Theor Appl Genet (1992) 85:353–359

THEORETICAL AND APPLIED GENETICS

Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus

A. Darvasi and M. Soller

Department of Genetics, The Alexander Silberman Life Sciences Institute, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel

Received August 12, 1991; Accepted April 7, 1992
Communicated by A. L. Kahler

Summary. "Selective genotyping" is the term used when the determination of linkage between marker loci and quantitative trait loci (QTL) affecting some particular trait is carried out by genotyping only individuals from the high and low phenotypic tails of the entire sample population. Selective genotyping can markedly decrease the number of individuals genotyped for a given power at the expense of an increase in the number of individuals phenotyped. The optimum proportion of individuals genotyped from the point of view of minimizing costs for a given experimental power depends strongly on the cost of completely genotyping an individual for all of the markers included in the experiment (including the costs of obtaining a DNA sample) relative to the cost of rearing and trait evaluation of an individual. However, in single trait studies, it will almost never be useful to genotype more than the upper and lower 25% of a population. It is shown that the observed difference in quantitative trait values associated with alternative marker genotypes in the selected population can be much greater than the actual gene effect at the quantitative trait locus when the entire population is considered. An expression and a figure is provided for converting observed differences under selective genotyping to actual gene effects.

Key words: QTL – Selective genotyping – Genetic marker – Linkage

Introduction

Experimental designs for the determination of linkage between marker loci and quantitative trait loci (QTL) have been widely described (Elston and Stewart 1971; Geldermann 1975; Hill 1975; Jensen 1989; Knapp 1990; Lander and Botstein 1989; Lebowitz et al. 1987; Neimann-Sorensen and Robertson 1961; Simpson 1989; Soller et al. 1976; Soller and Beckmann 1983, 1990; Soller and Genizi 1978; Weller 1986), and a number of successful experimental studies have been carried out (Beever et al. 1989; Edwards et al. 1987; Gelderman et al. 1985; Gonyon et al. 1987; Haenlein et al. 1987; Kahler and Wehrhahn 1986; Paterson et al. 1988; Sax 1923; Tanksely et al. 1982, Weller 1987; Weller et al. 1988).

Although in the past experimental marker-QTL linkage studies have usually been aimed at carrying out a general linkage analysis for a variety of traits simultaneously, future studies will often be aimed at analyzing a single trait (Soller and Beckmann 1988). In this case, by genotyping only individuals from the high and low phenotypic tails of the entire sample population, the number of individuals genotyped for a given power can be decreased considerably, at the expense of an increase in the number of individuals phenotyped (Lebowitz et al. 1987). When analyzing traits with unknown distribution or with respect to qualitative traits (e.g. disease resistance) an approach based on the change in marker allele frequencies at the tails, as a consequence of linkage to the trait under consideration, is appropriate (Lebowitz et al. 1987). Lander and Botstein (1989) showed that for continuous traits, the power for such an analysis can be markedly increased when the analysis is based on the quantitative values of the individuals in the high and low tails of the population. They applied maximum likelihood methods to the analysis of the data and termed their approach "selective genotyping".

In the study presented here further theoretical aspects to selective genotyping are considered, including: choosing an appropriate model for use in analysis of the data, obtaining estimates of the gene effects at the QTL unbi-

Correspondence to: A. Darvasi

ased by the selection applied, and determining the optimum proportion to be retained in the selected high and low tails of the population from the point of view of minimizing total experimental cost for a given power.

Theory

Consider the case where the population consists of two groups, genotype A and genotype B, with equal expected sample size and each normally distributed with respective means and variances μ_A, μ_B and $\sigma_A^2=\sigma_B^2$. In many instances $(\mu_A-\mu_B)/\sigma_{A\text{ or }B}$ will be small, so that the QTL in question contributes only a small component to the overall population variance. In this case σ_A^2 and σ_B^2 are approximately equal to σ^2, the population variance. Thus, all three variances will henceforth be considered equal and will be denoted σ^2. For theoretical analysis, it is convenient to standardize population values, setting the population variance to 1 and the expectations of genotypes A and B, to δ and $-\delta$, respectively.

