Identifying local adaptation of quantitative traits in structured populations

(Using LAVA)

Montpellier, November 2025
Isabela do O

Illustration by Edward Donovan

Populations are spread across heterogeneous environments, thus likely heterogeneous selective pressures.

wikipedia

iNaturalist

Populations are spread across heterogeneous environments, thus likely heterogeneous selective pressures.

wikipedia

iNaturalist

1. There must be some structure

Populations are spread across heterogeneous environments, thus likely heterogeneous selective pressures.

wikipedia

iNaturalist

1. There must be some structure
2. Local adaptation is expected

Populations are spread across heterogeneous environments, thus likely heterogeneous selective pressures.

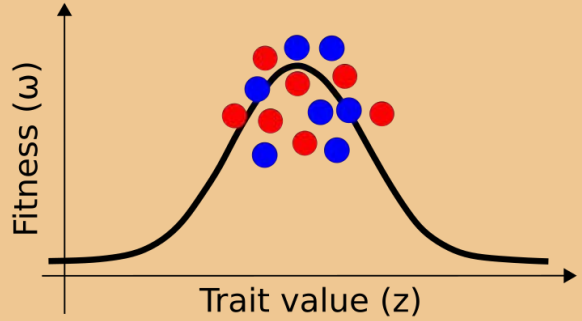
wikipedia

iNaturalist

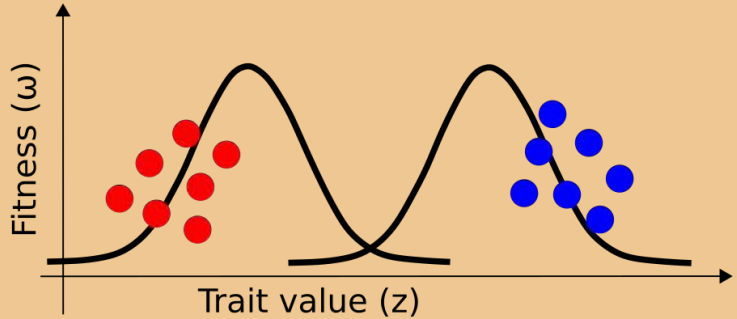
1. There must be some structure
2. Local adaptation is expected
3. Global adaptation as well

Sam Yeaman refers to as

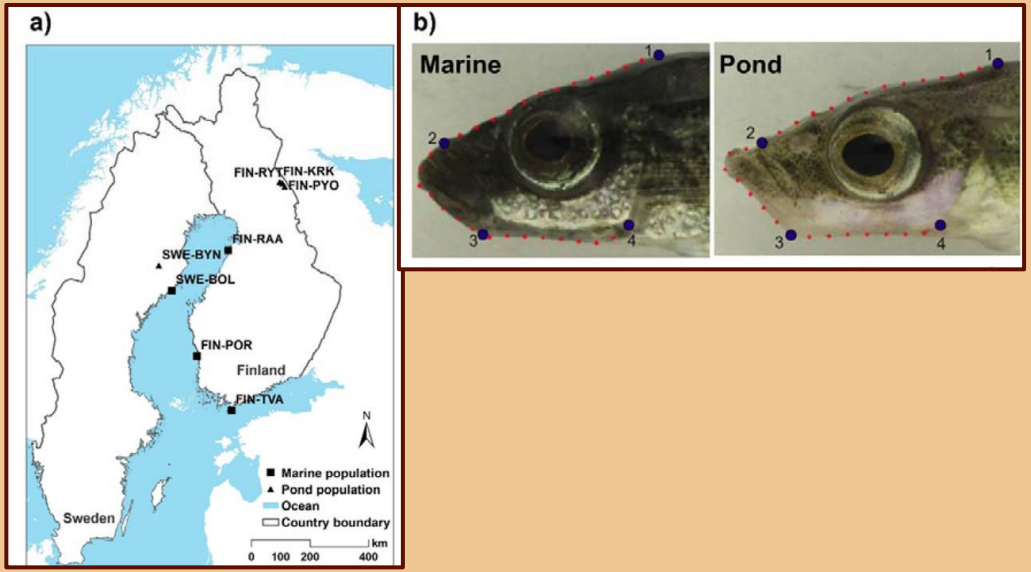
Global adaptation:



Local adaptation:



Example



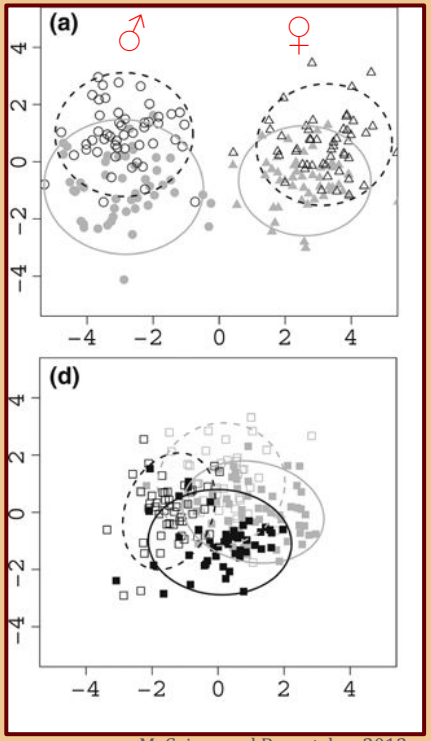
Fraimout et al (2023)

iNaturalist

Standardizing the environmental variance

marine

freshwater

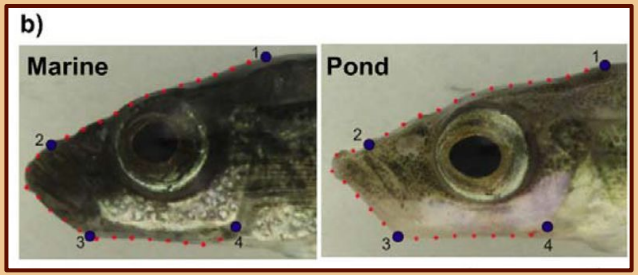


Wild individuals

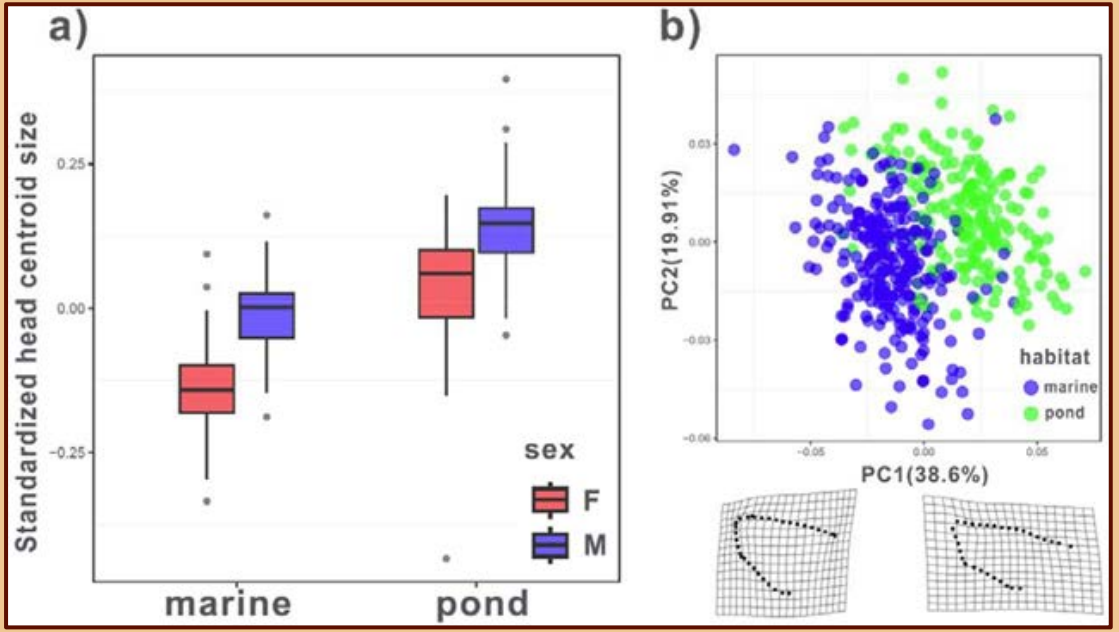
Lab-reared

McCairns and Bernatchez 2012

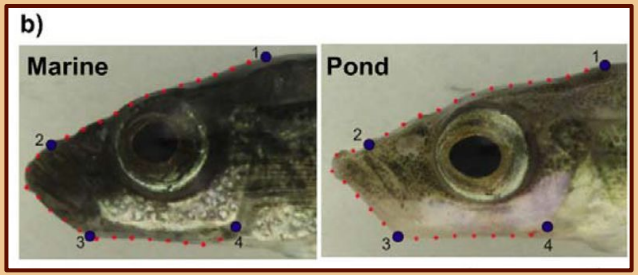
Fraimout *et al* (2023)



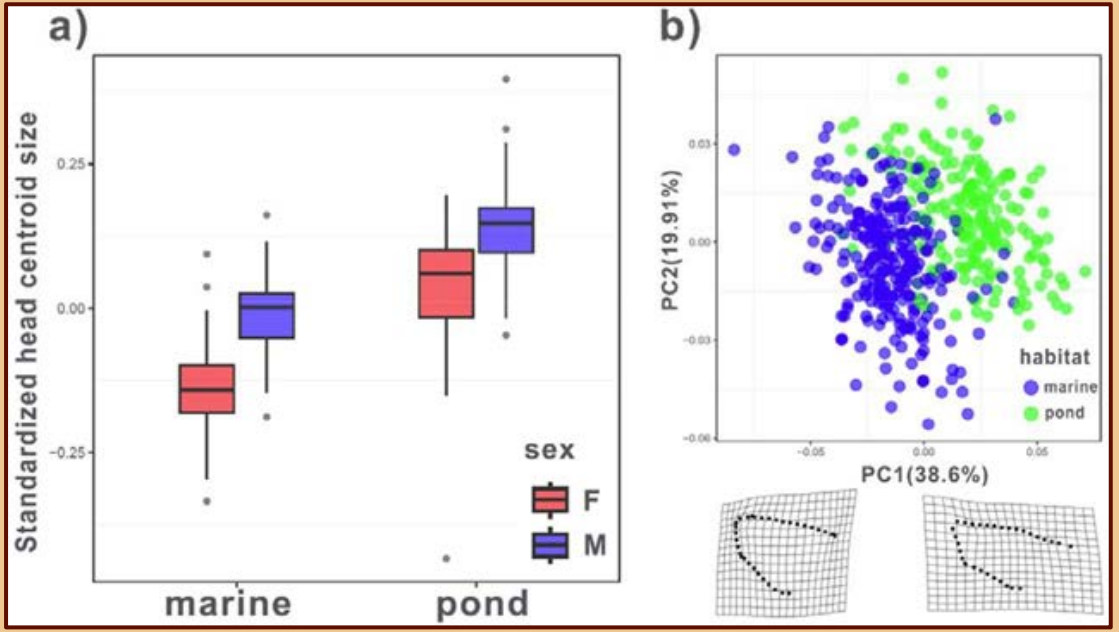
Fraimout et al (2023)



Is it local adaptation?



Fraimout et al (2023)



Drift alone will create divergence

$$V_B(t) = 2V_A$$

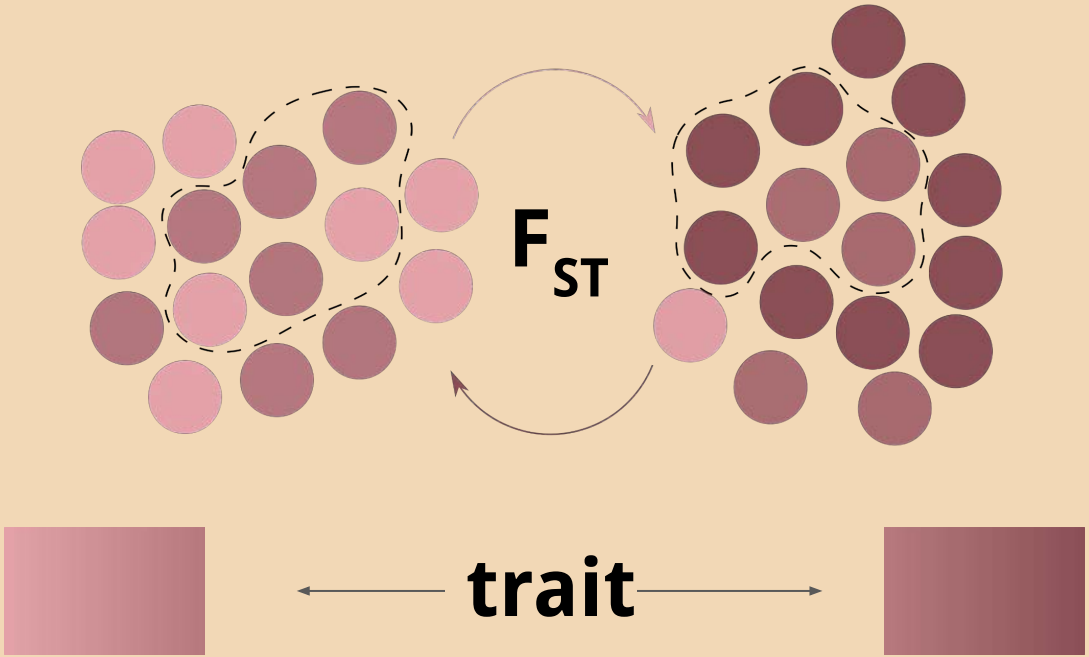
Walsh and Lynch, Chapter 12 (2018)

32

How to identify local adaptation?

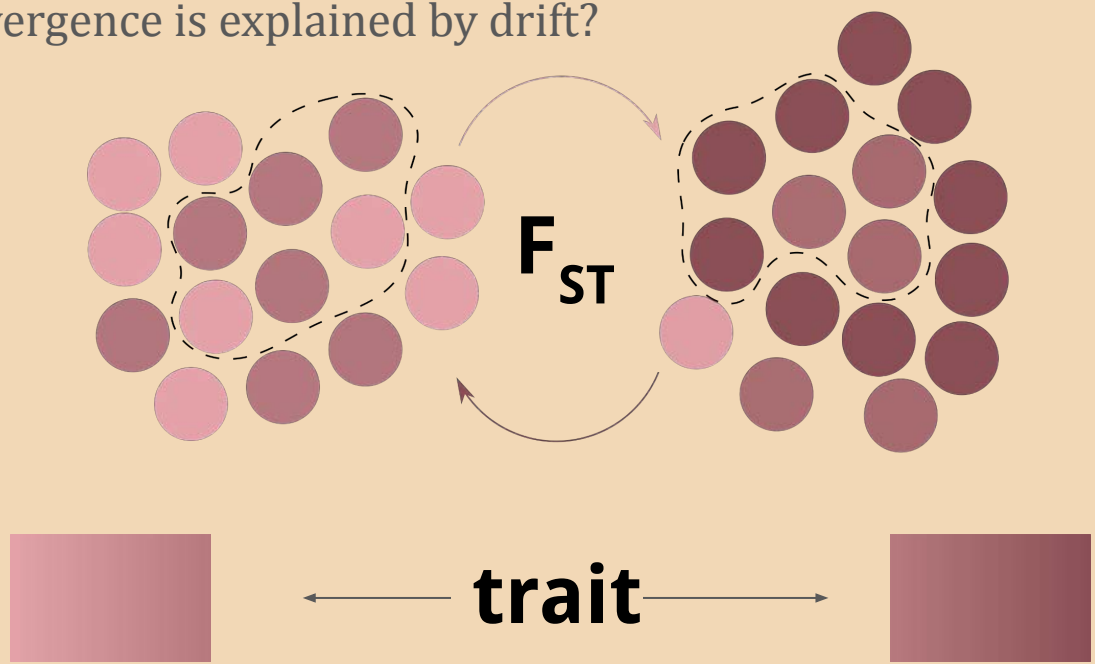
London. Pub.d as the Act directs by E. Donovan, & F. & C. Rivington. July 1. 1803.

Creating a neutral hypothesis

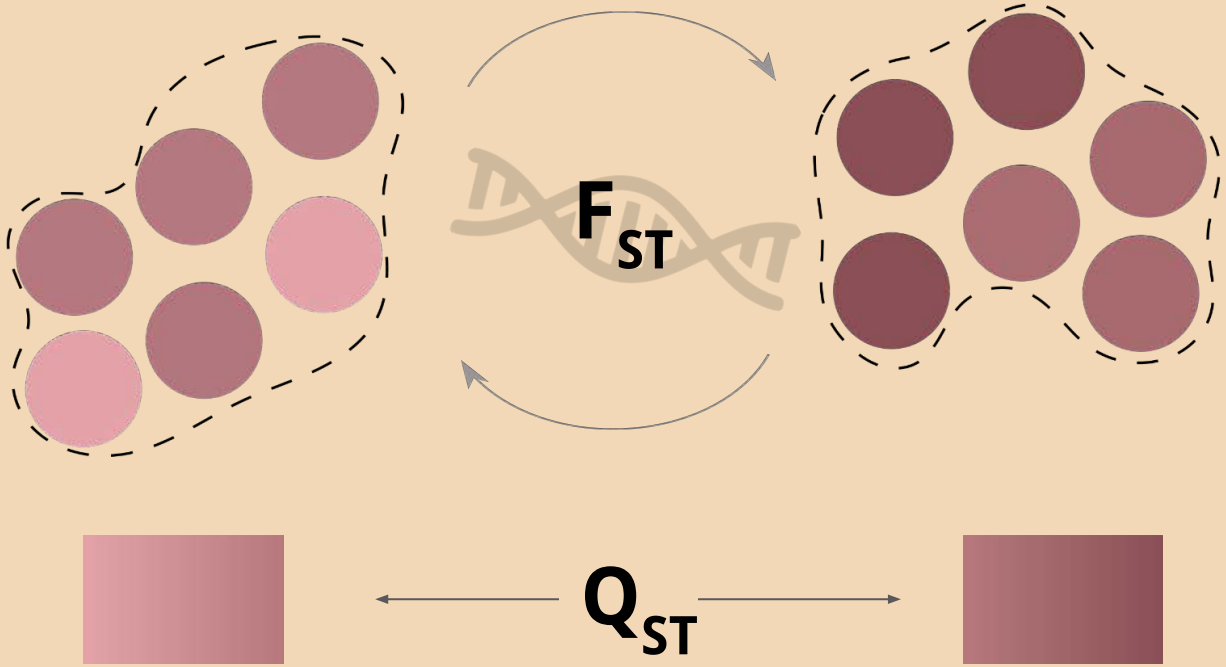


Creating a neutral hypothesis

How much divergence is explained by drift?



Q_{ST} - F_{ST}



Q_{ST} - F_{ST}

$$Q_{ST} = \frac{V_B}{V_B + 2V_W}$$

Spitze (1993)

Q_{ST} - F_{ST}

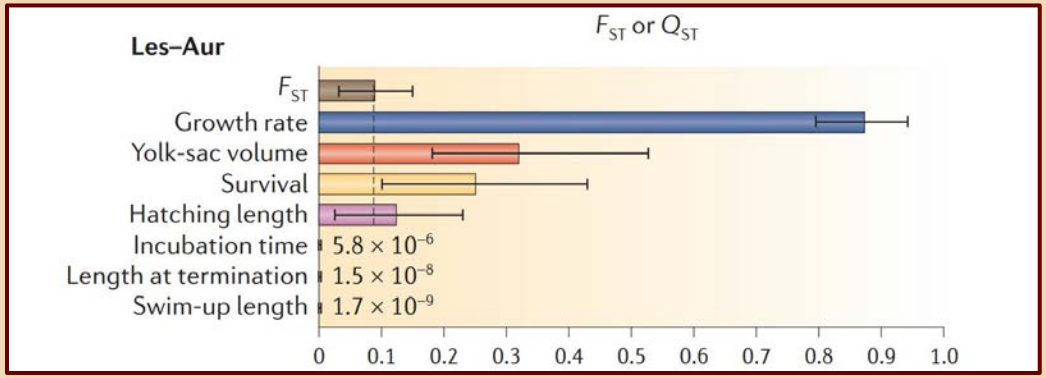
Additive genetic variance between

$$Q_{ST} = \frac{V_B}{V_B + 2V_W}$$

Spitze (1993)

Additive genetic variance within

Grayling, Thymallus thymallus - by Vera Kuttelvaserova



$$Q_{ST} > F_{ST} \rightarrow \text{Local adaptation}$$

$$Q_{ST} < F_{ST} \rightarrow \text{Global adaptation}$$

$$Q_{ST} = F_{ST} \rightarrow \text{Drift}$$

Q_{ST} - F_{ST} issues

$$Q_{ST} = \frac{V_B}{V_B + 2V_W}$$

How do we estimate these variances?

$$z \sim \mu + a_i + a_p + \epsilon$$

How do we estimate these variances?

trait value | individual-level random effect | residual

$$z \sim \mu + a_i + a_p + \epsilon$$

mean | subpop.-level random effect

How do we estimate these variances?

$$V_W$$

$$z \sim \mu + a_i + a_p + \epsilon$$

$$V_B$$

How do we estimate these variances?

