DETECTING LINKAGE BETWEEN A FULLY-INFORMATIVE MARKER LOCUS AND A TRAIT LOCUS IN OUTBRED POPULATIONS USING ANALYSIS OF VARIANCE

S. Kumar[1,2], S. D. Carson[3] & D. J. Garrick[2]

[1)] N. Z. Forest Research Institute, Private Bag 3020, Rotorua, New Zealand
[2)] Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Private Bag 11222, Palmerston North, New Zealand
[3)] Fletcher Challenge Forests Limited, RD 2, Whakatane, New Zealand
Correspondence: S. Kumar, Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Private Bag 11222, Palmerston North, New Zealand., E-mail: S.Kumar@massey.ac.nz

Received June 4, 1998; Accepted September 23, 1999

ABSTRACT

Analysis of variance can be used to detect the linkage of segregating quantitative trait loci (QTL) to molecular markers in outbred populations. Given a single fully-informative (FI) marker for independent full-sib families (with marker configuration: $M_1M_2 \times M_3M_4$) and assuming linkage equilibrium, variance components were derived to predict the power of detection of a QTL. These variance components are based on hierarchical analysis of variance assuming a completely random model. Formulae that relate power to the recombination frequency (r) between FI marker and the QTL, genetical properties of the quantitative trait controlled by the QTL and the design parameters are developed. The predicted powers using the FI marker configuration were compared to that obtained using pseudo-backcross (PBC : $M_1M_2 \times M_1M_1$) and pseudo-intercross (PIC : $M_1M_2 \times M_1M_2$) marker configurations. The effect of dominance properties of the QTL on power were also examined. The reliability of theoretical approximation of power was confirmed by computer simulations. The results showed that FI marker design is more efficient than PBC and PIC marker designs, few large families are better than many small families. Incomplete linkage and dominance of the QTL showed large effects on the power.

Key words: fully-informative marker, genetic linkage, statistical power, QTL

INTRODUCTION

The use of molecular markers as a complementary tool for breeding is based on linkage disequilibrium between marker and quantitative trait loci (QTL) involved in the control of quantitative characters. In most agricultural crops, inbreeding is followed by crossing between inbred lines to create disequilibrium for QTL detection. From population genetic studies, it is known that wild allogamous species like forest trees are often in linkage equilibrium and because of the long generation intervals and inbreeding depression, it is difficult to obtain inbred lines for QTL mapping experiments. Linkage disequilibrium between a marker and a linked QTL, however, can be found within families in outcross populations and it is increasingly common to carry out QTL detection in a pedigree of full or half-sib family.

In an outbred population different marker alleles will likely be associated with the same QTL allele in different families. Therefore, evidence for a linked QTL cannot be obtained at a population level from overall mean differences between marker genotypes. Using an hierarchical ANOVA, marker effects need to be analysed in each family separately and the test for a linked QTL comes from the comparison of the between-marker within-family mean squares with the residual mean squares and can be tested as an F-ratio (HILL 1975, SOLLER & GENIZI 1978). Under the null hypothesis (marker is not linked to the QTL i.e., recombination (r) between marker and QTL is 0.5), this ratio is distributed as a central F-variable; whereas this ratio will be a noncentral F-variable when r is less than 0.5 (JAYAKAR 1970, LUO 1993). Hence, given the pedigree structure, it is possible to predict the power of detection of a given QTL (HILL 1975, SOLLER & GENIZI 1978, LUC 1993, KNOTT 1994).

The informative full-sib families considered in previous simulation studies of outbred populations (HILL 1975, SOLLER & GENIZI 1978, LUO 1993, KNOTT 1994) were of two types with respect to the marker genotypes of the parents. First, those where one parent

is homozygous at the marker locus and one is heterozygous (pseudo-backcross or PBC families) and second, those where both parents are heterozygous for the same genotype at the marker locus (pseudo-intercross or PIC families). Both of these strategies suggest that only two alleles are segregating in full-sib progeny. However, for an outbred *Pinus radiata* pedigree, as many as four alleles may be segregating at a locus. With the continued development of multiallelic codominant markers (for example, microsatellites), the exclusive use of fully-informative markers (i.e., $M_1M_2 \times M_3M_4$) is becoming possible. This creates an additional family type (fully- informative or FI families) with respect to the markers; that is, one where all four marker genotypes can be distinguished in the offspring.

HILL (1975) and LUO (1989, 1993), assuming a bi-allelic marker and the linked bi-allelic QTL, derived the expressions for expected variances for ANOVA of PBC and PIC marker designs in a segregating population. MURANTY (1996) derived the expressions for noncentrality parameter of different mating schemes assuming FI marker design. However, these expressions could not relate the power directly to different recombination rates and non-additive gene action at QTL. So far, the theoretical expressions of variance components, to predict the power of using FI marker type families using ANOVA, have not been derived.

The present study was focused on deriving the expressions of expected variances for ANOVA of a FI marker design ($M_1M_2 \times M_3M_4$), in two-generation pedigrees of outbred populations, and relating the power directly to genetic parameters at the QTL and the relevant design parameters. This will allow factors affecting the power to be investigated comprehensively. A second objective of this study was to compare the power obtained from using FI marker design to that obtained from PBC and PIC type marker strategies.

