A statistical model for mapping quantitative trait loci in multivalent autotetraploids underlying double reduction

Yafei Lv[1*], Xiaoxia Yang[1*], Xin Li[1], Sisi Feng[1], Chunfa Tong[2], Zhong Wang[2], Xiaoming Pang[1], Christian M. Tobias[3] and Rongling Wu[1,2]

[1]Center for Computational Biology, Beijing Forestry University, Beijing 100083, China; [2]Center for Statistical Genetics, The Pennsylvania State University, Hershey, PA 17033, USA; [3]Genomics and Gene Discovery Research Unit, Western Regional Research Center, 800 Buchanan St., Albany, CA 94710 USA

*These authors contributed to this work equally.

Corresponding author: Rongling Wu, rwu@bjfu.edu.cn

Abstract

Multivalent tetraploids that include many plant species, such as potato, sugarcane and rose, are of paramount importance to agricultural production and biological research. Quantitative trait locus (QTL) mapping in multivalent tetraploids is challenged by their unique cytogenetic properties, such as double reduction. We develop a statistical method for mapping multivalent tetraploid QTLs by considering these cytogenetic properties. This method is built in the mixture model-based framework and implemented with the EM algorithm. The method allows the simultaneous estimation of QTL positions, QTL effects, the chromosomal pairing factor and the degree of double reduction as well as the assessment of the estimation precision of these parameters. We used simulated data to examine the statistical properties of the method and validate its utilization. The new method and its software will provide a useful tool for QTL mapping in multivalent tetraploids that undergo double reduction.

Key words: Quantitative trait loci, Multivalent tetraploid, EM algorithm, Quantitative genetic model.

1 Introduction

Genetic analysis in polyploids has received considerable interest in recent years because of the biological and economic importance [1–3]. Genetic linkage maps constructed from molecular markers have been published for several major polyploids [4–10]. Statistical models for linkage analysis and map construction that consider unique biological properties of polyploids have been developed [11–14]. For bivalent polyploids, Wu et al. [15,16] incorporated the so-called chromosomal pairing preference [17] into the linkage analysis framework, to increase the biological relevance of linkage mapping models. There have been several statistical models developed to map quantitative trait loci (QTLs) in bivalent polyploids [18,19].

There is also a group of polyploids, called multivalent polyploids, in which chromosomes

pair among more than two homologous copies at meiosis, rather than only two copies as like in bivalent polyploids. The origin of multivalent polyploids is mostly from the duplication of similar genomes and, for this reason, are called autopolyploids [20,21]. The consequence of multivalent pairing in autopolyploids is the occurrence of double reduction, i.e., two sister chromatids of a chromosome sort into the same gamete [22]. Fisher [23] proposed a conceptual model for characterizing the individual probabilities of 11 different modes of gamete formation for a quadrivalent polyploid in terms of the recombination fraction between two different loci and their double reductions. Wu et al. [24] used Fisher's model to derive the EM algorithm for the estimation of the linkage between fully informative markers. Wu and Ma [25] extended this model into analyze any type of markers, regardless of their informativeness and dominant or codominant nature. The significant advantage of the models by Wu and colleagues directly lies in their generality, flexibility and robustness.

In this chapter, we develop a statistical method for QTL mapping in multivalent tetraploids by considering Fisher's [23] 11 classifications of gamete formation. The method allows the estimation and test of not only the QTL-marker linkage, but also the extent of double reduction of the QTL. Because of the inherent complexity of classification analyses of gamete formation, we will focus on the modeling and analysis of one-marker/one-QTL associations. A two-stage hierarchical model is derived to estimate the probabilities of gamete formation modes and therefore double reduction in the upper hierarchy and estimate the marker-QTL recombination fraction in the lower hierarchy within the maximum likelihood context implemented with the EM algorithm. The method is used to analyze a simulated data set, with results demonstrating statistical properties of the method and its analytical and biological merits.

Model

Genetic Design

We consider a so-called pseudotest backcross population which is generated by hybridizing a heterzogyous multivalent tetraploid line and a homozygous line. In such a population, the genotypes of progeny can be identified by examining the genotypes of gametes produced by the heterozygous parent. Therefore, the statistical models for mapping QTLs in tetraploids can be established on the basis of the segregation of gametes. Assume that n individuals are sampled from the pseudotest backcross population, and each is typed by codominant markers and measured for a trait that is controlled by QTLs. A linkage map is supposed to be constructed with all the available markers, on which we scan QTLs by the new QTL mapping method developed in this paper.

Frequencies of Gametes

To implement interval mapping QTLs in tetraploids, we need to know the frequencies of gametes that contain two markers and a potential QTL between them, and hence to derive the conditional probabilities of QTL genotypes on the marker genotypes. Let $a_1, \cdots, a_4$ be the four alleles at marker A, $b_1, \cdots, b_4$ be the four alleles at marker B, and $q_1, \cdots, q_4$ be the four alleles at a QTL between markers A and B. Because of the double reduction, the heterozygous parent produces 10 diploid gametes at each locus, which are arrayed as (a_1a_1, a_2a_2, a_3a_3, a_4a_4, a_1a_2, a_1a_3, a_1a_4, a_2a_3, a_2a_4, a_3a_4) for marker A, (b_1b_1, b_2b_2, b_3b_3, b_4b_4, b_1b_2, b_1b_3, b_1b_4, b_2b_3, b_2b_4, b_3b_4) for marker B, and (q_1q_1, q_2q_2, q_3q_3, q_4q_4, q_1q_2, q_1q_3, q_1q_4, q_2q_3, q_2q_4, q_3q_4) for the QTL. The first four gametes in each case are created due to the double reduction, while the others are generated from the chromosome pairing. Let r, r_1, and r_2 to be the recombination fractions between markers A and B, marker A and the QTL, and the QTL and marker B, respectively. Under the assumption of no interference in recombination,

we have the relationship for the three recombination fractions as

$$r = r_1 + r_2 - \frac{4}{3} r_1 r_2 \tag{1}$$

by noticing that there is analogous to Haldane's mapping function in tetraploids as

$$r = \frac{3}{4}\left(1 - e^{-4x/3}\right) \tag{2}$$

where x is the genetic distance in centimorgans (Luo et al. 2006).

