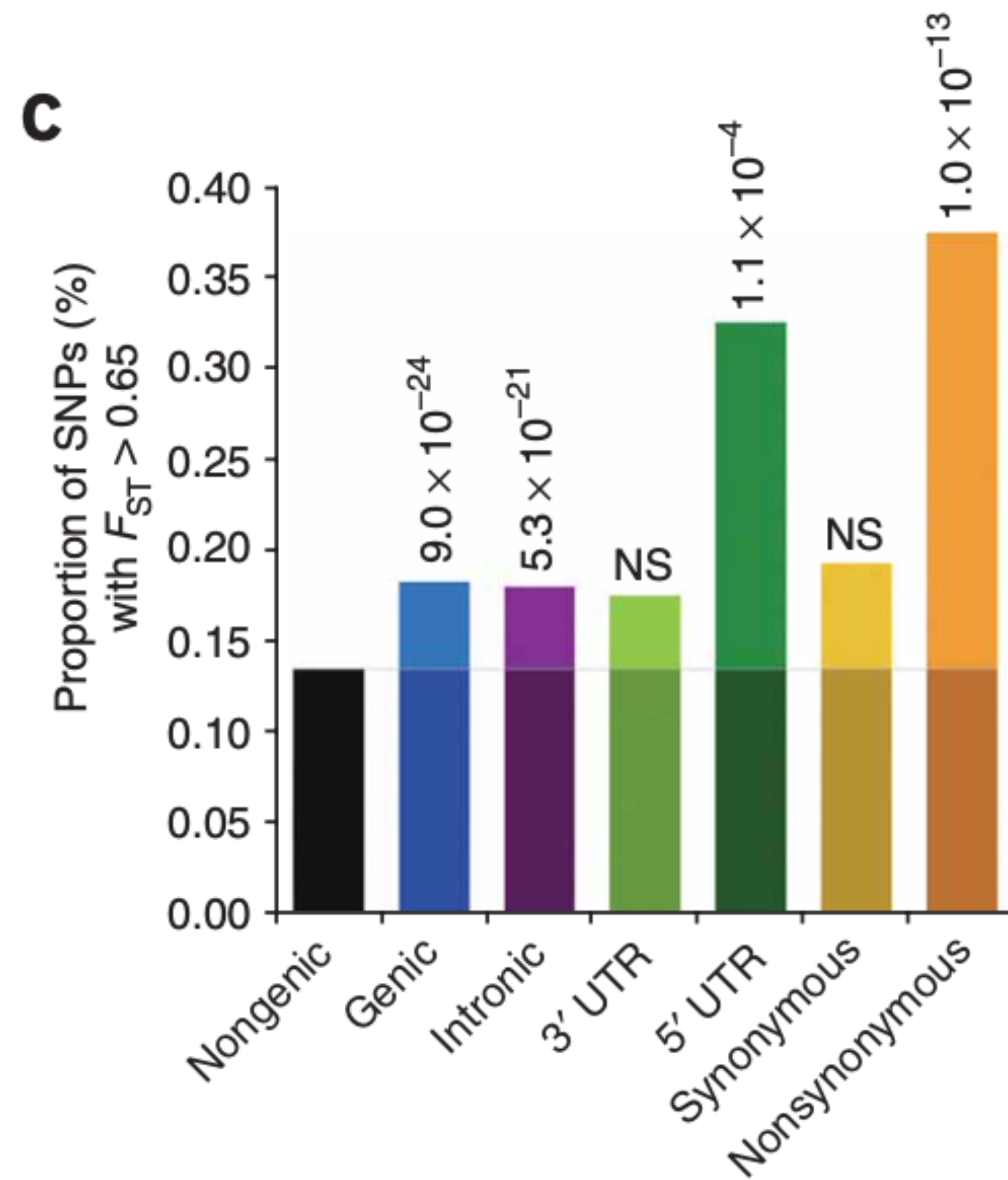


GCTTACCGTAATTATGCG
GCTTACAGGAATTATGCG
GCTCACCGGATTATGCG
GCTTACAGGATTATGCG
GCTTACCGGAATTATGG
GCTTACCGGATTATGCG
CCTTACCGGAATTATTCG
CCTTACCGGAATTATTCG
CCTTACAGGAATTATGCG
GCTTACCGGATTATGCG