In the following analysis, detection of a marker-QTL linkage will be based on a t-test of the difference between the sample average of individuals with genotype A (denoted $\bar{A}$) and the sample average of individuals with genotype B (denoted $\bar{B}$) as found in the population obtained by pooling both selected tails to form one unified sample population, denoted "pooled-tails population". In order to determine the total number, N_T, of individuals in the pooled-tails population needed to detect a QTL of given effect, the respective expectations of individuals with genotypes A and B in the pooled-tails population, E(A) and E(B); their difference, D_T; and the within-genotype variance in the pooled-tails population, σ_T^2, will be derived.

Mean difference between the two genotypes in the pooled-tails population (D_T)

The distributions of the trait value for the two genotypes are presented in Fig. 1. The total proportion of the sample analyzed is denoted p ($p/2$ at each tail), while p_1 and p_2 are the proportions of genotypes A and B, respectively, in the upper selected tail (equal to proportions of genotypes B and A, respectively, in the lower tail, see Fig. 1). In general, subscript notation "1" will refer henceforth to genotype A in the upper tail and genotype B in the lower tail, while subscript "2" will refer to genotype B in the upper tail and genotype A in the lower tail. Consequently p_1 and p_2 are:

$$p_1=\phi(Z_{p/2}+\delta) \tag{1}$$

$$p_2=\phi(Z_{p/2}-\delta) \tag{2}$$

where $\phi(x)$ is the cumulative probability function of the standard normal distribution, and $Z_{p/2}$ is the value of x that solves $\phi(x)=p/2$. Note that for small values of δ, as considered in this study,

$$p_1+p_2=p.$$

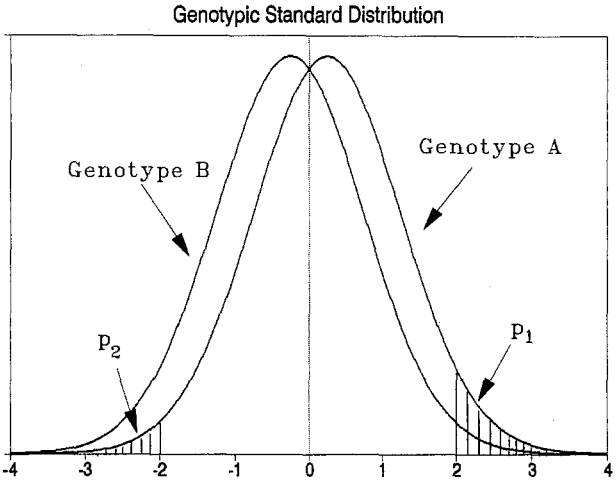

Fig. 1. The distribution of trait values of two genotypes (genotype A and genotype B, having mean standardized values $+\delta$ and $-\delta$, respectively) under selective genotyping, showing the proportion of individuals of genotype A present in the upper and lower tails of the population (p_1 and p_2, respectively)

Denote:

$$p_1^*=\frac{p_1}{p_1+p_2} \tag{3}$$

$$p_2^*=\frac{p_2}{p_1+p_2} \tag{4}$$

Then the probability density functions of individuals with genotype A and B in the upper and lower tails of the population are, respectively:

$$f_A(x)=p_1^*\cdot f_{A,U}(x)+p_2^*\cdot f_{A,L}(x) \tag{5}$$

$$f_B(x)=p_2^*\cdot f_{B,U}(x)+p_1^*\cdot f_{B,L}(x) \tag{6}$$

where the probability density functions of each genotype, A and B, at the upper (U) and lower (L) tails are:

$$f_{A,U}(x)=\frac{1}{p_1}\cdot\frac{1}{\sqrt{2\pi}}e^{-(x-\delta)^2/2}\cdot I_{(Z_{1-p},\infty)}(x) \tag{7}$$