For three linked loci in a multivalent tetraploid, Fisher (1947) summarized that there are a total of 2080 gametes which can be classified into 107 gamete modes. We rearrange these gamete modes for the two markers and QTL and find that there are only 59 typical genotypes each with a frequency denoted by $g_i (i = 1, \cdots, 59)$ (Table 1). Some of the typical genotypes contains 2 or 4 gamete modes with relative proportions determined by the recombination fractions r_1 and r_2. For example,the sum of frequencies of modes 15 and 16 (denoted by g_{15a} and g_{15b}) composes the frequency of the typical genotype $a_1 a_1 q_1 q_2 b_1 b_2$, i.e., $g_{15} = g_{15a} + g_{15b}$. All the frequencies of the genotypes at the two markers and QTL form a $10 \times 10 \times 10$ cubic matrix G with its element denoted by G_{ijk} where the subscript i, j, and k are for the genotypes at marker A, marker B, and the QTL, respectively. By shrinking the matrix G, we can get the expected frequencies of the observable genotypes at the two markers in

matrix form expressed as

$$M = \begin{array}{c} \\ a_1a_1 \\ a_2a_2 \\ a_3a_3 \\ a_4a_4 \\ a_1a_2 \\ a_1a_3 \\ a_1a_4 \\ a_2a_3 \\ a_2a_4 \\ a_3a_4 \end{array} \begin{array}{c} \begin{array}{cccccccccc} b_1b_1 & b_2b_2 & b_3b_3 & b_4b_4 & b_1b_2 & b_1b_3 & b_1b_4 & b_2b_3 & b_2b_4 & b_3b_4 \end{array} \\ \begin{pmatrix} \frac{1}{4}\pi_1 & \frac{1}{12}\pi_2 & \frac{1}{12}\pi_2 & \frac{1}{12}\pi_2 & \frac{1}{12}\pi_3 & \frac{1}{12}\pi_3 & \frac{1}{12}\pi_3 & \frac{1}{12}\pi_4 & \frac{1}{12}\pi_4 & \frac{1}{12}\pi_4 \\ \frac{1}{12}\pi_2 & \frac{1}{4}\pi_1 & \frac{1}{12}\pi_2 & \frac{1}{12}\pi_2 & \frac{1}{12}\pi_3 & \frac{1}{12}\pi_4 & \frac{1}{12}\pi_4 & \frac{1}{12}\pi_3 & \frac{1}{12}\pi_3 & \frac{1}{12}\pi_4 \\ \frac{1}{12}\pi_2 & \frac{1}{12}\pi_2 & \frac{1}{4}\pi_1 & \frac{1}{12}\pi_2 & \frac{1}{12}\pi_4 & \frac{1}{12}\pi_3 & \frac{1}{12}\pi_4 & \frac{1}{12}\pi_3 & \frac{1}{12}\pi_4 & \frac{1}{12}\pi_3 \\ \frac{1}{12}\pi_2 & \frac{1}{12}\pi_2 & \frac{1}{12}\pi_2 & \frac{1}{4}\pi_1 & \frac{1}{12}\pi_4 & \frac{1}{12}\pi_4 & \frac{1}{12}\pi_3 & \frac{1}{12}\pi_4 & \frac{1}{12}\pi_3 & \frac{1}{12}\pi_3 \\ \frac{1}{12}\pi_5 & \frac{1}{12}\pi_5 & \frac{1}{12}\pi_6 & \frac{1}{12}\pi_6 & \frac{1}{6}\pi_7 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 & \frac{1}{6}\pi_9 \\ \frac{1}{12}\pi_5 & \frac{1}{12}\pi_6 & \frac{1}{12}\pi_5 & \frac{1}{12}\pi_6 & \frac{1}{24}\pi_8 & \frac{1}{6}\pi_7 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 & \frac{1}{6}\pi_9 & \frac{1}{24}\pi_8 \\ \frac{1}{12}\pi_5 & \frac{1}{12}\pi_6 & \frac{1}{12}\pi_6 & \frac{1}{12}\pi_5 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 & \frac{1}{6}\pi_7 & \frac{1}{6}\pi_9 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 \\ \frac{1}{12}\pi_6 & \frac{1}{12}\pi_5 & \frac{1}{12}\pi_5 & \frac{1}{12}\pi_6 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 & \frac{1}{6}\pi_9 & \frac{1}{6}\pi_7 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 \\ \frac{1}{12}\pi_6 & \frac{1}{12}\pi_5 & \frac{1}{12}\pi_6 & \frac{1}{12}\pi_5 & \frac{1}{24}\pi_8 & \frac{1}{6}\pi_9 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 & \frac{1}{6}\pi_7 & \frac{1}{24}\pi_8 \\ \frac{1}{12}\pi_6 & \frac{1}{12}\pi_6 & \frac{1}{12}\pi_5 & \frac{1}{12}\pi_5 & \frac{1}{6}\pi_9 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 & \frac{1}{24}\pi_8 & \frac{1}{6}\pi_7 \end{pmatrix} \end{array} \tag{3}$$

where

$$\begin{cases} \pi_1 = g_1 + g_2 + g_3 + g_4 \\ \pi_2 = g_5 + g_6 + g_7 + g_8 + g_9 + g_{10} + g_{11} \\ \pi_3 = g_{12} + g_{13} + g_{14} + g_{15} + g_{16} + g_{17} + g_{18} \\ \pi_4 = g_{19} + g_{20} + g_{21} + g_{22} + g_{23} + g_{24} + g_{25} \\ \pi_5 = g_{26} + g_{27} + g_{28} + g_{29} + g_{30} + g_{31} + g_{32} \\ \pi_6 = g_{33} + g_{34} + g_{35} + g_{36} + g_{37} + g_{38} + g_{39} \\ \pi_7 = g_{40} + g_{41} + g_{42} + g_{43} + g_{44} \\ \pi_8 = g_{45} + g_{46} + g_{47} + g_{48} + g_{49} + g_{50} + g_{51} + g_{52} + g_{53} + g_{54} \\ \pi_9 = g_{55} + g_{56} + g_{57} + g_{58} + g_{59} \end{cases} \tag{4}$$

and the (i,j)th element of M is derived as $M_{ij} = \sum_{k=1}^{10} G_{ijk}$. If we define

$$\begin{cases} n_1 = n_{11} + n_{22} + n_{33} + n_{44} \\ n_2 = n_{12} + n_{13} + n_{14} + n_{21} + n_{23} + n_{24} + n_{31} + n_{32} + n_{34} + n_{41} + n_{42} + n_{43} \\ n_3 = n_{15} + n_{16} + n_{17} + n_{25} + n_{28} + n_{29} + n_{36} + n_{38} + n_{3,10} + n_{47} + n_{49} + n_{4,10} \\ n_4 = n_{18} + n_{19} + n_{1,10} + n_{26} + n_{27} + n_{2,10} + n_{35} + n_{37} + n_{39} + n_{45} + n_{46} + n_{48} \\ n_5 = n_{51} + n_{52} + n_{61} + n_{63} + n_{71} + n_{74} + n_{82} + n_{83} + n_{92} + n_{94} + n_{10,3} + n_{10,4} \\ n_6 = n_{53} + n_{54} + n_{62} + n_{64} + n_{72} + n_{73} + n_{81} + n_{84} + n_{91} + n_{93} + n_{10,1} + n_{10,2} \\ n_7 = n_{55} + n_{66} + n_{77} + n_{88} + n_{99} + n_{10,10} \\ n_8 = n_{56} + n_{57} + n_{58} + n_{59} + n_{65} + n_{67} + n_{68} + n_{6,10} + n_{75} + n_{76} + n_{79} + n_{7,10} \\ \quad + n_{85} + n_{86} + n_{89} + n_{8,10} + n_{95} + n_{97} + n_{98} + n_{9,10} + n_{10,6} + n_{10,7} + n_{10,8} + n_{10,9} \\ n_9 = n_{5,10} + n_{69} + n_{78} + n_{87} + n_{96} + n_{10,5} \end{cases} \tag{5}$$