[illegible]



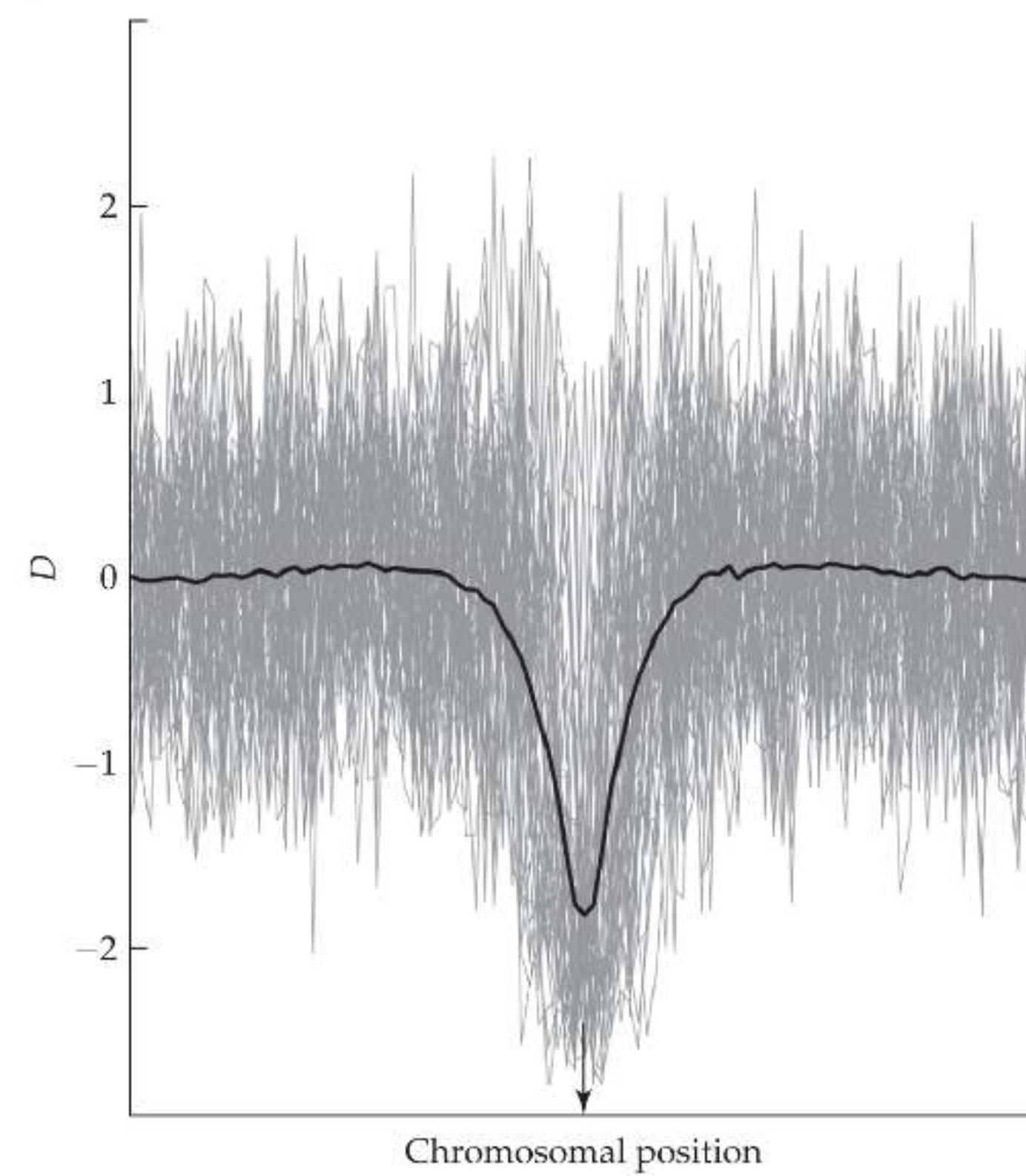


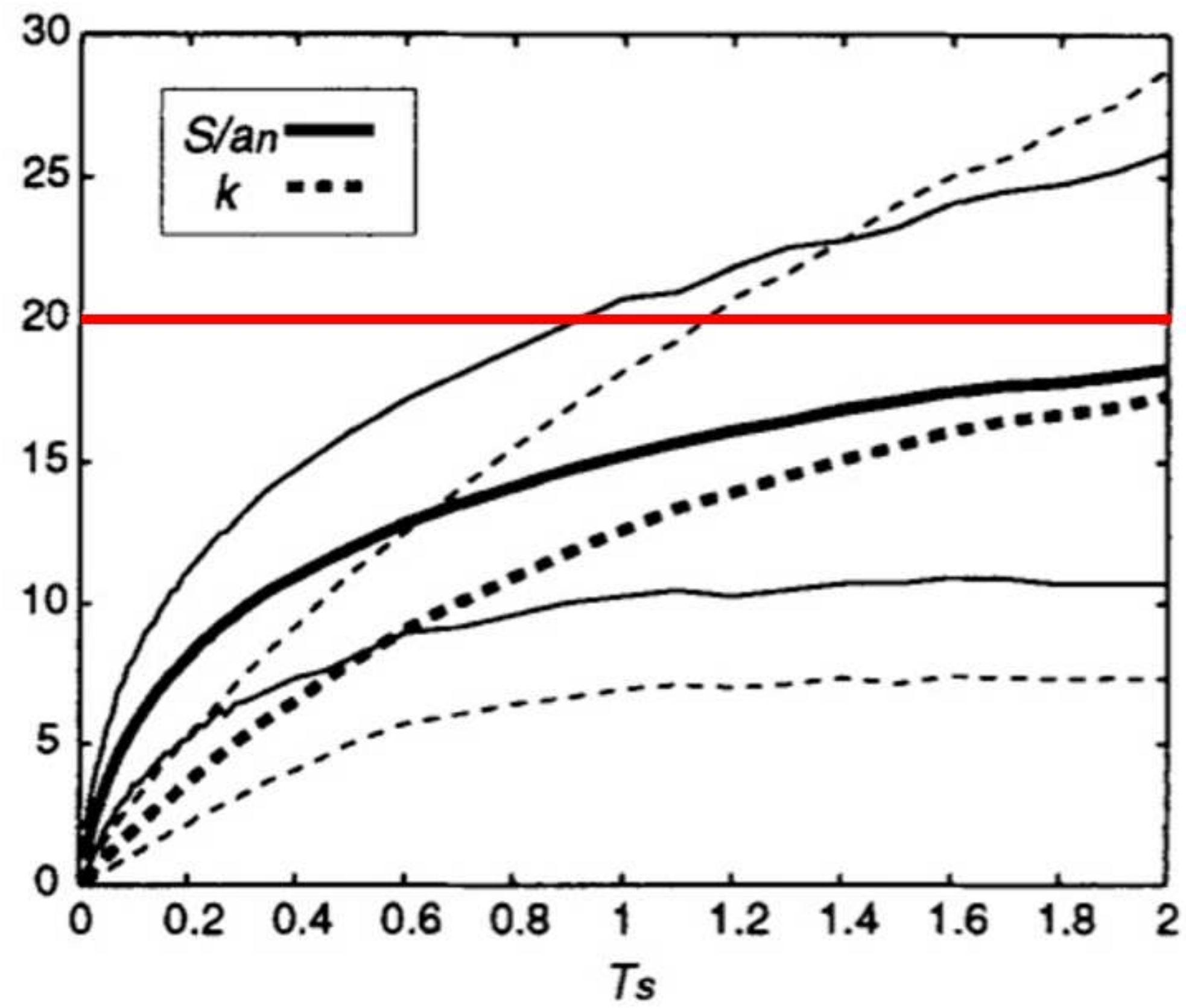
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG



GCTTGCCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTCTGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATAATATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GTTTACCGTAATTATGCG

(A)





[illegible]

GCTTACCGGAATTATGCG
GCTTACAGGAATTATGCG
GCTCACC GGATT TATGCG
GCTTACAGGATT TATGCG
GCTTACCGGAATTATGG
GCTTACCGGATT TATGCG
CCTTACCGGAATTATTCG
CCTTACCGGAATTATTCG
CCTTACAGGAATTATGCG
GCTTACCGGATT TATGCG



GCTTACCGGAATTATGCG
GCTTACAGGAATTATGCG
GCTCACC GGATT TATGCG
GCTTACAGGATT TATGCG
GCTTACCGGAATTATGG
GCTTACCGGATT TATGCG
CCTTACCGGAATTATTCG
CCTTACCGGAATTATTCG
CCTTACAGGAATTATGCG
GCTTACCGGATT TATGCG



GCTTACAGTAATTATGCG
GCTTACAGTAATTATGCG
GCTTACAGTAATTATGCG
GCTTACAGTAATTATGCG
GCTTACAGTAATTATGCG
GCTTACAGTAATTATGCG
CCTTACAGTAATTATGCG
CCTTACAGTAATTATGCG
CCTTACAGTAATTATGCG
CCTTACAGTAATTATGCG

