L5. Quantitative population genetics

OVERVIEW

. L1. Approaches to ecological modelling

. L2. Model parameterization and validation

. L3. Stochastic models of population dynamics (math)

. L4. Animal movement (math + stat)

. L5. Quantitative population genetics (math + stat)

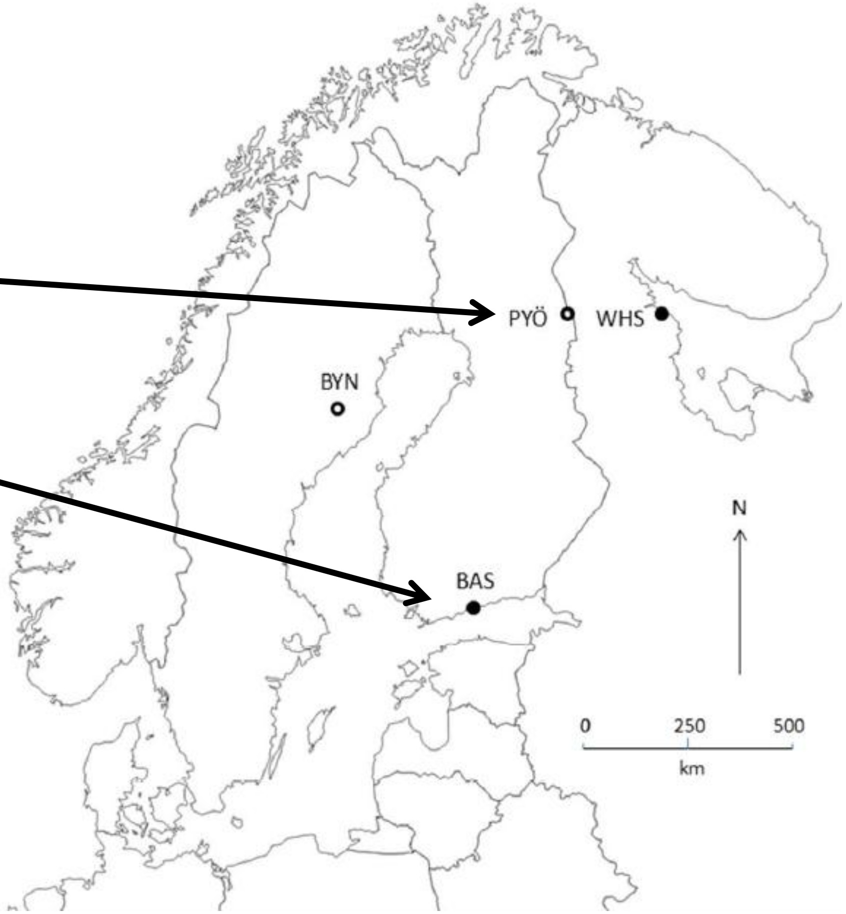
. L6. Community ecology (stat)

There is variation among individuals

This nine-spined stickleback originates from Pond Pyöreälampi population

This nine-spined stickleback originates from the Baltic Sea population

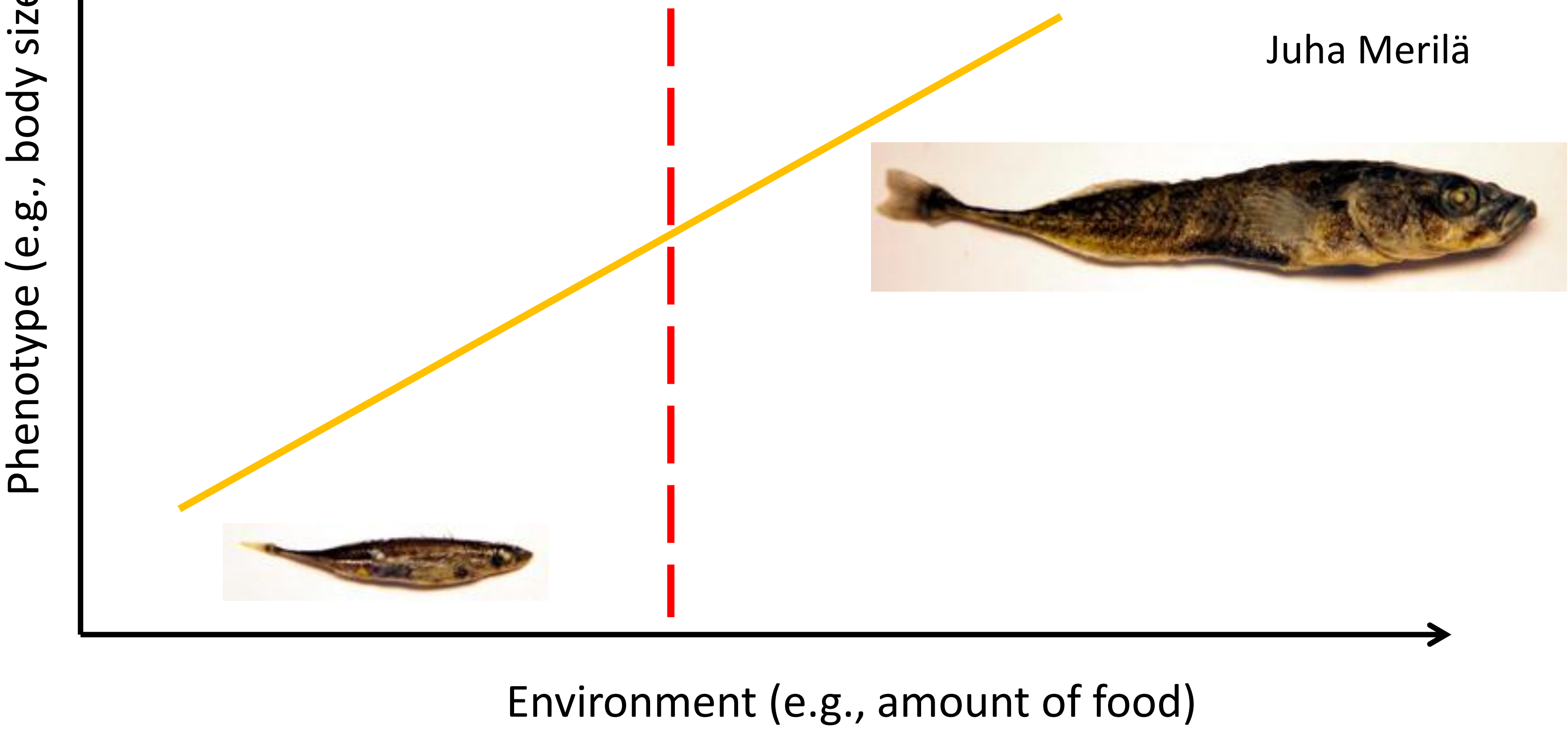
Why are they different?



