Interval Mapping of Multiple Quantitative Trait Loci

Ritsert C. Jansen

Centre for Plant Breeding and Reproduction Research (CPRO-DLO), P.O. Box 1b, 6700 AA Wageningen, The Netherlands

ABSTRACT

The interval mapping method is widely used for the mapping of quantitative trait loci (QTLs) in segregating generations derived from crosses between inbred lines. The efficiency of detecting and the accuracy of mapping multiple QTLs by using genetic markers are much increased by employing multiple QTL models instead of the single QTL models (and no QTL models) used in interval mapping. However, the computational work involved with multiple QTL models is considerable when the number of QTLs is large. In this paper it is proposed to combine multiple linear regression methods with conventional interval mapping. This is achieved by fitting one QTL at a time in a given interval and simultaneously using (part of) the markers as cofactors to eliminate the effects of additional QTLs. It is shown that the proposed method combines the easy computation of the single QTL interval mapping method with much of the efficiency and accuracy of multiple QTL models.

CONVENTIONAL methods for the detection of quantitative trait loci (QTLs) are based on a comparison of single QTL models with a model assuming no QTL. For instance in the "interval mapping" method (LANDER and BOTSTEIN 1989) the likelihood for a single putative QTL is assessed at each location on the genome. However, QTLs located elsewhere on the genome can have an interfering effect. As a consequence, the power of detection may be compromised, and the estimates of locations and effects of QTLs may be biased (LANDER and BOTSTEIN 1989; KNAPP 1991). Even nonexisting so-called "ghost" QTLs may appear (HALEY and KNOTT 1992; MARTINEZ and CURNOW 1992). Therefore, it is obvious that multiple QTLs could be mapped more efficiently and more accurately by using multiple QTL models. KNAPP (1991), HALEY and KNOTT (1992) and MARTINEZ and CURNOW (1992) developed approximate methods for mapping QTLs using the information in the expected values of marker genotype means. JANSEN (1992) described a general mixture model for the case of multiple QTLs. Unfortunately, the computation involved with all these methods is almost infeasible when the number of QTLs is large. Also, standard multiple linear regression procedures are used in mapping QTLs (COWEN 1989, STAM 1991). The regression method is available in many statistical packages but suffers from the relative lack of interpretability in terms of genetic models. In these standard multiple linear regression procedures, the quantitative trait is regressed on the markers so that all markers are treated as if they are QTLs themselves. The effects of QTLs will be absorbed (partially) by linked markers. STAM (1991) showed that in a backcross population of infinite size QTL effects are fully absorbed by their flanking markers when these are

used as regressors. Although this will rarely be the case in finite populations (due to random deviations from the theoretical cosegregation ratio of markers), flanking markers will tend to absorb the effects of nearby QTLs. JANSEN (1992) suggested a detection and mapping approach that is basically a hybrid between the interval method and the multiple regression method. It was proposed to fit single QTL models (one per marker interval) and use (selected) markers to eliminate the effects of possible QTLs in other intervals. This can be achieved by using markers as cofactors in the regression of phenotype on genotype. Again, single QTL models may be compared with the model assuming no QTL, but now markers are used as cofactors. In the present paper this hybrid approach is worked out and illustrated for backcross populations, but the same ideas apply to other types of population; emphasis will be on detection aspects. A simple simulation study with three QTLs, two of them located on the same chromosome, is presented to illustrate the potential use of marker cofactors in the detection of multiple QTLs. A simulated example concerning detection of 11 QTLs on a genome of 10 chromosomes is also included.

SOME PRELIMINARY INVESTIGATIONS

A genome of two chromosomes was simulated 100 times in a backcross of F_1-individuals to one of the parental lines with two markers (M) and a single QTL (Q_1) on the first chromosome (MQ_1M/mq_1m), and with three markers and two other QTLs $(Q_2$ and $Q_3)$ on the second chromosome (MQ_2MQ_3M/mq_2mq_3m). The markers were set at a distance of 20 cM apart. The QTLs were located halfway between their flanking markers. The environmental contribution was normally distributed. The effects of the genes at the

R. C. Jansen

TABLE 1

Outline of the models fitted to compare different strategies for detection of QTL 1 (see also Figure 1)

Model	QTL fitted			Marker cofactors fitted
	1	2	3	
A	Yes	Yes	Yes	No
B	No	Yes	Yes	No
C	Yes	No	No	No
D	No	No	No	No
A′	Yes	No	No	⎡Flanking markers
B′	No	No	No	⎣of QTLs 2 and 3