$$f_{A,L}(x)=\frac{1}{p_2}\cdot\frac{1}{\sqrt{2\pi}}e^{-(x-\delta)^2/2}\cdot I_{(-\infty,Z_p)}(x) \tag{8}$$

$$f_{B,U}(x)=\frac{1}{p_2}\cdot\frac{1}{\sqrt{2\pi}}e^{-(x+\delta)^2/2}\cdot I_{(Z_{1-p},\infty)}(x) \tag{9}$$

$$f_{B,L}(x)=\frac{1}{p_1}\cdot\frac{1}{\sqrt{2\pi}}e^{-(x+\delta)^2/2}\cdot I_{(-\infty,Z_p)}(x) \tag{10}$$

where $I_{(a,b)}(x)$ is an indicator function that equals 1 when $a<x<b$ and 0 otherwise. Each of these distributions is obtained from the appropriate tail of the normal distribution.

The mean of each genotype, over the pooled sample formed of the upper and lower tails of the entire population, are obtained directly from (5) and (6) and equal:

$$E(A)=p_1^*\,\mu_1-p_2^*\,\mu_2 \tag{11}$$

$$E(B)=-p_1^*\,\mu_1+p_2^*\,\mu_2 \tag{12}$$

where μ_1 and μ_2 are the population means of genotypes A and B in the upper tail and $-\mu_1$ and $-\mu_2$ are the population means of genotypes B and A in the lower tail, and equal:

$$\mu_1=i_{p_1}+\delta \tag{13}$$

$$\mu_2=i_{p_2}-\delta \tag{14}$$

where, $i_t=x_t/t$ (t stands for either p_1 or p_2) is the mean of an upper tail of a standard normal distribution, t is the area under the tail and x_t is the ordinate of the standard normal distribution at the point Z_t (Falconer 1989).

Using (13), (14) and the appropriate substitutions in (11) and (12) yields:

$$E(A)=\frac{x_{p_1}-x_{p_2}}{p}+\delta \tag{15}$$

$$E(B)=\frac{x_{p_2}-x_{p_1}}{p}-\delta \tag{16}$$

Then D_T is:

$$D_T\equiv E(A)-E(B)=2\left(\frac{x_{p_1}-x_{p_2}}{p}+\delta\right) \tag{17}$$

Using Taylor's expansion and defining:

$$\gamma_p=1+Z_{1-p/2}\,i_{p/2} \tag{18}$$

we obtain:

$$D_T=2\delta\cdot\gamma_p \tag{19}$$

Consequently, estimating D_T through $\hat{D}_T=\bar{A}-\bar{B}$, it is possible to estimate the actual gene effect:

$$\hat{\delta}=\frac{\hat{D}_T}{\gamma_p} \tag{20}$$

Within-genotype variance in the pooled-tails population (σ_T^2)

From (5), (6) and using the well-known equation:

$$Var(x)=E(x^2)-[E(x)]^2 \tag{21}$$

it is possible to obtain Var(A) and Var(B) (the within-genotype variance σ_T^2):

$$\sigma_T^2\equiv Var(A)=Var(B)=1+\frac{1}{p}(Z_{1-p_1}\,x_{p_1}+Z_{1-p_2}\,x_{p_2})-\left(\frac{x_{p_1}-x_{p_2}}{p}\right)^2 \tag{22}$$

When considering relatively small gene effects this expression can be very closely approximated by:

$$\sigma_T^2=\gamma_p \tag{23}$$

Thus, relative to the mean and the variance in the unselected population, both the difference between means and the variance in the pooled-tails population are increased by the same factor, γ_p.

