Article
Chinese Science Bulletin
SPECIAL TOPIC
Quantitative Genetics
July 2012 Vol.57 No.21: 2701–2710
doi: 10.1007/s11434-012-5239-3

Estimation of statistical power and false discovery rate of QTL mapping methods through computer simulation

LI HuiHui, ZHANG LuYan & WANG JianKang*

Institute of Crop Sciences, The National Key Facility for Crop Gene Resources and Genetic Improvement, and CIMMYT China, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Received February 20, 2012; accepted April 23, 2012

Many QTL mapping methods have been developed in the past two decades. Statistically, the best method should have a high detection power but a low false discovery rate (FDR). Power and FDR cannot be derived theoretically for most QTL mapping methods, but they can be properly evaluated using computer simulations. In this paper, we used four genetic models (two for independent loci and two for linked loci) to illustrate power and FDR estimation for interval mapping (IM) and inclusive composite interval mapping (ICIM). For each model, we simulated 1000 populations each of 200 doubled haploids. A support interval (SI) was first defined to indicate to which predefined QTL the significant QTL belonged. Power was calculated by counting the number of simulation runs with significant peaks higher than the logarithm of odds (LOD) threshold in the SI. Quantitative trait loci not identified in any SIs were viewed as false positives. The FDR is the rate at which QTLs are identified as significant when they are actually non-significant. Simulation results allowed us to estimate power and FDR of IM and ICIM for two independent and two linkage genetic models. Our estimates allowed us to readily compare the efficiencies of different statistical methods for QTL mapping, including the ability to separate linkage, under a wide range of genetic models. We used IM and ICIM as examples of how to estimate power and FDR, but the principles shown in this paper can be used for power analysis and comparison of any other QTL mapping methods, especially those based on interval tests.

false discovery rate (FDR), inclusive composite interval mapping (ICIM), interval mapping (IM), power simulation

Citation: Li H H, Zhang L Y, Wang J K. Estimation of statistical power and false discovery rate of QTL mapping methods through computer simulation. Chin Sci Bull, 2012, 57: 2701–2710, doi: 10.1007/s11434-012-5239-3

Quantitative trait locus (QTL) mapping has become routine for genetic studies of complex traits in plants, animals, and humans. Interval mapping (IM) has been viewed as a milestone in QTL mapping methods; IM uses a likelihood ratio test for the existence of QTLs based on maximum likelihood parameter estimation [1]. However, conventional IM cannot properly separate linked QTLs [2,3], and its power of detection is low, because of the lack of a background control. Composite interval mapping (CIM), proposed by Zeng [4], is one of the most commonly used methods, but CIM can result in biased mapping because it simultaneously estimates QTLs and background effects [5]. Inclusive composite interval mapping (ICIM) is a critical step forward that highlights the importance of model selection and interval testing in QTL linkage mapping, and is able to separate linked QTLs through a two-step mapping strategy [5–8]. In addition, ICIM can be readily extended to epistasis mapping and QTL mapping in multi-parental populations [9–11].

There are many other QTL mapping methods in addition to IM, CIM, and ICIM [5,12–14]. To appropriately evaluate the efficiency of these methods, we need to compare the true QTL positions and effects with their estimate and calculate the statistical power and false discovery rate (FDR). As with any statistical test, two types of error, Type I and Type II, can occur in QTL mapping [5,8,11]. (1) A Type I error is a false positive, in which a segregating QTL is detected when in fact it is not present. (2) A Type II error is a false negative, in which a QTL is not detected when it actually exists.

*Corresponding authors (email: wangjk@caas.net.cn; jkwang@cgiar.org)

Type I error can be controlled by choosing proper critical values, whereas Type II error is determined by the experiment design and the size of the QTL effect. Statistical power is the probability that the null hypothesis is rejected when it is indeed false. In other words, if β is the probability of Type II error, the power equals 1−β. In QTL mapping, power indicates the probability that a real QTL is detected and is therefore the most important indicator of the method's efficiency. The FDR offers an intuitive balance between false and true positives [15]. The FDR is defined by the rate at which significant features are truly null.

The asymptotical properties of test statistics used in most QTL mapping methods are hardly known, therefore the statistical power and the FDR cannot be estimated theoretically [8,16]. When using actual mapping populations, the true QTL positions and effects are usually unknown. Another problem with using real data to compare methods is that the parameter ranges are usually limited by experimental conditions, genetic architectures, and species. Collecting real data occupying the full range of parameter spaces is usually impossible.

Monte Carlo computer simulation is often used in physical and mathematical systems when computing an exact result with a deterministic algorithm is infeasible or impossible. It provides an efficient way to calculate the power and FDR of different QTL mapping methods to evaluate and compare their efficiencies. Our objective in this paper was to illustrate how detection power and FDR can be estimated in simulated populations under various genetic models, using two QTL mapping methods (i.e., IM and ICIM) as examples.

