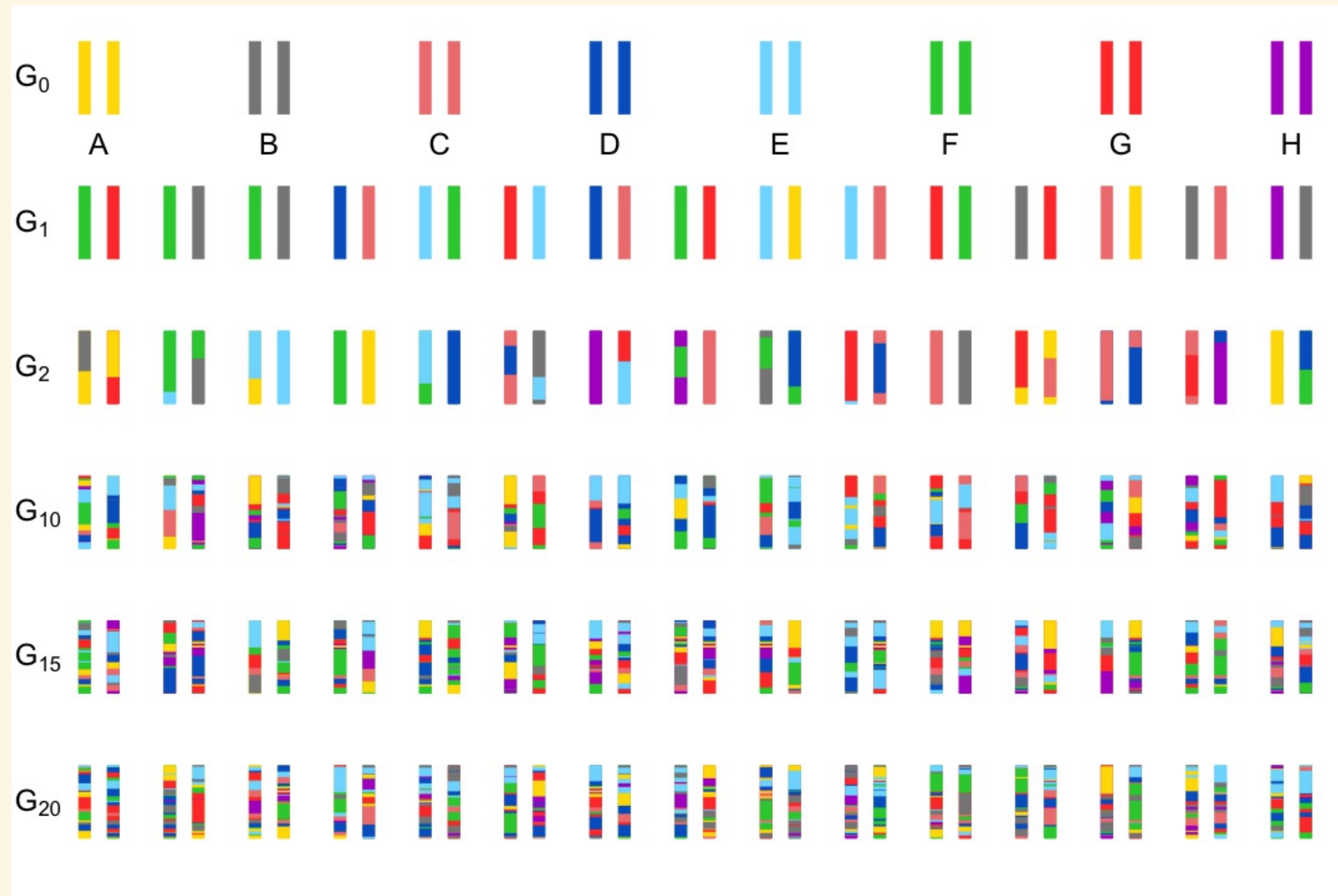


Testing pleiotropy in multiparental populations

Frederick Boehm

April 22, 2019

Multiparental populations



Background

- 10,000+ traits with RNA sequencing and mass spectrometry
- Multiparental populations offer high-resolution QTL mapping
- New analysis tools, such as a pleiotropy test for multiparental populations, are needed



