

Using Molecular Markers to Estimate Quantitative Trait Locus Parameters: Power and Genetic Variances for Unreplicated and Replicated Progeny

S. J. Knapp* and W. C. Bridges†

*Department of Crop Science, Oregon State University, Corvallis, Oregon 97731, and †Department of Experimental Statistics, Clemson University, Clemson, South Carolina 29634

Manuscript received February 5, 1990
Accepted for publication August 6, 1990

ABSTRACT

Many of the progeny types used to estimate quantitative trait locus (QTL) parameters can be replicated, *e.g.*, recombinant inbred, doubled haploid, and F_3 lines. These parameters are estimated using molecular markers or QTL genotypes estimated from molecular markers as independent variables. Experiment designs for replicated progeny are functions of the number of replications per line (r) and the number of replications per QTL genotype (n). The value of n is determined by the size of the progeny population (N), the progeny type, and the number of simultaneously estimated QTL parameters ($q - 1$). Power for testing hypotheses about means of QTL genotypes is increased by increasing r and n , but the effects of these factors have not been quantified. In this paper, we describe how power is affected by r , n , and other factors. The genetic variance between lines nested in QTL genotypes ($\sigma_{n;q}^2$) is the fraction of the genetic variance between lines (σ_n^2) which is not explained by simultaneously estimated intralocus and interlocus QTL parameters (ϕ_Q^2); thus, $\sigma_{n;q}^2 = \sigma_n^2 - \phi_Q^2$. If $\sigma_{n;q}^2 \neq 0$, then power is not efficiently increased by increasing r and is maximized by maximizing n and using $r = 1$; however, if $\sigma_{n;q}^2 = 0$, then r and n affect power equally and power is efficiently increased by increasing r and is maximized by maximizing $N \cdot r$. Increasing n efficiently increases power for a wide range of values of $\sigma_{n;q}^2$. $\sigma_{n;q}^2 = 0$ when the genetic variance between lines is fully explained by QTL parameters ($\sigma_n^2 = \phi_Q^2$). This can be achieved by fitting $N - 1$ independent QTL parameters. Significant power increases are seldom achieved by using replicated progeny *unless* a significant fraction of the genetic variance between lines is explained by *simultaneously* estimated QTL parameters. QTL parameter estimation algorithms are proposed which maximize power by minimizing $\sigma_{n;q}^2$.

SEVERAL methods have been described for estimating means of quantitative trait locus (QTL) genotypes using unreplicated progeny (WELLER 1986, 1987; JENSEN 1989; LANDER and BOTSTEIN 1989; SIMPSON 1989; KNAPP, BRIDGES and BIRKES 1990; KNAPP 1990). These methods use marker genotypes or QTL genotypes estimated from marker genotypes as independent variables, genetic models for progeny derived from crosses between inbred lines, *e.g.*, F_2 , backcross (BC), recombinant inbred (RI), doubled haploid (DH), and F_3 progeny, and statistical models suitable for unreplicated individuals or lines. Statistical models and QTL parameter estimation methods have not been described for replicated progeny; however, COWEN (1987) and LANDER and BOTSTEIN (1989) have proposed using replicated progeny to increase power for mapping quantitative trait loci.

The power of tests of hypotheses about means of QTL genotypes are partly determined by the genetic model (LANDER and BOTSTEIN 1989) and the parameter estimation method (SIMPSON 1989). These factors do not determine how experiment design factors affect power; however, they do determine the sample size needed to achieve a given power.

The power of the test of a difference between means of QTL genotypes is increased by increasing the number of replications of lines (r) (LANDER and BOTSTEIN 1989). LANDER and BOTSTEIN (1989) proposed using "progeny testing" ($r > 1$) to increase the "power of QTL mapping" and compared using replicated recombinant inbred (RI) and unreplicated backcross (BC) progeny. They proposed, "recombinant inbred strains will thus typically be more efficient for QTL mapping than an equal number of backcross progeny" by "reducing the environmental variance through replicate phenotypic measurements within each recombinant inbred line" (LANDER and BOTSTEIN 1989). Power is obviously increased by increasing r , but the effect of r has not been quantified. Furthermore, power for replicated progeny is determined by n and other factors in addition to r , and how these factors affect power has not been quantified. In this paper, we estimate power for unreplicated and replicated progeny and show the standard error of a QTL genotype mean is a function of the environmental variance and unexplained genetic variance. Increasing r decreases the standard error of a QTL genotype mean, and consequently increases power for testing

hypotheses about means of QTL genotypes, but this standard error is determined by r and n . We examine the consequences of this and describe ways to maximize the power of unreplicated and replicated progeny.

SAMPLE SIZE DEFINITIONS

Experiment designs for replicated progeny are defined by the number of replications of QTL genotypes (n_i), the number of replications of individuals or lines (r_{ij}), and the way these replications are physically arranged. We use n_i to index the number of individuals or lines of the i th QTL genotype (number of replications of the i th QTL genotype) where the QTL genotype of an individual or line is estimated using molecular marker phenotypes (LANDER and BOTSTEIN 1989; KNAPP, BRIDGES and BIRKES 1990). If marker phenotypes are used as independent variables, then n_i indexes the number of individuals or lines of the i th marker genotype. LANDER and BOTSTEIN (1989) and KNAPP, BRIDGES and BIRKES (1990) described methods which use the EM algorithm to estimate QTL genotypes using linked marker loci flanking a QTL. These values are estimated with error; however, to estimate power, we ignored these errors by assuming the QTL genotypes are known. If misestimation of QTL genotypes causes over- or underestimation of the means of these genotypes, then power is decreased or increased accordingly, but this does not affect how experiment design factors affect power.

We use r_{ij} to index the number of replications of the j th individual or line of the i th QTL genotype. For balanced experiment designs ($n_1 = n_2 = \dots = n_q$ and $r_{11} = r_{12} = \dots = r_{1n} = r_{21} = r_{22} = \dots = r_{qn}$), we write n and r instead of n_i and r_{ij} .

Distinguishing between replications of QTL genotypes and replications of individuals or lines is necessary to define treatment and experiment designs for replicated progeny. Replications of individuals or lines are nested within replications of QTL genotypes and are equivalent to subsamples within replications of QTL genotypes; thus, the treatment design is two-factor nested where the factors are QTL genotypes and lines nested in QTL genotypes. The genotypes for a particular QTL are usually replicated, even when individuals or lines are not; however, a particular multilocus QTL genotype may be missing or unreplicated, as so often happens when the number of simultaneously estimated QTL parameters ($q - 1$) is great. The maximum number of QTL parameters which can be simultaneously estimated is $N - 1$ where N is the size of the progeny population.

