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

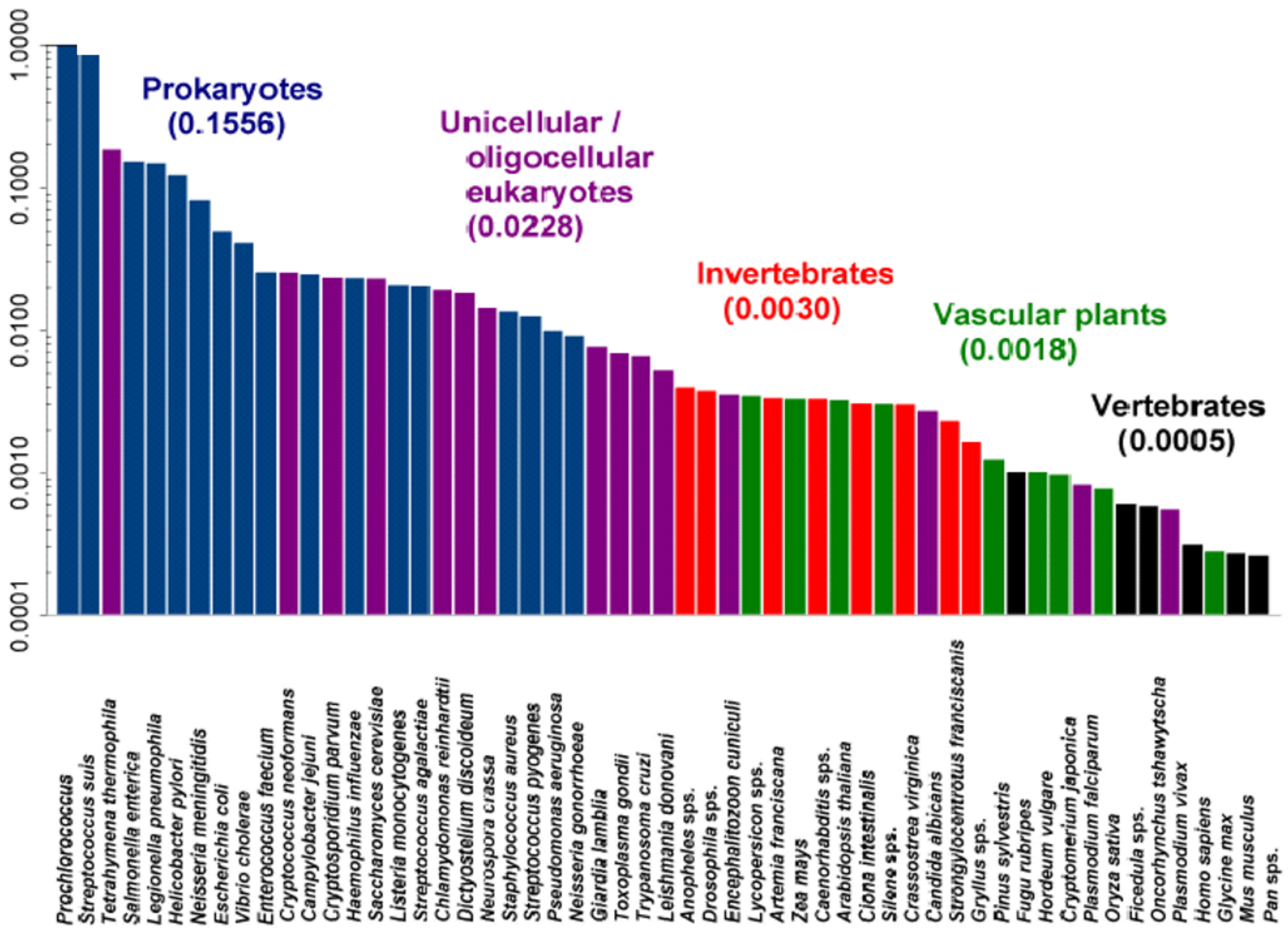
TCACAATGCGATGGA

TTACGATGCGCTCGT

```
TTACAATCCGATCGT
--ACGATGCGCTCGT
TCACAATGCGATGGA
TTACGATG--CTCGT
```

Levels of diversity in different species

π



Lynch (2006)

1 GCTCACCGGAATTATCCGATATGCTAGTA

2 GCTTACCGGAATTATGCGATATGCTTGTA

3 GCTCACCGGAATTATGCGATATGGTAGAA

4 GCTCACCGGAATTATGCGATATGGTAGAA

5 GCTCACCGGGATGATGCGATATGCTAGTA

6 GCTCACCGGAATTATGCGATATGCTAGAA

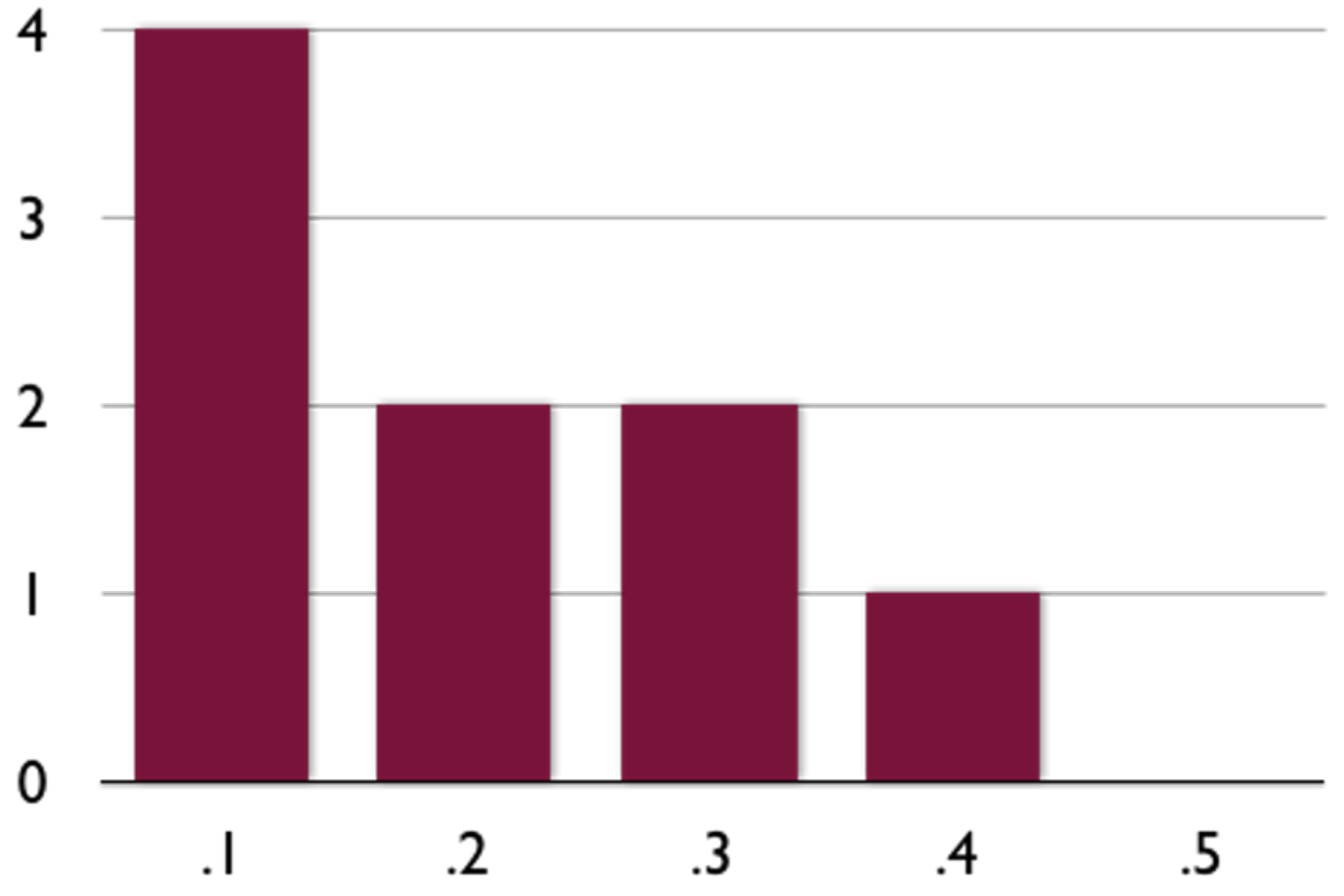
7 GCTTACCGGAATTATCCGATATGCTAGTA

8 GCTCACAGGGATTATGCGCTATGCTAGTA

9 GCTCACCGGAATTATGCGATATGGTAGAA

10 GCTCACCGGAATTATCCGATATGCTAGTA

Allele frequency spectrum (Minor allele frequency)