1 Materials and methods

1.1 Genetic model

In our simulation studies, we considered four QTLs with different additive effects affecting a quantitative trait and a genome of six chromosomes (Table 1). Each chromosome was 120 cM in length with 13 evenly-distributed markers.

Table 1 Characteristics of four genetic models for four QTLs used to estimate the power and false discovery rate of interval mapping and inclusive composite interval mapping

Model	Chromosome	Position (cM)	Additive effect	PVE (%)[a)]
Independent I				
Q1	1	35	0.316	5.0
Q2	2	35	0.447	10.0
Q3	3	35	0.548	15.0
Q4	4	35	0.633	20.0
Genetic variance	1.000			
Error variance	1.000			
Heritability	0.500			
Independent II				
Q1	1	35	0.316	5.0
Q2	2	35	–0.447	10.0
Q3	3	35	0.548	15.0
Q4	4	35	–0.633	20.0
Genetic variance	1.000			
Error variance	1.000			
Heritability	0.500			
Linkage I				
Q1	1	35	0.316	3.9
Q2	1	65	0.447	7.9
Q3	2	35	0.548	11.8
Q4	2	65	0.633	15.8
Genetic variance	1.535			
Error variance	1.000			
Heritability	0.606			
Linkage II				
Q1	1	35	0.316	6.8
Q2	1	65	–0.447	13.7
Q3	2	35	0.548	20.5
Q4	2	65	–0.633	27.3
Genetic variance	0.465			
Error variance	1.000			
Heritability	0.317			

a) PVE, phenotypic variance explained by each QTL.

Each marker interval was 10 cM. Two independent models (Independents I and II) and two linkage models (Linkages I and II) were defined. When there was no linkage, the four QTLs were located on the first four chromosomes at position 35 cM. Each independent model represented a different scenario of QTL effects. In Independent I, the four QTLs had additive effects 0.316, 0.447, 0.548, and 0.633, respectively, indicating that one parent had the four favorable alleles, and the other parent had the four non-favorable alleles. When error variance was 1.00, they explained 5%, 10%, 15%, and 20% of phenotypic variance, respectively. In Independent II, the four QTLs had additive effects 0.316, –0.447, 0.548, and –0.633, respectively, indicating that one parent had the favorable alleles for two loci, and the other parent had the favorable alleles for the other two loci. They explained the same amount of phenotypic variance as in Independent I. The total genetic variance is sum of individual QTL variances, i.e., $V_G = \sum_{i=1}^{4} a_i^2$, where a_i is the additive effect of the ith QTL. When the error variance was fixed at 1 in the independent models, genetic variance was 1 and heritability was 0.5.

For the two linkage models, Q1 and Q2 were located at 35 and 65 cM on chromosome 1, and Q3 and Q4 were located at 35 and 65 cM on chromosome 2 (Table 1). In Linkage I, the QTLs had the same effects as in Independent I, representing coupling linkage. In Linkage II, the QTLs had the same effects as in Independent II, representing repulsion linkage. If r is the recombination frequency between two linked QTLs having a genetic distance of d cM, $r = \frac{1}{2}(1 - e^{-d/50})$ in Haldane's mapping function. The total genetic variance is

$$V_G = \sum_{i=1}^{4} a_i^2 + 2(1 - r_{12})a_1 a_2 + 2(1 - r_{34})a_3 a_4,$$

where a_i is the additive effect of the ith QTL, r_{12} is the recombination frequency between Q1 and Q2, and r_{34} is the recombination frequency between Q3 and Q4. When the error variance was fixed at 1.00 in the linkage models, genetic variance was 1.535 for Linkage I and 0.465 for Linkage II, resulting in heritabilities of 0.606 and 0.317, respectively.

1.2 Simulation of genetic populations and their QTL mapping

One thousand doubled haploid populations, each of size 200, were simulated for each model, and QTL mapping was completed using the QTL IciMapping software package, which is freely available from http://www.isbreeding.net. IM and ICIM were used on each simulated population. In the first step of ICIM, marker selection was conducted through stepwise regression by considering all marker information simultaneously. The two probabilities for entering and removing variables in the first step were set at 0.001 and 0.002, respectively. Phenotypic values were then adjusted for all markers retained in the regression equation, except for the two markers flanking the current mapping interval. In the second step, the adjusted phenotypic values are used in one-dimensional scanning. The logarithm of odds (LOD) threshold was set at 2.5 for IM and ICIM.

1.3 Estimation of statistical power and false discovery rate

Both IM and ICIM are based on an interval test, which is not a point estimation procedure [5]. The LOD score is the test statistic used in IM and ICIM and is calculated for each position in the genome through one-dimensional scanning. Example LOD profiles from IM and ICIM in five simulated populations are shown in Figures 1 and 2 for the independent and linkage models, respectively. QTLs are identified at the peaks on the LOD profiles. One QTL was unlikely to be located exactly at the predefined position in each simulated population. The multiple non-independent and non-point tests complicate the calculation of power, even though simulation. In particular, when QTLs are closely linked, determining which putative QTL the LOD peak belongs to is difficult.