The value of n_i for a particular genotype is partly determined by the observed segregation ratio of QTL genotypes. For BC individuals and RI and DH lines, the expected number of replications per QTL geno-

type is $E(n_i) = N/2^k$ where k is the number of independent QTL; thus, for k independent loci, the $E(n_i)$ are equal. The n_i are bound to be unequal in practice because of sampling, natural selection, and linkage between QTL; nevertheless, for unlinked QTL, the arrays of QTL genotypes for BC, RI, and DH progeny are expected to be approximately balanced. The power estimates we make for balanced experiment designs closely approximate the power expected for unlinked QTL for these progeny types.

For F_2 individuals or F_3 lines and $k = 1$, the expected numbers of replications of QTL homozygotes and heterozygotes are $N/2$ and $N/4$, respectively. For k independent loci, the expected numbers of replications of QTL genotypes are products of expected numbers for $k = 1$, e.g., the numbers of replications of the two-locus 11/11 and 11/12 genotypes are $\frac{N}{4} \cdot \frac{N}{4} = \frac{N}{16}$ and $\frac{N}{2} \cdot \frac{N}{4} = \frac{N}{8}$, respectively; thus, the QTL genotype arrays for F_2 or F_3 progeny are unbalanced, and the unbalance becomes increasingly exaggerated as the number of loci increases. We make power estimates for F_3 progeny to show how unbalanced QTL genotype arrays affect power.

LINEAR MODELS AND EXPECTED MEAN SQUARES

The one-factor linear model has been used to develop QTL parameter estimation theory (LANDER and BOTSTEIN 1989; KNAPP, BRIDGES and BIRKES 1990). This model, which is suitable for experiments done using replicated QTL genotypes ($n > 1$), unreplicated lines ($r = 1$), and the completely randomized experiment design, is

$$y_{ij} = \mu + \tau_i + e_{ij} \quad (1)$$

where y_{ij} is an observation of the j th line of the i th QTL genotype, μ is the population mean, τ_i is the effect of the i th QTL genotype, e_{ij} is the ij th residual effect, $i = 1, 2, \dots, q$, $j = 1, 2, \dots, n_i$, q is the number of QTL genotypes, and n_i is the number of replications of the i th QTL genotype. Model (1) is suitable for unreplicated BC, F_2 , DH, RI, and F_3 progeny where progeny and QTL genotypes are randomly assigned to experimental units, as is frequently done in practice.

We expressed the expected mean squares (EMS) of model (1) as a function of the observed variances and the genetic and environmental variances underlying the observed variances (Table 1). The EMS show the error variance is a function of the environmental variance (σ_e^2) and unexplained genetic variance ($\sigma_G^2 = \sigma_G^2 - \phi_Q^2$). σ_G^2 is the fraction of the genetic variance which is not explained by QTL parameters.

The values of ϕ_Q^2 and σ_G^2 are determined by the QTL parameters which are simultaneously estimated.

TABLE 1

Analysis of variance of a completely randomized experiment design for replicated QTL genotypes ($n_i > 1$) and unreplicated lines ($r = 1$) (effects of QTL genotypes are fixed)

Source of variation	Degrees of freedom ^a	Expected mean square ^b
Between QTL genotype	$df_q = q - 1$	$E(M_q) = \sigma_e^2 + \bar{n}\phi_Q^2$ $= [\sigma_E^2 + (\sigma_G^2 - \phi_Q^2)] + \bar{n}\phi_Q^2$
Residual	$df_r = N - q$	$E(M_r) = \sigma_e^2 = \sigma_E^2 + (\sigma_G^2 - \phi_Q^2)$

^a q is the number of QTL genotypes and $N = \sum_{i=1}^q n_i$ is the number of lines where n_i is the number of lines of the i th QTL genotype (number of replications of the i th QTL genotype). If the data are balanced ($n_1 = n_2 = \dots = n_q = n$), then $df_r = q(n - 1)$ where n is the number of replications per QTL genotype.

^b σ_e^2 is the residual variance, σ_E^2 is the environmental variance, σ_G^2 is the genetic variance, ϕ_Q^2 are the fixed effects of QTL genotypes, $\sigma_G^2 - \phi_Q^2$ is the genetic variance which is not explained by QTL

parameters, and $\bar{n} = \frac{\sum_{i=1}^q n_i^2}{N - \frac{N}{q}}$. If the data are balanced, then $\bar{n} = n$.

The total degrees of freedom of (1) are $N - 1$ (Table 1). These degrees of freedom can be partitioned many different ways by fitting different QTL models. q QTL genotypes are needed to simultaneously estimate $q - 1$ QTL parameters or, equivalently, to test $q - 1$ independent hypotheses about intralocus and interlocus QTL effects. The value of q for a given model is limited by N . Using model (1), N must be greater than q because the error degrees of freedom are $N - q$ and the error is estimated using replications of QTL genotypes (Table 1). For balanced arrays of QTL genotypes ($n_1 = n_2 = \dots = n_q$), independent QTL, and RI, DH, or BC progeny, the maximum value of q is $N/2 = n = \bar{n}$, otherwise \bar{n} must be greater than 1 because $\lim_{\bar{n} \rightarrow 1} q = N$ (Table 1).

If QTL parameters explain the genetic variance among unreplicated individuals or lines ($\sigma_G^2 = \phi_Q^2$), then $\sigma_e^2 = \sigma_E^2$, otherwise $\sigma_e^2 = \sigma_E^2 + (\sigma_G^2 - \phi_Q^2)$ (Table 1). $H_0: \sigma_G^2 = \phi_Q^2$ cannot be tested using (1); nevertheless, it is important to understand the consequences of the structure of σ_e^2 . σ_e^2 is minimized and power is maximized by simultaneously estimating those intra-locus and interlocus QTL parameters (ϕ_Q^2) which explain a maximum of σ_G^2 (Table 1).

Even though $H_0: \sigma_G^2 = \phi_Q^2$ cannot be tested using (1), a hypothesis can be tested to determine if σ_e^2 is decreased by adding certain QTL parameters to the model. Let the full model be a linear model which has the QTL parameters of the reduced model plus additional QTL parameters. A test is needed to determine whether or not additional genetic variance is explained by the QTL parameters which have been added to the reduced model. The null and alternate hypotheses are $H_0: \phi_Q^2 = 0$ and $H_a: \phi_Q^2 \neq 0$ where ϕ_Q^2

are the fixed effects of the QTL parameters added to the reduced model.