where n_{ij} is the number of individuals with the ijth marker genotype, then π_i can be estimated as

$$\hat{\pi_i} = \frac{n_i}{n} \quad (i = 1, \cdots, 9). \tag{6}$$

In addition, the coefficients of double reduction at marker A (α), marker B (β), and the QTL (γ) can be expressed as

$$\alpha = \pi_1 + \pi_2 + \pi_3 + \pi_4, \tag{7}$$

$$\beta = \pi_1 + \pi_2 + \pi_5 + \pi_6, \tag{8}$$

$$\gamma = g_1 + g_2 + g_5 + g_6 + g_7 + g_{12} + g_{13} + g_{14} + g_{19} + g_{20} + g_{21} + g_{26} + g_{27} + g_{28} + g_{33} + g_{34} + g_{35} + g_{40} + g_{41} + g_{45} + g_{46} + g_{47} + g_{48} + g_{55} + g_{56}. \tag{9}$$

Let $\phi_1 = \frac{r_1^2}{9-18r_1+10r_1^2}$, $\phi_2 = \frac{r_2^2}{9-18r_2+10r_2^2}$, $\psi_1 = \frac{r_1}{3-2r_1}$, and $\psi_2 = \frac{r_2}{3-2r_2}$, then the frequencies of those gamete modes that have the same genotype can be expressed in terms of g_i as Table 2. These formulae are derived by the principle that the expected number of recombination events in each gamete mode should be the weighted average of the number of recombination

events for each typical genotype. If the frequencies of all the 107 gamete modes are known, we can express the recombination fractions r_1, r_2, and r by directly counting method as

$$\begin{aligned}
r_1 =& g_2 + g_4 + g_6 + g_7 + g_{10} + g_{11} + g_{13} + g_{14} + g_{17} + g_{18} + g_{20} + g_{21} + g_{24} + g_{25} \\
&+ g_{28} + g_{32} + g_{33} + g_{35} + g_{37} + g_{41} + g_{44} + g_{47} + g_{48} + g_{54} + g_{56} + g_{59} \\
&+ \frac{1}{2}\left(g_3 + g_8 + g_9 + g_{15} + g_{16} + g_{22} + g_{23} + g_{26} + g_{27} + g_{34} + g_{40} + g_{45} + g_{46} + g_{55}\right) \\
&+ \left(g_{29} + g_{38} + g_{42} + g_{49} + g_{57}\right)\phi_1 \\
&+ \frac{1}{2}\left(g_{30} + g_{31} + g_{36} + g_{39} + g_{43} + g_{50} + g_{51} + g_{52} + g_{53} + g_{58}\right)(1 + \psi_1),
\end{aligned} \tag{10}$$

$$\begin{aligned}
r_2 =& g_2 + g_4 + g_5 + g_7 + g_9 + g_{11} + g_{14} + g_{18} + g_{19} + g_{21} + g_{23} + g_{27} + g_{28} + g_{31} \\
&+ g_{32} + g_{34} + g_{35} + g_{38} + g_{39} + g_{41} + g_{44} + g_{46} + g_{48} + g_{53} + g_{55} + g_{57} \\
&+ \frac{1}{2}\left(g_3 + g_8 + g_{10} + g_{12} + g_{13} + g_{20} + g_{29} + g_{30} + g_{36} + g_{37} + g_{40} + g_{45} + g_{47} + g_{56}\right) \\
&+ \left(g_{15} + g_{24} + g_{42} + g_{50} + g_{59}\right)\phi_2 \\
&+ \frac{1}{2}\left(g_{16} + g_{17} + g_{22} + g_{25} + g_{43} + g_{49} + g_{51} + g_{52} + g_{54} + g_{58}\right)(1 + \psi_2),
\end{aligned} \tag{11}$$

$$\begin{aligned}
r =& \pi_2 + \pi_4 + \pi_6 + \pi_9 + g_{52} + \frac{1}{2}(\pi_3 + \pi_5 + g_{40} + g_{41} + g_{44} + g_{49} + g_{50} + g_{51}) \\
&+ \frac{3}{4}(g_{45} + g_{46} + g_{47} + g_{48} + g_{53} + g_{54}) + \frac{1}{2}(2g_{42} + g_{49})\phi_1 + \frac{1}{2}(2g_{42} + g_{50})\phi_2 \\
&- 2g_{42}\phi_1\phi_2 + \frac{1}{2}(2g_{43} + g_{50} + g_{51} - g_{52})\psi_1 - g_{50}\phi_2\psi_1 \\
&+ \frac{1}{2}(2g_{43} + g_{49} + g_{51} - g_{52})\psi_2 - g_{49}\phi_1\psi_2 - (2g_{43} + g_{51} - g_{52})\psi_1\psi_2.
\end{aligned} \tag{12}$$

Parameter Estimation

Suppose that n individuals are sampled from the pseudo-backcross population in the multivalent tetraploid. The phenotypic value of the lth individual who has the ith genotype at marker A and the jth genotype at marker B can be expressed in terms of the QTL effect and residual error as

$$y_{ijl} = \sum_{k=1}^{10} x_{ijl}^k \mu_k + e_{ijl} \tag{13}$$

where x_{ijl}^k is an indicator variable with a value of 1 if the ijlth individual has the kth QTL genotype, and 0 otherwise, μ_k is the genotypic value of the kth QTL genotype, and e_i is the residual error normally distributed with mean 0 and variance σ^2. The genotypic value of μ_k can be partitioned into additive and dominant genetic effects of different types indicated by Li et al.(2010) as

$$\begin{cases} \mu_1 = \mu + a_1 \\ \mu_2 = \mu + a_2 \\ \mu_3 = \mu + a_3 \\ \mu_4 = \mu - a_1 - a_2 - a_3 \\ \mu_5 = \mu + a_1 + a_2 + d_{12} \\ \mu_6 = \mu + a_1 + a_3 + d_{13} \\ \mu_7 = \mu - a_2 - a_3 + d_{14} \\ \mu_8 = \mu + a_2 + a_3 + d_{23} \\ \mu_9 = \mu - a_1 - a_3 + d_{24} \\ \mu_{10} = \mu - a_1 - a_2 + d_{34} \end{cases} \tag{14}$$

where μ is the overall mean, a_1, a_2, and a_3 are the additive genetic effects of alleles q_1, q_2, and q_3 relative to allele q_4, and d_{12}, d_{13}, d_{14}, d_{23}, d_{24}, and d_{34} are the dominant genetic effects between alleles q_1 and q_2, q_1 and q_3, q_1 and q_4, q_2 and q_3, q_2 and q_4, and q_3 and q_4, respectively.

