# The effects of epistasis and pleiotropy on local adaptation and the detection of adaptive outlier loci

**Reinhard Bürger**

Department of Mathematics

universität wien

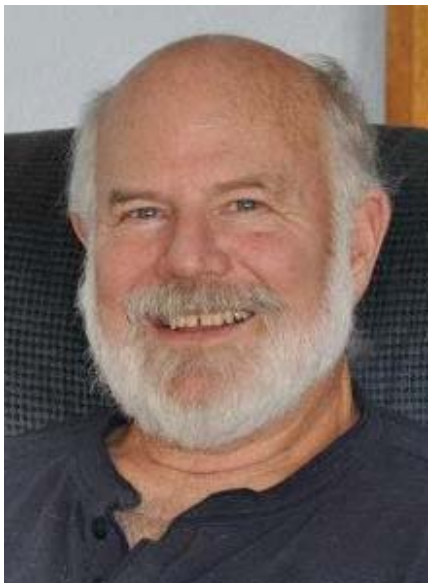
Vienna, 14 February 2019, SMBE Satellite Meeting

'Towards an integrated concept of adaptation: uniting molecular population genetics and quantitative genetics'

# Collaborators[1]

Adam G. Jones

Stevan J. Arnold



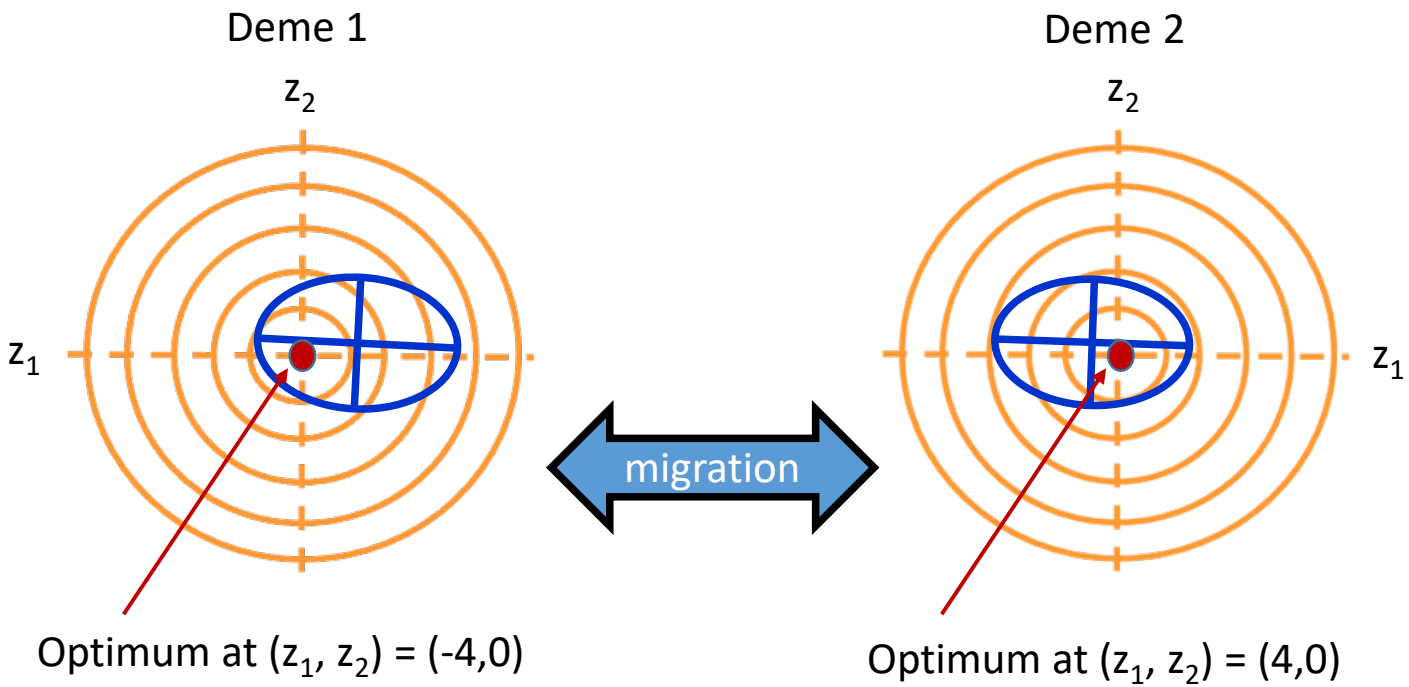
[1] AG Jones, SJ Arnold, R Bürger: The effects of epistasis and pleiotropy on genome-wide scans for adaptive outlier loci. J. of Heredity, online (2019)

# Goals

- Model the evolution of two quantitative traits in a pair of populations that are subject to selection towards different phenotypic optima and exchange migrants
- Investigate how quantitative genetic architectures that include pleiotropy and epistasis affect
  - Ability of populations to adapt to their local optima
  - Patterns of differentiation between locally adapted populations
  - Efficacy of genome-wide scans for selection based on outlier loci

# Selection and migration

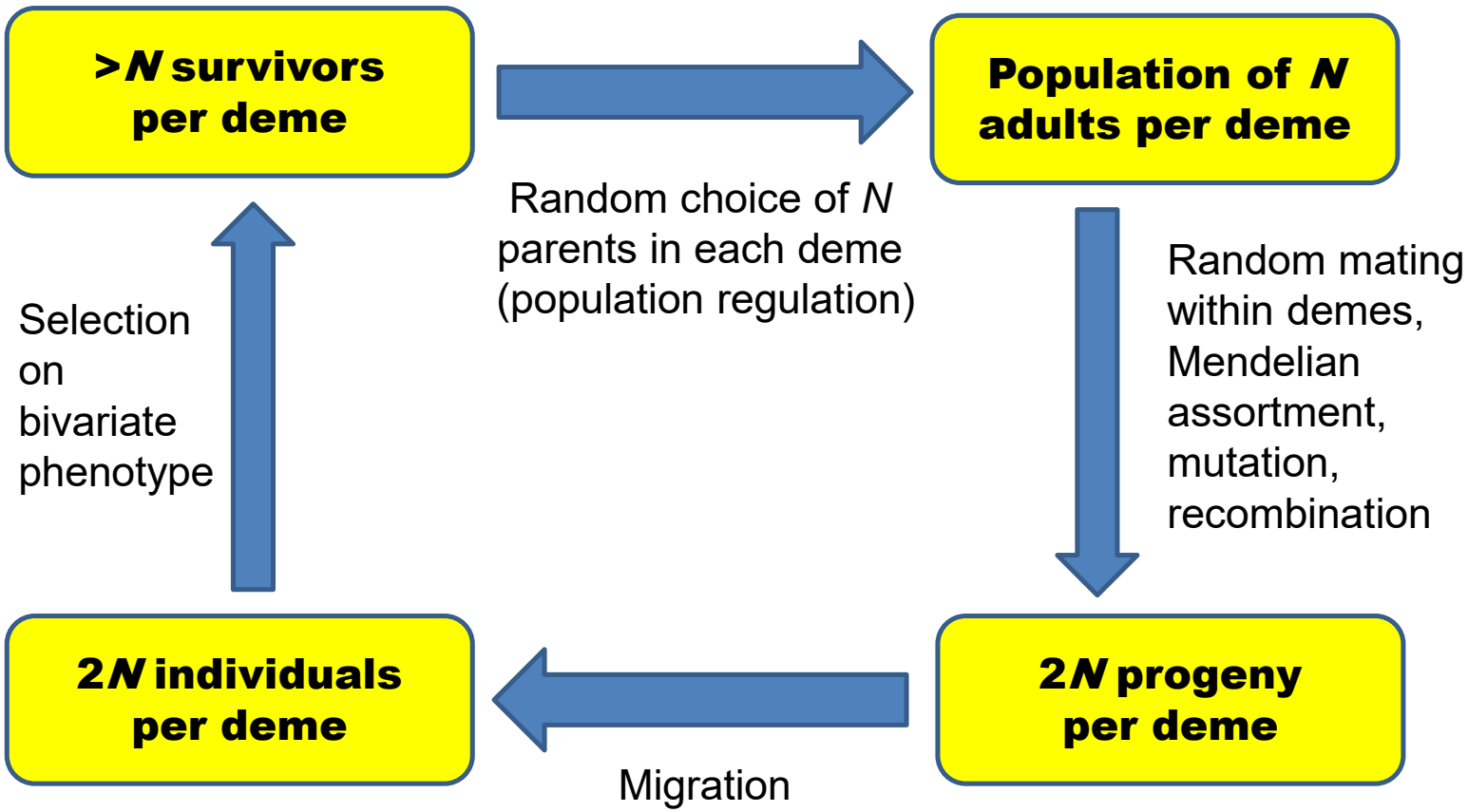
- In each of two demes, there is a symmetric bivariate Gaussian selection surface with no correlational selection
- Migration is symmetric and the migration probability $m$ is varied between 0 and 0.256




(Units are environmental standard deviations)

# Simulation model

- Individual-based forward-in-time simulation of diploids with separate sexes and polygyneous mating system
- Life cycle:



# Genetic system

- Genome consists of marker loci and of QTL
- A QTL may affect only trait 1, only trait 2, or be pleiotropic
- Marker loci are arranged in linkage groups, each of which has a specified recombination rate $R$
- $R =$ Expected number of recombination events per meiosis per individual; $0.1 \leq R \leq 4$; typical value $R = 0.25$
- Each linkage group contains between 500 and 10000 evenly spaced markers; typical value: 2000
- Each linkage group contains 1 - 5 QTL (at random positions)

# Genetic system

- Mutations at QTL are drawn from univariate or bivariate normal distributions and added to existing effects
- Markers are allowed to have up to four alleles
- Mutations at markers result in one of the other allelic types
- Epistasis is implemented according to the multivariate version of the multilinear model of Hansen and Wagner (2001); see Jones et al. (2014)
- In the multilinear model a gene substitution can change the phenotypic effect of any other gene or genotypic substitution, but only as a linear function of its own phenotypic effect

# The multilinear model for a bivariate trait and pleiotropic effects

- Let $(\xi_1, \xi_2)$ be the bivariate value of an arbitrary (multilocus) reference genotype (e.g., the population mean).
- Let $(y_1^{(i)}, y_2^{(i)})$ be the effect of a genotype at locus $i$ if substituted into the reference genotype.

