CHAPTER 9

POPULATION SUBDIVISION

There is abundant geographical variation in both morphology and gene frequency in most species. The extent of geographic variation results from a balance of forces tending to produce local genetic differentiation and forces tending to produce genetic homogeneity.

Montgomery Slatkin (1987)

The term "species" includes any subspecies of fish or wildlife or plants, and any ***distinct population segment*** *of any species of vertebrate fish or wildlife which breeds when mature.*

Endangered Species Act of the U.S.A.

Assumptions of Hardy-Weinberg model

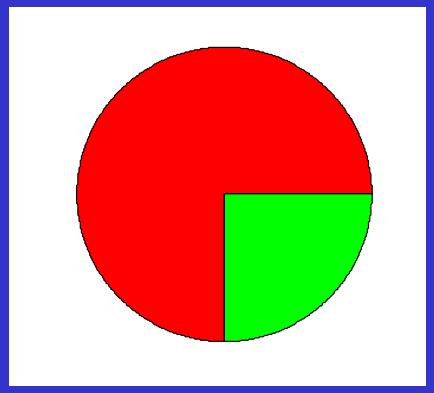
1. Random mating.
2. No mutation.
3. Large (infinite) population size.
4. No differential survival or reproduction (i.e., no natural selection).
5. No immigration

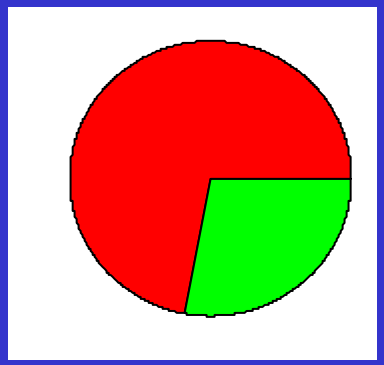
The total amount of genetic variation within a species (H_T) can be partitioned into genetic differences among individuals within populations (H_S) and genetic differences among different populations. The proportion of total genetic variation within a species that is due to differences among populations is represented by F_{ST} :

$$F_{ST} = 1 - \frac{H_S}{H_T}$$

Pop A Pop B

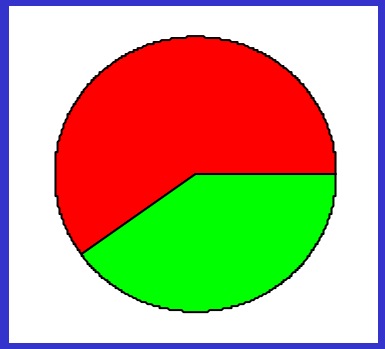
(1)

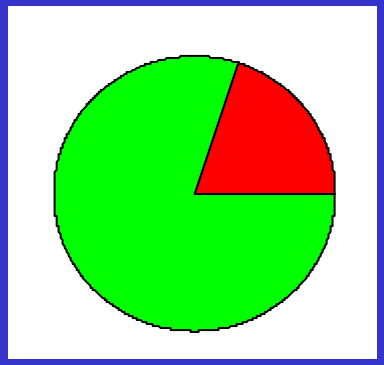




$F_{ST} \sim 0$

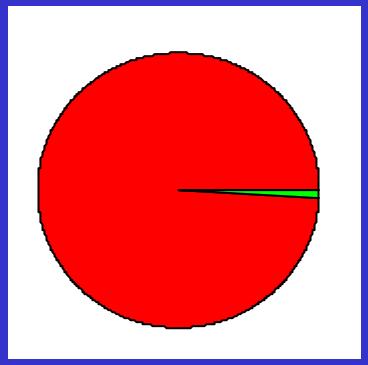
(2)

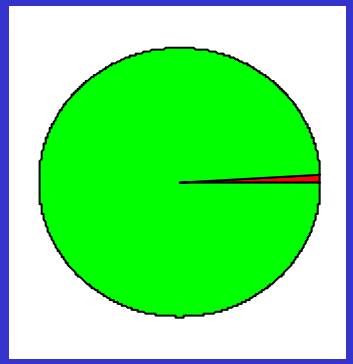




$F_{ST} \sim 0.3$

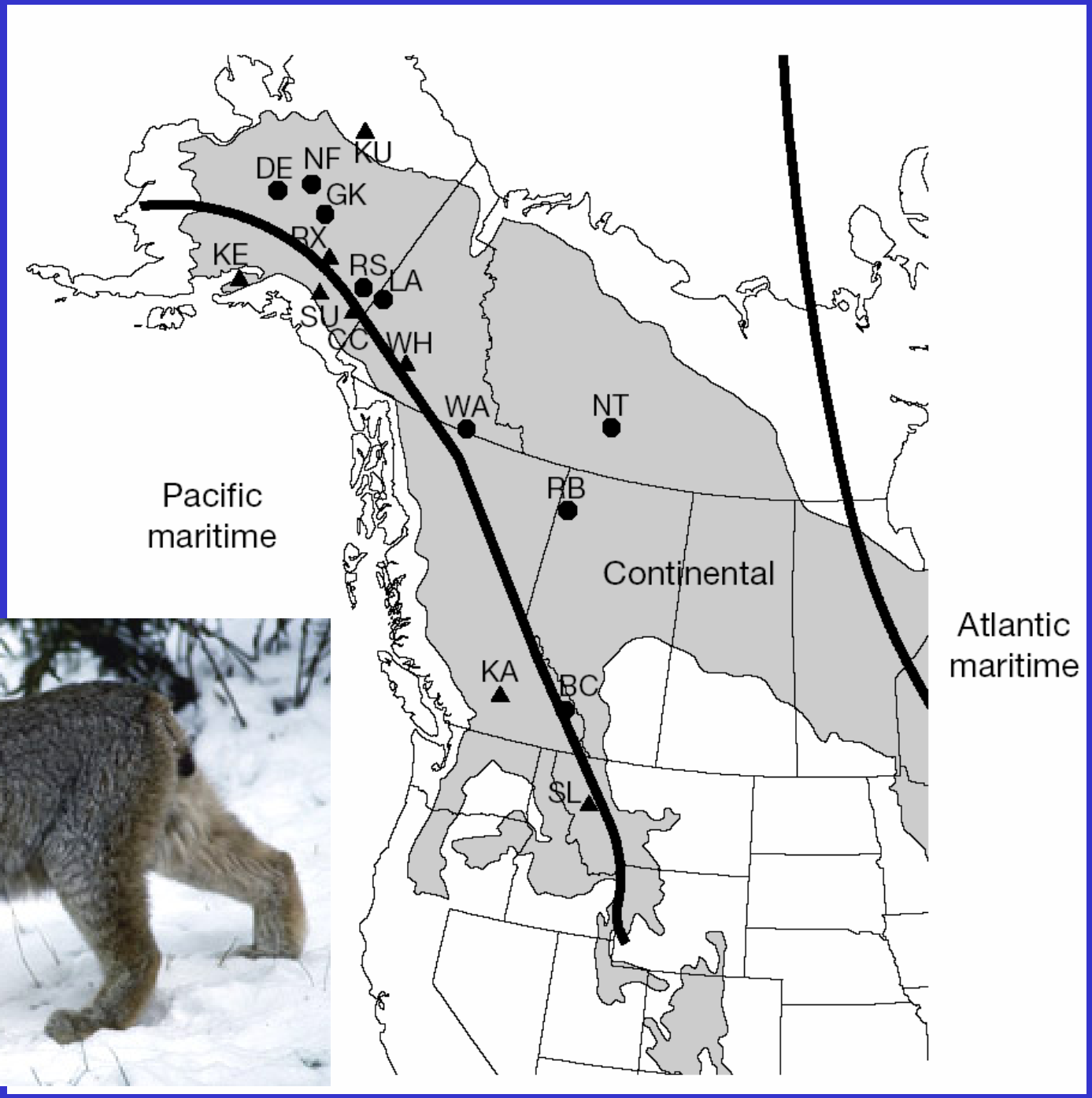
(3)

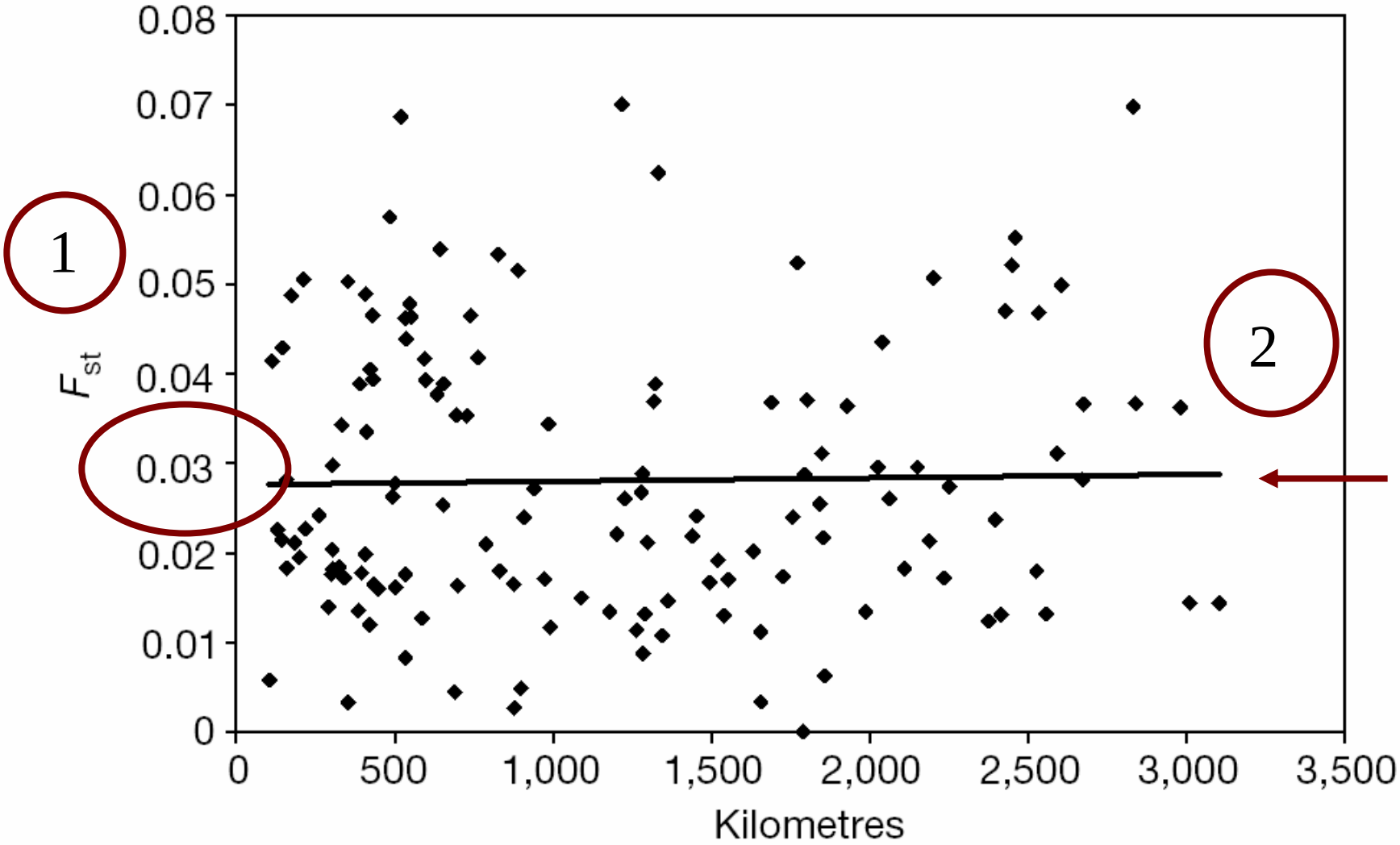


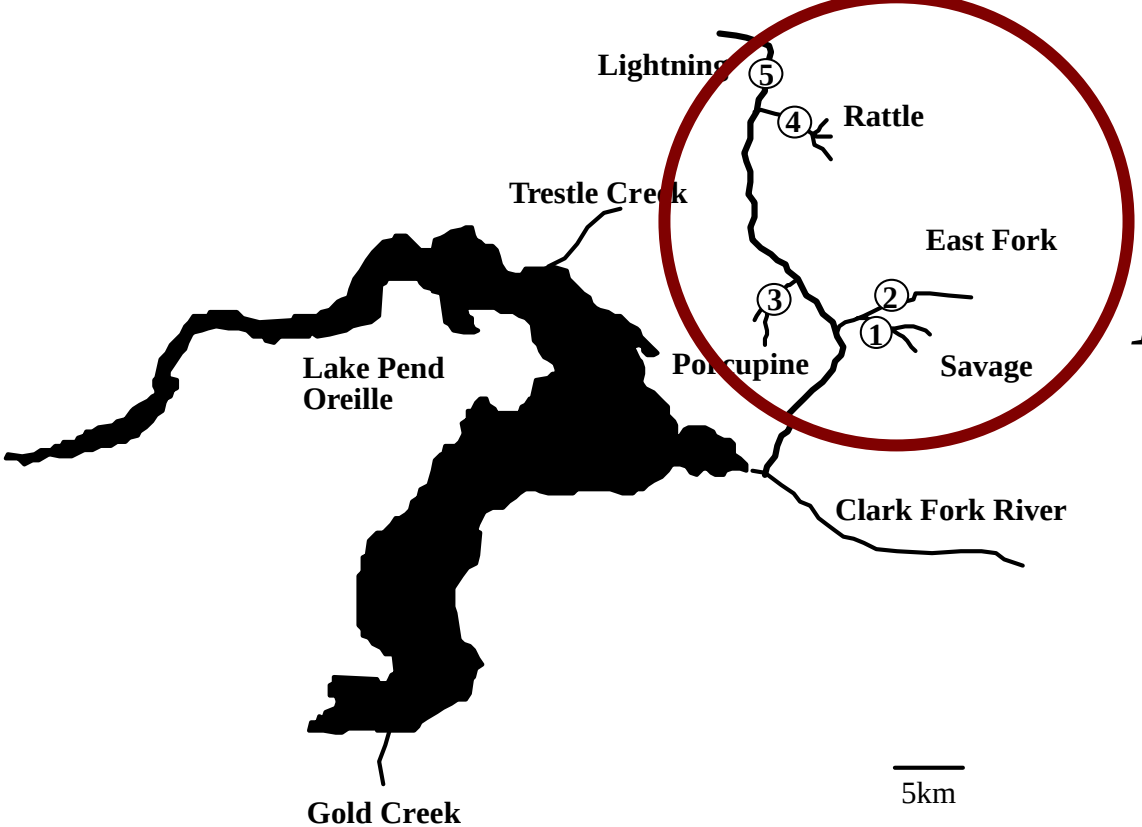


$F_{ST} \sim 1.0$

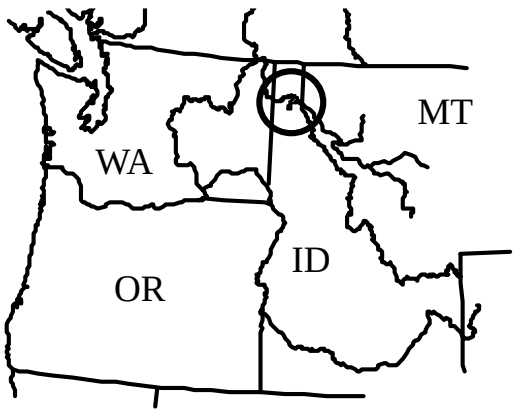
Taxa	H_T	H_S	F_{ST}	No. species
Amphibians	0.136	0.094	0.315	33
Birds	0.059	0.054	0.076	16
Fish	0.067	0.054	0.135	79
Mammals	0.078	0.054	0.242	57
Reptiles	0.124	0.090	0.258	22
Crustaceans	0.088	0.063	0.169	19
Insects	0.138	0.122	0.097	46
Mollusks	0.157	0.121	0.263	44

