Estimating variance effect of QTL: an important prospect to increase the resolution power of interval mapping

A. B. KOROL1*, Y. I. RONIN1, Y. TADMOR2, A. BAR-ZUR2, V. M. KIRZHNER1 AND E. NEVO1

1Institute of Evolution, University of Haifa, Mount Carmel, Haifa 31905, Israel
2Corn Breeding Unit, Neve Yaar Research Center ARO, P.O.B. 90000, Haifa 31900, Israel

(Received 9 January 1995 and in revised form 6 September 1995)

Summary

Equal variances within quantitative trait locus (QTL) groups in the segregating population are a usual simplifying assumption in QTL mapping. The objective of this paper is to demonstrate the advantages of taking into account potential variance effect of QTLs within the framework of standard interval mapping approach. Using backcross case as an example, we show that the resolution power of the analysis may be increased in the presence of variance effect, if the latter is allowed for in the model. For a putative QTL (say, A/a) one can compare two situations, (i) $\sigma^2_{Aa} = \sigma^2_{aa}$ and (ii) $\sigma^2_{Aa} + \sigma^2_{aa}$. It was found that, if the variance effect of A/a is large enough, then in spite of the necessity to evaluate an increased number of parameters, the more correctly specified model provides an increase in the resolution power, as compared to the situation (i). This is not unexpected, if either $\sigma^2_{Aa}$ or $\sigma^2_{aa}$ in (ii) is lower than $\sigma^2_a$ from (i). But our conclusion holds even if $\sigma^2_{Aa} > \sigma^2_{aa} = \sigma^2_0$ or $\sigma^2_{Aa} > \sigma^2_{aa} = \sigma^2_a$. These advantages are illustrated on sweet corn data ($F_3$ families of $F_2$ genotypes). In particular, the log-likelihood test statistics and the parameter estimates obtained for a QTL locus in the distal region of chromosome 2 show that the allele enhancing the trait is recessive over the opposite allele simultaneously for the mean value and variance.

1. Introduction

The resolution capacity of marker analysis of quantitative traits (QTs) is the major factor affecting the practical importance of QT loci (QTL) mapping. A detailed discussion of the issues concerning the power of tests for detecting linkage can be found in many publications (e.g. Demenais et al. 1988; Lander & Botstein, 1989; Soller & Beckmann, 1990; Carbonell et al. 1993). The precision of the parameter estimates depends on the effect of the QTL in question relative to the total phenotypic variance of the trait in the mapping population. In other words, the higher the discrepancy between the distribution densities of the QTL groups (say, $f_{Aa}(x)$ and $f_{aa}(x)$, for a backcross), the better is the expected resolution. Among several possibilities to improve the precision of mapping, it is worth mentioning selective sampling (Lebowitz et al. 1987; Carey & Williamson, 1991; Darvasi & Soller, 1992), replicated progeny testing (Soller & Beckmann, 1990), sequential experimentation (Boehnke & Moll, 1989; Motro & Soller, 1993), multi-interval analysis (Jansen & Stam, 1994; Zeng, 1994) and multi-trait analysis (Korol et al. 1994, 1995; Ronin et al. 1995).

With few exceptions (Zhuchenko et al. 1979; Korol et al. 1981, 1994; Weller, 1986, 1987), variance effects of QTLs have not been considered previously. However, as rightly pointed out by Weller & Wyler (1992), the effect of a QTL on the variance is sometimes likely to be economically more critical than on the mean (e.g. for earliness, flowering time, ripening time under machine harvesting, time to hatching in chicken). The same applies to QTLs related to fitness traits in natural populations, e.g. seed dormancy, or flowering time. Clearly, this aspect of the problem might be very important in supporting climatic adaptive radiation into increasingly hostile conditions.

The objective of this paper is to demonstrate the advantages of taking into account potential variance effects of QT loci within the framework of the standard interval mapping approach. Using backcross case as an example, we show that the resolution power of the
markeral analysis of QTL may be increased in the presence of the variance effect, if the latter is allowed for by the model. This advantage is demonstrated on sweet corn data (F3 families of F2 genotypes) (Tadmor et al. 1994).