Number of individuals required (N_T)

Using the described procedure, and the standard expression for sample size for a t-test as a function of type I and type II errors (Soller et al. 1976), the number of genotyped individuals (when genotyping a fraction p of the population sample) required to attain a given power $1-\beta$ with a given type I error, α, is:

$$N_T=\frac{4\sigma_T^2\,(Z_{\frac{\alpha}{2}}+Z_\beta)^2}{D_T^2} \tag{24}$$

Consequently, N_T relative to the required population size, N, for the case where all the population is genotyped, is:

$$\frac{N_T}{N}=\frac{(2\delta)^2\,\delta_T^2}{D_T^2}=\frac{1}{\gamma_p} \tag{25}$$

Clearly, N_T decreases as p decreases. However, the number of individuals raised and phenotyped (N_{ph}) increases as p decreases since,

$$N_{ph}=\frac{N_T}{p} \tag{26}$$

Therefore, N_{ph} relative to N would be:

$$\frac{N_{ph}}{N}=\frac{1}{p\gamma_p} \tag{27}$$

Thus, savings in genotyping by decreasing p result in an increased total number of individuals raised and phenotyped. Finding the optimal proportion of selection, in terms of reducing the overall cost of such an experiment, is considered in the following section.

Optimal selection

The total cost of an experiment aimed at detecting linkage between a marker and a QTL as a function of the proportion of the population selected, p, is:

$$F(p)=c_g\,N_T+c_{ph}\,N_{ph} \tag{28}$$

where c_g is the cost of completely genotyping an individual for all markers followed in the experiment (including costs of obtaining a DNA sample for that individual) and c_{ph} is the cost of rearing and trait evaluation of an individual. Incorporating (19) and (23) in (24), substituting the resultant N_T in (26) and both N_T and N_{ph} in (28), dividing

by c_{ph} and all the constants in N_T and defining $C = c_g/c_{ph}$ yields a relative cost function, $F^*(p)$, which has the same optimum as $F(p)$,

$$F^*(p) = \frac{1}{\gamma_p}\left(C + \frac{1}{p}\right) \tag{29}$$

Since p ranges only from 0 to 1, the value for p which minimizes the relative cost function, $F^*(p)$, was readily found numerically for a wide range of cost ratios, C.

In some cases the cost of genotyping x individuals is not a linear expression, $c_g x$, and the cost of rearing y individuals is not a linear expression, $c_{ph}\, y$. Instead, costs increase or decrease when exceeding a certain number of markers per individual or individuals per experiment. Costs may increase, particularly for raising and phenotyping additional animals, since additional or more expensive facilities and efforts might be required. Costs may decrease, particularly for genotyping individuals, since a major effort is required to set up an efficient system for genotyping each marker, but once the system is in place, genotyping additional individuals reduces unit costs. In these cases, the optimum is obtained by following the same procedure as above, except that C equals the marginal cost ratio, i.e. c_g is the cost of genotyping the last individuals, and c_{ph} is the cost of raising and phenotyping the last individuals (Darvasi and Soller 1992).

Effect of recombination

The analysis presented so far considers the case of a pleiotropic effect of the marker (the marker is the quantitative trait locus) or a marker tightly linked to a QTL. However, in most cases there will be an appreciable proportion of recombination, r, between the QTL and the marker. The probability density functions, (5) and (6), adjusted for recombination become:

$$f_A(x) = (1-r)\, p_1^*\, f_{A,U}(x) + r\, p_2^*\, f_{B,U}(x) + (1-r)\, p_2^*\, f_{A,L}(X) + r\, p_1^*\, f_{B,L}(x) \tag{30}$$

$$f_B(x) = (1-r)\, p_2^*\, f_{B,U}(x) + r\, p_1^*\, f_{A,U}(x) + (1-r)\, p_1^*\, f_{B,L}(X) + r\, p_2^*\, f_{A,L}(x) \tag{31}$$

Following the same argument as above, it can be readily shown that this will cause a decrease in D_T by a factor of $(1-2r)$, so that N_T will increase by a factor of $1/(1-2r)^2$. Consequently, in the presence of recombination the estimate of the gene effect, δ, will be biased, because under the above analysis it is impossible to estimate r. This problem can be overcome by using maximum likelihood techniques to estimate r (see following sections).