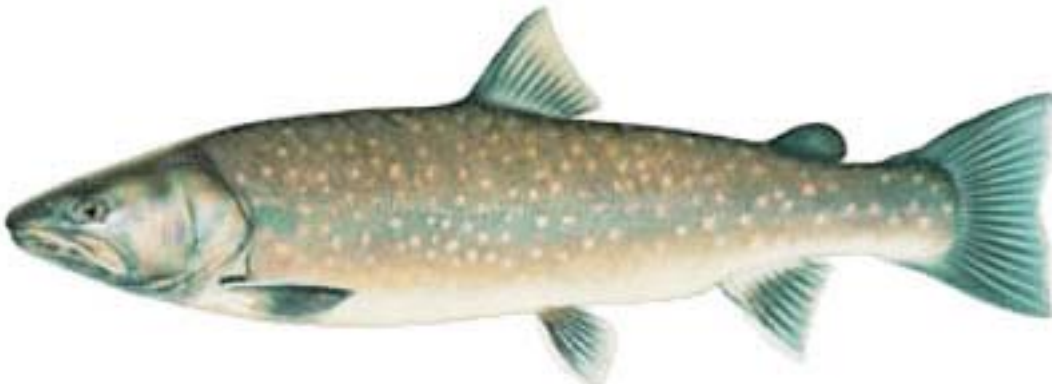

KU
DE
NF
GK
RX
KE
RS
LA
SU
CC
WH
WA
NT
RB
Pacific
maritime
Continental
Atlantic
maritime
KA
BC
SL

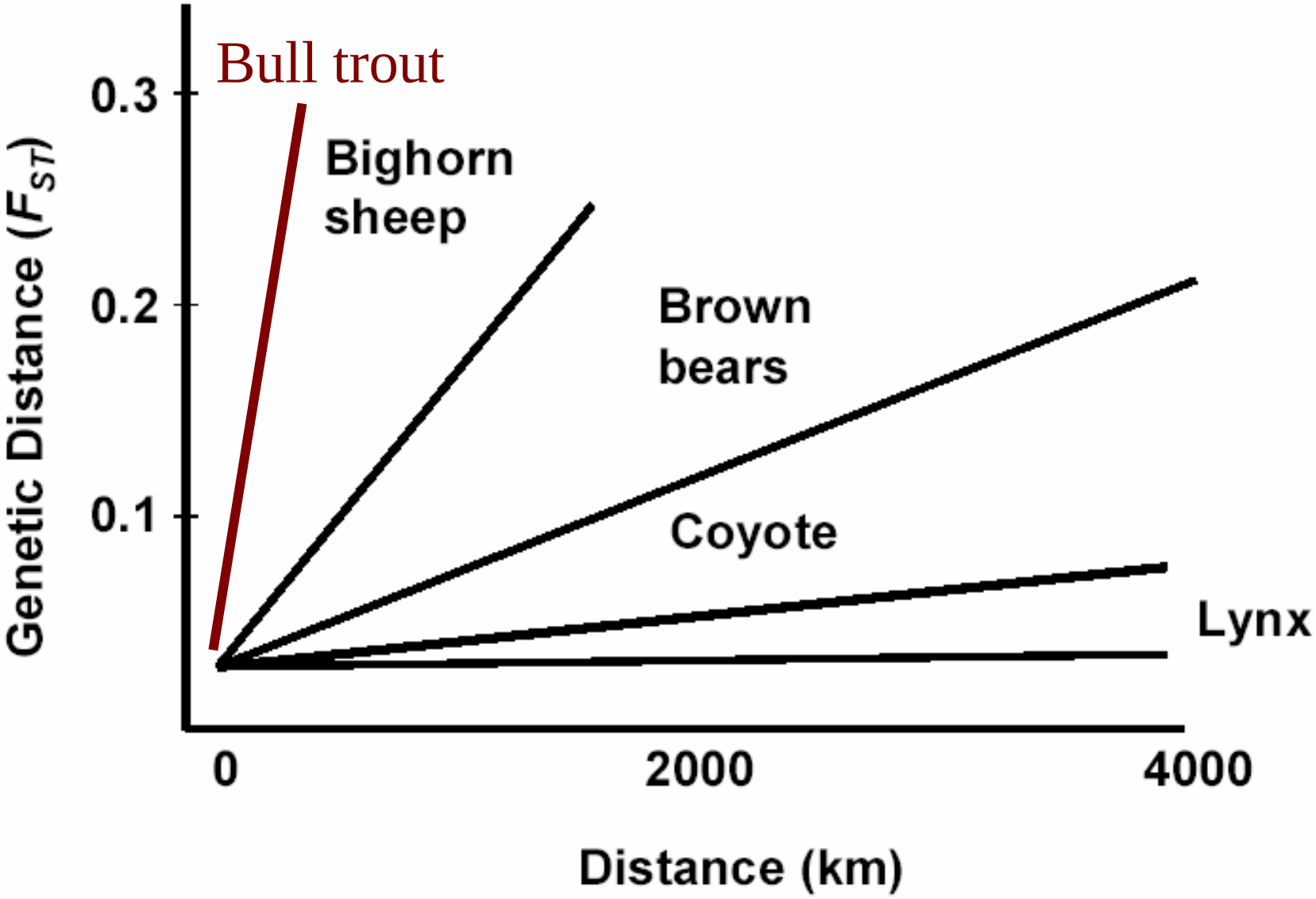

0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
F_{st}
0
500
1,000
1,500
2,000
2,500
3,000
3,500
Kilometres
1
2


Lightning
5
4
Rattle
Trestle Creek
East Fork
3
2
1
Porcupine
Savage
Lake Pend
Oreille
Clark Fork River
Gold Creek
5km

$$F_{ST} = 0.063$$


MT
WA
ID
OR



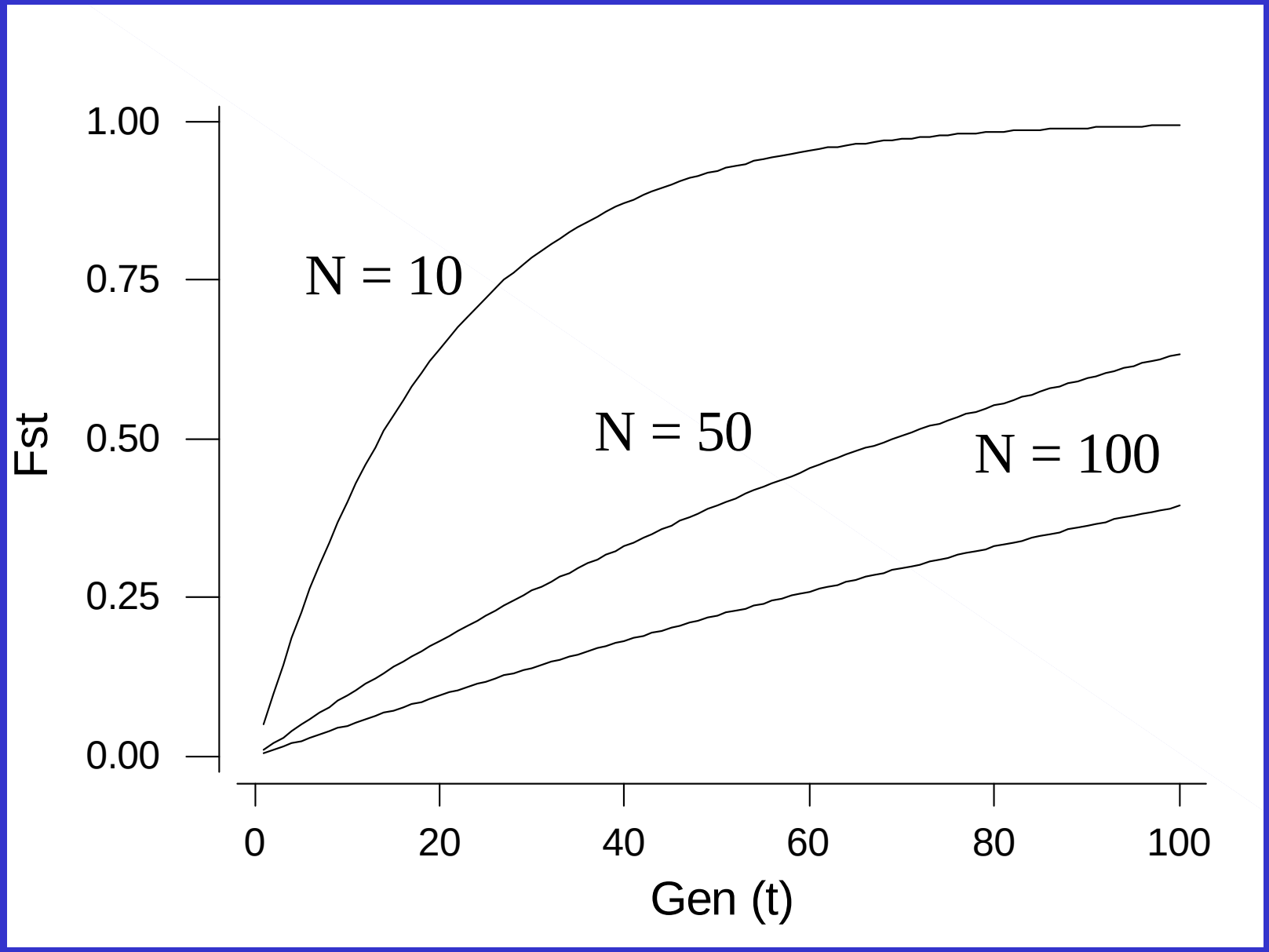

Bull trout
Bighorn sheep
Brown bears
Coyote
Lynx
0.3
0.2
0.1
Genetic Distance (F_{ST})
0
2000
4000
Distance (km)

