Behavior Genetics, Vol. 35, No. 2, March 2005 (© 2005)

Mapping Quantitative Trait Loci Using Linkage Disequilibrium: Marker- versus Trait-based Methods

Albert Tenesa,[1,*] Peter M. Visscher,[1] Andrew D. Carothers,[2] and Sara A. Knott[1]

Received 25 May 2004—Final 28 Sep. 2004

Two approaches for mapping quantitative trait loci (QTL) using linkage disequilibrium at the population level were investigated. In the trait-based (TB) approach, the frequencies of marker alleles (or genotypes) are compared in individuals selected from the two tails of the trait distribution. The TB approach uses phenotypic information only in the selection step. In the marker-based (MB) approach, the quantitative trait values for the marker genotypes in the selected individuals are compared. The MB approach uses both the difference in marker allele (or genotype) frequencies and the phenotypic values of each marker genotype in the selected samples. We quantify the power of each approach and show that the power of the MB approach is greater than or equal to that of the TB approach. The advantage of the former is expected to increase with increasing number of traits phenotyped. Our accurate predictions obviate the need for elaborate simulation studies.

KEY WORDS: Association; LD; mapping; power; QTL.

INTRODUCTION

In recent years, human geneticists have advocated the use of linkage disequilibrium (LD) at the population level to fine-map genes associated with complex diseases. The reasons are that traditional linkage methods offer poor resolution of the trait locus position (due to the small number of recombination events available in most human pedigrees), that they have low power to detect associations between genes of small effect and complex traits, and that technological advances, such as high-throughput genotyping methods, are now available for typing large numbers of genetic markers (e.g. SNPs) in large numbers of individuals, making it feasible to use population-wide LD for mapping. At the same time there has been an increased interest in quantitative traits that are genetically correlated with disease status, because they are generally more easily and objectively measured than are binary traits (such as disease status). However, geneticists sometimes dichotomize continuous traits in an attempt to classify individuals as affected or unaffected. Osteoporosis is a good example of this (Langdahl *et al.*, 2003). According to World Health Organization criteria, a person has osteoporosis if they have a bone mineral density of less than two and a half standard deviations below the young population mean. In this case, people in the lower tail of the trait distribution would be treated as cases and the rest of the population as controls. This dichotomizing effect is sometimes taken to the extent that only individuals with very extreme phenotypes are used (Angius *et al.*, 2002; Little *et al.*, 2002). These are then treated as disease phenotypes, and the data analyzed using the appropriate linkage approach. This "dichotomizing" approach may be favoured because it mimics traditional disease mapping methods, allowing similar interpretation of results

[1] School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, Scotland, UK.
[2] MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh, Scotland, UK.
* To whom correspondence should be addressed at The University of Edinburgh, MRC Human Genetics Unit-Western General Hospital, Crewe Road, South Edinburgh, EH4 2XU Scotland, UK. Tel: +44 131 3322471. Fax: +44 131 4678450. E-mail: albert.tenesa@ed.ac.uk

and the use of readily-available software. However, the price paid (as shown below), in terms of loss of statistical power, might in some cases, be too high. Although the "dichotomizing" approach might be justified from a practitioner's point of view, allowing a decision to be made on whether a patient needs treatment, it is not always justified when trying to map the relevant trait loci.

Following Lebowitz *et al.*, (1987), we will refer to the "dichotomizing" approach for mapping quantitative trait loci (QTL) using LD at the population level as the trait-based (TB) approach, and to the "non-dichotomizing" approach as the marker-based (MB) approach. TB methods dichotomize individuals into two classes and, therefore, information contained within each class is lost. In this study, we quantify this loss in terms of statistical power and show that MB methods either outperform or, at worst, are equivalent to TB methods.

METHODS

The TB approach compares allele (or genotype) frequencies in individuals selected from the two extremes of the trait distribution, whereas the MB approach compares the mean phenotypic value for each marker allele (or genotype) in the same individuals. For each of the TB and MB approaches, we performed two tests, based on an additive and a dominant model of analysis, respectively.

For the additive model of analysis, the TB and MB tests were respectively a χ^2 test, and an F-test from the regression of phenotype on the numbers of a given marker allele present (0, 1 or 2). The χ^2 test (based on a 2×2 contingency table; 2 marker alleles and 2 tails) compared the allele counts in individuals selected from the two tails of the trait distribution. If the distribution of allele frequencies differed significantly in the two tails, then this suggested that the marker was in LD with the locus affecting the trait. In the regression analysis, the presence of a QTL in LD with the marker locus would lead to a non-zero slope of the regression line.

For the dominant model of analysis, the TB and MB tests were respectively a χ^2 test (based on a 3×2 contingency table; 3 possible marker genotypes and 2 tails) and an F-test from an ANOVA, respectively. The χ^2 test was based on the comparison of genotype counts in individuals selected from the two tails of the trait distribution. A significant difference in the frequency distribution of the marker genotypes between the upper and lower tails suggested the presence of a QTL influencing the trait in LD with the marker locus. The ANOVA tested whether the quantitative trait values for the marker genotypes of the selected individuals were different. Under the null hypothesis of no QTL in LD with the marker locus, the phenotypic values for the different marker genotypes would not differ significantly. In order to make a fair comparison of the two approaches (MB vs. TB), comparisons should be made only for tests with the same degrees of freedom. Comparisons were made between tests with two degrees of freedom (the ANOVA F-test and the χ^2 test based on genotype counts) and between tests with one degree of freedom (the regression F-test and the χ^2 test based on allele counts).

