Theor Appl Genet (2005) 110: 1061–1067
DOI 10.1007/s00122-005-1926-7

ORIGINAL PAPER

J. Yu · M. Arbelbide · R. Bernardo

Power of in silico QTL mapping from phenotypic, pedigree, and marker data in a hybrid breeding program

Received: 6 September 2004 / Accepted: 6 January 2005 / Published online: 8 March 2005

Abstract Most quantitative trait locus (QTL) mapping studies in plants have used designed mapping populations. As an alternative to traditional QTL mapping, in silico mapping via a mixed-model approach simultaneously exploits phenotypic, genotypic, and pedigree data already available in breeding programs. The statistical power of this in silico mapping method, however, remains unknown. Our objective was to evaluate the power of in silico mapping via a mixed-model approach in hybrid crops. We used maize (*Zea mays* L.) as a model species to study, by computer simulation, the influence of number of QTLs (20 or 80), heritability (0.40 or 0.70), number of markers (200 or 400), and sample size (600 or 2,400 hybrids). We found that the average power to detect QTLs ranged from 0.11 to 0.59 for a significance level of $\alpha = 0.01$, and from 0.01 to 0.47 for $\alpha = 0.0001$. The false discovery rate ranged from 0.22 to 0.74 for $\alpha = 0.01$, and from 0.05 to 0.46 for $\alpha = 0.0001$. As with designed mapping experiments, a large sample size, high marker density, high heritability, and small number of QTLs led to the highest power for in silico mapping via a mixed-model approach. The power to detect QTLs with large effects was greater than the power to detect QTL with small effects. We conclude that gene discovery in hybrid crops can be initiated by in silico mapping. Finding an acceptable compromise, however, between the power to detect QTL and the proportion of false QTL would be necessary.

Communicated by H.C. Becker

J. Yu · M. Arbelbide · R. Bernardo (✉)
Department of Agronomy and Plant Genetics,
University of Minnesota, 411 Borlaug Hall,
1991 Upper Buford Circle,
St. Paul, MN 55108, USA
E-mail: berna022@umn.edu
Tel.: +1-612-6256282
Fax: +1-612-6251268

Present address: J. Yu
Institute for Genomic Diversity, Cornell University,
157 Biotechnology Building, Ithaca, NY 14853, USA

Introduction

Quantitative trait loci (QTLs) mapping is useful for dissecting complex traits [Lander and Botstein 1989; Lynch and Walsh 1998, (p 379); Mackay 2001]. Most QTL mapping studies in plants have used designed mapping populations, such as F_2 or backcross populations between two inbreds (Kearsey and Farquhar 1998). As an alternative approach, in silico mapping aims to exploit existing phenotypic and genomic databases to discover QTLs (Grupe et al. 2001).

The mixed-model approach, which was developed to exploit massive amounts of phenotypic and pedigree data in animal breeding (Henderson 1984), has been successfully adapted in plants (Panter and Allen 1995; Bernardo 1996). The integration of genomic data in the mixed-model approach for the purpose of QTL mapping in hybrid crops was first suggested by Bernardo (1998). Recently, Parisseaux and Bernardo (2004) found that in silico mapping via a mixed-model approach can detect repeatable associations across different populations. Specifically, they attempted to identify simple sequence repeat (SSR) markers associated with different traits in maize (*Zea mays* L.) by utilizing the following data already available in a private breeding program: (1) multilocation phenotypic data for 22,774 single-cross hybrids; (2) SSR marker data at 96 loci for the 1,266 parental inbreds of the single-cross hybrids; (3) pedigree records for the 1,266 parental inbreds. By in silico mapping via a mixed-model approach, they detected 37 SSR markers with significant effects for plant height, 24 for smut [*Ustilago maydis* (DC.) Cda.] resistance, and 44 for grain moisture.

The statistical power of in silico mapping via the Parisseaux and Bernardo (2004) approach, however, remains unknown. It has been shown that the heritability and genetic architecture (e.g., number of QTLs and distribution of effects) of the trait and the resources available for QTL mapping (e.g., sample size and number of markers) affect the statistical power of

designed QTL mapping experiments (Lander and Thompson 1990; Beavis 1994; Utz and Melchinger 1994). These genetic and non-genetic factors are also expected to affect the power of in silico mapping via a mixed-model approach. Moreover, the false discovery rate (FDR), which is the probability of a QTL being false given that a QTL has been declared, has recently been proposed as another criterion to assess the ability to detect QTLs (Benjamini and Hochberg 1995; Fernando et al. 2004). Our objective was to evaluate, by computer simulation, the power of in silico mapping via a mixed-model approach in hybrid crops. We used maize as a model species, but the results should generally apply to other hybrid crops.