For marker-based QTL mapping method, we need to know the conditional probability of QTL genotype given the marker genotype for an individual. The conditional probability of the kth QTL genotype on the ijth marker genotype is calculated as G_{ijk}/M_{ij}, i.e., the frequency of joint marker-QTL genotype ijk, expressed in terms of the 59 g probabilities, divided by the frequency of marker genotype ij. Therefore, given the phenotype and marker

data of the n individuals, the likelihood of the unknown parameters can be expressed as

$$L(\Omega|y,M)=\prod_{i=1}^{10}\prod_{j=1}^{10}\prod_{l=1}^{n_{ij}}\sum_{k=1}^{10}\frac{G_{ijk}}{M_{ij}}f_k(y_{ijl};\mu_k,\sigma^2) \tag{15}$$

where $\Omega=(g_1,\cdots,g_{59},\mu_1,\cdots,\mu_{10},\sigma^2)$ is the vector of the unknown parameters, and $f_k(y_{ijl};\mu_k,\sigma^2)$ is the probability density function of normal distribution with mean μ_k and variance σ^2. Moreover, the parameters of g_i in equation (15) are required to satisfy the linear restrictions of the expression (4),and the equations (10)–(12) if the recombination fractions of r_1, r_2, and r are given.

To find the maximum likelihood estimates (MLEs) of the unknown parameters, we proposed an EM algorithm under the linear restrictions on the parameters. This algorithm applies the interval mapping strategy to scan QTLs along the genome by the step of 1 cM. For a specific position of the QTL flanked by two markers, the iterative procedures to calculate the parameters are as following:

(i) Take initial values of μ_1, $\cdots$,μ_{10}, and σ^2, and then find the MLEs of g_i that satisfy the linear constrains of the equations (4), (10), (11), and (12). This procedure can be implemented with the optimization toolbox function of *fmincon* in MATLAB.

(ii) Given the estimates of g_i in the procedure (i), the estimates of μ_k and σ^2 can be obtained through EM algorithm as

$$\mu_k=\frac{\sum_{i=1}^{10}\sum_{j=1}^{10}\sum_{l=1}^{n_{ij}}P_{ijl}^k y_{ijl}}{\sum_{i=1}^{10}\sum_{j=1}^{10}\sum_{l=1}^{n_{ij}}P_{ijl}^k},\quad k=1,2,\cdots,10, \tag{16}$$

$$\sigma^2=\frac{1}{n}\sum_{i=1}^{10}\sum_{j=1}^{10}\sum_{l=1}^{n_{ij}}\sum_{k=1}^{10}P_{ijl}^k(y_{ijl}-\mu_k)^2, \tag{17}$$

where

$$P_{ijl}^k=\frac{G_{ijk}f_k(y_{ijl};\mu_k,\sigma^2)}{\sum_{k'=1}^{10}G_{ijk'}f_{k'}(y_{ijl};\mu_{k'},\sigma^2)}. \tag{18}$$

(iii) Update μ_1, $\cdots$, μ_{10}, and σ^2 with the new estimates in the procedure (ii), and repeat the procedures (i)–(iii) until all the estimates are convergent.

Hypothesis Testing

The hypothesis of testing if there exists a QTL in the pseudotest backcross can be formulated as

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 = \cdots = \mu_{10} \\ H_1 &: \text{at least one of the equalities above does not hold} \end{aligned} \tag{19}$$

where the H_0 corresponds to the reduced model assuming that no QTLs exists, and the H_1 corresponds to the full model (15). The log-likelihood ratio of the full model over the reduced model to test the hypothesis is calculated as

$$LR = -2 \log \left[\frac{L_0(\tilde{\Omega})}{L(\hat{\Omega})} \right] \tag{20}$$

where the tilde and hat are the MLEs under the H_0 and H_1, respectively. The critical threshold for asserting the existence of a QTL is determined by performing permutation tests (Churchill and Doerge 1994).

To test if the additive genetic effect of the QTL is significant, we form the hull hypothesis as

$$H_0 : a_1 = a_2 = a_3 = 0 \tag{21}$$

under which the parameter estimates can be obtained with the algorithm as describe above by shrinking the number of QTL genotypic values with $\mu_1 = \mu_2 = \mu_3 = \mu_4$. Similarly, we can also test the dominant genetic effect by the null hypothesis,

$$H_0 : d_{12} = d_{13} = d_{14} = d_{23} = d_{24} = d_{34} = 0 \tag{22}$$

under which only four genotypic values need to be estimated because all the genotypic values can be uniquely expressed in terms of μ_1–μ_4.

Application

Okada et al. [17] reported a linkage map for tetraploid Switchgrass (*Panicum virgatum* with a full-sib family of 238 progeny derived from two heterozygous parents. The mapping

population was genotyped for 1509 dominant markers (each with a dominant allele 1 and recessive allele 0), of which 606 are the testcross markers segregating for the female parent (K5), 667 are the testcross markers for the male parent (A4) and 126 are intercross markers for both parents. All these markers can be single-dose amplicons (simplex), double-dose amplicons (duplex) or triple-dose amplicons (triplex).

Consider two dominant markers which bracket a putative QTL. Given that limited informativeness the dominant markers possess is inadequate to identify and estimate a fully informative QTL, we assume that the QTL bracketed has two alleles Q and q. Because the two markers can be simplex (S), duplex (D), or triplex (T), there are $3 \times 3 = 9$ possible genotype combinations. For the simplex $\times$ simplex combination, there are four marker-QTL configurations:

SSS1			
1	0	0	0
Q	q	q	q
1	0	0	0

SSS2			
1	0	0	0
q	Q	q	q
1	0	0	0

SSS3			
1	0	0	0
Q	q	q	q
0	1	0	0

SSS4			
1	0	0	0
q	Q	q	q
0	1	0	0

The duplex $\times$ duplex combination has nine marker-QTL configurations:

DDD1			
1	1	0	0
Q	Q	q	q
1	1	0	0

DDD2			
1	1	0	0
Q	q	Q	q
1	1	0	0

DDD3			
1	1	0	0
q	q	Q	Q
1	1	0	0

DDD4			
1	1	0	0
Q	Q	q	q
1	0	1	0

DDD5			
1	1	0	0
Q	q	Q	q
1	0	1	0

DDD6			
1	1	0	0
Q	q	Q	q
1	0	0	1

DDD7			
1	1	0	0
Q	Q	q	q
0	0	1	1

DDD8			
1	1	0	0
Q	q	Q	q
0	0	1	1

DDD9			
1	1	0	0
q	q	Q	Q
0	0	1	1

The triplex $\times$ triplex combination has four marker-QTL configurations:

TTT1			
1	1	1	0
Q	Q	Q	q
1	1	1	0

TTT2			
1	1	1	0
Q	Q	q	Q
1	1	1	0

TTT3			
1	1	1	0
Q	Q	Q	q
1	1	0	1

TTT4			
1	1	1	0
Q	Q	q	q
1	1	0	1

The other six combinations each with a different marker-QTL configuration can also be provided. In each configuration, the QTL produces three genotypes QQ, Qq and qq. The conditional probabilities of the three QTL genotypes given the interval marker genotyps can

be derived in terms of g's (Text S1). In this case, genetic effects of the QTL are parameterized as $\mu_1 = \mu + a$ for QQ, $\mu_2 = \mu + d$ for Qq, and $\mu_3 = \mu - a$ for qq, where the additive (a) and dominant genetic effects (d) are estimated and tested.

Monte Carlo Simulation

A simulation study was performed to examine the statistical properties of our QTL mapping method for the pseudotest backcross population in a multivalent tetraploid. A QTL was assumed at the middle of a marker interval with a genetic distance of 20 cM. Marker and QTL genotypes were simulated by setting the coefficients of double reduction at the two markers and QTL and the QTL genotypic values as in the second column of Table 3 or 4. The population was sampled with different sample sizes ($n = 200, 400, 800$) and different QTL heritabilities ($H^2 = 0.1, 0.4$). The phenotype of an individual was drawn from the normal distribution with the mean as its genotypic value and the variance determined according to the heritability of QTL.

The means of the MLEs of the parameters and their standard errors based on 1000 simulation replicates are shown in Tables 3 and 4. We found that the QTL position and the double reduction coefficients at the two markers can well be estimated even with a small sample size (200) and a low heritability (0.1). The coefficient of double reduction at the QTL was estimated accurately with a modest sample size (400) when the heritability is at a low level of 0.1. When the heritability is at a higher level (0.4), the precision of the estimation of the QTL double reduction increased significantly even when the sample size is small (200). The main factor that affected the estimate of QTL double reduction was the heritability, followed by the sample size. Generally, a sample size of 400 can guarantee the precision for estimating the double reduction of a QTL that explains over 10% of the phenotypic variation.

The precision of estimates for the QTL effects was affected greatly by the heritability and sample size. As the heritability or the sample size increased, the QTL genotypic values

were more accurately estimated, which led to the improvement of estimating the additive and dominant effects. We found that the genotypic values of QTL genotypes from the chromosome pairing (μ_5–μ_{10}) can be reasonably estimated whatever the sample size and the heritability are. However, the genotypic values of QTL genotypes generated due to the double reduction (μ_1–μ_4) were influenced heavily by the sample size when the heritability is 0.1. This influence directly affected the estimation precision of the additive and dominant effects. For a QTL that explains a small proportion of the phenotypic variance (0.1), a large sample size is needed to obtain reliable estimates of the QTL effects. For a QTL with a high level of heritability, the QTL effects can be reasonably estimated even with a small sample size.

2 Discussion

A statistical method for genetic mapping of quantitative trait loci (QTLs) in a multivalent tetraploid undergoing a double reduction process is described. As an important cytological characteristic of polyploids, double reduction may play a significant role in plant evolution and maintenance of genetic polymorphism in natural populations. Also, because double reduction affects the result of linkage analysis through the crossing-over events between different chromosomes [24,25], it is important to incorporate double reduction into a QTL mapping framework. This method provides a powerful tool for QTL mapping and understanding the genetic control of a quantitative trait in an multivalent tetraploid.

The method capitalizes on 11 different classifications of two-locus gamete formations, derived by Fisher [23], during multivalent tetraploid meiosis and has proven to be powerful for simultaneous estimation of the frequencies of double reduction and the recombination fraction between different loci. Although a couple of statistical approaches have been proposed to map multivalent tetraploid QTLs [26,27], this method has for the first time incorporated Fisher's tetrasomic inheritance into the mapping framework, thus enhancing the cytological

relevance of QTL detection. Results from simulation studies showed that the method can be used to map QTLs in a controlled cross of multivalent tetraploids when the mapping population is adequately large (say 400). When a QTL undergoes double reduction, traditional mapping approaches will incorrectly estimate the position and effects of the QTL, proportional to the degree of double reduction. The new method can estimate the double reduction of a QTL, an important parameter related to the genetic diversity and evolution of polyploids [28,29].

Because of the high complexity of the mixture model implemented with tetrasomic inheritance, we only considered a one-marker model for QTL mapping. Interval mapping, which localizes a QTL with two flanking markers, has proven to be more advantageous in parameter estimation over the one-marker model [30]. It will be worthwhile to integrate components of our model into the interval mapping framework to fully explore the statistical merits of interval mapping for QTL mapping in multivalent tetraploids. Furthermore, the model proposed in this article assumes the segregation of fully informative codominant loci, each with 10 distinct genotypes, in a controlled cross of multivalent tetraploids. For partially informative codominant markers, a two-stage hierarchical mixture model will be needed to model the different allelic configurations for a phenotypically identical genotype. Although molecular marker technologies have improved in recent years, dominant markers may still be used in genetic mapping projects of some underrepresentative species including polyploids. Thus, it is also important to extend our model to map QTLs with dominant markers. For partially informative loci, the number of QTL genotypes may be unknown and, thus, a model selection procedure should be incorporated to determine the optimal number of genotypes at a QTL.

The genetic mapping of polyploids is complex because of their complex inheritance modes. Sophisticated statistical models are required to tackle genetic problems hidden in the polysomic inheritance of polyploids. Currently, there are some debates on the optimal modeling of tetrasomic inheritance in linkage analysis [13,25] and QTL mapping [18,31] partly

because of our limited knowledge about these fascinating species. Before a detailed understanding of the cytological mechanisms for meioses in multivalent polyploids is obtained, this type of debate will continue. In any case, the development of powerful statistical models for polyploid mapping continues to be a pressing need. The application of these models to real-world data will not only test their usefulness, but also provide an unprecedented opportunity to understand the genetic differentiation among polyploid genomes and characterize the genetic architecture of quantitatively inherited traits for this unique group of species. Software for the method described is available at http://statgen.psu.edu.

Table 1: Modes of gamete formation for a QTL flanked by two markers in a tetrasomic organism.