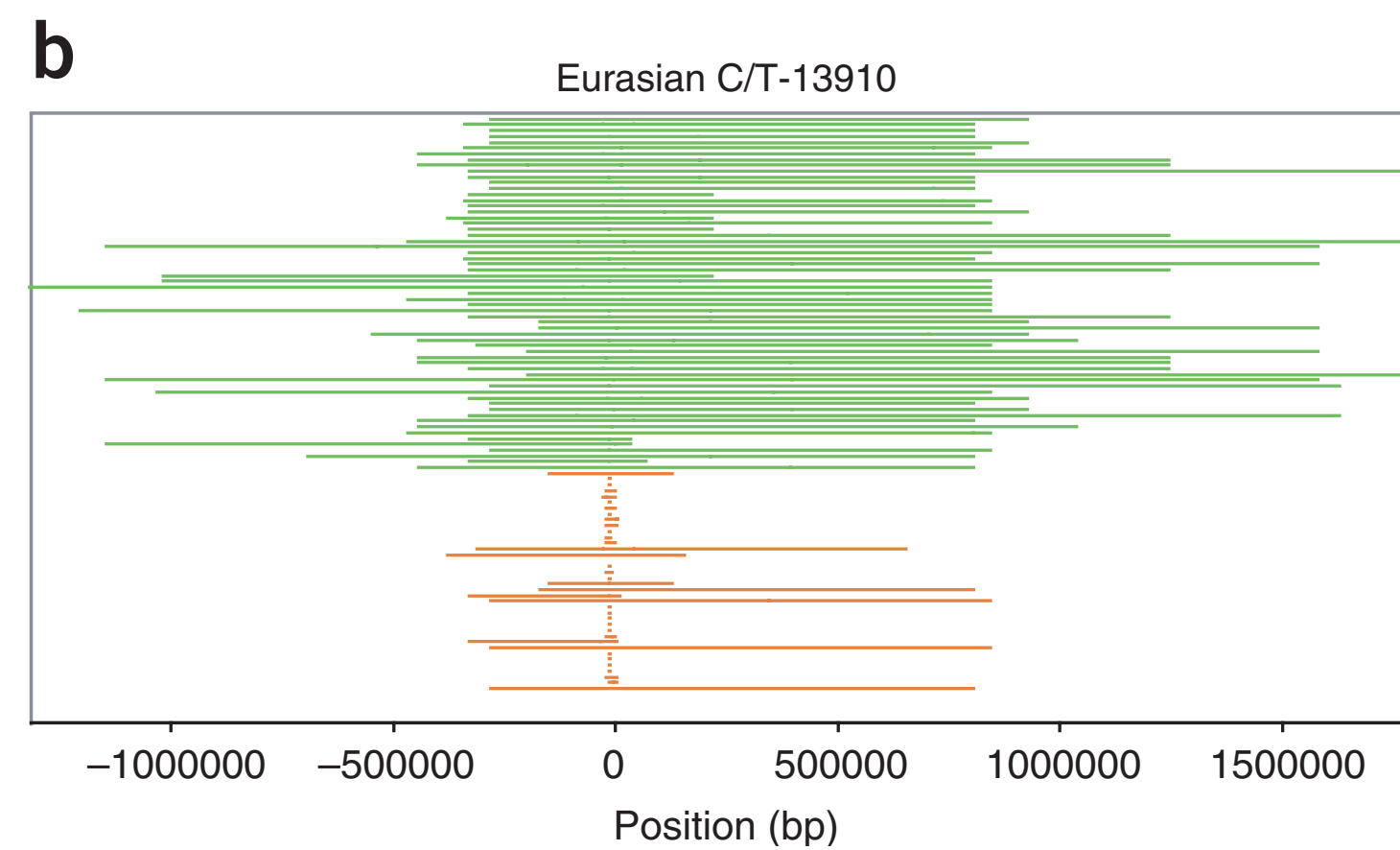
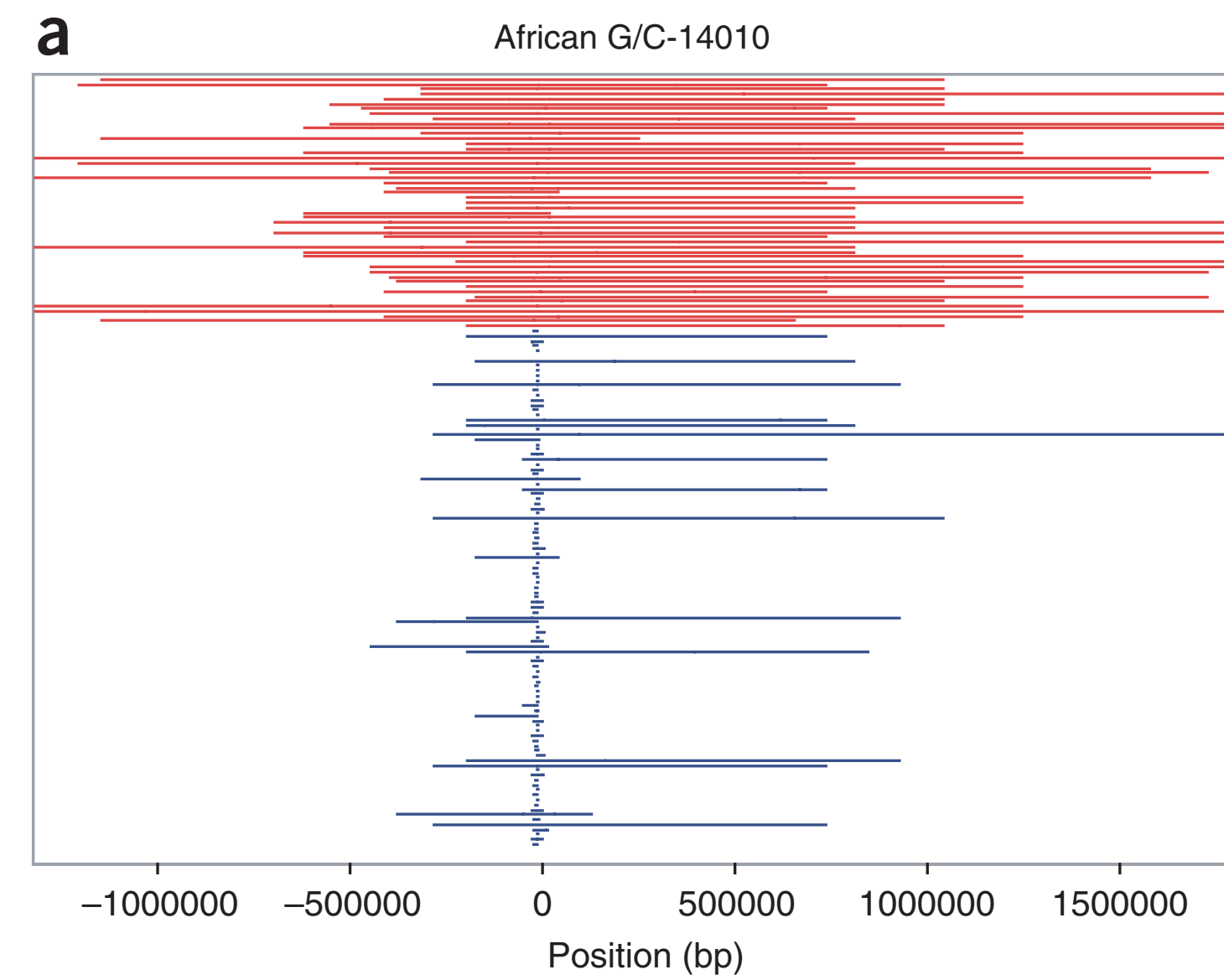
GCTTACCGGAATTATGCG
GCTTACAGGAATTATGCG
GCTCACCGGATTATGCG
GCTTACAGGATTATGCG
GCTTACCGGAATTATGG
GCTTACCGGATTATGCG
CCTTACCGGAATTATTCG
CCTTACCGGAATTATTCG
CCTTACAGGAATTATGCG
GCTTACCGGATTATGCG