Materials and methods

Maize breeding comprises two stages, inbred development and hybrid testing (Hallauer 1990). During inbred development, pairs of elite inbreds that belong to the same heterotic group are crossed to form an F_2 or backcross population from which new inbreds are developed. Lines are developed by selfing and are crossed to one or two inbred testers from a different heterotic group. The testcrosses are then evaluated in field trials at several locations (Smith et al. 1999). During hybrid testing, single-cross hybrids made from pairs of inbreds from different heterotic groups are evaluated in multilocation performance trials. The tested hybrids, however, often account for only 10% to 15% of all potential combinations of single-crosses (Bernardo 1996).

Our simulation mimicked this two-stage breeding process in maize. First, we considered two opposite heterotic groups, each having a total of $n_1 = n_2 = 112$ inbreds developed from different ancestral inbreds. Second, we assumed that $n = 600$ or 2,400 hybrids, among all potential single-cross hybrids ($112 \times 112 = 12{,}544$) between the two heterotic groups, had data available from multilocation performance trails. The number of inbreds in each heterotic group and the number of hybrids with available phenotypic data were chosen to agree with the empirical data of Parisseaux and Bernardo (2004).

We conducted a total of 64 simulation experiments. These 64 experiments had contrasting values of six different parameters: level of initial linkage disequilibrium ($t = 10$ or 20 generations of random mating), significance level ($\alpha = 0.01$ or 0.0001), number of QTLs ($l = 20$ or 80), heritability ($H = 0.40$ or 0.70), number of markers ($m = 200$ or 400), and sample size ($n = 600$ or 2,400 hybrids). For each experiment, 50 runs were conducted with different locations of QTLs and markers on the genetic map and different inbreds and hybrids. Data from each run of a particular experiment were individually analyzed with the mixed-model method and the results from 50 runs were then summarized. We wrote a simulation program in C++ and conducted the simulation and data analysis on an IBM Power4 supercomputer at the Supercomputing Institute for Digital Simulation and Advanced Computation, University of Minnesota.

Inbred development

Two ancestral inbreds in heterotic group 1 were crossed and random-mated for $t = 10$ or 20 generations. Likewise, two ancestral inbreds in heterotic group 2 were crossed and random-mated for $t = 10$ or 20 generations. The purpose of considering both $t = 10$ and $t = 20$ was to create different levels of initial linkage disequilibrium (i.e., high for ten generations of random mating and low for 20 generations of random mating) between the QTL and markers. A total of 16 founder inbreds were obtained through single-seed descent from the random-mated F_2 population of each heterotic group. These founder inbreds were denoted by I_1^1, I_1^2, ..., I_1^{16} for heterotic group 1 and I_2^1, I_2^2, ..., I_2^{16} for heterotic group 2. For heterotic group 1, chain crosses were made among the 16 founder inbreds, i.e., $I_1^1 \times I_1^2$, $I_1^2 \times I_1^3$, ..., $I_1^{16} \times I_1^1$. The recombinant inbreds from $I_1^1 \times I_1^2$ were then testcrossed to I_2^1, the recombinant inbreds from $I_1^2 \times I_1^3$ testcrossed to I_2^2, and so on. Heritability on a testcross mean-basis (across testing locations) during inbred development was 0.40 or 0.70. Based on testcross performance, the best two recombinant inbreds from each cross were selected, resulting in a total of 32 second-cycle inbreds. Using the same chain-cross and testing system, a total of 64 third-cycle inbreds were developed. The above process for inbred development was simultaneously performed for the heterotic group 2. Coefficients of coancestry among inbreds within each heterotic group were calculated from pedigree records by tabular analysis (Emik and Terrill 1949).

Hybrid testing

A total of $n = 600$ or 2,400 hybrids were assumed evaluated in P performance trials. Each performance trial included 30 different hybrids. The effects of performance trials were randomly drawn from a normal distribution with zero mean, and the variance of performance trial effects was scaled to account for 70% of the total variation (Delacy and Cooper 1990). The heritability was adjusted upward from $H = 0.40$ or 0.70 during inbred development to $H = 0.67$ or 0.88 during hybrid testing. This adjustment was necessary given that the number of locations in each performance trial is about three times greater during hybrid testing than during inbred development (Smith et al. 1999).

Genetic model

We considered a published maize linkage map with 1,749 cM for ten chromosomes (Senior et al. 1996).