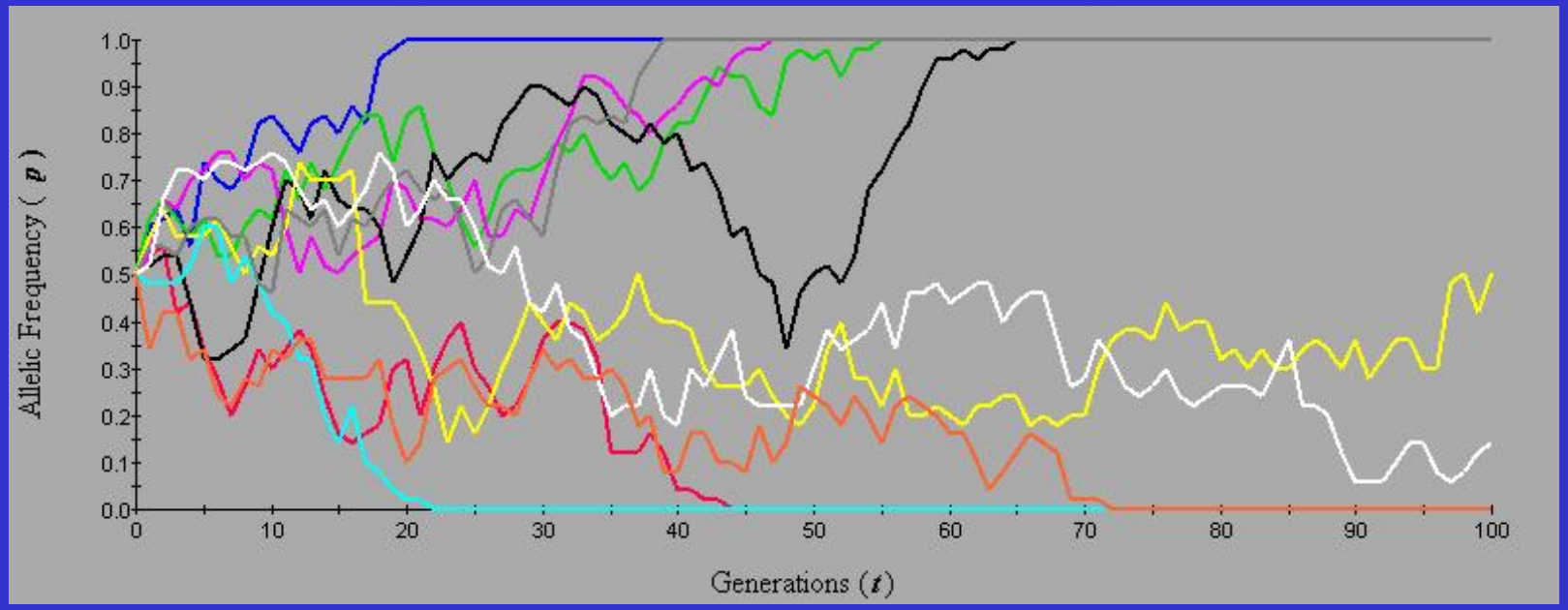
Complete Isolation

$$F_{ST} = 1 - (1 - \frac{1}{2N_e})^t$$

(1) Eventual fixation

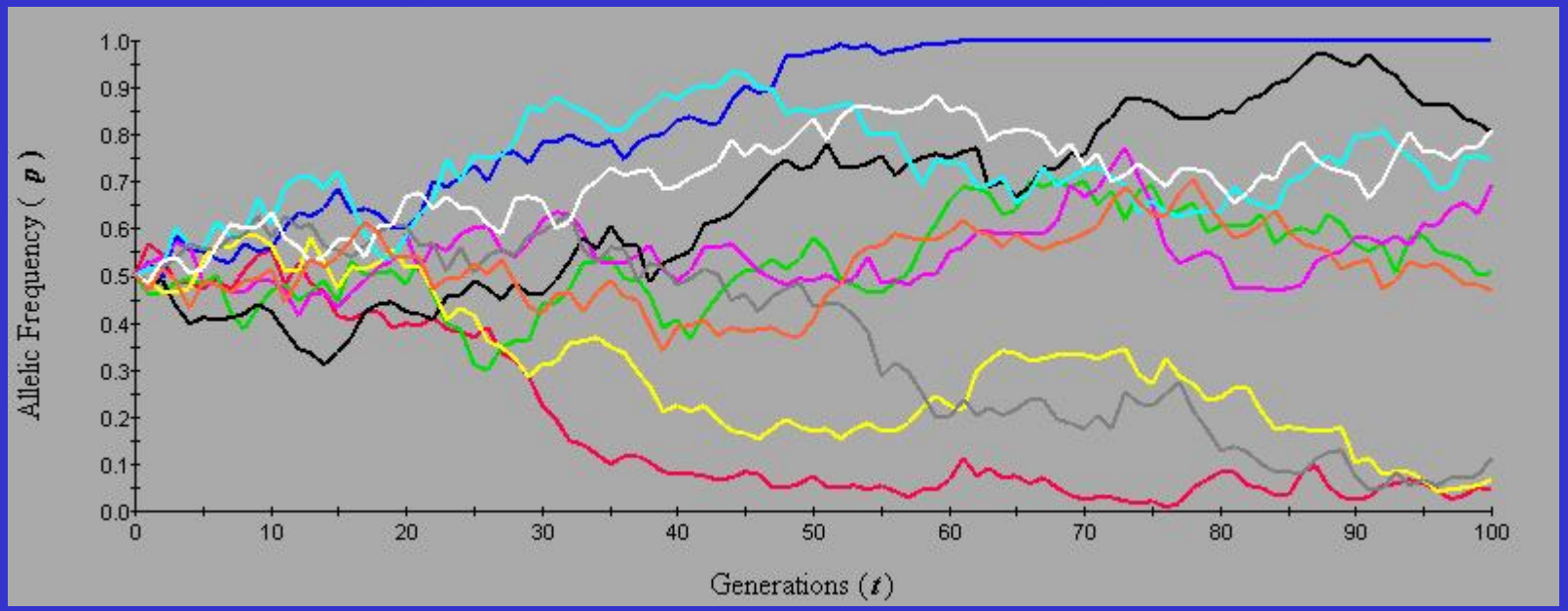
(2) Rate dependent upon N_e


1.00
0.75
0.50
0.25
0.00
Fst
N = 10
N = 50
N = 100
0
20
40
60
80
100
Gen (t)



N=25

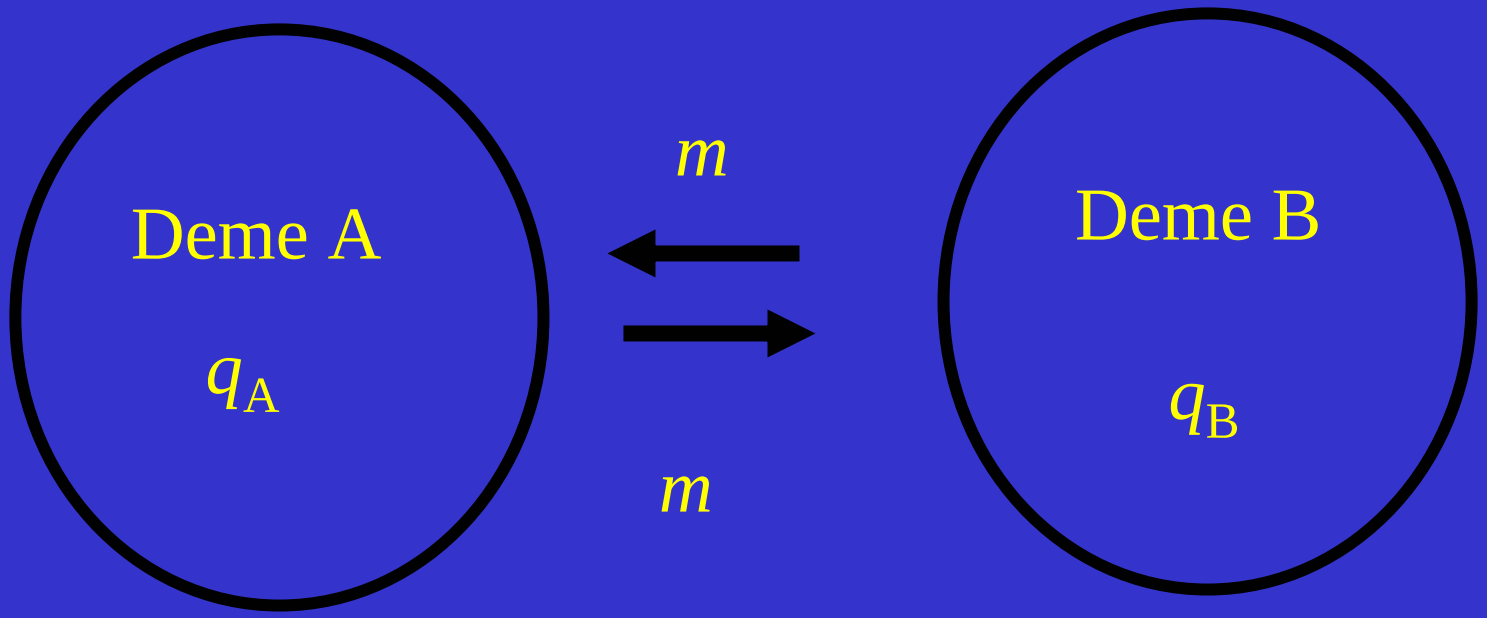
mN=0



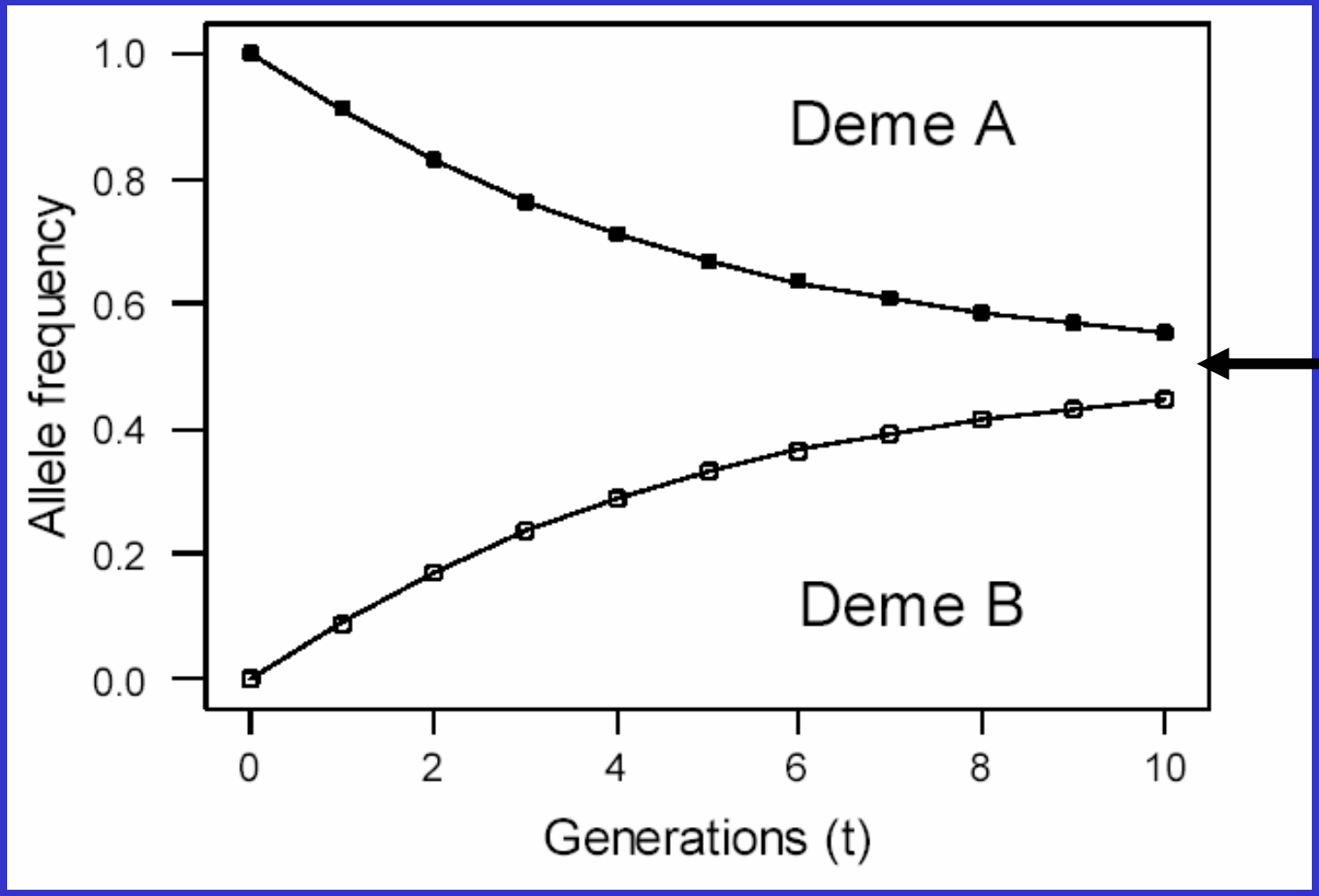
N=100

mN=0

Gene Flow



$$q'_A = (1 - m)q_A + mq_B$$


1.0
0.8
0.6
0.4
0.2
0.0
Allele frequency
Deme A
Deme B
0
2
4
6
8
10
Generations (t)

q^*

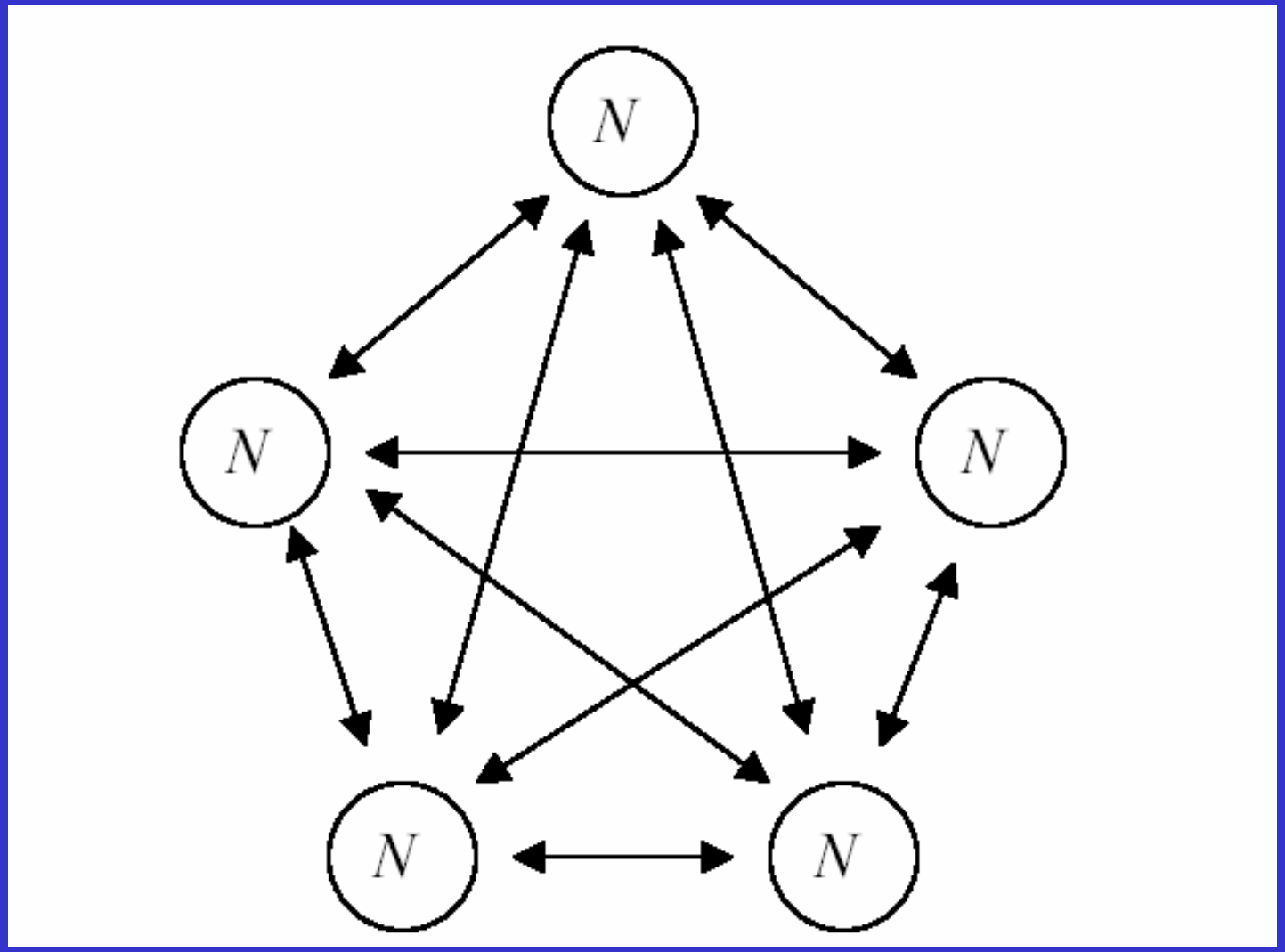
In general, there are two primary effects of gene flow:

(1) Gene flow reduces genetic differences between populations.

