Theor Appl Genet (2004) 109: 419–424
DOI 10.1007/s00122-004-1639-3

ORIGINAL PAPER

R. Bernardo

What proportion of declared QTL in plants are false?

Received: 21 August 2003 / Accepted: 23 February 2004 / Published online: 14 April 2004

Abstract The false discovery rate (FDR) is the probability that a quantitative trait locus (QTL) is false, given that a QTL has been declared. A misconception in QTL mapping is that the FDR is equal to the comparison-wise significance level, α_C. The objective of this simulation study was to determine the FDR in an F_2 mapping population, given different numbers of QTL, population sizes, and trait heritabilities. Markers linked to QTL were detected by multiple regression of phenotype on marker genotype. Phenotypic selection and marker-based recurrent selection were compared. The FDR increased as α_C increased. Notably, the FDR was often 10–30 times higher than the α_C level used. Regardless of the number of QTL, heritability, or size of the genome, the FDR was $\leq$0.01 when α_C was 0.0001. The FDR increased to 0.82 when α_C was 0.05, heritability was low, and only one QTL controlled the trait. An α_C of 0.05 led to a low FDR when many QTL (30 or 100) controlled the trait, but this lower FDR was accompanied by a diminished power to detect QTL. Larger mapping populations led to both lower a FDR and increased power. Relaxed significance levels of α_C=0.1 or 0.2 led to the largest responses to marker-based recurrent selection, despite the high FDR. To prevent false QTL from confusing the literature and databases, a detected QTL should, in general, be reported as a QTL only if it was identified at a stringent significance level, e.g., $\alpha_C \cong 0.0001$.

Communicated by H.C. Becker

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

Introduction

The development of molecular markers has led to more than 700 published studies to map QTL (Coors 2001). In plants, more than 900 QTL have been reported across QTL studies that involved relatively large (i.e., at least 250 families) mapping populations (Bernardo 2002), and the total number of reported QTL would undoubtedly be much larger if studies with smaller mapping populations are also considered. The total number of reported QTL will increase as resources continue to be devoted to finding markers associated with quantitative traits.

Declaring the presence of a QTL always carries some risk that such declaration is false. Suppose that out of 64 independent markers, 60 are unlinked to QTL in a mapping population. Out of these 60 markers, three are incorrectly declared to be linked to a QTL (i.e., false positives, Fig. 1) and 57 are correctly declared to be unlinked to QTL (i.e., true negatives, Fig. 1). The comparison-wise significance level or type I error rate, denoted by α_C, is equal to (number of false positives)/[(number of false positives) + (number of true negatives)]. In the example in Fig. 1, α_C is equal to 3/(3+57)=0.05. Studies to map QTL have differed in the significance levels used. Some investigators have used stringent significance levels of $\alpha_C \cong 0.0001$, as suggested by a permutation test to control the experiment-wise error rate (Churchill and Doerge 1994), whereas other investigators (Openshaw and Frascaroli 1997) have used a relaxed significance level of α_C=0.1.

Regardless of the significance level used, a misconception is that α_C is equal to the proportion of false positives among all declared marker-QTL linkages. In other words, if 20 QTL have been declared at a significance level of α_C=0.05, a misconception is that only 20×0.05=1 out of the 20 declared QTL is false. The false discovery rate (FDR) is defined as the probability that a QTL is false, given that a QTL has been declared (Benjamini and Hochberg 1995); it is equal to (number of false positives)/[(number of false positives) + (number of true positives)]. In the example in Fig. 1, the FDR is equal to 3/(3+1)=0.75

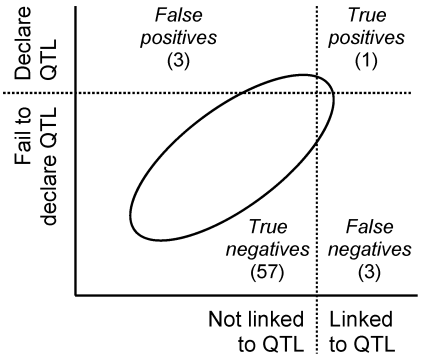

Fig. 1 Outcomes of a significance test for detecting QTL

rather than 0.05. The FDR can therefore be much greater than α_C (Fernando 2002).