THEORY

Basic assumptions and experimental design

The underlying assumptions of the method are those commonly made by researchers. The method involves analysing progeny from controlled mating in a population. Two autosomal loci are considered, one of them affects a quantitative trait (QTL) while the other is a fully-informative marker. The two loci are linked with a recombination frequency of r ($s = 1 - r$). Let the frequency of allele Q_1 at the QTL be denoted as p ($q = 1 - p$), and the phenotypic distributions of the 3 genotypes at the QTL i.e., Q_1Q_1, Q_1Q_2 and Q_2Q_2 are assumed to be $N(a, \sigma^2)$, $N(d, \sigma^2)$ and $N(-a, \sigma^2)$ respectively, where a and d represent the additive and dominance effects at the QTL. With just one QTL and no other gene effects, σ^2 will be only the environmental variance, whereas in the presence of unlinked QTLs, it will also include genetic variance at these loci (i.e., polygenic variance).

Let the parental genotypes at the marker locus be M_1M_2 and M_3M_4 and four marker genotype classes ($M = 4$) are distinguishable in the offspring: M_1M_3, M_1M_4, M_2M_3 and M_2M_4 segregating with a 1:1:1:1 ratio. We assume that the QTL and the marker gene are in linkage equilibrium in the population. Let n_{ij} denote the number of sibs within the j^{th} marker class within the i^{th} sibship. Also each sibship (N_f) has a constant size of N_o and thus the total experimental size is $N_f \times N_o$.

Statistical Model

The linear model for the phenotype of the quantitative trait measured on the k^{th} sib ($k = 1, 2, \ldots, n_{ij}$) with the j^{th} marker genotype ($j = 1, 2, \ldots, M = 4$) within the i^{th} sibship ($i = 1, 2, \ldots, N_f$) can be written as:

$$y_{ijk} = \mu + \alpha_i + \beta_{ij} + e_{ijk} \qquad [1]$$

where μ is an overall mean, α_i, β_{ij} and e_{ijk} are contributions from the sibship, from the marker genotype within sibship and from within-marker within-sibship residual, respectively. They are assumed to be independently and normally distributed with zero means and variances σ^2_α, σ^2_β, and σ^2_e, respectively. Similar assumptions have been made in several studies (e.g., HILL 1975, LUO 1993, LYNCH & WALSH 1997). The assumption of functional independence of quantitative trait from the marker locus was also made. We have considered only one QTL and all other " background" genetic variation is considered as environmental. The ANOVA for this model is given in Table 1.

Under the assumption of a constant size of sibship (N_o) and

$$n_{i1} : n_{i2} : n_{i3} : n_{i4} \approx 1 : 1 : 1 : 1 \qquad [2]$$

the approximation for n_o will be:

$$n_o \approx 0.25\, N_o. \qquad [3]$$

The expression for mean squares and the general version of n_0 (Table 1) can be found, for example, in HILL (1975). All possible marker-QTL genotypes of parents and the gametes inherited by the offspring are given in Table 2 with their probabilities. Using these probabilities, the expected values of the quantitative

Table 1. ANOVA for a two-factor completely nested design.

Source	Degrees of freedom	MS	EMS
Between sibships	N_f -1	MS_s	-
Between marker genotypes within sibship	$\Sigma(M_i - 1)$	MS_m	$\sigma_e^2 + n_0\sigma_\beta^2$
Within marker genotype within sibship	$\Sigma(n_{ij} - 1)$	MS_w	σ_e^2

Table 2. Probabilities of various gametes inherited from parents to progeny. Parental marker genotypes are: $M_1M_2 \times M_3M_4$. We assumed that the QTL and the marker genes are in linkage equilibrium in the population. Recombination rate between the marker and the QTL is r.

Parental genotype	From first parent			
	Gametes			
	M_1Q_1	M_1Q_2	M_2Q_1	M_2Q_2
M_1Q_1 / M_2Q_1	1/2	0	1/2	0
M_1Q_1 / M_2Q_2	$(1 - r) / 2$	$r/2$	$r/2$	$(1 - r) / 2$
M_1Q_2 / M_2Q_1	$r/2$	$(1 - r)/2$	$(1 - r) / 2$	$r/2$
M_1Q_2 / M_2Q_2	0	1/2	0	1/2
Parental genotype	From second parent			
	Gametes			
	M_3Q_1	M_3Q_2	M_4Q_1	M_4Q_2
M_3Q_1 / M_4Q_1	1/2	0	1/2	0
M_3Q_1 / M_4Q_2	$(1 - r) / 2$	$r/2$	$r/2$	$(1 - r) / 2$
M_3Q_2 / M_4Q_1	$r/2$	$(1 - r) / 2$	$(1 - r) / 2$	$r/2$
M_3Q_2 / M_4Q_2	0	1/2	0	1/2

trait value, y, were obtained for different marker genotypes within sibship (Appendix 1). Similarly, the variances of the trait value within sibship within marker genotypes were derived (Appendix 2). Finally, the variance between marker genotype classes within sibship were obtained (Appendix 3).

The variance expressions, given in Appendix 3, were averaged by using the corresponding probabilities as weights and it gives us the expected variance between marker genotypes within sibships (σ_β^2) as:

$$\sigma_\beta^2 = pq[(s-r)^2[(d+a)^2 - 2pqd^2 - 4pad] + 2pqd^2[(s^2+r^2)^2 + 4r^2s^2 - 0.5]] \quad [4]$$

Using Appendix 2, first the average variance within each marker genotype was obtained by using the corresponding probabilities as weights. After this, the variances within each marker genotypes were averaged using equal probabilities (because four marker genotypes are assumed to be segregating with a 1:1:1:1 ratio) and it gives us the expected variance within marker genotypes within sibships (σ_e^2) as:

$$\sigma_e^2 = \sigma^2 + 4pqrs[p^2(a-d)^2 + q^2(a+d)^2 + 2pqa^2 + pqd^2(1-2rs)] + 2p^2q^2d^2(r^2+s^2)[1-(r^2+s^2)] \quad [5]$$