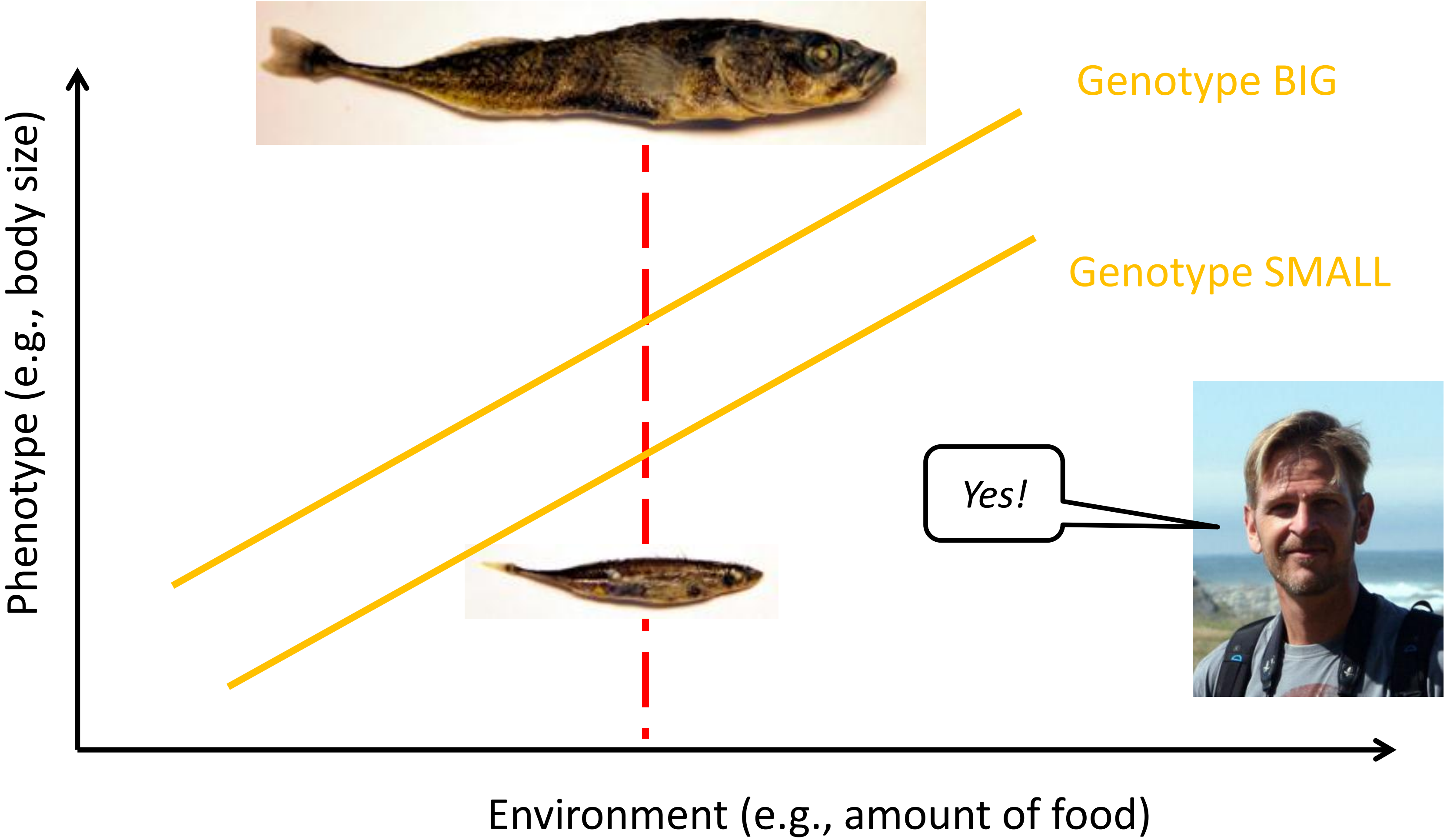
Environmental effect?