A total of $l = 20$ or 80 QTLs and $m = 200$ or 400 markers were randomly located across the genome. Four alleles were present at each QTL or marker locus, but each heterotic group has two alleles at a locus. The two ancestral inbreds in heterotic group 1 had a QTL genotype of Q_1Q_1 or Q_4Q_4 and a marker genotype of M_1M_1 or M_4M_4 at the odd-numbered loci, and a QTL genotype of Q_2Q_2 or Q_3Q_3 and a marker genotype of M_2M_2 or M_3M_3 at the even-numbered loci. The two ancestral inbreds in heterotic group 2 had a QTL genotype of Q_2Q_2 or Q_3Q_3 and a marker genotype of M_2M_2 or M_3M_3 at the odd-numbered loci, and a QTL genotype of Q_1Q_1 or Q_4Q_4 and a marker genotype of M_1M_1 or M_4M_4 at the even-numbered loci. For single crosses between heterotic groups, the four possible genotypes were Q_1Q_2, Q_1Q_3, Q_2Q_4, and Q_3Q_4 at each QTL locus, and M_1M_2, M_1M_3, M_2M_4, M_3M_4 at each marker locus.

The effects of QTL followed a geometric series. The effect of the ith QTL was a function of a^i where $a = 0.9$ for $l = 20$ QTL, and $a = 0.98$ for $l = 80$ QTL (Lande and Thompson 1990). The genotypic values for the four homozygous genotypes at each QTL were a^i for Q_1Q_1, $1/2\ a^i$ for Q_2Q_2, $-1/2\ a^i$ for Q_3Q_3 and $-a^i$ for Q_4Q_4. Assuming complete dominance among alleles at each QTL, the genotypic values at the ith QTL for single-cross hybrids were a^i for Q_1Q_2, a^i for Q_1Q_3, $1/2\ a^i$ for Q_2Q_4, and $-1/2\ a^i$ for Q_3Q_4. The differences in testcross general combining ability (GCA) effects were a^i between Q_1 and Q_4, and $1/2\ a^i$ between Q_2 and Q_3. The testcross dominance deviations were $-1/4\ a^i$ for Q_1Q_2, $1/4\ a^i$ for Q_1Q_3, $1/4\ a^i$ for Q_2Q_4, and $-1/4\ a^i$ for Q_3Q_4.

Mixed model

The mixed model used for analysis with $k < m$ markers was

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{M}_1\alpha_1 + \mathbf{M}_2\alpha_2 + \mathbf{M}_3\delta + \mathbf{Z}_1\mathbf{g}_1 + \mathbf{Z}_2\mathbf{g}_2 + \mathbf{e}$$

where $\mathbf{y} = n\times1$ vector of observed performance for a given trait; $\beta = p\times1$ vector of fixed effects associated with performance trials; $\alpha_1 = 2k\times1$ vector of GCA effects associated with the marker alleles in heterotic group 1; $\alpha_2 = 2k\times1$ vector of GCA effects associated with the marker alleles in heterotic group 2; $\delta = 4k\times1$ dominance deviation vector for four hybrid genotypes; $\mathbf{g}_1 = n_1\times1$ vector of background GCA effects, not associated with the marker being tested, of the inbreds in heterotic group 1; $\mathbf{g}_2 = n_2\times1$ vector of background GCA effects, not associated with the marker being tested, of the inbreds in heterotic group 2; $\mathbf{e} = n\times1$ vector of residual effects; $\mathbf{X}$, $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$, $\mathbf{Z}_1$, and $\mathbf{Z}_2$ were incidence matrices of 1s and 0s relating $\mathbf{y}$ to β, α_1, α_2, δ, $\mathbf{g}_1$ and $\mathbf{g}_2$, respectively.

The variances of the random effects were $\mathrm{Var}(\mathbf{g}_1) = \mathbf{G}_1V_{GCA(1)}$, $\mathrm{Var}(\mathbf{g}_2) = \mathbf{G}_2V_{GCA(2)}$, and $\mathrm{Var}(\mathbf{e}) = \mathbf{R}\ V_R$, where $\mathbf{G}_1 = n_1\times n_1$ matrix of coefficients of coancestry among group-1 inbreds; $\mathbf{G}_2 = n_2\times n_2$ matrix of coefficients of coancestry among group-2 inbreds; $\mathbf{R} = n\times n$ matrix with the off-diagonal elements being zero and the diagonal elements being the reciprocal of the number of locations from which each phenotypic data point was obtained. Best linear unbiased estimates (BLUE) of β, α_1, α_2, and δ (fixed effects), and best linear unbiased predictions (BLUP) of $\mathbf{g}_1$ and $\mathbf{g}_2$ (random effects) were obtained by solving the mixed-model equations for single crosses (Henderson 1985). Restricted maximum likelihood (REML) estimates of the variances were obtained through iteration (Henderson 1985). We assumed equal numbers of testing locations (s) among performance trials. Correspondingly, $\mathbf{R}$ became an identity matrix ($\mathbf{I}$) and $\mathrm{Var}(\mathbf{e})$ was equal to $\mathbf{I}(V_R/s)$.