The rational used for the derivation of [4] and [5] is similar to that of HILL (1975). From equation [4] it can be easily shown that the expected variance between marker genotypes within sibship (σ_β^2) will be zero if there is no linkage between the marker and the QTL, i.e., $r = s = 0.5$. Under the null hypothesis ($H_0 : r = 0.5$) the ratio MS_m / MS_w has an expected value of 1 and is distributed as a central F-variable; whereas this ratio has an expected value of more than one and will be a noncentral F-variable when r is less than 0.5 (JAYAKAR 1970, LUO 1993). Using the standard definition, the

power function for linkage detection with the design under study (FI families) can be written in the following general form:

$$\text{Power} = Pr\,[F_{(v1,\,v2;\,\delta)} > F_{(\alpha;\,v1,\,v2)}\,], \qquad [6]$$

where $F_{(v1,\,v2;\,\delta)}$ is a noncentral F-variable with degrees of freedom v_1 and v_2 and noncentrality parameter δ, while $F_{(\alpha;\,v1,\,v2)}$ is the upper α point of a central F-variable with degrees of freedom v_1 and v_2. The value of noncentrality parameter, δ, was calculated as (LUO 1993):

$$\delta = (M \times N_f - 1)\, n_o \sigma_\beta^2 / \sigma_e^2 \qquad [7]$$

Power calculation

The power of a test is defined as the probability of rejecting the null hypothesis when its alternative is true. The power of QTL mapping experiment is the probability that the null hypothesis (no linked QTL) is rejected when its alternative (presence of a linked QTL) is true. The formulae developed for expected variance between marker genotypes within sibships (σ_β^2) and the expected variance within marker genotypes within sibships (σ_e^2) were used in theoretical prediction of the powers of QTL detection for a wide range of combinations of parameters (i.e., genetic parameters at the QTL and design parameters). In order to derive the parameters, the total genetic variance, V_G, arising from one locus (QTL) can be written as (FALCONER 1989):

$$\begin{aligned} V_G &= V_A + V_D \\ &= 2pq[a + d(q - p)]^2 + [2pqd]^2 \\ &= 2pq[a^2 + (1 - 2pq)d^2 + 2(q - p)ad] \end{aligned} \qquad [8]$$

By assuming the phenotypic variance (V_P) to be unity, the V_G (or QTL variance) becomes the broad-sense heritability (H^2) at the QTL. Also, $\sigma^2 = 1 - V_G$. To determine the value of parameter a and d at the QTL, we take following steps (LUO 1993):
Assume the dominance ratio (f) = d/a, then

$$a = \sqrt{\frac{V_G}{2pq[1 + (1 - 2pq)f^2 + 2(q - p)f]}}, \text{ and} \qquad [9]$$

$$d = f \times a \qquad [10]$$

Using different combinations of design parameters (N_f and N_o), genetic parameters at the QTL (p, f, and H^2) and recombination frequency (r), the noncentrality parameter can be calculated. After that power can be easily calculated using [6].

Power evaluation from simulations

Since approximations [2] and [3] were made in deriving the power function, the reliability of these approximations was checked by comparing the theoretical predictions of the power to the powers calculated from simulation experiments. A program was written in SAS(1989) for simulating the inheritance of marker-QTL linkage for any combination of experimental design and genetic parameters. The simulated data was analysed using SAS PROC GLM and the frequency of significant F-values in replicated simulation trials was calculated as in CARBONELL *et al.* (1992) and LUO (1993), which gives the empirical power.

Comparison of power

The power, calculated using FI families ($M_1M_2 \times M_3M_4$) in this study, were compared to those obtained from using the families where parents are $M_1M_1 \times M_1M_2$ (PBC families) or $M_1M_2 \times M_1M_2$ (PIC families). The power for these two designs (PBC and PIC) in a segregating population were evaluated by LUO (1993). The results for the PBC and PIC type marker configurations in our study are solely based on the formulae derived by LUO (1993).

RESULTS

Theoretical powers of linkage detection were calculated for wide range of genetic parameters at the QTL and design parameters. Empirical powers, based on 500 replications, are presented along with those obtained from theoretical approximation. When assuming gene action at the QTL to be purely additive, the power of QTL detection for three types of marker loci varies substantially (Table 3). The power is the highest with FI markers (both parents have different heterozygote genotypes at marker locus) and is lowest for PIC markers (parents are heterozygous for the same genotype at marker locus). The power of linkage detection increases as the number of offspring per family increase. Keeping the number of offspring genotyped fixed (say, 1000), then having fewer larger families clearly increases power relative to many small families.

The various levels of genetic variance or the broad-sense heritability at the QTL (H^2) and different recombination rates between marker and the QTL has significant impact on power of QTL detection for all three marker configurations (Table 4). As the heritability at the QTL increase the power also increases but a decreasing trend in power was obtained for a increase in

Table 3. Theoretical prediction (PR) of powers of 3 marker designs for a QTL that has a heterozygosity of 50%, for various number of families (N_f), various number of offspring per family (N_o). The other assumptions were: broad-heritability at the QTL (H^2) = 0.05, recombination rate (r) = 0.10, type-I error = 0.01 and dominance ratio (f) = 0.0. The powers evaluated from simulation experiments (SI) are also given. PBC = pseudo-backcross, PIC = pseudo-intercross, FI = fully-informative.

