Single- and Multiple-Trait Mapping Analysis of Linked Quantitative Trait Loci: Some Asymptotic Analytical Approximations

Yefim I. Ronin, Abraham B. Korol and Eviatar Nevo

Institute of Evolution, University of Haifa, Haifa 31905, Israel

Manuscript received January 20, 1998
Accepted for publication September 28, 1998

ABSTRACT

Estimating the resolution power of mapping analysis of linked quantitative trait loci (QTL) remains a difficult problem, which has been previously addressed mainly by Monte Carlo simulations. The analytical method of evaluation of the expected LOD developed in this article spreads the "deterministic sampling" approach for the case of two linked QTL for single- and two-trait analysis. Several complicated questions are addressed through this evaluation: the dependence of QTL detection power on the QTL effects, residual correlation between the traits, and the effect of epistatic interaction between the QTL for one or both traits on expected LOD (ELOD), etc. Although this method gives only an asymptotic estimation of ELOD, it allows one to get an approximate assessment of a broad spectrum of mapping situations. A good correspondence was found between the ELODs predicted by the model and LOD values averaged over Monte Carlo simulations.

MANY efforts have been devoted to increasing the efficiency of marker analysis of quantitative traits, including interval analysis (Lander and Botstein 1989; Knott and Haley 1992), selective sampling (Lebowitz *et al.* 1987; Darvasi and Soller 1992, 1994; Weller *et al.* 1997), replicated progeny testing (Soller and Beckmann 1990), and sequential experimentation (Boehnke and Moll 1989; Motro and Soller 1993). Recently, a general method to improve the efficiency of quantitative trait loci (QTL) mapping was proposed by taking into account simultaneous segregation at many genomic segments that affect the trait in question (Jansen and Stam 1994; Zeng 1994). A situation in which one QTL (or a chromosome segment) affects several traits simultaneously can also be considered to result in increased power (Korol *et al.* 1987, 1994, 1995, 1998; Jiang and Zeng 1995; Ronin *et al.* 1995; Zeng 1997). Such an analysis may be important in marker-assisted breeding strategies, dissecting heterosis as a multilocus multitrait phenomenon, developing optimized programs for evaluation and bioconservation of genetic resources, and revealing genetic architecture of fitness systems in natural populations, etc. Multiple-trait mapping analysis proved to be very useful within the framework of the selective genotyping design (Weller *et al.* 1997; Ronin *et al.* 1998).

The multiple-trait approach may help in coping with a complicated problem arising when the considered chromosome contains several QTL (*e.g.*, Jiang and Zeng 1995; Korol *et al.* 1998). If one tries to fit a single-locus model to such a case, a ghost QTL can be detected in an interval that has no effect on the trait (Knott and Haley 1992; Martinez and Curnow 1992; Wright and Kong 1997). Especially difficult are situations with *trans* effects of linked QTL (Knott and Haley 1992; Luo and Kearsey 1992). That *trans*-association of QTL could be a common phenomenon even in interspecific crosses has been demonstrated by DeVicente and Tanksley (1993) in tomato: they found that up to 36% of the detected QTL had alleles with effects opposite to the direction expected from the parental differences.

The usual way of dealing with several linked QTL is multiple regression analysis or mixture model analysis that includes markers as regression cofactors to account for segregation of QTL of the same chromosome (Jansen and Stam 1994; Zeng 1994). The third possibility is to construct two- to three-interval mixture models, although this approach is rather cumbersome and needs intensive calculations. Employing Monte Carlo simulations with mixture models, we demonstrated recently the advantage of multiple trait analysis in detection of linked QTL effects (Korol *et al.* 1998). The goal of this article is to elaborate an analytical model enabling us to evaluate in a general form the expected LOD values in cases of two linked QTL. Such a model can be used as a tool to predict the expected resolution in different complicated situations. As a practical application one can consider the possibility of calculating

Corresponding author: A. B. Korol, Institute of Evolution, University of Haifa, Mt. Carmel, Haifa 31905, Israel.
E-mail: korol@esti.haifa.ac.il

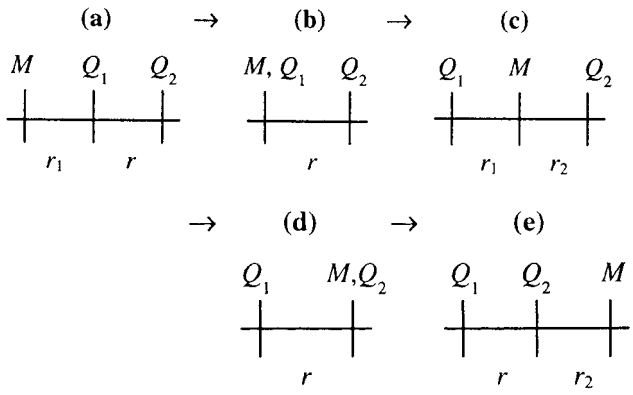
the minimum sample size needed to detect linked QTL with certain effects on either of the correlated quantitative traits or to prove the existence of epistasis for any of the traits. Likewise, the proposed analysis allows us to predict situations where a ghost QTL will be detected using interval analysis and to evaluate the minimum marker density needed to prevent such a possibility for given effects of the linked QTL. Recently, a similar technique, referred to as "deterministic sampling," was applied to single-QTL situations in single-trait analysis, with the expected LOD values calculated numerically (Mackinnon and Weller 1995; Mackinnon *et al.* 1996; Wright and Kong 1997). Our major target here is analytical and numerical deterministic sampling for *two-trait analysis* with linked QTL. We first treat the case of a single-trait analysis and then generalize the results for the two-trait analysis. The consideration will be based on a modification of the maximum-likelihood technique relevant to asymptotic properties of the LOD test, which will be referred to as "regression of the log-likelihood function." For the case of single-marker analysis this modification is equivalent to the usual procedure of expected LOD (ELOD) calculation (Lander and Botstein 1989) with the only difference that it is a function of the variable position of marker.

SINGLE-TRAIT ANALYSIS

The major target of our analysis is analytical and numerical deterministic sampling with linked QTL. Therefore, analytical expression of ELOD should be obtained that allows us to compare H_2 (two linked QTL) and H_1 (single QTL) for any set of parameter values.

Single-QTL models: Let a trait x be dependent on two linked loci Q_1/q_1 and Q_2/q_2 and let the trait values in the four QTL groups $Q_1Q_1Q_2Q_2$, $Q_1Q_1q_2q_2$, $q_1q_1Q_2Q_2$, and $q_1q_1q_2q_2$ of a mapping population have normal densities $f_{11}(x)$, $f_{12}(x)$, $f_{21}(x)$, and $f_{22}(x)$ with (unknown) means μ^*_{11}, μ^*_{12}, μ^*_{21}, and μ^*_{22} and standard deviations σ^*_{11}, σ^*_{12}, σ^*_{21}, and σ^*_{22}, respectively. Usually, the mapping procedure is started with the assumption of one QTL in the chromosome and then one can try to apply some versions of single-marker or interval analysis. Reduced test power, biased parameter estimates, and detection of ghost factors may result from this simplification, as demonstrated by simulation studies (Knott and Haley 1992; Luo and Kearsey 1992; Martinez and Curnow 1992; Korol *et al.* 1998). This question is treated here analytically for both single-marker and single-interval analysis (see also Wright and Kong 1997).