With two loci and, therefore, only pairwise epistatic interactions, the genotypic value $x_1$ of trait 1 is

$$x_1 = \xi_1 + y_1^{(1)} + y_1^{(2)}$$

# The multilinear model for a bivariate trait and pleiotropic effects

- Let $(\xi_1, \xi_2)$ be the bivariate value of an arbitrary (multilocus) reference genotype (e.g., the population mean).
- Let $(y_1^{(i)}, y_2^{(i)})$ be the effect of a genotype at locus $i$ if substituted into the reference genotype.

With two loci and, therefore, only pairwise epistatic interactions, the genotypic value $x_1$ of trait 1 is

$$x_1 = \xi_1 + y_1^{(1)} + y_1^{(2)} + \epsilon_{111}^{(1,2)} y_1^{(1)} y_1^{(2)}$$

# The multilinear model for a bivariate trait and pleiotropic effects

- Let $(\xi_1, \xi_2)$ be the bivariate value of an arbitrary (multilocus) reference genotype (e.g., the population mean).
- Let $(y_1^{(i)}, y_2^{(i)})$ be the effect of a genotype at locus $i$ if substituted into the reference genotype.

With two loci and, therefore, only pairwise epistatic interactions, the genotypic value $x_1$ of trait 1 is

$$x_1 = \xi_1 + y_1^{(1)} + y_1^{(2)} + \epsilon_{111}^{(1,2)} y_1^{(1)} y_1^{(2)} + \epsilon_{112}^{(1,2)} y_1^{(1)} y_2^{(2)} + \epsilon_{121}^{(1,2)} y_2^{(1)} y_1^{(2)} + \epsilon_{122}^{(1,2)} y_2^{(1)} y_2^{(2)}$$

where $\epsilon_{abc}^{(i,j)}$ measures the epistatic effect on trait $a$ of the interaction between the effects of locus $i$ on trait $b$ and locus $j$ on trait $c$.

# The multilinear model

- For two traits and $n$ pleiotropic QTLs, there are $4n(n-1)$ pariwise epistatic coefficients $\epsilon_{abc}^{(i,j)}$
- We draw the coefficients $\epsilon_{abc}^{(i,j)}$ from a normal distribution with

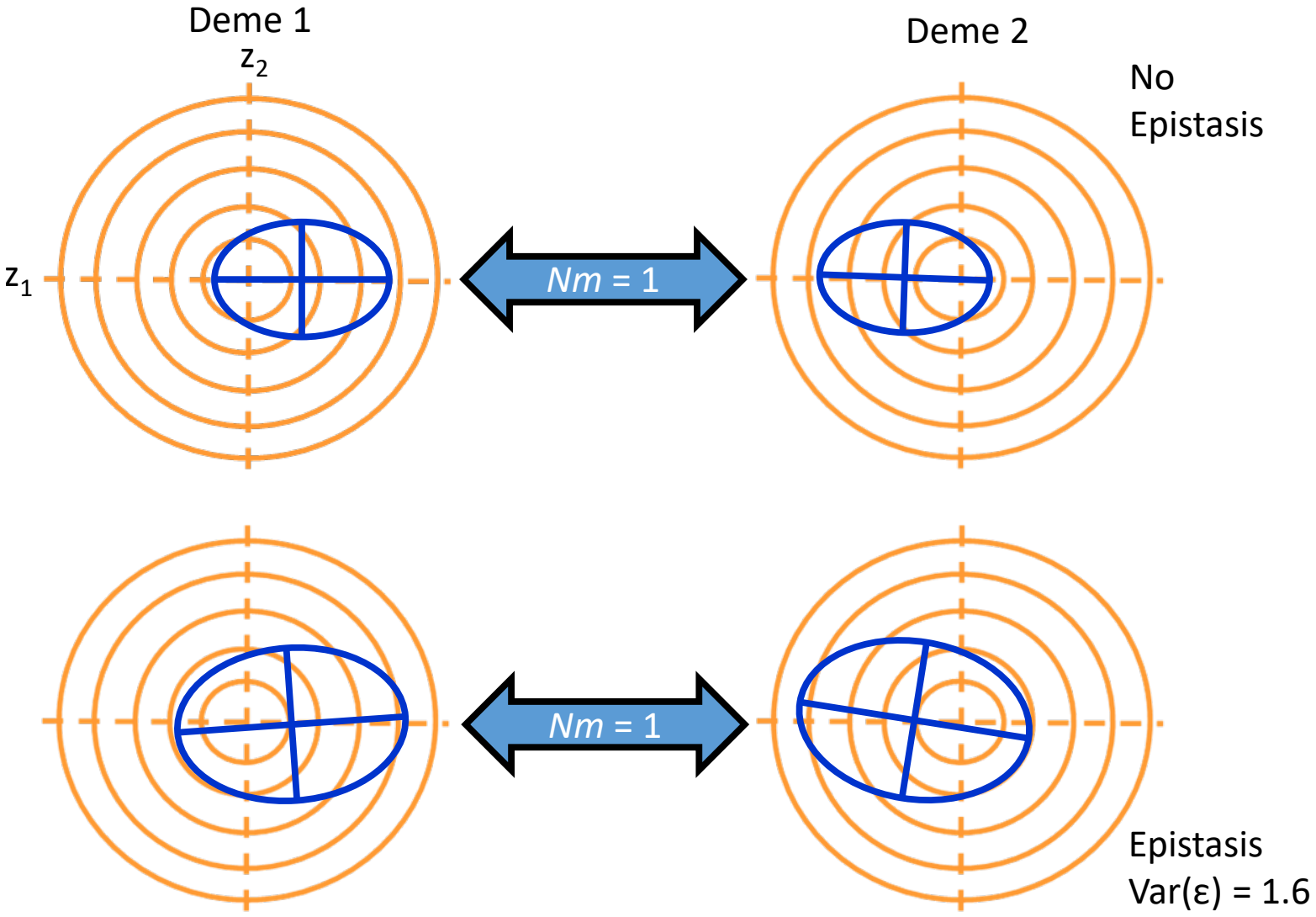
$$\mathrm{E}(\epsilon) = 0 \quad \text{and} \quad \mathrm{Var}(\epsilon) = \sigma_\epsilon^2$$

  (on average, positive and negative epistatic effects cancel)
- Epistasis coefficients remain constant during each run
- An independent environmental effect (from a standardized normal distribution) is added to the genotypic values

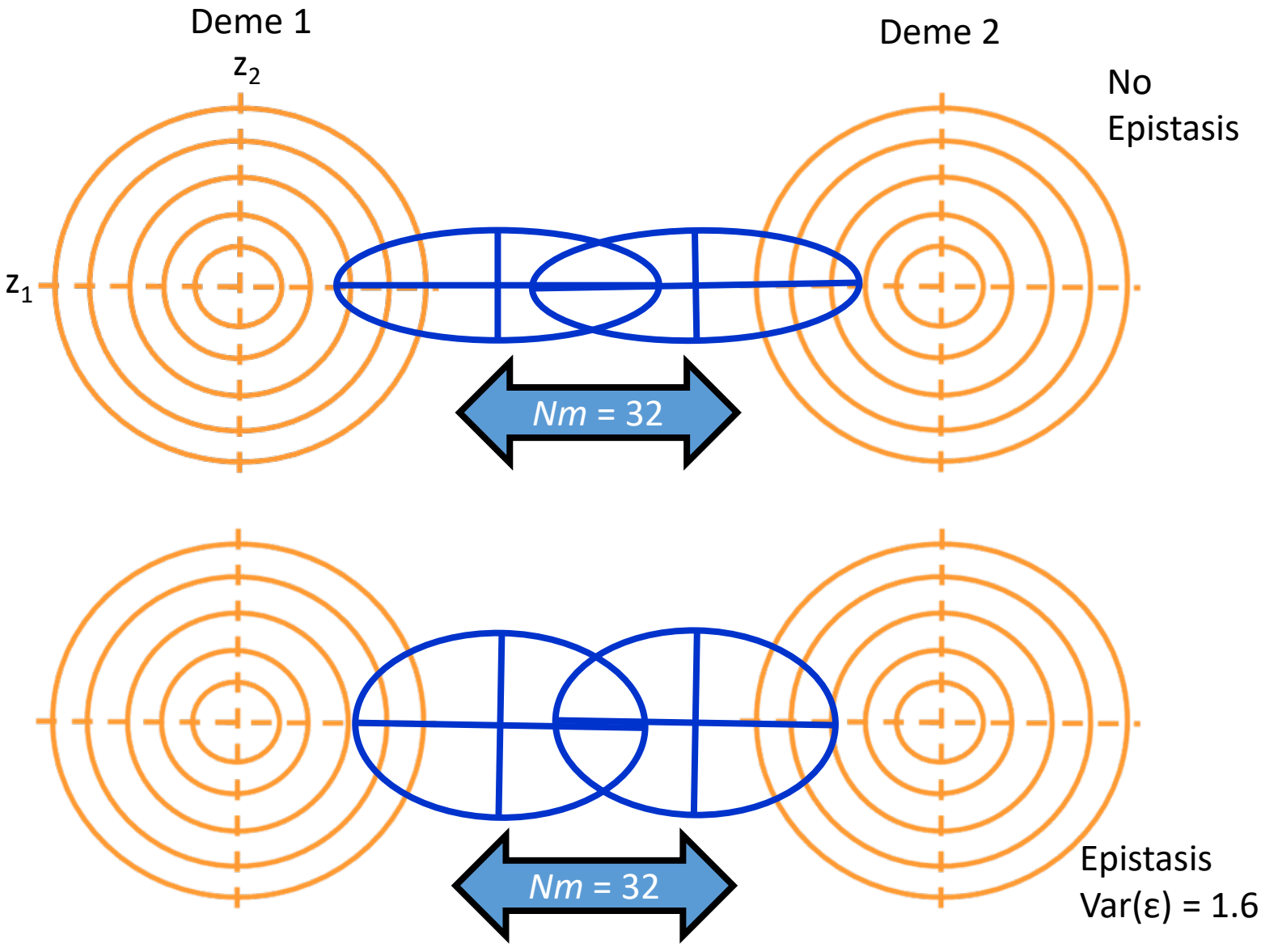
# Implementation

- Long burn-in period, so that initial populations are approximately in migration-selection-mutation-drift balance
- Quantities of interest are measured during 2000 experimental generations, and then averaged
- There are 30 replicate runs for each parameter combination; each replicate run starts from new allelic values, new randomly chosen epistatic parameters, and new locations for the QTLs