N_f	N_o	PBC		PIC		FI	
		PR	SI	PR	SI	PR	SI
5	50	0.04	0.05	0.03	0.04	0.05	0.05
	100	0.11	0.12	0.07	0.08	0.15	0.15
	200	0.29	0.32	0.18	0.24	0.46	0.48
10	50	0.07	0.07	0.04	0.05	0.09	0.07
	100	0.21	0.21	0.12	0.12	0.30	0.28
	200	0.57	0.55	0.36	0.39	0.79	0.79
20	50	0.13	0.12	0.07	0.08	0.17	0.17
	100	0.41	0.41	0.22	0.22	0.58	0.57
	200	0.87	0.83	0.66	0.67	0.98	0.98

Table 4. Comparison of theoretically predicted (PR) powers of linkage detection of 3 marker designs for varying number of families (N_f) and number of offspring per family (N_o) where H^2 and r represent the broad-sense heritability at the QTL and recombination frequency between marker and the QTL. The other assumptions were: type-I error = 0.01, dominance ratio (f) = 0.0 and p = 0.50. The powers evaluated from simulation experiments (SI) are also given. PBC = pseudo-backcross, PIC = pseudo-intercross, FI = fully-informative.

H^2	r	$N_f = 10, N_o = 100$						$N_f = 25, N_o = 40$					
		PBC		PIC		FI		PBC		PIC		FI	
		PR	SI	PR	SI	PR	SI	PR	SI	PR	SI	PR	SI
0.05	0.0	0.41	0.43	0.24	0.27	0.60	0.58	0.23	0.27	0.12	0.11	0.32	0.30
	0.1	0.21	0.21	0.12	0.11	0.30	0.28	0.11	0.12	0.06	0.06	0.14	0.14
	0.3	0.03	0.03	0.02	0.01	0.04	0.03	0.02	0.04	0.02	0.02	0.02	0.02
0.10	0.0	0.87	0.87	0.69	0.67	0.98	0.97	0.68	0.65	0.42	0.46	0.86	0.87
	0.1	0.59	0.56	0.38	0.39	0.81	0.77	0.36	0.37	0.19	0.18	0.51	0.51
	0.3	0.08	0.08	0.04	0.05	0.10	0.06	0.04	0.05	0.03	0.03	0.05	0.04
0.15	0.0	0.99	0.96	0.94	0.86	0.99	0.99	0.94	0.90	0.76	0.76	0.99	0.99
	0.1	0.86	0.88	0.68	0.67	0.98	0.96	0.67	0.66	0.41	0.40	0.85	0.84
	0.3	0.14	0.15	0.08	0.09	0.20	0.21	0.08	0.10	0.04	0.06	0.09	0.09

recombination rate. For a larger $H^2 = 0.15$, the theoretical powers of linkage detection when $r = 0.10$ were 0.68, 0.86 and 0.98 for PIC, PBC and FI marker loci, respectively, for a sample size of 10 families with 100 offspring each (Table 4). It also shows that once r is greater than 0.10 the power of linkage detection is very low even if the broad-sense heritability at the QTL is 0.15.

Powers were also evaluated with varying dominance ratio at the QTL (Table 5). It shows that the theoretical power of linkage detection increase, in general, for PIC and FI marker designs whereas it remains constant for PBC design as the dominance increase. However, the rate of increase is large when small number of families with large number of offspring are used. For example the power increase from 0.18 to 0.24 and 0.46 to 0.53 for PIC and FI marker designs with a sample size of 5 families each having 200 offspring. The effect of different QTL allele frequencies on the power of linkage detection is shown in Table 6. For an additive

Table 5. Effect of gene action at the QTL on power of linkage detection of 3 marker designs. The symbols N_f, N_o, and f (= d/a) represents the number of families, number of offspring per family and the dominance ratio at the QTL. The powers given here were evaluated from theoretical prediction (PR) and simulation (SI) at $H^2 = 0.05, r = 0.10, p = 0.50$ and type-1 error = 0.01. PBC = pseudo-backcross, PIC = pseudo-intercross, FI = fully-informative.

N_f (N_o)	f	PBC		PIC		FI	
		PR	SI	PR	SI	PR	SI
5 (200)	0.0	0.29	0.32	0.18	0.24	0.46	0.48
	0.5	0.29	0.35	0.20	0.26	0.49	0.49
	1.0	0.29	0.36	0.24	0.31	0.53	0.60
10 (100)	0.0	0.21	0.21	0.12	0.12	0.30	0.28
	0.5	0.21	0.23	0.13	0.15	0.32	0.29
	1.0	0.21	0.24	0.16	0.18	0.36	0.38
20 (50)	0.0	0.13	0.12	0.07	0.08	0.17	0.17
	0.5	0.13	0.15	0.08	0.08	0.18	0.17
	1.0	0.13	0.15	0.09	0.11	0.20	0.22

Table 6. Effect of gene frequency (p) at the QTL on power of linkage detection of 3 marker designs. The symbols N_f and N_o, represents the number of families, number of offspring per family and the dominance ratio at the QTL. The powers given here were evaluated from theoretical prediction at r (recombination rate) = 0.00, f (= d/a) = 0 and type-1 error = 0.01. Half the difference between QTL homozygotes (i.e., a) = 0.30 SD. PBC = pseudo-backcross, PIC = pseudo-intercross, FI = fully-informative.