Model and Derivations

We suppose that the trait is influenced by a bi-allelic QTL with alleles Q_1 and Q_2, having frequencies q_1 and q_2 $(= 1 - q_1)$, respectively. We assume Hardy–Weinberg equilibrium and that phenotypic values for the three genotypes Q_1Q_1, Q_1Q_2 and Q_2Q_2 are normally distributed about mean values of $\mu_{11}(= -a)$, $\mu_{12}(= ad)$ and $\mu_{22}(= a)$, respectively, and with equal variances (taken to be 1). The QTL genotype Q_1Q_2 is considered to be phenotypically identical to Q_2Q_1 $(\mu_{12} = \mu_{21})$. However, we distinguish them to clarify the mathematical expressions below. The QTL narrow sense heritability is defined as $h^2_{\mathrm{QTL}} = V_{\mathrm{A}}/(V_{\mathrm{A}} + V_{\mathrm{D}} + 1)$ where $V_{\mathrm{A}}(= 2q_1q_2a^2[1 + \{q_1 - q_2\}d]^2)$ and $V_{\mathrm{D}}(= [2q_1q_2da]^2)$ are respectively the additive and dominant variances for the QTL. If x denotes the phenotypic value of an individual, then the probability density function of x is

$$\rho(x) = \sum_{i=1}^{2} \sum_{j=1}^{2} q_i q_j \varphi(\mu_{ij}, \sigma^2_{ij}),$$

where $\varphi(\mu, \sigma^2)$ denotes the Gaussian distribution with mean μ and variance σ^2 (assumed, without loss of generality, to be 1).

The upper and lower tails of the distribution are defined, respectively, as the proportions α_U and α_L of the individuals phenotyped that are to be genotyped. We determined the upper and lower cutoffs τ_U and τ_L by solving:

$$\alpha_U = \int_{\tau_U}^{\infty} \rho(x)dx; \quad \alpha_L = \int_{-\infty}^{\tau_L} \rho(x)dx$$

If $\Phi(\tau - \mu_{ij}) = \Phi_{ij} = \int_{-\infty}^{\tau} \varphi(x|\mu = \mu_{ij}, \sigma^2 = 1)dx$, then the probability that an individual selected from one of the tails of the trait distribution has a given QTL genotype is:

$$P(Q_iQ_j|x > \tau_U) = \frac{q_iq_j(1 - \Phi(\tau_U - \mu_{ij}))}{\alpha_U};$$
$$P(Q_iQ_j|x < \tau_L) = \frac{q_iq_j\Phi(\tau_L - \mu_{ij})}{\alpha_L}; \ i,j \in [1,2]$$

and the frequency of the QTL alleles in the two tails of the trait distribution is:

$$P(Q_i|x > \tau_U) = \frac{q_i \sum_{j=1}^{2} q_j(1 - \Phi(\tau_U - \mu_{ij}))}{\alpha_U};$$
$$P(Q_i|x < \tau_L) = \frac{q_i \sum_{j=1}^{2} q_j(\Phi(\tau_L - \mu_{ij}))}{\alpha_L}; \ i \in [1,2]$$

where i, j denote the possible QTL alleles.

The expected values for QTL genotype Q_iQ_j, in the upper and lower tails are η_{ij}^U and η_{ij}^L, respectively, where

$$\eta_{ij}^U = \frac{1}{(1 - \Phi(\tau_U - \mu_{ij}))} \int_{\tau_U}^{\infty} \frac{x}{\sqrt{2\pi}} e^{\frac{-(x-\mu_{ij})^2}{2}} dx$$
$$= \frac{z_{ij}^U}{(1 - \Phi(\tau_U - \mu_{ij}))} + \mu_{ij} = i_{ij}^U + \mu_{ij},$$
$$\eta_{ij}^L = \frac{1}{\Phi(\tau_L - \mu_{ij})} \int_{-\infty}^{\tau_L} \frac{x}{\sqrt{2\pi}} e^{-\frac{(x-\mu_{ij})^2}{2}} dx$$
$$= \frac{-z_{ij}^L}{\Phi(\tau_L - \mu_{ij})} + \mu_{ij} = i_{ij}^L + \mu_{ij},$$

where z_{ij}^U and z_{ij}^L are the ordinates of the appropriate Gaussian distribution [$\phi(\mu_{ij}, 1)$] at the cut-offs τ_U and τ_L, respectively.

Since we do not usually genotype the QTL itself, we assume that there is a linked bi-allelic marker locus in LD with the trait locus, and that the marker locus does not have an independent effect on the trait. The marker locus has alleles M_1 and M_2 with frequencies m_1 and $m_2 (= 1 - m_1)$, respectively.

The disequilibrium parameter (D) between marker allele M_2 and QTL allele Q_2 is defined as $D = f_{Q_2M_2} - q_2m_2$, where $f_{Q_2M_2}$ is the population frequency of the haplotype Q_2M_2. We expressed our results as a function of Lewontin's normalized measure of disequilibrium D' (Lewontin, 1964). D' is the value of D expressed as a fraction of its maximum possible value, that is, $D'(= D/D_{\max})$, where $D_{\max}$ is the minimum value of q_2m_1 or q_1m_2 (as D is assumed, without loss of generality, to be positive between alleles M_2 and Q_2).