A likelihood ratio (LEHMANN 1986) for testing $H_0: \phi_Q^2 = 0$ is

$$L = \frac{(S_{e_R} - S_{e_F})/(df_{e_R} - df_{e_F})}{S_{e_F}/(df_{e_F})} = \frac{S_{\hat{Q}}/(df_{\hat{Q}})}{S_{e_F}/(df_{e_F})}$$

where S_{e_F} and S_{e_R} are the residual sums of squares and df_{e_F} and df_{e_R} are the residual degrees of freedom for the full and reduced models, respectively, and $S_{\hat{Q}}$ and $df_{\hat{Q}}$ are the sums of squares and degrees of freedom for the QTL parameters added to the model (LEH-

MANN 1987). $\frac{S_{\hat{Q}}}{E(M_{\hat{Q}})}$ and $\frac{S_{e_F}}{E(M_{e_F})}$ are independent χ^2 random variables under normality of the random effects of the full and reduced models; thus, the ratio

$$\frac{M_{\hat{Q}}/[E(M_{\hat{Q}})]}{M_{e_F}/[E(M_{e_F})]} \sim F_{df_{\hat{Q}}, df_{e_F}}$$

is distributed as an F random variable (MOOD, GRAYBILL and BOES 1974; LEHMANN 1987). If $L > F_{\alpha, df_{\hat{Q}}, df_{e_F}, 0}$ where $F_{\alpha, df_{\hat{Q}}, df_{e_F}, 0}$ is a critical value from a central F -distribution, then there is evidence the QTL parameters added to the model explain a significant fraction of the genetic variance.

Testing $H_0: \phi_Q^2 = 0$ is equivalent to testing $H_0: \sigma_{\hat{G}_R}^2 \leq \sigma_{\hat{G}_F}^2$ where $\sigma_{\hat{G}_F}^2$ and $\sigma_{\hat{G}_R}^2$ are unexplained genetic variances for the full and reduced models. The residual variances for the full and reduced models are $\sigma_{e_F}^2 = \sigma_E^2 + \sigma_{\hat{G}_F}^2 = \frac{S_{e_F}}{df_{e_F}} = M_{e_F}$ and $\sigma_{e_R}^2 = \sigma_E^2 + \sigma_{\hat{G}_R}^2 = \frac{S_{e_R}}{df_{e_R}} = M_{e_R}$ where M_{e_F} and M_{e_R} are estimates of $\sigma_{e_F}^2$ and $\sigma_{e_R}^2$. If meaningful QTL parameters are added to the reduced model, then the residual variance for the full model should be less than the residual variance for the reduced model. This is the alternate hypothesis ($H_a: \sigma_{e_R}^2 > \sigma_{e_F}^2$). The null hypothesis is $H_0: \sigma_{e_R}^2 \leq \sigma_{e_F}^2$. These hypotheses are equivalent to $H_0: \sigma_{\hat{G}_R}^2 \leq \sigma_{\hat{G}_F}^2$ and $H_a: \sigma_{\hat{G}_R}^2 > \sigma_{\hat{G}_F}^2$ (Table 1). The F -statistic for testing $H_0: \sigma_{e_R}^2 \leq \sigma_{e_F}^2$ is

$$F = \frac{M_{e_R}}{M_{e_F}} = \frac{\hat{\sigma}_{e_R}^2}{\hat{\sigma}_{e_F}^2} = \frac{\hat{\sigma}_E^2 + \hat{\sigma}_{\hat{G}_R}^2}{\hat{\sigma}_E^2 + \hat{\sigma}_{\hat{G}_F}^2} = \frac{\hat{\sigma}_{\hat{G}_R}^2}{\hat{\sigma}_{\hat{G}_F}^2}$$

(MOOD, GRAYBILL and BOES 1974; SOKAL and ROHLF 1981). If $F > F_{\alpha, df_{\hat{G}_R}, df_{\hat{G}_F}, 0}$ where $F_{\alpha, df_{\hat{G}_R}, df_{\hat{G}_F}, 0}$ is a critical value from a central F -distribution, then H_0 is rejected and there is evidence additional genetic variance is explained by the QTL parameters added to the model (MOOD, GRAYBILL and BOES 1974; SOKAL and ROHLF 1981). This shows how the residual or error variance of model (1) is decreased by adding QTL parameters to the model which explain additional genetic variance.

This principle extends to other models, e.g., those

TABLE 2

Analysis of variance of a balanced completely randomized experiment design for replicated QTL genotypes ($n > 1$) and replicated lines ($r > 1$) (effects of QTL genotypes and lines are fixed and random, respectively)

Source of variation	Degrees of freedom ^a	Expected mean square ^b
Line (L)	$df_n = qn - 1 = N - 1$	$E(M_n) = \sigma_e^2 + r\sigma_n^2 = \sigma_e^2 + \sigma_{C_W}^2 + r\sigma_{C_B}^2 = \sigma_e^2 + \sigma_{C_W}^2 + r(\sigma_{C_B}^2 - \phi_Q^2) + \left[\frac{(q-1)rn}{(qn-1)} \right] \phi_Q^2$
QTL Genotype (Q)	$df_q = q - 1$	$E(M_q) = \sigma_e^2 + r\sigma_{n:q}^2 + rn\phi_Q^2 = \sigma_e^2 + \sigma_{C_W}^2 + r(\sigma_{C_B}^2 - \phi_Q^2) + rn\phi_Q^2$
$L:Q$	$df_{n:q} = q(n-1) = N - q$	$E(M_{n:q}) = \sigma_e^2 + r\sigma_{n:q}^2 = \sigma_e^2 + \sigma_{C_W}^2 + r(\sigma_{C_B}^2 - \phi_Q^2)$
Residual	$df_e = qn(r-1) = N(r-1)$	$E(M_e) = \sigma_e^2 = \sigma_e^2 + \sigma_{C_W}^2$

^a q is the number of QTL genotypes, n is the number of lines per QTL genotype, r is the number of replications per line, and N is the number of lines ($N = qn$).

