HAL open science

Distinguishing shared ancestral polymorphism from recent introgression in genes with recombination

Miguel Navascués, Frantz Depaulis

To cite this version:

Miguel Navascués, Frantz Depaulis. Distinguishing shared ancestral polymorphism from recent introgression in genes with recombination. Réunion 2008 du GDR 1928 Génomique des populations et génomique évolutive, GDR 1928 Génomique des populations et génomique évolutive., Oct 2008, Sète, France. hal-02814597

HAL Id: hal-02814597

https://hal.inrae.fr/hal-02814597v1

Submitted on 10 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

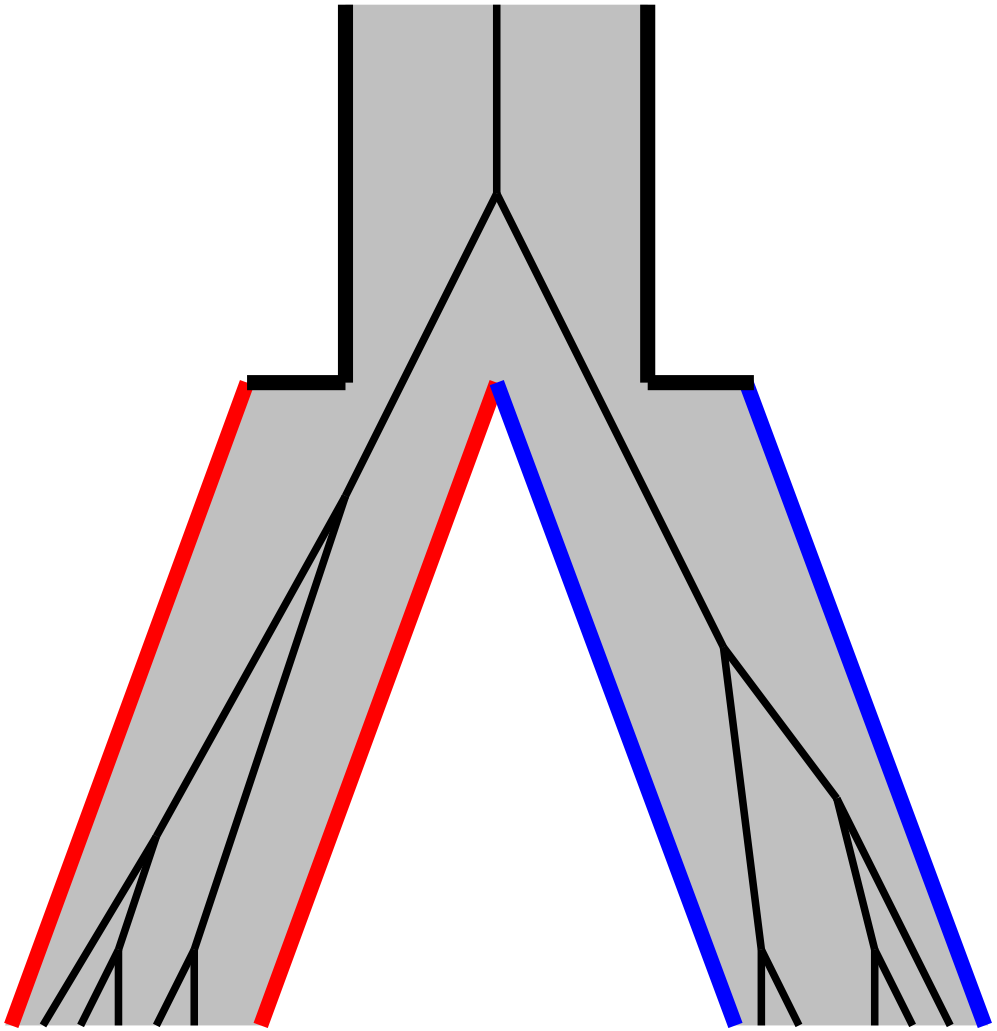
Distinguishing shared ancestral polymorphism from recent introgression

Miguel Navascués & Frantz Depaulis
CNRS UMR 7625 Écologie et Évolution (ENS/UPMC)

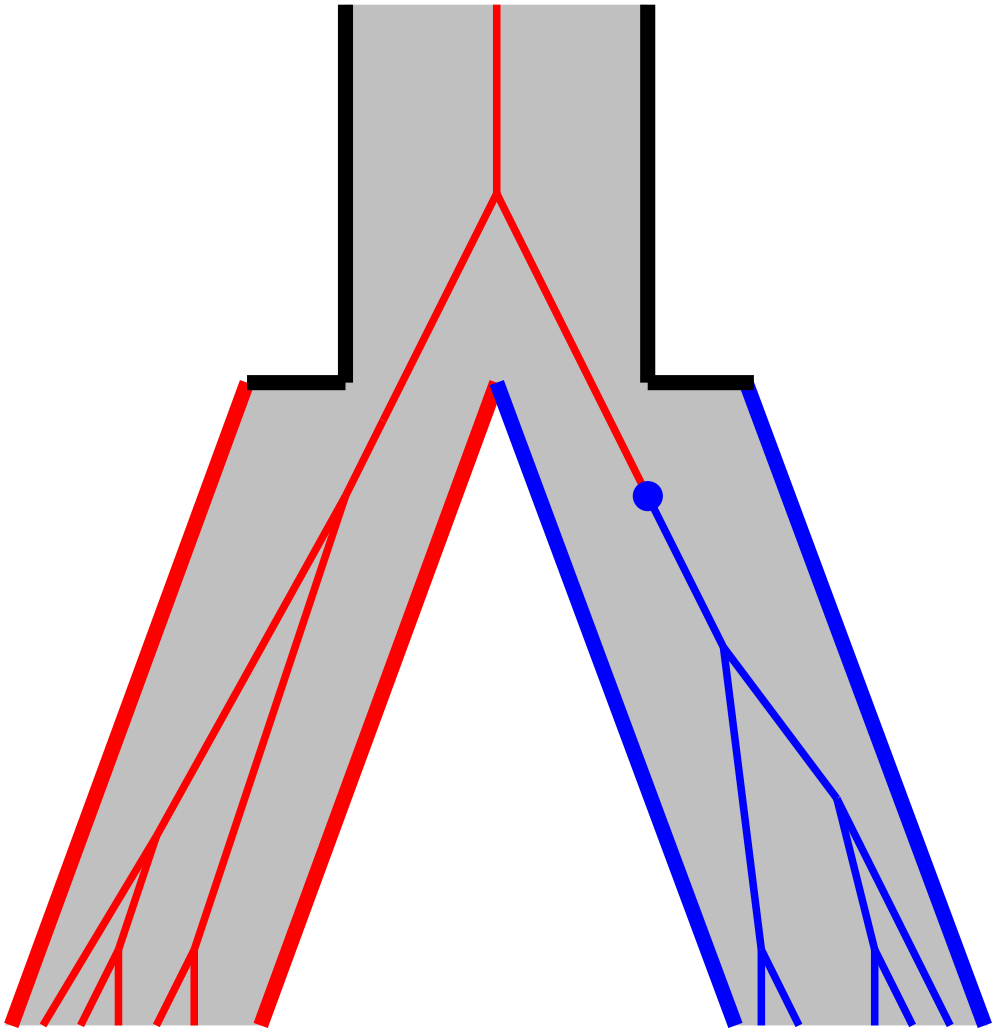
Biodiversité: TRANSBIODIV

Réunion GDR 1928 Génomique des populations et génomique évolutive – Sète, 16 October 2008