A support interval (SI) has to be used to indicate to which predefined QTL the significant QTL belonged in the simulations. Each predefined QTL was assigned to a predefined interval length centered at the true QTL location, and then the power was determined by counting the number of simulation runs with significant peaks higher than the LOD threshold along the LOD profile in this interval. QTLs not identified in the corresponding SI were viewed as false positives. Five simulated populations from Independent I were used to illustrate how the detection power and FDR were calculated through computer simulation.

When an LOD threshold value of 2.5 was used, IM identified a total of 16 and ICIM identified a total of 20 QTLs in five simulation runs (Table 2). Using the first run as an example, IM did not detect Q1, so its power was 0. One QTL was identified at 25 cM on chromosome 2, but this QTL was not located within the 10 cM SI of the actual position of Q2, i.e., from 30 to 40 cM, and was treated as a false positive. IM identified one QTL at 35 cM, within the 10 cM SI, on chromosome 3, so the detection power for Q3 was 1. The detection power for Q4 was 1 as well, because one QTL was identified on chromosome 4 at 40 cM, within the 10 cM SI of Q4. Across the five runs, the QTLs 1–4 were detected 1, 3, 4, and 5 times, respectively, with SI=10 cM (Table 3). Therefore, the detection powers for the four QTL were 20%, 60%, 80%, and 100%, respectively. In total, there were three false QTLs and 16 positive ones. Therefore, FDR equaled 3/16=19% for IM.

When ICIM was used, the four QTLs were detected 4, 4, 5, and 4 times, respectively (Table 3). Therefore, detection

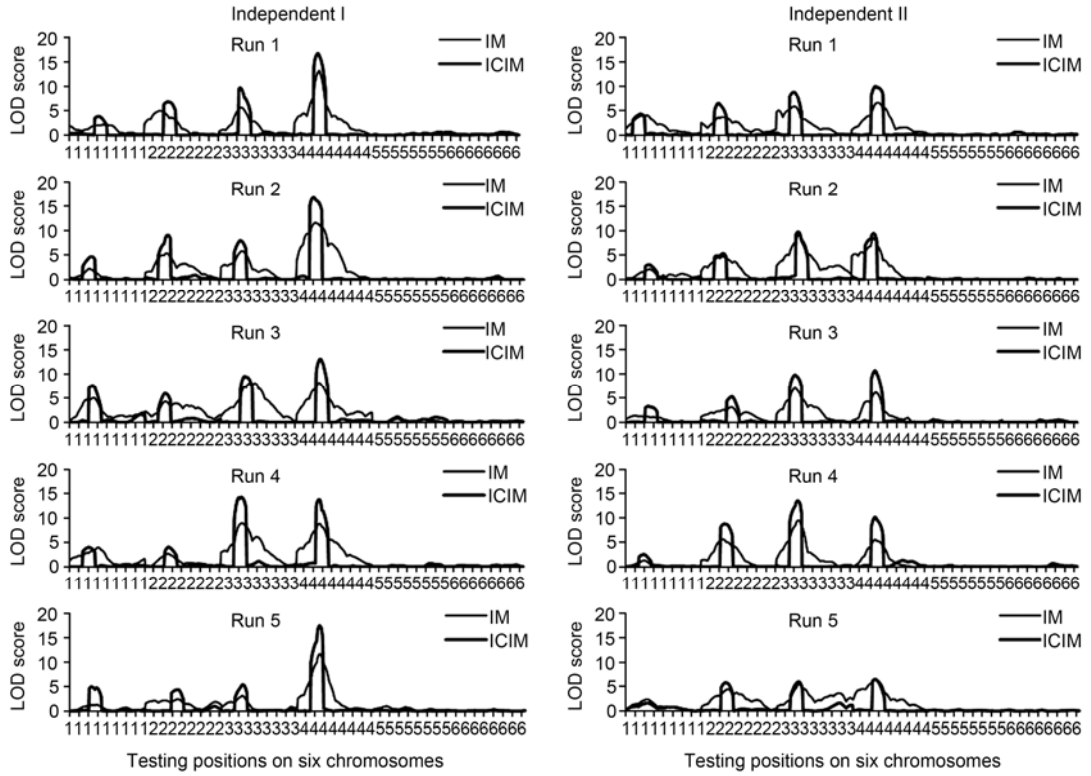

Figure 1 Example logarithm of odds (LOD) profiles of interval mapping (IM) and inclusive composite interval mapping (ICIM) on six chromosomes in five simulated populations under two models of independence.