^b σ_e^2 is the residual variance, $\sigma_{n:q}^2$ is the line nested in QTL genotype variance, σ_n^2 is the between line variance, ϕ_Q^2 are the fixed effects of QTL genotypes, $\sigma_{C_B}^2 = \sigma_{C_B}^2 + \sigma_{C_W}^2$ is the genetic variance, $\sigma_{C_B}^2$ and $\sigma_{C_W}^2$ are between and within line genetic variances, respectively, and σ_e^2 is the environmental variance.

for replicated progeny. Using replicated QTL genotypes ($n_i > 1$), replicated progeny ($r_{ij} > 1$), and the completely randomized experiment design, a suitable linear model is

$$y_{ijk} = \mu + \tau_i + \gamma_{ij} + e_{ijk} \quad (2)$$

where y_{ijk} is the phenotypic value of the k th replication of the j th line of the i th QTL genotype, μ is the population mean, τ_i is the effect of the i th QTL genotype, γ_{ij} is the effect of the j th line of the i th QTL genotype, e_{ijk} is the ijk th residual effect, $i = 1, 2, \dots, q$, $j = 1, 2, \dots, n_i$, $k = 1, 2, \dots, r_{ij}$, q is the number of QTL genotypes, n_i is the number of replications of the i th QTL genotype, and r_{ij} is the number of replications of the j th line of the i th QTL genotype. Model (2) is suitable for replicated BC, F_2 , DH, RI, and F_3 progeny where $q < N$ and progeny and QTL genotypes are randomly assigned to experimental units.

We expressed the EMS of (2) as a function of the genetic and environmental variances underlying the observed variances (Table 2). The sum of τ_i and γ_{ij} of (2) is equal to the effects of lines (v_m); thus, rewriting (2) we get

$$y_{mk} = \mu + v_m + e_{mk} \quad (3)$$

where y_{mk} is the phenotypic value of the k th replication of the m th line, μ is the population mean, v_m is the effect of the m th line, e_{mk} is the mk th residual effect, $m = 1, 2, \dots, N$, $k = 1, 2, \dots, r_m$, N is the number of lines, and r_m is the number of replications of the m th line (Table 2). Models (2) and (3) show $v_m = \tau_i + \gamma_{ij}$ and $e_{mk} = e_{ijk}$, which means the sum of the sums of squares between QTL genotypes (S_q) and the sums of squares between lines nested within QTL genotypes ($S_{n:q}$) are equal to the sums of squares between lines (S_n) (Table 2).

The expected sums of squares between lines, QTL

genotypes, and lines nested in QTL genotypes are

$$E(S_n) = (qn - 1)[\sigma_e^2 + r\sigma_n^2],$$

$$E(S_q) = (q - 1)[\sigma_e^2 + r\sigma_{n:q}^2 + rn\phi_Q^2],$$

and

$$E(S_{n:q}) = q(n - 1)[\sigma_e^2 + r\sigma_{n:q}^2],$$

respectively (Table 2). Pooling $E(S_q)$ and $E(S_{n:q})$, we get

$$\begin{aligned} E(S_q) + E(S_{n:q}) &= q\sigma_e^2 - \sigma_e^2 + qr\sigma_{n:q}^2 - r\sigma_{n:q}^2 + qnr\phi_Q^2 - nr\phi_Q^2 \\ &\quad + qn\sigma_e^2 - q\sigma_e^2 + qnr\sigma_{n:q}^2 - qr\sigma_{n:q}^2 \\ &= (qn - 1)\sigma_e^2 + (qn - 1)r\sigma_{n:q}^2 + (q - 1)rn\phi_Q^2. \end{aligned} \quad (4)$$

The pooled expected mean square is found by dividing (4) by the degrees of freedom for lines ($df_n = N - 1 = qn - 1$); thus,

$$\begin{aligned} &\frac{(qn - 1)\sigma_e^2 + (qn - 1)r\sigma_{n:q}^2 + (q - 1)rn\phi_Q^2}{qn - 1} \\ &= \sigma_e^2 + r \left[\sigma_{n:q}^2 + \frac{(qn - n)\phi_Q^2}{(qn - 1)} \right] \end{aligned} \quad (5)$$

and

$$\begin{aligned} \sigma_n^2 &= \sigma_{C_B}^2 = \sigma_{n:q}^2 + \frac{(qn - n)}{(qn - 1)} \phi_Q^2 \\ &= (\sigma_{C_B}^2 - \phi_Q^2) + \frac{(qn - n)}{(qn - 1)} \phi_Q^2. \end{aligned} \quad (6)$$

Equations 5 and 6 show the genetic variance between lines is equal to the fraction of the genetic variance explained by the QTL model (ϕ_Q^2) and the residual genetic variance between lines ($\sigma_{n:q}^2$); thus, power is determined by the experiment and treatment

designs. The experiment design is a function of r , whereas the treatment design is a function of N , q , n , and the QTL model. n is determined by the size of the progeny population (N), the number of simultaneously estimated QTL parameters ($q - 1$), progeny type, marker model, and, for multilocus models, whether or not there is linkage between QTL. n is increased by increasing N .

The standard error of a QTL genotype mean for model (2) is

$$\begin{aligned}\sigma_{\bar{y}} &= \sqrt{\frac{\sigma_e^2}{rn} + \frac{\sigma_{n:q}^2}{n}} \\ &= \sqrt{\frac{\sigma_E^2 + \sigma_{Gw}^2}{rn} + \frac{\sigma_{Gg}^2 - \phi_Q^2}{n}}\end{aligned}\quad (7)$$

(Table 2). $\sigma_{\bar{y}}$ is decreased by increasing r and n , but increasing r solely decreases the effect of σ_e^2 , whereas increasing n decreases the effects of σ_e^2 and $\sigma_{n:q}^2$; thus, how r and n affect power is a function of the size of σ_e^2 and $\sigma_{n:q}^2$. The sizes of these variances are determined by the size of ϕ_Q^2 (Table 2). $\sigma_{\bar{y}}$ is minimized by maximizing ϕ_Q^2 . An estimation algorithm for minimizing $\sigma_{\bar{y}}$ is described later.