(2) Gene flow increases genetic variation within populations.

What happens if we combine gene flow and genetic drift?

Island Model of Migration



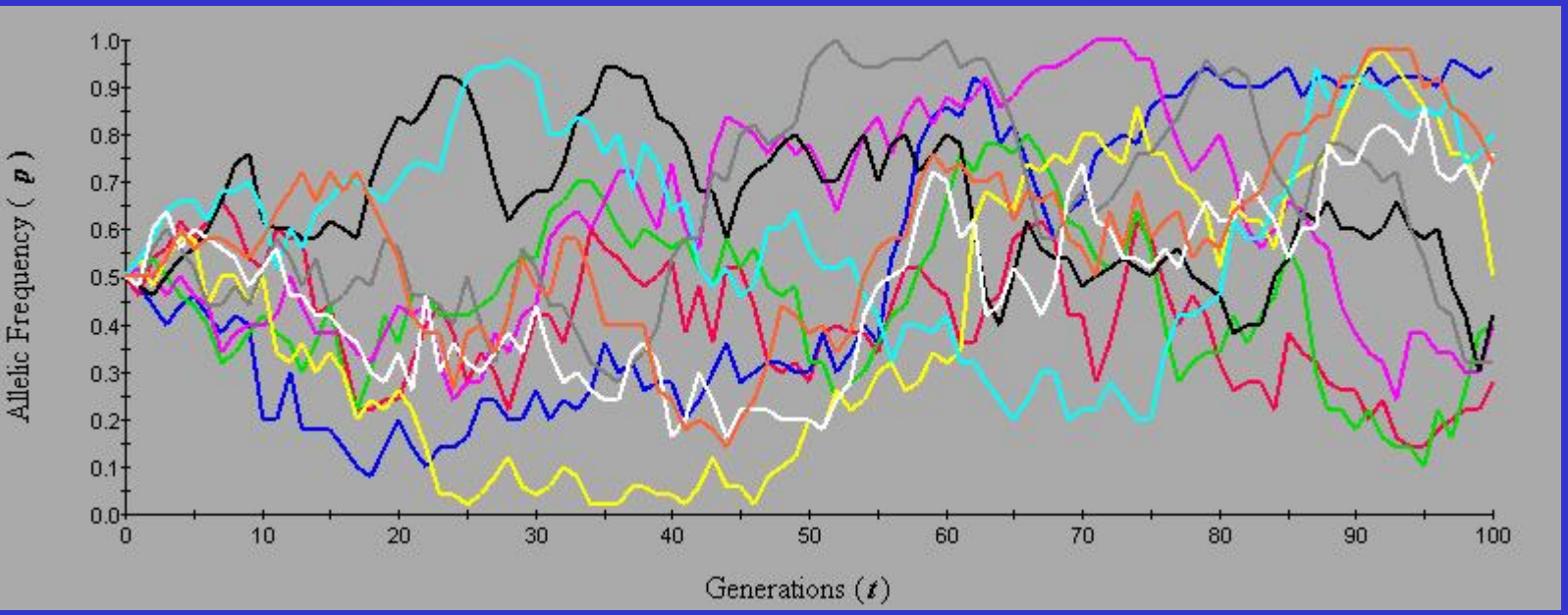
$$F^*_{ST} = \frac{(1-m)^2}{[2N-(2N-1)(1-m)^2]}$$

$$F^*_{ST} \approx \frac{1}{(4mN+1)}$$

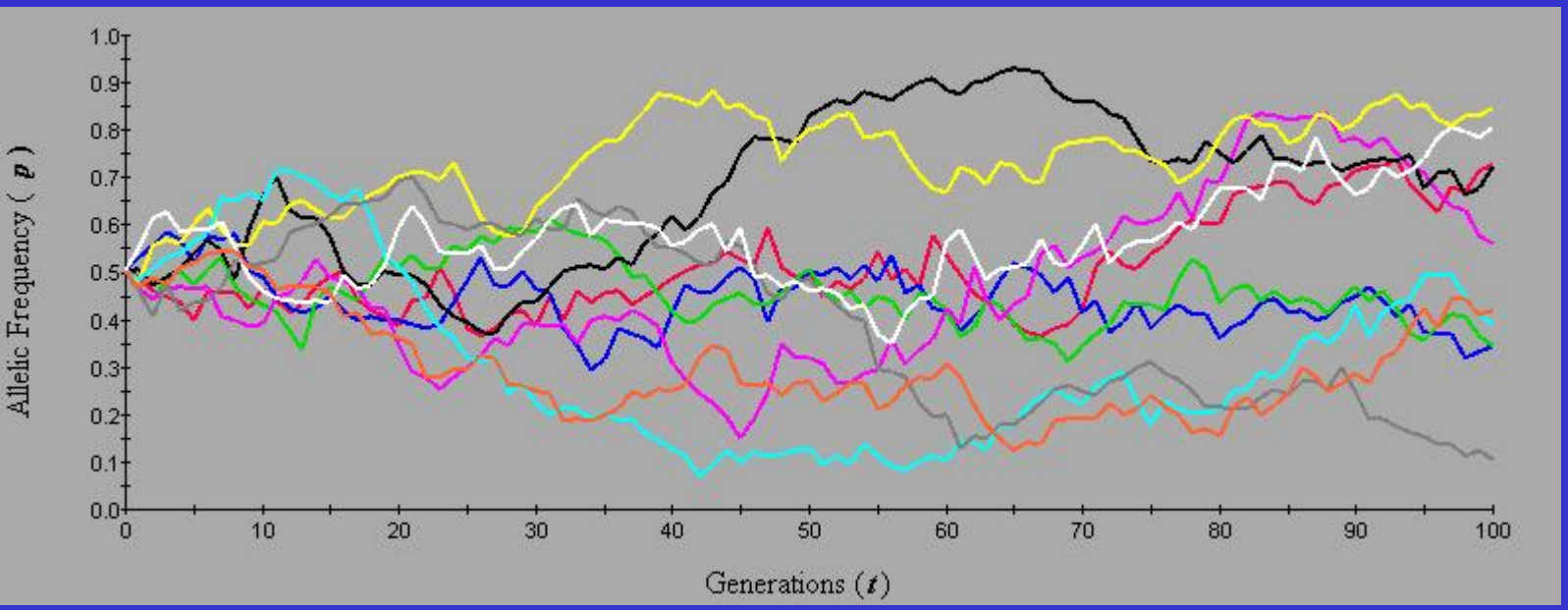
(1) Divergence is dependent only on *mN*, not on the proportional contribution of migrants (*m*).

(2) Even very little gene flow (*mN*~1) can have a major effect on population structure.

(1) Divergence is dependent only on mN, not on the proportional contribution of migrants (m). mN is the number of migrant individuals per individual per generation.



N=25
mN=1
m=0.04



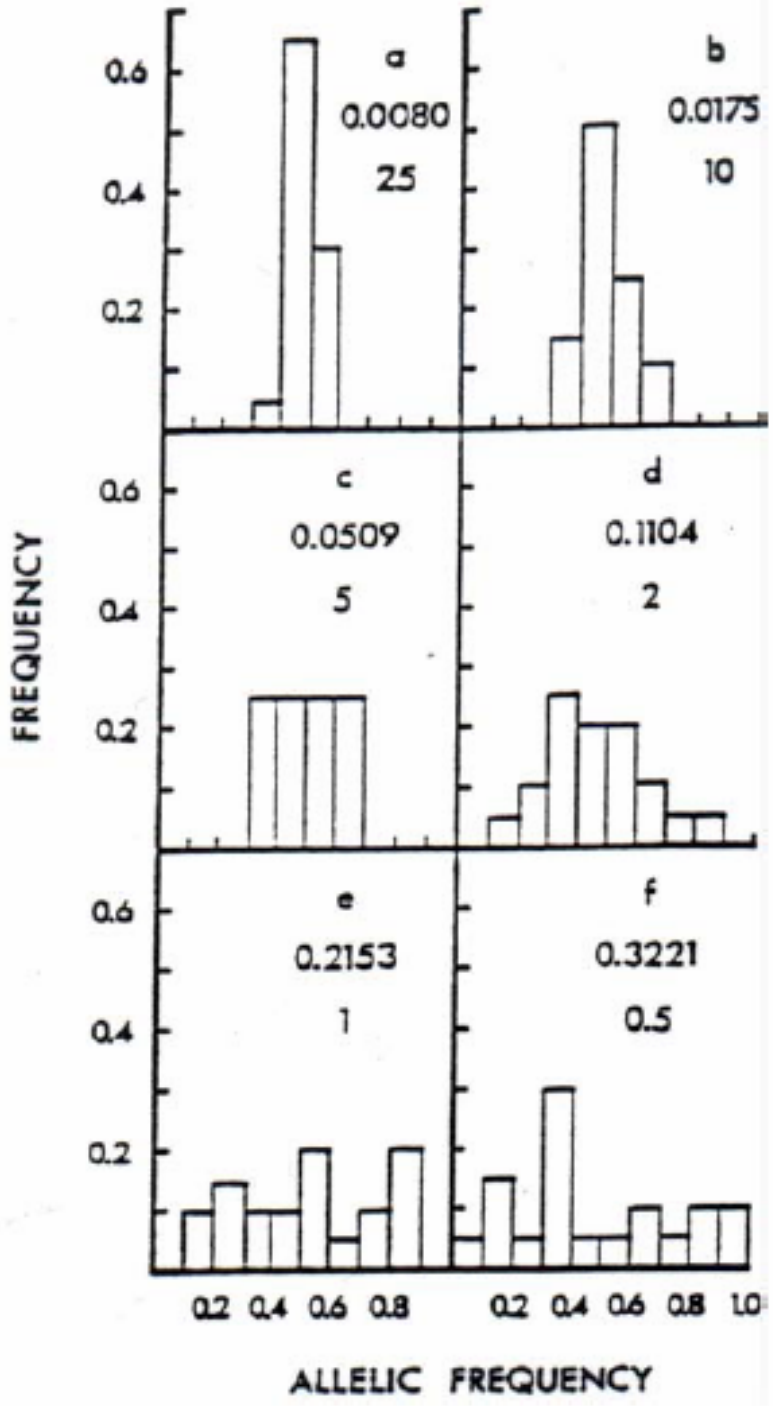
N=100
mN=1
m=0.01

(2) Even very little gene flow (mN~1) can have a major effect on population structure.

25

5

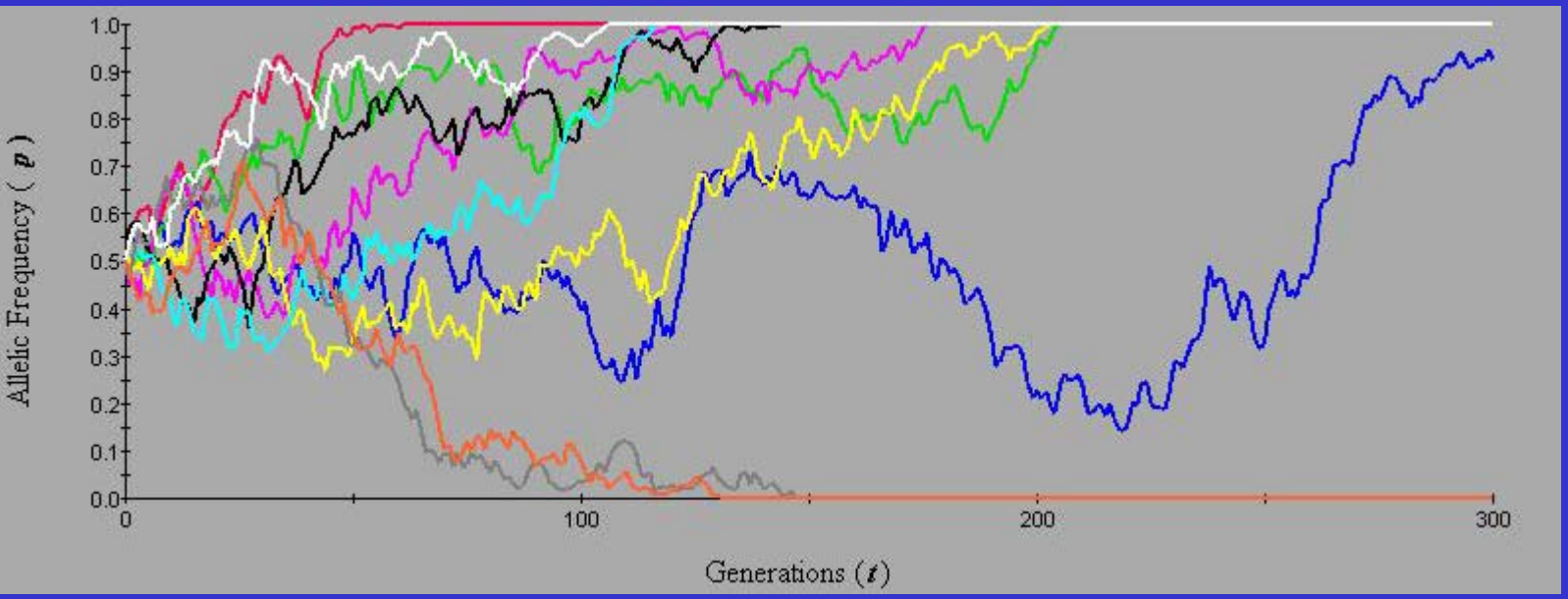
1


a
0.0080
25
b
0.0175
10
c
0.0509
5
d
0.1104
2
e
0.2153
1
f
0.3221
0.5
0.6
0.4
0.2
FREQUENCY
0.2 0.4 0.6 0.8
0.2 0.4 0.6 0.8 1.0
ALLELIC FREQUENCY