The four possible QTL-marker haplotype frequencies are $P(Q_1M_1) = q_1m_1 + D' \times D_{\max}$, $P(Q_1M_2) = q_1m_2 - D' \times D_{\max}$, $P(Q_2M_1) = q_2m_1 - D' \times D_{\max}$ and $P(Q_2M_2) = q_2m_2 + D' \times D_{\max}$. Assuming random mating, the probability of each of the possible QTL-marker diplotypes is equal to the product of their component haplotypes (e.g., $P(Q_iM_l, Q_jM_n) = P(Q_iM_l) \times P(Q_jM_n)$). The probabilities of occurrence of each marker genotype in the upper and lower tails are obtained using Bayes' theorem. After some algebra:

$$P(M_lM_n|x > \tau_U) = \sum_{i=1}^{2}\sum_{j=1}^{2} P(Q_iM_l, Q_jM_n)$$
$$(1 - \Phi(\tau_U - \mu_{ij}))/\alpha_U; l,n \in [1,2], \quad (1)$$

$$P(M_lM_n|x < \tau_L) = \sum_{i=1}^{2}\sum_{j=1}^{2} P(Q_iM_l, Q_jM_n)\Phi(\tau_L - \mu_{ij})/\alpha_L;$$
$$l,n \in [1,2], \quad (2)$$

where l, n denote the possible marker alleles. From equations (1) and (2) we obtain the probabilities of occurrence of the kth marker allele in the two tails.

$$P(M_k|x > \tau_U) = \sum_{l=1}^{2}\sum_{n=1}^{2} P(M_k|M_lM_n)P(M_lM_n|x > \tau_U),$$
$$P(M_k|x < \tau_L) = \sum_{l=1}^{2}\sum_{n=1}^{2} P(M_k|M_lM_n)P(M_lM_n|x < \tau_L),$$

where $P(M_k|M_lM_n)$ is 1,1/2 and 0 for $k = l = n$, $k = l \neq n$ or $k = n \neq l$ and $k \neq l = n$, respectively.

The expected quantitative trait value for marker genotype M_lM_n in the selected sample is equal to:

$$E(x|M_lM_n) = \frac{\alpha_U \times P(M_lM_n|x > \tau_U) \times E(x|M_lM_n)^U + \alpha_L \times P(M_lM_n|x<\tau_L) \times E(x|M_lM_n)^L}{\alpha_U \times P(M_lM_n|x > \tau_U) + \alpha_L \times P(M_lM_n|x<\tau_L)}; \quad l,n \in [1,2]$$

where $E(x|M_lM_n)^U$ and $E(x|M_lM_n)^L$ are the expected quantitative trait values for marker genotype M_lM_n in the upper and lower tails (see Appendix).

Derivations of the non-centrality parameters for the four tests studied are shown in the Appendix. The MB tests were based on an F_{n_1,n_2} distribution (note that n_1 and n_2 are just general labels and do not mean anything in the context of the model). However, when the denominator degrees of freedom are large ($n_2 \to \infty$) this distribution can be approximated to n_1^{-1} times a chi-squared distribution with n_1 degrees of freedom. The sample sizes required to detect a QTL with small effect (as considered here) are large, and we therefore considered the approximation to be valid, and referred all the results to a chi-squared distribution. This makes the comparison of the two approaches easier. Simulations were carried out to check that the approximations, shown in the Appendix, were very close (results not shown).

RESULTS

All the results shown assumed that selection was symmetric, so that $2\alpha_U = 2\alpha_L = P$. Figure 1 shows that with equal degrees of freedom MB methods always performed better than TB methods. This was so regardless of the genetic model considered for the generation of the data. Power was practically the same for both approaches when selection was sufficiently intense. However, differences in power were important when the whole population was genotyped as shown in Table I.

Figure 2 shows the effect of the amount of LD on power for three different intensities of selection. For extreme selection the power curves for the MB

Table I. Comparison of the power obtained with the TB or MB approach for different levels of disequilibrium when the whole population is genotyped (P = 1)

		Marker-based		Trait-based	
Model	D'	Anova (2df)	Regression (1 df)	χ^2 Genotype (2 df)	χ^2 Alleles (1 df)
Dominant (d = 1)	1.00	0.98	0.93	0.64	0.51
	0.75	0.37	0.35	0.08	0.08
	0.50	0.01	0.01	0.01	0.01
Additive (d = 0)	1.00	0.88	0.93	0.38	0.48
	0.75	0.26	0.35	0.04	0.07
	0.50	0.01	0.01	<0.01	0.01
Recessive (d = −1)	1.00	>0.99	0.96	0.91	0.23
	0.75	0.82	0.37	0.12	0.02
	0.50	0.03	0.01	0.01	0.01

The marker and QTL were assumed to be in varying levels of disequilibrium (D'), $m_2 = q_2 = 0.3$ and $h^2_{QTL} = 0.05$. The significance level was 10^{-8} and the total number of individuals genotyped was 1000.

