

# qt12pleio: Testing pleiotropy vs. separate QTL in multiparental populations

Frederick Boehm<sup>1</sup>, Brian Yandell<sup>1</sup>, and Karl W. Broman<sup>2</sup>

<sup>1</sup> Department of Statistics, University of Wisconsin-Madison <sup>2</sup> Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison

DOI: [10.21105/joss.01435](https://doi.org/10.21105/joss.01435)

## Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted: 28 April 2019

Published: 30 June 2019

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

## Summary

Modern quantitative trait locus (QTL) studies in multiparental populations offer opportunities to identify causal genes for thousands of clinical and molecular traits. Traditional analyses examine each trait by itself. However, to fully leverage this vast number of measured traits, the systems genetics community needs statistical tools to analyze multiple traits simultaneously (Jiang & Zeng, 1995; Korol, Ronin, & Kirzhner, 1995). A test of pleiotropy vs. separate QTL is one such tool that will aid dissection of complex trait genetics and enhance understanding of genetic architecture.

Jiang & Zeng (1995) developed a pleiotropy test for two-parent crosses. For a pair of traits that map to a single genomic region, they formulated the test with the null hypothesis being pleiotropy (the two traits are affected by a single QTL) against the alternative hypothesis of two separate QTL, with each QTL affecting exactly one trait in the pair.

The test of Jiang & Zeng (1995) doesn't directly apply to multiparental populations because

1. Multiparental populations have more than two founders
2. Multiparental populations have complicated pedigrees

Additionally, the test statistic distribution, under the null hypothesis of pleiotropy, doesn't follow a distribution with tabulated quantiles, like the chi-square with 1 degree of freedom. Thus, we need to implement a method for determining p-values.

We addressed the first two challenges by adding a fixed effect for every founder line and incorporating a multivariate polygenic random effect into the linear model, which resulted in a multivariate linear mixed effects model (Kang et al., 2008; Zhou & Stephens, 2014). We implemented a parametric bootstrap procedure to determine p-values for test statistics (Efron, 1979; Tian et al., 2016). We describe details of our statistical methods elsewhere (Boehm, Chesler, Yandell, & Broman, 2019).

qt12pleio offers a convenient interface for those already analyzing data with [qt12](#). The primary functions in qt12pleio are `scan_pvl`, to perform the multivariate multi-QTL scan, and `boot_pvl`, to obtain bootstrap samples. We also include functions for visualizing results. qt12pleio features three R package vignettes that demonstrate these and other qt12pleio functions. One vignette provides examples for performing bootstrap analysis with a computing cluster. For quality assurance purposes, we incorporated unit tests into qt12pleio via the R package `testthat` (Wickham, 2011).

qt12pleio uses C++ code for model fitting via generalized least squares. We use the R package `Rcpp` to interface with our C++ code (Eddelbuettel et al., 2011). We also make use of the C++ library `Eigen` via the R package `RcppEigen` (D. Bates & Eddelbuettel, 2013).

## Acknowledgments

This work was supported in part by National Institutes of Health grant R01GM070683 (to K.W.B.).

## References

- Bates, D., & Eddelbuettel, D. (2013). Fast and elegant numerical linear algebra using the RcppEigen package. *Journal of Statistical Software*, 52(5), 1–24. doi:[10.18637/jss.v052.i05](https://doi.org/10.18637/jss.v052.i05)
- Boehm, F. J., Chesler, E. J., Yandell, B. S., & Broman, K. W. (2019). Testing pleiotropy vs. separate QTL in multiparental populations. *G3: Genes, Genomes, Genetics*. doi:[10.1534/g3.119.400098](https://doi.org/10.1534/g3.119.400098)
- Eddelbuettel, D., François, R., Allaire, J., Ushey, K., Kou, Q., Russel, N., Chambers, J., et al. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8), 1–18. doi:[10.18637/jss.v040.i08](https://doi.org/10.18637/jss.v040.i08)
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7(1), 1–26. doi:[10.1214/aos/1176344552](https://doi.org/10.1214/aos/1176344552)
- Jiang, C., & Zeng, Z.-B. (1995). Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics*, 140(3), 1111–1127.
- Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., & Eskin, E. (2008). Efficient control of population structure in model organism association mapping. *Genetics*, 178(3), 1709–1723. doi:[10.1534/genetics.107.080101](https://doi.org/10.1534/genetics.107.080101)
- Korol, A. B., Ronin, Y. I., & Kirzhner, V. M. (1995). Interval mapping of quantitative trait loci employing correlated trait complexes. *Genetics*, 140(3), 1137–1147.
- Tian, J., Keller, M. P., Broman, A. T., Kendzierski, C., Yandell, B. S., Attie, A. D., & Broman, K. W. (2016). The dissection of expression quantitative trait locus hotspots. *Genetics*, 202(4), 1563–1574. doi:[10.1534/genetics.115.183624](https://doi.org/10.1534/genetics.115.183624)
- Wickham, H. (2011). testthat: Get started with testing. *The R Journal*, 3(1), 5–10. doi:[10.32614/rj-2011-002](https://doi.org/10.32614/rj-2011-002)
- Zhou, X., & Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods*, 11(4), 407. doi:[10.1038/nmeth.2848](https://doi.org/10.1038/nmeth.2848)