2. Interval analysis of QTL

(i) Mixture-model formulation

Assume that the putative QT locus A/a resides in an interval flanked by two marker loci, M1/m1 and M3/m3, with recombination rates r1 and r2 in intervals M1/m1—A/a and A/a—M3/m3. Many modes of interference in the interval could be considered, but we assume no interference, so that r = r1 + r2 — 2r1r2, where r is the recombination rate between M1/m1 and M3/m3. Based on the marker scores and measurements of the QT of interest (say, x) for individuals from the mapping population, we should test whether or not variation of x indeed depends on the interval M1/m1—M3/m3 and, if so, identify the corresponding locus A/a. For a backcross case, the expected distributions of the trait in each of the four marker groups, U1(x) = U(x), U2(x) = U(x), U3(x) = U(x), and U4(x) = U(x), can be written as:

\[ U_i(x) = \pi_i f_{aa}(x) + (1-\pi_i)f_{aa}(x), \quad i = 1, 4, \]  

the proportions \( \pi_i = \pi_i(r_1, r_2) \) depending on the unknown rates of recombination \( r_1 \) and \( r_2 \). With no interference,

\[ \pi_1 = (1-r_1)(1-r_2)/(1-r), \quad \pi_2 = r_1(1-r_2)/r, \quad \pi_3 = 1-\pi_2, \quad \text{and} \quad \pi_4 = 1-\pi_1. \]

The specification of the densities \( f_{aa}(x) \) depends on the assumptions made about the genetic control of the trait. Thus, if one anticipates that no other oligogenes affecting x are segregating or, by contrast, that the number of such loci is not too small, then normal density could be a good approximation, \( f_{aa}(x) = N(x_{aa}, \sigma_{aa}) \) and \( f_{aa}(x) = N(x_{aa}, \sigma_{aa}) \), where \( x_{aa} \) and \( x_{aa} \) are the expected mean values of x in groups aa and Aa, and \( \sigma_{aa} \) and \( \sigma_{aa} \) are the standard deviations.

(ii) Lod-score test and parameter estimation

Assuming that locus A/a belongs to the interval M1/m1—M3/m3, the log-likelihood for a sample of measurements \( x_i \) in marker groups with sizes \( n_i \) (i = 1, 4) can be written as:

\[ \ln L(\Theta_{st}) = \sum_{i=1}^{4} \sum_{k=1}^{n_i} \ln U_i(x_k) = \sum_{i=1}^{4} \sum_{k=1}^{n_i} \ln \left[ \pi_i f_{aa}(x_k) + (1-\pi_i)f_{aa}(x_k) \right], \]

where \( \Theta_{st} \) is the vector of \( s_t \) unknown parameters, specifying recombination rates and distribution of the trait x in the QTL groups aa and Aa. The assumption of no effect of genes from the interval M1/m1—M3/m3 on x can formally be presented by another set of parameters, \( \Theta = \Theta_{st} \) (the null hypothesis \{H0; \( \Theta = \Theta_{st} \)\}) as contrasting to the alternative one \{H1; \( \Theta = \Theta_{st} \)\}). According to the likelihood ratio test approach (Wilks, 1962), if H0 is true, the statistic

\[ \chi^2 = 2 \ln \left[ \max L(\Theta_{st})/\max L(\Theta_{st}) \right] \]

is distributed asymptotically as chi-square with \( s_t - s_0 \) degrees of freedom, where \( s_0 \) and \( s_t \) are the parameter spaces corresponding to \( H_0 \) and \( H_1 \), respectively. The null hypothesis \( H_0 \) is rejected if \( \chi^2 \) exceeds some critical value, corresponding to a preset level of significance. Then, the numerical values of the parameters that provide maximum to \( L(\Theta_{st}) \) could be considered as maximum likelihood estimates characterizing our QT locus A/a (Knott & Haley, 1992). However, the suitability of the chi-square approximation for the above test statistic remains an open question (e.g. Churchill & Doerge, 1994). In the multi-interval formulation of the QTL mapping problem, the exact asymptotic distribution of the log-likelihood ratio is also unknown (Lander & Botstein, 1989; Zeng, 1994). In such a case, Monte Carlo simulation for a given \( H_0 \) allows us to obtain an empirical critical value of the test statistics (2) (Knott & Haley, 1992; Ooijen, 1992; Zeng, 1994).

