Theor Appl Genet (1996) 93:691–702

P. M. Visscher · C. S. Haley

Detection of putative quantitative trait loci in line crosses under infinitesimal genetic models

Received: 22 May 1995 , Accepted: 19 January 1996

Abstract Quantitative trait locus (QTL) mapping studies often employ segregating generations derived from a cross between genetically divergent inbred lines. In the analysis of such data it is customary to fit a single QTL and use a null hypothesis which assumes that the genomic region under study contributes no genetic variance. To explore the situation in which multiple linked genes contribute to the genetic variance, we simulated an F_2-mapping experiment in which the genetic difference between the two original inbred strains was caused by a large number of loci, each having equal effect on the quantitative trait. QTLs were either in coupling, dispersion or repulsion phase in the base population of inbred lines, with the expected F_2 genetic variance explained by the QTLs being equivalent in the three models. Where QTLs were in coupling phase, one inbred line was fixed for all plus alleles, and the other line was fixed for minus alleles. Where QTLs were in dispersion phase, they were assumed to be randomly fixed for one or other allele (as if the inbred lines had evolved from a common ancestor by random drift). Where QTLs were in repulsion phase alleles within an inbred line were alternating plus and minus at adjacent loci, and alternative alleles were fixed in the two inbred lines. In all these genetic models a standard interval mapping test statistic used to determine whether there is a QTL of large effect segregating in the population was inflated on average. Furthermore, the use of a threshold for QTL detection derived under the assumption that no QTLs were segregating would often lead to spurious conclusions regards the presence of genes of large effects (i.e. type I errors). The employment of an alternative model for the analysis, including linked markers as cofactors in the analysis of a single interval, reduced the problem of type I error rate, although test statistics were still inflated relative to the case of no QTLs. It is argued that in practice one should take into account the difference between the strains or the genetic variance in the F_2 population when setting significance thresholds. In addition, tests designed to probe the adequacy of a single-QTL model or of an alternative infinitesimal coupling model are described. Such tests should be applied in QTL mapping studies to help dissect the true nature of genetic variation.

Key words Quantitative trait locus · Genetic mapping · Significance threshold · F_2 population · Infinitesimal model

Communicated by C. Smith

P. M. Visscher[1] (✉) · C. S. Haley
Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, Scotland

Present address:
[1] Institute of Ecology and Resource Management, University of Edinburg, School of Agriculture Building, West Mains Road, Edinburgh EH9 3JG, Scotland

Introduction

Animal and plant breeding has got along very well with a theory based on an infinitesimal model, i.e. that genetic variance is controlled by an infinite number of independent genes of infinitesimally small effect. Although obviously ultimately unrealistic, this model has provided a useful and relatively accurate framework for designing breeding programmes and predicting their results, to which the many successes of animal and plant breeding testify. Recently, with the advent of genetic marker information, methods to detect individual quantitative trait loci (QTLs) segregating in crosses between inbred lines, crosses between divergent outbred populations and within outbred populations have been developed (Lander and Botstein 1989; Weller et al. 1990; Haley and Knott 1992; Haley et al. 1994) and have been applied to animal and plant experimental populations and outbred populations using field data (e.g. Paterson et al. 1988; Andersson et al. 1994; Georges et al. 1995).

There is some controversy over the correct threshold to use when testing for the presence of a single QTL because multiple correlated tests are performed (Lander and Botstein 1989; Knott and Haley 1992; Churchill and

Doerge 1994; Jansen 1994; Rebai et al. 1994). However, in all studies about the appropriate threshold for the test statistic so far, the null hypothesis has been that there is no genetic variance attributable to the segment of chromosome under consideration. It could be argued that we are often interested in a null hypothesis based upon the infinitesimal model (or, more realistically, we are interested in testing the hypothesis that there is one, or a limited number, of QTLs of relatively large effect in a region versus the hypothesis that there are many of small effect). Since the strains for the experimental crosses on which QTL mapping studies are based are usually chosen because of their genetic difference, a priori under an infinitesimal model we might expect each chromosomal segment to explain a part of the strain difference in proportion to its size. Alternatively, we expect each chromosomal region to explain a proportion of the genetic variance in the F_2 or backcross population or in any outbred population under study. We are then faced with the dilemma of distinguishing this model from one in which only one or a few QTLs are segregating.

The aim of this study is (1) to present infinitesimal models which explain differences between two inbred lines, (2) to show that these genetic models give results in terms of average test statistics when testing 'no QTL versus a single QTL' which are similar to results from populations with segregating single QTLs of large effects and (3) to explore alternative analytical models which address this problem of inference and propose an alternative null hypothesis for the detection of QTLs.

Methods

Genetic models

For illustration purposes, we only consider an F_2 population which originates from two inbred lines. Extension to other populations derived from an inbred line cross, e.g. backcross populations, is straightforward. We also assume throughout that there is no recombinational interference, so the mapping function of Haldane (1919) applies, and that all genetic variance is additive.

Single-QTL model

In the case of a single QTL explaining the genetic variance in the F_2 population, the relationship between the line difference (D) and the variance ($\mathrm{var}_A(F_2)$) is simply

$$\mathrm{var}_A(F_2) = D^2/8 = (1/2)\alpha^2 \quad (1)$$

with α the allele substitution effect, i.e. one breed has a QTL with alleles each of effect $\alpha/2$, and another has alleles of effect $-\alpha/2$, so that $D = 2\alpha$. Wright's number of effective loci (Wright 1968) for an F_2 population

$$k = D^2/(8\,\mathrm{var}_A(F_2)),$$

is of course 1 for this model.

Infinitesimal coupling model

Again, consider two inbred lines which differ by D for some quantitative trait. If we assume that a very large number of QTLs, each with the same effect and in coupling phase, explain the strain difference, the relationship between D and the genetic variance in an F_2 population can be derived from Hill (1993) and is detailed in the Appendix

$$\mathrm{var}_A(F_2) \cong (1/2)(D/2)^2[1 - (1/L)\sum_{i=1}^{k} r_i]\,L \quad (2)$$

with k = number of chromosomes, L = the total map length (in Morgans), r_i = the recombination rate between the chromosome ends for the ith chromosome

Equation 2 was verified both by analytical methods and by simulation, and was found to be a very good approximation for the case of finite loci (even with as few as five evenly spaced loci on a single chromosome the approximation using an infinite number of loci works very well). For illustration purposes, we will only consider results from a single chromosome, i.e. a genome with only one chromosome, and assume that all the genetic variance and the strain difference are from loci residing on one chromosome. Then

$$\mathrm{var}_A(F_2) \cong (1/2)(D/2)^2(1 - r/L)/L \quad (3)$$

In the present study we use an arbitrary environmental variance of unity and a chromosome length of 1.0 Morgan ($L = 1.0$). The heritability, h^2, is defined as the ratio of the genetic variance and the total phenotypic variance in the F_2 population, $h^2 = \mathrm{var}_A(F_2)/(\mathrm{var}_A(F_2) + 1.0)$. For a chromosome of 100 cM, $r = (1/2)(1 - e^{-2}) = 0.43$, and

$$D = [8h^2/\{(1 - h^2)(1 - r)\}]^{1/2} = 3.75\,[h^2/(1 - h^2)]^{1/2} \quad (4)$$