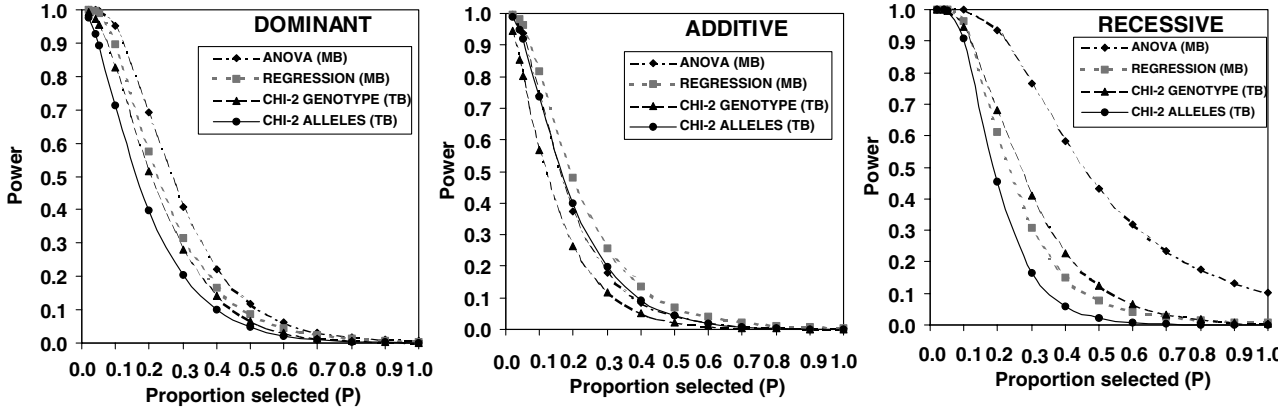

Fig. 1. Comparison of the power obtained when using the TB and MB approach for different proportions selected (P). The marker is assumed to be the QTL ($D' = 1$) with allele frequency 0.3, $h^2_{QTL} = 0.05$ and d = 1, 0, −1 respectively for the dominant, additive and recessive models. The significance level was 10^{-8}, the total number of individuals genotyped was fixed to 200 and number phenotyped was 200/P.

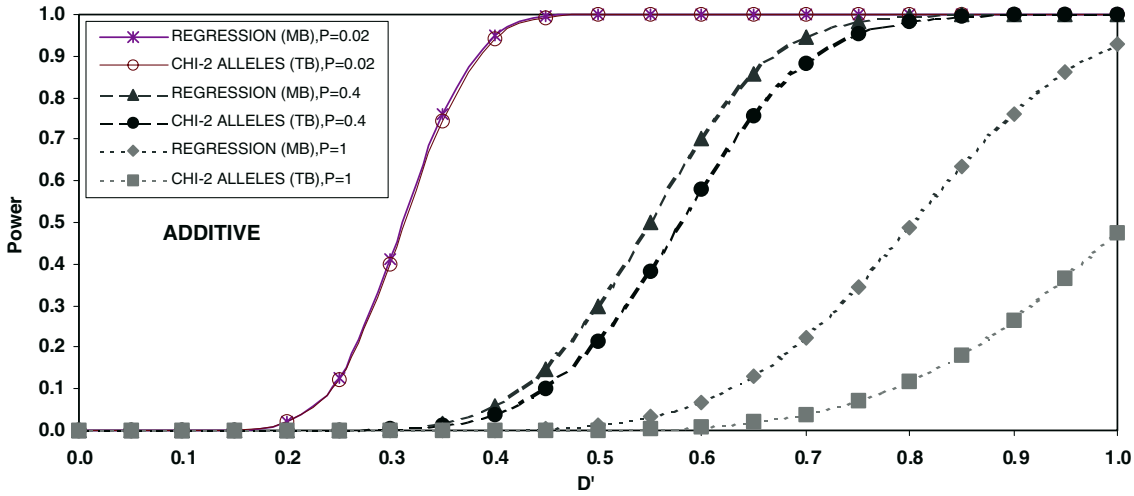

Fig. 2. Effect of the amount of LD on power of the MB and TB approach. The marker and the QTL both have allele frequency equal to 0.3, $h^2_{QTL} = 0.05$ and the model is additive. The significance level was 10^{-8} and the total number of individuals genotyped was 1000. The proportion selected (P) is shown in the Figure.

and TB approach almost overlapped regardless of the amount of LD. The difference between the two approaches was largest when no selection was applied.

Figure 3 shows the power obtained for the MB and TB approach with 1 *df* when we fixed the number of individuals phenotyped. The TB approach provides maximum power when about 27.5% of individuals are genotyped in each tail, and is less powerful than the MB approach. This level of selection ($p = 55\%$) provides the maximum power for the TB approach (in accordance with Lebowitz *et al.*, (1987) and Bader *et al.*, (2001)) but not for the MB approach, for which power increases monotonically with the number of individuals genotyped and included in the analysis.

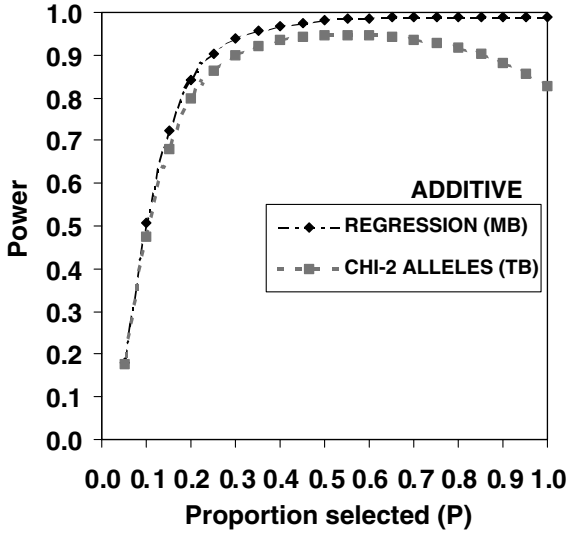

Fig. 3. Comparison of the power obtained when using the TB and MB approach with 1 df for different proportions selected (P) when the total number of phenotyped individuals is fixed. The total number of phenotypes is 4500 and the total number of genotypes is 4500×P. The marker is assumed to be the QTL ($D' = 1$) with allele frequency 0.3, $h^2_{QTL} = 0.01$ and the model is additive. The significance level was 10^{-5}.