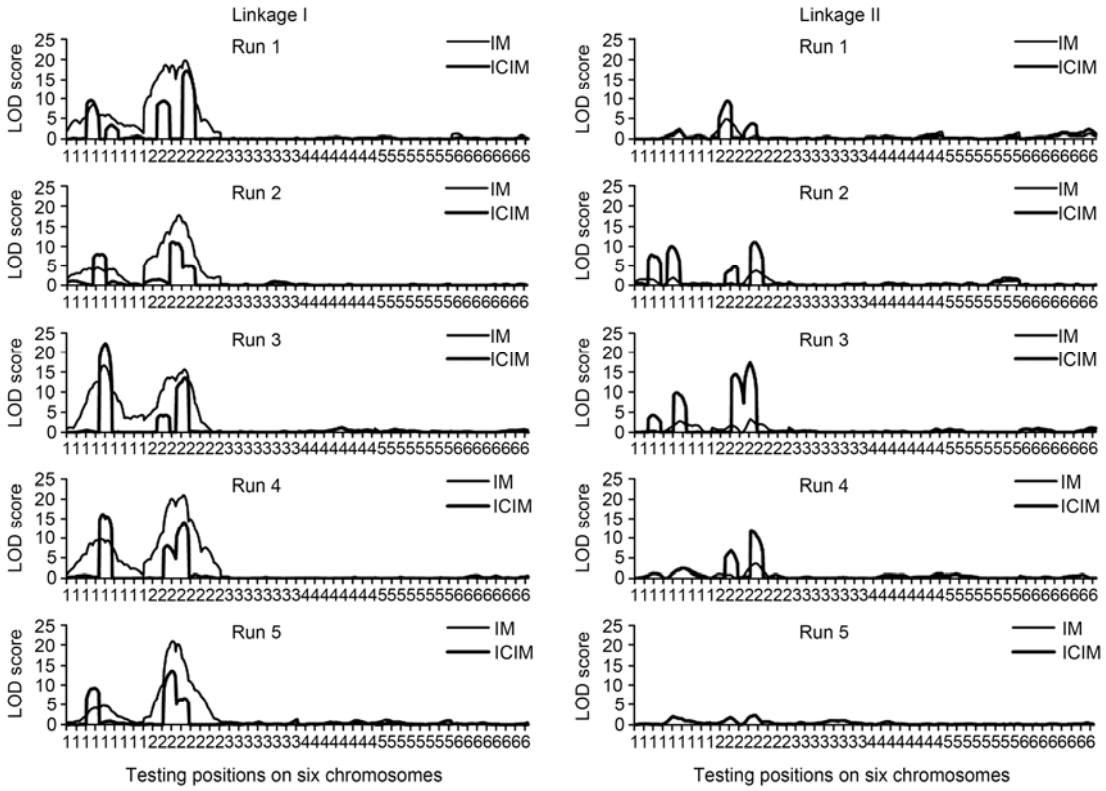

Figure 2 Example logarithm of odds (LOD) profiles of interval mapping (IM) and inclusive composite interval mapping (ICIM) on six chromosomes in five simulated populations under two linkage models.

Table 2 QTL mapping and power counts for five populations simulated using the Independent I model

Method	Simulation	QTL identified in simulated population					Power counting	
		Chromosome	Position	LOD[a)]	PVE (%)[b)]	Effect	SI=10 cM	SI=20 cM
IM	1	2	25	4.97	11.44	0.503	False	Q2
		3	35	5.61	13.35	0.541	Q3	Q3
		4	40	13.21	26.22	0.761	Q4	Q4
	2	2	34	5.36	13.01	0.509	Q2	Q2
		3	34	5.82	13.72	0.521	Q3	Q3
		4	30	11.59	23.43	0.682	Q4	Q4
	3	1	39	5.05	11.22	0.508	Q1	Q1
		2	32	4.30	10.09	0.482	Q2	Q2
		3	54	8.03	18.42	0.651	False	False
		4	36	8.06	18.55	0.653	Q4	Q4
	4	1	45	3.97	10.21	0.420	False	Q1
		2	36	2.69	6.81	0.343	Q2	Q2
		3	34	8.92	19.66	0.583	Q3	Q3
		4	36	8.79	20.15	0.591	Q4	Q4
	5	3	33	3.08	8.16	0.389	Q3	Q3
		4	35	11.71	26.65	0.701	Q4	Q4
ICIM	1	1	47	3.80	5.06	0.335	False	False
		2	38	6.79	9.11	0.448	Q2	Q2
		3	33	9.70	13.81	0.551	Q3	Q3
		4	38	16.72	25.50	0.753	Q4	Q4
	2	1	35	4.65	6.26	0.352	Q1	Q1
		2	36	9.07	12.56	0.500	Q2	Q2
		3	31	7.93	10.41	0.454	Q3	Q3
		4	27	16.77	24.93	0.703	False	Q4
	3	1	36	7.52	10.23	0.486	Q1	Q1
		2	32	6.00	8.10	0.432	Q2	Q2
		3	38	9.52	13.63	0.560	Q3	Q3
		4	38	13.05	19.18	0.664	Q4	Q4
	4	1	30	3.99	5.13	0.298	Q1	Q1
		2	37	4.04	5.89	0.319	Q2	Q2
		3	33	14.21	21.68	0.613	Q3	Q3
		4	36	13.73	21.23	0.607	Q4	Q4
	5	1	35	4.91	8.04	0.384	Q1	Q1
		2	51	4.35	6.87	0.356	False	False
		3	34	5.35	9.45	0.419	Q3	Q3
		4	35	17.46	31.65	0.764	Q4	Q4

a) Logarithm of odds; b) phenotypic variance explained by each identified QTL.

