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

Reinhard Bürger

Department of Mathematics

Vienna, 14 February 2019, SMBE Satellite Meeting

‘Towards an integrated concept of adaptation: uniting molecular population genetics and quantitative genetics’
Collaborators

Adam G. Jones

Stevan J. Arnold

1AG 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.

![Diagram showing selection and migration](image)

Optimum at $(z_1, z_2) = (-4,0)$
Optimum at $(z_1, z_2) = (4,0)$

(Units are environmental standard deviations)
Simulation model

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

- Life cycle:

  >\( N \) survivors per deme
  
  Random choice of \( N \) parents in each deme (population regulation)
  
  Random mating within demes, Mendelian assortment, mutation, recombination
  
  Migration

  Population of \( N \) adults per deme
  
  Population of \( N \) individuals per deme
  
  \( 2N \) progeny per deme
  
  \( 2N \) individuals per deme
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 $R$
- $R = \text{Expected number of recombination events per meiosis per individual; } 0.1 \leq R \leq 4; \text{ typical value } R = 0.25$
- 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 \((y^{(i)}_1, y^{(i)}_2)\) 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

\[
x_1 = \xi_1 + y^{(1)}_1 + y^{(2)}_1 + \epsilon^{(1,2)}_{111} y^{(1)}_1 y^{(2)}_1
\]
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 $4n(n - 1)$ pairwise epistatic coefficients $\epsilon_{(i,j)}^{abc}$
- We draw the coefficients $\epsilon_{(i,j)}^{abc}$ from a normal distribution with
  
  $\text{E}(\epsilon) = 0$ and $\text{Var}(\epsilon) = \sigma^2_{\epsilon}$

  (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

Deme 1

\[ z_2 \]

Deme 2

No Epistasis

\[ Nm = 1 \]

Epistasis

\[ \text{Var}(\varepsilon) = 1.6 \]

\[ z_1 \]
Epistasis and local adaptation: strong migration

Deme 1

$z_2$

$z_1$

Deme 2

No Epistasis

Epistasis

$\text{Var}(\varepsilon) = 1.6$

$Nm = 32$

$Nm = 32$

$Nm = 32$
Local adaptation: epistasis but no pleiotropy

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

\[ Z_1 \]

\[ m \]

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:

\[
\bar{Z}_1 = \frac{1}{N} \sum_{i=1}^{N} Z_i
\]

- \( \sigma^2_{\epsilon} \) represents the variance of the error term:
  - 0
  - 1.6

- 0, pleiotropy
- 1.6, pleiotropy

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

\[ V_1, V_2 \]

\[ \sigma^2 \]

\[ \sigma^2_{\epsilon} \]

\[ \text{Variances of traits 1 and 2} \]

\[ \text{trait optima at } (z_1, z_2) = (\pm 4, 0); N = 500; \text{ each trait determined by 4 QTL.} \]
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 $(z_1, z_2) = (±4, 0)$; $N = 500$; 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 $m = 0.016$

$\sigma^2_\epsilon = 0$

$\sigma^2_\epsilon = 1.6$
Number of outliers and “true” outliers

30 simulation runs with 4 linkage groups and 1 QTL per group → 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

![Graph showing patterns of $F_{ST}$](image)