Consider a random sample of individuals genotyped for marker loci from the chromosome that carries the two QTL. With a dense molecular map, one analyzes consequently a series of markers with five different locations relative to the linked QTL:



Clearly, situations (d) and (e) are equivalent (up to parameter replacement) to (b) and (a), respectively. Our intention is to evaluate how misspecification of the model (assumption of one QTL when actually two linked QTL reside on the chromosome) affects the parameter estimation. This is done by scanning across a large number of markers, so that besides situations (a), (c), and (e), one could also encounter situations close to those of (b) and (d). Moreover, in all of the cases we assume that the trial marker exactly coincides with the putative (single) QTL. Due to the foregoing assumptions, the true expected densities of the trait distribution in the alternative marker groups for an arbitrary marker will be

$$h_{MM}(x) = h_1 = \alpha_1 f_{11}(x) + \alpha_2 f_{12}(x) + \beta_2 f_{21}(x) + \beta_1 f_{22}(x),$$
$$h_{mm}(x) = h_2 = \beta_1 f_{11}(x) + \beta_2 f_{12}(x) + \alpha_2 f_{21}(x) + \alpha_1 f_{22}(x),$$

where in case

$$\begin{aligned}
&(a)\quad \alpha_1 = (1-r_1)(1-r), && \alpha_2 = (1-r_1)r,\\
&\qquad \beta_1 = r_1(1-r), && \beta_2 = r_1 r,\\
&(b)\quad \alpha_1 = 1-r, && \alpha_2 = r,\\
&\qquad \beta_1 = 0, && \beta_2 = 0,\\
&(c)\quad \alpha_1 = (1-r_1)(1-r_2), && \alpha_2 = (1-r_1)r_2,\\
&\qquad \beta_1 = r_1 r_2, && \beta_2 = r_1(1-r_2),\\
&(d)\quad \alpha_1 = 1-r, && \alpha_2 = 0,\\
&\qquad \beta_1 = 0, && \beta_2 = r,\\
&(e)\quad \alpha_1 = (1-r)(1-r_2), && \alpha_2 = rr_2,\\
&\qquad \beta_1 = (1-r)r_2, && \beta_2 = r(1-r_2). \quad (1)
\end{aligned}$$

Then the expected mean values and variances in the alternative groups can be represented as

$$\begin{aligned}
\tilde{\mu}_1 &= \alpha_1\mu^*_{11} + \alpha_2\mu^*_{12} + \beta_2\mu^*_{21} + \beta_1\mu^*_{22},\\
\tilde{\mu}_2 &= \beta_1\mu^*_{11} + \beta_2\mu^*_{12} + \alpha_2\mu^*_{21} + \alpha_1\mu^*_{22},\\
\tilde{\sigma}^2_1 &= \alpha_1\sigma^{*2}_{11} + \alpha_2\sigma^{*2}_{12} + \beta_2\sigma^{*2}_{21} + \beta_1\sigma^{*2}_{22} + G(\alpha_1, \alpha_2, \beta_1, \beta_2),\\
\tilde{\sigma}^2_2 &= \beta_1\sigma^{*2}_{11} + \beta_2\sigma^{*2}_{12} + \alpha_2\sigma^{*2}_{21} + \alpha_1\sigma^{*2}_{22} + G(\beta_1, \beta_2, \alpha_1, \alpha_2), \quad (2)
\end{aligned}$$

TABLE 1

Asymptotic estimates of parameters and max ELODs (mELODs) of a single-QTL model when applied to the case of two linked QTL ($r = 0.25$) for coupling and repulsion phases

Phase	d_1^*	d_2^*	d_1	d_2	σ_1	σ_{2q}	$mELOD_{11}$	$mELOD_{12}$	$mELOD_2$
C		0.5	0.75	0.75			10.11	10.11	
	0.5				0.83	0.83			3.84
R		−0.5	0.25	−0.25			1.22	1.22	
C		0.3	0.86				13.19	—	
	0.7			—	0.81	—			1.41
R		−0.3	0.55				5.92	—	

The results of application of single marker sliding are presented. It appears that the ELOD function may have two local maxima ($mELOD_{11}$ and $mELOD_{12}$) coinciding with the true positions of the QTL (see also Hyne and Kearsey 1995; Wright and Kong 1997), resulting in two sets of parameter estimates. For comparison, we also provide the results of two-QTL sliding, $mELOD_2$ for the hypothesis H_2 (two linked QTL) *vs.* H_1 (single QTL), assuming the sample size $n = 250$. C and R stand for coupling and repulsion phases, respectively; d_1^*, d_2^*, d_1, and d_2, are the true and estimated values of the effects of the first and second QTL, respectively; σ_1 and σ_2 are the estimates of the residual standard deviations (the true value was $\sigma^* = 0.8$). Note that the same values of $mELOD_2$ are presented for the two linkage phases, C and R (for further details see the two-trait analysis section).

where

$$\begin{aligned} G(a, b, c, d) = {} & a(c(\mu_{11}^* - \mu_{22}^*) + d(\mu_{11}^* - \mu_{21}^*))^2 \\ & + b(c(\mu_{12}^* - \mu_{22}^*) + d(\mu_{12}^* - \mu_{21}^*))^2 \\ & + d(a(\mu_{21}^* - \mu_{11}^*) + b(\mu_{21}^* - \mu_{12}^*))^2 \\ & + c(a(\mu_{22}^* - \mu_{11}^*) + b(\mu_{22}^* - \mu_{12}^*))^2 \\ & + a\,b(\mu_{11}^* - \mu_{12}^*)^2(1 + c + d) \\ & + c\,d(\mu_{22}^* - \mu_{21}^*)^2(1 + a + b). \end{aligned}$$

Assuming that our trial marker is tightly linked to (or coincides with) the putative (single) QTL and that the trait distributions in alternative groups are normal, we can calculate the regression of the log-likelihood as a function of parameter set $\theta = (\mu_1,\mu_2,\sigma_1,\sigma_2)$:

$$\begin{aligned} U_1(\theta) = {} & \mathbf{E}\sum_{i=1}^{n_1}\log\{(\sigma_1\sqrt{2\pi})^{-1}\exp[-(x_i - \mu_1)^2/(2\sigma_1^2)]\} \\ & + \mathbf{E}\sum_{i=n_1+1}^{n}\log\{(\sigma_2\sqrt{2\pi})^{-1}\exp[-(x_i - \mu_2)^2/(2\sigma_2^2)]\}, \end{aligned}$$

where **E** stands for expectations. To obtain the asymptotic estimates of parameter values, one can calculate max $U_1(\theta)$:

$$\begin{aligned} \max_\theta U_1(\theta) &= U_1(\tilde\mu_1, \tilde\mu_2, \tilde\sigma_1, \tilde\sigma_2) \\ &= -0.5n(1 + \log(2\pi\tilde\sigma_1\tilde\sigma_2)) = V_1(\tilde\sigma_1, \tilde\sigma_2), \end{aligned}$$

where $\tilde\sigma_1$ and $\tilde\sigma_2$ are as defined in (2).

