Theor Appl Genet (1993) 86:1014–1022

THEORETICAL AND APPLIED GENETICS

Power of tests for QTL detection using replicated progenies derived from a diallel cross

A. Rebai, B. Goffinet

Laboratory of Biometry, National Institute of Agronomy Research, Centre de Toulouse, Chemin Borde Rouge, Auzeville BP 27, 31326 Castanet-Tolosan, France

Received: 7 October 1992 / Accepted: 28 January 1993

Abstract. In crop species, most QTL (quantitative trait loci) mapping strategies use segregating populations derived from an initial cross between two lines. However, schemes including more than two parents could also be used. We propose an approach using a high-density restriction fragment length polymorphism (RFLP) map established on six F_2 populations derived from diallel crosses among four inbred lines and the phenotypic performances of two types of replicated progenies (F_3 and topcross). The QTL is supposed to be on the marker locus considered. Three linear model tests for the detection of QTL effects (T_1, T_2 and T_3) are described and their power studied for the two types of progeny. T_1 tests the global genetic effects of the QTL (additivity and dominance) and T_2 tests only additive effects assuming dominance is absent when it could exist. The models of these two tests assume that the main effects of QTL alleles are constant in different genetic backgrounds. The additive model of test T_3 considers the six F_2 populations independently, and T_3 is the equivalent of the classical mean comparison test if we neglect dominance; it uses only contrasts between the homozygote marker classes. The results show that T_2 is much more powerful than T_3. The power of T_1 and T_2 depends on the relative sizes of the additive and dominance effects, and their comparison is not easy to establish. Nevertheless, T_2 seems to be the more powerful in most situations, indicating that it is often more interesting to ignore dominance when testing for a QTL effect. For a given size of genetic effects, the power is affected by the total number of individuals genotyped in F_2 and the recombination rate between the marker locus and the putative QTL. The approach presented in this paper has some drawbacks but could be easily generalized to other sizes of diallels and different progeny types.

Key words: Diallel – Restriction fragment length polymorphism markers – Replicated progenies – Linear models – Power of test

Communicated by A. R. Hallauer
Correspondence to: A. Rebai

Introduction

Most traits of economic importance in plants show a continuous variation in phenotype that is the result of the collective action of multiple genetic factors and environmental effects. These factors are located at different quantitative trait loci (QTLs) and, in general, have not been resolved individually. Quantitative genetic studies have dealt largely with a global characterization of genetic factors by developing powerful biometrical approaches.

Sax (1923) was the first to propose the detection of specific genes affecting quantitative traits by studying their associations with marker genes. Numerous contributions to the theory and applications of the marker-based detection of QTLs have since been made (reviewed by Thompson and Thoday 1979), but the restricted number of effective markers rapidly limited this kind of investigation until the development, in the 1980s, of a new biomolecular tool: restriction fragment length polymorphism (RFLP). This technique provides a potentially unlimited number of codominant markers dispersed along the length of the genome (see Beckmann and Soller 1986 for review).

In crop species, most QTL mapping strategies start with the recognition of two parental genotypes, gen-

erally inbred lines homozygous for alternative alleles at the loci of interest, followed by the generation of hybrid F_1 progeny. The classical schemes continue by either selfing the F_1 or backcrossing it to one of the parental lines to generate a segregating F_2 or backcross (BC) population. The latter is then genotyped for the markers and scored for phenotypic performance in order to examine marker-phenotype relationships. Several methods have been advanced in the literature for the detection of linkage between a marker locus and a QTL and their theoretical bases discussed by many authors. Individual marker models, aiming at identifying associations between a single marker locus and a QTL by linear model or maximum likelihood analysis, have been described for various progeny types (Jayakar 1970; Soller et al. 1976; Soller and Genizi 1978; Weller 1986; Luo and Kearsey 1989, 1991). The traditional method uses contrasts among marker genotype means (mean comparison method) or a one-way analysis of variance (ANOVA) to test for the presence of a QTL in the vicinity of the marker locus (Soller et al. 1976; Tanksley et al. 1982; Edwards et al. 1987). It has been widely applied in crop species (Edwards et al. 1987, 1992; Stuber et al. 1987; Weller et al. 1988) and extended to other types of progeny (Bechmann and Soller 1988; Cowen 1988; Ellis 1986; Simpson 1989; Soller and Beckmann 1990). Even if the ANOVA approach (contrary to maxiumum likelihood methods) provides no information about the recombination between markers and putative QTLs, it does provide simple tests the powers of which can be easily compared for various experimental designs (Soller et al. 1976; Knapp and Bridges 1990).

Lander and Botstein (1989) proposed a promising method (interval mapping) based on maximum likelihood analysis that examines intervals between neighbouring markers to provide a confidence interval within which the QTL may be found. Statistical details of their approach and its advantages relative to the ANOVA method in F_2 populations are given by Carbonell et al. (1992) and van Ooijen (1992). Although no complete studies of the power of interval mapping have been done, many authors state that it has a greater power than traditional methods. Whatever the method used for QTL detection [ANOVA, interval mapping, method of Knapp et al. (1990), etc.], the power of a test can be calculated by considering, first, that there is an infinite number of markers along the genome. Consequently, one will calculate the power of a test by assuming that the QTL is on the marker and using an appropriate threshold which takes into account the fact that we will perform this same test all along the genome. In the present case we are concerned with an ANOVA test. In the discussion we give some indications about the exact power of different tests. We can say that the loss in power, compared to the case where we have an infinite number of markers, will be approximately 10–20% and that the proportion between the power of different tests will be conserved.

The aim of this paper is to compare, by analytical calculations, the powers of different linear model tests for QTL mapping using two progeny types derived from a diallel cross between four inbred lines.

