

# Testing pleiotropy in multiparental populations

Frederick Boehm

March 6, 2019

# Introduction

- 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

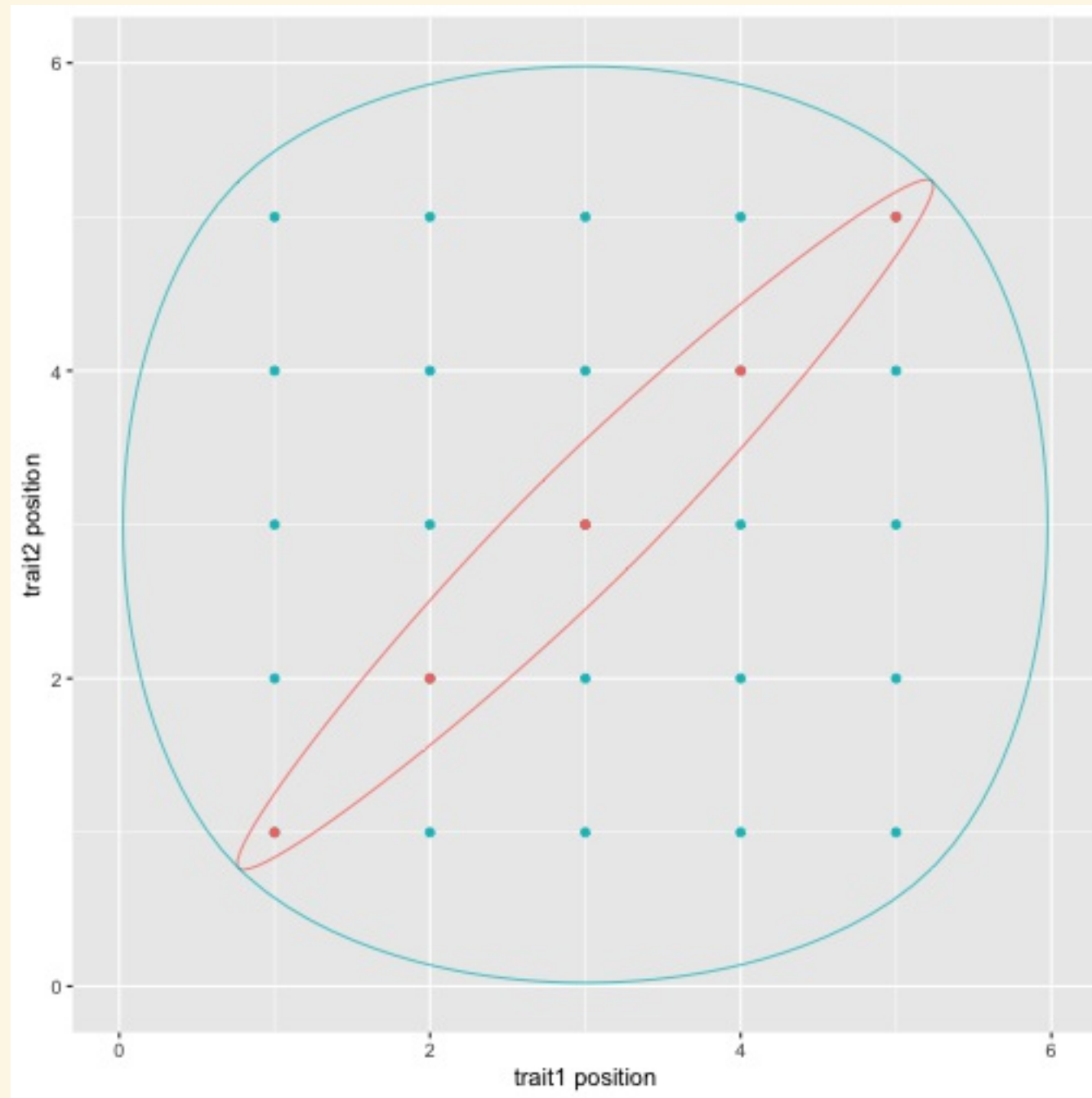
# Jiang and Zeng (1995) test

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

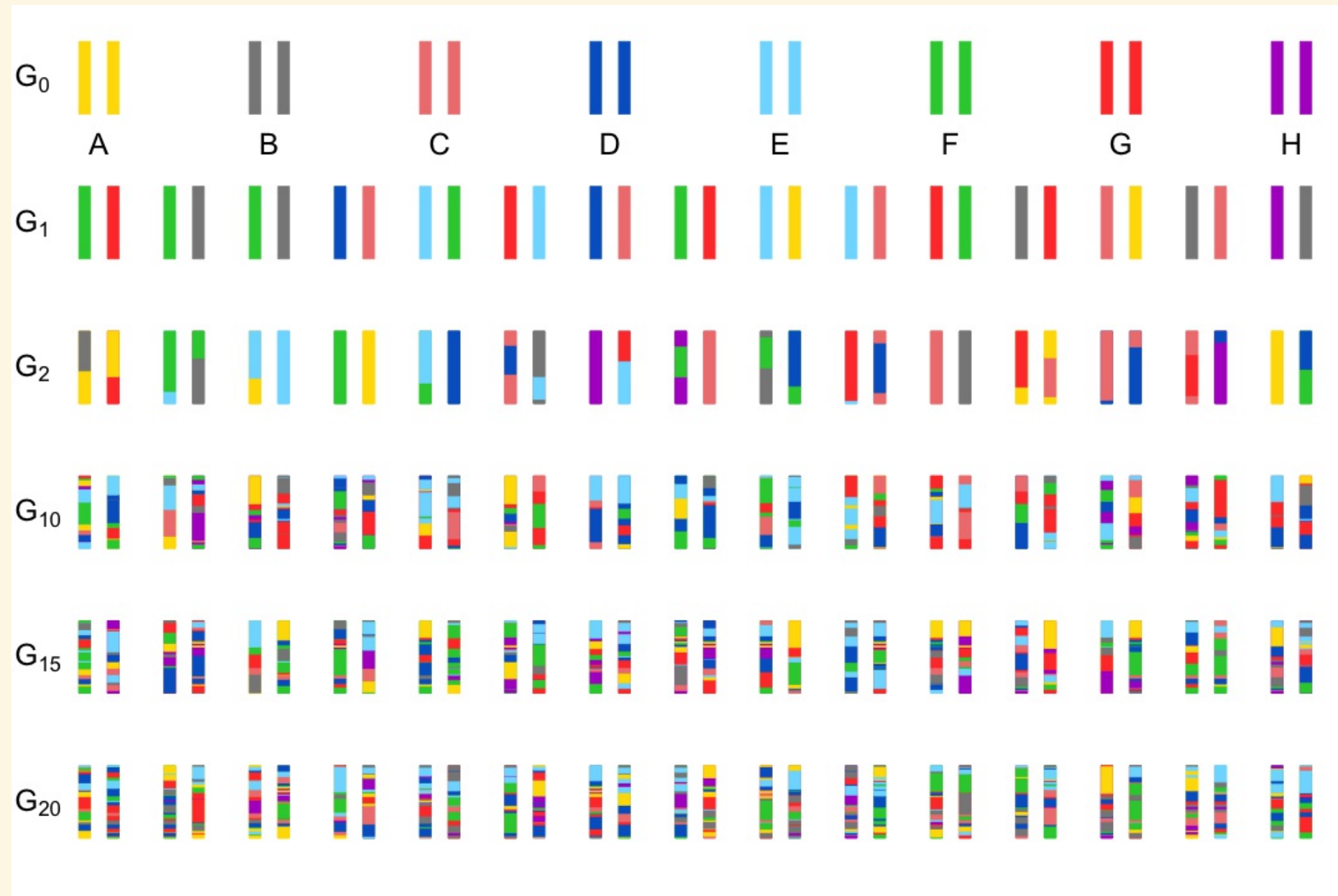
# Jiang and Zeng (1995) test

- Perform a two-dimensional QTL scan
- $Y = XB + E$
- Calculate likelihood ratio test statistic

# Jiang and Zeng (1995) test



# Multiparental populations



# Challenges in multiparental populations

- Complex patterns of relatedness
- Multiple founder lines
- Determining statistical significance

# Challenges in multiparental populations

- Complex patterns of relatedness
- Multiple founder lines
- Determining statistical significance



# Test procedure

- Model:

$$\text{vec}(Y) = X\text{vec}(B) + \text{vec}(G) + \text{vec}(E)$$

- Calculate likelihoods

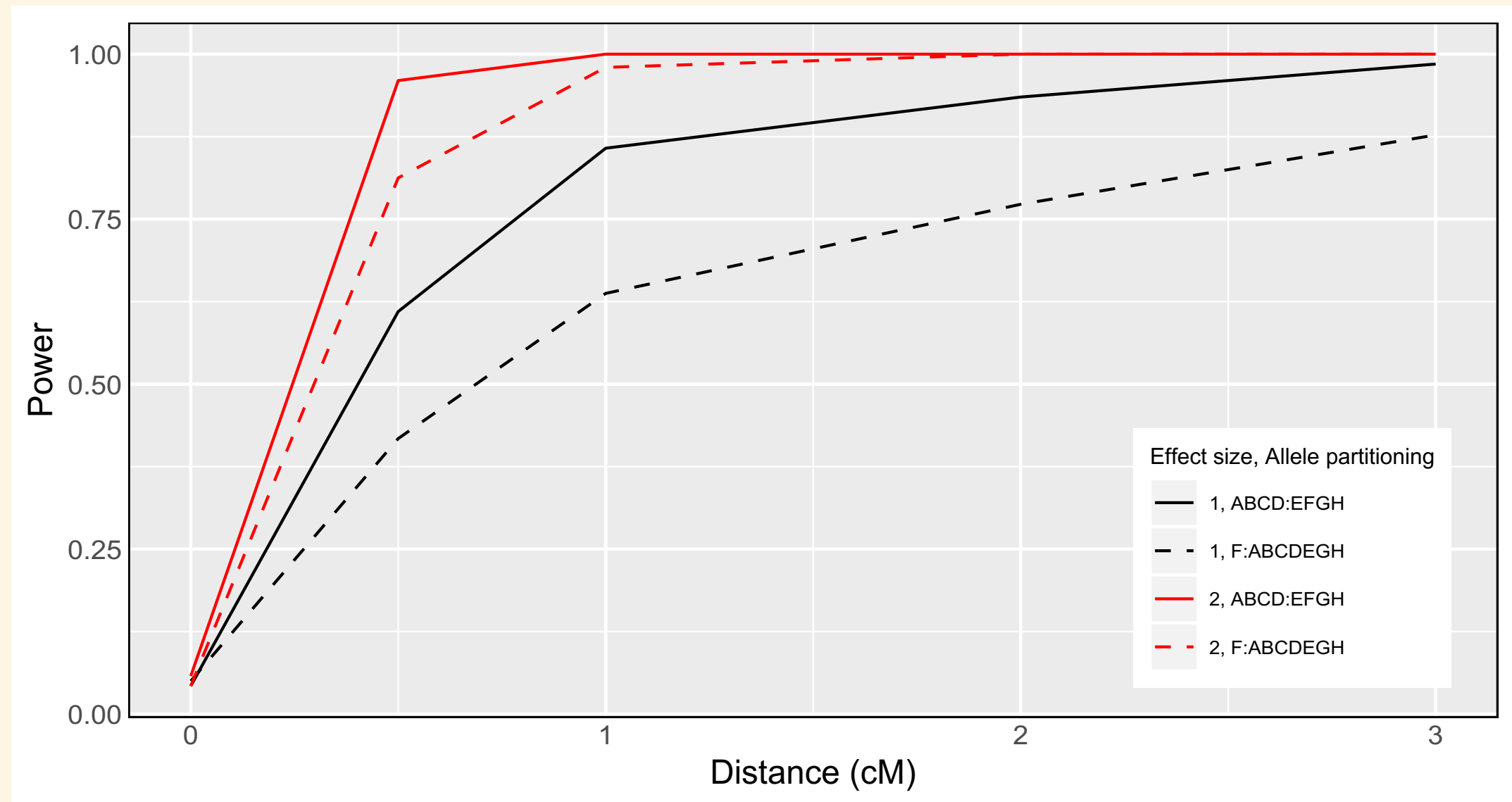
# Test procedure

- Test statistic:

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

- Parametric bootstrap to get p-value

# Test characteristics



# Diversity Outbred Mice

- Multiparental population arising from 8 inbred lines

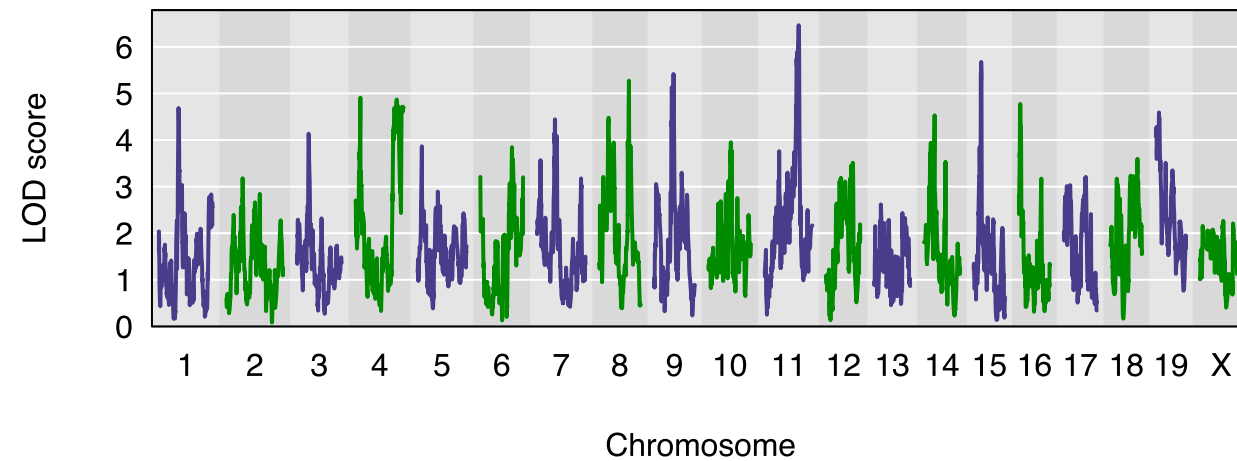


# Application

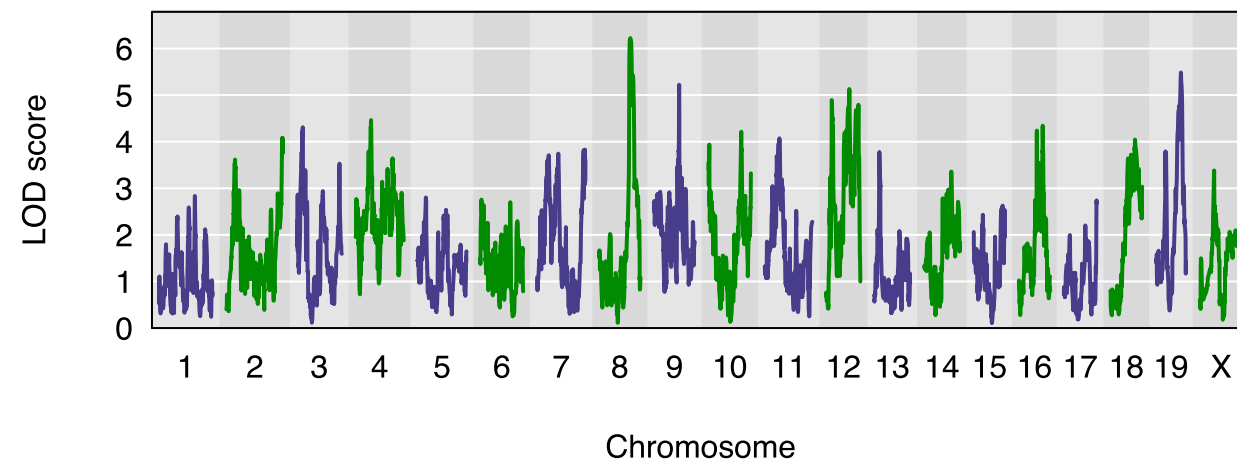
- Logan, et al. (2013) and Recla, et al. (2014) studied 261 Diversity Outbred mice
- Two traits map to Chr 8:
  - "hot plate latency" (57 cM)
  - "percent time in light" (55 cM)