TABLE 2

Outline of the models fitted to compare different strategies for detection of QTL 2 (see also Figure 2)

Model	QTL fitted			Marker cofactors fitted
	1	2	3	
A	Yes	Yes	Yes	No
B	Yes	No	Yes	No
C	No	Yes	No	No
D	No	No	No	No
A′	No	Yes	No	⎡Flanking markers
B′	No	No	No	⎣of QTLs 1 and 3

QTLs were additive; the additive deviations (half the differences between the homozygotes) were set to 1 SD. In all simulations data were generated for 200 individuals assuming absence of interference.

Tables 1 and 2 show the specification of the various models that were fitted to the simulated data. Expressions for the simultaneous likelihood of the observed phenotypic and genotypic (marker) data are given by JANSEN (1992). Let $\mathscr{L}_A$, $\mathscr{L}_B$, $\mathscr{L}_C$, $\mathscr{L}_D$, $\mathscr{L}_{A'}$ and $\mathscr{L}_{B'}$ denote the maximum log-likelihoods of the corresponding models. For instance, $\mathscr{L}_A$ can be written as follows

$$\mathscr{L}_A = \sum_{i=1}^{N} \log P(h)$$
$$+ \sum_{i=1}^{N} \log\left[\frac{1}{\sqrt{2\pi\sigma^2}} \sum_g P(g \mid h)\exp\left(-\frac{(y - m(g))^2}{2\sigma^2}\right)\right]$$

where y is the phenotype, h is the observed marker genotype with probability $P(h)$, g is the complete genotype (markers and QTLs) with conditional probability $P(g|h)$, $m(g)$ is the normal mean and σ^2 the normal variance of individuals with genotype g. In general $m(g) = m + A + M$, where A is now the additive component of the QTLs and M is the component for the marker cofactors with two levels per marker.

In conventional interval mapping, the detection of a QTL is based on $\mathscr{L}_C$-$\mathscr{L}_D$. $\mathscr{L}_A$-$\mathscr{L}_B$ is a similar expression, but now the QTLs on the other chromosome are also accounted for. Figure 1a shows that $\mathscr{L}_C$-$\mathscr{L}_D$ is less than $\mathscr{L}_A$-$\mathscr{L}_B$ in almost all simulations. Thus, higher power for detection of QTL 1 is achieved when taking
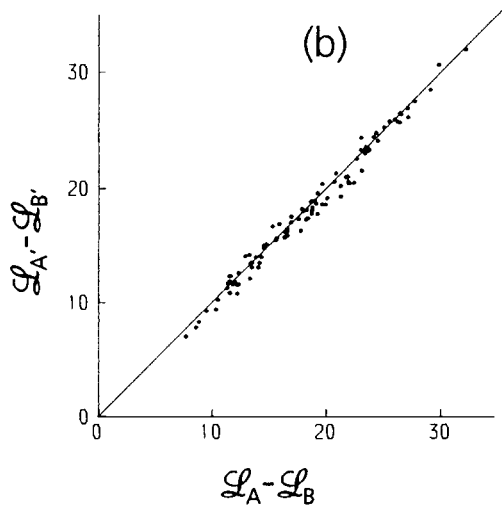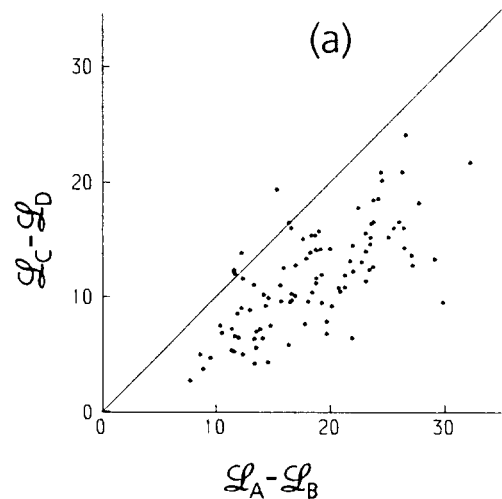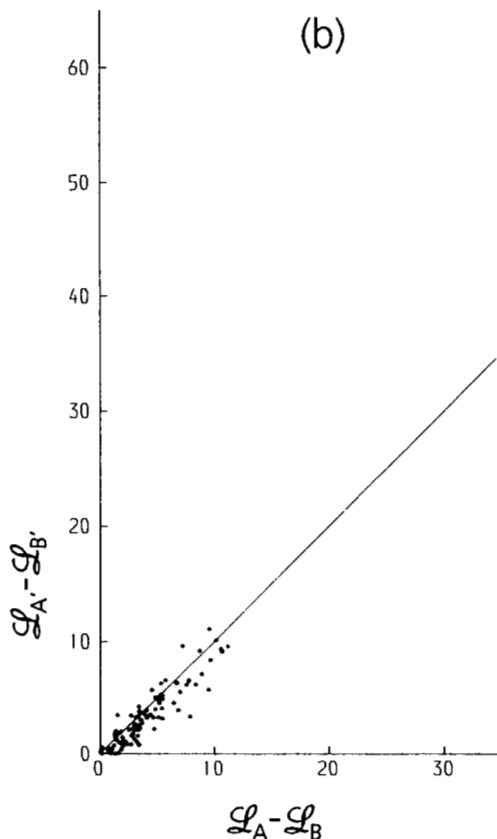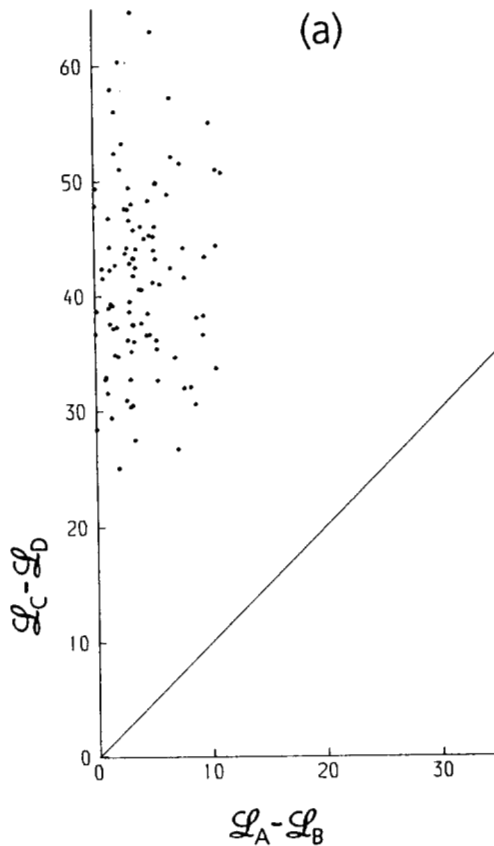