Modelling

Assumptions

We consider a diallel cross between four homozygous lines L_1, L_2, L_3, L_4, (without selfings nor reciprocals). The six F_2 populations obtained by self-fertilizing F_1 hybrids are genotyped with RFLPs in order to construct a high-density marker linkage map. F_2 progenies are then selfed and crossed to two testers, which are the complementary parental lines (the F_2 coming from the cross $L_1 \times L_2$ is crossed with L_3 and L_4 as testers). The F_3 and topcross progenies are then grown in replicate and scored for a quantitative trait. We suppose that the number of individuals is the same for each F_2 population, and we denote by N the total F_2 population size. The polymorphic marker locus studied (M) is supposed to have four alleles, M_1, M_2, M_3, M_4. This assumption is quite restrictive because only a limited number of marker loci will have four alleles (about 50% of the RFLP markers in maize [*Zea mays* (L.) if the parents are well sampled] but does facilitate the statistical treatments. The segregation of the markers in each F_2 is supposed to be Mendelian (an F_1 which if M_iM_j gives a F_2 descendance that is $\frac{1}{4}M_iM_i$, $\frac{1}{2}M_iM_j$, $\frac{1}{4}M_jM_j$). We made no assumption about the number of QTL alleles. This point will be discussed later.

Genetic and statistical models

Two models that consider marker-linked effects as being independent of the genetic background are described. A more complete model will be discussed later.

F_3 progenies

We used a linear model neglecting genotype × environment interactions and assuming no epistasis. So, we wrote (1):

$$Y_{ijk.}(M_iM_i) = \mu_{ij} + 2a_i + g_{ijk} + e_{ijk.}$$

$$Y_{ijk.}(M_iM_j) = \mu_{ij} + a_i + a_j + \theta^*_{ij} + g_{ijk} + e_{ijk.}$$

$$Y_{ijk.}(M_jM_j) = \mu_{ij} + 2a_j + g_{ijk} + e_{ijk.}$$

where $Y_{ijk.}(M_iM_i)$ is the phenotypic mean (over all replications) value of the F_3 offsprings of the k^{th} indi-

vidual F_2 deriving from the cross $L_i \times L_j$ and having the genotype M_iM_i; μ_{ij} is an unknown parameter traducing a genetic background-dependent mean; a_i (respectively a_j) is the QTL additive effect linked to marker allele M_i (respectively M_j); θ^*_{ij} is the linked dominance effect between QTL alleles; g_{ijk} is the random genetic effect of loci other than the QTL considered, of variance σ^2_g; and $e_{ijk.}$ is the mean of random environmental effects with expectation 0 and variance σ^2_e. g_{ijk} and $e_{ijk.}$ are independent variables and could be regrouped into a residual of variance, $\sigma^2_{F_3} = \sigma^2_g + (1/L)\sigma^2_e$, where L is the number of replications, assumed the same for each progeny.

Topcross progenies (TC)

Since we know exactly the marker allele coming from the tester, we used the following model (2):

$$Y_{ijtk.}(M_iM_i) = \eta_{ijt} + a_i + a_t + \theta_{it} + h_{ijtk} + e_{ijtk.}$$

$$Y_{ijtk.}(M_iM_j) = \eta_{ijk} + \tfrac{1}{2}(a_i + a_j) + a_t + \tfrac{1}{2}(\theta_{it} + \theta_{jt}) + h_{ijtk} + f_{ijtk.}$$

$$Y_{ijtk.}(M_jM_j) = \eta_{ijt} + a_j + a_t + \theta_{jt} + h_{ijtk} + f_{ijtk.}$$

The notations are almost the same than those in model (1) with η equivalent to μ, h to g and f to e; parameters indexed in t are relative to the tester, and the genotype between brackets corresponds to the F_2 individual crossed to the tester (genotype M_tM_t). Note that $\theta^*_{ij} = \theta_{ij}/4$ (θ_{ij} is the true domainance effect) because the number of heterozygotes for the marker locus is 4 times larger in TC than in F_3. In the same way we note $\sigma^2_{TC} = \sigma^2_h + (1/L)\sigma^2_f$.

The six F_2 populations could be also analysed independently, so that, for an allele M_i, each effect a_i is different according to the cross $L_i \times L_j$ considered and is indexed in (ij). The following model could be used for F_3 progenies: model (3)

$$Y_{ijk.}(M_iM_i) = \mu^*_{ij} + 2a_{i(ij)} + g^*_{ijk} + e^*_{ijk.}$$

$$Y_{ijk.}(M_iM_j) = \mu^*_{ij} + a_{i(ij)} + a_{j(ij)} + \theta^*_{ij} + g^*_{ijk} + e^*_{ijk.}$$

$$Y_{ijk.}(M_jM_j) = \mu^*_{ij} + 2a_{j(ij)} + g^*_{ijk} + e^*_{ijk.}$$

with the notations equivalent to those of model (1).

Parameters estimations

For simplification we will use, in the following paragraphs, the same general linear model notation for models (1) and (2):

$$Y = Z\gamma + r$$

where Y is a $N \times 1$ observation vector, Z is the $N \times p$ design matrix, γ is the $p \times 1$ vector of parameters and r is a $N \times 1$ random vector of residuals with mean 0 and variance $\sigma^2 I$ with σ^2 being $\sigma^2_{F_3}$ for model (1) and σ^2_{TC} for model (2). Z and γ could be partitioned as follows: $Z = [Z_0|Z_1|Z_2]$ and $\gamma' = [\gamma'_0|\gamma'_1|\gamma'_0]$ where $\gamma'_0 = [\mu_{ij}]$ for (1) and $[\eta_{ijt}]$ for (2), $\gamma'_1 = [a_i]_{i=1\ldots4}$, $\gamma'_2 = [\theta_{ij}]_{i=1\ldots4;j=1\ldots4\neq i}$ and x' designates the transpose of matrix or vector x. Since we suppose $\theta_{ii} = 0$ and $\theta_{ij} = \theta_{ji}$ we have 4 parameters a_i, 6 parameters θ_{ij} and 6 parameters μ_{ij} in the F_3 model (1) (16 in all) against 4, 6 and 12 (η_{ijt}), respectively, in the TC model (2) (22 in all). In order to determine the models we used the classical constraints $\sum_{i=1}^{4} a_i = 0$ and $\sum_{j=1}^{4} \theta_{ij} = 0$ which give five independent equations. The transformed Z matrix is written as $X = [X_0|X_1|X_2]$ where X_0 is a $N \times 6$ (respectively $N \times 12$) matrix in the F_3 model (respectively TC model), X_1 is a $N \times 3$ matrix and X_2 is a $N \times 2$ matrix, and γ' as $\beta' = [\beta'_0, \beta'_1, \beta'_2]$ with $\beta_0 = \gamma_0$, $\beta'_1 = [a_1, a_2, a_3]$ and $\beta'_2 = [\theta_{12}, \theta_{13}]$. All these parameters are estimable.