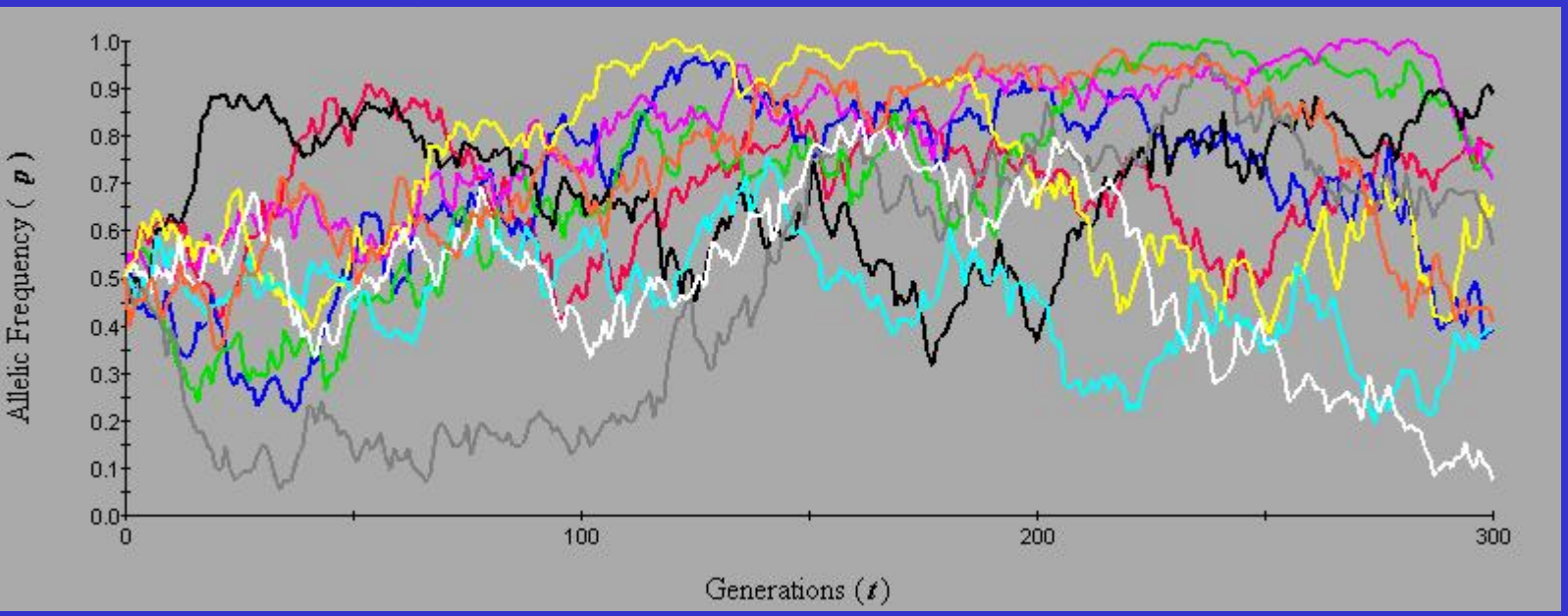
10

2

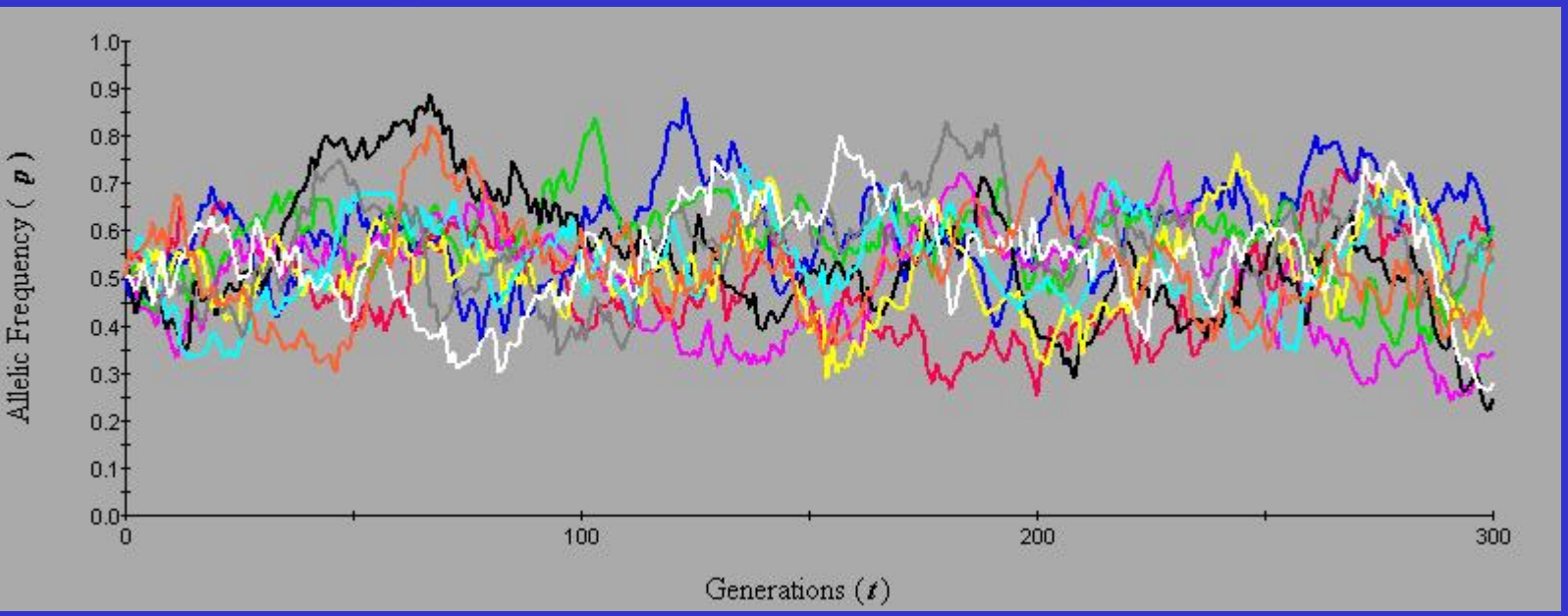
0.5



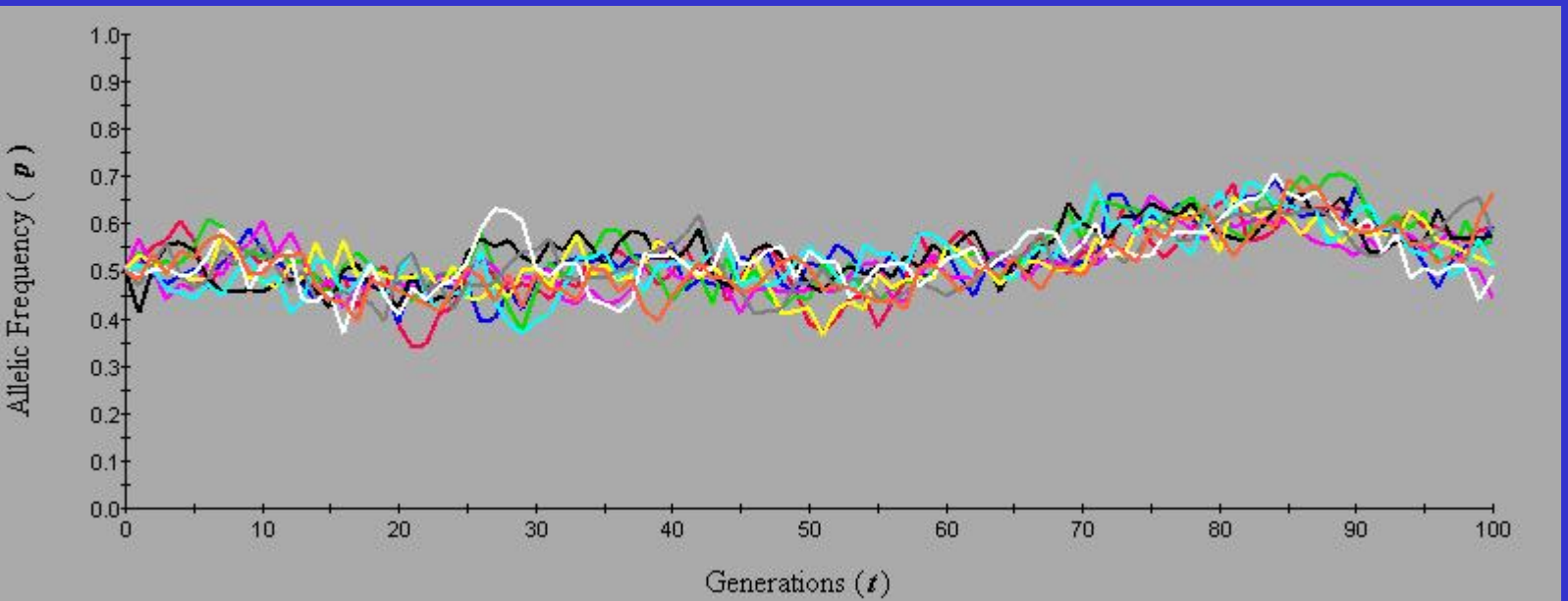
N=100
mN=0
m=0



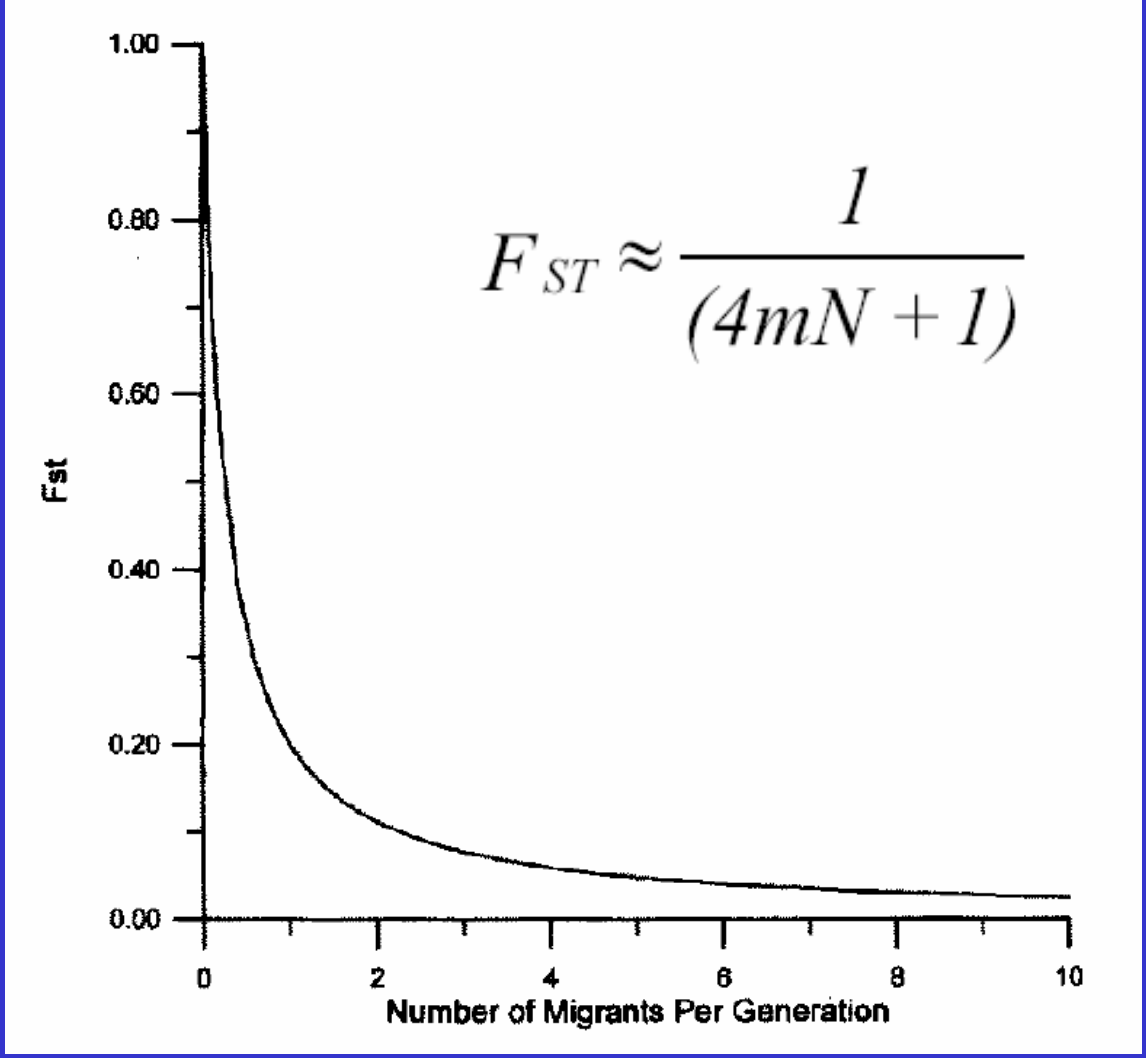
N=100
mN=1
m=0.01



N=100
mN=5
m=0.05



N=100
mN=50
m=0.50

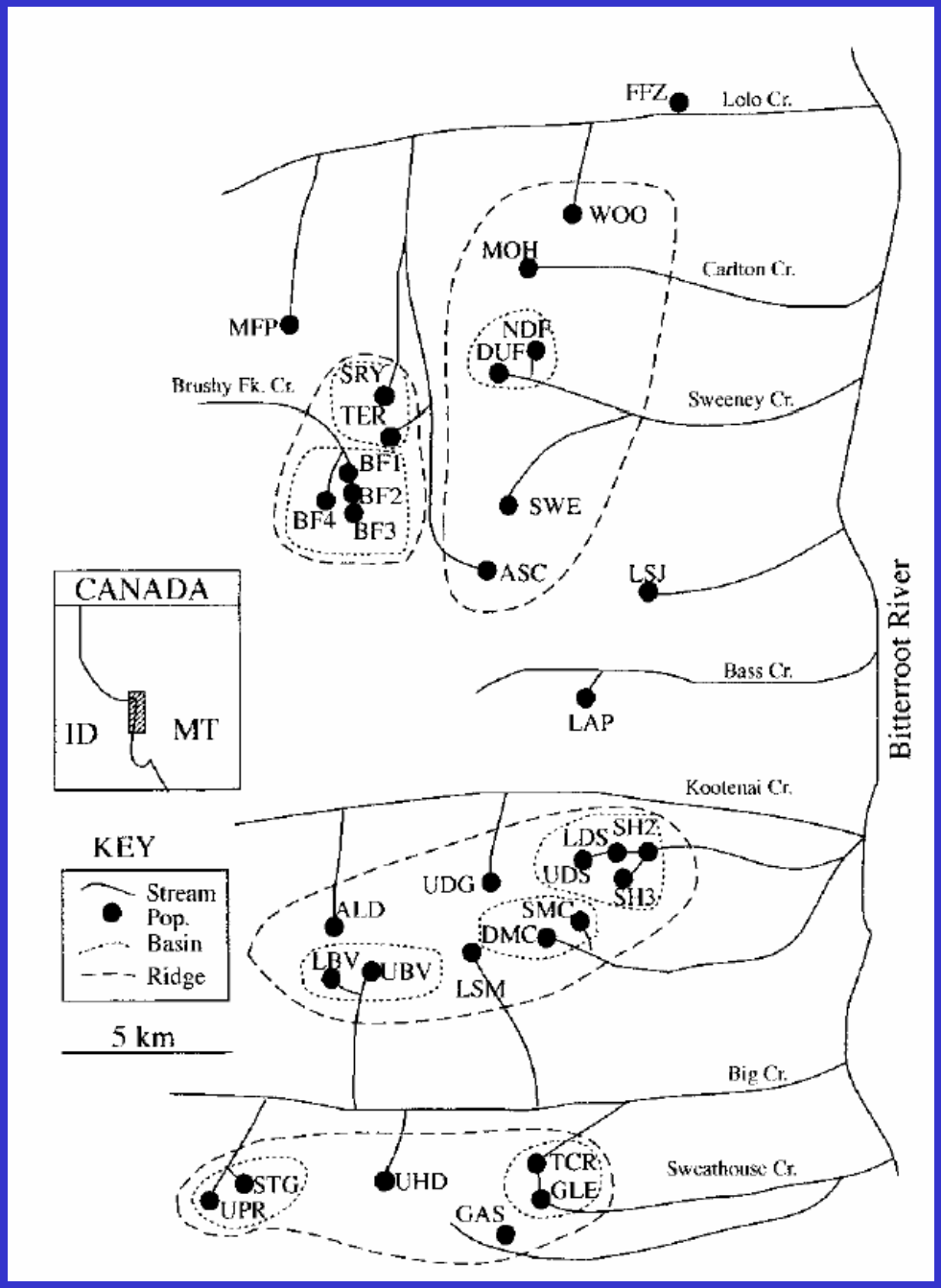


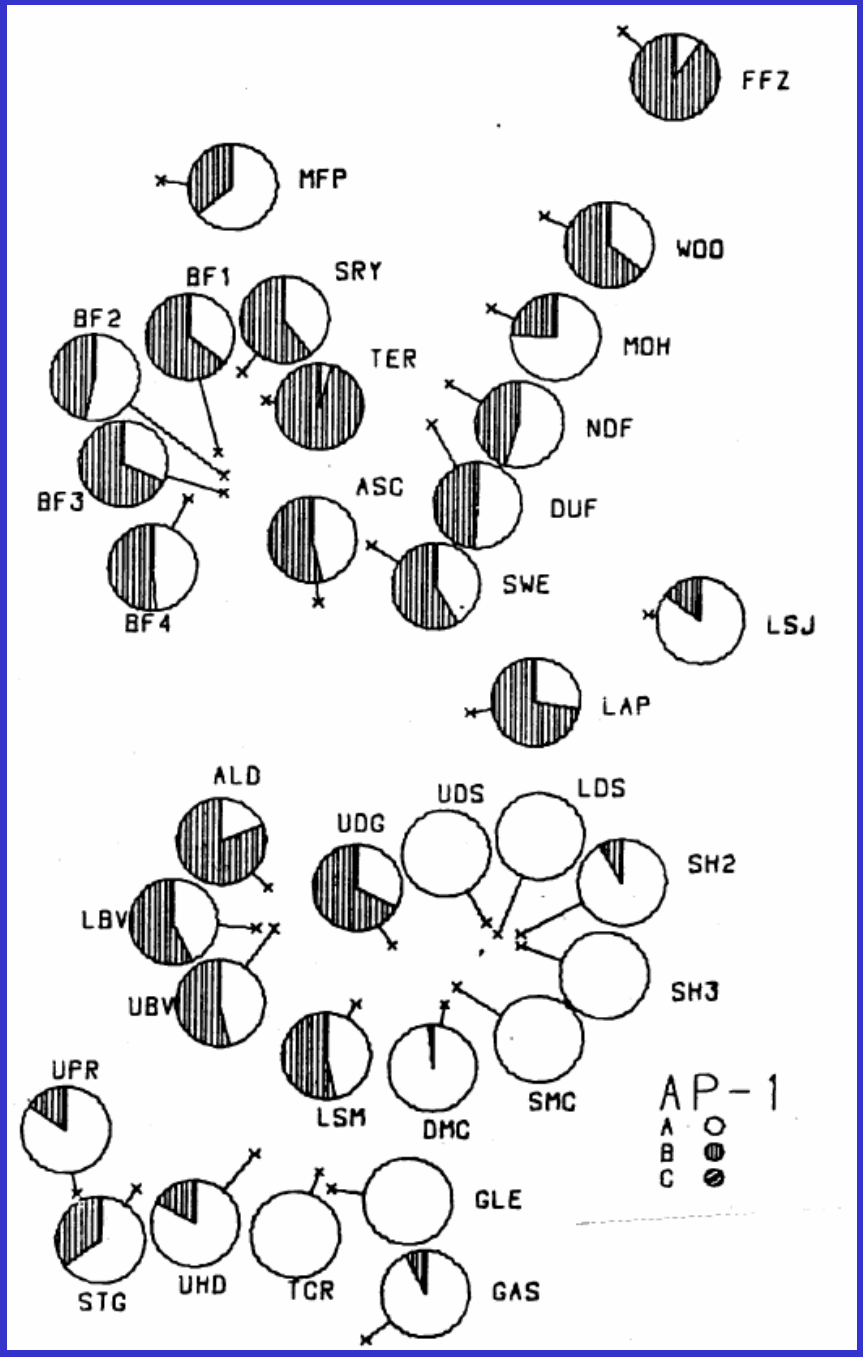