# QTL scan results

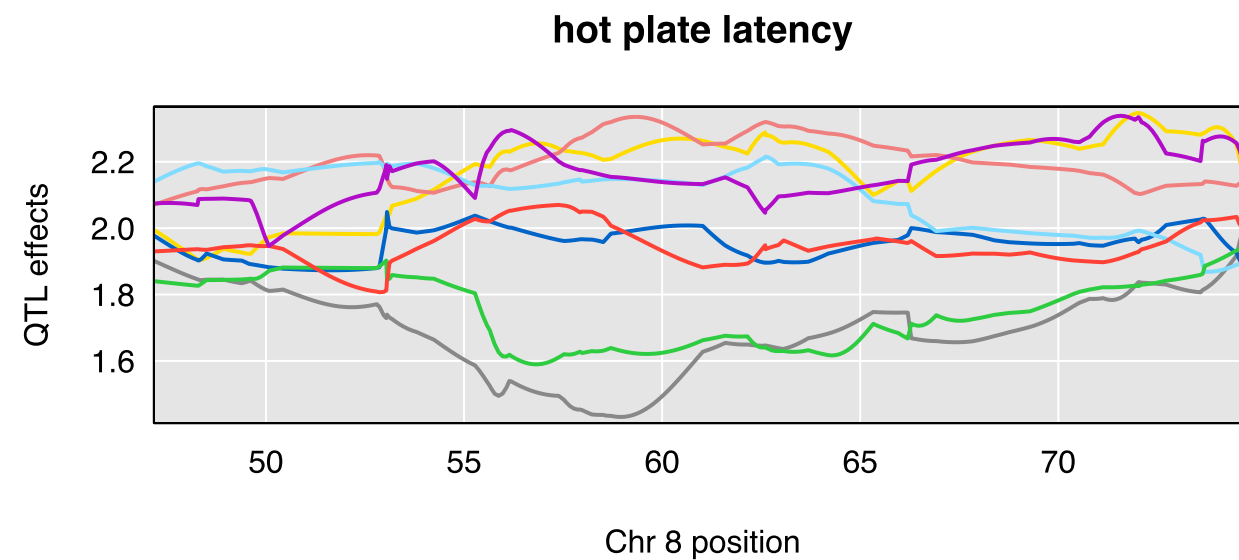
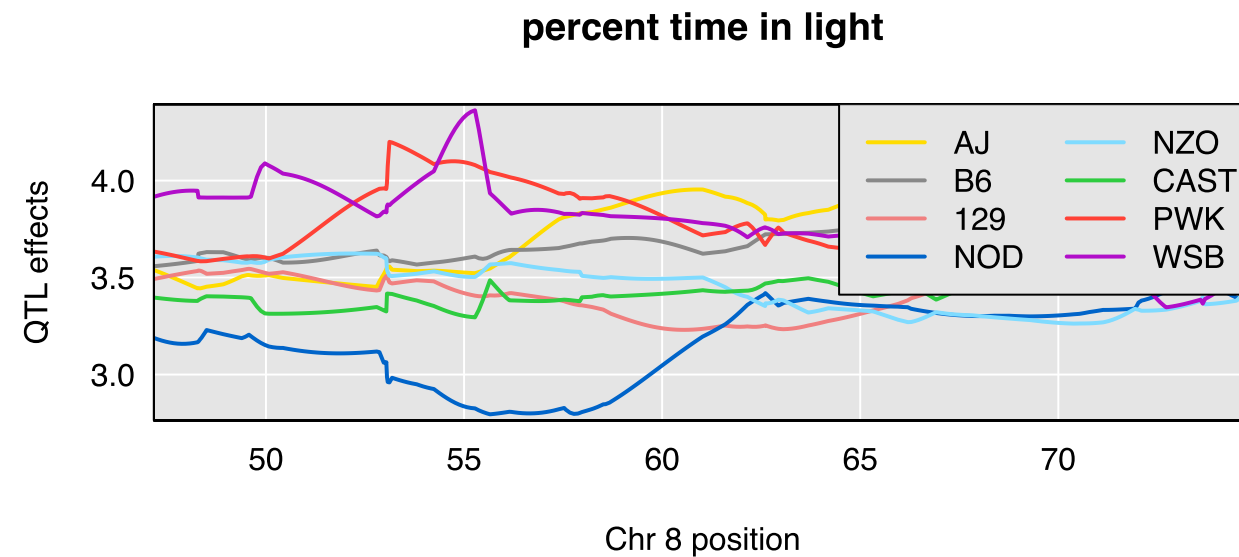
**percent time in light**