Clearly, $\tilde\sigma_1$ and $\tilde\sigma_2$ depend on the position l of the trial marker with respect to Q_1/q_1 and Q_2/q_2 loci. For any trial marker position l the conditional max of $U_1(\theta)$ is $V_1(\tilde\sigma_1(l),\tilde\sigma_2(l))$. We now maximize $V_1(\tilde\sigma_1(l),\sigma_2(l))$ with respect to marker position l. To save space we skip the details and provide here only the final results. Our task was to analyze the consequences of application of a single-QTL mapping model to a situation with two linked QTL. In particular, we compare the expected LOD scores corresponding to the hypotheses H_1, "one QTL in the considered chromosome" and H_0, "no QTL in the chromosome." Then, we have consequently

$$\begin{aligned} &\text{for } H_1 \rightarrow \max_l V_1(\tilde\sigma_1(l), \tilde\sigma_2(l)) = V_1\text{; for } H_0 \rightarrow U_0(\mu, \sigma) \\ &\quad = \mathbf{E}\log\prod_{i=1}^{n}(\sigma\sqrt{2\pi})^{-1}\exp(-(x_i - \mu)^2/(2\sigma^2)), \end{aligned}$$

and

$$\begin{aligned} V_0 &= \max_{\mu,\sigma} U_0(\mu, \sigma) \\ &= \mathbf{E}\log\prod_{i=1}^{n}(\hat\sigma\sqrt{2\pi})^{-1}\exp(-(x_i - \hat\mu)^2/(2\hat\sigma^2)), \end{aligned}$$

where

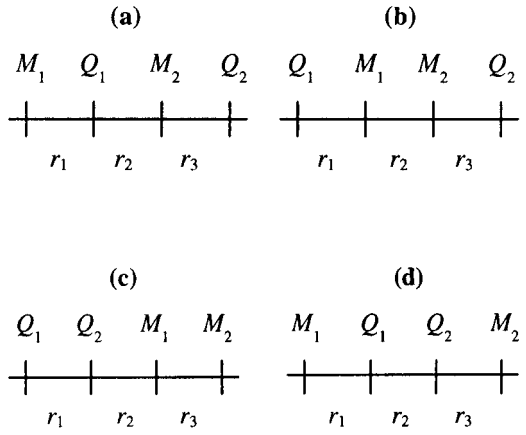
$$\begin{aligned} \hat\mu &= \tfrac{1}{2}(1 - r)\mu_{11}^* + \tfrac{1}{2}r\mu_{12}^* + \tfrac{1}{2}r\mu_{21}^* + \tfrac{1}{2}(1 - r)\mu_{22}^*, \\ \hat\sigma^2 &= \tfrac{1}{2}(1 - r)\sigma_{11}^{*2} + \tfrac{1}{2}r\sigma_{12}^{*2} + \tfrac{1}{2}r\sigma_{21}^{*2} + \tfrac{1}{2}(1 - r)\sigma_{22}^{*2} \\ &\quad + G(\tfrac{1}{2}(1 - r), \tfrac{1}{2}r, \tfrac{1}{2}(1 - r), \tfrac{1}{2}r). \end{aligned}$$

It is easy to show that $V_0 = -0.5n(1 + \log(2\pi) + \log\hat\sigma^2)$ and one can obtain the expression for the maximum expected LOD value, max ELOD $= V_1 - V_0$.

These results allow us to evaluate the consequences of model misspecification. The behavior of the score ELOD $= V_1(\cdot) - V_0$ as a function of trial marker position and the parameters characterizing the effect of Q_1/q_1 and Q_2/q_2 are represented in Table 1 and Figure 1. Clearly, $V_1(\cdot) - V_0$ reaches a local maximum when the marker coincides with one of the QTL. One can easily see from the presented illustrations that the possibility of finding an indication of the existence of two QTL by revealing two local maxima depends on linkage phase (coupling or repulsion), distance between the QTL,

and magnitudes of the QTL effects and their ratio (see Figure 1 and Table 1).

Two linked QTL: *ELOD for testing H_2 vs. H_1:* As before, consider a situation when the target trait x depends on the two linked loci Q_1/q_1 and Q_2/q_2 with normal trait densities $f_{11}(x)$, $f_{12}(x)$, $f_{21}(x)$, $f_{22}(x)$ in the QTL groups $Q_1Q_1Q_2Q_2$, $Q_1Q_1q_2q_2$, $q_1q_1Q_2Q_2$, and $q_1q_1q_2q_2$ of the dihaploid mapping population characterized by unknown means μ_{11}^*, μ_{12}^*, μ_{21}^*, and μ_{22}^*, and (residual) standard deviations σ_{11}^*, σ_{12}^*, μ_{21}^*, and σ_{22}^*. Employment of two markers instead of one allows us to take into account both QTL. Several basic situations of marker loci positioning relative to the QTL could be considered:



Clearly, other possible situations are equivalent to these four, up to a replacement of parameters. In the foregoing single-marker sliding, we had two discrepancies between the model specification and the real situation: (i) only one QTL was assumed, and (ii) the trial marker was treated as if its position coincides with that of the putative QTL. Now the model is improved, because the first assumption is removed. Therefore, we can consider a process of sliding with a pair of markers along the chromosome as a tool to locate the pair of QTL. Such a procedure is equivalent to two-interval mapping analysis (Haley and Knott 1992; Martinez and Curnow 1992; Jansen 1993; Korol *et al.* 1998) with vanishing lengths of the trial intervals. Because of the foregoing assumptions, the true expected densities of the trait distribution in four alternative marker groups for an arbitrary pair of trial markers can be written as

$$\begin{aligned} h_{11}(x) &= \alpha_1 f_{11}(x) + \alpha_2 f_{12}(x) + \beta_2 f_{21}(x) + \beta_1 f_{22}(x) \\ &= H(\alpha_1, \alpha_2, \beta_1, \beta_2), \\ h_{12}(x) &= H(\tau_1, \tau_2, \delta_1, \delta_2), \\ h_{21}(x) &= H(\sigma_1, \delta_2, \tau_1, \tau_2), \\ h_{22}(x) &= H(\beta_1, \beta_2, \alpha_1, \alpha_2), \end{aligned}$$

with corresponding mixture parameters

$$\begin{aligned}
\text{(a)}\ \alpha_1 &= (1-r_1)(1-r_2)(1-r_3)/s_{\alpha\beta}, \\
\alpha_2 &= (1-r_1)(1-r_2)r_3/s_{\alpha\beta}, \\
\beta_1 &= r_1r_2r_3/s_{\alpha\beta}, \\
\beta_2 &= r_1r_2(1-r_3)/s_{\alpha\beta}, \\
s_{\alpha\beta} &= r_1r_2 + (1-r_1)(1-r_2); \\
\tau_1 &= (1-r_1)r_2r_3/s_{\tau\delta}, \\
\tau_2 &= (1-r_1)r_2(1-r_3)/s_{\tau\delta}, \\
\delta_1 &= r_1(1-r_2)(1-r_3)/s_{\tau\delta}, \\
\delta_2 &= r_1(1-r_2)r_3/s_{\tau\delta}, \\
s_{\tau\delta} &= (1-r_1)r_2 + r_1(1-r_2); \\
\text{(b)}\ \alpha_1 &= (1-r_1)(1-r_3), \quad \alpha_2 = (1-r_1)r_3, \\
\beta_1 &= r_1r_3, \quad \beta_2 = r_1(1-r_3), \\
\tau_1 &= (1-r_1)r_3, \quad \tau_2 = (1-r_1)(1-r_3), \\
\delta_1 &= r_1(1-r_3), \quad \delta_2 = r_1r_3; \\
\text{(c)}\ \alpha_1 &= (1-r_1)(1-r_2), \quad \alpha_2 = r_1r_2, \\
\beta_1 &= (1-r_1)r_2, \quad \beta_2 = r_1(1-r_2), \\
\tau_1 &= (1-r_1)(1-r_2), \quad \tau_2 = r_1r_2, \\
\delta_1 &= (1-r_1)r_2, \quad \delta_2 = r_1(1-r_2); \\
\text{(d)}\ \alpha_1 &= (1-r_1)(1-r_2)(1-r_3)/s_{\alpha\beta}, \\
\alpha_2 &= (1-r_1)r_2r_3/s_{\alpha\beta}, \\
\beta_1 &= r_1(1-r_2)r_3/s_{\alpha\beta}, \\
\beta_2 &= r_1r_2(1-r_3)/s_{\alpha\beta}, \\
s_{\alpha\beta} &= (1-r_1)(1-r_2)(1-r_3) \\
&\quad + (1-r_1)r_2r_3 + r_1r_2(1-r_3) \\
&\quad + r_1(1-r_2)r_3, \\
\tau_1 &= (1-r_1)(1-r_2)r_3/s_{\tau\delta}, \\
\tau_2 &= (1-r_1)r_2(1-r_3)/s_{\tau\delta}, \\
\delta_1 &= r_1(1-r_2)(1-r_3)/s_{\tau\delta}, \\
\delta_2 &= r_1r_2r_3/s_{\tau\delta}, \\
s_{\tau\delta} &= (1-r_1)(1-r_2)r_3 \\
&\quad + (1-r_1)r_2(1-r_3) + r_1r_2r_3 \\
&\quad + r_1(1-r_2)(1-r_3). \qquad (3)
\end{aligned}$$