In model (3) there are 6 parameters μ^*, 12 $a_{i(ij)}$ and 6 θ^*_{ij}. Only 6 of the 12 $a_{i(ij)}$ are estimable, and the constraints $a_{i(ij)} + a_{j(ij)} = 0$ were used. For the next developments, parameters of dominance in this model are not considered ($\theta^*_{ij} = 0$).

Hypothesis testing

To test for the presence of a QTL linked to the marker, we chose to test two hypotheses using F_3 or TC progenies:

$H_{01}: \beta_1 = \beta_2 = 0$ (no QTL), the test will be noted as T_1,

$H_{02}: \beta_1 = 0$ under the assumption "$\beta_2 = 0$", test noted as T_2.

The full model: $Y = X\beta + r = X_0\beta_0 + X_1\beta_1 + X_2\beta_2 + r$ (4) is used for test T_1. T_1 tests both additive and dominance effects as a whole. For T_2, a model reduced by the hypothesis $\beta_2 = 0$ is used: $Y = X_0\beta_0 + X_1\beta_1 + r$ (5). T_2 is then the test of model (6): $Y = X_0\beta_0 + r$ against (5) when (4) is true. This means that T_2 tests only additive effects supposing that dominance is absent when it could be present.

For each test we considered the likelihood ratio statistic which is, in our case, the classical F statistic. For T_1, this statistic is (see Graybill 1976):

$$W_1 = \frac{Y'(XX^- - X_0X_0^-)Y}{Y'(I - XX^-)Y}\,\frac{N-p}{q_1}$$

where $X^- = (X'X)^{-1}X'$, $q_1 = 5$ (number of tested parameters), $p = 11$ for F_3 and 17 for TC (total number of estimated parameters) and N, the total number of observations equal to the F_2 population size. Under the alternative H'_{01} ($\beta_1 \neq 0, \beta_2 \neq 0$). $Y'(I - XX^-)Y/\sigma^2$ is distributed as a χ^2 with $(N-p)$ degrees of freedom (noted df) and $Y'(XX^- - X_0X_0^-)Y/\sigma^2$ as a noncentral

χ^2 with q_1 df and a noncentrality parameter

$$\lambda_1 = \frac{E(Y')(XX^- - X_0X_0^-)E(Y)}{q_1\sigma^2}$$

where $E(Y) = X\beta$. These two quadratic forms are statistically independent. W_1 is then distributed as a non-central F random variable with $(q_1, N-p)$ df and a noncentrality parameter λ_1: $W_1 \sim F'(q_1, N-p, \lambda_1)$. Under H_{01}, $W_1 \sim F(q_1, N-p)$. So, if $W_1 > F_\alpha(q_1, N-p)$ we reject H_{01} (at least one QTL is present near the marker locus), where $F_\alpha(q_1, N-p)$ is a critical value from a central F distribution for a siginificance level α.

For T_2, the statistic is

$$W_2 = \frac{Y'(X_{01}X_{01}^- - X_0X_0^-)Y}{Y'(I - X_{01}X_{01}^-)Y}\frac{N-p}{q_2}$$

where $X_{01} = [X_0|X_1]$ and $q_2 = 3$. Under the alternative $H'_{02}(\beta_1 \neq 0$ under $\beta_2 = 0)$, $Y'(X_{01}X_{01}^- - X_0X_0^-)Y/\sigma^2 \sim \chi'^2(q_2, \lambda_2)$ with

$$\lambda_2 = \frac{E(Y')(X_{01}X_{01}^- - X_0X_0^-)E(Y)}{q_2\sigma^2}.$$

$Y'(I - X_{01}X_{01}^-)Y$ could be approximated by its expectation when the number of df is large ($N > 300$). It is easy to show that

$$E\left[\frac{1}{N-p}Y'(I - X_{01}X_{01}^-)Y\right] = \sigma^2 + \frac{1}{N-p}E(Y')(XX^- - X_{01}X_{01}^-)E(Y).$$

It follows that

$$\frac{1}{N-p}Y'(I - X_{01}X_{01}^-)Y \simeq \sigma^2 + \frac{1}{N-p}E(Y')(XX^- - X_{01}X_{01}^-)E(Y).$$

Consequently $W_2 \sim \left(\frac{C}{q_2}\right)\chi'^2(q_2, \lambda_2)$ under H'_{02}, with

$$C = \frac{\sigma^2}{\sigma^2 + \frac{1}{N-p}E(Y')(XX^- - X_{01}X_{01}^-)E(Y)}.$$

Under H_{02}, $Y'(I - X_{01}X_{01}^-)Y/\sigma^2 \sim \chi^2_{N-p}$ and $W_2 \sim F(q_2, N-p)$. So, we reject H_{02} at a significance level α if $W_2 > F_\alpha(q_2, N-p)$.

