A Simple Regression-Based Method to Map Quantitative Trait Loci Underlying Function-Valued Phenotypes

Il-Youp Kwak,* Candace R. Moore,† Edgar P. Spalding,† and Karl W. Broman‡,1
*Department of Statistics, †Department of Botany, and ‡Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, Wisconsin 53706

ABSTRACT Most statistical methods for quantitative trait loci (QTL) mapping focus on a single phenotype. However, multiple phenotypes are commonly measured, and recent technological advances have greatly simplified the automated acquisition of numerous phenotypes, including function-valued phenotypes, such as growth measured over time. While methods exist for QTL mapping with function-valued phenotypes, they are generally computationally intensive and focus on single-QTL models. We propose two simple, fast methods that maintain high power and precision and are amenable to extensions with multiple-QTL models using a penalized likelihood approach. After identifying multiple QTL by these approaches, we can view the function-valued QTL effects to provide a deeper understanding of the underlying processes. Our methods have been implemented as a package for R, funqtl.

THERE is a long history of work to map genetic loci (quantitative trait loci, QTL) influencing quantitative traits. Most statistical methods for QTL mapping, such as interval mapping (Lander and Botstein 1989), focus on a single phenotype. However, multiple phenotypes are commonly measured, and recent technological advances have greatly simplified the automated acquisition of numerous phenotypes, including phenotypes measured over time. Phenotypes measured over time, an example of a function-valued trait, have a number of advantages, including the ability to dissect the time course of QTL effects.

A simple and intuitive approach to the analysis of such data is to perform QTL analysis at each time point, individually, to identify QTL that affect the phenotype at each time point. This method is simple; however, it does not consider the smooth association across time points, and so it may have less power to detect QTL. Moreover, it can be difficult to combine the results across time points into a consistent story.

A second approach is to fit parametric curves to the data from each individual and treat the parameter estimates as phenotypes in QTL analysis (*e.g.*, see Kendziorski *et al.* 2002). Ma *et al.* (2002) expanded this approach by fitting a logistic growth model, $g(t) = a/(1 + be^{-rt})$, at each putative QTL position, with parameters depending on QTL genotype. This approach can have high power if the model is correct, but it can be difficult to interpret the results if QTL have pleiotropic effects on multiple parameters, and the parameters may have no obvious biologic or mechanistic interpretation.

Another natural approach is to use a nonparametric method so that we do not need to specify the functional shape. For example, Yang *et al.* (2009) proposed a nonparametric functional QTL mapping method that used a certain number of basis functions to fit a function-valued phenotype. For example, we might use 10 basis functions. This reduces the dimension from the number of time points to 10, and this is done in a flexible way, guided by the data. Min *et al.* (2011) extended this method to multiple-QTL models, using Markov chain Monte Carlo (MCMC) techniques.

Xiong *et al.* (2011) proposed an additional nonparametric functional mapping method based on estimating equations (EE). This method is fast and allows the selection of multiple QTL by a test statistic that they proposed. Sillanpää

doi: 10.1534/genetics.114.166306
Manuscript received May 15, 2014; accepted for publication June 8, 2014; published Early Online June 14, 2014.
Available freely online through the author-supported open access option.
Supporting information is available online at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.166306/-/DC1.
[1]Corresponding author: Department of Biostatistics and Medical Informatics, University of Wisconsin, 2126 Genetics–Biotechnology Center, 425 Henry Mall, Madison, WI 53706. E-mail: kbroman@biostat.wisc.edu