For any pair of markers, one can assume that they coincide with (or are closely linked to) the corresponding QTL. If so, the parameter values characterizing these QTL are easily derived by employing maximization of regression of the log-likelihood function analogous to the procedure described in the previous section. Consequently, one can calculate, for the current pair of markers, the expected LOD assuming two linked QTL (H_2 hypothesis), which can be compared to the expected log-likelihood obtained under the assumption of one QTL (H_1). We found that given independent variance effects of the linked QTL, the maximum of ELOD over possible locations (l_1,l_2) of the trial pairs of markers is attained exactly in the case when these locations coin-

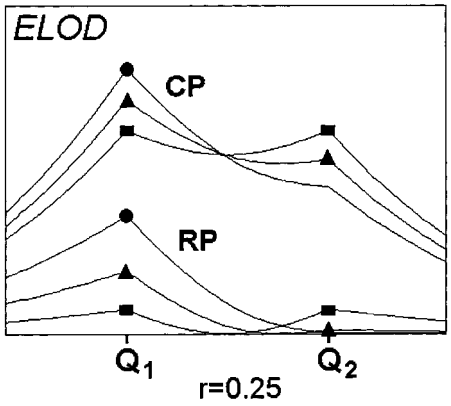


Figure 1.—Behavior of ELOD for a single-marker sliding in the single-QTL model when applied to situations with two linked QTL, as a function of the QTL positions and effects. CP and RP, coupling and repulsion phases. The effects of the QTL are denoted by the following: ■, $d_1^* = 0.5$ and $d_2^* = 0.5$; ▲, $d_1^* = 0.6$ and $d_2^* = 0.4$; ●, $d_1^* = 0.7$ and $d_2^* = 0.3$. The residual standard deviation was 0.8.

cide with those of the QTL (the proof is available from authors upon request). The surface ELOD = ELOD(l_1,l_2) represented in Figure 2 manifests an important asymptotic property of interval QTL mapping with vanishing interval length and increasing sample size: a faster-than-linear growth of the criterion when approaching the true position of the QTL (the second derivatives are positive). Note that the same max ELOD is predicted for coupling and repulsion phases.

Comparison of the analytical and simulation results: The foregoing model allows us to deduce the expected LOD values in the QTL mapping analysis in the case of two linked QTL. However, these results are essentially asymptotic and may be biased at small samples. Therefore, it is important to assess how the obtained estimates converge to the expected parameter values when the sample sizes and marker density are increasing. To do that we employed Monte Carlo simulations. Chromosomes with two linked QTL were modeled for two population sizes (n = 500 and 2000). No crossing-over interference (Haldane mapping function) was assumed. For each sample, we employed two subsets of markers, using the information on intervals 12 and 48. Table 2 shows the behavior of the average LOD values and the discrepancy between the estimated and simulated QTL positions as dependent on sample size and number of markers. The main conclusion from the simulations is that the proposed method can indeed serve as a basis to get an approximate prediction of the expected LOD for interval mapping of two linked QTL (compare the average max LODs with max ELODs).

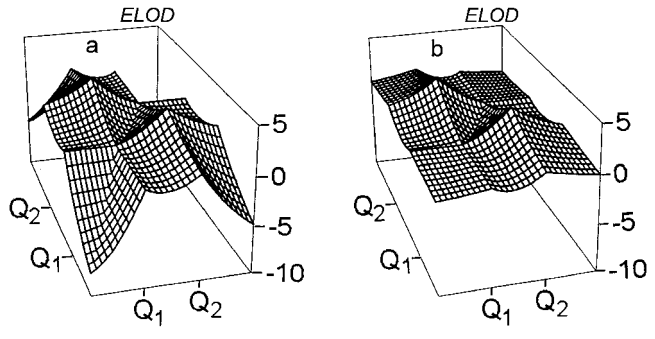


Figure 2.—The ELOD surface for the alternative H_2 (two linked QTL) *vs.* H_1 (one QTL in the chromosome). (a) Coupling phase; (b) repulsion phase.

It follows from the presented results that the difference between predicted max ELOD and the averaged over simulations max LOD in repulsion phase is smaller than that in coupling phase, in spite of the fact that our theory predicts the same value for the two phases. In both cases the experimental LODs are smaller than the predicted ones; *i.e.*, for the same combinations of parameters the simulated LODs were higher in repulsion phase. A simple explanation can be proposed for this effect. The simulated procedure includes analyses for two hypotheses, H_1 and H_2. In Monte Carlo experiments with two linked QTL, we can consider two options for fitting parameters of the maximum-likelihood function to the H_1 hypothesis (Lander and Botstein 1989; Haley and Knott 1992; Korol *et al.* 1998): (i) *fixed position* of the putative QTL, when its position is assumed to be known and coincides with either of the two simulated positions (which would not necessarily be true in the practical data analysis when these positions are unknown); (ii) *variable position* of the putative QTL that is assumed unknown, but can be found because it provides maximum value of the maximum-likelihood function. Certainly, the achievable maximum is higher in the second situation resulting in an underestimation of the LOD value for H_2 *vs.* H_1 (not shown).

Applications: The proposed analytical tool allows us to evaluate easily, without the necessity of Monte Carlo simulations, the behavior of the ELOD values across all possible locations of the putative QTL, for any fixed sets of parameters (see Figure 2), which is important for designing mapping experiments. For example, using the obtained expression of max ELOD, we can get an estimate of the *minimum sample size* needed to discriminate between H_1 and H_2, when H_2 is true (*i.e.*, when we have a pair of linked QTL with some effects d_1 and d_2), with a certain preset test power. This is based on the fact that the expected LOD value is distributed as noncentral chi-square with degrees of freedom equal to the difference in the number of parameters specifying the alternatives (H_2 and H_1) (Wald 1943). This tool enables us to compare different practical situations with respect to the foregoing prediction of the minimum sample size (see Lander and Botstein 1989). The usefulness of such an option is especially obvious for mapping of linked QTL, where the efficiency of the experimental design depends on many factors characterizing the un-

known "configuration" of the problem: the distance of the putative QTL, their relative effects on trait mean value and variance, linkage phase (coupling *vs.* repulsion), and presence or absence of epistatic interaction, etc. We now consider two examples to illustrate the possibilities of the proposed analysis: the dependence of ELOD for H_2 *vs.* H_1 on epistasis and the detectability of epistasis provided H_2 is already proved.