A high FDR can result in false leads and wasted resources in characterizing and exploiting genes for quantitative traits, as well as confuse the QTL literature and databases. Knowledge of the magnitude of the FDR would be helpful for designing QTL mapping experiments and for properly interpreting their results. Methods have been proposed to specify a level of α_C that controls the FDR (Benjamini and Hochberg 1995; Weller et al. 1998). On the other hand, little is known about the FDR that would arise from a typical QTL mapping experiment in plants, where the mapping population usually comprises a cross between two inbreds. The objective of this simulation study was to determine the FDR in an F_2 mapping population given different numbers of QTL, population sizes, and trait heritabilities. I discuss the implications of the FDR with regard to gene discovery, QTL introgression, and marker-based recurrent selection.

Materials and methods

Mapping population

I wrote a Fortran program to simulate QTL mapping and marker-based selection in a maize (*Zea mays* L.) F_2 population. The simulation experiment was repeated 2,000 times. Each repeat differed in the location of QTL and markers and in the genotypes, genotypic values, and phenotypic values of the F_2 individuals.

Two parental inbreds were crossed to form an F_1, and the F_1 was selfed to form an F_2 population. Sizes of mapping populations in plants have typically ranged from 100 to 250 (Beavis 1994; Lynch and Walsh 1998), and in this study a population size of 150 was considered. The F_2 population was segregating at 100 codominant marker loci and at l =1, 5, 10, 30, or 100 QTL. Linkage among the QTL was generated by randomly locating the l QTL on 10 chromosomes. The first parent had the favorable allele at an odd-numbered QTL, and the less-favorable allele at even-numbered QTL. The sizes of the chromosomes (ranging from 128 to 241 centiMorgans, cM) and of the entire genome (1,749 cM) corresponded to those in a published maize linkage map (Senior et al. 1996). The genome was divided into 100 bins of approximately 1,749/100=17 cM. A marker was assumed to be present within ±5 cM of the midpoint of each bin.

Genotypic values were generated according to metabolic control theory (Kacser and Burns 1981). This method was chosen because it is biologically meaningful and it leads to a distribution of gene effects that is consistent with theoretical arguments (Thompson 1975; Lande and Thompson 1990) and with experimental data (Kearsey and Farquhar 1998; Bost et al. 1999; Bernardo 2002), i.e., few genes with large effects and many genes with small effects. A linear metabolic pathway with l enzymes, corresponding to the l QTL, was assumed. The first QTL coded for the enzyme that converted substrate 1 into substrate 2; the second QTL coded for the enzyme that converted substrate 2 into substrate 3; and the lth QTL coded for the enzyme that converted substrate l into the final product.

At the ith QTL, the enzyme activity for the allele from the first parental inbred was denoted by E_{i1}, whereas the enzyme activity for the allele from the second parental inbred was denoted by E_{i2}. At the odd-numbered QTL, these enzyme activities were $E_{i1}=m_i+b_i$ and $E_{i2}=m_i-b_i$, where m_i was the midparent enzyme activity and b_i was half the difference between enzyme activities of the homozygotes. At the even-numbered QTL, the enzyme activities were $E_{i1}=m_i-b_i$ and $E_{i2}=m_i+b_i$. The m_i values followed a geometric series. Specifically, m_i was equal to $1-a^{i/2}$, where the value of a for an effective number of loci equal to l was calculated using Eq. 11 of Lande and Thompson (1990). The value of b_i was calculated assuming a coefficient of variation of 0.15 relative to m_i. Specifically, b_i was calculated as $(0.15)m_i\sqrt{2}$ (Bost et al. 1999). The enzyme activity of the heterozygote was equal to m_i, i.e., no dominance for enzyme activity.

The metabolic flux was considered as the quantitative trait being studied. For individual k, the genotypic value (G_k) was equal to the flux (J_k, Kacser and Burns 1981):

$$G_k = J_k = c/\sum (1/E_{ik})$$

where E_{ik} was the activity of the ith enzyme in the kth F_2 individual, and c was a constant which did not affect the relative values of G_k. To reduce rounding errors, c was assumed equal to the square of the number of QTL.