1 GCTCACCGGAATTATCCGATATGCTAGTA

2 GCTTACCGGAATTATGCGATATGCTTGTA

3 GCTCACCGGAATTATGCGATATGGTAGAA

4 GCTCACCGGAATTATGCGATATGGTAGAA

5 GCTCACCGGGATGATGCGATATGCTAGTA

6 GCTCACCGGAATTATGCGATATGCTAGAA

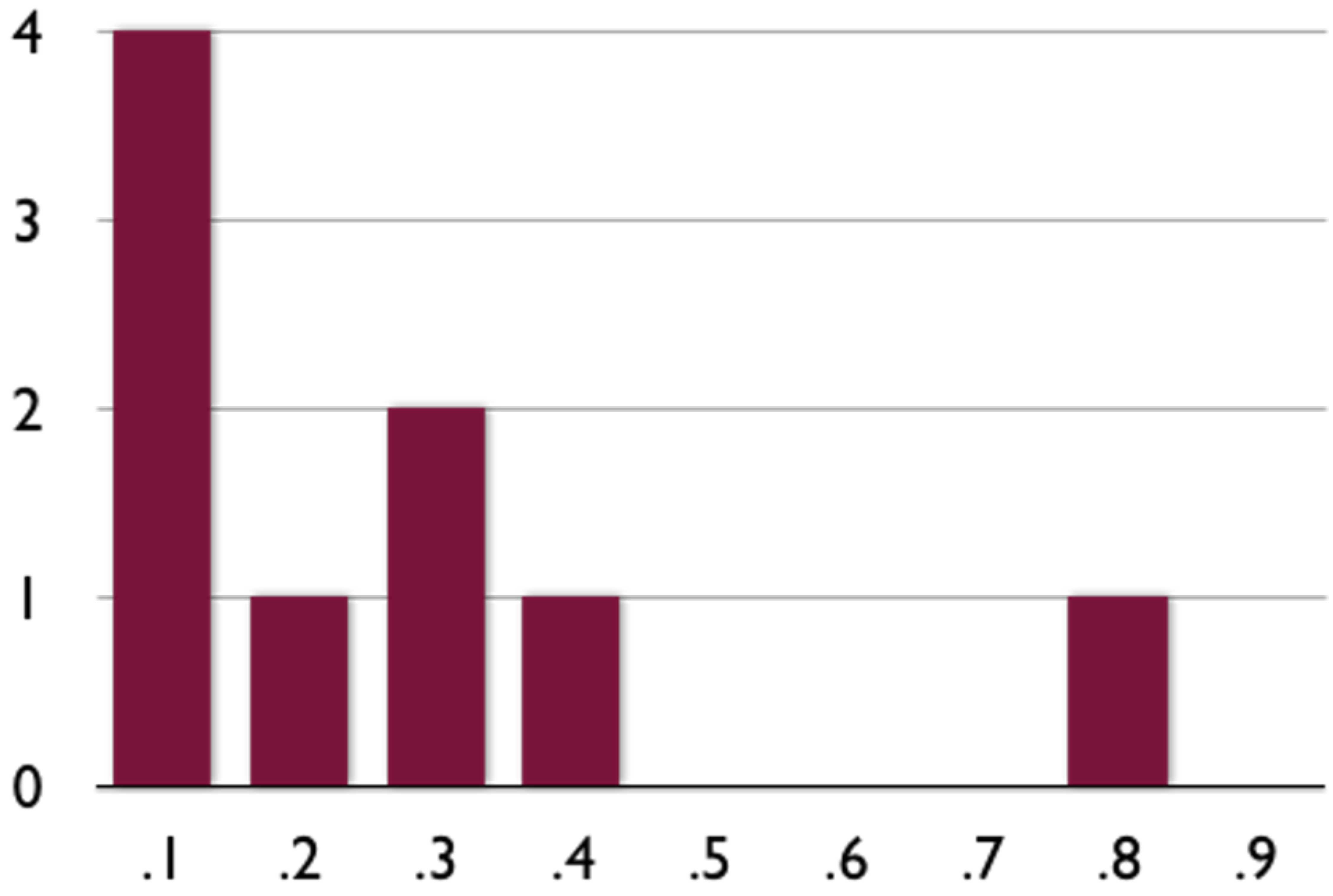
7 GCTTACCGGAATTATCCGATATGCTAGTA

8 GCTCACAGGGATTATGCGCTATGCTAGTA

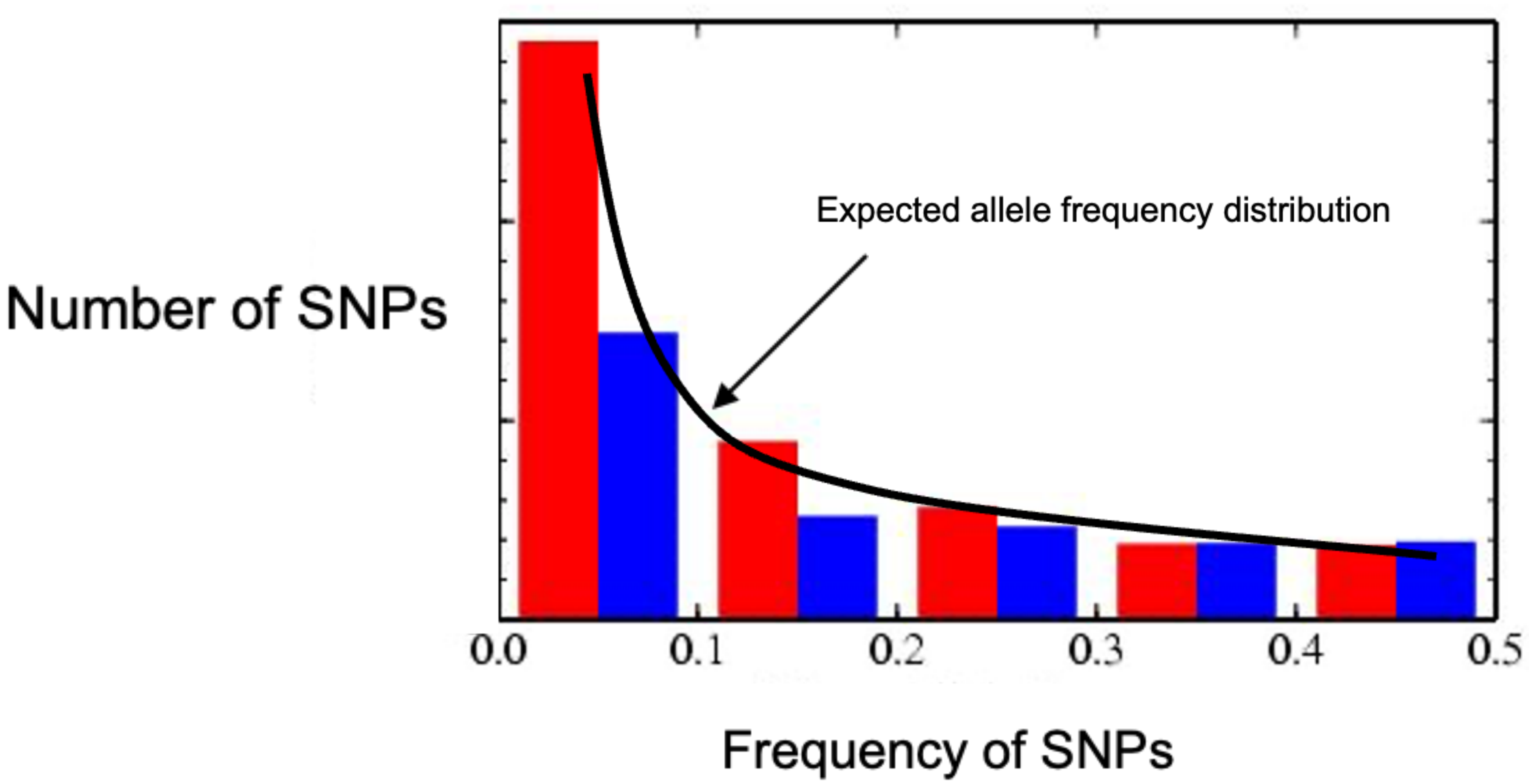
9 GCTCACCGGAATTATGCGATATGGTAGAA

10 GCTCACCGGAATTATCCGATATGCTAGTA

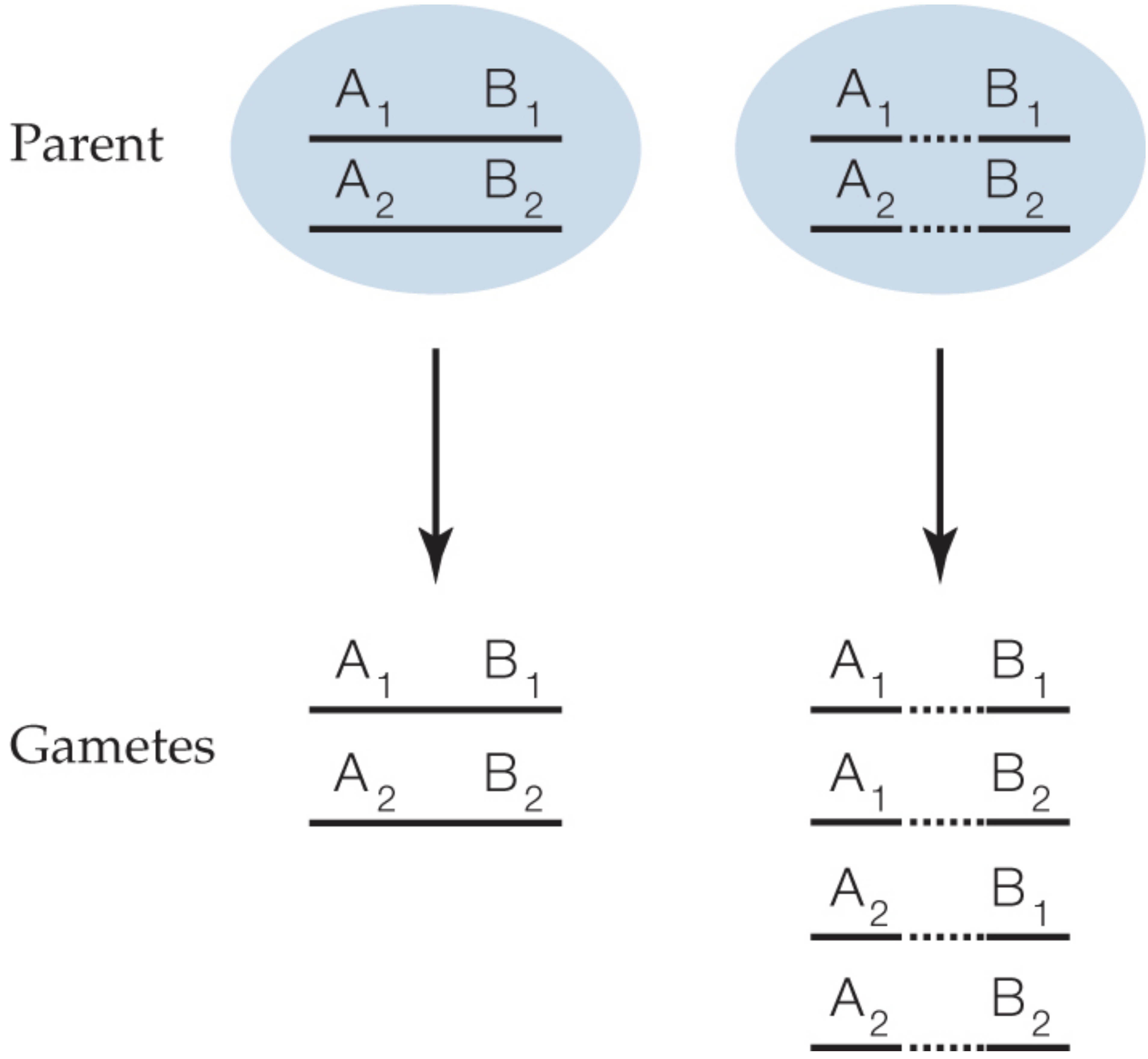
Allele frequency spectrum (Derived allele frequency)