$$F_{ST} \approx \frac{1}{(4mN+1)}$$

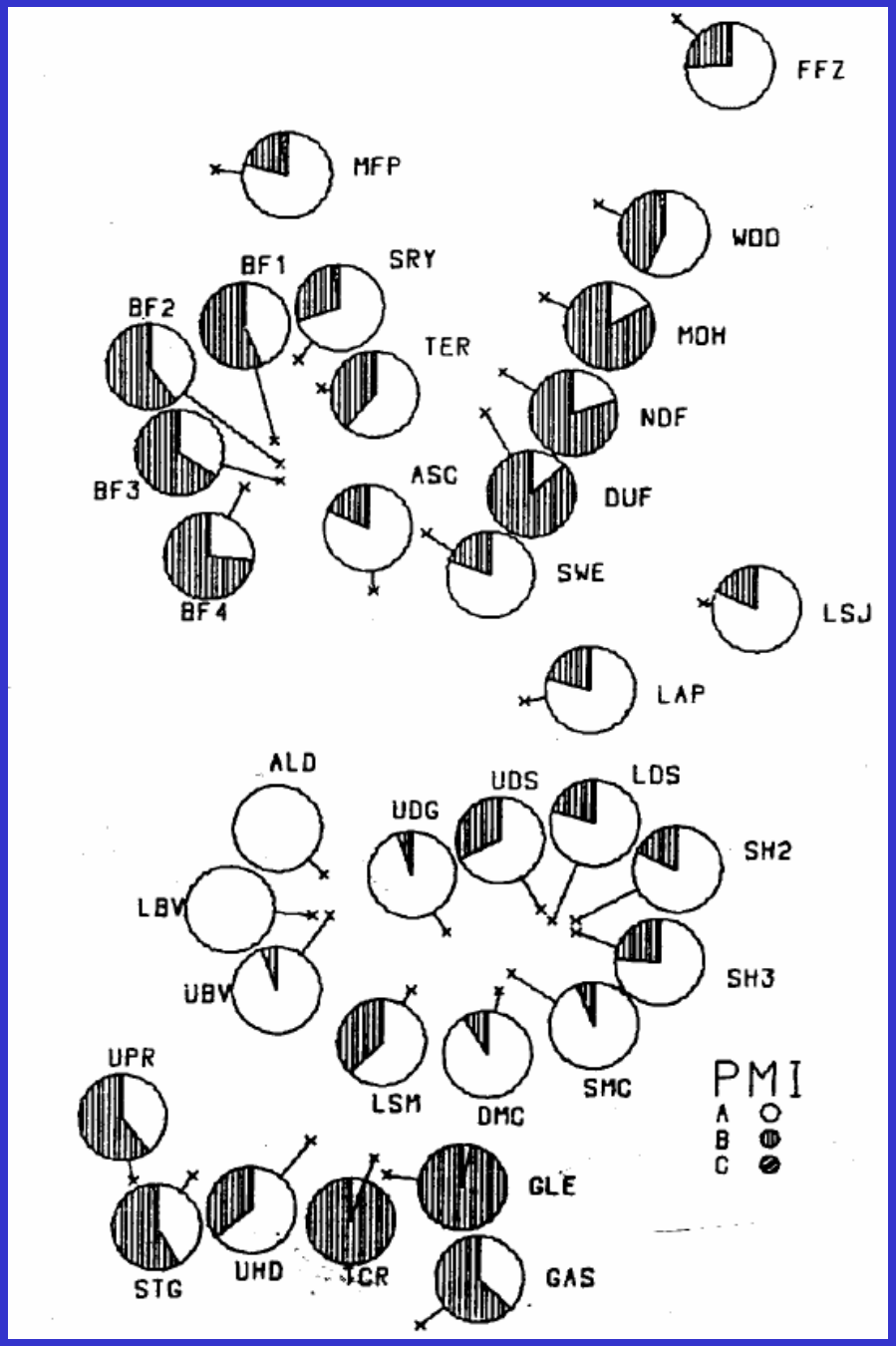
We can use this relationship to estimate gene flow among natural populations.

Long-toed salamander

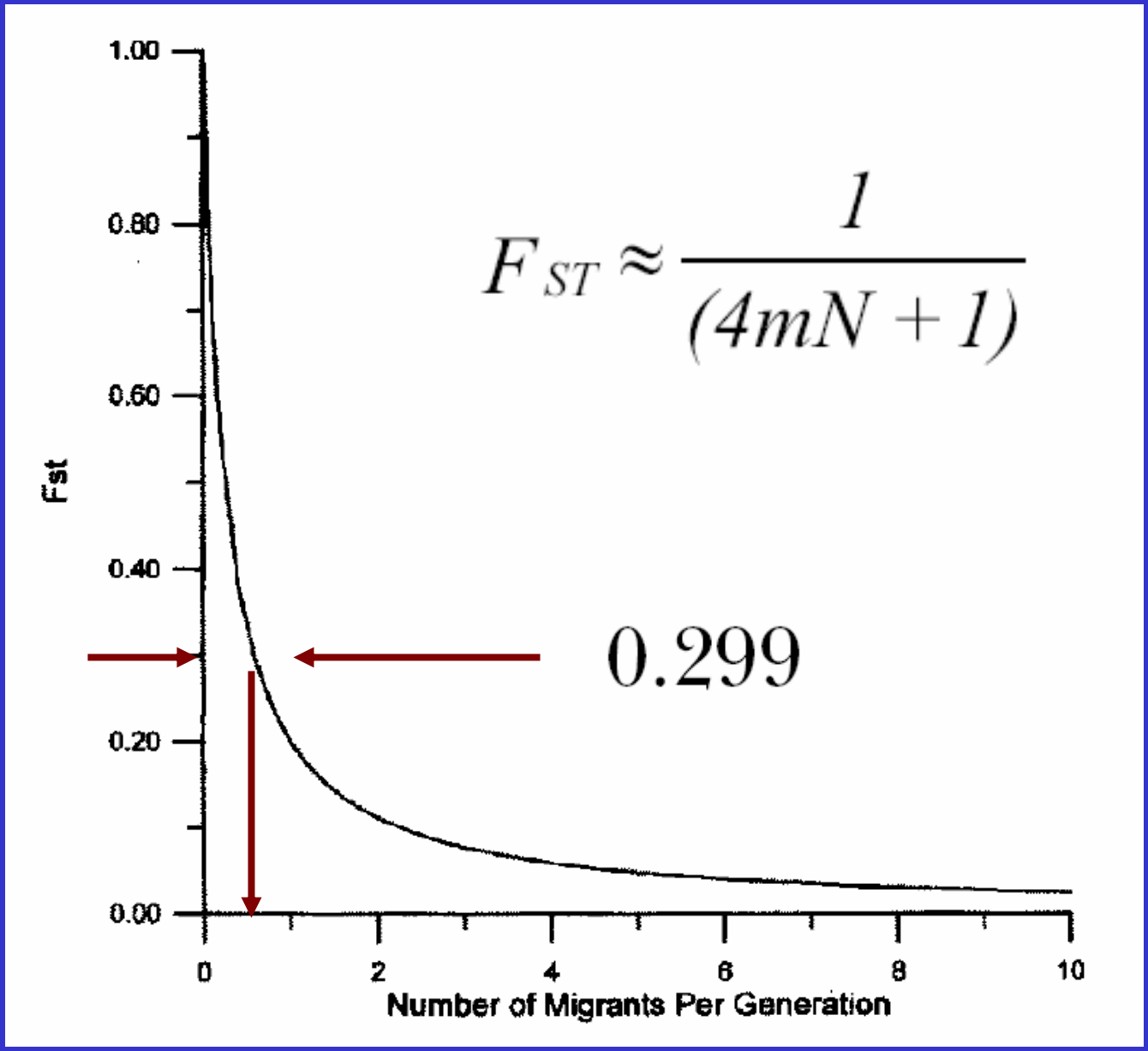
Tallmon et al. Genetic differentiation among long-toed salamander (*Ambystoma macrodactylum*) populations. Copeia 27-35.