Even though the QTL in metabolic control analysis exhibit physiological epistasis, the amount of non-additive genetic variance for the metabolic flux is often small, particularly when many QTL are involved (Keightley 1989). Negligible effects of dominance and epistasis in a metabolic flux indicate that although genotypic values in this study were simulated for F_2 individuals, such genotypic values also apply to F_2-derived families (e.g., $F_{2:3}$, $F_{2:4}$, $F_{2:5}$, etc.). Furthermore, testcross means of families in a cross-pollinated crop behave in an additive manner (Bernardo 2002) even if dominance is present. The results from this study therefore also apply to mapping QTL for testcross performance in a cross-pollinated crop.

The broad-sense heritability was $H=V_G/(V_G+V_E)$, where V_E was the nongenetic variance. The values of H were 0.20, 0.50, or 0.80. The phenotypic value of the kth F_2 individual (or F_2-derived family or testcross) was obtained as the sum of J_k and a random nongenetic effect, drawn from a normal distribution with a mean of zero and a variance of V_E. The true value of the genotypic variance (V_G) was needed to specify V_E for a given H, but the true value of V_G was unknown. The true value of V_G was therefore determined empirically by generating a large F_2 population (20,000 individuals) and calculating V_G as the variance among G_k values.

Detection of marker-QTL linkage

The QTL were mapped by multiple regression of phenotype on marker genotype. This method was chosen because it simultaneously considers several linked markers, and the multiple regression coefficients are expected to directly provide information on the presence or absence of a linked QTL (Doerge et al. 1994; Whittaker et al. 1996). The QTL mapping procedure was performed in two steps. In the first step, multiple regression of phenotypic value on the number of marker alleles (0, 1, or 2) from the first parental inbred was performed on a chromosome-by-chromosome basis. A backwards elimination procedure was used; this procedure was chosen because it allowed the examination of the full model (i.e., all markers) for each chromosome. The comparison-wise significance

level for retaining a marker in the model was α_C=0.0001, 0.001, 0.1, 0.05, 0.1, 0.2, 0.3, or 0.4. In the second step, multiple regression coefficients were obtained by jointly analyzing all the markers that were found significant in the per-chromosome analysis.

Suppose three markers are on the same chromosome in the order M_1–M_2–M_3. The regression coefficient for M_2 is expected to be significant if (1) a QTL is present in the M_1–M_2 interval, (2) a QTL is present in the M_2–M_3 interval, or (3) QTL are present in both the M_1–M_2 and M_2–M_3 intervals (Doerge et al. 1994; Whittaker et al. 1996). In other words, a significant marker regression coefficient indicates the presence of a flanking QTL. A true positive was declared whenever a marker (e.g., M_2) had a significant regression coefficient and a QTL was present in either or both of the marker's adjacent intervals (e.g., M_1–M_2 or M_2–M_3). Conversely, a false positive was declared whenever a marker had a significant regression coefficient but a QTL was absent in either of the marker's adjacent intervals. Whether or not a QTL was detected within a specific marker interval was considered as sufficient information for declaring a true positive or a false positive. The position of the QTL within the interval, which could be estimated by interval mapping methods (Lynch and Walsh 1998), was not considered as necessary information given the objectives of the study.

Three parameters were estimated. First, the total number of false positives was determined, and the average FDR was calculated across the 2,000 repeats of the experiment. Second, the power of the experiment was calculated as average number of QTL detected divided by the number of QTL present. Third, the average number of markers with significant effects was calculated.

Other models

Three variations of the main model were considered to examine the robustness of the results. First, the use of 50 or 200 markers (instead of 100) was considered. Second, a mapping population with 400 or 2,000 individuals (instead of 150) was studied. Third, a smaller genome (rye, *Secale cereale* L) and larger genome (wheat, *Triticum aestivum* L. em Thell) relative to maize were considered. The chromosome and genomes sizes for rye (seven chromosomes that comprised 727 cM) and wheat (21 chromosomes that comprised 3,436 cM) were obtained from published linkage maps (Ma et al. 2001; Gupta et al. 2002).

