3FunMap: full-sib family functional mapping of dynamic traits

Chunfa Tong[1], Zhong Wang[2], Bo Zhang[1], Jisen Shi[1] and Rongling Wu[2],*

[1]Key Laboratory of Forest Genetics and Biotechnology of Ministry of Education, Nanjing Forestry University, Nanjing 210037, China and [2]Center for Statistical Genetics, The Pennsylvania State University, Hershey, PA 17033, USA

ABSTRACT

Motivation: Functional mapping that embeds the developmental mechanisms of complex traits shows great power to study the dynamic pattern of genetic effects triggered by individual quantitative trait loci (QTLs). A full-sib family, produced by crossing two heterozygous parents, is characteristic of uncertainties about cross-type at a locus and linkage phase between different loci. Integrating functional mapping into a full-sib family requires a model selection procedure capable of addressing these uncertainties. 3FunMap, written in VC++ 6.0, provides a flexible and extensible platform to perform full-sib functional mapping of dynamic traits. Functions in the package encompass linkage phase determination, marker map construction and the pattern identification of QTL segregation, dynamic tests of QTL effects, permutation tests and numerical simulation. We demonstrate the features of 3FunMap through real data analysis and computer simulation.

Availability: http://statgen.psu.edu/software.

Contact: rwu@hes.hmc.psu.edu

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 INTRODUCTION

Many traits important in agriculture, biology and medicine change with time or other independent variables. The genetic control of these dynamic complex traits should be accordingly expressed as a function of the independent and continuous variable (Atchley, 1984). A new breakthrough in mapping the dynamic change of quantitative trait locus (QTL) effects has been recently made through the invention of a series of statistical models, called functional mapping (He *et al.*, 2009; Li and Wu, 2010; Wu and Lin, 2006). Functional mapping integrates mathematical aspects of trait development into a QTL mapping framework, enabling geneticists to test the interplay between gene action and development. Ma *et al.* (2004) packed functional mapping into a web-based platform, FunMap, for a mapping population derived from inbred lines.

FunMap cannot be applied directly to a full-sib family produced by crossing heterozygous parents. Inevitably, such a single full-sib family that may generate an almost unlimited number of progeny has been commonly used for QTL mapping in outcrossing species, like forest trees and wildlife species (Zhang *et al.*, 2009). Since several assumptions essential for linkage mapping, including a fixed marker cross-type, known linkage phase and known parental origin

of QTL alleles, do not exist in outcrossing populations, a model selection procedure should be implemented to make an inference about these uncertainties from observed marker and phenotypic data. Here, we report a new package of Windows software, 3FunMap, which conducts functional mapping of dynamic QTLs segregating in a full-sib family derived from two heterozygous parents.

Since the genotypes of the two parents are unknown, the segregation of a QTL in a full-sib family may obey the following possible types: full cross (segregating 1:1:1:1), pseudo-F_2 cross (segregating 1:2:1) and pseudo-backcross (segregating 1:1). By incorporating model selection criteria, such as Akaike Information criterion (AIC) and Bayesian information criterion (BIC), 3FunMap can discern an optimal segregation type based on a given dataset, thus meeting the urgent need for analysis tools of functional mapping capable of handling the complexities of genetic segregation associated with crossed heterozygous parents. The software integrates a comprehensive suite of biological merits of functional mapping in terms of testing and estimating the developmental pattern of genetic effects triggered by dynamic QTLs. The software enables users to test when a QTL is switched on to perform its function, when the QTL is switched off to cease the function, how long it is expressed in a time course and with which temporal pattern is the QTL effect operational? Wu and Lin (2006) classified four temporal patterns of genetic expression for a QTL, i.e. long, early, late and inverse. By asking and addressing biologically meaningful questions, 3FunMap will help to better understand the genetic and developmental architecture of complex traits.

2 EXAMPLE

In Supplementary Material 1, we provided the implementation of 3FunMap that performs functional mapping in a full-sib family by taking into account the characteristics of outcrossing populations. A full-sib family of 86 hybrids was derived from the hybridization between *Populus deltoides* and *P.euramericana*. A panel of codominant and dominant markers was genotyped to construct an integrative linkage map of 19 groups (Zhang *et al.*, 2009). Ramets from the genotyped hybrids were water cultured in a randomized complete block design in a green house. During 65 day culture in water, the total number of roots (TNR) per cutting was repeatedly measured once every other week (Supplementary Material 2).