In model (3), assuming $\theta^*_{ij} = 0$, the tested hypothesis is: H_{03}: $\beta_3 = 0$ where $\beta'_3 = [(a_{i(ij)})i, j]$, and the test is noted T_3. T_3 has the same properties as T_1, and its statistic is distributed as a non-central F with 6 and $(N-p)$ df and a noncentrality parameter λ_3 that could be expressed as λ_1. However, when dominance is absent the heterozygotes in model (3) do not bring any information. T_3 is then equivalent to the classical test (Soller et al. 1976) using contrasts between homozygote marker class means in each F_2 population: $\hat{\delta}_k = Y_{ij\ldots}(M_iM_i) - Y_{ij\ldots}(M_jM_j)$ and $\delta_k = 2a_{i(ij)} - 2a_{j(ij)}$ with $i = 1\ldots4$, $j = 1\ldots4$, $j > i$ and $k = 1\ldots6$. H_{03} is then equivalent to the hypothesis "$\delta_1 = \delta_2 = \ldots = \delta_6 = 0$" and could be written as $H_{03} = H_1 \cap H_2 \cap \ldots \cap H_6$ where H_k: $\delta_k = 0$. The test statistic for a single hypothesis H_k is

$$T_{3k} = \frac{\hat{\delta}_k - \hat{\delta}_k}{\sqrt{\frac{2\hat{\sigma}^2_\omega}{v}}}$$

where σ^2_ω is the within marker genotypes classes variance ($\sigma^2_\omega = \sigma^2$) and v, the number of indivduals in each homozygote marker class (which gives a total number $4v$ for each F_2 population). For $v > 30$, T_k is distributed as a standard normal distributions $N(0, 1)$. So, for a significance level α

T_3 accepts H_{03} if $|T_{3k}| \leq Z_{\alpha/12}$ for every $k = 1\ldots6$.

This means that we reject H_{03} (a QTL is declared present) if at least one H_k is false, i.e. if there is at least one δ_k such as $|T_k| \geq Z_{\alpha/12}$ where $Z_{\alpha/12}$ is the critical value of a two-sided Normal-test for a level $\alpha/6$; to ensure an overall false positive rate of α a significance level of approximately $\alpha/6$ is required for each single test. Note that the extact level for a single test is rather $1 - (1-\alpha)^{1/6} = 0.0085$ for $\alpha = 0.05$, which is nearly equal to $\alpha/6 = 0.00833$.

Power calculations

The power of a test is the probability of rejecting the null hypothesis when its alternative is true. It is a function of the specific values of the parameters. In our case the power of T_1 and T_2 is the probability of detecting an effect when it exists and is a function of the only noncentrality parameter λ. Its expression for T_1 is:

$$\Pi_1(\lambda_1) = Pr(F'(q_1, N-p, \lambda_1) > F_\alpha(q_1, N-p)) = \int_{F_\alpha(q_1, N-p)}^{\infty} F'_x(q_1, N-p, \lambda_1)\,dx.$$

For T_2 the power is calculated by

$$\Pi_2(\lambda_2) = Pr\left(\frac{C}{q_2}\cdot\chi'^2(q_2, \lambda_2) > F_\alpha(q_2, N-p)\right) = Pr\left(\chi'^2(q_2, \lambda_2) > \frac{q_2}{C}F_\alpha(q_2, N-p)\right).$$

Since the power is a direct function of λ, it is important to give an analytical expression of λ according to the

parameters β_1 and β_2. Using linear algebra calculations, we have shown that

$$\lambda_1 = \frac{\beta'_{12}X'_{12}(I - X_0X_0^-)\beta_{12}X_{12}}{q_1\sigma^2}$$

$$\lambda_2 = \frac{\beta'_{12}X'_{12}(I - X_0X_0^-)\beta_{12}X_{12} - \beta'_2X'_2(I - X_{01}X'_{01})\beta_2X_2}{q_2\sigma^2}$$

$$C = \frac{\sigma^2}{\sigma^2 + \frac{1}{N-p}\beta'_2X'_2(I - X_{01}X'_{01})\beta_2}$$

where $X_{12} = [X_1|X_2]$ and $\beta'_{12} = [\beta'_1|\beta'_2]$. So, knowing X, λ_1 and λ_2 could be simply expressed in terms of a_i, θ_{ij} and σ^2. Their expressions are given in Table 1 for both the F_3 and TC models.

For the T_3 test, the power of a single test T_{3k} at a significance level α is

$$\Pi_{3k} = 1 - \beta_k = 1 - Pr\left(-Z_{\alpha/12} - \frac{\delta_k}{\sqrt{2\hat{\sigma}^2_\omega/v}} \right.$$

$$\left. \leq T_{3k} \leq Z_{\alpha/12} - \frac{\delta_k}{\sqrt{2\hat{\sigma}^2_\omega/v}} \right)$$

where β_k is the second-type error. The power of T_3 could be calculated as follows: the second-type error is defined as $\beta = Pr(\text{at least one } \delta_k \neq 0/\delta_1, \delta_2, \ldots, \delta_6) = 1 - Pr(\delta_1 = \delta_2 = \ldots = \delta_6 = 0) = 1 - \Pi_{k=1}^6 Pr(\delta_k = 0)$. $Pr(\delta_k = 0)$ is the second-type error β_k for a single test and is calculated at the specific level $\alpha/6$ by $\int_a^b 1/\sqrt{2\pi}\exp(-x^2/2)\,dx$ where $a = -Z_{\alpha/12} - \delta_k/\sqrt{2\sigma^2_\omega/v}$ and $b = Z_{\alpha/12} - \delta_k/\sqrt{2\sigma^2_\omega/v}$. The power of T_3 is then given by $\Pi_3 = 1 - \beta$, where $\beta = 1 - \Pi_{k=1}^6 \beta_k$ is the global second-type error, so $\Pi_3 = \Pi_{k=1}^6 \beta_k$. When computing the power we supposed that $a_{i(ij)} = a_i$ for all ij. Only 4 parameters a_i were used.

For power computations, we have written a fortran F77 program which for any F_2 population size and each set of parameter values given by the user calculates the power of the five tests: T_3 in F_3, T_1 and T_2 in F_3 and TC. We calculated these powers for different sizes of additive and dominance effects with $N = 600$.