Marker-based recurrent selection

The responses to phenotypic selection and to marker-based recurrent selection, with different levels of α_C, were compared under the main model. In phenotypic selection, the best 10% of individuals in the initial F_2 population were inter-mated to form the next cycle of selection. In marker-based recurrent selection, marker scores for each F_2 individual (Lande and Thompson 1990) were calculated from the multiple regression coefficients for the markers with significant effects. The best 10% of individuals according to their marker scores were inter-mated to form the next cycle. Four cycles of phenotypic and marker-based recurrent selection were simulated. The same values of the multiple regression coefficients were used for each cycle of marker-based recurrent selection. Selection responses were calculated for each cycle, and the experiment was repeated 2,000 times.

Results

Main model

The FDR increased (Fig. 2) as the comparison-wise significance level increased. For a trait controlled by a single QTL and with a heritability of H=0.20, the FDR

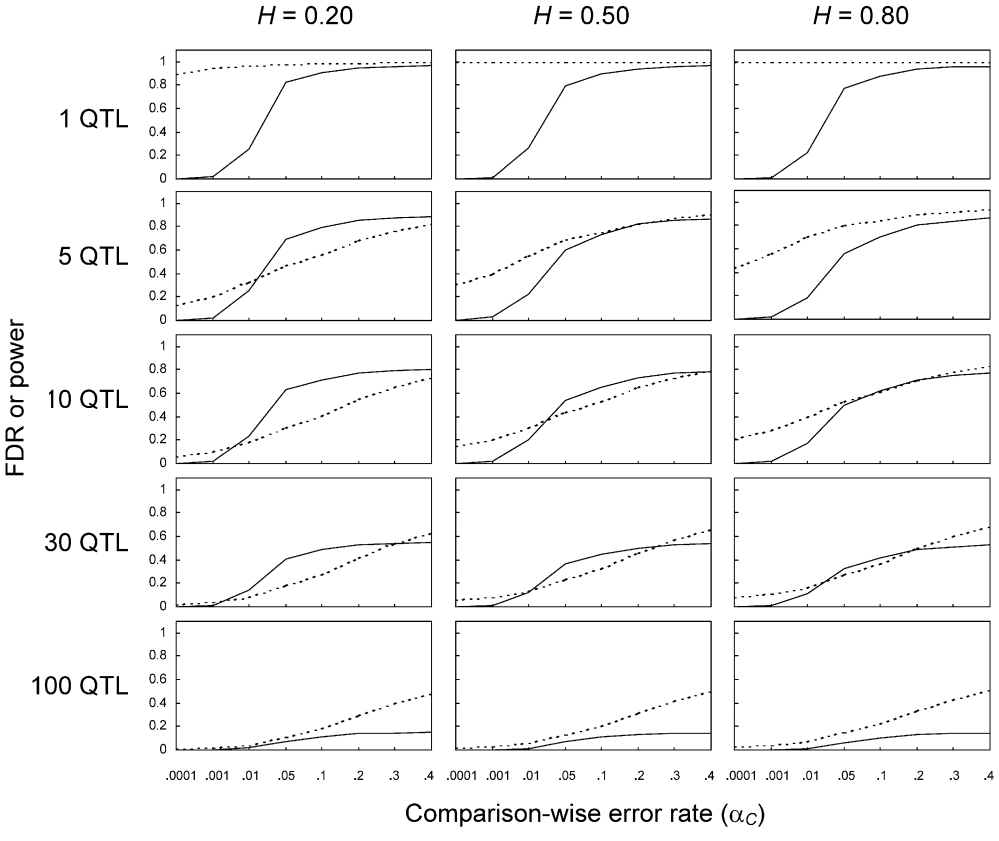

Fig. 2 False discovery rate (FDR, *solid lines*) and power for identifying QTL (*dashed lines*) for different comparison-wise significance levels (α_C), numbers of QTL and trait heritabilities (H)

increased from <0.01 for α_C=0.0001, to 0.82 for α_C=0.05. The FDR was greater than 0.90 when α_C was 0.10 or greater. The FDR increased as α_C increased regardless of the number of QTL. The effects of α_C on FDR were less dramatic, however, when the trait was controlled by a larger number of QTL. For a trait controlled by 100 QTL and with a heritability of 0.20, the FDR increased from 0 for α_C=0.0001, to 0.07 for α_C=0.05. When α_C ranged from 0.1 to 0.4, FDR was about 0.10–0.15.