Under this model, the equilibrium genetic variance, i.e. the genetic variance after many generations of random mating, is zero, because all variance is caused by covariances among pairs of loci, and the genic variance (the sum of the variances at individual loci) goes to zero for an infinite number of loci. Individual locus effects are of order (O) $O(1/n)$, and the variance is $O(1/n^2)$. The effective number of loci for the single chromosome is

$$k = D^2/[8(1/2)(D/2)^2(1 - r/L)/L] = L/(1 - r/L)$$

Infinitesimal dispersion model

We now consider a genetic model which gives the same average heritability in the F_2 as before, but without an initial linkage disequilibrium in the inbred lines. At each locus for each inbred line, an effect of either α or $-\alpha$ was simulated with equal probability, so that alleles at loci had effects either $(\alpha/2, \alpha/2)$ or $(-\alpha/2, -\alpha/2)$ for either inbred line. An interpretation of this model is that inbred lines have evolved by random drift from a common ancestor for which all loci were in linkage equilibrium, i.e. inbreeding without selection, and that we take a random sample of two such inbred lines. For n loci, paramters for the line difference and the genetic variance in the F_2 are now related to individual locus effects as

$$E(D) = 0$$

$$\mathrm{var}(D) = 2n\alpha^2$$

$$E[\mathrm{var}_A(F_2)] = n\alpha^2/4 \quad (5)$$

This model was termed the dispersion model. For a single locus ($n = 1$), the probability that the inbred lines are fixed for the same alleles is 1/2, so that the expected variance is $1/2(0 + \alpha^2/2) = \alpha^2/4$ under this model. To obtain a finite variance in the F_2 population, individual locus effects are of order $O(1/n^{1/2})$. The equilibrium genetic variance under this model is the same as the variance in the F_2 population. The effective number of loci is

$$k = E(D^2)/[8\,\mathrm{var}_A(F_2)] = \mathrm{var}(D)/[8n\alpha^2/4] = 1$$

This model is similar to the model of Robertson (1977), except for the assumption of inbred populations. Robertson (1977) sampled alleles at a particular locus independently of each other.

Infinitesimal repulsion model

In the previous model the line difference is on average zero, but its variance is not, so that by chance one line (or chromosome) may have a larger effect than the other. The same applies to the genetic variance: this variance varies depending on which inbred lines are sampled out of the hypothetical pool of many inbred lines. To ensure that the line difference is always zero and the genetic variance is constant, we define a model in which + and − alleles alternate and in which the lines are fixed for alternative alleles.
So, for n loci, the first inbred line is

$(+\alpha/2, +\alpha/2), (-\alpha/2, -\alpha/2), (+\alpha/2, +\alpha/2), (-\alpha/2, -\alpha/2)$, etc.,

and the second inbred line is

$(-\alpha/2, -\alpha/2), (+\alpha/2, +\alpha/2), (-\alpha/2, -\alpha/2), (+\alpha/2, +\alpha/2)$, etc.

The variance in the F_2 population is

$$\mathrm{var}_A(F_2) = (1/2)\alpha^2 \sum_{i=1}^{n} \sum_{j=1}^{n} (-1)^i (-1)^j (1 - 2r_{ij})$$

with r_{ij} the recombination fraction between loci i and j. It can be shown that if n becomes large and is even, that

$D = 0$,

$\mathrm{var}(D) = 0$,

$$\mathrm{var}_A(F_2) = (1/2)\alpha^2 [r + L], \quad (6)$$

with L the length of the chromosome and r the recombination rate between the chromosome ends. This model is not strictly an infinitesimal model because the effects of individual loci do not become infinitesimal when n tends to infinity (α is not a function of n, so that a particular F_2 variance can be obtained with any number of loci).

The above infinitesimal genetic models differ from the 'standard' Fisher-Bulmer infinitesimal model used in plant and animal breeding (e.g. Bulmer 1980) because of linkage and initial linkage disequilibrium. Our many models represent three extremes if we consider the correlation between effects of loci which are very close together: in the coupling model, this correlation tends to + 1, in the dispersion model the correlation tends to 0 and in the repulsion model it tends to − 1.

Simulation

Chromosomes of 100 cM with 11 evenly spaced fully informative markers (including markers at the ends) were simulated for an F_2 population, with individual loci at 1-cM intervals. Hence, 101 linked loci were simulated. Depending on the genetic model (single QTL, many QTLs in coupling phase, many QTLs in dispersion phase, repulsion model), individual locus additive effects were determined so as to obtain heritabilities in the F_2 of 0%, 1%, 2%, 5% and 10% for a single chromosome. For the single-QTL model, using Eq. 1, these heritabilities correspond to $D = 0$, 0.28, 0.40, 0.65 and 0.94, with D in units of environmental standard deviations. For the coupling model, using Eq. 3, these heritabilities correspond to $D = 0$, 0.38, 0.54, 0.86 and 1.25, with D in units of environmental standard deviations. No dominance or epistatic effects were simulated. For the single gene model data were generated with a single QTL at 55 cM from the start of the chromosome.

Models for analysis

We used several models to analyse the data. Test statistics were calculated using a likelihood ratio test, assuming that residuals are normally distributed, i.e.

test statistic = N log(Residual SS reduced model / Residual SS full model),

with N the number of observations (Haley and Knott 1992). The full model contains both a mean effect and QTL effects (an additive effect in the backcross population, and both an additive and dominance effect in F_2 populations), and in the reduced model only an overall mean is fitted. It had been shown by Knott and Haley (1992) that using either a likelihood ratio test or an F-test gives nearly identical results. Under the null hypothesis of no QTL on a chromosome, the test statistic follows a multiple of a χ^2 distribution (Haley and Knott 1992; Rebai et al. 1994). If we were to test at a specific location on the chromosome rather than searching along the chromosome for the best model fit, the test statistic would be distributed (asymptotically) as a χ^2 wih 1 or 2 degrees of freedom, depending on whether only an additive effect was fitted or both an additive and a dominance effect (Haley and Knott 1992).

Regression interval mapping

Data were analysed with the regression method of Haley and Knott (1992), which fits a putative QTL at different places along the chromosome (e.g. at 1-cM intervals) and calculates the test statistic at each point. The position giving the largest test statistic is the most likely position for a QTL. Apart from an overall mean, either an additive effect, or both an additive and a dominance effect were fitted in the regression. Only a single putative QTL was postulated in the model fitted.

Fitting selected markers as cofactors

For interval mapping markers were fitted in the model as cofactors following Jansen (1993, 1994) and Zeng (1993, 1994). Cofactors will absorb some of the variance due to additional (linked) QTL when mapping in a particular interval, and this model was fitted so as to try to distinguish between the single-QTL and the multiple-QTL models.