Results

In this section we note by $T_l^{F_3}$ (resp. T_l^{TC}), $l = 1, 2, 3$ the test T_l when it is applied to F_3 (respectively TC) progenies, $\sigma_a^2 = (\sum_{i=1}^4 a_i^2)/4$ and $\sigma_d^2 = \sum_{j=2}^4 \theta_{ij}^2)/3$. σ_a^2 and σ_d^2 represent, respectively, the additive and dominance variances due to the QTL. Such a QTL explains $100(\sigma_q^2/\sigma_q^2 + \sigma^2)$ % of the total pheotypic variance, where $\sigma_q^2 = \sigma_a^2 + \sigma_d^2$ is the total genetic variance due to the QTL. For every QTL underlying the trait, one can show that the additive variance σ_a^2 is 4 times larger in F_3 than in TC. If the number of replications is large ($L > 10$) then $\sigma^2_{TC} \simeq \sigma_h^2$ and $\sigma^2_{F_3} \simeq \sigma_q^2$ and under additivity, $\sigma_g^2 = 4\sigma_h^2$. When dominance is present and the number of replications is small, $\sigma^2_{F_3} < 4\sigma^2_{TC}$ and σ_q^2 in F_3 is less than 4 times larger than in TC. This implies that the part of phenotypic variance explained by a single QTL will be nearly the same for both kinds of progeny. For power calculations, we chose $\sigma^2_{F_3} = 3\sigma^2_{TC}$ and $\sigma^2_{TC} = 1$.

Comparison between T_3 and additive T_2 tests applied in F_3

For 'reasonable' values of θ_{ij} (relative to the phenotypic variance), C will be equal to 1 (Table 1), and then Π_2 is a function of only a_i. For several values of the a_i corresponding to different values of σ_a^2, we have calculated the power of tests T_2 and T_3 applied on F_3 progeny with two significance levels, $\alpha = 0.05$ and $\alpha = 0.001$. The first level occurs when we test the mar-

Table 1. Expressions of decentrality coefficients for tests T_1 and T_2 calculated for both F_3 and TC progenies

For test T_1[a]

TC

$$\lambda_1^{TC} = \frac{n_1}{q_1\sigma^2_{TC}}\left[4\sum_{i=1}^4 a_i^2 + 6\sum_{j=2}^4 \theta_{1j}^2 + \frac{3}{2}(\theta_{13}^2 - \theta_{12}^2) + (\theta_{12} - \theta_{13})(a_2 - a_4) \right]^b$$

F_3

$$\lambda_1^{F_3} = \frac{n_2}{q_1\sigma^2_{F_3}}\left[8\sum_{i=1}^4 a_i^2 + \frac{1}{4}\sum_{j=2}^4 \theta_{1j}^2 \right]$$

For test T_2

TC

$$\lambda_2^{TC} = \frac{n_1}{q_2\sigma^2_{TC}}\left[4\sum_{i=1}^4 a_i^2 + \frac{1}{8}(\theta_{13}^2 - \theta_{12}^2)^2 + (\theta_{12} - \theta_{13})(a_2 - a_4) \right]$$

$$C = \frac{\sigma^2_{TC}}{\sigma^2_{TC} + \frac{n_1}{N}(4.25\theta_{12}^2 + 7.25\theta_{13}^2 + 6.125\theta_{14}^2)}$$

F_3

$$\lambda_2^{F_3} = \frac{n_2}{q_2\sigma^2_{F_3}}\left[8\sum_{i=1}^4 a_i^2 \right]$$

$$C = \frac{\sigma^2_{F_3}}{\sigma^2_{F_3} + \frac{n_2}{N}\sum_{j=2}^4 \theta_{1j}^2/4}$$

[a] n_1 and n_2 are the number of individuals in each marker genotype class in TC and F_3 progenies, respectively. In this case $N/n_1 = 48$ ($4n_1$ individuals in each of the 12 marker genotype classes) and $N/n_2 = 24$, so $n_2 = 2n_1$

[b] $a_4 = -(a_1 + a_2 + a_3)$ and $\theta_{14} = -(\theta_{12} + \theta_{13})$

kers individually, at a precise location in the genome, for a 5% false positive rate. To test on the entire genome with the same global rate of false positives, we used an approximate level of 0.001 at each single marker (based on approximations for a high-density linkage map; see Lander and Botstein 1989).

Figure 1 shows that $T_2^{F_3}$ is always much more powerful than $T_3^{F_3}$. With $\alpha = 5\%$, the power of $T_2^{F_3}$ rapidly reaches high values (> 0.9) for effects more than 5%, whereas Π_3 equals 0.9 for an effect of 11%. With $\alpha = 0.001$, the size of detectable effects with a given power is considerably reduced (Fig. 1); effects of about 9% and 20% are needed to achieve 90% power for $T_2^{F_3}$ and $T_3^{F_3}$, respectively. Variation in the minimum size of detectable genetic effects with a power of 90% according to the number of F_2 individuals, N, is represented in Fig. 2. For a given false positive rate, the required progeny size (for 90% power) scales are essentially conversely correlated with the variance explained by the QTL, i.e. the size of the effects detected with a given power decreases when the population size increases. We have calculated that approximately 6300 F_2 individuals are required to detect a QTL having an effect of about 1% of the total phenotypic variance using $T_3^{F_3}$, with a significance level $\alpha = 5\%$ and a power of 90%. This corresponds exactly to the number given by Soller et al. (1976) (1050 for a single F_2 population) in the case of complete linkage between marker and QTL. The disadvantage of $T_3^{F_3}$ compared to $T_2^{F_3}$ is due to the fact that the former uses only the contrasts between the homozygote classes and does not exploit the repetitions (six populations and four parents) of additive effects in the contrasts.

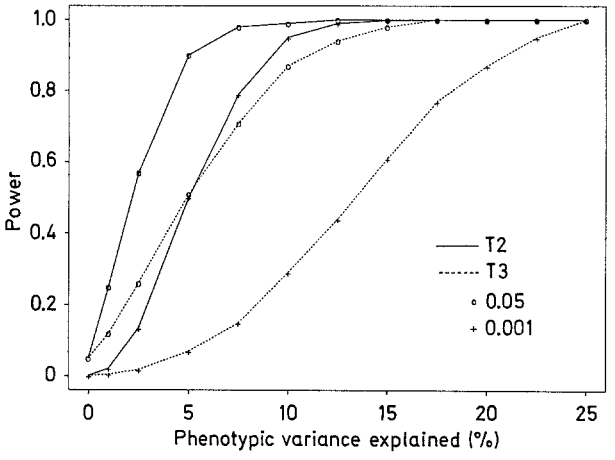

Fig. 1. Power of tests $T_2^{F_3}$ $T_3^{F_3}$ according to the percentage of phenotypic variance explained by the QTL, with two significance level $\alpha = 0.05$, 0.001 and a population size $N = 600$