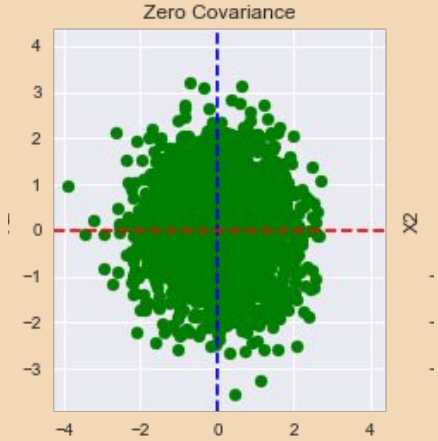
$$V_W$$

$$z \sim \mu + a_i + a_p + \epsilon$$

$$V_B$$

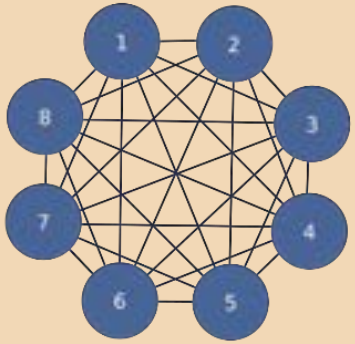
Q_{ST} - F_{ST} issues

$$a_p \sim \mathcal{N}(0, IV_B)$$



From: https://magic-with-latents.github.io/latent/

Island Model

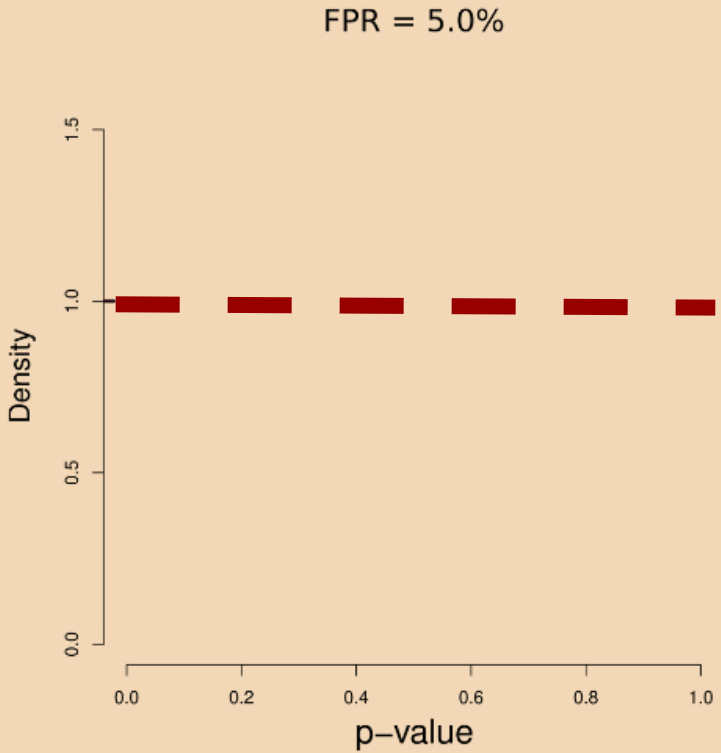


Stepping Stones

Issue: Qst-Fst makes strong assumptions on population structure

We tested the impact of this assumption with simulated data

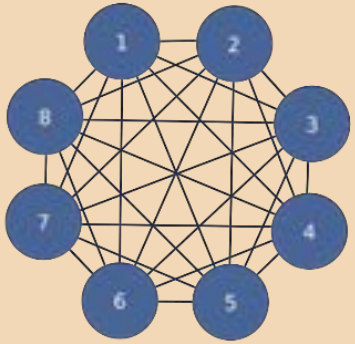
Since they are neutral populations, we expect:



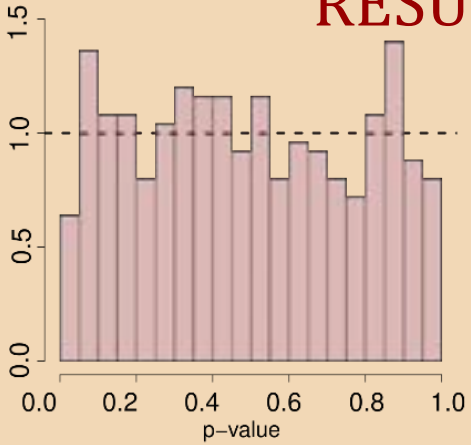
Q_{ST} - F_{ST} issues

RESULTS

Island Model



Density

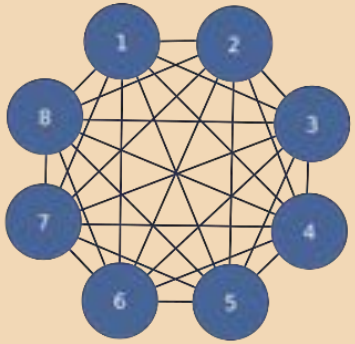


Stepping Stones

Q_{ST} - F_{ST} issues

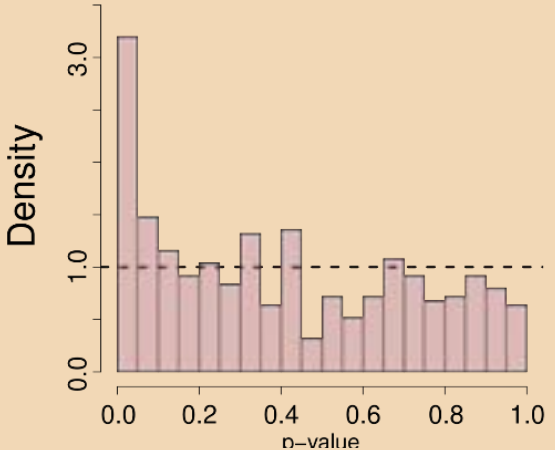
RESULTS

Island Model



**** FPR = 16%

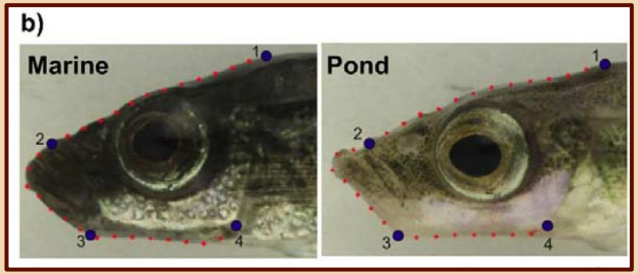
Stepping Stones



A New Method to Uncover Signatures of Divergent and Stabilizing Selection in Quantitative Traits FREE

Otso Ovaskainen ✉, Markku Karhunen, Chaozhi Zheng, José Manuel Cano Arias, Juha Merilä Author Notes

Genetics, Volume 189, Issue 2, 1 October 2011, Pages 621–632, https://doi.org/10.1534/genetics.111.129387



Fraimout et al (2023)

$$a_p \sim \mathcal{N}(0, IV_B)$$

Coancestry matrix

$$a_p \sim \mathcal{N}\left(0, 2G^A \otimes \mathbf{\Theta^P}\right)$$

Ancestral G matrix

$$S = \Pr(f_P(a_p) < f_P(a^R))$$

Randomly generated pattern
sampled from normal distribution

$$S = \Pr(f_P(a_p) < f_P(a^R))$$

Observed pattern

S represents the **probability** that a randomly generated pattern (through sampling from the random realizations) would show **equal or greater deviation from neutrality** than the observed pattern

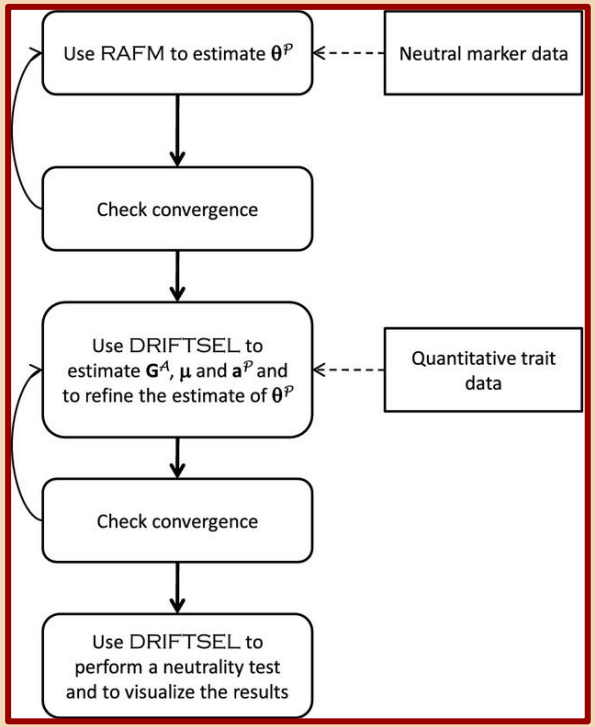
$$S = \Pr(f_P(a_p) < f_P(a^R))$$

S ~ 0.5 → Drift alone can explain the pattern

S ~ 0 → Global adaptation

S ~ 1 → Local adaptation

Using “Driftsel”



MOLECULAR ECOLOGY RESOURCES

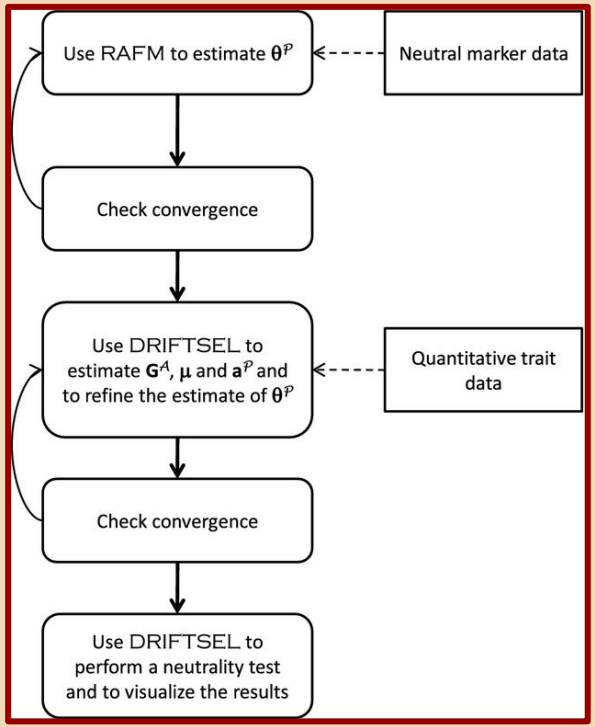
Resource Article | Full Access

DRIFTSEL: an R package for detecting signals of natural selection in quantitative traits

M. Karhunen, J. Merilä, T. Leinonen, J. M. Cano, O. Ovaskainen

First published: 09 May 2013 | https://doi.org/10.1111/1755-0998.12111 | Citations: 44

Using “Driftsel”



MOLECULAR ECOLOGY RESOURCES

Resource Article | Full Access

DRIFTSEL: an R package for detecting signals of natural selection in quantitative traits

M. Karhunen, J. Merilä, T. Leinonen, J. M. Cano, O. Ovaskainen

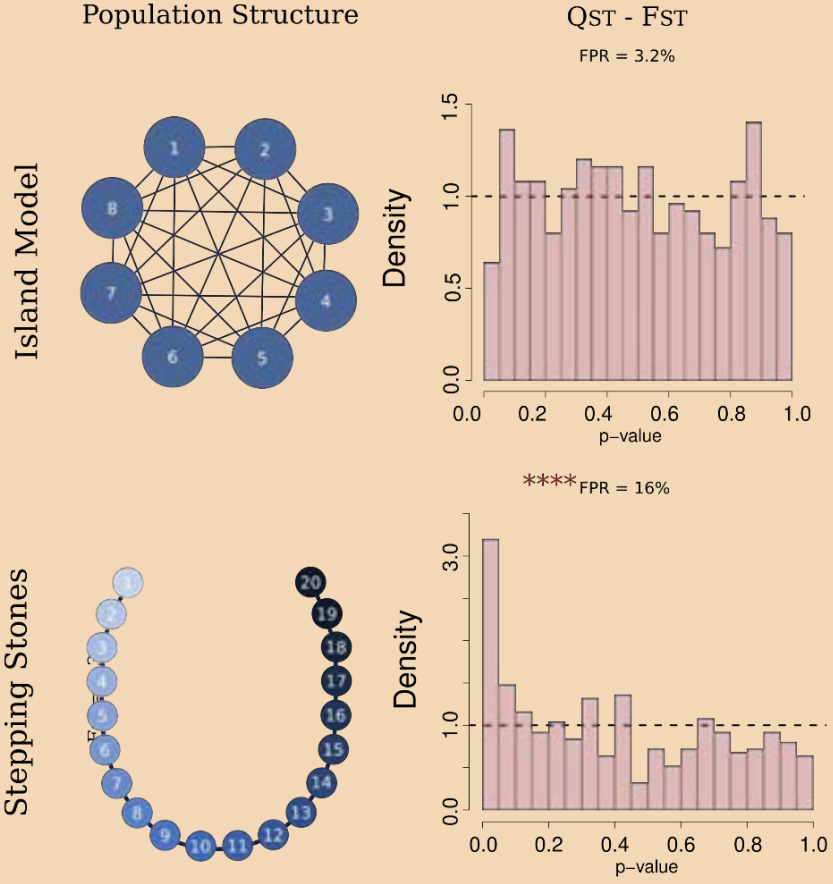
First published: 09 May 2013 | https://doi.org/10.1111/1755-0998.12111 | Citations: 44

An allele-sharing, moment-based estimator of global, population-specific and population-pair F_{ST} under a general model of population structure

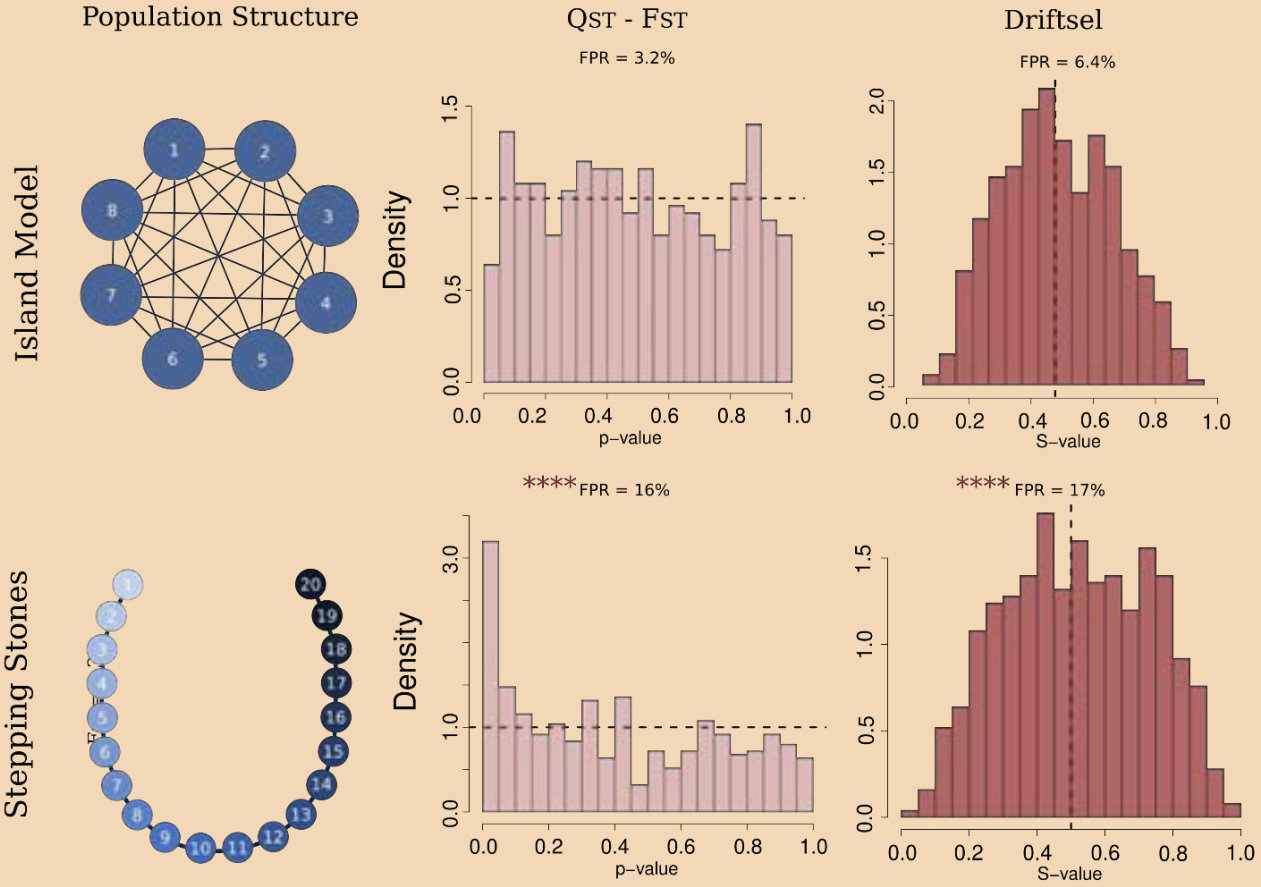
Jerome Goudet, Bruce S. Weir

The **AFM** assumes individuals within each population mate randomly (no inbreeding, no substructure, no assortative mating etc).