By means of backward selection (e.g. Draper and Smith 1966), individual markers were selected to be included in the model as cofactors depending on their F-ratio. Each marker has three genotypic classes, hence two additional effects are fitted for each marker. Probabilities [$p(F)$], corresponding to an F-ratio with (2, number of F_2 individuals minus 2) degrees of freedom, used to select cofactors in alternative models were 1, 0.20, 0.10, 0.05 or 0.01. A probability of 1 corresponds to the case of fitting all markers as cofactors, which was suggested by Zeng (1994). After markers had been selected as cofactors, the search was for a single QTL on the chromosome while fitting the selected markers as cofactors. Markers flanking the interval under consideration were not fitted, even if they had been selected (or pre-selected, in the case of 'selecting' all markers), because they would absorb some or all of the variance associated with the putative QTL (Jansen 1993; Zeng 1993).

Alternative statistical tests

Before trying to map QTLs on chromosomes or chromosomal regions, we may wish to apply some statistical tests to find out whether there is evidence of genetic variance which is associated with a chromosome, and if there is, whether the findings can be explained by an infinitesimal model.

Chromosomal test

The first test of whether there is genetic variance associated with the chromosome is performed by fitting all markers in the model of analysis and testing for significance using an F-test or a likelihood ratio test. Hence, we test for the joint marker effects.

Coupling test

The second test is whether any genetic variance associated with the chromosome is consistent with the infinitesimal coupling model. For

this latter test, if we assume the additive coupling model and perform multiple regression of phenotypes on all (evenly spaced) markers on a chromosome, we expect a priori that the relative marker estimates, up to a constant of proportionality, are 1/2 for the end markers, and 1 for the other markers (Visscher 1996). Hence, a priori weights for markers are proportional to the length of chromosome segment that they mark. If we use the results from Visscher (1996), the constant of proportionality can be derived easily,

$$\gamma = D[2r/(1-2r)]^2/L \quad [7]$$

so

$$\hat{D} = \hat{\gamma}[(1-2r)/(2r)]^2(L) \quad [8]$$

with r the recombination rate between the markers flanking the (evenly spaced) intervals. This suggests the test for the adequacy of an infinitesimal coupling model by comparison of the full regression model, in which all markers are fitted (with $m+1$ degrees of freedom, where m is the number of markers), to the model

$$y = \mu + \gamma[(1/2)x_1 + x_2 + \ldots + (1/2)x_m] + e,$$

with 2 degrees of freedom. x_i is the additive effect of marker i (for an additive model, x_i was -1 if both marker alleles originated from line 1, $+1$ if they orginated from line 2 and 0 for the heterozygotes). The line difference can be estimated using Eq. 8. The resulting likelihood ratio test has $(m+1)-2 = m-1$ degrees of freedom. Note that in this case we fit only a single degree of freedom per marker, whereas previously (when fitting markers as cofactors) we fitted 2 degrees of freedom per marker. This was done because we wish to test an additive (coupling) model and were not interested in possible dominance effects for these analyses, although in practice a dominance effect could be included in analyses of F_2 data.

Single-QTL test

In some cases we may be interested in testing whether genetic variance explained by a single chromosome is caused by a single QTL, or by two or more QTLs. For an additive model, this suggests testing a model in which a single QTL is fitted on the chromosome versus a model in which all markers are fitted as fixed effects. If the null hypothesis (a single QTL) is true, the test should be close to a chi-square with $(m-2)$ degrees of freedom (m degrees of freedom in the full model for fitting all markers and 2 degrees of freedom in the reduced model for fitting a position effect and an additive effect). Again, extension to the case where we are interested in dominance would be straightforward.

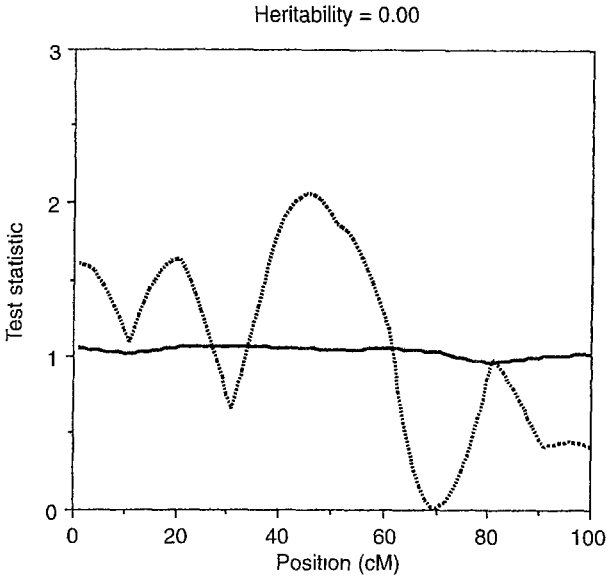

Fig. 1 Mean test statistic (*solid line*) per 1-cM intervals averaged over 1000 replicates and the test statistic for a random replicate (*dashed line*) when there are no QTLs segregating in the F_2 population

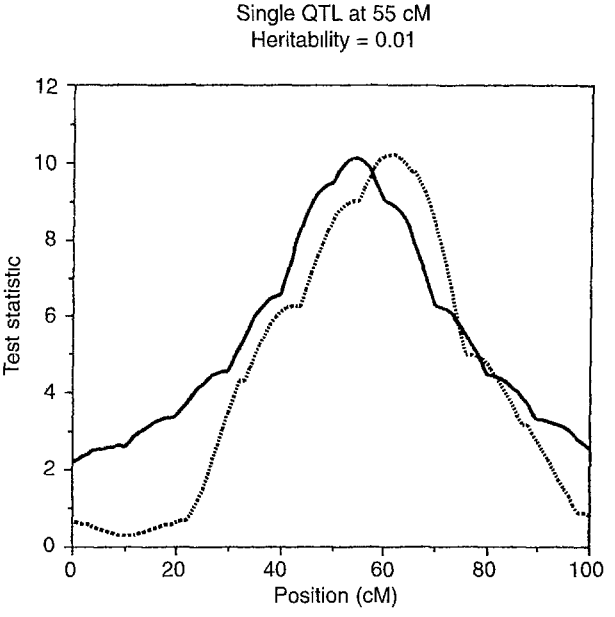

Fig. 2 Mean test statistic (*solid line*) per 1-cM intervals averaged over 1000 replicates and the test statistic for a random replicate (*dashed line*) for a single QTL at 55 cM from the start of a 100-cM chromosome. The effect of the QTL corresponds to a heritability of 0.01 in the F_2 population

Results

Results for the four genetic models are given in Tables 1–4. These results are summarized in terms of the mean likelihood ratio test statistic over all chromosomal positions over replicates, the mean over replicates of the maximum test statistic on each chromosome and the threshold over which the maximum test statistics were observed in the top 5% of replicates, i.e. the 95th percentile of the maximum test statistic. These values can be compared directly to significant thresholds based on likelihood ratio tests when testing 'no QTL versus a single QTL' (e.g. Lander and Botstein 1989; Rebai et al. 1994). Standard errors on the average maximum test statistic ranged from 0.2 to 0.8. Results from the coupling model (Table 2) were similar to those of the single QTL model for the same heritability in the F_2 (Table 1), implying that it may be difficult to distinguish between a single QTL or a very large number of QTLs linked in coupling spread along the chromosome, although we did not test how well those models fitted at this stage. Average and average maximum test statistics for the dispersion model were slightly lower, but not by much (Table 3). The smallest inflation of the test statistic, relative to the case of no genetic variance (Table 1) was from the repulsion model (Table 4). For realistic heri-