Juha Merilä

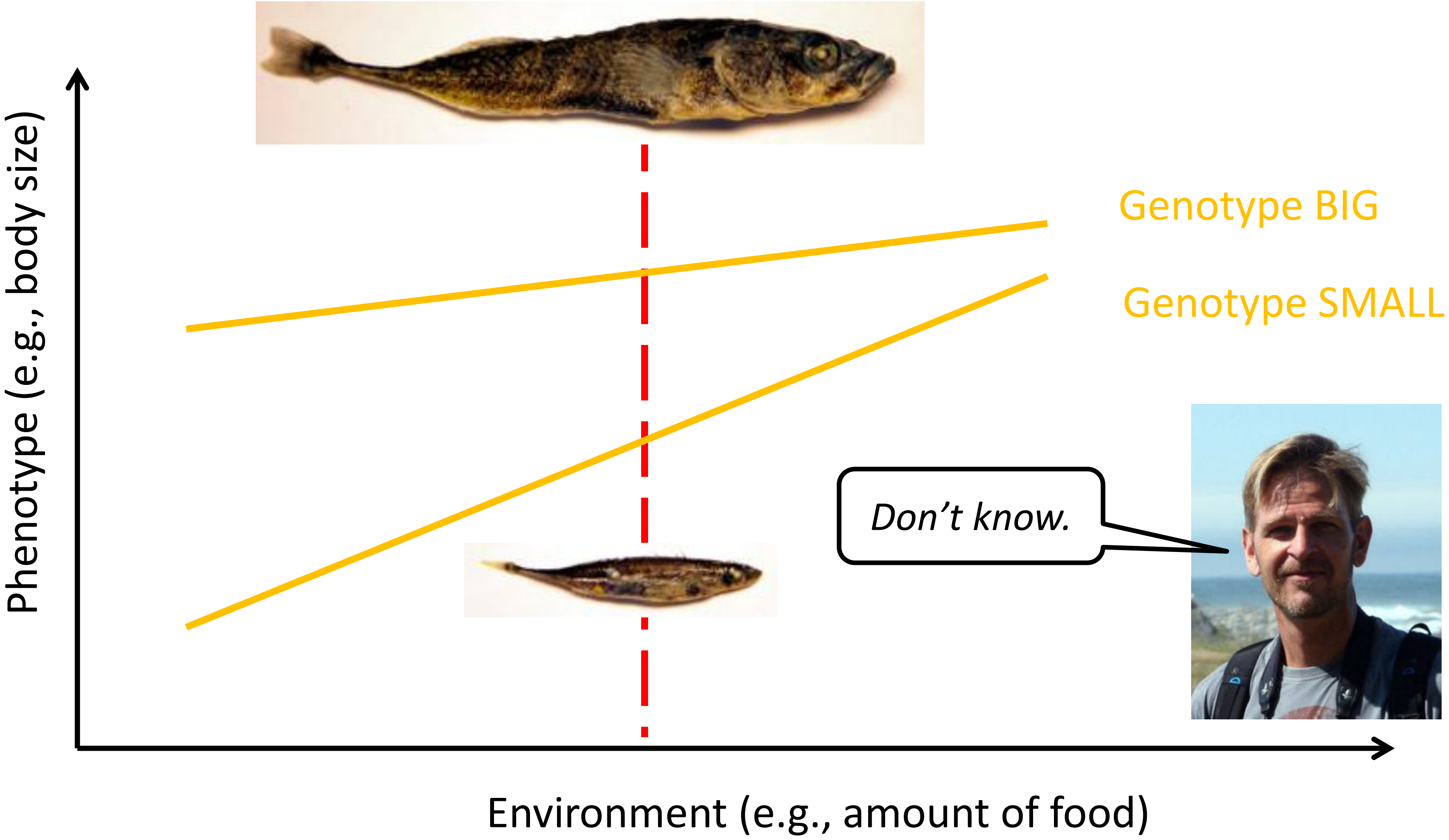
No! We did a common garden experiment.



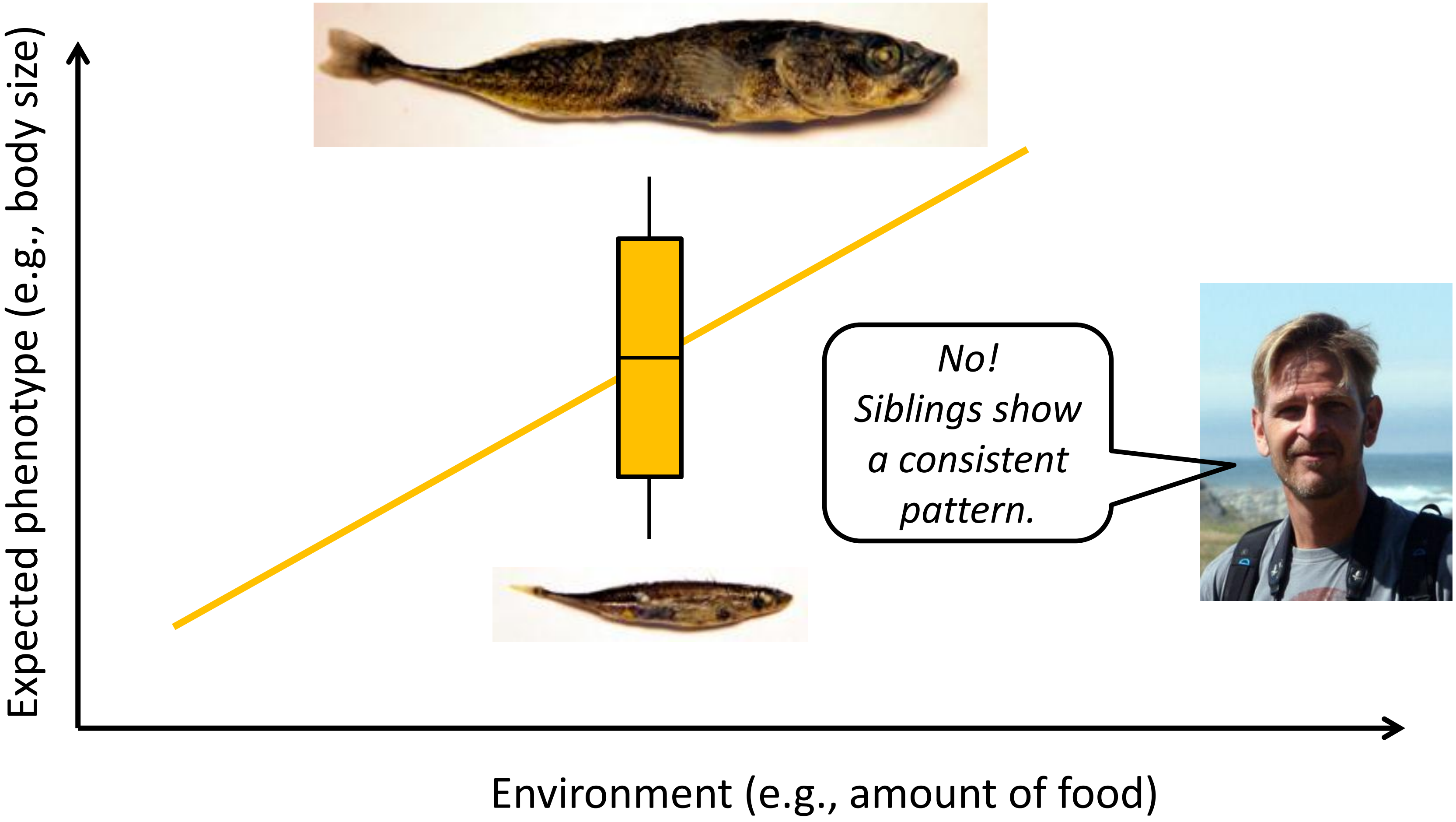
Environmental + genetic effect?