N_f (N_o)	p	PIC	PBC	FI
5 (200)	0.20	0.15	0.25	0.39
	0.40	0.29	0.44	0.67
	0.50	0.31	0.46	0.70
	0.60	0.29	0.44	0.67
	0.80	0.15	0.25	0.39
10 (100)	0.20	0.10	0.17	0.25
	0.40	0.19	0.33	0.49
	0.50	0.20	0.35	0.52
	0.60	0.19	0.33	0.49
	0.80	0.10	0.17	0.25

gene action at the QTL, it shows that power of linkage detection is highest when $p = 0.50$.

DISCUSSION

Power of QTL detection in two-generation outbred pedigree for varying dominance ratios, size of the QTL, recombination rates between the marker and the QTL and design parameters, was predicted in the present study. Three types of marker configurations were investigated. Derivations in the present paper have shown that the power of detecting linkage between a fully-informative marker and a QTL can be expressed as function of design parameters and parameters describing genetic properties of the QTL. A very close agreement was found between the powers from theoretical evaluation and stochastic simulation under wide range of situations, suggesting reliability of the theoretical analysis.

Effect of family- type on power

Three types of informative marker configurations in full-sib families were considered in the present study. First, those families where one parent is homozygous for marker ($M_1M_1 \times M_1M_2$, backcross-type or PBC family); second, those where both parents are heterozygous, with the same genotype at marker locus ($M_1M_2 \times M_1M_2$, intercross-type or PIC family); and third, those where both parents have different genotypes at marker locus ($M_1M_2 \times M_3M_4$, fully-informative or FI family). The power of the third-type of maker configuration was clearly the highest compared to the other two designs for all parameter combinations considered in this study (Table 3). This is because the use of a fully-informative marker allows all four genotypic classes to be distinguished. If any classes were confounded, then power would decrease (MURANTY 1996). GÖTZ & OLLIVIER (1992) and KNOTT & HALEY (1992) using sib-pairs analysis and maximum likelihood analysis, respectively, also showed that the use of fully informative markers would greatly increase the power of QTL

detection. In general, the power of backcross-type families was higher than intercross-type families. Similar results were obtained by LUO (1993) and SOLLER & GENIZI (1978).

Effect of sample size

The full-sib families were assumed to be independent, which can be thought as a single-pair mating design structure. Increasing the number of offspring per family was found to be more efficient than increasing the number of families for a fixed total population size (Table 3). Table 3 shows that for a given experimental size of 1000, the 5 families with 200 offspring each (5 × 200) gave higher power compared to 10 × 100 and 20 × 50 combinations. Similar results were obtained by several researchers (HILL 1975, SOLLER & GENIZI 1978, WELLER *et al.* 1990, LUO 1993, VAN DER BEEK *et al.* 1995).

MURANTY (1996) found that the power increases when variance explained by QTL and/or population size increase, and when these factors determine a low power level, the power decreases as the number of parents increases. However, at a high power level, the power increases as the number of parents increases. As a result, MURANTY (1996) suggested that the use of only one full-sib family for QTL detection is often less powerful, especially when QTL effects to be detected explain more than 10 per cent of phenotypic variance. The reason for this is that the total variance in a population attributable to QTL is better sampled with more than two parents than with only two parents. However, single full-sib family are being used for QTL mapping studies, for example, in eucalyptus (GRATTAPAGLIA *et al.* 1995) and loblolly pine (KNOTT *et al.* 1997).

Effect of gene action and allele frequency

In our study we evaluated the effect of additive and non-additive QTL effects on the power of linkage detection. REBAÏ & GOFFINET (1993) suggested that at the QTL detection step, it is better to neglect dominance if it is not very large. However, recent study by LI *et al.* (1996) reported that the dominance variance contributes significantly to variation in tree height and diameter in loblolly pine. At $p = 0.5$, the power of PIC and FI family type designs increases as the dominance ratio increase. However, at $p = 0.5$, there was almost no effect of dominance on the power of PBC-type family design (Table 5). Similar results were reported by LUO (1993) for PIC and PBC-type family designs. Table 5 also showed that power of QTL detection using FI markers is greater compared to other two marker designs, at different levels of dominance ratio. For an additive gene action, we also evaluated the power of linkage detection at different allele frequencies (Table 6). It shows that the power is highest when $p = 0.50$. The effect of gene frequency and dominance becomes important when number of families is small. This is because the probability that the marker contrast in each of the family be zero is so large that even an infinite number of offspring will not meet the power requirement (SOLLER & GENIZI 1978). This effect is generally unimportant except when a dominant allele is also the more frequent. However, with the method presented in our paper, it is likely that those families with zero contrast will nevertheless contribute to the significance of the variance between marker types within families. Thus, the loss in power due to probability of sampling families with zero marker contrast can be reduced.

In this study, only bi-allelic QTL was considered. The use of fully-informative markers permits the assessment of multiple-allele QTL. Evidence of existence of more than two QTL alleles has been reported in loblolly pine (GROOVER *et al.* 1994). Degree of dominance must be estimated separately from the original QTL analysis. Pedigree and population level studies are needed to determine the prevalence of multiple-allele QTL (WILLIAMS 1996). However, the levels of power obtained for detection of linkage between a FI marker and a QTL, bi-allelic or multi-allelic, are quite similar under the given conditions (MURANTY 1996).

Comparison of power of 3 marker designs revealed that FI marker design was more powerful than PBC and PIC deigns. It would be quite useful to consider using information from the whole population rather than subsets of it (i.e., combining PIC, PBC and FI family types). The joint analysis of any informative family types can be done following the suggestions of KNOTT (1994). Many alleles in a population are necessary to obtain a fully-informative marker for crosses among several parents. Isoenzymes and restricted fragment length polymorphisms (RFLP) have seldom met this criteria, but micro-satellite (SSR) and expressed sequence tag (EST) techniques promise to provide enough alleles and are currently being developed in *Pinus radiata*.