Allele frequency spectrum

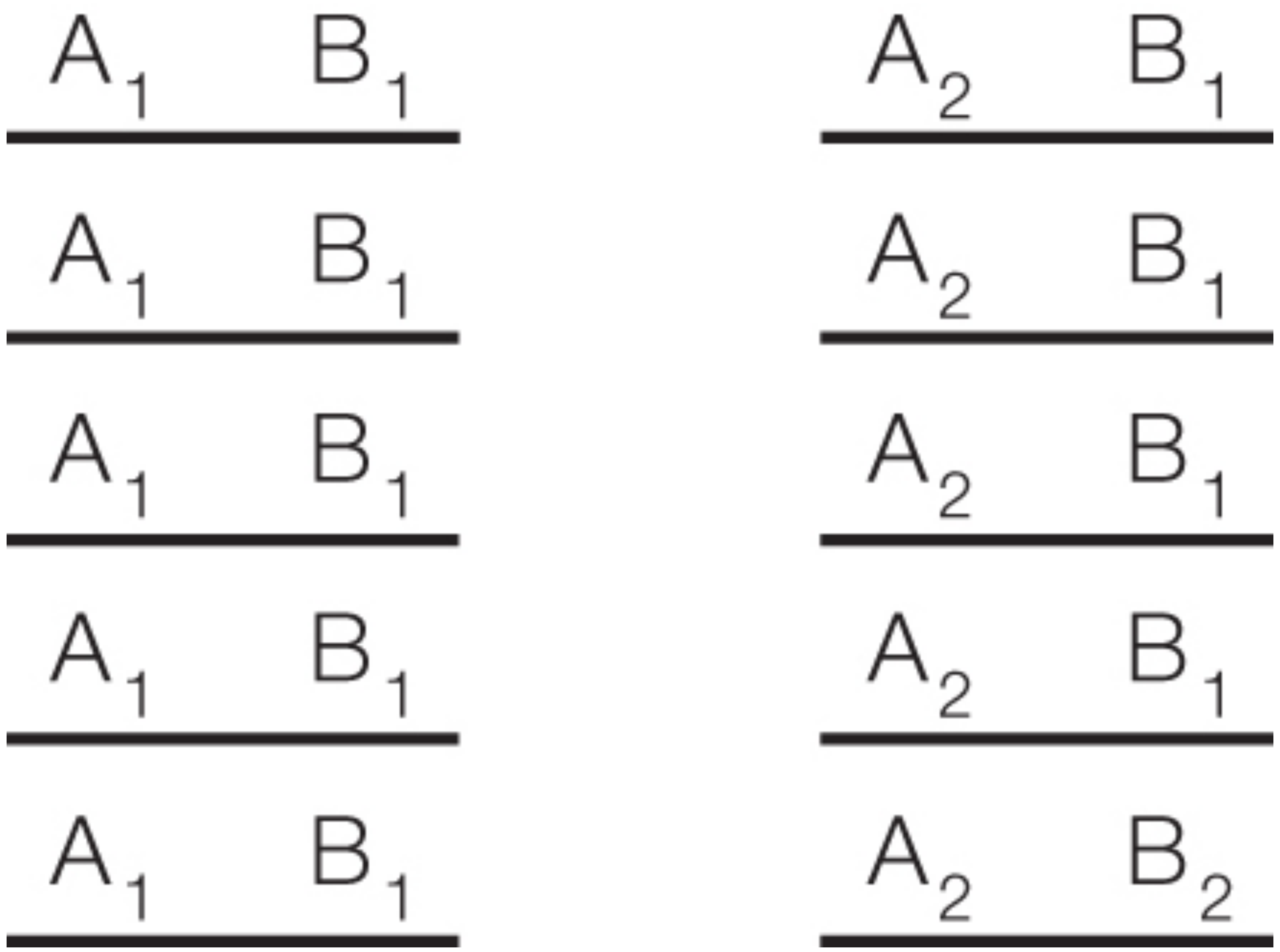


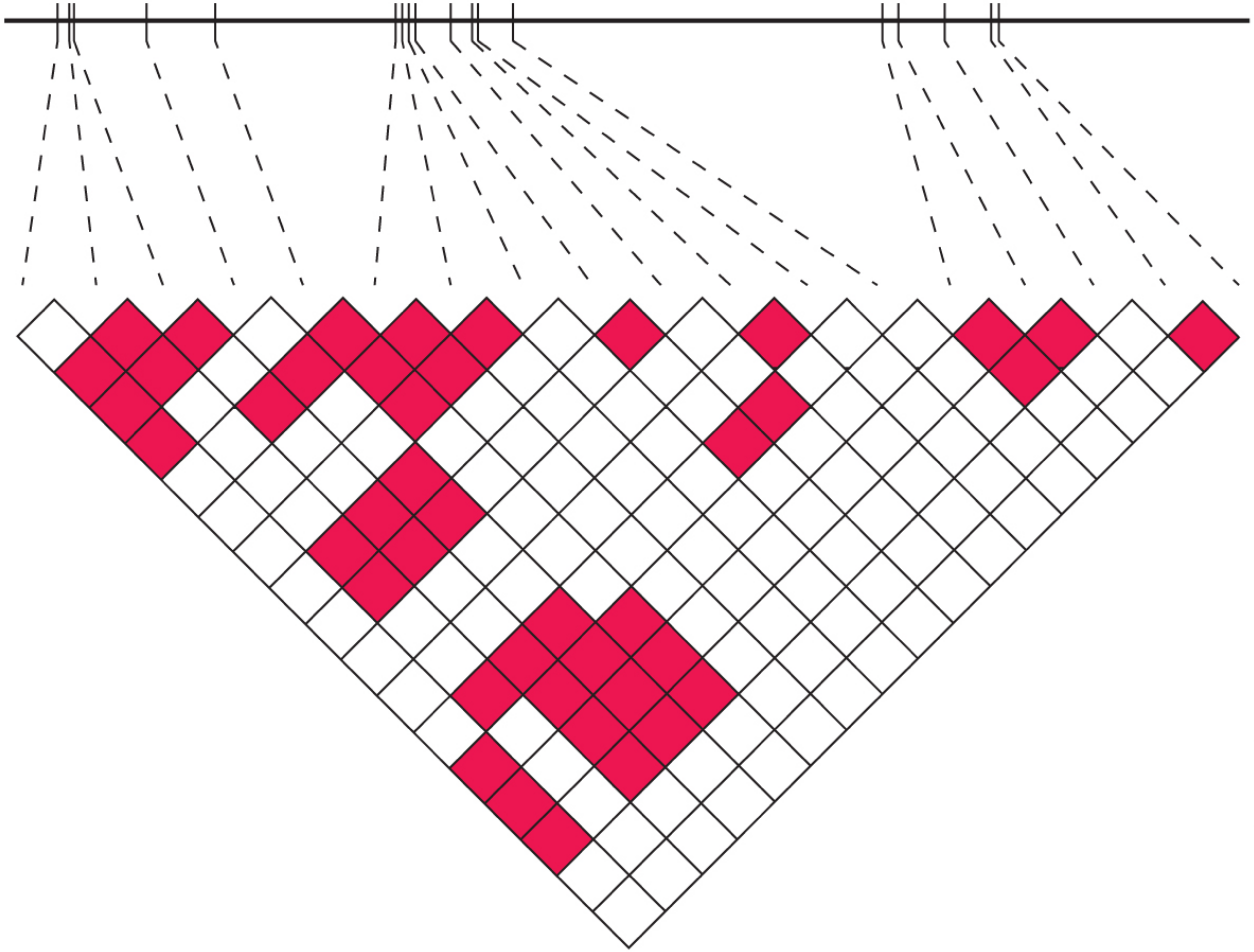
Linkage



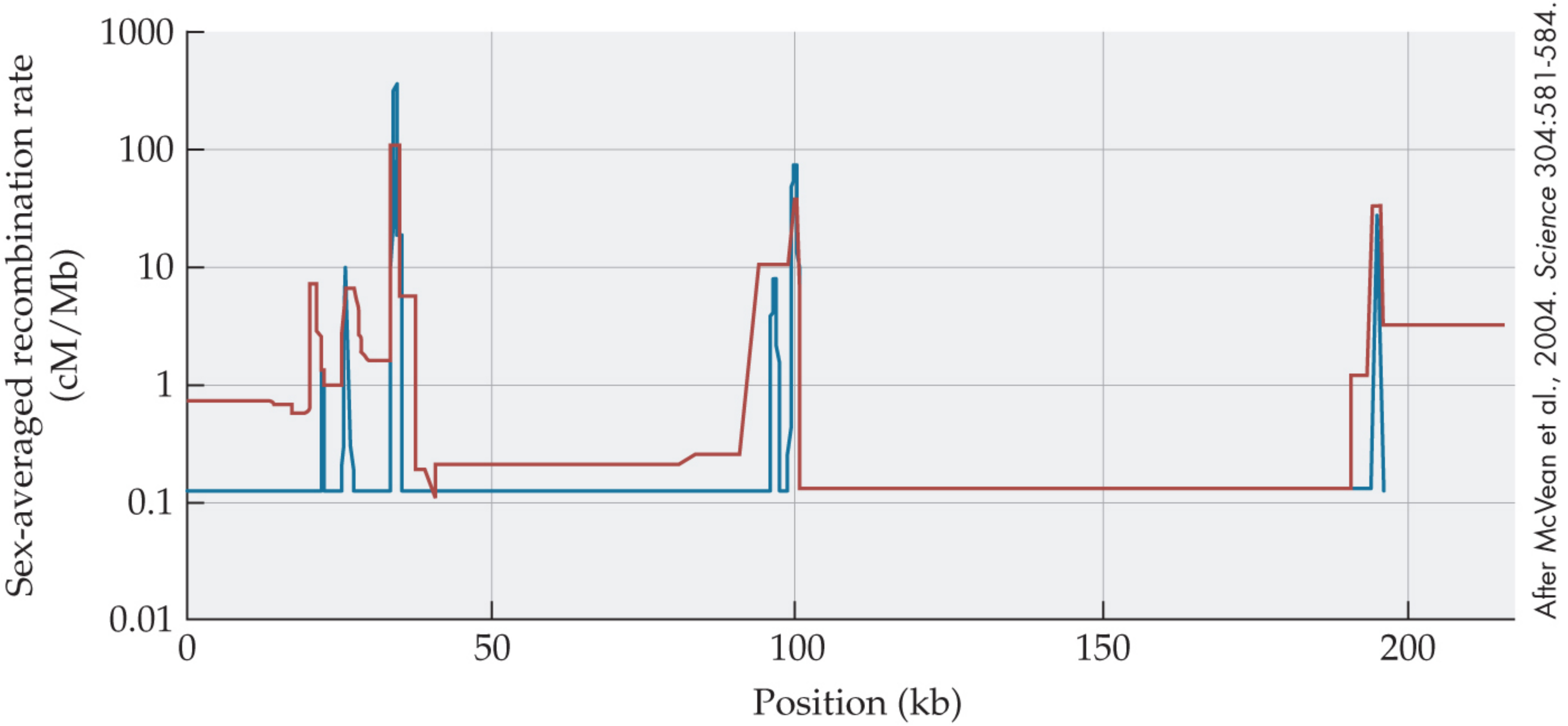
Parent
A1 B1
A2 B2
A1 B1
A2 B2
Gametes
A1 B1
A2 B2
A1 B1
A1 B2
A2 B1
A2 B2

Linkage disequilibrium





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


Sex-averaged recombination rate (cM/Mb)
1000
100
10
1
0.1
0.01
0
50
100
150
200
Position (kb)
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

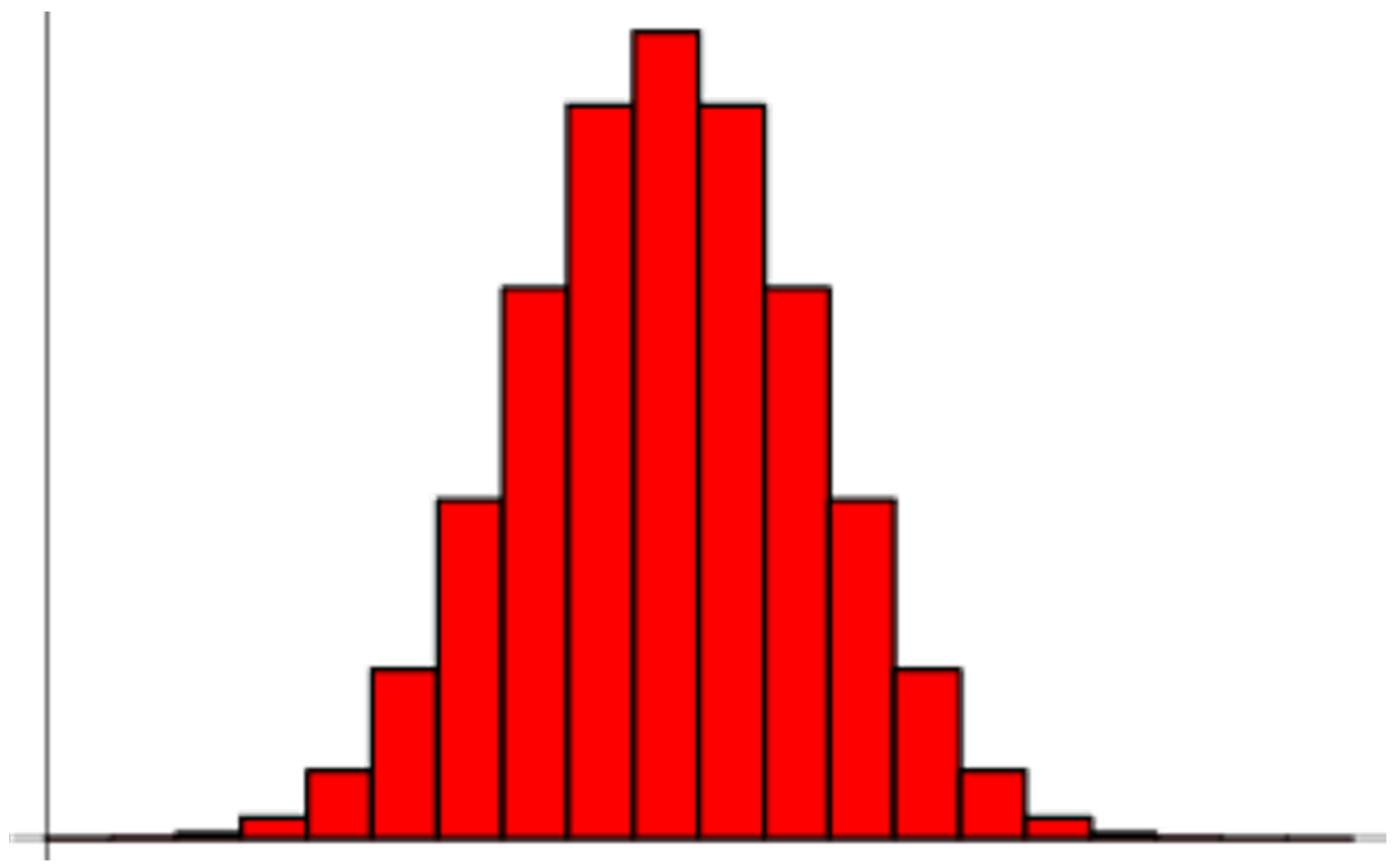
Forces that change genotype/haplotype frequencies:

mating
recombination

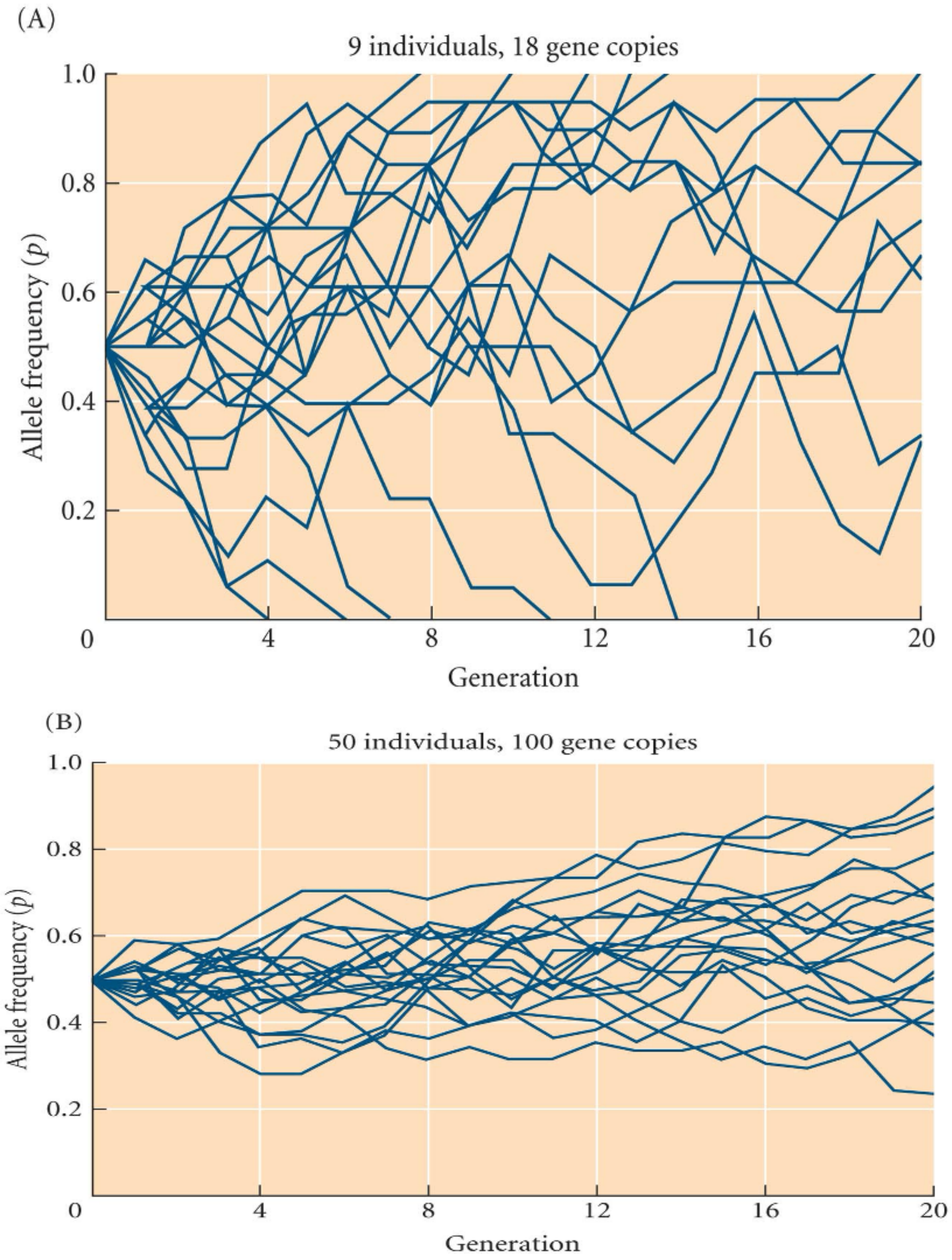
Genetic drift:

change in allele frequencies due to sampling in a finite population

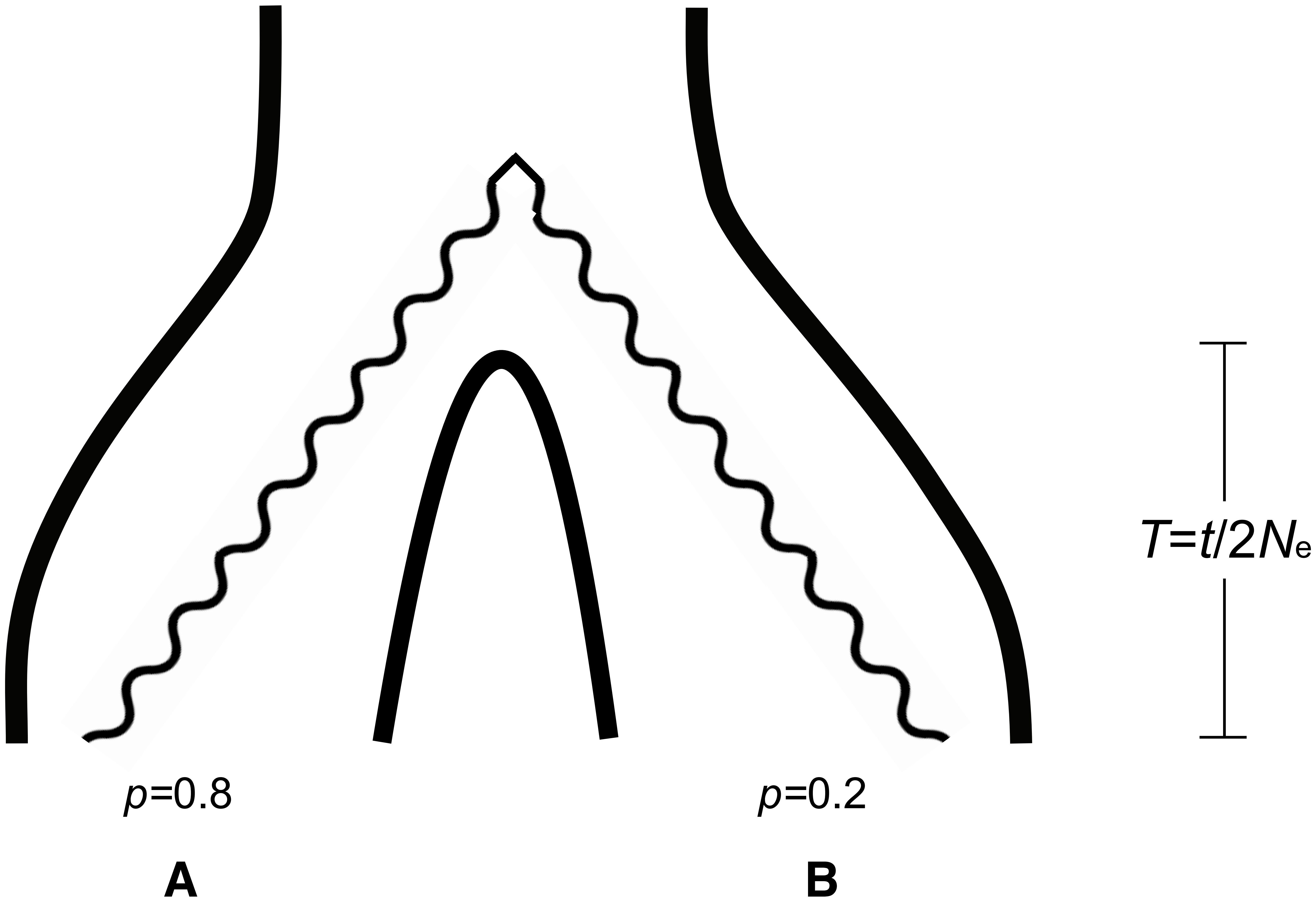
$$V = \frac{2pq}{2N_e}$$



2N=20 and p=0.5



(A)
9 individuals, 18 gene copies
Allele frequency (p)
1.0
0.8
0.6
0.4
0.2
0
4
8
12
16
20
Generation
(B)
50 individuals, 100 gene copies
Allele frequency (p)
1.0
0.8
0.6
0.4
0.2
0
4
8
12
16
20
Generation