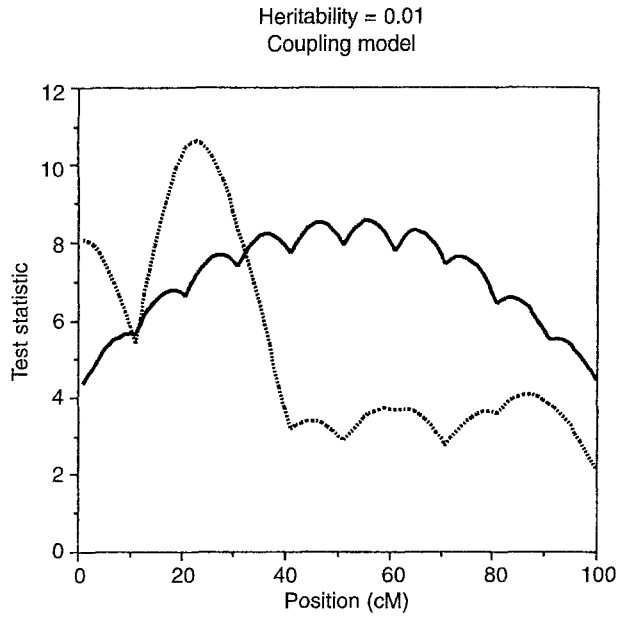

Fig. 3 Mean test statistic (*solid line*) per 1-cM intervals averaged over 1000 replicates and the test statistic for a random replicate (*dashed line*) when the heritability in the F_2 population is 0.01 based on the coupling model

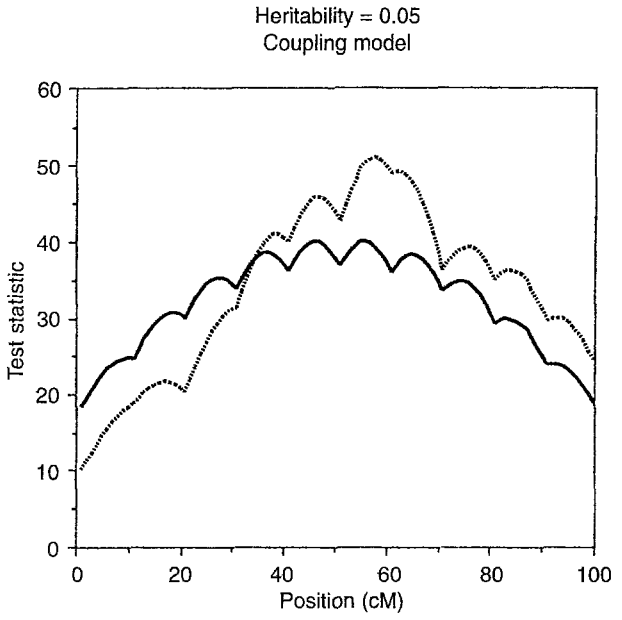

Fig. 4 Mean test statistic (*solid line*) per 1-cM intervals averaged over 1000 replicates and the test statistic for a random replicate (*dashed line*) when the heritability in the F_2 population is 0.05 based on the coupling model

tabilities per chromosome in livestock species (1 or 2%), the average maximum test statistics are similar for the first three models. Hence, for such cases it is the variance in the F_2 population which largely determines the average maximum test statistics, rather than the individual locus effects. Figure 1 shows the test statistic at each position along the chromosome for the case of no segregating QTLs, both for a single replicate and the average across replicates. The single replicate was the first replicate in a simulation run and was thus randomly chosen. As expected, the average test statistic is approximately 1, and deviations from the average for the single replicate is due to the sampling variation. Figure 2 shows the test statistic plotted against location on the chromosome for the case of a single QTL at 55 cM, which explains 1% of the F_2 variance. The average peak

Table 1 Test statistic summary from 100-cM chromosome for 1000 replicate F_2 populations of 1000 individuals. The marker map used has 11 equally spaced fully informative markers, and a single QTL at 55 cM explained the genetic variance

	Fit additive effect only h^2(%) in F_2 population					Fit additive and dominance effect h^2(%) in F_2 population				
	0	1	2	5	10	0	1	2	5	10
T^a	1.0	5.4	10.2	23.7	47.6	2.0	6.4	11.0	24.6	48.1
T^b_{max}	3.0	11.6	21.1	48.5	96.7	5.3	13.1	21.6	49.3	98.0
$T^c_{0.95}$	7.3	22.8	37.1	71.9	129.6	10.5	24.9	37.1	73.4	129.6

[a] Mean test statistic of all positions along the chromosome
[b] Maximum test statistic per chromosome
[c] 95th percentile of the maximum test statistic

Table 2 Test statistic summary for data from a coupling model. Simulations were of a 100-cM chromosome for 1000 replicate F_2 populations of 1000 individuals. The marker map used has 11 equally spaced fully informative markers, and 101 loci of equal effect coupling phase explained the heritability

	Fit additive effect only h^2(%) in F_2 population				Fit additive and dominance effect h^2(%) in F_2 population			
	1	2	5	10	1	2	5	10
T^a	7.1	12.9	31.7	64.1	8.1	14.1	32.4	64.6
T^b_{max}	12.7	21.2	45.9	87.1	14.2	22.5	46.9	88.5
$T^c_{0.95}$	24.2	35.9	67.4	116.2	25.8	37.3	68.7	118.0

[a] Mean test statistic of all positions along the chromosome
[b] Maximum test statistic per chromosome
[c] 95th percentile of the maximum test statistic

Table 3 Test statistic summary for data from a dispersion model. Simulations were of a 100-cM chromosome for 1000 replicate F_2 populations of 1000 individuals. The marker map used has 11 equally spaced fully informative markers, and 101 loci were in dispersion

	Fit additive effect only h^2(%) in F_2 population				Fit additive and dominance effect h^2(%) in F_2 population			
	1	2	5	10	1	2	5	10
T^a	4.9	8.5	19.7	38.3	5.7	9.8	22.2	41.5
T^b_{max}	10.4	17.0	37.5	72.9	11.8	18.7	41.2	77.4
$T^c_{0.95}$	28.3	49.3	112.2	213.6	29.3	52.2	122.8	237.3

[a] Mean test statistic of all positions along the chromosome
[b] Maximum test statistic per chromosome
[c] 95th percentile of the maximum test statistic

Table 4 Test statistic summary from a repulsion model. Simulations were of a 100-cM chromosome for 1000 replicate F_2 populations of 1000 individuals The marker map used has 11 equally spaced fully informative markers, and 101 loci were in repulsion

	Fit additive effect only h^2(%) in F_2 population				Fit additive and dominance effect h^2(%) in F_2 population			
	1	2	5	10	1	2	5	10
T^a	1.4	1.8	3.1	5.1	2.5	3.0	4.3	6.5
T^b_{max}	4.8	6.4	11.2	19.0	6.8	8.3	12.7	20.5
$T^c_{0.95}$	10.6	13.9	20.7	31.6	13.4	15.8	22.3	33.8

[a] Mean test statistic of all positions along the chromosome
[b] Maximum test statistic per chromosome
[c] 95th percentile of the maximum test statistic

is at 55 cM, although a single replicate can vary considerably from the average curve. Figures 3 and 4 show the test statistic plotted against the chromosomes for data generated with heritabilities of 1 and 5%, respectively, under the coupling model and analysed fitting a single QTL with a additive effect. Clearly, the test statistics are inflated by the presence of many QTLs each with small effects. If the figures were produced by real data, and the maximum test statistic compared with a threshold derived under the assumption that no QTLs were segregating (e.g. using 7.3 as the 5% threshold for the situation with no variance attributable to the chromosome, as given in Table 1) evidence for single QTLs (of large effect) would have been reported for Figs. 3 and 4.