FIGURE 1.—A comparison of strategies for detection of QTL 1 (see also Table 1). A simulated backcross of F_1 individuals with MQ_1M/mq_1m on the first chromosome and MQ_2MQ_3M/mq_2mq_3m on the second chromosome to the parent with a normally distributed trait. The effects of the QTLs were additive and set to 1 SD. All distances between QTLs and flanking markers were set to 10 cM. $\mathscr{L}_A$, $\mathscr{L}_B$, $\mathscr{L}_C$, $\mathscr{L}_D$, $\mathscr{L}_{A'}$ and $\mathscr{L}_{B'}$ denote the maximum log-likelihoods of the corresponding models (Table 1). Differences $\mathscr{L}_A$-$\mathscr{L}_B$, $\mathscr{L}_C$-$\mathscr{L}_D$ and $\mathscr{L}_{A'}$-$\mathscr{L}_{B'}$ represent the contribution of QTL 1 to the log-likelihood. (a) Interval mapping approach vs. a multiple QTL approach. (b) Interval mapping approach using marker cofactors vs. a multiple QTL approach.

the QTLs on chromosome 2 into account. Figure 1b shows that taking the QTLs on chromosome 2 into account by using multiple QTL models results in about the same power as using the markers of chromosome 2 as cofactors. Contrary to Figure 1a, Figure 2a shows now that $\mathscr{L}_C$-$\mathscr{L}_D$ exceeds $\mathscr{L}_A$-$\mathscr{L}_B$ in all simulations. The single QTL model (model C) now absorbs the simultaneous effect of QTL 2 and 3, so that $\mathscr{L}_C$-$\mathscr{L}_D$ represents approximately the simultaneous contribution of QTLs 2 and 3, whereas $\mathscr{L}_A$-$\mathscr{L}_B$ represents the contribution of an additional QTL in the model. The fact that $\mathscr{L}_C$-$\mathscr{L}_D$ exceeds $\mathscr{L}_A$-$\mathscr{L}_B$ therefore indi-

cates the possible presence of multiple QTLs on chromosome 2. Figure 2b shows that again $\mathscr{L}_A$-$\mathscr{L}_B$ may be well approximated by $\mathscr{L}_{A'}$-$\mathscr{L}_{B'}$ using the flanking markers of QTLs 1 and 3 as cofactors.