Photo by Alan Attie

Benefits of a new pleiotropy test

- Insights into genetic architecture
- Tool for expression trait hotspot dissection
- Complements mediation analysis

Jiang and Zeng (1995) test

- Two-parent crosses
- Applies to two traits that co-map
- H_0 : Pleiotropy
- H_A : Two separate QTL

Jiang and Zeng (1995) test

- Perform a two-dimensional two-QTL scan
 - $\text{vec}(Y) = X\text{vec}(B) + \text{vec}(E)$
 - Calculate likelihood at each ordered pair of positions
- Calculate likelihood ratio test statistic

Challenges in multiparental populations

- Complex patterns of relatedness

Multivariate random effects

- Multiple founder lines

Fixed effect for each founder allele



Photo by UNC Computational Genetics

Test procedure

- Perform a two-dimensional two-QTL scan
 - $\text{vec}(Y) = X\text{vec}(B) + \text{vec}(G) + \text{vec}(E)$
 - Calculate likelihood at each ordered pair of positions
- Calculate likelihood ratio test statistic

Test procedure

- Test statistic:

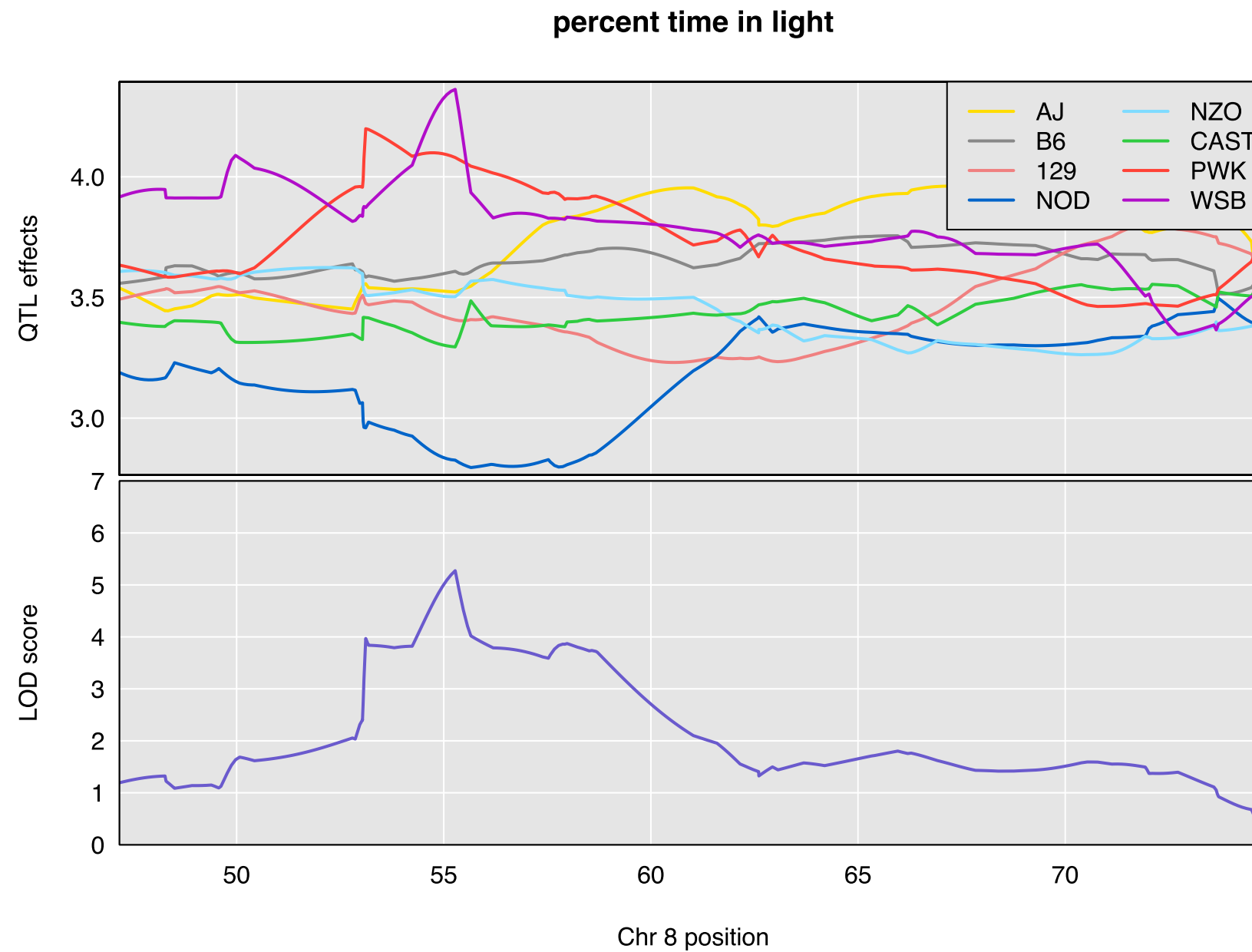
$$-\log_{10} \frac{\max(\text{likelihood under pleiotropy})}{\max(\text{likelihood for separate QTL})}$$