ACKNOWLEDGEMENTS

S Kumar thanks N. Z. Forest Research Institute, Rotorua, New Zealand, for financial support. We thank Drs J. I. WELLER, R. D. BURDON and Z. W. LUO for helpful discussion.

REFERENCES

CARBONELL, E. A., GERIG, T. M., BALANSARD, E. & ASINS, M. J. 1992: Interval mapping in the analysis of nonadditive quantitative trait loci. *Biometrics* **48**:305–313.

FALCONER, D. S. 1989: Introduction to Quantitative Genetics. Longman, Scientific & Technical, London, 438 pp.

GÖTZ, K. W. & OLLIVIER, L. 1992: Theoretical aspects of applying sib-pair linkage tests to livestock species. *Genetics Selection Evolution* **24**:29–42.

GRATTAPAGLIA, D., BERTOLUCCI, F. L. & SEDEROFF, R. R. 1995: Genetic mapping of QTL controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using pseudo-testcross strategy and RAPD markers. *Theoretical and Applied Genetics* **90**: 933–947.

GROOVER, A., DEVEY, M., FIDDLER, T., LEE, J., MEGRAW, R., MITCHEL-OLDS, T., SHERMAN, B., VUJCIC, S., WILLIAMS, C. & NEALE, D. 1994: Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. *Genetics* **138**:1293–1300.

HILL, A. P. 1975: Quantitative linkage: a statistical procedure for its detection and estimation. *Annals of Human Genetics* **38**:439–449.

JAYAKAR, S. D. 1970: On the detection and estimation of linkage between a locus influencing a quantitative character and a marker locus. *Biometrics* **26**:451–464.

KNOTT, S. A. 1994: Prediction of the power of detection of marker-quantitative trait locus linkages using analysis of variance. *Theoretical and Applied Genetics* **89**:318–322.

KNOTT, S. A. & HALEY, C. S. 1992: Maximum likelihood mapping of quantitative trait loci using full-sib families. *Genetics* **132**:1211–1222.

KNOTT, S. A., NEALE, D. B., SEWELL, M. M. & HALEY, C. S. 1997: Multiple marker mapping of quantitative trait loci in an outbred pedigree of loblolly pine. *Theoretical and Applied Genetics* **94**:810–820.

LI, B., MCKEAND, S. E. & WEIR, R. J. 1996: Genetic parameter estimates and selection efficiency for the loblolly pine breeding in south-eastern U. S. *In*: Proceedings QFRI-IUFRO conference on Tree Improvement for Sustainable Tropical Forestry, Caloundra, Queensland, Australia, 27 October–1 November 1996, p. 164–168.

LUO, Z. W. 1989: Polygene location and selection for heterotic traits in *Drosophila melanogaster*. Ph D thesis, Univ. Birmingham, UK.

LUO, Z. W. 1993: The power of two experimental designs for detecting linkage between a marker locus and a locus affecting a quantitative character in a segregating population. *Genetics Selection Evolution* **25**:249–261.

LYNCH, M. & WALSH, B. 1997: Genetics and Analysis of Quantitative Traits. Sinauer Associates, Massachusetts, USA. 980 p.

MURANTY, H. 1996: Power of tests for quantitative trait loci detection using full-sib families in different schemes. *Heredity* **76**:156–165.

REBAÏ, A. & GOFFINET, B. 1993: Power of tests for QTL detection using replicated progenies derived from a diallel cross. *Theoretical and Applied Genetics* **86**:1014–1022.

SAS INSTITUTE INC. 1986: "SAS User's Guide: Statistics", Version 5 edition. SAS Institute Inc., Cary, NC. 956 pp.

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.

VAN DER BEEK, S., VAN ARENDONK, J. A. M. & GROEN, A. F. 1995: Power of two-and three-generation QTL mapping experiments in an outbred population containing full-sib or half-sib families. *Theoretical and Applied Genetics* **91**:1115 -1124.

WELLER, J. I., KASHI, Y. & SOLLER, M. 1990: Power of daughter and granddaughter designs for determining linkage between marker locus and quantitative trait loci in dairy cattle. *Journal of Dairy Science* **73**:2525–2537.

WILLIAMS, C. G. 1996: How will genomic mapping shape forest tree breeding strategy? *In*: Proceedings QFRI-IUFRO conference on Tree Improvement for Sustainable Tropical Forestry, Caloundra, Queensland, Australia, 27 October – 1 November 1996, p. 464–466.

Appendix 1. Expected values (E_{ij}) of the quantitative trait value (y) within observed marker genotype (M_iM_j) of offspring within sibship. Parental marker genotypes are: $M_1M_2 \times M_3M_4$. The cross represents all possible parental genotypes at the QTL along with their probabilities (Prob.). It is assumed that $Q_1Q_1 \sim N(a, \sigma^2)$, $Q_1Q_2 \sim N(d, \sigma^2)$, $Q_2Q_2 \sim N(-a, \sigma^2)$. The p and q represents QTL alleles frequencies and r ($s = 1 - r$) is the recombination rate between marker and the QTL.