Specific designs

The method we have just presented has been described in a general manner (i.e. in terms of "genotype A", "genotype B") so that it can be applied to a variety of populations and designs. As a specific example we consider a backcross population originating from a cross between two inbred lines. It is assumed that one line has genotype M/M at a marker linked to a QTL with genotype A/A; whereas the other line has genotypes m/m and a/a, respectively. Genotype A of the general design would now be M/M, and genotype B of the general design would now be M/m.

In this simple example the results of the general design apply directly. In more complex designs adjustments might be needed. For example, in an F_2 population there will be three genotypes M/M, M/m and m/m. The analysis can then be based on the two homozygous marker genotypes only. In this case genotype A would be M/M, genotype B would be m/m and N_T would be the total number of homozygous marker genotypes in the pooled-tails needed to attain a given power. Alternatively, the analysis could be carried out by a maximum likelihood approach that takes into account all of the available genotypes.

Maximum likelihood approach

Using maximum likelihood techniques it is possible to estimate both the proportion of recombination and the gene effect (Weller 1986, 1987; Lander and Botstein 1989) and to increase the overall power of detecting a QTL (Simpson 1989; Lander and Botstein 1989). Maximum likelihood can also be applied, as in the F_2 design, when more than two genotypes are present in each tail and their distributions are known.

The likelihood function would be:

$$L = \prod_{i=U,L} \prod_{j=1}^{M} \prod_{k=1}^{N_{i,j}} f_{i,j}(x_{i,j,k}) \tag{32}$$

where $f_{i,j}(x)$ is the probability density function of an individual with genotype j at the i^{th} tail (Upper, U, or Lower, L). M is the number of marker genotypes at each tail, $N_{i,j}$ is the number of individuals at the i^{th} tail with genotype j and $x_{i,j,k}$ is the trait value of the k^{th} individual at the i^{th} tail with the j^{th} marker genotype. Maximizing L in relation to the unknown parameters, i.e. the means and variances of the trait value for the QTL genotypes and proportion of recombination between marker and QTL, provides maximum likelihood estimators (MLE) of these parameters.

A likelihood ratio test can be carried out to test if $\delta = 0$. This test is expected to be more powerful then the t-test presented in the general analysis.

The likelihood function, presented here in a general way, must be applied and solved separately for each design. Maximizing the likelihood function is not a simple task, and specific computer programs are needed.

Numerical results

For convenience, Fig. 2 presents 2δ as a function of D_T and p. When p is small, D_T is many fold greater than 2δ. Thus, estimates of gene effect obtained from selective genotyping experiments must be appropriately adjusted using expression (19) or Fig. 2. From (19) it is seen that 2δ is a linear function of D_T and that as p increases the function approaches the values for $p=1$ at which point $D_T=2\delta$.

Over the entire range of values of p, as the proportion selected decreases, the relative number of progeny genotyped decreases linearly (Fig. 3) while the relative number of progeny phenotyped increases exponentially. Over the range $p=1.0$ to 0.3, the increase in number of progeny phenotyped, although exponential, is slow and moderately linear in form. For $p<0.3$, however, the increase becomes dramatically more rapid. Consequently, when p is small, although the number of progeny genotyped is markedly reduced, this is only at the expense of a major increase in the number of progeny phenotyped. For example, at $p=0.05$ (0.025 selected in each tail), the number of progeny genotyped for a given power is smaller by a factor of 5.6 relative to genotyping the entire population, but the number of progeny phenotyped is increased by a factor of 3.6 (similar results are presented by Lander and Botstein 1989). When p is large, however, the number of progeny genotyped can be appreciably reduced at the expense of only a negligible increase in total number of progeny phenotyped. For example, at $p=0.5$ (0.25 selected in each tail), the required number of progeny phenotyped for a given power increases by only 8%. Considered another way: for $\delta=0.125$, with a total population size of 500 and genotyping only half of the population (i.e. $p=0.5$), the power of detecting marker-QTL linkage will only drop from 0.80 to 0.77. This remarkable result means that the middle 50% of the experimental population contributes essentially no information to the analysis.