# Epistasis and local adaptation: weak migration

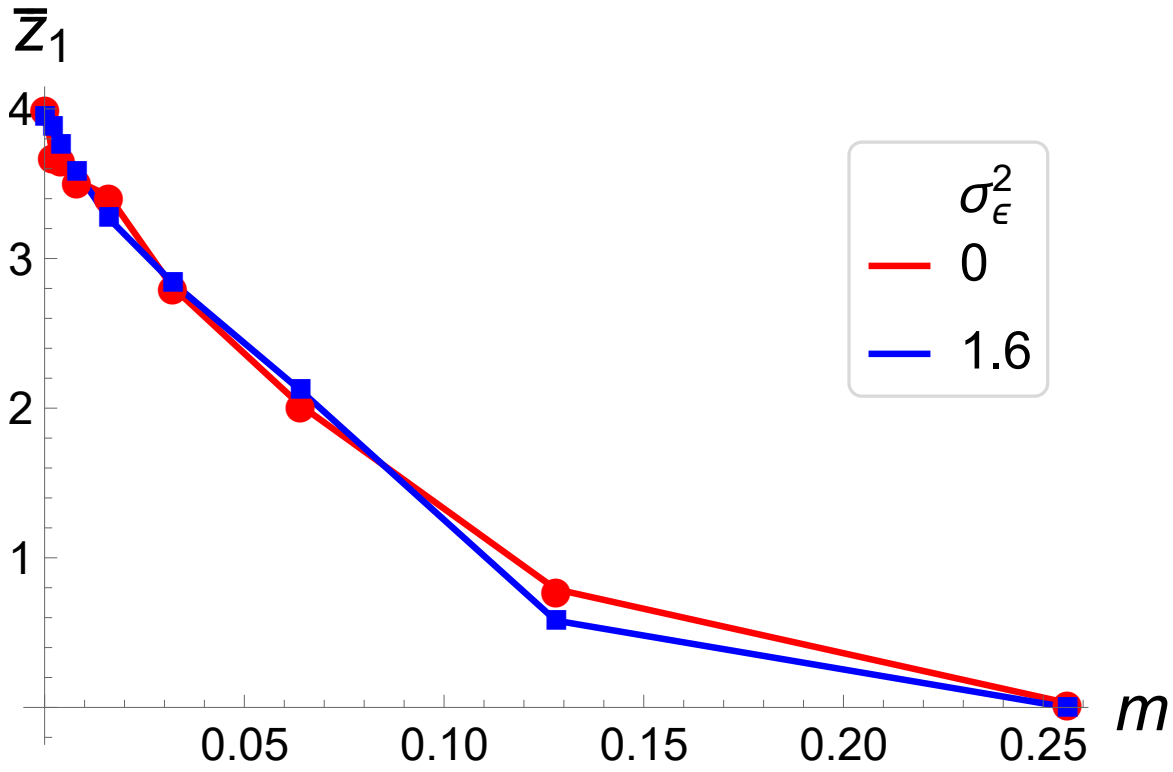


# Epistasis and local adaptation: strong migration



# Local adaptation: epistasis but no pleiotropy

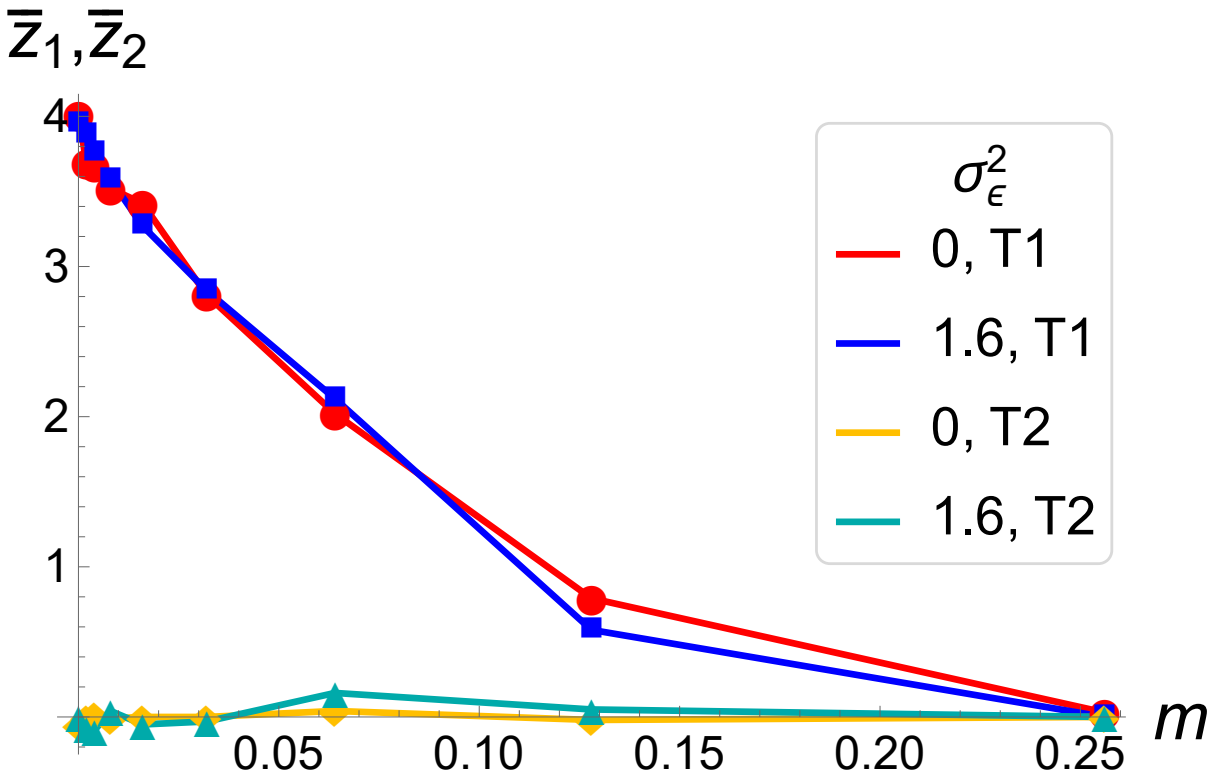
Mean of trait 1 as a function of the migration rate:




Trait optima at $(z_1, z_2) = (\pm 4, 0)$; $N = 500$; each trait determined by 4 QTL.

# Local adaptation: epistasis but no pleiotropy

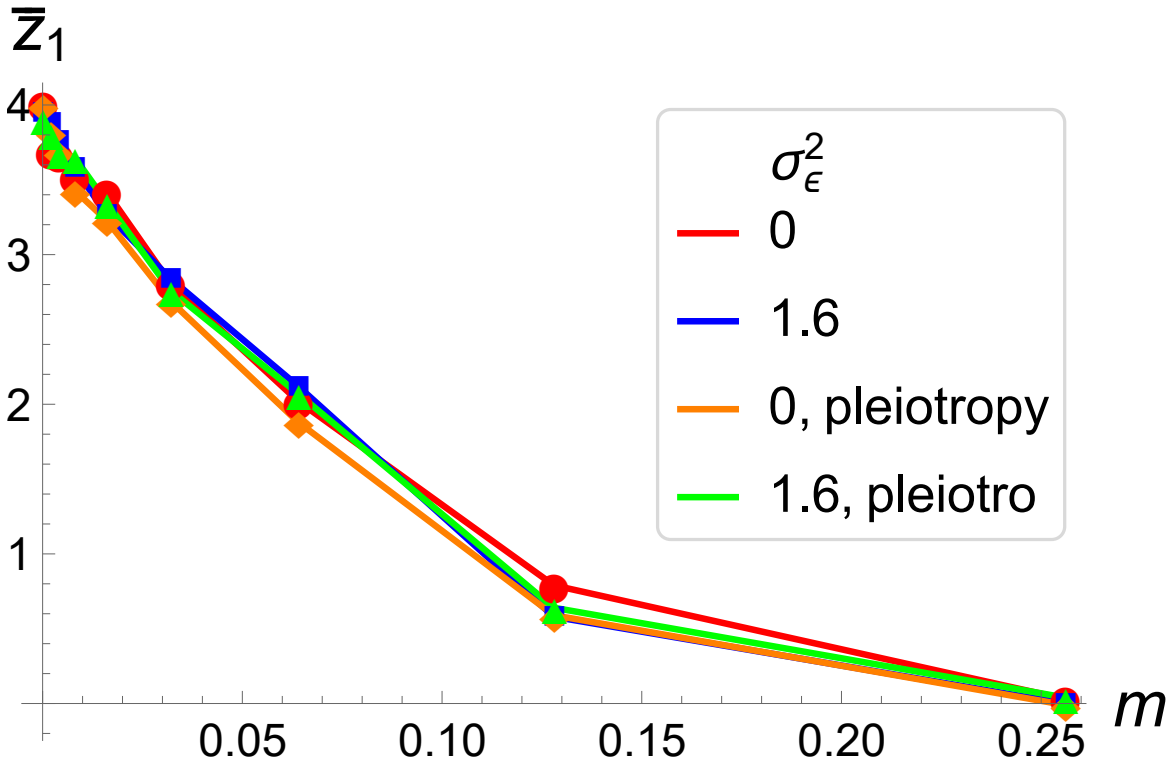
Means of traits 1 and 2 as functions of the migration rate




Trait optima at $(z_1, z_2) = (\pm 4, 0)$; $N = 500$; each trait determined by 4 QTL.