Environmental | genetic effect?

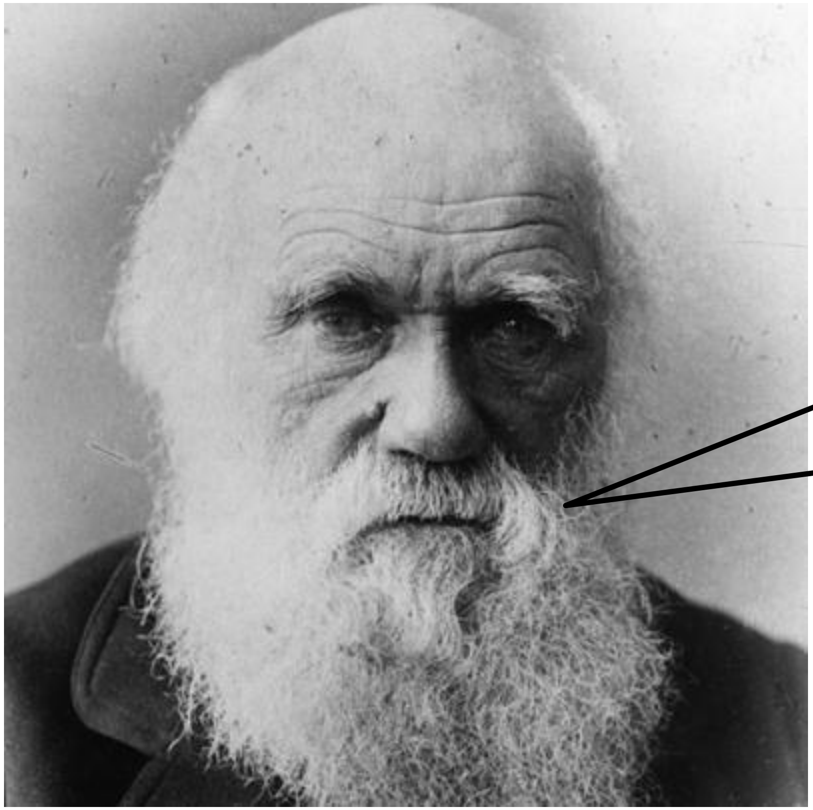


Developmental instability?



No!
Siblings show
a consistent
pattern.

Why is there a genetic difference?



Natural selection.

Survival of the fittest!

Charles Darwin(1809-1882)

Natural selection?

Juha Merilä

The fish are different because of local adaptation.

Small fish are better in escaping predators (important in sea),
big fish are better in competing for food (important in ponds).

Not only natural selection?

Sewall Wright (1889-1988)

Mutation, migration and genetic drift.

Adaptive landscapes.

Genetic drift?

Juha Merilä

Maybe the fish are different just by chance?

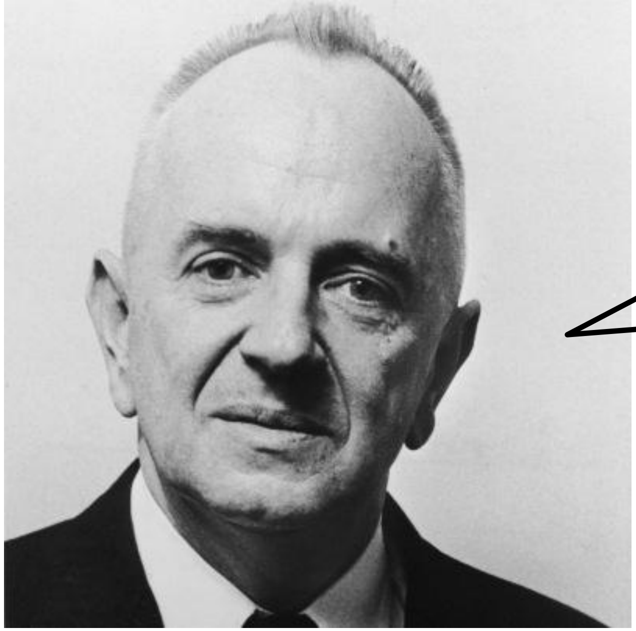
Could the difference be generated by the random assortment of genes from parents to offspring?

Is drift a plausible hypothesis?

Ronald Fisher (1890-1962)

Most natural populations are too large for drift to be important.

Is drift a plausible hypothesis?



Theodosius Dobzhansky (1900-1975)

Drift can still be important in genetically isolated sub-populations.

Genetic drift?

Juha Merilä

I am studying genetically isolated sub-populations, so drift could be a plausible explanation…

To separate drift and selection, we need quantitative tools

Russel Lande

Quantitative genetics for evolutionary biology of natural populations.

How does one apply quantitative genetics theory in practice?

Juha Merilä

Hire a sta

PhD stud

Markku Karhunen

Statistical methods for detecting signals of natural selection in the wild

(PhD dissertation in Helsinki, 18th October 2013)