- Parametric bootstrap to get a p-value

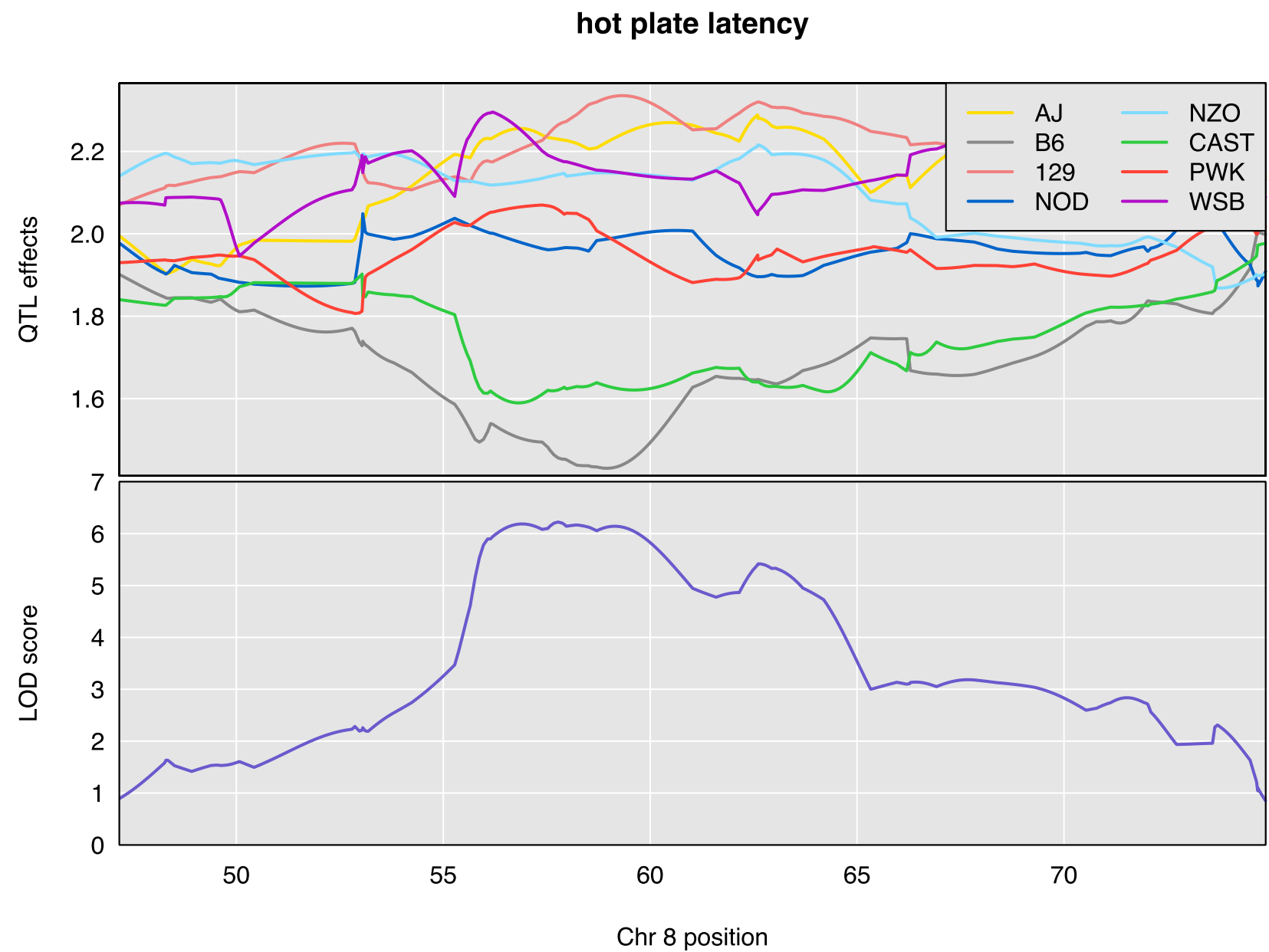
Application

- Logan, et al. (2013) and Recla, et al. (2014) studied 261 Diversity Outbred mice
- Measured about two dozen behavioral traits
- Two traits map to Chr 8:
 - "hot plate latency" (57 cM)
 - "percent time in light" (55 cM)

