Detecting adaptive differentiation in structured populations with genomic data and common gardens

Emily Josephs

@emjosephs
Detecting local adaptation requires knowing about (1) genetic variation in traits

Photo credits: IRRI, Childs et al., Poland et al

Northern leaf blight in maize

Pop. A

Pop. B

Common garden
Detecting local adaptation requires knowing about (1) genetic variation in traits ($V_A$).

[Image of plants representing populations A and B and a common garden]

Crosses to get $V_A$
Detecting local adaptation requires knowing about (1) genetic variation and (2) relatedness.
Detecting local adaptation requires knowing about (1) genetic variation and (2) relatedness.
$Q_{st} - F_{st}$ comparisons test for excess trait divergence

<table>
<thead>
<tr>
<th>$Q_{st}$</th>
<th>vs.</th>
<th>$F_{st}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>btw-pop genetic var. for trait</td>
<td></td>
<td>btw-pop neutral genetic var.</td>
</tr>
<tr>
<td>total genetic var. for trait</td>
<td></td>
<td>total neutral genetic var.</td>
</tr>
</tbody>
</table>

Prout and Barker 1993, Spitze 1993, Whitlock 1999
$Q_{st} - F_{st}$ comparisons test for excess trait divergence

\[
\frac{Q_{st}}{F_{st}} = \frac{\text{btw-pop genetic var. for trait}}{\text{total genetic var. for trait}} = \frac{\text{btw-pop neutral genetic var.}}{\text{total neutral genetic var.}}
\]

Prout and Barker 1993, Spitze 1993, Whitlock 1999
$Q_{st} - F_{st}$ comparisons test for excess trait divergence

\[
\frac{Q_{st}}{\text{btw-pop genetic var. for trait}} \ > \ \frac{F_{st}}{\text{btw-pop neutral genetic var.}}
\]

\[
\frac{\text{total genetic var. for trait}}{\text{total neutral genetic var.}}
\]

Prout and Barker 1993, Spitze 1993, Whitlock 1999
Lots of datasets that have genomes + phenotypes from diversity panels
How can we systematically detect local adaptation...

In quantitative traits?

In gene expression?

For GxE?
Relatedness between populations can be summarized with a kinship matrix.

Lighter colors = more related
The eigenvectors of the kinship matrix (PCs) summarize relatedness.
How do PCs relate to traits?
A correlation btw PC & trait can be consistent with drift
A correlation btw PC & trait can be consistent with drift
Modelling the slope expected due to drift

\[ q_m \sim N(0, \lambda_m V_A) \]

Slope of trait against PC m
Modelling the slope expected due to drift

\[ q_m \sim N(0, \lambda_m V_A) \]

Slope of trait against PC m  Mean slope

Size
Modelling the slope expected due to drift

\[ q_m \sim N(0, \lambda_m V_A) \]

Slope of trait against PC m
Mean slope
Amount of relatedness explained by PC m
Modelling the slope expected due to drift

\[ q_m \sim N(0, \lambda_m V_A) \]

- Slope of trait against PC m
- Mean slope
- Can estimate \( V_A \) from a subset of PCs
- Amount of relatedness explained by PC m
Selection can increase trait divergence

\[ q_m \sim N(0, \lambda_m V_A) \]

Slope of trait against PC \( m \)
Mean slope
Amount of relatedness explained by PC \( m \)

Can estimate \( V_A \) from a subset of PCs
Selection can increase trait divergence

\[ q_m \sim N(0, \lambda_m V_A) \]

Slope of trait against PC m
Mean slope
Can estimate \( V_A \) from a subset of PCs
Amount of relatedness explained by PC m
Testing for diversifying selection

Size

\[ q_m \sim N(0, \lambda_m V_A) \]

Slope of trait against PC m

Mean slope

Amount of relatedness explained by PC m

Can estimate \( V_A \) from a subset of PCs

\[ \frac{q_m^2}{V_A \lambda_m} \sim F(1, N) \]

Number of PCs used to estimate \( V_A \)
How can we systematically detect local adaptation...

In quantitative traits?

In gene expression?