For a given α_C, the decrease in the FDR for large numbers of QTL was accompanied by a decrease in power (Fig. 2). For a trait controlled by 100 loci and with a heritability of 0.20, the proportion of QTL detected increased from 0.01 for α_C=0.0001, to 0.11 for α_C=0.05. In other words, a significance level of α_C=0.0001 enabled the detection of only 1 out of 100 QTL, whereas a significance level of α_C=0.05 enabled the detection of 11 out of 100 QTL. In contrast, when the trait was controlled by a single QTL, the power to detect this QTL was 0.89–0.98 when α_C ranged from 0.0001 to 0.05.

An increase in heritability from 0.20 to either 0.50 or 0.80 led to a decrease in the FDR and an increase in power (Fig. 2). The effects of heritability, however, on the FDR and power were less than the effects of the number of QTL. When an α_C of 0.05 was used to identify QTL for a trait controlled by 100 loci, an increase in heritability from 0.20 to 0.80 increased the power only from 0.11 to 0.15. When many QTL controlled the trait, the power therefore remained low even when heritability was high.

The number of markers with significant effects increased as α_C increased, and was not affected substantially by the number of QTL or by heritability (results not shown). On average, about 1–2 markers (out of 100) had significant effects when α_C was 0.0001, regardless of the number of QTL or heritability. The number of markers with significant effects increased to 1–3 for α_C=0.001; 2–5 for α_C=0.01; 8–12 for α_C=0.05; 14–18 for α_C=0.1; 26–29 for α_C=0.2; 36–39 for α_C=0.3; and 45–48 for α_C=0.4.

Other models

An increase in the size of the mapping population led to a decrease in the FDR and an increase in power. Consider a trait controlled by 10 QTL, a heritability of 0.50, and α_C values from 0.0001 to 0.05. For a mapping population of 150 F_2 individuals, the FDR for these values of α_C was <0.01–0.54 (Fig. 2). The FDR decreased to <0.01–0.45 when the size of the mapping population was increased to 400 (Fig. 3A), and to <0.01–0.33 when the size of the mapping population was further increased to 2,000 (Fig. 3B).

The FDR was lower when 50 markers (Fig. 3C) rather than 200 markers (Fig. 3D) were used to map QTL (i.e., for a trait controlled by 10 QTL and with a heritability of 0.50). This result, however, must be interpreted with caution because the probability that a QTL is linked to a marker (as defined in this study) differs according to the number of markers used. Consider a trait controlled by only one QTL and a genome size of 1,749 cM for maize. When 50 markers are used to map QTL, 2/50=4% of the markers (i.e., the two markers that flank the QTL) can be correctly declared as linked to the QTL via multiple regression. Given that the size of a bin is approximately 1,749/50=35 cM, the maximum distance between the QTL and its nearest marker is 17.5 cM. When 200 markers are used, only 2/200=1% of the markers can be correctly declared as linked to the QTL, but the maximum distance between the QTL and its nearest marker decreases to about 4 cM. The increase in the FDR with more markers was therefore accompanied by a tighter linkage between a true QTL and its flanking markers.

The FDR was lower for a smaller genome (rye, 727 cM) than for a larger genome (wheat, 3,436 cM). For a trait controlled by 10 QTL and with a heritability of 0.50, the FDR in rye ranged from <0.01 to 0.25 when α_C ranged from 0.0001 to 0.05 (Fig. 3E). In contrast, the FDR in wheat for these values of α_C ranged from <0.01 to 0.72 (Fig. 3F). The power, however, remained similar between the two genomes.