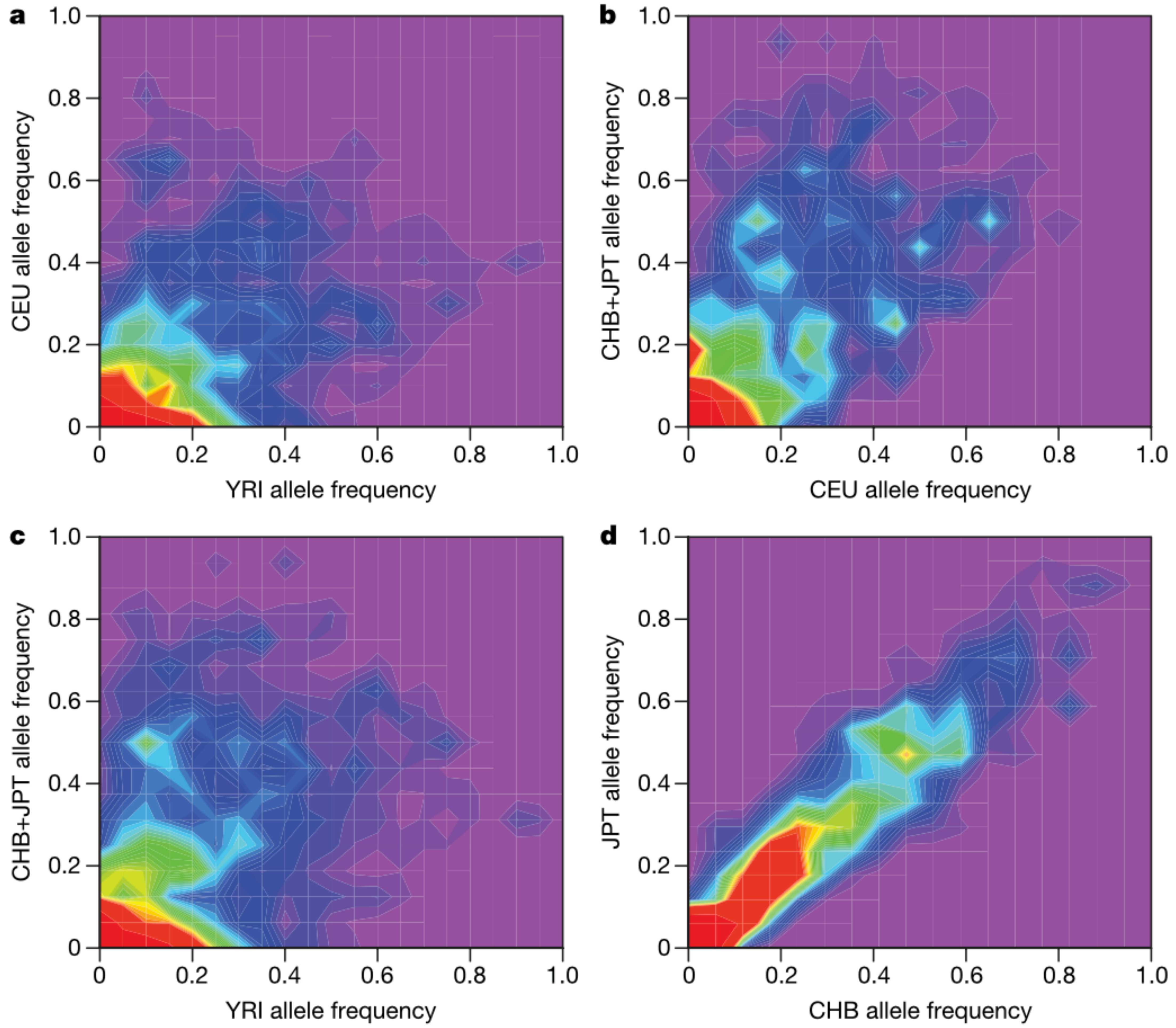


T=t/2Ne
p=0.8
p=0.2
A
B

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

F_{ST}=0.11

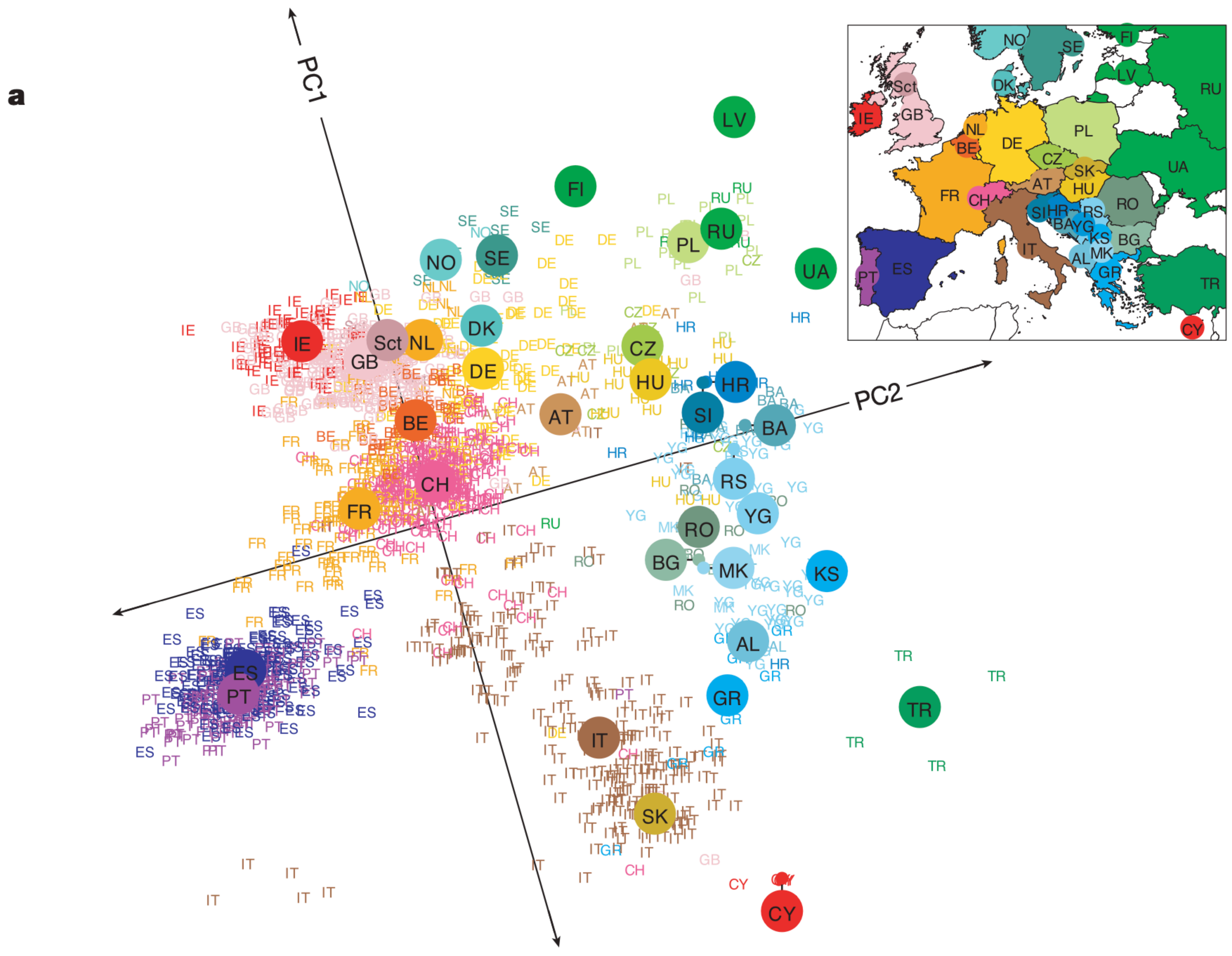
F_{ST}=0.07



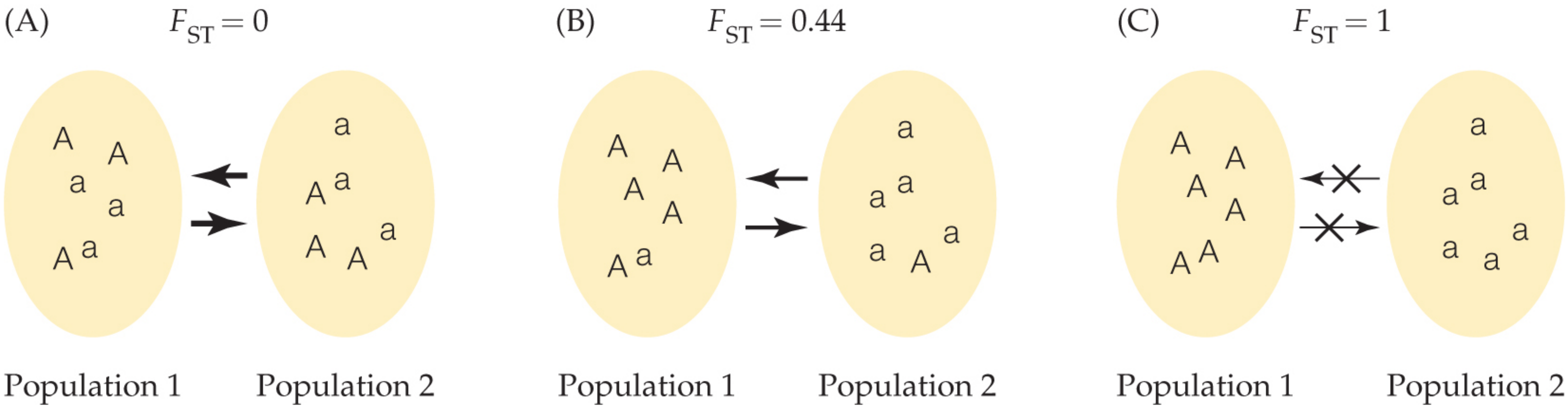
F_{ST}=0.12

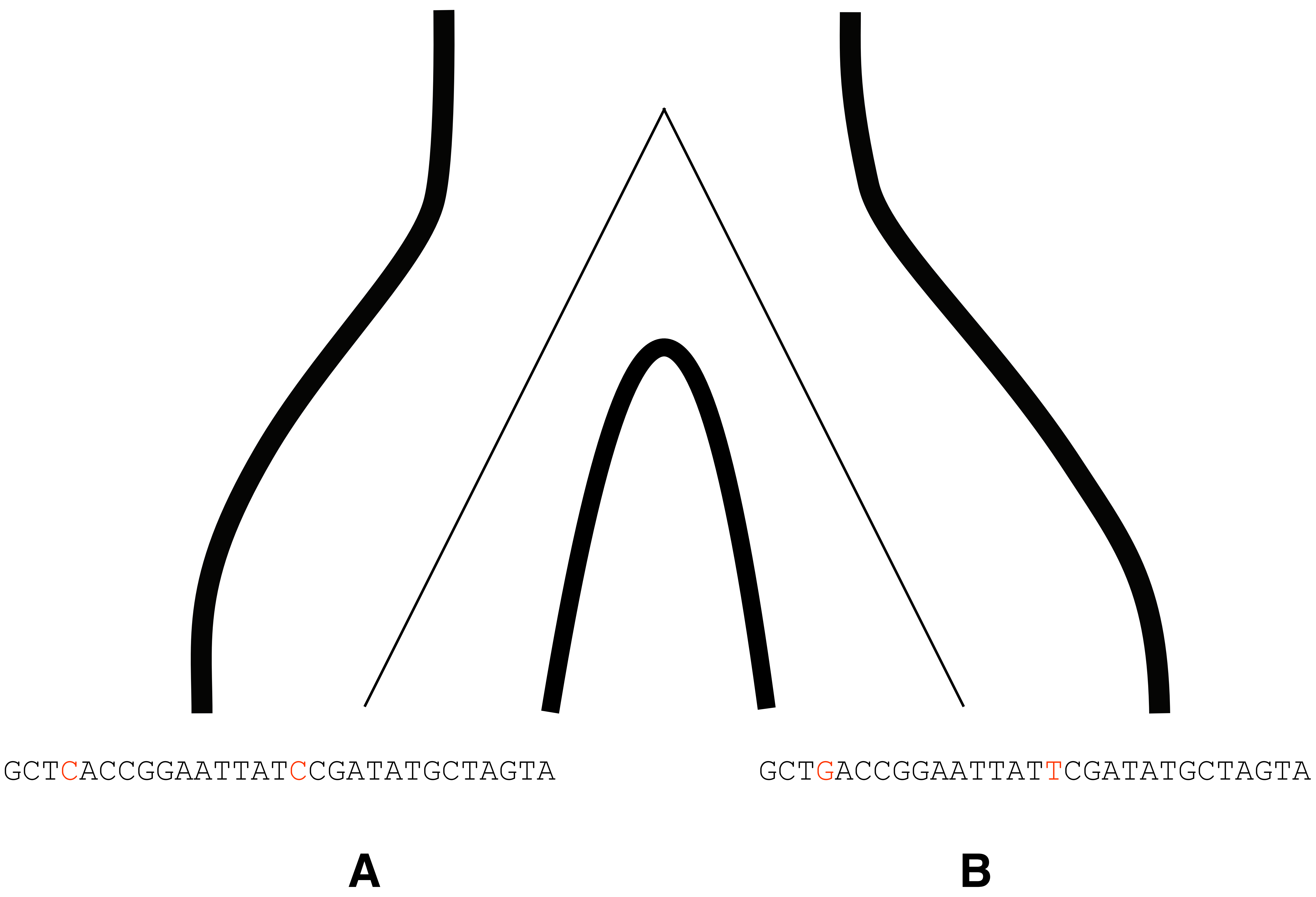
F_{ST}=0.01

HapMap (2005)

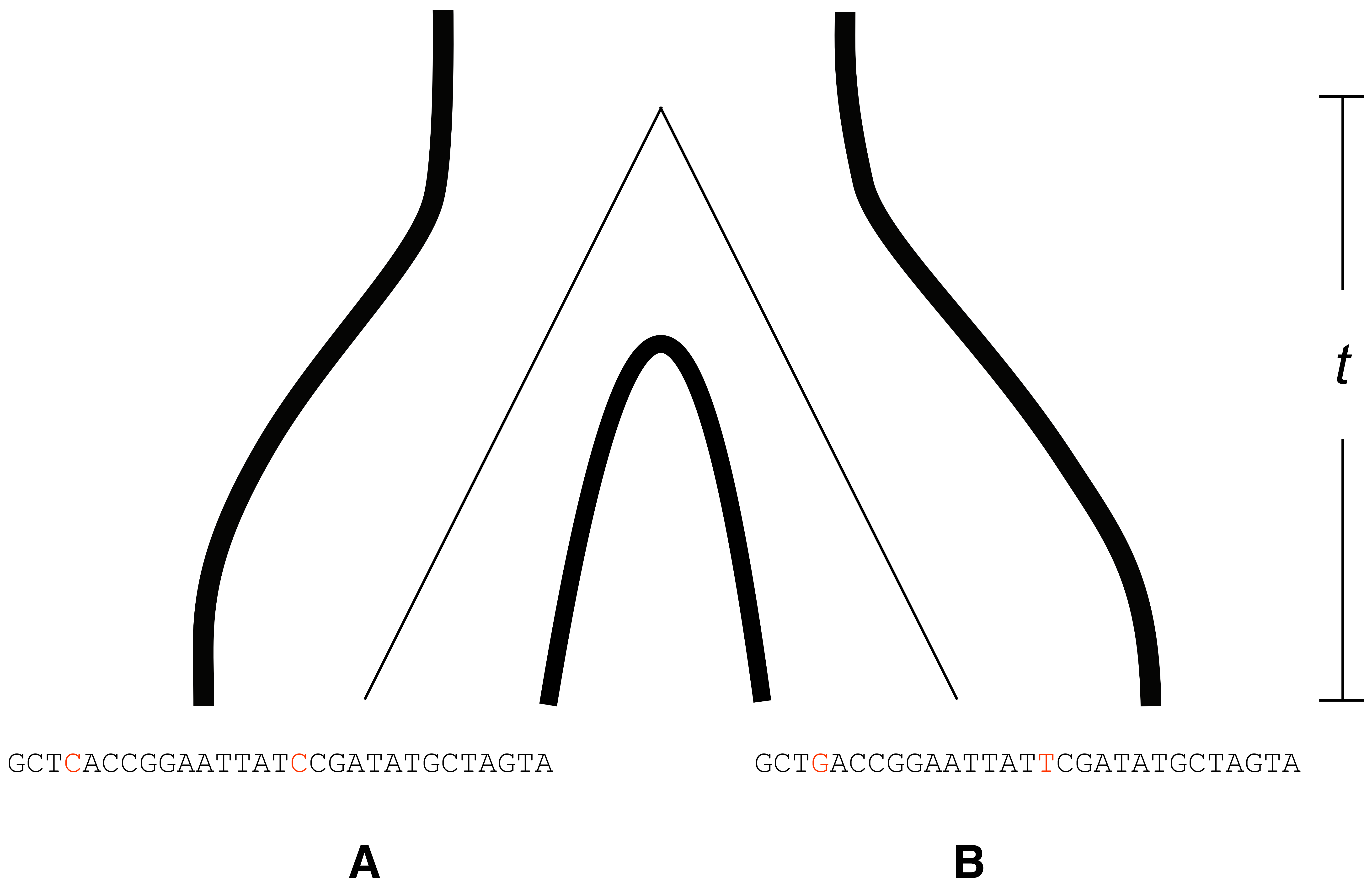


Novembre et al. (2008)