By fitting time-varying TNR values using Legendre polynomials (Lin and Wu, 2006), 3FunMap scans the linkage map under different types of QTL segregation; each generates an LR profile for QTL detection (Fig. 1). By comparing the critical thresholds from permutation tests, 3FunMap detected three QTLs: *TNRQ*1 on

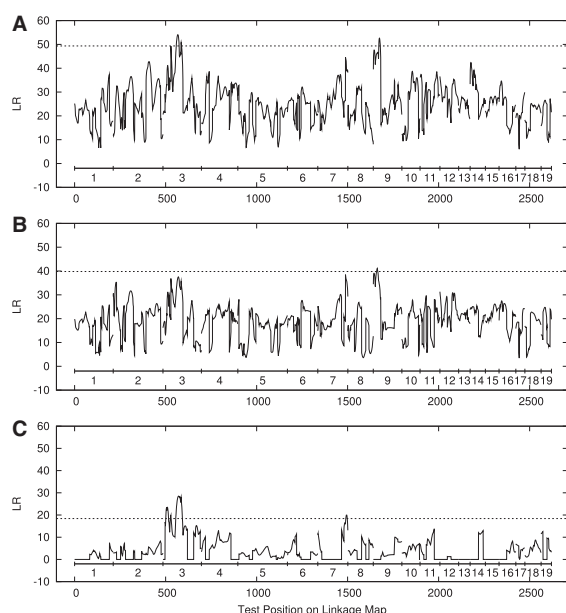*To whom correspondence should be addressed.

Fig. 1. The profiles of log-likelihood (LR) ratios for TNR trajectories across all the 19 linkage groups in the integrated map of *P.deltoides* and *P.euramericana* assuming that a QTL is segregating in the type of full cross (**A**), pseudo-F$_2$ cross (**B**) and pseudo-backcross (**C**). Short horizontal line segments represent 19 linkage groups. Threshold values for asserting the existence of a QTL were determined from 1000 permutation tests, which are 18.42, 39.81 and 49.35 at $P = 0.05$ (indicated by dot horizontal lines) for three segregation types, respectively. The peaks of LR profiles beyond the thresholds under a specific QTL cross-type indicate the genomic locations of the QTLs detected.

Fig. 2. Growth trajectories of TNR (normalized by a square root transformation) under water culture for different QTL genotypes at a QTL detected by a pseudo-F$_2$ model on linkage group 9 (**A**), a pseudo-backcross model on linkage group 3 (**C**) and a pseudo-backcross model on linkage group 7 (**E**). In (**B**), (**D**) and (**F**), time-varying genetic effects of the detected QTLs are shown.

linkage group 9 that was identified by the full cross and pseudo-F$_2$ model; *TNRQ*2 on linkage group 3 identified by the full cross and pseudo-backcross model; and *TNRQ*3 on linkage group 7 identified by the pseudo-backcross model. Based on both the AIC and BIC, it is more likely that *TNRQ*1 is a pseudo-F$_2$ QTL, whereas *TNRQ*2 is a pseudo-backcross QTL. The estimates of QTL locations, QTL segregation types and heritabilities at each time point are given in Supplementary Material 3.

3FunMap drew developmental curves of TNR for different genotypes at each QTL (Fig. 2). *TNRQ*1 exerts both additive and dominant effects on TNR growth. The additive effect of this QTL is quite stable, but its dominant effect changes dramatically with time and even alters the direction in an early stage. This QTL is operational in an additive to dominant manner, but after about 90 days in water culture, it is switched to perform its overdominant effect. At a beginning of growth, *TNRQ*2 shows a decreasing effect with time, but after a particular time its effect increases gradually. The effect of *TNRQ*3 decreases slightly with time and shortly increases exponentially with time. All the QTLs detected show complex gene × time interactions.