Markku Karhunen

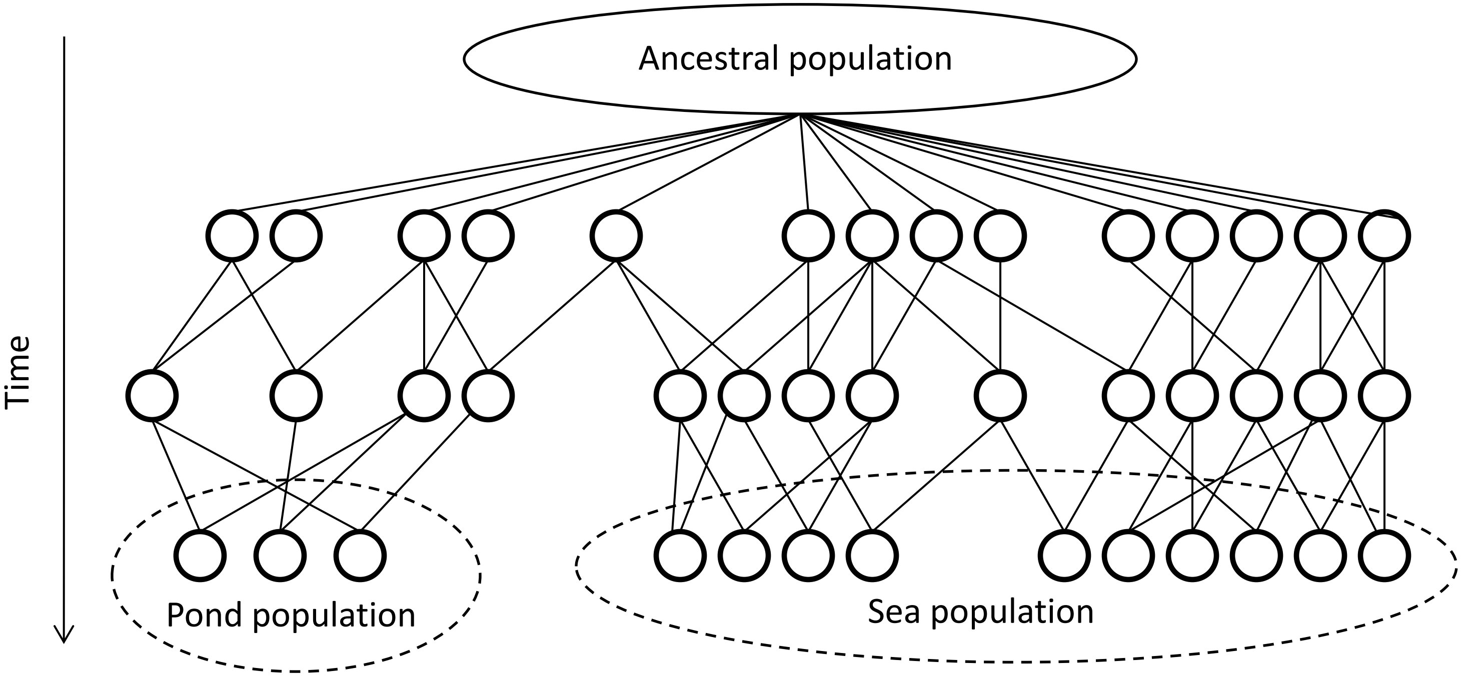
I. Ovaskainen, Karhunen, Zheng, Cano Arias and Merilä, *Genetics 2011*

II. Karhunen and Ovaskainen, *Genetics 2012*

III. Karhunen, Merilä, Leinonen, Cano Arias and Ovaskainen, *Molecular Ecology Resources 2013*

IV. Karhunen, Ovaskainen, Herczeg and Merilä, *Evolution (in press)*

Genetic differentiation by drift or selection?


Time

Ancestral population

Pond population

Sea population

Null model: what happens under neutrality?

Related individuals resemble each other:

Breeding values of individuals *i* and *j*

Amount of additive variance

$$\mathrm{Cov}[\mathbf{a}_i, \mathbf{a}_j] = 2\theta_{ij}\mathbf{G}$$

Coancestry (relatedness) between individuals *i* and *j*

Related populations resemble each other:

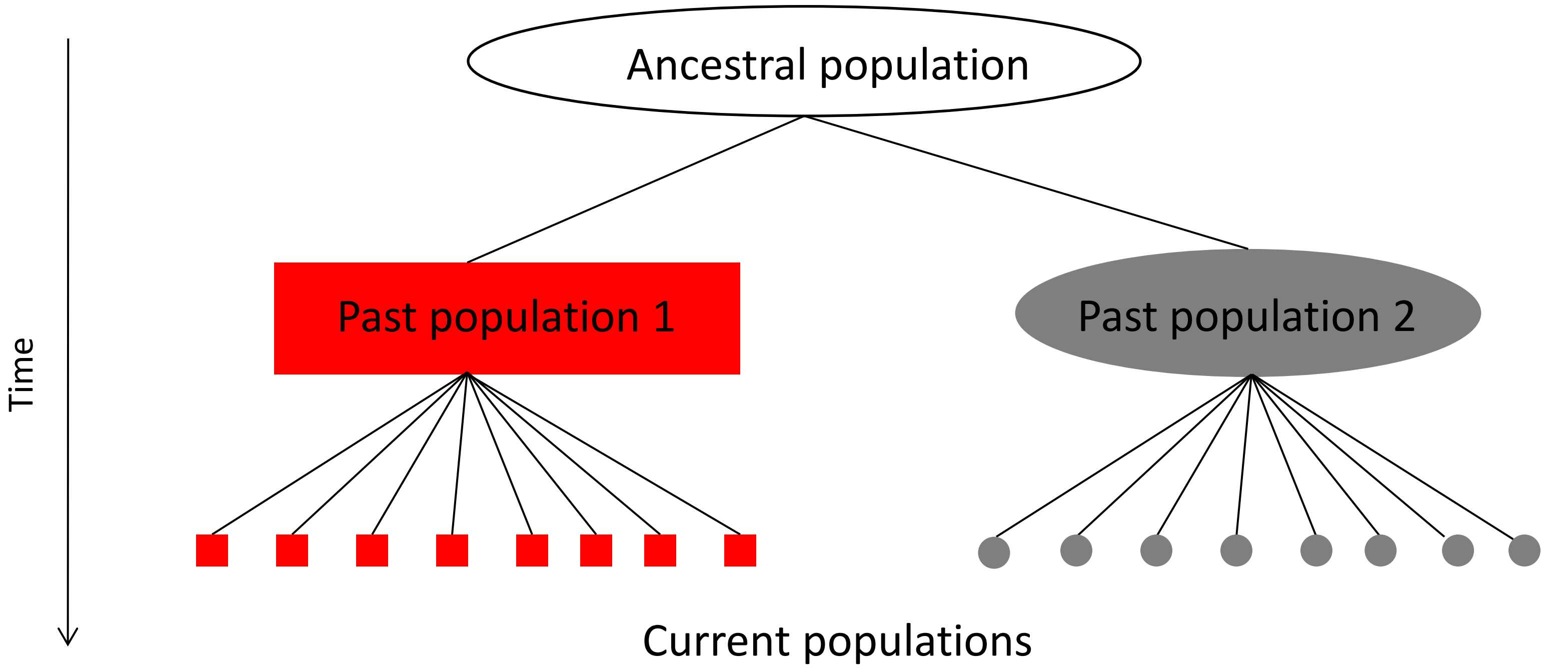
MEAN breeding values in POPULATIONS *A* and *B*

$$\mathrm{Cov}[\mathbf{a}_\mathrm{A}, \mathbf{a}_\mathrm{B}] = 2\theta_{\mathrm{AB}}\mathbf{G}$$