FFZ
Lolo Cr.
WOO
MOH
Carlton Cr.
MFP
NDF
DUF
Brushy Fk. Cr.
SRY
TER
Sweeney Cr.
BF1
BF2
BF4
BF3
SWE
ASC
LSJ
CANADA
ID
MT
Bass Cr.
LAP
Bitterroot River
Kootenai Cr.
KEY
Stream
Pop.
Basin
Ridge
5 km
LDS
SH2
UDS
SH3
UDG
ALD
SMC
DMC
LBV
UBV
LSM
Big Cr.
STG
UPR
UHD
TCR
GLE
Sweathouse Cr.
GAS


FFZ
MFP
WOO
BF1
SRY
BF2
TER
MOH
NDF
BF3
ASC
DUF
SWE
BF4
LSJ
LAP
ALD
UDS
LDS
UDG
SH2
LBV
SH3
UBV
UPR
LSM
DMC
SMC
AP-1
A
B
C
GLE
STG
UHD
TCR
GAS



Locus	H_t	G_{st}
AAT	0.493	0.350
PGM-1	0.242	0.210
MPI	0.481	0.340
G6PGH	0.493	0.212
ACP-1	0.477	0.357
PGDH	0.091	0.322
Overall		0.299

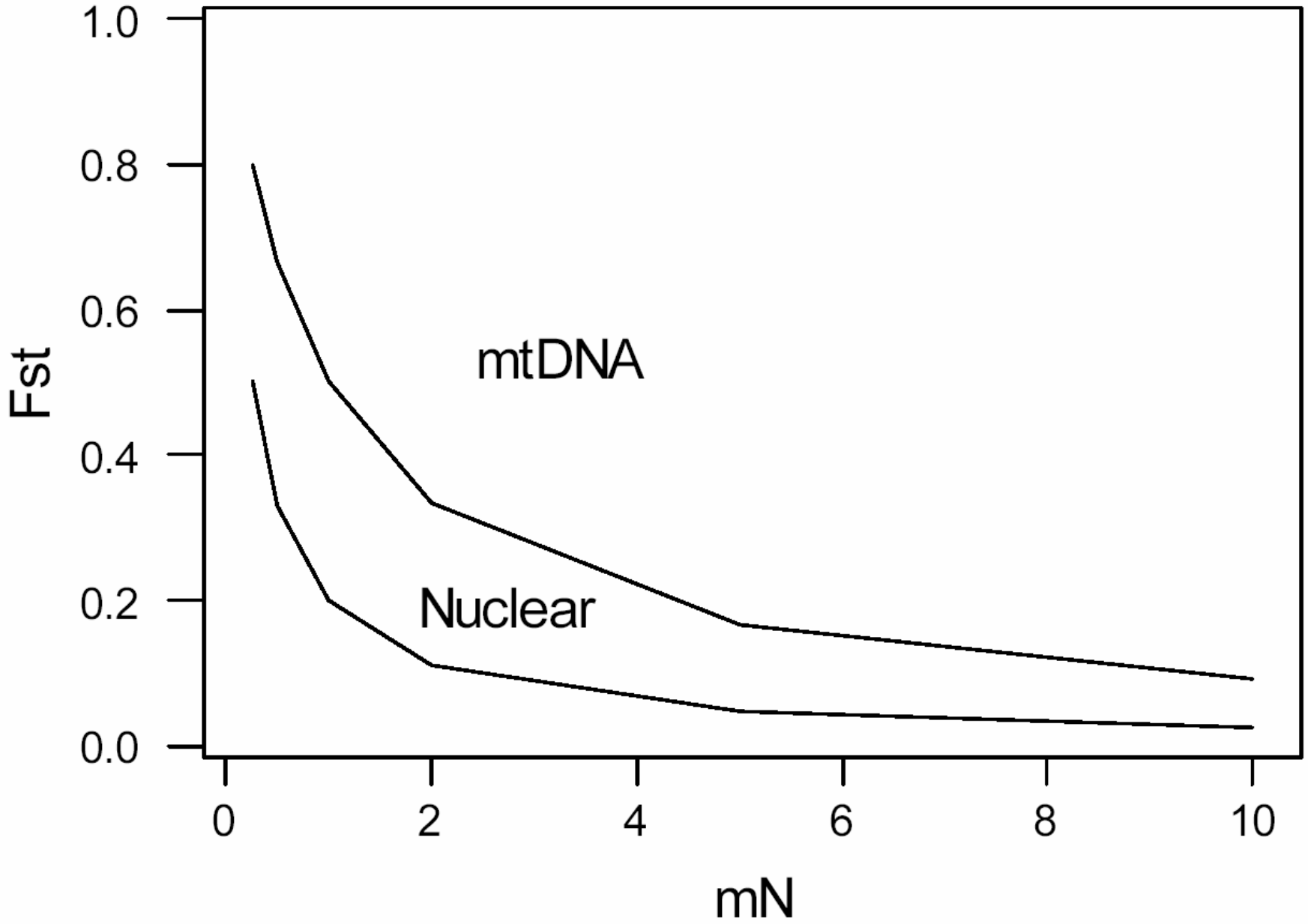

1.00
0.80
0.60
0.40
0.20
0.00
Fst
0
2
4
6
8
10
Number of Migrants Per Generation
$F_{ST} \approx \frac{1}{(4mN+1)}$
0.299

Mitochondrial DNA

The N_e of mtDNA is only one-quarter that of nuclear genes. Therefore,

$$F_{ST} \approx \frac{1}{(mN+1)}$$

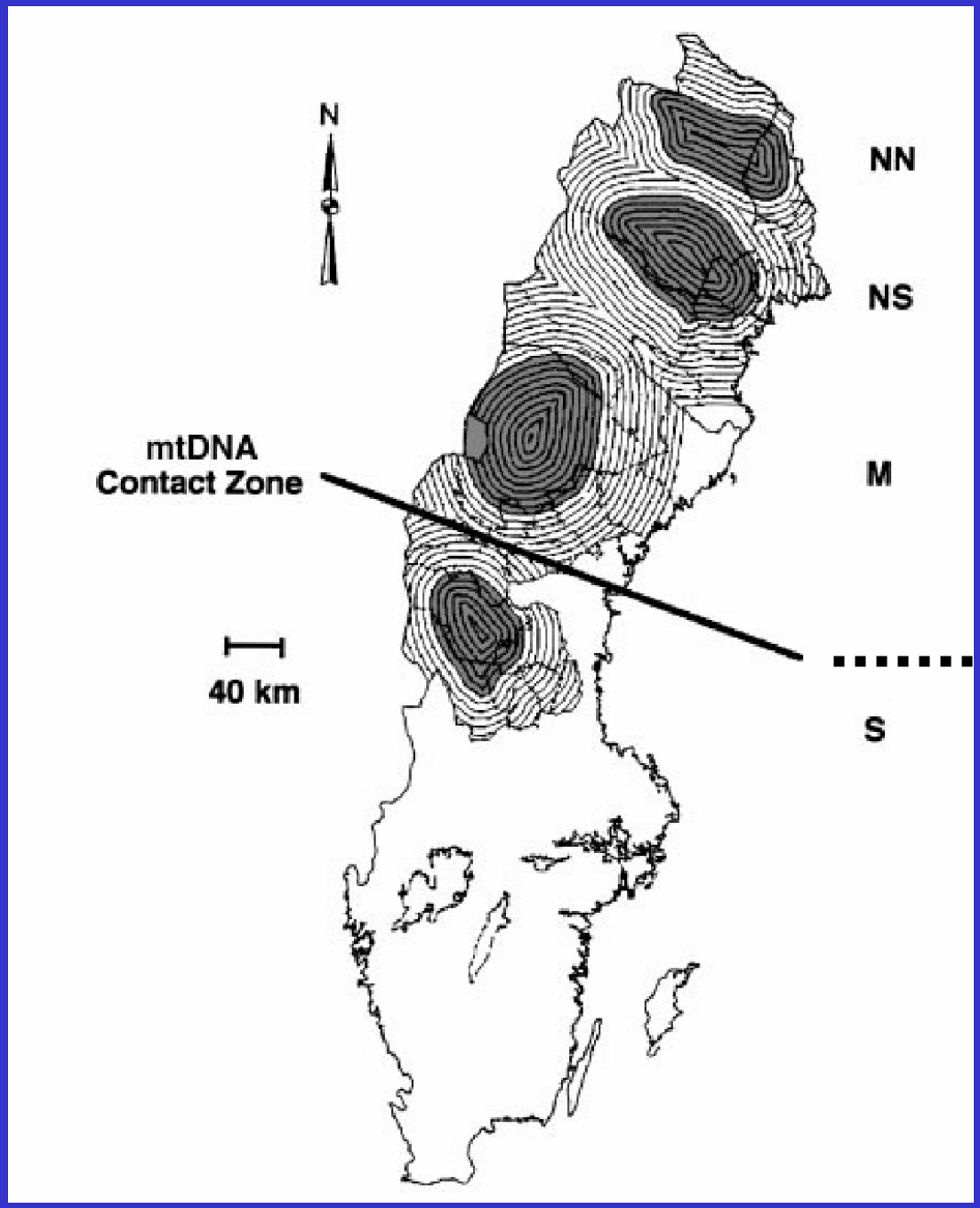
$$F_{ST} \approx \frac{1}{(2mN_F+1)}$$


1.0
0.8
0.6
0.4
0.2
0.0
Fst
mtDNA
Nuclear
0
2
4
6
8
10
mN

mtDNA versus nucDNA

Nuclear DNA microsatellite analysis of genetic diversity and gene flow in the Scandinavian brown bear (*Ursus arctos*)

LISETTE WAITS,*† PIERRE TABERLET,† JON E. SWENSON,‡§ FINN SANDEGREN¶ and ROBERT FRANZÉN**



mtDNA suggested subspecies

$F_{ST} \sim 1.0$

$mN = 0$

What happens if we combine natural selection along with genetic drift and gene flow?

Local adaptation

	A_1A_1	A_1A_2	A_2A_2
Demes 1-10	1	$1-t/2$	$1-t$
Demes 11-20	$1-t$	$1-t/2$	1

Local adaptation

	mN						
	0.5	1	2	5	10	25	N
Expected	0.3333	0.2000	0.1111	0.0476	0.0244	0.0099	
t=0.00	0.3070	0.2043	0.1245	0.0418	0.0198	-	25
	0.3350	0.1826	0.1077	0.0484	0.0264	0.0120	50
	0.3216	0.1884	0.1061	0.0437	0.0251	0.0095	100
t=0.01	0.3343	0.1703	0.1070	0.0556	0.0220	-	25
	0.2979	0.1192	0.1000	0.0381	0.0256	0.0099	50
	0.2997	0.1850	0.1146	0.0354	0.0229	0.0105	100
t=0.05	0.3560	0.1857	0.1204	0.0497	0.0217	-	25
	0.4618	0.2679	0.1489	0.0550	0.0265	0.0113	50
	0.5950	0.4230	0.1982	0.0632	0.0207	0.0118	100
t=0.10	0.4700	0.2446	0.1632	0.0473	0.0289	-	25
	0.6242	0.3653	0.2611	0.0771	0.0356	0.0128	50
	0.8054	0.6575	0.4432	0.1589	0.0632	0.0193	100