Cross	Probability	E_{13}	E_{14}	E_{23}	E_{24}
$Q_1Q_1 \times Q_1Q_1$	p^4	a	a	a	a
$Q_1Q_1 \times Q_1Q_2$	p^3q	$a - r(a - d)$	$d + r(a - d)$	$a - r(a - d)$	$d + r(a - d)$
$Q_1Q_1 \times Q_2Q_1$	p^3q	$d + r(a - d)$	$a - r(a - d)$	$d + r(a - d)$	$a - r(a - d)$
$Q_1Q_1 \times Q_2Q_2$	p^2q^2	d	d	d	d
$Q_1Q_2 \times Q_1Q_1$	p^3q	$a - r(a - d)$	$a - r(a - d)$	$d + r(a - d)$	$d + r(a - d)$
$Q_1Q_2 \times Q_1Q_2$	p^2q^2	$(s^2 - r^2)a + 2rsd$	$(s^2 + r^2)d$	$(s^2 + r^2)d$	$(r^2 - s^2)a + 2rsd$
$Q_1Q_2 \times Q_2Q_1$	p^2q^2	$(s^2 + r^2)d$	$(s^2 - r^2)a + 2rsd$	$(r^2 - s^2)a + 2rsd$	$(s^2 + r^2)d$
$Q_1Q_2 \times Q_2Q_2$	pq^3	$d - r(a + d)$	$d - r(a + d)$	$-a + r(d + a)$	$-a + r(d + a)$
$Q_2Q_1 \times Q_1Q_1$	p^3q	$d + r(a - d)$	$d + r(a - d)$	$a - r(a - d)$	$a - r(a - d)$
$Q_2Q_1 \times Q_1Q_2$	p^2q^2	$(s^2 + r^2)d$	$(r^2 - s^2)a + 2rsd$	$(s^2 - r^2)a + 2rsd$	$(s^2 + r^2)d$
$Q_2Q_1 \times Q_2Q_1$	p^2q^2	$(r^2 - s^2)a + 2rsd$	$(s^2 + r^2)d$	$(s^2 + r^2)d$	$(s^2 - r^2)a + 2rsd$
$Q_2Q_1 \times Q_2Q_2$	pq^3	$-a + r(d + a)$	$-a + r(d + a)$	$d - r(a + d)$	$d - r(a + d)$
$Q_2Q_2 \times Q_1Q_1$	p^2q^2	d	d	d	d
$Q_2Q_2 \times Q_1Q_2$	pq^3	$d - r(a + d)$	$-a + r(a + d)$	$d - r(a + d)$	$-a + r(a + d)$
$Q_2Q_2 \times Q_2Q_1$	pq^3	$-a + r(a + d)$	$d - r(a + d)$	$-a + r(a + d)$	$d - r(a + d)$
$Q_2Q_2 \times Q_2Q_2$	q^4	$-a$	$-a$	$-a$	$-a$

Appendix 2. Variance [$Var\ (i,j)$] of the quantitative trait value (y) within observed marker genotype (M_iM_j) of offspring within sibship. Parental marker genotypes are: M_1M_2 x M_3M_4. The cross represents all possible parental genotypes at the QTL along with their probabilities (Prob.). It is assumed that $Q_1Q_1 \sim N(a, \sigma^2)$, $Q_1Q_2 \sim N(d, \sigma^2)$, $Q_2Q_2 \sim N(-a, \sigma^2)$. The p and q represents QTL alleles frequencies and r ($s = 1 - r$) is the recombination rate between marker and the QTL.

Cross	Prob.	Var (1,3)	Var (1,4)	Var (2,3)	Var (2,4)
$Q_1Q_1 \times Q_1Q_1$	p^4	σ^2	σ^2	σ^2	σ^2
$Q_1Q_1 \times Q_1Q_2$	p^3q	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a - d)^2$
$Q_1Q_1 \times Q_2Q_1$	p^3q	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a - d)^2$
$Q_1Q_1 \times Q_2Q_2$	p^2q^2	σ^2	σ^2	σ^2	σ^2
$Q_1Q_2 \times Q_1Q_1$	p^3q	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a - d)^2$
$Q_1Q_2 \times Q_1Q_2$	p^2q^2	$\sigma^2 + 2rs\,[a^2 + d^2 (1-2rs) - 2da\,(1-2r)]$	$\sigma^2 + 2rsa^2 + d^2 (r^2 + s^2)[1 - (r^2 + s^2)]$	$\sigma^2 + 2rsa^2 + d^2 (r^2 + s^2)[1 - (r^2 + s^2)]$	$\sigma^2 + 2rs\,[a^2 + d^2 (1-2rs) - 2da\,(2r-1)]$
$Q_1Q_2 \times Q_2Q_1$	p^2q^2	$\sigma^2 + 2\,rsa^2 + d^2\,(r^2 + s^2)\,[1 - (r^2 + s^2)]$	$\sigma^2 + 2\,rs\,[a^2 + d^2\,(1-2rs) - 2da(1-2r)]$	$\sigma^2 + 2\,rs\,[a^2 + d^2\,(1-2rs) - 2da\,(2r-1)]$	$\sigma^2 + 2\,rsa^2 + d^2\,(r^2 + s^2)[1 - (r^2 + s^2)]$
$Q_1Q_2 \times Q_2Q_2$	pq^3	$\sigma^2 + rs\,(d + a)^2$	$\sigma^2 + rs\,(d + a)^2$	$\sigma^2 + rs\,(d + a)^2$	$\sigma^2 + rs\,(a + d)^2$
$Q_2Q_1 \times Q_1Q_1$	p^3q	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a - d)^2$	$\sigma^2 + rs\,(a + d)^2$
$Q_2Q_1 \times Q_1Q_2$	p^2q^2	$\sigma^2 + 2\,rsa^2 + d^2\,(r^2 + s^2)\,[1 - (r^2 + s^2)]$	$\sigma^2 + 2\,rs\,[a^2 + d^2\,(1-2rs) - 2da\,(2r-1)]$	$\sigma^2 + 2\,rs\,[a^2 + d^2\,(1-2rs) - 2da\,(1-2r)]$	$\sigma^2 + 2\,rsa^2 + d^2\,(r^2 + s^2)[1 - (r^2 + s^2)]$
$Q_2Q_1 \times Q_2Q_1$	p^2q^2	$\sigma^2 + 2rs\,[a^2 + d^2 (1-2rs) - 2da\,(2r-1)]$	$\sigma^2 + 2rsa^2 + d^2 (r^2 + s^2)[1 - (r^2 + s^2)]$	$\sigma^2 + 2rsa^2 + d^2 (r^2 + s^2)[1 - (r^2 + s^2)]$	$\sigma^2 + 2rs\,[a^2 + d^2 (1-2rs) - 2da\,(1-2r)]$
$Q_2Q_1 \times Q_2Q_2$	pq^3	$\sigma^2 + rs\,(a + d)^2$	$\sigma^2 + rs\,(a + d)^2$	$\sigma^2 + rs\,(a + d)^2$	$\sigma^2 + rs\,(a + d)^2$
$Q_2Q_2 \times Q_1Q_1$	p^2q^2	σ^2	σ^2	σ^2	σ^2
$Q_2Q_2 \times Q_1Q_2$	pq^3	$\sigma^2 + rs\,(a + d)^2$	$\sigma^2 + rs\,(a + d)^2$	$\sigma^2 + rs\,(a + d)^2$	$\sigma^2 + rs\,(a + d)^2$
$Q_2Q_2 \times Q_2Q_1$	pq^3	$\sigma^2 + rs\,(a + d)^2$	$\sigma^2 + rs\,(a + d)^2$	$\sigma^2 + rs\,(a + d)^2$	$\sigma^2 + rs\,(a + d)^2$
$Q_2Q_2 \times Q_2Q_2$	q^4	σ^2	σ^2	σ^2	σ^2