Data analysis

In silico mapping via a mixed-model approach comprised three steps. In the first step, a mixed model ignoring the marker data (i.e., without α_1, α_2, and δ in the model) was fitted to obtain estimates of V_R, $V_{GCA(1)}$, and $V_{GCA(2)}$ using the above formulas. In the second step, single-marker analysis was performed for each marker, using $V_R/V_{GCA(1)}$ and $V_R/V_{GCA(2)}$ ratios obtained from the first step rather than from new estimates obtained by iteration. Using REML estimates of $V_R/V_{GCA(1)}$ and $V_R/V_{GCA(2)}$ ratios obtained from the first step reduced the computational time 200- to 400-fold, with little impact on the markers chosen for the multiple-marker analysis (i.e., the third step). An F-test was performed to test the significance ($\alpha = 0.01$ or 0.0001) of the marker effects (α_1, α_2, and δ) as described by Kennedy et al. (1992). To reduce multicollinearity, only the marker with the most significant P-value was chosen if several adjacent markers were significant.

In the third step, multiple-marker analysis was performed by simultaneously fitting effects for those markers retained from the single-marker analysis. The BLUE of fixed effects and BLUP of random effects were obtained by solving the mixed-model equations until convergence of V_R, $V_{GCA(1)}$, and $V_{GCA(2)}$. An F-test at the $\alpha = 0.01$ or 0.0001 significance levels was performed for the fixed marker effects. Based on the complete dominance model, $\mathbf{a} = 1.5\times(\alpha_1 + \alpha_2 + \delta)$ was considered the estimated effect for the significant marker. The power of in silico mapping via a mixed-model approach was evaluated based on the results from the final multiple-marker analysis. A marker is expected to have a significant regression coefficient only if it is adjacent to a QTL (Doerge et al. 1994; Whittaker et al. 1996). A true positive was therefore declared if a marker had at least one significant regression coefficient for α_1, α_2, or δ and a QTL was present in either or both of the marker's adjacent intervals. A false positive was declared if a marker had at least one significant regression coefficient for α_1, α_2, or δ but no QTL was present in either of the marker's adjacent intervals.

In this manuscript, power refers to the ability to detect a given QTL, whereas average power refers to the

mean power across QTL. Power was calculated as the proportion of the number of times a particular QTL was detected out of 50 runs. The correlation was calculated between the true effect of a QTL and the power to detect that QTL. Average power was calculated as the number of true positives divided by the total number of QTLs simulated, averaged across 50 runs. The FDR was calculated as the number of false positives divided by the total number of significant markers detected, averaged across 50 runs.

For a true positive, the true effect of a significant marker was assumed to be equal to the effect of the adjacent QTL (a^i). The true effect was zero for a false positive. The correlation was calculated between the estimated effects and the true effects of significant markers for each run whenever possible—i.e., more than two significant markers were identified for that particular run. The mean and median of these correlation coefficients across 50 runs were calculated. A chi-square test was applied to the ratio between the number of times the effect was overestimated and the number of times the effect was underestimated, given the null hypothesis of a 1:1 ratio.

Results

The average map distance between loci (both QTLs and markers) under different combinations of the number of QTLs and the number of markers ranged from 3.6 cM to 8.0 cM (Table 1). The corresponding recombination frequencies (r) in the initial F_2 population between the ancestral inbreds, for the Kosambi mapping function, ranged from 0.04 to 0.08. The linkage disequilibrium before random mating (D_0) at the genome level was similar among different combinations of the number of QTLs and markers. After $t = 10$ generations of random mating, the linkage disequilibrium [$D_t = D_0(1 - r)^t$] varied from 0.09 to 0.16. After $t = 20$ generations of random mating, D_t varied from 0.04 to 0.11.

Both the number of QTLs controlling a trait and the heritability of the trait had a prominent effect on the average power of in silico mapping (Table 2). The average power decreased as the number of QTLs increased. The average power across experiments (i.e., averaged across experiments at a specified level of a factor) decreased from 0.34 when 20 QTLs controlled the trait, to 0.17 when 80 QTLs controlled the trait. This decrease in the average power was more evident at more stringent significance levels. The FDR across experiments, on the other hand, decreased from 0.41 when 20 QTLs controlled the trait, to 0.23 when 80 QTLs controlled the trait. Moreover, the average power increased as the heritability increased. Increasing heritability from 0.40 to 0.70 led to an increase in the average power across experiments from 0.21 to 0.29, but it led to only a slightly increase for the FDR, from 0.30 to 0.34. Notably, the gain in average power (37%) outweighed the loss in FDR (15%).