Marker-based recurrent selection

The responses to marker-based recurrent selection were generally highest when α_C was at least 0.1. Across cycles of selection, the differences in the response to marker-based recurrent selection were small for α_C values between 0.05 and 0.2. For a trait controlled by 10 QTL and with a heritability of 0.20, the responses (in $\sqrt{V_G}$) in

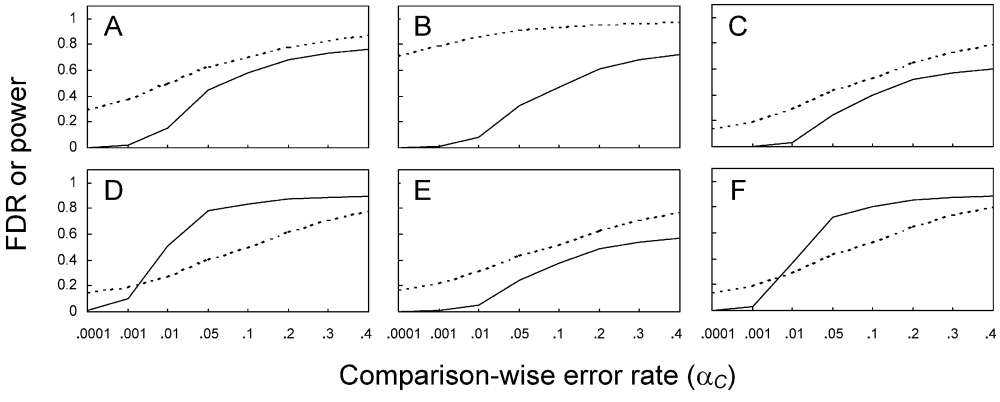

Fig. 3A–F False discovery rate (FDR, *solid lines*) and power for identifying QTL (*dashed lines*) for a trait controlled by 10 QTL and with a heritability of 0.50. **A** mapping population of 400, **B** mapping population of 2,000, **C** 50 markers, **D** 200 markers, **E** rye genome, **F** wheat genome

cycle 1 were similar across different levels of α_C, except for $\alpha_C \leq 0.001$ (Fig. 4). By cycle 2, however, the responses to marker-based recurrent selection favored liberal α_C values of 0.1 or 0.2.

By cycle 3, any advantage of marker-based recurrent selection over phenotypic selection disappeared (Fig. 4). This result held true for larger numbers of QTL (l=30 or 100) and higher heritabilities (H=0.50 or 0.80), for which any advantage of marker-based recurrent selection over phenotypic selection was small to begin with (results not shown).

Discussion

The results indicated that the FDR can be much higher than the comparison-wise significance level used to detect QTL. Regardless of the number of QTL controlling the trait and the size of the genome, a maximum α_C level of 0.0001 should be used to guarantee a FDR of 0.01 or less. The use of stringent α_C levels becomes particularly important in efforts to discover genes for traits that are likely controlled by few loci (e.g., l =1–10). Prior information on the number of loci controlling a trait may be unavailable, but in certain instances, such as resistance to Fusarium head blight (*Fusarium graminearum*) in wheat or to cyst nematode (*Heterodera glycines* Inchinoe) in soybean [*Glycine max* (L.) Merrill], researchers believe that relatively few genes control the trait. Studies to map QTL for these two traits have used α_C levels of ≤0.0001 (Anderson et al. 1999; Mudge et al. 1997).

The use of stringent α_C levels would reduce the number of reported QTL, but it should not be viewed as an impediment to further gene discovery. Much of QTL mapping will remain exploratory, and putative QTL detected at a less stringent α_C level should be subjected to further analysis. For example, fine-scale mapping or candidate gene analysis could be used for genomic regions where putative QTL have been detected at less stringent α_C levels.