powers of ICIM were 80%, 80%, 100%, and 80%, respectively, for the four QTLs. There were a total of three false QTLs and 20 positives. Therefore, FDR of ICIM equaled 3/20=15%.

1.4 Position and effect estimates

Precisely locating a QTL's position and estimating its effect are both important in QTL mapping. For a significant QTL detected by test statistics that exceed a predetermined threshold, the estimate of its genetic effects must be assessed in terms of accuracy and precision. Precision is another criterion by which different statistical methods for QTL detection can be compared, in addition to high power and low FDR. In simulation studies, if the estimates of effect are calculated only from significant QTLs, the effects are usually over-estimated, because simulated QTLs with LOD scores below the threshold may have smaller effects

Table 3 QTL detection frequency and power of interval mapping (IM) and inclusive composite interval mapping (ICIM) estimated from five populations simulated using the Independent I

Method	QTL	Detection times in five runs		Power (%)	
		SI=10 cM	SI=20 cM	SI=10 cM	SI=20 cM
IM	Q1	1	2	20.0	40.0
	Q2	3	4	60.0	80.0
	Q3	4	4	80.0	80.0
	Q4	5	5	100.0	100.0
	False (FDR[a)])	3	1	19.0	6.0
ICIM	Q1	4	4	80.0	80.0
	Q2	4	4	80.0	80.0
	Q3	5	5	100.0	100.0
	Q4	4	5	80.0	100.0
	False (FDR)	3	2	15.0	10.0

a) False discovery rate.

but will not be counted. Unbiased estimates can be achieved if all runs with peaks are counted, whether or not the peaks are above the LOD threshold.

2 Results

2.1 Power and false discovery rate of IM at a predefined support interval

The QTL detection power of IM was calculated for SI=10 and 20 cM (Table 4). Similar power and FDR were observed for both independent models, indicating that scattering the distribution of favorable and non-favorable alleles between parents had no effect on QTL detection when the QTLs were genetically independent. For the four QTLs considered, those with large genetic effects also had high detection power. For example, Q1 explained only 5% of phenotypic variance, and the power of IM to locate it in the 10 cM SI was 25.8%. In contrast, Q4 explained 20% of phenotypic variance, and the power of IM to locate it in the 10 cM SI was 85.4%.

Linkage significantly reduced detection power and increased FDR (Table 4). Linkage in repulsion (Linkage model II; Table 1) represented the worst case of QTL mapping. Both the LOD score and power were the lowest, and the standard errors of both QTL position and effects were the highest among the four models (Table 4). That larger SI will result in higher power and lower FDR (Tables 3 and 4) is understandable. In this sense, the SI length must be specified when estimating QTL detection power; otherwise, positive and false QTLs cannot be distinguished.

2.2 Power and false discovery rate of ICIM at a predefined support interval

The QTL detection power of ICIM was calculated for SI=10 and 20 cM (Table 5). Almost identical power and FDR were observed for both independent models. QTLs with large genetic effects also had high detection power. For example, Q1 explained 5% of phenotypic variance, and the power of ICIM to locate it in the 10 cM SI was 49.5%. In contrast, Q4 explained 20% of phenotypic variance, and the power of ICIM to locate it in the 10 cM SI was 89.0%.

Linkage reduced ICIM's detection power and increased FDR (Table 5). Linkage in repulsion (Linkage model II) resulted in the lowest LOD score and power and the highest standard error of QTL position and effects among the four models. For Linkage II, IM had powers of 0.3%, 25.35%, 6.85%, and 40.6% to detect the respective QTLs and FDR= 38.8%. In contrast, ICIM had powers 11.65%, 33.05%, 56.25%, and 60.9%, respectively, and FDR=23.8%. In addition, ICIM resulted in smaller standard errors for QTL position and effect.

2.3 Detection power of IM and ICIM on marker intervals

The detection power could be calculated for each interval defined by markers. Power thus determined allows monitoring of QTL locations if not on the putative intervals. For the two independent models, IM and ICIM located QTLs at their flanking intervals in most times. However, there were chances that QTLs were located in other neighboring intervals (Figure 3). For the two linkage models, IM located the two QTLs on chromosome 1 at the middle of the two positions in most times, and so as the two QTL on chromosome 2. ICIM located the two linked QTL on chromosomes 1 and 2 in their flanking intervals in most times, indicating its ability in separating linkage. Generally speaking, linkage complicates the QTL mapping procedure. In other words, much larger population is needed to dissect close linkage.