RESULTS


Population Structure
QST - FST
Island Model
FPR = 3.2%
Density
p−value
Stepping Stones
****FPR = 16%
Density
p−value

RESULTS

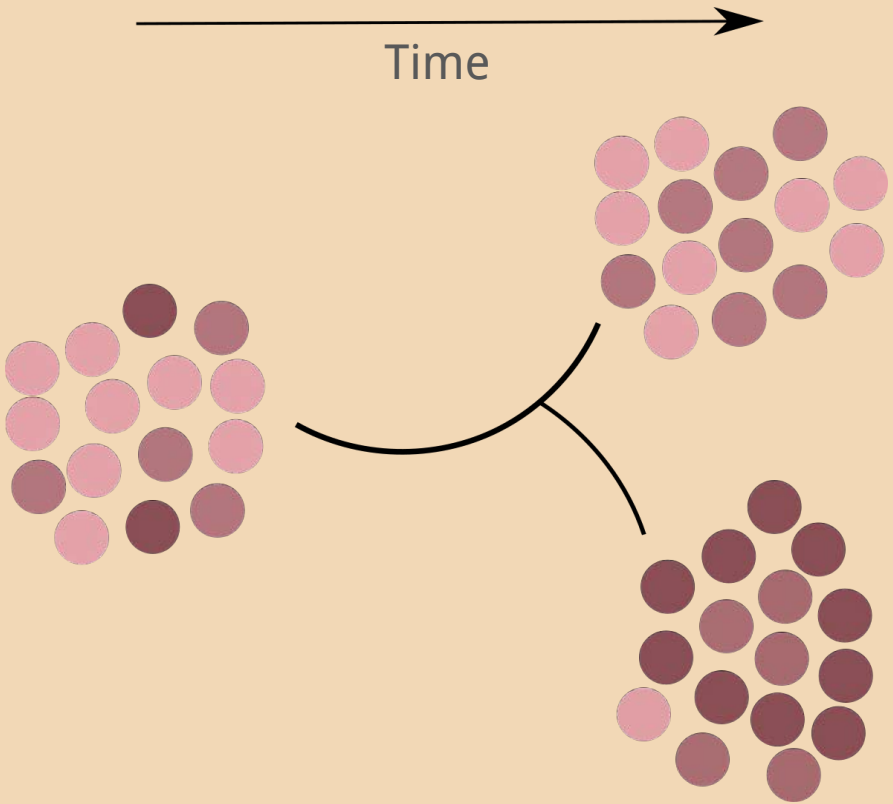

Population Structure
QST - FST
Driftsel
Island Model
Stepping Stones
Density
FPR = 3.2%
FPR = 6.4%
****FPR = 16%
****FPR = 17%
p−value
S−value

A method for identifying local adaptation in structured populations

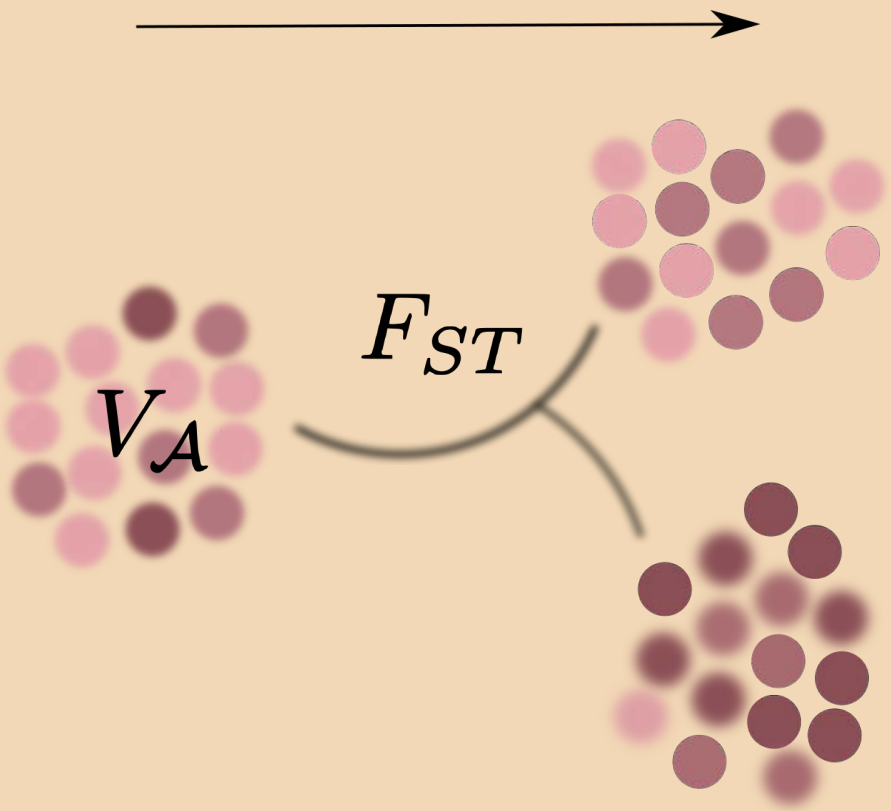
Isabela do O, Oscar E Gaggiotti, Pierre de Villemereuil, Jerome Goudet

Published: September 23, 2025 • https://doi.org/10.1371/journal.pgen.1011871

Quant Gen theory behind the method



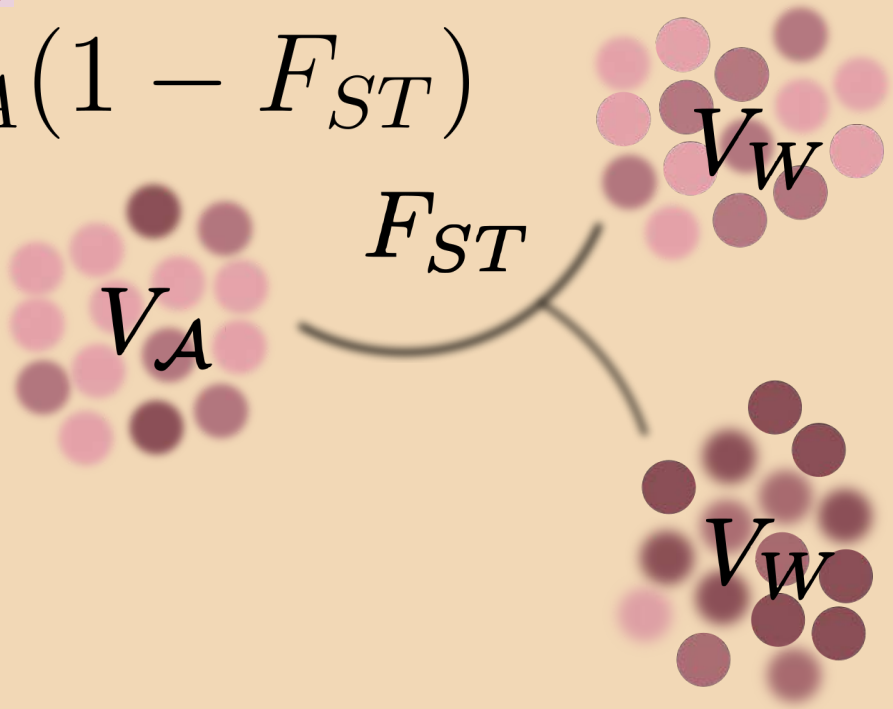
Quant Gen theory behind the method



Quant Gen theory behind the method

Under neutrality we can derive:

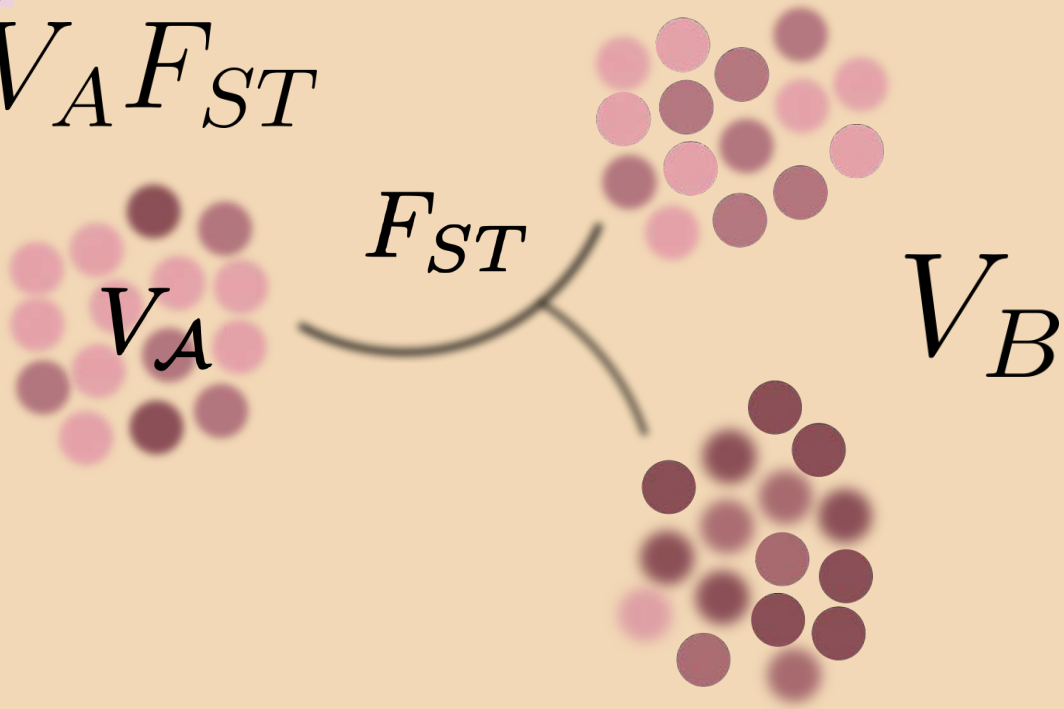
$$V_W = V_A(1 - F_{ST})$$



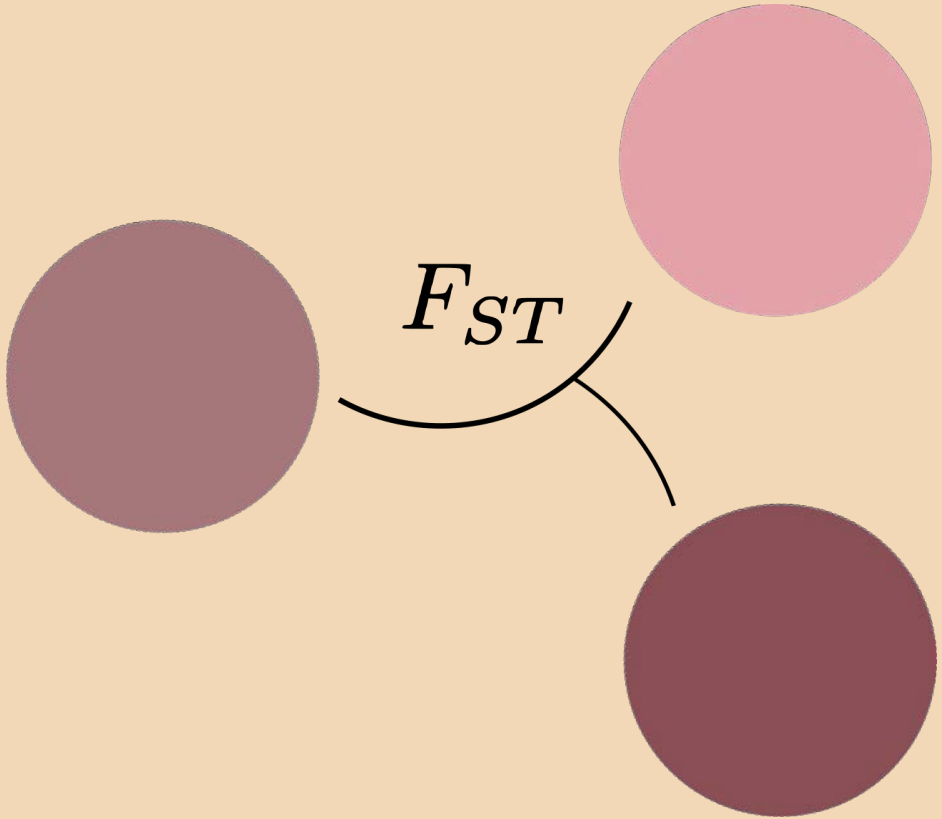
Quant Gen theory behind the method

Under neutrality:

$$V_B = 2V_A F_{ST}$$

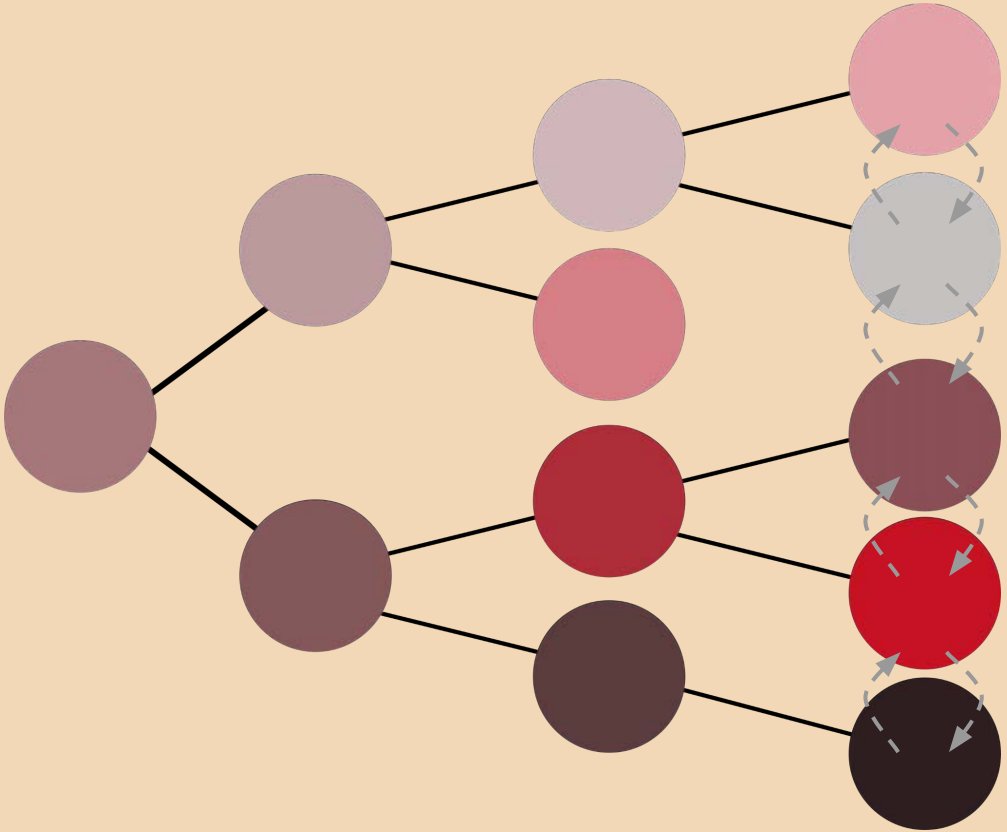


Quant Gen theory behind the method

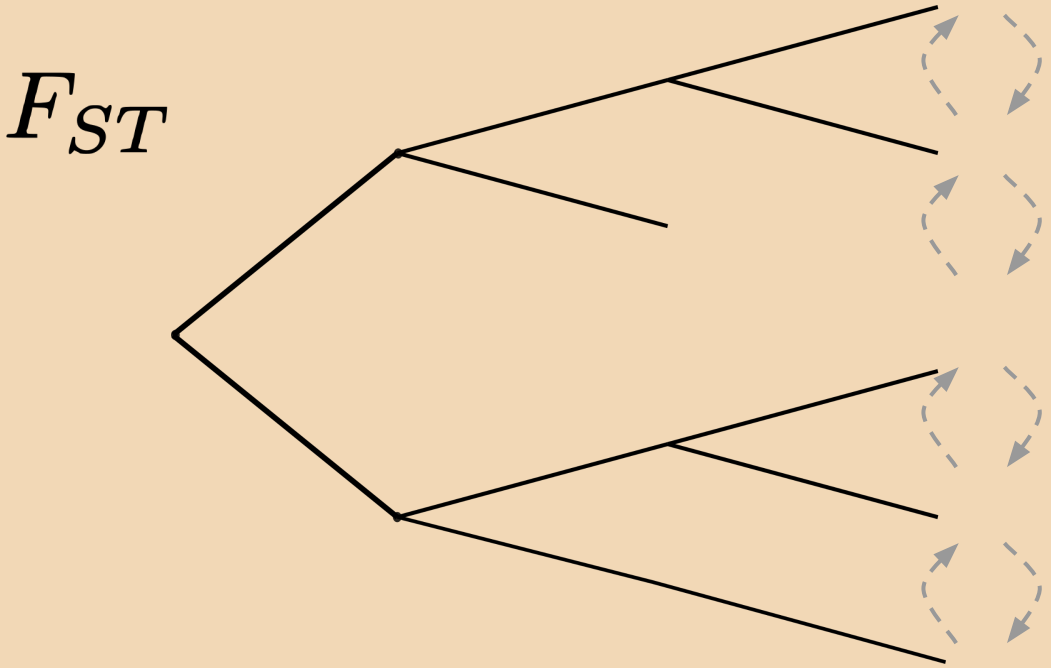


Quant Gen theory behind the method

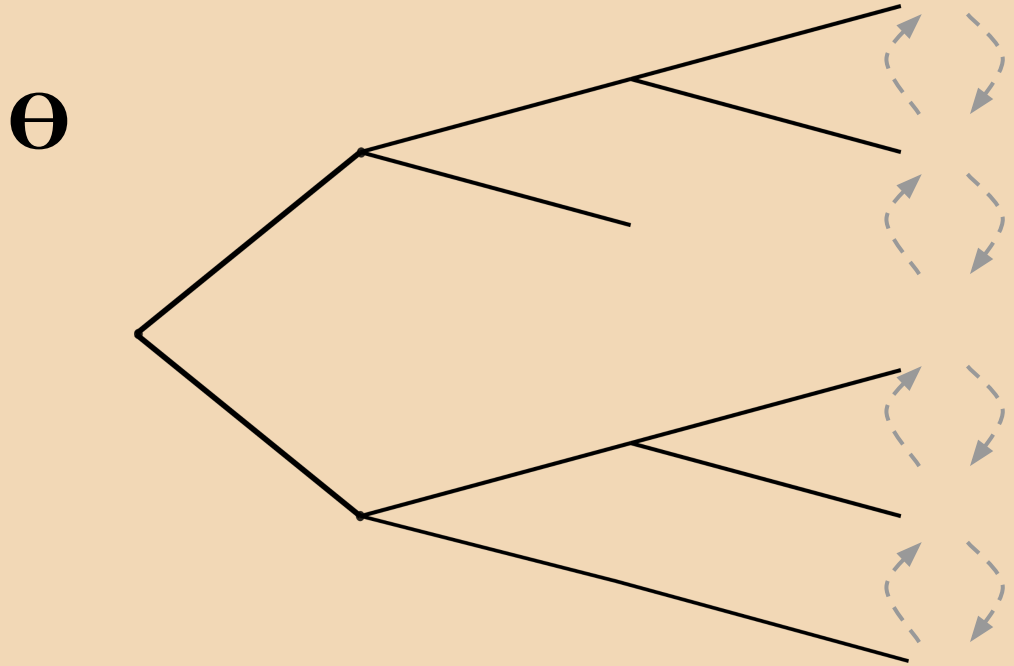
~~F_{ST}~~



Issue: Fst is not sufficient to describe complex structures

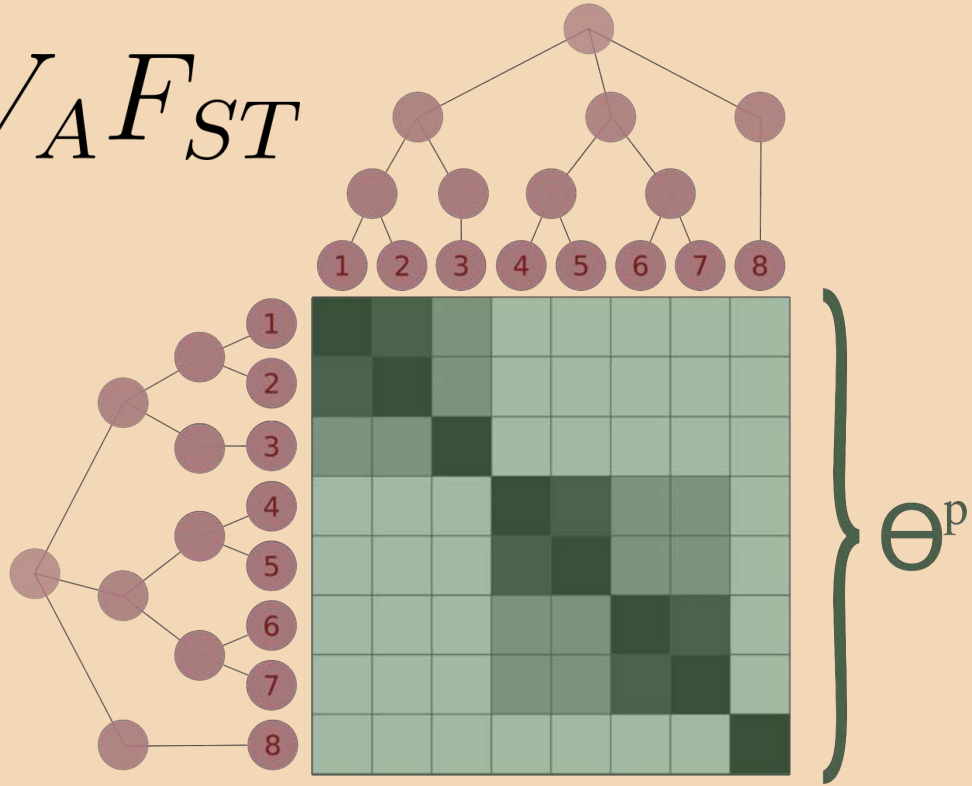


Solution: We use a matrix!