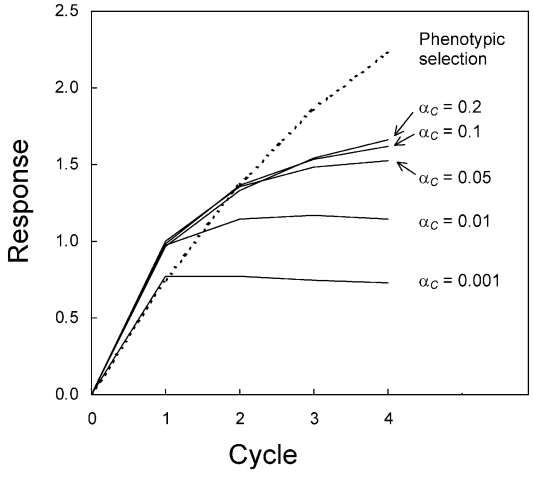

Fig. 4 Response to phenotypic selection and marker-based recurrent selection with different comparison-wise significance levels (α_C) for detecting QTL. The trait was controlled by 10 QTL and had a heritability of 0.20

In breeding for a trait controlled by a few genes, the QTL are likely to be exploited by introgressing the QTL with significant effects into elite germplasm (Dudley 1993). This approach has been tried for different traits in several species [see Bernardo (2002) for a review]. In this situation, the FDR should be kept low so that resources are not wasted in introgressing false QTL. Perhaps the success or failure in attempts to introgress QTL may be partly due to the α_C level used to identify QTL. In rice, for example, Shen et al. (2001) introgressed four QTL for deeper roots from one parent to the other parent of the mapping population. Three QTL had been detected at a significance level of α_C=0.05, whereas the fourth QTL had been detected at α_C=0.001 (Yadav et al. 1997). Fewer than 50% of the comparisons involving near-isogenic lines had improved root depth. While the use of a relaxed α_C level of 0.05 could not be construed as having directly caused the inconsistent results, one may speculate whether part of the problem was in identifying the QTL in the first place, rather than in introgressing and validating the four QTL.

For a trait controlled by many loci, such as yield in maize, QTL have been exploited not by introgressing the QTL into elite germplasm, but by marker-based recurrent selection (Edwards and Johnson 1994; Johnson 2001). In this procedure, which was simulated in this study, the mean of the population is improved by a few cycles of selection based on marker information. An improved base population leads to a better chance of obtaining superior families or inbreds, and no direct attempt is made at developing an inbred that has the favorable allele at all markers with significant effects. Hospital et al. (1997) found that for a trait controlled by 25 QTL, a liberal α_C level should be used in marker-based selection. This result was confirmed in the current study, which considered different numbers of QTL. The long-term advantage of phenotypic selection over marker-based recurrent selection also agreed with the results of Hospital et al. (1997).

The optimum α_C level for marker-based recurrent selection was 0.1–0.2. Such relaxed α_C levels led to the largest responses to selection and a greater power for the experiment, but at the expense of a large number of markers that were incorrectly declared to have significant effects. The use of a stringent α_C level of 0.0001 led to insufficient power (e.g., 0.01–0.02 for a trait with 100 QTL and a heritability of 0.20–0.50) for detecting QTL and, consequently, poor response to selection. The results confirmed that a type I error has little impact on marker-based recurrent selection, and that the main drawback of a type I error is the additional expense in scoring markers that need not have been scored.

Increasing the size of the mapping population leads to both increased power (Lande and Thompson 1990, Beavis 1994) and a lower FDR. In practice, however, large mapping populations may not be feasible because breeders tend to conduct marker-based recurrent selection in several populations simultaneously. The use of larger mapping populations would lead to fewer populations being

improved, and many breeders prefer to select in a large number of populations with relatively few progenies, instead of in a few populations with many progenies (Baker 1984; Hallauer 1990).

In conclusion, the question of "What proportion of declared QTL in plants are false?" cannot be answered definitely because QTL studies have used different significance levels, traits differ in the number of underlying QTL, and experiments have used different types of mapping populations (e.g., backcross instead of F_2 populations). The results do indicate, however, that for a trait controlled by few genes, the proportion of falsely declared QTL can be 10–30 times higher than the comparison-wise significance level used. In general, a QTL should be reported in the literature only if it has been identified at a stringent significance level (e.g., $\alpha_C \cong 0.0001$). The FDR, for a given α_C, is affected by the number of QTL and trait heritability (Fig. 2), as well by as the size of the mapping population, distance between markers, and size of the linkage map (Fig. 3). Methods proposed by Benjamini and Hochberg (1995) and Weller et al. (1998) can be used to estimate the level of α_C that maintains a specific FDR. Relaxed significance levels (e.g., α_C=0.1–0.2) are preferred in marker-based recurrent selection, but in this situation only those QTL that meet a more stringent significance level should be declared as QTL.