Equations 5 and 6 show the connection between classical quantitative genetic and QTL parameters, and lead to useful hypothesis tests. Using (2) (Table 2), $N - 1$ orthogonal contrasts among the means of QTL genotypes must exist which explain 100% of the genetic variance between lines; of course, 100% of σ_n^2 may be explained by fewer than $N - 1$ orthogonal contrasts. Under the restriction of balance among QTL genotypes ($n = n_1 = n_2 = \dots = n_q$), we know $n = 1$, $N = q$, $\hat{\sigma}_{n:q}^2 = 0$, and, as a consequence, $\hat{\sigma}_n^2 = \hat{\sigma}_Q^2$ when $N - 1$ orthogonal contrasts among the means of QTL genotypes are estimated. $\hat{\phi}_Q^2 = \hat{\sigma}_n^2$ does not imply meaningful QTL parameters have been estimated because $N - 1$ orthogonal contrasts explain 100% of $\hat{\sigma}_n^2$, even when none of the individual QTL parameters explain a significant fraction of the genetic variance. Thus, while it is technically feasible to simultaneously estimate QTL parameters which explain 100% of the genetic variance between lines ($\phi_Q^2 = \sigma_n^2$), this does not imply the parameters of every important QTL have been estimated. This feature of model (2) is nonetheless useful for devising algorithms which maximize power for testing hypotheses about means of QTL genotypes.

The basic principle of the proposed estimation algorithm is to fit saturated QTL models. A model is saturated when $\sigma_{n:q}^2 = 0$. The QTL model may be iteratively built by subtracting nonsignificant effects, retaining significant effects, and estimating a maximum number of independent QTL parameters at every stage. This is an effective strategy for maximizing power and finding a QTL model when the number

of observations (N) is less than the number of QTL parameters which might be estimated. The strategy of simultaneously estimating $q - 1 = N - 1$ independent QTL parameters forces $\sigma_{n:q}^2 = 0$ and leads to models different from (2). Other strategies might be devised to achieve this end, e.g., estimating $< N - 1$ independent QTL parameters and testing whether or not $S_{n:q}$ and S_e of (2) should be pooled.

Whether or not pooling should be done is determined by the strength of the evidence about the size of the variance to be pooled ($H_0: \sigma_{n:q}^2 = 0$) and how pooling affects power for testing hypotheses about the effects of QTL genotypes (PAULL 1950; DUNN and CLARK 1987). Liberal type I error probabilities (α) have been proposed for testing $H_0: \sigma_{n:q}^2 = 0$ to minimize the probability of type II errors, e.g., $\alpha = 0.25$ has been widely used (PAULL 1950; DUNN and CLARK 1987). Pooling is typically justified when

$$Pr\left[\frac{M_{n:q}}{M_e} > F_{\alpha, df_{n:q}, df_e, 0}\right]$$

ranges from 0.00 to 0.25 where $F_{\alpha, df_{n:q}, df_e, 0}$ is a critical value from a central F -distribution and $\frac{M_{n:q}}{M_e}$ is the F -

statistic used to test $H_0: \sigma_{n:q}^2 = 0$ (Table 2) (PAULL 1950; DUNN and CLARK 1987). The rationale for using the $\alpha \leq 0.25$ threshold as a pooling criteria is to maximize power for testing $H_0: \mu_{11} = \mu_{22}$. This is justified because there usually are no serious negative ramifications of type I errors for $H_0: \sigma_{n:q}^2 = 0$.

This is known as the *sometimes pool* criteria, which states model effects should be pooled when the evidence justifies it (KIRK 1982; DUNN and CLARK 1987). The *never pool* criteria states model effects should never be pooled. Both criteria have their proponents (KIRK 1982). We favor the *sometimes pool* criteria. The $\alpha \leq 0.25$ threshold increases the probability of a type I error but decreases the probability of a type II error for $H_0: \sigma_{n:q}^2 = 0$. Other type I error probability thresholds could be used to define the *sometimes pool* criteria for $H_0: \sigma_{n:q}^2 = 0$.

If $H_0: \sigma_{n:q}^2 = 0$ is accepted, then model (2) is not valid, $S_{n:q}$ and S_e should be pooled, and a suitable linear model is

$$y_{im} = \mu + \beta_i + e_{im} \quad (8)$$

where y_{im} is the phenotypic value of the m th replication of the i th QTL genotype, μ is the population mean, β_i is the effect of the i th QTL genotype, e_{im} is the im th residual effect, $i = 1, 2, \dots, q$, $m = 1, 2, \dots, s$, q is the number of QTL genotypes, $s = nr$ is the number of replications of per QTL genotype, n is the number of lines per QTL genotype, and r is the number of replications per line. Using model (8), there is no distinction between n and r .

Thus, to determine whether or not model (8) is

TABLE 3

Analysis of variance of a balanced completely randomized experiment design for replicated lines ($r > 1$) where the line nested in QTL genotype and residual sums of squares of model (2) have been pooled (effects of QTL genotypes are fixed)

Source of variation	Degrees of freedom ^a	Expected mean square ^b
Between QTL genotype	$df_q = q - 1$	$E(M_q) = \sigma_e^2 + r\phi_Q^2$ $= \sigma_e^2 + \sigma_{C_W}^2 + r\phi_Q^2$
Residual	$df_e = Nr - q$	$E(M_e) = \sigma_e^2 = \sigma_e^2 + \sigma_{C_W}^2$

^a r is the number of replications per line, q is the number of QTL genotypes, and N is the number of lines. $q = N$, $n = 1$, and $df_e = N(r - 1)$ when $N - 1$ independent QTL parameters are estimated where n is the number of replications per QTL genotype.

^b σ_e^2 is the residual variance, σ_e^2 is the environmental variance, $\sigma_{C_W}^2$ is the within line genetic variance, and ϕ_Q^2 are the fixed effects of QTL genotypes.

valid, the hypothesis $H_0: \sigma_{n:q}^2 = \sigma_{C_B}^2 - \phi_Q^2 = 0$, which is equivalent to $H_0: \sigma_{C_B}^2 = \phi_Q^2$, should be tested to determine whether or not $S_{n:q}$ and S_e of (2) should be pooled. This hypothesis need not be tested if $N - 1$ orthogonal contrasts among the means of QTL genotypes are estimated because model (8) is valid under this restriction (Table 3).