3 Discussion

Several QTL mapping methods have been developed in the

Table 4 Power (%) and false discovery rate (FDR, %) of interval mapping estimated from 1000 simulated populations under the four genetic models defined in Table 1

	SI=10 cM							SI=20 cM						
	Power (%)	Position	SE[a)]	LOD[b)]	SE[a)]	Effect	SE[a)]	Power (%)	Position	SE[a)]	LOD[b)]	SE[a)]	Effect	SE[a)]
	Independent I							Independent I						
Q1	25.8	35.182	3.461	3.849	1.003	0.421	0.056	33.7	35.315	4.977	3.801	1.068	0.422	0.062
Q2	62.6	34.861	3.014	5.084	1.692	0.480	0.082	76.9	34.802	4.671	5.066	1.591	0.481	0.080
Q3	77.7	35.006	2.669	7.013	2.178	0.560	0.092	91.6	35.073	4.119	6.981	2.228	0.558	0.096
Q4	85.4	35.067	2.464	9.205	2.584	0.635	0.095	97.0	35.025	3.662	9.272	2.558	0.636	0.094
FDR (%)	32.4							18.6						
	Independent II							Independent II						
Q1	27.3	34.835	3.414	3.865	1.151	0.424	0.062	35.9	34.978	4.818	3.792	1.031	0.422	0.060
Q2	64.2	35.062	3.035	5.006	1.607	–0.478	0.078	78.5	34.967	4.625	5.059	1.698	–0.481	0.083
Q3	78.6	34.956	2.706	6.963	2.143	0.558	0.090	93.5	34.645	4.196	7.020	2.067	0.560	0.087
Q4	84.6	34.865	2.481	9.374	2.441	–0.640	0.089	97.0	34.967	3.605	9.204	2.504	–0.635	0.093
FDR (%)	31.1							18.1						
	Linkage I							Linkage I						
Q1	24.1	35.448	2.757	6.944	2.092	0.625	0.099	41.5	38.096	5.251	7.356	2.227	0.646	0.104
Q2	49.0	64.790	2.549	7.859	2.278	0.665	0.104	72.2	62.731	4.669	7.906	2.402	0.667	0.110
Q3	40.2	35.759	1.648	17.017	3.105	0.937	0.090	71.5	39.469	4.566	17.680	3.340	0.948	0.093
Q4	56.5	64.165	1.624	18.682	3.359	0.971	0.093	87.3	61.190	4.282	19.073	3.337	0.977	0.093
FDR (%)	53.1							19.2						
	Linkage II							Linkage II						
Q1	0.3	33.333	4.714	2.965	0.431	0.313	0.013	0.4	31.750	4.918	3.242	0.467	0.345	0.033
Q2	25.3	66.534	3.220	3.667	1.029	–0.354	0.050	31.1	67.987	3.985	3.763	1.063	–0.355	0.049
Q3	6.8	31.691	2.608	3.176	0.553	0.334	0.031	9.9	29.748	2.706	3.394	0.789	0.339	0.040
Q4	40.6	67.746	2.680	4.169	1.302	–0.373	0.061	52.6	69.029	3.402	4.208	1.258	–0.377	0.059
FDR (%)	38.9							22.0						

a) Standard error of estimated position, LOD, or effect; b) logarithm of odds.

Table 5 Power (%) and false discovery rate (FDR, %) of inclusive composite interval mapping (ICIM) estimated from 1000 simulated populations under the four genetic models defined in Table 1