(A)
$F_{ST} = 0$
A
A
a
a
A
a
a
A
a
A
A
a
Population 1
Population 2
(B)
$F_{ST} = 0.44$
A
A
A
A
A
a
a
a
a
a
A
a
Population 1
Population 2
(C)
$F_{ST} = 1$
A
A
A
A
A
A
a
a
a
a
a
a
Population 1
Population 2


GCTCACCGGAATTATCCGATATGCTAGTA
GCTGACCGGAATTATTCGATATGCTAGTA
A
B

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


t
GCTCACCGGAATTATCCGATATGCTAGTA
GCTGACCGGAATTATTCGATATGCTAGTA
A
B

Outline for today

1. Measures of variation

2. Population structure

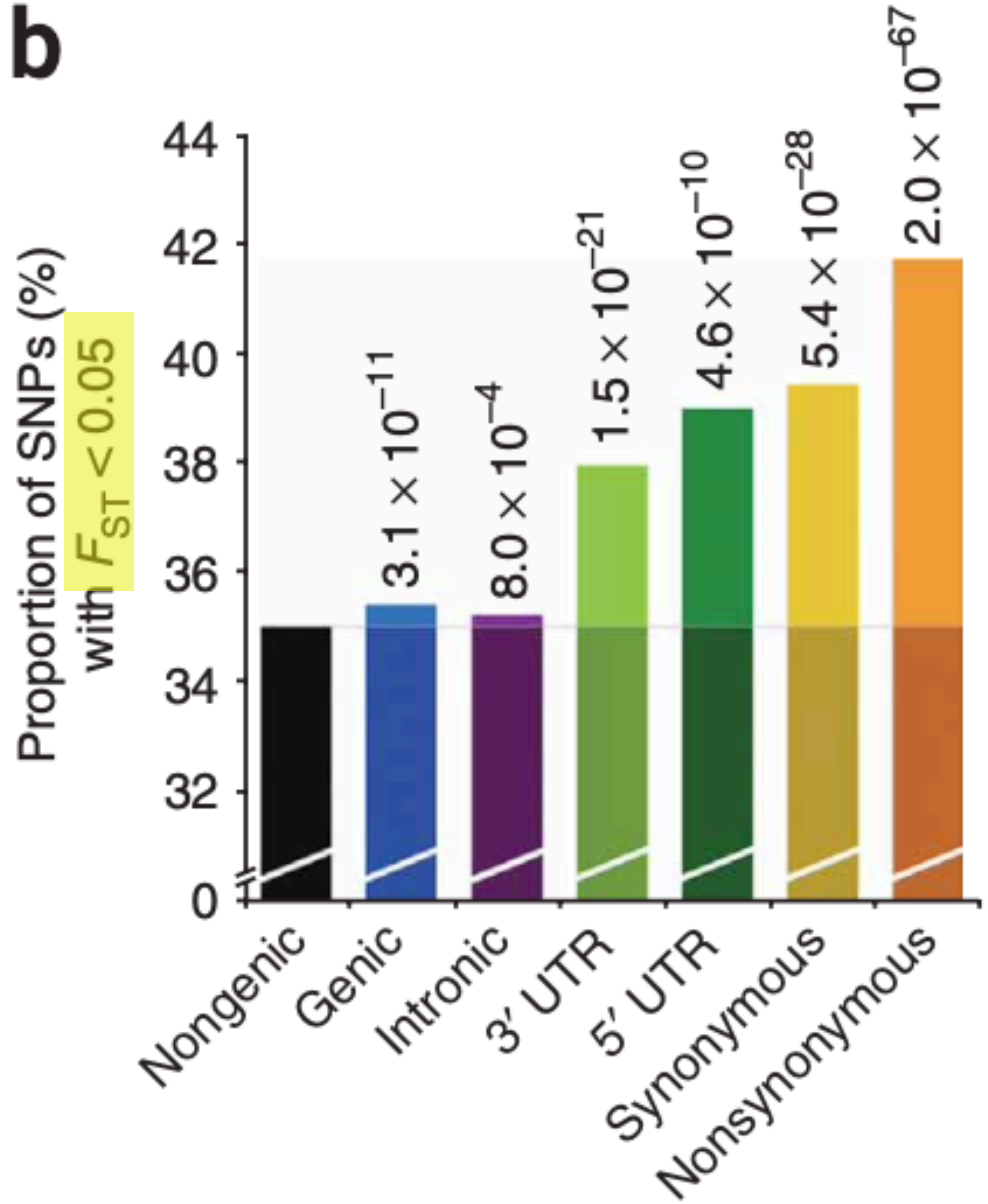
3. **Natural selection**

Matthew Hahn
mwh@iu.edu
@3rdreviewer

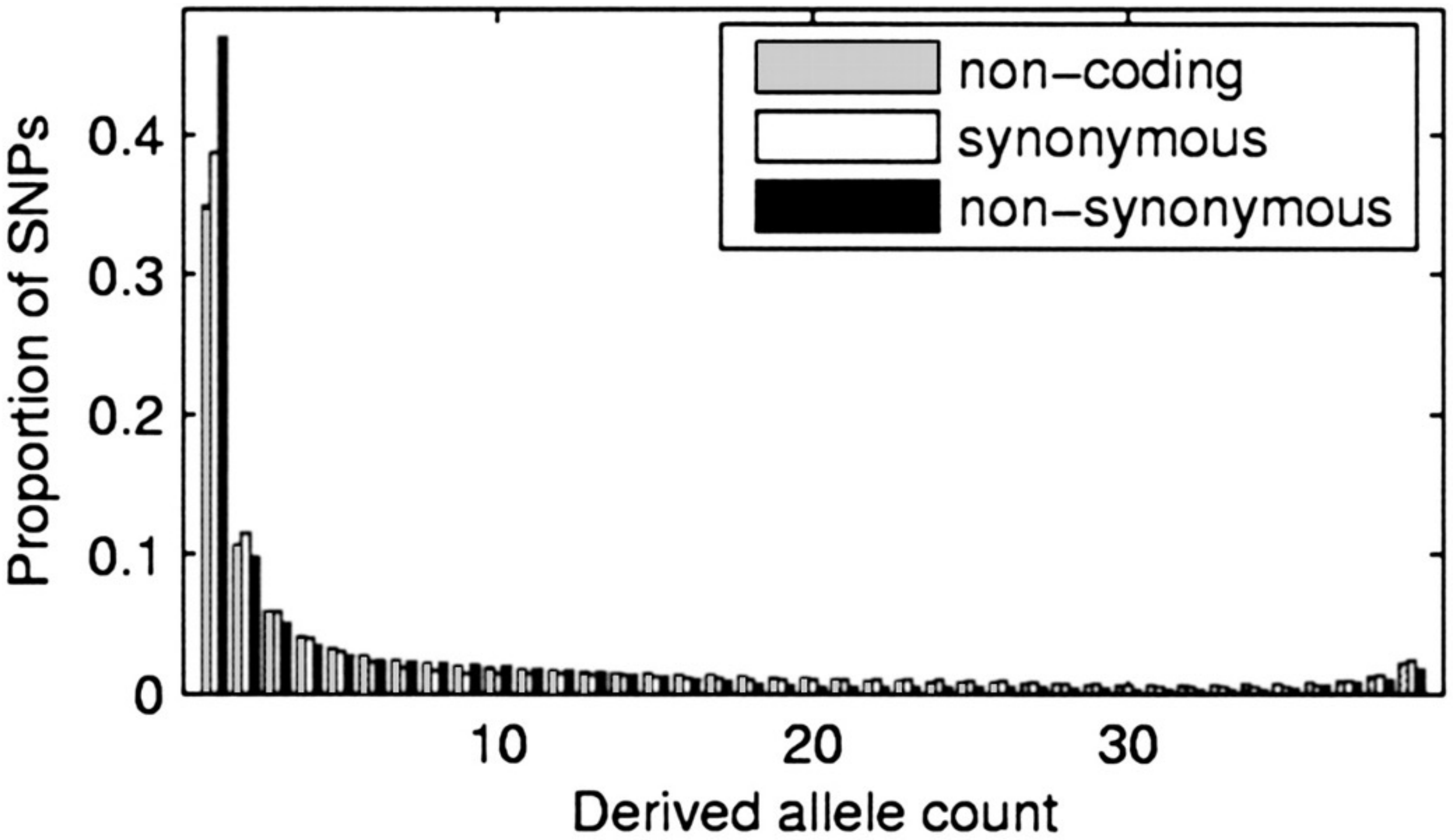
Natural selection:

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

GCTTACCGGAATTATGCG		GCTTACCGTAATTATGCG
GCTTACAGGAATTATGCG		GCTTACAGGAATTATGCG
GCTCACCGGATTTATGCG		GCTCACCGGATTTATGCG
GCTTACAGGATTTATGCG		GCTTACAGGATTTATGCG
GCTTACCGGAATTATGGG	→	GCTTACCGGAATTATGGG
GCTTACCGGATTTATGCG		GCTTACCGGATTTATGCG
CCTTACCGGAATTATTCG		CCTTACCGGAATTATTCG
CCTTACCGGAATTATTCG		CCTTACCGGAATTATTCG
CCTTACAGGAATTATGCG		CCTTACAGGAATTATGCG
GCTTACCGGATTTATGCG		GCTTACCGGATTTATGCG