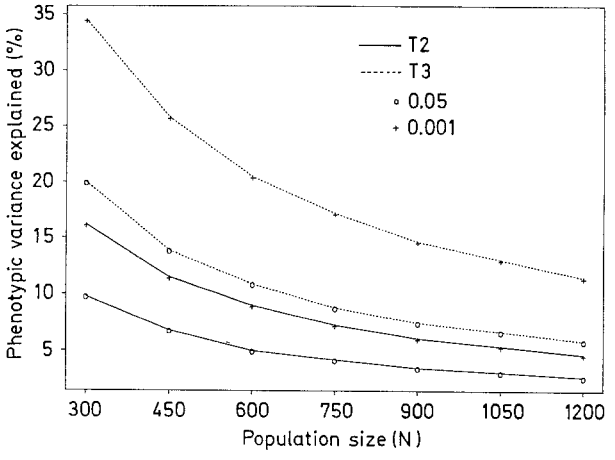

Fig. 2. Genetic effect sizes (expressed in % of phenotypic variance) detectable with 90% power by $T_2^{F_3}$ and $T_3^{F_3}$ according to the F_2 population size (N), for two significance levels $\alpha = 0.05$, 0.001

Comparison between tests T_1 and T_2 applied in F_3 and TC

The decentrality parameters (Table 1) for the F_3 tests are functions of σ_a^2 and σ_d^2 only because of the orthogonality of the model contrasts. In the TC case, the parameter spaces are no longer orthogonal, and λ varies according to σ_a^2, σ_d^2 and non-orthogonality terms including products between a_i and θ_{ij}. If an additive model ($\theta_{ij} = 0, C = 1$) is assumed a direct comparison between T_1 and T_2 is possible (by comparing their λ). In fact, for both progenies $\lambda_1 = \frac{3}{5}\lambda_2$ and T_2 is better than T_1. Even in the case where $\lambda_1 = \lambda_2$, T_1 will be less powerful than T_2 because the test T_2 consummates fewer degrees of freedom (only 3). Moreover, $\lambda_1^{TC}/\lambda_1^{F_3} = \sigma_{F_3}^2/4\sigma_{TC}^2$, which gives, under additivity and large number of replications, $\lambda_1^{TC} = \lambda_1^{F_3}$. Generally, $\sigma_{F_3}^2$ is less than $4\sigma_{TC}^2$, and $T_1^{F_3}$ would be, in most genetic situations, more powerful than T_1^{TC}.

Comparison between T_1 and T_2 was achieved by trying many sets of parameter values that cover several genetic situations. $T_2^{F_3}$ appears to have the greatest power unless dominance values are excessively large relative to additive effects (data not shown). With TC progenies, T_2 seems better than T_1 for small values of dominance relative to additive effects. T_1 becomes more powerful than T_2 if the QTL expresses very strong dominance effects ($\sigma_d^2 \gtrsim 4\sigma_a^2$) as shown in Figs. 3–5. We also verfied that T_2 is still more powerful than T_1 when we use recombinant inbred lines derived from the same diallel. We think that the superiority of the T_2 test is related to the experimental design used and would be observed to be more or less large according to the quality of the estimation of dominance effects allowed by the design. This would explain why, in F_3, T_2 stays more powerful than T_1 for values of dominance larger than those in TC.

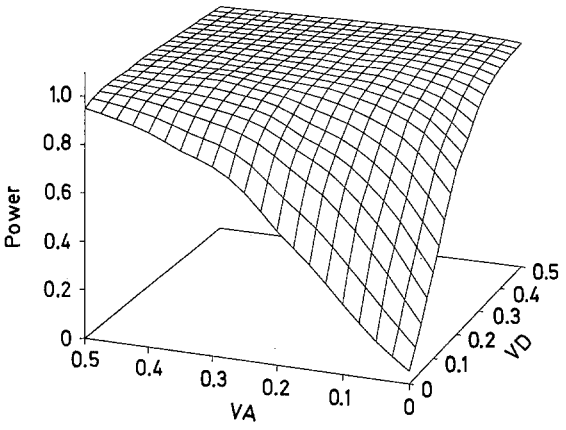

Fig. 3. Variation in the power of T_1^{TC} according to additive (VA) and dominance (VD) variances explained by the QTL, with a significance level $\alpha = 5\%$ and $N = 600$

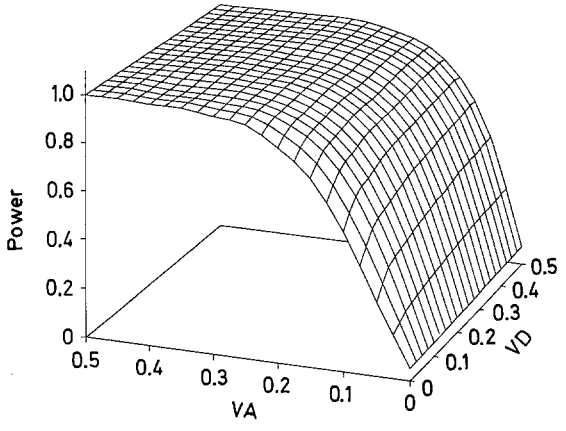

Fig. 4. Variation in the power of T_2^{TC} according to additive and dominance variances explained by the QTL, with a significance level $\alpha = 5\%$ and $N = 600$

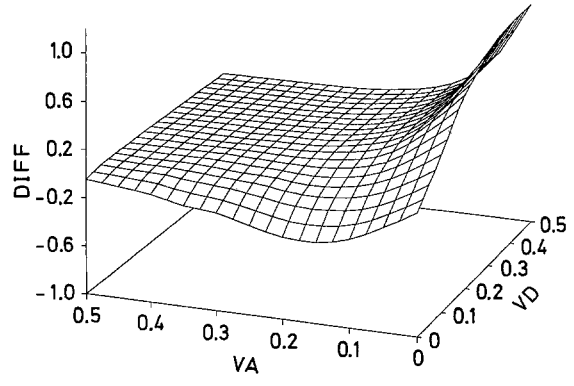

Fig. 5. Variation in the power difference DIFF $= [Pow(T_1^{TC}) - Pow(T_2^{TC})]$ according to additive and dominance variances explained by the QTL, with a significance level $\alpha = 5\%$ and $N = 600$