GCTTACCGTAATTATGCG
GCTTACAGGAATTATGCG
GCTCACCGGATTATGCG
GCTTACAGGATTATGCG
GCTTACCGGAATTATGG
GCTTACCGGATTATGCG
CCTTACCGGAATTATTCG
CCTTACCGGAATTATTCG
CCTTACAGGAATTATGCG
GCTTACCGGATTATGCG



GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG
GCTTACCGGATTATGCG
CCTTACCGGAATTATTCG
CCTTACCGGAATTATTCG
CCTTACAGGAATTATGCG
GCTTACCGGATTATGCG



Tishkoff et al. (2007)

GCTTACCGGAATTATGCG
GCTTACCGGAATTATGCG
GCTTACCGGAATTATGCG
GCTTACCGGAATTATGCG
GCTTACCGGAATTATGCG
GCTTACCGG GATTATGCG
GCTTACCGG GATTATGCG
GCTTACCGG GATTATGCG
GCTTACCGG GATTATGCG
GCTTACCGG GATTATGCG



GCTCACCGGAATTATGCG
GCTTACCGGAATTATGCG
GCTTACCGGAATTATGCG
GCTTACCGGAATTATGCG
GCTTACCGGAATTATGCG
GCTTACCGGAATTATGCG
GCTTACCGG GATTATGCG
GCTTACCGG GATTATTCG
GCTTACCGG GATTATGCG
GCTTACCGG GATTATGCG
GCTTACCGG GATTATGCG



GCTCACCGGAATTATGCG
GCTTACAGGAATTATGCG
GCTCACCGGAATTATGCG
GCTTACAGGAATTATGCG
GCTTACCGGAATTATGG
GCTTACCGG GATTATGCG
CCTTACCGG GATTATTCG
CCTTACCGG GATTATTCG
CCTTACCGG GATTATGCG
GCTTACCGG GATTATGCG