3. Dependence of the likelihood ratio on variance effect of the QT-locus

Let us estimate the joint effect of non-equal variances and averages in the putative QT locus groups of a backcross on the resolution of the lod-score test. For the sake of simplicity we consider the case of a very small marker interval, so that an approximation to no-recombination between marker loci and the putative QT is suitable (r = 0). We assume also that the estimates \( x_{aa} \), \( x_{aa} \), \( x_{aa} \), and \( \sigma_{aa} \) are replaced by their expectations. For such a simplified situation the likelihood ratio will be as follows:

\[
\begin{align*}
\prod_{i=1}^{n} \left\{ \frac{1}{\sqrt{2\pi\sigma_{aa}}} \exp\left[ \frac{-(x_{ti} - x_{aa})^2}{2\sigma_{aa}^2} \right] \right\} \prod_{i=1}^{n} \left\{ \frac{1}{\sqrt{2\pi\sigma_{aa}}} \exp\left[ \frac{-(x_{ti} - x_{aa})^2}{2\sigma_{aa}^2} \right] \right\} \\
\prod_{k=1}^{2} \left\{ \exp\left[ \frac{-(x_{tk} - 0.5(x_{aa} + x_{aa}))^2}{2\sigma_{aa}^2} \right] \right\} = B/C,
\end{align*}
\]
Variance effect in interval mapping of QTL

Fig. 1. The expected likelihood ratio (L) of QTL-marker analysis as a function of the variance effect \( \alpha = (\sigma_{AA}/\sigma_{aa}) \) and mean value effect \( \delta = d/\sigma_{aa} \) of the putative QT locus, assuming very close linkage between marker loci spanning the QT, and 250 individuals in the sample: (a) dependence of \( L \) on \( \alpha \), given fixed value of \( \delta \); (b) dependence of \( L \) on \( \delta \), given fixed value of \( \alpha \).

where \( d = x_{AA} - x_{aa} \), \( 2n = N \) is the size of the mapping population, so that \( n \) is the expected size of the two marker subgroups, and \( G = (\sigma_{aa}^2 + \sigma_{Aa}^2 + 0.5d^2) \) is twice the phenotypic variance of the trait in the population.

Let us now calculate the expected mean value of \( \log B/C = \log B - \log C = L \). We will have separately:

\[
\log B = -\log (2n\sigma_{aa}\sigma_{Aa}) + \sum_{k=1,2,2;1=1, n} [x_{ak} - 0.5(x_{aa} + x_{Aa})]^2 / G
\]

\[
\log C = -n \log \pi G
\]

Finally,

\[
L = \log(B/C) = n \log \left( \frac{\pi(\alpha + \alpha^{-1}) + \frac{1}{2} \delta^2/\alpha}{\alpha} \right) = L(\alpha, \delta),
\]

where \( \delta = d/\sigma_{aa} \) and \( \alpha = \sigma_{AA}/\sigma_{aa} \). This measure shows a monotonic increase with distance between the trait means of the QTL groups (\( \delta \)) and a non-monotonic (in general) dependence on the variance ratio of the groups (\( \alpha \)) (Fig. 1).

4. Simulation study

The above analysis, conducted for the case of a very close linkage between the marker loci spanning the putative QTL, revealed a surprising dependence of the test statistics (lod-score) on trait variance ratio. Namely, given any fixed value of the variance in one of the QTL groups, the test statistic may grow with increasing values of the variance in the other groups. It is of primary interest to see whether this effect holds also under looser linkage between the flanking markers. Such an examination was done by Monte Carlo simulations.

(i) Generating the data

Monte Carlo simulations were used to produce the ‘observations’. For each situation studied, 200 repeated backcross populations were generated. Normal distribution was used for the trait groups \( aa \) and \( Aa \). For comparative analysis of different situations we used, where possible, one and the same set of data. The composition of the marker groups (mixtures \( U_i, i = 1,4 \)) were modelled as binomial distributions with proportions \( p_i \). Clearly, this restriction is not principal. For most of the experiments, the parameter values used were in the range: \( 0.5 \leq d = x_{AA} - x_{aa} \leq 0.75 \), \( \sigma_{aa} = 1 \), \( 1 \leq \sigma_{AA} \leq 2 \), sample size \( N = 250 \); the length of the marker interval was 20 cM with the QT locus in the middle. No interference was assumed in our model (and Haldane’s mapping function is suitable).