The effect of epistatic interaction on ELOD for QTL detection: In the example on epistatic interactions the trait values in the four QTL classes were modeled as:

$$\mu_{11} = \mu_0 + 0.5(d_1 + d_2) + \varepsilon,$$
$$\mu_{21} = \mu_0 - 0.5(d_1 - d_2) - \varepsilon,$$
$$\mu_{12} = \mu_0 + 0.5(d_1 - d_2) - \varepsilon,$$
$$\mu_{22} = \mu_0 - 0.5(d_1 + d_2) + \varepsilon,$$

where ε is the epistatic effect. It is of high practical importance to predict the expected power of detection of epistatic interactions within the framework of QTL mapping analysis (Haley and Knott 1992; Eaves 1994; Korol *et al.* 1994; Fu and Ritland 1996). This problem is addressed in the next section. Here we consider first the question of how the presence of epistasis may affect the power of QTL detection. The fitted single-QTL model of two-QTL data may depend on the peculiarities of the effects and positions of the involved QTL. Clearly, if epistasis is present it also may affect the results of such a fitting. Then, when fitting H_2 for data with epistasis, one can either adopt or ignore epistasis, which may affect the final result. Figure 3 shows the behavior of max ELOD for testing H_2 *vs.* H_1 (testing for the presence of two linked QTL as compared to one QTL). Two

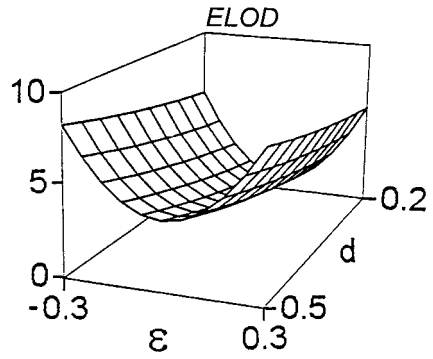


Figure 3.—The behavior of the maximum ELOD (H_2 *vs.* H_1) as a function of the effects of the QTL (d) and the level of epistasis (ε). Here d is the effect of the first QTL and we assumed that the effect of the second QTL also varies, in such a manner that the sum of the absolute values of both effects is equal to 1. The numerical values of the parameters were $n = 250$, $r = 25\%$, and $\sigma = 0.8$.

TABLE 2

Comparison of the average LOD values (over 200 Monte Carlo runs) with the asymptotic values of max ELOD for comparing the hypotheses H_2 (two linked QTL in the chromosome) *vs.* H_1 (one QTL in the choromosome)

			Number of intervals					
					12		48	
			12	48	Average estimated positions of the QTL			
n	max ELOD	Phase	Average LOD value		L_1	L_2	L_1	L_2
500	7.68	C	5.26	5.88	14.37	49.66	14.10	49.47
			(1.46)	(1.49)	(3.36)	(3.22)	(2.38)	(2.52)
		R	6.64	7.15	14.80	49.48	14.29	49.49
			(2.17)	(2.25)	(3.64)	(3.50)	(2.60)	(2.66)
2000	30.72	C	23.46	26.84	14.29	49.15	14.29	49.02
			(3.55)	(3.61)	(1.54)	(1.12)	(0.56)	(0.62)
		R	27.33	29.07	14.72	49.12	14.41	49.19
			(4.64)	(4.73)	(1.59)	(1.18)	(0.70)	(0.73)

Monte Carlo simulations (200 runs) were conducted for doubled haploid mapping populations with n = 500 and 2000 genotypes. A single chromosome was modeled with two QTL (Q_1/q_1 and Q_2/q_2 of equal additive effects ($d_1 = d_2 = 0.5$) in coupling (C) and repulsion (R) phases; the residual standard deviations were all 0.8. Chromosomes with 49 equidistant markers with a recombination rate of 0.02 between adjacent markers were simulated assuming no interference (hence Haldane mapping function). The simulated positions of the QTL coincided with markers 8 and 25. For the considered two marker densities (*i.e.*, with 12 and 48 intervals), correspondingly 13 and 49 markers were used in two-interval analysis. Thus, the simulated positions of the QTL were $L_1 = 14.29$ cM and $L_2 = 48.98$ cM. For both LOD values and estimated positions of the QTL, we provide the averages (first rows) and standard deviations (second rows in parentheses). The hypothesis H_2 (two linked QTL) was compared to H_1 (a single QTL) by fitting the H_1 model with no presumed constraints on the position of the putative single QTL; correspondingly, the two-QTL model was fitted by scanning over all possible pairs of intervals (see Halley and Knott 1992; Korol *et al.* 1998).

TABLE 3

Comparison of the average LOD values (over 100 Monte Carlo runs) with the asymptotic values of max ELOD for comparing the hypotheses H_2 (two linked QTL with epistasis), H_2 (two linked QTL with no epistasis), and H_0 (no QTL in the chromosome)

n		H_2 ($\varepsilon \neq 0$) *vs.* H_0	H_2 ($\varepsilon = 0$) *vs.* H_0	H_2 ($\varepsilon \neq 0$) *vs.* H_2 ($\varepsilon = 0$)	Power ($\alpha = 5\%$) H_2 ($\varepsilon \neq 0$) *vs.* H_2 ($\varepsilon = 0$)
250	*s*	16.25	15.06	1.19	55%
	t	14.71	13.74	0.97	
	$s - t$	1.54	1.32	0.22	
500	*s*	30.48	28.45	2.02	76%
	t	29.42	27.48	1.94	
	$s - t$	1.06	0.97	0.08	

A situation with two linked epistatically interacting QTL was considered for coupling phase. The positions, the additive effects, and the residual variances of the QTL are as described in Table 2. The epistatic effect was $\varepsilon = 0.125$. Here *s* denotes the average LOD resulting from simulation experiments whereas *t* denotes the predicted max ELOD for the compared alternatives H_2 ($\varepsilon \neq 0$) *vs.* H_2 ($\varepsilon = 0$) (d.f. = 1).

conclusions can be derived from our analysis: (i) the effect of epistasis on detection power is symmetric with respect to the sign of ε, and (ii) epistasis may have either a positive or negative effect on max ELOD of the test of H_1 *vs.* H_0 (not shown) and always a positive effect when testing H_2 *vs.* H_1 (Figure 3; see also Eaves 1994).

The detectability of epistasis (comparison of H_2 under $\varepsilon \neq 0$ vs. H_2 under $\varepsilon = 0$): In the foregoing section we could see how epistasis affects the expected LOD values when single-QTL and two-QTL models are applied to the analysis. The proposed tool allows us also to predict the expected LOD for the situation when one wants to contrast two versions of the hypothesis H_2 (two linked QTL in the chromosome): H_2 ($\varepsilon = 0$), *i.e.*, additive effects of the QTL, and H_2 ($\varepsilon \neq 0$), *i.e.*, assuming epistasis. Testing for epistasis (coadaptation) and evaluating the magnitude of epistasis have recently become an important component of QTL mapping analysis (Doebley *et al.* 1995; Li *et al.* 1997). This meaningful subject has a long history in both evolutionary genetics (Dobzhansky 1970; Wright 1977), theories of recombination and sex evolution (Barton 1995; Otto and Feldman 1997), and agricultural genetics (Yu *et al.* 1997). However, only with QTL mapping can epistatic effects be objectively detected and evaluated. Each of the foregoing alternative versions of H_2, without and with epistasis, can be compared to H_0 (no QTL in the chromosome), using the proposed approximation. The difference between the resulting max ELODs will give us max ELOD for the presence of epistasis. An example presented in Table 3 illustrates the closeness between the predicted LOD values and the average LODs obtained in direct Monte Carlo simulations.

TWO-TRAIT ANALYSIS

As was shown in our previous simulation study (Korol *et al.* 1998), joint analysis of correlated quantitative traits may increase the mapping resolution in situations with linked QTL, *i.e.*, when H_2 (two linked QTL) and H_1 (one QTL) are compared. The higher the residual correlation the better the expected LOD. In two-trait analysis, the residual correlation between the traits in the QTL groups may be caused by nongenetic mechanisms, pleiotropy, or linkage of genes from other chromosomes affecting either of the traits, and by pleiotropy and linkage of genes from the chromosome under consideration.