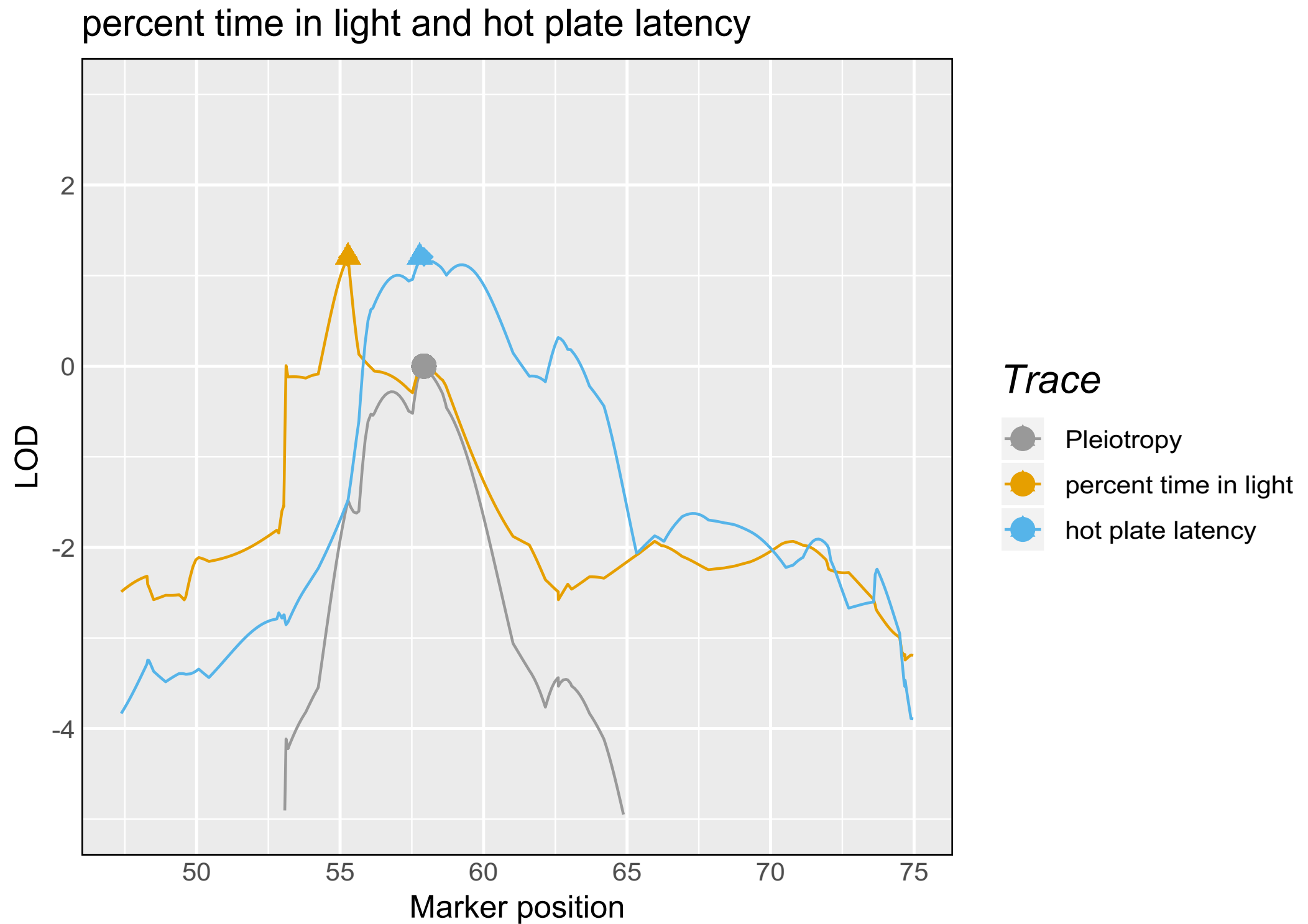
**hot plate latency**



# Allele effects plots



# Profile LOD plot





# Test results

- $p = 0.11$  (1000 bootstrap samples)

# Conclusions

- Weak evidence for two separate QTL
  - One QTL affects "distance traveled in light"
  - Second QTL affects "hot plate latency"

# Mediation analysis and pleiotropy testing

# Mediation analysis

- Is one molecule an intermediate in a causal pathway?

# Mediation analysis example

- QTL near Dhtkd1 gene affects Dhtkd1 RNA levels and DHTKD1 protein levels
- Is Dhtkd1 RNA an intermediate between QTL and DHTKD1 protein levels?

```
graph LR; A[Dhtkd1_DNA] --> B[Dhtkd1_RNA]; B --> C[DHTKD1_Protein]
```

Dhtkd1\_DNA → Dhtkd1\_RNA → DHTKD1\_Protein

# Mediation analysis methods

- LOD score with and without conditioning on candidate intermediate
- LOD difference
- LOD difference proportion