POWER FOR BALANCED LINEAR MODELS

We estimated the power of the test of the additive effect of a QTL ($\mu_{11} - \mu_{22}$) for balanced models (2) and (8) where μ_{11} and μ_{22} are means of QTL genotypes 11 and 22, respectively (Tables 2 and 3). We set $q = 2$ and n_1 and n_2 where $n_1 = n_2$ are the number of replications of QTL genotypes 11 and 22, respectively. The power of the test of $H_0: \mu_{11} = \mu_{22}$ for model (2) is

$$Pr[F_{df_q, df_{n:q}, \lambda} > F_{\alpha, df_q, df_{n:q}, 0}]$$

where $F_{df_q, df_{n:q}, \lambda}$ is a random variable from a noncentral F -distribution ($\lambda \neq 0$), $F_{\alpha, df_q, df_{n:q}, 0}$ is a critical value from a central F -distribution ($\lambda = 0$), λ is the noncentrality parameter, and α is the probability of a type I error (Table 2) (O'BRIEN 1986). The F -statistic for testing this hypothesis is $F = \frac{M_q}{M_{n:q}}$, and the noncentrality parameter is

$$\lambda = \frac{E(S_q)}{E(M_{n:q})} = \frac{df_q[\sigma_e^2 + r\sigma_{n:q}^2 + rn\phi_Q^2]}{\sigma_e^2 + r\sigma_{n:q}^2} \quad (9)$$

(Table 2). Because $df_q = 1$, the noncentrality parameter for the test of $H_0: \mu_{11} = \mu_{22}$ is

$$\begin{aligned} \lambda &= \frac{E(M_q)}{E(M_{n:q})} = \frac{\sigma_e^2 + r\sigma_{n:q}^2 + rn\phi_Q^2}{\sigma_e^2 + r\sigma_{n:q}^2} \\ &= 1 + \frac{rn\phi_Q^2}{\sigma_e^2 + \sigma_{C_W}^2 + r(\sigma_{C_B}^2 - \phi_Q^2)} \end{aligned} \quad (10)$$

(Table 2). Critical values and probabilities for F -distributions were determined using the SAS (1987)

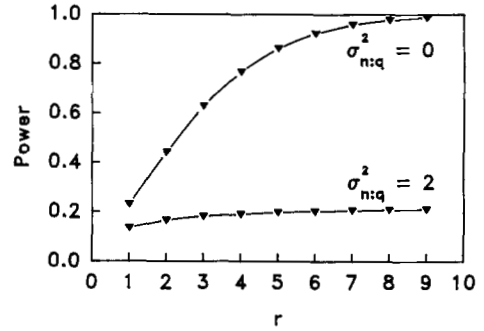


FIGURE 1.—Power of the test of the difference between means of QTL homozygotes. Power was estimated using models (2) ($\sigma_{n:q}^2 \neq 0$) or (7) ($\sigma_{n:q}^2 = 0$), $\alpha = 0.01$, $\sigma_e^2 = 2$, $\phi_Q^2 = 0.1$, $df_q = 1$, $q = 2$, $n = 50$, and $N = 100$.

functions FINV and PROBF.

If, in addition to testing $H_0: \mu_{11} = \mu_{22}$, an additional $N - 2$ orthogonal contrasts are simultaneously tested, then $\sigma_{n:q}^2 = 0$ and model (8) is used. Power for model (8) is

$$Pr[F_{df_q', df_e, \lambda} > F_{\alpha, df_q', df_e, 0}]$$

where $F_{df_q', df_e, \lambda}$ is a random variable from a noncentral F -distribution ($\lambda \neq 0$), $F_{\alpha, df_q', df_e, 0}$ is a critical value from a central F -distribution ($\lambda = 0$), and $df_q' = df_q - N + 2 = 1$ is the degree of freedom for testing $H_0: \mu_{11} = \mu_{22}$ (Table 3). The F -statistic for testing $H_0: \mu_{11} = \mu_{22}$ is $F = \frac{M_{q'}}{M_e}$ where $M_{q'}$ is the between QTL genotype mean square for this hypothesis (Table 3). The noncentrality parameter for this test uses $df_q' = 1$, whereas the full QTL model uses $df_q = N - 1$ (df_q' and an additional $N - 2$ degrees of freedom). The noncentrality parameter is

$$\begin{aligned} \lambda &= \frac{E(S_{q'})}{E(M_e)} = \frac{\sigma_e^2 + r\phi_{Q'}^2}{\sigma_e^2} = \frac{\sigma_e^2 + \sigma_{C_W}^2 + r\phi_{Q'}^2}{\sigma_e^2 + \sigma_{C_W}^2} \\ &= 1 + \frac{r\phi_{Q'}^2}{\sigma_e^2 + \sigma_{C_W}^2} \end{aligned} \quad (11)$$

where $\phi_{Q'}$ are the fixed effects of QTL genotypes 11 and 22 and $E(S_{q'}) = df_{q'}E(M_q)$. This noncentrality parameter is quite different from (10) for model (2). The hypotheses being tested are equivalent ($\phi_{Q'}$ of (8) and ϕ_Q^2 of (2) are equal), but the error variances are different. The denominator of (11) is strictly a function of environmental and within line genetic variance, whereas the denominator of (10) is a function of environmental, within line genetic, and unexplained between line genetic variance.

The power of the test of $H_0: \mu_{11} = \mu_{22}$, estimated using models (2) and (7) (Tables 2 and 3), is determined by n , r , q , σ_e^2 , $\sigma_{n:q}^2$, and $\mu_{11} - \mu_{22}$. If $\sigma_{n:q}^2 \neq 0$, then power is not efficiently increased by increasing the number of replications of lines (Figures 1 and 2); however, if $\sigma_{n:q}^2 = 0$, then power is efficiently increased by increasing r (Figure 1). This is illustrated by the

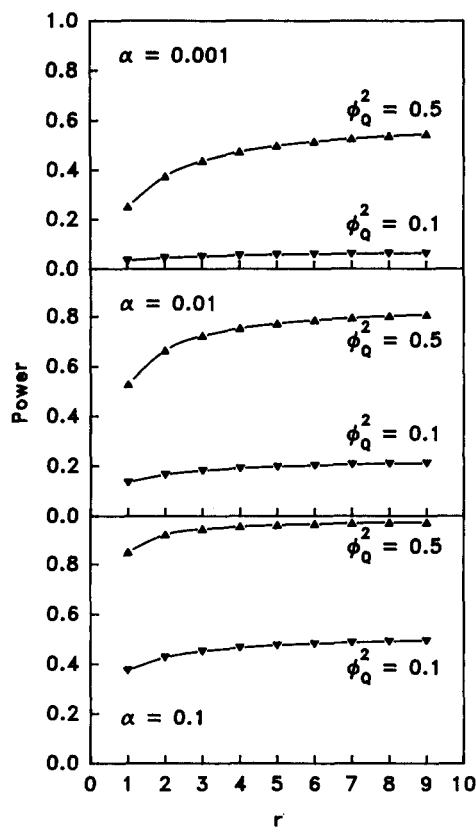


FIGURE 2.—Power of the test of the difference between means of QTL homozygotes. Power was estimated using model (2), $\sigma_e^2 = 2.0$, $\sigma_{n;q}^2 = 2.0$, $df_q = 1$, $q = 2$, $n = 50$, and $N = 100$.

power curves for $\sigma_{n;q}^2 = 0$ and $\sigma_{n;q}^2 = 2$ (Figure 1). The power curve for $\sigma_{n;q}^2 = 0$ sharply increases as r increases, whereas the power curve for $\sigma_{n;q}^2 = 2$ is nearly flat (Figure 1).