References

Anderson JA, Stack RW, Liu S, Waldron BL, Fjeld AD, Coyne C, Moreno-Sevilla B, Mitchell Fetch J, Song QJ, Cregan PB, Frohberg RC (2001) DNA markers for Fusarium head blight resistance QTLs in two wheat populations. Theor Appl Genet 102:1164–1168

Baker RJ (1984) Quantitative genetic principles in plant breeding. In: Gustafson JP (ed) Gene manipulation in plant improvement. Plenum, New York, pp 147–176

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: a practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300

Bernardo R (2002) Breeding for quantitative traits in plants. Stemma, Woodbury, Minn.

Bost B, Dillmann C, de Vienne D (1999) Fluxes and metabolic pools as model traits for quantitative genetics. I. The L-shaped distribution of gene effects. Genetics 153:2001–2012

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Coors JG (2001) Changing role of plant breeding in the public sector. Proc Corn Sorghum Ind Res Conf 56:48–56

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 breeding symposium series, Corvallis, Ore., 5–6 August 1994, pp 15–26

Dudley JW (1993) Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. Crop Sci 33:660–668

Edwards M, Johnson L (1994) RFLPs for rapid recurrent selection. In: Analysis of molecular marker data. Joint plant breeding symposium series, Corvallis, Ore., 5–6 August 1994, pp 33–40

Fernando RL (2002) Methods to map QTL. http://meishan.ansci.iastate.edu/rohan/notes-dir/QTL.pdf

Gupta PK, Balyan HS, Edwards KJ, Isaac P, Korzun V, Roder M, Gautier M-F, Joudrier P, Schlatter AR, Dubcovsky J, de la Pena RC, Khairallah M, Penner G, Hayden MJ, Sharp P, Keller B, Wang RCC, Hardouin JP, Jack P, Leroy P (2002) Genetic mapping of 66 new microsattelite (SSR) loci in bread wheat. Theor Appl Genet 105:413–422

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

Hospital F, Moreau L, Lacoudre F, Charcosset A, Gallais A (1997) More on the efficiency of marker-assisted selection. Theor Appl Genet 95:1181–1189

Johnson L (2001) Marker assisted sweet corn breeding: A model for specialty crops. Proc Corn Sorghum Ind Res Conf 56:25–30

Kacser H, Burns JA (1981) The molecular basis of dominance. Genetics 97:639–666

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

Keightley PD (1989) Models of quantitative variation of flux in metabolic pathways. Genetics 121:869–876

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

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

Ma X-F, Wanous MK, Houchins K, Rodriguez Milla MA, Goicoechea PG, Wang Z, Xie M, Gustafson JP (2001) Molecular linkage mapping in rye (*Secale cereale* L.). Theor Appl Genet 102:517–523

Mudge J, Cregan PB, Kenworthy JP, Kenworthy WJ, Orf JH, Young ND (1997) Two microsatellite markers that flank the major soybean cyst nematode resistance locus. Crop Sci 37:1611–1615

Openshaw S, Frascaroli E (1997) QTL detection and marker-assisted selection for complex traits in maize. Proc Corn Sorghum Ind Res Conf 52:44–53

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

Shen L, Courtois B, McNally KL, Robin S, Li Z (2001) Evaluation of near-isogenic lines of rice introgressed with QTLs for root depth through marker-aided selection. Theor Appl Genet 103:75–83

Thompson JN (1975) Quantitative variation and gene number. Nature 258:665–668

Weller JI, Song JZ, Heyen DW, Lewin HA, Ron M (1998) A new approach to the problem of multiple comparisons in the genetic dissection of complex traits. Genetics 150:1699–1706

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

Yadav R, Courtois B, Huang N, McLaren G (1997) Mapping genes controlling root morphology and root distribution in a doubled-haploid population of rice. Theor Appl Genet 94:619–632