(ii) Obtaining numerical solutions

The target of this work was to estimate the gain in the test power and estimation accuracy when the variance effect of the putative QTL is taken into account. Therefore, we do not dwell in this study on problems of numerical procedures of multidimensional optimization. The main objective here was to check how the difference of the QTL groups with respect to trait variance affects the detection power of the likelihood ratio test and closeness of the optimal points (representing the estimate of the parameter vector \( \Theta \)) to the true parameter set. For this specific goal, we do not have to search the solution starting from arbitrary points. The simplest way to obtain the necessary
estimates is to use as an initial point in the optimization procedure the parameter values equal to the ‘true’ ones of the considered sample (e.g. Titterington et al. 1985). Based on numerical analysis, we found that for the studied combinations of the model parameters this initial point lies in the domain of the attraction of the global maximum of the ML-functional. Of course, it could not be true for small sample sizes (Titterington et al. 1985). In case of simulation of the test statistic distribution under $H_0$ (see below) different starting points were used to ensure convergence to the global maximum. As tools for local optimization we employed different modifications of the gradient and Newton methods.

(iii) Estimation of the power of the test

In order to estimate approximately the power of the log-likelihood ratio test we used the critical level of the test statistics ($2 = \chi^2_{critical}$ based on the asymptotic distribution (chi square with D.F. = $s_1 - s_0$), where $s_1$ is the number of parameters in the model corresponding to $H_1$ and $s_0$ the number of parameters for the model $H_0$ (Wilks, 1962; Knott & Haley, 1992). In order to check whether the real distribution of the test statistic under $H_0$ is approximated by the asymptotic one, Monte Carlo simulations were conducted allowing us to obtain an empirical significance threshold. The proportion of cases where the QT locus was revealed
when it really exists was measured using both the asymptotic and empirical critical values.

(iv) Estimating the accuracy of obtained solutions

Usually, standard errors of the estimates are employed as a means for accuracy comparison of the estimation procedures. However, in addition to random fluctuations around the mean, another possible source of disturbances, the bias of the estimates, should also be taken into account. Thus, one should simultaneously take care of the estimation variance and estimate bias. These two measures are reflected in the average and standard deviation of the estimates.

(v) Resolution power as a function of the variance effect of QTL

For the backcross case, we have simulated and analysed some situations when a QT locus (A/a) residing in a marked interval \( (M_1/m_1 - M_2/m_2) \) affects both mean and variance of the QTL in question. In order to show the advantage of including the variance effect into the QTL detection and estimation model, we compared the resulting characteristics (test power and precision of estimates) with those obtained under assumption of equal variances and with the results when the putative QTL does not affect the trait variance (Table 1).

Consider the situation when equal variance assumption \( (\sigma^2_{aa} = \sigma^2_{Aa}) \) is made. Then, as one could expect, an increase in one of the variances (say, \( \sigma^2_{Aa} \)) leads to reduction in the power of the LOD test and in estimation precision of the genetic parameters. But much less expected are the results obtained for the same variants when the fact that \( \sigma^2_{aa} \neq \sigma^2_{Aa} \) is taken into account. In such circumstances a substantial increase in resolution can be achieved, as compared both with the last result and with the situation when A/a does not affect the trait variance (compare the corresponding columns in Table 1).

It is worthwhile to recall the foregoing conclusion for the limiting case of no recombination between the putative QT locus and flanking markers (see Fig. 1). Here the prediction of an increase in the resolution power of the LOD test for QTL detection due to the variance effect of the QT locus in question is verified by simulations for the more general case of non-zero recombination. It is interesting that the expected values of LOD obtained for \( r = 0 \) (see the above expression \( L(\alpha, \delta) \)) are, in fact, a good representation of the corresponding situations with non-zero recombination. One could easily come to this conclusion by comparison of LODs \( L(\alpha, \delta) \) values in the last column of Table 1 with the LODs obtained from Monte Carlo simulations when the variance effect \( (\sigma^2_{aa} \neq \sigma^2_{Aa}) \) is included into the model. Thus, besides biological importance (see Introduction), variance effect may help to improve the sensitivity and precision of the interval analysis, of course with an appropriate model. However, one may ask whether strong enough variance effects are possible in practice, to allow for their detection and a rise in the resolution power in spite of increased number of parameters in the model specifying the variance effect. An example of application of the proposed procedure to real data provided below shows that both questions can be answered positively.