# Local adaptation: epistasis and pleiotropy

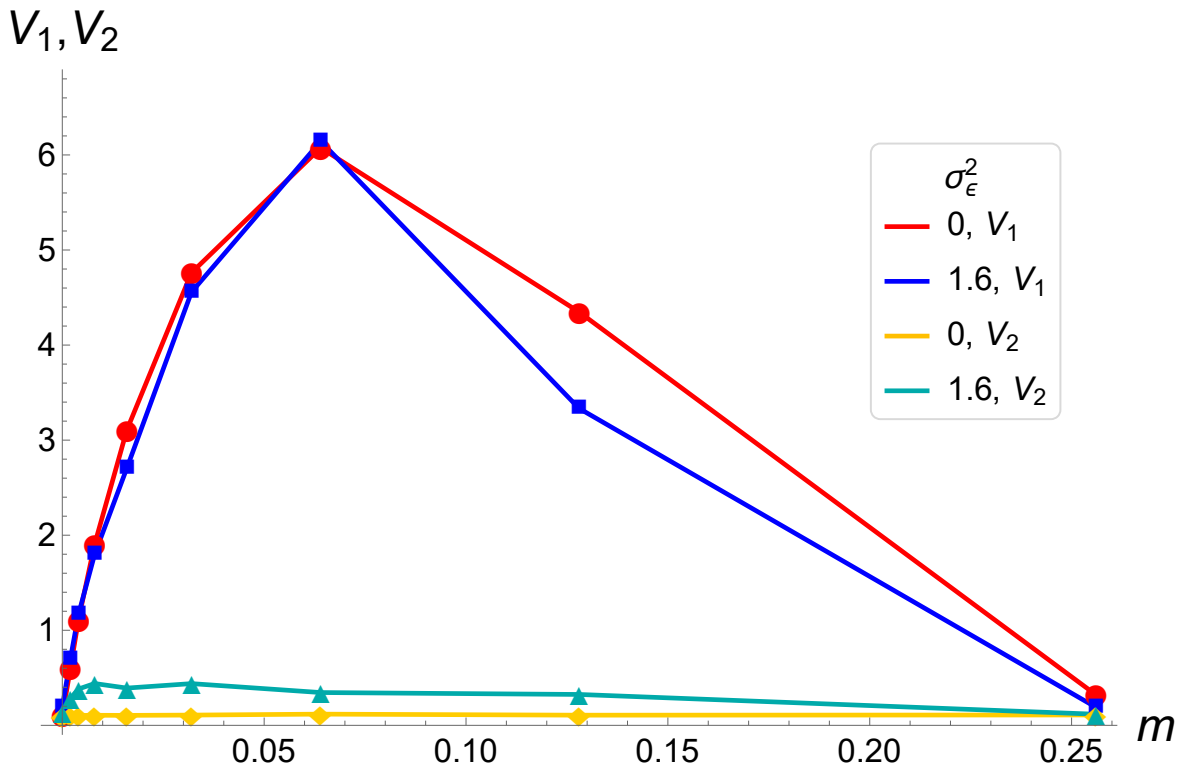
Mean of trait 1 as a function of the migration rate:




Trait optima at $(1, 2)$ at $(z_1, z_2) = (\pm 4, 0)$; $N = 500$; 4 pleiotropic QTL.

# Local adaptation: epistasis but no pleiotropy

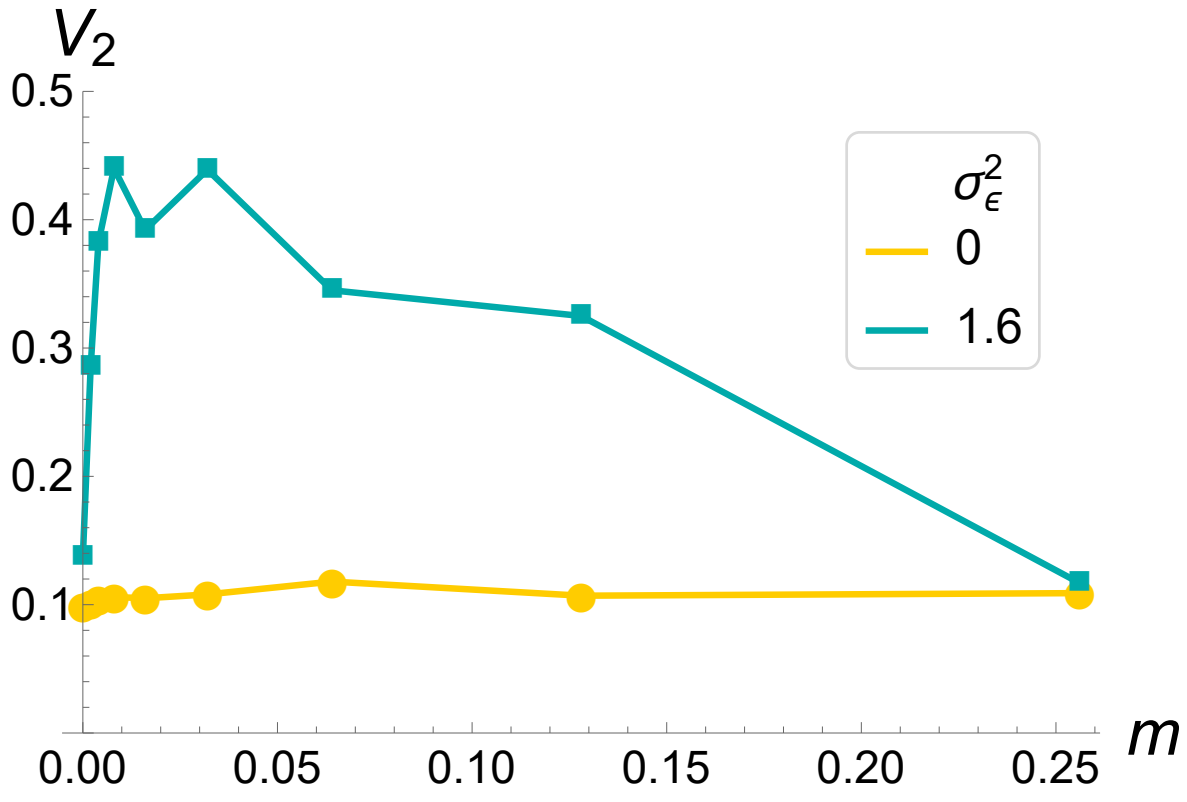
Variances of traits 1 and 2 as functions of the migration rate




Trait optima at $(z_1, z_2) = (\pm 4, 0)$; $N = 500$; each trait determined by 4 QTL.

# Local adaptation: epistasis but no pleiotropy

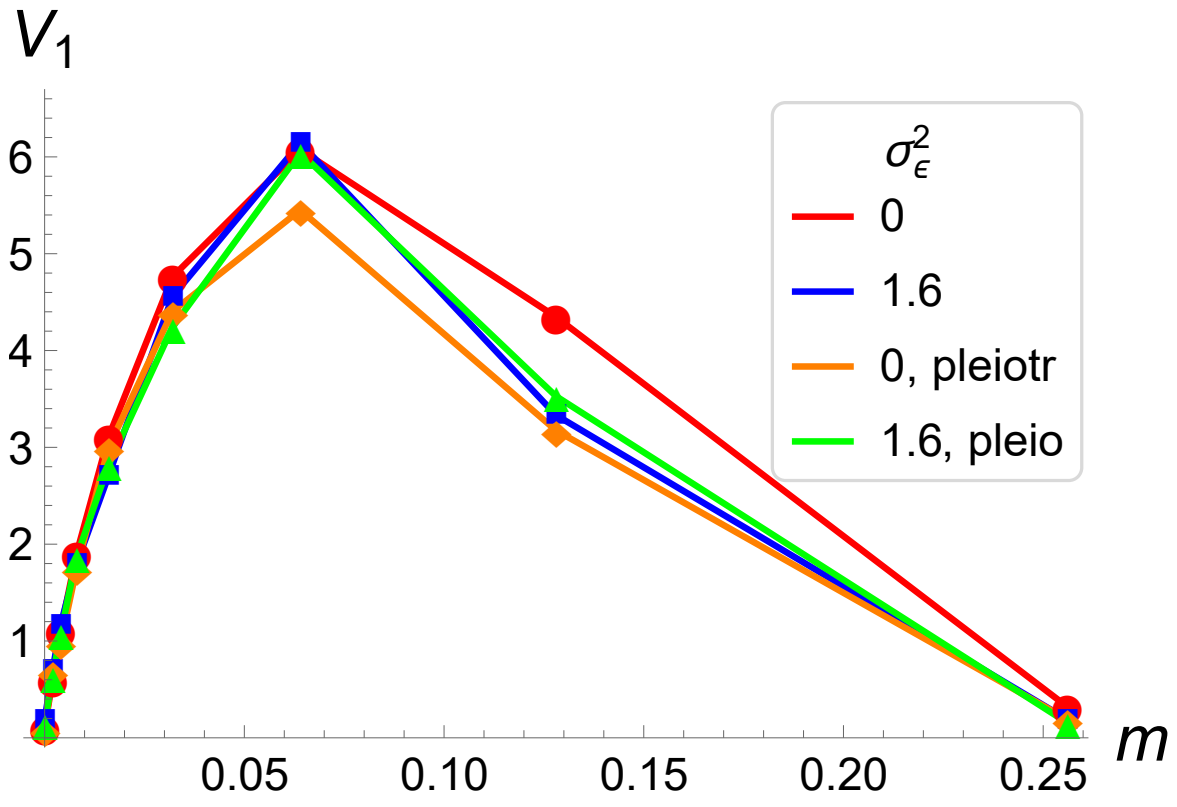
Variance of trait 2 as a function of the migration rate




Trait optima at $(z_1, z_2) = (\pm 4, 0)$; $N = 500$; each trait determined by 4 QTL.