Fitting all markers on the chromosome as cofactors in the model of analysis reduces the test statistic dramatically for the coupling model. For example, the 95th percentile for $h^2 = 0.05$ decreased from 71.9 (Table 2) to 12.5 (Table 5), which is not much larger than the 95th percentile for $h^2 = 0.0$ of 9.5. For the dispersion model the values for the 95th percentile with $h^2 = 0.05$ were 112.2 without cofactors (Table 3) and 22.7 fitting all markers as cofactors (Table 5). However, the average maximum test statistic and the 95th percentile values for the single QTL model were also greatly reduced when fitting all additional markers as cofactors, with the resulting values being similar to those from the dispersion model with the same F_2 heritability (Table 6). Hence, fitting all markers as cofactors reduces the power when only a single QTL is segregating. For example, for a single QTL explaining 5% of the variance in the F_2, the 95th percentile of the maximum test statistic was reduced from 37.1 (Table 1) to 24.1 (Table 6). Selecting cofactors using F-thresholds from multiple regression of phenotypes on markers generally resulted in only a few markers being selected. For example, when an F-threshold corresponding to a probability of 0.20 was used, the largest probability used for selection (apart from selecting all markers, which corresponds to a probability of 1), 2–3.8 markers were selected for the coupling and dispersion model.

Testing for genetic variance per chromosome ('chromosomal test') and for the infinitesimal coupling genetic ('coupling test') model was performed for the single QTL, the coupling and the dispersion models, and the results are presented in Table 7. The mean test statistic and its 95th percentile when testing for variances associated with the chromosome, i.e. chromosomal test, approximately followed a chi-squared distribution with $m = 11$ degrees of freedom when the null hypothesis was true (mean of 11.2 and 95th percentile of 20.3). When there was genetic variance associated with the chromosome, mean test statistics were much larger, and for a particular h^2, similar for all models. For example, mean test statistics for $h^2 = 0.02$ were 29.0, 31.2 and 29.1 for the single QTL, coupling, and dispersion models, respectively.

For the coupling model, the test statistic for all heritabilities (i.e. testing for an infinitesimal coupling model) seemed to behave like the expected chi-squared distribution with 10 degrees of freedom, and the estimates of the breed differences were unbiased. For the single-QTL model, test statistics were larger, so that the hypothesis of a large number of QTLs explaining the chromosomal variance was likely to be rejected. If the

Table 5 Test statistic when fitting selected markers as cofactors in the analysis of data from coupling and dispersion models. Simulations were of a 100-cM chromosome for 1000 replicate F_2 populations of 1000 individuals

			Coupling				Dispersion			
			h^2							
p(F)		0.00	0 01	0.02	0 05	0.10	0.01	0.02	0 05	0.10
1.0	T	1.0	1 1	1.3	1.6	2.4	1.4	1.7	2.7	4.5
	T_{max}	4.6	5.2	5.8	7 0	9.2	6.1	7.5	11.6	19.9
	$T_{0\,95}$	9.5	10.1	11.1	12.5	15 8	12.2	15.6	22.7	49.6
	Cof.[a]	11	11	11	11	11	11	11	11	11
0 20	T	1.1	2.4	3 3	5.2	7.8	2.4	3.4	6 0	9.6
	T_{max}	4.4	10.2	13.2	18.5	25.2	9.6	14.0	24.2	40.1
	$T_{0\,95}$	10.6	19.4	25.7	34 2	49.4	21.0	31.9	63.1	95.6
	Cof.[a]	1.5	2.0	2.4	2.8	3.4	2.2	2.5	3.2	3.8
0.10	T	1.0	2 8	4.0	6 5	9.5	2.5	3.8	6.8	11.0
	T_{max}	3.7	11.5	15.8	23.2	30.2	10.1	15 4	27.0	44.5
	$T_{0\,95}$	10.0	21.6	28.5	43.0	58.4	23.3	35.4	71 2	114.9
	Cof.[a]	0 6	1.3	1.7	2.2	2.9	1.3	1.7	2.4	3 0
0 05	T	1.0	3.2	4.6	7.5	11.0	2.7	4.1	7.3	12.0
	T_{max}	3.4	12.2	17.8	26.6	34.6	10.3	16.1	29.1	47.9
	$T_{0\,95}$	9.5	23.0	30.9	50.8	64.9	25.8	40 5	74.9	127.8
	Cof.[a]	0.3	1.0	1.3	1 9	2 5	0.9	1.2	1.9	2.5
0.01	T	1.0	4.1	5.6	9.4	13.7	3.0	4.5	8.3	13.6
	T_{Tmax}	3.1	12.6	20 0	33.8	44.0	10.3	17.2	32.5	53.2
	$T_{0\,95}$	7.8	23.8	32.9	61.1	81.8	28.1	43.8	84.8	140.7
	Cof.[a]	0.0	0 6	1 0	1.5	2.1	0.4	0.7	1.3	1 9

[a] Average number of cofactors fitted in the model

Table 6 Test statistic when fitting selected markers as cofactors in the analysis of data from the single QTL genetic model. Simulations were of a 100-cM chromosome for 1000 replicate F_2 populations of 1000 individuals

		h^2 *Single QTL model*				
p(F)		0 00	0.01	0.02	0.05	0 10
1.0	T	1.0	1.3	1.6	2 4	3.8
	T_{max}	4.6	5.9	7.5	12.8	23.2
	$T_{0\,95}$	9.5	12.1	15 6	24.1	40.2
	Cof.[a]	11	11	11	11	11
0.20	T	1 1	2.5	3.5	6.4	10.2
	T_{max}	4.4	11.1	16.3	34.8	69.2
	$T_{0\,95}$	10 6	22.5	32.1	64.7	117.9
	Cof.[a]	1.5	2.0	2.2	2.3	2.8
0.10	T	1.0	2.8	4.0	7.8	12.8
	T_{max}	3.7	12.1	18 8	40.5	79.2
	$T_{0\,95}$	10.0	23.5	34.3	69.8	124.4
	Cof.[a]	0.6	1.3	1.4	1 6	2.0
0.05	T	1.0	3.0	4.3	8.6	14.6
	T_{max}	3.4	12.2	19.7	43.7	84.7
	$T_{0\,95}$	9.5	23.5	34.7	70.5	125.9
	Cof.[a]	0 3	1.0	1.2	1.2	1.6
0.01	T	1 0	3.6	4.7	9.3	17.6
	T_{max}	3.1	12.2	20.3	47.0	92.6
	$T_{0\,95}$	7.8	23.6	35.4	71.0	127.8
	Cof.[a]	0.0	0.5	0.9	1.0	1.1