5. An example of application

The following analysis of real data obtained in a study of sweet corn quantitative traits illustrates the points. A part of that study was devoted to reveal QTLs responsible for differences in sucrose content (Tadmor et al. 1994). Only an isolated example of Tadmor et al. (1994) results are presented here; the full analysis will be published elsewhere.

(i) Experimental design and the estimation procedure

Plants of an F2 population resulting from a cross between two sweet corn inbreds (IL731a and W6786) were scored for 93 markers (mainly RFLPs) from all 10 chromosomes. In order to obtain reliable estimates of the quantitative traits in question, each of the 214 F2 genotypes was represented by mean trait values of its F3 family (no less than 30 plants per family). However, beside an increased resolution, such a design complicates to some degree the estimation procedure.

Indeed, let us consider some QT locus A/B from an interval \( M_1/m_1 - M_2/m_2 \). If one takes from F2 a homozygote for this locus, either AA or BB, then for this locus the genotype of the entire F3 family will be the same, with the expected variance \( \sigma^2_{AA}/p \) or \( \sigma^2_{BB}/p \), respectively, where \( p \) is the family size. But for a heterozygote AB, the selfed F3 progeny will involve a mixture of three classes. Assuming no interaction between A/B and loci from other chromosomes, the expected mean value of the QT in question is

\[
X_{F2(F2-AB)} = 0.25x_{AA} + 0.5x_{AB} + 0.25x_{BB}.
\]

The expected variance can then be presented as:

\[
\sigma^2_{X_{F2(F2-AB)}} = \frac{(\sigma^2_{AA} + 2\sigma^2_{AB} + \sigma^2_{BB})}{2} + \frac{4(x_{AB} - x_{AA})^2 + 4(x_{AB} - x_{BB})^2 + (x_{BB} - x_{AA})^2}{32}/p.
\]

Thus, each of the nine marker groups for the interval \( M_1/m_1 - M_2/m_2 \) of our mapping F2 population is a mixture of three components, \( f_{AA}(x) \), \( f_{AB}(x) \), and \( f_{BB}(x) \), with means \( x_{AA} \), \( x_{F2(F2-AB)} = x(x_{AA}, x_{AB}, x_{BB}) \), and \( x_{BB} \), respectively, and variances \( \sigma^2_{AA}/2p \), \( \sigma^2_{AB}/2p \) (F2 = AB), \( \sigma^2_{BB}/2p \):

\[
U_i(x) = \pi_{AA}f_{AA}(x) + \pi_{AB}f_{AB}(x) + \pi_{BB}f_{BB}(x), \quad i = 1, \ldots, 9.
\]

Here the proportions \( \pi_i \) of the components depend on the position of A/B within the interval \( M_1/m_1 - M_2/m_2 \), and the chosen mapping function. Then

https://doi.org/10.1017/S0016672300033632 Published online by Cambridge University Press
for any set of parameters \( \Theta_{x_2} \) specifying the recombination rates and distribution of the trait \( x \) in the QTL groups \( AA, AB, \) and \( BB \), one can define the log-likelihood as

\[
\log L(\Theta_{x_2}) = \sum_{i=1}^{n} \sum_{k=1}^{2} \ln U_i(x_k),
\]

and employ the log-likelihood ratio test (see eqn 2).

The above procedure was applied to our data on sucrose content in dry kernels (40 d after pollination). The importance of the variance effect is demonstrated in the behaviour of the respective statistics obtained for the interval from umc36a to sel which was scored phenotypically (La Bonte & Juvik, Tadmor et al. 1994). This interval was shown to explain a major portion of sucrose content variation in dry seeds and was suggested to include the \( se \) gene itself (Tadmor et al. 1994). Our numerical analysis showed that along the considered portion of the chromosome 2 the log \( L \) is a unimodal function of \( \Theta \). The extremum can easily be obtained by an algorithm which involves scanning for \( r \) within the umc36a-sel interval with gradient optimization for other parameters of the set \( \Theta \).