In this example the "saturated" multiple QTL model still involves only three QTLs, which can be dealt with satisfactorily in terms of computational efforts. However, when the number of QTLs to be fitted simultaneously increases, the computational complexity quickly becomes prohibitive. Though representing a simple situation, the example clearly demonstrates the following points. First, searching for one QTL at a time by using markers as cofactors to absorb the effects of additional QTLs is (approximately) as powerful as searching for QTLs by dropping a single QTL from the full multiple QTL model. Second, the comparison of (a) the difference between the full multiple QTL model and one from which a single QTL is dropped, and (b) the difference between the conventional single and no-QTL model, is indicative of the presence of multiple QTLs on the same chromosome. In the next sections these ideas are extended to a general strategy for the detection of multiple QTLs.

A GENERAL STRATEGY FOR THE DETECTION OF MULTIPLE QTLS

The log-likelihoods of various models when maximized over unknown parameters provide a basis for choosing the genetic model that best fits the data: the genetic model that gives rise to the largest likelihood is the best fitting one. However, it is clear that, for instance, by adding an extra QTL or an extra marker cofactor to the model, the likelihood will increase. To allow for the fact that different genetic models depend on different numbers of parameters, we choose the genetic model that leads to the largest value of the log-likelihood ($\mathscr{L}$) minus a penalty for the number of free parameters (k) in the model. Equivalently, Akaike's Information Criterion (AIC)

$$AIC = -2(\mathscr{L} - k),$$

may be minimized (SAKAMOTO, ISHIGURO and KITAGAWA 1986). If the difference between AICs for two models is larger than 2, then the difference is considered to be significant (SAKAMOTO, ISHIGURO and KITAGAWA 1986). A single QTL model with the QTL located at a marker, position is equivalent to the model with that specific marker as cofactor, *i.e.*, that marker is also considered to represent a QTL. We impose no penalty on the AIC for the additional recombination

FIGURE 2.—A comparison of strategies for detection of QTL 2 (see also Table 2). A simulated backcross of F_1 individuals with MQ_1M/mq_1m on the first chromosome and MQ_2MQ_3M/mq_2mq_3m on the second chromosome to the parent. Differences $\mathscr{L}_A$-$\mathscr{L}_B$, $\mathscr{L}_C$-$\mathscr{L}_D$ and $\mathscr{L}_{A'}$-$\mathscr{L}_{B'}$ represent the contribution of QTL 2 to the log-likelihood, otherwise as Figure 1.
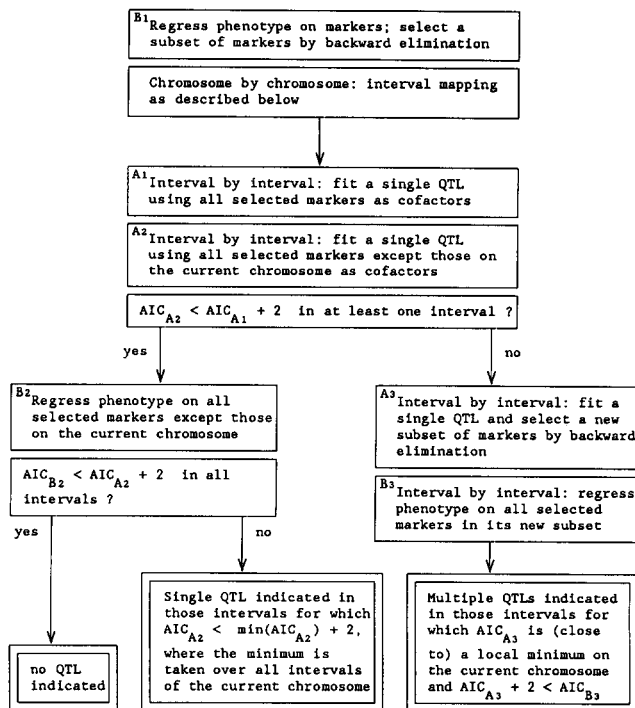
FIGURE 3.—Flow diagram for interval mapping of multiple QTLs.

parameter in a single QTL model. Then the above single QTL model and marker cofactor model have the same AIC. Our detection procedure consists of two stages: (1) selection of markers located closely to QTLs and (2) interval mapping using (subsets of) the selected markers to absorb effects of other QTLs. An example in the next section serves to illustrate the procedure. Figure 3 shows a flow diagram for the detection procedure; the details will now be described below.