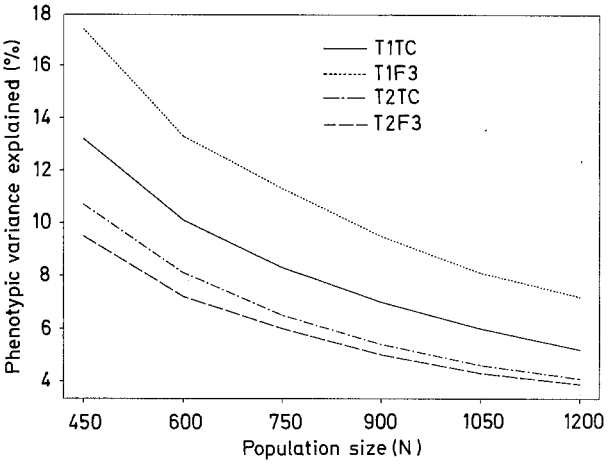

Fig. 6. Genetic effect sizes (expressed in % of phenotypic variance detectable with 90% power by T_1 and T_2 according to the F_2 population size (N), with a significance level $\alpha = 5\%$

$T_2^{F_3}$ seems to be the most powerful test in many situations but there is not an absolutely best test. We can say that tests for QTL detection will be often more powerful with F_3 progenies than with TC, but TC would provide more accurate estimations of dominance effects. When the environmental variance for the trait considered is large (low heritability) and the number of replications too small, $\sigma^2_{F_3}$ gets closer to σ^2_{TC}, and the difference in power between the F_3 and TC tests will be reduced, especially if dominanceis strong. Figure 6 shows the evolution of minimum detectable effects required to achieve 90% power according to the total individual number in the F_2 in the case: $\sigma^2_d = \sigma^2_a/2$. We see that T_2 is better than T_1 whatever the kind of descendance we use and that $T_2^{F_3}$ is the best test.

The results of this section suggest that taking dominance into account when testing QTL effects corresponds to a loss of power (unless dominance is excessively large). So, it is almost always better to consider only additive effects when mapping QTLs (at least in a first approach) in a diallel cross by using tests like T_2. It would be interesting to see if this property (of test T_2) holds for other designs (such as a single F_2 population) and other other types of tests, especially for the likelihood ratio test used by Lander and Botstein (1989).

Discussion

We have considered that the additive effect linked to a marker allele is the same in all the genetic backgrounds. However, if we consider that additive effects are different according to the context, the models become much more complicated. Considering only additivity, one can see that the complete model, equivalent to model

(3) (each allele M_i has three additive linked-effects, one for each context), could be reduced to a model with interaction if we subdivide the additive effects into a constant term (whatever the context is) and a deviation term due to the context (interaction). So, for F_3 one could write $a_{i(ij)}$, the additive effect linked to allele M_i in the genetic context ij as follows: $a_{i(ij)} = \alpha_i + d_{ij}$ where α_i is the mean additive effect and d_{ij}, the interaction term. The results of the previous paragraph show that even when dominance is present neglecting it does not decrease dramatically the power of test T_2. As we have several reasons to think that the interaction term d_{ij} is not large relative to α_i, the same kind of results apply. Therefore we can consider that the additive effect of an allele is the same in different genetic backgrounds without risking an appreciable loss in power.

With the classical approach (Soller et al. 1976), for a single F_2 population derived from a cross between two inbred lines of genotypes $M_1M_1Q_1Q_1$ and $M_2M_2Q_2Q_2$, where M is a marker locus (with two codominant alleles M_1, M_2) and Q a QTL (alleles Q_1, Q_2), it is easy to prove that $\overline{M_1M_1} - \overline{M_2M_2} = (1-2r)(\mu_{11} - \mu_{22})$ where $\overline{M_1M_1}, \overline{M_2M_2}$ are means of marker genotype classes M_1M_1, M_2M_2, respectively, μ_{11}, μ_{22} are means of QTL genotypes Q_1Q_1, Q_2Q_2, respectively, and r, the recombination rate between M and Q. This yields a confounding between recombination and QTL effects that affects the power of the test. Power decreases for an effect of fixed size as r increases, and the total number of progenies required for the same power is increased by a factor $1/(1-2r)^2$.

In our statistical models we are confronted with the same type of loss in power when using the statistics T_1, T_2 and T_3 at each of the positions of the markers. In this case, genetic effects (a_i and θ_{ij}) are multiplied by a factor $(1-2r)$, and the decentrality parameter λ is multiplied by $(1-2r)^2$. For a recombination rate more than 0.15–0.20 the power of the tests decreases considerably. We have calculated that with $r = 0.2$, power is reduced of about 50% (data not shown). We believe that the use of interval mapping method or flanking marker models (Lander and Botstein 1989; Knapp et al. 1990) to construct tests similar to T_1, T_2 and T_3 will limit the loss of power due to recombination between markers and QTL. In a backcross, Lander and Botstein (1989) showed that power is increased by about 20% when we use interval mapping rather than single marker tests for a QTL in the middle of an interval between two markers at $d = 25$ cM ($r \simeq 0.2$ according to Haldane's mapping function) from each other. We think that for such markers and QTL, the power of the 'interval mapping version' of tests T_1, T_2 and T_3 will be about 25% less than that calculated in this paper (supposing the QTL on the marker).

We have also assumed that there are four alleles at each marker locus. In reality, most markers would have two or three alleles, and the power would consequently decrease. The strains should be chosen to differ in the quantitative trait studied and to carry different alleles at a large number of RFLP markers; in particular, the choice could be based on classifications of lines according to publicly known RFLP markers that are now available for many maize inbreds (e.g. Smith et al. 1990). The power of T_1 and T_2 depends essentially on the additive and dominance variances, regardless of the number of alleles at the QTL. However, a QTL with two alleles will segregate in at least three of the six F_2 populations and would explain a smaller part of the total phenotypic variance (depending on the genotypes of parent lines). It would therefore be detected with less power. With TC progenies, as we can see in Table 1, the power depends also on the distribution of marker-linked effects (a_i, θ_{ij}) i.e. on the number of QTL alleles. The comparison between tests presented above remains fairly valid whatever the number of alleles at the QTL.