Figure 4 shows the effect of a discrepancy between marker and QTL allele frequency has on power. When the discrepancy between QTL and marker allele frequency increases power is smaller, this is so regardless of the level of disequilibrium ($D' = 1$ and $D' = 0.7$ in figure 4). The discrepancy between marker and trait allele frequency has similar effect for the TB and MB approach and the MB is always more powerful.

DISCUSSION

Quantitative traits are of interest for human genetics because they are often correlated with disease traits. Schork *et al.*, (2000) proposed the use of threshold-defined case/controls (that is, the TB approach with 1 *df*) for mapping loci influencing quantitative traits using LD at the population level. The objective of the present study was to assess how much information is lost when analyzing a quantitative trait as a threshold-defined binary trait as opposed to analyzing it using all the information available. The information lost in the former case is

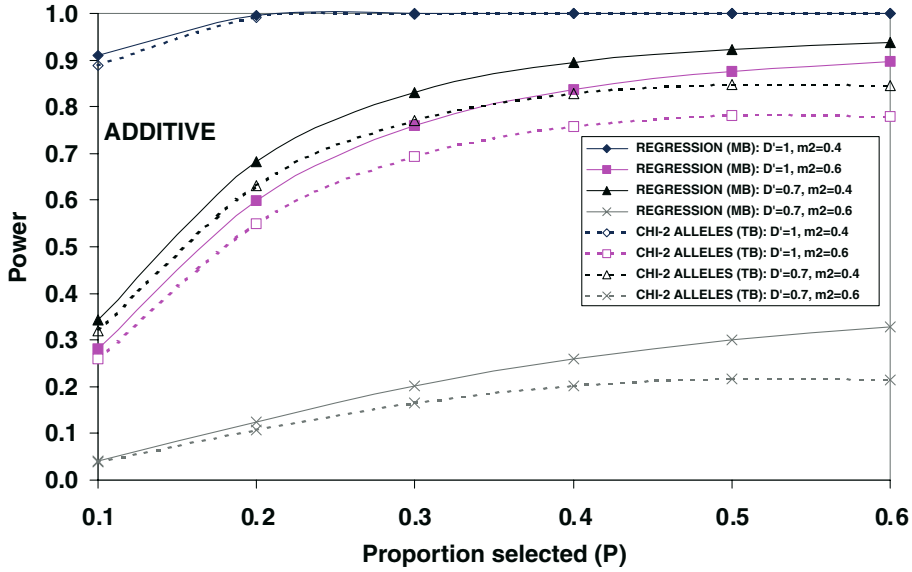

Fig. 4. Effect of the discrepancy between marker and QTL allele frequency. A total number of 10000 individuals' phenotypes are assumed to be measured and a proportion (P) of them to be genotyped for analysis. Two different marker allele frequencies are considered ($m_2 = 0.4$ or 0.6) to detect a QTL allele frequency (q_2) of 0.2. Two different levels of disequilibrium are shown ($D' = 1$ or 0.7). The genetic model is additive and $h^2_{\mathrm{QTL}} = 0.02$. The significance level was 10^{-5}.

clearly reflected in a loss of statistical power to detect an association between marker genotype and phenotype. Our proposed method of analysis was more powerful, except under very extreme selection, when both methods performed similarly. Our method can be implemented with standard statistical packages or spreadsheets.

Sirota *et al*., (1999) studied the power of selective genotyping for a TB method (i.e. they compared the genotypic frequencies in the upper and lower groups). They typed the 5-HTTLPR polymorphism in the promoter of the serotonin transporter gene in a sample of 902 individuals and obtained the expected power that one would obtain at different selection intensities given the difference in allele frequencies between tails they observed in their sample (the empirical power). In addition, they simulated a large dataset based on the parameters they empirically observed (i.e., the observed genotype frequencies and effect size in the 902 individuals) and obtained the power (the theoretical power). They found that empirical differences in allele frequencies between the two selected groups decreased as selection increased and that empirical power decreased with increasing selection intensity. In contrast, the theoretical power increased with increasing selection intensity. Sirota *et al*., (1999) concluded that using extreme samples does not give increased power, at least not for their trait and polymorphism, and suggested that the reduction in power is caused by their observed non-linear increase in allele frequency difference in the tails with increasing selection intensity. From the results presented here with a fixed number of individuals phenotyped, greatest power would be obtained by genotyping as many individuals as possible (as long as the appropriate test was used) (See Figure 3). If a TB method is used for a fixed number of phenotypes, in agreement with the observations of Sirota *et al*., (1999), the optimum will be intermediate because of the trade-off between increased allele frequency difference and decrease in the accuracy of the allele frequency estimate due to small sample sizes. The apparent discrepancy between the empirical and theoretical results shown by Sirota *et al*. (1999) is because they used a fixed number of individuals genotyped in their theoretical approach and a fixed number of individuals phenotyped in their empirical results. In addition their observed non-linear change in HTTLPR allele frequency with increasing selection intensity, may not be inconsistent with a linear relationship between genotype and phenotype as they suggest, but could purely be due to sampling. For example, if one were to phenotype 902 people and select for genotyping those above a deviate of ± 1.75 ($Z >= 1.75$ or $Z < 1.75$) one would genotype about 36 people (N) from each tail. If the allele frequencies (f) were 0.5 and 0.82 in the lower and upper tail respectively,

then the sampling variance $[f*(1-f)/N]$ for f would be 0.007 and 0.004 in the lower and upper tail, respectively. Hence, any sample of size 36 with allele frequency between 0.33–0.66 and 0.69–0.94 would be consistent with the allele frequencies being 0.5 and 0.82.