Mode	Typical Gamete	Number	Frequency
1	$a_1q_1b_1/a_1q_1b_1$	4	g_1
2	$a_1q_2b_1/a_1q_2b_1$	12	g_2
3	$a_1q_1b_1/a_1q_2b_1$	12	g_3
4	$a_1q_2b_1/a_1q_3b_1$	12	g_4
5	$a_1q_1b_2/a_1q_1b_2$	12	g_5
6	$a_1q_2b_2/a_1q_2b_2$	12	g_6
7	$a_1q_3b_2/a_1q_3b_2$	24	g_7
8	$a_1q_1b_2/a_1q_2b_2$	12	g_8
9	$a_1q_1b_2/a_1q_3b_2$	24	g_9
10	$a_1q_2b_2/a_1q_3b_2$	24	g_{10}
11	$a_1q_3b_2/a_1q_4b_2$	24	g_{11}
12	$a_1q_1b_1/a_1q_1b_2$	12	g_{12}
13	$a_1q_2b_1/a_1q_2b_2$	12	g_{13}
14	$a_1q_3b_1/a_1q_3b_2$	24	g_{14}
15	$a_1q_1b_1/a_1q_2b_2$	12	g_{15a}
16	$a_1q_2b_1/a_1q_1b_2$	12	g_{15b}
17	$a_1q_1b_1/a_1q_3b_2$	24	g_{16a}
18	$a_1q_3b_1/a_1q_1b_2$	24	g_{16b}
19	$a_1q_2b_1/a_1q_3b_2$	24	g_{17a}
20	$a_1q_3b_1/a_1q_2b_2$	24	g_{17b}
21	$a_1q_3b_1/a_1q_4b_2$	24	g_{18}
22	$a_1q_1b_2/a_1q_1b_3$	12	g_{19}
23	$a_1q_2b_2/a_1q_2b_3$	24	g_{20}
24	$a_1q_4b_2/a_1q_4b_3$	12	g_{21}
25	$a_1q_1b_2/a_1q_2b_3$	24	g_{22a}
26	$a_1q_2b_2/a_1q_1b_3$	24	g_{22b}
27	$a_1q_1b_2/a_1q_4b_3$	24	g_{23}
28	$a_1q_2b_2/a_1q_3b_3$	12	g_{24a}
29	$a_1q_3b_2/a_1q_2b_3$	12	g_{24b}
30	$a_1q_3b_2/a_1q_4b_3$	24	g_{25a}
31	$a_1q_4b_2/a_1q_3b_3$	24	g_{25b}
32	$a_1q_1b_1/a_2q_1b_1$	12	g_{26}
33	$a_1q_2b_1/a_2q_2b_1$	12	g_{27}
34	$a_1q_3b_1/a_2q_3b_1$	24	g_{28}
35	$a_1q_1b_1/a_2q_2b_1$	12	g_{29a}
36	$a_1q_2b_1/a_2q_1b_1$	12	g_{29b}

(To be continued)

Table 1: (Continued)

Mode	Typical Gamete	Number	Frequency
37	$a_1q_1b_1/a_2q_3b_1$	24	g_{30a}
38	$a_1q_3b_1/a_2q_1b_1$	24	g_{30b}
39	$a_1q_2b_1/a_2q_3b_1$	24	g_{31a}
40	$a_1q_3b_1/a_2q_2b_1$	24	g_{31b}
41	$a_1q_3b_1/a_2q_4b_1$	24	g_{32}
42	$a_2q_1b_1/a_3q_1b_1$	12	g_{33}
43	$a_2q_2b_1/a_3q_2b_1$	24	g_{34}
44	$a_2q_4b_1/a_3q_4b_1$	12	g_{35}
45	$a_2q_1b_1/a_3q_2b_1$	24	g_{36a}
46	$a_2q_2b_1/a_3q_1b_1$	24	g_{36b}
47	$a_2q_1b_1/a_3q_4b_1$	24	g_{37}
48	$a_2q_2b_1/a_3q_3b_1$	12	g_{38a}
49	$a_2q_3b_1/a_3q_2b_1$	12	g_{38b}
50	$a_2q_2b_1/a_3q_4b_1$	24	g_{39a}
51	$a_2q_4b_1/a_3q_2b_1$	24	g_{39b}
52	$a_1q_1b_1/a_2q_1b_2$	12	g_{40a}
53	$a_1q_1b_2/a_2q_1b_1$	12	g_{40b}
54	$a_1q_3b_1/a_2q_3b_2$	12	g_{41a}
55	$a_1q_3b_2/a_2q_3b_1$	12	g_{41b}
56	$a_1q_1b_1/a_2q_2b_2$	6	g_{42a}
57	$a_1q_2b_1/a_2q_1b_2$	6	g_{42b}
58	$a_1q_1b_2/a_2q_2b_1$	6	g_{42c}
59	$a_1q_2b_2/a_2q_1b_1$	6	g_{42d}
60	$a_1q_1b_1/a_2q_3b_2$	24	g_{43a}
61	$a_1q_3b_1/a_2q_1b_2$	24	g_{43b}
62	$a_1q_1b_2/a_2q_3b_1$	24	g_{43c}
63	$a_1q_3b_2/a_2q_1b_1$	24	g_{43d}
64	$a_1q_3b_1/a_2q_4b_2$	12	g_{44a}
65	$a_1q_4b_2/a_2q_3b_1$	12	g_{44b}
66	$a_1q_1b_1/a_2q_1b_3$	24	g_{45a}
67	$a_1q_1b_3/a_2q_1b_1$	24	g_{45b}
68	$a_1q_2b_1/a_2q_2b_3$	24	g_{46a}
69	$a_1q_2b_3/a_2q_2b_1$	24	g_{46b}
70	$a_1q_3b_1/a_2q_3b_3$	24	g_{47a}
71	$a_1q_3b_3/a_2q_3b_1$	24	g_{47b}
72	$a_1q_4b_1/a_2q_4b_3$	24	g_{48a}
73	$a_1q_4b_3/a_2q_4b_1$	24	g_{48b}

(To be continued)

Table 1: (Continued)