Optimal proportion selected, p^*, depends strongly on the cost ratio, C. Figure 4 presents the optimal selection for a wide range of C. It can be seen that the optimal proportion selected, p^*, is a very sensitive function of C, ranging from $p^*=0.8$ for $C=0.01$ to $p^*=0.02$ for $C=100$.

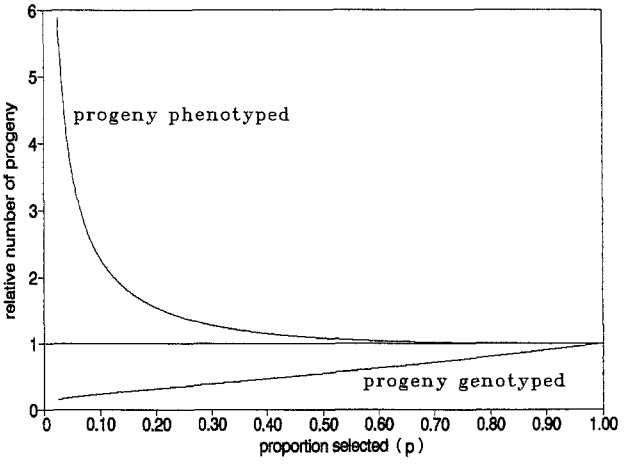

Fig. 3. Numbers of progeny genotyped and phenotyped required to attain a given power with a given type I error (relative to an unselected population) as a function of the proportion of the sample analyzed (p)

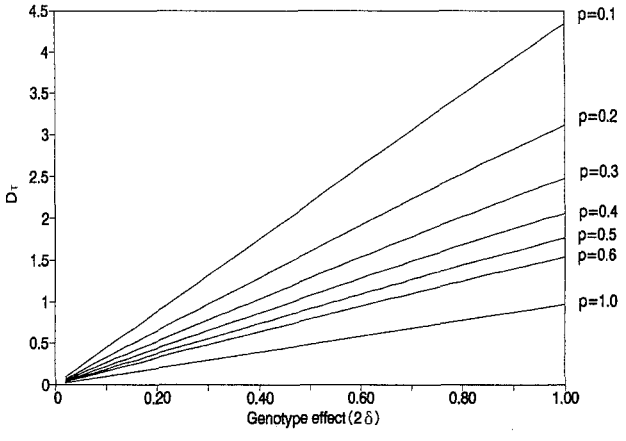

Fig. 2. The observed difference in quantitative trait value associated with alternative marker genotypes in the selected population (D_T) as a function of the actual gene effect at the quantitative trait locus (2δ) and total proportion selected over both tails (p)

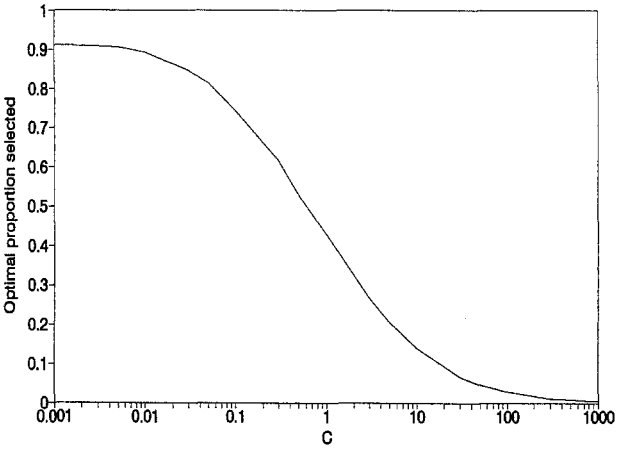

Fig. 4. Optimal proportion selected over both tails of the population in order to minimize costs for given experimental power as a function of C, the cost of completely genotyping an individual for all markers included in the experiment (including costs of obtaining a DNA sample) relative to the cost of the rearing and trait evaluation of an individual. C is given in logarithmic scale