	SI=10 cM							SI=20 cM						
	Power (%)	Position	SE[a)]	LOD[b)]	SE[a)]	Effect	SE[a)]	Power (%)	Position	SE[a)]	LOD[b)]	SE[a)]	Effect	SE[a)]
	Independent I							Independent I						
Q1	49.5	34.867	3.184	4.667	1.656	0.354	0.062	60.5	35.017	4.714	4.617	1.475	0.354	0.057
Q2	73.9	34.874	2.769	7.156	2.295	0.450	0.077	92.2	34.904	4.250	7.072	2.304	0.448	0.078
Q3	82.8	34.958	2.521	10.161	2.710	0.548	0.078	95.9	34.776	3.781	10.109	2.830	0.546	0.082
Q4	89.0	35.160	2.278	13.087	3.229	0.632	0.083	98.6	35.046	3.239	12.864	3.212	0.626	0.085
FDR (%)	22.6							8.3						
	Independent II							Independent II						
Q1	49.2	34.831	3.204	4.589	1.640	0.352	0.063	66.5	35.138	4.665	4.562	1.502	0.353	0.059
Q2	76.1	35.030	2.861	7.142	2.328	–0.448	0.076	92.7	35.121	4.141	7.109	2.281	–0.450	0.075
Q3	85.6	35.051	2.484	10.193	2.755	0.548	0.081	96.8	34.810	3.750	10.051	2.779	0.544	0.080
Q4	90.0	34.939	2.325	13.203	3.221	–0.634	0.082	98.9	35.073	3.089	12.835	3.247	–0.625	0.085
FDR (%)	21.3							7.4						
	Linkage I							Linkage I						
Q1	26.9	35.353	3.051	7.335	3.466	0.449	0.118	41.7	36.1175	5.2289	7.6173	3.7349	0.4606	0.128
Q2	55.5	64.872	2.701	10.519	4.184	0.558	0.133	78.3	63.5019	4.6213	10.5503	3.8173	0.5598	0.1227
Q3	77.0	34.952	2.618	10.560	3.890	0.559	0.113	89.7	34.9666	3.9823	10.6777	4.2	0.5629	0.1237
Q4	84.2	64.828	2.533	13.668	4.761	0.649	0.130	94.5	64.9164	3.5018	13.8805	4.8132	0.6562	0.1335
FDR (%)	26.4							8.2						
	Linkage II							Linkage II						
Q1	11.6	34.216	3.615	5.100	1.915	0.370	0.066	12.6	33.675	4.548	5.462	2.003	0.387	0.072
Q2	33.0	66.179	3.053	5.872	3.188	–0.402	0.108	40.0	67.478	4.029	5.949	3.156	–0.404	0.107
Q3	56.2	34.383	2.894	8.332	3.381	0.492	0.104	62.5	33.562	3.593	8.230	3.311	0.492	0.101
Q4	60.9	65.984	2.429	11.413	4.131	–0.591	0.114	69.8	66.781	3.573	10.934	3.850	–0.577	0.112
FDR (%)	23.8							11.8						

a) Standard error of estimated position, LOD, or effect; b) logarithm of odds.

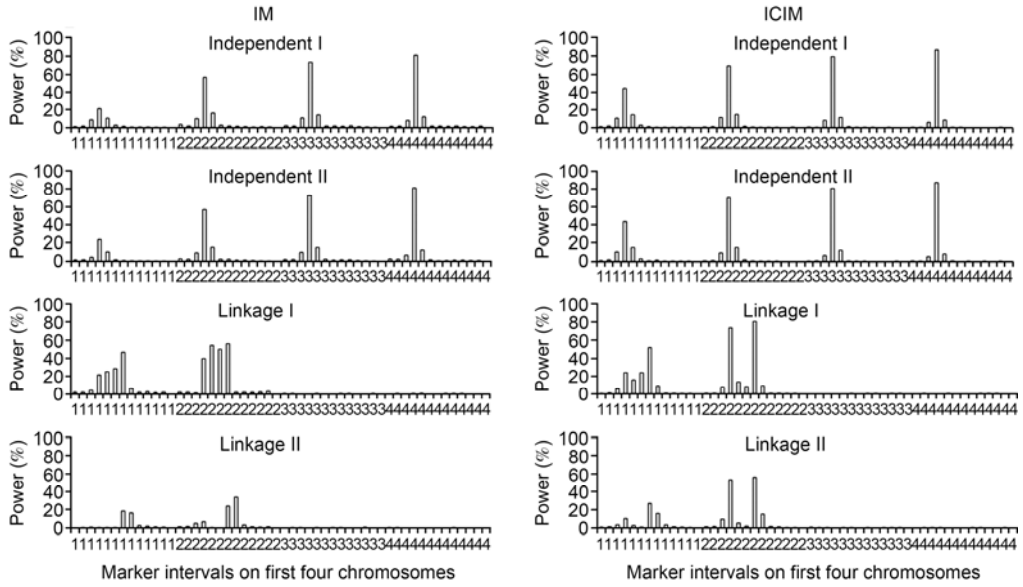

Figure 3 Power of interval mapping (IM) and inclusive composite interval mapping (ICIM) on each marker interval on the first four chromosomes. Power was estimated from 1000 populations simulated using each of two independent and two linkage genetic models.

past two decades, and the most efficient and powerful method should be applied in genetic studies. Statistically, the best QTL mapping method should meet two criteria: high detection power and low FDR. The asymptotic properties of most test statistics used in QTL mapping are barely known theoretically. Computer simulation may be the only way to effectively compare different methods. We used IM and ICIM as examples, but the principles shown in this paper are applicable to any other QTL mapping methods, especially those based on interval tests.