3 DISCUSSION

We described a Windows software, 3FunMap, to perform functional mapping in a full-sib family by considering unique outcrossing characteristics of heterozygous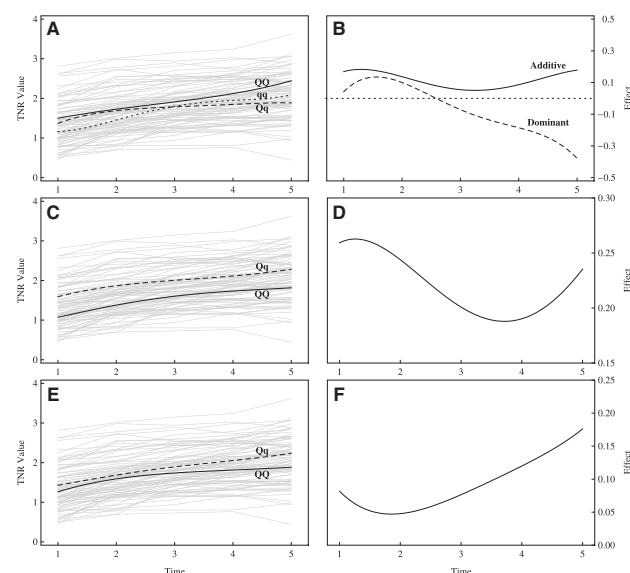 parents. With 3FunMap, users can construct an integrative linkage map by using all types of molecular markers (Lu *et al.,* 2004), map-specific QTLs that govern dynamic changes of a complex trait, select an optimal type of QTL segregation and estimate the temporal pattern of genetic effects triggered by individual QTLs. To our best knowledge, this is a first package of software that can handle dynamic QTL segregation in a full-sib family in a comprehensive way. It equips breeders and geneticists for outcrossing species with a powerful tool to dissect the genetics of dynamic complex traits.

3FunMap has incorporated a function to run simulation studies to investigate the statistical behavior of functional mapping for a full-sib family. Results from computer simulation under different heritability levels and sample sizes (Supplementary Material 4) suggest that a modest heritability ($H^2 = 0.1$) would provide a reasonably accurate estimate of the genetic control of trait trajectories when a sample size $n = 400$ is used. It is important to emphasize that a better management of plants that minimizes environmental and phenotyping errors may be more useful for the enhancement of QTL mapping precision than augmenting the size of roughly managed samples.

We also tested model misspecification by simulating and analyzing data reciprocally using different segregation models. Under $H^2 = 0.1$ and $n = 400$, 3FunMap has full power to identify a correct model.

We recognize the importance of epistasis in trait control. Epistasis has been incorporated into our software. For a single QTL model, 3FunMap uses ∼20 min on a PC desktop to draw the LR profiles for the *Populus* example shown in Figure 1. Yet, computing time on for epistatic detection for a full-sib family increases exponentially because there are many combinations of segregation types between any two QTLs. Also, two full-cross QTLs produce 16 genotypes,

thus increasing the dimension of epistasis. Although analysis and discrimination processes of epistasis in a full-sib family can be accelerated by parallelizing a cluster of machines, we are currently working to incorporate score statistics approaches (Change *et al.*, 2009) for functional mapping to enhance computing efficiency.

Conflict of Interest: none declared.

REFERENCES

Atchley,W.R. (1984) Ontogeny, timing of development, and genetic variance-covariance structure. *Am. Nat.*, **123**, 519–540.

Chang,M.R. *et al.* (2009) Score statistics of quantitative trait locus mapping. *Stat. App. Genet. Mol. Biol.*, **8**, Article 16.

He,Q.L. *et al.* (2009) Mapping genes for plant structure, development and evolution: functional mapping meets ontology. *Trends Genet.*, **26**, 39–46.

Li,Y. and Wu,R.L. (2010) Functional mapping of growth and development. *Biol. Rev.*, **85**, 207–216.

Lin,M. and Wu,R.L. (2006) A joint model for nonparametric functional mapping of longitudinal trajectories and time-to-events. *BMC Bioinformatics*, **7**, 38.

Lu,Q. *et al.* (2004) A multilocus likelihood approach to joint modeling of linkage, parental diplotype and gene order in a full-sib family. *BMC Genet.*, **5**, 20.

Ma,C.-X. *et al.* (2004) FunMap: functional mapping of complex trait. *Bioinformatics*, **20**, 1808–1811.

Wu,R.L. and Lin,M. (2006) Functional mapping – how to map and study the genetic architecture of complex dynamic traits. *Nat. Rev. Genet.*, **7**, 229–237.

Zhang,B. *et al.* (2009) Detection of quantitative trait loci influencing growth trajectories of adventitious roots in *Populus* using functional mapping. *Tree Genet. Genom.*, **5**, 539–552.