The results obtained from the TB approach are valid only under asymptotic assumptions, that is, for large sample sizes. Small sample sizes and low frequency alleles might lead to sparse contingency tables and hence spurious results. On the other hand, the MB approach does not assume large samples and is quite robust to departures of normality (simulation results not shown).

In the present study, we have assumed that the three genotypic values have equal variances and are normally distributed. As discussed by Schork *et al.*, (2000), these assumptions might not be completely realistic. Usually, the larger the mean the larger the variance around this mean. In addition, complex multi-locus effects might lead to departures from normality (Allison *et al.*, 1998) especially if the loci effects are big. We do not, however, expect these two factors to be of huge importance for the QTL of small effect shown above.

Currently, large numbers of SNPs are available, and researchers are interested in exploiting them by using multi-locus haplotypes instead of single marker alleles. Using multi-locus haplotypes might have higher power than using single marker alleles for detecting an association. However, the relative efficiency of the MB and TB approach would be the same, regardless. On the other hand, in the final stage of a study (i.e., after the haplotype analysis) researchers would like to know which unidentified variant or variants are causing the phenotypic differences, and this would involve testing each variant independently (as assumed in our study).

Finally, our results demonstrate the advantage of the MB over the TB approach. Although both approaches would be similarly efficient when just one trait was phenotyped and high selection performed for just this trait, this would not be so when the selection intensity was low. The most realistic scenario would be one in which a number of individuals had been phenotyped for a number of traits. If selection had to be applied for a large number of traits, then all or almost all of the individuals phenotyped would eventually be genotyped. The MB approach would then clearly be the most powerful.

APPENDIX

The expected values for marker genotype M_lM_n in the upper tail and lower tails are $E(x|M_lM_n)^U$ and $E(x|M_lM_n)^L$, respectively.

$$E(x|M_lM_n)^U = \sum_{i=1}^{2}\sum_{j=1}^{2}\eta_{ij}^U \times \frac{P(x>\tau_U|Q_iQ_j)\times P(Q_iM_l,Q_jM_n)}{P(M_lM_n|x>\tau_U)\times\alpha_U};\ \ l,n\in[1,2]$$

$$E(x|M_lM_n)^L = \sum_{i=1}^{2}\sum_{j=1}^{2}\eta_{ij}^L \times \frac{P(x<\tau_L|Q_iQ_j)\times P(Q_iM_l,Q_jM_n)}{P(M_lM_n|x<\tau_L)\times\alpha_L};\ \ l,n\in[1,2]$$

The within-genotype variance for the marker genotypes in both tails combined is:

$$\mathrm{var}(x|M_lM_n) = E(x^2|M_lM_n) - E(x|M_lM_n)^2$$
$$= \frac{\sum_{i=1}^{2}\sum_{j=1}^{2}\left[P(Q_iM_l,Q_jM_n)\times\left((1+\mu_{ij}^2)(\Phi_{ij})+z_{ij}^U(\tau_U+\mu_{ij})-z_{ij}^L(\tau_L+\mu_{ij})\right)\right]}{\sum_{i=1}^{2}\sum_{j=1}^{2}\left[P(Q_iM_l,Q_jM_n)\times(\Phi_{ij})\right]}$$
$$-\left\{\frac{\sum_{i=1}^{2}\sum_{j=1}^{2}\left[P(Q_iM_{l,}Q_jM_n)\times\left(z_{ij}^U-z_{ij}^L+\mu_{ij}\times\Phi_{ij}\right)\right]}{\sum_{i=1}^{2}\sum_{j=1}^{2}\left[P(Q_iM_{l,}Q_jM_n)\times\Phi_{ij}\right]}\right\}^2$$

The X^2 statistics, obtained from contingency tables of 3 ×2 and 2 ×2 for the counts of genotypes and alleles, respectively, are distributed under the

null hypothesis (H_0) of no association as chi-squared with 2 and 1 degrees of freedom, respectively. Under the alternative hypothesis (H_1), X^2 is asymptotically distributed as non-central chi-squared with respectively 2 and 1 degrees of freedom and non-centrality parameter $\lambda_{\text{Genotypes}}$ and λ_{Alleles} given by

$$\lambda_{\text{Genotypes}} = N_T \times \left[\sum_{l=1}^{2}\sum_{n=1}^{2}\frac{\left(P(M_lM_n, x<\tau_L|H_1) - P(M_lM_n, x<\tau_L|H_0)\right)^2}{P(M_lM_n, x<\tau_L|H_0)}\right] + N_T \times \left[\sum_{l=1}^{2}\sum_{n=1}^{2}\frac{\left(P(M_lM_n, x>\tau_U|H_1) - P(M_lM_n, x>\tau_U|H_0)\right)^2}{P(M_lM_n, x>\tau_U|H_0)}\right]$$

which after some algebra reduces to

$$= \frac{\alpha_U\alpha_L N_T}{(\alpha_U+\alpha_L)^2} \times \sum_{l=1}^{2}\sum_{n=1}^{2}\frac{\left(\sum_{i=1}^{2}\sum_{j=1}^{2}P(Q_iM_l, Q_jM_n) \times \left[\left(1-\Phi(\tau_U-\mu_{ij})\right)/\alpha_U - \Phi(\tau_L-\mu_{ij})/\alpha_L\right]\right)^2}{P(M_lM_n)}$$