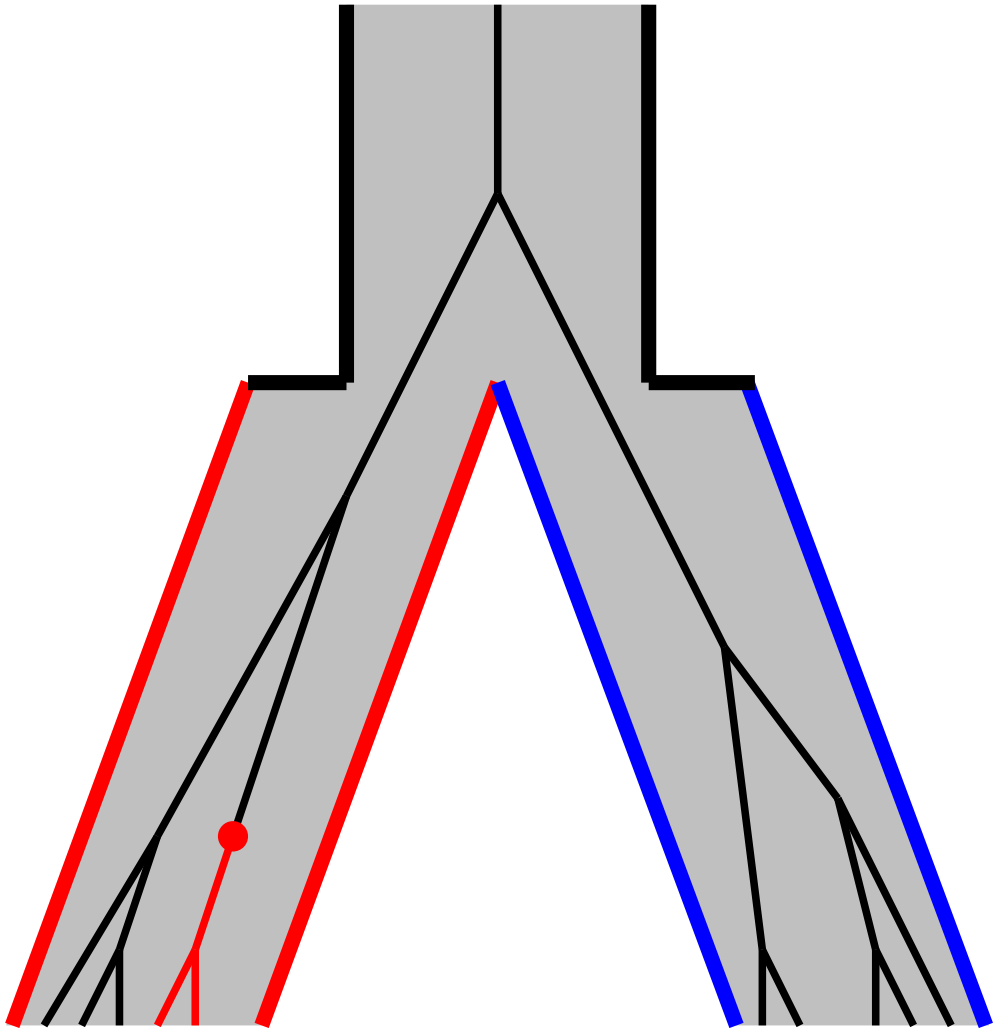
Polymorphisms between two species



Fixed polymorphism

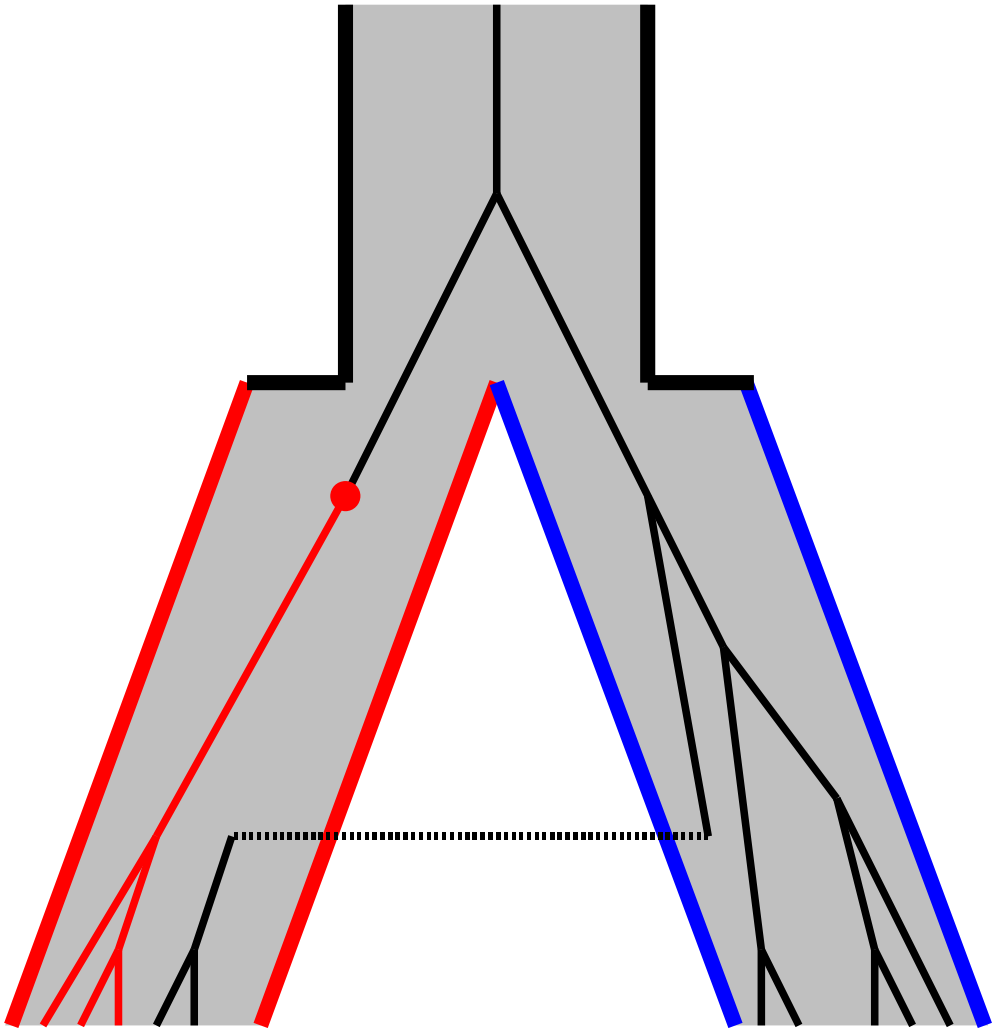


Exclusive polymorphism



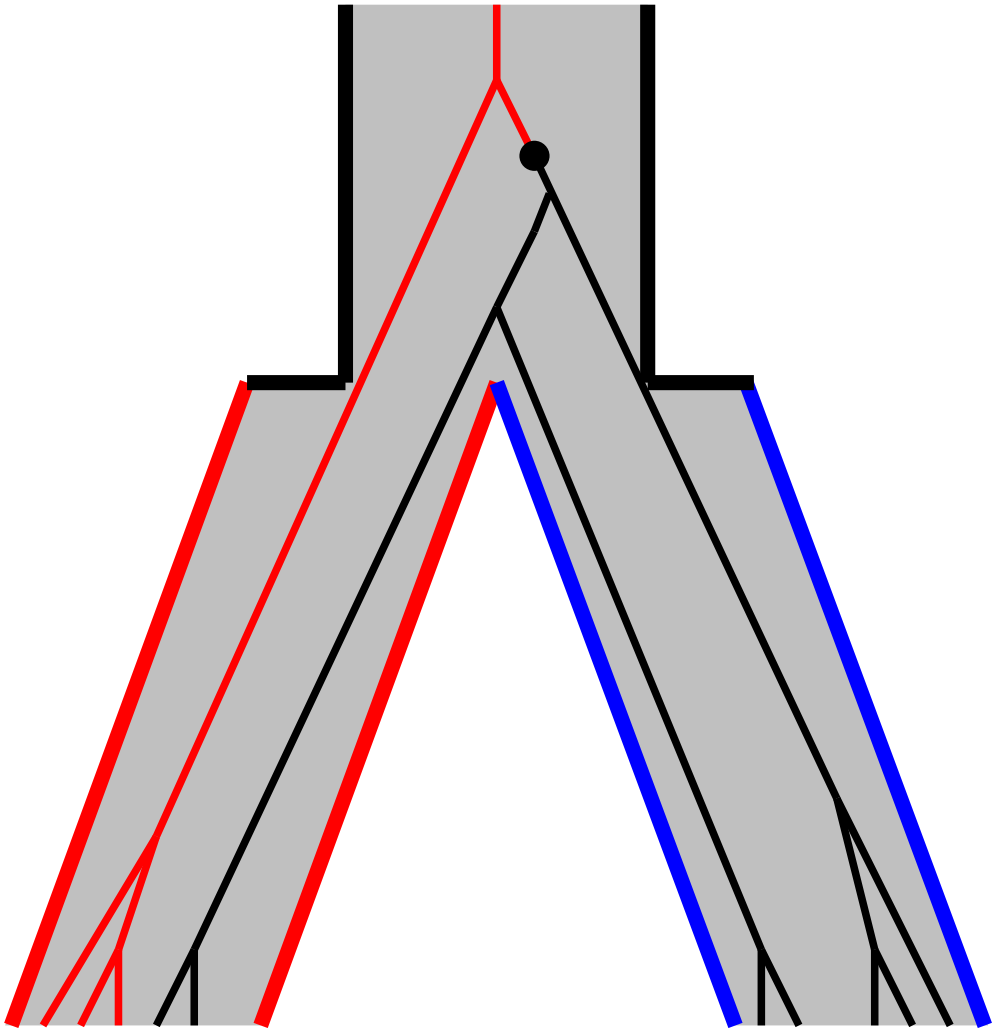
Exclusive polymorphism

Introgression



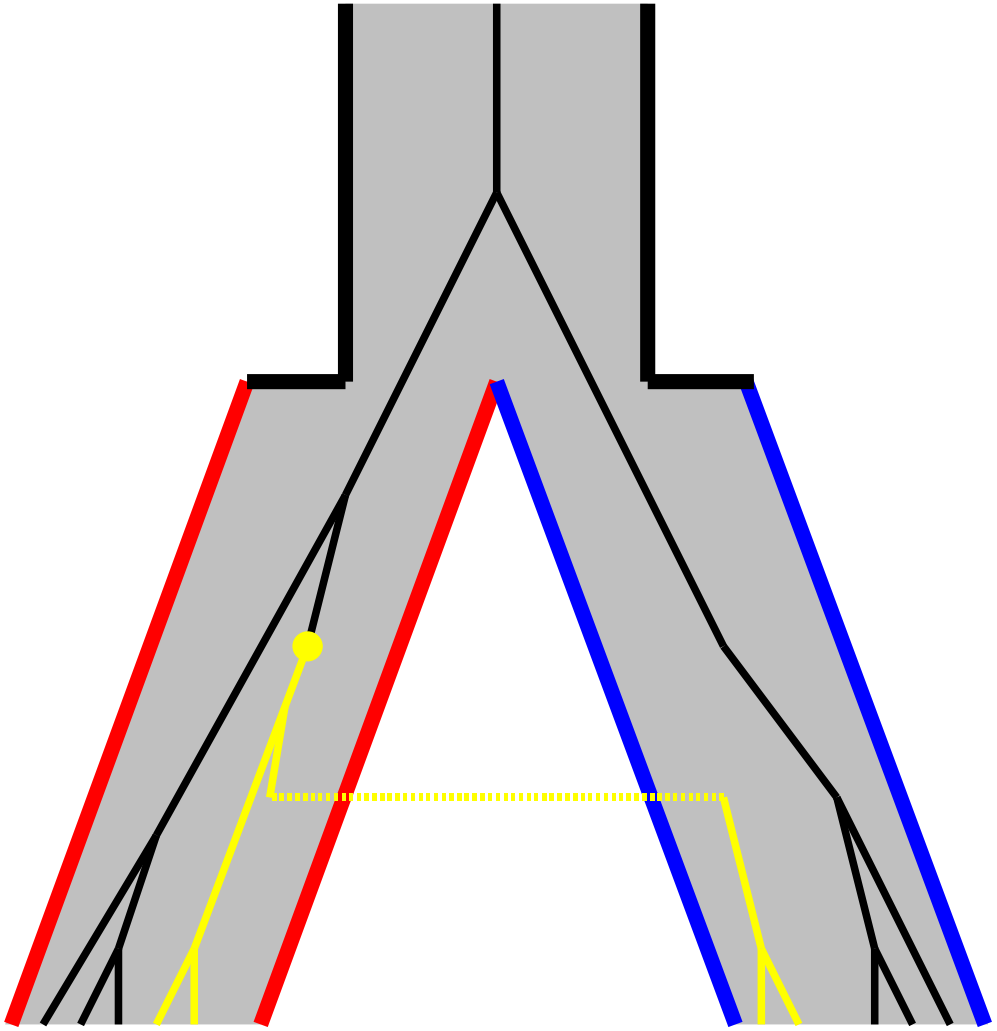
Exclusive polymorphism

Ancestral



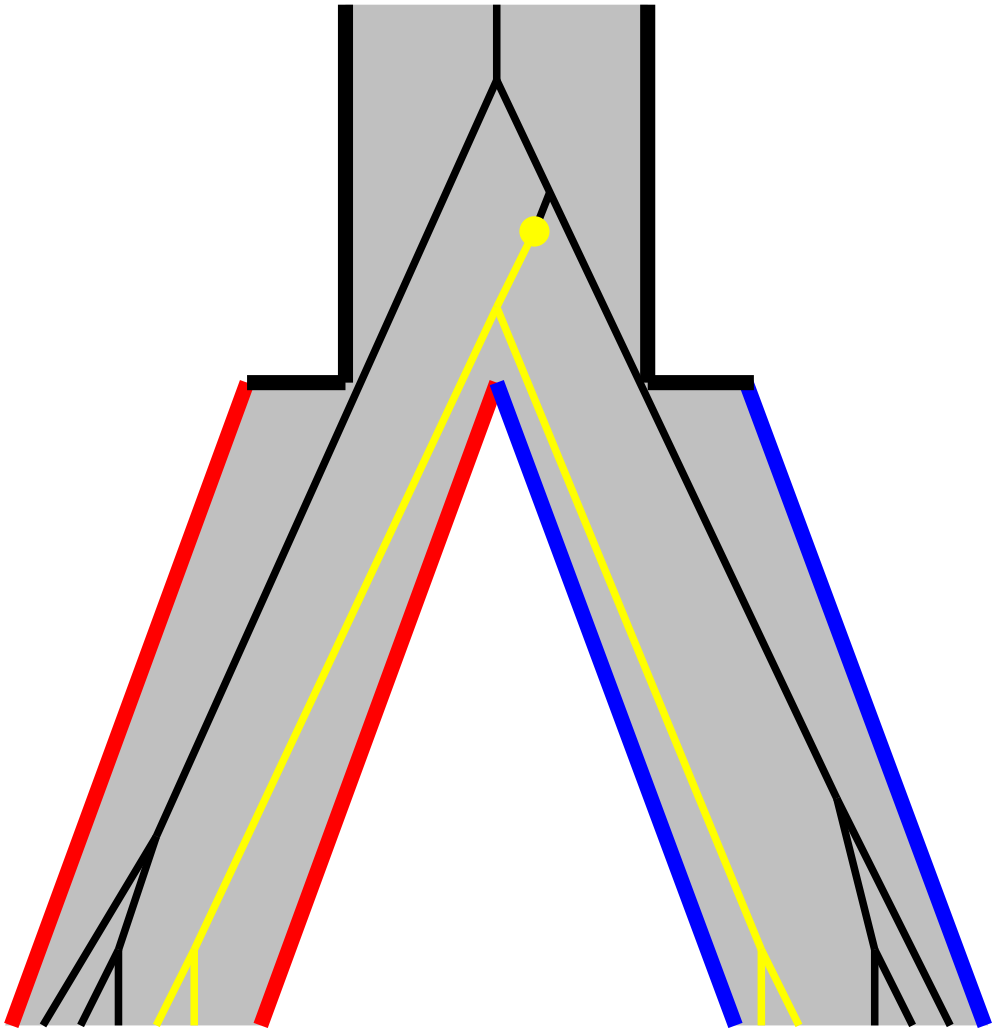