MEAN coancestry (relatedness) between individuals in POPULATIONS *A* and *B*

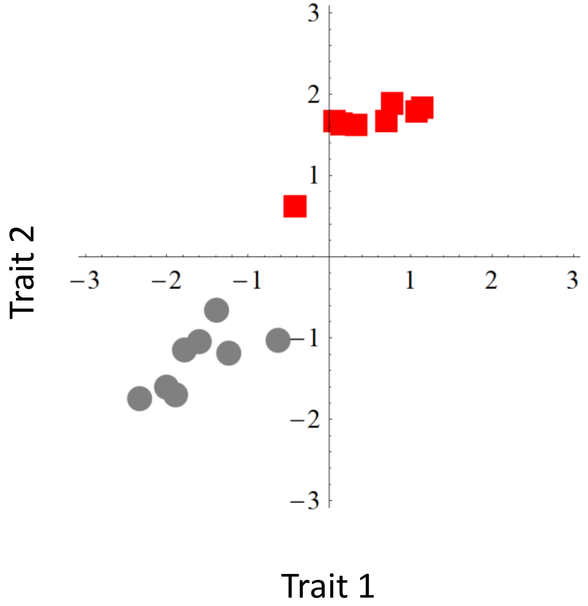
Exercise: what happens under drift?

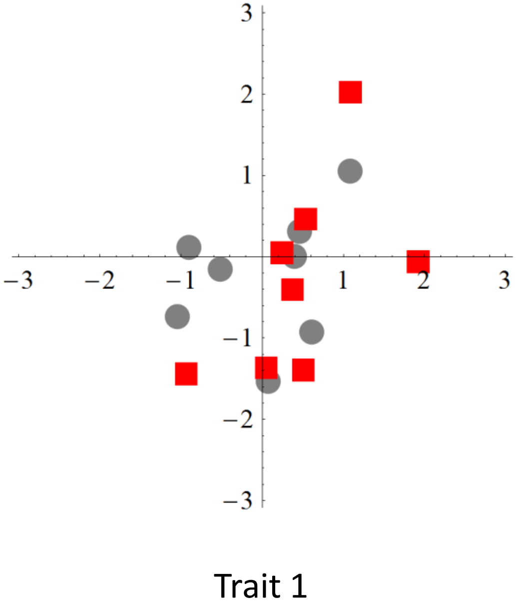
Assume that neutral molecular data tells that in terms of coancestry the populations form 2 groups:



Exercise: what happens under drift?

Which of these patterns is more likely to have evolved due to random genetic drift?


CASE 1


CASE 2

Related populations resemble each other: $\mathrm{Cov}[\mathbf{a}_{\mathrm{A}}, \mathbf{a}_{\mathrm{B}}] = 2\theta_{\mathrm{AB}}\mathbf{G}$

How to turn the eyeballing exercise into a statistical test?

Assume $\mathrm{Cov}[\mathbf{a}_{\mathrm{A}}, \mathbf{a}_{\mathrm{B}}] = 2\theta_{\mathrm{AB}}\mathbf{G}$ and e.g. normally distributed traits. Then

Vector of population means

Matrix of ancestral genetic variances and co-variances

$$\mathbf{A}^{\mathcal{P}} \sim N\left(\boldsymbol{\mu}\otimes\mathbf{I}_{n_{\mathcal{P}}}, 2\mathbf{G}^{\mathcal{A}}\otimes\boldsymbol{\theta}^{\mathcal{P}}\right)$$

Overall mean

Population-to-population relatedness matrix

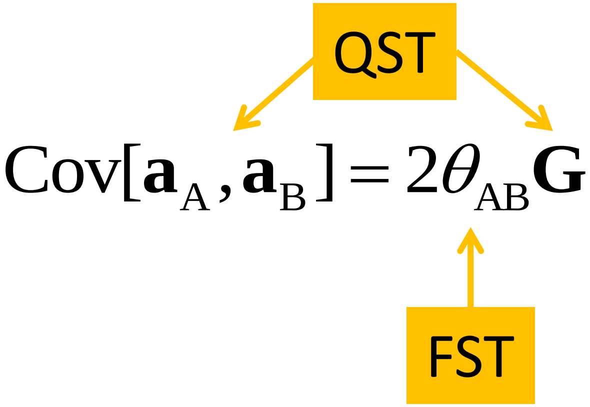
S-statistic (based on Mahalanobis distance): does the observed vector of population means fit into the “core” of this distribution or is it an “outlier”?

S close to 0: stabilizing selection
S close to 1: diversifying selection
S close to 0.5: drift plausible

How does this relate to FST-QST tests?