$$\lambda_{\text{Alleles}} = 2N_T \times \left[\sum_{l=1}^{2}\frac{\left(P(M_l, x<\tau_L|H_1) - P(M_l, x<\tau_L|H_0)\right)^2}{P(M_l, x<\tau_L|H_0)}\right] + 2N_T \times \left[\sum_{l=1}^{2}\frac{\left(P(M_l, x>\tau_U|H_1) - P(M_l, x>\tau_U|H_0)\right)^2}{P(M_l, x>\tau_U|H_0)}\right]$$

which after some algebra reduces to

$$= \frac{2\alpha_U\alpha_L N_T}{(\alpha_U+\alpha_L)^2} \times \sum_{k=1}^{2}\frac{\sum_{l=1}^{2}\sum_{n=1}^{2}\left\{\left(\sum_{i=1}^{2}\sum_{j=1}^{2}P(Q_iM_l, Q_jM_n) \times \left[\left(1-\Phi(\tau_U-\mu_{ij})\right)/\alpha_U - \Phi(\tau_L-\mu_{ij})/\alpha_L\right]\right) \times P(M_k|M_lM_n)\right\}^2}{\sum_{l=1}^{2}\sum_{n=1}^{2}P(M_lM_n) \times P(M_k|M_lM_n)}$$

where N_T denotes the number of individuals genotyped (Kendall and Stuart, 1961). Power is then defined as the probability that a non-central χ^2 with respectively 2 and 1 degrees of freedom and non-centrality parameters $\lambda_{\text{Genotypes}}$ and λ_{Alleles} is greater than the critical value defined by a central χ^2 with 2 and 1 degrees of freedom and significance level α.

Testing for association between marker genotype and phenotype using ANOVA requires specifying the model. Our model is $y_{gz} = \mu + \tau_g + e_{gz}$ where y_{gz} is the phenotype for individual z with marker genotype g ($=1$ if M_1M_1 $=2$ if M_1M_2 or M_2M_1, $=3$ if M_2M_2); μ is the mean of the selected individuals from both tails; τ_g is the gth marker genotype effect taken to be fixed (its effect is constrained so that $\sum_{g=1}^{3} n_g\tau_g = 0$); and e_{gz} is the residual effect for individual z with genotype g. The total number of individuals sampled is $N_T = \sum_{g=1}^{3} n_g$ where n_1, n_2, n_3 are respectively the numbers with genotypes M_1M_1, M_1M_2, or M_2M_1 and M_2M_2 (strictly speaking, n_g are random variables but here we treat them as fixed and equal to their expected values).

The between marker genotype sum of squares is $SS_B = \sum_{g=1}^{3} n_g \times \left(\bar{y}_{g\bullet} - \bar{y}_{\bullet\bullet}\right)^2$ and the within marker genotype sum of squares is $SS_W = \sum_{g=1}^{3}\sum_{z=1}^{n_g}\left(y_{gz} - \bar{y}_{g\bullet}\right)^2$ where $\bar{y}_{g\bullet}$ and $\bar{y}_{\bullet\bullet}$ are the mean phenotypic values for the g marker genotype of the selected individuals and for the selected individuals, respectively. When selection is applied, the within genotypic variance (σ^2_W) is not equal for all genotypes. Therefore, we use the weighted average (the weights being n_g).

Under H_0, the statistic $F = \frac{MS_B}{MS_W} \sim \frac{\chi^2_2}{2}$ for large N_T.

Under H_1, $E(SS_B/\sigma^2_W) = 2 + \lambda_{\text{ANOVA}} = 2 + \frac{\sum_{x=1}^{3} n_x\tau_x^2}{\sigma^2_W}$, where λ_{ANOVA} is the non-centrality parameter

and $F \sim F_{2,N_T-3,\lambda_{\text{ANOVA}}}$ (Kendall and Stuart, 1961). We express the non-centrality parameter as:

$$\lambda_{\text{ANOVA}} = \sum_{l=1}^{2}\sum_{n=1}^{2} \frac{N_T \times P(M_lM_n) \times (E(x|M_lM_n) - \mu)^2}{\sigma_W^2},$$

where

$$\begin{aligned}\mu &= \sum_{i=1}^{2}\sum_{j=1}^{2} E(x|Q_iQ_j, x > \tau_U \text{ or } x<\tau_L) \\ &\quad \times P(Q_iQ_j, x > \tau_U \text{ or } x<\tau_L) \\ &= \sum_{l=1}^{2}\sum_{n=1}^{2} \frac{q_iq_j \times \left(z_{ij}^U - z_{ij}^L + \mu_{ij}\Phi_{ij}\right)}{\alpha_U + \alpha_L}\end{aligned}$$

Power is then defined as the probability that a non-central $F_{2,N_T-3,\lambda_{\text{ANOVA}}}$ with 2 and $N_T - 3$ degrees of freedom and non-centrality parameter λ_{ANOVA} is greater than the critical value defined by an F_{2,N_T-3} with 2 and $N_T - 3$ degrees of freedom and significance level α.