Having more resources available for in silico mapping led to a higher average power. When the number of markers increased from 200 to 400, the average power across experiments increased from 0.21 to 0.29 (Table 2). The FDR across experiments, on the other hand, increased from 0.27 to 0.37. With a sample size of 600 tested hybrids, the average power across experiments was 0.19 and the FDR across experiments was 0.26. Increasing the sample size to 2,400 tested hybrids led to a higher average power across experiments (0.32), but also to a higher FDR across experiments (0.38). The gain in the average power across experiments (69%), however, still outweighed the loss in the FDR across experiments (45%).

We found a strong association between the average power and the FDR. An increase in the average power at different levels of each parameter generally led to an increase in the FDR, and vice versa. However, when a high marker density and a large sample size were used, a more stringent α level allowed a compromise between the average power and the FDR. With $\alpha = 0.01$, the average power across experiments was 0.33 and the FDR across experiments was 0.47 (Table 2). A more stringent significance level of $\alpha = 0.0001$ led to an average power across experiments of 0.17, half of that at $\alpha = 0.01$, and an FDR across experiments of 0.17, a 63% decrease from $\alpha = 0.01$ (Table 2). At both significance levels, the maximum average power (0.59 for $\alpha = 0.01$, and 0.47 for $\alpha = 0.0001$) was achieved when a trait was controlled by 20 QTLs and had a heritability of 0.70, linkage disequilibrium was high (i.e., ten generations of random mating), and in silico mapping was conducted with 400 markers and 2,400 hybrids. The minimum average power (0.11 for $\alpha = 0.01$, and 0.01 for $\alpha = 0.0001$) was achieved when a trait was controlled by 80 QTL and had a

Table 1 Genome-wide linkage disequilibrium before and after 10 or 20 generations of random mating with different numbers of QTLs and markers

Number of QTLs	Number of markers	Map distance (cM)	Recombination frequency	Linkage disequilibrium		
				Before random mating (D_0)	After 10 generations (D_{10})	After 20 generations (D_{20})
20	200	8.0	0.08	0.21	0.09	0.04
	400	4.2	0.04	0.23	0.15	0.10
80	200	6.2	0.06	0.22	0.12	0.06
	400	3.6	0.04	0.23	0.16	0.11

Table 2 Average power, false discovery rate (FDR), mean and median (in parentheses) of the correlation between estimated and true effects, and χ^2 test for the number of underestimates versus the number of overestimates from a ratio of 1:1, using in silico mapping via a mixed-model approach