Mode	Typical Gamete	Number	Frequency
74	$a_1q_1b_1/a_2q_2b_3$	24	g_{49a}
75	$a_1q_2b_1/a_2q_1b_3$	24	g_{49b}
76	$a_1q_1b_3/a_2q_2b_1$	24	g_{49c}
77	$a_1q_2b_3/a_2q_1b_1$	24	g_{49d}
78	$a_1q_1b_1/a_2q_3b_3$	24	g_{50a}
79	$a_1q_3b_1/a_2q_1b_3$	24	g_{50b}
80	$a_1q_1b_3/a_2q_3b_1$	24	g_{50c}
81	$a_1q_3b_3/a_2q_1b_1$	24	g_{50d}
82	$a_1q_1b_1/a_2q_4b_3$	24	g_{51a}
83	$a_1q_4b_1/a_2q_1b_3$	24	g_{51b}
84	$a_1q_1b_3/a_2q_4b_1$	24	g_{51c}
85	$a_1q_4b_3/a_2q_1b_1$	24	g_{51d}
86	$a_1q_2b_1/a_2q_3b_3$	24	g_{52a}
87	$a_1q_3b_1/a_2q_2b_3$	24	g_{52b}
88	$a_1q_2b_3/a_2q_3b_1$	24	g_{52c}
89	$a_1q_3b_3/a_2q_2b_1$	24	g_{52d}
90	$a_1q_2b_1/a_2q_4b_3$	24	g_{53a}
91	$a_1q_4b_1/a_2q_2b_3$	24	g_{53b}
92	$a_1q_2b_3/a_2q_4b_1$	24	g_{53c}
93	$a_1q_4b_3/a_2q_2b_1$	24	g_{53d}
94	$a_1q_3b_1/a_2q_4b_3$	24	g_{54a}
95	$a_1q_4b_1/a_2q_3b_3$	24	g_{54b}
96	$a_1q_3b_3/a_2q_4b_1$	24	g_{54c}
97	$a_1q_4b_3/a_2q_3b_1$	24	g_{54d}
98	$a_1q_1b_3/a_2q_1b_4$	24	g_{55}
99	$a_1q_3b_3/a_2q_3b_4$	24	g_{56}
100	$a_1q_1b_3/a_2q_2b_4$	12	g_{57a}
101	$a_1q_2b_3/a_2q_1b_4$	12	g_{57b}
102	$a_1q_1b_3/a_2q_3b_4$	24	g_{58a}
103	$a_1q_3b_3/a_2q_1b_4$	24	g_{58b}
104	$a_1q_1b_4/a_2q_3b_3$	24	g_{58c}
105	$a_1q_3b_4/a_2q_1b_3$	24	g_{58d}
106	$a_1q_3b_3/a_2q_4b_4$	12	g_{59a}
107	$a_1q_4b_3/a_2q_3b_4$	12	g_{59b}

Table 2: Frequencies of the mixed gamete modes in terms of g_i.

$g_{15a} = (1-\phi_2)g_{15}$	$g_{15b} = \phi_2 g_{15}$
$g_{16a} = (1-\psi_2)g_{16}$	$g_{16b} = \psi_2 g_{16}$
$g_{17a} = \psi_2 g_{17}$	$g_{17b} = (1-\psi_2)g_{17}$
$g_{22a} = \psi_2 g_{22}$	$g_{22b} = (1-\psi_2)g_{22}$
$g_{24a} = (1-\phi_2)g_{24}$	$g_{24b} = \phi_2 g_{24}$
$g_{25a} = \psi_2 g_{25}$	$g_{25b} = (1-\psi_2)g_{25}$
$g_{29a} = (1-\phi_1)g_{29}$	$g_{29b} = \phi_1 g_{29}$
$g_{30a} = (1-\psi_1)g_{30}$	$g_{30b} = \psi_1 g_{30}$
$g_{31a} = \psi_1 g_{31}$	$g_{31b} = (1-\psi_1)g_{31}$
$g_{36a} = \psi_1 g_{36}$	$g_{36b} = (1-\psi_1)g_{36}$
$g_{38a} = (1-\phi_1)g_{38}$	$g_{38b} = \phi_1 g_{38}$
$g_{39a} = (1-\psi_1)g_{39}$	$g_{39b} = \psi_1 g_{39}$
$g_{40a} = \frac{1}{2}g_{40}$	$g_{40b} = \frac{1}{2}g_{40}$
$g_{41a} = \frac{1}{2}g_{41}$	$g_{41b} = \frac{1}{2}g_{41}$
$g_{42a} = (1-\phi_1)(1-\phi_2)g_{42}$	$g_{42b} = \phi_1\phi_2 g_{42}$
$g_{42c} = (1-\phi_1)\phi_2 g_{42}$	$g_{42d} = \phi_1(1-\phi_2)g_{42}$
$g_{43a} = (1-\psi_1)(1-\psi_2)g_{43}$	$g_{43b} = \psi_1\psi_2 g_{43}$
$g_{43c} = (1-\psi_1)\psi_2 g_{43}$	$g_{43d} = \psi_1(1-\psi_2)g_{43}$
$g_{44a} = \frac{1}{2}g_{44}$	$g_{44b} = \frac{1}{2}g_{44}$
$g_{45a} = \frac{1}{2}g_{45}$	$g_{45b} = \frac{1}{2}g_{45}$
$g_{46a} = \frac{1}{2}g_{46}$	$g_{46b} = \frac{1}{2}g_{46}$
$g_{47a} = \frac{1}{2}g_{47}$	$g_{47b} = \frac{1}{2}g_{47}$
$g_{48a} = \frac{1}{2}g_{48}$	$g_{48b} = \frac{1}{2}g_{48}$
$g_{49a} = (1-\phi_1)(1-\psi_2)g_{49}$	$g_{49b} = \phi_1\psi_2 g_{49}$
$g_{49c} = (1-\phi_1)\psi_2 g_{49}$	$g_{49d} = \phi_1(1-\psi_2)g_{49}$
$g_{50a} = (1-\psi_1)(1-\phi_2)g_{50}$	$g_{50b} = \psi_1\phi_2 g_{50}$
$g_{50c} = (1-\psi_1)\phi_2 g_{50}$	$g_{50d} = \psi_1(1-\phi_2)g_{50}$
$g_{51a} = (1-\psi_1)(1-\psi_2)g_{51}$	$g_{51b} = \psi_1\psi_2 g_{51}$
$g_{51c} = (1-\psi_1)\psi_2 g_{51}$	$g_{51d} = \psi_1(1-\psi_2)g_{51}$
$g_{52a} = \psi_1(1-\psi_2)g_{52}$	$g_{52b} = (1-\psi_1)\psi_2 g_{52}$
$g_{52c} = \psi_1\psi_2 g_{52}$	$g_{52d} = (1-\psi_1)(1-\psi_2)g_{52}$
$g_{53a} = \frac{1}{2}\psi_1 g_{53}$	$g_{53b} = \frac{1}{2}(1-\psi_1)g_{53}$
$g_{53c} = \frac{1}{2}\psi_1 g_{53}$	$g_{53d} = \frac{1}{2}(1-\psi_1)g_{53}$
$g_{54a} = \frac{1}{2}\psi_2 g_{54}$	$g_{54b} = \frac{1}{2}(1-\psi_2)g_{54}$
$g_{54c} = \frac{1}{2}(1-\psi_2)g_{54}$	$g_{54d} = \frac{1}{2}\psi_2 g_{54}$
$g_{57a} = (1-\phi_1)g_{57}$	$g_{57b} = \phi_1 g_{57}$
$g_{58a} = (1-\psi_1)\psi_2 g_{58}$	$g_{58b} = \psi_1(1-\psi_2)g_{58}$
$g_{58c} = (1-\psi_1)(1-\psi_2)g_{58}$	$g_{58d} = \psi_1\psi_2 g_{58}$
$g_{59a} = (1-\phi_2)g_{59}$	$g_{59b} = \phi_2 g_{59}$

Table 3: Average parameter estimates and their standard errors of the QTL mapping model for a pseudotest backcross population in a multivalent tetraploid based on 1000 repeat simulations. One QTL with the heritability of 0.1 is set at the center between two markers with 20 cM genetic distance and the true genetic parameters are set as in the second column.