Percent time in light



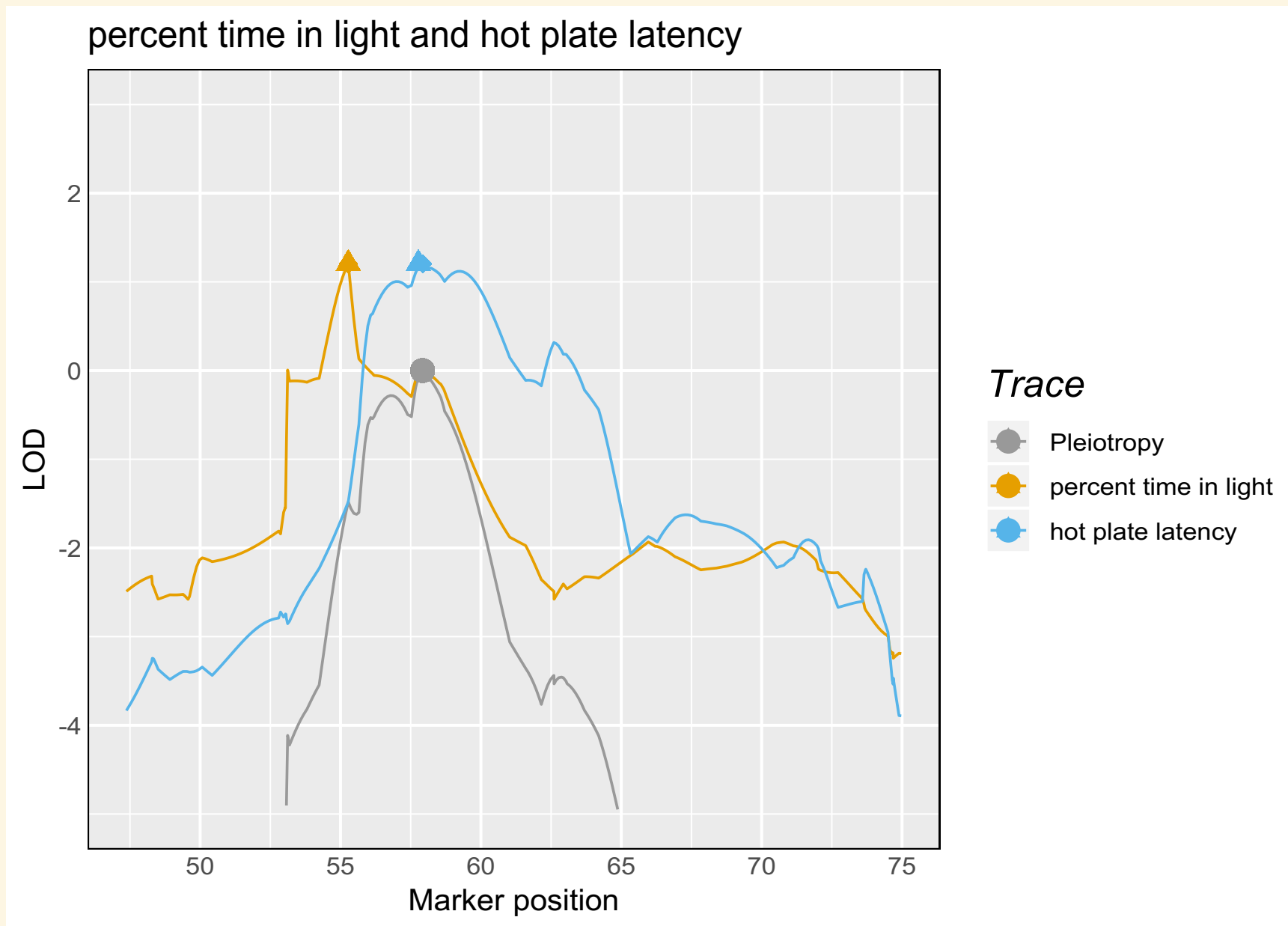
Hot plate latency



LOD definitions

- $\text{LOD}(\lambda_1, \lambda_2) = l_{10}(\lambda_1, \lambda_2) - \max l_{10}(\lambda, \lambda)$
- $\text{profile LOD}_{\text{trait } 1}(\lambda_1) = \max_{\lambda_2} \text{LOD}(\lambda_1, \lambda_2)$
- $\text{LOD}_p(\lambda) = l_{10}(\lambda, \lambda) - \max l_{10}(\lambda, \lambda)$

Profile LOD



Test results

- $\log_{10} \Lambda = 1.2$
- $p = 0.11$ (1000 bootstrap samples)