<table>
<thead>
<tr>
<th>No. QTL</th>
<th>$F_{ST}$</th>
<th>$F_{ST}$ with Trait 1</th>
<th>$F_{ST}$ with Trait 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>30</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>50</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Variable of Interest and Its Value</th>
<th>$\sigma^2_e$</th>
<th>Mean Marker $F_{ST}$</th>
<th>Mean Trt 1 QTL $F_{ST}$</th>
<th>Mean Trt 2 QTL $F_{ST}$</th>
<th>No. Smoothed $F_{ST}$ Outliers</th>
<th>No. Near Trt 1 QTL</th>
<th>No. Near Trt 2 QTL</th>
<th>No. W&amp;L $F_{ST}$ Outliers</th>
<th>No. Near Trt 1 QTL</th>
<th>No. Near Trt 2 QTL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Carrying Capacity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 250$</td>
<td>0</td>
<td>0.0172</td>
<td>0.1046</td>
<td>0.0175</td>
<td>4.800</td>
<td>1.700</td>
<td>0.133</td>
<td>3.533</td>
<td>1.433</td>
<td>0.067</td>
</tr>
<tr>
<td>$K = 250$</td>
<td>1.6</td>
<td>0.0172</td>
<td>0.0461</td>
<td>0.0484</td>
<td>5.300</td>
<td>0.700</td>
<td>0.867</td>
<td>3.500</td>
<td>0.633</td>
<td>0.733</td>
</tr>
<tr>
<td>$K = 4000$</td>
<td>0</td>
<td>0.0380</td>
<td>0.1415</td>
<td>0.0382</td>
<td>4.333</td>
<td>2.433</td>
<td>0.200</td>
<td>5.000</td>
<td>2.333</td>
<td>0.200</td>
</tr>
<tr>
<td>$K = 4000$</td>
<td>1.6</td>
<td>0.0409</td>
<td>0.1499</td>
<td>0.0741</td>
<td>3.967</td>
<td>1.700</td>
<td>0.467</td>
<td>4.867</td>
<td>1.533</td>
<td>0.667</td>
</tr>
<tr>
<td><strong>Sample Size</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S = 10$</td>
<td>0</td>
<td>0.0433</td>
<td>0.1641</td>
<td>0.0445</td>
<td>4.767</td>
<td>1.733</td>
<td>0.300</td>
<td>3.500</td>
<td>1.200</td>
<td>0.100</td>
</tr>
<tr>
<td>$S = 10$</td>
<td>1.6</td>
<td>0.0453</td>
<td>0.0868</td>
<td>0.0810</td>
<td>4.500</td>
<td>0.800</td>
<td>0.733</td>
<td>3.667</td>
<td>0.500</td>
<td>0.467</td>
</tr>
<tr>
<td>$S = 500$</td>
<td>0</td>
<td>0.0271</td>
<td>0.1271</td>
<td>0.0284</td>
<td>4.700</td>
<td>1.967</td>
<td>0.333</td>
<td>6.333</td>
<td>2.067</td>
<td>0.567</td>
</tr>
<tr>
<td>$S = 500$</td>
<td>1.6</td>
<td>0.0300</td>
<td>0.0765</td>
<td>0.0795</td>
<td>4.767</td>
<td>0.900</td>
<td>0.667</td>
<td>6.000</td>
<td>0.933</td>
<td>0.833</td>
</tr>
<tr>
<td><strong>Selection Strength</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\omega_{11} = 19$</td>
<td>0</td>
<td>0.0922</td>
<td>0.3091</td>
<td>0.0757</td>
<td>4.500</td>
<td>1.633</td>
<td>0.267</td>
<td>2.133</td>
<td>0.600</td>
<td>0.167</td>
</tr>
<tr>
<td>$\omega_{11} = 19$</td>
<td>1.6</td>
<td>0.1081</td>
<td>0.2356</td>
<td>0.1925</td>
<td>4.267</td>
<td>1.000</td>
<td>0.633</td>
<td>1.767</td>
<td>0.067</td>
<td>0.067</td>
</tr>
<tr>
<td>$\omega_{11} = 99$</td>
<td>0</td>
<td>0.0177</td>
<td>0.0761</td>
<td>0.0163</td>
<td>4.633</td>
<td>1.567</td>
<td>0.133</td>
<td>6.967</td>
<td>1.733</td>
<td>0.400</td>
</tr>
<tr>
<td>$\omega_{11} = 99$</td>
<td>1.6</td>
<td>0.0182</td>
<td>0.0550</td>
<td>0.0382</td>
<td>4.467</td>
<td>0.867</td>
<td>0.600</td>
<td>6.067</td>
<td>0.867</td>
<td>0.667</td>
</tr>
<tr>
<td><strong>Epistasis Amount</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma^2_e = 0.4$</td>
<td>0.4</td>
<td>0.0306</td>
<td>0.1075</td>
<td>0.0480</td>
<td>4.667</td>
<td>1.200</td>
<td>0.433</td>
<td>5.467</td>
<td>1.200</td>
<td>0.567</td>
</tr>
<tr>
<td>$\sigma^2_e = 0.8$</td>
<td>0.8</td>
<td>0.0296</td>
<td>0.1005</td>
<td>0.0605</td>
<td>4.900</td>
<td>1.167</td>
<td>0.567</td>
<td>6.000</td>
<td>1.233</td>
<td>0.600</td>
</tr>
<tr>
<td>$\sigma^2_e = 3.2$</td>
<td>3.2</td>
<td>0.0313</td>
<td>0.0752</td>
<td>0.686</td>
<td>4.433</td>
<td>0.900</td>
<td>0.633</td>
<td>5.867</td>
<td>0.900</td>
<td>0.600</td>
</tr>
<tr>
<td>$\sigma^2_e = 6.4$</td>
<td>6.4</td>
<td>0.308</td>
<td>0.0879</td>
<td>0.0710</td>
<td>4.433</td>
<td>0.933</td>
<td>0.633</td>
<td>5.933</td>
<td>0.933</td>
<td>0.633</td>
</tr>
</tbody>
</table>
Table 6