In statistics, Type I error can be controlled by choosing the proper critical values, whereas Type II error is determined by the experimental design and the size of the QTL effect. Statistical power is the probability that the null hypothesis is rejected when it is indeed false. Before conducting a potentially costly experiment, an investigator would like to be certain that the design ensures sufficiently high power given the study's objectives. Statistical power depends not only on sample size and the actual values of the unknown distribution parameters being estimated, but also on the assumed level of the significance threshold. In QTL mapping, power tells the likelihood that a real QTL was detected and is therefore the most important indicator of a method's efficiency.

The Type I error rate and FDR are often mistakenly equated, but their difference is actually very important. Given a rule for calling features significant, the Type I error rate is the probability that truly null features are called significant. FDR is the rate that statistically significant features are truly null. For example, a Type I error rate of 5% means that, on average, 5% of the truly null features in the study will be called significant. A FDR of 5% means that, among all the features identified as significant, 5% on average are truly null. In addition, a much higher FDR can be tolerated than a *P*-value. For instance, a *P*-value of 0.30 is statistically unacceptable in any situation; but an FDR as high as 0.50 or even higher could be quite meaningful.

Quantitative traits are normally polygenic. Therefore, a number of QTLs must be considered simultaneously in power simulation. To tell which identified QTL belongs to which predefined QTL, and which are true or false, an SI must be predefined. We used both SI=10 and 20 cM in this study, but other SIs can be used as long as consistent standards are used to compare different methods. When more than one QTL are identified in a predefined SI, the one with the highest LOD score is chosen to ensure that the estimated power will not exceed 100%. Any QTL that falls beyond all predefined SI is considered false. In this sense, more than one false QTL may exist in a population, and the total number of false QTLs may exceed the number of runs, especially when a narrow SI is used. Defined by the proportion of false QTLs to all positives (true+false QTLs), FDR avoids a false rate greater than 100%, and has been widely used in QTL simulation studies.

High power and low FDR are the two most important statistical requirements for efficient QTL mapping methods. However, precisely estimating QTL position and effect is also important. These factors must also be considered when conducting simulation studies to provide perspective on the mapping methods. Extensive simulation studies previously completed have proved the great advantage of ICIM in improving QTL detection power, separating linked QTLs, mapping interacting QTLs, and assessing multi-parental populations and QTL-by-environment interaction [5–10]. The ICIM method has been implemented in the freely-available public software QTL IciMapping (available from

www.isbreeding.net). In addition to QTL mapping, the software can build high-density linkage genetic maps as well. The simulation functionality in QTL IciMapping provides a useful tool to compare different mapping methods using a wide range of genetic models.

This work was supported by the NationalBasic Research Program of China (*2011CB100100*) *and the National Natural Science Foundation of China* (*31000540*).

1 Lander E S, Botstein D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics, 1989, 121: 185–199
2 Haley C S, Knott S A. A simple regression method for mapping quantitative loci in line crosses using flanking markers. Heredity, 1992, 69: 315–324
3 Martinez O, Curnow R N. Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor Appl Genet, 1992, 85: 480–488
4 Zeng Z B. Precision mapping of quantitative trait loci. Genetics, 1994, 136: 1457–1468
5 Li H, Ye G, Wang J. A modified algorithm for the improvement of composite interval mapping. Genetics, 2007, 175: 361–374
6 Zhang L, Li H, Li Z, et al. Interactions between markers can be caused by the dominance effect of QTL. Genetics, 2008, 180: 1177–1190
7 Wang J. Inclusive composite interval mapping of quantitative trait genes. Acta Agron Sin, 2009, 35: 239–245
8 Li H, Hearne S, Bänziger M, et al. Statistical properties of QTL linkage mapping in biparental genetic populations. Heredity, 2010, 105: 257–267
9 Li H, Ribaut J M, Li Z, et al. Inclusive composite interval mapping (ICIM) for digenic epistasis of quantitative traits in biparental populations. Theor Appl Genet, 2008, 116: 243–260
10 Li H, Bradbury P, Ersoz E, et al. Joint QTL linkage mapping for multiple-cross mating design sharing one common parent. PLoS ONE, 2011, 6: e17573
11 Li H, Zhang L, Wang J. Analysis and answers to frequently asked questions in quantitative trait locus mapping (in Chinese). Acta Agron Sin, 2010, 36: 918–931
12 Wright A J, Mowers R P. Multiple regression for molecular-marker, quantitative trait data from large F_2 populations. Theor Appl Genet, 1994, 89: 305–331
13 Whittaker J C, Thompson R, Visscher P M. On the mapping of QTL by regression of phenotype on marker-type. Heredity, 1996, 77: 23–32
14 Doerge R W, Zeng Z B, Weir B S. Statistical issues in the search for genes affecting quantitative traits in experimental populations. Statist Sci, 1997, 12: 195–219
15 Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Series B, 1995, 57: 289–300
16 Mayer M, Liu Y, Freyer G. A simulation study on the accuracy of position and effect estimates of linked QTL and their asymptotic standard deviations using multiple interval mapping in an F_2 scheme. Genet Sel Evol, 2004, 36: 455–479