Shared polymorphism

Introgression



Shared polymorphism

Ancestral

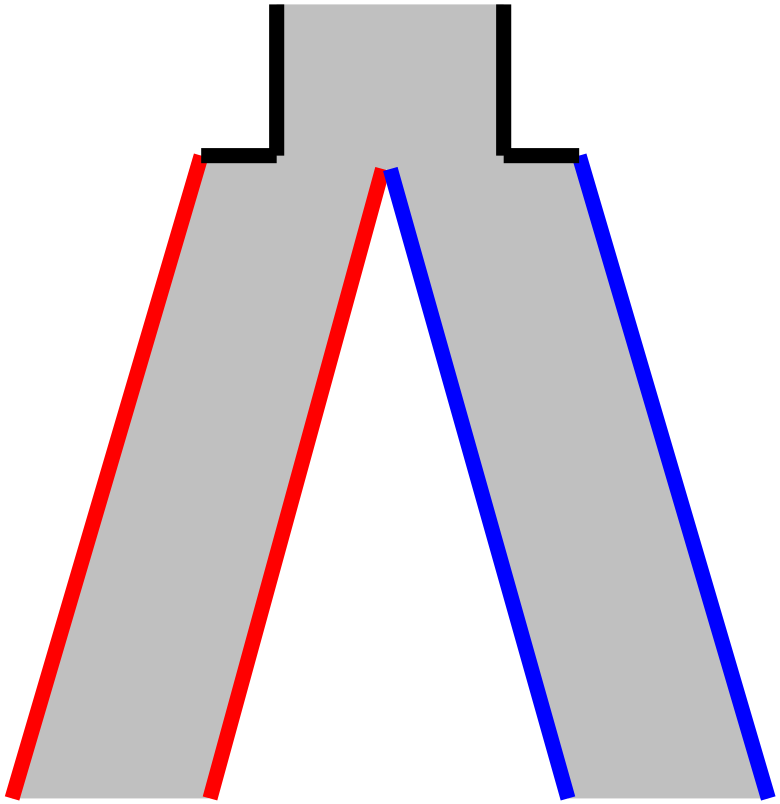


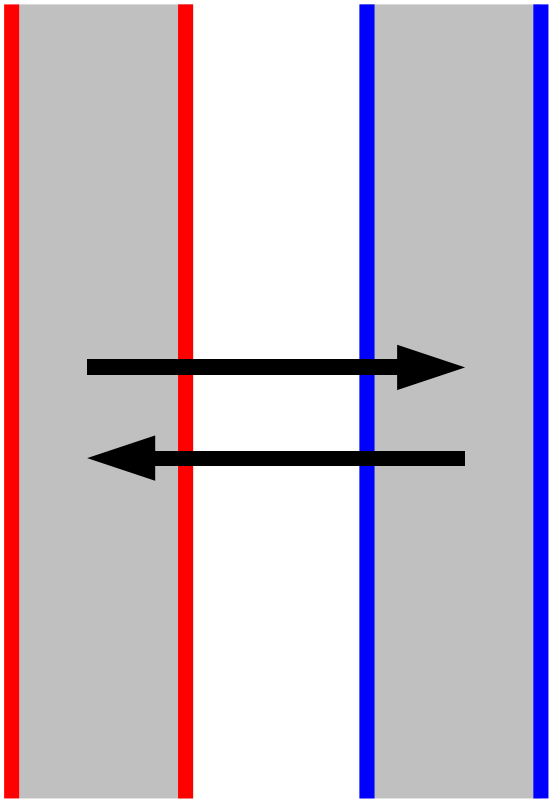
Interpretation of F_{ST}

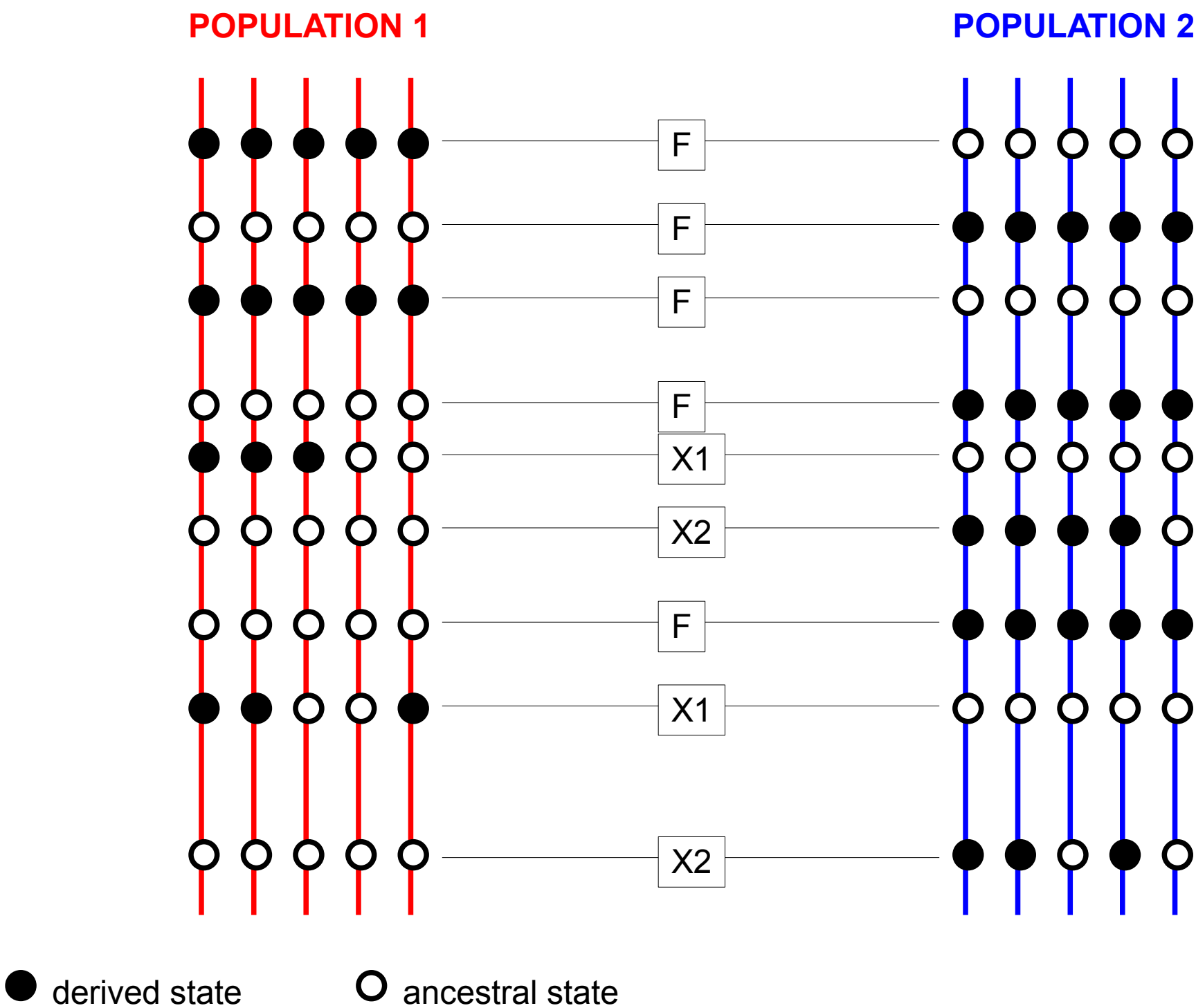
Isolation vs. Migration

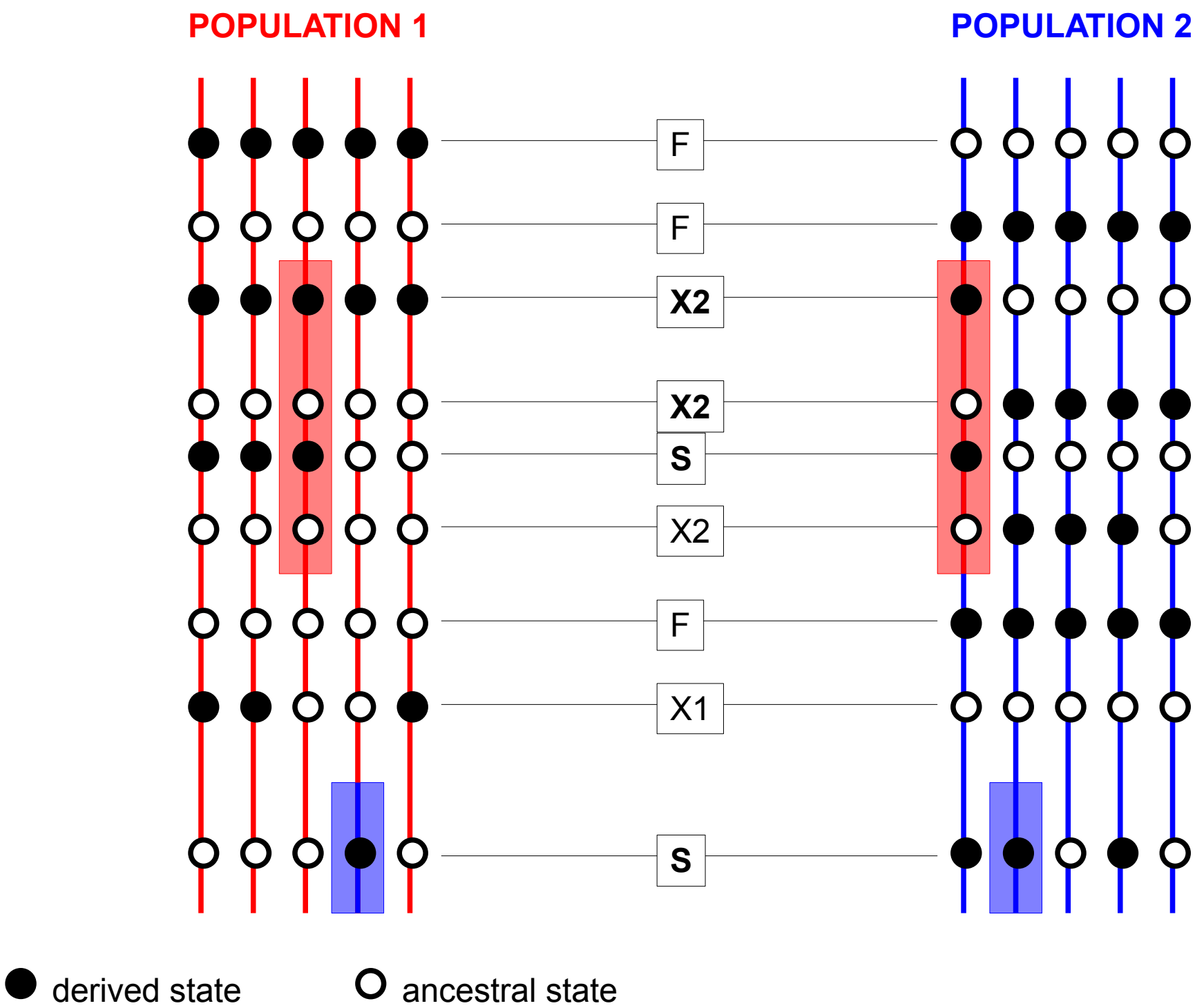
$$F_{ST} \approx 1 - e^{-t/2N}$$

$$F_{ST} \approx \frac{1}{1+4Nm}$$






POPULATION 1
POPULATION 2
F
F
F
F
X1
X2
F
X1
X2
derived state
ancestral state


POPULATION 1
POPULATION 2
F
F
X2
X2
S
X2
F
X1
S
derived state
ancestral state

Statistics: number of runs

$n^{*}=8$, $n^{-}=7$

```
**--*-****-*---    NR=8    random
1  2  345    678

********-------    NR=2    clustered
1         2
```