Significance level (α)	Linkage disequilibrium	Number of QTLs	Number of markers	Heritability	Number of hybrids	Average power	FDR	Correlation	χ^2
$\alpha = 0.01$	High[a]	20	200	0.40	600	0.28	0.44	0.68 (0.71)	NS[b]
					2400	0.43	0.54	0.72 (0.76)	U**
				0.70	600	0.41	0.45	0.72 (0.76)	NS
					2,400	0.49	0.62	0.74 (0.76)	U**
			400	0.40	600	0.38	0.58	0.72 (0.76)	NS
					2,400	0.53	0.67	0.77 (0.79)	NS
				0.70	600	0.48	0.59	0.78 (0.80)	NS
					2,400	0.59	0.73	0.77 (0.77)	NS
		80	200	0.40	600	0.13	0.24	0.31 (0.38)	**
					2,400	0.23	0.24	0.32 (0.34)	**
				0.70	600	0.20	0.22	0.34 (0.36)	**
					2,400	0.29	0.28	0.36 (0.38)	**
			400	0.40	600	0.18	0.39	0.34 (0.34)	**
					2,400	0.31	0.41	0.36 (0.38)	**
				0.70	600	0.26	0.35	0.36 (0.36)	**
					2,400	0.37	0.45	0.42 (0.43)	**
	Low	20	200	0.40	600	0.20	0.52	0.67 (0.67)	NS
					2,400	0.36	0.59	0.65 (0.65)	NS
				0.70	600	0.30	0.49	0.69 (0.73)	NS
					2,400	0.43	0.64	0.67 (0.73)	NS
			400	0.40	600	0.33	0.62	0.68 (0.70)	*
					2,400	0.48	0.68	0.73 (0.74)	*
				0.70	600	0.42	0.63	0.71 (0.73)	NS
					2,400	0.57	0.74	0.73 (0.74)	NS
		80	200	0.40	600	0.11	0.24	0.30 (0.36)	**
					2,400	0.21	0.30	0.35 (0.35)	**
				0.70	600	0.17	0.23	0.34 (0.34)	**
					2,400	0.26	0.34	0.36 (0.37)	**
			400	0.40	600	0.15	0.39	0.35 (0.35)	**
					2,400	0.31	0.42	0.38 (0.40)	**
				0.70	600	0.23	0.39	0.39 (0.39)	**
					2,400	0.39	0.48	0.39 (0.39)	**
$\alpha = 0.0001$	High	20	200	0.40	600	0.12	0.11	0.67 (0.87)	**
					2,400	0.25	0.21	0.62 (0.73)	NS
				0.70	600	0.21	0.09	0.56 (0.71)	**
					2,400	0.36	0.36	0.73 (0.75)	NS
			400	0.40	600	0.19	0.08	0.47 (0.64)	**
					2,400	0.37	0.31	0.73 (0.78)	*
				0.70	600	0.30	0.13	0.59 (0.70)	**
					2,400	0.47	0.46	0.81 (0.85)	NS
		80	200	0.40	600	0.02	0.07	0.21 (0.48)	**
					2,400	0.09	0.09	0.24 (0.21)	**
				0.70	600	0.05	0.05	0.16 (0.25)	**
					2,400	0.16	0.13	0.31 (0.37)	**
			400	0.40	600	0.03	0.17	0.17 (0.46)	**
					2,400	0.15	0.14	0.34 (0.38)	**
				0.70	600	0.08	0.12	0.19 (0.17)	**
					2,400	0.25	0.24	0.44 (0.47)	**
	Low	20	200	0.40	600	0.09	0.08	0.50 (0.77)	**
					2,400	0.19	0.23	0.70 (0.80)	NS
				0.70	600	0.13	0.11	0.64 (0.82)	**
					2,400	0.28	0.40	0.73 (0.82)	**
			400	0.40	600	0.15	0.10	0.49 (0.77)	NS
					2,400	0.30	0.24	0.75 (0.85)	NS
				0.70	600	0.24	0.17	0.64 (0.70)	**
					2,400	0.42	0.45	0.81 (0.83)	NS
		80	200	0.40	600	0.01	0.05	−0.47 (−0.90)	**
					2,400	0.07	0.11	0.28 (0.25)	**
				0.70	600	0.03	0.06	0.34 (0.43)	**
					2,400	0.14	0.18	0.38 (0.38)	**
			400	0.40	600	0.02	0.09	0.01 (0.18)	**
					2,400	0.12	0.17	0.34 (0.39)	**
				0.70	600	0.06	0.12	0.35 (0.37)	**
					2,400	0.23	0.26	0.44 (0.46)	**

*, ** χ^2 test for a ratio of 1:1 significant at the $P < 0.05$, and $P < 0.01$ levels, respectively

[a] Founder inbreds developed by ten generations (high linkage disequilibrium) or 20 generations (low linkage disequilibrium) of random mating

[b] NS, Not significant; U, more underestimates, otherwise more overestimates

heritability of 0.40, linkage disequilibrium was low (i.e., 20 generations of random mating), and in silico mapping was conducted with 200 markers and 600 hybrids. The FDR closely followed this pattern for average power.

The mean of correlation coefficients between the true effects and the estimated effects was moderately high when 20 QTLs controlled a trait but was low when 80 QTLs controlled a trait (Table 2). The median of correlation coefficients followed the same pattern. Average power was low when a trait was controlled by 80 QTLs and had a heritability of 0.40, linkage disequilibrium was low (i.e., 20 generations of random mating), and in silico mapping was conducted with 200 markers and 600 hybrids at the $\alpha = 0.0001$ significance level (Table 2). Consequently, the mean and median of correlation coefficients were negative.

The correlations between the true effects of individual QTLs and the power to detect the QTLs ranged from 0.69 to 0.98 and were all highly significant (results not shown). In our model, the upper quartile QTLs (e.g., QTLs 1 to 20 for $l = 80$) had larger effects than the lower quartile QTLs (e.g., QTLs 61 to 80 for $l = 80$). With $\alpha = 0.01$, the average power across experiments for the upper quartile QTLs (0.53) was 3.5-fold that for the lower quartile QTLs (0.15). With $\alpha = 0.0001$, the average power across experiments for the upper quartile QTLs (0.37) was about tenfold that for the lower quartile QTLs (0.04).

Discussion

In silico mapping has four advantages over designed mapping experiments (Parisseaux and Bernardo 2004). First, in silico mapping exploits larger populations than designed mapping experiments. Second, the phenotypic data used in in silico mapping are obtained through more extensive testing under multiple, diverse environments. Third, the hybrids and inbreds tested typically represent wider genetic backgrounds. Fourth, the data used for in silico mapping are available without extra cost.