# Dissecting an eQTL hotspot

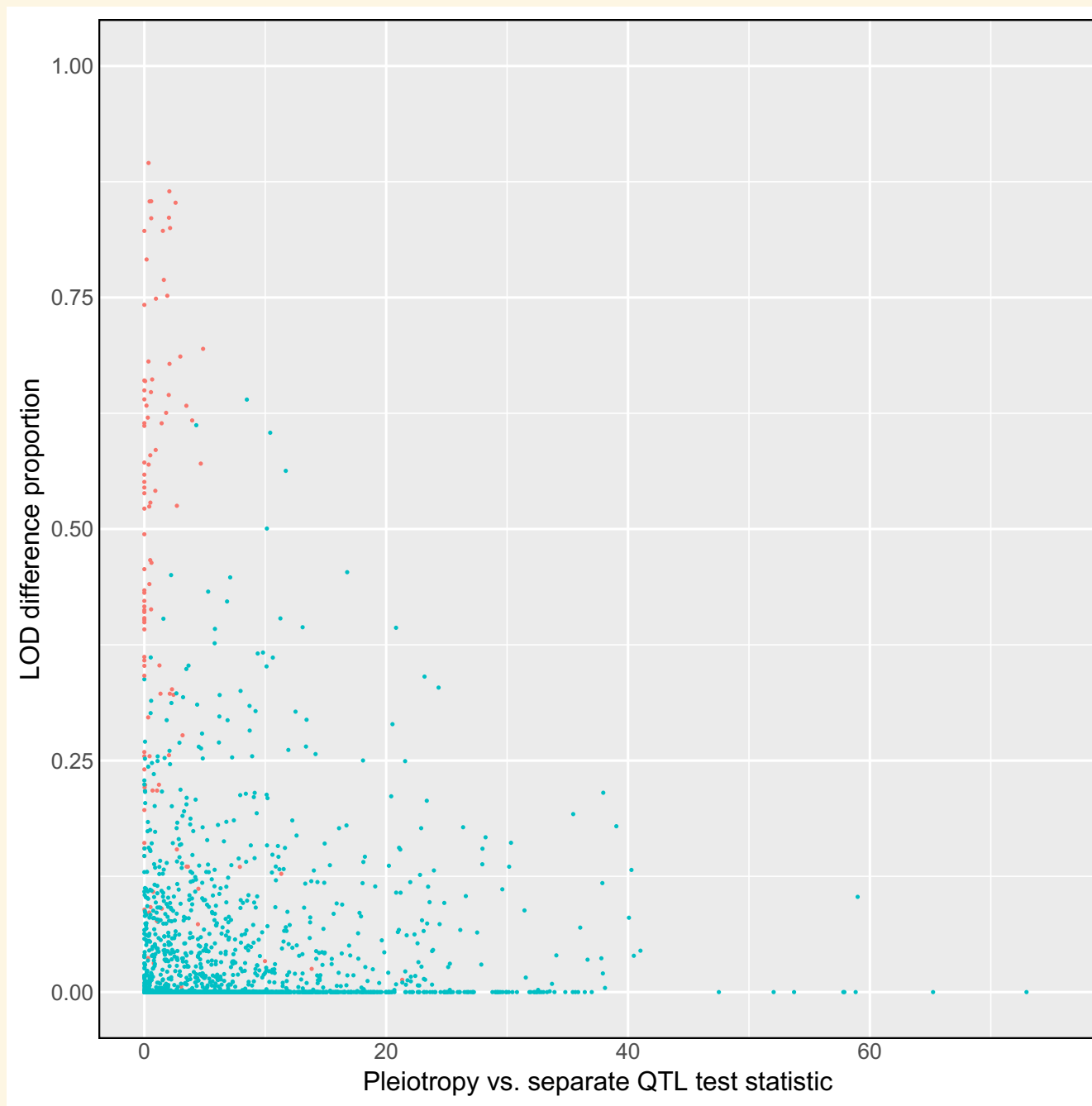
- Keller, et al. (2018) identified 147 nonlocal traits that map to Chr 2 (163.5 Mb to 167.5 Mb)
- expression levels mediate 88 of 147 traits
- Do other local genes act as intermediates?

# Dissecting an eQTL hotspot

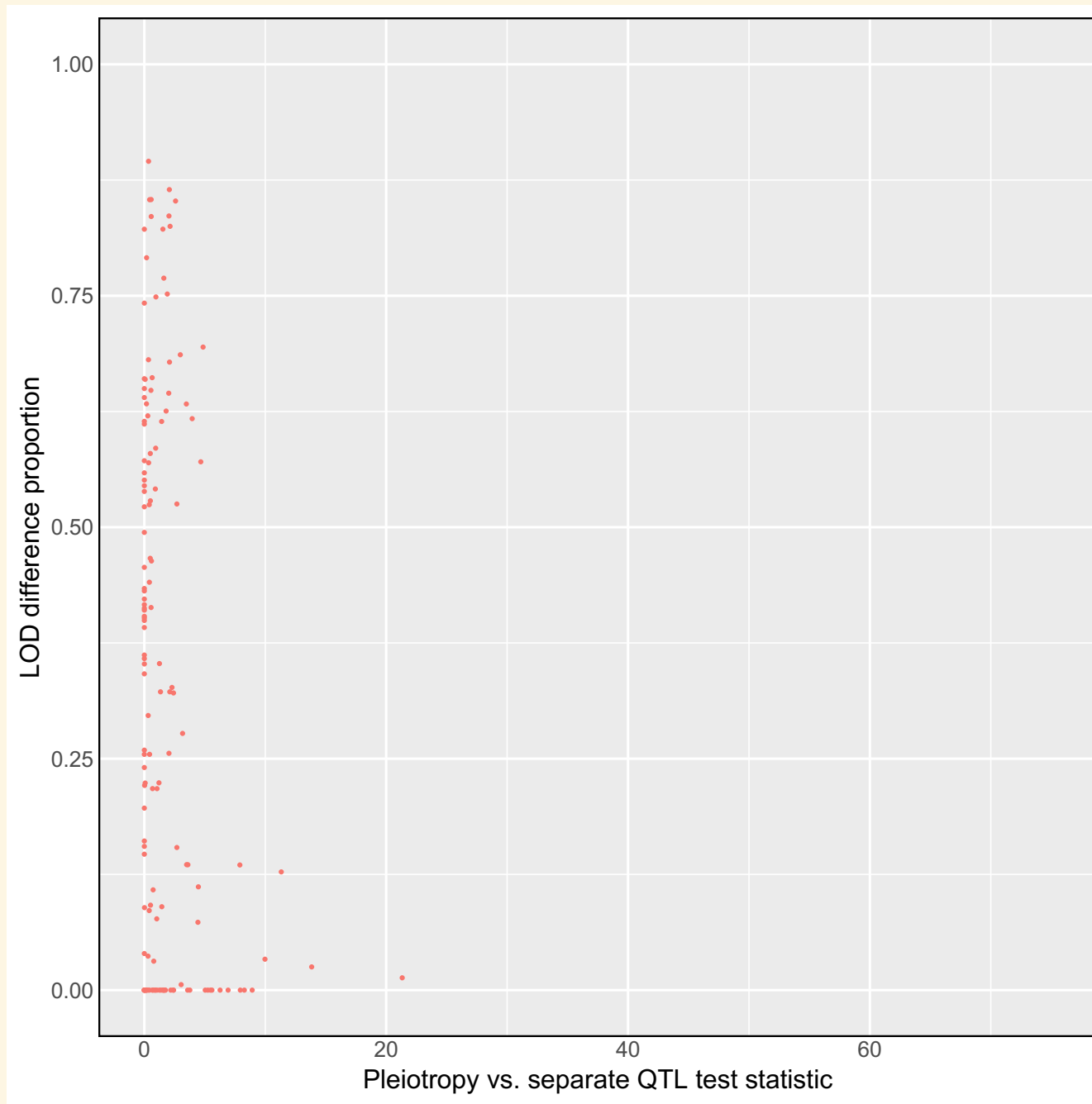
- 13 local gene expression traits as candidate mediators
- 147 nonlocal gene expression traits
- Pair one local trait with one nonlocal trait