Discussion

The present study emphasizes the efficiency of selective genotyping when analyzing marker-QTL linkage data with respect to a single trait. The results show that for this situation it will almost never be useful to genotype more than the upper and lower 25% of the population, even if, for some reason, it is not possible to increase population size accordingly. When population size can be suitably increased, a much smaller proportion of selection will generally be appropriate.

Complete exploitation of the information in the selected sample requires application of maximum likelihood techniques. However, the ANOVA-based analysis presented throughout this paper can be easily applied in all cases with reasonable power. The analysis presented here also provides an expression relating the observed difference in quantitative trait value associated with alternative marker genotypes (D_T) in the selected pooled-tails population, to the actual gene effect at the quantitative trait locus. As pointed out in the Theory section, due to recombination this expression will generally underestimate the actual gene effect. However, analyses of this sort will generally be carried out at rather close marker spacings in regions of interest. Consequently, the average and even maximum distance of a QTL from the nearest marker, and resultant bias, will be small.

The distribution of values found for individuals having genotype A or B is a consequence of residual effects of genes at other loci and environmental factors. Individuals in the upper tail of their genotype distribution are those that received positive residuals. Individuals in the lower tail of their genotype distribution are those that received negative residuals. Consequently, a powerful positive correlation between residual efforts and genotpye value will be present in the pooled-tails population. It is this positive correlation (analogous to a positive genotype × environment correlation) that magnifies the actual difference in genetic value between the two genotypes.

The powerful dependence of optimum proportion selected on total costs of genotyping an individual for all markers relative to total costs of rearing and phenotyping an individual requires detailed consideration of these costs when planning a marker-QTL analysis based on selective genotyping. Relative costs in the range of $C=1$ to 10 would be appropriate when considering experiments specifically carried out for purposes of QTL mapping and involving fruit trees, large farm animals or traits that are costly to measure. In such cases, optimal selection would be between 0.1 and 0.4. In contrast, C can be very large for traits that are readily evaluated in annual plants or small animals (e.g. body weight in poultry). In such cases optimal selection would be in the range 0.05 to 0.1. Finally, in some cases, e.g. dairy herd recording programs, large numbers of individuals are routinely evaluated with respect to a large number of agrotechnical traits for management, disease control or other purpose. In this case, rearing and phenotyping costs for the QTL mapping program are nil, and optimum proportion selected can be extremely small, limited only by the general caution that extreme phenotypic outliers may represent errors of recording or other artifacts.

Most experimental studies of marker-QTL linkage to date have been of a general nature, aimed at exploring the genetic architecture of quantitative traits. In the future, it can be expected that a greater proportion of marker-QTL linkage studies will be carried out for purposes of providing basic information for marker-assisted introgression (Soller and Beckmann 1988). In this case, studies will generally be focused on single traits. For example, experiments are currently under way in a number of institutes in Africa aimed at mapping loci conferring trypanotolerance through the analysis of crosses between the trypanotolerant N'Dama cattle of West Africa and trypanosensitive Zebu animals (Soller and Beckmann 1987). F_1, F_2 and backcross animals are being produced specifically for this purpose and will be evaluated for trypanotolerance and associated traits only. Even in this case, although the costs of rearing and phenotyping the experimental animals are very great, selective genotyping can still make a major contribution to reducing genotyping costs.