- FST: population divergence in neutral markers
- QST: population divergence in quantitative traits

- FST<QST: diversifying selection
- FST>QST: stabilizing selection
- FST=QST: drift plausible



QST

$$\mathrm{Cov}[\mathbf{a}_{\mathrm{A}},\mathbf{a}_{\mathrm{B}}]=2\theta_{\mathrm{AB}}\mathbf{G}$$

FST

Unlike FST-QST tests, this equation

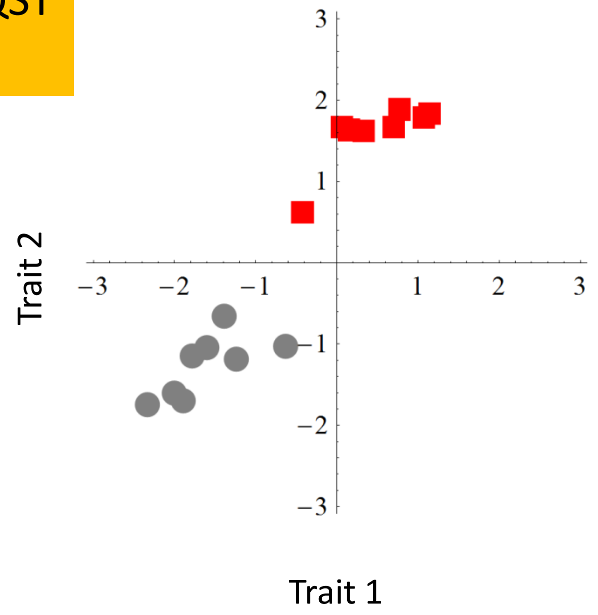
- accounts for evolutionary stochasticity
- utilizes population-specific information
- extends naturally to multivariate traits

Exercise: what happens under drift?

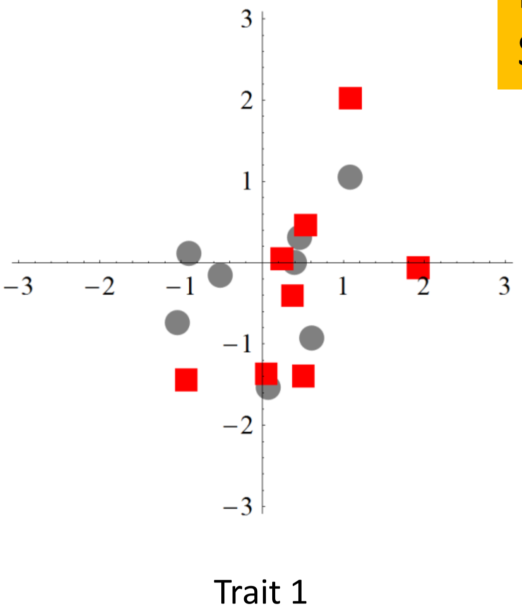
Which of these patterns is more likely to have evolved due to random genetic drift?