Regression of phenotype on marker genotype is the last type of analysis considered here. The model is $y_z = a + bx_z + e_z$ where y_z is the phenotype for individual z, a is the intercept, b is the slope, x_z is a dummy variable for individual z (taking values -1, 0 or 1 depending on whether the individual's genotype is M_1M_1, M_1M_2 ($= M_2M_1$ or M_2M_2, respectively) and e_z is the residual for individual z. Regression tests for marker–trait association with one degree of freedom (i.e., ignores non-additivity), while the ANOVA based test has 2 degrees of freedom. The expected value for the estimate of b equals:

$$\begin{aligned}E(\hat{b}) &= E\left(\frac{SS_{xy}}{SS_{xx}}\right) \\ &= \frac{\sum_{l=1}^{2}\sum_{n=1}^{2} P(M_lM_n) \times (x_{ln} - \bar{x}) \times (E(x|M_lM_n) - \mu)}{\sum_{l=1}^{2}\sum_{n=1}^{2} P(M_lM_n) \times (x_{ln} - \bar{x})^2},\end{aligned}$$

where SS_{xx}, SS_{xy} are respectively the sum of squares and the sum of products, $x_{11} = -1$, $x_{12} = x_{21} = 0$, $x_{22} = 1$ and

$$\begin{aligned}\bar{x} &= (P(M_2M_2|x > \tau_U) + P(M_2M_2|x<\tau_L)) \\ &\quad - (P(M_1M_1|x > \tau_U) + P(M_1M_1|x<\tau_L)) \\ &= \frac{1}{\alpha_U + \alpha_L}\sum_{i=1}^{2}\sum_{j=1}^{2}\{P(Q_iM_{2,}Q_jM_2) - P(Q_iM_{1,}Q_jM_1)\} \\ &\quad \times (\Phi_{ij}).\end{aligned}$$

Under H_0 the expected value of $\hat{b}$, b, is zero and the statistic $\text{T} = \hat{b}^2/\widehat{\text{var}}(\hat{b})$ is $F_{1,N_T-2} \sim \chi_1^2$ (for large N_T) distributed. Under H_1 the statistic T is non-central F distributed ($F_{1,N_T-2,\lambda_{\text{Regression}}}$), where $\lambda_{\text{Regression}} = SS_{xx} \times b^2/\sigma_W^2$ (Lynch and Walsh, 1998).

Power is then defined as the probability that a non-central $F_{1,N_T-2,\lambda_{\text{Regression}}}$ with *1* and $N_T - 2$ degrees of freedom and non-centrality parameter $\lambda_{\text{Regression}}$ is greater than the critical value defined by an F_{1,N_T-2} with *1* and $N_T - 2$ degrees of freedom and significance level α.

ACKNOWLEDGMENTS

We thank Eco de Geus for helpful discussions. SAK acknowledges support from the Royal Society and PMV from the Biotechnology and Biological Sciences Research Council.

REFERENCES

Allison, D.B., Heo, M., Schork, N. J., Wong, S. L., and Elston, R. C. (1998). Extreme selection strategies in gene mapping studies of oligogenic quantitative traits do not always increase power. *Hum. Hered.* **48**:97–107.

Angius, A., Petretto, E., Maestrale, G. B., Forabosco, P., Casu, G., Piras, D., Fanciulli, M., Falchi, M., Melis, P. M., Palermo, M., and Pirastu, M. (2002). A new essential hypertension susceptibility locus on chromosome 2p24-p25, detected by genomewide search. *Am. J. Hum. Genet.*. **71**:893–905.

Bader, J. S., Bansal, A., and Sham, P. (2001). Efficient SNP-based tests of association for quantitative phenotypes using pooled DNA. *Genescreen*. **1**:143–150.

Kendall, M. G., and Stuart, A. (1961). *The advanced theory of statistics. Vol. 2. Inference and relationship (4th ed.)*. Great Britain:Charles Griffin and Company Limited.

Langdahl, B. L., Carstens, M., Stenkjaer, L., and Eriksem, E. F. (2003). Polymorphisms in the transforming growth factor beta 1 gene and osteoporosis. *Bone* **32**:297–310.

Lebowitz, R. J., Soller, M., and Beckmann, J. S. (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.

Lewontin, R. C. (1964). The interaction of selection and linkage. I. General considerations; Heterotic Models. *Genetics* **49**:49–67.

Little, R. D., Carulli, J. P., Del Mastro, R. G., Dupuis, J., Osborne, M., Folz, C., Manning, S. P., Swain, P. M., Zhao, S., Eustace, B., Lappe, M. M., Spitzer, L., Zweier, S., Braun-

schweiger, K., Benchekroun, Y., Hu, X., Adair, R., Chee, L., FitzGerald, M. G., Tulig, C., Caruso, A., Tzellas, N., Bawa, A., Franklin, B., McGuire, S., Nogues, X., Gong, G., Allen, K. M., Anisowicz, A., Morales, A. J., Lomedico, P. T., Recker, S. M., Van Eerdewegh, P., Recker, R. R., and Johnson, M. L. (2002). A mutation in the LDL receptor-related protein 5 gene results in the autosomal dominant high-bone-mass trait. *Am. J. Hum. Genet.* **70**:11–19.

Lynch, M., and Walsh, B. (1998). *Genetics and analysis of quantitative traits*. USA: Sinauer Associates.

Schork, N. J., Nath, S. K., Fallin, D., and Chakravarti, A. (2000). Linkage disequilibrium analysis of biallelic DNA markers, human quantitative trait loci, and threshold-defined case and control subjects. *Am. J. Hum. Genet.* **67**:1208–1218.

Sirota, L. A., Greenberg, D. B., Murphy, D. L., and Hamer, D. H. (1999). Non-linear association between the serotonin transporter promoter polymorphism and neuroticism:a caution against using extreme samples to identify quantitative trait loci. *Psychiatr. Genet.* **9**:35–38.

Edited by Stacey Cherney