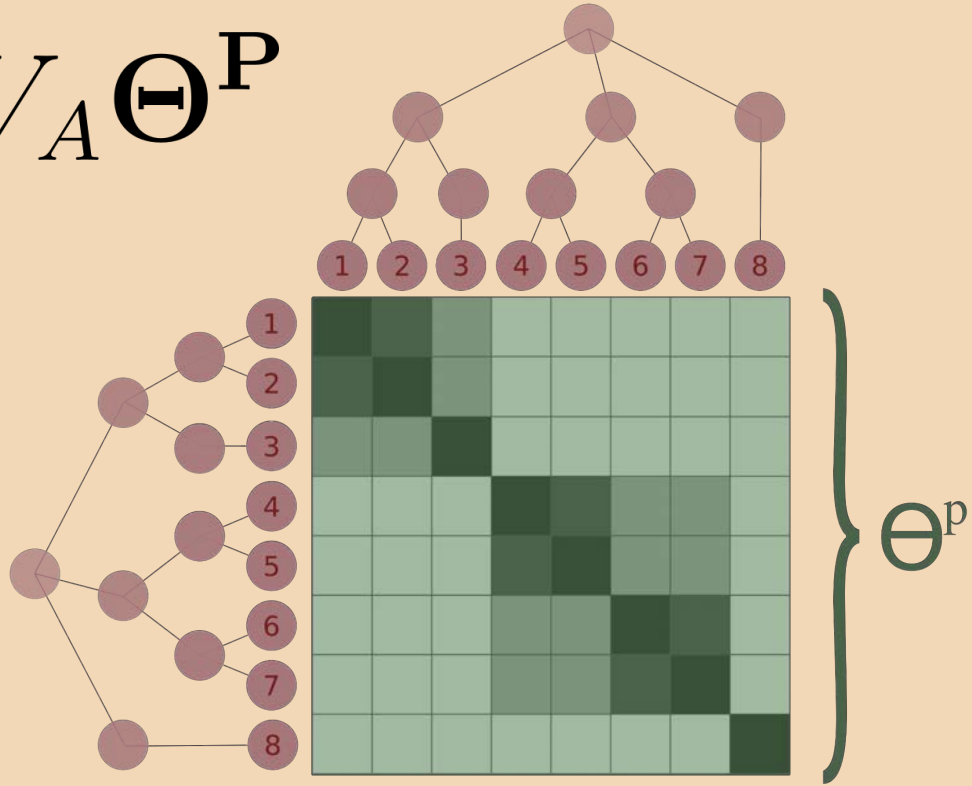
Quant Gen theory behind the method

$$V_B = 2V_A F_{ST}$$



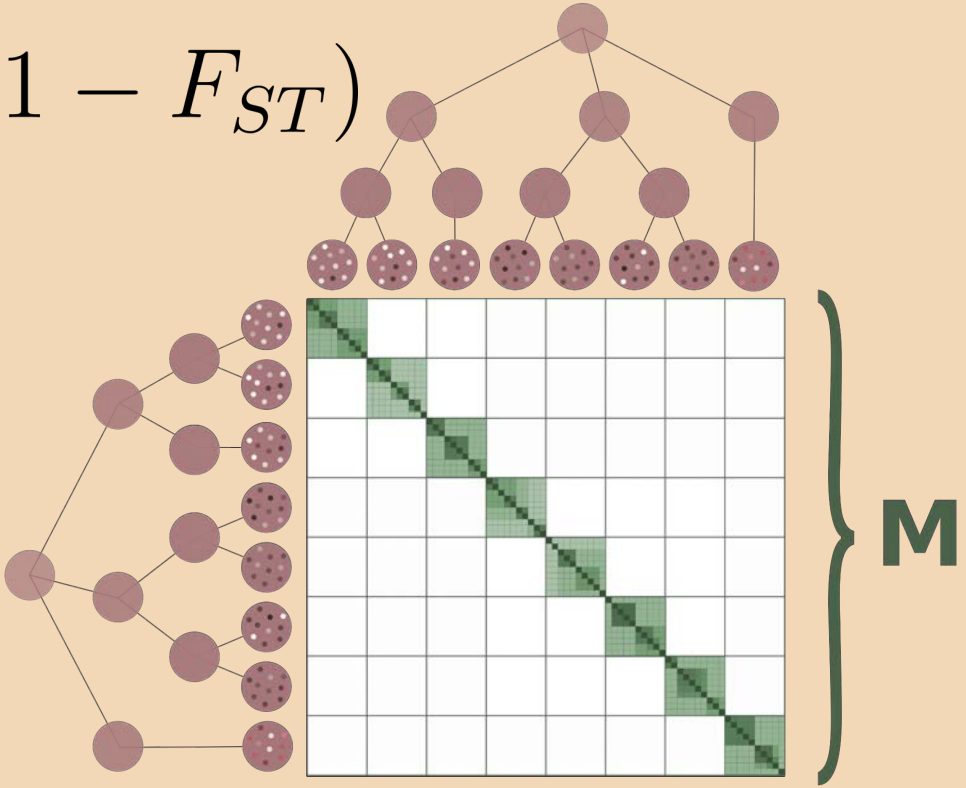
Quant Gen theory behind the method

$$V_B = 2V_A\mathbf{\Theta^P}$$



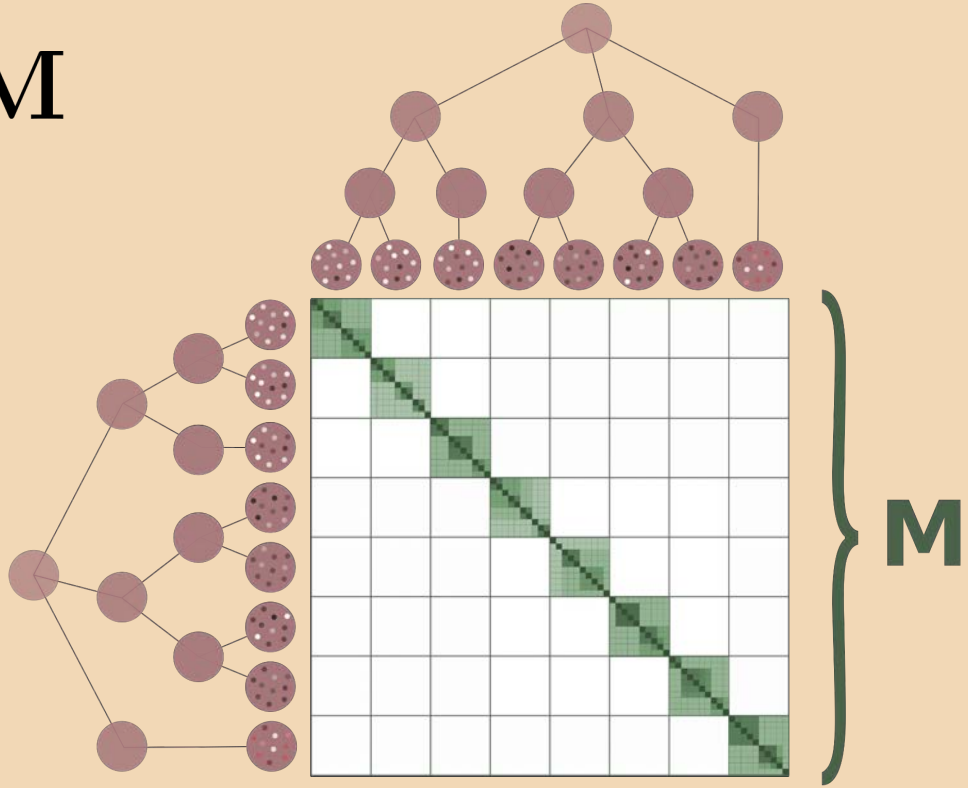
Quant Gen theory behind the method

$$V_W = V_A(1 - F_{ST})$$



Quant Gen theory behind the method

$$V_W = V_A \mathbf{M}$$



Quant Gen theory behind the method

$$z \sim \mu + a_i + a_p + \epsilon$$

$$a_i \sim N(0, \mathbf{M}V_{A,W}) \quad a_p \sim N(0, 2\mathbf{\Theta^P}V_{A,B})$$

LogAV method

Under neutrality:

$$V_B = 2V_A F_{ST} \longrightarrow \sum_B = 2V_A \Theta^P$$

$$V_W = V_A(1 - F_{ST}) \longrightarrow \sum_w = V_A \mathbf{M}$$

LogAV method

Provided $\mathcal{Z}$

$$\Sigma_B = 2V_A\Theta^P \xrightarrow{\text{Infer}} \hat{V}_{A,B}$$

$$\Sigma_w = V_A\mathbf{M} \xrightarrow{\text{Infer}} \hat{V}_{A,W}$$

LogAV method

Under neutrality:

$$\hat{V}_{A,W} = \hat{V}_{A,B}$$

LogAV method: **Hypothesis testing**

Under neutrality:

H0:

$$\hat{V}_{A,W} = \hat{V}_{A,B}$$

LogAV method: **Hypothesis testing**

Under neutrality:

Test statistics:

$$LogAV = log\left(\frac{V_{A,W}}{V_{A,B}}\right)$$

LogAV → LAVA

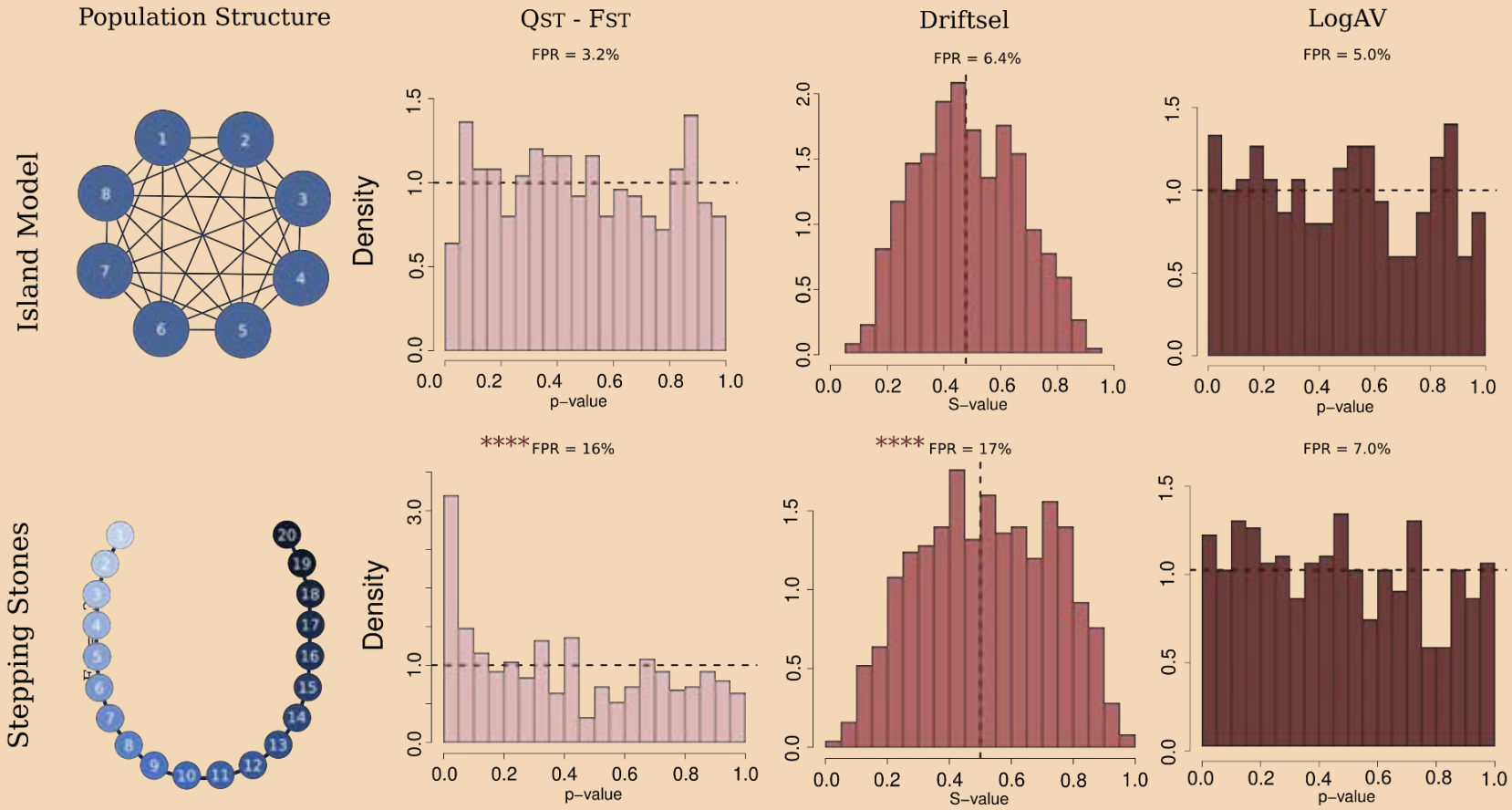
$$z \sim \mu + a_i + a_p + \epsilon$$

Preliminary tests:

RESULTS

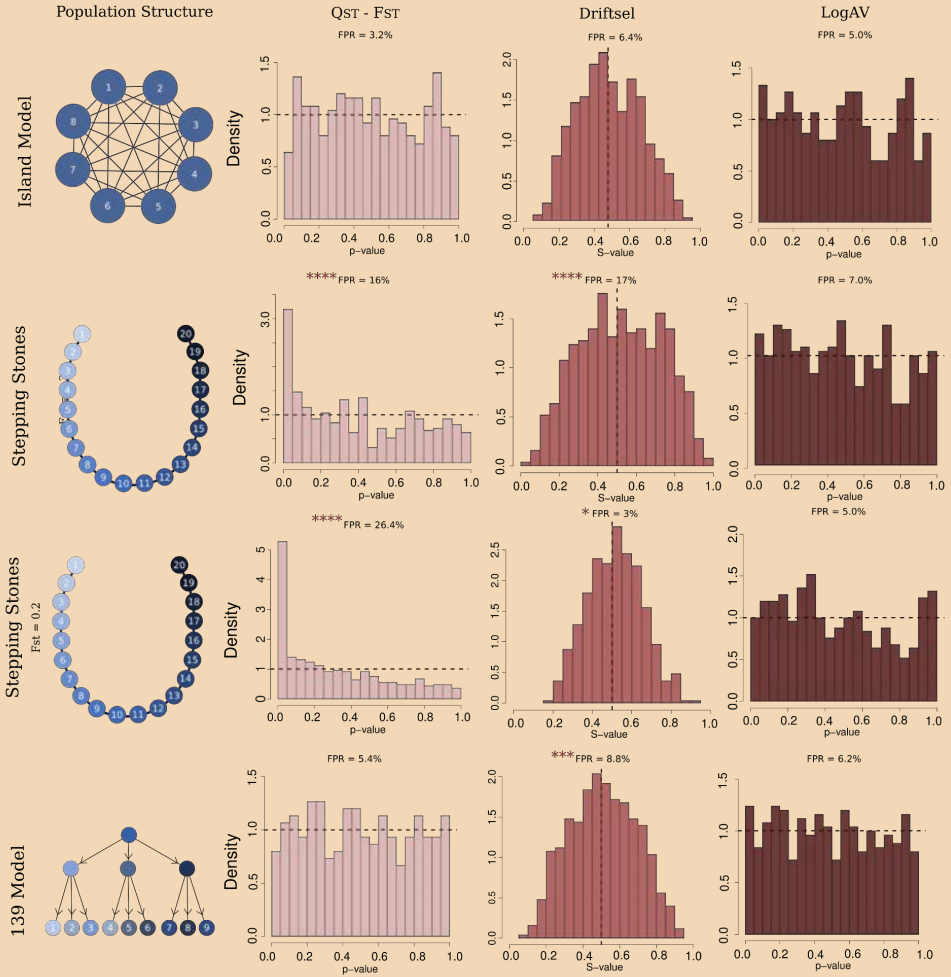
Preliminary tests:

RESULTS



Preliminary tests:

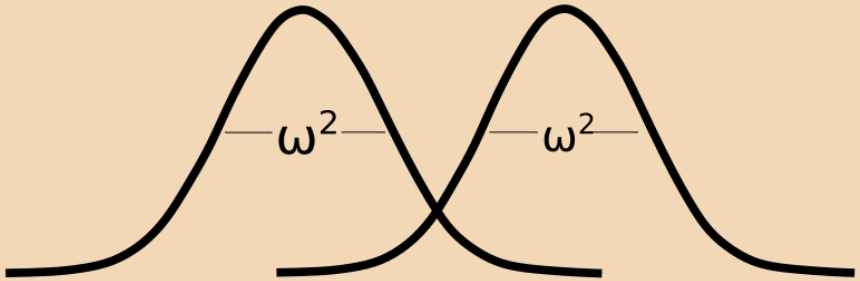
RESULTS



Preliminary tests:

RESULTS

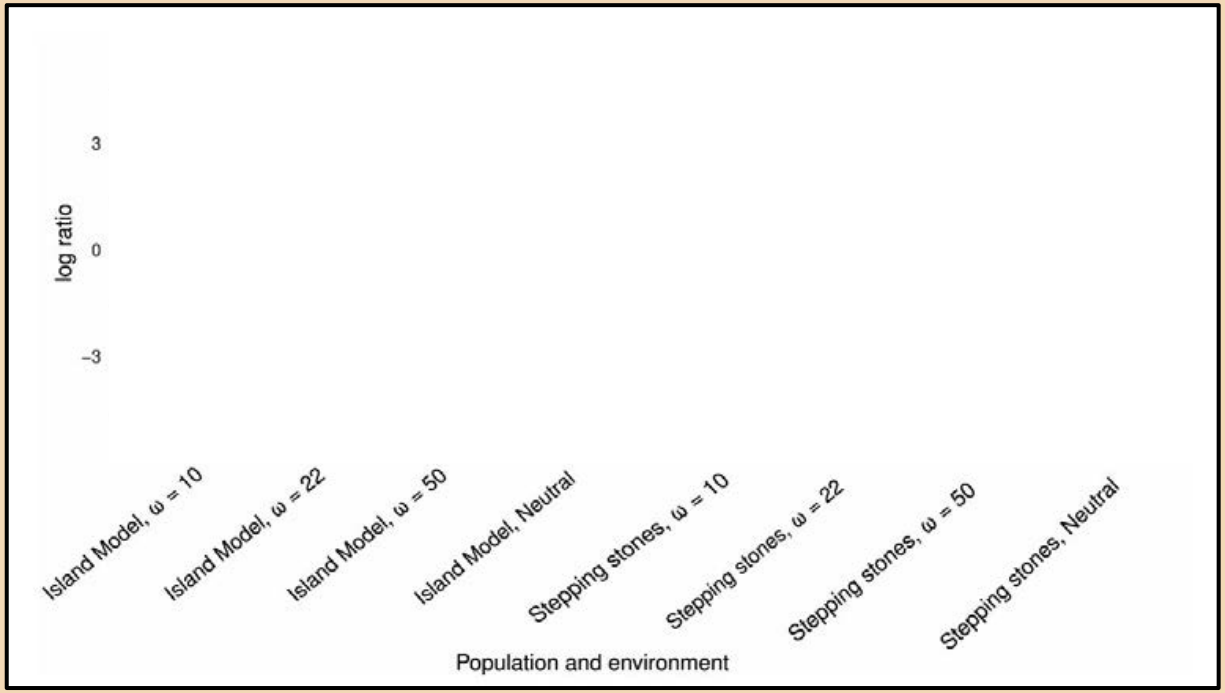
Selection scenarios



Preliminary tests:

RESULTS

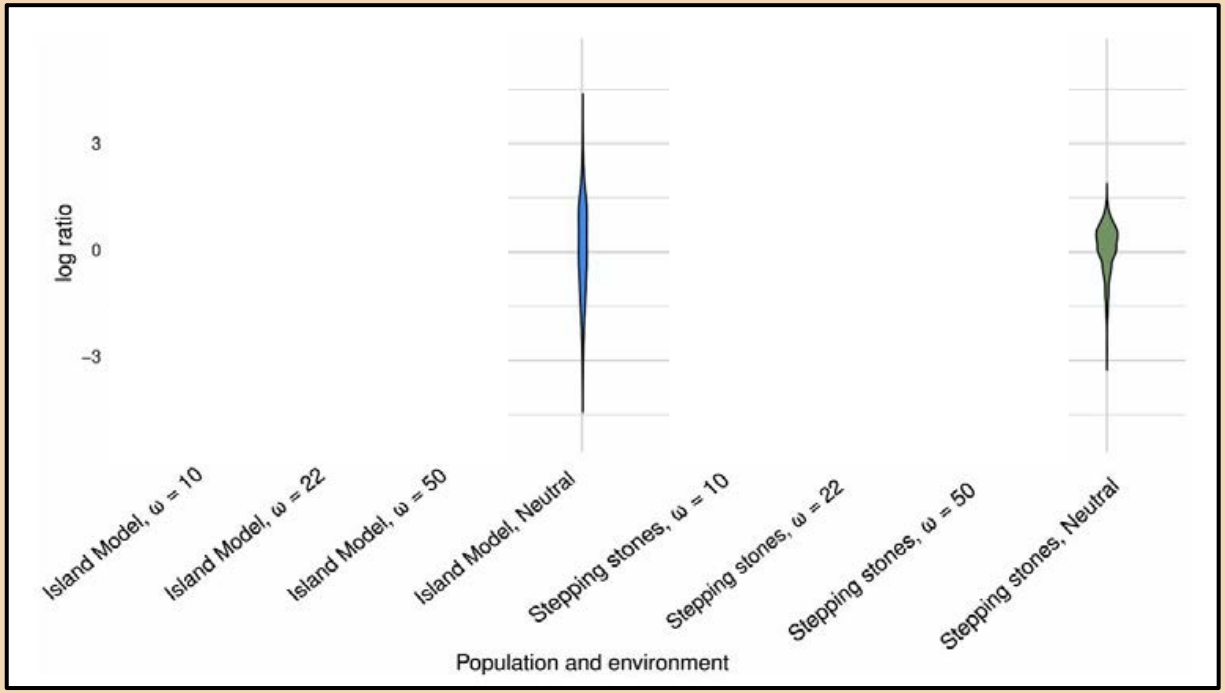
Selection scenarios



Preliminary tests:

RESULTS

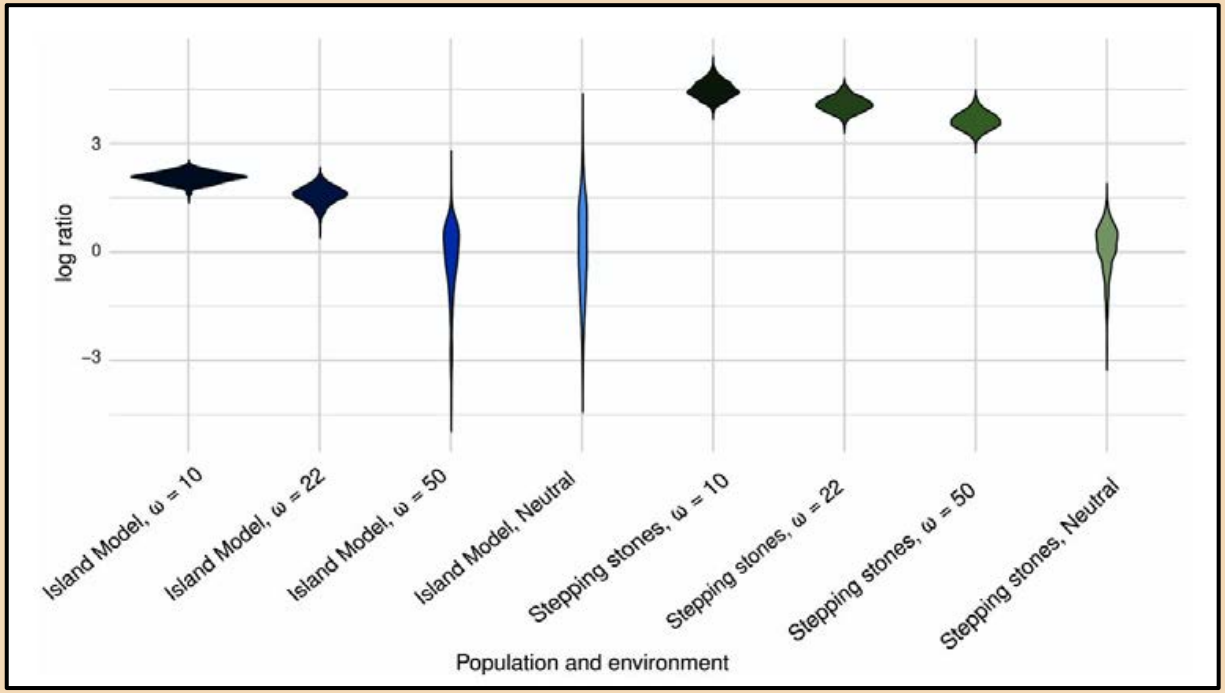
Selection scenarios



Preliminary tests:

RESULTS

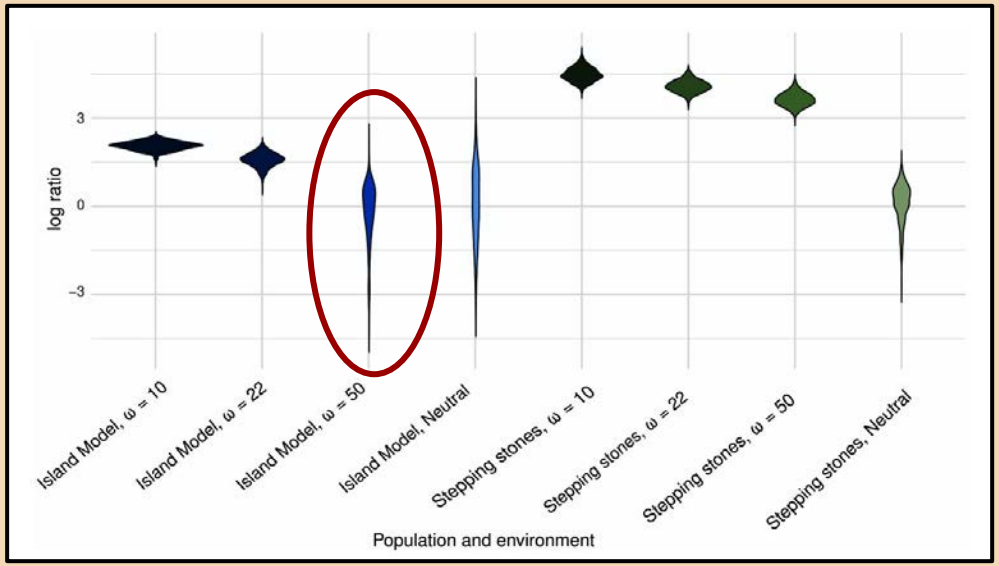
Selection scenarios

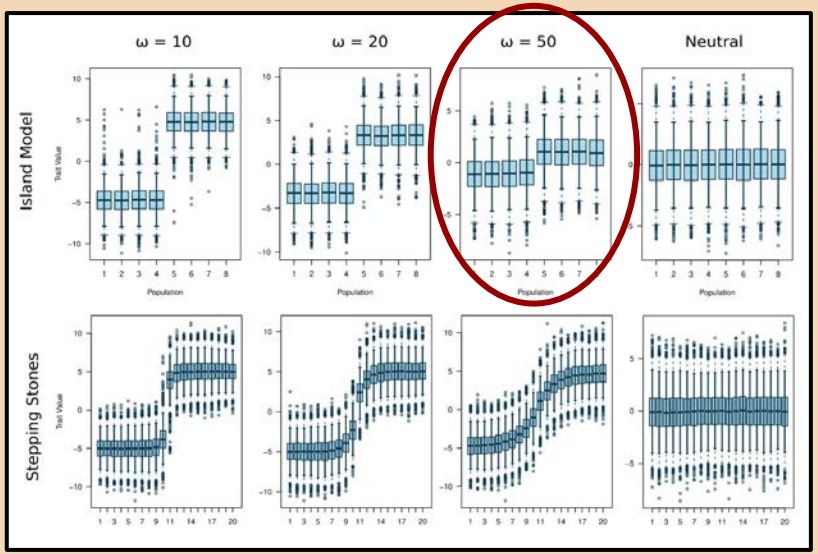


Preliminary tests:

RESULTS

Selection scenarios



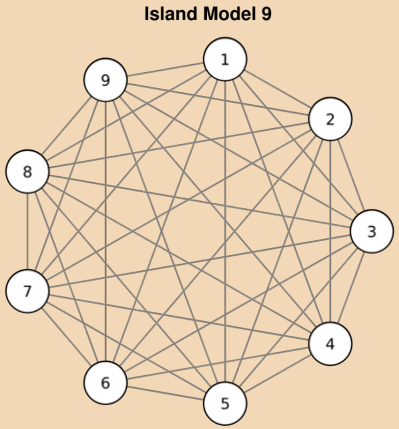


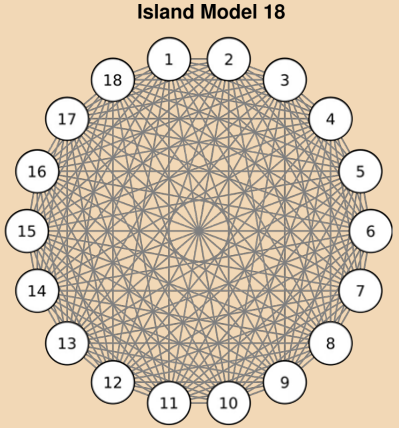
Thorough testing

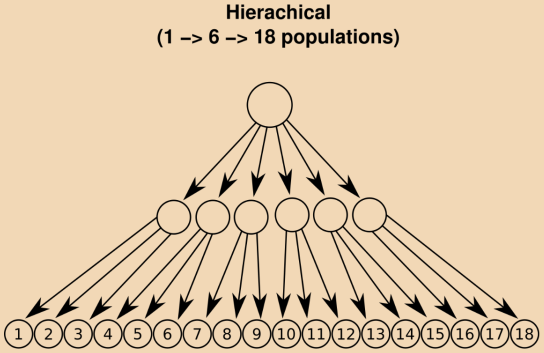
- Population structure
- Fitness variance
- Optimum difference
- Selection to migration correlation
- Fst
- Number of subpopulations

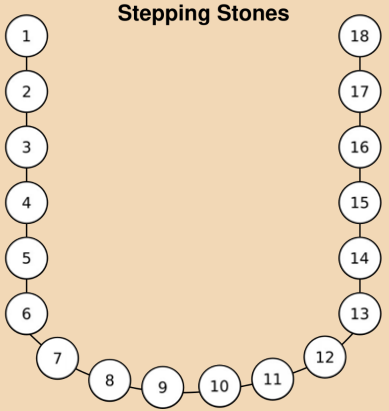
Thorough testing

- Population structure
- Fitness variance
- Optimum difference
- Selection to migration correlation
- Fst
- Number of subpopulations



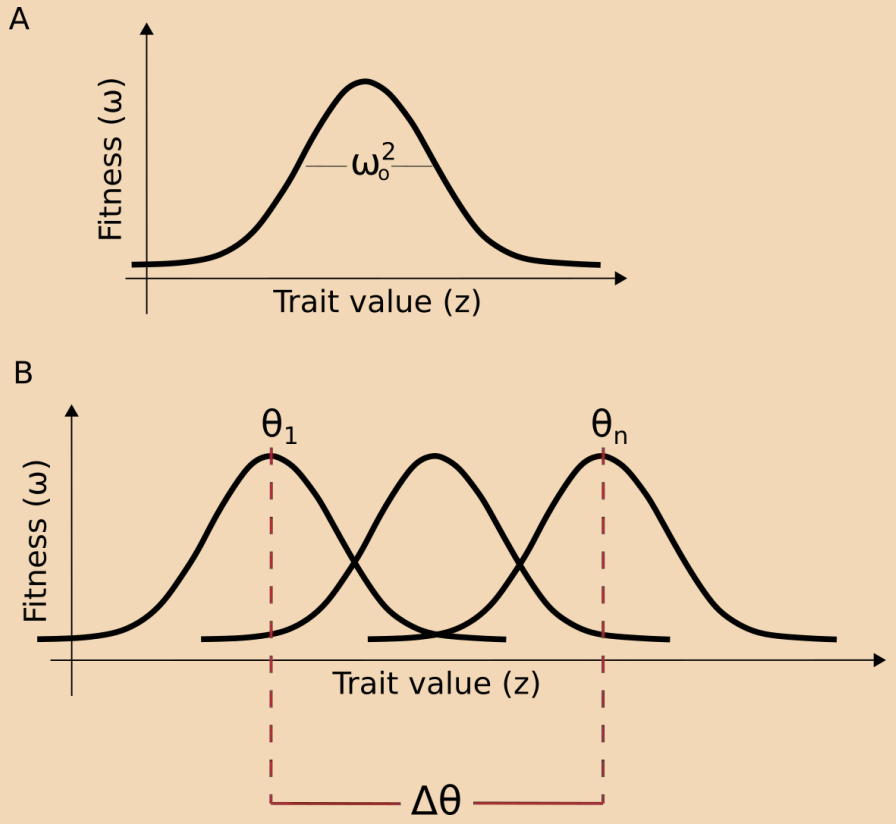






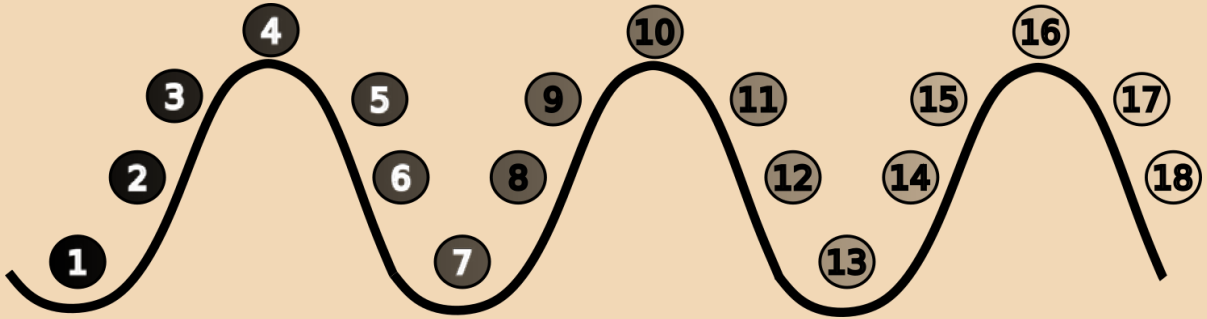
Thorough testing

- Population structure
- Fitness variance
- Optimum difference
- Selection to migration correlation
- Fst
- Number of subpopulations



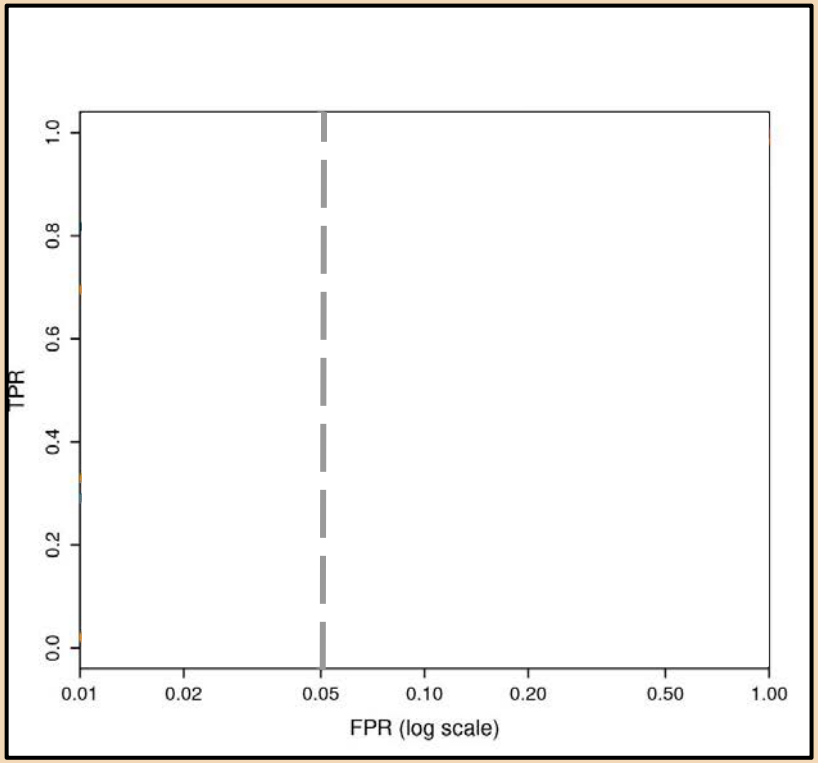
Thorough testing

- Population structure
- Fitness variance
- Optimum difference
- Selection to migration correlation
- Fst
- Number of subpopulations