<table>
<thead>
<tr>
<th>No. Marker Loci per Linkage Grp</th>
<th>$\sigma_e^2$</th>
<th>Mean Marker $F_{ST}$</th>
<th>Mean Trt 1 QTL $F_{ST}$</th>
<th>Mean Trt 2 QTL $F_{ST}$</th>
<th>No. Smoothed $F_{ST}$ Outliers</th>
<th>No. Near Trt 1 QTL $F_{ST}$</th>
<th>No. Near Trt 2 QTL $F_{ST}$</th>
<th>No. W&amp;L $F_{ST}$ Outliers</th>
<th>No. Near Trt 1 QTL</th>
<th>No. Near Trt 2 QTL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_m = 500$</td>
<td>0</td>
<td>0.0305</td>
<td>0.1459</td>
<td>0.0270</td>
<td>1.700</td>
<td>1.100</td>
<td>0.467</td>
<td>2.333</td>
<td>1.533</td>
<td>0.533</td>
</tr>
<tr>
<td>$n_m = 500$</td>
<td>1.6</td>
<td>0.0306</td>
<td>0.1035</td>
<td>0.0633</td>
<td>1.533</td>
<td>0.833</td>
<td>0.533</td>
<td>2.467</td>
<td>1.233</td>
<td>0.867</td>
</tr>
<tr>
<td>$n_m = 10,000$</td>
<td>0</td>
<td>0.0285</td>
<td>0.1399</td>
<td>0.0326</td>
<td>14.800</td>
<td>1.833</td>
<td>0.167</td>
<td>17.067</td>
<td>1.633</td>
<td>0.133</td>
</tr>
<tr>
<td>$n_m = 10,000$</td>
<td>1.6</td>
<td>0.0313</td>
<td>0.1021</td>
<td>0.0575</td>
<td>16.333</td>
<td>0.767</td>
<td>0.433</td>
<td>15.400</td>
<td>0.800</td>
<td>0.333</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. QTL per Linkage Grp</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_{q1} = n_{q2} = 2$</td>
<td>0</td>
<td>0.0289</td>
<td>0.1372</td>
<td>0.0291</td>
<td>4.900</td>
<td>2.033</td>
<td>0.267</td>
<td>5.967</td>
<td>2.100</td>
<td>0.333</td>
</tr>
<tr>
<td>$n_{q1} = n_{q2} = 2$</td>
<td>1.6</td>
<td>0.0301</td>
<td>0.0984</td>
<td>0.0601</td>
<td>4.933</td>
<td>0.933</td>
<td>0.667</td>
<td>5.733</td>
<td>0.900</td>
<td>0.833</td>
</tr>
<tr>
<td>$n_{q1} = n_{q2} = 5$</td>
<td>0</td>
<td>0.0273</td>
<td>0.0730</td>
<td>0.0284</td>
<td>4.467</td>
<td>2.333</td>
<td>0.533</td>
<td>5.267</td>
<td>2.333</td>
<td>0.700</td>
</tr>
<tr>
<td>$n_{q1} = n_{q2} = 5$</td>
<td>1.6</td>
<td>0.0287</td>
<td>0.0544</td>
<td>0.0407</td>
<td>4.900</td>
<td>1.467</td>
<td>0.900</td>
<td>5.267</td>
<td>1.367</td>
<td>1.100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Recomb. Rate</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$R = 0.10$</td>
<td>0</td>
<td>0.0394</td>
<td>0.1442</td>
<td>0.0350</td>
<td>4.300</td>
<td>1.500</td>
<td>0.167</td>
<td>2.933</td>
<td>0.900</td>
<td>0.033</td>
</tr>
<tr>
<td>$R = 0.10$</td>
<td>1.6</td>
<td>0.0401</td>
<td>0.1104</td>
<td>0.0687</td>
<td>4.267</td>
<td>0.867</td>
<td>0.633</td>
<td>3.200</td>
<td>0.467</td>
<td>0.333</td>
</tr>
<tr>
<td>$R = 4.00$</td>
<td>0</td>
<td>0.0209</td>
<td>0.1166</td>
<td>0.0180</td>
<td>4.133</td>
<td>1.567</td>
<td>0.100</td>
<td>9.567</td>
<td>1.700</td>
<td>0.633</td>
</tr>
<tr>
<td>$R = 4.00$</td>
<td>1.6</td>
<td>0.0222</td>
<td>0.0827</td>
<td>0.0466</td>
<td>3.567</td>
<td>0.967</td>
<td>0.633</td>
<td>9.300</td>
<td>1.200</td>
<td>0.767</td>
</tr>
</tbody>
</table>
Relationship between the between-population difference in mean allelic effects and the within-population variance in allelic effects with and without pleiotropy and epistasis