Power study with local expression QTL

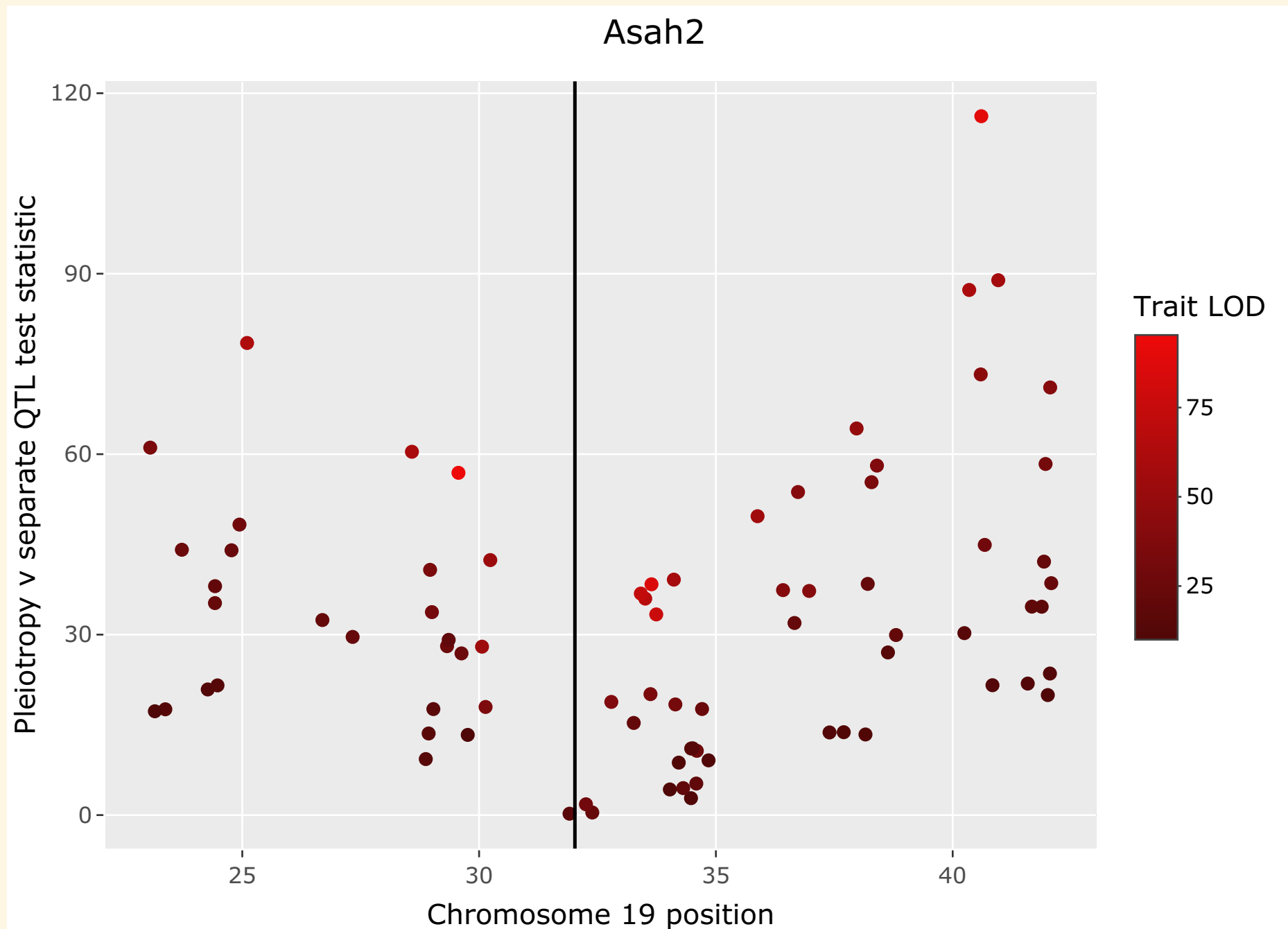
Data

- Keller, et al. (2018) measured pancreatic islet gene expression levels in 378 Diversity Outbred mice
- 80 local expression QTL were identified in a 20-Mb region of Chr 19
 - local: expression trait QTL is near gene position