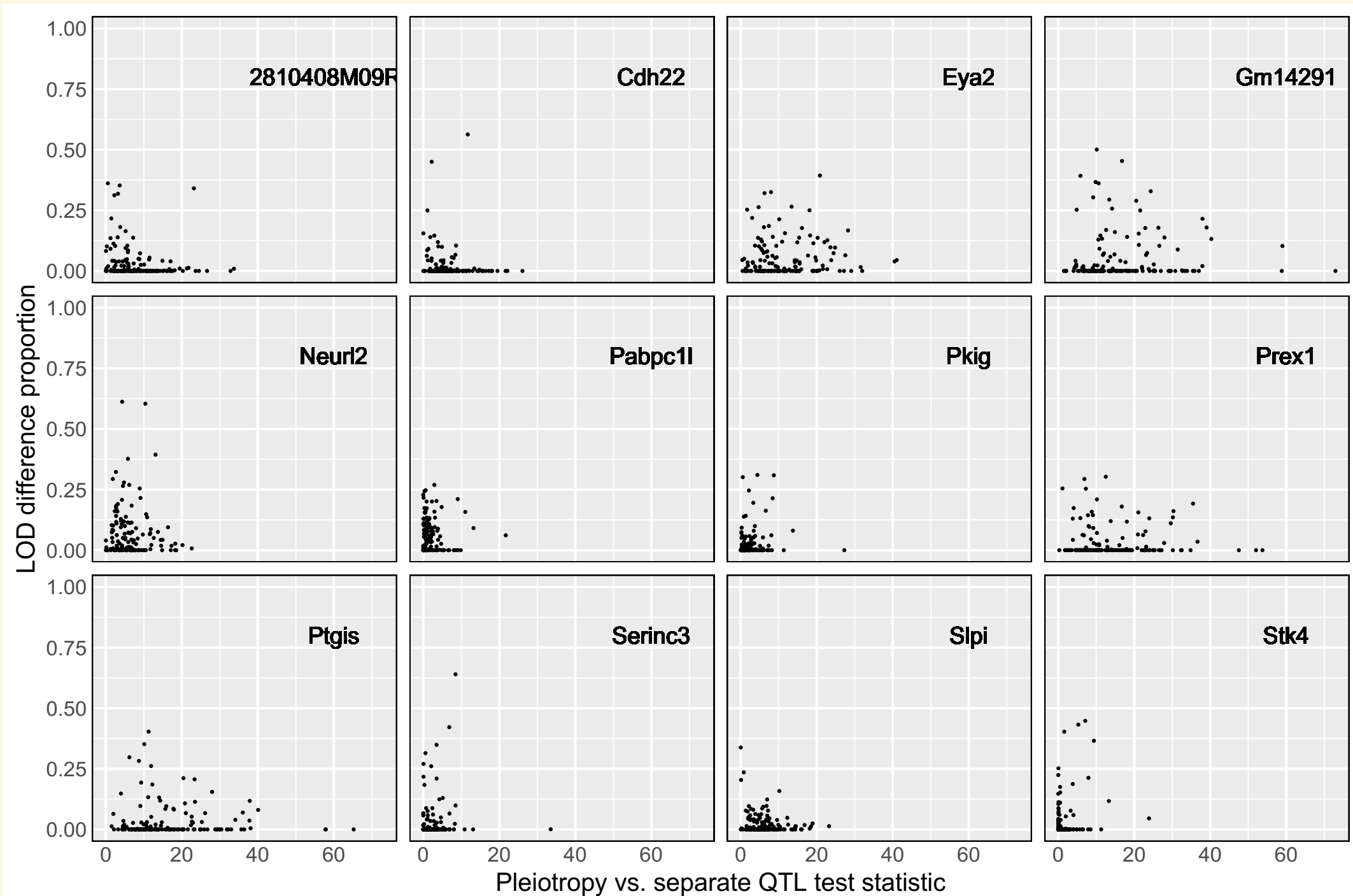
# Dissecting an eQTL hotspot



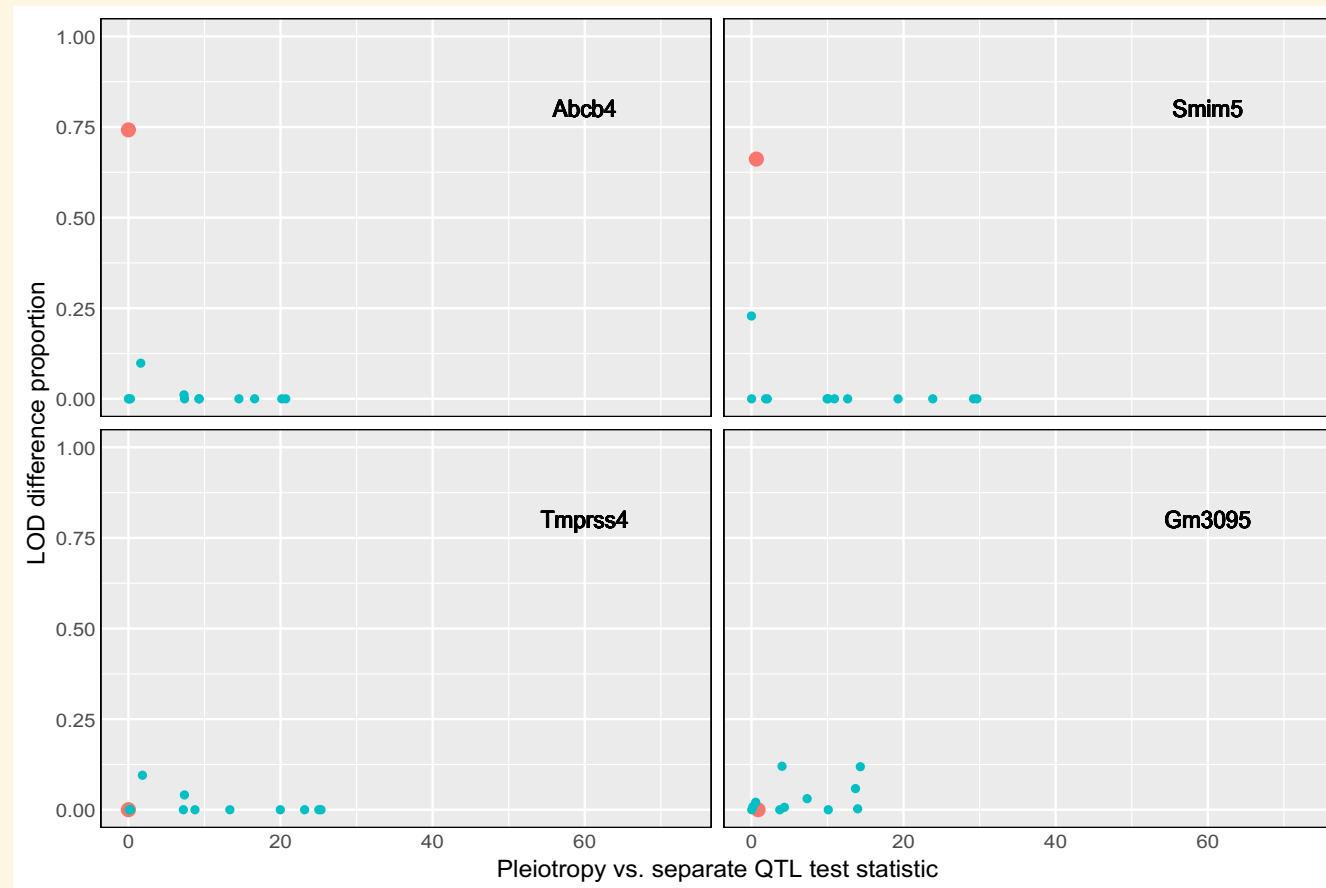
# Dissecting an eQTL hotspot



# Dissecting an eQTL hotspot



# Dissecting an eQTL hotspot



qtl2pleio R package development

# Design elements

- Written mostly in R
  - Some calculations in C++, via Rcpp package
  - C++ Eigen library and RcppEigen for matrix algebra
  - User interacts only with R functions

# Design elements

- Tutorial vignettes demonstrate functionality
- Unit tests included via R package
- Documentation with R package
- Package website with R package

# R packages book

- Hadley Wickham's "R packages" book as a guide





# Open-source R package

- Hosted on github
- Install via `install.packages("rstanarm")` R package
- `commit` commits publicly visible
- Travis CI for testing after every commit

# Binder integration

- Enhances reproducibility
- Uses Docker images
- Downloads R packages for a specified date (or commit on Github)

# Binder integration: 2 new files

- `install.R`
  - contains code to install R packages
- `runtime.txt`
  - contains one line of text to indicate date for CRAN packages

# Future directions

# Contact information

- frederick.boehm@gmail.com
- <https://fboehm.us/>
- qtl2pleio R package:  
<https://github.com/fboehm/qtl2pleio>
- Biorxiv pre-print: <http://bit.ly/boehm-biorxiv>

# References

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

Keller, M. P, D. M. Gatti, et al. (2018). "Genetic Drivers of Pancreatic Islet Function". In: , 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: 12.4, pp. 424-437.

# References

Recla, J. M, R. F. Robledo, et al. (2014). "Precise genetic mapping and integrative bioinformatics in Diversity Outbred mice reveals Hydin as a novel pain gene". In: 25.5-6, pp. 211-222.