In plant breeding programs, the phenotypic data are highly unbalanced and the inbreds and hybrids have a pedigree structure. The original in silico mapping procedure proposed by Grupe et al. (2001) does not consider pedigree structures and becomes less powerful when data are unbalanced. In contrast, in silico mapping via a mixed-model approach accommodates unbalanced data, pedigree relationships, and different heterotic groups of parental inbreds by fitting relevant terms in the mixed model. Furthermore, the relative effects of the QTLs are measured by the regression coefficients of the significant markers, and the approximate positions of the QTL are indicated by the location of the significant markers. Interval mapping, within the method we described, can be conceivably used to estimate the location of a QTL within a marker interval. An interval mapping approach, however, might be computationally prohibitive.

The power of in silico mapping via a mixed-model approach in hybrid crops was affected by the heritability and genetic architecture of the trait (e.g., number and effects of QTLs), the resources available for mapping (e.g. number of markers and sample size), and the genetic structure of the particular breeding population (e.g., initial linkage disequilibrium among the founder inbreds). As with designed mapping experiments (Haley and Knott 1992; Beavis 1994), a large sample size, high marker density, high heritability, and small number of QTLs led to the highest power for in silico mapping via a mixed-model approach. A higher power to detect QTLs is expected for a trait with a high heritability and supposedly controlled by few QTLs (e.g., grain moisture and plant height in maize). A moderate power to detect QTLs is expected for a complex trait with a low to medium heritability and supposedly controlled by few QTLs (e.g., most disease resistance traits in maize). A low power to detect QTLs is expected for a complex trait with a low heritability and supposedly controlled by many QTLs (e.g., stalk lodging and root lodging in maize).

The strong association between the average power and FDR reflected a main challenge in dissecting complex traits. A compromise between the power to detect QTLs and the risk of having false positives can be reached by choosing an appropriate significance level for a given level of resources (e.g., sample size and marker density) and trait complexity (e.g., number of QTLs and heritability).

It would be useful to directly compare the power of a designed QTL mapping experiment (e.g., Haley and Knott 1992; Beavis 1994) versus the power of in silico mapping via a mixed-model approach. Previous simulations of designed QTL mapping experiments (Haley and Knott 1992; Beavis 1994), however, involved less realistic assumptions—i.e., fewer, unlinked QTL with equal effects underlying a complex trait. In contrast, we considered 80 randomly linked QTLs whose effects followed a geometric series. Assuming few, unlinked QTLs with equal effects optimizes the power to detect QTLs. A recent study by Bernardo (2004), however, considered a typical designed QTL mapping experiment with more realistic assumptions of genetic architecture of a complex trait, as in our study. Specifically, Bernardo (2004) considered a trait controlled by 30–100 QTLs, a heritability of 0.20 to 0.80, a sample size of 150 F_2-derived families, and 100 markers. For a significance level of $\alpha = 0.0001$ to 0.01, the average power to detect QTLs ranged from 0.01 to 0.17 (versus 0.01 to 0.59 in the current study). These results suggest that the power of in silico mapping via a mixed-model approach compares favorably with the power of designed QTL mapping experiments. Although this comparison involves different population sizes used in the two studies, they do reflect the population sizes available for the two approaches.

Previous studies using designed mapping populations have found that if only a small proportion of underlying QTLs were detected, the total variation accounted for by

these detected QTLs was overestimated (Beavis 1994; Utz and Melchinger 1994). Our results agreed with this general finding. While a lower power was achieved when 80 QTLs controlled a trait, the effects of significant markers were generally overestimated. Two factors presumably led to differences between the estimated effects and true effects of QTL in our study: (1) a QTL effect being confounded with the recombination distance between the QTL and a significant marker, and (2) a QTL effect being confounded with the effects of linked QTLs. The balance between these confounded factors determines the estimates of marker effects. As the number of QTLs increased, separating the effects of different linked QTLs became more difficult.

In gene discovery, one might want to identify primarily those QTLs with large effects. For traits controlled by few QTLs, the estimated effects of the significant markers were highly correlated with the effects of the underlying QTLs. This was not the case for traits controlled by many QTLs. The significant correlation between the true effects of QTLs and the power to detect QTLs, however, suggested that if an identified QTL is verified to be true, it is more likely to be one with large effect than with small effect, regardless of the number of QTLs underlying the trait.