Natural selection + drift + gene flow

1. Divergence no longer a function of only mN.

2. Large differences among loci can be produced by even fairly weak natural selection.

3. The larger the local N_e, the more effective natural selection is.

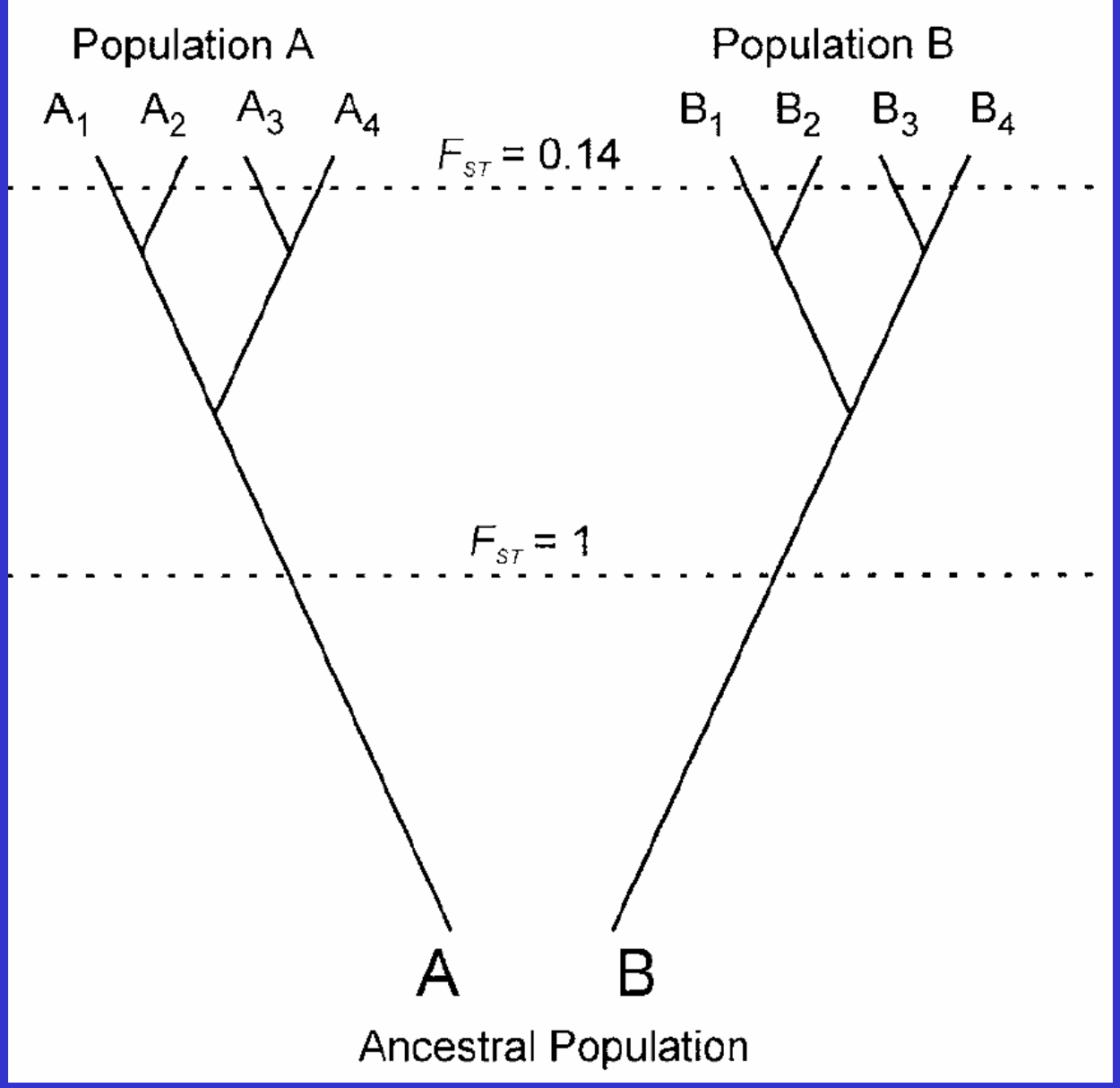
Population “Genomics”

Estimate divergence (F_{ST}) at many loci. Outlier loci are “candidates” for natural selection.

Beyond F_{ST}: Other measures of population divergence

F_{ST} is not a good measure if within population variation (H_S) is high, as with microsatellites.

$$F_{ST} = 1 - \frac{H_S}{H_T}$$

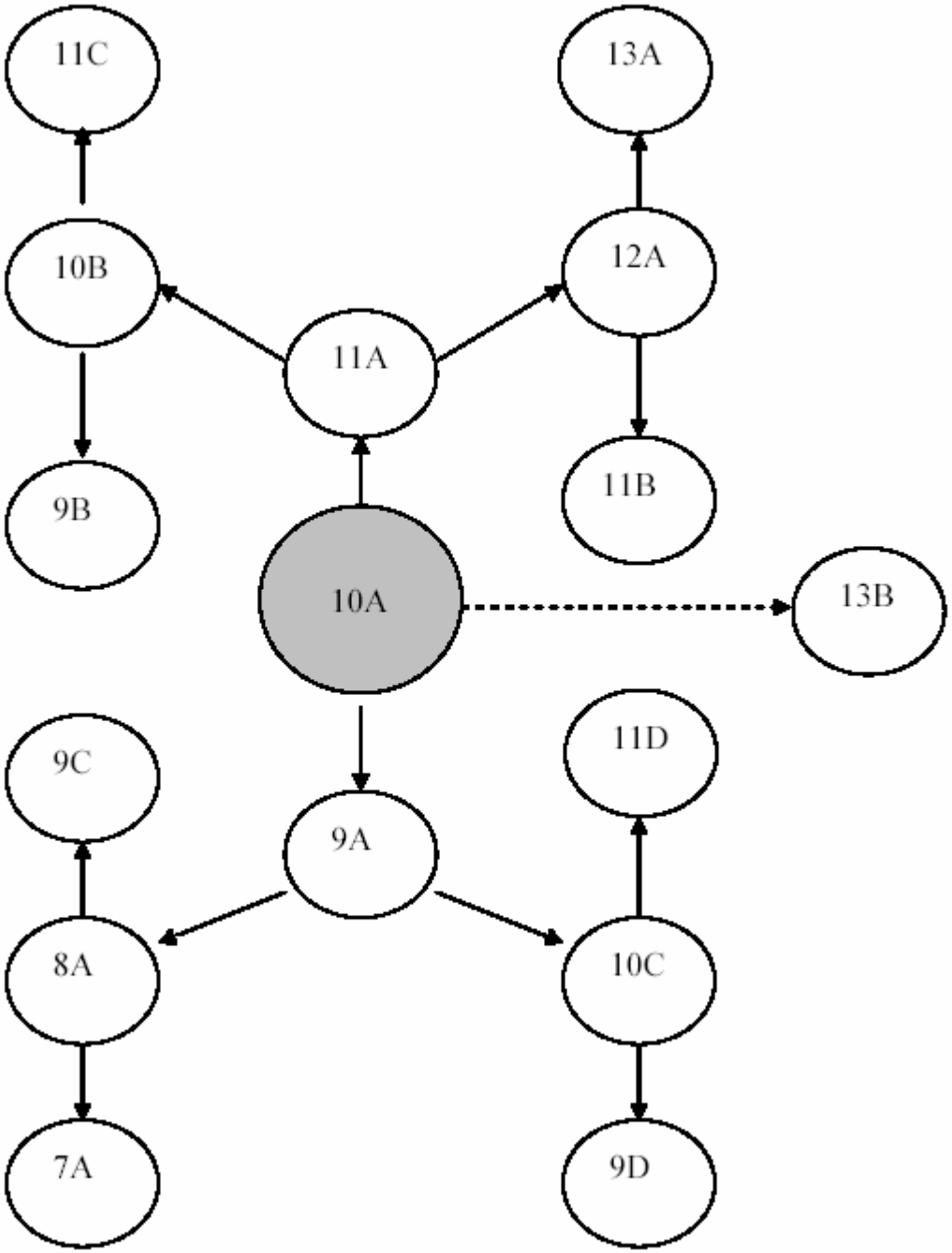


Allele genealogy

(coalescence)

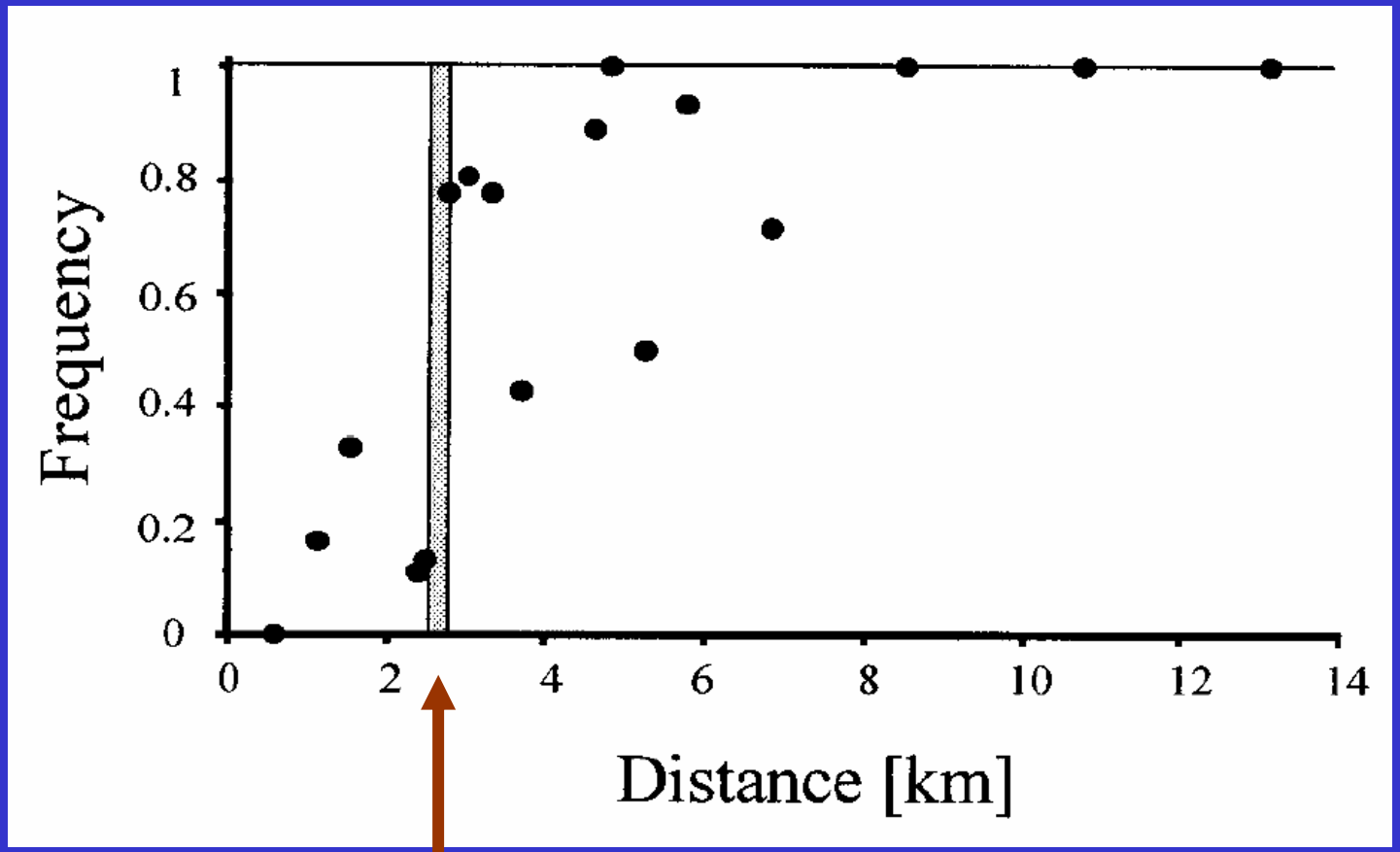
R_{ST} is an analogue of F_{ST} that takes the relative size of alleles (i.e., allelic state) into consideration.

Microsatellite mutation network



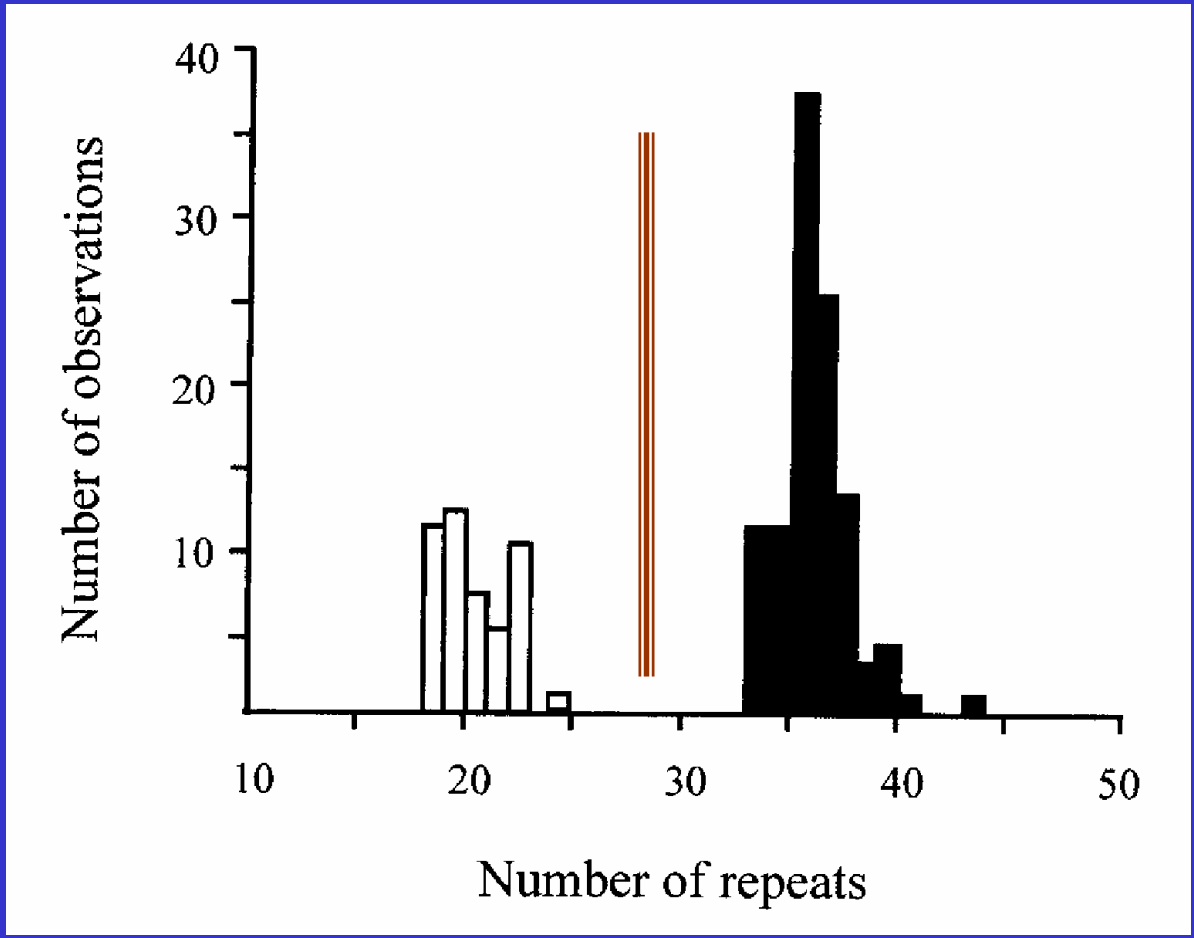
Common shrew in western France

mtDNA



Hybrid zone

Y-linked microsatellite locus



$H_S = 0.71$

$F_{ST} = 0.19$; $R_{ST} = 0.98$