(ii) Results

We considered 12 versions of the QT-locus dominance effect specification derived from three possibilities for inter-allele relationships with respect to trait mean value and four possibilities for the variance effect (Table 2). In each of these formulations the \( H_a \) hypothesis assumes ‘no effect on the trait mean value and trait variance’. Several conclusions may be derived from the presented tests and estimates. Let us denote by \( A \) the allele of the putative QTL from the umc36a-sel interval corresponding to the lower trait value, and by \( B \) the enhancing allele.

Consider first the general model, with no constraints either for trait mean effect or for the variance effect (model 1, or \( M_1 \), Table 2). The obtained estimates allow one to assume that the trait value of heterozygotes lies between the homozygotes with a clear tendency of \( A \) to dominate. For the variance effect, the QTL group homozygote for the enhancing allele \( BB \) has higher variance than the alternative homozygote group \( AA \), while the heterozygote group \( AB \) has the lowest variance. The last fact may reflect higher homeostatic ability of heterozygotes (Lerner, 1954). Comparison of the three cases presented in the first column of Table 2 allows one to conclude that with respect to the effect on the trait mean value the attenuating allele \( A \) is dominant over the enhancer \( B \) (the respective model 2 fits the data as well as the full model 1). The opposite assumption of dominance of \( B \) is rejected at high level of significance (\( \chi^2 \) for \( M_1 \) as against \( M_2 \) is 81.70 - 53.93 = 27.77, D.F. = 1).

The next step was to consider the variance effect. As before, we can start with the full model for mean value effect, i.e. to contrast different models of the first row. Comparison of \( M_1 \) and \( M_2 \) clearly demonstrates the importance of the variance effect: \( \chi^2 = 81.70 - 42.83 = 38.87 \) (D.F. = 2). Thus, beside losing biologically important information, neglecting this effect results in a very serious reduction in the resolution power. In addition, some bias in the estimation of the QTL position within the interval is also possible. Further, comparison of the full model \( M_3 \) with \( M_4 \) and \( M_5 \) leads to the conclusion that allele \( A \) tends to be dominant over \( B \) with respect to the variance effect. Taking into account also the above result concerning the mean trait effect, we can expect that the most economic description of the putative QT locus will be a model with allele \( A \) dominant for both trait mean and variance effects. And indeed, the results presented in Table 2 and Fig. 2 confirm this expectation (the model \( M_6 \) provides nearly the same LOD value as the full model \( M_1 \) employing a lower number of parameters).

Notably, the \( se \) phenotype was originally described as dependent on a recessive gene (Ferguson et al. 1978). In our LOD score based identification of the putative QT (\( A/B \)) the interval umc36a-sel was the most distal one available on chromosome 2. The estimates obtained indicate that the revealed locus \( A/B \) lies within this interval close to sel. To come to a final decision whether or nor \( A/B \) indeed coincides with sel, one more interval, distal to sel, is required.

6. Discussion

Equal variances of the trait in question within the QT locus groups are a usual simplifying assumption in marker analysis of quantitative traits, both for single marker analysis and interval mapping. The dependence of resolution power of marker analysis of QTs on the assumption of ‘no effect of the putative QT locus on trait variance’ was considered recently for single marker cases (Weller & Wyler, 1992; Korol et al. 1994; Ronin et al. 1995). The conclusions we drew for such situations earlier, and confirmed here for interval mapping, seem to contradict the intuitive expectations. Indeed, consider a simple example of a QT segregating in a backcross progeny. One can compare two situations, \( \sigma^2_{AA} = \sigma^2_{BB} = 1 \) and \( \sigma^2_{AA} + \sigma^2_{BB} = 1 \). Clearly, if the last inequality is \( \sigma^2_{AA} < \sigma^2_{BB} \), and the inequality of the variances is allowed by the model, then an increase in resolution is expected (compared to the situation with \( \sigma^2_{AA} = \sigma^2_{BB} = 1 \)) and this definitely will be the case. But what will be the consequences of an opposite situation, \( \sigma^2_{BB} > \sigma^2_{AA} = 1 \), if the variance effect is included in the model? Clearly, the situations \( \sigma^2_{AA} < \sigma^2_{BB} = 1 \) and \( \sigma^2_{BB} > \sigma^2_{AA} = 1 \) are non-equivalent: in the first one the relative distance of the means of the QTL groups is larger and in the second one smaller than in case of \( \sigma^2_{AA} = \sigma^2_{BB} = 1 \). We found that if the effect of the putative QT on trait variance is large, then allowing for \( \sigma^2_{AA} = \sigma^2_{BB} \) in the model increases the
Table 2. LOD score analysis of 12 models specifying mode of action of a putative QT locus (interval umc36a-se/ of the chromosome 2) affecting sucrose content in dry kernels of sweet corn. The most suitable marginal models (M_2 and M_7) as well as the best final model (M_f) are boxed.