FST=QST
S=0.5

CASE 1



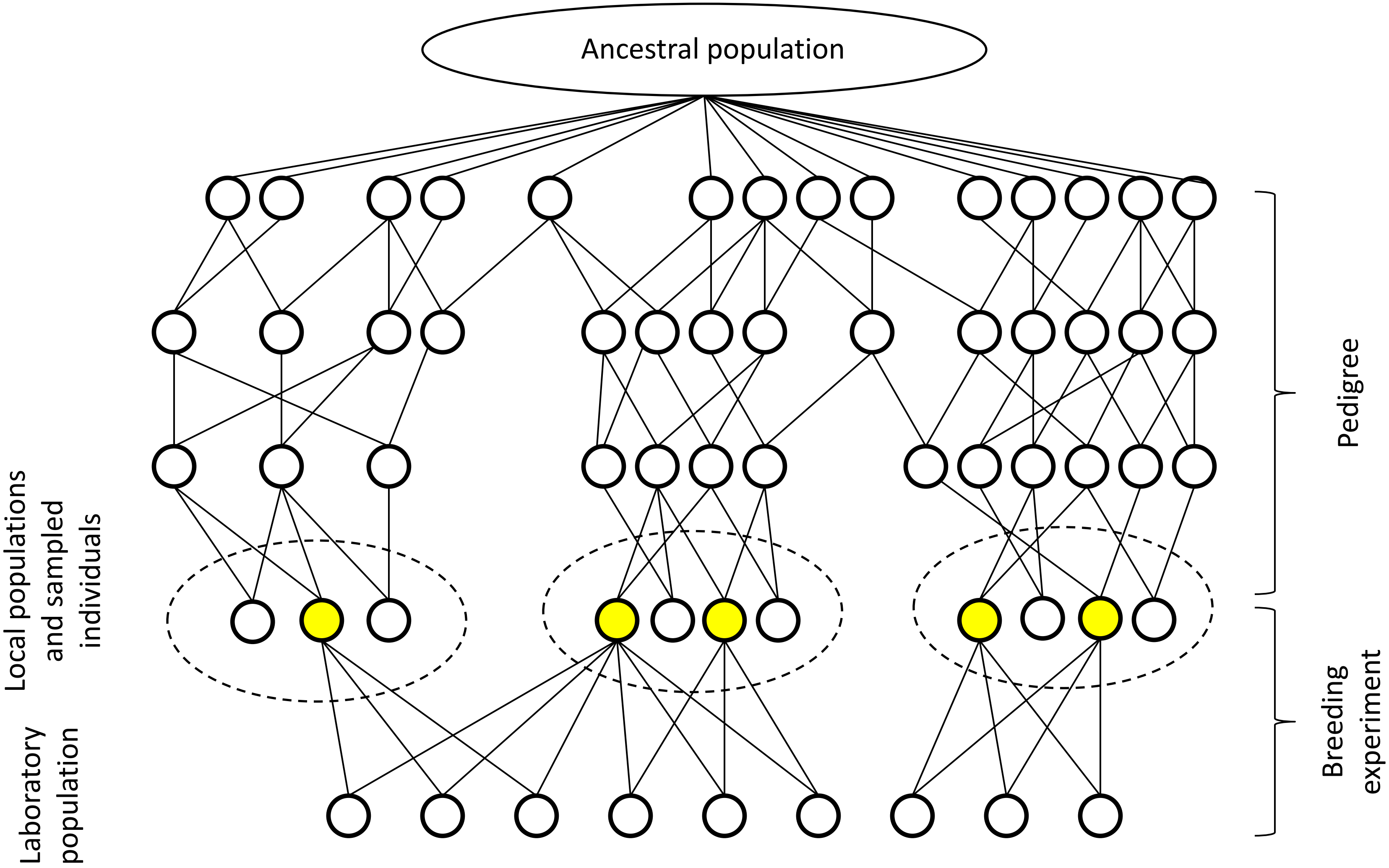
CASE 2



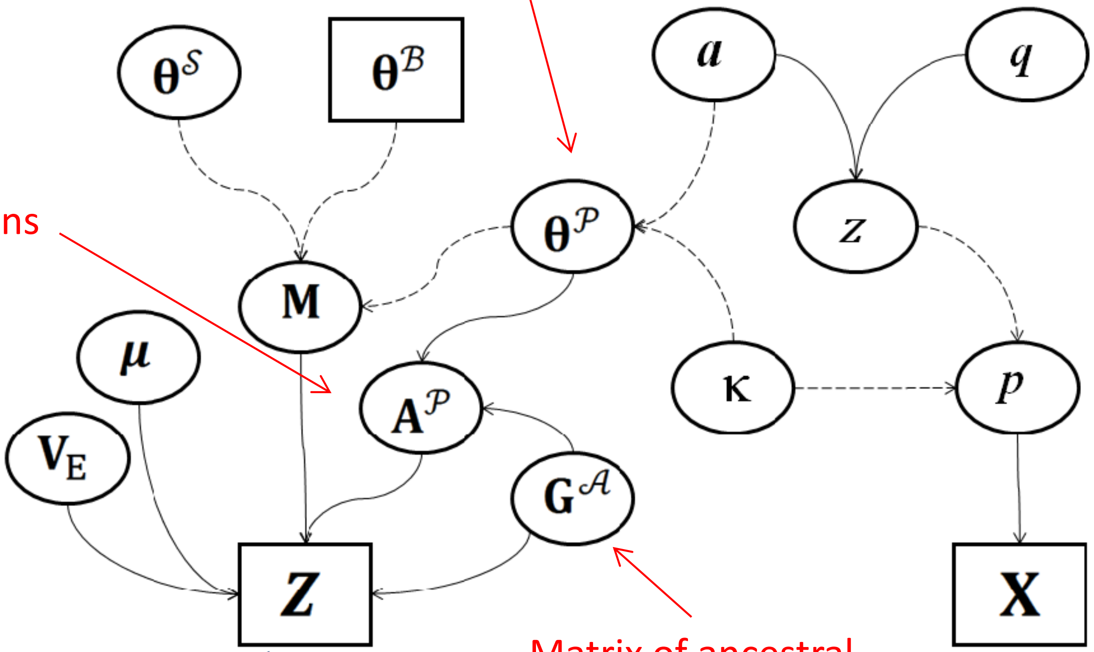
FST=QST
S>0.99

Related populations resemble each other: $\mathrm{Cov}[\mathbf{a}_{\mathrm{A}}, \mathbf{a}_{\mathrm{B}}] = 2\theta_{\mathrm{AB}}\mathbf{G}$

Data: neutral markers + quantitative traits



Parameter estimation with a Baysian approach



Population-to-population relatedness matrix

Vector of population means

Matrix of ancestral genetic variances and co-variances

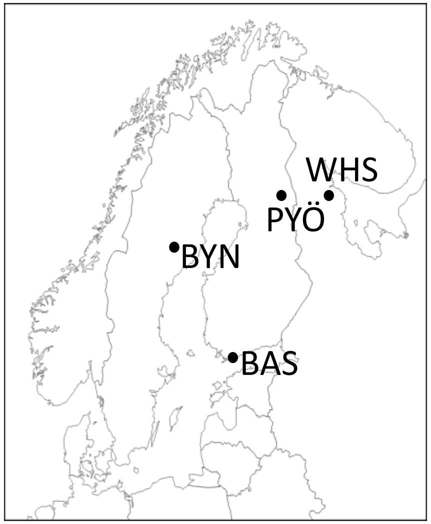
Quantitative trait data from breeding experiments

Neutral marker data

Directed Acyclic Graph (DAG) representing a hierarchical Bayesian model

Example results on nine-spine sticklebacks

Morphological traits: *S*=1.00
Behavioural traits: *S*=0.91



S close to 0: stabilizing selection
S close to 1: diversifying selection
S close to 0.5: drift plausible

Signature of local adaptation from habitat information?

Juha Merilä

We found small fish from BOTH sea populations, and large fish from BOTH pond populations – this must give additional evidence of selection operating.

H-statistic: are populations from similar habitats more similar than expected by random drift?

Mantel test statistic

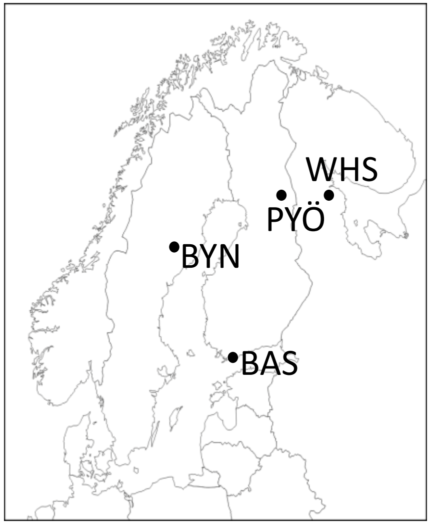
$$H := \mathrm{P}\{\mathrm{M}(\mathbf{D}^E, \mathbf{D}^T) > \mathrm{M}(\mathbf{D}^E, \mathbf{D}^{T,R})\}.$$

Distance matrices of environmental covariates and mean trait values

Example results on nine-spine sticklebacks

Morphological traits: *S*=1.00
Behavioural traits: *S*=0.91

Morphological traits: *H*=1.00
Behavioural traits: *H*=0.99



L5: take home messages

- Major forces behind evolution include selection, drift, mutation and migration (gene-flow)

- Showing that two populations are genetically different does not necessarily mean that the populations have been influenced by different selection pressures.

- To rule out the alternative explanation of genetic drift, the pattern of genetic divergence among the populations can be compared to patterns that would be generated by drift alone.

- Quantitative genetic theory provides theoretically well-founded and statistically powerful approaches for testing hypotheses related to selection and drift, as well as estimating related key parameters (e.g. G-matrices, gene-flows, coancestry matrices at the individual and population levels, etc.)