Appendix 3. Mean and variances between marker genotype classes within sibships. Parental marker genotypes are: $M_1M_2 \times M_3M_4$. The cross represents all possible parental genotypes at the QTL along with their probabilities (Prob.), E_{ij} represents the expected value of offspring having marker genotype M_iM_j. It is assumed that $Q_1Q_1 \sim N(a, \sigma^2)$, $Q_1Q_2 \sim N(d, \sigma^2)$, $Q_2Q_2 \sim N(-a, \sigma^2)$. The p and q represents QTL alleles frequencies and r ($s = 1 - r$) is the recombination rate between marker and the QTL.

Cross	Probability	$\frac{1}{4}(E_{13} + E_{14} + E_{23} + E_{24})$	Variance between E_{13}, E_{14}, E_{23}, E_{24}
$Q_1Q_1 \times Q_1Q_1$	p^4	a	0
$Q_1Q_1 \times Q_1Q_2$	p^3q	$\frac{1}{2}(a + d)$	$\frac{1}{4}(s - r)^2(a - d)^2$
$Q_1Q_1 \times Q_2Q_1$	p^3q	$\frac{1}{2}(a + d)$	$\frac{1}{4}(s - r)^2(a - d)^2$
$Q_1Q_1 \times Q_2Q_2$	p^2q^2	d	0
$Q_1Q_2 \times Q_1Q_1$	p^3q	$\frac{1}{2}(a + d)$	$\frac{1}{4}(s - r)^2(a - d)^2$
$Q_1Q_2 \times Q_1Q_2$	p^2q^2	$\frac{1}{2}d$	$\frac{1}{2}[(s - r)^2a^2 + (r^2+s^2)^2d^2 + 4r^2s^2d^2] - \frac{1}{4}d^2$
$Q_1Q_2 \times Q_2Q_1$	p^2q^2	$\frac{1}{2}d$	$\frac{1}{2}[(s - r)^2a^2 + (r^2+s^2)^2d^2 + 4r^2s^2d^2] - \frac{1}{4}d^2$
$Q_1Q_2 \times Q_2Q_2$	pq^3	$\frac{1}{2}(d - a)$	$\frac{1}{4}(a + d)^2(s - r)^2$
$Q_2Q_1 \times Q_1Q_1$	p^3q	$\frac{1}{2}(a + d)$	$\frac{1}{4}(s - r)^2(a - d)^2$
$Q_2Q_1 \times Q_1Q_2$	p^2q^2	$\frac{1}{2}d$	$\frac{1}{2}[(s - r)^2a^2 + (r^2+s^2)^2d^2 + 4r^2s^2d^2] - \frac{1}{4}d^2$
$Q_2Q_1 \times Q_2Q_1$	p^2q^2	$\frac{1}{2}d$	$\frac{1}{2}[(s - r)^2a^2 + (r^2+s^2)^2d^2 + 4r^2s^2d^2] - \frac{1}{4}d^2$
$Q_2Q_1 \times Q_2Q_2$	pq^3	$\frac{1}{2}(d - a)$	$\frac{1}{4}(a + d)^2(s - r)^2$
$Q_2Q_2 \times Q_1Q_1$	p^2q^2	d	0
$Q_2Q_2 \times Q_1Q_2$	pq^3	$\frac{1}{2}(d - a)$	$\frac{1}{4}(a + d)^2(s - r)^2$
$Q_2Q_2 \times Q_2Q_1$	pq^3	$\frac{1}{2}(d - a)$	$\frac{1}{4}(a + d)^2(s - r)^2$
$Q_2Q_2 \times Q_2Q_2$	q^4	$-a$	0