<table>
<thead>
<tr>
<th>Assumption</th>
<th>σ_{AA} ≠ σ_{AB} ≠ σ_{BB}</th>
<th>σ_{AA} ≠ σ_{AB} = σ_{BB}</th>
<th>σ_{AA} = σ_{AB} ≠ σ_{BB}</th>
<th>σ_{AA} ≠ σ_{AB} = σ_{BB}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M_1</td>
<td>M_4</td>
<td>M_7</td>
<td>M_{10}</td>
</tr>
<tr>
<td></td>
<td>Mean σ</td>
<td>Mean σ</td>
<td>Mean σ</td>
<td>Mean σ</td>
</tr>
<tr>
<td>AA</td>
<td>35.7 10.8</td>
<td>35.4 11.8</td>
<td>34.9 8.3</td>
<td>35.0 8.3</td>
</tr>
<tr>
<td>AB</td>
<td>38.4 7.4</td>
<td>39.6 11.8</td>
<td>38.7 8.3</td>
<td>39.9 12.7</td>
</tr>
<tr>
<td>BB</td>
<td>57.4 14.4</td>
<td>55.1 11.8</td>
<td>57.4 17.5</td>
<td>54.7 12.7</td>
</tr>
<tr>
<td>r</td>
<td>0.080</td>
<td>0.090</td>
<td>0.077</td>
<td>0.088</td>
</tr>
<tr>
<td>LOD</td>
<td>17.74</td>
<td>9.30</td>
<td>16.81</td>
<td>10.61</td>
</tr>
<tr>
<td>χ^2(D.F.)</td>
<td>81.7(5)</td>
<td>42.8(3)</td>
<td>77.4(4)</td>
<td>48.8(4)</td>
</tr>
<tr>
<td></td>
<td>M_2</td>
<td>M_5</td>
<td>M_8</td>
<td>M_{11}</td>
</tr>
<tr>
<td></td>
<td>Mean σ</td>
<td>Mean σ</td>
<td>Mean σ</td>
<td>Mean σ</td>
</tr>
<tr>
<td>AA</td>
<td>38.0 11.4</td>
<td>38.4 11.8</td>
<td>37.6 8.5</td>
<td>37.9 8.9</td>
</tr>
<tr>
<td>AB</td>
<td>38.0 7.3</td>
<td>38.4 11.8</td>
<td>37.6 8.5</td>
<td>37.9 12.6</td>
</tr>
<tr>
<td>BB</td>
<td>57.3 17.4</td>
<td>55.6 11.8</td>
<td>57.4 17.4</td>
<td>55.1 12.6</td>
</tr>
<tr>
<td>r</td>
<td>0.076</td>
<td>0.086</td>
<td>0.078</td>
<td>0.083</td>
</tr>
<tr>
<td>LOD</td>
<td>17.41</td>
<td>8.69</td>
<td>15.98</td>
<td>9.38</td>
</tr>
<tr>
<td>χ^2(D.F.)</td>
<td>80.2(4)</td>
<td>40.0(2)</td>
<td>73.6(3)</td>
<td>43.2(3)</td>
</tr>
<tr>
<td></td>
<td>M_3</td>
<td>M_6</td>
<td>M_9</td>
<td>M_{12}</td>
</tr>
<tr>
<td></td>
<td>Mean σ</td>
<td>Mean σ</td>
<td>Mean σ</td>
<td>Mean σ</td>
</tr>
<tr>
<td>AA</td>
<td>39.9 11.1</td>
<td>35.1 13.2</td>
<td>34.8 8.4</td>
<td>35.0 8.1</td>
</tr>
<tr>
<td>AB</td>
<td>39.1 7.5</td>
<td>43.7 13.2</td>
<td>39.7 8.4</td>
<td>43.8 14.3</td>
</tr>
<tr>
<td>BB</td>
<td>39.1 25.2</td>
<td>43.7 13.2</td>
<td>39.7 24.9</td>
<td>43.8 14.3</td>
</tr>
<tr>
<td>r</td>
<td>0.076</td>
<td>0.078</td>
<td>0.075</td>
<td>0.074</td>
</tr>
<tr>
<td>LOD</td>
<td>11.71</td>
<td>2.37</td>
<td>11.02</td>
<td>4.90</td>
</tr>
<tr>
<td>χ^2(D.F.)</td>
<td>53.9(4)</td>
<td>10.9(3)</td>
<td>50.8(3)</td>
<td>22.6(3)</td>
</tr>
</tbody>
</table>