# Local adaptation: epistasis and pleiotropy

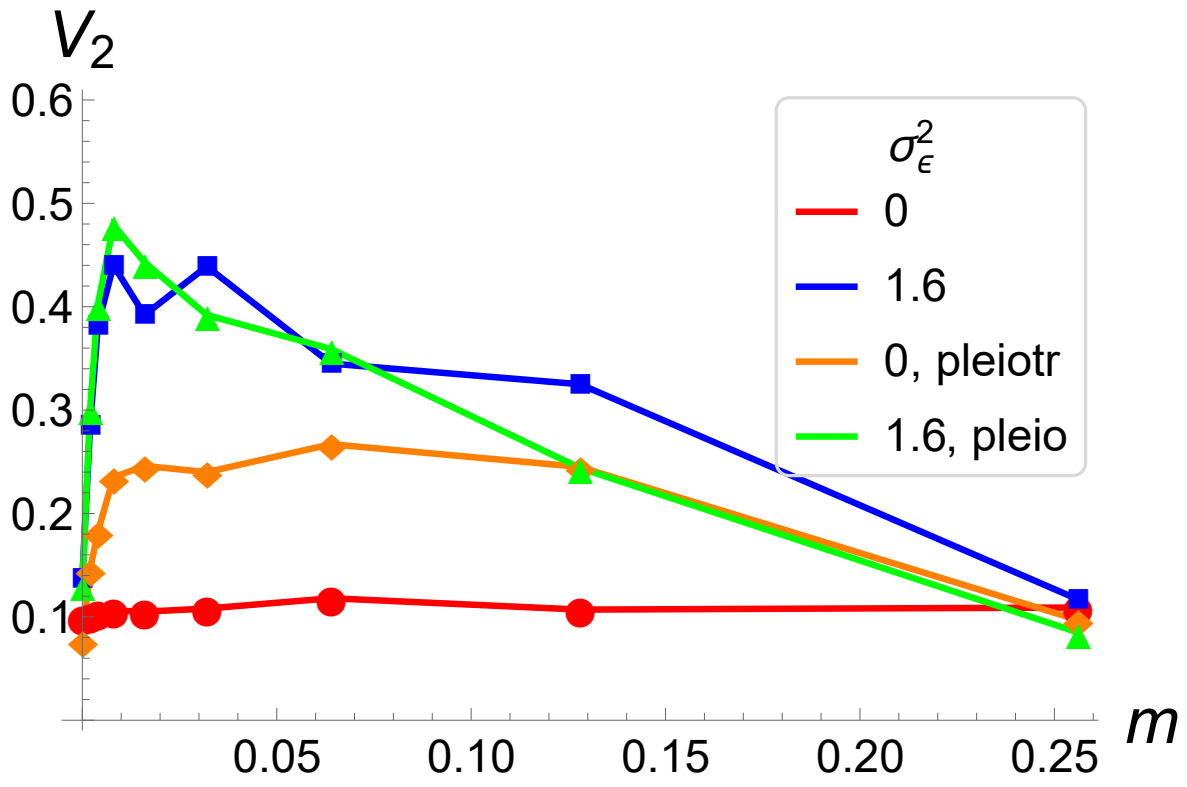
Variance of trait 1 as a function of the migration rate




Trait optima at $(1, 2)$ at $(z_1, z_2) = (\pm 4, 0)$; $N = 500$; 4 pleiotropic QTL.

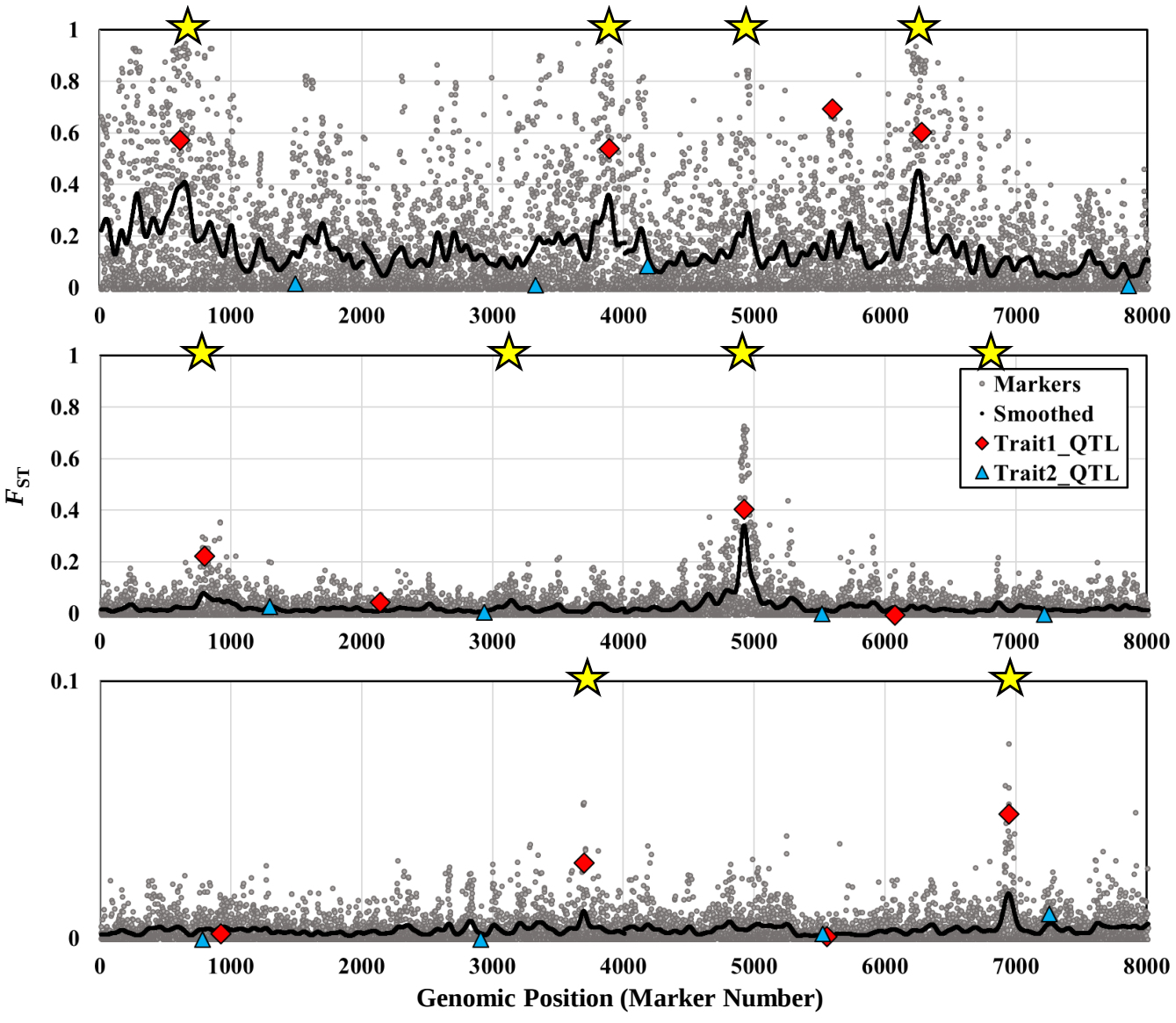
# Local adaptation: epistasis and pleiotropy

Variance of trait 2 as a function of the migration rate




Trait optima at $(1, 2)$ at $(z_1, z_2) = (\pm 4, 0)$; $N = 500$; 4 pleiotropic QTL.