Statistics: length of runs

$n^{*}=7$, $n^{-}=8$

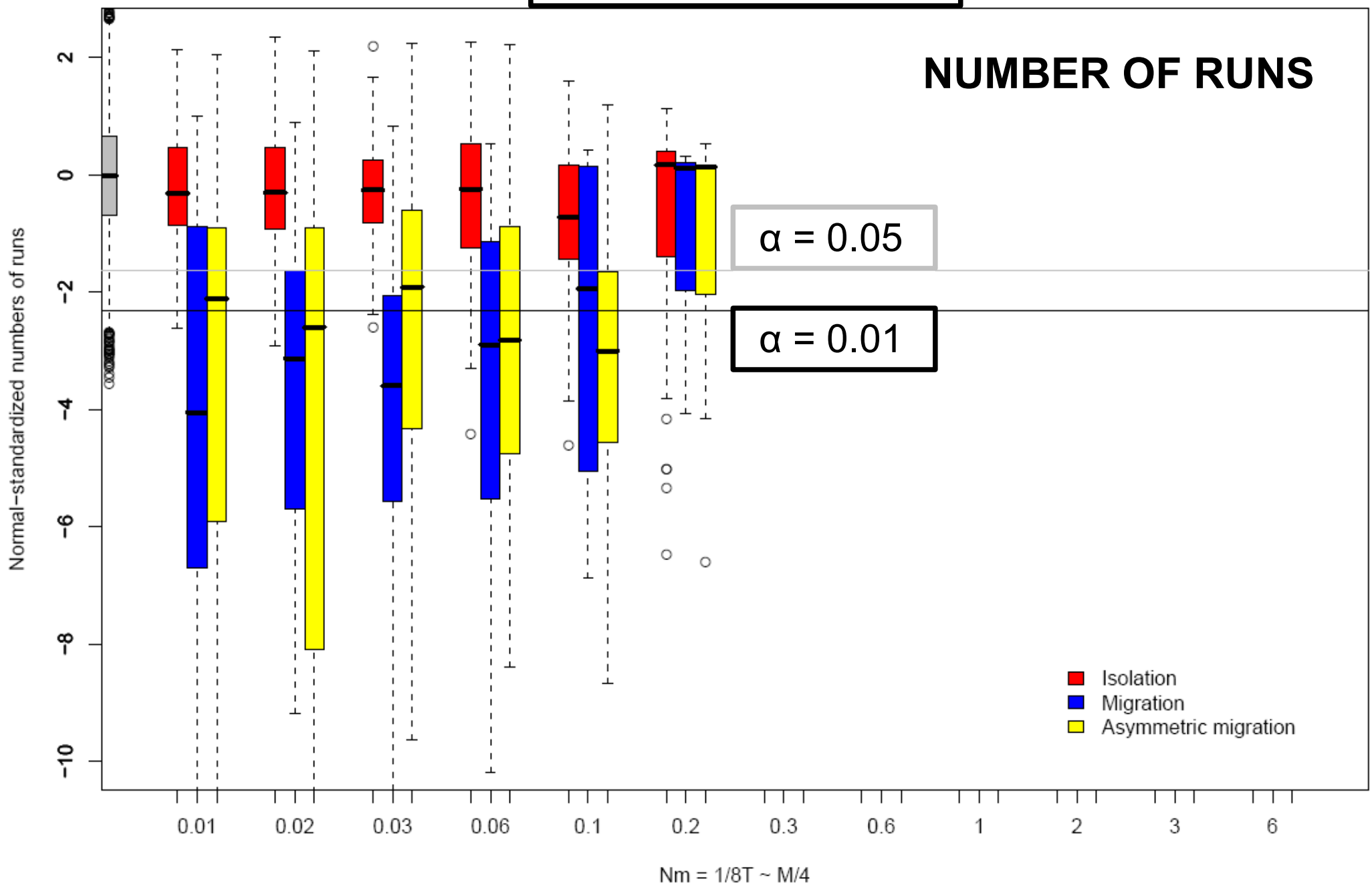
"−−−−−−−−−" + 7 random cuts (*) =

−*−**−−−*−−**−* Lengths (−): 1, 1, 0, 3, 2, 0, 1, 0

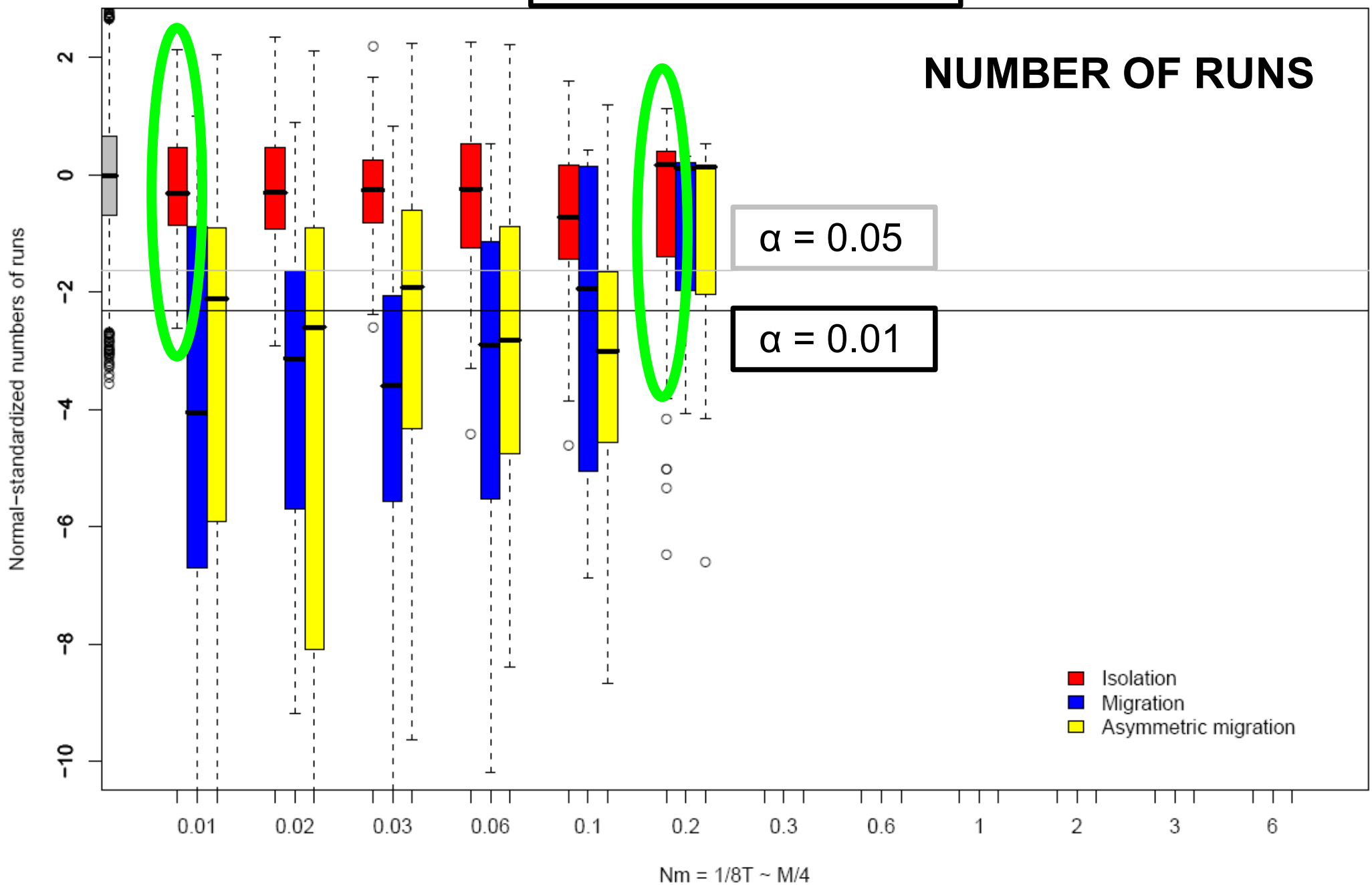
"−−−−−−−−−" + 7 clustered cuts (*) =

−−−−−*******−−− Lengths (−): 5, 0, 0, 0, 0, 0, 0, 3

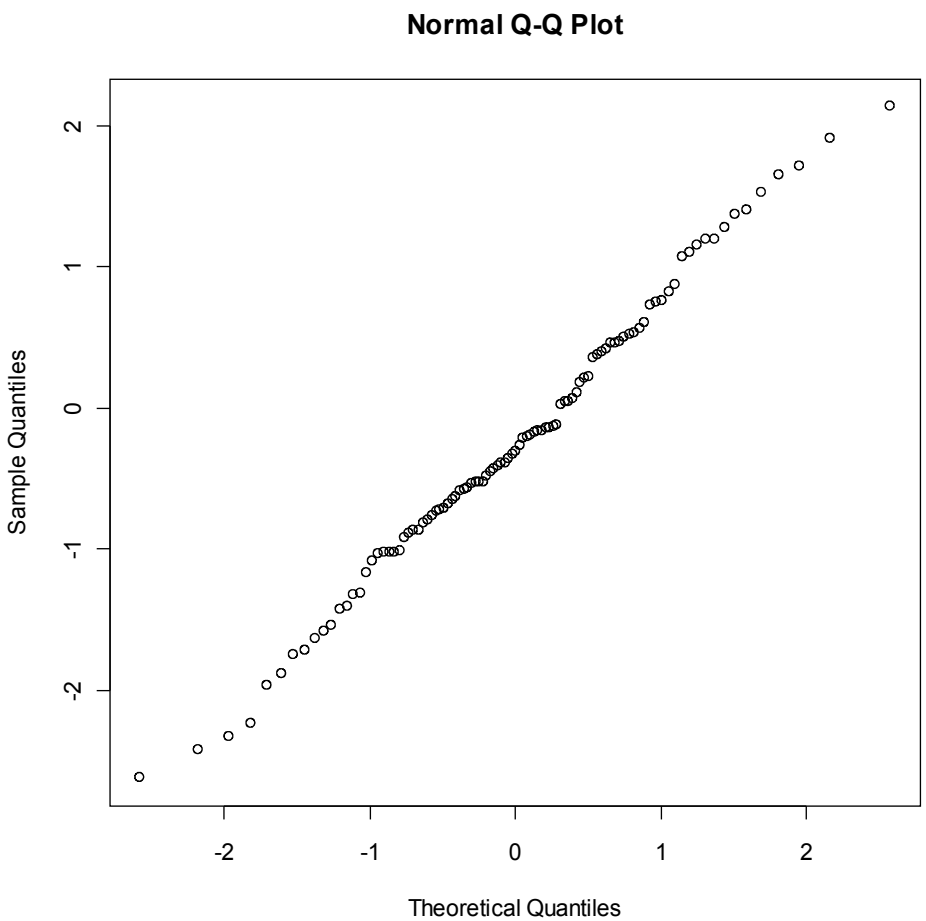
F(*) vs. S+X(−)

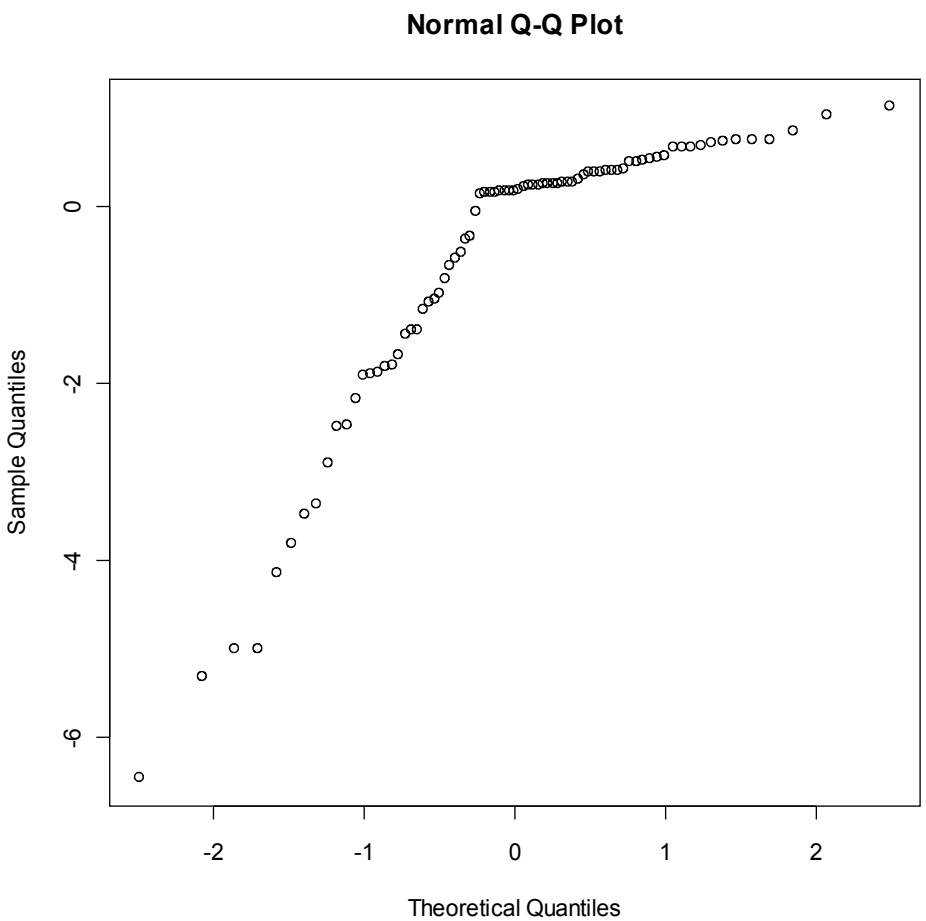

NUMBER OF RUNS
α = 0.05
α = 0.01
Isolation
Migration
Asymmetric migration
Normal−standardized numbers of runs
2
0
−2
−4
−6
−8
−10
0.01
0.02
0.03
0.06
0.1
0.2
0.3
0.6
1
2
3
6
Nm = 1/8T ~ M/4