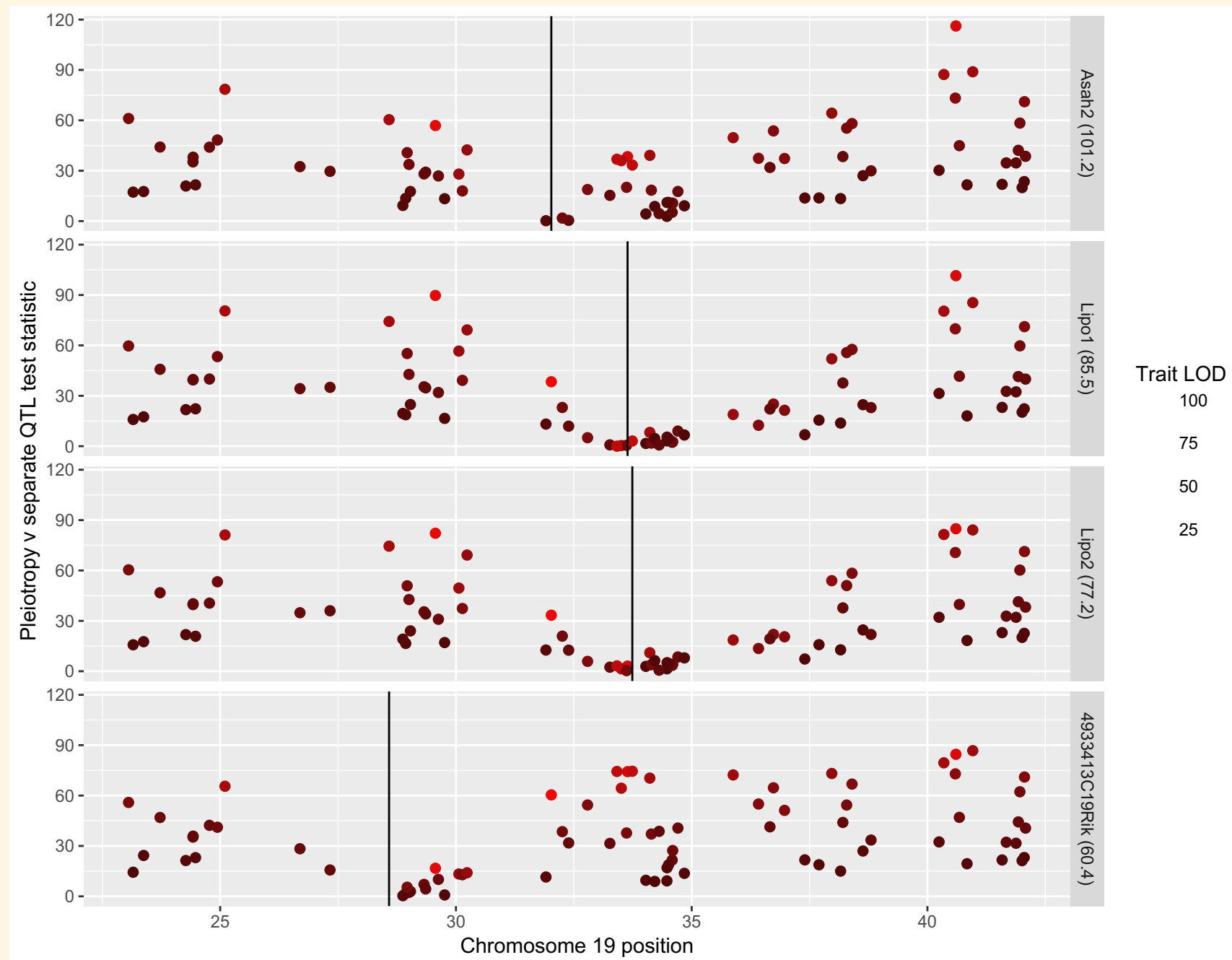
Design

- Examine power when expression trait QTL locations are known
 - Interlocus distance
 - Univariate LOD
- 4 strong, centrally located, local eQTL chosen as "anchor" traits
 - *Asah2*, *Lipo1*, *Lipo2*, *4933413C19Rik*
- Pairwise pleiotropy tests each involving one anchor and one of 79 other traits