# Manhattan plots: no epistasis or pleiotropy

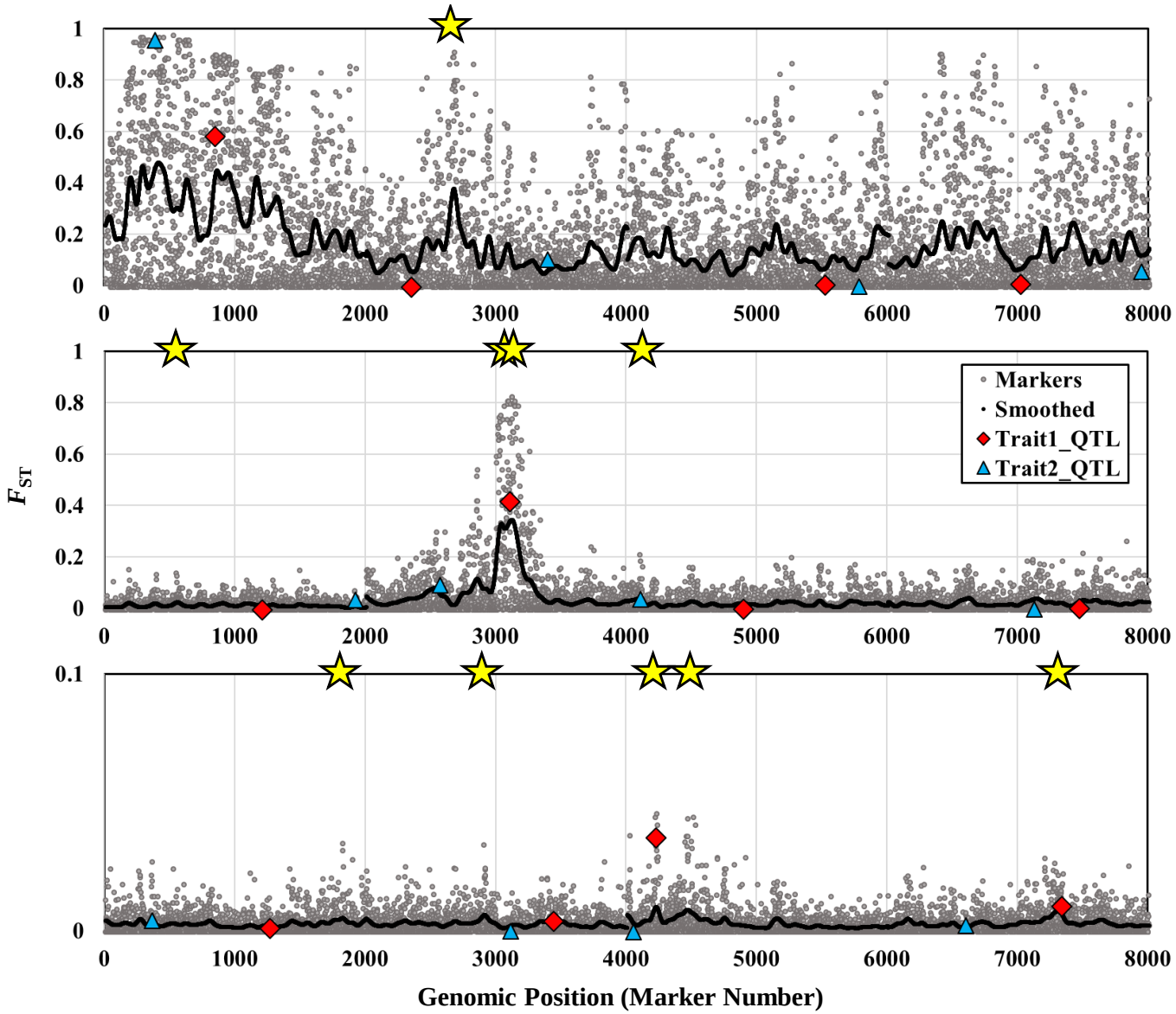



$m = 0.002$

$m = 0.016$

$m = 0.128$

# Manhattan plots: epistasis but no pleiotropy

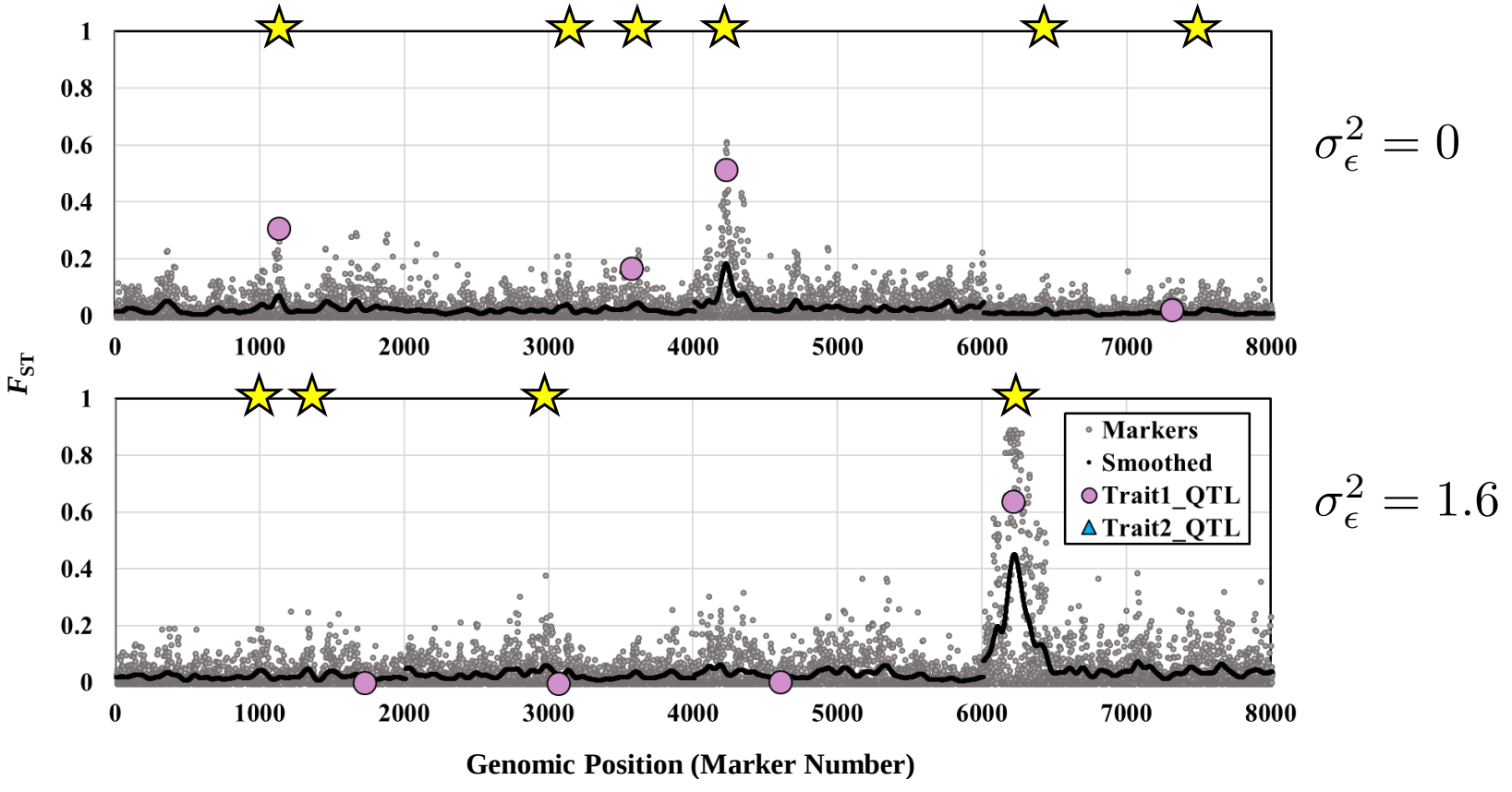



$m = 0.002$

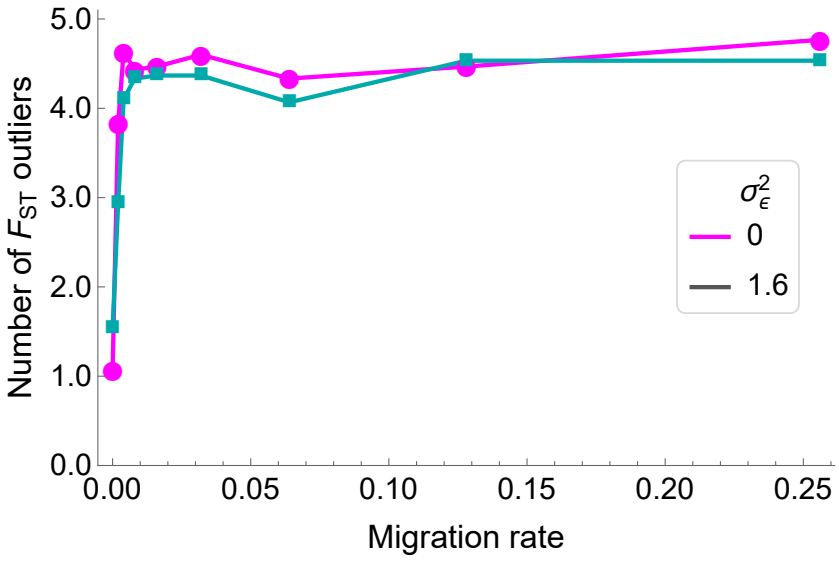
$m = 0.016$

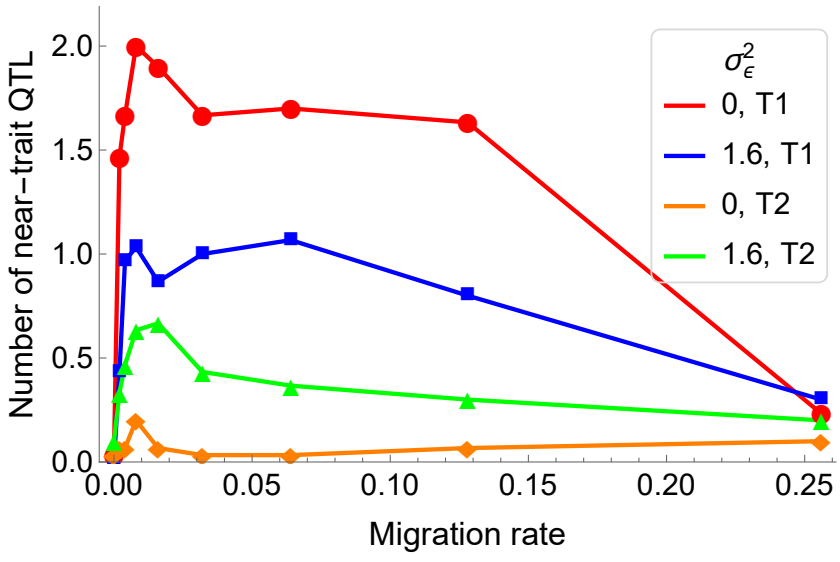
$m = 0.128$

# Manhattan plots: pleiotropy and $m = 0.016$



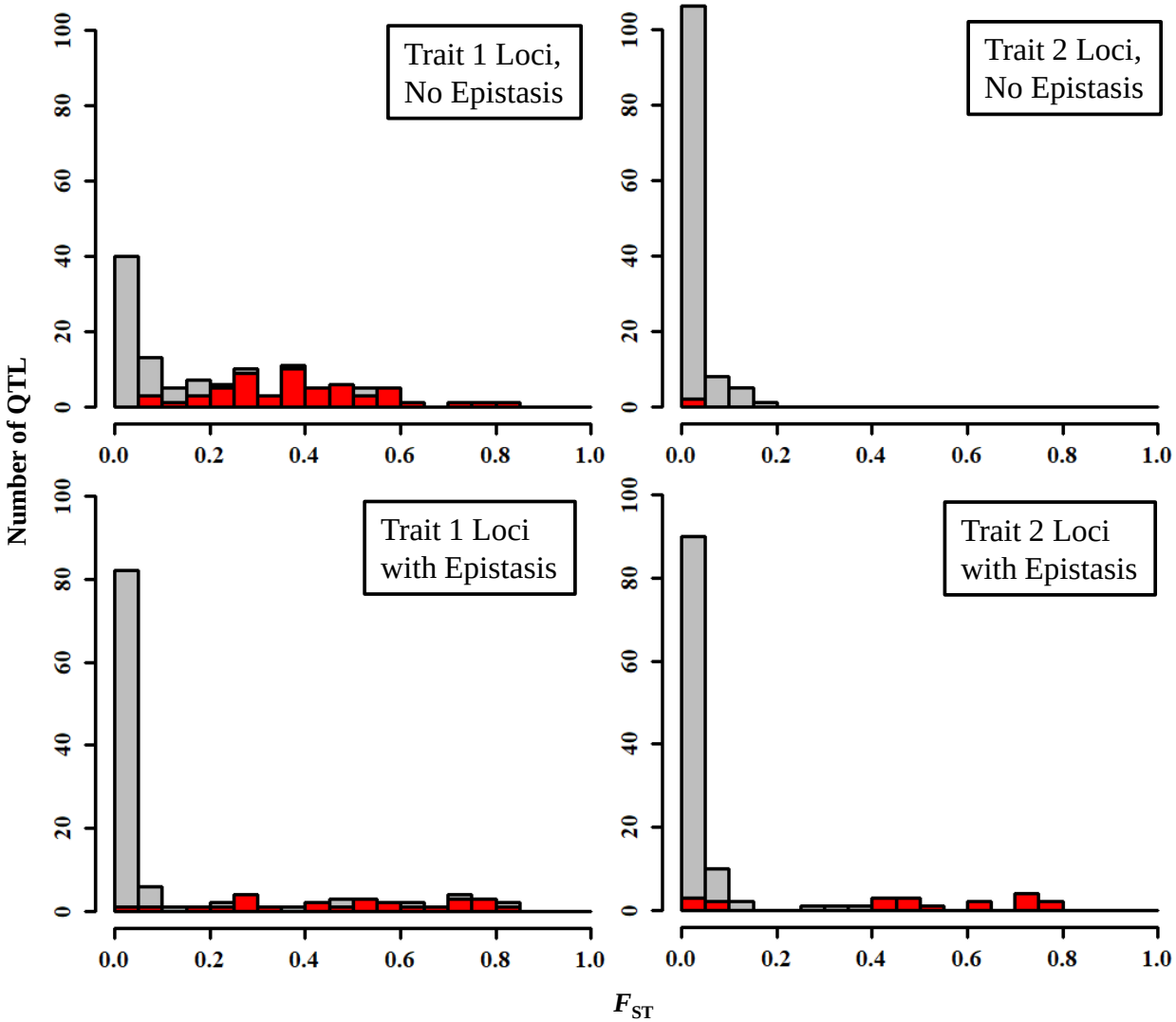
# Number of outliers and "true" outliers



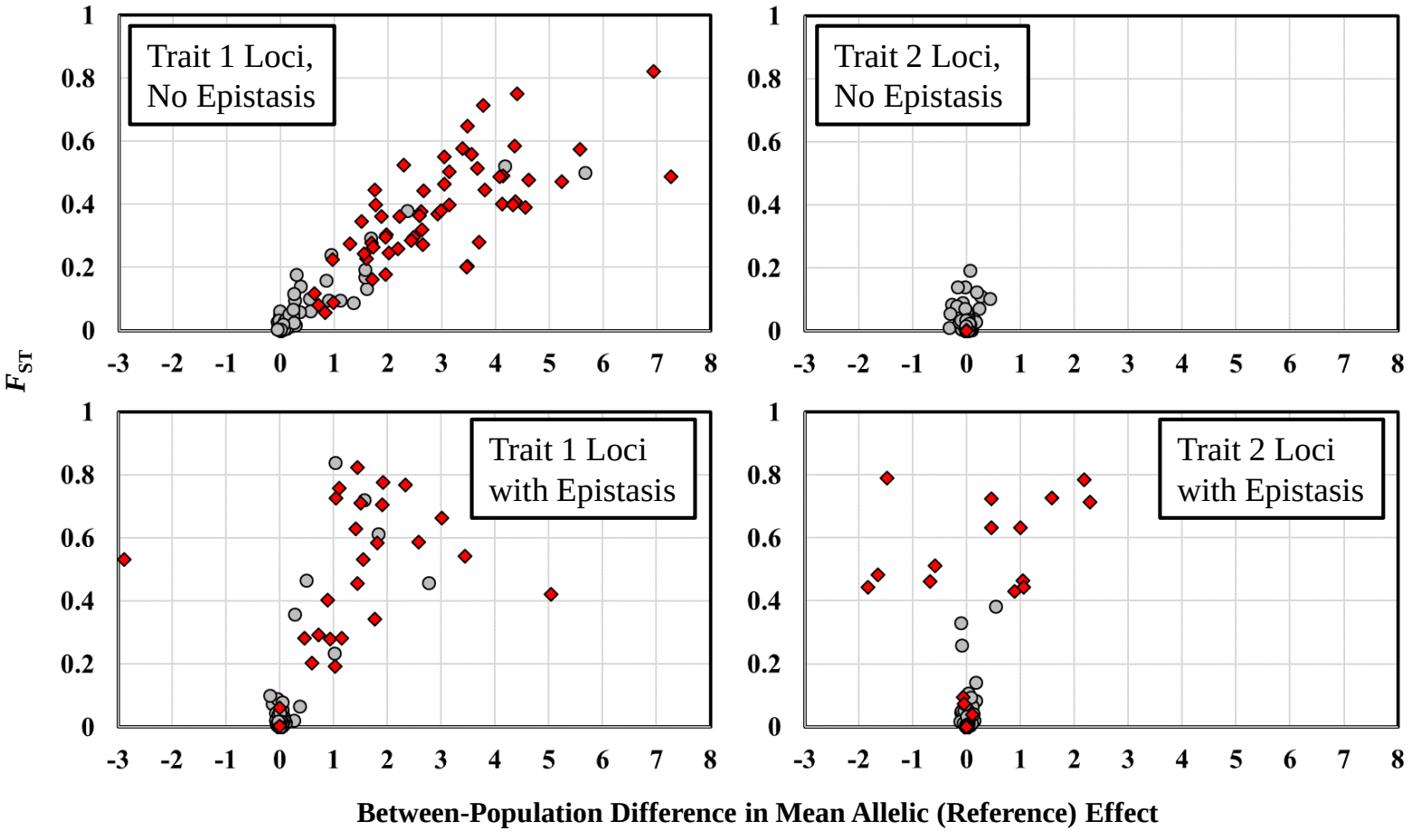