In this paper we have presented a preliminary approach using linear model tests to detect QTL effects in a diallel cross between four inbred lines. We have demonstrated that these tests are more powerful than the test considering the crosses independently. We have also proven that the test assuming the absence of dominance effects even when they exists is, in general, more powerful than the test of global effects (both additive and dominance). This indicates that it is better to neglect dominance when testing for the presence of a QTL. If this hypothesis is accepted one can then estimate and test dominance effects. The major advantage of diallel schemes (as the one we have described) over many independent F_2 populations is that they provide more powerful tests. They seem to be rather efficient approaches especially when one is interested in detecting QTLs with consistent effects across populations.

In spite of its several imperfections, our approach has the advantage of being easy to generalize to other order diallels (if we suppose that each parent line has a different allele at the marker loci considered, which is an acceptable premise for diallels of sizes fewer than 5) and other replicated progenies, such as recombinant inbred lines, doubled haploid lines and top-cross progenies with testers at several stages of the scheme. Power computations of different tests will enable the best one for a particular scheme to be chosen and the efficiency of different experimental strategies to be compared. There are several possible refinements to the approach of this paper. Particularly, the use of flanking marker models would provide estimates of putative QTL parameters that are more efficient and unbiased by the recombination. In addition, with these models, the problem of markers having fewer than four alleles could be quite easily solved. At the present time we are trying to develop an interval mapping method for diallel schemes that would permit the many short-

comings of the models proposed in this paper to be overcome.

Acknowledgments. We are grateful to Patrick Vincourt (Research Director, Rustica) for his scientific suspport and helpful comments on this work. This research was partly subsidized by Rustica Semences (Domaine de Sandreau, 31700 Mondonville), which since 1989 has been carrying out a four-parent diallel in maize for QTL mapping and which is financing the PhD thesis of the first author.

References

Beckmann JS, Soller M (1986) Restriction fragment length polymorphism in plant genetic improvement. In: Miflin BJ (ed) Oxford surveys of plant molecular and cell biology vol 3. Oxford Press, Oxford, pp 197–250

Beckmann JS, Soller M (1988) Detection of linkage between marker loci and loci affecting quantitative traits in crosses between segregating populations. Theor Appl Genet 76:228–236

Carbonell EA, Gerig TM, Balansard E, Asins MJ (1992). Interval mapping in the analysis of nonadditive quantitative trait loci. Biometrics 48:305–315

Cowen NM (1988) The use of replicated progenies in marker-based mapping of QTLs. Theor Appl Genet 75:857–862

Edwards MD, Stuber CW, Wendel JF (1987) Molecular-marker-facilitated investigation of quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action. Genetics 116:113–125

Edwards MD, Helentjaris T, Wright S, Stuber CW (1992) Molecular-markar-facilitated investigations of quantiative trait loci in maize. 4. Analysis based on genome saturation with isozyme and restriction fragment length polymorphism markers. Theor Appl Genet 83:765–774

Ellis THN (1986) Restriction fragment length polymorphism markers in relation to quantitative characters. Theor Appl Genet 72:1–2

Graybill FA (1976) Theory and application of the linear model. Wadsworth, Belmont, Calif.

Jayakar SD (1970) On the detection and estimation of linkage between a locus influencing a quantitative character and a marker locus. Biometrics 26:466–479

Knapp SJ, Bridges WC (1990) Using molecular markers to estimate quantitative trait locus parameters: power and genetic variances for unreplicated and replicated progeny. Genetics 126:769–777

Knapp SJ, Bridges WC, Brikes D (1990) Mapping quantitative trait loci using molecular marker linkage maps. Theor Appl Genet 79:583–592

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Luo ZW, Kearsey MJ (1989) Maxiumum likelihood estimation of linkage between a marker gene and a quantitative trait locus. Heredity 63:401–408

Luo ZW, Kearsey MJ (1991) Maximum likelihood estimation of linkage between a marker gene and a quantitative trait locus. II. Application to backcross and doubled haploid populations. Heredity 66:117–124

Ooijen JW van (1992) Accuracy of mapping quantitative trait loci in autogamous species. Theor Appl Genet 84: 803–811

Sax K (1923) The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgarus*. Genetics 8:552–560

Simpson SP (1989) Detection of linkage between quantitative trait loci and restriction fragment length polymorphisms using inbred line. Theor Appl Genet 77:815–819

Smith OS, Smith JSC, Bowen SL, Tenborg RA, Wall SJ (1990) Similarities among a group of elite maize imbreds as measured by pedigree, F_1 grain yield, grain yield, heterosis and RFLPs. Theor Appl Genet 80:833–840

Soller M, Beckmann JS (1990) marker-based mapping of quantitative trait loci using replicated progenies. Theor Appl Genet 80:205–208

Soller M, Genizi A (1978) The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating populations. Biometrics 34:47–55

Soller M, Genizi A, Brody T (1976) On the power of experimental designs for the detection of linkage between marker loci and quantiative loci in crosses between inbred lines. Theor Appl Genet 47:35–39

Stuber CW, Edwards MD, Wendel JF (1987) Molecular-marker-facilitated investigations of quantitative-trait loci in maize. II. Factors influencing yield and its component traits. Crop Sci 27:629–648

Tanksley SD, Medina-Filho H, Rick CM (1982) Use of naturally-occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49:11–25

Thompson JN, Thoday JM (1979) Quantitative genetic variation. Academic Press, London

Weller JI (1986) Maximum likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic markers. Biometrics 42:627–640

Weller JI, Soller M, Brody T (1988) Linkage analysis of quantitative traits in an interspecific cross of tomato (*Lycopersicon esculentum* × *Lycopersicon pimpinellifolium*) by means of genetic markers. Genetics 118:3229–339