[a] Average number of cofactors fitted in the model

effect of a single QTL is small (say, $h^2 < 0.01$), the power of a test which aims to distinguish between a single QTL and many linked QTLs is likely to be small. For example, the mean coupling test statistic and its 95th percentile for $h^2 = 0.01$ and data from a single-QTL model were 12.4 and 22.4, not too different from the values under the genetic model of a large number of QTLs in coupling (10.2 and 18.5). The estimate of the

Table 7 Summary of test statistics from the chromosomal test (TC) and the coupling test (TI) Simulations were of a 100-cM chromosome for 1000 replicate F_2 populations of 1000 individuals. The marker map used has 11 equally spaced fully informative markers

Model		h^2(%) in F_2 population				
		0	1	2	5	10
Single QTL	TC[a]	11.2	20.5	29.0	56 2	105.1
	$TC_{0\,95}$[b]	20.3	34.0	45.9	81.7	138.8
	TI[a]	10.1	12.4	14.2	20.7	32.8
	$TI_{0\,95}$[b]	18.9	22.4	24.6	35.4	51.3
	D[c]	0.00	0.28	0.40	0.65	0.94
	$\hat{D}$[d]	0.00	0.32	0.45	0 72	1.06
Coupling	TC[a]		21.3	31.2	61.6	115.6
	$TC_{0\,95}$[b]		35.7	48.8	87.1	149.7
	TI[a]		10.2	10.2	10.1	10.3
	$TI_{0\,95}$[b]		18 5	18.8	18.9	19.4
	D[c]		0.38	0.54	0.86	1.25
	$\hat{D}$[d]		0.38	0.54	0.87	1.27
Dispersion	TC[a]		20 3	29 1	57.6	106.2
	$TC_{0\,95}$[b]		42.7	63.0	140.0	268.0
	TI[a]		13.6	17.5	29.1	49.8
	$TI_{0\,95}$[b]		24.9	35.5	66.1	117.7
	D[c]		−0.01	−0.02	−0.01	−0.05
	$\hat{D}$[d]		−0.01	−0.02	−0.02	−0.04

[a] Average test statistic
[b] 95th percentile of the test statistic
[c] True breed difference (for single QTL and coupling model) or average breed difference from simulation (for dispersion model)
[d] Estimate of the breed difference assuming the infinitesimal coupling model

Table 8 Summary of test statistics from the single QTL test. Simulations were of a 100-cM chromosome for 1000 replicate F_2 populations of 1000 individuals. The marker map used has 11 equally spaced fully informative markers

Model		h^2(%) in F_2 population				
		0	1	2	5	10
Single QTL	T[a]	8.2	8.4	8.5	8.8	9 0
	$T_{0.95}$[b]	15.9	16.4	16.6	17.1	16.8
Coupling	T[a]		8.4	10.3	16.3	28.7
	$T_{0\,95}$[b]		16.2	18.4	26.8	45.3
Dispersion	T[a]		10.0	12.3	19.7	31.9
	$T_{0\,95}$[b]		20.2	25.2	42.7	77.2

[a] Average test statistic
[b] 95th percentile of the test statistic

line difference was biased upwards for the single-QTL genetic model when assuming an infinitesimal model. Mean and 95th percentile for the dispersion model did not follow a chi-squared distribution with 10 degrees of freedom. This is because by chance chromosomal intervals may differ between the lines because of a cluster of QTLs, with the effects being positive in some intervals and negative in others. Hence, by chance we effectively create a multiple-QTL genetic model, so that the hypothesis of a large number of linked QTL in coupling phase should be rejected.

The null hypothesis of a single additive QTL explaining the genetic variance was tested for a single chromosome for the single QTL model and the coupling and dispersion genetic models. Results are in Table 8. For the single QTL model the mean test statistic and its 95th percentile look very similar to expected values from a chi-squared distribution with $11 - 2 = 9$ degrees of freedom (Table 8). For small heritabilities, results tend towards a chi-squared with 8 degrees of freedom. For heritabilities of 1% and 2%, it still seems difficult to distinguish between the different genetic models. However, for larger heritabilities, it is clearly possible to separate the single QTL from the coupling and dispersion models. For example, for $h^2 = 0.05$, the mean test statistic for the single-QTL model is 8.8, whereas the mean values for the coupling and dispersion models are 26.8 and 42.7, respectively.

Discussion

The results presented here demonstrate the difficulty in using interval mapping approaches to distinguish be-

tween models with a single QTL of relatively large effect and those with multiple QTLs of small effect. Different genetic models which explained the same amount of variance in an F_2 population derived from inbred lines were compared when analysing data of 1000 individuals. The genetic variance was either explained by a single QTL, by many linked QTLs in coupling phase, by many QTLs in linkage equilibrium in the inbred lines or by a model for which inbred lines were fixed for alternative alleles and within a line adjacent loci had opposite effects.

We do not claim that any of these many-loci models truly represent strain differences in practice. However, the genetic nature of strain differences in terms of the distribution and frequencies of QTLs is not known, and these models were selected to cover the range of possibilities.

It was shown that for heritabilities in the F_2 of 1% or 2%, the average test statistics were similar for widely different genetic models if we only fit a single QTL in the analysis. For a single chromosome, thresholds pertaining to a Type-I error of 5% for a QTL of additive effect were in the range of 11 to 36 for the coupling and repulsion infinitesimal models (Tables 2 and 4). Here we use type-I error in the sense that a single QTL model is not rejected when the data were simulated under alternative genetic models. These log-likelihood thresholds correspond to LOD scores of 2.4–7.8, well above the LOD threshold used in practice for a single chromosome (Lander and Botstein 1989; Paterson et al. 1988). Thus, when testing for the presence of a single QTL on a chromosome, if there was 'polygenic' variance associated with that chromosome the null hypothesis would be rejected too frequently. In practice, when data from crosses between inbred lines are analysed, the relevant statistical test should perhaps be 'Is there evidence for more genetic variance explained by a chromosome segment than expected by chance (assuming an infinitesimal model), taking the strain difference into account'. Finding segments of the chromosome which explain some of the genetic variance, but attributing this to QTLs of large effects will be inefficient if the aim is to find 'the' QTL itself or to use the finding for marker-assisted selection. Hence, one should be careful in interpreting results from QTL mapping experiments. This argument is similar to that of Dekkers and Dentine (1991), who presented an infinitesimal model for outbred populations (similar to that proposed by Robertson 1977) and showed that chromosome segments of apparent large effects are to be expected under their model. McMillan and Robertson (1974) have also highlighted that it is very difficult to distinguish models with single QTL versus models with many linked QTLs.