F(*) vs. S+X(–)


NUMBER OF RUNS
α = 0.05
α = 0.01
Isolation
Migration
Asymmetric migration
Normal−standardized numbers of runs
2
0
−2
−4
−6
−8
−10
0.01
0.02
0.03
0.06
0.1
0.2
0.3
0.6
1
2
3
6
Nm = 1/8T ~ M/4

Isolation as null hypothesis



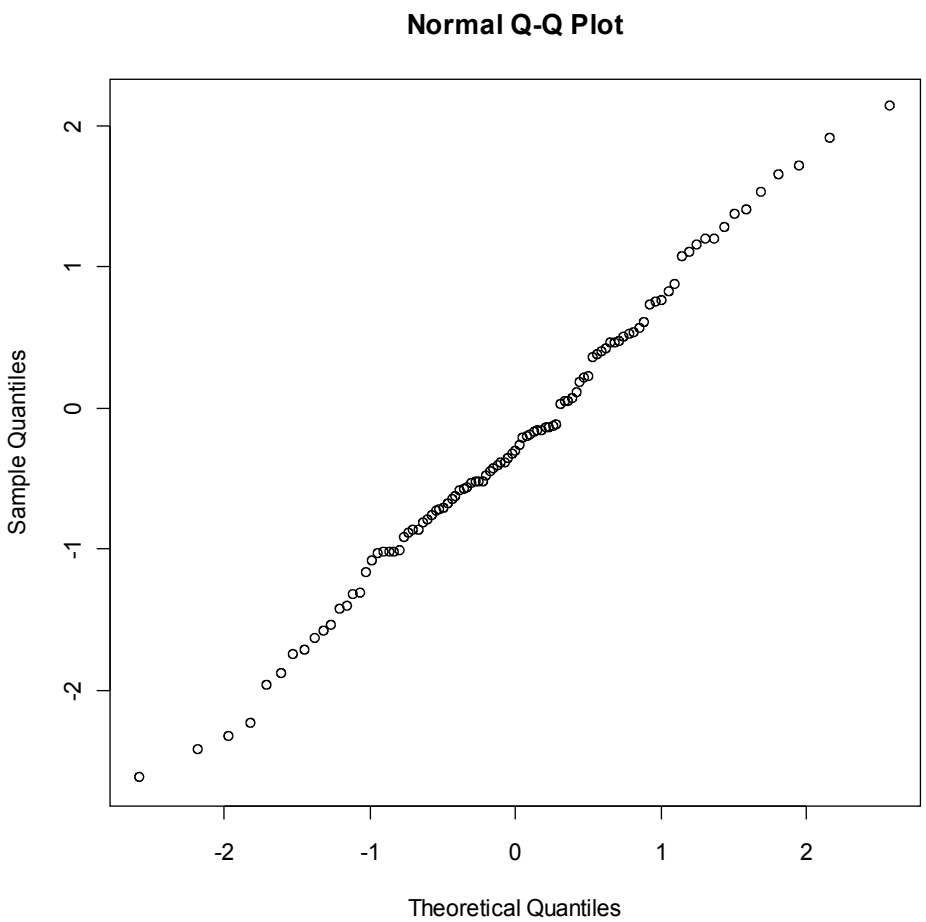


Isolation: T=12.5

Isolation: T=0.625

(T: time scaled to population size)