The first stage starts with multiple regression of the quantitative trait on all markers. By a subset selection method for multiple regression (the method of backward elimination) markers are dropped from the model until no further reduction in AIC can be achieved. The final model is denoted by B_1. The final subset of markers will be used in the second stage of the procedure. Models A_1, A_2, A_3, B_1, B_2, B_3, C and D as used below, now refer to the models specified in Figure 3 and in the example given in the next section. Models A_1, A_2 and A_3 are single QTL models, models B_1, B_2 and B_3 are "no QTL" models. Models C and D are the commonly used models with and without a single QTL (no marker cofactors used). It is the difference among AICs that matters and not the actual values themselves. Therefore, we present AICs relative to the AIC of the multiple regression model using all selected markers (model B_1).

Selected markers (hopefully) indicate locations of QTLs or at least regions where QTLs are located. Important QTLs are located on those chromosomes for which the dropping of markers from the final multiple regression model results in a large increase

of AIC (model B_2 is compared with model B_1).

In the second stage (the interval mapping stage) selected markers are used as cofactors in the regression of phenotype on genotype (JANSEN 1992). Interval by interval, the AICs of several models are calculated. First, a single QTL model is fitted using all selected markers as cofactors (model A_1). However, by fitting the putative QTL, some (or all) of the selected markers on the current chromosome may now be redundant. This may be studied by dropping some or even all selected markers on the current chromosome (model A_1 is compared with models A_3 and A_2, respectively). Suppose that for some interval all selected markers on the current chromosome may be dropped without a loss in AIC (model A_2 fits better than model A_1). In that case a single QTL suffices to take over the role of these markers. However, if the putative QTL may also be dropped (model B_2 fits better than model A_2) without a loss in AIC, then no QTL is detected on the current chromosome. Alternatively, suppose that the AIC of model A_2 exceeds the AIC of model A_1 in all intervals. This indicates that a single QTL cannot take over the role of the markers, and the presence of multiple QTLs on the current chromosome is indicated. Then a second selection procedure is carried out interval by interval. Starting from the single QTL model using all selected markers, which markers still may be dropped (those markers previously explained the effect of the putative QTL) and which markers cannot be dropped (these markers possibly absorb the effects of other QTLs on the current chromosome) are studied. Dropped markers will often be the markers flanking the interval. Interval by interval, detection of the putative QTLs is now carried out by dropping the QTL (model B_3 is compared with model A_3) using for each interval its own subset of selected markers.

EXAMPLE

A simulated backcross example will be worked out in the case of a genome of 10 chromosomes and a quantitative trait that is affected by 11 QTLs spread over the chromosomes (Figure 4). The example serves to illustrate the behavior of our new interval mapping approach and to compare this approach with the traditional interval mapping method. Data were simulated for 500 individuals. Genotypes were generated assuming absence of interference. The markers were set at a distance of 20 cM apart; the QTLs were located halfway between their flanking markers. The environmental contribution was normally distributed. The effects of the genes were additive and the additive deviations were set to either 1 or −1 SD. Expressions for the simultaneous likelihood of the observed phenotypic and genotypic (marker) data are given by JANSEN (1992).

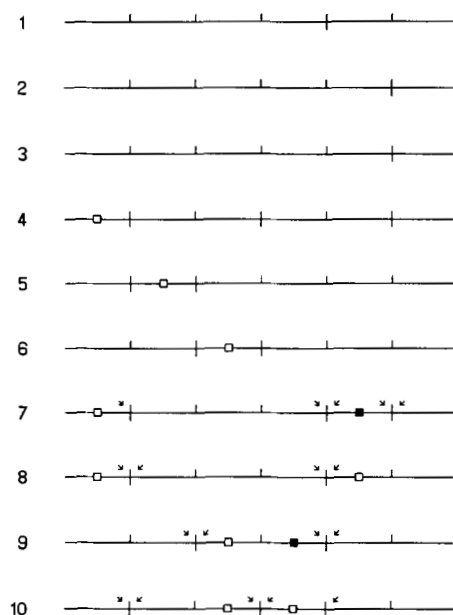Table 3 shows the specification of the various