30 simulation runs with 4 linkage groups and 1 QTL per group $\rightarrow$ 120 QTL in total

# Histograms showing distributions of $F_{ST}$ of trait 1 and trait 2 QTL (no pleiotropy, $m = 0.016$)



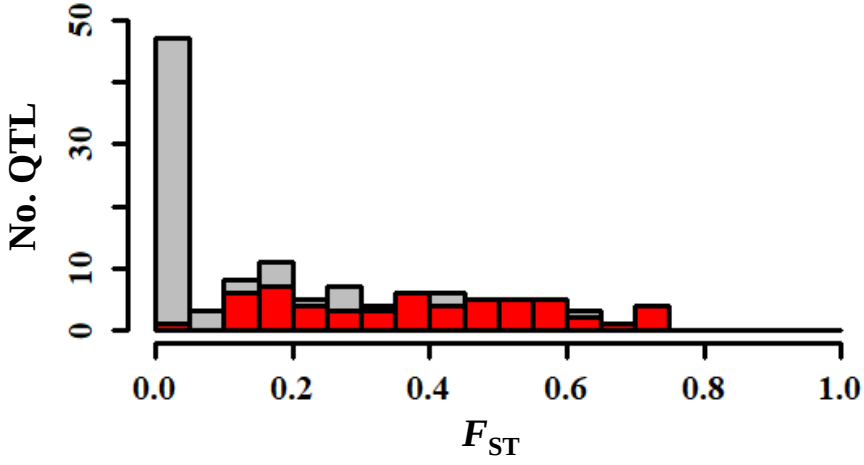

30 simulation runs with 4 linkage groups and 1 QTL per group $\rightarrow$ 120 QTL in total;
Red: QTL identified as outlier (by an outlier marker in the vicinity)

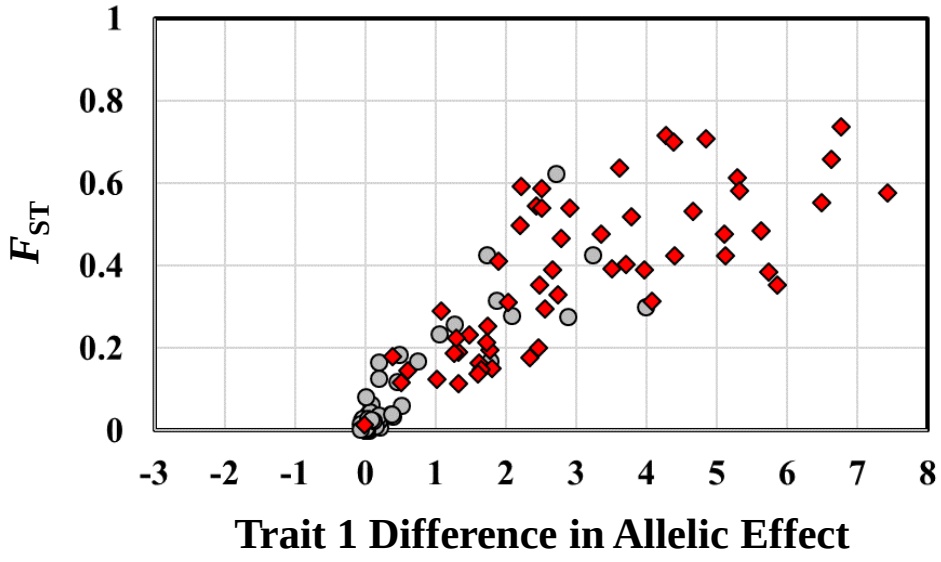
# Relationship between QTL $F_{ST}$ and between-population difference in mean allelic effects (no pleiotropy)