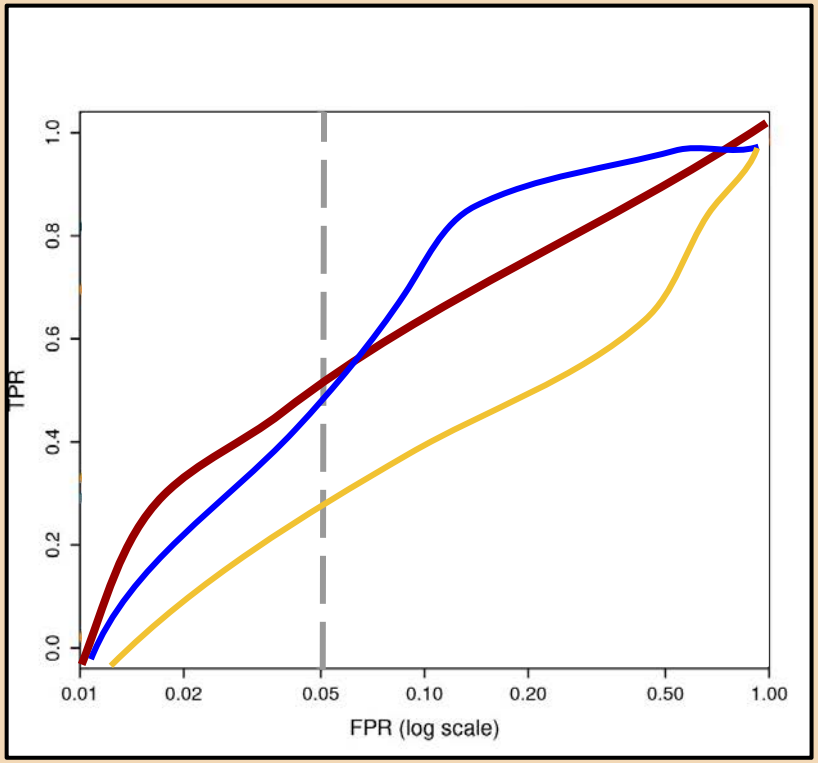
Description

True positive rate



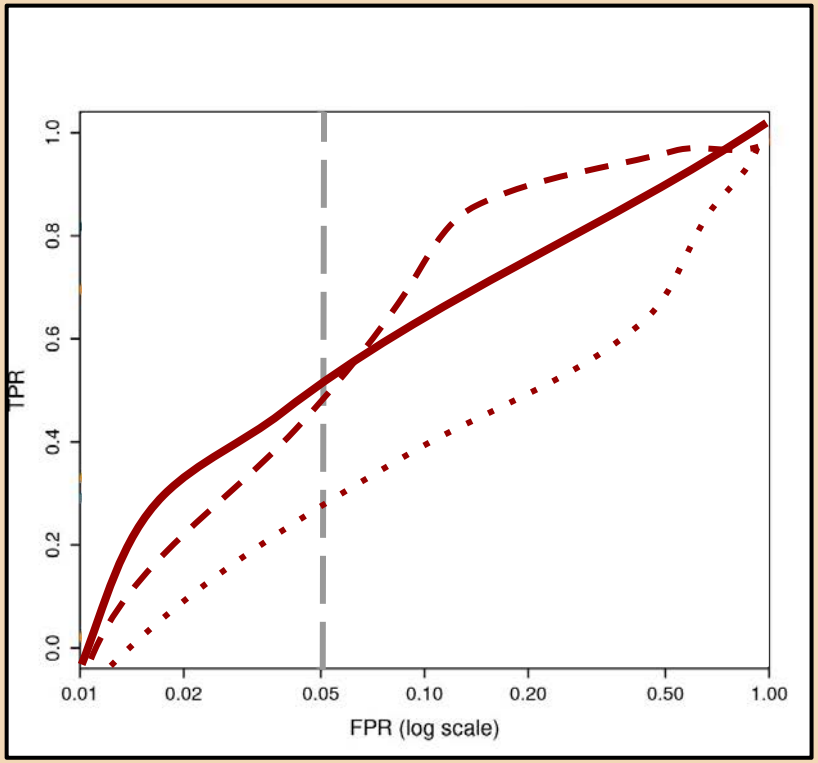
False positive rate

Description



Different methods:
Qst-Fst
Driftsel
LAVA

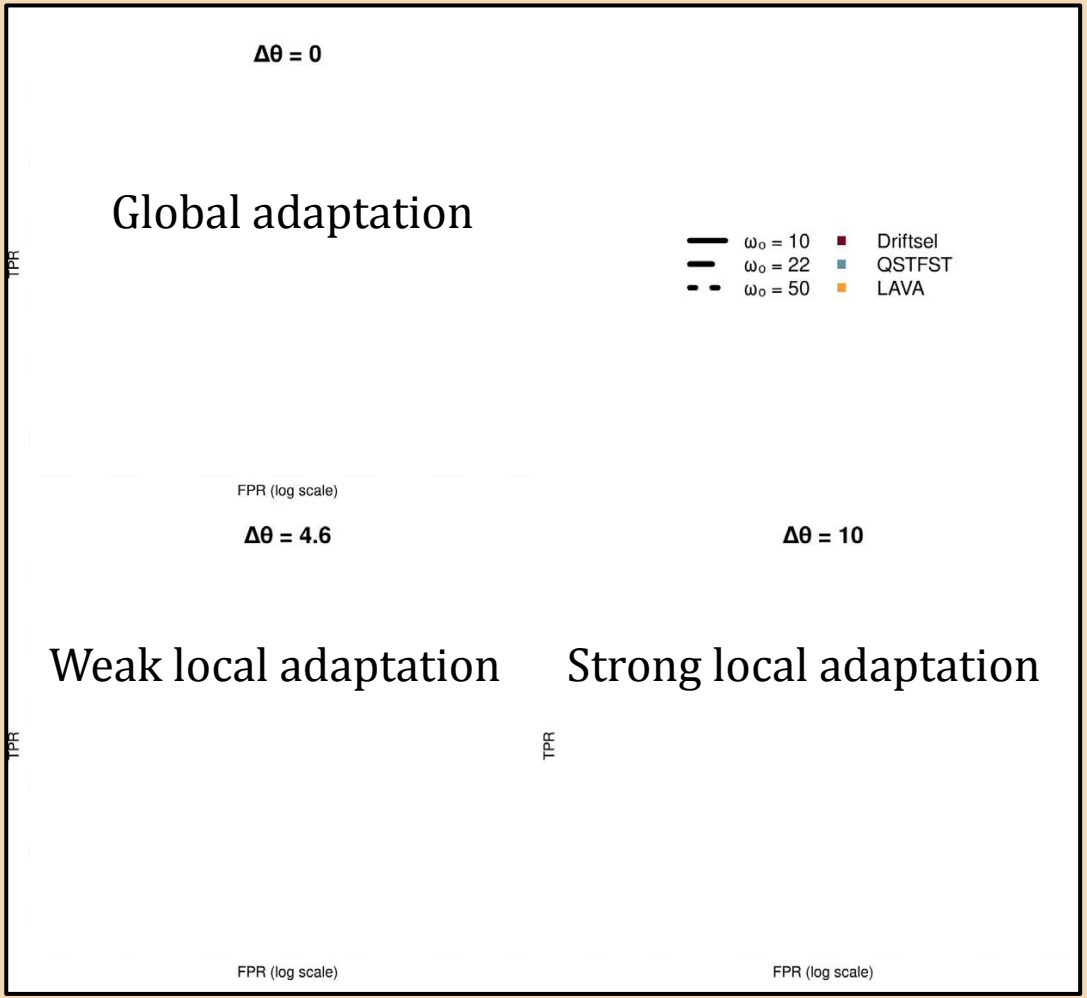
Description



Different strength of selection/ Width of the fitness function

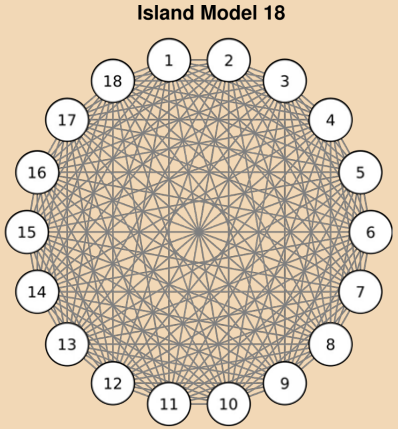
RESULTS

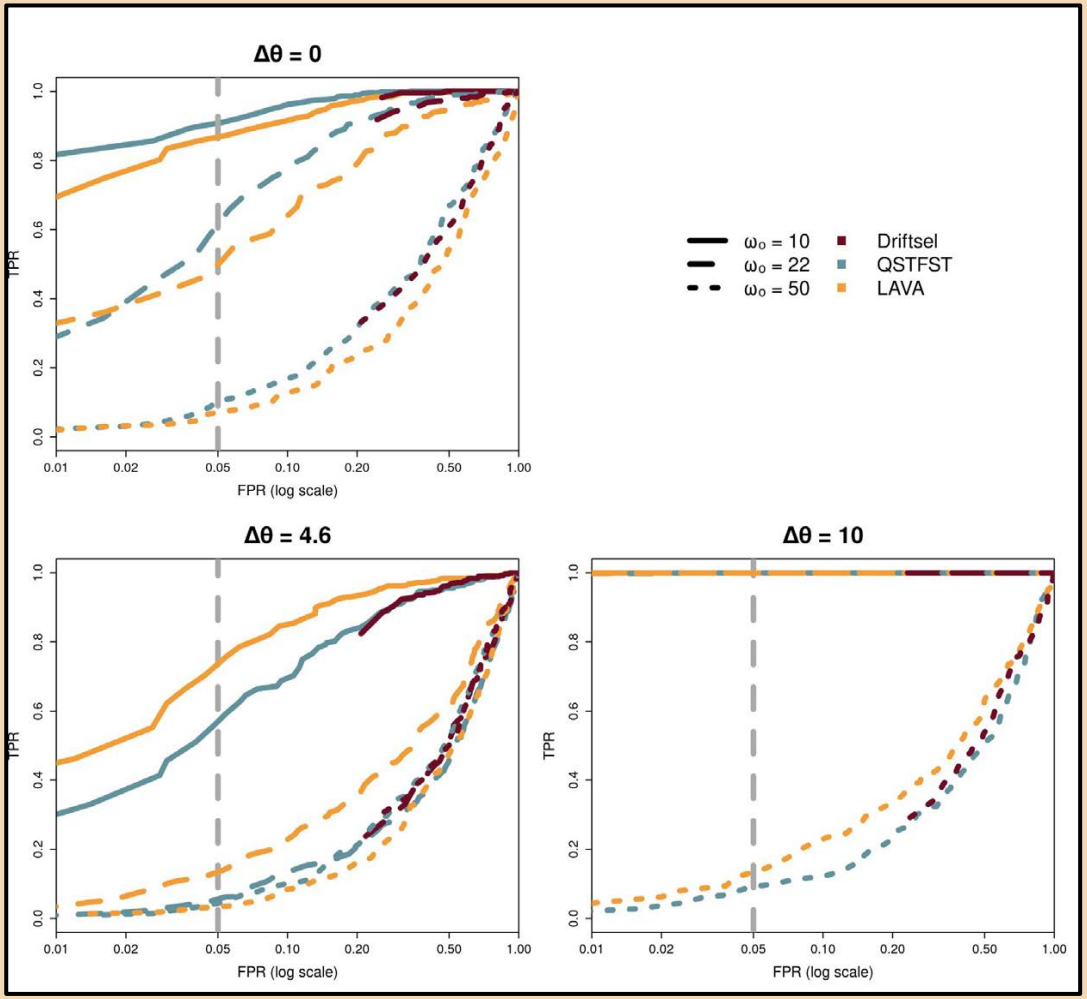
Per population structure



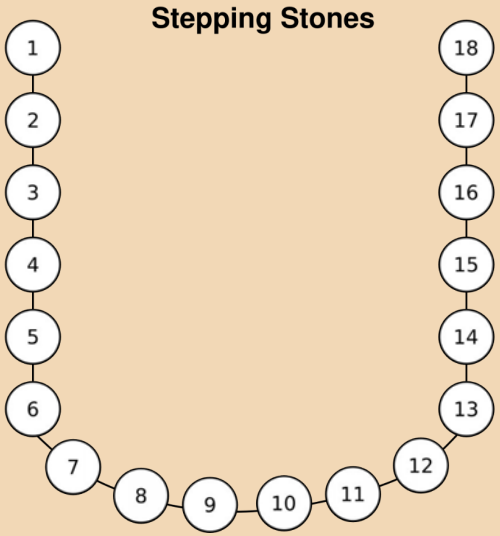
RESULTS

Island Model



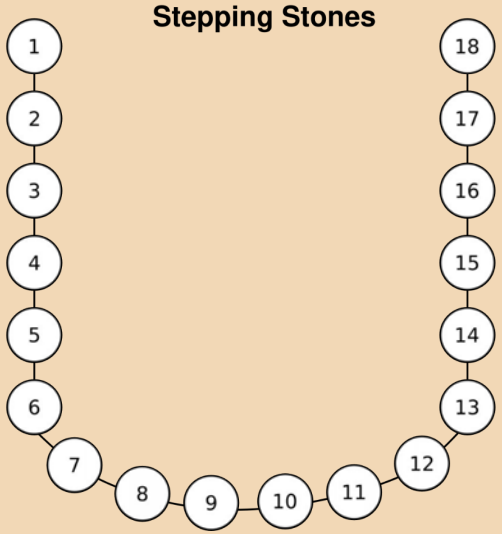


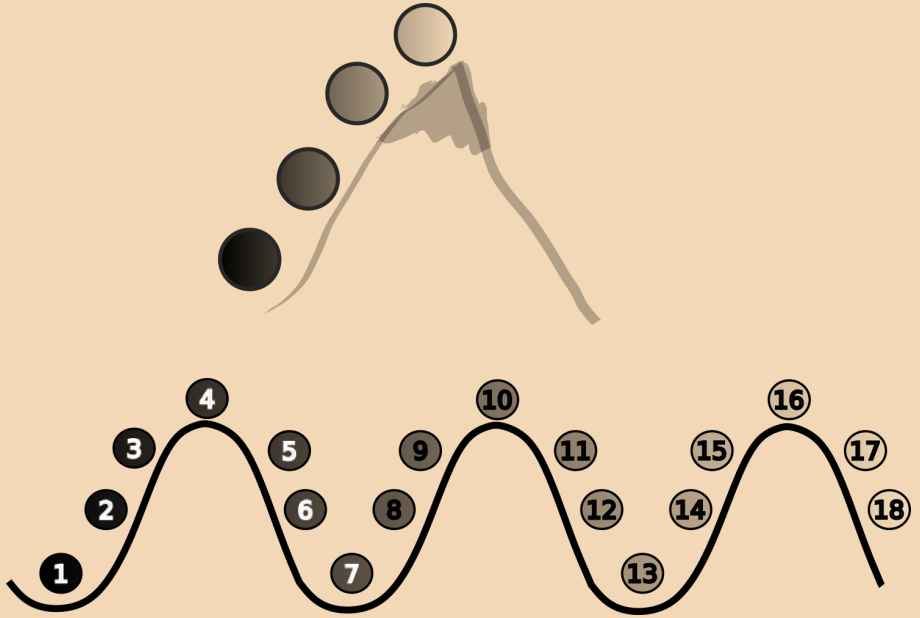
Stepping stones



Stepping stones

Correlation between selection and relatedness

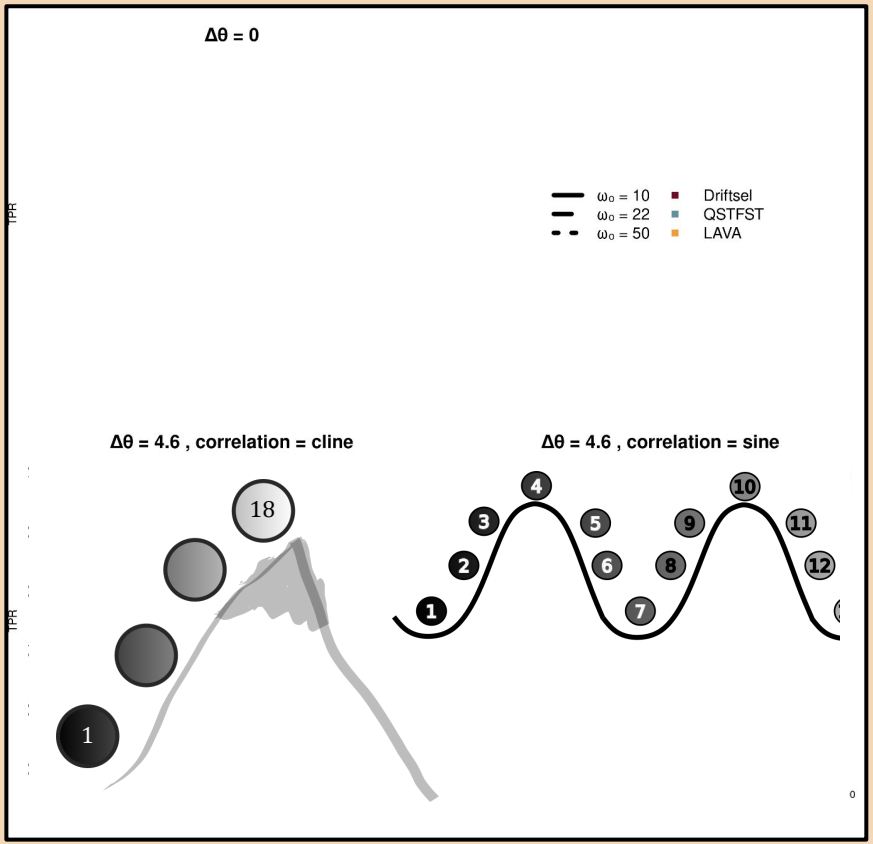


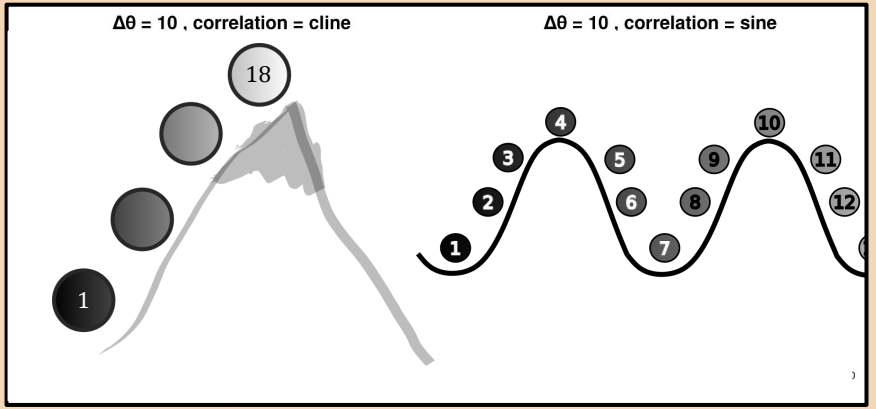


Description

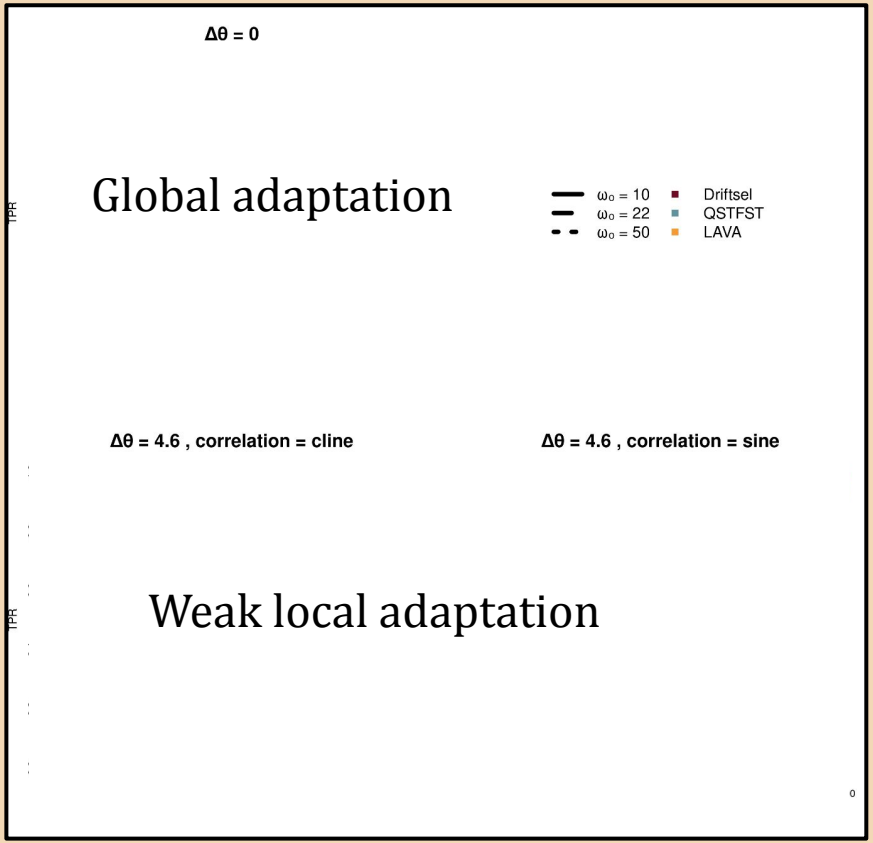
Correlation between selection and relatedness