If $\sigma_{n;q}^2 \geq \sigma_e^2$, then power curves expressed as a function of r are fairly flat (Figure 2). Meaningful power increases are seldom achieved by increasing r unless $\sigma_{n;q}^2 = 0$ (Figure 2). The merit of increasing r under model (2) ($\sigma_{n;q}^2 \neq 0$) increases as ϕ_Q^2 increases and α decreases; however, meaningful increases are seldom achieved by using $r > 3$ (Figure 2). If $N - 1$ independent QTL parameters are simultaneously estimated or if the simultaneously estimated QTL parameters eliminate $\sigma_{n;q}^2$, then power can be substantially increased by increasing r (Figure 1).

Power is efficiently increased by increasing the number of replications of QTL genotypes (n) or the size of the progeny population (N) independent of the value of $\sigma_{n;q}^2$ (Figure 3). n is much less affected by $\sigma_{n;q}^2$ than r (Figure 3); nevertheless, the effects of n and r are maximized by satisfying $\sigma_G^2 - \phi_Q^2 = 0$ or, equivalently, $\sigma_{n;q}^2 = 0$ and $\sigma_e^2 = \sigma_E^2 + \sigma_{G_w}^2$. If $\sigma_{n;q}^2 = 0$, then model (8) is used and r and n affect power equally. Under model (8) there is no functional difference between r and n or between increasing the progeny population size (N), which is equivalent to increasing n , or the number of replications of progeny

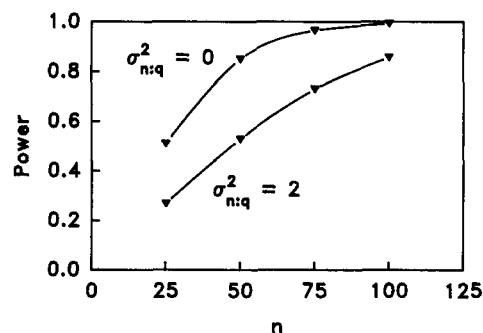


FIGURE 3.—Power of the test of the difference between means of QTL homozygotes. Power was estimated using models (2) ($\sigma_{n;q}^2 \neq 0$) or (7) ($\sigma_{n;q}^2 = 0$), $\alpha = 0.01$, $\sigma_e^2 = 2$, $\phi_Q^2 = 0.5$, $df_q = 1$, $q = 2$, and $r = 1$.

(r) (Table 3). Either increases the number of replications per QTL genotype under model (8) (Table 3).

If the number of experimental units is fixed and $\sigma_{n;q}^2 \neq 0$, then increasing r decreases n and power, and power is maximized by maximizing n and using $r = 1$ (Figure 4). The power achieved using $r > 1$ never exceeds but may be equal to the power achieved using $r = 1$ for a fixed number of experimental units (Figure 4); however, an estimation algorithm should be used which seeks to minimize $\sigma_{n;q}^2$, thus maximizing power for a given sample size.

We have shown using replicated progeny does not significantly increase power for estimating the effects of QTL *unless* an estimation algorithm is used which minimizes or eliminates $\sigma_{n;q}^2$ (Figures 1 and 2). If such an algorithm is used, then power can be greatly increased by increasing r . The cost of r and N determines the efficiency of a particular sampling strategy. The cost of N is a function of the number of marker loci assayed and the cost of the technology used to assay marker phenotypes, *e.g.*, restriction fragment length polymorphism (RFLP) assays. The cost of r is a function of the number of environments used, the number of dependent variables measured, and the cost of measuring dependent variables. Because the cost of r is usually substantially less than the cost of N , sample size ($N \cdot r$) is usually more efficiently increased by increasing r than by increasing N . Maximum efficiency can be achieved by coupling the use of replicated progeny with an estimation algorithm which minimizes or eliminates $\sigma_{n;q}^2$. This maximizes the power of replicated progeny and power for a given sample size ($N \cdot r$).

Technological advances may decrease the cost of DNA marker assays, which should greatly affect experiment design and planning. Equal costs of r and N , for example, favor increasing N at the expense of r since N determines the number of QTL parameters which can be simultaneously estimated. It may be necessary, for example, to increase N to simultaneously estimate the parameters of QTL affecting a trait. This problem is exacerbated by unbalanced arrays of

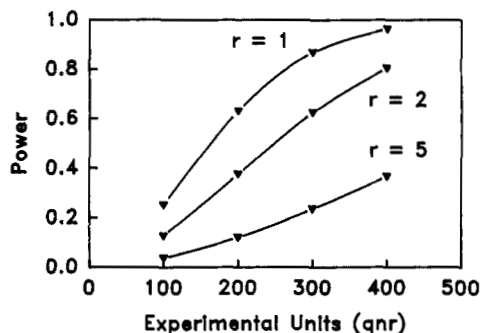


FIGURE 4.—Power of the test of the difference between means of QTL homozygotes. Power was estimated using model (2), $\alpha = 0.001$, $\sigma_e^2 = 2$, $\sigma_{n;q}^2 = 2$, $\phi_Q^2 = 0.5$, $df_q = 1$, and $q = 2$.

QTL genotypes. Using unreplicated lines greatly diminishes the versatility of progeny which can be replicated, e.g., by replicating lines, genetic variances between lines can be estimated and used to test hypotheses about classical quantitative genetic and QTL parameters and, if every line is replicated, entire lines and the investment in marker genotype assays for those lines are seldom lost to natural causes. Every line need not be replicated if the problem of losing lines is unimportant.

POWER FOR UNBALANCED LINEAR MODELS

Although the n_i for RI and DH progeny are usually not equal in practice, the weighted mean number of replications of QTL genotypes (\bar{n}) for these progeny is not expected to be greatly different from n . If, for model (1), $N = 100$, $n_1 = 46$, and $n_2 = 54$, then the mean number of replications is

$$\bar{n} = \frac{N - \sum_{i=1}^q n_i^2}{q - 1} = \frac{100 - \frac{(46^2 + 54^2)}{100}}{1} = 49.68,$$

which is not greatly different from the value for balanced data ($n = n_1 = n_2 = 50$); thus, power for balanced models closely approximates the power expected for BC, DH, and RI progeny using independent QTL.