FIGURE 4.—A simulated backcross of 500 individuals with a genome of 10 chromosomes and 11 QTLs spread over the chromosomes. The markers were set at a distance of 20 cM apart. The QTLs were located halfway between their flanking markers. Lines indicate chromosomes. Blocks indicate QTL positions, the effect of a QTL is either 1 (□) or −1 (■). Marker locations are indicated by a ⊥ and +. A subset of all markers was selected by backward elimination in multiple linear regression of the trait on the markers; selected markers are indicated by +, the other markers by ⊥. Arrows indicate per marker interval when the left (↙) or the right (↘) flanking marker is dropped in the second marker subset selection (see also text, Table 3 and Figure 3). Markers are numbered 1 to 5 from the left to the right on each chromosome.

TABLE 3

Outline of the models fitted in the example

Model	QTL fitted	Selected marker cofactors used on this/other chromosome(s)	
		This	Other
A₁	Yes	All	All
B₁	No	All	All
A₂	Yes	None	All
B₂	No	None	All
A₃	Yes	See Figure 4	All
B₃	No	See Figure 4	All
C	Yes	None	None
D	No	None	None

models fitted. Some results are presented in Table 4 for chromosomes 1, 6, 7, 8, 9 and 10; emphasis is on detection aspects. Results for other chromosomes were similar. A complete overview of the results would contain not only AICs as given in Table 4, but also a monitoring of AICs and parameter estimates during the whole detection process and for the complete genome.

The total phenotypic variance in the simulated data was equal to 3.76, which consisted of environmental variance (1.02) and genotypic variance (2.74). The explained (genotypic) variance was in the range of 0.00 to 0.74 when using single QTL models (conventional interval mapping), and in the range of 2.18 to 2.36 when using single QTL models with marker cofactors. This clearly demonstrates that a considerable part of the genotypic (QTL) variance was absorbed by marker cofactors. Table 4 shows that the AICs of single QTL models (conventional interval mapping) were large relative to the single QTL models with marker cofactors. Thus, the better fit of the model to the data is achieved when using marker cofactors.

The procedure indicates correctly the presence of no QTL on chromosomes 1, 2 and 3, the presence of one QTL on chromosomes 4, 5 and 6, and the presence of multiple QTLs on chromosomes 7, 8, 9 and 10. The multiple QTLs could be well separated on chromosomes 7 and 8. The estimates of the QTL effects on chromosome 9 take values −0.06, 0.36, 0.45, −0.67, −0.67 and 0.01 in the first up to the sixth interval (model A₃), which shows a clear changeover at the third marker. No clear separation of the two QTLs on chromosome 10 could be obtained. The estimates of the QTL effects on chromosome 10 are 1.27, 1.20, 1.63, 1.60, 1.11 and 0.94 in the first up to the sixth interval when using model A₂, and they are 0.44, 0.38, 1.24, 1.22, 0.30 and 0.05 when using model A₃. This change clearly represents the effect of the marker cofactors. However, the selected marker 1 was not replaced by the single QTL in the third interval. Nevertheless, the AIC for the model without marker 1 was close to the AIC for the given optimum model. Similar results hold for the selected marker 4 and the single QTL in the fourth interval. Thus, discrimination between the various models was poor.

The choice of a genetic model may also be based on additional considerations. For instance, a QTL is indicated in the fourth, fifth or sixth interval on chromosome 7, but the markers 4 and 5 are simultaneously redundant as cofactors only when fitting a QTL in the fifth interval (model A₃). Therefore, only a QTL in the fifth interval can take over the role of these markers. It may also be worthwhile to force a marker cofactor to be included in the selected subset. For instance, when observing the changeover in the estimated QTL effects at marker 3 of chromosome 9, a marker cofactor for marker 3 can be reentered into the model. The estimates of the QTL effects on chromosome 9 now take the slightly better values 0.52 and −0.77 in the third and the fourth interval, respectively (model A₃).