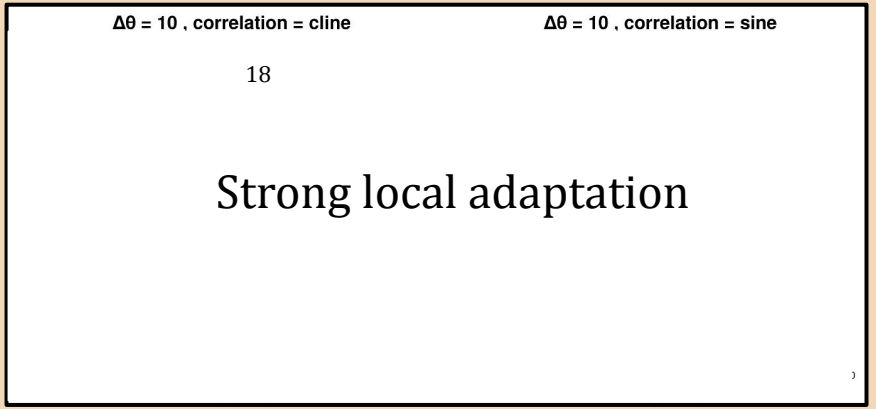
Stepping stones



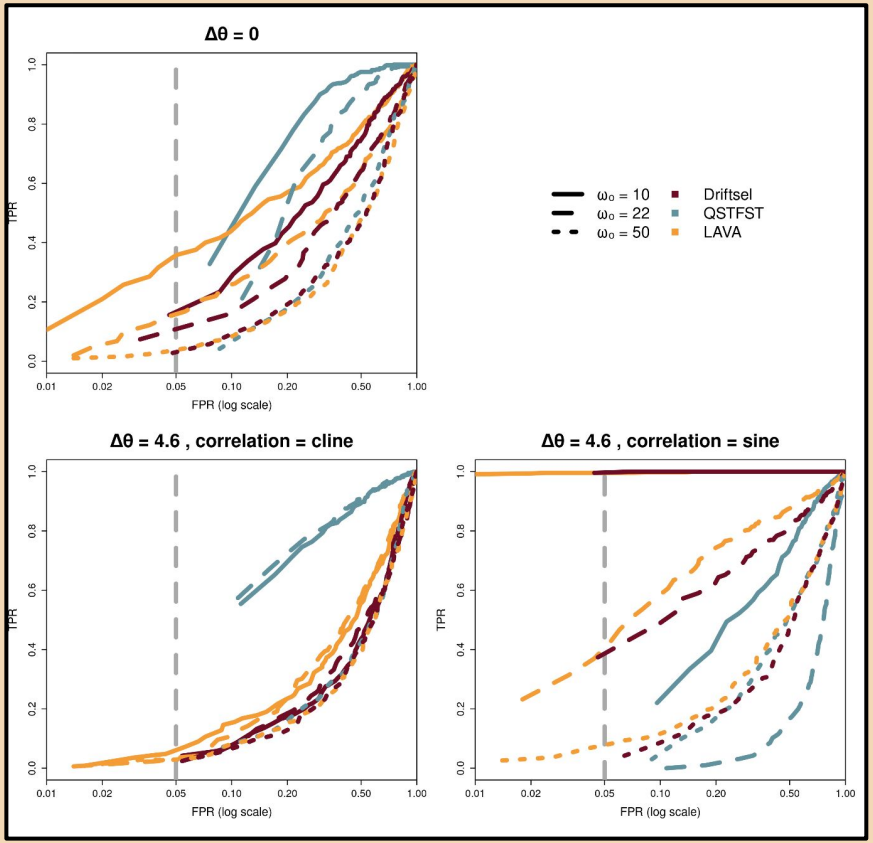


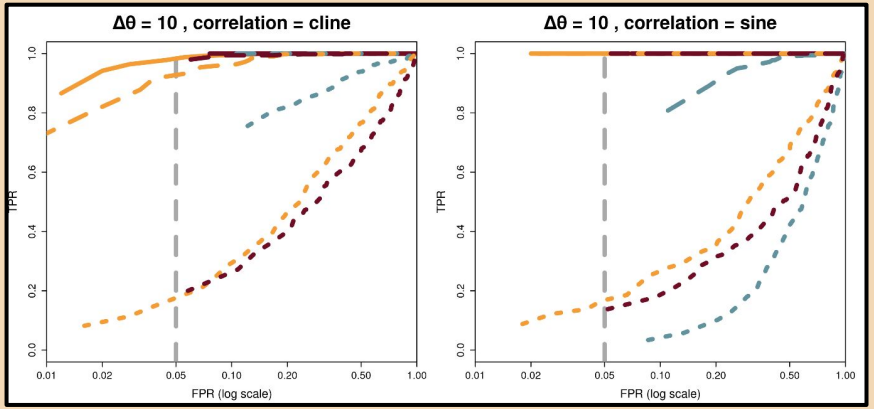
Stepping stones



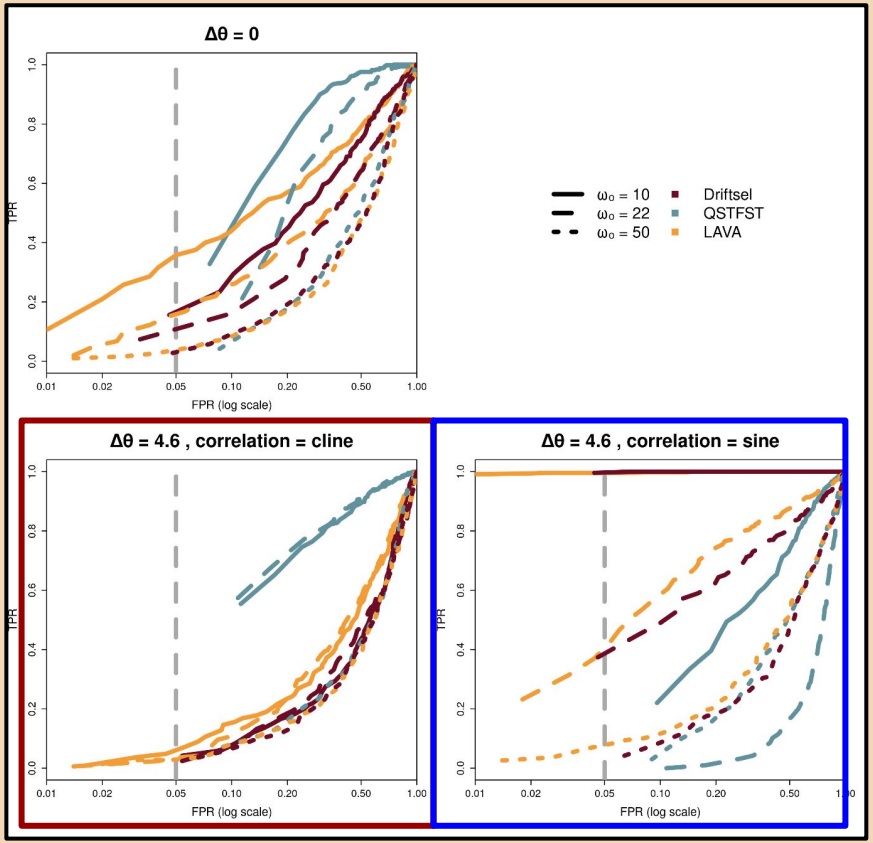


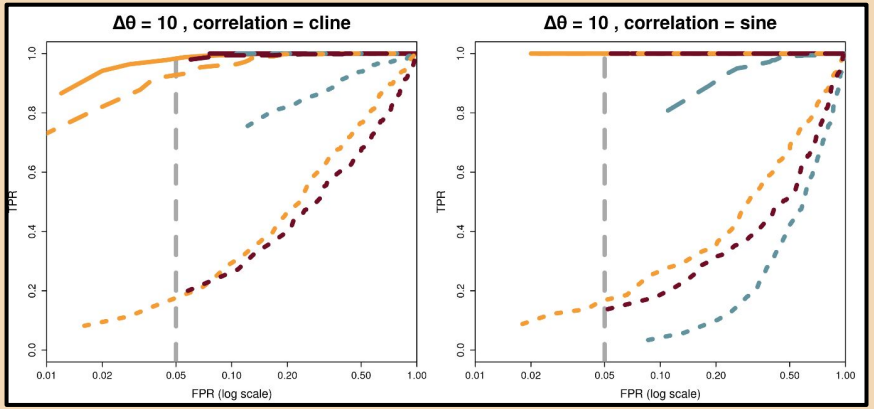
Stepping stones





Stepping stones





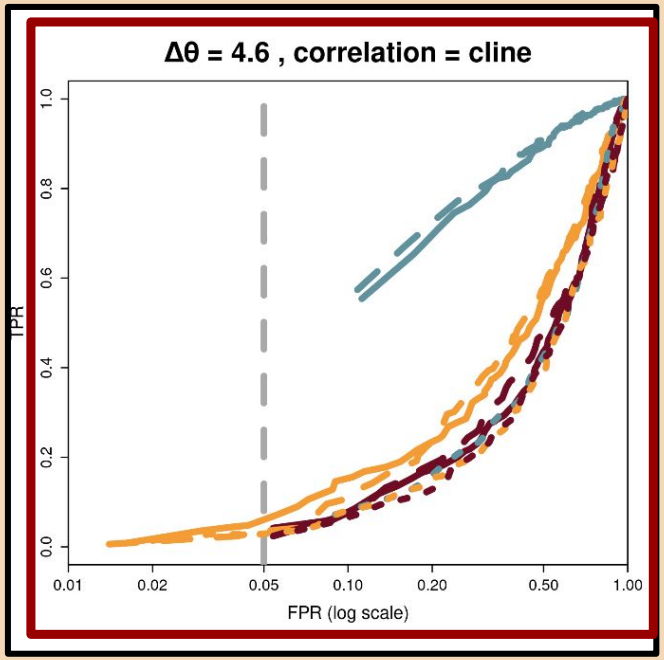
RESULTS

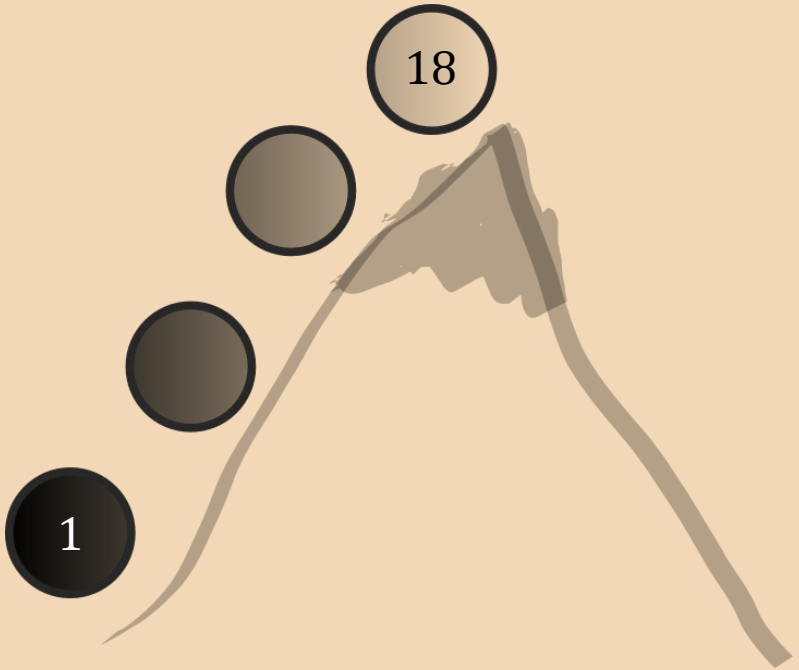
Stepping stones

ORIGINAL ARTICLE (2015)

Detecting adaptive evolution based on association with ecological gradients: Orientation matters!

E Frichot[1], SD Schoville[2], P de Villemereuil[3], OE Gaggiotti[3,4] and O François[1]





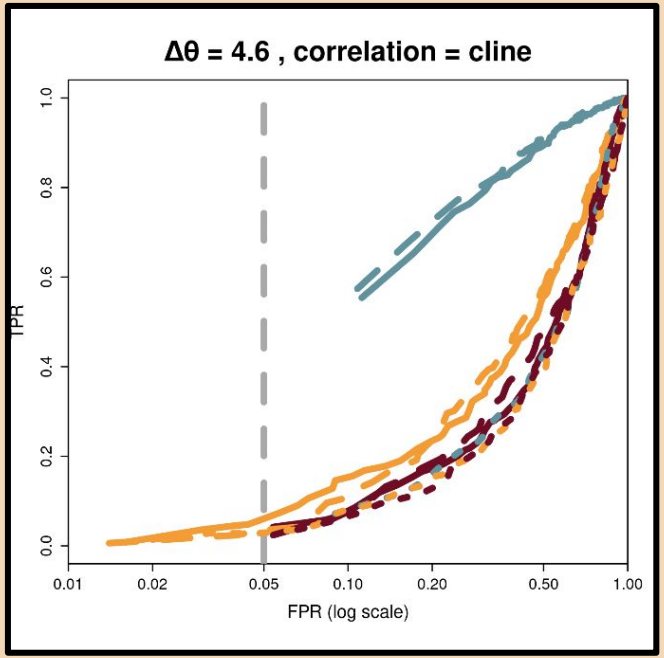
RESULTS

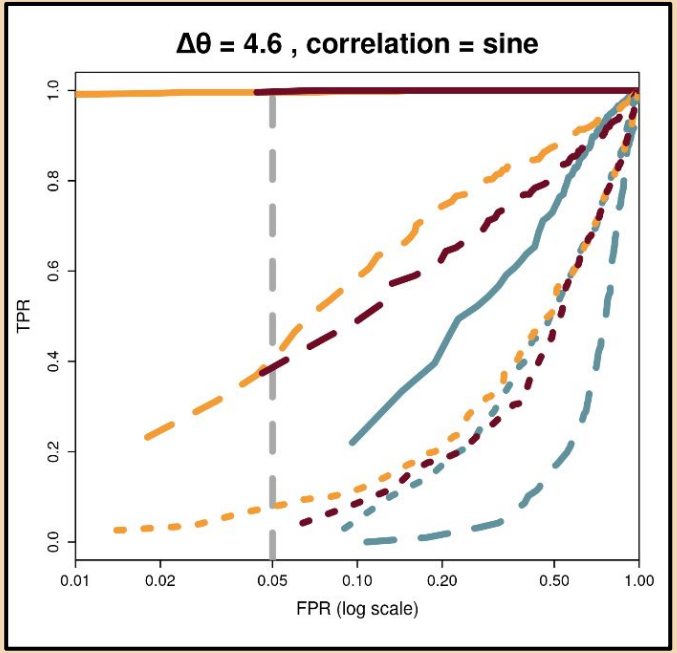
Stepping stones

ORIGINAL ARTICLE (2015)

Detecting adaptive evolution based on association with ecological gradients: Orientation matters!

E Frichot[1], SD Schoville[2], P de Villemereuil[3], OE Gaggiotti[3,4] and O François[1]



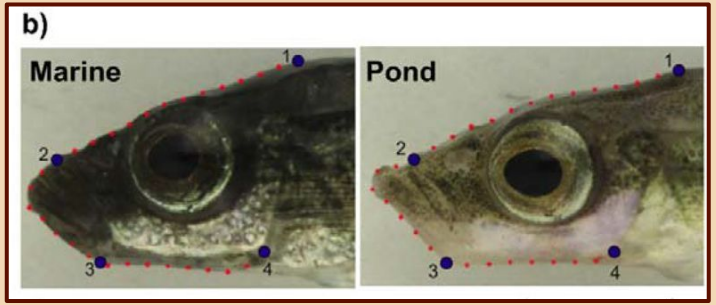


No correlation between selection and structure

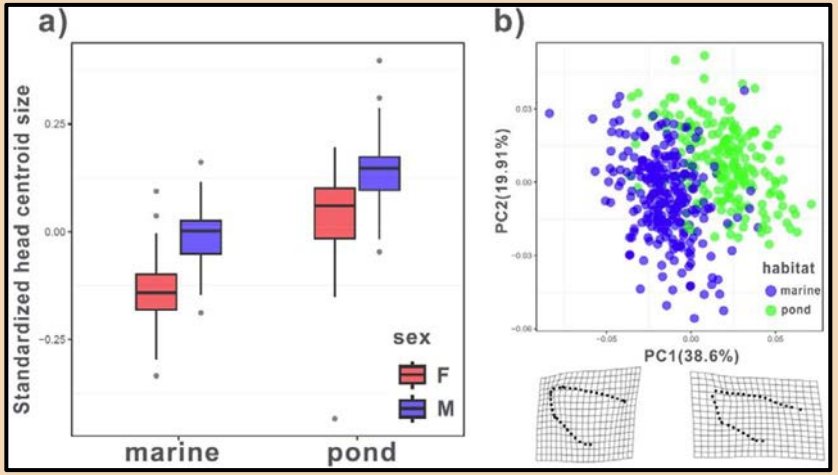
Thorough testing - Common Gardens and Sampling

- Number of sampled F1 individuals
- Number of sampled subpopulations
- Breeding design

Let’s not forget our ninespined stickleback



Driftsel: first 3 PCs, and head size when considering habitat

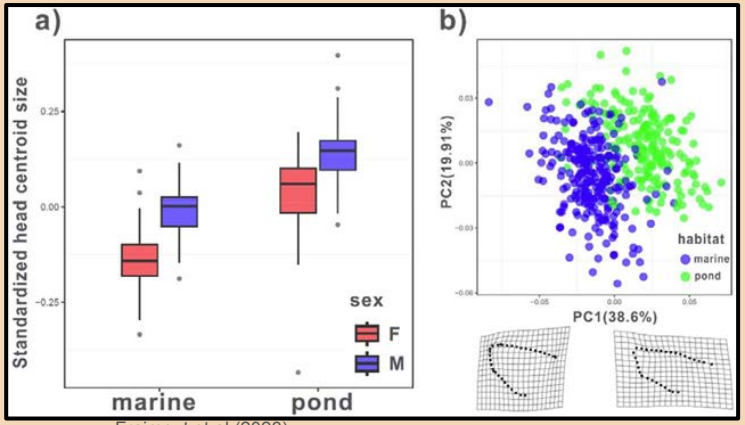


Is this trait under local adaptation?

Applying LAVA

$$Y \sim 1 + \text{sex} + (1|gr(\text{pop}, \text{cov} = 2\boldsymbol{\Theta}^{\mathbf{P}})) + (1|gr(\text{ind}, \text{cov} = \mathbf{M})$$

$$Y \sim 1 + \text{sex} + \text{hab} + (1|gr(\text{pop}, \text{cov} = 2\boldsymbol{\Theta}^{\mathbf{P}})) + (1|gr(\text{ind}, \text{cov}) = \mathbf{M}))$$



Fraimout et al (2023)

Is this trait under local adaptation?

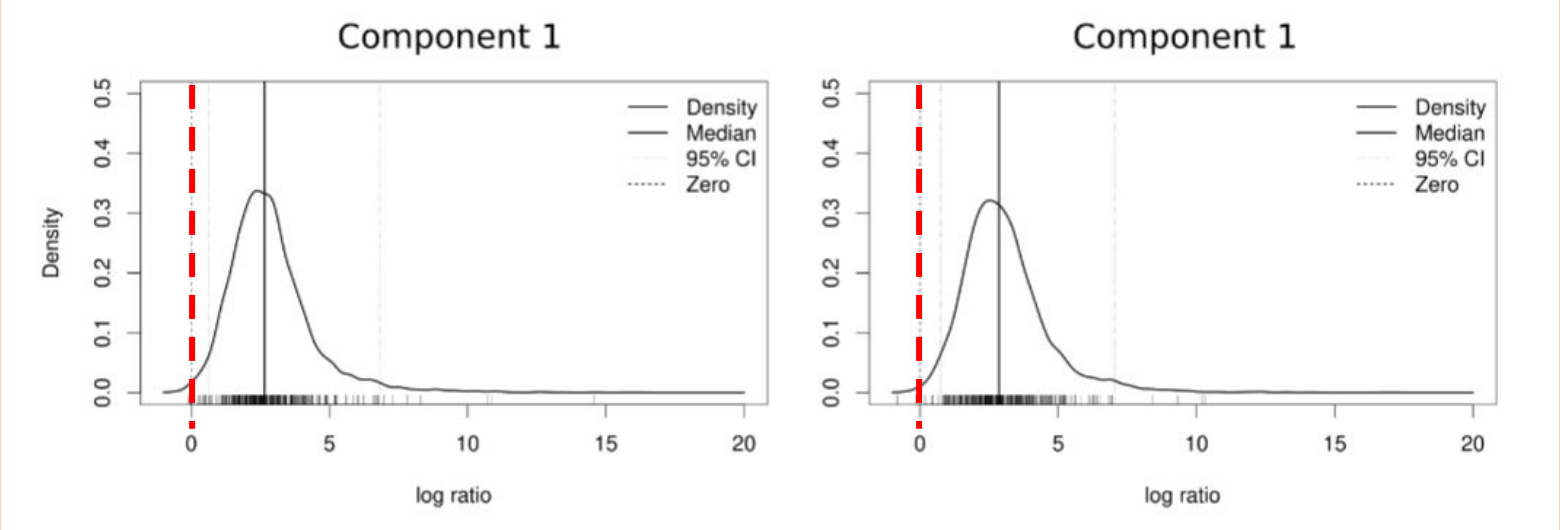
Applying LAVA

Posterior distribution of

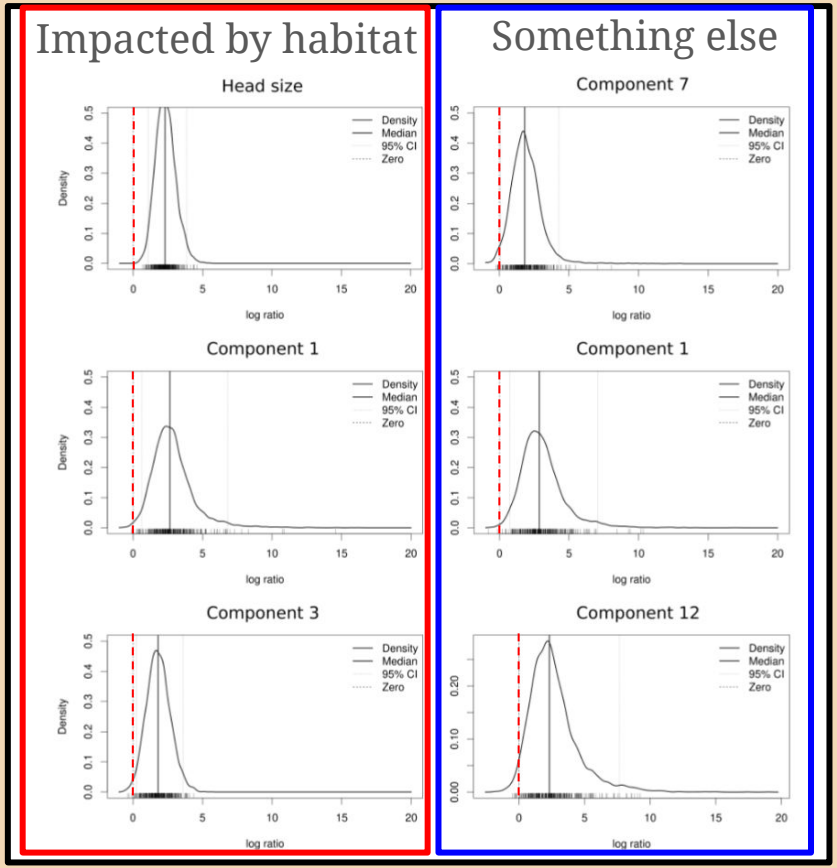
$$log\left(\frac{V_{A,W}}{V_{A,B}}\right)$$

Without habitat

With habitat



Is this trait under local adaptation?