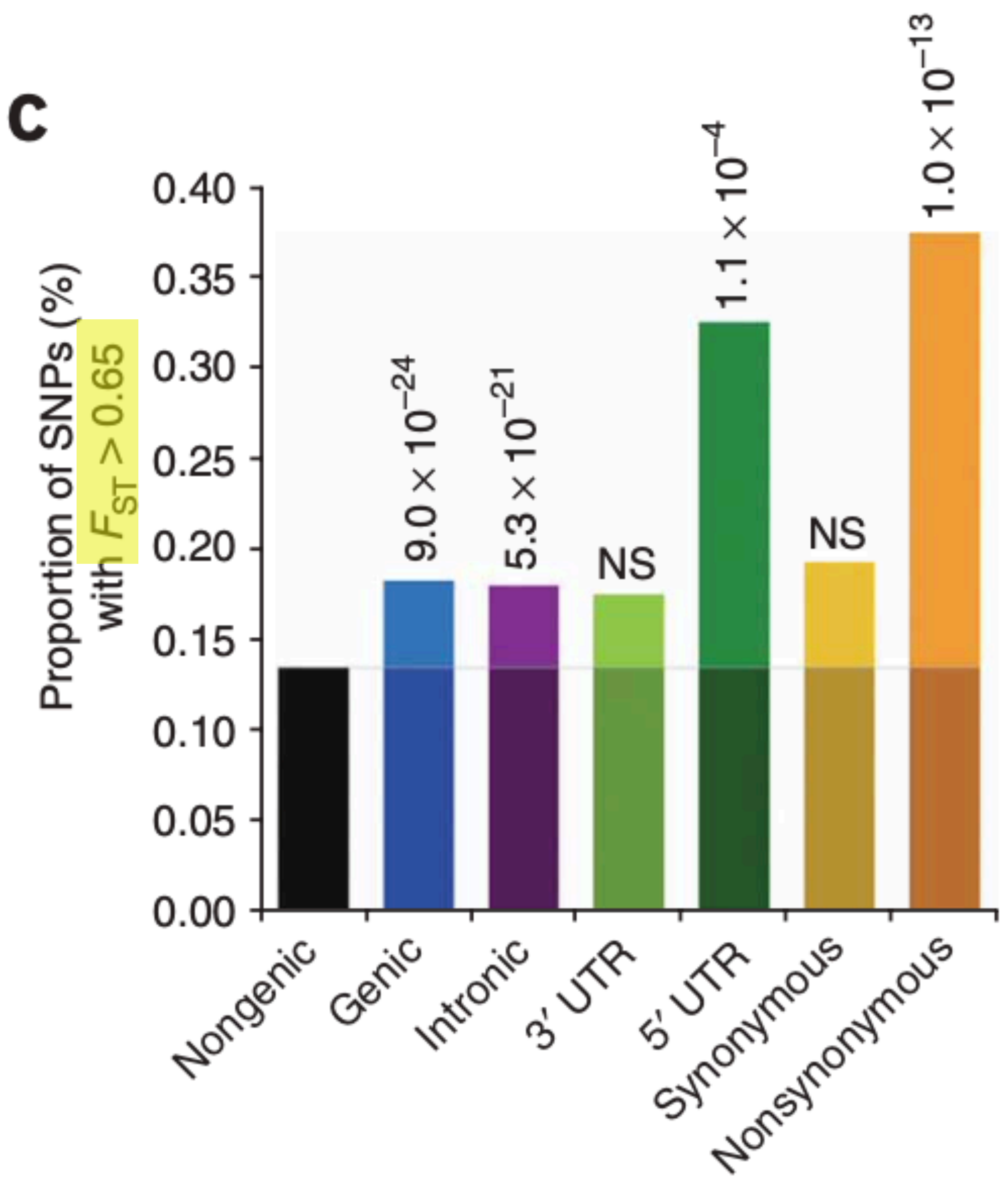


Barreiro et al. (2008)



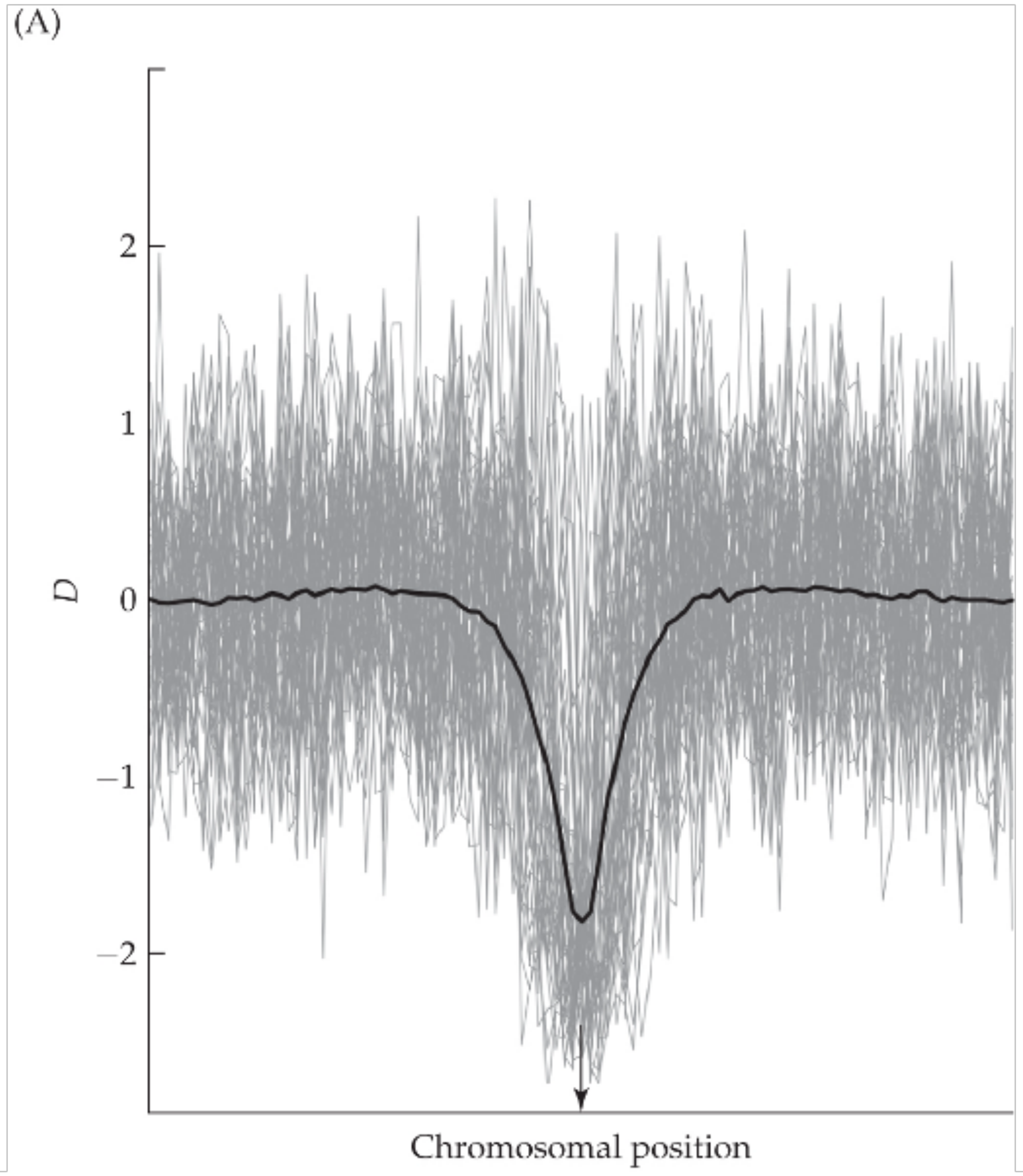
Williamson et al. (2007)

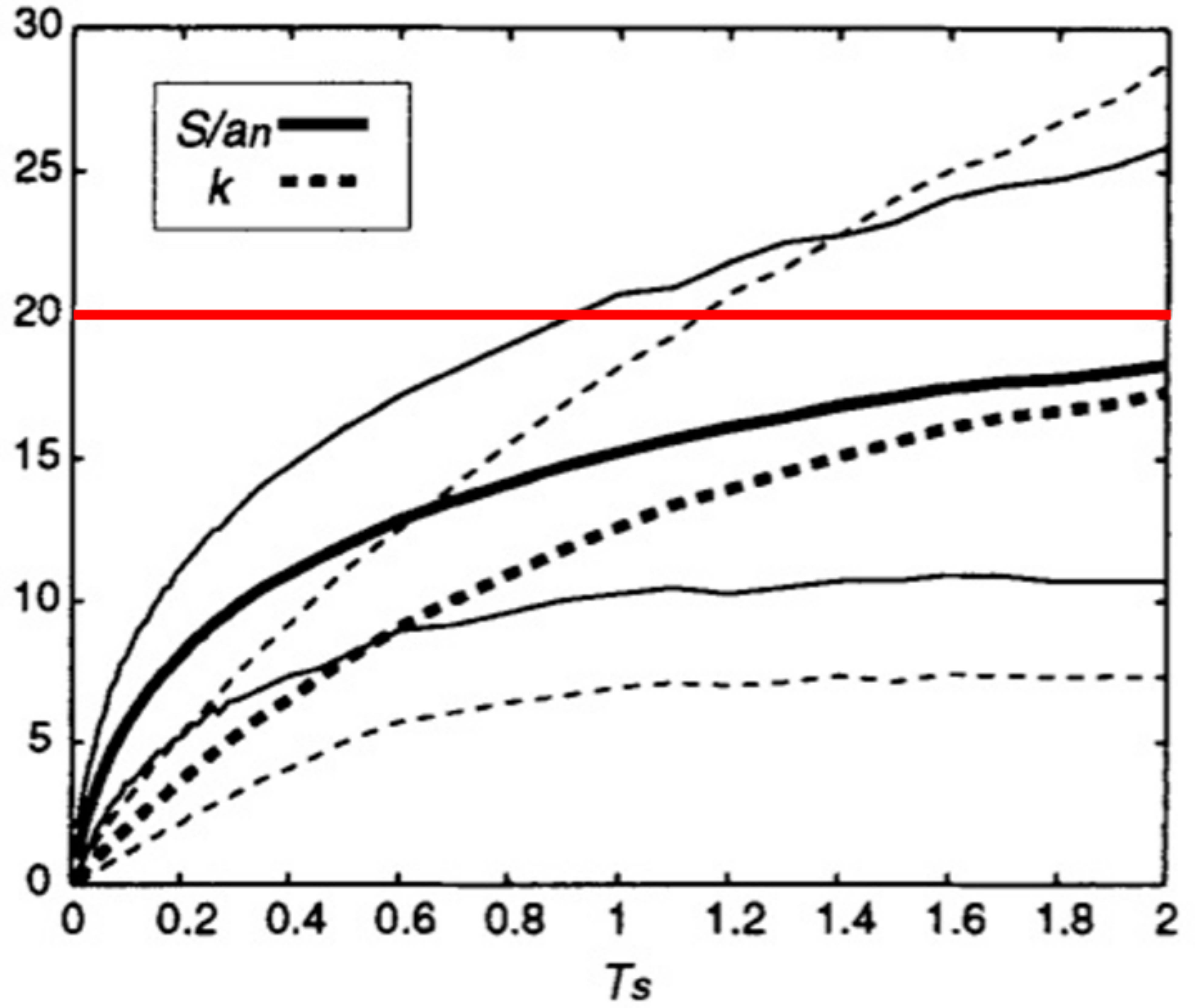
GCTTACCGGAATTATGCG		GCTTACCGTAATTATGCG		GCTTACCGTAATTATGCG
GCTTACAGGAATTATGCG		GCTTACAGGAATTATGCG		GCTTACCGTAATTATGCG
GCTCACCGGATTTATGCG		GCTCACCGGATTTATGCG		GCTTACCGTAATTATGCG
GCTTACAGGATTTATGCG		GCTTACAGGATTTATGCG		GCTTACCGTAATTATGCG
GCTTACCGGAATTATGGG	→	GCTTACCGGAATTATGGG	→	GCTTACCGTAATTATGCG
GCTTACCGGATTTATGCG		GCTTACCGGATTTATGCG		GCTTACCGTAATTATGCG
CCTTACCGGAATTATTCG		CCTTACCGGAATTATTCG		GCTTACCGTAATTATGCG
CCTTACCGGAATTATTCG		CCTTACCGGAATTATTCG		GCTTACCGTAATTATGCG
CCTTACAGGAATTATGCG		CCTTACAGGAATTATGCG		GCTTACCGTAATTATGCG
GCTTACCGGATTTATGCG		GCTTACCGGATTTATGCG		GCTTACCGTAATTATGCG



Barreiro et al. (2008)

GCTTACCGTAATTATGCG	GCTTGCCGTAATTATGCG
GCTTACCGTAATTATGCG	GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG	GCTTACCGTAATTCTGCG
GCTTACCGTAATTATGCG	GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG	GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG	GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG	GCTTACCGTAAATATGCG
GCTTACCGTAATTATGCG	GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG	GCTTACCGTAATTATGCG
GCTTACCGTAATTATGCG	GTTTACCGTAATTATGCG