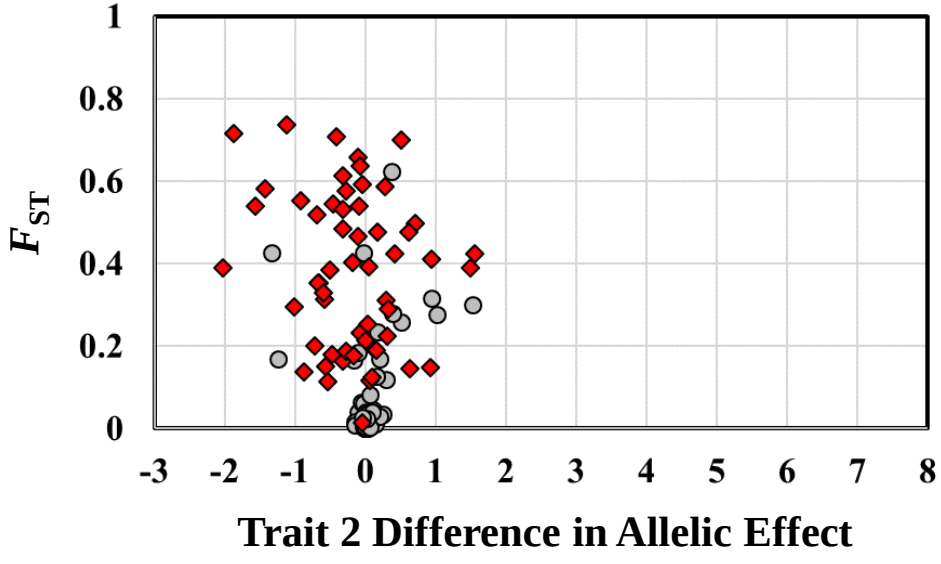



Data as above; red diamonds indicate QTL identified as outlier

# Patterns of $F_{ST}$ for pleiotropic QTL without epistasis




30 simulations, $m = 0.016$; red diamonds indicate loci identified as outlier

# Conclusions: Means

- Population means converge to values determined by migration-selection balance that are essentially independent of epistasis and pleiotropy
- Increasing migration leads to strong displacement from the optimum and to a dramatic increase in genetic variance
- Most divergence and increase in variance is caused by a small number of QTL, as the majority of QTL have small $F_{ST}$ values and contribute little to (additive) variance.

# Conclusions: Variances

- Most of the genetic variance is additive genetic variance, even if epistasis is very strong (more than 85%)
- Epistasis and pleiotropy cause a considerable increase of the variance of trait 2, in particular, for weak to moderate migration
- Epistasis and pleiotropy may cause a slight decrease of the variance of trait 1

# Conclusions: Outliers

- Under most circumstances, some outlier loci were detected, in no case all
- Even with 20 QTL per trait (5 per linkage group), our analyses never identified more than 2.5 true QTL
- More markers per linkage group increase the number of false positives substantially, but the number of true positives only slightly.
- The number of detected QTL depends only weakly on the recombination rate (0.1 - 4), selection intensity, population size (250 - 4000), and sample size (10 - 500)
- Pleiotropy had little influence on outlier detection
- Epistasis tends to reduce $F_{ST}$ and makes the causal QTL less detectable

THANK YOU!

# Table 6

| Variable of Interest and Its Value | $\sigma_\varepsilon^2$ | Mean Marker $F_{ST}$ | Mean Trt 1 QTL $F_{ST}$ | Mean Trt 2 QTL $F_{ST}$ | No. Smoothed $F_{ST}$ Outliers | No. Near Trt 1 QTL | No. Near Trt 2 QTL | No. W&L $F_{ST}$ Outliers | No. Near Trt 1 QTL | No. Near Trt 2 QTL |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Carrying Capacity*** | | | | | | | | | | |
| $K = 250$ | 0 | 0.0172 | 0.1046 | 0.0175 | 4.800 | 1.700 | 0.133 | 3.533 | 1.433 | 0.067 |
| $K = 250$ | 1.6 | 0.0172 | 0.0461 | 0.0484 | 5.300 | 0.700 | 0.867 | 3.500 | 0.633 | 0.733 |
| $K = 4000$ | 0 | 0.0380 | 0.1415 | 0.0382 | 4.333 | 2.433 | 0.200 | 5.000 | 2.333 | 0.200 |
| $K = 4000$ | 1.6 | 0.0409 | 0.1499 | 0.0741 | 3.967 | 1.700 | 0.467 | 4.867 | 1.533 | 0.667 |
| | | | | | | | | | | |
| ***Sample Size*** | | | | | | | | | | |
| $S = 10$ | 0 | 0.0433 | 0.1641 | 0.0445 | 4.767 | 1.733 | 0.300 | 3.500 | 1.200 | 0.100 |
| S = 10 | 1.6 | 0.0453 | 0.0868 | 0.0810 | 4.500 | 0.800 | 0.733 | 3.667 | 0.500 | 0.467 |
| $S = 500$ | 0 | 0.0271 | 0.1271 | 0.0284 | 4.700 | 1.967 | 0.333 | 6.333 | 2.067 | 0.567 |
| $S = 500$ | 1.6 | 0.0300 | 0.0765 | 0.0795 | 4.767 | 0.900 | 0.667 | 6.000 | 0.933 | 0.833 |
| | | | | | | | | | | |
| ***Selection Strength*** | | | | | | | | | | |
| $\omega_{11} = 19$ | 0 | 0.0922 | 0.3091 | 0.0757 | 4.500 | 1.633 | 0.267 | 2.133 | 0.600 | 0.167 |
| $\omega_{11} = 19$ | 1.6 | 0.1081 | 0.2356 | 0.1925 | 4.267 | 1.000 | 0.633 | 1.767 | 0.067 | 0.067 |
| $\omega_{11} = 99$ | 0 | 0.0177 | 0.0761 | 0.0163 | 4.633 | 1.567 | 0.133 | 6.967 | 1.733 | 0.400 |
| $\omega_{11} = 99$ | 1.6 | 0.0182 | 0.0550 | 0.0382 | 4.467 | 0.867 | 0.600 | 6.067 | 0.867 | 0.667 |
| | | | | | | | | | | |
| ***Epistasis Amount*** | | | | | | | | | | |
| $\sigma_\varepsilon^2 = 0.4$ | 0.4 | 0.0306 | 0.1075 | 0.0480 | 4.667 | 1.200 | 0.433 | 5.467 | 1.200 | 0.567 |
| $\sigma_\varepsilon^2 = 0.8$ | 0.8 | 0.0296 | 0.1005 | 0.0605 | 4.900 | 1.167 | 0.567 | 6.000 | 1.233 | 0.600 |
| $\sigma_\varepsilon^2 = 3.2$ | 3.2 | 0.0313 | 0.0752 | 0.686 | 4.433 | 0.900 | 0.633 | 5.867 | 0.900 | 0.600 |
| $\sigma_\varepsilon^2 = 6.4$ | 6.4 | 0.308 | 0.0879 | 0.0710 | 4.433 | 0.933 | 0.633 | 5.933 | 0.933 | 0.633 |
| | | | | | | | | | | |

# Table 6

| ***No. Marker Loci per Linkage Grp*** | $\sigma_\varepsilon^2$ | **Mean Marker $F_{ST}$** | **Mean Trt 1 QTL $F_{ST}$** | **Mean Trt 2 QTL $F_{ST}$** | **No. Smoothed $F_{ST}$ Outliers** | **No. Near Trt 1 QTL** | **No. Near Trt 2 QTL** | **No. W&L $F_{ST}$ Outliers** | **No. Near Trt 1 QTL** | **No. Near Trt 2 QTL** |
|---|---|---|---|---|---|---|---|---|---|---|
| $n_m = 500$ | 0 | 0.0305 | 0.1459 | 0.0270 | 1.700 | 1.100 | 0.467 | 2.333 | 1.533 | 0.533 |
| $n_m = 500$ | 1.6 | 0.0306 | 0.1035 | 0.0633 | 1.533 | 0.833 | 0.533 | 2.467 | 1.233 | 0.867 |
| $n_m = 10{,}000$ | 0 | 0.0285 | 0.1399 | 0.0326 | 14.800 | 1.833 | 0.167 | 17.067 | 1.633 | 0.133 |
| $n_m = 10{,}000$ | 1.6 | 0.0313 | 0.1021 | 0.0575 | 16.333 | 0.767 | 0.433 | 15.400 | 0.800 | 0.333 |
| | | | | | | | | | | |
| ***No. QTL per Linkage Grp*** | | | | | | | | | | |
| $n_{q1} = n_{q2} = 2$ | 0 | 0.0289 | 0.1372 | 0.0291 | 4.900 | 2.033 | 0.267 | 5.967 | 2.100 | 0.333 |
| $n_{q1} = n_{q2} = 2$ | 1.6 | 0.0301 | 0.0984 | 0.0601 | 4.933 | 0.933 | 0.667 | 5.733 | 0.900 | 0.833 |
| $n_{q1} = n_{q2} = 5$ | 0 | 0.0273 | 0.0730 | 0.0284 | 4.467 | 2.333 | 0.533 | 5.267 | 2.333 | 0.700 |
| $n_{q1} = n_{q2} = 5$ | 1.6 | 0.0287 | 0.0544 | 0.0407 | 4.900 | 1.467 | 0.900 | 5.267 | 1.367 | 1.100 |
| | | | | | | | | | | |
| ***Recomb. Rate*** | | | | | | | | | | |
| $R = 0.10$ | 0 | 0.0394 | 0.1442 | 0.0350 | 4.300 | 1.500 | 0.167 | 2.933 | 0.900 | 0.033 |
| $R = 0.10$ | 1.6 | 0.0401 | 0.1104 | 0.0687 | 4.267 | 0.867 | 0.633 | 3.200 | 0.467 | 0.333 |
| $R = 4.00$ | 0 | 0.0209 | 0.1166 | 0.0180 | 4.133 | 1.567 | 0.100 | 9.567 | 1.700 | 0.633 |
| $R = 4.00$ | 1.6 | 0.0222 | 0.0827 | 0.0466 | 3.567 | 0.967 | 0.633 | 9.300 | 1.200 | 0.767 |
| | | | | | | | | | | |

# Relationship between the between-population difference in mean allelic effects and the within-population variance in allelic effects with and without pleiotropy and epistasis