Overall, our results indicated that gene discovery in hybrid crops can be initiated by in silico mapping via a mixed-model approach. It would be necessary, however, to find an acceptable compromise between the power to detect QTLs and the FDR. As with other QTL mapping methods, the results from in silico mapping should be followed by fine-mapping at the target regions, sequence analysis, and functional tests of gene effects (Glazier et al. 2002). In hybrid crops for which multiple heterotic groups exist, in silico mapping via a mixed-model approach can be applied to different heterotic patterns. Subsequently, the markers or the genomic regions that show a repeatable association with the trait of interest across different populations can be considered as the prime targets for further analysis (Parisseaux and Bernardo 2004). Cross validation by conducting in silico mapping in multiple heterotic patterns would result in a better control in overall FDR and provide increased confidence in conducting further investigation in putative QTL regions.

Acknowledgments This research was funded by the United States Department of Agriculture National Research Initiative Competitive Grants Program (Plant Genomics - Bioinformatics) and supported in part by the University of Minnesota Supercomputing Institute.

References

Beavis WD (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. Proc Corn Sorghum Ind Res Conf 49:250–266

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: lessons from comparative QTL approach to multiple testing. J R Statist Soc B 57:289–300

Bernardo R (1996) Best linear unbiased prediction of maize single-cross performance. Crop Sci 36:50–56

Bernardo R (1998) Predicting the performance of untested single crosses: trait and marker data. In: Lamkey KR, Staub JE (eds) Concepts and breeding of heterosis in crop plants. Crop Science Society of America, Madison, pp 117–127

Bernardo R (2004) What proportion of declared QTL in plants are false? Theor Appl Genet 109:419–424

Delacy IH, Cooper M (1990) Pattern analysis for the analysis of regional variety trials. In: Kang MS (ed) Genotype-by-environments interaction and plant breeding. Louisiana State University Agriculture Centre, Louisiana, pp 287–300

Doerge RW, Zeng Z-B, Weir BS (1994) Statistical issues in the search for genes affecting quantitative traits in populations. In: Analysis of molecular marker data. Joint Plant Breed Symp Ser. American Society for Horticulture Science, Crop Science Society of America, Madison, pp 15–26

Emik LO, Terrill CE (1949) Systematic procedures for calculating inbreeding coefficients. J Hered 40:51–55

Fernando RL, Nettleton D, Southey BR, Dekkers JCM, Rothschild MF, Soller M (2004) Controlling the proportion of false positives in multiple dependent tests. Genetics 166:611–619

Glazier AM, Nadeau JH, Aitman TJ (2002) Finding genes that underlie complex traits. Science 298:2345–2349

Grupe A, Germer S, Usuka J, Aud D, Belknap JK, Klein RF, Ahluwalia MK, Higuchi R, Peltz G (2001) In silico mapping of complex disease-related traits in mice. Science 292:1915–1918

Haley CS, Knott SA (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69:315–324

Hallauer AR (1990) Methods used in developing maize inbreds. Maydica 35:1–16

Henderson CR (1984) Application of linear models in animal breeding. University of Guelph, Ontario

Henderson CR (1985) Best linear unbiased prediction of nonadditive genetic merits in noninbred populations. J Anim Sci 60:111–117

Kearsey MJ, Farquhar AGL (1998) QTL analysis in plants; where are we now? Heredity 80:137–142

Kennedy BW, Quinton M, van Arendonk JAM (1992) Estimation of effects of single genes on quantitative trait. J Anim Sci 70:2000–2012

Lande R, Thompson R (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743–756

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Assoc, Sunderland

Mackay TFC (2001) The genetic architecture of quantitative traits. Annu Rev Genet 35:303–339

Panter DM, Allen FL (1995) Using best linear unbiased predictions to enhance breeding for yield in soybean: I. Choosing parents. Crop Sci 35:397–405

Parisseaux B, Bernardo R (2004) In silico mapping of quantitative trait loci in maize. Theor Appl Genet 109:508–514

Senior ML, Chin ECL, Lee M, Smith JSC, Stuber CW (1996) Simple sequence repeat markers developed from maize sequence found in the GENEBANK database: map construction. Crop Sci 36:1676–1683

Smith OS, Hoard K, Shaw F, Shaw R (1999) Prediction of single-cross performance. In: Coors JG, Pandey S (eds) The genetics and exploitation of heterosis in crops. Am Soc Agron Crop Sci, Crop Science Society of America, Madison, pp 277–285

Utz HF, Melchinger AE (1994) Comparison of different approaches to interval mapping of quantitative trait loci. In: van Ooijen JW, Janse J (eds) Biometrics in plant breeding: applications of molecular markers. CPRO-DLO, Wageningen, pp 195–204

Whittaker JC, Thompson R, Visscher PM (1996) On the mapping of QTL by regression of phenotypes on marker-type. Heredity 77:23–32