As in single-trait analysis, to analyze two-QTL situations we calculate max ELOD for the alternative hypotheses: H_2 *vs.* H_1. This means that we need to develop bivariate analogues of the foregoing single-QTL and two-QTL models based on single- and two-marker sliding procedures. Hence, the goal of the first part of this section is to obtain the regression of the log-likelihood function assuming that only one QTL resides in the chromosome that in fact carries two linked QTL. Let the traits *x* and *y* be dependent on two loci, Q_1/q_1 and Q_2/q_2, residing in the marked chromosome and let the bivariate trait distributions in the four QTL groups $Q_1Q_1Q_2Q_2$, $Q_1Q_1q_2q_2$, $q_1q_1Q_2Q_2$, and $q_1q_1q_2q_2$ of dihaploid mapping population be normal densities $f_{11}(x, y)$, $f_{12}(x, y)$, $f_{21}(x, y)$, and $f_{22}(x, y)$ with unknown vectors of means $\{\mu_x^*\} = \{\mu_{11x}^*, \mu_{12x}^*, \mu_{21x}^*, \mu_{22x}^*\}$ and $\{\mu_y^*\} = \{\mu_{11y}^*, \mu_{12y}^*, \mu_{21y}^*, \mu_{22y}^*\}$, residual standard deviations $\{\sigma_x^*\} = \{\sigma_{11x}^*, \sigma_{12x}^*, \sigma_{21x}^*, \sigma_{22x}^*\}$ and $\{\sigma_y^*\} = \{\sigma_{11y}^*, \sigma_{12y}^*, \sigma_{21y}^*, \sigma_{22y}^*\}$, and correlations $\{\rho_{xy}^*\} = (\rho_{11}^*, \rho_{12}^*, \rho_{21}^*, \rho_{22}^*)$, respectively. Usually, the mapping procedure is started with the assumption of one QTL in the chromosome, and then one could try to apply some versions of single-marker or interval analysis. Reduced test power, biased parameter estimates, and detection of ghost factors may result from this simplification, as demonstrated by many simulation studies. The foregoing analytical treatment of this problem developed for a single-trait analysis is expanded now to two-trait analysis.

Consider a random sample of individuals each characterized for traits x and y and a set of marker loci from the chromosome in question. For an arbitrary marker, we take into account the same five situations (a–e) as those considered above for the single-trait analysis. Then the true expected densities of the bivariate trait distribution in the alternative marker groups for an arbitrary scanning marker will be

$$\begin{aligned} h_{MM}(x, y) &= h_1 = \alpha_1 f_{11}(x, y) + \alpha_2 f_{12}(x, y) \\ &\quad + \beta_2 f_{21}(x, y) + \beta_1 f_{22}(x, y), \\ h_{mm}(x, y) &= h_2 = \beta_1 f_{11}(x, y) + \beta_2 f_{12}(x, y) \\ &\quad + \alpha_2 f_{21}(x, y) + \alpha_1 f_{22}(x, y), \end{aligned}$$

where α_i and β_j are as defined in (1). To proceed with the analysis we need to make the following note.

Consider an arbitrary bivariate distribution with finite central moments (up to the fourth). Then the maximum of log-likelihood per individual for the Gaussian model will converge in probability to the maximum of the regression of the log-likelihood function per individual. Assume that the trial marker is exactly at the same position as our putative QTL and the trait distributions in the alternative groups *MM* and *mm* are bivariate normals. The regression of the log-likelihood will take the form

$$\begin{aligned} U_{1xy}(\theta) = \mathbf{E} \log\Big(&\prod_{i=1}^{n_1} N(x_{1i}, \mu_{1x}, \mu_{1y}, \sigma_{1x}, \sigma_{1y}, \rho_1) \\ &\times \prod_{j=n_1+1}^{n} N(x_{2j}, \mu_{2x}, \mu_{2y}, \sigma_{2x}, \sigma_{2y}, \rho_2)\Big), \end{aligned}$$

where $\theta = (\mu_{1x}, \mu_{2x}, \mu_{1y}, \mu_{2y}, \sigma_{1x}, \sigma_{2x}, \sigma_{1y}, \sigma_{2y}, \rho_1, \rho_2)$. Then,

$$\begin{aligned} \max_{\theta} U_{1xy}(\theta) &= -n\Big(1 + \tfrac{1}{2}\log\big(4\pi^2 \tilde{\sigma}_{1x}\tilde{\sigma}_{2x}\tilde{\sigma}_{1y}\tilde{\sigma}_{2y} \\ &\quad \times \sqrt{(1 - \tilde{\rho}_1^2)(1 - \tilde{\rho}_2^2)}\big)\Big) \\ &= V_{1sy}(\tilde{\sigma}_{1x}, \tilde{\sigma}_{2x}, \tilde{\sigma}_{1y}, \tilde{\sigma}_{2y}, \tilde{\rho}_1, \tilde{\rho}_2), \end{aligned}$$

where $\tilde{\sigma}_{iu}$ ($u = x$ or y) and $\tilde{\rho}_i$ ($i = 1,2$) are some functions of the main parameters.

Consider now a process of scanning with a pair of markers along the chromosome. Because of the foregoing assumptions, the true expected densities of the trait distribution in four alternative marker groups for an arbitrary pair of trial markers can be written as

$$\begin{aligned} h_{11}(x, y) &= \alpha_1 f_{11}(x, y) + \alpha_2 f_{12}(x, y) + \beta_2 f_{21}(x, y) \\ &\quad + \beta_1 f_{22}(x, y) = H(\alpha_1, \alpha_2, \beta_1, \beta_2), \\ h_{12}(x, y) &= H(\tau_1, \tau_2, \delta_1, \delta_2), \quad h_{21}(x, y) = H(\delta_1, \delta_2, \tau_1, \tau_2), \\ h_{22}(x, y) &= H(\beta_1, \beta_2, \alpha_1, \alpha_2), \end{aligned}$$

where the mixture parameters α_i, β_i, τ_i, and δ_i ($i = 1, 2$) are as defined in (3). For any pair of markers, one can assume that they coincide with (or are closely linked to) the corresponding QTL. If so, the parameter values characterizing these QTL are easily derived by employing maximization of the regression of the log-likelihood function. The regression of the log-likelihood now looks like

$$\begin{aligned} U_{2xy}(\{\mu_x\}, \{\sigma_x\}, \{\mu_y\}, \{\sigma_y\}, \{\rho_{xy}\}) &= U_{2xy}(\cdot) \\ = \mathbf{E} \log\Big(&\prod_{j,k=1}^{2} \prod_{i_{jk}=1}^{n_{jk}} N(x_{ijk}, \mu_{jkx}, \mu_{jky}, \sigma_{jkx}, \sigma_{jky}, \rho_{jk})\Big), \end{aligned}$$

where $N(\cdot)$ is a bivariate normal density, $\{\mu_x\} = \{\mu_{11x}, \mu_{12x}, \mu_{21x}, \mu_{22x}\}$, $\{\mu_y\} = (\mu_{11y}, \mu_{12y}, \mu_{21y}, \mu_{22y})$, $\{\sigma_x\} = (\sigma_{11x}, \sigma_{12x}, \sigma_{21x}, \sigma_{22x})$, $\{\sigma_y\} = (\sigma_{11y}, \sigma_{12y}, \sigma_{21y}, \sigma_{22y})$, $\{\rho_{xy}\} = (\rho_{11}, \rho_{12}, \rho_{21}, \rho_{22})$, and n_{kl} are the frequencies of the four marker classes for the current pairs of markers, $n_{11} + n_{12} + n_{21} + n_{22} = n$. Then,