(A)
2
1
0
−1
−2
D
Chromosomal position

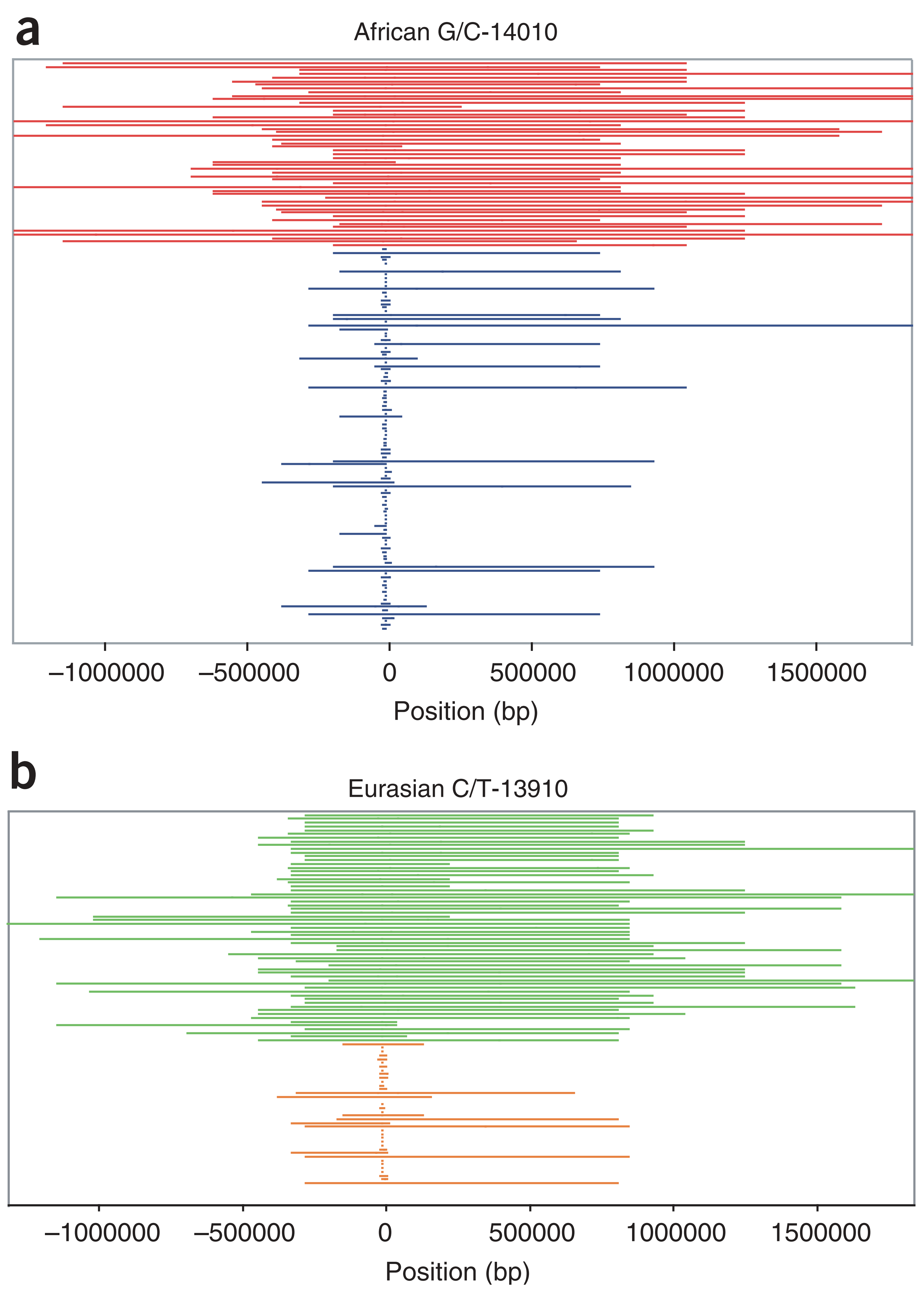


Braverman et al. (1995)

GCTTACCGGAATTATGCG		GCTTACCGTAATTATGCG		GCTTACCGTAATTATGCG
GCTTACAGGAATTATGCG		GCTTACAGGAATTATGCG		GCTTACCGTAATTATGCG
GCTCACCGGATTTATGCG		GCTCACCGGATTTATGCG		GCTTACCGTAATTATGCG
GCTTACAGGATTTATGCG		GCTTACAGGATTTATGCG		GCTTACCGTAATTATGCG
GCTTACCGGAATTATGGG	→	GCTTACCGGAATTATGGG	→	GCTTACCGTAATTATGCG
GCTTACCGGATTTATGCG		GCTTACCGGATTTATGCG		GCTTACCGTAATTATGCG
CCTTACCGGAATTATTCG		CCTTACCGGAATTATTCG		GCTTACCGTAATTATGCG
CCTTACCGGAATTATTCG		CCTTACCGGAATTATTCG		GCTTACCGTAATTATGCG
CCTTACAGGAATTATGCG		CCTTACAGGAATTATGCG		GCTTACCGTAATTATGCG
GCTTACCGGATTTATGCG		GCTTACCGGATTTATGCG		GCTTACCGTAATTATGCG

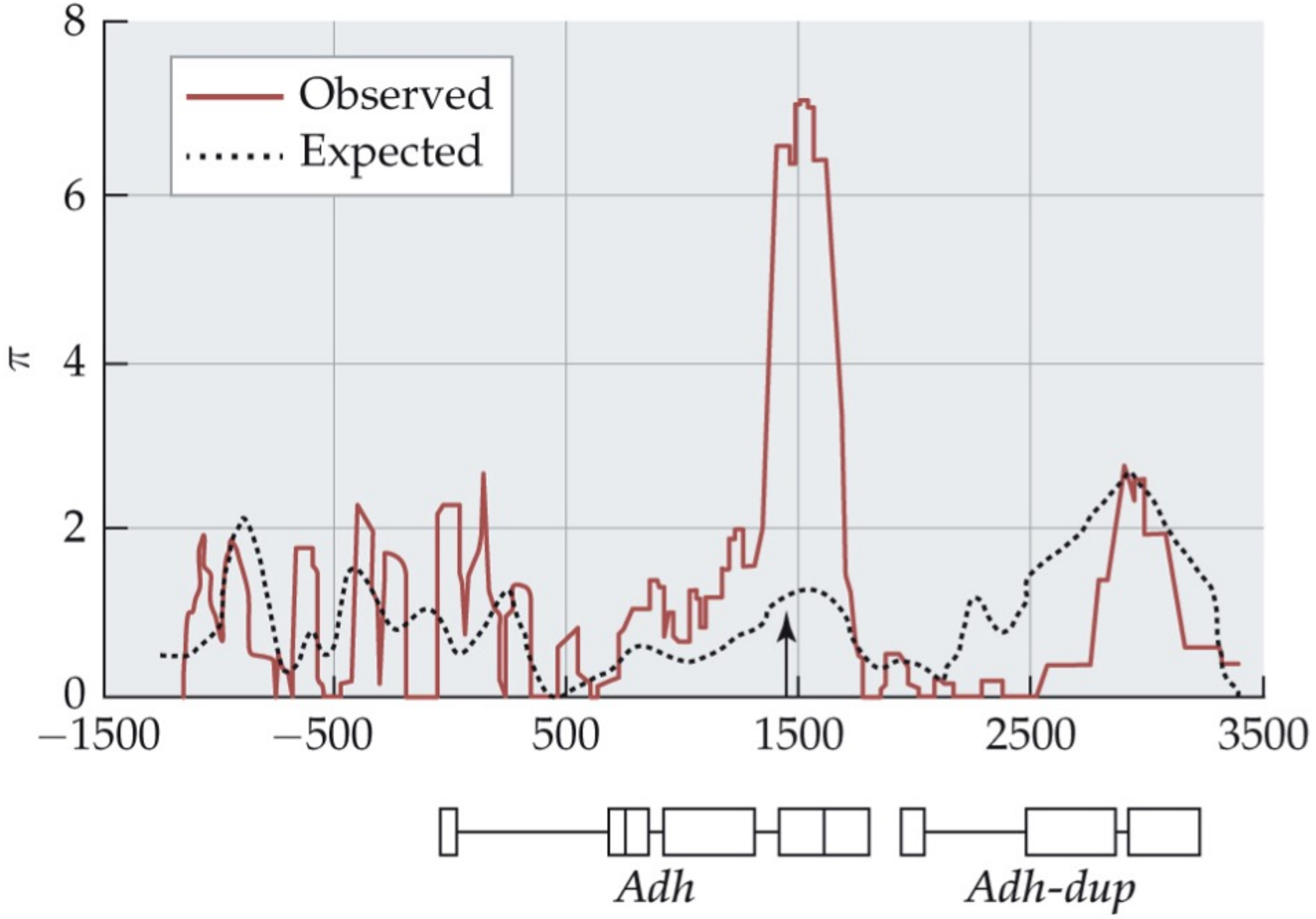

GCTTACCGGAATTATGCG
GCTTACAGGAATTATGCG
GCTCACCGGATTTATGCG
GCTTACAGGATTTATGCG
GCTTACCGGAATTATGGG
GCTTACCGGATTTATGCG
CCTTACCGGAATTATTCG
CCTTACCGGAATTATTCG
CCTTACAGGAATTATGCG
GCTTACCGGATTTATGCG
GCTTACCGGAATTATGCG
GCTTACAGGAATTATGCG
GCTCACCGGATTTATGCG
GCTTACAGGATTTATGCG
GCTTACCGGAATTATGGG
GCTTACCGGATTTATGCG
CCTTACCGGAATTATTCG
CCTTACCGGAATTATTCG
CCTTACAGGAATTATGCG
GCTTACCGGATTTATGCG
GCTTACAGTAATTATGCG
GCTTACAGTAATTATGCG
GCTTACAGTAATTATGCG
GCTTACAGTAATTATGCG
GCTTACAGTAATTATGCG
GCTTACAGTAATTATGCG
CCTTACAGTAATTATGCG
CCTTACAGTAATTATGCG
CCTTACAGTAATTATGCG
CCTTACAGTAATTATGCG

GCTTACCGGAATTATGCG		GCTTACCGTAATTATGCG		GCTTACCGTAATTATGCG
GCTTACAGGAATTATGCG		GCTTACAGGAATTATGCG		GCTTACCGTAATTATGCG
GCTCACCGGATTTATGCG		GCTCACCGGATTTATGCG		GCTTACCGTAATTATGCG
GCTTACAGGATTTATGCG		GCTTACAGGATTTATGCG		GCTTACCGTAATTATGCG
GCTTACCGGAATTATGGG	→	GCTTACCGGAATTATGGG	→	GCTTACCGTAATTATGCG
GCTTACCGGATTTATGCG		GCTTACCGGATTTATGCG		GCTTACCGGATTTATGCG
CCTTACCGGAATTATTCG		CCTTACCGGAATTATTCG		CCTTACCGGAATTATTCG
CCTTACCGGAATTATTCG		CCTTACCGGAATTATTCG		CCTTACCGGAATTATTCG
CCTTACAGGAATTATGCG		CCTTACAGGAATTATGCG		CCTTACAGGAATTATGCG
GCTTACCGGATTTATGCG		GCTTACCGGATTTATGCG		GCTTACCGGATTTATGCG



Tishkoff et al. (2007)

GCTTACCGGAATTATGCG		GCTCACCGGAATTATGCG		GCTCACCGGAATTATGCG
GCTTACCGGAATTATGCG		GCTTACCGGAATTATGCG		GCTTACAGGAATTATGCG
GCTTACCGGAATTATGCG		GCTTACCGGAATTATGCG		GCTCACCGGAATTATGCG
GCTTACCGGAATTATGCG		GCTTACCGGAATTATGCG		GCTTACAGGAATTATGCG
GCTTACCGGAATTATGCG		GCTTACCGGAATTATGCG		GCTTACCGGAATTATGGG
GCTTACCGGGATTATGCG	→	GCTTACCGGGATTATGCG	→	GCTTACCGGGATTATGCG
GCTTACCGGGATTATGCG		GCTTACCGGGATTATTCG		CCTTACCGGGATTATTCG
GCTTACCGGGATTATGCG		GCTTACCGGGATTATGCG		CCTTACCGGGATTATTCG
GCTTACCGGGATTATGCG		GCTTACCGGGATTATGCG		CCTTACCGGGATTATGCG
GCTTACCGGGATTATGCG		GCTTACCGGGATTATGCG		GCTTACCGGGATTATGCG



From Kreitman and Hudson, 1991. *Genetics* 127:565-582.