		Estimate ($H^2 = 0.1$)		
Parameter	True Value	$N = 200$	$N = 400$	$N = 800$
Position (cM)	10.00	10.3833 (4.5672)	10.3609 (4.2539)	9.8640 (3.4637)
α	0.20	0.2057 (0.0274)	0.2015 (0.0203)	0.2003 (0.0141)
β	0.30	0.3002 (0.0320)	0.2998 (0.0233)	0.2993 (0.0161)
γ	0.25	0.2323 (0.0957)	0.2435 (0.0856)	0.2486 (0.0725)
μ_1	1.40	1.3844 (1.5359)	1.3977 (1.0259)	1.3953 (0.5976)
μ_2	1.50	1.5262 (1.6630)	1.5349 (1.0210)	1.4702 (0.6420)
μ_3	1.60	1.6103 (1.6160)	1.6262 (0.9470)	1.6204 (0.6126)
μ_4	-0.50	-0.6922 (1.5663)	-0.6582 (0.9799)	-0.5853 (0.6376)
μ_5	2.20	2.1970 (0.6605)	2.2326 (0.4623)	2.2076 (0.3174)
μ_6	2.40	2.4325 (0.6819)	2.4070 (0.4673)	2.4095 (0.3330)
μ_7	0.40	0.3796 (0.6694)	0.4169(0.4648)	0.3927 (0.3346)
μ_8	2.50	2.5693 (0.6695)	2.5198 (0.4558)	2.5039 (0.3133)
μ_9	0.50	0.4772 (0.6651)	0.4754 (0.4432)	0.5072 (0.3150)
μ_{10}	0.60	0.6058 (0.6628)	0.5639 (0.4773)	0.5795 (0.3289)
σ^2	7.9706	7.2230 (0.8012)	7.5642 (0.5626)	7.7776 (0.4070)
μ	1.00	0.9572 (0.7829)	0.9751 (0.4930)	0.9752 (0.3172)
a_1	0.40	0.4272 (1.3372)	0.4226 (0.8826)	0.4202 (0.5370)
a_2	0.50	0.5691 (1.4557)	0.5597 (0.8747)	0.4951 (0.5387)
a_3	0.60	0.6531 (1.3736)	0.6510 (0.8459)	0.6453 (0.5400)
d_{12}	0.30	0.2436 (2.0350)	0.2751 (1.3405)	0.3172 (0.7940)
d_{13}	0.40	0.3950(1.9783)	0.3582 (1.2350)	0.3690 (0.8003)
d_{14}	0.50	0.6445(2.0169)	0.6525 (1.2923)	0.5579 (0.8053)
d_{23}	0.40	0.3899 (2.0609)	0.3339 (1.2589)	0.3884 (0.7982)
d_{24}	0.50	0.6003 (1.9119)	0.5739 (1.2320)	0.5974 (0.8692)
d_{34}	0.50	0.6449 (2.0174)	0.5711 (1.3135)	0.5195 (0.8142)

Table 4: Average parameter estimates and their standard errors of the QTL mapping model for a pseudotest backcross population in a multivalent tetraploid based on 1000 repeat simulations. One QTL with the heritability of 0.4 is set at the center between two markers with 20 cM genetic distance and the true genetic parameters are set as in the second column.

		Estimate ($H^2 = 0.4$)		
Parameter	True Value	$N = 200$	$N = 400$	$N = 800$
Position (cM)	10.00	11.0063 (3.4189)	10.3719 (2.4893)	10.1372(1.7101)
α	0.20	0.2039 (0.0272)	0.2025 (0.0193)	0.2010 (0.0142)
β	0.30	0.3027 (0.0320)	0.3005 (0.0239)	0.3009 (0.0160)
γ	0.25	0.2549 (0.0664)	0.2567 (0.0485)	0.2557 (0.0336)
μ_1	1.40	1.3797 (0.5068)	1.3879 (0.3260)	1.3954 (0.2103)
μ_2	1.50	1.4619 (0.5180)	1.4957 (0.3273)	1.5087 (0.2129)
μ_3	1.60	1.5751(0.5042)	1.5860 (0.2999)	1.6017 (0.2100)
μ_4	-0.50	-0.5649 (0.5379)	-0.5511 (0.3334)	-0.5128 (0.2235)
μ_5	2.20	2.2136 (0.2736)	2.2131 (0.1869)	2.2131 (0.1286)
μ_6	2.40	2.4140 ((0.2568)	2.4116 (0.1812)	2.4085 (0.1340)
μ_7	0.40	0.3819 (0.2611)	0.3890 (0.1817)	0.3987 (0.1324)
μ_8	2.50	2.5146 (0.2671)	2.5222 (0.1875)	2.5103 (0.1262)
μ_9	0.50	0.4747 (0.2644)	0.5032 (0.1829)	0.5046 (0.1241)
μ_{10}	0.60	0.5936 (0.2682)	0.5964 (0.1843)	0.5931 (0.1228)
σ^2	1.3284	1.1881 (0.1341)	1.2603 (0.1015)	1.2909 (0.0694)
μ	1.00	0.9629 (0.2635)	0.9796 (0.1660)	0.9983 (0.1103)
a_1	0.40	0.4168 (0.4434)	0.4083 (0.2804)	0.3971 (0.1822)
a_2	0.50	0.4989 (0.4400)	0.5161 (0.2810)	0.5105 (0.1849)
a_3	0.60	0.6121 (0.4412)	0.6064 (0.2647)	0.6035 (0.1822)
d_{12}	0.30	0.3350 (0.6807)	0.3092 (0.4293)	0.3072 (0.2907)
d_{13}	0.40	0.4222 (0.6300)	0.4174 (0.4093)	0.4097 (0.2809)
d_{14}	0.50	0.5300 (0.6530)	0.5319 (0.4365)	0.5144 (0.2961)
d_{23}	0.40	0.4406 (0.6535)	0.4202 (0.4085)	0.3980 (0.2796)
d_{24}	0.50	0.5407 (0.6758)	0.5382 (0.4318)	0.5069 (0.2832)
d_{34}	0.50	0.5463 (0.6795)	0.5412 (0.4078)	0.5024 (0.2833)