$$\begin{aligned} \max U_{2xy}(\cdot) &= U_{2xy}(\{\tilde{\mu}_x\}, \{\tilde{\sigma}_x\}, \{\tilde{\mu}_y\}, \{\tilde{\sigma}_y\}, \{\tilde{\rho}_{xy}\}) \\ &= -n(1 + \log 2\pi) \\ &\quad + \tfrac{1}{2}\Big((1 - \tilde{r})\log\big(\tilde{\sigma}_{11x}\tilde{\sigma}_{11y}\tilde{\sigma}_{22x}\tilde{\sigma}_{22y} \\ &\quad \times \sqrt{(1 - \tilde{\rho}_{11}^2)(1 - \tilde{\rho}_{22}^2)}\big) \\ &\quad + \tilde{r}\log\big(\tilde{\sigma}_{12x}\tilde{\sigma}_{12y}\tilde{\sigma}_{21x}\tilde{\sigma}_{22y}\sqrt{(1 - \tilde{\rho}_{12}^2)(1 - \tilde{\rho}_{21}^2)}\big)\Big) \\ &= V_{2xy}(\tilde{r}, \{\tilde{\sigma}_x\}, \{\tilde{\sigma}_y\}, \{\rho_{xy}\}) = V_{2xy}(\cdot), \end{aligned}$$

where the components of vector $(\{\tilde{\mu}_x\},\{\tilde{\sigma}_x\},\{\tilde{\mu}_y\},\{\tilde{\sigma}_y\},\{\tilde{\rho}_{xy}\})$ and $\tilde{r}$ are calculated routinely. This presentation allows us to obtain the following results.

Assume independent variance effects of the linked QTL for each of the traits (*i.e.*, $\sigma^*_{11x} = \sigma^*_x$, $\sigma^*_{12x} = a_x\sigma^*_x$, $\sigma^*_{21x} = b_x\sigma^*_x$, $\sigma^*_{22x} = a_xb_x\sigma^*_x$, $\sigma^*_{11y} = \sigma^*_y$, $\sigma^*_{12y} = a_y\sigma^*_y$, $\sigma^*_{21y} = b_y\sigma^*_y$, $\sigma^*_{22y} = a_yb_y\sigma^*_y$) and equal residual correlation between the target traits in all of the four QTL groups (*i.e.*, $\rho^*_{ij} = \rho^*$; $i, j = 1.2$). Then, if one of the traits, x, depends additively on both linked QTL (Q_1/q_1 and Q_2/q_2), whereas the correlated trait y is independent of the considered QTL, then the global maximum of $V_{2xy}(\cdot)$ over all possible locations of the trial pair of markers in the chromosome is attained exactly when these locations coincide with those of the QTL. Likewise, if each of the two linked QTL (i) affects one and only one of the traits or (ii) has a pleiotropic effect on both traits, x and y, but in such a manner that $d_{1x}/d_{2x} = d_{1y}/d_{2y} = c$, then the maximum of $V_{2xy}(\cdot)$ over possible locations of the trial pairs of markers in a sufficiently small neighborhood of the QTL is attained exactly when these locations coincide with those of the QTL (the proof is available from the authors upon request).

To illustrate how the proposed model works we now address two questions concerning the dependence of the expected LOD value for discrimination between H_1 and H_2 on (i) the residual correlation between the traits and (ii) epistasis. Let us fix the effect of one of the QTL (say Q_1/q_1) and consider how max ELOD depends on the effect of the second QTL (Q_2/q_2) and on the residual correlation (ρ) between the quantitative traits. We are interested here in testing H_2 *vs.* H_1, assuming additive effects of the two QTL. Two situations are considered: in the first, each trait depends on only one of the linked

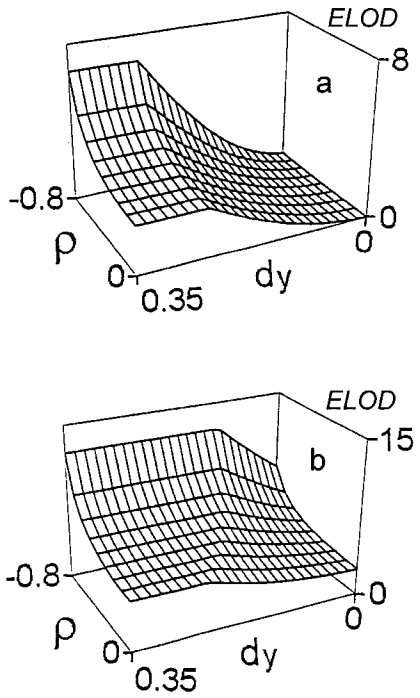


Figure 4.—The behavior of the maximum ELOD (H_2 *vs.* H_1) in the two-trait model with two linked QTL as a function of the residual correlation and the effect of one of the QTL on one of the two traits (y). (a) A situation with the first QTL affecting the trait x ($d_{1x} = 0.25$), whereas the second QTL affects the trait y. (b) The first QTL affects only the trait x ($d_{1x} = 0.35$), whereas the second QTL affects both traits ($d_{2x} = 0.25$; d_{2y} varies, as shown in the figure). (a and b) $r = 25\%$, $\sigma = 0.5$, and $n = 250$.

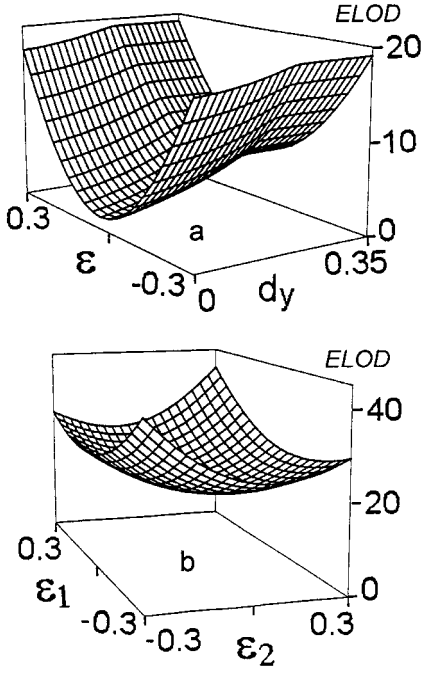


Figure 5.—The behavior of the maximum ELOD (H_2 *vs.* H_1) in the two-trait model with two linked QTL as a function of the parameters. In (a), the first QTL affects only the trait x ($d_{1x} = 0.35$), whereas the second QTL affects both traits: $d_{2x} = 0.15$, and d_{2y} varies, as shown in the figure. The residual correlation was $\rho = -0.5$; the variables ε and $d_y = d_{2y}$ denote the epistatic interaction between the QTL for the trait x and the effect of the second QTL on trait y, respectively. In b, both QTL affect both traits (all effects were equal to 0.35). The residual correlation was $\rho = -0.5$. (a and b) $r = 25\%$, $\sigma = 0.5$, and $n = 250$.

QTL (Figure 4a), whereas in the second case both QTL affect the first trait and one of the QTL affects the second trait (Figure 4b). One can conclude that the detection power increases with the residual correlation between the analyzed traits. An additional conclusion is that the power increases with the effect of Q_2/q_2 up to some "saturation" point. In the first situation the saturation is reached when the effect of Q_2/q_2 becomes equal to that of Q_1/q_1 (see Figure 4a). The only difference in the second situation is that the saturation point depends on ρ: the larger abs(ρ) the earlier (at lower effects of Q_2/q_2) the saturation (see Figure 4b). For the second situation, let us consider the complication caused by epistasis. Namely, we allow for epistatic interaction between the QTL with respect to the trait x. As in the foregoing example on single-trait analysis, it is interesting here to evaluate how epistasis affects the expected detection power. Figure 5a demonstrates that epistasis may be helpful in discriminating between H_2 (two linked QTLs) and H_1 (only a single QTL in the chromosome). This effect is manifested for both positive and negative residual correlations, but the sign of ρ is important in determining the details of the behavior of max ELOD as a function of epistasis and the effect of Q_2/q_2 on trait y (not shown). The second example (Figure 5b) demonstrates a situation in which the QTL interact epistatically for both traits. Clearly, the provided examples are not more than illustrations of the possibilities of the proposed analytical tool. Each of the questions discussed in these illustrations can be dealt with in necessary detail.