In the conventional interval mapping approach, the detection and mapping of QTLs is based on models with a single QTL (model C) and without a single QTL (model D). The AIC of the latter model is equal to 390.3. A LOD threshold of about 2.4 (LANDER and BOTSTEIN 1989; their Figure 4), or equivalently, an AIC threshold of about $2(2.4/\log_{10}e - 1) \approx 9.1$ is commonly used as a threshold for the detection of

TABLE 4

Interval mapping multiple QTLs: AIC values for various models in a simulated backcross (see Figure 4 for a description of the backcross and see Table 3 and Figure 3 for an outline of the models A_1, A_2, A_3, B_1, B_2, B_3, C and D)

Chromosome	Model	Marker interval					
		-1	1-2	2-3	3-4	4-5	5-
1	A_1	1.9	2.0	1.9	1.9	1.5	1.5
(0.6)	A_2	2.5	2.2	1.1	0.0	0.0	2.2
	C	391.5	390.9	385.7	385.6	391.8	392.3
6	A_1	2.0	2.0	2.0	2.0	1.8	1.6
(43.4)	A_2	28.3	2.5	-1.4	11.1	30.4	35.4
	C	385.0	377.4	374.0	377.6	388.0	389.0
7	A_1	2.0	2.0	1.0	0.0	2.0	2.0
(82.8)	A_2	60.2	60.3	79.4	61.5	54.7	57.4
	A_3	-2.7	-3.5	0.5	-2.2	0.0	0.0
	B_3	55.9	0.0	0.0	8.9	60.2	10.9
	C	371.1	371.3	392.3	381.0	374.7	375.8
8	A_1	1.5	1.9	1.4	1.1	-0.7	0.6
(108.0)	A_2	43.6	42.6	53.8	41.0	41.3	69.4
	A_3	-0.3	0.0	1.4	0.0	-2.5	0.6
	B_3	42.4	42.4	0.0	43.5	43.5	0.0
	C	346.5	341.7	346.9	348.4	349.3	369.3
9	A_1	1.9	1.9	1.4	1.5	2.0	2.0
(21.9)	A_2	23.9	23.9	22.1	5.3	5.7	16.9
	A_3	1.9	0.0	-0.5	0.0	0.0	2.0
	B_3	0.0	5.3	5.3	23.9	23.9	0.0
	C	392.2	391.7	392.0	389.3	389.0	391.2
10	A_1	2.0	2.0	2.0	2.0	1.9	1.9
(157.3)	A_2	106.9	79.4	5.2	7.0	87.7	125.0
	A_3	0.0	0.0	0.0	0.0	0.0	1.9
	B_3	6.1	6.1	55.0	55.0	2.1	0.0
	C	349.5	333.9	297.0	297.6	340.3	363.2

A subset of markers was selected by backward elimination in multiple linear regression of the trait on the markers. Only selected markers (or subsets of selected markers) were used as cofactors. All AICs are relative to the AIC for the multiple regression of the trait on the selected markers (model B_1). AICs for model B_2 are printed between brackets below the chromosome number, the AIC for model D equals 390.3.

QTLs. A QTL would be indicated then in those intervals for which the AIC of model C is less than 381.2. Following this approach, QTLs would be indicated on all chromosomes but chromosome 9. The putative QTL on chromosome 8 is most likely (but incorrectly) located in the second interval.

The example clearly demonstrates the following points. First, the AIC profile is much steeper around QTLs when using model A_2 instead of when using model C (see for instance chromosome 6 in Table 4). Therefore, the locations of the QTLs can be assessed more accurately when using marker cofactors. Second, the difference for AIC between model A_2 and model B_2 is often (much) larger than the difference for AIC between model C and model D in case a single QTL is indicated on a specific chromosome (see chromosome 6, results were similar for chromosomes 4 and 5). This difference is indicative for the effect of dropping the QTL, so that detection is more powerful when using marker cofactors. Finally, contrary to our method, conventional interval mapping does not indicate the presence of multiple QTLs on chromosomes 8, 9 and 10. In conclusion, the example demonstrates that more efficient detection and more accurate mapping can be achieved by the interval

mapping approach proposed here than by conventional interval mapping.

DISCUSSION

Detection of multiple QTLs is hampered by two main problems. First, though exact models for mapping multiple QTLs can be formulated (JANSEN 1992), the computational work involved is almost infeasible for large numbers of QTLs. Second, many genetic models have to be compared; thus, problems of model selection arise. In the present paper an approach is developed to get around these problems. In this approach only single QTL models are used, whereas effects of other QTLs are (hopefully well) eliminated by their flanking markers. A small simulation study demonstrated the usefulness of this approach for the detection of multiple QTLs. The Akaike Information Criterion (AIC) is used to evaluate the goodness of the assumed models (SAKAMOTO, ISHIGURO and KITAGAWA 1986). A model that minimizes the AIC, or models for which the AIC is close to the minimum, are considered to be the most appropriate. This procedure shows promise, as is suggested by the example: the results indicate that more efficient detection and more accurate mapping can be achieved by using our

approach than by using the conventional single QTL interval mapping approach. However, it should be noted that, even when it is detected that a specific chromosome contains multiple QTLs, large data sets may still be required to unravel the separate effects of closely linked QTLs.