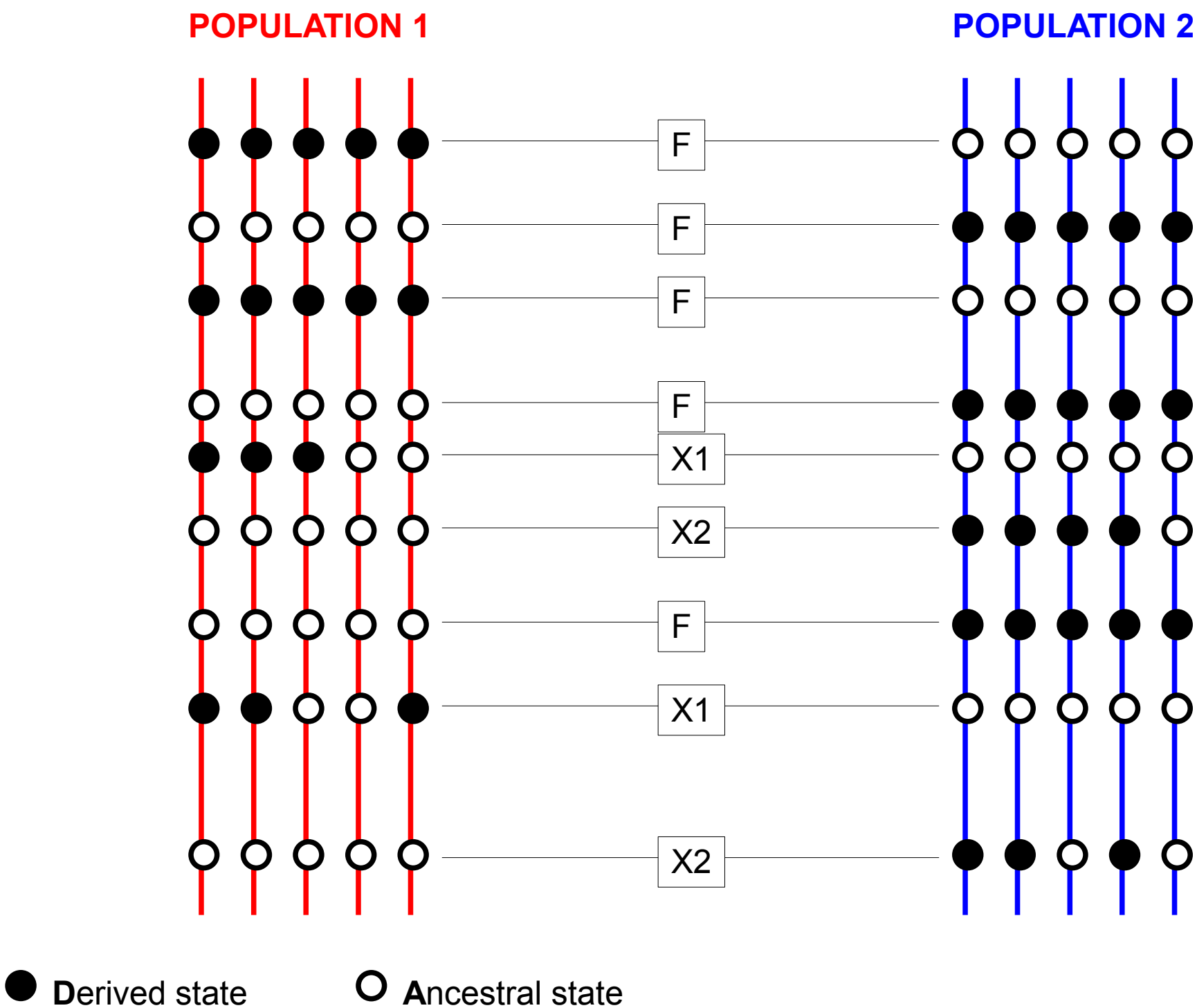
Test with several loci

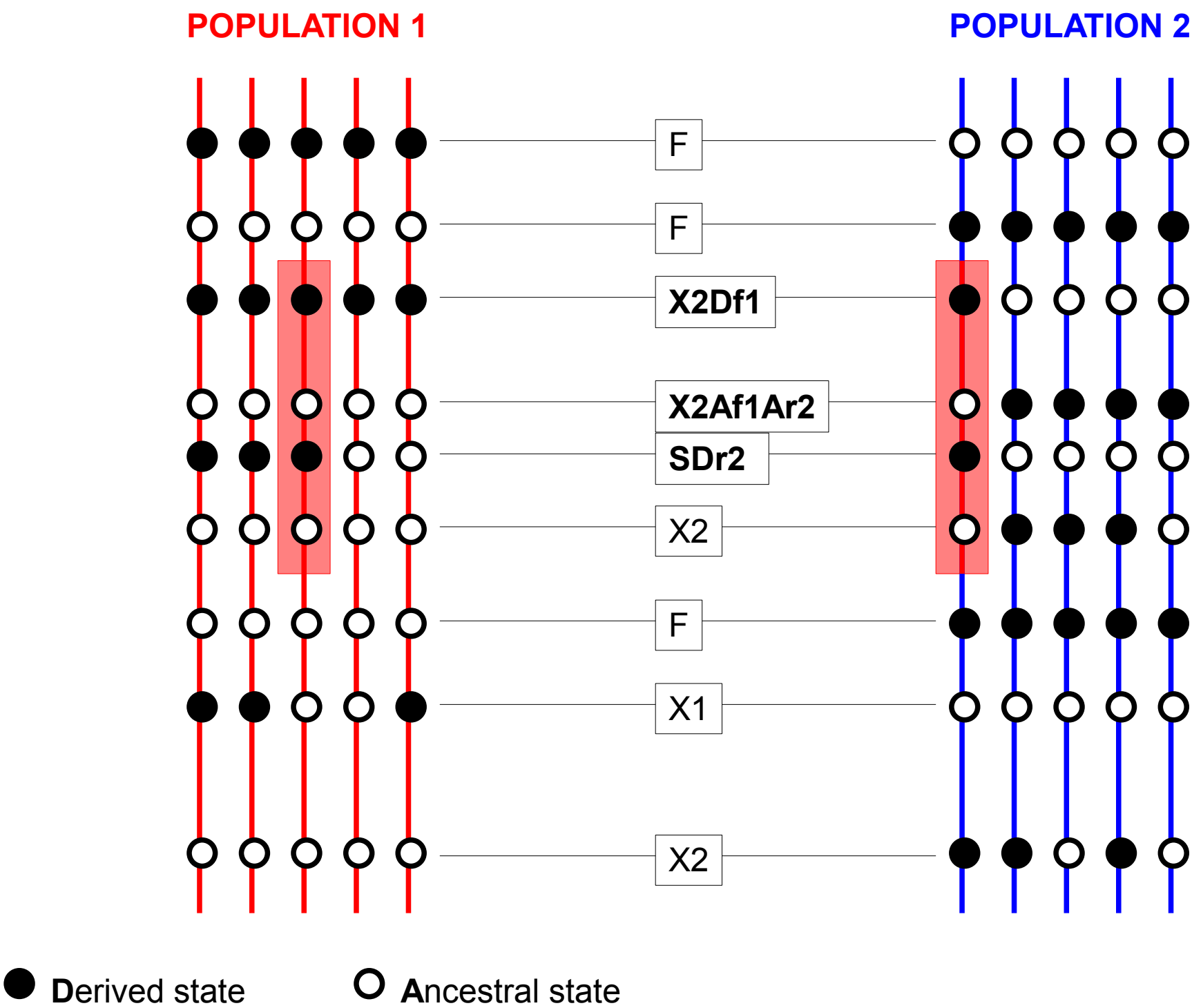


Kolmogorov–Smirnov test (on 10 loci)

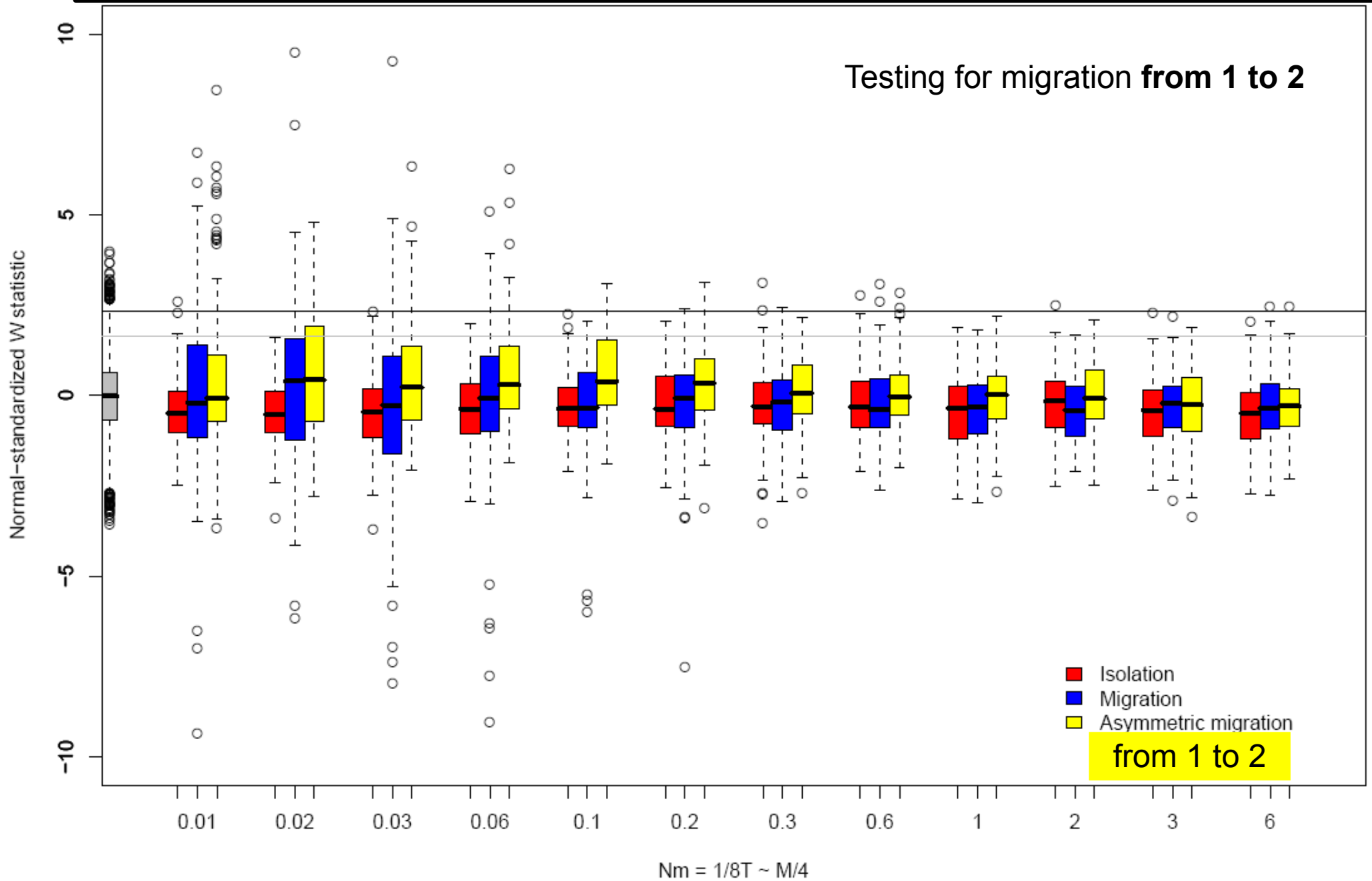
Isolation:	1/10
Migration:	9/10
Asymmetric mig.:	10/10

Asymmetric migration

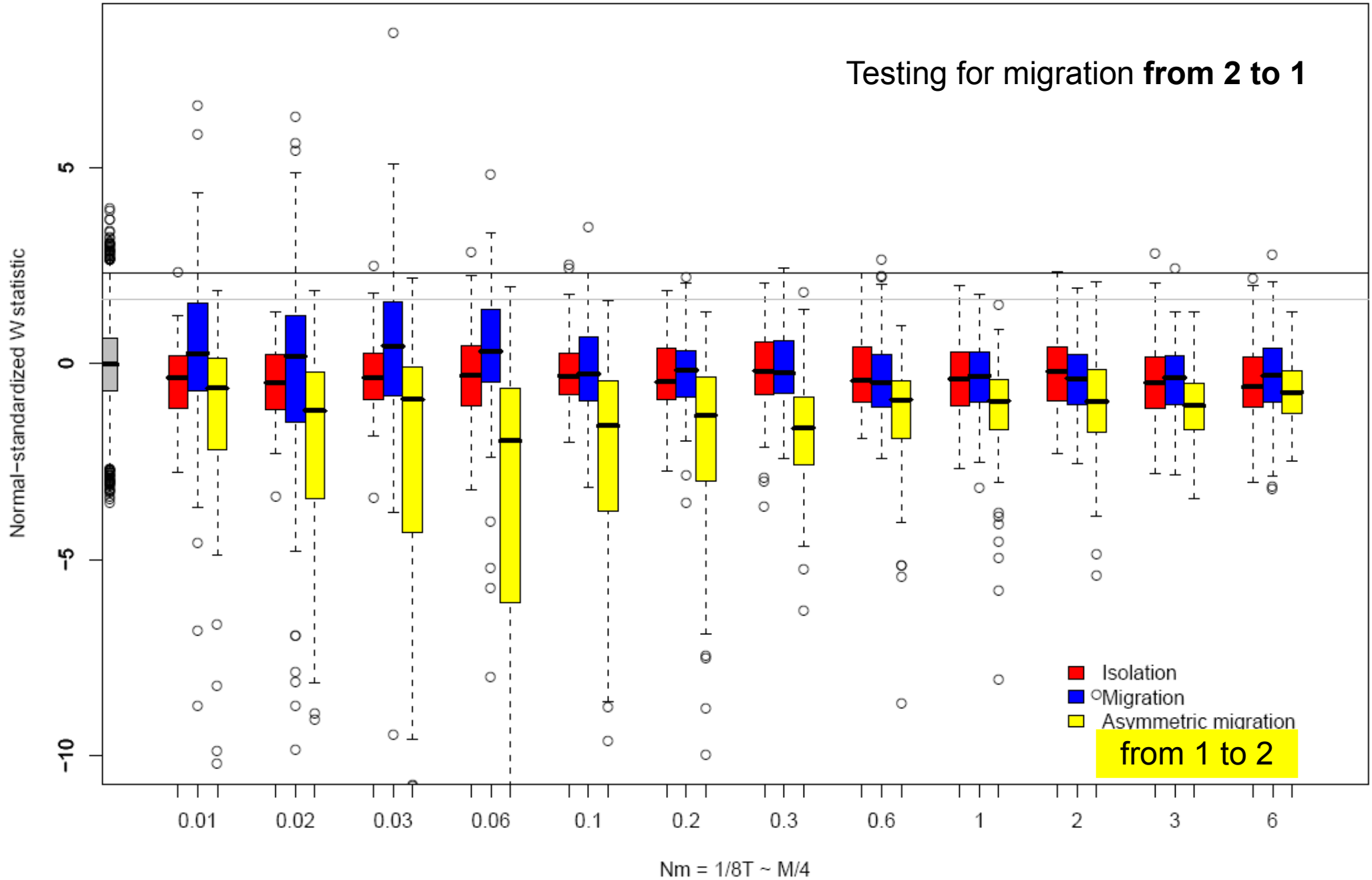

POPULATION 1
POPULATION 2
F
F
F
F
X1
X2
F
X1
X2
Derived state
Ancestral state


POPULATION 1
POPULATION 2
F
F
X2Df1
X2Af1Ar2
SDr2
X2
F
X1
X2
Derived state
Ancestral state