Although presented here in terms of marker-QTL linkage analysis, the statistical benefits of basing an analysis on selected tails of a population are perfectly general for any situation where it is desirable to determine a relationship between a two-class variable (such as sex, habitat, treatment versus control group, etc.) and a continuous variable. Clearly, in most cases of this sort there is no difficulty in defining the class to which an individual belongs, so that little is to be gained by reducing classification ("genotyping") costs. However, the method may be useful in those instances where the classification procedure is complicated or expensive, or when a new classification is required for an existing sample, requiring that the sample be revisited for reclassification purposes. In this latter case it would be appropriate to classify and analyze a sub-sample of the original sample consisting of only the pooled population tails.

Acknowledgements. We thank J. Weller and Margaret Mackinnon for helpful comments. This study was supported by grants from the US-Israel Agricultural Research and Development Fund (BARD) and from the United States-Israel CDR program of the United States Agency for International Development (AID).

References

Beever JE, George PD, Fernando RL, Stormont DJ, Lewin HA (1989) Associations between genetic markers and growth and carcass traits in a paternal half-sib family of Angus cattle. J Anim Sci 68:337–344

Darvasi A, Soller M (1992) Optimal spacing of genetic markers for determining linkage between marker loci and quantitative trait loci. (Submitted)

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

Elston RD, Stewart J (1971) A general method for the genetic analysis of pedigree data. Hum Hered 21:523–542

Falconer DS (1989) Introduction to quantitative genetics 3rd edn. Longman, New York

Geldermann H (1975) Investigations on inheritance of quantitative characters in animals by genetic markers. I. Methods. Theor Appl Genetics 46:319–330

Geldermann H, Pieper U, Roth B (1985) Effects of marked chromosome sections on milk performance in cattle. Theor Appl Genet 70:138–146

Gonyon DS, Mather RE, Hines HC, Haenlein GFW, Arave CW, Gaunt SN (1987) Associations of bovine blood and milk polymorphisms with lactation traits: Holstein. J Dairy Sci 70:2585–2598

Haenlein GFW, Gonyon DS, Mather RE, Hines HC (1987) Associations of bovine blood and milk polymorphisms with lactation traits: Guernseys. J Dairy Sci 70:2599–2609

Hill AP (1975) Quantitative linkage: A statistical procedure for its detection and estimation. Ann Hum Genet 38:439–449

Jensen (1989) Estimation of recombination parameters between a quantitative trait locus (QTL) and two marker gene loci. Theor Appl Genet 78:613–618

Kahler AL, Wehrhahn CF (1986) Associations between quantitative traits and enzyme loci in the F_2 population of a maize hybrid. Theor Appl Genet 72:15–26

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

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

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

Neimann-Sorensen A, Robertson A (1961) The association between blood groups and several production characters in three Danish cattle breeds. Acta Agric Scand 11:163–196

Paterson AH, Lander ES, Hewitt JD, 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

Sax K (1923) The association of size difference with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetic 8:552–560

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

Soller M, Beckmann JS (1983) Genetic polymorphism in varietal identification and genetic improvement. Theor Appl Genet 67:25–33

Soller M, Beckmann JS (1987) Toward an understanding of the genetic basis of trypanotolerance in the N'Dama cattle of West Africa. Consultation report submitted to FAO, Rome, March 1987

Soller M, Beckmann JS (1988) Genomic genetics and the utilization for breeding purposes of genetic variation between populations. In: Weir BS, Eisen EJ, Goodman MM, Namkoong G (eds) Proc 2nd Int Conf Quant Genet. Sinauer Assoc, Sunderland, Mass., pp 161–188

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

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

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

Tanksley SD, Medina-Filho H, Rick CM (1982) Use of naturally occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49:11–15

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

Weller JI (1987) Mapping and analysis of quantitative trait loci in Lycopersicon (tomato) with the aid of genetic markers using approximate maximum likelihood methods. Heredity 59:413–421

Weller JI, Soller M, Brody T (1988) Linkage analysis of quantitative traits in an interspecific cross of tomato (*Lycopersicon esculentum* × *Lycopersicon pimpinellifolium*) by means of genetic markers. Genetics 118:329–339