These results are in agreement with our previously suggested explanation of an analogous effect of non-equal variances σ_{AA}^2 ≠ σ_{BB}^2 in single marker analysis (Korolev et al. 1994). It employs a notion of 'discrepancy of the QT locus group distributions', D(μ_{aa}(x),μ_{Aa}(x)), as a function of d = x_{aa} - x_{Aa} and σ_{aa}^2/σ_{Aa}^2. We found that both D(μ_{aa}(x),μ_{Aa}(x)) and the resolution power may grow not only with increasing d = x_{aa} - x_{Aa} (which is quite expected), but also with increasing σ_{aa}^2/σ_{Aa}^2, provided d is relatively small. As one could easily see from Fig. 1B, for any large enough α = σ_{aa}^2/σ_{Aa}^2 > 1, a range of values δ = d/σ_{aa} could be found, so that within this range the power of the test is higher than for the corresponding case with α = 1 (in spite of the increased number of parameters). The last fact may help to understand better the seemingly paradoxical behaviour of the resolution power as a function of α = σ_{aa}^2/σ_{Aa}^2. The putative QT locus could affect either the mean value of the trait x or its variance, or both. Thus, even if d ≈ 0, one could try to identify the effect of a —> A substitution on the trait variance. It is natural to expect that with small d the resolution should grow with increasing effect of the locus A/a on the trait variance, i.e. with increased deviation of α from unity, e.g. due to increase of σ_{aa}^2 given fixed σ_{Aa}^2.

Using sweet corn data on sucrose content in dry kernels also in the case σ_{aa}^2 > σ_{Aa}^2 = 1, as compared to σ_{aa}^2 = σ_{Aa}^2 = 1.

Resolution also in the case σ_{aa}^2 > σ_{Aa}^2 = 1, as compared to σ_{aa}^2 = σ_{Aa}^2 = 1.

Fig. 2. Comparison of several models specifying the mode of action of a QTL affecting sucrose content in dry kernels. The log-likelihood test and the estimation procedure were applied to the most distal available interval (umc36a-se/) of the chromosome 2. For the description of the models see Table 2.

See PDF for the image.
kernels, we demonstrated the possibility of extracting important biological information concerning the mode of QTL effect on trait mean and variance. In particular, the conducted log-likelihood tests and obtained parameter estimates show that the allele enhancing the trait is recessive over the opposite allele simultaneously for the mean value and variance. In spite of the necessity to evaluate an increased number of parameters, the more correctly specified model allowed us to achieve a serious increase in the power of the test and better precision of the resulting estimates.

It is likewise important that variance effect of QTLs may be of primary biological interest and the target of mapping efforts (Zhuchenko et al. 1979; Korol et al. 1981, 1994; Weller & Wyler, 1992). Of special interest may be variance effects of loci involved in control of fitness related quantitative traits in natural populations, like seed dormancy, or flowering time. Clearly, fitness related quantitative traits in natural populations may be variance effects of loci involved in control of fitness related quantitative traits in natural populations, like seed dormancy, or flowering time. Clearly, fitness related quantitative traits in natural populations may be variance effects of loci involved in control of fitness related quantitative traits in natural populations, like seed dormancy, or flowering time.

References