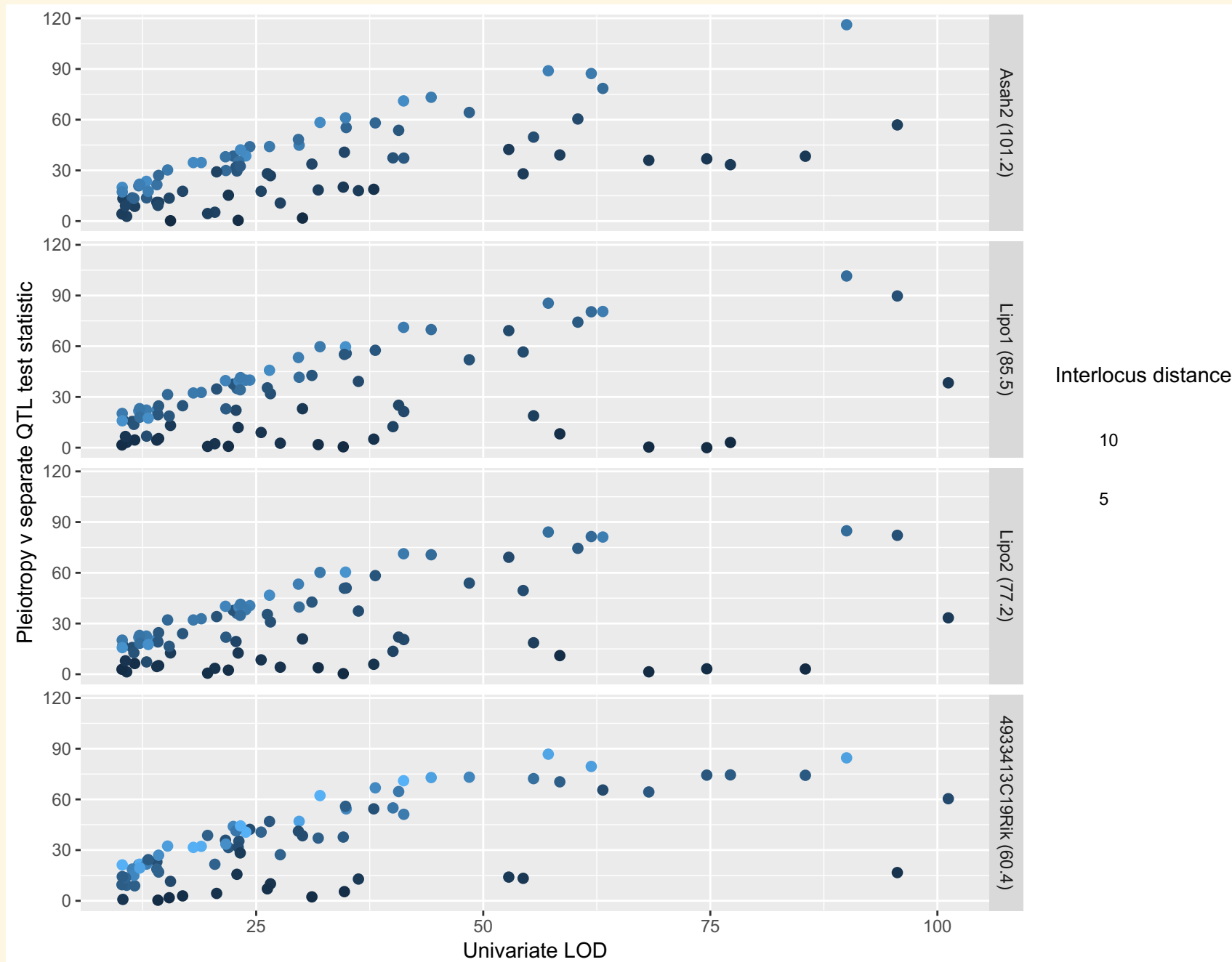
Interlocus distance



Interlocus distance



Univariate LOD



Conclusions

- ↑ Pleiotropy test statistics
 - ↑ Interlocus distance
 - ↑ Univariate LOD

qtl2pleio R package

qtl2pleio R package

- Functions for d-variate, d-QTL scan & profile LOD plots
- Uses C++ for matrix calculations (via Rcpp and RcppEigen)
- Uses [gemma2](#) R implementation of GEMMA EM algorithm for multivariate random effects
- Unit tests, vignettes, and version control

Summary

1. Background
2. Methods
3. Applications
 - a. Pleiotropy testing and mediation analysis
 - b. Power in pleiotropy testing
 - c. Microbiome case study
4. Software
5. Conclusions

References

Jiang, C. and Z. Zeng (1995). "Multiple trait analysis of genetic mapping for quantitative trait loci." In: *Genetics* 140.3, pp. 1111-1127.

Keller, M. P, D. M. Gatti, et al. (2018). "Genetic Drivers of Pancreatic Islet Function". In: *Genetics*, pp. genetics-300864.

Logan, R. W, R. F. Robledo, et al. (2013). "High-precision genetic mapping of behavioral traits in the diversity outbred mouse population". In: *Genes, Brain and Behavior* 12.4, pp. 424-437.