The fitting of all cofactors in the model reduced the test statistic (and Type-I error) substantially for the multiple-QTL models, but it is known that the power to detect QTLs will also be reduced (Zeng 1994), and the reduced test statistics for data simulated under the single-QTL model found in this study provide further evidence of this problem. The selection of markers for use as cofactors on the basis of their F probabilities might reduce the problem slightly, in that for a given F_2 heritability the difference between the single-QTL and the coupling model test statistics was increased. This increase was greatest with cofactors selected on the basis of F probabilities of around 0.05. However, the difference between the single-QTL and the coupling models for the average maximum and 95th percentile threshold test statistics remained slight except at high heritabilities. Furthermore, the difference between the single-QTL model and the dispersion model test statistics were even less affected by fitting cofactors. Thus, cofactors are of limited use for distinguishing a single QTL per chromosome from multiple linked QTLs, although this does not devalue their use for distinguishing one QTL from two or a few linked QTL as advocated by Jansen (1993, 1994) and Zeng (1993, 1994).

For practical purposes, one could set the appropriate threshold by simulation, assuming that the strain difference is caused by many genes smeared out along the chromosomes. Once a QTL of large effect has been detected, the threshold for testing for remaining QTLs on other chromosomes should be reduced, taking into account that a smaller amount of genetic variance has to be explained by the remaining chromosomes. A simple iterative scheme which uses the peaks from all chromosomes and the initial strain difference (or variance in the F_2) could then be used to distinguish QTLs of large effects from lots of QTLs of small effects. Note that with real data, a permutation test (Churchill and Doerge 1994) to set the threshold for the test statistic will not reduce the problem of single-QTL versus many-QTL models. A permutation test is performed by repeatedly shuffling phenotypes among marker genotypes, and for each sampled dataset mapping QTLs. The empirical distribution of test statistics from the permuted datasets may then be used to determine whether the test statistic obtained from the original unshuffled dataset was significant. Permutation may help ameliorate the problem of non-normality in the distribution of the segregating generation, but permutation at random with respect to the marker data is equivalent to assuming that there is no genetic variance in the segregating generation.

A simple test per chromosome based on the multiple regression of phenotypes on all the markers on the chromosome is easily performed and seemed to work well for testing for the presence of genetic variance associated with a chromosome (Table 6). Furthermore, the number of independent tests performed would be equal to the number of chromosomes, and hence the overall 5% significance threshold (at least for a single trait) can be simply derived from the Bonferroni adjustment as approximately 0.05/(the number of chromosomes). We have not, however, explored the power of this test here, and there is further work to be done to optimize its performance in various situations (e.g. should all markers or a well-spaced subset be used, and how is this related to sample size and other design factors?).

Table 9 Statistical tests for the dispersion model restricting the heritability in any F_2 cross to be within 1% of the average heritability (h^2_{ave}). The tests were: (1) single QTL versus no QTL, (2) variance associated with the chromosome versus no effects on the chromosome, (3) single QTL versus other genetic model, (4) infinitesimal coupling model versus other genetic model. Simulations of a 100-cM chromosome for 1000 replicate F_2 populations of 1000 individuals

h^2_{ave} (%)	Statistical test							
	No QTL vs. single QTL		Chromosomal		Single QTL		Infinitesimal coupling	
	T	$T_{0.95}$	T	$T_{0.95}$	T	$T_{0.95}$	T	$T_{0.95}$
1	10.1	20.3	20.0	32.3	10.6	19.4	14.3	26.2
2	17.6	33.4	29.8	48.1	13.1	24.6	18.9	35.8
5	38.6	62.4	58.7	82.9	21.0	40.1	33.6	66.0
10	74.0	108.5	107.7	141.7	34.6	69.9	58.6	114.4

Testing directly for an infinitesimal coupling model is again easily performed in practice and worked very well for the simple cases we studied. Again, the significance threshold for this test can be derived from the Bonferroni adjustment. In practice, markers will not be evenly spaced and fully informative, and the relative marker effects under an assumed infinitesimal model will not reduce to 1/2 for markers at the ends and 1 for all other markers. However, using the results from Visscher (1996), we can easily derive the relative weights for the markers using a simple selection index procedure.

The testing for a single-QTL versus some other genetic model seemed to work well and is recommended for use in practice. This test was the only one which clearly separated a single-QTL from the coupling model. However, the power of this test is only likely to be high with single QTL of large effect in reasonably sized populations.

The comparison between the alternative genetic models was based on a constant heritability in the F_2 population. However, for the dispersion model in particular the heritability varied over replicates because at every replicate two inbred lines were sampled from a hypothetical large pool of inbred lines. This sampling process mimics to some extent current QTL mapping projects in many plant species because different research groups look for QTLs in crosses between different inbred lines. It can be questioned whether the comparison with the other genetic models is the most appropriate one because the maximum test statistic will be correlated with the amount of genetic variance and the breed difference. For example, the correlations between the test statistic (QTL vs. no QTL) and the absolute breed difference and genetic variance in the F_2 were 0.74 and 0.79 for an average heritability of 0.01, and 0.90 and 0.95 for a heritability of 0.10. To compare the results for the dispersion model with the other models at a constant heritability, we performed additional simulation for which pairs of inbred lines were used only if the heritability in the F_2 population was within 1% of the desired average heritability. For example, for an average heritability of 0.10, data were selected from those simulated only for those crosses which resulted in heritabilities between 0.099 and 0.101. On average, approximately 100 000 crosses were simulated to obtain 1000 crosses which satisfied the heritability criterion. For any simulated data set, several different statistical tests were performed. The results are shown in Table 9 and generally show little change in the average maximum test statistic relative to values in Table 7 and 8 for unselected datasets, although the 95th percentile was reduced substantially for the test 'no QTL vs. single QTL' and the chromosomal test. For both the single-QTL test and the infinitesimal coupling test, average maximum test statistics were slightly increased. For example, when the heritability in the F_2 was within 1% of an average heritability of 5%, the average maximum test statistic for the single-QTL test and infinitesimal coupling test were 21.0 and 33.6, respectively (Table 9), whereas the maximum test statistics for the unrestricted case were 19.7 (Table 8) and 29.1 (Table 7), respectively.

For data analyses in practice there is no easy single solution. We would recommend that the data be explored by simple tests based only on a multiple regression of observations on markers as a first step before attempting to map QTLs. The careful interpretation of results from a chromosomal test, an infinitesimal test and interval mapping using cofactors (using carefully chosen, and perhaps simulation derived, significance thresholds) may help limit the number of false trails that are followed. However, the distinguishing of single QTL from linked groups of QTL (as occur in the dispersion model used in this study) will always be difficult if not impossible, without resort to more markers and more meioses.

Appendix

The infinitesimal coupling model

Model for many linked QTL

The aim of this Appendix is to present a genetic model of many linked loci which explains breed differences and genetic varainces in F_2 and backcross populations derived from inbred lines.