The expected value of \bar{n} for 100 F_2 individuals or F_3 lines, model (1), and $k = 1$ is

$$\bar{n} = \frac{100 - \frac{(25^2 + 50^2 + 25^2)}{100}}{2} = 31.25;$$

thus, for a given value of N , \bar{n} is less for F_2 and F_3 progeny than for BC, DH, or RI progeny.

To show how decreasing \bar{n} affects power for a fixed population size (N), we estimated the power of the additive effect of a QTL using unreplicated RI and F_3 lines. This was done using model (1) and expected values of n_i . The power of the test of $H_0: \mu_{11} = \mu_{22}$ for

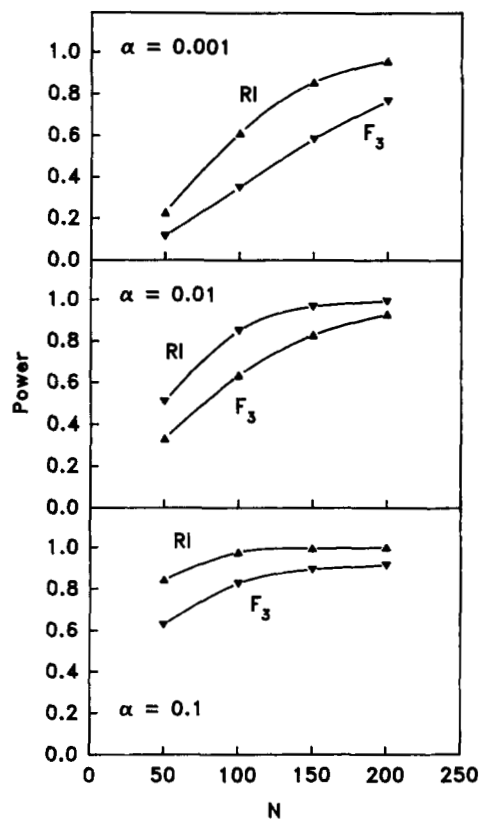


FIGURE 5.—Power of the test of the difference between means of QTL homozygotes. Power was estimated for unreplicated RI and F_3 lines using model (1), $\sigma_e^2 = 2$, $\phi_Q^2 = 0.5$, $df_q = 1$, $q = 2$, $r = 1$, and expected values of n_i for RI and F_3 progeny.

model (1) is

$$Pr[F_{df_q, df_e, \lambda} > F_{\alpha, df_q, df_e, 0}]$$

where $F_{df_q, df_e, \lambda}$ is a random variable from a noncentral F -distribution ($\lambda \neq 0$) and $F_{\alpha, df_q, df_e, 0}$ is a critical value from a central F -distribution ($\lambda = 0$) (Table 1). The F -statistic and noncentrality parameter for model (1)

are $F = \frac{M_q}{M_e}$ and

$$\lambda = \frac{E(S_q)}{E(M_e)} = \frac{\sigma_e^2 + \bar{n}\phi_Q^2}{\sigma_e^2} = 1 + \frac{\bar{n}\phi_Q^2}{\sigma_e^2 + (\sigma_G^2 - \phi_Q^2)}, \quad (12)$$

respectively (Table 1).

Power for estimating the additive effect of a QTL is greater for RI than for F_3 progeny if equivalent population sizes (N) are used (Figure 5). This is strictly a function of \bar{n} . Power is less for F_3 progeny because the average number of replications per QTL genotype is less (Figure 5). Restated, the average number of replications per contrast is less for F_3 progeny than for RI progeny; nevertheless, the principles needed to maximize power for unbalanced models are equivalent to those for balanced models. Unbalance decreases \bar{n} for a given N , but the estimation and experiment design strategies needed to maximize power are

not affected by this. Unbalance does, however, affect the sample size needed to achieve a given power.

This work was supported by grants to S. J. K. from the U.S. Department of Agriculture and Medical Research Foundation of Oregon. We are grateful to BILL BEAVIS and ERIC LANDER for criticisms which significantly improved the manuscript. Oregon Agricultural Experiment Station Technical Paper No. 9147.

LITERATURE CITED

- COWEN, N. M., 1988 The use of replicated progenies in marker-based mapping of QTL's. *Theor. Appl. Genet.* **75**: 857-862.
- DUNN, O. J., and V. CLARK, 1987 *Applied Statistics: Analysis of Variance and Regression Analysis*. John Wiley & Sons, New York.
- JENSEN, J., 1989 Estimation of recombination parameters between a quantitative trait locus (QTL) and two marker gene loci. *Theor. Appl. Genet.* **78**: 613-618.
- KIRK, R. E., 1982 *Experimental Design: Procedures for the Behavioral Sciences*. Wadsworth, Belmont, Calif.
- 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.* (in press).
- KNAPP, S. J., W. C. BRIDGES and D. BIRKES, 1990 Mapping quantitative trait loci using molecular marker linkage maps. *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.
- LEHMANN, E. L., 1987 *Testing Statistical Hypotheses*. John Wiley & Sons, New York.
- MOOD, A. M., F. A. GRAYBILL and D. C. BOES, 1974 *Introduction to the Theory of Statistics*. McGraw-Hill, New York.
- O'BRIEN, R. G., 1986 Power analysis for linear models, in *Proceedings of the Eleventh Annual SAS Users Group Conference*. SAS, Cary, N.C.
- PAULL, A. E., 1950 On a preliminary test for pooling mean squares in the analysis of variance. *Ann. Math. Stat.* **21**: 539-556.
- SAS, 1987 *SAS STAT Guide for Personal Computers*. Version 6 Edition. SAS, Cary, N.C.
- SIMPSON, S. P., 1989 Detection of linkage between quantitative trait loci and restriction fragment length polymorphisms using inbred lines. *Theor. Appl. Genet.* **77**: 815-819.
- SOKAL, R. R., and F. J. ROHLF, 1981 *Biometry*. Freeman, New York.
- WELLER, J. I., 1986 Maximum likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic markers. *Biometrics* **42**: 627-640.
- WELLER, J. I., 1987 Mapping and analysis of quantitative trait loci in *Lycopersicon* (tomato) with the aid of genetic markers using approximate maximum likelihood methods. *Heredity* **59**: 413-421.

Communicating editor: S. D. TANKSLEY