CONCLUSION

Resolution power of mapping analysis of linked QTL remains a difficult problem, which was previously addressed mainly in terms of Monte Carlo simulations. This has restricted the possibilities of detailed evaluation and comparison of different mapping situations and experimental designs. The proposed analytical method of evaluating the expected LOD generalizes for the case of two linked QTL the corresponding estimates derived by Lander and Botstein (1989), Mackinnon and Weller (1995), and Mackinnon *et al.* (1996) (referred to as "deterministic sampling"). Our model allows us to analyze situations with variance effect and epistatic interaction between the putative QTL. We developed here also a two-locus analogue of our previous analytical predictor (see Korol *et al.* 1995) of the expected LOD for two-trait mapping analysis. And again, many compli-

cated questions can be addressed, like dependence of the QTL detection power on residual correlation between the traits, accounting of epistatic interaction between the QTL for one or both traits, and the influence of variance effect for one or both traits on ELOD. Although this method gives only an asymptotic estimation of ELOD, it allows one to get an approximate assessment of a broad spectrum of specific mapping situations. Clearly, any asymptotic effect found by the proposed tool can (and should) be checked by Monte Carlo simulations for given sample sizes. Our comparisons made for a series of situations indeed show a good correspondence between the predicted max ELODs and LOD values averaged over Monte Carlo runs. An important point is that our results prove (for the considered class of situations) the important theoretical fact of consistency of interval mapping analysis with two linked QTL (convergence of the parameter estimates to the true values with increasing sample size and vanishing interval length).

We thank Z-B. Zeng for valuable comments. Two anonymous reviewers provided useful criticisms and suggestions on the earlier version of the manuscript. This research was supported by the Israeli Ministry of Absorption and Ministry of Science.

LITERATURE CITED

Barton, N. H., 1995 A general model for the evolution of recombination. Genet. Res. **65:** 123–144.

Boehnke, M., and P. Moll, 1989 Identifying pedigrees segregating at a major locus for a quantitative trait: an efficient strategy for linkage analysis. Am. J. Hum. Genet. **44:** 216–224.

Darvasi, A., and M. Soller, 1992 Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. Theor. Appl. Genet. **85:** 353–359.

Darvasi, A., and M. Soller, 1994 Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. Genetics **138:** 1365–1373.

DeVicente, M. C., and S. D. Tanksley, 1993 QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics **134:** 585–596.

Dobzhansky, T., 1970 *Genetics of the Evolutionary Process*. Columbia University Press, New York.

Doebley, J., A. Stec and C. Gustus, 1995 teosinte brandched1 and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics **141:** 333–346.

Eaves, L. J., 1994 Effect of genetic architecture on the power of human linkage studies to resolve the contribution of quantitative trait loci. Heredity **72:** 175–192.

Fu, Y.-B., and K. Ritland, 1996 Marker-based inferences about epistasis for genes influencing inbreeding depression. Genetics **144:** 339–348.

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.

Hyne, V., and M. J. Kearsey, 1995 QTL analysis: further uses of 'marker regression.' Theor. Appl. Genet. **91:** 471–476.

Jansen, R. C., 1993 Interval mapping of multiple quantitative trait loci. Genetics **135:** 205–211.

Jansen, R. C., and P. Stam, 1994 High resolution of quantitative traits into multiple loci via interval mapping. Genetics **136:** 1447–1455.

Jiang, C., and Z.-B. Zeng, 1995 Multiple trait analysis and genetic mapping for quantitative trait loci. Genetics **140:** 1111–1127.

Knott, S. A., and C. S. Haley, 1992 Aspects of maximum likelihood methods for mapping of quantitative trait loci in line crosses. Genet. Res. **60:** 139–151.

Korol, A. B., I. A. Preygel and N. I. Bocharnikova, 1987 Linkage between loci of quantitative traits and marker loci: 5. Simultaneous analysis of a set of marker and quantitative traits. Genetika (USSR) **23:** 1421–1431.

Korol, A. B., I. A. Preygel and S. I. Preygel, 1994 *Recombination Variability and Evolution*. Chapman & Hall, London.

Korol, A. B., Y. I. Ronin and V. M. Kirzhner, 1995 Interval mapping of quantitative trait loci employing correlated trait complexes. Genetics **140:** 1137–1147.

Korol, A. B., Y. I. Ronin, E. Nevo and P. Hayes, 1998 Multi-interval mapping of correlated trait complexes: simulation analysis and evidence from barley. Heredity **80:** 273–284.

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

Lebowitz, R. J., M. Soller and J. S. Beckmann, 1987 Trait-based analyses for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines. Theor. Appl. Genet. **73:** 556–562.

Li, Z., S. R. Pinson, W. D. Park, A. H. Paterson and J. W. Stansel, 1997 Epistasis for three grain yield components in rice (Oryza sativa L.). Genetics **145:** 453–465.

Luo, Z. W., and M. J. Kearsey, 1992 Interval mapping of quantitative trait loci in an F2 population. Heredity **69:** 236–242.

Mackinnon, M. J., and J. I. Weller, 1995 Methodology and accuracy of estimation of quantitative trait loci parameters in a half-sib design using maximum likelihood. Genetics **141:** 755–770.

Mackinnon, M. J., S. van der Beek and B. P. Kinghorn, 1996 Use of deterministic sampling for exploring likelihoods in linkage analysis of quantitative traits. Theor. Appl. Genet. **92:** 130–139.

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

Motro, U., and M. Soller, 1993 Sequential sampling in determining linkage between marker loci and quantitative trait loci. Theor. Appl. Genet. **85:** 658–664.

Otto, S. P., and M. W. Feldman, 1997 Deleterious mutations, variable epistatic interactions, and the evolution of recombination. Theor. Pop. Biol. **51:** 134–147.

Ronin, Y. I., V. M. Kirzhner and A. B. Korol, 1995 Linkage between loci of quantitative traits and marker loci: multi-trait analysis with a single marker. Theor. Appl. Genet. **90:** 776–786.

Ronin, Y. I., A. B. Korol and J. J. Weller, 1998 Selective genotyping to detect quantitative trait loci affecting multiple traits: interval mapping analysis. Theor. Appl. Genet. (in press).

Soller, M., and J. S. Beckmann, 1990 Marker-based mapping of quantitative trait loci using replicated progeny. Theor. Appl. Genetics **80:** 205–208.

Wald, A., 1943 Tests of statistical hypotheses concerning several parameters when the number of observations is large. Trans. Am. Math. Soc. **54:** 426–482.

Weller, J. I., J. Z. Song, Y. I. Ronin and A. B. Korol, 1997 Experimental designs and solutions to multiple trait comparisons. Anim. Biotechnol. **8:** 107–122.

Wright, S., 1977 Experimental results and evolutionary deductions, in *Evolution and the Genetics of Populations*, Vol. 3. University of Chicago Press, Chicago.

Wright, F. A., and A. Kong, 1997 Linkage mapping in experimental crosses: the robustness of single-gene models. Genetics **146:** 417–425.

Yu, S. B., J. X. Li, C. G. Xu, Y. F. Tan, Y. J. Gao *et al.*, 1997 Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. Proc. Natl. Acad. Sci. USA. **94:** 9226–9231.

Zeng, Z-B., 1994 Precise mapping of quantitative trait loci. Genetics **136:** 1457–1468.

Zeng, Z-B., 1997 Combining information from data in mapping analysis: use of multiple markers and multiple traits. Anim. Biotechnol. **8:** 145–150.

Communicating editor: Z-B. Zeng