For GxE?
Detecting local adaptation in maize
Detecting local adaptation in maize

240 domesticated inbred maize lines (Flint-Garcia et al. 2005).

Whole genome sequence data (Panzea).

22 traits measured in common garden.
The kinship matrix of 240 maize lines

Lighter colors = more related
The kinship matrix of 240 maize lines

Lighter colors = more related
Neutral expectations of relationship between PC and trait

![Graph showing the relationship between PC 1 and kernel number with neutral expectations.](image-url)
Divergence along PC1 consistent with drift
Neutral expectations of relationship between PC and trait

[Graph showing the relationship between PC 1 and Flowering time]
Adaptive divergence in flowering time along PC 1

FDR < 0.05
Signatures of adaptation across traits

White circles mean FDR < 0.05
How can we systematically detect local adaptation...

In quantitative traits?

In gene expression?

For GxE?
Testing for selection on gene expression

Gene expression data (RNAseq) for 208 maize lines in 7 tissues

Kremling et al. 2018
How many genes show evidence of selection on expression?
Which tissues show strongest evidence of adaptation?

- Kernel
- Leaf (day)
- Leaf (night)
### Evidence of local adaptation for gene expression

<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaf night (day 8)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leaf night (day 26)</td>
<td>4</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Leaf day (day 8)</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Leaf day (day 26)</td>
<td>1</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>3rd leaf tip</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd leaf base</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kernel</td>
<td></td>
<td>15</td>
<td>21</td>
</tr>
<tr>
<td>Seedling shoot</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seedling root</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of genes that show evidence of selection on expression.

Blanc et al. in prep

Jennifer Blanc
How can we systematically detect local adaptation...

In quantitative traits?

In gene expression?

For GxE?
The environment shapes traits
Genetic variation for plasticity (‘GxE’)

Clausen, Keck, and Heisey
A. thaliana flowering time as a model for GxE
Flowering time at 16°C is locally adapted

$p = 0.0002$

Bonferoni $p = 0.0015$
Flowering time at 10°C is also locally adapted

\[ p = 0.006 \]
\[ \text{Bonferoni } p = 0.03 \]
Using the reaction norm to measure plasticity
Using the reaction norm to measure plasticity

Days to flower

Temperature

Slopes
Variation in reaction norm consistent with drift

Later flowering at 16°C than 10°C

Earlier flowering at 16°C than 10°C

$p = 0.03$

Bonferoni p = 0.16
How to deal with trait correlations?
How to deal with trait correlations?
We can use PC + trait relationships to detect adaptation
We can use PC + trait relationships to detect adaptation

Local adaptation shapes trait divergence
We can use PC + trait relationships to detect adaptation

Local adaptation shapes trait divergence

These methods can be used in additional systems!!!
Thanks!

Jeremy Berg  Graham Coop  Karl Kremling
Jennifer Blanc  Jeff Ross-Ibarra  Ed Buckler

Cinta Romay  Kate Crosby  Coop Lab  Ross-Ibarra Lab

UC DAVIS UNIVERSITY OF CALIFORNIA
The Josephs lab at Michigan State University

Contact me! Josep993@msu.edu, http://JosephsLab.github.io/
We can use PC + trait relationships to detect adaptation

Local adaptation shapes trait divergence

These methods can be used in additional systems!!!

Detecting Adaptive Differentiation in Structured Populations with Genomic Data and Common Gardens
Emily B. Josephs, Jeremy J. Berg, Jeffrey Ross-Ibarra and Graham Coop
GENETICS Early online January 28, 2019; https://doi.org/10.1534/genetics.118.301786

https://github.com/emjosephs/quaint
Adaptive divergence along additional PCs

Flowering time

FDR < 0.05
What about detecting adaptation in genotypes that haven’t been phenotyped?
Testing for adaptation with polygenic scores

\[ Z = 2 \sum_{l=1}^{L} \alpha_l p_{ml} \]

- Polygenic score
- Effect size
- Allele frequency
- At all GWAS loci

Berg and Coop 2014
Evidence for polygenic adaptation in European maize

Josephs et al. 2018 Biorxiv