Suppose we have n equally spaced loci on a chromosome (or interval) of length L. If alternative alleles are fixed in the two breeds for all loci, and all loci have equal additive effects (allele substitution effect α), then

D = trait difference for parental breeds $= 2n\alpha$, or

$\alpha = D/(2n)$

We first consider a backcross population which represents the simplest case. For backcross (BC) generation t, the total additive genetic variance is

$$\begin{aligned}\mathrm{var}_A(BC_t) &= \sum_{i=1}^{n}\sum_{j=1}^{n}[(1/2^t)(1-r_{ij})^t\alpha^2] - n^2(1/4^t)\alpha^2 \\ &= \alpha^2 n^2\left[(1/2^t)\sum_{i=1}^{n}\sum_{j=1}^{n}(1-r_{ij})^t(1/n)(1/n) - (1/4^t)\right] \qquad [A1] \\ &= \left(\frac{D^2}{4}\right)\left[(1/2^t)\sum_{i=1}^{n}\sum_{j=1}^{n}(1-r_{ij})^t(1/n)(1/n) - (1/4^t)\right]\end{aligned}$$

with r_{ij} the recombination fraction between loci i and j. The term inside the square brackets is similar to the equations of Stam and Zeven (1981) and Hill (1993) for the variance of the proportion of the genome from the recurrent population. Essentially, the genetic variance under the assumed model is just a scaled version of the variance of the proportion of the genome from one of the inbred lines. For large n, and assuming Haldane's (1919) mapping function without interference, A1 can be approximated by (see also Hill 1993)

$$\mathrm{var}_A(BC_t) = \frac{D^2}{4}\left[1/(2L^2)(1/4^t)\sum_{i=1}^{t}\binom{t}{i}1/(i^2)(2iL - 1 + e^{-2iL})\right] \qquad [A2]$$

with L the length of the chrosomosome (block) in Morgans. For the first BC generation, A2 may be written as

$$\mathrm{var}_A(BC_1) = (1/4)(D^2/4)(1 - r_m L)/L$$

with r_m the recombination fraction between the chromosome ends.

For an F_k population with random mating, the additive variance is (using results from Kempthorne 1957)

$$\mathrm{var}_A(F_k) = (1/2)\alpha^2\Sigma\Sigma(1-2r_{ij})(1-r_{ij})^{k-2}/n^2$$

For an F_2 population this reduces to

$$\begin{aligned}\mathrm{var}_A(F_2) &= 1/2\alpha^2\Sigma\Sigma(1-2r_{ij})/n^2 \\ &\cong (1/2)(D/2)^2(1 - r_m/L)/L = 2\,\mathrm{var}_A(BC_1)\end{aligned}$$

Extension to multiple chromosomes is straightforward (see Hill 1993). for v chromosome, the additive genetic variance in the F_2 population is

$$\mathrm{var}_A(F_2) \cong (1/2)(D/2)^2\left[1 - (1/L)\sum_{i=1}^{v} r_i\right]/L$$

with r_j the recombination fraction between the chromosome ends for chromosome j, and L the total map length.

With the 'standard' infinitesimal model (e.g. Bulmer 1980) the genotypic value at any locus is of order (O) $1/n^{1/2}$, and the variance at each locus is O$(1/n)$. With a very large number of linked loci on a chromosome (block), the variance in backcross populations is of O$(1/n^2)$, so values at individual loci should be of order $(1/n)$. However, we cannot have both finite genic and total variance for infinite n (the genic variance goes to zero for large n with finite genetic variance).

Acknowledgements We thank Robin Thompson for helpful comments and discussions. Thanks to Bill Hill, Peter Keightley, John Whittaker, Robert Curnow and an anonymous referee for many useful comments on an earlier version of the manuscript. PMV was funded by the Marker Assisted Selection Consortium of the U K. pig industry (Cotswold Pig Development Company Ltd., J.S.R Farms Ltd., Newsham Hybrid Pigs Ltd., Pig Improvement Company, National Pig Development Company and the Meat and Livestock Commission) and by MAFF, DTI and the BBSRC. CSH acknowledges support from MAFF, BBSRC and the European Commission.

References

Andersson L, Haley CS, Ellegren H, Knott SA, Johansson M, Andersson K, Andersson-Eklund L, Edfors-Lilja I, Fredholm M, Hansson I, Håkansson J, Lundström K (1994) Genetic mapping of quantitative trait loci for growth and fatness in pigs. Science 263:1771–1774

Bulmer MG (1980) The mathematical theory of quantitative genetics Oxford University Press. Oxford

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Dekkers JCM, Dentine MR (1991) Quantitative genetic variance associated with chromosomal markers in segregating populations. Theor Appl Genet 81:212–220

Draper NR, Smith H (1966) Applied regression analysis. John Wiley & Sons, New York

Georges M, Nielsen D, Mackinnon M, Mishra A, Okimoto R, Pasquino AT, Sargeant LS, Sorensen A, Steele MR, Zhao X, Womack JE, Hoeschele I (1995) Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing Genetics 139:907–920

Haldane JBS (1919) The combination of linkage values and the calculation of distances between the loci of linked factors. J Genet 8:299–309

Haley CS, Knott SA (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69:315–324

Haley CS, Knott SA, Elsen JM (1994) Mapping quantitative trait loci in crosses between outbred lines using least squares. Genetics 136:1195–1207

Hill WG (1993) Variation in genetic composition in backcrossing programs. J Hered 84:212–213

Jansen RC (1993) Interval mapping of multiple quantitative trait loci. Genetics 135:205–211

Jansen RC (1994) Controlling the type I and type II errors in mapping quantitative trait loci. Genetics 138:871–881

Kempthorne O (1957) An introduction to genetic statistics. John Wiley & Sons. New York

Knott SA, Haley CS (1992) Aspects of maximum likelihood methods for the mapping of quantitative trait loci in line crosses. Genet Res 60:139–151

Lander ES, Botstein DB (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps Genetics 121.185–199

McMillan I, Robertson A (1974) The power of methods for the detection of major genes affecting quantitative characters. Heredity 32:349–356

Paterson AH, Lander ES, Hewitt D, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726

Rebai AR, Goffinet B, Mangin B (1994) Approximate thresholds of interval mapping tests for QTL detection. Genetics 138:235–240

Robertson A (1977) Artificial selection with a large number of linked loci. In: (eds) Proc Int Conf Quant Genet. Iowa State University Press, Ames, Iowa, pp 307–322

Stam P, Zeven AC (1981) The theoretical proportion of the donor genome in near-isogenic lines of self-fertilizers bred by backcrossing. Euphytica 30:227–238

Visscher PM (1996) Proportion of the variation in genetic composition in backcrossing programs explained by genetic markers. J Hered 87:136–138

Weller JI, Kashi Y, Soller M (1990) Power of daughter and granddaughter designs for determinign linkage between marker loci

and quantitative trait loci in dairy cattle. Journal of Dairy Science 73:2525–2537
Wright S (1968) Evolution and the genetics of populations, vol 1: genetic and biometric foundations. University of Chicago Press, Chicago
Zeng Z-B (1993) Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci Proc Natl Acad Sci. USA 90:10972–10976
Zeng Z-B (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468