Conventional interval mapping starts with a "no QTL" model and compares this model with a single QTL model. The test statistic shows the improvement in fitting a single QTL over fitting no QTL. If the improvement is significant, a second test may be carried out and the test statistic shows the improvement in fitting two QTLs over a single QTL, and so on. However, the first test may not be significant due, for instance, to linked genes with opposite effects or to unaccounted segregation on other QTLs. In conventional interval mapping the error of "missing an existing QTL" is uncontrolled and may therefore be high. It has also been reported that nonexisting "ghost" QTLs can appear, due to interference between undetected multiple QTLs (HALEY and KNOTT 1992; MARTINEZ and CURNOW 1992). The interval mapping method proposed in this paper starts with a hypothetical "polygenic" model to get around such detection and mapping problems concerned with interfering QTLs. This method has like multiple regression methods the advantage of controlling the error of "missing an existing QTL." In conventional interval mapping the probability of "detection by error of a QTL somewhere on the genome, whereas no QTL is actually present" is controlled. However, the assumption that "no QTL is actually present" makes no sense whenever a QTL is detected. In that case the significance level of the test is no longer known. Probabilities (and costs) of both error-types ("missing an existing QTL" or "detecting by error a QTL") may be balanced by the researcher; he may prefer to choose an AIC-threshold with a value other than the one used here (=2) for the comparison of models. Further research has to be done to study the probabilities of both error-types under various circumstances (*e.g.*, for different levels of heritability, different numbers of multiple QTLs, linked or unlinked QTLs, linked QTLs in repulsion or coupling phase, different population sizes and so on).

The general and flexible facilities of the mixture model approach described by JANSEN (1992) also apply to the interval mapping method proposed in this paper. For instance, it is possible to analyze non-normally distributed traits in addition to normally distributed traits, to take experimental design factors into ac-

count, or to carry out a (combined) analysis of different population types. Furthermore, the interval mapping method can readily be programmed in statistical computer packages that have facilities for generalized linear models. The observed quantitative trait and the observed marker genotypes should be specified by the user and standard output may then be produced. But a general procedure would make it also possible to specify the type of distribution for the trait or to include the experimental design. More advanced users may also be able to leave the beaten track and may try to fit alternative models. For instance, specific markers that were dropped during the process may be added again to the model. Alternatively, specific markers may now be excluded. For instance, selected markers that are located on chromosomes for which no QTL is detected may be dropped. The advanced user may also want to fit multiple QTL models, for instance two or three QTLs simultaneously on one chromosome, while taking into account additional QTLs on other chromosomes by marker cofactors. This is possibly the most accurate, efficient and still feasible way to unravel the separate effects of closely linked QTLs.

LITERATURE CITED

COWEN, N. M., 1989 Multiple linear regression analysis of RFLP data sets used in mapping QTLs, pp. 113–116 in *Development and application of molecular markers to problems in plant genetics*, edited by T. HELENTJARIS and B. BURR. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

HALEY, C. S., and S. A. KNOTT, 1992 A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity **69:** 315–324.

JANSEN, R. C., 1992 A general mixture model for mapping quantitative trait loci by using molecular markers. Theor. Appl. Genet. **85:** 252–260.

KNAPP, S. J., 1991 Using molecular markers to map multiple quantitative trait loci: models for backcross, recombinant inbred, and doubled haploid progeny. Theor. Appl. Genet. **79:** 583–592.

LANDER, E. S., and D. BOTSTEIN, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121:** 185–199.

MARTINEZ, O., and R. N. CURNOW, 1992 Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor. Appl. Genet. **85:** 480–488.

SAKAMOTO, Y., M. ISHIGURO and G. KITAGAWA, 1986 *Akaike Information Criterion Statistics*. KTK Scientific Publishers, Tokyo

STAM, P., 1991 Some aspects of QTL analysis, in *Proceedings of the Eighth Meeting of the Eucarpia Section Biometrics in Plant Breeding*. BRNO, July 1991.

Communicating editor: W. G. HILL