Driftsel: first 3 PCs, and head size when considering habitat

LAVA: PC 1, 7, 12.
Habitat impacts: head size, 1, 3.

Conclusion:

- Populations structure matters,
- Detecting local adaptation is challenging,
- Detecting global adaptation is even harder!

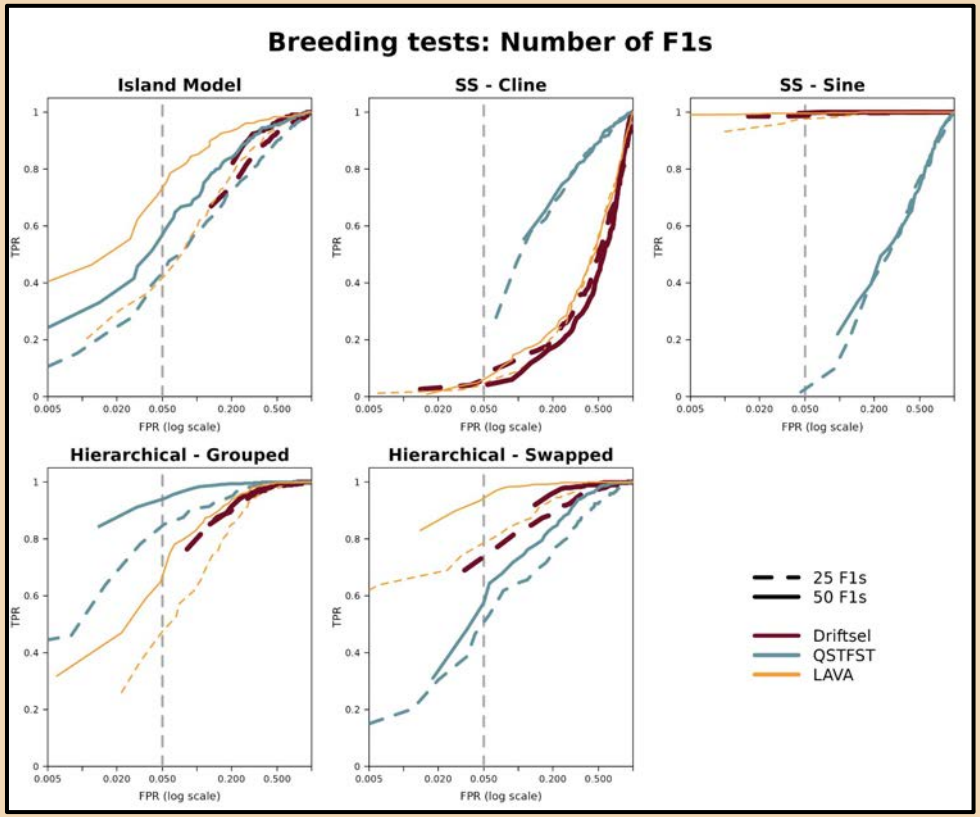
Coming soon:

Thank **you**!

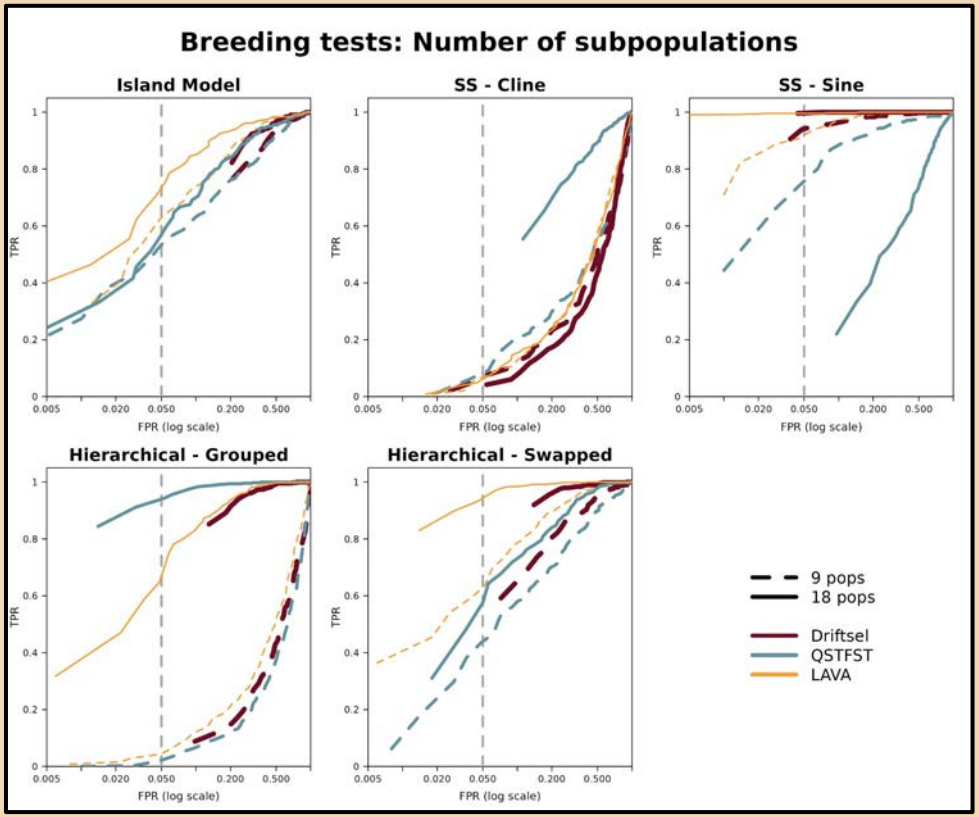
UNIL | Université de Lausanne



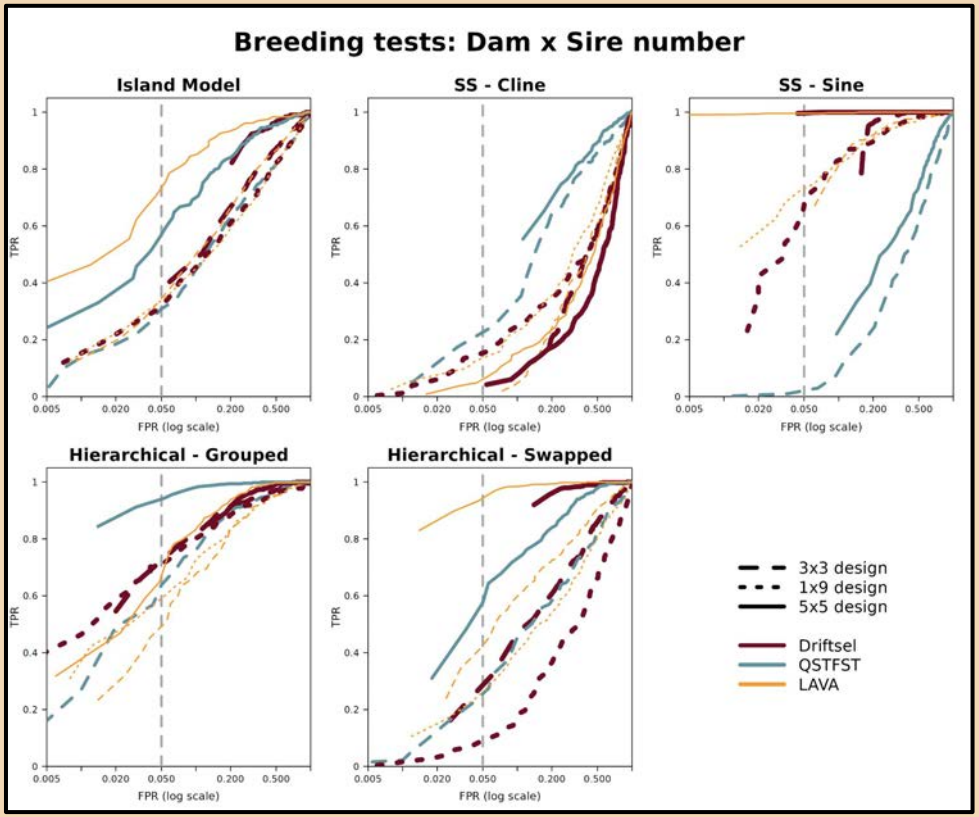

Thorough testing - Common Gardens and Sampling



Thorough testing - Common Gardens and Sampling



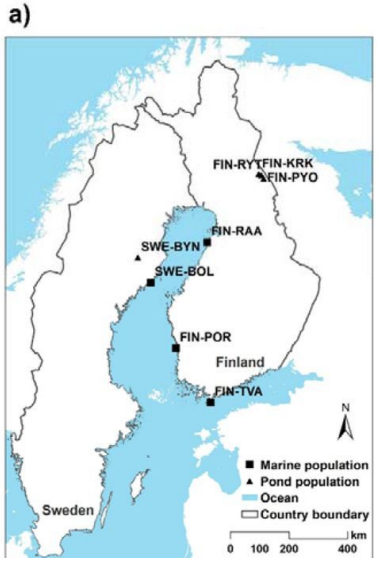
Thorough testing - Common Gardens and Sampling

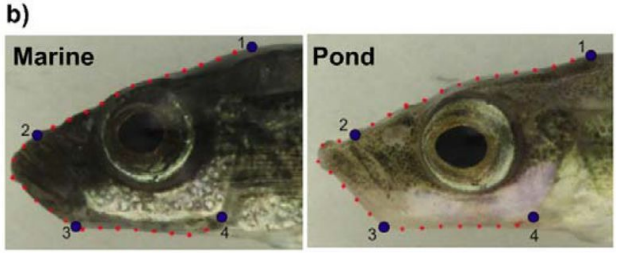


Results

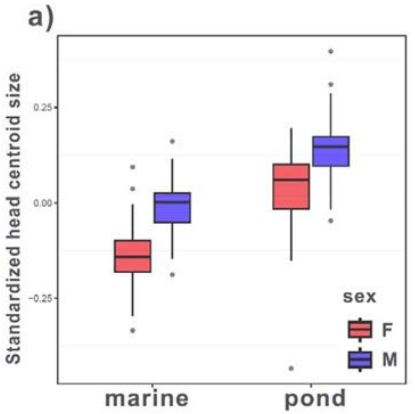
Breeding results

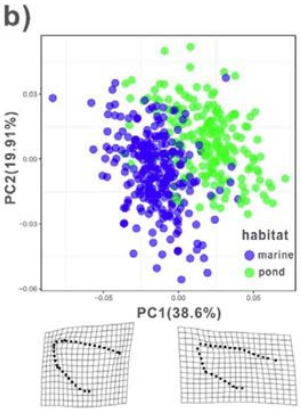
Is this trait under local adaptation?

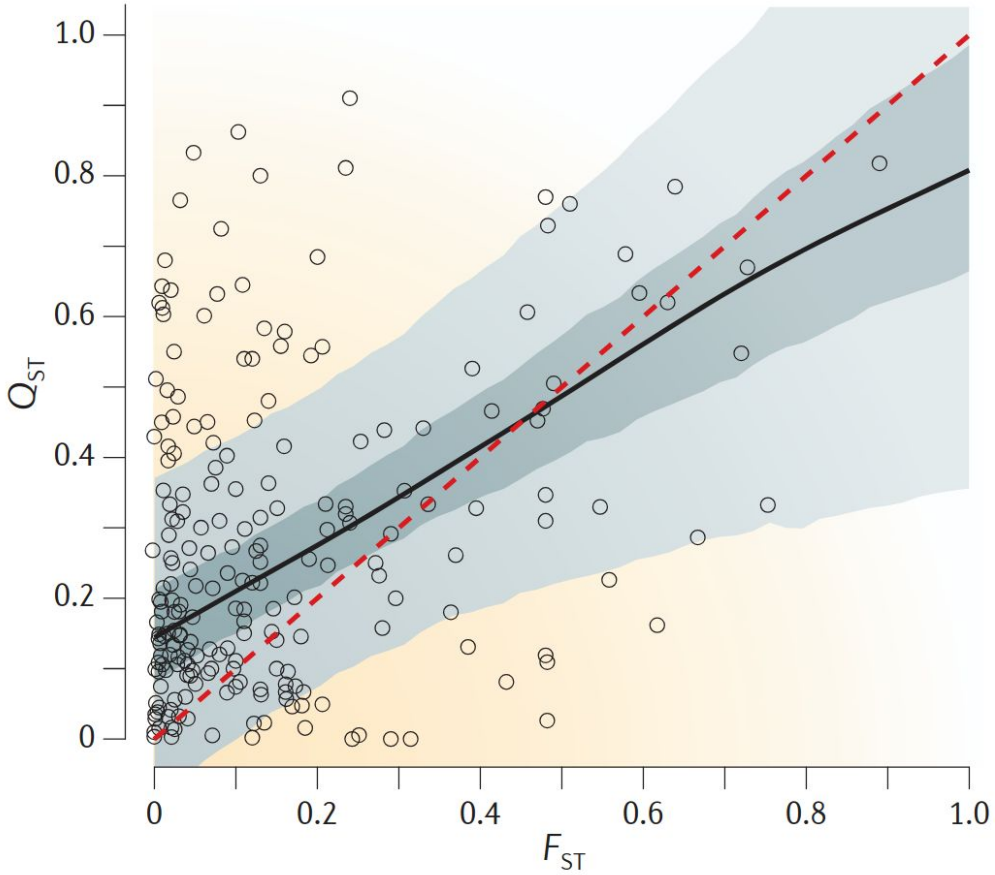




Fraimout et al (2023)






1.0
0.8
0.6
0.4
0.2
0
Q_{ST}
0
0.2
0.4
0.6
0.8
1.0
F_{ST}

Nearly every plant or animal species includes many partially isolated populations. As a result of genetic drift or divergent natural selection, such populations become genetically differentiated over time

… including humans

(Holsinger and Weir 2009)

only 5–10% of human genetic diversity is accounted for by genetic differences among populations from major geographical regions

F_{ST} is directly related to the variance in allele frequency among populations and, conversely, to the degree of resemblance among individuals within populations. If F_{ST} is small, it means that the allele frequencies within each population are similar; if it is large, it means that the allele frequencies are different.

$$x_{11,1} = p_1^2 + f_1 p_1 (1 - p_1)$$
$$x_{12,1} = 2 p_1 (1 - p_1)(1 - f_1)$$
$$x_{22,1} = (1 - p_1)^2 + f_1 p_1 (1 - p_1)$$

$$x_{11} = \pi^2 + F\pi(1 - \pi)$$
$$x_{12} = 2\pi(1 - \pi)(1 - F)$$
$$x_{22} = (1 - \pi)^2 + F\pi(1 - \pi)$$

in which σ_π^2 is the variance in allele frequency among populations. $\pi(1 - \pi)$ is the variance in the allelic state for

$$\theta = \frac{\sigma_\pi^2}{\pi(1 - \pi)}$$

When we go from f to theta matrix

belong. θ and F_{ST} can be thought of either as the proportion of genetic diversity due to allele frequency differences among populations or as the correlations between alleles within populations relative to the entire population. F and F_{IT} can be thought of either as the depar-

However, the degree of genetic differentiation among subpopulations also depends on the strength and nature (for example, diversifying or balancing selection) of the predominant selective pressures experienced by the populations or demes under study. In the case of adaptive population divergence, directional selection is expected to increase F_{ST} of selected or linked loci, relative to that of neutral loci[4,5]. Yet, because most quantitative traits of evolutionary, ecological, economic and even of medical interest — such as body size and intelligence quotient — are known or thought to have a polygenic basis[6,7], distinguishing neutral and selective patterns of population differentiation at the phenotypic level is not easily accomplished with standard F_{ST} estimates. Trait-based inference, however, can be accomplished under a related analytical framework.

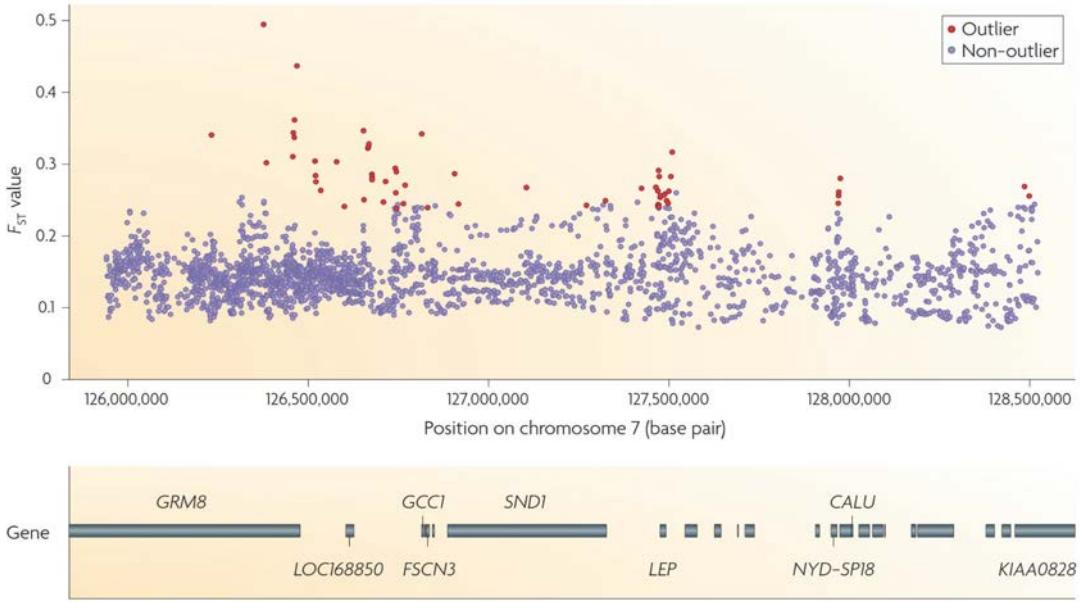


Figure 1 | **Locus-specific estimates of F_{ST} on human chromosome 7.** Estimates are as inferred from the phase II HapMap data set[95]. Horizontal bars indicate the locations of known genes. The red circles are posterior means for SNPs with estimates that are detectably different from the genomic background (purple circles). All 'outliers' show significantly more differentiation among the four populations in the sample than is consistent with the level of differentiation seen in the genomic background. The excess differentiation suggests that these SNPs are associated with genomic regions in which loci have been subject to diversifying selection among populations. *CALU*, calumenin; *FSCN3*, ascin homolog 3; *GCC1*, GRIP and coiled-coil domain containing 1; *GRM8*, glutamate receptor, metabotropic 8; *LEP*, leptin; *SND1*, staphylococcal nuclease and tudor domain containing 1. Figure is modified, with permission, from REF. 8 © (2009) American Statistical Association.

Q_{ST} *and polygenic variation.* Spitze[42] noted that another quantity analogous to θ can be estimated for continuously varying traits. Specifically, we can define:

$$Q_{ST} = \frac{\sigma^2_{GP}}{\sigma^2_{GP} + 2\sigma^2_{GI}} \qquad (6)$$

$$Q_{ST} = \frac{V_B}{V_B + 2V_W}$$

QST is the
proportion of additive genetic variance in a trait that
is due to among-population differences. If the trait is
selectively neutral, if all genetic variation is additive
and if the mutation rates at loci contributing to the
trait are the same as those at other loci, we expect QST
and FST to be equal 43,44 . Comparing the magnitude of
QST and FST may therefore indicate whether a particu-
lar trait has been subject to stabilizing selection (QST
<FST
)
or diversifying selection (QST >FST). However, because
of the uncertainties associated with estimates of QST
and FST
, such comparisons are likely to be useful only
when they are available for a moderately large number
of populations (>20)45 . Furthermore, caution is neces-
sary when suggesting that a comparison of QST and FST
provides evidence for stabilizing selection because non-
additive genetic variation tends to change QST , even for
a neutral trait