SDr2+X2Df1+X2Af1Ar2 vs. SDr1+X1Df2+X1Af2Ar2

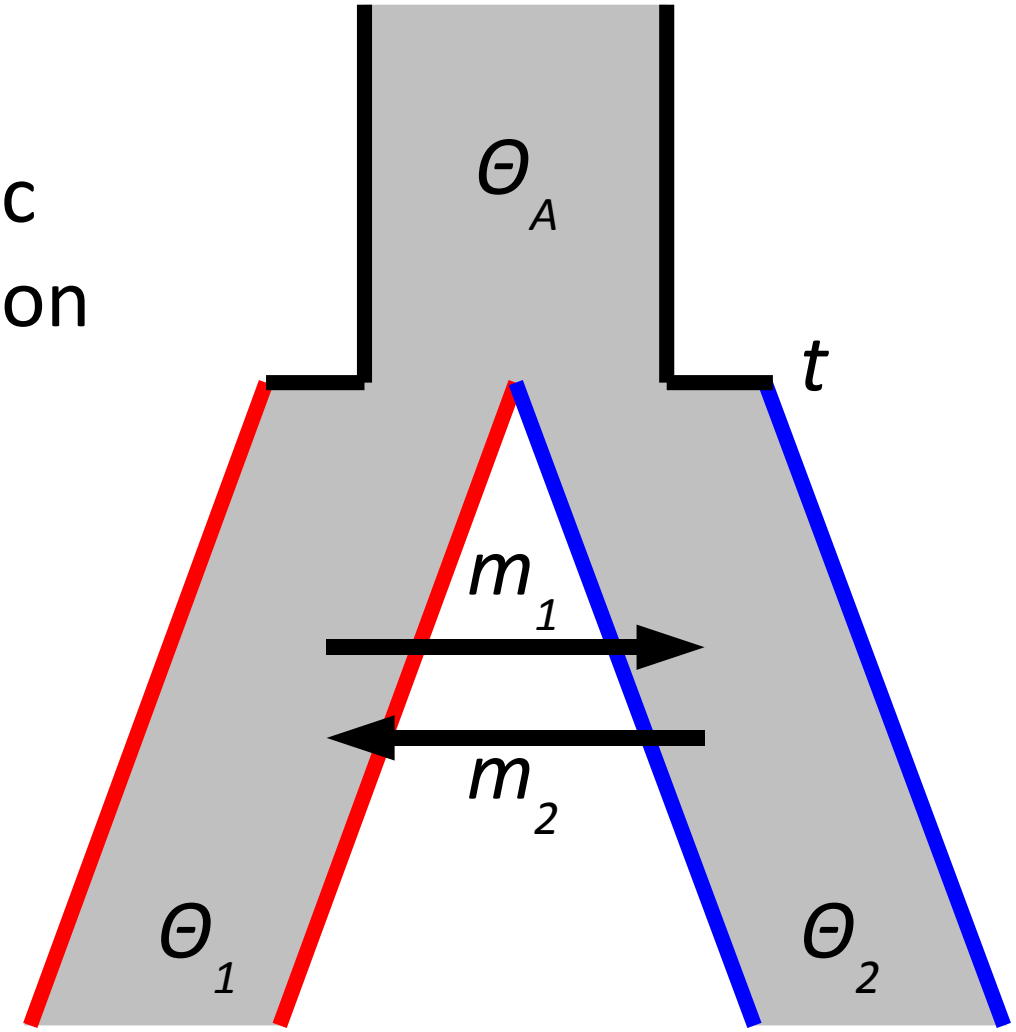


SDr1+X1Df2+X1Af2Ar1 vs. SDr2+X2Df1+X2Af1Ar1



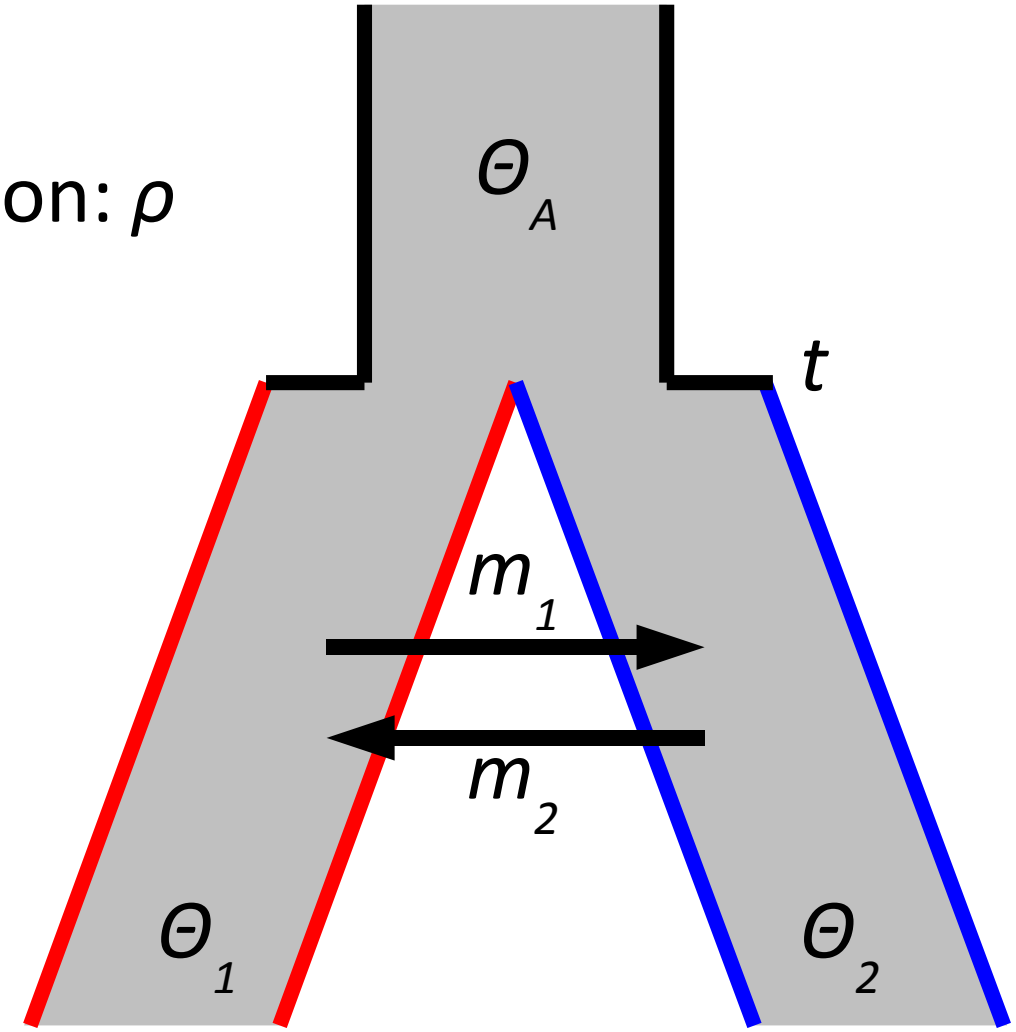
Likelihood Coalescent MCMC
(Nielsen & Wakeley 2001, Hey & Nielsen 2004)

no intragenic recombination



Summary Statistics Coalescent MCMC
(Becquet & Przeworski 2007)

recombination: ρ

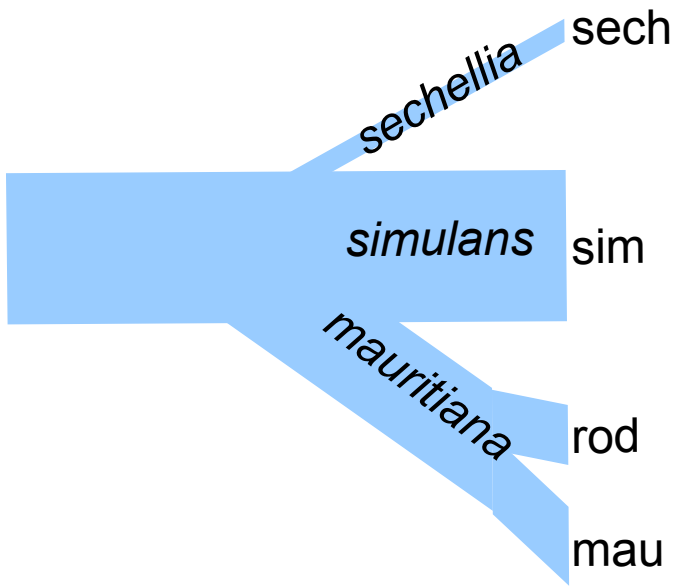


Summary

- Clustering of certain types of polymorphism
- Non-parametric statistics: based on runs
- Statistics sensitive to migration
- Statistics sensitive to direction of migration
- Summary statistics replacing likelihood (e.g. ABC)

Drosophila simulans, sechellia, mauritiania

Marie-Louise Cariou and team



amyrel
joc
notch
obp
otu
period
vermilion
white

AGENCE NATIONALE DE LA RECHERCHE ANR Biodiversité: TRANSBIODIV

Arabidopsis: Xavier Vekemans and team

Oaks: Antoine Kremer and team

Millet: Thierry Robert and team

Mice: Pierre Boursot and team